ABSTRACT

This article presents an intelligent corporate governance analysis and rating system, called IDA System, capable of retrieving SEC required documents of public companies and performing analysis and rating in terms of recommended corporate governance practices. With the techniques of analogical learning, local knowledge bases, databases, and question-dependent semantic networks, the IDA System is able to automatically evaluate the strengths, deficiencies, and risks of a company’s corporate governance practices based on the documents stored in the “SEC EDGAR database by (U.S. Securities and Exchange Commission 2013)”. A produced score reduces a complex corporate governance process and related policies into a single number which enables concerned government agencies, investors and legislators to assess the governance characteristics of individual companies.

Keywords: Analogical Learning, Information Retrieval, Knowledge Base, Semantic Net, Text Mining.
costs, it also has the risk associated with human error and bias in their ratings. To reduce the cost, some companies outsource the job to other countries with a low labor cost. But the problem of how to timely respond to a changing world and constantly updates information remains unanswered. With a research grant from the Maryland Industrial Partners Board, we developed and tested an intelligent Knowledge-based Corporate Governance Practice Analysis System capable of analyzing the documents in EDGAR and ranking companies accordingly. The IDA system provides trusted, independent information and resources for government agencies, investors, security analysts, lawyers, accountants and financial institutions. This paper is organized as follows. Section 2 describes the problem domain of the IDA System. Section 3 describes the IDA System description, with experimental cases. Section 4 evaluates the results and performs comparison between different approaches. Section 5 summarizes the research and provides recommendations for future studies.

THE PROBLEM DOMAIN

The required documents by SEC include 8-k, 10-K and DEF 14A (proxy statement) contain passages that are of interest in terms of corporate governance practice. Among the key elements included in the corporate governance section are: bylaws, code of conduct and ethics, board of directors’ information, board committee mandates, basis for executive compensation, and information on insider purchases and sales. Relevant corporate governance and disclosure data should be woven into the company’s 10-K and proxy statement. These filings are important source documents for the rating system. We designed and selected 200 questions based on “Corporate Governance Handbook 2005 Developments in Best Practices, Compliance, and Legal Standards by Carolyn Kay (2005)”. Based on the handbook, these questions are used to retrieve valuable information about the companies to help investors understand corporate governance of the companies they’re willing to invest. Some of these 200 questions are listed below:

1. Does the company provide employment agreements that protect executive producers?
2. Does the management is required to hold accumulated stock for the long-term?
3. Does the compensation committee have policies and programs to recapture incentives from management in case of malfeasance?
4. Does the company use industry standards to decide on type of compensation and level of compensation for executives?
5. Does the company measures and tracks adherence to ethical conduct?

In order to answer these questions, the IDA System analyzes and retrieves answers from companies’ 8-k, 10-k and DEF 14a. A system with natural language processing capabilities must address the problems associated with processing very large unstructured text documents. Natural language processing programs struggle with terminology because of term variation, when a concept is expressed in several different ways; and term ambiguity, when the same term is used to refer to multiple concepts. Term variation and ambiguity may cause irrelevant information to be retrieved and relevant information to be overlooked. The most serious problem faced by a natural language processing system is the time complexity problem. That is, the computational time grows exponentially when the number of words increases. It is a common belief that such a system is very difficult, even impossible, to be developed. In response to these research challenges, we developed the IDA System that employs machine Learning, question-dependent knowledge bases, databases and question-specific semantic networks in order to reduce the problem to a manageable size.
Trust Building Process for Global Software Development Teams: A Review from the Literature
www.igi-global.com/chapter/trust-building-process-global-software/77736?camid=4v1a