Accountability for Data Governance in Cloud Ecosystems

Massimo Felici, Theofrastos Koulouris, Siani Pearson
Hewlett-Packard Laboratories
Security and Cloud Lab
Bristol BS34 8QZ, United Kingdom

***Abstract*— Accountability has emerged as a critical concept related to data protection in cloud ecosystems. It is necessary to maintain chains of accountability across cloud ecosystems. This is to enhance the confidence in the trust that cloud actors have while operating in the cloud. This paper is concerned with accountability in the cloud. It presents a conceptual model, consisting of attributes, practices and mechanisms for accountability in the cloud. The proposed model allows us to explain, in terms of accountability attributes, cloud-mediated interactions between actors. This forms the basis for characterizing accountability relationships between cloud actors, and hence chains of accountability in cloud ecosystems.**

Keywords-accountability; cloud computing; data governance

I. INTRODUCTION

Cloud computing is transforming the way Information and Communication Technology (ICT) is deployed and consumed across different application domains. Cloud computing is defined as *"a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction"* [1]. Different service and deployment models enable various consumer-provider relationships depending on the type of cloud service procured (for example, some generic cloud use cases are described in [2]).

One of the most relevant shifts is concerned with data governance in the cloud. As increasingly large amounts of personal and confidential data are transferred to the cloud, stakeholders' interactions change and responsibilities are allocated across the entire cloud supply-chain. This manifests in different ways to the stakeholders involved. Data subjects are no longer co-located with their data, leading to uncertainty due to lack of transparency and loss of control. Cloud service consumers and providers, who may act as a controller and/or processor of personal and confidential data (Article 29's Opinion 1/2010 clarifies the concepts of controller and processor, and provides some examples, e.g. telecom operators, e-government portals, social networks [3]), have the responsibility of protecting such data from privacy and security breaches as well as unintended usage [4]. Governance in the cloud therefore requires understanding, moderating and regulating the relationships between cloud consumers and providers – Roles and Responsibilities: *"The partnership between providers and consumers in designing, building, deploying, and operating clouds presents new challenges in providing adequate security and privacy protection. It becomes a collaborative process between providers and consumers to share the responsibilities in implementing the necessary controls."* [5]. Fig. 1 shows a generic representation of the data governance problem in a cloud ecosystem.

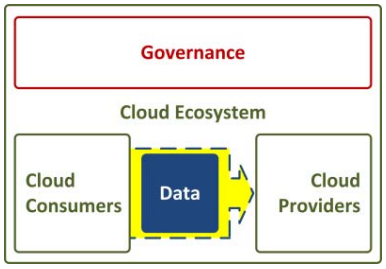

Figure 1. Data governance problem in a cloud ecosystem

Different data protection requirements arise in the consumer-provider relationship [6]. Accountability is among the identified *"technical and organizational measures of data protection and data security"* [6]. The Organisation for Economic Co-operation and Development (OECD) guidelines on privacy and data protection highlight the *"accountability principle"* – that *"a data controller should be accountable for complying with measures which give effect to the"* privacy principles (i.e. collection limitation, data quality, purpose specification, use limitation, security safeguards, openness and individual participation) [7]. Unfortunately, despite its relevance for supporting governance of privacy, data protection and security in the cloud [8], the concept of accountability is difficult to define and operationalize (that is, to put it into practice) uniformly across cloud ecosystems as *"defining what exactly accountability means in practice is complex"* [9]. However, various elements have been identified as characterizing a general accountability principle.

The Article 29 Data Protection Working Party in [9] identifies various elements defining the general principle of accountability (e.g. reinforcing existing obligations, appropriate measures to implement the provisions of the Data Protection Directive, role of data protection authorities, sanctions, development and regulation of certification schemes). Similarly, the Galway Project has identified five essential elements of an accountable organization [10].

978-0-7695-5095-4/13 $31.00 © 2013 IEEE
DOI 10.1109/CloudCom.2013.157

This paper addresses the need to establish a shared understanding of accountability, tailored to the cloud. It introduces a conceptual definition of accountability based on a critical analysis and review of related work conducted within the *Cloud Accountability Project* (which focuses on the *Accountability For Cloud and Other Future Internet Services* as the most critical prerequisite for effective governance and control of corporate and private data processed by cloud-based services). It tailors the conceptual definition of accountability to the data governance problem in cloud ecosystems. These definitions underlie a model addressing the complexity of accountability by structuring and identifying alternative perspectives (i.e. accountability attributes, practices and mechanisms). This accountability model enables us to critically analyze accountability in the cloud. In particular, it allows us to explain accountability in terms of accountability relationships between actors (e.g. consumers and providers) in cloud ecosystems. This paper is structured as follows. Section II introduces the definitions of accountability, i.e. conceptual definition of accountability and definition of accountability for data stewardship in the cloud. Section III describes the accountability model based on such definitions. Section IV analyzes sample cloud use cases from an accountability viewpoint. Section V highlights the complexity of chains of accountability, hence the need to support accountability governance. Section VI highlights some concluding remarks.

II. ACCOUNTABILITY IN THE CLOUD

The following definition captures a shared understanding of accountability based on extensive review of previous multi-disciplinary related work and discussion within the *Cloud Accountability Project*:

> **Conceptual Definition of Accountability:** *Accountability consists of defining governance to comply in a responsible manner with internal and external criteria, ensuring implementation of appropriate actions, explaining and justifying those actions and remedying any failure to act properly.*

Governance here comprises the processes which devise ways of achieving accountability. This conceptual definition of accountability encompasses different understandings drawn from different disciplines. It is intentionally phrased to be generally applicable across different domains. Further to this generic definition, we tailor the conceptual definition of accountability to the domain of focus of the *Cloud Accountability Project*, i.e. data stewardship in the cloud. Accountability therefore is understood and analyzed in the context of protecting data processed in service provision ecosystems. Our focus is accountability towards cloud users. Hence, the focus is on accountability of the custodians of user data in the cloud. We will consider accountability of organizations using the cloud, cloud service providers (who may or may not also be cloud users) and accountability agents that may assess organizations' trustworthiness, towards stakeholders including data subjects and regulators (assessment of accountability of individuals in a private context is excluded, although accountability of employees of service providers is within scope). A major focus is on personal data in the cloud (therefore, accountability will be understood in the data protection law context). However, we will not confine ourselves exclusively to scenarios focused on personal data, but also for specific cases we will consider how our proposed mechanisms would apply to confidential data such as business secrets (government surveillance, including government acquisition of data from cloud service providers, is outside the scope of the *Cloud Accountability Project*, except where it relates specifically to a data protection law accountability mechanism). The following definition contextualizes the conceptual definition of accountability and makes it relevant to the data governance problem in cloud ecosystems:

> **Definition of Accountability for Data Stewardship in the Cloud:** *Accountability for an organization consists of accepting responsibility for the stewardship of personal and/or confidential data with which it is entrusted in a cloud environment, for processing, storing, sharing, deleting and otherwise using the data according to contractual and legal requirements from the time it is collected until when the data are destroyed (including onward transfer to and from third parties). It involves committing to legal and ethical obligations, policies, procedures and mechanisms, explaining and demonstrating ethical implementation to internal and external stakeholders and remedying any failure to act properly.*

These definitions highlight the main conceptual aspects of accountability. They characterize the necessary practices emerging in organizations that adopt an accountability-based approach for data stewardship. The next section defines a model characterizing accountability in cloud ecosystems.

III. ACCOUNTABILITY MODEL

Building on the definitions of accountability introduced in the previous section, we introduce a model of accountability for data stewardship in the cloud. The model expands upon the definitions using accountability practices, attributes and mechanisms. Accountability, a complex concept related to privacy and data protection, encompasses different attributes, hence accountability attributes. Accountability practices characterize organizational behavior, hence what defines accountable organizations. Diverse mechanisms are used in order to support such practices. The accountability model consists of:

- **Accountability attributes** – conceptual elements of accountability as used across different domains (i.e. the conceptual basis for our definition, and related taxonomic analysis)
- **Accountability practices** – emergent behavior characterizing accountable organizations (that is, how organizations operationalize accountability or put accountability into practices)
- **Accountability mechanisms** – diverse processes, non-technical mechanisms and tools that support accountability practices (that is, accountability practices use them).

Fig. 2 illustrates how attributes, practices and mechanisms form a model of accountability for cloud ecosystems. Accountability is interpreted in terms of attributes – *accountability attributes*. These accountability attributes are operationalized (that is, put into practices) by organizational practices – *accountability practices*. In order to implement such practices, organizations use different mechanisms tailored to their domains – *accountability mechanisms*. On the other hand, these mechanisms constrain and support accountability practices, and the operational interpretation of the accountability attributes.

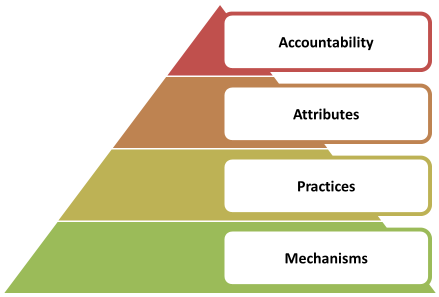

Figure 2. Accountability Attributes, Practices and Mechanisms

The emerging relationships between accountability attributes, practices and mechanisms give rise to an operational interpretation of accountability. This characterization explains how organizations may attain accountability in different ways, that is, instantiate this accountability model differently according to their particular contexts. The remainder of this section describes the accountability attributes, practices and mechanisms.

A. Accountability Attributes

Accountability attributes capture concepts that are strongly related to and support the principle of accountability. These include: key properties of accountability (e.g. transparency); conceptual elements (e.g. remediation); consequences (e.g. sanctions); related objects (e.g., obligations). There exist emerging relationships (e.g. implication and inclusion) among attributes dependent on different viewpoints of analysis (which are related to societal, legal and ethical aspects of accountability). For instance, from a legal perspective, responsibilities imply obligations, which consequently may involve sanctions. From a social perspective, transparency implies both observability and verifiability (and vice versa, transparency is obtained by combining observability and verifiability). This section defines and focuses on accountability attributes: *observability*, *verifiability*, *attributability*, *transparency*, *responsibility*, *liability* and *remediability*.

- **Observability** is a property of an object, process or system which describes how well the internal actions of the system can be described by observing the external outputs of the system.
- **Verifiability** is a property of an object, process or system that its behavior can be verified against a requirement or set of requirements.
- **Attributability** is a property of an observation that discloses or can be assigned to actions of a particular actor (or system element).
- **Transparency** is the property of an accountable system that it is capable of 'giving account' of, or providing visibility of, how it conforms to its governing rules and commitments.
- **Responsibility** is defined as the state of being assigned to take action to ensure conformity to a particular set of policies or rules.
- **Liability** is the state of being liable (legally responsible).
- **Remediability** is the state of being able to be remedied.

A critical analysis of such attributes and their definitions (tailored to accountability based on relevant literature and discussion within the *Cloud Accountability Project*) has been presented in [11]. Structuring accountability (in terms of attributes) allows us to interpret relationships between actors in the cloud. Fig. 3 illustrates the scope and inter-relationships among the defined accountability attributes in the context of a cloud-mediated interaction between two generic actors (Actor A and Actor B). The actors are intentionally kept generic to allow for generalizations where one of the actors is actually an oversight or enforcement entity (e.g. regulator and auditor).

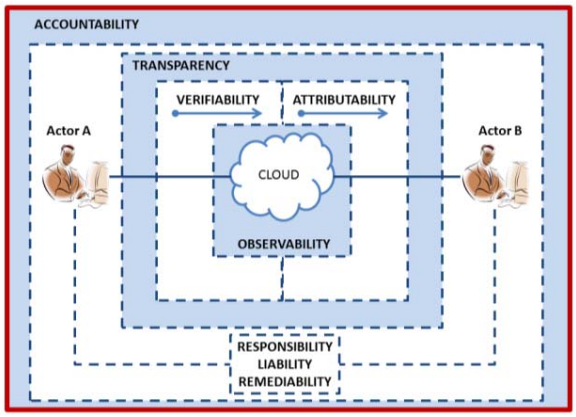

Figure 3. Accountability relationships between actors

Fig. 3 highlights how the attributes fit together to enable (evidence of) accountability. Transparency relies on verifiability and attributability, which in turn rely on observability. Responsibility, liability and remediability rely on transparency. Other aspects of accountability are also relevant. For instance, sanctions are (legal) consequences of failing to fulfill responsibilities. Assurance (resulting from responsibility and liability) is a positive declaration intending to give confidence. Assurance can take the form of evidence, which can be used to convince a third party about, for example, the reason for a failure that has happened. Remediation is the act or process of correcting, for example, a failure or deficiency. Some of these concepts (e.g. obligations, sanctions and holding to account) are further discussed within accountability practices.

B. *Accountability Practices*

Accountability practices, derived directly from the given definitions, characterize emerging behavior (highlighting operational and organizational objectives to be met) manifested in accountable organizations. Specifically, an accountable organization:

- Defines governance to responsibly comply with internal and external criteria, particularly relating to treatment of personal data and/or confidential data
- Ensures implementation of appropriate actions
- Explains and justifies those actions, namely, demonstrates regulatory compliance that stakeholders' expectations have been met and that organizational policies have been followed
- Remedies any failure to act properly, for example, notifies the affected data subjects or organizations, and/or provides redress to affected data subjects or organizations, even in global situations where multiple cloud service providers are involved.

Accountable organizations need to define and implement appropriate governance mechanisms relating to treatment of personal and/or confidential data in cloud environments. The actions in question pertain to the collection, storage, processing and dissemination of personal and/or confidential data. Fig. 4 shows the interaction between two organizations (as a continuous process) driven by *accountability governance* (constrained by external criteria and regulatory regimes but orchestrated independently by organizations). Organization A could be part of a service provision chain that involves cloud service providers and Organization B is actually an oversight and enforcement actor (e.g. regulators and accountability agents) in the chain. Organization A defines and implements appropriate governance mechanisms, which enable to demonstrate governance. Organization B, holding to account Organization A, can ask for further clarification, engage in discussions and also apply sanctions. As a result, Organization A may modify organizational governance.

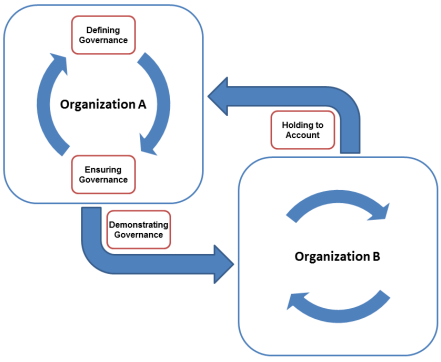

Figure 4. Accountability Governance

Organizations need to provide transparency of those actions taken in order to show that stakeholders' expectations have been met and that organizational policies have been followed. They also need to remedy any failure to act properly (e.g. by notifications, remedies, sanctions) even in cloud-supply chains involving multiple service providers. Accountability governance redefines interactions between providers and regulators as well as between providers themselves. The ethical nature of an accountability-based approach and the organizational obligations that result from taking this approach represent a shift from reactive to proactive governance of personal and/or confidential data. Organizations commit to the stewardship of personal and/or confidential data by addressing legal, contractual and ethical obligations. In order to do so, organizations deploy and use different mechanisms (e.g. policies, standards), take into account social norms, provide evidence to internal and external stakeholders, and remedy any failure to act properly.

C. *Accountability Mechanisms*

The accountability model highlights 'what' needs to be implemented. Within the model, accountability mechanisms (cf. the 'how') are instances of tools and techniques supporting accountability practices (that is, high level objectives that accountable organizations need to achieve). Organizations can adopt different available accountability mechanisms as appropriate for their contexts. They will use what is best for their particular processes (but of course, they also need to demonstrate that they have used appropriate mechanisms). Accountability mechanisms focus on the core aspects of accountability (e.g. remediation, notification and risk assessment). In addition, privacy mechanisms need to be used to reduce privacy risk as necessary [12].

Mechanisms (e.g. security controls, policies, tools, standards, legal mechanisms, penalties), from a social science viewpoint, are accountability objects *("that both inhabit several communities of practice and satisfy the information requirements of each of them"* [13]). Accountability mechanisms (developed by the *Cloud Accountability Project*) will complement others that are available from third parties. They may be used individually or in combination. Organizations may select from different alternatives. For example, they may choose to use the Privacy Level Agreement format specified by the Cloud Security Alliance (CSA) to express privacy-related obligations [14], or the Cloud Trust protocol [15] to ask for and receive information from cloud service providers about the elements of transparency, or they may take another approach to do so.

IV. CLOUD USE CASES

The NIST recommendations use the different cloud actors in order to discuss cloud usage scenarios [5]. Depending on the deployment models (e.g. private or public), cloud providers and consumers interact differently. Their security boundaries would define control and visibility over deployed resources. They might also be exposed differently to emerging threats in the cloud [16]. The Cloud Computing Use Case Discussion Group characterized some generic use cases [2]. This section uses the definition of accountability for data stewardship in the cloud, the accountability model and the characterization of relationships between cloud actors to discuss accountability perspectives of some cloud use cases drawn from [2].

A. End User to Cloud

Fig. 5 shows the interaction between an end user and applications running in a public cloud. End users access such applications and usually have little idea how the underlying architecture works. This use case would apply for most generic SaaS applications (e.g. email services and social networks). It has implications in terms of transparency, verifiability and observability requirements. Information about how cloud services manage user data would be needed in order to guarantee an extent of obervability over the behavior of any cloud service. Evidence based on such observability would enable assessing (verifiability) whether or not cloud services fulfill user-defined policies. Information on how cloud services operate on user data as well as comply with user-defined policies would be made available to end users with a degree of transparency. This identifies additional requirements for transparency, verifiability and observability to the ones (i.e. Identity, Open technology platform, Security and Clearness of Service Level Agreements) identified in [2].

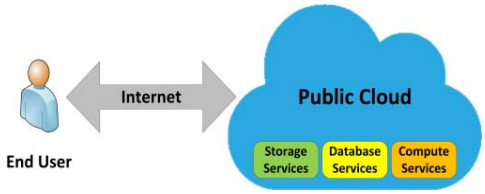

Figure 5. End User to Cloud

B. Enterprise to Cloud to End User

Fig. 6 shows the case of applications running in the public cloud and that are accessed by individuals, who might be either employees or external customers of the enterprise (e.g. governmental cloud services). From a viewpoint of accountability analysis, alongside the other already identified requirements, end users would benefit from attributability between cloud services (in this case acting as data processor) and the enterprise (in this case acting as data controller) using or offering such services. Supporting attributability in combination with the other accountability attributes (in particular, observability, verifiability and transparency) would enable end user to assess responsibilities clearly (who did what) between the enterprise and the cloud services in case of data breaches or other potential threats [16].

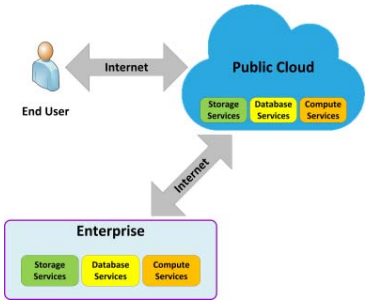

Figure 6. Enterprise to Cloud to End User

C. Enterprise to Cloud

Fig. 7 shows the case of an enterprise that uses (or switches to) cloud services for its internal processes or ICTs (for example, using cloud storage for backups, or storage of seldom-used data, or SaaS for email, calendar applications). NIST has developed several use cases (of organizations moving to the cloud) in different application domains [17]. These use cases inform an analysis that defines and prioritizes requirements for interoperability, portability, and security in order to support secure and effective adoption of cloud computing. An accountability analysis of such use cases elicit further requirements concerned with data governance in the cloud. For instance, an accountability analysis would elicit requirements concerned with the accountability attributes dealing with emerging hazards and vulnerabilities in the cloud [18]. The problem is to guarantee policies throughout chains of accountability while services are accessed by multiple cloud users.

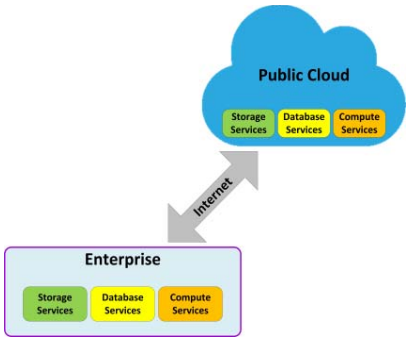

Figure 7. Enterprise to Cloud

D. Enterprise to Cloud to Enterprise

Fig. 8 shows two enterprises that use the same cloud (within a supply chain) in an interoperable manner. Within cloud service deployments there may be horizontal chains of SaaS providers, as well as vertical chains down to PaaS and IaaS providers. Some services might span multiple roles. Within service provision chains, CSPs may have multiple roles (cloud service consumers as well as providers). The accountability analysis of the different actors involved in any particular cloud supply-chain would identify specific activities as well as responsibilities and any associated liabilities and remedies.

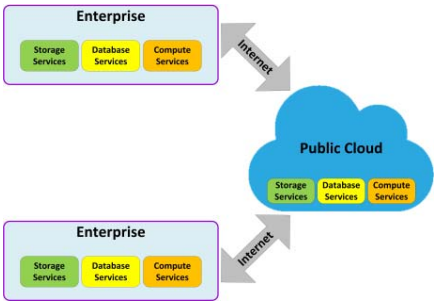

Figure 8. Enterprise to Cloud to Enterprise

V. CHAINS OF ACCOUNTABILITY

Cloud services are defined in terms of different essential characteristics, service models and deployment models [1]. Combining such features enables different business models and cloud ecosystems involving various stakeholders. Organizations that use or provide cloud services operate in a complex dynamic environment, use cloud services within a supply-chain, and need to feel confident that providers further down that chain are accountable for how they manage personal and/or confidential data. Fig. 9 shows a cloud ecosystem involving different actors who contribute to data governance in the cloud. A chain of accountability should exist that extends all the way back to the cloud users.

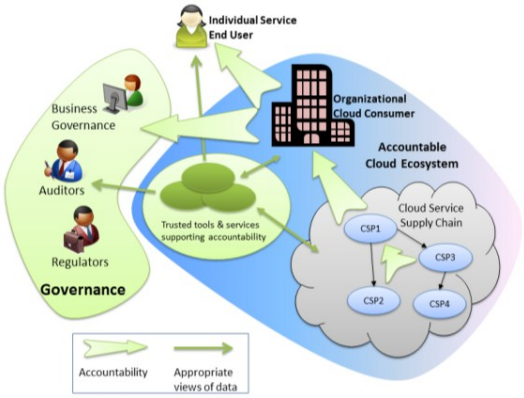

Figure 9. Sample Actors in a Cloud Ecosystem

Accountability is central to a trustworthy cloud. Without accountability, cloud consumers will lack confidence to put personal and/or confidential data in the cloud. Switching to the cloud model involves changes in control, in trust and security boundaries and, may be also, in legal regulatory requirements. In order to enhance trustworthiness of cloud ecosystems, it is necessary to have a thorough understanding of the potential benefits and risks of adopting the cloud [18]. Moving data to the cloud involves a shift in responsibilities across organizational boundaries. This redistribution of responsibilities across the cloud ecosystem changes risk fundamentals (e.g. likelihood of occurrence and severity) as well as risk perceptions of such threats. It becomes necessary to understand vulnerabilities as well as to identify new mechanisms enhancing trustworthiness of cloud ecosystems. Accountable organizations ensure that obligations to protect data are observed throughout service supply-chains. Accountability enhances the confidence of service providers, regulators and end users to deploy, use, and monitor cloud services. It enables cloud ecosystems to position themselves with respect to regulatory regimes. Accountability provides a means for achieving compliance with respect to regulatory regimes, enabling (transparency, security and privacy) mechanisms tailored to protect data and data subjects.

VI. CONCLUDING REMARKS

In conclusion, this paper has discussed accountability as a means to support and achieve data governance in cloud ecosystems. The main contributions are:

- Accountability definitions tailored for characterizing data stewardship in cloud ecosystems
- Accountability model consisting of accountability attributes, practices and mechanisms
- Cloud use cases analyzed for the identification of accountability requirements
- Accountability Governance enabling actors in cloud ecosystem to hold to account for data stewardship and to comply with regulatory regiments.

These contributions provide new insights on accountability for data governance in cloud ecosystems.

ACKNOWLEDGMENT

This work has been partly funded from the European Commission's Seventh Framework Programme (FP7/2007-2013) under grant agreement no: 317550 (A4CLOUD) *Cloud Accountability Project* – http://www.a4cloud.eu/.

REFERENCES

[1] P. Mell, T. Grance, "The NIST Definition of Cloud Computing", NIST Special Publication 800-145, 2011.

[2] Cloud Computing Use Case Discussion Group, "Cloud Computing Use Cases White Paper", Version 4.0, 2010.

[3] Article 29 Data Protection Working Party, "Opinion 1/2010 on the concepts of "controller" and "processor", 00264/10/EN, 2010.

[4] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Official Journal L 281, 23/11/1995, p. 0031-0050, 1995.

[5] L. Badger et al., "Cloud Computing Synopsis and Recommendations", NIST Special Publication 800-146, 2012.

[6] Article 29 Data Protection Working Party, "Opinion 05/2012 on Cloud Computing", 01037/12/EN WP 196, 2012.

[7] Organisation for Economic Co-operation and Development (OECD), "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data", 1980.

[8] D. Guagnin et al. (eds), "Managing Privacy through Accountability", Plagrave Macmillan, 2012,

[9] Article 29 Data Protection Working Party, "Opinion 3/2010 on the principle of accountability", 00062/10/EN WP 173, 2010.

[10] Galway Project, "Accountability: A Compendium for Stakeholders", The Centre for Information Policy Leadership LLP, 2011.

[11] S. Pearson et al., "Towards a Model of Accountability for Cloud Computing Services", In International Workshop on Trustworthiness, Accountability and Forensics in the Cloud (TAFC), 2013.

[12] S. Pearson, "On the Relationship between the Different Methods to Address Privacy Issues in the Cloud", In: OTM 2013 Conference, R. Meersman, H. Panetto et al (eds.), Springer, LNCS, 2013.

[13] G.C. Bowker, S.L. Star, "Sorting things out: classification and its consequences", The MIT Press, Cambridge, 1999.

[14] CSA, "Privacy Level Agreement Outline for the Sale of Cloud Services in the European Union", Cloud Security Alliance, Privacy Level Agreement Working Group, 2013.

[15] R. Knode, D. Egan, "Digital Trust in the Cloud: A Precis for the CloudTrust Protocol (V2.0), Computer Science Corporation, 2010.

[16] CSA, "The Notorious Nine: Cloud Computing Top Threats in 2013", Cloud Security Alliance, Top Threats Working Group, 2013.

[17] L. Badger et al., "US Government Cloud Computing Technology Roadmap, Volume II Release 1.0 (Draft), Useful Information for Cloud Adopters", NIST Special Publication 500-293 (Draft), 2011.

[18] D Catteddu, G. Hogben (Eds.), "Cloud Computing: Benefits, risks and recommendations for information security", ENISA, 2009.