

Analysing narratives in the Corporate Financial Information Environment

11

CASS: Briefings

CASS

Corpus Approaches
to Social Science

CASS: Briefings

Published 2015 by

The ESRC Centre for Corpus
Approaches to Social Science (CASS),
Lancaster University, UK

Research

Martin Walker, Steven Young, Paul
Rayson, Mahmoud El-Haj, Vasiliki
Athanasakou, Thomas Schleicher

CASS Centre Director

Tony McEneaney

Series Editing & Design

Robbie Love

Contents

About CASS...

The ESRC funded Centre for Corpus Approaches to Social Science (CASS) is a research centre based at Lancaster University which aims to bring the methods and benefits of the corpus approach to other disciplines.

From the Centre Director

The corpus approach harnesses the power of computers to allow analysts to work to produce machine aided analyses of large bodies of language data - so-called *corpora*. Computers allow us to do this on a scale and with a depth that would typically defy analysis by hand and eye alone.

In doing so, we gain unprecedented insights into the use and manipulation of language in society. The centre's work is generating such insights into a range of important social issues like climate change, hate crime and education. This series of briefings aims to spread the social impact and benefits of the work being done by the centre and, in so doing, encourage others to use our methods in future.

Introduction.....	4
The CFIE project.....	5
Research objectives.....	6
Key findings.....	7
Find out more.....	10



Professor Tony McEnery

Transparent and effective communication between firms and the investment community is a key determinant of corporate success. Audited financial statements and associated narrative disclosures are among the main methods that firms use to communicate with investors and analysts. These disclosures combine with information from financial journalists and other market commentators to form the Corporate Financial Information Environment (CFIE).

Textual disclosures help to clarify issues obscured by complex accounting methods and footnotes. Narratives also help to contextualize financial results and summarize key elements of business activity that are difficult to quantify such as corporate strategy, governance arrangements, and corporate social responsibility policy.

While a considerable body of work exists on financial narratives, research has been limited by the methods used for measuring the characteristics and quality of such disclosures. In particular, the need to hand-collect relevant data from firms' annual reports and the subjectivity of textual scoring based on manual methods has restricted progress. Recent advances in computational and corpus linguistics provide a basis for undertaking more sophisticated analyses.

The Corporate Financial Information Environment project ran for two years (2013-14) and is continuing as part of CASS. It brings together a multidisciplinary team of researchers from the **University of Manchester** (Business School), **Lancaster University** (Accounting and Finance; School of Computing and Communications), and **London School of Economics** (Accounting). The determinants of financial reporting quality and the factors that influence the quality of information disclosed to investors beyond the financial statements are the main focal points of our work.

In particular, we have been developing new ways of measuring the quality and tone of company narratives using computer-based ratings of annual reports. Both the ratings and linguistic techniques on which these ratings are based have been made available to those seeking information on corporate disclosure policy, or wishing to undertake their own analysis of specific narrative statements.

The CFIE project was funded by the **Economic and Social Research Council** (Ref. ES/J012394/1) and **The Institute of Chartered Accountants in England and Wales** (ICAEW). For more details about the project, please see the website:

ucrel.lancs.ac.uk/cfie/

The aim of our work is:

1. To advance research on the lexical properties and narrative aspects of corporate disclosures by developing a suite of corpus-based natural language processing (NLP) tools for analysing firms' narrative communication practices;
2. To use the resulting tools to measure the linguistic characteristics of key corporate disclosures (both mandatory and voluntary), to identify determinants of cross-firm variation in these characteristics, and to relate these characteristics to disclosure informativeness;
3. To apply the tools to advance research on the interactions between corporate voluntary disclosures and accounting quality, including new work on the joint effects of corporate disclosure and earnings management practices on share price anticipation of earnings and companies' standing in published rankings of investor relations quality;
4. To apply scoring methods to UK corporate news stories in the financial media with the aim of developing a more complete measure of corporate financial communications quality;
5. To use the methods and insights from objectives 1 to 4 above to provide new evidence on the links between earnings quality, disclosure quality, and cost of capital.

To date our work has yielded four primary outputs:

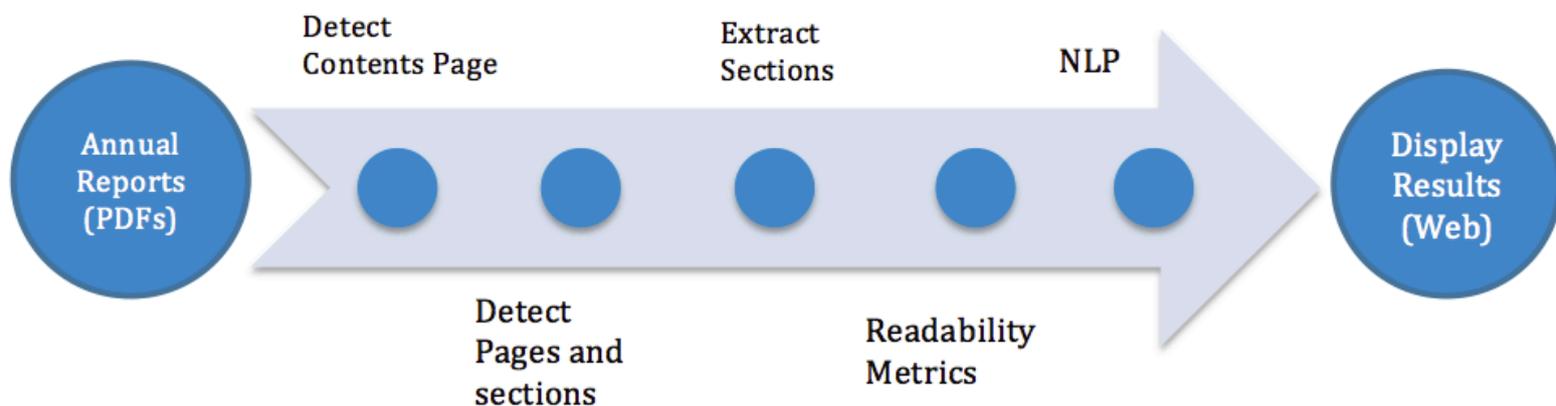
Output 1

- **The creation of a publicly available web tool to permit large batch scoring of UK annual report narratives**

Details of the web tool are available at:

<https://cfie.lancaster.ac.uk:8443/>

The diagram below shows the process embedded in the CFIE web tool.

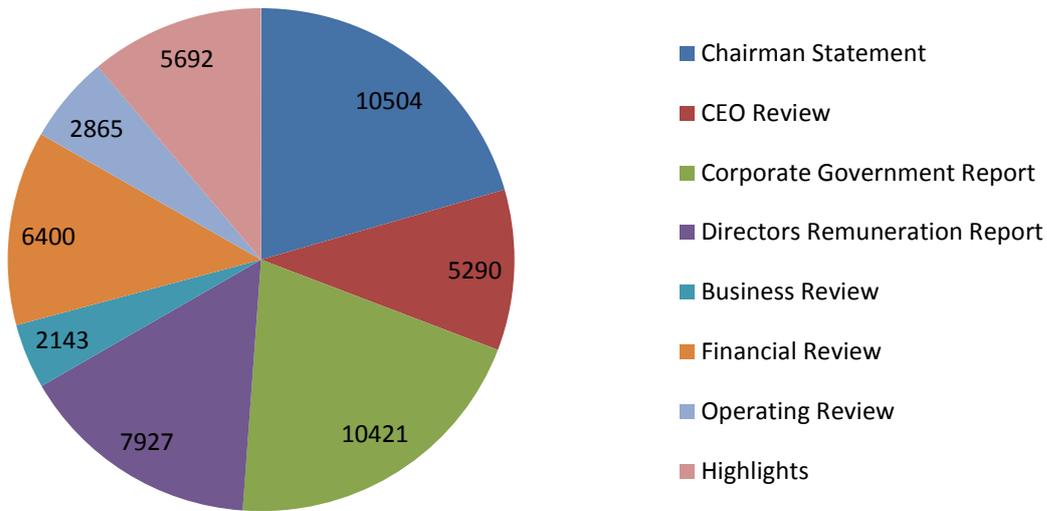


Key features include its use of novel hybrid text analysis methods to detect the structure of UK annual reports, count word frequencies, calculate readability and other quality metrics to score individual sections of annual reports such as strategy, performance, governance and remuneration. It also carefully distinguishes between the front-end annual report narratives, and the mandatory financial statements and corresponding notes. The tool links with the Wmatrix corpus tool for further comparative language analysis:

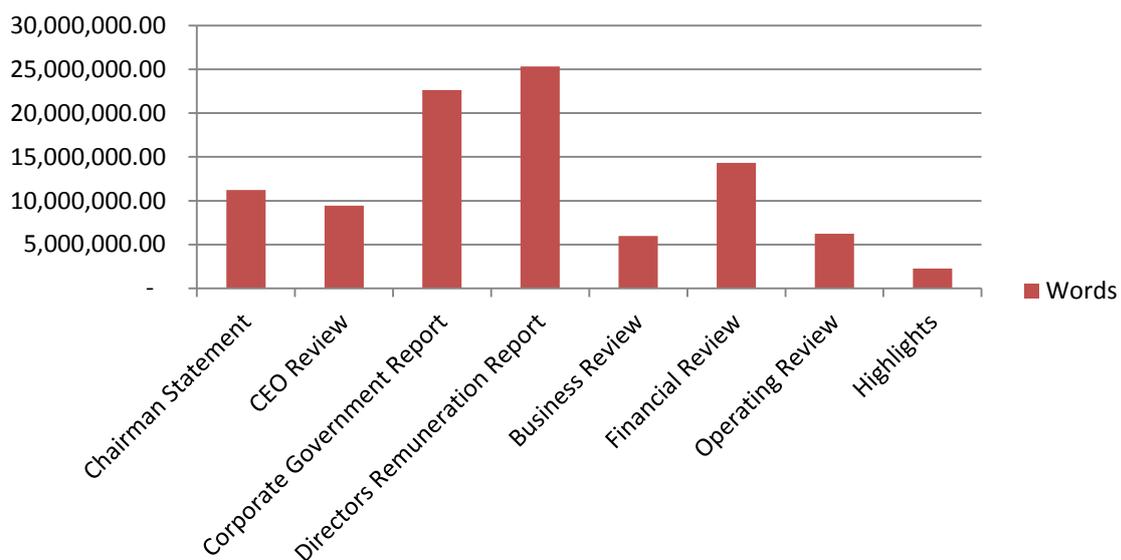
ucrel.lancs.ac.uk/wmatrix/

The web tool has been used by the team to score the linguistic characteristics of the narratives in over 10,000 UK annual reports, spanning the period 2003 to 2013. These scores and reference corpus profiles are also being made publicly available.

Key Sections Count



Key Sections Word Count



Output 2

- **The large scale analysis of the cross-sectional variation, and variation over time in annual report narratives across several dimensions**

We have carried out the first large scale analysis of the link between narrative disclosure and UK firm performance and the first detailed empirical analysis of the development of disclosures relating to the business strategies of UK firms in their annual report narratives. We have produced the first comprehensive index of narrative disclosure quality for a very large panel of UK listed companies and its relation to time series and cross section variation in the cost of equity capital.

Output 3

- **Significant progress towards understanding and analysing the content and impact of other corporate disclosures relating to the financial performance of UK firms**

This includes Interim Management Statements, and Preliminary Earnings Announcements (PEAs). This body of work includes findings from interviewing key capital market stakeholders for their views on the relative timeliness of PEAs and Annual Report narratives, and the usefulness of Interim Management Statements.

Output 4

- **Significant new research on understanding the causes and consequences of the quality of the corporate financial information environment**

Work on the interaction between investors' perceived credibility of forward-looking disclosures and earnings quality, original work on earnings quality and earnings management for both US and UK firms; work on the interaction between company press releases and financial media articles in the USA, and the creation of a novel, stock returns based, measure of asymmetry in the flow of good news versus bad news that is complementary to other measures of disclosure quality.

Find out more

Project conferences in 2013 and 2014 brought together leading practitioners and international researchers to shape the "narratives" research agenda. Details of the conferences, published outputs and working papers are available on the project website at

ucrel.lancs.ac.uk/cfie/

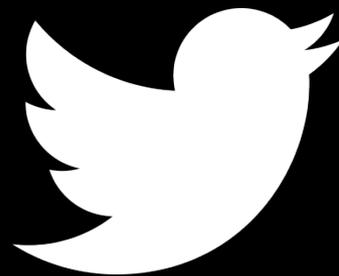
Part of our aim at CASS is to make Corpus Linguistics accessible, which is why we have created our **free online FutureLearn course**. With the course, we aim to demonstrate that corpus approaches can offer researchers from all disciplines unique, valuable insights into the use and manipulation of language in society. We provide all you need to start 'doing' Corpus Linguistics yourself.

This briefing should act as an introduction and companion to the course where you will begin to apply the concepts and methods mentioned here in a practical way relevant to your field of interest.

The course is free, can be done from home, and comes with a whole range of content and support from world-leading scholars in the field of Corpus Linguistics. For more, visit:

futurelearn.com/courses/corpus-linguistics

For more about CASS and our freely available resources, please visit: **cass.lancs.ac.uk**



CASS

@CorpusSocialSci

CASS: Briefings is a series of short, quick reads on the work being done at the ESRC/CASS research centre at Lancaster University, UK. Commissioning work from internationally recognised academics in the field of Corpus Linguistics, *CASS: Briefings* set out to make cutting edge research easily accessible, providing a good introduction to the variety of vital and exciting research going on in the area of Corpus Linguistics.

