Chapter 4
Retrieving HTML and XBRL Data With Microsoft Excel 2010

Jianing Fang
Marist College, USA

ABSTRACT
The Securities and Exchange Commission (SEC) has upgraded the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system, to the Interactive Data Electronic Applications (IDEA) platform, or the Next-Generation EDGAR (New EDGAR). The SEC issued its final mandate for XBRL adoption and the conversion target dates for all firms in January 2009. With this conversion, users can retrieve the financial statement information of listed companies at both the document level and data element level. This chapter reviews the fundamental concepts of XBRL and reports on the current compliance status of the SEC XBRL conversion mandate. The main task is to demonstrate how to retrieve data from the New EDGAR and how to process the data with Microsoft Excel 2010.

INTRODUCTION
Since the Securities and Exchange Commission’s (SEC) eXtensible Business Reporting Language (XBRL) conversion mandate was issued in 2009, most of the registrants have successfully followed suit. Many have converted their annual financial reports (10K), quarterly financial reports (10Q), and other required financial reports from the Hyper Text Markup Language (HTML) format under the old Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system, which

DOI: 10.4018/978-1-5225-3616-1.ch004

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
Retrieving HTML and XBRL Data with Microsoft Excel 2010

reported only at the document level, to the XBRL format under Interactive Data Electronic Applications (IDEA)—otherwise known as the Next-Generation EDGAR (New EDGAR) – which reports at both document level and data element level (Du, Vasarhelyi & Zheng, 2013). This provided a more efficient and cheaper source of data for financial analysts, investors, students, and other users needing information on the annual or quarterly financial reports of all publicly traded companies (Jones & Willis, 2003). XBRL has since been “ranked as one of [the] top ten technologies for accounting and auditing professionals” (Liu, 2013, p. 60).

This new data element level data depository is a very valuable addition to the existing and ever enlarging stores of qualitative and quantitative data online and available for researcher use publicly. While these public resources can be invaluable to the skilled researchers in many specific fields, there are inherent technical challenges and differences between data reuse and traditional researcher-collected qualitative and quantitative data. Especially in the XBRL domain—due to the lacking efforts and limited opportunities for education and training, many users do not have the necessary knowledge or skills to make full use of this new source of financial data (Harris & Morsfield, 2012; Fang, 2013). In this chapter, the author will review the SEC’s XBRL conversion mandate and the compliance status. Additionally, this chapter will walk through the entire process of retrieving financial data on the New EDGAR system with a simple and practical example and step-by-step instructions with plenty of screenshots for Windows operating systems. The detailed instructions will not only show how to search for companies and their data on IDEA, but it will also demonstrate how to correctly set up Excel worksheets to import and process the HTML and XBRL data.

BACKGROUND


The new rules aimed to render financial information into easier formats for investors to analyze and to assist in automating regulatory filings and business information processing. The New EDGAR system provides information at both the