



# IMS Vocabulary Definition Exchange Information Model

## Version 1.0 Final Specification

### IPR and Distribution Notices

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the specification set forth in this document, and to provide supporting documentation.

IMS takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Information on IMS's procedures with respect to rights in IMS specifications can be found at the IMS Intellectual Property Rights web page: [http://www.imsglobal.org/ipr/imsipr\\_policyFinal.pdf](http://www.imsglobal.org/ipr/imsipr_policyFinal.pdf).

Copyright © 2004 IMS Global Learning Consortium. All Rights Reserved.

If you wish to copy or distribute this document, you must complete a valid Registered User license registration with IMS and receive an email from IMS granting the license to distribute the specification. To register, follow the instructions on the IMS website: <http://www.imsglobal.org/specificationdownload.cfm>.

This document may be copied and furnished to others by Registered Users who have registered on the IMS website provided that the above copyright notice and this paragraph are included on all such copies. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to IMS, except as needed for the purpose of developing IMS specifications, under the auspices of a chartered IMS project group.

Use of this specification to develop products or services is governed by the license with IMS found on the IMS website: <http://www.imsglobal.org/license.html>.

The limited permissions granted above are perpetual and will not be revoked by IMS or its successors or assigns.

THIS SPECIFICATION IS BEING OFFERED WITHOUT ANY WARRANTY WHATSOEVER, AND IN PARTICULAR, ANY WARRANTY OF NONINFRINGEMENT IS EXPRESSLY DISCLAIMED. ANY USE OF THIS SPECIFICATION SHALL BE MADE ENTIRELY AT THE IMPLEMENTER'S OWN RISK, AND NEITHER THE CONSORTIUM, NOR ANY OF ITS MEMBERS OR SUBMITTERS, SHALL HAVE ANY LIABILITY WHATSOEVER TO ANY IMPLEMENTER OR THIRD PARTY FOR ANY DAMAGES OF ANY NATURE WHATSOEVER, DIRECTLY OR INDIRECTLY, ARISING FROM THE USE OF THIS SPECIFICATION.

**Copyright © 2004 by IMS Global Learning Consortium, Inc.**  
**All Rights Reserved.**

The IMS Logo is a trademark of IMS Global Learning Consortium, Inc.  
Document Name: IMS Vocabulary Definition Exchange Information Model  
Date: 23 February 2004

# Table of Contents

<b>1.</b>	<b>INTRODUCTION .....</b>	<b>3</b>
1.1	OVERVIEW AND RELATIONSHIP TO OTHER SPECIFICATIONS.....	3
1.2	IMS VOCABULARY DEFINITION EXCHANGE SPECIFICATION COMPONENTS.....	3
1.2.1	<i>Information Model (this document).....</i>	<i>3</i>
1.2.2	<i>XML Binding .....</i>	<i>4</i>
1.2.3	<i>Best Practice and Implementation Guide.....</i>	<i>4</i>
1.2.4	<i>Conformance Requirements .....</i>	<i>4</i>
1.3	NOMENCLATURE.....	4
1.3.1	<i>Abbreviated Terms .....</i>	<i>4</i>
1.3.2	<i>Definitions .....</i>	<i>4</i>
1.4	REFERENCES.....	6
<b>2.</b>	<b>INFORMATION MODEL .....</b>	<b>7</b>
2.1	INFORMATION ITEMS PERTAINING TO THE VOCABULARY AS A WHOLE.....	7
2.2	INFORMATION ITEMS PERTAINING TO TERMS IN THE VOCABULARY .....	8
<b>3.</b>	<b>VDEX ELEMENTS.....</b>	<b>10</b>
3.1	LANGSTRING ELEMENTS .....	12
3.2	OBJECT MODEL .....	12
<b>4.</b>	<b>SEMANTICS AND VALUE SPACE FOR “PROFILE TYPE” PROPERTY OF VDEX VOCABULARIES .....</b>	<b>14</b>
4.1	RATIONALE AND INTRODUCTION .....	14
4.2	DEFINITION OF THE PROFILE TYPES .....	14
<b>5.</b>	<b>PREFERRED VALUE DOMAINS FOR EXPRESSING THESAURUS RELATIONSHIPS .....</b>	<b>17</b>
	<b>ABOUT THIS DOCUMENT .....</b>	<b>18</b>
	LIST OF CONTRIBUTORS .....	18
	<b>REVISION HISTORY .....</b>	<b>19</b>
	<b>INDEX.....</b>	<b>20</b>

# 1. Introduction

This section is not normative.

## 1.1 Overview and Relationship to Other Specifications

The IMS Vocabulary Definition Exchange (VDEX) specification defines a grammar for the expression of value lists of various classes: collections often denoted “vocabulary”. Section 1.3.2 provides some definitions of the classes of vocabulary that VDEX can be used for. The ISO prefers the terms “value domain” and “permitted value” for the data VDEX principally relates to.

VDEX **does not** seek to permit the expression of all possible vocabularies: rather to permit the light-weight expression of relatively simple value lists. Furthermore, VDEX is not a modelling language for vocabularies; there may be some properties of a term of interest to the vocabulary creator, an expert in a domain, but that are not required by the end user.

VDEX **does** permit the expression of simple machine-readable lists of human language terms together with information that may aid a human being in understanding the meaning or applicability of the various terms. Collections of such human language terms are often collected in the form of thesauri; VDEX enables thesauri to be expressed using arbitrary relationship types, inclusive of those specified in ISO 2788 and ISO 5964 standards.

VDEX may be also used to express valid data for use in instances of, for example: IEEE LOM, IMS Meta-Data, IMS Learner Information Package, ADL SCORM. In these specifications, the elements using value lists are often described as being of “vocabulary data type” and expressed through separate domain and value identifiers. The permitted values are formally codes to identify a concept and they may not be always be textually identical to human language terms

Whereas vocabulary data types typically have an associated unstructured list of terms, meta-data schemas often permit the expression of subject classifications from hierarchical vocabularies (taxonomies), for example Dewey Decimal (DDC) and specialist schemes such as the Medical Subject Headings (MeSH). VDEX can express strictly hierarchical schemes in a compact manner and allows for more loose networks of relationship to be expressed if required.

Polyhierarchical schemes and faceted schemes, such as Universal Decimal Classification (UDC), are a complex area. VDEX provides support for several use cases through the Information Model (this document) and makes some recommendations in the Best Practices and Implementation Guide. Experts in these fields with specialist use cases may need to apply techniques such as:

- defining a custom set of relationship types and using the VDEX model;
- using VDEX along side certain working practices;
- extending the VDEX model;
- embedding VDEX within a more elaborate model.

## 1.2 IMS Vocabulary Definition Exchange Specification Components

The VDEX documents permit simple collections of terms to be expressed and optionally exchanged. The specific types of terms are defined in section 1.3.2. This document is the IMS Vocabulary Definition Exchange Specification version 1.0. As such it forms one of a set that comprise the specification, each with distinct scope:

### 1.2.1 Information Model (this document)

Describes the core aspects of the specification and contains parts that are normative for any binding claiming to use this information model. It contains details of: semantics, structure, data types, value spaces, multiplicity, and obligation (i.e., whether mandatory or optional).

### 1.2.2 XML Binding

Describes a binding of the Information Model to XML version 1.0 and is normative for any XML instance that claims to use this binding, whether by reference to the specification or by declaration of the namespace reserved by the specification. In cases of error or omission, the Information Model takes precedence. The VDEX XML Binding is released with a control document using W3C XML Schema 1.0 that should be used in implementations.

### 1.2.3 Best Practice and Implementation Guide

Provides non-normative guidance on application of the Information Model and XML Binding. This includes reference to existing practice in handling information that this specification seeks to support and guidance on practices that will promote interoperability and durability. It also includes examples to illustrate how the conceptual framework maps to practical uses and to identify the relationship between this specification and related IMS specifications. Implementers are encouraged, but not required, to follow guidance in this part of the specification.

### 1.2.4 Conformance Requirements

Provides a set of testable statements that may be used in relation to applications of the information model and binding. These statements may form the basis of formal conformance testing and certification or informal assertions. This document makes no statement about the formal processes or method of certification or assertion; it only deals with criteria.

## 1.3 Nomenclature

### 1.3.1 Abbreviated Terms

ADL	Advanced Distributed Learning
DDC	Dewey Decimal Classification Scheme
IEEE	Institute of Electrical and Electronics Engineers Inc.
IETF	Internet Engineering Task Force
ISO	International Organization for Standardisation
LOM	Learning Object Metadata (usually used in “IEEE LOM”)
LTSC	IEEE Learning Technology Standards Committee
RFC	Request for Comment (usually used in “IETF RFC xxxx”)
SCORM	Shareable Content Object Reference Model
UDC	Universal Decimal Classification Scheme
VDEX	IMS Vocabulary Definition Exchange Specification (this specification)
W3C	World Wide Web Consortium
XML	Extensible Mark-up Language

### 1.3.2 Definitions

We adopt some widely accepted definitions for “vocabulary” and related terms. A useful collection of definitions is given in [CVLOM], drawing upon well-known authorities, and contains the following summary of various authorities:

**Vocabulary:**

*“dictionary of terms related to a particular subject field.” [ISO 5127/2-1983]<sup>1</sup>*

*“1.a A collection or list of words with brief explanations of their meanings; now esp. a list of this kind given in an elementary grammar or reading book of a foreign language. 2.a. The range of language of a particular person, class, profession, or the like. 3. The sum or aggregate of words composing a language.” [OED]<sup>2</sup>*

**Classification:**

*“arrangement of concepts into classes (3) and their subdivisions to express the semantic relations between them; the classes are represented by means of a notation”. [ISO 5127/6-1983]<sup>3</sup>*

*“1) The action of classifying or arranging in classes, according to common characteristics or affinities; assignment to the proper class. 2. The result of classifying; a systematic distribution, allocation, or arrangement, in the class or classes; esp. of things which form the subject-matter of a science or of a methodic enquiry.” [OED]<sup>4</sup>*

**Taxonomy:**

*“1. Classification, esp. in relation to its general laws or principals; that department of science, or of a particular science or subject, which consists in or relates to classification; esp. the systematic classification of living things. 2. A classification of anything.” [OED]<sup>5</sup>*

**Thesaurus:**

*“The vocabulary of a controlled indexing language, formally organized so that the a priori relationships between concepts (for example as broader and narrower) are made explicit.” [ISO 2788:1986; and ISO 5964:1985]<sup>6</sup>*

*“2.b. A collection of concepts or words arranged according to sense; also (U.S.) a dictionary of synonyms and antonyms. 2.c. A classified list of terms, esp. key-words, in a particular field, for use in indexing and information retrieval.”<sup>7</sup>*

**Glossary:**

*“dictionary of ancient, rare, little-known or new words in a given language.” [ISO 5127/2-1983]<sup>8</sup>*

*“A collection of glosses; a list with explanations of abstruse, antiquated, dialectical or technical terms; a partial dictionary.”<sup>9</sup>*

**Dictionary:**

*collection of words or a category of words from a language and explained or translated in one or more languages, arranged alphabetically or systematically. [ISO 5127/2-1983]<sup>10</sup>*

*“1.a. A book dealing with the individual words of a language (or certain specified classes of them), so as to set forth their orthography, pronunciation, signification, and use, their synonyms, derivation, and history, or at least some of these facts: for convenience of reference, the words are arranged in some stated order, now, in most languages, alphabetical; and in larger dictionaries the information given is illustrated by quotations from literature; a word-book, vocabulary of lexicon. 1.b. The vocabulary or whole list of words used or admitted by anyone. 2.a. By extension: A book of information or reference on any subject or branch of alphabetical knowledge, the items of which are arranged in alphabetical order; as an alphabetical encyclopaedia.” [OED]<sup>11</sup>*

1. ISO 5127/2-1983: Documentation and information vocabulary part 2 (traditional documents).

2. The Oxford English Dictionary. Second Edition, Oxford: Clarendon Press, 1989.

3. ISO 5127/6-1983: Documentation and information: vocabulary; part 6 (documentary languages)

4. The Oxford English Dictionary. Second Edition, Oxford: Clarendon Press, 1989.

5. The Oxford English Dictionary. Second Edition, Oxford: Clarendon Press, 1989.

6. ISO 2788:1986, Guidelines for the establishment and development of monolingual thesauri. ISO 5964:1985, Guidelines for the establishment and development of multilingual thesauri.

7. The Oxford English Dictionary. Second Edition, Oxford: Clarendon Press, 1989.

8. ISO 5127/2-1983: Documentation and information: vocabulary; part 2 (traditional documents)

9. The Oxford English Dictionary. Second Edition, Oxford: Clarendon Press, 1989.

10. ISO 5127/2-1983: Documentation and information: vocabulary; part 2 (traditional documents).

11. The Oxford English Dictionary. Second Edition, Oxford: Clarendon Press, 1989.

## 1.4 References

- [CVLOM] “Controlled Vocabularies for Learning Object Metadata. – Typology, impact analysis, guidelines and a web based Vocabularies Registry”. F. Van Assche, L. Anido-Rifón, L. M. Campbell and M. Willem, CEN/ISSS Learning Technologies Workshop, DRAFT, <http://office.eun.org/kms/sites/cenisss/index.html>
- [DDC] Dewey Decimal Classification Scheme, OCLC Forest Press, <http://www.oclc.org/dewey/>
- [IEEE LOM] IEEE 1484.12.1-2002, Standard for Learning Object Metadata, <http://www.ieee.org>
- [IMSBUND] *Using IMS Content Packaging to Package Instances of LIP and Other IMS Specifications v1.0*, B.Olivier, M.McKell, *IMS Global Learning Consortium, Inc.*, August 2001 ([http://www.imsglobal.org/implementationhandbook/imspack\\_handv1p0.html](http://www.imsglobal.org/implementationhandbook/imspack_handv1p0.html)).
- [IMSPLID] *IMS Persistent, Location-Independent, Resource Identifier Implementation Handbook v1.0*, M.McKell, *IMS Global Learning Consortium, Inc.*, April 2001 ([http://www.imsglobal.org/implementationhandbook/imsrid\\_handv1p0.html](http://www.imsglobal.org/implementationhandbook/imsrid_handv1p0.html)).
- ISO 2788 ISO 2788:1986, Guidelines for the establishment and development of monolingual thesauri.
- ISO 5964 ISO 5964:1985, Guidelines for the establishment and development of multilingual thesauri.
- ISO 11404 ISO 11404, Language-independent Datatypes, <http://www.iso.ch/cate/d19346.html>
- [MESH] Medical Subject Headings, US National Library of Medicine, <http://www.nlm.nih.gov/mesh/meshhome.html>
- [RFC 1766] Tags for the Identification of Languages, <http://www.ietf.org/rfc/rfc1766.txt>
- [RFC 2119] IETF RFC 2119 – Key words for use in RFCs to Indicate Requirement Levels
- [RFC 2396] IETF RFC 2396 Uniform Resource Identifiers (URI): Generic Syntax
- [RFC3066] RFC 3066, Tags for the Identification of Languages, <http://www.ietf.org/rfc/rfc3066.txt>
- [UDC] Universal Decimal Classification Scheme, UDC Consortium, <http://www.udcc.org/>
- [Z39.19] ANSI/NISO Z39.19-1993, Guidelines for the Construction, Format, and Management of Monolingual Thesauri, ISSN 1041-5653

## 2. Information Model

This section is normative. The terms defined in RFC 2119 are used with the meaning there defined.

The overall structure of the information model is shown in the following diagram and explained in the following subsections. The term structure may be repeated and nested. It should be noted that the ordering of elements at any level in the *information model* diagram and in the textual description of *information model* elements is not significant unless a statement is made to the contrary. NB: some *bindings* of the information model may require particular ordering.

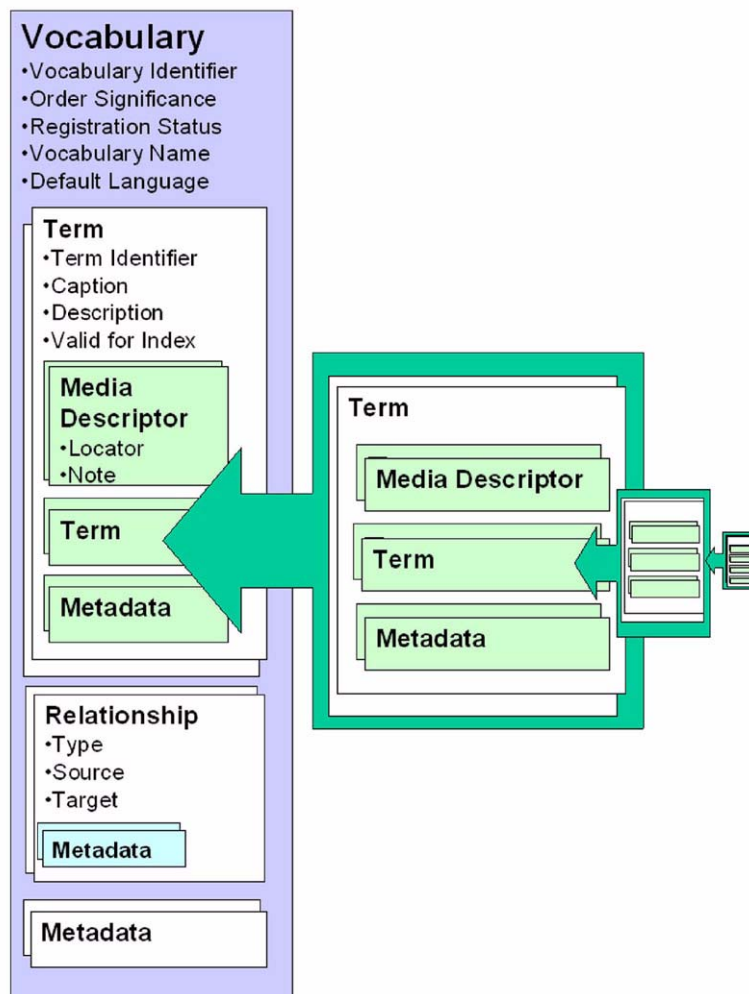


Figure 2.1 - Schematic of the Information Model.

### 2.1 Information Items Pertaining to the Vocabulary as a Whole

This section relates to properties of the collection of terms, the “value domain”.

- 1) **Identifier:** A universally unique identifier controlled by registration authority or an unregistered identifier that the creator warrants to be persistent<sup>12</sup> and takes reasonable care to make specific enough to avoid conflicts. The vocabulary identifier is formally a string but for exchange it is constrained to use the characters and encodings permitted by RFC 2396, as amended by RFC 2732, such that it is syntactically a URI. It must not contain a URI fragment reference. The VDEX Best Practices and Implementation Guide considers some application issues.

12. I.e., the creator will not use the same identifier for another purpose and will not discard in favour of an alternative identifier.



- 2) **Registration Status:** A simple flag to declare whether the identifier is registered or not. Meta-data could be used to identify the registration authority etc. Vocabularies are considered unregistered unless explicitly stated as being registered.
- 3) **Name:** Natural language name for the vocabulary. This can be expressed in any number of human languages. The name of a vocabulary should not be used as a surrogate for an identifier.
- 4) **Profile Type:** This property may be used to make explicit the class of vocabulary that an instance represents. Section 3.2 provides more information on the use of profile type. If no profile type is explicitly stated then it is assumed to be “lax”.
- 5) **Order Significance:** This property is used to state whether the ordering of terms has any significance. In cases where there is a hierarchy of terms, it only applies to the ordering of top level terms. The default information is that order is not significant. It may not always be meaningful for order to be significant.
- 6) **Default Language:** This optional property may be used to declare a default language for all elements that represent human language strings. This may be overridden at the level of the individual language string.
- 7) **Meta-data:** Optional meta-data records that further describe the vocabulary. This is for convenience in cases where it is desirable to closely associate meta-data with the data it describes. More than one meta-data element may be used but each should only contain one meta-data record.

## 2.2 Information Items Pertaining to Terms in the Vocabulary

A vocabulary must consist of a non-empty set of permitted values. These are modeled as **terms** and **relationships** with properties as below.

- 1) **Identifier:** An identification string that must be unique among terms in the vocabulary. This may be used as the referent for relations defined from other terms (see “relationship” below).
- 2) **Valid Index:** An indication of whether or not this term is valid for use by indexing applications.
- 3) **Caption:** Natural language caption(s) that apply to the term.
- 4) **Description:** Textual notes for the use of the term or description of the term. These may be presented in any number of human languages, including languages not used for the caption element.
- 5) **Media Descriptor:** Media file (or files) to provide additional clarification of the description. This may also be used as a substitute for the textual description of a term.
- 6) **Media Locator:** A URL to locate the media file.
- 7) **Interpretation Note:** Textual information to explain how the media relates to the term. This note may appear in multiple human languages.
- 8) **Order Significance:** This property applies to *child* terms of the term on which it is declared. In the absence of a statement of order significance on a term then the order or children is not deemed significant. Note that there is no “inheritance” of this property down a hierarchy of terms.
- 9) **Term:** Terms may be arranged hierarchically in the information model – i.e., without the use of the relationship element. Such arrangements should be used for taxonomies where the only relation types are parent-child relations. They should not be used to express narrower/broader term relations in a thesaurus; the relationship element should be used for this purpose.
- 10) **Meta-data:** A container is available at the level of the vocabulary term as it is for the vocabulary as a whole.

### Relationship:

- 1) **Identifier:** An optional identification string that (if it exists) must be unique among terms in the vocabulary.
- 2) **Source Term:** The identifier of the term that is the source of the relationship. The term identified by this property must exist but may be in another vocabulary. See the Binding document for more information.
- 3) **Source Vocabulary:** The identifier of the vocabulary containing the source term. If this information is not present then the {source|target} vocabulary is identical to the vocabulary containing the relationship.
- 4) **Target Term:** The identifier of the related term. The term identified by this property must exist but may be in another vocabulary. See the Binding document for more information.



- 5) **Target Vocabulary:** The identifier of the vocabulary containing the related term. If this information is not present then the {source|target} vocabulary is identical to the vocabulary containing the relationship.
- 6) **Relationship Type:** Relationship types are described using a sourced vocabulary. VDEX does not mandate any vocabulary for expressing relationship types but does make a recommendation in section 5 if a monolingual or multilingual thesaurus is being expressed.
- 7) **Meta-data:** A container is available at the level of the relationship as it is for the vocabulary as a whole.

### 3. VDEX Elements

This section is normative. The terms defined in RFC 2119 are used as defined therein.

The names used for elements in the VDEX information model are not normative names for any binding; named elements in a binding must only identify which of the elements described in this document they relate to.

Most of the VDEX information elements are optional to reflect the differing classes of vocabulary that VDEX can be used to express. Section 3.2 describes how the main classes handled by VDEX apply further constraints on the general grammar described in the following table.

**Table 3.1 - Information Elements Comprising a Vocabulary.**

	Name	Description of Use	Value Space	Data Type	SPM	Mult <sup>a</sup>
1	Order Significant	Whether the ordering of terms is significant. Default value is false.		Boolean	-	0..1
2	Profile Type	Describes the type of vocabulary being expressed. Can be used to support restrictive validation. Default value is "lax".	See section 3.2	String (enum)	-	1
3	Vocabulary Name	Human language name for vocabulary. May be expressed in multiple different languages.	Repertoire of ISO/IEC 10646-1:2000	Langstring See Note *	1000 chrs	0..1
4	Vocabulary Identifier	Identifier for the vocabulary as a whole. This identifies the value domain. This element is optional to support cases where no identifier is known. If there is a known identifier then it should be provided.	See notes	URI	4096 chrs <sup>b</sup>	0..1
4.1	Registration Status	Whether the vocabulary and its identifier is registered with some authority. Default value is false.		Boolean	-	0..1
5	Default Language	The default language for all elements of type "langstring" that do not independently declare a language.	Language Code (see section 3.1)	String	see 3.1	0..1
6	{Term}	See Table 3.2		Container		1..*
7	{ Relationship }	See Table 3.3		Container		0..*
8	Metadata	Meta-data for the vocabulary as a whole. This is a container for arbitrary meta-data.		Container	10 reps	0..*

a. In UML notation, summarized below the table

b. To support URN identifiers.

**Table 3.2 - Information Elements Comprising a Term.**

	Name	Description of Use	Value Space	Data Type	SPM	Mult
6	Term	A vocabulary item.		Container		1..*
6.1	Identifier	An identifier for the value within the scope of the vocabulary.	See notes	URI	100 chrs	1
6.2	Valid Index	A flag indicating whether or not this is a valid term for use in indexing. Default value is true.		Boolean		0..1
6.3	Caption	A human language caption.		Langstring See Note *	1000 chrs	0..1

	Name	Description of Use	Value Space	Data Type	SPM	Mult
6.4	Description	A human language description of the meaning of the vocabulary term and its usage.		Langstring See Note *	2000 chrs	0..1
6.5	Media Descriptor	URLs of an image (static or dynamic), non-verbal audio, etc. that provide a non-textual description. Useful in clarifying the precise meaning.		Container		0..*
6.5.1	Media Location	The location of a media file.	See notes	URI	4096 chrs	1 <sup>a</sup>
6.5.2	Interpretation Note			Langstring See Note	2000 chrs	0..1
6.6	Order Significant	Whether the ordering of child terms is significant. Default value is false.		Boolean	-	0..1
6.7	Metadata	Meta-data for the term. This is a container for arbitrary meta-data.		Container		0..*
6.8	{Term}	Nested Term elements for use in expressing taxonomies.		Container		0..*

a. A mandatory child of an optional element should be taken to mean that the child must appear *if* the optional element is used.

**Table 3.3 - Information Elements Comprising a Relationship.**

	Name	Description of Use	Value Space	Data Type	SPM	Mult
7.1	Identifier	An identifier for the value within the scope of the vocabulary.	See notes	URI	100 chrs	0..1
7.2	{Source Term}	The term from which this relationship originates. See table 3.4		Container		1
7.3	{Target Term}	The term which this relationship targets. The term from which the relationship originates is related TO the target term. See table 3.4		Container		1
7.4	Relationship Type	Identifies the type of relationship. Vocabularies of relationship types can be sourced as required by an application.		Container	-	0..1
7.4.1	Source of type	Identifier of the source of the relationship type vocabulary	See notes	URI	4096 chrs	1
7.4.2	Value of type	The term for the relationship type	See notes	URI	100 chrs	1
7.5	Metadata	Meta-data for the relationship. This is a container for arbitrary meta-data.		Container		0..*

**Table 3.4 - Information Elements Comprising a Term Reference.**

	Name	Description of Use	Value Space	Data Type	SPM	Mult
1	Term Identifier	The identifier of the term.	See notes	URI	1000 chrs	1
2	Vocabulary Identifier	The identifier of the vocabulary containing the term. If this is not present, the term is in the same vocabulary as the relationship.	See notes	URI	1000 chrs	0..1

**Types:**

Langstring - defined in this spec

URI - conformant to RFC2396 as amended by RFC 2732 in the value space according to ISO/IEC 10646-1:2000

String - Value space according to ISO/IEC 10646-1:2000

Boolean - as defined by ISO 11404

Language Code - a valid language code as defined by RFC 1766

**UML multiplicity notation:**

0..1 - optional and not repeatable

0..\* - optional and repeatable

1..\* - required at least once and repeatable

1 - required once only

**Note on Smallest Permitted Maximum (SPM):**

This is expressed in the format [m chrs].

“m chrs” denotes the minimum number of unicode *characters* that implementations must support within an element with text content. For the cases where an element is of type “langstring”, the SPM requirement applies to the number of characters in each langstring.

**Note on repeatable elements:**

Elements marked “Multiple” may be repeated.

### 3.1 Langstring Elements

Elements with type “langstring” and multiplicity “single” may only occur once but may contain multiple langstring constructs, each with an equivalent content. The SPM columns in Table 3.1 and Table 3.2 specify the individual constraints, which should be read in the context of the accompanying note on SPM notation and interpretation.

**Table 3.5 - Information Elements Comprising Langstring.**

	Name	Description of Use	Value Space	Data Type	SPM	Mult
A	Language string	A String in a human language	Repertoire of ISO/IEC 10646-1:2000	String	Defined by the context of use	-
A.1	Language	The human language of the string	Language Code	String	-	0..1

If an individual langstring element does not declare a language property then the Default Language of the instance must be taken to be the language of the langstring. If there is no Default Language then the language of the string is undefined.

Implementers must take care that:

- Two langstring constructs inside the same parent element must not have the same language property, including an undefined value;
- Undefined languages should be avoided;
- Users are discouraged from placing non-equivalent langstring constructs inside the same container.

The value space of the language property and the instance default language property is declared as “Language Code”. This must conform to the requirements of RFC1766<sup>13</sup>.

### 3.2 Object Model

This sub-section is not normative.

13. The Best Practice and Implementation Guide explains this requirement.

Figure 3.1 shows a class diagram following UML notation that may help some readers in interpreting the information model descriptions elsewhere in section. Types are not provided except where this specification defines a type: for vocabulary data types and langstring collections. This is an informative diagram in that it does not define the information model but supplements Table 3.1 and Table 3.2.

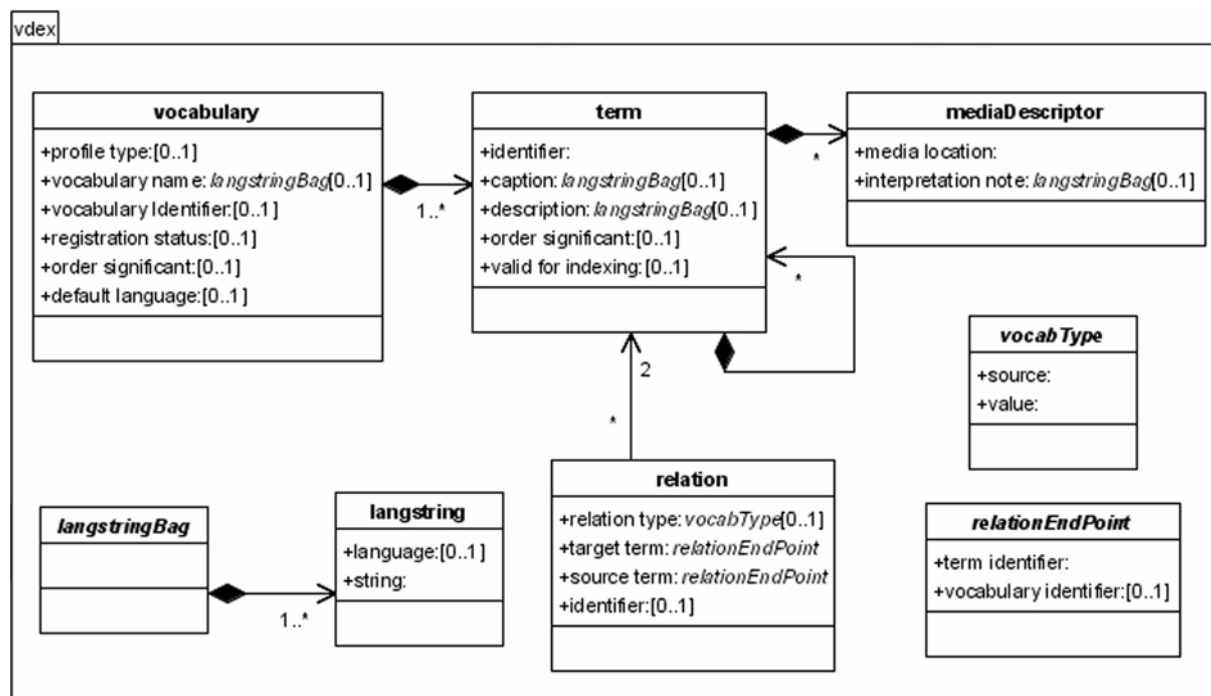


Figure 3.1 - UML representation of the VDEX model.

Relations are shown as always being binary even though the source or target terms may be in a different vocabulary to the relationship.

## 4. Semantics and Value Space for “Profile Type” Property of VDEX Vocabularies

### 4.1 Rationale and Introduction

This section is not normative.

The terminology “profile type” may be equivalently expressed as a “vocabulary type” but both terms run the risk of misinterpretation. We are not describing an application profile here; the VDEX “profile type” is more akin to a label for a subset of a more general information model. The subsets are intended to provide the necessary information elements for various well-known classes of vocabulary (see section 1.3.2). The VDEX “profile type” is not a property of the vocabulary to identify the class of vocabulary being expressed, hence we prefer not to use the terminology “vocabulary type” either.

The profile type is formally a property of the VDEX instance and makes explicit the subset of the VDEX grammar that is in use. This property is used to:

- Permit implementations to support minimal required subsets of the VDEX grammar in order to be as light-weight as reasonable;
- Promote interoperability by allowing exchangers of data to voluntarily reduce the degrees of freedom;
- Allow VDEX to provide a common grammar for several classes of vocabulary, each of which shares a great deal of information concepts but each of which has specific requirements. The alternative approaches of creating small models for each class or modularizing VDEX to a finer granularity are felt to be less maintainable, less intelligible and probably less interoperable than the proposed approach.

Profiling allows a single flexible information model with shared semantic units to be applied selectively if required. Making the profile type an explicit part of the model allows for easy profile-specific validity checking of, for example, XML instances by the use of more restrictive XML Schema control documents.

### 4.2 Definition of the Profile Types

This section is normative.

The profile types defined by VDEX are: “lax”, “thesaurus”, “flatTokenTerms”, “hierarchicalTokenTerms”, “glossaryOrDictionary”. The non-profile, i.e., complete model, is identified by profileType=“lax”.

A profile type other than one of the specified profile types above should be declared as being “lax”, even if only a small number of elements are used. If it is desired to express the identity of such a custom type then the information model should be extended to permit this information to be passed to systems that understand its meaning. Systems that do not understand the significance of a custom type would consider the type to be “lax” since that is the declared type as far as the VDEX information model is concerned.

**Table 4.1 - Summary of profile types.**

The profile type is declared as:	The data:	
lax	Exploits arbitrary elements or cannot be classified as one of the more restrictive profile types. The “lax” profile type makes available the full VDEX model.	Least restrictive
thesaurus	Is a thesaurus. Nesting of “term” data containers is not permitted. Multiple languages can be expressed by using multiple “caption” and “description” langstring elements.	.

**The profile type is declared as:****The data:**

hierarchicalTokenTerms	Is a taxonomy where the primary identifier for the term is a language independent token – the “term” “identifier”. The following elements must not be present: “relatedTerms” (the hierarchical structure is expressed by element containment) and “mediaDescriptors” (a light-weight model is preferred).	.
glossaryOrDictionary	Is an unrelated list of human language terms with some explanatory information. Nesting of “term” containers is not permitted. Multiple language equivalents may be expressed through “caption” and “description” langstrings. The “identifier” of a “term” is present but not functional in the model.	.
flatTokenTerms	Is a flat list of tokens such as used in IEEE 1484-12.1-2002 (the IEEE LOM). This profile imposes the additional constraint to the “hierarchicalTokenTerms” profile: that no “term” may be contained within another “term”.	Most restrictive, most light-weight

**Table 4.2 - Permitted, and required elements and multiplicities for the profile types.**

Profile type: Element	lax	glossaryOrDictionary	flatTokenTerms	hierarchicalTokenTerms	thesaurus
<b>Elements pertaining to the vocabulary as a whole</b>					
Order Significant	0..1	0..1	0..1	0..1	0..1
Profile Type <sup>a</sup>	0..1	1	1	1	1
Default Language	0..1	0..1	0..1	0..1	0..1
Vocabulary Name	0..1	0..1	0..1	0..1	0..1
Vocabulary Identifier	0..1	0..1	0..1	0..1	0..1
Registration Status	0..1	0	0..1	0..1	0..1
Term	1..*	1..*	1..*	1..*	1..*
Relationship	0..*	0..*	0	0	0..*
<b>Elements pertaining to the term</b>					
Identifier	1	1	1	1	1
Caption <sup>b</sup>	[0..1]*	1	[0..1]*	[0..1]*	1
Valid Index	0..1	0	0..1	0..1	0
Description	0..1	0..1	0..1	0..1	0..1
Media Descriptor	0..*	0..*	0	0	0..*



Profile type: Element	lax	glossaryOrDictionary	flatTokenTerms	hierarchicalTokenTerms	thesaurus
Media Location	1	1	-	-	1
Interpretation Note	0..1	0..1	-	-	0..1
Order Significant	0..1	-	-	0..1	-
{ Term }	0..*	0	0	0..*	0
Metadata	0..*	0..*	0..*	0..*	0..*

- a. Profile types with a required value must use the value specific to the relevant profile type. The full VDEX model assumes lax if no value is provided.
- b. 1 indicates a single langstring must be used as the content of the caption. [0..1]\* indicates any number langstrings may be used as the content of the caption (with equivalent content).

**UML multiplicity notation:**

0..1 - optional and not repeatable

0..\* - optional and repeatable

1..\* - required at least once and repeatable

1 - required once only

0 - must not be used

**Notes on using profiles:** The most restrictive profile type should be declared that is sufficient to permit the elements in use.

## 5. Preferred Value Domains for Expressing Thesaurus Relationships

This section is normative. The terms defined in RFC 2119 are used as specified therein.

Monolingual thesauri expressed using VDEX should follow the conventions of d 2788. Multilingual thesauri should follow the recommendations of ISO 5964.

Monolingual thesauri following ISO 2788 should use the following value domain to express the relationship types defined in ISO 2788:

**Vocabulary Identifier:** “[http://www.imsglobal.org/vocabularies/iso2788\\_relations.xml](http://www.imsglobal.org/vocabularies/iso2788_relations.xml)”

Used for the “source” of a relationship type

**Permitted terms:** “USE”, “UF”, “RT”, “BT”, “NT”, “TT”

Used for the “value” of a relationship type

**Meaning of terms:** Defined by ISO 2788

Multilingual thesauri following ISO 5964 should use the following value domain to express the degree of language equivalence:

**Vocabulary Identifier:** “[http://www.imsglobal.org/vocabularies/iso5964\\_equivalences.xml](http://www.imsglobal.org/vocabularies/iso5964_equivalences.xml)”

Used for the “source” of a relationship type

**Permitted terms:** “exact”, “inexact”, “partial”, “singleToMultiple”, “NonEquivalent”

Used for the “value” of a relationship type

**Meaning of terms:** Defined by ISO 5964

IMS will maintain a persistent copy of definitions of these value domains at a URL identical to the vocabulary identifiers above. These definitions will be expressed according to the most recent VDEX version XML binding.

VDEX instances using the above vocabulary identifiers must only use values from the relevant list of permitted terms. IMS reserves all vocabulary identifiers starting “<http://www.imsglobal.org/vocabularies/>”; such identifiers may not be used without the authority of the Board of Directors of IMS Global Consortium. Implementers should also refer to the “IMS Persistent, Location-Independent, Resource Identifier Implementation Handbook v1.0” [IMSPLID].

The preferred use of these value domains is described in the Vocabulary Definition Exchange Specification Best Practice and Implementation Guide.

### Note on other relationship types:

The recommendations of this section apply only in relation to thesauri that follow the ISO standards mentioned. Other types of relationship models are permitted in VDEX and VDEX imposes no requirements as to the value domains and permitted terms for the associated relationship types. Specific mention is made of Z39.19 for monolingual thesauri since it extends the relationship indicators used in ISO 2788.

### Note on multilingual thesauri and exact equivalence:

Multilingual thesauri that are essentially sets of equivalent terms and these sets are related in various ways should probably avoid using the Relationship element but make use of the fact that the Caption element may contain multiple language strings. This makes for greater clarity and economy in expressing the information.

## About This Document

<b>Title</b>	IMS Vocabulary Definition Exchange Information Model
<b>Editor</b>	Adam Cooper
<b>Version</b>	1.0
<b>Version Date</b>	23 February 2004
<b>Status</b>	<b>Final Specification</b>
<b>Summary</b>	This document describes the Vocabulary Definition Exchange Information Model
<b>Revision Information</b>	February 2004
<b>Purpose</b>	Defines the VDEX Information Model
<b>Document Location</b>	<a href="http://www.imsglobal.org/vdex/vdexv1p0/imsvdex_infov1p0.html">http://www.imsglobal.org/vdex/vdexv1p0/imsvdex_infov1p0.html</a>

To register any comments about the VDEX specification, please visit:  
<http://www.imsglobal.org/developers/ims/imsforum/categories.cfm?catid=18>

## List of Contributors

The following individuals contributed to the development of this document:

<b>Name</b>	<b>Organization</b>
Lorna Campbell	CETIS
Adam Cooper	FD Learning
Alex Jackl	IMS Global Learning Consortium, Inc.
Claude Ostyn	Click2learn, Inc.
Michael Pelikan	Penn State University
Brendon Towle	Thomson NETg

## Revision History

Version No.	Release Date	Comments
Public Draft 1.0	4 August 2003	Minor revisions.
Final 1.0	23 February 2004	a) Made minor editorial changes and added new elements to the element section b) Included additional information in section 2.2 about relationship.

# Index

## G

grammar 3, 10, 14

## I

identifier 7, 8, 9, 10, 11

IMS Specifications

Content Packaging 6

Learner Information Package 6

ISO 3, 4, 17

## L

LOM 3, 4, 15

## M

meta-data 3, 8, 10, 11

## N

namespace 4

normative 3, 4, 7, 10, 14, 17

## P

profile 8, 14, 15, 16

## R

Relationship 11

relationship 3, 4, 8, 9, 11, 15, 17

## S

schema 3, 14

## T

taxonomy 3, 8, 15

term 3

thesauri 3, 6, 17

thesaurus 5, 8, 9, 14

## U

URI 7, 10, 11, 12

## V

valid 3, 8, 10

## X

XML 14

*IMS Global Learning Consortium, Inc. (“IMS”) is publishing the information contained in this IMS Vocabulary Definition Exchange Information Model (“Specification”) for purposes of scientific, experimental, and scholarly collaboration only.*

*IMS makes no warranty or representation regarding the accuracy or completeness of the Specification.*

*This material is provided on an “As Is” and “As Available” basis.*

*The Specification is at all times subject to change and revision without notice.*

*It is your sole responsibility to evaluate the usefulness, accuracy, and completeness of the Specification as it relates to you.*

*IMS would appreciate receiving your comments and suggestions.*

*Please contact IMS through our website at <http://www.imsglobal.org>*

*Please refer to Document Name: IMS Vocabulary Definition Exchange Information Model*

*Date: 23 February 2004*

