Corporate Social Responsibility (CSR) or Sustanibility Text Analysis

For Energy Utilities

Abid khan
Department of Micro Data Analysis
Dalarna University
Borlänge, Sweden
h18abikh@du.se

***Abstract*—Corporate Social Responsibility (CSR) is gaining popularity in recent times. The traditional profit-based strategy considering only financial aspects of the organizations owns benefits are not appreciated by society. That why Corporate Social Responsibility (CSR) is taken more seriously and more companies are publishing CSR reports. CSR reports reflect the commitment of organizations towards well-being of all stockholders like employees, customers and environment by balancing their business interests. The aim of this report is to understand the term Corporate Social Responsibility CSR or Sustainability. We want to know; whether companies CSR's reflect some serious strategy or false promises toward well-being of society. The text mining technique of data mining will be used to explore the contents of reports submitted by companies in GRI's Sustainability Disclosure Data launched in 2000. We will try to find out similar patterns or clusters in relation to the terms of Corporate Social Responsibility. The aim of this study to analyze the reports of Energy Utilities companies. For this analysis we will be using R Studio text mining package (tm) with ggplot2 for visualization.**

Keywords-component; Corporate Social Responsibility, stakeholders, Sustainability, text mining, visualization

1. INTRODUCTION

The basic concept of Corporate Social Responsibility (CSR) is growth with sustainability. This term can be elaborated as development by caring about environment and all the stakeholders. There are several ways to understand the sustainable development. The most recent CSR strategies given by Garriga and Mele[1] are based on four points

1. economic goals for achieving profits
2. business decision making with responsibility
3. considering social requirement of the society
4. ethically good measures

The quality of CSR reports depends upon the commitment of companies and stakeholders. The building blocks of CSR reports are mainly socio-economic and elaborations about environment. The aim of this study is to look into CSR reports and try to find out the promised objectives are meet or not. More companies are revealing their commitment towards sustainability-based strategies related to environment, employee's well-being and stockholders' interest [2]. Compiling CSR reports enables companies to take a look into their business strategies and increase their competitiveness [3]. GRI's Sustainability Disclosure Data contains CSR reports of different industries, our focus of study will remain for Energy Utilities companies.

2. REVIEW OF LITERATURE

The Corporate Social Responsibility (CSR) is becoming essential part of business plan. The idea or logic behind CSR is to generate profits for enterprises by taking care of stockholders, employees, customers, environment and society. Study show that more and more companies are including their CSR activities in annual reports [4]. There different angle to look at CSR reports, whereas Roca and Searcy's [5] studied about the main points or indicators of CSR reports and their similarity. In this study by going through more than 90 reports, they discover wide range of indicators. The most common indicators are funding, donations, employees, customers, community well-being and environment. Weber [6] thinks that CSR uplift the image of companies and help to improve their business. Some researchers believe that CSR reports base on GRI standards are not reflecting the true pictures. Boiral's [7] thinks that 90% of the negative information is not disclosed. There are companies who claims in CSR reports their

exaggerated positive image. The purpose of this article is to explore the basic statistics about CSR reports published by Energy Utilities companies.

3. METHODOLOGY AND DATA

3.1 Content Analysis

This paper is supposed to investigate the CSR report published by Energy Utilities companies, for this purpose content analysis is the perfect method to look into CSR reports for some patterns or clusters. Content analysis is carried out on the bases of words termed as called bag of words. Where bag of words considers every word, as group of words with frequencies [8].

3.2 Data Description

GRI's Sustainability Disclosure Data launched in 2000, till 2018 there are more than 9500 companies representing 164 countries submitted reports on Corporate Social Responsibility (CSR). The data base store the reports from 1999 to July 4th, 2018. We acquired the data from CSR data course (AMI23A) work folder. We are doing content analysis on CSR report of Energy Utilities companies. Overall information about GRI reports submitted by companies is given in excel file. The content of reports are extracted in text files.

3.3 Data Exploration

To understand the content of data set data exploration essential part. Exploration and visualization of acquired data set was started by John Tucky [9]. Data exploration is part of data/text analysis by understanding and transforming data for research problem.

4. TEXT ANALYSIS

4.1 Descriptive Statistics

In this part we discuss the result of data analysis by using graphic information about the reports of corporate social responsibility (CSR). The graph given below in Figures 4.1.2, 4.1.2 and 4.1.3 showing maximum number reports submitted by companies. Performance of Europe and Asia far better than rest of contents. It seems European and Asian are companies are more concerned about customer social responsibility.

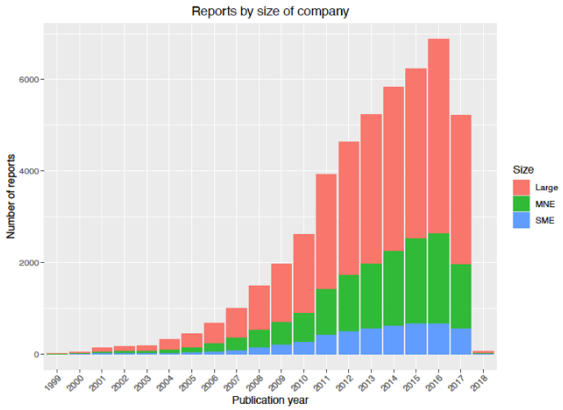

Fig. 4.1.1 Report with respect to size of organization.

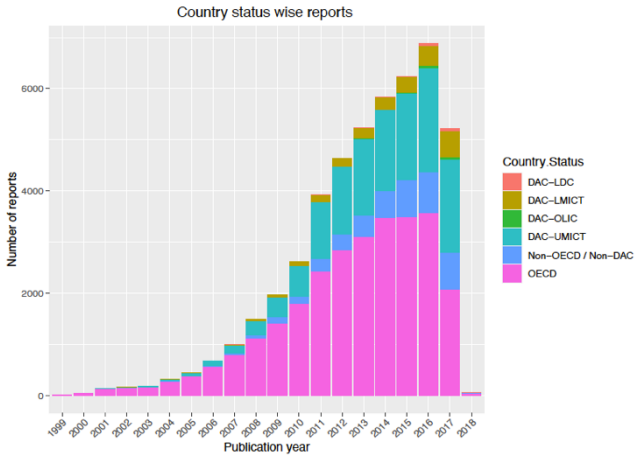

Fig. 4.1.2 Report with respect to the status of country.

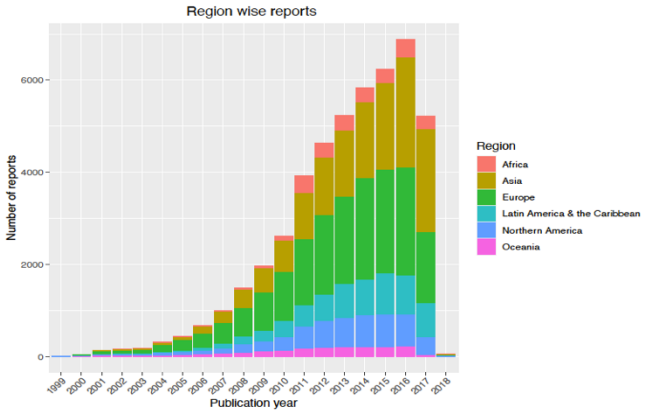

Fig. 4.1.3 Report with respect to the status of country.

4.2 Text Analysis

To extract information from large amount of text, we need to prepare text for further analysis. For this we have to go through following steps: importing text, string manipulation and pre-processing for construction of document term matrix (DTM). Reading Energy Utilities text into RStudio by using VCorpus which can handle data more than 2.5Mb. Once VCorpus is created the other string operations take place like removal of meta data, special character, stop words, numbers and white space. Also, covert the text to lower case and plain text. By doing all these steps we convert text to bag-of-words or document term matrix [10].

4.3 Word Clouds and Frequency

Word clouds is visualization type based on frequencies. In figure (4.1.4) word cloud of most 200 words occurring in bag-of-words. Also, word frequency graph given in figure 4.1.5. From these two graphs one can observe that key term like environment, employee, health, responsibility etc are given importance in reports of CSR published by Energy Utilities companies.

Fig. 4.1.4 Try to show word clouds

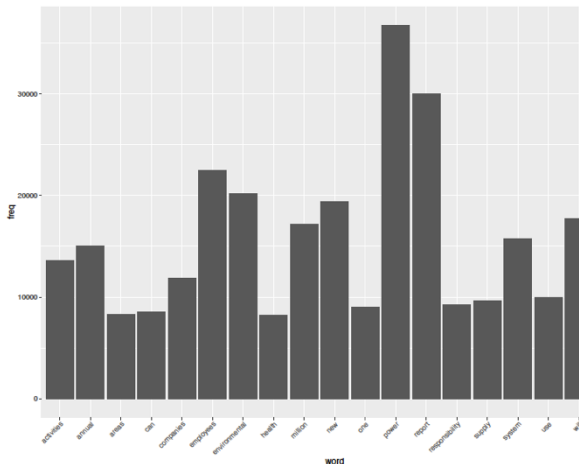

Fig. 4.1.5 Word frequency graph

4.4 Word Association

The word association is similar to correlation and ranges between 0 to 1 excluding negative part. Corporate is one the word from Corporate Social Responsibility and try to explore with the whole of VCorpus or bag-of-words. Graph of association is given in figure 4.1.6

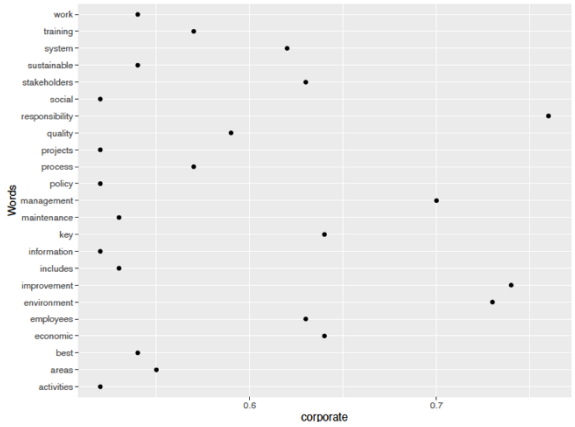

Fig. 4.1.5 Word association graph

4.5 Cluster Analysis

To identify patterns or cluster of words, we are going to use Hierarchical dendrograms. A dendrogram is tree like structure for visualization base on frequency distance. This is technique for information reduction, in other words it leads us to population average. For our problem we are supposed to find any certain patterns form CSR reports. In figure 4.1.6, we can observe six clusters like

1. Cluster (annual, system, activities, million, will)
2. Cluster (new, environment, employee)
3. Cluster (supply, use, one, responsibility)
4. Cluster (health, areas, can, employee)
5. Cluster (work, companies)
6. Cluster (report, power)

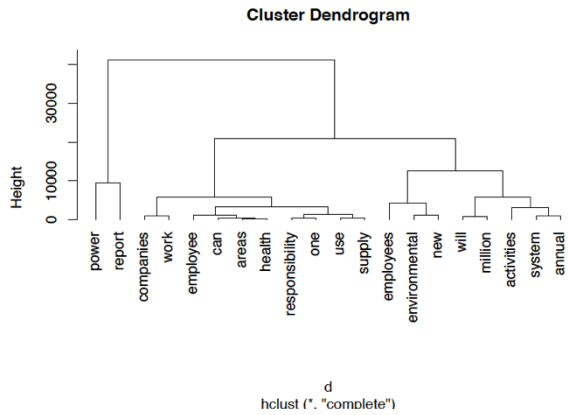

Fig. 4.1.6 Hierarchical dendrogram

5. CONCLUSION

By looking at graphs word clouds, word frequency, word association and dendrogram, all these representing almost same words which hierarchical dendrogram showing us in figure 4.1.6. In concluding remarks, we can say that corporate social responsibility reports of Energy Utilities companies giving emphasis on environment, employee, health, responsibilities, will, activities etc. From the data we can't decide that organization's policies contributed towards well-being of stakeholders except their profit. For future study we should look into corporate social responsibility (CSR) from the perspective each stakeholder separately.

References

[1] Garriga, E.; Melé, D. Corporate Social Responsibility Theories: Mapping the Territory. J. Bus. Ethics 2004, 53, 51–71.

[2] Tate, W. L., Ellram, L. M., & Kirchoff, J. F. (2010). Corporate social responsibility reports: A thematic analysis related to supply chain management. Journal of Supply Chain Management, 46(1), 19–44. http://doi.org/10.1111/j.1745-493X.2009.03184.x

[3] Liu, Sheng-Hung, Sih-Yu Chen, and Sheng-Tun Li. "Text-Mining Application on CSR Report Analytics: A Study of Petrochemical Industry." 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI). IEEE, 2017.

[4] Boli J, Hartsuijker D (2001) world culture and transnational corporations: sketch of a project. International Conference on Effects of and Responses to Globalization, Istanbul

[5] Roca, L.C.; Searcy, C. An analysis of indicators disclosed in corporate sustainability reports. J. Clean. Prod. 2012, 20, 103–118.

[6] Weber M (2008) The business case for corporate social responsibility: a company-level measurement approach for CSR. Eur Manag J 26:247–261

[7] Boiral, O. Sustainability reports as simulacra? A counter-account of A and A+ GRI reports. Account. Audit. Account. J. 2013, 26, 1036–1071.

[8] Kwartler, Ted. Text Mining in Practice with R, John Wiley & Sons, Incorporated, 2017, 20-21.

[9] Tukey, J.W.: Exploratory Data Analysis. Pearson, London (1977)

[10] Arnold, T. (2017a). cleannlp: A tidy data model for natural language processing [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=cleanNLP (R package version 1.9.0)