TripCom

Triple Space Communication

FP6 – 027324

Deliverable

D7.1

Analysis of EDIFACT and other standards

doug foxvog [NUIG]
Juan Pablo Palacios [TID]
Elena Paslaru Bontas Simperl [FUB]

October 16, 2006

Executive Summary

The aim of the Ontological Infrastructure work package (WP 7) is to offer a means for a semantically rich definition of business processes for the express purposes of overcoming heterogeneity problems. The concrete goal of Work Package 7 is to ontologize a significant portion of the current EDIFACT EDI standard as a basis for business-to-business process integration. The main objective of this document is to analyze this standard for the purpose of identifying relevant and important subsets which should be included in future ontologization efforts. After a general survey on EDI standards and EDIFACT (Chapters 2 and 3), we elaborate on major EDIFACT sub-standards and perform the selection of those appropriate for our needs (Chapter 4). The document concludes with two chapters dedicated to similar and relevant initiatives in the eBusiness sector. First, a closer look at ebXML, as the most representative example of the new generation of EDI products, is provided (Chapter 5). The UN/CEFACT consortium has remarked on the lack of interoperability between eBusiness applications and has taken initial steps to alleviate this situation, a goal which is shared by this work package. Their experience and methods are a valuable source of information for our future work. The same applies for complementary initiatives designed to transform (parts of) established EDI standards to more structured representation forms, such as XML, RDF, or WSML. We report on such projects in Chapter 6.

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**	7.1	**Title**	Analysis of EDIFACT and other standards
Work Package	**Number**	7	**Title**	Ontological Infrastructure for Business Processes and Data

Date of Delivery	**Contractual**	M6	**Actual**	15-9-06
Status	version 0.6		final ☐	
Nature	prototype ☐ report ☒ dissemination ☐			
Dissemination Level	public ☒ consortium ☐			

Authors (Partner)	doug foxvog (NUIG), Juan Pablo Palacios (TID), Elena Paslaru Bontas Simperl (FUB)			
Resp. Author	doug foxvog (NUIG)		**E-mail**	doug.foxvog@deri.org
	Partner	NUIG	**Phone**	+353 (91) 495 150

Abstract (for dissemination)	The aim of the Ontological Infrastructure work package is to offer a means for a semantically rich definition of business processes for the express purposes of overcoming heterogeneity problems. The concrete goal of the Work Package is to ontologize a significant portion of the current EDIFACT EDI standard as a basis for business-to-business process integration. The main objective of this document is to analyze this standard for the purpose of identifying relevant and important subsets which should be included in future ontologization efforts. After a general survey on EDI standards and EDIFACT, we elaborate on major EDIFACT sub-standards and perform the selection of the appropriate ones. The document concludes with examining similar and relevant initiatives in the eBusiness sector. First, a closer look at ebXML, as the most representative example of the new generation of EDI products, is provided. The UN/CEFACT consortium has acknowledged the lack of interoperability between eBusiness applications and has taken initial steps to alleviate this situation, a goal which is shared by this work package. Their experience and methods are a valuable source of information for our work. We also examine complementary initiatives designed to transform (parts of) established EDI standards to more structured representation forms, such as XML, RDF, or WSML.
Keywords	Ontology, EDI, EDIFACT, TripCom, business standards

Version Log			
Issue Date	**Rev No.**	**Author**	**Change**
2006-06-02	1	doug foxvog	Initial version created
2006-06-07	2	JP Palacios	Standards Analysis document incorporated
2006-08-11	3	JP Palacios	Analysis of mayor EDIFACT subsets incorporated
2006-08-17	4	Elena Simperl	Revision of Chapter 5
2006-08-18	5	doug foxvog	As discussed at teleconference; Chapter 7 removed
2006-08-28	6	Elena Simperl	Latex-related revision of the main chapters and the references.
2006-10-13	7	doug foxvog	Modifications due to QA.

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	Sara Carro Martinez Telefonica I+D Madrid, España E-mail: tripcom@tid.es

Table of Contents

List of Abbreviations

ANSI	American National Standards Institute
B2B	Business to Business
BIC	Book Industry Communication
CCL	Core Components Technical Specificdation
CCTS	Core Component Library
CECED	European Committee of Domestic Equipment Manufacturers
CEFACT	Centre for the Facilitation of the Administration, Commerce and Transport
CEFIC	European Chemical Industry Council
CIDX	Chemical Industry Data Exchange
CPA	Collaboration Protocol Agreement
CPP	Collaboration Protocol Profile
DOM	Document Object Model
ebXML	Electronic Business using eXtensible Markup Language
EDI	Electronic Data Interchange
EDIFACT	United Nations Directories for Electronic Data Interchange for Administration, Commerce and Transport
EDIFICAS	EDI in the Finance, Information, Cost, Accounting/Auditing and Social areas
EDIFICE	EDI Forum for companies with Interest in Computing and Electronics
EDItEUR	European Book Sector Electronic Data Interchange Group
EMEDI	European Medical Electronic Data Interchange
ETIS	European Telecommunications Information Service
FTP	File Transfer Protocol
GS1	The Global language of Business
ICEDI	International Committee on EDI for Serials
ISO	International Organization for Standardization
MIME	Multipurpose Internet Mail Extensions
OASIS	Organization for the Advancement of Structured Information Standards
TSC	Triple Space Computing
TTI	Trravel Technology Initiative
UBL	Universal Business Language
UNICORN	United Nations Interactive Communications Over Reservations Networks
UPC	Universal Product Code
VAN	Value Added Network
WSML	Web Service Modeling Language
WSMO	Web Service Modeling Ontology
XML	Extensible Mark-up Language

1 Introduction

1.1 Business Data Interchange

It is widely expected that the impact of computerization will be as great as the impact of the industrial revolution. Evident immersion of computer network technologies, software, data bases, and Internet within the enterprise world has caused a great change both in the way of working and in their negotiation mechanisms.

It is common today that in companies, even the smallest ones, activities are computerized by means of systems with high levels of automation in administrative-accountancy tasks. For example, accounting programs and payroll programs are frequently integrated, so data is processed very quickly.

In the inter-enterprise environment the reality is often very different: companies maintain intense commercial relationships with each other, but in spite of having advanced computational systems (to automate all of their economic transactions, payroll, invoices, inventory, etc.) they have many problems interchanging documents and messages with other companies. Documents are still often physically mailed, undergoing delays or loss information.

A somewhat more advanced solution is to send these documents by fax, which speeds up communication. If the company uses a fax-modem connected to a computer and managed by a software program, documents need not be printed by the sending party - they are posted directly from one company's computer to another company's fax machine.

A more sophisticated, but less frequent, solution consists of sending these documents via electronic mail. Electronic mail allows the sending of messages between different computers that are connected to a computer network. This procedure has its own disadvantages, because partners are free to create their own standards for electronic mail, or use none at all. Unless client and supplier previously decided the way of composing their messages so that they can be read by a computer program, the receiver needs to manually input received documents into their computer system.

The use of un-encrypted electronic mail is neither a safe media for conducting commercial transactions nor guarantees their confidentiality. Encrypted email is possible, but is often hard to use between commercial staff of two companies.

Even at these levels, computers provide a variety of services at rising speed and diminishing cost. Regular communication dealing with trade, whether domestic or international, can benefit from greater standardization. Not only do paper documentation and procedures represent as much as 10 per cent of goods value; also they are slow, insecure, complicated and growing. The possibilities of cost reduction are in the order of 50 per cent to the benefit of not only the main parties, but everyone involved, not least the authorities [29].

A solution that has been developed to solve these problems is the provision of a platform for electronic business data interchange. This allows electronic transmission of business documents between computer programs in a standardized form.

Starting in the 1970's such electronic data interchange (EDI) systems have been developed – first in the United States and later in Europe and Asia – for speeding up and standardizing regular inter-business communication.

Focused on Europe, the major activity of the Working Party on Facilitation of International Trade Procedures of the United Nations Economic Commission for Europe

(UN/CEFACT) [12], over the last decade and a half has been the development and enhancement of tools that would make electronic interchange of data in international trade be a secure, effective and cheap alternative.

A series of technologies have appeared with the objective to solve the electronic business data interchange problem. In this paper, we analyse the major European EDI system (UN/EDIFACT) and its variations.

1.2 EDI

1.2.1 Concept

Electronic Data Interchange is the computer-to-computer exchange of structured information, by agreed message standards, from one computer application to another by electronic means and with a minimum of human intervention [8]. In common usage, EDI is understood to mean specific interchange methods agreed upon by national or international standards bodies for the transfer of business transaction data, with one typical application being the automated purchase of goods and services.

Early developments of EDI were driven primarily by banks and large buying organizations - supermarkets, chain stores, health services - which had the financial muscle to influence their trading partners to adopt a standard method of electronic trading, initially largely to the benefit of the buyer, though ultimately with benefits for both sides. The typical application would be based on a "hub and spoke" model, the hub being the buying organization, and the spokes its suppliers. Messages would be transported over dial-up or, for large users, leased line connection, to a value-added network (VAN), using proprietary protocols for assuring integrity, security, and an end-to-end audit trail. Each party would have a mailbox on the VAN.

The economics of this form of communication involved a volume charge for communication, so that the message standards aimed for conciseness, the fullest possible use of coded data, and minimal use of text.

Since the real payoff from EDI came not just from more efficient communication, but also from faster and more efficient processing and turn round of transactions at each end, it was important that a message could be processed automatically, and this was and is another powerful factor favoring coded data rather than text.

Early applications did nothing more than automate the existing way of doing simple transactions such as orders, delivery notes and invoices, so that the benefits came (at least for the "hub") from speed, accuracy, removal of duplicate keyboarding and elimination of paper and postage. The "spokes" often had employees key in the EDI messages they sent and printed out received messages so that they could be keyed into their own systems.

More sophistication has crept in as companies have realized that continuous electronic communication between supplier and customer can bring additional benefits, e.g. by anticipating and planning for demand rather than waiting for orders, or even by allowing the supplier to take control of the maintenance of appropriate stock levels in stores (so-called "vendor-managed inventory").

The standards are broad enough to encode variations in message requirements used by different industries. Industry bodies define subsets of and modifications to the EDI standard for messages which they use. Individual companies add further restrictions

to the EDI messages which they send and accept. These modifications are applied in order to adjust the standard defintiion to specific needs of the companies.

The use of different standards is so prevalent that when two corporations wish to start up EDI communications they must negotiate a technical agreement to define exactly what subset of EDI they will use [19]. Such an agreement is called an implementation guideline in the industry.

These EDI varients define some optional EDI components as mandatory and others as forbidden, specify additional inter-component restrictions, identify a subset of codes within used code sets that will be accepted and used, may use additional codes, and restrict the message types that will be used.

The consequence of the widespread adoption and divergence of EDI messaging standards is that there is an equally wide range of software and systems available to support their use. Most commercial users do not develop their own EDI front-ends. They buy a package which typically handles mapping between the EDI standard and internal file formats defined by the user; the management of trading partner relationships, by maintaining a database of who's who and what messages are enabled as part of the trading agreement; and the timetabling and automatic running of online sessions to send and receive messages [23].

Increasingly, the Internet is now being used for sending and receiving EDI messages, either by FTP or by email transfer. A simple MIME protocol for encapsulating an EDI message was defined in 1995 [4]; and, under the aegis of the Internet Engineering Task Force, a much more comprehensive standard is being finalized which is intended to enable an EDI message to be sent over the Internet with the same levels of security and audit trail which Value Added Networks have traditionally provided.

1.2.2 Standards

Recent years have seen an explosion of interest in EDI between national and international trade participants. The technology is available; however, the entry cost for using traditional EDI is prohibitive for many small and medium-scale enterprises.

Nowadays, there are two major sets of traditional EDI standards. UN/EDIFACT is a United Nations recommendation [1] and is predominant outside of North America. ANSI ASC X12 (X12) is predominant in North America and used worldwide. Other systems are locally prominent in Japan and South Korea. These standards prescribe the formats, character sets, and data elements used in the exchange of EDI documents and forms.

EDI documents contain the same data that would normally be found in a paper document used for the same organizational function. For example, an X12 940 ship-from-warehouse order is used by a manufacturer to tell a warehouse to ship product to a retailer. It typically has a ship-to address, a bill-to address, a list of product numbers (usually UPC codes), and quantities. It may have other information if the parties agree to include it. However, EDI is not confined to just business data related to trade but encompasses fields such as medicine (patient records, laboratory results, etc.), transport (container and modal information, etc.), engineering, and construction.

An EDI standard says which pieces of information are mandatory for a particular document, which pieces are optional, and gives the rules for the structure of the document. The standards are like building codes. Just as two kitchens can be built "to code" but look completely different, two EDI documents can follow the same

standard and contain different sets of information. For example, a food company may indicate a particular product expiration date while a clothing manufacturer would send color and size information.

Organizations that send or receive documents from each other are referred to as "trading partners" in EDI terminology. The trading partners agree on the specific information to be transmitted and how it should be used. This is done in human-readable specifications (also called specs or spec/specification sheets).

While the standards are analogous to building codes the specifications are analogous to blueprints. (The specification may also be called a mapping but the term mapping is typically reserved for specific machine-readable instructions given to the translation software.) Larger companies have existing specification sheets and are usually unwilling to negotiate. Often in a large company these sheets will be written to be used by different branches or divisions and therefore will contain information not needed for a particular exchange. Deviations from and clarification of the specification sheets should always be obtained in writing.

1.3 Overview of Rest of Deliverable

Chapter 2 discusses major EDI standards in general. Chapter 3 goes more into detail on the UN/EDIFACT standard. Chapter 4 analyzes major EDIFACT subsets used by different industries, the coverage of these subsets, and the selection of EDIFACT message types and subsets to be ontologized. Chapter 5 investigates the relationship of EDIFACT with more recent UN/CEFACT standards. Chapter 6 discusses related and prior work with ontologizing EDI standards. Chapter 7 concludes the discussion.

2 EDI Standards

2.1 UN/EDIFACT

The United Nations Centre for Trade Facilitation and Electronic Business, (UN/CEFACT) created the UN/EDIFACT standard, which it officially defined 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."

2.1.1 Concept

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 United Nations Electronic Data Interchange for Administration, Commerce and Transport (UN/EDIFACT) syntax rules. The maintenance and extension of this standard is done by UN/CEFACT, the United Nations Centre for Trade Facilitation and Electronic Business under the UN Economic Commission for Europe [12].

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 data elements definitions which are grouped in sets, according with some rules, forming segments. 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 need to be exchanged between these 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 [9].

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

2.2 X12

The Accredited Standards Committee (ASC) of the American National Standards Institute (ANSI) created the X12 standard, publishing it in 1979. The standard comprises a set of agreed standards and guidelines for the electronic interchange of structured data between independent, computerized information systems.

2.2.1 Concept

ASC X12 is a standard designed to cover EDI trading operations. The maintenance and extension of this standard is done by ANSI's Accredited Standards Committee.

ASC X12:

- Provides a set of syntax rules to define structure data.
- Provides an interactive exchange protocol.
- Provides standard messages (allowing multi-country and multi-industry exchange).

X12 contains data elements definitions which are grouped in structures, according with some rules, forming data segments. Message types, called "transaction sets", are defined as structures of data segments and groups of data segments, again with associated rules. These structures and rules are published in proprietary documents.

X12 standardizes the messages needed 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 X12.

X12 is a flexible standard in order to cover all possible restrictions and business domains needed by users.

The X12 standard changes periodically, with message types and message type components being added, modified, and sometimes removed.

2.3 XML-EDI Standards

When electronic transfer of information is performed, machines must be aware of the format of the data being exchanged, but people would like this information to be human-readable as well. XML is the most popular human-readable data exchange language nowadays, used in conjunction with XML Schemas and XML Stylesheets technologies to separate data format from data presentation layers.

While an equivalent EDIFACT message will be smaller than an XML message, the XML message will be easier for a person to read. This has some benefits in practical

application when developers from disparate companies attempt to establish automated communications and troubleshoot systems.

Different XML standards are being developed to apply to different business domains, while EDIFACT tries to cover all domains. The standards being developed for different industries are generally not designed to be compatible with either each other or legacy EDI systems. One of the emerging XML standards for business messages is RosettaNet, widely used in the semiconductor and high tech industry in general. Another is UBL, currently being adopted by Scandinavian governments as a legal requirement to send invoices to governments. For example, all invoices to the Danish government have had to be in electronic format since February 2005.

2.4 Applicability Domains

Although XML-based languages seem to be proliferating, traditional EDI (EDIFACT and X12) is still widely used in the high tech, civil aviation, and tourism industries and is likely to remain so for some time due to the amount of software making use of it and the need for newer systems to be able to integrate with legacy systems. Europe has a large installed base of EDIFACT users, where as the Asian region started later with electronic B2B implementations and is therefore using more XML standards.

EDIFACT plays an important role dealing with international business in certain business domains. Chapter 3 describes some of the most common areas of use.

3 EDIFACT

3.1 Development and maintenance processes

EDIFACT is developed and maintained under the aegis of UNECE, the United Nations Economic Commission for Europe, whose remit in this respect runs globally - in spite of the name. The responsible agency is UN/CEFACT, the Centre for the Facilitation of Procedures and Practices in Administration, Commerce and Transport.

CEFACT is supported by national and regional bodies worldwide; and through a series of working groups and plenary meetings, it publishes successive revisions of the EDIFACT Directory. It is necessary to check these bodies to design a new EDIFACT message, in this way, it is ensured that similar messages, which could fit the requirements of the extensions already defined, can be made.

The EDIFACT Directory is the authority which specifies data elements and code lists, segments and messages. New directories are issued normally twice a year, and can be downloaded as text files from a UNECE Web site [8]. A number of commercial suppliers of EDI software also maintain websites from which the directories are available in other formats, together with software for viewing them and using them to derive specific application subsets.

The standards are, however, more stable than a twice-yearly revision process might seem to imply. Since 1992, when there was a rather major review of the basic messages, almost all change has been upwards-compatible. Most changes are additions or clarifications.

Successive EDIFACT directories are developed within the current version of EDIFACT syntax. New versions of the syntax are infrequent, and when they are required they go through a full ISO approvals process.

3.2 Syntax and structure

3.2.1 Interchange Structure

EDIFACT provides a standardized way to exchange information between trading partners. This is accomplished through an EDIFACT transmission, or interchange. This structure is organized on three levels: interchange, group, and message, as shown Figure 3.1.

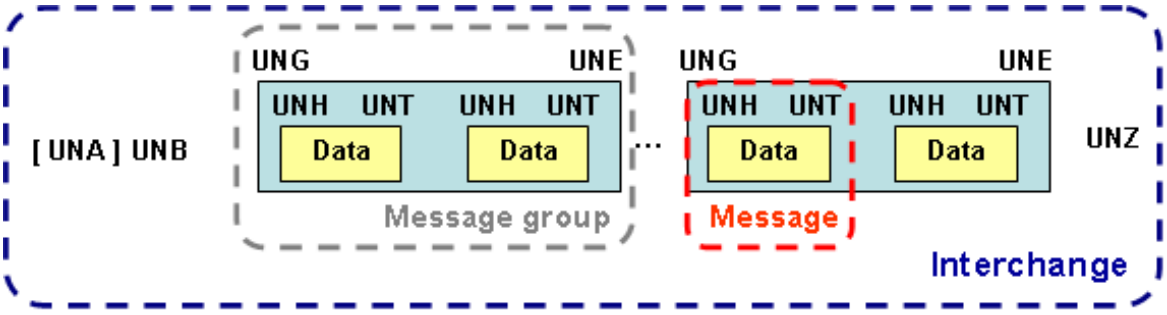

Figure 3.1: EDIFACT Interchange Structure.

As EDIFACT syntax specification (Version 3) defines, an EDIFACT interchange begins with a `UNA` (optional) or `UNB` (mandatory) segment and ends with a `UNZ` segment. These segments constitute the "envelope" for all the messages of the transmission.

The `UNA` segment defines which codes are going to be used as separators between segments used during the transmission. If no `UNA` segment is specified, a default value is used. The `UNB` segment is a header which identifies the sender and addressee for the transmission, specifies the character set used, and carries other transmission data. The `UNZ` segment ends the interchange structure, and provides a first level of error checking of the interchange.

Inside this envelope, many messages can be transmitted (although this is not mandatory and your can send a single message in each interchange). These messages are grouped relating all the messages of the same type. A message group begins with a `UNG` Functional Group Header segment and is ended by a `UNE` Functional Group Trailer segment.

Each message is introduced by a mandatory `UNH` Message Header segment, which identifies the type and version of the message, and ends with a mandatory `UNT` Message Trailer segment, which provides a basic error checking function. In some messages there is an additional service segment, `UNS`, which acts as an internal separator between two major sections of the message.

Segments defined inside an EDIFACT interchange can be service or user segments. Both service segments and user segments begin with a three-letter identifier, and end with a segment separator. Segments with names starting with `UN` are service segments. User data segments contain the information itself, in a format specific to each message type.

3.3 Structure of a message

Each EDIFACT message is composed of three sections (see Figure 3.2): header, detail and summary. Each section is made up of segment groups and segments. Each segment is made up of data elements which may be composite.

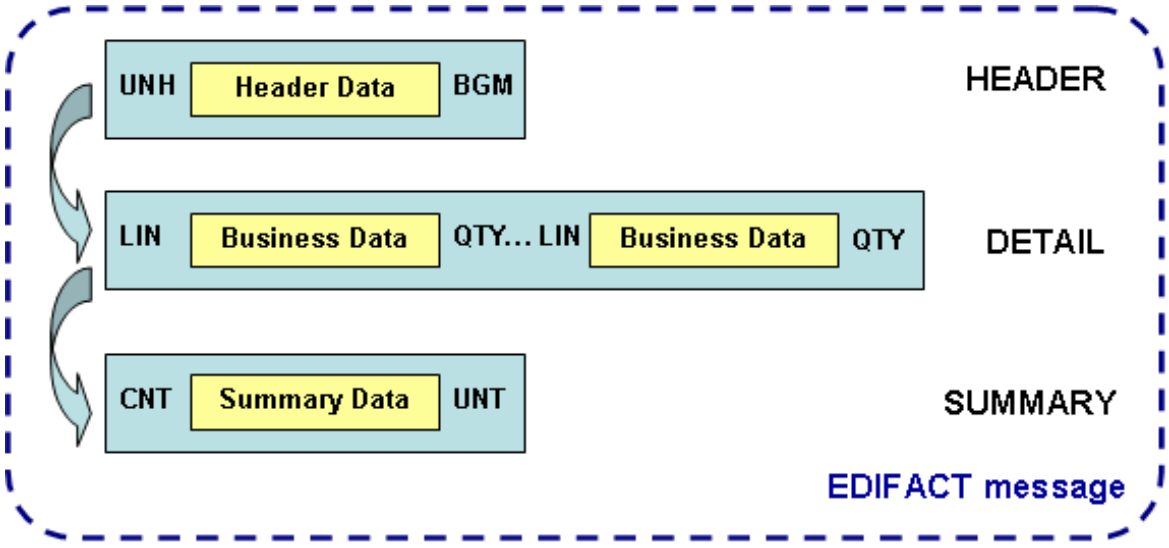

Figure 3.2: EDIFACT Message Structure.

The header section is a group of segments containing information which relates to the entire message. The complete header section can not repeated, though it might contain some segments or groups of segments which may be individually repeated.

The detail section is one or more groups of segments carrying information which relates to repeating elements in the message, such as order lines or invoice lines. By analogy with paper documents, in many of the simpler standard transaction messages each repeat of the detail section is referred to as a line. In more complex messages the structure of the detail section may involve nested repeats at several levels.

The summary section is a group of segments containing totals or control information, e.g. invoice total amount or number of lines in a purchase order. Like the header section, it occurs only once per message.

The same segment type may occur in more than one of the message sections, e.g. in the header and in the detail section, and/or more than once in the same section.

Some segments may be repeated a certain number of times at their specific location in the message. The status (mandatory or conditional), and the maximum number of repetitions of segment types are indicated in the message specification through a branching diagram and a segment table which defines the information of a segment.

Within a message, specific groups of functionally related segments may be repeated; these groups are referred to as segment groups. The maximum number of repetitions of a particular segment group at a specific location is indicated in the message specification.

A segment group may be nested within other segment groups, provided that the inner segment group terminates before any outer segment group terminates.

3.4 Structure of an EDIFACT segment

A segment consists of:

- A three-letter segment tag, which identifies the segment type and is mandatory in every segment.
- Component data element separators, which are optional.
- Nesting and repeating indicators, which are optional.
- Data element separators, which are mandatory.
- Simple, composite, or component data elements, which are optional or mandatory depending what is specified in segment directories.
- A segment terminator, which is mandatory.

Data elements may be either simple or composite. Composite data elements contain two or more simple data elements, called component data elements used inside a composite data element.

Simple data elements can be defined as having fixed or variable length. Data elements may be filled by coded values, numeric values, or text strings (see Figure 3.3).

A data element can be qualified by another data element, the value of which is expressed as a code that gives specific meaning to the data. The data value of a qualifier is a code taken from an agreed set of code values.

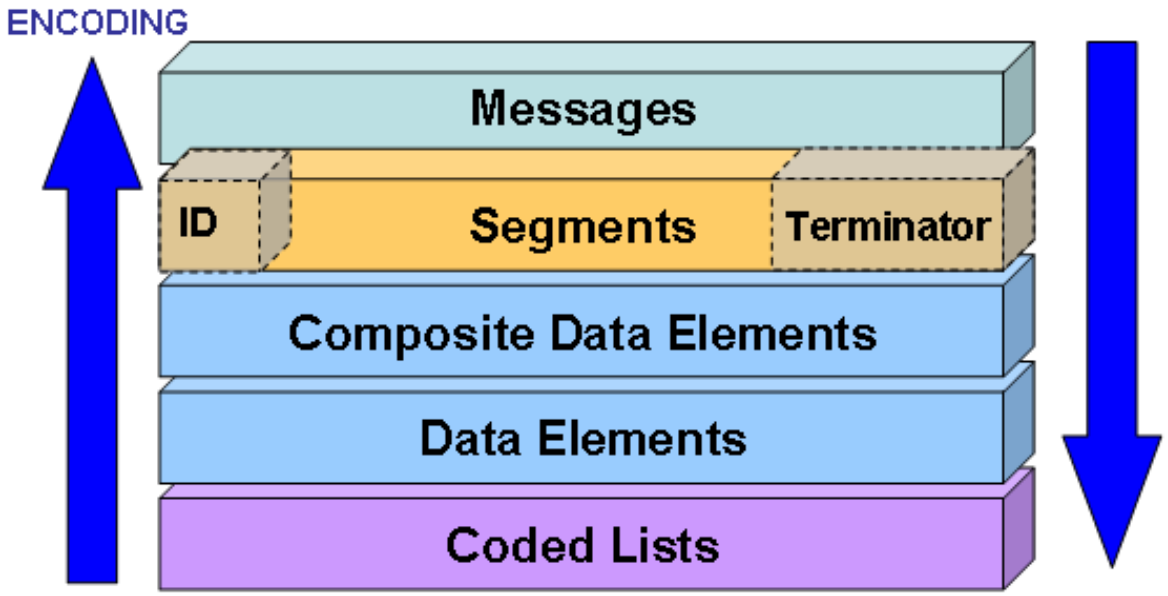

Figure 3.3: EDIFACT message layers.

Here is a simple EDIFACT segment example: `DTM+137:19940121:102'`

- `DTM` = segment tag, identifying the "Date/time/period" segment
- `+` = segment tag and data element separator
- `137` = date qualifier, code indicating that the date is the document/message date/time
- `:` = separator for component data elements within a composite (here, date qualifier and date)
- `19940121` = date, in the format specified by the date format qualifier
- `:` = separator for data elements within a composite (here, date and date format qualifier)
- `102` = date format qualifier, code indicating the format of the date (here, CCYYMMDD)
- `'` = segment terminator

3.5 Representation

3.5.1 Character sets and separators

EDIFACT standards define a number of character sets, coded in the `UNB` segment at the beginning of each interchange as UNOA, UNOB, UNOC, etc. UNOA and UNOB correspond to the basic ASCII character sets of ISO 646 and ISO 6937. UNOC to UNOF correspond with the alphabets defined in Parts 1, 2, 5 and 7 of ISO 8859.

Book Industry Communication (BIC) in the UK and EDItEUR internationally have adopted UNOC as the standard character set for book and serials trading within

EDIFACT syntax version 3. This character set permits the representation of a full repertoire of special characters, and corresponds to the international standard character set ISO 8859.1 [16].

Version 4 of EDIFACT syntax, most of which has recently completed the ISO approval process, will extend the character set options in order to include UNICODE support.

There is a default set of separators associated with each standard character set UNOA, UNOB etc. If the default is not used, the `UNA` segment must be sent.

The set shown here is the default for UNOA:

- Apostrophe ' segment terminator
- Plus sign + segment tag and data element separator
- Colon : component data element separator
- Period . decimal point (in numeric data elements only)
- Question Mark ? release character

The function of the release character is to allow a symbol which has been defined as a separator to be used with its normal meaning in a text data element. The release character precedes the character which is to be returned to its normal meaning. These release characters are not taken into account within the maximum length defined for an EDIFACT data element.

3.5.2 Representation of numeric values

Numeric data elements in EDIFACT are sent as real, positive or negative, decimal numbers, with a decimal point and/or a minus sign if required, and without leading or trailing zeros. The only exception to this rule is a single zero before the decimal point if the value is less than 1. The decimal point and the minus sign are not counted within the maximum length of the data element.

Within the overall maximum length of a numeric data element, EDIFACT itself does not define how the number may be divided between an integer part and a decimal fraction part. However, BIC and EDItEUR implementation guidelines include recommendations to be followed for book and serials applications.

3.5.3 Omission of unused elements

In order to keep messages concise, EDIFACT omits unused elements wherever possible. If any complete segment or segment group is not required in the message it is omitted. A segment is terminated immediately after the last data element with non-empty content. Empty data elements are represented by including the required separators, and nothing else.

4 Analysis of major EDIFACT subsets

Due to the vast number of segments defined in EDIFACT, most industries and organizations 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 applications; it is in contrast to the original idea of a common standard for electronic data interchange.

In this chapter some of these subsets will be described in order to find some aspects that could be interesting for our project.

4.1 Common Uses of EDIFACT

EDIFACT is a world-wide, cross-industry standard for electronic communications dealing with trade of all kinds. One typical characteristic of EDIFACT is that "standard EDIFACT messages" are normally used in a far more limited manner than their definitions indicate. Since standard messages have evolved through accretion of optional data elements to handle the information requirements of every conceivable business relationship, they contain vastly more information than is typically necessary in any particular case. As a result, the message forms that are exchanged between trading partners are always substantially reduced parties of the full message specification which are heavily customized to the relationship between the partners.

This way, EDIFACT has a self-fulfilling bias against enabling spontaneous commerce in open trading communities. Besides, because of the historically high cost of EDI integration, companies do not use it unless they have entered into a long term, high volume, or high value business arrangement.

When a point-to-point relationship exists, it is possible to optimize it by encoding in EDI messages any information that is specific to that relationship, such as contract numbers, buyer catalog numbers, and so on. These optimizations make these messages at worst unintelligible and at best bloated from the perspective of other potential business partners, even in the same trading community.

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. This report describes several of these subsets – typically called Implementation Guidelines (cf. Table 4.1).

4.2 The Major Subsets

4.2.1 CIDX - CEFIC

CIDX (Chemical Industry Data eXchange) [6] is an organization of chemical companies and their trading partners, whose purpose is to improve profitability by promoting business process simplification and innovation using EDI and related technologies.

CIDX's main objective focuses on the development of eBusiness standards for its knowledge area, called Chem eStandards, which have become the facto standards for transacting business electronically in this industry.

Standard	Industry	Publicly available	Versioned
CIDX - CEFIC	Chemical Industry	yes	yes
EANCOM - GS1	International Article Numbering Association	yes	yes
EDIFICAS	Accounting	yes	yes
EDIFICE - ICEDIS/BIC	IT, Electronics, Phone	yes	yes
EDItEUR	Library supply	no	yes
EDIWHITE - CECED	Appliances	yes	yes
EMEDI	Medical	no	yes
ETIS	Telecommunications	yes	yes
UNICORN	Travel bookings	no	yes

Table 4.1: EDIFACT subsets

CIDX is active world-wide organization with member based in Europe and North America, in Japan through a close alliance with the Japanese Petrochemical Association's Chemical EDI (CEDI) committee, and with a developing initiative in Latin America. We can say that CIDX mirrors the global nature of the chemical industry with substantial work groups and efforts around the world, including the European Chemical Industry Council - CEFIC [3]. From here, the name of this subset for Chemical Industry.

CIDX coordinates the Chem eStandards Initiative on behalf of its member companies and the chemical industry at large. CIDX and its members identify projects needed to enhance industry use of eBusiness. Industry volunteers form work teams to produce project deliverables that are made available to the entire chemical industry.

Chem eStandards are, concretely, the eBusiness standards developed specifically for the buying, selling and delivery of chemical products. Chem eStandards were developed through a cooperative effort of more than 20 chemical companies. Chem eStandards are comprised of business process guidelines, message specifications, envelope and security specifications, and various implementation tools and technical white papers. Both current and archived versions of these publications are available for download, some only to CIDX members. Efforts are underway with RAPID (agricultural inputs industry), PIDX (petroleum industry), papiNet (paper and forest products industry), AIAG (automotive industry), RosettaNet (electronics industry), and GUSI (consumer packaged goods industry).

4.2.2 EANCOM - GS1

GS1 is a leading 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 [21].

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** [22] for XML message formats
- **GS1 EANCOM** [20] 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 [20]:

- 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.

4.2.3 EDIFICAS

The French organization EDIFICAS Europe (EDI in the Finance, Information, Cost, Accounting/Auditing and Social areas) [5] 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 enterprizes.
- 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.

4.2.4 EDIFICE

EDIFICE [11] 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.

In 2004, EDIFICE started the development of EPIGs (EDIFICE PIP Implementation Guidelines) for those RosettaNet PIPs relevant to the EDIFICE Business scope.

4.2.5 EDItEUR

In the library supply business, EDIFACT is leading the trade in its adoption. The main reason is that book and journal supply is an international business and book editors want to implement a single standard for as many of their customers as possible. That is why there is a need for a standard application in this area.

EDItEUR [7] is the international group coordinating development of the standards infrastructure for electronic commerce in the book and serials industries. It provides its international membership with research, standards, and guidance in such diverse areas as:

- EDI and other eCommerce standards for book and serial transactions.
- Bibliographic and product information.
- The standards infrastructure for digital publishing.

- Radio frequency identification (RFID) tags.
- Rights management and trading.

EDItEUR has more than 90 members from 17 countries, including Australia, Canada, Japan, South Africa, the United States, and most European countries.

A considerable number of library system suppliers have installed basic EDIFACT book ordering functions in existing systems, and these are in daily use with suppliers such as Dawson's and Blackwell's, Harrassowitz, and Casalini. Currently, many if not most international library systems suppliers are working on more comprehensive implementations, for both books and serials. It has been particularly encouraged to EDItEUR that US libraries and their systems suppliers have been among the first to adopt these standards, in spite of the existence and widespread trade use of the US national EDI standard, X12.

Membership of EDItEUR is open to individual enterprises with an interest in EDI in the book trade and to relevant associations. Only the members of EDItEUR have free access to all EDItEUR standards and reports.

Two important members of EditEUR involved with EDI standardizations are:

- **ICEDIS** [29]. The International Committee on EDI for Serials is an independent organization which has commenced work on standards for orders and order responses. This standard will include activities to handle multiple currencies, VAT, and other sales taxes required by law. It is not intended to replace the standard already developed for handling machine-readable orders but rather to automate the handling of late renewals placed after the main renewal activity each year.
- **BIC** [2]. Book Industry Communication is an independent organizatio nsponsored and partially-funded by the Publishers Association, Booksellers Association, the Chartered Institute of Library and Information Professionals and the British Library to promote increased efficiency in the book and serials supply chain - physical and electronic - through the application of standard processes and procedures and e-commerce.

4.2.6 EDIWHITE- CECED

CECED (European Committee of Domestic Equipment Manufacturers) [27] was founded in 1958 as the European Association of National Associations. CECED represents the European household appliance industry. The member companies produce the following products:

- **Large household appliances**, including refrigerators, freezes, dishwashers, washing machines, clothes dryers, ranges, ovens, hoods.
- **Small household appliances**, covering a wide range of products from shavers to vacuum cleaners.

The main objective of CECED is to represent and defend European Industry interests. The Association acts as a partner in dialogue with the EU's political and regulatory

institutions and also interacts with the press and media, non-governmental organizations representing a range of different interests (including consumers and environmental groups) and other interested parties.

CECED provides a forum for the industry to reach consensus on issues of common interest such as product/process standardization and regulatory initiatives. This consensus-building role can also lead to concrete commitments or agreements within the industry that can become the basis for EU policy initiatives.

CECED developed a EDIFACT subset called EDIWHITE [28], in order to improve the industry's electronic transactions in this area. A number of messages were agreed upon, such as order, order response, order change, invoice, dispatch advice, price catalogue, remittance advice, receiving advice, transport instruction, sales and stock report, partner identification, delivery forecast and schedule, and just-in-time delivery.

4.2.7 EMEDI

EMEDI (European Medical Electronic Data Interchange association) [18] is the pan-European initiative for the promotion and awareness of EDI in health care. The EMEDI organisation was launched in Brussels in 1990 as a not-for-profit, user-driven group with the aim of promoting and developing the use of standardized EDI between all partners in the healthcare sector.

EMEDI's mission statement is as follows:

- To act as a European forum for users of EDI in health care, including healthcare agencies and other organizations that serve the healthcare community.
- To promote the implementation of EDI in health care using international standards, such as UN/EDIFACT.
- To represent users of EDI in the health sector in the EBES (European Board of EDI Standardization) , in the EBES Expert Group 9 for Health Care Messages, as well as in other agencies active in the field of EDI in Europe.

EMEDI works closely with the European standardization procedures managed by the European Committee for Standardization (CEN) and in a particular has a formal liaison with the key healthcare standards making committee TC 251 (Healthcare Informatics).

EMEDI has over 100 member organizations from 16 European countries, including healthcare providers, health insurance companies, pharmaceutical industries, IT suppliers, government bodies, and national standards organizations.

User groups like EMEDI offer the opportunity to promote the implementation of EDI using their in-depth knowledge of a specific sector. EMEDI brings together national and regional healthcare providing organizations, leading healthcare insurance companies and healthcare suppliers from various European countries. EMEDI provides the means to promote the needs of healthcare users to the EDI standardization process. In addition, EMEDI can, through its wide membership and promotion and awareness activities, help achieve the critical mass required for the use of standardized EDI in health care. EMEDI message types have been accepted into EDIFACT unchanged.

The EMEDI organization has three task groups related to the areas in health care where EDI is applied: medical, finance and logistics. EMEDI has an open structure that enables interested parties to participate at a level that suits their requirements.

Members may participate in various task groups or attend the annual meeting to gather or exchange information.

4.2.8 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 [10].

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 such as CEN/ISSS, TMF, OMG, TOG, EURESCOM, ETSI in the development and harmonization of work in architectures, particularly those related to IT in telecommunications.

ETIS currently maintains the following six working groups:

- Information Security
- Electronic Billing
- Enterprise Architecture
- International Settlements
- TeBIT survey
- CIO Executive Forum

Two of these working groups are of interest for this report:

- **International Settlements**. ETIS International Settlements Working Group is a group of experts from different telecommunications operators dealing 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 is 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.

4.2.9 UNICORN

The UNICORN (United Nations Interactive Communications Over Reservations Networks) messaging guidelines [17] were conceived by the European ferry industry to process and respond to travel requests between host reservation systems and travel agents. Since the first messages were developed in 1985, UNICORN has been expanded to cover other sectors of travel, such as air, general sales, insurance, rail and accommodation. The ability to provide travel documentation, such as paper tickets and magnetically encoded ATB2s (Automated Ticket and Boarding Cards) is also provided in the message set.

The 2006 release of UNICORN has nearly a hundred and forty message types. These messages use EDIFACT message segments and follow the EDIFACT guidelines, but are distinct from EDIFACT messages.

UNICORN was developed within the guidelines defined in the United Nations Trade Data Interchange Standards which included a body of existing applications, principally within freight transport. UNICORN adopts the flexible approach of the UNTDI framework that allows EDI to expand in the travel industry and to cross trading boundaries to other related business sectors.

Typically, UNICORN is used to pass messages between host computer systems, although there are applications between PCs and hosts. UNICORN is a set of EDI messages that sit above the user application and have no influence on the presentation of data at the receiving system and is not designed to interact directly with non-intelligent terminals.

UNICORN messages support the following business applications:

- Product information, enquiries, tariffs, schedules, and availability
- Making of reservations
- Enquiry on, amendment to, or cancelation of reservations
- Pricing, ticketing, and production of similar documents
- Free text

The UNICORN Application Message Standard is issued and maintained by TTI (Travel Technology Initiative). Development and modification of UNICORN messages is an on-going process through TTI and membership is recommended for any user. There are four levels of TTI membership: Executive, Associate, Academic, and Trade Body. For non-members of TTI the cost of obtaining a copy of UNICORN is over 2,000 euros.

4.3 Criteria for Selecting Subsets to Ontologize

The EDIFACT standard is so large; the number of subsets is so great, and the need for EDI messages is so limited for TripCom use cases, that we intend to ontologize

appropriate subsets of the standard, broadening the coverage as circumstances allow. This ontologization process is discussed in detail in the original TripCom work plan (Annex I) and in the Detailed Work Plan (Deliverable 10.1).

We select subsets of EDIFACT to ontologize first on several grounds:

- Overlap with other subsets – A type of message used in several industries is likely to be more useful than one used in a single industry.
- Utility of message types for TripCom use case in WP 8A (Enterprise Integration) – This work package is desiging its communications to fit standard EDIFACT messages. WP7 has identified possible message types for use by WP8A given their initial plan for the use case. Mutual feedback between the two work packages will continue.
- Utility of message types for TripCom use case in WP 8B (European Patient Summary) – WP7 has been in discussion with WP8B about the use of EMEDI medical messages.
- Use of subset by Partners' industries (IT, Telecommunications)
- Being actual subsets of EDIFACT, not merely using EDIFACT format – Message types of interest in non-subset variants are likely to have similar message types in standard EDIFACT and its subsets.

As time allows, additional subsets may be ontologized.

The message types used by Use Cases 8A and 8B will be completely ontologized by month 18 so that they may use the developed ontologies in their prototypes. Work will continue on other parts of the ontology until month 24.

4.4 The Major Message Types

CEFIC (the chemical industry) and EDItEUR (library supply) have little to do with the use cases of Work Packages 8A (EIA) and 8B (medical records) or with the industries of the consortium's products.

UNICORN, although it uses EDIFACT format in its messages, has created a completely different set of messages. The message types of the other listed subsets have more bearing on the TripCom use cases and/or overlap to a greater or lesser extent.

We propose to start the EDIFACT ontologization project with these message types and subsets. The EDIFACT message types included in these subsets are shown in Table 4.2.

Acronym	Description	EANCOM	EDIFICAS	EDIFICE	EMEDI	ETIS
ATHSTS	Authorization status	no	no	yes	no	no
AUTACK	Secure authentication and acknowledgement	yes	no	no	no	no
BALANC	Balance	no	yes	no	no	no
BANSTA	Banking status	yes	no	no	no	no
BORDERS	Blanket order	no	no	yes	no	no
CHACCO	Chart of accounts	no	yes	no	no	no

CHAMAP	Structure of accounts	no	yes	no	no	no
COACSU	Commercial account summary	yes	no	no	no	no
COMDIS	Commercial dispute	yes	no	no	no	no
CONTRL	Syntax and service report message	yes	no	yes	no	no
CREMUL	Multiple credit advice	yes	no	no	no	no
CUSCAR	Customs cargo report	no	no	yes	no	no
CUSDEC	Customs declaration	no	no	yes	no	no
CUSRES	Customs response	no	no	yes	no	no
DEBMUL	Multiple debit advice	yes	no	no	no	no
DELFOR	Delivery schedule	yes	no	yes	no	no
DEMDOC	Digital Document Request	no	yes	no	no	no
DESADV	Despatch advice	yes	no	yes	no	no
DIRDEB	Direct debit	yes	no	no	no	no
ENTREC	Accounting entries	no	yes	no	no	no
FINSTA	Financial statement	yes	no	no	no	no
HANMOV	Goods handling and movement	yes	no	no	no	no
IFCSUM	Forwarding and consolidation summary	yes	no	no	no	no
IFTMAN	Arrival notice	yes	no	yes	no	no
IFTMBC	Booking confirmation	yes	no	no	no	no
IFTMIN	Instruction	yes	no	yes	no	no
IFTSTA	Int'l multimodal status report	yes	no	yes	no	no
IHCEBI	Health insur. eligib. and benefit inquiry	no	no	no	yes	no
IHCLME	Health care claim or encounter request	no	no	no	yes	no
INFENT	nterprise accounting information	no	yes	no	no	no
INSDES	Instruction to despatch	yes	no	no	no	no
INVOIC	Invoice	yes	no	no	no	yes
INVRPT	Inventory report	yes	no	yes	no	no
INVRPTDS	Inventory report (distributor to supplier)	no	no	yes	no	no
INVRPTFI	Inventory report (forecast and inventory management)	no	no	yes	no	no
INVRPTSD	Inventory report (supplier to distributor)	no	no	yes	no	no
LEDGER	Ledger	no	yes	no	no	no
MEDPID	Person identification	no	no	no	yes	no
MEDPRE	Medical prescription	no	no	no	yes	no
MEDREQ	Medical service request	no	no	no	yes	no
MEDRPT	Medical service report	no	no	no	yes	no
MEDRUC	Medical resource usage and cost	no	no	no	yes	no
MSCONS	Metered services consumption report	yes	no	no	no	no
ORDCHG	Purchase order change request	yes	no	yes	no	no
ORDERS	Purchase order	yes	no	yes	no	yes
ORDRSP	Purchase order response	yes	no	yes	no	no
OSTENQ	Order status enquiry	yes	no	no	no	no
OSTRPT	Order status report	yes	no	no	no	no
OSTRPTDS	Order status report (distributor to supplier)	no	no	yes	no	no
OSTRPTSD	Order status report (supplier to distributor)	no	no	yes	no	no
PARTIN	Party Information	yes	no	no	no	no
PAYMUL	Multiple payment order	yes	no	no	no	no
PRICAT	Price/sales catalogue	yes	no	yes	no	yes
PRODAT	Product data	yes	no	no	no	no
PRODOS	Key client Information	no	yes	no	no	no
PROINQ	Product inquiry	yes	no	no	no	no
QALITY	Quality	yes	no	no	no	no
QUOTES	Quote	yes	no	yes	no	no
RECADV	Receiving advice	yes	no	yes	no	no
REMADV	Remittance Advice	yes	no	yes	no	no
REQOTE	Request for quote message	yes	no	yes	no	no
RETANN	Announcement for returns	yes	no	no	no	no
RETINS	Instruction for returns	yes	no	no	no	no
SBIINV	Self-billing Invoice	no	no	yes	no	no

SLSFCT	Sales forecast	yes	no	no	no	no
SLSRPT	Sales data report	yes	no	yes	no	no
SSDCLM	Ship from stock and debit claim	no	no	yes	no	no
SSDRSP	Ship from stock and debit claim response	no	no	yes	no	no
TAXCON	Tax control	yes	no	no	no	no

Table 4.2: Major EDIFACT Message Types

4.5 Formats of Available Standards Listings

EDIFACT formats are provided in HTML files. Segment groups start on separate lines. Repetitions, optionality, and nesting level are indicated by trailing strings. Plus (+) signs indicate start/end of segment groups. Some commercial vendors (see Section 6.3) provide XML Schema (.xsd) files describing the format of EDIFACT message types.

EANCON formats are provided in HTML tables. Images indicate start and end of segment groups. Repetitions and optionality are indicated by trailing TD segments.

EUROSTAT provides a single message type which is specified in .pdf format. The message format probably is equivalent to that of the message in the full EDIFACT standard, which is in HTML. Restrictions on message subcomponents are only provided in .pdf format.

EDIFICE has most message types in HTML tables; however, a few are in .pdf format. Images indicate start and end of segment groups. Repetitions and optionality are indicated in sequential TD segments.

EDIFICAS has the format of only one of its seven message types publicly available. This format is only available as a diagram.

EMEDI does not provide separate message formats. However, this is only source of medical messages, so basic definitions from EDIFACT will suffice. Any further restrictions on segment and element formats are not publicly available.

ETIS publicly provides the formats of both of its message types in .pdf format. GREFIS provides a single message, whose format is publicly available in .pdf format. This is the only source of this message, so basic definitions from EDIFACT will suffice.

UNICORN is not freely available to the public. Its message types are distinct from EDIFACT message types; it merely uses EDIFACT message structure and subcomponents (data segments and data elements).

4.6 Ease of Automating Ontologization of Syntax of Standards

The process of ontologizing EDIFACT subsets will be performed in two steps:

- Ontologize the syntax (format) of the subset
- Ontologize the semantics (meaning) of the subset

Automating the ontologization of formats for messages whose syntax is provided in HTML should be relatively easy through the use of multiple editor scripts. Those formats which are provided only in .pdf format would need to be handled manually.

The use of a commercial tool to describe the format in XML will be explored. The output of such a tool might not be in a form that we can directly use, but conversion of this output to a desired form might be more straight forward than conversion of HTML encoded files. This output would probably be available only for the complete EDIFACT standard, so that it would have to be manually edited down to the restricted message formats from the individual subsets.

5 Relationship with ebXML

In this chapter we investigate the relationship between EDIFACT/X12 with the new generation of XML-driven EDI standards. For this purpose we exemplarily give an overview of ebXML, as one of the most representatives approaches in the field, analyze commonalities and differences to the EDIFACT series and outline several initiatives coping with their interoperability.

5.1 ebXML Overview

ebXML (electronic business XML) was created by OASIS and the United Nations/ECE agency CEFACT in collaboration with a large body of public institutions and major IT companies as an open technical framework enabling the consistent and systematic usage of XML for the exchange of electronic business data in B2B and B2C interactions. Following up the XML-centered advances in IT in the nineties, ebXML was originally developed for the purpose of providing a more scalable and cost-saving counterpart to EDI standards such as X12 and EDIFACT, which have predominantly become accepted in large companies and for point-to-point business interaction patterns.

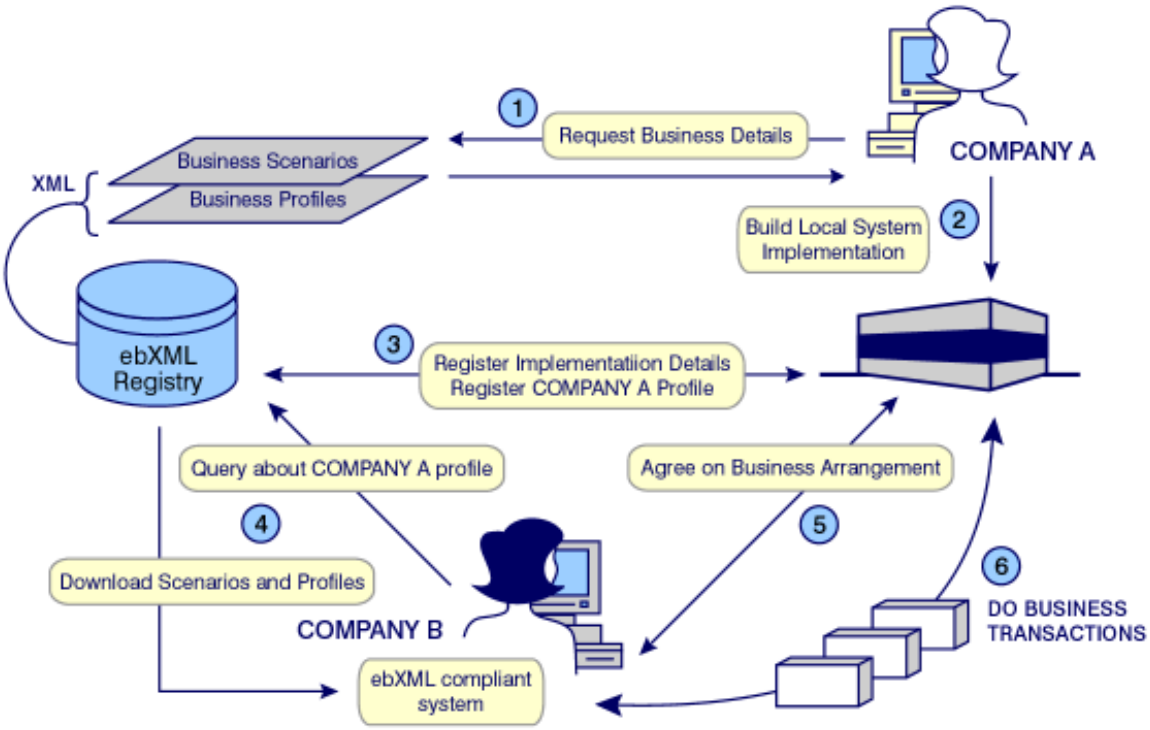

Figure 5.1: High-level Overview of ebXML-based Business Interactions cf. [24]

Figure 5.1 illustrates the main phases in the order in which they are supposed to be executed to perform business transactions on the basis of ebXML. The architecture depicted in this figure is based on the following concepts:

- a standard business process model described by the UN/CEFACT Modeling Methodology (UMM) [26].

- a registry and a repository storing business processes and information describing them, thus enabling their reusability. Additionally the repository captures so-called collaboration protocol profiles (CPPs). These contain information about each business party in terms of the supported processes, the interfaces by which these can be accessed, as well as messages to be exchanged and various security, encoding and transport constraints.
- a mechanism for discovering information about business participants and/or business processes and core components.
- collaboration protocol agreement (CPA): a means to automatically derive a contract between two or more businesses from the CPPs of the respective companies.
- core component library (CCL): a set of standard core components that may be used to assemble ebXML documents for arbitrary sectors and business transactions. They are described in a comprehensive technical specification by the UN/ECE CEFACT [25].
- business messaging service ensuring a secure and reliable exchange of messages between partners.

Company A in Figure 5.1 reviews the contents of an ebXML registry in order to determine the requirements for an ebXML implementation and the feasibility of an ebXML-based approach for their business needs. If found appropriate, the next step after implementing the ebXML system at Company A is to create and register a CPP with the registry. The CPP will contain the information necessary for a potential partner to identify the type of business in which Company A is interested, and the technical details required for a business interaction to be established. Company B, which is potentially interested in a collaboration with the newly registered Company A, is able to automatically negotiate a protocol agreement with Company A, given the profiles stored in the repository and the ebXML specifications which are realized in the ebXML implementations. In a final step, the two business parties exchange messages using the messaging service provided by the ebXML systems and specified in the ebXML recommendations.

5.2 ebXML vs. EDIFACT/X12

Compared to traditional EDI approaches, ebXML provides a series of important technical, economic and process-related benefits:

Technology: XML and XML Schema form the basis for a structured, easily processable, maintainable and extendible description of the syntax and semantics of the exchanged data. Tools required to process and manage the ebXML data are available at negligible costs and high quality due to the popularity of XML-based technologies and the openness of the standards.

The architecture of ebXML is explicitly targeted at business relationships situated in the dynamic environment of the Web: it scales with respect to the number of parties involved in the business transaction by limiting the number of messages to be exchanged and supports the modular assembly of business documents using automatic discovery mechanisms for core components (see below).

Economic: Apart from the high level of expertise required to use EDIFACT/X12 and the costs associated with the deployment of the associated software, the costs involved in establishing many-to-many business relationships grow quadratically with respect to the number of business parties involved. The high start-up and deployment costs made EDIFACT impracticable for small and medium-sized enterprises (SMEs), which is one of the target audience groups for XML-based EDI.

Process: The systematic interaction between business partners is supported in ebXML by means of an explicitly defined process model, a dedicated messaging/transport standard and a standard for binding business agreements. This complies with the rapid business process re-engineering practices to date.

From a content perspective, ebXML proposes standard instruments for exchanging business messages and for negotiating agreements between business partners.

5.2.1 UMM

The Unified Modelling Methodology is intended to provide a framework for the systematic development of UN/CEFACT core components and business information entities. [1] UMM contains several components:

- methodology to capture business process knowledge at a conceptual level and independently of any implementation details.
- business process and a comprehensive business information meta-model which formally describe the modelling process and are used by the
- process analysis methodology, which allows for a controlled process operation and quality assurance.
- methods to discover and define reusable process and information descriptions, thus enforcing reusability.

The language used in the methodology is an extension of UML.

Many of the most recent UN/CEFACT standards are the product of using UMM or one of its components. The UN/CEFACT Core Component Library is an implementation of the UN/CEFACT Modelling Methodology dictionary concept. The business components collected in the CCL are instantiations of the abstract UMM information entity concept.

5.2.2 CCTS

CEFACT's Core Components Technical Specification is an attempt to cope with the widely known lack of information interoperability between eBusiness application systems. By contrast to the traditional approach, according to which an EDI standard is a collection of static message types, which need to be mapped to message types covered

[1] Annex I used the term "CEFACT" to refer to an ontology. CEFACT developed the ebXML ontology as well as CCTS. These are discussed separately in this document (the detailed work plan) and the term "CEFACT" is no longer used to mean an ontology.

by external resources, ebXML is built upon a common set of precisely defined business components. These represent the general types of business data currently in use and can be assembled and refined to create new message types for various purposes.

The modularity of this CCTS approach allows the identification of similarities and differences between message types at a higher level of accuracy, while parts which are not compliant can be replaced instead of dismissing the entire artefact as being incompatible.

The specification differentiates between basic and association components and between components and business entities, which are to be understood as refinements of the former in a particular business context. Furthermore, the document provides details on how to build new components and entities, focusing on the reusability of the existing library. For this purpose it describes how to look for these and to decide whether they are appropriate in a new application setting.

5.3 Interoperability Initiatives

To the knowledge of the authors the question of interoperability between EDIFACT/X12 and ebXML has been marginally addressed by eBusiness solution providers so far. A series of initiatives address the related issue of transforming the ASCI-based EDI standards to an equivalent XML format.

The European ISIS XML/EDI initiative investigated means to map EDI standards such as EDIFACT to XML[2]. Several pilot projects in sectors such as health care and transportation created example XML-based representations of several EDIFACT message types and a data validation module for the associated DTDs. They summarized their experience in a suite of guidelines for creating XML DTDs to represent EDI messages. These are likely to be useful for the objectives of this workpackage, at least at the level of syntax.

The XML::EDIFACT project provides a `perl` module which transforms UN/EDIFACT messages into human readable XML. It uses the original terms from the EDIFACT batch directories as markup and the defining document as namespaces[3].

The commercial software suite Stylus Studio provides similar functionality.[4] The tool generates XML Schemas from EDIFACT message types and validates EDIFACT data with respect to these schema definitions.

Other commerical tools such as ExtremeTranslator and Altova's MapForce EDI Mapping [5] provide similar capabilities.

[2] `http://palvelut.tieke.fi/edi/isis-xmledi/synopsis.htm` last visited in September, 2006

[3] `http://www.xml-edifact.org` last visited in September, 2006

[4] `http://www.stylusstudio.com/edifact_to_xml.html` last visited in September, 2006

[5] `http://www.xtranslator.com/`, `http://www.altova.com/products/mapforce/edi_to_xml_mapping.html` last visited in September, 2006

6 Related and Prior Work Ontologizing EDI Systems

6.1 Not "Worth the Effort"

In the mid-1990s the task of ontologizing EDI systems was judged too great to be worth the effort [23]. Additional arguments against the task were the instability of EDI systems and the multiplicity of variations of the standards.

6.2 Ontologizing X12

6.2.1 Overview

DERI Galway started to ontologize the ANSI X12 EDI system in 2004. They designed ontologies for expressing both the syntax of EDI messages [14] and templates for the meaning encoded in messages [15].

An ontology for specifying X12 formats was encoded in WSML and other semantic languages (OWL, TRIPLE, CycL, FLORA). An ontology for specifying templates for the syntax of X12 messages, based on the syntactic (format) ontology, was encoded in two of these languages (WSML and CycL). These ontologies are general enough to be used for specifying EDIFACT formats as well as X12 formats.

The syntax of over 100 X12 messages was encoded using the syntactic ontology. This encoding was performed to a great extent through the use of shell scripts to edit the HTML descriptions of the formats. A manual effort was needed to clean up the message formats, although the formats of standardized message components, "data segments" and "data elements" was purely automatic.

Semantic templates for the meaning of three of the message types (along with their subcomponents) were created manually as a proof of concept. The operation of a data mediator using these templates on an X12 message was simulated. No data mediator was written.

6.2.2 Output Data Formats

The WSML ontology for X12 EDI uses the concepts `TransactionSet` (Message type), `DataSegment`, `DataElement`, `FormatDescriptor`, `Format`, `FormatCode`, `FormatRestriction`, and `SegmentElementRestriction`.

An example description of a short message type and each type of component using these concepts (adapted from [13]) follows

```
// define a message type (TransactionSet) with given Format Descriptor
instance X12_TS_997 memberOf edi:TransactionSet
    nonFunctionalProperties
      dc:description hasValue "Functional Acknowledgement Transaction Set"
    endNonFunctionalProperties
    hasFormatDescriptor hasValue X12_TS_997FD

// define format for the five components of the Format Descriptor
```

```
instance X12_TS_997FD memberOf edi:FormatDescriptor
    hasFormat hasValue X12_TS_997F
    formats1stComponent hasValue ds:X12_ST_DS   // Data Segment
    formats2ndComponent hasValue ds:X12_AK1_DS
    formats3rdComponent hasValue X12_AK2_Loop   // Data Segment Loop
    formats4thComponent hasValue ds:X12_AK9_DS
    formats5thComponent hasValue ds:X12_SE_DS

// define format for the five components of the Format
instance X12_TS_997F memberOf edi:Format
  formatFor1stComponent hasValue fc:FC_M1
  formatFor2ndComponent hasValue fc:FC_M1
  formatFor3rdComponent hasValue fc:FC_O1
  formatFor4thComponent hasValue fc:FC_M1
  formatFor5thComponent hasValue fc:FC_M1
// define a Data Segment and relate it to its Format Descriptor

// From the Data Segment (ds:) ontology:
instance X12_AK1_DS memberOf edi:DataSegment
    hasFormatDescriptor hasValue X12_AK1_DSFD

// define format for the three components of the DSs Format Descriptor
instance X12_AK1_DSFD memberOf edi:FormatDescriptor
    hasFormat hasValue X12_AK1_DSF
    formats1stComponent hasValue de:X12_DE_479
    formats2ndComponent hasValue de:X12_DE_28
    formats3rdComponent hasValue de:X12_DE_480

// From the Data Element (de:) ontology
instance X12_DE_479 memberOf edi:DataElement
    hasDataElementFormat hasValue _string
    minLength hasValue 2
    maxLength hasValue 2

// From the FormatCondition (fc:) ontology
instance FC_M1 memberOf edi:FormatCode
    optionality hasValue edi:Mandatory
    maxRepeats hasValue 1

// A data segment with format restrictions
// From the Data Segment (ds:) ontology:
instance X12_CUR_DS memberOf edi:DataSegment
    hasFormatDescriptor hasValue X12_CUR_DSFD
    formatRestriction hasValue fr:FR_C1110
    formatRestriction hasValue fr:FR_L101112

// From the FormatRestriction (fr:) ontology
instance FR_C1110 memberOf edi:FormatRestriction
```

```
    restrictionType hasValue edi:FR_Conditional
    restrictionLine1 hasValue 11
    restrictionLine2 hasValue 10

instance FR_Conditional memberOf edi:FormatRestrictionType
// If the first mentioned component is present, then the rest must be also.


// format restrictions also apply to message types:
instance X12_TS_855 memberOf edi:TransactionSet
    hasFormatDescriptor hasValue X12_TS_855FD
    formatRestriction hasValue fr:X12_FR_TS855a

instance X12_FR_TS855a memberOf edi:SegmentElementRestriction
    restrictionType hasValue fr:FR_ExclusiveUse
    segmentNumber hasValue 17
    elementNumber hasValue 1
// If the segment is present, its nth element is absent.
```

6.2.3 Applicability

The ontologies created for specifying EDI formats and semantic templates appear to be usable in their current forms to specify EDIFACT message format and meaning. A data mediator would have to be written to use the generated ontologies to create WSML versions given EDIFACT messages. As neither use case (WP 8A or WP 8B) will be working with legacy systems which transmit EDIFACT messages, the creation of such a data mediator is not necessary for TripCom.

6.3 Commercial EDIFACT to XML Conversion

Several vendors provide software to translate EDIFACT (or other format) EDI messages into XML. The XML which they create could be converted into a semantic ontology used by Web services through the use of an adaptor.

One such tool is discussed here.

6.3.1 Stylus Studio

One such company is DataDirect Technologies. Its Stylus Studio® Enterprise Edition includes an EDIFACT to XML Schema conversion tool which creates annotated XML Schema (.xsd) files describing EDIFACT message types and components and which converts EDIFACT messages to XML as defined by these files. The output is in terms of data structures, not the semantics of the message.

The tool does not include descriptions of the standard subsets of EDIFACT, but provides an XML editing tool to allow a user to convert the xsd file from a generic description of an EDIFACT message type to that of a desired subset. It does have each of the full versions of each message type as released (normally) twice a year.

Other conversions performed by the tool are XML to EDIFACT as well as conversions between X12 and XML. The XML for converted EDIFACT and X12 files are different.

Validation of correct EDI message formats are also included.

A system of semantic templates for conversion of XML to the target language would have to be developed to be used by the adaptor.

6.3.2 Output Data Formats

Stylus Studio® enables the creation of XML schema documentation (.xsd) files for individual message types, creating complex types for each data element, data segment, data segment group, and message type.

An extract from the sample .xsd file they provide for the INVOIC message[1] follows:

```
<xsd:complexType name="INVOIC">
<xsd:annotation>
<xsd:appinfo>D96A INVOIC</xsd:appinfo>
<xsd:documentation>Invoice message</xsd:documentation>
</xsd:annotation>
<xsd:sequence>
<xsd:element name="UNB" type="UNB" minOccurs="0" maxOccurs="1"/>
<xsd:element name="UNH" type="UNH" minOccurs="1" maxOccurs="1"/>
<xsd:element name="BGM" type="BGM" minOccurs="1" maxOccurs="1"/>
<xsd:element name="DTM" type="DTM" minOccurs="1" maxOccurs="35"/>
<xsd:element name="PAI" type="PAI" minOccurs="0" maxOccurs="1"/>
<xsd:element name="ALI" type="ALI" minOccurs="0" maxOccurs="5"/>
<xsd:element name="IMD" type="IMD" minOccurs="0" maxOccurs="1"/>
<xsd:element name="FTX" type="FTX" minOccurs="0" maxOccurs="10"/>
<xsd:element type="GROUP_1" name="GROUP_1" minOccurs="0" maxOccurs="unbounded"/>
<xsd:element type="GROUP_2" name="GROUP_2" minOccurs="0" maxOccurs="unbounded"/>
<xsd:element type="GROUP_6" name="GROUP_6" minOccurs="0" maxOccurs="5"/>
<xsd:element type="GROUP_7" name="GROUP_7" minOccurs="0" maxOccurs="5"/>
<xsd:element type="GROUP_8" name="GROUP_8" minOccurs="0" maxOccurs="10"/>
<xsd:element type="GROUP_9" name="GROUP_9" minOccurs="0" maxOccurs="10"/>
<xsd:element type="GROUP_12" name="GROUP_12" minOccurs="0" maxOccurs="5"/>
<xsd:element type="GROUP_13" name="GROUP_13" minOccurs="0" maxOccurs="unbounded"/>
<xsd:element type="GROUP_15" name="GROUP_15" minOccurs="0" maxOccurs="unbounded"/>
<xsd:element type="GROUP_22" name="GROUP_22" minOccurs="0" maxOccurs="unbounded"/>
<xsd:element type="GROUP_23" name="GROUP_23" minOccurs="0" maxOccurs="1"/>
<xsd:element type="GROUP_24" name="GROUP_24" minOccurs="0" maxOccurs="1"/>
<xsd:element type="GROUP_25" name="GROUP_25" minOccurs="0" maxOccurs="unbounded"/>
<xsd:element name="UNS" type="UNS" minOccurs="1" maxOccurs="1"/>
<xsd:element name="CNT" type="CNT" minOccurs="0" maxOccurs="10"/>
<xsd:element type="GROUP_48" name="GROUP_48" minOccurs="1" maxOccurs="unbounded"/>
<xsd:element type="GROUP_50" name="GROUP_50" minOccurs="0" maxOccurs="10"/>
<xsd:element type="GROUP_51" name="GROUP_51" minOccurs="0" maxOccurs="15"/>
<xsd:element name="UNT" type="UNT" minOccurs="1" maxOccurs="1"/>
```

[1] http://www.stylusstudio.com/examples/invoice.zip

```
<xsd:element name="UNZ" type="UNZ" minOccurs="0" maxOccurs="1"/>
</xsd:sequence>
</xsd:complexType>
<xsd:complexType name="GROUP_1">
<xsd:sequence minOccurs="1" maxOccurs="1">
<xsd:element name="RFF" type="RFF" minOccurs="1" maxOccurs="1"/>
<xsd:element name="DTM" type="DTM" minOccurs="0" maxOccurs="5"/>
</xsd:sequence>
  ...
<xsd:complexType name="RFF">
<xsd:sequence>
<xsd:element name="RFF01" minOccurs="1" maxOccurs="1">
<xsd:complexType>
<xsd:sequence>
<xsd:element name="RFF0101" type="type1153" minOccurs="1" maxOccurs="1"/>
<xsd:element name="RFF0102" type="type1154" minOccurs="0" maxOccurs="1"/>
<xsd:element name="RFF0103" type="type1156" minOccurs="0" maxOccurs="1"/>
<xsd:element name="RFF0104" type="type4000" minOccurs="0" maxOccurs="1"/>
</xsd:sequence>
</xsd:complexType>
</xsd:element>
</xsd:sequence>
</xsd:complexType>
  ...
<xsd:simpleType name="type1153">
<xsd:annotation>
<xsd:documentation>Reference qualifier</xsd:documentation>
</xsd:annotation>
<xsd:restriction base="xsd:string">
<xsd:minLength value="0"/>
<xsd:maxLength value="3"/>
<xsd:enumeration value="AAA">
<xsd:annotation>
<xsd:documentation>Acknowledgement of order number</xsd:documentation>
</xsd:annotation>
</xsd:enumeration>
<xsd:enumeration value="AAB">
<xsd:annotation>
<xsd:documentation>Proforma invoice number</xsd:documentation>
</xsd:annotation>
</xsd:enumeration>
      ...
</xsd:restriction>
  </xsd:simpleType>
  ...
```

6.3.3 Applicability

A tool such as this for transforming EDIFACT to XML can be beneficial for WP7. This could be a first step towards the final goal of this workpackage, since translating from XML to the ontology language we decide to use could be more comfortable for programmers who have experience in converting XML sources and are used to thinking in terms of DOM objects, but may have difficulties in writing arbitrary parsers.

However, such a tool is by no means a panacea. The semantics of the messages in the selected EDIFACT subsets is far simpler than that produced by such a tool, which provides all options used by all varients of the messages. Such tools do not assist in defining the semantics of messages or their parts.

7 Conclusion

EDIFACT standardizes messages for a wide variety of industries and business purposes. Individual industries have defined subsets of these standards, limiting the types and options in the messages which they use. The majority of these message types and industries are not applicable for the TripCom use cases, although they might be applicable for future use of a widely installed triple space based on TripCom.

A number of more recent standards are being based on XML, allowing them to be more readily readable by people at the expense of increasing the message size. Different standards are being created for different industries, generally without an attempt at compatibility with either each other or legacy EDI systems.

We have selected to ontologize subsets of the EDIFACT standard related to the TripCom use cases, broadening the coverage as circumstances allow. The standards that we have selected are publicly available and presented in a format which allows for partial automatization of the definition of their formats. The subsets of EDIFACT selected appear to cover the messages needed by the EIA use case (WP 8A). Patient summary data (needed by WP 8B) is not covered by EDIFACT, but the medical messages developed by EMEDI for EDIFACT cover some of this material.

Commercial tools for converting EDIFACT messages to XML could be beneficial for specifying formats and for generating output for use by a data mediator, however, they would not be useful in ontologizing the meanings of EDIFACT messages. As the use cases do not envision using a legacy system which transmits and/or receives EDIFACT messages, no data mediator needs to be written at this point.

An existing system for specifying templates for the meanings of EDI messages appears to be usable in its current form to specify EDIFACT message format and meaning.

The ontologization process for Work Package 7, based on this analysis is discussed in detail in the original TripCom work plan (Annex I) and in the Detailed Work Plan (Deliverable 10.1).

Relevant Web Sites

EANCOM	http://www.ean.se/EANCOM_2002/ean02s4/experts
EDIFICAS	http://www.edificas.org
EDIFICE	http://repository.edifice.org/migs/index.htm
UN/EDIFACT	http://www.unece.org/trade/untdid/welcome.htm
ETIS	http://www.etis.org/activities/ebg.asp
ExtremeTranslator	http://www.xtranslator.com/
ISIS XML/EDI	http://palvelut.tieke.fi/edi/isis-xmledi/synopsis.htm
Stylus Studio	http://www.stylusstudio.com/xml_product_index.html
TripCom	http://www.tripcom.org
UNICORN	http://www.analysis.gr/english/html/news/unicorn.htm
WSMO	http://www.wsmo.org
XML::EDIFACT	http://www.xml-edifact.org

References

[1] J. Berje. The EDIFACT Standards. *NCC Blackwell*, 1991.

[2] Book Industry Communication. BIC-EDI Standards and Implementation Guidelines. http://www.bic.org.uk, 2005.

[3] European Chemical Industry Council. Cefic Research and Science - Annual review 2005. http://www.cefic.org, 2005.

[4] D. Crocker. Mime Encapsulation of EDI Objects, TC RFC1767. http://www.rfc-editor.org, 1995.

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

[6] Chemical Industry Data Exchange. Chem eStandards v3.0. http://www.cidx.org, 2005.

[7] The International Group for Electronic Commerce in the Book and Serial sectors. EDItEUR Implementation Guidelines. http://www.editeur.org, 2005.

[8] United Nations Economic Commission for Europe. http://www.unece.org, 2005.

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

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

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

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

[13] D. Foxvog. D27v0.1 EDI Ontology. http://www.wsmo.org/TR/d27/v0.1/, 2005.

[14] D. Foxvog and C. Bussler. Ontologizing EDI: First Steps and Initial Experience. Proceedings International Workshop on Data Engineering Issues in E-Commerce (DEEC 2005), 2005.

[15] D. Foxvog and C. Bussler. Ontologizing EDI Semantics. Proceedings First International Workshop on Ontologizing Industrial Standards (OIS 2006), 2006.

[16] T. Harding and R. Scott. FTP Transport for Secure Peer-to-Peer Business Data Interchange over the Internet. *www.ietf.org*, 2006.

[17] Travel Technology Initiative. The Unicorn Application Message Standard. http://www.tti.org, 2005.

[18] European Medical Electronic Data Interchange. EMEDI overview. http://www.accart.nom.fr, 2003.

[19] ISO. Open-EDI Reference Manual. ISO/IED JTD 1/SC30 ISO Standard 14662, 1997.

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

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

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

[23] F. Lehmann. Machine-negotiated Ontology-based EDI. Electronic Commerce: Current Research Issues and Applications, 1996.

[24] OASIS. ebXML Technical Architecture Specification v1.0.4. www.ebxml.org/specs/ebTA.pdf, 2001.

[25] OASIS. Core Components Technical Specification Part 8 of the ebXML Framework v2.0.1. http://www.unece.org/cefact/ebxml/CCTS_V2-01_Final.pdf, 2003.

[26] OASIS. UN/CEFACT Modeling Methodology (UMM) User Guide. http://www.unece.org/cefact/umm/UMM_userguide_220606.pdf, 2003.

[27] European Committee of Domestic Equipment Manufacturers. CECED statement press release. http://www.ceced.org, 2005.

[28] European Committee of Domestic Equipment Manufacturers. EDI-White messages Guidelines. http://www.ceded.org, 2005.

[29] International Committee on EDI for Serials. ICEDIS Messages and Implementation Guidelines. http://www.ecedis.org, 2005.