

CEN

CWA 16100

WORKSHOP

February 2010

AGREEMENT

ICS 35.240.30

English version

Guidelines for the design, implementation and operation of a product property server (ePPS)

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, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland and United Kingdom.



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

Management Centre: Avenue Marnix 17, B-1000 Brussels

© 2010 CEN All rights of exploitation in any form and by any means reserved worldwide for CEN national Members.

Ref. No.:CWA 16100:2010 E

Contents

Page

Foreword	4
Scope	6
1 Introduction to Part 1	11
2 Data models for dictionaries	11
2.1 General aspects	11
2.2 ISO 13584 (PLIB) / IEC 61360	12
2.3 ISO 29002-6 (Terminology model)	13
2.4 Empirical results from the survey on dictionary models	15
3 Exchange formats	23
3.1 ISO 13584-32 (OntoML)	23
3.2 BMEcat	24
3.3 UN/CEFACT model	25
3.4 Referencing dictionary elements in catalogues	26
3.5 Empirical results from the survey on exchange formats	27
4 Multilingualism	34
4.1 Aspects of Multilingualism	34
4.2 Empirical results from the survey on Multilingualism (provider questionnaire)	34
5 System Architecture	44
6 Dictionary servers, Interoperability and ISO 29002	45
6.1 Interoperability	45
6.2 Overview of ISO 29002	46
6.3 Requirements and design principles of ISO 29002-20	48
6.4 Architecture and usage of the web service	48
6.5 Elements of the Web Service	49
6.6 Conclusion, status	51
6.7 Recommendations for dictionary servers	51
7 Requirements placed by ISO 8000 standards on the operation of dictionary servers	53
7.1 Overview on ISO 8000	53
7.2 Recommendations for dictionary server providers	55
8 Standards and formats used in HVAC industry	56
9 Summary and conclusion	57
Part 2 Organizational aspects of operating an ePPS	58
10 Electronic metadata dictionaries	58
10.1 Introduction	58
10.2 Problems	58
10.3 Consequences	58
11 Workflows	59
11.1 Introduction	59
11.2 Status quo	59
11.3 Analysis of current dictionaries implementations (provider)	62
11.4 Analysis of industry requirements regarding content update (industry)	76
11.5 Conclusion	77
11.6 Recommendations	77
12 Legal and economic requirements	78

12.1	Introduction.....	78
12.2	Licensing of dictionary content	78
12.3	Intellectual Property Rights (IP).....	79
12.4	Copyright.....	79
12.5	Public Domain.....	79
12.6	Free Licenses for software	80
12.7	Commercial licenses for content.....	80
12.8	Status Quo - Handling IPRs in current dictionary implementations	81
12.9	Status Quo - IPRs of Dictionary Content used in Industry	89
12.10	Conclusion	95
12.11	Recommendations	95
13	Skills and competences.....	96
13.1	Introduction.....	96
13.2	Business model	96
13.3	Developing and maintaining content	97
13.4	Localization	98
13.5	Education and Training.....	99
13.6	Support	99
13.7	Operation of Dictionary.....	100
13.8	Supporting users with special needs	100
Part 3	Usage of ePPS in industry	102
14	Introduction.....	102
15	Scenarios for use of ePPS in industry	103
15.1	Organization providing classification system.....	104
15.2	Company providing electronic catalogues.....	105
15.3	Company providing product descriptions for internal systems	106
16	Integration of ePPS into user processes	107
16.1	Formalization of catalogue requirements	107
16.2	Automated access to dictionary content	108
16.3	Integration of user processes and dictionary access via ISO 29002-20 Web Services.....	108
17	Requirements and usage of ePPS	111
17.1	Usage of company internal ePPS	111
17.2	Usage of external ePPS	114
17.3	Personal skills	121
18	Support concepts for implementation adapted to users and supporting software companies.....	123
18.1	Formulation of benefits of using dictionaries and supporting ePPS	123
18.2	Guidance for planning, establishment, maintenance and usage of an international PPS.....	125
19	Strategies and activities for spreading the knowledge into industry	130
19.1	Activities within project ePPS.....	130
19.2	Conceptual design of training and educational courses in the CEN Focus Group eSME.....	131
19.3	Building a concept for evaluating the results in industrial practice for a longer period of time.....	133
Annex A:	Survey on Dictionary Providers	134
1.	Organization of the Questionnaire	134
2.	Analysis of Dictionary Provider Questionnaire.....	135
3.	Provider Questionnaire.....	140
Annex B:	Survey on Dictionary Users.....	148
1.	Organization of the Questionnaire	148
2.	Participants	149

Foreword

This CEN Workshop Agreement on "Guidelines for the design, implementation and operation of a product property server" has been prepared by the CEN Workshop "Multilingual eCataloguing and eClassification in eBusiness" in the context of the project "electronic Product Property Servers" (ePPS). The production of this CWA was approved at the Workshop eCAT plenary meeting on 4 February 2008.

Over thirty companies and organizations from industry, SMEs, software developers, consulting companies, universities and research centers specialized in cataloguing and product property servers participated in the CEN Workshop eCAT – ePPS project.

This CWA is the seventh CWA developed within the Workshop eCAT. The list of already available CWAs is available under "References".

The document was prepared by the eCAT/ePPS Project Team:

- **Dr. Wolfgang Wilkes**, Fachhochschule Hagen, Germany,
- **Reinhard Pohn**, Paradine GmbH, Austria
- **Volker Schmitz**, University of Duisburg-Essen, Germany,

In his capacity as Chair of the Workshop **Dr. Christian Galinski**, Infoterm - International Information Centre for Terminology, Austria, has contributed greatly to the work with the CWA. The Secretary of the Workshop has been **Florian Rieger**, DIN.

The following companies/organizations endorsed the CWA:

- AFIM (France)
- CNCTST (China)
- DIN e.V. (Germany)
- ECCMA (USA)
- eCl@ss – International Classification System (Germany)
- ETIM e.V. (Germany)
- ETIM International (The Netherlands)
- FernUniversität in Hagen (Germany)
- funSTEP
- GS1 (USA)
- Infoterm – International Information Centre for Terminology
- Paradine GmbH (Austria)
- PIDX (USA)
- proficl@ss e.V. (Germany)
- Renault (France)
- SEEM (Czech Republic)
- Semaino Technologies GmbH (Germany)
- TermNet – the International Network for Terminology
- VdZ – Vereinigung der deutschen Zentralheizungswirtschaft e.V. (Germany)

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, HZN, LVS, LST, MSA, MSZT, NEN, NSAI, ON, PKN, SEE, SIS, SIST, SFS, SN, SNV, SUTN and UNI.

Introduction

A fundamental requirement for implementing full e-business is the ability to exchange information about products between different business partners and their software systems. As far as product-numbers, prices, delivery information, etc. are concerned, exchange is possible on the basis of general models which are applicable for any kind of product. However, the technical diversity of products leads to a diversity of technical descriptions. This makes it impossible to define a general model covering all aspects of all types of products in a concise way: The description of e.g. bolts and washers requires fundamentally different information than the description of integrated circuits or refrigerators.

Whereas STEP tries to define a big number of models for various domains to describe e.g., geometry or other representation models of a product, the exchange of technical product information in e-business and e-engineering is basically done by describing products by their characteristics or properties. This information exchange is normally based on a meta data approach: Information about a property of a product is exchanged as a pair (property_ref, property_value), where the property_ref is an identifier of a concept in a product dictionary (often called ontology or classification). For the correct interpretation of the property_value, the receiving system has to refer to the product dictionary, where the meaning of the property is defined (textually and possibly supported by graphical means) and further information is available like names (in different languages), synonyms, data type, unit of measure, relationship to other concepts, etc. Thus, we have a very simple structure for the exchange, but we need additional resources which are referenced from this exchange structure.

To provide such resources, a number of organizations have started to build dictionaries to provide the required meta data which can describe the content of a product catalogue. This started with the provision of paper lists of dictionary elements (product classes and their properties) in the standard IEC 61360-4, but today the trend is towards online databases containing the dictionary elements and providing web based and potentially automatic access to the dictionary elements. The range of meta data provided varies dramatically: Some provider only provide terminological information, others provide more information which is based on specific data models, and others provide quite different types of meta data like code sets or graphical symbols.

The provision of an electronic server which allows users to access data has a lot of implications: The meta data needs to be organized in a structured way, access mechanisms need to be provided to allow users to download data and to upload proposals, a business model has to be created to ensure the operation of the server, maintenance of the meta data has to be ensured in an open way to allow participation of the involved organizations and people, etc. Thus, the operation of an electronic Product Property Server is by no means an easy task and requires a lot of consideration.

Scope

The goal of this CWA is to consider various aspects of ePPS (electronic product property servers) which are operated to support the management, distribution and maintenance of meta data dictionaries. In part 1, technical aspects and standards will be discussed, in part 2, organizational aspects like requirements for all kinds of workflows, the evaluations of business models comprising also draft licensing agreements, copyright management issues, etc. are addressed, and in part 3, the user aspects are discussed. To prepare this CWA, the following activities were accomplished.

- Market analysis to identify current implementations and research on requirements and needs
- Analysis of current implementations including a survey with a standardized questionnaire
- Examination of the returned questionnaires and consolidation of learning into a list of requirements and needs

Part 1 of this CWA contains information about dictionary models which are in use for modeling dictionary information and exchange mechanisms for dictionary information. Specifically the standard ISO 29002 is described as a mechanism to exchange information from various dictionaries by means of a single interface. Since ISO 8000 puts specific requirements on the exchange of catalogues and makes dictionaries an indispensable element of this activity, the basic concepts of ISO 8000 are also discussed.

Part 2 of this CWA contains information about the workflows which were developed and are used by the organizations providing ePPS-systems as well as the legal requirements to distribute and use content. Two questionnaires were sent out to dictionary providers and to dictionary users to survey existing implementations and learn about requirements from providers and industry perspective. Also skills and competences required to participate in the development and maintenance of content are discussed. Dictionary providers offer different business models on how to participate in content generation and also how content of the dictionaries is made available to users.

Part 3 discusses ePPS from the point of companies and industries. It describes a number of scenarios in which ePPS can be used beneficially and gives information about organizational requirements for companies to use external ePPSs and illustrates how the use of and the direct connection to an ePPS can support the automation of catalogue related processes. In addition, in Part 3 a number of strategies and measures are proposed (which to a big degree have already been started) for spreading the knowledge on the use of ePPS in industry and in SMEs. Part 3 also contains checklists which may guide companies and standard providers in their way to realize an ePPS for their meta data.

Thus, the scope of this CWA contains:

- Overview on state of the art of product property servers in general
- Basic requirements concerning the data model under specific consideration of available standards, insurance of interoperability with users and other product property servers, multilingualism and requirements of people with special needs, etc.;
- Input/import and output/export formats, again under consideration of available standards
- Requirements for (web) services which are expected by users from a product property server
- Means to support a syntactic and semantic interoperability between different product property servers, in particular web services as defined in ISO 29002 and necessary specifications to relate other dictionary models, in particular ISO 13584, to the web service models and interfaces

- Requirements for keeping consistency between meta data and product data (validation of data, requirements coming e.g. from ISO 8000-110)
- Research of existing projects and data formats (e.g. VDI 3805) within HVAC industry dealing with metadata and compilation of a list of relevant aspects for a product property server in this field
- Requirements of the workflow management of development, distributed updating and maintenance of metadata content
- Requirements concerning workflows for the preparation of multilingual product property data, the control of correspondence between the data and the metadata, etc.
- Legal and economic requirements of a PPS (e.g. IPR, business model, licensing, etc.)
- Verification of available implementations regarding business model, licensing and IPRs
- Industry requirements on provision of metadata (e.g. access, format, tools, training, etc.)
- Types of skills and competences required
- Organizational requirements for companies to effectively use ePPS
- Activities to spread the knowledge on the use of ePPS and meta data in exchange processes
- Checklists to guide companies and standards organizations to develop their ePPS for managing their meta data

Normative References

CWA 15294:2005 — ePDC project — Dictionary of Terminology for Product Classification and Description

CWA 15295:2005 — ePDC project — Description of References and Data Models for Classification

CWA15556-1:2006 — Gen-ePDC project — Product Description and Classification — Part 1:New Property Library

CWA 15556-2:2006 — Gen-ePDC project — Product Description and Classification - Part 2: Product Classes with sets of properties

CWA 15556-3:2006 - Gen-ePDC project - Product Description and Classification - Part 3: Results of development in harmonization of product classifications and in multilingual electronic catalogues and their respective data modelling

ISO 13584-42:2008 Ed.2: Industrial automation systems and integration — Parts library — Part 42: Description methodology: Methodology for structuring part families (to be published)

ISO/TS 29002-5:2009: Industrial automation systems and integration — Exchange of characteristic data — Part 5: Identification scheme

ISO/TS 29002-20:2009: Industrial automation systems and integration — Exchange of characteristic data — Part 20: Concept dictionary resolution services (to be published)

ISO 8000-110:2009: Data Quality — Part 110: Master data: Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification

ISO/IEC Directives – Annex ST

<http://isotc.iso.org/livelink/livelink?func=ll&objid=6613777&objaction=open>

Informative References

BMEcat (Bundesverband für Materialwirtschaft und Einkauf), 2005 www.bmecat.org.

CWA 15045:2004, Multilingual catalogue strategies for eCommerce and eBusiness www.cen.eu/iss

CEN Focus Group eSMEs Final Report, 2010

<http://www.cen.eu/cenorm/sectors/sectors/iss/activity/fg-esmes.asp>

eCl@ss General Conditions of Use <http://www.eclass.eu>

Edwards, Mynatt — Enabling Technology for Users with Special Needs, CHI '95 Proceedings, 1995 Conference on Human Factors in Computing Systems, Denver, Colorado, USA

EIF v1.0 (European Interoperability Framework) for pan-European eGovernment Services, European Commission IDABC (Interoperable Delivery of European eGovernment Services to public Administrations, Businesses and Citizens), 2004 <http://ec.europa.eu/idabc/en/document/3761/5583>.

Draft document as basis for EIF 2.0, European Commission, IDABC, 2009

<http://ec.europa.eu/idabc/servlets/Doc?id=31597>.

GNU General Public License Version 3, 29 June 2007, <http://www.gnu.org/copyleft/gpl.html>

GNU Lesser General Public License, Version 3, 29 June 2007 <http://www.gnu.org/licenses/lgpl.html>

Regulation No 2195/2002 of the European Parliament and of the Council of 5 November 2002 on the Common Procurement Vocabulary (CPV)

http://ec.europa.eu/internal_market/publicprocurement/legislation_en.htm#standard

European Union Public License - EUPL v.1.1, July 2009, <http://ec.europa.eu/idabc/eupl>

Eurotermbank -Collection of Pan-European Terminology Resources through Cooperation of Terminology Institutions – Deliverable 2.2 – Standard document templates and procedures (2006)

Jérôme Euzenat, Pavel Shvaiko Ontology Matching, Springer Berlin Heidelberg 2007.

Galinski, Raupach - Business models as a means of IPR Protection of structured content – taking terminological data as an example: 2006, In: SCHAFFERT, S.; SURE, Y. Semantic systems - From visions to applications. Proceedings of the SEMANTICS 2006. Vienna

ISO 10303-11:2004 Industrial automation systems and integration — Product data representation and exchange — Part 11: Description methods: The EXPRESS language reference manual

ISO 13584-25:2004: Industrial automation systems and integration — Parts library — Part 25: Logical resource: Logical model of supplier library with aggregate values and explicit content

ISO 13584-32:2009: Industrial automation systems and integration — Parts library — Part 32: Implementation resources: OntoML: Product ontology markup language (to be published)

ISO 13584-501:2007: Industrial automation systems and integration — Parts library — Part 501: Reference dictionary for measuring instruments — Registration procedure

ISO 13584-511:2006: Industrial automation systems and integration -- Parts library — Part 511: Mechanical systems and components for general use — Reference dictionary for fasteners

ISO 15926-2:2003: Industrial automation systems and integration — Integration of life-cycle data for process plants including oil and gas production facilities — Part 2: Data model

ISO/TS 22745-1:2009 Industrial automation systems and integration — Open technical dictionaries and their application to master data — Part 1: Overview and fundamental principles

ISO 23584-1:2009: Optics and photonics — Specification of reference dictionary — Part 1: General overview on organization and structure

ISO/TS 29002-4:2009: Industrial automation systems and integration — Exchange of characteristic data — Part 4: Basic entities and types

ISO/TS 29002-6:2009: Industrial automation systems and integration — Exchange of characteristic data — Part 6: Concept dictionary terminology reference model (to be published)

ISO/TS 29002-10:2009: Industrial automation systems and integration — Exchange of characteristic data — Part 10: Characteristic data exchange format

ISO/TS 29002-31:2009: Industrial automation systems and integration — Exchange of characteristic data — Part 31: Query for characteristic data

ISO/TS 8000-100:2009: Data Quality — Part 100: Master data: Overview.

ISO 8000-102:2009: Data Quality — Part 102: Master data: Exchange of characteristic data: Vocabulary.

ISO/TS 8000-120:2009: Data Quality — Part 120: Master data. Exchange of characteristic data: Provenance

CWA 16100:2010 (E)

ISO/TS 8000-130:2009: Data Quality — Part 130: Master data. Exchange of characteristic data: Accuracy

ISO/TS 8000-140:2009: Data Quality — Part 140: Master data. Exchange of characteristic data: Completeness

ISO/NP 8000-150: Data Quality — Part 150: Master data. Quality management framework

ISO/IEC 11179-1:2004: Information technology — Metadata registries (MDR) — Part 1: Framework

ISO/IEC Guide 61: 1996 General requirements for assessment and accreditation of certification/registration bodies www.iso.org

ISO Guide 69:1999 - Harmonized Stage Code system (Edition 2) — Principles and guidelines for use

Universal Copyright Convention, with Appendix Declaration relating to Articles XVII and Resolution concerning Article XI 1952, <http://www.portal.unesco.org>

UN/CEFACT Catalogue Process. Business Requirements Specification (BRS), Document Identification: CEFACT/Forum/2006/TBG1 <http://www.uncefactforum.org/>

VDI 3805 – Produktdatenaustausch in der Technischen Gebäude Ausrüstung (Product data exchange for building services), VDI Recommendation, 2004.

W3C: OWL Web Ontology Language — Overview. W3C Recommendation, 10 February 2004 (<http://www.w3.org/TR/owl-features/>)

W3C: RDF Vocabulary Description Language 1.0: RDF Schema. W3C Recommendation, 10 February 2004 (<http://www.w3.org/TR/rdf-schema/>).

Wikipedia - <http://www.wikipedia.org>

Wolfgang Wilkes, Reinhard Pohn: Spezifikation und pilotmäßige Umsetzung von Geschäftsprozessen zum elektronischen Austausch von Merkmal-Beschreibungen und Produktdaten (Specification and pilot applications of business processes for the electronic exchange of property descriptions and product data). DIN – INS 2008.

WIPO, Berne Convention http://www.wipo.int/treaties/en/ip/berne/trtdocs_wo001.html

WIPO “Understanding Copyright and Related Rights” http://www.wipo.int/freepublications/en/intproperty/909/wipo_pub_909.html

WIPO Intellectual Property Handbook: Policy, Law and Use <http://www.wipo.int/about-ip/en/iprm/index.html>

WIPO Guide to Intellectual Property Worldwide <http://www.wipo.int/about-ip/en/ipworldwide/country.htm>

Part 1 Technical aspects of operating an ePPS

1 Introduction to Part 1

Over the last years, a number of standards have been developed to support the exchange of product data:

- product dictionary data models (e.g. PLIB / ISO 13584, OTD / ISO 22745, ISO 15926, OWL / RDFS),
- standardized product dictionaries developed by standardization bodies, e.g. for fasteners (ISO 13584-511), for electrical measuring instruments (ISO 13584-501), for optical instruments (ISO 23584), or the DIN Server covering a big number of technical domains (www.dinsml.net),
- product dictionaries developed by industry associations, e.g. eCI@ss (www.eclass.de), proficl@ss (www.proficlass.de), ETIM (www.etim.de and www.etim.nl), eOTD (www.eccma.org), RosettaNet Technical Dictionary (RNTD, www.rosettanet.org), ECALS (www3.elisasp.net/ecals), etc.,
- catalogue exchange formats like OntoML [ISO 13584-32], BMEcat [BMEcat 2005].

For the operation of a dictionary server it is important to provide a formal model of the dictionary which can be downloaded and accessed by catalogue tools. Only then a true system-to-system communication can be ensured which is the basis for an electronic exchange of product data. In addition, a formally described exchange format is required to allow the automatic processing of exchanged dictionary information. Therefore, this Part of the ePPS CWA will briefly describe some important standards for the description of products and the exchange of product information:

- ISO 13584 (PLIB)
- ISO 29002-6 (Terminology model)
- OntoML (ISO 13584 Part 32)
- BMEcat 2005
- UN/CEFACT BRS on Catalogue Exchange

To examine the state of the art of these dictionary servers and to recommend ways forward, the project ePPS was initiated in the context of the CEN/ISSS Workshop eCAT. The special goals of this part of the CWA are to analyze the standards in use for running dictionary servers, to give an overview about actual techniques used in running dictionary servers, and to highlight important technical aspects like interoperability of different dictionary servers and the requirements on dictionary servers placed by the upcoming standard ISO 8000. In addition, this CWA gives a brief overview about the German standard VDI 3805, which is used to exchange product information in the HVAC industry. This standard is currently lifted to an ISO standard in ISO TC 59.

2 Data models for dictionaries

2.1 General aspects

A dictionary model defines the means which are available for the specification of a concrete dictionary. It defines which dictionary elements are available (e.g. classes, properties, special kinds of relationships between dictionary elements, possibilities to specify constraints in the dictionary which then will be imposed on the actual values for describing products, like data types, specialized data types, etc.). The information which is used to describe a dictionary can be distinguished in 2 parts:

CWA 16100:2010 (E)

- Terminological information: Contains the information about dictionary elements which are basically required by humans to understand the dictionary element (names, definitions, graphical information, additional information like notes and remarks, etc.)
- Ontological information: Contains the information about dictionary elements which are necessary for further automatic processing (relationship to other dictionary elements, allowed data types, etc.)

One important element of dictionary models is the identification of their dictionary elements. This is of special importance because an identification mechanism which allows to uniquely identify each dictionary element, regardless of the underlying model and the provider of the specific dictionary, is one of the basic cornerstones for supporting the mutual exchange of dictionary information across dictionary providers and to allow the exchange of product information which is not only bound to a single dictionary.

2.2 ISO 13584 (PLIB) / IEC 61360

ISO 13584 is a standard that provides a formal data model for dictionaries and libraries, i.e. for both the metadata and the actual product data. The model is defined in the modeling language EXPRESS (ISO 10303-11), and the model is used also for IEC 61360. The two responsible committees, ISO TC 184 SC4 WG2 and IEC SC3D continue to work on the basis of the same model, also for the current edition 2 of PLIB which is going to reach the IS level by the end of 2009.

Although the PLIB model covers the dictionary level (i.e. the meta level which describes how products can be represented) and the product level (where concrete catalogue products are represented) we discuss in the following only the dictionary level.

The most important concepts which are provided by PLIB are classes and properties. Classes represent product groups or product types (which might be abstract), and properties specify the characteristics of products.

PLIB distinguishes between two fundamental types of classes:

- Categorization classes are used to group products into sets. An example is the UNSPSC code system (www.unspsc.org). Categorization classes cannot have properties.
- Characterization classes provide properties which can be used to characterize products. They are normally organized in a class hierarchy.

Class hierarchies can be built by two types of relationships:

- The "is-a" relationship builds a strict tree (or a set of trees) of characterization classes. "is-a" includes complete property inheritance (i.e. each property of the higher level class is inherited by the lower level class).
- The "case-of" relationship relates any kind of classes (characterization and categorization classes), and it allows an individual inheritance of properties. The relationship defines which properties are "imported" into the lower level class or which properties of the lower level class can be unified with properties of the higher level class.

Properties have always a definition class, i.e. they are defined for that class and all subclasses thereof. However, they may not be used automatically for characterizing products of that class – for that purpose, the property has to be made "applicable" for that class. This distinction between the class for which a property is defined and the applicability of that property allows for an exception mechanism and to have subclasses of a class which use a property defined in the super class and others which do not use it for characterizing products.

Properties are of various types which define the values which are allowed to be used for characterizing a product according to that property. Examples are simple types like integer, number, string, but there are also more complex types like measured types (which have a unit), aggregate types (arrays, lists and sets), and references to other classes which are used to establish relationships between products with the meaning of composition (i.e. the referenced instance is regarded as a component of the referencing instance).

Both classes and properties can be described by a number of human readable information: preferred name, synonyms, definitions, notes, remarks, graphical elements, etc. Most information can be presented in several languages (preferred name, definition, notes, remarks, etc.), and synonyms may also be of different languages.

2.3 ISO 29002-6 (Terminology model)

ISO 29002 is described in more detail in clause 6, here only the terminology model of ISO 29002-6 is discussed. The development of this model was led by the observation that all dictionary models describe their concepts (i.e. classes, properties, suppliers, etc.) by some terminological information like names, definitions, graphics, etc. ISO 29002-6 defines a model of these terminological descriptions of dictionary concepts with the goal to allow all dictionary models to map to this terminology model and to allow an exchange of human readable information about dictionary elements across dictionary models.

The basic idea of ISO 29002-6 is to define "concepts" and to associate various kinds of descriptive information to the concept. This is shown in Picture 1: The centre of the model is built by the entity concept, and the relationships show that a concept is named by a number of terms, defined by some definitions, designated by symbols and possibly illustrated by images. Definitions, terms, symbols and even images are terminological items which are associated with a language and may come from a specific document (e.g. a standard).

Thus, the ISO 29002-6 model only covers the human readable description of dictionary elements (concepts). Related to PLIB, it covers the elements which are used there for giving human readable information for classes and properties.

2.4 Empirical results from the survey on dictionary models

In the context of the empirical survey, the following five questions of the provider questionnaire dealt with data models of dictionary servers. The results are laid out in the following subsections. We have contrasted them with the results of the user questionnaire, where companies could also describe their own dictionary server which they either use for internal purposes or for supporting the exchange of product data with customers or suppliers.

The following questions from the provider questionnaire were related to data models:

- Do you have a formally specified data model which describes how your content is organized?
- Is your data model based on other models?
- Which mechanism do you use for specifying your data model (e.g. XML DTD, XML Schema, UML, tables, others)?
- If you do not have a formally specified data model, how is the structure of your content specified?
- Have you ensured that each element of your content has a world-wide unique identifier?

In the user questionnaire, the following questions are considered here:

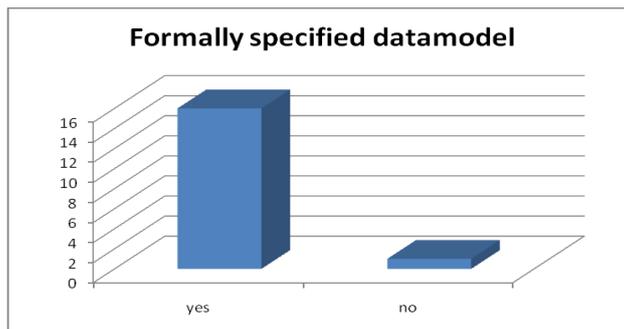
- Do you have a formally specified data model which describes how your content is organized?
- Is your data model based on other models?
- Which mechanism do you use for specifying your data model (e.g. XML DTD, XML Schema, UML, tables, others)?
- If you do not have a formally specified data model, how is the structure of your content specified?

2.4.1 Do you have a formally specified data model which describes how your content is organized?

From the provider questionnaire, 17 participants answered as follows.

No Answers	Option
16	Yes
1	No

Table 1: Providers with a formally specified data model



Picture 2: Provider result – Formally specified data model

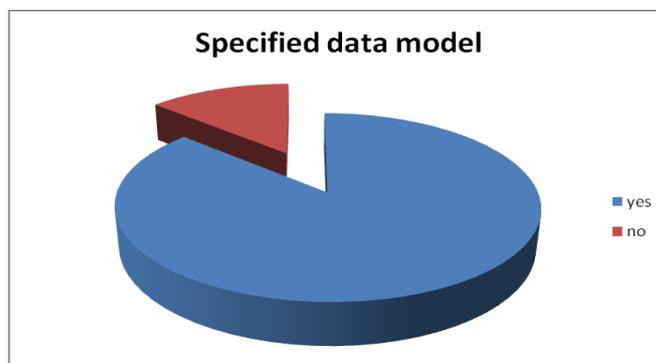
To really support the automatic processing of data and to allow for an easy integration of meta data into tools and applications, it is mandatory to have some model in which the meta data is formulated and thus process able by tools. All but one claim to have a formal model, the only provider without a formal data model is PIDX. But in one of the following questions they say that they use XML schemas for describing their data model – so we can assume that they also provide a data model.

The same question was asked in the user questionnaire with respect to companies' own dictionary and metadata providers. Here we got the following results:

Out of 26 possible answers the following replies were made.

No. of Answers	Option
13	Yes
2	No

Table 2: Users dealing with a formally specified data model



Picture 3: User result: Formally specified data model

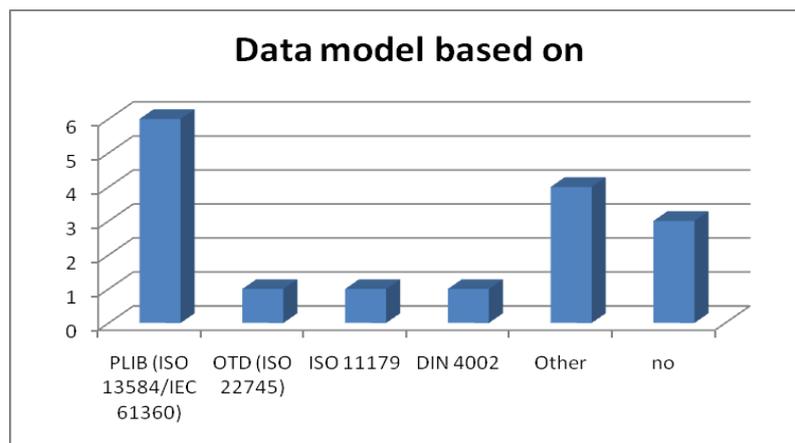
The majority of participants replying stated that a formally specified data model exists. More information on the data model was requested in the following questions.

2.4.2 Is your data model based on other models?

This question was supposed to be answered by those dictionary server providers who claim to have a formal data model as basis of their dictionary server. It was possible to make more than one statement as there may be an option of supporting different data models. Out of 11 potential answers (providers with a formal data model), the following statements were made.

No Answers	Option
6	PLIB (ISO 13584/IEC 61360)
1	OTD (ISO 22745)
1	ISO 11179
1	DIN 4002
4	No
3	Other

Table 3: Standardized data models used by dictionary providers



Picture 4: Provider questionnaire: Data model based on standards

Obviously, the biggest group is still the group of providers who use their homegrown data model. Another big group is build by the dictionaries that base their description on the ISO 13584 and its derivative DIN 4002. One dictionary provider uses the ISO 22745, another dictionary data model of ISO. Thus, it seems to be a challenge to identify how the big number of self-made models can be used together with the standard based data models.

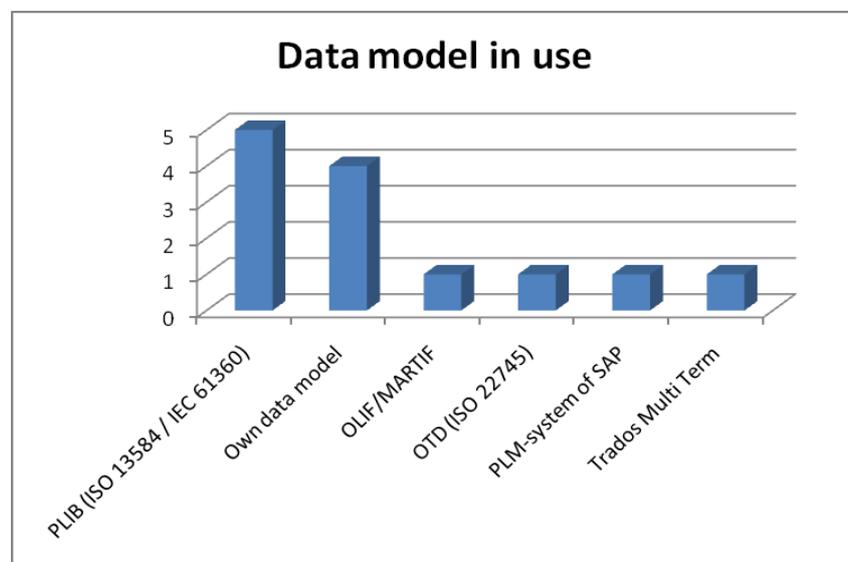
CWA 16100:2010 (E)

The same question in the user questionnaire for users providing their own in-house dictionary server yielded the following results:

Out of 26 possible answers the following replies were made.

No. of Answers	Option
5	PLIB (ISO 13584 / IEC 61360)
4	Own data model
1	OLIF/MARTIF
1	OTD (ISO 22745)
1	PLM-system of SAP
1	Trados Multi Term

Table 4: Dictionary models used in company specific dictionary servers



Picture 5: User Questionnaire – Data model in use

The analysis of the replies shows that the PLIB (ISO 13584/IEC61360) data model was named most from the participants. The second most named is an own data model. The other data models named were only named once. Additionally to the offered data models the users mentioned OLIF/MARTIF, SAP specific format and the Trados Multi Term data model.

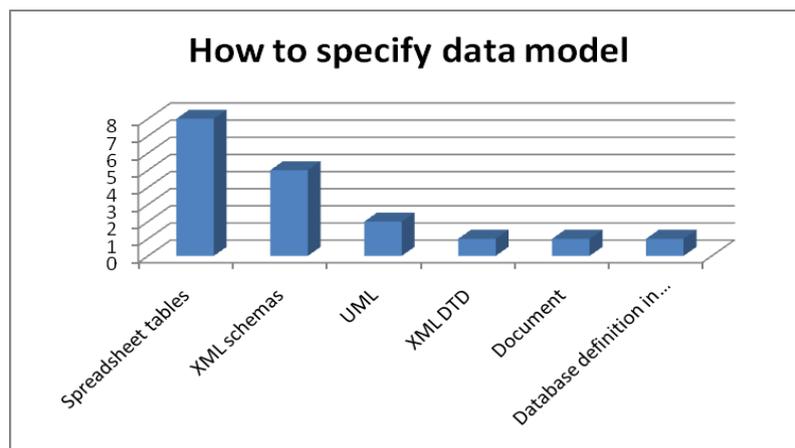
2.4.3 Which mechanism do you use for specifying your data model (e.g. XML DTD, XML Schema, UML, tables, others)?

It was possible to make more than one statement as there may be more than one options to specify the data model. The result of this question from the provider questionnaire was as follows:

11 potential answers (providers with a formal data model), thereof

No Answers	Option
8	Spreadsheet tables
5	XML schemas
2	UML
1	XML DTD
1	Document
1	Database definition in application

Table 5: Mechanisms used by dictionary providers to specify their data model



Picture 6: Provider questionnaire – How to specify data model

Obviously tables (which might be a set of CSV tables or Excel tables or tables in a relational database like Access) are still the most used way to disseminate the meta data content. On the other hand, XML has already reached a good position, and sometimes it is supported by an UML model describing the structures of the XML schema. Of course, table structures are different and as well XML structures are different, so again we see potential problems if more than a single meta data set needs to be used in a single application.

CWA 16100:2010 (E)

In the user questionnaire we also asked this question for those participants who stated that they use a formally specified data model and for those who mentioned that they plan to use an ePPS in the future. Here the question was an open question without predefined answers.

The users mentioned

No. of Answers	Option
3	XML-schema
2	XML
2	xls, csv
1	XML-DTD
1	ABAP schemes
1	BMEcat
1	database tables
1	UML
1	EDIDATA (EDIFACT)
1	DATANORM

Table 6: Mechanisms used by company dictionaries to specify its data model

The analysis of the replies shows that the majority of the recipients use some form of XML-representation to specify the data model. xls/csv representation was mentioned twice and all other data models were mentioned only once.

2.4.4 How is the structure of your content specified (if you do not have a formal data model)?

From the **provider's questionnaire**, we did not get any response (including the responder who stated he had no formal data model).

In the **user's questionnaire**, two providers of company internal dictionary servers gave response to this question. One participant stated that he specifies content according to DIN 4000 and another participant stated that he structures content based on groups of customers.

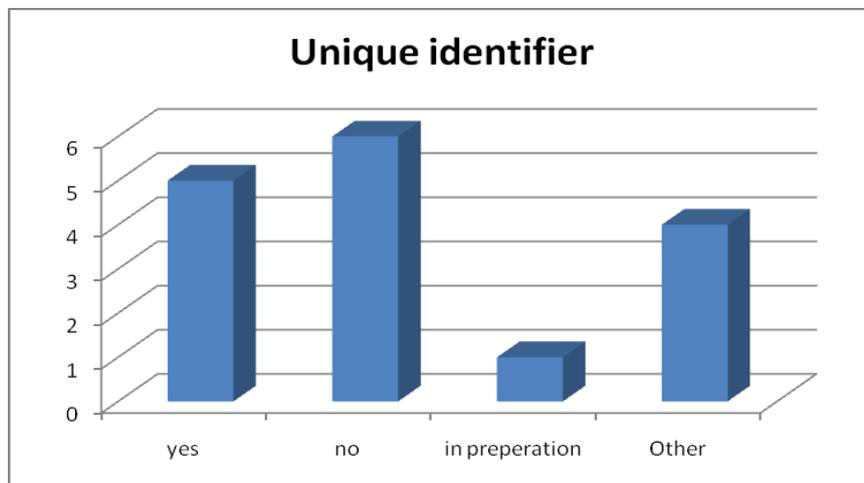
2.4.5 Have you ensured that each element of your content has a world-wide unique identifier?

This question was only asked in the provider questionnaire.

17 answered questionnaires, thereof

No Answers	Option
5	Yes
6	No
1	In preparation
4	Other

Table 7: World-wide unique identifiers in providers dictionaries



Picture 7: Provider Questionnaire – Unique identifier

The outcome of this question shows that indeed 11 of our 16 providers do not provide a worldwide unique identifier for their concepts:

Only 5 providers really offer a unique identifier for their concepts as e.g. has been defined in the ISO 29002-5. However, many answers claim that the provider guarantees a unique identifier within his set of concepts. According to the rules of ISO 29002, this is the fundamental requirement to generate the worldwide unique identifier (in combination with a code identifying provider itself). Because we assume that most of the dictionary providers use an internally unique identifier even if they have not explicitly mentioned it, it should be possible to build this worldwide unique identifier from the existing identifier schemes. Thus, most of the providers can probably support both ISO 29002 web services and product data exchange on the basis of their dictionary in an ISO 8000-1xx compliant way.

2.4.6 Summary of empirical results

To summarize the results of this survey, we can say:

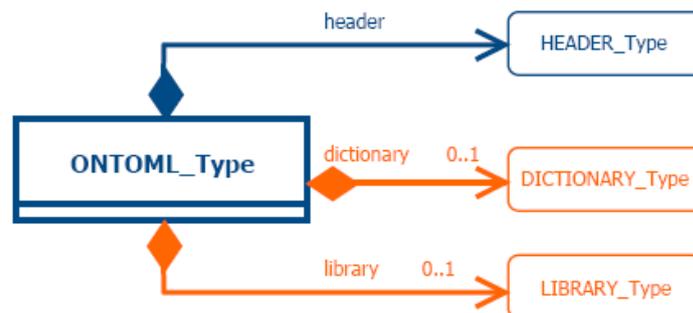
- Most of the providers use a data model to describe the structure of their standards. This is also true for providers of company internal dictionary servers.
- The data model is in almost half of the cases a data model defined in a standard, but in a number of cases, this is a homegrown data model.
- Most providers use spreadsheet tables to specify and distribute their data models. From the responding companies with company internal dictionary servers the majority used the PLIB model as their underlying data model.
- Only a few of the providers state that they provide a worldwide unique identifier for the identification of their dictionary elements. However, most of them provide a locally unique identifier, which could be easily upgraded to a unique identifier.

3 Exchange formats

3.1 ISO 13584-32 (OntoML)

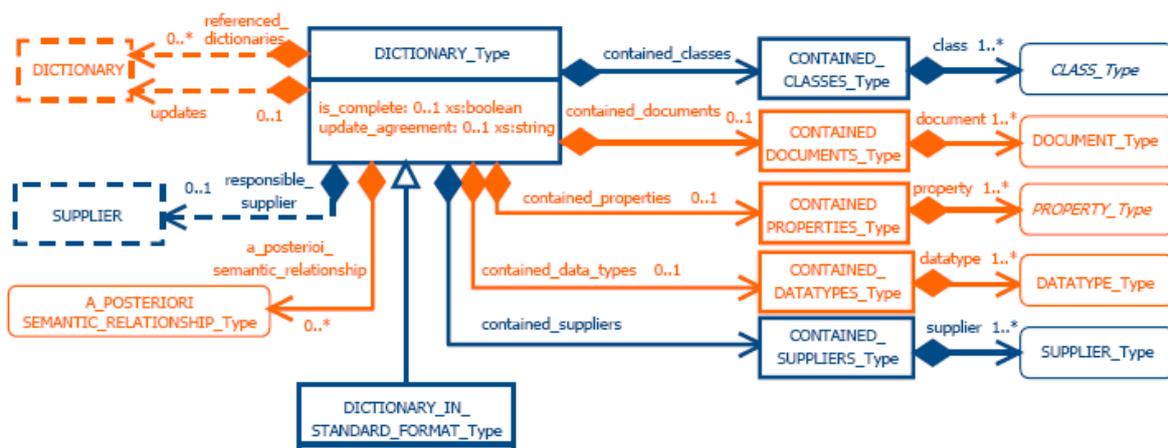
OntoML is an XML representation of PLIB dictionaries and PLIB catalogues. It contains all the modeling elements defined in the PLIB dictionary model in form of an XML schema. The original PLIB model has been created in EXPRESS, the information modeling language defined by STEP in ISO 10303-11. EXPRESS provides more modeling features than XML Schema, so that not all specifications of the EXPRESS model of PLIB can be handled by the XML schema (i.e., if a dictionary representation in OntoML passes an XML checker it might not pass an EXPRESS checker). In the standard it is defined, how OntoML constructs can be mapped to EXPRESS so that it is possible to implement a converter which would allow to convert an OntoML file into an EXPRESS representation and to check it with an EXPRESS checker.

OntoML distinguishes between 2 levels: simple level and advance level. The simple level contains those constructs of PLIB which are used in most application and which are very often sufficient for product data exchange. They cover mainly the features of PLIB described in clause 3.1. Elements of the advanced level are for instance functional class view classes and functional model classes which are used for describing different aspects of a class.



Picture 8: Basic structure of an OntoML file

Basically, an OntoML file has three parts (see Picture 8): A header which gives general information about the file, the dictionary description (which is optional), and an optional library description, i.e. the description of concrete products, e.g. in a catalogue.



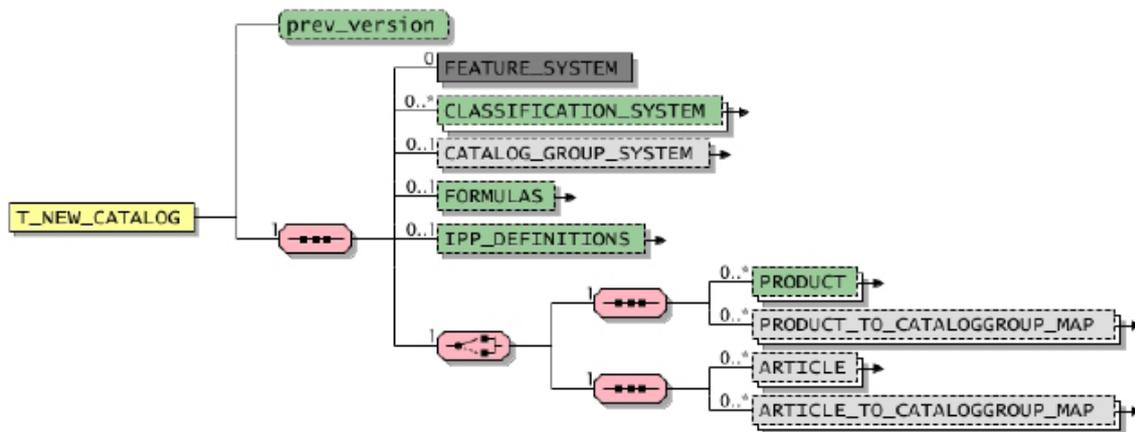
Picture 9: OntoML information about dictionaries

Picture 9 gives an overview about the elements of dictionary which are contained in an OntoML file: It contains classes and properties, additional information contained in documents, information about suppliers of the information and information about externally referenced dictionaries.

3.2 BMEcat

BMEcat is a catalogue exchange format which allows the exchange of product catalogues between business partners. Thus, the focus of BMEcat is on the exchange of product data. But BMEcat provides also elements for the exchange of dictionaries.

Basically, a BMEcat file consists of a header and the catalogue section. The header contains some general information, but it contains also information about supplier and receiver of the catalogue and it allows the definition of default values which are used in the catalogue section.



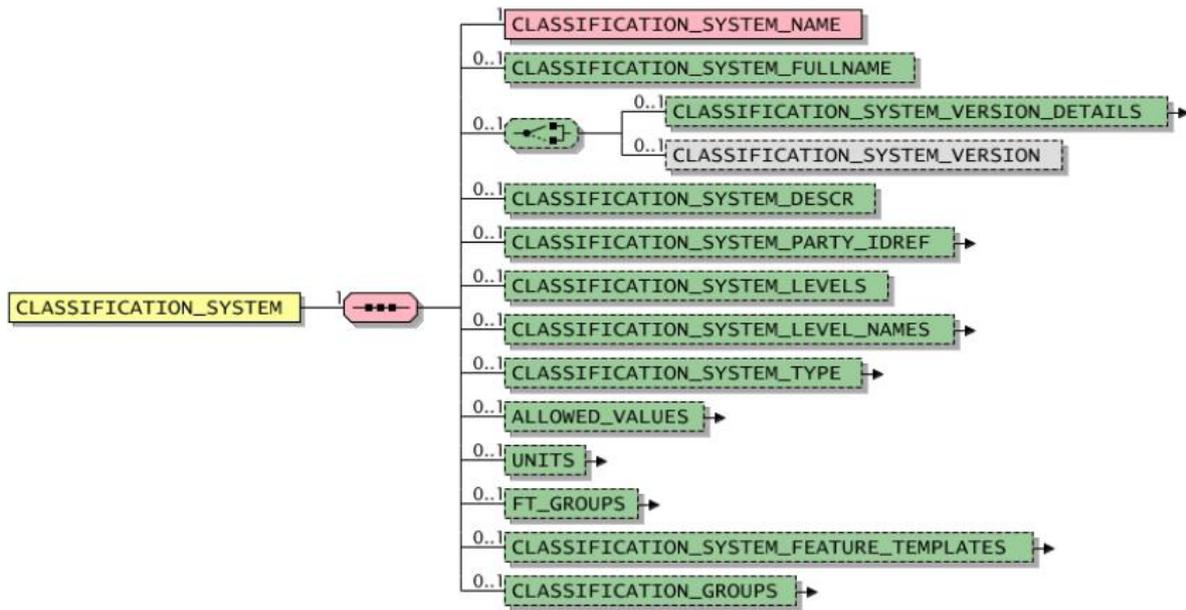
Picture 10: BMEcat catalogue element

Picture 10 shows the structure of the catalogue element in BMEcat. Most important for the normal use of BMEcat as a catalogue exchange format is the product section (or the alternative article section) in which the catalogue products are described. To describe a product, two different kinds of properties can be used:

- Properties which are pre-defined by the BMEcat XML structure. These are mainly commercial properties like prices, delivery information, etc. which are common for all kinds of products and which are hard-coded in the BMEcat XML structure.
- Properties which are defined in a dictionary or "classification". These are properties which normally are specific for the product group and are defined in dictionaries like eCI@ss.

It is possible, to reference the dictionary and the property externally (e.g. by a name or an identifier), but it is also possible to transmit in the BMEcat the dictionary structure which makes the BMEcat file self consistent.

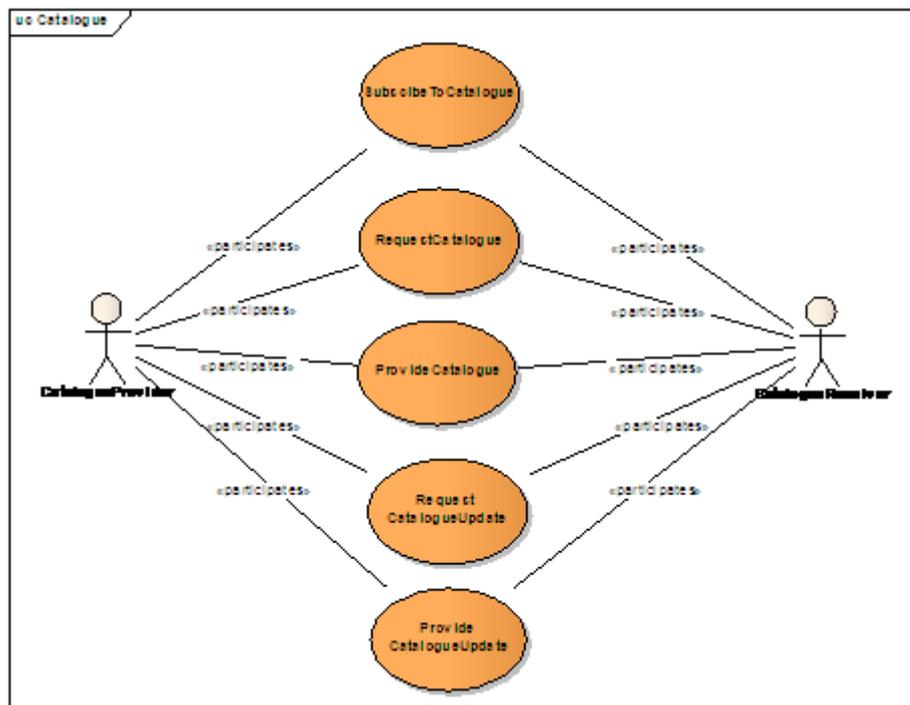
Picture 11 illustrates the top level structure of the classification description in BMEcat. The classification groups are the product groups or classes, and the classification_system_feature_templates define the properties which can be used to characterize the products in the various groups. There are a number of additional possibilities provided to further specify the classification system like allowed values and units used in the classification system.



Picture 11: BMEcat classification structure

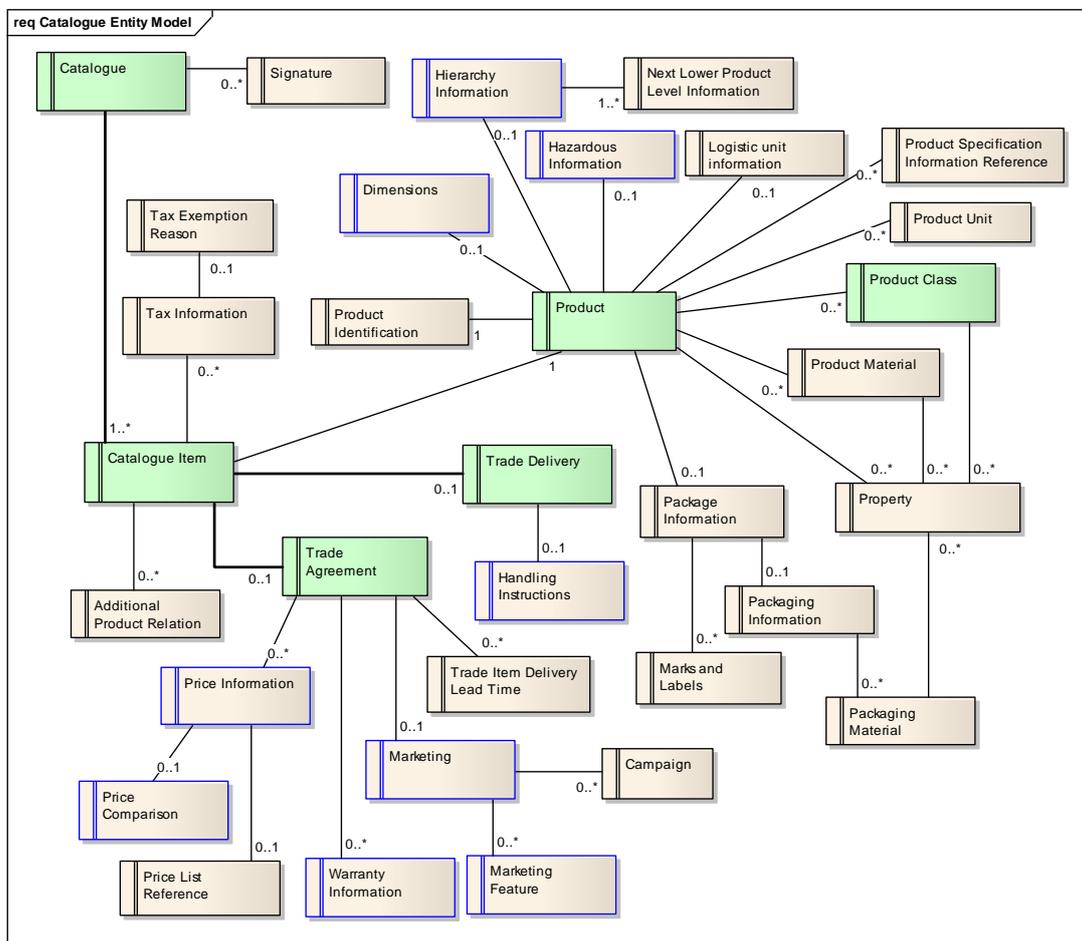
3.3 UN/CEFACT model

The UN/CEFACT is an organization which describes transactions and messages for the transmission of data in e-Business processes between organizations and companies. In TBG1, a business requirement specification has been defined which describes the processes, messages and data which are required for the exchange of catalogue data in different use cases.



Picture 12: Use cases of the UN/CEFACT BRS Catalogue Exchange

The use cases are illustrated in Picture 12. The most important cases are the Request Catalogue and the Provide Catalogue use cases. They describe the interaction of the catalogue requester and the catalogue provider: The requester sends a catalogue request to the provider; the provider decides whether he accepts or rejects the request, and in case of acceptance, he sends back a catalogue, which complies, with the requirements stated in the catalogue request.



Picture 13: Catalogue data model of UN/CEFACT BRS Catalogue Exchange

The model of the catalogue is illustrated in Picture 13. It is a set of catalogue items. The catalogue item describes the required trade information which covers commercial information like delivery information and price information. In addition, to the catalogue item the Product is attached which contains information about the packaging but also references to technical information. By the elements Product Class and Property a reference of external dictionaries is possible, both for classifying a product and for describing a product by means of properties. Thus, similar to the features of BMEcat, it is possible to incorporate property values into an exchange file which are related to classifications systems like UN/SPSC or eCI@ss.

3.4 Referencing dictionary elements in catalogues

In all three described catalogue formats, it is possible to provide values for externally defined properties. The (property-value) pairs do this where the property denotes a reference to the property in the external dictionary which is normally given by some kind of identifier, and the value denotes the actual value for this property. Two key concepts for making use of such a reference structure are (1) the availability of a worldwide unique identifier scheme and (2) a mechanism by which such an identifier can easily and instantly be de-referenced to the actual meaning, i.e. an access mechanism to the description of the concept with all the required information for its use. Such a mechanism is

provided in ISO 29002, and therefore, in the following, we will have a deeper look at this standard. This kind of reference is going to be used in all of the mentioned exchange formats.

3.5 Empirical results from the survey on exchange formats

For the empirical survey, both in the provider questionnaire and in the user questionnaire we asked for the exchange formats used by dictionary servers and the way, how dictionary information is exchanged. The results are laid out in the following subsections.

The following question of the provider questionnaire was used:

- In which format(s) can users receive your content? In which formats can users send proposals for content elements to the server?

In the user questionnaire, the following question was related to this topic:

- In which special data formats do you exchange the dictionary content with your SDO or business partner?

3.5.1 In which format(s) can users receive your content? In which formats can users send proposals for content elements to the server?

This question can be split into four categories:

1. Format for download of human readable information
2. Format for upload of human readable information
3. Format for download for computer readable information
4. Format for upload of computer readable information

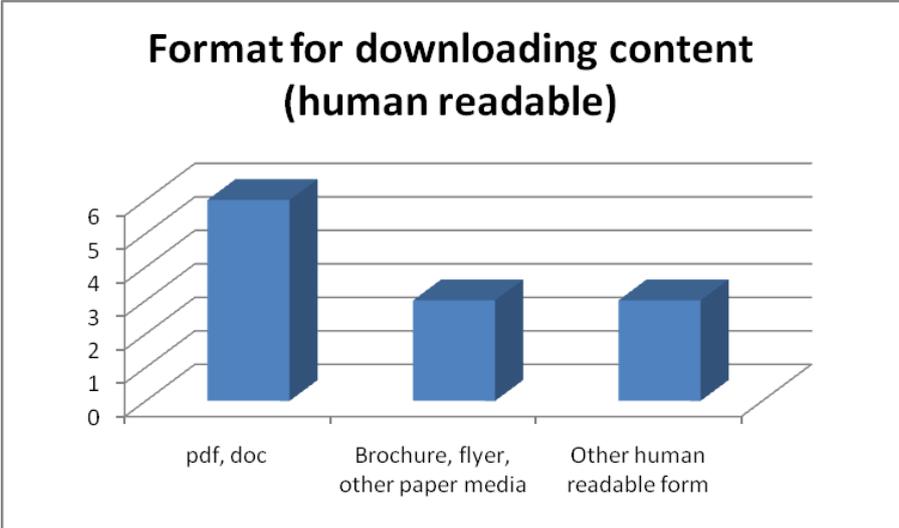
3.5.1.1 Format for download content (human readable)

The result of the survey was:

17 answered questionnaires (not all responded to this specific sub question), thereof

No Answers	Option
6	pdf, doc
3	Brochure, flyer, other paper media
3	Other human readable form

Table 8: Formats for download content (human readable) from provider dictionaries



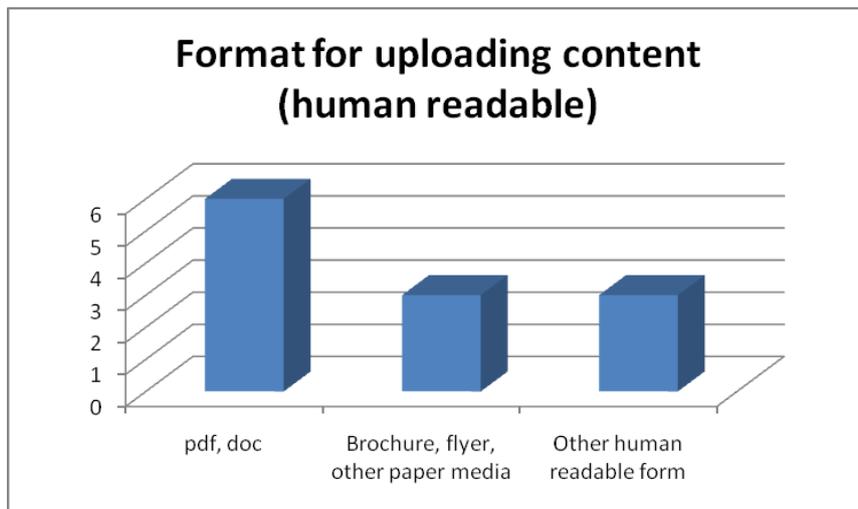
Picture 14: Provider Questionnaire – Formats for download content (human readable)

3.5.1.2 Format for upload content (human readable)

17 answered questionnaires (not all responded to this specific sub question), thereof

No Answers	Option
6	pdf, doc
3	Brochure, flyer, other paper media
3	Other human readable form

Table 9: Formats for upload of human readable dictionary content



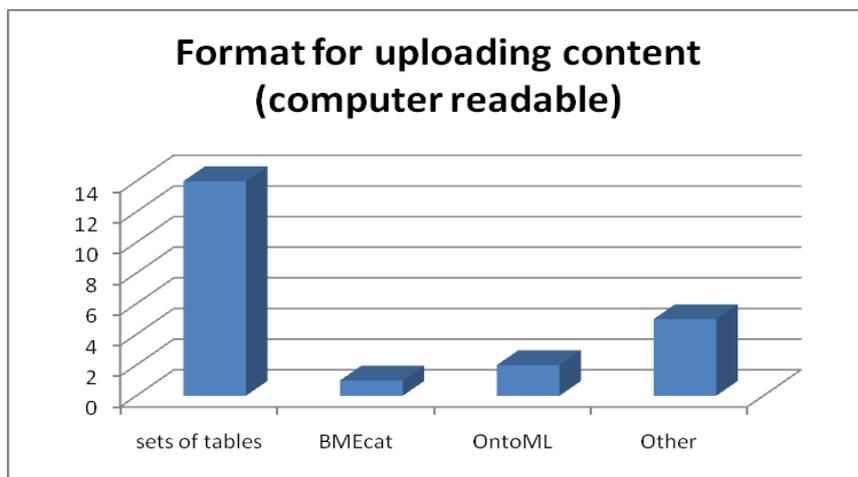
Picture 15: Provider Questionnaire – Format for upload content (human readable)

3.5.1.3 Format for download content (computer readable)

17 answered questionnaires (not all responded to this specific subquery), thereof

No Answers	Option
14	Set of tables
1	BMEcat
2	OntoML
5	Other

Table 10: Format for download of computer readable dictionary content



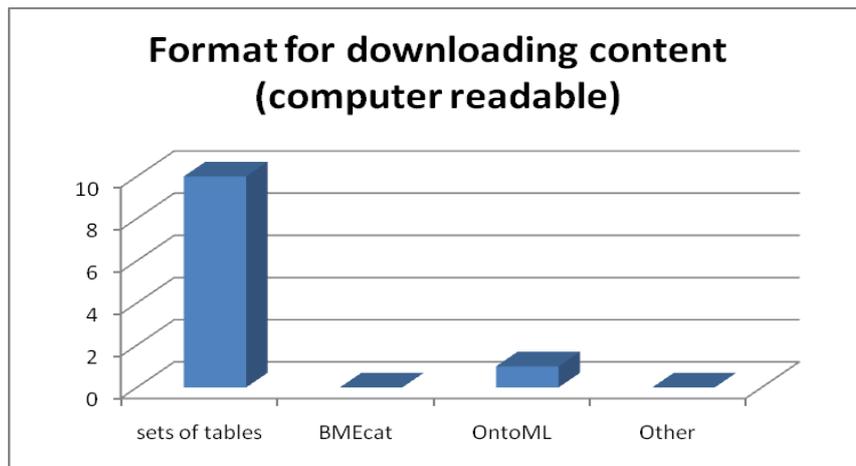
Picture 16: Provider Questionnaire – Format for download content (computer readable)

3.5.1.4 Format for upload content (computer readable)

17 answered questionnaires (not all responded to this specific subquery), thereof

No Answers	Option
10	Set of tables
0	BMEcat
1	OntoML
0	Other

Table 11: Format for upload of computer readable dictionary content



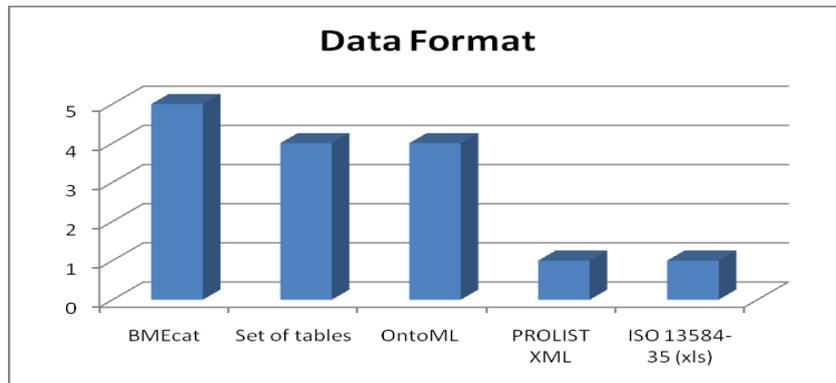
Picture 17: Provider Questionnaire – Format for upload content (computer readable)

3.5.2 In which special data formats do you exchange the dictionary content with your SDO or business partner?

We have contrasted the above provider view with a view of the users in the user questionnaire. In this question the users are asked which data format they use for data exchange of dictionary information with their dictionary providers. Multiple answers were possible. Out of 10 possible answers the following replies were made.

No. of Answers	Option
5	BMEcat
4	Set of tables
4	OntoML
1	PROLIST XML
1	ISO 13584-35 (xls)

Table 12: Formats for exchange of dictionary contents in the user’s view



Picture 18: User Questionnaire - Data formats in use

Half of the users stated that they use BMEcat for exchange dictionary content. With both 40 % replies sets of tables and OntoML were named. One user stated that he uses PROLIST-XML for data exchange and one user named ISO 13584-35, a standardized xls-format as data exchange format.

It seems that in this case many of the responses do not speak about the exchange of dictionary information but that they have spoken about the exchange of product data itself. In particular, the BMEcat answers can be interpreted like that because it is still quite unusual to really exchange dictionary information (apart from simple classification structures) via BMEcat.

3.5.3 Summary of empirical results for exchange formats

Not surprisingly, similar to the description of the data model as a set of tables, also the dissemination of the content is in most cases via tables (XLS or CSV files) possible. This is true for almost any of the dictionaries of our providers. Databases, XML structures or specific formats like BMEcat or OntoML are only rarely used today.

Again, this raises a number of interoperability issues because the different table structures and XML structures provided by dictionary provider are often home grown. They do not comply with standards like OntoML or BMEcat, so they differ in many details from each other, and the user of dictionaries has to take the burden of supporting many different formats. Thus, it has to be considered what can be done from a user's point of view to harmonize the different ways in which the content actually is disseminated.

The picture seems to look different in the perception of the users: even if we have also some users who mainly exchange dictionary content by tables, a number of users specify that they use xml formats like BMEcat or OntoML for the exchange of dictionary data. But this probably results from a misinterpretation of the question by user: Instead of looking at the exchange of dictionary information they looked at the exchange of catalogue data.

4 Multilingualism

4.1 Aspects of Multilingualism

Particularly in Europe, multilingualism is of utmost important for dictionaries to become a useful tool for supporting the language independent exchange of information about products. By identifying a concept by a unique identifier and associating it with names and definitions in various languages, translations are inherent to a dictionary.

Both PLIB and OntoML provide mechanisms for exchanging information in different languages. Two aspects have to be distinguished:

1. Exchange of dictionary information in different languages (e.g. names and definitions of concepts)
2. Exchange of product information in different languages (i.e. values of properties like string values which might be different in different languages, like a short description or a long description of a product).

In PLIB, both preferred names and description may be available in several languages. In addition, synonyms may exist in different languages. Here is a difference to preferred names: Each dictionary element (like a class or a property) has a single name which might be translated. Each dictionary has a list of synonyms, and the specialty is that some of the synonyms normally cannot be translated directly into another language, because they are regional names of the dictionary element which are used only locally, so that there does not exist a 1:1 mapping between synonyms of different languages. For the management of translations, PLIB introduces the “leading language” which has to be identified for any dictionary element and which provides the reference for all translations.

Different languages in product definition values are a different topic which is mainly to be considered in the exchange of product information itself. Again, the various values of a property have to be attached with an indication what language is used for this description.

4.2 Empirical results from the survey on Multilingualism (provider questionnaire)

The following questions from the provider questionnaire are related to this topic:

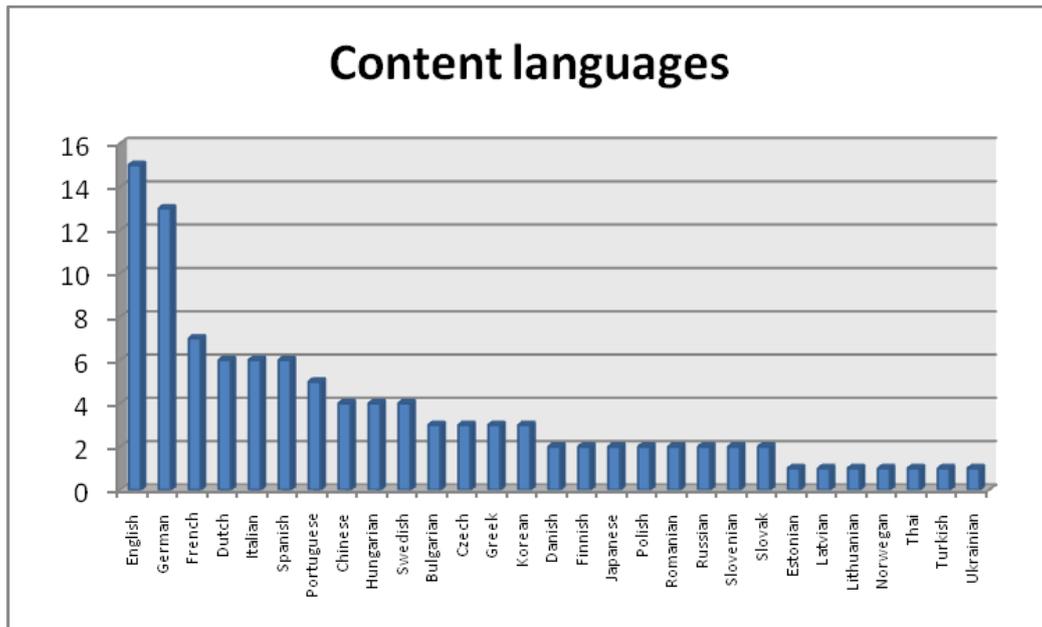
- Do you provide your content in several languages? Which ones?
- Do you provide different language variants? How do handle language variants (e.g. British and American English)?
- Do you have a leading (mandatory) language? Which one?
- Do you distinguish between translatable and non-translatable fields? (*e.g. codes may be not translatable, names may be translatable, please specify your rules*)
- Which kinds of character encodings do you support (e.g. UTF-8,...)?
- Have you encountered localization problems beyond translation? Examples: Need of different structuring of content for different countries / regions / cultures, different properties for different countries / regions / cultures, etc.
- Do you provide your content in a form, which supports users with special needs?

4.2.1 Do you provide content in several languages? Which ones?

It was possible to make more than one statement as each provider may make available several languages. Out of 17 potential answered the following statements were made:

No Answers	Option
15	English
13	German
7	French
6	Dutch
6	Italian
6	Spanish
5	Portuguese
4	Chinese
4	Hungarian
4	Swedish
3	Bulgarian
3	Czech
3	Greek
3	Korean
2	Danish
2	Finnish
2	Japanese
2	Polish
2	Romanian
2	Russian
2	Slovenian
2	Slovak
1	Estonian
1	Latvian
1	Lithuanian
1	Norwegian
1	Thai
1	Turkish
1	Ukrainian

Table 13: Languages provided in the various dictionaries



Picture 19: Provider Questionnaire – Content languages

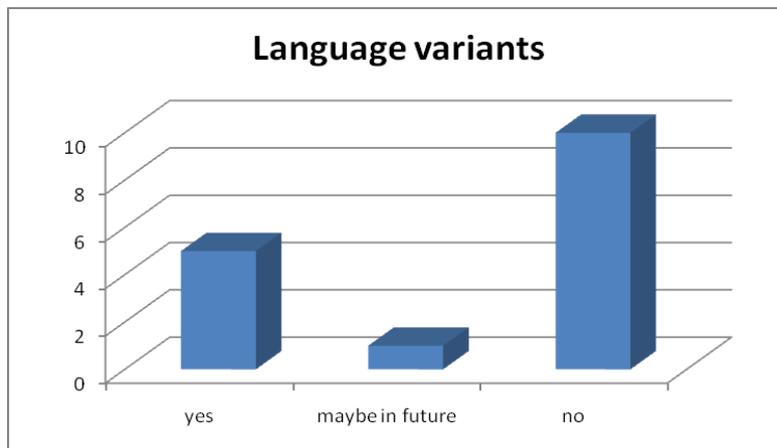
The variety of languages is pretty high. The organizations reported the availability of content in 29 languages. The most common languages for content are English and German. Furthermore content is made available in a wide variety of European languages as well as several Asian languages.

4.2.2 Do you provide different language variants? How do handle language variants (e.g. British and American English)?

Out of 17 potential answered the following statements were made:

No Answers	Option
5	Yes
1	maybe in future
10	No

Table 14: Provision of language variants



Picture 20: Provider Questionnaire – Language variants

Out of the sample of 17 questioned organizations 5 reported that they support different language variants and one may do so in the future. The majority of providers currently do not support language variants. Often American English is the most important English variant, and sometimes different Chinese languages are mentioned (traditional and simplified). Other variants are not mentioned, apart from the CPV which has to serve the EU with its various variants (German and Austrian German, French and Belgium French, Dutch and Belgium Dutch)

4.2.3 Do you have a leading language? Which one?

Out of 17 potential answered the following statements were made:

No Answers	Option
10	English
4	German
1	No

Table 15: Leading languages



Picture 21: Provider Questionnaire: Leading language

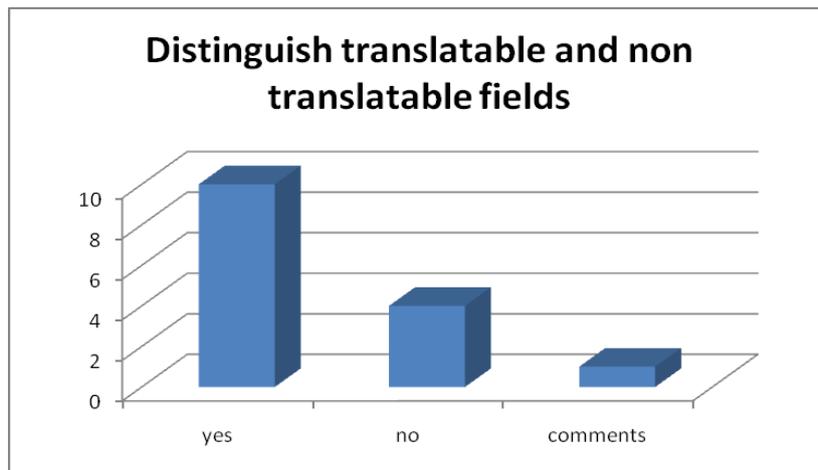
The majority of providers have English as the leading language for content. About a third name German as leading language.

4.2.4 Do you distinguish between translatable and non-translatable fields?

Out of 17 potential answered the following statements were made:

No Answers	Option
9	Yes
4	No
9	Comments

Table 16: Differentiation between translatable and non-translatable fields



Picture 22: Provider Questionnaire – Distinguish between translatable and non translatable fields

The majority of providers distinguish between translatable and non translatable fields in content. The comments specify more detailed how they deal with it and what part of the content can be translated. One organization mentions that rules for translation are in preparation while another one mentions that if no translations are available, the English content is shown.

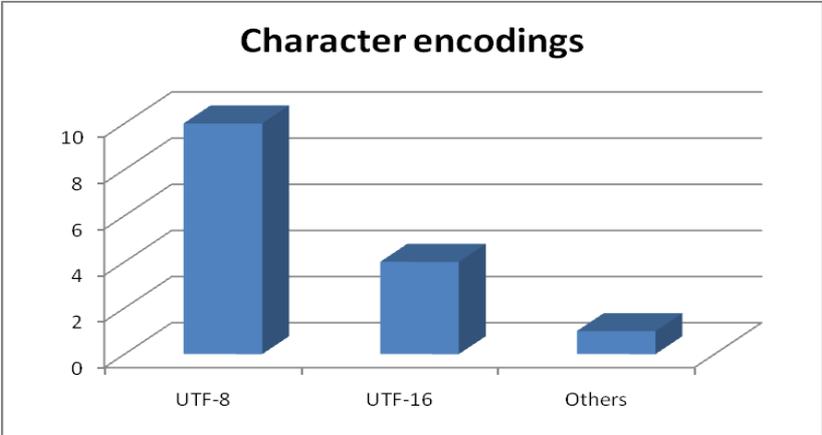
4.2.5 Which kinds of character encodings do you support?

Character encoding often plays an important role because inappropriate character codes may prevent to exchange of data in different languages. Thus, we have asked the providers, which character codes they support. It was possible to make more than one statement as each provider may provide more than one version of character encoding.

Out of 17 potential answered the following statements were made:

No Answers	Option
12	UTF-8
1	UTF-16
5	Others

Table 17: Supported character encodings



Picture 23: Provider Questionnaire – Character encoding

The majority of providers support UTF-8 for character encoding and some provide UTF-16.

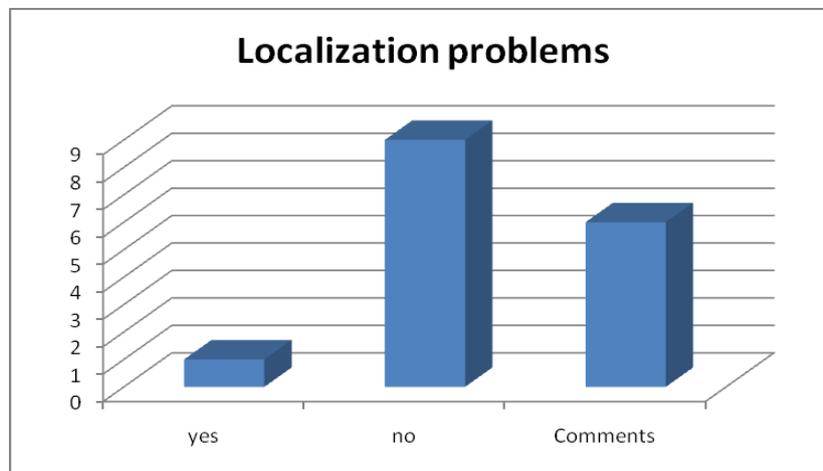
4.2.6 Have you encountered localization problems beyond translation?

We were interested in whether dictionary providers with dictionaries in different languages which are targeted to be used in different countries and cultures have found out special requirements for localization more than providing just translations.

Out of 17 potential answered the following statements were made:

No Answers	Option
1	Yes
9	No
6	Comments

Table 18: Localization problems?



Picture 24: Provider Questionnaire – Localization problems

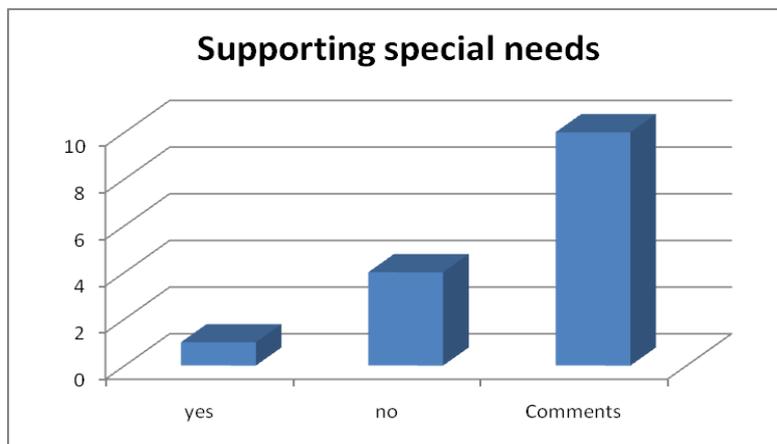
There was only one provider who reported that they encountered localization problems beyond translation. The majority of providers reported that they have not (yet) encounter localization problems. The comments included statements that there is no translation in content, that the provider is aware that there are problems but there is no progress on that so far and that there may be problems in country specific elements in the database.

4.2.7 Do you provide your content in a form, which supports users with special needs?

Out of 17 potential answered the following statements were made:

No Answers	Option
1	Yes
4	No
10	Comments

Table 19: Support for users with special needs?



Picture 25: Provider Questionnaire – Supporting special needs

There was only one provider who states that he supports users with special needs. 4 providers mentioned that they do not support users with special needs. Comments included statements that there shall be support in the future and that additional information like graphical data, hyperlinks and sound files shall be included.

4.2.8 Summary of empirical results on Multilingualism

The survey showed that a big number of languages are offered, but also that only a few of the dictionary providers provide many languages to their users. Most of the dictionaries select one language as leading language, which sometimes also serves as a fallback language to show a value in case that for the expected language no value exists. Most dictionaries distinguish between translatable and not translatable fields. There were no dramatic localization problems encountered other than translation between different languages. On the other hand, some of the providers try to support users, which handicaps or special needs by some means like graphical data and sound files.

Main character encodings are UTF-8 and UTF 16.

5 System Architecture

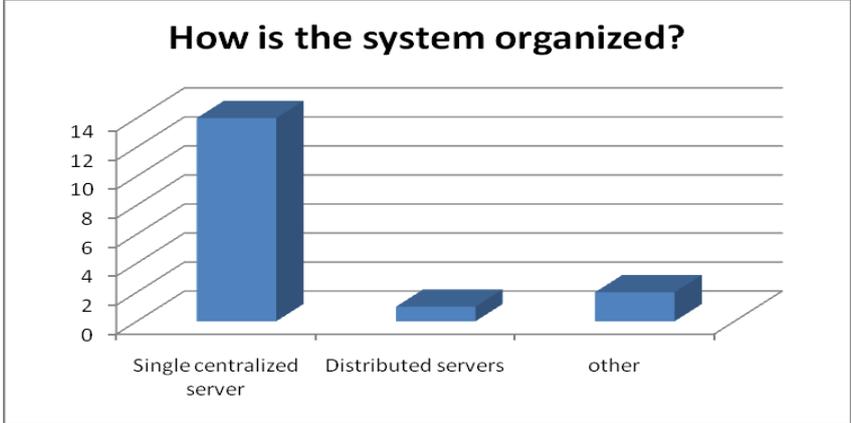
It was our interest to find out, how the dictionary providers organize their server system. In the provider questionnaire, this question deals with organization of the dictionary server.

Question: How is your system organized (e.g. single centralized server, distributed servers, several content collections on one server, etc.)?

Out of 17 potential answered the following statements were made:

No Answers	Option
14	Single centralized server
1	Distributed servers
2	other

Table 20: System architectures of dictionary servers



Picture 26: Provider Questionnaire – How is the system organized

The majority of the systems today is organized as a single server. There is only one reply of providing a system built on distributed servers. One provider has no server at all and one provider has legal requirements, which need a different organization of the system.

6 Dictionary servers, Interoperability and ISO 29002

6.1 Interoperability

One important aspect of dictionaries is the interoperability across different dictionary providers and dictionary models. For users it is important that they can use the same software for exchanging catalogues which refer to different dictionaries and possibly to refer in a single catalogue to different dictionaries. This is not given per se, even in the standardization area of ISO different dictionary data models have been developed which are in many respects not really compatible.

There exist different levels of interoperability, as e.g. defined in the European Interoperability Framework for pan-European eGovernment Services (EIF 2008):

- Syntactic or technical interoperability: Technical issues involved in linking computer systems and services, e.g. open interfaces, interconnection services, etc.)
- Semantic interoperability: Ensuring the precise meaning of exchange information is preserved and well understood
- Organizational interoperability: Alignment of processes by which organizations collaborate
- Legal interoperability: Synchronization of legislation so that electronic information can be exchanged across different countries and legislation systems

For the usage of dictionaries, the goal of the technical interoperability can be summarized as: Ensure that one software system can deal with different dictionaries. This requires, that dictionaries can be exchanged in a well defined manner and that there exist exchange interfaces by which it is possible to exchange dictionary elements of different underlying data models. This is exactly the purpose of the standard ISO 29002.

The provision of semantic interoperability for dictionaries requires to understand how the concepts defined in one dictionary (i.e. the classes and properties) are related to the concepts in another dictionary. This information would facilitate the exchange of data which has been described according to a dictionary also in relation to the other dictionary. Here an alignment of dictionaries is necessary, which can be supported by automatic tools and algorithms (see e.g. Euzenat and Shvaiko, 2007), but where a human decision is necessary to achieve approved mappings. The CEN/ISSS Workshop CC3P [CC3P 2009] is working on such a mapping for the classification structures of UNSPSC (www.unspsc.org), eCI@ss (www.eclass.eu), GPC (<http://www.gs1.org/gdsn/gpc>) and CPV (<http://www.epractice.eu/en/library/281871>).

In the future, semantic interoperability will become more and more important, and the users are keen on getting more harmonized dictionaries. That was the result of one question in the **user questionnaire** where we asked:

If you use more than one ePPS, please mention which kind of interoperability problems do you noticed?

Three users out of the 8 possible users replies on this question.

The problems identified were mapping problems, consistency problems and not full compatibility between the dictionaries.

6.2 Overview of ISO 29002

ISO 29002 has been developed with various goals in mind:

- **Interoperability:** Different dictionary models have been developed, and various dictionaries exist. Therefore it is important to support users who have to deal with several dictionaries simultaneously.
- **Global access to dictionary elements:** Dictionary providers provide information on their dictionaries in various databases – it is difficult for a user to get the right information about dictionary elements when needed. Therefore ISO 29002 describes a common access mechanism to dictionary information which is independent from the data model of the dictionary and provides even a common model for the terminology information about dictionary elements.
- **Identification Scheme for dictionary elements:** One of the backbones of ISO 29002 to achieve the first goals is the provision of an identification scheme which allows to identify dictionary elements uniquely worldwide.

ISO 29002 consists currently of the following parts:

ISO 29002-Part	Name	Description (Abstract of the standard)	Status
4	Basic entities and types	<p>ISO/TS 29002-4:2009 specifies a conceptual information model and a exchange file format for basic entities and types that are used as a resource by other parts of ISO/TS 29002. The conceptual model is in the Unified Modeling Language (UML). The physical data format is based on the Extensible Markup Language (XML).</p> <p>ISO/TS 29002-4:2009 defines types for language code, country code, currency code, uniform resource identifier, and Abstract Syntax Notation One (ASN.1) identifier. It defines an entity for human-readable text that has been localized into one or more languages.</p> <p>The resources in ISO/TS 29002-4:2009 are used by the models in ISO/TS 29002-6 and ISO/TS 29002-20.</p>	ISO Technical Specification
5	Identification scheme	<p>ISO/TS 29002-5:2009 specifies a format for identifiers for elements of a concept dictionary. The format specification is a concrete syntax that meets the general requirements for an international registration data identifier (IRDI) specified in ISO/IEC 11179-5.</p>	ISO Technical Specification
6	Concept dictionary terminology reference model	<p>ISO/TS 29002-6:2009 specifies a conceptual information model and an exchange file format for terminology information from a concept dictionary. The conceptual model is in the Unified Modeling Language (UML). The physical file</p>	ISO Technical Specification

		<p>format is based on the Extensible Markup Language (XML).</p> <p>The XML structures defined in ISO/TS 29002-6:2009 are intended to be returned by calls to the Terminology Service defined in ISO/TS 29002-20. The Terminology Service is an interface for retrieving terminological information about a concept from a concept dictionary.</p> <p>ISO/TS 29002-6:2009 was developed as a common resource to provide a level of interoperability between implementations of ISO 13584 and ISO/TS 22745 for exchange of information from a concept dictionary. Its use is not restricted to these standards, however. It can be restricted through implementation profiles by standards that reference it. It can also be implemented without restriction.</p>	
10	Characteristic data exchange format	<p>ISO/TS 29002-10:2009 specifies a conceptual information model and an exchange file format for characteristic data. The conceptual model is in the Unified Modeling Language (UML). The physical file format is based on the Extensible Markup Language (XML).</p>	ISO Technical Specification
20	Concept dictionary resolution services	<p>ISO/TS 29002-20:2009 specifies an interface to retrieve information from a concept dictionary. Three sets of capabilities, called "services", are defined:</p> <ul style="list-style-type: none"> • Location Service: given a registration authority identifier, returns location on the Internet of Terminology and Ontology servers for the registration authority; • Terminology Service: search for concepts, languages, and other elements of the dictionary based on a rich set of keywords, search criteria, and expressions; given an international registration data identifier (IRDI) in the format defined in ISO/TS 29002-5, retrieve terminological information; • Ontology Service: retrieve ontological information about a concept based on the concept's IRDI. 	ISO Technical Specification
31	Query for characteristic data	<p>ISO/TS 29002-31:2009 specifies a conceptual information model and a exchange file format for a query that requests product and service characteristic data and supplier identification from item and information suppliers. The conceptual model is in the Unified Modeling Language (UML). The physical file format is based on the Extensible Markup Language (XML).</p>	ISO Technical Specification

Table 21: Parts of ISO 29002

In the following, an overview about the Web services defined in ISO 29002-20 is given.

6.3 Requirements and design principles of ISO 29002-20

The requirements which guided the development of the web services can be summarized as follows:

- The web service has to be a distributed system, i.e. there does not exist a central access point where the service is provided, but it is the idea that each provider of a product dictionary can offer his own web service according to this standard. Thus, a service is needed that finds the server which can resolve a given identifier.
- The resolution service should not be bound to a specific product dictionary or a specific product dictionary data model. The service should be implementable and usable for any product dictionary provider, regardless of the used data model and regardless of the implemented business model used to maintain the product dictionary. The web service should provide a level of information which is common to all the different product dictionary providers.
- Currently, product dictionary information is normally only transmitted by product dictionary providers in bulk form, i.e. users can download the whole product dictionary, and if a new version comes out, then again, the whole product dictionary (version) has to be downloaded and somehow integrated into the processing systems. The web service should provide means for more fine grained exchange of product dictionary information, e.g. about a single property or a single class with all its properties.

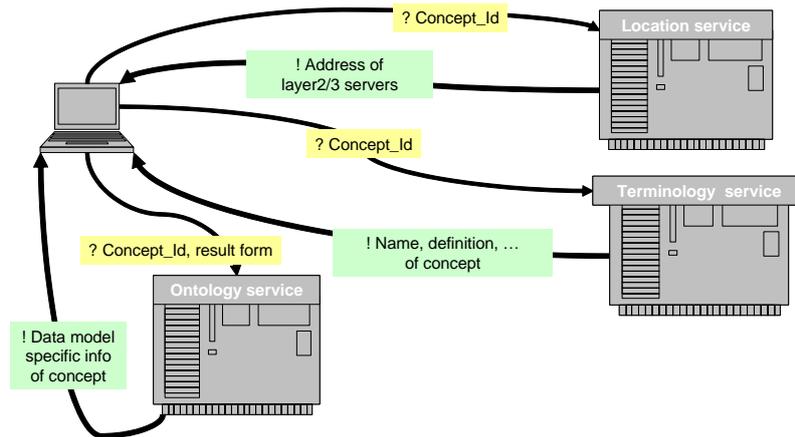
It is expected that such a service will improve the way how product dictionary information is exchanged and maintained:

- By having a "light-weight" procedure to get information on single product dictionary items like properties or classes, processing tools will become more flexible because they can rely on information in the internet and download on demand the necessary piece of information. Thus, the exchange process between product dictionary provider and user will change from a bulk download of the complete product dictionary every year to a steady flow of small information pieces as needed. This may even change the way how product ontologies evolve: instead of providing extensions and corrections only once a year, it might become more appropriate to publish modifications in a product dictionary as soon as they are approved.
- By providing means to get information about different types of product ontologies, the flexibility of tools dealing with product data can be increased. They can pick the information from any product dictionary and thus process product data which is described according to different product ontologies. This will be a boost for the interoperability of product ontologies and allow users to deal more easily with different product ontologies (e.g. when making business with different communities using different ways to organize their product information). It could even be possible to mix different product ontologies for the description of a single product.

6.4 Architecture and usage of the web service

The web service is organised in a 3 level architecture (see Picture 27):

- The location service provides the address of a server which can resolve a given concept identifier.
- The terminology service provides terminological information about the concept reference by the concept identifier.
- The ontology service delivers the complete model specific information about the concept in a model specific format.



Picture 27: Overall architecture of the web service

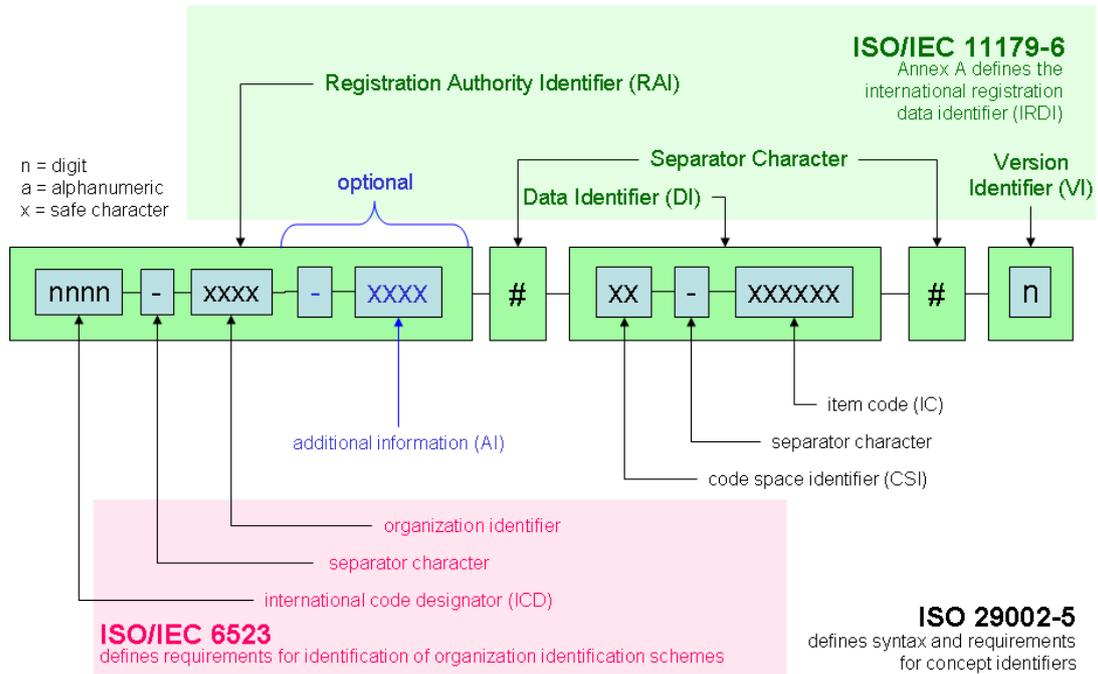
A processing tool, which tries to understand the meaning of a property would first send the property identifier to a known location service. The location service provides addresses of web servers which can resolve the given identifier. The technical basis for the localization service is the structure of the concept identifiers according to ISO 11179: It contains the identifier of the product dictionary provider.

Afterwards, the tool knows which terminology and/or ontology server it can approach to get the requested information. If it is only interested in terminology information like name, definition, synonyms, etc., it will get this information from a server providing the terminology service. ISO 29002-6 contains a terminology model which covers the terminological information found in many product dictionary models (see clause 3.1.3), so that the answer can always be given in the same form – independent of the underlying model of the product dictionary. Thus, on this level true interoperability has been achieved. If the tool needs more information, e.g. about the class hierarchy, then it has to approach an ontology service. Here it will get a data model specific result. In case of a PLIB (ISO 13584) product dictionary this could be e.g. in the form of OntoML (ISO 13584-32) or the spreadsheet format (ISO 13584-35). Thus, the service needs as parameter the model and the format in which the requested information has to be supplied.

6.5 Elements of the Web Service

6.5.1 Location service

The basic function of the location service is to identify a server which provides level 2 or level 3 services which can resolve the given identifier. As such, this service has to deduce the service provider information from the identifier.



Picture 28: Structure of concept identifiers according to ISO/IEC 11179 and ISO 29002-5

This can be done due to the organization of the concept identifier according to ISO 11179: Each identifier consists of three parts (see Picture 28):

- A registration authority identifier which identifies the organization that provides this concept in one of its product ontologies,
- a data identifier which identifies the concept uniquely across all concepts provided by that organization, and
- a version identifier identifying the version of the concept.

This scheme has been further specified in ISO 29002-5, which describes in detail the identification of dictionary elements in ISO 29002.

Thus, the location service can extract the owning organization of the concept from its identifier. Then it needs to have access to a registry where the addresses of web servers of product dictionary providers are registered which offer level 2 or level 3 services. This allows the location service to send the respective addresses to the originator of the query.

6.5.2 Terminology service

The basic task of the terminology service is to provide terminological information to the requester. The terminological information has been defined in a terminology model based on various terminology standards (as described in clause 3.3). Elements of this model are e.g. term, definition, graphical element, abbreviation, symbol, document, etc. which can be presented in different languages.

Based on this terminology model, a number of functions have been defined, for example:

- Retrieval of complete concepts, i.e. of all its attributes
- Retrieval of specific concept attributes like names, definitions, etc.
- Search for concepts

To connect a web service to a product dictionary, the provider has to build a mapping of his terminology model to the ISO 29002-6 terminology model. In ISO 29002, the terminology model has been designed with the goal to make this as easy as possible for product dictionary providers, and we believe that the entities contained in the model are widely used in product dictionary models. It has been verified for ISO 22745 and ISO 13584 that their terminology elements can be mapped to the ISO 29002-6 model.

6.5.3 Ontology service

The ontology service delivers the complete information about a concept which is available in the product dictionary. The requester has to specify in which format this information is to be delivered, e.g. in case of a PLIB based product dictionary this could be OntoML, the XML representation of ISO 13584.

It is clear that with different data models different information about a concept can be delivered. To allow for a specification of the "query" in a generic form, ISO 29002-20 provides only a generic function with a generic parameter. The proposed mechanism is to base the query methodology on a generic graph concept of dictionaries (edges representing any kind of relationship) and only identify explicitly the "is-a" relationship as a known concept. Then the queries can identify a node (class, property, unit, etc.) and specify how many "relationship layers" around this node have to be returned.

6.6 Conclusion, status

The ISO 29002 web service for resolving concept identifiers will

- support the common use of product ontologies based on different product dictionary models by providing a common terminology exchange platform and providing common (but semantically specialized) ontology functions,
- change the way how product dictionary information is brought into application systems from a bulk data exchange to a fine grained exchange of necessary information on demand,
- allow different ways how product ontologies can evolve: It supports the release process, where a release is published every year (or in longer or shorter periods) as well as the publication of sub-releases in shorter intervals down to the continuous publication of approved modifications and extensions.

The standard ISO 29002-20 is being developed in ISO TC184 SC 4. It has been published as a Technical Specification in 2009, and it is further developed and promoted to an IS probably in 2010.

In parallel to the standardization work, first implementations have already started by DIN for their DIN server and by ECCMA for their eOTD. An integration into a catalogue generation program has been done in a DIN-INS project funded by the German Ministry of Economics [Wilkes+Pohn 2008].

6.7 Recommendations for dictionary servers

For dictionary servers, it is important to support the mechanisms which are defined in the ISO 29002 standards. Since ISO 29002 is a standard which is supported by different standardization groups and which has the goal to unify the access to dictionary elements independently of the data model used by dictionary providers, it is expected that this standard will be heavily used in the future. Thus, dictionary providers should ensure that their users can reference dictionary elements in their catalogues and libraries by ISO 29002 compliant identifiers and that they can resolve these identifiers and obtain the information about their underlying concepts by using the interfaces and the web services defined in ISO 29002.

Thus, our recommendations are:

Recommendation 1:

Dictionary server providers should ensure that their dictionary elements can be identified according to the identification scheme specified in ISO 29002-5.

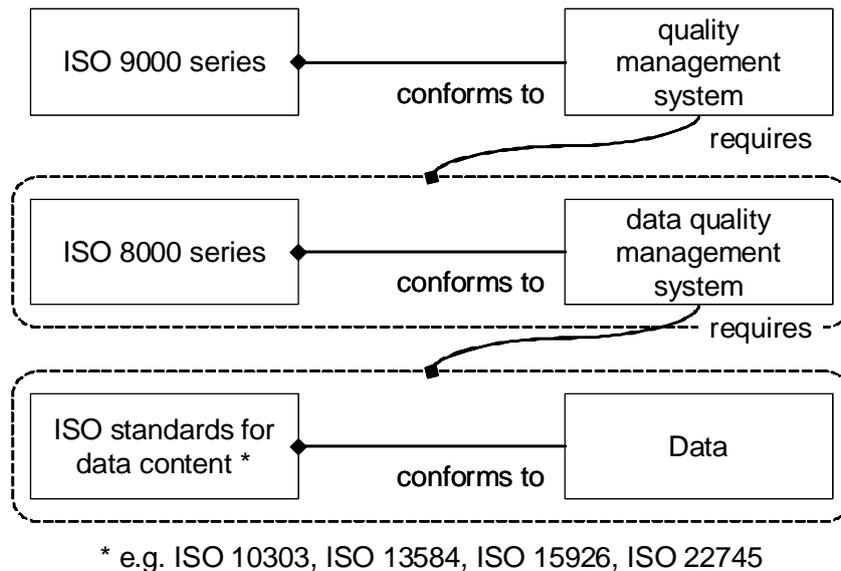
Recommendation 2:

Dictionary server providers should provide web service functions as defined in ISO 29002 to ensure that the identifiers of their dictionary elements can be resolved by software tools. This includes a mapping of their dictionary model to the terminology model of ISO 29002-6 and the provision of the functions defined in ISO 29002-20.

7 Requirements placed by ISO 8000 standards on the operation of dictionary servers

7.1 Overview on ISO 8000

ISO 8000 is an upcoming standard about "data quality". It has been initiated by the US department of defense (DOD), in particular DLIS together with ECCMA for the area of catalogue data exchange, and its scope has been extended by subcommittee 4 of ISO TC184 to cover general data quality aspects in the context of technical data exchange.



Picture 29: The data quality stack

Picture 29 shows the "quality stack" which is supposed to illustrate the relationship of ISO 8000 to ISO 9000 and to the technical product standards: ISO 8000 builds on the product standards which define models and formats to exchange data and it adds additional requirements on the usage of these standards to ensure a high quality of the exchanged data. On the other hand, ISO 9000 cannot really achieve its goal if it cannot rely on data quality. As such, ISO 8000 fills a gap between ISO 9000 and the data handling standards.

Of most interest for ePPS is the ISO 8000 – 100 series which describes "master data quality". It comprises the following parts (with the current status of standardization):

ISO 29002-Part	Name	Description (Abstract of the standard)	Status
100	Master data: Overview	<p>This part of ISO 8000 describes fundamentals of master data quality, defines related terms, and specifies requirements on both data and organizations to enable master data quality.</p> <p>The following are within the scope of this part of ISO 8000:</p> <ul style="list-style-type: none"> specification of the scope of the master data quality series of parts of ISO 8000; 	ISO TS

		<ul style="list-style-type: none"> • introduction to master data; • description of the data architecture; • overview of the content of the other parts of the series. 	
102	Master data: Exchange of characteristic data: Vocabulary	ISO/FDIS 8000-102:2009 contains the vocabulary relating to master data quality, for use in connection with the master data quality series of parts of ISO 8000. ISO/FDIS 8000-102:2009 includes those terms and definitions referenced in published parts of ISO 8000 dealing with master data quality. It is therefore not a comprehensive lexicon for master data quality. This edition of ISO/FDIS 8000-102:2009 includes terms and definitions related to syntax and semantics of master data messages, terms and definitions related to characteristic data and data provenance, and terms and definitions related to items of production and items of supply.	ISO FDIS
110	Master data: Syntax, semantic encoding, and conformance to data specification	ISO 8000-110:2009 specifies general, syntax, semantic encoding and data specification requirements for master data messages between organizations and systems. The focus of ISO 8000-110:2009 is on requirements that can be checked by computer.	ISO TS, DIS ballot
120	Master data: Exchange of characteristic data: Provenance	ISO/TS 8000-120:2009 specifies requirements for capture and exchange of data provenance information and supplements the requirements of ISO 8000-110. This part of ISO 8000 includes a conceptual data model for data provenance.	ISO TS
130	Master data: Exchange of characteristic data: Accuracy	ISO/TS 8000-130:2009 is an optional addition to ISO/TS 8000-120 and specifies requirements for representation and exchange of information about accuracy of master data that consists of characteristic data.	ISO TS
140	Master data: Exchange of characteristic data: Completeness	ISO/TS 8000-140:2009 is an optional addition to ISO/TS 8000-120 and specifies requirements for representation and exchange of information about completeness of master data that consists of characteristic data.	ISO TS
150	Master data: Quality management framework	This part of ISO 8000 specifies a framework that can be used to measure and improve master data quality. This framework can be used in conjunction with ISO 9000 or independently.	NWI

Table 22: Parts of ISO 8000

The most important part of this 100 series is the Part 110. Its basic requirements can be summarized as follows:

- Basically, ISO 8000-110 assumes that the underlying exchange schema is the (property, value) pair, i.e. a value which is associated with a description of the meaning of the value.
- Each exchange of catalogue data has to contain all information which are necessary to understand the meaning of the data. This means, for the receiver of an information it must be possible to get information of the concept which describes the meaning of the value. This can be done either by
 - Adding the dictionary information to the exchange file, or by
 - Having uniquely resolvable references to the concept in some dictionary which can be accessed automatically by software tools.
- Product data has to be exchanged according to a formally described syntax (=exchange format) the description of which is openly accessible.
- There must be no copyright problems for users who integrate references to some dictionaries into their catalogues: The dictionary providers have to take legal means to ensure that the integration of references to their dictionary elements does not create a common work of the dictionary provider and the catalogue creator.

7.2 Recommendations for dictionary server providers

The short description of ISO 8000's 100 series shows that the use of dictionaries is an important corner stone of ISO 8000. Basically, ISO 8000 requires that dictionary servers provide means to access the dictionary information either by a complete download or by specific access to selected dictionary elements, for instance via Web Services as described in ISO 29002. Furthermore, ISO 8000 requires that dictionary providers identify their dictionary elements by an unambiguous and world-wide unique identifier scheme.

Thus, if a dictionary provider follows the ISO 29002 requirements, it fully complies to the technical ISO 8000-110 requirements. Thus our recommendation is:

Recommendation 3:

To be fully compliant to the technical requirements of ISO 8000-110, dictionary providers should act according to recommendations 1 and 2 and follow the specifications of ISO 29002, which provide an implementation path to ISO 8000-110 compatibility.

8 Standards and formats used in HVAC industry

One aspect of the ePPS project is the demonstration of the concepts and recommendations developed in this CWA in the context of the HVAC industry. As a basis for this demonstration, a short overview of the German standard VDI 3805 will be given which is currently being transferred to an ISO standard in ISO TC 59.

VDI 3805 is a standard which allows planning systems in the building services industry to receive information about devices directly from the manufacturers of the devices. For this purpose, VDI 3805 provides a standardized data structure which captures the fundamental information necessary for describing devices and which can be extended and adapted specifically for concrete product classes. Thus, the entire general and the product class specific properties have a place in this data structure. In addition, a number of further descriptive elements can be transferred:

- complete technical data in standardized form
- complete verbal description as a basis for announcements
- complete geometric depiction in 3D for direct use in CAD systems
- complete photographic depiction of products
- complete product hierarchy for the product search process

The basis of VDI 3805 is a hierarchical data structure which describes a composition hierarchy. Its purpose is the definition of variants of products: By selecting some of the listed potential components or by selecting some property values, a specific product can be defined. Thus, it can be seen as a basic data structure describing the information which is necessary to configure a product from a number of components and properties.

VDI 3805 is targeted to the small and medium companies which are active in the HVAC industry. Thus, all information is represented in a very simple exchange format based on an ASCII format, and for the representation of the various elements simple methods were chosen. For instance, the graphical elements are built on 2D and 3D primitives that can be combined according to simple rules, and functions can be defined by a simple subset of the Fortran language.

A big number of companies uses VDI 3805 in Germany and it has been incorporated into a number of tools so that users can directly make use of the exchange capabilities of VDI 3805.

It is the goal of ISO 3805 to rely on properties which are available from standardized dictionaries. As such, the definition of the specific HVAC related properties in a dictionary like eCI@ss is of interest for the HVAC industry, and part of the demo will show the definition of such a properties and their automatic use in applications for product data exchange by use of access mechanisms like ISO 29002 web services. Such a demonstration will be shown at the final ePPS conference in November in Rotterdam.

9 Summary and conclusion

Part 1 of the CWA has given an overview about technical aspects related to dictionary servers, focusing on data models and exchange formats and ISO29002 for interoperability and ISO 8000 for data quality. It has become clear that dictionary providers who follow the requirements of ISO 29002 give their users a lot of advantages:

- Users can access their dictionary elements on demand via Web services
- Users need only to support one interface / web service functionality to access different dictionary servers
- Users get a unique identification of the dictionary elements (which might even lead to catalogues which describe products according to several dictionaries)
- Users can more or less automatically become ISO 8000 compatible because the basic requirements (formal exchange format, unique identification of dictionary elements) are fulfilled by ISO 29002.

Thus, the basic recommendation of this Part of the ePPS-CWA is that dictionary server providers should implement the ISO 29002 requirements in their dictionary servers.

Part 2 Organizational aspects of operating an ePPS

10 Electronic metadata dictionaries

10.1 Introduction

Today there are already several electronic metadata dictionaries available on the market. The number is increasing steadily and especially content-wise there are a lot of ongoing activities. The industry is supporting these activities but the deployment-ratio still has to increase. Awareness about the need and the advantages of reliable electronic metadata dictionaries is also increasing steadily. Industry currently is still heavily using „home made“ solutions. Own metadata dictionaries are built up in company internal “master data” projects.

Within the last years we see an increasing number of national and international projects on electronic metadata dictionaries being made available or are under development. National and international standards development organizations (SDOs), industry associations and companies are developing content and using it within electronic business processes.

10.2 Problems

Industry and trade partners recognize the need for structured and standardized electronic metadata. Companies and organizations are heavily investing in the development and implementation of content. Within these projects they are facing several challenges and problems.

- No common processes and workflows for developing and maintaining electronic metadata are available
- Content is made available under various licensing agreements and business models (free of charge, commercial licenses, limited licenses, ...)
- Handling of IPRs for content is difficult as there are many national laws applicable both for the provider of content and the (potential) user

10.3 Consequences

Facing these issues and problems in implementing structured and standardized electronic metadata may result in the following consequences

- Potential users are unconfident and confused and may decide to not implement at all
- „Home-made solutions“ are preferred as there are no (or less) potential legal risks involved in using electronic content
- Costs and complexity of data exchange to customers and suppliers are still extremely high and need a lot of resources

11 Workflows

11.1 Introduction

A specified group of people is developing the content of electronic metadata dictionaries. Depending on the definitions made by the publisher or owner of the content this group can be a closed community where membership to the community is a requirement for participation in content development. An example for such a closed community is ISO where only members (the national standardization bodies) may participate in the development of content. Content development can also be carried out in an open community where every interested party may contribute to the content. An example for an open community is eCI@ss where everybody can enter change requests to create a new element, change or withdraw an existing one. In both cases, it is necessary to define rules and processes how new content is being developed, existing content is maintained, outdated or void content is withdrawn and the publication of content is organized.

In this chapter we will look into some workflows currently used to develop, maintain and publish content. We will learn about the results of the provider questionnaire how the dictionary providers currently support workflows, which standards they have implemented and for what operations they supply support functions to their users. Providers also replied on how they are dealing with content modifications and dictionary users were questioned how they are dealing with content updates made available by the dictionary providers.

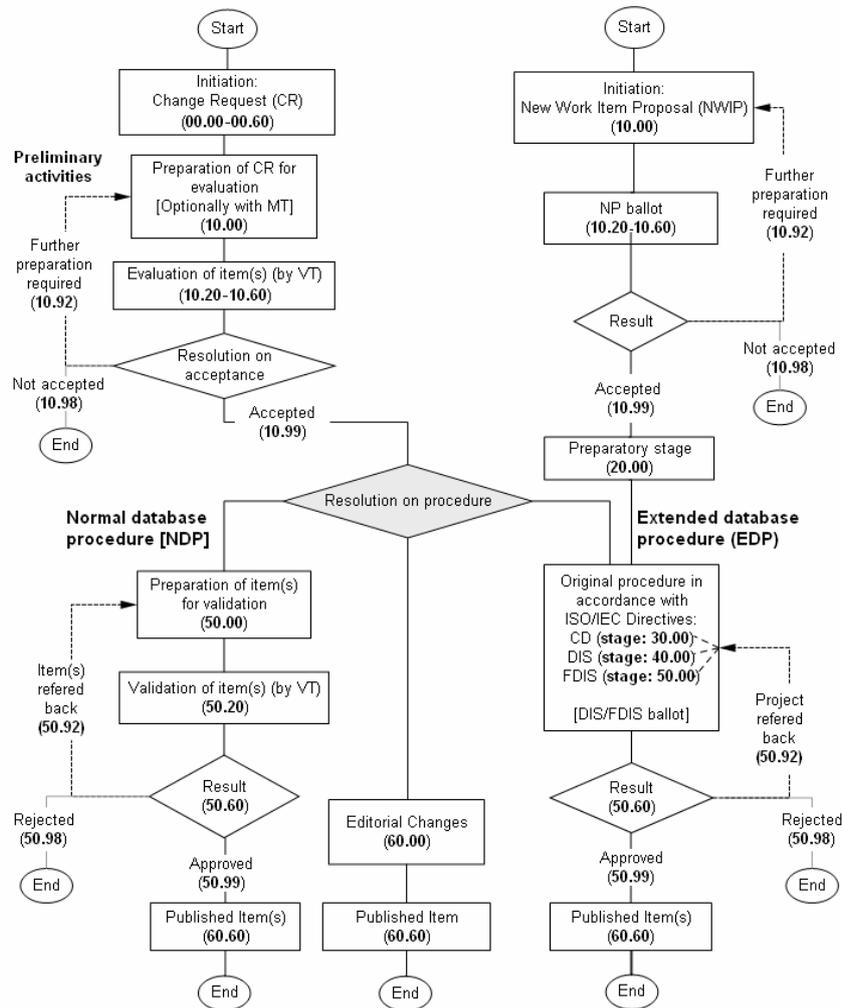
11.2 Status quo

When we talk about workflow in the context of electronic metadata dictionaries, we usually mean the workflow to develop, release and maintain metadata within the electronic dictionaries. Content offered in electronic metadata dictionaries is developed by a group of interested parties who want to create standardized content, which may be reused by the users of the dictionary.

Currently we see many different workflows for dealing with content development. Each organization defined its own workflows depending on the structure of the organization and the target group of the content. Organizations with a long history in workflows are national and international standardization bodies.

One example for a workflow is Annex ST of the ISO supplement to the ISO/IEC Directives [ISO Annex ST], which describes a procedure for the development, maintenance, review, and withdrawal of any international standard consisting of “collections of items” managed in a database. Such items may include graphical symbols of all kinds, sets of definitions, sets of dimensions, dictionaries of data element types with associated classification schema and other standards in which collections of objects require maintenance (addition or amendment) on a continual basis. In compliance with the ISO/IEC Directives, the development of a new standard requires a new work item proposal (NP), whereas an NP is not needed for the maintenance of an existing standard.

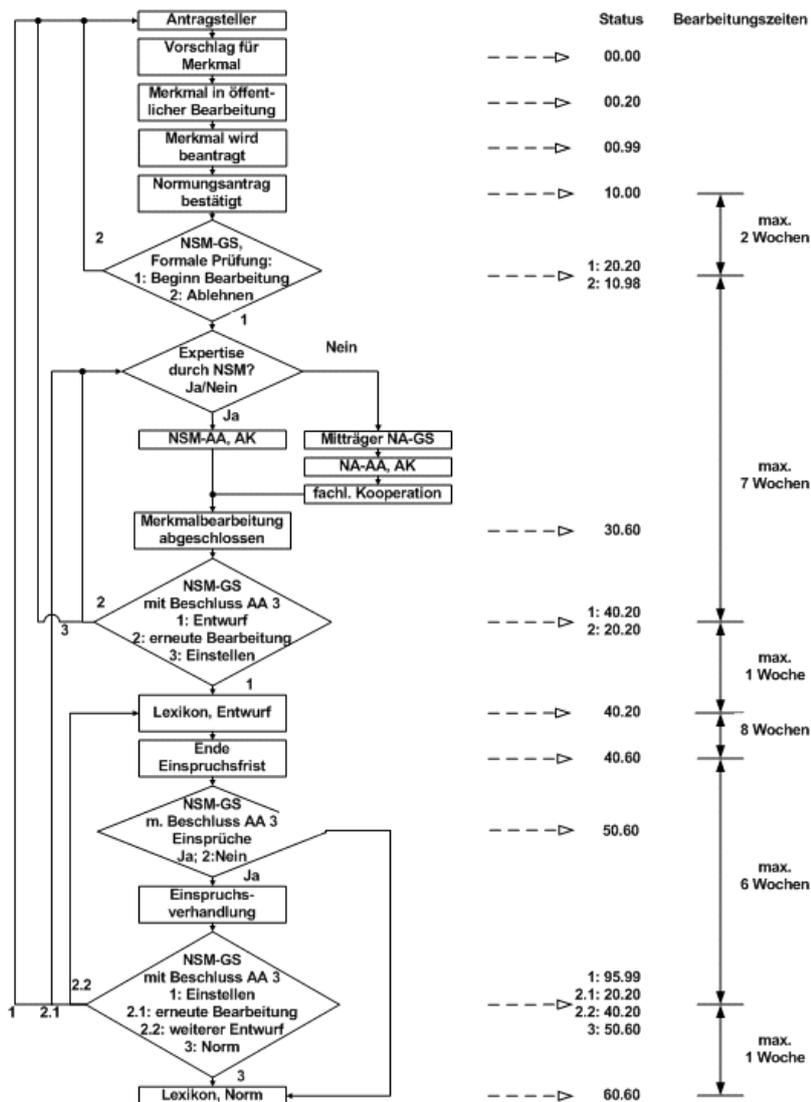
Below picture shows the workflow for the development of new and the maintenance of existing standards



Picture 30: Workflow for the development of new and the maintenance of existing standards from Annex ST of ISO/IEC Directives

The ISO/IEC workflow shows two options to develop a standard. There is the Normal Database Procedure (NDP) and an Extended Database Procedure (EDP).

Another example for an electronic workflow is shown below for the DIN Property Dictionary (www.DINsml.net) the online dictionary of product properties from the German national Standardization body DIN e.V. This workflow is defined in DIN 4002 part 100 [DIN 4002-100].



Picture 31: Workflow for the development of content in the DIN Property Dictionary (www.DINsml.net)

The workflow for elements developed and maintained within the DIN Property Dictionary also shows maximum durations for the single stages. Both the DIN and ISO/IEC workflow refer to ISO Guide 61 [ISO Guide 61] which defines the principles and usage of the Harmonized Stage Code System. Both examples may deal with an ongoing release process of new content. Other dictionary providers like eCI@ss or UNSPSC have own workflows where content is published in fixed releases (e.g. twice a year).

11.3 Analysis of current dictionaries implementations (provider)

Many of the organizations developed an own workflow which reflects the size and structure of the organization. In the provider questionnaire we asked the content providers about the availability and the implementation of workflows.

In the provider questionnaire the following questions regarding workflow were included:

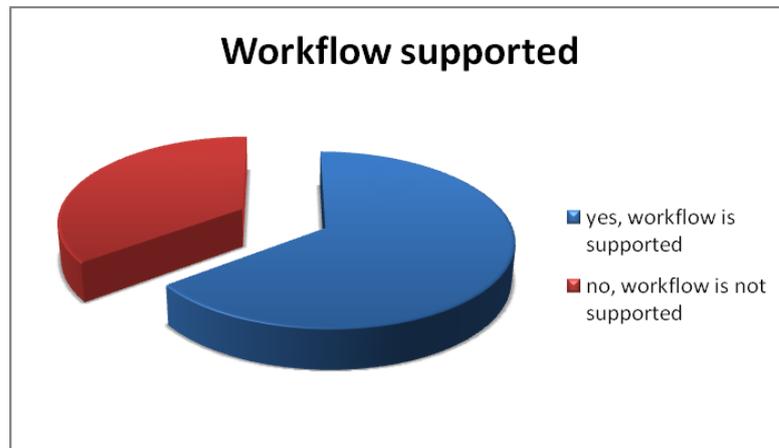
- Do you support a workflow for development and maintenance of content?
- What kind of workflow is implemented?
- Is the maintenance procedure supported by your ePPS system?
- Do you provide means for upload of proposals for new content elements?
- How do you deal with modifications of your content?
- Do you have a fixed or scheduled release cycle where new versions of the content are published?
- If you publish your content in a release cycle, how often do you provide new versions of your content?
- If you do not have a release cycle, how do you publish your content?
- Which form do you publish your content?
- In which granularity do you provide your content?
- Which additional information do you provide for download?
- Do you provide means for community building?
- Do you provide additional services for your users or your committee members?

11.3.1 Do you support a workflow for development and maintenance of content?

In this question the dictionary providers are asked to state whether they support a workflow for development and maintenance of content. Only one answer was possible. Out of 17 possible answers the following replies were made

No. of Answers	Option
11	yes workflow is supported
6	no, workflow is not supported

Table 23: Workflow supported



Picture 32: Workflow supported

The analysis of the results shows that the majority of the providers currently support a workflow for development and maintenance of the dictionary content. Depending on the answer in this question (yes, workflow is supported/no, workflow is not supported) the providers were requested to provide further details about their implementations.

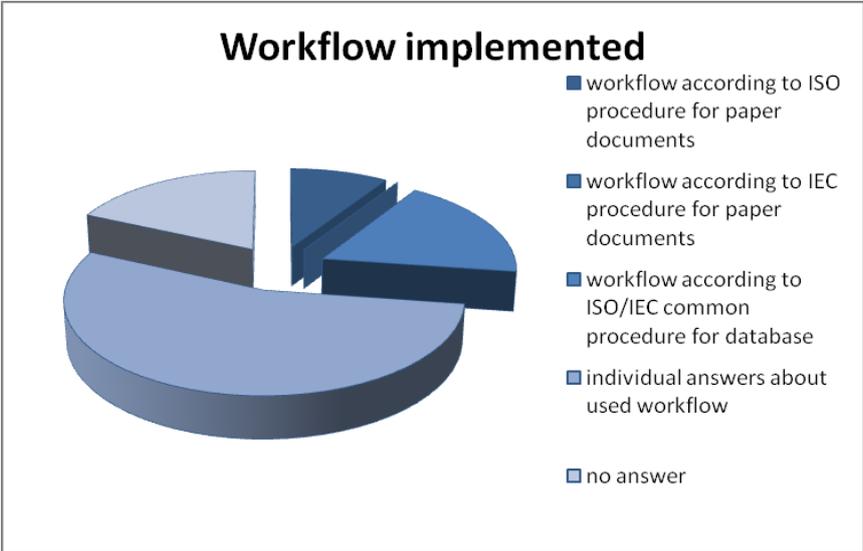
The following three questions are feedback from providers of dictionaries who answered that they support workflow for development and maintenance of content.

11.3.2 What kind of workflow is implemented?

In this question the dictionary providers are asked what kind of workflow they have implemented. This question is only relevant for those providers who answered that they supply workflows. Out of 11 possible answers the following replies were made

No. of Answers	Option
1	workflow according to ISO procedure for paper documents
0	workflow according to IEC procedure for paper documents
2	workflow according to ISO/IEC common procedure for database
6	individual answers about used workflow
2	no answer

Table 24: Workflow implemented



Picture 33: Workflow implemented

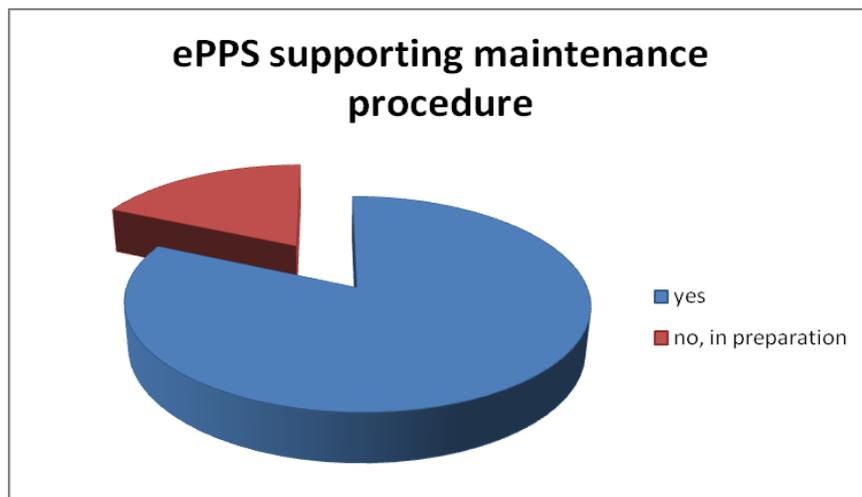
The analysis of the results shows that one provider supports workflow according to ISO procedure for paper documents, two provider support workflow according to ISO/IEC common procedure for database and the majority of the providers replied with individual answer about the workflow supported. These answers referred to specific workflows for the respective dictionary as well as to ISO procedures for paper document, and to ISO 22745 [ISO 22745]. 3 dictionary providers mentioned DIN 4002 as the basis for the workflow specification.

11.3.3 Is the maintenance procedure supported by your ePPS system?

In this question the dictionary providers are asked if the maintenance procedure is supported by their ePPS system. This question is only relevant for those providers who answered that they supply workflows. Out of 11 possible answers the following replies were made

No. of Answers	Option
9	yes
2	no, in preparation

Table 25: ePPS supporting maintenance procedure



Picture 34: ePPS supporting maintenance procedure

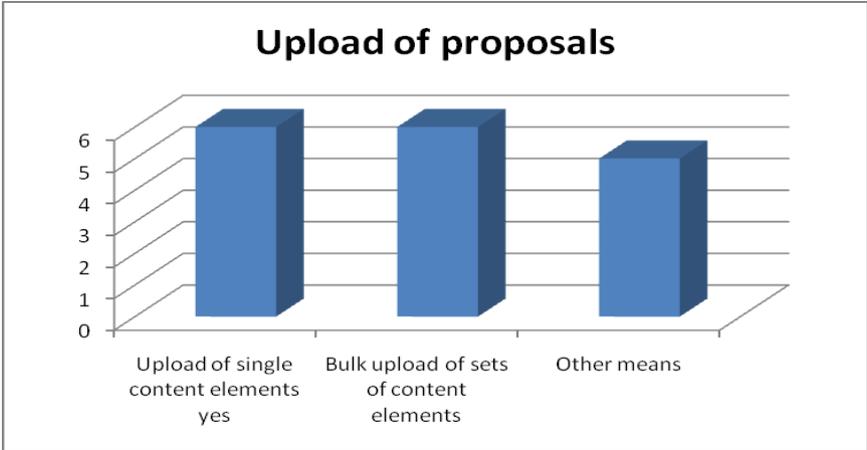
The analysis of the results shows that the majority of the providers (9 providers) have already integrated a maintenance procedure into their ePPS system. The remaining two providers state that they have a maintenance procedure in preparation.

11.3.4 Do you provide means for upload of proposals for new content elements?

In this question the dictionary providers are asked if they provide means for upload of proposals for new content elements for their ePPS systems. This question is only relevant for those providers who answered that they supply workflows. Multiple answers were possible. Out of 11 possible answers the following replies were made

No. of Answers	Option
6	Upload of single content elements yes
6	Bulk upload of sets of content elements
5	Other means

Table 26: Provide means of upload for proposals



Picture 35: Provide means of upload for proposals

The analysis of the results shows that nearly two third of the providers offer various kinds of upload possibilities for proposing new elements or modifications to existing elements. Other means include collection of proposals via web-page, wiki and xls-sheets.

11.3.5 How do you deal with modifications of your content?

In this question, the dictionary providers are asked to state how they deal with modifications of the content of their ePPS systems. This question is only relevant for those providers who answered that they DO NOT supply workflows. The question was an open question and the providers were asked for a verbal statement. Out of 6 possible answers the following replies were made

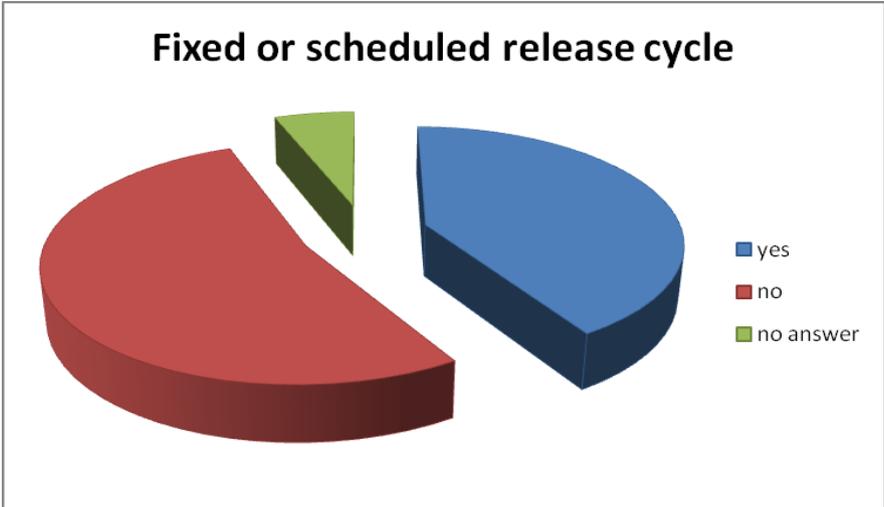
The analysis of the results shows that the answers on this question were individual. The processes to handle updates are very different between the providers. One provider stated that he is changing and publishing data in short cycles without notification to the users about the specific change. For CPV the change of content has big implications as the CPV is based on a binding Community legal act, Regulation 2195/2002 [EU Regulation 2195/2002] from November 5th, 2002. This means any modification to the content or structure of the CPV follows a codified legislative process at EU level. One of the main features, and advantages, of the CPV is its stability over time, with modifications having been implemented in 2003 and 2007. Another organization reported that content changes are under responsibility of the members themselves and are being developed in local databases.

11.3.6 Do you have a fixed or scheduled release cycle where new versions of the content are published?

In this question the kind of release cycle for dictionary content is questioned. Only one answer was possible. Out of 17 possible answers the following replies were made.

No Answers	Option
7	Yes
9	No
1	no answer

Table 27: Fixed or scheduled release cycle



Picture 36: Fixed or scheduled release cycle

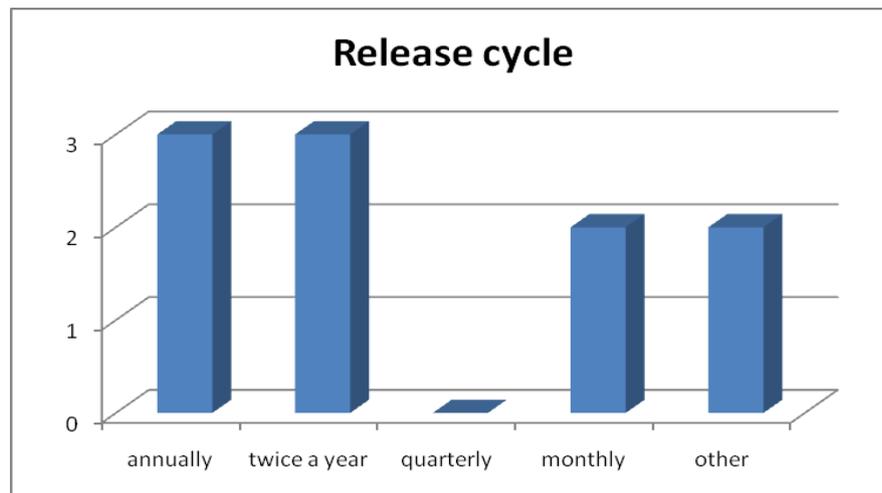
The analysis of the results shows that more than half of the providers answer that they do not have a fixed or scheduled release cycle for publishing new versions. About 40 % of the providers do have a planned and scheduled release policy.

11.3.7 If you publish your content in a release cycle, how often do you provide new versions of your content (e.g. monthly, quarterly, annually, or other period)?

In this question the number and periods of release cycle for dictionary content is questioned. It was possible to make more than one statement as some of the providers have different release cycles for different geographic areas or distinguish between internal release and external release to customers. Out of 17 possible answers the following replies were made.

No Answers	Option
3	Annually
3	twice a year
0	Quarterly
2	Monthly
2	Other

Table 28: Release cycle for content



Picture 37: Release cycle for content

The analysis of the results shows that there is no general rule for release cycles to publish content. The providers plan their release cycle due to the requirements of their content developing community and users.

11.3.8 If you do not have a release cycle, how do you publish your content?

In this question, the providers which do not have fixed release cycles were asked to make a statement how they publish their content. This was an open question and the providers were asked for a verbal statement. This question was relevant for 9 providers whereof 8 made statements on the way of publication of content processed by their organizations.

The analysis of the results shows that the majority of statements is individual to the organization. The answers given include the following statements:

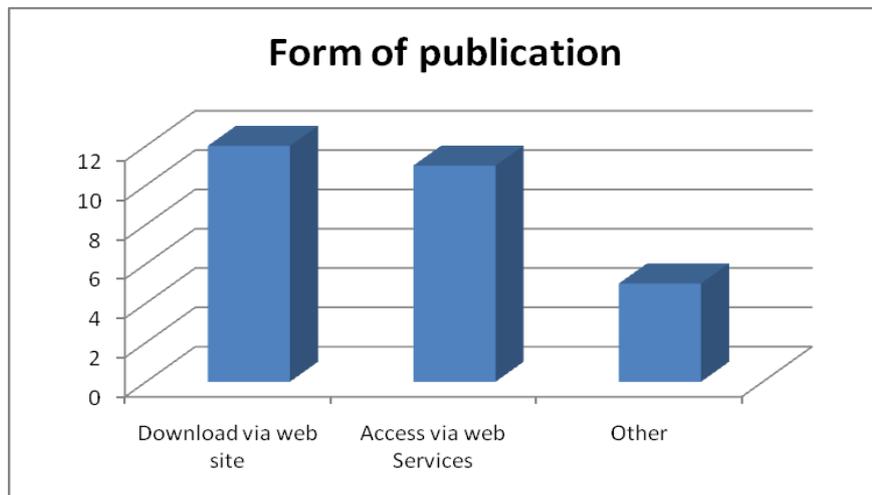
- Publishing is done permanently when the elements passed the workflow.
- The members of the organization decide, at which date a new release will be published (usually every 2 years).
- Modification and publication is a legal act to be published in the Official Journal of the European Communities (JOUE).
- Publication becomes visible on wiki.
- Publication is based on changes requested by the membership.
- Whenever a significant number of updates is available, they will be published.

11.3.9 In which form do you publish your content?

In this question the form of publication of the dictionary content is questioned. Multiple answers were possible. Out of 17 possible answers the following replies were made.

No Answers	Option
12	Download via web site
11	Access via web Services
5	Other

Table 29: Form of publication



Picture 38: Form of publication

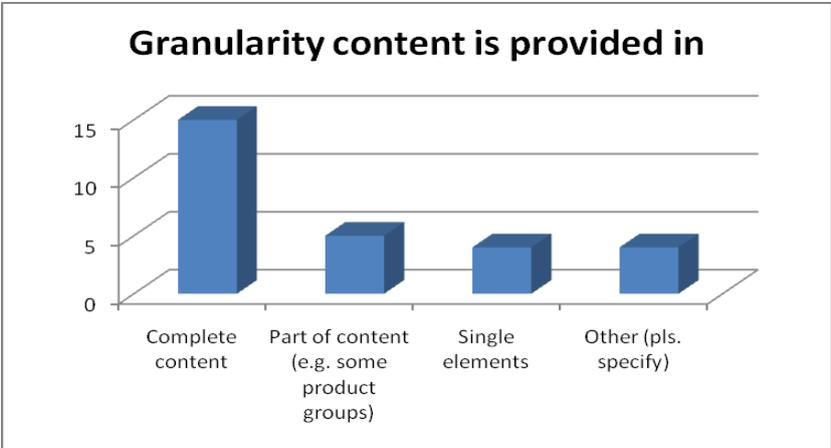
The analysis of the results shows that the majority of the providers both publish and make their content available for download via their web-site and also allow users to access content via web services. Other means of publications include CD, paper based publication. Some organizations give members and customers direct access to a web-based database or publish content in combination with PC based database application.

11.3.10 In which granularity do you provide your content?

In this question the granularity of the provided dictionary content is questioned. It was possible to make more than one statement as some of the providers offer content in different levels of granularity. Out of 17 possible answers the following replies were made.

No Answers	Option
15	Complete content
5	Part of content (e.g. some product groups)
4	Single elements
4	Other (pls. specify)

Table 30: Granularity of content



Picture 39: Granularity of content

The analysis of the results shows that the big majority of providers make the content available as a complete package. Some also provide the possibility to access only parts of the content or single elements. The statements to provide “Other” granularity include:

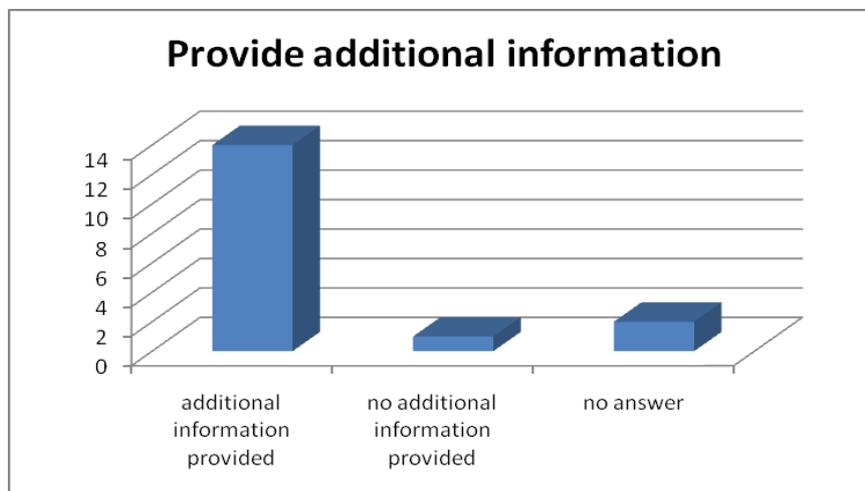
- Order entry collections which are possible in different formats
- Complete new version of the code set with a change/audit log which is only available to its members.

11.3.11 Which additional information do you provide for download?

In this question the dictionary providers were asked to make statements if they provide additional information (e.g. help information, white papers, support material for version upgrade (delta files), information about the maintenance process and the release planning, etc.) for download to their users. This was an open question and the providers were asked for a verbal statement. Out of 17 possible answers the following replies were made.

No Answers	Option
14	additional information provided
1	no additional information provided
2	no answer

Table 31: Provide additional information



Picture 40: Provide additional information

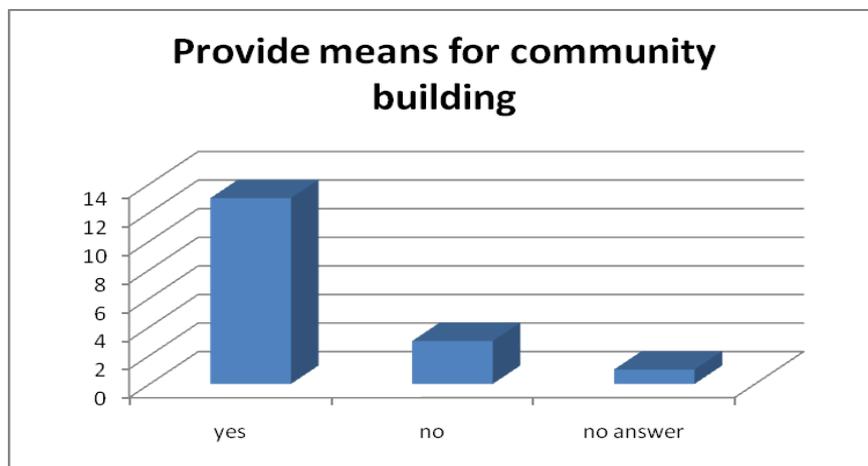
The analysis of the results showed that the majority of the providers make additional information available for download to their users. The additional information provided includes help information, various guides (user, data elements, implementation, xml-exchange formats,) notes, development specifications and delta files in different languages. They are made available on web-portals or for download. For some organizations the provision of information is dependent on the status of the user (members get more information than the public).

11.3.12 Do you provide means for community building?

In this question the dictionary providers were asked to make statements if they provide means for community building (e.g. news, dates of events, forums for discussions, contact means to committees in your organization, etc.) to their users. This was an open question and the providers were asked for a verbal statement. Out of 17 possible answers the following replies were made.

No Answers	Option
13	Yes
3	No
1	no answer

Table 32: Provide means for community building



Picture 41: Provide means for community building

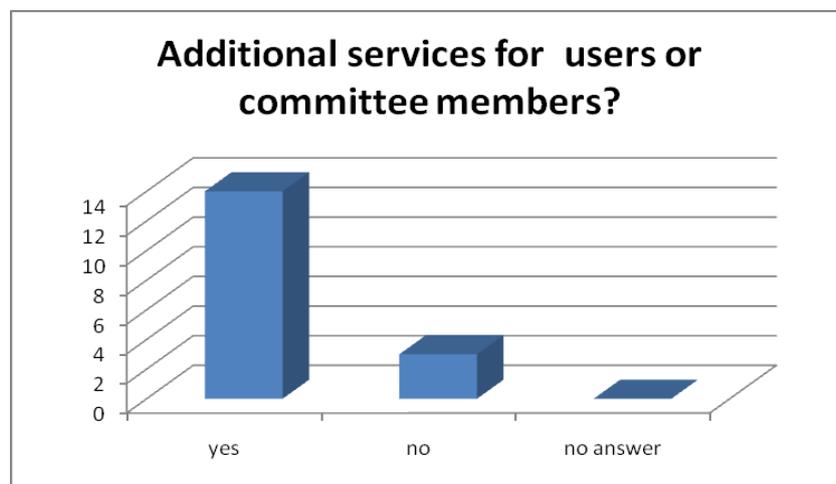
The analysis of the results showed that the majority of the providers provide means for community building. Most of the organizations, which stated that it provides means for community building, use electronic media to do so. Provision of information via a web-page is used the most. In addition, newsletters, forums, open portals, event calendars, events, Online Help-Service, and web-based seminars are used.

11.3.13 Do you provide additional services for your users or your committee members?

In this question the dictionary providers were asked if they provide additional service for their users or community members. This was an open question and the providers were asked for a verbal statement. Out of 17 possible answers the following replies were made.

No Answers	Option
14	Yes
3	No
0	no answer

Table 33: Provide additional services for users or committee members



Picture 42: Provide additional services for users or committee members

The analysis of the results showed that the majority of the organizations provide additional services for users or committee members. The statements about the services provided include:

- Information about the structure of elements is given, success stories, FAQs, licensing information, etc.
- Make to order entry collections / comments on draft standards
- Download support
- Registry of identification guides
- Specialized search functionality
- Automatic alerting on selected categories
- Business process standards for transactions
- Resource page where service providers can expose their value added services
- Newsletter
- White papers
- Business cases

11.4 Analysis of industry requirements regarding content update (industry)

After analyzing the dictionary provider’s perspective of workflows and the ways, frequencies and tools provided for making available the content of their dictionaries to the users, also the users perspective is of interest. In this chapter the requirements of the industrial user of dictionary content were analyzed regarding the frequency of update processes.

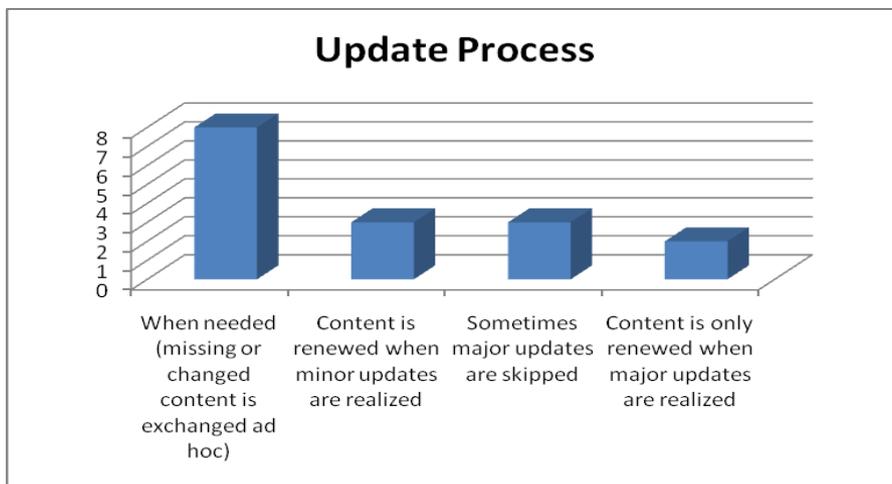
Due to the number of responses and the structure of questions (very similar) the answers of users working with dictionaries from SDOs and the ones using dictionaries from business partners were merged for the analysis. The following question was included in the industry questionnaire to get direct feedback from the users to this topic.

How often do you start update processes with your SDOs or business partners?

In this question the users are asked to state the frequency of update processes for dictionary content. Multiple answers were possible. Out of 10 possible answers the following replies were made.

No. of Answers	Update process
8	When needed (missing or changed content is exchanged ad hoc)
3	Content is renewed when minor updates are realized
3	Sometimes major updates are skipped
2	Content is only renewed when major updates are realized

Table 34: Update process for content



Picture 43: Update process for content

The analysis of the replies show that the vast majority of the users process updates only if there is a specific need for it. The availability of new releases of the dictionary content is for most of the users no reason to perform a content update within their systems as well.

11.5 Conclusion

In the past dictionary was usually related to publication of content in form of a book. The processes were developed in a way that at a certain point of time the whole content had to be “frozen” and published. Publication meant printing the book or a brochure. Corrections in the content, which had to be made after the “GO” decision for printing, were very expensive as this normally meant that the whole publication had to be printed again. Today we still see these release-oriented workflows in electronic dictionary implementations.

Workflows in databases today would technically allow a much more flexible way in development, maintenance and publication of content. An ongoing release process of new and reworked content is possible as single elements can be handled within the workflow.

The results of the survey show that most of the organizations state that they are supporting workflows to develop and maintain content. The kind of workflow, which is supported, is individual to the organizations. On level of standardization bodies, there is a defined workflow in Annex ST of the ISO supplement to the ISO/IEC Directives while national standardization bodies also do have their own workflows. Development of content is both done online and offline. Nearly two thirds of the participants provide means to upload content to propose new content to the dictionaries. Currently we see:

- Dictionary provider see the need to support content development by defined workflows and also already implemented workflows
- Most of the workflows currently in use are individual workflows and often represent the structure of the organization
- Industry needs detailed information about updated content to be able to verify and deploy the changes in their own business processes and electronic catalogues

11.6 Recommendations

As content development and maintenance is the most time and budget consuming activity within a dictionary project it is essential that these investments are being secured. Reuse of content and controlled exchange of content between different dictionaries can also help to work efficiently. As workflows and procedures are currently, very much tailor made for the specific dictionary we see much sense in the coordination of workflows and the development of common standards for workflows to develop, maintain and publish content. Together with the workflow also stage codes which define the maturity of content could be used in the same way. This would support the activities to make content interoperable. Maybe the existing workflows available from the standardization bodies could be a starting point for discussion.

- Intense work to develop common workflows for the development, maintenance and publication of dictionary content
- Define common standards about the maturity of content to enable the users to verify how reliable the content is they want to use
- Develop common standards to document the changes in new dictionary content releases to enable the industry to automatically verify the changes and automate updates of their electronic catalogues.

12 Legal and economic requirements

12.1 Introduction

To provide content electronically to other parties means that the provider has to collect or develop content, verify it and then make it available to the user base. The reasons why to provide content are manifold. It can be declared by law to provide information as done with the Common Procurement Vocabulary (CPV) from the European Commission or the result of some standardization work as the DIN Property Dictionary (DINsml.net) or the result of the work of industry associations like eCI@ss or ETIM.

As a potential user of dictionary content you may also be forced by law to use the content or you get the requirement from one or some of your business partners. In other cases it might be your own free decision to use available dictionary content.

In any cases it is important for the potential users of dictionary content to get information about what he is allowed to do with the offered content and also what are the costs for it.

The costs for a license are usually an easy to answer question by looking up the price list. The terms of usage are sometimes quite tricky and not that clear to the users. Intellectual property rights (IPR) are a very sensitive issue as failures in its usage may involve court procedures and high penalties.

The complexity of intellectual property rights is high as there are many different national laws. Working in the world wide web also sometimes does not make it easy to verify who is your business partner you are getting content from, where is he (or his server) located and which law is applicable.

Intellectual property right for dictionary content may be related to copyright, trademark rules, patents laws and all kinds of national law depending where the content is published. Other laws apply if content is published in the US than if it is published in Germany.

The goal of this chapter is to collect the status quo of legal and economic requirements of current metadata dictionaries and to discuss the requirements and needs. It is not intended to be a complete overview of possible handling of IPRs nor will it give legal advice how to handle IPRs.

The providers of dictionary content do have various different approaches regarding the handling of IPRs and they also have different economic models. Some content is provided free of charge to the users and other providers do have commercial licenses where the users have to pay license fees to be able to access, download and use the content. Most of the dictionary providers are non for profit organizations, standardization bodies or public bodies. Users more and more demand to get content free of charge. But also for free of charge content different types of licenses may apply. In many cases the license fees are needed for financing content development and maintenance and running the organization.

12.2 Licensing of dictionary content

For using content of an electronic dictionary the user normally has to buy, sign or accept a license statement. For some dictionaries it is already necessary to get a license for viewing the content. The license agreement states the terms of usage. The terms of usage define what the user may do with the content he gets access to. The license may include the right to view, download, use it company internally, use it in business processes with suppliers and customers, pass the content on to suppliers and customers, change content or even resell the content. In general there is a differentiation between licenses which are free of charge licenses and commercial licenses.

12.3 Intellectual Property Rights (IP)

Intellectual property (IP) refers to creations of the mind: inventions, literary and artistic works, and symbols, names, images, and designs used in commerce. The World Intellectual Property Organization (WIPO) provides a guide for “Understanding Copyright and Related Rights” [WIPO 1]. Intellectual property is usually divided into two branches, namely industrial property, which broadly speaking protects inventions, and copyright, which protects literary and artistic works. Other interesting sources are the WIPO Intellectual Property Handbook: Policy, Law and Use [WIPO2] and WIPO Guide to Intellectual Property Worldwide [WIPO 3].

Dr. Christian Galinski and Inke Raupach discuss business models and means of IPR protection on structured content on the example of terminology [Galinski, Raupach] in a paper which was presented on the Semantics 2006 in Vienna, Austria.

Extensive work on the procedural and legal framework for data exchange and exploitation was performed in the EUROTERMBANK project [EUROTERMBANK]. Besides the discussion also several sample contracts and agreements for exchanging and using terminology content were developed.

12.4 Copyright

The Berne Convention [Berne Convention] is the very basis for copyright for the protection of Literary and artistic work in most parts of the world. It was first established in 1886 and then revised several times. Copyright gives the author of an original work exclusive right for a certain time period in relation to that work, including its publication, distribution and adaptation. Copyright laws have been standardized to some extent through international conventions such as the Berne Convention and Universal Copyright Convention [Universal Copyright Convention].

The problem with copyright is that there are differences in the conditions depending on the country we are looking at. Most copyright laws state that the author or rights owner has the right to authorize or prevent certain acts in relation to a work. The rights owner of a work can prohibit or authorize:

- the reproduction in various forms, such as printed publications or sound recordings;
- the distribution of copies;
- a public performance;
- the broadcasting or other communication to the public;
- the translation into other languages;
- the adaptation, such as a novel into a screenplay.

12.5 Public Domain

According to Wikipedia [Wikipedia] public domain is a range of abstract materials—commonly referred to as intellectual property—which are not owned or controlled by anyone. The term indicates that these materials are therefore "public property", and available for anyone to use for any purpose.

As we have the situation of national differences with Copyright the same is true for public domain. Public domain has different meanings depending on the applicable national law. In Germany, for example it is not possible to cut the connection between an author and his work, which is possible in the US. Difference are depending both on the applicable law for the author and for the user of some work. To give an example: Pictures from US-governmental organizations may be in public domain when used within the US and may be copyright protected if used in Europe.

12.6 Free Licenses for software

Even if this type of licenses is not directly applicable to dictionary content a short overview about the most popular licenses for free software was compiled.

It became more and more popular in the last years that software is distributed free of charge. But this does not mean that there are no rules for exchanging the software. For software licenses which are free of charge there are some very popular license which are used when exchanging and distributing the software.

These kinds of licenses may serve as a basis for future free licenses for content available in dictionaries.

12.6.1 GNU General Public License (GPL)

The GNU General Public License [GNU General Public License].is a widely used free software license. It was originally written by Richard Stallman for the GNU project. It is currently available in version 3 which was published in June 2007. The GPL grants the recipients of a computer program the rights of the free software definition and uses copyleft to ensure the freedoms are preserved, even when the work is changed or added to. This is in distinction to permissive free software licenses, of which the BSD licenses are the standard examples.

There are several other licenses and derivatives which include:

12.6.2 GNU Lesser General Public License (LGPL)

The GNU Lesser General Public License (LGPL) [GNU Lesser General Public License] is a derivative from the GNU General Public License. It was originally intended for some software libraries and is a modified, more permissive, version of the GPL. In addition, there is a GNU Free Documentation License, which was originally intended for use with documentation for GNU software. This license has also been adopted for other uses, like the Wikipedia project.

12.6.3 European Union Public License (EUPL)

The European Commission also has created and approved a free software license called the European Union Public License (EUPL) [European Union Public License]. Its first version was approved in 2007 and version 1.1 was approved on January 9th, 2009. The EUPL was approved as a license to be used for the distribution of software developed in the framework of the IDA and IDABC programs. Nevertheless, the text of the license is drafted in general terms, so the license may be used for other software applications, as the case may be, by other European Institutions, by national, regional or local administrations, other public entities as well as private entities and natural persons.

12.7 Commercial licenses for content

Of course also commercial licenses play an important role when offering and using dictionary content. Some of the organizations today are depending to some extent on the earnings which are generated by the commercial licenses they are selling to content users. A commercial license means that a user purchases a license from the dictionary provider to get the right to use some parts or the whole content of the dictionary. What kind of content usage is allowed is defined in the terms of the license. Users may be allowed to use the content internally for own purposes, they may also exchange it to business partners or the even may resell content.

12.8 Status Quo - Handling IPRs in current dictionary implementations

12.8.1 Provider Questionnaire

To get a better understanding of the current use of licenses the following questions regarding terms of usage, type of licensing and pricing models were developed for the provider questionnaire:

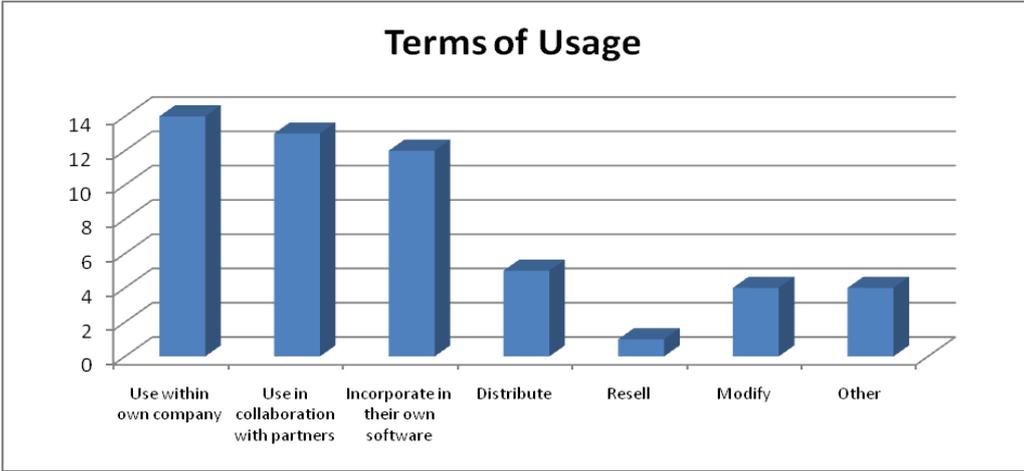
- What are the terms of usage of the content?
- What kind of license do you provide for accessing and using the content to users and committees?
 - Public Domain
 - Free license (e.g. GNU license)
 - Commercial license
 - Other form of license
 - No license
- What is your price model (e.g. free of charge, license fee: one-time, annual, etc.), registration with registration fees, others, please specify)?

12.8.2 What are the terms of usage of the content?

In this question the dictionary providers were asked under what terms of usage they provide the content of their dictionary. More than one answer was possible Out of 17 possible providers 14 participants replied as follows:

No Answers	Option
14	Use within own company
13	Use in collaboration with partners
12	Incorporate in their own software
5	Distribute
1	Resell
4	Modify
4	Other

Table 35: Terms of Usage



Picture 44: Terms of Usage

The analysis of the replies show that the majority of the providers allow users to use the content within their own corporation as well as use it in collaboration with partners and also to incorporate the content into their own software.

The right to distribute and modify the content is granted by 5 providers. 4 providers allow modification of the content while only one organization grants the right to resell the content.

Comments on other forms of granting rights include the right to search the data content free of charge, create mappings to other content and distribution of content to non organizational members under close restrictions.

12.8.3 What kind of license do you provide for accessing and using the content to users and committees?

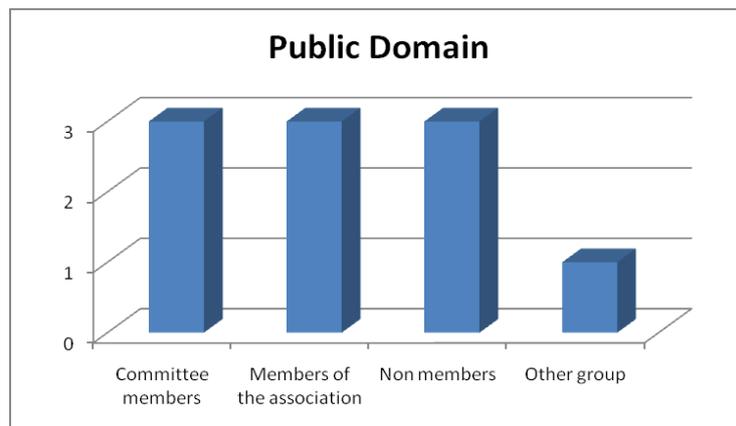
The following questions were developed to learn more about the different licensing schemas in use for publishing dictionary content. This question is a collection of 5 questions regarding licensing. Out of the 17 questionnaires returned only 11 organizations answered this question. Multiple answers were possible.

12.8.3.1 Public Domain

In this question the dictionary providers were asked state for which user group they provide content under a public domain license. Multiple answers were possible. Out of the 17 questionnaires returned the following answers were given.

No Answers	Option
3	Committee members
3	Members of the association
3	Non members
1	Other group

Table 36: Licensing (Public Domain)



Picture 45: Licensing (Public Domain)

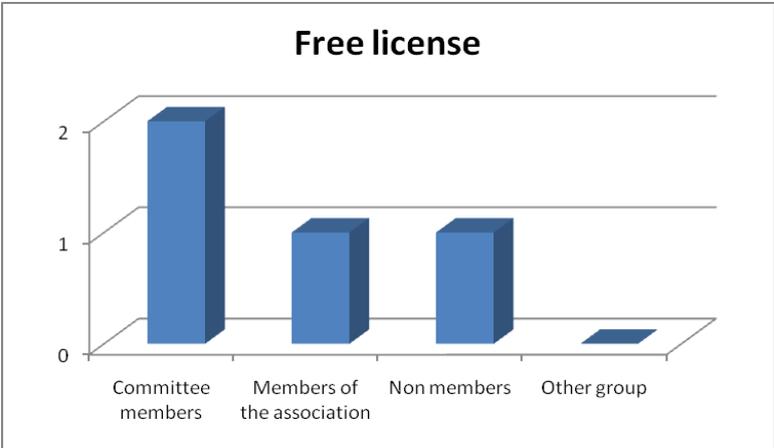
The analysis of the replies shows that in each case three of the providers state that they provide content under the conditions of a public domain license for committee members, members of the organization and to non-members. One of the providers states that he also provides content to other groups, without specifying the group.

12.8.3.2 Free license (e.g. GNU license)

In this question, the dictionary providers were asked state for which user group they provide content under a free license. Multiple answers were possible. Out of the 17 questionnaires returned the following answers were given.

No Answers	Option
2	Committee members
1	Members of the association
1	Non members
0	Other group

Table 37: Licensing (Free license)



Picture 46: Licensing (Free license)

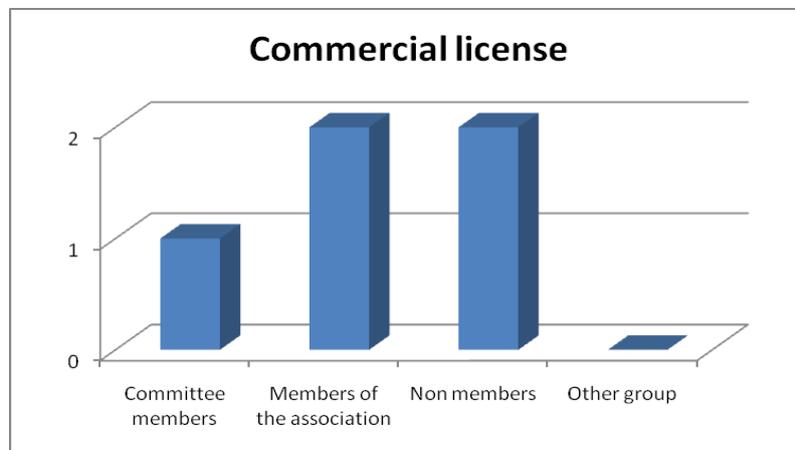
The analysis of the replies shows that two organizations provide content in form of a free license to committee members. In each case, one provider stated that he provides content in form of a free license to members of the association and non members. Another organization provides content only for its committee members under a free license while others may use it under a commercial license.

12.8.3.3 Commercial license

In this question the dictionary providers were asked state for which user group they provide content under a commercial license. Multiple answers were possible. Out of the 17 questionnaires returned the following answers were given.

No Answers	Option
1	Committee members
2	Members of the association
2	Non members
0	Other group

Table 38: Licensing (Commercial license)



Picture 47: Licensing (Commercial license)

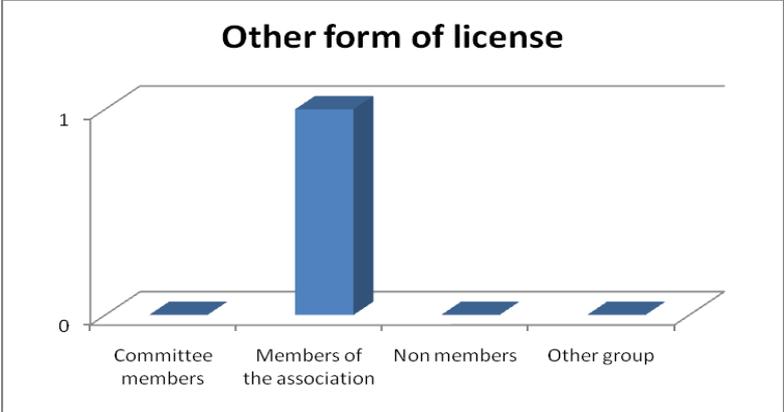
The analysis of the replies shows that in each case two organizations provide content under a commercial license to both members and non members of the organization. One organization states that it provides content under a commercial license for committee members.

12.8.3.4 Other form of license

In this question the dictionary providers were asked state for which user group they provide content under another form of license. Multiple answers were possible. Out of the 17 questionnaires returned the following answers were given.

No Answers	Option
0	Committee members
1	Members of the association
0	Non members
0	Other group

Table 39: Licensing (Other form of license)



Picture 48: Licensing (Other form of license)

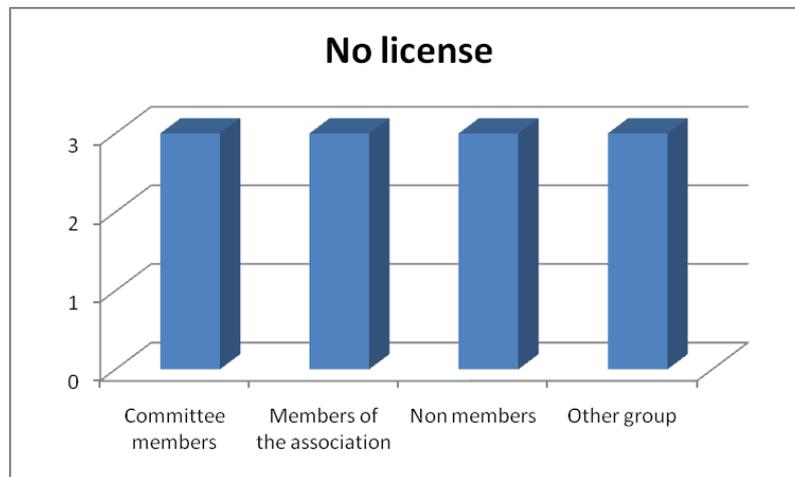
The analysis of the replies shows that only one organization replied and this organization states that it offers content to members of the association under another form of license without further comments on the form of license.

12.8.3.5 No license

In this question the dictionary providers were asked state for which user group they provide content with no need for a license. Multiple answers were possible. Out of the 17 questionnaires returned the following answers were given.

No Answers	Option
3	Committee members
3	Members of the association
3	Non members
3	Other group

Table 40: Licensing (Other form of license)



Picture 49: Licensing (Other form of license)

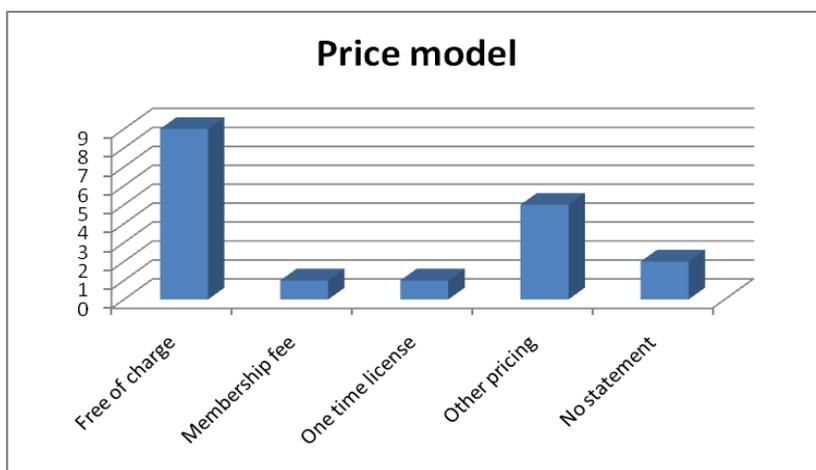
The analysis of the replies shows that three organizations stated that no license at all is needed when using content from their dictionary.

12.8.4 What is your price model?

In this question the dictionary providers were asked what pricing model they use for providing content. This was an open question and the providers were asked for a verbal statement and they were also asked to attach a price list when returning the questionnaire. More than one answer was possible. Out of the 17 questionnaires returned 15 providers made statements and the following answers were given.

No Answers	Option
9	Free of charge
1	Membership fee
1	One time license
5	Other pricing
2	No statement

Table 41: Price model



Picture 50: Price model

The analysis of the replies shows that the majority of the questioned providers offer their content free of charge. One organization offers a membership fee based pricing and optional a one-time license. Five organizations offer other pricing models mostly depending on company size, content volume, data format or locations where the content shall be used.

The pricing for non free of charge licenses ranges from about EUR 40,-- for an xls-based license up to about EUR 1.850,-- for a network-license of content. The pricing models also differ in the approach. Some providers base the licensing on a content release (e.g. license for xxx release 4.0). Other providers base the terms of usage on an annual payment (e.g. access to the dictionary for 2009).

12.9 Status Quo - IPRs of Dictionary Content used in Industry

12.9.1 Industry Questionnaire

The last chapter showed the analysis of the handling of IPRs from provider's perspective. This chapter will focus on the status quo of IPR handling in dictionary implementations in the industry. Under which conditions does the industry use dictionary content provided by their business partners and how does industry handle dictionary content, which they develop on their own.

Due to the number of responses and the structure of questions (very similar), the answers of users working with dictionaries from SDOs and the ones using dictionaries from business partners were merged for the analysis.

To get a better understanding of the current handling of IPRs the following questions regarding terms of usage were developed for the industry questionnaire:

- Which ePPS are you using?
- What types of content do you use via ePPS with your SDO or business partner?
- What types of content do you use via ePPS within your own company?
- How are the terms of usage when you use the content of a dictionary of your SDO or business partner?
- How are the terms of usage when others use your dictionary content?

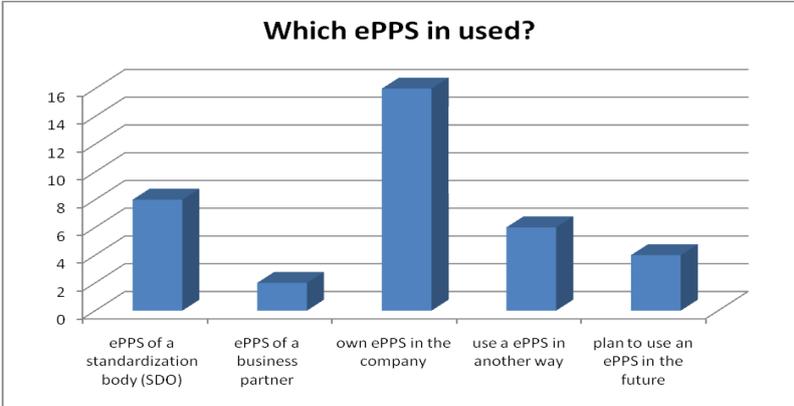
The discussions with organizations providing corporate internal dictionaries showed that there is usually no pricing of content no questions on that topic were included in the industry questionnaire.

12.9.2 Which ePPS are you using?

To understand better the current use and the needs of dictionary users the participants were asked to declare which kind of dictionaries they are using. Multiple answers were possible. The results to the question which dictionaries are used were:

No. of Answers	Which ePPS is used
8	ePPS of a standardization body (SDO)
2	ePPS of a business partner
16	own ePPS in the company
6	use a ePPS in another way
4	plan to use an ePPS in the future

Table 42: Which dictionaries are used?



Picture 51: Which dictionaries are used?

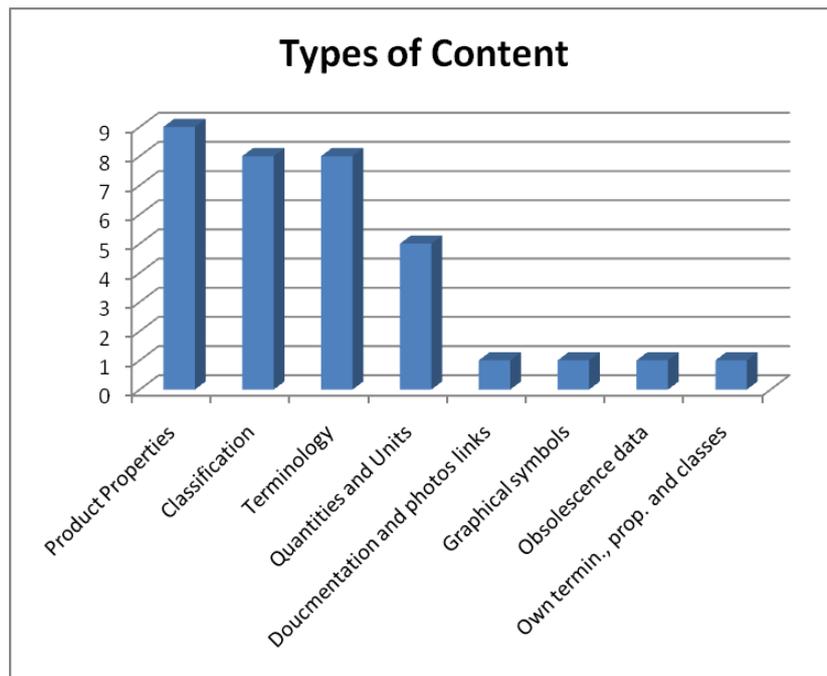
After this question the online survey continued depending on the answers of this question. Due to the number of responses and the structure of questions (very similar), the answers of users working with dictionaries from SDOs and the ones using dictionaries from business partners were merged for the analysis.

12.9.3 What types of content do you use via ePPS?

This question was about the content which is requested when using a dictionary from an SDO or a business partner. Multiple answers were possible. Out of 10 possible answers the following replies were made.

No. of Answers	Types of content used via ePPS
9	Product Properties
8	Classification
8	Terminology
5	Quantities and Units
1	Documentation and photos links
1	Graphical symbols
1	Obsolescence data
1	Own terminology, properties and classes

Table 43: Types of content used



Picture 52: Types of content used

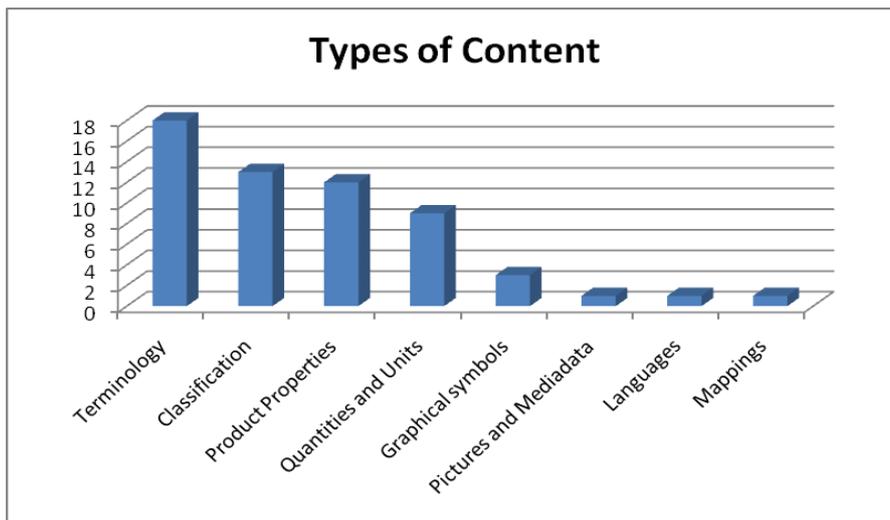
The analysis of the results of the survey shows that a vast majority of the participants use product properties, classification and terminology. We also see that quantities and units are more and more of interest for the industrial users. Users also exchange other kind of data like graphical symbols, documentation and photo links, obsolescence data or own (private) information by means of a dictionary.

12.9.4 What types of content do you use via ePPS within your own company?

This question was about the content which is used in the company’s own dictionary. Multiple answers were possible. Out of 26 possible answers the following replies were made.

No. of Answers	Types of content
18	Terminology
13	Classification
12	Product Properties
9	Quantities and Units
3	Graphical symbols
1	Pictures and Mediadata
1	Languages
1	Mappings

Table 44: Types of content used



Picture 53: Types of content used

The analysis of the results of the survey shows that the majority of users have terminology dictionaries. About half of the users use classification and product properties within their dictionaries followed by quantities and units. Three users have graphical symbols in their dictionary while one reply was on pictures and media data, languages and mappings, respectively.

12.9.5 Which functions are missing within your own ePPS?

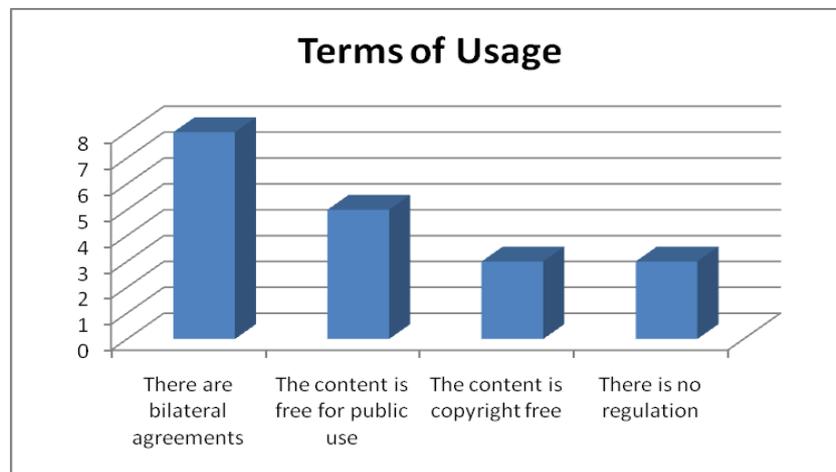
On the question, which functions are, missing within your own dictionary two users stated that no function is missing. One statement of missing functionality was made on data exchange with other companies, direct links or import/export to business partners and graphical information (pictures; drawings). One user stated that a general philosophy is missing.

12.9.6 How are the terms of usage when you use the content of a dictionary of your SDO or business partner?

In this question, the users are asked to state under which terms of usage they use dictionary content. Multiple answers were possible. Out of 10 possible answers, the following replies were made.

No. of Answers	Terms of usage
8	There are bilateral agreements
5	The content is free for public use
3	The content is copyright free
3	There is no regulation

Table 45: Terms of usage of dictionary content used



Picture 54: Terms of usage of dictionary content used

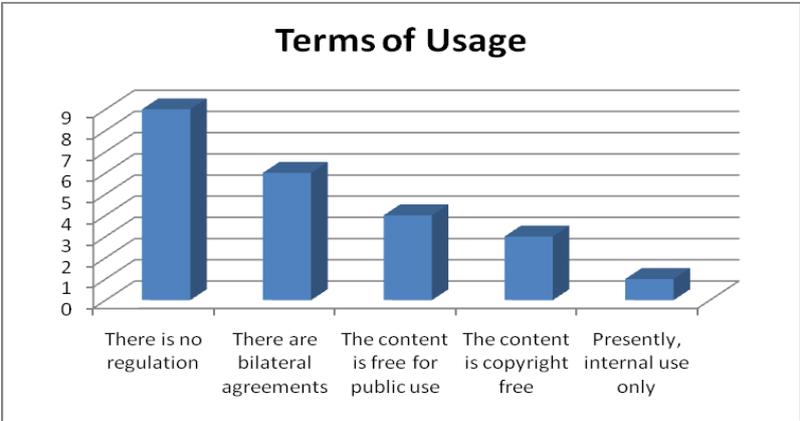
The analysis shows that the vast majority of the users have bilateral agreements with the SDOs and business partners about usage of content. Half of the participants use content which is free for public use. About a third of the users use content which is copyright free and also a third of the users state that they use content without a regulation.

12.9.7 How are the terms of usage when others use your dictionary content?

In this question the users are asked to state under which terms of usage they use their dictionary content. Multiple answers were possible. Out of 26 possible answers the following replies were made.

No. of Answers	Terms of usage
9	There is no regulation
6	There are bilateral agreements
4	The content is free for public use
3	The content is copyright free
1	Presently, internal use only

Table 46: Terms of usage for own dictionary content



Picture 55: Terms of usage for own dictionary content

The analysis of the replies shows that the majority of the organizations, which replied, do not have a formal regulation for using content on internal dictionaries. Some organizations have bilateral agreements and there are also several implementations with free content or content which is copyright free. One remark from a user stated that the dictionary presently is only for internal use.

12.10 Conclusion

Both the research in applicable rules, regulations and laws as well as the results from the questionnaires show, that the handling of IPRs for dictionary content currently is very heterogeneous. The situation of different national laws and regulations for copyright, public domain, etc. make it also very difficult for a potential user to understand the situation and possible legal implications. Dictionary providers have the same problems of handling various IPRs.

In times of a raising number of lawsuits this situation may lead companies to decisions to not implement meaningful dictionary content because of potential legal risks if the conditions for usage are not clear. The increasing number of dictionaries comes along with more and more content overlapping with other dictionaries as whole or in parts. This increases the need for data harmonization and data consolidation between different dictionaries. The outcome of consolidating and harmonizing dictionary content automatically raises the question of at least harmonization of IPRs used for publication of content.

Today we see no general type of a free of charge license or commercial license for content. In current implementations the licenses are usually specific to the provider which offers the license. An example of a commercial license is the download license offered by eCI@ss [eCI@ss General Conditions of Use] As alternative to a download license some organizations also offer the right to access and use the content as benefit coming along with the membership to the organization (e.g. RosettaNet and eCI@ss). Also different membership levels going along with different licenses are possible.

Discussions with industrial users show that industry is getting more and more sensitive regarding IPRs of content they intend or do use. If terms of usage are not clear or not available the industry will refuse to use this content. As many business processes will be based on such dictionary content, the industry will not take the risk of a potential lawsuit. The industry is looking for long term secure and reliable dictionary content as it is business critical.

Today the situation is that:

- IPRs are handled individually in most of the current implementations
- Unclear IPRs for dictionary content may lead to a no go decision for using dictionary content
- Pricing of content and related terms of usage is also individual to the dictionary today

12.11 Recommendations

Industry is increasingly depending on the availability of high quality meta data content to be used in business processes internally and in electronic data exchange with suppliers and customers. It is essential for the companies that they have legal certainty in the use of dictionary content. As the terms of usage for dictionary content today is very inconsistent and individual for each provider it creates a lot of efforts for the users to ensure that the terms fulfill their needs.

Therefore we recommend to:

- work on the development of general terms of use for dictionary content
- initiate a discussion about IPR handling for dictionary content in the web
- include business cases to enable non-lawyers to understand if the Terms of Use fulfill their needs

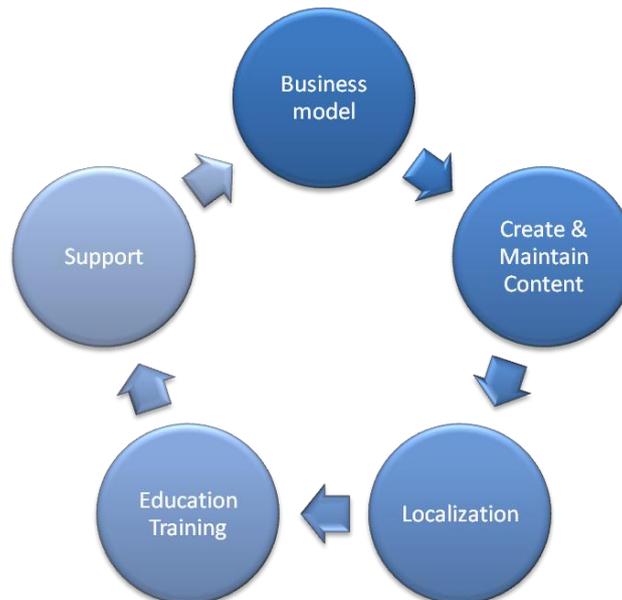
13 Skills and competences

13.1 Introduction

For setup and operation of an online dictionary as well as providing content to the users the organization offering these services has to think about the necessary skills and competences to be able to provide a comprehensive service offering. This chapter provides information about necessary skills and competences also taking into account people with special needs.

Looking at today's way of doing business we have integrated the internet fully into our business processes. We need information to be provided at our fingertips as soon as we have a specific need. However, the information needs to be reliable. For describing products, creating electronic catalogues and delivering product and safety related data we even need data which is clearly identifiable and persistent on a long-term scale. Changes of data need to be traceable and information about compatibility needs to be provided. When providing data dictionaries we need to deal with these issues.

For organizations providing content in form of an online dictionary, we see the following areas of skills and competences:



Picture 56: Areas of Competence

The following chapters discuss necessary skills and competences.

13.2 Business model

To start a dictionary project it is necessary to think about the basic business setup and the underlying business model. This work is usually done by developing a business plan. The basic conclusion of a business plan is to get a basis for the decision to go for the project or not. This includes the needed budget to setup and maintain the content as well as to setup and maintain the needed dictionary infrastructure.

Needed skills and competences:

- good economic background

- understand online dictionaries
- understanding business processes of target groups (content development, content users)
- writing business plans

Needed skills and competences regarding users with special needs:

- basis understanding on how to support users with special needs

13.3 Developing and maintaining content

Development and maintenance of content can be organized in different ways within an electronic dictionary. The workflows of processing the elements can be more or less sophisticated depending on the community size, budget and organization of the dictionary provider. Usually development and maintenance of content is organized in three groups.

13.3.1 Proposer

A proposer is a user who proposes new elements or change requests to existing element in the dictionary. Often these users are domain experts who want to use the published dictionary content within their business processes.

Needed skills and competences:

- domain knowledge of elements or change requests proposed
- basic information about data model and formal requirements for a proposal
- basic information about process of proposing, processing and publicizing elements in the dictionary

Needed skills and competences regarding users with special needs:

- basic information on which additional information may be provided in a proposal to support users with special needs. (e.g. not only provide a definition for a product property but also provide a mp3-file to be able to listen to a definition of a property)

13.3.2 Expert

An expert is a user processing applications and change requests within the dictionary to develop a draft for publication. Depending on the volume of a dictionary and diversity of domains experts are usually organized in one or more groups. Each group is responsible for a defined domain.

Needed skills and competences:

- domain expert
- detailed knowledge about formal requirements of the data proposed in the dictionary
- detailed knowledge of processes for content development

Needed skills and competences regarding users with special needs:

- good understanding on processing information to be provided for users with special needs

13.3.3 Quality experts

A quality expert is a user who verifies the drafts developed by the experts from the proposals and change requests. After verification, the drafts will be rejected for rework or released for publication.

Needed skills and competences:

- Process expert
- Detailed knowledge about formal requirements of the data proposed in the dictionary
- Detailed knowledge of processes for content publication

Needed skills and competences regarding users with special needs:

- Detailed information on release processes for information to be provided for users with special needs

13.4 Localization

In a global environment it is increasingly necessary to provide multilingual content. The analyses from the dictionary provider questionnaire showed that we already have implementations with more than 20 language versions supported. As the content of an electronic dictionary is changing all the time it is necessary to both define processes how the development of localized content can be organized and how to process the localization work itself. Therefore, we see two groups of users.

13.4.1 Language manager

The language manager is responsible to define and organize the processes to develop and maintain the various language versions of the content. Definitions have to be developed about the handling of languages, responsibilities and rules for the localization work. This includes also the release process.

Needed skills and competences:

- Process expert
- Detailed knowledge about formal requirements of the data proposed in the dictionary
- Detailed knowledge of processes for content publication

Needed skills and competences regarding users with special needs:

- Detailed information on release processes for information to be provided for users with special needs

13.4.2 Translator

The translator is responsible to translate content from one language into another. To deliver high quality translation results it is necessary that the translator has both language and domain knowhow.

Needed skills and competences:

- Language competence
- Domain competence
- Basic knowledge of translation processes and formal requirements

Needed skills and competences regarding users with special needs:

- Basic information on translation processes for information to be provided for users with special needs

13.5 Education and Training

To motivate people to participate in the development of dictionary content education and training shall be offered. To provide this information we see two groups involved.

13.5.1 Developer for Training-Material and Documentation

These are persons, who develop training material, online trainings or documentation. Members of this group have an excellent understanding about handling the dictionary, have good know how on the rules of content development, maintenance and publication. Training and documentation is made available more and more on an online basis, so handling of IT-tools for development of the courses is also a requirement.

Needed skills and competences:

- Excellent knowledge in handling the dictionary
- Good knowledge about the processes for content development and maintenance
- Good educational/writing skills to develop training and documentation
- Good knowledge on IT-tools used for develop (Online-) trainings and documentation.

Needed skills and competences regarding users with special needs:

- Good knowledge about training methodologies for users with special needs

13.5.2 Trainer/Coach

These are persons, who perform face to face trainings, workshops and coaching. Members of this group conduct lessons and workshops in classrooms and coach individual users. In addition, online-trainings and coaching via the internet gets increasingly popular.

Needed skills and competences:

- Excellent knowledge in handling the dictionary
- Good knowledge about the processes for content development and maintenance
- Excellent educational and communication skills

Needed skills and competences regarding users with special needs:

- Good knowledge about training methodologies for users with special needs

13.6 Support

A provider of an online dictionary also has to make sure that his users have access to a proper and professional support organization. This includes the definition of Service Level Agreements and it's communication to the users.

Needed skills and competences:

CWA 16100:2010 (E)

- Excellent knowledge in handling the dictionary
- Good knowledge about the processes for content development and maintenance
- Good communication skills and competence in problem solving

13.7 Operation of Dictionary

Providing a dictionary includes also the operation of the dictionary. As the operation of a dictionary is often outsourced to professional system operators and heavily depending on the software we will not discuss the needed skills and competences here.

13.8 Supporting users with special needs

Dictionary providers shall also take into account users with special needs. Online dictionaries are available everywhere via the internet and therefore e.g. an excellent source for people with impairments in mobility.

In the lesson “Enabling Technology for Users with Special Needs” [Edwards, Edwards, Mynatt] the authors discuss the need for people with mobility impairment, vision impairment, speech impairment, language impairment, hearing impairment, and cognitive impairment.

Many challenges in addressing the needs of people with physical and cognitive impairments who need or want to use computers still need to be overcome. These challenges conceal numerous basic issues in human- computer interface design.

Many possible additions to dictionary content, such as voice input or switching the size of texts to display content are required by people with disabilities.

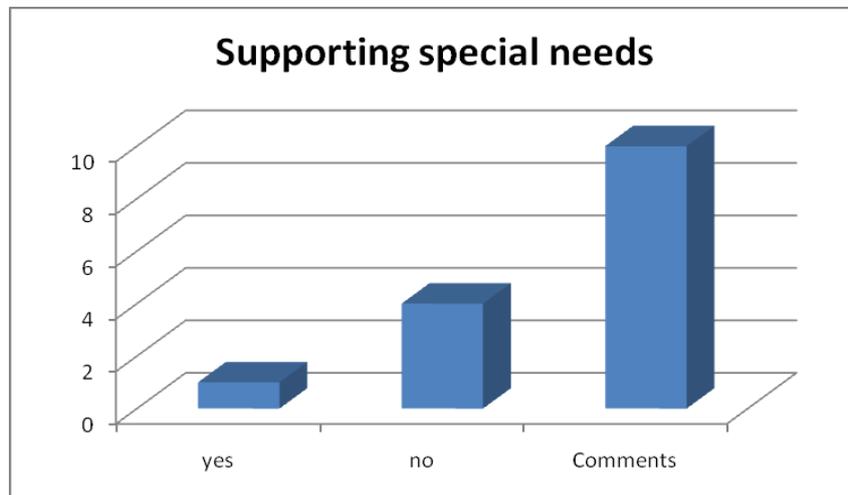
To get a picture on the status quo of dictionary implementations the following question was included in the dictionary provider questionnaire.

- Do you provide your content in a form, which supports users with special needs?

In this question the users are asked to state if they provide dictionary content in a form which supports users with special needs. It was an open question and if the providers support such content they should make a verbal statement about what kind of content they support. Out of 17 possible answers the following replies were made.

No Answers	Option
1	Yes
4	No
10	Comments

Table 47: Supporting special needs



Picture 57: Supporting special needs

The analysis of the replies shows that there was only one provider who states that he provides content which supports users with special needs. 4 providers mentioned that they do not provide content which supports users with special needs. 10 providers made comments. These comments included statements that there shall be support in the future and that additional information like graphical data, hyperlinks and sound files shall be included.

Part 3 Usage of ePPS in industry

14 Introduction

Part 3 of this CWA discusses aspects of using an ePPS. It starts with a number of scenarios in which ePPS can be used beneficially. To show the potential benefits of using an ePPS and supporting company internal business processes, one scenario has been elaborated in more detail: The generation of catalogues on demand of a catalogue receiver. This scenario illustrates the potential for automation if meta data can be accessed directly through a global ePPS and if standards are used both for describing and querying the content of a catalogue and for accessing the ePPS.

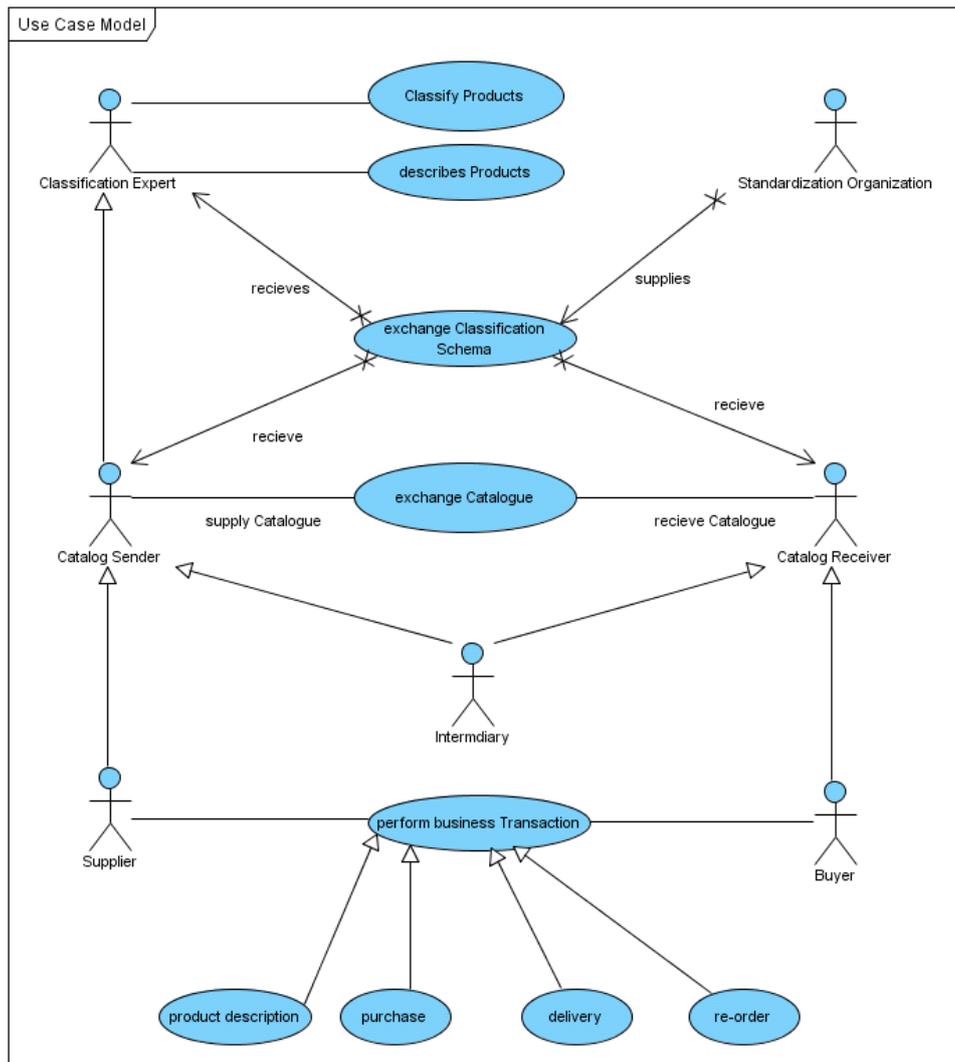
Another aspect described in this Part are the experiences made by companies who are using external ePPSs (either from their business partners or from associations or standardization organizations) and what they regard as personal skills which people working with an ePPS need to have. To support the implementation of ePPS and response to the experiences, two checklists have been developed in this project which may guide companies and standard providers in their way to realize an ePPS for their meta data.

Many strategies and measures are proposed in this Part for spreading the knowledge on the use of ePPS in industry and in SMEs. Many of the proposed activities have already been launched during the ePPS project, (both within ePPS or within related other projects), and the Part concludes with a list of activities which should be pursued after ePPS to ensure a sustainable knowledge transfer to industry.

15 Scenarios for use of ePPS in industry

The results of the questionnaires as well as the results from part 1 and part 2 of this CWA show, that the use of dictionaries is always related to specific scenarios. The scenarios are based on business processes of the respective organization.

In the previous project of CEN/ISSS Workshop eCAT ePDC a reference model for the integration of electronic catalogues was described in CWA 15295. The picture below shows this reference model.



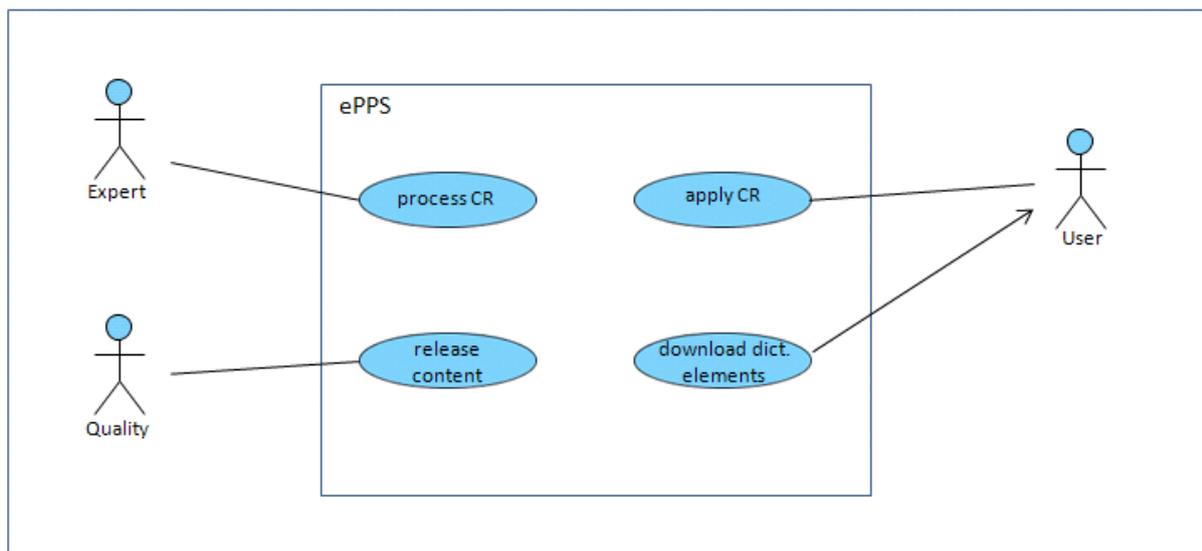
Picture 58: Reference model for the integration of electronic catalogues

The following scenarios are based on this general model and show examples about how ePPS systems can be implemented and what requirements have to be fulfilled.

15.1 Organization providing classification system

An organization providing a multilingual classification system operates an ePPS to develop, maintain and publish the multilingual classification system. The ePPS has to support the following processes:

- Collaborative online development of the classification system by expert-groups (support workflow)
- Online capturing and administration of change requests from classification system users (support change management)
- Publishing of new releases of the classification system (support release management)



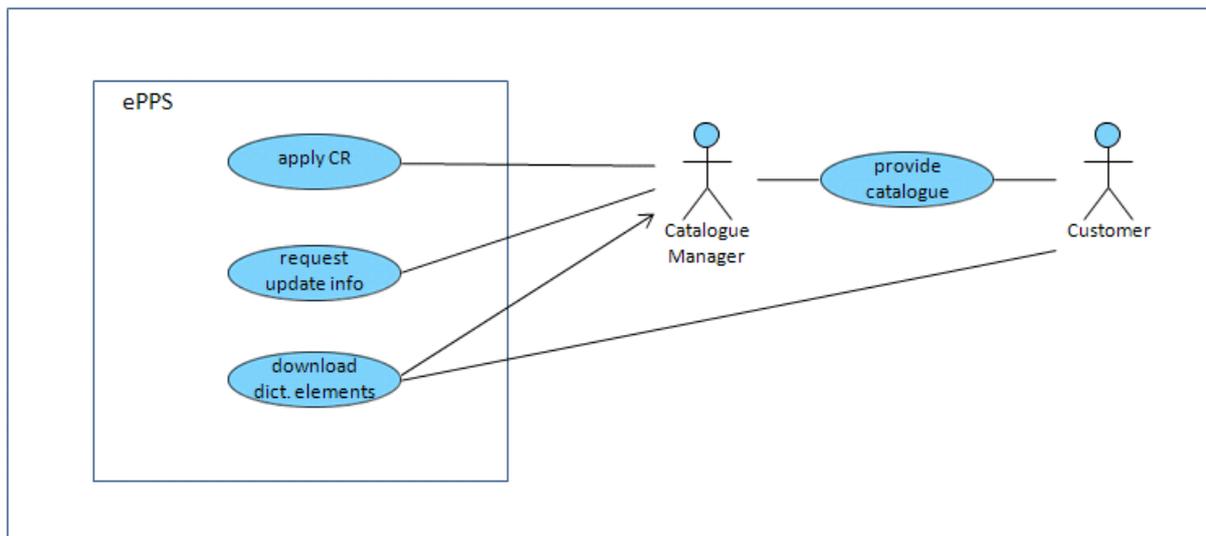
Picture 59: Organization providing classification system

The system has to be operated in a secure and highly available environment on a 24/7 basis and serve users with different needs. It shall enable dedicated experts to collaboratively develop and maintain assigned domains of the classification system. The development and maintenance of the classification system has to follow a defined online-workflow, which includes the quality team to check and release changes of the classification. The users of the classification system may contribute to the development and maintenance of the classification system by entering online change requests. To provide different language versions it is necessary that language and domain experts can be qualified and registered to do translation work and release the language versions.

15.2 Company providing electronic catalogues

A company acting as supplier of products to other companies is requested by its customers to provide its product catalogues in a standardized electronic format. The company is requested to use standardized product descriptions. These product descriptions can be accessed from an ePPS. The company therefore has to use an ePPS to access and download product specifications to fulfill its customers requirements. The ePPS has to support the following processes:

- Requesting and retrieving (download) of product specifications
- Providing different versions of product descriptions
- Providing update or mapping information between different versions of product descriptions
- Entering of change requests from users to address necessary changes of product descriptions
- Automatic notifications on changes of product descriptions



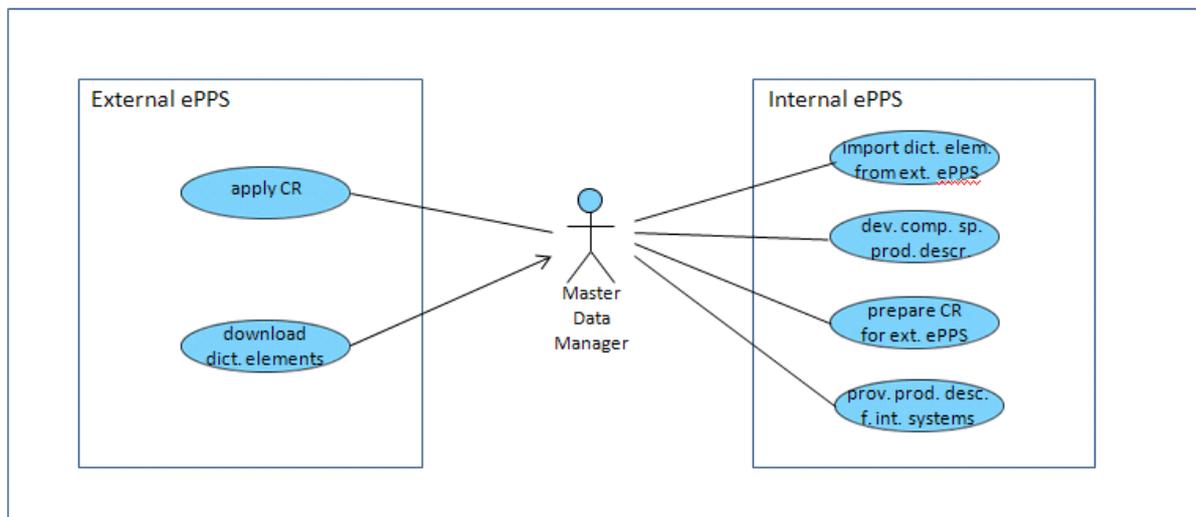
Picture 60: Company providing electronic catalogues

Providing electronic catalogues is not a one-time activity as products as well market needs continuously develop. On the one hand the content of the dictionary will change due to new or discontinued products from the supplier. On the other hand the product structure will change due to new features of the products or new information needs from customer side. Therefore it is necessary that users are supported in this adaptations which means that they should be automatically notified when a new version of a product description is available as well as there should be information provided to automatically transfer product catalogues from one version of the product description to the next release. This scenario is described in more detailed in Chapter 16.3 to show the potential of process automation if ePPSs provide common access interfaces to the dictionary content.

15.3 Company providing product descriptions for internal systems

During the life-cycle of a product data about the product is stored in many different systems within the company (product development, production planning, procurement, marketing, sales, maintenance, recycling). Data about the product is exchanged to suppliers and customers as well. To assure that product data are consistent and redundancy of product data maintained in the various systems is reduced to a minimum, many companies setup their own ePPS to administrate product descriptions. These ePPS systems have to support the following processes:

- Requesting and retrieving (download) of dictionary elements and updates of dictionary elements from external ePPS
- Provide functionality for adapting standard dictionary elements to company specific needs by internal expert groups
- Provide adapted and released internal product descriptions for various internal systems
- Collect and process internal change requests on product descriptions
- Apply company change requests to external ePPS
- Provide product specifications to suppliers and partners



Picture 61: Company providing product descriptions for internal systems

Company internal ePPS systems are getting more and more to business critical systems to consolidate product descriptions on a company wide basis. If the companies are able to implement a consistent description of their products and services throughout the various systems in use, data exchange between systems will become easier and the complexity of business processes will be reduced.

The company internal ePPS has to act as a central system to provide structural data (master data) to various internal systems. It will take over e.g. classification data from external ePPS as a basis for the internal description of the products. Internal expert groups will adapt and enrich the standard classification to company specific needs. Before providing product descriptions to internal systems the product-views for the systems have to be defined. For instance, a system for product development uses a different set of properties for the same product as a procurement system or a spare marketing catalogue. The ePPS may also serve to retrieve structural data for product catalogues for external users from suppliers or customers. Therefore the system has to provide functionality for external users to login and download catalogue specification. Also automatic notifications on changes are necessary.

16 Integration of ePPS into user processes

The use of dictionaries is important for the exchange of product data between business partners because the dictionary provides definitions of the structure and of the meaning of data which is used in the exchange process. Nevertheless, the participating companies often see the use of dictionaries as a burden: They have to provide data in compliance to a dictionary (e.g. eCI@ss) which requires them to either restructure their internal databases or to provide mappings from their internal product descriptions to the “external” product description according to the external standards. The situation is complicated by the fact that suppliers often have to provide their product data according to several dictionaries, e.g. eCI@ss, UNSPSC, CPV, GPC, ETIM, proficl@ss, etc. The situation is further hampered by the fact that these dictionaries and classifications evolve over time and exist in various versions, and different companies (i.e. customers of the supplier) use different versions of the dictionaries. On the other hand, access to the dictionary servers is often only possible by manual download of the complete dictionary, and very often this is only obtainable in a dictionary provider specific form.

In this situation it is very difficult for users to automate the process of electronic catalogue generation and to integrate it into other company specific processes. In the following, some requirements for such automation are specified. Some key factors are:

- Formalization of catalogue requirements
- Automated access to the content of a dictionary server

16.1 Formalization of catalogue requirements

An automation of the generation of e-catalogues is only possible, if the specification of the expected catalogue can be specified in a form which is understandable by software. Whereas currently often text documents are used to specify how a catalogue has to look like, it is important to move to formal specification of the content of the expected catalogue. Such a specification can be understood similar to a query to a database which tells the database two things:

- Which objects have to be delivered
- Which information has to be delivered per selected object

If we translate this to the catalogue specification, it means:

- Which products have to be put into the catalogue
- Which properties need to be given per product (in a specific product class)

The selection of products can be done in various ways:

- Specification of the requested product groups (classes)
- Specification of characteristics of the products to be delivered, e.g.
 - Pumps with a specific pumping capacity
 - Laptops with a specific processor performance and disc capacity

CWA 16100:2010 (E)

- Specification of requirements which have to be fulfilled by the products. This is a more indirect specification of the product and does not ask for specific characteristics but is based on the situation in which the product will be used, e.g.
 - Temperature range in which the pump has to work
 - Purpose for which the laptop is needed

The specification of the information which is expected to be contained in the catalogue may comprise a number of different elements, e.g.

- Specification of properties to be delivered (normally based on a standardized dictionary)
- Specification whether a property has to be delivered or may be optionally delivered
- Specification of units and value lists (possibly deviating from the underlying standard)

The use of standards is recommended in this area, unfortunately, currently no generally accepted standard exists. Nevertheless, there are several activities ongoing. Examples are the work of ProList [...] which uses specific categories of properties to describe the requirements and the characteristics for products and which uses views to specify requested property lists. In ISO 22745-30, identification guides are defined which allow to specify the information which has to be delivered for the selected products. Moreover, in ISO 29002-31 a query mechanism is defined which allows to specify the products which are to be put into a catalogue.

In the future, it is desirable that these and similar standards are integrated to provide a common way of expressing the content of electronic catalogues.

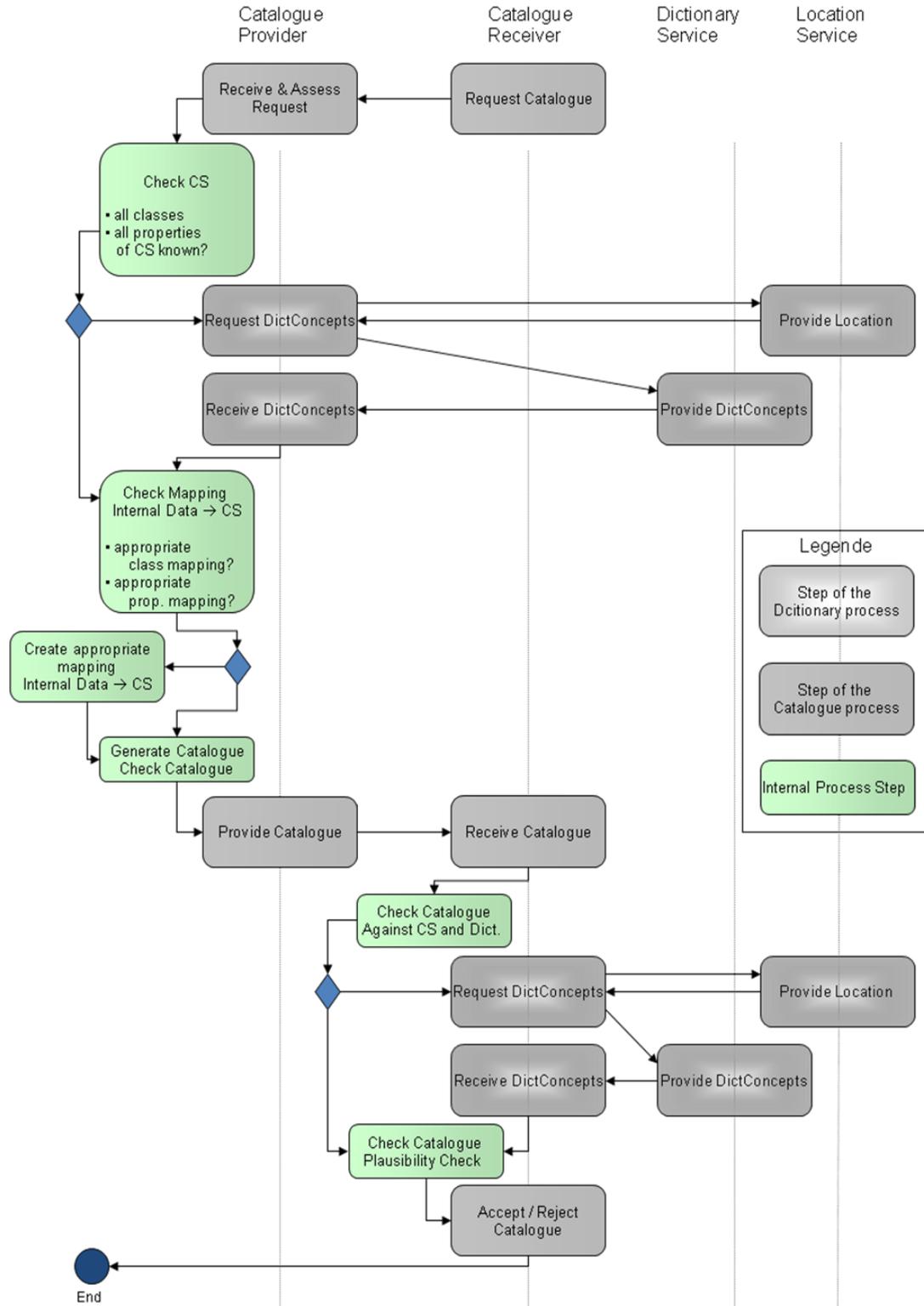
16.2 Automated access to dictionary content

The second requirement is the automated access to the content of dictionary by the catalogue software of the dictionary user. Such a mechanism is provided by ISO 29002-20 in form of Web services specifications, which define a common interface for tools to access a dictionary server regardless of the underlying dictionary data model, provided in the ePPS. This is described in much more detail in first part of this CWA.

16.3 Integration of user processes and dictionary access via ISO 29002-20 Web Services

An integration of a catalogue generation process with the automated access to a dictionary server is illustrated in a scenario in Picture 62. It shows the communication between catalogue requester and catalogue provider and integrates the communication of the catalogue provider and the dictionary server. In addition, the scenario shows some of the internal steps which are done by the catalogue provider to generate the catalogue and by the catalogue requester to check and understand the catalogue.

The scenario assumes that a catalogue supplier gets a formal request from one of its customers to provide a catalogue. The formal request contains information about the classes (of a specific classification system) for which products are to be delivered, it specifies, which information has to be delivered for these classes, etc. These specifications are done by references to the respective classes and properties using the ISO 29002 identification scheme.



Picture 62: Integration of user processes and dictionary server processes

CWA 16100:2010 (E)

The advantage of the formal specification of the request is that the catalogue management system can automatically check and process the request:

Step1: Receive and import catalogue request into internal database; match information in the request with internal representation of the used classification scheme.

Step2: Resolve unknown concepts from the catalogue request via the ISO 29002 web services.

Step 3: Make sure that all requested information can be delivered, i.e. for each target property a mapping from the internal product database exists. If necessary, define additional property mappings.

Step 4: The catalogue can be generated automatically and will be sent to the customer.

It should be noted, that only in step 3 human intervention may be required, and this is only the case if the system recognizes that mappings are missing from the internal database to the required properties in the catalogue request. All the other steps can be performed totally automatically, including the extension of the internal representation of the classification by use of the ISO 29002 web services.

On the requester's site, the received catalogue is processed in a similar way: If it should contain unknown properties, this information is obtained from the dictionary server via ISO 29002 web services, and based on this information, the validity and correctness of the catalogue can be checked automatically.

With such an integration of dictionary access mechanisms and formal specifications of catalogue requirements, a high degree of automation can be achieved which allows for a cost reduction and shorter time to customer specific catalogue delivery.

17 Requirements and usage of ePPS

17.1 Usage of company internal ePPS

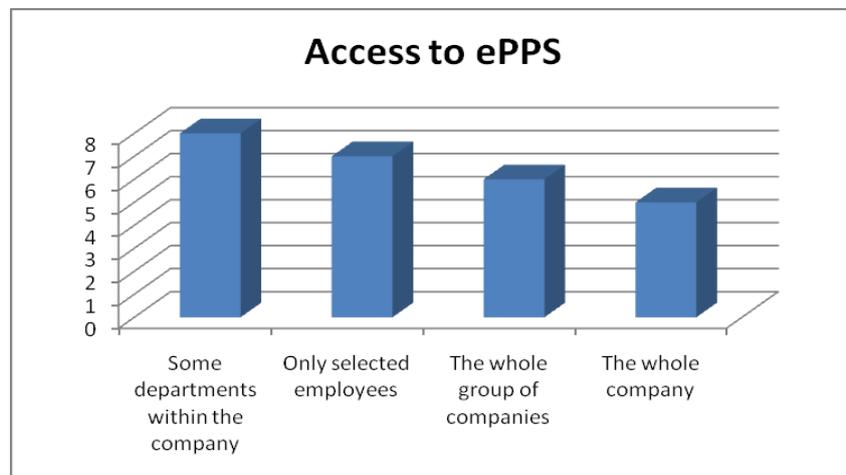
This chapter describes requirements on the organization when using an ePPS system. In the user survey three questions were dedicated to topics regarding the organizational requirements. Below the results of the questions are shown.

17.1.1 Who has access to the ePPS content?

In this question the users shall state which groups of people within the organization have access to the dictionary. Multiple answers were possible. Out of 26 possible answers the following replies were made.

No. of Answers	Access to ePPS
8	Some departments within the company
7	Only selected employees
6	The whole group of companies
5	The whole company

Table 48: Access to ePPS



Picture 63: Access to ePPS

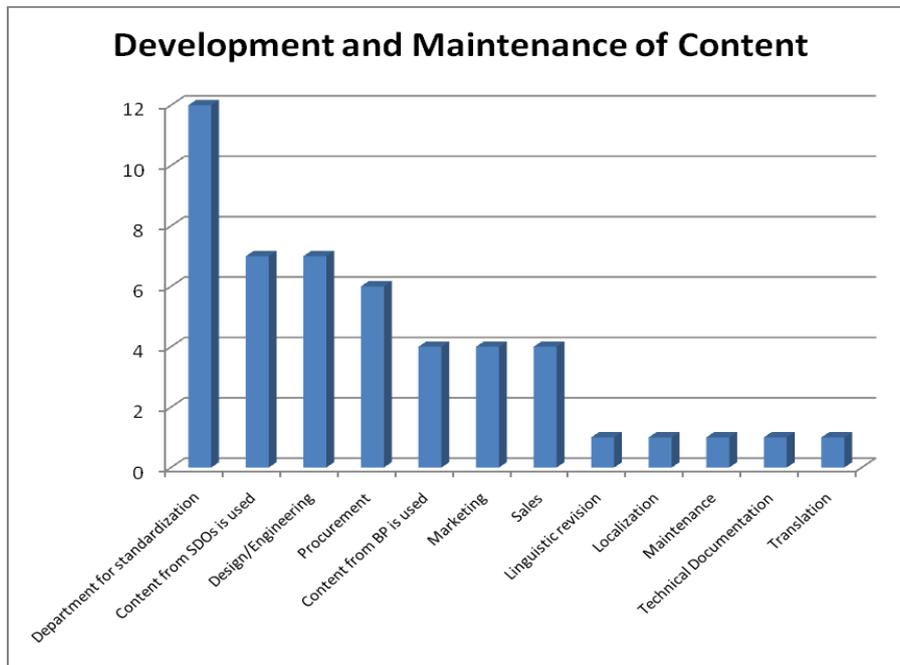
The analysis shows that 19 out of the 26 participants of this part of the survey replied. The companies tend to open the dictionaries to a wider group of people.

17.1.2 Which departments are responsible for defining and maintaining your content?

In this question the responsibility within the organization for development and maintenance of content is questioned. Multiple answers were possible. Out of 26 possible answers the following replies were made.

No. of Answers	Development and Maintenance of Content by:
12	Department for standardization
7	Content from SDOs is used
7	Design/Engineering
6	Procurement
4	Content from BP is used
4	Marketing
4	Sales
1	Linguistic revision
1	Localization
1	Maintenance
1	Technical Documentation
1	Translation

Table 49: Development and Maintenance of Content



Picture 64: Development and Maintenance of Content

18 participants answered to this question. The analysis of the replies shows that in about two thirds of the organizations the standardization department has responsibility for the dictionary content. About a third of the responses state that they take over content from SDOs and therefore the SDOs are responsible for content. Also about a third of the participants state that design/engineering and procurement have content responsibility. Less than a quarter of the statements, name marketing and sales department as responsible or state that they use content from business partners. Users also

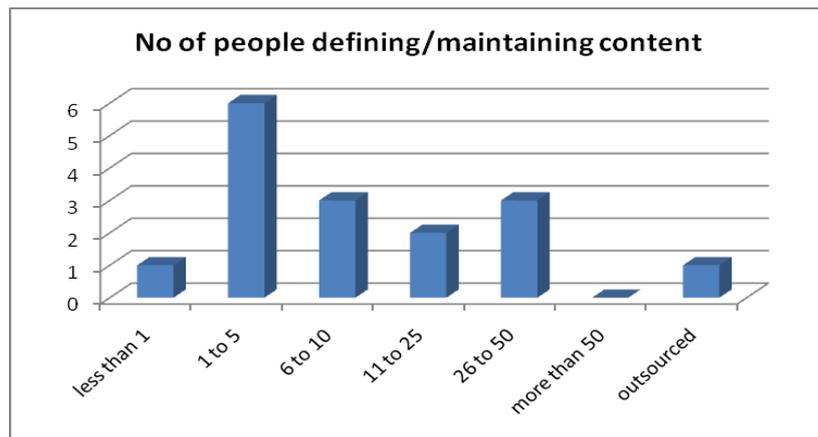
added that people working in the linguistic, localization, maintenance, technical documentation or translation domain are responsible for content within their organizations.

17.1.3 How many employees are working on defining and maintaining the content?

In this question the users shall state how many persons are working on content development and maintenance. The users entered the exact number. For better presentation, the replies were categorized. Out of 26 possible answers the following replies were made.

No. of Answers	Number of employees defining/maintaining dictionary content
1	less than 1
6	1 to 5
3	6 to 10
2	11 to 25
3	26 to 50
0	more than 50
1	outsourced

Table 50: Development and Maintenance of Content



Picture 65: Number of people defining/maintaining content

The analysis of the replies shows that 16 out of 29 potential users responded. Even if the most replies were stated between 1 and 5 users we can see that bigger organizations have up to 50 persons working on content development and maintenance. Therefore it is necessary to have a more detailed look into the various skills necessary to develop and maintain content. This information was requested from the users in the following questions.

17.1.4 Summary

The results of the survey show that development, maintenance and use of dictionary content is a companywide spanning issue. The organizations have learned that well structured product data as well as common terminology reduces complexity of business processes. The data has to be made available throughout the company. Therefore different departments are using ePPS content and are involved in developing and maintaining it.

Besides internal use of ePPS content the companies also need to provide content to suppliers and customers. The importance of standardized interfaces to external organizations and companies increases. As a result to this, also mapping of dictionary data to external structures needs to be improved.

17.2 Usage of external ePPS

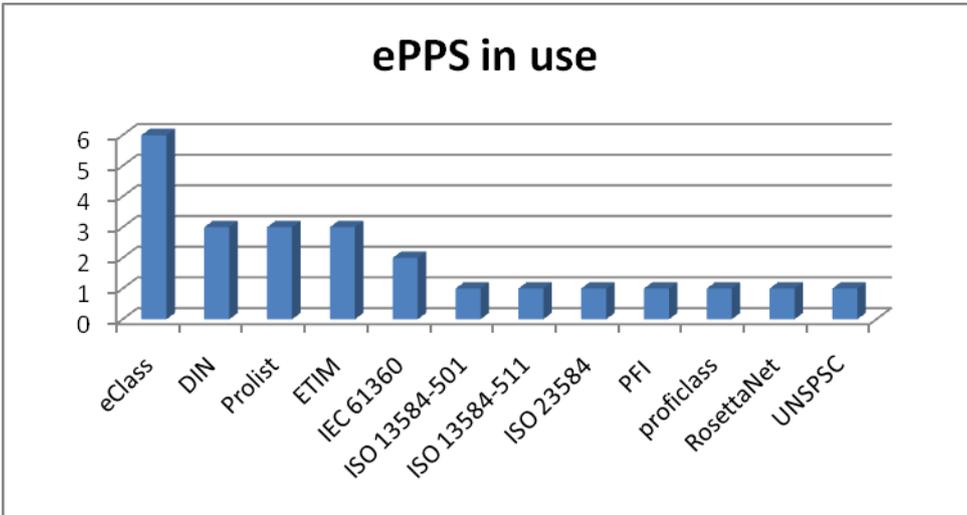
In this chapter, the results from the survey have been collected which are related to the use of ePPS which are provided either by dictionary providers or by business partners.

17.2.1 Which standardization organizations' ePPS do you use?

This question was included to find out which dictionaries are currently in use by the industry. The users were asked to state, which SDO's dictionaries they are using. Multiple answers were possible. Out of the 8 possible responses the following replies were made.

No. of Answers	ePPS in use
6	eCl@ss
3	DIN
3	Prolist
3	ETIM
2	IEC 61360
1	ISO 13584-501
1	ISO 13584-511
1	ISO 23584
1	PFI
1	proficlass
1	RosettaNet
1	UNSPSC

Table 51: ePPS in use



Picture 66: ePPS in use

The analysis of the results of the survey shows that that the majority of the responders use the eCl@ss dictionary. This is not surprising as the majority of the responded came from Europe, and eCl@ss is a horizontal classification system addressing most industry segments. The DIN Property Dictionary, PROLIST dictionary and ETIM dictionary were mentioned three times and IEC 61360

database was mentioned by two participants. The following dictionaries were mentioned once: ISO 13584-501, ISO 13584-511, ISO 23584, PFI, proficl@ss, RosettaNet and UNSPSC.

17.2.2 Which standardization organizations' ePPS do you plan to use?

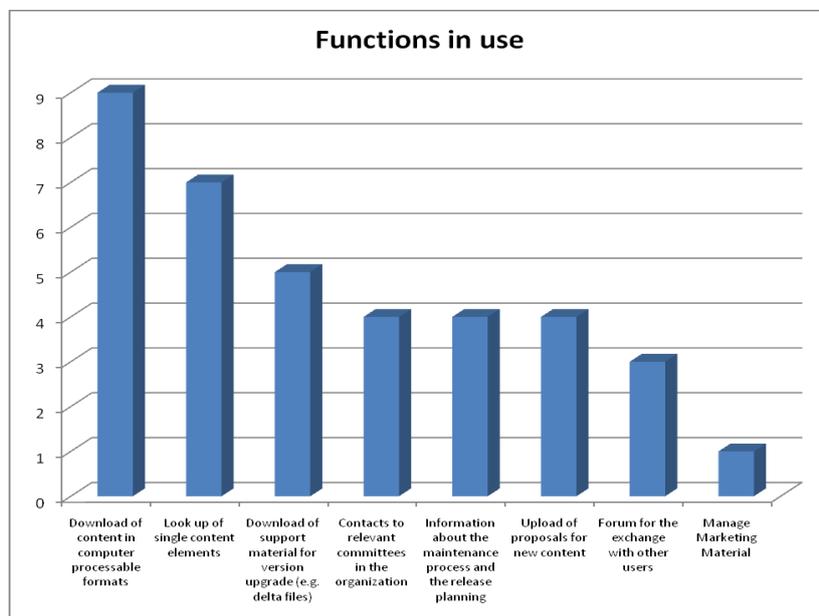
There was one reply from industry to this question stating that one user plans to use eCI@ss in the future.

17.2.3 Which functions of the ePPS of your SDO or business partners do you use?

This question asks for information about the usage of the dictionary in industry. Multiple answers were possible. Out of the 10 possible answers the following replies were made.

No. of Answers	Functions in use
9	Download of content in computer processable formats
7	Look up of single content elements
5	Download of support material for version upgrade (e.g. delta files)
4	Contacts to relevant committees in the organization
4	Information about the maintenance process and the release planning
4	Upload of proposals for new content
3	Forum for the exchange with other users
1	Manage Marketing Material

Table 52: Functions of the dictionary in use



Picture 67: Functions of the dictionary in use

The analysis shows that almost all users download content from the ePPS in computer processable formats. In addition, a majority of users does lookup of single elements. Half of the users mention that they download support material for version upgrades, 40 % use the ePPS to contact relevant committees in the organization, to get information about the maintenance process and to upload proposals for new content. The ePPSs are also used as a forum for the exchange with other users. As an additional entry one user mentioned that he manages marketing material through a dictionary.

17.2.4 Which functions are missing in the ePPS of your SDO or business partner?

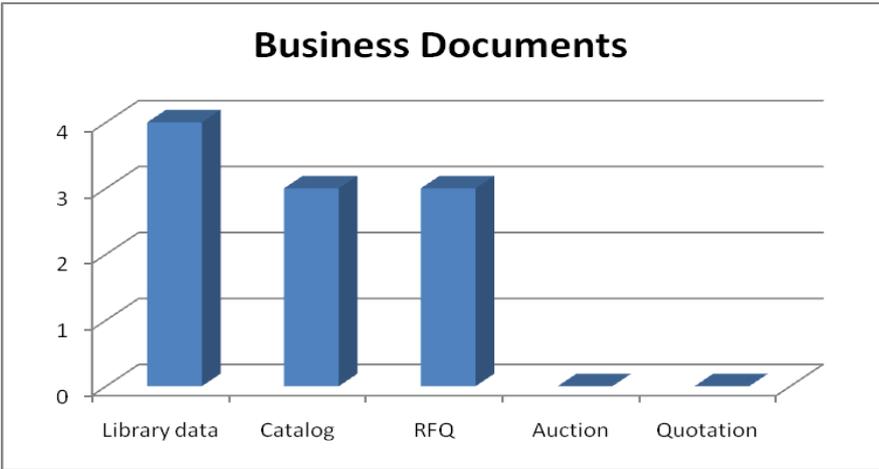
There was only one reply from industry on this question and this user stated that he misses “Data sheet templates for frequently used product classes”.

17.2.5 In which business documents, such as request for quotations [RFQ], quotations, auctions, catalogues do you exchange the dictionary content with your SDO or business partner?

In this question the industrial users were asked to state in which business documents they use dictionary content from SDOs and business partners. Multiple answers were possible. Out of 10 possible answers the following replies were made.

No. of Answers	Business Documents
4	Library data
3	Catalog
3	RFQ
0	Auction
0	Quotation

Table 53: Exchange dictionary content in business documents



Picture 68: Exchange dictionary content in business documents

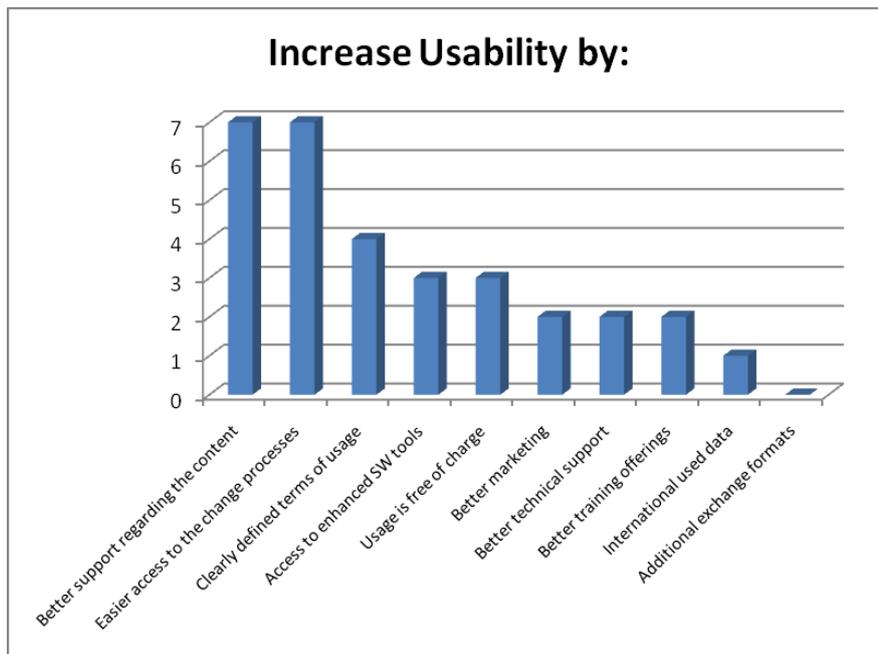
The analysis shows that dictionary content is mostly used with library data followed by Catalogs and Requests for Quotation (RFQ).

17.2.6 What would increase the usability of your SDOs or business partners ePPS?

In this question the industrial users are asked to make statements about how to increase the usability of the dictionaries. Multiple answers were possible. Out of 10 possible answers the following replies were made.

No. of Answers	Increase usability by:
7	Better support regarding the content
7	Easier access to the change processes
4	Clearly defined terms of usage
3	Access to enhanced SW tools
3	Usage is free of charge
2	Better marketing
2	Better technical support
2	Better training offerings
1	International used data
0	Additional exchange formats

Table 54: How to increase usability of the dictionaries used?



Picture 69: How to increase usability of the dictionaries used?

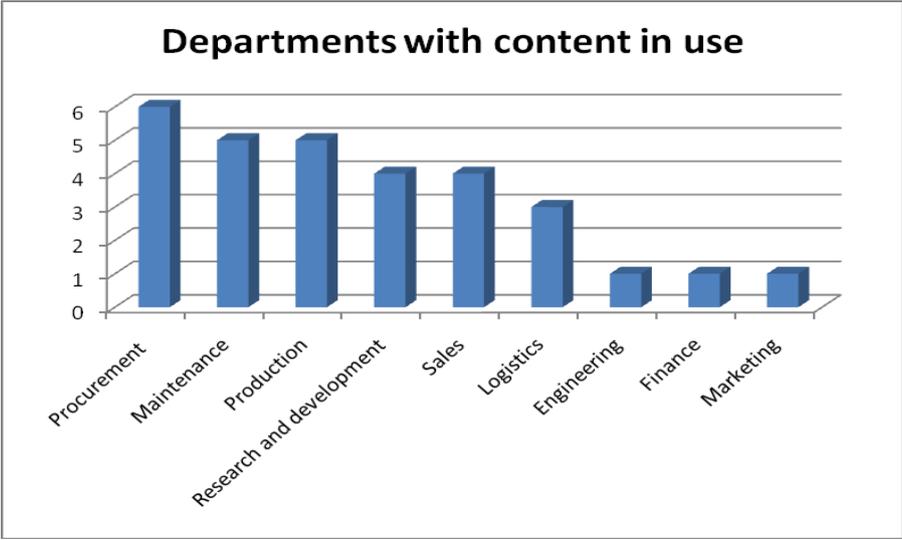
The analysis of the replies shows that the majority of the users both want better support regarding the content as well as easier access to the change process of the dictionary. In addition, the terms of usage were mentioned by 40 % of the users. About one third feels that the possibility to use the content free of charge will increase the usability of the dictionary. Furthermore, access to enhanced SW tools, better technical support, better training offerings and better marketing were mentioned. One user added the provision of international used data would help to increase usability. Interesting is the fact, that nobody requested additional exchange formats.

17.2.7 In which departments of your company do you use the content of the SDO or business partner?

The answers to this question to industry should show in which departments of the organizations dictionary content is used today. Multiple answers were possible. Out of 10 possible answers the following replies were made.

No. of Answers	Departments with dictionary content in use
6	Procurement
5	Maintenance
5	Production
4	Research and development
4	Sales
3	Logistics
1	Engineering
1	Finance
1	Marketing

Table 55: Departments with Dictionary Content in use



Picture 70: Departments with Dictionary Content in use

The analysis of the replies shows that the majority of users use dictionary content in procurement. Dictionary content is also often used in maintenance, production, R&D, sales and logistics. Only one statement was made for usage in Engineering, Finance and marketing.

17.2.8 Which additional functions do you regard as important for an ePPS?

In the provider questionnaire the providers were requested to make statements about what additional functions they see important and which are not yet implemented in their ePPS. From the 17 participants 7 providers replied to this question. The question was formulated as an open question. The comments on which additional functionality the providers regard as important for an ePPS included:

- search functions for content in other databases
- representation of units of measurement
- import and representation of international standardization elements
- logically structured content for better administration
- information, for what purpose the classification standard is really used and detailed remark of the license costs (or remark, that it is free of charge)
- Enable bi-directional links to domain ontology

The statements from providers show, that they see a need for more collaboration on content between different ePPS and the necessary for better statements regarding IPRs and costs.

17.2.9 What are your plans for your ePPS in the future?

In the provider questionnaire the providers were requested to make statements about what future plans they have for their ePPS. From the 17 participants 11 providers replied to this question. The question was formulated as an open question. The comments on which additional functionality the providers regard as important for an ePPS included:

- cooperation with content provider and classification standards
- additional export formats
- content harmonization functions
- Further European expansion
- Increase content of dictionary (classes, properties)
- Improve data quality
- Provide maintenance workflow which enables user suggestions
- Enable bi-directional links to domain ontology
- Move to an enhanced platform and language to facilitate enhancements and user experience.

Also here the replies show that the majority of the providers will invest in new and enhanced functionality for their ePPS and that collaboration, harmonization and process enhancements are key topics.

17.2.10 Summary

The results of the survey show that users of ePPS use a wide variety of dictionary and associated ePPS. Because mostly European users answered the survey, it is not surprising that the European ePPSs are dominating in the result: eCI@ss, DIN property server, Prolist and ETIM are mostly mentioned. The download of computer processable information is the dominant use case (which is in most cases the download of a whole dictionary), but also the lookup of single dictionary elements (probably in most cases via web interface) is mentioned quite often. The ePPS is also used for upload of proposals of new content. Beside these dictionary exchange mechanisms, ePPS are used also for “human related information”, like support material and networking with other users. The dictionary content is mainly reflected in libraries and catalogues and RfQs.

The users mention a number of possibilities for improvement of dictionary servers. In particular, the support regarding the content and the access to the change process seems to need improvements, also the terms of usage seem to be not yet clearly defined. The technical issues seem to be of less importance, e.g. access to SW tools or additional exchange formats are not that often required. It seems, that it is most important for users to clarify organizational and content problems.

The requirements and future plans of the ePPS provides show that they the users can expect both improvements in functionality, content and also processes. Also cooperation and harmonization of content can be expected.

17.3 Personal skills

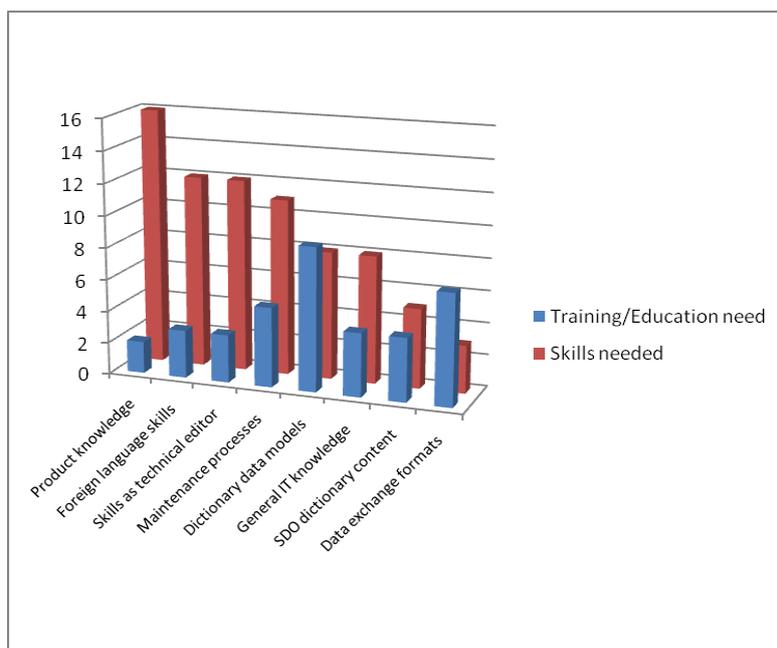
An ePPS system is as good as the processes implemented and the people running it as well as processing content. In this chapter we look into the requirements on the personal skills for operating an ePPS. The user survey had one question to this topic. The results show that the skill profiles for users involved in dictionary work is manifold. The results also show that there is need for education and various trainings to be offered. The conceptual design for trainings and education courses are discussed in the next main chapter.

17.3.1 What kinds of skills are necessary for defining and maintaining the content? Are there still training needs in these fields?

In these questions the participants are asked to state which skills are necessary for content creation and where they still see needs for training in these fields. Multiple answers were possible. Out of 26 possible answers the following replies were made.

Option	Skills needed (No. of Answers)	Training/Education needed (No. of Answers)
Product knowledge	16	2
Foreign language skills	12	3
Skills as technical editor	12	3
Maintenance processes	11	5
Dictionary data models	8	9
General IT knowledge	8	4
SDO dictionary content	5	4
Data exchange formats	3	7

Table 56: Skills and training needs for content definition and maintenance



Picture 71: Skills and training needs for content definition and maintenance

The analysis of the replies shows that the users see many skills necessary for content creation and maintenance. The majority of the users named product knowledge as skills needed. More than 40 %

CWA 16100:2010 (E)

of the possible users for a reply named foreign language skills, an indication that multilingual content is of importance to the organizations, as well as skills as technical editor. Nearly 40 % named knowhow about maintenance processed as important skill for content maintenance. Knowhow on dictionary data model and general IT knowledge were mentioned by about 25 % of the users while knowledge on SDO dictionary content and knowledge on data exchange format was mentioned by less than 20 % of the participants. As additional needed skills the participants mentioned communication skills, business experience, linguistics, analysis of concepts and background in eCI@ss and DINsml.net.

18 Support concepts for implementation adapted to users and supporting software companies

A fundamental prerequisite for the implementation of ePPS accesses and dictionary usage is the awareness of the requirements, future plans and the benefits of using an ePPS. We look in this chapter at the comments which we got from providers and users as a result of the surveys, and we will highlight some benefits which we regard as essential and most important. The second part of this chapter is devoted to two checklists which summarize the important issues to be considered when setting up an ePPS. These checklists are supposed to be used as guidelines for such projects.

18.1 Formulation of benefits of using dictionaries and supporting ePPS

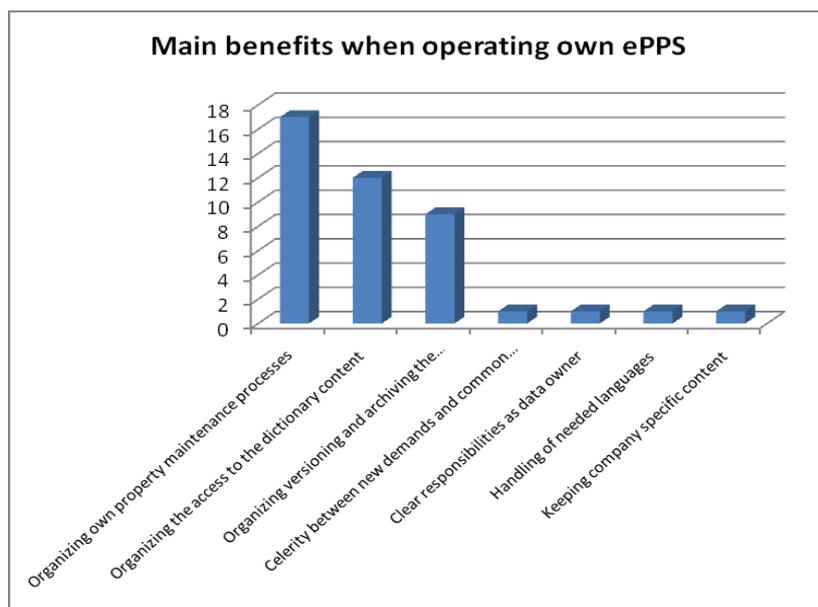
The basic argument for convincing people to change some of their current habits is to identify the benefits of this change. In our survey, this was exactly one of the questions which we asked the ePPS users:

18.1.1 ePPS User Survey: Which main benefits do you see operating an own ePPS?

In this question the participants were asked to make statements what main benefit they see in operating their own ePPS. Multiple answers were possible. Out of 26 possible answers the following replies were made.

No. of Answers	Main benefits when operating own ePPS
17	Organizing own property maintenance processes
12	Organizing the access to the dictionary content
9	Organizing versioning and archiving the dictionaries
1	Celerity between new demands and common solution
1	Clear responsibilities as data owner
1	Handling of needed languages
1	Keeping company specific content

Table 57: Main benefits when operating own ePPS



Picture 72: Main benefits when operating own ePPS

The analysis of the replies shows that the majority of participants see the organization of their own property maintenance process as the main benefit for operating their own ePPS. Also the possibility to organize the access to their dictionary content on their own and the organization of versioning and archiving of dictionaries was regarded as major benefits. In addition, the participants mentioned celerity between new demands and common solutions, clear responsibilities as data owner, possibility of own handling of needed languages and the possibility to also keep company specific content.

Thus, there is a clear list of benefits which is seen by users in running their own dictionary server but what if they use a public ePPS of a dictionary provider? We have not asked this question explicitly, but we assume that similar results would come up. Probably, some of the mentioned benefits would get more important (e.g. management of different languages is definitely easier in the context of collaborative work for a standardized dictionary than within a single company where possibly not enough resources are available).

From our point of view, there are some additional benefits for the case of using the ePPS of a dictionary provider which we see as essential in the future:

- The use of an ePPS for accessing dictionary content provides a number of opportunities for an increase of automation in catalogue exchange and product data related business processes. This was detailed in the section about integration of automatic accesses to dictionary server into business processes.
- An ePPS allows a much broader participation of companies and domain experts in the development of a standard. By providing a web based means for proposing modifications of dictionaries which can be used by any user in any language, the barriers for participating in these processes can be considerably lowered.

18.2 Guidance for planning, establishment, maintenance and usage of an international PPS

Setting up an ePPS system means dealing with technical, organizational, commercial and legal issues. In the analysis of the current implementations the authors of this report found several issues which should be answered more precisely by the dictionary providers. Especially the question about how to deal with intellectual property rights (IPR) is not always answered comprehensively. This might prevent companies and users from using the dictionary content.

In this section we provide some guidance for overcoming these obstacles. We cannot provide the ultimate answers to issues which are not solved yet on a universal basis, but we try to provide guidelines how an international PPS can be planned, established and maintained. We do this both from the provider's and the user's point of view by providing checklists which can be used as a guideline to take the most important issues into consideration when building or using a PPS.

The provider of a PPS has to consider

- how to plan, establish and maintain a PPS,
- which rules and procedures are necessary for running the PPS (e.g. by running a secretariat, hosting the PPS system, supporting online functionalities for maintenance, etc),
- which process to follow to ensure that the PPS can be established as an international product property server

The user of a PPS has to consider

- goals and purposes of using a PPS
- integration of the PPS usage into the internal processes and alignment with the internal data
- legal and financial issues

The following checklists shall help potential dictionary providers and dictionary users as a guideline through the relevant issues when deciding to implement or to use a community or company internal ePPS system.

18.2.1 Checklist dictionary provider

The following checklist was developed to help organizations which intend to implement a dictionary to provide content to their community. By following the checklist, they are guided in developing answers to the various aspects of their ePPS system.

Dictionary providers could be organizations developing and providing classification systems, industry associations providing product descriptions, organizations providing terminology, standardization bodies providing property catalogues, etc.

No	Topic	Comment
1	General	
1.1	What is the objective target of the dictionary?	
1.2	What is the business goal of the dictionary?	
1.3	What business processes are supported by the dictionary?	
1.4	What concepts shall be provided in the dictionary?	e.g. terminology, product classification, product descriptions, product properties, graphical symbols, unit of measurement, codes, ...
1.5	What content shall be covered by the dictionary?	industry spanning, industry segment, application specific,
1.6	Is the target group already using the dictionary content?	
2	IPR/Terms of Usage	
2.1	What kind of license shall be offered?	e.g. public domain, free of charge, commercial license, ...
2.2.	What are the Terms of Usage for dictionary content?	
2.3	May a user also exchange dictionary content to his business partners/customers/suppliers?	
2.4	What information may be exchanged (Ids, names, complete content)	e.g. to internal users and/or to customers and suppliers
2.5	May the user provide the dictionary content in a company internal dictionary?	
3.	Costs	
3.1	What is the pricing model?	e.g. one time license, license per version, license per item, subscription, yearly license, maintenance contract, ...
3.2	What are the costs for different language versions of content?	
4	Content	
4.1	What content is/shall be offered?	
4.2	Is also content from other organizations offered?	If yes, under which conditions/Terms of Usage?
4.3	What is the versioning concept?	e.g. ongoing release of content, fixed versions, different versions (major, minor, service pack), ...
4.4	What is the time planning for new versions of content?	
4.5	Shall update information from one version to another be provided to users?	Useful for users to update content from one version to the next?
4.6	Are there unique identifiers for content elements?	e.g. according to ISO 29002-5, URI, ...
4.7	In which form will the content be offered for users?	e.g. paper (pdf), CD-ROM, local database, online access, ...
4.8	What standardized data models are supported?	e.g. ISO 13584, ISO 8000, ISO 29002, IEC 61360, national standards, industry standards,...

4.9		Content is offered in which data formats?	
4.10		What languages versions are provided?	
4.11		How is the handling of language versions?	
5	Maintenance / Workflow		
5.1		Is there a defined workflow for developing / maintenance of content?	This could be a description of the steps and responsible people/roles of the maintenance process.
5.2		Can users contribute in content development/maintenance?	e.g. by sending change requests per email, per letter, etc.
5.3		What is the technical platform for running the maintenance process?	This might be a manual performance of a process according to a description or an automated workflow run by a system
5.4		How is transparency ensured?	e.g. which kind of feedback will proposers of changes get, how open is the decision process, etc.
5.5		How is the availability of expertise for a high quality development and maintenance ensured?	There is a need to involve domain experts and modeling experts in the development and evaluation processe
6.	Special needs		
6.1		How does the dictionary support users with special needs?	
7	Hosting/Operating		
7.1		What platform do you use for the dictionary?	
7.2		What are the HW/SW requirements?	
7.3		Who is operating the dictionary?	
7.4		What are the service levels for operating the dictionary server?	
8	Secretariat		
8.1		Which organization shall act as secretariat for the dictionary?	
8.2		What kind of roles are needed to run the dictionary server?	e.g. project manager, experts groups, quality group, support, operation, sales/marketing, ...
8.3		What resources are needed for running the dictionary?	
8	Documentation / Training		
8.1		What kind of documentation shall be provided?	
8.2		What kind of training shall be provided?	

Table 58: Questionnaire dictionary provider

18.2.2 Checklist dictionary user

The following checklist was developed to help companies which intend to use content of a dictionary to develop answers to the different aspects for the ePPS system.

Dictionary users could be companies describing products, suppliers delivering electronic catalogues, companies using terminology, etc.

No	Topic	Question	Comment
1	General		
1.1		What is the objective target of the dictionary?	
1.2		What ist the business goal of the dictionary?	
1.3		What business processes are supported by the dictionary?	
1.4		What concepts are provided in the dictionary?	e.g. terminology, product classification, product descriptions, product properties, graphical symbols, unit of measurement, codes, ...
1.5		How is the fit of the provided content to my product/service offering	
1.6		Are my competitors/suppliers/customers using this dictionary?	
1.7		What is the current user basis of the dictionary?	
1.8		What is the relevance of the dictionary in my industry?	
2	IPR/Terms of Usage		
2.1		Under which license the content is offered?	e.g. public domain, free of charge, commercial license, ...
2.2		What are the Terms of Usage for the content?	
2.3		May I also exchange dictionary content to my business partners/customers/suppliers?	
2.4		What information may be exchanged (Ids, names, complete content	e.g. to internal users and/or to customers and suppliers
2.5		May I use the dictionary content in my company internal dictionary?	
3	Costs		
3.1		What is the pricing model?	e.g. one time license, license per version, license per item, subscription, yearly license, maintenance contract, ...
3.2		What are the costs for different language versions of content?	
4	Content		
4.1		What content is/shall be offered?	
4.2		Is also content from other organizations offered?	If yes, under which conditions/Terms of Usage?
4.3		What is the versioning concept?	e.g. ongoing release of content, fixed versions, different versions (major, minor, service pack), ...
4.4		What is the time planning for new versions of content?	
4.5		Is update information from one version to another offered?	e.g. mapping tables, transaction update files, ...

4.6		Are there unique identifiers for content elements?	e.g. according to ISO 29002-5, URI, ...
4.7		In which form is the content offered?	e.g. paper (pdf), CD-ROM, local database, online access, ...
4.8		What standardized data models are supported?	e.g. ISO 13584, ISO 8000, ISO 29002, IEC 61360, national standards, industry standards,...
4.9		Content is offered in which data formats?	
4.10		What languages versions are offered?	
4.11		How is the handling of language versions?	
4.12		How can we contribute in content development/maintenance to better support our needs?	
5	Implementation		
5.1		Are interfaces offered to import content to my systems?	
5.2		What is the offering for supporting system implementation?	e.g. training, documentation, implementation partners, tools, ...
5.3		How can you integrate the ePPS into your internal processes?	
5.4		How can you integrate and align the ePPS data into your internal data structure?	
5.5		Who will be involved internally in implementing dictionary content?	e.g. product development, procurement, material management, sales/marketing, ...
6	Special needs		
6.1		How does the dictionary support users with special needs?	
7	Documentation / Training		
7.1		What kind of documentation has to be provided internally?	
7.2		What kind of training has to be provided internally?	
8	Other		
8.1		Who is the organization providing the dictionary?	
8.2		What is the long term business strategy of for the dictionary	

Table 59: Questionnaire dictionary user

19 Strategies and activities for spreading the knowledge into industry

In this chapter we describe the strategy which has been followed in the ePPS project to spread this knowledge about dictionaries and dictionary servers into industry. An important aspect is the training of people who have to deal with meta data and ePPS. This aspect is addressed by the CEN Focus Group eSME launched in April 2009. The CEN Focus Group eSMEs aims at providing training material and tutorial on eBusiness standards. In the second part of this chapter, we will give brief overview about the eSME training material strategy.

19.1 Activities within project ePPS

The use of dictionaries and classifications, in particular those which are defined in a consensual process, is still not exercised by all companies and all industries. Single companies can gain a number of benefits by implementing their internal ePPS or by adopting standards like eCI@ss for internal product organization. The most benefits for industry as whole however, can only be gained if dictionaries, classifications and ePPS' are used on a broader scale and if the use of dictionaries is taken for granted in any exchange process between companies. Therefore, it is important to organize the knowledge transfer from experts to the industry on a European scale in order to improve the awareness and usage of these methods in industry practices.

The ePPS strategy to transfer this knowledge into industry is twofold:

- **Industry seminars:** Together with the sister Focus Group CDC, ePPS conducted 3 industry seminars with the goal to present success stories of ePPS systems which are provided by dictionary and classification providers.
- **Collaboration with projects and associations:** ePPS has worked together with a number of projects and associations in order to build a basis for a sustainable information flow into industry.

Some existing examples of the activities of ePPS are the following:

- A sister project within the eCAT Workshop has been defined with the support of ePPS which works on the topic of interoperability among the most important dictionaries and classifications. The worry of many users is that they have to support a big number of different dictionaries in their exchange processes. By providing mechanisms which allow the "navigation" from one dictionary to another one in an easy way a very important obstacle for entering the dictionary use can be removed.
- Collaboration with the Peppol project on European e-procurement. The ePPS experts have worked together with the Peppol experts on ePPS related matters acquainted them with the CC3P project and established a collaboration between Peppol and CC3P on the mapping of dictionaries and classification systems.
- ePPS has used the eBIF Focus Group to discuss the importance of dictionaries in the context of e-business. This will be a permanent topic in this group.
- ePPS experts have been involved in the constitution of two further CEN projects which are very important for spreading the knowledge and the use of dictionaries in industry and in particular by SMEs:
 - eSME is a CEN Focus Group which has the goal collect material for raising awareness of the importance of standards in e-business, and its focus is on bringing this information to SMEs. A specific topic here is the development of tutorial material about standards in the area of dictionaries and product data exchange. This is described in more detail in the following subsection.

- Orchid is a CEN Workshop which has the goal to provide a roadmap and an implementation about the use of standards in the process industry for all product and life cycle data related processes. One key element is the support of dictionaries and dictionary servers and the specification of their role in the various exchange processes.
- In cooperation with the Focus Group CDC, ePPS has supported the implementation of an implementation forum in ISO TC184 SC4 where implementations of SC4 standards including dictionary standards and dictionary servers can be discussed. The final ePPS conference took place in the context of the first implementation forum at the ISO TC184 SC4 meeting in Rotterdam on November 11/12.

19.2 Conceptual design of training and educational courses in the CEN Focus Group eSME

In the CEN Focus Group eSME, a methodology for training courses in the area of product data exchange by use of dictionaries and classifications is under development which is very much influenced by the activities of the ePPS project. Thus, this chapter has been developed in close cooperation with the eSME experts in this area.

To define a tutorial for a specific group of people, a number of considerations have to be taken with respect to purpose, target group, level of expertise of the tutorial participants, topic to be discussed, etc. These elements can be abstracted into a number of dimensions, and the following four dimensions have been identified:

19.2.1 Dimension D1: Topic

The first dimension defines the knowledge area where the participants need to get information. Our target area is product data modeling and exchange, so that we have to identify subtopics. Some examples of potential topics are the following

- Product Classes, properties and their relationships
- Blocks / complex properties
- Visible and applicable properties
- Relationships between classes (is-a, case-of, component relationships)
- Identification mechanisms for classes and properties
- Management of units
- Dealing with multiple languages
- Web Services to access dictionary
- ISO 8000 requirements for electronic catalogues
- Exchange formats for exchanging catalogue data

The elements in the list may overlap, they may vary in their scale, and they may be related to standards or to use cases of users. The list is tentative, open and will be extended and refined in the future.

19.2.2 Dimension D2: Target Group

The second dimension defines the group of learners. Depending on this target group different information has to be brought across. Some examples:

- **Managers, decision makers:** Here it is important to highlight the purpose of the standard or standard elements, in which context can they be used, what is the benefit of using the standards, etc.
- **Standards users:** Users need information how they can integrate the standards in their processes and what they have to do to adapt or migrate to standards. Users need an “operating” knowledge about the standards so that they can use tools which are based on the standards.
- **Developers of standards:** Developers of standards (in our case e.g. members of committees who define a specific part of a content standard like IEC 61360-4 or eCI@ss) need knowledge about the elements they can use to model their product classes of interest. Thus they need examples in which case which modeling feature should be used, and which modeling features can be used for which modeling cases.
- **Software developers:** Software developers need a very deep understanding of the details of standards and their underlying concepts, in particular they need knowledge about formats, concepts and representations which are required for the implementation of the standard into a system. As such they often have also to know what a specific modeling element is used for from a user’s perspective in order to be able to implement a meaningful user interface.

The basic idea is to define modules which describe a specific topic from the viewpoint of a specific target group. To produce a complete tutorial it is very likely that modules have to be combined to provide the complete set of information required by a specific audience. For instance, it might be sufficient for a short tutorial for managers to just use the managerial modules, but for a user it might be reasonable to combine the managerial modules with the user modules (i.e. the user modules should avoid as far as possible to repeat content which is already part of the managerial modules). Similarly, a tutorial for standard developers could possibly contain in addition to the standard developer modules the managerial modules and possibly also the user modules. A tutorial for software developers could possibly contain all modules about the topic. These examples show that by providing target group oriented modules a big flexibility can be achieved, and many more combinations are possible.

19.2.3 Dimension D3: Level of expertise of audience

The level of expertise may be another dimension. It mainly describes (1) how much does the audience already know about the topic and (2) what is the technical background of the attendees. It might need different kinds of material for computer scientists than for mechanical engineers or procurement experts. Thus, the modules should indicate what level of pre-knowledge and which kind of background they expect from the participants.

19.2.4 Dimension D4: Context of tutorial

It is planned to reuse the modules for different contexts. Contexts may comprise various variations of the environment of the tutorial or the special goals of the tutorial. For example, a tutorial which is produced for PLIB users should be easily adaptable to a tutorial for eCI@ss developers because eCI@ss has based its underlying data models on PLIB, but uses some renaming or restrictions. In a similar way, tutorials may be adapted for different industry branches (e.g. by using different examples), or they may be adapted by translation to other countries. Thus, this dimension can be spread into several sub dimensions like language, target standard variant, industry branch, etc.

19.2.5 Matrix of dimensions to position tutorial modules and materials

Based on these dimensions, a matrix could be drawn which allows to position each tutorial element and allows a trainer to select the appropriate modules which then can be combined to prepare a tutorial which best fits the actual requirements. We can also see it the other way round: Each module becomes a profile which indicates where it can be used and for which listeners it is best suited.

One of the tasks of eSME is the development of tutorial elements according to this methodology for the area of product data exchange and description, and PLIB as a basic data model standard for dictionaries and exchange formats is the basic tutorial topic. By the end of eSME, both an organization structure and a number of tutorials will be available, and in the context of the eBSN network, this will hopefully grow and been used for training in the area of dictionary management and use.

19.3 Building a concept for evaluating the results in industrial practice for a longer period of time

During the work for this project the project team spent a lot of work on discussions with dictionary providers and dictionary users to better understand their needs and experiences in operating or using an ePPS. Also two surveys were developed, conducted and analyzed. It was challenging to contact the right people outside the known community to reply on the provider questionnaire. For the user questionnaire it was not possible to track the final number of people who were addressed, as the information was spread through different channels and due to privacy laws both companies and organizations do not give feedback whom they addressed. Sometimes not even the numbers of receivers were reported.

The people and organizations who replied on the surveys gave very valuable feedback on their usage, lessons learned and needs in running an ePPS or operating one. The results are included in this report. To make this information available for the whole community and to continuously add new findings and experiences we suggest to take action for building a community able to continuously provide and exchange information and findings. We suggest the following actions:

- Develop and implement an online platform for the ePPS community. On this platform both ePPS providers and ePPS users should be represented. Within a moderated environment information and findings about providing and using content in an ePPS should be made available. Within discussion forums the community can discuss issues of interest. A monthly newsletter should inform about news, publications and events to ePPS topics.
- Develop a structured online questionnaire to survey the learnings and requirements from the industry. The survey shall be sent out periodically (e.g. once a year). Due to the periodical feedback the results may also tell the developments in usage, satisfaction, needs and business benefits. As basis for this questionnaire the questionnaires developed for this project may be used. Interested ePPS users may register on the online-platform to be included in the survey. Additionally calls for participation on the survey may be issued through press releases and organizations, standardization bodies and EU information channels.
- Organize an annual conference where both ePPS providers and ePPS users are invited to present their experiences with a focus on business benefits. In addition, the results of the annual survey shall be presented at the conference.
- Setup and organize regular online meetings where ePPS users and providers present their projects and benefits. The meetings may take place on a monthly basis and last for 45 to 60 minutes consisting of the introduction by a moderator, the actual presentation and a Q&A session respectively a discussion at the end.
- Offer workshops for ePPS providers and users on focus topics like IPR, workflow, data exchange formats, release management, etc. Within the workshop, the participants may also present their status quo in using or operating an ePPS.

The development and setup of the platform may be initially funded through an EU project while the ongoing operation may be financed by sponsoring, organization of workshops and the annual conference.

Annex A: Survey on Dictionary Providers

1. Organization of the Questionnaire

The goal of the survey was to get some empirical results about the actual practices in the operation of dictionary servers, or more general, of meta data servers. For this survey, a questionnaire has been developed by the ePPS project team. The questionnaire addressed a number of different aspects which are related with a meta data server, including:

- data models of meta data servers
- exchange formats used for exchanging dictionary information
- (Web-) services provided by dictionary servers to users of the server
- means provided for interoperability with other dictionary servers and users of the dictionary server
- means and experiences with respect to multilingual aspects and access for users with special needs
- means to support the maintenance of the meta data
- copyright issues
- licenses
- business model

The sending out of the questionnaire started in summer 2008. It was sent to 50 meta data providers (see list in appendix 1), and it took three months to get feedback from 12 providers. Over the next 4 months (up to March 2009), another 5 responses were received by the project team. The questionnaire is attached.

The questionnaire is organized as follows:

1. Information about the organization providing the dictionary and the kind of data captured in the database.
2. Relationships to other organizations: Here the goal is to get information about the current need for interoperability support.
3. Functions for update and maintenance of the dictionaries: Is there a workflow for maintenance defined? How does it look like? Is it supported by the system? In which granularity can users upload their maintenance requests and proposals for new content?

4. Functions for publication of content: In this section, information about all services towards the users is gathered: release cycles, ways to access the content, services provided to help users and user communities, etc.
5. Terms of usage / copyright: Collects information about the rights which users gain when they obtain the content of the dictionary.
6. Data models and formats: Information about the data model which is used to represent the dictionary, about the identification of dictionary elements (e.g. with respect to the requirements of ISO 8000-110), and about formats in which dictionary content can be downloaded and uploaded.
7. Multilingualism, users with specific needs: Information about languages provided, elements and their translatability (translatable and non-translatable fields), problems encountered by the dictionary providers with multilingualism beyond translation (e.g. non existing concepts in a country or a community), and means provided for users with special needs.
8. Additional functions and plans for the future: here the dictionary providers are asked to give information on additional functions they provide and they regard as important, and on where they plan to go in the future.

2 Analysis of Dictionary Provider Questionnaire

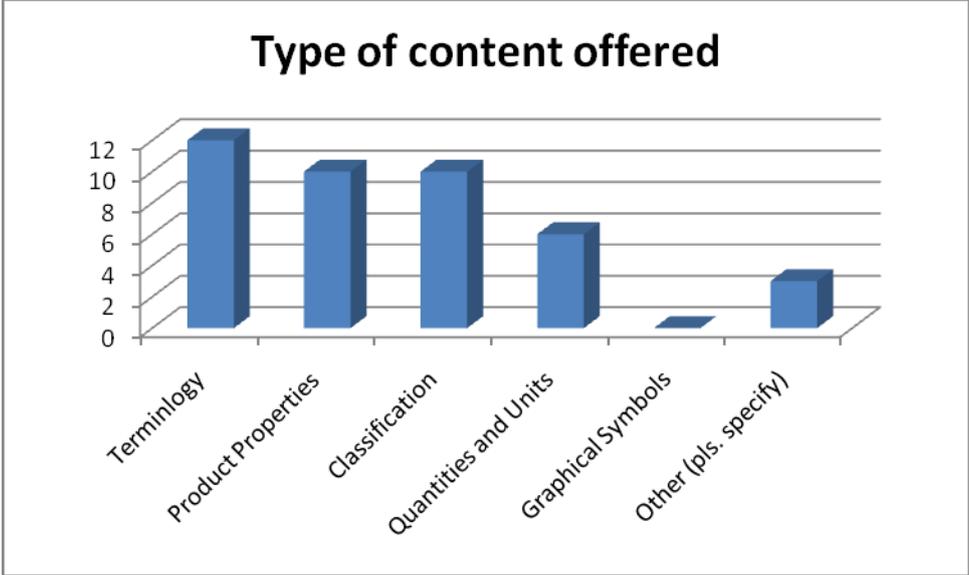
This chapter includes general information about the systems and environment the participating ePPS providers are working in.

What type of content is offered?

The result of the survey was that all 17 providers replied on this question. They currently offer the following content in their systems:

No Answers	Option
12	Terminology
10	Product Properties
10	Classification
6	Quantities and Units
0	Graphical Symbols
3	Other (pls. specify)

Table 60: Type of content offered



Picture 73: Type of content offered

Even if ten providers claim to describe properties and ten to describe classifications, there is not a total match between these both sets. Thus, some describe product properties without having a classification, and some provide a classification without describing properties. But, of course, many provide both a classification of products and a description of products by properties. Almost all providers see themselves as terminology databases, Twelve of the responders claim that they describe terminology in their database. Here different providers of similar nature seem to interpret their role differently: Some who provide product properties and/or classifications do not see this as terminology information, others do. Some of the providers are pure terminology providers, they provide more or less only terms with their definitions.

What is the volume of the offered content (e.g. no of elements handled by your ePPS)?

The result of the survey was that all 17 participants replied. The number of entries in the meta databases vary from a few hundred entries to hundreds of thousands of entries. Pure terminology databases tend to have the biggest volume of entries, but also where numbers of classes and properties are counted separately, class numbers of more than 30.000 and property numbers of more than 10.000 exist. Of course, branch-specific meta data sets are smaller but may still contain a couple of thousand classes and more than 1000 properties.

How many users use the content of your ePPS?

The result of the survey was that in summary all 17 participants replied.

No Answers	Option
6	Replies on number of downloads
6	Replies on number of registered users
4	Replies on number of purchased licenses
6	Other replies

Table 61: Use of content

A big variation of answers has been given. Some providers do not know how many companies and users they have, simply because they provide their meta data set freely on the web for download and do not track who downloads the data. Or the providers do not know the exact user number because the meta data is distributed via several servers and integrated in other software platforms. Other providers can give different kind of information about their user base depending on the business model: some providers just count downloads and other providers know their user base because users have to register or even to buy licenses.

The number of downloads is normally in the order of some thousands, in some cases even up to 6000 per week.

The number of registered users (if applicable) is in the order form about 25 companies (very specific dictionary) to a few hundred (consortium specific) up to a few thousands for horizontal dictionaries.

In a few cases, licenses are sold. It is not always clear whether this is related to the registration or whether the registration implies the achieving of a license. Here up to a about 1000 licenses have been sold for some dictionaries.

Do you have connections/cooperation with other providers of content?

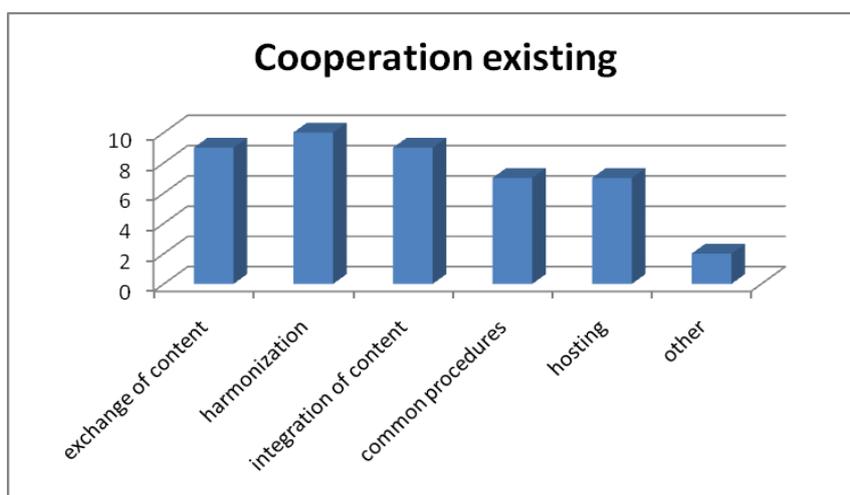
The question was split into two sections to get answers on co-operations which already exist and those which are planned for the future.

Existing co-operations

The result of the survey was that 13 providers replied on this question. They made the following statements on existing co-operations:

No Answers	Option
9	Exchange of content elements
10	Harmonization of parts of the content
9	Integration of parts of the content into the content of the other organization
7	Common procedures for maintenance and evolution
7	Hosting
2	Other (pls. specify)

Table 62: Cooperation existing



Picture 74: Cooperation existing

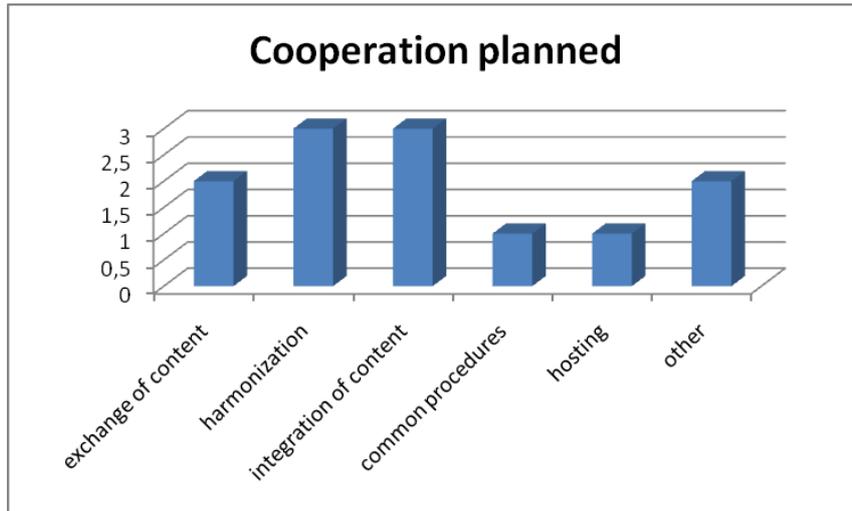
The results show that the major focus on existing cooperations are content wise either to exchange, integrate or harmonize content. Furthermore cooperations exist to work on common procedures and to host the systems. Two providers made other statements related to their cooperations with standardization bodies and other organizations.

Planned co-operations

The result of the survey was that 13 providers replied on this question. They made the following statements on planned co-operations

No Answers	Option
2	Exchange of content elements
3	Harmonization of parts of the content
3	Integration of parts of the content into the content of the other organization
1	Common procedures for maintenance and evolution
1	Hosting
2	Other (pls. specify)

Table 63: Cooperation planned



Picture 75: Cooperation planned

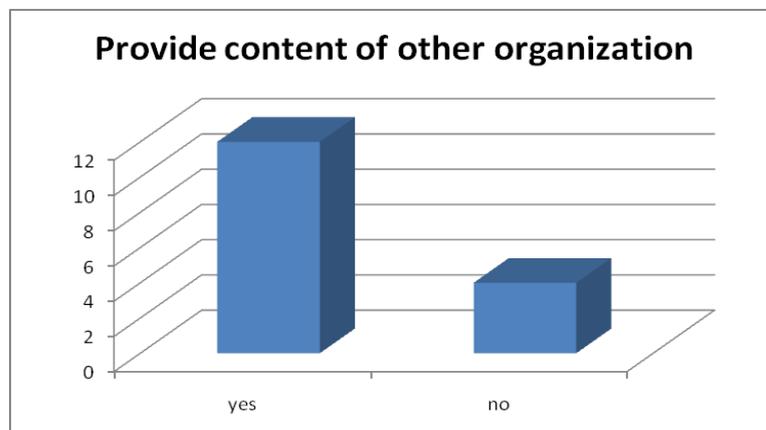
Many providers have a lot of cooperation with various other dictionary providers. There seem to exist different strategies: Some have established formal relationships (e.g. eCI@ss, proficl@ss, ETIM, Prolist, DIN, UNSPSC, IEC, ECALS, RosettaNet), others provide platforms for collecting different elements in a single system (like ECCMA, e.g. integrating the FUNSTEP content). In some cases one dictionary provider is responsible for a chunk of the dictionary of a bigger provider (e.g. PIDX for UNSPSC or Prolist for eCI@ss). Some mention that they have a co-operation with other organizations (no dictionary providers), e.g. for hosting or for delivering common procedures.

Do you provide content of other organizations (e.g. ISO/IEC standards, national standards, or industry standards)?

The result of the survey was that 16 providers replied on this question. They made the following statements:

No Answers	Option
12	Yes
4	No

Table 64: Provide content of other organization



Picture 76: Provide content of other organization

There is some reuse of content from other organizations which goes along the lines of question 8: Some of the relationships and co-operations include the sharing of content (sometimes the harmonization of content). Sometimes this means pointing to other dictionaries for specific product areas. ECCMA, to name one special case, looks for collecting metadata from different sources from standardization organizations and industrial consortia.

Request on Reference customers

The ePPS providers were also requested to provide reference users of the dictionary. Only 6 providers made references. Some providers mentioned that they do not track usage. Other providers stated that they are not providing data of users but they are willing to forward questions to their community.

3. Provider Questionnaire

CEN/ISSS Questionnaire:

How do Providers of dictionaries support their users with electronic systems?

**Version 0.8
2008-06-06**

General Questions

Organization: _____

Organization residence: _____

Questionnaire answered by: _____

We are pleased to provide the results of the survey. Please enter your contact details.

Contact Details (optional):

Telephone: _____

Email*: _____

* necessary if you like to receive the results of the survey

1. What is the name of your dictionary / content?

2. How to access your content? Is there a Web address?

3. What is your target group?

4. Who is the owner / supporter / sponsor of your organization?

Questions regarding the content

5. What type of content is offered?

- Terminology
- Product properties
- Classification
- Graphical symbols
- Quantities and units
- Other (please specify) _____

6. What is the volume of the offered content (e.g. number of elements handled by your ePPS)?

7. How many users use the content of your ePPS? (Please specify appropriate figures)

- Number of downloads _____
- Number of registered users _____
- Number of purchased licenses _____
- Other (please specify) _____

Inter-organizational relationships

8. Do you have connections / cooperation with other providers of content?

Cooperation about	Existing	Planned	With which organization?
- the exchange of content elements?	<input type="checkbox"/>	<input type="checkbox"/>	_____
- harmonization of parts of your content?	<input type="checkbox"/>	<input type="checkbox"/>	_____
- integration of parts of the content of your organization into the content of the other organization (or vice versa)?	<input type="checkbox"/>	<input type="checkbox"/>	_____
- common procedures for maintenance and evolution	<input type="checkbox"/>	<input type="checkbox"/>	_____
- Hosting	<input type="checkbox"/>	<input type="checkbox"/>	_____
- Other (please specify)	<input type="checkbox"/>	<input type="checkbox"/>	_____

If possible, provide some information on the nature of the cooperation.

9. Do you provide content of other organizations (e.g. ISO/IEC standards, national standards, or industry standards)?
-

Functionality

Update and maintenance

10. Do you support a workflow for development and maintenance of content?

yes no

If yes:

11. What kind of workflow is implemented?
(If possible, give (link to) more detailed information)

- Workflow according to ISO procedure for paper documents
 Workflow according to IEC procedure for paper documents
 Workflow according to ISO/IEC common procedure for database standards
 Other (please provide description)
-

12. Is the maintenance procedure supported by your ePPS system?
(If yes, please add some information on how this done)

yes no

13. Do you provide means for upload of proposals for new content elements?

- Upload of single content elements
 Bulk upload of sets of content elements
 Other means, e.g. CD (please specify)
-

If no:

14. How do you deal with modifications of your content?
-

Publishing of content

15. Do you have a fixed or scheduled release cycle where new versions of the content are published?

yes no

16. If you publish your content in a release cycle, how often do you provide new versions of your content (e.g. monthly, quarterly, annually, or other period)?
-

17. If you do not have a release cycle, how do you publish new or modified content:
-

18. In which form do you publish your content?

Download via web site
 Access via web Services
 Other publishing means, e.g. CDs (*please specify*)

19. In which granularity do you provide your content?

Complete content
 Part of the content (e.g. some product groups)
 Single elements
 Others (*please specify*)

20. Which additional information do you provide for download (e.g. help information, white papers, support material for version upgrade (delta files), information about the maintenance process and the release planning, etc.)?
-

CWA 16100:2010 (E)

21. Do you provide means for community building (e.g. news, dates of events, forums for discussions, contact means to committees in your organization, ...)?

22. Do you provide additional services for your users or your committee members?

23. How is your system organized (e.g. single centralized server, distributed servers, several content collections on one server, etc.)?

Terms of Usage/Copyright

24. What are the rights of your users? What are they allowed to do with the content they have downloaded / incorporated into their own system?

- Use within own company
- Use in collaboration with partners
- Incorporate in their own software
- Distribute
- Resell
- Modify
- Others (*please specify*)

25. What kind of license do you provide for accessing and using the content to users and committees? (*please attach license agreement, if possible*)

For which user group	Committee members	Members of the association	Non members	Other group	Other group	Other group
Kind of license				_____	_____	_____
- Public domain?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- Free license (e.g. GNU license)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- Commercial license?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- Other form of license	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- No license	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

26. What is your price model (e.g. free of charge, license fee: one-time, annual, etc.), registration with registration fees, others, please specify)?

If possible, please attach the price list.

Data models and formats

27. Do you have a formally specified data model which describes how your content is organized?

yes no

If yes:

28. Is your data model based on other models?

- PLIB (ISO 13584 / IEC 61360)
 OTD (ISO 22745)
 Other (please specify)

29. Which mechanism do you use for specifying your data model (e.g. XML DTD, XML Schema, UML, tables, others)?
(Please provide a description or a link to a description, if possible)

If no:

30. How is the structure of your content specified?

31. Have you ensured that each element of your content has a world-wide unique identifier? *(Please specify the organization of your identifiers. This is one of the prerequisites to support the data quality of catalogs as defined in ISO 8000-110)*

33. In which format(s) can users receive your content? In which formats can users send proposals for content elements to the server?
(Please provide descriptions of your formats or a link to such a description.)

Usable for Format	download	upload
Human readable forms		
- PDF		
- brochure, flyer, other paper media		
- Other human readable form		
Computer readable formats		
- Set of tables		
- BMEcat	<input type="checkbox"/>	<input type="checkbox"/>
- OntoML	<input type="checkbox"/>	<input type="checkbox"/>
- Others		

Multilingualism, users with specific needs

34. Do you provide your content in several languages? Which ones?

35. Do you provide different language variants? How do handle language variants (e.g. British and American English)?

36. Do you have a leading (mandatory) language? Which one?

37. Do you distinguish between translatable and non-translatable fields? (e.g. codes may be not translatable, names may be translatable, please specify your rules)

38. Which kinds of character encodings do you support (e.g. UTF-8, ...)?

39. Have you encountered localization problems beyond translation? Examples: Need of different structuring of content for different countries / regions / cultures, different properties for different countries / regions / cultures, etc. If yes, how have you dealt with them?

40. Do you provide your content in a form, which supports users with special needs. If yes, please specify

Miscellaneous

41. Which additional functions do you regard as important for an ePPS? Please specify.

42. What are your plans for your ePPS in the future

Please provide names of some of your reference users:

(answering this question will help us to do a similar survey about the users of ePPS)

Company name	Contact person	Email address	remarks

Annex B: Survey on Dictionary Users

1. Organization of the Questionnaire

To examine the state of the art of the current use of dictionary servers in the industry and get the requirements from industry on the provision of dictionaries a questionnaire was developed and a survey was accomplished. The goal of the survey was to get some empirical results about the actual practices in the usage of dictionary servers, or more general, of meta data servers from industry perspective as well as get information about needs and future requirements from industry perspective. For this survey, a questionnaire has been developed by the ePPS project team. The questionnaire addressed a number of different aspects related to meta data servers, including:

- Information about the organization using dictionaries
- Kind of dictionaries in use and planned to use
- Functions used and usage of dictionary data
- IPR issues
- Requirements on further functionality
- Who is using dictionary data and who has access to
- What kind of skills are necessary to define content
- Benefits of operating an own dictionary
- Information about own operated dictionaries

The questionnaire was developed and then implemented as an online-survey. An online survey tool was made available by courtesy of USPI-NL. The online survey was made available to participants in April 2009 and was online until End of August, 2009.

An invitation to participate in the survey was sent out to probably more than 3000 potential dictionary users by different organizations. The organizations informed their users directly through publication on their web-pages, newsletters and other communication. A detailed number is not available as the organizations did not give feedback about to whom the information was distributed.

To motivate the users to participate in the survey a drawing was included in the online-survey. The prizes were an EUR 50,- voucher for a big online shop (books, music, etc.) and 5 USB-Memory-Sticks with 4 GB.

By August 25th, 133 results were entered in the online-survey from which 36 users responded on the complete questionnaire. After analyses of the incomplete answers it was decided to only use the full responses for the analysis. In the incomplete responses only a few entries were made and so it was not useful for the analysis.

The Questionnaire is attached and organized as follows:

- 1 Information about the organization using the dictionaries
- 2 The information requested from the users was questioned depending if the dictionary in use is:
 - a. from a SDO (standards developing organization)

- b. a business partner or
 - c. an own dictionary
- 3 When a dictionary from a SDO or a business partner -is in use we asked:
- a. What kind of content is in use
 - b. Content from which SDOs is in use
 - c. Kind of use of dictionary (functions)
 - d. Functions missing in dictionary
 - e. Formats used for data exchange
 - f. Usage of dictionary data
 - g. Update processes
 - h. Terms of usage
 - i. Additional requirements for dictionaries
 - j. Which departments use the content from the dictionaries
- 4 When an own dictionary is in use we questioned:
- a. What kind of content is in use
 - b. Functions missing in dictionary
 - c. Terms of usage
 - d. Who has access to the dictionary content
 - e. Which departments are responsible for defining and maintaining content
 - f. How many employees are working in content development and maintenance
 - g. What kind of skills are necessary to do content work
 - h. Main benefit from operating own dictionary
 - i. Which data model is in use

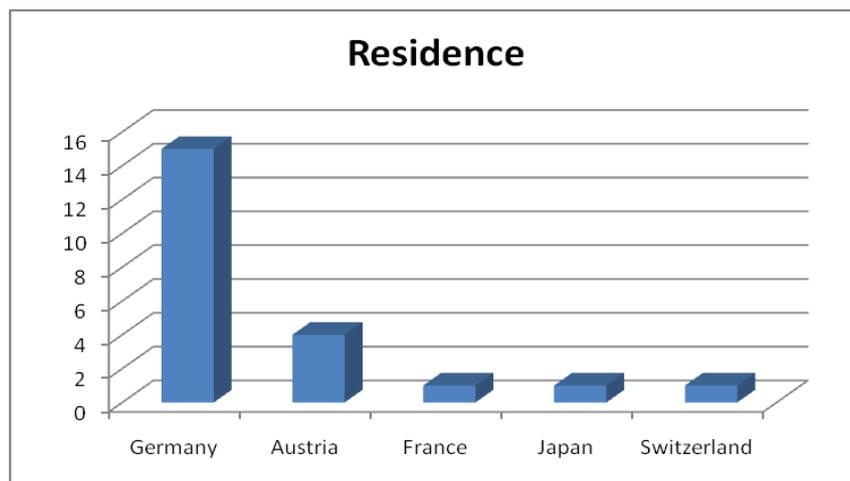
2. Participants

In this section the information on the participant is summarized. This document contains the analysis as of End of August 2009.

The first question was to request the country of residence of the organization the participant works for. 22 out of 36 participants replied.

No. of Answers	Company Residence
15	Germany
4	Austria
1	France
1	Japan
1	Switzerland

Table 65: Country of residence of participants



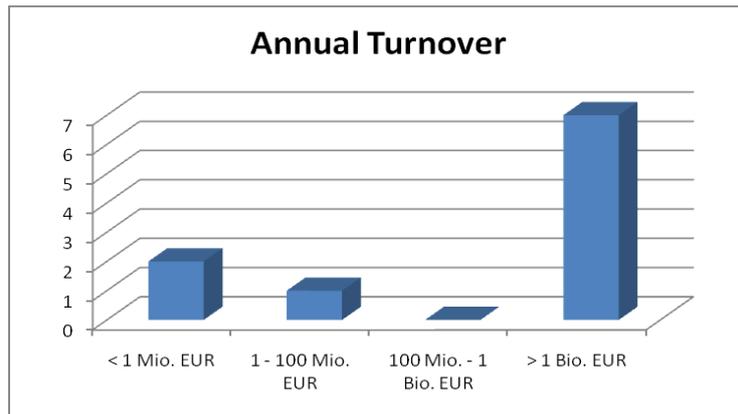
Picture 77: Country of Residence of participants

The majority of the participants who declared the country of residence of their organization came from companies located in Germany. Together with the Austrian participants more than half of the overall participants came from German speaking countries.

The second question about the organization was about the annual turnover. Only 10 out of 36 participants replied on this question.

No. of Answers	Annual Turnover
2	< 1 Mio. EUR
1	1 - 100 Mio. EUR
0	100 Mio. - 1 Bio. EUR
7	> 1 Bio. EUR

Table 66: Annual turnover of participants



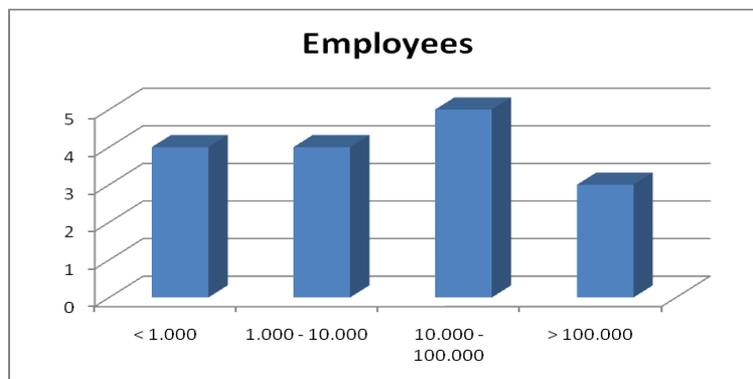
Picture 78: Annual turnover of participants

The result shows that the big majority of the participants answering to this question are very large organization with an annual turnover of more than 1 billion EURO.

The third question about the organization was about the number of employees. 16 out of 36 participants replied on this question.

No. of Answers	Number of employees
4	< 1.000
4	1.000 - 10.000
5	10.000 - 100.000
3	> 100.000

Table 67: Number of employees of participants

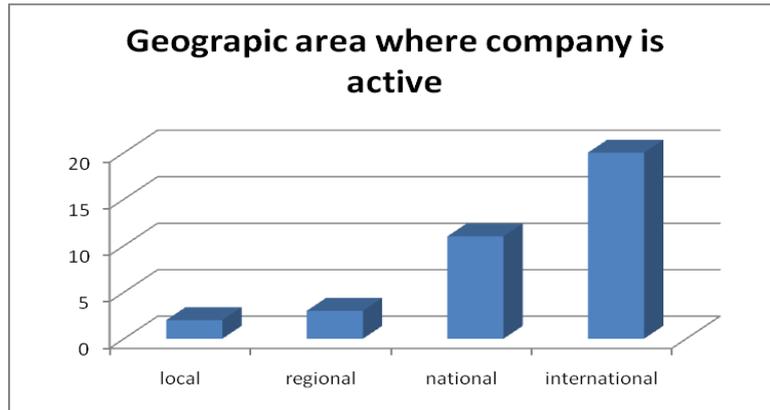


Picture 79: Number of employees of participants

The next question was in which geographical area the company is active. Multiple answers were possible. Out of the 36 participants 26 replied on this question.

No. of Answers	Geographic area where company is active
2	local
3	regional
11	national
20	international

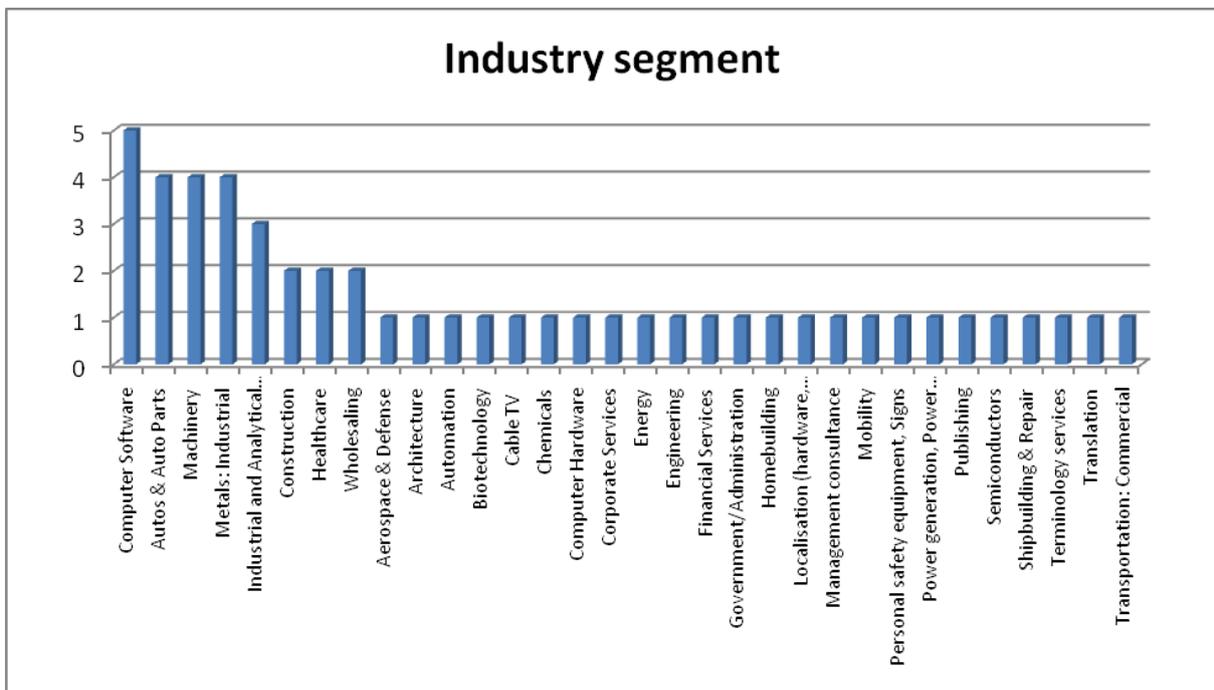
Table 68: Geographical area the participants are active



Picture 80: Geographical area the participants are active

The results of the survey show that the big majority of participants are internationally active.

The next question was related to the organization the participant is working for. The question was: “To which part of the industry does your company belong to?” Multiple answers were possible.



Picture 81: Industry segment where the participants are active

The results of the survey show that dictionaries are used in many segments of the industry. There are some segments like Software, Automotive, Machinery, Metals and Industrial and Analytical Instruments with more replies because these segments are either directly related to the use of electronic data (Software) or there are activities in this sectors to actively build or support electronic dictionaries.

CEN/ISSS eCAT WS Questionnaire: How is industry using electronic Product Property Server (ePPS)?

Thank you for your interest in answering this questionnaire.

Answered questionnaires will be analyzed and evaluated; the resulting survey will become part of the ePPS CEN Workshop Agreement (CWA – a kind of pre-standard like PAS in ISO).

With saving the answers in this questionnaire you accept the use of the answers in a survey. In the publication of this survey anonymity of the answers is guaranteed.

Every seriously answered questionnaire will take part in a drawing for an amazon.com coupon worth 50 Euros.

Thanks to USPI for providing the technical platform for this questionnaire.

Background:

CEN/ISSS is the name given to CEN's ICT sector activities. It provides market players with a comprehensive and integrated range of standardization services and products, in order to contribute to the success of the Information Society in Europe.

The ePPS (electronic Product Property Server) project was started within the CEN/ISSS Workshop eCAT (Multilingual eCataloguing and eClassification in eBusiness) in February 2008.

Catalogues and classification systems are key issues for product data management industry. Their importance is even higher in the electronic marketplace, where companies cooperate with each other and with public administrations based in different countries using electronic tools. There is a link between catalogues and classification systems as products need to be grouped (by a classification) in order to set up a catalogue. Moreover in Europe catalogues and classification schemes are required to be multilingual. The ePPS project focuses on systems handling metadata which are being used to handle and implement master data (such as product properties, classification systems, catalogue metadata).

These systems can be:

- available for cross industry use (e.g. ISO dictionaries, IEC dictionaries),
- used within a specific industry or business domain (e.g. optical industry) or
- accessible only within one company or organization (e.g. internal master data server or property dictionary)

Besides property servers we are also interested in metadata systems handling other concepts like terminology, graphical symbols, units of measurement, codes of all types, databases,...)

The ePPS project will produce one CWA (CEN Workshop Agreement), which will contain 'Guidelines for the design, implementation and operation of a product property server'. The CWA will be published in 2009. ePPS - Questionnaire:

This questionnaire was developed in order to obtain an overview about the state-of-the-art in the use of Product Property Servers/Metadata Dictionaries in industry. We are highly interested in your experience in using external or internal Product Property Servers/Metadata Dictionaries, which will help us to improve pertinent systems.

General questions

General (optional)	
Organization	<input type="text"/>
Organization residence (country)	<input type="text"/>
Questionnaire answered by (name)	<input type="text"/>
Annual turnover in the year 2007	<input type="text"/>
Number of employees	<input type="text"/>
Telephone	<input type="text"/>
E-Mail (necessary if you like to receive the results of the survey or participate in the lottery)	<input type="text"/>

Do you want to receive the results of the survey?	
<input type="radio"/>	Yes
<input type="radio"/>	No
<input checked="" type="radio"/>	No answer

In which geographic area is your company active? (Please specify)

Check any that apply

- | | |
|--|----------------------|
| <input type="checkbox"/> local | <input type="text"/> |
| <input type="checkbox"/> regional | <input type="text"/> |
| <input type="checkbox"/> national | <input type="text"/> |
| <input type="checkbox"/> international | <input type="text"/> |

Which part of industry does your company belong to?
Check any that apply

- | | |
|--|----------------------|
| <input type="checkbox"/> Advertising | <input type="text"/> |
| <input type="checkbox"/> Aerospace & Defense | <input type="text"/> |
| <input type="checkbox"/> Agribusiness | <input type="text"/> |
| <input type="checkbox"/> Airlines | <input type="text"/> |
| <input type="checkbox"/> Alcoholic Beverages & Tobacco | <input type="text"/> |
| <input type="checkbox"/> Apparel & Footwear | <input type="text"/> |
| <input type="checkbox"/> Autos & Auto Parts | <input type="text"/> |
| <input type="checkbox"/> Banking | <input type="text"/> |
| <input type="checkbox"/> Biotechnology | <input type="text"/> |
| <input type="checkbox"/> Broadcasting & Cable | <input type="text"/> |
| <input type="checkbox"/> Cable TV | <input type="text"/> |

<input type="checkbox"/> Chemicals	
<input type="checkbox"/> Coal Mining	
<input type="checkbox"/> Computer Hardware	
<input type="checkbox"/> Computer Software	
<input type="checkbox"/> Educational Services	
<input type="checkbox"/> Environmental & Waste Management	
<input type="checkbox"/> Financial Services	
<input type="checkbox"/> Foods & Nonalcoholic Beverages	
<input type="checkbox"/> Government/Administration	
<input type="checkbox"/> Healthcare	
<input type="checkbox"/> Homebuilding	
<input type="checkbox"/> Household Durables	
<input type="checkbox"/> Household Nondurables	
<input type="checkbox"/> Industrial and Analytical Instruments	
<input type="checkbox"/> Insurance	
<input type="checkbox"/> Investment Services	
<input type="checkbox"/> Lodging & Gaming	
<input type="checkbox"/> Machinery	

- | | |
|--|--|
| <input type="checkbox"/> Metals: Industrial | |
| <input type="checkbox"/> Movies & Home Entertainment | |
| <input type="checkbox"/> Natural Gas Distribution | |
| <input type="checkbox"/> Packaging & Containers | |
| <input type="checkbox"/> Paper & Forest Products | |
| <input type="checkbox"/> Publishing | |
| <input type="checkbox"/> Recreation | |
| <input type="checkbox"/> Restaurants | |
| <input type="checkbox"/> Retailing | |
| <input type="checkbox"/> Savings & Loans | |
| <input type="checkbox"/> Securities, Mutual Funds, and Commodity Futures Trading | |
| <input type="checkbox"/> Semiconductors | |
| <input type="checkbox"/> Shipbuilding & Repair | |
| <input type="checkbox"/> Supermarkets & Drugstores | |
| <input type="checkbox"/> Telecommunications | |
| <input type="checkbox"/> Textile | |
| <input type="checkbox"/> Transportation: Commercial | |
| <input type="checkbox"/> Travel & Tourism | |
| <input type="checkbox"/> Wholesaling | |

<input type="checkbox"/> Wood Products	<input type="text"/>
<input type="checkbox"/> Others, please specify	<input type="text"/>
<input type="checkbox"/> Others, please specify	<input type="text"/>
<input type="checkbox"/> Others, please specify	<input type="text"/>

Which ePPS are you using?
Check any that apply

- We are using the ePPS of a standardization body in order to access a product dictionary
- We use the ePPS of a business partner to access their product dictionary
- We have our own ePPS in the company to manage product dictionary content
- We use a ePPS in another way
- We plan to use an ePPS in the future for the following purpose(s)

The next parts of the questionnaire are depending on the answers of this question.



European Committee for Standardization
Comité Européen de Normalisation
Europäisches Komitee für Normung

Questions for users of ePPS of SDOs

What types of content do you use via ePPS with your SDO?
Check any that apply

- Terminology
- Product Properties
- Classification
- Graphical symbols
- Quantities and Units
- Others, please specify
- Others, please specify
- Others, please specify

SDO = Standards Developing Organization

Which standardization organizations' ePPS do you use? (Part 1)

	in use	planned	No answer
CPV	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
DIN	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
ECALS	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
eCI@ss	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
ETIM	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
GPC	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
ISO 13584-501 RA (Jemima)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
ISO 13584-511	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
ISO 13399	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
ISO 15926	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
ISO 23584	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
ISO/CDB	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
NCS	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
OTD	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
PFI	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
PI	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
proficl@ss	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Prolist	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

RosettaNet	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
UNSPSC	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
VDZ	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

In the next question you can add free text if there are options missing.

Which standardization organizations' ePPS do you use? (Part 2)

in use

planned

If you use more than one, please mention which kind of interoperability problems do you noticed?

Which functions of the ePPS of your SDO partners do you use? (Part 1)
Check any that apply

Look up of single content elements

Download of content in computer process able formats

Upload of proposals for new content

<input type="checkbox"/>	Download of support material for version upgrade (e.g. delta files)	<input type="text"/>
<input type="checkbox"/>	Information about the maintenance process and the release planning	<input type="text"/>
<input type="checkbox"/>	Contacts to relevant committees in the organization	<input type="text"/>
<input type="checkbox"/>	Forum for the exchange with other users	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>

Which functions are missing in the ePPS of your SDO?



In which formats do you exchange the content with your SDO?
Check any that apply

- We use special data formats to exchange the dictionary content.
- We exchange dictionary content, such as properties, within other business documents, such as request for quotations [RFQ], quotations, auctions, catalogues.

In which special data formats do you exchange the dictionary content with your SDO?
Check any that apply

- Set of tables
- BMEcat
- OntoML
- Others, please specify
- Others, please specify
- Others, please specify
- Other:

In which business documents, such as request for quotations [RFQ], quotations, auctions, catalogues do you exchange the dictionary content with your SDO?
Check any that apply

- RFQ - Data Format
- Quotation - Data Format
- Auction - Data Format
- Catalog - Data Format
- Library data - Data Format
- Others, please specify Document Type and Data Format separated by comma
- Others, please specify Document Type and Data Format separated by comma
- Others, please specify Document Type and Data Format separated by comma
- Others, please specify Document Type and Data Format separated by comma

How often do you start update processes with your SDO?
Check any that apply

- When needed (missing or changed content is exchanged ad hoc)
- Content is renewed when minor updates are realized
- Content is only renewed when major updates are realized
- Sometimes major updates are skipped
- All (needed) content is included in every exchange
- Others, please specify
- Others, please specify
- Others, please specify

How are the terms of usage when you use the content of a dictionaries of your SDO?
Check any that apply

- There is no regulation
- There are bilateral agreements
- The content is copyright free
- The content is free for public use
- Others, please specify

<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>

What would increase the usability of your SDOs ePPS?
Check any that apply

- Clearly defined terms of usage
- Usage is free of charge
- Additional exchange formats
- Easier access to the change processes
- Better support regarding the content
- Better technical support
- Better training offerings
- Better marketing
- Access to enhanced SW tools
- Others, please specify
- Others, please specify
- Others, please specify

In which departments of your company do you use the content of the SDO?
Check any that apply

--

<input type="checkbox"/>	Research and development	<input type="text"/>
<input type="checkbox"/>	Production	<input type="text"/>
<input type="checkbox"/>	Procurement	<input type="text"/>
<input type="checkbox"/>	Sales	<input type="text"/>
<input type="checkbox"/>	Maintenance	<input type="text"/>
<input type="checkbox"/>	Marketing	<input type="text"/>
<input type="checkbox"/>	Logistics	<input type="text"/>
<input type="checkbox"/>	Finance	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>



European Committee for Standardization
Comité Européen de Normalisation
Europäisches Komitee für Normung

Questions for users of ePPS of business partners

What types of content do you use via ePPS with your business partners?
Check any that apply

- Terminology
- Product Properties
- Classification
- Graphical symbols
- Quantities and Units
- Others, please specify
- Others, please specify
- Others, please specify

Which functions of the ePPS of your business partners do you use?
Check any that apply

- Look up of single content elements
- Download of content in computer process
able formats

<input type="checkbox"/>	Upload of proposals for new content	<input type="text"/>
<input type="checkbox"/>	Download of support material for version upgrade (e.g. delta files)	<input type="text"/>
<input type="checkbox"/>	Information about the maintenance process and the release planning	<input type="text"/>
<input type="checkbox"/>	Contacts to relevant committees in the organization	<input type="text"/>
<input type="checkbox"/>	Forum for the exchange with other users	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>

Which functions are missing in the ePPS of your business partners?



In which formats do you exchange the content with your business partners?
Check any that apply

- We use special data formats to exchange the dictionary content.
- We exchange dictionary content, such as properties, within other business documents, such as request for quotations [RFQ], quotations, auctions, catalogues.

In which special data formats do you exchange the dictionary content with your business partners?

Check any that apply

- Set of tables
- BMEcat
- OntoML
- Others, please specify
- Others, please specify
- Others, please specify
- Other:

In which business documents, such as request for quotations [RFQ], quotations, auctions, catalogues do you exchange the dictionary content with your business partners?

Check any that apply

- RFQ - Data Format
- Quotation - Data Format
- Auction - Data Format
- Catalog - Data Format
- Library data - Data Format

<input type="checkbox"/>	Others, please specify Document Type and Data Format separated by comma	<input type="text"/>
<input type="checkbox"/>	Others, please specify Document Type and Data Format separated by comma	<input type="text"/>
<input type="checkbox"/>	Others, please specify Document Type and Data Format separated by comma	<input type="text"/>
<input type="checkbox"/>	Others, please specify Document Type and Data Format separated by comma	<input type="text"/>

How often do you start update processes with your business partners?
Check any that apply

<input type="checkbox"/>	When needed (missing or changed content is exchanged ad hoc)	<input type="text"/>
<input type="checkbox"/>	Content is renewed when minor updates are realized	<input type="text"/>
<input type="checkbox"/>	Content is only renewed when major updates are realized	<input type="text"/>
<input type="checkbox"/>	Sometimes major updates are skipped	<input type="text"/>
<input type="checkbox"/>	All (needed) content is included in every exchange	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>

How are the terms of usage when you use the content of a dictionaries of your business partners?

Check any that apply

- There is no regulation
- There are bilateral agreements
- The content is copyright free
- The content is free for public use
- Others, please specify
- Others, please specify
- Others, please specify

What would increase the usability of your business partners ePPS?

Check any that apply

- Clearly defined terms of usage
- Usage is free of charge
- Additional exchange formats
- Easier access to the change processes
- Better support regarding the content

<input type="checkbox"/>	Better technical support	<input type="text"/>
<input type="checkbox"/>	Better training offerings	<input type="text"/>
<input type="checkbox"/>	Better marketing	<input type="text"/>
<input type="checkbox"/>	Access to enhanced SW tools	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>

In which departments of your company do you use the content of the business partners?
Check any that apply

<input type="checkbox"/>	Research and development	<input type="text"/>
<input type="checkbox"/>	Production	<input type="text"/>
<input type="checkbox"/>	Procurement	<input type="text"/>
<input type="checkbox"/>	Sales	<input type="text"/>
<input type="checkbox"/>	Maintenance	<input type="text"/>
<input type="checkbox"/>	Marketing	<input type="text"/>
<input type="checkbox"/>	Logistics	<input type="text"/>
<input type="checkbox"/>	Finance	<input type="text"/>

<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>

Do you have a formally specified data model which describes how your content is organized?
Choose one of the following answers – YES

Is your data model based on other models? Check any that apply		
<input type="checkbox"/>	PLIB (ISO 13584 / IEC 61360)	<input type="text"/>
<input type="checkbox"/>	OTD (ISO 22745)	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>
<input type="checkbox"/>	Others, please specify	<input type="text"/>

Which mechanism do you use for specifying your data model (e.g. XML DTD, XML Schema, UML, tables, others)? (Please provide a description or a link to a description, if possible)

Do you have a formally specified data model which describes how your content is organized?
Choose one of the following answers – NO

How is the structure of your content specified?