An Ontology-based Knowledge Management System for Software Testing

Shanmuganathan Vasanthapriyan†, Jing Tian*, Dongdong Zhao†, Shengwu Xiong§ and Jianwen Xiang∗

Hubei Key Laboratory of Transportation Internet of Things
School of Computer Science and Technology
Wuhan University of Technology
Wuhan, P.R. China

Email: priyan@appsc.sab.ac.lk , {jtian*,zdd,xiongsw,jwxiang*}@whut.edu.cn

Abstract—Software testing is an important activity in quality assurance and it generates large amount of knowledge. Software testers need to gather domain knowledge to be able to successfully conduct a software testing activity. Not having a proper knowledge base within its own context by software testing environments cause software testers to query limited knowledge available or consult peer software testers, which would greatly impact on their decision-making process. Ontologies emerge as one of the more appropriate knowledge management tools for supporting knowledge representation, processing, storage and retrieval. Given great importance to knowledge for software testing, and the potential benefits of managing software testing knowledge, using semantic web technologies, ontology based knowledge management system is developed. A Software testing knowledge sharing ontology is designed to describe software testing domain knowledge. SPARQL is used as the query language to retrieve software testing knowledge from the semantic storage. Both Ontology experts and non-experts evaluated the developed ontology. We believe our software testing ontology can support other software organizations to improve the sharing of knowledge and learning practices.

Keywords—software testing ontology, software testing knowledge, ontology based knowledge management system, knowledge sharing.

1  2

I. INTRODUCTION

Software testing is a sub area of software engineering which is also a knowledge intensive and collaborative activity[1][2]. Knowledge can be applied to different testing tasks and purposes. Since software development is an error prone task, to achieve quality software products, Validation and Verification should be carried throughout the development[3]. As software testers, they would be familiar with the several software testing methods and considerably aware of the software development models, need information relevant to their context. For instance, software testers may either need assistance on a test case design information relevant to a similar project which was handled previously for testing purposes or to design a test case. Moreover, this information would also have a greater impact on their decision-making process. Importantly, during the software testing activities a huge amount of knowledge is being continuously produced and consumed. But, the approaches are limited and less employed to capture this knowledge and manage inside the organization to facilitate software testers. Our previous study results revealed some of the issues such as, outdated knowledge in the repositories, unstructured internal documents and varied formats, less accessing facilities and lack of targeted delivery methods. Hence, efficient mechanisms for capturing, representing, reusing, and sharing the software testing knowledge involved are sorely needed.

According to Gruber[4], an ontology is an explicit specification of a conceptualization. Ontology provides a structured view of domain knowledge and acts as a repository of concepts in the domain. Recently, ontologies and Semantic Web technologies have received more attention and been gradually used in the knowledge representation[5]. Given great importance to knowledge for software testing, and the potential benefits of managing software testing knowledge, using semantic web technologies, ontology based knowledge management system is developed.

The context has also been decided to confine the study to a particular Sri Lankan software development company. The key reasons are based on the geographical location of the researcher, practicality and ease of access to those software development companies and comparability of research data due to company’s same jurisdiction, same economic and regulatory regimes governing their operation. Further, we briefly explain each type of the high-level concepts based on IEEE 829-2008[6], also known as the 829 Standard for Software and System Test Documentation and ISTQB (International Software Testing Qualifications Board)[7]. Even though the standard specifies the procedures of software testing, we have also included what software company is stipulating in it’s practice.

The remainder of this paper is organized as follows. Succinct analysis of related research is presented in the second section. Section 3 discusses software testing ontology in detail. Development of knowledge management portal to manage software testing knowledge is discussed in Section 4. Section 5 discusses the evaluation of the ontology developed under two points-of-view: domain experts and non-experts. Finally, Section 6 concludes the paper and presents directions for future work.

1DOI Reference Number: 10.18293/SEKE2017-020
2*Corresponding Authors
II. RELATED WORK

A number of overviews of work on knowledge management in software testing have previously been published. The research carried out by Wei and Ying[8] discussed the proposal of implementing test knowledge management framework in iDEN phone software system testing and how the knowledge management approach can benefit the testing team in terms of cost and productivity. A reusable test case knowledge management model is proposed by Li and Zhang [9] to support the knowledge reuse based on the ontology representation of reusable test cases so that the test engineers can retrieve and reuse test cases flexibly. Douglas [10] proposed to investigate an open-standard based approach to the sharing of test results and it requires some means to represent and to associate such digital objects. Such an approach would not only reduce the needless replication of tests that occurs when there are no public records of previous tests conducted, but it would also allow the accumulation of a good deal of evidence to support certain usability design patterns and guidelines. Reference Ontology on Software Testing (ROoST) which was developed managing relevant knowledge to reuse is difficult and it requires some means to represent and to associate semantics to a large volume of test information[11]. Bajnaid[12] proposed an SQA ontology that represents both domain and operational knowledge which provides consistent support to communicate between people and software agents but does not eliciting anything related to software testing process.

III. DESIGNING OF SOFTWARE TESTING DOMAIN ONTOLOGY

A. Describing the Software Testing Process

Software testing process contains TestPlanning for planning tests, TestCaseDesign for test case construction, and test execution TestExecution for execution of test cases and producing ActualResult, and TestResultAnalysis for analysis and evaluate the test results [13]. In addition, our software testing process includes Test Design Techniques, Test Levels, Artifacts, TestEnviroment ( includes hardware, software and Human resources) and Static Testing Techniques.

A TestPlan is produced during the TestPlanning activity. TestCaseDesign activity targets to design a TestCase. During the TestCaseDesign, if a Test Design Techniques is used then the following axiom can be defined for TestCaseDesign to design any TestCase as follows.

\[ \forall tc:TestCase, tdt:TestDesignTechnique, tcd:TestCaseDesign \] 
\[ \text{hasDesignAccordingTo}(tc, tdt) \text{ and } \] 
\[ \text{isGeneratesTestCase}(tcd, tc) \rightarrow \text{isAdopts}(tcd, tdt) \]

Several Artifacts have been used to derive test cases in software testing which describes the functionalities, architecture, and design of software. Such Artifacts are used as TestCaseDesignInput during the software TestCaseDesign activity and the output is test cases. Test cases can be documented as described in the IEEE 829-2008 documentation. The document that describes the steps to be taken in running a set of tests (and specifies the executable order of the tests) is called a test procedure in IEEE 829. Besides, Test case contains a set of input values (TestCaseInput), execution preconditions, expected results (Expected Result) and execution post conditions, developed for a particular objective or test condition, such as to exercise a particular program path or to verify compliance with a specific requirement.

That is, a test case targets to test a Code To be Tested . Code To be Tested can be any programs, modules, and the whole system code. Furthermore, Test Code is a portion of code that is to be run for executing a given set of test cases, contain three subtypes such as Test Script, Driver and Stub.

Test Execution activity executes any Test case. Test Execution requires as input the Test Code to be run and the Code To Be Tested. Notably, both Test Code and Code To Be Tested are needed for Test Execution activity. This can be illustrated in an axiom as follows.

\[ \forall te:TextExecution, tc:TestCase \] 
\[ \text{hasExecutesTestCases}(te, tc) \rightarrow (\exists tcode:TestCode, CodeTobeTested: CodeToBeTested) \] 
\[ \text{uses}(te, tcode) \land \text{isContainedOf}(tcode, tc) \land \text{uses}(te, CodeTobeTested) \land \text{hasCodeTobeTested}(tc, tcode) \]

The output of the Test Execution activity is the Test Result. Each of the Test Result is particularly related to a TestCase. Test Result may be related to an Actual Result, for a particular TestCaseInput and Expected Result. Test Execution also can be elaborated necessary and sufficient axioms in Protégé 5.1 as follows.

\[ \text{hasExecutesTestCases only}(TestCase \text{ and } \text{hasTestCaseExpectedResult some ExpectedResult}) \]

A test execution can run and achieve a result (Actual Result), but it can also fail, and generating a Test Issue. Thus, a Test Result contains either an Actual Result, or Test Issue or both. This can be expressed using the following axiom.

\[ \forall TestR:TestResult \rightarrow (\exists ActResult:ActualResult, Issue: TestIssue) (ActualResult(ActResult) \lor (TestIssue(Issue)) \land \text{isGeneratesTest}(ActResult, Issue)) \]

Further, Test Issue Report, could report this event in detail, which requires investigation. Finally, during a Test Result Analysis, Test Results are analyzed and a Test Analysis Report is produced.

B. Ontology Engineering Approach

Ontology engineering approach investigates the principles, methods and tools for initiating, developing and maintaining ontologies[14]. In literature, many methodologies have been proposed until now to build an ontology[15], [16], [17], but we considered the Grüninger and Fox methodology[18], which considered a formal approach to design ontology as well as providing a framework for evaluating the adequacy of the
developed ontology. This methodology focuses on building ontology based on first-order logic (FOL) by providing strong semantics. In our scenario, we introduce Description Logic (DL) which is a decidable fragment of FOL since we are designing with OWL 2 Web Ontology Language [19] for semantic web. The widely-used Protégé system (http://protege.stanford.edu) has recently been extended with support for the additional constructs provided by OWL2 [20].

C. Contextual Information

We describe "context specific" [21] to the software testers belonging to a leading software company in Sri Lanka and the approach which will be used to design the ontology to provide context-specific information and knowledge to software testers. To identify software tester’s context clearly, we have extracted domain specific knowledge of software testing ontologies from existing literature and interviewed the software testing experts belonging to a particular company.

D. Competency Questions (CQs)

Competency Questions (CQ) are a set of questions that the ontology must be capable of answering using its axioms [18]. Thus, these CQs work as requirement’s specification of the software testing ontology. With a set of CQs at hand, it is possible to know whether an ontology was created correctly, if it contains all the necessary and sufficient axioms that correctly answer the CQs. Some of the CQs used are shown in Table III.

### Table I

<table>
<thead>
<tr>
<th>Tester’s Information Needs in Context (i.e. Competency Questions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>List out the human resources available in a testing activity?</td>
</tr>
<tr>
<td>Suitable test case design techniques for a given scenario?</td>
</tr>
<tr>
<td>What are the Test Levels described in software testing?</td>
</tr>
<tr>
<td>To which Test levels a particular TestDesignTechnique could be applied?</td>
</tr>
<tr>
<td>What kind of test automated software tools are available for use in your testing process?</td>
</tr>
<tr>
<td>Which Suitable Hardware Resources available for a given TestPlan?</td>
</tr>
<tr>
<td>What is the project name in which particular testing activity occurred?</td>
</tr>
<tr>
<td>What are the testing artifacts used in a testing activity?</td>
</tr>
<tr>
<td>What are the testing artifacts produced in a testing activity?</td>
</tr>
<tr>
<td>What are the testing objects used in a testing activity?</td>
</tr>
</tbody>
</table>

IV. DEVELOPMENT OF SOFTWARE TESTING KM PORTAL

As mentioned before, ontologies are powerful mechanism for representing knowledge presented in semantic web. The ontology and semantic web technologies provide powerful reasoning capabilities. In this section, we describe the ontology based knowledge management system to manage software testing knowledge, built upon the Java J2EE distributed component environment.

### Table II

<table>
<thead>
<tr>
<th>Necessary and Sufficient Conditions Written in Protégé 5.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing Activities</td>
</tr>
<tr>
<td>Axioms written in Protégé 5.1 for Equivalent Classes</td>
</tr>
<tr>
<td>TestCaseDesign</td>
</tr>
<tr>
<td>(isGeneratesTestCase only TestCase) and (isGeneratesTestCase min 1 TestCase)</td>
</tr>
<tr>
<td>TestExecution</td>
</tr>
<tr>
<td>hasExecutesTestCases only (TestCase and (hasTestCaseExpectedResult some ExpectedResult))</td>
</tr>
<tr>
<td>TestPlanning</td>
</tr>
<tr>
<td>(hasCreatesTestPlan some TestPlan) and (hasCreatesTestPlan only TestPlan)</td>
</tr>
<tr>
<td>TestResultAnalysis</td>
</tr>
<tr>
<td>hasEvaluatesTestCases some (TestCase and (hasTestActualResult some ActualResult))</td>
</tr>
</tbody>
</table>

Our KM Portal design consists of Experience Sharing, Ontology, Knowledge Retrieval, Storage and Reasoning Layer and is shown in Figure 1. The class diagram of Knowledge Sharing Portal is shown in Figure 2. We have implemented as a starting point of specifying the Knowledge Management Portal. Importantly, our presented concepts, properties and
relationships here are identified according to the characteristics of the particular organization’s software testing environment.

A. Ontology Layer

The ontology layer has our ontology which includes domain rules, axioms etc. Using the Protége Ontology Editor 5.1, these concepts and their relationships were partly described in section III.

B. Experience Sharing Layer

Through the Experience Sharing layer, software testers can annotate their testing knowledge with the support of ISTQB and IEEE 829-2008 terms. Once the knowledge is shared, such knowledge is transformed by the semantic data generator into the semantic data in a machine-understandable format of triple structure.

C. Storage Layer

We used Triple-store, which stores RDF triples and are queried using SPARQL. Jena TDB[22] has been selected to use in this study because it is a component of Jena for RDF storage and query. Importantly, it supports the full range of Jena APIs and can be used as a high performance of RDF store on a single machine.

D. Reasoning Layer

The Semantic Web Rule Language (SWRL) is based on a combination of OWL with the Rule Markup Language which provides inference capabilities[23] from existing OWL ontology. Rules in SWRL reason about OWL instances in terms of OWL classes and properties. Importantly, such rules express more complex relationships and restrictions between concepts. Software Testing rules were generated with Protége SWRL Editor that is a plugin in Protége environment and with the support of the Jess Rule Engine.

E. Knowledge Retrieval Layer

To show how our ontology can be used to share software testing knowledge collected from software testers, the Knowledge Retrieval Layer includes two functionalities that use Semantic Web technologies: (1) basic search, and (2) Advanced Search. SPARQL(SPARQL Protocol And RDF Query Language) has been used as the query language to retrieve software testing knowledge from the semantic data storage. The basic search provides a simple triple pattern matching service, which is one of the most frequently used functions for searching documents in the Semantic Web. Besides, Advanced Search Option includes, logical operators (AND or NOT or OR), so that user can combine different options to retrieve knowledge.

V. Ontology Evaluation

The quality of an ontology should be verified and validated before it is used in practice to avoid defects[24]. Further, such validation process will prevent contain ontology with anomalies or pitfalls, inconsistent incorrect and redundant information. Our evaluation of the ontology is carried out in three methods such as internal(using reasoners and OOPS!), ontology expert method and non-expert methods. Importantly, Protége inbuilt reasoners such as FaCT++ 1.6.5 and HermiT 1.3.8.413 were used to check the internal consistency and inferences. OOPS!( http://oops.linkeddata.es/) is an online ontology evaluator has been used for our context to detect potential pitfalls that could lead to modelling errors, before ontology has been deployed in the end-user application. This method evaluates human understanding, logical consistency, modelling issues, ontology language specification, real world representation and semantic applications from the developed ontology[24]. Table III summarizes the pitfalls encountered, along with a brief description and the way each one of them was handled. There were three levels such as Critical, Important and Minor. Critical level is very crucial and it must be corrected in order to avoid ontology inconsistency. To make ontology nicer both Minor and Important cases were corrected.

Table III: Pitfall Description and Solution Proposed

<table>
<thead>
<tr>
<th>Pitfall</th>
<th>Description</th>
<th>Solution Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing Annotation (256 cases</td>
<td>Ontology terms lack annotation properties that could improve the understanding of the ontology</td>
<td>Included the ontology annotation properties</td>
</tr>
<tr>
<td>Missing domain or range in properties (7 cases</td>
<td>Relationships and (or) attributes without domain or range are included.</td>
<td>Restored missing domains.</td>
</tr>
<tr>
<td>Missing Inverse Relationships (14 cases</td>
<td>When a relation has non-inverse relationship defined</td>
<td>Included Missing inverse relationships</td>
</tr>
<tr>
<td>Defining wrong inverse relationships (6 cases</td>
<td>Relationships are defined as inverse relations when they are not necessarily inverse.</td>
<td>They were removed to improve the expressiveness</td>
</tr>
</tbody>
</table>

Expert evaluation activity is performed by two ontology experts who have a good understanding of software testing. Even though there were many methods discussed to evaluate ontologies, our ontology experts considered Vocabulary, Syntax, Structure, Semantics, Representation and Context to perform evaluation[25]. The main objectives of expert evaluation are (a) whether the software testing ontology meets its requirements, standards, representation of concepts, relationships among them (b) coverage of the software testing domain and (c) checking for internal consistencies. The following suggestions and improvements were highlighted by the ontology experts: (a) Need of Annotations, renaming of concepts to standard methods. (b)Improving the relationship names, removing redundant relationships, some of the relationships were not mentioned, adding of association relationships which were missing. All such suggestions, comments were taken into consideration and implemented in the software testing ontology.

To carry out our first industrial application based evaluation from software testing experts, the developed Knowledge Management Portal was hosted locally inside one software company in Sri Lanka. A separate questionnaire was designed.
To what extent is the software testing process incorporated?

Results illustrates that responded software testers have different views about the software testing process.

To what extend is the software testing process incorporated?

More than (75%) of software testers agreed that in the knowledge repositories is precise.

Is the knowledge in the knowledge repositories precise?

The large majority (66.67%) of the respondents from group A believed (Strongly Agree or Agree) that the provided content in Web Portal is precise. Only 44.44% of the respondents from group B believed (Strongly Agree or Agree) that in the knowledge repositories is precise.

To what extend is the software testing process incorporated?

Results illustrates that responded software testers have different views about the software testing process.

To what extend is the software testing process incorporated?

More than (75%) of software testers agreed that in the knowledge repositories is precise.

Is the knowledge in the knowledge repositories precise?

The large majority (66.67%) of the respondents from group A believed (Strongly Agree or Agree) that the provided content in Web Portal is precise. Only 44.44% of the respondents from group B believed (Strongly Agree or Agree) that in the knowledge repositories is precise.

To what extend is the software testing process incorporated?

Results illustrates that responded software testers have different views about the software testing process.

To what extend is the software testing process incorporated?

More than (75%) of software testers agreed that in the knowledge repositories is precise.

Is the knowledge in the knowledge repositories precise?

The large majority (66.67%) of the respondents from group A believed (Strongly Agree or Agree) that the provided content in Web Portal is precise. Only 44.44% of the respondents from group B believed (Strongly Agree or Agree) that in the knowledge repositories is precise.

To what extend is the software testing process incorporated?

Results illustrates that responded software testers have different views about the software testing process.

To what extend is the software testing process incorporated?

More than (75%) of software testers agreed that in the knowledge repositories is precise.

Is the knowledge in the knowledge repositories precise?

The large majority (66.67%) of the respondents from group A believed (Strongly Agree or Agree) that the provided content in Web Portal is precise. Only 44.44% of the respondents from group B believed (Strongly Agree or Agree) that in the knowledge repositories is precise.

To what extend is the software testing process incorporated?

Results illustrates that responded software testers have different views about the software testing process.

To what extend is the software testing process incorporated?

More than (75%) of software testers agreed that in the knowledge repositories is precise.

Is the knowledge in the knowledge repositories precise?

The large majority (66.67%) of the respondents from group A believed (Strongly Agree or Agree) that the provided content in Web Portal is precise. Only 44.44% of the respondents from group B believed (Strongly Agree or Agree) that in the knowledge repositories is precise.

To what extend is the software testing process incorporated?

Results illustrates that responded software testers have different views about the software testing process.

To what extend is the software testing process incorporated?

More than (75%) of software testers agreed that in the knowledge repositories is precise.

Is the knowledge in the knowledge repositories precise?

The large majority (66.67%) of the respondents from group A believed (Strongly Agree or Agree) that the provided content in Web Portal is precise. Only 44.44% of the respondents from group B believed (Strongly Agree or Agree) that in the knowledge repositories is precise.

To what extend is the software testing process incorporated?

Results illustrates that responded software testers have different views about the software testing process.

To what extend is the software testing process incorporated?

More than (75%) of software testers agreed that in the knowledge repositories is precise.

Is the knowledge in the knowledge repositories precise?

The large majority (66.67%) of the respondents from group A believed (Strongly Agree or Agree) that the provided content in Web Portal is precise. Only 44.44% of the respondents from group B believed (Strongly Agree or Agree) that in the knowledge repositories is precise.

To what extend is the software testing process incorporated?

Results illustrates that responded software testers have different views about the software testing process.

To what extend is the software testing process incorporated?

More than (75%) of software testers agreed that in the knowledge repositories is precise.

Is the knowledge in the knowledge repositories precise?

The large majority (66.67%) of the respondents from group A believed (Strongly Agree or Agree) that the provided content in Web Portal is precise. Only 44.44% of the respondents from group B believed (Strongly Agree or Agree) that in the knowledge repositories is precise.

To what extend is the software testing process incorporated?

Results illustrates that responded software testers have different views about the software testing process.

To what extend is the software testing process incorporated?

More than (75%) of software testers agreed that in the knowledge repositories is precise.

Is the knowledge in the knowledge repositories precise?

The large majority (66.67%) of the respondents from group A believed (Strongly Agree or Agree) that the provided content in Web Portal is precise. Only 44.44% of the respondents from group B believed (Strongly Agree or Agree) that in the knowledge repositories is precise.
two software testing groups working on a similar information system project for comparison and in depth understanding. Importantly, a prior training session was conducted through recorded video to make software testers familiar with the Knowledge Management Portal.

The results are summarized in Table IV. When asked to evaluate the content of the KM Portal, Group B believes that correct contents have been included than the Group A. All of the software testers participated believed that the standards have been maintained, but a few have included some extra comments, such as inclusion of some vocabularies, parameters in Test Level and Test Types. Moreover, when asked about their satisfaction on the factors related to user manipulation, personalization and knowledge community, both groups have mostly responded neither satisfied nor unsatisfied. Notably, a very few have been very much satisfied with the user manipulation or knowledge community of the KM Portal. To conclude, overall 80.95% (See Figure 3) of the respondents would like to recommend the use of such KM Portal among the software testers for knowledge sharing and knowledge retrieval.

VI. DISCUSSION AND CONCLUSION

This research presents software testing ontology to represent software testing domain knowledge which includes software testing concepts, properties and their relationships. The implemented an ontology-based KMS based on semantic web technologies provides facilities for software testers to share their knowledge and experiences. Basic search and advanced search operations are used to retrieve knowledge. The implemented ontology was validated and evaluated before it is used. Ontology experts opinion was received to improve the ontology. The results from the industrial experimental investigation show that the proposed ontology-based KM Portal is adequate to support knowledge sharing and reuse, allowing: (a) knowledge representation and organization; (b) distributed knowledge inference and retrieval; (c) management of organizational knowledge on software testing. Our Portal addresses the existing key issues such as knowledge is not reaching the software testers due to its unstructured, incomplete, varied formats, and lack of targeted delivery methods. We believe our software testing ontology can support other software organizations to improve the sharing of knowledge and learning practices. In the future work, the reasoning engine with Query-enhanced Web Rule Language (SQWRL) will be incorporated into knowledge searching to support more precise and effective knowledge sharing.

ACKNOWLEDGMENT

The authors would also like to thank the anonymous referees for their valuable comments and helpful suggestions. This work was partially supported by the National Natural Science Foundation of China (Grant No. 61672398), the Key Natural Science Foundation of Hubei Province of China (Grant No. 2015CFC069), and the Applied Fundamental Research of Wuhan (Grant No. 20160101010004).

REFERENCES