A System for Analyzing Human Capability at Scale using AI

To be presented at Intelligent Systems Conference (IntelliSys)
7 & 8 September 2023 in Amsterdam

Daniel Zhu[1], Dave Ulrich[2], Sanjiv Das[3,4], and Norm Smallwood[2] *

[1] University of California, Berkeley, CA 94701
[2] The RBL Group, https://www.rbl.net
[3] Amazon Web Services, Santa Clara, CA 95054
[4] Santa Clara University, CA 95053

Abstract. Over 80% of corporate value is now comprised of intangibles, of which a large component is human capability (HC). Reflecting this, the SEC has recently mandated HC reporting requirements (SEC, Q4 2020). We use machine learning to build a prototype system to analyze HC using SEC filings and applied it to 5760 companies. The approach algorithmically generates lexicons for HC concepts, and then applies machine learning to extract the relevant text on HC and business outcomes from annual reports, to create a dashboard for each firm on the quantity of reporting over four dimensions of HC: talent, leadership, organization, and human resources operations. This system links HC reporting to measurable business outcomes such as revenue per employee, earnings, Tobin's Q, and social citizenship. This will enable companies to improve the quality of reporting and governance of HC as well as guide investments in specific areas of HC.

Keywords: human capability, human capital, AI, natural language processing, multi-modal machine learning

1 Introduction

In the United States, more than a third of employees work for big firms[5], and human capital management has become increasingly important as a component of corporate value. It is now possible to use data science techniques to assess human capability, leveraging text and tabular data from regulatory reports. This paper describes a system to do so using multimodal machine learning.

Attention to human capability has increased dramatically in recent years due to contextual challenges around the global pandemic, racial and social injustice, digital and technological advances, political divisiveness, and economic shifts. In this article, we intentionally use the term human capability (HC) rather than

* We are grateful to Bratin Saha for his ideas, advice, and support on making the system in this paper a socially beneficial one.

[5] https://www.wsj.com/graphics/big-companies-get-bigger/

human capital. Our working definition of HC focuses on four pathways: (1) talent, circumscribed with concepts like people, individual competence, employees, HC, or workforce, (2) organization, with concepts like culture, organization ca- pability, agility, systems, or workplace, (3) leadership, including leader, manager, boss, supervisor, and (4) human resources, with concepts like HR practices, de- partments, operations, people, services.

Research has shown the impact of HC activities on firm performance. Huselid [11] found that a one standard deviation increase in the use of high performing work systems was associated with a per employee increase in market value of $18,641, an increase of $27,044 in sales (on a mean sales per employee of $171,099), and an increased cash flow of $3,814. This early work has expanded dramatically and shows that HC improvements deliver financial returns as well Huselid [12].

As knowledge work increases in every economy, HC becomes an important part of corporate investment. HC enhances the intangible value of all companies. The share of intangible assets in corporate value has increased from 17% in 1975 to 84% in 2015 [15]. Tangible physical technology is greatly enhanced by HC around it, for which Tambe et al [21] coined the term "digital capital." A number of studies have shown the importance of HC as a core intangible. Smallwood and Ulrich [19] show how organization capabilities such as agility, culture, innovation, collaboration, and strategic clarity shape shareholder value. Ulrich [22] proposes a leadership capital index to help investors identify leadership qualities that will increase investor confidence. Schneider et al [18] find evidence for the impact of talent or workforce engagement on performance. Ulrich and Brockbank [23] show how the human resource function can deliver value to all stakeholders. Amazon has formally recognized the value of HC by adding a new leadership principle in 2021, to "Strive to be the Earth's Best Employer."

Most of the work showing the impact of HC on business outcomes relies on surveys or work within specific companies. Storey et al [20] summarize this work as showing a positive relationship between HR practices and firm performance across industries and geographies, particularly when HR practices are bundled together to deliver individual competencies, organization capabilities, and leadership. The RBL Group[6] found that survey results with key informant data from over 1200 organizations show that investments in the four pathways of (1) Talent, (2) Leadership, (3) Organization, and (4) HR can be linked to five stakeholder outcomes: employee well-being/productivity, strategic reinvention, customer value, investor performance, and community reputation. The work in this paper presents a system for organization guidance, i.e., to scale these survey- based studies through application of machine learning and AI to SEC reporting, using large amounts of text and tabular data.

The SEC mandated human capital reporting via the Federal Register – Final Rule: Modernization of Regulation S-K Items 101, 103, and 105; Release Nos. 33-
10825; 34-89670; File No. S7-11-19 on November 9, 2020. These rules modernize

[6] https://www.rbl.net/

the requirements of Regulation S-K applicable to disclosure of the description of the business (Item 101), legal proceedings (Item 103) and risk factors (Item 105). This greatly expands HC management disclosures. The wording of the rule is as follows:

> Item 101(c)(2)(ii): Provide "A description of the registrant's human capital resources, including the number of persons employed by the registrant, and any human capital measures or objectives that the registrant focuses on in managing the business (such as, depend- ing on the nature of the registrant's business and workforce, measures or objectives that ad- dress the development, attraction and reten- tion of personnel)."

Because of the breadth and ambiguity in this definition of "human capital," early SEC reporting of HC varies dramatically both in length (ranging from un- der 200 words to over 2000) and in content covered—from safety to unions to broad axioms to specific quantitative data. The reporting of HC content is scat- tered around the annual report, and requires specialized information retrieval.

To help frame the reporting on HC using SEC filings, machine learning can be used to glean an evaluation of the four pathways of talent, organization, leader- ship, and human resources. This enables automated evaluation of all companies that make 10-K filings, which account for 7 to 8 thousand firms per year; thus, HC evaluation may be scaled using machine learning. Further, by standardiz- ing the framework along the four pathways, it is possible to create a system to support how all HC reporting companies standardize their reporting, which may then drive management and investor decisions.

There have been preliminary attempts to use econometric approaches to ex- tend survey methodology and show the impact of HC on firm performance. Guiso et al [10], in an examination of the value of corporate culture, look at S&P 500 companies (from June to October 2011) and show that proclaimed corporate values may be relevant. Notably, firms in which top managers are seen as trustworthy and ethical show strong financial performance, whereas gover- nance structures do not appear to matter. Of corporate web sites, 85% explicitly stipulate some statements about their corporate culture, an important organiz- ing principle of HC. The relation of these statements to corporate performance is tenuous, to say the least. But, Guiso et al [10] find that responses to surveys by employees are more revealing: improvements in reported management integrity scores are strongly correlated with increases in Tobin's Q and a decline in the fraction of unionized workers.[7]

In another study, Li et al [14] also focus on value words to define culture and draw on data from earnings calls between 2001 and 2018. They find that

[7] https://www.investopedia.com/terms/q/qratio.asp

these words correlate with many aspects of business performance including op- erational efficiency, risk-taking, earnings management, executive compensation design, and firm value. In a recent survey of 1,348 North American executives, Graham et al [9] find that 84% of them strongly believe that culture impacts cor- porate value. Popadak [17] constructed an innovative measure of corporate cul- ture at the firm level by utilizing insider reviews from popular online job boards and forums, such as Glassdoor.com and Payscale.com. She measures six elements of corporate culture on an annual basis: adaptability, collaboration, customer- orientation, detail-orientation, integrity, and results-orientation. These are found to be related to firm value. Gorton et al [8] offer a comprehensive survey of the work on corporate culture.

This paper describes an AI/ML system to extend prior econometric work to: (1) create a more comprehensive model of HC as comprised of four dimensions: talent, organization, leadership, and human resources, (2) influence and stan- dardize more effective and transparent reporting of HC activity in corporations that informs executive decisions and investor confidence, and (3) scale studies of HC beyond surveys to large databases that show relationships between HC and employee, business, investor, and community results. For the machine learning field, we demonstrate how machine learning technologies can define the HC field with broader and more accurate definitions of HC and of its impact on business outcomes.

Our system offers a common typology for HC reporting so that companies and investors learn from each other using a common framework and vocabulary. At present, there is extensive debate about what "HC" refers to. For example, some work focuses extensively on HR practice areas. ISO-30314[8] titled "Human resource management — Guidelines for internal and external HC reporting" suggests core HC areas range from general ideas to organization practices to specific metrics. At present the terminology of HC is nebulous. Our system offers word lists to define the breadth and vocabulary of HC and serve as a reference glossary, thesaurus, or lexicons for the HR industry. We then organize this lexicon into the four HC pathways to provide a comprehensive and cohesive framework for the HC industry.

Other studies use employee surveys (e.g., [10], for S&P 500 companies in 2011) or use earnings calls, which vary dramatically by company. With the new SEC regulation, we are able to extract all text in the SEC filings that relate to HC and to have a common corpus for data (SEC filings). Because SEC data is bound by regulatory requirements, it offers a comparable and reliable source of HC information. We have automated the process to collect all 10-K filings and extract and analyze HC text, to generate a report. Using this approach, our algorithm retrieved and culled HC text for the calendar year 2021, for more than 7,000 filings, and after culling some companies on account of missing data, we are able to undertake analysis on 5,760 companies to demonstrate the prototype.

[8] https:/www.iso.org/standard/69338.html

Given this seamless automation, this analysis may be re-run for any period at any time and on other data sources.

The main contributions of the system are described in the following sections. Section 2 describes the benefits of the new system for HC analysis. Section 3 describes the SEC filing data and the downloading and processing of the reported text data at scale. Section 4 discusses how ML is used to extract the relevant HC text from SEC filings that comprise thousands of words. This forms an essential first step in scaling the analysis of HC. Section 5 explains how ML is used to create lexicons for scoring the various attributes of HC. Section 6 fits ML models to the extracted text and scores to link HC reporting to business models. These models may be used to understand what aspects of HC drive outcomes such as revenues, earnings, etc. Engineering details are in Section 7 and concluding discussion is in Section 8.

2 System Implications

This proposed system has implications for both the overall "HR" industry vertical and for individual firms. For the overall industry, the system:

1. Develops a typology for what constitutes "HC" into four pathways. In almost every field, typologies become the foundation for organizing disparate activities and events into accepted categories or patterns: food typologies (four food groups), political typologies (political parties), biology typologies (kingdom, class, order, genus), employee typologies (full time, part time, con- tract), industry typologies (farming, manufacturing, service, etc.). This work provides a conceptual and empirical frame that defines the HR industry. (See the impact of framing in Cukier et al [5]).
2. The research defines the breadth and vocabulary of HC and serves as a reference glossary, thesaurus, or wordlist for the HR industry.
3. Offers an overall measure of HC for SEC (and other) reporting. This overall indicator could become an accepted standard/metric for HC like Tobin's Q for intangibles or Treadway Commission for risk with four risk categories (compliance, strategic, operational, and financial).

For a specific firm, the system:

1. Develops a HC score as a ranking on how a firm compares to the overall sample as well as to the industry, etc. This benchmark score can become part of the firm's overall performance scorecard used by investors, regulators, customers, media, boards, executive teams, investor relations, and internal human resource groups.
2. Enables each company to assess their public reporting and likely internal actions in HC. Business and HR leaders will be able to determine how they perform on each of the four HC pathways. This will help them either [a]

better report what they are doing since they will now have a framework and language to do so and [b] prioritize where they should focus to improve in each of the four pathways.

The system described here has two stages: (1) governance (scoring and reporting of HC for regulators, shareholders, and communities), and (2) guidance on improving HC towards improving business outcomes, of which we focus on:
(i) revenue per employee, (ii) Tobin's Q, (iii) Earnings before interest, taxes, depreciation, and amortization, i.e., EBITDA,[9] (iv) social responsibility based on fraud and litigiousness scoring of firms. This article describes how machine learning is used to implement these two stages.

3 Data

The primary data source for this analysis is 10-K SEC filings. These are annual reports filed by all publicly traded firms as well as private firms that have exceeded a threshold of stock ownership (500 shareholders) and assets ($10 mil- lion) as mandated by the Securities and Exchange Act of 1934. These filings are public record and may be downloaded by anyone freely. We built an API[10] to download the filings in XML and parse them into plain text.

Since being mandated by the SEC, HC reporting has been varied. Some firms created a new section titled "Human Capital" in their 10-Ks, whereas others reported the content in various places in the filing, often in the Management Discussion & Analysis (MD&A) section. Since the HC reporting is not uniform, we cannot just search for and extract a section on HC. Instead, we used a word- based approach to detect the relevant sentences and paragraphs with HC content. We augmented this approach with a machine learning model trained to detect sentences related to HC content. Our HC text extractor attains a high level of accuracy (details in the following section).

Li et al [14] analyze earnings calls to score five attributes of corporate culture: innovation, integrity, quality, respect, and teamwork. We also apply a similar ap- proach with a much broader set of HC concepts. Whereas they score 5 attributes, we score 14 and combine them into the four pathways (Section 5).

4 Human Capability Text Extraction

The 10-K (annual report) filed by companies with the SEC is an extensive document, comprising tens of thousands of words. Within the 10-Ks, since the reporting of HC by firms is varied, we used word-based and machine learning

[9] https://www.investopedia.com/terms/e/ebitda.asp

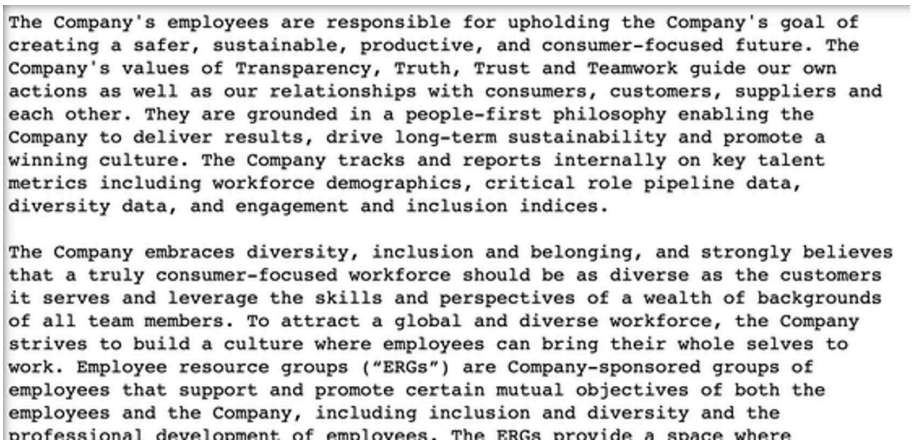
[10] https://sagemaker-jumpstart-industry-pack.readthedocs.io/en/latest/notebooks/index.html

approaches to extract HC related text from the SEC filings. The various approaches are described here.

We extracted sentences containing a preponderance of HC words using a keywords-based extractor. The word lists were generated using an automated algorithm [6] and further refined by human curation. However, manually checking extracted sentences revealed that this method resulted in many false positives.

We then trained a machine learning model to choose sentences in the 10-K filings that are HC related and/or related to business outcomes. This was under- taken with few-shot learning on the 10-Ks from a few companies, from which we manually extracted all sentences that were HC related and consequential business outcomes (the remaining sentences are negative samples). The chosen companies are: Amazon, Applied Materials, BK Technologies, Borg Warner, CEVA, Dell, FCCN, Intel, Interdigital, and Walgreens. This machine learning approach does better and extracts HC sentences with a test accuracy of 88%, with a F1 score of 88.5%, precision of 89.6%, and recall of 87.5%. The trained classifier is then used to extract HC text for all the companies in the sample.

A two-step approach, where we first use the word lists to run a coarse filter on the 10-K filings and extract sentences that are likely to be HC related, does not result in significant reduction in the amount of text that the ML model must process. Thus, our final approach for extracting HC text is the one-step machine learning model. An example of extracted HC text is shown in Figure 1.



The Company's employees are responsible for upholding the Company's goal of creating a safer, sustainable, productive, and consumer-focused future. The Company's values of Transparency, Truth, Trust and Teamwork guide our own actions as well as our relationships with consumers, customers, suppliers and each other. They are grounded in a people-first philosophy enabling the Company to deliver results, drive long-term sustainability and promote a winning culture. The Company tracks and reports internally on key talent metrics including workforce demographics, critical role pipeline data, diversity data, and engagement and inclusion indices.

The Company embraces diversity, inclusion and belonging, and strongly believes that a truly consumer-focused workforce should be as diverse as the customers it serves and leverage the skills and perspectives of a wealth of backgrounds of all team members. To attract a global and diverse workforce, the Company strives to build a culture where employees can bring their whole selves to work. Employee resource groups ("ERGs") are Company-sponsored groups of employees that support and promote certain mutual objectives of both the employees and the Company, including inclusion and diversity and the professional development of employees. The ERGs provide a space where

Fig. 1. Example text extracted using machine learning.

5 HC Lexicons

In this section, we briefly discuss the lexicons used in the project. Using "seed" words drawn from domain expertise, we used the algorithm in Das et al [6] to automatically extract words that are conceptually related to the seed words.

A brief description of the mechanics of this approach is as follows. The user provides a pair of words that are either synonyms or antonyms.

1. If the words are synonyms, we generate two word lists with numerical vector representations of words (embeddings, based on the word2vec algorithm of [16] that are closest to the two words, using the cosine similarity metric on pre-trained word vectors. These word lists are then intersected with a dictionary to keep only the words that are valid in English, and then the algorithm returns the union set of both word lists.
2. If the words are antonyms, we generate two word lists with embeddings that are closest to the two words, intersect these lists with a dictionary to keep only the ones that are valid words, and then return two separate word lists. If a word appears in both lists, then we keep the word only in the list in which it has highest similarity with the concept word.

In short, with synonyms, the algorithm returns a single list (support for the concept) and with antonyms, it generates two lists (support for, as well as against the concept). We generated 14 such word lists using the following seed words: capability, vision, talent, organization, mission, management, leadership, human resources, human capital, employee, develop, culture, competence, agility. These lists were further triaged (using human curation) to construct a final set that was used for scoring.

These fourteen word lists are aggregated into the 4 pathways for Talent, Leadership, Organization, and HR as needed for coarser granularity of HC text scoring. These word lists are assigned to the pathways as follows:

1. Talent = talent + employee + competence
2. Leadership = leadership + management + develop
3. Organization = organization + culture + agility + mission + vision + capability
4. HR = human capital + human resources

Using these word lists, we compute the fraction of the HC text that contains the words in a given list. This operation is compute-intensive and therefore we use special purpose APIs developed in AWS SageMaker JumpStart for the financial sector.[11] These scores are then normalized across the dataset to put each company's score on each attribute into a range from 1 to 10. This scoring

[11] https://sagemaker-jumpstart-industry-pack.readthedocs.io/en/latest/smjsindustry.nlp scorer.html

table permits ranking and filtering companies on one or more attributes, and enables an analysis of where a company stands in relation to others based on their HC reporting. An example of this table is shown in Figure 2.

Show 10 entries Search:

	ticker	Leadership_score	Talent_score	HR_score	Org_score	Overall_score
	All	All	All	All	All	All
1	pcyg	4.4858	3.7181	6.7378	8.625	5.82972367598278
2	tex	5.9434	5.3334	5.2093	6.7054	5.73353912223555
3	amtbb	6.1271	5.4309	5.5815	6.9085	5.95304875870415
4	alrs	6.7215	5.7083	6.829	7.1823	6.56634241046699
5	ttsi	9.5978	3.3675	3.1026	4.2361	4.99360674138334
6	nwyu	6.3357	5.6505	5.0945	6.649	5.87149073715841
7	fccn	6.5017	6.3468	10	6.8238	7.39437843350165
8	gkos	4.644	5.6474	5.7357	4.6383	5.08618273105273
9	mlnd	4.6979	7.4544	4.0824	3.7723	4.91747524735101
10	pkoh	4.1812	5.533	7.9476	7.9684	6.35852838914525

Showing 1 to 10 of 5,760 entries Previous 1 2 3 4 5 … 576 Next

Fig. 2. Table of HC scores. The user can filter this table using the filter template above each column.

6 HC Reporting and Business Outcomes

Does the new reporting mandated by the SEC matter? Does it reflect how corporate value is impacted by HC, and does it help analysts to understand how HC relates to the value of corporate intangibles? To assess this question, we fit machine learning models to the dataset comprising around 5,760 firms. For each firm, we have the four pathway text scores discussed earlier as numerical fea- tures. We also have a column of HC text, extracted using our few-shot trained model that recognizes sentences related to HC. Our machine learning is there- fore multi-modal, yet parsimonious in the number of features (a text column and four tabular columns).

We focus on the following outcomes:

1. Employee: productivity (revenue/employee), etc.
2. Financial: operations, profitability (e.g., EBITDA) or intangible value (Tobin's Q).
3. Community: reputation and social citizenship (e.g., litigiousness scores, fraud scores, etc.)

These outcomes form the labels for our analysis. When the label is continuous, we fit regression models as well as break the outcomes into categories and

fit classifiers. Our models are fit using AWS AutoGluon,[12] which supports the fitting of accurate machine learning models on multi-modal (text plus tabular) data. These are not causal models, but indicate how HC reporting co-varies with business outcomes in the cross-section of firms.

6.1 Revenue per employee

This is a common metric used to assess the productivity of HC. The distribution (in log values) is seen in Figure 3.

The results of the regression model are shown in Table 1. The errors may be assessed against the spread of the distribution above. For the classification problem, we split revenue per employee into 4 quartiles to build a multi-category classifier. The regression and classification models are both stack-ensembled ma- chine learning models. The regression model is not ordinary least squares. The approach ensembles regression versions of ML models such as K nearest neigh- bors, XGBoost, LightGBM (gradient boosted models), CatBoost, Random For- est, Extra Trees, and Neural Networks, etc. More than one of these model forms may be ensembled.

Balanced accuracy is the average of recall across all four classification categories. The Matthews Correlation Coefficient (*MCC*) is a metric[13] that consolidates all values in the confusion matrix into a single score that lies in the range (−1, +1). When the *MCC* is zero, it implies no classification ability. When *MCC* > 0, the model demonstrates classification ability, with *MCC* = 1 being perfect ability. There are several advantages to using MCC [2]. The *MCC* = 0.41, which is evidence of good fit of the classification model. The R^2 = 0.45 is also
better (by ~2x) than studies in this area of work, for example, in comparison to canonical papers such as Combs et al [3]; Crook et al [4]; and Jiang et al [13]. The good fit of this model may partly be attributed to the use of text in a multi-modal model, a new approach in comparison to previous work in this area, where only tabular data is used.

6.2 Tobin's Q (Price to Book value)

This metric is widely used to assess if a firm is undervalued or overvalued. In its pure form, as envisaged by James Tobin, this ratio is market value to intrinsic value, but the latter is not always easy to define and measure, so in practice book value is used in place of intrinsic value. This "market-to-book" ratio proxies for the growth prospects of a company. Hence, it is widely used in forward-looking analyses of corporations. For our sample of firms, we display the Q ratio in logs, shown in Figure 3.

[12] https://auto.gluon.ai/

[13] https://en.wikipedia.org/wiki/Phi coefficient

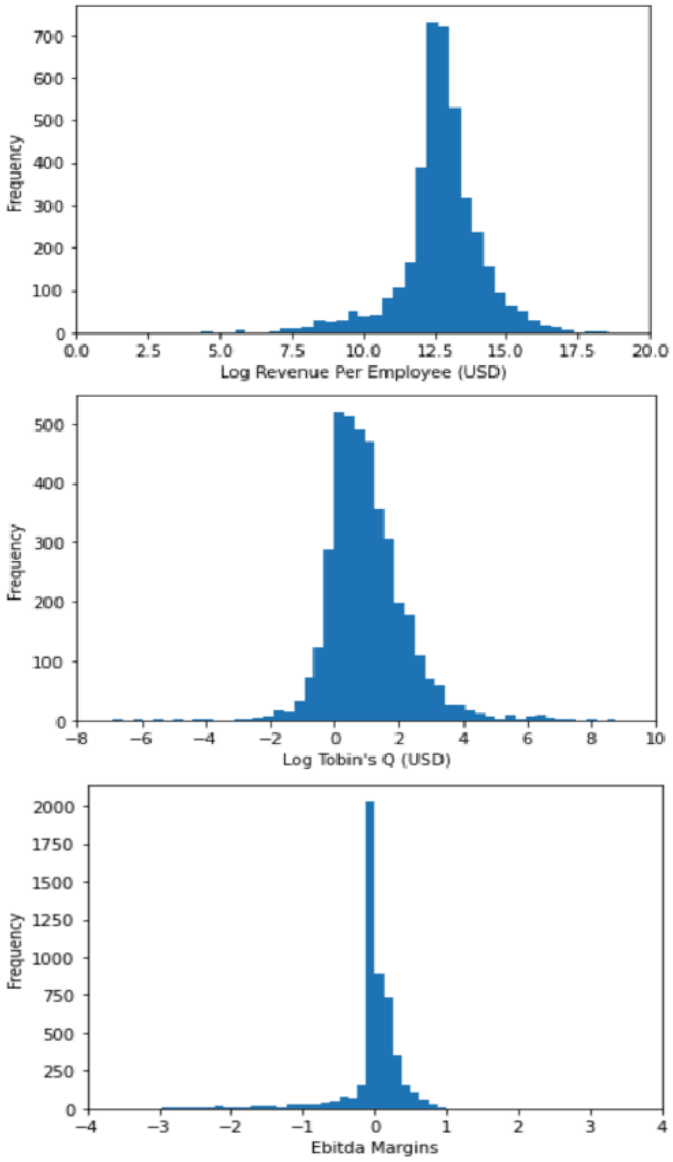


Fig. 3. Distributions of (a) revenue per employee, in log values; (b) of Tobin's Q, in log values; (c) of EBITDA margin.

Regression Metrics	Label				
	Revenue per Employee	Tobin's Q	EBITDA	Fraud	Litigiousness
Root mean-squared error	1.117	0.983	0.336	0.008	0.009
Mean absolute error	0.712	0.707	0.181	0.005	0.006
Median absolute error	0.431	0.533	0.100	0.003	0.004
R^2	0.445	0.253	0.261	0.359	0.484
Classification Metrics	4-way	4-way	2-way	4-way	4-way
Accuracy	0.556	0.404	0.897	0.472	0.554
Balanced accuracy	0.551	0.411	0.897	0.464	0.554
MCC	0.407	0.216	0.793	0.295	0.407

Table 1. ML models fitted to HC text and scores for various business outcomes. This table shows regression and classification results. The feature set comprises a column of HC text and four columns of scores, one each for talent, leadership, organization, and HR. The column header "2-way" stands for binary classification and "4-way" stands for classification into four categories.

The R^2 = 0.25 from the regression model and the *MCC* = 0.22 in the classification model suggest that the fit to the data supports a connection between HC features and Tobin's Q.

6.3 Earnings before interest, taxes, depreciation, and amortization (EBITDA)

EBITDA is an important measure of firm profitability and operational efficiency. It ignores non-operational expenses and is hence a better metric to use when assessing the impact of HC. EBITDA is also often used to generate baseline firm valuations, as a multiple of EBITDA. EBITDA margin is used, i.e., EBITDA divided by revenue.[14] The range of EBITDA margins in our sample is shown in Figure 3.

Interestingly, the figure above displays the classic cliff to the left of the peak, evidencing earnings manipulation as first highlighted in the paper by [7], and more recently in work by [1]. This shows that firms that are about to report barely negative EBITDA, may be undertaking window-dressing of their accounts to push EBITDA to the positive region.

We fitted both, a regression model and a classification model. For the latter, we created a binary split of the data for positive versus negative EBITDA (notice that the data has a pronounced left skew). For both models, we report the results in Table 1. The R^2 = 0.26 from the regression model suggests that the fit to the data supports a connection between HC features and EBITDA. The results

14 https://corporatefinanceinstitute.com/resources/knowledge/valuation/ebitda-multiple/

from the classification model are very strong with an accuracy level of 89% and an area under the curve (AUC) from ROC analysis of 0.95. We see a high *MCC* of 0.79 as well. The model fit to this earnings metric strongly relates to HC reporting.

6.4 Social Responsibility

Using a lexicon of words related to two concepts, fraud and litigiousness, we score the MD&A section of the 10-K filings to get proxies for social responsibility from the management discussion, because the absence of fraud and litigious reporting suggests a good level of corporate responsibility. We then see if the feature set (HC text and four pathways) provides a good fit to the social responsibility outcomes.

The distribution of these scores in the dataset is shown in the histograms in Figure 4. The plot on the left is for fraud and the one on the right is for litigiousness. The x-axis values represent the fraction of words in the MD&A section that are matched to the fraud and litigiousness word lists.

For both variables, the following is the fit of the regression model, which delivers good R^2 (0.36 for fraud and 0.48 for litigiousness) and *MCC* values (0.30 for fraud and 0.41 for litigiousness). This suggests a relationship between HC activity and business outcomes in the cross-section of firms. This confirms both the validity of the HC framework we propose and its impact on key business outcomes.

7 Engineering the System

We provide a brief description of the engineering pipeline built to implement the analytic system described in the paper. We leverage AWS SageMaker[15] for building the system. The pipeline comprises several subsystems/modules, which are as follows: (1) Module to download and parse SEC 10-K filings, packaged into an SDK.[16] (2) Module to extract HC text from the SEC filings using a trained ML model based on hand-labeling and few-shot learning. (3) Module to generate HC word lists for scoring the HC text, using [6]. (4) Module to score HC and create a dashboard, using SageMaker JumpStart[17] with a special purpose API.[18] (5) Multi-modal ML Training modules to fit business outcomes to HC text

[15] https://aws.amazon.com/sagemaker/

[16] https://sagemaker-jumpstart-industry-pack.readthedocs.io/en/latest/notebooks/index.html

[17] Scoring to prepare a dashboard is discussed here: https://aws.amazon.com/blogs/machine-learning/create-a-dashboard-with-sec-text-for-financial-nlp-in-amazon-sagemaker-jumpstart/.

[18] https://sagemaker-jumpstart-industry-pack.readthedocs.io/en/latest/smjsindustry.nlp scorer.html

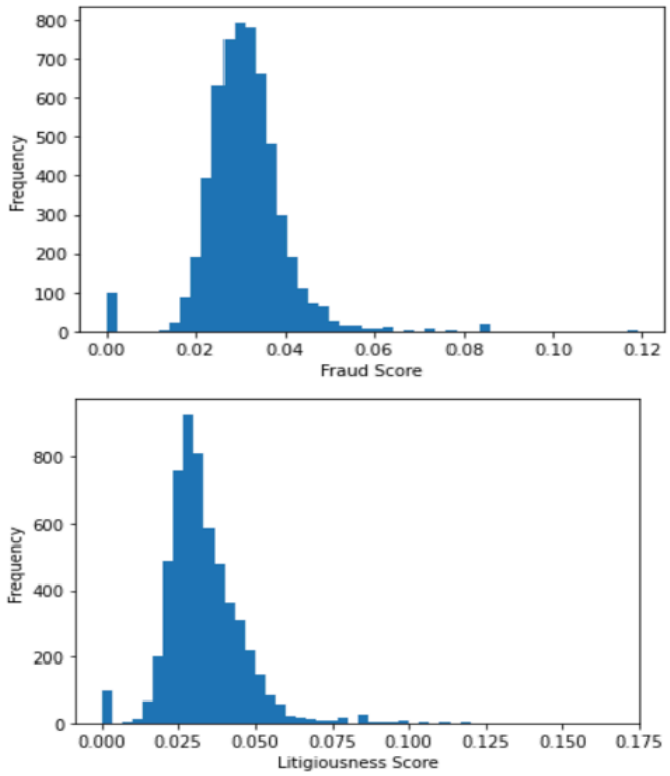


Fig. 4. Distribution of social responsibility scores. The plot on the left is for fraud and the one on the right is litigiousness. The x-axis values represent the fraction of words in the MD&A section that are matched to the fraud and litigiousness word lists.

and HC scores using AutoGluon.[19] (6) ML explainers linking the predictions of the trained models to underlying features using SageMaker Clarify.[20] Remaining work would entail integration of these components into a workflow, UX additions, and report generation.

8 Conclusions

The SEC mandated HC reporting by companies in their 10-K filings. In the absence of a standardized reporting template, companies reported HC activity in many diverse ways throughout their 10-Ks. The system outlined in this paper enables HC assessment at scale applying AI/ML to a four pathways framework and incorporates managerial guidance to enhance business outcomes through better use of human capital. It uses a trained machine learning model to extract text from the filings that relates to HC activity and business outcomes. It devolves HC activity into four categories: (i) talent, (ii) leadership, (iii) organization, and (iv) human resource processes, and scores HC reporting for these attributes using machine learning generated dictionaries for 14 sub-attributes of the four main activities. The system relates reported HC activity to business outcomes using machine learning models, establishing a link to financials, con- comitant with the idea that HC forms a material share of corporate intangible value. In a feedback loop, these analyses will also help companies improve their reporting on HC. Productionizing this work may be supported by artifacts on Amazon SageMaker.

[19] https://auto.gluon.ai/stable/tutorials/tabular prediction/tabular-multi-modal.html

[20] https://aws.amazon.com/sagemaker/clarify/

References

[1] Caramanis C, Lennox C (2008) Audit effort and earnings management. Journal of Accounting and Economics 45(1):116–138, DOI 10.1016/j.jacceco.2007.05.002, URL https://www.sciencedirect.com/science/article/ pii/S0165410107000432

[2] Chicco D, Jurman G (2020) The advantages of the Matthews correla- tion coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 21(1):6, DOI 10.1186/s12864-019-6413-7, URL https://doi.org/10.1186/s12864-019-6413-7

[3] Combs J, Liu Y, Hall A, Ketchen D (2006) How Much Do High-Performance Work Practices Matter? A Meta-Analysis of Their Effects on Organizational Performance. Personnel Psychol- ogy 59(3):501–528, DOI 10.1111/j.1744-6570.2006.00045.x, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-6570.2006.00045.x, eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1744-6570.2006.00045.x

[4] Crook TR, Todd SY, Combs JG, Woehr DJ, Ketchen Jr DJ (2011) Does human capital matter? A meta-analysis of the relationship between human capital and firm performance. Journal of Applied Psychology 96(3):443– 456, DOI 10.1037/a0022147, place: US Publisher: American Psychological Association

[5] Cukier K, Mayer-Sch¨onberger V, V´ericourt Fd (2021) Framers: Human Ad- vantage in an Age of Technology and Turmoil. Dutton

[6] Das SR, Donini M, Zafar MB, He J, Kenthapadi K (2022) FinLex: An effec- tive use of word embeddings for financial lexicon generation. The Journal of Finance and Data Science 8:1–11, DOI 10.1016/j.jfds.2021.10.001, URL https://www.sciencedirect.com/science/article/pii/S2405918821000131

[7] Degeorge F, Patel J, Zeckhauser R (1999) Earnings Management to Exceed Thresholds. The Journal of Business 72(1):1–33, DOI 10.1086/209601, URL https://www.jstor.org/stable/10.1086/209601, publisher: The University of Chicago Press

[8] Gorton GB, Grennan J, Zentefis A (2021) Corporate Culture. SSRN Scholarly Paper ID 3928972, Social Science Research Network, Rochester, NY, DOI 10.2139/ssrn.3928972, URL https://papers.ssrn.com/abstract= 3928972

[9] Graham JR, Grennan J, Harvey CR, Rajgopal S (2021) Corporate Culture: Evidence from the Field. SSRN Scholarly Paper ID 2805602, Social Science Research Network, Rochester, NY, DOI 10.2139/ssrn.2805602, URL https://papers.ssrn.com/abstract=2805602

[10] Guiso L, Sapienza P, Zingales L (2015) Corporate Culture, Societal Culture, and Institutions. American Economic Review 105(5):336–339, DOI 10.1257/aer.p20151074, URL https://www.aeaweb.org/articles?id=10. 1257/aer.p20151074

[11] Huselid MA (1995) The Impact of Human Resource Management Practices on Turnover, Productivity, and Corporate Financial Performance. Academy of Management Journal 38(3):635–672, DOI 10.2307/256741, URL http://amj.aom.org/cgi/doi/10.2307/256741

[12] Huselid MA (2018) The science and practice of workforce an- alytics: Introduction to the HRM special issue. Human Re- source Management 57(3):679–684, DOI 10.1002/hrm.21916, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hrm.21916, eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hrm.21916

[13] Jiang K, Lepak DP, Hu J, Baer JC (2012) How Does Human Resource Man- agement Influence Organizational Outcomes? A Meta-analytic Investigation of Mediating Mechanisms. Academy of Management Journal 55(6):1264– 1294, DOI 10.5465/amj.2011.0088, URL https://journals.aom.org/doi/abs/ 10.5465/amj.2011.0088, publisher: Academy of Management

[14] Li K, Mai F, Shen R, Yan X (2021) Measuring Corporate Culture Using Machine Learning. The Review of Financial Studies 34(7):3265–3315, DOI 10.1093/rfs/hhaa079, URL https://doi.org/10.1093/rfs/hhaa079

[15] McMurrer D, Bassi L (2021) Human Resources Gets Its ISO Approval. URL https://www.workforce.com/news/human-resources-gets-its-iso-approval

[16] Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient Estimation of Word Representations in Vector Space. arXiv:13013781 [cs] URL http://arxiv.org/abs/1301.3781, arXiv: 1301.3781

[17] Popadak J (2016) Balancing governance and culture to cre- ate sustainable firm value. Governance Studies, The Initia- tive on 21st Century Capitalism, No 27, 1-13, Brookings In- stitution 27:1–13, URL https://www.brookings.edu/research/balancing-governance-and-culture-to-create-sustainable-firm-value/

[18] Schneider B, Yost AB, Kropp A, Kind C, Lam H (2018) Workforce engage- ment: What it is, what drives it, and why it matters for organizational per- formance. Journal of Organizational Behavior 39(4):462–480, DOI 10.1002/ job.2244, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/job.2244, eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/job.2244

[19] Smallwood N, Ulrich D (2004) Capitalizing on Capabilities. Harvard Busi- ness Review URL https://hbr.org/2004/06/capitalizing-on-capabilities, section: Accounting

[20] Storey J, Wright P, Ulrich D (eds) (2008) The Routledge Compan- ion to Strategic Human Resource Management, 0th edn. Routledge, DOI 10.4324/9780203889015, URL https://www.taylorfrancis.com/books/9781135981983

[21] Tambe P, Hitt L, Rock D, Brynjolfsson E (2020) Digital Capital and Super- star Firms. Working Paper 28285, National Bureau of Economic Research, DOI 10.3386/w28285, URL https://www.nber.org/papers/w28285, series: Working Paper Series

[22] Ulrich D (2015) The Leadership Capital Index: Realizing the Market Value of Leadership, 1st edn. Berrett-Koehler Publishers

[23] Ulrich D, Brockbank W (2005) The HR Value Proposition, 1st edn. Harvard Business Review Press