

DATA COLLECTION FROM LEGISLATION – AN ONTOLOGICAL APPROACH

Rosetta Romano, Faculty of Government, Business and Law, University of Canberra,
Canberra, Australia, rosetta.romano@canberra.edu.au

Abstract

Legislation connects government service providers (agencies, and their partners) to government service consumers (individuals and organisations). While legislation is a rich source, it is difficult to navigate, complex to understand and is subject to frequent change. This research-in-progress paper offers a novel way using ontologies to model the relationship terms existing in government service delivery legislation.

An ontology provides the agreed definitions and describes how the terms in a subject area, or a domain, are related. If an ontology of government service delivery terms and relationships was available, then, both government service providers and service consumers could use the model to assess the impact of legislative change.

The problem is that intense manual processing is required to detect the terms and the relationships existing in legislation, and there is scant coverage of the process in the literature that can be used to develop an ontology of government relationships. This paper describes the data collection method being taken to detect, extract and analyse the terms and relationships existing in legislation using a case study of an Australian government family tax benefit.

Keywords: Ontology, Legislation, Government service delivery, Online government claim form.

1 INTRODUCTION

Complex government information systems are required to manage complex legislation. The February 2015 McClure Report to the Minister for Social Services, calls for government to simplify the welfare payment system (Department of Social Services, 2015). McClure offers a number of recommendations to simplify the system, many of which will require changes to legislation, and the government information systems that operationalise the legislation. A government information system is the application of people, technologies and procedures to solve government problems (Lynch et al. 2013).

In Australia, there is no single instrument being maintained that provides government service providers and service consumers with a model of the terms in legislation relating to government service delivery, or of the terms used in information systems delivering government services. In 2012 there was a plea for the development of instruments to assess the organisational and technology impacts, and the implementation risks of policies on organisations prior to implementation (Gong & Janssen 2012, p. 570). As multiple agencies and partners work together to deliver government services, the need for a single instrument is becoming more urgent.

In the absence of such an instrument, it is difficult to understand how a whole of government response to the impact of legislative change on government information systems can be managed and communicated. At a minimum, the recommendations of the McClure report will require a whole of government impact assessment of the legislative changes to the information systems in the policy and service delivery departments where the functions are managed separately. While individually, each department's assessment may be valid, there may be little consistency in a whole of government impact assessment. An incomplete or inefficient process at the departmental level would be carried through to a whole of government impact assessment.

The research being conducted is a novel use of ontologies contributing to Australia's move to e-government. E-government is a way to use new technologies to provide people with more convenient access to government information and services, to improve the quality of services and to provide greater opportunities to participate in the democratic institutions and processes (Li & Dai, 2011). An ontology is an artefact that provides a community with the agreed definitions and describes how the terms in a subject area or domain, are related. Besides being useful as an agreed dictionary, its strength lies in the way that technology is able to consume it. As a model, ontology can be read by humans and coded for computers. Once coded, or formalised, ontologies allow sophisticated machine manipulation, aggregation of information, pattern analysis and inferences from vast quantities of data that humans would not be able to handle (Lambe 2007).

New technologies can be used to consume service delivery ontologies of complex legislation describing the providers, consumers, services, and the relationships between them. If an ontology describing the legislative service relationships with government was available, then, an evidence-based method to identify the stakeholders for a whole of government legislative impact assessment would exist. A query would return information about the policy, legislation, responsible minister, department, information system (i.e., people, technologies, procedures) impacted by the change.

While there is no one correct way to model a domain as ontology, there is agreement that the important terms in a domain need to be identified (Noy & McGuinness 2001), (Buitelaar et al. 2005). Like ontologies, legislation provides definitions of terms in a domain and describes the relations between the terms. Legislation is a primary source for government agencies to harvest terms to build ontologies.

When legislation fails to provide definitions of key terms, then, enumerating the domain becomes complicated. For example, key terms such as adult, carer, child, claimant or parent are not defined in the legislation related to the payment the family tax benefit (FTB). FTB supports the government's family policy, and is an income-tested payment for eligible parents and carers (Australian government,

Australian Law Reform Commission, 2014). Besides the Act for general interpretation of all Commonwealth Acts the *Acts Interpretation Act, 1991*, there are two pieces of legislation specifically related to the FTB payment. The legislation providing the eligibility rules is, *A New Tax System (Family Assistance) Act 1999*; and the legislation providing the payment rule is, *A New Tax System (Family Assistance) (Administration) Act 1999*.

While there is coverage of legal ontology term extraction (Despres & Szulman 2004), (van Engers, et al. 2008), (Hassan & Logrippo 2009), the literature does not provide a process to detect relationships between the terms. The relationships are embedded in complex legislation and, they are key to understanding how government providers and consumers are connected by the legislation. Detecting these relationships requires human effort, and this should not be under-estimated (van Heel & van Engers 2003). The representation of regulation requires total human interpretation and subject matter expertise to make them (1) accessible electronically, (2) structured and understandable by machines, and (3) represented in a standard format and interoperable for human interpretation (Hassan & Logrippo 2009).

This research-in-progress describes a process to detect, extract and analyse the relationships that exist in legislation. The context for the related work in ontologies extracted from policy or legal texts related to eGovernment services is depicted in Figure 1. In two government service meta-ontologies for e-government services, Vassilakis & Lepouras (2006), and Charalabidis & Metaxiotis (2009) capture service consumer and service provider information (see 3 in Figure 1), but neither provide an explanation of how the relationships are captured. A meta-ontology contains entities needed to describe and discover the service in an ontology (Stojanovic et al. 2004). The research being reported provides a deeper understanding of the meta ontology concepts: service consumer and services provider. This depth emerges from the ontological assessment made possible by the process proposed to detect, extract and analyse the relationships that exist in legislation.



Figure 1 eGovernment service ontology research context.

This paper describes the data collection process for proposed research. The remainder of the paper is organised as follows: in Section 2, the research setting and case study is described. In Section 3, the steps to detect the data are explained. Section 4 describes the expected outcomes of the research, and finally, Section 5, outlines some early observations and challenges that are emerging from the preliminary attempts to detect the relationships between the terms.

2 RESEARCH SETTING AND CASE STUDY

The research setting is the Australian Public Service, and the scope is the information systems bound by the legislation related to the FTB payment in normal circumstances. The first column in Figure 2 depicts the three arms of law making in Australia: the parliament that makes the law; the executive that operationalises the law; and the judiciary that interprets the law. The second column indicates

the types of APS departments, and information systems involved in Australian law making. The third column describes the particular artefacts used, or not used, in the case study.

The three arms of law making in Australia	GOVERNMENT INFORMATION SYSTEMS	CASE STUDY INFORMATION SYSTEM ARTEFACTS
<p>SCOPE</p> <p>The Parliament makes the law</p>	<p>Legislative department information systems used to draft, introduce and manage legislation.</p>	<p>The FTB legislation</p> <p><i>Acts Interpretation Act 1991</i> : Dictionary of terms for Commonwealth Acts <i>A New Tax System (Family Assistance) Act 1999</i> : Eligibility for FTB <i>A New Tax Systems (Family Assistance)(Administration) Act</i> : Payment for FTB</p>
<p>The Executive operationalises the law</p>	<p>Policy department information systems used to communicate and measure the effectiveness of the policy.</p> <p>Service delivery information systems used to administer the service.</p>	<p>The policy – Helping families with the cost of raising children</p> <p>Department of Social Services - Family Assistance Policy</p> <p>The FTB claim form</p> <p>Department of Human Services</p>
<p>The Judiciary interprets the law</p>	<p>Courts' information systems</p> <p>Administrative Appeals Tribunal systems</p>	<p>Case law</p> <p>Decisions</p>

Is there conceptual alignment ?

SCOPE

Figure 2 Research setting and scope.

The FTB payment is used as a case study in the research-in-progress to demonstrate how the legislation connects different service providers to the government information systems residing in those departments. In Australia, the prime minister of the day assigns the responsibility for legislation to Ministers in a legal instrument known as the Administrative Arrangement Orders. All legislation is drafted by the Office of the Parliamentary Counsel. The implementation or operationalisation of the legislation requires a policy focus and a service delivery focus. These may be offered by the same agency, but this is not the case for the FTB payment. For the FTB, the Department of Social Services communicates the policy and measures its effectiveness, and the service delivery for the payment is undertaken by the Department of Human Services (DHS). DHS is responsible for developing and maintaining the government information systems that assess eligibility and, payment of the FTB. Claimants seeking an FTB payment must complete a DHS online claim form to be assessed for the payment in accordance with the legislation.

The legislation also creates many relationships to other ministers through cross-references to other legislation. These Acts connect the ministers who have responsibility for the legislation, and their departments. In the research to date, cross-references to another 13 pieces of legislation have been identified including the *Migration Act, 1958* and the *Health Insurance Act, 1973*. The legislatively connected departments for these two acts are: the Department of Immigration and Border Protection and, the Department of Health respectively. The information systems in the cross-referenced or legislatively-connected departments will not be reviewed, although the relationships with the Departments will be included in the developed ontologies.

The research being reported in this work-in-progress applies an instrumental case study (Yin, 1994) to gain a broader and deeper appreciation of the process to develop an ontology of legislation. It is expected that the legislation and the claim form applying to the payment of the FTB payment in ‘normal circumstances’ will be sufficient to demonstrate the process used to model the legislation and claim form artefacts as ontologies.

3 CASE STUDY DATA COLLECTION

While the steps to develop an ontology have been described in the literature, actually deriving the terms from the legislation is a process that is not covered in sufficient detail to construct one from

scratch. The researcher commenced the data collection by considering traditional modelling and constructed an entity relationship model (ER-model) (See Figure 3A for an extract). An ER-model provides a limited representation of entities in a domain and their relationships to each other.

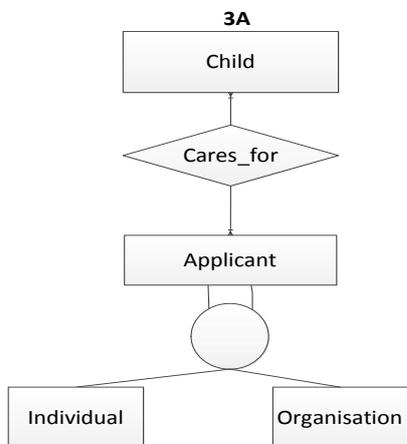


Figure 3A. An ER-Model extract.

3B

Business Requirements Analysis						
Qualification Id	Qualification			Boolean Operator	Legal Document Title	Authority
	(Subject	Operation			
Normal circumstances					A New Tax System (Family Assistance) Act 1999 - C2013C00649	Part 3, Division 1, Subdivision A Section 21
	((the individual	has at least one FTB child	OR	A New Tax System (Family Assistance) Act 1999 - C2013C00649	Part 3, Division 1, Subdivision A Section 21(1)(a)(i)
	(the individual	is not an absent overseas recipient	AND	A New Tax System (Family Assistance) Act 1999 - C2013C00649	Part 3, Division 1, Subdivision A Section (1)(a)(ii)
		the individual	has at least one regular care child	AND	A New Tax System (Family Assistance) Act 1999 - C2013C00649	Part 3, Division 1, Subdivision A Section (1)(a)(iii)
		child	the regular care child is also a rent assistance child)	A New Tax System (Family Assistance) Act 1999 - C2013C00649	Part 3, Division 1, Subdivision A Section (1)(a)(iii)

Figure 3B. Rules Modelling Diagram.

Another model is required to express the complex legal rules. To do this, a template capturing the rules of an FTB child in normal circumstances has been developed (see Figure 3B). The complexity of the legislative definitions describing relationships is exacerbated by the nesting within other definitions. Often these definitions are nothing more than a series of rules. The complexity of this process is overwhelming and modelling does little more than confirm that definitions are being reused from other parts of the same legislation, and at other times from other legislation.

Both the ER-model and the modelling of the rules provide useful domain context. However, neither the approach, nor the artefacts, could be easily shared with non-technical and non-business experts such as service consumers. As this research is attempting to support eGovernment, then, being able to share the models with both government service providers and service consumers is important. The development of an artefact in the form of an ontology that models the complexity of the legislation may assist others to simplify it, because it can be shared by all parties and be used to engage them in democratic processes such as impact assessment following a legislative change.

Each of the two pieces of legislation, and the claim form provide different perspectives of the same domain. The legislation provides the eligibility and the payment perspective, and it is assumed therefore, that the claim form represents the set, of the terms in both pieces of legislation. The claim form is used as a window to the government information systems that process the information provided by the applicants. All of the data items provided by the applicant should be consumed by the government information systems to determine the eligibility and the payment. No more, and no less. The FTB claim form currently asks the applicant to answer questions that generates 946 data items about themselves; 145 about their partner; and 52 items about each of their children. The research is seeking to confirm that information systems and legislation are aligned by comparing the underlying conceptual structure of the FTB legislation and government claim form. An important requirement in the data collection is to ensure that the legislation is being reflected in its application and administration to meet the stated policy objectives.

The researcher has conducted a manual exercise to identify the relationships in both the legislation and the claim forms related to FTB in normal circumstances. A manual scan of the nouns was useful to determine the entities nominated in the legislation. These entities are represented as: individuals, organisations, or positions. Some examples of individuals include: adult, child, citizen, parent, and partner. Some examples of organisations include: approved care facilities, bank, college, government departments; and some examples of position include the: Australian Statistician, Minister, Secretary, and the Child Support Registrar. This exercise provides some early analysis opportunities. For

example, across the domain made up of the eligibility legislation, the payment legislation and the claim form, there are 20 different descriptions of the term ‘child’ (see Figure 4).

The iterative development of legislation by different legal drafters means that sometimes, inconsistency emerges. Figure 4 shows an extract of a model developed to capture the instances detected manually, where the term [child] is used in the legislation and the claim form. In the first column, those detected in the eligibility legislation are captured; in the second column those in the payment legislation; and in the third column those in the claim form are captured. While there is an assumption that, the claim form should use all the terms from the legislation, the extract shows that there are only three instances, marked with an (*), to denote instances where the claim form uses the term used in the legislation. Using this type of analysis to identify similarities may be appropriate however there is always a danger in attempting to relate legal concepts in one system to a seemingly similar situation in another system (Ault & Arnold 2004) without an ontological assessment. For example there is no school child in the claim form, but there is a child in education or training. The outcome of the ontological assessment will be an explicit determination in the model that the terms are, or are not, the same (synonyms) within the domain. This type of analysis may result in suggestions for amendments to the legislation to remove ambiguity.

Future extension of the process could potentially include other entitlements for families with children, and the ontological assessment and explicit determination may even extend the ontology across other legislation in the government service delivery domain. This type of analysis could be used to identify gaps between the law and the administrative processes. There are two issues for operationalisation of the law. The first is that the claim form has not used the term provided in the legislation, and the second is that the claim form introduces new terms. The significance of these two issues will be assessed as part of the research. Clarification of these gaps could lead to a closer legislative alignment.

Ref #	(1) Eligibility legislation	(2) Payment legislation	(3) Claim form
1	adopted [child]	adopted [child]	
2	dependent [child]		dependent [child]*
3	[child] from		[child] from*
4			[child] in full-time education or training
5	FTB [child]	FTB [child]	FTB [child]*
6			[child] having a
7	immediate [child]		
8	[child] in the adult's care		
9			[child] living independently
10	natural [child]		
11	[child] of	[child] of	
12			orphan
13	other [child]		
14	qualifying [child]		
15	regular care [child]	regular care [child]	
16	relationship [child]		
17	rent assistance [child]	rent assistance [child]	
18	school [child]	school [child]	
19	[child] who has died		
20	[child] without the adult's consent		

Figure 4. The different use of the term ‘child’ in the legislation and the government claim form.

By using some of these processes it will be possible to develop a set of relationships existing in service delivery legislation and identify possible candidates for inclusion in the *Acts Interpretation Act, 1901*, the government’s dictionary. By examining one domain at a time, using a bottom-up approach, it should be possible to develop a rich model of the government service delivery domain using ontologies. Each relationship will be modelled using the Resource Description Framework (RDF). RDF provides a mechanism for allowing anyone to make a basic statement about anything and layering these statements into a single model (Allemang & Hendler 2011, Page 28). Using RDF, all relationships will be expressed in the form of a triple i.e., subject/predicate/object. For example in Figure 3A the Applicant cares_for Child would be represented. All triples will then be captured as an

ontology using the Web Ontology Language, OWL. OWL enables the processing and interpretation of information so that it is understandable to software (Lacy 2005).

In the case study, only the terms and relationships in the FTB in normal circumstances are modelled. A broader view of the entire legislation would enrich the representation of the term 'child' as it is being used in government service delivery. This is expected because the legislation supporting the eligibility and payment for FTB also provides another three family payments being the: maternity immunisation allowance; Child care benefit; and Schoolkids bonus. The relationship between the claimant and the child is a fundamental prerequisite to each of these payments. The ontological analysis will determine whether terms are synonyms, or, legitimate extensions.

The relationships also provide richness for the development of broader ontologies. For example the term 'child' is used in 20 different ways within the FTB government service domain. If government was developing family service ontology, then, other services would be included and the 'child' concept could be extended. If the Child Support payment was the next service to be included, then the ontology would be extended with the concept 'child eligible for administrative assessment' from the *Australian Child Support (Assessment) Act, 1989*.

If an ontology describing the relationship of child to government was available, service providers, services consumers, and government could benefit. Service providers could determine which government information systems impact on different service consumers who may be responsible for a child. While, a service consumer with a child could, use the ontology to, self-determine which, of the many different pieces of legislation, in different service delivery domains, applies to them. Government would have a process to identify stakeholders that exist in legislatively-connected departments, and partners who exist outside of the policy/service delivery arrangement, so a more complete impact assessment process could be facilitated.

4 EXPECTED OUTCOMES

By conducting an ontological investigation of the relationships expressed in the legislation, it is expected that two views of the underlying structure of the FTB domain relating to normal circumstances will emerge. The first view will represent the eligibility legislation, and the second will represent the payment legislation. These structures will be modelled as ontologies. A similar investigation will be undertaken to detect the terms existing in the claim form. The three ontologies will provide the data to enable an assessment of the alignment of the government information systems to the legislation. The unique terms across all three ontologies will form the FTB service delivery in normal circumstances *domain ontology*. Repeating the process to capture the relationship in more, if not the entire, two pieces of legislation will provide a deeper understanding of the relationships existing in the family assistance domain. The processes to detect the underlying structure of the domain, and the modelling as ontology are expected to be generalizable across legislation being administered by service delivery departments in government. Potentially, the government relationship modelling could provide a process to develop a whole of government view of service delivery domains as prescribed in the legislation. With this view, it should be possible to determine the extent to which the terms of legislation are being reflected in the service applications and administration to, meet the stated policy objectives.

5 EARLY OBSERVATIONS AND CHALLENGES

The data collection for this research requires an investigation of the legislation, to detect the underlying conceptual structure of the FTB in normal circumstances domain. A bottom-up approach demonstrated by using the term 'child' in this paper develops an understanding of a relationship of 'child' in the government service delivery FTB domain. Furthermore, it is developing an understanding of the Family services domain within the broader payments offered by the Australian government to service consumers. While legislation does provide many of the relationships between

the service providers and the service consumers, it does not provide all of them. In a social welfare setting, any misalignment may result in a disadvantage to the very service consumer that the payment is intended to support. This paper describes some very early research exploring how the terms and relationships existing in a government claim form can be detected and assessed for alignment. In the future the researcher will, undertake work to investigate text mining as an approach to conduct automatic detection and extraction of terms from the legislation.

6 CONCLUSION

This research describes the data collection of terms and relationships existing in service delivery legislation relating to an Australian government family tax benefit payment. This paper reports on research that is exploring the use of ontologies to manage frequent and complex legislative change, and to understand the whole of government impact of change on service providers and service consumers. This paper has outlined some preliminary attempts to model the legislation, and has pointed out some reasons why ontologies would be a better model than other traditional methods for government service providers and consumers. The development of government service delivery ontology will support the adoption of e-government in Australia by modelling the terms across legislation that is shared by separate agencies. The research-in-progress reported in this paper is expected to demonstrate how ontologies of the complex service delivery legislation can be used by government to engage service providers and service consumers in a discussion about simplifying the legislation. It should also support whole of government impact assessments of proposed or enacted legislative change. The novel use of ontologies as described in this paper should be considered as another way to advance e-Government for service delivery in Australia.

REFERENCES

- Allemang, D. and Hendler, J. (2011). *Semantic web for the working ontologist: effective modelling in RDFS and OWL*. Elsevier.
- Ault, H. J. & Arnold, B. J. (2004). *Comparative Income Taxation - A Structural Analysis*. Kluwer Law International.
- Apostolou, D. Stojanovic, L. Lobo, T.P. Miro, J.C. and Papadakis, A. (2005). Configuring e-government services using ontologies. In *Challenges of Expanding Internet: E-Commerce, E-Business, and E-Government*, 141-155. Springer US.
- Apostolou, D. Stojanovic, L. Lobo, T.P. and Thoenssen, B. (2005). Towards a semantically-driven software engineering environment for e-government. In *E-Government: Towards Electronic Democracy* (pp. 157-168). Springer Berlin Heidelberg.
- Australian Government, Australian Law Reform Commission. (2014, 07 22). Australia's social security system. Retrieved 03 17, 2015, from http://www.alrc.gov.au/publications/5-social-security/australia%E2%80%99s-social-security-system#_ftn5
- Buitelaar, P. Cimiano, P. and Magnini, B. (2005). *Ontology learning from text: An overview*. Vol. 123.
- Charalabidis, Y. and Metaxiotis, K. (2009). *Ontology-Based Management of e-Government Knowledge*. Semantic Interoperability Centre Europe, IGI Global.
- Department of Social Services. (2015). *A New System for Better Employment and Social Outcomes*. Canberra: Commonwealth of Australia.
- Despres, S. and Szulman, S. (2004). Construction of a Legal Ontology from a European Community Legislative Text. *Legal Knowledge and Information System Juris - The Seventeenth Annual Conference*. Amsterdam: IOS Press.
- Feltus, C. Khadraoui, A. Yurchyshyna, A. Leonard, M. and Dubois, E. (2012). Responsibility Aspects in Service engineering for e-Government. In *Interoperability for Enterprise Systems and Applications conference (I-ESA'12) Workshop Service Science and the Next Wave in Enterprise Interoperability*, Valencia, Spain.
- Gomez-Perez, A. Ortiz-Rodriguez, F. and Villazon-Terrazas, B. (2006). Legal ontologies for the Spanish e-government. In *Current Topics in Artificial Intelligence*, 301-310. Springer Berlin Heidelberg.
- Gong, Y. and Janssen, M. (2012). From policy implementation to business process management: Principles for creating flexibility and agility. *Government Information Quarterly*, 29, S61-S71.
- Gong, Y. and Janssen, M. (2013). Adaptive and Compliant Policy Implementation: Creating Administrative Processes Using Semantic Web Services and Business Rules. In *Collaborative, Trusted and Privacy-Aware e/m-Services* (pp. 298-310). Springer Heidelberg.
- Hassan, W., and Logrippo, L. (2009). Governance requirements extraction model for legal compliance validation. *Second International Workshop on Requirements Engineering and Law (RELAW) (7-12)*. IEEE.
- Khadraoui, A. Opprecht, W. Leonard, M., and Aidonidis, C. (2011). Law-Based Ontology for E-Government Services Construction – Case Study: The Specification of Services I Relationship with the Venture Creation in Switzerland. In *Practical Studies in E-government* (149-166). Springer New York.
- Lacy, L.W. (2005). *OWL: Representing Information Using the Web Ontology Language*. Trafford Publishing (UK) Ltd.
- Lambe, P. (2011). *Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness*. Chandos Publishing (Oxford) Limited, Oxford.
- Li, L. and Dai, H. (2011). Building a Change Management Model for e-Government Services Evolution. *Fifth International Conference on Management of e-Commerce and e-Government (ICMeCG)*. IEEE.
- Lynch, K. Baltzan, P. and Blakey, P. (2013). *Business Driven Information Systems* (2nd ed.). Australia: McGraw-Hill Education- Europe.

- Noy, N. F. and McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*. Stanford Medical Informatics.
- Stojanovic, L, Abecker, A., Stojanovic, N. and Studer, R. (2004). In Proceedings of the 3rd International Conference on Ontologies, Databases and Application of Semantics (ODBASE 2004), number 3291 in Lecture Notes in Computer Science, pages 1080-1097, Agia Napa, Cyprus, November 2004. Springer Verlag.
- van Engers, T. Winkels, R. and Kordelaar, P. (2008). *The Know and the Flow*. Proceedings of the 5th International Symposium on the Management of Industrial and Corporate Knowledge. Columbus, Ohio.
- van Heel, E. and van Engers, T. M. (2003). *Goldmine Power: Decisions in Governmental Process Design*. The Sixteenth Annual Conference Legal Knowledge and Information Systems Jurix 2003. Amsterdam.
- Vassilakis, C. and Lepouras, G. (2006). *An Ontology for e-Government Public Services*. Encyclopedia of E-Commerce, E-Government and Mobile Commerce.
- Yin, R. K. (1994). *Case Study Research (Second edition)*. California: SAGE Publications Inc.