ISO/IEC JTC 1/SC 32 N 0462

Date: 2000-

REPLACES: --

ISO/IEC JTC 1/SC 32

Data Management and Interchange

Secretariat: United States of America (ANSI)
Administered by Pacific Northwest National Laboratory on behalf of ANSI

DOCUMENT TYPE	Other Document (Open)
TITLE	Making Standards Work in Electronic Commerce and Among Jurisdictions: IT-Enablement of Data Element-based Standards - Presentation at the Open Forum on Metadata Registries in Santa Fe
SOURCE	Dr. Jake Knoppers
PROJECT NUMBER	1.32.18.01.00.00
STATUS	This input to the New Project "Information technology - Identification, Mapping and IT-enablement of Existing Standards for Widely Used Encodeable Value Domains"
REFERENCES	
ACTION ID.	FYI
REQUESTED ACTION	
DUE DATE	
Number of Pages	24
LANGUAGE USED	English
DISTRIBUTION	P & L Members SC Chair WG Conveners and Secretaries

Douglas Mann, Secretariat, ISO/IEC JTC 1/SC 32

Pacific Northwest National Laboratory *, 901 D Street, SW., Suite 900, Washington, DC, 20024-2115, United States of America

Telephone: +1 703 575 2114; Facsimile; +1 703 681 9180; E-mail: MannD@battelle.org

available from the JTC 1/SC 32 WebSite http://bwonotes5.wdc.pnl.gov/SC32/JTC1SC32.nsf

*Pacific Northwest National Laboratory (PNL) administers the ISO/IEC JTC 1/SC 32 Secretariat on behalf of ANSI

Making Standards Work in Electronic Commerce and Among Jurisdictions: IT-Enablement of Data Element-based Standards

presentation at the

Open Forum on Metadata Registries

Santa Fé, NM (USA)
19 January, 2000

by

Dr. Jake V. Th. Knoppers, Senior Vice-President
Canaglobe International Inc.
(Tel:+1-613-234-3244; Fax:+1-613-234-3935;
E-mail:mpereira@istar.ca)

OUTLINE

standards, as well as “new new” standards).

1.2 WHAT IS "E-COMMERCE"? STANDARDIZATION & BUSINESS OPERATIONAL VIEW (BOV) PERSPECTIVE

Already Defined Terms

ISO/IEC 14662 Information technology - Open-edi reference Model

"Open-edi", "business", "business transaction", "electronic data interchange (EDI)" [as well as "business operational view (BOV), functional services view (FSV)]

ISO/IEC 6523-1 Information technology - Structure for the identification of organizations and organization parts -

"organization", "organization part"

ISO/IEC CD 15944-1 Information technology - Business semantics descriptive techniques - Part 1: Open-edi implementations

"person", "individual", "buyer", "seller", "regulator"

Utilizing these standards with their already defined terms to develop a definition for e-commerce results in the following definition for "electronic commerce":

> ***Electronic commerce****: a category of business transactions, involving two or more persons, enacted through (open) electronic data interchange using computer and telecommunications technologies.*
>
> *[Note: Persons can be individuals and/or organizations.]*

[commerce: a category of *business transaction* which involves the exchange of goods or services, between *buyers* and *sellers*, most often on a for-profit basis]

Other categories of "business transaction" include: "e-administration" (mostly on a not-for-profit basis), e-business (a general term), e-government, e-tailing, e-medicine, e-etc.

2.0 HORIZONTAL ISSUES (1/2)

A. Four horizontal issues were identified by the ISO/IEC JTC1 Business Team on Electronic Commerce (BT-EC) [ISO/IEC JTC1 N5296] which are of general relevance for all scenarios involving electronic commerce. They need to be addressed by all standardization work in support of e-commerce. The four horizontal issues are:

1. IT-enablement
2. Localization and Multilingualism
3. Cross-sectorial
4. Cultural adaptability

Issues 1) and 2) are the most "do-able". Resolving them will help resolve 3) & 4).

B. The ISO/IEC JTC1 BT-EC ordered these four horizontal issues on the basis of:

1. the need to go from simpler to more complex challenges;
2. placing priority on the "do-able" and immediately useful in the context of increasing resource constraints in standardization work; and,
3. promotion and visibility of ISO, IEC and ITU especially outside of these standardization activities.

C. A key characteristic of commerce world-wide, in particular in the organization-to-organization and organization-to-public administration domains, is that it consists of business transactions which:

1. are rule-based, i.e., mutually understood and accepted sets of business conventions, practices, procedures, etc.; and,
2. make extensive use of "lists", "codes", often table-based, representing pre-defined possible choices for common aspects of business transaction, (e.g., countries, languages, currencies, etc.). The problem is that most of these lists/code sets in use worldwide are paper-based and lack a computer-processable version.

D. What is needed is a standard to be used as a tool for (1) creating IT-enabled versions of existing sets of "codes representing X"; and, (2) for building new "Codes representing X" standards.

2.0 **HORIZONTAL ISSUES (2/2)**

Working definitions for the four horizontal issues developed by the BT-EC are:

- **IT-enablement** : *"the transformation of current standards utilized in commerce, (e.g., code tables), from a manual to a computational perspective so as to be able to support computational integrity".*

  ("Computational integrity: the expression of standards in a form that ensures precise description of behaviour and semantics in a manner that allows for automated processing to occur, and the managed evolution of such standards in a way that enables dynamic introduction by the next generation of information systems").

- **Localization**: *"pertaining to or concerned with anything that is not global and is bound through specified sets of parameters of:*

  (a) a linguistic nature including natural and special languages and associated multilingual requirements;
  (b) a jurisdictional nature, i.e., legal, regulatory, geopolitical, etc.
  (c) a sectorial nature, i.e., industry sector, scientific, professional, etc.
  (d) a human rights nature, i.e., privacy, disabled/handicapped persons, etc.; and/or,
  (e) consumer behaviour requirements".

  Within and among these parameters, interoperability and harmonization objectives apply.

  [This working definition for "localization" incorporates existing JTC1 standardization work of POSIX on "locale" and SC22/WG20 on "internationalization" a.k.a. "i18n"]

- **Multilingualism:** *"the ability to support not only character sets specific to a (natural) language (or family of languages) and associated rules, but also localization requirements, i.e., use of a language from jurisdictional, sectorial, and consumer marketplace perspectives."*

3.0 TWO JTC1/SC32 PROPOSED NEW WORK ITEMS (NWIs)

3.1 OVERVIEW - SCOPE AND STATUS

[Note: Both NWIs are in response to high priority work items requiring standardization work identified by the ISO/IEC JTC1 BT-EC, from which the titles are taken]

NWI - IDENTIFICATION AND MAPPING OF VARIOUS CATEGORIES OF JURISDICTIONAL DOMAINS [ISO/IEC JTC1 N5846]
[Responsibility of ISO/IEC JTC1/SC32/WG1]

SCOPE: ISO standardization in the field of identification and mapping of various categories of jurisdictional domains to be done in conformity with existing requirements and practices of international, regional, and national authorities (with a priority on those impacting several sectors of electronic commerce).

[Note: Jurisdictions are the major source of external constraints on business transactions.]

JTC1 Ballot Results (N6001) 18 in favour + 1 negative, but only 4 P-members initially agreed to actively participate in work. Thus ballot "failed" until 1 or more P-members of those voting in favour agree to participate. [As of today, 19 January, 2000, the UK and Australia have now voted to actively participate in the work and the USA is in the process of doing the same. So the ballot has now "passed"].

NWI - IDENTIFICATION, MAPPING AND IT-ENABLEMENT OF EXISTING STANDARDS FOR WIDELY USED ENCODABLE VALUE DOMAINS [ISO/IEC JTC1 N5847]
(Responsibility of ISO/IEC JTC1/SC32/WG2 in close liaison with SC32/WG1)

SCOPE: ISO standardization in the field of identification, mapping, and IT-enablement of existing encodable value domains resulting from common business practices which are rule-based. This work is to be done in conformity to existing requirements with a priority on the code sets used in several sectors of electronic commerce.

JTC1 Ballot Results (N6002) Ballot passed. 19 in favour + 1 negative. The minimum of 5 P-members participating was achieved. [France felt this NWI should be part of SC32/WG1 work. Note: There is and will be close liaison with SC32/WG1 on this NWI].

The rest of this presentation focuses on this NWI, a.k.a, "IT-enablement of Codes Representing X"

3.2 PRACTICAL APPROACH TO NWI: IT-ENABLEMENT OF "CODES REPRESENTING X"

1. Many of the standards (ISO, IEC, ITU and others) for use in E-commerce are value domains. They are of the nature of value domains which are bounded sets in the sense that the value domain and the set of permitted values in that domain are pre-defined and enumerated in the standard.

Most of these are of a "Codes representing X" nature. From a global E-commerce perspective, standardization work is required for the identification and referencing of such objects in an unambiguous, linguistically neutral, IT-processable and in an E-commerce facilitated manner. These standards need to be: (1) re-cast in a computer processable form in order to support more fully the objective of computational integrity, (a key part of IT-enablement); and, (2) in a manner which supports localization and multilingual requirements. In short, a standard which will serve as a tool to be used by others.

2. The codes sets represent agreed upon common business rules and practices. However, these code sets represent only the "tips of icebergs" of the common business rules and practices, many of which are not explicitly stated. Even if these code sets are distributed in electronic form they cannot be "plugged in" for use in Open-edi implementations including E-commerce. Much of the intelligence in ISO and non-ISO code sets is human understandable or discernable, explicitly or implicitly. From an analytical/logical IT perspective, (e.g., entity-relationship, object-oriented modeling), these code sets are not IT-enabled and lack computational integrity. Nor have such code sets been formally described using Formal Description Techniques (FDTs).

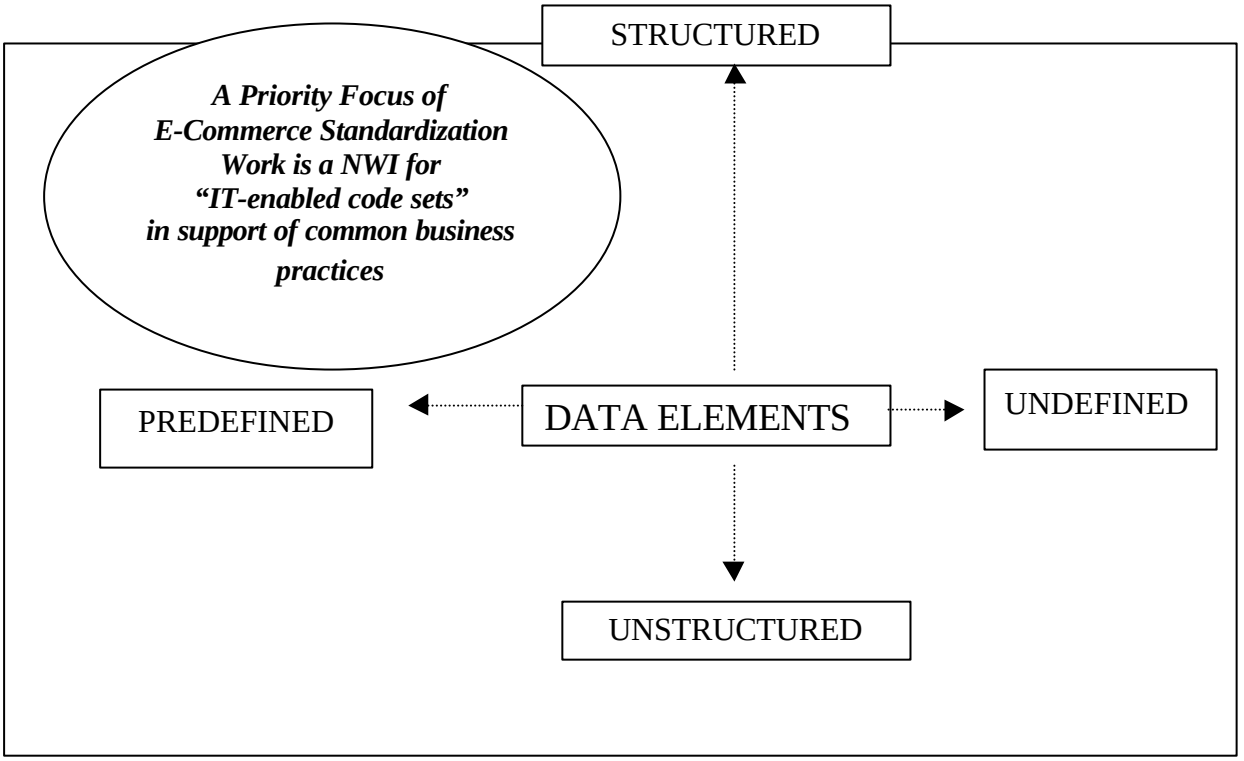
Consequently, those implementing these code sets have to spend considerable resources: (1) to figure out and interpret them; (2) build applications; and, (3) pray that their interpretation and use, (as imbedded in the application), is interoperable when they interwork with heterogeneous information systems of autonomous enterprises on a global and multi-sector basis.

3. There are well over 150 ISO standards which involve the use of "code sets", these need to be mined for content elements, rules and practices for building this new standard.

4. Existing standard terms and definitions will be utilized, when and wherever applicable. To avoid the existing standards problem of polysemy, when required, new terms will be invented, i.e., "coined", to accompany any needed new concepts/definitions.

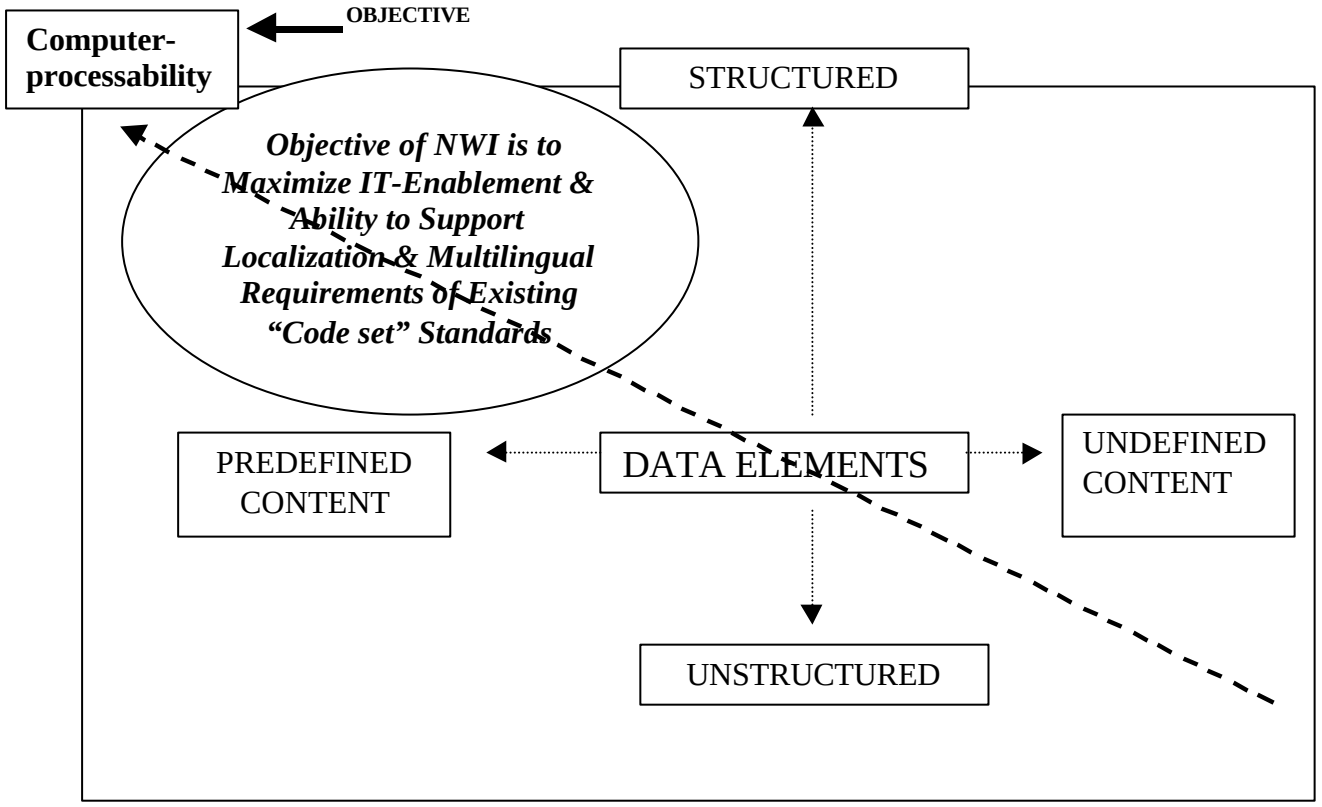
5. Those responsible for existing "code set" standards, would use this new standard as a tool to develop the "equivalent" IT-enabled version. Their participation in the NWI standards development work should generate accompanying case studies.

3.3 POSITIONING THIS NWI RELATIVE TO BOTH SC32/WG1 - OPEN-EDI AND SC32/WG2 - METADATA (1/2)

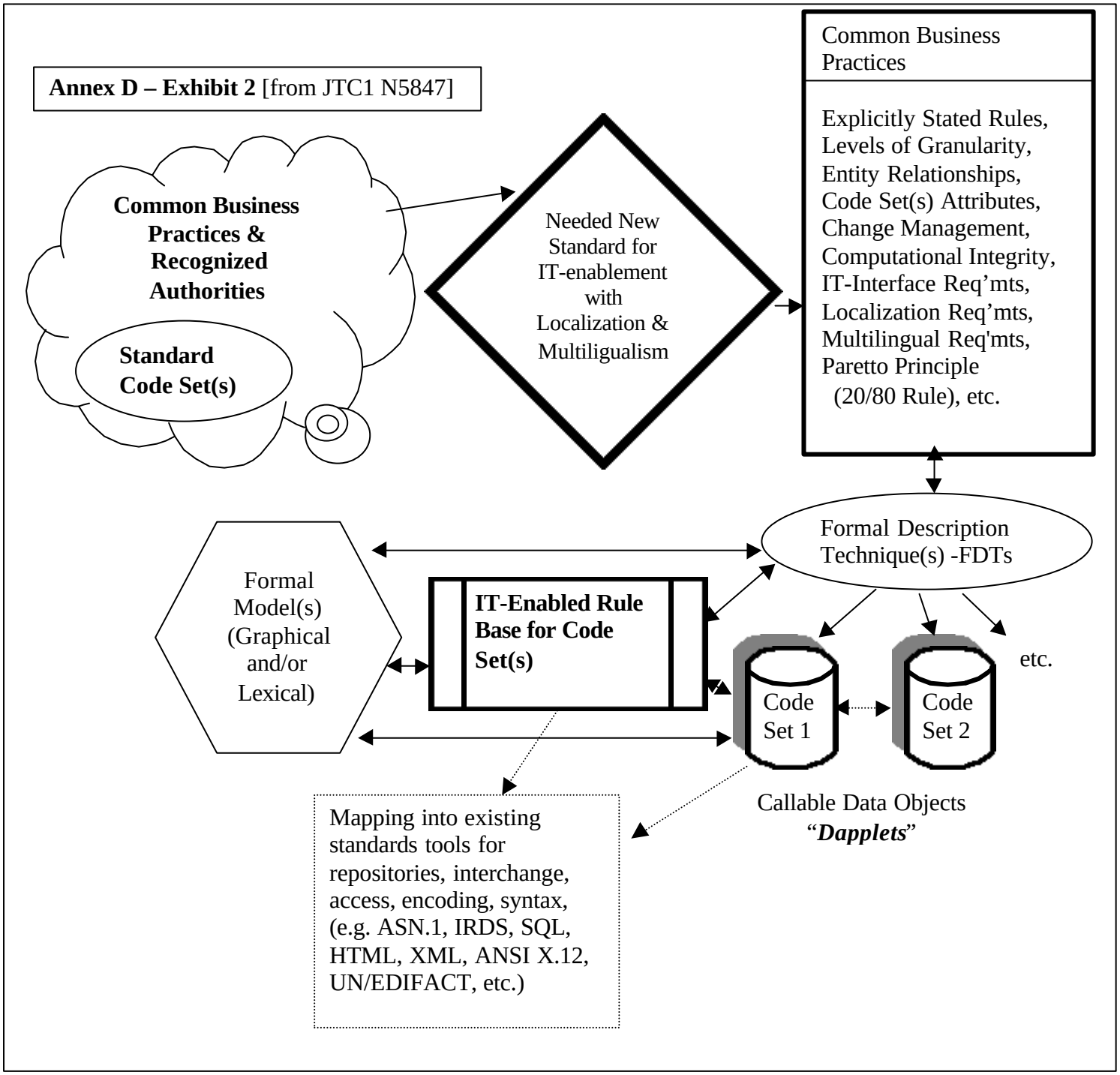


Annex D – Exhibit 1 [from NWI Proposal JTC1 N5847]

3.3 POSITIONING THIS NWI RELATIVE TO BOTH SC32/WG1 - OPEN-EDI AND SC32/WG2 – METADATA (2/2)



3.4 OVERVIEW OF KEY ELEMENTS TO BE ADDRESSED FOR PROGRESSING WORK ON THIS NWI



4.0 CASE STUDIES IN PROGRESS - REAL WORLD EXAMPLES OF IT-ENABLEMENT WITH LOCALIZATION AND MULTILINGUALISM (IT-INTERFACE & HUMAN INTERFACE REQUIREMENTS)

4.1 EXAMPLE 1: INTERNATIONAL COMMERCE - “POTATO”

4.2 EXAMPLE 2: COUNTRY CODES BASED ON ISO 3166

- 4.2.1 LEVEL 1 - GENERAL
- 4.2.2 LEVEL 1 - ZAIRE
- 4.2.3 LEVEL 2 - CANADA & NUNAVUT

4.3 EXAMPLE 3: SIMPLE TOPOLOGY BASED ON ISO CD 19107 (TC211-GEOMATICS) USING UML IN ENGLISH, FRENCH AND MANADARIN CHINESE

4.1 EXAMPLE 1: INTERNATIONAL COMMERCE - "POTATO"

Example 1 focuses on human understandable representation of what should be an IT-enabled global standard for international trade in goods based on the Harmonized System (HS) of the World Customs Organization (WCC). The example here is "potato" (fresh or chilled) for which the HS code is "0701".

The example demonstrates:

- a jurisdiction, in this case a country, having more than one language of use and thus multilingual equivalents;
- differences in uses of the same natural language in various countries and thus different multilingual equivalences within a natural language.

Common IT Interface	**Country Code - Numeric Code & Short Name (en) Equivalent**	**Human Interface: Localization and Multilingual Equivalents**
HS: 0701	124 CANADA	(en): potato (fr): pomme de terre (ik): patiti
HS: 0701	464 MEXICO	(es): papa
HS: 0701	724 SPAIN	(es): patata
HS: 0701	040 AUSTRIA	(de): erdapfel
HS: 0701	276 GERMANY	(de): kartoffel
HS: 0701	056 BELGIUM	(fr): pomme de terre (nl): aardappel
HS: 0701	246 FINLAND	(fi): peruna (sv): potatis

4.2 EXAMPLE 2: COUNTRY CODES BASED ON ISO 3166

4.2.1 LEVEL 1 - GENERAL EXAMPLE (taken from ISO 3166-1)

Example 2 demonstrates not only the six (6) ISO human interface linguistically equivalent expressions (of which two are also used as ISO alpha codes) to the single one (1) linguistically neutral common numeric code (the pivot) for each permitted instance in ISO 3166-1 , but also the "official" human interface linguistic equivalent in the source jurisdictions.

Common IT Interface	**Human Interface : Localization and Multilingual Equivalents**			
3166:246	ISO Alpha-2:	FI	ISO Alpha-3:	FIN
3166:246	ISO Short Name(en):	Finland	ISO Long Name (en):	Republic of Finland
3166:246	ISO Short Name(fr):	Finlande	ISO Long Name (fr):	République de Finlande
3166:246	Local Short Name(fi):	Suomi	Local Long Name(fi):	Suomen tasavalta
3166:246	Local Short Name(sv):	Finland	Local Long Name(sv):	Republiken av Finland
3166:056	ISO Alpha-2:	BE	ISO Alpha-3:	BEL
3166:056	ISO Short Name(en):	Belgium	ISO Long Name (en):	Kingdom of Belgium
3166:056	ISO Short Name (fr):	Belgique	ISO Long Name (fr):	Royaume de Belgique
3166:056	Local Short Name(nl):	Belgie	Local Long Name(nl):	Koninkrijk van Belgie
3166:056	Local Short Name(fr):	Belgique	Local Long Name(fr):	Royaume de Belgique
3166:792	ISO Alpha-2:	TR	ISO Alpha-3:	TUR
3166:792	ISO Short Name(en):	Turkey	ISO Long Name (en):	Republic of Turkey
3166:792	ISO Short Name (fr):	Turquie	ISO Long Name (fr):	République turque
3166:792	Local Short Name(tr):	Turkiye	Local Long Name(tr):	Turkiye Cumhuriyeti

4.2 EXAMPLE 2: COUNTRY CODES BASED ON ISO 3166

4.2.2 LEVEL 1 – ZAIRE

Example of a change in all six of the ISO human interface equivalent linguistic expressions with (the pivot) numeric code staying the same. The numeric code for ZAIRE is "3166:180".

	Former	New
Alpha-2	ZR	CD
Alpha-3	ZAR	COD
Short Name (en)	Zaire	Congo, Democratic Republic of the
Long Name (en)	Republic of Zaire	The Democratic Republic of Congo
Short Name (fr)	Zaïre	Congo, la République démocratique du
Long Name (fr)	République de Zaïre	La République démocratique du Congo

4.2 EXAMPLE 2: COUNTRY CODES BASED ON ISO 3166

4.2.3 LEVEL 2 - CANADA & NUNAVUT – ISO 3166-2 SUB-DIVISIONS

This example demonstrates the requirements for human interface linguistic equivalent terms: (1) of a multilingual nature; (2) using non-Latin 1 character sets; and, (3) supporting a language that have two writing systems in a set of codes representing X.

On 1 April, 1999, the Nuvavut Territory was established through an Act of the Canadian Parliament, joining the existing provinces and territories. The establishment of Nunavut included the introduction of Inuktitut as an official legal language in this part of Canada. Nunavut has chosen "NU" as its two alpha code under ISO 3166-2:1998 Subdivisions (as well as for its Internet 2nd TLD ID). This required a change to the ISO 3166 Part 2 Subdivisions entry for Canada as well as in various Canadian standards. The example given here is that taken from the **Canadian National Standard for geographical metadata – CAN/ CGSB 171-3 "Directory Information for Geo-Referenced Datasets".**

Inuktitut is an oral language with a long history. In the 18th century, two writing systems became associated with the oral language. One is syllabic based and the other (Roman) Latin-1 alphabet based. The names of the Canadian provinces and territories in Inuktitut are presented in these two writing systems. Note that the Inuktitut names using Latin-1 are the same as in English except for the Northwest Territories which is "Nunattiaq". Development of Col. (5) included determining appropriate character map reference based on ISO/IEC 10646-1:1993 *Information Technology -- Universal Multiple-Octet Coded Character Set (UCS) -- Part 1: Architecture and Basic Multilingual Plane* and its amendments up to Amendment 11:1997 Unified Canadian Aboriginal Syllabics, and the Unicode Standard Version 3.0.

IT Interface		Human Interface/Linguistic Equivalent Terms			
Table ID	Code	English (en)	French (fr)	Inukitut (ik)	
				Syllabic Equivalent	(Roman) Latin-1 Equivalent
(1)	(2)	(3)	(4)	(5)	(6)
171.3:03	AB	Alberta	Alberta	ᐊᓪᐴᕐᑖ	Alberta
171.3:03	BC	British Columbia	Colombie-Britannique	ᐳᕆᑎᔅ ᑲᓛᒻᐱᐊ	British Columbia
171.3:03	MB	Manitoba	Manitoba	ᒫᓂᑑᐸ	Manitoba
171.3:03	NB	New Brunswick	Nouveau-Brunswick	ᓂᐅ ᐳᕉᓐᓯᕕᒃ	New Brunswick
171.3:03	NF	Newfoundland	Terre-Neuve	ᓂᐅᕙᐅᓐᓛᓐ	Newfoundland
171.3:03	NT	Northwest Territories	Territoires du Nord-Ouest	ᓄᓇᑦᓯᐊᖅ	Nunattiaq
171.3:03	NS	Nova Scotia	Nouvelle-Écosse	ᓅᕙ ᓯᑰᔅ	Nova Scotia
171.3:03	NU	Nunavut	Nunavut	ᓄᓇᕗᑦ	Nunavut
171.3:03	ON	Ontario	Ontario	ᐋᓐᑎᕆᐅ	Ontario
171.3:03	PE	Prince Edward Island	Île-du-Prince-Édouard	ᐴᕆᓐᔅ ᐊᑦᕗᐊᕐᑦ ᐊᐃᓚᓐ	Prince Edward Island
171.3:03	QC	Quebec	Québec	ᑯᐸᐃᒃ	Quebec
171.3:03	SK	Saskatchewan	Saskatchewan	ᓴᔅᑳᑐᕚᓐ	Saskatchewan
171.3:03	YT	Yukon Territory	Territoire du Yukon	ᔫᑳᓐ	Yukon Territory

4.3 EXAMPLE 3: SIMPLE TOPOLOGY BASED ON ISO CD 19107 (TC211- GEOMATICS) (1/6)

ISO CD 19107 Geographic Information – Spatial Schema [TC211 N818]

At the Kyoto 9th Plenary of TC211 (September, 1999), an example was presented at the Cultural and Linguistically Adaptability Workshop of a simple topology as specified in ISO CD 19107 (of TC 211 – Geomatics). The example provides for a table containing the objects, their relations and the role names presented in three equivalent different languages along with their modeling using a formal description technique (FDT), i.e., in this example the use of UML based modeling. The example includes the equivalents in the English and French editions of the standard as well as one in Mandarin Chinese.

It should be noted here that the underlying UML was at the outset designed to be linguistically neutral so that the human interface equivalent expressions could be generated automatically through reference the appropriate column in a table.

English edition of standard	
	Objects
1	TS_Root
2	TS_Node
3	TS_Edge
4	TS_DirectedEdge
5	TS_Face
	Relations
6	Within
7	NodeRose
8	Orientation
9	Boundary
	Role Names
10	isolatedNode
11	theFace
12	center
13	spoke
14	edge
15	side
16	bound
17	center
	Constraints
18	(Circular List)

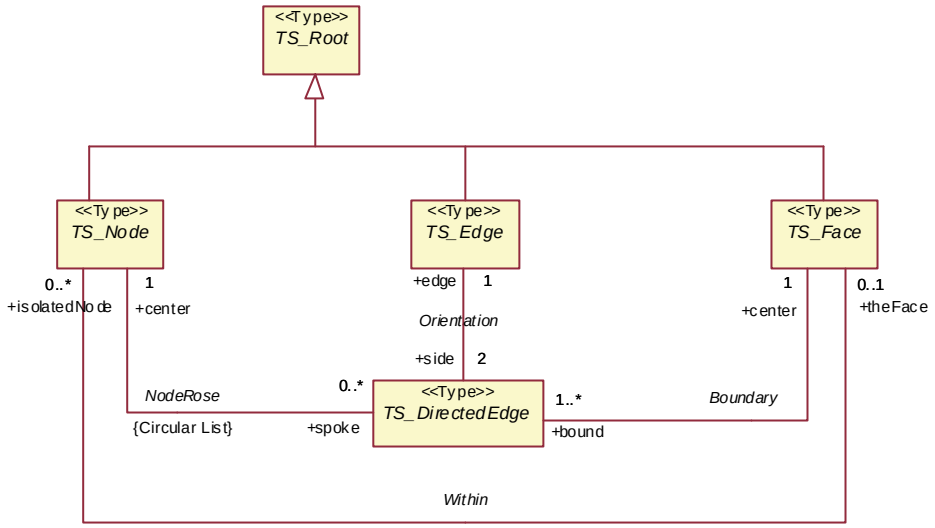
endoftable

4.3 EXAMPLE 3: SIMPLE TOPOLOGY BASED ON ISO CD 19107

French edition of standard	
	Objets
1	TS_Racine
2	TS_Noeud
3	TS_Arête
4	TS_ArêteOrientée
5	TS_Face
	Relations
6	Intérieur
7	NoeudRose
8	Orientation
9	Frontière
	Noms des rôles
10	noeudIsolé
11	laFace
12	centre
13	rayon
14	arête
15	côté
16	limite
17	centre
	Contraintes
18	(Liste circulare)

一	几何体
二	节点
三	线
四	有向线
五	面
六	内点
七	点线连接
八	方向
九	边界

十	特征点
十一	表面
十二	中心
十三	线点连接
十四	线
十五	线侧
十六	区域范围
十七	中心
十八	图形链



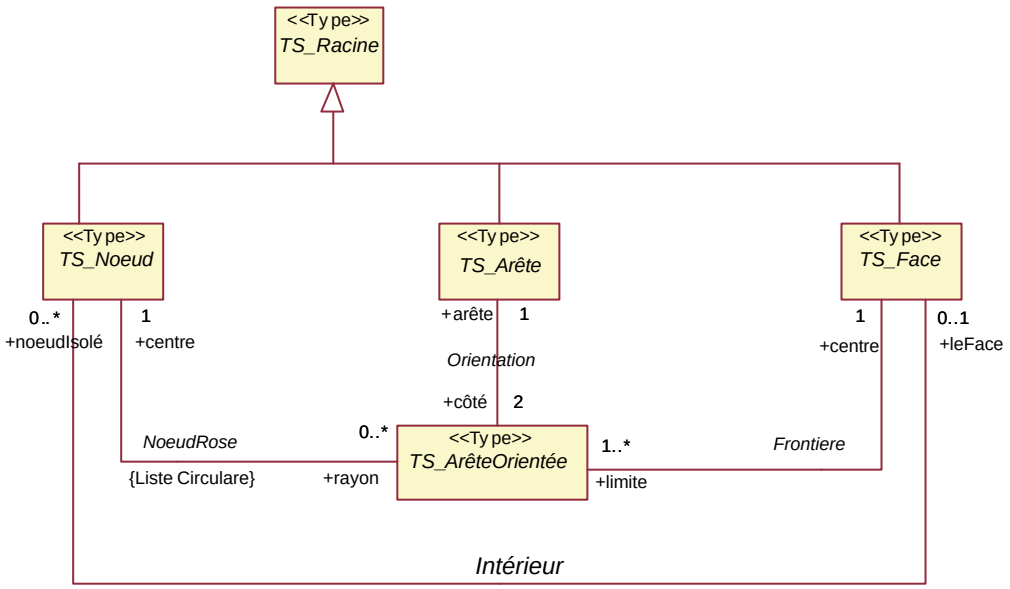
Example of Simple Topology from ISO CD 19107-1

Example from ISO CD 19107 (English edition).This was formerly ISO CD 15046-7.

4.3 EXAMPLE 3: SIMPLE TOPOLOGY BASED ON ISO CD 19107 (TC211-GEOMATICS) FRENCH UML (5/6)

Examples of a "Simple Topology" including a UML model and associated "permitted value domain" presented in three (3) different languages (from a human interface perspective).

UML Model – French language representation (from French language edition of standard)



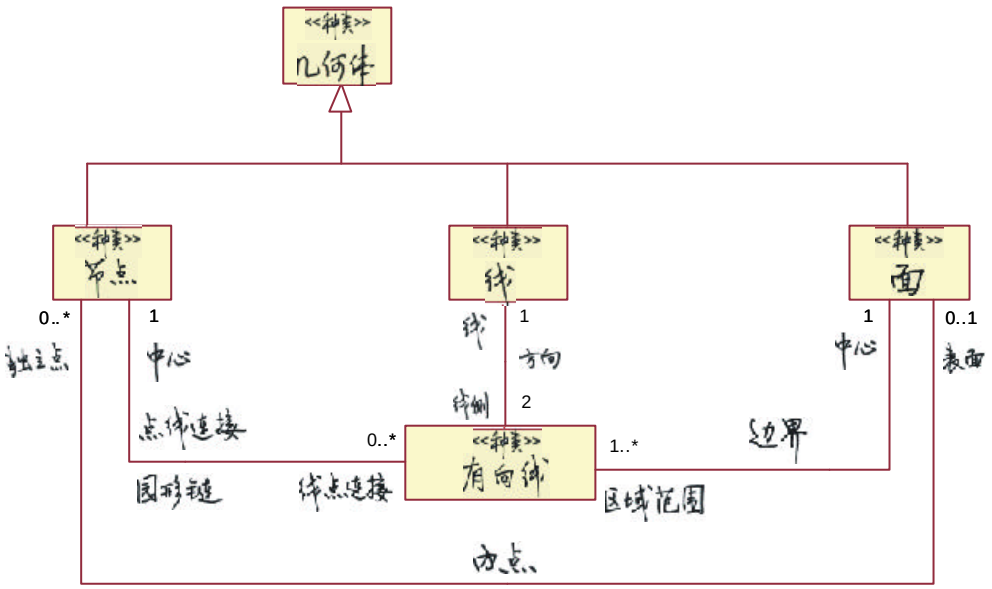
Example of Simple Topology from ISO CD 19107-1 in French

endf

4.3 EXAMPLE 3: SIMPLE TOPOLOGY BASED ON ISO CD 19107 (TC211-GEOMATICS) MANDARIN CHINESE UML (6/6)

Simple topology example from ISO 19107-1 modeled in UML with the textual information represented in Mandarin Chinese as equivalents to those above presented in English and French languages.

[Note: Even though all the textual tables identifying the objects in this UML model are "hand written" here, ITC software and hardware which can support ISO/IEC 10646-1:1993 *Information technology – Universal Multiple-Octet Coded Character Set* (and its amendments), a.k.a. "Unicode" is already available and soon will be made more widely available, (e.g., via Windows 2000). Consequently, unlike a decade ago, there are no longer any technical barriers to being able to support localization and multilingual requirements of the global economy, especially via the Internet].



Example of Simple Topology from ISO CD 19107-1 in Manadrin Chinese

endch

5.0 CONCLUDING REMARKS

(1) Many of the standards used today in commerce worldwide are likely to be utilized in support of electronic commerce. Electronic commerce, particularly business-to-business applications, consist of rule-based transactions that make extensive use of sets of codes, often through tables, representing pre-defined possible choices of common aspects of business transactions.

(2) Standards whose essential nature is that of a list or set of "codes representing x" (or several sets of codes) represent a particular category of data element-based standards. The list of permitted values, i.e., contents, is already pre-defined. However, most of these existing lists and sets of codes are neither IT-enabled nor structured in a manner which can support the multi-jurisdictional and multilingual requirements of the global marketplace.

(3) This presentation has:

- (a) positioned this NWI relative to existing standards and standardization work of JTC1/SC32/WG1 Open-edi JTC1/SC32/WG2 Metadata;
- (b) provided an overview of key elements to be addressed for progressing work;
- (c) presented several real world case studies of existing "codes representing x"; and,
- (d) provided a "first level" representation of the same in the context of IT-enablement and at the same time supporting localization and multilingualism requirements.

(4) This presentation demonstrated that making standards work in electronic commerce and among jurisdictions is a "do-able" and achievable standardization project. The key here is to separate: (1) the IT-interface requirements with the need for unique, and unambiguous (and where possible linguistically neutral) identifiers, i.e., ID for the schema & permitted value instance in a schema; and, (2) the multiple human interface linguistic equivalent expressions (including those which are pre-defined on a jurisdictional basis).

(5) Widespread participation in this standards development work is encouraged. The responsible standardization committee is ISO/IEC JTC1/SC32/WG2. The project editor is Dr. Jake V. Th. Knoppers (CANADA). {E-mail: mpereira@istar.ca}