OMAK System: Ontology-based Mathematical Knowledge Management System

Alan Ramírez-Noriega [a], Yobani Martínez-Ramírez [a], Manuel Quintero-Fonseca[a], Samantha Jiménez[b], José Mendivil-Torres [a], José Emilio Sanchez García [c].

[a] Universidad Autónoma de Sinaloa, Facultad de Ingeniería Mochis, {alandramireznoriega,yobani,manuel,jose.mendivil}@uas.edu.mx, Los Mochis, Sinaloa, México.
[b] Universidad Autónoma de Baja California, Facultad de Ciencias Químicas e Ingeniería, samantha.jimenez@uabc.edu.mx, Tijuana, Baja California, México.
[c] Universidad Autónoma Intercultural de Sinaloa, esanchez@uais.edu.mx, Los Mochis, Sinaloa, México.

Abstract

Knowledge management systems based on ontologies are important software tools to maintain the knowledge of experts. However, the mathematical field needs to improve aspects such as creating repositories of formalized mathematics, mathematical search and retrieval, as well as implementing math assistants for education. This research proposes a Mathematical Knowledge Management system based on Ontologies (OMAK system) to provide support in solving mathematical problems. In essence, the OMAK system maintains a knowledge base of problems, organized with the support of an ontology. The case study in this research considered Linear Equations Systems (LES). Users search LES problems solved in the OMAK system, to serve as a guide to solve other similar equations systems. The OMAK system was used with a student group to determine the quality of searches and the efficiency of results. Results show that the OMAK system provides support in the resolution of mathematical problems to 92.5% of students, preserving the teaching experience in a knowledge repository that can be updated through the system.

Keywords— Information Retrieval, Linear Equation, Mathematical Knowledge Management System, Math Assistant, Ontology.

1. INTRODUCTION

Knowledge management takes an important function in the modern society [1]. A Knowledge Management System (KMS) makes the use of efficient knowledge while providing the access of information to users when required [2]. The general goal of a KMS is to provide the right knowledge to the right people, at the right time, and in the right format [3]. In addition, a KMS supports the process of creation, storage, recovery, transfer, and application of knowledge [2], [4]. Research of these systems focuses on the development of concepts, methods, and tools supporting the management of human knowledge [4].

In education, a KMS is employed as activity of strategic management supporting educators in the gathering of information or making use of the knowledge resource to maximize their teachings and tasks. This practice can help capture, codify, and distribute knowledge through the application of information, or human interaction [5].

A KMS retains the expertise of experienced teachers, strengthens the novice teacher's knowledge through knowledge transfer in administrative work and teaching, increases their effectiveness in terms of teaching and learning performance, supports the development of a knowledge community in schools and fosters a culture of learning [6]. Mathematics as well as other fields, need abstract models to represent their core concepts and relationships [7].

Therefore, we can represent and recover mathematical knowledge to solve problems efficiently by emerging the Mathematical Knowledge Management (MKM) [8]. MKM is an interdisciplinary field of research in the intersection of mathematics, computer science, library science, and scientific publishing. The purpose of MKM is to develop a better approach for mathematical knowledge management using sophisticated software tools. MKM is expected to serve mathematicians, scientists, and engineers who produce and use mathematical knowledge. Educators and students who teach and learn mathematics, publishers who offer textbooks and disseminate new mathematical results, librarians and math experts who catalog and organize mathematical knowledge will also find a great use of MKM [8], [9]. To systematize the previous work arise the MKM System (MKMS).

The development of repositories that contain mathematical knowledge is necessary for the search and retrieval of information to support students the learning needs. Through these repositories, knowledge is stored and maintained, preventing it from being lost when a teacher is not present. Since the ontology domain is powerful in knowledge representation and associated inference, the emerging ontology-based KMSs can find the content-oriented knowledge that users want. Ontologies give an understanding of the static knowledge domain that facilitates knowledge retrieval, storage, sharing, and dissemination. For KMSs, ontology can be regarded as the classification of knowledge; that means, ontology defines shared vocabulary by facilitating knowledge communication, storage, searches and sharing in KMSs [3], [10].

Certain pieces of the literature have addressed KMSs, as presented by [11]–[13]. These authors present proposals for structures and implementations that serve as an example for the creation of systems based on knowledge. On the other hand, the article [1] proposes the creation of a knowledge system based on ontologies, and makes use of natural language processing although it is not focused on the mathematical area. Moreover, [10] presents a system of mathematical knowledge based on ontologies with very similar principles as our proposal, but with a focus away from education, an aspect that is a fundamental element in our research.

This research proposes an ontology-based MKMS (OMAK System) to provide support in solving Linear Equations Systems (LES) with two variables. In general, the OMAK system maintains a knowledge base of mathematical problems organized through an ontology. Users search LES problems solved in the OMAK system, to serve as a guide to solve other similar equations systems. The OMAK system

was used with a student group to determine the quality of searches and the efficiency of results.
The remainder of the paper is organized as follows: Section 2, which refers to related work of this study. Section 3 explains related concepts to the article. Section 4 describes the OMAK system as proposed by the authors. Section 5 presents the experimental methodology. Finally, section 6 describes a conclusion and suggests future work.

2. RELATED WORK

This research is based on the work presented in [14]. The article mentioned proposes an ontology to be used in an MKMS with the objective of storing and retrieving LES, these equations are used to teach students to solve problems with an approach based on cases. Also, the article superficially describes the architecture of an ontology-based MKMS. Considering this work, we build and test the OMAK system presented in the research.
There are several pieces in the literature that have addressed the KMS. Liao et al. [12] proposed an architecture of an MKS that takes advantage of the infrastructure of a cloud computing to support the knowledge management services. The proposal is limited to a model of exploration of the architecture of an MKS without the implementation, nor corresponding evaluation.
Chua [11] developed a KMS model which analyzes different technologies that can support knowledge management processes, and proposes different technologies that contribute to knowledge management. However, the author does not fully implement the proposed architecture; therefore, its evaluation is not fulfilled.
Fu et al. [13] proposed a KMS architecture that uses software agents to support knowledge management processes; in addition, it uses ontologies for the exchange of information between agents.
These works include aspects related to structure and how to deal with knowledge management, which has proven useful for the basis of our proposal; however, they do not address the mathematical aspect.
Jeschke et al. [1] extracted linguistic relationships from texts, created ontologies, and combined them to form a knowledge base to build a retrieval system based on natural language processing techniques. Their proposal is a comprehensive MKMS to accumulate, store, merge, evaluate and visualize mathematical information for applications such as: encyclopedias, context sensitive library search systems, intelligent book indexes, and e-learning software using standardized methods. Unlike this proposal, our intention is not to analyze text and build ontologies. Our research focuses on finding similar mathematical problems in the knowledge database to support the student in solving their equation. However, the research by Jeschke et al. presents interesting aspects that can be applied in a future work of our research.
Elizarov et al. [10] presented the architecture of the OntoMath digital ecosystem. The system aims to retrieve mathematical scientific articles based on ontologies. The implementation of the system includes a semantic search service for mathematical formulas and a recommendation system based on the user's profile and work scenario. Although the work is good, we focused our research on education. The work of Elizarov and ours have the same ideas with a different vision. We employ an MKMS based on ontologies but focused on solving mathematical problems based on similar cases.

3. OMAK SYSTEM

This section describes the OMAK system; the architecture, the ontology used, technologies employed to implement the proposal, and system interfaces.

3.1. OMAK System Architecture

The OMAK system maintains a database with problems related to LES. The main concept is that the student defines his own LES based on the problem statement. The student will access the system to find similar problems and solve their own. Figure 1 represents the OMAK system architecture. Students interact with the system through the interface (View). The system break down the user's equation to search based on knowledge of ontology, this is supported by the JENA API[1] and SPARQL (Model). The system then shows solutions similar to the student's problem.

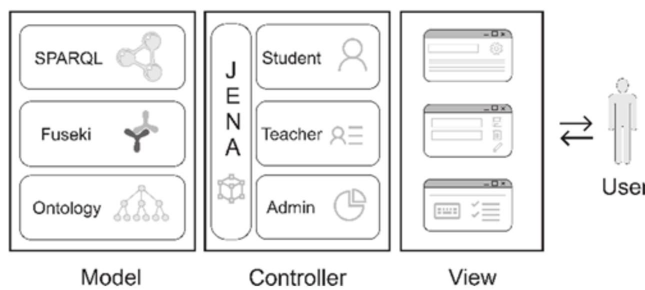

Figure 1. The OMAK System architecture.

The responsibility of the teacher is to feed the instances of the ontology and provide documentation and necessary support; this way, the system will be able to give valid answers to users. This research focuses on recovering the information, however, future versions consider giving support in solving the user's problem. In addition, teacher provide support for any other subject of mathematics through the reuse or design of new ontologies.
The OMAK system is based in the Model-View-Controller (MVC) architectural pattern that divides the system into three interconnected parts (see Figure 1). Each part of the architecture is described below.

View: This part generates the interfaces of the system, depending on the user (student or teacher) is the graphic environment displayed.
Controller: This section handles four main submodules:

- **Student Module:** This module processes the students' requests, mainly the consultations to individuals in the ontology.

[1] https://jena.apache.org/

- **Teacher Module:** This module adds the example cases to feed the ontology to generate the knowledge base of the system.
- **Admin Module:** This model is responsible for the management of students and teachers, it is a secondary process to the system.
- **JENA:** It is a free and open source Java framework for building Semantic Web and Linked Data applications. The OMAK system uses this framework to send data between the model and the controller.

Model: This handles the information that contains the ontology, through the Fuseki server and SPARQL queries.

3.2. OMAK System Ontology

An ontology development method is comprised of a set of established principles, processes, practices, methods, and activities used to design, construct, test, and deploy ontologies [15]. Methontology is one of most comprehensive methodology to build ontologies [15] and includes a set of activities, techniques, and deliverables [16]. Different investigations [17]–[20] applied Methontology to their proposals. This study developed the OMAK system ontology through Methontology for educational purposes based on [14].

To build the hierarchy of classes, we defined the most general concepts of a LES (equations and variables); then, using a top-down process, a combination of top-down/bottom-up, we improved the ontology.

A mathematical problem of a real situation constructs a LES, where equations form the LES, in this case, two equations and two variables. We define the problem by the problem statement, complexity, and solution; where a solution is represented by a method to solve the LES. An equation has a result represented by an equalization. It has a variable with a coefficient and a value as an attribute. The value is the number that represents the variable. Based on the previous definition, the most important terms of LES are: Linear equation, Equation, Variable, and Problem.

The designed ontology is shown in Figure 2. Mainly, the ontology has classes to represent equations, variables, and the problem. Because two variables are handled, *x* and *y* are defined as subclasses of *variable*. The same thing happens with *equation1* and *equation2* which are subclasses of *equation*. The object properties represent the relation mereology, which denotes how a set of objects are combined to form a bigger one [21].

Attributes allow the storage and retrieval information, with each attribute having its data type (data type column). Hence, main classes contain attributes to store information of the problem equation. The class *variable* has attributes to store the solution, the numerator, and the denominator. The class *equation* stores the result. The class *problem* stores the problem statement, the document path, the complexity, and the identifier.

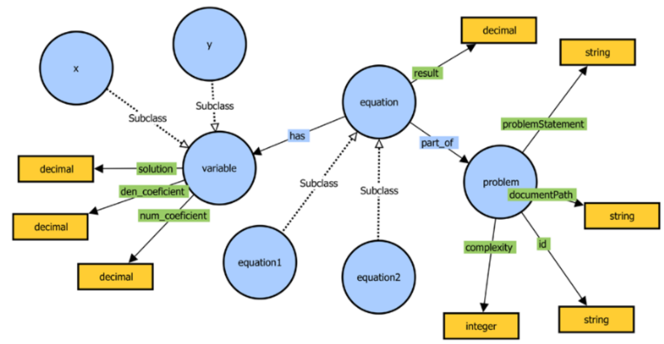

Figure 2. The OMAK System Ontology.

We developed the ontology using the Protégé[2] software, a free open-source ontology editor and framework for building intelligent systems. Protégé is supported by an active community of academic, government, and corporate users. The implemented ontology is shown in Figure 2 through VOWL [22]. This software simulates the ontology before implementing into a system, making it easy to change the ontology structure by adding, eliminating, or modifying classes. Also, we can work with instances and execute SPARQL queries, to analyze the ontological reasoning.

3.3. Implementation technologies

The OMAK system uses the JSP (JavaServer Pages) programming language[3], the HTML language (HyperText Markup Language), and the Apache Tomcat[4] to develop and implement the system on the web platform. The OMAK system uses Apache Jena Fuseki to maintain the knowledge repository and make SPARQL queries.

3.4. OMAK System web app

This subsection shows some features of the web app[5] implementing the OMAK system (see Figure 3). The OMAK system has a section to show cases of registered LES. At the moment, there are only cases in the area of mathematics in linear algebra, however, future work will provide support for other areas. The system allows to filter the cases in three levels of difficulty.

OMAK System problem description shows contents of an example case. The information in the document focuses on the description of the problem and solutions. This document supports the student to answer similar exercises. The search section retrieves problems based on the input equations, returning a series of cases similar to the search equation (see Figure 4). This last section represents the main functionality based on queries to the ontology.

[2] https://protege.stanford.edu/

[3] http://www.oracle.com/technetwork/java/index-jsp-138231.html

[4] http://tomcat.apache.org/

[5] http://siace.uas.edu.mx/Omak/

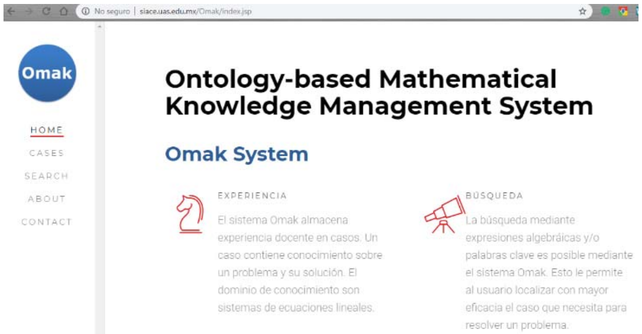

Figure 3. OMAK System web app.

Figure 4. OMAK System search result.

4. METHODOLOGY FOR EVALUATING THE OMAK SYSTEM

4.1. Participants

The authors used a convenience sampling (non-probabilistic) because of the availability and easy access of participating students [23]. The sample included thirty undergraduate students belonging to engineering programs at the Autonomous University of Sinaloa, enrolled in Linear Algebra (LA) courses. The participants had prior experience with the use of a computer and Internet. The average age of participants was 20.2 years. Among the participants were fourteen men and twelve women, while the remaining participants chose to omit this information.

4.2. Problems description

This research defined three mathematical problems printed on paper, each with different levels of complexity. For instance, a problem is: *1/9 of the age of A exceeds 2 years to 1/5 of the age of B, and twice the age of B is equal to the age of A 15 years ago*. Find the current ages.

The document given to students to solve the mathematical problem with the OMAK system (experimental group) contained questions about the interaction with the system, focusing on the quality of results returned by searches and the usefulness. These questions are analyzed in the results section. Mathematical problems were applied at the beginning of the semester.

The document delivered to the student to solve the mathematical problem without the OMAK system (control group) contained the following sections: variable identification, LES deduction, method development, and results verification. This document has the objective of separating phases of knowledge applied to solve the problem.

4.3. Procedure

The procedure followed to evaluate the OMAK system (experimental group) is described below:

1. We assign the problem statement to the student in a random way.
2. The student develops the LES for the given problem.
3. The student introduces the LES developed to the OMAK System to search similar problems.
4. The Omak system displays a set of similar LES.
5. The student chooses the most similar LES to solve the problem.
6. The Omak system gives the student a solved LES in a document.
7. The student solves the problem using the information on the recovered document.

After solving the problem with the document recommended by the OMAK System, the student answers the survey to determine the usefulness of the system.

4.4. Results

We distributed 30 surveys to participating students, however, 3 students did not answer. There were three different mathematical problems dispersed equally. Due to three blank submissions, the experiment obtained 27 exercises solved by the students. Cronbach's alpha revealed a survey evaluation of 0.88.

One of questions in the survey was whether or not students used the document suggested by OMAK system to solve the problem. Results indicated that 92.59% of the students used it, while the rest did not (see Figure 5).

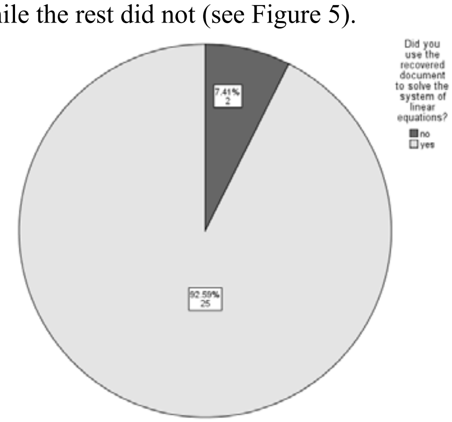

Figure 5. Did you use the document recovered by the OMAK system?

Figure 6 displays results obtained by the students in the resolution of the LES. There were 14 (51.8%) students who answered the problem correctly. The remainder of the students were evaluated according to the percentage of solution they offered; although they did not answer the problem correctly, they did a solution procedure, which in some cases was more accurate than others. The experiment obtained on average an evaluation of 77.77 on a scale of 0 to 100, with a standard deviation of 24.96.

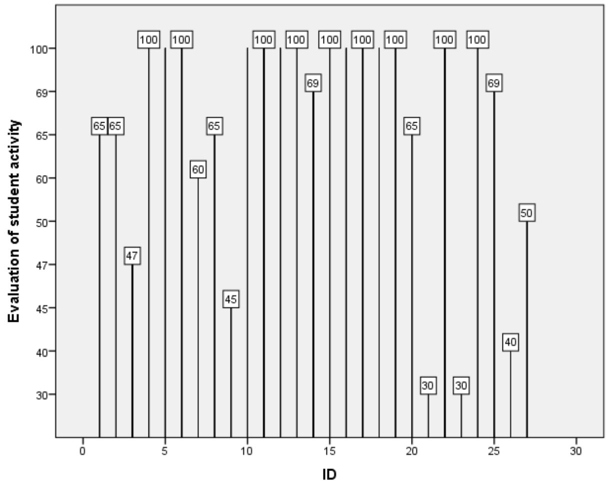

Figure 6. Evaluation of problems solved by students.

Table 1 represents the response frequencies for the remaining questions. There was a 74% of students who agreed that the first option offered by the OMAK system was relevant to solving their LES. In addition, a question from the survey asked if the recovered document by the OMAK system proved useful in solving their LES. This question was asked before students used the document, and once again after using the document, in which results showed similar responses.

Table 1. Some of questions applied to students.

Item	Strongly disagree (%)	Disagree (%)	Neutral (%)	Agree (%)	Strongly agree (%)
1.- The first option of the list of results is relevant	2 (7.4)	0 (0)	5 (18.5)	12 (44.4)	8 (29.6)
2.- The recovered document is useful	1 (3.7)	0 (0)	4 (14.8)	10 (37.0)	12 (44.4)
3.- The recovered document to solve the system of linear equations was useful	1 (3.7)	1 (3.7)	3 (11.1)	12 (44.4)	10 (37.0)

4.5 Discussion

Figure 7 shows a boxplot with the question: In which position of the list of results is the answer closest to your LES? Figure shows 2 outliers that alter the result of this question. The mean obtained considering all values is 3.48 with a standard deviation of 4.17; however, if we discard the outliers, we obtain an average of 2.01 with a standard deviation of 2.39. Results improve considerably without outliers. Therefore, it is important to improve the search algorithm to control outliers, improving the accuracy of searches.

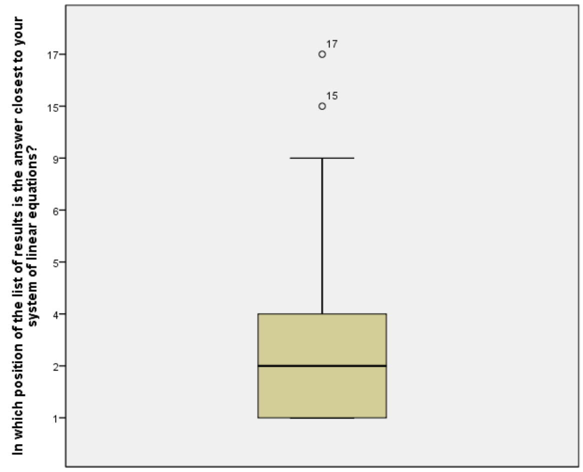

Figure 7. Boxplot of a question of the survey.

Based on item 2 and 3 of the Table 1, we can determine that students appreciate a document that provides support in solving an equation. Items 2 and 3 are the same, only applied before and after using the document returned by the OMAK system. Table 1 shows similar results; to avoid subjectivity, we apply the Spearman correlation test. This test shows whether two variables are related and to what extent they are. The Pearson test showed a correlation of 0.835 between both items, and a P-value below the significance level (0.001 < 0.05). Therefore, we can determine that student responses have a high degree of reliability.

The problems evaluation represented in Figure 8 shows an average score of 77.77, where 92.59% of the students were supported by the OMAK system documents. Students who used the OMAK system scored higher on their results than the students who did not have the support. The experimental group obtained an average score of 77.77 against 47.87 of the students in the control group. This implies that OMAK System support students in solving mathematical problems.

In this proposal, two important characteristics were used to make problem-solving easier. First, the use of domain ontologies for the representation of knowledge and associated inference in knowledge management systems facilitate the storage, search and transfer of knowledge. Second, problem-solving based on similar cases, a strategy that uses a previously documented experience or situation that can be referenced for future problems. Both characteristics have been combined, forming the OMAK system.

5. CONCLUSION AND FUTURE WORK

This research developed a knowledge system based on ontologies to solve LES. The system provides help to the student through queries on the knowledge base to generate similar problems for the student. The OMAK system showed acceptable results with the experiments executed in this article; when the OMAK system was used, the

group's scores increased significantly compared to when they did not use it.
The knowledge recovery provided by the OMAK system allows knowledge to be applied to the student's learning with a positive impact. In addition, OMAK system has the following advantages: 1) The knowledge generated by the teacher in the institution is conserved and can be used by students or novice teachers. 2) A knowledge management system in the mathematical area to facilitate the maintenance of knowledge. In this way, knowledge can be extended or eliminated.
The OMAK system shows support for students with LES problems. For now, the system is designed for a single topic; however, in future versions, the system will be generalized to adapt to any mathematical problem. This entails a new and better search algorithm in the ontology, and new strategies to automatically generate the code to register the information in the ontology.

REFERENCIAS

[1] S. Jeschke, N. Natho, and M. Wilke, "KEA - A Knowledge Management System for Mathematics," in *IEEE International Conference on Signal Processing and Communications*, 2007, pp. 1431–1434.

[2] M. Alavi and D. E. Leidner, "Knowledge management systems: emerging views and practices from the field," in *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences.*, 1999, pp. 1–11.

[3] Y. L. Y. Li, Y. S. Y. Sheng, L. L. L. Luan, and L. C. L. Chen, "A Text Classification Method with an Effective Feature Extraction Based on Category Analysis," *2009 Sixth Int. Conf. Fuzzy Syst. Knowl. Discov.*, vol. 1, pp. 95–99, 2009.

[4] I. Jurisica, J. Mylopoulos, and E. Yu, "Using Ontologies for Knowledge Management: An Information Systems Perspective," in *In Proceedings of the 62nd Annual Meeting of the American Society for Information Science*, 1999, pp. 482–496.

[5] E. C. K. Cheng, "Knowledge Management for School Education," in *Knowledge Management for School Education*, 1st ed., no. 7, Springer Singapore, 2015, pp. 11–23.

[6] C.-H. Leung, "Critical Factors of Implementing Knowledge Management in School Environment: A Qualitative Study in Hong Kong," *Res. J. Inf. Technol.*, vol. 2, no. 2, pp. 66–80, 2010.

[7] F. C. Coelho, R. R. Souza, and C. T. Codeço, "Towards an ontology for mathematical modeling with application to epidemiology," *Adv. Knowl. Organ.*, vol. 13, no. January, pp. 138–144, 2012.

[8] M. Kohlhase, "Mathematical Knowledge Management: Transcending the One-Brain-barrier with Theory Graphs," *EMS Newsl.*, pp. 1–6, 2014.

[9] W. Farmer, "MKM: a new interdisciplinary field of research," *SIGSAM Bull.*, vol. 38, pp. 47–52, 2004.

[10] A. Elizarov, A. Kirillovich, E. Lipachev, and O. Nevzorova, "OntoMath Digital Ecosystem: Ontologies, Mathematical Knowledge Analytics and Management," *CoRR*, pp. 1–18, 2017.

[11] A. Chua, "Knowledge management system architecture: a bridge between KM consultants and technologists," *Int. J. Inf. Manage.*, vol. 24, no. 1, pp. 87–98, 2004.

[12] C.-N. Liao, I.-L. Chih, and Y.-K. Fu, "Cloud computing: A conceptual framework for knowledge management system," *Hum. Syst. Manag.*, vol. 30, no. 3, pp. 137–143, 2011.

[13] R. Fu, X. Yue, M. Song, and Z. Xin, "An architecture of knowledge management system based on agent and ontology," *J. China Univ. Posts Telecommun.*, vol. 15, no. 4, pp. 126–130, 2008.

[14] A. Ramírez-Noriega *et al.*, "Towards a mathematical knowledge management system: Ontology to model linear equations," in *Advances in Intelligent Systems and Computing*, vol. 745, A. Rocha, H. Adeli, L. P. Reis, and S. Costanzo, Eds. Napoli, Italia: Springer, 2018, pp. 518–527.

[15] D. Gasevic, D. Djuric, and V. Devedzic, *Model Driven Engineering and Ontology Development*. Springer-Verlag, 2009.

[16] M. Fernandez-Lopez, A. Gomez-Perez, and N. Juristo, "METHONTOLOGY: from Ontological Art towards Ontological Engineering," in *Proceedings of the AAAI97 Spring Symposium*, 1997, pp. 33–40.

[17] S. John, "Development of an Educational Ontology for Java Programming (JLEO) with a Hybrid Methodology Derived from Conventional Software Engineering Process Models," *Int. J. Inf. Educ. Technol.*, vol. 4, no. 4, pp. 308–312, 2014.

[18] A. C. Muñoz Garcia, B. Sandia Saldivia, and G. Monzòn Pàez, "An ontological model of collaborative learning in interactive distance education," *Red Rev. Científicas América Lat. el Caribe, España y Port.*, vol. 18, no. 61, pp. 449–460, 2014.

[19] P. Pribyl, V. Fábera, and V. Faltus, "Domain-Oriented Ontology for ITS System," in *2012 ELEKTRO*, 2012, pp. 364–368.

[20] J. J. Tabares García and J. A. Jiménez Builes, "Ontology for the evaluation process in higher education," *Rev. Virtual Univ. Católica del Norte*, no. 42, pp. 68–79, 2014.

[21] M. Fernández-López, A. Gómez-Pérez, and M. C. Suárez-Figueroa, "Selecting and customizing a mereology ontology for its reuse in a pharmaceutical product ontology," *Front. Artif. Intell. Appl.*, vol. 183, no. 1, pp. 181–194, 2008.

[22] S. Lohmann, S. Negru, F. Haag, and T. Ertl, "Visualizing ontologies with VOWL," *Semant. Web*, vol. 7, no. 4, pp. 399–419, 2016.

[23] R. Hernández Sampieri, C. Fernández Collado, and M. Baptista Lucio, "Análisis de datos cuantitativos," in *Metodología de la Investigación*, 6ta ed., M. Hill, Ed. 2014, p. 632.