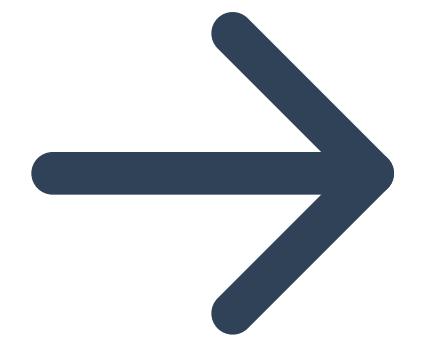




BRAINSTATION APR 2023

Shining A Light On ESG With NLP



HENRY PAREMAIN



The Problem

ESG scoring

is time consuming, costly and inconsistent



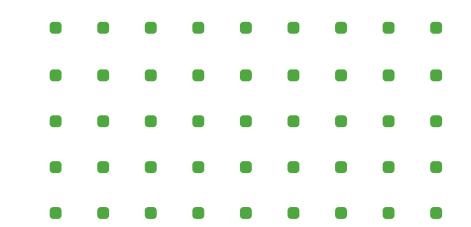
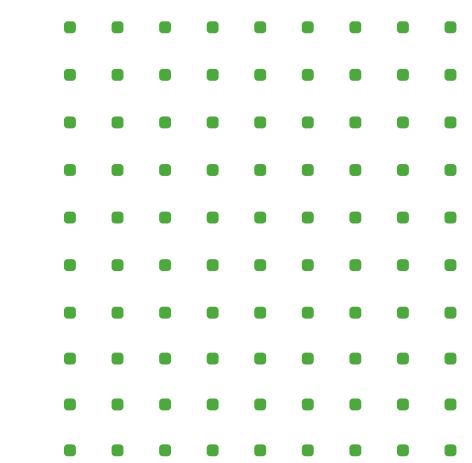
The Question

How can NLP be used to summarise annual ESG reports
and facilitate the ESG scoring process?



03

Data Collection



01

responsibilityreports.com



02

web-scrape 1200 pdfs



03

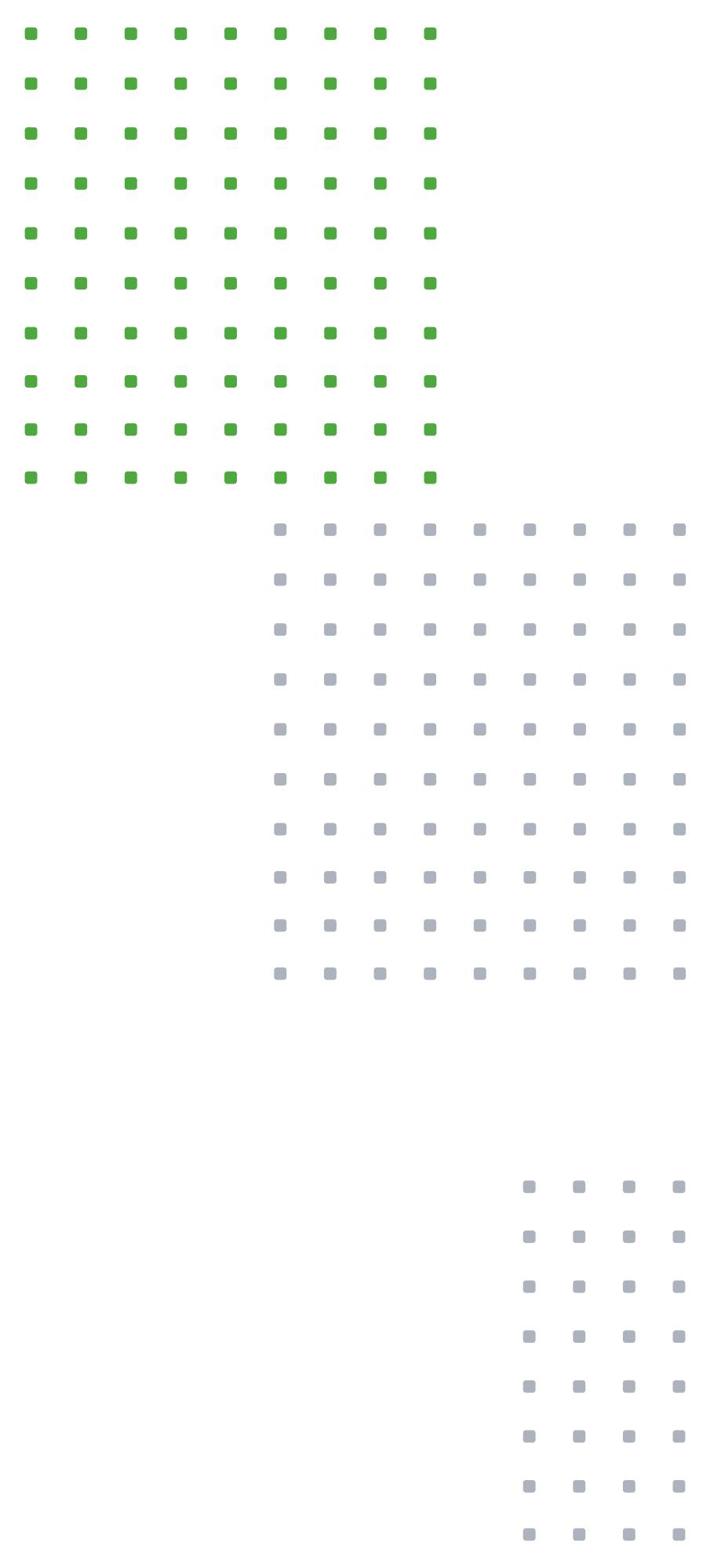
extract raw text from pdfs



Pre Processing

Cleaning Step

Result



Split Camel Case

This isA messyTest, to DEMONSTRATE the \n
100 ofCleaning *£ steps!!

Remove numbers, unicode
and punctuation

This is A messy Test, to DEMONSTRATE the \n
100 of Cleaning *£ steps!!

Lowercase

This is A messy Test to DEMONSTRATE the
of Cleaning steps

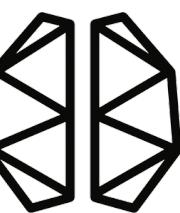
Remove stopwords, adjectives
and verbs

this is a messy test to demonstrate the
cleaning steps

Lemmatize and Tokenize

test steps

[test, step]

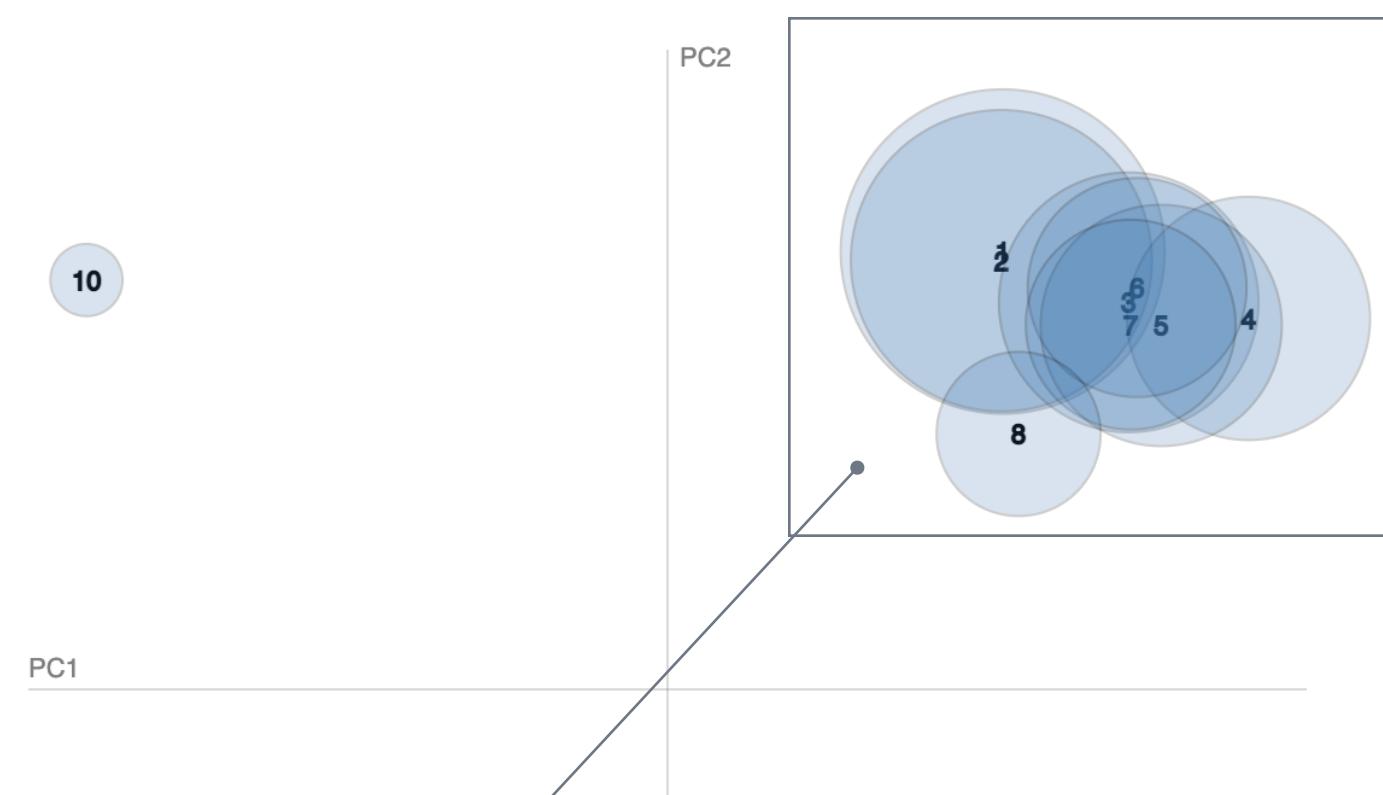


LDA Input Choice

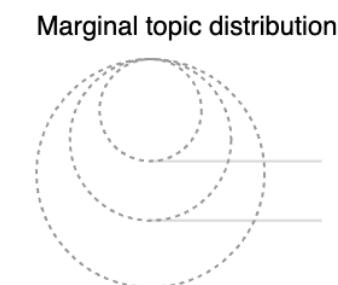
Report by Report

Document = 1 report
Corpus = 1200 reports

Intertopic Distance Map (via multidimensional scaling)



Large overlap,
indicating common
words between topics



Marginal topic distribution

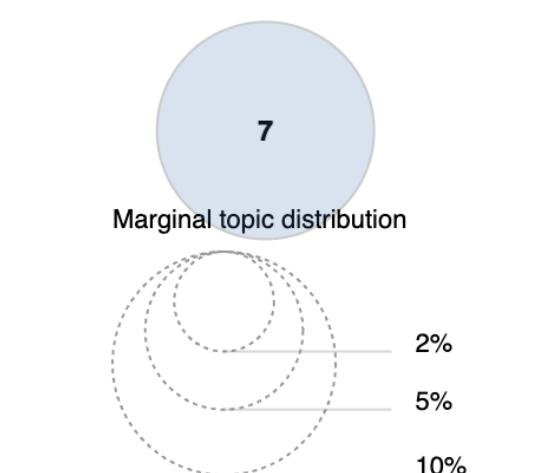
Paragraph by Paragraph

Document = 1 paragraph
Corpus = 1200 reports

Intertopic Distance Map (via multidimensional scaling)



Well separated
clusters, indicating
words are unique to
individual topics



Marginal topic distribution

vs



Results

E S G

Example Paragraphs

Cummins believes a dynamic work environment begins with authentic leaders, who create outstanding places to work by encouraging all employees to achieve their full potential. The company urges leaders to connect employees and their work to Cummins' mission, vision, values, brand promise and company strategies (page 14), giving them a higher sense of purpose.

0.02 0.96 0.02

At the 2020 Annual Meeting in Davos, 120 of the world's largest companies supported efforts to develop a core set of common metrics and disclosures for their investors and other stakeholders. Below is our alignment against the Core metrics from this framework developed by the World Economic Forum, as well as select disclosures from the Expanded metrics.

0.02 0.21 0.77

As part of our plan for carbon footprint reduction, we lead the cruise industry's use of Liquid Natural Gas (LNG) powered cruise ships with a total of 11 next-generation cruise ships that are expected to join the fleet through 2025, including six ships already in operation.

0.96 0.02 0.02

Future work



Network Analysis

to weight the scoring based on the activity of a company's strongest connections

Social Media

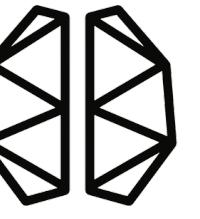
to explore how semantic analysis of posts can be used to weight the scores

LDA Optimisation

through further exploration of hyper-parameter tuning

Explore real world data

to identify if there's a link between actual ESG performance and the topics mentioned in the report

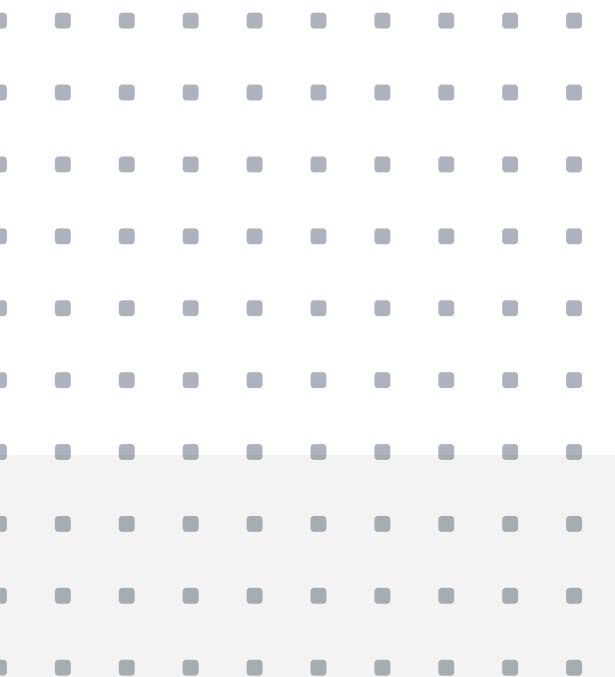


Thank you



Henry Paremain

Data scientist



Presentation template:

- hi@hanihusam.com
- [@bapak2developer](https://twitter.com/@bapak2developer)