Achieving Cross-Country Electronic Documents Interoperability with the help of a CCTS-based Modelling Framework

Fenareti Lampathaki[1], Spiros Mouzakitis[1], Till Janner[2], Christoph Schroth[2], Dimitris Askounis[1], Volker Hoyer[2]

[1] Decision Support Systems Laboratory, National Technical University of Athens,

9 Iroon Polytechniou Str., 15773 Zografou, Greece,

{flamp, smouzakitis, askous}@epu.ntua.gr

[2] SAP Research CEC, St. Gallen, Switzerland,

{till.janner, christoph.schroth, volker.hoyer}@sap.com

Abstract. The dawn of the XML era appears insufficient to guarantee a holistic approach in electronic transactions, as long as the semantics confusion, the lack of a common understanding of the underlying data and the business standards dilemma prevail and prevent data interoperability issues from being resolved. As various e-Government initiatives are being realized throughout the world, an imperative need for creating unified governmental data models that will facilitate the seamless exchange of information and the deployment of interoperable systems in Enterprises and Central, Regional and Municipal Government emerges. After attempting to identify relevant research that has been undertaken in the field of data modeling, the present paper analyzes the scope and the objectives of the business information modeling oriented towards governmental data. Based on UBL (Universal Business Language) and CCTS (Core Components Technical Specification), a component-based data modelling methodology enhancing business-to-government interoperability is proposed, creating the basis for a repository of governmental data models. The approach adopted may also serve as guidelines for creating and transitioning between "generic - harmonized" and "specific - contextualized" documents. A real world paradigm extracted from the Periodic VAT Statement document finally comes to prove the soundness of the proposed methodology.

Keywords. Semantic Interoperability, Data Modelling, Business and Government Semantics, Core Components Technical Specification (CCTS), Universal Business Language (UBL)

1. Introduction

Nowadays, Enterprises and Governmental Organizations seem to be, more than ever, challenged by the accelerating pace of change and innovation since globalization is putting increasing pressure on their operational efficiency and is demanding decrease in their integration costs bringing in the limelight the multi-disciplinary issue of interoperability. According to the Enterprise Interoperability roadmap [[13]], "*Successful enterprises of the future will be characterized by their ability to collaborate, their ability to adapt, and their ability to interoperate.*" Achieving interoperability, though, requires the resolution of syntactic and semantic interoperability issues [[22]], [[23]]. Syntactic interoperability is achieved through agreeing on common data structures and/or on transposing mechanisms between different structures, while semantic interoperability is obtained when stakeholders are able to share meaning, which is to understand each other. This implies sharing some context (e.g.: definition of terms, examples and counter examples, translations, etc.), which enables common interpretation of the data that is exchanged. It also implies that stakeholders have some joint objectives, which justify the data exchange, hence contributing to give it some additional contextual meaning.

To this direction, Data Modelling issues have aroused the interest of the research community since

the late 1960s when EDI (Electronic Data Interchange) [[38]] was introduced. In progress of time e-Business modelling frameworks went through an evolutionary path from monolithic and proprietary standards (e.g. TRADACOMS [[36]], ANSI X.12 [[2]]) towards flexible and standardized XML-based stacks covering the requirements from different industries [[34]]. These modern XML-based standards have indeed solved major technical issues of traditional EDI, but appear insufficient to provide a common understanding of the underlying data, arrange the semantics of the business information and achieve real interoperability [[19]]. The prevalent "business standards dilemma", defined as the diversity of standards that address particular data requirements, but are designed on such a different basis that make the choice of a specific standard to be adopted a new challenge, is compounding the problem.

In the e-Government sphere, a struggle to overwhelm the existing hurdles is observed and can be attributed to the following key-barriers: first, the majority of the standards applies to e-Business and does not take into account the specific needs of governmental processes and data, second, e-Government Interoperability Frameworks though often characterized as the e-Government Bibles have not achieved to address cross-country governmental data interoperability issues [[39]]. Further reactions after the publication of the first version of the European Interoperability Framework for pan-European eGovernment Services (EIF) [[22], [15]] have also demonstrated that a more profound understanding of eGovernment interoperability and its impact is needed. In order to deliver cross-border services, the European institutions and the Member States will have to agree on a multitude of semantic specifications, such as descriptions of people, products, processes, forms, etc. They will also have to agree on how to formulate these descriptions and where to store them for public use.

It is often stated that as long as the semantic discourse of data and standards exists, a holistic approach in e-Government cannot be achieved [[31]]. In this context, the present paper is oriented towards presenting an innovative solution, which combines the most promising concepts and technologies of the e-Business data standardization together with a comprehensive modelling environment. With regard to data modelling and the definition of reusable data components to obtain a common understanding of the business semantics the ISO 15000-5 standard (formerly UN/CEFACT Core Components Technical Specification - CCTS) [[38]] defines a state-of-the-art standardized methodology to define such components on a syntax independent level. An implementation of such data components is met in the UBL common library [[29]], which, however, has a clear focus on e-Business. The extension of scope of the UBL components to fulfil the requirements of the e-Government sector will be supported through our approach, which leverages the advantages of CCTS and UBL together with methodological support for an efficient modelling environment for business and governmental data. The component-based data modelling methodology aiming at enhancing business-to-government interoperability enables the creation of a common, standard-based repository of data components that conforms to UBL common components but with a support of the government sector's requirements.

The remainder of the paper is structured as follows: In the second chapter related work of other research groups, projects and also of the relevant standardization bodies is discussed. An introduction on Business Information and Governmental Data Modelling is given in chapter 3. Chapter 4 proceeds with the concepts of CCTS and UBL and the proposed data modelling framework. The application of the proposed approach is validated by a real-world scenario – the modelling of a Greek Periodic VAT statement – which follows in chapter 5 and leads to the creation of a repository of e-Government Components to be presented in chapter 6. A short summary of our results and further research activities required towards the next generation of enterprise interoperability complement this work.

2. Relevant Research Review

In the span of this work related research efforts which reference Business Information Modelling and Data Modelling Repositories were examined and reviewed. The main findings upon which the proposed data modelling approach builds originate from the following areas: eGovernment Interoperability Frameworks, Standardization Efforts and Research Results.

First, major initiatives being carried out by e-government agencies in the interoperability arena and having produced corresponding interoperability frameworks were pondered over [[39]]:

- In the United Kingdom, the e-Government Interoperability Framework [[6]] and its relevant specifications (e.g. the e-Government Metadata Standard [[7]] and the Schema Library [[8]]) as issued by the e-Government Unit,
- In Germany, the Standards and Architectures for e-Government Applications (SAGA) [[25]] by the KBSt,
- In Denmark, the Interoperability Framework [[10]] and the collaboration tool Infostructurebase [[11]]
- In France, the "Référentiel Général d'Interopérabilité (RGI)" [13] complementing the "Le cadre commun d'interopérabilité des systèmes d'information publics" [**Error! Reference source not found.**] issued by ADAE
- In Belgium, the Belgian Interoperability Framework (BELGIF) [[3]]
- In Greece, the e-Government Interoperability Framework [[16]] has delivered the Certification Framework for Public Sites and Portals, the Interoperability and Services Framework and the Digital Authentication Framework and a Registry of Process Models, reusable Core Components, XML Schemas and Web Services
- In a pan-European level, the European Interoperability Framework [[22]] which is currently being revised published by IDABC [[15]] and the Architecture Guidelines For Trans-European Telematics Networks for Administrations (Version 7.1) [[20]] are met

The main findings of standardization bodies appear as the second area our work is closely related to. To enhance business interoperability on a semantic level, the ISO 15005-5 Core Component Technical Specification (CCTS) [[38]] and numerous of closely affiliated standards issued by the UN/CEFACT [[41]], as well as the Core Component Library (CCL) [[42]] representing the repository for generic business data components, the so called Core Components, have been thoroughly investigated. UBL [6] with its library of pre-defined business document structures and data components is a further central reference part. Since the concept of the repository can be associated with predominant e-Business standards like ebXML [[13]] and RosettaNet [[30]], the ebXML Registry and the RosettaNet Dictionaries are also taken into account during this work.

Further relevant work analysis in the area of data modelling standardisation for inter-organisational business processes has also emerged from research papers, such as [[5], [21], [24], [32], [40]].

Besides this related work from the research and industrial community this work is embedded in the context of the EU-funded "GENESIS: Enterprise Application Interoperability via Internet-Integration for SMEs, Governmental Organisations and Intermediaries in the New European Union" Project [[17]]. Its aim is the research, development and pilot application of the needed methodologies, infrastructure and software components that will allow the typical, usually small and medium, European enterprise to conduct its Business transactions over Internet, by interconnecting its main transactional software applications and systems with those of collaborating enterprises, governmental bodies, banking and insurance institutions with respect to the EC current legal and regulatory status and the existing one in the new EU, candidate and associate countries.

3. Data Modelling in the context of e-Government

3.1 Background on Business Information Modelling

In the present paper, business information modelling is conceived as the process of structuring and organising data. Information modelling strives to bring the data structures of interest together into a cohesive, inseparable whole by eliminating unnecessary redundancies and by relating information structures with relationships.

Information models exist at multiple levels including [[18]]:

- The Conceptual Data Model describes data from a high level. It defines the problem rather than the solution from the business point of view. It includes entities and their relationships.
- The Logical Data Model describes a logical solution to a data project. It provides more details than the conceptual data model and is nearly ready for the creation of data structures. These details include attributes, the individual pieces of information that will be included. The outcome of the model is in essence encapsulated into an ontology. In general, ontologies are defined by Gruber as "formal, explicit specifications of a shared conceptualization" [[35]], which in the case of GENESIS project takes the form of documents, e.g. invoices, orders, etc.
- The Physical Data Model describes the implementation of data in a physical data structure. It is the technological blueprint for data schemas.

The different standards for data modelling existing can be classified on the basis of the three abovementioned levels. The UN/EDIFACT standard is a representative for a comprehensive and static data modelling approach. EDIFACT messages are not based on XML syntax and do not consider human-readable tags. Instead, they consist of a sequence of ASCII-coded segments, data element groups, data elements and values that are separated and structured by symbols such as colons, semicolons and apostrophes. These symbols are also utilized to realize a hierarchical structuring of the messages, which comprise IDs for sender and receiver, payload data, time stamps and check codes. EDIFACT offers inflexible, out-of-the box business document descriptions which tightly couple the conceptual, logical and physical data models introduced above.

As a second example, the XML-based RosettaNet standard goes beyond the mere modelling of data and specifies both the structure of the business documents to be exchanged during a business process and also the message choreography (sequence) in which they are sent between the trading partners. It also links physical representation of data and its underlying concepts and semantics and thus does not provide technological flexibility. For each of the business documents, there is static XML representation that can only hardly be adapted to actual business requirements.

In Chapter 4, the advantages of the emerging UN/CEFACT approach for business data modeling shall be presented. One of its main advantages is the loose coupling between the semantic, conceptual and the physical levels as introduced above. Technical representation is independent of the conceptual model, thereby contributing to an improved cross-organizational interoperability.

3.2 Modelling of governmental data – issues and challenges

Unified governmental data models for facilitating the seamless exchange of information and the

deployment of interoperable systems in Central, Regional and Municipal Government appear today as critical yet less touched issues that deserve more in depth exploration.

In general, modelling of governmental data is confronted with the following key-barriers that answer for the heterogeneity generally observed:

- The plethora of standards in data modelling (UBL [[29]], xCBL [[43]], eBIS-XML [[12]], OAGIS [[28]] and cXML [[9]] to name a few) apply exclusively to the field of e-Business transactions and a lack of standardization efforts for governmental documents, addressing how the underlying governmental information must be modelled, named and structured, is detected.
- The current European or National e-Government Interoperability Frameworks have developed repositories of XML schemas for the exchange of specific-context information throughout the public sector within the country borders. For example, the United Kingdom has developed the XML Schema Library, Greece has deployed an Interoperability Registry with service and document models, Denmark has designed the InfoStructureBase system and Germany the XML Infopoint which is to be replaced by the oncoming XML Repository. However, the lack of aggreement upon a universal language to describe the semantics of governmental data in unambiguous terms has led to tremendous confusion when attempting to create cross-country data interoperability since each element will most likely have the same semantic meaning but be expressed by synonyms (e.g. Surname, Last Name, and Family Name) and/or different meaning of terms (e.g. bank as a financial institution and bank as the side of a body of water). Syntactic, semantic and lexical rules also mostly vary per country context.

The cornerstone of government information exchanges has been a common understanding of semantics at a syntax-independent level. However, no government has been able to achieve this even in a national level since every governmental authority is focused on their own requirements and defining their specific, inflexible business data models according to their own assumptions and interpretations. Despite following the guidelines posed by the e-Government Interoperability Frameworks, each authority's XML vocabularies are often based on different methodologies for representing the semantics of governmental information, with the result that similar entities like for example the address in the VAT Statement and the INTRASTAT Statement are designed quite differently.

Another barrier in modelling governmental data is raised by the legal rules originating from each country's underlying legal framework. A complete set of rules holding information about actions that affect the data in the full spectrum of governmental transactions – specifically format and content of the exchanged documents – can be detected in national legislations. However, laws often define a set of required fields and leave the corresponding authorities a significant amount of lee-way as to exact supplementary data to be exchanged, with the result that discrepancies and conflicts in governmental information ultimately exchanged between different authorities not being a rare phenomenon.

When it comes to governmental data modelling in a pan-European level, the country dimension appears in the limelight. Despite the formulation of a rigid EU Legal Framework with which national frameworks must align and which constantly evolves with self-executing regulations and directives defining the format and the procedure for the transmission of the required data in specific documents (for example in a VAT Statement), a highly harmonized approach among the Member States has not yet been achieved.

In conclusion, the lack of standard governmental data models yields considerable lee-way in implementation efforts, yet in order to address the growing heterogeneity – in terms of the different complexities and the different structures for identical entities – in existing data models, modelling of governmental data requires a straightforward approach that will be based on context-neutral conceptual data models.

4. Component-based Data Modelling Framework

4.1 A Primer in CCTS and UBL Concepts

4.1.1 Core Components Technical Specification & Core Components Library

The Core Component Technical Specification (CCTS) [[38]] is the associated method comprising meta-models and rules for the semantically unambiguous definition of business information on a syntax-independent level. This second building block of the UN/CEFACT stack provides users with guidelines for correctly naming and combining Core Components and introduces ways to apply context-specific restrictions on the generic data templates by adding semantic precision for a given business requirement. The generic data templates can thus be adapted and restricted to the individual requirements of users, thereby ensuring the scalability of the whole system.

The Core Component Library (CCL) [[42]] represents the repository for generic business data components, the so called Core Components. Based on the experiences gained in previous data standardization efforts, the CCL does not provide pre-determined, static or industry-specific data definitions, but comprises a huge set of context-agnostic, generally valid data templates (e.g. postal address, personal information) that are syntax-independent and represent the general business data entities which are commonly used in today's business processes. Major benefits of leveraging such a Core Component Repository include an increase of reuse of data elements during modelling and improved enterprise interoperability due to a common basis for business information description. UN/CEFACT envisions this library to grow and also change over time as users can either modify existing components or design and submit new Core Components in case the existing ones are not sufficient to fulfil the actual business requirements.

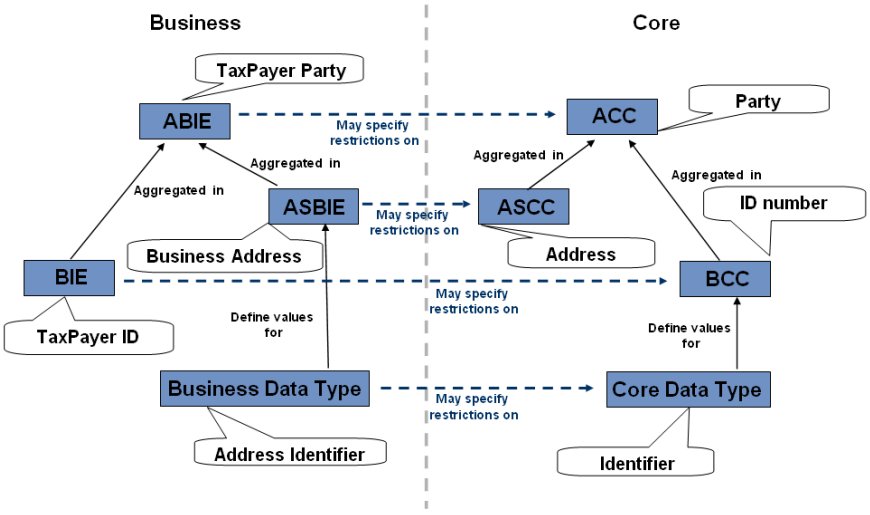

Fig. 1: CCTS Meta Model Entities (adapted to [[38]])

An example for the use of syntax-free Core Components is as follows: a postal address is a so-called

Aggregate Core Component (ACC) that consists of numerous components for representing data at leaf-level, such as "Street Name", "Zip Code" etc. These basic data items that cannot be further decomposed into other components are called Basic Core Components (BCCs). The generic address template comprises all BCCs that could ever occur in an arbitrary user context and is stored as an unstructured, syntax-free list. In case a Greek user desires to model a business document that comprises a delivery address, he is able to access the CCL, apply his Greek context and is provided an address component that features only those BCCs that are relevant for his specific situation. This contextualized subset of the generic template is then called Business Information Entity (BIE). To semantically enrich the BIE even more, it is also possible to add qualifiers and, for example, transform a "Zip Code" into a "Post Office Box_ Zip Code" to allow for differentiation from other "Zip Code" instances. In this way, all users have a common understanding of how the different parts of an address are named and are yet able to use arbitrary syntactical representations for their business documents.

The UN/CEFACT Naming and Design Rules (NDR) [[41]] define a set of guidelines for transforming CCTS based artefacts into XML Schema and XML based instances. UN/CEFACT however does not prescribe users to build upon this technical representation of its semantic data models but also allows for leveraging other XML languages such as UBL.

4.1.2 Universal Business Language (UBL)

The Universal Business Language [[29]] is a royalty-free library of standard electronic XML business documents, designed to provide a universally understood and recognized commercial syntax for legally binding business documents. Developed by the Organisation for the Advancement of Structured Information Standards (OASIS), UBL's purpose is to provide:

- A library of XML schemas for reusable data components, such as Address, Item and Payment, which are the common data elements of everyday business documents.
- A set of XML schemas for common business documents such as Order, Despatch Advice and Invoice that are constructed from the UBL library components and can be used in generic procurement and transportation contexts.
- A set of processes and business rules associated with the business documents that define a context for their use.

UBL operates within a standard business framework such as ISO 15000-5 in order to provide a complete, standards-based infrastructure that can extend the benefits of existing EDI systems to businesses of all sizes. As the first standard implementation of UN/CEFACT CCTS, the UBL Library is based on a conceptual model of information components known as Business Information Entities. These components are assembled into specific document models such as Order and Invoice, which are then transformed in accordance with UBL Naming and Design Rules into W3C XSD schema syntax.

UBL can be characterized as an emerging standard that has the credentials to dominate the area of Data Modelling, but to date its scope is rather limited as it covers only the basic business documents involved in a few common B2B processes – UBL Version 1.0 supported 8 business documents, while UBL Version 2.0 provides schemas for 29 business documents. For instance, there is a whole range of business documents before and after a payment scenario that UBL does not cover. Moreover, UBL currently does not support transactions between businesses and the government or banking institutions. Therefore, taking into account that UBL's customization is relatively rigid, the need for a innovative modelling approach, that on the one hand takes advantage of UBL' s offerings but on the other, extends its scope and usage is identified.

In the span of the GENESIS project, we decided to choose UBL instead of the UN/CEFACT NDR due to the rich UBL Core Component Library already existing.

4.2 GENESIS Data Modelling Framework

The integrated modelling approach of GENESIS aims at reusing the existing elements of UBL and CCTS as templates and the definition of new elements where neither UBL nor CCTS do provide sufficient support. Therefore the meta model must be able to depict the existing UBL templates in the modelling tool ADONIS® [[4]] and the definition of similar but new templates. Furthermore the meta model has to provide easy-to-use and easy-to-understand modelling entities. In this way the GENESIS integrated modelling approach is able to overcome the gaps described in chapter 4.1 and to master the challenges involved.

A main target of the integrated GENESIS modelling approach is to refrain from creating redundant information. The key element in terms of describing business information in the context of the GENESIS framework is the *"Specific Business Document"* (SBD), which serves as an entry point as well as a data repository. Even though the objects (re-)using the SBD may vary in context of the modelling stage it always references the same SBD.

According to the different levels of IT knowledge of each modeller, the SBD distinguishes between *"Unstructured Documentation"* and *"Structured Documentation"*. The unstructured documentation requires only basic input on a business document like "Business information entity" (the name of the data field), "Mandatory", and "Formatting Rules". It allows a user with a limited technical know-how to easily describe a business document. Derived from the unstructured documentation, the structured information is created. The creation of the structured documentation requires knowledge of CCTS and UBL. The basis for the creation is provided by the "UBL Common Library" (UBL CL) and the "Core Component Library" (CCL). In a first step the user copies the UBL CL and the CCL into a user specific model group to create the foundation for the user specific libraries which are not necessarily UBL compliant but based on UBL. In a next step the user analyses the information of the unstructured documentation and converts the information into the CCTS elements "Aggregated Business Information Entity" (ABIE), "Associate business Information Entity" (ASBIE) and/or "Basic Business Information Entity" (BBIE). For the conversion process the user tries to identify the unstructured elements in the UBL CL and copy the required information from the user specific UBL template into the SBD. If UBL cannot provide a satisfying template the user is allowed to create an own definition in the user context based upon the CCL. When all unstructured information of all SBD of an end user has been transformed into structured information, the end user provides a user specific library based upon UBL and extended as far as necessary. As a last step the similar SBD of different users are assigned to one "Generic Business Document" (GBD) via inter model reference. The GBD incorporates the information superset of all SBD referenced and provides a generic library of documents which can be easily reused and extended.

Due to the fact that core entities only describe an abstract formulation without business semantics they are not realized in the GENESIS meta-model directly but indirectly through the business entities offered to the modeller.

The model shown in the real world scenario of Chapter 5 (Fig. 2) describes entities interpreted as relations and aggregated into an ABIE. The entities aggregated into ABIE are therefore realized as table entities called records in ADONIS® and not as own objects. This solution allows the user to easily structure information without overloading the modelling space. The relations from ASBIE to an ABIE are solved by inter model references which can be understood as virtual connectors. The "Business Data Type" (BDT), used to define a BBIE in an ABIE or "Business Document", is based upon the CCL and also connected via inter model reference. ABIE and BDT are directly implemented into the "Business Information Model" (BIM). In addition to the CCTS related objects the BIM provides the SBD and GBD for a holistic business information modelling. To allow a user to focus on either the

CCTS elements or the business document modelling the BIM offers a visualization mode to filter either CCTS related objects or business document elements or shows all objects.

5. A real scenario

A real world paradigm applying the GENESIS Modelling Methodology is built on a specific country's – Greece – VAT Statement. The decision was reached upon the criterion that it is one of the most critical documents exchanged between citizens and businesses and the State and appears in all countries of the European Union.

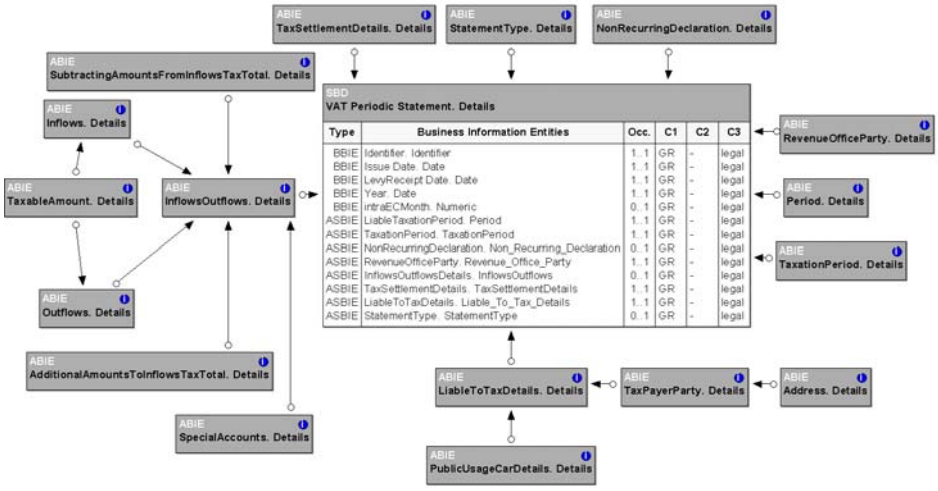

Fig. 2: The Greece VAT Statement, according to the proposed integrated modelling approach

VAT Statement declaration in Greece takes place every 3 months. Citizens working as freelancers and businesses are obliged to submit their VAT Statement declaration and to pay or be credited the subsequent tax amount. The transaction can be fully automated with the use of TAXISnet [[33]] or can take place directly in the authorized Tax Agencies in the jurisdiction of which the liable to tax person or business belongs. TAXISnet (the word stands for TAXation Information System) is a set of electronic services for taxation accessible through Internet to all citizens and enterprises in Greece. Its main services are: *e-VAT*, e-filling of VAT forms and VAT payment through Banking System, *e-Income Tax service*, e-filling of Income Tax Forms and the *e-Income Tax-Assessment info service* which provides personalized information for Income Tax Assessment.

Adhering to the aforementioned GENESIS modelling methodology, the fields of the VAT Statement were analyzed and categorized. In the next step the necessary BIEs were identified and then the core components provided by the UBL common library which could be reused in the VAT Statement context were assembled and customized according to the particularities of the VAT Statement document in Greece and finally the specific business document (see Fig. 2) was created. For visibility reasons, only the table representing the specific business document (SBD) VAT Statement is explicitly presented.

In the proposed approach, the type (Basic Information Entity – BBIE, Association Business Information Entity – ASBIE mapping to an Aggregate Business Information Entity – ABIE), the name complemented with the data type, the occurrence and the context are placed among the crucial metadata for an information entity. The country dimension is apparent from the field C1, with set value for our example GR – Greece.

The relation between the UBL Business Information Entities (BIEs) and the Business Information Entities customized to the needs of the VAT Statement and ultimately adopted is evident from Figure 3. As depicted, the Tax Payer Party ABIE extends its UBL corresponding ABIE, named Person in the sense that it creates two additional information entities: Telephone (BBIE) and Address (ASBIE), while the BIE Period appears as a "light" version of the UBL Period BIE upon deleting all unnecessary (and non required) fields.

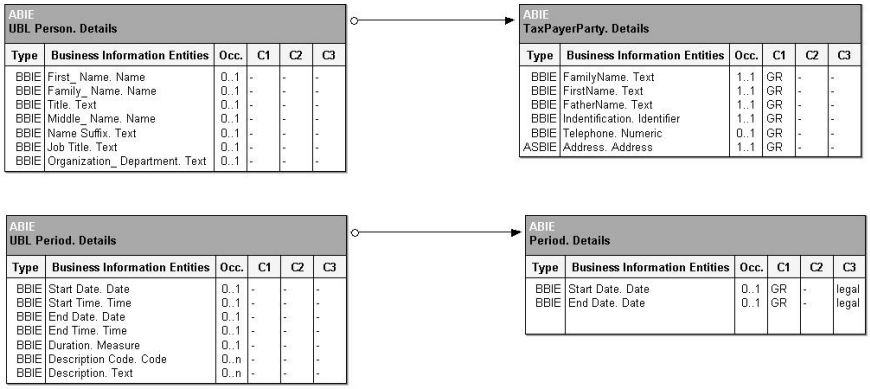

Fig. 3: The relation between UBL BIEs and VAT Statement's BIEs

6. Repository of e-Government Components

In order to create the first release of an e-Government Repository, the documents exchanged during several business-to-government transactions, which were modelled in the span of the GENESIS Project, were examined. On the basis of the documents' prestige and frequency, it was concluded to initially apply the proposed Data Modelling Methodology to: Periodic VAT Statement, Annual VAT Statement INTRASTAT Statement - Arrivals, INTRASTAT Statement – Dispatches, Social Security Statement and Declaration of a new employee in the span of four countries, i.e. Greece, Italy, Cyprus and Turkey. The Aggregate Business Information Entities (ABIEs) identified in the aforementioned documents were assembled into generic Components, leading to the repository of e-Government Components depicted in the following figure. The generic Components may have been extracted by more than one ABIE, in the case that conflicts in the names or the context were detected (e.g. the Component Person represents the common denominator of the Periodic and Annual VAT Statement, the INTRASTAT Statement and the Social Security Statement).

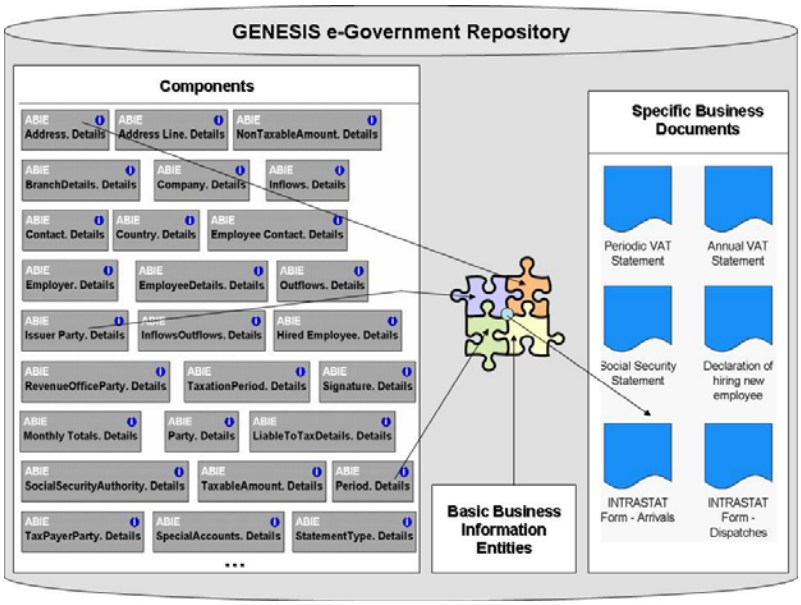

Fig. 4: The e-Government Repository

7. Conclusions

The scarcity of standard governmental data models and the fact that most of the data modelling specifications do not support business to government transactions create the need for a thorough data modelling framework that will provide an approach for modelling e-Government documents and at the same time will generate a reusable repository of e-Government Core Components.

In contradiction to most other data modelling approaches, that start from scratch and propose their own methodologies and schemas, leading to a war of conflicting and recriminating specifications that eternalize the interoperability problem, our data modelling framework suggests an innovative approach: By reusing the existing elements of UBL, the emerging widely respected data modelling standard, our framework takes advantage of UBL's undeniable offerings and expertise and promotes a creative freedom that offers flexibility and extensibility, without being confined to UBL's scope and customization limitations. Furthermore, the proposed approach yields considerable lee-way in implementation efforts and respects each country's and public authority's autonomy in defining the exact data interchanged, yet in order to address the growing heterogeneity – in terms of the different complexities and the different structures for identical entities – in existing data models, it adopts a straightforward approach that will be based on context-neutral conceptual data models for modelling of governmental data.

Future steps of our work include further research on customization issues regarding the Generic and Specific Documents as well as expansion of the involved countries.

References

[1] ADAE, Le cadre commun d'interopérabilité des systèmes d'information publics, Version 2.1, 2003

[2] ANSI Accredited Standards Committee (ASC) X12, 2007, http://www.x12.org/

[3] BELgian Governement Interoperability Framework, http://www.belgif.be/index.php/Main_Page.

[4] BOC ADONIS Business Process Management Suite, 2007, http://www.boc-eu.com/index.jsp;jsessionid=21546A540030913013EF4407FBD00C0D?file=WP_582571cc1ed802de.46e381.f59775478f.-7f17&lg=en

[5] Brahim Medjahed, Boualem Benatallah, Athman Bouguettaya, Anne H. H. Ngu, Ahmed K. Elmagarmid; Business-to-business interactions: issues and enabling technologies, The VLDB Journal (2003) 12: 59–85

[6] Cabinet Office – e-Government Unit: e-Government Interoperability Framework, Version 6.1, Retrieved February 5, 2007 from http://www.govtalk.gov.uk/documents/eGIF%20v6_1(1).pdf.

[7] Cabinet Office – Office of the e-Envoy, e-Government Metadata Standard, Version 3.1, Retrieved February 5, 2007 from http://www.govtalk.gov.uk/documents/eGMS%20version%203_1.pdf

[8] Cabinet Office, UK GovTalk Schema Library (2007), http://www.govtalk.gov.uk/schemasstandards/schemalibrary.asp

[9] cXML Version 2.0.17, Retrieved June 2007 from http://cxml.org/

[10] Danish Interoperability Framework, Version 1.2.14, Retrieved February 5, 2007 from, http://standarder.oio.dk/English/Guidelines

[11] Danish e-Government Project, InfoStructureBase (2007), http://isb.oio.dk/info

[12] eBIS-XML Suite Version 3.09, Retrieved May 2006 from http://www.basda.org/VD65/default.asp?PSID=51

[13] ebXML, 2007, http://www.ebxml.org/

[14] Enterprise Interoperability Research Roadmap, http://cordis.europa.eu/ist/ict-ent-net/eiroadmap_en.htm

[15] Gartner Group, Preparation for Update European Interoperability Framework 2.0 - Final Report, June 2007, http://ec.europa.eu/idabc/servlets/Doc?id=29101

[16] Greek Interoperability Framework, 2007, http://www.e-gif.gov.gr

[17] GENESIS Project, 2007, http://www.genesis-ist.eu

[18] Graeme C. Simsion, Graham C. Witt, Data Modelling Essentials, Third Edition, Morgan Kaufmann Publications, Elsevier, 2005

[19] G. Stuhec, “How to solve the business standards dilemma - CCTS key model concepts,” SAP Developer Network (SDN), http://sdn.sap.com, (2006)

[20] IDA Architecture Guidelines For Trans-European Telematics Networks for Administrations, Version 7.1, September 2004, http://ec.europa.eu/idabc/en/document/3485/5585

[21] IDA Project: XML-Based Business Frameworks, Deliverable 2.1 Evaluation of XML Frameworks, Version 2.1, May 2003

[22] IDABC, European Interoperability Framework for pan-European e-Government Services, Version 1.0, Retrieved February 5, 2007 from http://europa.eu.int/idabc/en/document/3761

[23] IDABC Content Interoperability Strategy, Working Paper, http://ec.europa.eu/idabc/servlets/Doc?id=24405

[24] Juha-Miikka Nurmilaakso, Paavo Kotinurmi, Hannu Laesvuori; XML-based e-business frameworks and standardization, Computer Standards & Interfaces 28 (2006) 585–599

[25] KBSt unit at the Federal Ministry of the Interior, SAGA Standards and Architectures for e-Government Applications Version 3.0, Retrieved February 5, 2007 from http://www.kbst.bund.de/cln_046/nn_945224/SharedDocs/Anlagen-kbst/Saga/standards-and-Architectures-for-_20e-Government-applications-version-3__0-

[26] Ministère délégué au budget et à la réforme de l'Etat, Direction Générale de la Modernisation de l'Etat, Référentiel Général d'Interopérabilité (RGI), Version 0.90, 2006

[27] MODINIS Study on Interoperability at Local and Regional Level, Version 2.0, April 2007, http://www.epractice.eu/files/media/media1309.pdf

[28] OAGIS Version 9.1, Retrieved May 2007 from http://openapplications.org/oagis/9.1/index.html

[29] OASIS, Universal Business Language (UBL) Version 2.0, Standard December 2006, http://docs.oasis-open.org/ubl/os-UBL-2.0.zip

[30] RosettaNet, 2007, http://www.rosettanet.org/index.html

[31] Spiros Mouzakitis, Fenareti Lampathaki, Christoph Schroth, Ulrich Scheper, Till Janner, Towards a common repository for governmental data: A modeling framework and real world application, in Enterprise Interoperability II: New Challenges and Approaches (Springer), Proceedings of the 3rd International Conference Interoperability for Enterprise Software and Applications I-ESA 2007, Funchal (Madeira Island) Portugal, March 2007

[32] S. Androutsellis-Theotokis, V. Karakoidas, G. Gousios, D. Spinellis, Y. Charalabidis: «Building an e-Business Platform: An Experience Report», e-Challenges 2005 Conference, European Commission, Luibliana, 21 - 25 October 2005

[33] TAXISnet, https://www.taxisnet.gr/, 2007

[34] T. Janner, A. Schmidt, C. Schroth, and G. Stuhec, "From EDI to UN/CEFACT: An evolutionary path towards a next generation e-business framework," in Proceedings of The 5th International Conference on e- Business 2006 (NCEB2006), Bangkok, (2006)

[35] T. R. Gruber, "A translation approach to portable ontologies", Knowledge Acquisition, Vol. 5, No. 2, (1993).

[36] TRADACOMS, 2007, http://en.wikipedia.org/wiki/TRADACOMS

[37] United Nations Directories for Electronic Data Interchange for Administration, Commerce and Transport (UN/EDIFACT), 2007, http://www.unece.org/trade/untdid/welcome.htm

[38] UN/CEFACT Core Components Technical Specification, Part 8 of the ebXML Framework, Version 2.01 (November 2003), http://www.unece.org/cefact/ebxml/CCTS_V2-01_Final.pdf

[39] Yannis Charalabidis, Fenareti Lampathaki, Dimitris Askounis, Antonis Stassis, Shifting to Second Generation E-Government Interoperability Frameworks, Proceedings of the International EGOV 2007 Conference, Regensburg (Germany), September 3-7, 2007

[40] Y. Charalabidis, S. Pantelopoulos, Y. Koussos: "Enabling Interoperability of Transactional Enterprise Applications", Workshop on Interoperability of Enterprise Systems, 18th European Conference on Object-Oriented Programming (ECOOP), Oslo, 14–18 June 2004

[41] UN/CEFACT XML Naming and Design Rules, Version 2.0 (February 2006), http://www.unece.org/cefact/xml/XML-Naming-and-Design-Rules-V2.0.pdf

[42] UN/CEFACT Core Component Library (UN/CCL), version 1.0, http://www.unece.org/cefact/codesfortrade/codes_index.htm, 2006

[43] xCBL Version 4.0, Retrieved May 2006 from http://www.xcbl.org/