Strategic Human Resource Management Content in the Annual Report of Companies: An Analysis through Text Mining

Kumar Kunal Kamal and Ashok Kumar

Abstract

In the light of different annual report studies, the paper argues that annual reports, considered to be a rich source of information about the future strategy of any company, can be analyzed for finding out the importance that top management places on SHRM concepts. By employing frequency analysis under text mining approach, the paper first identifies top ranked SHRM concepts by probing SHRM literature and then compares their ranking in the text of annual reports. The findings suggest that so far as annual report narratives are concerned, the areas of interest of SHRM researchers are disjoint with those of the top management.

Keywords: SHRM, Annual Reports, Text Mining, Concept Mining, Frequency Analysis

1. Introduction

Apart from supplying useful quantitative information, annual reports are a rich source of qualitative data that permit a researcher to do content analysis (Bowman, 1978). The present study employs a text mining approach to analyze the content of the narratives of texts found in Strategic Human Resource Management (SHRM) literature and compares them with those found in annual reports. As literature on Human Resource Management (HRM) is full of arguments in support of strategic importance of human resources for getting any kind of success in a variety of organizations, the paper accentuates SHRM literature. Moreover, as a company's vision, strategies, policies and their execution cannot be thought of without making a reference to strategic human resources, it is expected that annual reports, considered to be a rich source of information about the future strategy of any company, would put due emphasis on their SHRM content, and that this emphasis would get a reflection in the frequency of SHRM concepts mentioned in the annual reports.

The paper is divided into multiple sections: starting with a brief review of literature, it moves on to elaborate the focus areas and then to the paradigms involved in analyzing annual reports. The paper argues that annual reports, considered to be one of the most important documents for strategic communication with the shareholders, can be analyzed for probing the value that top management places on particular 'areas of interest' of SHRM literature. The methodology of finding areas of interest through a text mining approach is detailed, followed by a comprehensive assessment of the issues and challenges involved.

2. Literature Review

Annual reports have been analyzed by researchers for a very wide range of topics: while Bowman (1978) studied the role of annual reports in corporate strategy; Revsine (1991) probed their role in what he termed 'selective financial misrepresentation'; working on a different note, Alessandri (2001) examined the use of annual reports in projecting a positive corporate identity. The use of annual reports as persuasive rhetorical devices has also attracted the attention of researchers (Graves, Flesher, & Jordan, 1996; Thompson, 1991), who have emphasized the application of annual reports as a tool of legitimating process for symbolic management (Tsang, 2001), and their employment in construction of company's social image (Gray, Kouhy, & Lavers, 1995; Hines, 1988). Their use as a tool for managing public impressions (Neu, Warsame, & Pedwell, 1998) and as a marketing

and branding tool (Bekey, 1990; Judd & Tims, 1991) has also been given due emphasis by researchers. On a wider note, researchers have also studied the role of annual reports in democratic processes (Butterworth, Gray, & Haslam, 1989) and have analyzed them as documents that reflect national traits (Dawson & Steward, 1991). Across all these studies, the methodology of research has primarily been content analysis in which the authors have focused on the qualitative aspects of annual reports.

2.1 Focus areas of annual reports analysis

Annual reports have been analyzed by researchers in a very minute way; the prime focus areas of annual reports that have gone under the lens of researchers are: pictorial information, graphical information, president's/chairman's message, the narrative part, and the form and structure of the whole text. While discussing pictorial information, researchers have focused on issues such as: sex-role stereotyping (Anderson & Imperia, 1992); male-female depiction in annual reports and their relation to perception of corporate climate (Kuiper, 1988); ways in which one can look at the images, viz. representational, constitutive, and ideological (Preston, Wright, & Young, 1996), and have even linked the representation of images in annual reports with the representation in other media (Graves, Flesher, & Jordon, 1996). The focus on graphical information has attracted a lot of attention from researchers who have underlined selective misrepresentation of information through graphs: the use and abuse of graphs in representing information is a well-researched topic (Beattie & Jones, 1992; Johnson, Rice, & Roemmich, 1980; Mather, Mather, & Ramsay, 2005; Mather, Ramsay, & Serry, 1996; Mather, Ramsay, & Steen, 2000). Researches done on graphical information are also rich with longitudinal (Bartlett & Jones, 1997; Beattie, Dhanani, & Jones, 2008; Beattie & Jones, 2000) and comparative studies across nations (Beattie & Jones, 1997, 2001).

Turning to president's/chairman's message in annual reports: Abrahamson and Amir (1996) examined them for relationship between the use of words with negative connotations and their relationship with financial measures like sales, book rate of return, etc. In a similar vein, Smith and Taffler (2000) explored for associations between the narrative of chairman's message and the event of firm failure. While Hyland (1998) examined the role of metadiscourse to show the persuasive nature of the messages that try to realize rational, credible, and affective appeals, Hooghiemstra (2008) examined the narratives for presence of self-serving attribution bias across cultures in explaining company results to shareholders. Researchers have also looked at the linguistic structures of the management messages to establish a link between the changes in linguistic structures and the kind of information conveyed (Thomas, 1997).

The narrative part of annual reports has also attracted attention from many researchers: while Lajili and Zegal (2005) examined the risk information disclosures in the narratives for their usefulness in gaining insights into risk environment, Tennyson, Ingram and Dugan (1990) have analyzed the content in explaining bankruptcy. The narratives have also been explored to gauge the ethical identity of Islamic banks (Haniffa & Hudaib, 2007) and have been the subject of analysis in understanding the role of ideological belief in issues of environmental reporting (Yapa, Harvey, & Ellis, 2005). Working on a different note, researchers like Zeghal and Ahmed (1990) have argued that narratives in annual reports alone are not adequate to represent the information disclosure activities of a firm or an industry, and have emphasized the utility of disclosures through other mass media vehicles.

The overall form and structure of the annual reports has also been analyzed in detail, with most studies focusing on longitudinal analysis to find out the differences in form and structure of annual reports published through a time period (Bartlett & Jones, 1997; Beattie, Dhanani, & Jones, 2008; Lee, 1994; McKinstry, 1996). Researchers have also attempted to arrive at a generic model of annual reports that can be used by companies to either confirm to a 'typical' annual report or to separate one's annual report from the 'herd' (Marino Jr, 1995).

2.2 Paradigms in the analysis of annual reports

The paradigms applied to the analysis of annual reports have in general focused on syntactic probes utilizing application of readability formulas and frequency analysis; however, semantic and stylistic probes have also been undertaken. It might be noted that researchers have often applied multiple forms of analysis to probe the texts.

While the focus of syntactic analysis has been the analysis of various readability indices like Flesch and Fog (Parker, 1982; Pashalian & Crissy, 1950; Smith & Smith, 1971; Soper & Dolphin Jr, 1964; Subramanian, Insley, & Blackwell, 1993), researchers have done semantic analysis in different ways like: investigation of variation in the language of annual report and verification through utilization of a general semantic model of extensional-intentional paradigm (Seidler & Lebar, 1982); putting annual reports under lens of Haberman's norms of communicative action (Yuthas, Rogers, & Dillard, 2002); and, inquiring the text of annual reports for the application of 'Pollyanna Principle' (Kohut & Segars, 1992). While doing stylistic analysis, researchers have examined the word choice, frequency of use, and other aspects of linguistic information like verb structures and thematic structure (Hyland, 1998; Rutherford, 2005; Thomas, 1997). The focus of frequency analysis has been the computation of frequency with which a word, concept, lines of prose or disclosers are mentioned (Bowman, 1978, 1984).

2.3 Text mining annual reports

Annual reports are considered to be the most important external document of any company as they contain crucial information about their financial performance and their future strategies (Kloptchenko, et al., 2004). However, as Kloptchenko, et al. (2004) argue, the quantitative part of an annual report, along with the tables and graphs that visualize the quantitative information, reflect only the past and not the future - it is the linguistic structure and the written style of the text that indicates how a company will do. Rutherford's (2005) analysis of annual reports

to identify the genre to which the annual reports belong is an important study in which the researchers applied a 'text mining approach' to find out the coverage given to particular words based on the frequency of their occurrence.

As evident through review of literature, no study has yet attempted to analyze company annual reports using text mining tools to probe their content in order to find out the relative importance that the top management places on SHRM concepts.

3. Methodology

As the present study requires finding out the focus areas of researchers in the field of SHRM and inquiring if they get a reflection in the text of the annual reports, the focus areas of SHRM literature are found first: to do so, measurable surrogate variables are established. Grounding on the argument of Bowman (1978), that "surrogate variables which can be measured are loosely coupled to and can be used as approximations for the variables of interest", the paper identifies 'frequency of occurrence of phrases or concepts in texts' as the surrogate variable and the 'focus areas of top management' as variables of interest. The paper argues that the frequency of occurrence of concepts would capture the relative importance of concepts in the texts probed. It might be noted that even though researchers have differed in their use of the terms 'phrase' and 'concept', the terms are used interchangeably in this paper.

Two studies require elaborate mention; both studies are recent and have applied a text mining approach to analyze narratives: while Peslak (2005) analyzed the narratives of annual reports, Singh, Hu, and Roehl (2007) probed research articles. Peslak's study was exploratory in nature and examined the relative importance of information technology through content analysis of annual reports. The methodology applied included identification of important terms and examining their inclusion in annual reports: while frequent and common IT terms were identified through analysis of various sources, their inclusion in the annual reports was examined through two ways - a discrete yes/no variable (whether the term was

mentioned or not) and the frequency of the mentions. A similar approach was applied in Singh et al.'s (2007) study but to an altogether different dataset: they selected 40 research papers in the field of human resource management published in the International Journal of Hospitality Management between the years 1994 - 2003. The authors did a textual analysis through CATPAC software to identify high frequency keywords that clustered together to present the prominent themes within the field of human resources. The study identified nine major research themes and the authors discussed each thematic development with respect to time and country. As the authors asserted, the study demonstrated the usefulness of text-mining application in the field of human resource management.

Both the studies discussed above applied a text mining approach to discover knowledge from narratives. Taking clue from these studies, the present study applies a text mining approach to discover patterns from one dataset comprising narratives of books, and looks out for their occurrence in another dataset comprising narratives of annual reports. This way the present study does a comparative analysis of the narratives of the two datasets.

3.1 Applying text mining to probe narratives

Text Analysis is a form of content analysis that deals with the textual form of information; while content analyses can be performed on any symbolic material, be it visual, acoustic, tactile or whatever, text analyses are only of text and transcripts (Roberts, 1997). The approach in this study is to analyze the text through 'mining' of the narratives. To put objectivity in the study, the narratives are mined through the text mining software SPSS Clementine 12.0 (with add-on module for text mining) that uses LexiQuest, considered a strong tool with solid linguistic foundation (Sullivan, 2004). The study is divided into two parts: first the important concepts from SHRM literature are identified; next their importance in the text of annual reports is examined.

3.2 Application of a concept based approach

The present study is focused on concept-based approach: major phrases used in the academic environment (i.e. books) are found out and compared with those found in the business environment (i.e. annual reports). The focus on concept approach is useful as it excels over both rule-based natural language processing and word-based approach: while rule-based approach is marred by exceptions, researchers have argued that using phrases/concepts as a basic unit is more practical and useful than using words as basic unit (Cohen & Hersh, 2005; Majoros, Subramanian, & Yandell, 2003). The present study therefore uses concepts as the basic unit of analysis. To authors' knowledge, this is the first study that uses a concept-based approach to mine the texts of annual reports and academic writings in the field of SHRM.

3.3 Concept extraction

The primary subtask in the present study was 'concept extraction'. The text-mining program used in this study helps find the concepts and their frequency of occurrence in any given dataset. It has been argued that in order to discover the hidden patterns of information one should combine mining function with a trial-and-error approach (Uramoto, et al., 2004). The choice of various settings during the computerized text mining, e.g. giving inputs to the program in terms of the number of concepts (based on global frequencies) it should return for further analysis, is an example of the application of trial-and-error approach used in the present study.

A major challenge of text mining approach is considered to be 'term ambiguities'. Term ambiguity occurs when the same term is used to denote multiple concepts: it has been argued that researchers would have to use their discretion to resolve any term ambiguities that might creep in during the whole process (Spasic, et al., 2005). However, as the textual databases used in the present study refer to specialized domains (the annual reports referring to a particular domain of business writing, and the

books focusing on issues related to SHRM) term ambiguity is expected to play a restricted role in the present study, more so, because the language used in particular domains are supposed to contain lesser ambiguities because of the use of explicit framework and the closed vocabulary (Harris, 2002).

Another problem that affects text mining is the problem of mapping synonyms to a single concept. It has been argued that if all of the synonyms and abbreviations for an entity (the concept in the present study) are mapped to a single entity representing the concept then text-mining tasks would become more efficient (Cohen & Hersh, 2005). This particular problem of mapping synonyms to a single concept is taken care of by the text-mining tool which the present study employs as it matches different 'entities' (i.e. phrases) to a singular concept.

3.4 Selection of samples

As in the field of information retrieval, it is impractical to observe entire populations (Cohen & Hersh, 2005), the following sample populations were selected for the present study: a) Annual reports of companies belonging to BSE Sectoral Indices representing eleven different sectors (Auto, Capital Goods, Consumer Durables, FMCG, Healthcare, IT, Metal, Oil & Gas, Power, and Realty); b) Electronic format of books related to the field of human resource management and dealing specifically with its strategic nature. This way, a total of 11 books (representing SHRM literature) and annual reports of 142 companies (representing BSE sectoral indices) for a period of six years (2003-2008) comprised the sample data for the study. For the list of 11 books used in the study please refer to Appendix I.

3.5 Data processing

Although the text-mining tool used in the study supports multiple file formats, to decrease the processing time of data, all the documents were converted to TXT format. This format is the simplest format and the most efficient for text mining purposes. This process also ensured that the concepts identified would be with respect to 100% of the contents of the

documents concerned, i.e. books and the annual reports. Next, the data comprising individual files were put into different folders to assist in proper analysis. The data was run through the SPSS Clementine 12.0 program under the following model: creating a concept model based on global frequencies of top 10000 concepts in the documents-sets. The frequency based concept model was used to identify both the 'areas of interest' of the researchers as well as probe the areas identified in the annual reports. It might be noted here that as the whole process of concept identification and extraction was completed by the software, there was no need to test inter-coder reliability (Stathopoulou, 2005).

4.   Results

The data forming the 11 books were analyzed to extract top 50 concepts based on their frequency of occurrence. Table 1 contains the list of the top 50 concepts identified from the books (representative of SHRM literature). The top 50 concepts taken together accounted for a total of 7.87 % (rounded off to two decimal places) of all the concepts extracted from the books. On the contrary, these concepts could account only for a total of 0.12 % of total concepts extracted from the annual reports. It was surprising to find that some of the concepts among the top 50 concepts of SHRM literature could not find a place in the top 10000 concepts extracted from annual reports, thus getting no representation in the final percentage outcomes calculated on the basis of the frequency of occurrence. The details are provided in exhibit 1. Removing those concepts that did not find a place in top 10000 concepts extracted from annual reports, a final list of concepts was drawn to calculate the ranks of the concepts based on their representation in annual reports. A list of the top concepts of SHRM literature based on their representation in annual reports is given in exhibit II.

The highest rank-getter was the concept of 'human resources' followed by 'information technology'. Concepts like human resource planning, HR strategies, and strategic learning, which have often been used by the researchers in the SHRM field, could

not make it to top 10000 concepts. Even concepts like people management, personnel management, human resource practices, and HR function, considered to be strategically very important to companies, did not find a place in the top 10000 concepts.

As the present study involved analysis of annual reports of companies belonging to different sectors of the economy, it was possible to conduct a cross sectoral inquiry: the representations of the top 50 concepts (from the books) in the annual reports of the different companies representing different sectors were found to be largely non-variant with IT sector taking a lead in mention of SHRM concepts (see Exhibit III).

Next, the study focused on application of a time-based approach to examine if there have been any major variations in the mention of SHRM concepts by the companies through the time period of Year 2003-2008. For this, the period was divided into three time frames, each representing a time period of two years (2003-2004, 2005-2006, and 2007-2008). Apart from doing a consolidated analysis on all the sectors (wherein all the data was put under three headings representing the three time periods), a time based analysis of the sectoral indices was also performed: the outputs showed no significant differences except for the realty sector which showed a marked increase in mention of concepts related to the field of SHRM in the annual reports during the period 2007-2008; IT sector dominated all sectors in the three time periods with Oil & Gas sector closely following the lead (for details please refer to Exhibit IV and V).

5.   Discussion

The findings suggest that top management prefers to give the important concepts of the field of SHRM a miss when it comes to communicating through one of its most strategically important documents. More so, even in terms of time-periods it was found that the mention of SHRM concepts in annual reports has more or less remained the same: in fact a decrease in the mention of SHRM concepts was indicated in comparing the annual reports of the recent time period of 2007-2008 to the time period of 2003-2004. The

dismal representation of the Top 50 concepts of SHRM in the annual reports raise an important question: are the areas of interest of SHRM researchers disjoint with the areas of top management as reflected through annual reports? The findings suggest that there indeed is dissociation: however, this conclusion is based only on the annual reports of the companies and does not include communication through other media. Even so, this is an alarming situation in a country like India that has strategic advantage of the size and variety of skilled human resources. This should also be seen in the light of the fact that annual reports are considered an important tool of communication for putting across future strategies to the stakeholders of any organization.

Text mining is not without its challenges. One of the major challenges, as identified by Cohen and Hersh (2005), is the lack of user-friendly tools and restricted access to full texts. Another challenge is the lack of updated lexical databases - as most text mining tools use the lexical database of 'WordNet' which is not domain specific, in order to improve the efficiency and effectiveness of the text mining tools, researchers need to come up with updated lexical databases containing domain specific terms. This challenge can however be attended with ease as researchers in the field of human resource management have already built domain specific dictionaries and encyclopedias (Cartwright, 2005; Heery & Noon, 2008; Ivanovic & Collin, 2006); this way, creating a domain specific lexical database compatible with WordNet would not be that challenging a task.

Another important problem concerns the nature of the Data. The two biggest hindrances in text mining approach are: a) the availability of the document, and b) its readiness for text mining purposes, i.e. it should be 'readable' for the program. Most of the data available, even when in electronic form, are not scanned to suit text mining purposes, or are sometimes encrypted - two biggest hindrances to data processing when it comes to data mining, as OCR (optical character recognition) scanning of the non-OCR documents is a necessity of almost all the text mining tools. In non-OCR scanned documents, the text in the

images/files cannot be identified, rendering text mining an impossible task. Similar is the case with encrypted documents - encryption being a major roadblock in OCR conversion that impedes the entire process of text mining. Researchers who apply text-mining approach to data analysis should not overlook these issues.

In the modern day where rate of publication of scholarly texts itself is so high that it is virtually impossible for any researcher to keep abreast with the huge information overload that the avid researcher faces, text mining would come handy by giving a sneak preview of the happenings in the texts. Text mining would help the researchers identify new concepts being discussed and would help them locate documents that would be most relevant. Such an approach would be even more important for a researcher who is working on an interdisciplinary research as it would help identify important documents without much hindrance; such approaches are already being used by bibliographic databases that help researchers identify the relevant scholarly texts; moreover a detailed text mining approach will help researchers identify relations which they could not find in a restricted environment. This is an approach that needs more focus from researchers, and is waiting to be exploited for the benefit of the research community and for the society at large. The areas of strategic importance such as SHRM should be analyzed through text mining tools to draw the attention of the top management. The periodical feedback of such a nature will definitely lead to action and communication about SHRM by the top management in giving it due importance.

References

Abrahamson, E., & Amir, E. (1996). The information content of the president's letter to shareholders. *Journal of Business Finance & Accounting,* 23(8), 1157-1182.

Alessandri, S. (2001). Modeling corporate identity: A concept explication and theoretical explanation. Corporate Communications: *An International Journal,* 6(4), 173-182.

Anderson, C. J., & Imperia, G. (1992). The corporate annual report: A photo analysis of male and female portrayals. *Journal of Business Communication,* 29(2), 113-128.

Bartlett, S., & Jones, M. (1997). Annual reporting disclosures 1970-90: An exemplification. *Accounting, Business & Financial History,* 7(1), 61-80.

Beattie, V., Dhanani, A., & Jones, M. (2008). Investigating presentational change in UK annual reports: A longitudinal perspective. *Journal of Business Communication,* 45(2), 181.

Beattie, V. A., & Jones, M. J. (1992). The use and abuse of graphs in annual reports: Theoretical framework and empirical study. *Accounting & Business Research,* 22(88), 291-303.

Beattie, V. A., & Jones, M. J. (1997). A comparative study of the use of financial graphs in the corporate annual reports of major U.S. and U.K. companies. *Journal of International Financial Management & Accounting,* 8(1), 33-68.

Beattie, V. A., & Jones, M. J. (2000). Changing graph use in corporate annual reports: A time-series analysis. *Contemporary Accounting Research,* 17(2), 213-226.

Beattie, V. A., & Jones, M. J. (2001). A six-country comparison of the use of graphs in annual reports. International Journal of Accounting, 36(2), 195-222.

Bekey, M. (1990). Annual reports evolve into marketing tools. *Financial Manager,* 3(1), 50-60.

Bowman, E. H. (1978). Strategy, annual reports, and alchemy. California Management Review, 20(3), 64-71.

Bowman, E. H. (1984). Content analysis of annual reports for corporate strategy and risk. *Interfaces,* 14(1), 61-71.

Butterworth, P., Gray, R. H., & Haslam, J. (1989). The local authority annual report in the UK: An exploratory study of accounting communication and democracy. *Financial Accountability & Management,* 5(2), 73.

Cartwright, S., & Blackwell Publishing, L. (2005). *The Blackwell encyclopedia of management : Human resource management. Malden,* MA: Blackwell Pub.

Cohen, A. M., & Hersh, W. R. (2005). A survey of current work in biomedical text mining. *Briefings in Bioinformatics,* 6(1), 57-71.

Dawson, S., & Steward, K. (1991). Annual reports and 1992: Cultural differences and the creation of a Pan-European marketing tool. *European Business Journal,* 3(4), 48-52.

Graves, O., Flesher, D., & Jordan, R. (1996). Pictures and the bottom line: The television epistemology of US annual reports. *Accounting, Organizations and Society,* 21(1), 57-88.

Gray, R., Kouhy, R., & Lavers, S. (1995). Corporate social and environmental reporting. *Accounting, Auditing and Accountability Journal,* 8(2), 47-77.

Haniffa, R., & Hudaib, M. (2007). Exploring the ethical identity of Islamic banks via communication in annual reports. *Journal of Business Ethics,* 76(1), 97-116.

Harris, Z. S. (2002). The structure of science information. *Journal of Biomedical Informatics,* 35(4), 215-221.

Heery, E., & Noon, M. (2008). *A dictionary of human resource management.* Oxford: Oxford University Press.

Hines, R. (1988). Financial accounting: In communicating reality, we construct reality. *Accounting, Organizations and Society,* 13(3), 251-261.

Hooghiemstra, R. (2008). East-West differences in attributions for company performance: A content analysis of Japanese and U.S. corporate annual reports. *Journal of Cross-Cultural Psychology,* 39(5), 618-629.

Hyland, K. (1998). Exploring corporate rhetoric: Metadiscourse in the CEO's letter. *Journal of Business Communication,* 35(2), 224.

Ivanovic, A., & Collin, P. H. (2006). *Dictionary of human resources and personnel management.* London: A & C Black.

Johnson, J., Rice, R., & Roemmich, R. (1980). Pictures that lie: The abuse of graphs in annual reports. *Management Accounting,* 62(4), 50-56.

Judd, V., & Tims, B. (1991). How annual reports communicate a customer orientation. I*ndustrial Marketing Management,* 20(4), 353-360.

Kloptchenko, A., Eklund, T., Karlsson, J., Back, B., Vanharanta, H., & Visa, A. (2004). Combining data and text mining techniques for analysing financial reports. *Intelligent Systems in Accounting, Finance & Management,* 12(1), 29-41.

Kohut, G., & Segars, A. (1992). The president's letter to stockholders: An examination of corporate communication strategy. *Journal of Business Communication,* 29(1), 7-21.

Kuiper, S. (1988). Gender representation in corporate annual reports and perceptions of corporate climate. *Journal of Business Communication,* 25(3), 87-94.

Lajili, K., & Zeghal, D. (2005). A content analysis of risk management disclosures in Canadian annual reports. *Revue Canadienne des Sciences de l'Administration/Canadian Journal of Administrative Sciences,* 22(2), 125-142.

Lee, T. (1994). The changing form of the corporate annual report. *The Accounting Historians Journal,* 21(1), 215-232.

Majoros, W. H., Subramanian, G. M., & Yandell, M. D. (2003). Identification of key concepts in biomedical literature using a modified Markov heuristic. *Bioinformatics,* 19(3), 402-407.

Marino Jr, A. J. (1995). Separating your annual report from the herd. *Public Relations Quarterly,* 40(2), 44-47.

Mather, D., Mather, P., & Ramsay, A. (2005). An investigation into the measurement of graph distortion in financial reports. *Accounting and Business Research,* 35(2), 147-159.

Mather, P., Ramsay, A., & Serry, A. (1996). The use and representational faithfulness of graphs in annual reports: Australian evidence. *Australian Accounting Review,* 6(12), 56-63.

Mather, P., Ramsay, A., & Steen, A. (2000). The use and representational faithfulness of graphs in Australian IPO prospectuses. *Accounting, Auditing & Accountability Journal,* 13(1), 65-83.

McKinstry, S. (1995). Designing the annual reports of Burton PLC from 1930 to 1994. *Economics and management working papers-University of Paisley.*

Neu, D., Warsame, H., & Pedwell, K. (1998). Managing public impressions: Environmental disclosures in annual reports. *Accounting, Organizations and Society,* 23(3), 265-282.

NSE. (2009). Standard & Poor's: S&P CNX Nifty methodology. Retrieved September 20, 2009, from http://www.nse-india.com/content/indices/nifty_freefloat_method.pdf

Parker, L. (1982). Corporate annual reporting: A mass communication perspective. *Accounting and Business Research,* 12(48), 279-286.

Pashalian, S., & Crissy, W. (1950). How readable are corporate annual reports. *Journal of Applied Psychology,* 34(4), 244-248.

Peslak, A. R. (2005). The importance of information technology: An empirical and logitudinal study of the annual reports of the 50 largest companies in the United States. *Journal of Computer Information Systems,* 45(3), 32-42.

Preston, A., Wright, C., & Young, J. (1996). Imag[in]ing annual reports. *Accounting, Organizations and Society,* 21(1), 113-137.

Revsine, L. (1991). The selective financial misrepresentation hypothesis. *Accounting Horizons,* 5(4), 16-27.

Roberts, C. W. (1997). Introduction. In C. W. Roberts (Ed.), *Text analysis for the social sciences : Methods for drawing statistical inferences from texts and transcripts* (pp. 1). Mahwah NJ.: Erlbaum.

Rutherford, B. (2005). Genre analysis of corporate annual report narratives. *Journal of Business Communication,* 42(4), 349-378.

Seidler, L. J., & Lebar, M. A. (1982). A general semantics analysis of selected sections of the 10-K, the annual report to shareholders, and the financial press release. *Accounting Review,* 57(1), 176.

Singh, N., Hu, C., & Roehl, W. (2007). Text mining a decade of progress in hospitality human resource management research: Identifying emerging thematic development. *International Journal of Hospitality Management,* 26(1), 131-147.

Smith, J., & Smith, N. (1971). Readability: A measure of the performance of the communication function of financial reporting. *Accounting Review,* 552-561.

Smith, M., & Taffler, R. (2000). The chairman's statement: A content analysis of discretionary narrative disclosures. *Accounting, Auditing & Accountability*

Journal, 13(5), 624-647.

Soper, F., & Dolphin Jr, R. (1964). Readability and corporate annual reports. *Accounting Review,* 358-362.

Spasic, I., Ananiadou, S., McNaught, J., & Kumar, A. (2005). Text mining and ontologies in biomedicine: Making sense of raw text. *Briefings in Bioinformatics,* 6(3), 239-251.

Stathopoulou, T. (2005). Using text mining tools for event data analysis. *Studies in Fuzziness and Soft Computing,* 185, 239.

Subramanian, R., Insley, R. G., & Blackwell, R. D. (1993). Performance and readability: A comparison of annual reports of profitable and unprofitable corporations. *Journal of Business Communication,* 30(1), 49-61.

Tennyson, B., Ingram, R., & Dugan, M. (1990). Assessing the information content of narrative disclosures in explaining bankruptcy. *Journal of Business Finance & Accounting,* 17(3), 391-410.

Thomas, J. (1997). Discourse in the marketplace: The making of meaning in annual reports. *Journal of Business Communication,* 34(1), 47-66.

Thompson, G. (1991). Is accounting rhetorical? Methodology, Luca Pacioli and printing. *Accounting, Organizations and Society,* 16(5/6), 572-599.

Tsang, E. W. K. (2001). Annual report disclosure and corporate legitimacy management: A study of Singapore companies. *Asia Pacific Journal of Management,* 18(1), 27.

Uramoto, N., Matsuzawa, H., Nagano, T., Murakami, A., Takeuchi, H., & Takeda, K. (2004). A text-mining system for knowledge discovery from biomedical documents. *IBM Systems Journal,* 43(3), 516-533.

Yapa, P. W. S., Harvey, D., & Ellis, G. (2005). Disclosure of corporate environmental policies in annual reports: Further evidence incorporating a national ideology - an Australian study. *Journal of Asia-Pacific Business,* 6(1), 75-90.

Yuthas, K., Rogers, R., & Dillard, J. F. (2002). Communicative action and corporate annual reports. *Journal of Business Ethics,* 41(1/2), 141-157.

Zeghal, D., & Ahmed, S. (1990). Comparison of social responsibility information disclosure media used by Canadian firms. *Accounting, Auditing & Accountability Journal,* 3(1), 38-53.

APPENDIX I

List of Books

1. Armstrong, M. (2008). Strategic human resource management : a guide to action. London; Philadelphia: Kogan Page.

2. Eigenhuis, A., & Dijk, R. v. (2008). HR strategy for the high performing business : inspiring success through effective human resource management. London; Philadelphia: Kogan Page.

3. Farazmand, A. (2007). Strategic public personnel administration : building and managing human capital for the 21st century. Westport, Conn.: Praeger Publishers.

4. Gratton, L. (1999). Strategic human resource management : corporate rhetoric and human reality. Oxford; New York: Oxford University Press.

5. Greer, C. R. (2001). Strategic human resource management. Upper Saddle River, NJ: Prentice Hall.

6. Grieves, J. (2003). Strategic human resource development. London; Thousand Oaks, Calif.: Sage Publications.

7. Ingham, J. (2006). Strategic human capital management : creating value through people. Amsterdam; Boston: Butterworth-Heinemann.

8. Lawler, E. E., Boudreau, J. W., Mohrman, S. A., Mark, A. Y., Neilson, B., Osganian, N., et al. (2006). Achieving strategic excellence : an assessment of human resource organizations. Stanford, Calif.: Stanford Business Books, an imprint of Stanford University Press.

9. Lawler, E. E., & Mohrman, S. A. (2003). Creating a strategic human resources organization : an assessment of trends and new directions. Stanford, Calif.: Stanford University Press.

10. Storey, J., Wright, P. M., & Ulrich, D. (2009). The Routledge companion to strategic human resource management. London; New York: Routledge.

11. Vince, R. (2004). Rethinking strategic learning. London; New York: Routledge.

EXHIBIT I
Top 50 SHRM Concepts and their Occurrences

Rank	Concept	% of Occurrence (SHRM Literature)	% of Occurrence (Annual Reports)
1	strategic human resource management	0.747	Not Represented
2	human resource management	0.555	0.003
3	human resources	0.374	0.036
4	business strategy	0.351	0.007
5	hr function	0.350	Not Represented
6	human capital	0.331	0.004
7	line managers	0.233	Not Represented
8	competitive advantage	0.223	0.008
9	hr strategies	0.220	Not Represented
10	strategic public personnel administration	0.201	Not Represented
11	human resource planning	0.167	Not Represented
12	hr professionals	0.165	Not Represented
13	performance management	0.160	0.002
14	organizational learning	0.159	Not Represented
15	hr practices	0.152	Not Represented
16	strategic hrm	0.152	Not Represented
17	people management	0.149	Not Represented
18	public administration	0.146	Not Represented
19	strategic learning	0.146	Not Represented
20	strategic human capital management	0.140	Not Represented
21	hr effectiveness	0.136	Not Represented
22	strategic partner	0.135	0.001
23	senior managers	0.129	Not Represented
24	hr executives	0.112	Not Represented
25	strategic planning	0.112	0.002
26	business units	0.111	0.008
27	human resource practices	0.104	Not Represented
28	team members	0.104	0.001
29	organizational performance	0.103	Not Represented
30	strategic human resource development	0.101	Not Represented
31	best practices	0.100	0.012
32	organizational dynamics	0.100	Not Represented
33	change agent	0.099	Not Represented

34	change initiatives	0.096	Not Represented
35	organizational culture	0.092	Not Represented
36	performance appraisal	0.092	Not Represented
37	public sector	0.016	0.007
38	human resource function	0.089	Not Represented
39	strategic human resources organization	0.084	Not Represented
40	firm performance	0.081	Not Represented
41	change management	0.079	Not Represented
42	information technology	0.078	0.016
43	organizational change	0.078	Not Represented
44	business partner	0.077	Not Represented
45	personnel management	0.077	Not Represented
46	hr organization	0.075	Not Represented
47	organizational design	0.074	Not Represented
48	industrial relations	0.073	0.013
49	hr activities	0.073	Not Represented
50	organizational behavior	0.071	Not Represented

EXHIBIT II
Representation of Top Concepts of SHRM Literature in Annual Reports

Rank (SHRM Literature)	Rank (Annual Reports)	Concept	% of Occurrence (SHRM Literature)	% of Occurrence (Annual Reports)
3	1	human resources	0.374	0.036
42	2	information technology	0.078	0.016
48	3	industrial relations	0.073	0.013
31	4	best practices	0.100	0.012
26	5	business units	0.111	0.008
8	6	competitive advantage	0.223	0.008
37	7	public sector	0.016	0.007
4	8	business strategy	0.351	0.007
6	9	human capital	0.331	0.004
2	10	human resource management	0.555	0.003
13	11	performance management	0.160	0.002
25	12	strategic planning	0.112	0.002
22	13	strategic partner	0.135	0.001
28	14	team members	0.104	0.001

EXHIBIT III
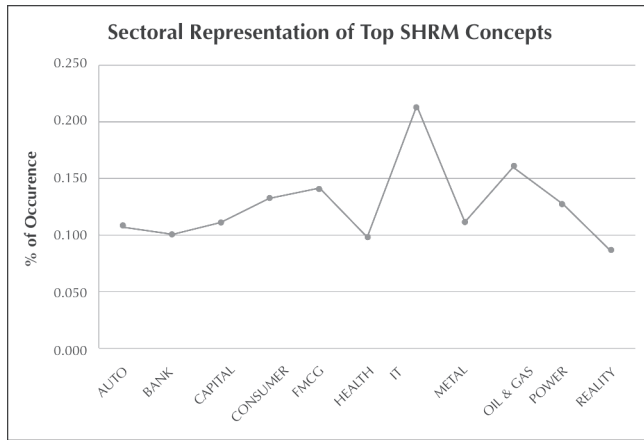**Sectoral Representation of Top SHRM Concepts**

EXHIBIT IV
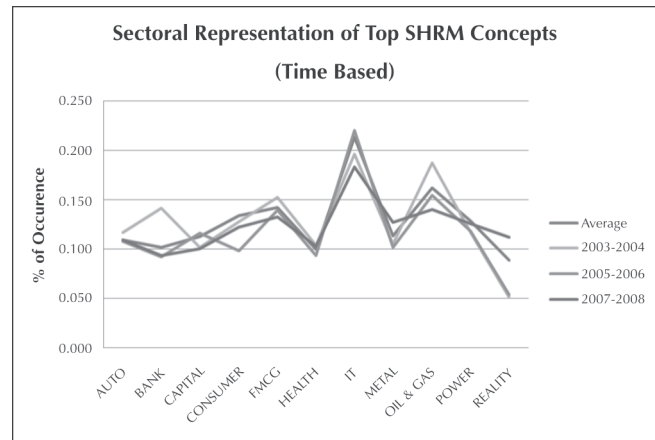**Sectoral Representation of Top SHRM Concepts**
(Time Based)

EXHIBIT V
Time Based Analysis of Top Concepts of SHRM Literature

	Average	**2003-2004**	**2005-06**	**2007-08**
Consolidated	0.122	0.129	0.121	0.123
AUTO	0.109	0.117	0.108	0.109
BANK	0.101	0.142	0.092	0.094
CAPITAL	0.112	0.102	0.116	0.101
CONSUMER	0.134	0.128	0.098	0.122
FMCG	0.142	0.152	0.139	0.133
HEALTH	0.099	0.104	0.093	0.103
IT	0.213	0.196	0.220	0.183
METAL	0.113	0.105	0.102	0.127
OIL&GAS	0.162	0.187	0.154	0.140
POWER	0.128	0.117	0.118	0.126
REALTY	0.089	0.051	0.054	0.112

Kumar Kunal Kamal is a Doctoral Student in the field of Organizational Behavior and Human Resource Management at Indian Institute of Management Indore. His research interests include linguistic diversity, linguistic identity, and organizational communication. He has presented papers in many national and international conferences around the world and is a member of many professional bodies like Academy of Management, Academy of International Business, Indian Association for Commonwealth Literature and Language Studies, Linguistic Society of India, and the National HRD Network.

Ashok Kumar is Professor and Head in HR & OB area of the Amity Business School, Lucknow. Prior to this, he was a Professor at Indian Institute of Management Indore. He has experience of 40 years in Academics and Industry. He has served as General Manager in Steel Authority of India Ltd. He has many publications to his credit.