JURNAL TATA KELOLA DAN AKUNTABILITAS KEUANGAN NEGARA

Check for updates

p-ISSN 2460-3937
e-ISSN 2549-452X

Volume 10 Number 2, 2024: 191-213

https://jurnal.bpk.go.id/TAKEN

Using LDA for audit risk assessment of the Indonesian BOS fund: Insights from news analysis

Iis Istianah[1], Nia Pramita Sari[1], Afrialdi Syahputra Butar Butar[2], Bonar Cornellius Pasaribu[3]

Faculty of Economics and Business, Gadjah Mada University, Yogyakarta, Indonesia[1]

The Audit Board of the Republic of Indonesia, Jakarta, Indonesia[2]

Master of Public Policy and Management, University of Melbourne, Melbourne, Australia[3]

ABSTRACT

This study explores the implementation of text mining in audit risk assessment. We use the latent Dirichlet allocation (LDA) algorithm to reveal hidden topics representing risks in the management of the Indonesian School Operational Assistance Fund (BOS Fund). Using 1,460 news data points from a leading Indonesian news portal, this study proves that using text mining with the LDA algorithm effectively identifies the risks of an audit object. This study makes two important contributions to the information systems and audit literature. First, it provides evidence from online news archives to facilitate a more reliable, current, and comprehensive selection of potential audit areas by encompassing evolving social realities and facts. In the contemporary era, the accelerated and precise dissemination of information via the Internet renders the LDA approach feasible and prudent. Second, it provides a practical and applicable framework for audit risk assessment using nonfinancial sources from independent parties, which can be used as a guide for the development of audit models in the public and private sectors.

KEYWORDS:

Risk-based audit; text mining; topic modeling; LDA; BOS Fund

HOW TO CITE:

Istianah, I., Sari, N. P., Butar, A. S. B., & Pasaribu, B. C. (2024). Using LDA for audit risk assessment of the Indonesian BOS fund: Insights from news analysis. *Jurnal Tata Kelola dan Akuntabilitas Keuangan Negara*, *10*(2), 191-213. https://doi.org/10.28986/jtaken.v10i2.1803

*Corresponding author's
Email: iisistianah1981@mail.ugm.ac.id

ARTICLE HISTORY:

Received	: 2 October 2024	Accepted	: 23 September 2024
Revised	: 31 October 2024	Published	: 26 December 2024

INTRODUCTION

This study explores the implementation of text mining, specifically topic modeling with latent Dirichlet allocation (LDA) algorithms on online news, to facilitate auditors in assessing risks surrounding their audit objects. Although big data analytics has developed immensely in auditing and a noticeable implementation of risk-based auditing (RBA), the application of topic modeling using the LDA algorithm on online news to assist auditors in a risk-based audit is underexplored. Hence, our interest is in exploring the use of the LDA algorithm on online news in audit risk assessment. The use of the LDA algorithm in topic modeling has been effective in discovering latent data in text documents (Blei, 2012). Accordingly, its application to online news is expected to reveal potential areas with higher risks, thus benefiting auditors in conducting a risk-based audit. This study aims to evaluate the effectiveness and demonstrate the use of topic modeling with the LDA algorithm to identify issues related to an audit subject.

The development of RBA emphasizes the importance for auditors to focus their resources on examining areas exposed to higher risks (Nugraheni & Pamungkas, 2021) to minimize failures in achieving audit objectives (ACCA, 2008). One of the crucial steps in the RBA is defining areas with risks or potential risks through the analysis of news and social media (BPK, 2020). This study is motivated by the development of online news in the digital era, which presents various growing issues in society at a near real-time speed. Another motivation lies in the massive development of big data analytics that helps alter unstructured data, such as news media information, into a structured format. These developments are beneficial for auditors in identifying potential areas for risk-based audits.

Despite the growing body of research on RBA, there remains a significant gap in the examination of the use of media as an essential source of information in the RBA process. To assess risk, existing research generally focuses on internal data, regulatory requirements, and financial information (Arens et al., 2021). However, dynamic external risks, such as political instability, regulatory changes, and market fluctuations often communicated through various media platforms, are underexplored. While the role of the media in disseminating risk-related information has been recognized in broader risk communication research (Perko, 2012), its integration into the audit process, particularly RBA, has not been adequately addressed.

Previous literature tends to consider traditional risk factors without incorporating real-time data from news reports, social media, and other media outlets that provide early warning and insight into emerging risks (Wendling et al., 2013). This gap highlights the potential benefits of further research into how media can elevate risk identification and audit management, mainly by providing auditors with timely external data that can influence audit strategy and decision-making processes. The integration of media into the audit process can create new avenues for risk identification and management, sparking optimism for the future of RBA.

Several accounting studies have proven the effectiveness of the text mining approach in analyzing textual data (e.g., Boskou et al., 2019; Caylor et al., 2017; Cheong et al., 2020; Craja et al., 2020; Hájek & Henriques, 2017; Zhaokai & Moffitt, 2019). Topic modeling is widely considered a powerful technique in text mining for various purposes, such as mining textual data, defining relationships among data and text documents, also discovering latent data. Focusing on one of the techniques in text mining, the LDA is generally used in topic modeling. Jelodar et al. (2019) systematically reviewed scholarly articles with LDA from 2003 to 2016 and mapped the use of LDA in various subjects (social network, software engineering, crime science, political science, medical/biomedical, and linguistic science). In the accounting domain, the use of LDA is very limited and still in the embryonic phase (Gandía & Huguet, 2021). However, Dyer et al. (2017) proved that LDA was useful in marking trends in 10-K disclosures.

The development of news articles marks a potential outlet to identify growing issues in society. In addition, the unstructured format of news articles makes them suitable for text-mining applications. This has been proven by the existing literature, which provides evidence on the use of news articles in text mining for multiple data analytics objectives, from descriptive to predictive analysis. For example, Re Lee et al. (2022) comprehensively described the extent of start-up innovations in East Asia and North America; Chuluunsaikhan et al. (2020) forecasted a commodity price using deep learning and news topic modeling; also Ertek and Kailas (2021) explored news articles with topic modeling to predict workplace safety risks in the turbine industry.

The abovementioned progress in the literature provides proof of the advantages provided by text mining, specifically topic modeling with online news and using LDA in various data analytics studies. However, the application of LDA to online news in RBA is underexplored. Therefore, the main research question in this study is how topic modeling with the LDA algorithm can be incorporated into the audit risk-assessment process to identify potential areas.

The identification of potential areas is imperative for an efficient and effective audit process. By identifying potential areas, an auditor can understand audit objects and associated problems (BPK, 2020). In Indonesia, the Audit Board of the Republic of Indonesia (BPK) conducts audits for the government's public programs. The management of the School Operational Assistance Fund (BOS Fund) became the focus of the BPK annual audit as it is related to the achievement of Sustainable Development Goal 4 at the national level. The BOS Fund is an educational intervention program by the government of Indonesia in the form of a special allocation fund to support non-personnel operational costs in educational units or schools providing primary and secondary education. Although the program has been running for a long time, there were many problems with its management. This urged BPK to carry out a special audit on the management of the BOS Fund in the second semesters of 2008 and 2018. The audit results revealed many problems related to the mismanagement of the BOS Fund in every stage (i.e., planning and budgeting, implementation and administration, reporting and accountability, and guidance and supervision). These problems existed because BPK was unable to appropriately select audit areas with a high probability of mismanagement during initial audits. This has prompted the need for alternative ways to improve audit quality, one of which is better identifying potential areas in BOS Fund management. Although text analytics holds significant potential as a robust method in the audit process, it has not been integrated into any phase of the BOS Fund audit program.

This study demonstrates the effectiveness of employing topic modeling using the LDA algorithm to identify issues related to an audit object. Our model successfully identifies two potential areas in the BOS Fund management—implementation and administration and guidance and supervision. This study makes two important contributions to the information systems and auditing literature. First, the study provides evidence of the additional benefits of employing topic modeling with the LDA algorithm in areas that have not been widely explored. Second, this study offers a framework that auditors can use for risk assessments using nonfinancial sources issued by independent parties. Practically, this study serves as a guide for model development by auditors in both the public and private sectors in determining potential areas and the scope of audits.

The remainder of the paper is organized as follows. The next section outlines the literature discussing risk-based audits, text mining, and topic modeling using LDA. The third section highlights the institutional background of this study, covering the BOS Fund as an education intervention program by the government of Indonesia and the program's audit. The fourth section describes the method employed. The fifth section presents the data and application of the LDA model in extracting textual information useful for audit, the empirical results, and the discussion. The last section presents the conclusions and suggestions for future research.

Risk Based Audit (RBA)

RBA was defined by Johnstone-Zehms et al. (2015) as a method of conducting audits that concentrates audit resources on areas with the highest risk. It entails the identification, evaluation, and response to potential risks of significant inaccuracies in financial statements by utilizing appropriate audit procedures according to specific circumstances. RBA aims to achieve an effective and efficient audit process (Nugraheni & Pamungkas, 2021). The application of the RBA approach has been emphasized in public sector audits because they are faced with relatively shorter audit times compared with those in the private sector. Completion of audits in the public sector is regulated by law, making time constraints more apparent. In Indonesia, the RBA approach allows BPK to understand current risks and assess the effectiveness of existing controls. In financial auditing, the basic premise of implementing RBA is that the auditor needs to allocate more resources to testing accounts with a greater risk of misstatement and vice versa (Bell et al., 2005; Knechel, 2007; Rittenberg & Schwieger, 1994).

Van Buuren et al. (2014) reveal that the use of RBA is based on considering the trade-off between audit effectiveness and efficiency. Auditors with larger and more complex clients apply RBA more broadly. A number of studies have emphasized the benefits of incorporating RBA in audit procedures. For example, RBA was found to correlate with higher audit quality (Nazmi et al., 2017) and was effective in detecting high-risk areas that could affect the fairness of local government financial reports (Sastra et al., 2018).

In the process, RBA usually comprises several critical stages as follows:

1. Scope Determination and Risk Identification

   At this stage, the auditor determines the scope of the audit and identifies potential risks an entity faces based on a comprehensive range of external and internal information, including financial statements, organizational policies, and current market conditions (Arens, 2012). This risk identification is crucial to determining areas or aspects that require more attention. Information from the mass media or external data plays a significant role in identifying potential risks that are not visible from internal data alone, thereby ensuring a more comprehensive risk assessment (McComb & Shaw, 1972).

2. Risk Assessment

   After identifying risks, an auditor conducts an assessment to measure the extent to which the risk can affect the objectivity and results of the audit. This assessment includes evaluating the impact and likelihood of the risk occurring (Bell et al., 2005). At this stage, the use of text mining on external data, such as news, can provide additional context regarding external issues or trends that may affect risks faced by the entity. For example, news about regulatory changes or economic conditions can indicate specific risks impacting an entity's activities (Feldman & Sanger, 2007).

3. Risk Response

   Based on the risk assessment, an auditor is responsible for determining the audit strategy or procedures to address the risks. The auditor must plan an effective and efficient approach to provide optimal audit results. These audit procedures can be substantive tests or tests of controls, depending on the type and level of risks identified (Knechel, 2007; Rittenberg & Schwieger, 1994).

4. Monitoring and Evaluation of Audit Results

   The final stage in the RBA is monitoring and evaluating audit results and responses to the risks identified. The auditor plays a crucial role in assessing the effectiveness of the audit procedures performed and determining whether further actions need to be taken. This stage includes preparing an audit report that provides recommendations for management to manage risks more effectively (Van Buuren et al., 2014).

Role of Media Information and Text Mining in RBA

Mass media, including traditional and digital platforms, play an important role as a source of information in risk assessment. It helps disseminate risk-related information significantly influencing individuals' understanding of external risks. Different sources of media influence exist (Wendling et al., 2013). Thus, newspapers possess the power to shape or establish individuals' perspectives about the world, their understanding of what is considered normal or appropriate, and their perception of significant public issues. By focusing on media studies in the digital era, Vogler and Eisenegger (2021) find that the prominence and sentiment of news media coverage of corporate social responsibilities have a positive correlation with corporate reputation.

RBA prioritizes the allocation of audit resources to areas that pose the most significant risk to financial statements and the achievement of organizational objectives. This proactive approach allows auditors to identify and assess risks that can significantly impact business continuity and the reliability of financial reporting. In the RBA process, the risk identification and evaluation steps are crucial because they determine audit procedures that must be applied to each risk area. To make audits more effective and efficient, the implementation of RBA in the public sector is increasing, especially in overcoming relatively short time constraints, compared with the private sector (Nugraheni & Pamungkas, 2021). In this context, external information from the mass media, including digital and print platforms, is valuable because it can provide additional insights for auditors regarding external risks that impact operational activities, including economic changes, regulations, and global events (BPK, 2020; Arens, 2012).

At the risk assessment stage in RBA, mass media are an important external source of information for identifying and understanding risks that may not be visible in an entity's internal data. For example, news about changes in economic policy, the introduction of new regulations, or global events such as pandemics or international conflicts can be early signals of risks that will significantly affect an entity. With its real-time information, the mass media keeps auditors informed about public perception and industry trends, which can strengthen auditors' risk assessment. Using information from these media helps auditors conduct a more comprehensive analysis and improve the accuracy of risk assessments so that overall audit quality can be improved. Thus, integrating mass media information in the risk-based audit process enriches an auditor's understanding of external risks and strengthens the audit's capabilities in detecting and anticipating risks more proactively (Arens, 2012).

Text mining, a method that enables the automatic analysis of textual data to uncover patterns and relationships in large volumes of unstructured text (Feldman & Sanger, 2007), has significant potential in auditing. It can be a powerful tool for detecting suspicious financial transactions (Hájek & Henriques, 2017) and supporting risk assessment and substantive testing (Zhaokai & Moffitt, 2019). Gandía and Huguet (2021) suggested that text mining can revolutionize accounting research by analyzing text data from diverse sources such as the Internet, social media, and corporate reports, thereby addressing public and private issues.

In RBA, text mining is a potent tool for analyzing news articles as an external data source to identify audit risks. The LDA algorithm, a key method in topic modeling, is particularly effective in uncovering hidden themes in a collection of documents. Its use is instrumental in helping auditors identify areas of risk that may not be apparent from internal data alone (Blei, 2012; Jelodar et al., 2019). With LDA, auditors can categorize texts based on content similarities and associate them with the management stages of a particular audit object (Dyer et al., 2017). This process aids in prioritizing areas with the highest potential risks, enhancing the quality of a risk-based audit.

In the context of risk assessment in the audit of the BOS Fund in Indonesia, the use of LDA for text mining is a game-changer. It allows the classification of news articles based on keywords related to BOS Fund management, such as planning, implementation, and supervision. This approach enables the identification of topics that can be sources of risk, thereby helping auditors determine areas to focus on during an audit (BPK, 2020). This study demonstrates that text mining and topic modeling with LDA can provide valuable insights for auditors, empowering them to better understand audit objects and identify external risk areas.

Overall, text mining, especially with the LDA algorithm, plays an important role in the RBA risk assessment stage through media information analysis. Utilizing media as a source of external information allows auditors to understand risks from a broader perspective, not limiting them to the internal data of an audited entity. Thus, text mining is a potential tool for supporting a more efficient audit process and improving audit quality through the early identification of potential risks that are not easily identified from internal data alone.

Based on the findings demonstrating the power of text mining and topic modeling using LDA, the authors argue that online news articles can provide valuable information for auditors in identifying risky areas around audit objects. This leads to two main propositions as follows:

Proposition 1: Online news contains information that auditors can use to assess risks associated with audit objects.

Proposition 2: LDA in topic modeling is useful for identifying areas of higher risk in risk-based audits.

These two propositions confirm that LDA plays a crucial role in strengthening risk analysis and improving the quality of risk-based audits by enabling auditors to use online news effectively.

Institutional Setting

BOS fund

The BOS Fund is a program by the government of Indonesia that focuses on education and aims to cover nonpersonnel operating costs of primary and secondary schools in Indonesia. Since its launch, the Indonesian government has been providing the BOS Fund to gradually reduce the annual average education unit cost per student required to support the teaching and learning process in accordance with established service standards (BPK, 2009). The budget for the BOS Fund is set by the central government and is transferred to the regional governments using the regional transfer expenditures account. Through the Ministry of Education and Culture, the government provides technical guidelines for managing the fund. The BOS Fund allocated for each school comprises the Regular BOS Fund and Performance BOS Fund. The Regular BOS Fund is allocated based on the unit cost of the Regular BOS Fund in each region multiplied by the number of students, while the BOS Performance Fund allocation is determined by a Ministerial Decree.

The BOS Fund is managed by the receiving school and the regional government. The management of the funds by a school comprises three main activities—a) planning and budgeting, that is, preparing the school work plan and budget; b) implementation and administration, that is, using funds and inputting the use of funds into the application system for activity plans and the school's budget provided by the Ministry; and c) reporting and accountability, that is, preparing reports on the results of the administration of the BOS Fund and examination and verification of the completion of the procurement of goods and services for the year in question. Further, the management cycle at the regional level includes four activities—a) planning and budgeting, b) implementation and administration, c) reporting and accountability, and d) guidance and supervision. This study discusses the regional government's BOS Fund management cycle.

BOS fund audit

The BOS Fund is a government educational intervention program that has been running for 18 years in Indonesia. According to the latest data from the Ministry of Education and Culture (2023), the budget for the BOS fund in 2022 was IDR 54.1 trillion or USD 3.6 million and allocated to 216,483 schools in Indonesia. The large amount of the fund makes accountability strongly demanded so that the BOS Fund can bring optimum benefit to society (Rahayu et al., 2015). This demand is further strengthened by the occurrence and spread of irregularities and corruption cases related to the BOS Fund (Herrudi et al., 2018). This finding is supported by the fact that the BOS Fund is one of the 17 objects in Indonesia prone to corruption (Ramadhani et al., 2022). One of the reasons might be the managerial challenges related to the institutional and technical aspects of its operations (Haniatun et al., 2022).

Based on the above-mentioned issues, BPK conducts audits of the BOS Fund through financial audits as part of audits of the financial reports of the central and regional governments, as well as thematic audits through special purpose and performance audits. BPK has been constantly conducting audits for the BOS Fund since the first launch of the program. The first audit was in the second semester of 2008, which was conducted together with the other education programs funded by the state and regional budget for the 2007 fiscal year and the first semester of the 2008 fiscal year. The audit was carried out for all 33 provincial governments and 62 municipalities, with a total of 4,127 schools being sampled. The results of the audit revealed that several improvements in the design of internal controls and operational matters of the BOS Fund need to be made by the fund managers (BPK, 2009).

In the second semester of 2018, BPK conducted an audit on the management of the BOS Fund through a thematic performance audit with 54 objects at various levels of regional governments in Indonesia, including 13 objects at the provincial level, 30 at the district level, and 11 at the municipality level. The audit revealed that regarding the management of education funding for students through the BOS Fund, among the sample, 72.22% (39) were not effective and 27.78% (15) were not fully effective (BPK, 2018). The audit further found that significant problems in local governments that can hinder the effectiveness of managing education funds for students include problems with the accuracy of the amount, timeliness, and accuracy of targets, as well as monitoring and evaluation problems. Regarding these problems, the BPK provided recommendations for improvement.

RESEARCH METHOD

Text Mining

Text mining is a versatile approach to analyzing textual data for research in various domains, whether natural science (e.g., biomedics, computation sciences, and neurosciences) or social science (e.g., psychology, linguistics, marketing, finance, accounting, and auditing). One of the reasons is its applicability to a wide range of text documents, reaching a higher level of development. In finance and accounting disciplines, text mining analysis has great potential for research, following an ever-growing number of disclosures and other relevant documents, news, and online media conversations. Gandía and Huguet (2021) noted that corporate disclosures were the most used source for text analysis, followed by Internet and social media, press news, also analysts' reports.

Text mining is the process of searching for information in large amounts of text by automatically identifying patterns and relationships among data. This concept grew along with the increasing number of documents available from both online and offline sources (Feldman & Sanger, 2007). Tan (1999) defined text mining as the process of extracting patterns of importance or knowledge from unstructured textual data. This technique is commonly used to deal with the problems of feature extraction, classification, clustering, and information retrieval of text data. Because text mining stems from data mining, the logic behind it is very similar to that of data mining, except that data mining patterns are taken from structured data (Miner, 2012). The basic text mining steps, widely used in various applications, are text preprocessing and feature selection. Text preprocessing is the first step of text mining and an important process because it produces a standardized and consistent form of data that affects all experiments (Vijayarani et al., 2015), while feature selection aims to select a small subset of relevant features from all original features by eliminating irrelevant features, redundancy, and noise.

In accounting research, text mining offers opportunities to detect fraudulent financial transactions (Craja et al., 2020; Hájek & Henriques, 2017), assess the quality of internal audits (Boskou et al. (2019); evaluate recommendations given by stock analysts (Caylor et al., 2017); assess the informativeness of disclosures (Cheong et al., 2020); and assist auditors in risk assessment, substantive testing, and audit review (Zhaokai & Moffitt, 2019). Gandía and Huguet (2021) systematically reviewed accounting articles in top journals in this field from 2008 to 2019 and recommended extending accounting research by exploiting textual analysis on the Internet and social media sources addressing not only corporate accounting topics but also emerging problems in the public sector domain.

Topic Modeling

Topic modeling is one of the research methods that utilizes textual analysis (Gandía & Huguet, 2021). It focuses on finding the underlying topic or hidden cues of a set of text documents based on the correlations between the words in the document (Blei et al., 2003; Huang et al., 2022). Finding a pattern in a document can reflect the topic underlying the formation of the text document, namely, the probability distribution of the words in the document (Blei & Lafferty, 2009). The basic idea of topic modeling is that a topic comprises certain words that make up the topic. Therefore, it can be used to find topics based on words contained in the text documents (Blei, 2012). Topic modeling is one of the prominent techniques for text mining in the data mining domain, data discovery, exploration of the relationship between data and text documents (Jelodar et al., 2019), and analyzing a set of text documents and clustering them into several representative topics. Therefore, this method is classified as a clustering approach in the machine-learning literature (Nurlayli & Nasichuddin, 2019). The underlying concept of topic modeling is presented in Figure 1. Figure 1 depicts several topics that constitute the distribution of words (left side). Each document is assumed to contain a distribution of topics (histogram on the right side), and each word is then assigned to a topic. Topic selection is based on the choice of words contained in each topic.

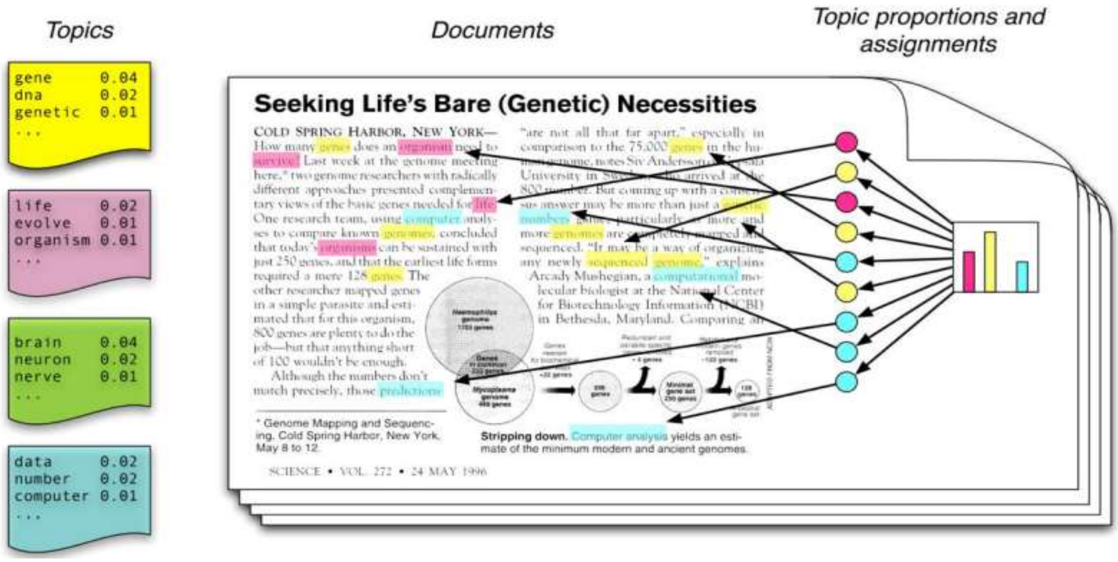

Figure 1. Topic Modeling Concept (Blei, 2012)

Latent Dirichlet Allocation (LDA)

Topic modeling can be carried out in research using the LDA algorithm, which has been documented as the most widely used and empirically tested algorithm in research (Sun & Yin, 2017). The LDA is a generative probabilistic model of a collection of documents called a corpus. The basic premise of the LDA method is that each text document is represented as a random mix of hidden topics, where each topic has a character determined by the distribution of the words contained in it (Jelodar et al., 2019). Blei (2012) visually represented the LDA method as a probabilistic model, as depicted in Figure 2.

As depicted in Figure 2, there are three levels in the LDA modeling. The parameters α and β are topic distribution parameters at the corpus level—a collection of M documents. A "corpus" refers to a large and structured set of texts, while a "document" is a single piece of text in the corpus. The parameter α determines the topic distribution in a document, where the higher the alpha value in a document, the more diverse the mix of topics discussed in the document. The parameter β is used to determine the word distribution in a topic, where the higher the β value, the more words are included in the topic, making the topic contain more specific words. A "topic" is a collection of words that represent a particular theme or subject, and a "word" is a unit of language that carries meaning. The variable θ is at the document level (M). The variable θ represents the topic distribution for a specific document. The higher the value of θ, the more topics are in the document. The lower the value of θ, the more specific the document is to a particular topic. The variables Zn and Wn are at the word level (N). The variable Z represents the topic of a specific word in a document, while the variable W represents the word associated with a particular topic in a document.

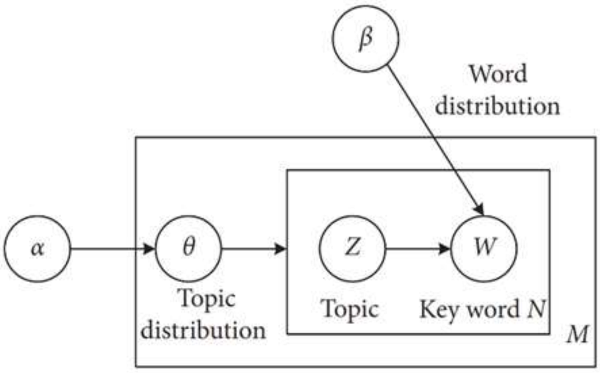

Figure 2. Topic Modeling Visualization with LDA (Blei, 2012)

In general, LDA works by inputting individual documents and several parameters. The resulting output is a model that carries weights and can be normalized based on probability. This probability refers to two types, namely, (a) the probability of a certain specific document producing a specific topic and (b) the probability of a specific topic producing specific words from a set of vocabulary. Probability type (a) documents, that is, those that have been labeled with a list of topics, often continue to produce probability type (b), resulting in specific words (Bird et al., 2015).

Our analysis is based on case studies of problems related to the management of the BOS Fund. We obtained the data from *detik.com*—a leading Indonesian online news portal that extensively covers issues related to government programs across regions in Indonesia. Although we did not impose a time restriction, our data revealed that the earliest news about BOS Fund management was reported in January 2008 and the latest was in May 2023, retrieving 1,460 news articles. We used a keyword approach to generate news articles that discuss the BOS Fund. The keyword approach was adopted because it focused on a small range of words or phrases instead of a wide list of expressions, which can cause ambiguity (Loughran et al., 2009; Loughran & Mcdonald, 2011). As the BOS Fund is an educational intervention specific to Indonesia, our keyword used on the online news portal was in Indonesian, and it was “Bantuan Operasional Sekolah.” This keyword was chosen because it is the common term that describes the BOS Fund. A search using “BOS” yielded news and “conversations about superlatives” at workplaces.

This study proposes the use of text mining with the LDA algorithm to identify problems surrounding an audit object so that it can help auditors determine potential areas for a risk-based audit. Text mining is one of the subdomains in data mining that analyzes natural language text (Daróczi, 2015) by collecting, cleaning, processing, analyzing, and gaining knowledge or useful insights from data (Aggarwal, 2015). Text mining requires extracting meaningful numeric keys from unstructured textual information so that the information can be applied to various data mining algorithms. Text mining is an interdisciplinary field in the sense that it is developed with contributions from various fields, such as statistics, machine learning, pattern recognition, information retrieval, and bioinformatics.

Text mining was applied following data mining approaches, such as document clustering and classification (Zorio-Grima & Carmona, 2019). The notion is to alter unstructured text into structured data based on term frequencies and apply protocol data mining techniques accordingly. This study focused on understanding issues related to the BOS Fund by identifying topics that emerged from the data. First, it employed the bag-of-words reference used in the model development. A bag-of-words model provides a sufficient understanding of the textual data's content, and cluster analysis allows text to be classified according to similarities in the content and subsequently given the appropriate label. The labels reflected topics discovered in the text and were developed by a subject matter expert. Using Jupyter Notebook (Anaconda 3) to run Python version 3.9.7 software enabled us to develop diagrams based on clusters of topics identified and make a good analysis.

We followed text mining procedures proposed by Vijayarani et al. (2015), which involve data collection, text preprocessing, tokening, and feature extraction, before conducting the analysis using the LDA algorithm. The research procedure in this study also involved mapping the topics by an expert auditor with experience in auditing the BOS Fund. In brief, we developed and used the framework presented in Figure 3 to implement topic modeling with LDA to identify potential risk areas in BOS Fund management.

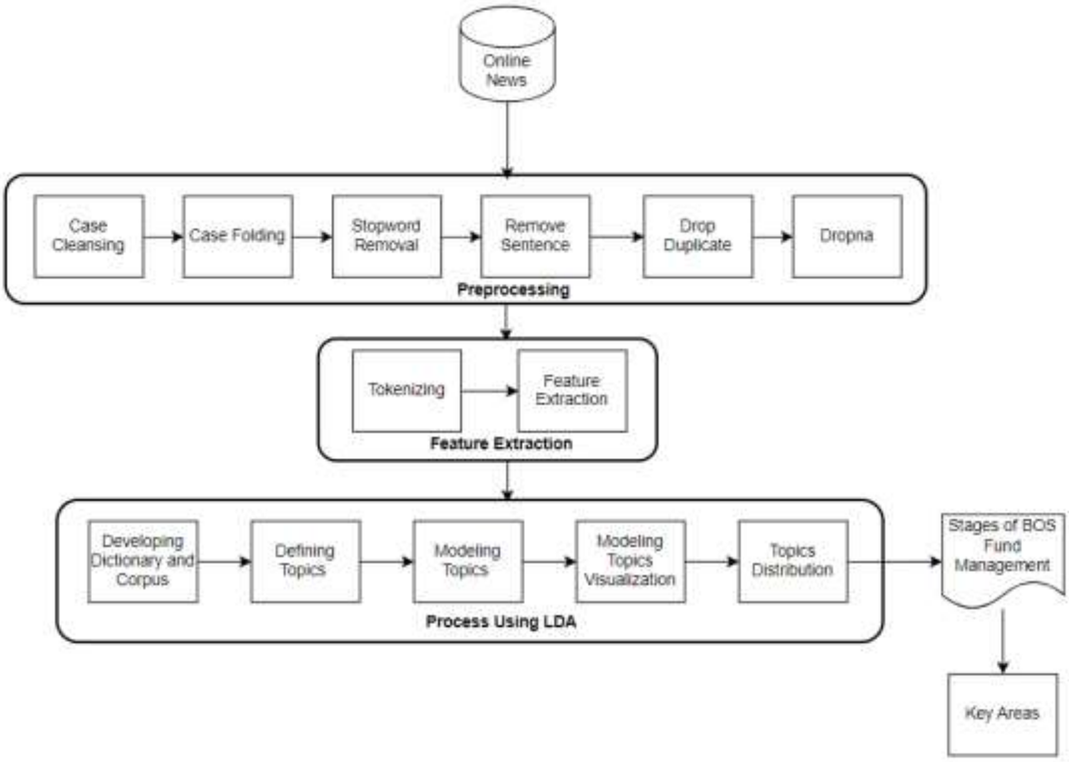

Figure 3. Framework for Topic Modeling Using LDA for Potential Risk Areas Identification

Based on Figure 3, the stages of topic modeling with LDA are as follows:

1. Preprocessing
   This process aims to clean and prepare the text data for further analysis. The steps are as follows:
   a. Case Cleansing: Cleansing the data from unwanted characters or symbols, such as punctuation marks or special symbols
   b. Case Folding: Converting all text to lowercase to ensure consistency
   c. Stopword Removal: Removing common and insignificant words for analysis, such as "dan" and "oleh"

   d. Sentence Removal: Deleting sentences or parts of the text that are irrelevant to the analysis purpose
   e. Dropping Duplicates: Removing duplicated data, ensuring no repeated articles or texts
   f. Performing Dropna: Deleting empty or null data to avoid disruption in the analysis
2. Feature Extraction
   This stage aims to transform the text into a format that can be analyzed. The steps are as follows:
   a. Tokenizing: Splitting the text into smaller units such as words or phrases (tokens)
   b. Feature Extraction: Extracting features from the text to be used as input in the topic modeling process
3. Process with LDA
   The core process in this text analysis is using the LDA algorithm to find hidden topics in the text dataset. The steps include the following:
   a. Developing Dictionary and Corpus: Create a dictionary and corpus that represent all words and documents in a format that the algorithm can process
   b. Defining Topics: Determine the number and type of topics to be extracted from the text
   c. Modeling Topics: Use the LDA algorithm to divide the documents into topics based on word distribution in the text
   d. Modeling Topics Visualization: Visualize the discovered topics to make them easier to understand
   e. Topics Distribution: Determine the distribution or proportion of topics in each document
4. Stages of BOS Fund Management
   After the LDA process, we used expert auditors to map each term generated by LDA to the BOS Fund management cycle (i.e., planning and budgeting, implementation and administration, reporting and accountability, and guidance and supervision). The results provide insights into the stages of BOS Fund management discussed in online news articles. This helps identify the key areas of BOS Fund management.

RESULT AND DISCUSSION

The authors divide our results presentation and discussion into two main subsections: the LDA topic modeling process and the identification of potential areas based on the LDA topic modeling results.

LDA Topic Modeling Process

We performed topic modeling by using the LDA algorithm to form a topic model. The three main steps used to obtain the topic model are developing a dictionary and corpus, determining the number of topics analyzed, and visualizing the LDA results.

1. Formation of the dictionary and corpus

   The formation of a dictionary involves the identification of a data format containing a unique set of words with their respective index numbers. This process utilized the *corpora.dictionary* function from the *genesis* library. The results of the formation of the dictionary are as follows:

```
Dictionary (3483 unique tokens: [‘akhir’, ‘antiseptic’, ‘aplikasi’, ‘aset’, ‘aset_bpkad’]...)
```

The corpus is a data format in the form of a bag-of-word reference that is used for model building. The formation of this corpus produced a document in the form of a list comprising *token_id* and *token_count* for each word as follows:

```
[(0, 1), (1, 1), (2, 3), (3, 2), (4, 1), (5, 1), (6, 3), (7, 3), (8, 8), (9, 1), (10, 2), (11, 1), (12, 1), (13, 1), (14, 1), (15, 1), (16, 1), (17, 1), (18, 1), (19, 3), (20, 1), (21, 3), (22, 1), (23, 1), (24, 1), (25, 1), (26, 1), (27, 1), (28, 2), (29, 1), (30, 1), (31, 1), (32, 4), (33, 1), (34, 1), (35, 1), (36, 1), (37, 1), (38, 1), (39, 2), (40, 5), (41, 1), (42, 6), (43, 7), (44, 3), (45, 1), (46, 2), (47, 1), (48, 2), (49, 1), (50, 1), (51, 1), (52, 16), (53, 9), (54, 1), (55, 1), (56, 2), (57, 1), (58, 2), (59, 6), (60, 5), (61, 1), (62, 1), (63, 1), (64, 1), (65, 3), (66, 2), (67, 1), (68, 3), (69, 1), (70, 1), (71, 1), (72, 1), (73, 1), (74, 3)]
```

2. Determination of the number of topics presented

   We determined the number of topics as the primary experiment to define the best topic model. This step is critical to produce a model with a high coherence score that translates to a quality result of the topic modeling interpretation. Determining the number of topics began with an assumption of an initial value of 10 topics, which were then tested with 10 experiments. This experiment resulted in coherence scores, which were then visually analyzed for the trend of values.

3. Visualization of the LDA topic modeling results

   The number of topics selected was based on those with the highest average value according to the visualization. Based on the experimental results, we used a total of 10 topics, having the highest average topic coherence score of 0.4360. An example of the visualization results of the limit of the number of topics (t = 10) is presented in Figure 4.

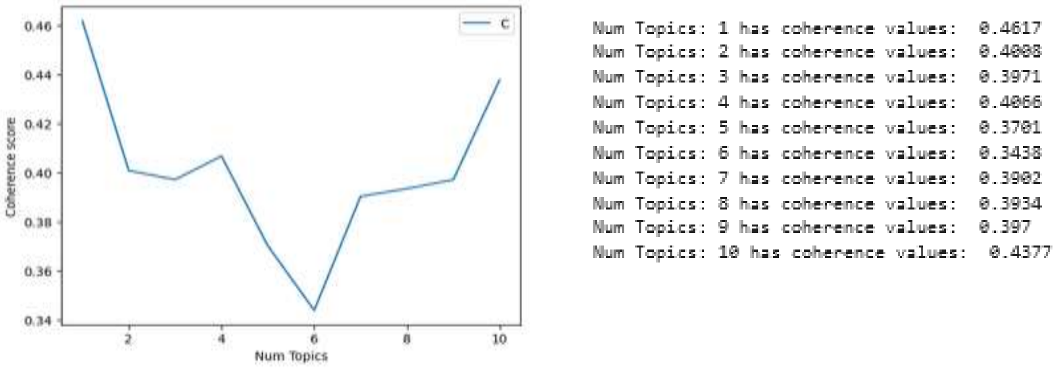

Figure 4. LDA Topic Modeling Results Visualization

Modeling is done with the *LdaModel* function from the *genesis* library. This function generates the order of words in a topic according to the highest probability value in each topic. The output of this function is visualized with the *PyLDAvis* function in the *genesis* library, resulting in a distance map and a workload for each topic. Figures 5 and 6 depict an example.

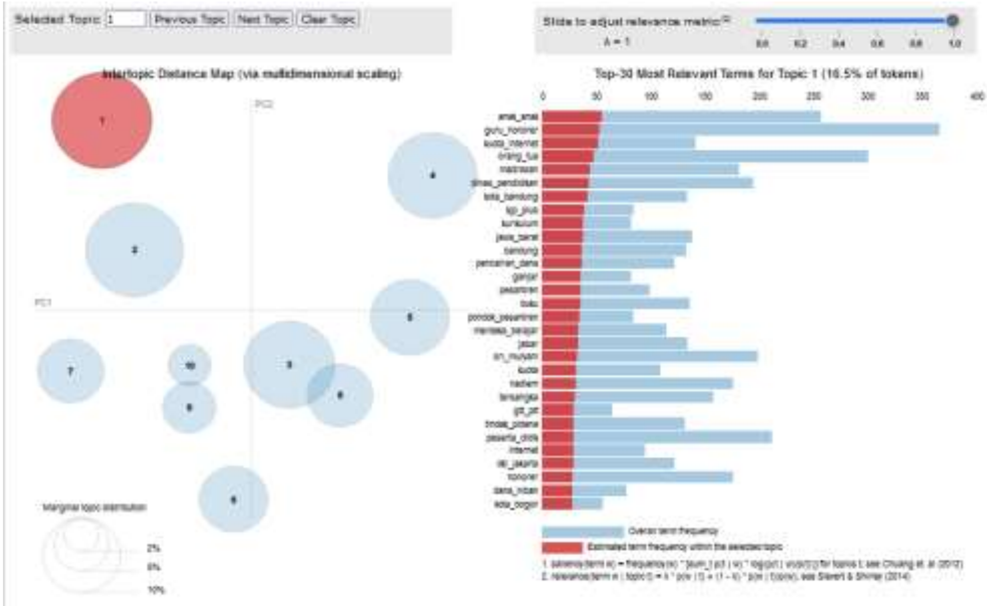

Figure 5. PyLDAvis for Topic 1

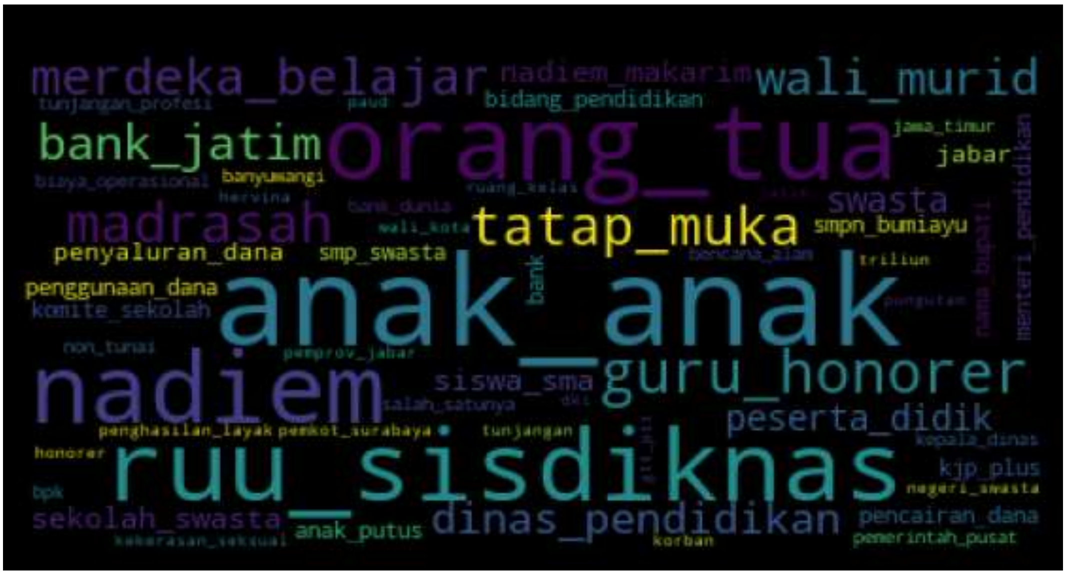

Figure 6. Wordcloud for Topic 1

Identification of potential areas based on the LDA topic modeling results

The previous step provided a list of 10 topics arranged in order based on the probability of occurrence in the news articles. We collected ten terms of the ten topics and interpreted them to identify potential areas. This process involved judgment from an expert to adequately map each word (term) into the BOS Fund management cycle (i.e., planning and budgeting, implementation and administration, reporting and accountability, and guidance and supervision). Table 1 in the Appendix presents all terms on the generated topics mapped to the BOS Fund management cycle.

The resulting map above indicates that only three of the four steps in the BOS Fund management cycle were reported in the news articles. Therefore, we excluded one topic that was not covered, reporting and accountability, from our further analysis. Based on these three steps, Figure

7 defines the number of occurrences of each stage in all topics.

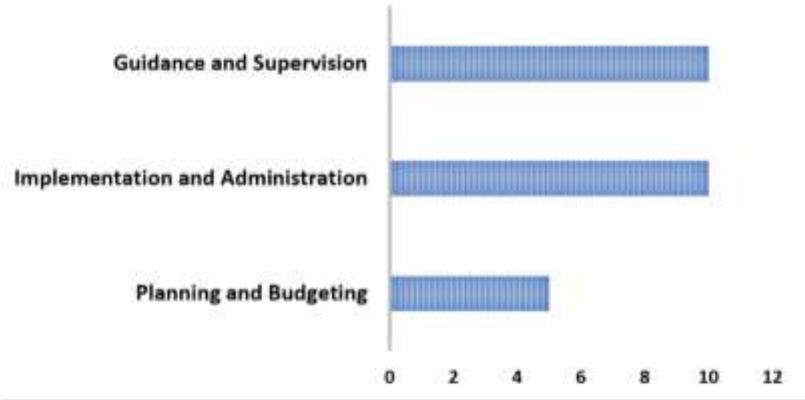

Figure 7. Number of Occurrences of BOS Fund Management Cycle

This study adopted the criteria developed by Ibtida and Pamungkas (2017) in deciding whether a stage in the BOS Fund management cycle is a potential area in a risk-based audit. To be judged as a potential area, a stage must occur in more than 50% of the total topics (five occurrences). Based on the visual calculation presented in Figure 7, we found two potential areas, namely, implementation and administration, as well as guidance and supervision. The implementation and administration stage is the core of the BOS Fund management cycle. Activities in this stage include the distribution of funds, the use of funds, and the administration of the use of funds. Potential deviations that can occur in this stage include delays in the disbursement of funds, use of funds that are not in accordance with the provisions, misuse of funds for personal gain, and failure to record the use of funds. Guidance and supervision are equally important stages in managing a BOS Fund. Activities in this stage include data verification and validation, provision of training and technical guidance, as well as coordination and consolidation. Potential deviations in this stage include the absence of verification and validation of school data, inappropriate training and technical assistance, and lack of coordination between regional and central governments.

Our results are validated by the summary of the BPK Audit Report for the first semesters of 2021 and 2022. According to the reports, the findings in the management of the BOS Fund included the following: expenditures were overstated; the process of validating the BOS Fund receipt and disbursements was not carried out periodically; goods and services expenditure did not comply with provisions; and schools did not receive their entitled BOS Fund amounts. These findings are related to the implementation and administration stage. Other important findings are related to the guidance and supervision stage, including poor cash management at the BOS treasury, errors in the distribution of the BOS Fund, accountability for the BOS Fund not matching the actual conditions, and the lack of reports validations to ensure accountability for the BOS Fund. These research findings imply that news topic modeling is effective in helping auditors identify potential areas in a risk-based audit. Therefore, this study successfully provided alternative efforts for auditors to efficiently and effectively assess the risk process.

LDA in topic modeling effectively identifies areas with higher risks in RBA. The LDA algorithm allows auditors to uncover latent data in text documents, which helps them understand the audit object and related issues (Blei, 2012). Thus, LDA plays a crucial role in improving audit quality by helping auditors focus on areas with a high probability of deviation, thereby reassuring the audience about the method's effectiveness.

This study provides strong evidence that using text mining with the LDA algorithm in topic modeling of online news effectively identifies risk areas in the risk-based audit process. The LDA method allows auditors to analyze large volumes of text data (e.g., news articles) and group hidden topics relevant to certain risks. In this study, an LDA analysis of online news related to the BOS Fund successfully identified potential areas for risk audits, such as the implementation and administration stages as well as guidance and supervision stages, which are often vulnerable aspects of the BOS Fund management. This implies that online news contains information on which auditors can focus when they conduct an audit risk assessment.

The effectiveness of LDA in RBA makes it a comprehensive approach that can significantly improve the accuracy of risk assessments by auditors. This approach allows for the collection of information from external data to identify trends or issues that are not always visible from internal data. Dyer et al. (2017) demonstrated that LDA effectively tracks trends in annual financial statements (10-K), supporting the use of LDA to identify and prioritize risk areas in audits. In addition, the use of LDA in this study provides practical guidance for auditors to develop a more comprehensive risk assessment model by utilizing independent nonfinancial information sources, such as online media.

CONCLUSION

Authors explored the possibility of implementing text mining, especially topic modeling using LDA algorithms on online news, to facilitate public sector auditors in the risk assessment process. Using 1460 online news data from more than 14 years of publication in a leading online news portal in Indonesia, we textually analyze each news covering concerns over an Indonesian massive educational program. The result of this study suggested that online news can be textually analyzed using LDA algorithms to generate meaningful topics. Gaining insights from an expert to map the topics into an object of interest is critical in identifying potential areas in a risk-based audit program.

The LDA algorithm can be effectively used by following two sequential stages, namely the LDA topic modeling process and the identification of potential areas according to the results of the first stage. This suggests that online news contains information auditors can use to assess the risks associated with audit objects. The initial stage of LDA is designed to ascertain the news topics frequently emerging on the internet and categorize them according to their recurrence. Afterward, all topics are classified according to the policy's management cycle. Subsequently, according to the LDA topic modeling process, we can determine the potential area by using judgment with the provision that such area must occur in more than 50% of the total topics. In the end, we can identify potential areas of particular issues. This method provided different risk assessment efforts as an alternative to non-financial sources. In addition, the results of the LDA algorithm method in this study align with those obtained by not using this method. In conclusion, our result suggests that the LDA topic modeling can be used on online news to uncover risks associated with audit object, thus benefiting auditors to enhance the audit quality through a better risk assessment.

The main contribution of this study lies in the idea of using topic modeling in the risk assessment process of risk-based audits. This study provides a framework that can be used by auditors to better identify areas surrounding their auditees exposed to relatively higher risks. This careful identification is imperative to conduct a risk-based audit to achieve the goal of efficient allocation of audit resources and effective overall audit programs. A practical implication that can be drawn from this study is that it provides auditors in both the public and private sectors with guidance for developing models to identify potential audit areas.

While the scope of our study is limited to the textual analysis of online news, further research could explore social media to identify potential areas, such as X, a.k.a Twitter, with a careful procedure as X is exposed to noise. We also limit our exploration only to the LDA algorithm; the presence of other algorithms, such as Latent Semantic Analysis, Non-Negative Matrix Factorization, Probabilistic Latent Semantic Analysis, and Hierarchical Dirichlet Process, calls for evaluations by further research. Discovering what algorithm gives the best results for a range of audit purposes would be beneficial to enriching the existing accounting information system literature.

REFERENCES

ACCA. (2008). *A risk-based approach to auditing financial statements*. https://www.accaglobal.com/content/dam/acca/global/PDF-students/2012s/sa_feb08_pine.pdf

Aggarwal, C. C. (2015). Data mining: The textbook. *Springer International Publishing*, pp. 285–344. https://doi.org/10.1007/978-3-319-14142-8_10

Arens, A. A. (2012). *Auditing and assurance services: An integrated approach*. Pearson Education.

Arens, A. A., Elder, R. J., Mark S, B., Hogan, C. E., & Jones, J. C. (2021). *Auditing: The Art and Science of Assurance Engagements*. Pearson Education Canada.

Bell, T. B., Peecher, M. E., & Solomon, I. (2005). *The 21st-century public company audit: Conceptual elements of KPMG's global audit methodology*. KPMG International.

Bird, C., Menzies, T., & Zimmermann, T. (2015). Chapter 1 - Past, present, and future of analyzing software data. In C. Bird, T. Menzies, & T. Zimmermann (Eds.), The art and science of analyzing software data (pp. 1–13). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-411519-4.00001-X

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84. https://doi.org/10.1145/2133806.2133826

Blei, D. M., & Lafferty, J. D. (2009). Topic models. In A. Srivastava, & M. Sahami (Eds.), *Text mining*: *Classification, clustering and applications* (pp. 71-93). Cambridge: Chapman and Hall/CRC.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Boskou, G., Kirkos, E., & Spathis, C. (2019). Classifying internal audit quality using textual analysis: The case of auditor selection. *Managerial Auditing Journal*, *34*(8), 924–950. https://doi.org/10.1108/MAJ-01-2018-1785

BPK. (2009). *Summary of Semester Audit Results II Year 2009 (Ikhtisar Hasil Pemeriksaan Semester II Tahun 2009)*. https://www.bpk.go.id/ihps

BPK. (2018). *Summary of Semester Audit Results II Year 2018 (Ikhtisar Hasil Pemeriksaan Semester II Tahun 2018)*. https://www.bpk.go.id/ihps

BPK. (2020). *Performance Audit Guidelines (Petunjuk Pelaksanaan Pemeriksaan Kinerja).* Supreme Audit Board of The Republic of Indonesia, Jakarta.

Caylor, M., Cecchini, M., & Winchel, J. (2017). Analysts' qualitative statements and the profitability of favorable investment recommendations. *Accounting, Organizations and Society*, *57*, 33–51. https://doi.org/10.1016/j.aos.2017.03.005

Cheong, A., Yoon, K., Cho, S., & No, W. (2020). Classifying the contents of cybersecurity risk disclosure through textual analysis and factor analysis. *Journal of Information Systems*, *35*. https://doi.org/10.2308/ISYS-2020-031

Chuluunsaikhan, T., Ryu, G. A., Yoo, K. H., Rah, H., & Nasridinov, A. (2020). Incorporating deep learning and news topic modeling for forecasting pork prices: The case of South Korea. *Agriculture (Switzerland)*, *10*(11), 1–22. https://doi.org/10.3390/agriculture10110513

Craja, P., Kim, A., & Lessmann, S. (2020). Deep learning for detecting financial statement fraud. *Decision Support Systems*, *139*, 113421. https://doi.org/10.1016/j.dss.2020.113421

Daróczi, G. (2015). *Mastering data analysis with R*. Packt Publishing Ltd.

Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, *64*(2–3), 221–245. https://doi.org/10.1016/j.jacceco.2017.07.002

Ertek, G., & Kailas, L. (2021). Analyzing a decade of wind turbine accident news with topic modeling. *Sustainability (Switzerland)*, *13*(22). https://doi.org/10.3390/su132212757

Feldman, R., & Sanger, J. (2007). *The text mining handbook: Advanced approaches in analyzing unstructured data*. Cambridge University Press.

Gandía, J. L., & Huguet, D. (2021). Textual analysis and sentiment analysis in accounting. *Revista de Contabilidad-Spanish Accounting Review*, *24*(2), 168–183. https://doi.org/10.6018/RCSAR.386541

Hájek, P., & Henriques, R. (2017). Mining corporate annual reports for intelligent detection of financial statement fraud – A comparative study of machine learning methods. *Knowledge-Based Systems*, *128*. https://doi.org/10.1016/j.knosys.2017.05.001

Haniatun, H., Islahuddin, I., & Abdullah, S. (2022). Influence of management competence, utilization of information technology and stakeholder engagement on accountability of management of BOS funds with transparency as a moderating variable (Study on SMAN and SMKN in Aceh Selatan district). *International Journal of Business, Economics, and Social Development*, *3*(3), 110–123.

Herrudi., Handajani, L., & Surasni, N. K. (2018). Fraud financial management of school operational assistance (BOS): Study at state elementary school in East Lombok, Indonesia. *International Journal of Economics, Commerce and Management*, *VI*(7).

Huang, F., Yuan, C., Bi, Y., Lu, J., Lu, L., & Wang, X. (2022). Multi-granular document-level sentiment topic analysis for online reviews. *Applied Intelligence*, 52, 7723–7733. https://doi.org/10.1007/s10489-021-02817-1

Ibtida, R., & Pamungkas, B. (2017, August). *The design of a risk-based performance audit program for court-fee management for the comptroller of supreme court of Indonesia* [Paper presentation]. The 6th International Accounting Conference. https://doi.org/10.2991/iac-17.2018.36

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: Models, applications, a survey. *Multimedia Tools and Applications*, *78*(11), 15169–15211. https://doi.org/10.1007/s11042-018-6894-4

Johnstone-Zehms, K. M., Gramling, A. A., & Rittenberg, L. E. (2015). *Auditing: A risk based-approach to conducting a quality audit* (10th ed.). Cengage Learning.

Knechel, W. R. (2007). The business risk audit: Origins, obstacles and opportunities. *Accounting, Organizations and Society*, *32*(4), 383–408. https://doi.org/10.1016/j.aos.2006.09.005

Loughran, T., & Mcdonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Loughran, T., McDonald, B., & Yun, H. (2009). A wolf in sheep's clothing: The use of ethics-related terms in 10-K reports. *Journal of Business Ethics*, *89*, 39–49.

McComb, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quaterly*, *36*(2), 176–187. http://www.jstor.org/stable/2747787

Miner, G. (2012). *Practical text mining and statistical analysis for non-structured text data applications*. Academic Press.

Ministry of Education and Culture. (2023). *Distribution report of the school operational assistance fund (Laporan penyaluran BOS)*. https://bos.kemdikbud.go.id/rekap/penyaluran

Nazmi, E., Arori, I. M. S., & Ibrahim, M. R. (2017). The factors affect business risk audit and their impact on the external auditing quality in Jordanian commercial banks (Case study). *European Journal of Accounting, Auditing and Finance Research*, *5*(5), 1–17.

Nugraheni, N. M. T., & Pamungkas, B. (2021). Analysis of the RBA implementation and the preparation of an audit program at the Ministry of Villages, Development of Disadvantaged Regions and Transmigration. *Jurnal Tata Kelola Dan Akuntabilitas Keuangan Negara*, *7*(1), 77–93. https://doi.org/10.28986/jtaken.v7i1.489

Nurlayli, A., & Nasichuddin, M. A. (2019). Topik modeling penelitian dosen jptei uny pada google scholar menggunakan latent dirichlet allocation. *Elinvo (Electronics, Informatics, and Vocational Education)*, *4*(2), 154–161.

Perko, T. (2012). The role of mass media and journalism in risk communication. *Journal of Mass Communication and Journalism*, *2*(2). https://doi.org/10.4172/2165-7912.1000e110

Rahayu, S., Ludigdo, U., & Irianto, G. (2015). Budgeting of school operational assistance fund based on the value of gotong royong. *Procedia-Social and Behavioral Sciences*, *211*, 364–369. https://doi.org/10.1016/j.sbspro.2015.11.047

Ramadhani, R. S., Lawita, N. F., & Anriva, D. (2022). The factors affecting fraud prevention in BOS fund management. *Economic Education Analysis Journal*, *11*(2), 144–154. https://doi.org/10.15294/eeaj.v11i2.54135

Re Lee, K., Hyun Kim, J., Jang, J., Yoon, J., Nan, D., Kim, Y., & Kim, B. (2022). News big data analysis of international start-up innovation discourses through topic modelling and network analysis: Comparing East Asia and North America. *Asian Journal of Technology Innovation*, *31*(3), 581–603. https://doi.org/10.1080/19761597.2022.2134154

Rittenberg, L. E., & Schwieger, B. J. (1994). *Auditing: Concepts for a changing environment*. Fort Worth, Tex.: Dryden Press.

Sastra, C. D., Yuhertiana, I., & Budiwitjaksono, G. S. (2018). The use of risk based audit techniques in government entities: Indonesia case. *Account and Financial Management Journal*, *3*(6), 1581–1586. https://doi.org/10.31142/afmj/v3i6.04

Sun, L., & Yin, Y. (2017). Discovering themes and trends in transportation research using topic modeling. *Transportation Research Part C: Emerging Technologies*, *77*, 49–66. https://doi.org/10.1016/j.trc.2017.01.013

Tan, A.-H. (1999). *Text mining: The state of the art and the challenges*. https://www.researchgate.net/publication/2471634_Text_Mining_The_state_of_the_art_and_the_challenges

Van Buuren, J., Koch, C., van Nieuw Amerongen, N., & Wright, A. M. (2014). The use of business risk audit perspectives by non-big 4 audit firms. *AUDITING: A Journal of Practice & Theory*, *33*(3), 105–128. https://doi.org/10.2308/ajpt-50760

Vijayarani, S., Ilamathi, M. J., & Nithya, M. (2015). Preprocessing techniques for text mining - An overview. *International Journal of Computer Science & Communication Networks*, *5*(1), 7–16.

Vogler, D., & Eisenegger, M. (2021). CSR communication, corporate reputation, and the role of the news media as an agenda-setter in the digital age. *Business and Society*, *60*(8), 1957–1986. https://doi.org/10.1177/0007650320928969

Wendling, C., Radisch, J., & Jacobzone, S. (2013). *The use of social media in risk and crisis communication* (OECD Working Papers on Public Governance No. 24). https://doi.org/10.1787/5k3v01fskp9s-en

Zhaokai, Y., & Moffitt, K. C. (2019). Contract analytics in auditing. *Accounting Horizons*, *33*(3), 111–126. https://doi.org/10.2308/acch-52457

Zorio-Grima, A., & Carmona, P. (2019). Narratives of the big-4 transparency reports: Country effects or firm strategy? *Managerial Auditing Journal*, *34*(8), 951–985. https://doi.org/10.1108/MAJ-09-2018-1994

APPENDIX

Mapping Topics Related to the BOS Fund Management Cycle

Topic	BOS Fund Management Cycle
Topic 1 • Anak-Anak (Children), Guru Honorer (Honorary Teacher), Kuota Internet (Internet Quota), Madrasah, KJP Plus (Jakarta Smart Card Plus), and Kurikulum (Curriculum) are interpreted as matters related to implementation and administration. Children, honorary teachers, and Madrasahs are all key players in the teaching and learning process. Madrasahs, as educational units, are instrumental in organizing and conducting teaching and learning activities, and their role is further strengthened by the BOS funds they receive. KJP Plus (Jakarta Smart Card Plus) is a form of BOS companion funds that complements these efforts. As a guide, the curriculum ensures a structured and effective teaching and learning process. • Orang Tua (Parents), Dinas Pendidikan (Education Office), Kota Bandung (Bandung City), and Jawa Barat (West Java) are interpreted as matters related to guidance and supervision. Parents and the education office in Bandung City, West Java, supervise the management of BOS funds.	• Implementation and Administration • Guidance and Supervision
Topic 2 • Peserta Didik (Students) are interpreted as matters related to planning and budgeting. The number of students is used to calculate the allocation of BOS funds per school. • Saturn Pendidikan (education units), Tersangka (suspect), Madrasah, Penggunaan Dana (use of funds), Anak-Anak (children), and Guru Honorer (honorary teacher) are interpreted as matters related to implementation and administration. Education units and Madrasahs are the recipients and users of BOS funds. Children and honorary teachers are beneficiaries of BOS funds. The suspect is a party suspected of misusing BOS funds. The use of BOS funds is one part of the implementation of the BOS Fund management. • Dinas Pendidikan (education units), Orang Tua (parents), and Wali Murid (student guardian) are interpreted as matters related to guidance and supervision. Students' guardians play a role in supervising the management of BOS funds.	• Planning and Budgeting • Implementation and Administration • Guidance and Supervision
Topic 3 • Orang Tua (parents) and Nadiem (Minister of Education) are interpreted as matters related to guidance and supervision. The Minister of Education plays a role in supervising the management of BOS funds. • Guru Honorer (honorary teacher), Ruang Kelas (classroom), Paket Data (data package), Anak-Anak (children), and Sekolah Swasta (private school) are interpreted as matters related to implementation and administration. Data packages and classrooms are the outcomes of using BOS funds. Private schools are recipients and users of BOS funds. • Sri Mulyani (Minister of Finance), Peserta Didik (students), and Triliun (trillion) are interpreted as matters related to planning and budgeting. Sri Mulyani, as Minister of Finance, is the party authorized to approve the allocation of the BOS Fund budget.	• Planning and Budgeting • Implementation and Administration • Guidance and Supervision
Topic 4 • Guru Honorer (honorary teacher), Honorer (honorary), and Penyaluran Dana (fund distribution) are interpreted as matters related to implementation and administration. Fund distribution is one of the stages in the implementation of BOS Fund management.	• Planning and Budgeting • Implementation and Administration

Topic	BOS Fund Management Cycle
• Sri Mulyani (Minister of Finance) and Triliun (trillion) are interpreted as matters related to planning and budgeting. • Orang Tua (parents), Kaligis, Kabupaten Kota (city district), Pemerintah Pusat (central government), and Jawa Barat (West Java) are interpreted as matters related to guidance and supervision. Kaligis is a community figure who plays a role in overseeing the management of BOS funds.	• Guidance and Supervision
Topic 5 • Guru Honorer (honorary teacher), Gaji Guru (teacher salary), Honorer (honorary), Sekolah Swasta (private school), Gaji (salary), and Tersangka (suspect) are interpreted as matters related to implementation and administration. • Kaligis, Orang Tua (parents), Komite Sekolah (school committee), and Kabupaten Serang (Serang Regency) are interpreted as matters related to guidance and supervision.	• Implementation and Administration • Guidance and Supervision
Topic 6 • Guru Honorer (honorary teacher), Honorer (honorary), Daerah Tertinggal (underdeveloped regions), and Penyaluran Dana (fund distribution) are interpreted as matters related to implementation and administration. • Triliun (trillion) is interpreted as a matter related to planning and budgeting. • DKI, Pemprov DKI (DKI Provincial Government), Kota Bandung (Bandung City), Jawa Timur (East Java), and Khofifah (Governor of East Java) are interpreted as matters related to guidance and supervision.	• Planning and Budgeting • Implementation and Administration • Guidance and Supervision
Topic 7 • Anak-Anak (children), RUU Sisdiknas (national education system draft law), Guru Honorer (honorary teacher), Madrasah, Tatap Muka (face to face), Merdeka Belajar (independent learning), and Bank Jatim (East Java Bank) are interpreted as matters related to implementation and administration. • Orang Tua (parents) and Nadiem (Ministry of Education) are interpreted as matters related to guidance and supervision.	• Implementation and Administration • Guidance and Supervision
Topic 8 • Guru Honorer (honorary teacher), Anak-Anak (children), and Honorer (honorary) are interpreted as matters related to implementation and administration. • Orang Tua (Parents), Kota Bandung (Bandung City), Bandar Lampung, Kaligis, and Kabupaten Kota (City District) are interpreted as matters related to guidance and supervision. • Sri Mulyani and Triliun are interpreted as matters related to planning and budgeting.	• Planning and Budgeting • Implementation and Administration • Guidance and Supervision
Topic 9 • Orang Tua (Parents), Pemprov Jabar (West Java Provincial Government), ICW, Jabar (West Java), and Jawa Barat (West Java) are interpreted as matters related to guidance and supervision. • Anak-Anak (children), Penyaluran Dana (distribution of funds), Tatap Muka (face to face), Madrasah, and Siswa SMA (senior high school students) are interpreted as matters related to implementation and administration.	• Implementation and Administration • Guidance and Supervision
Topic 10 • Sekolah Swasta (private school), Guru Honorer (honorary teacher), Buku (books), Anak-Anak (children), Penyaluran Dana (distribution of funds), Bank BTN Bank of BTN), BOS	• Implementation and Administration

Topic	BOS Fund Management Cycle
(School Operational Assistance), BOP (Education Operational Assistance), and Madrasah are interpreted as matters related to implementation and administration. ● Kaligis and Agus are interpreted as matters related to guidance and supervision.	● Guidance and Supervision


Rp
BOS Fund
LOW
HIGH
AUDIT RISK ASSESMENT