A RESEARCH-BASED ONTOLOGY FOR COLLABORATIVE INNOVATION: A METHODOLOGY LEVERAGING AI AND DOMAIN EXPERT KNOWLEDGE

Faten Kharbat [1], Abdallah Alshawabkeh[2] and Mohammad Sharairi [2]

(Received: 17-Feb.-2024, Revised: 22-Apr.-2024, Accepted: 10-May-2024)

ABSTRACT

This paper introduces a method, for creating research-driven ontology to foster collaboration and innovation. The concept of collaborative innovation implies a process where multiple stakeholders work together to generate novel ideas, solutions or products. The suggested approach combines Artificial Intelligence (AI) and expert knowledge to build a comprehensive model encompassing various aspects of research, development and innovation. To demonstrate the feasibility of this method, the paper showcases its implementation in the field of accounting science. First, AI-powered machine-learning algorithms and text-mining techniques are used to extract the main ontological elements from a large corpus of accounting literature. Subsequently, expert knowledge is utilized to refine and validate these identified elements. The resulting ontology can be used as the foundation of a knowledge-based system to promote collaboration and analyze the state of innovation.

KEYWORDS

Artificial intelligence, Expert knowledge, Ontology, Research-based, Accounting, Text mining.

1. INTRODUCTION

The most commonly cited problem in developing expert systems is acquiring specific knowledge for a well-defined domain from experts and representing it in the appropriate digital format. Within the Artificial Intelligence (AI) field, this has been called the knowledge acquisition's problem and has been identified as a bottleneck in the process of building expert systems [1]-[2]. The exponential growth of data due to Industry 4.0 technologies necessitates capturing and transforming data into useful information for organizations [3]. It is worth mentioning that ontologies and knowledge graphs are used to represent complex knowledge. Knowledge graphs enable efficient querying and reasoning about complex data, supporting the development of intelligent agents that can learn from the represented knowledge [4]. A well-designed ontology describing the expert knowledge of a domain is at the core of many knowledge-based systems [5] to "support large-scale data/information interoperability, sharing of information and ontology-supported processes" [6]. According to Zhang and Li [7], a well-designed ontology is the key to building a successful knowledge-based system. However, [8] argued that systematic literature review highlighted the need for more structured development processes in knowledge-based systems, emphasizing knowledge elicitation and formalization, which could potentially involve ontology.

Ontology is a formal, explicit specification of a shared abstract representation of a real-world phenomenon by describing its relevant concepts, relations and axioms of a domain of interest [9]. At the beginning of the 1990s, computer science began to recognize ontology. It became an important area to investigate, especially in the area of artificial intelligence AI, because it was proposed as an effective method for creating representations of reality to be used later in the process of AI [10]-[11]. Ontologies enable inferences based on their content and relationships, simulating human inference capabilities [12].

Different ontology-development methods have been proposed, such as methods for new ontology development, new ontology alignment and merging ontology learning and re-engineering existing ontologies [13]-[14]. However, building and maintaining an ontology is considered a "labor-intensive process" [15] that costs time, money and effort. Different techniques and methods have been proposed for building new ontologies, including new-ontology development, alignment, merging, ontology

1. F. Kharbat is with Department of Computer Science, College of Engineering, Al Ain University, Abu Dhabi, UAE. Email: faten.kharbat@aau.ac.ae
2. A. Alshawabkeh and M. Sharairi are with College of Business, Al Ain University, Abu Dhabi, UAE. Emails: {Abdallah.Alshawabkeh, Mohammad.Sharairi}@aau.ac.ae

learning and re-engineering existing ontologies [3], [12]-[13], [15]-[16]. This research is developed to support collaborative innovation in ontology development by establishing a common language and conceptual framework that bridges interdisciplinary boundaries and enhances communication among diverse stakeholders. A well-designed ontology captures expert knowledge in a usable format, empowering stakeholders to leverage domain-specific insights, identify synergies and co-create innovative solutions. The ontology's capacity to organize and structure information improves information retrieval, idea generation and decision-making processes, fostering a collaborative environment conducive to innovation.

1.1 Current Status of Accounting Ontology

Organizational accounting knowledge is vital; therefore, Aparaschivei [17] outlined the academic and commercial advantages of creating an accounting ontology. The transactional accounting model with its associated value constraint as well as the asset liability equity resource types was examined in detail by Shannaq and Fatima [18], who conducted a hierarchical observation of accounting ontology. Developing a functional ontology for accounting is the first step in establishing a knowledge base for a discipline within an organization. A concept hierarchy was created to fully understand and simplify the accounting system [18].

Currently, there are several economic exchange ontologies, such as OntoREA [19], COFRIS [20] and ATE [21]. Most of them are derived from or refer to McCarthy’s REA accounting model, which is commonly used as a reference for accounting ontologies focusing on economic exchanges [19]. The REA is designed to unify accounting and management perspectives on accounting information systems (AISs) [22]. However, it has been argued that accounting is more about reporting economic exchanges. The REA accounting model refers to AISs that record these exchanges. The processing that is done in AISs, the way data is aggregated into financial reports and how the quality of data is assured are not included in the REA model [23].

As part of REA business ontology, Geerts and McCarthy [24] added a policymaking architecture to the REA accounting framework. This way of thinking allows for the consideration of company policies regarding acquisition, transformation, revenues, banking and investment transactions by extending economic reasoning from a prospective to a context and setting viewpoint. The consequences for the evolution of EIS, EIS built using social-networking sites, EIS built on the network and EIS compatibility between different types of businesses are brought into sharp focus by the comprehensive general framework of REA ontology (REA2) [25].

Any occurrence in a company’s operations that has a financial impact on its accounting information is considered as a financial-reporting transaction. This information can be found in a company’s books. Financial accounting requires that, after adjustments have been made, an entity’s revenues equal its total liabilities and its stakeholders’ equity. The accountability equation is the foundation of the dual-entry accounting information system. The basic accounting formula is assets = liabilities + shareholders’ equity [26]. To provide a uniform meaning for the subtraction and addition of assets, liabilities and equity inside the financial statement, the concepts of both debit and credit are required.

Accounting experts may utilize their expertise to automate accounting activities. To characterize the types, qualities and interactions of ideas in an area is the goal of the ontology method [17]. Intelligent applications can be used in the field of accounting to automate accounting information. The ontology-based automated transaction financial-reporting system was presented by Shen and Tijerino [27], with a particular emphasis on processing total count documentation, such as receipts, to satisfy the corporation requirements. This is time-consuming, because it requires physically sorting receipts into relevant categories for company-expenditure reports.

There are several clear benefits to developing an ontology for accounting transactions. Users of accounting records may benefit from this ontology by better comprehending the specific meaning of the accountancy-operation terminology. Developing an accounting-related understanding may also be grounded in the ontology of accounting transactions. Even though its theoretical roots are in accounting, REA economic ontology does not address core needs in that discipline. Schwaiger et al. [19] found that the REA commercial ontology lacked the necessary traditional accounting logic, such as documenting credited and crediting modifications in assets and obligations, all of which provided appropriate categories as part of accounting records. A gap between scientific work and actual accounting practices has been noted [23]. By using domain-specific expertise in accounting, one may streamline the overall bookkeeping process. As an example of such an intelligent user's use in the field of accounting, we consider the automation of financial statements.

Therefore, the key challenges that specialists confront are: (1) locating a complete domain ontology and (2) locating domain specialists to enlist in collecting the necessary domain ontology. Cognitive computing and information-extraction methods allow for the semi-automatic identification of ontology when reading subject materials, providing insight into these topics. Similar semi-automatically and intelligently produced accountancy subjects and ideas are essential for something like an agile framework that develops and expands a global financial intelligent system. It also eliminates proprietary information that is either too costly to protect or forbidden by law. Consequently, this study aims to develop a low-effort, high-return strategy for extracting the most value from a previously specified domain ontology.

2. METHODOLOGY

Owing to the nature of the research, the developed ontology requires a methodology that supports the integration of several components from different sub-results within the project. This research adopts the agile methodology for building an ontology proposed by Abdelghany et al. [28]. This supports the core activities of building ontologies [28]. In a comparative analysis of methodologies for domain ontology development conducted by Sattar et al. [29], various aspects of ontology development and documentation were compared, including ontology construction strategies and support for integration and merging. The adopted methodology had a high rank among other methodologies to build an ontology. [29] highlighted the importance of selecting a methodology that adequately addresses both core ontology-building activities and project-management requirements.

The research-based ontology crafted for collaborative innovation acts as a foundational framework that amalgamates intelligence from AI algorithms, text-mining techniques and expert knowledge to facilitate knowledge sharing, interdisciplinary collaboration and innovation management. This ontology serves as a structured representation of domain-specific concepts, relationships and constraints, enabling stakeholders to access, interpret and contribute to a shared-knowledge repository.

The methodology process [28] consists of three main stages: the pregame stage, the development stage and the postgame stage, as shown in Figure 1. The three stages involved a team of experts consisting of five experts in the fields of knowledge management, artificial intelligence and three from the accounting. The first (pregame) stage allows the team to identify the goal of the ontology, tools and techniques to be used while building it and to select the data-collection methodology. The next step was to formulate ontology requirements. Because the main aim of ontology is to build a research-based ontology, the research dataset was identified and selected in this phase, as explained in the data collection part.

The second (development) stage executes an iterative process for multiple unsupervised cycles of development and evaluation. During the second stage, different meetings were held to discuss the ontologies produced. At each meeting, changes in the ontologies were verified. These steps led to the final set of ontologies approved by all participants.

The last (postgame) stage should be followed by another iterative process for multiple supervised cycles, per experts' feedback. This cyclic process and results are an essential step in contributing to improving knowledge acquisition for the representation of scientific knowledge in ontologies. The final stage (postgame stage) allows verification and validation processes, along with documentation.

To elucidate the relationship between lexicon elicitation and ontology development, it is crucial to emphasize how defining domain-specific terms and concepts contributes to building a robust ontology. Lexicon elicitation involves extracting and defining domain-specific terms, essential for constructing an ontology that accurately represents the domain's knowledge landscape. By structuring expert knowledge, the ontology can organize information, facilitate knowledge sharing and support collaborative innovation. Utilizing AI-powered machine-learning algorithms and text-mining techniques in lexicon elicitation helps extract ontological elements from domain literature, identifying key concepts and relationships foundational to the ontology. Expert-knowledge validation further refines these elements, ensuring the ontology's accuracy and relevance in fostering collaboration and innovation.

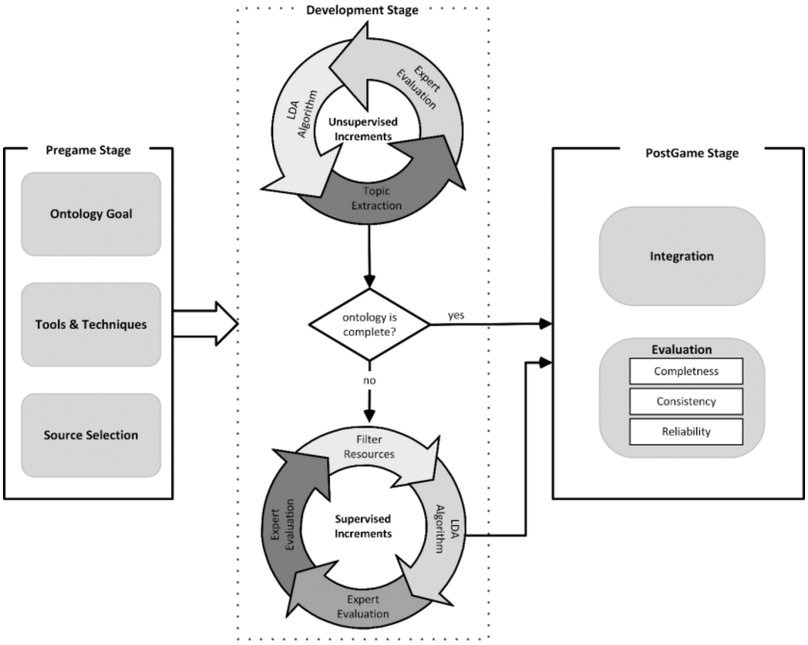

Figure 1. The research methodology as adopted from [28].

3. Results and Discussion

3.1 Pregame Stage

3.1.1 Goal & Scope

The fundamental scope of accounting has not changed over time; nevertheless, new innovative accounting directions have evolved [30]. The accounting domain was selected due to its inherent complexity, evolving nature and critical role in various industries. The compilation of reports, transaction recording, cyber-auditing and other technological advancements have grown. As a result, different specialized accounting topics have emerged. Thus, by developing an accounting ontology, the paper aims to establish a structured framework that captures and organizes domain-specific knowledge, facilitating collaboration and innovation within the accounting field. The goal of building the proposed accounting ontology is to establish a common understanding of the meaning of the terms used by researchers to support the knowledge-acquisition process of innovative directions related to accounting in the current-research directions. [31] explores the accounting dimensions of innovations, highlighting the interconnectedness of accounting practices with innovative processes. In addition, [32] discusses the identification of misstatement accounts in financial statements through ontology reasoning. It demonstrated the application of ontology-based decision-support systems in the accounting domain, emphasizing the importance of ontology reasoning in financial-statement analysis. In fact, the selected literature reinforces the importance of ontology-based approaches in addressing accounting challenges, thereby justifying the choice of the accounting domain.

The proposed accounting ontology can be used as a tool to map the concepts and services used by expert systems to align with the most recent information. The scope of accounting ontology focuses on the main specialized topics of accounting in the recent literature: financial accounting, management accounting [33], cost accounting [33], tax accounting [34] and auditing [34]. Other topics may have been included; however, the experts preferred not to expand the search for more detailed ones.

3.1.2 Tools

The Protégé ontology editor and framework in the OWL language were selected as the main tools to formalize the accounting ontology. The selection was based on previous research, such as [28] and its well-known ability to customize and extend.

3.1.3 Source Selection

To align with the research aims, the proposed ontology is based on a research-based source and the research dataset is identified and selected as follows: To collect all research written in "Accounting," three main sources were identified [33], [35]: ABDC, Scopus and Web of Science. All journals indexed in ABDC, Scopus and the Web of Science related to this field were identified and selected by September 2021. From the Scopus index, 154 journals related to the accounting discipline were published by 46 publishers. The journal citation scores ranged from 0 (e.g. Journal of Taxation, SJR=0.1) to 10.3 (e.g. Journal of Finance, SJR= 17.134). In the ABCD list, 2679 journals related to the accounting discipline were published by 744 publishers. The journal ranks ranged from A* (199 journals, e.g. Australian Tax Forum) to C (982 journals, e.g. Real Estate Taxation). In the Web of Science, 18 journals related to the accounting discipline were published by 12 publishers. Of course, there were duplicate findings between the three main sources, which will be discussed later.

All research published in journals from 1946 to 2021 was automatically collected using a script written in Python 3.8.5. The total number of papers is 209,345. However, the data collected was filtered, because some documents were not journal papers or their abstracts were not electronically available. After filtering, the total number of journal papers was 159,239, with available abstracts. All collected titles and abstracts were injected into an empty dataset to be processed *via* machine-learning and text-mining techniques.

3.2 Development Stage

The development stage aims to produce pieces for the accounting ontology from the final dataset processed in the previous stage to integrate them in the third stage. The development stage consisted of several increments to allow different levels of information extraction from the available dataset containing titles and abstracts. This process was introduced and implemented in [36], where the iOntoBioethics ontology for Bioethics Ontologies in Pandemics was proposed and evaluated.

The relation between the extracted ontology topic/concept and the dataset is generally based on proportion. For example, the topic of "cost accounting" is extracted as an important topic based on the number of times it is mentioned in different abstracts. If mentioned in one abstract twice, this does not mean that it is as important as if mentioned once in two abstracts. Each given abstract relates to the discovered accounting topic/concepts with different proportions. To work with these proportions, certain factors must be investigated: accounting topics per abstract and assignment of accounting concepts per accounting topic in an abstract. Hence, in the current development stage, the latent Dirichlet allocation (LDA) [37]-[38] algorithm in the field of machine learning (text mining) was utilized to illustrate the topics and their related concepts within the abstracts. The stage started with unsupervised increments, followed by supervised increments, as shown below. To illustrate the outcomes from the supervised and unsupervised iterations, Protégé was used, as mentioned in sub-section 3.1. Every topic is mapped to a class and every concept is mapped to a property associated with the class.

3.2.1 Unsupervised Increments

This stage consisted of three iterations, in which the LDA algorithm automatically identified accounting topics and their associated concepts. An iterative automated process was conducted to determine the maximum coherence value for the best number of topics to avoid bias in executing the LDA algorithm.

Iteration 1: Preparation. The first increment was a preparation increment, in which the dataset containing all abstracts and titles was fed into a customized Python tool, the Standardization Text Characteristics (SRC) with a reliance process [39] was directed and the LDA algorithm was executed to extract the general topics and their associated concepts. The SRC process converted all upper-case characters into lower-case characters, removed all stop words such as "the," "on," …etc., removed all the white spaces and removed punctuations. The output of the first iteration shows that six topics were extracted, as shown in Figure 2.

Three accounting-domain specialists were involved in investigating the six topic structures to validate the output of the first iteration, arriving at a consensus on the extent which these topics represent. Three experts evaluated the six topics separately and the research team collected their responses. The main comments from the three experts focused on unrelated concepts within the extracted topics. For example, in topic 1, the words “research”, “new” and “case” were considered irrelevant. Topic 3 has the words “study” and “purpose” that “weaken the aggregation concept,” as one of the experts expressed.

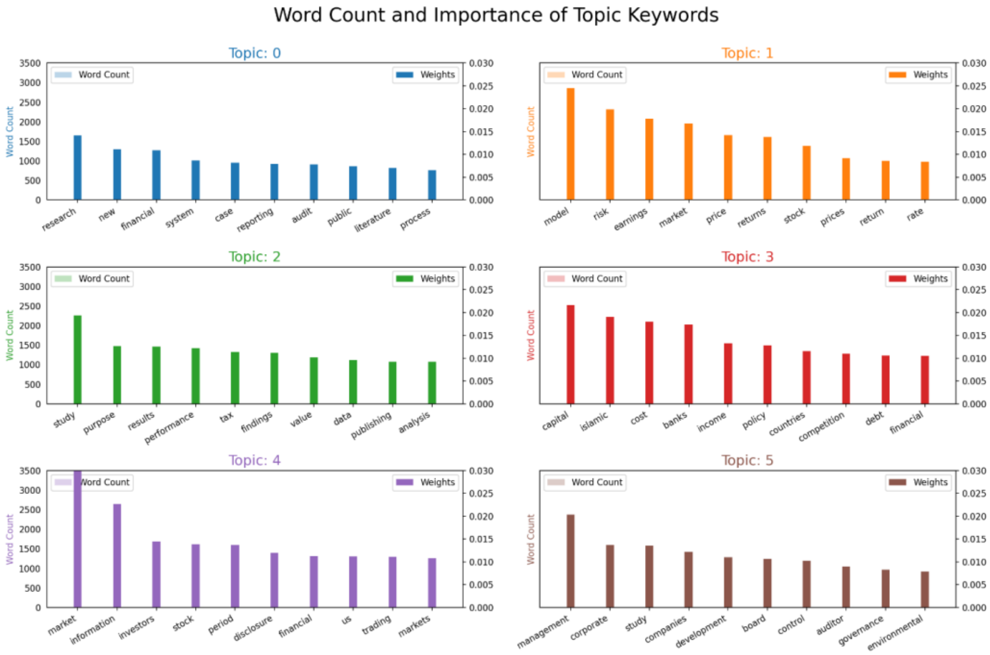

Figure 2. The topics extracted from the preparation interaction.

Iteration 2: Enhancement. The second iteration involved enhancing the basic environment for the dataset of abstracts and titles. The LDA algorithm was executed several times to detect new unrelated words and phrases to enhance the final results. All intermediate results were provided to the three experts for validation and recommendations. All unrelated words recommended by the experts from the first iteration were added to the set of stop words. The final stop-word list is presented in Table 1.

Table 1. The stop words included in the experiments.

```
'i', 'me', 'my', 'myself', 'we', 'our', 'ours', 'ourselves', 'you',
"you're", "you've", "you'll", "you'd", 'your', 'yours', 'yourself',
'yourselves', 'he', 'him', 'his', 'himself', 'she', "she's", 'her',
'hers', 'herself', 'it', "it's", 'its', 'itself', 'they', 'them',
'their', 'theirs', 'themselves', 'what', 'which', 'who', 'whom', 'this',
'that', "that'll", 'these', 'those', 'am', 'is', 'are', 'was', 'were',
'be', 'been', 'being', 'have', 'has', 'had', 'having', 'do', 'does',
'did', 'doing', 'a', 'an', 'the', 'and', 'but', 'if', 'or', 'because',
'as', 'until', 'while', 'of', 'at', 'by', 'for', 'with', 'about',
'against', 'between', 'into', 'through', 'during', 'before', 'after',
'above', 'below', 'to', 'from', 'up', 'down', 'in', 'out', 'on', 'off',
'over', 'under', 'again', 'further', 'then', 'once', 'here', 'there',
'when', 'where', 'why', 'how', 'all', 'any', 'both', 'each', 'few',
'more', 'most', 'other', 'some', 'such', 'no', 'nor', 'not', 'only',
'own', 'same', 'so', 'than', 'too', 'very', 's', 't', 'can', 'will',
'just', 'don', "don't", 'should', "should've", 'now', 'd', 'll', 'm',
'o', 're', 've', 'y', 'ain', 'aren', "aren't", 'couldn', "couldn't",
'didn', "didn't", 'doesn', "doesn't", 'hadn', "hadn't", 'hasn', "hasn't",
'haven', "haven't", 'isn', "isn't", 'ma', 'mightn', "mightn't", 'mustn',
"mustn't", 'needn', "needn't", 'shan', "shan't", 'shouldn', "shouldn't",
'wasn', "wasn't", 'weren', "weren't", 'won', "won't", 'wouldn',
"wouldn't",'from','subject','astract','avaible','a','find','firm','firms'
,'accounting','paper','accountent','Research','literature',      'new',
'study',   'purpose','Findings',’rights','reserved','well','methodology',
'design', research', 'approach','findings','finding', 'company','used',
'use','uses','also','de','part','parts','find','white','wisdom','elsevier
ltd', 'academic press limited', 'american accounting association', 'by de
la salle university','all rights reserved'
```

Iteration 3: Output Production. The execution of the third iteration results in four topics, as shown in Figures 3 and 4. In Figure 3, the number of documents related to the topics is shown, whereas in Figure 4, the concepts related to each topic are illustrated.

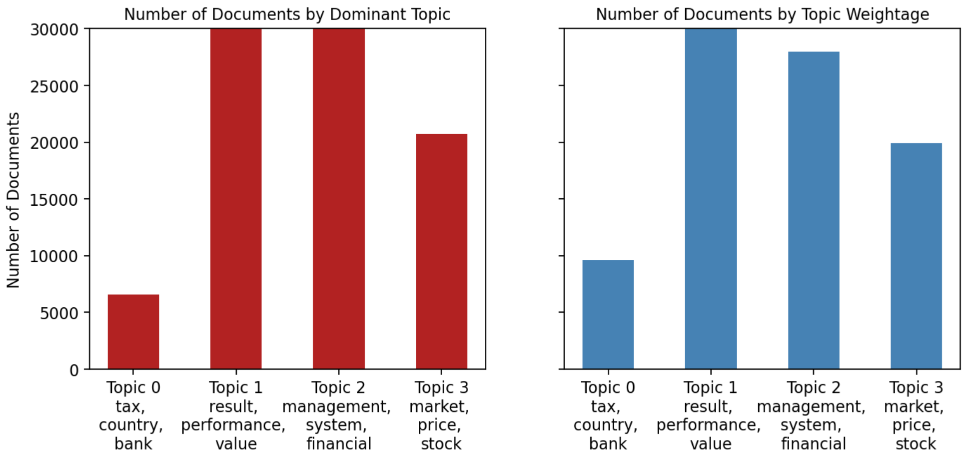

Figure 3. The number of documents related to each extracted topic (iteration 3).

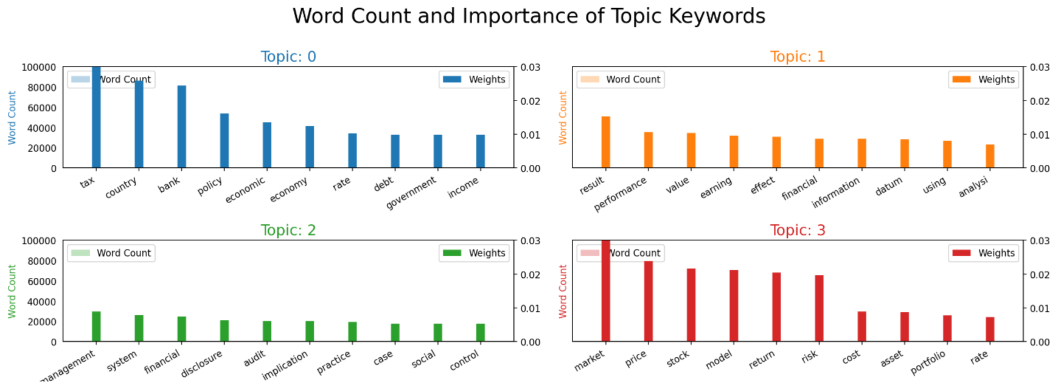

Figure 4. The main four extracted topics with their associated concepts (iteration 3).

It is worth noting that the LDA algorithm does not name the extracted topics, but assigns them numbers within the range of related topics. Therefore, the three experts were involved later in independently determining the titles of the topics that aggregated the shown concepts.

At the end of the third iteration, the three experts were given the generated topics and the related concepts to evaluate them independently. They were also asked to suggest a title for each topic. Consequently, the results of this process are as follows:

- ***Topic 0: Tax Accounting***

The terminology on this topic refers to tax accounting. Tax accounting checks whether businesses adhere to all rules set out by tax authorities. Economic data and other revenue data are subject to reporting by tax accountants to tax authorities. The accounting and tax standards used by each nation are different. Long-term economic advancement and resource allocation are susceptible to taxation considerations [40].

- ***Topic 1: Financial Accounting***

The terminology for this topic expresses the topic of financial accounting. Financial accounting is a sub-field of financial reporting that focuses on disseminating a firm's financial information to stakeholders, including stockholders, researchers, suppliers and regulatory agencies. Statement of income, financial statements, cash-flow statements and cash-position declarations are the four fundamental audited financials. This accounting information is the foundation for evaluating a corporation's economic health and competitive position. In addition, people use such data to decide whether or not to invest in a firm. Investors and traders utilize such accounting records in the financial sector to evaluate the health and potential returns of publicly-traded businesses and the stock market and brokerage houses [41].

- ***Topic 2: Auditing***

The terminology for this topic mainly indicates auditing. An audit examines and assesses the effectiveness of internal company controls and financial-analysis processes. External and internal accountants perform this function. The public relies on a firm's publicly released income statement when making a variety of financial choices. Conversely, internal audits directly report to managers and supervisors inside an organization. A company's auditors examine whether senior-management directives are being followed. When conducting an internal review, it is crucial to determine whether a company's actions are in accordance with its own stated objectives [42].

- ***Topic 3: Financial Accounting - Portfolio Investment***

The terminology of this topic points toward portfolio investment (finance), which relies on financial-accounting information. Sector-specific breakdowns in private investment, investment spending and foreign reserves are all part of a complete financial statement. When investing in stocks, bonds and other marketable securities to earn a profit, grow in valuation or combine them, you are creating a portfolio. This implies a less proactive managerial position than direct investment by way of the passive ownership of assets. When valuing a company or conducting a financial analysis, shareholders rely heavily on the data provided in its financial reports. As a result, it is crucial to understand the fundamentals of business accounting as well as the rules governing the creation of financial statements. Accounting is advantageous to investors, because it allows them to assess the worth of a firm's profits, learn about its financial products, measure its financial performance and gauge the risks inherent in the income statement [43].

Visual Illustration

To illustrate the topics and concepts from the third iteration of the unsupervised increments, the topics and concepts are implemented in Protégé as follows.

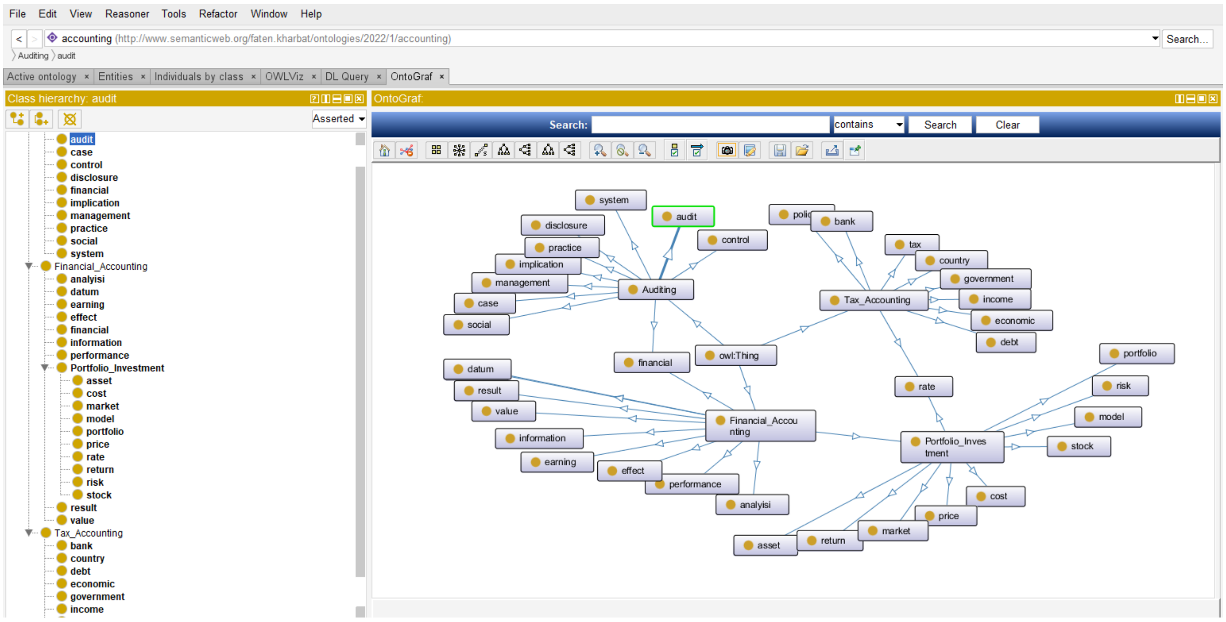

Figure 5. The draft of the ontology from the unsupervised increments.

3.2.2 Supervised Increments

Ontology validation was discussed by Quinn and McArthur [44] to evaluate whether the ontology matches the world model as demonstrated by the industry dataset [44]. Different methods have been used to measure the validity of an ontology [44]–[46]. However, a very strong approach was proposed and implemented by Quinn and McArthur [44], which included qualitative and quantitative approaches to express ontology completeness and expressiveness. Completeness indicates the sufficiency of the semantic relationship between the extracted topics/concepts and the accounting domain. To measure the completeness of the resulting concepts, an anonymous survey was distributed to the accounting experts that included a table with all the topics and related concepts to state whether the topics/concepts were related to the accounting domain and to what extent (0 = not related and 10 = entirely related). The average number of experts was then calculated on the agreement that any average less than 7.0 would be excluded from the results.

Regarding the topics, the average number of experts was 9.4/10. The minimum average mark was given to the topic of portfolio investment (7.6/10). Regarding the concepts, the average from the experts was 8.4/10, indicating that the extracted concepts were highly related to the topics and accounting domain.

The next step was to measure the expressiveness of the current version of the ontology. The expressiveness of an ontology, the so-called "coverage percentage" [28], is defined as "quantifying the number of key relationships required" in a domain [44]. To conduct such measurements, a list of known accounting topics and some emerging ones were listed from different sources, such as Shkulipa [33] and the experts were asked whether they would think that any of them would be essential to be added. At least 70% of the experts agreed to manually add the main topics: cost, managerial accounting and forensic accounting.

Therefore, a supervised iteration was implemented to extract related concepts for missing topics. To conduct supervised iteration, a separate iteration for each topic was executed to extract the related concepts. For each iteration, the dataset, including all titles and abstracts, was filtered to include only that topic in particular. The results of each iteration are as follows.

- **Cost Accounting**

The three experts were again given the topics generated from the supervised iteration to independently evaluate the topics and their concepts and suggest a couple of titles for each topic. Consequently, the conclusion from the supervised iterations for the "cost accounting" was identified as follows:

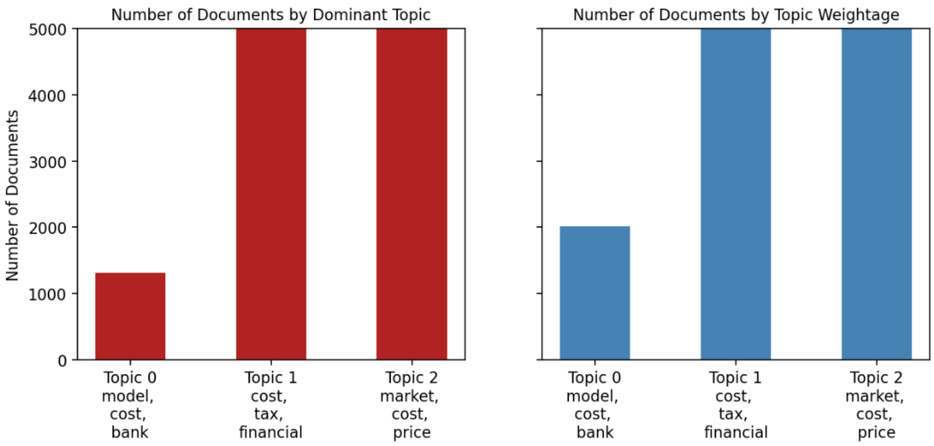

Figure 6. The number of documents related to each extracted topic of cost accounting.

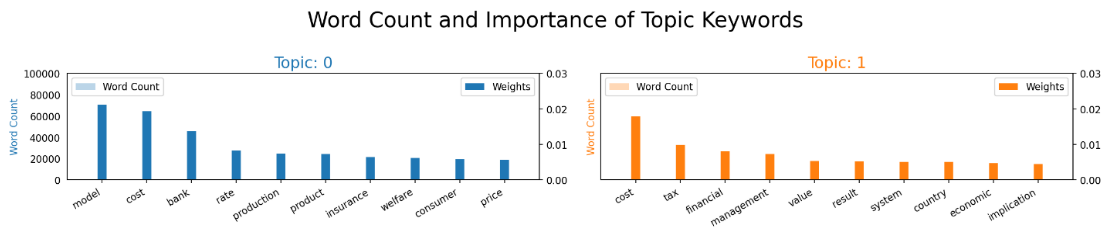

Figure 7. The generated topics for cost accounting.

- Topic 0: Cost Accounting

The terminology used here refers to cost accounting. In cost accounting, to maximize revenue and enhance the efficacy of business processes, standard costing (also known as control accounting) generates data to be utilized in the improvement of the organization. Cost accounting entails taking care of all the money that comes up in the operation of a company. Management uses cost information to prepare and manage various cost activities. Standard costing is concerned with collecting, categorizing and interpreting cost data in a quantifiable manner. The primary objective of cost accounting is to collect and analyze a firm's variable and constant expenses. From an administrative perspective (marketing, transportation organization, protection, manufacturing, …etc.), both direct and indirect materials and direct and indirect employees are the main components of the costing system [47].

○ *Topic 1: Miscellaneous Topics*

It is difficult to give a specific title to this group, as the terminologies are related to different general accounting topics, such as cost, tax, financial and management.

- **Managerial Accounting**

The three experts were again given the topics generated from the supervised iteration to independently evaluate the topics and their concepts and suggest a couple of titles for each topic. As a result, the conclusions from the supervised iterations for "managerial accounting" were identified as follows:

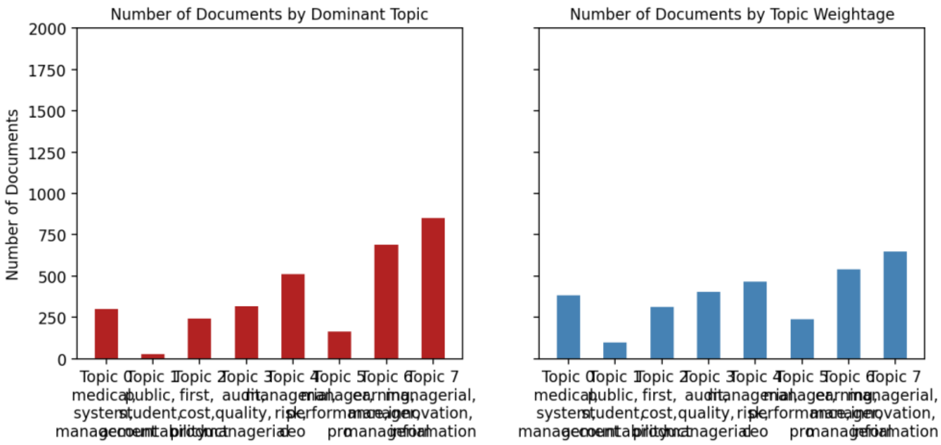

Figure 8. Number of documents related to each extracted topic for managerial accounting.

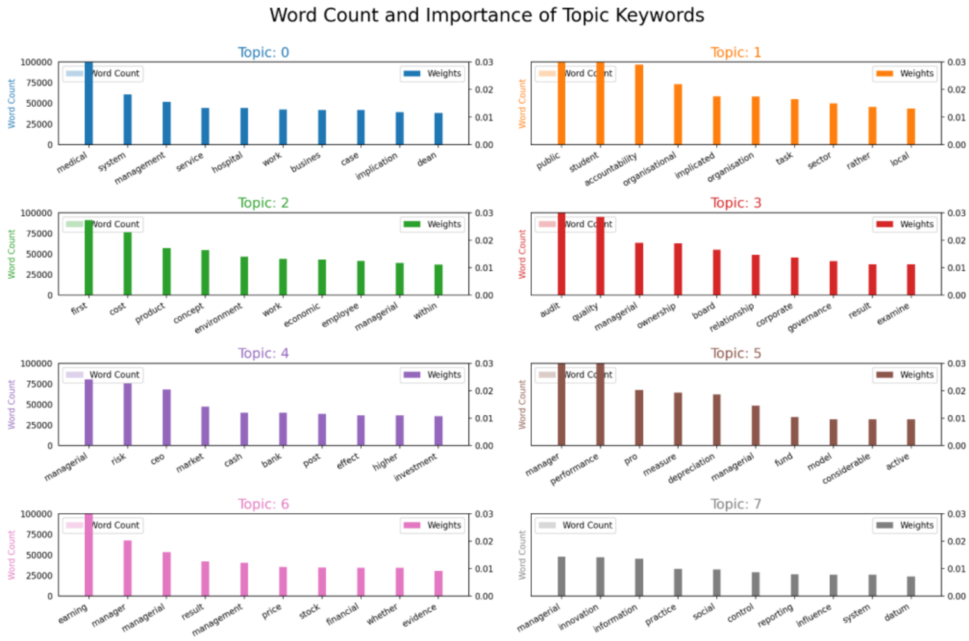

Figure 9. The topics extracted for managerial accounting.

○ Topics 0, 1, 4, 5, 6 and 7

It is difficult to give a specific accounting title for this topic, as the results are not accounting-related terminologies or are related to different general accounting topics.

○ *Topic 2: Managerial Accounting*

The terminology on this topic refers to managerial accounting. Managerial accounting, as opposed to financial accounting, includes the dissemination of financial information to a company's operational units. It is a sub-set of accountancy that focuses on analyzing financial data to generate financial accounting documents and reports to aid in the judgment procedure of heads of departments as well as top management. Managerial accounting clarifies the company's monetary information and delivers meaningful numbers and statistics to higher management and decision-makers.

Executives and departments, such as sales, marketing and production, may request custom reports that meet their unique reporting requirements. These reports combine actual and predicted data to provide managers with a wealth of information to make better business choices. In contrast to financial accounts, which are made public as well as publicized, data packets are used internally to enhance procedures, including total profit appraisal, departmental planning and other similar activities [48].

○ *Topic 3: Auditing*

Discussed previously. Adding some concepts: ownership, governance and quality.

- **Forensic Accounting**

Once again, the three experts were given the topics generated from the supervised iterations to independently evaluate the topics and their concepts and suggest a couple of titles for each topic. As a result, the conclusion from the supervised iterations for "forensics accounting" was identified as follows:

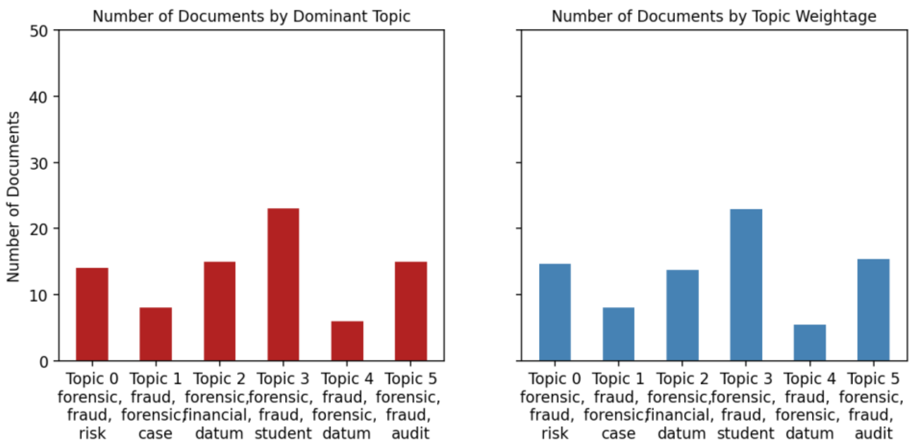

Figure 10. Number of documents related to each extracted topic for forensic accounting.

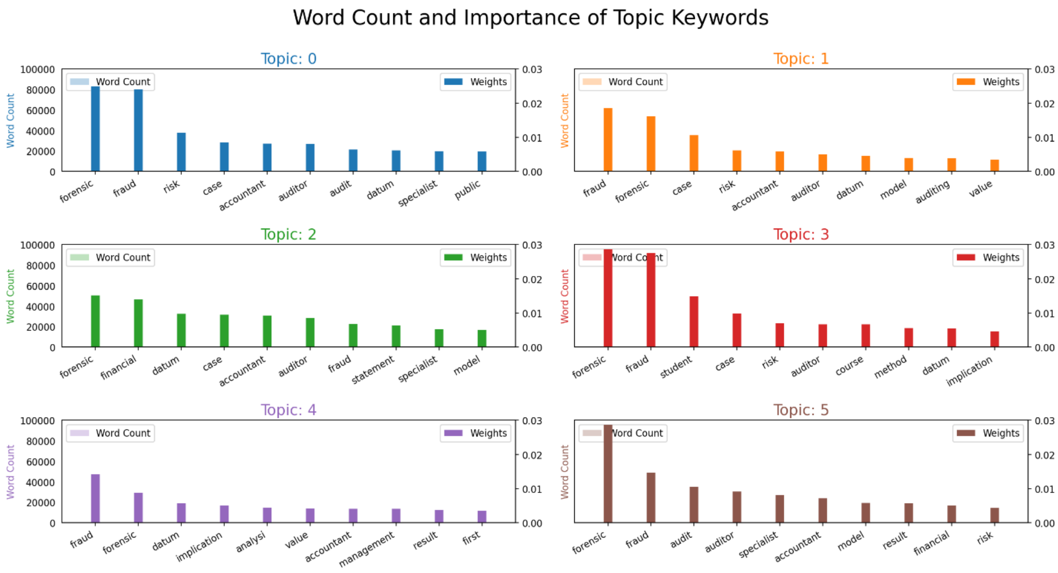

Figure 11. The topics extracted for forensic accounting.

○ *Topic 0, 1 & 2: Forensics Audit*

The terminology for this topic is related to forensic auditing. During a forensic audit, experts apply their knowledge of accounting to investigate cases that may have criminal consequences. The scope of the full investigation exceeds that of standard forensic accounting. Forensic auditing investigates the nature of the transactions themselves and identifies signs of possible asset theft [49].

○ *Topic 3: Accounting Education-Forensics*

The terminology used here refers to accounting education, specifically in forensics, as course content. Accounting education aims to prepare students for jobs, as accounting professionals are the ultimate goal of financial reporting. There has been an increase in interest in the accounting profession from the ranks of professional accounting organizations, which seek to blend theoretical studies with hands-on experience as well as outline specifics of required coursework [50]-[51].

○ *Topic 4 & 5: Forensic Accountant*

The terminology of this topic mainly refers to forensic accountants. Forensic accounting is used to investigate a person's or company's financial situation. Forensic accountants use a wide range of techniques in the accounting, reporting and investigation industries. Accounting information is frequently referred to as a sub-set of accounting. Experts witness that testimony is common for forensic accountants, who also conduct investigations of financial information that could be admissible in court. In court, professionals may demonstrate the monetary elements of crimes fraudulently. There is a growing need for forensic-accounting professionals in various fields, including law-enforcement agencies, auditing firms and insurance organizations [52].

3.3 PostGame Stage

The final stage of the adopted methodology (as described in Section 2) consists of two main steps: integration and evaluation. The integration process in this context is simple, because the outputs from the supervised and unsupervised iterations are based on the same dataset and for the same domain. It is unlikely that there is any inconsistency between topics or concepts. Every topic is mapped to a class and every concept is mapped to a property associated with the class. The only relation demonstrated was "is-a" from the main domain. The new version of the ontology is demonstrated using Protégé, as shown in Figure 12:

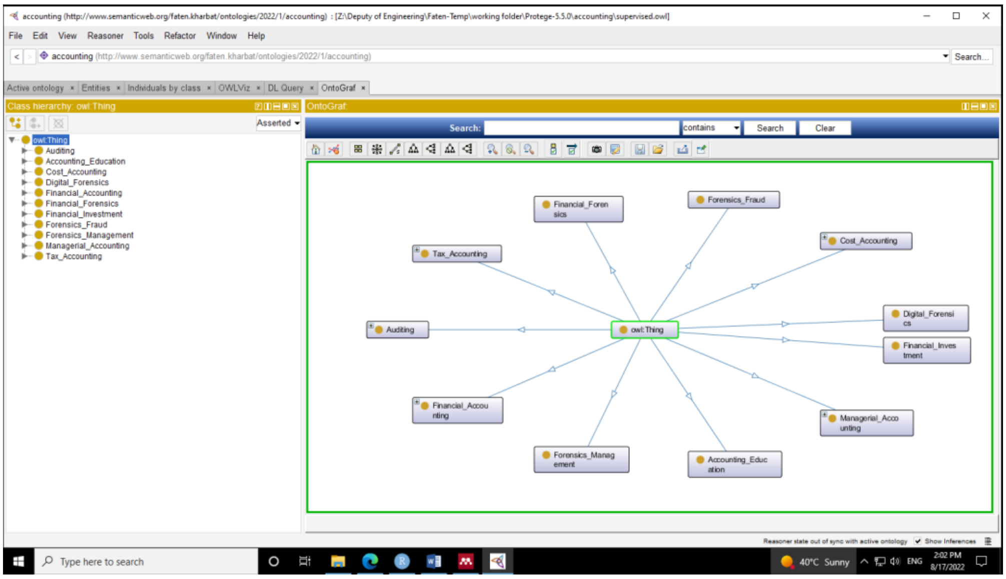

Figure 12. The general view of the new version of the ontology.

While the development stage is pivotal in establishing the ontology's foundation, the postGame stage plays a critical role in assessing the ontology's utility, identifying areas for improvement and ensuring its alignment with the intended objectives. Figure 13, created towards the end of the postGame stage, serves as a visual representation of key components, relationships or insights derived during the ontology-construction process. Drawing insights from the research-based ontology methodology for collaborative innovation, the postGame stage involves a comprehensive assessment of the ontology's utility in facilitating knowledge sharing, interdisciplinary collaboration and innovation management. Figure 13 encapsulates the culmination of the ontology-development process, showcasing the refined ontology structure, validated concepts and the integration of AI-powered machine-learning algorithms and expert knowledge to create a robust knowledge-based system.

The evaluation process is implemented again as in the previous stage: qualitative and quantitative approaches to express ontology completeness and expressiveness. To measure the completeness of the final draft of the ontology, a concise anonymous survey was distributed to a group of accounting experts that included a table with all the topics and related concepts to state whether the topics/concepts were related to the accounting domain and to what extent (0 was not related and 10 was entirely related). The average number of experts was then calculated. Regarding topics, the average number of experts was 8.8/10. Regarding the concepts, the average from the experts was 7.5/10, indicating that the extracted concepts were highly related to the topics and accounting domain.

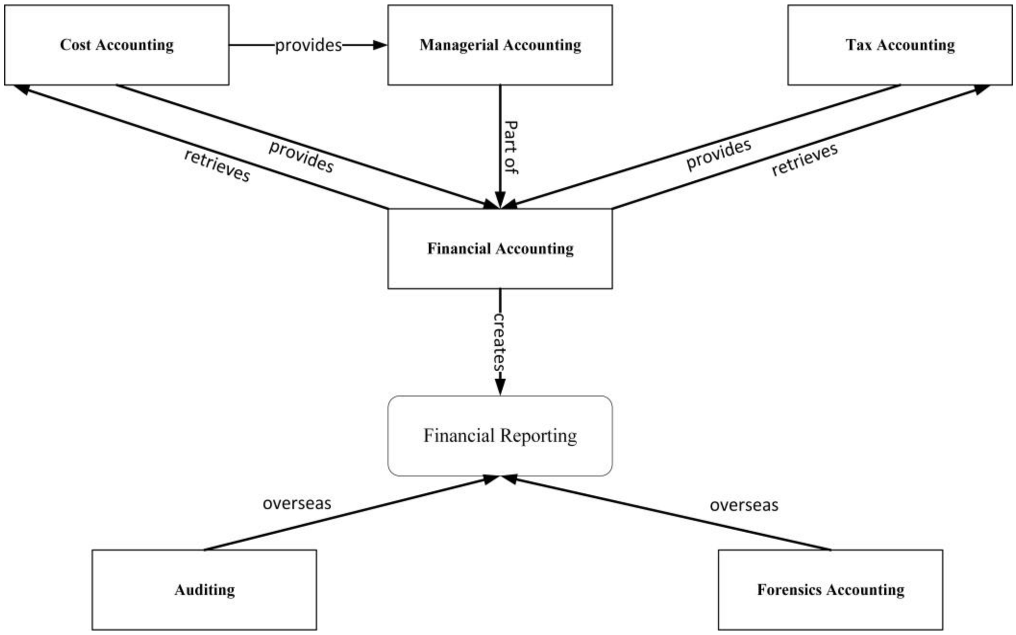

Figure 13. The final version of the generated ontology including the supervised and unsupervised iterations.

4. LIMITATIONS

While the integration of machine-learning techniques within the AMOD methodology presents promising opportunities for enhancing ontology development, several limitations warrant consideration. Firstly, as highlighted in previous research, the knowledge-acquisition process remains a critical challenge in ontology engineering. Despite leveraging machine learning for automated knowledge extraction, the reliance on existing data sources and expert input may limit the scalability and generalizability of the ontology across diverse domains.

Secondly, as discussed in the literature, the balance between automated machine learning -driven processes and expert validation is crucial for ensuring the ontology's relevance and reliability. The potential risk of overlooking domain difficulties or context-specific traces in favour of automated ML algorithms underscores the importance of maintaining a robust expert-driven validation process.

Lastly, the feasibility and effectiveness of the proposed methodology may vary across different domains and data-availability scenarios. While the research demonstrates the integration of machine learning within AMOD in a specific context, the transferability and performance of the approach in domains with limited data or distinct characteristics require further investigation. The adaptability of the methodology to diverse domains and the robustness of the machine-learning components in handling domain-specific complexities represent areas for future exploration and refinement.

5. CONCLUSIONS

Different ontology development methods have been proposed in the literature to build and maintain a comprehensive ontology, such as new ontology alignment, merging ontology learning and re-engineering existing ontologies. However, these methods are considered labor-intensive, and are incapable of describing the specific requirements of different new-research directions. Therefore, this research utilizes one of the Agile Methodologies for Ontology Development (AMOD) to develop an ontology and integrate it with machine-learning techniques. This adaptation signifies an evolution of the existing methodology to leverage machine learning capabilities, highlighting the importance of clearly articulating the degree of innovation in the research. The incorporation of machine-learning components within AMOD can enhance the ontology-development process by introducing automation, predictive analytics or pattern recognition, thereby improving efficiency and accuracy. The developed ontology for collaborative innovation plays a pivotal role in facilitating knowledge sharing, interdisciplinary collaboration and innovation management within the research domain. By serving as a structured-knowledge repository, promoting communication among stakeholders and aligning with agile ontology-development principles, the ontology contributes to creating a dynamic ecosystem that nurtures collaborative innovation and propels research advancement.

To achieve this, as a proof of concept, the related literature on accounting was collected and analyzed from among the most influential and well-cataloged works in the accounting field since their inception in 1945. This has helped gain comprehensive coverage of the main concepts required for optimized text analysis and computational-modeling technology. The proposed method automatically detects accounting-related topics and their associated concepts, thereby empowering the derived ontology. As recent findings, regulations, laws, quality assessments, …etc. have been released, an ontology of this kind has to adapt and accommodate fresh accounting topics and concepts that arise. These intelligently generated accounting topics and concepts established by artificial intelligence are fundamental to developing an intelligent accounting system.

For future work, the research can explore the integration of knowledge graphs alongside ontologies to enhance the representation and utilization of complex, heterogeneous data in the accounting domain. Leveraging knowledge graphs can provide a more comprehensive and interconnected view of accounting concepts, enabling advanced querying, reasoning and knowledge discovery. Additionally, incorporating knowledge graphs can support the development of intelligent systems that can learn and reason from the rich knowledge encapsulated in the graph, thereby enhancing the paper's focus on ontology development for collaborative innovation. Lastly, the feasibility and effectiveness of the proposed methodology may vary across different domains and data-availability scenarios. While the research demonstrates the integration of machine learning within AMOD in a specific context, the transferability and performance of the approach in domains with limited data or distinct characteristics require further investigation. The adaptability of the methodology to diverse domains and the robustness of the machine-learning components in handling domain-specific complexities represent areas for future exploration and refinement.

REFERENCES

[1] J. A. Oravec, "Experts in a Box : Expert Systems and Knowledge-based Engineering (1984–1991)," Chapter in Book: Historical Instructional Design Cases, pp. 253–270, DOI: 10.4324/9780429330995-14, Nov. 2020.

[2] W. P. Wagner, J. Otto and Q. B. Chung, "Knowledge Acquisition for Expert Systems in Accounting and Financial Problem Domains," Knowledge-based Systems, vol. 15, no. 8, pp. 439–447, Nov. 2002.

[3] D. Spoladore, E. Pessot and A. Trombetta, "A Novel Agile Ontology Engineering Methodology for Supporting Organizations in Collaborative Ontology Development," Computers in Industry, vol. 151, p. 103979, DOI: 10.1016/j.compind.2023.103979, Oct. 2023.

[4] A. Holzinger et al., "Human-in-the-Loop Integration with Domain-knowledge Graphs for Explainable Federated Deep Learning," Proc. of the Int. Cross-Domain Conf. for Machine Learning and Knowledge Extraction, Part of the Book Series: Lecture Notes in Computer Science, vol. 14065, pp. 45–64, 2023.

[5] T. R. Gruber, The Acquisition of Strategic Knowledge, ISBN: 978-0-323-16258-6, Elsevier, 2013.

[6] K. I. Kotis, G. A. Vouros and D. Spiliotopoulos, "Ontology Engineering Methodologies for the Evolution of Living and Reused Ontologies: Status, Trends, Findings and Recommendations," Knowledge Engineering Review, vol. 35, p. e4, DOI: 10.1017/S0269888920000065, Jan. 2020.

[7] L. Zhang and J. Li, "Automatic Generation of Ontology Based on Database Trajectory Analysis View Project Automatic Generation of Ontology Based on Database," J. of Computer Information Systems, vol. 7, pp. 1148–1154, 2011.

[8] P. Kügler, F. Dworschak, B. Schleich and S. Wartzack, "The Evolution of Knowledge-based Engineering from a Design Research Perspective: Literature Review 2012–2021," Advanced Eng. Informatics, vol. 55, p. 101892, DOI: 10.1016/j.aei.2023.101892, Jan. 2023.

[9] R. Studer, V. R. Benjamins and D. Fensel, "Knowledge Engineering: Principles and Methods," Data

Knowl. Eng., vol. 25, no. 1–2, pp. 161–197, DOI: 10.1016/S0169-023X(97)00056-6, Mar. 1998.

[10] F. M. Mendonça, A. M. P. Cardoso and E. Drumond, "Ontology of Application into the Domain of Mortality: A Tool Support for Filling out the Death Certificate," Ciência da Informação, vol. 39, no. 3, pp. 23–34, DOI: 10.1590/S0100-19652010000300002, 2010.

[11] R. Kishore and R. Sharman, "Computational Ontologies and Information Systems I: Foundations," Communications of the Association for Information Systems, vol. 14, pp. 158–183, 2004.

[12] D. Spoladore et al., "A Review of Domain Ontologies for Disability Representation," Expert Systems with Applications, vol. 228, p. 120467, DOI: 10.1016/j.eswa.2023.120467, Oct. 2023.

[13] O. Corcho, M. Fernández-López and A. Gómez-Pérez, "Methodologies, Tools and Languages for Building Ontologies. Where Is Their Meeting Point?," Data Knowl. Eng., vol. 46, no. 1, pp. 41–64, DOI: 10.1016/S0169-023X(02)00195-7, Jul. 2003.

[14] A. Gómez-Pérez, "Ontology Evaluation," Chapter of Book: Handbook on Ontologies, pp. 251–273, DOI: 10.1007/978-3-540-24750-0_13, 2004.

[15] N.-W. Chi, Y.-H. Jin and S.-H. Hsieh, "Developing Base Domain Ontology from a Reference Collection to Aid Information Retrieval," Automation in Construction, vol. 100, pp. 180–189, Apr. 2019.

[16] S. Al-Fedaghi, "Toward Flow-based Ontology," Studies in Computational Intelligence, vol. 653, pp. 125–137, Springer, 2016.

[17] F. Aparaschivei, "The Importance of an Accounting Ontology," Economic Informatics, vol. 1, no. 1–4, 2007.

[18] B. Shannaq and K. Fatima, "Hierarchy Concept Analysis in Accounting Ontology," Asian J. of Computer Science and Technology, vol. 2, no. 2, pp. 13–20, 2012.

[19] W. S. A. Schwaiger et al., "The OntoREA©Accounting and Finance Model: Inclusion of Future Uncertainty," Proc. of IFIP Working Conf. on the Practice of Enterprise Modeling, Part of the Book Series: Lecture Notes in Business Information Processing, vol. 369, pp. 53–67, Nov. 2019.

[20] I. Blums and H. Weigand, "Towards a Core Ontology of Economic Exchanges for Multilateral Accounting Information Systems," Proc. of the 2020 IEEE 24th Int. Enterprise Distributed Object Computing Conf. (EDOC), pp. 227–232, Eindhoven, Netherlands, Oct. 2020.

[21] D. Porello, G. Guizzardi, T. P. Sales and G. Amaral, "A Core Ontology for Economic Exchanges," Proc. of the 39th Int. Conf. on Conceptual Modeling, pp. 364–374, DOI: 10.1007/978-3-030-62522-1_27, Vienna, Austria, Nov. 2020.

[22] W. E. McCarthy, "The REA Accounting Model: A Generalized Framework for Accounting Systems in a Shared Data Environment," Accounting Review, vol. 57, no. 3, pp. 554-578, 1982.

[23] H. Weigand, P. Johannesson and B. Andersson, "An Artifact Ontology for Design Science Research," Data Knowl. Eng., vol. 133, p. 101878, DOI: 10.1016/J.DATAK.2021.101878, May 2021.

[24] G. L. Geerts and W. E. McCarthy, "An Ontological Analysis of the Economic Primitives of the Extended-REA Enterprise Information Architecture," Int. J. of Accounting Information Systems, vol. 3, no. 1, pp. 1–16, DOI: 10.1016/S1467-0895(01)00020-3, Mar. 2002.

[25] W. Laurier, J. Kiehn and S. Polovina, "REA 2: A Unified Formalisation of the Resource-Event-Agent Ontology," Applied Ontology, vol. 13, no. 3, pp. 201–224, DOI: 10.3233/AO-180198, Jan. 2018.

[26] J. J. Wild, K. W. Shaw, B. Chiappetta, K. Dahawy and K. Samaha, Fundamental Accounting Principles, McGraw-Hill, 2007.

[27] Z. N. Shen and Y. Tijerino, "Ontology-based Automatic Receipt Accounting System," Proc. of the 2012 IEEE/WIC/ACM Int. Conf. on Web Intelligence and Intelligent Agent Technology (WI-IAT 2012), pp. 236–239, DOI: 10.1109/WI-IAT.2012.265, Macau, China, 2012.

[28] A. Abdelghany, N. Darwish and H. Hefni, "An Agile Methodology for Ontology Development," Int. J. of Intelligent Engineering and Systems, vol. 12, no. 2, pp. 170–181, Apr. 2019.

[29] A. Sattar, E. Salwana, M. Nazir, M. Ahmad and A. Kamil, "Comparative Analysis of Methodologies for Domain Ontology Development: A Systematic Review," Int. J. of Advanced Computer Science and Applications, vol. 11, no. 5, DOI: 10.14569/IJACSA.2020.0110515, 2020.

[30] S. Kruskopf, C. Lobbas, H. Meinander, K. Söderling, M. Martikainen and O. Lehner, "Digital Accounting and the Human Factor: Theory and Practice," ACRN J. of Finance and Risk Perspectives, vol. 9, pp. 78–89, DOI: 10.35944/jofrp.2020.9.1.006, 2020.

[31] A. Dyhdalewicz and U. Widelska, "Accouting and Marketing Dimensions of Innovations," e-Finanse, vol. 13, no. 2, pp. 1–13, DOI: 10.1515/fiqf-2016-0018, Dec. 2017.

[32] L. Chen, B. Xiu and Z. Ding, "Finding Misstatement Accounts in Financial Statements through Ontology Reasoning," IEEE Access, Early Access, DOI: 10.1109/ACCESS.2020.3014620, 2024.

[33] L. Shkulipa, "Evaluation of Accounting Journals by Coverage of Accounting Topics in 2018–2019," Scientometrics, vol. 126, no. 9, pp. 7251–7327, Sep. 2021.

[34] H. Khlif, "Hofstede's Cultural Dimensions in Accounting Research: A Review," Meditari Accountancy Research, vol. 24, no. 4, pp. 545–573, 2016.

[35] P. Mongeon and A. Paul-Hus, "The Journal Coverage of Bibliometric Databases : A Comparison of Scopus and Web of Science," Proc. of METRICS 2014 Workshop at ASIS & T, pp. 1751–1577, 2014.

[36] M. Odeh et al., "iOntoBioethics: A Framework for the Agile Development of Bioethics Ontologies in Pandemics, Applied to COVID-19," Frontiers in Medicine, vol. 8, DOI: 10.3389/fmed.2021.619978, 2021.

[37] D. M. Blei, "Probabilistic Topic Models," Communications of the ACM, vol. 55, no. 4, pp. 77–84, DOI: 10.1145/2133806.2133826, Apr. 2012.

[38] D. M. Blei, A. Y. Ng and M. I. Jordan, "Latent Dirichlet Allocation," J. of Machine Learning Research, vol. 3, no. Jan, pp. 993–1022, 2003.

[39] T. Kwartler, Text Mining in Practice with R, ISBN: 9781119282013, John Wiley & Sons, 2017.

[40] F. Moisescu, "Issues Concerning the Relationship between Accounting and Taxation in Determining Financial Result," European J. of Sustainable Development, vol. 7, no. 1, pp. 287–297, Feb. 2018.

[41] R. M. Bushman and A. J. Smith, "Financial Accounting Information and Corporate Governance," J. of Accounting and Economics, vol. 32, no. 1–3, pp. 237–333, Dec. 2001.

[42] M. K. Power, "Auditing and the Production of Legitimacy," Accounting, Organizations and Society, vol. 28, no. 4, pp. 379–394, DOI: 10.1016/S0361-3682(01)00047-2, May 2003.

[43] R. A. G. Monks et al., Corporate Valuation for Portfolio Investment: Analyzing Assets, Earnings, Cash Flow, Stock Price, Governance and Special Situations, ISBN: 1576603172, John Wiley & Sons, 2010.

[44] C. Quinn and J. J. McArthur, "A Case Study Comparing the Completeness and Expressiveness of Two Industry Recognized Ontologies," Advanced Eng. Informatics, vol. 47, p. 101233, Jan. 2021.

[45] E. Rohrer, P. Severi and R. Motz, "Applying Meta-modeling to an Accounting Application," ceur-ws.org, vol. 2228, pp. 92–103, 2018, Accessed: Nov. 21, 2021.

[46] M. von Rosing and W. Laurier, "An Introduction to the Business Ontology," Chapter in Book: Sustainable Business: Concepts, Methodologies, Tools and Applications, pp. 1–24, IGI Global, 2020.

[47] C. M. Drury, Management and Cost Accounting, ISBN: 978-1-4737-4887-3, Springer, 2013.

[48] M. M. Mowen and Don R. Hansen, Managerial Accounting, ISBN: 1337116009, South-Western, 2007.

[49] K. Mircheska et al., "The Importance of Forensic Audit and Differences in Relation to Financial Audit," Int. J. of Sciences: Basic and Applied Research, vol. 54, no. 2, pp. 190–200, 2020.

[50] B. Nikkel, "Fintech Forensics: Criminal Investigation and Digital Evidence in Financial Technologies," Forensic Science Int.: Digital Investigation, vol. 33, p. 200908, Jun. 2020.

[51] K. Farmer, "A Qualitative Study: Forensics Coaches' Perceptions of Administrators' Leadership Styles and the Impact within Their Professional Learning Communities," Electronic Theses & Dissertations Center, [Online], Available: https://digitalcommons.acu.edu/etd/234, Apr. 2020.

[52] W. D. Huber, "Is Forensic Accounting in the United States Becoming a Profession?," J. of Forensic Investigation, vol. 4, no. 1, DOI: 10.1002/9781119488552, Apr. 2012.

ملخص البحث:

تقدّم هذه الورقة طريقةً لبناء نموذجٍ وجودي مبنيٍ على البحث لدعم التّعاون والإبداع. ويتضمّن مفهوم الإبداع التّعاوني عمليةً تسمح لأصحاب المصلحة المتعدّدين بالعمل معاً لإنتاج أفكار وحلول ومنتجات مبتكرة.

يدْمج النّموذج المقترح بين الذّكاء الاصطناعي ومعرفة الخبراء لبناء نظامٍ شاملٍ يضمّ جوانب متعدّدة من البحث والتطوير والإبداع. ولبيان جدوى الطّريقة المقترحة، تُبين هذه الورقة تطبيقها على مجال المحاسبة. فأوّلاً، يتمّ استخدام خوارزميات تعلُّم الآلة وتقنيات التّنقيب عن النُّصوص لاستخلاص العناصر الأساسية من مجموعةٍ ضخمةٍ من الأبخاث والدّراسات في علم المحاسبة. وبعد ذلك، تجري الاستفادة من معرفة الخبراء في حقل المحاسبة للتحقُّق من تلك العع صر وتنقيحها. ويُمكن استخدام النّموذج النّاتج كأساسٍ لنظامٍ قائمٍ على المعرفة لتعزيز التّعاون وتحليل حالة الإبداع