TripCom

Triple Space Communication

FP6 – 027324

Deliverable

D9.1

Continuous Report on Standards Activities in Related Fields

Daniel Martin (USTUTT)
Daniel Wutke (USTUTT)
Dario Cerizza (Cefriel)
Alessio Carenini (Cefriel)
Jacek Kopecký (LFUI)
Janne Saarela (Profium)
David de Francisco (TID)
Doug Foxvog (NUIG)

April 6, 2009

Executive Summary

Work Package 9 provides continuous technology watch on emerging standards in Internet Technology, Web services Technology as well as on standards and formats used in Business Process Management. In WP9, we monitor ongoing standardization activities such as W3C, OASIS, and GGF. Furthermore, the work package identifies modifications and extensions of such existing standards that will enable them to support Triple Spaces. This document reports on standards related to TripCom, as well as assesses them for potential use in our project. Furthermore, we provide a report on our own standardization activities.

Document Information

IST Project Number	FP6 – 027324	**Acronym**	TripCom
Full Title	Triple Space Communication		
Project URL	http://www.tripcom.org/		
Document URL			
EU Project Officer	Werner Janusch		

Deliverable	**Number**	9.1	**Title**	Continuous Report on Standards Activities in Related Fields
Work Package	**Number**	9	**Title**	Dissemination, Standardisation and Exploitation

Date of Delivery	**Contractual**	M36	**Actual**	31-March-09
Status	version 1.0		final ☒	
Nature	prototype ☐ report ☒ dissemination ☐			
Dissemination Level	public ☒ consortium ☐			

Authors (Partner)	Daniel Martin (USTUTT) Daniel Wutke (USTUTT) Dario Cerizza (Cefriel) Alessio Carenini (Cefriel) Jacek Kopecký (LFUI) Janne Saarela (Profium) David de Francisco (TID) Doug Foxvog (NUIG)			
Resp. Author	Daniel Martin (USTUTT)		**E-mail**	daniel.martin@iaas.uni-stuttgart.de
	Partner	USTUTT	**Phone**	+49 (711) 7816-485

Abstract (for dissemination)	Work Package 9 provides continuous technology watch on emerging standards in Internet Technology, Web services Technology as well as on standards and formats used in Business Process Management. In WP9, we monitor ongoing standardization activities such as W3C, OASIS, and GGF. Furthermore, the work package identifies modifications and extensions of such existing standards that will enable them to support Triple Spaces. This document reports on standards related to TripCom, as well as assesses them for potential use in our project. Furthermore, we provide a report on our own standardization activities.
Keywords	triplespace, standards, W3C, OASIS, EDIFACT, web service, rdf, wsml, sparql, sa-wsdl

Version Log			
Issue Date	**Rev No.**	**Author**	**Change**
2006-12-15	1	Daniel Martin	Document created
2007-04-22	2	Daniel Martin	Initial Outline
2007-06-12	3	Daniel Martin	Outline Refined
2007-08-06	4	Daniel Martin	Partner Assignments
2007-10-17	5	Daniel Martin	Outline Refinement and Partner Assigments as discussed in Vienna meeting
2008-02-30	6	Daniel Martin	Drafts for Standard Assesments
2008-06-22	7	Daniel Martin	Revised Texts
2008-07-03	8	Daniel Martin	Document structure revised
2009-01-26	9	Janne Saarela	Updated W3C and Semantic Web specific parts
2009-01-26	10	Alessio Carenini	Updated eHealth ontologies part
2009-02-03	11	Daniel Martin	Added Conclusion
2009-02-23	12	Daniel Martin	Finalized

Project Consortium Information

Acronym	Partner	Contact
Leopold Franzens University Innsbruck http://www.deri.at	LFUI	Prof. Dr. Dieter Fensel Digital Enterprise Research Institute (DERI) Innsbruck, Austria E-mail: dieter.fensel@deri.org
National University of Ireland, Galway http://www.deri.ie	NUIG	Dr. Laurentiu Vasiliu Digital Enterprise Research Institute (DERI) Galway, Ireland Email: laurentiu.vasiliu@deri.org
University of Stuttgart http://www.iaas.uni-stuttgart.de/	USTUTT	Prof.Dr. Frank Leymann Inst. für Architektur von Anwendungssystemen (IAAS) Stuttgart, Germany E-mail: frank.leymann@informatik.uni-stuttgart.de
Vienna university of Technology http://www.complang.tuwien.ac.at/	TUW	Prof.Dr. eva Kühn Institut für Computersprachen Vienna, Austria E-mail: eva@complang.tuwien.ac.at
Free University Berlin http://www.ag-nbi.de/	FUB	Prof. Dr.-Ing. Robert Tolksdorf AG Netzbasierte Informationssysteme Berlin, Germany E-mail : tolk@inf.fu-berlin.de
Ontotext Lab, Sirma Group Corp. http://www.ontotext.com/	ONTO	Atanas Kiryakov, Vassil Momtchev, Ontotext Lab, Sirma Group Corp. Sofia, Bulgaria E-mail: vassil.momtchev@ontotext.com
Profium OY http://www.profium.com/	Profium	Dr. Janne Saarela Profium OY Espoo, Finnland E-mail: janne.saarela@profium.com
CEFRIEL SCRL. http://www.cefriel.it/	CEFRIEL	Davide Cerri CEFRIEL SCRL. Milano, Italy E-mail: cerri@cefriel.it
Telefonica I+D http://www.tid.es/	TID	Noelia Pérez Crespo Telefonica I+D Madrid, España E-mail: npc@tid.es

Table of Contents

1 Introduction

Triple Space Communication strives towards providing the base for the next significant step of the Internet. Thus, the project inherently deals with heterogeneity, and consequently, standards play a fundamental role in the project: Triples have to be accessible wherever they are a particular install of the Triple Space middleware on a certain hosting platform must interoperate with Triple Space middleware installs on other platforms. This requires standards at various levels, for example, at the structural level for messages exchanged by the various Triple Space middleware installs, at the protocol level by the different installs to ensure quality of services like guaranteed delivery, or at the interface level to access Triple Space functionality. Work Package 9 provides continuous technology watch on emerging standards in Internet Technology, Web services Technology as well as on standards and formats used in Business Process Management. In WP9, we monitor ongoing standardization activities such as W3C, OASIS, and GGF. Furthermore, the work package identifies modifications and extensions of such existing standards that will enable them to support Triple Spaces. This document reports on standards related to TripCom, as well as assesses them for potential use in our project. Furthermore, we provide a report on our own standardization activities.

2 Assessment of Related Standards

In this chapter, we analyse standards in related areas of TripCom and check if these standards can be used as is or need to be extended in some way. We conclude each section with an evaluation giving advice on how to proceed.

2.1 Standards in eHealth

2.1.1 Electronic Health Record Standards

To address the EHR interoperability problem, several standards and technical specifications are currently under development. These specifications aim to structure and markup the clinical content for the purpose of data exchange.

ASTM Continuity of Care Record (CCR)

The Continuity of Care Record (CCR) is a standard specification being developed jointly by several standardization entities led by ASTM International[1].

The CCR is being developed and enhanced in response to a perceived need to organize and make transportable a set of basic patient information, aiming at representing the most relevant and timely facts about a patient's condition.

CCR is designed and implemented as a standard for a comprehensive data summary that aggregates data from multiple sources, such as health care records, medical legal documents, and health care encounters, to form a comprehensive overall clinical picture of a patient's current and relevant historical health care status. Briefly, these include patient and provider information, insurance information, patient's health status (e.g., allergies, medications, vital signs, diagnoses, recent procedures), recent care provided, recommendations for future care (care plan) and the reason for referral or transfer. This minimum data set will enhance the continuity of care by providing a method for communicating the most relevant information about a patient and providing both context and support for the electronic health record (EHR) through extensions [30].

EHRcom and CEN/TC 251

The CEN standard EN 13606 "Electronic Healthcare Record Communication" (EHRcom) is a comprehensive EHR standard that is currently under development at the technical committee on Health Informatics of the European Committee for Standardization (CEN/TC 251). The mission of this technical committee is to achieve compatibility and interoperability between independent health systems and to enable modularity by means of standardization. This includes requirements on health information structure to support clinical and administrative procedures, technical methods to support interoperable systems as well as requirements regarding safety, security and quality.

The EHRcom is a message-based standard for the exchange of electronic healthcare records. The standard defines an EHR information model, called the "extended architecture" since it is an extension of an earlier pre-standard. It also defines a list of machine-readable domain terms that can be used to structure EHR content, a method of specifying "distribution rules" (rules under which certain EHR content may be

[1] http://www.astm.org/

shared with other systems) and request and response messages that allow systems to exchange subsets of an EHR. ENV 13606 does not attempt to specify a complete EHR system, instead it focuses on the interfaces relevant for communication between EHR systems [13].

GEHR/openEHR

The GEHR/openEHR initiative was started in 1992 and currently is maintained by the openEHR Foundation, a non-profit organization whose aim is to promote and facilitate progress toward electronic health care records.

This EHR specification introduces an *archetype*, which is a reusable, formal expression of a distinct, domain-level concept such as "blood pressure", "physical examination", or "laboratory results", expressed in the form of constraints on data whose instances conform to some reference model [5].

The openEHR approach uses a two-level methodology to model the EHR structure[12]. In the first level, a generic reference model that is specific to the healthcare domain but still very general is developed. In the second level, healthcare and application specific concepts such as blood pressure, lab results etc. are modeled as archetypes, constraint rules that specialize the generic data structures implemented using the reference model.

Health Level Seven (HL7)

Founded in 1987, HL7 (Health Level Seven) is a non-profit, ANSI accredited Standards Developing Organization that provides standards for the exchange, management and integration of data that supports clinical patient care and the management, delivery and evaluation of healthcare services.

HL7 Message Protocols The HL7 standard is developed with the assumption that an event in the healthcare world, called the trigger event, causes the exchange of messages between applications. When an event occurs in an HL7 compliant system, an HL7 message is prepared by collecting requested data from the underlying application systems and then passed to the requester.

Up to the current Version 2.5, the scope of the HL7 standard was limited to the exchange of messages between medical information systems. Starting with Version 3.0, a document markup standard, called the Clinical Document Architecture has been proposed.

HL7 Clinical Document Architecture (CDA) Clinical Document Architecture, Release One, became an American National Standards Institute (ANSI)-approved HL7 Standard in November 2000, representing the first specification derived from the Health Level 7 (HL7) Reference Information Model (RIM). CDA, Release Two (CDA R2) [19], became an ANSI-approved HL7 Standard in May 2005. CDA is a document markup standard that specifies the structure and semantics of a clinical document (such as a discharge summary or progress note) for the purpose of exchange. A CDA document is a defined and complete information object that can include text, images, sounds, and other multimedia content. It can be transferred within a message and can exist independently, outside the transferring message. CDA documents are encoded in

Extensible Markup Language (XML), and they derive their machine processable meaning from the RIM, coupled with terminology. The CDA R2 model is richly expressive, enabling the formal representation of clinical statements (such as observations, medication administrations, and adverse events) such that they can be interpreted and acted upon by a computer. On the other hand, CDA R2 provide a mechanism for simply wrapping a non-XML document with the CDA header. It also allows creating a document with a structured header and sections containing only narrative content. The intent is to facilitate widespread adoption, while providing a mechanism for incremental semantic interoperability.

IHE Cross-Enterprise Document Sharing (XDS)

Integrating the Healthcare Enterprise (IHE) is a not-for-profit initiative founded in 1998 that does not develop standards, but selects and recommends an appropriate usage of existing standards for specific use cases, stimulating the integration of healthcare information resources [31].

Cross-enterprise Document Sharing (XDS) is an IHE specification aimed at providing a document archive for the "longitudinal", that is, life-long, cross-institutional healthcare record. XDS is document centric and content agnostic in the sense that any kind of document can be stored in an XDS archive, provided that the metadata for the document (for which XDS has a detailed specification) is available.

Evaluation

HL7 is one of the most prominent standards for exchanging clinical messages, and part of the WP8B work deals with enabling native use of this standard in order to ease the integration of existing medical systems. Moreover, the structures of the messages can provide basic concepts that can be used to design the European Patient Summary ontology. The same applies for CCR, as this standard is closely related to the Patient Summary that is developed in WP8B context.

2.1.2 Medical Coding Systems

Medical coding systems are fundamental to medical record keeping as well as to gathering and communicating public health statistics. They are used for a variety of purposes, such as recording causes of death, coding diseases and procedures and also physician billing and reimbursement.

ATC

The Anatomical Therapeutic Chemical (ATC)[2] Classification System aims to classify drugs for therapeutic utilizations according to the organ or system on which they act and/or their therapeutic and chemical characteristics. It born in 1976 under the control of the WHO Collaborating Center for Drug Statistics Methodology and is currently widely used in European countries.

[2]http://www.who.int

ICD: International Classification of Diseases

The work that led to International Classification of Diseases (ICD[3]) started in 1893. The standard is systematically updated by the World Health Organization (WHO) and is designed to describe and classify diseases and injuries. Currently, ten versions have been defined and the most widely used are the 9th and the 10th (known as ICD-9 and ICD-10, respectively). ICD-10 is a great expansion of the coding system, allowing more than 155,000 different codes (compared to 17,000 codes that are available in ICD-9) and allowing tracking many new diagnoses and procedures, including codes for diagnoses, procedures, symptoms, social circumstances, etc. which were not included in ICD-9. ICD-10 is not a superset of ICD-9, using different codes than ICD-9 does for the concepts that are included in both.

ICD is widely used at international scale for indexing healthcare record with respect to topics related to mortality and morbidity. This wide adoption is surely a result of the continuous process of enhancements of the classification to to better fit the needs of particular health care institutions in various countries.

LOINC: Logical Observation Identifiers Names and Codes

The Logical Observation Identifiers Names and Codes (LOINC)[4] data base provides a standard set of universal names and codes for identifying individual laboratory results (e.g. Hemoglobin, Serum Sodium concentration), clinical observations (e.g. Discharge Diagnosis, Diastolic blood pressure) and diagnostic study observations, (e.g. PR-interval, Cardiac echo left ventricular diameter, Chest x-ray impression).

LOINC was developed as an encoding system for identifying laboratory observations. The code names and identifiers in LOINC are intended to cover terminology related to the context of Electronic Health Records, thus enhancing the collection, processing and exchange of clinical results such as laboratory tests and clinical observations. With this respect LOINC since 1999 has been the preferred code set for laboratory test names in transactions between health care facilities of the HL7 Development Consortium.

OpenGALEN

OpenGALEN[5] is an open source medical terminology that aims to promote healthcare through stimulating the use and development of the results of the GALEN programme. Galen experience and technology work as a basis for teaching, training and services in the area of medical terminology, language, knowledge and information and in anything directly or indirectly related in the widest sense.

The GALEN Common Reference Model is a clinical terminology that can represent medical concepts. The medical concepts represented using that scheme are accessible and manipulable by computers, as well as being accessible to clinicians. OpenGALEN uses GRAIL (GALEN Representation And Integration Language) as a representation scheme to build the GALEN Common Reference Model. All this approach produces a computer-based multilingual coding system for medicine.

[3] `http://www.who.int/classifications/icd/en/`
[4] `http://www.regenstrief.org/loinc`
[5] http://www.opengalen.org/

SNOMED: Systematized Nomenclature of Medicine

SNOMED[6] [29] is a clinical classification system published and maintained by SNOMED International, a division of the College of American Pathologists (CAP). SNOMED aims to provide a common language enabling a consistent way to capture, share and aggregate data across medical disciplines and within healthcare delivery networks.

SNOMED's design is based on the premise that a detailed and specific nomenclature is essential to accurately reflect, in computer readable format, the complexity and diversity of information found in a patient record. Independent studies have demonstrated SNOMED to be the most complete reference terminology in existence today for the clinical environment.

One of the main component of SNOMED, is the SNOMED Clinical Terminology (SNOMED CT) that is a medical terminology resource that combines the content of SNOMED Reference Terminology with the United Kingdom's Clinical Terms Version 3. This joint effort, covering most aspects of clinical medicine with over 344,000 concepts, employs a reference terminology model to develop a comprehensive multilingual controlled medical vocabulary; SNOMED is available in various languages, while multilingual versions are released and updated by independent organizations.

UMLS: Unified Medical Language System

In 1986, the National Library of Medicine (NLM), began a long term research and development project to build a Unified Medical Language System (UMLS)[7]. The purpose of the UMLS is to aid the development of systems that help health professionals and researchers retrieve and integrate electronic biomedical information.

Currently, UMLS provides a unified view upon a collection of controlled vocabularies in the field of eHealth, and it works as a compendium of many controlled vocabularies in the biomedical sciences. The terms have roots in about 100 heterogeneous concept-based classification systems and medical vocabularies in 15 languages at present. It provides a mapping structure between these vocabularies and thus allows translating between the various terminology systems; it may also be viewed as a comprehensive ontology of biomedical concepts (as it is a shared and formal conceptualization). UMLS further provides facilities for natural language processing and is intended to be used mainly by developers of systems in medical informatics.

2.1.3 Evaluation

UMLS is a key standard in the context of the eHealth scenario, as it already provides an integrated view over different coding system standards. Exploiting the work already done for this standard by ontologizing meaningful subset would automatically result in obtaining a coherent ontology set of the integrated coding system standards. However, a complete ontologization of such a huge number of concepts is not straightforward. An ontology resulting from a meaningful subset of UMLS anyway enables the possibility of expressing medical terms using different coding systems, using UMLS as a bridge to convert from a coding system to coding system.

[6] http://www.snomed.org
[7] http://umlsinfo.nlm.nih.gov/

2.2 UN/EDIFACT

The United Nations Centre for Trade Facilitation and Electronic Business, (UN/CEFACT) created the United Nations Electronic Data Interchange for Administration, Commerce and Transport (UN/EDIFACT) standard, which it officially describes as "United Nations rules for Electronic Data Interchange For Administration, Commerce and Transport. The standard comprises a set of internationally agreed standards, directories and guidelines for the electronic interchange of structured data, and in particular that related to trade in goods and services between independent, computerized information systems."

EDIFACT is an international standard designed to cover EDI trading operations. It was created by the UN-JEDI group, which contained European and American representatives. This commission created the first version of the UN/EDIFACT syntax rules. The maintenance and extension of this standard is done by UN/CEFACT, under the UN Economic Commission for Europe [17].

EDIFACT has been adopted by the International Organization for Standardization (ISO) as the ISO 9735 which:

- Provides a set of syntax rules to defines structure data.
- Provides an interactive exchange protocol (I-EDI).
- Provides standard messages (allows multi-country and multi-industry exchange).

EDIFACT contains definitions for data elements and sequences of data elements structured by rules forming data. These rules are published under the Trade Interchange Directory (UNTID) following these sections:

- Part 2: UNCID: Uniform Rules of Conduct for Interchange of Trade Data by Teletransmission.
- Part 3: terminology used and a list of technical standards and recommendations.
- Part 4: general UN/EDIFACT rules and guidelines for syntax and message designing.
- Part 5: Approved UNSM specifications in message type directory (UNEDMD).
- An additional part is meant to guide UN/EDIFACT database users.

EDIFACT standardizes the messages to be exchanged between EDI partners when accomplishing electronic business transactions. This standard specifies neither communication platforms nor communication protocols needed to perform these exchanges, so any concrete implementation can be used beneath the message format specified by EDIFACT.

EDIFACT is a flexible and extensible standard in order to cover all possible restrictions and business domains needed by users. To maintain a coherent syntax and a controlled extension mechanism, EDIFACT defines some extension rules and message designing guidelines [14].

The EDIFACT standard changes twice a year, with message types and message type components being added, modified, and sometimes removed.

2.2.1 Major EDIFACT subsets

Due to the vast number of segments defined in EDIFACT, most industries and organizations using the system have developed specialized standards focused in their business domain, based on EDIFACT.

The result is what is known as subsets or sub-standards of EDIFACT. These kinds of standards are only suitable for special application areas; in contrast to the original idea of a common standard for electronic data interchange.

Some industries that are heavy users of EDI have attempted to combat the proliferation of customized EDI subsets and have developed standard subsets for the most commonly used EDI messages in their communities. Work Package 7 selected several of these subsets – typically called Implementation Guidelines – to focus on in constructing their ontologies. The selection process was described in Deliverable 7.1 [9]. These industry sub-standards are described below.

2.2.2 EANCOM - GS1

GS1 is a global organization dedicated to the design and implementation of global standards and solutions to improve the efficiency and visibility of supply and demand chains globally and across sectors [24].

GS1's main activity is the development of the GS1 System: a series of standards designed to improve supply chain management. GS1's interests are represented at meetings with official bodies (such as the United Nations and the European Commission), international associations and other institutions.

Much of the development is initiated by its member organizations. Member organizations are usually national associations which provide tools and support that enable their own member companies to manage their supply chains and trade processes far more efficiently. One of the most important tasks for the member organizations is providing training and support for eCom - EDI (Electronic Data Interchange).

GS1-eCom provides two global standards for electronic business messaging that allow rapid, efficient and accurate automatic electronic transmission of agreed business data between trading partners:

- **GS1 XML** [25] for XML message formats
- **GS1 EANCOM** [23] for traditional textual message formats

EANCOM is fully based on the UN/EDIFACT (United Nations Electronic Data Interchange for Administration, Commerce and Transport) standard. Since EDIFACT message formats are often complex, users may easily misunderstand the principles and original intentions of message's designers. The EANCOM guideline is a subset of EDIFACT, providing clear definitions, explanations and examples to allow trading partners to exchange commercial documents in a simple, accurate, and cost effective manner.

EANCOM also incorporates the GS1 standards of physical identification of trade items, logistics units and the Global Location Numbers identifying the trading partners into the electronic messages. This allows integrating the physical flow of goods with related information sent by electronic means.

The main benefits of EANCOM are [23]:

- It is a standard totally compatible with the United Nations EDIFACT.

- International and multi-sectorial application.
- It supports ECR (Effective Consumer Response) applications.
- It reduces the inherent complexity of EDIFACT by using explanations and examples.
- It is translated into 21 different languages.
- It is used by around 30,000 companies in all over the world.
- It supports the commercial globalization.

2.2.3 EDIFICAS

The French organization EDIFICAS Europe (EDI in the Finance, Information, Cost, Accounting/Auditing and Social areas) [11] is in charge of developing messages related with financial information.

EDIFICAS currently relies on the EDIFACT standard while preparing migration towards a new XML standard. It is employed in order to:

- Identify common patterns of data flow as related to accounting, auditing, and related information between enterprises.
- Consider and advise on the practical implication of EDI as regards accounting and auditing, with particular reference to legal, telecommunications, security, and general internal control requirements.
- Promote the use of EDI in all operations relevant to e-accountancy, using UN/EDIFACT standards or XML recommendations as the means to manage these data flows.
- Initiate proposal of new messages or changes to existing messages to international standardization bodies; to liaise with the other sectorial groups on proposal of UML model.
- Ensure effective communication with other sectorial bodies groups such as retail, construction, automotive, transport, insurance, public bodies, banks, with foreign accounting professional institutions and with international groupings.

2.2.4 EDIFICE

EDIFICE [16] is the European B2B forum for the Electronics Industry, open for companies with interests in computing, electronics, and telecommunications. EDIFICE is also the European RosettaNet User Group. Representatives of over 50 member companies meet quarterly for plenary conferences at which the task groups present the results of their work. These task groups currently include:

- Automatic Data Capture (ADC).
- Forecast and Inventory Management (FIM).
- SAP Integration.

- Supplier/Distributor (S&D).
- Technical Support and Quality Assurance (TSQ).
- Portals and MarketPlaces.

The EDIFICE Message Implementation Guidelines (MIGs) are a subset of the EDIFACT standard, based on Directory D.97A, and describe the EDI messages that occur specifically in business transactions in the electronics industry.

2.2.5 ETIS

ETIS (European Telecommunications Information Service) is a membership-based organization which brings together the major telecommunications providers in Europe on key information and communication technology issues [15].

The mission of ETIS is to enable its members to improve their business performance by personal exchange of information on using ICT effectively.

ETIS works together with other industry bodies in the development and harmonization of work in architectures, particularly those related to IT in telecommunications.

ETIS maintains the following six working groups:

- Information Security
- Electronic Billing
- Enterprise Architecture
- Customer Self-Service
- Business Intelligence & Data Warehouse
- Procurement and Vendor Management
- TeBIT survey
- CIO Executive Forum

Two of these working groups were of interest for Work Package 7:

- **International Settlements**. ETIS International Settlements Working Group (discontinued) was a group of experts from different telecommunications operators that dealt with the developments and implementation of standards for electronic data interchange of accounts, invoices, and settlement statements for intercarrier billing and accounting. This development work was done in close cooperation with the ITU-T (Standardization Sector of International Telecommunication Union). The development result of ETIS IS WG can be found in the relevant Recommendation of ITU-T. On basis of this recommendation telecom operators are able to exchange traffic accounts, invoices and settlement data with each other by electronic means.

- **Electronic Billing Group**. In recognition of the growing interest in electronic commerce, and of billing as an integral part of this, ETIS set up a working group to develop and promote the use of a standard subset of the UN EDIFACT INVOIC message. This subset was completed some years ago and is now in everyday use within several ETIS member companies, and is in a trial phase within other companies. In 2003 the group finished an XML standard for the telecommunications industry in addition to the existing EDIFACT messages. The XML schema and related software for implementation is available for telecommunications operators who can now choose if they want their invoices in EDIFACT or XML. The ETIS EDIFACT and XML standard and solution are the only standard on the market, there are currently no other developments in XML to define telecommunications invoices for end-customers. All documentation in EDIFACT format are available for free and the documentation in XML format requires a licence for non-members.

2.3 Web Service related standards

2.3.1 SOAP

SOAP [28] defines a standardized XML-based message format, a set of rules that describe how services should process a message, a mechanism to bind SOAP messages to different network transport protocols, and a way to encode non-XML information in SOAP messages through SOAP attachments [34]. SOAP messages (envelopes) are transferred from an origin (initial sender) to a destination (ultimate receiver) via potentially multiple SOAP processing nodes (intermediaries). Messages consist of two parts: the *header* and the *body*. While the body is directed to the ultimate receiver of a message and contains the actual message payload, the headers are primarily directed at the intermediaries. Certain SOAP headers (roles) define how a message is to be processed by the SOAP nodes along the message path. If a node chooses to implement a role which is defined in the header part of a message, it can process the message according to its role, remove the header and forward the message on the message path. SOAP bindings allow SOAP messages to be transferred over different network protocols. They define a serialization of the SOAP infoset in such a way, that it can be transferred to the next SOAP node in the message path. Along the message path multiple different serializations are possible.

Evaluation

This standard is highly related to the work going in Work Package 4, “Web service integration in TripCom”. SOAP is one of the central Web service specifications. Part of the work of WP4 is to enable Triple Space as a transport mechanism for Web services, thus communicating SOAP envelopes over the Triple Space API. This also includes the definition of a bidirectional mapping of SOAP envelopes to RDF, the triplespaces’ native storage format.

2.3.2 WSDL

Web Services Description Language (WSDL) is an XML format for describing network services as a set of endpoints operating on messages containing either document-oriented or procedure-oriented information [6]. There are currently two versions of WSDL in widespread use: WSDL 1.1 [6] and WSDL 2.0 [32]. Version 1.1 was prepared by a group of companies to capture a rough industrial consensus, which served as the input to standardization in W3C, resulting in version 2.0. In this document, we use the terminology of version 2.0.

An XML document specifying a Web service in WSDL defines services as collections of network endpoints. In WSDL, the abstract definition of endpoints and messages is separated from their concrete network deployment or data format bindings. This allows the reuse of abstract definitions: schemas, which are abstract descriptions of the data being exchanged, and interfaces which are abstract collections of operations. The concrete protocol and data format specifications constitute a reusable binding. An endpoint is defined by associating a network address with a binding, and a collection of endpoints defines a service.

WSDL is extensible to allow description of endpoints and their messages regardless of what message formats or network protocols are used to communicate. WSDL can be used in conjunction with SOAP 1.1 and 1.2 and with HTTP. Messages are described using XML schema as the preferred type definition language, other languages can be used via extensibility. Operations refer to input and output messages and describe the available communication actions of the service. Communication with a web service can be asynchronous or synchronous. Asynchronous communication is realized with one-way (service receives a message) and notification (service sends a message) operations. Synchronous communication is realized with Request-response (service receives a message and sends a correlated message) and Solicit-response (service sends a message and receives a correlated message) operations. Synchronous operations also support remote-procedure-call bindings and can reflect the original function signature with possibly several parameters.

Implementing Web services with WSDL requires to also adopting standards such as SOAP, HTTP, XML, XML Schema and (optionally) UDDI. As each of these standards exists in several versions, some of which are specified in ambiguous ways, it was necessary to define a combination of the necessary standards in specific versions, which was achieved in the Web Services Interoperability Organization (WS-I). The WS-I Basic Profile 1.0 defines a set of specifications at specific version levels to ensure interoperability [3]. It defines SOAP 1.1, WSDL 1.1, UDDI 2.0, XML 1.0, XML Schema and HTTP 1.1 and resolves more than 200 interoperability issues. This was fundamental for the adoption of WSDL by the industry.

Along with the core XML language WSDL, the W3C has also produced a WSDL RDF Mapping [33], providing an OWL ontology for WSDL data. The existence of this RDF mapping effectively adds WSDL documents to the Semantic Web.

Evaluation

WSDL is the fundamental technical specification for the application of Web services in the industry. Any semantic annotations of Web services proposed and used by TripCom must be compatible with both versions of WSDL, and in particular they must

consider the fundamental concepts of WSDL such as synchronous and asynchronous communication with single input and output messages.

WSDL is the key standard for two important tasks within Work Package 4:

WSDL binding As mentioned in the evaluation of SOAP before, a key deliverable of Work Package 4 is the integration of Web service technology with triplespace. This is achieved through the definition of a WSDL binding for Triple Space, that allows to define TripCom as the transport mechanism for the invocation of the described service.

WSDL RDF mapping Another scenario in WP4 is the use of triplespace as a registry for (semantic) Web services. As Web service interfaces are described using WSDL and the native data format of Triple Space is RDF, two-way mapping between the two systems is necessary and should be built on the W3C WSDL 2.0 RDF mapping.

2.3.3 SAWSDL

Semantic Annotations in WSDL and XML Schema (SAWSDL) [27] is a W3C specification that defines how to add semantic annotations to WSDL and XML Schema descriptions. The purpose of SAWSDL is to link WSDL with semantic descriptions. It defines extension attributes that can be applied to elements in both WSDL and XML Schema in order to annotate WSDL interfaces, operations and their input and output messages. SAWSDL is the first step towards standardization in the area of Semantic Web Services.

Semantic annotations in WSDL and XML Schema can be used for example for these purposes:

- associating WSDL interfaces with some taxonomical categories to help semantic Web service discovery,
- describing the purpose or applicability of WSDL operations to help discovery or composition,
- linking and mapping inputs, outputs and faults of WSDL operations to semantic concepts to help facilitate invocation, mediation and service discovery and composition.

While the semantic annotations are used to point to taxonomies, ontologies or mappings, SAWSDL is independent of any particular ontology language or mapping language. The mechanism only requires that the concepts in the semantic models can be identified with URIs.

SAWSDL defines two kinds of annotations: semantic model references from elements in WSDL or XML Schema to concepts in a semantic model (usually an ontology or taxonomy), and data lifting and lowering mappings between XML and semantic models. The first kind of SAWSDL annotation is an attribute called `modelReference`. The value of the attribute is a list of URIs that reference concepts in a semantic model. SAWSDL defines how model references can be used on WSDL interfaces, operations, faults, and on XML Schema element declarations or type definitions. In general, model

references can have many uses, and indeed, SAWSDL does not limit the applicability of the attribute.

Since Web services are described in terms of XML messages, and semantic models usually deal with knowledge representation languages such as RDF, it is necessary to add data mappings alongside schema annotations. For data lifting (from XML to a semantic representation) and lowering (from a semantic representation to XML) mappings, SAWSDL defines two extension attributes: `liftingSchemaMapping` and `loweringSchemaMapping`. These attributes are used to point from a schema element declaration or type definition to a mapping that specifies (in any suitable mapping language, e.g. XSLT) how data is lifted from XML to the semantic level or lowered back. Mappings can be used for example to support invocation of a Web service from a client that works natively with semantic data.

Similarly to WSDL 2.0, SAWSDL also provides an RDF mapping – a small ontology that captures the SAWSDL annotations – which can readily be combined with the WSDL RDF mapping.

Evaluation

TripCom does not work on semantic automation of the use of Web services, which would readily be supported by SAWSDL. Nevertheless, SAWSDL acts as a suitable form for linking Web service descriptions to semantics in an RDF form, which can then be stored in and communicated through triple spaces. Such a combination of semantic spaces and Web service automation is being investigated in a new project SOA4All.[8]

2.3.4 WS-Addressing

Endpoint references as defined in WS-Addressing [20] are a general-purpose way of encoding the addressing information needed to reach a service endpoint (i.e., any potential source or destination of Web service messages) at runtime [34].

Endpoint references consist of runtime interaction information which enables routing message to the endpoint and associated metadata. A set of message headers is defined to allow messages to be directed to service endpoints and to support bidirectional and asynchronous interaction. Mandatory `To` and `Action` headers define the URI of the target endpoint and the intent or semantics of the message. Optional `ReplyTo` and `FaultTo` headers contain the endpoints which should receive response and fault messages and are only mandatory in the case of request-reply interaction. Correlation of messages in complex message exchanges can be achieved via `MessageId` and `RelatesTo` headers. The optional `From` header encodes the source endpoint from which a message originates. A SOAP binding [21] describes how endpoint references are mapped to SOAP messages.

Evaluation

The SOAP binding for Triple Space (see Section 3.6) makes heavy use of the SOAP message headers defined by WS-Addressing for addressing of communication partners, message identification and message correlation; a detailed description is provided as part of the Triple Space Web service binding defined in Chapter 9 of D4.4.

[8] `http://www.soa4all.eu/`

2.3.5 WS-Notification

WS-Notification [18] is a family of related specifications (WS-BaseNotification, WS-BrokeredNotification, WS-Topics) that define a standard Web services approach to notification using a topic-based publish/subscribe pattern, i.e. they standardize the way Web services interact using Notifications or Events. Thus a Web service may disseminate information to a set of other Web services, without having to have prior knowledge of these other Web services. The standards define message exchanges to be implemented by service providers that wish to participate in notifications, standard message exchanges for a notification broker service provider (allowing publication of messages from entities that are not themselves service providers), operational requirements expected of service providers and requestors that participate in notifications, and an XML model that describes topics.

Evaluation

Triple Space provides – among other things – a platform for asynchronous communication based on asynchronous notification; this motivates an analysis, how WS-Notification-compliant notification can be realized using Triple Space. Subscriber and notification consumer could be implemented as Triple Space external entities; the notification producer is located at the Triple Space kernels and exposes the necessary operations for subscription. Since data in Triple Space is always in RDF-format, filter expressions can be written using SPARQL.

Following these ideas, a more thorough analysis of WS Notification and its application to Triplespace can be found in D4.4.

2.3.6 WS-Security (SOAP Message Security 1.0)

WS-Security [26] provides a common syntax and processing model for carrying security information in SOAP envelopes, enabling the integration of existing enterprise and internet security mechanisms. WS-Security (and SOAP) is the base standard for all further Web Service security standards. Other major building blocks are XML Encryption [22] and XML Signatures [4]. WS-Security describes enhancements to SOAP messaging to provide quality of protection through *message integrity* and *message confidentiality*. The specification defines how to attach and include security tokens within SOAP messages. A mechanism is also provided for specifying binary encoded security tokens (e.g. X.509 certificates). These mechanisms can be used independently or in combination to accommodate a wide variety of security models and encryption technologies. WS-Security provides a general-purpose mechanism for associating *security tokens* with messages. No specific type of security token is required by WS-Security. It is designed to be extensible (e.g., support multiple security token formats). For example, a requester might provide proof of identity and proof that they have a particular business certification. *Message integrity* is provided by leveraging XML Signature in conjunction with security tokens (which may contain or imply key data) to ensure that messages are transmitted without modifications. The integrity mechanisms are designed to support multiple signatures, potentially by multiple actors, and to be extensible to support additional signature formats. The signatures may reference (i.e., point to) a security token. Similarly, *message confidentiality* is provided by leveraging XML Encryption in conjunction with security tokens to keep portions of SOAP

messages confidential. The encryption mechanisms are designed to support additional encryption technologies, processes, and operations by multiple actors. The encryption may also reference a security token. Finally, WS-Security describes a mechanism for encoding binary security tokens. Specifically, the specification describes how to encode X.509 certificates and Kerberos tickets as well as how to include opaque encrypted keys. It also includes extensibility mechanisms that can be used to further describe the characteristics of the security tokens that are included with a message.

Evaluation

Web Service Security standards can provide a general solution to the security requirements in TripCom. It is possible to use WS-Trust, WS-SecureConversation, and WS-Security to map a TLS handshake to Web Service standards. This solution has the advantage of reusing the structure of the TLS security handshake, which is very well established, rather than requiring the design of a new security handshake. Because it completely maps TLS logic and functionality to Web service security standards, this solution can be used to replace the use of TLS in the Triple Space without requiring any additional modification. In particular, it is compatible with keeping the TS API as currently specified, i.e., with the assertions and cookies passed as additional parameters in the payload. As a result, the application of security-related Web service standards is discussed further in TripCom deliverable D4.4.

2.3.7 WS-Metadata Exchange

Web Services Metadata Exchange defines a mechanism to retrieve metadata from a Web service. It can be used to discover WS-Policies, WSDL, and other metadata about a Web service, of which only its end-point address is known. Web services have several types of metadata. WSDL documents describe the functional interface; WS-Policy can be used to describe non-functional requirements or capabilities. To bootstrap communication with Web services, WS-Metadata Exchange [2] provides two request-response interactions to enable service consumers to retrieve the needed metadata by the service provider itself. When the type of the metadata is known (e.g. WSDL or WS-Policy), a requester may indicate what type of metadata should be returned. Where additional types of metadata are being used, a requestor may retrieve all available metadata. The interactions defined are intended for the retrieval of metadata only.

Evaluation

As there is very little acceptance of this Web service standard, we do not investigate its further application within TripCom.

2.3.8 WS-BPEL

The Business Process Execution Language (BPEL4WS, WS-BPEL or BPEL for short) is a standard to describe Web service flows. The Business Process Execution Language is the de facto standard to describe Web service flows. It was developed out of two competing languages for Web Service orchestration, namely WSFL and XLANG. BPEL describes the process from the view of one single participant of a business process

that communicates with other partners (Web services) to achieve a business goal. A BPEL process not only uses Web services but is also presented to the outside as a Web service (via a WSDL-file). The BPEL specification not only allows for describing executable processes, it explicitly introduces an abstraction of executable processes, so called abstract BPEL. Via abstract BPEL the internal procedures of a process can be hidden. Thus it is possible to only describe its external behaviour for business protocol modelling. "BPEL is layered on top of several XML specifications: WSDL 1.1, XML Schema 1.0, and XPath1.0. WSDL messages and XML Schema type definitions provide the data model used by BPEL processes. XPath provides support for data manipulation. All external resources and partners are represented as WSDL services. BPEL provides extensibility to accommodate future versions of these standards, specifically the XPath and related standards used in XML computation" [1]. The dependency on WSDL is the major drawback of BPEL, because it means that Web services used are restricted to a special kind of interface (WSDL) and not to the business goal they serve for.

Evaluation

WS-BPEL has proven to have a large application area within Triple Space. Deliverable D4.4 identifies four different usescases:

1. Orchestration of Triple Space API methods
2. Triple Space as a storage mechanism for process models
3. Invocation of Web services over Triple Space
4. Decentralized navigation of BPEL processes based on Triple Space

2.4 Semantic (Web) related standards

2.4.1 RDF/ RDFS

Resource Description Framework (RDF) and RDF Schemas (RDFS) are technologies that have reached recommendation status at the World Wide Web Consortium (W3C). They are thus considered mature for implementation and the marketplace provides tools for creating, processing and validating such metadata descriptions.

RDF establishes a formal data model of 3-tuples (subject,predicate,object) that are equal to binary relationships encoded often as `predicate(subject,object)`. RDF schemas introduces a type system with some constraints and some inference rules which enable RDFS capable software systems to check the validity of RDF data models and infer implicit metadata descriptions from the explicit descriptions with the help of RDF schemas.

2.4.2 SPARQL

SPARQL technology has been developed at the W3C by the Data Access Working Group (DAWG) and the technology has reached recommendation status in January 2008.

SPARQL consists of multiple specifications, namely one that defines a query language syntax and semantics, one that defines XML encoded result sets and one that defines a protocol for interacting with SPARQL services.

SPARQL is positioned similarly to SQL in that SPARQL is a declarative query language that operates on the RDF data model and enables application to make use of a multitude of SPARQL services. (SQL is a declarative query language that operates on the relational data model.)

W3C has presented a candidate draft charter for continued work on SPARQL. The charter explicitly states the syntax and the semantics of the existing SPARQL Recommendation are not going to be changed. The future work is going to address, depending on the participating companies' interests, data manipulation operations, aggregate functions and interoperability mechanisms.

Evaluation

TripCom project has made use of SPARQL query language and explored its subsets and supersets as a mechanism to enable more efficient and/or complete reasoning in a query service. The project partners are thus well positioned to further evolve SPARQL to a direction which solves problems seen in practical use of the language.

3 Standardization Activities

3.1 Standardization Committees

3.1.1 W3C

Founded in 1994 by Tim Berners-Lee and others, the World Wide Web Consortium develops interoperable technologies. To initiate a new standardization activity, W3C members submit a so called Submission (mostly specifications, guidelines, software, and tools). Then the consortium publishes a note and provides the infrastructure to enable collaborative work within a working group. The publishing process defined by W3C consists of the four maturity levels Working Draft, Candidate Recommendation, Proposed Recommendation, and Recommendation. The last can be considered a de-facto standard specification. Best established standards: (X)HTML/CSS, XML, URI/URL, SOAP.

3.1.2 OASIS

The Organization for the Advancement of Structured Information Standards (OASIS), founded in 1993, is a global consortium that drives the development and adoption of e-business and web service standards. OASIS focusses on the following domains: Web services, e-commerce, security, law & government, supply chain, computing management, application focus, document-centric, XML processing, conformance/interoperability, and industry domains. The best established standards are: OpenDocument, UDDI, WSRF, WSS, Relax NG, (ebXML).

3.1.3 OMG

The Object Management Group (OMG), founded in 1989, aims to define standards for object oriented and distributed systems. The charter of the organisation includes the establishment of industry guidelines to provide a common framework for the development of applications. Primary objectives are the reusability, portability and interoperability in distributed, heterogeneous environments. In 2005, the Business Process Management Initiative (BPMI) and the OMG merged their Business Process Management (BPM) activities. Together, they will continue BPMI's and OMG's work and focus on Business Process Management. Best established standards: UML, MOF, MDA, CORBA.

3.2 Identification of TripCom Results Suitable for Standardization

Of the standards and technologies analysed above, WSDL provides the most benefit when adapted to be used with Triple Space. WSDL is of great significance in the stack of Web service specifications since is one of the basic Web service standards that many others build upon. As a consequence, integration with higher level standards such as *WS Security*, *WS Trust*, *WS Coordination* or *WS Business Activity* can be implemented with little effort or in some cases is even provided out-of-the-box. The

WSDL specification contains communication protocol bindings to support different technologies and protocols that transport messages between Web service implementations. The goal is to define the bindings of Triple Space communication protocol in WSDL that should enable Triple Space clients to communicate and invoke Web services. It will enable Web services based on current standards to be invoked with Triple Space-based communication protocol. It concerns with finding a way to have Triple Space bindings in the service description, secondly it requires binding the communication protocol. There are two major steps involved in binding the Triple Space communication protocol with Web services. First is to provide protocol binding information in the description language of Web services which in case means providing Triple Space bindings in WSDL. Second is to provide communication protocol bindings the Triple Space communication via an RDF representation of SOAP.

Web service bindings describe how messages are exchanged between Web services. They provide a description of the format and encoding of the request message, the transport protocol used to transmit the request message, and a method to address a specific service provider. The most prominent Web service binding in use today is the SOAP/HTTP binding. Using HTTP as a transport for Web service communication has several drawbacks: (i) HTTP is based on a synchronous communication model, requiring both sender and receiver to be available at the same time. (ii) HTTP communication is based on the notion of a message exchange between a client and a server, which has to be addressed explicitly by a URI. Thus, both one-to-many communication and associative addressing are not possible on top of the HTTP transfer mechanism.

The resulting specification and the standardization process we employ is reported in Section 3.6.

3.3 Dissemination of Ontologies Developed

3.3.1 EDIFACT Ontologies

The dissemination of the ontologies derived from the industrial EDIFACT sub-standards is starting. The intended targets of this distribution is the bodies whose standards WP 7 has used in creating the ontologies. These include the EDIFACT sub-standard bodies and the organization responsible for the external ontology (OpenCyc) that WP 7 used.

The ontologies created from the EDIFACT system have been semi-mechanically converted into CycL [8] and submitted for inclusion in OpenCyc in order to encourage the reuse of the developed ontologies. Additional effort was put into grouping sets of type definitions into meta-classes – which Cyc uses, but which WP 7 avoided using for its WSML instantiation of the ontologies. OpenCyc is managed by The Cyc Foundation[1] with a charter to "grow the ontology and knowledge base ... with the help of volunteers". They have welcomed the submission from TripCom stating, "We'd be happy to work on putting your ontology through [their review] process and expect that it would be a great addition.".

The Global Standards Management Process (GSMP) is a worldwide collaborative forum which develops and manages the EANCOM EDIFACT standard and a related

[1] http://cycfoundation.org

GS1 XML standard. They state on their website that "[w]hile there is a strong preference that GS1 standards be usable without requiring the payment of royalties, [they] also seek to have standards that are "best of breed". TripCom presented the ontologies created in WP7, the encoding method, and the mappings of the standard to the ontologies as an invited talk at a GS1 "Global Standards Management" meeting.

EDIFICE is a "European B2B forum for the Electronics Industry", which, among other tasks, publishes UN/EDIFACT message guidelines approved by the electronics industry. EDIFICE also contributes to EDI over the Internet and XML versions of EDI, and RosettaNet. EDIFICE has been contacted and they are "looking into" the possibility of using our work.

ETIS (European Telecommunication Information Service) is a membership-based organization of major European telecommunication providers, which, among other missions developed and promotes a standard subset of a few EDIFACT messages (INVOIC, ORDERS, PRICAT) for electronic billing. We are investigating an interaction with this standards body.

The United Nations Electronic Data Interchange for Administration, Commerce and Transport (EDIFACT) group deals with many more messages than WP 7 worked on, and provides options in those messages not used in the subsets we worked on. However, they might be interested in the techniques we used and the method we have for encoding Message Implementation Guidelines (MIGs).

The latest version of the EDIFACT ontologies can be downloaded from `http://tripcom.org/ontologies/`.

3.3.2 Triple Space Ontology

For the purpose of an internal representation of the various components and data that make up triple spaces, the project has developed the Triple Space Ontology. This ontology consists of three parts:

- *Core ontology*[2] describes spaces and their hierarchy, repositories, kernels, data graphs and agents that act on the data, along with an access log.
- *Distribution extension*[3] deals with the relationships between kernels, repositories and spaces in a distributed triplespace deployment.
- *Security ontology*[4] captures the access policies for triple spaces, based on a subset of the policy standard XACML [10].

The core ontology and the distribution extension are very specific and internal to triplespaces, and therefore they can only be input to standardization together with a triplespace architecture specification. Such a standard may be investigated in the future when triplespaces start to be commercially successful.

The security ontology contains triplespace-specific terms alongside more general authorization concepts for XACML terms and authentication terms for user identification and authority trust. XACML is a finished OASIS standard; its ontologization is a straightforward process which so far has not required standardization. On the other hand, the authentication and trust model of Triple Space can be applicable outside

[2] `http://www.tripcom.org/ontologies/tsonto-core.php`
[3] `http://www.tripcom.org/ontologies/tsonto-distr.php`
[4] `http://www.tripcom.org/ontologies/ts-sec-onto.php`

of XACML as well, therefore the model and the ontology should be considered in any future efforts around standardization in this area.

In summary, we are not aware of any immediate opportunities for standardization around the Triple Space Ontology; however, we expect such opportunities to arise in the future and we plan to use the Triple Space Ontology as input where appropriate.

3.3.3 eHealth Ontology

The EPS ontologies [5] have the mission of formalizing the domain concepts, structures and relations, and of associating them with different standardized vocabularies used by various healthcare organizations in different countries (as depicted in fig. 3.1).

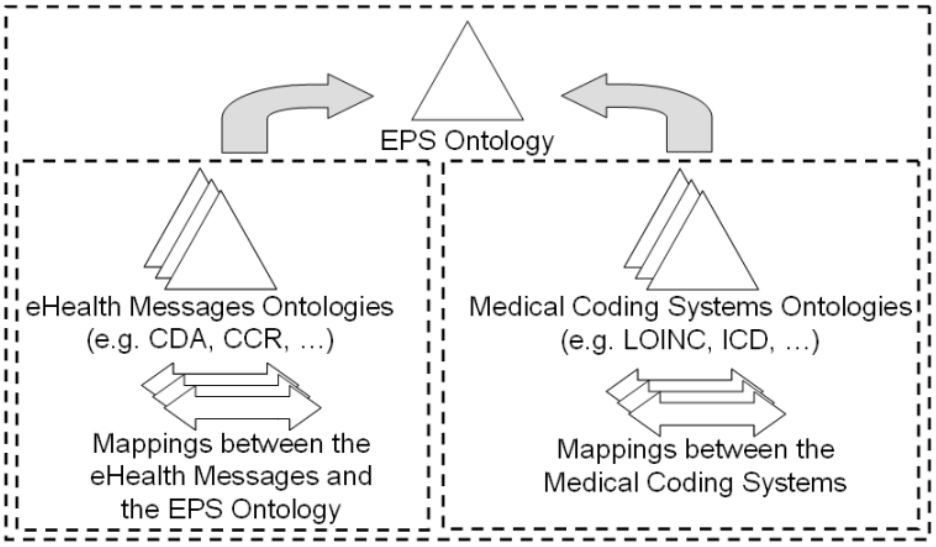

Figure 3.1: The ontologies developed for the EPS scenario.

At the top level there is the EPS ontology, an upper ontology which provides a skeleton into which the more detailed aspects of the problem domain are incorporated. To model the EPS ontology, we analyzed the structure of message standards (such as CDA and CCR) and medical coding systems (*LOINC*, *ICD*, *UMLS*, ...) and organized the portions we used in order to model an unified ontology. The EPS ontology includes demographic concepts such as patient and care provider information, insurance information, and medical terms related to patient health status such as medications, allergies, vital signs, and diagnoses. In order to support safety and continuity of care when the patient is cared for by another provider, the EPS Ontology also models concepts about medical procedures and care provided, as well as recommendations for future care and reasons for referral and transfer. These concepts are further broken down in more detailed descriptions and are more precisely described by properties and relations with other concepts on various levels.

At the second level of the EPS ontologies there are eHealth message ontologies and medical coding systems ontologies. Regarding messages, as we expect the data of many health records to be delivered in XML-based CDA or CCR formats, the EPS ontology framework contains ontologies that model these standards. Furthermore, the EPS is configured with a set of semantic mappings created by using an abstract

[5] `http://tripcom.org/ontologies`

mapping language in order to link CCR and CDA ontologies to the EPS upper ontology. These ontologies and mappings, other than enabling merging the content of messages expressed in different standards, can also be used to convert a meaningful set of information from CCR messages to CDA messages (and back).

As for the medical content of the messages, in the medical domain there are many sets of specialized vocabularies and it is in general highly complex to select the relevant ones and to correlate them. In consequence we make use of the laborious work already done in course of the Unified Medical Language System (UMLS).[6] In order to match the EPS ontology and the coding systems ontologies, meaningful subsets of different vocabularies from UMLS (NCI, LOINC, UMLS Metathesaurus, RXNORM, ICD-9 and ICD-10) were translated to ontologies too. For a full implementation of a European Patient Summary, these ontologies would have to be expanded to cover the portions of these standard vocabularies that might appear in patient records.

3.3.4 EAI Ontology

The goal of the Digital Asset Management use case [7] is to define a prototypical digital content marketplace implementation. The scenario is conceived as a realistic solution for a real business problem, in which the functionality of the Triple Space infrastructure is tested. For the semantic description of the use case, an EAI ontology has been developed [7].

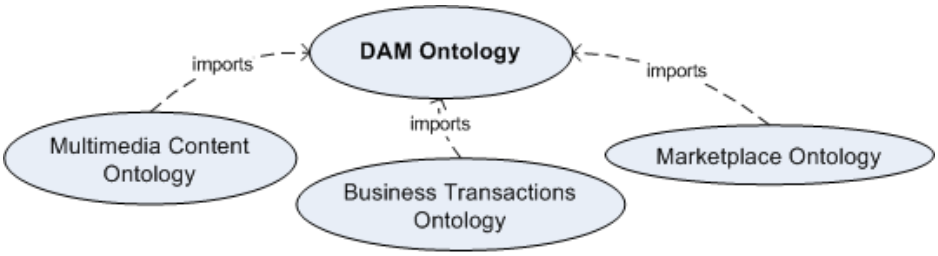

Figure 3.2: EAI ontologies for the DAM implementation

An important effort of reusing existing knowledge from the EDIFACT ontologies developed by the WP7. With the aim of presenting a clear ontological structure, the ontology could be presented as three conceptual sub-ontologies which are implemented together (in one file, as shown in Figure 3.2), which cover the different domains of the use case, providing a common vocabulary to all actors.

The Multimedia Content Ontology defines a taxonomy of assets, media works, awards, agents and any concepts related to multimedia contents. Its aim is the automatic classification of individual assets and the retrieval of inferred (not explicitly stated) information related to these assets. Since many of the concepts included in this ontology are not specific to the DAM use case, several were imported from other existing ontologies such as OpenCyc [8], which is one of the largest general knowledge ontologies.

The Business Transactions Ontology has the aim of solving the heterogeneity problems that arise when performing a business transaction, by standardizing the format of

[6]UMLS: www.nlm.nih.gov/research/umls/
[7]See http://www.tripcom.org/ontologies/dam.php
[8]http://www.opencyc.org/

the messages sent among the agents involved. To do so, the EDIFACT ontologies were kept as a reference during the design process, since most of the knowledge formalized by the EDIFACT ontologies is widely used by industry.

The Marketplace Ontology defines the knowledge related to the marketplace logic, such as film actors and auction related knowledge. Its objective is to facilitate the interoperability between communication partners and the semi-automatic handling of auctions (e.g: automatically starting an auction when a service is terminated by a content provider).

3.4 Standardization Process

To give a concrete example of how standardization works, we describe the standardization process of the W3C. Other standardization bodies generally have similar processes differing mainly in formalities. One significant difference between standardization body processes is the level of consensus (within the working groups and with external parties) that is sought in the process of standardization.

In the W3C, standards are called *Recommendation* and are developed by *Working Groups*. The process around Working Groups can be summarized as follows:

1. a Working Group is suggested by W3C *Members* (organization who have formally joined the Consortium) to work on a specific piece of technology, usually with some external specification submitted as an input;

2. the W3C *Team* creates a charter (possibly drafted in the WG proposal),

3. the charter is approved (or rejected) by representatives of the consortium members, the so-called *Advisory Committee*,

4. the Working Group is started, with well-defined timeline and deliverables,

5. and finally, after it finishes its work or its time (modulo possible extensions), the group is ended.

Working Groups can have normative and non-normative deliverables. The normative ones are said to be "on the Recommendation track", with the goal to get wide consensus and so to specify the standard technology. The steps of the Recommendation track are the following:

1. *First Public Working Draft* (FPWD) is the first time a Working Group formally publishes a document; this is announced in relevant channels such as mailing lists.

2. *Working Draft* (WD) is a formal update of a document. To demonstrate progress, working groups are required to update their deliverables at least every 3 months (a so-called heart-beat requirement).

3. *Last Call Working Draft* (LCWD) is an update that signifies that the Working Group is done with the document. At this stage, the group requests formal reviews from the public and selected relevant parties in particular, and there are stringent requirements for handling the feedback received for an LCWD. In case

any substantial changes are made to the document as result of the feedback, the next publication must again be a Working Draft or a Last Call Working Draft.

After all the Last Call feedback is handled, the WG submits the document to the W3C for publication as a Candidate Recommendation, and the WG must show that the document covers all the requirements, and that all feedback has been handled appropriately, with strong emphasis on reaching consensus, both within the group and with any external commenters.

4. *Candidate Recommendation* (CR) is a call for implementations. The working group enters CR with a set of *exit criteria*, specifying what and how many implementations are expected and how they should interoperate. In conjunction with the implementors, the WG performs any testing necessary to gather evidence of implementability of every feature of the specification.

   Implementors often uncover problems with specifications, and in case such reports lead to substantial changes, the specification must go back to Working Draft or Last Call.

5. *Proposed Recommendation* (PR) is published after the Working Group shows evidence of implementations of a stable Candidate Recommendation. A Proposed Recommendation is given to the Advisory Committee for formal approval, and when approved, the document becomes final.

6. *Recommendation* is a W3C standard, a stable technology specification. It can undergo maintenance, with errata etc.

This process, with specified minimum durations of the various review stages, takes no less than one year, commonly more. It is a process strongly focused on achieving wide consensus and a high level of stability and maturity of specifications.

In some cases, there is not enough consensus around a single technology on which the W3C could work, and yet the W3C would be a good environment for some work. Therefore, the W3C has established a lightweight *Incubator* process which allows its members to work on their technologies of interest on the neutral ground of the consortium, but with minimal support and without the ability to produce Recommendations.

An Incubator Group (XG) is established rapidly upon request of at least three W3C members, without the need for Advisory Committee approval, instead approved by the W3C Team. An XG gets the same infrastructure as a WG (Web pages, mailing lists etc.), for a duration initially not exceeding one year. XGs also publish Working Drafts, but the end product, possibly after Last Call(s), is a so-called XG Report, with no formal standards standing. However, an Incubator Group may request a simplified transition to a full Working Group; simplified by its already existing connection with the W3C, but still requiring full AC approval.

3.5 Report on Standardization Activities

In the duration of the project, several project partners have been involved in the standardization of semantic technologies important for triplespaces and their applications. This section summarizes this involvement.

The main query language for Triple Space (and triplespaces in general) is SPARQL, whose standardization in the W3C concluded in January 2008. The TripCom partner

Profium has been actively involved in this work, making sure that triplespace requirements are fulfilled by the query language and that the project is informed about important developments of the standard.

TripCom has also been concerned with interoperability with Web services, especially in Work Package 4. LFUI was involved in the W3C standardization of WSDL Version 2.0, concluded in June 2007. LFUI especially contributed to the work on the WSDL RDF Mapping that ties WSDL data into the Semantic Web.

The W3C also worked on semantic description of Web services; in particular, in August 2007 it produced the Recommendation of Semantic Annotations for WSDL and XML Schema (SAWSDL), the first step towards combining semantics with Web services. LFUI and NUIG were heavily involved in this standardization effort. While SAWSDL was not directly used in TripCom, it is the basis for follow-up work that combines the use of semantic spaces and semantic automation for Web services in the project SOA4All.[9]

Finally, also in the direction of Semantic Web Service automation, Since 2005, OASIS has been working on a specification of "guidelines, justifications, and implementation directions for deploying Semantic Web services in SOA" in the Semantic Execution Environment Technical Committee (SEE TC). LFUI is heavily involved in this effort, and NUIG is also a member of the committee; the results are again not directly applicable in TripCom, but they will be useful for follow-up work that will include semantic spaces and Web services.

Table 3.1 summarizes the involvement of consortium partners in standardization activities.

Table 3.1: TripCom Involvement in Standardization Activities

Standard	**Organization**	**Involved partners**
SPARQL	W3C	Profium
WSDL 2.0	W3C	LFUI
SAWSDL	W3C	LFUI, NUIG
SEE	OASIS	LFUI, NUIG

3.6 Triple Space Binding for Web Services

The Triplespace Web Service binding specification standardizes a Web service binding that allows to use the Triple Space as transport mechanism for SOAP messages. This way, Web service requesters and providers can leverage unique properties of Triple Space, such as extended Message Exchange Patterns and a semantics aware transport mechanism that avoids lowering and lifting of semantic data (information), while at the same time they integrate with existing Web service standards in order to be able to explicitly use other Web service technologies such as *WS-RM* or *WS-Security*.

The binding specification draft is part of TripCom deliverable D4.4 [35], and is available from the TripCom project website[10]. The binding specification is further reviewed currently and will be submitted to the STI International "Standardization

[9] `http://www.soa4all.eu/`
[10] `http://tripcom.org/deliverables.php`

and Reference Architectures" Service[11] for further development and support of the binding specification beyond the project duration of TripCom. STI International will also be responsible for a possible hand-over of the specification to a standardization body such as the W3C or OASIS.

[11] http://ras.sti2.org/

4 Conclusion

In this document, continuous technology watch on emerging and current standards related to technologies used within TripCom was documented. We have taken a closer look at standards in the area of semantic web, e-health, EAI and Web services, briefly summarizing each and providing an evaluation that gave suggestions on how to proceed, e.g. if the standard can be used as is, needs to be extended or that a new standard in the respective area is missing. Furthermore, TripCom's own standardisation activities were documented. This included ontology work in the area of EDIFACT, EAI and e-health, as well as concrete standards such as WSDL 2.0, SA-WSDL, SPARQL and SEE. Furthermore, a possible new standard, the "Triplespace Binding for Web Services", was discussed.

References

[1] T. Andrews, F. Curbera, H. Dholakia, Y. Goland, J. Klein, F. Leymann, K. Liu, D. Roller, D. Smith, S. Thatte, et al. Business Process Execution Language for Web Services. *URL: ftp://www6. software. ibm. com/software/developer/library/ws-bpel11. pdf, may*, 2003.

[2] K. Ballinger, D. Box, F. Curbera, S. Davanum, S. Graham, C. Liu, F. Leymann, B. Lovering, A. Nadalin, M. Notthingham, et al. Web Services Metadata Exchange, 2004.

[3] K. Ballinger, D. Ehnebuske, M. Gudgin, M. Nottingham, and P. Yendluri. WS-I Basic Profile Version 1.0 a. *online at: http://www. ws-i. org/Profiles/Basic/2002-10/BasicProfile-1.0-WGD. htm [accessed 17 December 2002]*, 2002.

[4] M. Bartel, J. Boyer, B. Fox, B. LaMacchia, and E. Simon. XML-Signature Syntax and Processing. *W3C Proposed Recommendation*, 20, 2001.

[5] T. Beale and S. Heard. Archetype Definitions and Principles. *OpenEHR.com: online at http://www.openehr.org/repositories/spec-dev/latest/publishing/architecture/archetypes/principles/REV HIST.html*.

[6] E. Christensen, F. Curbera, G. Meredith, and S. Weerawarana. Web Services Description Language (WSDL) 1.1, 2001.

[7] David de Francisco Marcos, Lyndon Nixon, and Germn Toro del Valle. Towards a multimedia content marketplace implementation based on triplespaces. In A. Sheth et al., editor, *Proceedings of the 7th International Semantic Web Conference (ISWC'08)*, number 5318 in Lecture Notes in Computer Science, pages 875–888. Springer-Verlag Berlin Heidelberg, October 2008.

[8] doug foxvog. Evaluation and Refinement of Ontologies D7.4. Deliverable, NUIG, January 2009.

[9] doug foxvog, Juan Pablo Palacios, and Elena Paslaru Bontas Simperl. Analysis of EDIFACT and other standards; TripCom D7.1. Deliverable, NUIG, September 2006.

[10] T. Moses (ed.). eXtensible Access Control Markup Language (XACML) Version 2.0. OASIS Standard, February 2005.

[11] Cost Accounting/Auditing EDI in the Finance, Information and Social areas. Guide de l'archivage lectronique scuris. http://www.edificas.org, 2005.

[12] Marco Eichelberg, Thomas Aden, Jörg Riesmeier, Asuman Dogac, and Gokce Laleci. Electronic health record standards a brief overview.

[13] Marco Eichelberg, Thomas Aden, Jörg Riesmeier, Asuman Dogac, and Gokce Laleci. A survey and analysis of electronic healthcare record standards. *ACM Comput. Surv.*, 37(4):277–315, 2005.

[14] United Nations Economic Commission for Europe. United Nations Directories for Electronic Data Interchange for Administration, Commerce and Transport. http://www.unece.org, 2005.

[15] The Global IT Association for Telecommunications. ETIS-EDIFACT User Guide. http://www.etis.org, 2005.

[16] The European B2B Forum for the Electronics Industry. UN/EDIFACT MIGs. http://www.edifice.org, 2005.

[17] United Nations Centre for Trade Facilitation and Electronic Business. Modeling Methodology (UMM), User Guide. http://www.ifs.univie.ac.at, 2003.

[18] S. Graham, P. Niblett, D. Chappell, A. Lewis, N. Nagaratnam, J. Parikh, S. Patil, S. Samdarshi, I. Sedukhin, D. Snelling, et al. Publish-Subscribe Notification for Web Services. *IBM DeveloperWorks Whitepaper*, 2004.

[19] HL7 Group. HL7 Version 3 Standard Clinical Data Architecture 2.0, 2005.

[20] M. Gudgin and M. Hadley. Web Services Addressing 1.0-Core. *W3C Candidate Recommendation, May*, 2006.

[21] M. Gudgin, M. Hadley, and T. Rogers. Web Services Addressing 1.0 SOAP Binding. *W3C Candidate Recommendation, May*, 2006.

[22] T. Imamura, B. Dillaway, E. Simon, et al. XML Encryption Syntax and Processing. *W3C Candidate Recommendation*, 4, 2002.

[23] The Global language of Business. EANCOM Guideline. http://www.gs1.org, 2005.

[24] The Global language of Business. Ecom standards in the GS1 Community 2005. http://www.gs1.org, 2005.

[25] The Global language of Business. GS1-XML user manual. http://www.gs1.org, 2005.

[26] A. Nadalin, C. Kaler, P. Hallam-Baker, and R. Monzillo. Web Services Security: SOAP Message Security 1.0. *OASIS specification*, 200401, 2004.

[27] Semantic Annotations for WSDL and XML Schema. Recommendation, W3C, August 2007. Available at `http://www.w3.org/TR/sawsdl/`.

[28] SOAP Version 1.2 Part 1: Messaging Framework. Recommendation, W3C, June 2003. Available at `http://www.w3.org/TR/2003/REC-soap12-part1-20030624/`.

[29] K.A. Spackman, K.E. Campbell, and R.A. Cote. SNOMED-RT - A Reference Terminology for Health Care. In *Proc. of the 1997 AMIA Annual Fall Symposium*, 1997.

[30] Claudia Tessier. The concept paper of the ccr.

[31] P. Vegoda. Introducing the IHE (Integrating the Healthcare Enterprise) concept. *Journal of Health care Information Management*, 2002.

[32] Web Services Description Language (WSDL) Version 2.0. Recommendation, W3C, June 2007. Available at `http://www.w3.org/TR/wsdl20/`.

[33] Web Services Description Language (WSDL) Version 2.0: RDF Mapping. Working Group Note, W3C, June 2007. Available at `http://www.w3.org/TR/wsdl20-rdf`.

[34] S. Weerawarana, F. Curbera, F. Leymann, T. Storey, and D.F. Ferguson. *Web Services Platform Architecture: SOAP, WSDL, WS-Policy, WS-Addressing, WS-BPEL, WS-Reliable Messaging and More*. Prentice Hall PTR Upper Saddle River, NJ, USA, 2005.

[35] Daniel Wutke, Davide Cerri, Michael Lafite, Daniel Martin, Jrg Nitzsche, and Christian Schreiber. Methodology for augmenting semantic web services with ws-* standards. TripCom Deliverable D4.4, 2008.