



An Ontology-Based Data Management Model Applied to a Real Information System

Raúl Comas Rodríguez¹ , Alfredo Simón-Cuevas² ,
Nayi Sánchez Fleitas³ , and María Matilde García Lorenzo⁴ 

¹ Research Directorate, Universidad Regional Autónoma de los Andes,
Vía a Baños km 5 ½, Ambato 180150, Tungurahua, Ecuador
raulcomasrodriguez@gmail.com

² Universidad Tecnológica de La Habana José Antonio Echeverría,
Ave. 114, No. 11901, 19390 La Habana, Cuba
asimon@ceis.cujae.edu.cu

³ Empresa de Tecnología de la Información y la Automática,
Ave. de los Mártires esq. Circunvalación, 60100 Sancti Spiritus, Cuba
nayi78@gmail.com

⁴ Universidad Central “Marta Abreu” de Las Villas,
Carretera de Camajuaní km 5,5/Santa Clara, 54830 Villa Clara, Cuba
mmgarcia@uclv.edu.cu

Abstract. In scenarios where capturing and efficiently managing data from heterogeneous data sources can be an advantage, the development and use of ontologies is increasingly common. In this paper, an ontology-based data management model applied to the Geographic Information System denominated SIGOBE as part of the Business Management System of the Cuban Electric Union (SIGE), is presented. The proposed data management model combines the use of a developed domain ontology (OntoSIGOBE), with an intelligent query answering process based on the case-based reasoning technic. OntoSIGOBE represents the conceptualization associated to the distribution and transmission processes of electrical energy, and the captured knowledge from the heterogeneous data sources of SIGE. OntoSIGOBE allows achieve the semantic interoperability between the data sources and the query answering process. The obtained results evidence that the proposed model provides a flexible and integrated data access in SIGOBE reaching high satisfaction to the end users.

Keywords: Ontology-based data management · Ontology engineering · Geographic information systems · Case-based reasoning

1 Introduction

The Electrical Union (UNE, Spanish acronym) in Cuba develops the Business Management System of the Electric Union (SIGE) for the automation of electrical processes [3]. SIGE integrates of two main subsystems: Integral System of Network Management (SIGERE) and Integral Management System of the Electric Industry Construction Enterprise (SIGECIE). The functions of SIGERE and SIGECIE are to collect technical,

economic and management data to convert them into key information for supporting the decision making.

The collected data facilitate and improve the efficiency in the analysis, planning, operation, and control of the distribution and transmission electricity networks. SIGERE and SIGECIE are considered complex systems because they have 36 modules and a database compose of: 716 tables, 1303 stored procedures and 74 functions. Both systems establish the databases of a Geographic Information System (GIS) of the SIGE, which is denominated SIGOBE. An average query in the SIGOBE involves approximately nine tables with different attributes and the queries on a specific topic requires knowledge of the database organization. Despite the number of stored queries in SIGOBE, they still do not cover the needs of the customer due to the operational dynamics of the national electro-energy system and several weaknesses identified, such as: absence of semantic correspondence between the databases of SIGE and the digital cartography of the electric system; low integration and heterogeneity of the key concepts for the electric system in the database, as well as between the geographical objects in cartography; and the develop of a static query for each problem which is very inefficient.

In order to address these challenges, the Ontology-Based Data Management (OBDM) paradigms [1, 7] were considered to define a new data management model approach for SIGOBE. OBDM has recently emerged as a general paradigm based on the assumption that a domain ontology capturing complex knowledge can be used for data management by linking it to data sources using declarative mappings [1, 12]. In these systems, the ontology is a key resource and constitutes a conceptual, formal description of the domain of interest to a given organization (for instance, the transmission and distribution processes for the Cuban Electric Union), expressed in terms of relevant concepts, attributes of concepts, relationships between concepts, and logical assertions characterizing the domain knowledge [2].

In this paper, an ontology-based data management model applied to SIGOBE is presented. This model combines the use of a manually developed domain ontology (denominated OntoSIGOBE) with an Intelligent Query Answering Process. OntoSIGOBE formalize the semantic meaning of the conceptualization associated to the processes of distribution and transmission of electrical energy, as well as represents the implicit captured knowledge from SIGERE and SIGECIE Databases, and SIGE Cartography, in order to achieve semantic interoperability between these data sources in the query answering process. The Intelligent Query Answering Process exploits conceptual schema in OntoSIGOBE as an intermediate layer for accessing and querying the data source, through the query preprocessing applying several basic natural language processing (NLP) tasks and a case-based reasoning process, for increasing the efficiency and flexibility of the querying process by end users. The proposed model and the constructed ontology were evaluated from different perspective.

OntoSIGOBE was evaluated considering a task-based approaches to measure how this ontology helps improving the querying process associated to customers' information needs, and using the OOPS! (OntOlogy Pitfall Scanner!) system¹ to diagnose

¹ <http://oops.linkeddata.es/>.

aspects such as: structural, functional, usability-Profiling, consistency, completeness, and conciseness [13]. Additionally, a software-quality evaluation process was carried out to evaluate the quality of the implemented system for applying the proposed model. The obtained results in the experimental and quality evaluations processes carried out evidence that the proposed model provides a flexible and integrated data access to the end users of SIGOBE, improving the performance of this system in the querying process and reaching high satisfaction of the end users.

The rest of the paper is organized as follows: Sect. 2 summarizes the theoretical and scientific foundation associated to the fundamental interest topics for the research carried out, Sect. 3 describes the proposed data management model for SIGOBE; Sect. 4 presents the experimental results and the corresponding analysis; and conclusions arrived and future works are given in Sect. 5.

2 Background

In the geographic data domain, ontologies are used to formalizing the semantic meaning of geospatial concepts and data, categories and spatial relations in a machine-understandable manner. Geospatial semantics can also facilitate the design of GIS by enhancing the interoperability of distributed systems and developing more intelligent interfaces for the user interactions [6]. Ontology-Based Data Management (OBDM) has recently emerged as a general paradigm.

The OBDM is based on the assumption that a domain ontology capturing complex knowledge can be used for data management by linking it to data sources using declarative mappings [1, 12]. The key idea of OBDM is to resort to a three-level architecture: the ontology, the sources, and the mapping between them, where the ontology is a formal description of the domain of interest, and is the heart of the whole system [7]. The ontology is a conceptual, formal description of the domain of interest to a given organization, expressed in terms of relevant concepts, attributes of concepts, relationships between concepts, and logical assertions characterizing the domain knowledge [2].

In OBDM, Ontology-Based Data Access and Integration (we refer to OBDI) is one of the distinctive dimensions, which deals with capturing implicit knowledge from heterogeneous data sources in order to achieve semantic interoperability between them at access and integration levels. Ontologies are the key resource in these systems, since they allow to management the implicit semantic meaning of the data source, at the same time that facilitate the retrieval, integration and maintenance of the information [17]. In recent years, research on OBDI applications from heterogeneous data sources in different real-world scenarios has been intensified. One of the most prominent scenarios concerns the geospatial data domain, in which ontologies deals with the totality of geospatial concepts, categories, relations and processes and with their interrelations at different resolutions [10]. According to the reported in [17], the semantic data integration can be defined in three variants, which are based on what kind of ontologies are used and how these ontologies relate to each other: (i) the single-ontology, (ii) the multiple-ontology, and (iii) the hybrid OBDI. Our proposed data

management model would classify to the first variant because a single domain ontology is developed and used to integrate all data sources.

The ontology offers the conceptualization and the required semantic description to reach a high level of data integration. However, we still have a problem associated to the static behavior of the actual querying process in the SIGOBE system. If a static query is developed for each problem that arises, the database begins to store a group of scarcely-used queries. In order solve the problem, the system must be able to generate intelligent queries in real time, in which the knowledge obtained from previous ones is used and the Case-Based Reasoning (CBR) constitute a promising alternative for this propose. CBR is based on the premise that similar previous problems will have similar solutions. CBR have three main components: a user interface, a knowledge base (cases database or cases repository) and an inference engine and contributes to progressive learning, so that the domain does not need to be fully represented. The CBR allow quickly solutions to the problems, because: the answers are not derived from scratch, but from previously solved cases; propose solutions in domains not fully understood; they offer a means of evaluating solutions when an algorithmic method is not available; focus on the most important characteristics or parts of the problem; and lighten the knowledge engineering required to build the knowledge base, because it works directly with cases or examples of the problem to be solved, without mediating a particular form of knowledge representation. An analysis of the queries carried out, including those for SIGERE, allows establishing the structure of a case to solve the problem, which is divided into predictive traits and objective traits; as it will be shown ahead.

3 Data Management Model Proposed for SIGOBE

In the proposed model, the information needs by end users are formulated in terms of natural language sentences and using a technical vocabulary, instead of the data sources. The sentences are automatically translated into operations (queries) over the data sources, which results are graphically visualized in the geographic interface of SIGOBE. A domain ontology, named OntoSIGOBE, was manually developed to represents the conceptualization associated to the distribution and transmission processes of electrical energy, as well as the captured knowledge from the heterogeneous data sources of SIGE, in order to achieve semantic interoperability between them in the query answering process. An Intelligent Query Answering Process exploits the conceptual schema (OntoSIGOBE) as an intermediate layer for accessing and querying the data source that given information to SIGOBE. This process is carried out combining the query preprocessing with the case-based reasoning, for increasing the efficiency and flexibility of the querying process by end users. An overview of the proposed model is shown in Fig. 1.

3.1 OntoSIGOBE Construction

OntoSIGOBE constitute a domain ontology, and formalize the semantic meaning of the conceptualization associated to the processes of distribution and transmission of electrical energy in the SIGE context. The conceptualization represents the implicit captured knowledge from the SIGERE and SIGECIE Databases, and the SIGE Cartography, in order to achieve semantic interoperability between these data sources in the query answering process in SIGOBE. This conceptualization was enriched and refined considering the captured knowledge from several domain experts in the energy-electro sector which were interviewees. The construction process of this ontology was carried out using Methontology [4] and Protégé, and OWL was used for coding the formalized knowledge.

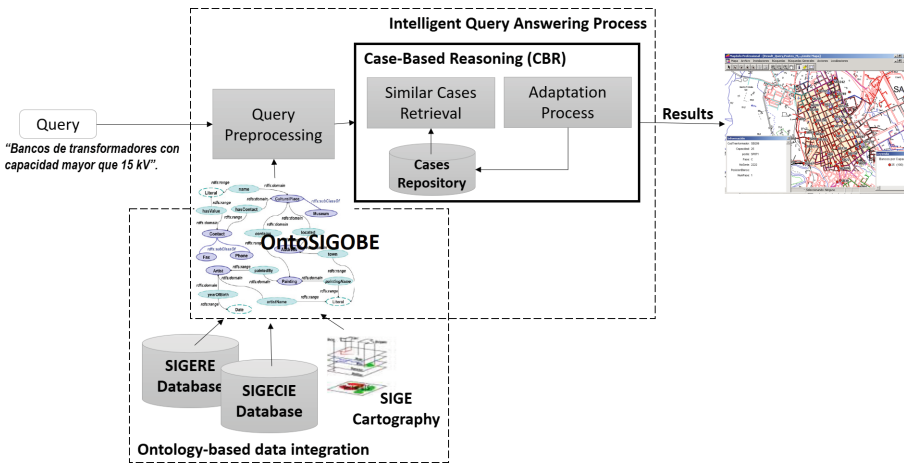


Fig. 1. Overview of the data management model proposed

In OntoSIGOBE, different concepts types are represented as *owl:Class* and integrated in the taxonomy: geographic and electrical domain concepts. Specifically, three concepts groups were represented and integrated: electrical concepts defined as nomenclatures in the databases, such as: administrative and political structures (according to the administrative organization of the Cuban Electrical Union), capacities, voltages, manufacturer, among others; electrical elements geographically represented in the cartography, which have spatial and alphanumeric information, such as: circuits, transformer installations, lamps, disconnectors, among others; and geographic concepts, such as: geographical areas, regions, natural geographical accidents or built by the man with impact in the electric system operation, among others. The basic elements of the electric network are represented in correspondence with their geographical characteristics (e.g. position in the geographic space). From the perspective of the knowledge representation architecture, OntoSIGOBE has been modelling according to the schema shown in Fig. 2.

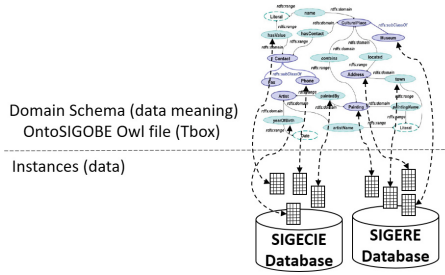


Fig. 2. Overview of the knowledge representation architecture

The TBox defines concepts and relationships between them (taxonomy and semantic relationship) and is stored in the OWL file, whereas the instances (data) are not included in the OWL file, these are stored in the data sources. The instances retrieval process in the proposed model is carried out through SQL queries, which are defined in the specification of each case stored in the knowledge base of the CBR Module (descriptions are given in above section). Some classes of the taxonomy of OntoSIGOBE are shown in Fig. 3.

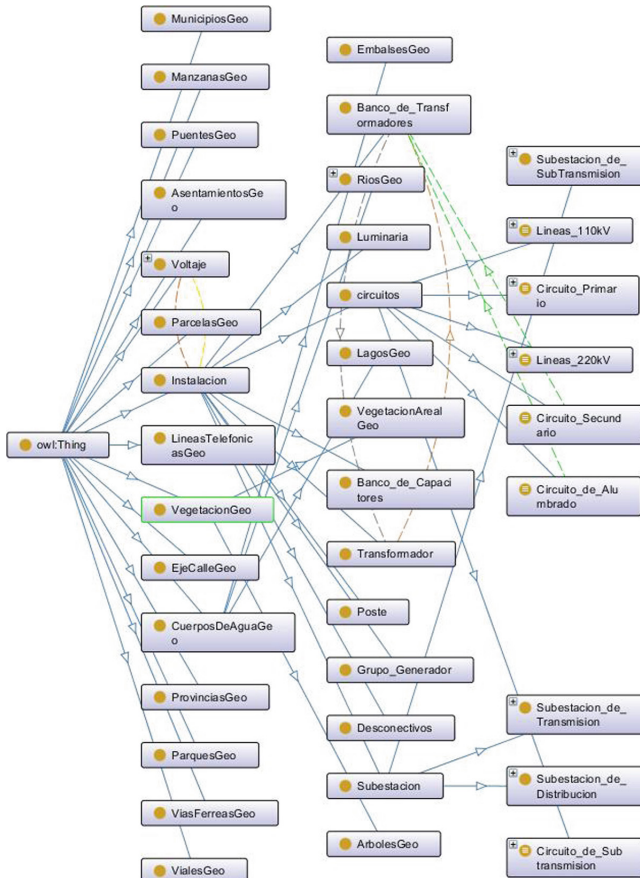


Fig. 3. Portion of the taxonomy of the OntoSIGOBE

In this taxonomy some examples of electrical elements are represented as *owl:Class*, such as: “Instalación” (*Installation*, in English) and represented as *rdfs:subClassOf* of it are “Circuitos” (*Circuits*, in English) and “Banco_de_Capacitores” (*Capacitors Bank*, in English), and from “Circuitos” others *rdfs:subClassOf* are derived: “Circuito_Primary” (*Primary Circuit*, in English) and “Líneas_220 kV” (*Line 220 kV*, in English). Several types of non-taxonomy or semantic relationship between concepts are represented in the conceptualization of OntoSIGOBE, which are formally defined as Object Properties and coding using *owl:ObjectProperty* specifications. These relationships have been defined from electric and geographic perspective. One example of electric domain relationship is “alimenta” (*feeds*, in English), useful to modelling the case of exists some sub-stations and transformers bank that feeds different circuits, and an example of geographic domain relationship is “intersecta” (*intersect*, in English), useful to modelling the case of the “RiosGeo”, “VialesGeo” or “ViasFerreasGeo” (captured concepts from the SIGE Cartography) that intersect circuits. The Object Properties in OntoSIGOBE are shown in Table 1. Several property restrictions, such as: *hasValue* and *someValuesFrom*, have been specified in the formalization of the defined Object Property (some of them are shown in Fig. 4),

Table 1. Examples of semantic relationship

Properties	Inverse (<i>owl:inverseOf</i>)	Examples of real situations
alimenta	alimentada_por	A circuit <u>feeds</u> to a transformers bank
contiene	esta_contenido_en	A Substation <u>contains</u> a control hut
dista	dista_de	A transformers bank is at a <u>distance</u> of a road
intersecta	intersectada_por	The line phone <u>intersect</u> the electric line
pertenece_a	tiene	A transformer <u>belongs</u> to a specific transformers bank

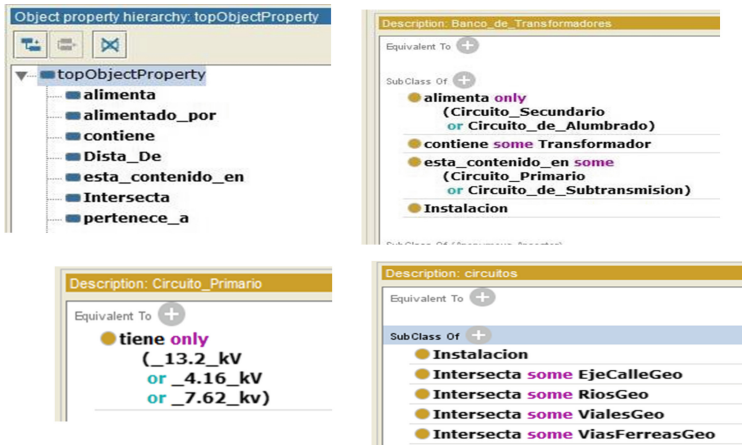


Fig. 4. Examples of object properties and property restrictions included in OntoSIGOBE

for increasing the semantic expressiveness of the OntoSIGOBE and the capabilities to answer more complex queries. Finally, in the Table 2 some composition elements of the OntoSIGOBE are shown.

Table 2. Composition of the OntoSIGOBE

Classes	SubClassOf relationship	Equivalent classes	Object properties	Axiom	Semantic relationship
53	63	6	10	155	929

3.2 Query Preprocessing

The users have the possibility of querying the SIGOBE system by free sentences in natural language using a technical vocabulary; for example: “*Bancos de transformadores con una capacidad mayor que 15 kV*”. In order to increase the efficiency in this querying process, the new proposed data management model for SIGOBE includes an intelligent query answering process, which is based on the use of case-based reasoning technic. In this sense, the query processing task is carried out for extracting the relevant information and predictive features from queries for understanding the information necessities of users based on the application of the case-based reasoning technic.

In this process, basic NLP tasks, like lexical-syntactic and semantic analysis, are applied to the query. The tokens without meaning are removed from the query sentence, e.g.: stop words, articles, and punctuation signs. Next, the relevant concepts and the relationship between them are identified through the syntactic and semantic analysis, which is supported by the OntoSIGOBE developed. These analyses include the identification of entities, the type of relationship between entities, and the search or filter requirements. This identification process is carried out by using the represented knowledge in the OntoSIGOBE as reference vocabulary to interpret the information available in the data sources. As a result of this process, the user’s query is transformed in a set of predictive features that considers the information elements captured from queries as values of those features. In Table 3, the specifications and meaning of the predictive and objective features are described.

The ON and OG ontological features are represented using a description logics approach. A possible value of the ON would be: $T \cap TPot \cap TMonophasic \cap \neg SSecondary$. This range expresses that the element is a monophasic primary transformer without secondary output. OG works similarly but considering spatial relationship. An example that refers to the location of an element would be: $P \cap Prov \cap Muncp$, which expresses that an element belongs to the country (P), to a province ($Prov$) and to a municipality ($Muncp$). The structure of the cases in this proposal is represented in Fig. 5, and it is based on the cases repository structure reported in [15].

Table 3. Universe of discourse of predictive and objective features

Features	Possible values	Type
Predictive features		
NV	Secondary, Primary, Subtransmission, Transmission	Symbolic and single-valued
EB	Posts, Transformer banks, Capacitor banks, Generator groups, Disconnectors, Structures, Lamps, Transmission Circuit, Subtransmission Circuit, Primary Circuit, Secondary Circuit, Lighting Circuit, Distribution Substation, Transmission Substation	Symbolic and single-valued
AT	Attributes to be returned by the inquiry (code, voltage, name, etc.)	Set
Tables	Tables of the SIGERE involved in the inquiry (Accessories, Actions, Connection, Interruptions, Line, CurrentSupplyPrimary)	Set
CA	Element to compare (Attribute being compared)	Symbolic and single-valued
OP	Operator (\cup , \cap , \leq , \geq , =, like, ...)	Symbolic and single-valued
ON	Ontology (description logics)	OntoSIGOBE
OG	Spatial constraint (description logics)	OntoSIGOBE
Objective features		
From	Returns the From of the inquiry to the SIGERE	String
Where	Returns the Where of the inquiry to the SIGERE	String
CE	Returns the GIS inquiry	String

Specifications: NV: voltage level; EB: Base element; AT: Attributes in analysis; T1... T10: Tables involved; CA: Analysis Conditions; OP: Operator analysis; ON: Ontology Base; OG: Ontology Space.

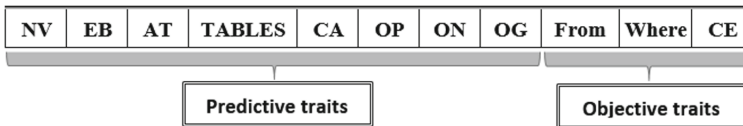


Fig. 5. Structure of the cases.

3.3 Case-Based Reasoning

Taking as a premise that similar problems will have similar solutions, Case Based Reasoning (CBR) is used as a tool in order to define an intelligent querying mechanism. The CBR is a method of artificial intelligence to solve unstructured problems, where the reasoning is made from an associative memory, which uses an algorithm to determine a measure of similarity between two objects. In cases-based systems, the domain does not have to be completely represented [8] and learning is progressive [16].

The CBR is a cycle so-called 4R that has the following stages [11]: retrieves, reuse, revise and retain.

Global dissimilarity is determined by Eq. 1. The distances are weighted considering the expert criterion, with a weight w_i , the greater w_i , the greater the importance of the trait. The most important traits are the ontological ones.

$$\text{DisSimGlobal}(X, Y) = \sum_{i=0}^m w_i * d_i(x_j, y_j) / n \quad (1)$$

where $\sum w_i = 1$.

Local distance $d_i = (x_i, y_i)$ is determined by the type of data x_i, y_i . In the case presented here there are three types of data for which different distance measures are used.

The defined recover process in Algorithm 1 obtains the k cases closest to the inquiry requested by the user, using Eq. 1 to calculate the distance. The result of the evaluation of the selected k system is 3 by default. The set of cases obtained by Algorithm 1 (Table 4) constitutes the input of the adapter module. Based on the transformational analogy, to propose an initial solution is developed [15]. The new case is evaluated and adapted to the conditions on the recovered cases. The pre-set consultations are not necessarily identical to those stored in previous cases. To develop an initial solution, all recovered cases are considered and a combination of the recovered solutions is taken as a starting point.

The input of the adaptation module is an initial solution of the three objective traits. This module allows to reuse and adapt based on transformational analogy, which implies structural changes in the solution. Transformational adaptation is guided by common sense where the rules were defined and used in the adaptation process. This process is considered a T-space, where the known solution (KS) is going to be transformed with the use of T-operators, until it becomes the solution of a new problem.

The retention is induced from the cases, so it will be necessary to redefine it periodically. The efficiency of the system is affected when the number of cases grows excessively, therefore, it is important to avoid including cases that do not contribute new information to the system. To carry out the retention of cases, the following steps are followed:

- The degree of information provided by the case to the system is calculated. This degree of information is estimated by the number of T-operators applied between the set of T-operators in the T-space.
- It is considered feasible to retain the case whose degree of information provided is greater than α (α represents an information threshold).
- If the case is feasible, it is retained in the corresponding sub-base according to the value of the *NV*, *EB* and *OP* predictive traits, given the calculation of the degree of information provided by an objective trait.

The degree of information provided to the system by the value of an objective trait is calculated as the minimum number of T-operators applied.

Table 4. Algorithm 1: Recover algorithm of k most similar cases

Input: P ; problem to solve and $BC' = \{bc_1, bc_2, \dots, bc_n\}$; sub-set of cases obtained from the sub-base that corresponds to the P case

Output: Sets of K more similar cases

$\forall bc_j \in BC'$ is calculated the $DisSimGlobal(bc_j, P)$. **Where:**

$$DisSimGlobal(bc_j, P) = \sum_{i=0}^8 w_i * d_i(bc_{ji}, P_i) / n$$

$$d_i(bc_{ji}, P_i) = \begin{cases} d_i(bc_{ji}, P_i) = \begin{cases} 0 & \text{Si } bc_{ji} == P_i \\ 1 & \text{Si } bc_{ji} \neq P_i \end{cases} & i = 1, 2, 5, 6 \\ d_i(bc_{ji}, P_i) = \frac{|bc_{ji} \cap P_i|}{|bc_{ji} \cup P_i|} & i = 3, 4 \\ d_i(bc_{ji}, P_i) = \frac{1}{3} \left(\frac{c}{|bc_{ji}|} + \frac{c}{|P_i|} + \frac{c-t/2}{c} \right) & i = 7, 8 \end{cases}$$

The most similar K cases are returned with a $K = 3$

4 Experimental Results and Quality Evaluation

SIGOBE is considered a support decision system with national scope, therefore applicable to different areas inside the Cuban National Electric Union. In order to apply the proposed data management model in SIGOBE, the module SICUNE (Intelligent System of Consultation for the UNE) was implemented. SICUNE implement the proposed model according to a structure of six packages: Set, String and Position, for grouping the functionalities and procedures associated to the sets, strings, and position similarity processing, respectively; Structure and Useful, for grouping the functionalities and procedures associated to the access and management of the cases base, respectively; and Visual CBR for establishing the link between the interface and the application. The system was designed with the possibility of adding new measures of similarity by attributes in the CBR process, although the Jaro-Winkler distance [18] was used in the experiment carried out (97% accuracy was obtained with this function).

Application or task-based evaluation constitutes one of those classification referred to mostly by recent research in ontology evaluations [5, 14]. Task-based approaches try to measure how far an ontology helps improving the results of a certain task [14]. In the proposed model, OntoSIGOBE was designed for improving the performance of the SIGOBE system in the querying process associated to customers' information needs about the distribution and transmission processes in the electrical sector, through a CBR approach and the SICUNE module.

Inspired in the task-based approach, we evaluate the proposed ontology-based data management model using SICUNE in the context of the daily operational work of three departments of the Electric Union: Office, Engineering Department, and Customer Service Area, and measuring the precision of the proposed model in the question answering process. These departments use information from different areas of the databases and achieve greater coverage in the information contained. A period of one month was considered as time interval for the evaluation tests; an average of 175 queries were carried out on SIGOBE from those departments. Additionally, a knowledge base composed of 265 cases was developed and used in this evaluation task.

Table 5 shows the obtained results. The engineering area was the one with the lowest representation in the knowledge base, because version one of SIGOBE focused on the Office, attention to complaints from the population and the investment area. However, 94.18% of precision was obtained.

Table 5. Results of the proposed ontology-based data management model using SICUNE

Areas	Queries	Correct predictions	Incorrect predictions	Learned	Precision (%)
Office	230	221	9	23	96.08
Engineering department	189	178	11	40	94.18
Customer service area	147	140	7	18	95.23
Total	566	539	27	81	
Precision Ave.					95.16%

OntoSIGOBE was also evaluated using the OOPS! (OntOlogy Pitfall Scanner!) system. OOPS! is actually the most complete available system for the (semi)automatically diagnose of ontologies; this system implements the quality model for ontology diagnose proposed in [13]. The quality model aligns the pitfall catalogue to the existing quality models for semantic technologies, which describe 41 evaluation pitfalls classified in the following dimensions: structural, functional, usability-profiling, consistency, completeness, and conciseness. The pitfalls report generated by the OOPS! system is shown in Table 6.

Table 6. Evaluation report of OntoSIGOBE from OOPS! system

Pitfalls	Definition	Category of evaluation	Important level	Argumentation
P08	Missing annotations	Usability-profiling	Minor	–
P10	Missing disjointness	Completeness	Important	OntoSIGOBE not include disjoint axioms because no classes fulfilled this characteristic
P22	Using different naming conventions in the ontology	Usability-profiling	Minor	–

Considering that SIGOBE constitutes a real system whose results have a direct impact in the decisions making of the Cuban Electric Union, a software-quality evaluation process was carried out. Specifically, the model reported in [9], which is based on ISO-9126:2002 was used, and the results are shown in Table 7.

Table 7. Results of the software-quality evaluation

External and internal evaluation		Usability evaluation	
Attributes	Values	Attributes	Values
Functionality	10,0	Efficacy	9,0
Reliability	10,0	Productivity	10,0
Usability	10,0	Satisfaction	10,0
Efficiency	10,0	Security	9,5
Maintenance capacity	10,0		
Portability	8,0		
Subtotal	58,0	Subtotal	38,5
Total			96,5

The results were computed from the captured valuation of five main specialists of technical departments of the Cuban Electric Union. The specialists gave a weight to each attribute in a value range of 1 to 10, being 10 the maximum score. ISO-9126:2002 is an official, approved and validated standard that aims to establish an international standard for the evaluation of the quality of computer systems using metric indicators. The obtained results are satisfactory because the system complies with 96.5% of the indicators.

5 Conclusions and Future Works

This paper presented an ontology-based data management model applied to the Geographic Information System denominated SIGOBE, which combines the use of a developed domain ontology with the application of the case-based reasoning technic for defining an intelligent query answering process and improving the performance of SIGOBE system. The application of the proposed model to SIGOBE allowed to increase the spectrum of successfully answered requests of the specialists and to facilitate several functions and operations such as: (1) to locate objections from the population associated to failed installation or abnormal parameters, (2) to organize the routes of the technical vehicles (displaying the customer voltages on the map, (3) to develop studies of equipment faults in rural areas, and (4) to optimize the electrical networks and their use.

Through the OntoSIGOBE and NLP tasks, applied to the query processing, comprehensive access to the database and dynamic queries are achieved, which are fundamental in high demand stages of services due to the speed of the link between alphanumeric and geographic information, increasing the ease and convenience with which requests are made. The obtained results in the evaluation process are promising because high accuracy rates and high satisfaction of the software quality indicators were achieved. On the other hand, satisfactory results were obtained in the automatic evaluation of OntoSIGOBE using OOPS! system.

As future work, several techniques that allow the ontology to be further exploited will be evaluated, as well as the increase of the represented knowledge, with the objective that the ontology itself becomes a query interface to satisfy more complex user requests. Furthermore, others alternative for obtaining the weights of the features will be analyzed in order to improves the global dissimilarity function and the results of the similar-cases retrieval process.

References

1. Abiteboul, S., Arenas, M., Barceló, P., Bienvenu, M., Calvanese, D., David, C., Hull, R., Hüllermeier, E., Kimelfeld, B., Libkin, L., Martens, W., Milo, T., Murlak, F., Neven, F., Ortiz, M., Schwentick, T., Stoyanovich, J., Su, J., Suciu, D., Vianu, V., Yi, K.: Research Directions for Principles of Data Management (Dagstuhl Perspectives Workshop 16151). [arXiv:1701.09007v1](https://arxiv.org/abs/1701.09007v1) [cs.DB], Dagstuhl Manifestos, pp. 1–28 (2017)
2. Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Rosati, R.: Ontology-based data access and integration. In: Liu, L., Tamer Özsu, M. (eds.) *Encyclopedia of Database Systems*, pp. 1–7. Springer, New York (2017)
3. Fernández, R.: *Informatización de la Gestión de las Redes Eléctricas*. Ph.D. thesis, Facultad de Ingeniería Eléctrica, Universidad Central “Marta Abreu” de Las Villas, Santa Clara, Cuba (2011)
4. Fernandez, M., Gómez, A.: Overview and analysis of methodologies for building ontologies. *Knowl. Eng. Rev.* **17**(2), 129–156 (2002)
5. Hlomani, H., Stacey, D.: Approaches, methods, metrics, measures, and subjectivity in ontology evaluation: a survey. *Semant. Web J.* **1**(5), 1–11 (2014)
6. Hu, Y.: Geospatial semantics. In: Huang, B., Cova, T.J., Tsou, M.-H., et al. (eds.) *Comprehensive Geographic Information Systems*. Elsevier, UK (2017)
7. Lenzerini, L.: Ontology-based data management. In: *Proceedings of the 6th Alberto Mendelzon International Workshop on Foundations of Data Management (AMW 2012)*, *CEUR Workshop Proceedings*, vol. 866, pp. 12–15 (2012)
8. Liu, X., Wang, X., Wright, G., Cheng, J.C., Li, X., Liu, R.: A state-of-the-art review on the integration of building information modeling (BIM) and geographic information system (GIS). *ISPRS Int. J. Geo-Inf.* **6**(2), 53–74 (2017)
9. Macías Rivero, Y., Sánchez, G., Victoria, M., Martínez Suárez, Y.: Modelo de evaluación para software que emplean indicadores métricos en la vigilancia científico-tecnológica. *ACIMED* **20**(6), 125–140 (2009)
10. Mostafavi, M., Edwards, G., Jeansoulin, R.: An ontology-based method for quality assessment of spatial data bases. In: *Proceedings of the 3rd International Symposium on Spatial Data Quality (Geoinfo series)*, vol. 1/28a, pp. 49–66 (2004)
11. Ortiz, Y.K., Bañuelos, P., Rodas, J.: Razonamiento basado en casos. *CULCyT* **14**(63), 48–56 (2017)
12. Poggi, A., Lembo, D., Calvanese, D., De Giacomo, G., Lenzerini, M., Rosati, R.: Linking data to ontologies. *J. Data Semant.* 133–173 (2008)
13. Villalón, M.P.: *Ontology evaluation: a pitfall-based approach to ontology diagnosis*. Ph.D. thesis, Departamento de Inteligencia Artificial, Escuela Técnica Superior de Ingenieros Informáticos, Universidad Politécnica de Madrid, España (2016)
14. Raad, J., Cruz, Ch.: A survey on ontology evaluation methods. In: *Proceedings of the International Conference on Knowledge Engineering and Ontology Development (IC3K 2015)*. *ACM DL*, pp. 179–186 (2015)

15. Sanchez, N., Comas, R., García, M., Carrera, F.: Geographical information system based on artificial intelligence techniques. In: Proceedings of International Conference on Technology Trends (CITT 2018), CCIS, vol. 895, pp. 446–461. Springer, Cham (2018)
16. Tah, J., Oti, A., Abanda, F.: A state-of-the-art review of built environment information modelling (BeIM). *Organ. Technol. Manag. Constr.* **9**(1), 1638–1654 (2017)
17. Wache, H., Voegelé, T., Visser, T., Stuckenschmidt, H., Schuster, H., Neumann, G., Huebner, S.: Ontology-based integration of information - a survey of existing approaches. In: IJCAI 2001 Workshop: Ontologies and Information, pp. 108–117 (2001)
18. Winkler, W.E.: Using the EM algorithm for weight computation in the Fellegi-Sunter model of record linkage. In: Proceedings of Section on Survey Research Methods. American Statistical Association (2000)