

What is...

IMS Vocabulary Definition Exchange?

Draft version, by Neil Fegen

In Brief

What is IMS Vocabulary Definition Exchange?

The IMS Vocabulary Definition Exchange (VDEX) specification allows the exchange and expression of simple machine-readable lists of human language terms, along with information that may assist a human in understanding the meaning of the various terms. Put more simply, it is a mark-up language – or grammar – for controlled vocabularies, with VDEX allowing the description and creation of controlled vocabularies that could be easily interchanged.

The term 'vocabulary' is defined in the OED (Oxford English Dictionary) online as 'a collection or list of words with brief explanations of their meanings'. The vocabulary that VDEX is capable of encoding can be of a number of types: a simple flat list of terms, a glossary or dictionary, a thesaurus or a hierarchy of terms. The VDEX specification does not attempt to accommodate all possible vocabularies and cannot support vocabularies of arbitrary complexity; however, there is limited allowance for complex vocabularies, such as faceted schemes, multi-language thesauri and polyhierarchical taxonomies.

VDEX was developed as an open specification by the IMS Global Learning Consortium, the Final Specification being approved in February 2004. The specification includes an extensive Best Practice and Implementation Guide, an Information Model and an XML Binding Guide that defines how to represent a vocabulary in XML.

What is IMS Vocabulary Definition Exchange for?

VDEX is designed to supplement other IMS specifications and the IEEE LOM standard by giving additional semantic control to tool developers. IMS VDEX could be used for the following purposes:

- **Distributing vocabularies among many users** – achieved by simple XML file sharing, or possibly a searchable repository or registry of vocabularies.
- **Interfaces providing pre-defined choices** – providing radio buttons and drop-down menus for interfaces such as metadata editors or a repository browse tool, based on the vocabulary allowed in the metadata profile used
- **XML stylesheets used to select and generate different views** – selecting an overview of an entire vocabulary as an HTML or PDF file, for example; providing scope notes for catalogues; or storing a glossary of terms which are called upon by hyperlinks within a document
- **Validation of metadata instances** – validated against an application profile, by comparison of the vocabulary terms used in certain metadata elements with those of the machine readable version of the vocabularies specified by the application profile.

Some IMS specifications and IEEE LOM contain elements where controlled terms should be used. These elements are often specified as being of a vocabulary data type, and a definition of the permitted terms and their usage may be expressed using VDEX. The content of the vocabulary data type element in LOM, for example, is split into two parts - often identified as the value of the term and the source of the term. The source is considered the equivalent of the value domain identifier (the VDEX vocabulary identifier), while the value should be the term identifier in the value domain (the VDEX term identifier).

Technical Details

How IMS VDEX works

The VDEX Information Model is represented in the diagram below. A VDEX file describing a vocabulary comprises a number of information elements, most of which are relatively simple, such as a string representation of the default (human) language or a URI identifying the value domain (or 'vocabulary'). Some of the elements are 'containers' – such as a **term** – that contain additional elements.

Elements may be required or optional, and in some cases, repeatable. Within a term, for example, a **description** and **caption** may be defined. Multiple language definitions can be used inside a description, by using a **langstring** element, where the description is paired with the language to be used. Additional elements within a term include **media descriptors**, which are one or more media files to supplement a term's description; and **metadata**, which is used to describe the vocabulary further.

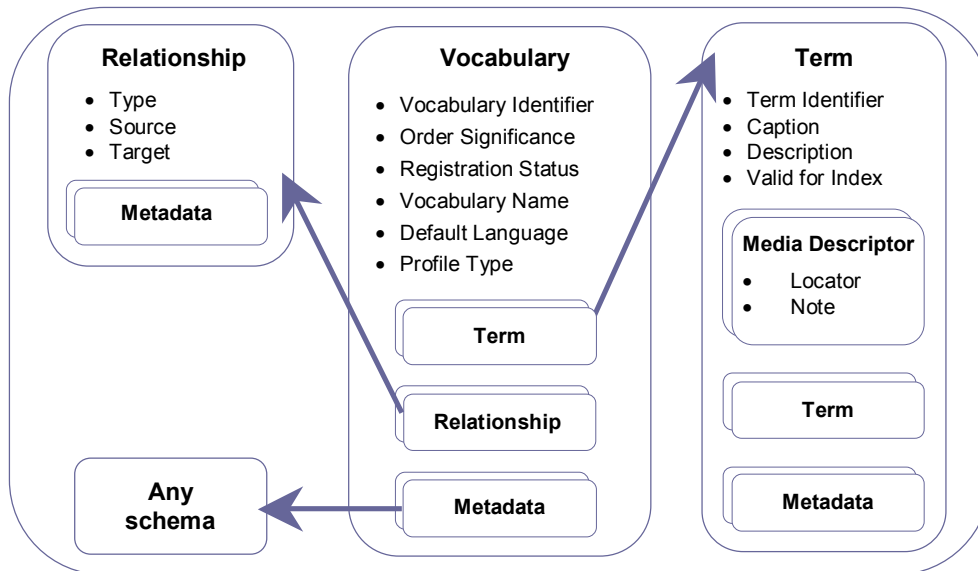
The **relationship** container defines a relationship between terms by identifying the two terms and the specifying type or relationship, such as a term being broader or narrower than another. The term used to specify the type of relationship may conform to the ISO standards for thesauri.

Vocabulary identifiers are unique, persistent URIs, whereas term or relationship identifiers are locally unique strings. VDEX also allows for a **default language** and **vocabulary name** to be given, and for whether the ordering of terms within the vocabulary is significant (**order significance**) to be specified.

A **profile type** is specified to describe the type of vocabulary being expressed; different features of the VDEX model are permitted depending on the profile type, providing a common grammar for several classes of vocabulary. For example, it is possible, in some profile types, for terms to be contained within one another and be 'nested', which is suited to the expression of hierarchical vocabularies. Five profile types exist: 'lax', 'thesaurus', 'heirarchicalTokenTerms', 'glossaryOrDictionary' and 'flatTokenTerms'. The lax profile is the least restrictive and offers the full VDEX model, whereas the flatTokenTerms profile is the most restrictive and lightweight.

VDEX also offers some scope for complex vocabularies, assuming the existence of a well-defined application profile (for exchange interoperability). Some examples are:

- **Faceted Schemes** – faceted vocabularies are possible with the definition of appropriate relationships
- **Multi-lingual thesauri** – metadata could be used within a relationship to achieve multi-lingual thesauri
- **Polyhierarchical Taxonomies** – can be expressed using the source/target value pairs in the relationship.



A simplified VDEX Information Model

Requirements

Identifiers in VDEX data should be persistent, unique, resolvable, transportable and URI-compliant. Specifically, vocabulary identifiers should be unique URIs, whereas term and relationship identifiers should be locally unique strings.

Related specifications

Zthes describes an abstract model for representing and searching thesauri, though it may be used for other types of vocabulary.

IEEE LOM can provide metadata on the vocabularies in a VDEX instance. VDEX can, in turn, be used to describe the controlled vocabularies which are the value space for many LOM elements.

IMS RDCEO – Reusable Definition of Competency or Educational Objective – can be used to create common understandings of terms describing competencies that appear as part of a learning plan, as learning pre-requisites or learning outcomes.

For more information on these specifications, see the ‘Vocabulary Management Technologies Review’ from the JISC Pedagogical Vocabularies Project, available from http://www.jisc.ac.uk/elp_vocabularies.html.

Implementations

The **ALOHA Metadata Tagging Tool** is a Java-based software project that can read IMS VDEX files.

<http://aloha2.netera.ca/>

IVIMEDS 1G v1.0 – from The International Virtual Medical School – includes VDEX instances in curriculum maps. Partners can create their own maps in VDEX format and use these to help students search the repository.

<http://www.ivimeds.org/news/demonstrator.html>

The **Skills Profiling Web Service** project implemented and demonstrated use of a skills profiling web service using open standards in a medical context. IMS VDEX files were used in the representation of the SPWS hierarchy skills framework.

<http://www.elframework.org/projects/spws/view>

The **Scottish Doctors** project used VDEX as a format for expressing curricular outcome systems.

<http://www.scottishdoctor.org/>

VDEX XSLT scripts were developed by The Higher Education Academy Centre for Philosophical and Religious Studies to convert VDEX to XHTML and PostgreSQL.

http://prs.heacademy.ac.uk/technical/vdex_scripts.html

The **VDEX Implementation Project** was carried out by the Institute for Computer Based Learning at Heriot-Watt University, with a primary objective of creating a tool for editing vocabularies in VDEX format. The project, which ended in January 2004, was based on the Public Draft (not the current Final Specification).

<http://www.icbl.hw.ac.uk/vdex>

Resources

People, Products and Services

The CETIS Metadata and Digital Repository Special Interest Group (SIG) has been set up for those in UK Higher and Further Education who are interested in creating, storing and serving educational metadata, and may be of interest to those working with IMS VDEX. Contact the co-ordinator (Phil Barker) at philb@icbl.hw.ac.uk or go to:

<http://metadata.cetis.ac.uk/>.

The CETIS Metadata and Digital Repository Special interest Group (SIG) list may be of interest.

<http://www.jiscmail.ac.uk/lists/CETIS-METADATA.html>

The Controlled vocabularies, thesauri & classification schemes list may be of interest.

<http://www.jiscmail.ac.uk/lists/VOCAB-HE.html>

Resources on the Internet

The IMS website contains both the specification documents and some supporting materials.

<http://www.imsglobal.org/vdex>

Comments

This is a draft version and feedback/comments are welcome.

E-mail me at nf16@macs.hw.ac.uk or use the CETIS-METADATA JISC mail list.