

## THE POSSIBILITIES FOR THE CREATION OF A LANGUAGE XML FOR THE FORMALIZATION OF THE ACCOUNTING RECORDS

Popescu Aurora

Doctoral school of the Academy of Economics Bucharest

---

### **Abstract:**

*During the nineties the main trend in the development of the applications was the supply of support and accessibility for the computers connected on the internet to a wide range of informational resources (data basis, applications). A witness in this are the numerous languages and technologies which permit an easy development of the applications for the processing of data bases with a simple web browser as, for example, the script languages ASP, PHP, JSP etc. Many changes took place in the last years regarding the informational needs or the equipments used by different users. So, today not only the computers are connected on the internet, but also a wide range of equipments as mobile phones and many home utility devices. As a result of these needs, it became an imperative necessity the conception of an universal language that be understood by all these diverse equipments. XML is the answer to this requirement, this language representing a new step in the development of the informational epoch. XML appeared as a consequence of the limits of the HTML (the language of the web pages), this last one being incapable to use data for other applications.*

---

Flexibility is one of the fundamental features of the present business environment. In this sense it is an imperative requirement that the economic information or other data be available on PC connected both to the company's network and to other locations. At the same time, the data supplied by these systems should be permanently actualized in concordance with the momentarily economic situation of the company. The world wide web is the service capable to offer the support for such informational systems . These are going to be accessed with the help of a traditional web browser. The present technologies allow the realization of some complex programs , that allow connection to data basis and that generate dynamically the web pages.

The development of XML began in 1996 and from February 1998 XML became a recommendation of W3C . But only apparently this technology is young. Before XML there was SGML, created at the beginning of 1980, an ISO standard from 1986 and used on a large scale in projects which involved big documentation. In 1990 HTML was born. So XML's designers took over the power of processing developed by SGML and the facility in use of HTML. But some evolutions are hardly distinguished

from revolutions...We should also mention that while SGML is mainly used for technical documentations and less for other kinds of data, in the case of XML takes place the opposite thing.

XML allows the definition of a new document format combining and reusing other formats. XML is a program capable to store data from various sources, including structured and semi-structured documents, relational data basis and containers of objects.

As a method of presentation for these data it is used a self-descriptive hierarchical representation which provides the possibility to present the data without losing their semantics and which can be easily interpreted by applications. So an application which receive an XML document can interpret it in different ways, can filter it by content or can restructure it depending on the needs of the respective application. This feature of self-description and the particularity of the structure make from XML an ideal language for representation and for data exchange. Because of the hierarchical modality of storage we can use the quality of the parents and children knots to inherit or to have its own attributes. When the documents are stored and represented in XML format, the content of it can be queried in order to be

extracted, analyzed and synthesized. But the efficiency of the extraction of these data depend on the way these are stored and on the application that perform this operation.

The documents entry-centered use the meta-language XML for the transport of data. They are designed to be understood at machine-level. It has no importance for the application or for the data base if these data are stored or not, for a period, in an XML file. Examples of entry-orientated documents are : orders of sale, flight hours, scientific data, stock-exchange quotations.

These entry-orientated documents have a regular structure, precisely defined data (the smallest independent unit of information is an attribute or a token – PCDATA) and could have a mixed content. The order in which the data appear has a significance only at the validation of the document.

The types of data from the documents which use an entry-centered representation can come both from the interior of the data base (in this case it is intended an exposition of the data base in a XML format) or from the exterior of the data base (in this case it is intended a storage of the information in a data base).

XML or eXtended Markup Language is in fact the brother of HTML (Hyper Text Markup Language). A number of similitude exists between the two languages – they have a commune origin ,SGML – Standard Generalized Markup Language, a general language, which offer to the users the possibility to organize the data in an electronic format , with the help of some mechanisms similar to those from a classical grammar.

In fact HTML uses a set of tags by which is defined the way how the information will be posted on the web browser used. The set of tags of a HTML language is finite and is generated with the help of SGML technology. In these conditions there is no possibility to chose a particular set of tags, an important limitation in some cases. So it is inadequate the use of HTML language to transport applications that work with data basis.

Unlike HTML, XML offer the user the possibility to define his particular set of tags in order to format a document. This flexibility is a necessary element in the definition of the electronically specialized documents.

SGML is a complex language, difficult to be used on a large scale, and XML represents a simplified solution. In XML a particular set of tags is possible and the significance of the information is more evident and can be easy taken up by other applications. For example, if from a catalogue it is wanted to be extracted the data related to products and prices , the method of work will be different in the two languages.

In the case of the use of HTML, the modality of the extraction of the data is artificial, without clear marks regarding the information regarding the price or the product name. In the case we use XML to define a catalogue of products, we will have fixed marks representing ‘price’ and ‘product’ tags, in order to extract the data . As a conclusion, we can say that HTML defines the way in which the information is displayed and points out the form and the appearance of information, while XML treats in particular the definition and storage of information.

HTML is used to create web pages and XML is going to be the technology that facilitate the exchange of information between data basis.

Like HTML files , XML uses the text format to store the information , in this way it is possible the transport of data on any platform or device. XML provides support for specification and validation of the complex documents with the help of DTD (Document Type Definition) schema format. The principal elements of a XML technology are the following:

- The XML file which includes the set of tags and the information transmitted with the help of it
- The declaration of the validity rules and of the types of data contented in the document with the help of DTD language (Data Type Declaration). With DTD will be defined the set of tags used to mark the document, as well as the validity rules for the data included in the document, if necessary.

DTD has the role to validate the content and the structure of the XML file. If for XML it is optional, in the case of SGML documents that section is compulsory. In present there is a number of DTD dedicated to specialized fields: HL 7 specific to the healthy field or MathML specific to the documents which include mathematical operations. With the help of these pre-created DTD the information can be organized more efficient in particular domains.

XSL style sheets are used to display the data from a XML file to a browser. The style sheets have a similar role to formatting declarations in the CSS files (Cascading Style Sheet).

In this way it can be established a tabular structure to display data, colors, fonts as well as other elements for formatting the data included in a XML file.

Some specifications regarding the XML file :

- The first line of the file specifies the version of the program and the codification of the characters used in the file

- At the beginning of the XML file , one can still set the definition and formatting rules of the data in DTD language, but this is an optional element. There are two possibilities to introduce the DTD declarations :directly in the XML file, or to insert them in a distinct file with this extension.

- There can be introduced, in the beginning of the file too, references towards the style sheet XLS, establishing that how to format the document.

- The proper beginning of a XML file is marked by the tag 'root', which has a similar function to the tag <HTML> specific to the HTML language(for example the tag root is represented by <Catalogue\_prices>).

- The XML file, similar to HTML file, is well-balanced as regards the tags used. In this sense, for every start-up tag it will be a similar one which will end the respective element (for example, in the case of a 'root' tag the XML will be ended by the declaration <Catalogue\_prices>).

What relation is between XML and the data basis? There is a simple answer:

XML is the easiest and most efficient way by which data can be transported on internet.

XML is a kind of envelope, a container where the information from the data base is encapsulated and distributed on the internet and the data stored in these files will be available to any device on any platform. In the past, for the transportation of data between different platforms there were used text files in which the data were separated by comma or by other special character. Beginning with XML this past way of transfer became old fashioned.

Compared to the other ways of data transfer XML has the advantage of simplicity, an efficient organization of information and the possibility to display the information on any platform or device.

Because of the penetration of the internet in the most diverse fields, we could conclude that there are no more independent data basis, but only interconnected data basis. For an example we propose the following scenario: a player on the stock exchange wants to obtain the daily stock exchange quotations displayed on the stock exchange site. The memorizing of data regarding the stock exchange quotations in a data base on a personal computer (for example in Access) can take place as follows:

- the user files a request to the data base web server asking for the desired stock quotations ;

- the web server transfer the request to the data base server in the form of a SQL inquiry and this last server sends back the specified set of entries;

- the data sent by the server are packed and delivered in XML format to the client via the web server;

- the data are taken up on the user's PC that sent up the request and transferred on the data base which contains statistical data related to the stock exchange quotations.

This scenario represents a simple example of interdependence and communication related to the internet and XML technology. The points 2 and 3 in this scenario tend to unify. In this sense the most data base administration systems have incorporated facility of export of data in XML

format and is no more necessary to use collateral applications.

Access XP provides the possibility to easily export data in XML format, including in this sense a set of facilities. To perform that operation one has to select a table or an inquiry in the data base, and then chooses from the File menu the option Export, selecting in the dialog box an XML file format. On the displayed window there is the possibility to establish a series of options regarding the export of the data in an XML file format as follows:

- the characteristics set used to create the file (UTF 8 represents the 8 bit characteristic set Unicode);
- establish if it will be created the scheme to describe the data included in the XML file, a function similar to the one of the DTD language, aiming to verify the data integrity in the file XML;
- there is also the possibility to establish if it will be created a style spreadsheet destined to establish the way the

information will be displayed from the XML file on the web browser; Access has a similar function by which one can import data sent in a XML format in an Access data base.

### **Conclusions**

XML is a high important technology which permits an increased interchangeability of information between different systems.

In a world in which the internet has an overwhelmed importance and where there is a multitude of devices connected to the internet, XML represents a technology with the help of it the data can be transferred on the most various platforms.

XML is independent of a hardware and software platform, being a meta-language easy to be used in the definition of a new types of documents by a formal specification.

Out of these reasons, XML is the central element for the development of many applications or for different data administration systems.

### **References:**

- [1] Buraga S., *Tehnologii Web*, Polirom, 2006;
- [2] Bourret, *XML and Database*, Campus XLM-2005
- [3] Obasanjo D, *An Exploration of XML In Database Management Systems*, 2001
- [4] Bansal V, Alam A, *Study and Comparison of Techniques to Efficiently Store and Retrieve XML Data*
- [5] Bray T, Paoli J, Sperberg-McQueen CM, and Maler I, *Extensible Markup Language (XML) 1.0 (Second Edition)*, October 2000, 2005
- [6] [www.25hoursaday.com/StoringAndQueryingXML.html](http://www.25hoursaday.com/StoringAndQueryingXML.html)
- [7] [www.w3.org](http://www.w3.org)
- [8] [www.xml.com](http://www.xml.com)