OntoExper-SPL: An Ontology for Software Product Line Experiments

Henrique Vignando[a], Viviane R. Furtado[b], Lucas O. Teixeira[c] and Edson OliveiraJr[d]
Informatics Department, State University of Maringá, Maringá - PR, Brazil

Keywords: Experiment, Ontology, Software Product Line.

Abstract: Given the overall popularity of experimentation in Software Engineering (SE) in the last decades, we observe an increasing research on guidelines and data standards for SE. Practically, experimentation in SE became compulsory for sharing evidence on theories or technologies and provide reliable, reproducible and auditable body of knowledge. Although existing literature is discussing SE experiments documentation and quality, we understand there is a lack of formalization on experimentation concepts, especially for emerging research topics as Software Product Lines (SPL), in which specific experimental elements are essential for planning, conducting and disseminating results. Therefore, we propose an ontology for SPL experiments, named OntoExper-SPL. We designed such ontology based on guidelines found in the literature and an extensive systematic mapping study previously performed by our research group. We believe this ontology might contribute to better document essential elements of an SPL experiment, thus promoting experiments repetition, replication, and reproducibility. We evaluated OntoExper-SPL using an ontology supporting tool and performing an empirical study. Results shown OntoExper-SPL is feasible for formalizing SPL experimental concepts.

1 INTRODUCTION

Experimentation is one of the most relevant scientific methods to provide evidence of a theory in a real-world scenario. There is a growing consensus that Software Engineering (SE) experimentation is fundamental to developing, improving and maintaining software, methods and tools. This allows knowledge to be generated in a systematic, disciplined, quantifiable and controlled way (Wohlin et al., 2012).

Experimentation is not a simple task, it requires careful planning and constant supervision to avoid any bias towards internal and external reliability. To reduce the planning burden and to mitigate some reliability issues SE researchers proposed protocols, guidelines, tools and others to aid the experimentation process. One of the approaches proposed is the formalization of the entire experiment (planning, execution, analysis and results) in order to facilitate the validation, accessibility and comprehension. This increases the chances of a successful replication and increases reliability in the results as well as the overall quality of the study (Wohlin et al., 2012). Nowadays, experiment replications in SE are almost non-existent which prevents any chance of a meta-analysis (Gómez et al., 2014).

Ontologies are among the most used methods to formalize information. Ontologies are formal representations of an abstraction containing formal definitions of nomenclature, concepts, properties and relationships between concepts. An ontology defines a controlled vocabulary of terms and relationships of concepts in a domain (Noy et al., 2001). Thus, the use of ontologies to represent information about experiments standardizes the data facilitating the interoperability, the exchange of information and the replication of experiments.

Given the particularities of each area in SE, one solution fits all might not work. The design of an ontology capable of representing all the particular characteristics of all SE areas is too complex. Because of that, we concentrated our efforts in the field of the Software Product Line (SPL) due to our group experience (Research Group on Systematic Software Reuse and Continuous Experimentation - GReater).

A SPL is determined by a set and products of a particular market segment (Pohl et al., 2005), as a system for cellular devices, where there are core assets

[a] https://orcid.org/0000-0003-3756-9711
[b] https://orcid.org/0000-0002-5650-4932
[c] https://orcid.org/0000-0003-3615-1567
[d] https://orcid.org/0000-0002-4760-1626

Vignando, H., Furtado, V., Teixeira, L. and OliveiraJr, E.
OntoExper-SPL: An Ontology for Software Product Line Experiments.
DOI: 10.5220/0009575404010408
In *Proceedings of the 22nd International Conference on Enterprise Information Systems (ICEIS 2020) - Volume 2*, pages 401-408
ISBN: 978-989-758-423-7

with the main functionalities that the software must implement, called similarities, and can have a number of functionalities specific to certain devices, called variabilities. Given the challenges due to the more flexible project, experiments in SPL are even harder to conduct. Therefore, the definition of an ontology specific for experiments in SPL might aid other researchers in the development and replication of new experiments, also assisting in the auditing and validation of experiments.

Furthermore, data formally represented allow the development of specialized systems, such as recommendation systems. A recommendation system in SE experimentation could be useful in two ways: (i) didactically, to aid students comprehend what is a high-quality experiment; and, (ii) practically, to aid ESE (Experiment in Software Engineering) practitioners plan and execute a high-quality experiment by following others experience.

Therefore, the goal of this paper is to propose an ontology, named OntoExper-SPL, to formally represent experiments in SPL. We expect to facilitate the experiment data representation in order to increase the overall experiment quality, raise the number of replications and data sharing. In order to design such ontology we considered guidelines found in the literature and an extensive systematic mapping study previously performed by our research group (Furtado, 2018).

Our initial feasibility evaluation indicates the possibility of inferences from the data which allows the creation of models for a recommendation system for example.

2 BACKGROUND AND RELATED WORK

2.1 Software Product Lines

A Software Product Line (SPL) is a set of products that address a particular market segment (Pohl et al., 2005). Such a set of products is called a product family, in which the members of this family are specific products generated from the reuse of a common infrastructure, called the core assets. This is formed by a set of common features, called similarities, and a set of variable characteristics, called variabilities (Linden et al., 2007).

Pohl et al. (2005) developed the framework for SPL engineering. The purpose of this framework is to incorporate the core concepts of SPL engineering, providing reuse of artifact and mass customization through variability.

The framework is divided into two processes: Domain Engineering and Application Engineering. Domain Engineering represents the process in which the similarities and variabilities of SPL are identified and realized. Application Engineering represents the process in which the applications of an SPL are built through the reuse of domain artifacts, exploiting the variabilities of a product line.

2.2 Experimentation in Software Product Lines

Furtado (2018) conducted a Systematic Mapping Study (SMS)[1] that extracted data of SPL-driven experiments from reliable sources, such as ACM, IEEE, Scopus and Springer, as well as prestigious journals and conferences in the area. She analyzed various important characteristics such as: the report template used, the experimental elements reported, the experimental design including the availability of its package to allow replication, among others.

Based on the SMS data, Furtado (2018) developed a conceptual model with a set of guidelines for the evaluation of SPL experiment quality. In order to create a representative formalization of the SPL domain, we clustered the items on the conceptual model based on Wohlin template. So our proposed ontology follows both the theoretical and practical sides of SPL.

2.3 Software Engineering Ontologies

An ontology is a set of entities that can have relationships with each other. Entities may have properties and constraints to represent their characteristics and attributes. Each entity has a population of individuals.

In the current scenario, the ontology in information science is used as a form of representation of logical knowledge, allowing the inference of new facts based on the individuals stored in the ontology (Gruber, 1993). These definitions follow the representation pattern known as descriptive logic. The major reasons for building an ontology are: i) the definition of a common domain vocabulary; ii) domain knowledge reuse; and, iii) information share.

The knowledge base has two main components: the concepts of a specific domain, called TBox (Terminological Box), and the individuals in that domain, called ABox (Assertion Box) (Calvanese et al., 2005). The individuals in the ABox must comply with all the

[1]Currently under review in a journal.

properties and restrictions defined in the TBox. Moreover, it is also possible to use an inference mechanism to query and extract new information from the knowledge base.

2.4 Related Work

We have found some studies that proposed approaches to formally represent data about SE experiments. However, none of the studies considered specifically SPL experiments.

The work of Garcia et al. (2008) proposes, through UML class diagrams, an ontology for controlled experiments in software engineering, called EXPEROntology. The work of Scatalon et al. (2011) is an evolution of the proposed ontology of Garcia et al. (2008). The work of Cruz et al. (2012) presents an ontology called OVO (Open proVence Ontology) developed inspired by three theories: (i) The life cycle of scientific experiments, (ii) Open Provent (OPM) and (iii) Unified Foundational Ontology (UFO). The OVO model is intended to be a reference for conceptual models that can be used by researchers to explore metadata semantics.

In a more broad topic, the work of Blondet et al. (2016) proposes an ontology proposal to numerical DoE (Design of Experiments) to support the process decisions about DoE. The work of Soldatova and King (2006) proposes a ontology for general experiments, called EXPO, it specifies the concepts of design, methodologies and representation of results. This work is the only one that uses the OWL-DL model to represent the ontology.

The work of Gelernter and Jha (2016) gives an overview on the challenges of evaluating an ontology, but does not mention ontology for experiments.

Finally, the work of Cruzes et al. (2007) deals with a technique to extract meta information from experiments in software engineering. This is especially important in our study because our proposed ontology must be able to represent that metadata.

3 AN ONTOLOGY FOR SPL EXPERIMENTS

3.1 Ontology Conception

Our proposed ontology[2] is a domain ontology following a semi-formal approach used for modelling application and domain knowledge (Gómez-Pérez, 2004).

[2]Complete diagrams of OntoExper-SPL at https://doi.org/10.5281/zenodo.3707797

The following ontology elaboration process was applied (Gómez-Pérez, 2004): (i) definition and structuring of terms in classes; (ii) establishment of properties (attributes) inherent to the concept represented by a term; (iii) population of the structure that satisfies a concept and its properties; (iv) establishment of relations between concepts; and (v) elaboration of sentences to restrict inferences of knowledge based on structure.

Initially, a graph was developed for the ontology model mainly to validate the concepts through classes, sub-classes and the relationships between them (Vignando et al., 2020). The creation of this initial ontology model was an exploratory step using the data collected on a systematic mapping of experiments in SPL (Furtado, 2018). The systematic mapping was based on the Wohlin experimental model, which describes five pillars for experiments: Definition, Planning, Operation, Analysis and Interpretation (Wohlin et al., 2012).

We then clustered that initial conceptual model based on the Wohlin pillars. This clustering was necessary for a more concise and abstract understanding of the relationships between domain terms raised in the original conceptual model. In this way it was possible to validate the structure of the initially proposed graph.

Next, we created a class diagram for a more formal representation of the initial modeling. In this representation, the relationship between the terms (classes) and their properties (attributes) was clearer. This form of representation highlighted the main relationship when we defined the composition of the *Experiment* and *ExperimentSPL* class in almost all other sub-classes. Figure 1 presents the clustered conceptual model.

3.2 Ontology Design

The OWL standard was used in OntoExper-SPL to define all elements, classes and sub-classes.

We used Protégé for the final design of the ontology. It can be used by both system developers and domain experts to create knowledge bases, allowing the representation of knowledge an area. We defined our entities based on the class diagram built in the concept phase, with the following order: (i) class definition (ii) definition of object properties, (iii) definition of data properties. Table 1 presents all elements defined in Protégé.

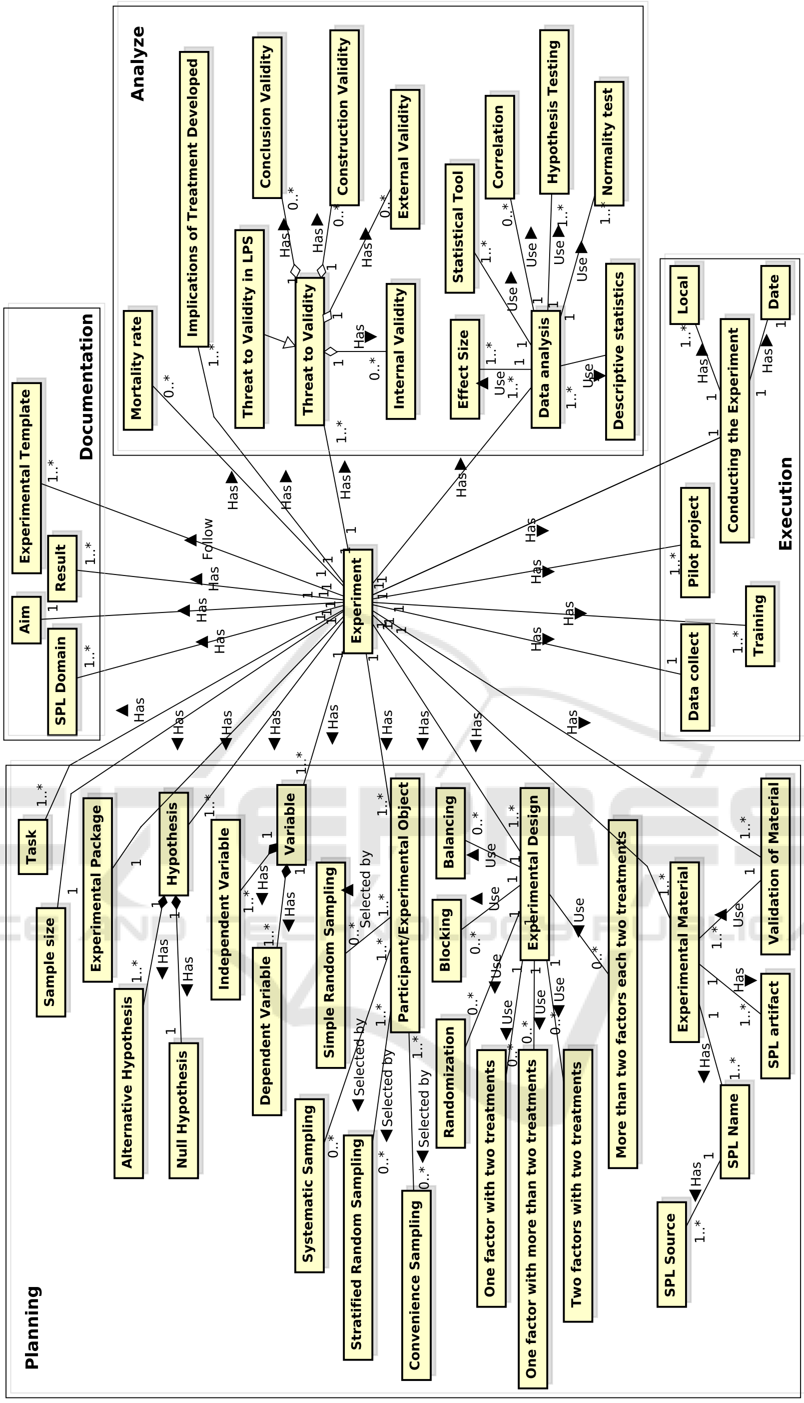

Figure 1: Conceptual Model Clustering of OntoExper-SPL.

3.3 Ontology Population

The next step, after the modeling of the ontology, was to insert the metadata from the SM into the ontology. Although Protégé is able to perform this operation, we chose to use a script to perform such a task, since in Protégé the process of inserting individuals into the ontology is done manually.

Thus, we opted for the use of a script in order to facilitate and automate the insertion of the individuals and later the insertion of new individuals. The script process resembles an Extract Transform Load (ETL)

Table 1: Ontology Design - Classes and Properties modeling.

Element	Definition
Classes	Abstract, Acknowledgments, Analysis, Appendices, ConclusionsFutureWork, Discussion, DiscussionSPL, Documentation, Evaluation, ExecutionSection, Experiment, ExperimentSPL, ExperimentPlanning, ExperimentPlanningSPL, Introduction, Package, References, RelatatedWork, TypeContextExperiment, TypeContextSelection, TypeDesignExperiment, TypeEsperiment, TypeEsperimentSPL, TypeSelectioParticipantObjects
Object Properties	documentation, experiment, typeContextxperiment, typeContextSelection, typeDesignExperiment, typeExperiment, typeExperimentSPL, typeSelectionOfParticipants
Data Properties	idExperiment, title, authorship, publicationYear, publicationType, publicationVenue, pagesNumber, idExperimentSPL, nameSPLUsed, wasTheSPLSourceUsedInformed, idDocumentation, useTemplate, template, observationsAboutTemplateUsed, idAbstract, objective, abstractBackground, methods, results, limitations, conclusions, keywords, idIntroduction, problemStatement, researchObjective, context, idRelatedWork, technologyUnderInvestigation, alternativeTechnologies, relatedStudies, relevancePractice, idConclusionsFutureWork, summary, impact, futureWork, idExperimentPlanning, goals, experimentalUnits, experimentalMaterial, tasks, hypotheses, parameters, variables, experimentDesign, procedureProcedure, explicitQuesiExperimentInStudy, isAQuasiExperiment, idExperimentPlanningSPL, artifactSPLused, idExecutionSection, preparation, deviations, pilotProjectCarriedOut, howManyPilotProjectCarriedOut, idAnalysis, descriptiveStatistics, datasetPreparation, hyp othesisTesting, whatQualitativeAnalysisPerformed, howDatahasBeenAnalyzed, experimentAnalysisBasedPValue, hasQualitativeAnalysisOfExperiment, studyHasPerformMetaAnalysis, idDiscussion, evaluationOfResultsAndImplications, inferences, lessonsLearned, threatsValidity, isFollowThreatsByWohlin, idDiscussionSPL, threatsValiditySPL, idAcknowledgements, acknowledgments, idReferences, references, idAppendices, appendicies, idEvaluation, theAuthorsConcernedEvaluatingTheQuality, idPackage, isExperimentalPackageInformed, url, isLinkAvailable

process.

3.4 Use Case Scenario

In order to illustrate a potential application for the proposed ontology, we present a simple use case scenario that extracts the most used experiment report template.

The query SPARQL (Prud'hommeaux and Seaborne, 2008) on Listing 1 returns all experiments template and how many times each have been used. From that, we can extract the most used template.

Listing 1: Example of an SPARQL query.

```
SELECT
   ?template
   (count(?template) as ?count)
WHERE {
   ?doc rdf:type :Documentation .
   ?doc :template ?template .
}
GROUP BY ?template
```

This example runs through one class (*Documentation*) of the 24 classes in the ontology, one data property (*template*) of 87 data properties. Based on that, in the example used 0.0004% of the response capacity that the model allows. This calculation checks the possibilities of paths between classes and ontology properties.

This initial query example shows how inference mechanisms can be created in the ontology model proposed in this work. Thus, it is possible to extract information about SPL experiments using OntoExper-SPL.

3.5 Preliminary Evaluation

The OOPS! tool was used to generate the assessment of the proposed ontology model. The tool helps to detect some of the most common pitfalls that appear when developing ontologies (Poveda-Villalón et al., 2014). The OOPS! tool has 41 evaluation points of which 34 points are semi-automatically run, as the others depend on a specific ontology domain and they encourage users to improve the tool. The result given by the tool suggests how the elements of the ontology could be modified to improve it. However, not all identified pitfalls should be interpreted as failures, but as suggestions that should be reviewed manually in some cases.

The tool lists the results of each trap as: critical, important and minor.

We summarized the results when running our ontology model proposal in the OOPS! tool in Table 2:

Table 2: OOPS! Traps.

Trap ID	Description	Level
P08	Missing annotations in 119 cases	Minor
P10	Missing disjointedness	Important
P13	Inverse relationships not explicitly declared in 8 cases	Minor
P19	Defining multiple domains or ranges in properties in 6 cases	Critical
P41	No license declared	Important

Given the analysis of the OOPS! tool, we corrected the pitfalls found.

4 EMPIRICAL EVALUATION OF OntoExper-SPL

Based on the Goal-Question-Metric model (GQM), this study aims at: **Evaluating** the OntoExper-SPL ontology, **with the purpose of** characterising its feasibility based on a set of criteria, **with respect to** formalising SPL experiment concepts and querying data from such formalization, **from the point of view of** SPL and Ontology experts, **in the context of** researchers from: State University of Londrina (UEL), University of São Paulo (ICMC-USP), Pontifical Catholic University of Paraná (PUCPR), State University of Maringá (UEM), Federal Universuty of Technology (UTFPR), Sidia Science and Technology Institute (SIDIA), Pontifical Catholic University of Rio Grande do Sul (PUCRS), Federal Institute of Paraná (IFPR) and Paraná University (Unipar).

4.1 Study Planning

Selection of Participants: the experts were invited in a convenient non-probabilistic way, all of them researchers in the SPL and/or Ontologies area. Of the participants, 4 are post-doctorate (23.5%), 6 are PhD candidates (35.3%), 6 are master's (35.3%), 1 is a Master's student (5.9%);

Training: because of the level of knowledge of the participants, it was not necessary to carry out this stage;

Instrumentation: all experts received the following documents: (i) a document containing a brief description of the evaluation, with links to the necessary tools and technologies; (ii) a copy of the Evaluation Instrument characterized by a questionnaire - Google Forms[3]; and (iii) a copy of the ontology metadata composed of: an OWL file of the OntoExper-SPL model, an OWL file of the OntoExper-SPL model populated with 174 individuals, an Excel spreadsheet file containing the original data of the experiments, a copy of all documents outlining the ontology conception;

Evaluation Criteria: The eight quality criteria adopted are based on the work of Vrandecic (2010): accuracy, adaptability, clarity, completeness, computational efficiency, concision, consistency and organizational ability.

To evaluate OntoExper-SPL based on the adopted criteria, we used the following Likert scale: 1 - Totally Disagree (TD); 2 - Partially Disagree (PD); 3 - Neither agree nor disagree (N); 4 - Partly Agree (PA); and 5 - Totally Agree (TA).

[3]https://www.google.com/forms

4.2 Study Execution

A pilot project was conducted with the intention of evaluating the instrumentation of the study, a pilot project was conducted in October 2019, with a Master and a PhD in Computer Science from the State University of Maringá (UEM). The data obtained were discarded, but considerations about errors and improvements in questionnaires were considered.

The full evaluation was conducted following the stages: (i) expert receives the documents, via e-mail; (ii) expert makes a preliminary study of the metadata, clarifies possible doubts; and (iii) expert reads and completes the Evaluation Form - Google Forms, according to their experience.

4.3 Analysis of Results

4.3.1 Profile of the Experts

Forty experts in ES were invited, but only 17 agreed to participate in the quantitative study. It was observed that the average experience in years of the participants is 7.2. The experts E3 and E12 with 15 years of experience and E1 with 0 years of experience stand out. Regarding the level of training, we have four Post-doctors, six doctors, six masters, and one graduate.

4.3.2 Likert Frequency

Table 3 presents the frequency of expert responses for each criterion in relation to the response scale.

Table 3: Frequency (%) of Experts' Responses (mode is in bold).

	%TD	%PD	%N	%PA	%TA
Precision	-	-	2 (11.76%)	6 (35.29%)	**9 (52.94%)**
Adaptability	-	-	1 (5.88%)	5 (29.41%)	**11 (64.71%)**
Clarity	-	1 (5.88%)	-	**8 (47.06%)**	**8 (47.06%)**
Complete	-	1 (5.88%)	5 (29.41%)	**6 (35.29%)**	5 (29.41%)
Computer Efficiency	1 (5.88%)	-	**8 (47.06%)**	3 (17.65%)	5 (29.41%)
Concision	-	-	2 (11.76%)	5 (29.41%)	**10 (58.82%)**
Consistency	-	-	2 (11.76%)	4 (23.53%)	**11 (64.71%)**
Organ. Capacity	-	1 (5.88%)	3 (17.65%)	4 (23.53%)	**9 (52.94%)**

As we can observe based on the empirically evaluated criteria, experts found OntoExper-SPL concepts formalization feasible. It means the organization and interrelationship of concepts makes sense for SPL experiments.

4.3.3 Normality Test

We adopted an statistical approach of the additive method with a continuous dependent variable (DV) to analyze data collected. The DV is the result of summing the value given to each criterion per expert (see Table 4).

Table 4: Sum of the Experts' Response.

Expert	Crit. Sum	Expert	Crit. Sum	Expert	Crit. Sum
E1	32	E7	37	E13	35
E2	37	E8	37	E14	34
E3	30	E9	32	E15	39
E4	32	E10	31	E16	37
E5	34	E11	30	E17	38
E6	40	E12	25		

To analyze the distribution of data we performed the Shapiro-Wilk. The result of the test was 0.94 and *p-value* was 0.44. Thus, we considered the DV distribution as **normal**.

4.3.4 Expert Experience Analysis

We performed a hypothesis test for analyzing data from Table 4 with regard to experts years of experience. To do so, we used the Mann-Whitney test even though the DV was considered normal. We decided for that due to our reduced sample size.

The hypotheses for **Experience in SPL and/or Ontologies (in years)** of the experts are: **Null Hypothesis (H0):** there is no difference in the expert's experience with relation to the observed DV values; and **Alternative Hypothesis:** there is a significant difference in the expert's experience with relation to the observed DV values.

We then created 3 groups representative of years of experience: **1:** 0 to 5 years of experience; **2:** 6 to 10 years of experiments; and **3:** 11 to 15 years of experiments.

The result of the Mann-Whitney test was u = 0, p-value = $0.00000027 < 0.05$, thus we could reject H0 and state **observed DV values depend on the SPL and/or Ontologies Experience**. It means the more experienced is the expert, the more OntoExper-SPL feasibility agreement is.

4.3.5 Correlation of Criteria

We performed correlation analysis between pairs of criteria to potentially make assumptions on the OntoExper-SPL feasibility and for prospective experiments. To do so, we applied the Spearman's correlation coefficient, also due to our reduced sample size.

According to Figure 2, we can state the relationships in Table 5.

Table 5: Quality criteria relationships.

The better	The better
Completenesss	Precision
Consistency	Concision
Organizational Capacity	Computer Efficiency
Computer Efficiency	Adaptability
Organizational Capacity	Adaptability
Concision	Adaptability
Consistency	Clarity
Consistency	Completeness

4.3.6 Validity Evaluation

Internal Validity - Differences between Experts: due to the size of the sample, the variations between the experts' skills were few, thus the experts were not divided into groups.

Internal Validity - Accuracy of Participant Responses: since the OntoExper-SPL related information is presented along with its metadata and considering that the participants are experts in SPL and Ontology, the responses provided are considered valid.

Internal Validity - Fatigue Effect: the OntoExper-SPL files and documentation are complex. In order to mitigate fatigue, the documentation was sent to the experts with a 40-day time frame.

External Validity: obtaining qualified experts in the areas of ES, SPL and Ontology was one of the difficulties encountered in this study and many invited experts could not participate in the study because of their commitments. Although few experts participated in the study, the quality of their profile is the most important variable in this assessment.

Constructo Validity: the instrumentation was evaluated and adequate, according to the pilot project carried out. As for the level of knowledge of the experts in ES, SPL and Ontology, they were satisfactory to evaluate the proposed guidelines.

4.4 Prospective Improvements

During the execution of the study, it was possible to notice that OntoExper-SPL needs to be improved in terms of its computational efficiency. Further improvement points are associate *SPL artifacts* to *SPL domain*, include the axioms *Results*, *Analyze*, *Threats to Validity*, *Subjects* and modify the axiom *Package* to *Replication-Package*.

5 CONCLUSION

OntoExper-SPL stands out for taking into account the SPL specific domain. SPLs are built through an application domain, similarities, core assessments, and variabilities, which distinguishes one product from the other within the product family.

	Precision	Adaptability	Clarity	Completeness	Computer Efficiency	Concision	Consistency	Organ. Capacity
Precision	X	0.27	-0.053	**0.74**	0.27	0.086	0.036	**0.67**
Adaptability	0.27	X	0.061	0.24	0.41	0.32	0.24	0.38
Clarity	-0.053	0.061	X	0.23	-0.063	0.24	0.31	0.21
Completeness	**0.74**	0.24	0.23	X	0.078	0.28	0.29	**0.73**
Computer Efficiency	0.27	0.41	-0.063	0.078	X	0.22	0.092	0.43
Concision	0.086	0.32	0.24	0.28	0.22	X	**0.7**	-0.079
Consistency	0.036	0.24	0.31	0.29	0.092	**0.7**	X	-0.011
Organ. Capacity	**0.67**	0.38	0.21	**0.73**	0.43	-0.079	-0.011	X

Figure 2: Criteria correlations.

We performed a preliminary evaluation of OntoExper-SPL using the OOPS! tool, which revealed several pitfalls. We fixed such pitfalls. Then we further evaluated OntoExper-SPL with an empirical study, which considered eight quality criteria. The results provide preliminary evidence that OntoExper-SPL is feasible to formalize SPL experimentation knowledge with satisfactory quality level.

As future work, we intend to standardize the ontology allowing further generalization. Another goal is to create a broader recommendation system to take into account the ontology formalization and possible inferences to recommend SPL experiments.

REFERENCES

Blondet, G., Le Duigou, J., and Boudaoud, N. (2016). Ode: an ontology for numerical design of experiments. *Procedia CIRP*, 50:496–501.

Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., and Rosati, R. (2005). Dl-lite: Tractable description logics for ontologies. In *AAAI*, volume 5, pages 602–607, USA. ACM.

Cruz, S. M. S., Campos, M. L. M., and Mattoso, M. (2012). A foundational ontology to support scientific experiments. In *ONTOBRAS-MOST*, pages 144–155, ACM. USA.

Cruzes, D., Mendonca, M., Basili, V., Shull, F., and Jino, M. (2007). Extracting information from experimental software engineering papers. In *XXVI International Conference of the Chilean Society of Computer Science (SCCC'07)*, pages 105–114, USA. IEEE.

Furtado, V. R. (2018). Guidelines for software product line experiment evaluation. Master's thesis, State University of Maringá, Maringá-PR. Brazil. in Portuguese.

Garcia, R. E., Höhn, E. N., Barbosa, E. F., and Maldonado, J. C. (2008). An ontology for controlled experiments on software engineering. In *SEKE*, pages 685–690, USA. ACM.

Gelernter, J. and Jha, J. (2016). Challenges in ontology evaluation. *Journal of Data and Information Quality (JDIQ)*, 7(3):11.

Gómez, O. S., Juristo, N., and Vegas, S. (2014). Understanding replication of experiments in software engineering: A classification. *Information and Software Technology*, 56(8):1033 – 1048.

Gómez-Pérez, A. (2004). Ontology evaluation. In *Handbook on ontologies*, pages 251–273. Springer.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2):199–220.

Linden, F. v. d., , Schmid, K., and Rommes, E. (2007). *Software product lines in action: the best industrial practice in product line engineering*. Springer Science & Business Media.

Noy, N. F., McGuinness, D. L., et al. (2001). Ontology development 101: A guide to creating your first ontology.

Pohl, K., Böckle, G., and van Der Linden, F. J. (2005). *Software product line engineering: foundations, principles and techniques*. Springer Science & Business Media, USA.

Poveda-Villalón, M., Gómez-Pérez, A., and Suárez-Figueroa, M. C. (2014). Oops!(ontology pitfall scanner!): An on-line tool for ontology evaluation. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 10(2):7–34.

Prud'hommeaux, E. and Seaborne, A. (2008). SPARQL Query Language for RDF. W3C Recommendation. http://www.w3.org/TR/rdf-sparql-query/.

Scatalon, L. P., Garcia, R. E., and Correia, R. C. M. (2011). Packaging controlled experiments using an evolutionary approach based on ontology (s). In *SEKE*, pages 408–413.

Soldatova, L. N. and King, R. D. (2006). An ontology of scientific experiments. *Journal of the Royal Society Interface*, 3(11):795–803.

Vignando, H., Furtado, V. R., Teixeira, L., and OliveiraJr, E. (2020). Ontoexper-spl - images. (version 1.0). http://doi.org/10.5281/zenodo.3707798.

Vrandecic, D. (2010). *Ontology Evaluation*. PhD thesis, Karlsruher Instituts für Technologie (KIT).

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., and Wesslén, A. (2012). *Experimentation in Software Engineering*. Springer Science & Business Media.