A systematic review process to software engineering

Article · January 2005

CITATIONS
118

READS
477

5 authors, including:

Tayana Conte
Federal University of Amazonas
184 PUBLICATIONS **898** CITATIONS
SEE PROFILE

Jorge Biolchini
Instituto Brasileiro de Informação em Ciência e Tecnologia
7 PUBLICATIONS **513** CITATIONS
SEE PROFILE

Guilherme Horta Travassos
Federal University of Rio de Janeiro
314 PUBLICATIONS **3,622** CITATIONS
SEE PROFILE

A Systematic Review Process for Software Engineering

Paula Mian, Tayana Conte, Ana Natali, Jorge Biolchini and Guilherme Travassos

COPPE / UFRJ – Computer Science Department
Cx. Postal 68.511, CEP 21945-970, Rio de Janeiro – Brazil
{pgmian, tayana, anatali, jorgebio, ght}@cos.ufrj.br

Abstract. *Usually researches start with their research work accomplishing a literature review of some sort. However, unless the review is true, far and repeatable, it is of little scientific value. In this sense, a systematic literature reviews aim at providing the means for carrying out literature reviews that are thorough and unbiased, such that produces scientific value results. This paper describes a systematic way to execute literature review on Software Engineering, presenting a protocol template and corresponding process capable of driving researchers throughout the review conduction process.*

Keywords: *Systematic Review, Secondary Studies, Scientific Methodology, Experimental Software Engineering.*

1. Introduction

A literature review is usually an initial step in any research and development enterprise. It is the means by which the researcher can perform a mapping of the existing and previously developed knowledge and initiatives in the field. Due to its important role in the scientific enterprise, general rules for performing literature overviews have been developed, in order to warrant the investigator good quality of information from the covered material [Biolchini et al. 2005].

The systematic review (SR) consists in a specific scientific methodology that goes one step further than the simple overview. A systematic review is a method that allows specialists obtain relevant and quantified results [Kitchenham 2004],[Kitchenham et al. 2004]. These can lead to the identification, selection, and production of evidence regarding research in a particular topic.

Based on these potential benefits, our research group carried out some literature reviews using this systematic methodology. Our experience, reported in [Mian et al. 2005], showed us that conducting a systematic review is not a simple task. In order to guide researchers to perform systematic reviews in the Software Engineering domain, a Systematic Review Conduction Process and also a Systematic Review Protocol Template to facilitate the conduction of this process have been described [Biolchini et al. 2005].

The objective of this paper is to illustrate the Systematic Review Conduction Process, presenting the model that guides its execution. This article is organized as follows: section 2 presents the SR process. Section 3 describes the protocol template used to guide this process execution. Finally section 4 presents our conclusions and future work.

2. Systematic Review Conduction Process

Some issues on conducting systematic reviews can be extracted from our first SR accomplishments [Mian et al. 2005]. These issues point out to the need of investing research efforts in developing systematic reviews planning and execution methodologies. Therefore, we have described a systematic review conduction process [Biolchini et al. 2005] which aims at guiding researchers to perform systematic reviews in the Software Engineering domain.

Usually, systematic reviews conduction is a three-step approach. The main steps composing the SR process (as shown in Figure 1) are regarding the Planning, Execution, and Result Analysis.

During the *planning phase*, research objectives are listed and a review protocol is defined. Such protocol specifies the central research question and the methods that will be used to execute the review. The *execution phase* involves primary studies identification, selection and evaluation in accordance with the inclusion and exclusion criteria established in the review protocol. Once studies were selected, data from the papers can be extracted and synthesized during the *result analysis phase*. Meanwhile which one of these phases is executed, their results must be stored. Therefore, *systematic review packaging activities* are performed throughout the whole process. There are two checkpoints in the proposed SR process. Before executing the systematic review, it is necessary to guarantee that the planning is feasible. The protocol must be evaluated and whether problems are found, the researcher must return to the planning phase to review the protocol. Similarly, when problems regarding web search engines are found during the execution phase, the systematic review must be re-executed. Despite the shape of the figure, the proposed SR process is not purely sequential having the phases a lot of iteration. . In particular, many activities are initiated during the protocol development activity (Planning phase), and refined when the review properly takes place [Biolchini et al. 2005].

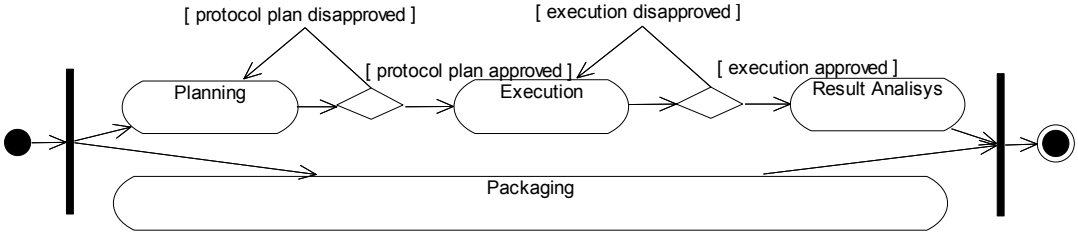

Figure 1. Systematic Review Conduction Process.

3. The Systematic Review Protocol Template

A review protocol template has been developed to reduce the overload of systematic reviews planning and execution. This template aggregates and formalizes ideas regarding SR protocols developed in the medical area, guidelines for systematic review in SE proposed by [Kitchenham, 2004] and the protocol example described in [Mendes and Kitchenham, 2004]. Besides, it also introduces a process description to support the template usage.

The objective of this template is to serve as a guideline to Software Engineering researchers when conducting the systematic review. Therefore, the template lead researchers through each step of the systematic review process presented previously,

defining clearly the content of each protocol section. In the next sections, we describe the SR process in terms of the protocol template items.

3.1. Review Planning

In this phase, it must be defined the research objectives and the way the review will be executed, which includes to formulate research questions and to plan how the sources and studies selection will be carry out. The sections of the protocol template that guide the planning phase are shown in Figure 2.

1. **Question Formularization:** the research objectives must be clearly defined, filling the items:
   - 1.1. **Question Focus:** defines the systematic review focus of interest, i.e., the review research objectives. Here, the researcher must decide what he/she expects to be answered in the SR end.
   - 1.2. **Question Quality and Amplitude:** this section aims at defining the research question syntax (the context in which the review is applied and the question the study must answer) and its semantics specificity (or question range) described by the remaining items of this section:
     - **Problem:** defines the systematic review target, describing briefly the research context.
     - **Question:** research question to be answered by the systematic review. It is important to highlight that, if the systematic review context is too wide, it may be necessary to decompose the research question in secondary questions to narrow the research target.
     - **Keywords and Synonyms:** list of the main terms that compose the research question. These terms will be used during the review execution (in case the search by keywords is chosen as study selection methodology).
     - **Intervention:** what is going to be observed in the context of the planned SR.
     - **Control:** baseline or initial data set that the researcher already posses.
     - **Effect:** types of results expected in the end of the systematic review.
     - **Outcome Measure:** metrics used to measure the effect.
     - **Population:** population group that will be observed by the intervention.
     - **Application:** roles, professional types or application areas that will benefit from the systematic review results.
     - **Experimental Design:** describes how meta-analysis will be conducted, defining which statistical analysis methods will be applied on the collected data to interpret the results.
2. **Sources Selection:** the objective of this section is to select the sources where searches for primary studies will be executed.
   - 2.1. **Sources Selection Criteria Definition:** defines which criteria are going to be used to evaluate studies sources, i.e., which characteristics make these sources candidate to be used in the review execution.
     - **Studies Languages:** it defines the languages in which obtained primary studies must be written. This item belongs to this section, and not to "Studies Selection", because the chosen language may constrain the sources identification.
   - 2.2. **Sources Identification:** this item aims at selecting sources for the review execution.
     - **Sources Search Methods:** describes how to execute the search for primary studies (for instance, manual search, search through web search engines).
     - **Search String:** case one of the selected search methods includes using keywords in search engines it is necessary to create search strings to be run at such engines. This item presents a set of logical expressions that combine keyword and its synonymous arranged in a way that highest amount of relevant studies is obtained from search engines.
     - **Sources List:** initial source list in which the systematic review execution will be run.
   - 2.3. **Sources Selection after Evaluation:** which element of the initial sources list must be evaluated according to the source selection criteria. If the source fits all criteria, it must be included in the final sources list, presented in this session of the protocol.
   - 2.4. **References Checking:** one or more experts must evaluate the sources list obtained from the previous item. Case the experts find the need to add new sources or to remove some of them, the result of such evaluation must be described in this item.
3. **Studies Selection:** once the sources are defined, it is necessary to describe the process and the criteria for studies selection and evaluation.
   - 3.1. **Studies Definition:** this item defines the way studies will be selected.
     - **Studies Inclusion and Exclusion Criteria Definition:** presents the criteria by which studies will be evaluated to decide if they must be selected or not in the SR context. It is necessary to define these criteria because a search executed in web engines may find a great number of papers that do not answer the research question. The main reason for this to happen is that a keyword may have different meanings or be used in studies that do not deal with the SR research topic. Therefore, it is necessary to define what makes a paper a potential candidate to be selected or to be excluded from the review.
     - **Studies Types Definition:** it defines the types of primary studies that are going to be selected during the systematic review execution: Qualitative or quantitative; observation; feasibility or characterization studies.
     - **Procedures for Studies Selection:** it describes the procedure by which the studies will be obtained and evaluated according to exclusion and inclusion criteria. If the selection process has more then one stage, all of them must be described. Studies selection procedures examples are reading the paper abstract or the full text.

Figure 2. SR protocol template: Planning phase.

3.2. Planning Evaluation

Before executing the systematic review, it is necessary to evaluate the planned review. A way to perform such evaluation is to ask experts to review the protocol. Another way to evaluate the planning is to test the protocol execution. The review is executed in a reduced set of selected sources. If the obtained results are not suitable, the protocol must be reviewed and a new version must be created [Biolchini et al. 2005].

3.3. Review Execution

After evaluating the planning, the systematic review execution can be initiated (Figure 3). During this phase, the search in the defined sources must be executed and the studies obtained must be evaluated according to the established criteria. Finally, the relevant information regarding the research question must be extracted from the selected studies.

3.2. Selection Execution: this section aims to register the primary studies selection process, reporting the obtained studies and the results of their evaluation.
- **Initial Studies Selection:** the search in itself is executed and all the obtained studies must be listed for further evaluation.
- **Studies Quality Evaluation:** the procedures for studies selection are applied to all obtained articles in order to verify if the studies fit the inclusion and exclusion criteria. Moreover, it must be checked if the studies belong to the types selected during the planning phase. The objective of this section is to register the results of this evaluation.
- **Selection Review:** studies selection must be reviewed to guarantee that the studies quality evaluation does not eliminate relevant articles. Here, independent reviewers may be useful. The results of the review must be recorded in this item.

4. Information Extraction: once primary studies are selected, the extraction of relevant information begins. In this protocol section, extraction criteria and results are described.

4.1. Information Inclusion and Exclusion Criteria Definition: criteria by which the information obtained from studies must be evaluated.

4.2. Data Extraction Forms: to standardize the way information will be represented, the researcher must create forms to collect data from the selected studies. These forms may vary depending on the systematic review's objective and context.

4.3. Extraction Execution: two kinds of results can be extracted from the selected studies: objective and subjective results.
- **Objective Results Extraction:** objective results are those that can be extracted directly from the selected studies. Such results must be organized as follow:
  - **i)Study Identification:** studies identification includes the publication title, its authors and the source from which it was obtained.
  - **ii)Study Methodology:** methods used to conduct the study.
  - **iii) Study Results:** effect obtained through the study execution.
  - **iv) Study Problems:** study limitations found by the article's authors.
- **Subjective Results Extraction:** subjective results are those that cannot be extracted directly from the selected studies. There are two ways to obtain such results:
  - **i) Information through Authors:** reviewers contact the study's authors to solve doubts or to ask for to more details about it.
  - **ii) General Impressions and Abstractions:** reviewers raise their own conclusions after the reading the study.

4.4. Resolution of divergences among reviewers: if reviewers don't agree on the information extracted from the studies, the divergences must be recorded. The reviewers must reach a consensus on this matter and register it in this section.

Figure 3. SR protocol template: Execution phase.

3.4. Execution Evaluation

Our experience on conducting systematic reviews has shown that, during the execution phase, several problems may occur due to web search engines limitations [Mian et al. 2005]. Therefore, the proposed systematic review process suggests evaluating web

search engines at the execution phase to verify whether they are able to perform the search strings previously defined during the planning phase. If so, the process may go on. Otherwise, it may be necessary to exclude a digital source selected or to rework the search strings to fit the search machine restrictions [Biolchini et al. 2005].

3.5. Result Analysis

After the systematic review execution, the results must be summarized and analyzed using the statistical methods defined during the planning phase (Figure 4).

5. Results Summarization: this systematic review protocol section aims to present the data resulting from the selected studies.

5.1. Results Statistical Calculus: statistical methods chosen in the "Experimental Design" section are applied to analyze data and to understand the complexity relations between obtained results.

5.2. Results Presentation in Tables: the results obtained from the systematic review must be displayed in tables to facilitate analysis. Tables allow to classify studies according to different criteria and to organize them under different perspectives.

5.3. Sensitivity Analysis: result robustness must be verified, investigating if there were uncertainties about including or excluding certain studies. Sensitivity analysis is more important when a complete meta-analysis is performed.

5.4. Plotting: a data plotting strategy may be chosen to present the results. Likewise sensitivity analysis, plotting is indicated when meta-analysis is performed.

5.5. Final Comments: this item presents reviewers final comments about the SR results.

- **Number of Studies:** quantity of obtained and selected studies.
- **Search, Selection and Extraction Bias:** if any search, selection or information extraction biases that can invalidate the systematic review results are identified by the reviewers, they must be described here.
- **Publication Bias:** it refers to the problem that positive results are more likely to be published than negative results since the concept of positive or negative results sometimes depends on the viewpoint of the researcher.
- **Inter-Reviewers Variation:** conflict resolution between reviewers regarding the systematic review results.
- **Results Application:** defines how the obtained systematic review results can be applied.

5.6. Recommendations: reviewers' suggestions on how the systematic review results must be applied.

Figure 4. SR protocol template: Result Analysis phase.

4. Conclusion and Future Work

This paper introduced a proposal for a systematic review template and conduction process to guide software engineering researchers in the accomplishment of secondary studies. Both SR template and process have been used by the Experimental Software Engineering team members at COPPE/UFRJ to perform systematic reviews concerned with different SE topics. To support the evaluation of such proposal, a pilot study had been conducted [Biolchini et al. 2005]. This study aimed at verifying the feasibility of using this proposal to drive researchers throughout the systematic review activities. Besides, it was also possible to observe how researches conducted their processes. The results of this study, fully described in [Biolchini et al. 2005], allowed making the improvements in the template structure, presented in this paper.

The use of this template to conduct systematic reviews and our experience have shown that it is required an additional conduction effort when compared to unsystematic reviews. The review must be planned prior its execution, and the whole process must be documented, including the intermediary results. However, most of the efforts are concentrated on the searching and retrieval activities, which still represent a bottleneck in the review process.

The obtained results make us ascribe the difficulties in the search and retrieval activities mainly to the current search mechanisms that do not adequately support the information representation necessary for conducting systematic reviews in Software Engineering. Scientific papers on SE are spread across different research organizations digital libraries. Many of these papers are not public, i.e., it is not possible to get their full texts unless the researcher has free access to the databases. Not having a sufficient access degree to these digital libraries can compromise the systematic review completeness and replication [Mian et al. 2005].

Besides, the methodological rigor in applying a systematic review requires more explicit and detailed documentation of all produced results, throughout systematic review process phases. Therefore, we believe that Software Engineering researches would benefit from a computational infrastructure to support the SR process and to package secondary studies results, such as those supported by the eSEE infrastructure [Mian et al. 2004].

Acknowledgments

The authors acknowledge CAPES, CNPq and FAPEAM for the financial support to this work. The authors especially thank professors Barbara Kitchenham and Emilia Mendes for sending the initial materials regarding this area.

References

Biolchini, J., Mian, P.G., Natali, A.C. and Travassos, G.H. (2005) “Systematic Review in Software Engineering: Relevance and Utility”, Technical Report ES67905, PESC - COPPE/UFRJ. Available at http://cronos.cos.ufrj.br/publicacoes/reltec/es67905.pdf

Kitchenham, B. (2004) “Procedures for Performing Systematic Reviews”, Joint Technical Report Software Engineering Group, Keele University, United Kingdom and Empirical Software Engineering, National ICT Australia Ltd, Australia.

Kitchenham, B. A., Dyba, T. and Jorgensen, M. (2004) “Evidence-based Software Engineering”, In Proc. of 26th International Conference on Software Engineering (ICSE 2004), Scotland.

Mendes, E. and Kitchenham, B. (2004) "Protocol for Systematic Review”, (http://www.cs.auckland.ac.nz/emilia/srspp.pdf). Last accessed by 05/10/2005.

Mian, P.G., Travassos, G.H and Rocha, A.R.C. (2004) “Towards a Computerized Infrastructure for Experimental Software Engineering”, In Proc. of the 1st Experimental Software Engineering Latin American Workshop (ESELAW'04), Brazil.

Mian, P., Conte, T., Natali, A., Biolchini, J., Mendes, E. and Travassos, G. (2005) “Lessons Learned on Applying Systematic Reviews to Software Engineering”, In Proc. of the 3rd International Workshop ‘Guidelines For Empirical Work’ in the Workshop Series on Empirical Software Engineering (WSESE2005), Finland.