An Approach and a Tool for Systematic Review Research

Manuel Gonçalves da Silva Neto, Walquiria Castelo Branco Lins, Eric B. Perazzo Mariz
Recife Center for Advanced Studies and Systems (CESAR)
Recife – PE, Brazil
Emails: {manuel.pi, wcbl, eric.perazzo}@cesar.edu.br

***Abstract*— Systematic Reviews and Systematic Mappings are widely used in medicine in an area called *evidence-based studies*. Recently, these techniques have been adapted and used in secondary studies in the area of Software Engineering and Systems. Sorting and synthesizing information in a particular research area by analysis of their primary studies, both involve both extensive work and researcher dedication. Adapting techniques applied in *evidence-based studies* in the medical field to software engineering led to an approach, which divides Systematic Review tasks into three main phases, namely, planning, conduct of review itself and reporting the results. Unlike in the area of medicine, in which there are many research groups and methodologies to support these tasks, researchers in the area of software engineering still lack tools and methods that support the implementation of these activities and, in general, they need to use software that was not designed for this purpose. This paper presents an approach based on Biolchini's proccess, using checkpoints techniques, to assist in maintaining of the main objectives of the review process; these tasks were supported by a management software. The software facilitates the execution of repetitive tasks of recording, quantifying and classifying of data in accordance with a predefined research protocol in the planning phase, thereby enabling studies to be better organized and an overview to be obtained in the early stages of the review. We used a Systematic Review theme to validate the approach and supporting tool. This article shows that by visualizing and classifying research data while still at the initial stages of a Systematic Review, problems may be identified in the design of the protocol (planning phase), which otherwise would only be detected in the final stages, when results are being generated.**

Keywords- Systematic Review; Systematic Mapping; Support Tool.

I. INTRODUCTION

A Systematic Literature Review (SLR) is a way to identify, evaluate and interpret all relevant research available on a particular research question, study area or phenomenon of interest. Individual studies that contribute to a Systematic Review are known as the primary studies; a Systematic Review is a kind of secondary study [1].

The term Systematic Review is used to refer to a specific research methodology developed to obtain and evaluate evidence in a particular topic or research area [2]. In general, a Systematic Review involves three phases: *i)* Planning a review or developing a research protocol, *ii)* Running or conducting the review, and *iii)* Reporting the results [1]–[3].

The area of Software Engineering has shown interest in evidence-based studies where the presence of experimental software engineering is becoming more and more common in large events of the area [2][5][6].

The main problem tackled in this article is in tasks related to planning and conducting the research protocol, which require the researcher to be extensively dedicated and highly organized since he/she must catalog and classify the primary research and perform quantitative and qualitative analyzes, in order to get a broad view of the object of study which will facilitate the generation of results [5][6].

This paper proposes an approach based on the process proposed by Biolchini *et al.* [2] for the tasks of planning and conducting Systematic Reviews. A computer system stores the data progressively during the review and summarizes the results by each research protocol. The system seeks to reduce the time and effort needed for this process by eliminating the need to transfer information between various software programs, a situation that arises when these have not been designed to support the review process appropriately.

This article shows that by visualizing and classifying research data while still at the initial stages of a Systematic Review, problems may be identified in the design of the protocol (planning phase), which otherwise would only be detected in the final stages when results are being generated. In order to validate the approach and the software system, these were used to conduct a Systematic Review of a Masters Dissertation [9].

This paper is organized as follows: Section 2 presents the theoretical framework for Systematic Reviews in the field of software engineering and related studies; Section 3 presents the approach, Section 4 presents a case study where the approach and the system support were experimentally used and Section 5 presents the conclusions and final considerations.

II. SYSTEMATIC REVIEWS

This section presents the concepts that comprise the use of Systematic Reviews in the field of software engineering and related studies.

A. Systematic Reviews in the Context of Software Engineering

Any consolidated research area ends up producing a lot of papers and results, which require summarization and classification, therefore, enabling a broader understandig of the field [3]. Although Systematic Reviews are widely used in medicine, they have only recently attracted the interest of researchers in software engineering [1]–[3].

In brief, the difference between Systematic Reviews and literature reviews can be seen in the way they are conducted. In a Systematic Review, there are a rigorous methodologies in which criteria for including or excluding a research study and steps for conducting the research are pre-specified through in a research protocol; unlike a literature review, which presents the studies analyzed without details of how they were chosen [2]. In literature reviews, there are no explicit explicit criteria for including and excluding studies in a review nor is other information given which wold enable a Systematic Review to be scientifically replicated or extended in a methodologically rigorous way.

The characteristic of Systematic Mappings is that they mainly focus on generating results in visual form, mapping itself being a particular area of research. On comparing Systematic Reviews and Systematic Mappings, it can be concluded that both involve the same methodological rigor, and they are often used, loosely, as synonyms. The main difference between them can be found in their goals and not in their methods [4].

Kitchenham [1] adapted the guidelines given to conduct reviews in medicine, including the best known one, "*The Cochrane Reviewer's Handbook*" [10], to the specific area of software engineering. Systematic Review were divided into three phases, namely, *i)* Planning, *ii)* Conducting, and *iii)* Reporting results. He also discusses the reasons for conducting a SRL in the area of software engineering:

- Summarize existing evidence about a technology.
- Identify gaps in current research in order to suggest areas for future research.
- Provide knowledge on new research activities.

The tasks related to the production of secondary studies by Systematic Reviews are carried out in three (3) distinct phases in several papers in the literature: In [5], a model, similar to those found in [1] and [2], was used. It is also performed in three phases, namely, developing a research protocol, conducting the research and reporting the results.

The model used by Montoni [5] was also used in the Systematic Review by Barcelos [6].

The phases of planning and implementing research which precede the step of generating the results, require greater manual effort by the empirical researcher. These phases involve defining of all the protocol items related to the research questions; the stages of selecting the criteria and excluding items. In addition, they include defining the sources in which searches for studies will be conducted, the primary language publications and the period the search will cover. The research has to strictly maintain the criteria defined in the protocol so as to avoid the search generating biased results.

In [2], Biolchini *et al*. presents a template for performing Systematic Reviews where the incremental use of the following process is recommended.

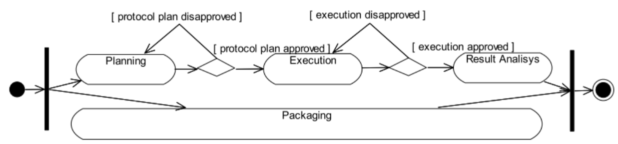

Figure 1. Proccess Overview [2].

The number of Systematic Reviews has grown insoftware engineering. The literature contains studies that propose solutions for problems found at various stages of the review process. Dieste *et al*. [7] gives strategies and methods for detecting relevant experiments, in particular draws attention to creating effective keywords for searches.

Montoni [5] and Barcelos [6], their implementation phase of the review, used a relational database to store information gradually with regard to title, author (s), year of publication, event, search source, classification and a brief summary of each article published is made while conducting the search so as to facilitate reaching the generation of the final results. Laguna and Crespo [8] used the *Mendeley* [12] software for managing data in review. Despite these studies having guidelines proposed by Kitchenham [1] in common, note that they use and adapt different tools and techniques for the same purpose, besides which their level of detail and the stages of the process are different in each study.

III. Proposed Approach

Based on three main steps, as in [1]–[3], wich comprise a Systematic Review, our approach proposes to conduct the review in compulsory and interdependent steps with the presence of checkpoints to guide inexperienced reviewers in a step-by-step style. We used a system for storing, retrieving and classifying information and tasks; that comprise a review, the *revision manager* [11].

The figure below shows an overview of the proposed methodology. A tool to support the process allows the insertion of checkpoints which enables the results of alignment with the objectives of the review and modification of the protocol for error correction to be aligned incrementally.

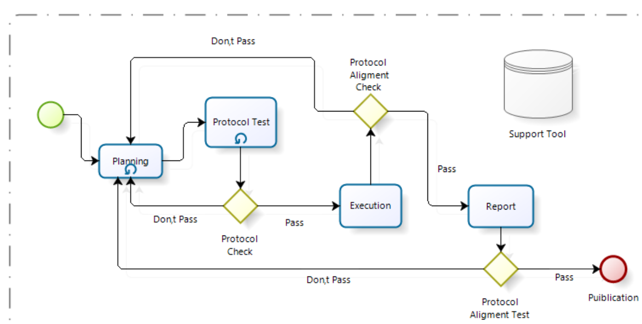

Figure 2. Proposed Approach [9].

The planning phase includes the construction of the research protocol itself, when its main items were defined. In [5], Montoni recommends defining the following items in the

 ISBN: 978-1-61208-367-4

planning phase: Context of the study, objectives to be achieved, research questions, research sources, languages, methods used in the search for primary publications, procedures and criteria for including studies, procedures for data extraction and analyzing results. It should also establish procedures and metrics for implementing the testing protocol.

The implementation phase is when the pre-defined protocol is followed so as to obtain primary research studies. Finally, we make a quantitative and qualitative analysis of the data in order to get an overview of the object of study and publish the results.

A tool was developed to support the phases of the Systematic Review. The Revision Manager (RM) [11], was developed using the Django web development framework [13], relational database MySQL [15] and a Javascript library for creating graphical reports, the Hightcharts [14].

The central idea of the tool is to avoid change in the working environment as well as to generate reports while still in the early phases, thereby enabling problems to be identified at an early stage.

IV. TOOL OVERVIEW

Currently in version 1.0, the Systematic Review Supporting Tool [11] is able to store and manage data of the review itself as well as information concerning the evaluation of each article, including information regarding the steps, research sources and selection criteria. This information is used to clustering and classify articles. The system is multi user, wich allows each user to access only data related to his/her own work. There is also the isolation of a review, where it is available only to the person who conducted the review.

Each stored item has a Create, Read, Update and Delete (CRUD) functionality to manage data thus allowing information regarding the protocol to be refined during the research. The figure shows the ERD (Entity Relationship Diagram) system.

Based on the process used, here is an overview of the workflow of the supporting tool:

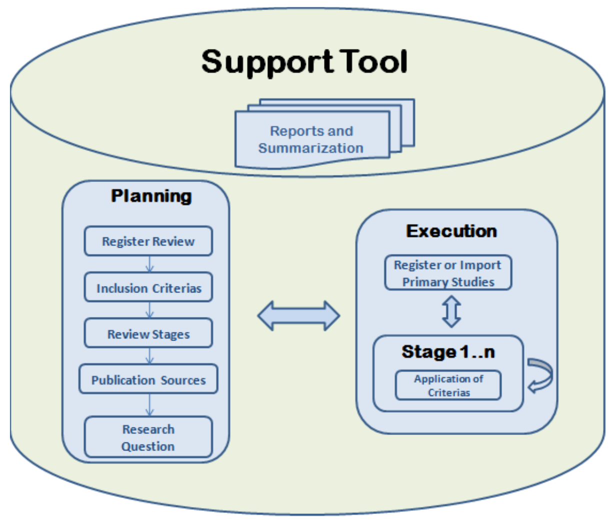

Figure 3. Tool Flow.

Figure 3 shows the incremental use of the tool by allowing reports to be generated at any point in the process. This procedure facilitated the identification of errors and the need to change the protocol.

Figure 4 reveals the presence of relationships which permit information of each step of the review to be stored and retrieved. This feature is essential so that reports can be generated and actions performed during the search traced.

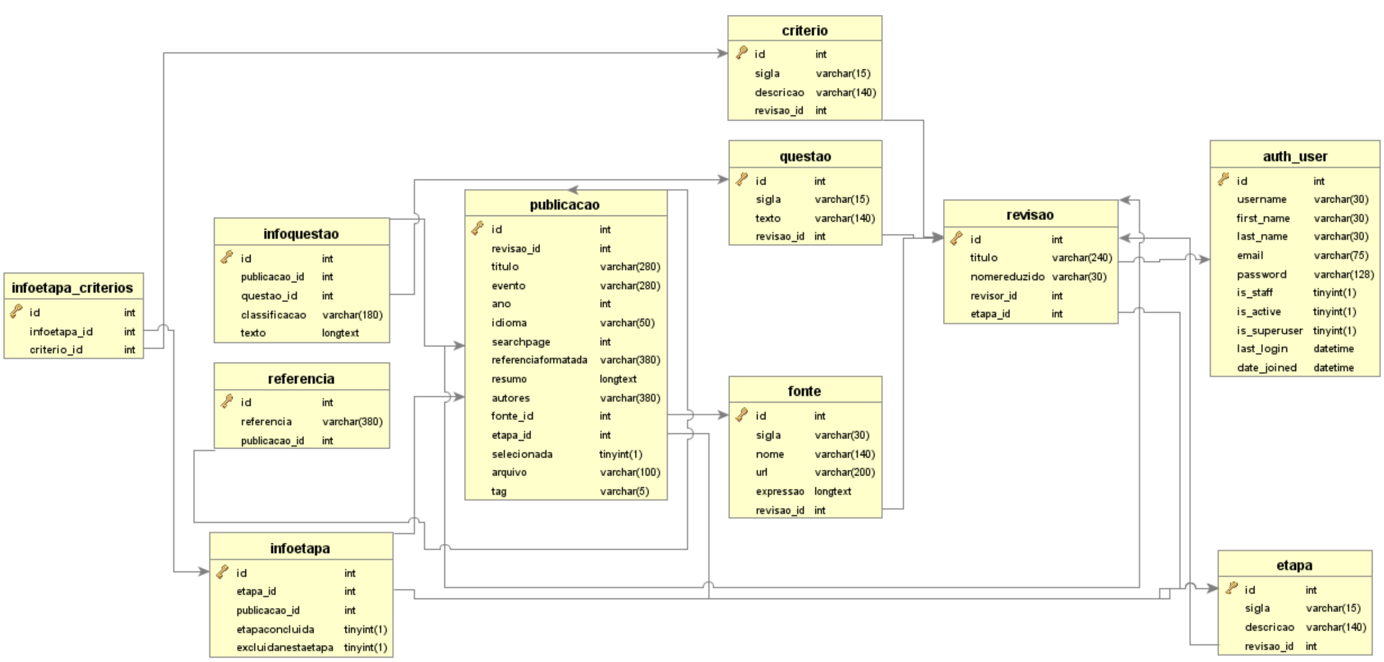

Figure 4. Entities and Relations Diagram.

V. USE CASE

We used the tool to support the management of a Systematic Review methodology of a final Master's project [9]. The planning phase followed the models of Systematic Reviews present in [1][2][5][6], was adapted for specific research items.

After proper registration in the review system, protocol data were stored and alignment criteria were specified. These criteria were evaluated at various points of the review. The alignment criteria are defined taking into account quantitative articles related to the theme and their relevance with regard to answering the research questions of the review.

In the planning phase of the protocol, it was decided that conduct phase would be divided into 3 stages (not to be confused with the general phases of the review). In each subdivision, a set of inclusion and exclusion criteria were applied to articles and every action was recorded by the system. The following is a summary of each subdivision:

a) Step 1 (E1): Obtaining articles by applying the search expression in research sources.

b) Step 2 (E2): Applying the selection criteria to the title and summary.

c) Step 3 (E3): Applying the selection criteria when reading the full text.

During the application of filters, checking every criterion was checked for alignment with the protocol as proposed in [6]. This action sought to identify if there was a need to refine the protocol. There follows a summary of the rules for alignment:

a) In relation to the number of articles before applying the filters: A high number may indicate that the search expression must be refined because it covers an area larger than the one desired. On the other hand, a very small volume may indicate the premature elimination of relevant publications.

b) In relation to the number of articles after applying the set of filters and selection criteria: A very high number or one very close to the number obtained in the early stages may indicate that there are unnecessary steps or that the criteria are too close to the original search expression.

The following is a summary of the results obtained at the beginning of the review:

TABLE I. SOURCES.

Source	Address
Compendex(CPX)	www.engineeringvillage.com
ScienceDirect(SCD)	www.sciencedirect.com

Table I shows the digital sources used. In the first iteration, we used the expression for searching databases. The period used as a search criterion was from January 2009 to January 2014. The results obtained are shown in Table II and Table III:

TABLE II. FIRST EXECUTION.

STAGE II			
Source	**Articles**	**Excluded**	**Approved**
CPX	152	50	102
SCD	483	433	50
Total	**635**	**483**	**152**

Due to the high rate of exclusions based on analyzing the abstract and title (E2), we opted not to proceed immediately to the next step, which comprises the full reading of the articles approved. We, therefore, interrupted the process in order to hone the search expression and selection criteria.

In the second iteration, after a terse expression search, and modifying the search period so that it ran from January 2010 to January 2014 and include new criteria for inclusion and exclusion the following quantitative data were yielded:

TABLE III. SECOND EXECUTION.

Stage (E1)			
Source	**Articles**	**Excluded**	**Approved**
CPX	94	0	94
SCD	188	0	188
Stage (E2)			
Source	**Articles**	**Excluded**	**Approved**
CPX	94	29	65
SCD	188	178	10
Stage (E3)			
Source	**Articles**	**Excluded**	**Approved**
CPX	65	36	29
SCD	10	10	0
Approved: 29			

Due to all articles obtained from *ScienceDirect* being excluded, we chose to remove it from the list of sources of the research protocol and to make refinements before starting a new implementation phase. Further details about the refinements and overall results can be found in [9].

VI. THREATS TO VALIDITY

The following two itens sumarises the main threats to the validity of the results of a this study:

(i) The proposed process was not measured, or no empirical experiments were conducted (apart from it being used in one case study). To further validate a process like this others factors should be considered as the number of articles in similar reviews. This can be used as a parameter.

(ii) The data in this study were not statistically analyzed with the standard measurements (mean, standard deviation, etc.) because the process was tested only once. Different iterations will help to understand the process and how it behaves.

VII. CONCLUSIONS AND FUTURE WORK

This paper presented an approach for conducting Systematic Reviews using checkpoints to check the alignment of the results obtained and expected. We used a tool to support data management and to generate quantitative results in the early stages of the review.

Given the excessive number of manual and repetitive tasks that are involved in this type of research it is believed that this approach together with the software support proposed contributes in particular to helping researchers who have no experience in Systematic Reviews. The proposed approach enable the researchers to focus their efforts on tasks related to qualitative empirical analysis while the quantitative analysis and classification is performed by the software support system.

We intend to provide the system and its user manuals available for use by the academic community. We also intend to validate it by conducting other reviews. We hope that users will add features they require, thereby contributing to the improvement of the system and the approach.

REFERENCES

[1] B. Kitthenham, Procedures for Performing Systematic Reviews. Technical Report TR/SE-0401: Keele University and National ICT Australia Ltd, 2004, p. 33.

[2] J. Biolchini, P. G. Mian, Ana Candida Cruz Natali, and G. H. Travassos, Systematic Review in Software Engineering, no. May. Technical Report RT - ES 679/05: COPPE/UFRJ/Programa de Engenharia de Sistemas e Computação/Rio de Janeiro-RJ, 2005, p. 31.

[3] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic mapping studies in software engineering," 12th Int. Conf. Eval. Assess. Softw. Eng., pp. 71–80, 2008.

[4] F. Q. B. da Silva, M. Suassuna, R. F. Lopes, T. B. Gouveia, a. C. a. Franca, J. P. N. De Oliveira, L. F. M. De Oliveira, and A. L. M. Santos, "Replication of Empirical Studies in Software Engineering: Preliminary Findings from a Systematic Mapping Study," 2011 Second Int. Work. Replication Empir. Softw. Eng. Res., pp. 61–70, Sep. 2011.

[5] M. A. Montoni, Uma investigação sobre os fatores críticos de sucesso em iniciativas de melhorias de processos de software. UFRJ/COPPE/ Programa de Engenharia de Sistemas e Computação, Rio de Janeiro - RJ: Tese (Doutorado), 2010, p. 400.

[6] M. P. Barcelos, Uma estratégia para medição de software e avaliação de bases de medidas para controle estatístico de processos de software em organizações de alta maturidade. UFRJ/COPPE/ Programa de Engenharia de Sistemas e Computação, Rio de Janeiro - RJ: Tese (Doutorado), 2009, p. 433.

[7] O. Dieste and A. G. Padua, "Developing Search Strategies for Detecting Relevant Experiments for Systematic Reviews," First Int. Symp. Empir. Softw. Eng. Meas. (ESEM 2007), pp. 215–224, Sep. 2007.

[8] M. a. Laguna and Y. Crespo, "A systematic mapping study on software product line evolution: From legacy system reengineering to product line refactoring," Sci. Comput. Program., vol. 78, no. 8, pp. 1010–1034, Aug. 2013.

[9] M. G. da S. Neto, Uso de SOA na modernização de Softwares Legados: Um Mapeamento Sistemático da Literatura. CESAR-edu/MPES/Recife-PE: Dissertação (Mestrado), 2014, p. 163.

[10] Cochrane Collaboration. Cochrane Reviewers' Handbook. Version 4.2.1. December 2003

[11] M. G. da S. Neto, (retrieved: 2014, August 20). Review Manager Source Code [Online]. Available: https://github.com/academicsgit/reviewmanager

[12] Mendely Inc., (retrieved: 2014, August 20). Mendeley Webpage [Online]. Available: http://www.mendeley.com

[13] Django Home, (retrieved: 2014, August 20). Django Framework Webpage [Online]. Available: http://www.djangoproject.com

[14] Hightcharts Home, (retrieved: 2014, August 20). Hightcharts Webpage [Online]. Available: http://www.hightcharts.com

[15] MySQL Inc., (retrieved: 2014, August 20). MySQL Webpage [Online]. Available: http://www.mysql.com

 ISBN: 978-1-61208-367-4