



Queensland University of Technology
Brisbane Australia

This is the author's version of a work that was submitted/accepted for publication in the following source:

Burdon, Mark, Gonzalez Nieto, Juan M., Christensen, Sharon A., Dawson, Edward P., Duncan, William D., & Lane, Bill B. (2007) Access Control in Federated Databases: How Legal Issues Shape Security. In Wimmer, Maria A., Scholl, Jochen, & Gronlund, Ake (Eds.) *Electronic Government, 6th International Conference, EGOV 2007*, 3-7 September, Regensburg, Germany.

This file was downloaded from: <http://eprints.qut.edu.au/9411/>

© Copyright 2007 Springer

Notice: *Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source:*

http://dx.doi.org/10.1007/978-3-540-74444-3_20

Access Control in Federated Databases: How Legal Issues Shape Security

Authors.

Institution
{authors} @email

Abstract. This paper will examine how legal considerations arising from the aggregation of data impact upon technical access control mechanisms. Research findings are based on a multi-disciplinary investigation of security issues regarding the aggregation of data in a governmental federated database system. The researchers conclude that the development of a federated architecture must consider technical security concerns within the context of legal risk management issues. As such, a holistic approach to the investigation of information security is required that incorporates the disciplines of information technology and law.

Keywords: Legal, societal and cultural aspects of eGovernment; Inter- and multidisciplinary research – issues and examples, Enterprise architectures and whole-of-government approaches; Trust and security: provisions and instruments.

Submitted to E-Gov Track: Completed Research Papers

1. Introduction

In this paper, we outline key issues regarding the development of a governmental federated database system to aggregate geo-spatial data. Specifically, we focus on how legal risk management concepts, such as, liability and compliance, impact upon technical architectures. The paper will then proceed to outline how particular legal issues, such as, information management concerns and public record keeping requirements can shape the form and location of access control measures.

This paper aims to contribute to the development of federated database systems in government by highlighting the interdependent relationship that the disciplines of law and information technology have on each other, which can affect overall architectural design and subsequent implementation of security measures. It is likely that federated database systems, to aggregate data of all kinds, will become a common feature of e-

government data sharing projects given the traditional “silo” based information structures of government departments. As such, it is important to identify all potential constraints to development and include technical, legal and other concerns that may inhibit the successful implementation of federated database projects. Whilst the paper is focused on research conducted within Queensland Government, many of the concepts outlined have potential application in other jurisdictions because the issues raised are not unique to the Australian situation. The issues highlighted in the paper are equally applicable to governments throughout the world, especially those that are aiming to establish federated database systems to aggregate data. The paper provides an indication of how fundamental legal concepts can shape security designs and future research will make further contributions to the literature by examining some key e-government issues, such as, multi-disciplinary research models, information sharing and implementation strategies.

The paper is structured as follows. In Section 2, we outline the background to the project including the purpose of the research and the methodology adopted. Section 3 details the proposed technical architecture. Section 4 provides a high level examination of legal concepts that are relevant to the technical architecture and the aggregation of data. Section 5 outlines key information security issues, with particular focus on access control measures and describes how legal considerations impact upon the design and implementation of security measures. Finally, Section 6 concludes the paper and briefly details future work.

2. Background

Governments throughout the world have been collecting geo-spatial information for a number of years. Despite this, it is not until relatively recently that governments have started to realize new applications for geo-spatial data held under their custodianship. This is primarily due to technological developments that have made the aggregation of data more feasible and more readily realizable. In particular, it is now widely recognized that benefits can be gained from the aggregation of geo-spatial data which provides new insight for policy making and opens up new commercial opportunities by bringing together different data sets and overlaying data into a single geo-spatial representation [14].

Problems can arise in the aggregation of geo-spatial data in government federated database systems because existing data has traditionally been collected and held within separate agency “data silos” [15]. Thus far, geo-spatial data has generally been collected for the individual purposes of different agencies rather than for the benefit of government as a whole. As such, individual agency data collection has been conducted independent of other government agencies. Concerns are further compounded because it is common for each individual data silo to have different data life-cycles and to be subject to different information management and security frameworks. The development of a government federated database system for the aggregation of geo-spatial data therefore has technical issues enmeshed with legal and risk management concerns involving information management [16].

In this context, [Organisation1] [9] is a major Queensland Government initiative to provide a central Internet portal for the dissemination of geo-spatial information held by different Queensland Government agencies. A federated database system is being developed to establish interoperability of different agency information systems and to enhance geo-spatial data sharing across Queensland Government. Essentially, [Organisation1] will lead to the development of a whole of government publication strategy for geo-spatial data [10].

Researchers from the [Institution] have embarked on a three year research project with [Organisation1] funded by the [Organisation2] [6] and the [Organisation3] [19]. The aim of the project is to develop a unified security and legal framework for [Organisation1]. The framework will incorporate multiple agency geo-spatial datasets even though each dataset will have their own individual security and information architecture.

We adopt a multidisciplinary methodological approach to security that encompasses technical analysis with legal and risk management issues. A comprehensive technical analysis involves consideration of possible security architectures for access control. Legal research entails an investigation of different areas of law that could impact upon the consideration of those security architectures, such as, information management concerns, public recordkeeping requirements and liability issues.

During the first phase of the project, each disciplinary group undertook reviews of the extant literature to obtain an understanding of key issues. Qualitative semi-structured interviews were conducted with key Queensland Government personnel to gain knowledge about the current technical/legal/risk environment at an agency level to further research input into the design of the federated technical architecture for the aggregation of data.

3. Technical Architecture

Figure 1 details the proposed technical architecture for the [Organisation1] federated database system. The architecture is composed of three tiers:

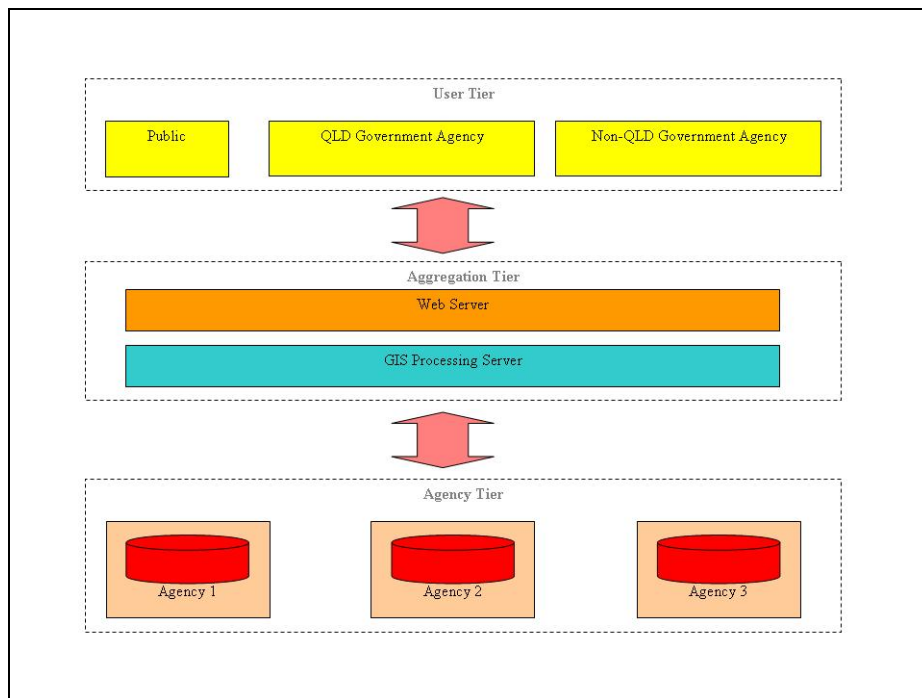
- The Agency Tier: At the lower level, the Agency Tier consists of the different government agencies that provide geo-spatial data to be aggregated. The agencies hold the data within their own information systems and have custodial responsibilities over the stored data. The data is transferred from the agencies to be aggregated in the next level of the architecture;
- The Aggregation Tier: At the mid-section, the actual aggregation of agency data takes place. It is likely that the tier will consist of a geo-data processing server that aggregates the data and a web server that publishes the aggregated output to the User Tier. Effectively, the Aggregation Tier mediates between user requests for data and agency responses. User requests are decomposed and data requests passed on to the appropriate agencies. Individual agency responses are aggregated and served back to users; and

4 Authors.

- The User Tier: The highest tier represents the users of the system. The project has been structured into different phases that gradually expand the user network from internal agencies to include other government agencies and eventually to the public. Users will be allowed to make requests, typically through a web browser, and view aggregated outputs.

The arrows represent the flow of data and communications between all three layers. In practice, request communications flow down through the architecture whilst data is pushed up to satisfy requests.

Fig. 1. Technical Architecture



4. Legal Concepts Arising from the Aggregation of Data

Two legal concepts are prevalent to the development of the technical architecture and the subsequent aggregation of data – future liabilities arising from the publication of erroneous information and compliance with existing relevant legislation and standards. Both of these concepts are relevant because the ultimate output for the aggregated data will be the public.

Traditionally, data access within government has been restricted to users within an agency, but with an integrated system, the pool of potential data users is widened substantially. It is clearly understood within government, that one agency can not sue another agency if the integrity of information is not maintained. If information published by one agency is published to another agency and the second agency acts upon or relies upon the information published by the first agency, the second agency cannot bring proceedings against the first agency. This changes when government publishes information to third parties outside of government and to the public because liabilities can flow from reliance upon inaccurate information. In turn, this can impact upon the location and the design of information security measures used to secure the technical architecture.

The publication of incorrect data gives rise to liability issues under contract law (i.e. misleading or deceptive behavior), tort law (i.e. negligence or negligent misstatement) or from a specific statutory provision that places certain obligations on the accuracy or correctness of data and how that data should be used [2]. Liability issues are further complicated in federated database systems used to aggregate data because different datasets and different map layers may have varying degrees of accuracy [11]. Given the federated nature of aggregated data distribution, it can be problematic to establish which particular piece of data is responsible for an error, and in turn, which organization is responsible for supplying the inaccurate data. From a government organizational and legal risk management perspective, the issue at the heart of liability is therefore who is legally accountable for providing the erroneous data [1]. This is an important point. In theory, a legal action will be brought against government as a whole, but in practice, funds to cover the legal action will have to be found from existing agency budgets. It is likely that the agency that provided the incorrect data will ultimately bear the burden of paying legal fees for a subsequent action.

Although there has not been an Australian case regarding liability from the aggregation of geo-spatial data, it is likely that a future legal action would refer to US case law regarding the accuracy of maps given the analogous nature of common law analysis [3]. Under US law, it is possible that liability for inaccurate information can arise from inaccurate maps where (a) a map is based on erroneous data and (b) where a map is based on accurate data but the representation of that data is inaccurate [7]. This is a distinction between issues of data content and issues of data context [1]. Issues of data content refer to the accuracy of data itself. Issues of data context refer to the notion that the aggregated representation properly represents the data upon which it is based, i.e. the map, or in this case, the aggregated spatial output is an accurate representation of the agency data provided. This point is critical because potential liabilities may arise in both the data held by agencies at the Agency Tier and by the subsequent aggregation of that data in the Aggregation Tier.

Information management structures are therefore a key concern in the legal analysis of the technical architecture and the aggregation of data. It is important that these structures are compliant with existing information management standards because this can be a method of mitigating potential liabilities [18]. An organization that has complied with recognized standards will have a stronger argument in any future legal action because it can claim that it took all possible actions to avoid a risk of harm from arising [4]. This is particularly relevant in Queensland because Section

35 of the Civil Liability Act 2003 (QLD) [5] acknowledges the financial constraints that public authorities face and recognizes that an agency can only provide a level of service that it is funded to provide. If an agency can show that it has properly exercised its functions and it can demonstrate that it has complied with general procedures and applicable standards, then Section 35 could have the effect of mitigating liability because the subsequent harm was beyond the resources available to the agency.

Compliance issues also arise in situations where government agencies are obliged by legislation or regulation to act in a specific way. For example, the public good emanating from accurate recordkeeping by governmental organizations is recognized by the statutory obligations placed on agencies to record, maintain and destroy records within certain guidelines. Queensland Government is no exception. The Public Records Act 2002 (QLD) [17] provides guidelines for agency recordkeeping which are supported by a range of information standards. Furthermore, the Queensland Financial Management Standard 1997 [8] engenders a governance framework that applies to all Queensland State Government agencies which requires that an agency develop a strategic and operational plan for each financial year. The Standard also requires agencies to implement key information security measures. Section 70 requires agencies to develop information systems to provide for certain fundamental elements such as access controls and audit trails. Section 71(1) requires that an agency must develop and implement internal controls to ensure the effective, efficient and economical management of the agency's resources and to accomplish the agency's strategic goals. With regards to security of information systems, Section 71(2) requires agencies to provide for certain mandatory internal controls such as authorization and authentication mechanisms.

An examination of these two underlying legal concepts reveals an understanding of how legal principles can impact upon the design of technical architectures for the aggregation of data. We will proceed now to show how legal issues – data custodianship, retaining public records and financial management – can affect the design of specific information security measures.

5. Legal Issues That Shape Security Measures

As highlighted above, the information management structure used within the technical architecture is a crucial concern both in terms of future liabilities and compliance with existing laws. In Australia, the custodianship model is becoming the prevailing information management system to co-ordinate and provide a control structure for the effective management of aggregated data [13] [12]. A data custodian can be defined as a public official who has physical and legal custody of data, and public records, and holds this information on behalf of a corporate entity or government agency [3]. Information management responsibilities are concentrated in data custodians and their role is essentially to be an information trustee that holds government data for the benefit of the public. Individual agencies retain custodianship over particular data sets but whole of government endeavors are made easier to realize

through the standardization of corporate-wide practices that reduce duplication of data and maximize value added product development [4].

On a day-to-day management level, data custodians ultimately decide what data is collected, aggregated and released to the public. Data custodians are based at the Agency Tier and are responsible for data quality (including the integrity, security and confidentiality of data), availability of data and access to data. As such, they play a key role in the development of a security structure for a governmental federated database system used to aggregate data. In turn, the development of a federated aggregation system impacts on the data custodian model because custodians do not have total control over the uses of their data.

The data custodianship concept remains unchanged if the agency has management control over the data retained in its possession. However, when the data leaves the agency, as it does in a federated database system, that management control passes from the agency because another organization now has the opportunity to modify, manipulate or delete their data. Furthermore, agency data custodians cannot be accountable for potential liability arising from data context, i.e. inaccurate spatial representations based on accurate data, because the aggregated data it is effectively outside the control of agencies. Therefore, it is unlikely or at least very difficult for the agency data custodian concept to apply to aggregated data outputs produced in the Aggregation Tier.

Limited or no custodian responsibilities at the Aggregation Tier also has consequences for the design, location and use of security measures because data custodians have a responsibility to ensure appropriate security procedures for their data. In Queensland, as highlighted above, this is mandated by Section 71(2) of the Financial Management Standard. Data custodianship is therefore legally relevant to security issues regarding authorization policies.

Authorization policies across different agencies may vary widely. Agencies are generally unaware of what other data will be used with their data to form an aggregated output. It is therefore difficult for an agency to devise authorization policies that predicate on aggregate data. As such, harmonizing established authorization rules at the Agency Tier for future adoption and use in the Aggregated Tier is a major challenge.

It is also likely that there will be different data classification schemes used by individual agencies. Inconsistent data classification is a potential problem because different agencies can apply different classifications to the same data or can use the same classifications for different data. Even assuming that different agencies could develop a consistent data classification scheme, another concern arises with the aggregated information itself. The aggregation of data contained in separate data sets may indicate information which is not intended for disclosure. For example, combining electricity grid and water reticulation maps may reveal information that would normally be made secret for defense and security related reasons.

Data custodianship concerns also impact on the positioning of access controls. As there are different government agencies, each with a significant number of data sets, it is important to consider where authorization and authentication should be performed:

- (a) At the Aggregation Tier; and/or

- (b) At the Agency Tier, either at the agency's perimeter (e.g. web services gateway) or at the point of access to the data (i.e. using built-in data base mechanisms).

If authorization policies are produced at the Aggregation Tier only, this may not be consistent with agency data custodial responsibilities. Furthermore, it is far from clear who will be accountable for devising policies as existing management structures are based at the Agency Tier. If authorization policies are created at the Agency Tier only, this precludes the application of authorization being considered for aggregated data. It seems unlikely that individual agencies would be considered wholly responsible for access control over data which they only partially hold, so they would have a limited role in determining access control rules for aggregated data. This responsibility would be better suited to the Aggregation Tier. It can be concluded that effective authorization policies should incorporate both policies from individual agencies, for data directly under their control, as well as policies by a different organization at the Aggregation Tier for aggregated data. However, this is an area requiring further research from both a legal and an information security point of view.

If authentication is performed at the Aggregation Tier, on behalf of the agencies, then the access control functions performed by agencies are critically dependent on the organization bearing management responsibility of the Aggregation Tier. Again, this may not be consistent with individual agency custodial responsibilities. Furthermore, the consequences of compromise of the authentication service are important. If the authentication service is compromised then this exposes the data of all of the agencies. If however authentication is performed at the Agency Tier, for aggregation requests which require data from multiple agencies, authentication is performed in each agency, reducing the efficiency of the request processing.

It is also questionable whether the instigation of authentication mechanisms solely at the Agency Tier will fulfill the requirements of the Financial Management Standard. Under the Standard, an agency must develop a strategic plan for the use of ICT resources within a whole of government context. In particular, the plan must evaluate the agency's requirements regarding existing and additional ICT resources and state how the agency will optimize the use of, and fund, existing and future ICT resources. It is not clear whether housing multiple authentication mechanisms at the Agency Tier, for each agency, would fulfill those financial management obligations that require a whole of government outlook.

The practical legal effects of data custodianship also manifest in recordkeeping and the retention of public records concerns. The recordkeeping and retention of public records is legally necessary to provide historical records, for example, evidence in the case of disputes arising over data alleged to have been obtained. Moreover, the Public Records Act places wide-ranging obligations on agencies to keep records of their activities. Section 6 of the Act defines a public record in a broad manner to effectively cover any information generated or received by an authority within its normal duties. A public record can also include a copy of a public record. Section 7 of the Act mandates agencies to keep full and accurate records of its activities. Section 7 also indicates that public recordkeeping and archiving activities should be in compliance with relevant standards and guidelines.

The legal issue of recordkeeping and retention of public records is probably of a lesser concern at the Agency Tier because the individual agencies should already have recordkeeping and record retaining functions in place as part of their normal day-to-day management activities. The supply of data from an agency to the Aggregation Tier could be classed as an activity of an individual agency. If that is the case, agencies may be required to keep full and accurate records of the data provided. Furthermore, given the potential liability issues arising from aggregated representations (i.e., data context accuracy), it would be legally advisable to keep records for every transaction between the Aggregation Tier and the User Tier, particularly involving members of the public, so it could be definitively proven which aggregated representation was provided to which user. This evidence would be crucial in any subsequent legal action.

A method to trace aggregated data is therefore required to identify and to correct source data. In other words, a mechanism may be needed, when it is provided with a piece of aggregated data, it identifies the component parts and the corresponding agencies from which that data was obtained. Being able to trace component data to its custodian may be essential in resolving liability disputes and who funds legal actions.

6. Conclusion

In this paper, we have outlined how fundamental legal concepts and interrelated legal issues can impact upon the design, the development and the location of security measures in a government federated database system for the aggregation of geo-spatial data. Legal and technical issues are enmeshed together because the legal concepts of liability and compliance need to be factored into the design of technical architectures. As such, it would be beneficial if the technical architecture took into account the possibility of future liabilities arising from aggregated publication at the very onset. Whilst it is possible to mitigate liability from the publication of incorrect data through purely legal mechanisms, such as disclaimers, the very structure of the technical architecture can also assist by acknowledging the crucial importance that information management structures have on technical and legal issues.

Specifically, this paper outlined how legal issues such as information management concerns, in the form of data custodianship, public recordkeeping requirements and financial management standards can impact on security mechanisms such as access controls, authorization policies and authentication mechanisms.

This paper represents research findings from the first phase of a three year project. Future work will continue to focus on the issues raised in this article and will ultimately seek to develop a multi-disciplinary methodological model that incorporates the academic disciplines of law, risk and information technology to provide a method of analysis, and a paradigm for discourse that frames research questions, regarding the aggregation of data in governmental federated database systems. This methodological model will provide a truly holistic outlook that recognizes and incorporates the different disciplinary requirements involved in the future development of governmental federated database systems and the subsequent aggregation of geo-spatial data.

References

- [1] Aronoff, S., *Geographic information systems : a management perspective*. Ottawa: WDL Publications, 1989.
- [2] Blakemore, M., "Access and security issues in the provision of geographic information," in *Metada in the Geosciences*, Mdeyckj-Scott, D. e. a., Ed. Loughbrough: Group D Publications, 1991, pp. 55-68.
- [3] Cho, G., *Geographic information science : mastering the legal issues*. Hoboken, NJ: Wiley & Sons Inc., 2005.
- [4] Cho, G., *Geographic information systems and the law : mapping the legal frontiers*. Chichester, England New York: J. Wiley & Sons, 1998.
- [5] *Civil Liability Act 2003 (QLD)*,
<http://www.legislation.qld.gov.au/LEGISLTN/ACTS/2003/03AC016.pdf>
- [6] Reference 1
- [7] Epstein, E. F., "Legal aspects of geographic information systems," in *Geographic information systems: principles and applications*, Maguire, D. J., Goodchild, M. F., and Rhind, D. W., Eds. London: Longman, 1991, pp. 489-502.
- [8] *Financial Management Standard 1997 (QLD)*,
http://www.austlii.edu.au/au/legis/qld/consol_reg/fms1997216/
- [9] Reference 2
- [10] Reference 3
- [11] Joffe, B., "How good are your maps," in *GeoWorld*, 2002.
- [12] Mason, R., "Developing Australian spatial data policies - existing practices and future strategies," in *School of Geomatic Engineering*. Sydney: UNSW, 2000, pp. 344.
- [13] Office of Spatial Data Management, "Australian Government Custodianship Guidelines,," 2001.
- [14] Onsrud, H. J., "Liability in the Use of Geographic Information Systems and Geographic Data Sets," in *Geographic Information Systems: Principles, Techniques, Management, and Applications*, Maguire, Goodchild, Rhind, and Longley, Eds. New York: Wiley, 1999.
- [15] Onsrud, H. J., "The Role of Law in Impeding and Facilitating the Sharing of Geographic Information," in *Sharing Geographic Information*, Onsrud, H. J. and Rushton, G., Eds. Rutgers: CUPR Press, 1995, pp. 292-306.
- [16] Onsrud, H. J., Poore, B., Rugg, R., Taupier, R., and Wiggins, L., "The future of spatial information infrastructure," in *A research agenda for information science*, McMaster, R. B. and Usery, E. L., Eds. Boca Raton, Fla.: CRC Press, 2005, pp. 225-255.
- [17] *Public Records Act 2002 (QLD)*,
http://www.austlii.edu.au/au/legis/qld/consol_act/prs2002153/
- [18] Reid, K., Clark, E., and Cho, G., "Legal risk management for geographic information systems," *Journal of Law and Information Science*, vol. 7, pp. 169-207, 1996.
- [19] Reference 4