

TBX-Basic
Translation-oriented Terminology Made Simple

Alan K. Melby
Professor of Linguistics, Brigham Young University at Provo, USA

Resumen: Este artículo resulta especialmente útil para todos aquellos traductores que quieran ir más allá del glossario de dos columnas y deseen profundizar en el lenguaje de intercambio de terminología sin ver XML. Describe un formato de tabla que representa el contenido de los archivos TBX-Basic. La información en este formato se puede convertir a TBX-Basic y, si se desea, procesarla a continuación a otro formato.
Palabras clave: Traducción, intercambio de terminología, XML, TBX, TBX-Basic, tabla, hoja de cálculo.

Resumen: Este artículo resulta especialmente útil para todos aquellos traductores que quieran ir más allá del glossario de dos columnas y deseen profundizar en el lenguaje de intercambio de terminología sin ver XML. Describe un formato de tabla que representa el contenido de los archivos TBX-Basic. La información en este formato se puede convertir a TBX-Basic y, si se desea, procesarla a continuación a otro formato.
Palabras clave: Traducción, intercambio de terminología, XML, TBX, TBX-Basic, tabla, hoja de cálculo.

Abstract: Translators wanting to go beyond two-column glossaries and learn about terminology exchange but not see any XML will find this article useful. It describes a tabular format that represents the content of TBX-Basic files. Information in this format can be converted to TBX-Basic and further processed, if desired.
Key words: Translation, terminology exchange, terminology interchange, XML, TBX, TBX-Basic, table, spreadsheet.

1. Introduction

Terminological databases (or termbases, for short) can be designed using a wide variety of data models and thus it may be difficult to exchange information among them. A neutral, XML-based intermediate representation of terminological information is provided by the TBX (TermBase eXchange) standard, which is the basis for this article. TBX is a joint endeavor of OSCAR, a standards body that is part of LISA (see www.lisa.org), and ISO Technical Committee 37 (www.iso.org). The first version of TBX was published by LISA in 2002. A revised version was submitted to ISO in 2007 and will likely be finalized by the end of 2008. Within the TBX framework users can define a variety of formats that all conform to the same abstract data model. One of these formats, described in this article, is called TBX-Basic.

What are some practices for translation-oriented terminology? At one extreme is a very simple two-column glossary in which each row consists of only a source term and a target term. At the other extreme, one can design a very complex system to manage an extensive termbase. This article is about following a middle road: the focus is on termbases that are more than two-column glossaries but are still relatively simple. Such termbases can be represented and exchanged using TBX-Basic.

Fundamental to TBX-Basic and indeed all terminology work are the notions of data category and structural level.

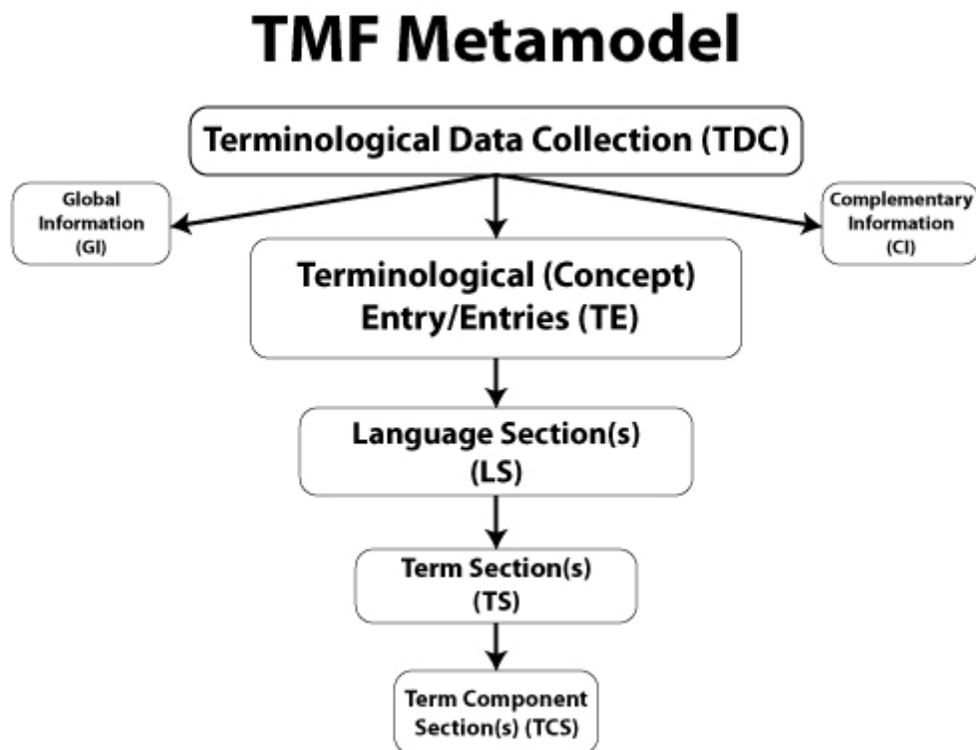
1.1. Data Categories

Each data item, whether it be a cell of a spreadsheet, a field in a database, or an element or attribute of an XML file, should contain just one piece of information. The type of information in a data item is called its data category in terminology work. Typical data categories are the name of a subject field that a concept is part of, the definition of a concept, the language of a set of terms, the part of speech of a term, and the term itself. Other data categories are administrative, such as the date of a modification to a termbase and the responsible party.

The approximately twenty data categories in TBX-Basic are taken from the larger inventory found in the default selection of data categories that is part of the TBX standard. The default data categories in TBX are in turn taken from the even larger inventory in the data category registry of ISO 12620.

1.2. Structural Levels

A well-structured termbase organizes data items according to the following structural levels dictated by the Terminological Markup Framework (TMF) standard: (1) terminological data collection (TDC), (2) terminological concept entry (TE), (3) language section (LS), and (4) term section (TS). These levels are illustrated in the following figure:



In this article, these four levels will be referred to by short forms: termbase, concept, language, and term levels. Each concept entry of a termbase describes one concept from a subject field and includes one or

more language sections. Each language section includes one or more terms -- in the language of the section, of course -- that describe the concept. Each term section includes one term and information about it, such as its part of speech. In the model above, there is an even lower term-component level, in which terms are broken down into words, morphemes, or syllables. This level is omitted from TBX-Basic.

Supporting the three main levels -- concept, language, and term -- are two other data sections: (1) information about the entire termbase is found in the Global Information section, and (2) reference material, including a list of persons and organizations cited as writers of entries or portions of entries, is found in the Complementary Information section. These sections complete the abstract, high-level structure of a termbase.

1.3. Relationship between Term Tables and TBX-Basic

TBX-Basic is one of the many formats that use the structural levels just described.

A TBX-Basic file is an XML document that conforms to the constraints of the TBX-Basic terminological markup language (TML) and contains terminological information. This article will show how to represent the same information in a table (called a term table) that can be stored and edited as a spreadsheet using spreadsheet software or as plain text using a text editor. When it is important to distinguish between the

tabular format described in this article and other tabular formats for terminological information, the specific name "MRCtermTable" can be used, where MRC stands for "Multiple Rows per Concept".

Free software that converts a term table into a TBX-Basic file will be made available (see Converter in References). From this it should not be inferred that the MRCtermTable format is presented as an alternative to commercial terminology management systems. Indeed, the author agrees with the arguments made against typical spreadsheet representations of terminological information found in Wetzel (2008). The principal purpose of the MRCtermTable format is to explain TBX-Basic in a concrete fashion. A secondary purpose is to provide a table format that allows data entry of terminological information and subsequent conversion to TBX-Basic, without working directly with XML. Once terminological information is in TBX-Basic, it can be imported into any TBX-Basic-aware terminology management system. There is no expectation that terminology management systems will import or export information in MRCtermTable format.

The central idea of an exchange format such as TBX-Basic is to provide a neutral intermediate representation. The origin of a TBX-Basic file could be any terminology management system that supports TBX-Basic import and export, a term table that has been converted to TBX-Basic, or a legacy dataset that has been converted to TBX-Basic. However, this article will focus on the MRCtermTable format.

2. Getting Started with TBX-Basic using Term Tables

Term tables separate information into the same four levels found in TBX-Basic: the termbase level (for information that applies to all concept entries in the termbase), the concept level, the language level, and the term level. Each concept is assumed to be part of some subject field. Within a particular subject field, experts can often agree on shared, language-independent and culture-independent concepts. In that case, we can define a multilingual concept entry and assume that all the terms in such an entry are synonymous, and that any one may serve as the source-language term. In other cases, a translator or terminologist will

judge that the two terms are related but do not designate the same concept, and will place them in separate monolingual concept entries and optionally cross-reference them to indicate their relationship.

Now it is time to describe the particular data categories included in TBX-Basic and explain how to put them into a term table. In order to use concrete examples, we will begin with a simple two-column glossary using terms that might be found on a bilingual menu at a restaurant.

French	English
boeuf	beef
déjeuner	lunch
fromage	cheese
moutarde	mustard
pain	bread
petits pois	green peas
pois chiches	garbanzo beans or chick peas
poulet	chicken
romarin	rosemary

The first row of a term table must consist of the keyword “=MRCtermTable” (all one word with no spaces). The second row indicates the default language for all text above the language-section level, unless overridden. For example, a term table using English as its working language would begin with:

=MRCtermTable		
A	workingLanguage	en

The third row of a term table indicates the source of the information in the term table:

A	sourceDesc	Gathered from actual menus
---	------------	----------------------------

Term tables have identifiers that indicate relationships among rows within a table. The next convention needed in order to represent this glossary in a term table is a systematic approach to creating element identifiers (IDs). TBX-Basic does not impose any particular system of creating element IDs except that they must be valid XML IDs. This means that they must begin with a letter (a-z or A-Z) and continue with letters and digits and a few punctuation marks such as hyphens, periods, and underscores. In the MRCtermTable format, we will use element IDs that reflect the level of the element. A concept entry ID will consist of the letter C followed by three digits, allowing up to a thousand concept entries in one file. Each language ID will consist of the concept ID followed by a two-letter or three-letter code taken from the international standard for language codes: ISO 639. In Part 1 of this standard (two-letter codes), "en" is used for English and "fr" is used for French. At the term level, an ID will consist of the language ID plus an integer.

Thus one minimal concept entry could be represented as follows in a term table:

C003fr1	term	poulet
C003en1	term	chicken

To indicate that all the terms (in this case, only two terms) are taken from the very narrow subject field "Restaurant Menus", simply add another row:

C003	subjectField	Restaurant Menus
C003fr1	term	poulet
C003en1	term	chicken

Each row in a term table is divided into three required columns. The first column is the ID (beginning with "C") of a term, language section, or concept, or it is a row that applies to the whole termbase ("A" for "All items"), or it is a reference ("R"). References will be explained in section 3.2. The second column states a data category, and the third column gives the value of that data category. In some cases there will be information in additional columns. Note that in term tables, data-category names have no spaces and use what is sometimes called "camel case" (e.g. in "subjectField", the "f" of "field" is upper case and there is no space between "subject" and "field").

So far we have used just two data categories: term and subjectField. The need for terms should be evident, but the need for subject fields may not be. Note that in the whimsical subjectField of "Dangerous Games that Young People Play", you would also find the term "chicken", but it would mean the sometimes fatal practice of two cars driving straight toward each other, hoping the other driver will turn away (i.e. "chicken

out") first. One would not want the two concepts for "chicken" confused. If this example seems far-fetched, consider the term "illegal operation". Without actual definitions (which are recommended for termbases but not required in TBX-Basic) the term "illegal operation" would still be understood differently in medicine (surgery performed by someone without a medical license) and computer software (a terrible malfunction at a very low level in a computer program). Homonymy like this is very common, so in terminology best practice the data category "subjectField" is strongly recommended.

The subject field should be recorded as part of each concept entry unless all the concepts in a termbase are taken from the same subject field and the name of that subject field is recorded at a higher level in a header about the entire termbase. As already seen, a row beginning with the letter "A" applies to all the concept entries of the termbase:

A	subjectField	Restaurant Menus
---	--------------	------------------

In practice, the subject field of a termbase might be implicit, but there is no such thing as a termbase for general vocabulary. In terminology work, a term always designates a concept in a particular subject field. This requirement is not imposed on general-language dictionaries.

There can be more than one term in a language section. For example, the food item called "pois chiches" in French can be called either garbanzo beans or chick peas in English:

C002fr1	term	pois chiches
C002en1	term	garbanzo beans
C002en2	term	chick peas

Just as each concept entry needs an implicit or explicit indication of subject field, each term needs an implicit or explicit indication of part of speech (noun, verb, etc.). We can add part of speech to our entry as follows:

C003	subjectField	Restaurant Menus
C003fr1	term	poulet
C003fr1	partOfSpeech	noun
C003en1	term	chicken
C003en1	partOfSpeech	noun

The above information represents one concept entry with three sections (an indication of the subject field and two term sections). The blank lines between sections in the above term table are allowed for readability but are not required.

In TBX-Basic, part of speech is usually required. Under certain conditions, it can be omitted. These conditions are given in section 4.

3. Additional Data Categories and Features of Term Tables

It is crucial to provide enough information in a concept entry so that it is clear which concept it is about. Sometimes identifying a term as a noun and indicating the subject field is sufficient. Other times it is important to describe terms using a definition. Suppose, for example, that not everyone knows that rosemary is the name of leaves of an evergreen bush in the mint family used as a seasoning. In TBX-Basic, a definition can appear at the concept level or the language level. It is often placed at the language level. The format in this article only allows a definition at the language level. The first concept placed in our sample termbase was: {chicken}, where the curly brackets indicate that we are talking about the concept of chicken (as used in restaurant menus) rather than the word "chicken". Let's add the concept {rosemary}, as just discussed, to our termbase:

C007	subjectField	Restaurant Menus
C007en	definition	Leaves...in the mint family...
C007en1	term	rosemary

These three lines mean that concept number seven in our system is part of the restaurant menu subject field and that the definition in English of this concept is "Leaves of an evergreen bush in the mint family sometimes used as a seasoning" (the definition has been shortened to avoid word wrap in the example). If a term table is stored in a spreadsheet program, it is likely possible to tell the program to wrap the text if it exceeds the width of a cell. This reduces the need for horizontal scrolling. Note that some spreadsheet software limits the length of text in a cell to 256 characters when exporting to tab-delimited format.

Often, a definition is borrowed from somewhere and not created from scratch. Suppose that this definition is taken, with permission, from a book called French Cooking for Beginners by Jean Beautodeaux, published by Flying Gourmet Press. This could be indicated as follows:

C007en	definition	Leaves...	Source: French Cooking...
--------	------------	-----------	---------------------------

The data-category "source" contains a bibliographic reference for the item at hand. Note that in this case the keyword "Source:" must appear in its own tab-separated column of the term table and be followed by a colon (spaces are allowed after the colon). This row thus has four columns (the element ID, the data category "definition", the definition itself, and the source of the definition) rather than three columns as in the previous examples.

This example shows how a source and its definition are linked in the MRCtermTable format. The other use for "source" in a term table is to indicate the source of a term. Instead of adding a fourth column to a term row, the source of a term (in this case an herb not used in cooking) is indicated by a second row as follows:

C004en1	term	comfry
C004en1	source	www.cookingindex.com/az/170/0/comfrey.htm

Some concepts are easier shown than told. We can link the concept entry to an image file containing a picture as follows:

C005	subjectField	Restaurant Menus	
C005	xGraphic	Peas in a pod	Link:peas.jpg

Here an image is used to describe the concept "peas". The text "peas in a pod" is a description of the image in the file named "peas.jpg". When only the file name is given, the file is expected to be in the same directory as the table. When the file is elsewhere, this can be indicated by using a URL (a Web address that can be found in a browser). Note that the xGraphic data category has two values: a description of the image and a link to the image file. As with the source indicator, the second value, the link, is separated by another tab, and preceded by a keyword and colon.

So far, we have given examples of two ways to describe a concept within a subject field: an image at the concept level and a definition at the language level. A concept can also be clarified by providing a piece of text that uses a term in context. For example, consider the following:

C005	subjectField	Restaurant Menus
C005en1	term	green peas
C005en1	context	Heat green peas in water not more than 3 minutes
C005fr1	term	petits pois

Another term-level data category is gender. Here the table shows that the French term for {chicken} is masculine:

C003	subjectField	Restaurant Menus
C003fr1	term	poulet
C003fr1	grammaticalGender	masculine
C003en1	term	chicken

An abbreviation is considered a term in its own right. Sometimes there is more than one term in the same language section of a concept entry. For example, suppose that the herb rosemary is abbreviated "rsy". This could be indicated as follows:

C007	subjectField	Restaurant Menus
C007en	definition	Leaves...in the mint family...

C007en1	term	rosemary
---------	------	----------

C007en2	term	rsy
C007en2	termType	abbreviation

A note can be added to the term, language, or concept level. For example:

C007en2	note	the abbreviation "rsy" is tentative
---------	------	-------------------------------------

3.1. Linking Two Rows

Sometimes you want to link together two items in a termbase. The need to link can arise in a case as simple as this one:

C015	subjectField	Restaurant Menus
C015en1	term	mustard
C015fr1	term	moutarde

Some might object that the default meaning of the term "mustard" in English and the term "moutarde" in French are different. The default mustard in America is a bright yellow concoction, while the default type of mustard in France is what Americans call "Dijon mustard". A terminologist must often decide whether two terms designate the same concept. This introduction to TBX-Basic will not attempt to instruct terminologists how to make such judgments. It will merely show how to link two monolingual concept entries (here, concepts 18 and 21, instead of only concept 15) if this is desired:

C018	crossReference	See "French" mustard	Link: C021
C018en1	term	mustard	

C021	crossReference	See "American" mustard	Link: C018
------	----------------	------------------------	------------

C021fr1	term	moutarde	
---------	------	----------	--

A link to an external resource (i.e. something outside the term table) may be made with the external cross reference data category:

C018en1	term	mustard	
C018en1	externalCrossReference	Discussion of mustard	Link: www.mustard.com

3.2. Transactions

So far, we have provided no method of indicating who is responsible for a given concept entry or some part of it. This is done in TBX-Basic using transactions. A transaction row in the table format describes the origination or modification of the concept entry, language section, or term whose ID it bears. For example, you can indicate who created a concept entry as follows (using a row with five columns):

C018	transactionType	origination	Responsibility: Jerry	Link: R005
------	-----------------	-------------	-----------------------	------------

This row introduces a new kind of "Link:", a link to a Reference in the Complementary Information of a TBX-Basic termbase. A responsible party row is identified by "R" followed by three digits. It takes several of these rows to identify a person. For example:

R005	type	person
R005	fn	Jerry Springerband
R005	title	Head chef at the Blue marlin restaurant in New York
R005	email	Jerry@example.com

Responsible party rows can describe an organization or a person, as determined by the value in the first row in the set (with data category "type"). The other data categories are adapted from the vCard standard (<http://www.imc.org/pdi/vcard-21.txt>). The main ones are:

- fn full name (of a person)
- org name (of an organization)
- title (e.g. Dr.)
- role (in the terminology management process)
- email
- uid (user ID)
- tel (telephone number)
- adr (postal address)

In addition to responsibility, we can also indicate the date of creation or modification. This is done with ISO date format (using hyphens to separate year, month, and day) in an additional field as follows:

C018	transactionType	modification	Date:2003-01-25
------	-----------------	--------------	-----------------

This date is January 25, 2003. A transaction can be annotated with date, responsibility, or both. The additional columns beyond the first three can appear in any order.

3.3. Customer and Project Subsets

Sometimes it is important to associate a term with a particular customer or project. Suppose we are part of a project to make it easier to understand the ingredients in the items on a restaurant menu. Further suppose the project is called "Know What You Order" with the project acronym "KWYO". We could indicate that the concept entry for {rosemary} is part of this project as follows:

C007	subjectField	Restaurant Menus
C007en	definition	Leaves...in the mint family
C007en1	term	rosemary
C007en1	projectSubset	KWYO

Indicating that a term is specific to a particular customer, rather than a particular project, would involve using the data category "customerSubset".

3.4. Inventory of Row Types in Term Tables

Each row of a term table begins with an A, C, or R and uses one of the following ID patterns:

- header: A (information that applies to All entries in a termbase)
- concept: C999 (information about a particular Concept, such as its subject field)
- language: C999nn (information about a language section in a particular concept entry)
- term: C999nn9 (a term, or information about it)
- reference: R999 (information about a Responsible party)

The row ID is followed by a data category and its value – and sometimes extra information. Non-blank rows should be sorted in ascending alphanumeric order according to this ID.

4. Required Data Categories

TBX-Basic does not impose heavy requirements on a user. The only data category that is always required is "term". If the termbase is intended to be machine processable, the data category partOfSpeech is also required. If machine processability is not a requirement, partOfSpeech is still recommended, and it must be present if there is neither a definition for the language section nor a contextual example for the term. An indication of subject field is strongly recommended, either at the concept level of each entry or in the header of the termbase. Other data categories need only be present as the function of the termbase and available information dictate.

5. Conclusion

This article has described a simple data format called a term table (short for a table in MRCtermTable format). Term tables communicate the information of the TBX-Basic format without using any XML. So long as a term table follows the layout rules in this article, it can be automatically converted to an XML file that conforms to the constraints of TBX-Basic. Software to perform this conversion, along with further details about the term table and TBX-Basic formats, will be made available free at the LISA website and the author's personal website (see note at beginning of References below for the URLs). However, this article and the further information at the LISA website are intended to be useful as a description of TBX-Basic and an introduction to TBX, even if the reader is not planning to create term tables.

The appendix to this article contains all the data categories allowed in a term table (and in TBX-Basic) in one list, intended to be used as a reference when building term tables and for implementing TBX-Basic in a terminology management system.

It is hoped that this article will promote the use of TBX-Basic and terminology management software that uses it, for interchange and integration with other components of a translation environment. The information in this article should also facilitate communication between users of translation-oriented terminology and developers of terminology management systems. Some systems will also support more comprehensive TBX formats, but it is hoped that translators and terminologists will insist that at least TBX-Basic import and export functions be implemented in every terminology management system.

References

Converter For further information about TBX-Basic and the MRCtermTable-to-TBX converter, see:

- LISA website: www.lisa.org (see TBX under Standards) or

- Author's personal website: www.ttt.org/tbx

The converter is free and open source software. The initial version is written in Perl. Files in the MRCtermTable format can be created in either a text editor or a spreadsheet and then exported to a tab-delimited text file. The Perl converter requires that input files use the UTF-8 encoding of Unicode.

ISO 639 (2002 and later) Codes for the representation of names of languages. [note: Part 1 provides two-letter codes; parts 2 and 3 provide three-letter codes]

ISO 8601 (2004) "date format" http://www.iso.org/iso/date_and_time_format (accessed June 30, 2008)

ISO 12620 (1999) Computer applications in terminology – Data categories [note: a new version of ISO 12620 is under development.]

ISO 16642 (2003) Computer applications in terminology – TMF (Terminological Markup Framework)

ISO 30042 (in press) TBX [note: an updated version of the TBX standard is available from the LISA website (<http://www.lisa.org/Term-Base-eXchange.32.0.html>)]

vCard (1996) "The electronic business card" (version 2.1) <http://www.imc.org/pdi/vcard-21.txt> (accessed June 30, 2008)

Wetzel, Michael (2008) "Structured Termbases – Why Spreadsheets Soon Fail" (in the presentation section of <http://www.iim.fh-koeln.de/AT08/> -- page consulted June 30, 2008; click the "Präsentationen" link and look for the article by Wetzel, which is in English even though the title is in German).

Appendix 1: TBX-Basic Data Categories used in MRCtermTable format

A Level: (these rows apply to all concept entries in the termbase)

sourceDesc: description of where this termbase came from (required)

workingLanguage: default language for text until overridden (required)

subjectField: implicit subject field for entire termbase (optional)

C Level: (a note row is allowed at the concept level)

subjectField: code for domain of concept (strongly recommended unless at "A" level)

optional: xGraphic: link name + link outside termbase (URL)

L Level: (a note row is allowed at the language level; a source is allowed on a definition row)

definition: suggested (can be omitted if partOfSpeech or context is included)

T Level: (these data categories follow a term row; a source row and a note row are allowed)

partOfSpeech: picklist A (required for machine processing)

optional:

administrativeStatus: picklist B

context: segment of text containing term (suggested)

geographicalUsage: region where term is used

grammaticalGender: picklist C

termLocation: where the term usually occurs (if software, see suggested values below)

termType: picklist D

customerSubset: code to identify a customer (often to indicate which synonym is preferred)

projectSubset: code to identify project

Multi-Level (allowed at C, L, and T levels unless otherwise indicated):

optional:

crossReference: link name + link within termbase (ID) (allowed at C and T levels only)

externalCrossReference: link name + link outside termbase (URL) (allowed at C and T levels only)

Transactions:

transactionType: picklist E (a transaction can include a responsibility or date or both)

Responsibility: name and contact info for responsible party

Date: date in ISO format (yyyy-mm-dd) transaction occurred

Picklists

picklist A: noun verb adjective adverb properNoun other

picklist B: preferredTerm-admn-sts admittedTerm-admn-sts deprecatedTerm-admn-sts supersededTerm-admn-sts

picklist C: masculine feminine neuter other

picklist D: fullForm acronym abbreviation shortForm variant phrase

picklist E: origination modification

suggested values for termLocation in software localization: menuItem dialogBox groupBox textBox comboBox comboBoxElement checkBox tab pushButton radioButton spinBox progressBar slider informativeMessage interactiveMessage tooltip tableText userDefinedType

Appendix 2: Sample term table

The following term table is in MRCtermTable format. It consists of header information and two concept entries. The first row of a term table is always the keyword that indicates it is an MRC-type terminology table. Then the two A rows indicate the working language of the table and where the information came from.

In this example, there are then two concept entries. The first one is minimal. It simply states that the English term chicken and the French term poulet designate the same concept. The second one (treating the

concept {garbanzo bean}) has additional information about the concept than just the terms that designate it and their part of speech. Blank lines are optional. Items shown in the fourth column must actually be multiple columns separated by tab characters. Indeed, pairs of items must always be separated by tabs.

Many of the data categories allowed in TBX-Basic are included in this example. One that is not included is termType. It can be used to indicate that a term is an abbreviation. For example, the term “MSG” could be given the termType acronym, and the term “monosodium glutamate” can be given the termType “fullForm” in the same English language section of a concept entry.

=MRCtermTable			
A	workingLanguage	en	
A	sourceDesc	a restaurant menu in English and French	
C003	subjectField	Restaurant Menus	
C003fr1	term	poulet	
C003fr1	partOfSpeech	noun	
C003fr1	grammaticalGender	masculine	
C003en1	term	chicken	
C003en1	partOfSpeech	noun	
C005	subjectField	Restaurant Menus	
C005	transactionType	origination	Responsibility: Jill Link: R007 Date: 2007-01-31

C005	xGraphic	garbanzo beans	Link: http://flickr.com/photos/lilgreen/432468210/
C005en	definition	an edible legume of the family Fabaceae, subfamily Faboideae	Source: http://en.wikipedia.org/wiki/Chickpea
C005en1	term	chick peas	
C005en1	partOfSpeech	noun	
C005en2	term	garbanzo beans	
C005en2	partOfSpeech	noun	
C005en2	geographicalUsage	southwest United States	
C005en2	customerSubset	AlmostRipe Foods	
C005fr1	term	pois chiches	
C005fr1	partOfSpeech	noun	
R007	type	person	
R007	fn	Jill Johnson	
R007	email	jill@example.com	
R007	title	bean expert	