2011 15th IEEE International Enterprise Distributed Object Computing Conference Workshops

# Privacy Preservation Approach in Service Ecosystems

Yiyun Shen*, Markus Miettinen†, Pirjo Moen*, Lea Kutvonen*
*Department of Computer Science, University of Helsinki, Finland
Email: firstname.lastname@cs.helsinki.fi
†Nokia Research Centre Lausanne, Switzerland
Email: firstname.lastname@nokia.com

***Abstract*—**
**Emergence of business networking and social networking increases the exchange of sensitive information and creation of behaviour traces in the network. However, the current computing and communication solutions do not provide sufficient conceptual, architectural or technical facilities to preserve privacy while collaborating in the network. This paper enhances definition on privacy-related concepts to become sufficient for open service ecosystems, and finally introduces a privacy-preservation architecture with emphasis on usability, sustainability against threats, and reasonable cost of establishment and utilisation. As this architecture introduces new categories of tools for privacy preservation, it is significant also as a roadmap or maturity model.**

***Keywords*-privacy, service ecosystem, architecture**

## I. Introduction

The present trend of computer supported networking is gaining ground in both social networking and business networking domains. The increasing use of social media, networked business and cloud computing add exposure of private information and private knowledge about behaviour, and therefore, the requirements for privacy-preservation in the formed communities are essential.

We find that the current open service ecosystems do not provide sufficient mechanisms, or even sufficient concepts to appropriate privacy management by users themselves. Traditionally, privacy has been defined to involve individuals – as the roots of privacy as a concept remains in social sciences – but open service ecosystems also involve groups, organisations, and collaborations between them. Therefore, relationships between these subjects must become first class concepts in the future architectures.

Furthermore, the privacy-preservation technologies are not mature enough as there is no commonly accepted global architecture framework in which these independent technology solutions should be placed. The missing privacy-preservation architecture must govern service ecosystems so that

- subjects can declare their own privacy requirements;
- infrastructure utilities govern these declarations, using them for explicit definitions of privacy violations; and
- tools support collaboration and interaction models where risk of privacy violations is known or minimized.

The validity of such privacy-preservation architectures depends on their usability, sustainability against threats, and cost of establishment and utilisation.

This paper contributes a *privacy-preservation architecture to enhance open service ecosystem architectures*, i.e. an architecture that regulates collaboration contracting and governance amongst services provisioned by independent organisations or individuals. We use Pilarcos ecosystem architecture [1], [2], [3], [4] as an example and enhance it with privacy tools.

The proposed privacy-preservation architecture involves subjects of privacy (individuals, organisations, groups, collaborations) in preserving not only their own privacy, but involving all ecosystem members to privacy preserving norms. These norms represent our ecosystem related privacy-preservation goals: i) each subject has the right to their privacy, and further, privacy expectations are associated as declared properties to relationships between subjects; ii) mutually beneficial behaviours in catching privacy violations and traces are thereof utilised; iii) automated control mechanisms for privacy preservation are provided for all categories of subjects; and finally, iv) subjects are provided with methods on managing these privacy controls.

In the following, Section II explores privacy threats and privacy needs from the Pilarcos ecosystem perspective to provide motivational background and evaluation criteria for privacy-preservation architectures. Section III summaries challenges identified for a generic privacy-preservation architecture, and requirements for users' ability to appropriate their privacy expectations. Section IV introduces the key functionalities and identifies privacy-preservation approaches for open service ecosystems.

## II. Privacy Threats in Open Service Ecosystem

The success and competitive edge of enterprises is increasingly dependent on the enterprises' agility to become members in business networks that are supporting their own business strategies. Enterprises must be able to participate in multiple business networks simultaneously, be quick in adopting new kinds of well-crafted business models, and establish new collaborations swiftly. Therefore, integration solutions with their well-weathered strategic networks are no more sufficient. Acquiring this kind of flexibility sets two

978-0-7695-4426-7/11 $26.00 © 2011 IEEE
DOI 10.1109/EDOCW.2011.59

requirements: first, routine decision-making on committing to a collaboration must be automated (similarly to automated broker agents dealing on stock exchanges), and second, situational information in the large service ecosystem (availability of services, reputation of partners, alliances, risk involved in the business model) must be made available to these automation tools.

In the following we focus on the Pilarcos ecosystem as an example of this trend, and elaborate on the support that is required for the privacy-preservation architecture and facilities. The privacy aspects of interest include selecting acceptable collaboration models, restricting information made available for partners within the collaborations or publicly, and detecting that information has leaked out of the trusted circle of information users and acceptable usages of information.

### A. Open service ecosystem and its management

Let us first become familiar with the open service ecosystems. The ecosystem is an environment, i.e. an open service market, where service providers and clients can meet, establish contract-governed collaborations and gain experience on the services and partners involved.

The open service ecosystem [5] is supported with infrastructure services to solve the evident problems of semantic and pragmatic interoperability and collaboration-governing contract management. First, we understand interoperability, or the capability to collaborate, as an effective capability to mutually communicate information in order to exchange proposals, requests, results, and commitments. This term covers technical, semantic and pragmatic interoperability. Technical interoperability is concerned with connectivity between the computational services, allowing messages to be transported from one application to another. Semantic interoperability means that the message content becomes understood in the same way by the senders and the receivers. This concerns both information representation and messaging sequences. Pragmatic interoperability in turn captures the willingness of partners to perform the actions needed for the collaboration. This willingness to participate refers both to the capability of performing a requested action, and to policies dictating whether it is preferable for the enterprise to allow that action to take place.

Second, the collaboration management goal is that in future, individual users, enterprises or public organizations can easily compose new services from open service markets, or establish temporary collaborations with complex peer relationships. Furthermore, these contract-governed collaborations can be managed by their partners. All this is supported by a global infrastructure with facilities for interoperability control and contract-based community management (establishment, control and breach recovery) among autonomous organizations; this infrastructure also takes responsibility of governing trust and privacy-preservation issues.

The Pilarcos architecture views inter-enterprise collaboration as a loosely-coupled, dynamic constellation of business services. The constellation is governed by an eContract that captures a business network model (BNM defines business processes and their interlinkage within the collaboration [1]), member services, and policies governing the joint behaviour [1], [2].

The Pilarcos architecture for the open service ecosystems (Figure 1) includes as actors or artefacts

1) participating enterprises, with their public business service portfolios exported [3];
2) business-domain governing consortia, with their public models of business scenarios and business models expressed as exported business network models (comprising of a set of business process descriptions and compulsory associations between roles in them, and governing policies about acceptable behaviour) [1];
3) a joint ontology about vocabulary to be used for contract negotiation, commitment and control [6], [7];
4) legislative rules to define acceptable contracts [7];
5) technical rules to define conformance rules over all categories of metainformation held as collaboration and interoperability knowledge [8];
6) infrastructure services to support partner discovery and selection, contract negotiation and commitment to new collaborations, monitoring of contracted behaviour of partners, and breach detection and recovery services; these services include trust aspects in decision-making on commitment and breaches [1], [2]; and
7) reputation information flow, collected from the past collaborations [4].

Figure 1 illustrates the ecosystem lifecycle. On the left, metainformation repositories and development flows are shown. The flows denote the publishing and exporting processes enlisted above as items 1 and 2. The repositories in particular contain public information about the available business network models, available services and reputation information about the available services. This information is stored to globally federated repositories, applying strictly specified structuring and conformance rules [8] created by the processes enlisted above as items 3, 4 and 5. The information is, in turn, utilised by the ecosystem infrastructure functions enlisted as item 6, e.g. service discovery and selection, eContracting functions, monitoring of business services and reporting of experience on the services when a collaboration terminates. These functions are further described below.

On the right, the lifecycle of independent collaborations is shown flowing from establishment to evaluation at the dissolution phase. The infrastructure functions provide support for the four phases of the collaboration: establishment, agreement, enactment and control, and evaluation.

*Service discovery and selection* supports the collaboration establishment phase. It is based on public business network

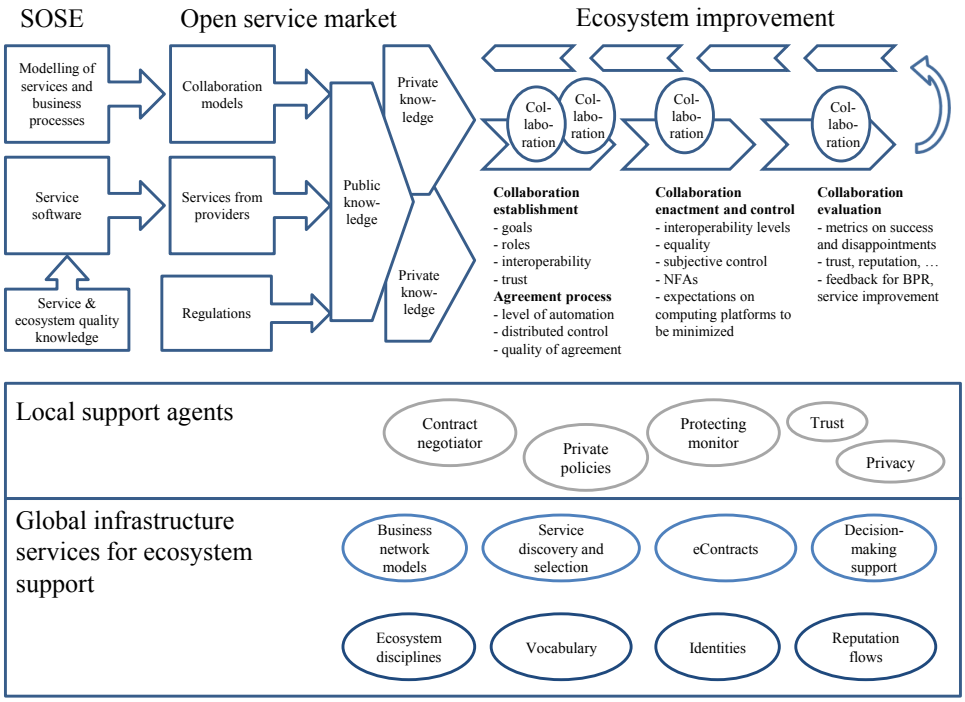


Figure 1. An overview of the Pilarcos open service ecosystem. Privacy-preservation architecture adds elements to modeling processes, quality of BNMs, quality of eContracts, decision-making support and ecosystem disciplines.

models describing the collaborations, and public service offers made by service providers [1], [6]. The business network models capture the best practices of a given field, and they are built from formally defined service types. The task of producing these models and types naturally falls to consortia and standardization bodies. Service selection includes automated static interoperability checking, which ensures that the service offers fit the model of the collaboration, and have terms that are compatible with other offers being selected into the proposed business network. As service discovery and selection is separate from contract negotiations, it can be done without access to sensitive information; this makes it possible to have this task implemented as a third-party service [1].

Automated *eContract establishment* supports the agreement phase of the collaboration [1]. The business network model and the proposed service offers to populate the roles in it are processed by an automated contract negotiation infrastructure, which is controlled locally by each collaboration partner. Contracts are based on templates specific to the collaboration model, and the terms of service provision given in service offers form the basis of negotiations. The negotiated eContract includes a model of the business process of the collaboration, as well as the finalized terms of service in the form of accepted service offers.

*Monitoring* supports the enactment and control phase of the collaboration in particular [2]. It is done subjectively by each collaborator to protect local resources, keep track of the progress of the collaboration, and to ensure that partners follow the collaboration model. The business process model and service provision terms set by the negotiated eContract form the specification of correct behaviour in the collaboration, which becomes relatively straightforward to monitor. In addition to the joint rules in the eContract, monitors also receive rules from local enterprise policies, therefore, monitors enable privacy enforcement as well.

*Experience reporting* has two roles. It supports the evaluation phase of the collaboration, although it also connects to the monitoring service during the enactment of the collaboration [4]. Moreover, the experience reporting forms the core of social control in the open service ecosystem. As contract violations are detected by monitors, they are published to other actors as well: it is important to directly react to privacy and data security violations (e.g. by reputation downgrading), in order to limit the damage that misbehaving actors can achieve in other collaborations. Against this background we can overlay the privacy challenges and threats involved.

### B. Privacy challenges in service ecosystems

As we have studied the threats in ecosystems [9], we found that the key questions to be asked are:

- Who can we trust sufficiently to exchange sensitive information with? Who can we collaborate with?
- What information can we expose to a trusted party?
- How can we define trust/exposure relation? How and why to commit to exposure? Who declares?
- How can we detect inappropriate exposure?

Therefore, we must extend the traditional definitions of privacy and the related concepts to involve not only humans, but also organisations and collaborations as independent actors. Moreover, the focal point in understanding privacy control is not in the agents themselves, but in the relationships between them: in exchange of benefit from the collaboration, how costly in terms of privacy the collaboration can be?

To illustrate the tradition of definitions on privacy, we chose two characterising definitions:

- Privacy denotes the persons right to be let alone [10]; and
- Privacy is the claim of individuals, groups, or institutions to determine for themselves, when, how, and to what extent information about them is communicated to others [11].

These definitions show the change on privacy-preservation approaches. In the first definition, the person is a subject on whom some acts can or should not be performed. In contrast, the second definition enhances the subject to cover organisations and groups as well, and changes their role to be a subject with decision power.

However, the definition does not yet explicate how the determined limits of privacy are controlled. In this paper, we adopt the definitions of privacy as follows:

- Privacy is the right of subjects to determine themselves for whom, for what purpose, to what extent, and how information about them or information held by them is communicated to others.
- A subject is a person, social group organisation or organisational group.
- Privacy control is the set of actions by which a subject makes decisions on refraining or involving in information exchange or sharing, is involved with privacy declaration management, and reacts on detected privacy violations.
- Privacy violation is circumstances where information is held or used in a way that breaches the privacy declaration by the information owning subject.
- Privacy declaration is an expression that gathers together rules on to whom, for what purpose, to what extent and how information can be made available.
- Privacy declaration management may involve the subject itself, other subjects due to hierarchical organization structures or agents governing the declarations on behalf of the subject and the ecosystem dynamically.

The ownership of the protected subjects or assets is not straightforwardly the creator of the information, nor is there necessarily a single owner. For example, in healthcare, owners of a created clinical statement (at least in Finland) are both the creating medical staff member and the patient. Different parts of this document may have different owners too, in terms of who can declare the usage of the document and for what purpose. For each ecosystem, it is necessary to make these rules explicit. Further, we must note that declarations by the owner can include implicit and delegated methods too, depending on the utilities provided by the ecosystem, organisation, default policies and technology.

These definitions must be considered against the set of assets or aspects that are considered needing privacy protection. We identify the following set of assets:

- *identity*: the exposure of subject's identity can be threatening and unnecessary for the purpose of the collaboration, and thus anonymity, pseudonymity and other techniques are preferable; this affects the collaboration contract nature by enforcing unidentified communication;
- *subject profile and context*: for example, the location of the subject could reveal sensitive information; in more general terms, these elements of information include metainformation about the subject itself and need to be protected;
- *collaboration relationships, service usage, behaviour, group membership*: these elements of information belong to metainformation relevant for the working of collaboration management, but also at this level, privacy must be preserved; and
- *privately held or created information*: this is the category of data traditionally considered as the target of privacy preservation solutions.

An important factor affecting the strength of privacy provisions are the associated control mechanisms. By control mechanisms we mean the framework that controls the sanctions against any subject that breaches the privacy or provisions set for specific data items. The sanctions may be directly implemented by the technical privacy enforcement mechanisms, or, they may be independent of the technical systems, e.g., in case of legal punishments. As the ecosystem structure already involves contractual structures, it is natural to expect the privacy control mechanisms to utilise the contractual control.

The presented Pilarcos ecosystem architecture addresses a number of privacy challenges [9] at each of its lifecycle steps. First, the populators are a central element in dealing with service offers and proposed contracts. It is essential, that the information made available for populators does not contain private information. The commitment decisions,

including trust-decisions and privacy-rule checking by each potential partner, must be protected from viewing (also as history chains). Thus, the publication of service offers must be carefully controlled by each organisation themselves, in order to prevent revealing of private enterprise policies, strategic network memberships, or preferences related to positioning on the marketplace.

For privacy preservation, the contract establishment phase concludes with a negotiation phase, in which each potential partner is able to make commitment decisions privately. At the negotiation phase, the to-be partners are already known, so decisions on revealing private information during the negotiation can be explicitly done.

As an additional aspect for the privacy-preserving qualities of the eContracts committed to, each organisation must utilise a mature set of policies guiding the commit-or-reject decision-making at the negotiation phase. This decision point is parallel with the decisions on suitability of the BNM type for the enterprise strategy and estimated risk-trust balance evaluated by the supporting trust management system.

Second, the eContracts govern mutual behaviour of participating services during collaborations. As eContracts are structured according to BNMs and carry associated policies, the nature of the BNM is a key in the privacy-concerned collaboration. For example, a BNM might require providing full access to all shared data with no revocation - clearly, that kind of collaboration would provide no privacy, and cannot be extended outside the data owning organisation.

For privacy preservation, the BNM design phase is critical. All expertise about good business practices, knowledge about privacy regulations and understanding of appropriate information flow patterns must be applied at that design step. Once a good design is reached, the monitoring facilities at each organisation are able to sufficiently enforce these rules.

Third, conforming to the eContract does not protect partners from revealing private information during the collaboration. The eContract rules focus on the viability of the collaboration, leaving it to the partners to determine whether they find themselves in non-acceptable situations.

For privacy preservation, enterprise policies must be monitored during the collaboration lifetime, and allowed to overrun eContract policies where privacy-preservation needs are more critical than the benefits earned from the collaboration. As this kind of decisions are often impossible to automate, this category of decisions belong to those that are either automatically refused or are directed for human intervention.

Fourth, the existence of the ecosystem repositories creates opportunities to detect information made available by any ecosystem member without the permission to use that information.

For privacy preservation, this metainformation must also be protected by declarations about its usage, and made available only for trusted infrastructure service providers with no other incentives in the ecosystem. The pressure for correct behaviour by partners must be based on threats of loosing their business. Especially, negative reputation information can be considered as a threat. Decisions on what part of reputation information can be revealed to the public should follow the rules defined by privacy-policies at each reputation network.

Finally, the automated binding mechanisms and relaxed matching of services create a risk of not noticing third-party elements in communication channels. While this can add privacy risks, the communication channel elements selection is agreed (or denied) by the organisations themselves. Thus, the trust towards these additional elements is equal to a systemic trust issue towards the infrastructure service providers.

For privacy preservation, the binding mechanisms, however, bring added value by providing room for third party components. This provides a practical declarative method for incorporating into communication channels tools for anonymisation and avoidance of identity-revealing cumulation of information from the use of several services.

### C. Related work

The need and development trend for more open business service ecosystems (or virtual breeding environments) is visible in many research and development activities in Europe, Asia and America, including projects like ECOLEAD [12], CrossWork [13], Pilarcos [3], FINeS cluster projects [14] and RMP project [15]. The main difference between virtual breeding environments (such as [12]) and Pilarcos type of open service ecosystems is that in the latter collaboration management does not rely on centralized control of the entire collaboration: participants remain autonomous and independent of the initiator of the collaboration. In addition, some of the management support needed can be offered as services by specialized third parties, rather than requiring one ultimately trusted actor to rule over everything.

Studies on privacy-preservation related questions are conducted in various fields. In the PRECIOSA project, privacy-preservation technology is integrated into co-operative systems to support collaboration between individual travellers, the operators of transport systems, and service providers [16]. The PRECIOSA project provides detailed insights into location privacy in mobile and ad-hoc networks, and contribute to the development of privacy-preserving vehicular communication systems. The PRISM project presents a framework for the protection of personal data in mobile networking, sensor networks, ubiquitous and context-aware computing. The central idea behind the framework is the integration of all privacy-critical functionality into a privacy-proof middleware, where proxy entities enforce the privacy legislation principles [17], [18]. The technologies

used in these projects can be applied in Pilarcos to enforce privacy-preserving functionalities.

In contrast to these typical examples from the related work, the Pilarcos ecosystem architecture goals are wider. It aims to provide a consistent mechanism to address several assets and privacy threat sources simultaneously. It also aims to bind together business level and technical level control to the same mechanisms, and automate it as far as possible to gain the expected flexibility on open ecosystems. The resulting mechanism has a number of characteristics, such as i) placing privacy enforcement responsibility on the ecosystem instead of on the subjects, still providing subjects with control facilities; ii) providing a "social control loop"; and iii) through trust management automatically forcing ecosystem members to apply the infrastructure mechanisms in an appropriate way.

## III. Privacy Preservation Architecture

We define the privacy-preservation architecture for open service ecosystems by i) definitions of privacy and privacy subjects, privacy-related actions of privacy controls, declarations and violations; ii) identification of privacy subjects and commitment to collaborative relationships between them, the collaborations becoming new privacy subjects themselves; and iii) ecosystem infrastructure functions involved with privacy declaration management, privacy discipline commitment, privacy enforcement by trusted infrastructure, and countermeasures against detected privacy violations. As i) and ii) are explained in the previous section, this section focuses on the four elements of ecosystem infrastructure for privacy preservation.

### A. Privacy Preservation Management

The privacy preservation management involves privacy subjects who define privacy policies to be used when interacting with their collaboration partners. The creation of privacy declarations requires supporting techniques, such as privacy-policy languages, interfaces, contracting policies and strategies, interfacing between users and administrators.

In the present architecture, several languages for privacy policies can be used in parallel to create privacy declarations. The languages found in related work have suitable expressive power, except that they rarely allow declaration of allowed use of private information to a limited purpose only. P3P [19], [20] is one of the most widely-deployed privacy-policy languages that are build into web browsers. However, they are the least expressive and has some semantic anomalies [21]. EPAL [22] is a privacy language designed by IBM to enforce privacy policies within the enterprise. EPAL and XACML [23] define semantics in terms of an authorization algorithm, making a utility extension difficult. Antón et al. [24] formalized privacy policies used by organizations but does not consider how these policies affect the design of organizational workflow.

We claim that the privacy-policy language used to create privacy declarations in ecosystems should be expressive and extensible with systematic semantics definitions, in order to enforce privacy declarations within collaboration enterprises.

Policy definition is clearly a challenge in complex scenarios, since the number of policies to be defined is large. Setting up policies is time-consuming and requires sufficient skills from the persons involved. This is even more emphasised in dynamic scenarios, where the privacy needs change over time, leading to a constant need to update the policies.

Due to the inherent difficulties related to the definition of policies, this challenge is especially relevant for scenarios, in which end-users are given responsibility for the policy definition. It is questionable, whether end-users have the willingness to sacrifice the required time and sufficient understanding to set-up and update their policies in a timely manner.

However, the duties of declaring policies can be eased by creating policy recommendations either within organisations or within groups or collaborations. Recommended policy sets can be, for example, inherited from domains of BNMs, national legislation or application domain. This information can be carried by ecosystem infrastructure services.

The subjects of these declarations can be learned from the behavioural patterns of the subject. For example, in mobile environments groups can automatically be based on people attending the same event. An important element in privacy declarations is that of not revealing identity in relationships with any or selected subjects. This rises technological requirements that are to be tested where privacy discipline commitment takes place.

Consistency of policy rules, which are set by a subject and the collaboration entered, does not need to be analysed beforehand. All necessary negotiation takes place either at commitment time or during collaborations. This is in parallel with other policies in Pilarcos [1].

### B. Privacy Discipline Commitment

During collaborations, privacy policies of subjects and the collaboration are simultaneously but independently enforced. The collaboration-wide privacy policies are declarations inherited from the BNM (and potentially incremented by partners). The discrepancies on major goals are then negotiated between presumably unbiased, professional BNM designers. However, not all aspects are necessarily covered by the policies in the BNMs, and at more detailed level discrepancies can happen, for example, because a document to be exposed would contain both the shared information and something restricted by a partner's privacy policy. In such a situation, the Pilarcos normal procedure is to be followed, overriding the contract and preserving the local policy.

The negotiation phase allows each partner to reject or agree the collaboration without exposing their private policies about the decision. Reasons for rejecting includes the

type of collaboration, partner identity, partner reputation, and strategy on committing to the collaboration load. Once committed to the collaboration rules governing privacy declarations of the collaboration itself, all information created by the collaboration becomes protected. This is necessary for being able to detect violations and trigger sanctions.

### C. Privacy Enforcement by Trusted Infrastructure

The detection of privacy risks is executed at the time of committing to a collaboration and for each step in the business network progress. Because humans tend to sacrifice their long-term privacy for short-term benefits [25], decision-making in open service ecosystems requires support from privacy-preservation functions.

The privacy-related decisions at this phase involve functions for testing whether the collaboration would hold information that combined with the intended new information exposures would reveal identity or other sensitive details. Existing techniques for this kind of precautions include unlinkability [26], unobservability [27] and l-diversity [28].

The privacy enforcement during collaborations is mainly performed by the trusted infrastructure, because in many scenarios, the actual enforcement of the policies is performed in systems that are not under the direct control of the owners of the processed private information. Therefore, it is required that the enforcement of the policies related to private information is monitored, so that any breaches against the policies can be detected. This is especially important in scenarios in which information is shared between parties based on an agreement or a policies, and remote parties are required to honour the policies or agreements with regard to policy enforcement. The privacy enforcement functionalities involved with the operational phase of the collaboration include techniques familiar from pseudonymity architectures [28], [29], location privacy [30] and communication unlinkability [31].

Common to these techniques is that proxies and additional protocol steps are required; these requirements can be passed to the communication channel configuration phases in the collaboration establishment, and thus the right technology elements to be utilised throughout the collaboration operation time can be enforced. A deviation from the selected protocol would cause a breach, and thus, lead to sanctions in the ecosystem.

The Pilarcos architecture includes local monitors to be used for all interactions from and to a services. Thus, the monitors can intercept (and stop) incoming requests that violate the privacy declarations at the contract level or cause a discrepancy against the receivers local policies, and outgoing information exchangers that violate or are at risk of indirectly compromising its local privacy declarations or those of the collaboration.

To understand the concept of the trusted infrastructure better, let us return to the idea of service ecosystem comprising of software-based application-level services representing subjects (i.e., individuals, groups and organisations) composed into collaborations, where exchange of potentially sensitive information takes place. Supporting the users and application elements resides first on a local platform with services for collaboration management, enterprise and user policies enforcement, monitoring and breach reporting. These local services utilise infrastructure services that are provided by external organisations trusted within the ecosystem. The infrastructure services govern contracts, supporting model repositories, and reputation flows. These platform and infrastructure services enable the ”correctness” of collaborations and the ”acceptability” of members in the ecosystem based on their ”decent” past behaviour.

The subjects themselves are involved only in declaring the privacy needs, negotiating and committing to those of the collaboration contract, and potentially terminating the collaboration due to a breach report. These declarations are forwarded to the local platform services. Eventually, all decision-making is supported by the underlying system, although in cases where policies are insufficient for automated decisions, the decision is relayed for human decision-making. Thus it is important to notice, that even on the privacy-preservation architecture view, the infrastructure has a *systemic trust* [32] position.

### D. Countermeasures

If parties that are involved in the process of private information and the enforcement of privacy policies and agreements do not honour the policies and agreements, it is important to ensure the possibility of taking efficient countermeasures against the party that is in breach of the agreement. Depending on the scenario at hand, the countermeasures can be technical, legal or even related to societal relations. The main purpose of the countermeasures is to de-incentivise behaviour which would be in breach with the privacy policies of individual parties or the agreements in force between them. The techniques have been discussed earlier in Section II.

Furthermore, all violations cannot be detected during the collaboration, but can be detected later by the involved partners. Privacy violation detection requires techniques at the individual, group, collaboration, or ecosystem level that catch a violating party when a privacy violation has already taken place. Information that has been leaked out can be detected by techniques like ”watermarking” [33] and DRM [34]. These techniques are used to solve the challenge: what can be done to avoid situations to reoccur, or to resolve the existing violation. For this, we extend the existing pattern to inter-enterprise collaborations governed by explicit contracts so that when privacy violations are detected, the ecosystem sanctioning mechanisms will be used.

## IV. Tools for Privacy in Open Service Ecosystem

As identified above, the sources of privacy violations can be categorized as follows:

- interaction patterns required in the collaboration enforce practices that violate privacy needs, even privacy-related regulations or best business practices;
- collaborating partners reveal their identities unnecessarily or accidentally to each other;
- in cases where collaboration benefits from extra information, partners provide it despite that the organisational policies deny the use of the extra information;
- information available in a collaboration can be combined so that "secrets" can be revealed; this can also take place when a subject participates several collaborations in sequence or in parallel; and
- a subject purposefully reveals information or metainformation owned by another subject.

The above threats are or will be addressed by the Pilarcos privacy-preservation architecture with the following tools.

First, an analysis tool is designed to validate BNMs against interaction patters to root out known anti-privacy patterns and to educate designers to use preferable patterns. This tool is an extension of basically any business process design tool that allows verification of the processes. The difference to traditional business process tools is that BNMs involve more than one process, and require selected roles to be simultaneously associated to a same subject.

When such analysis step is added to guard the BNM repository in the architecture, rooting out unacceptable BNM publications in the repository, all contracts are thus enforced to utilise only acceptable patterns. This is rather optimistic though, considering the business pressures, but at least it adds an opportunity for good design.

For each BNM it must be possible to specify and verify privacy goals of business process in individual steps. These steps include a set of information exchanges between collaborators and must be tagged for the intended and acceptable usage of the exchanged information. The question is not about the actual contents of the messages, but the intended use of the information exposed. This distinction is useful in evaluating which policy enforcement tasks can be carried out by ecosystems and which require human agents, as the privacy preservation facilities try to track the purpose for which data is used within a business process. A piece of information may be acceptable for statistics usage but not for identifying a user.

While designing such a tool the competing interests of privacy and utility should be reconciled with the principle of minimum necessary disclosure: disclose the minimum information necessary to achieve the utility goal [35]. It requires a commonly shared, clearly defined and comprehensive architecture for reducing high-level privacy preservation requirements to specific operating guidelines that can be applied at small steps in a business process or organizational workflow.

Useful privacy preservation tools include the two sets of general technical tools presented by Barth et al. [35]. The first set of tools concerns workflow design. Algorithms are used to check whether a workflow design achieves the privacy and utility goals, assuming all agents are responsible. A minimality condition is formulated on workflow designs that makes precise the informal principle of minimum necessary disclosure advocated by the privacy community. However, this set of tools consider only qualitative notions of privacy and utility, i.e. tagged privacy information. Quantitative notions for controlling information flows are missing. The other set of tools is for finding agents accountable for policy violations. It is aided by a human agents as auditors to access information through an oracle, but it aims to minimize the number of oracle calls. In addition, a heuristic for identifying irresponsible agents by analysing the audit log for suspicious events is also used as a tool for privacy management.

Second, we need to enhance the existing infrastructure solutions in Pilarcos to support privacy-preserving decision-making, according to the architectural requirements discussed in Section II-B. For this, we adopt the following principles:

- Privacy level expectations are always related to trust towards the partners or usages. The expectations balance between the available benefits from the collaboration and information exchange and the costs of relaxing privacy restrictions. However, the decisions on privacy requirements and trust are separate, both coexisting in the system independently and both mechanisms simultaneously involved in commitment decisions for collaborations.
- Privacy decisions are always subjective. The decisions must be adjustable and covered for the subject's and ecosystem's whole lifecycle.
- Trust management of Pilarcos provides a social control loop to create a "punishment" for those who forward private information.

Here, the existing decision-making and monitoring mechanisms are used, but policies and models guiding them require new, parallel refinements for the privacy-preservation goals.

Third, the communication channel types made available in the ecosystem members' platforms must include alternatives for supporting privacy and identity protection. The mechanisms are already in place for utilising these alternative channels when they become publicly available. A lot of standardisation work is to be seen before this become reality.

Fourth, techniques for detecting information leakage, such as watermarking [33] and DRM [34], must be used to avoid situations to reoccur or to resolve violations of privacy-policies. The infrastructure solutions in Pilarcos must ensure

the possibility of taking efficient countermeasures against the party that is in breach of the agreement.

The main future challenge is in the violation detection after collaborations have terminated.

## V. Conclusion

As part of research on privacy-preservation in open service ecosystems we have studied architectural requirements in the Pilarcos architecture of CINCO group (cinco.cs.helsinki.fi). We believe the study represents the demands of the present networked world sufficiently to provide us with solutions that are usable in other environments as well. Besides the prototype Pilarcos ecosystem approach, we have considered the applicability of the presented privacy-preservation architecture to cloud computing environment, social networking, or flexible usage of mobile solutions in everyday life.

We have found that main shortcomings in the present privacy architectures and techniques fall in to the following categories:

- Too restricted definition of privacy itself.
- Too weak concept of mutual commitment to privacy preservation.
- Lack of utilities for detecting privacy violations.
- Lack of tools to ensure committed collaborations become secure in privacy sense, too.

As contribution, we enhance the scope of privacy concerns from individuals to organisations, groups and collaborations. We also find privacy preservation to relationships between subjects, keep privacy as a subjective right, and define explicit rules for violations of privacy. On this basis we are able to develop utilities and tools to address the privacy violations and quality of collaboration commitments. Privacy-preservation tools embedded in ecosystems must be able to create privacy declarations, maintain trusted context and detect privacy anti-patterns and violations.

This paper serves as a roadmap for adding privacy-preservation functionality to the Pilarcos ecosystem architecture. It focuses on the additions required to the ecosystem processes and tools needed for privacy-preservation in the ecosystem. We have used the Pilarcos ecosystem as an example, but the results are applicable to other environments as well. The founding additions are elements for modelling processes, quality of BNMs, quality of eContracts, decision-making support and ecosystem disciplines.

As the main architectural structures for privacy-preservation needs are already presented in the Pilarcos architecture, in the future, we can concentrate on forwarding areas of BNM analysis, privacy declaration management processes and languages, privacy-aware communication channel architectures, and post-collaboration detection of leaked information.

## References


[1] L. Kutvonen, J. Metso, and S. Ruohomaa, "From trading to eCommunity management: Responding to social and contractual challenges," *Information Systems Frontiers (ISF) - Special Issue on Enterprise Services Computing: Evolution and Challenges*, vol. 9, no. 2–3, pp. 181–194, Jul. 2007. [Online]. Available: http://dx.doi.org/10.1007/s10796-007-9031-x

[2] L. Kutvonen, T. Ruokolainen, and J. Metso, "Interoperability middleware for federated business services in web-Pilarcos," *International Journal of Enterprise Information Systems, Special issue on Interoperability of Enterprise Systems and Applications*, vol. 3, no. 1, pp. 1–21, Jan. 2007. [Online]. Available: http://www.idea-group.com/articles/details.asp?id=6597

[3] L. Kutvonen, T. Ruokolainen, S. Ruohomaa, and J. Metso, "Service-oriented middleware for managing inter-enterprise collaborations," in *Global Implications of Modern Enterprise Information Systems: Technologies and Applications*, ser. Advances in Enterprise Information Systems (AEIS). IGI Global, Dec. 2008, pp. 209–241. [Online]. Available: http://www.igi-global.com/reference/details.asp?id=9648

[4] S. Ruohomaa and L. Kutvonen, "Making multi-dimensional trust decisions on inter-enterprise collaborations," in *Proceedings of the Third International Conference on Availability, Security and Reliability (ARES 2008)*. Barcelona, Spain: IEEE Computer Society, Mar. 2008, pp. 873–880. [Online]. Available: http://dx.doi.org/10.1109/ARES.2007.123

[5] T. Ruokolainen and L. Kutvonen, *Framework for Managing Features of Open Service Ecosystems*. IGI Global, 2011, in press.

[6] ——, "Service Typing in Collaborative Systems," in *Enterprise Interoperability: New Challenges and Approaches*, G. Doumeingts, J. Müller, G. Morel, and B. Vallespir, Eds. Springer, Apr. 2007, pp. 343–354. [Online]. Available: http://dx.doi.org/10.1007/978-1-84628-714-5_32

[7] J. Metso and L. Kutvonen, "Managing Virtual Organizations with Contracts," in *Workshop on Contract Architectures and Languages (CoALa2005)*, Enschede, The Netherlands, Sep. 2005. [Online]. Available: http://www.dstc.edu.au/Research/Projects/coala/2005/

[8] T. Ruokolainen and L. Kutvonen, "Managing Interoperability Knowledge in Open Service Ecosystems," in *Enterprise Distributed Object Computing Conference Workshops*, V. Tosic, Ed., 2009, pp. 203–211, in conjunction with the EDOC'09 conference. [Online]. Available: http://dx.doi.org/10.1109/EDOCW.2009.5331993

[9] P. Moen, S. Ruohomaa, L. Viljanen, and L. Kutvonen, "Safeguarding against new privacy threats in inter-enterprise collaboration environments," University of Helsinki, Department of Computer Science, Tech. Rep. C-2010-56, 2010.

[10] S. D. Warren and L. D. Brandeis, "The right to privacy," *Harvard Law Review*, vol. 4, no. 5, pp. 193–220, dec 1890.

[11] A. F. Westin, *Privacy and freedom*. New York: Atheneum, 1967.

[12] R. J. Rabelo, S. Gusmeroli, C. Arana, and T. Nagellen, "The ECOLEAD ICT infrastructure for collaborative networked organizations," in *Network-Centric Collaboration and Supporting Frameworks*, vol. 224. Springer, 2006, pp. 451–460.

[13] N. Mehandiev and P. Grefen, Eds., *Dynamic Business Process Formation for Instant Virtual Enterprises*, ser. Advanced Information and Knowledge Processing. Springer, Jun. 2010.

[14] M.-S. Li, M. Kürümlüoğlu, M. Mazura, and R. van den Berg, Eds., "Future Internet Enterprise Systems (FInES) cluster position paper," EU FP7 FInES cluster, Tech. Rep., Sep. 2009. [Online]. Available: http://cordis.europa.eu/fp7/ict/enet/fines-positionpaper_en.html

[15] T. Kang and L. Kagal, "Enabling privacy-awareness in social networks," *Intelligent Information Privacy Management Symposium at the AAAI Spring Symposium 2010*, 2010.

[16] Z. Ma, F. Kargl, and M. Weber, "Measuring long-term location privacy in vehicular communication systems," *Comput. Commun.*, vol. 33, pp. 1414–1427, July 2010.

[17] G. V. Lioudakis, E. A. Koutsoloukas, N. L. Dellas, N. Tselikas, S. Kapellaki, G. N. Prezerakos, D. I. Kaklamani, and I. S. Venieris, "A middleware architecture for privacy protection," *Computer Networks*, vol. 51, no. 16, pp. 4679–4696, 2007.

[18] G. Lioudakis, F. Gogoulos, A. Antonakopoulou, A. Mousas, I. Venieris, and D. Kaklamani, "An access control approach for privacy-preserving passive network monitoring," in *Internet Technology and Secured Transactions, 2009. ICITST 2009. International Conference for*, 2009, pp. 1–8.

[19] P. Samarati, "Protecting respondents' identities in microdata release," *IEEE Transactions on Knowledge and Data Engineering*, vol. 13, pp. 1010–1027, 2001.

[20] L. F. Cranor and L. Lessig, *Web Privacy with P3p*. Sebastopol, CA, USA: O'Reilly & Associates, Inc., 2002.

[21] A. Barth and J. C. Mitchell, "Enterprise privacy promises and enforcement," in *Proceedings of the 2005 workshop on Issues in the theory of security*, ser. WITS '05. New York, NY, USA: ACM, 2005, pp. 58–66. [Online]. Available: http://doi.acm.org/10.1145/1045405.1045412

[22] G. Karjoth and M. Schunter, "A privacy policy model for enterprises," in *Proceedings of the 15th IEEE workshop on Computer Security Foundations*, ser. CSFW '02. Washington, DC, USA: IEEE Computer Society, 2002, pp. 271–281. [Online]. Available: http://portal.acm.org/citation.cfm?id=794201.795180

[23] A. Anderson, "A comparison of two privacy policy languages: Epal and xacml," Mountain View, CA, USA, Tech. Rep., 2005.

[24] A. I. Antón, J. B. Earp, and A. Reese, "Analyzing website privacy requirements using a privacy goal taxonomy," in *Proceedings of the 10th Anniversary IEEE Joint International Conference on Requirements Engineering*, ser. RE '02. Washington, DC, USA: IEEE Computer Society, 2002, pp. 23–31. [Online]. Available: http://portal.acm.org/citation.cfm?id=647648.760605

[25] A. Acquisti and J. Grossklags, "Privacy and rationality in individual decision making," *IEEE Security and Privacy*, vol. 3, pp. 26–33, January 2005. [Online]. Available: http://portal.acm.org/citation.cfm?id=1048715.1048819

[26] S. Steinbrecher and S. Köpsell, 2003, ch. Modelling Unlinkability, pp. 32–47. [Online]. Available: http://www.springerlink.com/content/dxteg659uf2jtdd7

[27] A. Pfitzmann and M. Köhntopp, "Anonymity, unobservability, and pseudeonymity - a proposal for terminology," in *International workshop on Designing privacy enhancing technologies*. New York, NY, USA: Springer-Verlag New York, Inc., 2001, pp. 1–9.

[28] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "L-diversity: Privacy beyond k-anonymity," *ACM Trans. Knowl. Discov. Data*, vol. 1, March 2007.

[29] N. Li, T. Li, and S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity," in *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on*, 2007, pp. 106–115.

[30] A. R. Beresford and F. Stajano, "Mix zones: user privacy in location-aware services," *Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second IEEE Annual Conference on*, pp. 127–131, 14-17 March 2004.

[31] D. L. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Commun. ACM*, vol. 24, no. 2, pp. 84–90, 1981.

[32] D. H. McKnight and N. L. Chervany, "What is trust? a conceptual analysis and an interdisciplinary model," in *AMCIS 2000 Proceedings*, no. 382, 2000. [Online]. Available: http://aisel.aisnet.org/amcis2000/382

[33] P. Peng, P. Ning, and D. S. Reeves, "On the secrecy of timing-based active watermarking trace-back techniques," *Security and Privacy, IEEE Symposium on*, vol. 0, pp. 334–349, 2006.

[34] A. Becker, A. Arnab, and M. Serra, "Assessing privacy criteria for drm using eu privacy legislation," in *Proceedings of the 8th ACM workshop on Digital rights management*, ser. DRM '08. New York, NY, USA: ACM, 2008, pp. 77–86. [Online]. Available: http://doi.acm.org/10.1145/1456520.1456534

[35] A. Barth, J. Mitchell, A. Datta, and S. Sundaram, "Privacy and utility in business processes," in *Proceedings of the 20th IEEE Computer Security Foundations Symposium*. Washington, DC, USA: IEEE Computer Society, 2007, pp. 279–294. [Online]. Available: http://portal.acm.org/citation.cfm?id=1270382.1270658