CEN

WORKSHOP

AGREEMENT

CWA 15453

November 2005

ICS 01.020; 03.180

English version

Harmonisation of vocabularies for eLearning

This CEN Workshop Agreement has been drafted and approved by a Workshop of representatives of interested parties, the constitution of which is indicated in the foreword of this Workshop Agreement.

The formal process followed by the Workshop in the development of this Workshop Agreement has been endorsed by the National Members of CEN but neither the National Members of CEN nor the CEN Management Centre can be held accountable for the technical content of this CEN Workshop Agreement or possible conflicts with standards or legislation.

This CEN Workshop Agreement can in no way be held as being an official standard developed by CEN and its Members.

This CEN Workshop Agreement is publicly available as a reference document from the CEN Members National Standard Bodies.

CEN members are the national standards bodies of Austria, Belgium, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland and United Kingdom.

EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION
EUROPÄISCHES KOMITEE FÜR NORMUNG

Management Centre: rue de Stassart, 36   B-1050 Brussels

CONTENTS

FOREWORD

This CEN Workshop Agreement (CWA) provides concepts, techniques and guidelines for the harmonisation of *vocabularies*. The term *vocabulary* is used in this CWA in the broad sense, referring to value lists, classifications, taxonomies, glossaries, dictionaries, ontologies, and thesauri.

Different communities have their own identity, language, and *vocabularies* which are important to express exactly what is meant. For instance the educational system in North America is differently structured than in Europe and it uses different terminology. Even in Europe great differences exist between different sectors. For example, a thesaurus applicable for vocational education is different from the one for schools. These naturally grown differences make the application of standards more complex and often less effective. In addition *vocabularies* are created again and again sometimes with differences that are not essential but that inhibit interoperability.

The CWA also provides tools to specify mappings between learning technology vocabularies.

In parallel to this CWA, an online service that supports the analysis of different thesauri has been developed.

The decision for this work item was taken by the Learning Technologies Workshop at the 14th meeting on March 24/25, 2003. Work on the CWA actually started at the 17th meeting in January 2004.

The CWA was edited by Frans Van Assche with contributions of Sylvia Hartinger, Angelika Harvey, David Massart, Kateryna Synytsya, Anne Wanniart and Marc Willem.

The document has been developed through the collaboration of a number of contributing partners, representing a wide mix of interests, from universities to commercial companies representatives. The names of the individuals and their affiliations that have expressed support for this CWA is available from the CEN/ISSS Secretariat.

The final review/endorsement round for this CWA was started on 2005-06-16/17 and closed on 2005-07-15.

The final text of this CWA was submitted to CEN for approval and publication in 2005-09-06.

This CEN Workshop Agreement is publicly available as a reference document from the National Members of CEN: AENOR, AFNOR, BSI, CSNI, CYS, DIN, DS, ELOT, EVS, IBN, IPQ, IST, LVS, LST, MSA, MSZT, NEN, NSAI, ON, PKN, SEE, SIS, SIST, SFS, SN, SNV, SUTN and UNI.

Comments or suggestions from the users of the CEN Workshop Agreement are welcome and should be addressed to the CEN Management Centre.

ABBREVIATIONS

ADL	Advanced Distributed Learning
CEDEFOP	Centre européen pour le développement de la formation professionnelle
DDC	Dewey Decimal Classification Scheme
ECLAS	European Commission Library Automated System
ELR	European Learning Resource
EUN	European Schoolnet
IEEE	Institute of Electrical and Electronics Engineers Inc.
IETF	Internet Engineering Task Force
ISO	International Organization for Standardisation
ILO	International Labour Organisation
LOM	Learning Object Metadata (usually used in "IEEE LOM")
LTSC	IEEE Learning Technology Standards Committee
OECD	Organisation for Economic Co-operation and Development
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
VET	Vocational Education and Training
W3C	World Wide Web Consortium
XML	Extensible Mark-up Language
XVD	Exchange of Vocabulary Definition Specification (this specification)

1 INTRODUCTION

The sharing of education related data and services makes them less costly, increases the supply, enhances the quality through incremental improvements, and allows a shorter time delivery. However, sharing of data, content, tools and services can only be achieved when clear technical agreements are made between all parties concerned. The more global this agreement is, the greater the benefit. On the other hand different communities have their own identity, language, and *vocabularies* which are important to express exactly what is meant. For instance the educational system in North America is differently structured than in Europe and it uses different terminology. Even in Europe great differences exist between different sectors. For example, a thesaurus applicable for vocational education is different from the one for schools. These naturally grown differences make the application of standards more complex and often less effective. In addition *vocabularies* are created again and again sometimes with differences that are not essential but that inhibit interoperability.

This report provides concepts, techniques and guidelines for the harmonisation of *vocabularies*. The term *vocabulary* is used in this report in the broad sense, referring to value lists, classifications, taxonomies, glossaries, dictionaries, ontologies, and thesauri[1].

The harmonisation of vocabularies as described in this report has three categories of beneficiaries:
- Users who are indexing and searching. Harmonisation of vocabularies facilitates the indexing of resources when the indexer needs to deal with different application profiles. Users searching for resources will benefit when they are searching repositories that used different application profiles, as is typically the case in a federated search.
- Maintainers of *vocabularies*. Maintainers of *vocabularies* are always looking for the best way to represent the concepts that will be used by their audience. Harmonisation of *vocabularies*, such as terminological harmonisation or adopting the same (sub)*vocabularies* reduces the amount of work. Indeed work done by the maintainer of one vocabulary can be maintained by another vocabulary.
- Funding organizations. Funding organisations are evaluating costs versus benefits. On the cost side, harmonisation of vocabularies reduces costs since work of others can be reused more easily. On the benefit site harmonisation of vocabularies is enhanced for the users served by the funding organisations.

This introduction sets the stage in terms of background. Chapter 2 presents the levels of harmonisation for *vocabularies*. Chapter 3 explores concepts and guidelines for the harmonisation of *vocabularies*. More in particular it covers an analysis of harmonisation issues and options, concepts and guidelines of how to design *vocabularies* facilitating harmonisation, and requirements of methods and tools for harmonisation and sharing of *vocabularies*.

---

[1] A full discussion on the different thesauri can be found in the CEN Workshop Agreement 14871 Controlled Vocabularies for Learning Object Metadata - Typology, impact analysis, guidelines and a web based Vocabularies Registry.

Chapter 4 introduces a conceptual model and bindings for vocabularies. This conceptual model is, wherever possible, compatible with the IMS VDEX information model[2], Zthes[3], and the Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies[4]. One of the important motivations for this work is that the work referred to was not dealing with poly-hierarchies of terms for instance when using micro-thesauri, alternative structures, and mappings between thesauri. Hence a new comprehensive approach was required.

Chapter 5 introduces specific techniques for the harmonisation of vocabularies. Given that this CWA discusses harmonisation techniques especially in terms of thesauri, chapter 6 relates the presented work to other types of vocabularies.

The report focuses on thesauri as they are one of the more complex *vocabularies* available. The harmonisation of other *vocabularies* can be achieved by using subsets of the concepts, guidelines, and techniques presented in this report. Chapter 7 is devoted to a discussion of this.

---

[2] See http://www.imsglobal.org/vdex/

[3] http://zthes.z3950.org/xml/

[4] http://www.niso.org/standards/