This article was translated by the author and reprinted from the July 2019 issue of the Securities Analysts Journal® with the permission of the Securities Analysts Association of Japan (SAAJ).

Network Structure in ESG Ratings Suggests New Corporate Strategies: Evolving AI Technology to Quantify Qualitative Data

Yukihiro Aiba, CMA

Senior Quantitative Analyst in Quantitative Solution Research Department, Financial Engineering and Technology Research Center, Nomura Securities. Graduated in March 2006 with a PhD from the Department of Physics, Graduate School of Science, The University of Tokyo. Joined Nomura Securities in April 2006. With strong expertise in financial indices development, he has held a senior position in the AI Solution Research Group since July 2017.

Takeshi Ito

Executive Director of Quantitative Solution Research Department, Financial Engineering and Technology Research Center, Nomura Securities. Graduated in March 2008 with a PhD from the Department of Physics, Graduate School of Science, The University of Tokyo. Joined Nomura Securities in April 2008. After holding senior positions with Instinet's Global Trading Research (NYC, US), he has led the AI Solution Research Group since April 2018.

Yohei Ibe

Quantitative Solution Research Department, Financial Engineering and Technology Research Center, Nomura Securities. Graduated in March 2018 with an MS from the Division of Physics and Astronomy, Graduate School of Science, Kyoto University. Joined Nomura Securities in April 2018. He is a member of the AI Solution Research Group.

Contents

Based on environment, social, and governance (ESG) criteria, corporations are evaluated and receive ESG ratings which are largely based on analyst assessments of qualitative data. Therefore, it is considered rather implausible to quantitatively deduce assessment criteria and the areas that corporations should focus on in order to improve their ESG ratings. We have developed a quantitative model to predict a company's ESG ratings assigned by FTSE and MSCI, by systematically collecting corporate disclosure reports and processing them using AI technology. From the wide-ranging disclosure items related to ESG, our model allows us to quantitatively identify areas that each company should strategically prioritize and the information it should disclose. Our goal is that our continuing research will encourage companies to engage in more concrete action and appropriate information disclosure with regard to ESG.

1. Introduction

Recent years have brought numerous reports of efforts to forecast corporate earnings or macroeconomic performance through the use of artificial intelligence (AI) in analyzing text (from news feeds or Twitter, for example), satellite imagery, and voice data. Examples include efforts to forecast sales at retail outlets by applying image recognition AI to satellite photographs so as to calculate parking lot space utilization, and efforts to estimate macroeconomic trends by applying natural language analysis AI to social media postings (Suimon et al., 2015; Yamamoto & Matsuo, 2016; METI, 2017; Kotera et al., 2018; Cabinet Office, 2018; etc.).

The term *alternative data* has become the catch-all term for text data, image data, and other such types of information that fall outside the scope of the quantitative data that the finance industry has made its traditional realm, such as corporate financial information and share price data. Traditional data sources are still fundamental to understanding a given company's present circumstances, but it is thought that alternative data captures corporate information not necessarily reflected in the traditional data in a way that makes it essential to understanding an enterprise's growth potential and sustainability. The accumulated amount of alternative data has been steadily increasing—and its quality improving—in tandem with the leaps made in information technology. At the same time, AI technology has also become more capable and broadly applicable, so the idea of putting this data to use in the financial services industry has become the focus of a great deal of attention.

ESG scores are a nicely illustrative example of corporate assessments based on alternative data. A company's ESG efforts can be judged by means of a comprehensive analysis of the information disclosed on its website, reviews written by customers or employees, relevant news articles, and other such sources of information. Accordingly, the analysts employed by MSCI, FTSE Russell, and other such major ESG rating entities calculate ESG scores by sifting through and analyzing massive volumes of text.

There are three problems inherent to this manual approach. First, it makes ratings data quite costly (the cost problem). Second, the ratings assigned can differ hugely from one rating entity to another (the arbitrariness problem). Third, the assessment processes tend to be lacking in transparency, leaving companies in the dark as to what sorts of undertakings and disclosures would lift their scores (the opaqueness problem).

These three problems could potentially be resolved if it were possible to automatically aggregate and analyze text pertaining to ESG and then make assessments within a coherent framework. We have attempted to address these three problems by collecting a vast amount of alternative data and deploying AI engines tuned for image recognition or natural language processing.

In this paper, we take up the examples of the ESG ratings issued by FTSE and MSCI, and analyze them with an eye to clarifying where companies should strategically direct their ESG efforts (from among the huge range of ESG issues) and which ESG issues should be the subject of proactive disclosure, all from the standpoint of encouraging socially responsible investment through engagement between investors and business companies.

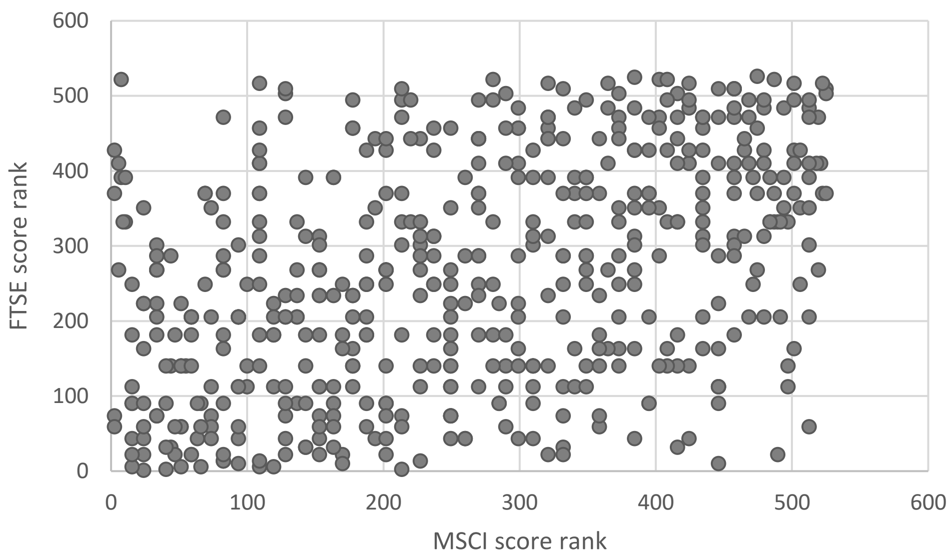
2. Assessment approaches taken by ESG rating entities

This paper deals with the ESG ratings issued by the ESG rating entities FTSE Russell (2017) and MSCI (2017). Both entities' ratings are referenced in the selection of stocks to include in the ESG index that has been adopted by Japan's Government Pension Investment Fund (GPIF) (GPIF, 2017), and are arguably the most referenced ESG ratings in Japan.

FTSE and MSCI reach different conclusions in their assessments

It has been pointed out that the ESG scores provided by FTSE and MSCI often differ considerably from each other (GPIF, 2017). **Figure 1** is a plot of the rank correlations between the two entities' ESG scores as of the end of February 2019, covering 526 Japanese companies for which data was available from both rating entities. This plot makes it readily apparent that the ESG scores issued by one entity can greatly differ from those issued by the other; there are even cases in which a company that earned the highest possible score from FTSE earned the lowest possible score from MSCI (and vice versa).

Figure 1. Relationship between FTSE scores and MSCI scores



Source: Created by the authors using the FTSE and MSCI databases.

ESG assessment approaches are broadly similar

While FTSE and MSCI sometimes come to remarkably different conclusions, the processes through which they arrive at their ESG scores are broadly similar, and look essentially like this:

1. ESG scores are composed of separate scores for the three pillars (E, S, and G)
2. Each pillar is subdivided into multiple themes
3. Each theme is further subdivided into multiple, distinct issues*
4. Scores are assigned for each issue based on the assessor's research

---

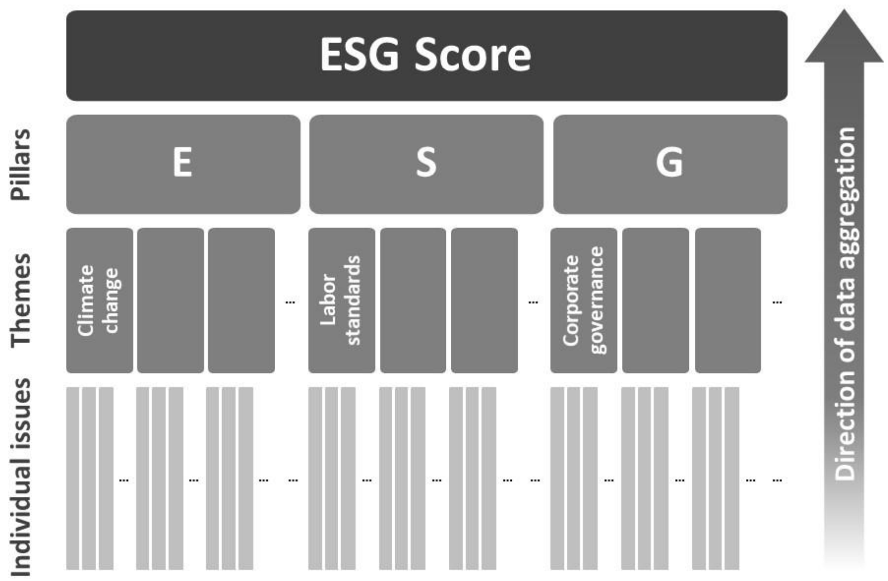
* FTSE calls these "indicators", while MSCI calls them "key issues".

5. The individual scores are added up to arrive at an overall ESG score, with some issues counting for more than others based on pre-determined weights for the company or industry in question

Figure 2 is a conceptual diagram of these pillars, themes, and issues. Information disclosed by companies is the most important component used in determining the issue scores that form the basis of the score calculation. If one were able to accurately estimate the scores that the rating entities assign to these individual issues based on the information disclosed, one would presumably then be able to add them up (in accordance with the rules used by the rating entity) to arrive at an accurate estimate of the overall ESG score as well.

In our analysis, we estimate individual FTSE and MSCI issue scores based on individual companies' disclosures, and analyze the impact of disclosures on ESG scores.

Figure 2. Conceptual diagram of how ESG scores are tabulated



Source: Created by the authors.

3. Giving structure to non-financial information

Companies disclose information about their ESG initiatives in their corporate social responsibility (CSR) reports, integrated annual reports, sustainability reports, or other such report formats. The location of these reports and the form that they take can vary from company to company, so normally, part of an ESG assessor's task is to seek out what are thought to be the most important reports issued by each company and then analyze the content. This requires an enormous amount of time and effort, and the reports selected for analysis and the assessment results can both be influenced by the biases of the person responsible for performing the assessment.

Using the information disclosure framework put together by the Global Reporting Initiative (GRI), we have created a database that makes disparate companies systematically comparable, using the automated collection and collation of reports as a way to remove arbitrariness from the process.

The GRI information disclosure framework

GRI is an international non-governmental organization (NGO) that has put together a framework for the production of the sustainability reports that companies use as a vehicle for reporting on their economic, environmental, and social track records; the organization counts more than one thousand companies and several thousand relevant stakeholders as members (CFA Society of Japan, 2010). According to a 2017 survey conducted by KPMG, the GRI framework is more commonly used worldwide than any other set of guidelines for the production of sustainability reports.

GRI identifies several hundred distinct items about which companies ought to disclose information, and recommends that companies disclose their activities with regard to each of these items in a report format called a "GRI content index". GRI content indices make it possible to see what companies have disclosed on an item-by-item basis, and, because of this, the format allows for systematic comparative assessments of multiple companies.

Collecting and collating GRI content indices

In theory, GRI content indices make for an orderly source of information, but in practice, different companies offer them in different file formats and different table layouts. It is not as though they can simply be gathered up and readily analyzed.

Figure 3 is a snippet from a GRI content index. On the left are the numbers and names of the items determined by GRI; on the right are the locations of where the relevant disclosures can be found (or page numbers pointing to where in a report the information can be found). The table looks quite organized at first glance, but upon a closer look, one discovers that the separate cells for items 102-3 through 102-6 on the left match up with one merged cell on the right. Some companies do the opposite and have single number call-outs on the left matched up with multiple cells on the right, or insert table-like structures within cells. Some companies leave out the table borders entirely, which can make it difficult for conventional image recognition systems to register the information as being a table at all.

The tables presented by companies can differ in other ways as well. Some are presented as PDFs and others in HTML, for example, and the files are located in website tree structures that vary from company to company. Often, the GRI content index is buried in a CSR report that can stretch to several hundred pages. Even with the latest technology available, arranging GRI content indices in a way that makes them amenable to analysis requires a number of creative workarounds.

Figure 3. Snippet from a GRI content index

Index No.	Index name	Related page
102-1	Name of the organization	Corporate Info Form 20-F Item 4
102-2	Activities, brands, products, and services	Form 20-F Item 4 Risk Management System Framework Crisis Management System Framework Approach to Supplier Relations Important Notice
102-3	Location of headquarters	Form 20-F Item 4 Risk Management System Framework Crisis Management System Framework Approach to Supplier Relations
102-4	Location of operations	
102-5	Ownership and legal form	
102-6	Markets served	
102-7	Scale of the organization	Form 20-F Item 6 Employee Data
102-8	Information on employees and other workers	
102-9	Supply chain	Form 20-F Supply Chain Management
102-10	Significant changes to the organization and its supply chain	Form 20-F
102-11	Precautionary principle or approach	Environment
102-12	External initiatives	Ethics and Compliance Approach to Sustainability

Source: Table created by the authors using materials disclosed by a major manufacturer of electronics.

We have independently assembled a system that incorporates a PDF analysis engine, an HTML analysis library, and other tools, and we deploy this in automatically collecting GRI content indices and collating the extracted data. In the section that follows, we delve into the analysis that we run on this structured database of information disclosed by companies worldwide.

4. Using natural language processing to match up disclosure items with assessment items

As discussed above in section 2, if one were able to accurately estimate the scores that rating entities assign to individual ESG issues, one should be able to add them up to arrive at an accurate estimate of the overall ESG score. It occurred to us that the scores for individual issues may be explained by the amount of information disclosed pertaining to that topic. Testing this hypothesis requires matching up ESG assessment items with the corresponding GRI indicators.

The approach that a human would naturally take to this matching task would probably consist of reading the explanations of the GRI indicators and the ESG assessment items and selecting the items that seem most closely related. This would be a very labor-intensive task, however, and it is doubtful that the results would be objective and replicable.

What we have done is to use natural language processing to automatically calculate degrees of similarity between pieces of text, and we have matched up each FTSE and MSCI assessment item with the 10 GRI indicators that show the most similarity to the assessment item.

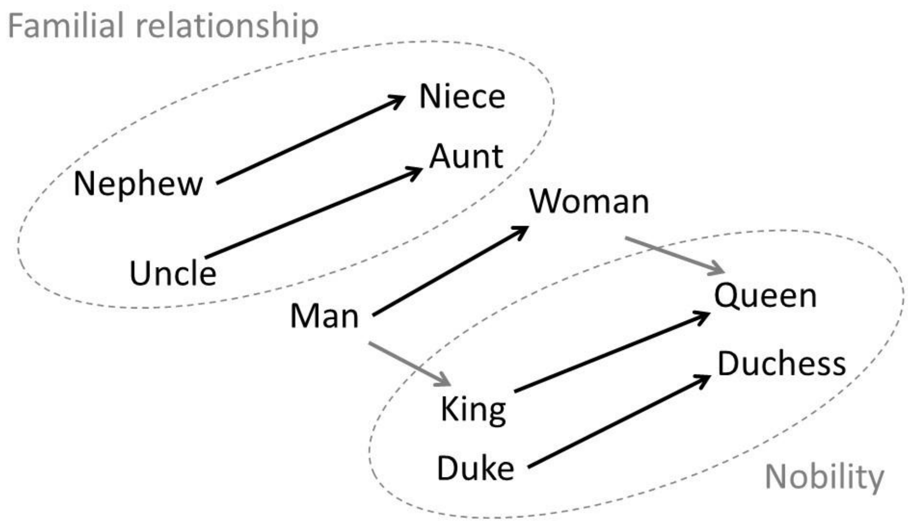
One way to deploy natural language processing in measuring the degree of similarity between two pieces of text is to convert the text into vectors and then treat the degree of similarity in orientation between vectors as an expression of the degree of similarity between the pieces of text. There are already multiple methods in widespread use by which to render text as vectors, and we have selected two of them for use in our research. The first method consists of splitting up a piece of text into its component words, vectorizing each word, weighting each word vector by its importance, and then summing these weighted vectors so as to arrive at a vector for the entire piece of text. The second method consists of vectorizing entire sentences within a piece of text and then summing these sentence vectors. We have chosen to use more than one method of vectorization so that we can test the robustness of our overall approach, aiming to confirm that different methods of vectorization do not ultimately lead us to substantially different conclusions.

1) Word-level vectorization

The algorithm we have selected for word-level vectorization is GloVe (Pennington et al., 2014). Word2vec (Mikolov et al., 2013) is probably the most well-known algorithm for word-level vectorization, but we have chosen to go with GloVe because the corpus used to train the publicly available pre-trained version of the algorithm is rich in ESG-related vocabulary.

Figure 4 is a conceptual diagram of how word-level vectorization works. When words are vectorized, it becomes apparent that the relationship between the vectors for “man” and “woman” (for example) resembles the relationship between the vectors for “king” and “queen”. Also, having words in vectorized form allows for the execution of word-level operations such as “queen = king – man + woman”. We want to draw particular attention to the fact that among groups of words that contain the idea of “man” or “woman”, words related to familial relationships (“nephew”, “niece”, “uncle”, “aunt”) clump together, while words related to nobility (“king”, “queen”, “duke”, “duchess”) form a separate clump. Although **Figure 4** was put together simply for the sake of illustrating the concept, actual calculations done with this algorithm indeed show that groups of words with similar meanings are expressed as vectors with similar orientations.

Figure 4. Conceptual diagram of vectorized words



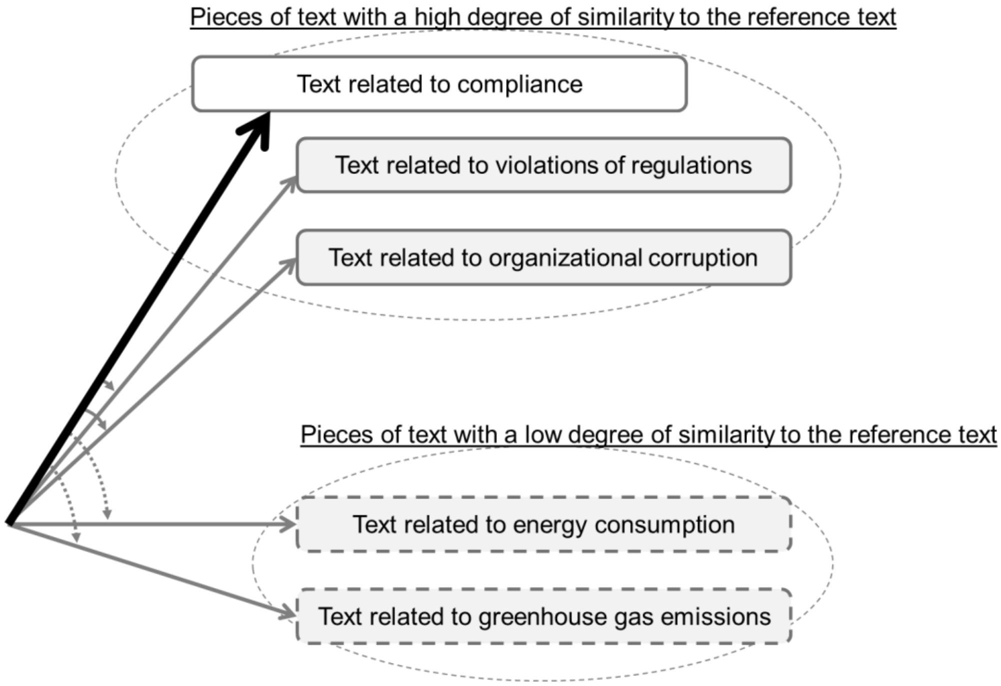
Source: Graphic created by the authors.

By summing the vectorized words, one can vectorize the entire piece of text from which they were taken. However, the thrust of the text will be poorly expressed if one simply adds up all of the words included in the text. This is because the outcome tends to be unduly influenced by the presence of words such as “a” and “the” that appear frequently in documents of all kinds. The usual approach to addressing this problem is to assign degrees of importance to words based on the extent to which they give a piece of text its particular meaning, and then weight those words accordingly when summing them.

In this paper, we use the *term frequency–inverse document frequency* (TF-IDF) algorithm (Jurafsky and Martin, 2000) as our means of weighting words by importance. The algorithm assigns higher TF-IDF values to words that appear frequently in a piece of text, but the TF-IDF values are lowered if the words are generic words that appear frequently in other documents as well. This two-stage filter design ends up assigning higher weights to the words that contribute the most to the sense of a piece of text.

Averaging out the TF-IDF-weighted vectors for all of the constituent words in a piece of text yields an aggregate vector for the entire piece of text. Similarities between two pieces of text, then, are defined as the similarity in orientation between the vectors. This is depicted in **Figure 5**.

Figure 5. Conceptual diagram of degrees of similarity between vectors for pieces of text



Note: The figure depicts how other pieces of text relate to a reference text ("text related to compliance") in their degree of similarity to it. Similar sentences are rendered as vectors with a similar orientation, resulting in relatively acute inter-vector angles. Under this schema, the more acute the angle, the more similar the pieces of text.

Source: Created by the authors.

2) Sentence-level vectorization

The second method we use is to vectorize a piece of text by summing vectors created for entire individual sentences within the text. For this we use BERT (Devlin et al., 2018). BERT, a technology revealed to the world by Google in October 2018, has attracted a great deal of attention for achieving higher degrees of accuracy in multiple natural language processing tasks than the best technologies previously available had managed to achieve.

BERT is seen as having arrived at a deep understanding of context as a result of having been run through two sorts of learning: one in which the algorithm has to fill in the blanks in a text in which words have been randomly hidden, and one in which the algorithm must determine whether a pair of sentences are sequentially connected (**Figure 6**). A feature of BERT is that it can guess at the meanings of words from context—that is, based on what comes before and after.

Using a pre-trained BERT model, we have vectorized the pieces of text we are examining and have measured the degrees of similarity between pieces of text in the same manner discussed above.

Figure 6. How BERT is trained

Task 1: Fill in the blank

System predicts the masked word based on what comes before and after

(1) The man went to _____ store. → Answer: the

(2) He bought a gallon ____ milk. → Answer: of

Task 2: Next sentence prediction

System judges which sentence follows from the first based on how the sentences relate to each other.

(1) The man went to the store.

(2) He bought a gallon of milk. → Answer: ✓ (follows)

(2') Penguins are flightless birds. → Answer: X (does not follow)

Source: Depiction created by the authors, using content adapted from Devlin et al., 2018.

Selection of similar items

Using the two methods detailed above (one at a time), we have calculated the degrees of similarity between GRI indicators and the ESG assessment items used by FTSE and MSCI. **Figure 7** shows actual examples of the GRI indicators that our system has determined most closely resemble certain ESG items, using the first of the two methods described. For the FTSE governance theme "the company has a corporate-wide approach to non-compliance", GRI indicators related to "incidents of non-compliance" and "anti-corruption" ranked as highly similar. Meanwhile, GRI indicators related to "GHG emissions" (greenhouse gas emissions) ranked as highly similar to the MSCI disclosure item "carbon emissions". Thus, the automated calculations of similarity between pieces of text appear to be making appropriate matches.

We have matched up the ESG assessment items in the standards issued by FTSE and MSCI with the top ten GRI indicator hits for each by measured similarity. In **Figure 8**, we draw tree networks representing what our study's approach discovers about how the FTSE and MSCI assessments work. The circles in the GRI indicator layer are proportional in size to the number of assessment items to which they connect, such that disclosure items with larger circles have a greater influence on the overall ESG score. In this way, the approach we have taken to structuring the data allows one to visualize qualitative judgments that have heretofore tended to be understood in only a vague way.

Figure 7. GRI indicators with high degrees of similarity to ESG rating entities' assessment items

a) GRI indicators measured as being highly similar to the FTSE indicator "The company has a corporate-wide approach to non-compliance"

Similarity rank	Disclosure code	Disclosure title
1	416-2	Incidents of non-compliance concerning the health and safety impacts of products and services
2	205-2	Communication and training about anti-corruption policies and procedures
3	417-3	Incidents of non-compliance concerning marketing communications

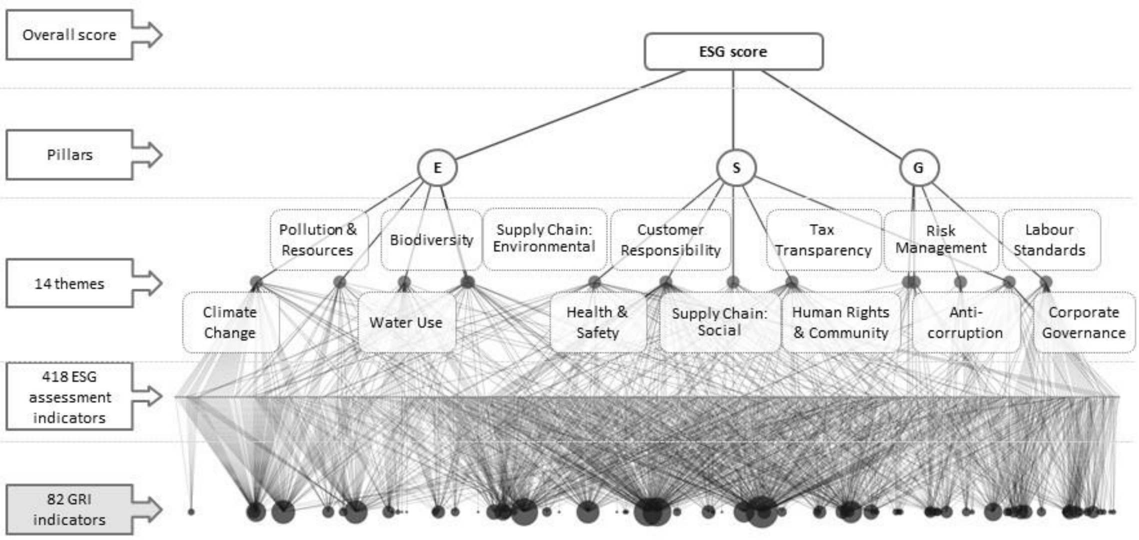
b) GRI indicators measured as being highly similar to the MSCI key issue "Carbon Emissions"

Similarity rank	Disclosure code	Disclosure title
1	305-1	Direct (Scope 1) GHG emissions
2	305-3	Other indirect (Scope 3) GHG emissions
3	305-2	Energy indirect (Scope 2) GHG emissions

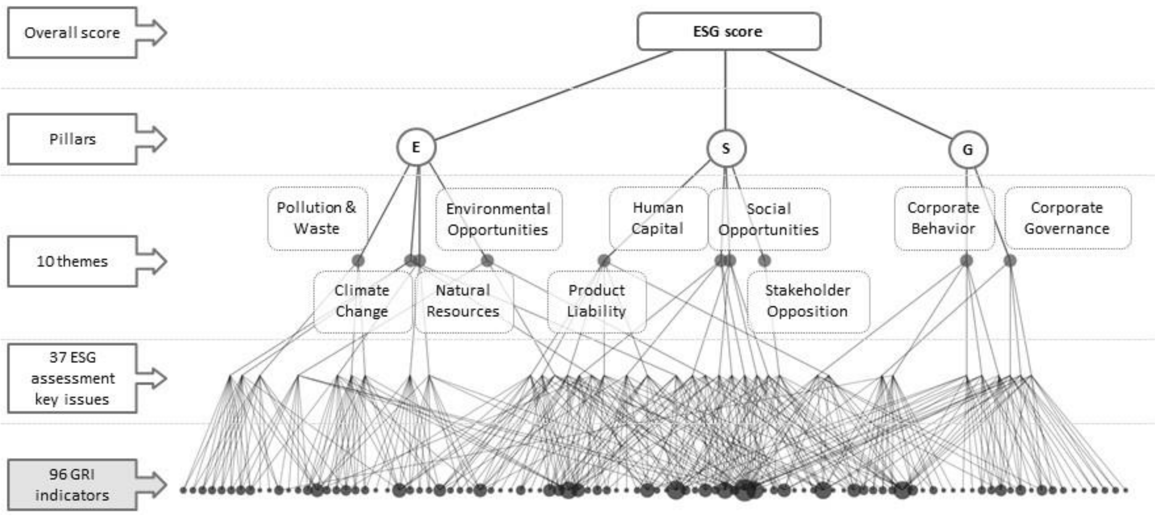
Source: Table created by the authors, using content from the FTSE and MSCI databases and from GRI (2016).

Figure 8. Tree networks for the assessment approaches of the two ESG rating entities studied

a) Structure of FTSE's ESG assessment



b) Structure of MSCI's ESG assessment



Notes: The tree networks shown here are for the assessment approaches of the two ESG rating entities studied, modeled using the approach outlined in this paper. The tree networks depict the network of relationships between GRI indicators and the multi-layered ESG assessment structures. The size of the circle for each GRI indicator is proportional to the number of individual ESG assessment items to which the GRI indicator connects. This gives one a rough idea of which indicators are emphasized most by which ESG rating entity.

Source: Figure created by the authors, using content from the FTSE and MSCI databases and from GRI (2016).

5. Information disclosure factors and ESG scores

Having established connections between ESG assessment items and GRI indicators using the methods spelled out in the preceding sections, we then calculate the fraction of GRI indicators pertinent to each ESG assessment item for which a company has disclosed information, and label this the "information disclosure factor". We use the four-step process spelled out below to work our way up from the information disclosure factors for individual assessment items to an aggregate disclosure factor.

1. Calculate information disclosure factors for each ESG assessment item.
2. Tally these to arrive at information disclosure factors for each ESG theme.
3. Tally these to arrive at information disclosure factors for each ESG pillar.
4. Tally these to arrive at an aggregate disclosure factor.

We have run regression analyses of these aggregate disclosure factors against actual ESG scores, for both the word-level and sentence-level vectorization approaches. The results are shown in **Figure 9**. Using the FTSE ESG scores for Japanese, American, and European companies for which data is available, sentence-level matching with Japanese companies excluded achieves the 5% significance threshold. In the case of the MSCI ESG scores (for which we had only data for Japanese companies available), both word-level vectorization and sentence-level vectorization met the 5% significance threshold. We believe the results show that we have identified a factor that can be effectively used to estimate ESG scores.

Figure 9. Regression analysis of ESG scores

a) FTSE

Region (no. of companies)	Regression coefficient	
	Word-level	Sentence-level
Japan (24)	0.52 **	0.43 *
US (21)	0.58 ***	0.58 ***
Europe (43)	0.40 ***	0.37 **
Total (115)	0.47 ***	0.48 ***

b) MSCI

Region (no. of companies)	Regression coefficient	
	Word-level	Sentence-level
Japan (24)	0.19 **	0.21 **

Notes: Companies analyzed are the set of TOPIX Core 30 constituents (for Japan), Dow Jones Industrial Average constituents (for the US), and Euro Stoxx 50 constituents (for Europe) that have disclosed GRI content indices. Analysis of ESG scores from MSCI is limited to Japanese companies for reasons of ESG score availability. Regression coefficients marked *** meet the 1% significance threshold, those ** the 5% significance threshold, and those * the 10% significance threshold. Corporate disclosures are as of the end of July 2018, and ESG scores are as of the end of February 2019.

Source: Table created by the authors, based on information disclosed on company websites, content from FTSE and MSCI databases, and content from GRI (2016).

Of particular interest is the fact that the measured significance of the regression coefficients does not depend on the ESG rating entity, method used to measure similarity between pieces of text, or nationality of the company in question. We take this as a testament to the robustness of our chosen approach of matching up disclosure items with the ESG assessment items they fit best and then measuring how extensively disclosures had been made.

Concerning Japanese companies specifically, the regression coefficients come out higher for the FTSE scores than for MSCI scores. This suggests that the FTSE ratings place a greater emphasis on *whether* a company has disclosed information. And it is indeed the case that the focus of the MSCI ratings is less about *whether* a company has disclosed information on some particular topic and more about *how well* risks are identified and managed. We find it intriguing that the analysis of text data throws this difference in the qualitative tilt of the two rating approaches into relief.

The weights that ESG rating entities assign to various items when tallying ESG scores differ depending on the company and industry in question. In other words, the tree networks depicted in **Figure 8** are not universal; there are as many distinct variants of these as there are combinations of ESG rating entities and rated companies. The approach that we have chosen makes it possible to take these particulars on board in producing estimates of the impact that each disclosure item has on ESG scores for different rating entities and different companies.

If one were to build a complete model using the approach we have outlined here, one could systematically generate estimates of which items any given company should prioritize in disclosing information for the sake of improving its ESG scores. We think this approach gives companies a path to follow in understanding the nature of ESG rating entities' assessment approaches and making informed, strategic decisions about information disclosure.

6. In conclusion

At a time when intangible assets have reportedly come to account for the greater part of corporate value (Ministry of Economy, Trade and Industry, 2014), anyone attempting to gauge a company's sustainability or growth potential needs to do more than examine traditional quantitative financial data—alternative data needs to be analyzed as well (Kato, 2018).

ESG criteria spring from the idea of incorporating principles of responsible investment into institutional investors' investment decision-making processes, so that companies can be vehicles for both sustainable business and solutions to societal challenges. However, the task of actually rating a company for its ESG performance requires the processing and analysis of large volumes of alternative data. As a result, the ESG scores calculated by major rating entities are both costly to produce and prone to arbitrariness, and the processes by which the scores are arrived at have tended to be opaque.

In this paper, we have made an attempt at estimating ESG scores in a low-cost and consistent way, analyzing large quantities of data using AI and other information processing technologies that have made tremendous gains in terms of usefulness in recent years. In particular, we have tried to identify which groups of GRI indicators (from among the huge number of them that broadly deal with ESG topics) companies should strategically select as priorities for information disclosure.

We expect to see increasingly rapid progress in the development of tools for data analysis built around AI and other such technologies, along with the increasing deployment of such tools in the financial services industry. Our hope is that by conducting research into the ESG assessments that benefit investors and companies alike, we can encourage companies to engage more actively in concrete ESG initiatives and appropriate disclosure of information.

References

Aiba, Y. and Yamamoto, Y. [2018] “Data Science and New Financial Engineering,” Zaikai-Kansoku, 2018 spring issue (in Japanese only.)

Buono, D., Mazzi, G. L., Kapetanios, G., Marcellino, M. and Pappailias, F. [2017] “Big data types for macroeconomic nowcasting,” EURONA - Eurostat Review on National Accounts and Macroeconomic Indicators, pp. 93-145.

Cabinet Office [2018] “FY2018 Annual Economic and Financial Report (Economic White Paper)” (in Japanese only.)

CFA Institute [2008] “Environmental, Social, and Governance Factors at Listed Companies—A Manual for Investors.”

Devlin, J., Chang M. W., Lee, K. and Toutanova K. [2018] “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” arXiv:1810.04805 [cs.CL].

FTSE Russell [2017] “ESG Data Model: 4th Research Cycle (2017/18).”

Jurafsky D. and Martin, J. H. [2000] “Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition (Prentice Hall Series in Artificial Intelligence),” Prentice Hall.

Kato, Y. (ed.) [2018] “Research of ESG Investment: The Forefront of Theory and Practice,” Ittosha (in Japanese only.)

Kotera, S., et al. [2018] “Consumption analysis using POS and text data –application of machine learning- ,” Economic and financial analysis discussion paper series (in Japanese only.)

KPMG [2017] “The road ahead: The KPMG Survey of Corporate Responsibility Reporting 2017.”

METI [2014] “Ito Review of Competitiveness and Incentives for Sustainable Growth: Building Favorable Relationships between Companies and Investors.”

METI [2017] “FY2016 New Industry Model Creation Platform Development Project (New Indicator Development Project Utilizing Big Data) for IoT Promotion Report” (in Japanese only.)

Mikolov, T., Chen, K., Corrado, G. and Dean, J. [2013] “Efficient Estimation of Word Representations in Vector Space,” arXiv:1301.3781 [cs.CL].

MSCI [2017] “MSCI ESG Ratings Methodology.”

Pennington, J., Socher, R. and Manning, C.D. [2014] “GloVe: Global Vectors for Word Representation,” Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1532-1543.

Suimon, Y., Kinoshita, T. and Yamamoto, Y. [2015] “Using AI to quantify economic sentiment: Development of Nomura AI economic sentiment index,” Nomura Japan macroeconomic insight.

Yamamoto Y. and Matsuo Y. [2016] “Sentiment Summarization of Financial Reports by LSTM RNN model with the Japan Economic Watcher Survey Data,” The 30th Annual Conference of the Japanese Society for Artificial Intelligence, 2016.