Pós-Graduação em Ciência da Computação

Vilmar Santos Nepomuceno

Guidelines for Maintaining Systematic Literature Reviews in Software Engineering

Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
http://cin.ufpe.br/~posgraduacao

Recife
2019

Vilmar Santos Nepomuceno

Guidelines for Maintaining Systematic Literature Reviews in Software Engineering

A Ph.D. Thesis presented to the Center for Informatics of Federal University of Pernambuco in partial fulfillment of the requirements for the degree of Philosophy Doctor in Computer Science.

Concentration Area: Software Engineering
Advisor: Sérgio Castelo Branco Soares

Recife
2019

Catalogação na fonte
Bibliotecária Mariana de Souza Alves CRB4-2106

N441g Nepomuceno, Vilmar Santos.
Guidelines for Maintaining Systematic Literature Reviews in Software Engineering/ Vilmar Santos Nepomuceno. – 2019.
103 f.: il., fig., tab.

Orientador: Sérgio Castelo Branco Soares.
Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da computação. Recife, 2019.
Inclui referências e apêndices.

1. Engenharia de Software. 2. Guideline. 3. Rastreabilidade. 4. Revisão Sistemática de Literatura. I. Soares, Sérgio Castelo Branco (orientador). II. Título.

005.1 CDD (22. ed.) UFPE-MEI 2019-159

Vilmar Santos Nepomuceno

“Guidelines for Maintaining Systematic Literature Reviews in Software Engineering”

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Ciência da Computação.

Aprovado em: 26/08/2019.

_____

Orientador: Prof. Dr. Sérgio Castelo Branco Soares

BANCA EXAMINADORA

_____

Prof. Dr. André Luis de Medeiros Santos
Centro de Informática/UFPE

_____

Prof. Dr. Leopoldo Motta Teixeira
Centro de Informática/UFPE

_____

Prof. Dr. Eduardo Henrique da Silva Aranha
Departamento de Informática e Matemática Aplicada/UFRN

_____

Prof. Dr. Alberto Cesar Cavalcanti Franca
Departamento de Computação/UFRPE

_____

Prof. Dr. Bruno Falcão de Souza Cartaxo
Instituto Federal de Pernambuco

I dedicate this result to my family, my advisor, all the friends who helped me in the construction of the thesis and to God.

ABSTRACT

Systematic Literature Review (SLR) is a process in which a representative set of relevant available research about a research question, or topic, or phenomenon of interest is identified, evaluated, and interpreted through their individual studies, called primary studies, during SLR process. Kitchenham's guidelines summarize systematic review process into three main phases: Planning, Conducting, and Reporting results. We argue SLR process should not end after reporting phase. The main purpose of maintain SLRs is to keep your evidence as up-to-date as possible. The main goal of this research is, based on researchers opinion who have already performed SLRs, to propose and evaluate a guideline to carry out systematic literature reviews maintenance in software engineering field. In this work we make a parallel with software maintenance concepts, adapting them to SLR concepts. Understanding what kind of changes can lead to a SLR maintenance, how to assess changes impact and how process should be conducted. Three steps were carried out: Investigation, Proposition, and Evaluation. In investigation step, two studies were conducted in parallel to serve as a basis for our work. First we conduct a systematic mapping (SM) to address the state-of-the-art on systematic literature reviews update, then we conduct a survey aimed to identify researchers opinion about maintenance of SLRs. In proposition step, we aim to define a guideline for conducting maintenance in systematic literature reviews. In evaluation step, to evaluate the proposed guideline, we conducted a series of semi-structured interviews. During SM, it was possible to check secondary studies updates since 2010, however if we compare with the amount of SLR studies already published, there is still a long way to go. We found that surveyed researchers have shown interest in keeping their SLRs up-to-date, but they have expressed concerns about the effort to be made to accomplish it. During interview the guideline was presented and process acceptance is perceived as basis for conducting SLR maintenance activities. Various contributions have been taken into account and incorporated into process. The guideline proposed was well accepted by the researchers, and besides the evaluation of its effectiveness as a guide for SLRs maintenance process, it was necessary to elaborate several discussions, involving the theme, for a better understanding about the topic. Important concepts were presented and validated during research. However, this is a first glimpse into SLRs maintaining process, and like any process formalization, still lacks depth. Some gaps in compliance or formalization may still be left blank. However, it is hoped with future community collaborations, this guideline may become a basis for conducting maintenance activities on SLRs.

Keywords: Guideline. Updates. Maintenance. Traceability. Systematic Literature Review. Evidence Based Software Engineering.

RESUMO

A Revisão Sistemática da Literatura (RSL) é um processo no qual grande parte das pesquisas relevantes disponíveis sobre uma questão de pesquisa, tópico ou fenômeno de interesse são identificadas, avaliadas e interpretadas através de seus estudos individuais durante o processo de RSL. O guideline proposto por Kitchenham resume o processo de revisão sistemática em três fases principais: Planejamento, Condução e Relatório de resultados. Argumentamos que o processo de RSL não deve terminar após a fase de relatórios. O principal objetivo de manter as RSLs é manter os achados o mais atualizados possível. O objetivo principal desta pesquisa é, com base na experiência de pesquisadores que já realizaram RSLs, propor e avaliar um guideline para realizar a manutenção de revisões sistemáticas de literatura no campo da engenharia de software. Neste trabalho fazemos um paralelo com os conceitos de manutenção de software, adaptando-os aos conceitos de RSL. Entendendo que tipo de alterações pode levar a uma manutenção de RSL, como avaliar o impacto das alterações e como o processo deve ser conduzido. Três etapas foram realizadas: Investigação, Proposição e Avaliação. Na etapa de investigação, dois estudos foram realizados em paralelo para servir de base para o nosso trabalho. Primeiro, conduzimos um mapeamento sistemático (MS) para abordar o estado da arte sobre atualizações de revisões sistemáticas de literatura, e então realizamos um survey com o objetivo de identificar a opinião dos pesquisadores sobre a manutenção de RSLs. Na etapa de proposição, definimos um guideline para a realização de manutenção em revisões sistemáticas da literatura. Na etapa de avaliação, para validar o guideline proposto, realizamos uma série de entrevistas semiestruturadas. Durante o MS, foi possível verificar atualizações de estudos secundários desde 2010, no entanto, se compararmos com a quantidade de RSLs já publicados, ainda há um longo caminho a percorrer. Quanto à definição de conceitos sobre atualizações de RSL, verificou-se uma falta de compreensão sobre este tema. Os pesquisadores entrevistados mostraram interesse em manter suas RSLs atualizadas, mas expressaram preocupação sobre o esforço a ser feito para realizá-lo. Durante a entrevista o guideline foi apresentado e, como resultado, a aceitação do processo é observada como base para a condução das atividades de manutenção de RSLs. Várias contribuições foram levadas em consideração e incorporadas ao processo. O guideline proposto foi bem aceito pelos pesquisadores, e além da avaliação de sua efetividade como guia para o processo de manutenção de RSLs, foi necessário elaborar diversas discussões, envolvendo o tema, para uma melhor compreensão sobre o tópico. Conceitos importantes foram apresentados e validados durante a pesquisa. No entanto, este é um primeiro vislumbre do processo de manutenção de RSLs e, como qualquer formalização de processo, ainda carece de aprofundamento. Algumas lacunas na conformidade ou formalização ainda podem ter sido deixadas em branco. No entanto, espera-se que com futuras colaborações da comunidade, esse guideline possa se tornar uma base para a realização de atividades de manutenção em RSLs.

Palavras-chaves: Guideline. Atualizações. Manutenção. Rastreabilidade. Revisão Sistemática de Literatura. Engenharia de Software Baseada em Evidência.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

CMMi	Capability Maturity Model integration
COPE	Committee on Publication Ethics
COTS	Commercial off-the-shelf
CVS	Concurrent Version System
EBSE	Evidence-Based Software Engineering
ESE	Empirical Software Engineering
GQM	Goal Question Metric
RQ	Research Question
SE	Software Engineering
SLR	Systematic Literature Review
SM	Systematic Mapping
SVN	Apache Subversion
TS	Tertiary Studies
VTM	Visual Text Mining
XP	eXtreme Programming

CONTENTS

1 INTRODUCTION

This chapter presents a thesis overview. Section 1.1 provides information about the problem being investigated and research conduction motivation. In sequence, Section 1.2 describes study goals and research questions. Then, research method is presented in Section 1.3. Section 1.4 is a summary of thesis' contributions. Finally, thesis structure is described in Section 1.5.

1.1 PROBLEM AND MOTIVATION

Systematic Literature Review (SLR) is a process in which a representative set of relevant available research about a research question, or topic, or phenomenon of interest is identified, evaluated, and interpreted through their individual studies. Systematic Literature Reviews, including Systematic Mappings (SM) and Tertiary Studies (TS) are becoming an important tool in Software Engineering (SE) since the guidelines were proposed by B. Kitchenham and S. Charters in 2004 and updated in 2007 in (KITCHENHAM; CHARTERS, 2007). Kitchenham's guidelines summarize systematic review process into three main phases: Planning, Conducting, and Reporting results. All phases are properly documented, enabling traceability and repeatability. We argue SLR, SM, and TS processes should not end after reporting phase.

During planning phase a protocol is generated to guide the SLR. This protocol defines a research question, search strategies to find relevant studies, inclusion and exclusion criteria, quality criteria, and information to be obtained from primary studies. During SLR conduction all information included in protocol is used to evaluate primary studies retrieved from search procedure. Primary studies selection, primary studies quality assessment, and data extraction are performed in this phase. The final SLR stage is reporting and at this point results are disseminated. The dissemination form varies according to SLR audience's target. According to (BRERETON et al., 2007), review authors need to keep a record of decisions taken while conducting SLRs, which is corroborated by (STAPLES; NIAZI, 2007). (BRERETON et al., 2007) also puts that SE community needs to establish mechanisms to become available the entire review process for their SLR's publication, not just a report.

According to Cochrane handbook (HIGGINS; GREEN(EDITORS), 2011): "Systematic reviews that are not maintained may become out of date or misleading". The main purpose of maintaining SLRs is to keep its evidence as up-to-date as possible. Research is constantly evolving, producing new evidence, which may corroborate with obtained results, or affect performed syntheses. To remain useful, it is inevitable that changes might be made in already reported SLRs. The number of systematic reviews published has been

increasing in recent years (BORGES et al., 2015). However, some of them are very similar to already existing ones. In health care, as early as 2008, it was estimated that 2500 SLRs were published per year (MOHER et al., 2008). Thus, Cochrane collaboration has a two years update policy, which may not happen, by justification (HIGGINS; GREEN(EDITORS), 2011). (SHOJANIA et al., 2007), when investigating SLRs survival time, identified that the average SLR lifespan is approximately 5.5 years. However, in 23% of cases SLRs showed signs of updating before two years, while in 15% these signs appeared in less than a year, and in 7% showed signs before SLR publication. No study like this has been performed in SE area yet, but due to a study detailed in Section 2.4, with only 22 updates found, become evident SLRs update are not a common practice in our area. Therefore, defining a process for maintaining SLRs is of great importance. An important part in SE is to describe processes that deal with how software should be maintained (SOMMERVILLE, 2011). SLRs must also follow certain processes to be maintained, however, there are still no well defined procedures to do so.

The general objective of this research work is to build guidelines for maintaining SLRs in SE context. The restriction to SE field is due to the conducted studies focus in papers and researchers in this area, however, nothing prevents the proposed guidelines from being used in other research areas. To do so, we have to understand researchers needs to evolve their SLRs and what researchers think about SLRs updates, formalizing what would be a SLR maintenance process, reusing methods from software maintenance. To accomplish this task a SM was performed, using guidelines proposed by (KITCHENHAM; CHARTERS, 2007), to understand how updates are being conducted and what kind of changes are being made. Then, we also produced a survey to investigate what Evidence-Based Software Engineering (EBSE) researchers think about updates in SLRs, what kinds of changes in SLR artifacts would generate a new SLR instead of an update, and who is conducting those updates. After producing the guidelines, interviews with researchers experienced in conducting SLRs were conducted to evaluate the proposed guidelines. During the research, questions were raised about possible plagiarism problems due to the characteristic of reuse present in maintenance processes. Therefore, the survey and interviews conducted investigated the plagiarism in SLRs experts' understanding of and the possible impact on the SLR maintenance process.

1.2 STUDY GOALS AND RESEARCH QUESTIONS

The main goal of this research is, based on researchers opinion who have already performed SLRs, to propose and evaluate guidelines to carry out SLRs maintenance in software engineering field. To achieve the main goal, the following objectives were defined:

- To define concepts about systematic literature reviews maintenance in software engineering;

- To guide researchers in systematic review maintenance process conduction;
- To avoid researchers always need to conduct systematic reviews from scratch;
- To develop guidelines conducting systematic review maintenance process, complementing SLRs conduction process proposed by (KITCHENHAM; CHARTERS, 2007), from existing software maintenance methods;
- To identify possible plagiarism problems with SLRs maintenance.

Based on research objectives, the following research questions were elaborated:

- **RQ1.:** What is the state-of-the-art on systematic literature reviews updates performed in SE?
- **RQ2.:** Is it possible to use software maintenance techniques to keep SLRs always up to date?

  RQ2.1.: Is it possible to reuse SLR maintenance methods from other areas?
- **RQ3.:** What do experienced researchers think about SLR maintenance?
- **RQ4.:** How should the SLR maintenance process be performed?
- **RQ5.:** Can SLR maintenance lead to plagiarism issues?

  RQ5.1.: How can plagiarism issues be avoided during SLRs maintenance?

1.3 RESEARCH METHOD

To answer the proposed research questions and achieve the study goals, three steps were carried out: Investigation, Proposition, and Evaluation. Research method overview is presented in Figure 1.

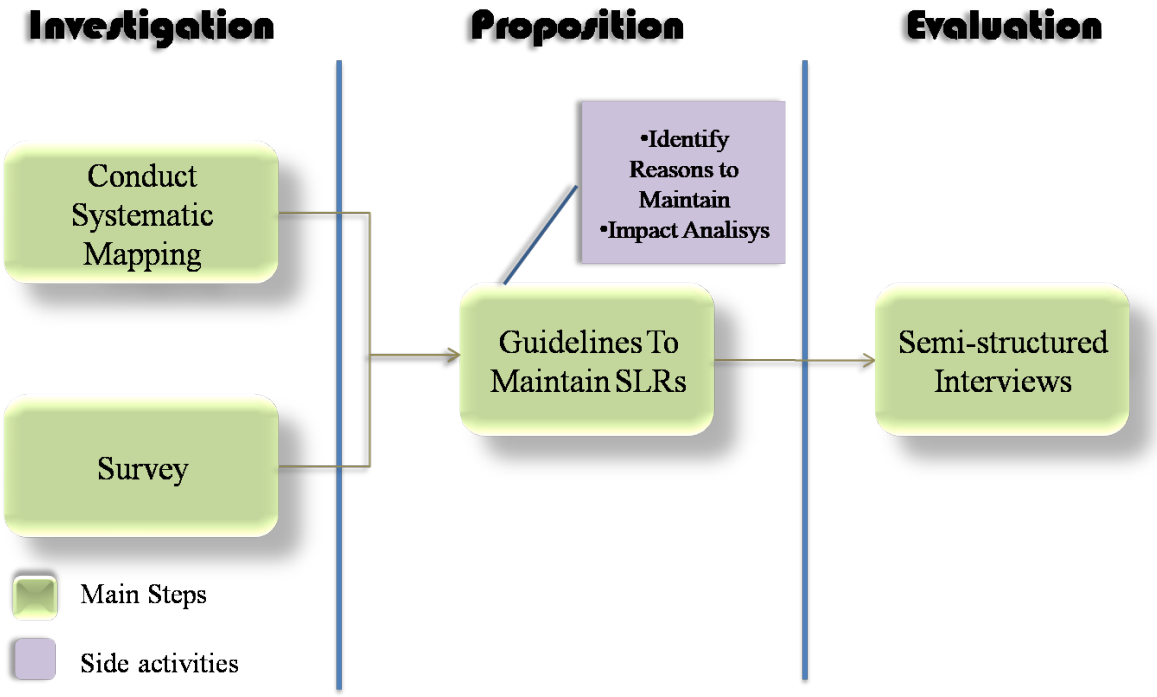
1.3.1 Investigation

Two studies were conducted in parallel to serve as a basis to define our guidelines.

Study 1: Systematic Mapping

A systematic mapping is a method that proposes to structure a research area (PETERSEN; VAKKALANKA; KUZNIARZ, 2015). Planning and conduction phases are very similar to systematic review method, however, synthesis and the way results are shown are different. Therefore, to answer the research question **RQ1**, whose goal is to show SLR updates state-of-the-art, this method was used.

Figure 1 – Research method summary.



Source: The author (2019).

For more details see Section 2.4.

Study 2: Survey

Survey is a methodology to obtain information from a target population and to compare or explain knowledge, attitudes and behavior from this population (PFLEEGER; KITCHENHAM, 2001-2003). This method was used to answer the research question **RQ2**: to obtain and analyze researchers opinion on systematic review updates and on whether it is possible to use software maintenance concepts to maintain SLRs.

For more details see Chapter 3.

1.3.2 Proposition

From the systematic mapping and survey, we collected base elements for the proposed guidelines construction. In this study stage, we aim to define guidelines for conducting maintenance in systematic literature reviews. It contains important terms definition for the process, why a maintenance activity should be performed and a form of analysis to decide whether a new revision should be created, or whether the original work should only be updated.

For more details see Chapter 4.

1.3.3 Evaluation

In order to evaluate the proposed guidelines, and answer **RQ3**, **RQ4** and **RQ5**, we conducted a series of semi-structured interviews. Semi-structured interviews are often used to collect data in qualitative research (MOLLéRI; PETERSEN; MENDES, 2019). Considering that the purpose of this study is to evaluate the proposed guidelines, through experienced researchers opinions in conducting SLRs, semi-structured interviews are presented as an adequate technique for this task.

For more details see Section 4.3.

1.4 CONTRIBUTIONS

This section summarizes this study contributions. As previously shown, the main goal of this study is to provide guidelines for SLRs maintenance conduction in SE. Therefore, it is expected that these guidelines should be broadly used by EBSE community to manage SLRs maintenance activities and be a way of spreading the idea that SLRs need to stay up-to-date to remain useful. The main research contributions can be summarized as:

- An overview of the current SLR updates state in SE;
- An analysis of what researchers think about updating SLRs.
- Guidelines to support the SLR maintenance process;
- Guidelines evaluation by experienced researchers;
- An analysis of possible plagiarism problems in SLRs maintenance.

So far, the results were published at ESEM 2018 (NEPOMUCENO; SOARES, 2018) and IST (NEPOMUCENO; SOARES, 2019), both related to Section 2.4 and Chapter 3.

1. **Maintaining systematic literature reviews: Benefits and drawbacks**. In proceedings of the 12th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM '18).
2. **On the need to update systematic literature reviews**. Information and Software Technology, v. 109, p. 40 – 42, 2019.

Both studies already have citations. (1) was cited by (DÍAZ; MEDINA; ANFURRUTIA, 2019) and (2) was cited by (MENDES et al., 2019).

1.5 THESIS STRUCTURE

We organize this document as follows:

- Chapter 2 presents concepts and studies that supported us to perform this research and also presents the SM results in Section 2.4.
- In order to obtain information about what researchers who have already developed SLRs think about updating SLRs, Chapter 3 presents the conducted survey.
- Chapter 4 presents guidelines to conduct a SLR maintenance process and also presents the interview results in Section 4.3.
- Discussions on SLR maintenance process key points are provided at Chapter 5.
- Finally, Chapter 6 presents the conclusions and future works.

2 BACKGROUND

This chapter presents concepts and studies that supported this research. These studies show background knowledge about Empirical Software Engineering (ESE), with special attention to Systematic Literature Reviews (SLR), including Systematic Mappings SM and Tertiary Studies TS, which are types of systematic reviews, and therefore, will be called Systematic Literature Reviews (SLR) from now on.

Section 2.1 describes ESE emergence and highlights secondary studies importance to aggregate evidence in Software Engineering (SE). Section 2.2 shows what is a SLR, how it is conducted, and what happen after conclusion. Section 2.3 presents studies that are related to SLRs maintenance process, including a perspective about the subject in other areas. Section 2.4 presents a SM about SLRs updates in SE and Section 2.5 shows a chapter summary.

2.1 EMPIRICAL SOFTWARE ENGINEERING

SE field is responsible for developing, maintain, and manage software systems. According to (SJOBERG; DYBA; JORGENSEN, 2007), research in ESE studies the real-world phenomena in SE, new technologies, process models, tools, etc., as well as evaluation and comparison between these new approaches and their use in organizations or academy by software developers, or regular individuals.

Empirical studies are responsible for the real-world phenomena studies in several research fields, with the objective of obtaining information that will provide a support for new researches, proposition of new and evolution of existing theories. To improve its scientific maturity, SE needs to use empirical studies.

ESE emerges from this need to elucidate phenomena being studied in SE. Although, even using ESE, we cannot guarantee the knowledge produced by an empirical research is certain (EASTERBROOK et al., 2008). According to (SJOBERG; DYBA; JORGENSEN, 2007) we can divide most relevant research methods in ESE in primary research and secondary research:

- **Primary Research**:

  Experimentation: an empirical method to investigate causal relations in a phenomenon.

  Surveys: an empirical method to investigate relationships and outcomes from phenomena in a retrospective way.

  Case Studies: an empirical method to investigate phenomena in real-life context.

Action Research: an empirical method to acquire theoretical knowledge, and, at the same time, provide practical value to organizations where a research is being conducted.

- **Secondary Research**:

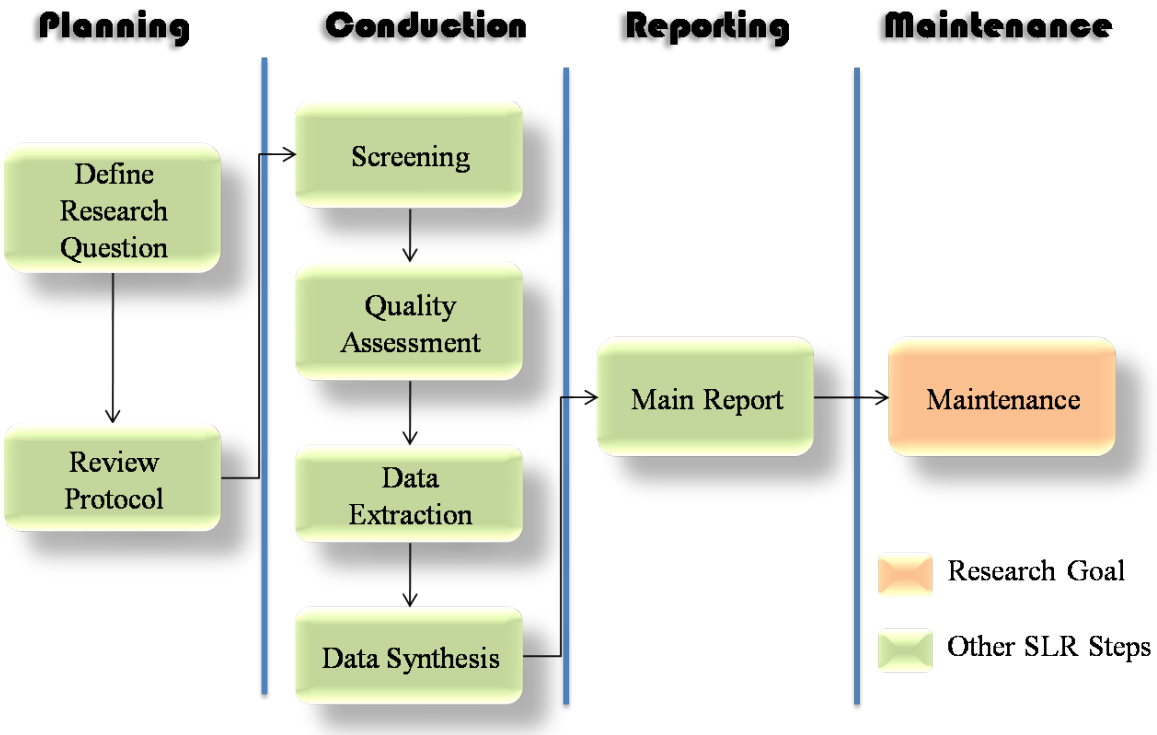
  Systematic Literature Reviews (SLR): an empirical method to acquire knowledge from synthesized data retrieved from primary research.

This research focuses on how investigators conduct a SLR and especially how SLRs should be maintained by researchers. We use primary studies when we conducted a survey (Chapter 3) and interviews (Section 4.3), and a secondary study when we performed a SM (Section 2.4).

2.2 SYSTEMATIC LITERATURE REVIEW

Figure 2 summarizes the SLR process that was proposed by (KITCHENHAM; CHARTERS, 2007). This research incorporates a fourth stage: SLRs maintenance.

Figure 2 – Systematic Literature Review Process.



Source: The author (2019).

2.2.1 Planning

In this phase, the need for a SLR is evaluated and a research question to be answered is defined and drives all the process. After that, a review protocol is developed. The

protocol defines steps that will be used (research question, study selection criteria, quality assessment procedures, etc.). A protocol is also important to guarantee SLR replicability by other researches and it is a critical element in SLR, therefore, its evaluation is usually performed before conduction phase begins.

2.2.2 Conducting

In this phase, all relevant studies to answer a research question are retrieved from literature using a search strategy, defined by the protocol. Second step is to look for evidence in these studies for suit selection criteria, including or excluding them from SLR. After that, a quality assessment is performed in remaining studies. Three concepts involved in quality assessment are: bias, internal validity, and external validity. Usually, a study with high quality intends to minimize bias and maximize internal and external validity. In fourth step, a data extraction must be realized to collect and record all information retrieved from primary studies. After that, the collected information is synthesized to fill the main report and answer a research question.

2.2.3 Reporting

In this phase, all synthesized data must be disseminated to all interested in research. A strategy to disseminate must be defined depending on your target audience. Disclosure can be made through technical reports, articles in conferences, or journals, as well as in course work completion (NAKAGAWA et al., 2017). However, if the dissemination is made only through use of these media, probably SLR results will be consumed only by academy. (CARTAXO; PINTO; SOARES, 2018) say practitioners are often neglected, which makes using current forms of SLR disclosure inaccessible. To better transfer knowledge, (CARTAXO; PINTO; SOARES, 2018) propose use of evidence briefings.

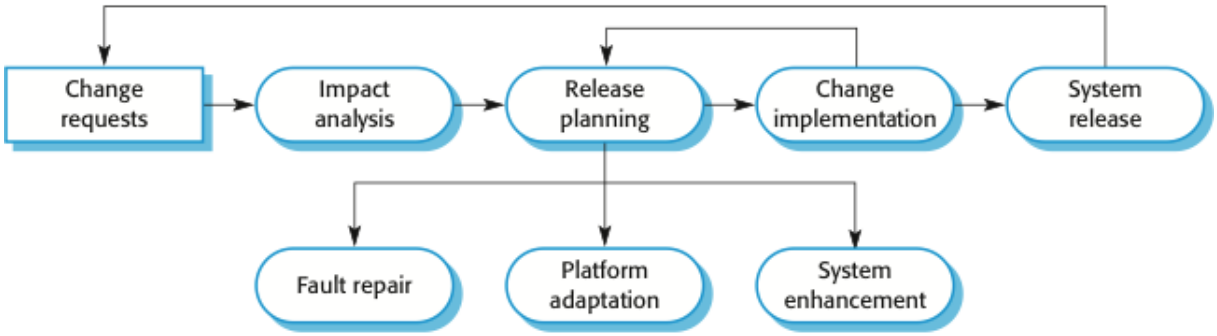
2.2.4 Maintenance

There are many papers and books that define what software maintenance is and how maintenance process can be conducted ((KITCHENHAM et al., 1999), (BENNETT; RAJLICH, 2000), (CHAPIN et al., 2001), (APRIL et al., 2005), (SOMMERVILLE, 2011), (SOCIETY; BOURQUE; FAIRLEY, 2014)). In (KITCHENHAM et al., 1999) an ontology was proposed to describe software maintenance process. Such need arose during a discussion in which a difficulty of talking about maintenance was perceived since there was no consensus on term. (BENNETT; RAJLICH, 2000) stresses the importance of improving maintenance process speed and accuracy, thus reducing costs, for which it has sought to identify key problems and to bring solutions to these problems. In (CHAPIN et al., 2001) is proposed a set of software maintenance and evolution types, it deepens work of (KITCHENHAM et al., 1999), as well as brings new forms of process definitions. In (APRIL et al., 2005) a Software

Maintenance Maturity Model is proposed. A model was created to complement Capability Maturity Model integration (CMMi). (SOMMERVILLE, 2011) defines software maintenance as the general process of changing a system after it has been delivered, however, (SOCIETY; BOURQUE; FAIRLEY, 2014) defines as software modification while preserving its integrity. These papers present their own views of different forms on same topic. In current work we bring several presented definitions by these works for SLRs' reality.

However, there is still no similar definition for SLRs maintenance. Building software process does not end after delivery. A set of actions should be taken to ensure software continues to evolve and remains useful (Fig. 3).

Figure 3 – Software evolution process. Source:



Source: (SOMMERVILLE, 2011) pag. 238.

In software maintenance we have basically three reasons for changes:

- Fault Repairs — problems correction encountered after delivery;
- Environmental adaptation — software adaptation to a new operating environment;
- Functionality addition — software updates due to new requirements needs.

However, these reasons do not fit well in SLRs and we need to adapt them. This work makes a parallel with software maintenance concepts presented, adapting them to SLR concepts. This is done through understanding what kind of changes can lead to a SLR maintenance, how to assess change impact, and how the process should be conducted.

2.3 RELATED WORK

Maintenance and versioning are constant concerns in SE, however, there were not found any work directly related to SLR maintenance and versioning in SE area, although some studies report SLRs updating importance and present some techniques used to perform updates. However, this topic is already addressed in other areas such as medicine, education, and psychology. We analyze how SLR maintenance process is performed in these areas and how we can improve our proposal with the presented definitions in these areas.

(MACDONELL et al., 2010) concluded systematic review is a robust tool, which even when applied by different teams must achieve same evidences. However, the way to get results can be, and probably will be, different. In this scenario, if any researcher wants to update a systematic review, whether or not this researcher took part on the original systematic review, they may encounter difficulties if they do not have the history of how that SLR was conducted.

According to (KITCHENHAM et al., 2011), inexperienced researchers may have even more difficulties to repeat a SLR result. (KITCHENHAM et al., 2011) exemplify that by showing two groups of inexperienced researchers selected different groups of primary studies, which had a direct impact on evidences found by SLR. They came to the conclusion that it is essential to fully describe the search process for a SLR to become truly repeatable. In this sense, in order to maintain a systematic review, search process description and also primary studies selection become essential, since the lack of information may lead to a lack of understanding on maintainers part, and, thus, bringing some new bias to search.

During our research, few works were found about how to update SLRs in SE. An update, in these cases, was considered if the studies used the original protocol, or at least what is available in the original protocol, but in a different time frame, or using different search techniques and/or different search sources. Thereafter, comparisons are made between new evidences found and the original study. Original study by (KITCHENHAM et al., 2009) was extended by (KITCHENHAM et al., 2010) and (SILVA et al., 2011). These three studies address SLRs use in SE. The original work (KITCHENHAM et al., 2009) performed search between January 1st 2004 and June 30th 2007, using a manual search in a restricted set of sources. In (KITCHENHAM et al., 2010), an automatic search was applied in some search engines and indexed systems, and in addition, time window was modified to January 1st 2004 and June 30th 2008. An extension protocol (KITCHENHAM; BRERETON; BUDGEN, 2008) was created to explain how the update would be performed. This same protocol was used by (SILVA et al., 2011). Third extension (SILVA et al., 2011) carried out, search between July 1st 2008 and December 31st 2009, performing manual and automatic searches.

In 2008 (DIESTE; LóPEZ; RAMOS, 2008) sought to formalize a way to update systematic reviews. While conducting his original work in 2005, where he did a systematic review with 26 primary studies, it was not possible to obtain 27 studies that would be interesting to work with. After publication, another 13 of the 27 studies were accessed and an update was conducted in 2008. This process resulted in an improvement in systematic review process taking updates into account. However, improvements placed by author are specific to reviews conducted to aggregate information from experiments, as was already done in other areas, which we understand to be a small set of what is done today with SLRs.

(FERRARI; MALDONADO, 2008) carried out a study in which he updated a systematic

review by several iterations looking to discuss SLRs replicability problems. For this, an adaptation was made in original SLR conduction process (BIOLCHINI et al., 2005). Three new steps were introduced: (1) Planning Update, (2) Filtering Results and (3) Merging Results. At (1), protocol is revisited for possible changes. In (2), search results are filtered to look for overlaps. In (3), new data is included within original dataset. Although it is an interesting approach and with some common points with our proposal, we understand this process takes into consideration only updating studies corpus for analysis. In present work, we present other reasons to update an SLR, and how maintenance process definition can make the method much more robust.

(FELIZARDO et al., 2014) proposes a VTM technique to support SLR updates, where selection of new primary studies is performed through a tool (Revis), which implements the proposed technique. (FELIZARDO et al., 2016), (FELIZARDO et al., 2018), and (WOHLIN, 2016) investigated snowballing use (JALALI; WOHLIN, 2012) to perform SLR updates. Both works are similar and come to same conclusion that snowballing is a technique to perform SLR updates. However, once again, such update is resulted by a new search for primary studies.

(RODRIGUEZ et al., 2017) is a report on two SLRs update. Study raises four important points to consider when conducting an update:

1. Tools usage to support updates;
2. Always include information about previous study;
3. Have a team member on update team that participated in previous study;
4. Reuse the protocol from previous study.

In our study, we asked researchers about points raised by Felizardo et al. leading us to similar conclusions. We understand updating SLRs is an important part of SLRs maintenance process. However, other points must be analyzed, such as:

- Information availability about original SLR;
- The actual need to update a particular SLR versus the effort needed;
- How to maintain SLR originally performed by other researchers;
- How to provide collaborative maintenance of SLRs;
- How to manage multiple SLR updates.

These studies show an interest from authors in keeping SLRs in SE updated, however, evidences are found in a dispersed way nowadays. Several different publications, published in different places, cause researchers to have difficulties in finding all work related to a

single SLR and then synthesize entire evidence found. Some of these problems have already been solved in other areas.

Systematic reviews are conducted in various knowledge areas beyond software engineering, such as health care and social sciences. In health care, systematic reviews are used to ensure the best scientific evidence for individual patients treatment, or for public policy (TACCONELLI, 2010). In social sciences, SLRs are being used for decision-making by politicians and policymakers. Evidence-based knowledge use serves to justify government social programs (PETTICREW; ROBERTS, 2006).

Systematic review is a long-standing tool, and its first paper was published in 1891 in American Journal of Psychology authored by Herbert Nichols. However its proliferation began by 1980s in health care, as early as 2007, approximately 2500 systematic reviews were produced annually (MOHER et al., 2007). In health care, SLRs updating is a consolidated process, including evaluation methods for when a SLR should be updated (SHEKELLE et al., 2011) (MOHER et al., 2008). However, in SE we do not find such processes, making necessary to evaluate if methods used in health area can be migrated, with adaptations, to SE.

(MOHER et al., 2008) gives the following definition of an SLR update:

> *"A distinguishing feature of an updated systematic review from a new review is that during updating the originally formulated protocol (e.g., eligibility criteria, search strategy) is retained, and sometimes extended, to accommodate newly identified information (e.g., new treatment type, diagnostic method, outcome, different population)"*

An example of how health care SLRs are maintained is Cochrane Collaboration, an international organization that was created in 1993. It has over 15,000 employees spread across more than 100 countries. Cochrane Handbook (HIGGINS; GREEN(EDITORS), 2011) has a chapter dedicated to SLR maintenance process, where several routines are defined to support procedures of updates, amendments, and feedback. Versioning issue is also presented in Cochrane Handbook, where each published review has a citation version. Future changes in reviews, or SLR protocol modifications, can generate new citation versions. However, there is still no such initiative in SE area. We can find other initiatives, including in other areas:

- Campbell Collaboration[1];
- Centre for Reviews and Dissemination[2];

[1] https://www.campbellcollaboration.org/
[2] https://www.york.ac.uk/crd/
[3] https://www.crd.york.ac.uk/PROSPERO/
[4] https://eppi.ioe.ac.uk/cms/

- PROSPERO International prospective register of systematic reviews[3];
- EPPI-Centre[4].

The Cochrane Opportunities Fund developed a tool to assist in assessing the need to update systematic reviews of Cochrane Database at the end of 2008 (HOPEWELL et al., 2008). Tool is composed of two steps: a decision tree presents a set of triggers that indicates the need to update a review, and then a checklist indicates which revision parts need to be updated. Triggers used in the first step are:

- New information — Any new information that may affect review results. E.g.: new treatments, new studies that were in progress while conducting review, etc.
- New methodology — Changes in Cochrane Handbook, or a new statistical method.
- Response to feedback from review users.
- Other factors — Age limit or imminent review use in any decision-making process.

The presented tool (HOPEWELL et al., 2008) cannot be used, directly, in SE, since they are intrinsically linked to health care. However, one can carry out a process adaptation to extend its use to SE, especially in relation to described triggers.

2.4 SLR UPDATES' SYSTEMATIC MAPPING

A Systematic Mapping SM is a method to categorize results giving a visual review about fields with a lack of information (PETERSEN; VAKKALANKA; KUZNIARZ, 2015). Due to the few numbers of SLR updates found in literature, we felt the need to perform a SM on this topic.

Research question addressing the goal of this SM is:

RQ.: What is the state-of-the-art on systematic literature reviews updates performed in SE?

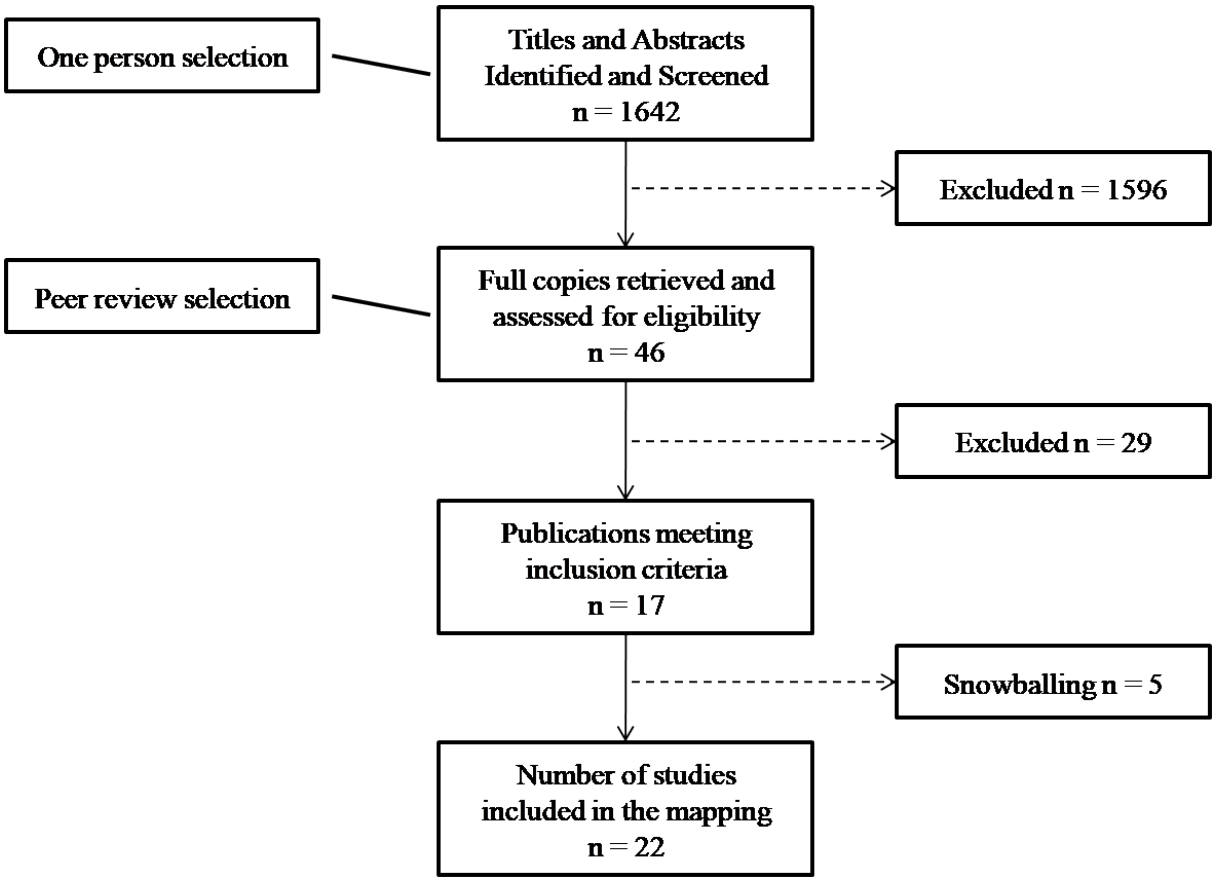
To perform SM, we used the tool StArt (State of the Art through Systematic Review) (FABBRI et al., 2016). StArt is a mature tool that has been available since 2010 and is constantly evolving. StArt has support for all steps of a SM.

Search procedure was performed using automatic search and snowballing (JALALI; WOHLIN, 2012). Search string used in selected sources is:

(update OR updated OR extended OR extension OR extend OR expanded OR expand) AND ("systematic review" OR "systematic literature review" OR "systematic mapping") AND ("software engineering")

Adaptations were made for each source (ACM, IEEE, Scopus, ScienceDirect and Springer). IEEE and ACM bibliographic sources use combined with two search engines is considered sufficient by (DYBA; DINGSOYR; HANSSEN, 2007). Search procedure returned 1642 studies. Selection process was carried out in three stages. In first stage only SM conductor applied selection criteria. In second step, complete texts were analyzed in pairs, one being the conductor, in case there was disagreement between conductor and the other researcher a meeting between was held to reach a consensus. Third step was applying the snowballing procedure (backward and forward) in the already selected studies by SM conductor. Fig. 4 summarizes screening process. Selected papers list is on Appendix A.

Figure 4 – Screening Process Summary.



Source: The author (2019).

Following the inclusion (I) and exclusion (E) criteria used for selection process:

- (I) Studies that are updates from systematic literature reviews in SE;
- (I) Studies that are updates from systematic mappings in SE;
- (I) Studies that are updates from tertiary studies in SE;
- (E) Short papers - less than 4 pages;
- (E) Written in any language but English;

- (E) Not accessible on Web;
- (E) Invited papers, keynote speeches, workshop reports, books, theses, dissertations, incomplete documents, drafts, slides of presentations, and extended abstracts;
- (E) Addressing update SLR or SM only as part of future work;
- (E) Not related with research topic.

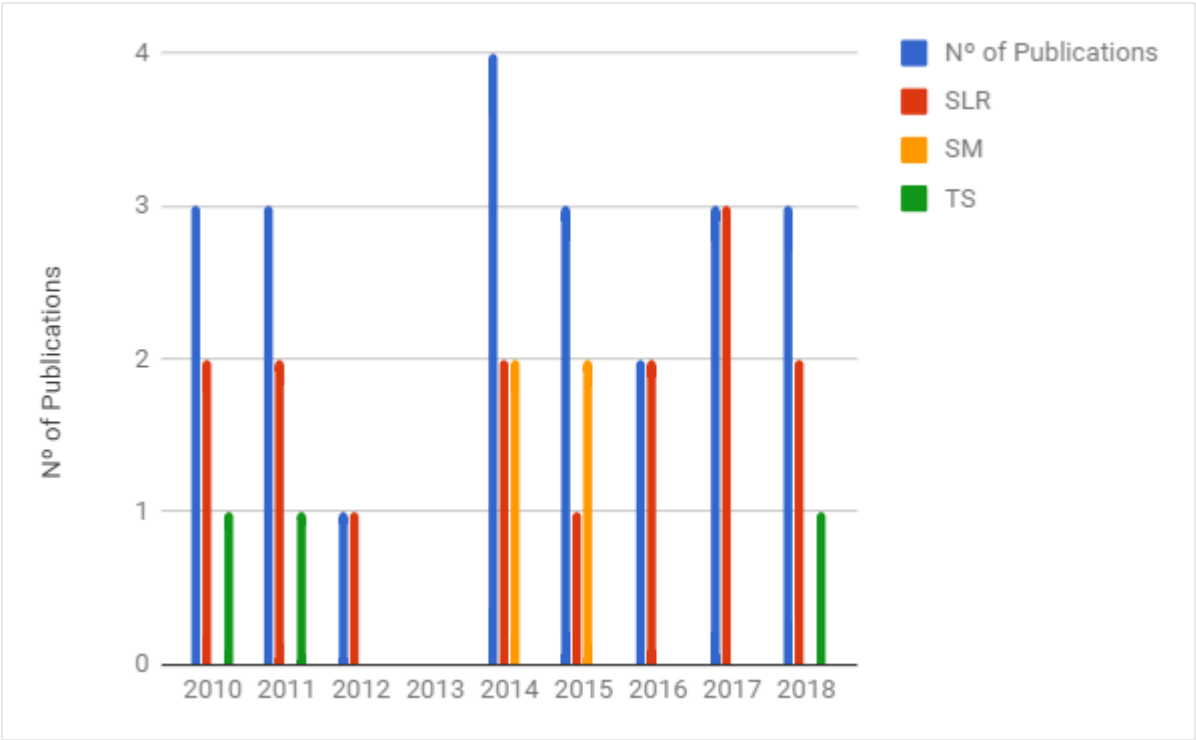
Google Spreadsheet was used to manage all data extraction and analysis procedure. One researcher was allocated to extract contents from papers. Data extraction form can be found on appendix B. Some metadata were extracted to identify studies, as well as important data to answer research question.

2.4.1 Synthesis

After extracting data, they were categorized and synthesized. General considerations were made about metadata and a coherent understanding on research question is provided.

Fig.5 presents publications by year and what kind of secondary review was used. It is possible to check secondary studies updates since 2010, however if we compare with the amount of SLR studies already published, there is still a gap between SLR updates number and SLR number.. We have not yet been able to see a trend in updates. By defining processes and tools that aid updates, this trend should be upward.

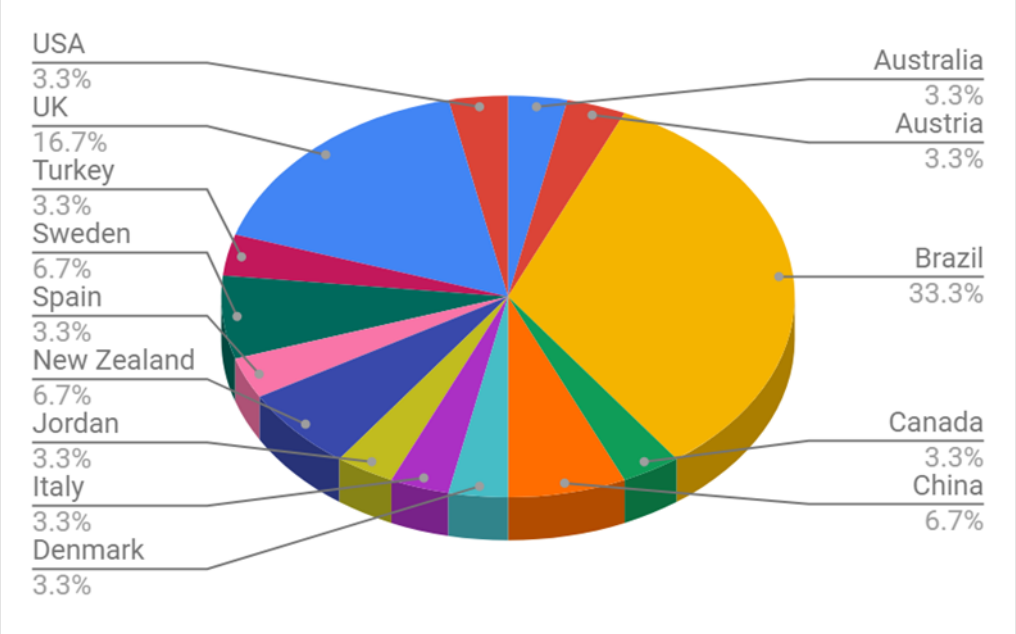
Figure 5 – Publications by year.



Source: The author (2019).

As demographic information, Fig.6 shows contributions by country. It is possible to see Brazil and United Kingdom are the main contributors.

Figure 6 – Contributions by country.



Source: The author (2019).

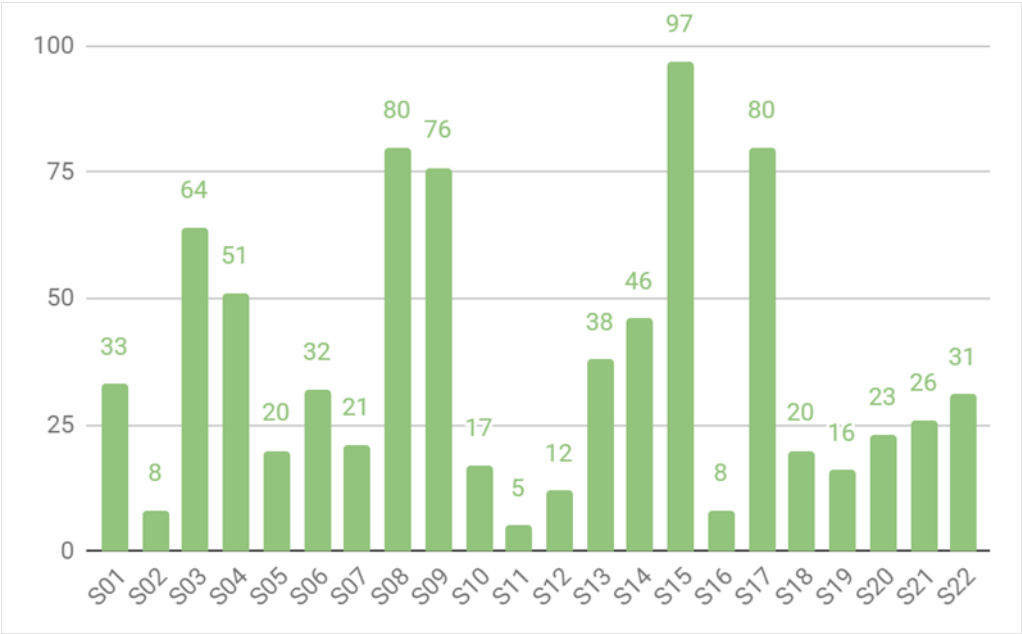
Fig. 7 exposes the lack of criterion, related to time, when performing an SLR update in SE. As put out by Cochrane Handbook (HIGGINS; GREEN(EDITORS), 2011), an interesting interval for updating an SLR is a maximum of two years for health care area. It is possible to verify situations in which the elapsed time is only 5 months, as well as situations in which the time is 97 months. The lack of a defined criteria in literature on when and why update SLRs in SE can cause this misunderstanding.

We investigated what reasons were put in place to carry out the update, we present below a summary of these reasons:

- Provide an updated field overview and document evolution;
- Increase our understanding of current state of practice;
- Describe a supplementary analysis of these studies;
- Extract more data and evaluate, analyze, and summarize unresolved problems and shortcomings about the topic;
- Report the experience in updating SLRs;
- From a previous SM result perform a deeper analysis through an SLR.

Table 1 lists primary studies and respective artifact modifications made by update. We are considering SLR artifacts as any documentation used to elaborate and conduct

Figure 7 – Elapsed time (months) per paper.



Source: The author (2019).

SLR, any extraction form used to store data from primary studies (quality assessment and data extraction steps), and SLR's reports. Changes were verified at critical points in an SLR, such as Research Question and search string, which are directly linked. In our research we discussed whether these modifications should be considered an update, or whether they should be considered a new SLR based on previous one. Table 2 and 3 lists some modification examples extracted from papers.

Table 1: Primary studies with modifications.

Modification	**Papers List**
Not Modified	S05, S07, S08, S11, S12 S13, S22
Modified Protocol	S01, S02, S03, S04, S06, S09, S10, S14, S15, S16, S17, S18, S19, S20, S21
Modified Research Question	S01, S10, S14, S17, S18, S19
Modified Search String	S02, S04, S06, S09, S16, S18, S20
Modified Search Strategy	S17, S19, S20, S21
Modified Inclusion/Exclusion Criteria	S03
Modified Sources	S15, S16, S18, S20, S21
Modified Quality Assessment	S06, S10, S19
Modified Data Extraction Form	S02, S04

Source: The author (2019).

Table 2: Modification examples.

Paper	Modification	Original Text	Modified Text
S01	Research Question	RQ 1: How is the term 'software ecosystem' defined? RQ 2: What is the research output per year in the SECO field? RQ 3: What is the type of result that software ecosystem research reports? RQ 4: What is the role of architecture in software ecosystem research? RQ 5: How is the connection between research and industry in the area of software ecosystems?	RQ 1: How has the number of publications evolved within the field? RQ 2: What can we extract about the field's evolution and maturity from studying the publication venues for ecosystem research? RQ 3: How have the types of research results evolved? RQ 4: How has research within software ecosystem groups evolved? RQ 5: Is software ecosystem research targeting real software ecosystems?
S09	Search String	(Agile, Scrum, XP, pair programming, Lean Development, Lean Software Development) AND (Global software engineering, global software development, distributed software engineering, distributed software development, GSE, GSD, distributed team, global team, dispersed team, spread team, virtual team, offshore, outsource, open source)	(agile OR scrum OR "extreme programming" OR "pair programming" OR "lean development" OR "lean software development") AND ("global software engineering" OR "global software development" OR "distributed software engineering" OR "distributed software development" OR GSE OR GSD OR "distributed team" OR "global team" OR "dispersed team" OR "spread team" OR "virtual team" OR offshore OR outsource)
S17	Search Strategy	-	It used forward snowballing from original paper.

Source: The author (2019).

Table 4 summarizes information about who conducts updates. From 22 selected studies, just six updates were conducted by a research group different from the one performing the original SLR (1). From these six, only one study made contact with original research group (2) and another reported difficulties to find some artifact from previous work, in this case, quality assessment criteria (3). One concern that can arise when updating a work done by another research group is plagiarism. To what extent we can use information contained in the original work without compromising the update. An interesting practice to reduce likelihood of encountering these problems is to contact original work's authors and always make clear differences between works.

Table 3: Modification examples.

Paper	**Modification**	**Original Text**	**Modified Text**
S03	Inclusion Exclusion Criteria	(I) SLR covered a topic addressed in the IEEE/ACM curriculum guidelines (SE2004); (E) SLRs that addressed research trends; (E) Mapping studies with no analysis of collected data; (E) SLRs on topics that were not deemed to be relevant to teaching (based on the content of four major textbooks).	(I). The paper is published in a journal, and either included in the three broad tertiary studies, or one of the five journals in the appropriate periods. (I). The topic of the paper is appropriate for introductory teaching of SE. (I). The paper contains conclusions or recommendations relevant to teaching and explicitly supported by the outcomes. (E). Systematic reviews addressing research trends. (E). Systematic reviews addressing research methodological issues. (E). Mapping studies with no synthesis of data. (E). Systematic reviews that address topics not considered relevant to introductory teaching of SE.
S15	Sources	-	Add search in SCOPUS and do not search in Individual journals and conference proceedings
S06	Quality Assessment	-	The adaptation made in this process was either to subdivide some scores in other yes/no questions or clarify the response options for score.
S02	Data Extraction Form	Publication Title Authors Year Research Organizations Country Replication Type Report Type Software engineering topic Research Method Unit of Analysis Confirmation of Original	Title Year Publisher Topic Replication definition Research problem Proposal Contribution

Source: The author (2019).

We investigated whether the updates change any result from original study (4). To fairly assign a confirmation, we used conclusions about confirmation as reported by authors whenever possible. Seven studies reported results with some difference from the original. In S20 study it was possible to verify the change in result comes from a more detailed selection process, which even led to papers exclusion that were included in original research. For other updates the differences come from new evidence.

Table 4: Collected information for other questions.

Number	Question	Answer	Papers List
1	Who performed the original study?	Different Group	S06, S08, S09, S14, S16, S17
2	Has there been any contact with the original author?	YES	S16
3	Does the author report difficulties in finding the artifacts of the original study?	YES	S09
4	Did the update change any result from the original study?	YES	S09, S16, S17, S18, S19, S20, S22
5	Did the author make available the update artifacts?	YES	S05, S13, S14

Source: The author (2019).

As we can see in Tab. 4 few papers published their artifacts outside paper report (5). One of major problems that can be encountered when updating a SLR is the lack of information about the original work. Study replicability depends on information provided by researchers and it is not always possible to put all information needed for replication in a paper report.

2.4.2 Threats to Validity

Some threats were identified during SM conduction:

- Lack of term "Secondary Study" in search string, however we did the search again in ACM, IEEE and Scopus sources and differences found were papers published outside our research time frame;
- First selection process was performed only for one person, the SM conductor. To reduce this threat impact, only papers that were not related to SLRs update were removed, any doubts about this relation led us to include the paper in selection second phase;

- Data extraction process was performed only for one person, SM conductor. To reduce the threat impact, all fields were double checked during extraction.

2.5 CHAPTER SUMMARY

In this chapter we have discussed important concepts in empirical software engineering field, especially about systematic literature reviews. Relevant studies about SLR maintenance in software engineering, as well as how other fields such as health care deal with SLRs were discussed. We also discussed the importance of keep SLRs up-to-date, so that SLR remain valid, a SM presented an state-of-the-art on systematic literature reviews updates performed in SE. With this knowledge, the reader will be able to understand context in which this research is inserted. In the next chapter a Survey is presented in order to obtain information about what researchers who have already developed SLRs think about maintain SLRs.

3 SURVEY

This chapter presents a survey conduction, in order to obtain information about what researchers who have already developed SLRs think about updating SLRs.

Section 3.1 provides information on Survey use as a methodology for obtaining desired information. Section 3.2 presents how survey was conducted. Obtained results are showed in Section 3.3. Section 3.5 discussed threats to validity in this survey and Section 3.6 brings a chapter summary.

3.1 METHODOLOGY

Survey is a methodology to obtain information from a target population and to compare or explain the knowledge, attitudes, and behavior from this population (PFLEEGER; KITCHENHAM, 2001-2003). In this research, we followed principles on how to conduct a survey proposed by (PFLEEGER; KITCHENHAM, 2001-2003).

A set of activities should be performed to ensure greater reliability in survey results, these activities are:

- Setting specific measurable objectives (The survey's expected results);
- Planning and scheduling the survey;
- Ensuring appropriate resources are available;
- Designing the survey;
- Preparing the data collection instrument;
- Validating the instrument;
- Selecting participants;
- Administering and scoring the instrument;
- Analyzing the data;
- Reporting the results.

This research aimed to identify researchers opinion about maintenance of SLRs, therefore it had an exploratory purpose (MOLLéRI; PETERSEN; MENDES, 2019). The survey was carried out to answer the following research question:

RQ.: Is it possible to use software maintenance techniques to keep SLRs always up to date?

Some questions have a descriptive character, that try to identify researchers profile in relation to a possible SLR update realization (MOLLéRI; PETERSEN; MENDES, 2019).

3.2 CONDUCTION

Survey has an observational characteristic, trying to understand retrospectively if participants have already performed an SLR update and how they conducted those updates. According to (PFLEEGER; KITCHENHAM, 2001-2003), present work can be fit as a cross sectional and case control study, due to the fact that participants were asked about their past experiences in conducting SLRs at a particular point in time. The survey was designed to be **unbiased**, trying to understand the population of our study, as well as **appropriate** so this population can answer questions, and **cost-effective** so it can be realized and analyzed within resources destined to accomplish it.

3.2.1 Instrument

Google forms was used for instrument construction. A search for research related surveys was performed, however none were found, probably due to characteristics of our population. Questions can be found at Appendix C.

The instrument was developed to fit two groups of respondents, first group (G1) for those who had already performed an SLR update, second group (G2) for those who had not performed, through the question *"Have you already performed a systematic literature review update?"*.

The remainder of the survey was split into three parts: 1) Maintenance – understanding what our population understands about update SLRs; 2) Traceability (Versioning) – questions were asked about use of versioning tools to maintain SLRs; 3) Plagiarism – investigating whether SLRs maintenance could present plagiarism problems, including self-plagiarism issues.

The survey has 22 questions for G1 and 17 questions for G2. It consists of closed and open questions. Only respondent's email was used for identification, since survey was resubmitted if no response was obtained after the first e-mail. To reduce instrument construction bias by researcher, his advisor validated instrument's clarity and questions coverage related to research question.

3.2.2 Population and Sample

Survey target population is software engineering researchers who have already conducted more than three SLRs, what we consider a more experienced group. Our sample was taken from (ALMEIDA, 2017) in which SLRs were identified between 2010 and 2015. More than 1000 researchers participate in SLRs during these years, but just 121 take part in three or more studies. 42 researchers were selected to participate in interview presented

in Chapter 4. Then, due to the size of our target population, we have sent an email to all 79 remaining researchers and 28 answers were obtained, a response rate of almost 36%.

Participants received an email explaining why they were chosen. To make the survey more personal and seeking an increase in response rate, emails were written individually, taking into account researcher name and one published paper identified in (ALMEIDA, 2017). After two weeks of waiting, a reminder was sent to researchers who had not yet answered the survey.

3.3 RESULTS

We first grouped participants in who already made a SLR update (G1), 12 answers, and who did not (G2), 16 answers. As mentioned before, we divided survey into three parts: Maintenance, Traceability, and Plagiarism. Therefore, results will be presented in that order.

3.3.1 Maintenance

In G1, only one participant related making changes in original SLR artifacts and just two participants updated a SLR of another research group. Both that updated another group's SLR reported problems to access original artifacts, but none contacted the original research group. We asked G2 if they would contact original SLR authors if they did an update of an SLR from another research group, 14 said yes and two said no.

We asked both groups if they would like to keep their published SLR always up-to-date, 19 answered yes (67,8%), and nine answered no (32,2%). From a positive point of view, we highlight following answers (our emphasis):

- *"This is important to keep the data updated and observe the searched* ***area evolution****."*
- *"****New studies are always published*** *after a SLR, so, I would like to have my SLRs updated."*
- *"This is* ***the aim*** *of any SLR protocol!"*

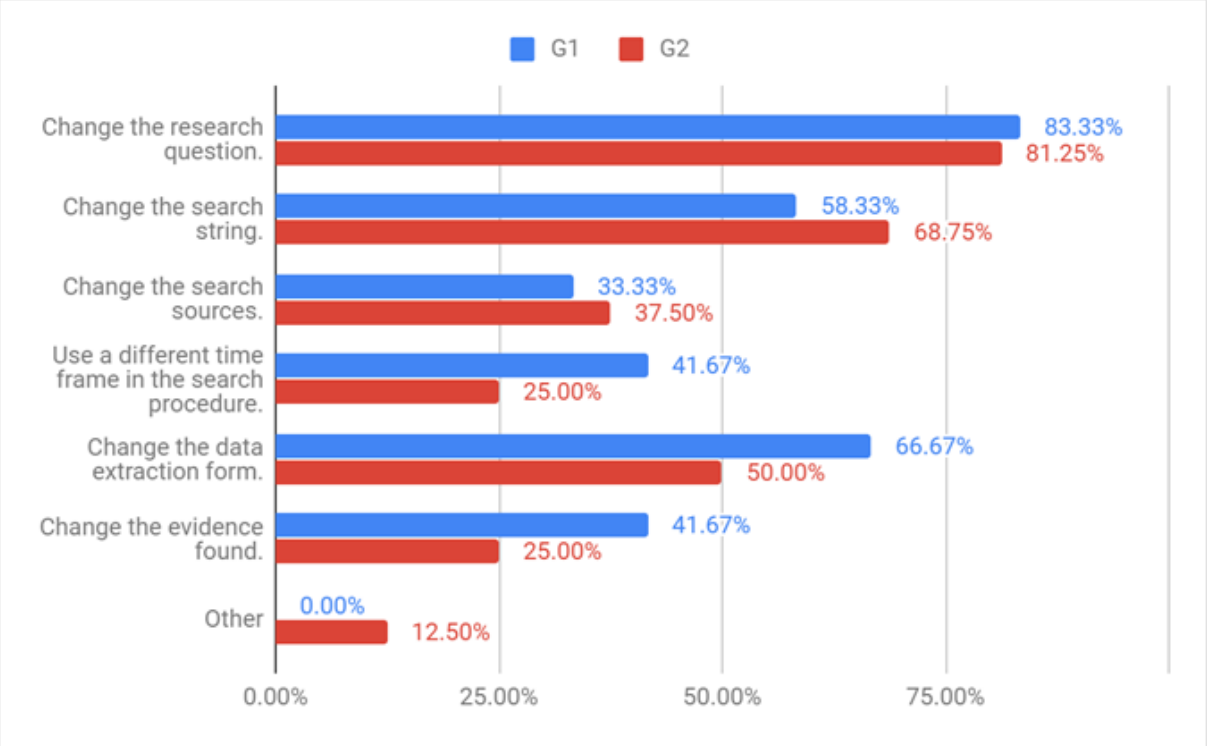
From a negative point of view, we highlight following answers (our emphasis):

- *"Well, of course that would be nice, but given* ***the effort involved*** *(primarily screening and synthesis), I don't really see how that could be feasibly done."*
- *"If* ***there's no additional effort****, of course it would be nice to have an SLR always updated as it is more useful. However, updating an SLR continuously and manually seems to be* ***a highly inefficient task*** *as one needs to keep up-to-date continuously with new research, which is connected to a lot of search effort."*

- *"I think it is not worth the effort of updating a SLR. It is hard* ***to publish it again****. The contributions may not be* ***significant****."*
- *"A SLR is a synthesis at one point in time on a specific topic.* ***Research evolves in various directions*** *and it is preferable to be flexible to tackle new perspectives, rather than stick with old ones as structured years ago.* ***New research questions*** *become more relevant that update to old ones, most of the time."*
- *"We would like to keep it up to date. However, most of the time, the researchers who did the SLR leave or switch to another topic after completing their thesis. This makes it difficult to update unless* ***another researcher*** *would like to continue on the same topic."*

Another question was if they would mind if another research group update their SLR, 27 did not mind, and just one did. Fig. 8 depicts answers about what differences from original SLRs will result a new SLR instead of an update. The question accepts multiple answers.

Figure 8 – Which differences from the original SLR will result a new SLR instead of an update?

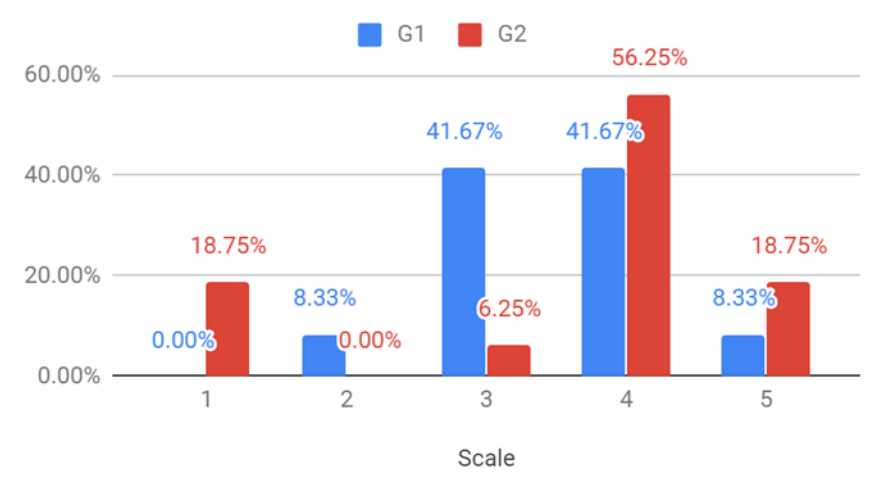


Source: The author (2019).

A possible definition of SLR maintenance was presented to participants. Answer was given using a five-point agreement scale, where 1 indicates *I Disagree* and 5 is *I Agree* (Fig. 9). Then, the participant could give suggestions for changes, or a reason why they disagree with definition. Proposed definition was: *"Modifying an existing SLR by altering*

some artifact, while reusing its original questions, keeping all changes and impacts caused by such changes tracked".

Figure 9 – Results for SLR maintenance definition question.



Source: The author (2019).

From a positive point of view, we highlight following answers (our emphasis):

- *"Make explicit what you mean by **artifact** and which artifacts could be **changed** without resulting in a new SLR."*
- *"I think it is a **good** definition, just need to **detail** more, for example, which **artifacts** will be changed."*
- *"Maybe it is more nuanced - in maintenance, we traditionally have classes of **corrective**, **perfective**, **preventative**. Maybe there is an **analogy** here?"*

From a negative point of view, we highlight following answers (our emphasis):

- *"It is **not clear** to me what "keeping all changes and impacts caused by such changes tracked" means. Do you mean that you will use another spreadsheet for example but the same graphics will be generated?"*
- *"The **research questions** could be **changed**."*
- *"Perhaps it is useful to have **types of maintenace**, as in software maintenace, corrective, perfective, etc ... In my opinnion, maintenance, is only related to the **time period, other changes are then considered a new SLR/SM**."*

Based on software maintenance definition presented in (SOMMERVILLE, 2011): *"Software maintenance is the general process of changing the system after it has been delivered"* and on survey comments, the definition was simplified, and presented to some interviewed researchers (Chapter 4). The new definition is: *"SLR maintenance is the general process of changing a SLR after it has been reported."*.

3.3.2 Traceability

Table 5 shows questions about repositories use presented on survey, and respective answers.

Table 5: Repository questions.

Question	Yes	No
Should SLRs be versioned using tools such as Git or SVN?	23 (82.1%)	5 (17.9%)
Would you keep your SLR artifacts in a common repository such as GitHub?	20 (71.4%)	8 (28.6%)

Source: The author (2019).

Related to question *"Should SLRs be versioned using tools such as Git or SVN?"*, participants were asked to detail they answers, then, we call attention to following answers:

- *"Versioning would be important for comparisons, or even reusing the same extracted data to perform other studies."*
- *"Configuration control (thus, version control and management of issue reports) is essential to track and control SLR protocol evolution/maintenance."*
- *"I think there are already tools that help in managing the artifacts and stages during the SLR, however I believe that a versioning tool can improve this management."*
- *"It is good to use these tools but it shouldn't be required"*
- *"SLRs is not software"*

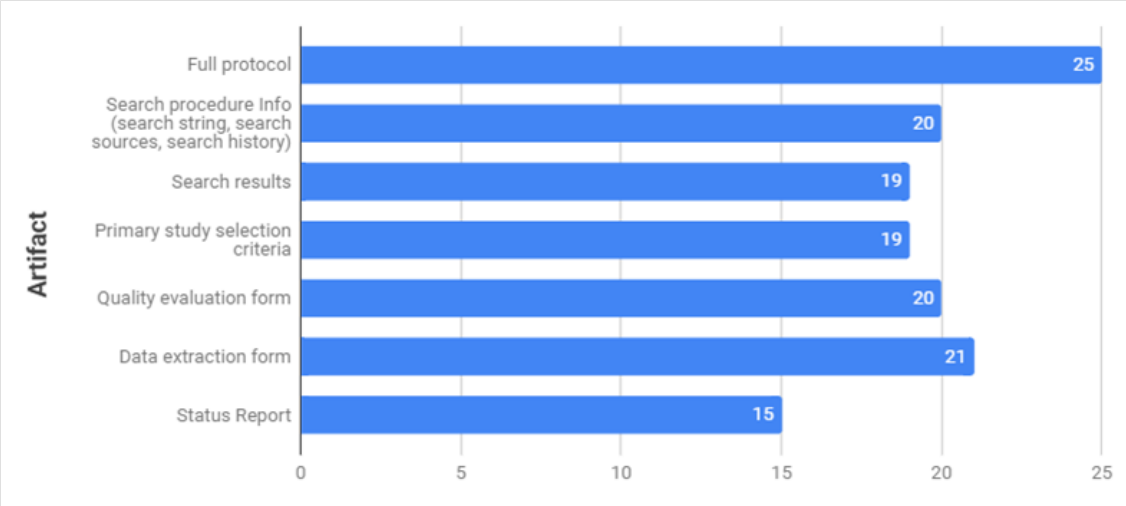
Related to question *"Would you keep your SLR artifacts in a common repository such as GitHub?"*, we also ask participants to detail their answers. We emphasize following detailed answers:

- *"I believe that it is important to make available the artifacts, to enable people to double check how one achieved the reported results."*
- *"Yes, this would make it much easier to manage and trace the updates, and identify the studies that might have brought new evidence or change the previous results of an SLR."*
- *"But only if I can control who has access to it while the review is still ongoing."*
- *"Again auditability and transparency are hallmarks of good research - we should be prepared to share."*

- *"When the knowledge has been building in systematic way is more consistent."*
- *"didn't appear important as we document this in the publication"*

Participants were asked about which artifacts were considered essential to be versioned, Fig. 10 presents results.

Figure 10 – Artifacts to be versioned.



Source: The author (2019).

3.3.3 Plagiarism

Participants were asked about what they consider plagiarism in SLRs. We highlight following answers (our emphasis):

- *"Really **not different** from any other study."*
- *"An SLR that has very similar questions to a previous SLR and the additional covered period **does not enable adding new evidence** (e.g., **less than 2 years**)."*
- *"Re-using text or the work of others without due credit. So in the context of SLR maintenance, I guess re-using the **data collection** and **analysis** of the results previous SLR without providing **ample references**."*
- *"use the same **string**, same **comments**, **conclusions**, same **RQs**"*
- *"Execute an SLR with the same **goals** and **research questions** from another SLR, at **close period**, in general, with the same parameters, obtaining **similar answers**. Or, perform a replication in the future period without **mentioning** that it is a **replication**."*

Other answers are in some way related to those already presented. We can verify there is a concern of citing original work, as in any other work. However, we have verified some

specific situations for SLRs, such as protocol parts reuse, data collected, and analysis performed without due credit. In addition, we asked if replications could be considered plagiarism, only three participants considered replications as plagiarism, in these three answers was placed, as the problem, new evidence lack.

The following scenario was used for participants assessment: *A researcher is updating an SLR, reusing its original protocol and reporting the same evidence, even with a different set of primary studies. Is the researcher committing plagiarism? Assume the researcher references the original work properly when reporting the results.* Only two participants pointed out this scenario should be considered plagiarism.

3.4 DISCUSSION

Regarding to problems accessing the original artifacts, researchers who reported difficulties claimed they were not accessible as mentioned in text. Therefore, a common repository, where SLRs could be available, would reduce this problem impact. Regarding to contacting the original authors during the SLR update, making this contact with researchers from original work can greatly reduce plagiarism issues chances and spent less effort to understanding the original artifacts. In general researchers did not present concerns if others conduct updates in their SLRs, but, even if update is done by others, contacting original group is be a good practice, preventing groups from working on same topic.

Related to answers in keep the SLRs up-to-date, in general, it is possible to check in researchers' comments the interest in keeping their SLRs up-to-date, but concerns about effort to be spent during the process, as well as, publishing results importance of these updates are visible. Related to effort, a well defined maintenance process would avoid a SLR original method replication decreasing effort and collaborative knowledge construction techniques use can help to distribute effort. Regarding to updates importance, academia might come to understand these updates should not necessarily generate new publications, but rather as a more robust knowledge construction.

Fig. 8 shows no consensus on what would be just an update and what would become a new SLR. For example, the response "Use a different time frame in the search procedure" was selected as a reason for creating a new SLR, however it is strange one considers this, since it is expected in an update process to use a new time frame searching for new primary studies. Therefore, creating a naming and criteria for defining when an SLR is an update and when it is a new SLR will improve discussion about the topic. Comparing how SLRs are updated in other fields such as medicine, education and psychology, and what reasons to carry out these updates, as well as questioning experts in SLRs area in SE can clarify what an SLR update really is. This discussion is further elaborated in Chapter 5.

Regarding the SLR update definition given in the survey and based on software maintenance definition presented in (SOMMERVILLE, 2011): *"Software maintenance is the general*

process of changing the system after it has been delivered", chosen because it is simple and easy to understand, and on survey comments, the definition was simplified, and presented to some interviewed researchers (Chapter 4). The new definition is: *"SLR maintenance is the general process of changing a SLR after it has been reported."*.

Related to answers given for traceability questions, the answers corroborates our belief that versioning SLRs and storing artifacts in common repositories can bring significant gains to SLR maintenance process. For example, the possibility of using mining repository techniques to search for new evidence, or to compare versions of different SLRs, looking for any kind of plagiarism, as well as, a possibility of performing maintenance in a collaborative way.

Regarding to plagiarism questions, it was verified the need to deepen the understanding of what is plagiarism in SLRs. Questions about this topic were made during interviews reported in Chapter 4. Some guides to avoid plagiarism are placed in Chapter 5.

3.5 THREATS TO VALIDITY

Some threats were identified during survey conduction:

- Using only publications number for interviewees' selection. Other factors, such as citations number should have been considered.
- Emphasized answers can generate a bias and therefore bring threats to validity to our work, however, we try to reduce this problem by choosing answers that represent all existing opinions.
- It is not possible to generalize results presented to all researchers who have ever conducted SLRs. However, it is believed response numbers to the defined population size is representative.

3.6 CHAPTER SUMMARY

In this chapter we have presented survey results about what researchers who have already developed SLRs think about updating SLRs. We also proposed a definition for SLR maintenance based on survey answers. We present several opinions about control version tools use and their use in SLR maintenance, and also opinions about plagiarism issues when a researcher is maintaining a SLR. With this knowledge, the reader will be able to understand our proposal for guide maintenance process. Next chapter presents a guide for maintaining SLRs. In order to validate our guide and bring some new contributions we conduct an interview with SLRs conductors.

4 SLRS MAINTENANCE GUIDELINES

This chapter presents guidelines to conduct a SLR maintenance process. Using results obtained from SM and survey, a proposal was built and evaluated with researchers through interviews. We improve our guidelines using considerations given by researchers.

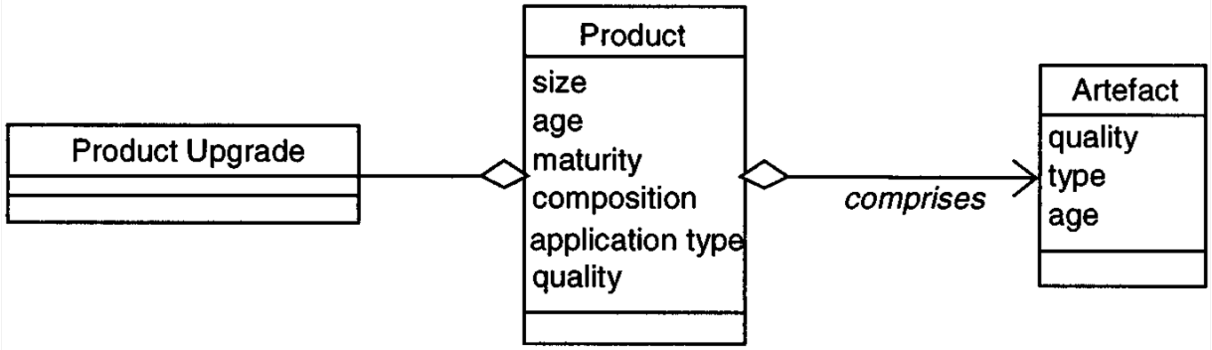
Section 4.1 provides definitions about software maintenance and how these definitions can be reused in SLR maintenance. Section 4.2 presents the proposed guidelines. Section 4.3 shows a evaluation process using a qualitative semi-structured interview study, and presents feedbacks and amendments given by interviewed researchers. Section 4.4 shows a chapter summary.

4.1 DEFINITIONS IN SLR MAINTENANCE

As pointed out previously in Chapter 2, there were no studies that deal with SLRs' maintenance in SE. Therefore, it is essential for this area development to define several terms that will be presented in proposed model. Researchers need well-defined naming to conduct their research, since it is possible to identify in other areas that lack of terms understanding can lead to wrong conclusions ((KITCHENHAM et al., 1999), (CHAPIN et al., 2001)).

Before presenting the proposed guidelines, we must first define some important points naming that are placed in the guidelines. Fig. 11 presents, through an ontology use, definitions about what is a product, what is an update of product and what are artifacts that make up product in software maintenance. Using definitions presented in Table 6, a terms correlation with what we verify in SLRs is made (Table 7).

Figure 11 – Maintained product ontology.



Source: Source: (KITCHENHAM et al., 1999).

Table 6: Maintained product ontology definitions.

Term	Definition
Product	The product is a software application, product or package that is undergoing modification. A product is a number of different artefacts conglomerate.
Product upgrade	A change to baseline product implementing or documenting a maintenance activity. An upgrade may be a new version of product, an object code patch, or a restriction notice.
Artefact	Artefacts that together correspond to a software product can be of the following types: documents that can be subdivided into textual and graphical documents, COTS products, and object code components. Textual documents include source code listings, plans, design and requirements specifications.

Source: (KITCHENHAM et al., 1999).

4.1.1 SLR Attributes

4.1.1.1 SLR Age

SLR age can affect how maintenance process should be conducted, so it is important to try to keep them alive. Possible effects of SLR age on the maintenance process are presented:

- It is possible the documentation created while conducting SLR has been lost, which makes its maintenance more complicated.
- People involved in original SLR may not be of more interest, and may not be available for possible inquiries.
- Other SLRs on same domain may have been conducted, i.e. it may have become obsolete.

4.1.1.2 SLR Quality

Following a process established for SLRs conduction, tends to increase reliability degree on the same (BRERETON et al., 2007). Therefore, original SLR quality impacts directly on new versions. Carefully evaluating SLR quality before performing a maintenance activity is essential to ensure new versions reliability. Unless maintenance activity is done to solve this problem in the original SLR.

Table 7: Correlation with SLRs.

Term Correlation	**Definition**
Product — SLR	SLR that is undergoing modification. A SLR comprises a set of artefacts, and has as attributes its age, quality, history and domain.
Product upgrade — SLR Change	A SLR change is any update on artifacts that comprise SLR itself. A change may be required for some reasons, such as a new search procedure or adjustment in method application, for example.
Artefact — Artifact	We are considering SLR artifacts as any documentation used to elaborate and conduct SLR, any extraction form used to store data from primary studies (quality assessment and data extraction steps), and SLR's reports.

Source: The author (2019).

4.1.1.3 SLR History

SLR history is linked to everything that has happened after it has been delivered. If it has already been replicated, if new versions have been created, what types of analysis have already been done using published data.

When a maintenance activity is performed, just like in software systems, we must be careful with rework. It is possible the update has already been performed by another team, which in this case can lead to plagiarism issues. Plagiarism will be better discussed in Chapter 5.

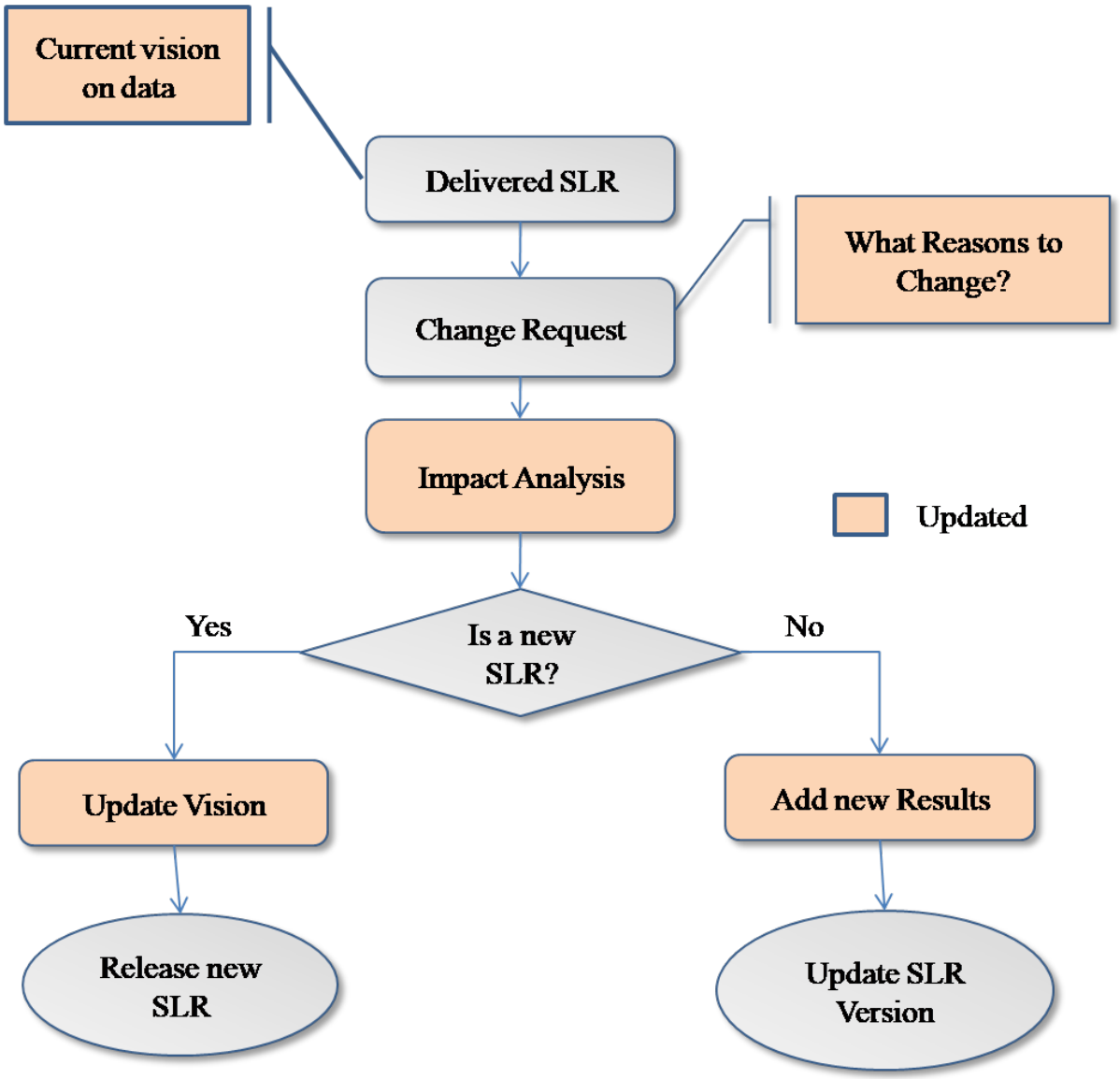
4.1.1.4 SLR Domain

Obsolescence of an SLR is tied to your research domain. Interest in a SLR, and its maintenance, is directly influenced by interest in SLR domain. For example, let us imagine an SLR that was performed on improvements in software development process using cascade methodology, it is common knowledge that this methodology is used today only for comparison purposes with new methodologies and therefore maintaining an SLR on this subject may not be practicable.

4.2 GUIDELINES PROPOSAL

Maintenance is required to ensure SLR continues to satisfy user needs. therefore, using results obtained by systematic mapping presented in Chapter 2, by survey results presented in Chapter 3 and concepts of software maintenance originated from works presented in Section 4.1, we propose guidelines for maintaining systematic literature reviews in software engineering (Fig. 12). These guidelines were the final version after contributions made by the interviewed researchers (Highlighted). The proposal presented to re-

Figure 12 – Guidelines to perform SLRs maintenance.



Source: The author (2019).

searchers during interviews (Section 4.3) and on which we obtained several improvement comments that were incorporated into our guidelines (Section 4.3.3.1) are in Appendix F. The guidelines were created using standard flowchart elements.

4.2.1 Delivered SLR

As stated in Table 7, delivered SLR is a product that is undergoing modification, with its respective attributes and represents the current vision on data. (BENNETT; RAJLICH, 2000) states after software first version delivery, there are two important outputs, the architecture built, which will remain throughout software life span and the knowledge acquired by the team during development. A parallel can be done in conducting an SLR. For a delivered SLR, the architecture, in this case, would be the protocol, which should

remain as faithful as possible during SLR lifetime and the researchers knowledge who carried out original version is of extreme importance to ensure a fair conduction in future versions (RODRIGUEZ et al., 2017). For a better maintenance activity conduction, it is expected a delivered SLR:

- has all artifacts available, since original SLR;
- has contact information from researchers participating in original SLR, or from the version which researcher is performing a maintenance activity.

4.2.2 Change Request

Change request is a process by which we come to a new maintenance activity. As SLR conduction is carried out by means of rigorous processes, and because an SLR represents a snapshot in time on a certain subject, until now, it was not expected it needs changes. Nevertheless, both in SM presented in Chapter 2, and in survey presented in Chapter 3, the need to perform maintenance activities was verified. Reasons why these activities are required vary greatly. Table 8 present five possible reasons for performing maintenance activities.

Table 8: Reasons to change a SLR.

Reason ID	Reason	Description
R1	Update Findings	Keep evidence up-to-date to track a particular area evolution.
R2	Methodological Issues	Any research is subject to methodological issues, so it would be interesting to fix these problems without necessarily realize a new SLR.
R3	Improve Reliability	An SLR that can have its data interpretation, or analysis, reviewed, improved and compared to other SLRs makes results more robust and reliable.
R4	SLR Extension	Perform a new analysis on existing data.
R5	Domain Search	Search for new results in domains other than original.

Source: The author (2019).

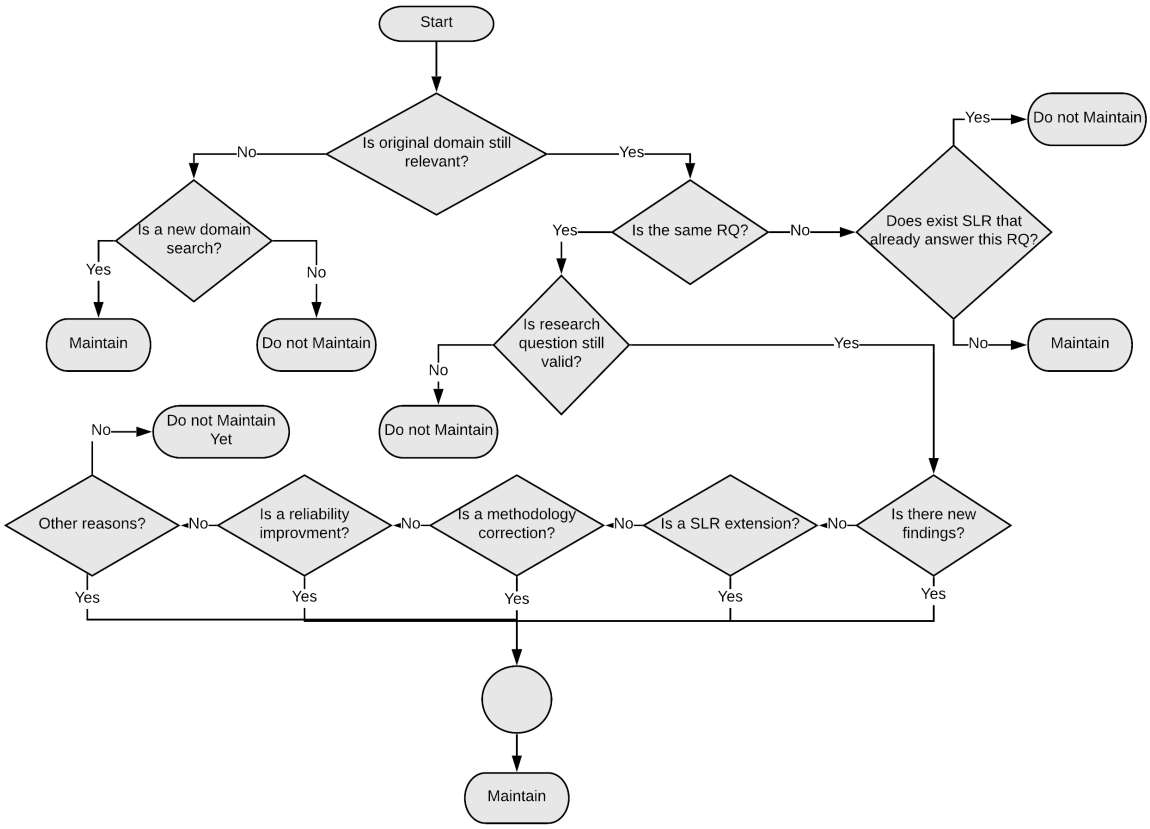
4.2.3 Impact Analysis

The cost and impact of these changes are assessed to see what gains are, and once the need is verified, a new SLR release is conducted. At this point researcher will need to

make a decision to just update SLR to a newer version, or changes' impact will lead to a new SLR.

As a process output, it is expected that, besides decision to perform or not maintenance activity, the need for modifications on artifacts is evaluated. After this evaluation we make a decision whether maintenance activity will result in a new SLR, or just an update. A decision-making process was built to assist impact analysis in SLR maintenance activities in SE (Fig. 13). This process is discussed in Section 5.1.

Figure 13 – Decision-making process to assist impact analysis.



Source: The author (2019).

Participation of a researcher, who has already contributed to a delivered SLR evolution process is of great importance to carry out analysis, just as it happens in software maintenance (SOCIETY; BOURQUE; FAIRLEY, 2014). Some criteria that can be analyzed for decision-making process:

- SLR artifacts availability;
- Presence of researchers who have contributed to SLR evolution in new version conduction;
- SLR domain importance;

- SLR conduction history, verifying impact new activities can have on new version quality.

4.2.4 Update Vision

During analysis phase, the data interpretation leads to a new vision of these data due to changes in SLR artifacts, provoking the need to create a new SLR, which is based on original SLR. Table 9 shows changes that can lead us to a new SLR.

Table 9: Artifact Change Evaluation.

Artifact	**Change Evaluation**
Protocol	• Has research question been changed? • Has search string changed? Adaptations for search engines are not considered. • Have changes in inclusion/exclusion criteria, which may have impacted primary studies of original SLR, been performed? • Have changes in quality assessment, which may have impacted primary studies of original SLR, been performed?
Data Extraction Form	• Are old extraction fields disregarded? • Are new extraction fields added? • Is this change caused by some high-impact change in protocol?

Source: The author (2019).

4.2.5 Add New Results

During analysis phase, it can be concluded that the protocol is faithfully followed, or the vision over data does not change even if there are changes in artifacts, causing previous version to no longer be used as a reference. A list of minor changes is showed:

- Using new sources in search procedure.
- Change time frame to fetch new primary studies.
- Perform quality assessment. If this process is not used in selection of primary studies.

- Report new results using available data.
- Report data in a different format from original.

4.3 EVALUATION

After model proposition is necessary its evaluation. There are several methods in literature to evaluate propositions in scientific research, such as experiments, case studies, surveys, interviews, among others. As the proposed guidelines involves many subjective questions about SLRs maintenance process understanding, a qualitative research was performed. In this section we present a set of semi-structured interviews, with experienced researchers in SLRs conduction, carried out with the proposed guidelines evaluation purpose.

Semi-structured interviews are often used to collect data in qualitative research ((SEAMAN, 1999), (HOVE; ANDA, 2005), (MYERS; NEWMAN, 2007), (MOLLéRI; PETERSEN; MENDES, 2019)). Considering this study purpose is to evaluate a model presented in Section 4.2, through experienced researchers opinions in conducting SLRs, semi-structured interviews are presented as an adequate technique for the task.

4.3.1 Interview Planning and Conduction

In this research, the focus is to evaluate a proposed SLR maintenance process according to point of view from experienced researchers in conducting SLRs. In order to do this, we used GQM goal template (SOLINGEN et al., 2002) to structure the following general study goal:

Analyze **systematic literature reviews maintenance process.**
For the purpose **of improving systematic literature review Process.**
With respect to **keep systematic literature reviews as up-to-date as possible.**
From the viewpoint of **experts in conducting systematic literature reviews in software engineering.**
In the context of **empirical software engineering researchers.**

4.3.1.1 Research Questions

Research's general goal was subdivided into three research questions:

RQ1.: *What do experienced researchers think about SLR maintenance?*
RQ2.: *How should SLR maintenance process be performed?*
RQ3.: *Can SLR maintenance lead to plagiarism issues?*

RQ1. is related to specialists point of view on the need to perform maintenance activities in SLR. Although the need has already been pointed out during survey conduction (Chapter 3), during interview it is possible to better understand researchers real needs.

RQ2. is related to the process of conducting maintenance. Why we need to perform maintenance, what infrastructure is needed to perform it, and how to leave the process uniform are questions to be answered so researchers know the way to do it.

RQ3. raises a complex discussion regarding plagiarism in SLRs. During survey (Chapter 3) we verified the need to deepen discussion with specialists.

4.3.1.2 Interview Design

Table 10 presents a relationship between research questions and questions asked during semi-structured interview. Questions were organized to give a fluent conversation. Questions were reviewed by advisor and validated with a pilot interview.

4.3.1.3 Procedures

Semi-structured interview was prepared in order to collect and analyse data about SLR maintenance process. Each interview duration was designed to last around 45 minutes. This time was respected with all interviewees, and it was not necessary to stop any interview for any reason.

At the interview begin, interviewees had opportunity to ask questions about interview topic, or any questions about interview process. A brief explanation of what interview is about and what was expected during interview was performed. In the end, interviewees had opportunity to put their additional thoughts, comments, or lessons learned during conducted interview.

4.3.1.4 Sample

Target population of this semi-structured interview are software engineering researchers who have already conducted more than five SLRs, what we consider a more experienced group. As pointed out in 3.2.2, our sample was taken from (ALMEIDA, 2017). 42 researchers were selected to participate in interview.

Initial objective was to interview 10 researchers, from a universe of 42. However we had great difficulty in contacting many of these researchers and we had the denial of some of them, in all negatives the claim was a lack of time. It was possible to perform two interviews with this group. For this reason, we invested our time by contacting researchers who participated in survey (Chapter 3) and who became willed to participate in interview, thus obtaining five more researchers, totaling seven interviews.

Table 10: Interview questions.

QID	RQ	Question	Sub-questions
Q1	RQ1	Have you already performed a systematic literature review update?	- Did you publish the update? How(Full paper, short paper, poster, etc)? - Why Not? - How the update was conducted (a new SLR)? - Do you consider the result important given the effort used?
Q2	RQ1	Considering the following software maintenance definition: "Software maintenance is the general process of changing the system after it has been delivered". Do you consider SLRs can be maintained?	- Why do you consider it important? - Why Not?
Q3	RQ2	Do you consider the reasons given below enough to perform maintenance of an SLR? Do you disagree with any presented reason?	
Q4	RQ2	To perform software maintenance we must evaluate the use of version control tools (Git, SVN, CVS), do you consider the use of these tools is also useful for maintaining SLRs?	
Q5	RQ2	Considering the presented process, is it able to guide the maintenance of SLRs?	
Q6	RQ3	What do you consider plagiarism?	
Q7	RQ3	Do you consider publishing SLR updates can lead to plagiarism (self-plagiarism)? What can be done to avoid this problem?	- What can be done to avoid this problem?

Source: The author (2019).

Four participants are Brazilian and had their interviews conducted in Portuguese, three are from other nationalities and had their interviews conducted in English. All interviews were conducted via Skype and stored in MP3 format, privately, in cloud, in a non-shared folder.

All interviewees were contacted by e-mail (Appendix D). Email contains reasons why researcher was chosen, interview subject, duration, and an attachment (Appendix E) containing some information that was used during interview. The benefit of participating

in this study was an opportunity to contribute to ESE development.

The main risk for respondents was fatigue. However, participants were free to stop at any time, including ending their participation. None of questions asked involved personal or user-sensitive issues, therefore, with a low risk of embarrassment. Participants were not paid for their contribution, participation was voluntary.

4.3.1.5 Piloting

A pilot study was conducted to test materials and procedures to check potential problems and interview fluidity. Pilot was conducted in Portuguese with a Brazilian researcher using Skype. We conducted pilot with a Ph.D. in empirical software engineering, and who has experience in conducting SLRs. After pilot, the participant gave feedback on related study, interview content, and interviewer fluency. Generated audio was stored with other interviews audios. Audio total time was 49 minutes. Two files were generated due to connection problems, but according to interviewee, there was no impact in interview fluidity. We conclude the study was designed according to study goals.

4.3.1.6 Data Collection

Interviews were scheduled according to researchers availability. All interviews were conducted using Skype. We used Movavi Screen Recorder (Version 9) software to record interviews and extract MP3 files.

Interviews were transcribed in same language as original recorded. Transcription process will be further explained in next section. Due to confidentiality issues, Table 11 was created to relate audios to interviewees. Organization was carried out following interviews chronological order. During analysis the codes presented will be used, participant names will not be provided. P1 and P2 are participants from first sample group.

Table 11: Participants ID.

ID	**Language**	**Interview Date**
P1	Portuguese	October 5, 2018
P2	Portuguese	November 21, 2018
P3	English	December 17, 2018
P4	Portuguese	December 17, 2018
P5	Portuguese	January 09, 2019
P6	English	January 15, 2019
P7	English	January 22, 2019

Source: The author (2019).

4.3.1.7 Data Analysis

Transcribing audio to text process is time-consuming, or generates a cost often prohibitive to be performed by third parties. In addition, transcription by third parties can cause errors that may not be perceived by researcher, leading to a rework that may not correct the problem (DUFFY; FERGUSON; WATSON, 2004). An existing concern when working with untranscribed audios is the difficulty in finding passages identified by researcher. The difficulty is even greater with large amount of data. However, it is possible to encode directly from audios, allowing researcher to focus on sections that really interest him. In addition, researcher will always need to revisit original material with all its imbued and non-verbal meaning (MARKLE; WEST; RICH, 2011).

A heuristic widely used to aid data analysis process is coding. In this process we create labels that symbolically relate a summative attribute to a particular piece of data. Coding can be considered a form of data pre-analysis (SALDAñA, 2009). Coding, in general, has two cycles. First cycle performs an initial coding, it may be sufficient depending on research goals. Second cycle comes from the need to create subcategories, but this cycle is not always necessary, it is commonly used to create theories from data (SALDAñA, 2009).

As research objective is to evaluate the proposed guidelines, without the need to create a theory, the approach carried out in this research was an implementation of structural coding process. Structural Coding seeks, from data coding, to create conceptual sentences that are related to interview research questions (SALDAñA, 2009). Always revisiting original files, for each iteration codes were placed in a data sheet. Table 12 shows last iteration results. Codes were grouped by research question.

4.3.2 Interview Results

This section presents interviews results. Results are presented grouped by research questions and identified by participant ID (Table 11). Excerpts extracted from interviews conducted in Portuguese were translated into English and reviewed by research advisor.

4.3.2.1 RQ1.: What do experienced researchers think about SLR maintenance?

Interviewees presented their interpretations on SLRs maintenance process. Excerpts have been extracted to help us get a better view on subject.

Published systematic reviews number has been increasing in recent years (BORGES et al., 2015). However, it is possible that many of them are investigations very similar to others already existing. P1 raises following situation:

> *"There are already too many revisions in our area, there are people doing things they do not need."*

Table 12: Coding Results.

Research Question	**Structural Codes**
RQ1.: What do experienced researchers think about SLR maintenance?	Less Effort
	Need to perform
	Need to publish
	Reuse of instruments
	New Results
	Difficulty in publishing
	Software Depreciation X SLR depreciation
	Economic Model
	Time to update
RQ2.: How should the SLR maintenance process be performed?	Agreement
	Update Findings
	Publish Results
	New Reasons for Maintenance
	Laboratory package
	Not necessarily as GIT
	Version Artifacts
	Transparency
	Reliability
	Impact Analysis
	Process Detailing
	New Review X Update
RQ3.: Can SLR maintenance lead to plagiarism issues?	Replication
	Credit
	Community
	Measurement
	Licensing/Policies
	Intentionality
	Collaboration
	New Publication

Source: The author (2019).

Maintenance process is not only to make changes in SLRs, it also serves as a way of publicizing data, and obtained results. Maintaining a SLR can help to preserve a research active, as it is placed in following snippets:

P1: *"... if you look at the systematic review as a snapshot, a snapshot of that*

time ... I do not think it makes much sense to maintain that. Now, if you view systematic review as a living artifact ... then it does. "

P1: *"A systematic review ... where I can go there and deposit the most recent data on what I did on that subject, it would never be outdated."*

P2: *"... I made a comparison ... I published and had another guy... that published, using even the same title as mine, then we did another study later."*

Even if maintenance is not the main goal, even if it is not desirable for researcher, for lack of interest, as in following fragments:

P2: *"... in those that are not in my area, I honestly have no interest in updating ... and in those that were in my area ... I would have to have a very great motivation ... because I changed a little bit of scenery ..."*

P3: *"I only did one update ... I sent a paper to a journal, and the reviewer says that I need to update ... but this was not my goal."*

P5: *"Hardly we publish twice ... so if it's published in the first version, it's not published on second... and vice versa."*

P6: *"You did not really like it, and maybe you want a completely retake of it."*

P7: *"My initial thought was that. If we need to correct something, we should not have published in the first place."*

They have realized the need for maintenance process, because they have already done it, or because they realize other researchers may want to perform an update on their SLRs, or they see potential need in a near future.

P1: *"... update ... that took the years from 2008 to 2012 ... and then we continued, maintained this thing until 2014, more or less, wrote a mega paper, sent to journal, but not passed."*

P2: *"The analysis was more in the methodology sense ... we researched on the same topic and we had some results that were different, so the analysis we did was: Why something came out different if the process is the same? ..."*

P4: *"I agree ... but ... I think it's not comparable, you do software maintenance with review maintenance."*

P5: *"... They can be maintained, that is, I do not know that there is any structured way to do it ... but it's what I said, we usually do it once, twice for a systematic review, which is different from what we do with software."*

Relative to a 2014 SM, P7 said: *"We are discussing right now if we should update or not."*

When asked if an SLR can be maintained as a software, the general view is it cannot be exactly like software, but there are similar situations between processes.

P2: *"These (software) updates should, and have to be done faster ... more urgently than an systematic review update."*

P4: *"If the review was as widely used as software in production, then it might make sense you do that, this comparison."*

P4: *"Software depreciates very quickly, a systematic review I do not know if the data depreciates quickly like this."*

P5: *"... in software we have a structured way, a technique to do maintenance is important because you will do it many times and this will cost you, so you do in an ad-hoc way it will probably increase its cost because you will often do this in an unstructured way, in case the systematic review does not happen, then you do it once, even if it is not in an structured way ... it may not be a problem ."*

P6: *"I suspect that the issues are slightly different of maintained software and maintained SLRs."*

P7: *"... yes, it could be the same, except for that we typically we do it some updates neither in the protocol or just to add new results."*

However, when we presented the proposed guide, acceptance was greater, because they realized intention is not to copy the process, but to adapt it to SLRs reality. Another point questioned, to those who have already performed an update, was connected to effort, below are the impressions:

P1: *"The effort is similar."*

P3: *"... same person that did the first searches, did the second, so it's not a hard work ..."*

P4: *"Extremely smaller ... smaller quantity ... as I had read so many previous works that I already knew where the information, I was looking for, was, and also that all the work I had to elaborate an instrument, how if data extraction and analysis were done, everything was ready, I just had to apply."*

P5: *"I think it's too smaller ... we do not have a methodology to do ... in an advisor figure ... I do not know that they(students) have applied any specific methodology to do the update. So basically it is to re-run the searches on the search engines, if you have the snowballing, redo, but I think the effort is much less ... Even because we do not need to re-consult, or re-analyze all those articles which were analyzed in the first version."*

The proposed guidelines should make it more visible to researchers that maintenance process is not a SLR realization from scratch but rather a complementation of activities already carried out.

P1 presents vision idea, which is something that transcends review. Original review gives us a first vision into a theme, a snapshot. Maintenance activities lead us to new looks about first vision presented.

"The value added by the review is in the vision. The review is the tool that we have today to retrospectively generate this vision, but if we maintain this vision alive and updated, in whatever way, the value will continue to be in the vision. Then you can do a parallel with the software, not in the review, but in the vision ... you have a vision maintenance ... "

For naming reasons, we continue to call the product to be maintained as systematic literature review. However, we understand what we are calling systematic literature review is actually a vision idea put forward by P1.

4.3.2.2 RQ2.: How should the SLR maintenance process be performed?

During interviews, reasons (Table 8) why a researcher might have the need to perform a maintenance activities were exposed. Interviewees made their comments on given reasons and made suggestions for new reasons. Then we asked about how SLRs should be maintained, regarding to tooling use, and finally we present the proposed guidelines. Researchers were asked to make their comments by presenting their opinion on the guidelines. Several suggestions for proposal improvements and changes have emerged.

4.3.2.2.1 Reasons to Maintain

In general, Given reasons were well accepted and understood.

P2: *"... I would not add any more ..."*

P3: *"The other issues (R1 and R2) are more evident ... the first one, it is the most evident and in some sense easier to publish."*

P4: *"For me, these 3 points are interesting ..."*

P6: *"I think these reasons are good. I agree with all of them."*

R3 (Improve Reliability) was the one that generated most doubts about the need, as putted by P3: *"The last one is difficult to appreciate without an example."*. Mainly, as for difference from R3 in relation to R2 (methodological issue), since a issue repair would be performed to increase reliability.

Another point raised by interviewees was regarding R1, with respect to term "Evidence". P1 and P3 have considered this term inappropriate use:

> P1: *"Updating evidence is not the correct term, because updating evidence means update the primary study. ... the interesting fact about this is that, in updating the systematic review, it happens, and I remember having happened to us, we have found evidence update ... we found a more recent study of the same primary studies set who had been indexed by the previous review and who brought more details about the previous study ..."*
>
> P3: *"Sometimes, when you are talking about systematic mappings, you are not always collecting evidence."*

Two new reasons have also been raised. P4 and P6 put SLR extension idea, as we can identify in following fragments.

> P4:*"Another reason why you can do a systematic review, which would be in the sense of not only doing an update, but doing an extension ... exploring other things based on the same data source ..."*
>
> P6: *"... you may perform at some point that you want to add some new analysis, without additional evidence, without additional primary studies."*

P7 points to the case where same search can be performed in several domains: *"The same search in a different domain."*. P7 also added: *"The actual reason is because it is my phd thesis."*

One issue raised during reasons discussion for performing maintenance activities is a trade-off between effort and publishing. Some researchers have put as process output, reasons regardless, the publication.

> P1: *"I'm updating the review and creating a new paper."*
>
> P3: *"I was thinking on the effort that you have to do, on you do one systematic update, and I was thinking that perhaps the first one, even that take more time, for update the evidences, perhaps, this kind of maintenance deserves a publication."*
>
> P3: *"You change some methodological issue, but don't the protocol, I don't know if this small change deserves a update ... sometimes ... we drive our research, because we want to publish."*

P4: *"... the first (R1) ... for me, it is the main reason, the part of methodological issues and improvement reliability, these two, I think they would be more complicated, because, normally, in systematic review, hardly, we have anything beyond article."*

P4: *"If you will improve the methodology, or improve reliability, I do not know if this will be enough to publish a new article."*

P5: *"Here the reviewer of your article says: No, you did not do it right and you have to redo it."*

4.3.2.2.2 Tooling

Respondents were asked about version control tools use for managing artifacts generated by SLR. Git tools use for version control support was given as an example, as it is already done for software. Version control idea was well accepted, however, Git-like tool usage raised concerns.

P1: *"I'm looking at the situation in which we consider the systematic review as the final report ... I think the final report does not need to. Now, looking at the situation that we think the systematic review is the data and the algorithms for analyzing that data ... this, I think it makes total sense to have a version control."*

P2: *"I find it interesting, versioning, now I find it very complicated to do this because in code we have a dependency between the lines of code, here it is very subjective ... I do not see a dependency ... the dependence that is going to happen. is in relation to methodology, or in relation to interpretation and results? ... it is quite complicated."*

P4: *"Certainly, I do not know if the Git logic itself would work."*

P5: *"It would be interesting, but I do not see these particular tools ... having the necessary functionalities."*

P6: *"I think that in a high level, yes ... but I am not fully sure that it would be the only way to do it."*

P7: *"I think it's important. It is that the old SLR does not get lost. ... If we would update a version of a SLR, we need to keep track on the second one."*

Some considerations about version control tools use have been made and are listed in snippets below. Discussion on these points will be further elaborated in Section 4.3.3.

P1: *"There are different artifacts and the maintenance value of these deals varies."*

P2: *"I do not know how often researchers do these updates."*

P2: *"This is essential, from the point of view that happens, at least with me, when I was doing the reviews, I would read an article, then I will look at the article. The guy says he made a thing and when I go look, look for that thing that he did to get results. he does not know where it is, he does not know who did it, he does not know how it was done."*

P2: *"I did a review, that the other reserachers needed my data, the search spreedsheet, ... all my analysis. I even had a spreadsheet here, and I sent, but if I had an online repository to put it in and someone would somehow contribute to update, it would be interesting."*

P2: *"... who would publish this? who would own this knowledge? The guy who created the repository, or the guy who inserted the update?"*

P3: *"Are you thinking in a similar repository like, do you know the cochrane?"*

P5: *"Although we do systematic reviews in a uniform way, it is not so uniform. To keep that inside a repository, you need somehow to have a uniform structure for systematic reviews."*

4.3.2.2.3 Guidelines

After guidelines presentation, interviewees were led to make considerations. The process presented was well accepted, however, researchers felt the need for a greater detail in some points.

P4: *"Honestly, this process for me is very good, I do not think you should improve, so improve in order to modify the process, but one thing is certain, what you're going to have to do is leave those steps more specific ."*

P5: *"On a more abstract level, I think it's perfect, ..., what I think I need to have is a detailing, especially these last two levels."*

P6: *"I think it's fair enough ... I think it's good enough to start with, definitilly"*

P7: *"I think that in general the structure is feasible, sounds plausible to me."*

P3, trying to summarize the guidelines' understanding, states that the central question is: *"Are you considering to change something in the protocol, or not?"*. It is understood that changes evaluation on protocol is a focal point on the guidelines. However, there are situations in which changes in protocol, which would lead to believe that should certainly be a new SLR, can still cause doubts, as putted by P1.

"If you change the research question and this research question does not change the search string, and the set of studies is the same, then you are reusing what

you have done before. You will just need to do, in the analysis, a different analysis to be able to answer the question. My question is: Does this set up a new review, or not? This argument makes sense and it strengthens the idea that, what has the major value of the whole review is the final vision that will be provide from those data ... when I change the research question, then yes, because the vision, at the end, will be completely different. ... looking at this sense, that the SLR can turn out to be a living business, the guy, who is doing an SLR, has to take a lot of responsibility with that business. Because, if the SLR in the future happens to be considered like a living organism ... if anyone likes it and wants to update it, you'll have to talk to me. Unless I publish in an opensource logic ... in a github line ... "

4.3.2.3 RQ3.: Can SLR maintenance lead to plagiarism issues?

Before deepening discussion about plagiarism issues in maintaining SLRs, we asked interviewees what they understood as plagiarism.

P1: *"Plagiarism is an infraction of intellectual rights, or copyright."*

P2: *"Copying ideas, copying questions, not just copying text."*

P4: *"Plagiarism, for me, is if I take your data and say that your data was mine, or else, I get your text and say that your text was mine."*

P6: *"... failure to give credit."*

P7: *"Repetition without proper citation."*

P7 poses as plagiarism is treated in Germany, taking into consideration the fact plagiarism can happen unintentionally.

"In germany, we have a very clear definition of a scientific miss contact ... whenever I detect plagiarism I have to also proof that is on purpose, that it is intentional ... if it is unintentional it should not be published, but if it is unintentional it also not be punished."

When questioned about possible plagiarism issues, including self-plagiarism, in maintaining SLRs, a concern about problem can be seen. However, it was pointed out that the question is more general, and not specifically for SLRs maintenance, as pointed out in fragments below.

P1: *"From the reader's point of view ... if I get an article called: an update of the systematic review X, and it has 90% the same text as the previous one, ... I delete the previous one, ... as a conference evaluator, it is a greater responsibility, if we see that the guy is sending an article to a conference and*

90% of that article replicates what he wrote at another conference, I would deny it. ... Unless it's 10% that contributes a lot, it's still a big responsibility to take the risk of plagiarism. ... If we are talking about systematic review as that living organism, which is constantly updated, this question loses its meaning. ... because the previous version does not matter anymore."

P3: *"I think that, now, the community is very worried about you are talking about, in some sense."*

P4: *"I reuse someone else's data, to me, this is not plagiarism. For example, picking up the studies you have collected and collecting other data about it, I present it and even compare it with yours, this does not mean plagiarism."*

P5: *"If the person wants to use it incorrectly, ie do the plagiarism, it is not a matter of being a systematic review, or not. ... you have published once, if you have a significant contribution, that justify your second publication, fine. As long as you make it clear, you had a systematic review with that result, we did an update and the new contributions are these ... so this avoids any blame for plagiarism."*

P6: *"... the open source software community has to be able to manage this fairly well, but they do not really leave on citations in publications the way we do, the researchers."*

P7: *"Update to SLR can never be plagiarism, if it is properly cited."*

One point raised by some interviewees was the criteria used to consider plagiarism. Sometimes, situation can become subjective, since just automatic tools use can cause misunderstanding. Therefore, a possible solution is to create a license to manage potential plagiarism issues in conjunction with creating an SLR repository, as is the case in other areas.

P1: *"Now comes the subjective question. Is a phrase a plagiarism?, or how serious is the plagiarism of a sentence? Because if I'm talking about a 5-page plagiarism, it's serious enough plagiarism for the newspapers to be upset with me. ... if you have one, or a set of little phrases, that I re-use to put in this second article, although yes, it is a plagiarism, it is a small enough plagiarism for nobody to care about it. ... It makes sense to think of an appropriate license to do this, or to adapt an existing license, to eliminate this problem from plagiarism."*

P3: *"Are you thinking in a similar repository like, do you know the cochrane? ... You have to change the policies of editors. ... When they compare with tools, to compare the file, without any semantic comparison."*

P6: *"I like the idea of keeping he entire process open, in the sense, such that it could be overlooked and transparent, even if the SLR has been conducted. ... It's a good example with the cochrane reviews."*

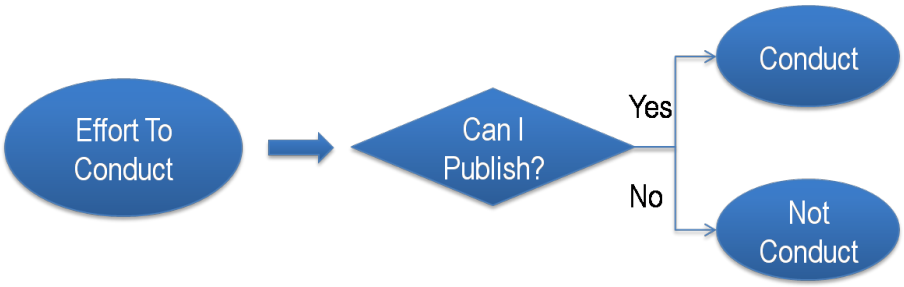
P7: *"... but many tools fail measurably ..."*

4.3.3 Discussion

No studies were found sought to understand the real need to maintain SLRs. Therefore, we sought to explore opinion of experienced researchers in SLRs conduction, to evaluate maintenance process formalization that assists this task conduction. First, it was tried to understand what researchers mean by maintaining SLRs, then how that task should be performed, and, finally, if maintenance process would lead to plagiarism problems.

At first, SLRs maintaining idea may not seem feasible, since at first, researcher only thinks if effort employed will result in a publishable study (Fig. 14). This problem was already identified during survey (Chapter 3) and was again highlighted during interviews. Following this principle, some problems can be seen:

Figure 14 – Effort X Publishing.



Source: The author (2019).

- Published works with overlapping contents;
- Researcher will always have effort of conducting an SLR from beginning;
- Different research groups doing the same work;
- Greater risks of committing plagiarism;
- SLR will be outdated and will lose its contribution level.

What needs to be understood is: Maintaining SLRs is not just an update issue. In parallel with software production and maintenance, we understand that, although an SLR does not need to be constantly evolved as a software, it may become obsolete and more and more complicated to be maintained over years, as well as software (BENNETT; RAJLICH, 2000).

Another point discussed is SLR definition. SLR, as it is understood today, is a report generated after process conduction, a publication, a thesis chapter. In this work it is understood SLR is a product, constructed from artifacts and that has its attributes, as presented in Section 4.1. If SLR is understood as a research final report, it really does not need to be maintained, but if we understand SLR as a living organism, which would be in constant evolution, as interviewee P1 well pointed out, then SLRs maintenance will make sense.

Therefore, bringing a structured way to carry out SLRs maintenance process, will give effort and time economy, as today happens in software production. It should be noted the way software is maintained is quite different from how a SLR should be, with different problems, different reasons, different costs. However, maintaining an SLR is not as simple as updating a document, in this case, a final report.

Impression given by researchers, who have already performed an update, its effort is similar, or less, than conducting from beginning. One cited point is number of primary studies to be analyzed (1), another point is same researcher performing update (2), as well as, reuse of instruments already created in original SLR (3). With a well-defined maintenance process, it is now possible to improve at least two (1 and 3) of the three presented points, because it will be possible to keep a set of primary studies always updated and evaluated, as well as, reuse all generated artifacts, to make changes in them, and, beyond that, to have a history of everything that has already been extracted and analyzed from studies set. Although maintenance process cannot improve Point 2, directly, it can improve maintenance activity done by third researchers, since it is not always possible to have researcher participation who was involved in original SLR. However, activities history performed by original conductor will be available and will guide others.

Regarding to reasons presented for maintaining SLRs, a conceptual doubt about difference between R2 and R3 was raised. Mainly because correcting methodological problems leads us to improve reliability, and in many cases, increase reliability means to correct methodological problems. However, it is understood increasing reliability pointed out in R3 is not related to method, but to data interpretation and analysis. Meaning R3 is a comparison between the published SLR with others already existing in literature, it is desired to improve, or even redo, data analysis. As exemplified by P2, in which performed an SLR on same subject as another researcher, including publishing work with same title, but reaching different results. Therefore, a third work was conducted to verify presented results differences, since both followed same method for conducting SLRs.

Still related to reasons, two interviewees pointed to misuse of term "evidences" in R1. Taking into consideration that in SLRs, evidence can be understood as a primary study retrieved in search process, update evidence, would mean update primary study, which is not a desired interpretation. Just as in systematic mapping, due to its exploratory characteristic, we do not look for evidence, but rather, for a broader presentation about a

particular research area. Therefore, term "evidence" has been replaced by term "findings", which fits better with the idea that one wants to update obtained results.

Two new reasons emerged during interviews: "SLR Extension", which is to perform a new analysis on existing data (RABISER et al., 2017), and "Domain Search", which is a search for new results in domains other than original. A third reason was that researcher is performing maintenance activity for being part of his thesis, however, it is understood this justification is quite personal, and it should be embedded in one of other reasons.

Regarding version control tools use, it is understood a tool like Git can generate an unnecessary overhead when maintaining SLRs. However, interviewed researchers perceived the need for a tool that not only can store artifacts generated during a SLR, also allowing to evaluate artifacts history, SLRs collaborative construction, managing contributions made by researchers, and creating a standard publication format for methodological process used, in which other researchers can easily find what interests them in SLR to perform a maintenance activity.

Still regarding version control, participant P2 asked a question related to SLR ownership: *"Who would own this knowledge?"*. We believe building an SLR should follow open source software fashion, in which third parties can study, change, and improve original SLR. Once published, results obtained and analyses carried out are owned by those who published them, however, artifacts generated would belong to community. A new way to understanding this approach is the open science movement, which is a knowledge construction in a transparent and accessible way through collaboration between researchers (VICENTE-SAEZ; MARTINEZ-FUENTES, 2018).

Concerning to the proposed guidelines comments, process acceptance is perceived as basis for conducting SLR maintenance activities. Various contributions have been taken into account and will be incorporated into process (Section 4.3.3.1). However, this is a first glimpse into SLRs maintaining process, and like any process formalization, still lacks depth. Some gaps in compliance or formalization can still be left blank. However, it is hoped, with future community collaborations, these guidelines may become a basis for conducting maintenance activities on SLRs.

Another question raised by interviewees was about what needs to happen so maintenance activity result is a new SLR, or just an original version update. According to the guidelines, this decision needs to be taken after impact analysis procedure. We understand evaluation may be subjective, but for process formalization, it is stated maintenance process will deliver a new SLR if data interpretation leads to a new vision of these data due to changes in SLR artifacts. An SLR version update happens when protocol is faithfully followed, or when vision over data does not change even if there are changes to artifacts, causing previous version to no longer be used as a reference. It is understood as vision the current knowledge about data, understanding SLR as a living organism, which is always being updated.

Finally, regarding contributions made by interviewees about SLR plagiarism, a discussion will be further elaborated in Section 5.4.

4.3.3.1 Guidelines Improvements After Interviews

This section incorporates contributions made by experienced researchers during interviews are into the guidelines. Table 13 shows a comparative picture of what we had before and after interviews. Appendix F contains the version presented to interviewed researchers and an example using the SM conducted in Section 2.4.

Table 13: Experienced researchers contributions.

Contribution	**Original Text**	**New Text**
Update Reason	Update Evidences: Keep evidence up-to-date to track evolution of a particular area.	Update Findings: Keep findings up-to-date to track evolution of a particular area.
Update Reason	Improve Reliability: A SLR that can be reviewed, improved and compared to others SLRs makes results more robust and reliable.	Improve Reliability: An SLR that can have its data interpretation, or analysis, reviewed, improved and compared to other SLRs makes results more robust and reliable.
New Reason	-	SLR Extension: Perform a new analysis on existing data.
New Reason	-	Domain Search: Search for new results in domains other than original.
Guidelines Step Update	Update Artifacts: Requested changes caused significant updates in original artifacts, provoking the need to create a new SLR, which is based on original SLR.	Update Vision: Data interpretation leads to a new vision of these data due to changes in SLR artifacts.
Guidelines Step Update	Add new Results: Requested changes caused minor updates in original artifacts.	Add new Results: Protocol is faithfully followed, or the vision over data does not change even if there are changes in artifacts, causing previous version to no longer be used as a reference.
New Concept	-	Vision: Current knowledge about data.

Source: The author (2019).

4.3.4 Trustworthiness of this Study

Possible threats that could affect trustworthiness of this interview study and corresponding mitigations are described below. Threats were classified according to Maxwell's categorization (MAXWELL, 1992).

4.3.4.1 Descriptive Validity

To avoid descriptive threats, all interviews were recorded, and all extraction and data analysis was done from original audios. With this, it is possible to capture all imbued and non-verbal meaning in interviewees' speech (MARKLE; WEST; RICH, 2011).

4.3.4.2 Interpretive Validity

In order to avoid interpretive threats, information placed in results (Section 4.3.2) are interviewees interpretations about questions asked, conducting researcher view is only presented in discussion (Section 4.3.3). In addition, participants with experience in SLRs conduction were selected and had necessary knowledge to present their points of view with clarity.

4.3.4.3 Theoretical Validity and Generalizability

As this study did not aim to create a theory based on information collected in interviews, but rather an evaluation of the proposed guidelines, we should note these threats do not apply to this study.

4.3.4.4 Evaluative Validity

In this research, there were no evaluations of what was said by interviewees, nor evaluations for respondents. Thus, we do not worry about this threat.

4.4 CHAPTER SUMMARY

The main objective of this chapter was to present a formalization for SLRs maintenance process. For this, software maintenance concepts that were reused for guidelines construction were presented, which allows a reader to better understand the proposed guidelines. In order to proposal evaluation, interviews were conducted with experienced researchers in SLRs conduction. During interviews, researchers were asked what they understand about SLR maintenance, as well as the proposed guidelines applicability, and whether SLRs maintenance process can lead to plagiarism issues. In next chapter, discussions are held on guidelines applicability, a comparison with other research areas, what to expect from a support tool for conducting maintenance activities, how to publish results of those activities, and an in-depth discussion on plagiarism in SLRs.

5 DISCUSSION

In this chapter discussions are held on SLR maintenance process key points. Key points emerged during research and require a deeper insight into a better process understanding.

In Section 5.1 a proposed model comparison with what is implemented in other areas of knowledge is carried out. Section 5.2 describes expected requirements for a tool to support SLR maintenance process. Section 5.3 discusses how maintenance activity results can be published. Section 5.4 puts the understanding obtained during research on what is plagiarism in SLRs and how it can affect SLR maintenance process. Section 5.5 presents a chapter summary.

5.1 SLR MAINTENANCE IN OTHER FIELDS

In order to answer RQ2.1 (*Is it possible to reuse SLR maintenance methods from other areas?*), a discussion was carried out on how SLRs maintenance is conducted in other areas of knowledge.

Systematic reviews are conducted in various knowledge areas beyond software engineering, such as health care and social sciences. In 1993, Cochrane Collaboration[1]was created to help health care researchers to prepare and maintain their systematic reviews. Most recently, Campbell Collaboration[2]was created to disseminate results from systematic reviews in social, educational, and criminological areas (PETTICREW; ROBERTS, 2006). There is no similar initiative in SE.

In our research we found a main reference for conducting systematic reviews in health care and for other areas of knowledge, the Cochrane Handbook for Systematic Reviews of Interventions (HIGGINS; GREEN(EDITORS), 2011). Cochrane Handbook was the only reference found that deals with systematic reviews maintenance, and, therefore, will be used for comparison purposes with the guidelines proposed in this work.

5.1.1 Updating SLRs

This work defines what is SLRs maintenance and shows differences from what would be just an update process. However, in addition to Cochrane Handbook, no papers that talked about SLRs maintenance were found. Although several studies use Cochrane Handbook as a basis for conducting SLRs, they only talk about updating SLRs ((PETTICREW; ROBERTS, 2006), (TACCONELLI, 2010), (COLLABORATION, 2019)).

Remembering the following definition of an SLR update given by (MOHER et al., 2008):

[1] https://www.cochrane.org/
[2] https://www.campbellcollaboration.org/

> *"A distinguishing feature of an updated systematic review from a new review is that during updating the originally formulated protocol (e.g., eligibility criteria, search strategy) is retained, and sometimes extended, to accommodate newly identified information (e.g., new treatment type, diagnostic method, outcome, different population)"*

This definition is in accordance with update idea presented in this thesis. An update must be performed with possible least impact in original protocol, only adjustments must be made to allow update to be conducted. However, as presented in Section 4.2, it is understood that other activities, in addition to update, can be performed after SLR delivery.

Performing an update is a process that requires less effort than starting from scratch ((PETTICREW; ROBERTS, 2006), (GARNER et al., 2016)). Questions raised by various researchers are about when and how to perform an SLR update. There are two main points when an update is considered: First point is whether original SLR theme remains relevant, second is original SLR quality, which can serve as a starting point for update (TACCONELLI, 2010).

Several methods that evaluate the need to update an SLR have been proposed in health care ((SHOJANIA et al., 2007), (SHEKELLE et al., 2011), (BASHIR; SURIAN; DUNN, 2018)). (SHOJANIA et al., 2007) has created a process to determine SLRs updating status. Quantitative and qualitative assessments were taken into account. In their results, 20% of studies evaluated presented an update quantitative indication, when considering qualitative evaluation, 54% of studies showed updating indications. (SHEKELLE et al., 2011) compared two methods that evaluate the need to update SLRs, they are RAND and Ottawa methods. RAND uses experts opinion in the area to reach a conclusion about update realization. Ottawa uses qualitative and quantitative evaluations on search results similar to that performed in original study. (BASHIR; SURIAN; DUNN, 2018) evaluated the time-to-update relationship between published SLR and subsequent updates in Cochrane database. It was concluded, due to effort to carry out an update, a careful evaluation is necessary, since in only eight of 204 reviews updated there was a change in conclusions.

The tool developed by the Cochrane Opportunities Fund to assist in assessing the need to update systematic reviews of Cochrane Database help us to elaborate a similar tool for maintaining SLRs in SE. The checklist to be evaluated in second step is extensive and divided by Cochrane review sections. Some of triggers can be reused, because they can be directly related to review sections at Kitchenham's guidelines (KITCHENHAM; CHARTERS, 2007), which is the most commonly used in SE. However, some points are specific to Cochrane reviews. A closer look at the checklist needs to be done for a possible adaptation.

The decision tool proposed in (HOPEWELL et al., 2008) was updated in (TAKWOINGI et al., 2013). Two new steps have been added. In the updated tool, first step is to answer

the following question: *"Is the clinical question already answered by the available evidence or is the clinical question deemed no longer relevant?"*. Second step is a decision tree application already presented in Hopewell's work. Third step is to verify existence of new primary studies that are relevant to study conclusions through a statistical prediction tool. In the next section, a decision support process, based on these tools, is presented.

5.1.2 When maintain SLRs in SE?

The tools presented in previous section cannot be used, directly, in SE, since they are intrinsically linked to health care. However, one can carry out a process adaptation to extend its use to SE, especially in relation to described triggers. After a proposed process evaluation in (TAKWOINGI et al., 2013), a decision-making process adaptation was built to assist impact analysis in SLR maintenance activities in SE (Fig. 13). We understand decision-making process is qualitative and, therefore, requires a specialist on review domain participation and, if possible, a previous maintenance processes member participation. An adaptation was carried out taking into account steps 1 and 2 proposed in (TAKWOINGI et al., 2013). Step 3 is very specific for Cochrane reviews reality, when performing a quantitative evaluation on new studies inclusion. Another point to consider is a proposed checklist adaptation from (HOPEWELL et al., 2008) to the guidelines proposed by Kitchenham, which will be addressed in a future work. When carrying out impact analysis, it is necessary to ensure:

- SLR artifacts are available;
- Members have sufficient knowledge about SLR domain;
- Decisions must be fully documented for transparency:

  What changes were made to artifacts;

  Another reason why maintenance was performed;

  Reasons for maintenance does not have to be performed.

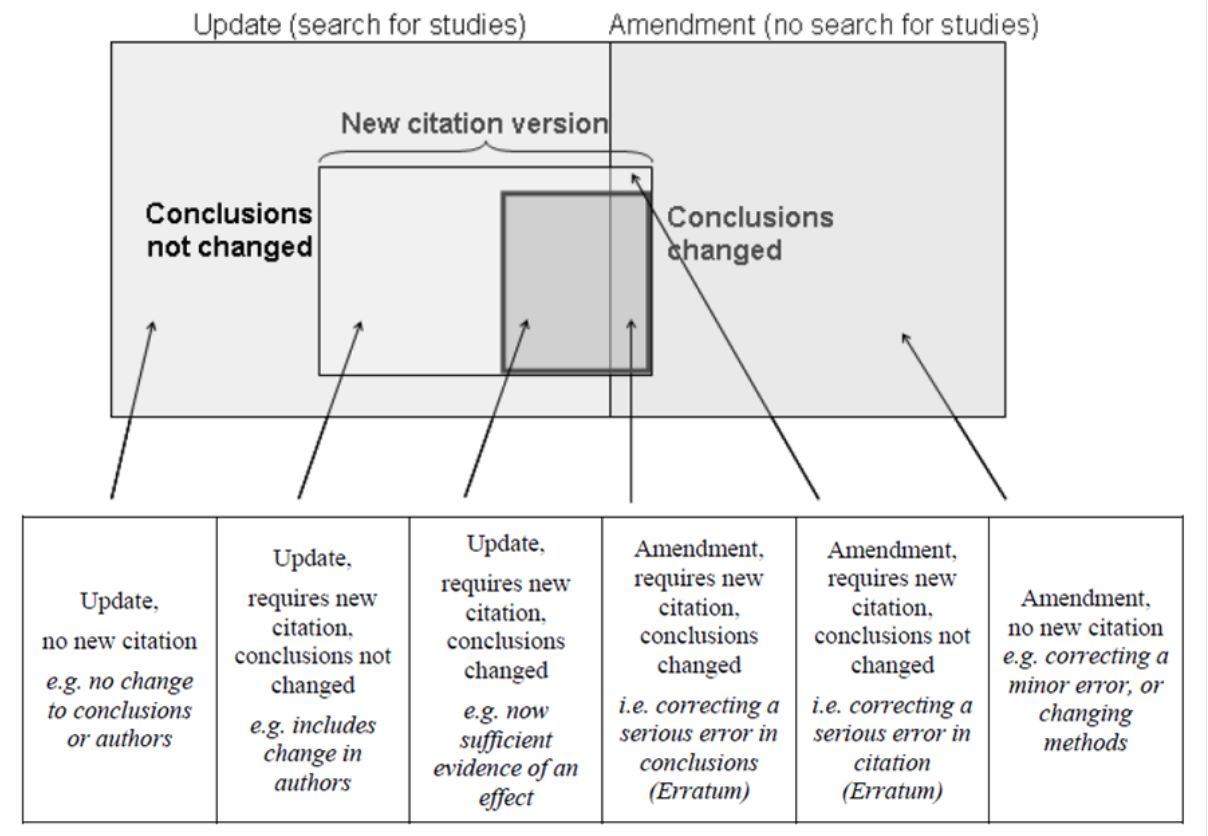
5.1.3 Maintaining Reviews in Cochrane Database

A Cochrane Database stored review important feature is they should always be up-to-date. For this, there is review authors commitment to maintain their reviews. Cochrane Collaboration stipulates a two-year interval for updating your reviews, or author should explain reasons why you did not update to Cochrane Review Group. There is not yet a standard way to evaluate when an SLR should be updated, but, as shown in previous section, there are several methods proposed in literature.

A change made in a Cochrane review should be considered as an update, or an amendment. Change will be considered an update when there is a new search for new primary

studies, and any other change will be considered an amendment. Like any maintenance process, Cochrane reviews use a versioning scheme, in this case a citation version. When published, review receives its citation version, and depending on changes made, review may, or may not, receive a new citation version (Figure 15).

Figure 15 – Summary of changes to Cochrane reviews.



Source: (HIGGINS; GREEN(EDITORS), 2011).

An important mechanism in Cochrane reviews maintenance process is feedback, which are comments made by reviews users. Feedback is received by Cochrane Review Group and forwarded to authors, who must respond clearly to review user, including whether there have been changes in reviews due to feedback.

5.1.4 Why not reuse Cochrane approach?

Cochrane Collaboration approach is focused on updating final report and protocol. We note an SLR consists of an artifacts set, of which final report and protocol are part. Maintenance of these various artifacts cannot be performed using only concepts presented in Cochrane Handbook.

Software engineering has great expertise in maintenance area due to complexity presented to accomplish this task in software. A key point for choosing to use concepts coming from SE is version control, which is much more complex than just using citation version.

In addition, Cochrane Collaboration's understanding is an SLR may undergo updates or amendments and only the latest version produced has validity. In this study, it is understood some changes in original SLR may lead to new views construction on data, without necessarily having previous vision become outdated. An example of this is a new research question completion on existing data, which will give us a new insight into data without necessarily change previous findings.

5.2 WHAT EXPECT FROM A SLR REPOSITORY

There are several tools to assist SLRs conduction in SE (MARSHALL; BRERETON, 2013). However, once conducted, SLRs have only their final report published. According to (BRERETON et al., 2007), review authors need to keep a record of decisions taken while conducting SLRs, which is corroborated by (STAPLES; NIAZI, 2007). (BRERETON et al., 2007) also puts SE community needs to establish mechanisms for their SLR's publication becoming available the entire review process, not just a report.

Before presenting what is expected from a SLRs' repository, we put what is not a necessity for a SLRs repository. However, there is nothing to prevent tools, which already perform functions listed below, from working as repositories.

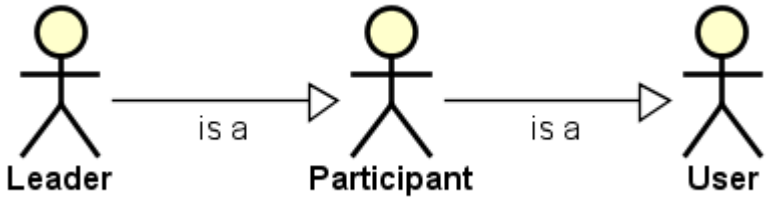
- Protocol Elaboration — Repository should not be concerned with protocol elaboration. However, we understand a standard format for artifacts needs to be built, including a standard format for protocol.
- SLR's Conduction — Repository does not need to assist SLR conduction. However, it is expected to store decisions made during this process.
- Generate reports — Repository does not have to be a reporting tool. However, it is understood stored artifacts need to be easily accessed, facilitating reports construction.

A SLR repository cannot serve only as a database, where users search for SLRs without any kind of iteration and participation in research. The desired repository has to work as a collaborative construction database, where data and respective findings are validated by users themselves through discussions. At the same time, the repository must be able to manage various SLRs versions while maintaining their change history. The repository should also support maintenance process steps defined in the proposed guidelines, always making clear those responsible for changes made. The repository essential characteristics are listed on Table 14 through user stories use, which is a technique originated in XP methodology for scope definition in software projects. User story default format is:

"As a <role> I can <capability>, so that <receive benefit>"

Basically, we have three roles in tool's use: user role, participant role in conducting an SLR, which is a user role specialization, and leader role in conduction process, which is a participant role specialization (Fig 16).

Figure 16 – SLR repository's roles.



Source: The author (2019).

Table 14: SLR repository's user stories.

ID	User Stories
US01	As a user I can access my personal information, so that I can manage them.
US02	As a user I can search for SLRs, so that I can contribute with them.
US03	As a user I can start a discussion in any SLR, so that I can provide feedback for SLR's participants.
US04	As a user I can sign up to participate in a SLR so that I can contribute as a participant.
US05	As a participant I can manage artifacts from a SLR, so that I can perform maintenance activities.
US06	As a participant I can create maintenance activities, so that any participant can perform them.
US07	As a participant I can change artifacts' version, so that I can keep change history.
US08	As a participant I can give reasons to perform maintenance activities, so that any participant can evaluate changes.
US09	As a leader, I can register a new SLR, so I can publicize SLR data.
US10	As a leader I can monitor a SLR participants team, so that I can check maintenance activities progress.
US11	As a leader I can accept new participants so that they can perform maintenance activities.

Source: The author (2019).

The repository must also have some characteristics that assist essential activities, and allow greater security for users, participants and leaders:

- The repository must have secure storage for its SLRs;
- The repository must have a conduct ruleset for users, with policies of impartiality and sources of information made available;
- The repository must make available on which license a SLR will be published;
- The repository must provide an easy access way to SLRs data. E.g.: git clone, or webservice.

5.3 PUBLISHING MAINTENANCE RESULTS

A major concern about how to publish SLR maintenance activities results was raised during survey (Chapter 3), such as, "I think it is not worth the effort of updating a SLR. It is hard to publish it again. The contributions may not be significant.", and later in interviews (Section 4.3), for example, "You change some methodological issue, but not the protocol, I don't know if this small change deserves an update ... sometimes ... we drive our research, because we want to publish.".

The proposed guidelines suggests two possible outputs for maintenance process: (1) A new SLR creation, or (2) An update to SLR version being maintained. When (1) happens, it is understood the concern described above should not occur, because we will have a new vision about data, which in itself justifies publication as in any other scientific study. When (2) happens we may have to think otherwise, since contributions may not really be sufficient for publication in journals and/or academic conferences. In this case, two scenarios are visualized: citation number, and practitioners influence.

Scientific work citation number importance has been studied since 1927 (GROSS; GROSS, 1927). It is common in scientific community to link study quality to its number of citations (PRADHAN; CHAKRABORTY; NANDI, 2019). There are several indexes, which calculate academic influence of a researcher using as a basis their citation number ((HIRSCH, 2005), (EGGHE, 2006)). A SLR that is no longer maintained becomes outdated, causing its number of citations to decrease over years after its publication. Already a constantly updated SLR provides greater assurance its results are still valid, which will increase a SLR life span, thus maintaining its academic citation indices.

(KITCHENHAM; CHARTERS, 2007) guideline's Section 7.1 presents strategy importance for disseminating SLR results, and it is already pointed out the need to transfer knowledge to practitioners different of what is done in academia. Recent studies have already realized this transfer importance, and means of transmitting knowledge have been studied and developed in these studies ((GORSCHEK et al., 2006), (GRIGOLEIT et al., 2015), (CARTAXO; PINTO; SOARES, 2018)). According to (GORSCHEK et al., 2006):

> *"Technology transfer happens over time — with small, incremental, and sometimes unplanned improvements to the overall research effort — and is adopted by practitioners continually."*

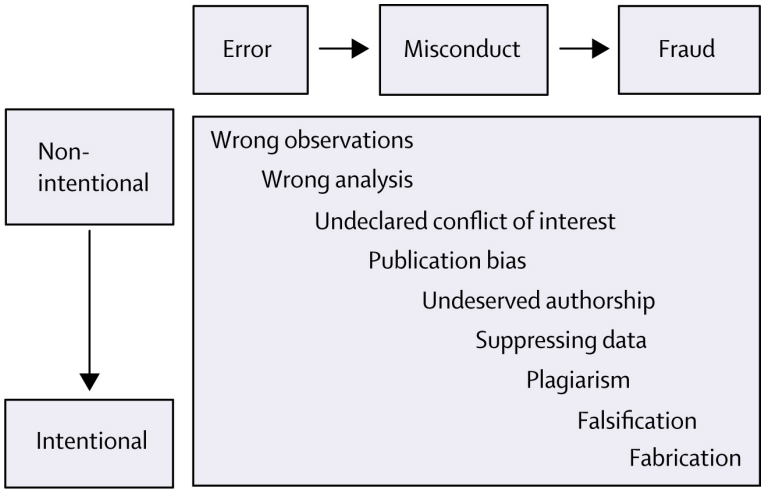
Conducting SLRs to support technology transfer proposal from (GORSCHEK et al., 2006) is impractical with existing guidelines. It is necessary a process allowing constant research evolution, with several views about the most updated data. The SLR maintenance guidelines proposed in this thesis, technically, makes possible incremental improvements. In practice, tool support, like the one described in Section 5.2, could make feasible a collaborative process in knowledge construction between researchers and practitioners.

5.4 PLAGIARISM

Section purpose is not to present a definition for plagiarism in systematic reviews. Also, because there is no single definition for plagiarism in scientific community. In this section we argue about several points in SLRs conduction, which can lead to problems with plagiarism, especially, related to SLRs maintenance.

One issue to be pointed out is plagiarism is not the only form of misconduct in academic research. We can have from unintentional errors to intentional frauds (NYLENNA; SIMONSEN, 2006). Figure 17 shows (NYLENNA; SIMONSEN, 2006) understanding about possible misconduct in scientific research.

Figure 17 – Slippery slope between honest errors and intentional fraud, with examples in the middle.



Source: (NYLENNA; SIMONSEN, 2006).

Several plagiarism definitions were found. Some of them are presented below:

- *"the action or practice of taking someone else's work, idea, etc., and passing it off as one's own; literary theft."*, Oxford English Dictionary.

- *"to steal and pass off (the ideas or words of another) as one's own; use (another's production) without crediting the source; to commit literary theft; present as new and original an idea or product derived from an existing source"*, Merriam-Webster online Dictionary.

- *"plagiarism ranges from the unreferenced use of others' published and unpublished ideas, including research grant applications to submission under "new" authorship of a complete paper, sometimes in a different language. It may occur at any stage of planning, research, writing, or publication: It applies to print and electronic versions."*, Committee on Publication Ethics (COPE).

In survey (Chapter 3) and in interviews (Section 4.3), researchers were asked what they understood plagiarism is. As in definitions presented, the general concern is with study proper citation and published by others. We understand, therefore, plagiarism in SLR is not different in relation to other areas of research.

Although we do not need a specific plagiarism definition for SLRs, when we think in SLR maintenance, we understand there is a great possibility of committing plagiarism if care is not taken. There are characteristics in SLRs maintenance, such as artifacts reuse, data reuse, results reuse, which can lead a researcher to commit plagiarism. In cases where original research is his/her own, it could be commenting self-plagiarism.

According to (LANCET, 2009), self-plagiarism is even less well defined than plagiarism, and self-plagiarism is even more difficult to take actions to penalize this behavior, as there is no theft action. In this case, it is up to journals and conferences to become increasingly able to verify their publications originality.

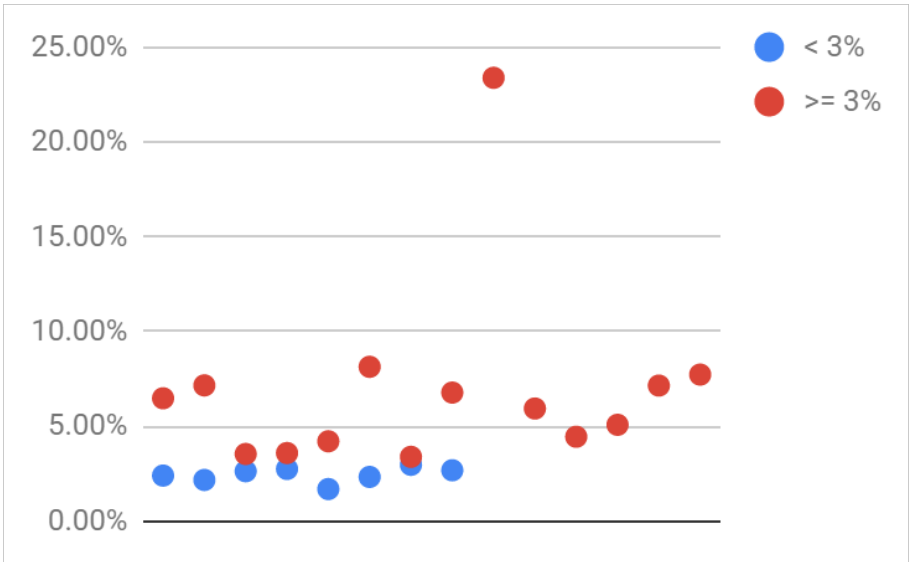
A very common way of performing verification is through plagiarism detection tools use. However, these tools use should not be reckless. Algorithms used by these tools, in general, makes comparison between texts, which can lead to a results masking. These tools results can serve as a starting point for a subjective evaluation.

To present an overview of what could happen in SLRs maintenance, a comparison was made between retrieved studies in systematic mapping on SLRs updates presented in Section 2.4 and their original studies. For this, we use a plagiarism verification tool (CopySpider[3]). According to the tool, its results should not serve as a plagiarism problem measurement, and a 3% limit, presented as possible existence of plagiarism, is based on anti-plagiarism studies. Comparison results are shown in Fig. 18.

It is possible to visualize 14 studies presented a possibility to contain plagiarism, being above 3% of chance, and eight studies are below. The tool failed to perform comparison with one study, and one article was compared with two original articles. The objective is not to verify plagiarism existence in these works, but only to discuss this possibility due to SLR maintenance process characteristics, where it is expected to reuse much of what

[3] https://copyspider.com.br

Figure 18 – Comparison between studies retrieved in SM (Section 2.4) and their original studies.



Source: The author (2019).

has already been done in previous works and which therefore needs to be assessed in a deeper way, not just by a pragmatic assessment tool.

It is believed, in addition to a set of good practices to avoid plagiarism in general, as presented in (KUMAR et al., 2014), we have specific situations for SLR maintenance process:

- Clearly disclose upfront the work is an update of another SLR;
- Clearly explain new primary studies added to original study;
- Clearly state new research questions and what is the impact on final data vision;
- Clearly state all changes made to original SLR artifacts;
- Clearly expose changes in final data vision, making clear differences with previous vision, and whether new vision produced should replace original vision.

5.5 CHAPTER SUMMARY

In this chapter, we have discussed about important topics related to maintaining SLRs. These topics emerged from discussions with experts during research. First, we discuss how other areas deal with maintaining SLRs and why not reuse their methods in SE. Afterwards, a set of user stories were presented as basis for building a SLRs' repository, which is a mandatory tool to perform maintenance activities. Then, a discussion on how to publish a SLR after a maintenance activity is made, showing importance in publish results. Finally, a deeper discussion about plagiarism and SLRs maintenance process impact was performed.

6 CONCLUSIONS AND FUTURE WORK

This chapter presents research conclusions, including answers to research questions in Section 6.1, and future works based on gaps found during research conduction in Section 6.2.

6.1 ANSWERS TO RESEARCH QUESTIONS

- **RQ1.: What is the state-of-the-art on systematic literature reviews updates performed in SE?**

  The goal of this research question was to map SLRs updates. Within this context, it was possible to verify works existence to update SLRs since 2010, however only 22 studies were found until the study conduction date. There is still a very large gap between SLR number of publications and updates. Probably, by an information lack to guide this kind of study and because the motivation to conduct SLRs is largely from academy to academy ((SANTOS; SILVA, 2013) and (CARTAXO et al., 2016)), often to find specific research gaps, frequently making research questions uninteresting to update.

- **RQ2.: Is it possible to use software maintenance techniques to keep SLRs always up to date?**

  The goal of this research question was to check with research community, who conduct SLRs, whether it is possible to create a way to assist SLR update process from procedures already known in software maintenance. In this context, it was possible to verify researchers showed interest in keeping their revisions updated, but they addressed concerns about effort required for this task. This is due to the fact that researchers believe that, to conduct an update, the whole process of conducting a systematic review should be carried out. There was also a lack of consensus on when an SLR needed to be only updated, or when a new revision would need to be created.

  When questioned about artifacts traceability, an essential condition for maintenance process, the majority of respondents consider version control tools use important, even if there is no specific tool for that.

  RQ2.1.: Is it possible to reuse SLR maintenance methods from other areas?

The goal of this sub research question was to discuss how SLRs maintenance is conducted in other areas of knowledge. The most well-known initiative in health care is the Cochrane Collaboration, which is also used as a basis for conduct SLRs in other areas. The Cochrane Collaboration approach is focused on updating final report and protocol. We note an SLR consists of an artifacts set, of which final report and protocol are part. Maintenance of these various artifacts cannot be performed using only concepts presented in Cochrane Handbook.

- **RQ3.: What do experienced researchers think about SLR maintenance?**

  The aim of this research question was to ascertain with experts the need to maintain SLRs. No studies were found sought to understand the real need to maintain SLRs. Therefore, we sought to explore experienced researchers in SLRs conduction opinion, to evaluate maintenance process formalization that assists this task conduction.

  One considered point was whether maintenance result would be publishable. In this context we must think in two points: citation number of an outdated review tends to decrease and for an SLR to keep interesting for practitioners, it needs to keep in constant evolution.

- **RQ4.: How should the SLR maintenance process be performed?**

  The goal of this research question was to present guidelines to maintain SLRs for experienced researchers and to obtain considerations about the process. In general, there was a process acceptance as basis for maintaining systematic reviews. Various contributions have been taken into account and were incorporated into guidelines. However, this is a first view into SLRs maintaining process, and like any process formalization, should have a deeper understanding. Some gaps in compliance or formalization can still be left blank. However, it is hoped, with future community collaborations, these guidelines may become a basis for conducting maintenance activities on SLRs.

- **RQ5.: Can SLR maintenance lead to plagiarism issues?**

  The aim of this research question was to check with researchers if maintenance process, as defined in guidelines, can lead to plagiarism issues. Although we do not need a specific plagiarism definition for SLRs, when we think in SLR maintenance, we understand there is a great possibility of committing plagiarism if care is not taken. There are characteristics in SLRs maintenance, such as artifacts reuse, data reuse, and results reuse, which can lead a researcher to commit plagiarism. In cases where original research is his/her own, it could be self-plagiarism.

RQ5.1.: How can plagiarism issues be avoided during SLRs maintenance?

The aim of this research question was to discuss how it is possible to avoid plagiarism problems during maintenance activities conduction. It is believed, in addition to a set of good practices to avoid plagiarism in general, as presented in (KUMAR et al., 2014), we have specific situations for SLR maintenance process as presented in Section 5.4.

6.2 FUTURE WORK

We present some future works that can be developed based on research conducted in this thesis as follows:

- Create a systematic review repository, verifying partnership possibility with journals and conferences that publish this type of study, trying to apply open science concepts;
- Create a SLR collaborative community aimed at the effort division, making conduction and maintenance processes less costly, and possibly leading to more significant and reliable studies.
- Building instruments bridging researchers and practitioners for collaborative knowledge construction with real industry cases using SLRs.
- Perform a study to assess decision-making process for impact analysis and adapt the checklist proposed in (HOPEWELL et al., 2008) to Kitchenham's guidelines;
- Developing instruments that automatically assess the need for SLR maintenance activities.
- Conduct a systematic mapping maintenance to see if new ways of updating SLRs are being performed;
- Conduct interviews with researchers who work with methodologies that support SLRs, to know their opinion about the guidelines;

REFERENCES

ALMEIDA, K. S. R. de. *Uma Avaliação do Crescente uso de Revisões Sistemáticas da Literatura em Engenharia de Software*. Dissertação (Mestrado) — Centro de Informática, Universidade Federal de Pernambuco (UFPE), 2017.

APRIL, A.; HAYES, J. H.; ABRAN, A.; DUMKE, R. Software maintenance maturity model (smmm): The software maintenance process model: Research articles. *J. Softw. Maint. Evol.*, v. 17, n. 3, p. 197–223, 2005.

BASHIR, R.; SURIAN, D.; DUNN, A. G. Time-to-update of systematic reviews relative to the availability of new evidence. *Systematic Reviews*, v. 7, n. 1, p. 195, 2018.

BENNETT, K. H.; RAJLICH, V. T. Software maintenance and evolution: A roadmap. In: *22th Conference on The Future of Software Engineering (ICSE '00)*. [S.l.: s.n.], 2000. p. 73–87.

BIOLCHINI, J.; MIAN, P. G.; NATALI, A. C. C.; TRAVASSOS, G. H. *Systematic review in software engineering*. 2005. (Tech. Report RT-ES 679/05, Systems Engineering and Computer Science Dept. COPPE/UFRJ).

BORGES, A.; FERREIRA, W.; BARREIROS, E.; ALMEIDA, A.; FONSECA, L.; TEIXEIRA, E.; SILVA, D.; ALENCAR, A.; SOARES, S. Support mechanisms to conduct empirical studies in software engineering: A systematic mapping study. In: *19th International Conference on Evaluation and Assessment in Software Engineering (EASE '15)*. [S.l.: s.n.], 2015. p. 1–14.

BRERETON, P.; KITCHENHAM, B. A.; BUDGEN, D.; TURNER, M.; KHALIL, M. Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, v. 80, n. 4, p. 571 – 583, 2007.

CARTAXO, B.; PINTO, G.; SOARES, S. Towards a model to transfer knowledge from software engineering research to practice. *Information and Software Technology*, v. 97, p. 80 – 82, 2018.

CARTAXO, B.; PINTO, G.; VIEIRA, E.; SOARES, S. Evidence briefings: Towards a medium to transfer knowledge from systematic reviews to practitioners. In: *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM '16)*. [S.l.: s.n.], 2016. p. 1–10.

CHAPIN, N.; HALE, J. E.; KHAM, K. M.; RAMIL, J. F.; TAN, W.-G. Types of software evolution and software maintenance. *Journal of Software Maintenance*, v. 13, n. 1, p. 3–30, 2001.

COLLABORATION, T. C. *Campbell systematic reviews: policies and guidelines*. 2019.

DÍAZ, O.; MEDINA, H.; ANFURRUTIA, F. I. Coding-data portability in systematic literature reviews: A w3c's open annotation approach. In: *Proceedings of the Evaluation and Assessment on Software Engineering*. [S.l.: s.n.], 2019. p. 178–187.

DIESTE, O.; LóPEZ, M.; RAMOS, F. Formalizing a systematic review updating process. In: *6th International Conference on Software Engineering Research, Management and Applications (2008)*. [S.l.: s.n.], 2008. p. 143–150.

DUFFY, K.; FERGUSON, C.; WATSON, H. Data collection in grounded theory-some practical issues. *Nurse researcher*, v. 11, p. 67–78, 2004.

DYBA, T.; DINGSOYR, T.; HANSSEN, G. K. Applying systematic reviews to diverse study types: An experience report. In: *1st International Symposium on Empirical Software Engineering and Measurement (ESEM 2007)*. [S.l.: s.n.], 2007. p. 225–234.

EASTERBROOK, S.; SINGER, J.; STOREY, M.-A.; DAMIAN, D. Selecting empirical methods for software engineering research. In: SHULL, F.; SINGER, J.; SJøBERG, D. I. K. (Ed.). *Guide to Advanced Empirical Software Engineering*. [S.l.]: Springer London, 2008. p. 285–311.

EGGHE, L. Theory and practise of the g-index. *Scientometrics*, v. 69, n. 1, p. 131–152, 2006.

FABBRI, S.; SILVA, C.; HERNANDES, E.; OCTAVIANO, F.; THOMMAZO, A. D.; BELGAMO, A. Improvements in the start tool to better support the systematic review process. In: *20th International Conference on Evaluation and Assessment in Software Engineering (EASE '16)*. [S.l.: s.n.], 2016. p. 1–5.

FELIZARDO, K. R.; MENDES, E.; KALINOWSKI, M.; SOUZA Érica F.; VIJAYKUMAR, N. L. Using forward snowballing to update systematic reviews in software engineering. In: *10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM '16)*. [S.l.: s.n.], 2016. p. 1–6.

FELIZARDO, K. R.; NAKAGAWA, E. Y.; MACDONELL, S. G.; MALDONADO, J. C. A visual analysis approach to update systematic reviews. In: *18th International Conference on Evaluation and Assessment in Software Engineering (EASE '14)*. [S.l.: s.n.], 2014. p. 1–10.

FELIZARDO, K. R.; SILVA, A. Y. I. da; SOUZA Érica Ferreira de; VIJAYKUMAR, N. L.; NAKAGAWA, E. Y. Evaluating strategies for forward snowballing application to support secondary studies updates: Emergent results. In: *32th Brazilian Symposium on Software Engineering (SBES '18)*. [S.l.: s.n.], 2018. p. 184–189.

FERRARI, F. C.; MALDONADO, J. C. Experimenting with a multi-iteration systematic review in software engineering. In: *5th Experimental Software Engineering Latin American Workshop (ESELAW 2008)*. [S.l.: s.n.], 2008.

GARNER, P.; HOPEWELL, S.; CHANDLER, J.; MACLEHOSE, H.; AKL, E. A.; BEYENE, J.; CHANG, S.; CHURCHILL, R.; DEARNESS, K.; GUYATT, G.; LEFEBVRE, C.; LILES, B.; MARSHALL, R.; GARCÍA, L. M.; MAVERGAMES, C.; NASSER, M.; QASEEM, A.; SAMPSON, M.; SOARES-WEISER, K.; TAKWOINGI, Y.; THABANE, L.; TRIVELLA, M.; TUGWELL, P.; WELSH, E.; WILSON, E. C.; SCHÜNEMANN, H. J. When and how to update systematic reviews: consensus and checklist. *BMJ*, v. 354, n. 3507, 2016.

GORSCHEK, T.; GARRE, P.; LARSSON, S.; WOHLIN, C. A model for technology transfer in practice. *IEEE Software*, v. 23, n. 6, p. 88–95, 2006.

GRIGOLEIT, F.; VETRO, A.; DIEBOLD, P.; FERNANDEZ, D. M.; BOHM, W. In quest for proper mediums for technology transfer in software engineering. In: *9th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM 2015)*. [S.l.: s.n.], 2015. p. 1–4.

GROSS, P. L. K.; GROSS, E. M. College libraries and chemical education. *Science*, v. 66, n. 1713, p. 385–389, 1927.

HIGGINS, J.; GREEN(EDITORS), S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. [S.l.]: The Cochrane Collaboration, 2011. Available from <www.cochrane-handbook.org>.

HIRSCH, J. E. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, v. 102, n. 46, p. 16569–16572, 2005.

HOPEWELL, S.; LOUDON, K.; CLARKE, M.; MOHER, D.; SCHOLTEN, R.; EISINGA, A.; FRENCH, S. *A decision tool for updating Cochrane reviews*. [S.l.]: Cochrane Centre UK. Cochrane Opportunities Fund, 2008.

HOVE, S. E.; ANDA, B. Experiences from conducting semi-structured interviews in empirical software engineering research. In: *11th IEEE International Software Metrics Symposium (METRICS'05)*. [S.l.: s.n.], 2005. p. 10–23.

JALALI, S.; WOHLIN, C. Systematic literature studies: Database searches vs. backward snowballing. In: *6th International Symposium on Empirical Software Engineering and Measurement (ESEM 2012)*. [S.l.: s.n.], 2012. p. 29–38.

KITCHENHAM, B.; BRERETON, O. P.; BUDGEN, D. *Protocol for extending a tertiary study of systematic literature reviews in software engineering*. [S.l.]: EPIC Technical Report, EBSE-2008-006, 2008.

KITCHENHAM, B.; BRERETON, O. P.; BUDGEN, D.; TURNER, M.; BAILEY, J.; LINKMAN, S. Systematic literature reviews in software engineering - a systematic literature review. *Information and Software Technology*, v. 51, n. 1, p. 7 – 15, 2009.

KITCHENHAM, B.; BRERETON, O. P.; LI, Z.; BUDGEN, D.; BURN, A. Repeatability of systematic literature reviews. In: *15th Annual Conference on Evaluation Assessment in Software Engineering (EASE 2011)*. [S.l.: s.n.], 2011. p. 46–55.

KITCHENHAM, B.; CHARTERS, S. *Guidelines for performing Systematic Literature Reviews in Software Engineering*. 2007.

KITCHENHAM, B.; PRETORIUS, R.; BUDGEN, D.; BRERETON, O. P.; TURNER, M.; NIAZI, M.; LINKMAN, S. Systematic literature reviews in software engineering - a tertiary study. *Inf. Softw. Technol.*, v. 52, n. 8, p. 792–805, 2010.

KITCHENHAM, B. A.; TRAVASSOS, G. H.; MAYRHAUSER, A. von; NIESSINK, F.; SCHNEIDEWIND, N. F.; SINGER, J.; TAKADA, S.; VEHVILAINEN, R.; YANG, H. Towards an ontology of software maintenance. *Journal of Software Maintenance*, v. 11, n. 6, p. 365–389, 1999.

KUMAR, P. M.; PRIYA, N. S.; MUSALAIAH, S.; NAGASREE, M. Knowing and avoiding plagiarism during scientific writing. *Annals of medical and health sciences research*, v. 4, n. 3, p. 193–198, 2014.

LANCET, T. Self-plagiarism: unintentional, harmless, or fraud? *The Lancet*, v. 374, n. 9691, p. 664, 2009.

MACDONELL, S.; SHEPPERD, M.; KITCHENHAM, B.; MENDES, E. How reliable are systematic reviews in empirical software engineering? *IEEE Trans. Softw. Eng.*, v. 36, n. 5, p. 676–687, 2010.

MARKLE, D. T.; WEST, R. E.; RICH, P. J. Beyond transcription: technology, change, and refinement of method. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, v. 12, n. 3, p. 21, 2011.

MARSHALL, C.; BRERETON, P. Tools to support systematic literature reviews in software engineering: A mapping study. In: *7th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM 2013)*. [S.l.: s.n.], 2013. p. 296–299.

MAXWELL, J. Understanding and validity in qualitative research. *Harvard Educational Review*, v. 62, p. 279–300, 1992.

MENDES, E.; FELIZARDO, K.; WOHLIN, C.; KALINOWSKI, M. Search strategy to update systematic literature reviews in software engineering. To appear at EuroMicro Conference on Software Engineering and Advanced Applications (SEAA 2019). 2019.

MOHER, D.; JENNIFER, T.; MARGARET, T. A. C.; MOHER, A. D. G. Epidemiology and reporting characteristics of systematic reviews. *PLOS Medicine*, v. 4, n. 3, p. 1–9, 2007.

MOHER, D.; TSERTSVADZE, A.; TRICCO, A.; ECCLES, M.; GRIMSHAW, J.; SAMPSON, M.; BARROWMAN, N. When and how to update systematic reviews. *Cochrane Database of Systematic Reviews*, n. 1, 2008.

MOLLéRI, J. S.; PETERSEN, K.; MENDES, E. Cerse - catalog for empirical research in software engineering: A systematic mapping study. *Information and Software Technology*, v. 105, p. 117 – 149, 2019.

MYERS, M. D.; NEWMAN, M. The qualitative interview in is research: Examining the craft. *Information and Organization*, v. 17, n. 1, p. 2 – 26, 2007.

NAKAGAWA, E.; SCANNAVINO, K.; FABBRI, S.; FERRARI, F. *Revisão sistemática da literatura em engenharia de software: Teoria e Prática*. [S.l.]: Elsevier Editora Ltda., 2017.

NEPOMUCENO, V.; SOARES, S. Maintaining systematic literature reviews: Benefits and drawbacks. In: *12th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM 2018)*. [S.l.: s.n.], 2018. p. 47:1–47:4.

NEPOMUCENO, V.; SOARES, S. On the need to update systematic literature reviews. *Information and Software Technology*, v. 109, p. 40 – 42, 2019.

NYLENNA, M.; SIMONSEN, S. Scientific misconduct: a new approach to prevention. *The Lancet*, v. 367, n. 9526, p. 1882 – 1884, 2006.

PETERSEN, K.; VAKKALANKA, S.; KUZNIARZ, L. Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, v. 64, p. 1 – 18, 2015.

PETTICREW, M.; ROBERTS, H. *Systematic Reviews in the Social Sciences: A Practical Guide*. [S.l.]: Blackwell Pub., 2006.

PFLEEGER, S. L.; KITCHENHAM, B. A. *Principles of Survey Research: Part 1 to 6*. 2001–2003.

PRADHAN, D. K.; CHAKRABORTY, J.; NANDI, S. Applications of machine learning in analysis of citation network. In: *6th ACM India Joint International Conference on Data Science and Management of Data (CoDS-COMAD 2019)*. [S.l.: s.n.], 2019. p. 330–333.

RABISER, R.; GUINEA, S.; VIERHAUSER, M.; BARESI, L.; GRüNBACHER, P. A comparison framework for runtime monitoring approaches. *Journal of Systems and Software*, v. 125, p. 309 – 321, 2017.

RODRIGUEZ, L. M. G.; FELIZARDO, K. R.; OLIVEIRA, L. B. R.; NAKAGAWA, E. Y. An experience report on update of systematic literature reviews. In: *29th International Conference on Software Engineering and Knowledge Engineering. (SEKE '17)*. [S.l.: s.n.], 2017. p. 78+.

SALDAñA, J. *The coding manual for qualitative researchers*. [S.l.]: SAGE London, 2009.

SANTOS, R. E. S.; SILVA, F. Q. B. da. Motivation to perform systematic reviews and their impact on software engineering practice. In: *2013 ACM / IEEE International Symposium on Empirical Software Engineering and Measurement*. [S.l.: s.n.], 2013. p. 292–295.

SEAMAN, C. B. Qualitative methods in empirical studies of software engineering. *IEEE Trans. Softw. Eng.*, v. 25, n. 4, p. 557–572, 1999.

SHEKELLE, P. G.; NEWBERRY, S. J.; WU, H.; SUTTORP, M.; MOTALA, A.; LIM, Y.-W.; BALK, E. M.; CHUNG, M.; YU, W. W.; LEE, J.; GAYLOR, J. M.; MOHER, D.; ANSARI, M. T.; SKIDMORE, R.; GARRITTY, C. *Identifying Signals for Updating Systematic Reviews: A Comparison of Two Methods*. Agency for Healthcare Research and Quality (US), Rockville (MD), 2011. Disponível em: <http://europepmc.org/books/NBK56774>.

SHOJANIA, K. G.; SAMPSON, M.; ANSARI, M. T.; JI, J.; DOUCETTE, S. P.; MOHER, D. How quickly do systematic reviews go out of date? a survival analysis. *Annals of Internal Medicine*, v. 147, p. 224, 2007.

SILVA, F. Q. B. da; SANTOS, A. L. M.; SOARES, S.; FRANCA, A. C. C.; MONTEIRO, C. V. F.; MACIEL, F. F. Six years of systematic literature reviews in software engineering: An updated tertiary study. *Inf. Softw. Technol.*, v. 53, n. 9, p. 899–913, 2011.

SJOBERG, D. I. K.; DYBA, T.; JORGENSEN, M. The future of empirical methods in software engineering research. In: *1st Future of Software Engineering (FOSE '07)*. [S.l.: s.n.], 2007. p. 358–378.

SOCIETY, I. C.; BOURQUE, P.; FAIRLEY, R. E. *Guide to the Software Engineering Body of Knowledge (SWEBOK(R)): Version 3.0*. [S.l.]: IEEE Computer Society Press, 2014.

SOLINGEN, R. van; BASILI, V.; CALDIERA, G.; ROMBACH, H. D. Goal question metric (gqm) approach. In: MARCINIAK, J. J. (Ed.). *Encyclopedia of Software Engineering*. [S.l.]: American Cancer Society, 2002.

SOMMERVILLE, I. *Software Engineering*. [S.l.]: Addison-Wesley Publishing Company, 2011.

STAPLES, M.; NIAZI, M. Experiences using systematic review guidelines. *Journal of Systems and Software*, v. 80, n. 9, p. 1425 – 1437, 2007.

TACCONELLI, E. Systematic reviews: Crd's guidance for undertaking reviews in health care. *The Lancet Infectious Diseases*, v. 10, n. 4, p. 226, 2010.

TAKWOINGI, Y.; HOPEWELL, S.; TOVEY, D.; SUTTON, A. J. A multicomponent decision tool for prioritising the updating of systematic reviews. *BMJ*, v. 347, 2013.

VICENTE-SAEZ, R.; MARTINEZ-FUENTES, C. Open science now: A systematic literature review for an integrated definition. *Journal of Business Research*, v. 88, p. 428 – 436, 2018.

WOHLIN, C. Second-generation systematic literature studies using snowballing. In: *20th International Conference on Evaluation and Assessment in Software Engineering (EASE '16)*. [S.l.: s.n.], 2016. p. 1–6.

APPENDIX A – LIST OF SELECTED PAPERS

ID	Title	Year
S01	Revisiting software ecosystems Research: A longitudinal literature study	2016
S02	Investigations about replication of empirical studies in software engineering: A systematic mapping study	2015
S03	The contribution that empirical studies performed in industry make to the findings of systematic reviews: A tertiary study	2018
S04	Cloud service evaluation method-based Multi-Criteria Decision-Making: A systematic literature review	2018
S05	Is There a Place for Qualitative Studies when Identifying Effort Predictors?: A Case in Web Effort Estimation	2014
S06	Motivation in software engineering: A systematic review update	2011
S07	An extended systematic review of software process improvement in small and medium Web companies	2011
S08	What is the Further Evidence about UML? - A Systematic Literature Review	2017
S09	Systematic literature review on agile practices in global software development	2018
S10	A systematic literature review on the applications of Bayesian networks to predict software quality	2017
S11	An experience report on update of systematic literature reviews	2017
S12	A systematic literature review on cloud computing adoption and migration	2017
S13	Replication of empirical studies in software engineering: An update of a systematic mapping study	2015
S14	Software process simulation modeling: Preliminary results from an updated systematic review	2014
S15	Cross- vs. Within-company cost estimation studies revisited: An extended systematic review	2014
S16	Six years of systematic literature reviews in software engineering: An updated tertiary study	2011

ID	Title	Year
S17	A Survey on Software Release Planning Models	2016
S18	Investigations about replication of empirical studies in Software Engineering: Preliminary Findings from a Mapping Study	2014
S19	Systematic literature reviews in software engineering – A tertiary study	2010
S20	Empirical evidence about the UML: a systematic literature review	2011
S21	Software Process Simulation Modeling: An Extended Systematic Review	2010
S22	Maintainability Prediction of Relational Database-Driven Applications: A Systematic Review	2012

APPENDIX B – DATA EXTRACTION FORM

Question	Answer
Title of publication	Text
Year of publication	Text
List of Authors (Separated by comma)	Text
Country	Text
What kind of systematic approach?	List (SLR, SM, TS)
Did the update change any result from the original study?	List (Yes or No)
What was the elapsed time between the udpate and corresponding original study?(publication)	Text
Have artifacts been changed from previous study?	List (Yes, No and "Answer not found")
What were the changed artifacts? (NA, if none.)	Text
Does the author report difficulties in finding the artifacts of the original study?	List (Yes or No)
What are the reported difficulties? (NA, if none.)	Text
Did the author make available the update artifacts?	List (Yes or No)
Who performed the original study?	List ("Other Research Group", "Same Research Group" and "Answer not found")
Has there been any contact with the original author?	List (Yes, No, "Answer not found" and "Same Research Group")
What are the reasons to update?	Text

APPENDIX C – SURVEY FORM

Systematic Literature Review (SLR) Maintenance Process

Dear Researcher,

You are receiving this form, because you had published a systematic literature review between 2011 and 2015.

This form is part of a Ph.D. research. Your personal information will be kept confidential.

With this questionnaire, we intend to collect the researcher's opinion about SLR maintenance process.

The queries presented next try to elucidate issues about maintenance, traceability, and plagiarism.

Feel free to ask any question about this questionnaire.

Thank´s in advance,

Vilmar Nepomuceno
vsn@cin.ufpe.br
PhD. candidate - Informatic Center at UFPE.

* Required

1. **Email address** *

Maintenance

In software engineering, software maintenance is the modification of existing software while preserving its integrity. P. Bourque and R.E. Fairley, eds., Guide to the Software Engineering Body of Knowledge, Version 3.0, IEEE Computer Society, 2014; www.swebok.org.

In this Section, trying to understand better what would be maintenance in systematic literature reviews, we are asking you to answer the following questions:

2. **Have you already performed a systematic literature review update?** *
   Mark only one oval.
   - ◯ Yes *Skip to question 2.*
   - ◯ No *Skip to question 15.*

Maintenance

In our research, we are considering that SLR's artifacts are the documentation used to elaborate and conduct the SLR, primary studies's result set take from the search procedure, any extraction form used to store data from the primary studies (quality assessment and data extraction steps), and the SLR's reports.

3. **Did you have any issue trying to access the original artifacts?** *
   Mark only one oval.
   - ◯ Yes
   - ◯ No
   - ◯ I did no try.

4. **Related with the previous question. What problems did you find to access the original artifacts?** *
   If you did not have any issues to acess or did not try to access just answer "NA".

5. **Did you change any artifact from the original SLR?** *
   Mark only one oval.
   - ◯ Yes
   - ◯ No

6. **Related with the previous question. What were the artifacts, and why?** *
   If you did not change any artifacts just answer "NA".

7. **Who performed the original SLR?** *
   Mark only one oval.
   - ◯ Your research Group
   - ◯ Other Research Group

8. **If you updated a SLR from another research group, did you contact the original research group?** *
   Mark only one oval.
   - ◯ Yes
   - ◯ No
   - ◯ N.A.

9. **Did you have any support from the original authors?** *
   Mark only one oval.
   - ◯ Yes
   - ◯ No
   - ◯ N.A.

10. **Would you like to keep your published SLR always up-to-date?** *
    Mark only one oval.
    - ◯ Yes
    - ◯ No

11. **Please, explain your answer to the previous question:** *

12. **Would you mind if other reasearchers update your SLR?** *
Mark only one oval.
- Yes
- No

13. **Which differences from the original SLR will result a new SLR instead of an update?** *
Check all that apply.
- [ ] Change the research question.
- [ ] Change the search string.
- [ ] Change the search sources.
- [ ] Use a different time frame in the search procedure.
- [ ] Change the data extraction form.
- [ ] Change the evidence found.
- [ ] Other:

14. **One possible proposed "SLR maintenance" definition could be: "Modifying an existing SLR by altering some artifact, while reusing its original questions, keeping all changes and impacts caused by such changes tracked"** *
Mark only one oval.

	1	2	3	4	5	
I disagree						I Agree

15. **What would you change in the previous definition?** *

Skip to question 22.

Maintenance

In our research, we are considering that SLR's artifacts are the documentation used to elaborate and conduct the SLR, primary studies's result set take from the search procedure, any extraction form used to store data from the primary studies (quality assessment and data extraction steps), and the SLR's reports.

16. **If you are updating a SLR from another research group, would you consider contact the original research group?** *
Mark only one oval.

- ◯ Yes
- ◯ No

17. **Would you like to keep your published SLR always up-to-date?** *
Mark only one oval.

- ◯ Yes
- ◯ No

18. **Please, explain your answer to the previous question:**

_____

_____

_____

_____

_____

19. **Would you mind if other reasearchers update your SLR?** *
Mark only one oval.

- ◯ Yes
- ◯ No

20. **Which differences from the original SLR will result a new SLR instead of an update?** *
Check all that apply.

- ☐ Change the research question.
- ☐ Change the search string.
- ☐ Change the search sources.
- ☐ Use a different time frame in the search procedure.
- ☐ Change the data extraction form.
- ☐ Change the evidence found.
- ☐ Other: _____

21. **One possible proposed "SLR maintenance" definition could be: "Modifying an existing SLR by altering some artifact, while reusing its original questions, keeping all changes and impacts caused by such changes tracked"** *
Mark only one oval.

	1	2	3	4	5	
I disagree	◯	◯	◯	◯	◯	I Agree

22. **What would you change in the previous definition?** *

Traceability (Versioning)

In software engineering, versioning is the management of multiple deliveries of a product, all deliveries have the same general function but with improvements.

In this section, we intend to analyze which/how artifacts of a SLR should be versioned.

23. **Should SLRs be versioned using tools such as Git or SVN?** *

Mark only one oval.

- ◯ Yes
- ◯ No

24. **Please, explain your answer to the previous question:** *

25. **What SLR artifacts do you consider essential to be versioned?** *

Check all that apply.

- ☐ Full protocol
- ☐ Search procedure Info (search string, search sources, search history)
- ☐ Search results
- ☐ Primary study selection criteria
- ☐ Quality evaluation form
- ☐ Data extraction form
- ☐ Status Report
- ☐ Other:

26. **Would you keep your SLR artifacts in a common repository such as Github?** *

Mark only one oval.

- ◯ Yes
- ◯ No

27. **Please, explain your answer to the previous question:** *

Plagiarism

The word plagiarism is derived from Latin. “Plagiare means to kidnap.”

"In simple words, plagiarism is the use of others ideas or work without any credit to the original authors. In other words, taking credit for others work whether intentionally or unintentionally." P. Mohan Kumar et al. in Knowing and Avoiding Plagiarism During Scientific Writing. (2014).

28. **What would you consider plagiarism in SLR?** *

29. **Are SLR replications* plagiarism? Why?** *

* "An exact replication is one in which the procedures of an experiment are followed as closely as possible to determine whether the same results can be obtained" Forrest J. Shull, Jeffrey C. Carver, Sira Vegas, and Natalia Juristo. 2008. The role of replications in Empirical Software Engineering. Empirical Softw. Engg. 13, 2 (April 2008), 211-218.

30. **A researcher is updating an SLR, reusing the original protocol and reporting the same evidence, even with a different set of primary studies. Is the researcher commiting plagiarism? Assume the researcher references the original work properly when reporting the results.** *

Mark only one oval.

- ◯ Yes
- ◯ No

Thank You!

31. **Are you willing to attend an interview about the subject presented in this questionnaire?** *

Mark only one oval.

- ◯ Yes
- ◯ No

APPENDIX D – E-MAIL TO INTERVIEWEES

Subject: Short interview invitation for a PhD Research

Body:

Dear XXXXXXXXX,

My name is Vilmar Nepomuceno, I am a PhD student at Federal University of Pernambuco - Brazil, advised by Prof. Sérgio Soares.

My PhD research addresses maintenance, traceability, and plagiarism issues in systematic literature reviews.

WHY YOU HAVE BEEN CONTACTED?

You are receiving this e-mail because in our research you were identified as a major contributor in software engineering systematic reviews publications.

THE IDEA

To interview experienced researchers about their opinion on maintaining systematic literature reviews. The interview can be made by Skype, hangout, or any other meeting tool of your preference.

HOW MUCH TIME WILL IT TAKE

The interview has few questions to be answered, nine in total. We will need about 45 minutes to complete the interview.

WHAT DO YOU HAVE TO KNOW TO PARTICIPATE?

We are sending you an attached file with some information that will be used during the interview, any questioning about the material can be done before the interview by email to vsn@cin.ufpe.br.

I will be very grateful if you can give us your opinion about the issues placed in the interview.

I'm looking forward to schedule our interview. Please, let me know if you have any questions and, if you are willing to help, when we could talk.

Thanks in advance.

Best Regards,

APPENDIX E – E-MAIL TO INTERVIEWEES ATTACHMENT

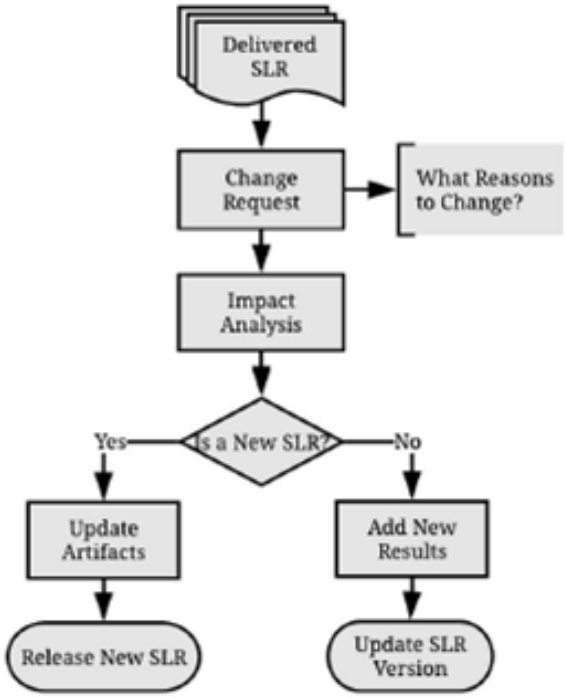
Software Maintenance Definition:

"Software maintenance is the general process of changing the system after it has been delivered"

Reasons to maintain SLRs:

- Update Evidences — Keep the evidence up-to-date to track the evolution of a particular area;
- Methodological Issues — Any research is subject to methodological issues, so it would be interesting to fix these problems without necessarily realize a new SLR;
- Improve Reliability — A SLR that can be reviewed, improve and compared to others SLRs makes the results more robust and reliable.

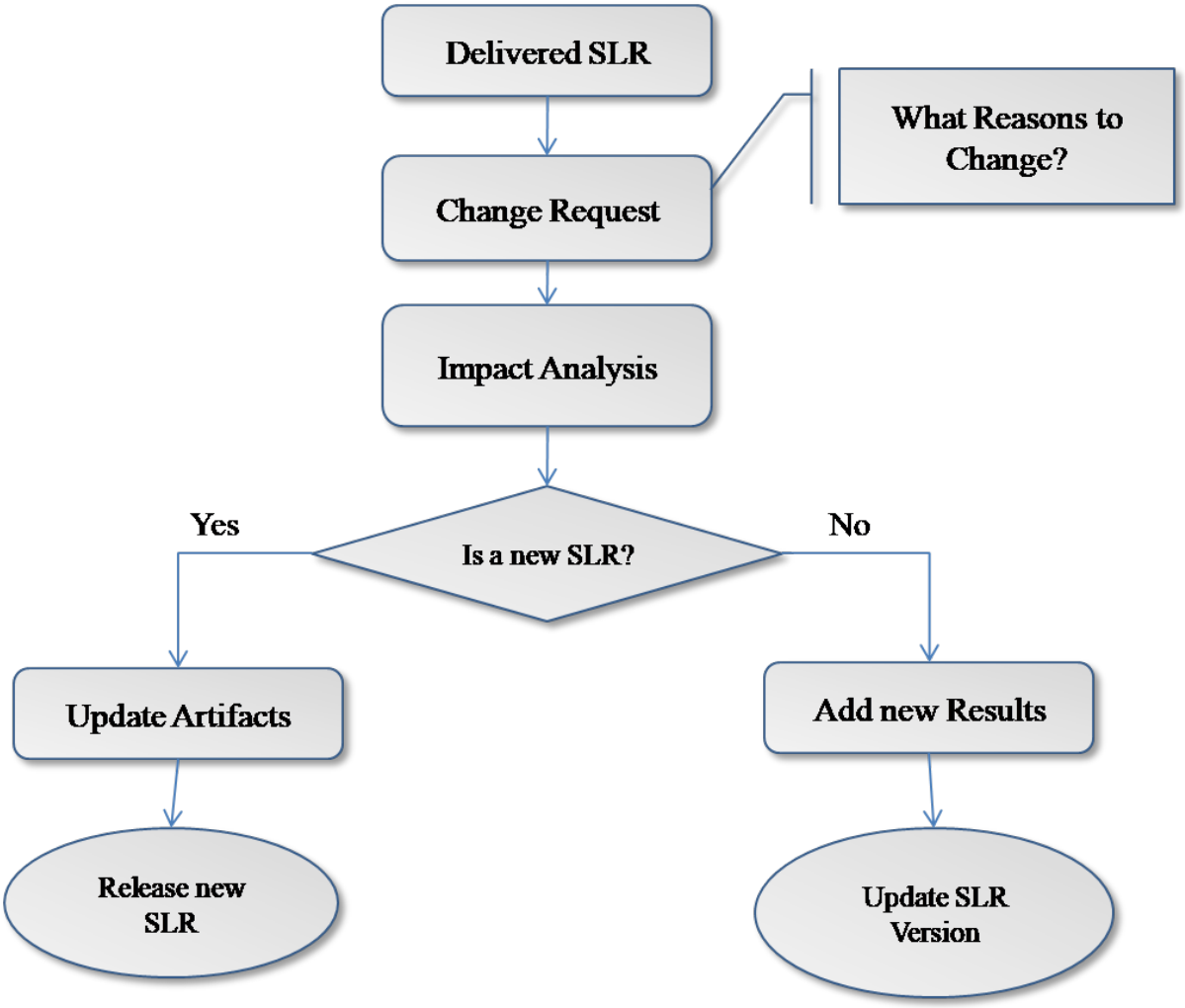
SLR maintenance process:



- Delivered SLR — The latest SLR version being maintained;
- Update Request—Starting a newmaintenance process, based on the presented reasons (Update Evidences, Methodological Issues, or Improve Reliability);
- Impact Analysis — The cost and impact of these changes are assessed to see what the gains are, and once verified the need, a new SLR release is conducted. At this point will need to make the decision to just upgrade the SLR to a newer version, or the changes' impact will lead to a new SLR;
- Update Artifacts — During the analysis phase, it can be concluded that the requested changes caused significant updates in the original artifacts, provoking the need to create a new SLR, which is based on the original SLR;
    - Release New SLR — A new SLR based on the original;
- Add New Results — During the analysis phase, it can be concluded that the requested changes caused minor updates in the original artifacts, meaning it is only a maintenance of the original SLR;
    - Update SLR Version — A new version from original SLR.

APPENDIX F – GUIDELINES ORIGINAL VERSION

Figure 19 – Guidelines to perform SLRs maintenance (version presented to interviewed researchers).



Source: The author (2019).

F.1 GUIDELINES APPLICATION EXAMPLE

Based on systematic mapping conducted in this work (Chapter 2.4).

F.1.1 Change Request

At this moment we check the reasons why a maintenance activity can be performed.

- Update findings: Check for new studies on SLR updates and add to existing results;
- Methodological issue: Correcting study threats to validity;

Lack of term "Secondary Study" in search string;

First selection process was performed only for one person, the SM conductor;

Data extraction process was performed only for one person, SM conductor.

- Improve reliability: No work was found allowing comparison;
- SLR extension: Check if primary studies report the decision process to perform an update and compare with the defined process;
- Domain search: Not currently applicable.

F.1.2 Impact Analysis

Based on impact analysis decision-making process and established reasons, we need to analyse:

- If there are new findings;
- If methodological issue may have impacted final report;
- If proposed SM extension is important to SE community.

If analysis has a positive result the maintenance should be performed, if not, the maintenance should be re-analysed at a future time.

F.1.3 Is a new SLR?

If the maintenance activity is:

- Update findings: **Add new results**
- Methodological issue:

  Lack of term "Secondary Study" in search string: Verify search string change had an impact on data vision, if a new vision is established, **Update vision**.

  First selection process was performed only for one person, the SM conductor. **Add new results**, because there is no change in the artifacts;

  Data extraction process was performed only for one person, SM conductor. **Add new results**, because there is no change in the artifacts;

- SLR extension: **Update vision**, due to new analysis over data that generates a new vision about these data.