Do e-catalog standards support advanced processes in B2B e-commerce? Findings from the CEN/ISSS workshop eCAT

Volker Schmitz, Joerg Leukel, Frank-Dieter Dorloff
University of Duisburg-Essen (Campus Essen), Germany
Institute for Computer Science and Business Information Systems (ICB)
[volker.schmitz, joerg.leukel, dorloff]@uni-essen.de

Abstract

This paper presents results from a European standardization project on electronic product catalogs and their application in business-to-business e-commerce. The CEN/ISSS Workshop eCAT aims at formulating a strategy for establishing a harmonized methodology for multilingual e-catalogues, and for implementing this methodology in a future full scale project. As part of this workshop, we will describe basic problems and standardization issues, followed by a detailed analysis of existing standards and standardization initiatives. This will help to reconstruct the state-of-the-art in e-catalog standardization, to identify problems and obstacles that hinder a broader acceptance and diffusion of standards, and finally, to argue on more suitable, harmonized standards.

1. Introduction

Driven by the success and importance of e-procurement and marketplace systems, electronic product catalogs (EPC, e-catalogs) and their supplementing product data have become a cornerstone for conducting e-business [1]. E-catalogs form the basis for buying decisions and the release of order transactions. Suppliers have to provide catalog data for their customers in defined quality and standardized formats. In contrast to business-to-consumer relationships, e-catalog usage in B2B is characterized by the fact that data of the catalog-creating enterprise is imported into an information system of the catalog-receiving enterprise. Hence data exchange is essential, and standards for this data exchange are necessary [2].

To tap the potential of automated, streamlined business transactions, the content of product catalogs has to be specified and transferred in a standardized way. This is already done in e-procurement since the advent of XML as a meta language for defining business vocabularies and machine-readable documents. However, the extensibility provided by XML has led to a variety of e-catalog standards so far. At least 25 e-catalog standards can be identified; 16 of them are based on XML (e.g., BMEcat, cXML, eCX, and xCBL). These standards differ in addressed markets, capabilities to represent product information, market acceptance, and standardization processes.

From a European point of view, a major challenge is to deal with different languages, legal requirements and cultural aspects. Multilingualism in the European Union (EU) 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. This is even more important since the EU has grown from 15 to 25 countries in May 2004. To meet these requirements, standards for e-catalogs must be capable of dealing with multi-lingual product descriptions.

If we look at current e-catalog standardization, we have to state that no standards of standard development organizations (SDOs, e.g., ISO, IEC, ITU) are available. Contrary, most standards are developed by industry consortia. Many standards address vertical or even country-specific needs, thus their relevance to global e-commerce is limited. Standardization processes are seldom transparent and open to new members. In addition, the participation of small and medium-sized companies in these processes is rather small.

2. Standardization and research design

In face of the situation described, CEN/ISSS as a European ICT standardization organization (European Committee for Standardization, Information and Communications Technologies), launched the eCAT Workshop in late 2002 [3]. eCAT aims at formulating a strategy for establishing a harmonized methodology for multilingual e-catalogs, and for implementing this methodology in a future full scale project. The work has been carried out by a project team of six experts from industry and academia, coming from five European countries. The full workshop consists of more than 50 persons and organizations, formally registered to the workshop. Based on their expertise and comments, the project report written by the experts became a CEN Workshop Agreement (CWA) [4]. It consists of an investigation and analysis of e-catalogs for e-business (work item 1), plans and concepts for a standardization strategy (work item 2) and for a pan-European implementation (work item 3).

We will focus on existing e-catalog standards and their capabilities to fulfill business requirements, which is part of work item 1. In Section 3 we will start with our empirical analysis of standards by defining an analysis framework, selecting relevant standards and applying the framework and its criteria to these standards. Section 4 will discuss the results of our analysis and draw conclusions for standardization from it.

Our approach relates to research work that has been carried out on e-business standards both in IS and computer science. Here we can adopt existing models for describing and classifying these standards ([5], [6], [7]). However, an in-depth analysis of e-catalog standards requires additional domain-specific criteria. We will apply criteria and requirements that we have developed in previous work. In particular, they are based on pricing strategies in B2B relationships ([8]), product models ([9], process automation ([2]), and formal specification languages ([10]). This paper presents the concepts and results of a comprehensive study, whereas its domain-specific foundation has been laid in previous work.

3. Analysis of e-catalog standards

3.1. Objects of investigation

E-catalog standards must not be seen as isolated from the higher context of standardization in e-business; therefore they are a specific component within this context. E-business standardization is a generic term for various standardizations in interorganizational and intraorganisational relationships. Here we limit the term e-business standard to those standards that explicitly address interorganizational business processes. Despite their high importance for e-business in general, we do not cover technological standards that deal with core services and infrastructure aspects only (e.g., web service standards like SOAP, UDDI, and WSDL); these standards are not specific for e-catalogs.

The goal of the following discussion is to develop a framework for classifying, describing, comparing and evaluating e-catalog standards. This framework is not limited to e-catalog standards, but covers e-business standards in general. Since we deal with data exchange and business communication, we can fall back on general concepts and models of communication, which is defined as an exchange process of information between a sender and a receiver. Communication models structure and explain communication processes. Many models describe communication by a set of different, hierarchical arranged levels. The definition of levels is a common instrument to structure complex systems. Each level fulfils defined tasks and provides services to higher levels. The best known model is the ISO/OSI reference model, though level models are seen in e-commerce and e-business also (e.g., [11], [12], [13]). They have in common that they assign applications and business rules to higher levels (e.g., e-markets, auctions, negotiation processes), while the lower levels are confined to technical aspects (e.g., internet protocols).

Documents are a key concept in every kind of business communication. This concept includes requirements concerning obligation, deliverability, readability, and storage. Document-orientation is a suitable foundation for the definition of a level model. This can be done coming from two directions. On one hand, the logical structure of documents can be formalized. On the other hand, the role of documents in business process can be determined. The result of this procedure is a level model that consists of the levels data types, vocabulary, documents, processes, framework and metamodel. It is shown in figure 1 and fulfills a second task, since we can classify e-business standards by assigning them to the levels they cover.

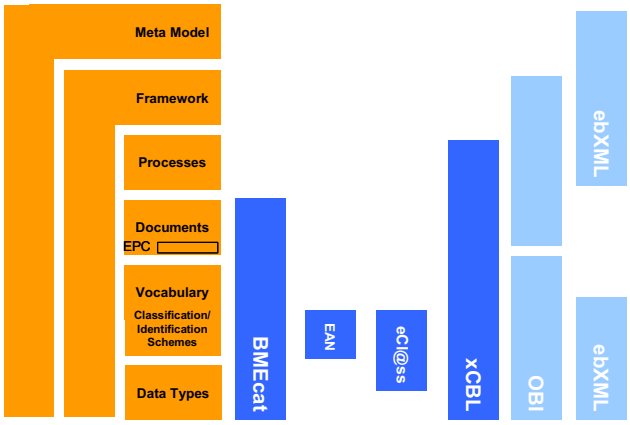

Figure 1. Level model of e-business standards

Literature shows a couple of alternative approaches that either propose models for describing standards or develop a set of criteria for classifying e-business standards. These model-oriented approaches differ in number, subject and definition of levels. They have in common that they propose a hierarchy that builds upon elementary constructs and leads to complete business processes (e.g., [5], [7]).

3.1.1 Data types. On the lowest level, data types are defined and standardized. They are used for typing atomic data elements. Data types are an essential requirement for every kind of electronic data processing. A data type determines the allowed values of a data element and the domain of values respectively. The task of a data type is to code the information that has to be represented by a data element. The codification transforms the information into a defined representation. This concept is a characteristic of all information systems and is implemented in programming languages and database systems as well. For these two areas different sets of data types are available. They differ in number of types and

degree of specialization. Besides general data types like 'string', 'integer' or 'float', data types for representing currencies, countries and date/time information have been standardized by ISO (e.g., ISO 4217, ISO 3166-1, ISO 8601); some of them were adopted by W3C.

3.1.2 Vocabulary. Based on data types, data elements that transfer pieces of information can be standardized. Hence the second level of the level model holds the data element definitions of a standard. In accordance with the language metaphor, the set of permitted data elements builds the vocabulary. It contains these words that are known to business partners in a business communication and therefore can be utilized in a communication process. The vocabulary level is in most cases the core component of document-oriented e-business standards. Developing a vocabulary is the most important domain-specific effort of standardization projects. Prime examples are elements for representing product identification number, product properties and price information.

3.1.3 Documents. Data elements and the implicitly or explicitly formulated conceptual data model form the basis for the definition of business documents that incorporate parts of the standardized vocabulary. The task of the document level is to define permitted business documents. To be more accurate, we have to speak of document types rather than documents, since a document is just an instantiation of a document type. Besides its main function to combine related data to a logical unit, each document type possesses an intended purpose. This means that we can draw a direct conclusion from the document type to the role of both the sender and receiver. For instance, the document type catalog is only meaningful in a communication between the creator and the user of a catalog. The purpose of this document type is providing product information of the sender to the receiver.

3.1.4 Processes. The sequence of exchanged documents and the underlying business logic between two companies are described by the process level. Thus a standard at this level models the order of documents and defines, if necessary, the rules how the receiver has to give an answer following an incoming document. A process can be defined as a transaction or a sequence of transactions between two business partners. Subject of each transaction is the exchange of a document according to agreed document types. With the help of these sequences it is possible to support inter-organizational procurement processes to a full extent. A specific sequence in catalog exchange might be as follows: request for catalog, catalog, catalog update, and catalog import response.

3.1.5 Framework. The framework level covers definitions that relate to technical and therefore domain-independent aspects of business communication. A framework defines the foundation for the communication and provides additional services. All domain-dependent aspects are strictly left to the lower levels. The aim of the framework level is to ensure a secure, dependable and structured exchange of business documents. One characteristic feature of framework services is independency from the content that has to be transferred and the logic that has to be followed. Rather supporting services are described, for example technical communication protocols (e.g., http, SMTP, and ftp), security issues (e.g., authentication, encryption) as well as message handling (e.g., queue management, notification, and acknowledgment services).

A basic concept to reach independency from business content and logic is expressed by the *envelope metaphor*. Very similar to a postal service, it says that the content of a message is kept in a sealed envelope which is the item that has to be transported. The transport requires a meaningful inscription only, which at least specifies the sender and receiver of each message clearly, or gives references to them. In this metaphor, the framework level describes a physical delivery system.

So called framework standards or B2B frameworks (e.g., RosettaNet) possess a close relation to the framework level ([14], [15]). These standards cover at least the framework level, but also integrate lower levels, or even build a customized level model to describe document-oriented business communication. In this interpretation, B2B frameworks are holistic models that support the implementation of e-business applications.

3.1.6 Meta model. The highest level is called meta model level. It aims at providing a generic model that describes the other levels and their relationships. Hence its instances are specific level models; in their most extensive form, these instances are framework standards. The number of standards that fulfill this sophisticated function is very small. A prominent standard is the ebXML standard which is not only a framework standard, but also has many features of a meta model, since it provides generic concepts and tools for modeling e-business communication [15].

3.2 Criteria of the analysis

From a user's point of view, many factors are relevant for choosing a catalog standard. The most important one is its current market penetration and its future potential. The second factor is the quality of the standard itself, in terms of satisfying the requirements from practice and the support which is given for adopting the standard. Because these two aspects can not easily be determined, the quality of the development process has to be taken into consideration, too. This leads to three main groups of criteria for our analysis: the standardization organization, the methodology used in the standardization process, and finally the content of the standard. Next, we explain these criteria in more detail.

3.2.1 Standardization organization. On one hand, the organizational criteria address the standardization organization as an entity that develops, publishes and maintains a standard. On the other hand, the standardization process has to be analyzed. Here we can rely on a set of domain-independent criteria. For the users it is important that the development process is ensured for a long period of time and that the standardization body has the power to bring the standard to a broad application, especially on an international base. In addition, there should be the possibility for the users to participate in the standardization process to bring their requirements in the standard.

3.2.2 Methodology used in the standardization process. The methodology relates to the documentation and the formal specification. The task of the documentation is to describe the content of a standard in such a manner that potential users can easily understand and eventually implement the standard. To achieve this goal, the documentation should meet user requirements. Especially the documentation has to be designed according to the user's knowledge level.

The documentation can be differentiated between the levels of standardization. Some parts of the documentation are often semi-formal or formal specifications in addition to textual descriptions. The close relationship to the formal specification lies in describing the semantics of the standards. Here we understand semantics as the meaning of defined document types and data elements. Only if users know this meaning, they are able to implement a standard correctly, because a common understanding of the semantics and syntax is crucial to e-business communication as it is crucial to any communication.

In view of the high complexity of catalog data, which results in extensive data models, it is suitable to introduce conceptual data models that visualize the general structure. Languages and notations such as extended entity-relationship models, UML (unified modeling language) and graphical representations of XML are used in practice. But these languages are not capable to describe all syntactical and semantic aspects of data elements; hence the most important instruments are data element dictionaries.

The formal specification describes the content of a standard also. This specification fulfils two important functions. First, the use of a formal language can result in more exact descriptions in comparison to non-formal languages. The understanding is facilitated and misinterpretations are prevented, or at least reduced. Second, formal specifications are machine-readable, which supports the implementation of standards in software systems.

According to the level model, different formal languages can be used; some of them are specific for one level only (e.g., process models); other languages cover two or more levels (e.g., data models). XML schema languages are available for specifying document types, data elements and data types. They aim at modeling the structure of XML documents, or parts of these documents. Schema languages provide a set of modeling concepts (e.g., user-defined data types, inheritance, default values, constraints), which are used to a greater or lesser extent by actual catalog standards [10]. XML Document Type Definition (XML DTD) and XML Schema (XSD) are the most important schema languages, since they are standardized by the W3C. XSD has become the prime schema language due to its high expressiveness.

3.2.3 Content. The content quality derives from the capabilities of a catalog standard. It can be assessed by asking whether the standard fulfils the requirements on catalog data. According to the level model this question can be answered by relating and modifying it to specific levels only: First, is the level covered by the standard? Second, what level-specific standardization objects are covered? Third, is the coverage right and satisfying? Checking these issues is a time-consuming task requiring a broad and deep domain knowledge, especially for the vocabulary level that calls for a detailed analysis of the syntax and semantic of all data elements.

When analyzing the *process level* the main question is how the e-catalog is embedded in the whole e-procurement or e-sales process? To answer this question, we rely on a catalog exchange model which contains of all catalog transactions and their respective document types (see figure 2).

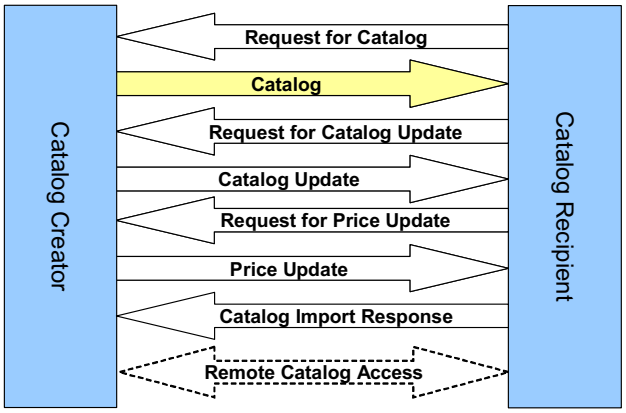

Figure 2. Catalog exchange model

On the *document level* we analyze which features are offered to meet requirements. The analysis is divided into the following parts analogous to the structure of most e-catalog standards.

Document information has to be capable of providing relevant information in the context of the scenarios in which e-catalogs are used. We must consider that catalog data does not mean the data of one specific catalog only. Rather catalog data represents the quantity of data from which catalogs can be created. The creation of specific

catalogs refers to an important characteristic: Each catalog possesses a validity, which can be determined by a set of parameters. Beside the customer, these are the validity period, the currency of prices and the language of all language-dependent data. In accordance with this multi-dependency concept each catalog can be described and identified by a tuple of parameters. It has to be considered that also multi-supplier as well as multi-vendor catalogs can be transferred. Furthermore, the representation of default values and contract information is part of document information.

The most extensive part of the analysis is the examination of *product information*. The way this information is modeled is vital for the use of e-catalogs. Especially product models and price models determine which products can be represented and, due to legal restrictions like taxes, in which countries these e-catalogs can be used.

The third part of the document layer analysis deals with *product relationships*. Three ways of structuring products can be distinguished: product classification systems, catalog group systems, and product references. Catalog group systems are hierarchical structures of product groups which enable an easy top-down navigation in a catalog. It differs from product classification systems by allowing that a product can be assigned to more than one product group and that therefore no group-specific sets of properties can be specified. To easily find related products in a catalog, often links between products are used. Sometimes these links are qualified to describe the relationships between products (e.g., product A is a spare part for product B).

Finally, the vocabulary and data type layer are analyzed. The question is answered to what extent existing standards are reused to prevent reinventing the wheel and introducing new proprietary solutions when there are standards available. The main focus is on language codes, currency codes, logistic information, package units, order units and other common data types.

3.3 Results of the analysis

In this chapter, we select e-catalog standards for the analysis according to the level model, and apply the criteria which were introduced in the previous chapter to them.

3.3.1 Objects of investigation. In order to reconstruct the state-of-the-art in e-catalog standardization, the eCAT workshop identified standardization organizations in this field, listed existing e-catalog standards, and selected relevant standards for a detailed comparative analysis. The attribute relevance was derived from a survey on e-catalog standard adoption by industry (online questionnaire plus 1,500 telephone interviews). This resulted in 251 participants and a return rate of 16% regarding the interviews).

While the survey identified standards actually used in practice, the detailed analysis had to be restricted to a smaller number of standards due to the time frame of the eCAT workshop. Regarding our analysis framework, only those e-business standards were taken into consideration which cover the document layer and provide specifications of e-catalogs (e.g., ebXML does not standardize documents). We emphasize this criterion, since the ISO standard for exchanging product model data, STEP is not considered, because its focus is not on providing product data for e-procurement and e-sales but for engineering and construction. Moreover, we excluded those vertical standards that are highly specific for one industry, or even one country (e.g., Eldanorm, GAEB). In addition, some standardization projects have discontinued their work (e.g., eCOS, OCP). Table 1 lists the remaining and therefore analyzed standards; it shows which level they cover.

Table 1. Coverage of the level model by catalog standards

	BMEcat 1.2	BMEcat 2.0 preliminary draft	catXML 0.21	CIDX 3.0	cXML 1.2009	DATANORM 4	EAN UCC 1.3	eCX 3.0	EDIFACT Pricat 96 A	EDIFACT Prodat 96 B	OAGIS 8.0	OCI 3.0	RosettaNet 2A1_MG_R02_00_00A	xCBL 4.0
Meta Model														
Framework												+	+	
Processes														
Other processes				+	+		+				+		+	+
Catalog exchange processes	+	+		+	+		+	+			+	+	+	+
Documents														
Other documents				+	+	+	+		+	+	+		+	+
Catalog documents	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Vocabulary														
Changeable attributes	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Rigid attributes				+		+	+		+	+			+	+
Datatypes	+	+	+	+	+	+	+	+	+	+	+	+	+	+

+ = level is covered
= level is not covered

The standards can be divided into the following groups:

- Exchange formats developed by e-business software vendors. These are actually no standards, but aim at establishing de-facto (industry) standards. The analysis includes cXML (Ariba, Inc.), catXML (XMLGlobal Technologies, Inc.), eCX (Requisite Technology, Inc.), OCI (SAP AG), and xCBL (CommerceOne, Inc.).
- Horizontal standards proposed by industry consortia: BMEcat (German initiative, leading standard in Europe) and OAGIS (US-dominated).
- Vertical standards proposed by industry consortia: CIDX (global, chemical industry), DATANORM (Germany, trade) and RosettaNet (global, IT &

electronic components industry). The latter is highly accepted in its domain.

- Standards developed by SDOs. The EDIFACT standard provides to message types for e-catalogs, Pricat transfers price information, Prodat is used for exchanging general product information. EAN.UCC is a new XML-based standard by EAN International.

3.3.2 Comparative analysis of the selected standards. Standardization is conducted by software companies, industry consortia and standardization bodies as shown in table 2.

Table 2. Analysis of the standardization organization

	BMEcat 1.2	BMEcat 2.0 preliminary draft	catXML 0.21	CIDX 3.0	cXML 1.2009	DATANORM 4	EAN UCC 1.3	eCX 3.0	EDIFACT Pricat 96 A	EDIFACT Prodat 96 B	OAGIS 8.0	OCI 3.0	RosettaNet 2A1_MG_R02_00_00A	xCBL 4.0
Standardization process														
Standardization body I = Industry consortium C = Company S = SDO	I	I	C	I	C	S	S	C	S	S	I	C	I	C
Supported industry branches N = Not specialized C = Chemical industry B = Building trade G = Consumer goods E = Electronic components / IT I = specialized for different industries	N	N	N	C	N	B	G	N	I G	I B	N	N	E	N
Release policy														
Change frequency L = Low M = Medium H = High	L	L	L	M	L	L	L	L	L	L	H	L	M	H
Draft versions Y = Yes	Y	Y	-	Y	-	Y	Y	-	-	Y	Y	Y	Y	Y
Legal aspects														
Document access F = Free P = Must be purchased	F	F	-	F	F	P	F	F	P	F	F	F	F	F
Input possibilities Y = Yes M = only for members	Y	Y	-	Y	Y	Y	-	-	M	Y	Y	-	Y	Y
Work within working groups M = Members & invited guests	M	M	-	M	-	M	M	-	M	M	M	-	M	-
Membership pre-condition B = Free for buying organizations A = Annual fee F = Fee	B	F	-	A	-	F	F	-	A	-	A	-	A	-
Services for users Y = yes														
Examples	Y	Y	-	Y	Y	Y	Y	-	Y	-	Y	-	Y	Y
Checklists, Guidelines	-	-	-	Y	-	-	-	-	-	-	Y	-	Y	Y
Training and consultancy	-	-	-	Y	-	-	-	-	-	-	Y	Y	Y	Y
SW-tools (API, converter, …)	-	Y	-	Y	-	-	-	-	-	-	Y	-	Y	-
Interactivity (discussion groups, feedback, faq, …)	Y	Y	-	Y	Y	-	-	-	-	-	Y	Y	Y	Y
Certification (SW, data, processes, …)	Y	Y	-	Y	-	-	-	-	-	-	-	Y	-	Y

The catalog standards being developed by software companies are based on the needs to exchange data between software products of these companies. Most of the standardization bodies have a well defined and transparent standardization process. However, the standardization process of the other standardization organizations is often not transparent and not very well documented. Except for the EDIFACT standards, most of the documentations can be downloaded for free over the internet – in some cases a registration is required. Participating in the standardization work is in many cases coupled with a membership; hence it requires paying a membership fee. The support for users in implementing the standards varies very much.

Table 3 presents the results of the analysis regarding the *methodology*.

The documentation is often very poor. There is no multi-lingual documentation or cultural adoption for any of the standards. cXML provided multi-lingual specifications in prior versions, but does only support English language in its current version. There is hardly any group-specific documentation. Especially helping beginners to get a first insight into a standard is a major problem when introducing the standard to the market. In most cases, no real life examples are provided.

Table 3. Analysis of the methodology used in the standardization process

	BMEcat 1.2	BMEcat 2.0 preliminary draft	catXML 0.21	CIDX 3.0	cXML 1.2009	DATANORM 4	EAN UCC 1.3	eCX 3.0	EDIFACT Pricat 96 A	EDIFACT Prodat 96 B	OAGIS 8.0	OCI 3.0	RosettaNet 2A1_MG_R02_00_00A	xCBL 4.0
Documentation														
Document types R = Reference B = Beginner's guide U = User's guide W = Whitepapers	R B	R B	-	U W	U	R	R	U	R	R U	R U	U	R U	R
Printable Y = Yes P = Paper based	Y	Y	Y	Y	Y	P	Y	Y	Y	P	-	Y	Y	-
File formats A = Adobe PDF H = HTML P = Paper based B = Book W = MS Word	A	A H	A	A	A	P	A	A	A H B	P	H	A W	A H W	H
Online documentation Y = Yes	-	Y	-	-	-	-	-	-	-	-	-	-	-	Y
Formal notation Y = Yes	-	-	-	Y	-	-	X	-	-	-	Y	-	Y	Y
Languages E = English G = German	E G	E G	-	E	E	G	E	E	E	G	E	E G	E	E
Formal specification														
Base technology X = XML E = EDIFACT	X	X	X	X	X	E	X	X	E	E	X	X	X	X
Formal specification language X = XSD D = XML DTD	X D	X D	D	D	D	-	X	D	-	-	X	-	D	X
Richness of specification Y = Yes X = in XSD														
Data types														
User defined types	X	X	-	-	-	-	Y	-	-	-	Y	-	-	Y
Use of domain constraints	X	X	-	-	-	-	Y	-	-	-	Y	-	-	Y
Attributes														
Use of domain constraints	X	X	-	-	Y	-	Y	-	-	-	Y	-	-	Y
Use of referential integrity	X	X	-	-	Y	-	-	-	-	-	-	-	-	-
Elements														
Use of cardinalities	X	X	-	-	-	-	Y	-	-	-	Y	-	-	Y
Use of referential integrity	X	X	-	-	-	-	-	-	-	-	-	-	-	-
Use of inheritance	X	X	-	-	-	-	Y	-	-	-	Y	-	-	Y
External datatypes	X	X	-		-	-	-		-	-	Y	-	-	Y
File splitting vs. all-in-one B = Both	B	B	A	S	A	-	S	A	-	-	S	-	S	S

Often even the data element specification (vocabulary) is ambiguous and hardly understandable, thus it can not be used as a basis for implementing a standard correctly. An increasing number of standards use formal notations like UML to specify conceptual data and process models; thus they provide a more precise specification.

Some of the standardization organizations define basic principles for developing their standards; a few even adhere to a custom methodology and meta model. But there is no common methodology or meta model used, therefore the comparison and integration of different e-catalog standards is very difficult. This is a critical drawback for converting e-catalog documents to a different exchange format.

The newer standards are all based on XML technology. These standards initially used XML DTD for their formal specification, but have already or will in the near future support or move to XSD. However, the capabilities of XSD are hardly used. Hence validating XML catalog documents is limited, and the processing in back-end systems is complicated. Especially the missing support of referential integrity by keys and uniqueness constraints is a major obstacle when converting XML data into relational databases [10].

As table 4 shows, most of the standards provide only document types for transferring complete or updating existing catalogs. cXML supports a bilateral coordination process between sender and receiver. Therefore it enables the specification of the requirements on the catalog from the receiver's point of view, and lets him send an import response message which helps to make the import processes easier. There is no continuous support for remote catalog access. cXML and OCI are specialized on this, but only cXML and BMEcat 2.0 provide a full integration.

Table 4. Analysis at process layer

	BMEcat 1.2	BMEcat 2.0 preliminary draft	catXML 0.21	CIDX 3.0	cXML 1.2009	DATANORM 4	EAN UCC 1.3	eCX 3.0	EDIFACT Pricat 96 A	EDIFACT Prodat 96 B	OAGIS 8.0	OCI 3.0	RosettaNet 2A1_MG_R02_00_00A	xCBL 4.0
Request for catalog					+		+				+		+	
New catalog	+	+	+	+		+	+	+	+	+	+	+	+	+
Request for catalog update							+				+		+	
Update catalog	+	+		+	+		+				+		+	+
Request for price update											+		+	
Update prices	+	+									+		+	
Catalog import response					+		+							
Remote Catalog Access		+			+							+		

+ = message is supported
= message is not supported

Besides different methodologies and meta models, each e-catalog standard has an individual semantic, and therefore different data model. Combined with the often unambiguous documentation, the comparison of the standards on the document layer is very difficult and could easily lead to interpretation errors.

Here we present some aggregated results of this comparison (table 5); the complete analysis can be found in [4].

There are two basic approaches to e-catalog data exchange. One approach is to split up the catalog for each scenario. This means that each catalog is used only between one supplier and one buyer in a clearly defined context; therefore the catalog contains only one language, is from one supplier for one buyer, refers to one availability territory and contains only prices valid for one period of time in one currency. The other approach aims at integrating all views of the same catalog in one catalog document. This means that the catalog may include data from two or more suppliers for two or more customers, covers multiple periods of time and different availability areas with associated prices, currencies and languages. The latter approach is especially preferable for marketplaces, because it can reduce the efforts for processing catalog data. The analysis shows that the standards handle this in different ways.

Table 5: Analysis at document layer

	BMEcat 1.2	BMEcat 2.0 preliminary draft	catXML 0.21	CIDX 3.0	cXML 1.2009	DATANORM 4	EAN UCC 1.3	eCX 3.0	EDIFACT Pricat 96 A	EDIFACT Prodat 96 B	OAGIS 8.0	OCI 3.0	RosettaNet 2A1_MG_R02_00_00A	xCBL 4.0
document information														
document identification	+	+	-	-	o	o	o	-	+	o	+	-	-	+
catalog parties	o	+	-	o	-	o	o	+	o	o	+	-	o	o
default values	o	+	-	-	-	-	-	o	o	-	o	-	o	+
scenario support	o	+	-	o	o	-	-	-	-	-	+	-	-	o
product information														
identification	+	+	o	+	+	+	+	-	+	o	o	+	+	+
properties	+	+	-	o	+	-	+	+	o	-	o	-	-	o
price model	+	+	o	+	-	+	+	o	+	+	+	-	-	+
order data	+	+	+	o	-	+	+	-	+	+	-	-	+	o
logistic data	-	+	-	+	-	-	+	-	+	+	-	-	+	-
product configuration	o	+	o	-	-	-	-	o	-	-	-	-	-	-
multimedia attachments	+	+	o	o	o	-	-	-	-	o	+	-	-	o
product relationship information														
hierarchical structures	+	+	o	o	o	o	o	+	o	+	+	-	-	+
product references	+	+	-	-	-	-	o	-	-	o	-	-	o	+

+ = good fulfilment of requirements
o = minimal fulfilment of requirements
- = poor or no fulfilment of requirements

Besides the product ID, a valid product price is the main condition for establishing ordering processes. Therefore the price models in e-catalog standards must meet market-oriented requirements. The characteristics of these price models are industry-specific. In particular, these models must be able to cope with the following aspects [8]: quantity scales, allowances and charges, taxes, different price types (e.g., list prices vs. customer-specific prices), and have to take into consideration the support of various scenarios as described before.

Further requirements arise if complex products should be represented. The coverage of these requirements varies from hardly usable (cXML) to nearly complete models (BMEcat 2.0), but no standard covers all requirements (see [9] for details).

Similar to price models, the specification of order relevant information is essential for product catalogs. The provided possibilities reach from allowing the transfer of the order unit only to specifying all relevant information (e.g., minimum and maximum quantity, quantity interval).

Speaking of logistics information, the different orientations with respect to the covered branches of industry are most clearly. Especially the industry standards that intend to make the order processing and product delivery more efficient include a lot of logistic information (e.g., product dimensions, customs and packing information), while some standards provide no possibilities in this area at all (e.g. cXML, OCI).

It can easily be seen that the use of enumeration types like country or currency codes is handled in different ways (table 6). BMEcat and OAGIS for instance do not define custom enumeration types but reference to other standards (e.g., ISO, UNECE). Contrary, RosettaNet defines some enumeration types on its own.

Table 6. Analysis at data type layer

	BMEcat 1.2	BMEcat 2.0 preliminary draft	catXML 0.21	CIDX 3.0	cXML 1.2009	DATANORM 4	EAN UCC 1.3	eCX 3.0	EDIFACT Pricat 96 A	EDIFACT Prodat 96 B	OAGIS 8.0	OCI 3.0	RosettaNet 2A1_MG_R02_00_00A	xCBL 4.0
reference to international standards	+	+	+	+	+			+	o		+	o	-	+

+ = yes
o = partly
- = no

4 Findings and recommendations

4.1 Enable a more efficient catalog processing

A major obstacle on the way efficient catalog creation processes is the long period of time until a first successful catalog has been transferred from the catalog creator into the catalog processing system (of the buyer or intermediary). The problems which arise on this way are based on inadequacies of the underlying catalog exchange formats.

4.1.1 Enhanced user support. The first problem for catalog creators or software companies which develop catalog processing software like catalog data management systems or e-procurement systems is, that due to the poor documentation, the correct understanding of the standard's semantics is difficult. This follows from the fact that hardly any catalog standard supports its users with documentations which are appropriate for their needs. Most standards provide only some kind of reference documents. These documents are not multilingual and thus especially for beginners hard to understand. Additionally, no real-world example catalog documents are provided. This makes the first catalog creation time-consuming and expensive, because an incorrect catalog is exchanged several times between the involved companies until a correct exchange process is established.

4.1.2 Improved formal specifications. The problems which are caused by these circumstances could be reduced if more standards would provide precise formal specifications. If a specification utilizes the full capabilities of the advanced modeling techniques of XML Schema, the catalog creators could easily use XML tools to validate catalog documents prior to the import process. Moreover, it would be of great benefit if catalog systems create reports that list all errors which occurred during a catalog import process in a qualified manner. These reports would help to decrease the number of circulations between catalog creator and catalog processor.

4.1.3 Support for the coordination phase. Even if the catalog format is well known to all involved parties, some coordination between catalog creators and processors is still necessary. Nearly all catalog standards provide options in the way the standard can be adopted. The agreements which have to be made cover the use of optional data elements, the fixing of enumerations like currencies or languages, and even the restriction of domains (like field length of descriptions). Today, this process is mostly handled on a non-formal way through the exchange of textual guidelines. If all catalog standards would provide a suitable "request for catalog" document type, it would be possible for the catalog processing companies to formulate their needs in a precise formal way.

All these enhancements could lead to straight and lean catalog exchange processes which would save time and money for all involved companies.

4.2 Support of different scenarios

Despite the fact that e-procurement develops towards the integration of global marketplaces, current e-catalogs are often not very suitable for these scenarios. There are special requirements which should be met by e-catalog standards to make the exchange of product data easier for both suppliers and buyers.

When delivering product information to a marketplace, the supplier has to take into account buyer-specific data, especially price information. Therefore he has to transfer some kind of core product data like product description, and additionally buyer-specific prices for each buyer on the marketplace. Especially the updating of this data could be reduced if only the modified data would be transferred, whereas the unchanged core data remains on the marketplace. E-catalog standards are able to implement the distinction between public and private data by providing some kind of multi-buyer capabilities. This concept can only be used effectively if there is a powerful

price model that copes with different requirements of buyers (e.g., territory-specific tax models).

In addition, marketplaces and catalog hubs often provide catalog aggregation services for buyers who want to import a single catalog only which incorporates the catalogs of different suppliers. To establish this in an effective way, the underlying e-catalog standards should have the possibilities to represent multi-supplier catalogs.

4.3 Enable a broader applicability

4.3.1 Complex products and services. So far, the main object of catalog-based e-commerce systems are standardized products of limited complexity. Among these products are primarily indirect goods that are not input factors for production processes and can not be attributed to manufactured final goods. A common term is MRO goods (maintenance, repair and operations). These indirect goods are characterized by a limited specification, low single values and high order frequencies as well as at the same time a low share in the procurement budget. However, they require a relevant amount of resources for procurement, order and stock receipt management.

These restrictions cause a limited area of application for current e-procurement systems. By extending the capabilities of e-catalog applications concerning product complexity, product models and product data exchange, e-procurement systems could reshape their role as tools for buying direct, complex or strategic goods as well. Therefore e-catalog standards need to broaden their product models in this direction. There are two different ways to integrate more complex products in the catalog-based procurement process.

On one hand, the product can be handled on the sell side and integrated through a remote catalog access. To fully use the possibilities of this approach, it has to be assured that the products handled by these special processes are fully integrated in the catalog along the "normal" products. Today there is no such integration in most of the catalog standards. This results in separated catalogs that fail to represent local and remote products in the same way; for instance, regarding keywords, classification, product search, and order process. This approach is not efficient because the cost reduction potentials through process optimization can not be achieved.

On the other hand, complex products could be represented in e-catalogs with a full description and specification. Therefore the e-catalog standard must provide product models that are capable of describing more complex products which may have to be configured. These product models must provide possibilities to describe the product structure and ensure that only valid products can be ordered through a validity concept based on constraints. In addition, the price models have to meet additional requirements of configuration processes.

4.3.2 Extending price models. E-catalogs contain a variety of product information, essential is price information. Prices are used for buying decisions and following order transactions. While simple price models are often sufficient for the description of MRO goods, other goods and lines of business make higher demands. Speaking of suppliers and buyers, it is necessary to represent more complex price models in e-catalogs. For example, the industrial trade uses multi-staged discount systems along the trade levels. Further requirements are dynamic prices being calculated at the time of order and different types of taxes according to legal conditions in the EU.

References

[1] Baron, J.P., Shaw, M.J., and Bailey, A.D., "Web-based E-catalog systems in B2B Procurement", *Communications of the ACM*, 43 (5), 2000, pp. 93-100.

[2] Leukel, J., Schmitz, V., and Dorloff, F.-D., Coordination and Exchange of XML Catalog Data in B2B, *Proceedings of the 5th International Conference on Electronic Commerce Research (ICECR-5)*, 2002.

[3] CEN/ISSS Workshop eCAT, Business Plan. Retrieved April 29, 2004, from http://linux.infoterm.org/termnet-e/ecatsec.htm.

[4] CEN/ISSS Workshop eCAT, CWA (CEN Workshop Agreement). Retrieved September 1, 2004, from http://linux.termnet.org/index.py?level=level3&id=2&lang=_en.

[5] Bussler, C., "B2B Protocol Standards and their Role in Semantic B2B Integration Engines", *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 24 (1), 2001, pp. 3-11.

[6] Dogac, A., amd Cingil, I., "A Survey and Comparison of Business-to-Business E-Commerce Frameworks", *SIGecom Exchange, 2 (2)*, 2001, pp. 16-27.

[7] Zhao, Y., and Sandahl, K., XML-Based Frameworks for Internet Commerce, *Proceedings of the 2nd International Conference on Enterprise Information Systems (ICEIS 2000)*, 2000, pp. 511-516.

[8] Kelkar, O., Leukel, J., and Schmitz, V., Price Modeling in Standards for Electronic Product Catalogs Based on XML, Proceedings of the 11th International World Wide Web Conference (WWW2002), 2002, pp. 366-375.

[9] Leukel, J., Schmitz, V., and Dorloff, F.-D., B2B E-Procurement Beyond MRO?, *Proceedings of the 6th*

International Conference on Electronic Commerce Research (ICECR-6), 2003, pp. 493-500.

[10] Schmitz, V., Leukel, J., and Dorloff, F.-D., Does B2B Data Exchange tap the full Potential of XML Schema Languages, *Proceedings of the 16th Bled Electronic Commerce Conference,* 2003, pp. 172-182.

[11] Zwass, V., "Electronic Commerce: Structures and Issues", *International Journal of Electronic Commerce*, 1 (1), 1996, pp. 3-23.

[12] Schmid, B., and Lindemann, M., Elements of a Reference Model for Electronic Markets, *Proceedings of the 31st Annual Hawaii International Conference on System Sciences (HICSS-31)*, 1998, pp. 193-201.

[13] WebMethods, *Creating a Standard*. Retrieved April 29, 2004, from http://www.cidx.org/Meetings/Meeting Presentations/0900_xml_standards.pdf.

[14] Shim, S.S., Pendyala, V.S., and Sundaram, M. et al., "Business-to-Business E-Commerce Frameworks", *IEEE Computer, 33 (10)*, 2000, pp. 40-47.

[15] Hofreiter, B., Huemer, C., and Klas, W., ebXML: Status, Research Issues and Obstacles, *Proceedings of the 12th International Workshop on Research Issues on Data Engineering (RIDE02)*, 2002, pp. 7-18.