Ontology Concepts for Requirements Engineering Process in E-Government Applications

Conference Paper · June 2010

DOI: 10.1109/ICIW.2010.66 · Source: IEEE Xplore

3 authors, including:

Ahmad Kayed
Sohar University
40 PUBLICATIONS **234** CITATIONS

SEE PROFILE

Mohammad Nezar
University of Bedfordshire
3 PUBLICATIONS **11** CITATIONS

SEE PROFILE

An Ontology for Requirements Engineering Process in E-Government Applications

Ahmad Kayed
Fahad Ben Sultan University
KSA
Kayed@FBSC.edu.Sa
Mohammad Nizar
mohdnizar11@yahoo.com

The main challenge for governments worldwide is to transform their systems to deliver efficient and cost effective services through information and communication technologies. Development of information and communication technologies catalyzed and led up to e-government (E-gov). Requirements engineering is a very crucial activity in system development process, especially in the development of large scale systems such as E-gov applications. Requirements engineering process includes domain analysis, elicitation, specification, assessment, negotiation, documentation and evolution. This paper develops ontology (mainly concepts) in the domain of requirements engineering process for E-gov applications. This contributes in enabling software engineers to find out shared-understandable and common concepts to describe requirements for different domain models used in developing E-gov applications. We built ontology for Requirements Engineering Process in E-gov Applications (REPEA). We conducted several experiments to extract the main concepts for REPEA. Several documents related to E-gov requirement are collected; the main concepts and relationships are extracted and refined. The results show that there are number of concepts are frequently used to describe these REPEAs. Summarizing and formalizing the semantic of these concepts will present a common understanding and agreement on the semantic of REPEA which can be used by software engineers, researchers, practitioners, and stakeholders.

Keywords
Software requirement engineering, Ontology, Semantic, concept, E-government.

1. Introduction

In E-government (E-gov) context, governments use the most innovative information and communication technologies, in particular web-based Internet applications, to provide their citizens and businesses with more convenient access to government information and services, to improve the quality of the services and to provide greater opportunities to participate in democratic institutions and processes **[1]**. Building and implementing applications for E-gov are increasing every day. Many governments are adopting E-gov approach to computerize their services.

The need to reinvent government systems in order to deliver efficient and cost effective services, information and knowledge through information and communication technologies is a big challenge for many governments worldwide. Software engineering had been widely affected by the appearance of the Internet, programmers and developers were required to deal with many new revealed issues and merge it within their developed software (images, maps, animations, web browsers usage, etc). Simpler and faster methodologies that developed running and inexpensive software products have been introduced to organizations and governments in order to satisfy their demand. Examples of these methodologies are: rapid prototyping, agile development, extreme programming, and others [2]. Software Engineering faces many key challenges such as heterogeneity, delivery, trust, cost, timelines, and quality challenges [3] [4].

Requirements engineering, an essential phase in the software life cycle, has been a challenging problem **[5]**. Software developers need to capture the rationale of the underlying business logic rather than merely operational procedures. They need to be able to capture the semantics of the required intelligent processing based on the application domain **[6]**. Requirements engineering is a very crucial activity in the development of large scale systems such as E-gov applications. It is concerned with domain specifications of software behavior and with their evolution over time **[7].** The processes involved in requirements engineering process include domain analysis, elicitation, specification, assessment, negotiation, documentation and evolution. This implies that

getting high quality and well-defined domain for software requirements process for E-gov application are difficult and critical **[8]**. Many surveys have confirmed the growing recognition of requirements engineering process as a new area of knowledge in software engineering research and practice. A survey over 3800 organizations in 17 countries concluded that most of the perceived software problems are in the area of requirements specification [**8]**.

Requirements engineering research spans a wide range of topics, but a topic of increasing importance is to require more specific focus on domain conceptualization and integrating different requirements engineering activities in an era obsessed with better, faster, and reusable software development scenarios **[9] [4][10]**. In this context, requirements engineering activities must be improved and integrated by applying new various support knowledge representation techniques **[11].** Each activity of requirements engineering process has its own characteristics, terms, and concepts. Each specific domain establishes its own vocabularies and semantics for the elements, processes, and relationships for software applications in that domain **[12].** This makes applying these various semantics in requirements engineering process is a complex process. The need to share understanding of domain specifications has becoming a critical issue. The share understating can be improved by building Ontology ideally captures more knowledge about the multiple aspects, concerns, and activities involved in the requirements engineering process **[8]**.

The main aim of this paper is to build shared concepts using ontological tools for E-gov applications. In this work we study and analyze 68 documents from reports, documents, and proposals concerned with software requirement for e-Gov applications. Then we extract various concepts, definitions, and terminologies from them. Our claim is that the semantic of these definitions can be condensed into small set of concepts. If these concepts have enough and clear semantic, they will achieve a common understanding for any other requirements for E-gov applications. This paper provides a brief introduction to Ontology and some of its practical usage which appears in section two. Section three describes our experiment in extracting and defining Ontology domain concepts to present common concepts for e-Gov application. Section four will include a discussion of our findings and results. Section five concludes the paper.

2. Ontology and Information Semantic

Ontology as seen from philosophical perspective is the science of studying beings (studying of what is, of the kinds and structures of objects, properties, events, processes and relations in every area of reality), this term which was coined in 1613 included in many philosophical areas from the metaphysics of Aristotle to the object-theory of Alexius Meinong [13]. Philosophical Ontology handles the precise utilization of words as descriptors of entities, it gives an account for those words that belong to entities and those do not [14]. In both computer science and information science, an ontology is a representation of a set of concepts within a domain and the relationships between those concepts. It is used to reason about the properties of that domain, and may be used to define the domain [15]. Recently, the term ontology has widely included in the field of computer and information science. Knowledge base systems designers use a wide variety of terms and concepts when they build frameworks for information representation. Studies showed that there are inconsistency problems in the semantic of the terms that used. For example, identical databases labels are used but with different meaning and the same meaning may be expressed using different names [16]. Ontology in this context is a dictionary of terms formulated in a canonical syntax and with commonly accepted definitions designed to yield a lexical or taxonomical framework for knowledge-representation which can be shared by different information systems communities [16]. Ontologies are used in many fields such as: artificial intelligence, software engineering, the semantic web, biomedical informatics, library science, information architecture, and as a form of knowledge representation about the world or some part of it.

Every domain researchers and practitioners need to share information to conduct their works. Inconsistencies between terms and concepts must be reduced. Ontology defines a common vocabulary and contains machine-interpretable definitions of basic concepts in the domain and relations among them**.** Ontologies have been used in different areas of software engineering. Ontologies can provide a general framework reference of an agreed concepts and terminologies among researchers, practitioners, and stakeholders. They enhance collaboration, communication, and knowledge sharing. They contribute in reducing gabs among researchers created by conceptual confusion [17,18,19].

Recently the semantic Web initiative, lead by W3C, has changed the ontology landscape, through the initiative, researchers and developers join forces to provide standard semantics markup languages based on XML, ontology management systems, and other useful tools. Web provides interesting applications of ontology critical to daily life such as search and

navigation. In addition, people rediscover the value of ontology in other important applications such as information and process integration [20][21]. Ontologies have been widely used during the last years in the area of e-gov. Balzer has started advocating the benefits of underlying ontologies of precise and formal specifications, notably for checking a specification adequacy through prototyping [22], others used ontology to propose multi-paradigm frameworks to combine multiple languages in a semantically meaningful way so that different facets can be captured by languages that fit them best [23]. Kim [24] proposes a formal model of enterprise quality, called "Ontology of enterprise quality modeling". This is a global ontology, whose main objective is to help evaluate the conformance of organizations to ISO/IEC 9000 standards. Although Kim's measurement ontology is not specific to software products and processes, it contains many concepts that can be applied within the context of e-gov applications. Under this perspective, Kim's proposal mainly focuses on *targets-and-goals*, including concepts such as "quality requirement", "entity", "enterprise model of quality", and "measured attribute". Contributions in Software Engineering have been made toward using ontologies. Devedzic has argued that ontologies are needed in all phases of software engineering life cycle, each of which must have knowledge, whether about data structure, methods and domain. This makes ontologies are everywhere and they make possible to smoothly integrate artificial intelligence with other software disciplines [25]. Luis Olsina [26], Weber [27], and Felix Garcia [28], showed that there is a lack of consensus on the concepts and terminologies that are used in this field.

In this paper, we are building ontology to capture the conceptualization knowledge about Requirements Engineering Process in E-Government Applications (REPEA) domain which will provide a solution for the semantic conflicts problem**.** REPEA has its own characteristics, terms, concepts, and terminologies that establishing the vocabulary and semantics for e-gov application, this makes applying and understanding these various semantics by software engineers as a common shared knowledge a complex issue. Ontology shows enormous potential in making software more efficient, adaptive, and intelligent. It is recognized as one of the areas which will bring the next breakthrough in software development [20].

3. Our Experiment

The main aim of this paper is to develop and produce a unified ontological framework that ideally captures and analyzes the domain knowledge about the multiple concepts, definitions, and terminologies that current software products for e-gov proposals. We are extracting the conceptualization of the REPEA in order to produce a coherent and consistent set of common terminologies and concepts as a common agreement in the E-gov domain**.** This contribution aims at enabling software engineers, researchers, practitioners, and stakeholders to find out shared, understanding, and common concepts to describe REPEA. This will remove gabs, inconsistencies, and terminology conflicts that can affect E-gov applications to reach to a consensus knowledge domain. In the following we will summarize our steps to develop this ontology:

1. We collected about 68 reports, documents, standards, and publications which are related to REPEA.

2. We used KAON's TextToOnto extension software [29] [30] to extract the terms, concepts, and relationships which are used in the selected REPEA context. There were 1358 concepts, sample of these concepts are shown in table 1.

3. We used a tool created by Kayed [31] which is a combination between MS Access tool and MS Visual Basic to refine these concepts. The 1358 concepts and their relationships entered to this tool. This tool provided us with 90 concepts. Table 1 below lists part of these concepts.

4. We eliminated the stopping words (extremely common words like use, can, the, of, etc) from these concepts.

5. The concepts set from the elimination process were sent to three experts (professors and practitioners) in the field of REPEA. The result was 25 concepts were common among the three experts.

6. To evaluate the 25 concepts we inserted all those concepts into a database with their definitions as in table 3. We extracted the related concepts for each 25 concept's definitions. For example the concept "Accuracy" are related with 11 concepts " Actual, Calculated , Closeness, Conformity, …etc.".
7. For each concept, we counted how many concepts in its definition are in the ontology (with their definitions). The results are summarized in table 4.
8. We computed the percentage of coverage by dividing the related concepts for each 25 concepts by the available concepts in the ontology. The average coverage for all 25 concepts was 74%.

Figure 1: The Term Extraction process using the TextToOnto tool.

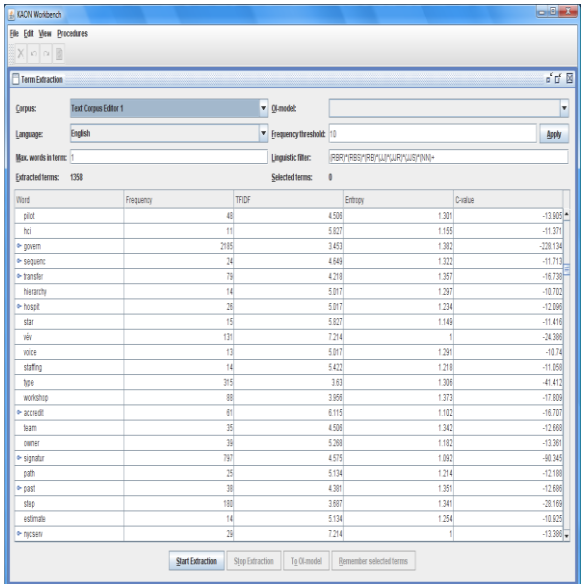

Table 1: Sample of the 90 concepts.

Concept	Freq	Concept	Freq	Concept	Freq
adoption	44	etax	52	scope	59
broker	47	principle	52	official	17
treatment	14	Tax	855	function	178
democracy	43	vote	655	confidentiality	67
client	132	categorization	18	channel	48
penetration	29	elicitation	13	consistency	18
feature	23	policy	234	accuracy	32
web	899	standard	119	awareness	50
vision	39	inspection	13	abuse	35
intrusion	53	competence	23	purpose	91
e-signature	84	clarity	10	stakehold	58

Table 2: The 25 ontology domain concepts.

Concept	Concept	Concept	Concept	Concept
adoption	Security	classification	confidentiality	trust
client	Elicitation	investment	accuracy	strategy
esignature	Policy	database	awareness	evaluation
citizen	Standard	govern	purpose	knowledge
sector	Reliability	scope	stakehold	identification

Table 3: Part of the REPEAs concepts and their definitions from various sources/references.

ID	Concept	Definition(s)	Source(s) Reference(s)
1	Accuracy	Conformity to fact.	www.answers.com
		Conformity to truth or to a standard or model	www.merriam-webster.com
		The degree of closeness of a measured or calculated quantity to its actual (true) value	www.wikipedia.org
2	Policy	A plan or course of action, as of a government, political party, or business, intended to influence and determine decisions, actions, and other matters.	www.answers.com
		etc	

Table 4: Coverage Process Results.

ID	Concept	Concepts from Definitions	Concepts from Ontology and Synonyms	Counts and Percentage
1	Accuracy	Accuracy	Accuracy	8 from 11
		Actual	Actual	0.73
		Calculated	Conformity	
		Closeness	Degree	
		Conformity	Fact	
		Degree	Quantity	
		Fact	Truth	
		Model	Value	
		Quantity		
		Truth		
		Value		
2	Policy	Action	Action	8 from 10
		Business	Business	0.80
		Government	Government	
		Guide	Guide	
		Influence	Outcome	
		Outcome	Plan	
		Overall	Policy	
		Plan	Term	
		Policy		
		Term		
		Speed		
The Average of Coverage is : 0.748 ≈ 0.75				

4. Discussion

We used the New Term Extraction function oh KAON in order to extract concepts from provided text corpus. This tool uses the natural language processing algorithms in addition to semantic lexicon filtering techniques. We set the frequency threshold to 5, but the result was more than 2000 concepts. We set the frequency threshold again to several values (10, 15, and 20). After many trials, we decided to choose the 10 frequencies and to retrieve concepts that consist of only one unique word as a term. 10 frequencies gave us a suitable number of concepts i.e. not too large or too small). Figure 1 illustrates this step.

The 68 documents have been reduced to 1358 concepts by KAON's tool, and reduced to 90 by Kayed's tools and reduced to 25 concepts by human experts. There are several methodologies to evaluate Ontology. In the following we will state some:

1. Evaluation based on the usage of ontology in a context of an application or project, to evaluate its effectiveness [32].
2. Evaluations based on the effort done by human experts [33].
3. Evaluations based on comparing the ontology with other ontology in the same domain [34].
4. Evaluations based on the formal representation of the ontology with other formal ontology [35].
5. Evaluations based on fitting or covering techniques between an ontology and a domain of knowledge that the ontology is created for [11] [36].

In this paper we decided to choose the covering approach to evaluate the proposed ontology for the following reasons:

1. It is beyond the scope of this paper to build an application to check the effectiveness of the proposed ontology.
2. We used the human expert approach to reduce the number of concepts this means that this approach will not give unbiased judgment to evaluate our ontology.
3. As far as we know that there is no available ontology (formal or informal) for REPEA. We can't compare our ontology with it.

We collected several definitions for each 25 concepts from several resources. Then we extracted the main concepts for these definitions. For each 25 concepts, we divided the number of concepts needed for each concept definitions by the number of concept available in the proposed ontology (the 25 concepts and their synonyms). Table 3.3 shows part of the definitions of the 25 concepts and the source of these definitions. We calculated the average coverage for each concept. Finally we calculated the average for all averages for all concepts. The result showed that 74% of the definitions concepts are covered by our ontology in average. The evaluation process confirmed that our ontology domain concepts covered almost 74% from the given knowledge domain. This result supports our claim; that we can condense the thousands of concepts used to define the most common E-gov applications requirements into a smaller set of concepts (25 concepts). Table 4 shows part of the results of the evaluation.

We have presented the conceptualization of the common REPEAs by the proposed ontology. We also have condensed the semantic of thousands of concepts used to define REPEAs into a smaller set of concepts consists of 25 concepts with a high percentage of coverage reached to 74%. Experts and practitioners in the field of software engineering (SWE) and E-gov can use our concepts (the 1385, the 90, or the 25) in their applications. Experts may use these concepts to define requirements for an e-Gov application. An e-gov expert can compare the concepts which are used in a writing proposal for an e-gov application's requirement with our ontology concepts. If these concepts are mapped with our proposed ontology, we can say that the requirement is strong, meaningful, and consistent. Our proposed ontology could be used as a base knowledge among participants who are developing e-gov applications. This knowledge base could be evolved and eventually we will reach to a common, agreed, and consistent semantic for e-gov applications. REPEAs are currently in the phase in which terminologies, principles, and methods are still being defined, consolidated, and agreed. In particular, there is a lack of consensus on the concepts and terminologies used in the semantic of this field.

5. CONCLUSIONS

In this paper we extracted concepts and terminologies from current e-gov. applications proposals, documents, and reports. We evaluated these concepts using the coverage technique. Sixty eight documents have been reduced to 1358 concepts by KAON's tool, the 1358 concepts have been reduced to 90 by Kayed's tools, and then these 90 concepts have been reduced to 25 concepts by human experts. In our future work, we are planning to provide a formal representation for the ontology, building relationships among these concepts, completing the lattice representation for the relationships, and using other approaches to evaluate and enhance the ontology. Finally we are

planning to conduct set of experiments to deploy the ontology in some applications to see its effectiveness, usefulness, and expressiveness to the audience in the context of software engineering and e-gov applications.

6. **Acknowledgment**

Authors would like to thank Mr. Mohammed Nezar for his effort to collect the data for our experiment. The first author would like to thank Fahad Bin Sultan University (FBSU) for their support for this research.

REFERENCES

[1] Zhiyuan F., "E-Government in Digital Era: Concept, Practice, and Development", International Journal of The Computer, The Internet and Management, Vol. 10, No.2, 2002, p 1-22.

[2] Randell B.; Naur P., "Software Engineering", Report on a conference sponsored by the NATO Science Committee, Garmisch, Germany, 7th to 11th Oct.1968.

[3] Sommerville I., "Software Engineering", Pearson Addison Wesely, Seventh Edition, ISBN: 020139815, 2007.

[4] Samhan A., "Master Thesis: Towards An Ontology for Software Product Quality Attributes", Middle East University for Graduate Studies, Amman Jordan, May, 2008.

[5] Liao L., Yuzhong Qu., Hareton, Leung K.N. "A software process ontology and its application", 2003.

[6] Brazier, F.M.T., Jonker, C.M., and Treur, J. (1998), Principles of Compositional Multi-agent System Development. In: J. Cuena (ed.), Proceedings of the 15th IFIP World Computer Congress, WCC'98, Conference on Information Technology and Knowledge Systems, IT&KNOWS'98, pp. 347-360.

[7] Scacchi W.; Jensen C.; Noll J.; Elliott M.," Multi-Modal Modeling, Analysis and Validation of Open Source Software Requirements Processes", Intern. J. Information Technology and Web Engineering, Vol. 1, Issue 3, PP: 49-63, 2006.

[8] Zhao G., Gao Y., Meersman R, "An ontology-based approach to business modeling" ACM.

[9] Nuseibeh B., Easterbrook S, "Requirements engineering: A Roadmap" ACM.

[10] Lamsweerde A., "Requirements engineering in the year 00: A research perspective", ACM, 2000.

[11] Girardi R., De faria C, "An ontology-based technique for the specification of domain and user models in multi-agent domain engineering, ACM.

[12] Greenspan S., Mylopoulos J., Bordiga A, "On formal requirements modeling languages: RML revisited", Sixth international conference on software engineering (ICSE-6), ACM.

[13] B. Smith,C. Welty, Ontology: Towards a New Synthesis, FOIS'01, October 17-19, 2001, Ogunquit, Maine, USA. Copyright 2001 ACM1-58113-377-4/01/0010.

[14] L. Floridi (ed.), Blackwell Guide to the Philosophy of Computing and Information,Preprint version of chapter "Ontology", Oxford: Blackwell, 2003, 155–166.

[15] T. Gruber,Ontology,to appear in the Encyclopedia of Database Systems, Ling Liu and M. Tamer Özsu (Eds.), Springer-Verlag, 2008.

[16] B. Smith, Ontology,Draft version of chapter published in Luciano Floridi (ed.), Blackwell Guide to the Philosophy of Computing and Information, Oxford: Blackwell, 2003, 155-166.

[17] R. William., S. Pawlowski, V. Volkov, "Requirements interaction management", ACM computing surveys, Vol35, No2, June 2003, pp 132-190.

[18] S. Walt., C. Jensen, J. Noll, M. Elliott,"Multi-Modal Modeling, Analysis and Validation of Open Source Software Requirements Processes" ACM 2005 .

[19] Z. Ganng, Y. Gao, R. Meersman, "An ontology-based approach to business modeling" ACM 2005.

[20] A. Kayed, R. Colomb. Extracting Ontological Concepts for Tendering Conceptual Structures. Eng. 40(1): 71-398(2002).

[21] A. Kayed, R. Colomb. Using BWW Model to Evaluate Building Ontologies in CGs Formalism. Inf.Syst. 30(5): 379-398(2005).

[22] R. Balzer, D. Dayer, M. Feahling, S. Sawnders. Specification Based Computing Environments. VLDB 1982, 237 – 279.

[23] J. Rumbaugh, M. Balha, W. Premelani. Object Oriented Modeling and Design. Printicehall 1991.

[24] H.M. Kim. Representing and Reasoning about Quality using Enterprise Models. PhD thesis, Dept. of Mechanical and Industrial Engineering, University of Toronto, Canada, 1999.

[25] V. Devedzic: Understanding Ontological Engineering. Commune ACM 45(4): 136-144, 2002.

[26] M. de los Martin, L. Olsina. Toward an Ontology for SW metrics and indicators as the Foundation For Catalog Web system. LA – WEB 2003: 103 -113, Santiago, Chile.

[27] Y. Wand and R. Weber. On the Ontological expressiveness of Information Systems analysis and design grammars. Journal of Information Systems, 3(4): 217-237, 1993.

[28] F. Garcia, M. Bertoa, C. Calero. Towards a Consistent Terminology for Software measurements, Information of Software Technology, 48(8): 631-644, 27 June 2005.

[29] Extensions to the Karlsruhe Ontology and Semantic Web Framework, KAON Extensions, Developer's Guide for KAON Extensions 0.6, August 2003.

[30] Maedche A.; Volz R.,"The ontology extraction and maintenance framework text-to-onto".In Proceedings of the ICDM'01 Workshop on Integrating Data Mining and Knowledge Management, 2001.

[31] Kayed A.; Colomb R., "Extracting Ontological Concepts for Tendering Conceptual Structures", Data & Knowledge Engineering, Vol. 40, Issue 1, Pages: 71 - 89, 2002.

[32] Brank J.; Grobelnik M.; Mladenić D., "A Survey of Ontology Evaluation Techniques", in proc. Of the 8th international multi conference information society, SIKDD, 2005.

[33] Parekh V.; Gwo J.; Finin T.; "Mining Domain Specific Texts and Glossaries to Evaluate and Enrich Domain Ontologies", International Conference on Information and Knowledge Engineering, PP: 533-540, USA, CSREA, ISBN 1-932415-27-0, 2004.

[34] Brewster C.; Alani H.; Dasmahapatra S.; Wilks Y., "Data-driven ontology evaluation". InProc. of the 4th International Conference on Language Resources and Evaluation, Lisbon,2004.

[35] V¨olker J.; Vrande D.; Sure Y., "Automatic evaluation of ontologies (AEON)".In Proc. Of the 4th International Semantic Web Conference (ISWC'05), Springer Verlag Berlin-Heidelberg, Vol. 3729, PP: 716-731, Nov. 2005.

[36] Kayed A., "Building e-laws ontology: New approach", Springer, Lecture Notes in Computer Science ,Springer-Verlag Vol. 3762 , Germany, ISBN/ISSN: 0302-9743, pp 826-830,2006.