

CEN/ISSS Workshop eCAT – A Step towards Multilingual Electronic Product Catalogues

V. Schmitz & J. Leukel

Institute for Computer Science and Business Information Systems (ICB), University of Essen, Germany



ABSTRACT: This paper presents the CEN Workshop "Multilingual Catalogue strategies for eCommerce and eBusiness" (CEN/ISSS Workshop eCAT). The eCAT Workshop aims at formulating a strategy for establishing a harmonised methodology for multilingual e-catalogues, and for implementing this methodology in a future full scale project on e-catalogues. The Workshop will not only lay the basis for a future implementation of a harmonised methodology for multilingual e-catalogues, but will also take into account the need for training experts and raise awareness on the subject. These are key issues for the future to help enable a high percentage of SMEs to become successful in e-commerce world-wide. The Workshop will closely cooperate with other initiatives in the field, including EU-funded RTD-projects – thus bundling experiences and competences from the private and public sectors.

1 INTRODUCTION

With the increasing use of e-business electronic product descriptions have gained central importance. The exchange of products descriptions are the base of inter-company processes along the whole product lifecycle. They are used for example to enable concurrent engineering as well as during the information phase in purchasing processes (Baron, Shaw & Bailey 2000).

To tap the potential of e-business in this context the product description has to be specified and transferred in a standardised way. This is already done in some of the sketched areas and has a wide appliance particularly in the field of B2B e-procurement. Especially the use of XML has led to a wide variety of electronic product catalogues standards with different levels of market penetration and a big range of capabilities.

The requirements on electronic product catalogues are based on different determinates, like involved software systems, range of represented products, underlying purchasing process and especially scenario of application. The scenario in which electronic product catalogues are used varies if an intermediary like a marketplace is involved and if the locations of the market participants are on a global base. The main challenge for the use of electronic product catalogues in a pan-european way is dealing with the different languages and pricing/taxes models.

Multilingualism in the Union is often seen as an obstacle for the European economy in terms of competition and the opening up of new markets, but it also has political dimensions relating to consumer protection, freedom to move, etc. However, new economies have emerged in the wake of trying to overcome the language barriers, such as the language industries (including activities and the language technologies for making language resources and terminologies available at a large scale), where Europe has a leading edge thanks to the R&D programmes of the EU Commission (CEN/ISSS Workshop eCAT 2002). To meet these requirements standards for electronic product catalogues as well as the underlying classification systems must be capable of dealing with multi-lingual product description.

This paper aims at presenting the work of the CEN/ISSS Workshop "Multilingual Catalogue strategies for eCommerce and eBusiness". To do so, our paper is structured as follows: Section 2 starts with a short introduction to CEN/ISSS and its mission. This is followed by a brief description what are the main objectives of the Workshop eCAT and how is the work structured. In Section 3 we take a closer look at the work item 1.3 "Analysis of existing e-catalogues". We will outline how the analysis is organised and which are the main subjects of research. We finish the paper with an outlook onto the ongoing work.

2 CEN/ISSS WORKSHOP ECAT

CEN/ISSS was created in mid-1997 by CEN (European Committee for Standardization) as the focus for its ICT (Information and Communications Technologies) activities. CEN recognised that the market needs of the Information Society could not be met through traditional standardisation methods alone, and that a new solution was required.

The mission of CEN/ISSS is to provide market players with a comprehensive and integrated range of standardisation-oriented services and products, in order to contribute to the success of the Information Society in Europe. The current major activities include work in the fields of data protection and electronic signature, e-business standardisation and e-commerce (CEN/ISSS 2003).

The Workshop "Multilingual Catalogue strategies for eCommerce and eBusiness" aims at formulating a strategy for establishing a harmonised methodology for multilingual e-catalogues, and for implementing this methodology in a future full scale project on e-catalogues. The Workshop will not only lay the basis for a future implementation of a harmonised methodology for multilingual e-catalogues, but will also take into account the need for training experts and raise awareness on the subject. These are key issues for the future to help enable a high percentage of SMEs to become successful in e-commerce world-wide. The Workshop is focussed on EU SMEs needs; it will also pay special attention to Central and Eastern European Countries and EFTA countries. The Workshop will closely cooperate with other initiatives in the field, including EU-funded RTD-projects – thus bundling experiences and competence from the private and public sectors. (CEN/ISSS Workshop eCAT 2002).

The project will be carried out by a project team of experts from Danterm & Copenhagen Business School, (Denmark), Fraunhofer IAO (Germany), ISO TC 37 (Norway), Paradine (Austria), Sonnenglanz (Netherlands) and University of Essen, (Germany). The secretariat of the workshop is managed by TermNet on behalf of the Austrian Standards Institute.

The work is divided up into three work items (see figure 1):

WI 1 "Investigation and analysis of e-catalogues for e-business. State-of-the-art concerning the application of e-catalogues for e-business and recommendations for future development".

WI 2 "Plans and concepts for a standardization strategy".

WI 3 "Plans and concepts for pan-European implementation".

The authors of this paper are responsible for work item 1.3 so the paper focuses on this work item.

The objective of work item 1 is to provide an analysis which shall give a comprehensive picture of

what happens in e-cataloguing globally and shall take into account a representative selection of institutions/organisations including e-marketplaces involved in the field, with a focus on the European situation.

The analysis shall identify the existing problems in e-cataloguing and come up with recommendations for a systematic approach to overcoming these problems (CEN/ISSS eCAT 2002).

The work item 1 is divided into the following five tasks:

WI 1.1 "Survey on the use of existing e-catalogues in e-business – schemes, systems/tools, way of application".

WI 1.2 "Survey on existing major institutions/organizations supporting e-catalogue development in e-business and state-of-the-art in pertinent research".

WI 1.3 "Analysis of existing e-catalogues".

WI 1.4 "Analysis of the relation between (as well as application of) existing product classification schemes/product identification schemes and e-catalogues".

WI 1.5 "Formulation of problems and recommendations for a systematic approach to implement highly interoperable e-catalogues at SME level".

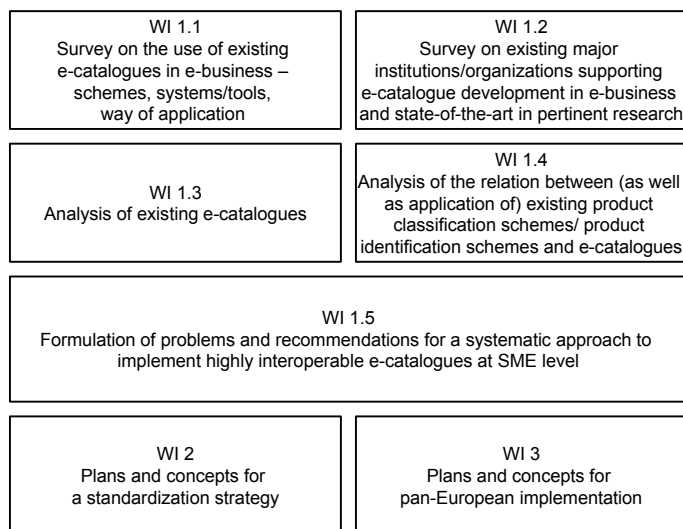


Figure 1. Work Items of WS eCAT.

The following chapter gives an outline over the approach which is followed in work item 1.3.

3 ANALYSIS OF EXISTING E-CATALOGUES

3.1 Determination of the objects of investigation

Standardised catalogue formats must not be seen as isolated from the higher context of standardisation in e-business; therefore they are a specific component within this context. E-business standardisation is a generic term for different standardisations in interorganisational and intraorganisational relationships. Here we limit the term e-business standard to those standards that explicitly address interorganisational

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; these standards are not specific for e-catalogues. To this group of standards belong for instance XML-based standards for describing, registering, calling and executing communication services, especially so called web services (e.g., web service standards like SOAP, UDDI, and WSDL). Contrary we focus on standards that are concerned with domain-specific aspects of business communication.

The goal of the following discussion is to structure the field of e-business standardisation. Since its object is 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. The information is coded according to an accepted language and transferred over a communication medium. Many communications models have been developed both in social science as well as in information systems theory. They structure and explain communication processes. Many models describe communication by a set of different, hierarchical arranged levels. The building of levels is a common instrument to structure complex systems. Each level fulfills a defined task and provides corresponding services for higher levels. The most known model is the ISO/OSI reference model, though level models are seen in e-commerce and e-business also (e.g., Zwass 1996; Schmid & Lindemann 1998). 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 more technical aspects (e.g., internet protocols).

When adopting characteristic features of level models to business communication we have to take into account that all relevant information is exchanged between business partners via paper-based or electronic messages respectively documents (exception: verbal business communication). 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 hierarchical level model. This can be done coming from two different directions. On the one hand the logical structure of business messages has to be formalized. On the other hand the role of messages in business process has to be determined. The result of this procedure is a level model that consists of the level data types, vocabulary, documents, processes, framework and meta-model. It is shown in figure 2 and fulfils a second task, since we can classify real-world catalogue standards by assigning them to those levels that are covered by the respective standard.

With reference to the level model the following groups of catalogue standards can be formed: The group of genuine catalogue standards contains those standards, whose origin is situated in the specification of catalogue documents for e-procurement. To this group belong for example BMEcat (Schmitz, Kelkar & Pastors 2001) and cXML (Ariba 2002). Meanwhile cXML has expanded its scope to further business messages; BMEcat is supplemented by the transaction standard openTRANS (Kelkar, Otto & Schmitz 2001). Transaction standards go a step further in standardizing a multiplicity of business messages; catalogue documents are just a part of it. Prominent members of this group are OAGIS (Open Applications Group 2002) and xCBL (CommerceOne 2002). The third group consists of e-business frameworks, which standardize a complete data and communication infrastructure. The most well-known frameworks are RosettaNet and ebXML. The latter does not standardise business documents.

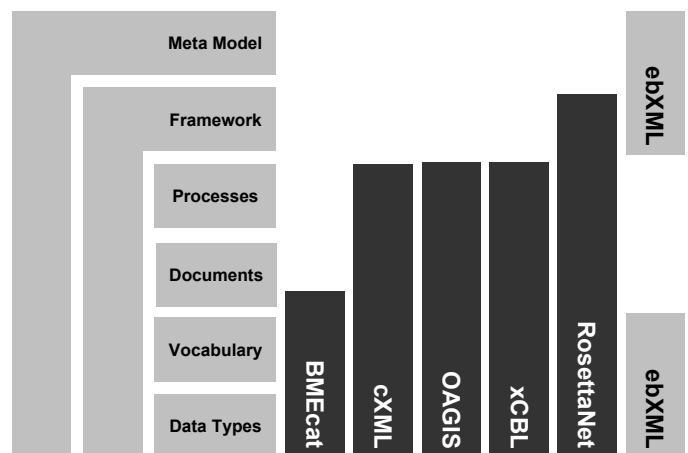


Figure 2. Level Model of E-Business Standardisation.

In the following we describe the levels and their relationships to other levels in a bottom-up sequence. Additionally, we point out those standardisation objects that are relevant for the domain e-cataloguing. 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. The model-oriented approaches differ in number, subject and definition of levels. They have in common that they build a hierarchy that builds upon elementary constructs and leads to complete business processes, (e.g., Zhao & Sandahl 2000; WebMethods 2000, and Bussler 2001).

3.1.1 Data Types

Data types are defined and standardised on the lowest level. 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 respectively the domain of values. The task of a data type is to code the information that has to be repre-

sented by a data element. The codification transforms the information into a defined representation. This concept is characteristic for all information systems and is implemented in programming languages and database systems. For these two areas different sets of data types are available. They differ in number of types and degree of specialisation.

3.1.2 *Vocabulary*

On the basis of data types a standard can specify typed data elements. Hence the second level of the level model holds the data element definitions of a standard. The set of permitted data element builds in accordance with the language metaphor the vocabulary. It contains these words that are known to business partners in a business communication and therefore can be utilised in a communication process. The vocabulary level is in most cases the core component of every document-oriented e-business standard. Developing a business vocabulary is with regard to complexity and resources the most important domain-specific effort of a standardisation project.

3.1.3 *Documents*

The data elements and the implicitly or explicitly formulated data model form the basis for the definition of business documents that incorporate parts of the standardised vocabulary. The document is the central concept of each business communication: It is transferred from a sender to a receiver and contains a set of data. 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. Above its main function to combine related data to a logical unit, each document type possesses an intended purpose. This means that we can already draw a direct conclusion from the document type to the relationship between and role of the sender and receiver. For instance, the document type catalogue is only meaningful in a communication between the creator and the user of a catalogue. The purpose of this document type is providing product information of the sender to the receiver.

3.1.4 *Processes*

The sequences of exchanged documents and the underlying business logic of a communication 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. The standardisation object process can be defined as a transaction or a sequence of transactions between two business partners. Subject of each transaction is the exchange of messages according to agreed document types. With the help of

document sequences it is possible to support inter-organisational procurement processes to a full extent. A specific sequence might be as follows: catalogue, request for availability information, order, order acknowledgement, delivery notification, invoice, and payment.

3.1.5 *Framework*

The framework level covers those definitions that concern technical and thus domain-independent aspects of business communication. A framework defines a foundation for the communication and provides parallel to the lower levels additional services. All domain-dependent aspects are strictly left to the levels data types up to processes; they describe the business content and business logic. The higher aim of the framework level is to ensure a secure, dependable and structured exchange of business documents. One characteristic feature of framework services is a far-reaching independency from the content that has to be transferred and the logic that has to be kept. Rather supporting services are described, for example basic communication protocols (HTTP, SMTP and FTP), security issues (e.g., authentication, encryption) as well as the handling of messages (e.g., management of queues, notification and acknowledgment services).

A main concept to reach the independency from business content and logic is expressed by the envelope metaphor. It says that very similar to a postal service the content of a message is kept in a sealed envelope which is the object that has to be transported. The transport requires a meaningful inscription only, which at least clearly specifies the sender and receiver (as an address) of each message, or gives references to them. In this metaphor, the framework level describes a physical distribution or delivery system and defines the rules which apply to this system.

So called framework standards or B2B frameworks (e.g., RosettaNet) indicate 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 further lower levels or even build an own level model to describe document-oriented business communication. In this interpretation B2B frameworks are holistic models that are sufficient to implement interorganisational components 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 fulfil this sophisticated function of the meta-model level is

very small. A prominent standard is the ebXML initiative which is not only a framework standard but also shows many features of a meta-model, since it provides generic concepts and tools for modelling e-business communication in an abstract way (Hofreiter, Huemer & Klas 2002).

3.2 *Definition of criteria for the analysis of e-catalogue standards*

3.2.1 *Standardisation Organisation*

The organisational criteria address on the one hand the standardisation organisation as an entity or body that develops, publishes and maintains the standard and on the other hand the standardisation process. Here we can rely on a wide range of mostly domain-independent criteria and consistent sets of criteria common to analysing standardisation in general.

3.2.2 *Methodology used in the analysis of the standardisation process*

The methodology in e-business standardisation can be subdivided in the formal specification and the documentation.

The task of a formal specification is to describe the content of a standard partly or completely by a formal specification language. This specification fulfils two important aims. Firstly, the use of a formal language can result in more exact and clear description in comparison to non-formal languages. The understanding of the standard's content is facilitated and misinterpretations can be prevented or at least reduced. Secondly, a formal language that can be processed in information systems supports the implementation of a standard in software applications.

According to the level model different specification languages can be used, some of them are specific for one level (e.g., process models); other languages cover two or more levels (e.g., data models). Most of the languages for the higher levels are non- or semi-formal; they differ in language concepts, they are less standardised and can also be counted as part of the documentation (see below). Contrary to this kind of heterogeneity, a limited number of so called XML schema languages are available for specifying document types, the vocabulary and data types. An XML schema language consists of language elements that can be used to model the syntax of XML documents or parts of these documents. The languages provide a set of modelling concepts (for example: user-defined data types, inheritance, default values), which are used to a greater or lesser extent by catalogue standards. At present a common language for the definition of business documents is the Document Type Definition (DTD) (Bray et al. 2000). It was already published by the W3C at the beginning of 1998. The DTD specifies XML documents by means of an own language in a document-

oriented view. However, the modelling concepts in XML DTD are rudimentary compared to other XML schema languages. In the paper of Lee and Chu six XML schema languages (e.g., XDR and XSD) are compared concerning their modelling concepts (Lee & Chu 2000).

The task of documentation is to describe the content of a standard in such a manner that users of the standard can easily understand and finally implement the standard. The documentation can be differentiated between the levels of standardisation. Parts of the documentation are often semi-formal and formal specifications in addition to verbal descriptions. The close relationship to the formal specification lies in describing the semantics of the standards, since semantics is hardly represented by formal languages so far. 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 and effectively, because a common and equal understanding of the semantics and syntax is crucial to e-business communication as it is crucial to any communication.

The main instruments for documenting the content of e-business standards are those instruments that are also applied in modelling data structures, since the core of documents are data elements, too. In view of the high complexity of catalogue data which results in extensive data models, it is suitable to introduce conceptual data models that show the general structure and its relationships. To do so, languages and notations such as extended ER models, UML and graphical representations of XML are used in practice. But these languages are not capable to describe all syntactical and especially no semantic aspects of data elements; hence the most important instrument of an e-business standard is a structured data dictionary.

3.2.3 *Content of e-catalogue standards*

The content quality and capabilities of a catalogue standard can be assessed by asking whether the standard fulfils the requirements on catalogue data as it is exchange between companies. The main question is, which issues and concepts are described by the standard and if they are suitable for the needs of suppliers, intermediaries and buying organisations. According to the level model this question can be answered only by relating and modifying it to specific levels: First, is the level covered by the standard; second, what level-specific standardisation objects are covered; third, is the coverage right and satisfying. Checking these content-oriented issues is a time-consuming task requiring a broad domain knowledge, especially for the vocabulary that calls for a detailed analysis of the syntax and semantic of all data elements (e.g., Kelkar, Leukel & Schmitz 2002; Leukel, Schmitz & Dorloff 2002).

3.3 *Special issues of interest*

The eCAT Workshop will address many issues which are important to assessing, evaluating and extending the current state of e-cataloguing from an international and especially European point of view. Standards have to implement concepts that fulfil these top-level requirements, for instance by being able to fully represent multilingual catalogue content, by providing product and price models that are suitable for many branches of industry in Europe and their specific requirements in cross-border trade. Altogether the range of application and the benefits of e-cataloguing can be widened and tapped by powerful and harmonised standards only. Three key issues are product and price models as well as the fact that catalogue content is diverse and dependent on many parameters.

So far the main object of catalogue-based transaction systems are standardised products of a limited specification and complexity. Among these products are primarily indirect goods that are not an immediate input factor for production processes and can not be attributed to manufactured final goods. A common term is MRO goods (maintenance, repair and operations). 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. Although they require a relevant amount of resources for procurement, order and stock receipt management. The described restrictions cause a limited area of application for e-procurement systems so far. By extending the capabilities of catalogue 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. Thus the analysis of existing e-catalogue standards has to look at product models closely.

E-catalogues 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-catalogues. 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 and forms of taxes according to legal conditions in the ~~WE~~ must consider that catalogue data does not mean the data of only one specific catalogue. Rather catalogue data in its whole represents the quantity of data from which catalogues can be created. The creation of specific catalogues refers to an important characteristic: Each catalogue possesses a validity, which can be determined by a set of parameters. Be-

side the customer, this is the validity period, the currency of prices and the language of all language-dependent data. In accordance with this multi-dependency concept each catalogue can be described and identified by a tuple of parameters. It has to be considered that also multi-supplier catalogues, catalogue documents with product data of several suppliers, as well as multi-vendor catalogues that contain customized data of several buyers, should be transferred. The catalogue exchange must guarantee that the right catalogue is transmitted in the required format to the addressed recipient.

4 OUTLOOK

In this paper we presented the CEN/ISSS Workshop "Multilingual Catalogue strategies for eCommerce and eBusiness", in particular the approach of the working item "WI 1.3 Analysis of existing e-catalogues". The Workshop will finish its work till the end of 2003. The (preliminary) results are presented in form of (draft) workshop agreements, which will be presented on plenary meetings and via the official website of the Workshop. To get more information about the Workshop and the ongoing work, please visit <http://linux.infoterm.org/termnet-e/ecatsec.htm>. All official documents and a contact list can be found there.

REFERENCES

- Ariba. 2002. cXML 1.2.008. URL: <http://xml.cxml.org>.
- Baron, J.P., Shaw, M.J. & Bailey, A.D. 2000. Web-based E-catalog systems in B2B Procurement, in: Communications of the ACM, Vol. 43 (2000), No. 5, pp. 93-100.
- Bray, T. et al. 2000. Extensible Markup Language (XML) 1.0 (Second Edition). W3C Recommendation. URL: <http://www.w3.org/TR/REC-xml>.
- Bussler, C. 2001. B2B Protocol Standards and their Role in Semantic B2B Integration Engines, in: Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, Vol. 24 (2001), No. 1, pp. 3-11.
- CEN/ISSS. 2003. Information Society Standardization System, URL: <http://www.cenorm.be/iss>.
- CEN/ISSS Workshop eCAT. 2002. Business Plan. URL: <http://linux.infoterm.org/termnet-e/ecatsec.htm>.
- CEN/ISSS Workshop eCAT. 2003. Introduction, Team Members & Work Programme. URL: <http://linux.infoterm.org/termnet-e/ecatsec.htm>.
- CommerceOne. 2002. XML Common Business Library (xCBL), Version 3.5 URL: <http://www.xcbl.org>.
- Dogac, A. & Cingil, I. 2001. A Survey and Comparison of Business-to-Business E-Commerce Frameworks, in: SIGecom Exchange, Vol. 2 (2001), No. 2, 2001, pp. 16-27.
- Hofreiter, B., Huemer C. & Klas W. 2002. ebXML: Status, Research Issues and Obstacles, in: Proceedings of the 12th International Workshop on Research Issues on Data Engineering (RIDE02), San Jose, CA, USA, pp. 7-18.
- Kelkar, O., Leukel, J. & Schmitz, V. 2002. Price Modeling in Standards for Electronic Product Catalogs Based on XML, in: Proceedings of the 11th International World Wide Web

- Conference (WWW2002), Honolulu, HI, USA, pp. 366-375.
- Kelkar, O., Otto, B. & Schmitz, V. 2001. Specification open-TRANS®, Version V1.0. URL: <http://www.opentrans.org>.
- Lee, D. & Chu, W.W. 2000. Comparative Analysis of Six XML Schema Languages, in: ACM SIGMOD Record, Vol. 29 (2000), No. 3, pp. 76-87.
- Leukel, J., Schmitz, V. & Dorloff, F.-D. 2002. Coordination and Exchange of XML Catalog Data in B2B, in: Proceedings of the 5th International Conference on Electronic Commerce Research (ICECR-5), Montreal, Canada.
- Open Applications Group. 2002. OAGIS - Open Applications Group Integration Specification, Release 8.0. URL: <http://www.openapplications.org>.
- Schmid, B. & Lindemann, M. 1998. Elements of a Reference Model for Electronic Markets, in: Proceedings of the 31st Annual Hawaii International Conference on System Sciences (HICSS-31), Kohala Coast, HI, USA, 1998, pp. 193-201.
- Schmitz, V. Kelkar, O. & Pastoors, T. 2001. Specification BMEcat Version 1.2. URL: <http://www.bmecat.org>.
- Shim, S. S., Pendyala, V. S. & Sundaram, M. et al. 2000. Business-to-Business E-Commerce Frameworks, in: IEEE Computer, Vol. 33 (2000), No. 10, pp. 40-47.
- WebMethods. 2000. Creating a Standard, Fairfax, USA, 2000.
- Zhao, Y. & Sandahl, K. 2000. XML-Based Frameworks for Internet Commerce, in: Proceedings of the 2nd International Conference on Enterprise Information Systems (ICEIS 2000), Staffordshire, UK, pp. 511-516.
- Zwass, V. 1996. Electronic Commerce: Structures and Issues, in: International Journal of Electronic Commerce, Vol. 1 (1996), No. 1, pp. 3-23.