Interoperability of XBRL Financial Statements in the U.S.

Hongwei Zhu, Old Dominion University, USA
Harris Wu, Old Dominion University, USA

ABSTRACT

In the wake of the global financial crisis, a pressing need exists for improving investor friendliness, especially the transparency and interoperability of the financial statements of public companies. eXtensible Business Reporting Language (XBRL) and XBRL taxonomies can accomplish this objective. In the U.S., the Securities and Exchange Commission (SEC) has mandated that all public companies must file their financial statements using XBRL and the U.S. Generally Accepted Accounting Principles (GAAP) taxonomy according to a phased-in schedule. Are the XBRL-based financial statements interoperable? This question is addressed by analyzing all of the annual XBRL financial statements filed to the SEC as of February 26, 2010. On average, 63% of data elements are not comparable between a pair of statements. The incomparability is partly caused by issues related to the GAAP taxonomy and misuse of the taxonomy by companies. The results have practical implications that will help improve the quality of financial data.

Keywords: Data Quality, GAAP Taxonomy, Interoperability, Long Tail, XBRL

INTRODUCTION

Publicly traded companies must disclose their financials quarterly and annually to regulators and investors. In the United States, companies file their financial statements to the Securities and Exchange Commission (SEC). In recent years, financial statements are usually filed in digital format such as PDF or HTML files. Although the digital filings are much better than the traditional paper filings, they are not easy to use due to substantial challenges in processing unstructured data. For example, it is often desirable to compare several companies on a set of metrics such as assets and net income. No tools are available to accurately and reliably automate such comparisons. In other words, the financial statements are not interoperable and cannot be conveniently analyzed using automated tools. To alleviate this problem, the filings are often semi-manually processed to have the data elements extracted and stored into structured databases. This process is inefficient and error-prone.

eXtensible Business Reporting Language (XBRL) has the potential to address the interoperability problem (Hoffman & Watson, 2009; XBRL International, 2006). Additionally, XBRL offers other benefits such as improving efficiency of the entire financial data supply chain and enabling continuous auditing (Du &
Financial statements in XBRL format can be easily parsed using software tools because of the syntactic uniformity of XBRL. Furthermore, XBRL can be used to define a taxonomy that specifies a set of data elements commonly used in financial statements and the relationship among these elements. If all companies use common elements specified in a standard taxonomy, their financial statements are semantically interoperable, which makes it easy to compare different companies’ financial data. Otherwise, semantically-equivalent data elements reported using different element names must be matched and semantically-different elements reported using the same element name by different companies must be distinguished. This is the well-known and extensively-researched semantic heterogeneity problem for which there is still no reliable, fully automated solution (Bernstein & Haas, 2008; Bernstein, Melnik, Quix, & Petropoulos, 2004; Rahm & Bernstein, 2001; Rahm, Do, & Maßmann, 2004). In this case, a significant amount of manual work is needed to enable semantic interoperability of financial statements from different companies.

In the U.S., the SEC has adopted the US GAAP Taxonomy as the standard base taxonomy for financial reporting. In addition to the data elements specified in the taxonomy, companies are allowed to introduce their own custom elements to produce their financial statements. The extensibility gives companies more flexibility, which can lead to financial statements that more accurately reflect company status. For example, a company may possess various types of assets that are not captured in the standard taxonomy. If only standard elements are allowed, all these types of assets will have to be aggregated as “other assets”, losing detailed information useful to investors. To avoid this problem, companies are allowed to introduce custom elements by extending the standard taxonomy. With custom elements, all the special types of assets can be properly represented in financial statements. However, it has been recognized that extensibility can have adverse effects on interoperability of financial statements from different companies (Debreceny et al., 2005).

Starting June 15, 2009, all companies with a public common equity float greater than $5B must submit their financial statements in XBRL to the SEC. XBRL will be used by all public companies in the U.S. by October 31, 2014. Has the use of XBRL and GAAP taxonomy made financial statement semantically interoperable? Motivated by this question, we analyzed all official XBRL filings submitted to the SEC as of February 26, 2010. We identify the patterns of how companies use or misuse the standard taxonomy as well as the issues related to the taxonomy itself.

RESEARCH METHOD

Metric Definition

In this paper, we focus on comparability aspect of semantic interoperability of financial statements. A set of financial statements are interoperable if they use the same set of data elements defined in a standard taxonomy. Interoperability measures the extent to which a set of financial statements have overlapping data elements defined in a standard taxonomy. We do not attempt to match custom elements with standard elements, nor do we match custom elements introduced by different companies. Thus this definition allows us to measure interoperability directly without relying on unreliable semantic matching techniques.

We introduce a number of metrics to measure interoperability of financial statements. Before defining the metrics, we need to introduce several notations. Let \( D = \{ d_i \mid i = 1 \text{ to } n \} \) be a set of financial statements. Since a statement is represented using a set of XBRL data elements, each \( d_i \) can be considered to be a set of XBRL elements each uniquely identified by a namespace and the element name. This abuse of notation should not introduce any confusion. Let \( | d_i^{GAAP} | \) be the number of GAAP elements used in \( d_i \), \( | d_i^{CUSTOM} | \) be the number of custom elements used in \( d_i \), and \( | d_i^{OTHER} | \) be the number of other elements used in \( d_i \).
Creating Competitive Markets for Small Businesses with New Media and E-Business Strategy
www.igi-global.com/article/creating-competitive-markets-small-businesses/59913?camid=4v1a