

CEN/ISSS WS/eCat

Multilingual catalogue strategies for eCommerce and eBusiness

Combined report WP 1

Version 2003-07-11

(with tracking of changes relative to the version 2003-06-25)

This combined report has been written by WS/eCat Project Team members responsible for tasks 1.1, 1.2, and 1.4. The report has been edited by Håvard Hjulstad.

Note: Clause numbers have been changed in editing. References to clause numbers may be incorrect in this version.

Members of the Project Team are:

Helmut Beckmann (1.1)
Fraunhofer IAO and Universität Stuttgart
helmut.beckmann@iao.fhg.de

Pim van der Eijk (1.2)
Sonnenglanz
pvde@sonnenglanz.net

Volker Schmitz (1.3)
Universität Essen
volker.schmitz@uni-essen.de

Nikolaus Ondracek (1.4 and 3)
Paradine
nikolaus.ondracek@paradine.at

Håvard Hjulstad (1.5 and coordinator)
havard@hjulstad.com

Bodil Nistrup Madsen (2)
Handelshøjskolen i København
bnm.id@cbs.dk

This title page will be replaced by the CWA Formal Title Page, created by the CEN Standards Production Department at the moment of producing the CWA for publication.

Contents

Contents	2
Foreword	6
Introduction	7
1 Scope	8
2 Normative References	9
3 Definitions, symbols and abbreviations	10
3.1 Definitions	10
3.2 Symbols	10
3.3 Abbreviations	10
4 Survey of existing e-catalogues for e-business [WP 1.1]	11
4.1 Method for creation of the survey of existing e-catalogues	11
4.1.1 Organizational Information	11
4.1.2 Usage of electronic product catalogues	11
4.1.3 e-catalogue formats	12
4.1.4 e-catalogue Content	12
4.1.5 System application	12
4.2 Survey of e-catalogues	13
4.2.1 Industry sectors	13
4.2.2 Turnover Distribution	13
4.2.3 Supply Chain Hierarchy	13
4.2.4 EDI	13
4.2.5 Requests for EDI	13
4.2.6 Departments involved in e-catalogue exchange	13
4.2.7 e-catalogue medium distribution	13
4.2.8 User estimation for standardisation need	13
4.2.9 Catalogue usage	13
4.2.10 Catalogue functionalities	13
4.2.11 Catalogue languages	13
4.2.12 System environment	14
4.2.13 System language support	14
5 Survey of organizations supporting e-catalogue development in e-business [WP 1.2]	15
5.1 Specifications and standards	15
5.1.1 International de jure standards organizations	15

5.1.1.1	ISO	15
5.1.1.2	UN/CEFACT	16
5.1.2	Other (standards) bodies and open (industry) organizations	16
5.1.2.1	BME	16
5.1.2.2	CEN ISSS	16
5.1.2.3	CIDX	17
5.1.2.4	CommerceNet	17
5.1.2.5	EAN International	17
5.1.2.6	ECCMA	18
5.1.2.7	eCI@ss	18
5.1.2.8	Edira	18
5.1.2.9	LISA	18
5.1.2.10	NATO	18
5.1.2.11	OASIS	19
5.1.2.12	ONCE	19
5.1.2.13	Open Application Group	19
5.1.2.14	RosettaNet	20
5.1.2.15	UCC	20
5.1.2.16	UNDP	20
5.1.3	Other industry organizations	21
5.1.3.1	Ariba	21
5.1.3.2	Commerce One	21
5.1.3.3	Martsoft	21
5.1.3.4	SAP	21
5.1.3.5	Requisite Technology, Inc.	21
5.2	Providers of e-Catalogue Software	21
5.2.1	Typology of e-Catalogue software	22
5.2.2	Geographic Origin	23
5.2.3	Standards supported	24
5.2.4	Multilinguality	24
6	State-of-the-art in e-Catalogue research [WP 1.2]	25
6.1	Literature Survey	25
6.1.1	Characterization of e-Catalogues	26
6.1.2	Content integration	26
6.1.3	Coverage of real-world business practice	28
6.1.4	Management issues	28

6.1.5	Database technology	29
6.1.6	Standards adoption	29
6.2	Organizations active in e-Catalogue Research	30
7	Analysis of existing e-catalogues [WP 1.3]	31
8	Analysis of the relation between (as well as application of) existing product classification schemes/product identification schemes and e-catalogues [WP 1.4]	32
8.1	Introduction	32
	Product Data Management today	32
	Product Identification Codes	33
	Product Classification	33
8.2	Criteria for Classification	34
8.2.1	multilingual aspects of existing product classification/identification schemes	34
8.2.2	international character and authoritative nature	34
8.2.3	speed of updating and change management	34
8.2.4	diffusion of support software	34
8.2.5	upward and downward compatibility of versions	34
8.2.6	application in e-catalogues and ontologies	34
8.2.7	relation to terminology and ontologies	34
8.3	Overview of existing product classification/identification schemes	34
8.3.1	CPV - Common Procurement Vocabulary	34
8.3.2	eCI@ss	35
8.3.3	ETIM	37
8.3.4	GCI / EAN	37
8.3.5	NATO Codification System (NCS)	37
8.3.6	Proficlass	38
8.3.7	RNTD -RosettaNet Technical Dictionary	38
8.3.8	UNSPSC Universal Standard Products and Services Classification	39
8.4	Comparison Table	40
8.5	Identification of Problems	40
8.6	Suggestion for Solution to existing (and foreseen future) Problems	40
9	Formulation of problems and recommendations for a systematic approach to implement highly interoperable e-catalogues at SME level [WP 1.5]	41
9.1	Summary of issues raised in analysis sections	41

9.2	Interoperable e-catalogues	41
9.3	Recommendations	41
10	Summary	42
11	Annex <yy> (normative):Title of normative annex	43
11.1	yy.1 First clause of this normative annex	43
11.2	yy.1.1 First subclause of this normative annex	43
12	Bibliography	44

Foreword

<XX>

Introduction

Multilingualism in the Union is often seen as an obstacle for the European economy in terms of competition and the opening up of new markets, but it also has political dimensions relating to consumer protection, freedom to move, etc. However, new economies have emerged in the wake of trying to overcome the language barriers, such as the language industries (incl. activities and the language technologies for making language resources and terminologies available at a large scale), where Europe has a leading edge thanks to the R&D programmes of the EU Commission.

It has been recognised that products and services must be sold in the language of the target market. There are strong indications that e-commerce and e-business can only function well, if the virtual marketplaces and all their major elements (such as product classification schemes, user interfaces, product catalogues etc.) are multilingual from the outset. This, however, would create insurmountable financial barriers for SMEs, if they cannot benefit from synergies through a systematic approach to multilingual data, the methods for their management and the respective tools, as well as the integration of these data, methods and tools into the company's whole ICT environment.

The Management Group (MoU/MG) of the Memorandum of Understanding of the ITU, UN/ECE, ISO ~~and IEC, CEN/ISSS, and (joined by~~ several other organizations, ~~for the sake of~~ coordinating all standardisation and harmonisation efforts with respect to e-commerce and e-business) ~~is~~ is convinced that e-commerce/e-business can only perform satisfactorily if it is multilingual and properly harmonised. That is the reason why the International Information Centre for Terminology (Infoterm), taking care of the Secretariat of ISO/TC 37 "Terminology and other language resources", has been invited to join the MoU/MG. ISO/TC 37 has recently decided on a NWI (New Working Item) "Basic principles for multilingual product classification schemes for electronic commerce".

The real challenges are establishing interoperable international multilingual product classification schemes for e-commerce and multilingual e-catalogues (fully integrated in terms of process chains into the whole information/knowledge management and data processing environment of enterprises). E-catalogues must be compatible or - what is more important - Interoperable with engineering systems (such as CAD/CAM, FIM, etc.) as well as with business software (such as ERP, etc.) and all kinds of information/knowledge management systems (incl. the language technology software supporting them).

1 Scope

The following document gives an overview and guidance in the usage of existing electronic product catalogues for eBusiness. On one hand an overview of relevant e-catalogue formats will be given on a quantitative level. On the other hand a qualitative comparison of existing catalogue formats has been done to show differences of the content in detail. In addition, this document will show future directions in e-catalogues and points out areas of standardization for e-catalogues.

The document shall be used as guideline for future selection, implementation and usage of electronic product catalogues.

This document is addressed to decision makers within organizations and companies as well as technically interested persons who want to get some knowledge about the technical differences between existing catalogue formats.

2 Normative References

This optional element shall give a list of any normative documents to which reference is made in the CWA in such a way as to make them indispensable for the application of the CWA. For dated references, each shall be given with its year of publication, or, in the case of enquiry or final drafts, with a dash together with a footnote "To be published", and full title. The year of publication or dash shall not be given for undated references. When an undated reference is to all parts of a standard, the publication number shall be followed by the indication "(all parts)" and the general title of the series of parts (i.e. the introductory and main elements).

It is suggested to start as follows: "The following normative documents contain provisions which, through reference in this text, constitute provisions of this CWA. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. However, parties to agreements based on this CWA are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies."

In case of a CWA that is containing informational material only, there are no Normative references. The references are in that case to be listed as "Informative references" in an Annex (see below). In addition, a CWA can contain both Normative AND Informative References.

In addition to these references, which are actually referenced somewhere, there may also be added a Bibliography (see below)

Examples are given below:

ISO 1234-1, Example Standard Reference - Part 1: First example

ISO 12345-1999, Second Example Standard Reference

ISO 6789, Third Example Standard - A non-published one, - 1

Note that normative references may be made not only to publications of formal standards bodies. For example, reference may be made to consortia specifications, provided these are publically available.

3 Definitions, symbols and abbreviations

[NOTE: title will have to be amended based upon the subsections that are present]

3.1 Definitions

Where applicable these are numbered 3.1. Defined terms should be ordered alphabetically.

For the purposes of the present document, the following terms and definitions apply:

example: this text is an example

3.2 Symbols

Where applicable these are numbered 3.1 (where there are no definitions) or 3.2.

For the purposes of the present document, the following symbols apply:

<Symbol1> <Explanation>

<Symbol2> <Explanation>

3.3 Abbreviations

eCat – electronic product catalogue

EDI – Electronic Data Interchange

XML – extensible Markup Language

4 Survey of existing e-catalogues for e-business [WP 1.1]

The investigation of electronic product catalogues, called e-catalogues, will be done within the first of three work packages of the workshop. The investigation of e-catalogues will be done on different levels. These levels will be partially reflected by the business plan of the workshop. The following section describes the results of work package WP 1.1. This description will be started with the subsection about the method for creating the survey on existing e-catalogue formats. After this, the results will be presented.

4.1 Method for creation of the survey of existing e-catalogues

The basic procedure for the creation of the survey consists of three different levels of investigation:

- Document research,
- Online questionnaire and
- Interviews.

The most important part of the method orchestration is the online questionnaire to get a founded analysis of the usage of catalogue formats for eBusiness and associated topics in the area. This questionnaire will be presented and explained in detail in the next subsections. Each of the following subsections is named accordingly to the sections within the questionnaire.

4.1.1 Organizational Information

In this section for organizational information some statistics information will be collected from the user to classify the information given later.

- Industry sector

This criteria asks for the industry sector the company of the user addresses.

- Turnover

This is the annual turnover of the user company which points out later if also smaller companies are using electronic product catalogues or even doing EDI.

- Supply Chain Hierarchy

The stage of the user within the supply chain gives information about existing standardisation effort since it will point out, where the main pain points concerning electronic product catalogues are.

4.1.2 Usage of electronic product catalogues

- Electronic Data Interchange (EDI)

The general usage of EDI gives hints about the fulfilment of the basic requirements for exchanging electronic data.

- Requests for EDI

According to the SCM Hierarchy this criteria also points out pain points for the usage of electronic product catalogues since it shows which party within the supply chain requests the exchange of electronic product data.

- Involved departments

Different departments have different views on electronic product catalogues. To take this into account for later standardisation efforts, the correlation between e-catalogues and departments will be questioned.

- Product catalogue medium

This criteria serves to get information about the ranking of different mediums in exchanging product information for eCommerce between trading partners.

- Influence of eCat Standards

To see the self estimation of the usefulness of standards in the area of e-catalogues, users give there personal score to this.

4.1.3 e-catalogue formats

- Number of used catalogue formats

One important benchmark for the usage of electronic catalogues is the number of catalogues which have to be exchanged between trading partners. This shows the urgency for the introduction of electronic catalogues.

- Used catalogue (industry-) standard formats

If catalogues are used within a company the formats used give important hints to functionalities needed by users and the strength of different products in the market.

- Usage of proprietary catalogue formats

This criteria shows if there is at least a proprietary catalogue format widely used by different users which should be taken into account (at least from a functional point of view) when talking about standardize on of the existing formats.

4.1.4 e-catalogue Content

- Content areas supported by catalogue formats

This criteria shows the functional areas needed by different user. It serves as the basic functional subset to define or evaluate catalogue formats.

- Support of languages by catalogue formats

Some questions extract the need of the users for multilingual product catalogues and there potential usage.

4.1.5 System application

- Supporting systems

Integration of e-catalogues in the company infrastructure is the most important criteria for the successful application of electronic catalogues. This questions shows the involved environment for catalogues.

- Language support of different systems

The support of different languages is not only a topic for electronic catalogues but also for associated systems. The facility of these systems to support multilingual product data gives an important hint about the integration capabilities of multilingual product catalogues.

4.2 Survey of e-catalogues

In this subsection the results of the investigation will be presented. The section is structured according to the questions that are included in the online questionnaire. These questions are the basic information for the survey creation.

After getting results from the questionnaire evaluation these results will be introduced in the appropriate subsection.

4.2.1 Industry sectors

<xx>

4.2.2 Turnover Distribution

<xx>

4.2.3 Supply Chain Hierarchy

<xx>

4.2.4 EDI

<xx>

4.2.5 Requests for EDI

<xx>

4.2.6 Departments involved in e-catalogue exchange

<xx>

4.2.7 e-catalogue medium distribution

<xx>

4.2.8 User estimation for standardisation need

<xx>

4.2.9 Catalogue usage

<xx>

4.2.10 Catalogue functionalities

<xx>

4.2.11 Catalogue languages

<xx>

4.2.12 System environment

<xx>

4.2.13 System language support

<xx>

5 Survey of organizations supporting e-catalogue development in e-business [WP 1.2]

In this section, we will provide an overview of organizations involved in the development of standards for e-catalogues in e-business. This section will cover the following topics:

- Organizations involved in the creation and/or maintenance of specifications and standards. This chapter only provides a listing of the various organizations and will list some of their products. Other chapters of this CWA will look in detail into the actual content of these specifications and standards.
- Organizations developing software products supporting the creation, management and deployment of e-Catalogues.

5.1 Specifications and standards

In this subsection we provide a listing of various organizations involved in creating specifications and standards in a number of areas of relevance for e-Catalogues:

- E-Catalogue vocabularies;
- Classification systems;
- Identification (numbering) systems;
- Product information and description systems;
- Terminology and multilinguality.

Rather than presenting the organizations by area, this section will group the organizations in three groups, as some organizations are involved in several of these areas:

- International de jure standards, organizations;
- Other (standards) bodies, open industry groups, user groups or trade associations;
- Commercial companies or closed consortia promoting proprietary specifications.

Note that we will use the term *vocabulary* in a generic sense to cover XML document formats, EDI message definitions, and other formats.

5.1.1 International de jure standards organizations

5.1.1.1 ISO

The International Standards Organization (URL: <http://www.iso.org/>) is a network of national standards institutes from 140 countries working in partnership with international organizations, governments, industry, business and consumer representatives. It has active Technical Committees in a number of areas related to aspects of e-catalogues, including language standards and product lifecycle management.

ISO Technical Committees relevant to language standards are TC 37 (Terminology and other language resources), ISO 639 (Language Codes), Terminology Data Categories (ISO 12620), MARTIF (ISO 12200) Machine-readable terminology interchange format,

Terminology Work (ISO 704), Terminological Markup Framework (ISO DIS 16642) and ISO639-1, a new ISO standard for the identification of language names.

ISO Technical Committees relevant to e-catalogues from the perspective of product data representation and exchange include ISO 10303, Step and ISO 13584, Parts library, concerned with the representation and exchange of libraries of manufactured parts data. ISO 15459 standardizes a unique identification system for transport items, representable in bar codes, and a procedure for the allocation of unique identification numbers for transport units through approved issuing agencies.

5.1.1.2 UN/CEFACT

UN/CEFACT (URL: <http://www.uncefact.org/>) is the United Nations Centre for Trade Facilitation and Electronic Business, located in the Economic Commission for Europe (UN/ECE). The mission is to “improve the ability of business, trade and administrative organizations, from developed, developing and transitional economies, to exchange products and relevant services effectively - and so contribute to the growth of global commerce”. Its focus is on worldwide facilitation of international transactions, through the simplification and harmonization of procedures and information flows. E-business is an important instrument in this context.

The best known deliverable of UN/CEFACT is the UN/EDIFACT framework for Electronic Data Interchange. UN/EDIFACT includes large libraries of data elements and business messages, including the PRICAT price/sales catalogue message.

With OASIS, UN/CEFACT has created the ebXML framework for electronic business (URL: <http://www.ebxml.org/>), and continues to be responsible for development of the ebXML business information and business process specifications. TBG1, “Purchasing”, has received eight Work Items including one item on Catalogues.

5.1.2 Other (standards) bodies and open (industry) organizations

5.1.2.1 BME

One of the more well-known e-catalogue vocabularies developed in Europe, BMEcat, is named after the Bundesverband Materialwirtschaft, Einkauf und Logistik e.V, the organization of procurement professionals in Germany (URL: <http://www.bme.de/>). The specification is in fact developed by larger consortium, which apart from BME includes the Fraunhofer IAO, Stuttgart Competence Center Electronic Business and the Universität Essen bli - Beschaffung, Logistik und Informationsmanagement (also see section Organizations active in e-Catalogue Research) as technology partners, as well as a number of enterprises.

BME offers a BME catalog certification service via a specialized subsidiary, BMEnet GmbH (URL: <http://www.bmenet.de/>).

5.1.2.2 CEN ISSS

CEN ISSS (URL: <http://www.cenorm.be/iss/>) was created in mid-1997 by CEN (European Committee for Standardization) as the focus for its ICT activities. The mission of CEN/ISSS is to “provide market players with a comprehensive and integrated range of standardization-oriented services and products, in order to contribute to the success of the Information Society in Europe”. CEN ISSS is hosting the e-Catalogue Workshop in the context of which this document is being prepared.

Related activities within CEN ISSS include a concluded workshop, e-NOM “European Electronic Commerce for Product Nomenclature”. This workshop was launched in April 2002, Brussels, and concluded its work on 31 December 2003. The final meeting was held on 28 November 2002 in Paris and approved the final text of the CWA. The CWA will be sent to the GCI (Global Commerce Initiative) as European input on classification of non-food products.

5.1.2.3 CIDX

Chemical Industry Data Exchange (URL: <http://www.cidx.org/>) is a trade association and neutral standards body engaged in the development of supply chain standards for chemical companies and their trading partners. Products of the organization include standards, guidelines, support materials and communications and are developed by CIDX member volunteers. The mission is “to improve the ease, speed and cost of conducting business electronically between chemical companies and their trading partners”.

The version 3.0 specifications of the Chem eStandards™ specifications includes XML messages for product catalogue updates and customer specific catalogue updates.

5.1.2.4 CommerceNet

Since 1994, CommerceNet (URL: <http://www.commerce.net/>) has been active as open, interoperable network linking many commerce communities through a common architectural framework, aiming at making it possible for companies to do business electronically through open Internet standards. It is well known for its association with initiatives like eCo (electronic commerce framework) and OBI (Open Buying on the Internet), a technical specification for the high-volume, low-amount purchase of non-strategic, indirect materials by large, distributed requisitioner populations, developed from 1996.

Current activities of CommerceNet include pilots, an extension of eCo that takes advantage of Web Services, research and ventures. CommerceNet collaborates with various organizations, including ONCE.

5.1.2.5 EAN International

EAN International (URL: <http://www.ean-int.org/>) aims at developing a set of standards enabling the efficient management of global, multi-industry supply chains by uniquely identifying products, shipping units, assets, locations and services. It facilitates electronic commerce processes including full tracking and traceability. It was formed as EAN association in 1977 to develop a uniform and standard numbering system for Europe, analogously to the North American Universal Product Code (UPC). It has a network of nearly a hundred Member Organizations in as many countries globally, supporting some 900000 member companies.

The mission of the organization is “to create Open, Global, Multisectoral Standards based on Best Business Practices, and by driving their implementation, play a leading role in Supply & Demand Chain Management improvement worldwide.” The EAN.UCC system is a series of standards designed to improve supply chain management. The system offers unique and unambiguous identification systems, such as the Global Trade Item Number (GTIN), as well as representation systems for these identification systems, such as bar codes.

The Global EAN Party Information Register, or GEPIR (URL: <http://www.gepir.org/>), is a worldwide online facility to allow reverse searches from trade items to their brand owners.

The search results in contact information, including (if provided by brand owner) email and web site address. The SINFOS system (URL: <http://www.sinfos.de/>), developed for the German EAN Member Organization, allows searches for companies on GLN and for trade items using GTIN. EAN Finland is starting a project using this SINFOS system.

5.1.2.6 ECCMA

The ECCMA (Electronic Commerce Code Management Association) is an organization based in the United States (URL: <http://www.eccma.org/>). Its Global Attribute Schema (EGAS) provides a dictionary of common descriptive terms used to describe the properties, characteristics or requirements of products, organizations or locations of supply. Its eCat is based on the data provided by the NATO Codification System and is comprehensive, including some 16 million items with 32 million part numbers.

5.1.2.7 eCI@ss

The eCI@ss consortium is a group of mainly German organizations. Its main product is a hierarchical classification system aimed at industrial buyers and supports engineers at development, planning and maintenance (URL: <http://www.eClass-online.com/>). It is hosted by the Institut der deutschen Wirtschaft Köln (Cologne Institute for Business Research) and made available free of charge.

5.1.2.8 Edira

The EDI/EC Registration Authorities Association (EDIRA) is a not-for-profit membership organization for the promotion of the International Code Designators (ICDs), initiated within the framework of the European Commission's TEDIS Programme. ICDs identify organizations that issue business identifier codes to legal entities that are engaged in e-commerce. The prefix joined with the business identifier is a worldwide unique identification. This allows the recipient of a business identifier to easily identify the issuing organization as well as any appropriate data validation and verification information.

The concept of ICDs offers an abstraction layer or namespace mechanism beyond facilities like the EAN GEPIR. The EDIRA Top Level Service (TLS) is designed to provide a unique lookup of Business Identifiers, combining ICD and business identifier into a Uniform Resource Identifier (URI) that can be looked up on the web. It provides access to multiple Code Designator systems, including Dun & Bradstreet Data Universal Numbering System (D-U-N-S®).

5.1.2.9 LISA

Founded in 1990, the Localization Industry Standards Association (LISA; URL: <http://www.lisa.org/>) is the premier professional organization for the globalization, internationalization, localization, and translation business communities. With more than 400 members from the high-tech and vertical market sectors, LISA is well-known for its best practice guidelines for language technology standards and enterprise globalization.

5.1.2.10 NATO

NATO is an example of a large user organization that has had to develop a classification system to support its logistics processing needs, the NATO Classification System (URL: <http://www.nato.int/structur/AC/135/>).

5.1.2.11 OASIS

OASIS (URL: <http://www.oasis-open.org/>) is a not-for-profit, global consortium that drives the development, convergence and adoption of e-business standards. It was founded in 1993 as SGML Open. Members themselves set the OASIS technical agenda, using a lightweight, open process expressly designed to promote industry consensus and convergence. OASIS produces worldwide *de facto* standards for security, Web services, XML conformance, business transactions, electronic publishing, topic maps and interoperability within and between marketplaces. OASIS has more than 600 members in 100 countries around the world.

OASIS and the United Nations jointly sponsor ebXML, a global framework for e-business data exchange (URL: <http://www.ebxml.org/>). The OASIS Universal Business Language (UBL) Technical Committee is closely linked to the ebXML framework. A subcommittee of the UBL Technical Committee, the Library Content subcommittee, aims to develop standard XML business library content by taking an existing library as a starting point and modifying it to incorporate the best features of other existing business and core component libraries. It aims at creating a BIE Catalogue - Identifying the BIE's out of the xCBL Library and XML (XSD) Schemas for business document types.

At the time of writing, the UBL LC TC is looking into developing business documents and BIEs for e-catalogues.

In the area of multilingual data management, the OASIS XLIFF (eXtensible Localization Interchange File Format) TC aims to define, through XML vocabularies an extensible specification for the interchange of localization information. The specification will provide the ability to mark up and capture localizable data and interoperate with different processes or phases without loss of information.

5.1.2.12 ONCE

The Open Network for Commerce Exchange (URL: <http://www.connect-once.org/>) is a global, not-for-profit organization. It is devoted to the development and growth of business-to-business (B2B) electronic trade. Its members constitute an alliance of electronic trading networks, customers, supplies, services and technology providers in the world working to provide a global platform for interoperability. It was formerly known as the Global Trading Web Association (GTWA). Among other activities, it encourages industry standardization for business interoperability, technical standards, and the public policies impacting the future growth of global B2B commerce.

Currently, ONCE committees are addressing four areas: Interoperability, Business Process or Service Syndication, Security and Market Education.

5.1.2.13 Open Application Group

The Open Applications Group (URL: <http://www.openapplications.org/>) is a non-profit consortium focusing on best practices and process based XML content for eBusiness and Application Integration. Its mission is to “define and encourage the adoption of a unifying standard for eBusiness and Application Software interoperability that reduces customer cost and time to deploy solutions”. The Open Application Group Integration Specification (OAGIS) is the largest collection of XML based content for business software interoperability in the world, and is aimed at both Business to Business (B2B) and Application to Application (A2A). It is widely adopted and continues to be developed and extended.

The OAGIS contains multiple Business Object Documents (BODs) related to e-catalogues. Its coverage of e-catalogue process models is one of the more complete ones.

5.1.2.14 RosettaNet

RosettaNet (<http://www.rosettanet.org/>) is a non-profit consortium of more than 400 of the world's leading Information Technology (IT), Electronic Components (EC), Semiconductor Manufacturing (SM) and Solution Provider (SP) companies working to create, implement and promote open e-business process standards. Its mission is to “[drive] collaborative development and rapid deployment of Internet-based business standards, creating a common language and open e-business processes that provide measurable benefits and are vital to the evolution of the global, high-technology trading network.” It is a subsidiary of the UC Council.

The RosettaNet specifications include industry standard dictionaries that provide unique identification of products. RosettaNet Partner Interface Processes (PIPs) are specialized system-to-system XML-based dialogs that define business processes between trading partners. The PIP segment “preparation for distribution” defines interface processes for buyers to include supplier catalogues into their own e-catalogues.

RosettaNet uses D-U-N-S, GTIN, and UN/SPSC for company, product and class identification, respectively.

5.1.2.15 UCC

The Uniform Code Council (URL: <http://www.uc-council.com/>) is a Member Organization of an expanded EAN International, located in North America. It originates from collaboration between trade associations representing the two branches of the food distribution system, which resulted in the creation of the Universal Product Code (U.P.C) system in 1974. The UPC system is the foundation for bar code scanning, which has had a major input on many industries.

The UC Council has initiated several initiatives related to e-catalogues. UCCnet (URL: <http://www.uccnet.org/>) is an initiative to support electronic trading by allowing trading partners served by the UC Council to have synchronized EAN-UCC item information, as well as access to compliant business applications and services. It serves both peer to peer and e-marketplace based commerce. UCCnet operates a GLOBAL registry for companies to store product data, ensuring that trading partners have access to accurate, up-to-date, industry-compliant trading information. This synchronized and standardized data enables companies to reduce the costs and delays associated with “dirty” supply chain information.

5.1.2.16 UNDP

UNDP is the UN's global development network, "advocating for change and connecting countries to knowledge, experience and resources to help people build a better life." The organization is involved in development projects in 166 countries, developing local capacity, drawing on UNDP staff and its range of partners. The UNSPSC (URL: <http://www.unspsc.org/>) is a global standard for classification of products and services that is designed to facilitate electronic commerce, resulting in reduced organization's costs, and improved supply chain efficiencies. It is available free of charge. As of May 2003, the UC Council has taken on the operation of the UNSPSC on behalf of UNDP.

5.1.3 Other industry organizations

5.1.3.1 Ariba

Ariba is known as the organization that has created and publishes the cXML specifications (URL: <http://www.cXML.org/>). It is well-known for its support for the *PunchOut* protocol concept, enabling access to remote supplier sites from a procurement site or system. The cXML specifications are made available on royalty-free basis, but it is stated that the organization may change the terms and conditions of the license agreement simply by posting an updated version on the Web site.

5.1.3.2 Commerce One

With the acquisition of Veo Systems, Commerce One (URL: <http://www.commerceone.com/>) acquired the intellectual property for XML Common Business Library (xCBL) specification. XCBL is a set of XML building blocks and a document framework that allows the creation of robust, reusable, XML documents to facilitate global trading. It is based on an analysis of existing e-commerce standards including Electronic Data Interchange (EDI) and aims to support global e-commerce including multi-company supply chain automation, direct and indirect procurement, planning, auctions, and invoicing and payment in an international multi-currency environment. While xCBL is subject to license terms set by Commerce One (published on the xCBL site, URL: <http://www.xcbl.org/>), these terms allow perpetual, nonexclusive, royalty-free, worldwide right and license to use, publish, copy, modify and distribute the Specification.

The xCBL specification is widely used and serves as input to the OASIS UBL specification, which does not yet offer equivalent functionality but is intended to supersede it in the future.

5.1.3.3 Martsoft

Martsoft Inc. has designed the OCP and OCF specifications for e-catalogues. The Open Catalogue Protocol (OCP) is a protocol to exchange catalogue information encoded in the Open Catalogue Format (OCF; URL: <http://www.martsoft.com/ocf/>).

5.1.3.4 SAP

The SAP Markets Enterprise Buyer product can access external catalog systems using an interface called OCI (Open Catalogue Interface) that offers a mechanism to connect to outbound catalog systems similar to the cXML *PunchOut* concept. SAP software also supports other e-catalogue formats.

5.1.3.5 Requisite Technology, Inc.

Requisite Technology provides widely used software for e-catalogues that uses the proprietary eCX-XML e-catalogue format, as well as many other e-catalogue formats. Use of the eCX-XML format is not exclusive to Requisite, it can also be used by, or in conjunction with SAP, Oracle or other products.

5.2 Providers of e-Catalogue Software

To support the creation, use, deployment and interchange of electronic catalogues, providers and users of electronic catalogues often resort to providers of software

products and services. The landscape of e-Catalogue software is comprised of a variety of solutions, which often address very different aspects of the use of e-Catalogues or support various types of business or technical functionality. In this section, we will provide an overview of these software systems. The questions we are asking are the following:

- What types of e-Catalogue systems can one distinguish? Where do the various products and solutions fit in this typology?
- What is the geographic origin of the providers of these solutions?
- Which existing e-Catalogue standards and formats are supported? What needs and requirements do vendors see among their customers and prospects and what does this mean for future requirements?
- How do these products handle a key requirement for the European market: multilinguality?

The following four sub-sections are devoted to each of these questions.

5.2.1 Typology of e-Catalogue software

E-Catalogues are machine-processable representations of business information, aimed at supporting electronic commerce. As such, the landscape of software systems support e-Catalogues in some way or another is almost as diverse as the overall landscape for electronic commerce. A report by Giga information group (Rugullies, 2001; updates available commercially) distinguishes five categories of vendors of e-commerce solutions that are relevant in the context of e-Catalogues.

- Commerce software, provided by vendors that have their roots in sell-side business-to-consumer (B2C) e-commerce that are moving into business-to-business (B2B) e-business territory;
- Web content management software, provided by vendors that view catalogue content management as a specific instance of more general content management functionality;
- Channel management software is the corollary to B2C commerce software in the B2B space;
- Catalogue aggregation service providers provide outsourced catalogue management and can make the aggregated catalogues available to market places;
- Specialized B2B catalogue management software;

The report relates these categories to a four-step process needed to produce an e-Catalogue:

- Extract data from the original format and aggregate in a catalog system; this requires file format conversion or integration with a back-end system.
- Transform catalogue data; this involves data cleansing, normalizing and categorization.
- Syndicate product information; here the transformed data is published on the supplier Web site, a marketplace, or transferred to a buy-side procurement system;
- Maintain the catalogue;

The first three categories of product only include e-Catalogue management as a component of a broader solution, but often do not provide full support for each of these

four steps. The syndication process is often missing with commerce software and Web content management systems. The B2B catalogue management software category offers the most complete range of support for catalogue creation, aggregation, distribution and syndication, both to Web-based sell-side systems, intermediaries, as well as buy-side procurement systems operated in-house by large customers

As in the general software industry, there is a trend towards consolidation, where smaller, specialized companies are increasingly partnering with, or being acquired by, larger companies, which are often usually more stable financially.

Apart from these five general categories of software for catalogue management, two related types of software functionality are:

- Data cleansing and normalization solutions;
- Software for product configuration;

The first software is often used at the stage of loading existing data into a catalogue management system. The second category refers to complex, configurable products (e.g. telecom routers, computers, automobiles) or services, where the number of different items may be limited but the ways to configure these items may be large and potentially huge.

In our survey of 83 vendors, the distribution over these categories is as follows:

Category	Count
Commerce Server	10
Web Content Management	7
Channel Management	10
Catalogue Aggregation Service	10
B2B Catalogue Management	23
Data extraction, transformation, cleansing	6
Product configuration software	3
<i>So far unclassified</i>	14

5.2.2 Geographic Origin

Apart from identifying the type of solution provided by a particular vendor, the geographic origin of the vendor has been identified. Note that this is a metric of limited use, as it does not show that a particular vendor may have subsidiaries in many countries, ranging from international sales organizations to product development and services being offered internationally. With this caveat, the following table presents the result for the companies surveyed so far.

Geography	Count
North America	66
Denmark	1
Netherlands	3
France	1

Geography	Count
Germany	7
Ireland	1
UK	2
Israel	1
Japan	1

5.2.3 Standards supported

This section will provide a similar mapping from companies to the various B2B XML e-Catalogue standards. At the moment the number of respondents to a request for information is too small to provide any insightful data here.

5.2.4 Multilinguality

One of the key requirements for e-Catalogues in the European market is support for multilinguality. The requirement for e-Catalogue software to be multilingual is not just that the product should be localized to the local market, providing a user interface, manuals etc. in the local language and providing support for local currencies. Instead, the software should provide support for users to sell or buy their products internationally, provide ways to translate product descriptions to the languages relevant to their international customer base.

This section will provide results from some discussions (interviews, email survey) with vendors on the capabilities of their products in this respect, and their views on and understanding of the requirements and state-of-the-art in this area. Again, so far there has not been enough response to a request for clarification.

6 State-of-the-art in e-Catalogue research [WP 1.2]

As a research topic, e-Catalogues have received the attention of researchers working in a variety of areas. They have approached the subject with different research questions and from a variety of backgrounds. This section contains two subsections, the first of which provides results of a literature survey on e-Catalogue research and a second, shorter, one, which lists some organizations engaged in doing this research.

6.1 Literature Survey

In this section, we will survey a sample of research papers discussing electronic catalogues drawn from the international refereed computer and information science research literature that are accessible via the Web. The overall research literature on e-Catalogues can roughly be grouped in two areas, one approaching the topic from a computer and information science perspective, and the other from a business management point of view. In the second case, e-Catalogues are usually only a secondary topic in discussions on more general procurement or e-business topics. For this reason this survey focuses on the first category.

Our primary source of information for this survey has been the Digital Library of the Association for Computing Machinery (ACM; URL: <http://www.acm.org/dl/>). The ACM Digital Library is a leading on-line subscription-based resource for computer and information science research. It provides access to hundreds of thousands of very high quality articles and has excellent search facilities. The set of references retrieved from this site was expanded with:

- Resources referenced by citations;
- Web sites of participants in the CEN ISSS Workshop e-Cat;
- Additional conference web sites like the annual World Wide Web conference;
- The Robin Cover XML bibliography (URL: <http://xml.coverpages.org/>);
- And pages retrieved using the Google search engine (URL: <http://www.google.com/>).

If closure under recursive citation link traversal is an indication, the overall resulting set seems to cover the more important research papers published in the research literature on e-Catalogues. Some of the more relevant resources are listed in the bibliography of this CWA.

A number of key research issues addressed in these documents are:

- Characterization of e-Catalogues
- Content Integration
- Coverage of real-world business practice
- Management issues
- Database technology
- Standards adoption

We will briefly discuss these items in the following subsections.

6.1.1 Characterization of e-Catalogues

The papers in this category focus on defining e-Catalogues and identifying the main issues and research topics associated with them, both from a technical point of view and from a business point of view.

A good general introduction to the topic of e-Catalogues is presented in the early paper by Sherif Danish, “Building database-driven electronic catalogues” (ACM SIGMOD Record December 1998, <http://doi.acm.org/10.1145/306101.306103>). It starts out by distinguishing the following types of products, which vary greatly in the associated types of catalogue:

- Component manufacturers. These provide many products, the attribute descriptions for which need to be standardized, with a need for central databases to share information.
- System manufacturers. These provide fewer, but more complex, configurable products. There are no standards for configuration rules yet; instead, catalogues refer to web site configuration tools.

Interchange with trading partners requires standardization of terminology and (XML) format. The data model for catalogues needs to be generic as product families differ in attributes. Data models need to support parametric search (search on attributes).

The paper identifies multilingualism as an issue with products that are developed and sold in an international market. The mechanism proposed is to use numeric codes for attributes, with mappings to descriptive text for each language.

External systems integration is needed in catalogue systems, for instance for VAT calculation or product availability checks. To facilitate such integration, component frameworks like COM are recommended as software architecture. Presentation to users (especially in a B2C context) can be customized using server pages technology.

Content management and initial catalogue creation can be handled using file format conversion, e.g. an import from Excel (with one sheet per product family), or a direct integration with an ERP system. Once created, minor incremental updates can be handled manually using a Web interface.

As products are organized in categories that come with their appropriate attribute sets, catalogue search needs to be a two-step process: first find the product family, then refine the search using attributes. Similar products can be compared using a pair-wise presentation of attribute values. Catalogue aggregation requires re-direction of searches to multiple remote web sites.

The paper by Segev, Wan, and Beam on “Electronic catalogs: a technology overview and survey results” (Proceedings of the fourth international conference on Information and knowledge management December 1995, <http://doi.acm.org/10.1145/221270.221284>) is an even earlier and somewhat dated paper. It seems to identify e-Catalogues with general sell-side B2C commerce site functionality, which nowadays seems to be just one aspect or specific use case for e-Catalogue systems and standards.

6.1.2 Content integration

Stonebraker and Hellerstein introduced the concept of *content integration* in an article for ACM SIGMOD (2001; URL: <http://doi.acm.org/10.1145/375663.375739>). It discusses catalogue integration as a special case of content integration, and as an issue that is prominent with the advent of the third generation of web usage, where the three generations of web usage are defined as:

- Web marketing: document publishing.
- E-Commerce: transaction processing, EAI.
- E-Business operations: B2B, marketplaces, exchanges, and procurement sites.

The characteristic features of the third generation are:

- Integration between enterprises;
- Operational data must be integrated;
- Structured and unstructured data must be easily integrated.

Three vignettes of Third Generation Business Information integration are:

- Integration for republishing: such as MRO supplier catalogue integration (catalogue management).
- Integration of availability and pricing: service catalogue management
- Integration for supply-chain management: forecasts, supply contracts

Content integration is said to deal “with the cross-enterprise integration of items such as product catalogues, product descriptions, product availability, product fulfilment and related information for custom and MRO products for E-Business purposes.” Content integration differs from several other types of integration.

- As opposed to Enterprise Application Integration (EAI), Content Integration also needs to cleanse, normalize and integrate into semantically useful cross-enterprise model. High-level data modelling, declarative languages, intelligent content integration are important issues.
- As opposed to Data Integration (Data Warehousing), Content Integration needs to address cross-enterprise, operational data integration
- As opposed to Content Management, it needs to look at cross-enterprise integration issues.

The authors identify eight main defining characteristics of content integration, grouped in two categories: mapping and query processing. The issue of mapping is linked to access and transformation issues:

- There are varying relations between information owner and integrator (technologies used range from EAI to web screen scraping)
- Different content formats need to be addressed. Beyond format issues there are semantic issues like terminology. This raises a requirement for transformation tools (GUIs, scripting) that are easy to use, even for non-technical people. Standards for catalogue representation are useful in this context.
- Multiple schemas and multiple taxonomies are a reality of life. Classification systems like UN/SPSC. Codes organized in hierarchies. New or custom taxonomies require mappings between taxonomies.
- Requirements for custom syndication (e.g. personalization, custom pricing). “Supplier enablement” using common standards.

The issue of query processing is discussed in section 6.1.5.

6.1.3 Coverage of real-world business practice

Electronic catalogues and the software systems and electronic processes that support them are electronic counterparts of paper artefacts and processes in the real world, so they need to be able to provide the same functionality as offered by paper catalogues and manual business processes.

One aspect of this general issue is price modelling. Kelkar, Leukel and Schmitz discuss how electronic product catalogues based on XML address this issue. They identify six determining factors (such as territory, customer, interval) and nine types of allowances and charges (such as relative percent basis, quantity scales, calculation order). They then evaluate 6 XML e-Catalogue vocabularies with respect to this model and looks at vocabulary inter-translation issues. This evaluates shows that many features are not implemented in any of the XML vocabularies. An example is multi-stage discount systems.

This paper also identifies representation of complex goods as a research topic.

6.1.4 Management issues

In a paper “Managing electronic interchange of business documents”, Banerjee and Kumar (2002; URL: <http://doi.acm.org/10.1145/514236.514241>) describe a framework to guide management decisions on electronic business document interchange applications. One table maps business and operational objectives to depth (of integration), breadth (coverage of trading partners) and diversity (of business documents). Other tables map infrastructure options (EDI, Web forms, outsourced services, intermediaries etc.) to impact on skills, costs, required management support for the dominant large (DLTP) and other trading partners (OTP), and indicate impact on diversity, breadth and depth for DLTP and OTP. These tables allow managers to decide, based on input on business objectives, readiness of trading partners, existing systems, competitive and legal environment, technical factors, to select the most appropriate technical solution.

An early paper by Baron Shaw and Bailey on “Web-based e-catalog systems in B2B procurement” (Communications of the ACM, May 2000, <http://doi.acm.org/10.1145/332833.332845>) defines concept and role of e-catalogues in procurement. It analyses the (generic) procurement process and distinguishes direct materials procurement from MRO (maintenance, repair and operations) procurement. A centralized Web-based procurement system (WBPS) communicating with suppliers using an Extranet can reduce complexity and costs. Case study analysis of factors involved in a move to a WBPS for MRO procurement, divided in various categories/dimensions:

- Internal/external
- Economic/social/legal/technological
- Tangible/intangible/critical success/managerial issues.

Five tangible factors are cycle time (can be reduced by 33%), transaction cost (reducible by 25%), error rate, inventory, item prices. Four intangible factors are business process integration, security, technology, and operations. The intangible factors were not found to be major impediments. A critical success factor is that conflicts may arise between stakeholders (user, procurement, supplier) in flexibility and control. This needs to be addressed before the project starts.

The case study already had a high degree of EC, which may impact generalization of the model. There is no established formula for measuring net value. Projects should first make

an analysis to avoid problems like supplier reluctance, culture clash, systems incompatibility, hidden costs, security holes and processing bottlenecks.

6.1.5 Database technology

Wang and Wang's paper on "Web and e-business application: Optimizing relational store for e-catalog queries: a data mining approach" (Proceedings of the 17th symposium on Proceedings of the 2002 ACM symposium on applied computing March 2002; <http://doi.acm.org/10.1145/508791.509015>) addresses the issue of attributes/values in e-catalogues in a relational database. As many attributes in catalogues are only appropriate for certain types of products, a general schema would result in many nulls (sparse tables). The authors propose to use data mining to find groups of co-occurring attributes and to create a storage schema that best matches the data.

In section 6.1.2, we mentioned an article by Stonebraker and Hellerstein that identifies a number of query processing-related issues:

- Content is operational and highly volatile. Therefore, there is a need to fetch data on-demand (for current information) and fetch in advance (for performance, cached). Data warehouses assume fetch in advance and therefore cannot retrieve some data on demand. ETL (Extract-Transform-Load) systems often use batch processes, which are hard to interface with.
- Catalogue systems must support ad hoc queries. With XML content, this uses XPath or Xquery.
- Need to support information retrieval functionality (synonyms, fuzzy search).
- High-availability, load balancing, and scalability. Most content should be available all of the time. So there is a need for an "adaptive, load-balancing federated query processor".

Abiteboul (1999) looks at XML data from a database management perspective. If an XML document (such as an e-Catalogue) is composed as a view on (potentially) several existing databases, mechanisms are needed to avoid having to recomputed a complete view if just a portion of the data changes. If the e-Catalogue is editable, mechanisms are needed to propagate modification back to the source data. Presentation issues arise when some (some of) the sources are temporarily unavailable.

A similar discussion of implications of e-Catalogues for database storage is given in Jhingran "Moving up the food chain: supporting e-commerce applications on databases" (ACM SIGMOD Record December 2000 <http://doi.acm.org/10.1145/369275.369287>).

6.1.6 Standards adoption

Several papers provide the rationale and design of specific standards and specifications for e-catalogues. These standards are described in other sections of this CWA. For instance, Meltzer and Glushko's "XML and electronic commerce: enabling the network economy" (ACM SIGMOD Record December 1998, <http://doi.acm.org/10.1145/306101.306105>) introduces CBL (the predecessor of the xCBL specification), originally designed by as a standard library for e-business by Veo Systems. Similarly, the BMEcat specification is discussed in various publications including Hümpel and Schmitz (2001).

The papers referenced in section 6.1.3 discuss catalogue standards with respect to specific requirements.

Other papers discuss issues with adoption of specific standards, such as migration. Sena Arpinar, Asuman Dogac's paper on "Provision of market services for eCo compliant electronic marketplaces" (ACM SIGMOD Record September 2000, <http://doi.acm.org/10.1145/362084.362098>) discusses the adoption of the Common Business Library (CBL) e-Catalogue with an existing e-marketplace. The CBL vocabulary is discussed in another section of the draft CWA.

6.2 Organizations active in e-Catalogue Research

The articles discussed in the survey presented in section 6.1 have been authored by researchers in industry or academia (George Mason University, IBM Almaden, IBM Watson Research lab, Middle East Technical University, Politecnico di Milano, Rutgers University, University of California, Berkeley, University of Illinois, University of North Carolina) and by experts working for companies developing e-Catalogue solutions and e-business products (Cohera, Saqqara, Veo Systems). Some of these papers have been authored by two active European research groups with a focus on standardization are:

- Fraunhofer IAO; URL: <http://www.e-business.iao.fhg.de/>
- University of Essen; URL: <http://www.bli.uni-essen.de/>

In addition to these organizations, research on e-Catalogues, procurement and e-business standardization is done by a number of academic organizations:

- The University of Edinburgh (UK), Delft University of Technology (Netherlands) and RWTH in Aachen (Germany) are setting up a European research network in the general area of Information Society Standards which is likely to have a heavy e-business focus. Web page forthcoming, temporary URL: <http://www.ed.ac.uk/~ehja36/>
- Donau University in Austria <http://www.donau-uni.ac.at/>; <http://www.bmoe.at/>
- Fachhochschule Darmstadt, Campus Dieburg, Fachbereich Wirtschaft, URL: <http://www.ifcc.de/>

More general academic e-procurement oriented research:

- UTIPS (Twente), URL: <http://www.sms.utwente.nl/frameset.asp?objectID=1141>
- IPSD (Eindhoven), URL: <http://www.tm.tue.nl/ipsd/>
- Groningen, URL: <http://www.griponpurchasing.com/>
- Centre for Research in Strategic Purchasing and Supply, URL: <http://www.bath.ac.uk/crisps/home.htm>

7 Analysis of existing e-catalogues [WP 1.3]

<XX>

8 Analysis of the relation between (as well as application of) existing product classification schemes/product identification schemes and e-catalogues [WP 1.4]

The chapters below will contain an analysis of the relation between (as well as application of) existing product classification schemes/product identification schemes and e-catalogues.

The focus in this chapter is on major product classification and product identification schemes and their relation to e-catalogues in virtual market places as well as in intra-company use, duly taking into account pan-European requirements.

8.1 Introduction

Nowadays globalization leads to increasing stress of competition. Enterprises therefore take steps towards reduction of costs and rationalization. In the past a wide diversification of all economic activities was a good strategy to ensure the success of an enterprise in the future. Today these enterprises are restructured so that their activities focus on their core competences again. Therefore complex products are as a rule not developed and manufactured by a single firm. Many globally distributed companies with different core competences cooperate in the design and manufacturing of a product. The success of global product development depends among other things on an easy exchange of design knowledge. The amount of design knowledge increases steadily. That is why it is imperative that design knowledge is formally documented so that it can be processed by a computer. To be able to exchange formally documented design knowledge between globally distributed enterprises, a standardized format for the knowledge documentation is needed. A promising possibility is the use of digital parts libraries according to the ISO 13584 PLIB (parts library) standard.

Integration in manufacturing industries needs the capability to model, generate, exchange and manage intelligent electronic catalogues of pre-existing components intended to be used in the design, manufacturing or operation of other products.

Such catalogues shall not only support all the mechanisms needed for selecting a particular component, they must also be able to generate any discipline-specific representation of these components (e.g. three-dimensional shape, simulation model) and they must be based on the same information model to enable their integration in the same repository.

Product Data Management today

The goal of Product Lifecycle Management (PLM) is to handle all information during product lifecycle which is generated in processes in the value chain through the extended enterprise.

Information used or served by one process of the value chain must be easily available for other processes.

In contrast, nowadays, Product Data Management is handled in most enterprises by encoding information about products in numbering systems and describing products in different database tables (“master data records”) leading to an overhead in data conversion efforts between systems, processes and enterprises and thus resulting in longer time-to-market.

Product Identification Codes

Identification codes are used to make an unambiguous identification of a thing. Some years ago, products were identified by identification codes which allowed to anticipate the characteristics of a product. Today identification codes have no special semantics; each time a new code for a product is needed, the code is simply incremented. The one-to-one correspondence between the code and the thing is very useful for recording and linking records of items and actions taken on the items (such as point-of-sale transactions, inventory management, record keeping).

Product Classification

The task of *product classification* is to assign each product to a product group (called “class”) corresponding to common attributes or application areas. Mostly each product gets a classification code assigned to group similar things into common categories. The values of the classification codes are arbitrary, i.e. are not deduced from the products properties. With product classification, similar things are members of a class. Similar classes are members of yet a more general class or family, and so on. Product classification is mainly used for identifying where expenditures are being made, and for searching and comparing offered products on electronic markets. Hence, product classification systems are heavily used in procurement.

To augment the usability of product classification systems for more general purposes, some product classification systems define sets of attributes [properties in our terminology] associated to classes (*Jörg Leukel, Volker Schmitz, Frank-Dieter Dorloff: “Referenzmodell für den Austausch von Produktklassifikationssystemen im E-Business”, 2002*). A set of attributes [properties] is assigned to a classification group [class] and contains therefore at least the set of the necessary product properties for this classification group [class].

In e-catalogues that claim to support a specific classification system each product has been augmented by the classification code and the set of group-dependent list of attributes/value pairs.

Today, almost every day, a new, more or less sophisticated product classification appears (each with a different viewpoint, each with different classes and each associated with a different set of attributes)

Mostly all currently available intra-company wide part management solutions structure their content in respect to similar company or domain specific hierarchical product classification methodologies. Changes in the classification structure, i.e. resulting from a company takeover or project cooperation, needs a complete rework of all products to get the classification keys updated.

Due to the huge amount of different products that are daily used and introduced, the goal of product classification can grow immensely complex.

A selection of parts on the basis of technical requirements is not possible (for example “bolt with anti-twist device and head without eyebolt-bore”).

A selection of objects in regards to a specific target is only possible for a limited amount of criteria. Parts can be collected and arranged in a classification system but not fully

described. The primary focus of a classification system is to supply an identification and structuring mechanism.

Supporting the exchange of data within the product life cycle is not the target of a product classification system.

8.2 Criterias for Classification

This subsection describes the criterias which were relevant for the survey.

8.2.1 multilingual aspects of existing product classification/identification schemes

One of the important topics which differentiates the European Union from the United States is the language. Europe has several languages which are actively used and therefore the multilingual aspect is one of the pan-european requirements in high focus.

8.2.2 international character and authoritative nature

A standard is as good as it is used. This topic deals with the role of the classification/identification schemes in the industry.

8.2.3 speed of updating and change management

<xx>

8.2.4 diffusion of support software

<xx>

8.2.5 upward and downward compatibility of versions

<xx>

8.2.6 application in e-catalogues and ontologies

<xx>

8.2.7 relation to terminology and ontologies

<xx>

8.3 Overview of existing product classification/identification schemes

The section below gives an overview of selected product classifications and product identification schemes.

8.3.1 CPV - Common Procurement Vocabulary

The Common Procurement Vocabulary CPV (URL: <http://simap.eu.int/EN/pub/src/main5.htm>) was created in 1996 by the European Community as a tool for improving transparency and efficiency in the field of public procurement. Use of standard terms in the CPV makes it easier for potential suppliers to

identify the procurement contracts in which they are interested. The CPV also facilitates fast and accurate translation of contract notices for publication in the EC Official Journal, and makes it easier to establish procurement statistics.

The Classification of Products by Activity (CPA) was chosen as a basis to build on. It was done so mainly because the CPA was a coherent nomenclature and was well adapted to the industrial structure of the European Community. However, it is supplier-oriented rather than purchaser-oriented. Consequently, many changes were needed.

CPA codes have six digits. Two more digits were added in order to make codes more accurate. Note too that there is a ninth digit, in order to check if the eight others are correct. These codes and the definitions of products (goods and services) in all the languages of the European Union constitute the main vocabulary. The supplementary vocabulary is designed to help contracting entities describe the subject matter of contracts comprehensively. It includes a letter that defines the general field concerned (transports, textiles etc.) and four digits, the fourth one being a check-digit.

Structure of the classification system

1. The CPV consists of a main vocabulary and a supplementary vocabulary.
2. The main vocabulary is based on a tree structure comprising codes of up to nine digits associated with a wording that describes the supplies, works or services forming the subject of the contract.

The numerical code consists of 8 digits, subdivided as follows:

- the first two digits identify the divisions;
- the first three digits identify the groups;
- the first four digits identify the classes;
- the first five digits identify the categories.

Each of the last three digits gives a greater degree of precision within each category. A ninth digit serves to verify the previous digits.

3. The supplementary vocabulary may be used to expand the description of the subject of a contract. The items are made up of an alphanumeric code with a corresponding wording allowing further details to be added regarding the specific nature or destination of the goods to be purchased.

The alphanumeric code is made up of:

- a first level comprising a letter corresponding to a section;
- a second level comprising four digits, the first three of which denote a subdivision and the last one being for verification purposes.

The use of the CPV when drafting public procurement notices in the EU remains optional. Its use will become mandatory with the entry into force of the forthcoming public procurement legislative package which is now in discussion in the Council and the European Parliament.

8.3.2 eCI@ss

eCI@ss (URL: <http://www.eClass-online.com>) was founded 1999 by leading german companies to developed a common classification system. The classification should set a standard for information exchange between suppliers and their customers. eCI@ss is

characterised by a 4-level hierarchical classification system with a key-word register of about 12,000 words. eCI@ss maps market structure for industrial buyers and supports engineers at development, planning and maintenance. The user can access eCI@ss either via the hierarchy or over the key words. Also lists of attributes for the description of the products are available.

Material Class Hierarchy

The Material Class Hierarchy is a four level tree. It is the supporting structure to which keywords and attribute sets are attached. The nodes of the tree are collectively called “Material Classes”. The Material Classes on the different levels have different names:

Each Material Class is associated with the following characteristic fields:

Identifying Key	Primary key that uniquely identifies a Material Class.
Material Class Code	<p>Eight-digit code that describes the position of the corresponding material class in the hierarchy. It is a concatenation of the parent’s code and a two-digit code relative to the parent.</p> <p>Segments have a two-digit code relative to the root, Main Groups have a two-digit code relative to their parent Segment and a four-digit code relative to the root, and so on. Material Class codes are right-padded with zeroes if necessary.</p> <p>A Material Class code uniquely identifies a Material Class in a given version of eCI@ss.</p>
Identifying Name	<p>A collective term that describes the goods and services belonging to the corresponding material class.</p> <p>An Identifying Name uniquely identifies a Material Class in a given version of eCI@ss.</p>
Short Name	<p>A variant of the Identifying Name whose length is limited to 40 characters.</p> <p>If the Identifying Name does not exceed 40 characters it will also be used as the Short Name.</p> <p>A Short Name uniquely identifies a Material Class in a given version of eCI@ss.</p>
Definition	A precise textual description of the goods and services belonging to the corresponding Material Class.
eCI@ss Version	Version of eCI@ss for which these values of the characteristic fields are valid.

Identifying Names, Short Names, and Definitions must be provided for each supported language. The above statements about the uniqueness of Names hold for each language separately. The same Names may appear in different languages, either with the same meaning (“Hammer” in English and German) or with different meanings (“Container” in English and German).

8.3.3 ETIM

<xx>

8.3.4 GCI / EAN

<xx>

8.3.5 NATO Codification System (NCS)

Over the last 50 years, the NATO codification system (NCS) (URL: <http://www.nato.int/structur/AC/135>) has been developed to support NATO logistic. It was very successful to define a common international understanding about items of supply, to reduce supplies of military forces, and to facilitate logistic data management.

The NCS is based on two main ideas:

- characterizing form, fit and function of each item of supply (IoS) by means of human readable and computer-sensible classification structures allowing to name items (approved item name: AIN) and to describe them by property values (Master Requirement Codes: MRCs); Item Identification Guides (IIGs) specify which properties shall be used to characterize any item belonging to one AIN among a set of similar AINs;
- identifying each IoS by an absolute identifier (NATO Stock Number: NSN), whoever is its manufacturer.

Implementation of the NCS resulted in a complete computer technology for modeling, storing and exchanging item data. Moreover, all kinds of items used by NATO, this means most existing kinds of products, were modeled using this technology.

NATO was a precursor. Only during the last ten years the requirement to systematically characterize product by names and properties emerged in the industry. The first reason was the need of components libraries for computer aided design systems, where the designer could search for component by fit and function. The latest, and not least reason, was the dramatic growth of e-commerce on technical products over the last period.

NCS is designed to achieve high effectiveness in logistics support and to facilitate material data management. The System has been agreed by all signatories of the NATO Alliance for use in identifying equipment and supplies (URL: http://www.nato.int/structur/AC/135/ncs_guide/e_guide.htm)

The principal objectives of the System are:

- To increase the effectiveness of the logistic systems
- To facilitate data handling
- To minimise logistics costs of user nations
- To increase efficiency in logistics operations

To achieve these objectives the System provides for each “Item of Supply” to be assigned with:

- A unique Item Name.
- A unique classification.
- A unique identification.

- A unique NATO Stock Number.

NATO STOCK NUMBER - NSN

When it is established that an Item of Supply is unique, its identity is fixed through the assignment of its own NATO Stock Number (NSN). NSNs are issued by NCBs.

The NSN is a 13 digit number and is divided into 3 parts: The first 4 digits are the NATO Supply Classification Code and relate the item to the group and class of similar items. The next 2 digits indicate the NCB assigning the NSN. The final 7 digits of a NSN are devoid of inherent significance. However, this number is assigned to one and to only one Item of Supply within the codifying country (URL:

<http://www.nato.int/structur/AC/135/NCS/nsn/nsn.htm>). Within the NCS the term MANUFACTURER covers the whole range of possible sources for technical data for items entering the supply systems of participating countries.

8.3.6 Proficlass

Proficlass (URL: <http://www.proficlass.org/>) is a german initiated sectoral spanning classification system for product data.

8.3.7 RNTD -RosettaNet Technical Dictionary

RosettaNet (URL: <http://www.rosettanel.org/>) is a self-funded, non-profit organization of more than 400 Information Technology (IT), Electronic Components (EC), Semiconductor Manufacturing (SM) and Solution Provider (SP) companies working to create, implement and promote open e-business process standards. These standards form a common e-business language, aligning processes between supply chain partners on a global basis. RosettaNet is a subsidiary of the Uniform Code Council (UCC) (URL: <http://www.ucc-council.org/>).

RosettaNet Standards

Besides several standards link defining business processes between trading partners RosettaNet has developed the RosettaNet Technical Dictionary (RNTD) and the RosettaNet Business Dictionary (RNBD).

RosettaNet dictionaries provide a common platform for conducting business within the supply chain, eliminate overlapping efforts by individual companies and reduce confusion in the procurement process due to each company's uniquely defined terminology.

- RosettaNet Business Dictionary:
During the development of RosettaNet standards, business properties are identified to define transactions between trading partners. The RosettaNet Business Dictionary designates the properties used in basic business activities, serving as a central repository of the properties for reference and reuse in additional standards.
- RosettaNet Technical Dictionary:
The RosettaNet Technical Dictionary (RNTD) is the reference model for the classification and characterization of all products manufactured in the supply chains in which RosettaNet is established. With hundreds of manufacturers, distributors and resellers describing thousands of products in different ways, the RosettaNet Technical Dictionary provides common language for defining products and services.

4.1.1 ECCMA Open Technical Dictionary (eOTD)

The Electronic Commerce Code Management Association (ECCMA) (URL: <http://www.eccma.org/>) is a new breed of standards organization focused on satisfying the needs of electronic commerce- the need for true international consensus in real time. ECCMA was formed in April of 1999 and has brought together experts from around the world and provides them a means of working together in the fair, open and fast environment of the Internet to build and maintain global, open standard dictionaries that are used to unambiguously label information.

Mission of ECCMA

To provide an open, transparent and efficient process for the development of open standards for content over the Internet.

The ECCMA Open Technical Dictionary (eOTD) is a collection of terms and definitions used to create encoded, language independent, description of products, organizations and location of supply to facilitate the unambiguous exchange of information between trading partners.

The EGAS is a dictionary of common descriptive terms used to describe the properties, characteristics or requirements of products, organizations or locations of supply. The EGAS is used to build standardized catalogs and product specifications as well as to support enhanced search engines.

The EGIS is a dictionary of names and definitions for common commercial concepts of products, organizations and locations. Also known as Standard Item Names(SIN), Approved Item Names (AIN), Product classes or Product families; a generic concept (good or service) that has a defined list (template) of attributes.

8.3.8 UNSPSC Universal Standard Products and Services Classification

In spring 2003 the Uniform Code Council, Inc. (UCC) (URL: <http://www.uc-council.org/>) has been selected by the United Nations Development Programme (UNDP) to serve as code manager of the United Nations Standard Products and Services Code (UNSPSC) (URL: <http://www.unspsc.org/>). The UNSPSC is a global standard for classification of products and services that is designed to facilitate electronic commerce, resulting in reduced organization's costs, and improved supply chain efficiencies.

The United Nations Standard Products and Services Code is a classification convention that is used to numerically identify products and services.

It is an efficient, accurate and flexible classification system for achieving company-wide visibility of spend analysis, as well as, enabling procurement to deliver on cost-effectiveness demands and allowing full exploitation of electronic commerce capabilities.

The UNSPSC™ was developed jointly by the UNDP (United Nations Development Programme) and D&B (Dun & Bradstreet Corporation) in 1998.

8.4 Comparison Table

Chapter		4.3.1	4.3.2	4.3.3	4.3.4	4.3.5	4.3.6	4.3.7	4.3.8	4.3.9
		CVP	eCI@ss	ETIM	GCI /EAN	NATO	Proficlass	RNTD	eOTD	UNSPSC
4.2.1	supported languages									
	czech		x							
	danish	x								
	dutch	x								
	english	x	x		x	x		x	x	x
	finnish	x								
	french	x	x							x
	german	x	x	x			x			x
	greek	x								
	italian	x	x							x
	japanese									x
	portuguese	x								
	spanish	x	x							x
	swedish	x								
						??				
4.2.2	international character and authoritative nature									
4.2.3	speed of updating and change management		max. once per year				once per year			
4.2.4	diffusion of support software									
4.2.5	upward and downward compatibility of versions									
4.2.6	application in e-catalogues and ontologies									
4.2.7	relation to terminology and ontologies									
	hierarchy levels		4	2		2	not defined up to 6 used			
	hierachy level 1	divisions	segmenet	main group		group				
	hierachy level 2	groups	main group	group		class				
	hierachy level 3	classes	group							
	hierachy level 4	categories	commodity class							
	hierachy level 5									
	hierachy level 6									
	hierachy level 7									
	attibutes		yes	yes		yes	yes			
	cornformity to standards			Edifact BMEcat					ISO 13584	
	properties available			yes			yes			
	no of classes						3760	943		
	no of attributes						3844	3788		
	synonyms		yes				yes			
	attribute lists		yes				yes	yes		
	Version		version 4.1				version 1.2	version 3		V5.0301

8.5 Identification of Problems

<XX>

8.6 Suggestion for Solution to existing (and foreseen future) Problems

<XX>

9 Formulation of problems and recommendations for a systematic approach to implement highly interoperable e-catalogues at SME level [WP 1.5]

<xx>

9.1 Summary of issues raised in analysis sections

<xx>

9.2 Interoperable e-catalogues

<xx>

9.3 Recommendations

<xx>

10 Summary

<xx>

11 Annex <yy> (normative): Title of normative annex

11.1 yy.1 First clause of this normative annex

<Text>

11.2 yy.1.1 First subclause of this normative annex

<Text>

12 Bibliography

The following material, though not specifically referenced in the body of the present document (or not publicly available), gives supporting information.

- Serge Abiteboul. "On views and XML" Proceedings of ACM SIGMOD Record December 1999, Volume 28 Issue 4.
URL: <http://doi.acm.org/10.1145/344816.344853>
- Sena Arpinar, Asuman Dogac. "Provision of market services for eCo compliant electronic marketplaces". ACM SIGMOD Record September 2000.
URL: <http://doi.acm.org/10.1145/362084.362098>
- Rakesh Agrawal, Ramakrishnan Srikant. "On integrating catalogs". Proceedings of the tenth international conference on World Wide Web May 2002.
URL: <http://www10.org/cdrom/papers/076/index.html>
- Sena Arpinar, Asuman Dogac. "Provision of market services for eCo compliant electronic marketplaces" Proceedings of ACM SIGMOD Record September 2000
URL: <http://doi.acm.org/10.1145/362084.362098>
- Snehamay Banerjee, Ram L. Kumar. "Managing electronic interchange of business documents" Communications of the ACM July 2002
URL: <http://doi.acm.org/10.1145/514236.514241>
- John P. Baron, Michael J. Shaw, Andrew D. Bailey. "Web-based e-catalog systems in B2B procurement" Communications of the ACM May 2000 Volume 43 Issue 5
URL: <http://doi.acm.org/10.1145/332833.332845>
- Sherif Danish. "Building database-driven electronic catalogs" Proceedings of ACM SIGMOD Record December 1998
URL: <http://doi.acm.org/10.1145/306101.306103>
- Frank Gilbane. "E-Catalogs: strategic issues for suppliers".
URL: <http://www.gilbane.com/>
- Claus Hümpel, Volker Schmitz. "BMEcat: Produktcataloge austauschen per XML".
URL: http://www.bli.uni-essen.de/publications/2001_XML_HuempelSchmitz.pdf
- Anant Jhingran. "Moving up the food chain: supporting e-commerce applications on databases" Proceedings of ACM SIGMOD Record December 2000
URL: <http://doi.acm.org/10.1145/369275.369287>
- Oliver Kelkar, Joerg Leukel, Volker Schmitz. "Price modeling in standards for electronic product catalogs based on XML". Proceedings of the eleventh international conference on World Wide Web May 2002.
URL: <http://doi.acm.org/10.1145/511446.511494>
- Joerg Leukel, Volker Schmitz, Frank-Dieter Dorloff. "Exchange of catalog data in B2B relationships. Analysis and Improvement".
URL: http://www.bli.uni-essen.de/publications/2002_LeukelSchmitzDorloff.pdf
- Bart Meltzer, Robert Glushko. "XML and electronic commerce: enabling the network economy" Proceedings of ACM SIGMOD Record December 1998
URL: <http://doi.acm.org/10.1145/306101.306105>

- Erica Rugullies. "Choosing the right B2B catalog management solution." Giga information group, 2001. RPA-062001-00002.
URL: <http://www.gigaweb.com/>
URL: <http://www.i2.com/assets/pdf/B2726C08-D8C6-4ACA-BB4DD5ADFBD92DC2.pdf>
- Arie Segev, Dadong Wan, Caroline Beam. "Electronic catalogs: a technology overview and survey results" Proceedings of the fourth international conference on Information and knowledge management December 1995
URL: <http://doi.acm.org/10.1145/221270.221284>
- Michael Stonebraker, Joseph M. Hellerstein. "Content integration for e-business". Proceedings of ACM SIGMOD 2001.
URL: <http://doi.acm.org/10.1145/375663.375739>
- Min Wang, X. Sean Wang. "Web and e-business application: Optimizing relational store for e-catalog queries: a data mining approach" Proceedings of the 17th symposium on Proceedings of the 2002 ACM symposium on applied computing March 2002
URL: <http://doi.acm.org/10.1145/508791.509015>
- eCI@ss: "eCI@ss White Paper 0.6", January 2000
URL: <http://www.eClass-online.com/informationen/download/eClassWhitePaper06.doc>
- NATO Group of National Directors on Codification (AC/135): "Guide to the NATO Codification System", September 2002
URL: http://www.nato.int/structur/AC/135/ncs_guide/e_guide.htm
- Jörg Leukel, Volker Schmitz, Frank-Dieter Dorloff: "Referenzmodell für den Austausch von Produktklassifikationssystemen im E-Business", September 2002
URL: http://www.bli.uni-essen.de/publications/info_d_2002_MKWI.htm
- Jörg Leukel, Volker Schmitz, Frank-Dieter Dorloff: "Exchange of Catalog Data in B2B Relationships - Analysis and Improvement", November 2002
URL: http://www.bli.uni-essen.de/publications/2002_ICWI_LeukelSchmitzDorloff.pdf
- Joachim Quantz, Thorsten Wichmann: "E-Business Standards in Deutschland, Bestandsaufnahme, Probleme, Perspektiven", April 2003
URL: <http://www.berlecon.de/studien/downloads/200304eStandards.pdf>
- DIN: "DIN 4000 Sachmerkmaleisten: Begriffe und Grundsätze", 1981
URL: <http://www.beuth.de/>
- Nikolaus Ondracek: "Überführung bestehender Sachmerkmale in ein Merkmal-Lexikon", Referatensammlung der DIN Tagung Merkmallexikon in der Anwendung, Beuth Verlag 1998
URL: <http://www.beuth.de/>