Information Systems 57 (2016) 20–37

Contents lists available at ScienceDirect

Information Systems

journal homepage: www.elsevier.com/locate/infosys

XBRL formula specification in the multidimensional data model

Ignacio Santos [a,*], Elena Castro [b], Manuel Velasco [b]

[a] Bank of Spain, Department of Information Systems, c/ Alcalá 522, 28027 Madrid, Spain
[b] Carlos III University of Madrid, Computer Science Department, Avda. de la Universidad No. 30, Madrid, 28911 Leganés, Spain

ARTICLE INFO

Article history:
Received 22 August 2012
Received in revised form
7 August 2015
Accepted 2 November 2015
Available online 17 November 2015

Keywords:
Extensible Business Reporting Language (XBRL)
Extensible Markup Language (XML)
Taxonomy
XBRL dimensional taxonomy (XDT)
Conceptual data model
Multidimensional data model (MDM)
XBRL data model (XBRLDM)
Model Driven Architecture (MDA)

ABSTRACT

XBRL is a specification used to exchange financial/economic information. It is actively used by many international institutions and agencies. In the USA, Canada, Europe, China, etc. all financial entities and companies quoted on the stock market have to report compulsorily to the supervisory and regulatory authority using the XBRL specification. XBRL consists of a set of taxonomies defining different accounting regulations for a specific statement and the statement itself. Reports are generated from various sources and are validated at origin. XBRL displays business information which is multidimensional and whose logical destination for storage is a data warehouse. The proposal presented here focuses on the automation of the mapping between XBRL and the multidimensional data model (MDM) and includes a formalization of the validation rules in the MDM. The approach is designed in accordance with the Model Driven Architecture (MDA) paradigm which consists of a new way to validate XBRL reports through an RDBMS, and offers a proof-of-concept. Additionally, the study aims to provide more clarity about XBRL, a highly complex language made by and for expert users, and to improve interoperability between applications. The proposal also analyses certain semantic questions associated with the XBRL formula specification and its performance.

1. Introduction

XBRL is a specification used to exchange financial/economic information. It is actively used by the Board of Governors of the Federal Reserve System (FED), the Securities and Exchange Commission (SEC), the ShenZhen Stock Exchange (SZSE), the Shanghai Stock Exchange (SSE) [47,48], the European Central Bank (ECB), the European Banking Authority (EBA), the European Insurance and Occupational Pensions Authority (EIOPA), the Deutsche Börse, the Deutsche Bundesbank, Companies House and HM Revenue & Customs (UK) and the Australian Prudential Regulation Authority (APRA), among many other institutions and agencies. The use of this specification is increasingly important. In the USA, Canada, Europe, China, etc. all financial entities and companies quoted on the stock market have to report compulsorily to the supervisory and regulatory authority using the XBRL specification. Financial statements are governed by strict requirements, such as the International Financial Reporting Standard (IFRS) or Generally Accepted Accounting Principles (GAAP). XBRL is used instead of XML, because the reports have specific semantics managed by IFRS and/or the GAAP, and these reports can be very big. In XBRL, an XML-based standard, the reports (XBRL or XML document instances) are generated from various sources and validated at origin, ensuring the correctness of its syntax (XBRL formula specification). Thus, the financial statements of a credit institution, for example, are specific statements defined by one or more taxonomies, including

* Corresponding author.
E-mail addresses: ignacio.santos@bde.es (I. Santos), ecastro@inf.uc3m.es (E. Castro), velasco@ia.uc3m.es (M. Velasco).

http://dx.doi.org/10.1016/j.is.2015.11.001
0306-4379/

their structures and semantics. As accounting directives are subject to modification over time, versioning and location (e.g., of a country, state or region) problems may often arise. To offer just one example, while accounting requirements on capital flows were, until recently, based on Basel II [18], as a result of the current financial crisis, these regulations have been amended in Basel III [19,33]. This large quantity of requirements and data to be managed has entailed extremely high expenditures on Information technology (IT) resources by many companies, as well as by economic and financial institutions.

UML/MDA [46] is a powerful tool that has helped in different areas of Information Technology (IT) to model structured and robust systems. However, the techniques of verification and validation of the software is not supported in the standard MDA [51,52]. For this reason, the main contribution of this paper related to Information Systems is focused to the automatic transformation of the models, formalising validation rules in each phase, and with the target for validating a high number of rules, which represent hundreds of concepts, dimensions, attributes of dimension, cubes, etc. In addition, this paper shows a new demonstration of software validation that can be applicable to any field in IT, where the semantic interoperability is required.

Following the bankruptcy of Enron Corporation in December of 2001, stock market regulators began to demand the reporting of much more business information and reduced the amount of time in which this reporting was to occur. The information provided had to be free of errors, delivered securely and easy to analyze. In addition, companies and financial institutions have to exchange financial and business information. While this swap of information may be done in paper, FAX or e-mail, all such reports are nevertheless becoming more complex, requiring more in-depth analysis and processing. Although for years there have been several solutions available for transmitting information, none are adequate from the viewpoint of business economics.

Given these requirements, the information included in such statements must be semantically well-built, free of syntax errors and must be prepared for immediate or subsequent processing and storage in a data warehouse (DW) or database management system (DBMS). Adding further complexity to the task, the analysis of the information provided for particular financial or business purposes sometimes requires contrast with additional information or reports [29,30,32,34,35].

In April of 1998, the accountant and auditor, Charles Hoffman, proposed the automation of the exchange of financial information, developing a prototype financial statement and auditing program recognized as the basis for XBRL [17]. As XBRL represents business information which is multidimensional, the logical location for its storage is a DW [21,28].

At the level of European Union (EU) institutions and projects, two specific examples of XBRL use can be found, namely, in the Guidelines on Financial Reporting (FINREP) and Common Reporting (COREP) [37] from the European Banking Authority (EBA) [36], as well as in Solvency II from the European Insurance and Occupational Pensions Authority (EIOPA) [38]. In FINREP, consolidated and sub-consolidated financial reporting for supervisory purposes based on IAS (International Accounting Standards)/IFRS is addressed. COREP, by contrast, focuses on the consolidated, sub-consolidated and solo reporting of capital requirements and capital and reserves based on EU directives. Adoption of both FINREP and COREP taxonomies in EU countries by 2013 has been made mandatory, although with limitations. With Solvency II, the harmonized rules for the EU insurance sector to be implemented in 2014, between 4000 and the 5000 of

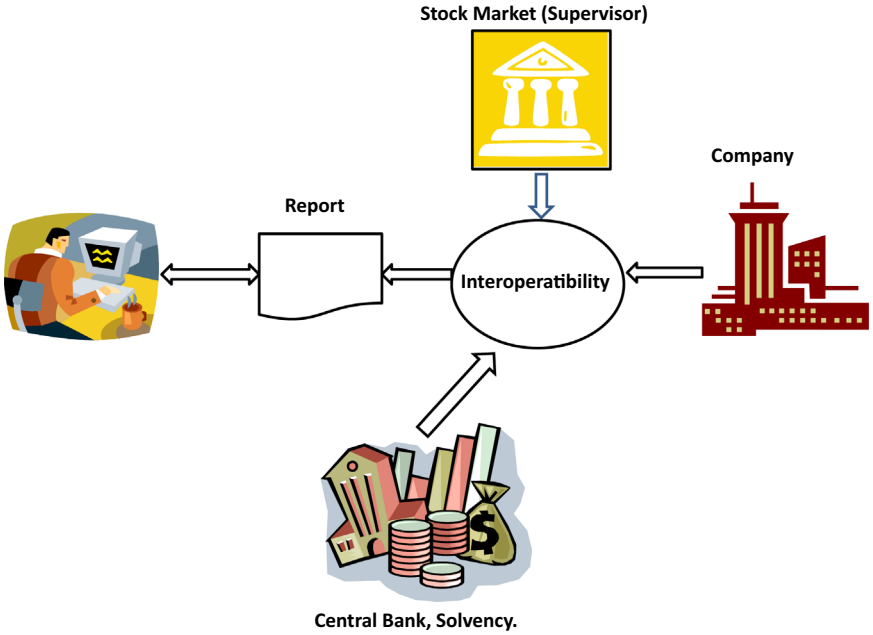

Fig. 1. Financial interoperability.