Ontology-Based System for Educational Program Counseling

Mamoona Majid¹, Muhammad Faisal Hayat², Farrukh Zeeshan Khan³, Muneer Ahmad⁴,*, NZ Jhanjhi⁵, Mohammad Arif Sobhan Bhuiyan⁶, Mehedi Masud⁷ and Mohammed A. AlZain⁸

¹University of Management and Technology, Faculty of Informatics and Technology (IT), Lahore, Pakistan
²University of Engineering and Technology Lahore, Faculty of Computer Engineering Department, Lahore, Pakistan
³University of Engineering and Technology Taxila, Faculty of Computer Science Department, Taxila, Pakistan
⁴Department of Information Systems, Faculty of Computer Science and Information Technology, Universiti Malaya, 50603 Kuala Lumpur, Malaysia
⁵School of Computer Science and Engineering, SCE, Taylor’s University, Malaysia
⁶Electrical and Electronics Engineering, Xiamen University Malaysia, Bandar Sunsuria, 43900 Sepang, Selangor, Malaysia, 21944, Saudi Arabia
⁷Department of Computer Science, College of Computers and Information Technology, Taif University, P. O. Box 11099, Taif 21944, Saudi Arabia
⁸Department of Information Technology, College of Computers and Information Technology, Taif University, P. O. Box 11099, Taif 21944, Saudi Arabia

*Corresponding Author: Muneer Ahmad. Email: mmalik@um.edu.my
Received: 14 February 2021; Accepted: 02 May 2021

Abstract: Choosing the right university program can be very challenging for students. This is especially the case in developing countries such as India and Pakistan, where university admission depends on not only the program of interest but also other factors such as the candidate’s financial standing. Since information on the Internet can be highly scattered, university candidates often need counseling from qualified people to decide their educational programs. Traditional database systems cannot effectively organize the large unstructured data related to university programs. It is challenging, then, for prospective students to acquire the information needed to make good decisions to consider factors such as personal preferences, available options, and the job market. This study proposes an integrated framework that takes advantage of the latest ontology-based semantic technologies. The proposed system can efficiently extract user-specific constraints in unambiguous queries and then retrieve more precise information. The system uses Web sources to ingest the right information and then extract explicit and implicit knowledge about educational domains and related university offerings. We believe the proposed system can effectively and efficiently choose an educational program by incorporating user constraints.

Keywords: semantics; ontology; knowledge-based system; educational counseling

1 Introduction

Choosing the right educational path is critically important for students. This issue is especially salient in developing countries where there is often a lack of sufficient education-counseling mechanisms [1].
Pakistan, for example, where the literacy rate is 60%, prospective students can face significant barriers when choosing a program of study [2]. Given the lack of sufficient information and counseling, students may be unaware of the options available to them based on their credentials. They might even miss admission deadlines as a result. Hampered by insufficient and incomplete knowledge, prospective students can make incorrect decisions that may lead to academic failure.

The methods of knowledge retrieval, production, and transmission are constantly changing due to the rapid growth of information and communication technologies. Database models enable relevant information to be efficiently stored and queried. As educational processes have changed and evolved, many new methods have been proposed to manage the relevant information [3,4], such as taxonomies, categories, vocabularies, thesauri, semantic network ontologies.

Ontologies create and classify domain knowledge and interact with each other while also reusing existing knowledge. From an application perspective, ontologies can be applied as knowledge-based tools to build, map, merge, and evaluate knowledge systems [5]. Ontologies have also appeared as an alternative to databases in applications that require a more “enriched” meaning [6]. It can be challenging to extract precise information without associating it with the appropriate meanings. Ontologies describe relationships between keyword objects and intelligently answer user queries. The semantic relationships between objects help to organize large and complex data and answer user queries efficiently. In databases, relationships between different concepts are made through foreign keys. In ontology, however, different concepts behave like nodes.

This study proposes an ontological model to support or enhance students’ ability to decide their educational paths. Many educational counseling websites have been developed to help students decide on their educational programs. Few, however, use real-time or email interactions with counselors to provide guidance [7]. Most of them use traditional database systems [8] and online documents to manage and store available information. Thus, users must sort through various tables and undertake manual searches to find answers to their questions. Some systems have been proposed [9–11] that use semantic tools to organize career-path information. However, the present study developed a system that uses semantic technologies to understand career counseling using various constraints (e.g., geographical area, economics, job-market trends) as preliminary information. The system uses a knowledge-management system for educational counseling to infer new knowledge from existing information. In addition to the semantic model, implicit knowledge is inferred using a semantic reasoner. Further, an inference engine was built to accommodate new rules as ontology becomes populated over time.

The integrated framework proposed in this study takes advantage of the latest ontology-based semantic technologies. We present domain ontology, which can be used to design/create new ontology models in many areas, such as course ontologies, school-system ontologies, and demographic-based educational-system ontologies. The proposed system can efficiently extract user-specific constraints in unambiguous queries and then retrieve more precise information. The system uses Web sources to ingest the right information and then extract explicit and implicit knowledge about educational domains and related university offerings.

The rest of this paper is organized as follows. Section 2 reviews the related work to highlight gaps in the research. Section 3 presents details about the proposed framework, while section 4 discusses the developed system. Challenges, future extensions, and limitations are covered in section 5. Section 6 discusses the system evaluation, while section 7 concludes.

2 Related Work

Various intelligent support systems have been proposed for educational planning [12]. Yang et al. [12] proposed an ontology-based personalized course recommendation (OPCR) framework that takes input data from a user profile and recommends suitable courses and jobs accordingly. That study used a hybrid filtering
approach that combines collaborative filtering and content-based filtering. Using the Protégé tool, three different ontologies (i.e., student, course, job) were built and then embedded into a single framework to create a recommender system.

Charbel et al. [13] proposed a semantic-based hybrid recommender system with improved machine-learning techniques that could identify students’ capabilities and recommend a major. After extracting information from the Internet, semantic ontology is used to manage the student’s profile information. That study used the open-source WEKA application for machine-learning algorithms to create profile models of graduate students.

Merlin et al. [14] developed a multilingual ontology system for the educational domain. Multilingual ontology receives input from users in different languages and extracts relevant information from a particular ontology. First, unstructured data are collected from closed and open corpuses, such as textbooks and the Internet. Then, natural language processing methods (e.g., morphological and syntactic analysis) are used to extract the structured data in the form of concepts and classes. Lastly, using the SMART algorithm, that study merged different language ontologies into a single ontology and exchanged concepts between different natural languages.

Conde et al. [4] developed a system called LiReWi for the educational domain that could extract relationships from electronic books to guide students in their learning processes. LiReWi consists of different knowledge-based systems called extractors (e.g., Wikipedia) used to extract the relationships among topics.

Gerard et al. [15] developed an ontology for electrochemistry to model the static and dynamic equations of chemical formulas. They developed a lightweight ontology to visualize the vocabularies in an organized form that is easily extendable, executable, and machine-readable. For this, the subject’s vocabulary is structured using XML and then transformed into RDF files using semantic Web language. That study implemented RDF/Owl code into the ontology hierarchy using Protégé.

Martinez et al. [6] developed an application that automates HR processes for recruitment using ontology. First, they collected data in video format from the HR departments of multiple companies and educational institutions. Then, they analyzed and extracted the information and monitored multiple features of video interviews using artificial intelligence (AI) techniques.

Kongsakun et al. [7] proposed a conceptual framework for an intelligent recommendation system (IRS). They proposed that this IRS could help students choose appropriate courses and provide job guidance. That study used different data-mining and AI techniques, such as clustering, association rules, and IRS development classification.

Monali et al. [16] developed an automated system based on analysis and suggestions from users to avoid physical interaction with counselors. They aimed to help students analyze their options and choose the right educational path by providing tests to gauge ability and interest for given fields.

Archana et al. [17] proposed an intelligent career-counseling bot that takes both voice and text as input, converts voice to text, and parses and processes the data using data-mining and AI techniques. The system is intended to help users choose a career path by answering users’ queries according to their interests.

Hokstad [18] proposed an ontology focused on educational planning and the classification of academic subjects at both the bachelor’s and master’s levels. That study developed relationships among two or more courses, course hierarchies, and the significance of those courses in an educational path. The British Broadcasting Corporation (BBC) [19] also created a course ontology that describes national curricula in the UK. That ontology used the geographical dimension and described levels of study, fields, and topics. Benelux [20], meanwhile, developed an ontology for advanced European education, focusing on fostering
open curricula and an agile knowledge base. It depicts various ideas and concepts, such as students, evaluations, professors, courses, and credits. Tab. 1 presents a comparison of different ontologies.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Field of study</td>
<td>Person</td>
<td>Course code</td>
</tr>
<tr>
<td>Topic</td>
<td>Professor</td>
<td>Course credits</td>
</tr>
<tr>
<td>Topic of study</td>
<td>Student</td>
<td>Course name</td>
</tr>
<tr>
<td>Program of study</td>
<td>Academic degree</td>
<td>Semester</td>
</tr>
<tr>
<td>–</td>
<td>Bachelor</td>
<td>–</td>
</tr>
<tr>
<td>–</td>
<td>Master</td>
<td>–</td>
</tr>
<tr>
<td>–</td>
<td>Study track</td>
<td>–</td>
</tr>
</tbody>
</table>

Chung et al. [21] proposed an ontology design for creating an adaptive learning path in an e-learning environment. That ontology was based on the combination of multiple ontologies, which further contained course sub ontologies (e.g., curriculum ontology, syllabus ontology, subject ontology).

Zeng et al. [22] proposed an ontology for university courses based on content, method, and process. Noting that the Internet lacks ontology content, the authors provided a baseline course ontology that could help institutions construct initial domain ontologies for the programs they offer. Meanwhile, Uma et al. [23] developed an automatic ontology-based on similarity measures and the reranking method. That study aimed to solve the problem of candidate selection for employers and suitable job selection for candidates. Natalya et al. [24] constructed a semiautomatic ontology alignment and merging technique using the PROMPT algorithm. That algorithm takes two ontologies and merges them into a single one based on linguistic similarity. Essaid et al. [25] proposed a multi expert system based on a multiagent model and Web semantics for educational guidance. Meanwhile, Helna et al. [26], describing the evolution of ontology, showed that ontology is a shared conceptualization of domain knowledge; thus, it needs to be evolved whenever a concept changes. Luis et al. [27] proposed an ontological structure to guide high school students to choose a path appropriate for their future careers. Information was collected from student profiles, and students were asked about their personal interests, preferences, and academic performance.

Shaik et al. [28] proposed architecture to convert natural language into a SPARQL query so that information could be fetched in a semantic manner that machines would easily understand.

To bridge gaps in the research, we extended [27,28] to process NLP queries into SPARQL queries and developed a framework that uses an ontological structure as the basic model. The proposed system can help students select an appropriate course of university study. The proposed decision-support system is evolutionarily in line with [27] and [29] in that it can grow automatically after the basic ontology model construction.

Tab. 2 presents a comparison of state-of-the-art systems discussed in this paper. While most are ontology based, none cover all the required features of a support system. Specifically, we have not found a comprehensive system that uses semantic technology to understand counseling for higher education using the various required constraints (e.g., geographical area, student merit, public/private schools) as preliminary information. Thus, this study developed a comprehensive intelligent decision-support framework for higher education counseling using an ontology-based structure.
Table 2: Overview of the related work

<table>
<thead>
<tr>
<th>Paper (author)</th>
<th>NLP Query Support</th>
<th>Reading Web Content</th>
<th>Ontology/Database/ML based</th>
<th>Inference &amp; Mining</th>
<th>Template Engine for Visualization</th>
<th>RESTful APIs Endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natalya et al. [24]</td>
<td>No</td>
<td>No</td>
<td>Ontology</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Zeng et al. [22]</td>
<td>No</td>
<td>No</td>
<td>Ontology</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Chung et al. [21]</td>
<td>No</td>
<td>No</td>
<td>Ontology</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Benelux [19,20]</td>
<td>No</td>
<td>No</td>
<td>Ontology</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Essaid et al. [25]</td>
<td>No</td>
<td>Yes</td>
<td>Ontology</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Uma et al. [23]</td>
<td>No</td>
<td>Yes</td>
<td>Ontology</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Boyce et al. [10]</td>
<td>No</td>
<td>No</td>
<td>Ontology</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Helna et al. [26]</td>
<td>No</td>
<td>Yes</td>
<td>Ontology</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Kongsakun et al. [7]</td>
<td>No</td>
<td>No</td>
<td>Ontology &amp; ML</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Monali et al. [16]</td>
<td>No</td>
<td>No</td>
<td>Database models</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Archana et al. [17]</td>
<td>No</td>
<td>No</td>
<td>ML</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Yan Yan Yung et al. [12]</td>
<td>No</td>
<td>Yes</td>
<td>Ontology</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Charbel Obeid et al. [13]</td>
<td>Yes</td>
<td>No</td>
<td>Ontology &amp; ML</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Merlin et al. [14]</td>
<td>Yes</td>
<td>No</td>
<td>Ontology</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Angel Conde et al. [7]</td>
<td>Yes</td>
<td>No</td>
<td>Ontology</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Gerard et al. [15]</td>
<td>No</td>
<td>No</td>
<td>Ontology</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Developed System</td>
<td>Yes</td>
<td>Yes</td>
<td>Ontology</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
3 Method

Ontology is defined as a body of knowledge in MeSH tree form describing a knowledge domain intelligible to humans [30]. The generated taxonomic tree consists of all evolved terms, which a user may select as preferences. The tree evolves as more data are ingested, based on a concept called a triplet. This study designed an educational-domain ontology using general concepts as well as the relationships among concepts. Semantically relevant data were captured by a single ontology representing all categories of concepts, as shown in Fig. 1. In the proposed system, context is like an operation, where statements from indexed Web sources are parsed and ingested into existing knowledge. We used a single ontology with five concept categories for a complete representation of educational domains. Five concepts were chosen because educational path selection is not just a matter of short-term personal preference but an activity oriented toward fulfilling lifelong goals. Educational path selection should align with the functions the educational system is based upon. A prospective student wants to develop lifelong learning, and career choice is the central concept in this regard. Once this is decided upon, an appropriate university needs to be selected based on several factors, such as region, cost, and the likelihood of acceptance. In this regard, program selection is also important insofar as different universities may offer the same knowledge domains using different nomenclature.

![Figure 1: Education ontology modelling](image)

Figure 1: Education ontology modelling
The concept categories are represented as metadata in the form of a class structure, where each class provides the abstract structure of a concept, such as knowledge area, program, province, city, university type, and merit (Fig. 1). The instances are the nodes, and they are stored in the class. The concept categories are related according to special relationships or rules, and new data are ingested into the ontology using first-order logic. The stored data are known as individuals in the ontology. This semantic system provides the dynamic structure; thus, new information can be added or removed from any instance at any stage.

We aimed to populate the ontology in an evolutionary way with user-friendly retrieval supporting natural-language queries. For this purpose, we propose an integrated framework that takes advantage of recent semantic technologies such as Protégé, Graph Database (GraphDB), and Eclipse IDE (integrated development environment) (Fig. 5). Protégé is an open-source framework used to build ontologies in OWL/RDF languages. GraphDB is a database that stores data items, their properties, and the relationships between these properties in a graphical structure and efficiently handles semantic queries. GraphDB has been used for ontology management, retrieval optimization, and connectivity endpoints. Based on these technologies, the developed system consists of four modules: ontology building, storage, mapping, and retrieval. These are described in detail below.

3.1 Ontology Building Module

In this module, the basic ontological domain model is developed using Protégé. The domain model is then imported as an OWL file into GraphDB. Fig. 2 shows the domain ontology developed in Protégé. The ontology metadata are also shown in the form of classes. Classes represent the nodes/concepts of a domain, such as university or country. Object properties describe the relationship between nodes and individuals in the classes. Some examples are “belongsTo,” “hasDept,” “locatedIn,” and “offeredBy.” Data properties describe the relationships between classes with literal values (e.g., “hasMerit,” “hasName”).

![Figure 2: Ontology structure](image)

3.2 Storage Module

Unstructured data extracted from websites and documents are preprocessed in this module. The obtained structured data are then inserted into the proposed ontology system using Eclipse RDF4J libraries.
The RDF4J libraries are Web APIs that take preprocessed data as inputs and convert them into triplet form by describing the semantics of the data. Semantics describe relationships in data that link nodes directly so they can be queried using an operation.

3.3 Mapping Module

In this module, a classifier consisting of an inference engine is used to infer new knowledge from inputted structured data. The ontology is populated by new incoming knowledge and inferred knowledge, which is a piece of implicit information drawn according to the rules of semantic Web rule language (SWRL). The inferred knowledge is mapped into the ontology repository. The SWRL rules are written in Prolog as pie files. Prolog explains the relationships among explicit statements, which are then used to infer new statements and facts. When data are ingested into the ontology, inferred statements are automatically mapped according to the ontological model of the new relationships. Fig. 3 shows two inference examples.

![Figure 3: Inference example](image)

3.4 Retrieval Module

A search query text box is provided in the retrieval module that allows users to enter text queries to search for information. This module has two user interfaces:

- Natural language processing (NLP)-based semantic search
- Template-based search

The NLP interface allows a user to enter a query in natural language through the GUI, as shown in Fig. 4. In an NLP-based semantic search, the system takes user queries in English, identifies the type of query, and parses it into triplet form using lexical, syntactic, and semantic analysis. The system can also extract the entity and semantic relationships from unstructured data by mapping given data to the ontological model. After parsing and relationship extraction, a SPARQL query is generated with the help of existing templates. Then, the SPARQL query is sent to GraphDB through the RESTful interface, where the data are fetched from the semantic store. This returns the Java service results and displays the results in human-readable form on the Java website after processing. In a template-based search, templates are provided showing different predefined standard semantic queries.
A user can also build a query using four drop-down menus to create a general question format. Fig. 5 shows a detailed description of the data flow between the semantic technologies involved in the system. Eclipse and GraphDB are the semantic technologies used to develop the system. In this system, statements are parsed from the Internet through Eclipse parsers and ingested into the ontology through the server. The number of added statements is 5290. After applying the inference technique, the number of statements increases to 13,037, which means 7747 statements have been inferred and added to the rule engine’s ontology. Therefore, about 140% of the workload has been lifted from the data-entry operator with a 2.46 expansion rate.

4 Results and Discussion

In this work, we developed a tool to help students in the early stages of their academic career choose the right path based on queries, which are generally based on financial, public/private, merit, and geographical constraints. This system can help academic counselors guide students in their decision-making. In this system, students choose a possible educational program based on their queries. Queries can include questions such as the following: “What options are available in Pakistan after secondary school?” “Which
program is suitable with a given level of merit?" “Which universities offer subjects based on such merit in a specified area?” “What programs are offered by universities in the preferred geographic area?”

We tested the system with different possible options. We discuss the results for NLP-based queries, template-based questions, and drop-down selection-based searches. Initially, the system was populated with information available on indexed websites with almost 5000 statements ingested by the system. The system can also self-sufficiently ingest data itself using a crawler. The system crawls data from various popular sites whenever information is updated on those sites. After that, the unstructured data are preprocessed into structured form, and the information is mapped onto the ontology in the form of different nodes.

When a user visits the website to interact with the system, he or she is provided with three types of search options. The user can select already-available questions as templates and receive answers based on their interests. The system also provides implicit information by inferring it along with explicit information. A strong feature of this system is its automatic data population and inferencing. Suppose, for example, a student asks about universities offering engineering programs. The system will automatically infer implicit information about the program (e.g., knowledge area, merit, domain, university location) using metadata along with the search query according to the student’s interest. The second option in the user interface is to select options from the drop-down menu, which works the same as above. In the case of an NLP-based query, the user types a question such as, “Which universities offer engineering programs in Punjab with a merit of less than 70%?” The system will then show a list of all universities meeting the criteria. In this query, the implicit statement is “DeptProgram.” We queried against the merit of an engineering program; it automatically infers the DeptProgram associated with merit, university name, and university location. From the results, it is clear that with such queries and constraints in mind, a user can choose options available across the country based on his or her credentials. With the exponential increase in Web content, it can be difficult to obtain information that meets one’s needs and requirements. Even if relevant content is found, assessing its suitability for a specific purpose can be a highly complex process.

5 Challenges, Future Extensions, and Limitations

We faced many difficulties in developing an educational ontology based on the Pakistani educational system. Interoperability among various educational systems is the key challenge. The creation and automation of the conceptual modeling of ontologies posed difficulties as well. We aimed to adopt a realistic semantic conceptualization and ontology modeling of support systems to enable interoperability. Reusability is another challenge one faces in ontology building. The developed semantic-based ontological model should be extendable for the purpose of building more complex systems and models.

In terms of technical aspects, the education-based ontology deals with three to four triplets in a statement. The developed system can extract all triplets from a statement, map them, and provide relevant information covering all aspects (e.g., merit, program, institution) using an inference engine. The second problem we faced concerns the concept of merit. Programs offered at universities have “merit,” which can thus be considered a “concept.” However, we could not make it into a concept (node) because many universities can offer the same program with different merits. We resolved this issue by using SWRL rules and object properties in the ontology modeling.

Our developed ontology can be used as a base ontology to develop other ontologies such as course ontology, student and instructor ontology, or curriculum ontology. It can also be used for managing course learning objectives or job-hunting ontology.
6 Performance and Evaluation

The proposed system is a restricted domain system that is generally a question–answer (QA) system. Elizabeth et al. [31] proposed evaluation and validation requirements for such systems. These evaluation criteria are user centered and task based; they are identified for the users of an operational QA system. The evaluation metrics include satisfaction, correctness, relevance, and usefulness.

Two experiments were performed to evaluate the performance of the developed system. In the first, we used Apache Lucene to implement a keyword-based text search engine. Apache Lucene is open-source software used to develop high-performance, full-feature, keyword-based text search engines. We had a group of 60 users enter six queries each, resulting in 360 queries. Users gave a score of 1–5 for each evaluation metric (1 = completely dissatisfied; 5 = fully satisfied). All evaluation metrics were calculated; Fig. 6 shows the results. Users gave a relatively low score of 50% for all evaluation criteria for the Apache Lucene system. The reason is that the queries were very specific, and users had to search through many retrieved documents to receive relevant answers.

In the second experiment, the developed system was evaluated by the same 60 users entering 360 queries. The selected metrics were evaluated using scores given by users regarding the responses to their queries. Fig. 7 shows that our system performed well and provided good results (80%) in terms of all criteria (i.e., user satisfaction, correctness, relevance, and usefulness [16]). We also compared the two systems (Apache Lucene and the developed system) in terms of the features shown in Tab. 3, such as extensibility, interoperability, inference engine, and exception handling.

![Figure 6: Evaluation metrics graph of apache lucene system and developed system](image)

In the second experiment, the developed system was evaluated by the same 60 users entering 360 queries. The selected metrics were evaluated using scores given by users regarding the responses to their queries. Fig. 7 shows that our system performed well and provided good results (80%) in terms of all criteria (i.e., user satisfaction, correctness, relevance, and usefulness [16]). We also compared the two systems (Apache Lucene and the developed system) in terms of the features shown in Tab. 3, such as extensibility, interoperability, inference engine, and exception handling.

![Figure 7: Validation of apache lucene and developed system](image)
Table 3: Features of apache lucene and developed system

<table>
<thead>
<tr>
<th></th>
<th>Apache Lucene</th>
<th>Developed Ontology based system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extensibility</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Interoperability</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Inference Engine</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Exception Handling</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Consistency Check</td>
<td>No (repetition of same data)</td>
<td>Yes</td>
</tr>
<tr>
<td>Search Nature</td>
<td>index and Keyword based</td>
<td>Semantic-Based</td>
</tr>
<tr>
<td>Visualization</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

7 Conclusion

This study proposed an integrated framework that takes advantage of the latest semantic-based technologies. The developed system efficiently captured semantically relevant data with a single proposed ontology representing all concept categories. It provides a dynamic structure where new information can be added or removed from any instance at any stage. The system can extract user-specific constraints in the form of unambiguous queries and then retrieve more precise information. The developed system worked very well (80%) compared to a common system (50%) as a result of the semantic approach to relationship extraction and the inferencing approach to statements. The proposed system extracts entity and semantic relationships from unstructured data by mapping data to the developed ontological model. It also uses queries to extract explicit and implicit knowledge about domains. Moreover, the proposed system is flexible in terms of providing NLP-based and template-based search options to users. The developed ontological model is concise and general with regard to education in developing countries. The developed ontology for higher education systems can also be used as a basis for the further development of other ontologies. It can, moreover, be stretched with more spatial and temporal information about the education domain to increase the system’s reasoning capabilities. The proposed system is suitable for different formats of educational queries, such as merit, city, and program. Lastly, this work can also be extended to model knowledge-management systems for other educational domains, such as medicine and media.

Funding Statement: This research received support from the Taif University Researchers Supporting Project no. TURSP-2020/98, Taif University, Taif, Saudi Arabia.

Conflicts of Interest: The authors have no conflicts of interests to declare regarding the publication of this paper.

References


