New software brings clarity to hard-to-decipher company annual reports

August 8 2019

New software from Lancaster University cuts through hard-to-understand financial reports, to help investors and regulators.

Researchers have developed an app to dissect and analyse narrative aspects of companies' annual reports, documents aimed at shareholders, but also used by other stakeholders, including financial analysts, prospective investors, journalists and regulators.

Annual report narratives contain commentary on financial performance, as well as supplementary information on topics such as principal risks and corporate social responsibility policies—but management has a high level of discretion over their content and structure. As a result, investors can struggle to find and understand the information they require.

At present, there is no uniform structure to such documents, making comparison and large-scale analysis severely challenging. The Corporate Financial Information Environment—Final Report Structure Extractor (CFIE-FRSE) app aims to cut through hard-to-understand annual report language and help users identify unusual patterns in corporate reports that may help to distinguish long-term financial strength from inflated short-term profits.

The app aids regulators in seeing where businesses may be trying to conceal information and where they need to intervene. It also allows them to see where regulation is working and where it may need to change.

Professor Steve Young, Head of Accounting in Lancaster University Management School, said: "Annual reports are highly unstructured, and different companies report in different ways, which makes extracting content and comparing reports very difficult. Almost every document is different.

"Many reports are almost impossible for non-specialists—a group which includes many investors—to read, which is at odds with the trend towards a broader model of stakeholder reporting.

"We have designed an app to extract commentary from these documents and normalise it across firms—making comparisons much easier. The procedure provides a reliable means of capturing and classifying these narratives."

The interdisciplinary project has involved academics from Lancaster University Management School, the School of Computing and Communications, and the Department of Linguistics & English Language.

There has already been interest in CFIE-FRSE from investment and hedge-fund managers, who would gain a greater insight into the status and stability of companies, as well as regulators seeking to assess the effectiveness of regulation.

More than 26,000 documents published between 2003 and 2017 by companies listed on the London Stock Exchange have been analysed by the app and scored on features such as length, readability and sentiment. Because the CFIE-FRSE app detects report structure, scores are available for each section listed in the table of contents.

Dr. Mahmoud El-Haj, Senior Research Associate in the School of Computing and Communications at Lancaster University, said: "The app

uses heuristic approaches and rule-based decision making to automatically detect the structure of an annual report. This helps the software to extract sections' text by knowing their start and end pages.

"The app was trained to identify a set of common section titles (types) based on a training list of synonyms generated by accounting and finance experts. For example, the app is able to identify that the 'Letter to shareholders' in one company's report is the same as the 'Chairman's statement' in another company's report."

Analysis of the annual reports processed by the app reveals a number of interesting features and reporting trends. For example, average report length has more than doubled over the last decade to almost 34,000 words.

Average report readability—measured using an algorithm that penalises long sentences and complex words—is also poor; and there has been no noticeable improvement over the sample period. Long, unstructured documents containing complex language mean many retail investors and other non-specialist stakeholders struggle to understand the typical annual report.

Sentiment also varies dramatically across different sections within the same report. For example, sections where regulation and compliance shape content—such as governance statements and remuneration reports—are characterised by neutral language. In contrast, the tone of language is up to four times more positive in sections where directors have more reporting discretion and where performance is the primary focus.