Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 65 (2015) 342 – 349

Procedia Computer Science

International Conference on Communication, Management and Information Technology (ICCMIT 2015)

Management of student records: Data Access Right Matrix and data sharing

Lionel Khalil [a] *, Marie Khair [a], Joseph Abi Nassif [a]

[a] *Notre Dame University – Louaize, Zouk Mosbeh, Lebanon*

Abstract

This work proposes a procedure to implement a Data Access Policy to ensure the protection of privacy rights of students' records within higher educational systems. First it reviews the general legal constraints applicable to higher Institutional Data management. Second it reviews the principles of a Coordinated Model. The procedure is based on an Access Right Matrix that assigns data access privileges to Data Users. The first purpose is to handle the common and regular access by rightful users to data needed in their daily routine job through operational interface. The main purpose of the procedure is to handle ad-hoc requests that come from outside the university or from some services which do not have formal access to the data. Data access right matrix is used to grant or reject ad hoc requests based on the following criteria: the degree of sensitivity of the data requested, the number of records requested the purpose of the usage, and finally the privileges and trustworthiness of the requester.

Peer-review under responsibility of Universal Society for Applied Research

Keywords: Data Access Policy; Coordinated Model; Access Right Matrix

1. Introduction

During the daily work of administrative support for a Higher Education Institution, it was noticed that the ad-hoc requests for data access is resulting an inefficient use of time and resources. For example, for accurate and efficient advising of students, chairpersons, coordinators, as well as advisors, need to get spontaneous up to date answers to different types of queries. Answers are needed quickly and efficiently taking into consideration all privacy and

Corresponding author. Tel.: +961 9-208118.
E-mail address: lkhalil@ndu.edu.lb

Peer-review under responsibility of Universal Society for Applied Research
doi:10.1016/j.procs.2015.09.094

security requirements. Considering the complexity and availability needs during the registration process, such requests should be optimized for time-efficiency but also should be straightforward enough that they can be easily and effectively managed by all stakeholders.

However, the current system is managing a too high number of ad-hoc request that results an inefficient use of the computer center‘s time and resources. By overloading employee man hours, the Computer center cannot concentrate on more strategic issues. Our solution to the problem of inefficient requests is to introduce this new procedure which will be described later on in this paper.

This work proposes a procedure to implement a Data Access Policy to ensure the protection of privacy rights of students’ records in higher education. First it reviews the general legal constraints applicable to data management in higher Institutions. Second it reviews the principles of a Coordinated Model based on data access policies adopted in five representative higher education institutions in the US. In contrast with the Centralized Model for data access, the chosen Coordinated Model is based on the coordination among a number of key university officials. In a higher education environment, many data types can be identified, depending on their nature and field. In addition, each type of data or part thereof can be assigned a security attribute based on the level of its sensitivity and protection. Under this framework the procedure is based on an Access Right Matrix that assigns data access privileges to Data Users. The first purpose is to handle the common and regular access by rightful users to information needed in their daily routine job through operational interface. The main purpose of the procedure is to handle ad-hoc requests that come from inside or outside the university or from one service that is not the owner of the data. Data Sharing being sensitive, these ad-hoc requests are managed according to the following criteria: the degree of sensibility of the data requested, the number of records requested (Individual record, Aggregated Report or set of records), the purpose of the usage (for internal/external administrative use or for external publication or presentation) and finally the privileges of the requester.

2. Legal and policy frameworks

First we review the general legal constraints applicable to Data management in higher Institutions. Second we do a literature review on the principles of a coordinated and centralized Models for data access in higher education institutions.

2.1. The Legal Framework in Lebanon

Principles of privacy protection are covered by many regional and international guidelines. In 1980 the Organization for Economic Cooperation and Development has issued the “Guidelines on the Protection of Privacy and Transborder Flows of Personal Data,” the Council of Europe has issued in 1981 the “Convention for the Protection of Individuals with Regard to the Automatic Processing of Personal Data”, and in 1999 the United Nations has issued “Guidelines for the Regulation of Computerized Personal Data Files” and the Health Insurance Portability and Accountability Act (HIPAA) of 1996. These guidelines have created the key principles to ensure minimum guarantees of privacy for personal information. It covers all actions that can be made on the data: collection, storage, use, transfer. Those guidelines recognize the right of each person to access its personal data, to update it, to be informed on the methods and the objectives of the data collection’s operations. Finally the guidelines establish the right of the person to have his data destroyed, after the purpose of its collection and processing is achieved.

Most countries have already defined national laws to cover the collection and the use of data privacy protection covering privacy in general and health data in particular.

Any user at the University who needs to access institutional data is required to comply with national laws and the pertinent rules and regulations set by the University on Data privacy and management. There are few Lebanese regulations that we can refer to. The Decree #179 (22/02/1979) prohibits the disclaiming of personal data without the expressed consent of the concerned persons. The Circular 4A (2006) covers data protection and piracy and includes penalties for software privacy and related issues. The Law #140 (27/10/1999) ensures the protection of the right of secrecy of communication. Finally the Lebanese Penal Code covers the violation of secrets by an employee in its Article 579: "anyone who, by reason of its status, function, profession or of his art, knowledge of a secret,

reveal it without just cause, or use it to his personal benefit or the benefit of a third party will be punished, if the offense is likely to cause even moral, to imprisonment for one year and a fine which shall not exceed four hundred thousand pounds". The general legal constraints applicable to higher Institutions for data management in Lebanon are not specifically related to the qualifications of the different IT users and their respective liabilities.

2.2. Coordinated and Centralized Models

In contrast with the Centralized Model for data access, the Coordinated Model is based on the coordination among a number of key university officials to allow access and privileges to different types of users of institutional data at higher education institutions. The Coordinated Model is based on the coordination among a number of data stakeholders including Data users, Data Trustees, Data Stewards, Data Processors, and Data Experts in charge of defining and implementing the access privileges of other users to data. In a higher education environment, Data types are identified according to their nature and field, including Academic Data, Personal Student Data, Administrative Data, Financial Data, Research and Development Data, HR Data. Each type of data or part thereof can be assigned a security attribute, such as Restricted, Limited Access, or Public, based on the level of its sensitivity and protection.

The Coordinated Model for Data Access is described in University report [1]. Several universities have adopted this model. The following five different higher education institutions constitute a representative sample of Coordinated Models for Data Access: Loyola Marymount University [2], Virginia Polytechnic Institute and State University [3], South Carolina University [4], the Ohio State University [5], the University of Virginia [6].

The University of Loyola Marymount Information Security Policy emphases on the management of data access security and the coordination of access including requests for review of data access restrictions and dispute resolution for data access. In addition, this policy covers data retention and physical security. The Virginia Polytechnic Institute and State University Administrative Data Management and Access Policy covers access security administration, but also data administration, including data capture, documentation, storage, validation, correction, collection, maintenance, reporting, archiving, and warehousing. The University of South Carolina Data Access Policy is a very concise six –pages policy covering the Data categories and types, and data stakeholders and the policy implementation.

The Ohio State University Policy on Institutional Data identifies the key responsibilities of Data Stewards as data classification and data access. The policy has a specific paragraph on the need for special protection of restricted data, for the reporting of breaches relating to it, and for proper management of requests for restricted data. The University of Virginia Administrative Data Access Policy covers only Administrative Data which are necessarily shared to ensure the activity of the University, and addresses the requests for its access.

3. Principle of Access Control in Higher Education

Most of the data in higher education institutions is unstructured and probably will be part of the big data evolution. However, currently, most of the data daily used for academic management is structured data. This supports our beginning point of building the access control policy. Previous researchers have already defined security management in higher education which will form the base to introduce the proposed state-of-the art in the access control matrix, and to present recent developments on full distributed access control especially with the introduction of digital rights management in electronic textbooks.

3.1. Security Management in Higher Education

Security is mainly defined as related to the three aspects of Confidentiality, Integrity, and Availability (CIA). Confidentiality means that only authorized parties have the ability to access protected data [7, 8]. Privacy is usually used in the context where the environment encompasses personal related data and deals with the possibility of the user being able to control access to his personal data. A typical example is ensuring that a student does not have read access to another student grades. Integrity means that assets can only be modified by authorized parties and only in

authorized ways. An example is making sure that no unauthorized person can alter a student's academic information. Availability refers to the possibility that authorized users can have access to their data upon need and on time. As an example a university manager has access to the latest report on registered students on time. In addition to support the above CIA security goals, there are additional goals: authentication, authorization and non-repudiation. Authentication is defined as verifying the identity of a requestor. Authorization is defined as allowing this requestor to properly use a specific service. An example is to give the needed credentials for a registrar's user to be able to update the grade for a specific student after properly authenticating him. Non-repudiation is concerned with preventing a denial by one of the users involved in a communication of having participated in all or part of the communication. An example could be the case where a student cannot deny receiving a notification of being on probation.

3.2. Access control policies: Discretionary Access Control, Mandatory Access Control and Role Based Access Control

Traditional access control policies included two major categories: Discretionary Access Control (DAC) and Mandatory Access Control (MAC). While the first policy, DAC, was mainly seen as an appropriate access control for commercial and civilian applications, the second policy MAC was mainly used and implemented in the military environments. DAC was proved to be flexible and easy to implement but it mainly suffers from the information flow vulnerability and it is difficult to manage. The most common model of MAC is the multilevel security policy where access from subjects to objects is based on classes or clearance levels assigned to subjects and labels assigned to objects [9].

Role Based Access Control (RBAC) is a current standard of managing access permissions. Despite the fact that RBAC was initially proposed in the 70's, it did not prevail until the 90's and later on when formally presented by Sandhu [10, 11]. The core idea in RBAC[12, 13] was that permissions are associated with roles, and users are made members of appropriate roles thereby acquiring the roles' permissions. In addition it was proved that RBAC is able to implement both DAC and MAC.

The core concept is based on the matrix model initially proposed by Lampson [14] to protect operating systems. The rows of the matrix list are the users or domains, the columns are the services to control, and in each matrix cell the entry defines the access rights of the respective user has over the respective service. The matrix is used to define all access rights including basic access rights, and administrative rights control including the control of the matrix itself.

3.3. Distributed Access control with public-key certificates

Recent key-oriented discretionary access control systems [15, 16] are based on delegation of access rights with public-key certificates. Decentralization of authority and management of operations allows to build local relations and to set up local authorities that might arise within groups of users in a University [17]. Indeed Haake et al [18] is proposing a dynamic access right management at the hand of the students. They are able to form groups and adapt access rights when changing groups and workspaces. This ability exists already in most of the social networks. Several critics have been raised that a total decentralization leads to anarchy. Furthermore the flexibility of the certificates cannot cover all the needs of a system manager and has to be completed with a centralized control system. For example separation-of-duty policy cannot be expressed with only certificates [19]. Thus introducing distributed access control can technically be done partially and the hybrid operational solution will be more complex to manage than using RBAC.

Nevertheless the current operational use of distributed access control in higher education is already well developed in some Universities to restrict access to a document by using Digital Rights Management (DRM). A DRM is a digital certificate embedded in the file which carries the rights attached within it. Any program supporting DRM is designed to prevent unauthorized user from reading to copying the document [20]. Within a university, DRM

are being used to promote the development of text eBooks instead of paper textbooks. Indeed intellectual property of the authors and editors are respected by restricting the access rights to text eBooks.

4. Implementation of Access Right Matrix procedure

The Access Right Matrix is based on concepts from DAC, MAC and RBAC. The ownership of data is initially attributed to its creator, but later on delegated to the authority that is academically responsible for this data. For example, when a student applies to the university, he needs to fill a form with all his academic, personal and financial data. Once a student is admitted, the ownership of this data will be delegated to the Registrar's office who will continue to be the responsible for his personal and academic data, while the business office will be responsible for the financial data. This continues until the student graduates where the ownership will be transferred to the alumni office.

Inspired from the multilevel security, data is classified to different levels and users are associated with rings of privileges which will govern the granting or revoking access. Finally RBAC is implemented through the application of static roles inspired from the concept of least privileges usually attributed to the users to perform their jobs. For the following, the procedure is divided into four steps: Data classification, User Roles classification, Access Right matrix (or assignment of data access privileges to user roles), and Data Access procedure.

4.1. Data Classification

Data type refers to the nature of information in the database field in which it is used. Typical institutional data types include (but are not limited):

a. Academic Data: Student-related (admissions, personal status, registration), Faculty data (qualifications, teaching, service), course data, Faculty instruction data, Library data.

b. Non-academic Student Data: Student housing, medical, counseling, and disciplinary.

c. Administrative Data: Human resources, facilities, services, maintenance, construction.

d. Financial Data: Payroll, financial aid, expenditures, revenues.

e. Research and Development Data: Faculty research data, research centers' data, publications data, alumni data, international affairs data.

4.2. Security categories

Each type of data or part thereof can be assigned a security attribute, such as "restricted," "limited access," or "public" based on data sensitivity level and the protection needed.

a. Public Data: This category is assigned to all data that is widely available for public use with no restrictions. The volume of this data category is to be maximized through the reduction of unnecessary limitations, in line with the principles of transparency and data availability.

b. Limited Access Data: This is the default data category, deemed by the corresponding Data Steward to be inappropriate for public access. It is made available only to a specific group of University community members, based on their job descriptions (on a need-to-know basis). Non-personal student admission and registration data are examples of limited access data.

c. Restricted Data: Data protected by law or considered critical to University operations. Protected health information, non-public personal information, and personally identifiable information are examples of restricted data.

To set up the Data types, Data types related to the Data Stewards were handled to the concerned offices in order to define the sensitivities of the different fields. By default data should be Limited access, and then it is up to the Data Stewards to increase to Restricted Access data or decrease to the level to Public Access.

4.3. Data Users Classification

Data Users are University community members who access institutional data as part of their job-related functions within the University. The privileges of Data Users are defined in onion rings as per the figure 1. This simple model implies that each level contains all privileges of the below level. This simplification offers more clarity to the multiple authorization levels of the users' roles in the University.

The definition looks straight forward for Administrative and Academic Data. The status of research data created by researcher is an Intellectual Property (IP) issue and depending on the IP Policy the Data Steward could be the researcher and the Data Trustee the head of the research Center. The data management involves three bodies.

a. Data Custodian: the Data Custodian is the Information Technology Department of the University hosting the data storage system.

b. Data Trustees: a Data Trustee is a senior University official, typically at the level of Vice-President who has planning and policy-making responsibilities related to the institutional data.

c. Data Steward: a Data Steward is a University official, typically at the level of Dean, Registrar etc. who has management responsibilities for the data handled in his/her own Unit. Generally reporting to a Data Trustee, the Data Steward overseas the data flow between the Unit and the Data Custodian.

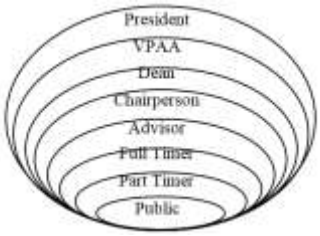

Fig. 1. Onion Rings privileges

4.4. Data Users Classification

For the daily work data is made available to predefined individuals for University-related uses according to their privileges. According to the Data policy the user agrees to make responsible and ethical usage of the data in cooperation with the offices that manage and maintain the data as well as data custodians. Internal or external users may request additional data access from the concerned data stewards. Once the Data Access Form is filled, it will be forwarded to the unit concerned. The reply to the request is taken from the data unit concerned (see Fig. 2).

There are legal issues surrounding the use of data for research purposes (which then may be published) versus administrative purposes: there is a clear distinction between the two usages of the data that should be taken into consideration in the criteria to give access to the data. Data used for research purposes relates to the supervision of the Institutional Research Board (IRB). The role of the IRB is to ask the relevant questions related to data used for research and publication. Any request for data which is publishable or especially requests from an external source (i.e. the public) must get approval from the concerned Data Trustee. Accordingly the conditions of approval/rejection are based on criteria related to both the requester and the sensitivity of the data, mainly:

a. Usage of the data: Will the data be restricted to internal administrative use or will it be disclosed outside? Will the data be used in presentations or publications? Did the study receive an approval from the Institutional Research Board? Where and when will this data be presented and/or published?

b. Security of the data management: How will the data be stored? Who will have access to the data? For how long will the data be stored?

Regarding combined data sets, it is up to the Data Stewards to evaluate and when appropriate reject requests of data combinations that could help identify individuals. Regarding the issue of what the procedure should be if

multiple data stewards receive requests from the same requestor, it is worth noting that he/she could infer sensitive data by combining less sensitive and even public data. This is why it is important to define operational areas so that when requests are submitted, all concerned Data Steward should agree; especially when data is shared between several offices.

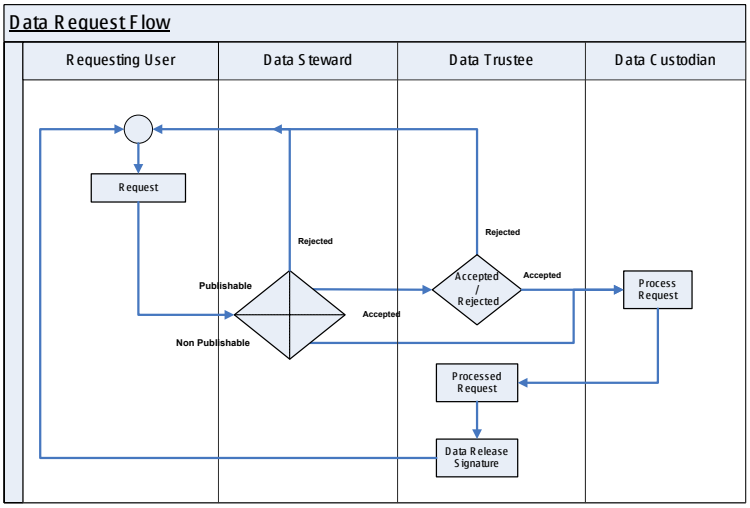

Fig. 2. Workflow of the Data request form

5. Conclusion

In future work, it is intended to extend the model to Human Resources data, and to Instructor files including Course portfolio and research portfolio; to finalize the procedure it is intended to develop a detailed informative and transparent appeal system which can be implemented when the request is rejected.

The legislation for Data Privacy in the Middle East is still ongoing. The implementation of the coordinated model reveals grey areas of decision making when the data sets requested belong to different Data Stewards and possible security holes. The flexibility to transfer data is a favorable environment for unwilling data spread. Each University has to develop awareness of the importance and the sensitivity of the data.

References

1. Hassoun G. *Data Access Policy*. Notre Dame University – Louaize; 2011. p.12-16
2. Administration Division. *Policies and Procedures Manual – Information Security Policy*. Loyola Marymount University; 2009. http://www.lmu.edu/AssetFactory.aspx?did=37798 (Retrieved on 20/4/2015)
3. Virginia Polytechnic Institute and State University. *Administrative Data Management and Access Policy*. Virginia Polytechnic Institute and State University; 2008. http://www.policies.vt.edu/7100.pdf (Retrieved on 20/4/2015)
4. President's Office, *Data Access Policy*, The University of South Carolina; 2010. http://www.sc.edu/policies/univ150.pdf (Retrieved on 20/4/2015)
5. Policy Review and Update Task Group. *Policy on Institutional Data*. The Ohio State University; 2007. https://ocio.osu.edu/sites/default/files/assets/Policies/InstitutionalData.pdf (Retrieved on 20/4/2015)

6. Information Technology and Communication. *University of Virginia Administrative Data Access Policy*. The University of Virginia; 2001. http://itc.virginia.edu/policy/admindataaccess.html (Retrieved on 20/4/2015)
7 Cheung SK. Information Security Management for Higher Education Institutions. *Intelligent Data analysis and its Applications*, 2014; 1: 11-19.
8 Patel A, Ghaghda S, Nagechac P. Model for security in wired and wireless network for education. *Computing for Sustainable Global Development (INDIACom);* 2014. p. 699-704
9. Joshi, JB, Aref WG, Ghafoor A., Spafford EH. Security models for web-based applications. *Communications of the ACM* 2001;44:38-44.
10. Sandhu, R. Lattice-based access control models. *Computer* 1993; 26:9-19.
11. Sandhu R, Coyne J, Feinstein H, You,an C. Role-based access control models. *IEEE Computer 1996; 29:38-47*
12. Sandhu R, Qamar M. How to do discretionary access control using roles. *Proceedings of the third ACM workshop on Role-based access control,* 1998.
13. Ferraiolo D, Sandhu R. Proposed NIST Standard for Role-Based Access Control. *ACM Trans. on Information and System Security (TISSEC)* 2001;4(3): 224-274
14. Lampson B. Protection. *ACM Operation Systems Review* 1974; 8:18–24.
15. Blaze M, Feigenbaum J, Strauss M. Decentralized Trust Management. *Proceedings 1996 IEEE Symposium on Security and Privacy;*1996. p.164-173.
16. Gasser M. Goldstein Q, Kaufmn C, Lampson B. The digital distributed system security architecture. *Proceedings National Computer security conference;* 1989. p.305-319.
17. Aura T. Distributed access-rights management with delegation certificates. *Secure Internet Programming;* 1999. p. 211-235. http://www.research.rutgers.edu/~serban/cs547/aura-lncs1603.pdf (Retrieved on 20/4/2015).
18. Haake J M, Haake A, Schümmer T, Bourimi M, Landgraf B. End-user controlled group formation and access rights management in a shared workspace system. *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, 2004. p. 554-563 http://www.fernunihagen.de/imperia/md/content/fakultaetfuermathematikundinformatik/forschung/berichte/bericht_315.pdf (Retrieved on 20/4/2015)
19. Brewer D. The Chinese wall security policy. *Processing IEEE Symposium on Research in Security and Privacy;* 1989. p. 206-214.
20. Liu Q, Safavi-Naini R, Sheppard N P. Digital rights management for content distribution. *Proceedings of the Australasian information security workshop conference on ACSW frontiers' 2003* 2003; 21:49-58. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.12.3484&rep=rep1&type=pdf (Retrieved on 20/4/2015).