Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 94 (2016) 160 – 167

Procedia Computer Science

The 13th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2016)

A Conceptual Framework for Designing Data Governance for Cloud Computing

Majid Al-Ruithe*, Elhadj Benkhelifa†, Khawar Hameed

Cloud Computing and Applications Research Lab, Staffordshire University, Stafford, ST18 0AD, UK

Abstract

Data complexity and volume continue to explode; businesses have grown more sophisticated in their use of data which drives new demands that require different ways to combine, manipulate, store, and present information. Forward thinking companies have recognised that data management solutions on their own are becoming very expensive and not able to cope with business reality, and that they need to solve the data problem in a different way through the implementation of effective data governance. Attempts in governing data failed before, as they were driven by IT, and affected by rigid processes and fragmented activities carried out on system-by-system basis. Up to very recently governance is mostly informal, in siloes around specific enterprise repositories, lacking in structure and the in wider support by the organisation. With the emergence of cloud computing and the increased adoption, data governance is receiving an increasing interest amongst specialist, but still under researched. This paper presents initial research towards developing an effective data governance programmes for the cloud paradigm. The paper discusses why it is essential to do so from both the cloud consumer and provider perspectives and proposes a conceptual framework and a five-step procedure for designing data governance for cloud computing.

Peer-review under responsibility of the Conference Program Chairs

Keywords: Data Governance; Cloud Computing; Framework; Data Management; Cloud Data Governance

* Corresponding author. Tel.: +447479471119.
E-mail address: mrowathi@gmail.com
* Corresponding author. Tel.: +447916706720.
E-mail address: e.benkhelifa@staffs.ac.uk

Peer-review under responsibility of the Conference Program Chairs
doi:10.1016/j.procs.2016.08.025

1. Introduction

A recent development in technology is the emergence of Cloud Computing. The National Institute of Standards and Technology (NIST) defined Cloud Computing 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]. Cloud Computing model enhances availability and is composed of five essential characteristics, four deployment models and three service models[2]. The essential characteristics of Cloud Computing include on-demand self-service, broad network access, resource pooling, rapid elasticity and measured service[3]. The Cloud deployment models are private, public, hybrid, and community model [7,8]. In addition, Cloud Computing includes three service delivery models which are: Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS)[7]. Cloud Computing offers potential benefits to public and private organisations [4,5] by making information technology (IT) services available as a commodity [9]. The general claimed benefits of Cloud Computing include [6]: cost efficiency, almost unlimited storage, backup and recovery, automatic software integration, easy access to information, quick deployment, easier scale of services and delivery of new services [10]. In addition, cloud computing enhances operational capabilities through increased storage and automation[5]. Furthermore, other benefits include: optimised server utilization, dynamic scalability and minimised life cycle development of new applications[7]. Though, Cloud Computing is still not quite widely adopted because of many factors, but mostly concerned with moving business data to be handled by third party[8], including loss of control on data, security and privacy of data, data quality and assurance, and data stewardship etc. Data lock-In is a potential risk whereupon cloud customers who can face difficulties in extracting their data from the Cloud[9]. Cloud consumers can also suffer from operational and regulatory challenges, as organisations transfer their data to third parties for storage and processing[5]. It may be difficult for the consumers to check the data handling practices of the cloud provider.

Cloud computing models are expected to be a highly disruptive technology and the adoption of its services will require an even more rigorous data governance strategies and programmes which can be more complex but necessary. There is very little research reported in literature on data governance for cloud computing and what is reported is still very superficial. This paper is presents an important contribution in this, so far neglected field. The next section reemphasises on the importance of implementing effective data governance for cloud computing. Section 3 proposes a conceptual framework for data governance design for cloud computing and Section 4 provides a step-by-step procedure in realising this design. Section 5 presents the conclusion and future work.

2. Importance of implementing Effective Data Governance for Cloud Computing

The most significant issues that are facing cloud consumers when adoption cloud computing is loss of control on their data since their data is stored on a computer belonging to the cloud provider[10,3]. Arguably, this loss of governance and control could have a potentially severe impact on the organisation's strategy, and therefore on the capacity to meet its mission and goals. The loss of control and governance can also lead to the impossibility of complying with the security requirements, a lack of confidentiality, integrity and availability of data, and a deterioration of performance and quality of service, not to mention the introduction of compliance challenges. Therefore, organisations need to be aware of the best practices for safeguarding, governing, and operating data in the cloud environment.

[6] NIST offers many recommendations to the cloud consumers, one being that organisations have to consider data governance strategy before they adopt cloud computing. This recommendation cements the argument that data governance is important for organizations who intend to move their data and services to cloud computing environment because it will set policies and rules, and distribution of responsibilities between cloud actors. Developed policies and data governance processes will help organisations to monitor compliance with approved standards, and technical and business guidance in cloud environments. Effective data governance may solve some of the challenges of cloud computing especially those issues related to data-ultimately offering many benefits to cloud consumers.

Ensuring security, privacy and quality of data is considered as one of the main benefits of data governance [11]. Another benefit of data governance concerns the improvement of the ability to manage and monitor data in the cloud environments[5]. Data governance also helps cloud consumers to classify their data based on level of sensitivity before moving to cloud computing environments. In addition, data recovery is enhanced by ensuring that the examination of data archiving, recovery and backup happens via data governance[12]. Moreover, reducing compliance risk and errors in the cloud environment is another advantage of data governance[6]. Collectively, all the benefits of data governance enables cloud consumers to enforce control on their data in the cloud environment in addition to ensuring that the cloud computing strategy aligns with the organisation's strategy.

However, implementing data governance for cloud computing will change according to the roles and responsibilities in the internal process of an organisation[13]. Thus, it will face many issues. Common barriers to implementing and sustaining data governance for cloud services which have been classified into four categories. Foremost, lack of data governance understanding from cloud consumer is one of these barriers, which includes not being part of organisation's culture, lack of training and lack of communication plan[14]. In addition, lack of support is the another barrier, which includes lack of top management support, lack of compliance enforcement and lack of cloud regulation[13]. Furthermore, lack of policies, process and defined roles in organization are one of the main barriers to implement data governance[15]. Finally, lack of resources is considered as another data governance barrier, this includes lack of funding, technology, people and employees' skills and experience[13].

3. A Conceptual Framework for Design Data Governance for Cloud Computing Services

A number of researchers have recognised a need for a data governance design framework for cloud computing[3,16]. The framework presented in this paper aims to supposed to support cloud consumers who need to design data governance for cloud computing in their organisations by covering general procedures in designing data governance for the cloud computing services. Designing data governance for cloud computing is potentially complex. In order to address this complexity, the proposed framework is based on the premises of *Analytic Theory*. According to Otto Boris (2011) analytic theory is useful for structuring the research topic of data governance. In this paper we use analytic theory for deducting and understanding the important processes necessary to construct an effective data governance design framework.

In the literature there are few emerging data governance frameworks designed by industry associations such as DAMA, DGI and IBM[17,18]. The Data Governance Institute Framework consists of three components: people and organizational bodies, rules and rules of engagement, and processes[17]. IBM's roadmap model for effective data governance consist of a fourteen steps phase, which includes ten steps which are required and four optional steps[18]. However, the aforementioned frameworks focus towards data governance for traditional IT, and they do not provide a comprehensive framework within which organisations can implement data governance for the Cloud services. Since the cloud computing differs from traditional IT, the design of data governance framework for the cloud will need to consider other aspects mostly related to features of the cloud computing itself. After careful analysis of existing literature, the proposed framework in this paper focuses on five key processes for data governance in cloud computing. These are:

1. Data Governance Structure.
2. Data Governance Assessment.
3. Data Governance Function.
4. Negotiation.
5. Data governance Level Agreement.

The conceptual framework is presented for designing data governance for cloud computing and considers five processes depicted in Figure 1 below.

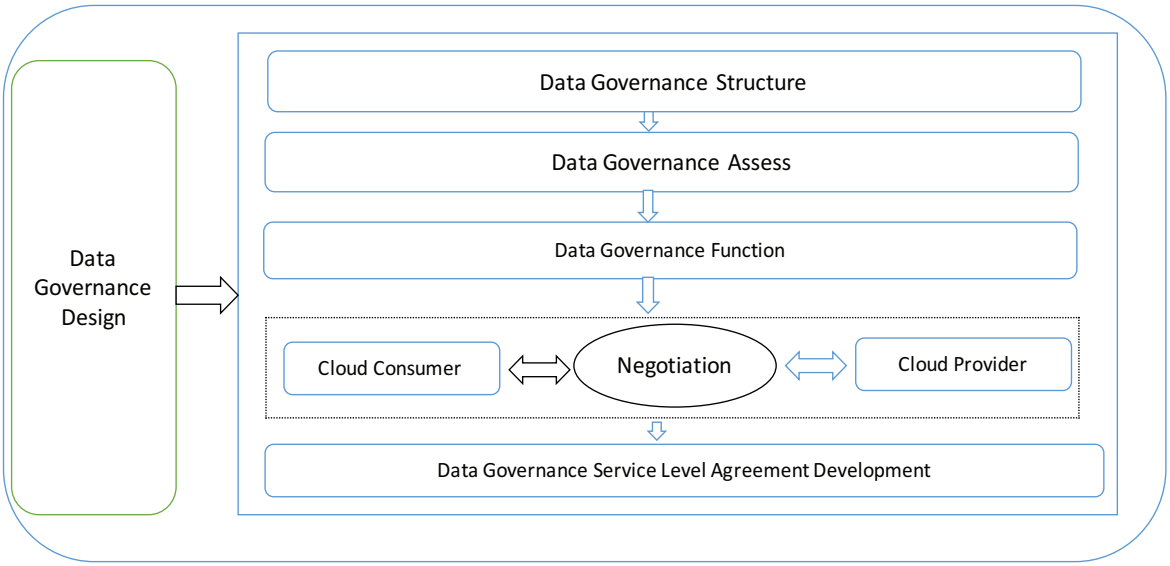

Fig. 1. Conceptual Framework for Design Data Governance for Cloud Computing Services.

3.1 Director Responsibilities for Cloud Data governance.

Recent studies on governance have raised the level of interest in directors' responsibilities on IT governance and the emerging technology of cloud computing[19]. Although cloud computing is recognized as an important issue for organisations, it also raises risks for organisations, and more specifically, in their data. Thus, to implement governance policies for data in cloud computing services has become an urgent necessity. Despite the recognition for the importance of data governance for cloud computing services, a survey by Judith R. Davis revealed that more than 43% of organisations lack knowledge about data governance[20]. This lack of understanding of data governance raises questions of governance and may put organisations in danger of losing their data when it moves to cloud computing environments.

Data governance has to align with the goals of the organisation as a whole, thus it will present the opportunity and add value to the organisation. It also contributes to the increased effective of cloud computing in organisation[15], and it contributes to increase adoption of cloud computing in organisations. Therefore, data governance for cloud computing needs to identify the necessary roles and responsibilities for supervising, implementing and monitoring an effective data governance program in organisation. The responsibilities of Board of Directors in an organisation are to supervise the operation and the management of business[5]. Thus, they have to evaluate the existing governance against data risks associated with cloud computing services and the impact on organisation strategy. As a result, organisations must be well prepared in implementing necessary policies to get control on their data over cloud computing. This needs more effort from organisation team through collaboration between business ,IT and legal members in the organisation.

However, data governance for traditional IT is different to data governance in cloud computing. In traditional IT, the audit committee is responsible for internal control to monitor data issues in the organisation[5]. In contrast, the cloud infrastructure is set up and maintained by a third party, thus the organisations need new roles and responsibilities to monitor and control their data in the cloud provider environment, thus setup roles and responsibilities for Cloud data governance is potentially complex process. Data governance issues for concern include risk management, disaster recovery plan, security, privacy, integrity, incident response, access management, and accountability[21]. Therefore, directors need to ensure that management is taking the steps necessary to ensure that effective data governance is in

place. As such, formalising roles and responsibilities creates transparency on data in the cloud environment. To setup an effective data governance roles and responsibilities for cloud actors, the data governance team need more efforts to setup them in direct way to achieve cloud data governance objectives.

4. Procedure for Designing Cloud Data Governance

In order to establish a framework for designing an effective cloud data governance, the study proposes a five-step procedure as illustrated in Fig 2.:

Step1: Set up data governance structure to enforce and identify roles and responsibilities between data governance teams. This will help cloud consumers to ensure that requisite roles and responsibilities for data governance are addressed throughout the enterprise at the right organisational level. There are many structures for data governance and not all will be a good fit for every organisation. A common model for data governance structure that takes a three-tiered approach includes a group of senior-level executives, a middle management group, the data governance office, and data governance working group. In the Cloud, the infrastructure is multi-site, and the management responsibilities are handled from cloud consumers and provider based on cloud services models. Thus, new members involved in data governance structures are the cloud manager, cloud provider, and cloud broker.

Step2: Evaluate and assess existing data governance in organisation. In data governance contexts, assessment refers to the the ability of the organisation to govern and to be governed[11] . It is useful to determine the current state of the data governance, mechanisms and capability of an organisation in order for it to change some its processes when implementing data governance for cloud computing services. Evaluation of risks associated with cloud computing will be considered in this step. In this step, the data governance committee should set up a data governance maturity model to assess the risks and opportunities data governance for cloud computing presents to organisation. Thus, the data governance maturity model will help cloud consumers to identify their targets for data governance before developing service level agreement with cloud providers. All of the assessment procedures and steps have to be documented. Through this step, the data governance requirements will be clear for cloud consumers.

Step3: Set up data governance functions for cloud computing services. Data governance functions refers to master activities for data governance which the data governance committees have to take in account when implementing data governance for the cloud[3]. The data governance functions consist of many activities that are: policies, principles, process, decision right, roles and responsibilities, communication and change management plan. The set up of data governance policies and standards (good standards and practices) will help cloud consumers to get control of their data in the cloud. Cloud business objectives and risk will also be considered in this step. Therefore, it is important that cloud consumers establish their data governance standards and policies at the very first stage before choosing a cloud provider. The whole procedure is a continuing process and is performed in the alignment with other strategy offers in the organisation. In addition, it is important to integrate with the cloud computing context, and its standards and policies. As results, the strong policies will lead to an effective data governance for the cloud services.

Step 4: Set up a negotiating contract for cloud data governance. It is important for cloud consumers to evaluate cloud providers, and to inform cloud providers of their requirements for the cloud in general and, more specifically, for data governance before moving their data to cloud provider environment [22]. Negotiation has been defined as "a process where two parties with differences which they need to resolve are trying to reach agreement through exploring for options and exchanging offers and an agreement"[23]. Cloud consumers have to know their requirements to govern data before negotiating with cloud provider because the success of negotiation strategies depends crucially on the planning preparations of the negotiators. Consumers should understand all the factors that may influence negotiation before starting negotiation. For example, complex infrastructure negotiations, context of the negotiation and negotiation culture. It is important that legal teams on both sides (consumer/provider) are fully involved in the negotiation [24]. This will result in more productive dialogue and negotiation between cloud consumer and provider when agreeing level of data governance and associated contracts.

Step 5: Develop data governance level agreement. Negotiations between cloud consumer and cloud provider should yield an appropriate service level agreement (SLA) encompass all data governance requirements. As the SLA should include a set of guidelines and policies to assist cloud consumers in defining governance plans for their data, and the cloud providers have to consider that when managing data in the cloud [25]. Therefore, the SLA for cloud data governance should includes data governance policy, data governance process, data governance principal, data governance procedure, roles and responsibilities, data governance metrics, data governance tool and techniques, and data governance monitor. All of these requirements have to comply with legal and regulatory requirements. In addition, the cloud consumers should take into account the realities of today's cloud landscape, and postulate how this space is likely to evolve in the future, including the important role that standards will play to improve interoperability and compatibility across providers. As results, the SLA are positively related to the trust and relationship commitment between the cloud consumer and provider. These agreements should protect both parties.

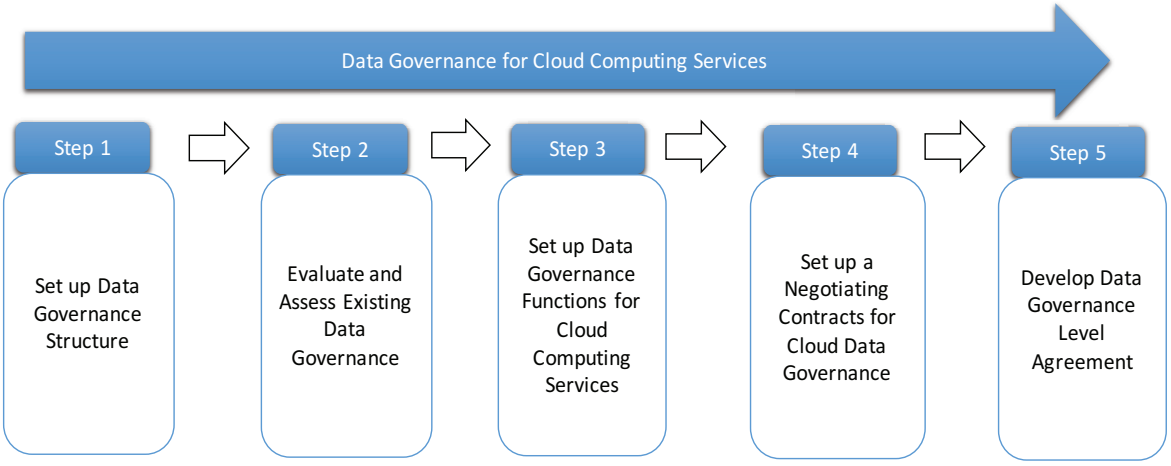

Fig. 2. A Five-Step Procedure for Designing Data Governance for Cloud Computing.

5. Conclusion

In the absence of enough literature on data governance in general and more particularly for the cloud paradigm, this paper presents a useful contribution to the relevant research communities. This paper further illustrates the importance of data governance for any organisation's success at the same time the lack of research and industrial initiative to drive this area forward. The paper presents an initial attempt to illustrate a conceptual framework for designing data governance for cloud computing. A step by step procedure is described to realize this framework. Future work will involve developing a more detailed holistic framework for cloud data governance strategy, including the main pillars, processes and attributes to design more specific data governance programs. The proposed framework will be validated with relevant stakeholders (cloud consumer / cloud provider) by case study in some countries.

Acknowledgements

The authors would like to thank the Government of Saudi Arabia for supporting this research through a PhD scholarship. The authors also extend their gratitude to Internal Ministry of Saudi Arabia, and to the Saudi cultural bureau in London for facilitating this scholarship and for all their support. The authors would like to thank the Staffordshire University for supporting this research.

References

1. Lui F. NIST Cloud Computing Reference Architecture Recommendations of the National Institute of Standards and. 2011. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:NIST+Cloud+Computing+Reference+Architecture+Recommendations+of+the+National+Institute+of+Standards+and#1.
2. Subashini S, Kavitha V. A survey on security issues in service delivery models of cloud computing. *J Netw Comput Appl*. 2011;34(1):1-11. doi:10.1016/j.jnca.2010.07.006.
3. Felici M, Koulouris T, Pearson S. Accountability for Data Governance in Cloud Ecosystems. *2013 IEEE 5th Int Conf Cloud Comput Technol Sci*. 2013:327-332. doi:10.1109/CloudCom.2013.157.
4. Modi C, Patel D, Borisaniya B. A survey on security issues and solutions at different layers of Cloud computing. 2013;4(07):561-592. doi:10.1007/s11227-012-0831-5.
5. Ioisu WL. Governance Model of Cloud Computing Service. *Natl Pingtung Univ Sci Technol*. 2012;2(7698).
6. Mell P, Grance T. The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology. *Nist Spec Publ*. 2011;145:7. doi:10.1136/emj.2010.096966.
7. Kshetri N. Cloud Computing in Developing Economies. *Computer (Long Beach Calif)*. 2010;43(October):47-55. doi:10.1109/MC.2010.212.
8. Niemi E. Designing a Data Governance Framework. 2011:14.
9. Li Y, Guo L, Guo Y. CACSS: Towards a generic cloud storage service. *CLOSER 2012 - Proc 2nd Int Conf Cloud Comput Serv Sci*. 2012:27-36. http://www.scopus.com/inward/record.url?eid=2-s2.0-84864871568&partnerID=40&md5=2febf3baaefbf549cbb3053d40f90ca6.
10. Catteddu D, Hogben G. The European Network and Information Security Agency (ENISA) is an EU agency created to advance This work takes place in the context of ENISA ' s Emerging and Future Risk programme . C ONTACT DETAILS : This report has been edited by. *Computing*. 2009;72:2009-2013. doi:10.1007/978-3-642-16120-9_9.
11. Wende K. A Model for Data Governance – Organising Accountabilities for Data Quality Management. *Corp Gov*. 2007:417-425. http://aisel.aisnet.org/acis2007/80.
12. Cloud Security Alliance. Cloud Data Governance Working Group. https://cloudsecurityalliance.org/group/cloud-data-governance/. Published 2015. Accessed June 12, 2015.
13. Joy Medved, SSBB,IQCP A. *Data Governance: The Four Critical Success Factors*. San Diego,CA USA; 2014. http://www.slideshare.net/Dataversity/enterprise-data-world-data-governance-the-four-critical-success-factors-42859742.
14. Beach T, Rana O, Rezgui Y. Governance Model for Cloud Computing in Building Information Management.pdf. *IEEE Trans Serv Comput*. 2015;8(2):314-327. doi:10.1109/TSC.2013.50.
15. Groß S, Schill A. Towards user centric data governance and control in the cloud. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2012;7039 LNCS:132-144. doi:10.1007/978-3-642-27585-2_11.
16. Tountopoulos V, Technology A. The problem of cloud data governance. 2014;317550.
17. The Data Governance Institute. Definitions of Data Governance. http://www.datagovernance.com/adg_data_governance_definition/. Published 2015. Accessed February 16, 2015.
18. IBM Institute for Business Value and IBM Strategy and Change. The IBM Data Governance Council Maturity Model : Building a roadmap for effective data governance. *Gov An Int J Policy Adm*. 2007;(October):1-16.
19. Becker JD, Bailey E, Proceedings A. IT Controls and Governance in Cloud Computing. 2014:1-20.
20. Kooper, M., Maes, R., and Roos Lindgreen E. Information Governance as a Holistic Approach to Managing and Leveraging Information Prepared for IBM Corporation. *Int J Inf Manage*. 2011;31.
21. Approach a H. Data Governance for Privacy, Confidentiality and Compliance: A Holistic Approach. *Cycle*. 2010;6:1-7.
22. More P, Lingayat M. Survey on Data Sharing in the Cloud Using Distributed Accountability. 2014;4(10):406-409.
23. Buyya R, Buyya R, Yeo CS, et al. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering

computing as the 5th utility. *Futur Gener Comput Syst*. 2009;25(6):17. doi:10.1016/j.future.2008.12.001.

24. Council CSC. Practical Guide to Cloud Service Level Agreements version 1.0. 2012:1-44. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Practical+Guide+to+Cloud+Service+Level+Agreements#0.

25. Cochran M, Witman PD. Journal of Information Technology Management A Publication of the Association of Management GOVERNANCE AND SERVICE LEVEL AGREEMENT ISSUES IN A CLOUD COMPUTING ENVIRONMENT. 2011;XXII(2).