Deploying a **Semantic-XML Message Builder Workbench** using TopBraid Suite™ at The Netherlands Ministry of Justice

SOLUTION CONTEXT AND REQUIREMENTS:

Enterprises (both commercial and government organizations) have a critical need to exchange information. Often these exchanges are implemented as electronic messages. Effective communication across company boundaries requires common business semantics. But the information comes from systems developed independently from each other to serve the distinctive needs of each community. They lack common semantics; that makes seamless data exchanges challenging.

XML is often seen as the best available and most heavily utilized data exchange mechanism and standard of choice. But with a continuing proliferation of XML standards, many organizations find that rather than it solving their data and message exchange problems, they become 'tangled up in XML'. They find that they are in a 'standards dilemma' with many standard XML schemas to choose from, but none fitting all the local needs. XML Schema mechanisms for extensions are highly limited. Enforcing conformance to XML naming and design rules across exchange partners is difficult. Resulting data models are too often brittle - becoming very complex, and sometimes nearly incomprehensible and non-implementable.

Examples abound of what can go wrong with attempting to match components of XML documents to achieve

CHALLENGES:

- Data is in many different systems not designed with sharing in mind
- Business documents/messages need to be tailored for local use
- There is a growing need for a common standard across organizations
- Without semantic and structural alignment, data exchanges result in significant failures and overheads
- Among the challenges to 'living in the XML ecology' for data exchange are:
  - How do you have common vocabularies?
  - How do you have a common way to construct schemas?
  - How do you represent data in a consistent way?
  - How do you represent data types in a consistent way?
  - How do you preserve semantics?

data exchange. Some simple causes of mismatching are: different information content in the 'same' tags, presence of different tags and meanings and different positions of information within tag hierarchies. There are many other common problems that become barriers to semantic interoperability.

Case study: The Metadata Workbench Solution at The Netherlands Ministry of Justice (MoJ), built by TopQuadrant with the MoJ

The organizations within the Ministry of Justice (MoJ) of the Netherlands need to share information covering many domains of law. Unique data requirements exist at courts, police, hospitals, border control, motor vehicle, and local and federal offices. For some time, the MoJ understood that the risks and costs of poor information sharing to society are high including:
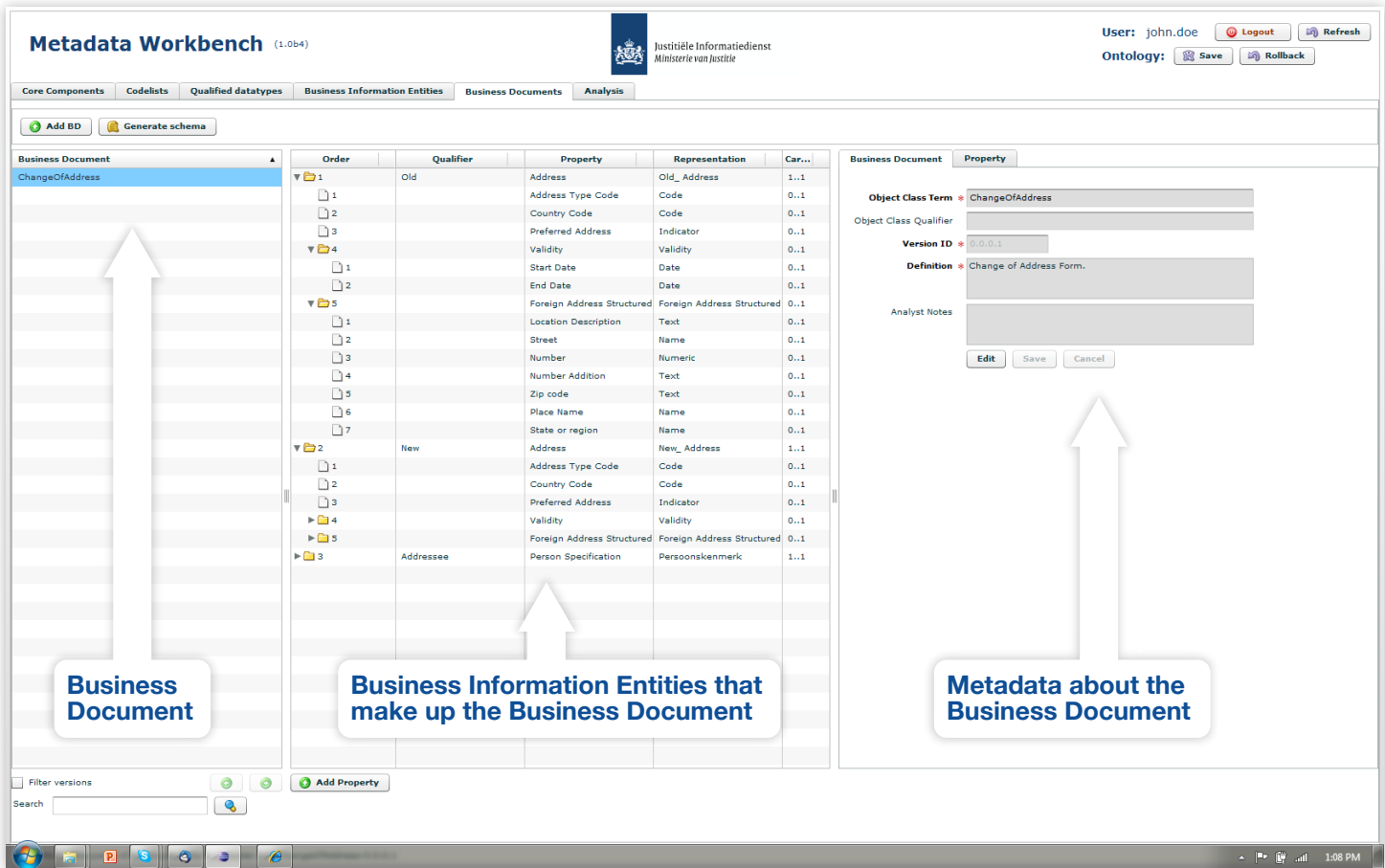
- Unique data requirements within and across the multiple domains of law are not addressed well
- Criminals not caught in time; security threats not prevented
- Mistakes happen due to incomplete data
- Inconveniences and costs to citizens of incorrect data, or having to provide the same data repeatedly

TopQuadrant™

TopQuadrant is a premier semantic web solutions company with offices in Alexandria, VA, and Mountain View, CA. Please contact us at:

TopQuadrant, Inc.
330 John Carlyle Street
Suite 180
Alexandria, VA 22314-5760

www.topquadrant.com
Phone: 703 299 9330
Fax: 703 299 8330

Metadata Workbench (1.0b4)

Justitiële Informatiedienst
Ministerie van Justitie

Core Components | Codelists | Qualified datatypes | Business Information Entities | Business Documents | Analysis

Add BD | Generate schema

Business Document
ChangeOfAddress

Order	Qualifier	Property	Representation	Car...
▼ 1	Old	Address	Old_ Address	1..1
1		Address Type Code	Code	0..1
2		Country Code	Code	0..1
3		Preferred Address	Indicator	0..1
▼ 4		Validity	Validity	0..1
1		Start Date	Date	0..1
2		End Date	Date	0..1
▼ 5		Foreign Address Structured	Foreign Address Structured	0..1
1		Location Description	Text	0..1
2		Street	Name	0..1
3		Number	Numeric	0..1
4		Number Addition	Text	0..1
5		Zip code	Text	0..1
6		Place Name	Name	0..1
7		State or region	Name	0..1
▼ 2	New	Address	New_ Address	1..1
1		Address Type Code	Code	0..1
2		Country Code	Code	0..1
3		Preferred Address	Indicator	0..1
▶ 4		Validity	Validity	0..1
▶ 5		Foreign Address Structured	Foreign Address Structured	0..1
▶ 3	Addressee	Person Specification	Persoonskenmerk	1..1

Business Document | Property

Object Class Term *	ChangeOfAddress
Object Class Qualifier	
Version ID *	0.0.0.1
Definition *	Change of Address Form.
Analyst Notes	

Edit | Save | Cancel

Business Document

Business Information Entities that make up the Business Document

Metadata about the Business Document

Filter versions
Search
Add Property

Figure 1: Example of a User Composing a Business Document from the CCTS Core Vocabulary.

- Higher costs for government to provide services

- Failure to have changes in legislation show up in data exchange representations

Various attempts were made by MoJ over time to address the problem, but they had not succeeded due to having:

- No generic reusable components - no reuse and much manual work (this only works for a short time)

- No conceptual model – everything lives in the head of the information model designer (ontologist)

- Version control existing only for each separate project

- No traceability – comparisons done manually

From these experiences, MoJ recognized that exchanging information between government parties requires a consistent, reusable and repeatable approach to specifying data exchanges as structured electronic business documents built from components. With the assistance of TopQuadrant, the MoJ resolved to design and implement a solution based on semantic web standards and technologies and an additional standard designed to bridge the conceptual world and the implementation-driven world of electronic messaging. The United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT) "Core Component Technical Specification (CCTS)" standard was chosen. CCTS describes an electronic message in a composable way. CCTS is a standard with growing adoption that provides:

- Reusable building blocks for building business documents

- A common semantic model

- Context support for industry/domain specific documents

Semantics of the business information is based on a standard grammar and library that is well known and understood by both people and computers. CCTS has been adopted in multiple industries – including automotive, retail, eCommerce, aerospace, manufacturing, and telecom, for government data exchange (as part of the OAGIS standard) at the USAF, Department of the Navy, DoD DCMS, and by major vendors including SAP and IBM.

As shown in Figure 1, the resulting Metadata Workbench, a semantic, model-driven solution built on the TopBraid Suite platform, is a new approach to the design of XML messages and generation of XML Schemas based on CCTS. Users can collaboratively design messages for data exchanges within and between organizations that are specific to the local context while remaining compliant with industry and enterprise standards.
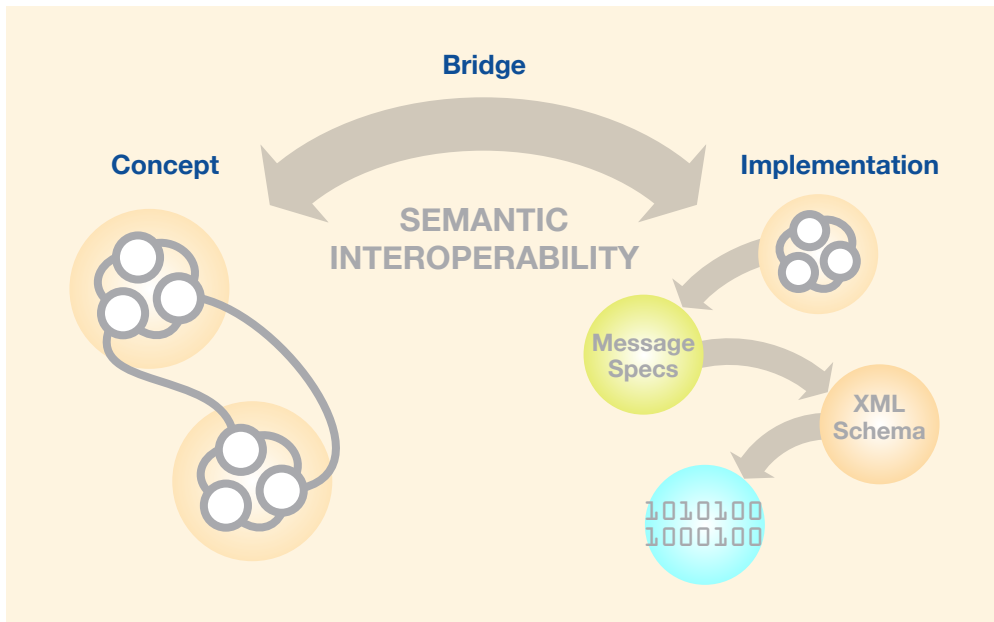
Figure 2: MoJ's Motivations for Using RDF / OWL with CCTS
As a principal member of the Central Information Systems of the Dutch government, with the Metadata Workbench solution, the MoJ is pioneering new approaches to business documents and message design with an emphasis on semantic checking, model-based generation of schemas and reuse of business components. As illustrated in Figures 2 and 3, domain models are transformed into CCTS Ontologies that allow business analysts to assemble business documents for electronic messages from component parts.

Using the CCTS ontology models, Information Analysts tailor and compose components to specify the business documents that make up the electronic messages. A web browser based user interface guides them in constructing the message exchange schemas. CCTS-conformant ontologies can be created directly in the workbench. Alternatively, they can be generated from already existing domain models including any RDF/OWL ontology. An XML Schema for each business document is generated using XML SchemaPlus (XSP), a technology developed by TopQuadrant in a project with NASA. XSP implements an approach to creating XML Schemas that preserve necessary semantics from RDF/OWL. This enables seamless round-tripping between an XML Schema and its representation in RDF/OWL.

Key Capabilities of the Solution

- Ability to create purpose-specific Business Information Entities from reusable core information components

- Construction of the Business Documents (electronic messages) by combining one or more Information Entities

- Import of the Domain Model from a variety of sources

- Automated generation of XML Schemas for Business Documents

- Flexible management of code lists, data types, terms, qualifiers and constraints in the form of metadata

- Support for versioning, governance and difference comparisons

Benefits of Data Exchange Solutions built with TopBraid Suite™

The case study of the MoJ Metadata Workbench exemplifies a class of solutions that we support and generally characterize as a Semantic-XML

Message Builder Workbench based on UN/CEFACT CCTS. Such solutions deliver multiple benefits. In the case of the MoJ, XML Schemas generated by the workbench get used by the various systems in the MoJ as the lingua franca. For instance, more than one system creates or processes change of address information. They can now either use the same XML Schema (when it fits) or use different schemas based on the same core components. Thus, when elements like address or city or name are used across different systems their semantics are unambiguous.

Furthermore, workbench generated XML Schemas have special "hooks" that allow conversion of XML messages based on the schema back to RDF without loss of meaning. This makes it possible to merge and aggregate message information across different systems.
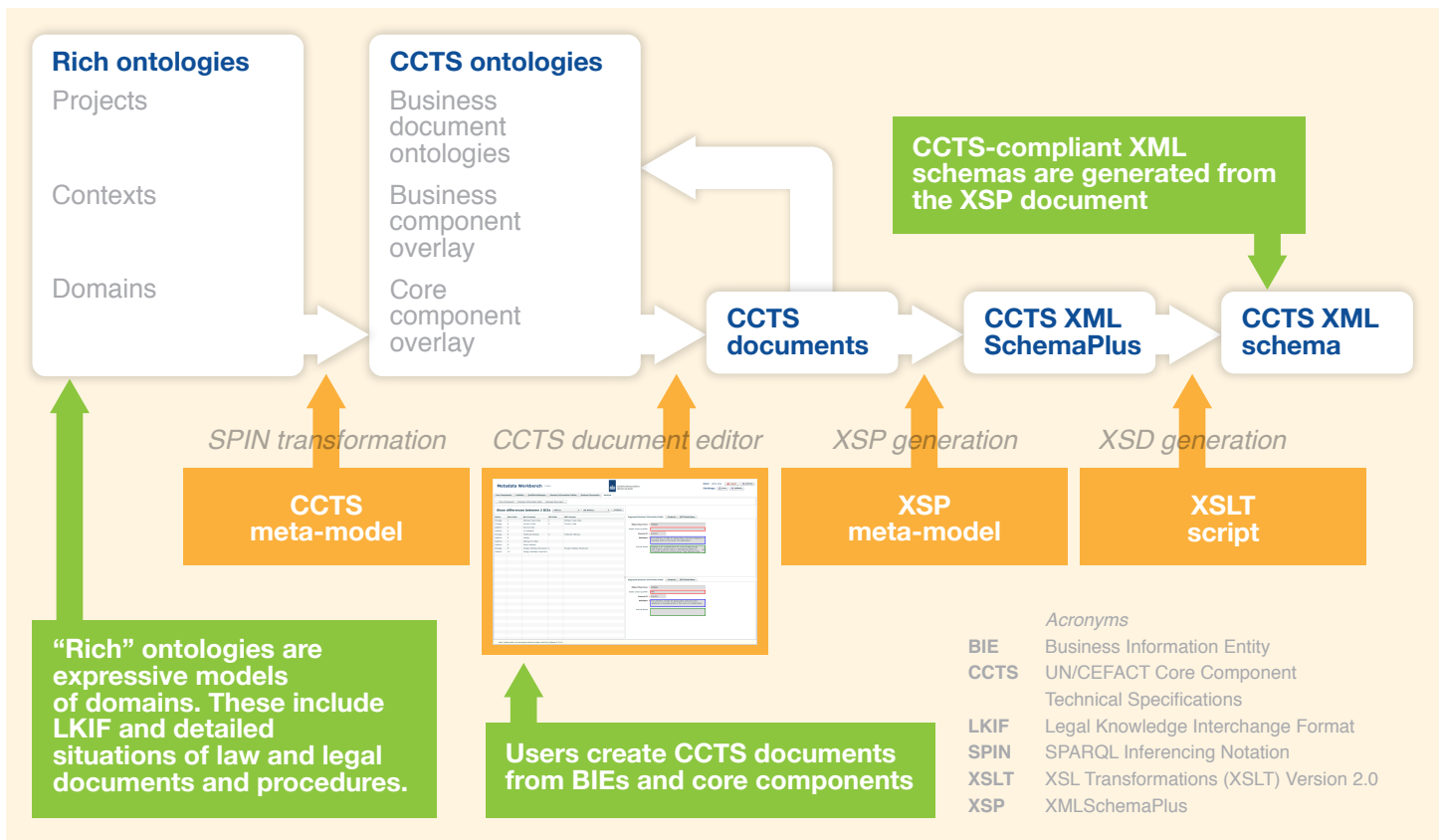
Business Benefits

- Accurate communication between organizations

- Agility in response to legislative changes

- Data quality guarantees

- Reduced message schema development costs

Technical Benefits

- Reuse

- Semantic consistency

- Traceability

- Version and change management

Semantic web technologies and TopBraid Suite provide key capabilities to the implementation of semantic-enabled XML message exchange solutions such as the MoJ Metadata Workbench. TopBraid Suite is a software platform that enables the rapid customization of knowledge applications enabling users to collaboratively

Rich ontologies
Projects

Contexts

Domains

CCTS ontologies
Business document ontologies

Business component overlay

Core component overlay

CCTS-compliant XML schemas are generated from the XSP document

CCTS documents

CCTS XML SchemaPlus

CCTS XML schema

SPIN transformation

CCTS ducument editor

XSP generation

XSD generation

CCTS meta-model

XSP meta-model

XSLT script

"Rich" ontologies are expressive models of domains. These include LKIF and detailed situations of law and legal documents and procedures.

Users create CCTS documents from BIEs and core components

Acronyms

BIE	Business Information Entity
CCTS	UN/CEFACT Core Component Technical Specifications
LKIF	Legal Knowledge Interchange Format
SPIN	SPARQL Inferencing Notation
XSLT	XSL Transformations (XSLT) Version 2.0
XSP	XMLSchemaPlus

create new information and efficiently use and interpret information created by others.

Capabilities of the TopBraid Suite platform include:

- Import and export to a variety of formats

- Data mappings and transformations

- Provenance and governance

- Change management and versioning

- Report generation

- Rich composable web user interfaces

- Semantic web services

Availability of this Solution and Where to Find Out More

The MoJ Metadata Workbench application will be made available by the MoJ in conjunction with TopQuadrant in the 3rd quarter of 2010 under an open source license. The TopBraid Live server is required to run the application. Contact TopQuadrant or visit: http://www.topquadrant.com/solutions/semantic_message_builder.html to find out more about our Semantic-XML Message Builder Workbench solution. The MoJ Metadata Workbench was featured in a webinar in May, 2010 and a recording of an extended demo is available for download from that page.

Figure 3: Generation of XML Message Schemas from User Composed Business Documents