Ontologies for Software Engineering: Past, Present and Future

M. P. S. Bhatia¹, Akshi Kumar² and Rohit Beniwal*¹

¹Division of Computer Engineering, Netaji Subhas Institute of Technology, Dwarka – 110078, Delhi, India; mpsbhatia@nsit.ac.in, rohitbeniwal@yahoo.co.in
²Department of Computer Engineering, Delhi Technological University, Delhi - 110042, India; akshi.kumar@gmail.com

Abstract

Background/Objectives: Research in recent years has probed integration amongst research field of Software Engineering and Semantic Web technology, addressing the advantages of applying Semantic techniques to the field of Software Engineering. Prolifically published studies have further substantiated the benefits of ontologies to the field of Software Engineering, which clearly motivate us to explore further opportunities available in this collaborated field. This paper is a survey expounding such opportunities while discussing the role of ontologies as a Software Life-Cycle support technology.

Method/Statistical Analysis: Survey centred on providing an overview of the state-of-art of all the ontologies available for Software Engineering followed by their categorization based on software life cycle phases and their application scope.

Findings: Characterization of ontologies as a Software Life-cycle support technology, instigated by the increasing need to investigate the interplay between Semantic Web & Software Engineering with the ultimate goal of enabling & improving Software Engineering capabilities.

Application/Improvements: This paper discusses the practical and potential applications of ontologies in the field of Software Engineering followed by the issues and challenges that will keep this field dynamic and lively for years to come.

Keywords: Life Cycle, Ontologies, Semantic Web, Semantic Web Enabled Software Engineering, Software Engineering, Support Technology

1. Introduction

“Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries”¹. It is a machine processable “Web of Data”². It is well recognized as an effective means of improving visibility of knowledge on the Web. At the core of Semantic Web is ontology that is used to explicitly represent our conceptualizations. Ontologies are built to model a domain and support reasoning over the concepts. Ontology engineering in Semantic Web is primarily supported by languages such as RDF, RDFS and OWL³.

Software Engineering is “the application of a systematic, disciplined, quantifiable approach to the development, operation, and maintenance of software, and the study of these approaches; that is, the application of engineering to software”⁴. Software development is a multifaceted task with new challenges being imposed frequently. The problem manifolds owing to overwhelming demands of the customers, numerous participants, shorter time frames, different geographical and virtual locations of software development teams, information overload and related issues. Sharing and reusing available information saves the effort in development and maintenance of software systems. As a major operative challenge this fosters the need to explore strategic and supporting technologies.

Recent studies have publicized the collaboration among research fields of Software Engineering (SE) and Semantic Web Technology (SW), which demonstrate the benefits of integrating semantic techniques with Software Engineering⁵-⁹. Ontologies have emerged as a key player in this direction⁷,¹⁰,¹¹. A rising trend to exploit ontologies to exchange and interconnect Software Engineering knowledge across Web has been recognized and accepted.

* Author for correspondence
by the Software Engineering community. This team-up of Software Engineering with Semantic Web has further drawn attraction of standardization bodies and opened up new research avenues. Ontology-Driven Architecture (ODA), which is W3C’s Software Engineering Best Practices Working Group’s effort, tries to bring out best practices for using ontologies in Software Engineering. Ontology Definition Metamodel (ODM) standard of the Object Management Group (OMG)'s permits integrating web ontology languages (i.e. ontologies) into the software development process that is based on model-driven engineering principles. Thus, motivation to lead from study-based research to acceptance in tool and practice is abundant.

In this paper, we characterize ontologies as a Software Life-cycle support technology, instigated by the increasing need to investigate the interplay between Semantic Web and Software Engineering with the ultimate goal of enabling and improving Software Engineering capabilities. This paper studies the role of ontologies in the field of Software Engineering in the past and present. It further discusses the practical and potential applications, followed by the issues and challenges that will keep the field dynamic and lively for years to come.

The rest of this paper is organized as follows: section 2 briefly discusses the core software life cycle phases followed by section 3 that demonstrate the use of ontologies as software life-cycle support technology depicting their phase-wise correspondence. Section 4 expounds ontologies and their application scope; section 5 illustrates the applications of ontologies in various Software Engineering areas; section 6 discusses the research directions for future work and finally section 7 concludes the paper.

2. Software Life Cycle Phases

The goal of this section is to describe the software life cycle phases, artifacts used and produced in them, and various modeling approaches that can be used during life cycle phases.

The most commonly used definition of software life cycle is the one given in the IEEE Standard Glossary for Software Engineering, where the software life cycle is defined as “The period of time that starts when a software product is conceived and ends when the product is no longer available for use. The software life cycle typically includes a requirements phase, design phase, implementation phase, test phase, installation and checkout phase, operation and maintenance phase, and sometimes, retirement phase”. This section discusses the imperative software life cycle phases, which are as follows:

2.1 Requirements Analysis and Specification Phase

This phase describes the requirements of a software system. The requirements describe “what” of a system, not the “how”. Once the requirements are gathered, they are analyzed for their validity. The possibility of incorporating them into the system to be developed is also studied. Finally the resultant document known as Software Requirement Specification (SRS) is developed, which may act as a legal document between the developer and customer. This process of gathering requirements from user and to document them into SRS is known as requirement engineering.

The SRS should be correct, complete, consistent, unambiguous, verifiable, modifiable, and traceable. Different approaches such as viewpoints-oriented, goal-driven, and scenario based approaches, or their combinations are used for the purpose of requirement engineering. Modeling approaches such as UML Use Cases and class diagrams are recommended for this phase. Formal approaches such as Petri nets are also recommended by some researchers.

2.2 Design Phase

This phase transforms the SRS into a structure, which can be implemented in some programming language. Here detail designs are produced for application domain, architecture, software components, interfaces and data. UML modeling approach is used for design purpose that results into a document known as Software Design Document (SDD). Object Constraint Language (OCL), which is a textual, declarative language for specifying (more formal) constraints and queries further supplements UML.

The more formal the designs are, the better the chance of automatic implementation. That’s why Software Engineering community is putting a lot of effort to the discipline called Model-Driven Engineering (MDE) to empower the Model-Driven development of software products.
2.3 Implementation Phase
This phase implements the SDD into some programming language that develops a software product. During this phase, various components are generated, which are later integrated into a final product.

2.4 Testing Phase
This phase tests the various components generated during implementation phase. It includes various levels of testing such as unit testing, integration testing, and system testing. It also includes regression testing as well as acceptance testing before the product is finally delivered to the customer. This phase is partially parallel with implementation phase. Model based testing is used to check whether implementation fully adhere with respect to models.

2.5 Operation and Maintenance Phase
This phase starts when the software is finally delivered to the customer and the users start using it. “Software Maintenance is a broad activity that includes error corrections, enhancements of capabilities, deletion of obsolete capabilities and optimization.” This phase highly depends on the quality of documentation in order to completely understand the software system for maintenance purpose. That’s why it is not surprising to see that 40-60% of software maintainer’s time is just spent in order to understand the software system being maintained. Any change in the software system should also lead to the change in the documentation manual for easy and better maintenance.

3. Ontologies: A Software Life-Cycle Support Technology
Pertinent literature addresses the potential advantages of applying Semantic Web techniques to enable new Software Engineering capabilities. While going through various studies, we identified twenty ontologies that have shown potential in the field of Software Engineering. The following sub-sections expound these ontologies. In each sub-section, we discuss a software activity and the related ontologies focusing on their definition and state-of-the-art. Here we would also like to mention that some issues which are resolved using multiple ontologies are discussed within the scope of last ontology dealing with it. Thus the state of art for previous ontologies has been limited to avoid repetition and missing links.

3.1 Upper Ontology
Definition: Upper ontology defines very general concepts that are identical across all the knowledge domains. These knowledge domains share common concepts form upper ontology. This ontology provides the semantic interoperability to all other ontologies that lies beneath it.

State of the art: There are several ontologies available that adheres to the properties of upper ontology. Example are Bunge-Wand-Weber (BWW) ontology, Business Objects Reference Ontology (BORO), Upper Mapping and Binding Exchange Layer (UMBEL) ontology, Unified Foundational Ontology (UFO), IDEAS, WordNet, CIDOC object-oriented Conceptual Reference Model (CRM) ontology, COmmon Semantic MOdel (COSMO) ontology, Object-Centered High-level Reference (OCHRE) ontology, Gist, MarineTLO, Yet Another More Advanced Top Ontology (YAMATO), Basic Formal Ontology (BFO), Cyc, Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE), General Formal Ontology (GFO), PROTo ONtology (PROTON), Sowa’s ontology, and Suggested Upper Merged Ontology (SUMO). A comparative study of last seven ontologies has been examined by.

3.2 Software Process Ontology
This sub-section discusses software process ontology, which describes the various processes that can be used for software development.

3.2.1 Software Process Ontology
Definition: Software process ontology defines the various software lifecycle models and their phases that are used to develop the software. Each phase can be defined by the sequence of activities. Waterfall Model, Prototype Model, Incremental Process Model, Iterative Enhancement Model, V Model, Rapid Application Development (RAD) Model, Evolutionary Development Model, Spiral Model, Agile Model, Unified Model, etc., are few examples of software life cycles models that we generally study.

State of the Art: While going through several process ontologies proposed by different authors, we found that each one has used different terminology by defining a
variable set of terms. Also, sometime the same term is used by different authors, but, has a different meaning, while different terms used by diverse authors, have same meaning. Such variation cannot be covered by one process ontology. Hence it is proposed to develop a general process ontology, which can be further extended to cover different existing process models. International standards such as CMMI and ISO/IEC 15504 can be referenced for developing such process models.

While worked upon implementing the process ontology on the software life cycle models, implemented the same ontology on software project planning activities. A review on ontologies for software project management process has been done by authors. Our study in this paper is limited to the life cycle models and their phases only.

3.3 Domain Ontologies
This sub-section illustrates application domain and application domain feature ontologies that can be used for specifying domain information of the software system being developed.

3.3.1 Application Domain Ontology
Definition: Application domain ontology represents the knowledge of a particular domain (for example, automobile sector domain ontology will represent the knowledge about automobile sector only) and the business information required for developing software applications in a particular domain. This ontology also defines the various relationships that exist between different concepts related to the domain. While describing relationship between different concepts, they are defined from the real world perspective.

State of the Art: To the best of our knowledge, we haven't seen this ontology for every domain. We found that most authors who have already built this ontology have just provided a module or only a section of it is implemented.

Literature survey further demonstrated that application domain ontology, once created, will be able to suggest elicitation questions pertaining to functional requirements, which in turn will provide help in requirement engineering.

Once all the functional requirement questions are elicited, the application domain ontology can be used to measure the quality of requirements. They used ontology reasoning with closed world assumption in detecting and correcting incomplete and inconsistent requirements.

According to, a transformation approach can be established in order to automatically generate source code from application domain ontology. Then any change in this ontology will also help in predicting the changes required in requirements specification. Hence, this ontology can serve as a base for change prediction in requirements specification.

3.3.2 Application Domain Feature Model Ontology
Definition: A feature is defined as a characteristic of a software system. Therefore, a software system can have lots of features. We can develop various products that share common features along with some variations. It means that all such products will have some common features along with some variations and they belong to the family of related programs. Hence, these products share common architecture with addition of few different features. This addition makes them a distinct product. Therefore a feature can be seen as an additional functionality to a given program. Feature Model is used to represent the commonalities and differences between software features in order to represent the variability within the software. Application domain feature model ontology models the features of software in the same application domain.

State of the Art: According to, this ontology can be used in Semantic Web based approach to provide the design guidance. This ontology also maintains the information features that shows which one is mandatory, optional, alternate, or, requires, and excludes.

Due to the large no. of features owned by software system (which could be up to a few thousand), and increased complexity of dependencies and relations between these features, the process of expeditiously insuring the correctness of the feature models representing the software system becomes difficult. Moreover the growth in size calls for creation of distributed feature models. In this regard, provided a framework that represents, integrates distributed feature models and validates them using OWL and Semantic Web Rule Language (SWRL).

3.4 Requirement Ontology
This sub-section expounds system behaviour ontology, which describes the requirements specified by the
customer. Requirements are of two types: functional and non-functional. Functional requirements can be described as the sequence of actions that a system performs. Non-functional requirements represent the quality related aspects of software system. They stipulate how well the software does, what it has to do.

3.4.1 System Behaviour Ontology

Definition: System behaviour ontology models the behaviour of the software system i.e. what actions a system is going to perform under a particular scenario. It also defines the conditions that need to be met before an action performs and will tell the status of the system once the action is performed\cite{42,56}.

State of the Art: As we know that UML Use Cases are used to capture behavioural knowledge of the software system such as who interacts with it, for what purpose one interacts with it, and what will be the outcome after such interactions without knowing internal details. As far as Use Case is concerned, the system is treated as ‘black box’ i.e. all the information that is captured belongs to the outside of the system. It means Use Cases are used to define overall required behaviour of the software system. In this regard, \cite{42} raised an issue to create a large library of Use Cases using system behaviour and application domain ontologies. While creating the library, we can store semantic information about Use Cases. Once the library is built, Semantic Web based approach can be used to retrieve them based on their semantic information.

3.5 Architecture and Design Ontologies

This sub-section examines software architecture, application logic, and object oriented design ontologies, which describe the architecture and design description of software system.

3.5.1 Software Architecture Ontology

Definition: Software architecture ontology models the architecture related concepts such as architecture styles (structure of the system), elements or components of the software system, element’s properties and their relationship\cite{57}.

3.5.2 Application Logic Ontology

Definition: Application logic ontology can be used to define logic (reason) behind the application behaviour. It defines the reasons for a particular behaviour performed by the software system.

While System behaviour ontology treats the system as ‘Black Box’, Application Logic Ontology treats it as ‘White Box’, which means, this ontology deals with the reason behind a particular behaviour and knows how and why the system is behaving in that way. This ontology can define and model diagrams such as data flow diagram, flow chart diagram, activity diagram, sequence diagram, class diagram state chart diagram, object diagram, component, deployment diagram, etc.\cite{58}.

3.5.3 Object Oriented Design Ontology (OO Design Ontology)

Definition: OO design ontology can be used to define the vocabulary for describing object oriented design principles. This ontology includes the concepts such as \textit{class}, \textit{subclass}, \textit{method or member function}, \textit{interface}, \textit{objects}, etc. It also includes relationships such as \textit{inheritance}, \textit{realization}, etc.\cite{59,60}.

3.6 Pattern Ontology

This sub-section discusses pattern ontology, which describes several kinds of software patterns. Software patterns capture the best practices of system design. There are three key characteristics of design pattern. First, it promotes design reuse and provides guidance. Second, it conforms to a literary style. Third, together, they define a vocabulary for discussing design\cite{61}. Presently, there are several kinds of patterns such as Gang of Four (GoF) design pattern, workflow pattern, web application pattern, and usability pattern. Semantic web techniques can be used to formally describe the patterns\cite{59,62}.

3.6.1 Pattern Ontology

Definition: This ontology provides a library of patterns that are commonly used to create the software system. This ontology can consist of design pattern, workflow pattern, web application pattern, usability pattern, etc.\cite{63,65}.

State of the Art: \cite{66} discussed about design pattern of ontologies. According to them, ontologies also have lifecycle phases like the softwares have. Ontologies are also designed, implemented and maintained like the softwares are. They considered ‘Reusability’ as one of the
most challenging and neglected area in ontology’s design process. They classified the design pattern of ontologies into six categories namely Structural, Correspondence, Reasoning, Presentation, Lexico-Syntactic, and Content Ontology Design Pattern (ODP). They specially focus on Content ODP that describes the conceptual pattern rather than logical. Content ODP helps in transforming Use Cases into design solutions and discusses the relationship between them. They also set a web portal, which discusses about Content ODP known as ontologydesignpatterns.org as a part of their research work which provides in-depth detail.

Also worked on Content Ontology Design Patterns (ODPs). They experimented with Content ODPs in order to show their advantages and it is found that Content ODPS are acknowledged beneficial by ontology developers. The quality of ontology improves in terms of task coverage, usability aspects, and fewer modeling mistakes when Content ODPS are reused. The ontology developers also perceived that modeling was made simpler when Content ODPS were used, but the hypothesis that it accelerates the development process wasn’t sustained. Instead, it was felt that though the development process is slow at the beginning, but it results in better constructed ontologies.

Usage of Content ODP is found in different domains and for different tasks. For example, Content Ontology Design Pattern called 'Participation' is found in different domain ontologies covering enterprise models, legal norms, software management, biochemical pathway, and fishery techniques.

Found the decent use of ODPS in a repository of biomedical ontologies. According to them, ODPS in biomedical ontologies may be due to age of ontologies in repository, deficiency of proper tooling, lack of user’s knowledge and education about them, specificity of domain and a tension between highly logical expressivity of ODPS with the objective of maintaining low expressivity of biomedical ontologies.

In the reference of pattern ontologies, provides description about Usability pattern, defines ontology for Web Application pattern and describes about ontology based automated discovery of Workflow pattern.

According to, once we create the well-defined pattern ontology, pattern training task can be facilitated by it and if any change occurs in the pattern, it will help in production of change adaptable code.

3.7 Implementation Ontologies

This subsection illustrates software artifact, object oriented source code, version, and configuration ontologies, which can be used for implementation purpose.

3.7.1 Software Artifact Ontology

**Definition:** Software artifact ontology defines the concepts of various artifacts generated during development of software system. These concepts allow us to classify them according to their type or formats such as text file, binary file, source code, image file, audio/video file, documentation manuals, etc.

This ontology can also define concepts, which contain metadata information such as during which phase a particular artifact is produced, by whom (person) it is created and in which project it is created. **State of the Art:** According to, once we have all the metadata information about the artifacts produced, this ontology along with application domain ontology can be used in automatic generation of documentation (up to some extent).

Metadata standards such as Dublin Core Metadata from Dublin Core Metadata Initiative and IEEE learning Object Metadata can be referenced for concepts containing information about metadata.

3.7.2 Object Oriented Source Code Ontology (OO Source Code Ontology)

**Definition:** OO source code ontology defines all the concepts related to Object Oriented Programming (OOP). OOP major concepts such as Abstraction, Inheritance, Polymorphism, and encapsulation will be defined by this ontology. It will also include concepts such as package, class, subclass, variable or data member, member function, object, etc.

3.7.3 Version Ontology

**Definition:** Version ontology defines the various versions of software, which are generated till date. It establishes the relationship between various generated artifacts to a particular version. It includes the concepts such as file, release, and revision.

3.7.4 System Configuration Ontology

**Definition:** System configuration ontology models the characteristics of various components belonging to different versions. The Ontology prescribes the kind of software/hardware configuration required to install the
software or which version is supported by the available configuration.\footnote{81}

3.8 Documentation Ontologies
This sub-section expounds documentation and document ontologies, which can be used for software's documentation purpose.

3.8.1 Documentation Ontology

Definition: Documentation ontology defines all the documents related to software system. It contains various documents generated during various phases such as formal specification, context diagram, and data flow diagram generated during requirement analysis and specification phase; flow chart and entity relationship diagram generated during design phase; source code listing and cross-reference listing generated during implementation phase; test data and test result generated during testing phase.\footnote{80-83}

State of the Art: In order to provide better documentation, Graaf\footnote{84} and \footnote{85}developed an annotating semantic wiki page with the help of light weight (with limited classes/properties) Software Engineering ontology, which contains Dublin core data properties to allow specification of metadata. They developed the wiki page with the help of ArchiMind and OntoWiki Semantic Wiki in which the annotated text can be searched. Both Archimind and Ontowiki semantic wiki provide advantages in versioning, responsibility, accessibility, and in traceability. However, issues such as synonym and homonym remain along with spelling error, abbreviation, ambiguity, and context dependent interpretation. The only difference between their work is that author \footnote{85}provides the documentation on software requirements and architecture design part only. An exploratory study on ontology engineering approach for software architecture documentation has been conducted by De Graaf et al.\footnote{86,87}.

According to \footnote{80} and \footnote{81}, this documentation ontology can be used to facilitate the task of re-establishing the lost information i.e. it can provide help in traceability. In a survey done by\footnote{88}, it was found that many practitioners don't document very frequently; reasons are “no time”, “no budget”, “no standards” and no suitable tool. As per\footnote{89}, it is also due to the unwelcoming processes that are not fully integrated into the software industry. To solve such problems we can also automate the documentation process, up to some extent, by using software artifact, domain, and software architecture ontologies.

As far as automatic documentation is concerned, \footnote{90}developed a Toeska Rationale Extraction (TREx) tool that is used for software architecture documentation part only and is based on the software architecture and design ontologies. According to them, due to lack of standard and guidelines, practitioners don't know what, why, when and how to document. It is also due to schedule and budget constraint that they don't document. They implemented the TREx on a case study and compared its outcome vs. plain text documents and it was found that humans are poorer than TREx in identifying rationale, but are better at dealing with “generic” components.

3.8.2 Document Ontology

Definition: Document ontology defines the relationship between various documents. It relates an updated document to its source document. It also defines that by what relationship the two documents are linked with each other.\footnote{82}

Document ontology differs from documentation ontology as the document ontology attempts to organize documents, whereas documentation ontology attempts to organize information recorded by documents. They also differ in the way the concepts and restriction are defined in both ontologies.

3.9 Quality Ontologies

This sub-section examines quality, testing, and defect ontologies, which can be used for describing various quality attributes and assessing those quality attributes that in turn enhances the quality of software system being developed. Presently, there are several tools that are widely used for software quality measurement such as Ishikawa's seven basic tools, scatter diagram, pareto diagram, histogram, control chart, run chart, and cause-effect diagram.\footnote{32} Several quality models and standards have been developed such as Capability Maturity Model (CMM), ISO 9001, and ISO/IEC 9126, which are used to improve the quality of software system being developed.\footnote{33}

3.9.1 Quality Ontology

Definition: Quality ontology defines various quality attributes related to software system such as correctness, consistency and precision, robustness, simplicity,
traceability, accuracy, completeness, efficiency, testability, modularity, readability, simplicity, modifiability, expandability, portability, etc.\textsuperscript{91}.

**State of the Art:** According to \textsuperscript{91}, this ontology once created will be able to suggest elicitation questions related to non-functional requirements.

### 3.9.2 Testing Ontology

**Definition:** Testing ontology defines the concepts of testing such as who is the tester, in what kind of environment testing is performed, what are the various testing mechanism available, on what artifacts testing is performed, etc.\textsuperscript{92}.

This ontology can further explain testing mechanism, for example, the various testing techniques that are available such as black box testing (functional testing) and white box testing (structural testing) techniques; levels of testing such as unit testing, integration testing, and system testing; kinds of testing such as validation and verification testing.

**State of the Art:** According to Zhu et al.\textsuperscript{92}, this ontology can provide help in automation of testing process. As far as automation is concerned, Nasser et al.\textsuperscript{93} presented a framework to automatically generate the executable test cases based on the ontologies on dissimilar domains with custom defined coverage rules. The framework uses various ontologies such as behavioural model ontology, domain ontology, and implementation ontology to generate the custom defined coverage criteria in order to generate the executable test suite. They also explained the limitations for practical implementation of this framework such as non-optimal redundancy checking constraint, time efficiency constraint, modeling complexity constraint, and the logic programming limitation constraint.

Paydar et al.\textsuperscript{94} presented a theoretical roadmap for ontology based web application testing. Here they discuss that first they require to capture all the knowledge related to testing process, second they require application domain ontology in order to have all knowledge about application under test, and third and last is to develop procedures, which are required in order to automate the testing process, with the help of previously gained knowledge. While creating the roadmap, they discuss the Software Engineering Body of Knowledge (SWEBOK) that keeps knowledge such as relationship, concepts, facts, and principles about various Software Engineering topics (known as knowledge area).

Li et al.\textsuperscript{95} used the testing ontology to automatically generate the test cases. They describe an ontology based approach for Graphical User Interface (GUI) testing in which ontology is built for GUI system. Using reverse engineering techniques based on the design of GUI ontology, all elements from source code are extracted and stored in GUI ontology. Then using the knowledge stored in GUI ontology, test case generation rules are extracted by analysing tester’s experience that is based on the GUI components sequences. At last, based on test case generated rules and GUI ontology, test cases are generated. Then Li et al.\textsuperscript{96} described the ontology based approach for reusing the already built test cases. It is found that the test case design effort is reduced and efficiency is increased as a result of their work. Dalal et al.\textsuperscript{97} also described an ontology based approach for test case reuse, which enhances flexibility and reusability during test case generation and handles users’ queries better.

To the best of our knowledge no work has been done on ontology based automatic testing, which gives researchers ample opportunity in this direction.

### 3.9.3 Defect Ontology

**Definition:** Defect ontology defines the concepts that describe the defects detected during the testing phase. It includes the concepts such as Issue, which defines the defect; Person, who detects and rectifies it; Comment to provide extra information; Activity, which is required to rectify it; Computer System, Product, Component to which a defect belongs\textsuperscript{99}.

This ontology can also be used for analysing detected defects. We can establish suitable metrics for the analysis purpose. With suitable metrics, severity level of the defect can be defined. Based on the above information, prioritization can be done. Higher priority means higher degree of risk to the software system, which means such a defect needs to be resolved first before mitigating others. Once the defect is detected, we can have information such as to which phase it belongs and by what method it is detected. Based on analysed information, we can also estimate the cost required to repair the detected defects.

**State of the Art:** As far as defect severity is concerned, Iliev et al.\textsuperscript{98} used ontology based approach for automatic prediction (classification) of defects severity that depends on the requirements, design, coding, and configuration. They focused on the classification of defects from User Point of View rather than developer’s and tester’s. To do this, they picked a case study from Industry Logic a (Organization Name) and collected defects
and classification as provided by them. Then using ontology based approach, they automatically classified the defects into three levels namely Major, Medium and Minor. Further they compared their work with original classification (as provided by the organization) and found that 58% defects matched with the same severity level.

According to Kiefer et al.\textsuperscript{79} also, this ontology can be used to detect the potential areas from the source code where defects can be found. The task of measuring defect density can also facilitated by this ontology in order to measure the quality of the software system being developed.

### 3.10 Maintenance Ontology

This sub-section discusses maintenance ontology, which describes various maintenance activities.

#### 3.10.1 Software Maintenance Ontology

**Definition:** Software Maintenance Ontology defines the maintenance related concepts and their relationship\textsuperscript{99}. It can define the concepts describing maintenance type such as perfective maintenance, adaptive maintenance, corrective maintenance, and any other type of maintenance; what procedures can be used to maintain the software i.e. maintenance activities; who does the maintenance. It includes the concepts like activity such as management activity and modification activity; person such as maintenance engineer and maintenance manager; procedure such as method, technique, and paradigm; resource such as maintenance human resource, client human resource, etc. It includes the relationship such as approves, performs, trains, uses, etc. Maintenance ontology can also be used to model various maintenance models such as Iterative enhancement Model, Reuse Oriented Model, Boehm's Model, Taute Maintenance Model, etc.\textsuperscript{100}

**State of the Art:** Reverse Engineering is also done for maintenance purpose that includes activities such as discovering unknown and hidden information of a software system. Pattern ontology along with software architecture and object oriented design ontologies can be used to facilitate the task of discovering such unknown and hidden information as patterns tends to reveal the style that was used to create the software system\textsuperscript{59}.

Semantic web approaches can also be facilitated for estimation of maintenance cost. We can use models such as ‘Belady and Lehman Model’, Boehm Model, etc. to estimate the maintenance cost.

April et al.\textsuperscript{99} discussed that a Semantic Web based software maintenance tool can be built. This tool will use software maintenance ontology. Hyland-Wood et al.\textsuperscript{50} also proposed an approach for maintaining the software system using Semantic Web techniques. They built an ontology that captures metadata information of software components such as requirements documentation (both functional and non-functional), metrics, test suite and the means by which various components interact with each other. If any change is done to any software component, it will trigger all the metadata information related to that changed software component in order to understand and maintain that component. For this purpose, they encoded the metadata information in RDF graph and SPARQL queries were used over them in order to facilitate the understanding and maintenance of the software component. SPARQL Queries were built to check whether re-validation of changed software component is required in order to know that which test case has failed and to which requirement that failed test case belongs\textsuperscript{101}.

### 3.11 Technology Ontology

This sub-section illustrates technology ontology, which describes the various existing tool and techniques that can be used for software development.

#### 3.11.1 Technology Ontology

**Definition:** Technology ontology is repository of various software development tools that can be used to develop a particular software\textsuperscript{43,58}.

**State of the Art:** Dinger et al.\textsuperscript{58} illustrates that using Semantic Web environment, a tool can be generated to select a particular technology for developing the software system. Metadata information about the tool can be provided using Semantic Web techniques such as information about its application scope, information about its required hardware-software combinations, and information about properties of particular tool. Metadata information can also include that what kind of information can be generated by a particular tool\textsuperscript{102}.

### 4. Categorization based on Application Scope of Ontologies

The aim of this section is to classify various ontologies...
based on software life cycle phases (as defined earlier) and their application scope. The categorization of ontologies based on software life cycle phases is illustrated in Figure 1.

![Figure 1. Categorization of ontologies based on software life cycle phases.](image)

In accordance to the categorization based on software life cycle phases, we identified that upper, process, quality, and technology ontologies can be exploited in all the life cycle phase; application domain, application domain feature model, and system behaviour ontologies have been utilized in requirement analysis and specification phase; software architecture, application logic, object oriented design, and pattern ontologies have been utilized in the design phase; software artifact, object oriented source code, version, system configuration, documentation, and document ontologies have been used in implementation phase; testing and defect ontologies have been utilized in testing phase and software maintenance ontology exploited in operation and maintenance phase.

The categorization of ontologies based on their application scope such as generic, domain specific, object-oriented system specific, and project specific is illustrated here. Generic scope means the concerned ontologies can be used in any Software Engineering process. Domain specific scope means the concerned ontologies can only be used in the domain being modelled. Object-oriented system specific scope means the concerned ontologies can only be used in the applications where OO programming is used. Project specific scope means the concerned ontologies can only be used in their concerned projects.

Table 1 shows ontologies and their respective

<table>
<thead>
<tr>
<th>Ontologies</th>
<th>Application Scope</th>
<th>Generic</th>
<th>Domain Specific</th>
<th>Object-Oriented System Specific</th>
<th>Project Specific</th>
<th>Related Work (References)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>(24-26)</td>
</tr>
<tr>
<td>Software Process</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(27-29)</td>
</tr>
<tr>
<td>Application Domain</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>(29, 34-45)</td>
</tr>
<tr>
<td>Feature Model</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>(46-49)</td>
</tr>
<tr>
<td>System Behaviour</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>(36)</td>
</tr>
<tr>
<td>Software Architecture</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>(50)</td>
</tr>
<tr>
<td>Application Logic</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>(51)</td>
</tr>
<tr>
<td>OO Design</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td>(52,53)</td>
</tr>
<tr>
<td>Pattern</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(56-60,66,68)</td>
</tr>
<tr>
<td>Software Artifact</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>(34,35,37)</td>
</tr>
<tr>
<td>OO Source Code</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td>(71-73)</td>
</tr>
<tr>
<td>Version</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>(71)</td>
</tr>
<tr>
<td>System Configuration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>(73)</td>
</tr>
<tr>
<td>Documentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>(72-78,81)</td>
</tr>
<tr>
<td>Document</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>(74)</td>
</tr>
<tr>
<td>Quality</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>(82)</td>
</tr>
<tr>
<td>Testing</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>(83-88)</td>
</tr>
<tr>
<td>Defect</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>(71,89)</td>
</tr>
<tr>
<td>Maintenance</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>(53,90,91)</td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>(37,51)</td>
</tr>
</tbody>
</table>
application scope to which they belong. A ✓ (Tick) mark against the ontology shows its corresponding application scope.

5. Application of Ontologies in Various Software Engineering Areas

This section illustrates the applications of ontologies in various Software Engineering areas. Through this literature survey, we found applications of ontologies in various Software Engineering areas such as requirement engineering, software design, implementation, testing, maintenance, documentation, traceability, change control, quality, reuse, and technology selection and process support along with 'Reverse Engineering' where very little work has been done limited only to web applications.

Table 2 shows ontologies and Software Engineering areas as which they address. A ✓ (Tick) mark against the ontology shows that its application is found in corresponding Software Engineering area. However, a ? (Question) mark against the ontology shows that its application is not found in existing literature, but can be used in the future.

At last, it can be summarised that if we traverse horizontally in the table, we will find that which ontology has addressed and can also address what Software Engineering areas. Further if we traverse vertically in the table, we will find that which Software Engineering area is addressed and can also be addressed by what combinations of ontologies.

6. Future Work (Open Issues)

The following open issues have been identified across
literature and clearly emerged as potential opportunities for future work.

1. Potential opportunities in Requirement Engineering: Using domain and quality ontologies to improve overall requirement engineering process (e.g. removing ambiguities) and assessing it with well-defined metrics can be a future work.

2. Potential opportunities in Software Design: Using software architecture, application logic, object oriented design and pattern ontologies to publicize architectural styles and information related to existing modeling languages. This may improve spreading architecture and design knowledge.

3. Potential opportunities in Software Testing: Work exhibiting semi or fully automatic testing of software using ontologies has though been presented, but much of the work is just theoretical or framework based, where real implementation is yet to be done. Also, another possible work in software testing is the use of domain, system behaviour and implementation ontologies can be explored to generate executable test cases with well-defined domain/system and custom specific coverage criteria rules.

4. Potential opportunities in Software Maintenance: The potential use of ontologies can be explored diversely to enrich the maintenance process as follows:
   • Using application domain, software artifact, version, quality, testing and maintenance ontologies can be used to represent software system components and information about them (information such as functional and non-functional requirements, metrics, test cases and how various components interact with each other) to facilitate understanding and maintenance of software system.
   • Improving semantic annotation process of software models and source code, integration of ontologies and meta-modeling architectures, and all-inclusive traceability model of software artifacts can be used to better understand software knowledge artifacts.
   • Using application domain, object oriented design and object oriented source code ontologies to automatically detect and update out-dated requirements during maintenance phase can be a potential future work. A framework has been proposed for the above issue by Bhatia et al., but real implementation is still to be done.

5. Potential opportunities in Software Documentation: Automation of documentation process using Application Domain (AD), Software Architecture (SArch), Software Artifact (SArt), and Documentation ontologies can be a possible future work. AD, SArch, and SArt ontologies can be merged together to develop documentation ontology. Any missing information in documentation ontology can be extracted from data corpus, if available, so that it has all necessary details. Then a transformation approach can be established to convert the documentation ontology into text/html document. Though a framework for this has been proposed by Bhatia et al., but implementation is yet to be done.

6. Potential opportunities to improve Software Quality: Using feature model and quality ontologies to detect inconsistencies in feature models along with reasons for inconsistencies from usability point can be explored. Also, detecting conflicts in feature models for accurate decision and error free configuration, creating repository of feature model can be possibly investigated.

7. Potential opportunities to Reuse software related knowledge: Reuse of all ontology design patterns to show their usefulness in terms of quality and maintainability issues can be a possible direction of future work. Assessment of design patterns to improve their quality using pattern and quality ontologies can also be examined. Further, Reuse of documentation ontology to create documents for software that reuses source code can also be explored.

8. Potential opportunities in Reverse Engineering: Using software architecture, object oriented design and pattern ontologies to facilitate the discovery of unknown and hidden information for the purpose of reverse engineering can be a prospective direction of work. Most of the work in this area is limited to ontology based web application reverse engineering, whereas the scope of using Semantic Web in conventional software reverse engineering is not yet examined.

7. Conclusion

This research paper has discussed about the past, present and future of ontologies for Software Engineering. The existing ontologies have been defined, their state-of-art reviewed to an extent where the past and present applications have been investigated. Also, a detailed description on open challenges that keep this area dynamic and alive have been analysed. We have attempted to provide a categorization of ontologies based on software life cycle phases and their application scope.
### 8. References

29. Hawker JS, Ma H, Smith RK. A Web-based process and process models to find and deliver information to improve the quality of flight software. The 22nd In Digital Avionics Systems Conference DASC’03. IEEE; 2003 Oct 12; 1:3-B.
Ontologies for Software Engineering: Past, Present and Future

84. de Graaf KA. Annotating software documentation in semantic wikis. InProceedings of the fourth workshop on Exploiting semantic annotations in information retrieval 2011 Oct 28 (pp. 5-6). ACM.
95. Li H, Chen F, Yang H, Guo H, Chu WC, Yang Y. An ontol-

96. Li X, Zhang W. Ontology-based testing platform for reusing. In Internet Computing for Science and Engineering (ICICSE), 2012 Sixth International Conference on 2012 Apr 21 (pp. 86-89). IEEE.


