The Journal of Values-Based Leadership

Volume 15
Issue 1 *Winter/Spring 2022*

Article 11

January 2022

# Codes of Ethics: Extending Classification Techniques with Natural Language Processing

Zachary Glass
zglass@alumni.princeton.edu

E. Susanna Cahn
*Pace University, Lubin School of Business*, ecahn@pace.edu



## Recommended Citation

Glass, Zachary and Cahn, E. Susanna (2022) "Codes of Ethics: Extending Classification Techniques with Natural Language Processing," *The Journal of Values-Based Leadership*: Vol. 15 : Iss. 1 , Article 11.
DOI: https://doi.org/10.22543/0733.151.1366
Available at: https://scholar.valpo.edu/jvbl/vol15/iss1/11

# Codes of Ethics:

## *Extending Classification Techniques with Natural Language Processing*

**ZACHARY GLASS**
**NEW YORK CITY, NY, USA**

**ELLEN SUSANNA CAHN**
**NEW YORK CITY, NY, USA**

*Abstract*

*Language is an indicator of how stakeholders view an ethics code's intent, and key to distinguishing code properties, such as promoting ethical-valued decision-making or code-based compliance. This article quantifies ethics codes' language using Natural Language Processing (NLP), then uses machine learning to classify ethics codes. NLP overcomes some inherent difficulties of "measuring" verbal documents. Ethics codes selected from lists of "best" companies were compared with codes from a sample of Fortune 500 companies. Results show that some of these ethics codes are sufficiently different from the norm to be distinguished by an algorithm — indicating as well that lists of "best" companies differ meaningfully from each other. Results suggest that NLP models hold promise as measurement tools for text research of corporate documents, with the potential to contribute to our understanding of the impact of language on corporate culture and enhance our understanding of relationships with corporate performance.*

## Introduction

Codes of ethics are written documents; their language is meant to influence both internal and external stakeholders and to convey various understandings of what is right and wrong (Winkler, 2011, p.654). The Sarbanes-Oxley Act of 2002, Section 406, requires an ethics code for senior officers. Businesses' ethics codes have since become common. Some corporations have ethics codes, others have codes of conduct, for employees, directors, and officers. Harris (2004) suggests that codes of conduct and codes of ethics are, by nature, different. He maintains that conduct and practice are linked with objective outputs, while principles and ethics are associated with justice and character. In this paper, we do not distinguish between codes of conduct and codes of ethics, treating both as "written, distinct, formal document[s] which consist of moral standards which help guide employee or corporate behavior" (Schwartz, 2002, p. 28).

Do ethics codes create an organizational standard that promotes consideration of ethics in decision making? Do ethics codes protect the company from litigation and control

compliance with company policies? Scholars struggle to work out which of these opposing possibilities, ethical-valued decision-making versus code-based compliance, is generally true. Ethics code language is not easily quantified or classified.

Corporate Social Responsibility (CSR) research on ethics codes often measures whether a code is present or absent, while studies of the impact of codes on corporate performance have had mixed results (Kaptein & Schwartz, 2008). Mixed results in studies on the impact of codes on corporate governance have led to search for a reason and a deeper examination of differences among ethics codes that, in turn, may be related to different impacts on corporate behavior. Presence of an ethics code alone may be insufficient as an indicator of corporate behavior. Perhaps mixed research results occur because ethics codes differ in meaningful ways. Consequently, some codes may have an impact while others do not; in the aggregate, results are mixed. Since codes are written documents, language is key, but measurement of language is difficult. This paper addresses the question of whether ethics codes differ by using a Natural Language Processing algorithm. The purpose is to take an important step in developing a potentially useful quantitative tool that can contribute to our understanding and analysis of corporate documents.

Our algorithm is distinctive in that it classifies ethics codes using Natural Language Processing to quantify the text data. As a proof of concept, we demonstrate that NLP can be used as a measurement tool for ethics codes. Advantages of quantitative, computational models include explicitness, known assumptions, and repeatability. Then again, quantitative models also have the shortcoming of being less nuanced than subjective judgment and may leave out information that is difficult to measure. Nevertheless, adding an unbiased objective model to the toolkit adds value by supporting intuition or by challenging intuitive assumptions.

The algorithmic measure of code content we develop is next used to test whether code language is associated with corporate behavior. We address the question: Do ethics codes make a difference in classifying corporate behavior as more ethical than the norm?

In the sections that follow, we review the literature on normative approaches to ethics codes, followed by a review of research describing how ethics codes appear in practice. Next, we consider research exploring the impact of ethics codes on corporate performance. We then focus on language differences in ethics codes and efforts to measure code differences. The NLP model we develop adds a novel quantitative measurement tool to those available for ethics code research. Finally, we use machine learning to classify the scored ethics codes into two categories which we have labeled *Ethical* and *Normal*.

## Normative Approach to Ethics Codes

Good ethics codes are written based on the principle that ethics is about right, as opposed to wrong, values and behavior. “Codes of ethics are intended to capture the key values of a firm and to convey those values to both internal and external stakeholders” (Coughlan, 2005, p. 45). Weaver (1993, p. 45) sees ethics codes as constraining behavior. He defines ethics codes as “distinct, formal documents specifying self-consciously ethical constraints on the conduct of organizational life.”

Reynolds and Bowie draw upon Kant’s moral principles as “an externally-established conception of what is right” (2004, p. 276). Kant considered moral principles of what is

ethically right to be independent of context. Following the Kantian framework, Reynolds and Bowie maintain that ethics codes should have the primary motive of doing what is right. Codes should respect the free will of individual employees, avoiding retaliatory language. The code should be written to be valuable to all employees, so organizations should provide opportunities for all members of the organization to contribution to it. "Employees need discretion in applying the policy, but they also need to be able to suggest changes and improvements in the policy. In so doing, the employees are exercising their rational and moral capacities. By actively participating in this way, their own ability to make better moral decisions is increased." In contrast, if measurable outcomes are emphasized rather than values, then codes tend "to legitimize the legalistic and symbolic benefits of an ethics program at the expense of the inherent value of moral behavior…. An ethics program that is adopted simply to support the bottom line will not have the best consequences" (Reynolds & Bowie, 2004, pp. 276-283). Since ethics code writers cannot anticipate future dilemmas, the point of the code is to guide decisions with value statements. If the code is written as a set of laws with punishments, employees will be motivated to adhere to the letter of the law to avoid punishment, rather than be thoughtful and adhere to the spirit of the value system.

Harris (2004) also advocates participation of stakeholders in the process of developing and implementing codes. He makes the case that ethics codes should be future-oriented, developing good habits, building trust, and encouraging decision-making based on principles. Language looms large, giving importance to narrative about principles and values along with objective quantitative measures. An ethics code can serve as one of the public pronouncements of espoused values of the corporation (Schein, 2016, p. 4). It becomes a reference for corporate behavior and choices.

Since the future is uncertain, ethics codes highlighting values and principles rather than attempting to address a list of potential scenarios, are more useful. They provide employees who are faced with ethical dilemmas a basis to justify the choices they make. Therefore, "if a code is meant to provide justifications for employees, it must specifically address important values" (Coughlin 2005, p. 48).

## Descriptive Approach to Ethics Codes — *Values versus Compliance*

In actual practice, the language of corporate ethics codes may or may not conform to normative ideals. Additionally, the actual reasons for introducing a corporate ethics code may be different from those announced with the code's introduction. At times the distinction between company policies and its ethics code is arbitrary. Weaver (1993) suggests that social desirability biases may lead to ambiguity in identifying which corporate documents are to be considered ethics codes. Managers may perceive the company's ethics code differently from employees. There may be multiple perspectives arising from individuals' multiple roles.

Farrell and Farrell (1998, p. 588) describe ethics codes as being either inspirational or prescriptive. Inspirational codes "in which code writers provide corporate values/principles only" thereby leave discretion in the application of those values to employees addressed by the code. Prescriptive codes develop expectations of employees for compliance; they "arise when code writers apply … corporate values and principles to perceived moral hazards that might occur … No discretion in the matter is expected." They analyzed the language in a small sample of Australian corporate codes of ethics examining linguistic structures of relational clauses, passive voice, nominalization, grammatical metaphor, and modality. They

concluded that the codes in their study primarily imposed conformity to rules, using language to maintain a hierarchical power relationship within the organization; they did not empower employees to make ethical decisions.

Winkler (2011) identified three parts of ethics codes: the introduction, the rules and regulations, and the code enforcement. Analyzing ethics codes of Dax30 companies – German blue chips listed at the Frankfurt Stock Exchange – he examined the role given to the actors addressed by the code. Did the codes ascribe any agency to the actors, or did the codes render them as being passive? The code introductions were seen to downplay the existence of hierarchy and asymmetries, literally elevating ordinary employees, in terms of social status and corporate responsibility. Considering the other parts of the codes, however, this initial attribution of agency quickly disappeared. The rules and regulation sections addressed employees as passive receivers of code instructions. "Compliance with the codes of ethics is usually enforced by creating a feeling of fear…the enforcement part of the codes of ethics once again fabricates employees as rather passive actors who are in need of guidance, assistance and control" (Winkler, 2011, p. 659). The codes studied by Winkler created a sense of ambiguity by placing a great deal of responsibility on the employees though denying them agency and competence.

The literature suggests that in practice, ethics codes follow one of two patterns. One pattern emphasizes values and leaves employees to make their own decisions about the ethical course of action when faced with a dilemma. The other emphasizes compliance with specific guidelines directing employees to seek guidance from a supervisor. Sometimes compliance codes offer scenarios of potential ethical dilemmas. Research indicates that more codes fall into the compliance category. In a review article, Babri, Davidson, and Helin (2019) found that the compliance orientation has increased over time.

## Code Impact Literature

Seeking a quantitative impact of corporate ethics codes, a body of literature researches the relationship between ethics codes and corporate performance. A review study by Kaptein and Schwartz (2008) of code impact on Corporate Social Responsibility (CSR) performance found mixed results. This is not surprising considering that ethics codes can differ from one another in a number of ways. Language varies considerably among different companies' ethics codes, potentially contributing to different perceptions and behavior among affected employees. Internal reasons for adopting an ethics code may differ from one company to another. Of course, there are numerous influences on corporate performance. Leaders modeling the way for others to behave is very important (see Kouzes and Posner, 2017, pp. 13-14). Corporate culture and structural features are also important, codes of ethics being the feature examined here. Different ways of measuring code content may contribute as well. Recently, Kaptein suggested myriad possible reasons for mixed research results on code impact: among them are differing topical content and level of prescription among codes, which may contribute to differing code effectiveness (2019, p. 3, 6).

While studies comparing companies with codes to those without have yielded mixed results as to whether a having code makes a significant difference in CSR performance, a few notable studies looked at code content in more detail. Erwin (2011) found a relationship between code quality and CSR performance. That study measured code quality based on

data from Ethisphere, which in turn rates corporate codes using a panel of experts.[1] He contrasted his results to the typical treatment by researchers that compare companies with codes of conduct to those without. Kaptein (2011) studied presence of a code along with the number of issues that are addressed by the code. Content was found to be one of the issues without which having a code could be counterproductive. It follows from the importance of quality and content that the intent of a code may be more important than the simple existence of a corporate ethics code, though "intent" is difficult to measure.

Coughlin's definition of impact follows from his view of an ethics code as a source of justifications for choice, rather than as specifying what choice to make. "A code's usefulness then is not gauged only by its effect on choice, but also by its effect on a decision-maker's justifications" (Coughlan 2005, p.46). He suggests that where laws are inconsistent, heavily legalistic codes are not useful guides for decision makers. Summarizing studies in a review from 2005-2016, Babri, Davidson, and Helin found that codes have both positive and negative outcomes (2019, p. 33).

## Importance of Language

Farrell and Farrell (1998) concluded that language could reinforce a hierarchical power relationship or free employees to be moral decision makers. Examining the language in ethics codes of five large Australian business enterprises, they even found conflicting messages, as employees were addressed as decision makers but then subsequently asked to conform to the hierarchy. Language might be the means for either empowering or constraining.

Béthoux, Didry, and Mias (2007) used software to perform a lexical analysis of a collection of 175 codes from 166 European and North American companies. The software created categories based on the words used frequently in the codes themselves. Their "analysis conveys the idea that codes of conduct are radically inconsistent with workers' participation in the management of the company" (p. 88).

Choice of language can influence code effectiveness. Rodriguez (2010, p. 36) claimed that watered-down language and "weasel wording" was used in some companies' ethics codes, which did not preclude unethical behavior. In those cases, ethics codes could lull investors into a false sense of security, providing assurance of ethical behavior but not actually delivering it. Shin and You (2020) studied the importance of language in CEOs' letters to shareholders. They found that language used affected CEO dismissal risk. Clearly, code language matters.

## Measurement

Measurement of code content or quality is an issue without a consistent solution because of the nature of ethics codes as written documents. Comparative analyses or impact studies often identify ethics codes as only being either present or absent (Kaptein & Schwartz, 2008).

[1] The method grades the codes of conduct from major corporations based on performance in eight categories: "Public Availability," "Tone from the Top," "Readability and Tone," "Non-Retaliation and Reporting," "Commitment and Values," "Risk Topics," "Comprehension Aids," and "Presentation and Style." A specific rating for each category is determined by a panel of experts from the Ethisphere Council. Ratings follow a standard letter grade scale (A = excellent, B = above average, C = average, D = below average, F = poor) (Erwin, 2011, p. 538).

A few researchers do more than simply note the presence or absence of a code. Gaumnitz (2004) measured code content by dimensions of length, focus, level of detail, thematic content (topic or topics), shape (breadth of theme coverage), and tone (positive vs. negative). He concedes that some professional judgment is involved in these measures. Lere and Gaumnitz (2007) expanded on these measures of code content to include disincentives to choose unethical alternatives. Farrell and Farrell (1998) studied linguistic structures of relational clauses, passive voice, nominalization, grammatical metaphor, and modality to distinguish between inspirational and prescriptive code intent. Preuss (2009) used content analysis of ethics codes to measure the frequency of topical coverage.

Harris debated explicitness in codes. "Objective and quantifiable measures are widely seen as essential if voluntary codes are to achieve community acceptance, fairness, and compliance. For those outside the organization, such measures may assist in the building of trust in the intention of the organization to implement the code and in its capacity to do so" (Harris, 2004). There is, however, potential danger that such measures will create false confidence in external stakeholders. Internally, overreliance on quantifiable measures avoids responsibility for consideration of values. Paradoxically, rigid rules may free employees to behave unscrupulously in the grey areas.

Erwin (2011) measured code quality based on benchmarking analyses by the Ethisphere Institute. In effect, an expert panel, opinion-based grading system takes into account various categories of ethical values to create the benchmarks. Kaptein (2011) measured content by the number of issues addressed by a code. Respondents were asked whether or not an issue was addressed in their own organization's code. This measure could be influenced by respondent perception. Wording and tone were not addressed by the survey, and appropriateness was addressed only indirectly.

All of these measurement approaches share the variability of human perceptions; people read ethics codes and come to subjective conclusions. Subjective measures, such as intentions, may be distorted by social desirability biases (Weaver, 1993).

## Natural Language Programming and Machine Learning

Recent and ongoing development of NLP promises to open text data to quantitative measurement. Basically, the idea is to use computer algorithms to find quantitative measures of text. NLP models, such as the model used here, build documents from words and score the documents. Machine learning can then be used to classify the documents, and take algorithmic action based on that classification. All models simplify reality to facilitate analysis. The type of model used here, while not all-encompassing, has yielded good results in a variety of cutting-edge applications where a more fully featured description of language would be too complex. One such instance is spam detection, where state-of-the-art systems rival humans in accuracy.

One of the earliest uses of software, rather than observation, to study corporate documents was a lexical analysis of codes and framework agreements done by Béthoux, Didry, and Mias (2007). Recently, text analysis using algorithms has been used by a few accounting researchers. For example, Baier, Berninger, and Kiesel (2020) use text analysis of annual reports. They algorithmically develop word frequencies to judge the environmental, social, and governance (ESG) content of annual reports. NLP is a tool still being developed. Ongoing

work in NLP aims to develop ever more sophisticated models that will be able to capture sentiment as well as naïve language (Peldszus & Stede, 2016; Stede, 2016).

In this paper, we add the use of NLP to earlier measurement devices as a tool for quantifying ethics codes. Earlier scholars have noted that ethics codes may be inspirational or prescriptive, values-based or compliance-based, moral decision-enabling or constraining. We start with the most basic measurement question: Can we measure differences among ethics codes? In particular, we are looking for a quantitative, objective measure. We go on to use those NLP measures to classify companies by their ethics code scores, assigning companies to an *Ethical* group or a *Normal* group. Then we test the predictive ability of our NLP algorithm by comparing the NLP classifications with a priori company classifications that used other measurement mechanisms.

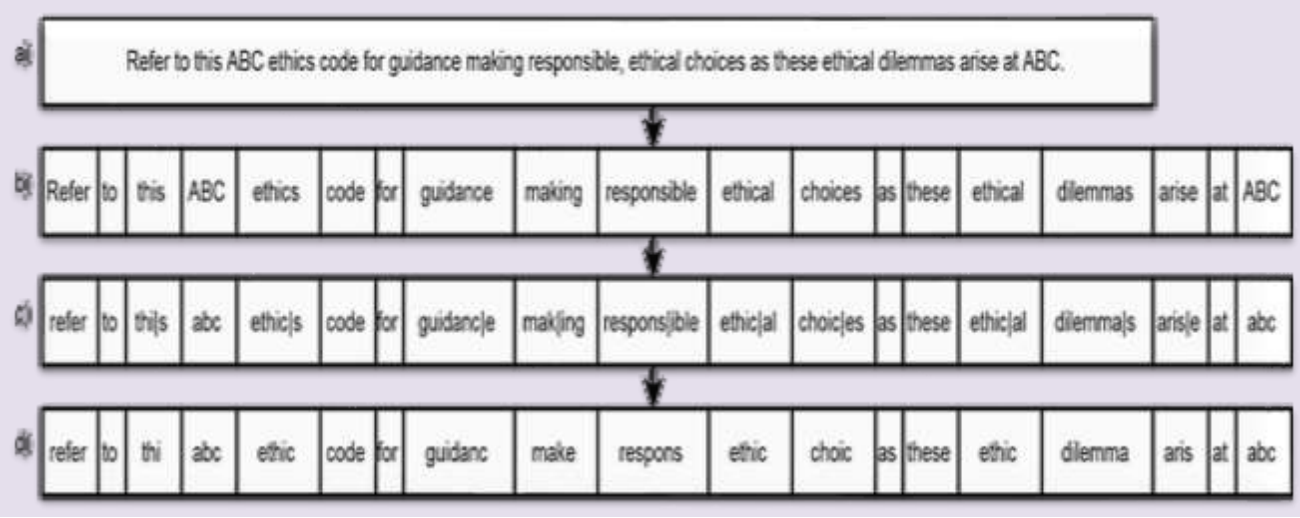
## NLP Model

We use NLP as a tool to develop an objective, quantifiable measure; then we apply machine learning to the problem of distinguishing or classifying the intent of ethics codes by this measure of their language. The first step of the process is to create a model that is representative of language, more specifically of ethics codes as documents.

### *What is a Word?*

This question is trivial to a human, but to a computer, every text is just a string of characters. In our program, word boundaries are marked by spaces. Since a main method will be automated word counting, once words are split by whitespace, word variations will be collapsed together by transforming all text to lowercase and *stemming* them to remove inflections. The stemming process used here is that defined by Porter (1980). See *Figure 1* for an illustrative example.

***Figure 1*: Stemming Example**



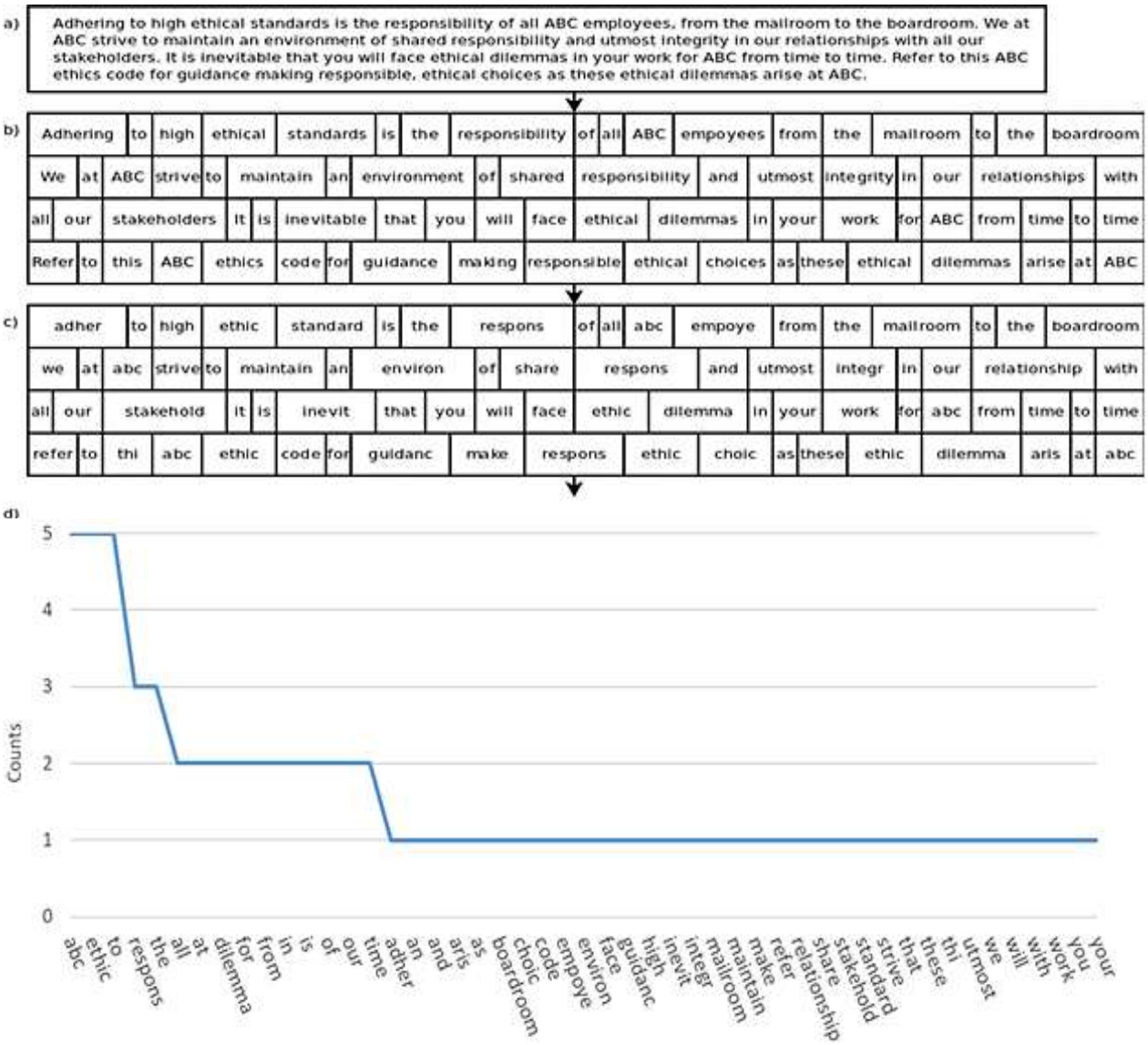

### *How do Words Form a Document?*

The next step in constructing the model is to define larger structures utilizing the "bag of words" model. It has the advantage of being highly descriptive in practice while also being straightforward. As the name implies, the bag of words model ignores sentence structure as if all the words of a document had been placed into a bag and shaken up. An observer could

pull individual words out of the bag, but would have no idea how those words once fit together. However, there is still a tremendous amount of information, especially about the document's topic, hidden in the frequency of words that the author chose (for example, see *Figure 2a)*. When transformed into a bag of words, the sentences yield the set of frequencies seen in *Figure 2d*. Words like "ABC," "ethics," "responsibility," and "dilemma" appear more frequently than most other words. An observer without prior knowledge of the underlying sentences could reasonably conclude that the document relates to ethical decisions concerning ABC company.

## *How is the Content of a Document Measured?*

***Figure 2:*** **Bag of Words Example**




Computer algorithms need to quantify each measurement, unlike humans who make qualitative distinctions intuitively. The frequency counts of words are the starting point for our model. Our model additionally penalizes words based on their commonness in the corpus — the full collection of documents under consideration. The weighting scheme utilized here is

called term frequency-inverse document frequency (TF-IDF) (Manning, Raghavan, & Schutze, 2008, p. 100). Mathematically, this is expressed by *Equation 1*. The intuition is that the TF-IDF score rewards words that occur often in the current document, but penalizes those that occur in many documents.

***Equation 1***

| | | |
|---|---|---|
| $TF - IDF(w) = F(w) * \log\left(\frac{1}{DF(w)}\right)$ | **Where:** | ***w* = a given word**<br>***F(w)* = the frequency of word *w* within a given document**<br><br>***DF(w)* = the fraction of all documents in the corpus in which word *w* occurs** |

Returning to the example in *Figure 2*, the words "ABC" and "ethics" occur in very few documents in the Brown corpus (a well-known topically-balanced corpus), while "the" and "all" occur in nearly all documents in the corpus. Reading from *Figure 3*, reweighting under TF-IDF the scores of "ABC" and "ethics" become much larger relative to "the" and "all." This outcome is in keeping with the descriptiveness of the individual words.

***Figure 3*: Word Frequency Scores Example**

| Word | TF (Term Frequency) | DF(Document Frequency)* | TF-IDF |
|---|---|---|---|
| ***Abc*** | 0.06579 | 0.004 | 0.15776 |
| ***Ethic*** | 0.06579 | 0.05 | 0.08559 |
| ***To*** | 0.06579 | 1.0 | 0.0 |
| ***Respons*** | 0.03947 | 0.32 | 0.01953 |
| ***The*** | 0.03947 | 1.0 | 0.0 |
| ***All*** | 0.02632 | 0.982 | 0.00021 |
| ***At*** | 0.02632 | 1.0 | 0.0 |
| ***For*** | 0.02632 | 1.0 | 0.0 |
| ***From*** | 0.02632 | 1.0 | 0.0 |
| ***Dilemma*** | 0.02632 | 0.036 | 0.03799 |
| ***Utmost*** | 0.01316 | 0.014 | 0.02439 |
| ***Code*** | 0.01316 | 0.056 | 0.01647 |
| ***…*** | | | |

*** Document frequency using the Brown corpus**

## Classifier

### *From Comparisons to Classifications*

The model developed thus far creates a method of quantifying and comparing documents. We take the TF-IDF measure one step further by developing a methodology for making predictions based on those quantified comparisons. Consistent with the argument made by code content researchers above, we hypothesize that there are two classes of ethics codes based on intent of their language. The first class we call *Ethical*, and a second more aggregated group from a background population we call *Normal*.

Published lists of select companies are used as proxy measures for CSR behavior. Relying on a third-party source also reduces the risk of confirmation bias. The lists chosen each has

some claim to identifying companies that are more ethical than is the corporate norm. We assume that inclusion in such lists is an indicator of the ethical corporate behavior. We use these lists in conjunction with NLP to test whether that laudatory corporate behavior is associated with differences in ethics code language. The datasets are described below.

## *Data*

Ethics codes from a sample of the largest *Fortune 500* companies are used as the reference benchmark for *Normal* ethics codes. Four alternative data sources are used as reference benchmarks for *Ethical* companies. The assumption noted above is that companies are chosen for selective listings because of corporate behavior that is exceptional in some way. Inclusion in a named list is also a way to create a convenient, one-dimensional composite measure for the inherently multi-criteria nature of ethical decision-making (Cahn, 2014). The companies tested here were included in: *Ethisphere*'s list of most ethical companies, *Corporate Responsibility Magazine's 100 Best Corporate Citizens*, *Fortune's* list of *Most Admired Companies*, and the *Fortune 100 Best Companies to Work For*. These lists each have a long history and a stable definition.

*Corporate Responsibility Magazine Year's 100 Best List* is created as follows. Its research team documents 260 data points of disclosure and performance measurements for the entire *Russell 1000*. The data is from publicly available information and each company is ranked in seven categories: environment, climate change, employee relations, human rights, corporate governance, financial performance, philanthropy, and community support. The *Corporate Responsibility* list is ranked. For the year in which our sample was taken, the top company in *Corporate Responsibility's* list was *Microsoft*.

*Ethisphere's Most Ethical Companies* program honors companies that excel in three areas – promoting ethical business standards and practices internally, enabling managers and employees to make good choices, and shaping future industry standards by introducing tomorrow's best practices today. The *Ethisphere* list is not ranked.

The *Fortune Most Admired* is a ranked list of fifty companies based on corporate reputations compiled by *Fortune* in partnership with *Korn Ferry Hay Group.* Executives, directors, and analysts are asked to rate companies in their own industry on nine criteria, from investment value to social responsibility. A company's score must rank in the top half of its industry survey to be listed. This ranking has been used empirically by Spencer and Taylor (1987); by McGuire, Sundgren, and Schneeweis (1988); by Wartick (1992); and by Mishra and Modi (2016) to measure corporate social responsibility. For the year of our sample, the top company in the *Fortune Most Admired* list was *Apple*.

*Fortune*'s *100 Best Companies to Work For*, is produced by *Fortune* in partnership with *Great Place to Work*. Two-thirds of a company's survey score is based on the results of the *Trust Index Employee Survey*, which is sent to a random sample of employees from each company. This survey asks questions related to employees' attitudes about management's credibility, overall job satisfaction, and camaraderie. The other third is based on responses to the *Culture Audit*, which includes detailed questions about pay and benefit programs and a series of open-ended questions about hiring practices, methods of internal communication, training, recognition programs, and diversity efforts. Some of these metrics reflect ethical management, but as an aggregate measure it is not exclusively about ethics. As the

respondents are from each company, the data may be subject to self-reporting bias. For the year of our sample, the top company in the *Fortune 100* list was *Google*.

The ethics codes were accessed from company websites and were the most current available at the time of this study, ranging from 2013 to 2016. Each dataset of *Ethical* codes is compared to a comparable number of codes from the set of *Normal* codes taken from the *Fortune 500* companies, excluding those from the companies in the corresponding *Ethical* set. Of *Ethisphere's Most Ethical Companies*, 89 were US companies with publicly available ethics codes. *Corporate Responsibility Magazine's Best Corporate Citizens* had 71 US companies with available ethics codes. *Fortune*'s list of *Most Admired Companies* had 48 with available ethics codes. Of the *Fortune 100 Best Companies to Work For,* 75 had ethics codes available. For each analysis, a subset of the largest *Fortune 500 Companies* containing a comparable number of company codes is used as the *Normal* set. For our sample, the top *Fortune 500* company was *Walmart*. In the public imagination "best" and "biggest" do not appear to coincide.

### *How Does an Algorithm Classify?*

The classifier described here quantifies documents in relation to a corpus (see section, *How is the Content of a Document Measured?).* In this case, that corpus is a combination of the set of ethics codes taken from the lists described above. The purpose is to demonstrate language that distinguishes ethics codes from each other.

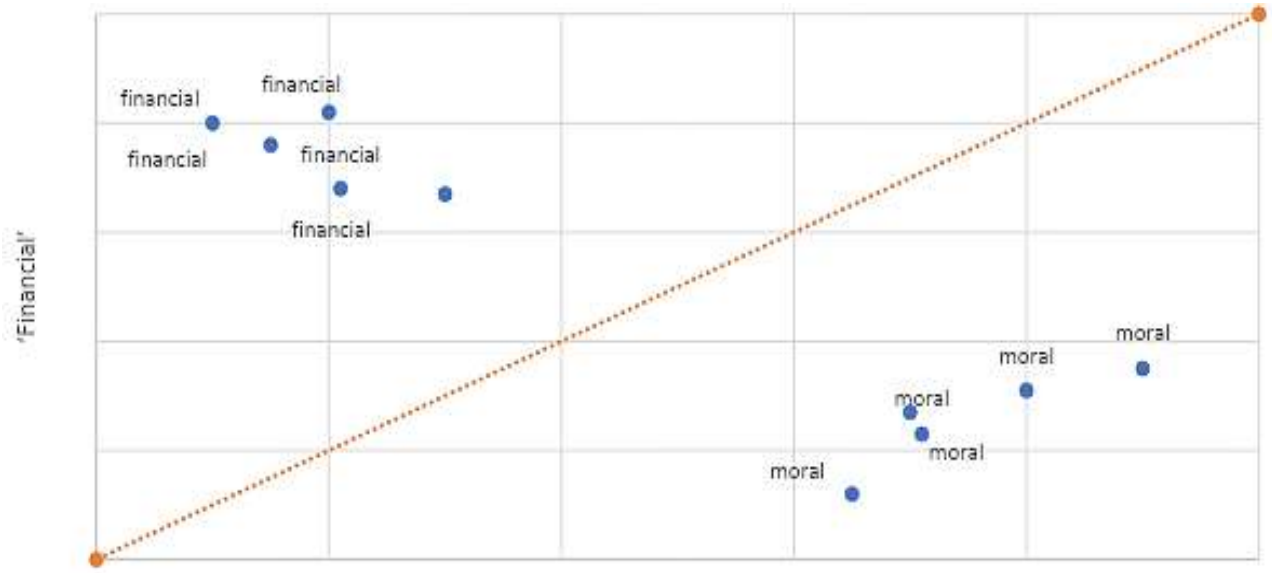
### *Building a Classifier*

In machine learning, classification is the problem of predicting the category of a new observation given previously observed data (and their categories), known as the training set. An individual observation in this problem is the content of a given ethics code, as measured using the model set forth earlier. The categories in our analysis are *Ethical* and *Normal*, known for each document based on the published rankings as explained in the *Data* section. Some of the data is held back from the training set to be used in later testing of the algorithm's predictions. The reason for holding back data is that training and testing an algorithmic classifier on the same observations skews the results.

### *Support Vector Machines*

The particular classifier utilized here is a *Support Vector Machine*. It represents each document as an n-dimensional vector of its TF-IDF word scores, where n is the number of unique words in the corpus. To illustrate, for a corpus with only two unique words (see *Figure 4*), the observations would lie in a 2-dimensional plane. The support vector machine then creates an (n-1)-dimensional boundary which divides the space into two, with one side corresponding to the *Ethical* section of the space and the other side corresponding to the *Normal* section. In the case of the two-word vocabulary of *Figure 4*, this boundary would be a line. The support vector machine chooses a boundary which maximizes the distance between the observations and the boundary. As our data has many more than the two unique words in *Figure 4*, so our results are in a much higher number of dimensions.

***Figure 4:*** **Hypothetical graph of a support vector machine with observed documents containing only the words "financial" and "moral"**




### *Software*

We developed the classifier described in this article in Python. In coding the model and classifier, we utilized mainly the natural language toolkit (Bird, Loper, & Klein 2009) and SciKit-Learn (Pedregosa et al., 2011).

## Results

### *Input Measurements*

In the first step towards building, or *training*, a classifier, each ethics code sampled from each of the *Ethical* listings is measured as compared to *Fortune 500* ethics codes using the TF-IDF measurement described previously. The word "ethics" does not rank highly in this measurement precisely due to this comparison since most ethics codes use the word "ethics." As such, "ethics" is a poor feature for distinguishing one class of ethics codes from another. In contrast, "ethics" is a strong distinguishing feature of ethics codes relative to general English. Words like "ethics" rank highly relative to the Brown corpus, which is representative of general English texts by virtue of including a variety of sources. Intuitively, "ethics" seems reasonably likely to derive from a common set of vocabulary shared among all ethics codes. As such, it is a useful feature in identifying that a document is in fact an ethics code, but not useful for distinguishing *between* ethics codes.

### *Training and Testing*

In training the classifier, each ethics code's measurements are taken to be the features of each observed document and the training label of *Ethical/Normal* is taken from whether the ethics code was drawn from an *Ethical* list or the *Normal Fortune 500*.

The dataset of ethics codes is split into two groups, one (larger) set for training the classifier and the other for testing it. In the training step, the model is fed a set of observed class labels (*Ethical/Normal*) to learn from, along with the measured features of the sample documents. In the testing steps, the class labels are withheld and the now-trained classifier

predicts the class of each document in the testing set. There are two basic summary measures used to gauge the effectiveness of the classifier: (1) how frequently the classifier correctly predicts each document's class is an indication of the overall *precision* of the classifier and (2) the proportion of each document class correctly predicted by the classifier is an indication of the *recall* (or coverage) of the classifier.

The number of ethics codes used here would be considered a small sample size in machine learning methodology. One drawback of a small sample size is that removing any piece of the dataset from training can significantly affect the resulting classifier. Another limitation is that the model may not have observed enough data to accurately predict classifications. For this reason, our analysis uses a technique called *cross-validation* in order to maximize the amount of training data available and find a smoothed estimate of the model's accuracy. In cross-validation, the dataset is divided into N equal groups. The classifier is then trained N times using all but one of the groups, with a different group being left out each time. The average accuracy of the N iterations is taken to be the overall model accuracy.

### *Output Measurements*

Once the ethics codes are divided into the two groups, they are inputted into the machine learning model previously described. *Tables 1* through *4* show the success of the model in classifying a given set of *Ethical* companies' ethics codes relative to the *Normal* set of companies in the *Fortune 500* group. In each case, the *Normal* set contained the same number of companies from the *Fortune 500* group matching the number of *Ethical* codes. For each corresponding *Ethical* dataset, any *Fortune 500* companies which were also in the *Ethical* group were excluded from the *Normal* group.

In *Tables 1* through *4*, the *Ethical* precision, or positive predictive value, reflects the percentage of codes that the model identified as *Ethical* which did in fact come from the set of *Ethical* companies. The *Normal* precision correspondingly represents the percentage of ethics codes that the model identified as *Normal* that did in fact come from the *Fortune 500* list. The complements of the precisions for each set would be false positives, that is codes classified as *Ethical* that were actually *Normal* or codes classified as *Normal* that were actually *Ethical*.

The *Ethical* recall, or sensitivity, reflects the probability of detecting the ethics codes that came from the total set of *Ethical* companies. That is, it is the percentage of correctly identified *Ethical* codes out of all *Ethical* codes. The *Normal* recall is the corresponding percentage of correctly identified *Normal* codes. The complement of the recall percentages are false negatives, that is codes that should have been classified as *Ethical* but were not, or codes that should have been classified as *Normal* but were not.

The F-measure is a weighted measure that includes consideration of both precision and recall. In fact, it is their harmonic mean, which is a useful average when dealing with rates. The key motivation in using both precision and recall (or the F-measure as a convenient combined measure) is that neither is fully indicative of a successful classifier. The best classifier will exhibit both high precision and high recall.

| *Table 1:* **Classification Accuracy for Ethisphere Codes** | | | |
|---|---|---|---|
| | Precision | Recall | F-measure |
| Ethical | 0.61 | 0.59 | 0.60 |
| Normal | 0.56 | 0.59 | 0.58 |
| Average | 0.59 | 0.59 | 0.59 |

## *Performance*

Results for the classifier trained on data from Ethisphere's Most Ethical Companies are shown in *Table 1*. Using this dataset, the classifier correctly identified the *Ethical* set of companies 61% of the time (precision), and managed to correctly identify 59% of all *Ethical* companies (recall). That is, 61% of the codes identified as *Ethical* were in fact from *Ethisphere's Most Ethical Companies* (and 39% were not). Of the codes in *Ethisphere's Most Ethical Companies*, 59% were among those identified as *Ethical* (and 41% were mislabeled as *Normal*).

Similarly, the classifier trained on data from *Ethisphere's Most Ethical Companies* identified the *Normal* set of companies 56% of the time (precision) and correctly identified 59% of all *Normal* companies (recall). That is, 56% of those companies labelled *Normal* were not in *Ethisphere's* dataset (while 44% were). Of the *Fortune 500* codes not in *Ethisphere's* set, 59% were correctly identified as *Normal* (while 41% were incorrectly labeled as *Ethical*). The combined F-measure for the classifier trained on *Ethisphere* data was 60% for *Ethical* and 58% for *Normal*. Since the task at hand requires both accuracy in classification (i.e., precision) and correctly covering as much of the data as possible (recall), these F-measure percentages are the most indicative of the behavior of the classifier. It should be noted that average precision, recall, and F-measure across both categories (*Ethical* and *Normal*) was 59% for this set of companies.

The *Ethisphere* dataset resulted in the best performance for the data we tested. An analogy would be a blind "taste test" where of the four sets of *Ethical* companies we tested, the blind tester (that is, the algorithm) observed a greater percentage of the *Ethisphere* set of companies to be ethical, as measured by the wording of their ethics codes, than for any of the other sets of *Ethical* companies. Compared to a truly blind guess (which would have 50% precision, recall, and F-measure), these results indicate that the algorithm identified key linguistic markers that distinguish codes of ethics written by *Ethisphere's Most Ethical Companies* from the remainder of the *Fortune 500*.

To consider these results from a perspective of significance, take for comparison the proportion of codes that would result if the algorithm were not at all discerning and the classification of codes in the dataset were completely random. Since the dataset had the same number of companies in each group, a random classification would be half *Ethical* and half *Normal.* Testing the significance of the average predictive ability of 59% for the *Ethisphere* data, using a test of proportion in comparison to 50% which would occur if the algorithm were not at all discerning, we found a z-score of 2.4015 which is significant at the 1% level (see *Table 5*).

| *Table 2*: Classification Accuracy for Corporate Responsibility Magazine Codes | | | |
|---|---|---|---|
| | Precision | Recall | F-measure |
| Ethical | 0.55 | 0.56 | 0.56 |
| Normal | 0.53 | 0.52 | 0.52 |
| Average | 0.54 | 0.54 | 0.54 |

*Table 2* shows results for the dataset which included ethics codes from *Corporate Responsibility Magazine's 100 Best Corporate Citizens* as the *Ethical* set. The comparison set of *Normal* codes was from a matched number of companies from the *Fortune 500* with any *Ethical* companies excluded. Using this dataset, the classifier accurately labelled an ethics code as *Ethical* 55% of the time (precision), while incorrectly labelling a *Normal* code as *Ethical* the other 45% of the time. Similarly, the classifier accurately labelled an ethics code as *Normal* 53% of the time (precision), with the remaining 47% being *Ethical* codes which were mislabeled as *Normal*. The classifier accurately identified 56% of the *Ethical* set of companies and 52% of *Normal* companies (recall). The average precision over both *Ethical* and *Normal* for this dataset was 54% and the average recall was also 54%. The combined F-measures were 56% for *Ethical* and 52% for *Normal*, with an average of 54% for both categories. While not as good an outcome as for the *Ethisphere* dataset, these results are still better than the 50%-50% outcome that would be expected from two groups of matched size. As such, the classifier results utilizing *Corporate Responsibility Magazine's 100 Best Corporate Citizens* also suggest that the classifier learned linguistic markers that distinguished the content of these ethics codes. Applying the same test of significance as above, however, the z-score for the dataset of *Corporate Responsibility Magazine's 100 Best Corporate Citizens* was 0.9533 which is significant at only the 17% level (*Table 5*).

| *Table 3:* Classification Accuracy for Fortune Most Admired Companies Codes | | | |
|---|---|---|---|
| | Precision | Recall | F-measure |
| Ethical | 0.50 | 0.49 | 0.50 |
| Normal | 0.49 | 0.50 | 0.50 |
| Average | 0.50 | 0.50 | 0.50 |

Results for the dataset including *Fortune's 50 Most Admired Companies* as the *Ethical* set of codes, shown in *Table 3*, stand in stark contrast with the two previous datasets. The classifier scored a precision of 50% and recall of 49% on the *Ethical* set. Similarly, the classifier scored a precision of 49% and recall of 50% on the *Normal* set. The F-measures for both sets and averages of precision, recall, and F-measure were all 50%. The results here are reminiscent of the coin toss analogy, and appear to be no better than chance. These results suggest that any further work in this area not focus on using *Fortune's 50 Most Admired Companies* as a benchmark. The companies in this list appear to differ from others by markers beyond the scope of the linguistic features available to the current classifier within their ethics codes.

| *Table 4:* **Classification Accuracy for Fortune Best Companies to Work for Codes** | | | |
|---|---|---|---|
| | Precision | Recall | F-measure |
| Ethical | 0.47 | 0.47 | 0.47 |
| Normal | 0.53 | 0.53 | 0.53 |
| Average | 0.50 | 0.50 | 0.50 |

*Table 4* shows the results for the dataset where the *Ethical* set of companies is taken from *Fortune 100 Best Companies to Work For*. Of the codes classified as *Ethical*, only 47% were actually from the *Ethical* set (precision). The remaining 53% were *Normal* companies mislabeled by the classifier as *Ethical*. The classifier has a similar recall for *Ethical* companies, correctly identifying only 47% of all *Ethical* companies. The balanced F-measure was likewise 47%. The results for the *Normal* group were a little better. Of the codes classified as *Normal* companies, 53% were correctly classified as *Normal* (precision) while 47% were *Ethical* codes mislabeled as *Normal*. The classifier also correctly recalled 53% of all *Normal* codes and had an overall F-measure of 53%. The average precision, recall, and F-measure are all 50%, indicative of an overall tepid performance of the classifier using this dataset. Any indications the classifier was performing better than chance on the *Normal* class were balanced by worse performance on the *Ethical* class. Overall, the classifier was not appreciably different in results from coin tosses. Similar to the results for *Fortune's 50 Most Admired Companies*, the tepid results for *Fortune 100 Best Companies to Work For* suggest that future work in this area not focus on using this particular ranking of companies as a benchmark.

| *Table 5:* **Tests of Significance of Average Recall** | | | | |
|---|---|---|---|---|
| | Recall Average (accuracy) | z-score | Sig. | n |
| Ethisphere | 0.59 | 2.4015 | 0.0082 | 178 |
| Corporate Responsibility Magazine | 0.54 | 0.9533 | 0.1711 | 142 |
| Fortune Most Admired | 0.50 | 0 | | |
| Fortune Best Companies to Work For | 0.50 | 0 | | |

Comparing the results for all four datasets tested, *Ethisphere's Most Ethical Companies* as well as *Corporate Responsibility Magazine's 100 Best Corporate Citizens* had ethics codes that were distinguishable by the model's algorithm from those of the corresponding *Normal Fortune 500* companies. As a quantitative model, the NLP algorithm shows that the two groups of ethics codes are different, although this model does not identify how they are different.

For the two selective *Fortune* lists, however, neither was distinguishable from the corresponding *Normal* companies' ethics codes by the model. While each of the data sources is a list comprised of companies that have distinguished themselves in some way, *Corporate Responsibility Magazine's* list and *Ethisphere*'s list are distinguishable by the algorithm measuring their ethics codes. The two *Fortune* lists, *Fortune's 50 Most Admired Companies* and *Fortune's 100 Best Companies to Work For*, were not distinguishable in this way.

## Conclusion and Implications

We introduce NLP as a tool for quantitatively measuring ethics code language. Ethics codes may in fact be distinguished based on their wording, as measured by a machine learning algorithm. The results above show this to be true for two of the four datasets tested, for the companies included in *Ethisphere's Most Ethical Companies* and to a lesser extent for those in *Corporate Responsibility Magazine's 100 Best Corporate Citizens* list. Considering the data sources, we conclude that these "most ethical" and "best corporate citizen" companies are laudable in ways that include having ethics codes that are different from the norm. The algorithm used here found that the laudatory corporate behavior reflected by these lists is associated with the language in their companies' ethics codes.

Companies tested here that were judged "best" and "most admired" by *Fortune* have ethics codes that are not distinguishable from the norm by the algorithm. Those companies are noteworthy in other ways, but their ethics codes are typical. Notably though, the tag of "best" has different meanings in the different datasets we studied, as the algorithm found these aggregated lists of noteworthy companies to be different. Why results were different for the different datasets is an interesting question that is beyond the scope of this paper.

Béthoux, Didry, and Mias (2007) suggest that collections of publicly available ethics codes can create a corpus to reference when constructing a new code or improving an existing code (p. 77). Ethics codes, particularly from the *Ethisphere* list of companies coded here as *Ethical,* may serve as such a corpus. The *Ethisphere* codes are measurably different from those of companies listed here as *Normal* large corporations. This measurable difference does not prove causality; we cannot say whether companies that are already more ethical write better codes or whether careful attention to code language improves those companies. But considering the demonstration that the codes are different, together with suggestions of earlier researchers about ethics code tone and intent, we can say that the codes of these *Ethical* companies might serve well as models on which to base writing of a new ethics code.

The implication that objectively measured content of ethics codes can sometimes, but not always, be distinguishable from one company to another indicates that not all ethics code content is the same. Further, the fact that differences studied here relate to inclusion in a list of exceptional companies supports the idea that code content and quality can make a difference in the kind of organizational behavior and performance that inclusion in such a list represents.

Good measurement is key to empirical analysis. Ethical values are notably difficult to measure, particularly because values are espoused by way of text. The ability to algorithmically analyze text therefore has implications for future business and society research. A frequent criticism of corporate language is that what companies say and what they do are not always consistent. NLP can be an objective tool for disentangling these concepts and an important addition to our toolkit. Algorithms will never replace human judgment but they can be valuable tools in a decision support system. The implication that NLP and machine learning can be used as research tools for studying business text documents will contribute to our understanding of the impact of language on corporate culture and enhance our understanding of relationships with corporate performance.

### *Limitations and Future Research*

A limitation of the classification methodology used here is that the TF-IDF score does not provide insight into how the ethics codes of listed companies included on the *Ethical* list differ from those on the *Normal* list. Further investigation into the details of how the recognized *Ethical* codes differ from the *Normal* codes may shed more light on the strategic impact of ethics code language.

Regarding impact studies' research, a quantitative classifier like that used here could be used in future research to measure ethics code "intent." The model results above demonstrate that NLP models can distinguish among companies' ethics codes. Using a quantitative classifier of ethics codes as a measure of intent together with financial data could advance research on the relation between corporate social responsibility and financial performance. Future research might investigate groups of companies whose ethics codes have been classified differently, like those studied here, and compare the relationship of those corporations' actions to the codes to which they say they adhere. More generally, these results suggest that NLP models may hold promise as measurement tools for text research to investigate corporate behavior.

---

## References


Babri, M., Davidson, B., and Helin, S. (2019). An Updated Inquiry into the Study of Corporate Codes of Ethics: 2005–2016. *Journal of Business Ethics*. https://doi.org/10.1007/s10551-019-04192-x

Baier, P., Berninger, M., and Kiesel, F. (2020). Environmental, Social and Governance Reporting in Annual Reports: A Textual Analysis. *Financial Markets, Institutions and Instruments*, *29*(3), 93–118. https://doi.org/10.1111/fmii.12132

Béthoux, É., Didry, C., and Mias, A. (2007). What Codes of Conduct Tell Us: Corporate Social Responsibility and the Nature of the Multi-National Corporation. *Corporate Governance: An International Review*, *15*(1), 77–90. https://doi.org/10.1111/j.1467-8683.2007.00544.x

Bird, S., Loper, E., and Klein, E. (2009). *Natural Language Processing with Python*. Sebastopol, CA: O'Reilly Media Inc.

Cahn, E. S. (2014). Measures of Corporate Social Performance and Ethical Business Decisions: A Review and Critique. *Southern Journal of Business and Ethics*, *6*, 142–152.

Coughlan, R. (2005). Codes, Values and Justifications in the Ethical Decision-Making Process. *Journal of Business Ethics* (Vol. 59, pp. 45–53). https://doi.org/10.1007/s10551-005-3409-9

Erwin, P. M. (2011). Corporate Codes of Conduct: The Effects of Code Content and Quality on Ethical Performance. *Journal of Business Ethics*, *99*(4), 535–548. https://doi.org/10.1007/s10551-010-0667-y

Farrell, H. and Farrell, B. J. (1998). The Language of Business Codes of Ethics: Implications of Knowledge and Power. *Journal of Business Ethics*, *17*(6), 587–601. https://doi.org/10.1023/A:1005749026983

Gaumnitz, B. R. and Lere, J. C. (2004). A Classification Scheme for Codes of Business Ethics. *Journal of Business Ethics*, *49*(4), 329–335. https://doi.org/10.1023/B:BUSI.0000021053.73525.23

Harris, H. (2004). Performance Measurement for Voluntary Codes: An Opportunity and a Challenge. *Business and Society Review*, *109*(4), 549–566. https://doi.org/10.1111/j.0045-3609.2004.00209.x

Jurafsky, D. and Martin, J. H. (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Kaptein, M. (2011). Toward Effective Codes: Testing the Relationship with Unethical Behavior. *Journal of Business Ethics*, *99*(2), 233–251. https://doi.org/10.1007/s10551-010-0652-5

Kaptein, M. (2019). Business Codes: A Review of the Literature. *Cambridge Compliance Handbook 2020*, (August), 1–22.

Kaptein, M. and Schwartz, M. S. (2008). The Effectiveness of Business Codes: A Critical Examination of Existing Studies and the Development of an Integrated Research Model. *Journal of Business Ethics*, *77*(2), 111–127. https://doi.org/10.1007/s10551-006-9305-0

Kouzes, J., and Posner, B. Z. (2017). *The Leadership Challenge : How to Make Extraordinary Things Happen in Organizations* (6th ed.). Hoboken, NJ: John Wiley & Sons.

Lere, J. C., and Gaumnitz, B. R. (2007). Changing Behavior by Improving Codes of Ethics. *American Journal of Business*, 22(2), 7–18. https://doi.org/10.1108/19355181200700006

Manning, C. D., Raghavan, P., and Schutze, H. (2008). *Scoring, Term Weighting, and the Vector Space Model. In Introduction to Information Retrieval*. Cambridge, England: Cambridge University Press, p. 100.

Manning, C. D. and Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, MA: The MIT Press.

McGuire, J. B., Sundgren, A., and Schneeweis, T. (1988). Corporate Social Responsibility and Firm Performance. *Academy of Management Journal*, *31*(4), 854–872. http://dx.doi.org/10.2307/256342

Mishra, S., and Modi, S. B. (2016). Corporate Social Responsibility and Shareholder Wealth: The Role of Marketing Capability. *Journal of Marketing*, *80*(1)(January), 26–46. http://www.jstor.org/stable/43785257

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., and Duchesnay, É. (2011). Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. https://arxiv.org/abs/1201.0490

Peldszus, A. and Stede, M. (2016). Rhetorical Structure and Argumentation Structure in Monologue Text. In *ACL, 103*, 103–112. https://doi.org/10.18653/v1/w16-2812

Porter, M. F. (1980). An Algorithm for Suffix Stripping. *Program: Electronic Library and Information Systems, 14*(3), 130–137. https://doi.org/10.1108/eb046814

Preuss, L. (2009). Ethical Sourcing Codes of Large UK-Based Corporations: Prevalence, Content, Limitations. *Journal of Business Ethics*, *88*(4), 735–747. https://doi.org/10.1007/s10551-008-9978-7

Reynolds, S. J. and Bowie, N. E. (2004). A Kantian Perspective on the Characteristics of Ethics Programs. *Business Ethics Quarterly*, *14*, 275–292. http://www.jstor.org/stable/3857 911

Rodrigues, U. and Stegemoller, M. (2010). Placebo Ethics. *Virginia Law Review, 96*(1), 1–68.

Russell, S. and Norvig, P. (2002). *Artificial Intelligence: A Modern Approach* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Schein, E. H., and Schein, P. A. (2016). *Organizational Culture and Leadership*. New York: John Wiley & Sons, Inc.

Schwartz, M. S. (2002). A Code of Ethics for Corporate Code of Ethics. *Journal of Business Ethics*, *41*, 27–43. https://doi.org/10.1023/A:1021393904930

Schwartz, M. S., Dunfee, T. W., and Kline, M. J. (2005). Tone at the Top: An Ethics Code for Directors? *Journal of Business Ethics*, *58*, 79–100. https://doi.org/10.1007/s10551-005-1390-y

Shin, T. and You, J. (2020). Changing Words: How Temporal Consistency in a CEO's Use of Language Toward Shareholders and Stakeholders Affects CEO Dismissal. *Corporate Governance: An International Review*, *28*(1), 47–68. https://doi.org/10.1111/corg.12302

Spencer, B. A. and Taylor, G. S. (1987). A Within and Between Analysis of the Relationship Between Corporate Social Responsibility and Financial Performance. *Akron Business and Economic Review*, *18*(3), 7–18.

Stede, M. (2016). Argumentation and Discourse Structure: Some Relationships. NLP talk presented at Columbia University, October 7, 2016.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., and Stede, M. (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, *37*(2), 267–307. https://doi.org/10.11 62/COLI_a_00049

Wartick, S. L. (1992). The Relationship between Intense Media Exposure and Change in Corporate Reputation. *Business & Society*, *31*(1), 33–49. https://doi.org/10.1177/000765039203100104

Weaver, G. R. (1993). Corporate Codes of Ethics: Purpose, Process and Content Issues. *Business & Society*, *32*(1), 44–58. https://doi.org/10.1177/000765039303200106

Winkler, I. (2011). The Representation of Social Actors in Corporate Codes of Ethics. How Code Language Positions Internal Actors. *Journal of Business Ethics*, *101*(4), 653–665. https://doi.org/10.1007/s10551-011-0762-8

## About the Authors


### ZACHARY GLASS

**Zachary Glass** holds a BSE from Princeton University in Mechanical and Aerospace Engineering. His research has focused on using natural language processing for domain-specific term extraction. He is also a licensed CPA with an interest in blockchain applications for business.

He can be reached at zglass@alumni.princeton.edu.

### DR. ELLEN SUSANNA CAHN

**Susanna Cahn** is Professor of Management and Management Science at Pace University, Lubin School of Business. She holds a PhD from Columbia University. Her research has focused on ethical decision making and quantitative methods within business ethics.

Dr. Cahn can be reached at ecahn@pace.edu