Privacy as an integral part of the implementation of cloud solutions

Evangelia Kavakli

Assistant Professor, Cultural Informatics Laboratory, Department of Cultural Technology and Communication, University of the Aegean, University Hill, GR81100 Mytilene, Greece, kavakli@aegean.gr

Christos Kalloniatis

Assistant Professor, Cultural Informatics Laboratory, Department of Cultural Technology and Communication, University of the Aegean, University Hill, GR81100 Mytilene, Greece, chkallon@aegean.gr

Haralambos Mouratidis

Professor, School of Computing, Engineering and Mathematics, University of Brighton, U.K., H.Mouratidis@brighton.ac.uk

Stefanos Gritzalis

Professor, Information and Communication Systems Security Laboratory, Department of Information and Communications Systems Engineering, University of the Aegean, GR83200 Samos, Greece, sgritz@aegean.gr

Category: Research Paper

Privacy as an integral part of the implementation of cloud solutions

Abstract

Bridging the gap between design and implementation stages has been a major concern of designers, analysts and developers of information systems and a major aspiration of a number of Information System (IS) engineering approaches. Cloud computing exacerbates the strain on traditional IS engineering approaches that service-oriented computing has started. At the same time, recent research has argued about the importance of security and privacy in a cloud environment and highlighted a number of security and privacy challenges that are not present in traditional environments and need special attention when implementing or migrating information systems into a cloud environment. This paper contributes to this direction. Specifically, it presents a number of privacy-related cloud properties that analysts need to consider when designing privacy-aware systems in a cloud environment. Also it indicates a number of implementation techniques that can assist developers in assuring the respective properties.

Keywords

Cloud Computing, Privacy Properties, Implementation Techniques, Software Engineering, PETs

Privacy as an integral part of the implementation of cloud solutions

1. Introduction

In recent years Cloud Computing has become an attractive IT paradigm to a broad range of users, from small and medium-sized enterprises (SMEs) and public administrations to end-users. The great demand for cloud services from online users, along with the reduced operational costs that these offer, have motivated many organizations to consider implementing new services or migrating existing applications on the cloud. However, despite the positive characteristics of cloud service models such as reduced costs, enhanced availability, on demand data storage and computing power, there are major concerns related to information privacy mainly due to the distributed character of cloud architectures, the involvement of different stakeholders in the operation of cloud services, and the limited (lack of) users' control of their data.

The Cloud computing paradigm is based on three delivery models: Infrastructure as a Service (SaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). Each providing, virtualized and on demand resources (computing power, network and storage), application development platforms and software services, respectively. Each delivery model is considered as a separate layer that is depended from the others with IaaS being the foundation, PaaS building upon IaaS and SaaS building upon PaaS. As a result, any attack to any cloud service layer can compromise the upper layers [1]. The service model also dictates end users' scope and control over the computational environment. In general, the higher the level of support available from a cloud provider, the narrower the scope and control the cloud user has over the system. IaaS is the model that enables more direct control but also leaves the cloud service user responsible for the implementation of privacy measures. Still the IaaS provider will typically take responsibility for securing the

data centres, network and systems, and will take steps to ensure that its employees and operational procedures comply with applicable laws and regulations [2]. Thus, the cloud provider has an important role on managing and implementing security and privacy measures in all three levels of abstraction.

Privacy is also affected by the selected cloud deployment model. The deployment model denotes the management and disposition of computational resources, as well as the differentiation between classes of users. In a private or community cloud for example, the computational resources are exclusive to a single organization or to a number of 'trusted' organizations that have common privacy considerations, thus reducing perceived privacy risks. In a public or hybrid cloud resources are shared between multiple users. However, in all cases the same threats related to the nature of cloud computing apply and therefore, privacy protection measures still need be considered.

The scope of the paper is twofold. First, it contributes to the existing literature through the identification of cloud specific privacy properties. Privacy properties are quality characteristics that the cloud service must demonstrate and which affect users' privacy (in terms of controlling how personal data is gathered, stored, processed or disseminated). Second, it advances the state of the art in privacy engineering for cloud computing, by introducing a number of implementation techniques that assure privacy properties in a cloud environment. These represent standard privacy enhancing mechanisms, which are able to prevent or detect privacy breaches. Although further analysis is necessary to specify the sub-mechanism(s) that can realize each implementation technique in the context of a specific cloud solution, we claim that identifying, early at the design stage, the possible means that contribute to the satisfaction of users privacy needs, can lead to implementations where privacy is an integral part.

The remainder of the paper is structured as follows. Section 2 presents related work in the area of privacy requirements engineering, within traditional software engineering methods as well as cloud computing. Section 3 discusses the privacy properties attributed to the cloud taking into consideration the relative security concerns and cloud service model. Section 4 describes a number of implementation techniques that realize the aforementioned properties. From a methodological perspective, section 5 illustrates a privacy driven way-of-working for defining appropriate implementation solutions using an eHealth example. Section 6 discusses the relation between privacy and trust issues in the cloud and how privacy-oriented and trust-oriented approaches can work together to bring new solutions to protecting user data. Finally, section 7 concludes the paper.

2. Related Work

A number of researchers have focused on requirements engineering methods that support the elicitation and modeling of privacy issues during the early stages of software systems development. The majority considers privacy as a security constraint. For example, Secure Tropos an extension of Tropos methodology proposed in [3] employs the concepts of security constraint, and secure dependency in order to model and analyze security issues during the requirements engineering phase. Similarly, the SecReq approach introduced in [4] describes a systematic approach to derive security requirements from system security objectives. In [5] misuse cases are used in order to represent security threats and to identify ‘‘security use cases’’, i.e., countermeasures that mitigate the threats.

PriS on the other hand, is a requirements engineering method that focuses specifically on privacy [6 – 8]. It makes the distinction between eight technical privacy requirements (such as anonymity and unlinkability) and adopts the use of process patterns as a way to: (a) describe the ef-

fect of privacy requirements on business processes; and (b) facilitate the identification of the system architecture that best supports the privacy-related business processes.

Work on privacy patterns focuses on reoccurring privacy-related problems and how these can best be solved with proven solutions. For example the privacy patterns for web-based activity, described in [9] document how to convey privacy policies to end users during online interactions. In [10] a pattern language is proposed, containing 12 patterns for developing anonymity solutions for various domains including anonymous messaging, anonymous voting and location anonymity.

Another line of work relates to legal compliance. Islam et al. [11] use natural language patterns and make use of the Hohfeld legal taxonomy, to extract security requirements from laws and combine them with the ISO/IEC policies. Finally they trace the identified requirements into secure system design. [12] describes an approach for evaluating the legal compliance of existing security and privacy requirements, by establishing traceability links from requirements to legal texts.

Recent research works deal with security and privacy issues related to the cloud-computing domain. Some identify existing cloud technology vulnerabilities where faults can occur. For example, [13] demonstrates that attackers can exploit data duplication techniques to access customer data by obtaining hash code of the stored file. Side-channel attack can instantiate new VMs of a target virtual machine so that the new VM can potentially monitor the cache hosted on the same physical machine as described in [14]. Other works focus on security and privacy attacks that may lead to a misuse of information or resources. For instance, [15] argues that privacy threats differ depending on the type of cloud scenario and lack of user control, potential unauthorized secondary usage and data proliferation. [1] summarizes the vulnerabilities and threats reported in the literature and it presents some countermeasures that solve or improve the identified problems.

In addition, [16] suggest tools for supporting the analysis of security and privacy risks from different perspectives in order to make informed decisions during the migration of IT systems to the cloud.

From the above discussion, it is clear that security and privacy modeling, in a cloud context, is attracting the attention of the research community. However, even though it is acknowledged that privacy threats and associated privacy concerns differ depending on the application context, most methods deal with privacy as a single requirement. Failing to understand the different facets of privacy leads to difficulty in identifying which privacy measures should be deployed. The privacy properties identified in this paper and their correspondence to specific privacy implementation techniques aim to aid designers to deal with the complex relationship between organizational privacy requirements and appropriate design solutions.

3. Privacy-Oriented Properties

Although privacy is common concern in distributed information systems, additional privacy issues arise due to the nature of cloud computing. In order to identify those particular issues we take into consideration (a) the security threats and vulnerabilities related to the nature of cloud computing and (b) the cloud service model employed.

The main advantages of cloud computing, its ability to scale rapidly, store data remotely and share services in a dynamic environment, have also created a number of vulnerabilities in terms of data protection. These vulnerabilities are reflected in a number of security threats reported in [1, 17, 18]. In [19] we have compiled a comprehensive list of 14 cloud related threats and vulner-

abilities[1] indicating the cloud service model, to which they apply. These are briefly demonstrated in Table 1 in terms of their privacy implications.

Table 1. Cloud threats and vulnerabilities and associated privacy implications

Cloud Threat	Cloud Vulnerability	Cloud Service Model	Privacy implication
Abuse of Cloud Services		IaaS, PaaS	Abuse relative cloud anonymity to hamper fraud detection such as password cracking.
Data Breaches		IaaS, PaaS, SaaS	Unauthorized access to user data referred also as data leakage.
Data Loss		IaaS, PaaS, SaaS	Accidental or malicious deletion of user data.
Account Hijacking		IaaS, PaaS, SaaS	Gaining access to user credentials.
Denial of Service		IaaS, PaaS, SaaS	Preventing users to access their data.
	Insecure APIs	IaaS, PaaS, SaaS	Implications arising from the set of software interfaces consumers use to manage and interact with the cloud service.
	Malicious Insiders	IaaS, PaaS, SaaS	Misuse of sensitive user information from entities having authorized access to it.
	Shared Technology Issue	IaaS	Vulnerabilities or misconfiguration of the underlying infrastructure that compromises privacy of user data.
	Insufficient Due Diligence	IaaS, PaaS, SaaS	Lack of understanding of the privacy issues relating to cloud adoption.
	Privileged User Access	IaaS, PaaS, SaaS	Unlimited access to user data

[1] We adopt the ISO IEC 27000 2014 definition, whereby a threat is a potential cause of an unwanted incident, which may result in harm to a system or organization, whereas vulnerability is a weakness that can be exploited by one or more threats.

			by authorized system administrators.
	Regulatory Compliance	IaaS, PaaS, SaaS	Difficulty to ensure compliance with different privacy legislations in multiple jurisdictions.
	Data Location	IaaS, PaaS, SaaS	Loss of control as to "where" user data is resided and in effect on the regulatory framework that applies.
	Lack of Data Segregation	PaaS, SaaS	Incomplete isolation of different users' data.
	Insufficient Investigate Support	IaaS, PaaS, SaaS	Difficulty in examining the causes and the circumstances of a privacy violation incident

As shown in Table 1, different cloud service models are not equally affected by all the above threats and vulnerabilities. For example, some of them are more applicable to IaaS such as shared technology issues, due to flaws of the virtualization technology. Lack of data segregation, relates to ineffective isolation of re-deployable platforms or multi-customer applications and mainly affects PaaS and SaaS service models. Privileged user access applies to all service models, because is referred to the fact that certain employees have authorized access to the cloud's services and as a result customer's data due to the nature of their work (e.g., a system administrator). Similarly, abuse of cloud services mainly affects IaaS and PaaS models since it refers to the misuse of computing power. As a result, the privacy concerns need to be determined on a case-by-case basis and in relation to the nature of the cloud services in question.

All of the above threats and vulnerabilities represent potential circumstances that may lead to misuse of information or resources. However, in order to deal with these circumstances, it is important to identify the privacy-related properties that are affected by each threat or vulnerability.

The aim of this section is twofold. Firstly, to identify and describe the privacy related properties associated to respective cloud threats and vulnerabilities. Secondly, to indicate the correspondence between privacy properties and cloud service models thus assisting stakeholders to decide which privacy properties need to be considered in order to satisfy their privacy needs on different cloud service model. The proposed set of privacy properties has been based on the European Commission Draft Report on Security Issues in Cloud Computing [20] as well as on our previous work presented in [3, 6, 21 – 25]. For each property, we provide a brief explanation and describe how it might affect user privacy. It should be noted that some properties are also relevant to 'traditional' distributed systems (e.g., accountability). In all cases however, the focus of the discussion bellow is on how these concepts are understood within the context of cloud computing.

3.1 Isolation

Multi-tenancy and shared resources are defining characteristics of cloud computing. Isolation refers to all these mechanisms aiming to the complete seal of user's data inside the cloud computing environment, and applies to the underlying multi-tenant architecture (IaaS), re-deployable platforms (PaaS) and multi-customer applications (SaaS).

The probability of isolation failures in cloud computing is considered very high [26], as isolation might be affected by the following threats and vulnerabilities: Abuse of Cloud Services, Account Hijacking, Data Breaches, Insecure APIs, Malicious Insiders, Shared Technology Issue, Privileged User Access and Lack of Data Segregation.

The impact of isolation failure on privacy protection is also very high, since it poses the risk of disclosure of personal identifiable information, thus it is an important privacy-related property.

3.2 Provenanceability

Cloud provenance studies the history of cloud resources activity [27]. Provenanceability provides awareness of what goes on in the back-end physical server (i.e., virtual and physical machines). Provenance data related to a cloud include customer identification information; information about the VMs and PMs where data is processed; and intra-cloud communication between different PMs within a cloud, transfers of data across VMs and PMs located across different geographies in a cloud.

Provenanceability is linked to tracing the origins of security and privacy violations of an entity [28]. Using provenance data one can check whether a file has been accessed by some malicious party using the same VM; whether a cloud administrator with full access rights tampering the file from anywhere inside the cloud; attest the integrity and enforcement of cloud resource policies; detect data transfer across geographic boundaries [29]. Provenanceability therefore, relates to the following vulnerabilities: Malicious Insiders, Privileged User Access, Regulatory Compliance, Data Location and Insufficient Investigate Support.

At the same time, the accumulation of provenance data (which potentially relate to sensitive information) might constitute a potential privacy violation in the case it is exploited in a malicious manner. Hence the cloud storage provider should ensure appropriate security for provenance itself.

3.3 Traceability

Traceability refers to the ability to track users' activity by means of recorded data and is an essential property of SaaS applications providing a trace of how user data is generated, used, stored and shared. Traceability enables users to trace the physical location of their data and to verify that they are processed according to their collection purpose. In addition, it can be helpful

in case of accidental deletion of user data. This property is matched with the following cloud threats and vulnerabilities: Data Loss, Data Breaches, Malicious Insiders, Regulatory Compliance and Data Location.

A main privacy concern linked to traceability is the privacy 'right to be forgotten', i.e., that no third parties are able to access personal data after being deleted. In fact several cases have been reported with respect to privacy violation due to improper data deletion (documents, photos, etc.). At the same time, traceability protects users' privacy through the ability of tracing data among the data repositories and reassuring that the data have been completely deleted or maintained invisible and anonymized after their deletion[2].

3.4 Intervenability

Intervenability refers to the fact that, users should be able to have access and process their data despite the cloud's service architecture. A cloud vendor may rely on other provider's subcontractor services in order to offer its services. That should not be an obstacle for users to intervene[3] with their data in case they suspects that their privacy is violated by the subcontractors. In fact cloud vendors must be able to provide all the technical, organizational and contractual means for accomplishing this functionality for the user including all respective subcontractors that the vendor cooperates and interrelates [20]. The same applies for the situation that a cloud vendor or the subcontractors are bankrupted and client's data are moved to another provider. This concept is matched with the cloud vulnerabilities of Insufficient Due Diligence, and Data Location.

[2] In some cases, certain cloud service providers apply retention policies as far as data are concerned. That means that for several reasons, that are stated inside the contract between the cloud provider and the client, the data remain at rest after the clients deletion request for some time and are strictly accessed form specific personnel and only for certain purposes.

[3] Intervention in this context includes access, rectification, erasure, blocking and objection

3.5 Accountability

According to the accountability property, cloud providers should be able to provide at any given time information about their data protection policies and procedures or specific cloud incidents related to users' data, irrespective to the cloud service model used and relates to all cloud threats and vulnerabilities. The cloud architecture is a complex type of information system. In terms of management and audit controls, this could result in very difficult manageability of the protections mechanisms and incidents. In case of a privacy violation, a cloud provider should be able in any given time to provide information about what, when and how an entity acted and which procedures were followed to tackle it [20].

3.6 Relation between privacy properties and cloud service models

As already mentioned, not all cloud service models suffer to the same extent from all threats and vulnerabilities. As a result different cloud models should place emphasis on providing appropriate measures for addressing different privacy issues in different levels of abstraction.

Table 2. Matching Privacy Properties with Cloud Threats, Vulnerabilities and Services Models

	IaaS	PaaS	SaaS	Cloud Threats	Cloud Vulnerabilities
Isolation	x	x	x	Abuse of Cloud Services Account Hijacking Data Breaches	Insecure APIs Malicious Insiders Shared Technology Issue Privileged User Access Lack of Data Segregation
Provenanceability	x				Malicious Insiders Privileged User Access Regulatory Compliance Data Location Insufficient Investigate Support
Traceability			x	Data Loss Data Breaches	Malicious Insiders Regulatory Compliance Data Location

Intervenability	x	x	x		Insufficient Due Diligence Data Location
Accountability	x	x	x	Abuse of Cloud Services Data Breaches Data Loss Account Hijacking Denial of Service	Insecure APIs Malicious Insiders Shared Technology Issue Insufficient Due Diligence Privileged User Access Regulatory Compliance Data Location Lack of Data Segregation Insufficient Investigate Support

Table 2 presents the correspondence between the identified privacy properties and the three cloud service models. Using this table, analysts can identify the privacy properties relevant to their system, and how these might constraint the design of the information system on a cloud environment.

4. Implementation Techniques

Implementation techniques are proactive measures aiming to promote effective privacy protection, which might be technically based, or in the form of contractual assurances. In this section we present a brief overview of standardized privacy solutions and research in progress aiming to prevent or detect privacy breaches.

4.1 Boundary protection techniques (Firewalls)

A firewall is a boundary protection mechanism that enforces access control policies and filtering rules in network environments, controlling the flow of information into or out of an interconnected system. In the cloud environment, firewalls tackle the problem of resource isolation from a networking perspective. Recent editions of firewalls are implementing intrusion detection and

prevention inside their core functions, which is in support of privacy preservation [30]. Of special interest are virtual firewalls that achieve isolation between virtual machines inside a virtual network through appropriate filtering of network data. Furthermore, they provide important logging function thus assisting analysis and detection of malicious traffic that is sent to and from a virtual machine through the router.

4.2 Hypervisor Hardening, Language, Sandbox, Virtual machine, OS – Kernel, and Hardware based Isolation

These techniques also implement access control on computational and storage resources providing logical isolation between different entities inside the cloud. Two types of isolation are implemented, software and hardware based isolation. The former is achieved by the first five techniques which aim to seal all the procedures, operations and data flows through the installation of multiple isolation layers using different programming techniques. For example, language based isolation ensures that programs can only access appropriate memory locations and that control transfers happen to appropriate program points. Sandbox based isolation creates confined execution environments for running untrusted programs on the same machine. OS-Kernel based isolation enforces policies that are required for isolation between applications. Virtual machines provide virtual isolated platforms for running operating systems. Hypervisor hardening secures the hypervisor management console used to configure and control all aspects of virtual machines, also monitoring administrators' operations. On the other hand, hardware based isolation is achieved through hardware controls that grant secure direct hardware access to virtual machines [31].

4.3 Encryption

Encryption mechanisms are used in order to ensure the secrecy of important information inside the cloud environment (including data kept for monitoring and tracking purposes) and to avoid inappropriate information disclosure [24]. Encryption techniques are implemented in various areas of the cloud, in order to encrypt data while it is being transmitted, stored or processed ([29, 32]). Ideally, a cloud customer should encrypt its data before moving it to a public cloud. If customers cannot do this themselves, then encryption should be outsourced to a security service offered by the cloud provider.

4.4 Privacy Policies and Contracts

Appropriate privacy policies and contracts can also be used in order to assure client's interest in terms of privacy protection. Machine-readable policies associated to personal information (including provenance data), are user preferences or conditions about how that information should be treated (for example, that it is only to be used for particular purposes, by certain people or that the user must be contacted before it is used).

Permitting users to state preferences for the management of their personal information and take account for this maximizes user control. Contractually fixed agreements (such as Security Service Level Agreements (SSLA) sometimes also referred to as Protection Level Agreements (PLA)), on the other hand, form the legal obligation of the cloud service provider. They include contractually fixed security restrictions, compliance checks, as well as security information and event management [33]. Cloud users must be very careful about the terms and conditions of the service they are using in order to ensure that their privacy is not violated in case of an incident or a situation that needs to be cleared, e.g. data hosting in foreign countries, what happens in case the cloud provider is bankrupted etc. [6, 20, 29].

4.5 Forensics

Cloud forensics is a subset of network forensics. The term was first introduced in [34], to designate the need for digital investigation in cloud environments. Forensics is meant to preserve the privacy of users from being exploited by detecting privacy violation incidents and identify who is accountable for them. Forensics is based on the appropriate collection of provenance data by cloud service providers. Whilst IaaS customers enjoy relatively unfettered access to the data required for forensic investigations, SaaS ans PaaS customers may have little or no access to such data. Therefore appropriate terms regarding forensic investigations should be included in SLAs [17, 28, 29, 35, 36]. On the other hand, the process of extracting data evidences raises privacy concerns in itself, making the balance between forensics and privacy a challenging issue [37].

4.6 Identity and Access Management (IAM)

In this category fall technologies which, combined or individually, protect the client's privacy through a solid framework that prevents unauthorized access to resources. According to [38] effective management of identities includes the following processes: Identity provisioning/deprovisioning; Authentication and federation; Authorization and user profile management; and Support for compliance. IAM employs several mechanisms including identity-based cryptography, federal identity management and role-based access control which can be accomplished with the use of formal internet standards such as SAML, XACML and SPML. Privacy-preserving/enhancing identity-management aims to maximize the control that individuals have over their identity information and to minimize the identity information that individuals have to release to the system. Anonymous credential systems and user-centric systems are example privacy-preserving identity-management systems [21, 39, 40].

4.7 Data tracking

Data tracking techniques monitor operations on data files at different levels of granularity (from the application level to the PM level) as well as file transfers both inter-cloud and intra-cloud. They capture and log for example, who, what (file operation), where, and at what time a file is accessed, or a table in a database were modified. Inquiring these logs allows users to verify properties of their data and increase awareness of its lineage, for example, in order to check whether their data is deleted or where it is located [29, 39]. Integrity and confidentiality of data traces should be ensured in terms of appropriate mechanisms in order to avoid unauthorized access (data leakage) or even tampering of log files from privileged users.

4.8 Process identification and validation

Complementary to the data-centric view described above these techniques monitor the processes that generate and modify data, aiming to capture provenance across applications, to ensure that the outcomes are reliable or that privacy is not violated by malicious processes [29]. Similar to data tracking, process-related data should also be protected in order to prevent data leakage and assure that privileged users will not tamper the contents of log files.

4.9 Privacy preserving data mining

In many situations service providers are using collected data from the users, e.g. data traffic, search history, configurations, in order to examine them and make a customer profile for marketing purposes. However, in order to protect users' privacy, such procedures should provide basic anonymization, de-personalising information prior to analyzing or transferring it across machines [41].

Privacy-preserving data mining techniques may also be used to mine the union of databases of different customers. In this case, the only information revealed to either of the database owners about the other's data should be the minimum amount of information that could possibly be provided by the customer for the service to be operable [42].

4.10 Monitor and Audit

Auditing is the process of tracing and logging significant events that could take place during a system run-time. It can be used for analysis, verification and validation of security measures to achieve overall security objectives in a system. Audit events can include, for example, password changes, failed logons, or failed accesses related to information systems, administrative privilege usage etc. Auditable events may also be dictated by applicable laws, directives, policies, regulations, and standards. Monitor and audit procedures are incorporated into security tools and help protecting client's privacy and provide information as to who is accountable for an event inside the cloud environment [43].

4.11 Matching Privacy Properties with Implementation techniques

The techniques described above fall in the category of privacy-enhancing technologies, aiming to prevent or detect privacy violations. In terms of prevention they aim to secure (IAM, encryption, firewalls, hypervisor hardening, isolation mechanisms) or minimize the amount of personal data being held or transferred (privacy preserving IAM, privacy preserving data mining). Furthermore they can provide individuals with control over their personal information that is being held in terms of privacy policies and contracts. With respect to detection they focus on the collection and analysis of historical data (IAM, data tracking, process identification, forensics, monitor and auditing). Sometimes a solution may address some concerns whilst raising new ones.

For example, detection solutions require the collection and analysis of additional data which might also contain or relate to personal information, which raises concerns with respect to who can access this data. Thus, additional prevention solutions should be used in combination.

In order to make an informed decision regarding which PETs should be used in a specific context one needs to understand the relationship between privacy needs and the capabilities of the implementation solution. To this end, the matching presented in Table 3, provides an initial step on how to bridge the gap between the desired privacy-related cloud properties and the technologies that might be used to assure these properties.

Table 3. Matching Privacy Properties with Implementation techniques

	Isolation	Provenanceability	Traceability	Intervenability	Accountability
Boundary protection techniques	x				x
Encryption	x	x	x		x
Hypervisor Hardening	x				x
Isolation techniques	x				
Privacy Policies and Contracts		x	x	x	
Forensics		x			x
Identity and Access Management	x	x			x
Data Tracking Techniques		x	x		
Process identification and validation		x			
Privacy Preserving Data Mining					x
Auditing					x

5. The EU eHealth example

The aim of this section is to illustrate how analysis of user privacy needs can guide the identification of an implementation solution, using an eHealth example described in [44].

The purpose of the system into consideration is to provide shared access and information about patients of a specific region, via a cloud infrastructure, to physicians and patients. The data to be moved to the cloud are:

- Physician's information (including address and contact details)
- Patients' generic health records (not relating to specific illnesses)
- General health information not related to any specific patient as well as details and locations of medical centers and pharmacies.

The cloud solution chosen by the consumer is a public cloud according to the IaaS service model where all data would be stored in public servers. The application will be made accessible to end users (physicians and patients) according to the SaaS model. The data to be shared and especially patients' health records fall in the category of sensitive information so the cloud provider should realize the appropriate technical means to assure the privacy of user data. Another main user concern was the issue of inter-border data flow. In particular, users wanted assurance that all data uploaded should be stored in servers inside the European Union and no data processing should be performed outside the European Union.

Inter-border data flow corresponds to two of the cloud vulnerabilities mentioned in Table 1 namely, Data Location and Regulatory Compliance. Both these issues can be extremely difficult to ascertain due to the dynamic nature of cloud computing. To guarantee that data will remain within the EU at all times the cloud provider should be able to derive the history of data objects, to assure that is, the privacy property of *provenanceability*. The use of non-dedicated-servers on

the other hand, raises concerns relating to Shared Technology Issues at the IaaS level and of Lack of Data Segregation at the SaaS level which in turn relates to the privacy property of *isolation*.

As a result the design of the cloud solution should provide appropriate mechanisms in order to assure both properties.

Based on the correspondence between privacy properties and implementation techniques provided in Table 3, a number of techniques can be applied however, application of each technique should be considered in combination with the cloud service layer. For example, since the solution combines the IaaS and SaaS models encryption mechanisms in both layers should be considered. The above example, though not exhaustive indicates how focusing on user needs can guide the definition of an appropriate implementation solution that addresses the associated privacy issues.

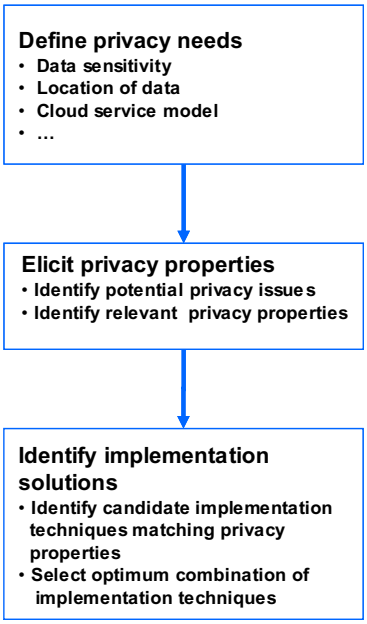

Figure 1. Privacy driven definition of implementation solution

In particular, the proposed way-of-working (shown in Figure 1) consists of the following steps:

1. *Define privacy needs of users*. To achieve this one should take into consideration among others the sensitivity of the data to be stored in the cloud, the privacy regulations that the

system should comply to (related also to the data location and cross-border data transfers) and the chosen cloud service model. Based on this analysis one can identify the privacy concerns of the users (e.g., transborder flaw and unauthorized access to sensitive information).

2. *Elicit related privacy properties*. In the proposed approach this is achieved by identifying the potential threats and vulnerabilities that hinder the satisfaction of the users needs (e.g., Shared Technology Issue, Lack of Data Segregation) and using the association Table 2 to identify relative properties (e.g. Isolation).
3. *Identify implementation solutions* that address each security property. To this one should identify the privacy objectives related to each property taking into consideration cloud service model as well as the association between privacy properties and implementation solutions in Table 3.

The result of this process might be expressed in terms of a secure Tropos reference model (see [3]), as shown in Figure 2.

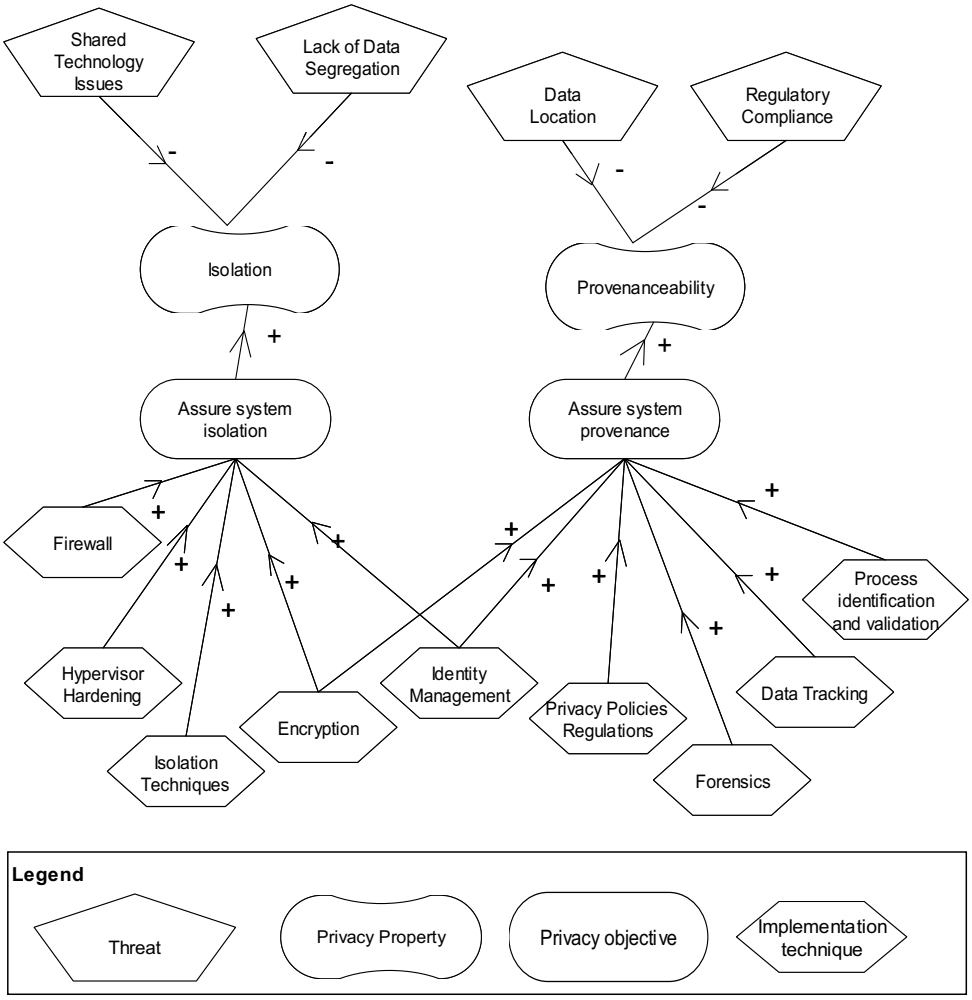

Figure 2. Security reference model for the EU eHealth example

This model is constructed during the initial stages of the system development and can be used later in the development process to identify privacy requirements that must be introduced to the system-to-be (by taking into account the privacy needs of the system) and also to identify possible means (implementation techniques) that contribute towards the satisfaction of the privacy requirements that are introduced to the system.

6. Discussion

The various innovations that cloud computing introduced in its operational environment vary from the traditional "trusted" environment where traditional information systems rely on. These innovations give rise to new privacy concerns that need to be taken into account during the design and implementation of new information systems or the migration of existing systems on cloud environments. The work presented in this paper forms an initial effort towards that direction. Furthermore, it helps to bridge the gap between design and implementation phases by suggesting a number of privacy-enhancing technologies specific to cloud environments. The identified privacy properties and associated implementation techniques can add on existing privacy requirements engineering approaches in order to better deal with the design of cloud oriented systems. This forms a future extension of our work, i.e. the transformation of the identified privacy properties into technical requirements and the formulation of appropriate design patterns specific to the cloud that will satisfy those requirements in different contexts, according to similar work on privacy patterns [9, 10, 45].

Allied to privacy is the issue of trust. Indeed recent surveys indicate privacy as one of the major factors affecting perceived trust on cloud services/providers [18]. The relation between trust and privacy is intricate. Privacy protection builds trust between cloud services and users. At the same time data privacy requirements depend on the trust in the entities collecting and processing it. Moreover, trust relationships can be at the center of certain privacy solutions involving for instance some form of key distribution. In many cases users may obtain cloud services (e.g. software applications) from a certain provider running on a cloud infrastructure that is provided by another provider, who might as well lease processing and storage capacity from other service providers. Apart from cloud providers and users, additional intermediary entities may be involved in the trust chain such as cloud brokers and cloud auditors, i.e., entities who conduct assessment

of cloud services [43, 46]. Understanding the trust relationships between the cloud entities involved is essential in order to establish that a cloud service will deliver the required privacy mechanisms needed to assure the privacy requirements of the cloud user.

The issues and challenges of trust in cloud computing have been widely discussed and a number of trust models and trust mechanisms have been proposed including the verification of Service Level Agreements (SLAs) or the measurement of Quality of Service (QoS) attributes. Each contributes a partial view of cloud trust however they do not illustrate how cloud entities work together or identify the chains of trust from cloud users to cloud services.

Huang and D.M. Nicol [46], suggest a 'societal' mechanism for identifying trust relations in the cloud. The proposed approach identifies the dependencies between cloud entities and the sources of evidence for trusting each entity. The framework aims to provide a systemic view of trust analysis; however it does not provide any tools for modeling trust dependencies or reasoning about these dependencies. In addition, whilst it acknowledges privacy as a trust parameter it does not specifically link nor how privacy concerns may affect this trust.

Another concern related to the selection of appropriate cloud solution is the selection of the service provider. A number of approaches focus on soft issues such as the reputation of a service provider, whilst others focus on hard trust factors (mainly security and performance). Most approaches aim on quantifying these factors and provide an aggregate score relating on how trusted a cloud provider is. However their evaluation is based on data on how well a cloud provider operates often provided by the cloud providers themselves and is usually hard to acquire.

Rather than relying on performance data, trustworthy selection of a cloud provider could be based on whether a cloud provider fulfils the customer's requirements in terms of the provision of appropriate mechanisms that achieve these requirements. To this end, the proposed privacy properties and implementation techniques could be used in order to elicit the specific privacy re-

quirements in a given context as well as the appropriate privacy mechanisms. Once these have been identified, the selection of an appropriate service provider can be based on the degree of satisfaction of these mechanisms by potential cloud providers.

7. Conclusion

There is increasing awareness that privacy should be integrated into the design of information systems delivered in the cloud, as dependence on the cloud provider to process and manage personal data leads to increased privacy concerns. Multi-tenancy and multi located data storage and applications in the cloud make privacy risk even more intense.

A number of recent surveys focus on the identification of privacy concerns related to different cloud service models in terms of potential attacks that exploit the vulnerabilities of cloud technologies [1, 14, 17, 19, 26, 47]. This paper extends the above by relating cloud threats and vulnerabilities to specific cloud quality properties that raise privacy concerns.

Parallel work in this area has identified technical countermeasures in order to address particular privacy issues in specific contexts. Examples include the dynamic credentials algorithm for mobile cloud computing systems described in [48] or the provenance-aware policy language for the cloud presented in [49]. In this paper the focus is in abstract solutions applicable to the cloud and their mapping to specific privacy-related properties.

Further work is required in order to identify and propose an appropriate combination of implementation solutions for the system under development due to the complexity of the cloud architecture and the combination of service models and this is the subject matter of our current work.

8. References

[1]. Hashizume, K., Rosado, D.G., Fernández-Medina, E. and Fernandez1, E. B. (2013) An analysis of security issues for cloud computing. *Journal of Internet Services and* Applications, 4, 1-13.

[2]. ITU Technology Watch (2012) Privacy in Cloud Computing. *International Telecommunications Union*, Geneva, Switzerland.

[3]. Mouratidis, H., and Giorgini, G. (2007a) Secure Tropos: A Security-Oriented Extension Of The Tropos Methodology, *International Journal of Software Engineering and Knowledge Engineering*, 17, 285-309.

[4]. Houmb, S. H., Islam, S., Knauss, E., Jürjens, J., and Schneider, K. (2010) Eliciting Security Requirements and Tracing them to Design: An Integration of Common Criteria, Heuristics, and UMLsec. *Requirements Engineering Journal*, 15, 63–93.

[5]. Sindre, G., and Opdahl, A. L. (2005) Eliciting security requirements with misuse cases, *Requirements Engineering Journal*, 10, 34–44.

[6]. Kalloniatis, C., Kavakli, E. and Gritzalis, S. (2008) Addressing privacy requirements in system design: The PriS method. *Requirements Engineering*, 13, 241-255.

[7]. Kalloniatis, C., Kavakli, E., and Gritzalis, S. (2009) Methods for Designing Privacy Aware Information Systems: A review. *Proceedings of the PCI 2009 13th Pan-Hellenic Conference on Informatics*, Corfu, Greece, 10 – 12 September, pp.185-194. IEEE CPS Conference Publishing Services, Washington, DC.

[8]. Kavakli, E., Gritzalis, S., and Kalloniatis, C., (2010) Protecting Privacy in System Design: The Electronic Voting Case, *Transforming Government: People, Process and Policy*, 1, 307-332.

[9]. Romanosky S., Acquisti A., Hong J., Cranor L. F., and Friedman B. (2006) Privacy patterns for online interactions. *Proceedings of the 2006 conference on Pattern languages of programs (PloP '06),* Portland, Oregon, 21-23 October, pp. 12:1 – 12:9. ACM New York, NY, USA.

[10]. Hafiz, M. (2013) A Pattern Language for Developing Privacy Enhancing Technologies, *Software Practice and Experience*. 43, 769-787.

[11]. Islam, S., Mouratidis, H., and Wagner, S. (2010) Toward a framework to elicit and manage security and privacy requirements from laws and regulation. *Proceeding of Requirements Engineering: Foundation for Software Quality(REFSQ)*, Essen, Germany, 30 June – 2 July, pp. 255 – 261. Springer-Verlag, Berlin, Heidelberg.

[12]. Massey, A.K., Otto, P.N., Hayward, L.J. and Antón, A. I. (2010) Evaluating existing security and privacy requirements for legal compliance, *Requirements Engineering Journal*, 15, 119-137.

[13]. Mulazzani, M., Schrittwieser, S., Leithner, M., Huber, M., and Weippl, E. (2011) Dark Clouds on the Horizon: Using Cloud Storage as Attack Vector and Online Slack Space. *Proceedings of the 20th USENIX conference on Security*, San Fransisco, CA, August 8-12, pp. 5 – 5. USENIX Association Berkeley, CA, USA.

[14]. Gong, C., Liu, J., Zhang, Q., Chen, H., and Gong, Z. (2010) The Characteristics of Cloud Computing. *Proceedings of the 2010 39th International Conference on Parallel Processing Workshop,* San Diego, CA, 13-16 September, pp. 275-279. IEEE Computer Society, Washington, DC, USA.

[15]. Pearson, S., and Benameur, A. (2010) Privacy, Security and Trust Issues Arising from Cloud Computing. *Proceedings of the 2nd IEEE International Conference on Cloud Computing Technology and Science*, Indianapolis, Indiana, USA, 30 November – 3 December, pp. 693 – 702. IEEE Computer Society, UK.

[16]. Islam, S., Mouratidis, H. and Weippl, E., (2012) A Goal-driven Risk Management Approach to Support Security and Privacy Analysis of Cloud-based System. In Rosado, D.G., Mellado, D., Fernandez-Medina E and Piattini M.,G. (eds), *Security Engineering for Cloud Computing: Approaches and Tools.* IGI Global, Hershey, PA.

[17]. CSA THREATS (2012) Top Threats to Cloud Computing Results update 2012, *Cloud Security Alliance,* Seattle, WA, USA.

[18]. Pearson, S. (2013) Privacy, Security and Trust in Cloud Computing. In Pearson, S.and Yee, G. (eds.), *Computer Communications and Networks.* Springer-Verlag, London.

[19]. Kalloniatis, C., Mouratidis, H., Manousakis, V., Islam, S., Gritzalis, S., Kavakli, E. (2014) Towards the design of secure and privacy-oriented Information Systems in the Cloud: Identifying the major concepts, *Computer, Standards and Interfaces*, 36, 759–775.

[20]. EU Draft (2012), EU Directive for Security issues in Cloud Computing. European Commission, Brussels, Belgium.

[21]. Kalloniatis, C., Kavakli, E., and S. Gritzalis (2005) PriS Methodology: Incorporating Privacy Requirements into the System Design Process. *Proceedings of the 13th IEEE International Requirements Engineering Conference – Symposium on Requirements Engineering for Information Security* (*SREIS 2005*), Paris, France, 29 August, pp. 1-9. IEEE Press, Los Alamitos, USA.

[22]. Kalloniatis, C., Kavakli, E. and Kontellis, E. (2010) PRIS tool: A case tool for privacy-oriented Requirements Engineering. *Journal of Information Systems Security*, 6, 3-19.

[23]. Kavakli, E., Kalloniatis, C. Loucopoulos, P. and Gritzalis, S. (2006) Incorporating Privacy Requirements into the System Design Process: The PriS Conceptual Framework. *Internet Research*, Special issue on Privacy and Anonymity in the Digital Era: Theory, Technologies and Practice, 16, 140-158.

[24]. Mouratidis, H., Kalloniatis, C., Islam, S., Huget, M. P. and Gritzalis, S. (2012) Aligning Security and Privacy to support the development of Secure Information Systems. *Journal of Universal Computer Science*, 18, 1608-1627.

[25]. Mouratidis, H., and Giorgini,P. (2007b) Security Attack Testing (SAT) - testing the security of information systems at design time. *Information Systems*. 32, 1166-1183.

[26]. ENISA (2012) Cloud Computing: Benefits, risks and recommendations for information security. Rev. B, V no. 2. *European Network and Information Security Agency*, Heraklion, Crete, Greece.

[27]. Imran, M. and Hlavacs, H. (2012) Provenance Framework for the Cloud Environment (IaaS). *Proceedings of the Third International Conference on Cloud Computing, GRIDs, and Virtualization*. Nice, France, 22-27 July, pp. 152-158. IARIA, NY, USA.

[28]. Jinpeng, W., Xiaolan, Z., Glenn, A., Vasanth, B. and Peng, N. (2009) Managing Security of Virtual Machine Images in a Cloud Environment. *Proceedings of the 2009 ACM workshop on Cloud computing security*, Chicago, IL, 13 November, pp. 91-96. ACM New York, NY, USA

[29]. HPL-2012-11 (2012) How To Track Your Data: The Case for Cloud Computing Provenance. *HP Labs,* Palo Alto, USA.

[30]. Wailly, A., Lacoste, M. and Debar, H. (2011) Towards Multi-Layer Autonomic Isolation of Cloud Computing and Networking Resources. *Proceedings of the 2011 Conference on Network and Information Systems Security (SAR-SSI)*, La Rochelle, France, 18- 21 May, pp. 1 – 9. IEEE CPS, Washington, DC, USA.

[31]. Viswanathan, A., and Neuman, B.C. (2012) A survey of isolation techniques, Draft Copy, University of Southern California, Information Sciences Institute, CA.

[32]. Sonehara, N., Echizen, I. and Wohlgemuth, S. (2011) Isolation in Cloud Computing and Privacy – Enhancing Technologies – Suitability of Privacy – Enchancing Technologies for Separating Data Usage in Business Processes. *Business & Information Systems Engineering Journal*, 3, 155-162.

[33]. Meixner F., and Buettner, R. (2012) Trust as an Integral Part for Success of Cloud Computing. *Proceedings of the Seventh International Conference on Internet and Web Applications and Services (ICIW 2012)*, Stuttgart, Germany, 27 May – 1 June, pp. 207 – 214, IARIA, NY, USA.

[34]. Ruan, K., Carthy, J., Kechadi, T., Crosbie, M. (2011) Cloud forensics: An overview. *Advances in Digital Forensics*, 7, 35–49.

[35]. Microsoft cloud_privacy_w_102809 (2009) Privacy in the cloud computing era, A Microsoft perspective, November 2009, *Microsoft Corp*, Redmond, USA

[36]. Simou S., Kalloniatis C., Kavakli E. and Gritzalis S. (2014) Cloud Forensics: Identifying the Major Issues and Challenges. *Proceedings of CAiSE 2014*, Thessaloniki, Greece. 16-19 June, pp 271-284. Springer-Verlag, Berlin.

[37]. Aminnezhad A., Dehghantanha, A. and Abdullah M.T. (2012) A Survey on Privacy Issues in Digital Forensics. *International Journal of Cyber-Security and Digital Forensics,* 1, 311-323.

[38]. CSA GUIDE-DOM12 (2012) Domain 12: Guidance for Identity & Access Management V2.1. *Cloud Security Alliance,* Seattle, WA, USA.

[39]. Gartner G00157782 (2008) Assessing the Security Risks of Cloud Computing. *Gartner, Inc.*, Stamford, USA.

[40]. Hansen, M., Berlich, P., Camenisch, J., Clauß, S., Pfitzmann, A., and Waidner, M. (2004) Privacy-enhancing identity management. *Information Security Technical Report*, 9, 35-44.

[41]. Meena, D.S, Radha K.P. and Ashutosh, S. (2010) A cryptography based privacy preserving solution to mine cloud a data. *Proceedings of the Third Annual ACM Bangalore Conference (COMPUTE '10)*, Bangalore, India, 22 – 23 January, pp. 14:1 – 14:4. ACM, NY, USA.

[42]. HPL-2009-54 (2009) Taking Account of Privacy when Designing Cloud Computing Services. HP Laboratories, Palo Alto, USA.

[43]. NIST-CCSRWG-092 (2011), NIST cloud computing standards roadmap, first edition. *National Institute of Standards and Technology*, Gaithersburg, MD, USA.

[44]. TR-933-EC (2010) The Cloud: Understanding the Security, Privacy and Trust Challenges. *European Commission*, Belgium, Brussels.

[45]. Pearson, S. and Shen, Y (2010) Context-Aware Privacy Design Pattern Selection. *Proceedings of TrustBus 2010*, Bilbao, Spain, 30-31 August, pp. 69-80. Springer-Verlag Berlin, Heidelberg.

[46]. Huang., J and Nicol, D.M. (2013) Trust mechanisms for cloud computing. *Journal of Cloud Computing: Advances, Systems and Applications*, 2, 1-14.

[47]. Grobauer, B., Walloschek, T., and Stocker, E. (2011) Understanding Cloud Computing Vulnerabilities, *IEEE Security & Privacy Magazine*, 9, 50-57.

[48]. Xiao, S., Gong, W. (2010) Mobility Can help: protect user identity with dynamic credential. *Proceedings of the Eleventh International conference on Mobile data Management (MDM)*. TBD, Kansas City, MO, USA, 23 -26 May, pp. 378–380. IEEE Computer Society, Washington, DC, USA.

[49]. Ali, M. and Moreau, L. (2013). A Provenance-Aware Policy Language (cProvl) and a Data Traceability Model (cProv) for the Cloud. *Proceedings of the Third International Conference on Cloud and Green Computing (CGC)*, Karlsruhe, Germany, 30 September – 2 October, pp. 479 – 486. IEEE Computer Society, Washington, DC, USA.