# Process Ontologies Facilitating Interoperability in eGovernment A Methodological Framework

Timo Herborn and Maria A. Wimmer

Institute for IS Research, Research Group eGovernment
Universitaetsstr. 1, 56070 Koblenz, Germany
{therborn|wimmer}@uni-koblenz.de

**Abstract.** eGovernment is characterized by the usage of multiple applications and heterogeneous data environments. Although the outcomes of administrative processes are defined quite good, the processes themselves are different depending on geographical, political or systemic factors. A vivid example for such a diversity are the business registers (BR) of the member states of the European Union (EU). They all have in common that they are gathering and providing information about companies and related data. But not only their data processed is totally different - also the underlying processes are. The project BRITE (Business Register Interoperability Throughout Europe) aims to build interoperability between the BRs in order to facilitate EU-wide transactional services for companies to e.g. register their branch in another country. These goals will be achieved by the application of ontology-driven semantics. This paper introduces the approach to develop a common BRITE domain ontology which links up national domain ontologies and BR processes. In this way, interoperability is reached among various national business registers and their respective services. The BRITE platform thereby serves as an intermediary to link up the diverging ontologies of national business registers.

**Keywords.** Interoperability, eGovernment Domain Ontology, High Level Domain Ontology (HLDO), Process Ontology, Business Registers, BRITE

## 1 Motivation

Governments are more and more using ICT to facilitate their tasks and responsibilities as well as to collaborate among public organizations more efficiently. In order to guarantee the free movement of European citizens and companies, the European Community perceives the imperative necessity to establish an appropriate and interoperable ICT basis [3]. To pave the way for interoperability by large, the EC co-funds a series of research and technology developments under IST[1], MODINIS [1] or IDABC [4].

[1] Examples of such projects from the 6th Framework Program IST of the EC are: GUIDE, INTELCITIES, ONTOGOV, QUALEG, Semantic-

In its visionary documents, the European Community (EC) attempts to become the world leading knowledge society [2] and [5]. Not only does this mean to enlarge the Community with new member states. Likewise importantly envisaged is a liberal market with companies being able to establish their branches in other EC Member State countries as quickly and simply as possible. A major aspect within this topic, which is, by the way one of the most important rights written down in the Roman Treaty, is the interoperability of registration systems. Therefore, the EU undertakes extensive homogeneisation efforts on political, economical and organizational levels. These efforts result in multiple bills and directives[2] which are demanding cooperating efforts of the Member State's governments.
This paper gives a short introduction to the scope of the EC co-funded project BRITE. Then, a methodological framework for developing a domain ontology for European Business Registers is presented. Thereby, different aspects of ontologies which are considered as key success factor in guaranteeing interoperation among Business Registers throughout Europe are discussed.

## 2 Context of Business Registers (BR) in Europe

In Europe, in order to run a business or company, the organization has to be registered at a business register office in the state, where the company or its seat is officially established. Legal forms of companies strongly depend on particular legislation of the country. The number of registration offices varies from country to country (depending on the state's constitutional distribution of responsibilities) as well as on the organization of the registration offices themselves. In some countries, the responsibilities for the company registration process are delegated to federal state or even district administration level[3]. With the vision of the EC becoming a world leading knowledge society with liberal market opportunities and simple registration procedures, the business registration processes need to be kept very simple, effective and without barriers. Consequently, the BRs need to smoothly interact. However in Europe, Business Register interoperability is currently not achieved. Especially in countries with BRs located at regional level, massive problems of interoperability may occur when interacting with BRs of other countries. In some cases, even within a country, interoperability among regional BRs is not reached yet. The barriers Business Registers are facing may start with language barriers, passing on to data and system incompatibility and ending in inconsistent registration processes as well as distinct strategic policies. Providing interoperability among BRs for the respective registration processes of a business are the topic of investigation in the Integrated Project BRITE,

Gov, SMARTGOV, TERREGOV, R4eGov, Access-eGov, etc. (see *http* : *//europa.eu.int/comm/research/fp6/index_en.htm*)

[2] These directives are to be seen as framework and are translated into national law by the member states

[3] In Germany, business registers are organized on "Länder" (federal state) level, not on national level

a project co-funded by the EC within the 6th framework program of IST[4]. Based on the legal groundings such as the 11th Directive of the EC, BRITE aims to build an agile and dynamic governmental environment for the emerging European market place. The approach pursued in BRITE will heavily build on semantic technology, especially ontologies, to facilitate interoperability and process-oriented information exchange [9]. The next section therefore introduces the methodological framework to develop a BR ontology. By developing this methodology we emphasizes the importance of not being too general and focusing on the migration tasks [8].

## 3 Methodological Frame for BR Ontology Development

In the context of guaranteeing interoperability among systems, services and organizations, ontology development has become an important issue. We understand ontologies as a key enabling concept for the Semantic Web. Ontologies 'interweave human understanding of symbols with their machine processability [and] promise a shared and common understanding of a domain that can be communicated between people and application systems' [6]. The general term *ontology* is further detailed in various literature. Guarino for example defines domain ontology as [7]:

- constituted by a specific vocabulary used to describe a certain reality
- a set of explicit assumptions regarding the intended meaning of the vocabulary.

According to that, a domain ontology is extended by some specifications:

- ontologies describe a formal specification of a specific domain
- share understanding of a domain of interest
- represent a formal and machine executable model of a domain of interest.

This article focuses on the use and development of ontologies in respect to the intended interoperability of eGovernment applications and public administration systems especially in the context of European Business Registers (EBR). In this context, already implemented domain ontologies based on different geographical, organizational and historical roots have to be faced and harmonized resp. linked up with each other. Instead of "reinventing the wheel", the re-usability of existing ontologies has to be checked and aggregation of existing data, document and process schemata should be aspired on an overarching level. Furthermore, an ontology has to fulfill criteria such as openness, dynamics and flexibility in order to allow for future changes and integration of laws to come. The aim of BRITE is to combine Domain Ontologies and Process Ontologies in a way to achieve maximum productivity. See fig. (**??** This combination is necessary in order to a) harmonize the vocabulary toward a common upper level conceptual standard, b) to get an understanding of the individual, national processes and

[4] see $http://europa.eu.int/comm/research/fp6/index\_en.htm$

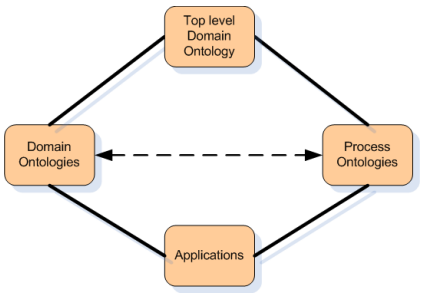


**Fig. 1.** Combination of Domain Ontologies and Process Ontologies

c) to integrate the corresponding processes correctly. In order to achieve the goals mentioned above, we introduce a methodological framework to secure real interoperability between different institutions, even with language barriers and massive process diversities. The approach has the advantage of not having to change legacy systems, but to link them up via standardized overall, domain and process ontologies. It consists of the following steps:

- Defining a high level domain ontology (HLDO)
  - Identifying Domain ontologies of interest
  - Selecting subset of vocabulary of Domain Ontology
  - Merging subsets
  - Mapping merged subsets to HLDO
  - Adding semantics to the HLDO
  - Provide interoperability layer through the usage of the HLDO
- Identifying national BR process ontologies (BRPO)
- Developing high level BR process ontologies
- Integrating national processes

## 4 Defining the High Level Domain Ontology (HLDO)

This paper is based on the idea that a general eGovernment ontology is considered as top level Domain Ontology (figure (reffig:hierarchy)). Technically, the HLDO is an aggregation of at least two Domain Ontologies based on similar semantic backgrounds. This definition bears on the idea of generating semantic interoperability between different Domains without destroying historical data architectures. Legacy systems and data structures can work without any restrictions or losses, because data structures are not changed but enlarged. Additionally, we see it as basis for the development of a standardized process ontology within BR environment. The methodology how to define an HLDO is described in the following figure (3. All steps mentioned in this model are to be done iteratively and are to be repeated until all participating project members feel comfortable with the solution.

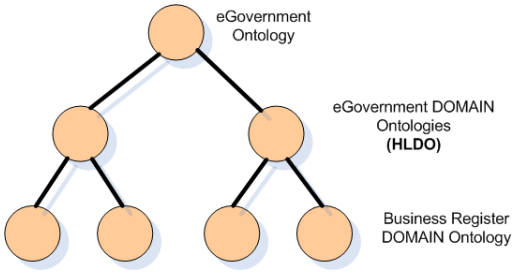


**Fig. 2.** Hierarchy of Ontologies within eGovernment

### 4.1 Identifying Domain Ontologies of interest

Domain Ontologies are - in most cases - not explicitly described in the Domains of interest. Mostly they hidden within legacy systems, databases and in people's minds. The last place mentioned is the most interesting but also most difficult to exploit. The methodologies to identify these various ontologies are as different as the ontologies themselves. Concerning the legacy systems and databases, field and data descriptions can be extracted and then imported into an ontology modeling tool. Getting ontologies out of people's head is much more difficult. Brainstorming, Soft System Methodologies, and Mind Maps are examples of how to retrieve knowledge residing in human brains.

### 4.2 Selecting subset of vocabulary of Domain of Interest

This task is easy to be done. Basically, all BR Domains are of interest for the development of the HLDO. Starting with these project partners that already have explicitly defined their very own ontologies seems to be a reasonable approach.

### 4.3 Merging subsets of national vocabulary

The idea of merging is to finally build an ontology combining all relevant subsets of vocabularies known in the Domain of BR within the project. Merging subsets includes deleting redundant entries. Historical sources are to be tracked by the help of semantics in order to re-assign them to their source of origin.

### 4.4 Mapping merged subsets to HLDO

After having defined the HLDO by merging different subsets, a mapping is to be done in order to validate the new version of the HLDO against the formerly existing Domain Ontologies. Basically, vocabularies and semantics from different domains are to be harmonized. The mapping process carrying out the development of the HLDO is described in figure 5.A simple example shall demonstrate this: "$< Firmenname >$" (domain ontology A) as well as "$< name\_of\_company >$" (domain ontology B) are mapped to "$< BRITE\_company\_name >$". In this

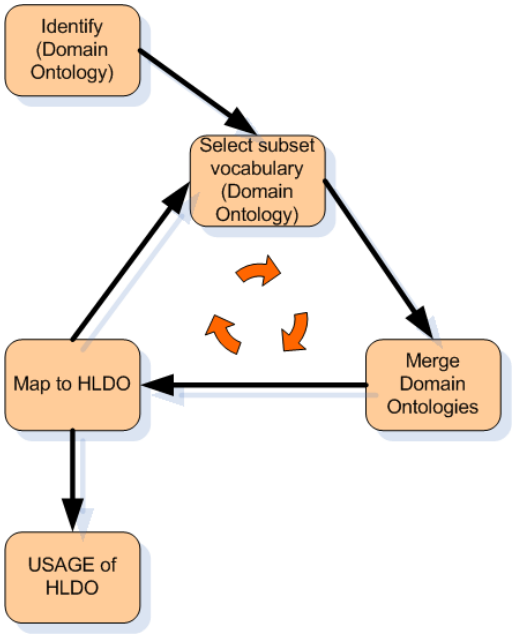


**Fig. 3.** Methodology to define a HLDO

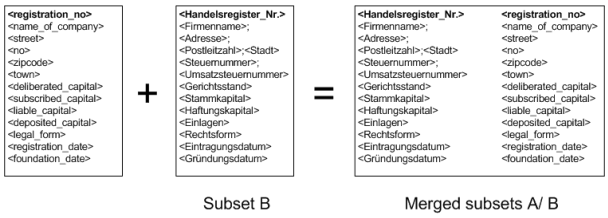


**Fig. 4.** Merging subsets of national ontologies

way, the new concept just needs to store the original identifier (in order to reverse the process) and its origin domain (for allocation reasons).

On the left side of the figure, a subset of a domain ontology is presented. One can easily recognize some major differences: a) linguistic (English and German), b) structural (database management) and c) organizational (mainly content based). According to the definition of the HLDO provided above, an ontology is more than just translating or mapping data and information. Making this information valuable for the processes to be executed, semantics need to be added. First of all, information of the mapping process needs to be stored.

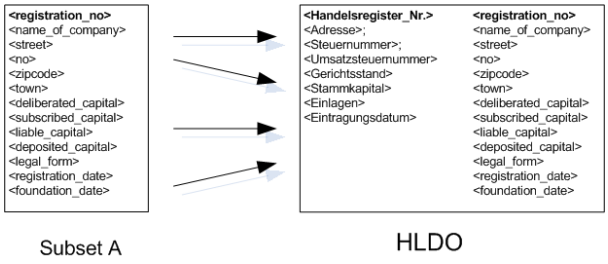


**Fig. 5.** Mapping national domain ontologies to HLDO

### 4.5 Adding Semantics to the HLDO

Next step is to add semantics to the HLDO. The reasons for adding semantics are several. First of all, these semantics will support in building inferences between independent data sets. Apart from that, advanced search mechanisms can be applied to get qualified retrieval results. Finally, semantics allow to introduce automatic machine processing of applications. A concept of Wimmer [10] is used to introduce the application of ontologies within eGovernment contexts. The concept presented in figure (6) consists of classes (topic) and their concrete instances (individuum). Adapting this concept to the vocabulary introduced

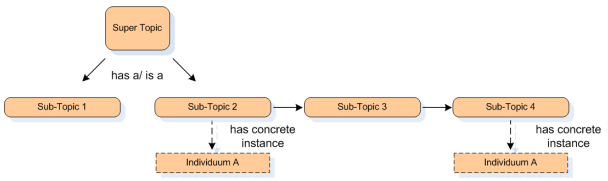


**Fig. 6.** Concept for the abstract ontology

before, a basic knowledge map can be build (7), in which simple inferences are possible. As one can easily see in the figure, the class "Supertopic" is -according

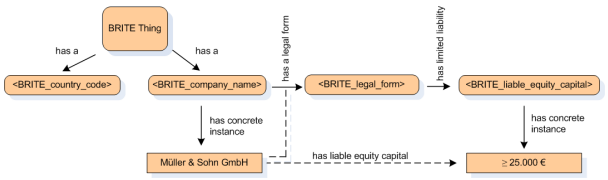


**Fig. 7.** Adding semantics to facilitate automatic machine processing and inferences

to the syntax of OWL- translated into a "BRITE Thing". "Subtopics" are named

"Company Name","Registration Number", etc. Important for the idea of an ontology concept is that we can now build inferences like: Instance A ("Müller und Sohn GmbH") *must* have Instance B ("a liable equity capital of at least 25.000 EUR"). This example does not only show the inference mechanism but also demonstrates the possibilities of using incremental logical conclusions (*here:* Country code AND Liable equity capital).

### 4.6 Provide interoperability layer through the usage of the HLDO

After having developed a HLDO, the domain ontologies need to be revisited because in some cases, adaptations may be needed to exploit the full potential of the common overall ontology. There are two ways to approach the interoperation of domain specific ontologies: In some cases, even legacy systems might have to be adapted or reengineered based on the new vocabulary in order to smoothly interoperate. However, this approach would take enormous amounts of person-power and time, apart from any technical and organizational constraints probably occurring. On top of that, actualizations of the HLDO could not be done synchronously. The second solution proposed is to make the common ontology available in a central storage and make it functioning as intermediary medium. This approach would allow the individuality of each domain ontology to be maintained. Next step is to integrate processes within a distributed eGovernmental environment.

## 5 Identifying national BR process ontologies (BRPO)

Identifying corresponding process ontologies is a key success factor for BRITE. In order to retrieve comprehensive information and generate applicable knowledge out of it, processes need to be well described and modeled. A business process is described as "[..] *representation of a business process in a form that supports automated computation, such as modeling, or enactment by a workflow management system. The process definition consists of a network of activities and their relationships, criteria to indicate the start and termination of the process, and information about the individual activities, such as participants, associated IT applications and data, etc.*" [5] This definition contains all relevant elements of a business process modeling approach. On another descriptive level, processes have a desired *output* generated by an *input* and a *throughput*.[6] This classical idea of processes makes clear distinctions between the process stages. Please refer to figure (**??**). Exploring and describing national BR processes is to be done along the cases provided in the description of work of the BRITE project. Important in the context of this paper are:

[5] WfMC Glossary - WfMCTC- 1011

[6] E.g. transfer of a registered office to another country within the EU. *Output*: establishment of registered office in another country. *Input*: all relevant data from source country. *Throughput*: all relevant processes and information enabling the transfer.

– Transfer of registered office
– Opening a branch in another EU Member State

### 5.1 Developing high level BR process ontologies

We define process ontology as a conceptual description framework for business processes. This definition is built on the following basic assumptions:

– Processes ontologies are abstract and generalized
– Processes have interfaces to any repository (databases, other processes etc.)
– Processes have formerly defined inputs [*throughputs*] and outputs
– All processes described are based on the HLDO

Process ontologies have clearly defined borders between the process stages (*input, throughput* and *output*). Dependent on national conditions, interfaces to adequate sources (databases, other process in different domains) are to be defined, as figure (8) illustrates. In this model, an input (e.g. application for a

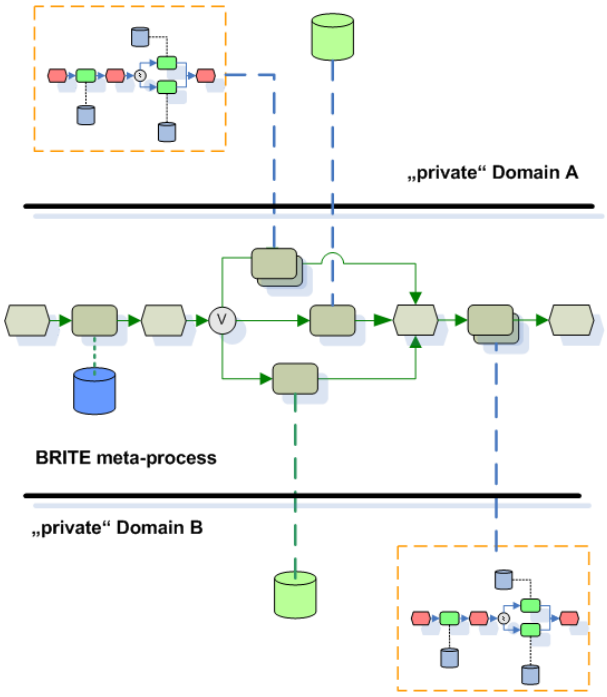


**Fig. 8.** Domains as boundaries

transfer of seat) and an output (e.g. transfer of seat) is given. All intermediary stages (interfaces to legacy systems, data storage, archiving) are defined in the

process ontology but executed on legacy systems. According to that idea, a limitation is also shown in the figure, one of the major threat to interoperability, the domain boundary. In the figure we refer to *"private" Domain A and B*, which are representatives for two separate Business Registers. The information flow is shown, but the integration level is very low, because of the clearly defined domain borders. In order to built real interoperability, these boundaries are to be broken - as far as possible. Therefore, the abstract process must be prepared for different kinds of input, due to particular situations and their context.

### 5.2 Integrating the processes

In the last step of our approach we are now integrating the process ontology into the heterogeneous landscape of European Business Registers. In order to protect the idea of an holistic integration and interoperability approach we are considering all integration possibilities. Depending on language, country, use case or any other influence, the abstract process can become concrete. By doing so, the process remains dynamic and adaptable. The HLDO is an indispensable prerequisite for this procedure. Figure 9 illustrates such an "opened-up process".

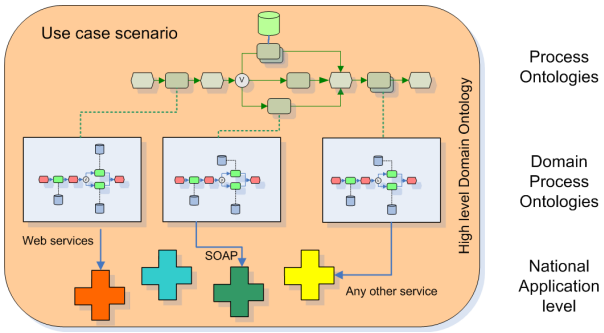


**Fig. 9.** Different interoperability layers of a Process Ontology

Figure 9 show three different layers: Process ontologies, domain process ontologies and national applications. They are linked together by the usage of different protocol and services, depending on the legacy systems and applications. All processes in this figure are based on the HLDO which is stored in the knowledge repository of BRITE. The figure shows a use case scenario which could be every use case, e.g. "transfer of seat". The domain process ontologies refer to the national application level using their specific protocols and service in order to retrieve, process and store data. Using this architecture, BRITE is enabled to use every source of all BR involved in the project - directly or indirectly.

- Databases and processes of BRITE
- Databases and processes of national Business Registers
- Third party repositories

**DB and processes of BRITE** The BRITE database contains the HLDO and all abstract processes. Only selected "live" data is stored centrally (as it is already done in the EBR) [7] in order not to get in conflict with the national regulations. The abstract processes are just the framework to enable the desired functionality described in the use cases.

**Databases and processes of national BR** These databases provide the essential information for the BRITE processes to work. They provide static information about companies as well as detailed information about processes (e.g. for registering, increase of capital stock etc.) which have important contributions to the BRITE process. Major constraints of proper integration are identification (authorisation/ authentication) of users, security constraints and legal restrictions. Detailed investigations in that field are done in literature.

**Third party repositories** Third parties can be assurances, finance providers, SCHUFA [8], IHK [9] and others. The (partial) integration of their databases is to be discussed. But the general architecture of the process ontology should be open enough for a latter integration of these.

These interfaces may vary from process to process, depending on all parameters possibly changing: country, legal form, use case scenarios etc. The usage of semantics (which are developed within in the HLDO) allows such a diversification. As a desired consequence, local variations are not affecting the functionality of the system. The depth of integration also depends on multiple parameters.

## 6 Summary

In this contribution, we have introduced an approach to develop a common ontology for European Business Registers that links up national Business Register ontologies. The development of a High Level Domain Ontology for European Business Registers allows to conciliate national Business Registers and, via the common high level concept, allows to fully interoperate among national or even local applications. The work is being carried out within the EC co-funded project BRITE, an integrated project within FP 6 of IST.
In order to design interoperability between different European Business Registers, the common high level ontology was used as a mapping tool. It was further used as a means to narrow down process ontologies and, based on that, to develop an intermediary that links up distinct national applications.
In our approach we firstly define a HLDO which enables a communication between all project partners and BR involved in the project. The development

[7] For an example browse to: http://www.ebr.org
Navigate to: products/view example
[8] Schutzgemeinschaft für allgemeine Kreditsicherung [german] Community to protect general capital interests
[9] Industrie und Handelskammer [german] Chamber of Industry and Commerce

is described as an iterative and self-repeating methodology consisting of selecting subsets of domain ontologies, merging them in order to delete redundancies and finally mapping the outcome against the domain ontologies. This approach maintains functionality of legacy systems and keeps them operating.

In the next step we are supporting the identification and description of processes related to the use cases. These domain process ontologies are abstracted to a general process ontology. By the help of this high level process ontology, legacy systems can be addressed over the borders of domains. The outcome of our approach is interoperability between different legacy systems on technical, organizational and semantical levels.

The first step we undertook was to resolve the differences within the vocabulary in order to get a common understanding of Business Register's topics. Then, semantics were added to the vocabulary in order to build a knowledge map and enable inferences. Finally, processes were modeled on a high level and were opened to other inputs (such as foreign or own data storages and processes) and created consequently the basis for interoperability. Final step will be to integrate these processes into the legacy systems, varying in depth and depending on legal, organizational and technical constraints.

## References


1. Modinis Programme. Study on Interoperability at Local and Regional Level. 2006. http://www.egov-iop.ifib.de/index.html.
2. eeurope 2005. an information society for all. eGovernment, *http : //europa.eu.int/ information_society/eeurope/newsl ibrary/ documents/eeurope*2005/*eeurope*2005_*en.pdf*, 2002.
3. Commission of the European Communities. Linking-up Europe: The Importance of Interoperability for e-Government Services. Staff Working Document, 2003.
4. Decision of the European Parliament and of the Council on Interoperable Delivery of pan European Services to Public Administrations, Businesses and Citizens (IDABC). Decision 2004/387/, *http : //europa.eu.int/idabc/*, 2005.
5. i2010 Eine europäische Informationsgesellschaft für Wachstum und Beschäftigung: Mitteilung der Europäischen Kommission und Nachfolgeprogramm von eEurope 2005. *http : //europa.eu.int/information_society/ eeurope/i*2010/*i*2010/*index_en.htm*, 2005.
6. J. Davies, D. Fensel, and F. Van Harmelen. In *Towards the Semantic Web Ontology-driven Knowledge Management. Sussex: John Wiley and Sons Ltd.*, 2003.
7. N. Guarino. Formal ontology and information systems. In *In Nicola Guarino, editor, Formal Ontology in Information Systems - Proceedings of FOIS'98, Trento, Italy, pages 3-15. IOS Press, Amsterdam*, 1998.
8. R. Klischewski. Migrating small governments' websites to the semantic web. In *Sematic Web meets eGovernment, Papers from the AAAI Spring Symposium, Technical report SS-06-06*, 2006.
9. L. e. a. Van Elst. Business register interoperability throughout europe: The brite project. In *Semantic Web Meets eGovernment - Papers from the AAAI Spring Symposium*, 2006.

10. M. Wimmer. Implementing a knowledge portal for egovernment based on semantic modelling: The e-government intelligent portal (eip.at). In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*, 2006.