

# XML PUBLISHING SOLUTIONS FOR A COMPANY

Claudiu Chiru, Liana Elefterie, Elena Iatan

Spiru Haret University, Unirii str. 32-34, Constanta

claudiu\_chiru@yahoo.com, office@audit.ro, iatan\_e@yahoo.com

**Abstract:** *The eXtensible Markup Language, or XML is a method to add structure and context to unstructured information. This paper presents some of the latest issues in using XML as a standard for information publishing. We present XBRL (eXtended Business Reporting Language) and possible applications in elaborating financial documents. We also present DITA (Darwin Information Typing Architecture) as a new XML based standard developed by IBM. DITA is used for technical information structuring and publishing. In the final part we present a DITA implementation of a procedures manual for a hypermarket.*

**Keywords:** XBRL, XML, DITA, HTML, taxonomy, instance document

## 1. Introduction

XML (EXtensible Markup Language) is an open standard that describes data from the Web. XML defines data elements on a Web page and business-to-business documents. XML uses tag structure like HTML but, whereas HTML defines how elements are displayed, XML defines what those elements contain. HTML uses predefined tags and XML allows tags to be defined by the developer of the page. Thus, data items can be described as database records. XML supports business-to-business transactions and has become a standard for electronic data interchange and Web services.

## 2. XML Based Languages

### 2.1. XBRL (eXtended Business Reporting Language)

XBRL (Extended Business Reporting Language) is a communication language that can be used to send/receive business or financial information in a standard form. It is based on XML technologies and is fast becoming a standard means of communicating information between businesses and on the Internet. XBRL can be used to efficiently :

- Elaborate reports
- Analyze financial data
- Share financial data
- Verify data

XBRL is a global effort for building a digital reporting language. This standard makes the global communication easier:

- XBRL creates a common vocabulary for financial information description
- Every country has a national chapter but the structure of the documents is the same.

Many European countries have developed taxonomies for XBRL implementation:

- Belgium-Belgium National Bank
- Denmark-Danish Commerce and Companies Agency (DCCA)
- European Union – European Committee of Central Balance Sheet Data Offices (ECCBSO)
- Germany – DeutscheBank, Bundesbank

### 2.1.1. XBRL structure

#### The taxonomy

A XBRL taxonomy [1] is a dictionary of terms built in a hierarchical structure. The taxonomy consists of two important parts: the schema file (or more schemas) and linkbases.

The schema file contains the elements and their characteristics such as names, ids and other attributes. A schema file has an .xsd extension and it is an XML schema file.

The root element of all schemas is called <schema>. The starting root element is <schema> and the ending root element is </schema>. XML documents and XBRL documents as well use namespaces. A namespace is used to distinguish between elements of different schemas. A namespace is like an Internet address (“http://xbrl.iasb.org/int/fr/ifrs-gp/”) but it’s not. The reason for using such names is their uniqueness and so unique elements in the schemas can be identified.

Elements are in fact business concepts like *Assets*. In the following example we present a concept named “Depreciable”. The most important parts provided in this example, from a business perspective, are name, type, balance and *PeriodType*.

```
<element id="ci_Depreciable" name="Depreciable"
type="xbrli:monetaryItemType" substitutionGroup="xbrli:item"
xbrli:periodType="instant" xbrli:balance="debit" />
```

*Name* is a unique element assigned to the concept, *type* is the type of the concept (in our case is a monetary type), *PeriodType* makes a distinction between flows and resources.

Payments, revenues or profit have the *PeriodType duration*.

XBRL has also to express the balance nature of the concept: *debit* or *credit*. The schema document reflects the nature of the concept by using the *balance* attribute: balance=”debit”.

*Linkbases* [1] are the components of the taxonomy that provide relationship between elements and link them with external resources.

The *linkbase files* are named using the name of the taxonomy followed by an underscore and then the type of the linkbase file :”*basicCalculation\_labels.xml*”. In this case the taxonomy file name is *basicCalculation.xsd*.

The taxonomy *linkbase files* [1] are:

- labels
- references
- presentations
- calculations
- definitions

The *labels file* [1] has the detailed description for all the elements in the taxonomy. There are at least four entries for each element if a single language is supported:

- locator entry
- label entry
- arc entry from the taxonomy to the label
- arc entry from the label to the taxonomy

The *reference file* [1] points to source documents ( documents issued by the authorities) indicating the relevant paragraphs and clauses.

The *presentation file* [1] is responsible for displaying the elements on the order in which it should appear in the financial statement.

The *calculation file* [1] improves the quality of the XBRL report. It has definitions of validation rules. These validation rules apply to all instance documents that refer to a specific taxonomy.

The *definition file* [1] creates all the parent-child relationships.

### Instance documents

XBRL instance documents are electronic reports created according to the XBRL taxonomy. An instance document contains concepts with a given value and a context in which those concepts are used.

An instance document defines contexts such as *period* or *measure*.

For example:

```
<context id="I-2002">
  - <entity>
    <identifier
      scheme="http://www.SampleCompany.com">SAMP
    </identifier>
  </entity>
  - <period>
    <instant>2002-12-31</instant>
  </period>
</context>
```

and

```
<context id="I-2003">
  - <entity>
    <identifier
      scheme="http://www.SampleCompany.com">SAMP</identi
    fier>
  </entity>
  - <period>
    <instant>2003-12-31</instant>
  </period>
</context>
```

define two instant moments.

The unit tag defines the type of the monetary item. In our case EURO.

```
<unit id="U-Monetary">
  <measure>iso4217:EUR</measure>
</unit>
```

Element's definition in the taxonomy:

```
<element id="ci_ComputerEquipment" name="ComputerEquipment"
  type="xbrli:monetaryItemType" substitutionGroup="xbrli:item"
  xbrli:periodType="instant" nillable="true" />
```

The following tags define the effective use and the contexts of the concept *ComputerEquipment* in the instance file.

```
<ci:ComputerEquipment contextRef="I-2002" unitRef="U-
Monetary">5313000</ci:ComputerEquipment>
<ci:ComputerEquipment contextRef="I-2003" unitRef="U-
Monetary">4169000</ci:ComputerEquipment>
```

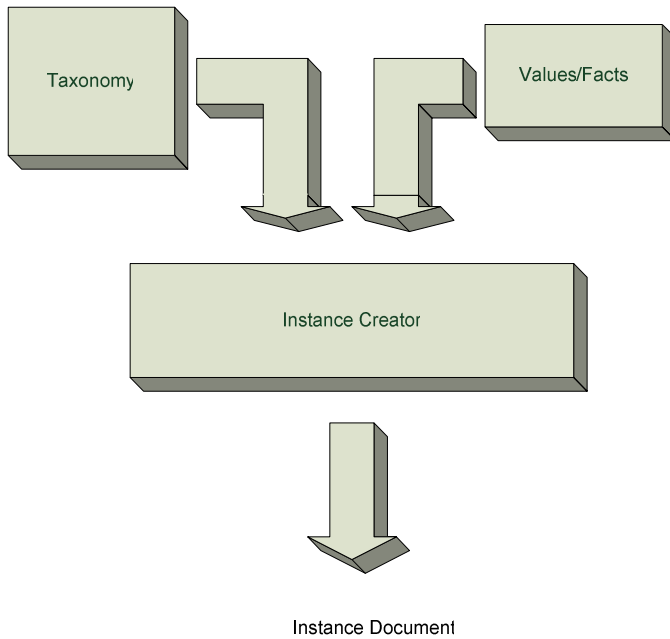


Fig. 1. The process of creating an instance document

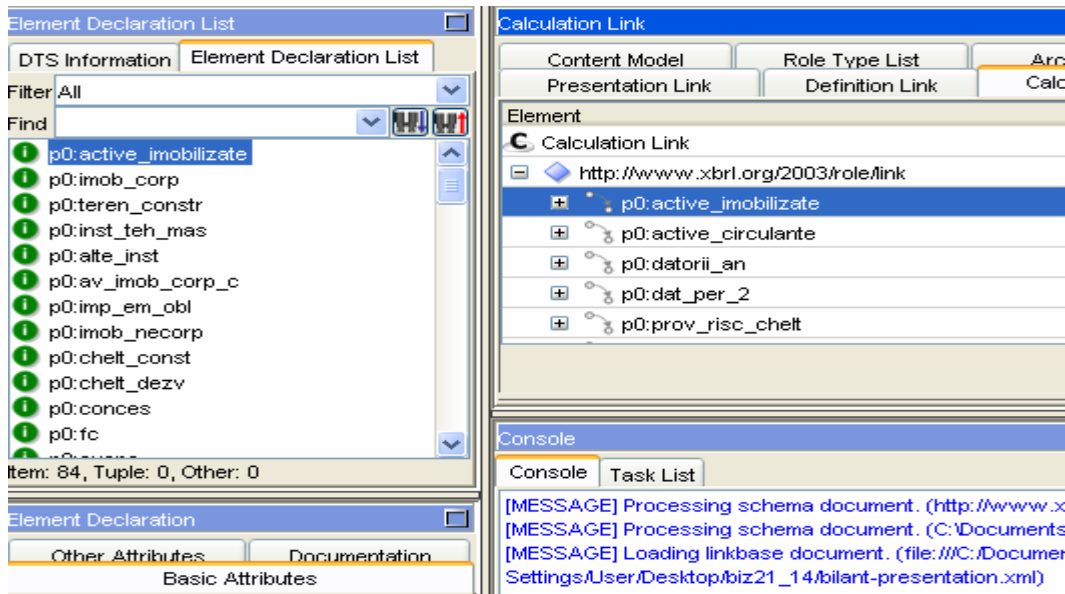


Fig 2. Fujitsu Taxonomy Creator Screenshot

The instance document may have the same layout as the original document. Reading an instance document is a process that extracts the data out of the document.

### 2.1.2. Sending XBRL Balance Sheet Documents to Local Financial Administration

The process of creating XBRL balance sheet documents needs a taxonomy and instance documents for every balance sheet. The taxonomy will contain all the concepts needed together with their attributes and relationships. The instance document was created using an instance

creator named Fujitsu Instance Creator. The instance document contains values from the original document.

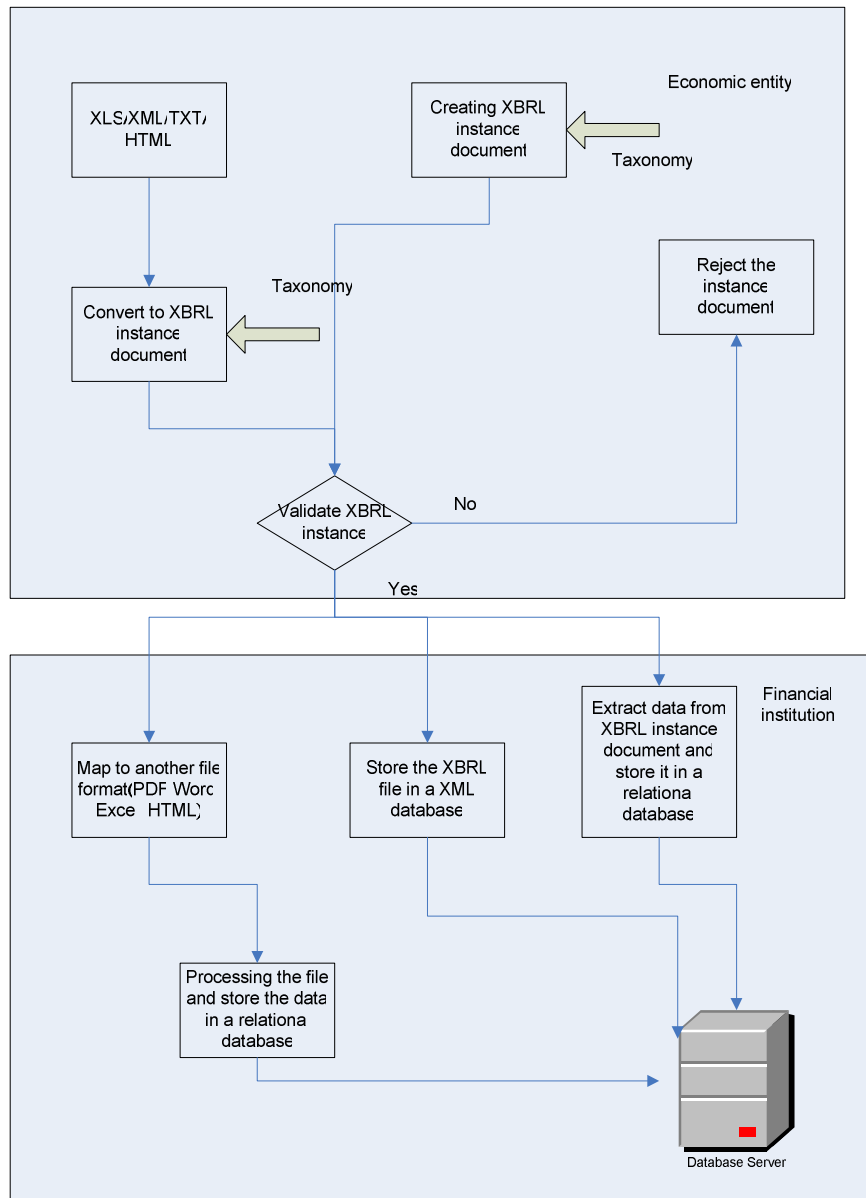


Fig. 3. XBRL Framework

The above XBRL framework can be used to send financial documents/reports to a financial institution. The XBRL documents can be created from Excel files. In this case we have to use a conversion software. A good solution might be Rivet’s software Dragon Tag.

The XBRL instance document can be created, as you had already seen, by using instance creator software or some Adobe based solutions like the one provided by Corefilling company.

Storage and processing data collected from XBRL documents is made by using several ways:

- storage in files
- storage in relational databases

- storage in native XML databases

There are several software vendors for the above storage solutions: Tamino XML server, Semansys Tamino XBRL Suite, Xindice..

## **2.2. Dita (Darwin Information Typing Architecture)**

The Darwin Information Typing Architecture (DITA) [2] is an XML-based architecture for authoring, producing, and delivering technical information. DITA includes topics like:Task, Concept, Reference. A Task topic describes how to accomplish a task. It displays a list of steps that the users must follow to produce a specified result. Concept contains definitions, rules and guidelines. A Reference topic describes command syntax, programming instructions, reference material. DITA is an emerging standard that provides an outstanding platform to promote information interchange, in an orderly fashion, within an enterprise or with other organizations. DITA provides several good methodologies for managing an “information supply chain” where supplier and consumer can easily interchange topics. DITA provides an excellent infrastructure for managing information at a topic level, getting away from document hierarchies that in the past have added complexities to interchange. The DITA architecture, a related DTD and a W3C-Schema was originally developed by IBM. DITA is now an OASIS standard.

### **Using DITA for publishing procedures manuals**

Procedures manuals are a collection of task to accomplish by the employees of a company. Structuring a procedure manual has become an easy task using DITA. Taking advantage of the DITA specification, the manuals can be structured using DITA topics: Task, Concept or Reference. On the other hand, the manuals may be print for different audience.

In this section we are trying to design a DITA structure for the procedures manual of a hypermarket.

### **The DITA Structure of a Procedure Manual for Goods Receiving in a Romanian Supermarket**

#### **Creating Content**

We used the XMLMind editor for creating the DITA procedures manual. XMLMind is a free XML editor which has a free DITA plug-in that provides user-friendly templates for standard DITA topic types and for DITA map files.

First, we created the concepts, then the tasks and the references. Each concept or task or reference is an xml file. The topics that we have created are:

#### **Topics**

##### **Concepts**

- Receptie marfa
- Comanda
- NIR
- ...

##### **Task**

- Lansare comanda
- Verificare existența comandă
- Verificarea mărfii la recepție
- ...

##### **Reference**

- Starea comenzii
- Inregistrarea receptiei

- Document privind returul/refuzul de marfa
- ...

### Audience

- Director magazin
- Sef departament
- Category manager
- Sef raion
- Operator calculator
- Lucrator comercial

After creating these topics we have created the ditamap file. Ditamap is a map file that puts together all the files into a single document. The map file describes the sequence and hierarchy of the topics and sometimes the relationship between them.

There can be created as many map files as one need. For example, one map file for a PDF output and one map file for the HTML output.



Fig. 4. Creating a DITA Concept with XMLMind

An important step in creating DITA documents is conditional processing or filtering logic [3]. To exclude topics that have certain attributes from the output we have created a ditaval file using *prop* elements with attributes *att*, *val*, *action*. For example:

```
<prop action="exclude" att="audience" val="director departament">
```

will exclude all topics that are related with the “director departament” duties.

The final step is creating output. The Dita Open Toolkit [3] can build output in HTML format or PDF format. The output can be customized using CSS styles or by modifying XSL templates. We could create output for every category of audience: *director departament*, *sef receptie marfa*,

...

### **3. Conclusion**

XML is fast becoming “the language” for information interchange and describing relationships. There is XML for biology, XML for chemistry, XML for rights management, XML for finance, XML for business rules, XML for publishing, etc. In this paper we tried to present some of the latest applications of XML. We have presented XBRL as the language for business reporting and DITA a new standard for writing documentation. We also have presented some adequate XML solutions for companies. In the near future, together with the Marklogic environment or TextML server we think that companies will implement XML solutions for document management as a standard.

### **4. References**

- [1] Charles Hoffman (2006), Financial Reporting Using XBRL, Published by Lulu
- [2] Don Ray (2005), Introduction to the Darwin Information Typing Architecture, IBM corporation
- [3] Louise Kasemeier(2007), Dita for Solo Writers, [www.lone-dita.com](http://www.lone-dita.com)