

Preliminary
EDI (X12 to UN/EDIFACT) Mapping Concept Paper
for the
DOD CALS IDE Project

June 16, 1997

Submitted by
ManTech Advanced Systems International, Inc.
West Virginia Technology Applications Operations Center
1000 Technology Drive, Suite 3310
Fairmont, West Virginia 26554

In support of
Contract DASW01-97-D-0006

Non-CDRL
Property of ManTech

MANAGEMENT
TECHNOLOGY MASI
ManTech Advanced Technology Systems

Robert S. Kidwell
Technical Director
DoD CALS IDE Project

Jack G. Richman
Project Manager
DoD CALS IDE Project

TABLE OF CONTENTS

LIST OF FIGURES	iii
LIST OF TABLES.....	iv
1.0 BACKGROUND.....	5
1.1 Introduction.....	5
2.0 COMPATIBILITY ISSUES AND BARRIERS.....	7
2.1 Direct Techniques for Comparing ANSI X12 to UN/EDIFACT.....	7
2.2 Establishing an Intermediate Foundation for Assessment.....	9
3.0 STANDARDIZED MAPPING METHODOLOGY.....	12
3.1 Universal Data Element Framework.....	12
3.2 Semantic Unit Identification and UDE Code Development.....	15
3.3 Semantic Equivalency Assessment Using UDEF.....	18
3.4 Semantic Equivalency Assessment Automation.....	25
4.0 SUMMARY AND CONCLUSIONS.....	26
REFERENCES	27
APPENDIX A: GLOSSARY OF TERMS.....	28
APPENDIX B: ABBREVIATIONS AND ACRONYMS.....	34
APPENDIX C: <i>CONVEYOR</i> APPLICATION.....	35
APPENDIX D: METHODOLOGY EVALUATIONS.....	39
D.1 Matrix Evaluation of Methodologies.....	42
D.2 Basic Semantic Repository.....	43
D.3 OPEN-EDI.....	44
D.4 Defense Data Dictionary System.....	45
D.5 Defense Procurement Corporate Information Management Systems Center.....	45
D.6 UDEF.....	47
APPENDIX E: REFERENCES.....	48
APPENDIX F: UDEF SHORTFALLS.....	50

LIST OF FIGURES

FIGURE 2.1-1: DIRECT MAPPING COMPARISON TECHNIQUE	8
FIGURE 2.1-2: DIRECT DATA ELEMENT COMPARISON TECHNIQUE	9
FIGURE 2.2-1: EQUATION FOR INFORMATION AS A FUNCTION OF DATA AND SEMANTICS	10
FIGURE 2.2-2: INTERMEDIATE DATA ELEMENT COMPARISON TECHNIQUE.....	11
FIGURE 3.1-1: ISO 11179 NAMING CONVENTION.....	13
FIGURE 3.2-1: INTERMEDIATE DATA ELEMENT COMPARISON TECHNIQUE.....	16
FIGURE 3.2-2: UDE CODE ASSIGNMENT PROCESS	17
FIGURE 3.2-3: UDE CODE TREE	18
FIGURE C-1: <i>CONVEYOR</i> APPLICATION	36
FIGURE C-2: <i>CONVEYOR</i> APPLICATION OPTIONS	37
FIGURE D-1: METHODOLOGY INTERRELATIONSHIPS	40

LIST OF TABLES

TABLE 3.1-1: UDEF OBJECT DEFINITIONS	14
TABLE 3.1-2: UDEF PROPERTY DEFINITIONS.....	14
TABLE 3.3-1: ANSI X12 BEG SEGMENT DATA ELEMENTS.....	19
TABLE 3.3-2: SEMANTIC MEANING DEFINED FOR ANSI X12 BEG SEGMENT ELEMENTS	19
TABLE 3.3-3: SEMANTIC UNIT LABELS.....	20
TABLE 3.3-4: UDE CODES.....	20
TABLE 3.3-5: UN/EDIFACT BGM SEGMENT DATA ELEMENTS	21
TABLE 3.3-6: SEMANTIC MEANING DEFINED FOR UN/EDIFACT BGM SEGMENT ELEMENTS	22
TABLE 3.3-7: UN/EDIFACT DTM SEGMENT DATA ELEMENTS	22
TABLE 3.3-8: SEMANTIC MEANING DEFINED FOR UN/EDIFACT DTM SEGMENT ELEMENTS	22
TABLE 3.3-9: UN/EDIFACT RFF SEGMENT DATA ELEMENTS	23
TABLE 3.3-10: SEMANTIC MEANING DEFINED FOR UN/EDIFACT RFF SEGMENT ELEMENTS	23
TABLE 3.3-11: UN/EDIFACT PAI SEGMENT DATA ELEMENTS.....	23
TABLE 3.3-12: SEMANTIC MEANING DEFINED FOR UN/EDIFACT RFF SEGMENT ELEMENTS	23
TABLE 3.3-13: UDE CODES FOR SEMANTIC UNIT LABELS.....	24
TABLE D.1: MATRIX EVALUATION OF METHODOLOGIES.....	42

Preliminary EDI (X12 to UN/EDIFACT) Mapping Concept Paper

1.0 BACKGROUND

There are currently two widely accepted EC/EDI standards in use in the global marketplace: ANSI X12 for use within the U.S., and UN/EDIFACT for international trade. Within the U.S., both industry and government are expressing interest in accelerating migration to or alignment with the United Nations EDIFACT standards (for international electronic data interchange) for business functions such as purchase orders, etc. In fact, the U.S. government has committed to an eventual migration to the UN/EDIFACT standard. This has resulted in a need for harmonization between these overlapping electronic data interchange standards. However because of fundamental differences between the data interchange format and the disparity of business functionality between the two standards, achieving compatibility between them presents a complex and challenging task.

1.1 Introduction

ManTech has been directed by the DOD, as part of the CALS IDE project, to assess the semantic equivalency between certain UN/EDIFACT messages with ANSI X12 transaction sets, and to test the feasibility of using the UN/EDIFACT ORDERS message for Basic Ordering Agreements and simple delivery orders within the DOD's Electronic Commerce (EC) infrastructure. These tasks have prompted the mapping methodology that is presented in this report. The report provides an overview of the compatibility issues associated with the ANSI X12 and UN/EDIFACT EDI standards and then proposes a methodology that will provide information analysts and subject matter experts a coherent approach to linking and eventually combining these two standards. The paper first introduces the concept of compatibility barriers between the two EC/EDI standards. These barriers are defined as those differences between the ANSI X12 and UN/EDIFACT standards that are significant enough to create an incompatibility between the two standards. The disparity between the current set of ANSI X12 standards and the UN/EDIFACT standards has hindered the interoperability across the two standards and is slowing the eventual migration to UN/EDIFACT. This paper therefore proposes a methodology for linking the two standards. The paper uses an example scenario for directly mapping from ANSI X12 to UN/EDIFACT to reveal some of the inherent pitfalls that necessitate a more sophisticated approach to this problem. The paper then discusses these pitfalls, e.g. accounting for the *One-to-Many* mapping relationships, etc., and then proposes an intermediate mapping solution. The importance of introducing an intermediate mapping step is emphasized in the paper, i.e. it enables comparing the data elements on a common ground using neutral terms that is independent from either standard. The paper then introduces the concept of semantic units, which is defined as the smallest unique and repeatable functional, physical, or conceptual unit of meaning that may be applied within a particular domain. It will be shown that these semantic units are the lowest common denominator for the interoperability of the two EDI standards, and how these semantic units can act as an intermediary for comparing the two standards.

Linking the two standards is a multi-step process that is accomplished by first analyzing the ANSI X12 functional data elements and then identifying semantic units from these data elements. This paper presents a detailed methodology for identifying, codifying, and mapping units of business functionality from the ANSI X12 standard to the UN/EDIFACT standard. The methodology we have selected is the Universal Data Element Framework (UDEF). The UDEF methodology is based on a predefined set of primary *objects* and *properties*, and *rules* that may be used to generate unique codes and encapsulate the semantic information pertaining to data element(s). The UDEF provides a standardized, consistent, compact, and object-oriented framework within which this work may be performed. Utilizing this methodology, the semantic compatibility of ANSI X12 transaction sets to UN/EDIFACT messages may be assessed in a uniform and simplified manner.

Once a semantic unit relationship exists from the ANSI X12 data element(s) to the semantic unit, a similar relationship may be established from the semantic unit to the UN/EDIFACT data element(s), that is assuming equivalent data element(s) exist in the UN/EDIFACT standard. The entire UDEF process is explained using the same example scenario referred to in the previously mentioned. Additionally, the design and development an automated software tool that assists in this mapping process is discussed in this report.

2.0 COMPATIBILITY ISSUES AND BARRIERS

In spite of the fact that the ANSI X12 and the UN/EDIFACT standards possess striking similarities such as: (1) both are comprised of tagged and delimited American Standard Code for Information Interchange (ASCII) character strings, (2) both utilize a data dictionary of standard (global) business data elements and data segments, and (3) both contain predefined message types (corresponding to particular business functions), which are composed of a specified sequence of data segments, which in turn contain combinations of related data elements, there still however remains significant barriers that stand in the way of compatibility between the two standards. The primary barrier between the two standards currently is the lack of direct semantic equivalence, which is illustrated in the following section.

“Although the two standards possess striking similarities, the subtle differences between the two are so significant the standards are currently not compatible.”

2.1 Basic Direct Techniques for Comparing ANSI X12 to UN/EDIFACT

Although it may seem instinctive to compare X12 transaction sets to UN/EDIFACT messages by selecting particular

“Simply selecting particular segments from each standard and comparing the data elements from one standard to the data elements of the other standard is not adequate for assessing the compatibility between the two, because semantic differences exist in the basic construction of information between the two.”

segments from each standard and simply comparing data elements from one standard to data elements of the other standard, this approach has inherent problems. The problem with this comparison technique is that there is not a direct correspondence of data elements across the two standards. Even attempting to compare the two standards at segment level reveals a lack of direct correspondence across the two standards. This compatibility problem exists for three reasons: (1) differences in design philosophies espoused by each of the standards bodies, (2) differing national and international business requirements, and (3) inherent linguistic differences between North America and Europe, and others.

In order to gain insight, into the aforementioned compatibility problems observe the following mapping scenario where we begin to compare the ANSI X12 Purchase Order transaction set to the UN/EDIFACT ORDERS message. The mapping comparison begins by identifying and selecting a key part of the Purchase Order transaction set that will be used for comparison in the ORDERS message. The BEG segment is used in the ANSI X12 standard to indicate the beginning of the Purchase Order transaction set and transmit identifying numbers and dates. This is an important segment, which is mandatory for every Purchase Order transaction. In the following example (Figure 2.1-1), an attempt is made to directly map the ANSI X12 BEG segment to the corresponding UN/EDIFACT segment(s). As a criteria for determining if the data elements from ANSI X12 are equivalent to those from UN/EDIFACT, we consulted the respective standards documents for their definitions (see Section 3.3 for definitions).

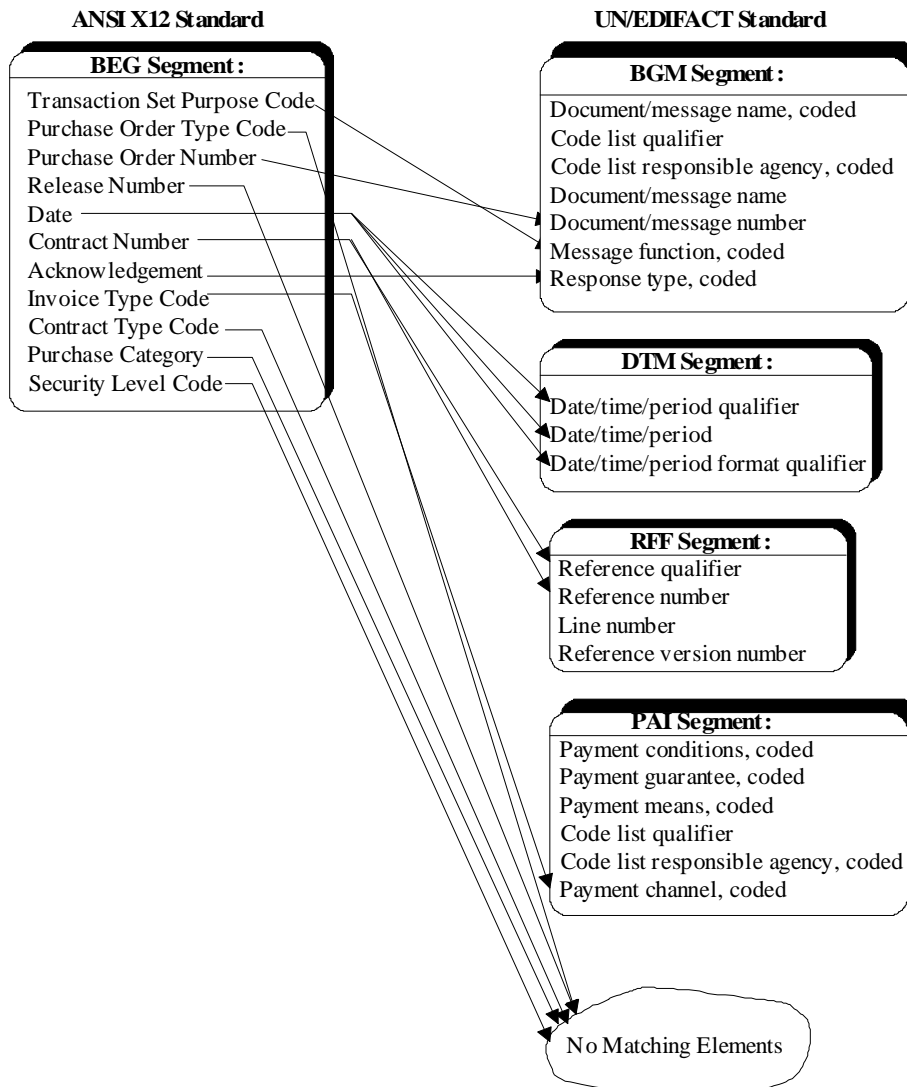


Figure 2.1-1: Direct Mapping Comparison Technique

When considering respective segments to map to in the UN/EDIFACT standard, the BGM segment, at first glance, appears to be equivalent to ANSI X12 BEG segment. The BGM segment is the beginning of message segment, which is used to indicate the type and function of a message and to transmit the identifying number for the message. However upon further investigation, it is revealed that while the BEG segment is only used for the ANSI X12 Purchase Order transaction set, the BGM segment is used more generically and is used in nearly all of the UN/EDIFACT messages. This fundamental message design difference necessitates including three additional segments from within the header section of the ORDERS message in order to account for all the respective data elements from the Purchase Order transaction set.

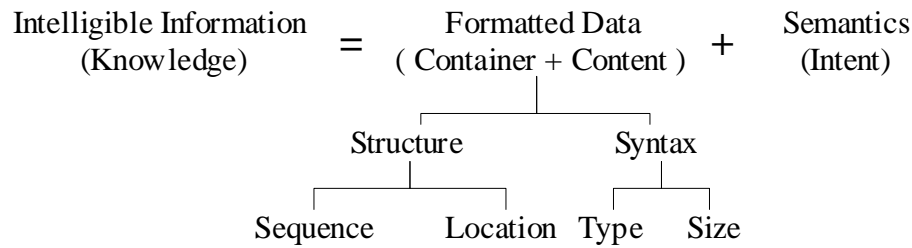


Figure 2.2-1: Equation for Information as a function of Data and Semantics

Considering the above definition, and recalling the mapping problems revealed in the previous section, the starting point for assessing the compatibility of the ANSI X12 standard to the UN/EDIFACT standard is to focus purely on the semantics. The semantics of the business functions and sub-functions are basically not affected by the constraints imposed by each of the EDI standards. In taking this approach, it is possible to assess the semantic equivalency between the two standards on even terms without giving up the ability to account for differences between corresponding data elements.

Establishing the relationship between data elements for the corresponding beginning message segments in the previous example (illustrated in Figure 2.1-1) was an arcane and arduous process. The core of this process involved determining the exact semantic definition of the ANSI X12 data element(s) and then searching through the UN/EDIFACT message for semantically equivalent data element(s). Both the ANSI X12 and UN/EDIFACT standards use data elements as their basic building blocks. These data elements are referenced by a data element name, e.g. Purchase Order Number, Transaction Set Purpose Code, etc. The words comprising the data element name generally provide a self-explanatory definition of the semantic purpose of the data element, however formal, more precise, semantic definitions do exist in the standards for each data element specified. Nevertheless, these definitions are often times not specific enough for the particular context in which the data element is used. This is exemplified in Figures 2.1-2 and 2.2-2 by the **date** data element. Neither the data element name or the definition is specific enough to convey the full intent of the data element, i.e. date for this example is really the purchase order date assigned by the purchaser. We will show how to capture the full semantic meaning of the **date** data element in Section 3.2 of this paper.

Several factors come in to play when determining the semantic intent of data elements in an EDI transaction, for instance which segment the element is in, the values for data elements, which transaction set or message the segment is in, etc. It is not uncommon to see the same data elements reused across multiple segments and or across multiple transactions, with slightly different semantics resulting from the context in which the element is used. Furthermore, to complicate matters, these factors vary from one standard to the other. However, basic business functions will always include a fundamental set of semantic elements that have the same meaning irrespective of the standard governing the data format.

This assessment will therefore proceed by first identifying intermediate semantic units that will be used to by-pass the compatibility barriers and link the two standards as illustrated in the Figure 2.2-2 below.

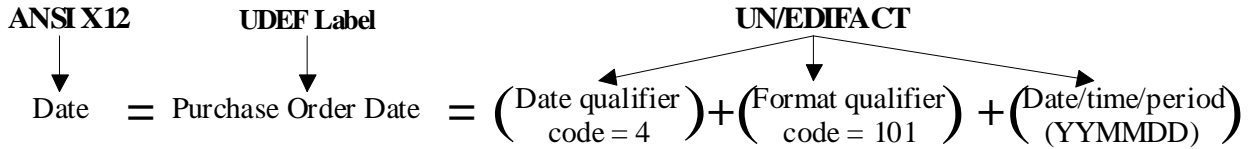


Figure 2.2-2: Intermediate Data Element Comparison Technique

Because the intermediate base selected to connect these two standards is fixed purely on semantics, the point of convergence spanning across the two standards will be focused on the units of “intent” or “meaning” within each standard message. For the purposes of this report, these units of “intent” or “meaning” are referred to as a semantic unit, which is defined as the smallest unique and repeatable functional, physical, or conceptual unit of meaning that may be applied within a particular domain. The methodology used to identify, classify, and attach meaning to these semantic units is defined in the next section.

“Semantics is the key point of convergence for equating the two standards.”

3.0 STANDARDIZED MAPPING METHODOLOGY

In order to conduct an analysis of semantic equivalence between the current set of ANSI X12 standards, and the current set of UN/EDIFACT standards, a

“The UDEF provides a well defined, standardized, and highly structured framework for assessing the semantic equivalency of the two EDI standards.”

clear, consistent, highly structured and standardized mapping methodology must be employed. The standard mapping methodology proposed for this project is the Universal Data Element Framework (UDEF). The background of the UDEF methodology and its applicability to our project is discussed and applied to an example mapping scenario in the following sub-sections.

3.1 Universal Data Element Framework

In 1988, the CALS Industry Steering Group began pursuing a framework known as the Universal Data Element Framework that would enable the Integrated Weapon System Database (IWSDB). Today, the Integrated Data Environment (IDE) has fundamentally the same objectives as the IWSDB. The IDE state is envisioned to allow "plug-in/plug-out" participation by all government and commercial entities. This plug-in/plug-out capability is considered essential for one to be considered a "virtual enterprise." To achieve this IDE state, a key presumption is the need for seamless interoperability between all information systems involved in the business and technical aspects of the complete life-cycle of a given product. [5]

In 1995 after many years and committee meetings, the CALS Industry Steering Group's UDEF began the process of migration toward an American National Standards Institute (ANSI) standard under the auspices of the Electronics Industries Association (EIA). Once standardized, this framework is considered to be an essential tool for an enterprise-wide data management function. In addition, a national or international data element registration authority function is required to enable the effective coordination and flow-down of the UDEF and the Universal Data Element (UDE) codes to those organizations wishing to keep their information systems interoperable with the systems within their customers and supplier based organizations. [5]

At the CALS EXPO '93, the concept demonstrated the feasibility of mapping data elements from the Standard for the Exchange of Product model data (STEP), EDI, UN/EDIFACT, the DoD Enterprise Model, and MIL-STD-1388 into the same framework. Through June 1995, approximately 50% of the over 600 Configuration Management (CM) Data Model data elements were analyzed and 95% were successfully mapped into the Universal Data Element Framework. The unmapped 5% require further coordination with the team that developed the data model. [5]

What is the Universal Data Element Framework?
The UDEF provides a common structure for mapping all shareable data elements across the

“Differences between the semantic levels of abstraction are equalized using UDEF.”

various standards and legacy systems. The methodology is built on the premise that one must take a concept to a level of abstraction that is high enough such that any concept can be categorized into one and only one choice of mutually exclusive options. In addition, the UDEF follows the data naming principles specified in the International Organization for Standardization / International Electrotechnical Commission (ISO/IEC) 11179-5 [7]. The same data naming principles are required for all "migration systems" to be used by the Department of Defense. The data element naming requirements are specified in DoD 8320.1-M-1 (Department of Defense Data Element Standardization Procedures). [5] The UDEF methodology is based on a predefined set of primary *objects* and *properties* (Figure 3.1-1) that may be used to classify data element(s) and their definitions (Tables 3.1-1 and 3.1-2).

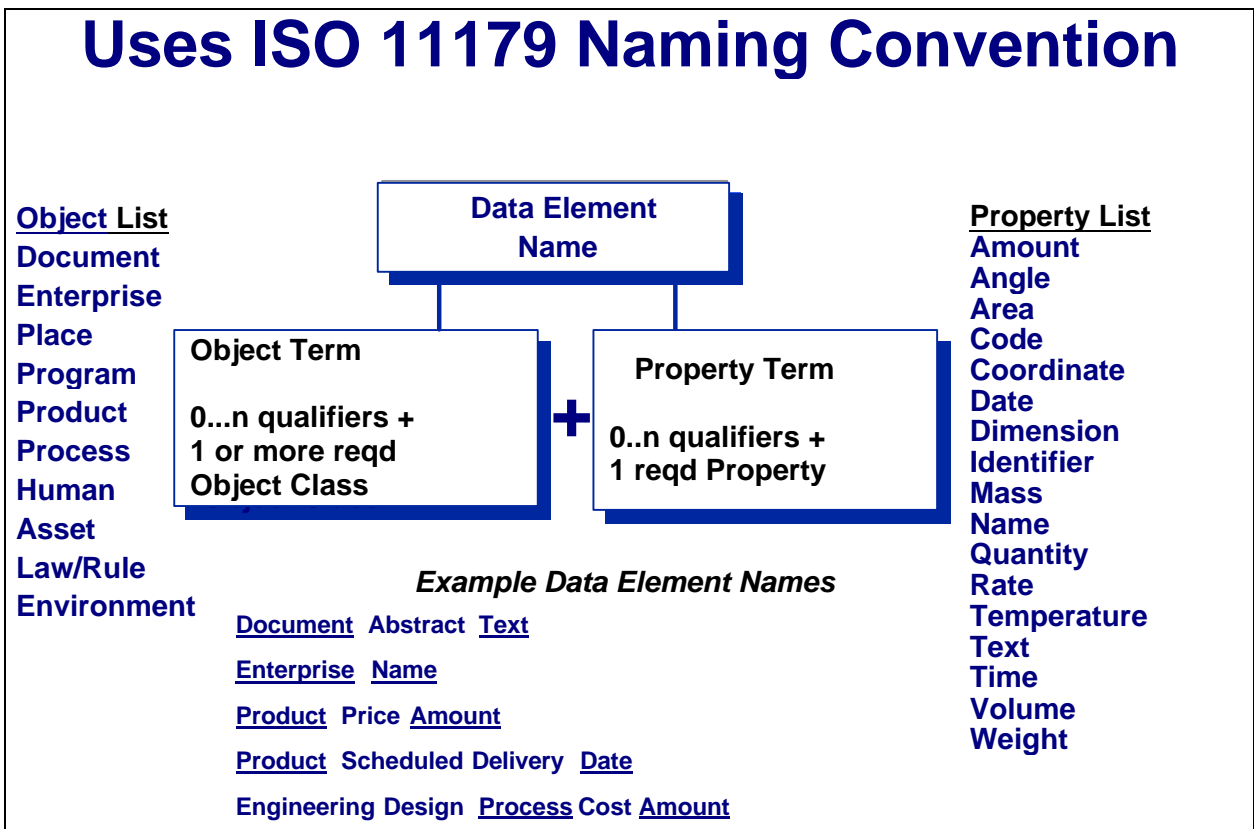


Figure 3.1-1: ISO 11179 Naming Convention

Notice from this figure, that the data element name, which for the purposes of this paper is the name of a semantic unit, is comprised of one or more *object* terms with zero or more *object type qualifiers*, and one *property* term with zero or more *property qualifiers*. These *objects* are defined as a set of ideas, abstractions, or things in the real world that can be identified with explicit boundaries and meaning and whose properties and behavior follow the same rules. Likewise, these *properties* are defined as the human perception of a single

“Semantic units are simply comprised of only an object and/or a property and any associated qualifiers.”

characteristic of an object in the real world. The following definitions for each of the 10 "object" words were developed by the CALS Industry Steering Group:

Table 3.1-1: UDEF Object Definitions

Object Definitions	
1. Asset -	Any data or information about any resource, other than human, which is used, consumed, or available for use/consumption by any process of an enterprise.
2. Document -	Any data or information about any collection of data or information, regardless of format, which has definable boundaries and is so designated for one or more purposes.
3. Enterprise -	Any data or information about any definable boundary collection of human and asset resources used to perform a collection of processes to create one or more products which are intended for use or consumption by outside activities.
4. Environment -	Any data or information about any natural or man-made condition that is relevant to the enterprise.
5. Human -	Any data or information about any human that is relevant to the enterprise.
6. Law/Rule -	Any data or information about laws (natural or man-made) or policies that govern any process of the enterprise.
7. Place -	Any data or information about any location that is relevant to the enterprise.
8. Process -	Any data or information about a definable course of events distinguishable by its purpose or by its effect, whether natural, manual, automated or machine supported and which is relevant to the enterprise.
9. Product -	Any data or information regarding something that is the result of a set of processes and which is intended to be used or consumed by activities outside of the enterprise.
10. Program -	Any data or information about any definable collection of enterprises bound by a common set of objectives.

Also notice that each of the objects (and properties below) are numbered in alphabetical order. These numbers are used in the definition and assignment of universal codes corresponding to each semantic unit, which will be explained in the next section. The following definitions for each of the 17 "property" words are taken directly from the DoD 8320.1-M-1, *Data Element Standardization Procedures*, which is a mandatory compliance requirement for all information systems used internal to DoD:

Table 3.1-2: UDEF Property Definitions

Property Definitions	
1. Amount	A monetary value. (Includes: Average, Balance, Deviation, Factor, Index, Level, Mean, Mode, Scale, and Yield)
2. Angle	The rotational measurement between two lines/planes diverging from a common point/line. (Includes: Azimuth, Heading)
3. Area	The measurement of a surface expressed in unit squares (two dimensional).
4. Code	A combination of one or more numbers, letters, or special characters substituted for a specific meaning. Represents finite, predetermined values. (Must have a specific domain.) (Includes category and status.)

5. Coordinate	Designation of the location of a line or plane. (Includes latitude and longitude.)
6. Date	The designation of a specific 24-hour period of time.
7. Dimension	A linear distance (one-dimensional). (Includes altitude, depth, diameter, elevation, height, length, radius, vertex, and width)
8. Identifier	A combination of one or more numbers, letters, or special characters that designate a specific object/entity but that have no readily definable meaning. (Must have a general domain.) (Includes designator, key, and number.)
9. Mass	The measurement of inertia of a body.
10. Name	The designation of an object and/or entity expressed in a word or phrase.
11. Quantity	A nonmonetary numeric value. (Includes average, balance, count, deviation, factor, index, level, mean, median, mode, and scale.)
12. Rate	A quantity or degree of something in relation to units of something else (e.g., miles per gallon). (Includes acceleration, density, factor, flow, force, frequency, humidity, impedance, inductance, intensity, magnitude, moment, percent, power, pressure, resistance, scale, speed, tension, torque, velocity, viscosity, and voltage.)
13. Temperature	The measure of heat in an object or space.
14. Text	An unformatted character string, generally in the form of words. (Includes category and comments.)
15. Time	A designation of a specified chronological point within a period.
16. Volume	Measurement of space occupied by a three-dimensional figure as measured in cubic units.
17. Weight	The force with which an object is attracted toward the earth and/or another celestial body by gravitation.

Now that definitions exist for both objects and properties, and that rules exist for naming semantic units using these objects and properties, we can explain the how to apply the UDEF in more detail. The next section discusses how to identify semantic units from the EDI standards, how to define unique and semantically “smart” codes for easily referring to these semantic units, and how to use these semantic unit codes as a basis for comparing and aligning the two EDI standards.

3.2 Semantic Unit Identification and UDE Code Development

In this section we will apply the UDEF methodology to the same example mapping scenario pictured in Figure 2.1-1, which involved mapping the ANSI X12 BEG segment to the corresponding UN/EDIFACT segment(s). Recall that it is important to establish a neutral reference point between both EDI standards in order to fairly asses the equivalency of the two. The UDEF approach breaks down the transactions sets and messages into semantic units that will function as these neutral reference points. Specifically in this section we will show how to identify and define semantic units, and how the UDEF methodology enables us to capture the semantic information using a universal code.

For the purposes of this paper a semantic unit may be thought of as being independent of both syntax and structural requirements, and is defined as the smallest unique and repeatable functional, physical, or conceptual unit of meaning that may be applied within a particular domain. Perhaps the best way to explain this is by example. Consider a product

code. The product code identifies a particular class of object. There are many synonyms: item number, part number, assembly code, raw material code, etc., and there can be several semantic contexts: customer’s product code, supplier’s product code, manufacturer’s product code, etc. All the synonyms, even though they convey different types of product code, in fact represent exactly the same thing in Information Technology (IT) terms - an identification for a specific class of objects.

The UDEF approach will allow for the assignment of codes for each semantic unit. Analysis of both transaction sets and messages can be performed relative to their respective semantic unit. The Universal Data Element (UDE) codes can be generated for the semantic units of each transaction set and message. These codes can then be used as the medium by which the two standards are compared and analyzed. The *One-to-One*, *One-to-Many*, and *One-to-None* assessments may therefore be classified as semantic matches and non-matches, which simplifies the overall equivalency assessment process.

“Using the UDEF, the number of relationship types is reduced to either a match or no match situation.”

The need for identifying semantic units for the EDI standards can be best illustrated by considering the following situation. Recall that in Section 2.2 we discussed the “vagueness” of the definition for the **date** data element. In the context of the BEG segment (see previous Figure 2.1-1) of the ANSI X12 850 Transaction Set (Purchase Order), the **date** data element has a more precise meaning than is conveyed by its data element definition. The actual intent of the **date** data element for this scenario is to convey the Purchase Order Date. Therefore, in this case, the term **Purchase Order Date** may be thought of as a semantic unit. This semantic unit maps to the UN/EDIFACT DTM segment of the UN/EDIFACT ORDERS message. However, the DTM segment has the Date/Time/Period data element along with two additional elements required to specify the type of date and the format for the date. Recall that the first qualifier data element is used to indicate the type of date (i.e. qualifier code value = ‘4’ for Purchase Order), while the second qualifier data element is used to specify the format for the date (i.e. qualifier code value = ‘101’ for the YYMMDD arrangement) as shown below.

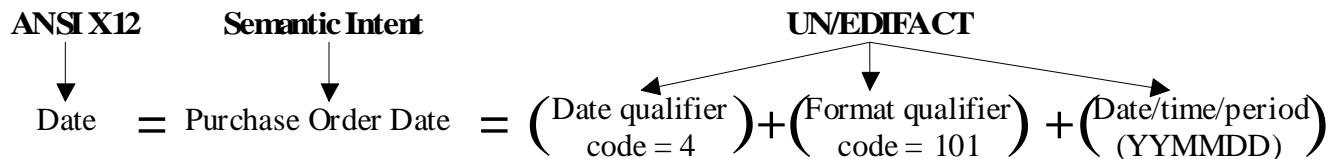


Figure 3.2-1: Intermediate Data Element Comparison Technique

The semantic unit must then be classified using the UDEF and assigned a UDE code (explanation of UDE code assignment will follow). A UDE code can be determined as indicated by using the methodology shown in Figure 3.2-2 below.

Universal Code

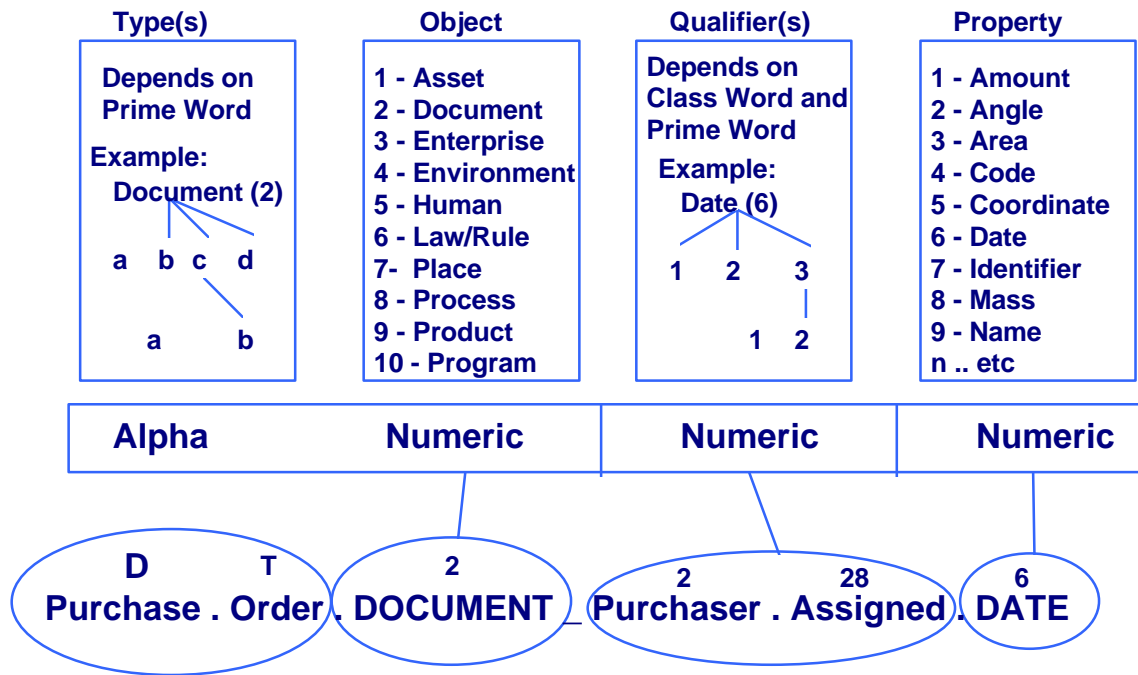


Figure 3.2-2: UDE Code Assignment Process

For example, the **Purchase Order Date** semantic unit can be classified using the object and property words contained in the UDEF. It is extremely important to accurately define the semantic intent of each semantic unit early on to avoid potential “downstream” implications. Recalling the definitions in Table 3.1-1, the *object* is determined to be a ‘**DOCUMENT**’. Following the UDEF approach, the *object type qualifier* for ‘**DOCUMENT**’ was determined to be an ‘**Order**’, which in turn is a ‘**Purchase**’ *object type qualifier* for this case. The resulting *object* for this semantic unit is therefore classified as **Purchase.Order.DOCUMENT**. Note that in order to preserve the neutrality of the semantic unit, we do not distinguish between the ANSI X12 Purchase Order and the UN/EDIFACT Orders message, i.e. we elected to not develop a Purchase.Order.ANSIX12.Document semantic unit separate from a Purchase.Order.EDIFACT.Document semantic unit. Using the same process and recalling that the definition of the **date** data element is “the purchase order date assigned by the purchaser,” we have defined the *property* as a **Purchaser.Assigned.DATE**. The full semantic unit may therefore be written as:

[Purchase.Order.DOCUMENT_Purchaser.Assigned.DATE],

which provides a clear and concise naming convention that incorporates the full meaning carried by this semantic unit.

Figure 3.2-3, shown on the next page, provides a full-scale tree view (Reference Appendix D: *CONVEYOR*) of an object tree and the corresponding property tree with the appropriate nodes expanded to show the **Purchase.Order.DOCUMENT_Purchaser.Assigned.DATE** semantic unit. Both “PROCESS” and “DATE” “prime words” pertaining to the UDEF *objects* and *properties* are indicated by

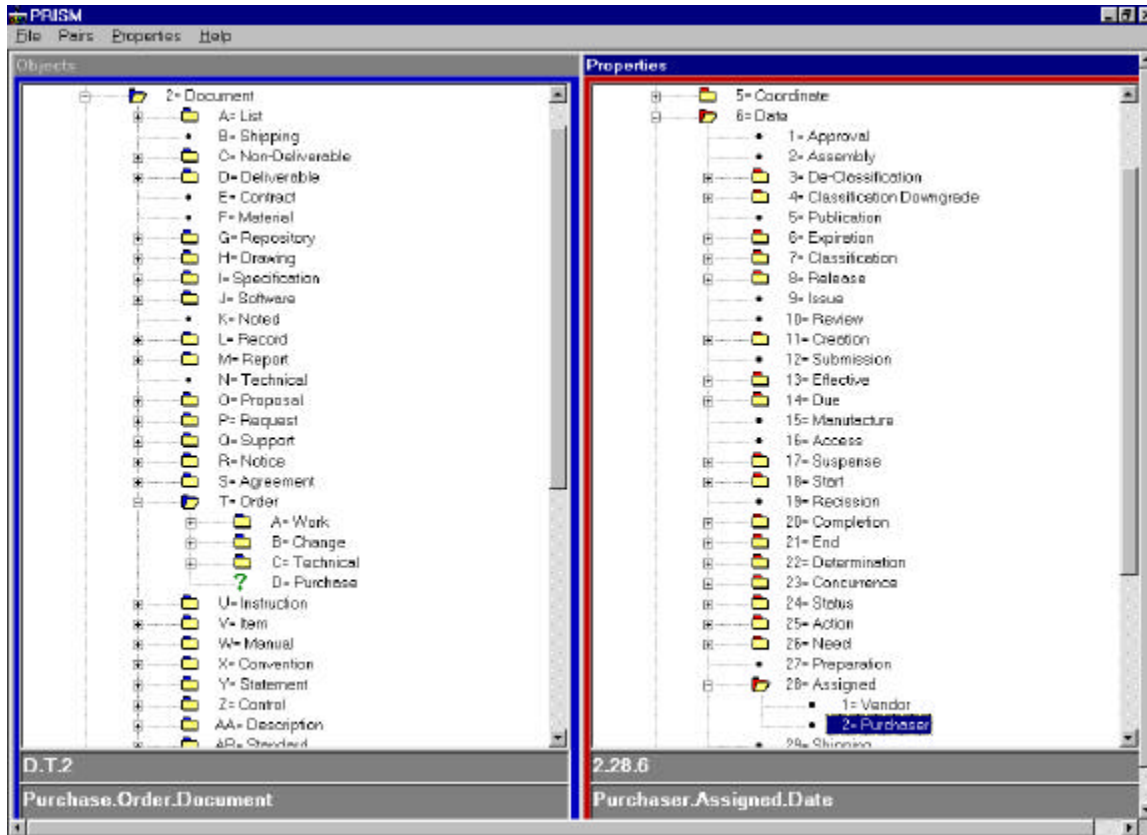


Figure 3.2-3: UDE Code Tree

Notice how the element numbers from both the *object* and the *property* sides are used to create the universal code of **D.T.2_2.28.6**, which corresponds to this semantic unit. These universal codes allow us to represent semantic units in a compact, complete, and efficient manner. With the semantic units defined and the UDE codes assigned, we may now begin to assess the equivalency of the ANSI X12 transaction sets with the targeted the UN/EDIFACT messages.

“Universal codes allow us to represent semantic units in a compact, complete, and efficient manner.”

3.3 Semantic Equivalency Assessment Using UDEF

Using the example shown in Figure 2.1-1, it is important to develop a concise semantic understanding of the intended purpose for these data elements in order to accurately

define corresponding semantic units. In this section we will define all the semantic units necessary to convey the intended purpose of the ANSI X12 BEG segment. From these semantic units we will assign UDE codes and then attempt to identify equivalent “units of meaning” in the corresponding pieces of the UN/EDIFACT ORDERS message. A detailed listing of all the data elements used in this example and definitions for these elements are also provided in this section.

The first step in analyzing the semantic equivalency between ANSI X12 transaction sets and UN/EDIFACT messages is to identify and define our source, which for the purposes of this example is the ANSI X12 BEG segment. This particular segment is used to indicate the beginning of the Purchase Order transaction set and transmit identifying numbers and dates. The data elements contained in this segment are listed in Table 3.3-1 below.

Table 3.3-1: ANSI X12 BEG Segment Data Elements

Element ID	Element Name
353	Transaction Set Purpose Code
92	Purchase Order Type Code
324	Purchase Order Number
328	Release Number
373	Date
367	Contract Number
587	Acknowledgment Type
1019	Invoice Type Code
1166	Contract Type Code
1232	Purchase Category
786	Security Level Code

Before attempting to link ANSI X12 data elements to the UN/EDIFACT data elements, the concise semantic meanings of the ANSI X12 data elements must be understood and defined. Definitions for the ANSI X12 data elements have been extracted from the ANSI X12 standard and the Federal/DoD implementation conventions and are listed below in Table 3.3-2.

Table 3.3-2: Semantic Meaning Defined for ANSI X12 BEG Segment Elements

ANSI X12 Semantic Unit (BEG Segment)	Semantic Definition
Transaction Set Purpose Code	Code identifying purpose of Purchase Order transaction set
Purchase Order Type Code	Code specifying the type of Purchase Order
Purchase Order Number	Identifying number for Purchase Order assigned by the orderer/purchaser
Release Number	Number identifying a release against a Purchase Order previously placed by the parties involved in the transaction
Date	Date assigned by the purchaser to Purchase Order

Contract Number	Contract number
Acknowledgment Type	Code specifying the type of acknowledgment
Invoice Type Code	Code defining the method by which invoices are to be processed
Contract Type Code	Code identifying a contract type
Purchase Category	Code identifying the broad category of products or services being acquired
Security Level Code	Code indicating the level of confidentiality assigned by the sender

With semantic units identified and defined in terms of the ANSI X12 data elements, we may now use the UDEF methodology to establish formal semantic unit names and universal data element (UDE) codes. Using the automated software tool pictured in Figure 3.2-3, which assists in this mapping process, both the semantic unit name and the UDE code may be generated simultaneously. The following semantic unit labels have been identified for the ANSI X12 data elements above:

Table 3.3-3: Semantic Unit Labels

Semantic Unit Labels	
ANSI X12 Data Element	Object Label Property Label
Transaction Set Purpose Code	Purchase.Order.Document _ Purpose.Identification.Code
Purchase Order Type Code	Purchase.Order.Document _ Type.Code
Purchase Order Number	Purchase.Order.Document _ Purchaser.Assigned.Identifier
Release Number	Purchase.Order.Document _ Release.Identifier
Date	Purchase.Order.Document _ Purchaser.Assigned.Date
Contract Number	Contract.Document _ Numeric.Identifier
Acknowledgment Type	Acknowledgement.Notification.Process _ Type.Code s
Invoice Type Code	Payment.Notification.Process _ Type.Code
Contract Type Code	Contract.Document _ Type.Code
Purchase Category	Asset _ Type.Code
Security Level Code	Information.Asset _ Security.Level.Code

Notice from the above table that the semantic unit corresponding to the ANSI X12 Date data element captures the true intent of this data element as defined in the ANSI X12 standard in this particular context. This “built-in” ability to understand the precisely intended meaning the EDI data elements is otherwise not achievable without utilizing the UDEF to define these semantic units.

“Semantic units inherently contain the precise definition of the intent of the EDI data element(s) as they are applied in various contexts.”

The corresponding UDE codes for the above semantic units are listed below in Table 3.3-4.

Table 3.3-4: UDE Codes

Semantic Unit	UDE Code
Purchase.Order.Document_Purpose.Identification.Code	D.T.2_11.8.4

Purchase.Order.Document_Type.Code	D.T.2_33.4
Purchase.Order.Document_Purchaser.Assigned.Identifier	D.T.2_2.35.8
Purchase.Order.Document_Release.Identifier	D.T.2_37.8
Purchase.Order.Document_Purchaser.Assigned.Date	D.T.2_2.28.6
Contract.Document_Numeric.Identifier	E.2_3.8
Acknowledgement.Notification.Process_Type.Code	A.P.8_33.4
Payment.Notification.Process_Type.Code	B.P.8_33.4
Contract.Document_Type.Code	E.2_33.4
Asset_Type.Code	1_33.4
Information.Asset_Security.Level.Code	M.1_4.22.4

Now that we have defined what we are mapping from, our source, we must identify and define equivalent semantic units within the corresponding UN/EDIFACT ORDERS message, which is what we are mapping to. Recall from Section 2.1 that when considering respective segments to map to in the UN/EDIFACT standard, the BGM segment, at first glance, appears to be equivalent to ANSI X12 BEG segment. The BGM segment is the beginning of message segment, which is used to indicate the type and function of a message and to transmit the identifying number for the message. However, upon further investigation, it is revealed that while the BEG segment is only used for the ANSI X12 Purchase Order transaction set, the BGM segment is used more generically and appears in nearly all of the UN/EDIFACT messages. This fundamental message design difference necessitates including three additional segments from within the header section of the ORDERS message in order to account for all the respective data elements from the Purchase Order transaction set. The BGM segment and the additional segments, which include the DTM, RFF, and PAI segments, along with the appropriate data elements and code qualifiers will be defined below.

The BGM (Beginning of Message) segment is defined by the UN/EDIFACT standard as a segment by which the sender must uniquely identify the order by means of its name and number and when necessary its function. The data elements contained in this segment are listed in Table 3.3-5 below.

Table 3.3-5: UN/EDIFACT BGM Segment Data Elements

Element ID	Element Name
1001	Document/message name, Coded
1131	Code list qualifier
3055	Code list responsible agency, Coded
1000	Document/message name
1004	Document/message number
1225	Message function, Coded
4343	Response type, Coded

Reiterating what we stated for defining the source data elements, the target data elements must be understood and defined. Definitions for these data elements have been extracted from the UN/EDIFACT standard and are listed below in Table 3.3-6 below.

Table 3.3-6: Semantic Meaning Defined for UN/EDIFACT BGM Segment Elements

UN/EDIFACT Semantic Unit (BGM Segment)	Semantic Definition
Document/message name, Coded	Document/message identifier expressed in code.
Code list qualifier	Identification of a particular code list.
Code list responsible agency, Coded	Code identifying the agency responsible for a code list.
Document/message name	Plain language identifier specifying the function of a document/message.
Document/Message Number	Reference number assigned to the document/message by the issuer.
Message Function, Coded	Code indicating the function of the message.
Response Type, Coded	Code specifying the type of acknowledgment required or transmitted.

The DTM (Date/Time/Period) segment is defined by the UN/EDIFACT standard as a segment specifying general dates and, when relevant, times related to the whole message. The segment must be specified at least once to identify the order date. Examples of the use of this DTM segment are: Lead times that apply to the whole of the Order and, if no schedule is to be specified, the delivery date. The Date/time/period segment within other Segment group should be used whenever the date/time/period requires to be logically related to another specified data item. e.g. Payment due date is specified within the PAT Segment group. The data elements contained in this segment are listed in Table 3.3-7 below.

Table 3.3-7: UN/EDIFACT DTM Segment Data Elements

Element ID	Element Name
2005	Date/Time/Period qualifier
2380	Date/Time/Period
2379	Date/Time/Period format qualifier

Definitions for these data elements have been extracted from the UN/EDIFACT standard and are listed below in Table 3.3-8 below.

Table 3.3-8: Semantic Meaning Defined for UN/EDIFACT DTM Segment Elements

UN/EDIFACT Semantic Unit (DTM Segment)	Semantic Definition
Date/Time/Period qualifier code	Code giving specific meaning to a date, time or period.
Date/Time/Period	The value of a date, a date and time, a time or of a period in a specified representation.
Date/Time/Period format qualifier	Specification of the representation of a date, a date and time or of a

code	period.
------	---------

The RFF (Reference) segment is defined by the UN/EDIFACT standard as a segment identifying the reference by its number and where appropriate a line number within a document. The data elements contained in this segment are listed in Table 3.3-9 below.

Table 3.3-9: UN/EDIFACT RFF Segment Data Elements

Element ID	Element Name
1153	Reference qualifier
1154	Reference number
1156	Line number
4000	Reference version number

Definitions for these data elements have been extracted from the UN/EDIFACT standard and are listed below in Table 3.3-10 below.

Table 3.3-10: Semantic Meaning Defined for UN/EDIFACT RFF Segment Elements

UN/EDIFACT Semantic Unit (RFF Segment)	Semantic Definition
Reference qualifier code	Code giving specific meaning to a reference segment or a reference number.
Reference number	Identification number the nature and function of which can be qualified by an entry in data element 1153 Reference qualifier.
Line number	Number of the line in the document/message referenced in 1154 Reference number.
Reference version number	To uniquely identify a reference by its revision number.

The PAI (Payment Instructions) segment is defined by the UN/EDIFACT standard as a segment requesting or confirming conditions of payment, guarantee and method of payment for the whole order. The data elements contained in this segment are listed in Table 3.3-11 below.

Table 3.3-11: UN/EDIFACT PAI Segment Data Elements

Element ID	Element Name
4439	Payment conditions, Coded
4431	Payment conditions, Coded
4461	Payment means, Coded
1131	Code list qualifier code
3055	Code list responsible agency, Coded
4435	Payment channel, Coded

Definitions for these data elements have been extracted from the UN/EDIFACT standard and are listed below in Table 3.3-12 below.

Table 3.3-12: Semantic Meaning Defined for UN/EDIFACT RFF Segment Elements

UN/EDIFACT Semantic Unit (RFF Segment)	Semantic Definition
Payment conditions, Coded	Identification of the method employed or to be employed in order that a payment may be made or regarded as made. The method may or may not be tied to a guarantee.
Payment conditions, Coded	Identification of the means of guarantee.
Payment means, Coded	Indication of the instrument of payment, which may include a guarantee.
Code list qualifier code	Identification of a code list.
Code list responsible agency, Coded	Code identifying the agency responsible for a code list.
Payment channel, Coded	Identification of the channel of payment.

These four segments (BGM, DTM, RFF, and PAI) were selected because they contain some semantic units that have an equivalent meaning to the UDEF semantic units defined in Table 3.3-3. These relationships among these semantic units and the various data elements are listed below in Table 3.3-13.

Table 3.3-13: UDE Codes for Semantic Unit Labels

ANSI X12 Semantic Unit (BEG Segment)	UDE Code	UN/EDIFACT Semantic Unit (BGM, DTM, RFF, PAI Segments)
1. Transaction Set Purpose Code	D.T.2_11.8.4	BGM Segment: Message function, Coded
2. Purchase Order Type Code	D.T.2_33.4	No match
3. Purchase Order Number	D.T.2_2.35.8	BGM Segment: Document/message number
4. Release Number	D.T.2_37.8	No match
5. Date (for purchase order)	D.T.2_2.28.6	DTM Segment: Date/Time/Period with qualifier code of 4 for PO date Date/Time/Period - (actual date value) Date/Time/Period with qualifier code of 101 for YYMMDD format
6. Contract Number	E.2_3.8	RFF Segment: Reference qualifier with qualifier code of CT for Contract number Reference number (contract number in this case)
7. Acknowledgment Type	A.P.8_33.4	BGM Segment: Response type, Coded
8. Invoice Type Code	B.P.8_33.4	PAI Segment: Payment channel, Coded
9. Contract Type Code	E.2_33.4	No match
10. Purchase Category	1_33.4	No match
11. Security Level Code	M.1_4.22.4	No match

From the results above of the example, it can be concluded that for the BEG segment of the ANSI X12 850 transaction set there exists six matches and five no-matches relative to the corresponding UN/EDIFACT ORDERS message. Semantic equivalence and nonequivalence has been determined. As these semantic units appear again within the transaction sets or messages, the existing UDE code may be reused, thus simplifying the mapping process.

“Semantic units may be identified once and may be used many times.”

3.4 Semantic Equivalency Assessment Automation

Performing a semantic equivalency assessment for these two standards is an enormous undertaking, because of the large number of data elements and codes contained in both the X12 transaction sets and their respective UN/EDIFACT counterparts. Any automation in this area would be beneficial in reducing the overwhelming work load of the assessment. As we described in Section 3.2, the method used to compare and assess the semantic units (UDEF) is very structured and uses a consistent numbering scheme that lends itself well to automation.

A quick breakdown of the comparison and assessment workflow is shown with areas of possible automation indicated below by a bold face font.

“Semantic unit identification and mapping using UDEF may be automated using a software tool.”

1. Determine the semantic unit within X12 transaction set
2. **Map the X12 semantic unit to the UDE Framework**
3. **Update the UDE Framework where necessary**
4. **Determine the UDE code for the X12 semantic unit**
5. Identify semantically equivalent UN/EDIFACT data elements when possible
6. **Associate the UN/EDIFACT semantic unit to the UDE Framework**
7. **Update the UDE Framework where necessary**
8. **If no equivalencies exist, record “No Match”**

Furthermore, transaction sets and messages will contain certain semantic units that appear again not only in the same transaction set / message, but also in other transaction sets / messages as well. Any semantic unit that has already been recorded within the UDE framework will already have been assigned a corresponding UDE code. The potential for reusing existing UDE codes could significantly reduce the level of effort required for assessing the equivalency of other transaction sets. As the UDEF matures, the possibility of having existing UDEF semantic units defined prior to mapping is greater. The administration and maintenance of the UDE framework can also be aided by automation in updating the hierarchical UDEF trees. Refer to Appendix C for a brief synopsis of an application prototype, *CONVEYOR* that assists in the determination of UDE codes based on the Universal Data Element Framework.

4.0 SUMMARY AND CONCLUSIONS

Although the DoD is currently migrating from ANSI X12 to UN/EDIFACT, there remains a need for the two standards to be interoperable. Subtle but fundamental differences between the two standards has led to compatibility barriers, which necessitates developing a standardized methodology for both assessing the semantic equivalency and aligning the two EDI standards. Through an analysis of the nuances of each standard, it was determined that the best starting point for assessing the potential “compatibility” of the two standards is to first focus purely on the semantics. By first focusing on only the semantics of the data, all of the peculiarities pertaining to the data format, structure, and syntax may be ignored, which allows for the standards to be compared on even terms without giving up the ability to account for differences between corresponding data elements.

Using the UDEF approach for assessing the equivalency of ANSI X12 transaction sets with UN/EDIFACT messages provides a formalized process for an otherwise disorderly and cumbersome task. Although the UDEF is based on an agreed upon set of defined *objects* and *properties*, additional sub-objects and sub-properties must be created to accurately define most, if not all, semantic units. Creating these sub-objects and sub-properties can be a very subjective process that can potentially result in definitions that are not universally applicable. It is therefore important for anyone using this approach to align these two standards to collaborate with industry / government experts when updating the UDEF database. Some additional benefits provided by using the UDEF approach include its compliance with DoD 8320.1-M-1 naming requirements, compliance with ISO/IEC 11179 requirements, reusability, and that it is object oriented in nature.

It is also worth mentioning that a similar project is ongoing in Europe dubbed the Basic Semantic Repository (BSR) project. This project is focused on using semantics as an intermediate connection between the two standards, however there are fundamental differences in their approach and objective. The underlying objective of the BSR project is to establish ISO as a registrar for BSR semantic units. Also the BSR’s approach to defining semantic units involves the use of *property classes*, but does not include *object classes*, which therefore reduces the measure of consistency provided by the UDEF approach. However, because the BSR project is similar in scope to the methodology proposed in this paper, the two approaches should be assessed in more detail to potentially harmonize the two efforts.

REFERENCES

- [1] *ASC X12 Version 003060*, [CD-ROM]. Washington Publishing Company. (1996).
- [2] *AVNET Implementation Guide for Electronic Data Interchange, A joint effort of the American Petroleum Institute, Air Transport Association of America and the International Air Transport Association*. (May, 1996) Washington Publishing Company.
- [3] *Defense Fuels Supply Center Homepage*, [on-line]. (August 31, 1996). Available: <http://www.DFSC.DLA.MIL>.
- [4] *UN/EDIFACT Version D.96A*, Pan American EDIFACT Board. (March, 1996). Data Interchange Standards Association.
- [5] *What is the Role of a Universal Data Element Framework in the Envisioned Integrated Data Environment?* (CAL S 1995 EXPO) Ronald L. Schuldt, Chairman, CAL S Enterprise Data Task Group
<http://arioch.gsfc.nasa.gov/wwwvl/de.html>
- [6] *The Integrated Data Environment (IDE) Continuous Acquisition and Lifecycle Support, Draft White Paper*. Office of the Under Secretary of Defense for Acquisition and Technology, 8 February 1995.
- [7] *ISO/IEC 11179*, Information Technology - Specification and Standardization of Data Elements, an evolving international standard that contains six parts, ISO/IEC JTC1/SC4.
- [8] *Multiple EDI Standards*, (CAL S 1995 EXPO) M.D. Conroy, DESCO.

APPENDIX A: GLOSSARY OF TERMS

Accredited Standards Committee (ASC) X12

The committee was chartered by ANSI in 1979 to develop uniform standards for the electronic interchange of business documents. Accredited Standards Committee X12 consists of U.S. industry and government members who develop Electronic Data Interchange (EDI) standards for submission to ANSI for subsequent approval as a United States Standard or to the UN/ECE for approval and submission of UN/EDIFACT international standards.

American National Standards Institute (ANSI)

A private, non-profit membership organization that coordinates the development and approval of voluntary consensus standards in the United States.

American Standard Code for Information Interchange (ASCII)

A standard character set developed to enable efficient data exchange and compatibility among different computer platforms and peripherals. ASCII is a seven bit code with an eighth bit used for parity. The term is used to describe the format for transmission and for storage.

Basic Semantic Unit (BSU)

A BSU is a "semantically pure" data element. Perhaps the best way to explain this is by example. Consider a product code. The Product Code identifies a particular class of object. There are many synonyms: item number, part number, assembly code, raw material code, etc., and there can be several semantic contexts: customer's product code, supplier's product code, manufacturer's product code, etc. All the synonyms, even though they convey different flavors of product code, in fact represent exactly the same thing in IT terms - an identification for a specific class of objects.

CALS Framework

A generic framework for creating a lifecycle enterprise based on an Integrated Technical Data Environment.

Data Element	A unit of data which, in a certain context, is considered indivisible. UN/EDIFACT: a unit of data for which the identification, description, and value representation have been specified. (ISO 9735)
Data Segment	A data segment is a group of related data elements in a transaction set. Each segment has a unique segment identifier, a combination of two or three uppercase letters and/or digits that serves as a name for the segment and occupies the first character positions of the segment. A segment is equivalent to a data record in a database.
DoD CALS Framework	An instantiation of the CALS Framework employed by the Department of Defense for constructing lifecycle enterprises for its weapon systems and other materiel.
DoD CALS Program	Any program or program element (i.e., funded activity with discrete milestones, resources, and deliverables) in the DoD, assigned to and managed by a DoD Programmatic Thrust, which is purposefully focused on developing or implementing aspects of the DoD CALS Framework; includes DoD computer system acquisition programs, weapon system implementation programs, information system migration programs, and research and development programs
EDI	Transactions that are structured in a highly standardized way that facilitates use over machine-to-machine interfaces with no human intervention. (Steel94a.)
Electronic Commerce (EC)	A concept and implementation that includes a range of enabling technologies supporting the timely digital exchange of business information and includes but is not limited to Electronic Data Interchange (EDI), Electronic Mail, bar-coding, electronic funds transfer, electronic bulletin boards, electronic security technologies (e.g., public key encryption), and electronic catalogs.
Electronic Data Interchange (EDI)	Exchange of formation from one computer system to another without human intervention.

Format	The organization, plan, type or style of an EC/EDI object.
Framework	<p>A structured set of standards used to integrate quasi-autonomous organizational modules into an integrated enterprise. A Framework comprises three architectures:</p> <ul style="list-style-type: none"> • Process Architecture - which defines the business processes of the integrated enterprise. • Information Technology Architecture - which defines the information technology used within the enterprise. • Control Architecture - which defines the functional, technical, and data standards controlling the process and information technology architectures of the enterprise.
Identifier	A character or group of characters used to identify or name an item of data and possibly to indicate certain properties of that data. (ISO 9735)
Implementation Convention (IC)	An agreement between trading partners on how they will interpret the meaning of data elements in an EDI transaction set or message.
Interchange	Communication between partners in the form of a structured set of messages and service segments starting with an interchange control header and ending with an interchange control trailer. (ISO 9735)
International Organization for Standardization (ISO)	An international organization with representatives from member countries that develops and approves international standards. International Organization for Standardization, based in Geneva, it is a federation of national standards organizations.
ISO 9735	International Standard issued by ISO which reproduces the UN/EDIFACT Syntax Rules as agreed by WP.4.

Mandatory	A statement in a segment or message directory that specifies that a segment, a data element, a composite data element, or a component data element must be used. (ISO 9735)
Message	An ordered series of characters intended to convey information. UN/EDIFACT: a set of segments in the order specified in a message directory starting with the message header and ending with the message trailer. (ISO 9735) Equivalent to a transaction set.
Object	An object is something of which one takes cognizance. The object taxonomy is an extensible, systematized, hierarchical structure that functions as an aid in associating, classifying, defining, distinguishing, identifying and ordering objects. (ISO 1179)
Object (our definition)	A label for the main physical, functional, or conceptual subject area of the semantic unit under consideration.
Object (from Ron Schuldt)	A set of ideas, abstractions, or things in the real world that can be identified with explicit boundaries and meaning and whose properties and behavior follow the same rules. One of the concepts which is linked to a data element. (Note: Within the scope of the Object Oriented paradigm, object is referred to as simply "class." In a data model, an object is an entity or a relationship.)
Property	A property is defined as that which identifies a quality specifically in the constitution of, or found in, the behavior of a species or type of thing by virtue of that thing's true or essential nature. (ISO 11179)
Property (our definition)	A defining quality that is associated with the object under consideration
Property (from Ron Schuldt)	The human perception of a single characteristic of an object in the real world. One of the concepts which is linked to a data element.
Qualifier	A data element whose value shall be expressed as a code that gives specific meaning to the function of another data element or a segment. (ISO 9735)
Repeating Segment	A segment that may repeat in a message as specified in the relevant message type specification. (ISO 9735)

Segment	A predefined and identified set of functionally related data element values that compose an X12 transaction set or an UN/EDIFACT message.
Semantic Unit	The smallest unique and repeatable functional, physical, or conceptual unit of meaning that may be applied within a particular domain.
Semantics	Pertains to the meaning of the information contained within a data structure, independent of the structure itself, such that the conversion from one structure to another would not change the meaning of the data. In this paper, the discussion focuses on what meaning can be inferred by the placement of information in the structure of EDI data.
Structure	The manner in which elements are organized or interrelated.
Syntax	The pattern and/or structure relative to an EC/EDI object.
Syntax Rules	Rules governing the structure of an interchange and its functional groups, messages, segments, and data elements. (ISO 9735)
Trading Partner	An organizational entity (government, private sector, association, small, or large business) that has a commercial relationship with another entity.
Trading Partner Agreement (TPA)	A legal document between two trading partners that defines general EDI procedures, terms, and conditions, and the EDI transaction sets that will be used.
Transaction Set	A complete business document such as an invoice, a purchase order, or a remittance advice.
UDEF	The Universal Data Element Framework provides a common structure for mapping all shareable data elements across the various standards and legacy systems.

UN/EDIFACT

United Nations rules for Electronic Data Interchange for Administration, Commerce, and Transport. They comprise a set of standards, directories, and guidelines for the electronic interchange of structured data, and in particular that related to trade in goods or services, between independent computerized information systems. Recommended within the framework of the United Nations, the rules are approved and published by the UN/ECE in the United Nations Trade Data Interchange Directory (UNTDID) and are maintained under agreed procedures. UNTDID includes:

- The UN/EDIFACT Application Level Syntax Rules (ISO 9735),
- The UN/EDIFACT Message Design Guidelines (MDG),
- The UN/EDIFACT Syntax Implementation Guidelines (SIG),
- The UN/EDIFACT Data Element Directory, EDED (a subset of UNTDED),
- The UN/EDIFACT Code List, EDCL,
- The UN/EDIFACT Composite Data Element Directory, EDCD,
- The UN/EDIFACT Segment Directory, EDSD,
- The UN/EDIFACT United Nations Standard Message Directory (UNSM), EDMD,
- Uniform Rules of Conduct for the Interchange of Trade Data by Teletransmission (UNCID), and
- Explanatory material.

Unit of Functionality

A logical grouping of information representing a particular function. (i.e. A unit of business functionality)

Universal Data Element Code

An alphanumeric representation of a semantic unit based on the Universal Data Element Framework concept.

APPENDIX B: ABBREVIATIONS AND ACRONYMS

ANSI	American National Standards Institute
ASC	Accredited Standards Committee
ASCII	American Standard Code for Information Interchange
AVNET	Aviation Network
BEG	Beginning of the Purchase Order Transaction Set Segment
BGM	Beginning of Message Segment
CALS	Continuous Acquisition and Life-Cycle Support
CM	Configuration Management
DFSC	Defense Fuels Supply Center
DLA	Defense Logistics Agency
DoD	Department of Defense
DTM	DATE/TIME/PERIOD Segment
EDI	Electronic Data Interchange
EDIFACT	Electronic Data Interchange for Administration, Commerce, and Transport
EIA	Electronics Industries Association
IDE	Integrated Data Environment
IMM	Integrated Materiel Manager
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
IT	Information Technology
IWSDB	Integrated Weapon System Database
PAI	Payment Instructions Segment
PAT	Payment Terms Basis
RFF	Reference Segment
STEP	Standard for the Exchange of Product model data
TPA	Trading Partner Agreement
UDE	Universal Data Element
UDEC	Universal Date Element Code
UDEF	Universal Data Element Framework
UN	United Nations

APPENDIX C: *CONVEYOR* APPLICATION

The administration and maintenance of the UDE framework can be aided by automation in updating the UDEF trees. An application prototype, *CONVEYOR*, that assists in the determination of UDE codes based on the Universal Data Element Framework is being developed to help in the automation of the UDEF methodology.

Mission Statement

CONVEYOR is an application developed to automate the mapping, code retrieval, and application maintenance for the UDE Framework methodology. It is to serve as a prototype for labeling and mapping system tool incorporating the UDEF concept.

Prototype

The following Figure C-1 is a view of the *CONVEYOR* prototype application. The application currently operates on 32-bit operating systems such as Windows NT 4.0 and Win95. The associated underlying database is MS-ACCESS. *CONVEYOR* is currently being developed in Visual Basic.

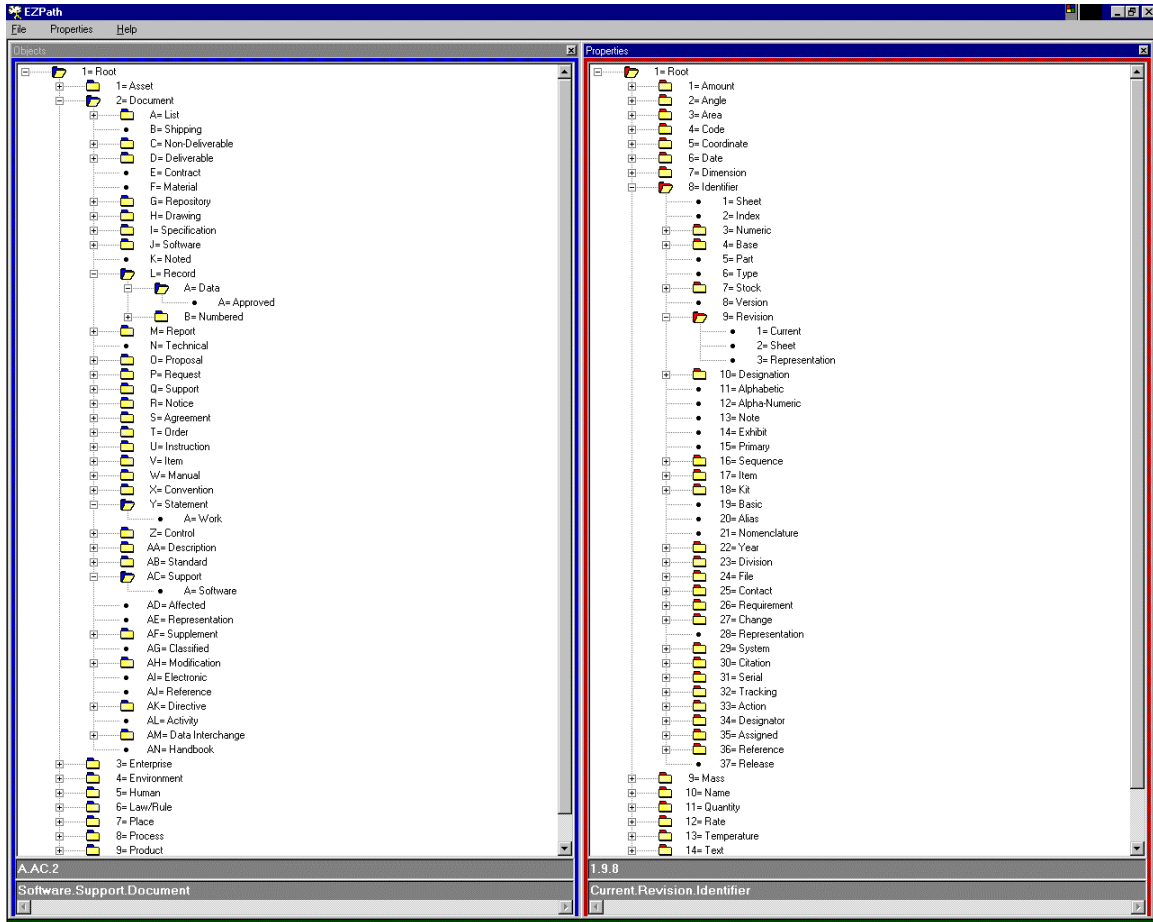


Figure C-1: CONVEYOR Application

The prototype will be developed in several phases as determined by input from designers, developers, and end users. The methodology will be developed based on the UDE Framework and it's respective automation needs.

Note: This application is a prototype going through many requirement and design changes. The application is still under major construction.

The following requirements were generated for the first version of *CONVEYOR*, with the highlighted (bold face font) items operational.

Requirements

- 1. Generate UDEF codes by selecting objects and/or properties**
- 2. Allow for add/delete/modify existing elements in framework**
- 3. Allow for installing updates to master database**
4. Generate “reports” representing tree structures (can be ASCII if needed)

5. Highlight unapproved changes, and allow for an administrator to approve changes
6. Allow for definitions to be added and associated with data elements
7. Printing capabilities
8. Re-sizable windows for object/property
9. Highlight the last marked selection in tree to get UDE code
10. Add full name to bottom of the window from the child up to the root
11. Add functionality to mark element and dump description and UDE codes to file/DB
12. Activate / Select and copy to clipboard

User options, as shown in Figure C-2, include tree operations to alter nodes (add, delete, rename, recode, realign, shrink, etc.) of the UDEF tree. As well, saving and opening new trees are user options. Administrative options have been included to approve and disapprove nodes. The user pull-down menu is the same for both the object and property operations dependent on the active window.

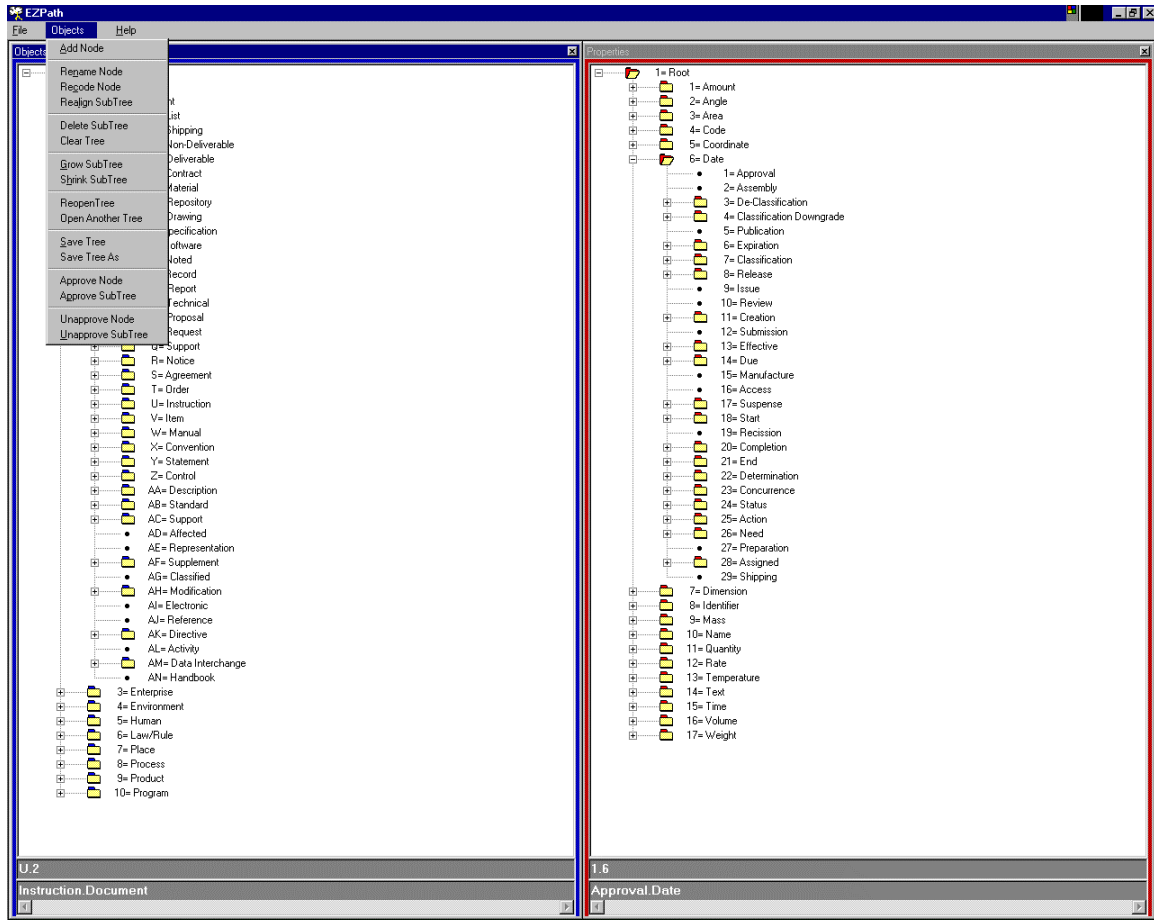


Figure C-2: CONVEYOR Application Options

Suggested future additions and changes for the next version of CONVEYOR are as follows:

1. **Do not show the exploding trees / initialization box/ open the application with just the tree shrunk**
2. Search capabilities – search from the point of cursor down through children
3. Add definitions to the elements – right click on node -> the definition if available, pops up.
4. Develop help files
5. Obtain a new name for the application
6. If closing through corner icon button, prompt for saving data.
7. Activate ‘Hot’ keys – particularly active the delete key for removing nodes when highlighted.
8. Write lock / read-only access issues for the application being executed simultaneously.

APPENDIX D: METHODOLOGY EVALUATIONS

Note: This appendix is a draft version and a living document. Changes within the methodologies below are continuous. This is simply a current snapshot assessment at this point in time.

Introduction

The fast paced world of electronic commerce and electronic data interchange has been evolving through the efforts of numerous organizations, agencies, and standards' bodies. This somewhat uncoordinated evolution has resulted in numerous methodologies and theories for the implementation of widespread EC/EDI. Current ANSI X12 users are considering the numerous approaches and needs to become UN/EDIFACT enabled for international EC/EDI capabilities. This appendix will explore the leading methodologies and current works related to the various backing organizations and schools of thought. The following methodologies or works are discussed and researched in later sections:

- Business Information Model (BIM)
- Basic semantic Repository (BSR)
- Defense Data Dictionary System (DDDS)
- OPEN-EDI -- Object-Oriented EDI (OO-EDI)
- Universal Date Element Framework (UDEF)

Current Assessment

The above listed bodies of work are all related to one another in several manners. The goal of these organizations is too advance the need and usefulness of electronic commerce and electronic data interchange. At the moment, it appears that many organizations and agencies are working together as well as independently to obtain the goal of electronic commerce interoperability. UN/ECE has backed numerous projects to advance the development of EC/EDI within the UN/EDIFACT messages. Areas of involvement include: Interactive EDI, Batch EDI, and OPEN-EDI. The open model is being development to allow for an open system of EC/EDI implementation. The OPEN-EDI work has produced an OPEN-EDI Reference Model to accomplish this task. Also, OPEN-EDI has laid out a future plan with three major tracks for EC/EDI :

1. Mainstream EDI
2. Simpler EDI
3. Object-Oriented EDI

Track 3: Object-Oriented EDI requires modeling based on the IDEF models which would incorporate the works of the Business Information Model (BIM).

The BIM Group is primarily concerned with developing a Framework for EDIFACT Message development that applies existing standards to improve the speed and quality of work. The BIM incorporates the OPEN-EDI Reference Model, input from the X12 BIM group via the Pan American EDIFACT Board (PAEB) BIM group, and ISO standards such as the ISO 11179 naming conventions. The BIM Warehouse utilizes the objects, entities, and attributes from the Basic Semantic Repository (BSR) as well.

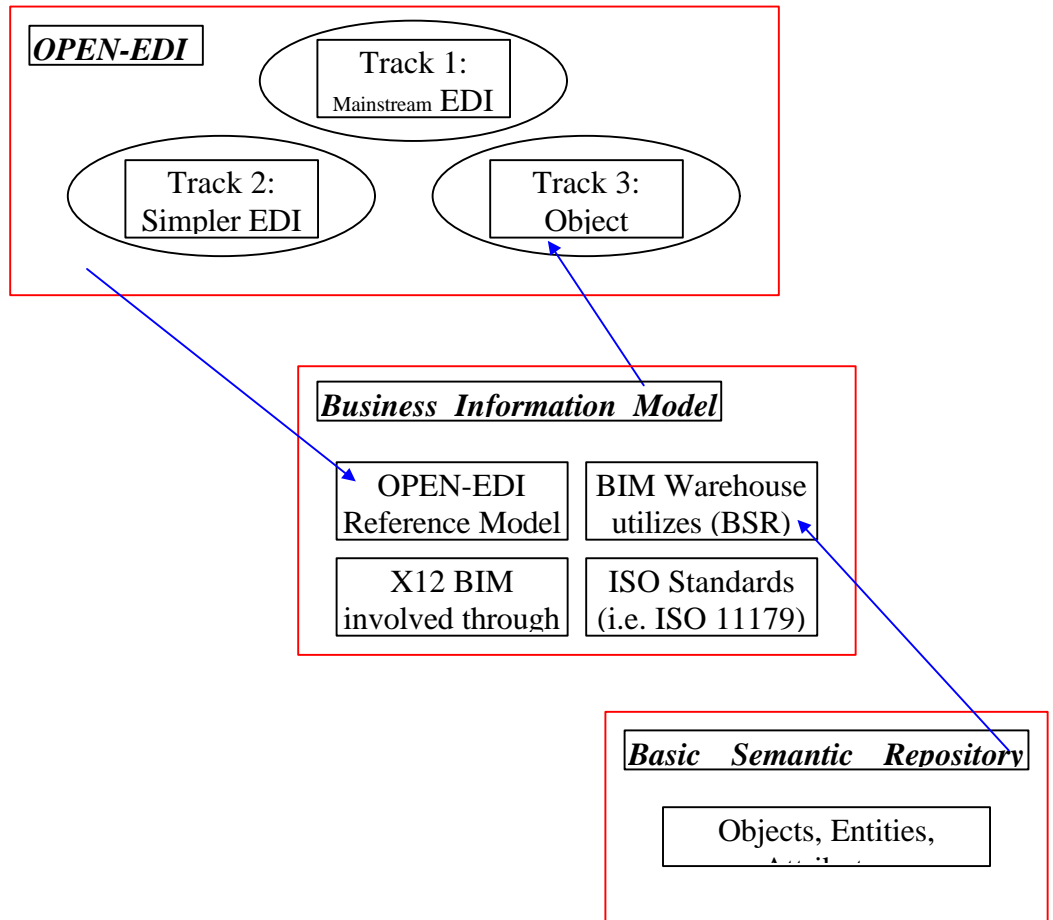


Figure D-1: Methodology Interrelationships

The question is raised as to whether this single plan is the correct model for future implementation of EC/EDI. Other works and means have been used to attack the problems of EC/EDI in today's society. One might try to utilize the Defense Data Dictionary System (DDDS) structure to move from an ANSI X12 transaction set system to an international usage of UN/EDIFACT messages. As well, another approach being developed is the Universal Date Element Framework (UDEF). This framework may lend to another methodology for future migration of ANSI X12 to UN/EDIFACT.

The section to follow provides a quick look methodology matrix for comparison of the aforementioned topics.

Table D.1: Matrix Evaluation of Methodologies

Methodology	ANSI X12 / EDIFACT	Maturity	Sponsoring Organization	Object naming convention standards	Weakness	Strength	Application to real world problems?	Reuse within scope	Automation tools	Status
BIM	UN/EDIFACT; consideration given to X12 concerns	1990s development; modeling techniques are more mature than methodology as a whole for use in EDI	ISO	ISO 11179 coordination	Limited scope to EDIFACT messages only.	Business functionality consideration given; Incorporation of standards communities	EDIFACT groups utilized in the development of messages.	√	Software suite of tools developed and available	Continued development of modeling frameworks and methodology maturity
BSR	Standard independent	1993 prototype; 1995 working group formed; new chair and direction being formulated today	ISO & UN/ECE	ISO 11179	Semantic guidelines and definition vagueness; object defined how?	Reduction to semantic units rather than data element analysis	Completed 60% mapping of ORDERS message. Working on Purchase Order transaction set.	√	None developed beyond database warehouse.	Continued mapping of remaining 40% of ORDERS message as well as other additional messages
DDDS	Standard independent	1995 directive initiated program	Defense Information Systems Agency; DoD	ISO1179 partially used	Not directly intended for EC/EDI purposes; Object Classification confusion; no properties; adds another standard for confusion	Consistent data element classification	Used in mapping the 850 transaction set to the ORDERS message	√	PCAT	Continued data element definition and database expansion
OPEN-EDI (Object Oriented-EDI)	Standard independent; EDIFACT used as basis	Early stage of development for OO-EDI (1996)	ISO & UN/ECE (CEFACT)	None proposed due to modeling methodology approach	Unproved; not well defined process yet	Future Vision; Object oriented approach	Tested theory on an OPEN-EDI scenarios.	√	Not yet; easily implemented with OO technology	Working group actions of 3 track systems for development; OO-EDI continuing scenarios assessment and development
UDEF	Standard independent	Initial prototype stages	CALS Steering Group	ISO 11179	Maturity and refinement; consistent definition of semantic units	Semantic assessment independent of standards.	Application to various ANSI X12 sets and UN/EDIFACT messages	√	Prototype software.	Continued prototype work and mapping scenarios

D.2 Business Information Model (BIM)

The BIM presents a new approach to the development and documentation of EDI messages. It outlines a framework for the application of structured modeling techniques which EDIFACT message development groups can use as the basis for the creation of messages and implementation guides. These techniques deliver formal specifications of Business requirements and EDI requirements and the framework links these to the message design through the application of EDIFACT Message Design Guidelines. Through the use of computer tools these processes can be carried out speedily and effectively. Although specifically directed to UN/EDIFACT the framework described is applicable to other EDI standards and facilitates a move towards "Open-EDI".

The BIM Framework is designed to be modeling method, technique and tool independent. A primary principle is the use of existing standards to develop standards.

The methods outlined are already being used in some EDIFACT groups (e.g. Finance, Construction and Health Care) to help with message design and by user groups (e.g. CEFACT, ODETTE and EDIFICE) to help guide implementation of existing messages.

The BIM Framework is based on a top down iterative process for modeling which converts business world perspectives into formal structured representations of activities and data. The models can be used to design consistent EDIFACT messages.

The framework consists of three phases:

1. Business Analysis Phase - model requirements for business functionality
2. EDI Requirements Phase - further modeling related to EDI requirements
3. EDIFACT Message Design Phase - message design and changes implemented

Other aspects of the framework:

- BIM Warehouse – incorporated for reusability.
- Different groups can perform the 3 phases so as to get experts in each area.

Incorporation of other standards

The BIM Group is primarily concerned with developing a Framework for EDIFACT Message development that applies existing standards to improve the speed and quality of work. Some of the relevant standardization activities carried out by various bodies are given below.

- ISO standardization work in EDI.

- Open-EDI -- "Open-EDI reference model" ISO/IEC CD 14662 for the coordination of Open-EDI standards
- Naming Standards etc. (ISO 11179)
- The Basic Semantic Repository - - These semantic units should include the Objects, Entities and Attributes defined by the application of the BIM methods to different Business Areas and thus should provide part of the BIM Warehouse facilities.
- EXPRESS and STEP – modeling standards
- ANSI-X12 -- The Business and Information Modeling Group in X12 are closely involved with the JRT BIM Group via the PAEB BIM Group.
- CEN --The European Standards Committee is currently publishing a series of UN/EDIFACT messages as European Norms.
- EWOS -- The European Workshop on Open Systems has an Expert Group on EDI and has published the EWOS Technical Guide on EDI.

D.3 Basic Semantic Repository (BSR)

The purpose of the BSR project is to construct a technical infrastructure which provides storage, maintenance and distribution facilities for reference data about semantic units and their links with operational directories. It will provide the following principal function which is to build cross references between different EDI directories which are maintained by different agencies. The founding partners of the BSR project are ISO and UN/ECE.

As a consequence of establishing the BSR, it will also be of use in:

- information modeling process within EDI
- defining new entries or redefining existing entries in existing directories,
- input for standards work in the relevant standards areas.

The BSR is seen solely as a tool (source of information and guidance) to be used by designers and implementers of EDI systems. The BSR is not envisioned as a new standardized data element directory or data concept repository.

The components of the Basic Semantic Repository (BSR) will be :

- Basic Semantic Units (BSUs) – (Intermediate Construct, Generic Intermediate Construct, Specific BSU, Generic BSU, and Process Level BSU)
 - Each BSU can be comprised of up to (4) Semantic Components
 - Concept
 - Generic – date format
 - Specific – most common: payment
 - Property Class – indicators, identifiers

- Process Type – purchase order, invoice
- Bridges
- List of concepts
- Glossary of controlled and agreed terms

What the BSR is not:

- the BSR is not a standard but a tool.
- the BSR is not a thesaurus.
- the BSR is not a system of concepts (as defined by ISO TC 37) as the BSR records only generic relations).
- the BSR is not an EDI directory/repository

The objective is to establish an operational BSR which will provide end users with the facilities they need for unambiguously linking their data representations in existing EDI directories. The BSR will be available to the public.

D.4 Defense Data Dictionary System (DDDS)

The DDDS, maintained by DISA, is the primary tool supporting DoD Data Administration. DDDS is a repository that houses, and enables centralized management of, the DDM and related information. DDDS contains all DoD standardized data elements and provides automated support for developing and managing standard data in accordance with DoD Directive 8320.1.

DDDS provides developers with access to approved standard elements and worldwide on-line query and reporting capability. DDDS also collects and stores standard data elements and attributes. It allows users to participate in the review and approval of data standards by functionality, identify DoD organizations and processes using standard elements, and track the status of data elements through the standardization process. DDDS also allows users to access metadata (data about data), including name, definition, data steward, field length, data type, and allowable domain values that are used by software developers in DoD application development.

DISA developed the PCAT software tool to provide DDDS users with a user-friendly method to access DDDS and to allow users to participate in augmenting the more than 11,070 DoD data standards available.

D.5 OPEN-EDI – to -- Object Oriented EDI (OO-EDI)

The EDI standards of tomorrow must be able to deal with current industry usage and all new industries that are adopting EDI. In order to accomplish this, the standards must be developed to utilize "across industry sectors" such as banking, manufacturing, chemical, insurance, transportation and healthcare, only to name a few. Further, tomorrow's EDI

must be able to not only deal with standard business data such as quantity and price, but also must deal with the related data such as electrical and mechanical characteristics.

EDI Problems Identified

1. EDI solves the problem of disparate operating systems, applications and computer models; but communications issues are still not being solved.
2. There are different interpretations on how to implement an EDI message. There is no agreed upon semantic and the standards contain every industry's solutions.

OPEN-EDI proposes to solve their defined problems through the following requirements:

Requirements on the Open-EDI Reference model:

- Organizational flexibility
- modeling of the Business Information
- plurality of sectors and information types
- public standards for interoperability
- easy for new partners to join in
- EDI-Transactions (scenarios)
- No prior agreement necessary and total integration.

ISO/IEC JTC 1/SC 30 efforts to CEFACT's future vision

SC 30 is the 30th Subcommittee of JTC 1, which in turn is the first (and only) Joint Technical Committee of ISO and IEC. Because of this, SC30 has a strong link to any of the Technical Committees of ISO, and Technical and Subcommittees of IEC. That does not preclude the link to other JTC 1 SCs. Further SC 30 has a strong liaison to UN/EDIFACT via T TN/F.CF./TR ADE/WP.4 and the ITU-TS .

JTC 1's technical strategy was to make the Open-EDI reference model its center and develop standards within that framework. The SC 30 produced three work items, the Open-EDI Reference Model, Business agreements using semantic descriptive techniques, and EDI support services. The Open-EDI Reference Model is the framework for coordinating standards development. It defines two views called the Business Operational View (BOV) and the Functional Service View (FSV). The model defines the main concepts of the two views and identifies the necessary standards. The added value of the Open-EDI Reference Model is the formalization of the concepts, identification of the classes of standards, and the basis for the coordination of future and existing standards. This includes work with other agencies outside of SC 30.

From the Open-EDI Reference Model approach the future of UN/EDIFACT is being charted with a strategy for the next phase. UN/EDIFACT at the strategic level has not had the impact that was predicted for it in two significant areas. The first is in the general rationalization and simplification of the business and administrative process and the second

is in the take up and use of the standard by small and medium sized enterprises (SME's). The EDIFACT Steering Group (ESG) has been working to prepare a coherent forward strategy for UN/EDIFACT which takes these into account and also places the strategy in the context of CEFACT's (Centre for the Facilitation of Administration, Commerce and Trade) Mission Statement. Accordingly the ESG now proposes, initially for consultation with the JRT and subsequently for incorporation into CEFACT's Work Program, which consists of three inter-linked strands which should be considered as parallel activities or "tracks.

- 1) Continuation of the development and maintenance of UN/EDIFACT as the global message standard based on both batch and interactive syntax's. (Track 1 - Mainstream UN/EDIFACT)
- 2) Encouragement of the development of simpler UN/EDIFACT messages relevant to the needs of SME's (Track 2 - Simpler UN/EDIFACT)
- 3) Full development of the Object Oriented approach to the design of future messages. (Track 3 -Object Oriented EDI, "OO/EDIFACT")

Committee/Working group activities will continue within the three track areas above to define the future direction of EC/EDI within the UN/EDIFACT framework.

D.6 UDEF

A complete discussion of the Universal Data Element Framework is provided within the main text of this paper. Please refer to the numerous sections detailing the concept and performance expectations of the methodology.

APPENDIX E: REFERENCES

UN / EDIFACT: The Business and Information Modeling Framework for UN/EDIFACT

Document Ref.: BIM2

Version Date : 26/4/95

Issued by : T5 BIM Group

Status : Draft for Comment

Universal Data Element Framework

Ronald L. Schuldt

Chairman, CALS Enterprise Data Task Group

<http://arioch.gsfc.nasa.gov/wwwvl/de.html>

BSR Basic Reference Document

Premenos Site

<http://www.premenos.com>

<http://www.iso.ch/BSR>

Basic Semantic Repository (BSR) Project

BSR/MC Executive Committee

Correction 1

1996-08-11

Basic Semantic Repository (BSR) Project

BSR/MC Executive Committee

Correction 1

1996-05-09

Open-EDI Reference Model Background information

Premenos Site

<http://www.premenos.com>

Open-EDI Reference Model

ISO/IEC DIS 14662: 1996 (E)

Reference Guide: The Next Generation of UN/EDIFACT

AC.1 Research Strategic Advice and Implementation Planning Group

PC Access Tool (PCAT) Software

Defense Information Systems Agency

Center for Computer Systems Engineering

DoD Data Administration

DDDS Reference

*<http://ssed1.ncr.disa.mil/srp/datadm/mission.html>
& <http://cfcse.ncr.disa.mil/jexhome/brochure/disafig4.gif>*

APPENDIX F: UDEF SHORTFALLS

TBD