Analysis of XBRL documents containing accounting information of listed firms using Semantic Web Technologies

Sheila Méndez, Jose Labra, Javier de Andrés, Patricia Ordóñez

Abstract This paper presents an approach to analyze XBRL documents using semantic web technologies. XBRL is an XML-based standard widely used in the exchange and representation of accounting and business information between different organizations. The proposed system takes an XBRL document and converts its information into RDF files. The obtained files are merged with an OWL ontology describing financial information domain. The system enables the formulation of SPARQL queries over the generated data which facilitate the analysis of the financial information. Currently the system has been applied to the XBRL reports generated by the Spanish Securities Commission on the basis of the accounting information submitted by listed societies with shares admitted to quotation.

1 Introduction

XBRL (eXtensible Business Reporting Language) is an XML based standard developed by a not-for-profit international consortium of approximately 450 organizations including regulators, government agencies and software vendors. It has successfully been applied in the exchange and representation of accounting and business information between different organizations. XBRL employs XML Schema and XLink technologies to describe different taxonomies for specific domains so that each XBRL document is an instance of a specific XBRL taxonomy.

Sheila Méndez
University of Oviedo. C/ Valdés Salas s/n 33007 Oviedo, sheilamendeznunez@gmail.com

Jose Labra, Javier De Andrés, Patricia Ordóñez
University of Oviedo. C/ Valdés Salas s/n 33007 Oviedo {labra, jdandres, patriop}@uniovi.es
In Spain, XBRL adoption was led by the Bank of Spain (Banco de España) and by the Spanish Securities Commission in 2005 (www.cnmv.es). In the case of the Spanish Securities Commission (Comisión Nacional del Mercado de Valores – CNMV), it publishes periodically via web, in XBRL format, accounting information regarding the listed firms. In our paper we will concentrate in the IPP (Información Pública Periódica - Periodic Public Information) taxonomy, which contains the main items of the quarterly and semester accounting reports.

On the other hand, the semantic web can be defined as a long-term vision which pursues the development of technologies that facilitate the automatic manipulation of data published on the Web. Led by the World Wide Web consortium (www.w3c.org), a number of semantic web technologies have appeared, like RDF to describe resources using a graph model, OWL to define ontologies based on description logics, and SPARQL, to define queries over data RDF graphs. One important aspect of these technologies is that they can be neatly combined using several tools and even allowing the system to infer new knowledge using description logics capabilities.

The remainder of the paper is structured as follows: the next section briefly describes the accounting standardization using XBRL. Section 3 explains the supervising function of the CNMV and why XBRL is important for the fulfillment of this function. Section 4 details the architecture of the proposed system for the analysis of the accounting information in XBRL format which can be downloaded from the CNMV website. Finally, sections 5 and 6 outline related works and the main conclusions of our research.

2 The XBRL standard

XBRL is a language for the electronic communication of business data. It is especially suitable for the standardization of accounting reports. It provides major benefits in the preparation, analysis and communication of financial information. It offers cost savings, greater efficiency and improved accuracy and reliability to all those involved in supplying or using accounting information.

XBRL stands for eXtensible Business Reporting Language. It is one of a family of "XML" languages which is becoming a standard means of communicating information between businesses and on the internet. XBRL is being developed by XBRL International, which is an international non-profit consortium (http://www.xbrl.org) of approximately 450 major companies, organisations and government agencies (march 2008). It is an open standard.

XBRL International is comprised of jurisdictions which represent countries, regions or international bodies and which focus on the progress of XBRL in their area. By now (march 2008) the following countries have established a XBRL jurisdiction: Australia, Belgium, Canada, Denmark, France, Germany, Ireland, Japan, Korea, Netherlands, New Zealand, Poland, South Africa, Spain, Sweden,