2008

The state of public data availability in Australia: a study of suppliers of critical infrastructure information

Roba Abbas
University of Wollongong, roba@uow.edu.au

The state of public data availability in Australia: a study of suppliers of critical infrastructure information

Abstract

The purpose of this study is to evaluate the public data availability situation in Australia, and the consequent impact on the nation's critical infrastructure (CI) and the critical infrastructure protection (CIP) process in general, through an evaluation of data supplying bodies in Australia. An assessment of data suppliers was conducted, in order to allow for the categorisation of data supplying entities, and to identify the critical infrastructure data that is available, The public data availability situation in Australia is described, and the concerns associated with having CI-related information available in the public domain are highlighted.

Disciplines

Physical Sciences and Mathematics

Publication Details

Abbas, R. (2008). The state of public data availability in Australia: a study of suppliers of critical infrastructure information. In K. Michael & M. G. Michael (Eds.), Australia and the New Technologies: Evidence Based Policy in Public Administration. The Third Workshop on the Social Implications of National Security (pp. 146-155). Wollongong: University of Wollongong Press.

17

The state of public data availability in Australia: A study of suppliers of critical infrastructure information

Roba Abbas

Graduate, School of Information Systems and Technology, Faculty of Informatics, University of Wollongong

Abstract

The purpose of this study is to evaluate the public data availability situation in Australia, and the consequent impact on the nation's critical infrastructure (CI) and the critical infrastructure protection (CIP) process in general, through an evaluation of data supplying bodies in Australia. An assessment of data suppliers was conducted, in order to allow for the categorisation of data supplying entities, and to identify the critical infrastructure data that is available. The public data availability situation in Australia is described, and the concerns associated with having CI-related information available in the public domain are highlighted.

Keywords: public data, critical infrastructure protection, threat, risks, information

1 Introduction

Critical infrastructure protection (CIP) has been a global concern since the Cold War. However, the issue has gained increased prominence in Australia since the incidents ofY2K, September 11 and Bali 2002 (Luiijf and Klaver, 2004; Emergency Management Australia, 2003). Additionally, the importance and increased use of the Internet and Information and Communication Technologies (ICT) have amplified the risks on critical infrastructure items (Popp et. al., 2004). These technologies provide outlets for data/information exchange, and have simplified the ability to transmit and access data; in particular, 'sensitive but unclassified' data that, when combined, enable inferences or previously undesired patterns to emerge.

Traditionally, the focus of CIP has been on the three major stages of vulnerability identification, risk assessment and risk management. A study conducted by Breeding in 2003 introduced the risk of 'sensitive but unclassified' data to America's infrastructure, viewing the threat on CIP from an alternative viewpoint. 'Sensitive but unclassified' data refers to information that may not on its own appear harmful, but when amalgamated with

additional data elements can be truly revealing about CI. The outcomes of Breeding's research indicated that openly available information concerning America's critical infrastructure could prove damaging, in that they allow inferences to be made, which could consequently compromise any protection efforts by providing valuable details relating to the weak points and interdependencies between infrastructure items.

The purpose of this study, conducted in late 2006, was to adapt Breeding's process to an Australian setting, by identifying the public data suppliers in Australia, and consequently the amount of data that can be gathered in the public domain relating to Australia's CI. A public data supplier in this instance refers to any individual, institution or body that supports or facilitates the open distribution and use of information concerning Australia's critical infrastructure. The data may be deliberately or indirectly provided. Of particular importance to this paper are the identification of relevant data sources, and the classification of the types of information that exist in the public domain.

2 The data collection process

Data can be categorised in many ways. This study is focused specifically on free and commercial public data, which can be accessed physically and/or online, and includes multiple formats such as images, text, video, maps, geographic coordinates and statistics. The data of interest is critical infrastructure-related data, which refers to data that reveal certain aspects about Australia's critical infrastructure. The aim of this research was primarily to collect CI-related public data from data supplying agencies in Australia, utilising an Internet-enabled computer, and word processing and spreadsheet software as the primary tools for collection. The initial stages of the study included the identification and categorisation of data supplying entities, after which a repository was created using the available data from the identified agencies.

3 Data supplier categories

An assessment of the CI-related public data landscape in Australia enabled the identification of a number of distinct data supplier categories. The classes primarily include the individual/physical data collection, Government bodies, commercial suppliers, and other entities such as utility companies and telecommunications providers.

An interesting observation made throughout this study is that a majority of the supplying bodies enable free access to information, in an attempt to increase knowledge and educate their intended audience. This can be referred to as an optimistic view of data provisioning, in that the data is provided merely as a tool to assist individuals in better accomplishing certain tasks. The potentially undesirable consequences are therefore disregarded to an extent.

The subsequent sections provide an introduction to each data supplier category, including an evaluation of the types of data that were obtained while profiling the suppliers.

4 The individual/physical

The *individual/physical* represents first-hand data collection, where an individual independently collects critical infrastructure data from their surroundings. It is perhaps the simplest method of data gathering and access, as in most cases it does not require the assistance of a third party that may influence or prohibit the collection process. Certain data may require the use of entities such as libraries, Councils and information desks; however, others such as the capturing of video, photographs, GPS points and audio are independent activities. Publicly available critical infrastructure data that may physically be collected includes: tourist guides/brochures; hardcopy statistics, books and magazines; maps; photographs; video and audio recordings; and geographic coordinates using a Global Positioning System (GPS) tool.

First-hand data collection may be considered the simplest method of data gathering, given that in most instances there are very few mechanisms in place to screen the individual gathering the data. Additionally, once the data has been collected, there are minimal (and almost non-existent) enforcement techniques that govern the manner in which the critical infrastructure data may be used.

This source of public data collection was included in the scope of the study to ensure its completeness. However, it is difficult to practically introduce stringent security mechanisms to prevent such data from being accessed. To do so would completely compromise the basic principles of a trusting and informative society and community. As a result, the focus of the remainder of this paper is on electronic public data access, which is facilitated by the supplying agencies discussed below.

5 Federal and State Government departments

Government departments in Australia are prominent suppliers of public data due chiefly to their focus on providing an open and accessible data network through an electronic-Government (or e-Government) portal.

The e-Government movement of recent years has resulted in the trend to provide effortless access to information, in addition to the need to process information in an electronic environment (Wunnava and Reddy, 2000). e-Government "refers to the use of ICTs to promote more efficient and effective government services, allow greater public access to information and make government more accountable to citizens" (Punia and Saxena, 2004, p. 500). Many western countries such as Britain, Canada, the United States, and most notably Australia have been actively involved in e-Government initiatives since the mid-1990s (Lee et. al., 2005).

Traditionally, the aim of e-Government was to allow the public to monitor the activities of its government. However, authors such as Givens (n.d.) feel that e-Government is now centred on public records access, resulting in a number of negative implications. That is, the available records may be used for secondary purposes, such as to make inferences and perform data mining activities, which can reveal undesired patterns about entities. However, positive implications do exist and signify a more informed and knowledgeable public, and a certain level of trust between Government and citizens, which is a desired and productive outcome. However, as Givens (n.d.) states, the negative consequences must also be addressed.

The Australian Government is becoming progressively sophisticated in its provision of e-Government services, with government departments at all levels moving towards interactive services delivery (Davey, 2005). Additionally, the number of citizens making use of the electronic portal is increasing, which is greatly attributed to the advanced and information-rich web pages.

Given this concept of e-Government, this study involved an examination of the Australian Government portal (Australia.gov.au), and its link with other CI-related Australian Government departments. Australia.gov.au provides a gateway that connects the government with Australian citizens; "It links to information and services on over 700 Australian Government websites as well as selected state and territory resources. Australia.gov.au also searches over five million government web pages" (Australian Government, 2006).

Australia.gov.au was utilised as an initial point of analysis; that is, it was used to locate independent government agency websites, and to profile each department with respect to the available critical infrastructure information.

Notably, this study involved identifying Federal and State Government departments that provide information relating to Australia's Critical Infrastructure, and profiling the websites of the respective agencies. These agencies are identified in Figure 1. At the State Government level, New South Wales (NSW) was used as a representative example (case study), as it was not feasible at the time of the study to extend the scope to include all Australian states and territories.

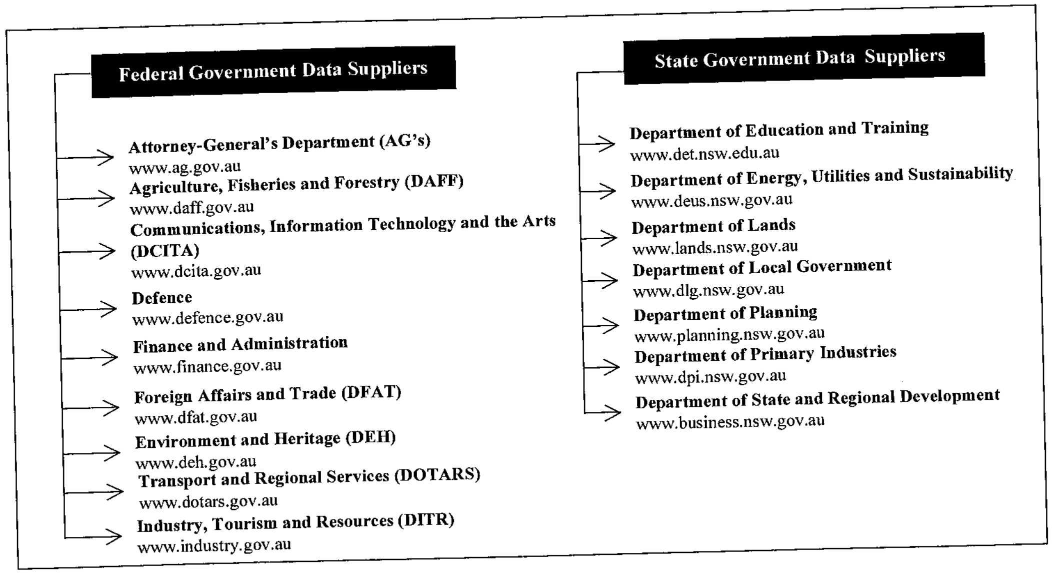
The Government data collection process highlighted the ease with which data can be collected from Government-related websites. Core components of the Government's role are to ensure the nation's security and facilitate access to information. The assessment presented the issue of whether the national security process and in particular critical infrastructure protection can potentially be impeded by the availability of critical infrastructure data. More specifically, where the data collection process presented can be carried out more exhaustively by an individual's intent in compromising Australia's CI for various reasons such as vandalism, competitive intelligence, theft, fraud and terrorism.

While it is important to achieve a balance between open information access and data access restrictions/censorship in the interest of national security, this study highlighted the need to address the following questions in the Government arena: How much data should be provided to the public? Is it necessary that Government be completely open to the public in view of data provisioning? Should data be provided to community members based on their profile or need for the data? Should data be openly available to all citizens? Can data be categorised based on its sensitivity, and the appropriate restrictions be applied? Is public data availability and e-Government impeding the critical infrastructure protection process? What are the future steps for Government with respect to this situation?

6 Commercial data sources

Commercial data suppliers are bodies that provide products and/or services to their customers for a certain price, and under particular conditions. More importantly, they are involved in providing Government with data, which is ultimately purchased for internal Government and public use. The commercial bodies of interest to this study are those that provide information about critical infrastructure, and physical entities at a specific location. That is, they are involved in the provision of spatial data that can be represented on a map, with the associated geographic coordinates.

At the time of this study, the four major data supplying agencies of interest were MapData Sciences, PSMA, Sensis, and MapInfo. As was the case with the Government assessment, this selection of commercial suppliers is not complete, but rather was used to illustrate the commercial public data situation, and to provide an overview of the types of products and information that can be acquired.



Federal Government Data Suppliers

Attorney-General's Department (AG's)
www.ag.gov.au
Agriculture, Fisheries and Forestry (DAFF)
www.daff.gov.au
Communications, Information Technology and the Arts (DCITA)
www.dcita.gov.au
Defence
www.defence.gov.au
Finance and Administration
www.finance.gov.au
Foreign Affairs and Trade (DFAT)
www.dfat.gov.au
Environment and Heritage (DEH)
www.deh.gov.au
Transport and Regional Services (DOTARS)
www.dotars.gov.au
Industry, Tourism and Resources (DITR)
www.industry.gov.au

State Government Data Suppliers

Department of Education and Training
www.det.nsw.edu.au
Department of Energy, Utilities and Sustainability
www.deus.nsw.gov.au
Department of Lands
www.lands.nsw.gov.au
Department of Local Government
www.dlg.nsw.gov.au
Department of Planning
www.planning.nsw.gov.au
Department of Primary Industries
www.dpi.nsw.gov.au
Department of State and Regional Development
www.business.nsw.gov.au

Figure 1 Federal and state government data supplying agencies

It is significant at this point to reiterate the link between the Government and Commercial data sectors. Whilst data is available from the commercial bodies for a fee, Government agencies heavily rely on the Commercial sector for their data, and particularly mapping needs. Consequently, data and maps produced by Government are in some instances widely available for public access, use and distribution. This creates a situation where the link between the Government and Commercial data supplying agencies is becoming less defined (or blurred), and data is increasingly being made available to the extent that purchasing location-specific and CI-related data seems unnecessary, as free data is made available on the websites of Government Departments.

An additional crucial point is that any individual can access commercial data, provided that they have the necessary funds. With respect to an individual intent on causing harm, it is clear that no mechanisms screen the individual, and prohibit them from accessing the datasets. The motivation for suppliers in this category is evidently monetary; however the following questions must be posed: does gaining profit from the distribution of CI-related datasets justify compromising Australia's critical infrastructure, and introducing national security concerns? Additionally, should there be stringent mechanisms in place to manage the individuals accessing CI-related data to ensure that any risks are minimised?

7 Other data supplying agencies

Government and Commercially accessed data can further be supplemented with information from other sources, specifically the owners and operators of critical infrastructure items, such as utility and telecommunications companies. These bodies generally seek to educate the community about their products/services, but are 'unintentionally' providing revealing information about their operations and infrastructure. Throughout this study, the major utility companies and telecommunications providers in Australia were assessed with respect to the amount of critical infrastructure data they offer.

Utility companies refer to organisations involved in providing energy (electricity and gas) and water-related services to the community. The major utility companies evaluated include Sydney Water, Hunter Water, Integral Energy, Energy Australia and AGL. The major telecommunications providers in Australia, Telstra and Optus, were also profiled in terms of their role in public data availability. As with the utility companies, telecommunications organisations provide public information solely for the role of educating the community, and promoting their services and networks.

However, the concerns previously raised, relating to national security and critical infrastructure protection are also applicable to this data supplier category. For instance, some agencies are making available information relating to energy networks in an attempt to be transparent with their customers, but are in turn creating a risk in providing such details, which can potentially comprise the CIP process.

8 The public data availability situation in Australia

The supplier identification and categorisation phase described above was employed primarily to assess the disparate sources of public data in Australia, in order to determine whether critical infrastructure data exists in the public domain. It is evident from the assessment that data can be provided deliberately or unintentionally, and can in both instances be reproduced. Data files, independently, may not appear particularly useful, but when stored with data from various other sources result in a structured repository of data relating to Australia's critical infrastructure, as was produced as part of this study (refer to Figure 2).

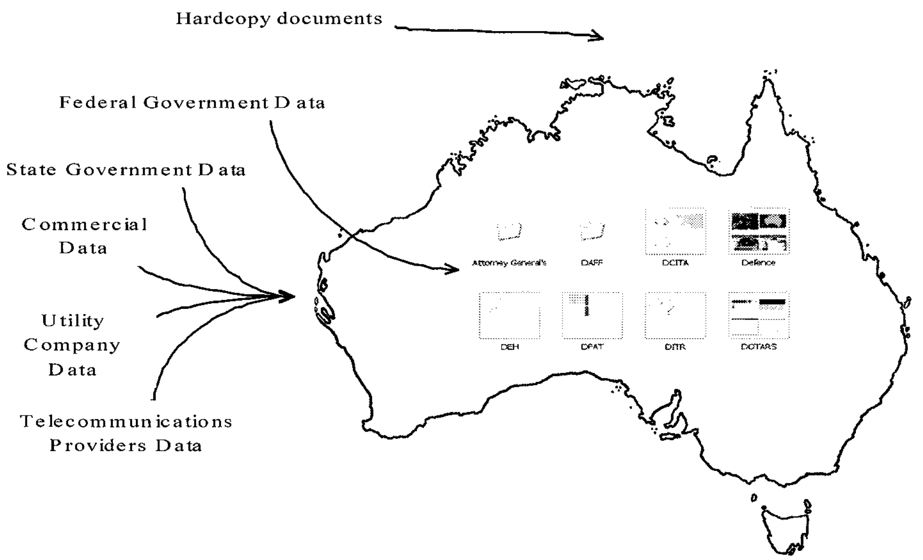


Figure 2 Australian critical infrastructure data repository

A repository of this nature allows for various characteristics of Australia's CI, both individual and collective, to emerge. Such characteristics include previously unconsidered patterns, interdependency information and vital data that may be revealing and compromise the CIP process, and thus affect national security. This structured process of collection can be replicated by any individual requiring little resources, and can be made more exhaustive and detailed with the appropriate funds and dedicated time. Therefore, an individual can engage in various preparations and activities with little inconvenience and detection, using publicly accessible information that was originally made available for the benefit of the public and to encourage openness in initiatives such as e-Government.

The data collected as a component of this study is essentially in its raw form; that is, no attempt has been made to manipulate the data and establish patterns or inferences for the purpose of creating scenarios and understanding the characteristics and linkages between infrastructure items. It is merely an initial repository of Australia's critical infrastructure-related data, collected from publicly available outlets. A selection of the data gathered from the identified bodies, the specific data elements, and the CI type(s) at risk is provided in Table 1.

In addition to the comparison in the table, the use of a geographic information system (GIS) can allow for the data elements to be further combined through the use of longitude and latitude coordinates, associating the data to a particular location. This allows for a graphical, map-based representation of CI-related data to be created, also allowing for detailed analysis of the data. Additional analysis and manipulation of the presented critical infrastructure data can also be conducted, to allow for more patterns and linkages to be revealed.

9 Conclusion

The outcome of this study is the identification of data supplying categories and agencies in Australia, and the creation of a data repository containing information relating to the country's critical infrastructure, by adapting the research conducted by Breeding (2003) in the United States. The examination highlighted the ease with which data could be collected, and the potential for more detailed gathering and analysis. It was apparent that a majority of the data has been provided for *positive* purposes; however, it appears that this level of transparency could result in negative implications for the CIP process. Furthermore, this assessment revealed that it is possible to simply engage in the creation of a comprehensive, sector-based repository of Australia's CI, with little detection or screening mechanisms. The Australian Government, through related Federal and State departments, offers a wealth of free data for public access. Other data supplying agencies are also present, such as commercial agencies, utility companies and telecommunications providers, who offer data both for profit and indirectly.

The data repository created, which was not exhaustive, demonstrated that revealing information can be publicly obtained in multiple formats, and can be stored for any purpose. While for research and emergency management activities this is important, a number of negative implications introduce the question of whether the positive uses of public data are overshadowed in certain situations. The potentially damaging effects of public data availability revolve around simplifying the ability to collect information that can compromise Australia's CI. These effects include terrorism, fraud, identity theft, vandalism, and competitive intelligence.

The ability to manipulate/present data in a form different from its original has demonstrated that 'sensitive but unclassified' data is in abundance in Australia. Data elements independently appear harmless, but techniques ranging in complexity and sophistication, such as geospatial information exploitation allow for inferences to be made and patterns to be extracted concerning Australia's CI, thus posing a threat to the CIP process.

The outcomes of this study were aligned with Breeding's, highlighting the potentially negative consequences enabled, and to an extent facilitated, by the provision of public data relating to Australia's CI. Further research into public data in Australia must be conducted, in an attempt to achieve a balance between open information access and restriction/censorship, to ensure that the threats associated with public data are minimised and eliminated where possible.

Table 1 Summary of public data and the threat to critical infrastructure

Supplier Type/Title	Major Available Data Elements (Attributes)	What CI is at Risk?
Federal/ Attorney General's Department	• Counter terrorism report & plan • Links to Emergency Management Australia & a set of publications/links	• All infrastructure types
Federal/ Agriculture, Fisheries & Forestry	• Food production & trade statistics	• Agriculture & the food supply
Federal/Communication Information Technology & the Arts	• Telephone services • Communications maps	• Communications • Cyber infrastructure
Federal/Defence	• Defence establishments • Video clips	• Government
Federal/Finance and Administration	• e-Government information	• Cyber infrastructure
Federal/ Foreign Affairs & Trade	• Trade statistics • Rail report • Rail network map • Supplier database	• Transport • All infrastructure types
Federal/Environment & Heritage	• Culture objects & places information • Power plant maps • Renewable energy	• Cultural icons • Energy
Federal/Transport & Regional Services	• 2006 transport statistics report • Transport map • Road, airport & rail network maps • Airline statistics • Shipping & ports statistics • Regional data	• Transport
Federal/ Industry, Tourism & Resources	• Energy supply details • Spreadsheet of renewable energy operators • Gas pipelines	• Energy
State/Department of Education & Training	• Searchable public school database • Statistical education information • Motorways in NSW (existing & proposed) • Freight & Rail infrastructure data	• Schools • Transport
State/Department of Energy, Utilities & Sustainability	• Local & metropolitan water utilities list • Sydney Water operations map • Energy Australia electricity network map • Gas network information • List of electricity suppliers	• Electricity • Gas • Water
State/Department of Lands	• Mapping tool for locating things such as roads, properties & national parks • Geographic names register CSV file	• All infrastructure types (particularly transport)
State/Department of Local Government	• Excel file of Council contact details • Detailed Council information by region • Councils map	• Government

Table 1 Summary of public data and the threat to critical infrastructure

Supplier Type/Title	Major Available Data Elements (Attributes)	What CI is at Risk?
State/Department of Planning	• Transport & population statistics, maps & graphs • Travel habits report	• Transport
State/Department of Primary Industries	• Map of mines • Petroleum, coal mines & energy maps & resources	• Energy
State/Department of State & Regional Development	• Economic, trade, infrastructure & business statistics & facts • Regional profile, with transport and major networks information • NSW train network • Ports data	• All infrastructure types (particularly transport)
Commercial/MapData Sciences	• Address locator • Street database, containing transport, towns & points of interest data	• All infrastructure types
Commercial/PSMA, G-NAF	• Points of interest, such as cultural, defence, emergency, medical, post offices, sewage, transport and utilities • Physical addresses datasets • Transport dataset	• All infrastructure types
Commercial/Sensis	• Telephone service • Business & residential information • Points of interest maps	• All infrastructure types
Commercial/MapInfo	• Streets, demographics, postal & administrative boundaries data	• All infrastructure types
Utility Companies/Sydney Water	• Area of operations map • Sewage treatment plants map • Water filtration & treatment plants information	• Water
Utility Companies/Hunter Water	• Supply & performance statistics • Dam fact sheets • Treatment plant diagrams • Area of operations map	• Water • Dams
Utility Companies/Integral Energy	• Network area maps	• Energy
Utility Companies/Energy Australia	• Network map • Proposed upgrades to electricity network map	• Energy • Electricity
Utility Companies/AGL	• Gas distribution network	• Gas
Telecommunications Providers/Telstra	• Network information • Interactive coverage maps	• Communications • Cyber infrastructure
Telecommunications Providers/Optus	• Network coverage maps • Broadband network map • Mobile coverage in NSW	• Communications

References

Australian Government (2006). 'About this Site' [Online]. Available: http://australia.gov.au/about-this-site [Accessed July, 2006].

Breeding, A. J. (2003). Sensitive but Unclassified Information: A Threat to Physical Security, SANS Institute [Online], Available: http://www.sans.org/rr/whitepapers/country/ [Accessed December, 2005].

Davey, S. (2005). 'Exploring e-democracy and Online Service Delivery for Australian Governments', CHISIG, Canberra, Australia, November 23-25.

Emergency Management Australia (2003). 'Mapping the Way Forward for Large-Scale Urban Disaster Management in Australia' [Online], Available: www.ema.gov.au [Accessed February, 2006].

Givens, B. (n.d.). 'Public Records on the Internet: The Privacy Dilemma' [Online], Available: www.privacyrights.org [Accessed March, 2006].

Lee, S. M., Tan, X. and Trimi, S. (2005). 'Current Practices of Leading e-government Countries', Communications of the ACM, 48(10): 99-104.

Luiijf, E. A. M. and Klaver, M. H. A (2004). Protecting a Nation's Critical Infrastructure: The First Steps. IEEE International Conference on Systems, Man and Cybernetics: 1185-1190.

Popp, R., Armour, T., Senator, T. and Nymrych, K. (2004). 'Countering Terrorism Through Information Technology', Communications of the ACM, 47(3): 36-43.

Punia, D. K. and Saxena, K. B. C. (2004). 'Managing Inter-organisational Workflows in eGovernment', Communications of the ACM: 500-505.

Wunnava, S.V and Reddy, M.V. (2000). 'Internet Based Digital Government Model Development', Florida International University College of Engineering, IEEE 2000: 205-208.