

Chapter XV

Financial News Analysis Using a Semantic Web Approach

Alex Micu

Erasmus University Rotterdam, The Netherlands

Laurens Mast

Erasmus University Rotterdam, The Netherlands

Viorel Milea

Erasmus University Rotterdam, The Netherlands

Flavius Frasincar

Erasmus University Rotterdam, The Netherlands

Uzay Kaymak

Erasmus University Rotterdam, The Netherlands

ABSTRACT

In this chapter we present StockWatcher, an OWL-based web application that enables the extraction of relevant news items from RSS feeds concerning the NASDAQ-100 listed companies. The application's goal is to present a customized, aggregated view of the news categorized by different topics. We distinguish between four relevant news categories: i) news regarding the company itself; ii) news regarding direct competitors of the company; iii) news regarding important people of the company; and iv) news regarding the industry in which the company is active. At the same time, the system presented in this chapter is able to rate these news items based on their relevance. We identify three possible effects that a news message can have on the company, and thus on the stock price of that company: i) positive; ii) negative; and iii) neutral. Currently, StockWatcher provides support for the NASDAQ-100 companies. The selection of the relevant news items is based on a customizable user portfolio that may consist of one or more of these companies.

INTRODUCTION

Unlike printed media or television programs, on the Web, news items can be made public as soon as they emerge. Simultaneously, Web coverage is continuously increasing. News websites provide RSS-feeds facilitating the public to remain up-to-date on nearly any topic of interest.

To better understand what the impact of the Internet is on our daily lives, we should first take a look at the main ideas behind its creation. The suggestion of social communications through networks dates from 1962, when J.C.R. Licklider, a professor at the Massachusetts Institute of Technology (MIT), suggested the “Galactic Network” theory. In this theory, he imagined a “globally interconnected set of computers through which everyone could quickly access data and programs from any site” (Licklider & Clark, 1962). The next step towards the Internet as we know it today was taken by the Defense Advanced Research Projects Agency (DARPA), which created The Advanced Research Projects Agency Network (ARPANET). This project was the first operational computer network in the world, and is seen as the ancestor of the Internet. As time progressed, different networks were created outside the ARPANET, and became eventually interconnected into one super network in 1990, creating the roots of today’s modern Internet. With the presence of this technological infrastructure, the next step towards public accessibility was the foundation of the World Wide Web (WWW). This project, led by Tim Berners-Lee, included the now so popular Hypertext Markup Language (HTML) being used for the creation of web pages, and the Hypertext Transfer Protocol (HTTP) being used to access Web content.

In our technology-driven society many people have a hard time to even imagine a world without the services and benefits of the Internet. To a certain degree it is safe to assume that the society

we live in is turning into an information technology society, characterized by the Internet use (Slabber, 2007). Presently more than 1.1 billion people (Miniwatts Marketing Group, 2007) make use of services provided by the Internet. Such services include e-mail messaging, file sharing, streaming media and voice communications. By making use of popular search engines such as Google, people all around the world have access to vast amounts of online information, provided by different Web sites. Simultaneously, the same people are even able to create Web content, and place information on the WWW without much effort. Eventually the success of the WWW has made it progressively more challenging to find, access, present, and maintain the information available on the Web.

In 1998 a new idea was born, under the name of Semantic Web (SW) (Berners-Lee, Hendler, & Lassila, 2001). This was supposed to be an extension to the current WWW as we know it. The SW would revolutionize the way in which data is described and presented, so that it can be read, interpreted and used by various software applications. There are three main goals that the SW seeks to achieve: i) provide common formats for integration and combination of data drawn from diverse sources; ii) record how the data relates to real world objects; and iii) semantically link documents.

Data on the Web is controlled by certain applications, and only useable by these applications. The SW tries to make this data neutral, available to all applications. One of the building stones to achieve this is the Resource Description Framework (RDF) (Brickley & Guha, 2004). RDF is a general-purpose language for representing information in the Web. Together with RDF Schema (RDFS), RDF can be used to code the data so that relations and information about the described entities get coded along (Brickley & Guha, 2004). This can be achieved by basing the representation

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/financial-news-analysis-using-semantic/28823

Related Content

A Knowledge-Engineering Approach to the Digital Heritage and Innovation of Traditional Chinese Ink Painting Techniques

Lei Xia (2026). *International Journal of Knowledge Management* (pp. 1-20).

www.irma-international.org/article/a-knowledge-engineering-approach-to-the-digital-heritage-and-innovation-of-traditional-chinese-ink-painting-techniques/402897

Implementing Business Processes: A Database Trigger Approach

Wai Yin Mok, Charles F. Hickman and Christopher D. Allport (2013). *International Journal of Knowledge-Based Organizations* (pp. 36-55).

www.irma-international.org/article/implementing-business-processes/77883

Document Description and Coding as Key Elements in Knowledge, Records, and Information Management

Olugbade Oladokun and Saul F. C. Zulu (2017). *Managing Knowledge Resources and Records in Modern Organizations* (pp. 179-197).

www.irma-international.org/chapter/document-description-and-coding-as-key-elements-in-knowledge-records-and-information-management/173804

Merging Controlled Vocabularies for More Efficient Subject-Based IR Systems

Ioannis Papadakis and Konstantinos Kyprianos (2013). *Dynamic Models for Knowledge-Driven Organizations* (pp. 254-267).

www.irma-international.org/chapter/merging-controlled-vocabularies-more-efficient/74081

Maximize Knowledge Sharing and Learning Opportunities

Rachad El Badawi El Najjar (2023). *Knowledge Management and Research Innovation in Global Higher Education Institutions* (pp. 31-60).

www.irma-international.org/chapter/maximize-knowledge-sharing-and-learning-opportunities/323663