**AUTHOR(S):** CLEVERLEY, P.H. and MUIR, L.

**TITLE:** Using Knowledge Organization Systems to automatically detect forward-looking sentiment in company reports to infer social phenomena.

**YEAR:** 2018

**Publisher citation:** CLEVERLEY, P.H. and MUIR, L. 2018. Using Knowledge Organization Systems to automatically detect forward-looking sentiment in company reports to infer social phenomena. Knowledge organization [online], 45(2), pages 152-169. Available from: https://doi.org/10.5771/0943-7444-2018-2-152.

**OpenAIR citation:** CLEVERLEY, P.H. and MUIR, L. [2018]. Using Knowledge Organization Systems to automatically detect forward-looking sentiment in company reports to infer social phenomena. Knowledge organization, 45(2), pages 152-169. Held on OpenAIR [online]. Available from: https://openair.rgu.ac.uk

# Using Knowledge Organization Systems to automatically detect forward-looking sentiment in company reports to infer social phenomena.

ABSTRACT

The study investigates whether existing Knowledge Organization Systems (KOS) for strong and hesitant forward-looking sentiment could be improved to detect social phenomena.

Five judges identified examples of strong/hesitant forward-looking sentiment which were used to compare the KOS developed in the study, to existing models. The 'composite' KOS was subsequently applied to annual company reports to generate word frequency and biologically inspired diversity ratios. Critical Realism was used as a philosophy to interpret word patterns.

Results indicate the composite KOS improved on existing models identified in the literature for strong forward-looking sentiment. In one company, a statistically significant association was found between increasing diversity of assertive forward-looking sentiment and subsequent declining relative business performance. This supported the Pollyanna effect: the social phenomena of over-positive business language in that company. Sharp increases in mentions of the 'future' and 'learnings' was discovered in another company which may be explained by an industrial disaster and subsequent crisis management rhetoric, supporting Discourse of Renewal Theory.

This study shows that improvements can be made to existing KOS used to detect forward-looking sentiment in reports. Adopting Critical Realism as a philosophy when analysing 'big data' may lead to improved theory generation and the potential for differentiating insights.

# 1. INTRODUCTION

## 1.1 Background

Exploiting its knowledge and competencies may be the only sustainable competitive advantage for an organization (Davenport and Prusak 2000). However, Pauleen et al. (2015) argue wisdom requires particular attitudes (ontologies, epistemologies and axiologies) towards knowledge, an approach largely absent from the Knowledge Management (KM) concept. Avoiding technological bias and determinism, the importance of human insight, 'data does not speak for itself', is highlighted by Floridi (2014), where opportunities posed by 'big data' may be as much about brainpower as computing power. Despite this, little research exists on the links between KM and 'big data' (Davenport 2013).

A research paradigm is a way of thinking about the world (a world view) in order to make sense of its complexities and considers issues such as 'what exists', 'how do I know' and 'what is valuable?' (Patton 2015). When knowledge is said to be 'hidden' in big data (Khan and Vorley 2017) it implies that knowledge exists separate to people, rather than knowledge being constructed through the minds of people. Adopting a philosophy towards 'big data' may be a key component of KM when examining the possibilities for action to exploit the wealth of information available to organizations. Statistical generalizations and demi-regularities in data may inform investigations. However, treating organizations as complex systems, identifying tendencies and seeking explanations may lead to more transformative outcomes (Boulton et al. 2015).

## 1.2 LIS, IR and Text Analytics

Much of the Library and Information Science (LIS) literature has been concerned with using Knowledge Organization Systems (KOS) such as taxonomies and dictionaries to manually index document objects so they can be stored in an information system for retrieval (Zeng et al. 2007). However, the LIS community has been slow to adopt automated methods (Ibekwe-Sanjuan and Bowker 2017) that deconstruct texts within the information aggregate (collection) where their sum may be greater than their parts (individual documents), leading to emergent properties (Aaltonen and Tempini 2014). The scope of KOS also extends beyond the narrow use of concepts to label information containers or things (physical or digital). Taxonomies and typologies have also been used historically as methodological tools to classify traits such as personality, cognition (Bloom et al. 1956) and psychological quantities such as the Five Factor Model (John 1990).

The Information Retrieval (IR) discipline focuses on matching a user's intent through a query to the most relevant document (information) acting through technology (Ruthven 2008). Within these documents, modal verbs are often used to show what we believe is possible or certain in the future (such as 'might', 'could' and 'will'). These are however, often treated as 'stop words' and removed from search indexes (Manning et al. 2008; Li 2010) within corporate 'Google-like' enterprise search deployments.

Text and Data Mining (TDM) is the use of automated analytical techniques to analyse text and data for patterns, trends and other useful information. These techniques can be used to summarize, synthesize and compare (Manning et al. 2008), supporting higher level thinking processes rather than simply retrieving (remembering) information (Bloom et al. 1956). Derived technology applications have recently been termed 'insight engines' or 'cognitive search' (Tetu 2016). One TDM technique is sentiment (tone) analysis, concerned with identifying meaning in text such as its polarity (positive or negative) and the strength or intensity of that opinion (Taboada 2015). They may reveal where individuals or entities are focusing their level of immersion and hidden motives. Emotions such as anger, calmness, fear, happiness and surprise have also been detected from text using algorithms (Pulman 2014).

Some scholars suggest we are at the cusp of a technology revolution linking word usage to real world intentions and behaviours (Tausczik and Pennebaker 2010) and they are gaining increasing importance within the enterprise (Kruschwitz and Hull 2017). There appears a growing realization that exploiting unstructured text can lead to potential insights on the future that cannot be gleaned from traditional numerical data and indices stored in structured databases. However, sentiment engines are likely to need customizations (Van Boeyen 2014). For example, using an off-the-shelf commercial sentiment analysis tool, it was reported that the American Red Cross found that only 21% of positive comments were successfully detected by the software (Grimes 2012).

1.3 Content Analysis

Content Analysis is a set of manually undertaken research methods for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use (Krippendorff 2004). It is a method for inquiring into social reality. It has only recently become more common with organizational and management scholars and challenges include the development of 'proxy dictionaries' lexicons (University of Georgia 2012). Modelling KOS is crucial for content analysis, as emphasized by Berelson (1952, 92), "Content analysis can stand and fall on its categories".

Another challenge for content analysis is the exponential increase in information volumes, 'big data' and 'big literature' (Gantz and Reinsell 2011) which preclude scholars from reading all the available material. This may be particularly significant in what is termed, a post-truth society (Brown 2016). For example, it has been reported that the 'lessons learnt' system in one organization would take a person over five years to read (Smith 2015).

Automated unsupervised machine learning techniques using various forms of complex word co-occurrence are capable of surfacing intricate structure within texts (clustered topics) without any external information (such as a KOS) being input into the original text corpus (Mikolov et al. 2013; Blei et al. 2003). However, Cambria and Hussain (2015) suggest a shift has been taking place in recent years in sentiment analysis, from syntax towards semantics. For example the phrase "The iPhone6 is expensive but nice" has the opposite polarity to the phrase "The iPhone6 is nice but expensive" in terms of consumer sentiment, evidencing the significance of small word order effects. Despite this, recent literature (Khan and Vorley 2017) tends to focus on inductive statistical clustering to glean insights from 'big data', largely neglecting the value that automatically applying deductive KOS/categories to text (Cleverley and Burnett 2015) might bring to facilitate discovery of new insights.

Michel et al. (2011) used the term 'culturomics' to define the study of word frequency patterns (corpus analysis) using automated techniques on 'big data'. Using a geological analogy, these word patterns may provide 'trace fossils' of social history within both society and organizations. The following section therefore reviews the landscape, models and studies where lexicons have been used to extract forward looking sentiment from texts, particularly related to organizations.

## 2. LITERATURE REVIEW

### 2.1. Annual Company Reports and modal verbs

Different nationalities and cultures (including native and non-native English speakers) may use modal verbs (e.g. 'could', 'must' and 'may') in different ways (Hinkel 1995). Research has also shown that non-native English speaking students used some modal verbs (such as 'can', and 'will') with twice the frequency of their English speaking professional counterparts, but used others (such as 'might') with half the frequency (Hykes 2000). Other studies of marketing disciplines in business however, have shown no statistically significant differences in modal word usage between native and non-native English speakers (Nathan 2010).

In a study of Annual Company Reports, Rutherford (2005) postulated that they were devices of stakeholder 'impression management'. Annual company accounting reports have been analysed through word frequencies of dictionary terms (Rutherford 2005), Natural Language Processing (El-Haj 2014) and collocation networks (Kloptchenko et al. 2004) using neural networks (Hajek and Olej 2013).

In addition to presenting indications of industry climate and company strategy, analysis of modal verbs used in Company Reports may also provide indications of organizational rhetoric concerned with persuasion (Ulmer et al. 2011) and underlying beliefs and ideologies.

Usage of modal verbs has been reported as being higher in business language (Yasumasa 2008). Findings in the literature include evidence in business communication for rhetoric and over-positive language known as ‘the Pollyanna effect’ (Hildebrandt and Snyder 1981). Previous studies have addressed genres such as charged words (e.g. losses versus profits) and financial position (e.g. assets versus liabilities). There is evidence for smaller and less profitable firms disclosing less information and companies disclosing more information during periods of increased earnings (Fisher et al. 2008).

Links have been shown for increases in risk sentiment words and lower future company earnings (Li 2006). In a study of financial news in the Wall Street Journal, researchers found high levels of pessimistic words in the column preceding lower stock market returns the next day (Tetlock 2007). Yuet-yung (2014) found that in a study of the worst performers and the top performers in the Fortune 500, top-performers were more assertive in their presentation of future possibilities.

Sentiment analysis has been applied to social media (He et al. 2017) and also to text such as movie reviews using SentiWordNet (2010), assigning a 0 (positive) or 1 (negative) score to words in the large lexical database WordNet (2010). Malhotra (2013) identified patterns and synonyms for detecting hypotheses in text using modal verbs, additional verbs and adjectives, although no division was made on the strength or intensity of conviction. Bochkay and Dimitrov (2014) assessed the positive (optimism) and negative (pessimism) tone of sentences of annual company filings. They found a systematic bias, when managers were more optimistic, future earnings were low and vice versa.

Finer grained continuous scales have been found to provide more accurate results than binary sentiment (Reagan et al. 2015). Continuous scales have been developed for sentiment such as the Semantic Orientation Calculator (Taboada et al. 2011).

The use of strong and weak modal verbs as part of an ensemble machine approach using word context has been applied to risk (Wang et al. 2013) and fraud detection in 10-K company filings (Humpherys et al. 2011). In the latter study, it was found that the ratio between hedge cues and total number of words in 10-K filings did not demarcate deceptive information. More frequent use of the hesitant modal verb ‘could’ and less frequent use of the strong modal verb ‘will’ did provide statistical significance for identifying fraud.

Loughran and McDonald (2011) found that companies with a higher proportion of weak or strong modal words, were more likely to have a material weakness in internal controls. It has been found that fraudulent reports tend to use more words (Bodnaruk et al. 2015) supporting Management Obfuscation Theory (Bloomfield 2002). However, it has also been reported that annual company reports have grown in size (number of words) by 50% between 2006-2015 due to the increasing complexities for regulation (Deloitte 2015). There is evidence that deception, or attempts to conceal information, lead to higher lexical diversity (Siegel et al. 2013). Minhas and Hussain (2016) found several constructs have been used successfully to identify deception using computational linguistics, including word quantity (to obscure truth) and more use of modal verbs.

Most modal verbs are theoretically polysemic (have multiple meanings depending on the context). For example, ‘could’, expresses a realization/possibility of an event occurring “Inflation could affect..” as well as requesting permission “Could I do this?” and ability “I could do this..”. However, in analysis of technical reports, the use of ‘can’ and ‘could’ in a permission (rather than possibility) sense, has been reported to be virtually absent, leading to effectively monosemic (single meaning) modal verb usage in certain contexts (Jaime and Perez-Guillot 2015). This was supported by Pique-Angordans et al. (2002) who found a propensity for many modal verbs (such as can, could, may, might, will, would) within documentation of a technical nature, to display almost 100% epistemic (a hedge) meaning. In analysis of 80 years of the TIME magazine, it was found that by the year 2006, the modal verb ‘may’ was associated with a hedge meaning 94% of the time, whilst ‘must’ and ‘should’ were deontic (about the future) 80% and 67% of the time respectively (Miller 2009). The significant increase in frequency of ‘may’ and ‘could’ in the TIME corpus from 1923 to the year 2006 was attributed to increased speculative reporting.

Dictionary (rule based) approaches automatically count the number of times words appear in texts or sentences. Li (2010) found no association between sentiment dictionaries (such as Diction, General Inquirer and Linguistic Inquiry and Word Count (LIWC)) and financial performance in company reports, with the assertion made that the dictionaries did not work well for the financial domain. This is supported by Loughran and McDonald (2011) who concluded that, when financial text is analysed, traditional words described as ‘negative’ in dictionaries are not negative in a financial sense, such as ‘liability’, ‘cost’ and ‘tax’. In general, dictionary approaches may not be transferable between domains, if they are domain specific.

Li (2010) classified 30,000 forward-looking statements in annual reports and used them to train a Naïve Bayes machine learning classifier. The derived model was subsequently applied to 140,000 annual reports. A positive tone was correlated positively with a 5% increase on return the following year. In general, machine learning approaches appear useful if existing dictionaries do not exist, or the domain scope is hard to pre-define.

Some methods include a hybrid approach and assess a specific industry vertical. For example, Gupta and Liu (2017) inferred organizational culture towards risk of Banks by analyzing annual reports of 578 banks using dictionaries of positive and negative words as well as words that represent risk categories. An unsupervised clustering algorithm was applied in order to group banks of similar types.

Many studies analysing word patterns in company reports seek to link statistically significant results to ‘universal laws’ that apply to all organizations (Bochkay and Dimitrov 2014; Humpherys et al. 2011; Fisher et al. 2008). Organizations are likely to be complex social systems. Rather than obeying ‘laws’, knowledge of organizational phenomena is likely to be more contextual and concerned with generative mechanisms and tendencies, rather than broad generalisations and absolute outcomes.

Studies have shown that various verbs (such as ‘shall’, ‘will’, ‘might’ and ‘could’) may provide an indication of assertiveness ‘sentiment’, a forward-looking opinion or hedge. An assumption made is that the use and distribution of particular verbs and adjectives in any manifesto is not random, but deliberate. Parts of speech such as modal verbs may indicate how definite or confident a company feels about a proposition; their attitude towards a state of affairs and possibility (certainty) of future events and outcomes. The next section reviews some of the existing dictionaries and markers used for identifying forward looking opinion in company reports assessing their strengths and weaknesses.

## 2.2 Lexical markers for forward looking sentiment

The LIWC dictionary (Tausczik and Pennebaker 2010) contains many informal ‘spoken’ words (such as ‘sure thing’ and ‘shoo-in’) which are unlikely to be present in formal corporate communications. The LIWC dictionary has two categories of interest to this study, ‘tentative’ (e.g. ‘maybe’, ‘perhaps’, ‘guess’) and ‘certainty’ (e.g. ‘always’, ‘never’). The dictionary is of a commercial nature so all the words are not available freely to the academic community. These ‘tentative’ and ‘certainty’ category descriptions however, do not necessarily translate into forward-looking ‘tentative’ and ‘certainty’ opinion.

Some studies have attempted to categorize lexical modality in Bio-medical texts (Thompson et al. 2008), as shown in Table 1, presenting the words used in respective categories.

Table 1 – Modality by lexical category defined from Biomedical texts (Thompson et al. 2008)

| | Category | Complete List of Words |
|---|---|---|
| Knowledge type markers | Speculative | Assume, assumption, belief, believe, claim, conceivable, estimate, expect, expectation, hypothesize, hypothesis, hypothetical, in principle, in theory, judge, model, notion, predict, prediction, proposal, propose, speculate, suggest, suggestion, suppose, suspect, theory, think, to our knowledge, view |
| | Deductive | Argue, argument, deduce, imply, indicate, indication, infer, interpret, interpretation, suggest |
| | Demon-strative | Conclude, conclusion, confirm, confirmation, demonstrate, find, finding, proof, prove, report, reveal, show |

| | | |
|---|---|---|
| | Sensory | Apparent, apparently, appear, observation, observe, evidence, evident, seem, see |
| Certainty markers | Absolute | Certainly, known |
| | High | Consistent with, clear, clearly, generally in agreement with, likelihood, likely, normally, obviously, probability, probable, strongly, support, would |
| | Medium | Can, could, feasible, may, might, perhaps, possibility, possible, potential, potentially |
| | Low | Unlikely, unknown |

The results indicate that the prediction of modality can be straight forward using lexical words with a small amount of contextual information. Critiquing Table 1, the use of just two words in the absolute and low categories for certainty markers may lead to sparse data and is unlikely to be comprehensive.

There are three main groupings of modal verbs (EOI 2012): epistemic (assessing confidence in propositions and speculation), deontic (generally about the future, how the world should be) and dynamic (ability & volition).

Table 2 (UNC Chapel Hill 2014) shows a grouping by the dimensions of 'strong' and 'weak' modal verbs by their typical frequency of use. They are used most frequently to indicate logical possibility.

Table 2 - Modal verbs by strength and frequency (UNC Chapel Hill 2014)

| | Most frequent | > | > | Least Frequent |
|---|---|---|---|---|
| | Logical possibility | Ability | Necessity | Permission |
| Strongest | Must | Can | Must | May |
| ^ | Will/would | Could | Should | Could |
| ^ | Should | | | Can |
| ^ | May | | | |
| Weakest | Can/could/might | | | |

Critiquing, Table 2 is missing 'ought' which is a modal verb and probably fits in between the strongest and weakest category. The modal verbs 'might', 'can' and 'could' are grouped equally weak which may miss finer variations in sentiment.

Piotti (2014) identified many devices indicative of 'hedging' with respect to a position on a future state of affairs (Table 3).

Table 3 **–** Hedging devices (Piotti 2014)

| Categories | Words |
|---|---|
| Modal auxiliaries | should, will, would, may, can, could, shall, might |
| Full verbs (reporting) | propose, imply, indicate, suggest |
| Full verbs (tentative cognition) | expect, assume, estimate, think, believe, evaluate, presume, allege |
| Adverbs of probability | likely, potentially, basically, possibly, reliably |
| Adverbs of indefinite frequency | generally, regularly, usually, normally, typically, occasionally, rarely |

Piotti (2014) complements modal verbs with additional verbs and adverbs that may be indicative of uncertainty. However, on their own many of these words may be predominantly used in both a past (backward looking) or present tense (such as typically) rather than an opinion about the future.

Modal verbs have also been grouped 'pragmatically' (TeachIT 2016) into three degrees of certainty (Table 4).

Table 4 – Modal verbs grouped by degree of certainty (TeachIT 2016)

| Degrees of certainty | Modal verbs |
|---|---|
| Strong | will, shall, must |
| Moderate | should, would, can, ought |
| Hesitant | might, may, could |

From a possibility perspective, 'must' is very strong (forcing something to occur) whilst 'may', 'could' and 'might' are suggested as the weakest, showing low commitment or confidence.

In an analysis of 10-K company filings, Bodnaruk et al. (2015) analysed US Securities and Exchange Committee (SEC) EDGAR filings and extracted what they believed to be words indicative of strong and weak modality (shown in Table 5).

Table 5 – Strong/weak modal words from SEC filings (Bodnaruk et al. 2015)

| Certainty level | Modal verbs |
|---|---|
| Strong | will, shall, must, undoubtedly, never, lowest, is, highest, definitely, clearly, best, always |
| Weak | might, may, could, uncertainty, suggest, sometimes, seldom, possibly, possible, perhaps, occasionally, maybe, depends, depending, could, conceivable |

Compared to TeachIT (2016) shown in Table 4, this adds assertive words, some of which may not always relate to sentiment about the future (e.g. clearly).

Cassidy (2013) organized hesitant words (also known as hedges) by function/categories to apply to a corpus (Table 6).

Table 6 – Hedging lexicon with descriptions (Cassidy 2016).

| Category | Description | Examples |
|---|---|---|
| Approximation | Indicates proposition is an estimate | About, almost, approximate, estimate, many, most, nearly, some |
| Degree | Indicates how well proposition fits into a category | Essentially, mostly, partially, quite, relatively, slightly, somewhat, virtually |
| Frequency | Indicates how often proposition occurs | Generally, normally, occasionally, often, rarely, usually |
| Intention | Indicates future plans | Intend, plan, propose, seek |
| Logic | Indicates proposition follows logically | Calculate, conclude, deductive, infer |
| Modality | Decreases a propositions certainty value | Could, may, might, ought, should, would |
| Objectivity | Extent to which data 'speaks for itself' | Apparent, appear, imply, indicate, show, suggest |
| Prediction | Judgement about the future | Eventually, expect, forecast, maybe, perhaps, predict, project, reckon, somehow, soon, speculate |
| Probability | Propositions likelihood | Likely, possible, possibility, potential, probable, probably, probability, unlikely |
| Subjectivity | Proposition based on assumptions | Assumptive, belief, believe, connotative, feel, felt, guess, however, presumably, presumptive, think |

In this hedging lexicon (Table 6), many of the words appear to fall into the ‘moderate’ strength category which are neither strong (confident) or weak (hesitant) words. For example, ‘generally’, ‘reckon’ and ‘somewhat’. In everyday parlance, some words in Table 6 may be polysemic: used to discuss the past or present as well as future, such as ‘plan’ and ‘project’. Inclusion of moderate or ambiguous words may smooth or mask subtle opinions at the edges around certainty and hesitancy.

Baker et al. (2012) also highlighted phrases such as ‘have to’, ‘need to’, ‘has to’ and ‘had to’ informally termed ‘semi-modals’ because although they differ syntactically from modal verbs, they share many of the same meaning characteristics. Whilst ‘had to’ does not convey an opinion regarding the future, ‘have to’, ‘need to’ and ‘has to’ may be useful as markers.

Muslu et al. (2015) used three methods in combination to identify forward-looking sentences in financial reports (Table 7).

Table 7 – Words and rules used to identify forward-looking statements (Muslu et al. 2015)

| Word Rule | Description |
|---|---|
| 1a Keywords | Will, future |
| 1b Keyword combination | Combining (next, subsequent, following, upcoming, incoming, coming) and (month, quarter, year, fiscal, period) |
| 2. Verb (including lemma’s) conjugation | Combining (aim, anticipate, assume, commit, estimate, expect, forecast, forsee, hope, intend, plan, project, seek, target) with (we, and, but, do not, company, corporation, management, does not, is, are, not, normally, currently, also) |
| 3. Mention of following year | For example, mentions of ‘2017’ in the 2016 annual report |

The modal verbs (‘should’, ‘would’, ‘can’, ‘could’, ‘may’ or ‘might’) were ignored as they were deemed to be of a ‘legal’ nature rather than forward-looking business opinions. The verb conjugation method (Table 7) was used to avoid false positives by picking up the noun versions of words.

Earlier sections identified how linking word patterns through various markers, to external numerical indices can support the inference of social phenomena (such as over positive business reporting). This section has identified a number of strengths and weaknesses in the dictionaries used for forward-looking (strong and hesitant) sentiment. This raises the possibility that several elements of the relevant KOS in the literature could be combined into an overall model, defined as a "composite KOS" to achieve improved performance for automated forward looking sentiment detection from company report text. This led to the development of the following three research questions:

Q1: Can a composite KOS be created for forward-looking assertive/hesitant sentiment which outperforms existing KOS models?

Q2: Is there an association between the use of strong and/or hesitant forward-looking word frequency and/or diversity and future business performance?

Q3: Do companies in the same industry exhibit different forward-looking word frequency and diversity patterns through time and what explanations could be postulated for those similarities and differences?

## 3. METHOD

A mixed methods Critical Realist (Sayer 2000) philosophy was adopted for this exploratory study. The adoption of a stratified ontology enables the hypothesis of unseen hidden motives inferred through their manifest effects (Wynn and Williams 2012), evidenced through word patterns. Explanations are therefore grounded in the data but are not constrained by empiricism.

The Oil and Gas (O&G) industry was purposefully chosen as it is commodity based (performance is linked to the oil price) and so nuances and differences between multinational companies are more likely to be related to individual strategies and culture rather than market effects. Four large multinationals from the same industry (O&G companies) were selected at random and their annual reports (including 20-F) downloaded from their websites for the years 2008-2015, 32 reports in total. All data is in the public domain, however the four companies are coded ‘Company A’, ‘Company B’, ‘Company C’ and ‘Company D’ to focus on the method and concepts, rather than specific company instances.

A dictionary (lexicon rule based) method was selected for this exploratory study for two reasons. Firstly, it is assumed that forward-looking strong and hesitant sentiment can be well defined by a composite dictionary from those that already exist in the literature. Secondly, for an exploratory study with a relatively small dataset, a statistical machine learning dataset may not be so well suited.

### 3.1 Word categories and types

To avoid smoothing out small patterns, only the categories of ‘strong’ (certain) and ‘hesitant’ (weak) forward-looking sentiment are used in this study to pick up extreme (edge) forward-looking opinions. The selection of extreme edges enables a simple counting polarity based approach, rather than any gradational continuous scale. The composite set is shown in Table 8.

Table 8 – Composite KOS

| Category | Words/concepts |
|---|---|

| | |
|---|---|
| Strong | will, won't, shall, must, certainly, known, definitely, always, is, undoubtedly, believe, has to, have to, need to, commit, aim, expect, anticipate, think, aspire, strive, optimistic, going to |
| Hesitant | might, may, could, unlikely, unknown, uncertain, suggest, sometimes, possibly, possible, perhaps, occasionally, depends, depending, seldom, conceivable, maybe, guess, speculate, hope, imaginably |

Most of the terms from Table 5 (Bodnaruk et al. 2015) were incorporated into the composite KOS, exceptions included the term 'is' which is not always associated with assessments regarding the future. All the absolute and low certainty markers from Table 1 (Thompson et al. 2008) were integrated into the composite along with selected verbs (such as 'hope', 'expect' and 'intend') from Table 7 (Muslu et al. 2015). Various prediction terms (e.g. 'speculate') from Table 6 (Cassidy 2016) were appended to the compoiste KOS with the exceptions of terms that were considered too poylsemic (e.g. project). Semi-modal concepts were included from Baker et al (2012) that were not present in any of the other models. The resultant terms in the 'composite KOS' were used to extract frequencies from the text corpus of company reports used in the study.

### 3.2 Validation with human judges

Accuracy for text classification appears to be typically in the range of 60-90% as a generalization (Jurka et al. 2013; Faith 2011; Sasaki 2008; Magnuson 2014; Miller 2014), although sentiment categorization may be particularly challenging due to the subtle and subjective nature of opinion (Pang and Lee 2008). Some studies such as Li (2010) rely on the researcher to identify sentences that are indicative of the sentiment being analysed. However, as stated by Grimes (2010), relying on a single human assessment of sentiment is likely to lead to bias and it has been reported that human agreement on sentiment is unlikely to be much better than 82% (Wilson et al. 2005).

Therefore, five independent human judges were purposefully recruited from the business social network site LinkedIN www.linkedin.com and personal networks of the researchers. Each was an experienced business professional or academic in an information based discipline, the judges acting effectively as informants due to their knowledge. To ensure some level of diversity, the judges were from Europe and North America, two women and three men.

Each judge was given a random company annual report and asked to identify thirty sentences from any part of the report, which for them represented strong forward-looking sentiment, hesitant forward-looking sentiment and neutral (neither) as a control. Each judge was therefore asked to identify ninety examples in total, to be cut and pasted to a notebook text file, labelled with the category and sent back to the researchers via email. Care was taken not to 'prime' the participants with trigger words to avoid biased data collection.

This exercise generated 450 test examples, with 150 in each of the three categories (strong, hesitant, neither). This was deemed sufficient for the study, based on heuristics for machine learning classification, which indicate that 50-100 labelled training examples are typically required to give good results per category (Hedden 2013; Faith 2011).

The composite dictionary (Table 8) was applied to these training examples to identify recall (how many of the strong and hesitant sentiment examples would be identified) and precision (how many incorrect categorizations were made). An F1 score was calculated (weighted average of precision and recall) which takes into account both false positives and false negatives. An average from these judges scores was used. These data would support a judgement on whether the composite dictionary (KOS) was a reasonable surrogate for determining forward-looking sentiment within company annual reports. Existing dictionary (lexicon) models from the literature were also applied automatically to the judges' examples in order to compare F1 scores to the composite KOS addressing RQ1.

### 3.3 Association between word frequency/diversity and benchmarked performance

From 2009 onwards, company annual reports use new 20-F reporting disclosure rules (Milbank 2009) making it difficult to compare word frequency ratios before and after 2009. Therefore, only data from 2010-2015 were analysed for RQ2.

Scatter plots and Spearman Rank Correlations were used to test for statistically significant associations between strong/hesitant (S/H) word frequency and diversity ratio's compared to subsequent relative business performance (see section 3.4) between 2010-2015 addressing RQ2.

3.4 Calculating word frequency and diversity data

OpenSource utilities including Python NLTK scripts were used to convert the 32 reports into a text corpus of over 5 million words. The composite KOS was automatically applied to that corpus, word frequencies were calculated per category, per company, per year, in order to address RQ3.

Mentions of the future included the explicit term 'future' and mentions of the following year in reports. For example, the mention of 2011 in the 2010 annual report would be a successful match.

Supported by the literature (Jaime and Perez-Guillot 2015; Pique-Angordans et al. 2002; Millar 2009), an assumption was made for this context that word usage is largely monosemic (single meaning). Analysis of concordance data however did parse for negation to remove these false positives from the respective categories. Negation is not straightforward as the phrase "not possible" changes the word 'possible' from hesitant/uncertain to strong/certain whilst the phrase "will not" does not change the certainty of the word 'will'. These natural language parts of speech concept differences were included in the rule set as a crude form of Word Sense Disambiguation (WSD), as a generic Bag of Words (BoW) approach would ignore these contextual differences. Pang and Lee (2008) indicate that catering for negation can improve accuracy by as much as 3%.

The total frequency of strong words ($S_f$) was divided by the total frequency of hesitant words ($H_f$) to create a ratio S/H. Any corresponding increase in assertive words and decrease in hesitant words would therefore be amplified.

A measure of diversity for the use of strong and hesitant words for each company is inspired from biological ecology, assessed using an Equitability ($E_D$) ratio which is the Simpson Diversity Index (Peet 1974) divided by the number of word types per category, given by the equation:

$$E_D = \frac{D}{D_{MAX}} = \frac{1}{\sum_{i=1}^{s} p_i^2} \times \frac{1}{s}$$

Where:

D = Simpson diversity index
S = Total number of species in the community (richness)
$p_i$ = Proportion of S made up by the ith species
$E_D$ = Equitability (evenness)

The $E_D$ ratio is a range between 0 and 1, with 1 representing true evenness. For example, if there were ten types of words (or categories) where individual words occurred 1,000 times, an evenness of 1 would equate to a frequency of 100 for each of the ten word type/categories.

The resulting strong to hesitant (S/H) and evenness ratio ($E_D$) were plotted for each of the four companies for the years (2008-2015).

Linear regression was performed to ascertain whether or not there were any strong associations over time. For small datasets, Moore et al. (2013) suggest an $R^2$>0.7 provides a strong correlation. This was tested for frequency and diversity of word usage over time and with respect to financial performance.

Revenue figures were extracted from each report (including 2005-2007) in order to create a three-year moving average for each company, reflecting percentage revenue change. The four companies were then compared to each other for each three-year window by calculating the percentage change compared to the mean of the four companies, given by the equations below:

$$AR_{cy} = \frac{\sum_{i=1}^{n} R_{cy}}{n}$$

$$ARPC_{cy} = \frac{AR_{cy}}{AR_{c(y-1)}} \times 100$$

$$ARPC\_DM_{cy} = \frac{\sum_{j=1}^{c} ARPC_{jy}}{c}$$

Where:

R = Yearly Revenue (for company c in year y)
n = 3 (size of moving average window)
c = 4 (number of companies studied)
$AR_{cy}$ = Average revenue (3 year moving average for company c in year y)
$ARPC_{cy}$ = Percentage revenue change 3 year moving average (y to y-1)

$ARPC\_DM_{cy}$ = Percentage revenue change above or below the mean for c

The $ARPC\_DM_{cy}$ measure was used as a surrogate for relative business performance amongst the four peers, a form of rolling benchmark.

3.5 Analytical Constructs

Based on the preceding information, the analytical constructs for this exploratory study are as follows:

- Increasing Strong/Hesitant Ratio (Frequency) = Increasing confidence about future.
- Decreasing Strong/Hesitant Ratio (Frequency) =Increasing uncertainty about future.
- Increasing Strong/Hesitant (Diversity) = Potential concealment of information

This was used as the basis for analysis and discussion of the results in section 4.

3.6 Study Limitations

A limitation of the method applied is generally assuming a Bag of Words (BOW) approach and monosemy (single meaning) for the modal verbs. However, natural language context negation was catered for in this study (section 3.4) and the inclusion of some concepts (Baker et al. 2012) rather than just single words, moving the KOS towards a Bag of Concepts (Cambria and Hussain 2015). Nevertheless, limitations of the BoW approach are well documented in the literature (Chan and Chong 2016). It is likely that in some contexts, the same modal verbs may be used with different meaning (polysemy) so would need to be disambiguated. However, supported by existing research, it is assumed the modal verbs analysed in this context have tendencies towards monosemic usage. The method was therefore deemed a valid approach for this exploratory study.

A small sample means statistical generalisation is not possible. However, generalisation of theoretical propositions (analytical generalization) can be proposed (Yin 2003) to stimulate further research.

## 4. RESULTS

Between 2008 and 2015, the total number of words used in annual company reports increased by 44% on average in the sample. Company C showed the greatest increase (71%) followed by Company D (56%), Company B (35%) and Company A (11%).

### 4.1 Evaluating the composite KOS to human judgements

The accuracy and completeness of the automatically applied rules-based composite dictionary (Table 8), compared to the 450 human sentiment judgements is shown in Table 9.

Table 9 – Recall, Precision and F1 scores for the Composite KOS applied to human judgements

| | Strong Forward Looking Sentiment | | | | Hesitant Forward Looking Sentiment | | |
|---|---|---|---|---|---|---|---|
| | **Recall** | **Precision** | **F1** | | **Recall** | **Precision** | **F1** |
| **Judge 1** | 0.89 | 0.89 | 0.89 | | 0.83 | 0.91 | 0.87 |
| **Judge 2** | 0.38 | 0.9 | 0.53 | | 0.92 | 0.92 | 0.92 |
| **Judge 3** | 0.64 | 0.62 | 0.63 | | 0.22 | 0.58 | 0.32 |
| **Judge 4** | 0.93 | 0.62 | 0.74 | | 0.4 | 0.8 | 0.53 |
| **Judge 5** | 0.86 | 0.93 | 0.89 | | 0.93 | 0.97 | 0.95 |
| **AVERAGE** | **0.74** | **0.792** | **0.77** | | **0.66** | **0.84** | **0.74** |

Average accuracy F1 scores for the composite KOS (Table 8) as applied to the judges examples were 77% for strong forward-looking sentiment and 74% for hesitant forward-looking sentiment (Table 9).

### 4.2 Comparing the composite KOS performance to existing models

The model from Thompson et al. (2008) in Table 1 found no examples from the 450 identified by the judges so is not included. The UNC Chapel Hill (2014) in Table 2 was combined with the TeachIT (2006) model in Table 4 as they both relate specifically to modal verbs. The models from Piotti (2014), Cassidy (2016) and Muslu et al. (2015) were not compared directly to the composite KOS, as it was not possible to extract 'strong' and 'hesitant' categories explicitly from these models without a significant amount of interpretation. The model from Muslu et al. (2015) depends on identifying the word 'future' or an associated time period (such as next quarter). The exclusion of modal verbs such as 'could', 'may' or 'might' is likely to have had a significant impact on recall. Table 10 shows the model comparison.

Table 10 – Comparison of KOS from literature to the composite KOS (Table 8). The green cells show where the composite KOS out-performed the existing models.

| | UNC Chappel Hill (2014) and TeachIT (2006) | | | | | | Bodnaruk, Loughlan and McDonald (2015) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | STRONG | | | HESITANT | | | STRONG | | | HESITANT | | |
| | Recall | Precision | F1 | Recall | Precision | F1 | Recall | Precison | F1 | Recall | Precison | F1 |
| Judge 1 | 0.69 | 0.9 | 0.781 | 0.83 | 1 | 0.907 | 0.72 | 0.66 | 0.689 | 0.83 | 1 | 0.907 |
| Judge 2 | 0.28 | 1 | 0.438 | 0.94 | 0.8 | 0.864 | 0.36 | 0.89 | 0.513 | 0.91 | 0.86 | 0.884 |
| Judge 3 | 0.36 | 0.82 | 0.5 | 0.25 | 0.73 | 0.372 | 0.6 | 0.4 | 0.48 | 0.19 | 0.66 | 0.295 |
| Judge 4 | 0.86 | 0.7 | 0.772 | 0.4 | 0.8 | 0.533 | 0.93 | 0.62 | 0.744 | 0.4 | 0.8 | 0.533 |
| Judge 5 | 0.5 | 1 | 0.667 | 0.9 | 0.97 | 0.934 | 0.63 | 0.86 | 0.727 | 0.9 | 1 | 0.947 |
| **AVERAGE** | **0.54** | **0.884** | **0.67** | **0.66** | **0.86** | **0.75** | **0.65** | **0.686** | **0.631** | **0.65** | **0.864** | **0.713** |
| **Compared to composite KOS (from Table 9)** | -0.2 | 0.092 | **-0.1** | 0.004 | 0.02 | **0.01** | -0.09 | -0.106 | **-0.14** | -0.01 | 0.024 | **-0.03** |

The pure modal verb (UNC Chapel Hill 2014) and TeachIT (2006) models performed slightly better for hesitant sentiment than the composite KOS (Table 10). One of the reasons was the modal verb 'would' which was omitted in the composite KOS as it was deemed half way between 'strong' and 'hesitant'.

For strong forward-looking sentiment, the composite's F1 score represented a 0.1 to 0.14 (Table 10) improvement over existing models. From analysis, the terms 'intend' and 'seek' may be useful additions to the strong category for forward-looking sentiment.

When applied to the corpus of 32 reports, the semi-modals such as "have to" accounted for 3% of all matches from the composite KOS (Table 8).

4.3 Correlating word patterns to future financial performance

Performing a Spearman Rank Correlation, there was no statistical association between the previous years S/H word frequency ratio and subsequent relative business performance for any company. For S/H diversity ratio's and subsequent relative business performance, Company C (r=-.857 and p=0.0137) showed a statistically significant association. No other statistically significant associations were identified.

4.4 Word frequency and diversity similarities and differences amongst companies

The strong/hesitant (S/H) word ratio's and diversity from Table 8 are shown in Figure 1.

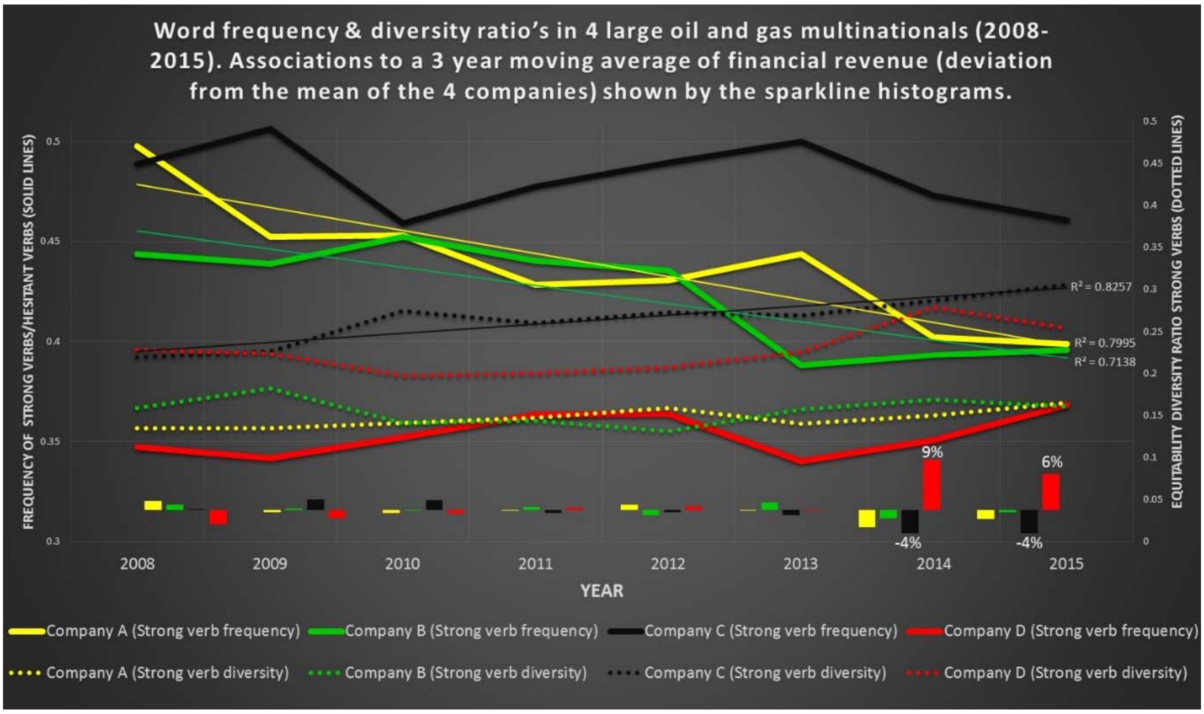


Figure 1 – Changes in the use of word ratios (Strong/Hesitant word frequency - solid lines and diversity – dotted lines) for four O&G companies (2008-2015). Histogram sparkline shows deviation from the mean for the 4 companies using a three year moving average revenue percentage change.

A thick solid line was used for the S/H frequency ratio and dotted line for the S/H ($E_D$) ratio. Strong linear regression trend lines are shown in a faint solid line with their $R^2$ on the right-hand side of Fig 1.

The Strong to hesitant (S/H) forward-looking word frequency ratio's (solid lines) for Company A (yellow) and Company B (green) showed a strong declining relationship with time (2008-2015) as shown in Figure 1. Assertive language reduced, with statistically significant correlations ($R^2$ of 0.79 and 0.71 respectively). Company C (black) showed the highest S/H ratio's and Company D (red) showed the lowest S/H ratio's in the time period studied.

The equitability diversity (evenness) ratio ($E_D$) for Company C (black dotted line) showed a strong increasing relationship with time, assertive language becoming more diverse ($R^2$=0.82). The diversity ratios for Company A and B (yellow/green dotted lines) were similar to each other, lower and remained virtually the same over the entire time period studied. The diversity of strong words for Company D (red dotted line) dissected the other companies (Figure 1).

At the base of Figure 1, the sparkline histogram ($ARPC\_DM_{cy}$) shows that in 2008, Company A (yellow) and Company B (green) performed above the average (of the four companies) with respect to revenue changes in a three-year moving average. Company C was neutral and Company D (red) was the worst performer. By 2014/2015, in relative terms, Company D (red) had become the top performer (9% and 6% above the average) with Company C the worst (4% below the average revenue percentage change).

Figure 2 shows the results of counting the mentions of 'future' or mentions of the following year in the annual reports of the 4 companies between 2008 and 2015.

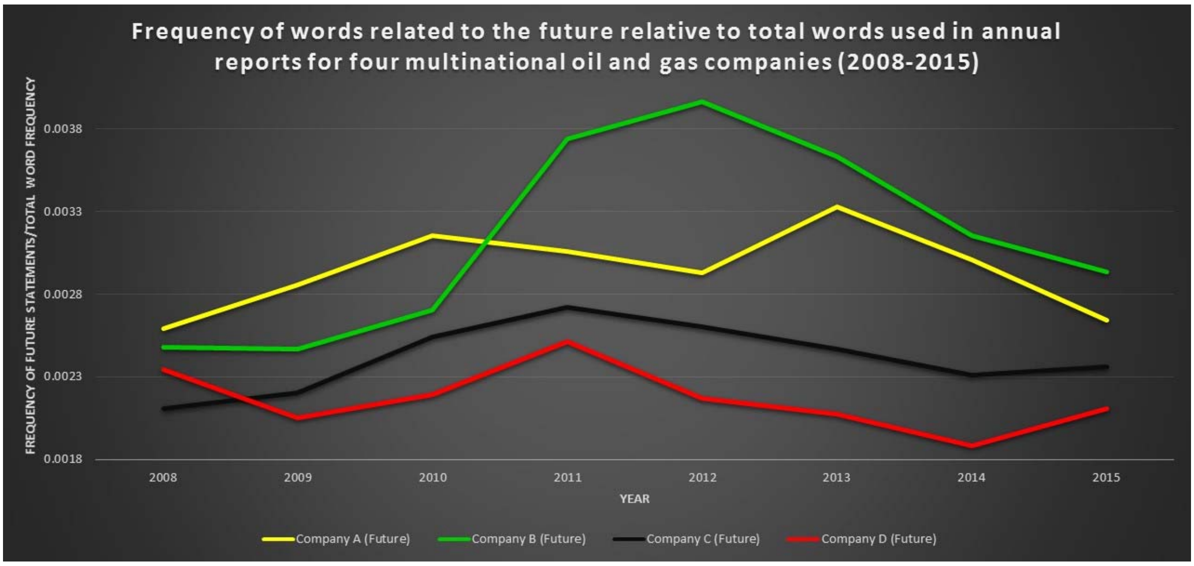


Figure 2 – Relative frequency of words mentioning 'the future' for the 4 companies over time

Company A, C and D appeared to have relatively uniform changes (although nothing statistically significant), whilst Company B exhibited a sharp gradient change from 2010 to 2011, with a 38% increase in mentions of the future. As part of the iterative process of discovery, further data were inductively collected on word frequency for Company B to explore this gradient change around 2010-2011. It was found that the concept of learning (represented by the stems 'learn' and 'lesson') also showed steep gradient changes in the previous year (Figure 3).

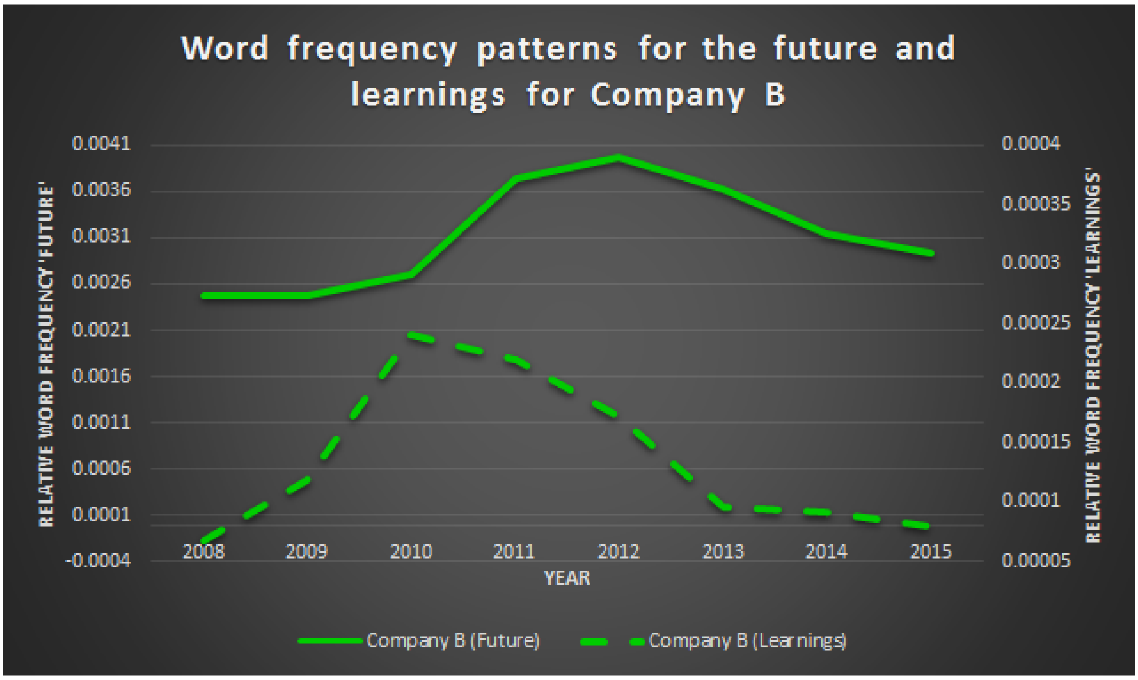


Figure 3 – Relative word frequencies for 'learning' compared to 'future' (from Fig 2) for Company B

For the 'learning' concept, the major increase in word frequency gradient change (over 100%) occurred from 2009-2010 (dotted line Figure 3), prior to the increases in the 'future' concept between 2010-2011 (solid line Figure 3). Potential mechanisms are discussed in section 5.

## 5. DISCUSSION

The average increase in the total number of words in annual reports between 2008 and 2015 was 44%, roughly in line with the 50% figure stated in the literature (Deloitte 2015).

When applied to the examples provided by the judges, the composite KOS created by analysing existing models (Table 1-6) generated F1 scores of 0.77 for strong and 0.74 for hesitant, forward-looking sentiment (Table 10), close to human levels of agreement. Some similar statements containing the word ‘expects’ were highlighted by judges as both examples of hesitant and also examples of strong forward-looking sentiment, supporting the findings of Wilson et al. (2005) that disagreement on sentiment is common. There was an improvement on existing models (Table 10) for strong forward-looking sentiment. Furthermore, none of the existing models identified (Tables 1-6), included appropriate semi-modal verbs (such as ‘have to’, ‘need to’, ‘has to’) so would have been projected to miss over 3% (section 4.2) of occurrences of assertive strong forward-looking sentiment when applied to the corpus of 32 reports and 5 Million words.

Contrary to the assertion made by Muslu et al. (2015) that modal verbs such as ‘could’ and ‘may’ were only used in a legal sense, many examples were provided by the human judges where they were deemed markers for forward-looking business sentiment.

It is therefore proposed that the final composite list of words (Table 8) adequately represents forward-looking sentiment and indicates that improvements can be made on existing models, addressing RQ1.

Company C exhibited a statistically significant association between increasing diversity of strong assertive forward-looking language and subsequent future decreasing relative business performance. No other company exhibited this association in the sample. Addressing RQ2, it is believed that this marker has not been reported before in the literature and presents an area for further research.

Addressing RQ3, both Company A and Company B showed a strong association over the period 2008-2015 with a decline in the use of strong words/assertive language about the future (Figure 1). It is possible that this reflected increasing business uncertainty. This could be explained through a narrative which ties word frequencies to global events. The financial crisis had just occurred (2007/2008) and the oil price had fallen to its lowest level for four years (end 2008). Whilst the oil price rose again, by 2015 it was less than half what it was in 2012. One explanation is that these two companies accurately assessed long term market trends and reflected these in their use of language in the annual report.

Mentions of the 'future' relative to total words in annual reports showed some major deviations between companies through time (Figure 2). In particular, the 38% increase in mentions of the 'future' by Company B from 2010-2011, where its frequency of word use had not dropped back to prior levels even five years after the initial rise. The annual reports and historical news archives provide evidence for a catastrophic event (crisis) that may have triggered these patterns. A crisis in this context is defined as, "specific, unexpected, and non-routine events or series of events that created high levels of uncertainty and threat or perceived threat to an organization's high-priority goals" (Seeger et al. 1998, 231).

In 2010 Company B was involved in a major industrial accident (crisis) that received significant news coverage. This may have influenced company attitudes and rhetoric to focus on the 'future' as a vision to move forward. The organization may have been participating in a future-based 'developmental conversation' with stakeholders, as a form of impression management, "While we can't ignore the past, we also can't change it. We can learn from it, but we shouldn't dwell on it" (Levin and Edwards 2007, 155). This proposition is supported by the increases in the word frequency of the 'learnings/lessons' concept (Figure 3). The word frequency patterns of the 'future' and 'learnings' concepts could support a narrative based on Discourse of Renewal Theory (DRT) as proposed by Ulmer et al. (2011). As part of an organizational rhetorical framework in a time of crisis, DRT focuses on renewal, growth and transformation. Reflecting on a crisis, it describes sense-making which contains a 'learning' component to gain confidence from stakeholders and then providing a 'future' prospective vision for moving forward as a response to a crisis. The word patterns observed and sequencing of them (Figure 2 and Figure 3) may support this theory.

Company C's use of strong forward-looking assertive language regarding the future was four times that of Company D which was the top performer in the sample by 2015. One explanation is that modal verb usage is influenced by nationality rather than corporate culture (Hykkes 2000), although differences in professional business contexts have been dismissed as negligible (Nathan 2010). A competing explanation is that Company C deploys more optimistic rhetoric in its annual report, supporting the Pollyanna effect (Hildebrandt and Snyder 1981). Due to greater consilience, this is proposed as the most plausible explanation based on the evidence collected in this study that links a number of markers for Company C to over-positive reporting. Four lines of evidence support this explanation:

(i) the greatest increase in company report word length from 2008 to 2015 (71%), over 6 times the increase of Company A for the same period. It has been reported that there is a tendency for more words to be used in reports that are trying to conceal or obscure information (Bodnaruk et al. 2015; Minhaus and Hussian 2016).

(ii) the highest frequency in the sample (Figure 1) of strong forward-looking sentiment modal words – a potential sign of material weakness in controls (Loughran and McDonald 2011).

(iii) the highest diversity in the sample (Figure 1) of forward-looking strong sentiment – a potential indicator of deception (Siegel et al. 2013).

(iv) a statistically significant association between increasing diversity of strong forward-looking sentiment and subsequently decreasing business performance (Section 4.3).

From Figure 1 (dotted lines) it can be seen that Company A and Company B showed a lower diversity of assertive language regarding the future than Company C and D. Company A and B may represent 'middle of the road' ideologies in their annual reports. They have neither the extreme use of strong words or diversity shown by Company C nor the minimal use shown by Company D.

Company D went from being the worst performer in the group to the best (Figure 1), despite having the lowest ratio through time for assertive/hesitant language. No strong correlations existed for Company D between word frequency/diversity and business performance, contradicting Yuet-yung

(2014) who suggested top performers would be more assertive in their presentation of future outcomes. One explanation could be the study sampling, which analysed companies that occupy the top echelons of financial market indices. Relative differences in word use between these companies may indicate other generative mechanisms, such as organizational cultural differences, rather than outright fraud/deception extremes or poor performance, that previous studies have focused on.

The findings from the study imply that companies may exhibit tendencies to behave in different ways (archetypes) rather than obey 'universal laws'. Figure 4 proposes a model based on the study findings.

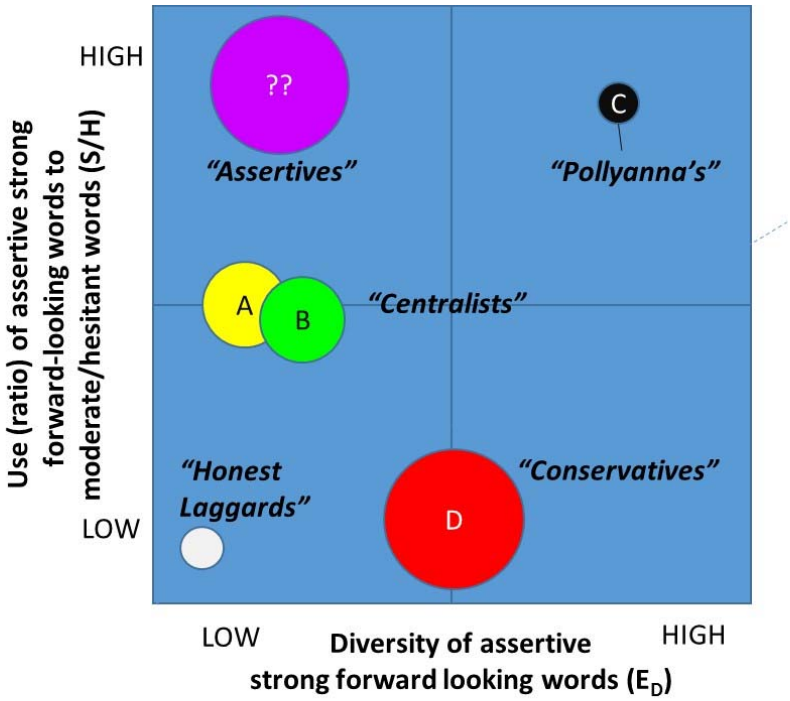


Figure 4 – A proposed model for differing underlying company cultural 'norms' inferred by frequency/diversity ratios. The size of the circle is relative business performance.

It may be possible for companies to 'shift' archetypes as their culture or intent changes. Company A and B are termed Centralists in that they occupy the middle ground compared to their peers for use of strong assertive forward-looking language and diversity related to performance. Company D is termed a Conservative, which is characterized by low use of assertive strong forward-looking words which has remained relatively unchanged, even when it performed well compared to its peers. This may evidence a more complex and perhaps cautionary or academic approach to communicating the future state of affairs than its peers. Company C is proposed as a Pollyanna where its high use of diverse assertive strong forward-looking language along with other markers, may not be justified by its benchmark performance. It may fit extreme impression management (Rutherford 2005).

Other postulated archetypes are Assertives (purple) which show high frequency of strong word use and superior business performance and Honest Laggards (white) - low performing companies that exhibit this through their hesitant word patterns. There is no explicit empirical evidence from this study for the existence of these two archetypes, presenting an area for further research.

## 6. CONCLUSION

This study has shown how KOS can be used to automatically detect 'mentions' of inferred intent within reports, as opposed to the traditional use of KOS to classify reports as a whole. There is evidence that the composite KOS developed from the existing literature may be an improvement over existing single models in the literature. This demonstrates the value of taking a 'composite' based methodological approach, integrating, testing and exploiting multiple KOS in the literature (rather than simply taking an existing model) for automated sentiment analysis, making a modest contribution in these fields. Applying more context disambiguation and Bayesian statistics may improve accuracy further and presents an area for further research.

No previous studies make an association between increasing diversity of words within forward-looking sentiment categories in reports and corresponding decreasing business performance. The findings from this initial study may therefore act as a catalyst to explore this marker in more detail.

Using proxy dictionaries to automatically discern word patterns in text is not new. However, it may be an increasingly useful epistemological tool in the big data, and post-truth society. Enabling the re-presentation of text which can facilitate the development of a non-obvious narrative.

Adopting Critical Realism as a philosophy when analysing ‘big data’ within organizations, may steer the practitioner towards explanation, mechanisms and tendencies rather than simply looking for statistical regularities. This may lead to theory generation and the potential for differentiating insights which go beyond regularities within the data.

Sentiment is typically applied with an a priori hypothesis in mind. Embedding these sentiment algorithms in standard enterprise search and discovery technology deployments may help facilitate the generation of differentiating insights and new knowledge from the most unexpected of places.

## REFERENCES


Aaltonen, Aleski and Niccoló Tempini 2014. Everything counts in large amounts: A critical realist case study on data-based production. *Journal of Information Technology*, 29(1): 97-110.

Baker, Kathryn et al. 2012. Modality and Negation in SIMT Use of Modality and Negation in Semantically-Informed Syntactic MT. *Computational Linguistics*, 38(2): 411-438.

Berelson, Bernard R. 1952. Content Analysis in Communication Research. *The Free Press*, Glencoe, IL.

Bholat, David et al. 2015. Text mining for central banks: handbook. *Centre for Central Banking Studies* (33): 1-19.

Blei, David M. et al. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3: 993–1022.

Bloom, Benjamin S. et al. 1956. *Taxonomy of educational objectives, handbook I: The cognitive domain*. New York: David McKay Co Inc.

Bloomfield, Robert J. 2002. The 'incomplete revelation hypothesis' and financial reporting. *Accounting Horizons*, 16: 233–243.

Bochkay, Khrystyna and Valentin Dimitrov 2014. Qualitative Management Disclosures and Market Sentiment (December 15, 2014). Rutgers University. Online Article: (https://ssrn.com/abstract=2538812, accessed March 2017).

Bodnaruk, Andriy et al. 2015. Using 10-K Text to Gauge Financial Constraints. Online Article (https://www2.warwick.ac.uk/fac/soc/wbs/subjects/finance/events/seminars/using_10-k_text_to_gauge_financial_constraints.pdf, accessed April 2017).

Boulton, Jean G. et al. 2015. *Embracing Complexity: Strategic perspectives in an age or turbulence*. USA: Oxford University Press.

Bozanic, Zahn et al. 2015. Management Earnings Forecasts and Other Forward-Looking Statements. Online Article (https://ssrn.com/abstract=2130145, accessed February 2017).

Brown, Tracey 2016. Evidence, expertise, and facts in a "post-truth" society. Online Article (http://www.bmj.com/content/355/bmj.i6467, accessed August 2017).

Cambria, Erik and Amir Hussain 2015. *Sentic Computing. A Common-Sense Framework for Concept-Level Sentiment Analysis*. Springer, Switzerland.

Cassidy, Caitlin 2013. Between the hedges: A Computational analysis of sentiment and linguistic hedging in financial documents. MSC Thesis. Online Article (https://www.ai.uga.edu/sites/default/files/theses/cassidy_caitlin_n_201505_ms.pdf, accessed January 2017).

Chan, Samuel WK and Mickey WC Chong 2016. Sentiment analysis in financial texts. *Decision Support Systems*. Online Article (http://dx/doi.org/10.1016/j.dd.2016.10.006, accessed June 2017).

Cleverley, Paul H and Simon Burnett 2015. The Best of Both Worlds: Highlighting the Synergies of Combining Manual and Automated Knowledge Organization Methods to Improve Information Search and Discovery. *Official Journal of the International Society for Knowledge Organization (ISKO*), 42(6): 428-445.

Davenport, Thomas H. and Laurence L Prusak 2000. *Working Knowledge: How Organizations Manage What They Know*. USA. Harvard Business School Press.

Davenport, Thomas H. 2013. Analytics 3.0. *Harvard Business Review*, 19(12): 64-72.

Deloitte 2015. Deloitte UK Annual Report Insights. Online Article (https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/audit/deloitte-uk-annual-report-insights-2015-full-survey.pdf, accessed May 2017).

El-Haj, Mahmoud 2014. Analysing UK Annual Report Narratives using Text Analysis and Natural Language Processing. Online Article (http://www.lancaster.ac.uk/staff/elhaj/docs/GlasgowTalk.pdf, accessed January 2017).

EOI 2012. Modal Verbs. Online Article (http://www.eoisabi.org/wp-content/uploads/2012/06/modality-modal-verbs.pdf, accessed January 2017).

Faith, Ashleigh. 2011. Linguistically Training Automatic Indexing Software for Complex Taxonomies. Semantic Technology & Business Conference, June 14-18 2013, San Jose, CA, USA.

Ibekwe-Sanjuan, Fidelia, and Geoffrey Bowker 2017. Implications of big data for knowledge organization.. Knowledge Organization, Ergon Verlag, 2017, *Special issue on New trends for Knowledge Organization*, Ed., Souza, R., (guest editor), 44 (3): 187-198.

Fisher, Ingrid E. et al. 2008. The Role of Text Analytics and Information Retrieval in the Accounting Domain. *Journal of Emerging Technologies in Accounting*, 7: 1-24.

Floridi, Luciano 2014. *The Fourth Revolution. How the Infosphere is reshaping human reality*. UK: Oxford University Press.

Gantz, John and David Reinsell 2011. Extracting Value from Chaos: Report ID 1142. [online]. International Data Corporation (IDC). Online Article (https://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf, accessed January 2014).

Grimes, Seth 2010. Expert Analysis: Is Sentiment Analysis an 80% Solution? InformationWeek. Online Article (http://www.informationweek.com/software/information-management/expert-analysis-is-sentiment-analysis-an-80--solution/d/d-id/1087919, accessed February 2017).

Grimes, Seth 2012. From Sentiment Analysis to Enterprise Applications. Greenbook Online Article (http://www.greenbookblog.org/2012/01/02/from-sentiment-analysis-to-enterprise-applications/, accessed July 2017).

Gupta, Aparna and Haochen Liu 2017. Addressing the Risk Culture Challenge in Banking using Text Analytics. European Financial Management Association (EFMA), Annual Meeting June 28-July 1 Athens, Greece.

Hajek, Petr and Vladimir Olej 2013. Evaluating Sentiment in Annual Reports for Financial Stress Prediction using Neural Networks and Support Vector Machines. EANN 2013, Part II, CCIS 384: 1-10.

He, Wu et al. 2017. Managing extracted knowledge from big social media data for business decision making. *Journal of Knowledge Management*, 21(2): 275-294.

Hedden Heather 2013. Taxonomies for Auto-Tagging Unstructured Content. Text Analytics World, October 1 2013, Boston USA.

Hildebrandt, Herbert W. and Richard D Snyder 1981. The Pollyanna Hypothesis in Business Writing: Initial Results, Suggestions for Research. *The Journal of Business Communication*, 18(1): 5-15.

Hinkel, Eli 1995. The Use of Modal Verbs as a Reflection of Cultural Values. In *TESOL Quarterly*. Ed McKay et al. 29(2): 100-118.

Humpherys, Sean L. et al. 2011. Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems*, 50(3): 585-594.

Hykes, Jenny M. 2000. A comparison of the use of modal verbs in research articles by professionals and non-native speaking graduate students. MA Thesis Online Article (http://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=8928&context=rtd, accessed January 2017).

Jaime, Asuncion and Cristina C Perez-Guillot 2015. A comparison analysis of modal auxiliary verbs in Technical and General English. Procedia – Social and Behavioural Sciences, 212: 292-297

John, Oliver P. 1990. The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In LA Pervin (Ed.), Handbook of personality: Theory and research . New York: Guilford.

Jurka, Timothy P. et al. 2013. RTextTools: A Supervisory Learning Package for Text Classification. *The R Journal*, 5(1): 6-12.

Khan, Zaheer and Tim Vorley 2017. Big data and text analytics: an enabler of knowledge management. *Journal of Knowledge Management*, 21(1): 18-34.

Kloptchenko, Antonina et al. 2004. Mining Textual Contents of Financial Reports. *The International Journal of Digital Accounting Research*, 4(7): 1-29.

Krippendorff, Klaus 2004. *Content Analysis. An Introduction to its Methodology*. 3rd Ed. Sage, 2013, USA.

Kruschwitz, Udo and Charlie Hull 2017. "Searching the Enterprise", *Foundations and Trends in Information Retrieval*, 11(1): 1-142.

Levin, Don and Terry Edwards 2007. *The leader coach: Exposing your soul*. Authorhouse, USA.

Li, Feng 2006. Do stock market investors understand the risk sentiment of corporate annual reports? Working paper. Online Article (http://www.cis.upenn.edu/~mkearns/finread/sentiment.pdf, accessed January 2017).

Li, Feng 2010. The Information Content of Forward-Looking Statements in Corporate Filings – A Naïve Bayesian Machine Learning Approach. *Journal of Accounting Research*, 48(5): 1049-1102.

Loughran, Tim and Bill McDonald 2011. When is a liability not a liability? Textual analysis, dictionaries and 10-K's. *The Journal of Finance*, 66(1): 35-65.

Magnuson, Doug 2014. Auto Classification and the Holy Grail for Records Managers. IBM Presentation as the Association or Records Managers and Administrators (ARMA), Houston, USA.

Malhotra, Ashutosh et al. 2013. 'HypothesisFinder:' A strategy for the Detection of Speculative Statements in Scientific Text. *PLoS Computational Biology*, 9(7). Doi 10.1371/journal.pcbi.1003117

Manning, Christopher D. et al. 2008. *Introduction to Information Retrieval*. Cambridge University Press.

Michel, Jean-Baptiste et al. 2011. Quantitative analysis of culture using millions of digitized books. *Science*, 331(6014): 176-182.

Mikolov, Tomas et al. 2013. Distributed representations of words and phrases and their compositionality. *Advanced in Neural Information Processing Systems*: 3111-3119.

Milbank 2009. Client Alert. Online Article (https://www.milbank.com/images/content/6/3/634/012609-Changes-to-SEC-Form-20F-Reporting-Obligations.pdf, accessed January 2009).

Millar, Neil 2009. Modal verbs in TIME: Frequency changes 1923-2000. *International Journal of Corpus Linguistics*, 14(2): 191-220.

Miller, David 2014. Just the facts: Auto-classification and Taxonomies. ConceptSearching Webinar, Online Article (https://www.conceptsearching.com/just-the-facts-auto-classification-and-taxonomies/, accessed July 2017).

Minhas, Saliha and Amir Hussain 2016. From Spin to Swindle: Identifying Falsification in Financial Text. *Cognitive Computing*, 8: 729-745.

Moore, David S. et al. 2013. *The basic practice of statistics*. 6th Ed. New York. W.H. Freeman and Company.

Muslu, Volkan et al. 2012. Forward-Looking Disclosures and the Information Environment. *Management Science*, 61(5): 931-948.

Narayanan, Ramanathan et al. 2009. Sentiment analysis of conditional sentences. Proceedings of the Conference on Empirical Methods in Natural Language Processing, August 6-7 2009: Singapore, 1(1): 180-189.

Nathan, Philip B. 2010. A Genre-based study of Pedagogical Business Case Reports. PhD Thesis. Online Article (http://etheses.bham.ac.uk/711/1/Nathan10PhD.pdf, accessed January 2017).

Pang, Bo and Lillian Lee 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2): 1-135.

Patton, Michael Q. 2015. *Qualitative Research & Evaluation Methods*. 4th ed. USA: Sage.

Pauleen, David et al. 2015. In Bed with Technology? Peril, Promise and Prudence. *Communications of the Association for Information Systems*, 37(38): 783-796.

Peet, Robert K. 1974. The Measurement of Species Diversity. Annual Review of Ecology and Systematics, 5(1-463). Online Article (http://www.annualreviews.org/doi/abs/10.1146/annurev.es.05.110174.001441?journalCode=ecolsys.1, accessed January 2017).

Piotti, Sonia R. 2014. *Exploring Corporate Rhetoric in English: Hedging in Company Annual Reports: A corpus assisted analysis*. EDUCatt: Milan, Italy.

Pique-Angordans, Jordi et al. 2002. Epistemic and Deontic Modality: A Linguistic Indicator or Disciplinary Variation in Academic English. *LSP & Professional Communication*, 2(2): 49-65.

Pulman, Stephen 2014. Multi-Dimensional Sentiment Analysis. University of Oxford and TheySay. Online Article: (http://www.theysay.io/wp-content/uploads/2016/05/SGPulman-LT-Accelerate-Brussels2014.pdf , accessed January 2017).

Reagan, Andrew J. et al. 2015. Benchmarking sentiment analysis methods for large-scale texts: a case study for using continuum-scored words and word shift graphs. Cornell University Library. Online Article (https://arxiv.org/abs/1512.00531 , accessed June 2017).

Rutherford, Brian A. 2005. Genre Analysis of Corporate Annual Report Narratives. *Journal of Business Communication*, 42(4): 349-378.

Ruthven, Ian 2008. Interactive Information Retrieval. *Annual Review of Information Science and Technology*. American Society for Information Science and Technology, 42(1): 43-91.

Sasaki, Yutaka 2008. Automatic Text Classification. University of Manchester. Online Article (http://www.nactem.ac.uk/dtc/DTC-Sasaki.pdf, accessed June 2017).

Sayer, Andrew 2000. *Realism and Social Science*. London, UK: Sage.

Seeger, Matthew W. et al. 1998. Communication, Organization and Crisis. *Annals of the International Communication Association*, 21: 231-276.

SentiWordNet 2010. Online Article (http://sentiwordnet.isti.cnr.it/, accessed January 2017).

Siegel, Jay A. et al. 2013. *Encyclopaedia of Forensic Sciences*, 2nd Edition. Elsevier, USA.

Smith, Paul 2015. Woodside Petroleum searches for data value with IBM's Watson cognitive computing. [online]. Australia: Financial Review. Fairfax Media Publications Pty Ltd. Online Article(http://www.afr.com/technology/woodside-petroleum-searches-for-data-value-with-ibms-watson-cognitive-computing-20150521-gh6un7, accessed November 2015).

Taboada, Maite et al. 2011. Lexicon-based methods for sentiment analysis. *Computational Linguist*. 37: 267–307.

Taboada, Maite 2015. Sentiment Analysis: An Overview from Linguistics. *Annual Review of Linguistics*. 2: 325-347.

Tausczik, Yia R and James W Pennebaker 2010. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, 29(1): 24-54.

TeachIT 2016. Modal verbs and Adverbs. Online Article (www.teachit.co.uk, accessed January 2017).

Tetlock, Paul C. 2007. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, 62: 1139–1168.

Tetu, Louis 2016. *Creating the Intelligent Workplace: Think Proficiency not Efficiency*. Enterprise Search Summit 14-17 November 2016: Washington DC, USA. Online Article (http://www.enterprisesearchanddiscovery.com/2016/presentations.aspx, accessed December 2016).

Thompson, Paul et al. 2017. Categorising Modality in Biomedical Texts. Online Article (http://personalpages.manchester.ac.uk/staff/paul.thompson/papers/LREC_modality_2008_final.pdf, accessed June 2017).

Ulmer, Robert R. et al. 2011. *Effective Crisis Communication: Moving from Crisis to Opportunity*, 2nd Edition. Sage, USA.

UNC Chapel Hill 2014. Modal Verbs. Online Article (http://writingcenter.unc.edu/handouts/modals/, accessed January 2017).

University of Georgia 2012. What is Content Analysis? Online Article (https://www.terry.uga.edu/management/contentanalysis/research/, accessed August 2017).

Van Boeyen, Scott 2014. Why Sentiment analysis engines need customization. Online Article (http://www.techradar.com/news/software/business-software/why-sentiment-analysis-engines-need-customization-1256701, accessed July 2017).

Wang, Chuan-Ju et al. 2013. Financial Sentiment Analysis for Risk Prediction. International Joint Conference on Natural Language Processing, Nagoya, Japan 14-18 October 2013: 802-808.

Wilson, Theresa et al. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, Oct 6-8 2005, Vancouver, Canada*: 347-354.

WordNet 2010. Princeton University. Online Article (http://wordnet.princeton.edu, accessed January 2017).

Wynn, Donald Jr and Clay K Williams 2012. Principles for Conducting Critical Realist Case Study Research in Information Systems. *Management Information Systems (MIS) Quarterly*, 36(3): 787-810.

Yasumasa, Someya. 2008. Modal Verbs and Their Semantic Functions in Business English. *Aoyama Journal of Business* (Aoyama Keiei Ronshu), 44(3).

Yin, Robert K. 2003. *Case Study Research: Design and Methods*. 3rd ed. Thousand Oaks, CA, USA: Sage.

Yuet-yung, Siu 2014. Forward-looking statements in annual reports: how is futurity expressed? MA Thesis Online Article (https://hub.hku.hk/bitstream/10722/207135/1/FullText.pdf, accessed March 2017).

Zeng, Marcia et al. 2007. Knowledge organization systems (KOS) standards. *Proceedings of the Association for Information Science and Technology*, 44(1): 1-3.