Computer Standards & Interfaces 46 (2016) 66–78

Contents lists available at ScienceDirect

Computer Standards & Interfaces

journal homepage: www.elsevier.com/locate/csi

A validated ontology for global software development

Aurora Vizcaíno [a,*], Felix García [a], Mario Piattini [a], Sarah Beecham [b]

[a] Alarcos Research Group, Institute of Information Technologies and Systems, Escuela Superior de Informática, University of Castilla-La Mancha, Spain
[b] Lero, The Irish Software Research Centre, Department of Computer Science and Information Systems, University of Limerick, Limerick, Ireland

ARTICLE INFO

Article history:
Received 31 March 2015
Received in revised form 11 December 2015
Accepted 2 February 2016
Available online 10 February 2016

Keywords:
Global software development (GSD)
Global software engineering (GSE)
Distributed software development (DSD)
Ontology
Meta-model

ABSTRACT

The global software development (GSD) paradigm has, over the last 15 fifteen years, shifted from being novel and ground breaking to being widely adopted and mainstream. This wide adoption is partly owing to the many benefits provided by GSD, such as reduced labour costs, proximity to new markets and access to a diverse and experienced skills pool. Yet taking advantage of these benefits is far from straightforward, and research literature now includes a proliferation of guidelines, reviews and models to support the GSD industry. Although this active area of study is firmly established as a research area in its own right, the boundaries between general software engineering and GSD are somewhat confused and poorly defined. In an effort to consolidate our understanding of GSD, we have developed an ontology in order to capture the most relevant terms, concepts and relationships related to the goals, barriers and features of GSD projects. The study we present here builds on research conducted in a collaboration project between industry and academia, in which we developed an ontology in order to provide practitioners with a "common language and conceptualisation". Its successful outcome encouraged us to create a broader ontology that captures the current trends in GSD literature. The key ontology, along with its three subontologies, are the result of a review of the relevant literature, together with several expert evaluations. This ontology can serve as a useful introduction to GSD for researchers who are new to the paradigm. Moreover, practitioners can take advantage of it in order to contextualise their projects and predict and detect possible barriers. What is more, using a common language will help both researchers and practitioners to avoid ambiguities and misunderstanding.

1. Introduction

There is currently a strong and growing trend that involves IT organisations employing multisite teams that span geographic boundaries to build their software. Organisations are adopting a Global Software Development (GSD) paradigm, either out of necessity or owing to the promise of competitive benefits. GSD can take various forms, and the proliferation of literature reviews on the subject reflects this active area of research and practise. Indeed, there are now even two tertiary research studies [1,2] that capture the groups of secondary literature reviews. Since GSD has now become a dominant paradigm, and thanks to the proliferation of research that the community can draw on, the research community now needs a common understanding of the terms and concepts related to GSD projects.

One way in which to achieve this goal is to create a GSD ontology. An ontology is defined by Gruber [3] as the formal specification of a shared conceptualisation. Ontologies provide a visual means to share a common understanding of the structure of information among people or software agents. Moreover, ontologies enable the reuse of domain knowledge, thus making domain assumptions explicit and helping us to clarify any ambiguities [3,4]. What is more, and as is illustrated by the successful applications of ontologies in the Software Engineering field, an ontology can provide numerous benefits, from serving as a basis for discussion to achieving the necessary agreement and consensus and contributing to the harmonisation of standards in the field, as a vehicle for achieving the interoperability required between the ever-increasing number of groups and organisations working on languages and tools, or to support model driven engineering initiatives (see [5,6,7,8]).

Bearing in mind the aforementioned issues, our main research objective has therefore been to build an ontology with which to determine the advantages, challenges and concepts related to GSD projects. This ontology is based on the relevant related GSD ontologies that exist in the field and has been agreed on by experts in the field.

The remainder of this paper is organised as follows. The background to this research is presented in the following section, while the research method followed to build the GSD ontology is described in Section 3. In Section 4, we focus on describing the systematic mapping carried out to discover the ontologies and taxonomies exist in the literature related to GSD, and Section 5 provides a description of the process carried out to discover new terms and relations and the results obtained. An evaluation of the first version of the ontology by experts is presented in Section 6 and the refined result obtained (GSD ontology v2) is then described in

* Corresponding author at: Escuela Superior de Informática, Paseo de la Universidad 4, 13071 Ciudad Real, Spain.

http://dx.doi.org/10.1016/j.csi.2016.02.004
0920-5489/

Section 7. In Section 8, we present the validation of GSD ontology v2. Finally, the conclusions, limitations and guidelines for using the ontology, along with an outline of future work, are presented in Section 9.

2. Background

There is currently an important interest in developing ontologies and using them in the software engineering field. A proof of this fact is the increasing number of workshops focused on this, such as the workshop on ontology, conceptualisation and epistemology for software and systems engineering (ONTOSE), the workshop on ontologies and metamodeling software and date engineering (WOMSDE) or the workshop on semantic web-enabled software engineering (SWESE). These and other efforts contribute towards ontologies attaining more maturity in the field as an engineering discipline.

Various works of interest exist in the field of Software Engineering. As a representative example of existing works, Wongthongtham et al. [9] described a software engineering ontology model based on the software engineering textbook [10] and SWEBOK with the aim of enabling communication between software engineers according to common software engineering knowledge.

With regard to GSD-related conceptualizations, Wongthongtham et al. [11] defined five ontologies to be used in a multisite software engineering environment: a business domain ontology to characterise the fundamental knowledge about a particular domain, since all software is designed to solve a business need such as accounting or a customer service etc.; a software engineering ontology in which software engineering principles and aspects are described; a project management ontology, created to enable all the stakeholders to have consistent knowledge when discussing project-related matters; an issues ontology, divided into ontological, technical and managerial issues; and the solution ontology, which is related to the knowledge of issues and solutions that drive the success of the project and product. These ontologies are developed to serve intelligent software agents, which support multisite software development. Wongthongtham et al. [12] proposed a software engineering sub-ontology to enable remote team members to browse, search for and share data in a distributed software engineering project environment. One of their main goals is therefore to enable communication between computer systems. The authors thus provide two subontologies, a generic ontology and an application-specific ontology. The generic ontology is a set of software engineering terms, including the vocabulary, the semantic interconnections and some simple rules of inference and logic for software development. It provides the vocabulary for the terms in software engineering, along with an application-specific ontology, which is an explicit specification of object-oriented development in software engineering. The approaches of both ontologies help to transform explicit knowledge into conceptual knowledge representation with the aim of using software agents to access data from this project ontology repository. In our previous research [13], we developed the O-GSD ontology with the objective of promoting a common understanding of GSD vocabulary based on a particular project called ORIGIN. This ontology achieved its purpose by facilitating communication between members of the ORIGIN project team and by helping members of the team to avoid misunderstandings when working in groups and to share a similar view of the project.

The analysis of relevant literature led to the discovery of certain limitations, particularly the lack of relevant terms that are specific to GSD projects or the need to reinforce the external validity of ontologies developed in the context of a single GSD project, such an in our previous research [13]. In addition, the benefits that the use of the O-GSD ontology [13] allowed us to obtain in this project encouraged us to consider the creation of a more general ontology for GSD projects during which the experts took will take part in the process of creating and evaluating it, since one important feature of an ontology is that of representing the consensus of the experts in a particular domain.

3. Research method for the creation of the GSD ontology

In this section, we present the research method followed to develop our validated GSD ontology. The general question used for this research was Is it possible to create an ontology with which to determine the advantages, challenges and concepts related to GSD projects? This research question was addressed by the following method shown below (Fig. 1).

As shown in Fig. 1, the GSD ontology was built by following three fundamental steps:

1) We augmented the initial version of the ontology (O-GSD) by means of the following:
   - o A rigorous analysis of GSD concepts and relationships found in the relevant literature. These concepts were found through the use of a systematic mapping of ontologies and/or taxonomies related to GSD and by searching for additional concepts in proceedings of the International Conference on Global Software Engineering (ICGSE) 2011, 2012 2013, 2014, as this is the most important international conference in this domain.
   - o Building on our previous O-GSD ontology. The concepts described in the O-GSD were also considered as a starting point for the GSD ontology, given that the building of this ontology was also based on the analysis of previous relevant literature and on a consensus between researchers and practitioners involved in the ORIGIN project. The aim was to discover the terms and relationships which were not specific to the project and that could be generalised to any GSD context, and focused mainly on the goals, barriers and features of a GSD project.
2) A survey was developed to serve as both a validation instrument and a means to review the ontology draught produced. As a result of this step, we gained very useful feedback which helped us to create the GSD ontology presented herein.
3) The new ontology was validated by experts by means of a survey in which Gruber's criteria for evaluating ontologies [3] were followed.

The first step is described in more detail in Sections 4 and 5. The evaluation of this ontology (draught) and the result GSD ontology are presented in Sections 6 and 7, respectively. Finally, the third step is described in Section 8.

4. GSD ontologies and taxonomies: a systematic mapping

In this section, we describe the protocol used to carry out the systematic mapping review. We first describe the design and then go on to present the results obtained.

4.1.1. Planning and design

One of the principal goals of ontologies is to formalise a domain in a way that reflects the consensus of the majority of the members of the given community. As this formalisation should promote a shared and common understanding of the domain, the first step was to explore what the other ontologies or taxonomies in GSD are proposing. This would then allow us to detect what topics are considered relevant by other authors and to determine whether these terms should be added to the ontology that we wished to extend. Furthermore, other researchers' previous work would help us discover the different views or focuses of GSD that we had not taken into account in our initial ontology (O-GSD). In summary, we analysed previous ontologies and taxonomies concerning GSD in order to discover new approaches or views.

- Detect new terms.
- Take advantage of other research efforts in which themes have been