The Combination of Ontology-Driven Conceptual Modeling and Ontology Matching for Building Domain Ontologies: E-Government Case Study

Shaimaa Haridy*, Rasha M. Ismail*, Nagwa Badr*, and Mohamed Hashem*
Ain Shams University, Cairo, Egypt

Abstract

During the last decade, ontology engineering has undoubtedly participated in a lot of beneficial applications in different domains. Nevertheless, ontology development still faces several significant challenges that need to be addressed. This study proposes an enhanced architecture for the ontology development lifecycle. With the help of this architecture, users can complete ontology development tasks since it provides guidance for all key activities, from requirement specification to ontology evaluation. Ontology-driven conceptual modeling (ODCM) and ontology matching serve as the foundation of this architecture. ODCM is defined as the application of ontological ideas from various fields to build engineering objects that improve conceptual modeling. Ontology matching is a promising approach to overcome the semantic heterogeneity challenge between different ontologies. The proposed architecture is applied to e-governance domain, which is one of the online services that gains a great attention worldwide, especially during the coronavirus pandemic. However, residents of Arab countries face numerous obstacles and do not receive the full benefits of e-governance. For these reasons, Egyptian e-government is selected as the suggested case study. The results are encouraging when the produced ontology is compared with 20 existing ontologies from the same domain. On the basis of OntoMetrics, the average values of metrics correlated to accuracy, understandability, cohesion and conciseness lie in the 95th, 95th, 95th and 57th percentiles respectively. The results can be further enhanced by defining more non-inheritance relations and distributing the instances across all classes.

Key Words: Artificial intelligence, digital government (e-government), ontology-driven conceptual modeling, ontology engineering, ontology enrichment, ontology matching, OntoUML, semantic web.

1 Introduction

In recent years the enormous growth of semantic web theories and techniques facilitates using ontologies extensively in numerous domains and applications [2]. Authors in [16] defined ontology engineering as "The set of activities that concern the ontology development process, the ontology life cycle, and the methodologies, tools and languages for building ontologies". It seeks to provide standard components for creating knowledge models. Conceptualization is one of the crucial activities in ontology engineering. Conceptualization focuses on recognizing the concepts in the real world to build the model of the relevant domain [40]. This activity has a significant impact on the quality of the final ontology, as the quality of any artifact based on a model is constrained by the model's quality [21]. Researchers proposed a new method known as ontology-driven conceptual modeling (ODCM) [13] that greatly aid in ontology conceptualization. ODCM is described as the application of ontological ideas from various fields, such as formal ontology, cognitive science, and philosophical logics, to build engineering objects that improve conceptual modeling theory and practice. One of the most used languages in ODCM is OntoUML [13], which is "a language whose meta-model has been designed to comply with the ontological distinctions and axiomatization of a theoretically well-grounded foundational ontology named UFO (Unified Foundational Ontology)" [20]. UFO is "an axiomatic formal theory based on contributions from Formal Ontology in Philosophy, Philosophical Logics, Cognitive Psychology, and Linguistics" [22].

Ontology engineering has undoubtedly participated in a lot of beneficial applications in the last decade, but there are still significant challenges about ontology development that need to be solved [34]. One of the challenges that face ontology-based applications is ontology heterogeneity which emerges from varying expertise of knowledge engineers who create and maintain ontologies in the same domain. For that reason, ontology matching is widely used in ontology engineering to solve this heterogeneity problem [25].

The digital government or electronic government (e-government) is one of the critical domains that could benefit from ontology engineering. It provides public services to citizens solely by means of information and communication technologies, such as computers and Internet [45]. However, the residents of Arab countries have faced numerous obstacles and have not received the full benefits of e-governance. Egypt's Ministry of Communications and Information Technology, in collaboration with the Ministry

* Faculty of Computer and Information Sciences. Email: shaimaaharidy@cis.asu.edu.eg, rashaismail@cis.asu.edu.eg, nagwabadr@cis.asu.edu.eg, mhashem100@yahoo.com.

ISCA Copyright© 2022

of State for Administrative Development, launched an e-government project. The first stage lasted from 2001 to 2007, and the second stage was from 2007 to 2012 [18]. Nevertheless, in 2020, the United Nations ranked Egypt as 111th out of 193 nations in terms of using e-government to provide public services; and the ninth in the Arab world [41]. This is due to the numerous obstacles faced in the growth of e-government [18]. Some of these challenges are:

1) There are no standards or specifications, which makes it difficult for government entities to communicate and integrate.
2) The government agencies are not able to share information that prohibits them from performing e-government efforts properly and effectively.
3) There is no unified standard for repeated inquiries made by citizens to the various government agencies.

Consequently, the goal of this research is to present an enhanced ontology engineering architecture for building domain ontologies from scratch. ODCM and ontology matching serve as the foundation of the proposed architecture. The e-government domain in Egypt is selected as the case study. The three main contributions of this study are: (1) the combination of ODCM conceptual modeling with ontology matching, (2) two new algorithms for the selection of the most relevant ontology and the extraction of new classes and relations, as well as (3) the new domain ontology for the Egyptian e-government. The remainder of the paper is organized as follows. Section 2 gives a review of the related work. A description of the proposed architecture is offered in Section 3. Section 4 discusses the results of the experiments. Finally, Section 5 states the conclusions and future work.

2 Related Work

There has been a tremendous amount of work proposed in the literature to cover various ontology-related aspects and their application in diverse disciplines. Some recent studies related to ontology-driven conceptual modeling, ontology matching, and e-government will be discussed in the following subsections.

2.1 Ontology-Driven Conceptual Modeling

The papers in this context are divided into two categories. OntoUML was used by the first group due of its widespread use in crucial and complex sectors, such as [2, 8-10, 17, 22, 36]. Whereas the second used alternative languages, like [1, 5, 26, 35, 39]. Table 1 provides a summary of the work done by these prior studies. Despite OntoUML's success in other domains, it has not yet been applied in e-government.

A. 2.2 Ontology Matching

The ontology matching task has been discussed for a decade. For example, authors in [42] acquired knowledge from groupware using facts enrichment approach (FEA). The advantage of this approach resides in its capacity to extract new concepts from unstructured text and insert those new concepts into an already-existing ontology. In [4], researchers improved the ontology matching task by merging several alignments generated by different matchers. Their approach used background knowledge (BK) resources to drive paths between source and target ontologies. Finally, they proposed a selection algorithm in order to determine the final mapping. In [23] used ontology matching to improve emotion detection from text. The ontology of the input statement was matched with emotion labeled ontology base, and the emotion with the highest matching score was selected. Matching ontologies algorithm were developed in [3]. The goal of this algorithm was to match entities and sub-entities in the Ontology document with sentences that have the same topic. Ontology matching and ODCM each have advantages that have been discussed in the literature, but little has been written on how to combine the two to promote ontology engineering.

2.3 E-Government

For years, many efforts have been undertaken to overcome the barriers related to e-governance. Authors of [6] addressed some variables affecting the adoption of cloud computing in government organizations, because Saudi Arabian government recognized the benefits of cloud computing and tried to create basic protocols for delivering government services via the cloud. As for [18] and [44], both conducted research on Egypt's e-government and its challenges. In addition, they proposed remedies for these problems. In [18] the researchers achieved some progress, but more needs to be done to address the problems faced in the development of e-government. Moreover, Egypt is still behind a few Arab countries in delivering online information and providing government services. In [44], the authors stated that the Egyptian information technology sector was impacted by the political turmoil in recent years. Egypt's e-government development index was downgraded from high to medium in the global rankings. Its position plummeted 28 places from the 80th place in 2014 to the 108th place in 2016. In [27], the researchers demonstrated how semantic technology is critical for the development of e-government services. Semantic web services and ontologies allow the automated processing of services and information and improve communication between the parties involved. They examined a set of available information, standards, existing referenced models and certain semantic web service formalisms. In addition, a knowledge-based modeling framework for Moroccan e-government services and its implementations of e-customs was suggested.

Since literature review reveals that Egypt's e-government area lacks an ontology, it is chosen as the suggested case study in this paper.

Table 1: Summary of ODCM utilization in various domains

Reference	Language / Tool	Conceptual Model	Ontology	Detailed Steps	Detailed Evaluation	Literature Review	Domain
[17]	OntoUML	✓		✓		✓	Human Genome
[8]	OntoUML	✓	✓	✓	✓	✓	Railway
[33]	OntoUML	✓		✓		✓	Economic Exchanges
[10]	OntoUML		✓	✓	✓		Marketplace
[9]	OntoUML	✓	✓	✓			Marketplace
[35]	-	✓	✓	✓			Farm
[26]	CoreWEB	✓					Enterprise Information Systems
[39]	Protégé	✓	✓			✓	Cybersecurity Vulnerability
[2]	Menthor & E-OntoUML & StarUML	✓	✓	✓		✓	Agriculture
[5]	-	✓	✓	✓			Transmedia Storytelling
[1]	Protégé	✓	✓	✓	✓	✓	Human Affective States
[36]	OntoUML	✓	✓	✓		✓	Higher Education

3 Proposed Architecture

This study introduces an enhanced architecture for the ontology development lifecycle. With the help of this architecture, users can complete ontology development tasks since it provides guidance for all key activities, from requirement specification to ontology evaluation. The domain of e-governance in Egypt is the suggested case study. Figure 1 depicts the proposed architecture, which is composed of four main modules that are thoroughly detailed in the following subsections.

3.1 Requirement Specification Module

This module seeks to provide the ontology's prospective uses. This can be accomplished by reviewing available domain documents in their various formats, in addition to related online resources. The output of this module is the use case diagram. UML (Unified Modeling Language) design tools are helpful for this. The diagram helps in the comprehension of system functionalities, the elimination of function redundancies, and the review of relationships between different actors and system functions.

In the proposed case study, due to the shortage of documents describing the Egyptian e-government in detail, the Egyptian web portal [11] serves as the lone source of information to establish system requirements. The portal connects 14 consumers with 30 service providers. The citizen is the main consumer who uses the greatest number of services through the portal. This paper focuses on the citizen module, which includes 35 different services. Their use cases are presented using the Rational Rose tool [7].

3.2 Ontology Development Module

Although there is no universal consensus on the technique that can be considered ideal for ontology development, the goal of creating an ontology might help select the optimal methodology. METHONTOLOGY is an ontology-building methodology that specifies the life cycle of the ontology development process. The full explanation of the conceptualization activity, which is the focus of this work, is the key strength of this methodology [15]. Therefore, it is followed in the ontology development module in the proposed architecture. This module accepts as input the use case diagram generated from the preceding module, as well as domain documents, online resources, and existing ontologies. As for output, a new domain ontology is produced. Subsequently subsections provide description of ontology development activities.

A written document in natural language is used to describe the ontology information in this activity. A common proposal for describing ontologies is the ontology metadata vocabulary (OMV). It allows for easy access to and exchange of ontologies across the Internet [24]. The OMV of the suggested ontology is given above.

The NeOn methodology [37] provides versatile possibilities for the reuse and re-engineering of knowledge sources for developing ontology networks. The following three scenarios are used by proposed ontology:

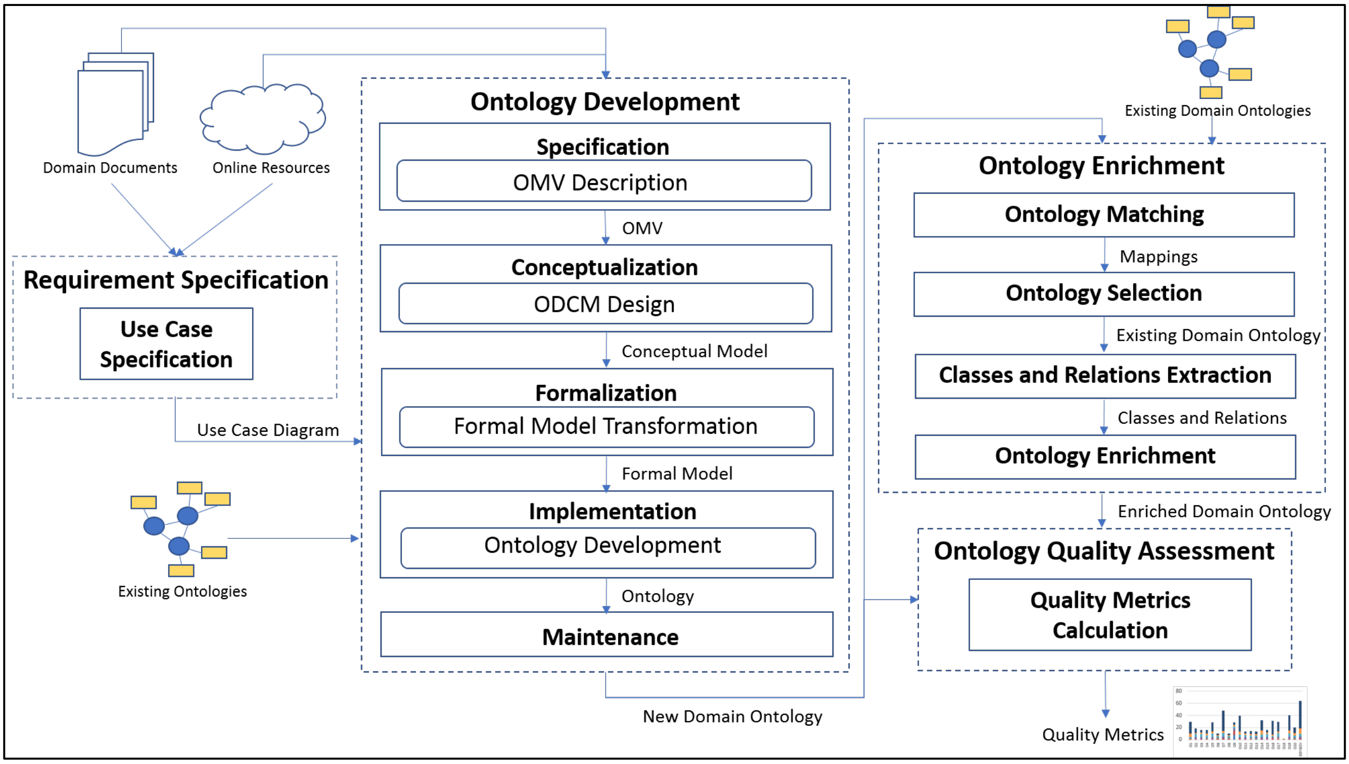

Figure 1: Proposed architecture

3.2.1 Specification

Egyptian E-government Ontology Metadata Vocabulary OMV

Ontology Name: Egyptian E-government Ontology (EGYGOV)

Location: Ain Shams University, Cairo, Egypt

Party (Organization): Ain Shams University

License Model: Academic research

Ontology Type: Domain Ontology

Ontology Domain: Electronic government (e-government)

Ontology Engineering Tool: OntoUML Lightweight Editor (OLED)

Ontology Language: OWL

Ontology Syntax: rdf xml Syntax

Ontology Task: Describes data and services of Egyptian e-government. EGYGOV represents a model for the semantic description of governmental features such as domain concepts, services, regulations, and organizational structures.

Ontology Engineering Methodology: NeOn Methodology for Building Ontology Networks; EGYGOV follows different scenarios:

Scenario 1: From specification to implementation.

Scenario 2: Reusing and re-engineering non-ontological resources.

Scenario 3: Reusing and re-engineering ontological resources.

Source of Knowledge: Non ontological resources (Egyptian e-government portal), Ontological resources (existing ontologies)

- Scenario 1: From specification to implementation. The ontology is developed from scratch.
- Scenario 2: Reusing and re-engineering non ontological resources. Egyptian e-government portal [11] is the non-ontological resource used to build the ontology. Re-engineering entails transferring this resource to an ontology format and possibly modifying the class name.
- Scenario 3: Reusing and re-engineering ontological resources. EGYGOV reused UFO-S [31], which is a core reference ontology that captures a clear account of services and service-related concepts.

3.2.2 Conceptualization. A model of the relevant domain knowledge is built in this step. This model can take any shape that domain experts accept and understand [15]. The proposed conceptual model is implemented using OLED [19], which is a model-based environment for formalizing, implementing, testing, and validating OntoUML models. Class and relationship stereotypes of OntoUML are described in depth in [38].

Two alternative portions of the designed conceptual model are shown in Figures 2 and 3. In Figure 2, the *Service Delivery* class consists of *Request* of the consumer, *Response* of the provider, *Description, Conditions* to be accepted, and *Actions* to be performed; as well as *Documents* to be extracted. For *Service Delivery*, two options are possible: *Free* or *Paid*. The *Request* has three possibilities: *Pending*, *Processed*, or *Cancelled*. The *Agent* category includes *Provider* and *Consumer*; *Provider* has many *Entities* and each one has a *Location*. Note that *Service*, *Agent*, *Provider*, and *Consumer* are reused and reengineered from the UFO-S core ontology. Figure 3 displays a portion of the *Citizen* class and *Violations* on his *Vehicle License* and *Driving License*. Both licenses are generalized from the *License* class. *Phone Line* has *Mobile Line* and *Land Line* as its descendants. Also, the citizen makes a *Phone Subscription* to his *Phone Line* which is associated with a *Phone Bill, a*s well as, the *Electricity Subscription* made to an *Electricity Meter* with its *Consumption Readout* and *Complaints* and *Inquiries* made by citizens.

3.2.3 Formalization. The aim of this activity is to output a model in an implementation language. Therefore, the preceding activity's well-founded conceptual model is transformed into a formal model using OLED code generation feature. As a result, the proposed model is converted from OntoUML model to OWL ontology.

3.2.4 Implementation. Using the protégé tool [309] the resulting OWL ontology is enhanced with data properties and individuals (instances), as displayed in Figures 4 and 5, respectively. For example, properties such as IDs, Texts are added to Condition, Complaint and Area classes. Dates are added to Birth Certificate, Document and Consumption Readout classes. Description is added to Service class.

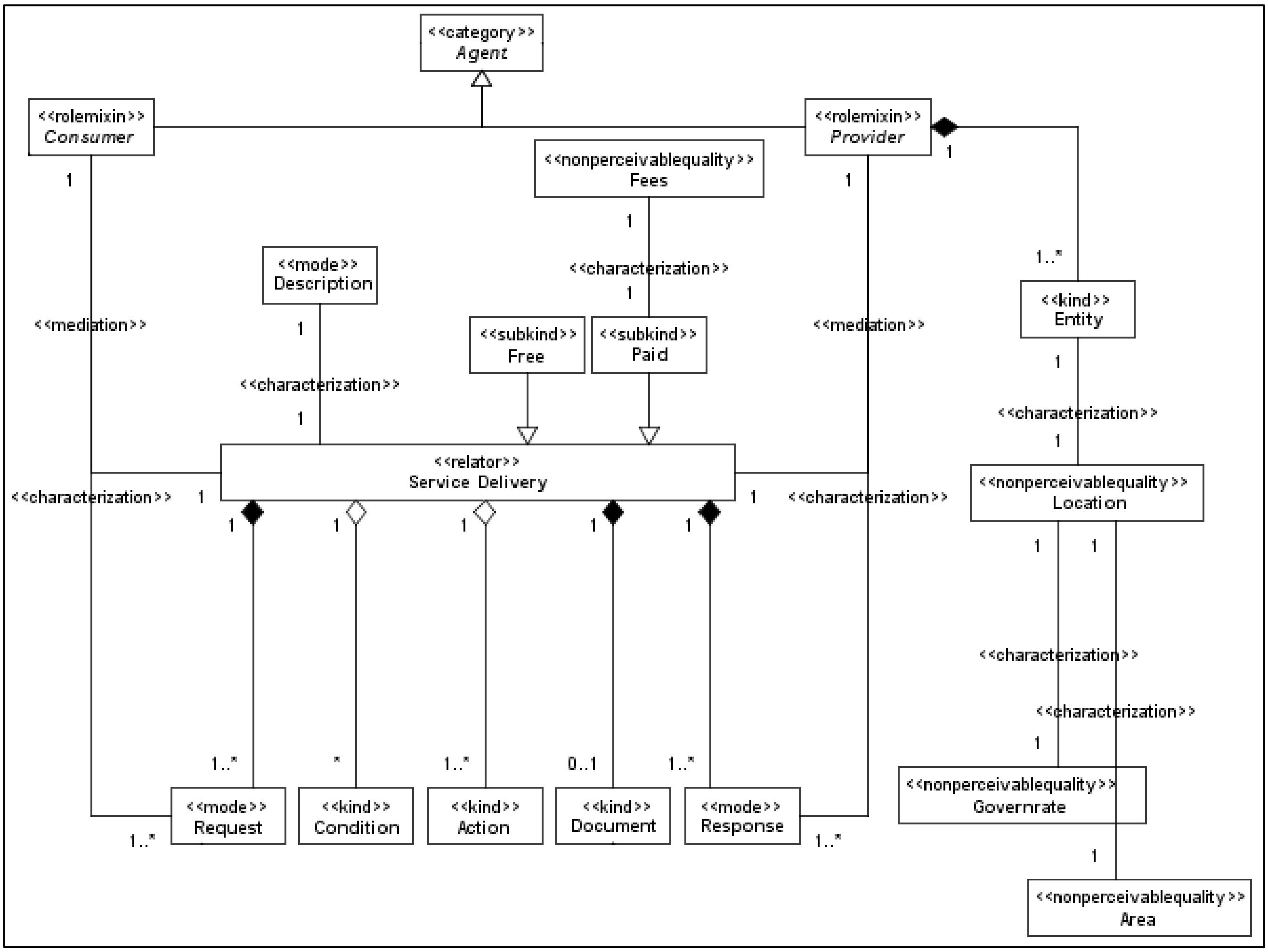

Figure 2: A fragment of the proposed conceptual model: The service delivery class

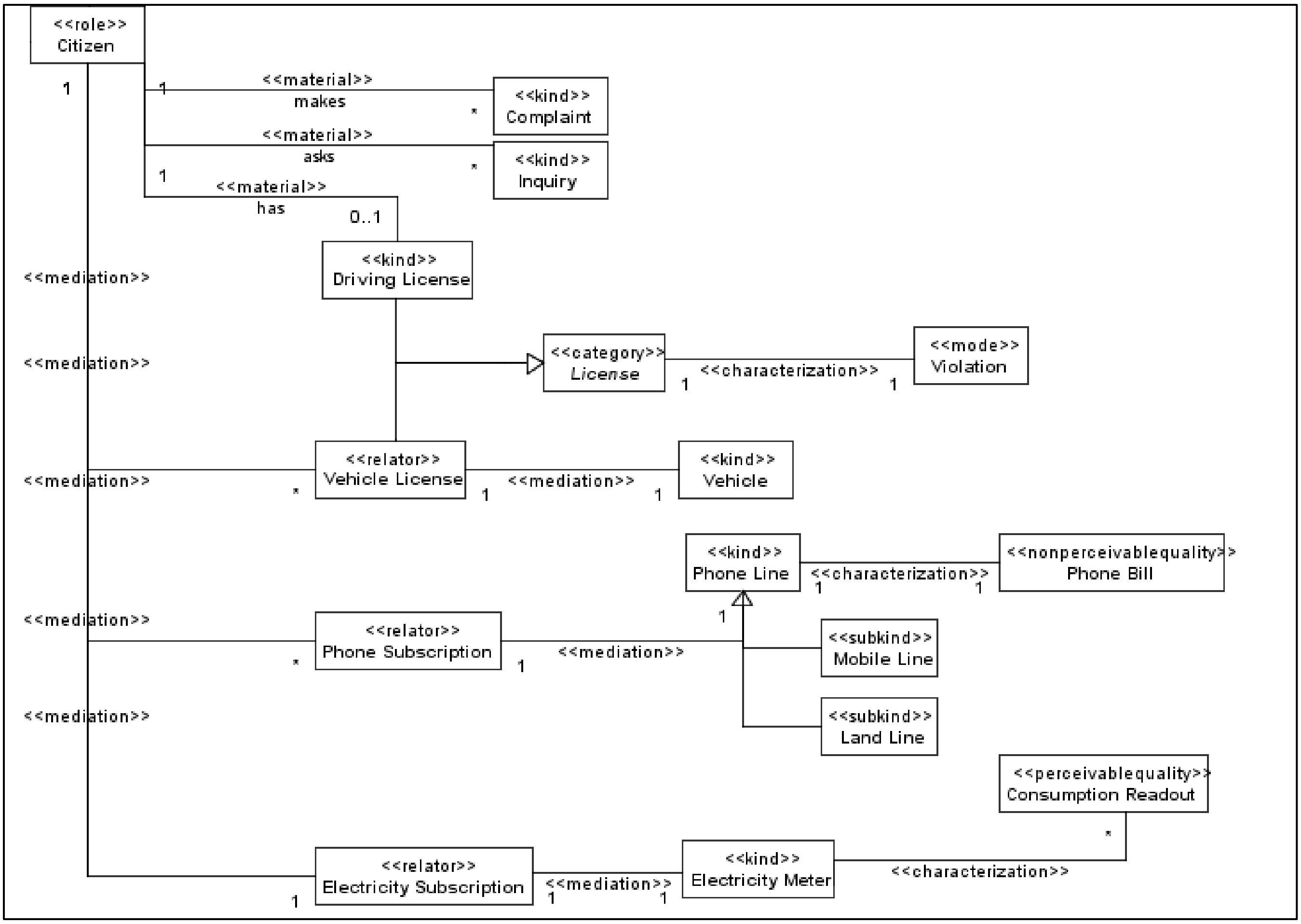

Figure 3: A fragment of the proposed conceptual model: The citizen class *documents*

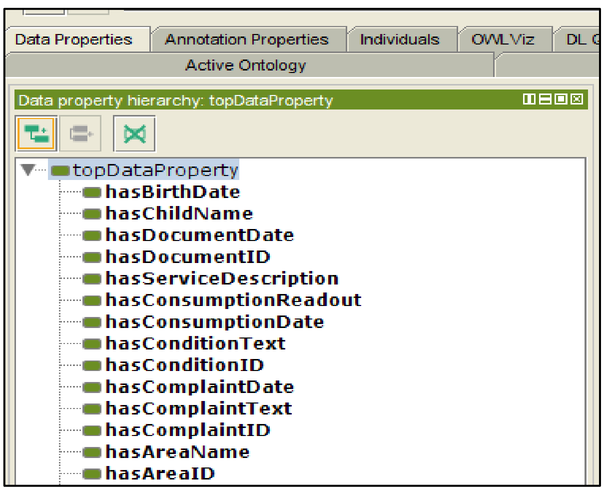

Figure 4: A fragment of data properties

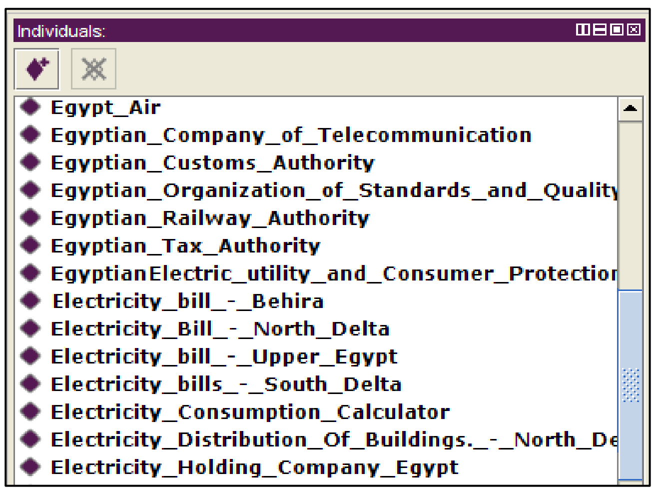

Figure 5: A fragment of individuals

Individuals are defined to classes such as *Ministry*, *Company, Authority, Office*, *Organization, Governorate* and *Request*.

3.2.5 Maintenance. This activity involves making any necessary updates or corrections to the ontology.

3.3 Ontology Enrichment Module

The ontology enrichment seeks to evolve their semantic content in order to cover new knowledge and enhance their semantic consistency 12]. In the current module, this is accomplished by finding new concepts and relations, then inserting them into the ontology. Although this new knowledge can come from a variety of sources, the existing domain ontologies are the main focus of this article. This enables the integration of domain knowledge from different perspectives. This module receives as input a set of existing domain ontologies in addition to the ontology created in the previous module. The output is an enriched domain ontology. This module consists of four activities described below. The paper proposes two novel algorithms. Algorithm 1 for the selection of the most relevant domain ontology, and Algorithm 2 for the extraction of relevant classes and relations.

3.3.1 Ontology Matching. This activity identifies correspondences between proposed ontology's entities and those in already existing domain ontologies. This can be accomplished by using one of the available matchers. AML [32] is one of the most effective matching systems. It has the advantage of the data structures generated for its word matcher. When an ontology is loaded, it builds a new lexicon with all class labels and synonyms using a bag-of-words technique. Additionally, a relationship map [13] data structure is created, which connects each class to the classes related to it via a part of relationships or disjoint clauses. In the domain of e-government, data of 20 existing ontologies were collected in the paper [14]. Unfortunately, only ten of them are accessible for download. So, in this case study AML matcher is used along with some customized code to match the proposed ontology with those ten ontologies. Therefore, results of their generated data structures (lexicon and relationship map) are displayed in Table 2. The output of this activity is mappings between the proposed ontology (EGYGOV) and each of the ten domain ontologies.

3.3.2 Ontology Selection. Based on the mappings produced by the previous activity, the most pertinent domain ontology is selected. one with the greatest number of mappings to the proposed ontology (EGYGOV). It is O_{18} as shown in Table 3.

3.3.3 Classes and Relations Extraction. The aim of this activity is to extract the list of classes and relations that will be injected later in the proposed ontology. To do this, there are two key steps: first, find the classes that exist in the mappings of the most relevant ontology and then look for the relationships of each class within the domain ontology itself. In this case the output list contains 748 classes and 1456 relationships.

3.3.4 Ontology Enrichment. This activity is the last step in the enriching process in which classes and relationships are inserted into the proposed ontology. To eliminate duplications or inappropriate insertions, the extracted list must first undergo a comprehensive editing. The list is thereby condensed to 703 classes and 830 relationships. They are injected in the proposed ontology using the protégé tool [30], then data properties and individuals for those new classes need to be defined.

Algorithm 1 - The Most Relevant Domain Ontology Selection

```
INPUT: Proposed ontology (proposedonto)

INPUT: List of existing domain ontologies
(domainontolist)

OUTPUT: The most relevant domain ontology and
its mappings with the proposed ontology

BEGIN
  SET maxmappcnt TO 0
  //ONTOLOGY MATCHING
  LOAD proposedonto INTO sourceonto
  FOR EACH onto IN domainontolist DO
    LOAD onto INTO targetonto
    CONSTRUCT ONTOLOGY OBJECTS
    BUILD LEXICON
    BUILD RELATIONSHIP MAP
    mappings ← ALIGN ONTOLOGY OBJECTS
    //ONTOLOGY SELECTION
    IF mappings.count > maxmappcnt THEN
      SET maxmappcnt To mappings.count
      SET mostrelvonto TO targetonto
      SET mostrelvmapps TO mappings
    END IF
  END FOR
  RETURN (mostrelvonto, mostrelvmapps)
END
```

```
Algorithm 2 – The Relevant Classes and Relations Extraction

INPUT: The most relevant domain ontology (mostrelvonto)

INPUT: The mappings between the proposed ontology with the most relevant domain ontology (mostrelvmapps)

OUTPUT: List of relevant classes and relations

BEGIN

    FOR EACH mapp IN mostrelvmapps DO
      class1id ← mapp.TARGETCLASSID
      class1name ← mostrelvonto.GETCLASSNAME(class1id)
      rels ← mostrelvonto.GETRELATIONS(class1id)
      FOR EACH r IN rels DO
        relname ← r.GETRELNAME()
        class2id ← r.GETCLASS2ID()
        class2name ← mostrelvonto.GETCLASSNAME(class2id)
        outputlist.ADD (class1id, class1name, relname, class2id, class2name)
      END FOR
    END FOR
    RETURN (outputlist)
END
```

3.4 Ontology Quality Assessment Module.

In this module, the proposed ontology is evaluated. This assessment process is applied twice, once following the development module and once following the enrichment module. The ontology's quality can be assessed in multiple ways. In this case study, to overcome the absence of gold standard ontology in the e-government domain, the OntoMetrics quantitative measures are applied. They were utilized in [14] and the evaluation results of 20 ontologies from the e-government area were documented. This makes it easier to assess and compare the outcome of the suggested architecture. The calculations of these metrics and their correlation to ontology assessment dimensions are displayed in Table 4 and Table 5, respectively.

Table 2: AML matcher results – lexicon and relationship map

Ontology	Classes	Labels	Properties	Direct is-a Relations	Disjoint Clauses
O_4	18	24	21	37	12
O_{11}	11	11	10	18	0
O_{12}	8	8	10	12	0
O_{13}	8	8	12	12	0
O_{14}	126	126	4	248	0
O_{15}	17	17	13	33	0
O_{16}	50	150	77	107	3
O_{17}	98	97	118	-	-
O_{18}	2509	13809	129	5637	0
O_{19}	209	458	198	327	0
EGYGOV	**125**	**122**	**95**	**200**	**39**

Table 3: AML matcher results – mappings

Ontology	No of Mappings
O_4	**0**
O_{11}	**8**
O_{12}	**1**
O_{13}	**1**
O_{14}	**4**
O_{15}	**4**
O_{16}	**33**
O_{17}	**21**
O_{18}	**122**
O_{19}	**42**

4 Experimental Results

This section presents the experimental results of the proposed case study. The dataset contains 20 of the existing ontologies in the domain of e-government. Their OntoMetrics measurements are outlined in [31]. Table 6

Table 4: OntoMetrics calculations [44]

Category	Metric	Equation	Description
Schema Metrics	Attribute Richness (AR)	$AR = \frac{\lvert att\rvert}{\lvert C\rvert}$ (1)	|att| is the total number of attributes |C| is the total number of classes in the ontology
Schema Metrics	Inheritance Richness (IR)	$IR = \frac{\lvert H\rvert}{\lvert C\rvert}$ (2)	|H| is the number of subclass relations |C| is the total number of classes
Schema Metrics	Relationship Richness (RR)	$RR = \frac{\lvert P\rvert}{\lvert H\rvert+\lvert P\rvert}$ (3)	|P| is the number of non-inheritance relations |H| is the number of inheritance relations
Knowledge Base Metrics	Average Population (AP)	$AP = \frac{\lvert I\rvert}{\lvert C\rvert}$ (4)	|I| is the total number of instances of the knowledge base |C| is the total number of classes
Knowledge Base Metrics	Class Richness (CR)	$CR = \frac{\lvert C'\rvert}{\lvert C\rvert}$ (5)	|C'| is the number of classes in the knowledge base |C| is the total number of classes
Graph Metrics	Absolute Root Cardinality (ARC)	$ARC = n_{ROO\subseteq g}$ (6)	$n_{ROO\subseteq g}$ represents the number of elements in the set of root nodes *ROO* in the directed graph *g*
Graph Metrics	Absolute Leaf Cardinality (AC)	$AC = n_{LEA\subseteq g}$ (7)	$n_{LEA\subseteq g}$ represents the number of elements in the set of leaf nodes *LEA* in the directed graph *g*
Graph Metrics	Average Depth (AD)	$AD = \frac{1}{n_{P\subseteq g}} \sum_j^P N_{j\in P}$ (8)	P represents the set of paths in the directed graph g $n_{P\subseteq g}$ is the number of elements in P $N_{j\in P}$ is the number of elements on the path j.
Graph Metrics	Maximum Depth (MD)	$MD = N_{j\in P} \forall i \exists j (N_{j\in P} \geq N_{i\in P})$ (9)	$N_{j\in P}$ is the number of elements on the path *j* $N_{i\in P}$ is the number of elements on the path *i* which belong to the set of paths *P* in the directed graph *g*
Graph Metrics	Average Breadth (AB)	$AB = \frac{1}{n_{L\subseteq g}} \sum_j^L N_{j\in L}$ (10)	*L* represents the set of levels in the directed graph g $n_{L\subseteq g}$ is the number of elements in L $N_{j\in L}$ is the number of elements on the level j.
Graph Metrics	Maximum Breadth (MB)	$MB = N_{j\in L} \forall i \exists j (N_{j\in L} \geq N_{i\in L})$ (11)	$N_{j\in L}$ and $N_{i\in L}$ are the number of elements on the level j and i respectively that belong to the set of levels L in the directed graph g

Table 5: OntoMetrics and ontology dimensions correlation [44]

Dimension	Description	Metrics
Accuracy	Determines how well the ontology represents the real-world domain	Equations (1), (2), (3), (8), (9), (10), and (11)
Understandability	Indicates the comprehension of the elements of the ontology	Equation (7)
Cohesion	Assesses how closely classes are related to one another	Equations (6) and (7)
Conciseness	Affects the extent to which the ontological information is beneficial	Equations (4) and (5)

illustrates the comparison between them and the proposed ontology's metrics. And to have an overall view, Table 7 summarizes the average value for all metrics and each dimension separately, moreover, the percentile [28] at which the average lies. The following sections discuss these findings.

4.1 Overall Metrics

In the case of EGYGOV's development outcome, the

Table 6: OntoMetrics results of e-government ontologies

Ontology	AR	IR	RR	AP	CR	ARC	AC	AD	MD	AB	MB
O_1	0.59099	0.136364	0.930233	0.681818	0.090909	19	19	1.136364	2	5.5	19
O_2	0.315789	0.736842	0.222222	0.526316	0.105263	7	10	3.033333	5	2.142857	7
O_3	0	1	0.346154	0.352941	0.058824	2	10	3.285714	5	2.33333	4
O_4	0.076923	1.230769	0.407407	0	0	3	10	2.076923	3	3.25	6
O_5	0.067797	0.779661	0.577982	0.016949	0.016949	15	42	2.836066	6	3.388889	15
O_6	0.666667	1.333333	0.5	1.166667	0.333333	3	2	1.5	2	1.5	3
O_7	0.033708	1.601124	0.2711	0.134831	0.02809	26	112	3.419162	6	3.604317	33
O_8	0.666667	1.333333	0.5	1.166667	0.333333	3	2	1.5	2	1.5	3
O_9	7.416667	7.333333	0.169811	26	0.583333	1	8	3.416667	4	2.4	4
O_{10}	0.106061	1.515152	0.602386	1.386364	0.530303	26	92	2.447368	4	4.956522	26
O_{11}	0.363636	0.818182	0.25	4.727273	0.272727	2	7	2.727273	4	2.2	3
O_{12}	0.375	0.75	0.4	1.375	0.125	2	6	2.375	3	2.666667	5
O_{13}	0.75	0.75	0.333333	1.875	0.125	2	6	2.25	3	2.666667	4
O_{14}	0	0.984127	0	0.992063	0.007937	2	107	3.801587	4	6.3	17
O_{15}	0	1	0.346154	0.352941	0.058824	2	10	3.285714	5	2.333333	4
O_{16}	0.36	2.12	0.341615	0.54	0.12	22	22	1.3125	2	2.909091	22
O_{17}	0.217391	1.978261	0.172727	0.652174	0.434783	16	56	3.009709	6	2.575	16
O_{18}	0.042231	1.130279	0.003862	0	0						
O_{19}	0.133333	1.148148	0.093567	0	0	8	112	3.639706	5	5.666667	25
O_{20}	0.4375	0.75	0.813956	1.90625	0.5	8	26	1.84375	3	4.571429	9
EGYGOV (Development Outcome)	**0.436508**	**1.571429**	**0.171548**	**0.880952**	**0.055556**	**1**	**114**	**3.188976**	**4**	**9.769231**	**45**
EGYGOV (Enrichment Outcome)	**0.772947**	**1.074879**	**0.044039**	**0.992754**	**0.310386**	**46**	**605**	**3.347107**	**8**	**8.144231**	**46**

average of all OntoMetrics measurements equals 16.37. This makes the proposed ontology the second one among the 20 ontologies and locates it in the 90th percentile, while O_7 comes in the first percentile (95th). On the other hand, after applying the enrichment module, results of eight metrics are enhanced. Therefore, the average is raised to 65.43. EGYGOV becomes the highest one among the 20 ontologies with a 95th percentile, while O_7 becomes the second one (90th percentile). These results are illustrated in Figure .6.

4.2 Accuracy

The average of accuracy-related metrics is the highest for both outcomes of EGYGOV and falls on the 95th percentile. The enrichment module enhances the results of four metrics out of seven while IR, RR and AB return lower values in case of enrichment outcome than development outcome. This low IR means that classes and relations injected into EGYGOV increase the depth and detailed coverage of the e-government domain. While low RR is brought on by the matcher returning only is-a relationships. As a result, inheritance relations are increased in contrast to other types of relations. So, it is better to define more non-inheritance relations in EGYGOV. Finally, the low value of AB indicates that EGYGOV's final outcome focuses on the vertical modeling of hierarchies rather than the horizontal modeling.

4.3 Understandability

The leaf nodes (AC) of the EGYGOV are the highest (95th percentile) in both outcomes. This large metric resolves the ambiguity of the ontology classes and relations. As a result, it leads to good understandability that will help domain experts to easily understand the ontology.

4.4 Cohesion

The average of both graph metrics ARC and AC in case of development outcome lies in the 80th percentile. This iscaused because of low number of root nodes in that outcome. But the average is moved to the 95th percentile after applying the enrichment module. The higher values of

Table 7: OntoMetrics - average and percentile

Ontology	All Metrics		Metrics correlated to Accuracy		Metrics correlated to Understandability		Metrics correlated to Cohesion		Metrics correlated to Conciseness	
	AVG	Percentile	AVG	Percentile	AVG	Percentile	AVG	Percentile	AVG	Percentile
O_1	6.19	57%	4.18	62%	19	50%	19	55%	0.39	43%
O_2	3.28	43%	2.64	43%	10	30%	8.5	45%	0.32	33%
O_3	2.58	29%	2.28	29%	10	30%	6	30%	0.21	24%
O_4	2.64	38%	2.29	38%	10	30%	6.5	40%	0	0%
O_5	7.79	67%	4.09	52%	42	65%	28.5	65%	0.02	14%
O_6	1.55	5%	1.50	5%	2	0%	2.5	0%	0.75	62%
O_7	16.92	90% or 95%	6.85	90%	112	85%	69	95%	0.08	19%
O_8	1.55	5%	1.50	5%	2	0%	2.5	0%	0.75	62%
O_9	5.85	52%	4.11	57%	8	25%	4.5	20%	13.29	95%
O_{10}	14.50	81%	5.66	81%	92	75%	59	85%	0.96	76%
O_{11}	2.49	24%	1.91	14%	7	20%	4.5	20%	2.50	90%
O_{12}	2.19	19%	2.08	24%	6	10%	4	10%	0.75	62%
O_{13}	2.16	14%	1.96	19%	6	10%	4	10%	1.00	81%
O_{14}	12.92	76%	4.58	76%	107	80%	54.5	75%	0.50	52%
O_{15}	2.58	29%	2.28	29%	10	30%	6	30%	0.21	24%
O_{16}	6.88	62%	4.43	71%	22	55%	22	60%	0.33	38%
O_{17}	9.37	71%	4.28	67%	56	70%	36	70%	0.54	57%
O_{18}	0.24	0%	0.39	0%					0	0%
O_{19}	14.61	86%	5.81	86%	112	85%	60	90%	0	0%
O_{20}	5.17	48%	2.92	48%	26	60%	17	50%	1.20	86%
EGYGOV (Development Outcome)	**16.37**	**90%**	**9.16**	**95%**	**114**	**95%**	**57.5**	**80%**	**0.47**	**48%**
EGYGOV (Enrichment Outcome)	**65.43**	**95%**	**9.63**	**95%**	**605**	**95%**	**325.5**	**95%**	**0.65**	**57%**

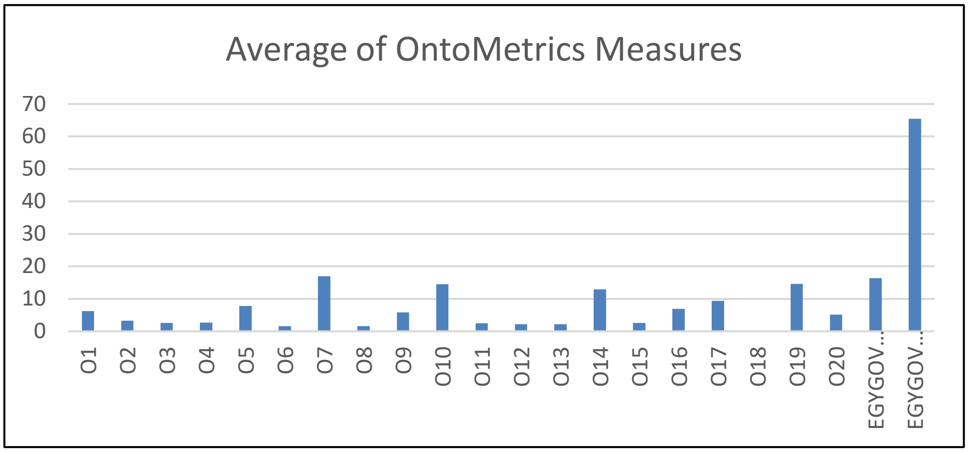

Figure 6: Average of onto metrics measures

both metrics in the EGYGOV's final outcome lead to the same conclusion, which is that the ontology accurately describes the e-government domain using a large number of inheritance relations. This also supports the idea that classes are well connected to each other.

4.5 Conciseness

This dimension is measured via the number of instances and their distribution among classes. The average of those metrics in the case of development and enrichment outcomes lies in percentiles 48% and 57%, respectively. Low AP and CR metrics are the reason why the average is relatively low. Low AP means that the instances defined into the knowledgebase is insufficient to populate all of the knowledge. While low CR indicates that the knowledgebase lacks data that fully demonstrates all of the available classes. This can be justified because EGYGOV has relatively high number of classes (125 classes in the development outcome and 828 classes in the enrichment outcome). So, to return high values in this dimension, a huge number of instances need to be defined and populated across all the classes. It is noticed that the same scenario applies to other ontologies that have a greater number of classes such as O_{14}, O_{18} and O_{19} (as displayed in Table 2). For two of them their average is located in the 0^{th} percentile, and the third one is located in 52^{nd} percentile.

5 Conclusion And Future Work

This research presents an enhanced ontology engineering architecture for building domain ontologies from scratch. This architecture is based on ODCM and ontology matching. With the help of this architecture, users can complete ontology development tasks since it provides guidance for all key activities, from requirement specification to ontology evaluation. It is composed of four main modules: Requirement Specification, Ontology Development, Ontology Enrichment, and Ontology Quality Assessment. The proposed architecture is applied to the domain of e-governance in Egypt. The results are encouraging when the produced ontology is compared to 20 existing ontologies from the same domain. On the basis of OntoMetrics, the average of metrics related to accuracy, understandability, cohesion and conciseness lies in 95^{th}, 95^{th}, 95^{th} and 57^{th} percentiles respectively. The results can be enhanced by increasing the non-inheritance relations and by distributing the instances across all classes. The three main contributions of this study are: (1) the combination of ODCM conceptual modeling with ontology matching, (2) two new algorithms for the selection of the most relevant ontology and the extraction of new classes and relations, as well as (3) the new domain ontology for the Egyptian e-government.

In the future, it is planned to 1) apply the proposed architecture to various domains. 2) enrich the ontology with new constraints to increase its expressiveness; and 3) employ other methods in the assessment process.

References

[1] R. Abaalkhail, *Ontology Based Framework for Conceptualizing Human Affective States and Their Influences*, Doctoral Dissertation, Université d'Ottawa/University of Ottawa, 2018.

[2] E. Abrahão and A. R. Hirakawa, "Complex Task Ontology Conceptual Modelling: Towards the Development of the Agriculture Operations Task Ontology," *KEOD*, 2:285-292, 2018.

[3] A. Q. Al-Namiy, "Algorithm to Match Ontologies on the Semantic Web," *Int J Adv Comput Sci Appl*, 4(3):221-227, 2013.

[4] S. Al-Yadumi, W. W. Goh, E. X. Tan, N. Z. Jhanjhi, and P. Boursier, "Multimatcher Model to Enhance Ontology Matching Using Background Knowledge," *Information,* Switzerland, 12(11):487, doi: 10.3390/info12110487, Nov. 2021.

[5] J. Antonio and T. Saorín Pérez, "A Conceptual Model for an OWL Ontology to Represent the Knowledge of Transmedia Storytelling," 15th International ISKO Conference, pp. 511-520, 2018.

[6] N. Alrebdi and N. Khan, "Core Elements Impacting Cloud Adoption in the Government of Saudi Arabia," *Int J Adv Comput Sci Appl*, 13(6):270-273, 2022.

[7] W. Boggs and M. Boggs, *Mastering UML with Rational Rose 2002,* Alameda: Sybex, Vol. 1, 2002.

[8] S. Debbech, S. Collart-Dutilleul, and P. Bon, "An Ontological Approach to Support Dysfunctional Analysis for Railway Systems Design," *Journal of Universal Computer Science*, 26(5):549-582, 2020.

[9] T. Derave, "A Reference Architecture for Customizable Marketplaces," International Conference on Conceptual Modeling, pp. 222–229. doi: 10.1007/978-3-030-34146-6_20, 2019.

[10] T. Derave, T. P. Sales, M. Verdonck, F. Gailly, and G. Poels, "Domain Ontology for Digital Marketplaces," International Conference on Conceptual Modeling, pp. 191-200, 2019.

[11] "Egyptian E-Government Portal." http://www.egypt.gov.eg/english/home.aspx (accessed Aug. 04, 2022).

[12] H. Elasri, A. Sekkaki, and L. Kzaz, "An Ontology-Based Method for Semantic Integration of Business Components," *2011 11th Annual International Conference on New Technologies of Distributed Systems, NOTERE 2011 – Proceedings*, pp. 1-8. doi: 10.1109/NOTERE.2011.5957993, 2011.

[13] D. Faria, C. Pesquita, E. Santos, M. Palmonari, I. F. Cruz, and F. M. Couto, "LNCS 8185 - The AgreementMakerLight Ontology Matching System," OTM Confederated International Conferences on the Move to Meaningful Internet Systems, Springer, Berlin, Heidelberg, pp. 527-541, September 2013.

[14] J. V. Fonou-Dombeu and S. Viriri, "OntoMetrics Evaluation of Quality of e-Government Ontologies," International Conference on Electronic Government

and the Information Systems Perspective, pp. 189-203, 2019.

[15] N. G'abor, "Ontology Development," Semantic Web Services, pp. 107-134, 2007.

[16] A. Gal, "Ontology Engineering," in *Encyclopedia of Database Systems*, New York, NY: Springer New York, pp. 2584-2585. doi: 10.1007/978-1-4614-8265-9_1315, 2018.

[17] A. García S, G. Guizzardi, O. Pastor, V. C. Storey, and A. Bernasconi, "An Ontological Characterization of a Conceptual Model of the Human Genome," *International Conference on Advanced Information Systems Engineering*, pp. 27–35. doi: 10.1007/978-3-031-07481-3_4, 2022.

[18] T. R. Gebba and M. R. Zakaria, "E-Government in Egypt: An Analysis of Practices and Challenges," *International Journal of Business Research and Development*, 4(2):11-25, 2015.

[19] J. Guerson, T. P. Sales, G. Guizzardi, and P. A. Almeida, "OntoUML Lightweight Editor A Model-Based Environment to Build, Evaluate and Implement Reference Ontologies," 2015 IEEE 19th International Enterprise Distributed Object Computing Workshop, pp. 144-147, 2015.

[20] G. Guizzardi, *Ontological Foundations for Structural Conceptual Models*, PhD Thesis, Centre for Telematics and Information Technology, University of Twente, 2005.

[21] G. Guizzardi, "Theoretical Foundations and Engineering Tools for Building Ontologies as Reference Conceptual Models," *Semant Web*, 1(1,2):3-10, 2010.

[22] G. Guizzardi, G. Wagner, J. Paulo, A. Almeida, and R. S. Guizzardi, "Towards Ontological Foundations for Conceptual Modeling: The Unified Foundational Ontology (UFO) Story," *Appl Ontol*, 10(3–4):259-271, 2015.

[23] M. Haggag, S. Fathy, and N. Elhaggar, "Ontology-Based Textual Emotion Detection," *Int J Adv Comput Sci Appl*, 6(9):239-246, 2015.

[24] J. Hartmann, R. Palma, and Y. Sure, "OMV-Ontology Metadata Vocabulary," *ISWC*, Vol. 3729, 2008.

[25] K. Hussein Shafa and J. Omer Atoum, "A Framework for Improving the Performance of Ontology Matching Techniques in Semantic Web," *Int J Adv Comput Sci Appl*, 3(1):8-14, 2012.

[26] T. Jonsson and H. Enquist, "Phenomenological Ontology Guided Conceptual Modeling for Enterprise Information Systems," International Conference on Conceptual Modeling, pp. 31-34, 2018.

[27] H. Lamharhar, D. Chiadmi, and L. Benhlima, "Moroccan e-Government Strategy and Semantic Technology," Government e-Strategic Planning and Management, pp. 323–343, 2014.

[28] E. Langford, "Quartiles in Elementary Statistics," *Journal of Statistics Education*, 14(3):1-27, doi: 10.1080/10691898.2006.11910589, Jan. 2006.

[29] X. Liu, Q. Tong, X. Liu, and Z. Qin, "Ontology Matching: State of the Art, Future Challenges, and Thinking Based on Utilized Information," *IEEE Access*, 9:91235-91243, doi: 10.1109/ACCESS.2021.3057081, 2021.

[30] M. A. Musen, "The Protégé Project: A Look Back and a Look Forward," *AI Matters*, 1(4):4-12, June 2015.

[31] J. C. Nardi, R. de Almeida Falbo, J. P. A. Almeida, G. Guizzaardi, L. F. Pires, M. J. van Sinderen, N. Guarino, and C. M. Fonseca, "A Commitment-Based Reference Ontology for Services," *Inf Syst*, 54:263-288, Dec. 2015.

[32] D. Oliveira and C. Pesquita, "Improving the Interoperability of Biomedical Ontologies with Compound Alignments," *J Biomed Semantics*, 9(1):1-13, doi: 10.1186/s13326-017-0171-8, Jan. 2018.

[33] D. Porello, G. Guizzardi, T. P. Sales, and G. Amaral, "A Core Ontology for Economic Exchanges," International Conference on Conceptual Modeling, pp. 364–374, 2020.

[34] M. Poveda-Villalón, A. Fernández-Izquierdo, M. Fernández-López, and R. García-Castro, "LOT: An Industrial Oriented Ontology Engineering Framework," *Eng Appl Artif Intell*, 111:104755, doi: 10.1016/j.engappai.2022.104755, May 2022.

[35] P. Sacco, R. Gallo, and F. Mazzetto, "Farm Ontology: A System Thinking Approach for Planning and Monitoring Farm Activities," *Proceedings of the 3rd LeNS World Distributed Conference*, pp. 429-434, 2019.

[36] C. Silva and O. Belo, "A Core Ontology for Brazilian Higher Education Institutions," *CSEDU*, pp. 377–383, 2018.

[37] M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López, "The Neon Methodology for Ontology Engineering," *Ontology Engineering in a Networked World*, Springer Berlin Heidelberg, pp. 9-34, 2012.

[38] M. Suchánek, "OntoUML Specification Documentation," https://ontouml.readthedocs.io/en/latest/, 2020.

[39] R. Syed and H. Zhong, "Cybersecurity Vulnerability Management: An Ontology-Based Conceptual Model," 24th Americas Conference on Information System*s, AMCIS*, pp. 16-18, 2018.

[40] J. Trujillo, K. C. Davis, X. Du, E. Damiani, and V. C. Storey, "Conceptual Modeling in the Era of Big Data and Artificial Intelligence: Research Topics and Introduction to the Special Issue," *Data and Knowledge Engineering*. Elsevier B.V., 135:101911, doi: 10.1016/j.datak.2021.101911 Sep. 01, 2021.

[41] United Nations, "UN E-Government Survey," https://publicadministration.un.org/egovkb/en-us/Data/Country-Information/id/53-Egypt/dataYear/2020 (accessed Aug. 04, 2022).

[42] C. F. Uwasomba, Y. Lee, Z. Yusoff, and T. M. Chin,

"Ontology-Based Methodology for Knowledge Acquisition from Groupware," *Applied Sciences,* Switzerland, 12(3):1448, doi: 10.3390/app12031448, Feb. 2022.

[43] M. Verdonck and F. Gailly, "Insights on the Use and Application of Ontology and Conceptual Modeling Languages in Ontology-Driven Conceptual Modeling," *International Conference on Conceptual Modeling*, LNCS, 9974:83-97, 2016.

[44] A. N. H. Zaied, A. H. Ali, and H. A. El-Ghareeb, "E-government Adoption in Egypt: Analysis, Challenges, and Prospects," *International Journal of Engineering Trends and Technology*, 52(2):70-79, Oct. 2017.

[45] M. R. Zein and H. Twinomurinzi, "Towards Blockchain Technology to Support Digital Government," International Conference on Electronic Government and the Information Systems Perspective, pp. 207–220, 2019.

SHAIMAA HARIDY received the B.Sc. and M.Sc. degrees in information systems from the Faculty of Computer and Information Sciences (FCIS), Ain Shams University (ASU), Egypt, in 2004 and 2011, respectively. Her master's was about Representation of Ontologies for Semantic Web Services. She is currently an Assistant Lecturer with FCIS, ASU. Her research interests include artificial intelligence, semantic web, ontology engineering, and software engineering.

RASHA ISMAIL received the Ph.D. degree in data mining and data warehousing from Ain Shams University, Egypt, in 2009. She was the Director of the credit hours programs, from 2013 to 2019. She has contributed in implementing of the current credit-hour curriculum for the undergraduate programs of bioinformatics and software engineering at Ain Shams University, where she is currently a Professor and the Vice Dean of postgraduate studies and research affairs with the Faculty of Computer and Information Sciences. In addition, she is also the Director of the Ain Shams Portal. Her current research interests include data science, software engineering, information retrieval, big data analytics, data mining, and bioinformatics.

NAGWA BADR received the M.S. degree in computer science and the Ph.D. degree in software engineering and distributed systems from Liverpool John Moores University, U.K., in 1996 and 2003, respectively. She had done postdoctoral studies with Glasgow University, U.K. She is currently a Professor and the Dean of the Faculty of Computer and Information Sciences (FCIS), Ain Shams University (ASU). For the last few years, she is the Head of committee that contributed to research projects funded by national and international grants in information systems, bioinformatics, business analytics, and health informatics. Her research interests include software engineering, cloud computing, big data analytics, social networking, Arabic search engines, and bioinformatics.

MOHAMED HASHEM was the Head of the Information Systems Department, FCIS. He was the Vice Dean of education and student affairs at the Faculty of Computer and Information Sciences (FCIS). He is currently a Professor in information systems with FCIS, Ain Shams University, Egypt. His research interests include modeling and simulation of computer networks, computer security, and data management.