IDEA GROUP PUBLISHING

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITJ2897

Findings and Recommendations from a Pan-European Research Project:

Comparative Analysis of E-Catalog Standards

Volker Schmitz & Joerg Leukel
University of Duisburg-Essen, Germany

ABSTRACT

This paper describes the methodology of a European research project on e-catalog standards and presents its major findings by formulating recommendations for future standardization work. The comparative analysis of 14 standards was conducted within the CEN/ISSS Workshop eCAT that aims at setting the basis for harmonized e-catalog standards. As part of this workshop, we present a framework and a complementing set of criteria for assessing the quality of e-business standards. Eventually, we apply this methodology to our object of interest. The findings can be regarded as an assessment of the state of the art in e-catalog standardization. By identifying problems and obstacles that hinder a broader acceptance and diffusion of standards, we argue on more suitable, harmonized standards.

Keywords: *B2B e-commerce; electronic commerce standards; electronic data interchange; electronic product catalogs; standards; XML*

INTRODUCTION

Electronic product catalogs (e-catalogs, EPC) have become a cornerstone for conducting e-business transactions both in B2B and B2C e-commerce (Baron, Shaw & Bailey, 2000). The reason is that e-catalogs form the basis for buying decisions and the release of order processes. Speaking of B2B e-commerce, e-catalogs are of special importance for the e-procurement and marketplace systems. 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 are imported into an information system of the catalog-receiving enterprise. Hence, data exchange is essential, and standards for this data exchange are necessary (Leukel, Schmitz & Dorloff, 2002).

Specifying and transferring the content of e-catalogs based on standards is essential in order to tap the full potential of automated, streamlined business transactions. 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 stan-

dards so far. At least 25 e-catalog standards can be identified; 16 of them are based on XML (e.g., BMEcat, cXML, OAGIS, and xCBL). These standards differ in addressed markets, capabilities to represent product information, market acceptance, and standardization processes.

A major challenge in e-catalog applications and their respective standards is dealing with different languages, legal requirements, and cultural aspects. This is especially true from a European point of view. Multilingualism in the European Union (EU) often is 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, and so forth. This is even more important since, as of May 2004, the EU has grown from 15 to 25 countries. To meet these requirements, standards for e-catalogs must be capable of dealing with multilingual product description.

E-CATALOG STANDARDIZATION

Looking at e-catalog data exchange in practice, we have to state that no standards published by standard development organizations (SDOs) (e.g., ISO, IEC, ITU) are available. On the 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.

In light of the situation described, CEN/ISSS (European Committee for Standardization, Information and Communications Technologies), as a European ICT standardization organization, launched the eCAT Workshop in late 2002. 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 in industry and academia from five European countries. The full workshop consists of more than 80 organizations and individuals formally registered in the workshop. Based on their expertise and comments, the project report written by the experts became a CEN Workshop Agreement (CWA) in February 2004 (CEN, 2004).

The remainder of this paper is structured as follows. In Section 3, we start with the definition of an analysis framework, followed by a detailed set of criteria in Section 4. Eventually, we apply these criteria to 14 selected e-catalog standards in Section 5. Finally, we draw conclusions from our analysis by formulating recommendations for future standardization work.

OBJECTS OF INVESTIGATION

E-business standardization is a generic term for various standardizations in interorganizational and intraorganizational relationships. Here, we limit the term *e-business standard* to standards that explicitly address interorganizational business processes. Despite their high importance for e-business in general, we do not cover technological standards that only deal with core services and infrastructure aspects (e.g., Web service standards such as SOAP, UDDI, and WSDL). E-catalog standards address the interorganizational exchange of catalog data; thus, they belong to the context of e-business standardization.

A Level Model of E-Business Standards

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 fall back on general concepts and models of communication, which are 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, hierarchically arranged levels. The definition of

levels is a common instrument to structure complex systems. Each level fulfills defined tasks and provides services to higher levels. The most well known model is the ISO/OSI reference model, although level models are seen in e-commerce and e-business also (Schmid & Lindemann, 1998; Zhao, 2001; Zwass, 1996). 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).

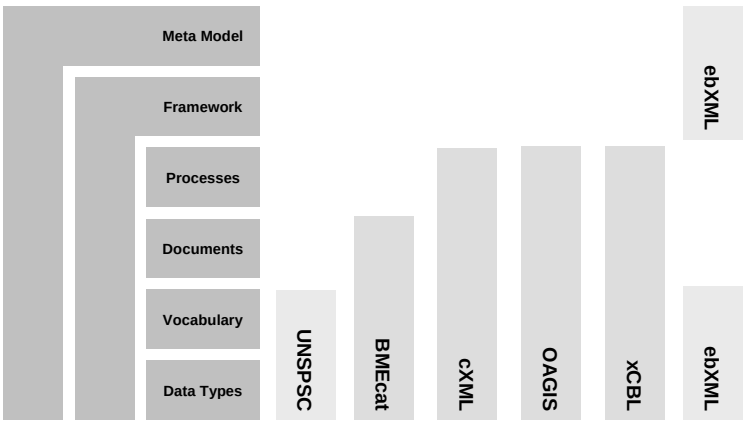
Documents are a key concept in every kind of business communication. This concept includes requirements concerning obligation, deliverability, readability, and storage. Document-orientation is a suitable foundation for the definition of a level model. This can be done coming from two directions. On the 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 (see Figure 1). Moreover, this model fulfills another task, since it allows us to categorize e-business standards (see Figure 3). For instance, product classification schemes, such as UNSPSC, define common terms about how to classify and describe products, but they are not catalog standards. However, the actual classification of specific products belongs to e-catalog documents.

Literature provides a couple of alternative approaches that either propose models for describing standards or develop criteria for classifying e-business standards. These 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 (Bussler, 2001; Zhao & Sandahl, 2000).

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 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 implemented in programming languages and database systems, as well. For these two areas, various sets of data types are available. They differ in number of types and degree of specialization. Besides basic data types like string, integer, or float, data types for representing currencies, countries, and date/time information have been stan-

Figure 1. Level model of e-business standards



dardized by ISO (e.g., ISO 4217, ISO 3166-1, ISO 8601); some of them were adopted by W3C for the needs of representation in the World Wide Web. While these codes are not specific for e-business, other standardized enumerations for packing, units of measurements, and logistics information (e.g., Incoterms) also are relevant to e-catalog standards.

Vocabulary

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

Documents

Data elements and the underlying 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. In addition to 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 the receiver. For instance, the document type catalog is meaningful only in a communication between the creator and the user of a catalog. The purpose of this document type is to provide product information of the sender to the receiver.

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 is defined as a transaction or a sequence of transactions between two business partners. The 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 interorganizational 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.

Framework

The framework level covers definitions that relate to technical and, therefore, domain-independent aspects of business communication. A framework defines the foundation for communication and provides additional services. It aims at ensuring the secure, dependable and structured exchange of business documents. All domain-dependent aspects are left strictly to the lower levels. One characteristic is the 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 address tag 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 (Bussler, 2001; Dogac & Cingil, 2001; Shim et al., 2000; Zhao & Sandahl, 2000). 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.

Meta Model

The highest level is called the meta-model level. It aims at providing a 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 ebXML, 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 (Hofreiter, Huemer & Klas, 2002).

CRITERIA FOR THE ANALYSIS

From a user's point of view, many factors are relevant for choosing the right 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 that is given for adopting the standard. Because these two aspects cannot be determined easily, 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.

Standardization Organization

On the 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 broad application, especially on an international basis. In addition, there should be the possibility for users to participate in the standardization process.

Methodology

The methodology relates to the documentation and the formal specification. The documentation describes 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. In particular, the documentation has to be designed according to the user's knowledge level (e.g., IT professionals, managers, and domain experts).

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 are they 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, resulting in extensive data models, it is suitable to introduce conceptual data models that visualize the general structure (e.g., UML, ERM, graphical representations of XML structures). But these languages are not capable of describing all syntactical and semantic aspects of data elements; hence, the most important instrument is the dictionary of data elements.

The formal specification also describes the content of a standard. This specification

fulfills two important roles. First, the use of a formal language results in precise and unambiguous descriptions compared to non-formal languages. Second, formal specifications are machine-readable, which supports the implementation of standards in software systems.

With regard to the level model, different formal languages can be used; some of them are specific for one level only (e.g., event-driven process chains for process modeling); 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. 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 (Schmitz, Leukel & Dorloff, 2003). XML Document Type Definition (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.

Content

The content quality derives from the capabilities of a catalog standard. It can be assessed by asking whether the standard fulfills the requirements of 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.

Process Level

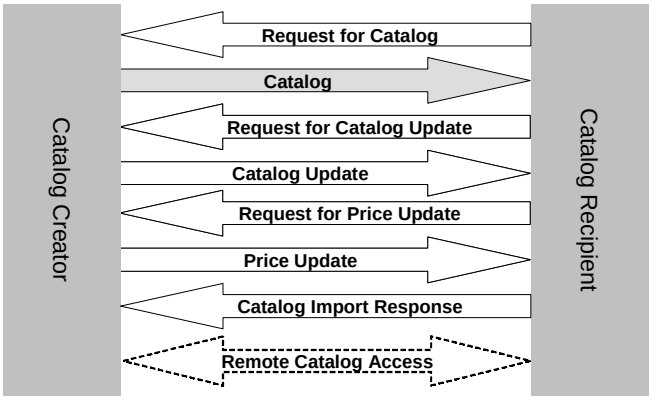
When analyzing the process level, the main question is how e-catalogs are embedded in the entire e-procurement or e-sales process. In this regard, we analyze in more detail which transactions of catalog data exchange are supported. To do this, we rely on a catalog exchange model, which consists of all catalog transactions and their respective document types (see Figure 2).

Document Level

On the document level, we analyze which features are offered to meet requirements. The analysis is divided into three 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 multiple catalogs can be cre-

Figure 2. Catalog exchange model



ated. The creation of specific catalogs refers to an important characteristic: Each catalog possesses a validity, which is determined by a set of parameters. Besides the customer, these are the validity period, the currency of prices, and the language of all language-dependent data. It also has to be considered that multi-supplier as well as multi-vendor catalogs may be transferred.

Product information and its modeling is essential for the quality of an e-catalog standard, because 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.

Product relationship information describes relationships between products. Three ways of structuring products can be distinguished: catalog group systems, product classification systems, and product references. Catalog group systems are hierarchical structures of product groups that enable easy top-down navigation in a catalog. They differ from product classification systems (e.g., eCl@ss, UNSPSC) by allowing one product to be assigned to more than one product group and that, therefore, no group-specific sets of properties can be specified. In order to easily find related products in a catalog, links between products often are used. Sometimes these links are qualified to describe the type of the relationship (e.g., product A is spare part for product B).

Vocabulary and Data Type Level

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 business data types.

RESULTS

In this section, e-catalog standards for the analysis are selected according to the level model. Eventually, we apply the criteria that were introduced in the previous sections.

Objects of Investigation

In order to reconstruct the state of the art in e-catalog standardization, the eCAT workshop identified relevant standardization organizations, listed existing e-catalog standards, and selected standards for a detailed comparative analysis. The attribute relevance was derived from a survey on e-catalog standards 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, due to limited resources. Regarding the analysis framework, only those e-business standards that cover the document layer and provide specifications of e-catalogs were taken into consideration (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). Table 1 lists the analyzed standards and shows which levels they cover.

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),

Table 1. Structuring the objects of investigation with the level model

	BMEcat 1.2	BMEcat 2.0 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 E	OAGIS 8.0	OCI 3.0	RosettaNet PIP2A1	xCBL 4.0
Meta Model														
Framework												+	+	
Processes														
Other processes				+	+		+				+		+	+
Catalog processes	+	+		+	+		+	+			+	+	+	+
Documents														
Other documents				+	+	+	+		+	+	+		+	+
Catalog documents	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Vocabulary														
Changeable attributes	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Rigid attributes				+		+	+		+	+			+	+
Data types	+	+	+	+	+	+	+	+	+	+	+	+	+	+

+ = level is covered
= level is not covered

DATANORM (Germany, trade), and RosettaNet (global, IT, and electronic components industry). The latter is highly accepted in its domain.

Standards developed by SDOs. The EDIFACT standard provides two 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.

Standardization Organization

Standardization is conducted by software companies, industry consortia, and standardization bodies, as shown in Table 2.

The catalog standards being developed by software companies are based on the need 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; in some cases, registering is required. Participation in the standardization work often is coupled with a membership.

Methodology

Table 3 presents the results of the analysis regarding the methodology. The documentation is often very poor. There is no multilingual documentation or cultural adoption for any of the standards. cXML provided multi-lingual specifications in prior versions, but only supports English in its current version. There is hardly any group-specific documentation. Helping beginners to get a first insight into a standard is a particularly major problem when introducing the standard to the market. In most cases, no real-life examples are provided.

Often, even the data element specifications (vocabulary) are ambiguous and hardly understandable; thus, they cannot be used as a basis for implementing a standard correctly. An increasing number of standards provides formal notations like UML for conceptual data and process models.

Some of the standardization organizations define basic principles for developing their standards; a few even adhere to a custom method-

Table 2. Analysis of the standardization organization

	BMEcat 1.2	BMEcat 2.0 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 PIP2A1	xCBL 4.0
Standardization process														
Standardization body	I	I	C	I	C	S	S	C	S	S	I	C	I	C
I = Industry consortium														
C = Company, S = SDO														
Supported industry branches	N	N	N	C	N	B	G	N	I G	I B	N	N	E	N
N = Not specialized														
C = Chemical industry, B = Building trade														
G = Consumer goods														
E = Electronic components / IT														
I = specialized for multiple industries														
Release policy														
Change frequency	L	L	L	M	L	L	L	L	L	L	H	L	M	H
L = Low, M = Medium, H = High														
Draft versions Y = Yes	Y	Y	-	Y	-	Y	Y	-	-	Y	Y	Y	Y	Y
Legal aspects														
Document access	F	F	-	F	F	P	F	F	P	F	F	F	F	F
F = Free, P = Must be purchsed														
Input possibilities	Y	Y	-	Y	Y	Y	-	-	M	Y	Y	-	Y	Y
Y = Yes, M = Members only														
Work within working groups	M	M	-	M	-	M	M	-	M	M	M	-	M	-
M = Members & invited guests														
Membership pre-condition	B	F	-	A	-	F	F	-	A	-	A	-	A	-
B = Free for buyers														
A = Annual fee														
F = Fee														
Services for users Y = yes														
Example catalogs	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, faq, …)	Y	Y	-	Y	Y	-	-	-	-	-	Y	Y	Y	Y
Certification (Software, catalogs)	Y	Y	-	Y	-	-	-	-	-	-	-	Y	-	Y

ology 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 hardly are used. Hence, validating XML catalog documents is limited, and the processing in back-end systems is complicated.

Content

Comparison at Process Layer

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 the receiver 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 spe-

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

	BMEcat 1.2	BMEcat 2.0 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 PIP2A1	xCBL 4.0
Documentation														
Document types	R	R	-	U	U	R	R	U	R	R	R	U	R	R
R = Reference, B = Beginner's guide	B	B		W						U	U		U	
U = User's guide, W = Whitepapers														
Printable	Y	Y	Y	Y	Y	P	Y	Y	Y	P	-	Y	Y	-
Y = Yes, P = Paper-based														
File formats	A	A	A	A	A	P	A	A	A	P	H	A	A	H
A = Adobe PDF, H = HTML		H							H			W	H	
P = Paper-based, B = Book									B				W	
W = MS Word														
Online documentation Y = Yes	-	Y	-	-	-	-	-	-	-	-	-	-	-	Y
Formal notation Y = Yes	-	-	-	Y	-	-	Y	-	-	-	Y	-	Y	Y
Languages	E	E	-	E	E	G	E	E	E	G	E	E	E	E
E = English, G = German	G	G										G		
Formal specification														
Base technology	X	X	X	X	X	E	X	X	E	E	X	X	X	X
X = XML, E = EDIFACT														
Formal language	X	X	D	D	D	-	X	D	-	-	X	-	D	X
D = XML DTD, X = XSD	D	D												
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	B	A	S	A	-	S	A	-	-	S	-	S	S
B = Both, S = file splitting, A = all-in-one														
Internal documentation Y = Yes	-	Y			-	-	-		-	-	Y	-	-	-

cialized on this, but only cXML and BMEcat 2.0 provide a full integration.

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 into one catalog document. This means that the catalog may include data from two or more suppliers for two or more customers and may cover multiple periods of time and different availability areas with associated prices, currencies, and languages. The latter approach is preferable for marketplaces, because it can reduce the efforts for processing catalog data. The analysis shows that the standards handle this in different ways.

Comparison at Document Layer

Besides different methodologies and meta models, each e-catalog standard has an indi-

Table 4. Analysis at process layer

	BMEcat 1.2	BMEcat 2.0 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 PIP2A1	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		+			+							+		

+ = transaction supported
= transaction not supported

Table 5. Analysis at document layer

	BMEcat 1.2	BMEcat 2.0 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 PIP2A1	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 fulfillment of requirements
o = minimal fulfillment of requirements
- = poor or no fulfillment of requirements

vidual semantic and, therefore, a different data model. Combined with the often unambiguous documentation, the comparison of the standards on the document layer is very difficult and easily could lead to interpretation errors. Here, we present some aggregated results of this comparison; the complete analysis can be found in CEN/ISSS (2004).

Besides the product ID, a valid product price is the main condition for executing order processes. Therefore, the price models in e-catalog standards must meet market-oriented re-

quirements. Many characteristics of these price models are dependent on the branch of industry. In particular, these models must be able to cope with the following aspects (Kelkar, Leukel & Schmitz, 2002): quantity scales, allowances and charges, taxes, price types (e.g., list prices vs. customer-specific prices), and support of various scenarios, as described previously.

Further requirements arise when complex products should be represented. The coverage varies from hardly usable (cXML) to nearly complete models (BMEcat 2.0), but no standard covers all requirements (see Leukel, Schmitz & Dorloff 2003 for details).

Similar to price models, order information is essential for e-catalogs. The provided possibilities extend 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 clear. In particular, 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 other standards provide no possibilities in this area at all (e.g., cXML, OCI).

Comparison at Vocabulary and Data Type Layer

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

RECOMMENDATIONS

Enhanced User Support

The first problem for catalog creators or software companies that 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 that are appropriate for their needs. Most standards provide only some kind of reference documents. 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.

Enriched Formal Specifications

The problems that 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 provided by XML schema, the catalog creators easily could use XML tools to validate catalog documents prior to the import process. Moreover, it would be of great benefit if catalog sys-

Table 6. Analysis at data type layer

	BMEcat 1.2	BMEcat 2.0 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 PIP2A1	xCBL 4.0
Reference to international standards	+	+	+	+	+			+	o		+	o	-	+

+ = yes
o = partly
- = no

tems created reports that list all errors that occurred during the import process in a qualified manner. These reports would help to decrease the number of circulations between catalog creator and catalog processor.

Coordinated Catalog Exchange

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 may be adopted. The agreements that 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 (e.g., field length of descriptions). Today, this process is handled in a non-formal way through the exchange of textual guidelines. If 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.

Support Multilingualism

Considering that a multilingual e-catalog is understood as a catalog and that its respective catalog system and all its information, interfaces, and operations for end users are available in more than one language (and typically more than two languages), e-catalog standards play an important role for enabling real multilingualism. Multilingual aspects of e-catalog standards appear on several levels. First, e-catalog standards should be able to transfer catalog content in multiple languages within a single document. All language-dependent data elements should be assigned with a standardized ISO code representing the actual language. Second, e-catalog standards should be able to transfer content of any possible language; this content has to be coded according to transnational encoding schemes. Third, international standards makers should consider providing their documentations, specifications, Web sites, and sample data not only in English but also in other important languages. This will help standards adopters in big markets where the knowledge of the English language is not that common (e.g., East Europe, China).

Support Different Scenarios

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

When delivering product information to a marketplace, the supplier has to take into account buyer-specific data, especially price information. Therefore, the supplier has to transfer some kind of core product data, such as product description and 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 would remain on the marketplace. E-catalog standards are able to implement the distinction between public and private data by providing multi-buyer capabilities.

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

Support Complex Products and Services

So far, the main object of catalog-based e-commerce is standardized products of limited complexity. Among these products are primarily indirect goods that are not immediate input factors for production processes and cannot 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.

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.

Provide Extended Price Models

E-catalogs contain a variety of product information; price information is essential. Prices are used for buying decisions and following order transactions. While simple price models often are sufficient for 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.

Consider Product Life Cycle

While the quality of current e-catalog standards in terms of coverage of business requirements and formal specification has developed in recent years, we have to emphasize that catalog standards address the late phases of the product life cycle only. The integration with standards for product data management (PDM) is very little or non-existent; hence, product-related management that arises during the early life cycle phases (e.g., product planning, design, construction, process planning, and manufacturing) is still the subject of other standards and information systems. With regard to the evolving concept of product life-cycle management (PLM) and respective information systems architectures, e-catalog standards makers at least should consider developments and existing standards for PDM and PLM. For instance, product description based on properties already has been standardized in ISO 13584. Therefore, product models in e-catalog standards can be based on these standards when defining data models and data elements for product properties.

REFERENCES

Baron, J.P., Shaw, M.J., & Bailey, A.D. (2000). Web-based e-catalog systems in B2B procurement. *Communications of the ACM, 43*(5), 93-100.

Bussler, C. (2001). B2B protocol standards and their role in semantic B2B integration engines. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 24*(1), 3-11.

CEN (2004). (2004) CEN workshop agreement 15045. *Proceedings of the 2004 E-Multilingual Catalogue Strategies for eCommerce and eBusiness*, Brussels, Belgium.

Dogac, A., & Cingil, I. (2001). A survey and comparison of business-to-business e-commerce frameworks. *SIGecom Exchange, 2*(2), 16-27.

Hofreiter, B., Huemer, C., & Klas, W. (2002). ebXML: Status, research issues and obstacles. *Proceedings of the 12th International Workshop on Research Issues on Data Engineering*.

Kelkar, O., Leukel, J., & Schmitz, V. (2002). Price modeling in standards for electronic product catalogs based on XML. *Proceedings of the 11th International World Wide Web Conference*.

Leukel, J., Schmitz, V., & Dorloff, F.-D. (2002). Coordination and exchange of XML catalog data in B2B. *Proceedings of the 5th International Conference on Electronic Commerce Research*.

Schmid, B., & Lindemann, M. (1998). Elements of a reference model for electronic markets. *Proceedings of the 31st Annual Hawaii International Conference on System Sciences*.

Schmitz, V., Leukel, J., & Dorloff, F.-D. (2003). Does B2B data exchange tap the full potential of XML schema languages. *Proceedings of the 16th Bled Electronic Commerce Conference*.

Shim, S.S., et al. (2000). Business-to-business e-commerce frameworks. *IEEE Computer, 33*(10), 40-47.

Zhao, Y. (2001). *XML-based frameworks for Internet commerce and an implementation of B2B e-procurement.* Doctoral Thesis. Sweden: Linköping University.

Zhao, Y., & Sandahl, K. (2000). XML-based frameworks for Internet commerce. *Proceedings of the 2nd International Conference on Enterprise Information Systems*.

Zwass, V. (1996). Electronic commerce: Structures and issues. *International Journal of Electronic Commerce, 1*(1), 3-23.

Volker Schmitz is a senior researcher and lecturer in the Institute for Computer Science and Business Information Systems at the University of Duisburg-Essen, Germany. He holds an MSc in computer science from RWTH Aachen University and is a member in several national and international standardization initiatives. Since 1999 he has been responsible for the development of the e-business standards BMEcat and openTRANS. His research and teaching interests include e-procurement, e-markets, catalog and product data management, product classification, business ontologies, and standardization.

Joerg Leukel is a senior researcher and lecturer in the Institute for Computer Science and Business Information Systems at the University of Duisburg-Essen, Germany. He has a PhD in business information systems from the same university, as well as substantial standardization experience in national and European projects on XML-based B2B e-commerce standards. His research and teaching interests are in the areas of e-procurement, e-markets, catalog and product data management, product classification, business ontologies, and standardization.