Ingénierie des Systèmes d'Information
Vol. 28, No. 1, February, 2023, pp. 31-40
Journal homepage: http://iieta.org/journals/isi

Efficient Criteria Based Method for Selection of Relevant Ontologies in Transport Domain

Sylvain G. Gbame[1*], Marcellin K. Brou[1], Bernard K. Saha[2]

[1] Mathematics and Computer Science Department, Institut National Polytechnique Félix Houphouët-Boigny (INP-HB), Yamoussoukro BP 1093, Côte d'Ivoire
[2] Science and Technology Department / Mathematics and Computer Science Section Ecole Normale Supérieure, Abidjan 08 BP 10, Côte d'Ivoire

Corresponding Author Email: gbede.gbame18@inphb.ci

https://doi.org/10.18280/isi.280104

Received: 27 September 2022
Accepted: 20 January 2023

Keywords:
criteria, relevant ontology, ontological layer, smart city, transport

ABSTRACT

The smart city is one of the most appropriate solutions to the challenges posed by uncontrolled urbanization, rapid population growth, waste of energy re-sources, and environmental pollution. As a result, many cities have undergone real transformation in the economic, civic, governmental, housing, environmental and especially transport domain. However, given the difficulties of representing the transport domain, several ontologies have been designed to overcome this difficulty. These ontologies enable requirements planning and decision making. Despite this fact, many of the proposed ontologies are difficult to reuse for various reasons such as their obsolescence or their specificity to a given city. Thus, in this study, we propose methods using criteria to determine the relevant ontologies among several others. Simulations confirmed that one these methods is very efficient.

1. INTRODUCTION

To preserve the planet, environmental experts and urban planners affiliated with the United Nations (UN) have thought of new innovative modes of social organizations [1]. This innovative conception of cities has also aroused various notions of smart cities. These cities are built using Information and Communication Technologies (ICT) to detect, analyse, and integrate key information from central systems in the cities in operation according to IBM [2]. A smart city is also a future-oriented high-performance city based on six characteristics, which are based on the intelligent combination of citizen activities [3].

Smart cities have risen owing to various needs and changes in cities. For example, cities are becoming increasingly congested, the environment is polluted by traffic jams, and energy and materials are not used efficiently.

Thus, several services have been established to create smart cities. These include the construction of advanced transport facilities such as railways, airports, ports, subways, and motorways. The need to live in these smart cities and consequently in a pleasant, healthy, structured and harmonious world has led researchers to use ontologies. An ontology is a set of hierarchically structured terms designed to describe a domain. An ontology can serve as a framework for a knowledge base [4]. The ontologies are used in several domains such as: information technology, education, health, the economy and transport [5]. In the domain of transport, ontologies such as SUMO [4], were designed to study the representation and management of related knowledge for highway systems.

In recent years, because of their role in the semantic web, ontologies evaluations have played an important role in knowledge description [6]. Indeed, evaluation has enabled the definition of a shared knowledge base that can be used by machines acting on behalf of humans [7]. Thus, the popularization of ontologies in the semantic web by developers depends on the issue of evaluation [8]. These evaluations have been carried out using several methods, including: gold standard based methods, corpus based methods, task-based methods and criteria based methods [9], [10]. Several effective ontologies have been proposed for transport domain. However, each of these transport ontologies is highly dependent on the characteristics or concepts of the city for which it was designed. As cities evolve rapidly, these ontologies become obsolete quickly. This implies regular updating of the ontologies. which is tedious. Therefore, in this study, we focus on the problem of determining the most relevant ontologies among several ontologies of the transport domain.

To solve this problem, we propose methods that takes into account criteria, and a method based on ontological layers [11]. The remainder of this paper is organized as follows. Section 2 presents related work. Section 3 gives a formulation the problem. Section 4 is a description of the proposed method. Sections 5 and 6 are related to the evaluated results, and the conclusion and future work respectively.

2. RELATED WORK

The frequent use of ontologies in decision making requires users to evaluate their robustness. This evaluation, which involves measuring the quality of semantic resources, has facilitated software development. In this work four different methods of ontology evaluation are presented [9]: gold standard, corpus-based, task-based, and criteria-based methods.

2.1 Gold standard method

An ontology can be evaluated using the gold standard method which compares the learned ontology with a previously created reference ontology, known as the "gold standard". The gold standard ontology is predefined by manual design by domain experts. Thus, the authors [12, 13] proposed to set up a meta-model for the analysis, comparison, and engineering of ontologies. This approach uses three steps: Firstly, a set of naming conventions is established for each comparable element in the two ontological models. Secondly, a comparison of the deeper structure of the models must be performed on the models as a whole, following a more complex form of analysis.

Finally, the models were compared as a whole, rather than as a selected part, to establish an accurate comparison of the completeness of the scope of the models. The authors [14] proposed to measure the performance of human experts in manually classifying classes in a general knowledge domain ontology with Basic Formal Ontology (BFO) entities. The tasks were conducted as part of a web-based survey. The finding indicates that even for a well-understood general knowledge domain such as travel, the results of manual classification tasks are inconsistent [15]. The number of participants was relatively low, given the number of BFO experts around the world and people willing to participate. Note that the Gold standard method is used in cases where the ontology is automatically generated. In many cases, the application of this method is difficult to apply, because such a gold standard does not exist.

2.2 Corpus-based method

A Corpus-based method is used to assess the extent to which an ontology sufficiently covers a domain. This is performed by comparing the learned ontology with the content of a corpus of texts that accurately covers a specified domain. This method proves to be efficient because it allows the comparison of one or more ontologies with a corpus, rather than comparing an ontology with an existing one. The authors [16] used Google browser to find a corpus based on a user's query. The introduction of the query through WordNet made it possible to obtain the first 100 pages of the results which formed the corpus to be evaluated. To this end, the authors [17] presented an ontology-based method for evaluating measures. This ontology provides useful guidance for the processing, conceptualization and representation of knowledge and the use of knowledge engineering.

This evaluation methodology involves the following processes: literature review, concept extraction by defining a set of criteria and sub-criteria, construction of the taxonomy and construction of the ontology. Task-based method attempts to measure how an ontology helps to improve the results of a certain task [18]. According to this method, an ontology is intended for a particular task. It is evaluated only in terms of its performance in solving the difficulties of this task. The authors [19] proposed an evaluation on the quality of an ontology to determine the efficiency of users to obtain relevant individuals during their searches. This efficiency was measured using a cost model to quantify the effort made by the user to obtain the desired information. Task-based methods were used to propose a PROMPTDIFF algorithm for maintaining an ontology [20]. The PROMPTDIFF algorithm integrates different heuristic matches to compare ontology versions. The authors [21] combined these matches in a fixed manner, using the results of one match as input for the others until they no longer produced changes. This algorithm is devided into three different definitions: structural difference, PROMPTDIFF table and the monotonicity principle. The evaluation of these ontologies is only relevant to their particular application. When you want to use these ontologies in other applications, evaluation is no longer relevant.

2.3 Task-based method

Task-based method attempts to measure how an ontology helps to improve the results of a certain task [21]. According to this method, an ontology is intended for a particular task. It is assessed only in terms of its performance in solving task difficulties. The author [22] presented a new method for evaluation and comparison. He proposes a new notion of risk. Then, he demonstrates that the Maximum Likelihood Estimate (MLE) of risk compared to a performance score is similar to the complement of the same score. This method provides important information on the performance of alignment and allows the comparison of alignment systems. In addition to the work [22], the work [19] consists of the implementation of a new evaluation and comparison method based on multiple performance measures. This method takes into account the preferences of experts using a Multi-Criteria Decision Making Method (MCDM). It allows experts to make judicious choices regarding the performance of a task or an application. The evaluation of all these ontologies is only relevant for their particular application. When one wants to use these ontologies in other applications, the evaluation is no longer relevant.

2.4 Criteria-based method

Finally, a qualitative criteria-based method consists of transport experts establishing a number of criteria that will allow them to assert the relevance of an ontology. The authors [23] classified the criteria into three categories: First, structural dimensions that refer to the hierarchical structure of ontologies represented by graphs. Next, functional dimension is focused on the selection, construction and exploitation of a given ontology and its components (concepts, hierarchical relations, roles). Finally, the applicative dimension is based on the usability of the ontology and depends on its annotation level. The author [22] evaluated and compared alignment systems. Furthermore, because of its various applications, ontology alignment has been the subject of much research, so that many alignment systems have been proposed to discover the correspondences of two given ontologies. A Bayesian test was designed to compare different systems (A1 and A2) based on their estimated risks, thereby calculating the confidence in the superiority of one system over another. The authors [7, 24] worked on the implementation of criteria from several other works. Indeed, the authors [24] proposed the ONTOMETRIC method to help the user to determine the most appropriate ontologies for a system by comparison. This method is based on the Analytical Hierarchy Process (AHP) by first comparing the relative importance of each criterion against all others.

This comparison method, which is inspired by the Analytical Hierarchy Process (AHP), calculates the relative weights of the criteria, and normalizes the weights to obtain the measures for the existing alternatives [25]. The authors [26] summarized these criteria in their work. Although this method can be applied to determine the most appropriate ontology for the analysis of business systems, it lacks precision in the fine details of the objectives of the evaluators' analysis. The specification of a certain ontology is also complicated and time consuming, and the comparison is subjective. Therefore, the authors [11, 27] addressed this limitation. The authors [11] conducted a comparative study to identify overlaps and establish the relationship between the different criteria. Their work focuses on the evaluation of qualitative criteria in relation to influencing factors, namely methods and levels of evaluation. The relationships between the criteria are summarized in Table 1. The proposed table matrix highlights the root of the ontology criteria, highlighting those influenced by the software engineering metrics and those proposed by the cited articles. As for [27] they presented a comprehensive survey of existing transport ontologies by comparing these ontologies to serve as a useful resource for the applied ontology and transport research communities. To describe the clarity and comprehensibility of one ontology compared to others, [27] extended the results on insight. A distinction is then made between ontologies that capture the same concepts, ontologies whose semantics are shared between them and ontologies whose semantics differ. Finally, the work [27] studies the relevance of an ontology according to each criterion. The work [27] does not make it possible to distinguish the set of relevant ontologies among several ontologies. Therefore, the purpose of our study is to find a set of relevant ontologies among several ontologies.

Table 1. Table of criteria and layers [11]

Ontological layers	Criteria
Semantic	"adaptability","analyzability","authority","clarity","cohesion","compatibility","adequacy","completeness","consistency","interoperability", "maintainability","replaceability","reconciliation","structural","testability","transparency","usability"(**18**)
context	"accuracy","clarity","completeness","conformance","extensibility","effectiveness","expressiveness","functionality","history","modification_stability","maturity","usability","quality_of_use", "readability","recall","relevance","reliability","replaceability","re-usability", "robustness"(**20**)
structure	"testability","interlocking","structural","redundancy","modularity","maintainability","disjoint", "cycle", "changeability"**(9)**
lexical	"clarity","cohesion","completeness","vocabulary_control","maintainability","precision","recall","reconciliation","usability"**(9)**
syntax	"transferability","testability","richness","reference","portability","maintainability","legality","formalisation","flexibility","cycle", "completeness" **(11)**

3. PROBLEM SPECIFICATION

In this section, the assumptions, notations, and definition of the problem are given.

3.1 Assumptions and notations

Note that the ontology criteria used are qualitative criteria. Indeed, they have proven to be empirically meaningful and useful in comparing ontologies, as opposed to metrics [27]. The objective of this study is to compare a number of transport ontologies using criteria defined [11]. In short, considering that all ontologies in the transport domain have been designed for specific purposes, the idea is to identify a set of relevant ontologies. The reified representation of an ontology is the transformation of objects into properties [28]. In other words, reification is a technique that allows a richer description of a property. Generally, reification has been used to support description of the source of knowledge; instead of statements of fact:

Ex: 'The father kicked the cat', we can say 'I saw the father kick the cat'. However, more generally, reification can support many different kinds of statements about properties [28].

We make the following assumptions:

- Assumption 1: All ontologies belong to the same domain, but have different specific purposes. This is because each of them has been designed for well-defined purposes in the transport domain, which is characterized by a set of criteria.
- Assumption 2: All criteria have the same value. More clearly, no one criterion is more important than others.

Useful notations are presented as follows:

The set of transport ontologies is denoted by OT and OT is given by Eq. (1):

$$\mathrm{OT}=\coprod_{k=1}^{p}\mathrm{OT}_k \tag{1}$$

where, OT_k is the k-th transport ontology belonging to OT and p is equal to |OT|.

The set of criteria of each ontology according to the research [11] is denoted by C and C is given by Eq. (2).

$$\mathrm{C}=\coprod_{j=1}^{n}\mathrm{c_j} \tag{2}$$

3.2 Problem formulation

Given: The set of transport ontologies OT, the set of criteria of each ontology C.

Goal: Find a small set of relevant ontologies $\mathrm{OT^r}$, with $\mathrm{OT^r} \subseteq \mathrm{OT}$.

Constraint: $\mathrm{OT^r} \neq \emptyset$.

4. PROPOSED APPROACHES

This section presents our algorithms to solve the problem. In section 4.1, Algorithm 1 determines the set of relevant ontologies of the domain as the set of ontologies that are relevant according to each criterion characterising the transport domain. Then, we propose Algorithm 2 in section 4.2 which allows the selection of relevant ontologies using ontological criteria and ontological layers. Finally, in section 4.3 Algorithm 3 allows us to regularise the number of criteria according to which the returned ontologies are relevant. Indeed, all relevant ontologies returned by Algorithm 3 are relevant according to the same number of transport domain criteria. This number is as close as possible to the cardinality of the set of criteria C.

4.1 Characterization of relevant ontology based on all criteria

The review of the literature revealed that in the study [29], an algorithm for determining relevant ontologies was proposed. However, the relevance of ontologies depends on each criterion. In other words, an ontology OT_1 relevant according to criterion c_1 is not necessarily relevant according to criterion c_2. Thus, the algorithm proposed [29] does not allow to determine a set of relevant ontologies of the transport domain.

Recall that in the study [11], the authors characterized the transport domain using a set of ontological criteria. Therefore, we propose here, algorithm 1 which is an extension of the algorithm proposed in the study [27]. In short, Algorithm 1 determines the set of relevant ontologies of the domain as the set of ontologies that are relevant according to each criterion characterising the transport domain.

Indeed, the algorithm initializes the set of relevant ontologies to the empty set: This is the initialization phase (see line 2 of Algorithm 1). Then, the construction phase of the set of relevant ontologies (or phase 2) comprising two steps is performed:

- Step 1 (see lines 4 to 10): For each ontology, if the latter is not relevant according to a criterion, then the relevance according to the other criteria is no longer evaluated. For the evaluation of the relevance of an ontology according to a criterion (see line 6), we rely on the approach [27]. This approach takes into account the weight of this criterion. Then this weight is in the range [0.5, 1], then the ontology is relevant according to this criterion. However, when the weight of this criteria is in range [0, 0.5], then this ontology is not relevant according to this criterion. If the first ontology is relevant according to the first criterion, the algorithm proceeds to the second criteria. If it is relevant it continues, otherwise the algorithm moves on to the second ontology in the set. Algorithm 1 performs the same process for all ontologies in the set as for the first ontology. Otherwise, the relevance according to the next criterion is evaluated and so on.

- Step 2 (see lines 11 to 13): At the end of the processing of each ontology, if it is relevant according to all the criteria (which corresponds to *is_relevant* unchanged by step 1) then it is considered relevant and the set of relevant ontologies is updated.

Algorithm 1: Algorithm for the Selection of Relevant Ontologies based on all Criteria (ASROC)

Input
OT // Set of transport ontologies
C //Set of criteria in OT
Output
OT^r// Set of relevant ontologies
1 **Begin**
2 $OT^r = \emptyset$; // Phase 1: Initialization Phase
3 **For each** $OT_k \in OT$ **Do** // Phase 2: Phase of building set
// Begining of Step 1 of Phase 2
4 i= 1; is_relevant = True ;
5 **While** i $\leq |C|$ and is_relevant == True **Do**
6 **if** OT_k is not relevant according to each $c_i \in C$ **Then**
7 is_relevant = $False$
8 **End if**
9 i=i+1
10 **End while //** End of step 1 of Phase 2
// Begining of Step 2 of Phase 2
11 **if** is_relevant == True **Then**
12 $OT^r = OT^r \cup \{OT_k\}$;
13 **End if //** End of Step 2 of Phase 2
14 **End for**
15 **End**

From the above, it follows that the set of ontologies returned by Algorithm 1 is either non-empty or empty. In the latter case, this means that no ontology is relevant according to all criteria of the transport domain. This case is real because the more criteria characterising the domain [11] the more difficult it is to ensure that at least one ontology relevant to all domain criteria is found by Algorithm 1. Thus, this extension (i.e. Algorithm 1) of the solution proposed [27] does not efficiently solve all instances of the formulated problem (see SECTION 3.2).

Ideally, a relevant ontology must be relevant according to all domain criteria. The main advantage of algorithm 1 is that it returns a set of ideal relevant ontologies. However, in pratice, the ideal definition of the concept of relevant ontology is not always applicable. Thus, the disadvantage of this algorithm is that it very often returns an empty set of relevant ontologies.

Therefore, in the next sections, we attempt to propose algorithms that guarantee that the set of relevant ontologies is always non-empty.

4.2 Characterization of relevant ontology based on ontological layers

Here, we provide some useful definitions (about relevant ontology) which taking account ontological layers. Then, we present our ontological layers-based algorithm (see Algorithm 2).

4.2.1 Definitions

Definition 1: A transport ontology OT_k with $OT_k \in OT$ is said to be relevant when for each ontological layer the number of criteria of the ontology OT_k belonging to this layer is greater than or equal to the product of the relevance coefficient α and the number of criteria supported by the ontological layer (according to Table 1) presented in [11]. Note that $\alpha \in [0,1]$. More formally, Count_Crit (OT_k, L_i) is the number of criteria of the ontology OT_k belonging to layer $L_i \in L$ and is given by Eq. (3).

$$\text{Count\_Crit}(OT_k, L_i) = \sum_{j=1}^{n=|C|} M(OT_k, L_i, C_j) \quad (3)$$

$$\text{with } M(OT_k, L_i, C_j) = \begin{cases} 1, \text{if criteria } C_j \in L_i \\ 0, \text{Otherwise} \end{cases} \quad (4)$$

OT_k is a relevant ontology means that Inequation (5) is fulfilled.

$$\forall L_i \in L, \qquad \text{Count\_Crit}(OT_k, L_i) \geq H(L_i) \quad (5)$$

with $H(L_i)$ is the product of the relevance coefficient α and the number of criteria supported by the ontological layer $L_i \in L$, with α[0,1]. In general, α=0.2. α indicates the degree of relaxation of the stric constraint on the ideal relevant ontology concept used by algorithm 1. The smaller (and non-zero) this value is, the better the degree of relaxation of this concept. Recall that an ideal relevant ontology is one that is rekevant according to all domain criteria.

Illustration: Let OT_1 and OT_2 be two ontologies of the transport domain, which are illustrated in Tables 2 and 3.

In Table 2, with L_2=c_t= 'Context', Count_Crit(OT_1, L_2)=9 and $H(L_2)$=10 (according to Table 1). So, for the ontological layer ‘Context’, Inequation(5) is not fulfilled. Therefore, according to Definition 1, it results that OT_1 is not a relevant ontology.

According to Table 3:

(i) With $L_1 = s_c$ = ‘Semantic’, Count_Crit(OT_2, L_1) = 9 and $H(L_1) = 9$ (See Table 1). So, for the ontological layer ‘Semantic’, Inequation (5) is fulfilled.

(ii) With$L_2 = c_t$ = ‘Context’, Count_Crit(OT_1, L_2) = 11 and $H(L_2) = 10$ (according to Table 1). So, for the ontological layer ‘Context’, Eq. (4) is fulfilled.

(iii) with $L_3 = s_e$ = ‘Structure’, Count_Crit(OT_2, L_3) = 4 and $H(L_3) = 4$ (according to Table 1). So, for the ontological layer ‘Structure’, Inequation (5) is fulfilled.

(iv) With $L_4 = l_a$ = ‘Lexical’, Count_Crit(OT_2, L_4) = 7 and $H(L_4) = 4$ (according to Table 1). So, for the ontological layer ‘Lexical’, Inequation (5) is fulfilled.

(v) with $L_5 = s_x$, = ‘Syntax’, Count_Crit(OT_2, L_5) = 5 and $H(L_5) = 5$ (according to Table 1). So, for the ontological layers ‘Syntax’, Inequation (5) is fulfilled.

From (i) to (v), it results that OT_2 is a relevant ontology.

Table 2. Ontology OT_1

s_c	c_t	s_e	l_a	s_x	**Criteria of OT_1**
0	0	0	0	1	Formalisation
1	0	0	0	0	Adaptability
1	1	1	0	1	Testability
1	0	1	1	1	Maintenability
1	0	1	0	0	Structure
1	1	0	1	0	Clarity
1	0	0	1	0	Cohesion
1	0	0	1	0	Completeness
0	0	1	0	1	Cycle
0	1	0	1	0	Accuracy
1	1	0	0	0	Replaceability
0	1	0	1	0	Recall
0	1	0	0	0	Efficiency
1	0	0	1	0	Reconciliation
0	1	0	0	0	Robustness
0	1	0	0	0	Reliability
0	1	0	0	0	Quality_of_use
9	**9**	4	7	**4**	

Table 3. Ontology OT_2

s_c	c_t	s_e	l_a	s_x	**OT_2**
0	0	0	0	1	Richness
1	0	0	0	0	Adaptability
1	1	1	0	1	Testability
1	0	1	1	1	Maintenabiliy
1	0	1	0	0	Structure
1	1	0	1	0	Clarity
1	0	0	1	0	Cohesion
1	0	0	1	0	Completeness
0	0	1	0	1	Cycle
0	1	0	1	0	Precision
1	1	0	0	0	Replaceability
0	1	0	1	0	Recall
0	1	0	0	0	Efficiency
1	0	0	1	0	Reconciliation
0	1	0	0	0	History
0	1	0	0	0	Reliability
0	1	0	0	0	Maturity
0	1	0	0	0	Transferability
0	1	0	0	0	Maturity
0	0	0	0	1	Transferability
9	11	4	7	5	

Definition 2: An ontological layer $L_i \in L$ is relevant for an ontology $OT_k \in OT$, when it fulfils Inequaton (5). So, an ontology is relevant if each $L_i \in L$ is relevant for it.

4.2.2 Presentation of Algorithm 2

To solve the formulated problem, we propose an Algorithm (see Algorithm 2) for the Selection of Relevant Ontologies based on ontological Layers(ASROL). Our algorithm uses two phases: Phase 1, which is the initialisation phase, and phase 2, which is the phase of building set of relevant ontologies.

Algorithm 2: Algorithm for the Selection of Relevant Ontologies based on ontological Layers (ASROL)

Input
OT // Set of transport ontologies
C // Set of criteria for each OT_k in OT
α // The relevance coefficient , with $\alpha \in [0,1]$
L // Set of ontological layers

Set_Crit // Set_Crit = { Set_Crit_1 , Set_Crit_2 , Set_Crit_3 , Set_Crit_4, Set_Crit_5}
//Set_Crit_i: Set of criteria supported by layer $L_i \in L$
Output
OT^r// The set of relevant ontologies
1 **Begin**
2 $OT^r = \emptyset$ // Phase 1: Initialization Phase
3 **For each** $OT_k \in OT$ **Do** // Phase 2: Phase of building set // Begining of Step 1 of Phase 2
4 i= 1; is_relevant= True
5 **While** i ≤ |L| and is_relevant == True **Do**
6 $H(L_i) = 0$; $Count\_Crit(OT_k, L_i) = 0$
7 **For each** $C_j \in C$ **Do**
8 $Count_{Crit(OT_k,L_i)} = Count\_Crit(OT_k, L_i) + M(OT_k, L_i, C_j)$
9 **End for**
10 $H(L_i) = Ent\ (|Set\_Crit_i| * \alpha)$ // the largest integer less than or equal to $|Set\_Crit_i| * \alpha$ is retained
11 **if** $Count\_Crit(OT_k, L_i) < H(L_i)$ **Then**
12 is_relevant= $False$
13 **End if**
14 i=i+1
15 **End while** // End of step 1 of Phase 2
// Begining of Step 2 of Phase 2
16 **if** is_relevant == True **Then**
17 $OT^r.Add(OT_k)$
18 **End if** // End of Step 2 of Phase 2
19 **End for**
20 **End**

The Initialisation phase (or phase 1) consists to initialise the set of relevant ontologies to the empty set (refer to line 2).

After initialisation phase, ASROL performs the phase of building a set of relevant ontologies for each ontology belonging to the set of transport ontologies. This phase consists of two steps:

- Step 1 (refer to the instructions from line 4 to line 15): For each ontology $OT_k \in OT$, each ontological layer is processed. Indeed, ASROL computes the number of criteria of OT_k, which belongs to the ontological layer L_i according to Eqns. (2) and (3) as specified in line 8. Then, $H(L_i)$ takes the value of the integer part of the product of the cardinality of the set of criteria Set_Crit_i belonging to the ontological layers L_i by the relevance coefficient α (refer to line 10). Then, the flag named *is_relevant* is eventually updated (refer to the instructions from line 11 to line 13).

- Step 2 (refer to the instructions from line 18 to line 20): This step consists of updating the set of relevant ontologies eventually according to *is_relevant*.

The main advantage of algorithm 2 is that it is based on ontological layers, which relaxes the strict constraint on ideally relevant ontologies. But, this first level on relaxation is not always sufficient. In other words, the disadvantage of algorithm 2 is that for some instances of the problem studied here, it can return an empty set of relevant ontologies.

4.2.3 Proof of algorithm 2 total correctness

Our algorithm takes as input the set of ontologies *OT*, the set of criteria *C*, the set of sets of citeria belonging to each layer *Set_Crit* and the relevance coefficient α. Then, it returns a set of relevant ontologies OT^r. Note that OT^r can be a empty set or not. In the following paragraphs, we prove the total correctness of the proposed algorithm.

The instructions belonging to the loop "while" allow the computation of the set of criteria of each $OT_k \in OT$ belonging to each layer L_i (from line 5 to line 15) according to Eqns. (4) and (5), and to check if Inequation (5) is not fulfilled. Thus, each OT_k that is added to OT^r fulfils Inequation (5) for each ontological layer L_i. According to Definition 2, this means that OT^r contains all relevant ontologies (i) if Inequation (5) is fulfils for at least one ontology. Otherwise, OT^r is a empty set. From (i), it results the partial correctness of our algorithm (ii).

The set of ontologies OT and the set of criteria are two finite sets. Hence, the two loops "for" (refer to line 3 and line 7 of Algorithm 1) stop (iii). According to the instructions from line 5 to line 15, the loop "while" of Algorithm 1 stops (iv). From (iii) and (iv), it results that our algorithm stops (v).

From (ii) and (v), it results the total correctness of our algorithm.

4.3 Characterization of relevant ontology whose criteria are closest to the set of criteria

Algorithm 3: Algorithm for the Selection of relevant Ontologies based on Criteria closest to the set of criteria (ASROC_2)

```
     Input
     OT // Set of transport ontologies
     C // Set of criteria for each OT_k in OT
     Output
     OT^r; // The set of relevant ontologies
     Begin
1      OT^r = ∅ ; nc = |C|; // Phase 1: Initialization Phase
2      While OT^r is empty do // Phase 2: Phase of building set
       // Begining of step 1 of Phase 2
3        if nc == |C| then
4          for each OT_k ∈ OT do
5              if OT_k is_relevant according to each c ∈ C
               then
6                OT^r = OT^r U {OT_k}
7              end if
9          end for // End of step 1 of Phase 2
           // Begining of step 2 of Phase 2
10       Else
11         C_subsets = list of subsets of C having nc elements
12         is_selected = false
13         while not is_relevant and C_subsets is not empty
           do
14             subset = first element of C_subsets
15             for each OT_k ∈ OT do
16               if OT_k is_relevant according to each c ∈
                 subset then
17                 OT^r = OT^r U {OT_k};
18               end if
19             end for
20               if OT^r not is empty them
21                   is_selected = true
22               end if
23             C_subsets = C_subsets \ subset
24         end while
25       end if // End of step 2 of Phase 2
26       nc = nc – 1
27     End while // End of Phase 2
28     Return OT^r
29   End
```

According to Section 4.2.3, Algorithm 2 can provide a non-empty set of relevant ontologies for some instances of the formulated problem. Moreover, Definitions 1 and 2 indicate that the relevant ontologies returned by Algorithm 2 are not relevant according to the same number of criteria. some relevant ontologies may be relevant according to a small number of criteria characterizing the transport domain. In other words, some relevant ontologies may weakly represent the transport domain. To overcome these issues, we propose Algorithm 3. Algorithm 3 was designed such that all relevant ontologies returned by this algorithm are relevant according to the same number of transport domain criteria. Moreover, this number is as close as possible to the cardinality of the set of criteria C.

Algorithm 3 begins with the initialization of the set of relevant ontologies to the empty set and the number of criteria *nc* by which an relevant ontology will be relevant is equal to the cardinality of the set criteria in [11]: this is phase 1 (see line 1 of Algorithm 3). Then, phase 2, which is the construction of the set of relevant ontologies comprising the following steps:

-Step 1 (see lines 3 to 9): If *nc* is equal to the number of domain criteria, then each ontology that is relevant according to all domain criteria is added to the set of relevant ontologies.

-Step 2 (see lines 10 to 25): If the condition (see line 3) is not fulfilled, then the subsets of criteria of *nc* elements are used to check if one of them allows the building of a set of relevant ontologies such that the elements of this set are relevant according to *nc* criteria.

Phase 2 is repeated until the set of relevant ontologies becomes a non-empty set. The goal of Algorithm 3 is to obtain a non-empty set of relevant ontologies whose number of criteria is closer to the number of criteria of transport domain.

Each ontology is defined for a specific purpose in the transport domain. This implies that an ontology is relevant to at least one criterion of the transport domain. Therefore, the set of relevant ontologies returned by Algorithm 3 is always non-empty (i). Steps 1 and 2 of phase 2 guarantee that all relevant ontologies are relevant according to the same number of criteria *nc*. *nc* is initialized to |C| and is decremented iteratirely. This ensures that the number of criteria nc according to which ontologies are relevant is as close as possible to the cardinality of C denoted by |C| (ii).

From (i), it results that while loop which starts at line 2 stops (iii). C_subsets (refer to line 11) is a finite set and this set is gradually reduced. Thus, it becomes empty at a time. It results that while loop which starts at line 13 stops (iv). From (iii) and (iv), it results that Algorithm 3 stops (v).

From (i), (ii) and (v), it results the total correctness of

Algorithm 3 is proven.

Algorithm 3 returns a set of relevant ontologies as close as possible to the number of domain criteria. Thus, the main advantage of algorithm 3 is that it always returns a non-empty set of relevant ontologies. Algorithm 3 proceeds by a naive search to find the set of relevant ontologies. Therefore, it requires more computational resources when dealing with large inputs.

5. PERFORMANCE EVALUATION

We performed experiments to compare the effectiveness of methods. Python was used as programming language during the simulation. This simulation was performed on a computer equipped with a CPU, i7-2620M @ 2.70GHz RAM, and 128 GO SSD with a windows 10 operating system.

5.1 Performance indicators

The Goal of this study is to find a small non-empty set of relevant ontologies OT^{r} of the transport domain from a given set of transport domain ontologies OT. Therefore, two performance indicators were used:

- Ontological relevance rate (Ro_r) is measured as the ratio between the number of relevant ontologies and the total number of transport domain ontologies to be evaluated. It is represented by Eq. (6):

$$\mathrm{Ro}_r = \frac{|OT^r|}{|OT|} * 100 \quad (6)$$

Ro_r is an indicator used to check the ability of methods to provide a small (non-empty) set of relevant ontologies.

- Relevance rate of the criteria (Rc_r) which is the ratio between the number of criteria according to which ontologies are relevant and the total number of criteria of transport domain C. It is represented by Eq. (7):

$$\mathrm{Rc}_{\mathrm{r}} = \frac{nc}{|C|} * 100 \quad (7)$$

This indicator is used to measure the ability of algorithms to provide relevant ontologies that are relevant according to a large number of transport domain criteria.

5.2 Simulation setup

The Simulation aimed to study the efficiency of the algorithms according to the type of transport domain considered. Table 4 Presents the types of domains considered during simulation.

Table 4. Transport domain configurations

Criteria / Type of domain	Number of domain criteria |C|
Transport domain with low number of criteria (TDLC)	7
Medium Criteria Transport Domain (TDMC)	16
High Criteria Transport Domain (TDHC)	25

For each domain type d∈{TDLC,TDMC,TDHC}:

1) We select n (n = $|C|$) criteria from the 53 criteria in [11] to from the set of criteria characterising domain d;
2) Then we form 100 ontologies based on the criteria of d:
- For any ontology OT_i (1<= i <= 100) to be trained, a weight p (with p ∈ {0,1}) is uniformly chosen and assigned to each qualitative criterion $\mathrm{C_j}$ (1 <= j <= n) according to whether $\mathrm{C_j}$ is relevant or not for $\mathrm{OT_i}$;
3) Then, each of the algorithms (i.e, ASROC, ASROL and ASROC_2) is run based on of the 100 ontologies and the set of n criteria of domain d. Note that for ASROL, the choice of the relevance coefficient has been fixed experimentally to α = 0.2.
4) Steps 1 to 3 were executed 1000 times to stabilise the values of the performance indicators.
5) At the end of this sequence of repetitions, the average of each performance indicator is kept.

5.3 Results analysis

- Ontological relevance rate (Ro_r)

In this section, the results of the first performance indicator called ontological relevance rate are presented. This rate measures the ratio between the number of relevant ontologies and the total number of transport domain ontologies to be evaluated. Thus, if an algorithm produces a low (and non-zero) ontological relevance rate then this is equivalent to saying that this algorithm produces a small set of relevant ontologies. In other words, the lower the ontological relevance rate produced by an algorithm, the more efficient is the algorithm according to this indicator.

After evaluating the three algorithms (i.e., ASROC, ASROL and ASROC_2) in the three transport domains TDLC, TDMC and TDHC which are characterized by 07 criteria, 16 criteria and 25 criteria respectively, the results concerning the Ontological Relevance Rate are presented in Figure 1 and Table 5. In Table 5, average (AVG), Standard Deviation (SD), minimum value (m) and maximum value (M) of this indicator are provided for each algorithm. Figure 1 shows the average of this indicator for each algorithm.

Figure 1 shows that for TDLC (having the smallest number of criteria), ASROL provides a null value (because the orange colour is absent for TDLC) as ontological relevance rate, unlike ASROL and ASROC_2. This is justified by the fact that ASROL considered as relevant ontology, an ontology that is relevant according to all ontological layers. Furthermore, ontologies from TDLC only take into account 7 criteria at most, so it is impossible for ASROL to provide a non-zero number of relevant ontologies because the number of criteria in all ontological layers is much higher than 7. However, as the number of domain criteria increases, it becomes difficult for ASROC to find a relevant ontology according to all domain criteria. This is why ASROC has a null value for TDMC and TDHC domains. Thus, except ASROC_2 provides a non-zero value regardless the domain configuration and the instance tested (for the last row of Table 5 we have 1<= Ro_r<=5) because it considers as relevant ontology an ontology which is relevant according to a number close to the number of criteria of the domain. Furthermore, Table 5 confirms that ASROC_2 provided a small (non-empty) set of relevant ontologies.

- Relevance rate of the criteria (Rc_r)

Table 5. Ontological relevance rates Ro_r (%) caused by the three methods

Parameters / Methods	TDLC				TDMC				TDHC			
	AVG	**SD**	**m**	**M**	**AVG**	**SD**	**m**	**M**	**AVG**	**SD**	**m**	**M**
ASROC	1	1	0	5	0	0	0	1	0	0	0	0
ASROL	0	0	0	0	1	1	0	7	12	3	4	24
ASROC_2	1	1	1	5	1	0	1	2	1	0	1	1

Table 6. Relevance rate of the criteria Rc_r (%) caused by the three methods

Parameters / Methods	TDLC				TDMC				TDHC			
	AVG	**SD**	**m**	**M**	**AVG**	**SD**	**m**	**M**	**AVG**	**SD**	**m**	**M**
ASROC	56	50	0	100	0	5	0	100	0	0	0	0
ASROL	0	0	0	0	76	43	0	100	100	0	100	100
ASROC_2	94	7	85,71	100	81	5	66,67	100	76	4	65,29	91,30

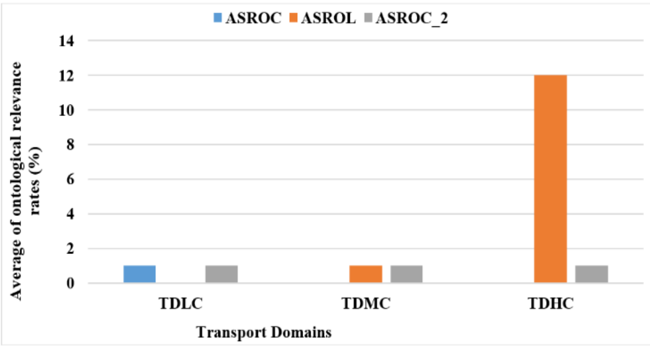

Figure 1. Comparison of ontological relevance rates (%)

The second indicator is the relevance rate of the criteria. This rate measures the ratio between the number of criteria according to which the ontologies are relevant and the total number of criteria of the domain considered. Ideally, a relevant ontology is a relevant ontology according to all the criteria of the domain. Thus, the higher the relevance rate of the criteria produced by an algorithm, the more efficient the algorithm is according to this indicator.

For the first indicator, we simulated the algorithms (ASROC, ASROL and ASROC_2) in the three configuration of transport domains TDLC, TDMC and TDHC. Recall that these domains are formed by 7 criteria, 16 criteria and 25 criteria respectively. In Table 6, the results of our simulations about the relevance rate of the criteria are given. For each algorithm, the average (AVG), standard deviation (SD), minimum value (m) and maximum value (M) of the relevance rate of the criteria are given. Figure 2 shows the average of this indicator for each algorithm.

Figure 2 shows that for the domains TDMC and TDHC, ASROC returns a zero value for the relevance rate of the criteria. This is not the case for ASROL and ASROC_2. This result can be explained by the fact that the ASROC algorithm returns as relevant ontologies, those ontologies where each criterion has a weight is equal to 1. However, TDMC (16 criteria) and TDHC (25 criteria) are domains with a large number of criteria.

Therefore, its is almost impossible for ASROC to find at least one ontology that is relevant according to 16 criteria or 25 criteria. It is obvious that the relevance rate of the criteria is zero if and only if there is no relevant ontology. It results that ASROC returns a zero relevance rate for TDMC and TDHC.

In short, it should be noted that:

-With the ontology relevance rate, regardless of the TDLC, TDMC, and TDHC domain, the ASROC_2 algorithm always finds a small (but not zero) number of relevant ontologies. This is not the case for ASROC and ASROL.

-With the relevance rate of the criteria, whatever the type of domain configuration (TDLC, TDMC and TDHC), the ASROC_2 algorithm always finds a small number of relevant ontologies which are relevant according to a high number of criteria.

These two remarks confirm that the ASROC_2 algorithm solves the studied problem more efficiently.

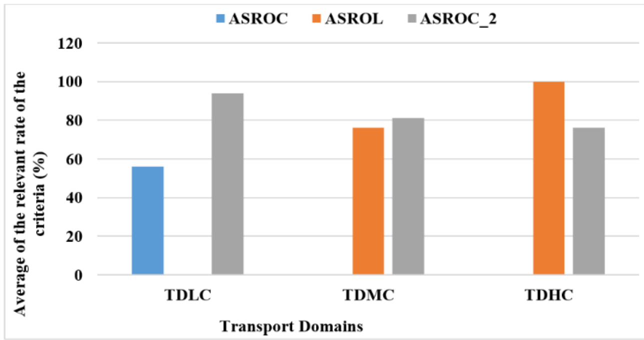

Figure 2. Comparison of the relevance rates of the criteria

6. CONCLUSION AND FUTURE WORK

The problem studied in this paper is to find a small set of relevant ontologies for a set of input transport ontologies [29], relevance is valuated according to one criterion at a time. However, ontologies depend on several criteria. To overcome this, we propose three methods: The first method, ASROC, is an extension of the algorithm [29] that allows to find a set of ontologies relevant to all domain criteria, which is not always possible. The second method, ASROL finds a set of relevant ontologies based on the criteria and ontological layers. Finally, the third method ASROC_2 allows finding of a set of relevant ontologies for many criteria close to domain criteria. Simulation confirms that ASROC_2 is more efficient to solve the studied problem.

Therefore, ASROC_2 is an essential tool that can be used by ontology developers to evaluate several ontologies in the decision process. As a future work, it will be interesting to propose a similarity method between concepts of relevant ontologies to obtain a single ontology of the transport domain.

ACKNOWLEDGMENT

This work was possible thanks to the contribution of the members of the Laboratory of Research in Computer Science and Telecommunications (LARIT) of the National Polytechnic Institute Félix Houphouët-Boigny of Yamoussoukro (INP-HB). Without forgetting the unfailing support of Dr Ferdinand A. Atta and the directors of this Laboratory.

REFERENCES

[1] Komninos, N., Kakderi, C., Panori, A., Tsarchopoulos, P. (2019). Smart city planning from an evolutionary perspective. Journal of Urban Technology, 26(2): 3-20. https://doi.org/10.1080/10630732.2018.1485368

[2] Cocchia, A. (2014). Smart and Digital City: A Systematic Literature Review. In: Dameri, R., Rosenthal-Sabroux, C. (eds) Smart City. Progress in IS. Springer, Cham. https://doi.org/10.1007/978-3-319-06160-3_2

[3] Caselli, B., Pellicelli, G., Rossetti, S., Zazzi, M. (2022). How are medium-sized cities implementing their smart city governance? Experiences from the Emilia-Romagna Region. Sustainability, 14(22): 15300. https://doi.org/10.3390/su142215300

[4] Li, Y., Jianhui, Z., Liu, J., Hou, Y. (2020). Matching large scale ontologies based on filter and verification. Mathematical Problems in Engineering, 2020: 1-10. https://doi.org/10.1155/2020/8107968

[5] Yoo, D., No, S., Ra, M. (2014). A practical military ontology construction for the intelligent army tactical command information system. International Journal of Computers Communications & Control, 9(1): 93-100.

[6] Monnin, A., Livet, P. (2014). Distinguer/Expliciter. L'ontologie du Web comme ontologie" d'opérations". Intellectica-La revue de l'Association pour la Recherche sur les sciences de la Cognition (ARCo), 1(61): 59-104. https://dx.doi.org/10.3406/intel.2014.1039

[7] Du Chateau, S., Mercier-Laurent, E., Bricault, L., Boulanger, D. (2020). Modélisation des connaissances et technologies du Web sémantique: deux applications au patrimoine culturel. Humanités Numériques, (2). https://doi.org/10.4000/revuehn.510

[8] Jindal, R., Seeja, K.R., Jain, S. (2020). Construction of domain ontology utilizing formal concept analysis and social media analytics. International Journal of Cognitive Computing in Engineering, 1: 62-69. https://doi.org/10.1016/j.ijcce.2020.11.003

[9] Raad, J., Cruz, C. (2015). A survey on ontology evaluation methods. In Proceedings of the International Conference on Knowledge Engineering and Ontology Development, part of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. https://dx.doi.org/10.5220/0005591001790186

[10] Algergawy, A., Babalou, S., Klan, F., König-Ries, B. (2020). Ontology modularization with OAPT. Journal on Data Semantics, 9: 53-83. https://doi.org/10.1007/s13740-020-00114-7

[11] Hooi, Y.K., Hassan, M.F., Shariff, A.M. (2015). Ontology evaluation—A criteria selection framework. In 2015 International Symposium on Mathematical Sciences and Computing Research (iSMSC), Ipoh, Malaysia, pp. 298-303. https://doi.org/10.1109/ISMSC.2015.7594069

[12] Hanelt, A., Bohnsack, R., Marz, D., Antunes Marante, C. (2021). A systematic review of the literature on digital transformation: Insights and implications for strategy and organizational change. Journal of Management Studies, 58(5): 1159-1197. https://doi.org/10.1111/joms.12639

[13] Avrunin, O., Kolisnyk, K., Nosova, Y., Tomashevskyi, R., Shushliapina, N. (2020). Improving the methods for visualization of middle ear pathologies based on telemedicine services in remote treatment. In 2020 IEEE KhPI Week on Advanced Technology (KhPIWeek), Kharkiv, Ukraine, pp. 347-350. https://doi.org/10.1109/KhPIWeek51551.2020.9250090

[14] Stevens, R., Lord, P., Malone, J., Matentzoglu, N. (2019). Measuring expert performance at manually classifying domain entities under upper ontology classes. Journal of Web Semantics, 57: 100469. https://doi.org/10.1016/j.websem.2018.08.004

[15] Jarosław, W. (2018). An attempt to knowledge conceptualization of methods and tools supporting ontology evaluation process. Procedia Computer Science, 126: 2238-2247. https://doi.org/10.1016/j.procs.2018.07.225

[16] Fonou-Dombeu, J.V. (2019). Ranking semantic web ontologies with ELECTRE. In 2019 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD), Winterton, South Africa, pp. 1-6. https://doi.org/10.1109/ICABCD.2019.8851051

[17] Lee, Y.H., Hu, P.J.H., Tsao, W.J., Li, L. (2021). Use of a domain-specific ontology to support automated document categorization at the concept level: Method development and evaluation. Expert Systems with Applications, 174: 114681. https://doi.org/10.1016/j.eswa.2021.114681

[18] Trojahn, C., Vieira, R., Schmidt, D., Pease, A., Guizzardi, G. (2022). Foundational ontologies meet ontology matching: A survey. Semantic Web, 13(4): 685-704. https://doi.org/10.3233/SW-210447

[19] Mohammadi, M., Rezaei, J. (2020). Evaluating and comparing ontology alignment systems: An MCDM approach. Journal of Web Semantics, 64: 100592. https://doi.org/10.1016/j.websem.2020.100592

[20] Cardoso, S.D., Pruski, C., Da Silveira, M. (2018). Supporting biomedical ontology evolution by identifying outdated concepts and the required type of change. Journal of Biomedical Informatics, 87: 1-11. https://doi.org/10.1016/j.jbi.2018.08.013

[21] Koonce, D., Sormaz, D. (2019). Metrics to gauge the success of a manufacturing ontology. Procedia Manufacturing, 38: 1678-1682. https://doi.org/10.1016/j.promfg.2020.01.116

[22] Mohammadi, M. (2019). Bayesian evaluation and comparison of ontology alignment systems. IEEE Access, 7: 55035-55049. https://doi.org/10.1109/ACCESS.2019.2903861

[23] Gangemi, A., Catenacci, C., Ciaramita, M., Lehmann, J. (2006). Modelling ontology evaluation and validation. In: Sure, Y., Domingue, J. (eds) The Semantic Web: Research and Applications. ESWC 2006. Lecture Notes in Computer Science, vol. 4011. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11762256_13

[24] Tartir, S., Arpinar, I.B., Moore, M., Sheth, A.P., Aleman-

Meza, B. (2005). OntoQA: Metric-based ontology quality analysis. IEEE ICDM Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources, Houston, TX, November 27, 2005.

[25] Mishra, S., Jain, S. (2020). Ontologies as a semantic model in IoT. International Journal of Computers and Applications, 42(3): 233-243. https://doi.org/10.1080/1206212X.2018.1504461

[26] Lozano-Tello, A., Gómez-Pérez, A. (2004). Ontometric: A method to choose the appropriate ontology. Journal of Database Management (JDM), 15(2): 1-18. https://doi.org/10.4018/jdm.2004040101

[27] Katsumi, M., Fox, M. (2018). Ontologies for transportation research: A survey. Transportation Research Part C: Emerging Technologies, 89: 53-82. https://doi.org/10.1016/j.trc.2018.01.023

[28] Louge, T., Karray, M.H., Archimède, B., Maamar, Z., Mrissa, M. (2019). Semantic web services composition in the astrophysics domain: Issues and solutions. Future Generation Computer Systems, 90: 185-197. https://doi.org/10.1016/j.future.2018.07.063

[29] Pramanik, P.K.D., Biswas, S., Pal, S., Marinković, D., Choudhury, P. (2021). A comparative analysis of multi-criteria decision-making methods for resource selection in mobile crowd computing. Symmetry, 13(9): 1713. https://doi.org/10.3390/sym13091713