Evaluating strategies for forward snowballing application to support secondary studies updates – Emergent Results

Katia Romero Felizardo
Federal University of Technology – Paraná
Cornélio Procópio, PR – Brazil
katiascannavino@utfpr.edu.br

Anderson Y. Iwazaki da Silva
Federal University of Technology – Paraná
Cornélio Procópio, PR – Brazil
anderson.1992@alunos.utfpr.edu.br

Érica Ferreira de Souza
Federal University of Technology – Paraná
Cornélio Procópio, PR – Brazil
ericasouza@utfpr.edu.br

Nandamudi L. Vijaykumar
National Institute for Space Research
Laboratory for Computing and Applied Mathematics (LAC)
São José dos Campos, SP – Brazil
vijay.nl@inpe.br

Elisa Yumi Nakagawa
Institute of Mathematics and Computer Sciences (ICMC/USP)
Department of Computer Systems
São Carlos, SP – Brazil
elisa@icmc.usp.br

ABSTRACT

Context: Secondary studies should be updated from time to time to include new evidence to preserve their value. It is recognized that one search technique to update secondary studies is forward snowballing and that the number of studies identified is dependent on the electronic databases selected. However, there is no consensus on what electronic database is most appropriate for applying forward snowballing. **Objective:** The main goal of this study is to evaluate the use of different electronic databases for applying forward snowballing to update secondary studies. **Method:** Six updates were performed using forward snowballing with support from two electronic databases, one specific (IEEE Xplore) and the other generic (Google Scholar) and three combinations were evaluated to obtain new evidence during secondary studies updating: (1) searching using Google Scholar as electronic database; (2) searching using IEEE Xplore as electronic database; and (3) searching using both, IEEE Xplore and Google Scholar as complementary electronic databases. **Results:** The use of a specific electronic database is not indicated for forward snowballing application to update SLRs, since many relevant studies may not be identified. However, the use of a generic database is sufficient to discover the majority of the studies. **Conclusions:** The emergent contribution of our work to the body of knowledge in the SLR field is to add empirical evidence regarding the use of different electronic databases to support forward snowballing application during secondary studies updates. These results should help reviewers when they decide to find evidences to update their SLRs.

SBES 2018, September 17–21, 2018, Sao Carlos, Brazil

ACM ISBN 978-1-4503-6503-1/18/09…$15.00
https://doi.org/10.1145/3266237.3266240

CCS CONCEPTS

• **Information systems** → *Digital libraries and archives*;

KEYWORDS

Systematic Literature Review, Secondary Studies, Forward Snowballing, Update, Search Evidence

ACM Reference Format:
Katia Romero Felizardo, Anderson Y. Iwazaki da Silva, Érica Ferreira de Souza, Nandamudi L. Vijaykumar, and Elisa Yumi Nakagawa. 2018. Evaluating strategies for forward snowballing application to support secondary studies updates – Emergent Results. In *XXXII BRAZILIAN SYMPOSIUM ON SOFTWARE ENGINEERING (SBES 2018), September 17–21, 2018, Sao Carlos, Brazil,* Editor 1, Editor 2, and Editor 3 (Eds.). ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3266237.3266240

1 INTRODUCTION

Secondary studies, including Systematic Literature Reviews (SLR) also known as Systematic Reviews (SR) and Systematic Mappings (SM) identify and summarize research evidence on several research topics in Software Engineering (SE), providing a complete and fair evaluation of the state-of-the-art of all relevant research available for a specific topic of interest [8].

The update of secondary studies is quite an important issue in SE since secondary studies that are not maintained (i.e., not updated) might become outdated or misleading. Secondary studies should be frequently updated with the purpose of identifying new evidence that has emerged after the completion of a review. An updated SLR has similar research question, objectives and inclusion criteria to the previous review. However, the main element of an update is the effort to identify new evidence. The new evidence can be identified for example, using the search strategy from the original review or adding a new database. Therefore, an update can include new data, new methods, or new analyses (adjusting the findings and conclusions as appropriate) [10]. Incorporating new evidence into existing secondary studies is therefore paramount in order to sustain their relevance.

Updating a secondary study is an arduous task because traditionally the search for new evidence involves the re-execution of search strings used in previous versions of the review. Other search techniques could be adopted, such as manual search or forward snowballing [6], [14].

The first step of forward snowballing involves the identification of a set of studies, defined as seed set [1]. The issue of applying the snowballing technique is to create this seed set, which is not an easy task. However, in the context of updating secondary studies the seed set can be formed by studies included in the previous version of the review. After defining the seed set, researchers analyze studies that have cited those belonging to the seed set to obtain further relevant studies. New evidence will certainly cite studies already included in earlier versions of a secondary study. It is worth mentioning that the citations are identified in electronic databases and the choice of these databases can impact the number of studies identified. One of the premises of secondary studies is to obtain the best evidence-based answer, so the loss of a study end up being a bias to the rigorous review process.

Different authors [6], [14] have investigated the use of forward snowballing to update secondary studies. They agree that the use of forward snowballing has a high precision (detection of relevant studies) and it also influences on reducing the effort (number of studies to be analyzed – time). The authors also alert that the risk of missing relevant papers should not be underrated. In this context, there are several questions with respect to the use of forward snowballing that are not extensively explored. For example, Wohlin [14] and Felizardo et al. [6] have separately updated an SLR on the topic of cross vs. within-company effort estimation using forward snowballing. While Wohlin chose Google Scholar to support applying the technique, Felizardo et al. made use of IEEE Xplore and ACM to identify citations. In this context, an issue that remains is "*What would be the most appropriate electronic database for forward snowballing application: searching using a generic database (Scopus, Google Scholar, etc.) a specific electronic database (IEEE Xplore, ACM Digital Library, etc.) or multiple specific databases (which ones?) unlike a generic database?*".

Although some authors have suggested methodologies to select databases for conducting SLRs [2], approaches, guidelines [11] and tools [12] for assisting the review process of search strategies used in SLRs, there is no consensus on what electronic database is the most appropriate to support employing forward snowballing to update secondary studies. The use of a specific database, such as IEEE Xplore or ACM Digital Library may lead to a lack of evidence. However, the adoption of generic databases, such as Scopus or Google Scholar may generate a large number of studies (not necessarily relevant) to be analyzed. Information that may assist in this decision is of fundamental importance.

Therefore, the goal of the research detailed herein is to evaluate the use of different electronic databases to support employing forward snowballing in updating secondary studies. This evaluation is assessed by means of updating three published secondary studies [3, 4, 13].

The remainder of this paper is organized as follows. In Section 2 we detailed the research methodology applied to evaluate electronic databases to identify citations for forward snowballing application, followed by the detailing of results. Finally, Section 3 concludes our work.

2 MULTIPLE CONDUCTIONS OF SYSTEMATIC REVIEW UPDATES TO EVALUATE ELECTRONIC DATABASES FOR FORWARD SNOWBALLING APPLICATION

The first step to apply forward snowballing is to identify the seed set, i.e., the initial set of studies. Felizardo et al. [6] recommend that the seed set could be formed by the primary studies included in a previous version of an SLR and the SLR itself. Following, for each one of the studies belonging to the seed set, reviewer identifies studies that cite it. The citations could be found in electronic databases, such as IEEE Xplore, Google Scholar, etc. Generally, the list of citations is provided automatically by these databases. The studies in this list should be classified as included or excluded, based on the inclusion and exclusion criteria, respectively. The challenge in applying forward snowballing is to define which electronic database should be used to identify the citations list. The choice of different databases can generate different lists of citations.

Our multiple conductions investigate the use of three combinations for forward snowballing to support looking for new evidence to update secondary studies in SE. The analyzed combinations are as follows:

(1) **Combination 1:** *Searching using Google Scholar as electronic database to obtain new evidence during secondary studies updating using forward snowballing*;
(2) **Combination 2:** *Searching using IEEE Xplore as electronic database to obtain new evidence during secondary studies updating using forward snowballing*; and
(3) **Combination 3:** *Searching using both, IEEE Xplore and Google Scholar as complementary electronic databases to obtain new evidence during secondary studies updating using forward snowballing.*

2.1 Instrumentation and Procedure

Multiple conductions were executed using: one search technique (forward snowballing); two electronic databases (IEEE Xplore and Google Scholar); and three secondary studies (two SLR and one SM).

A total of six updates were performed, as shown in Table 1.

Initially a secondary study was selected to be updated. The studies included in the previous version of this study and the study itself formed the seed set. In the sequence, an electronic database was selected to assist in the task of identifying studies that cite the studies contained in the seed set. The identified studies were analyzed by a researcher in two stages, initially reading the title/abstract followed by reading the full text, through the application of inclusion and exclusion criteria. For the studies selected as relevant a new search was performed.

Note that there are other databases (e.g. Scopus, Compendex, ScienceDirect, Springer Link, etc.) that were not selected because IEEE Xplore and Google Scholar were the databases chosen by Wohlin [14] and Felizardo [6], respectively, facilitating a comparison between the results of these studies. Moreover, Kitchenham

Table 1: Multiple Conductions Project

Round	Technique	Electronic Database	Secondary Study
Update 1	Forward Snowballing	Google Scholar	SLR [3] – S1
Update 2	Forward Snowballing	IEEE Xplore	SLR [3] – S1
Update 3	Forward Snowballing	Google Scholar	SM [4] – S2
Update 4	Forward Snowballing	IEEE Xplore	SM [4] – S2
Update 5	Forward Snowballing	Google Scholar	SLR [13] – S3
Update 6	Forward Snowballing	IEEE Xplore	SLR [13] – S3

and Brereton [7] identified the electronic databases IEEExplore, ACM Digital Library, Google Scholar, Citeseer library, INSPEC, ScienceDirect and EI Compendex as the most relevant sources to SE. In a further work, the evaluation will be expanded to other databases.

We chose three secondary studies to be updated using forward snowballing: (1) an SLR on ontologies in software testing [3], manually conducted and published in the literature; an SM on knowledge management initiatives in software testing [4], manually conducted and published in the literature; an SLR on towards cross-browser incompatibilities detection [13]; it has not been published yet and was conducted by one of the authors of this work for a bibliographic search of a doctoral research.

These studies were published in 2013 [3], 2015 [4] and one not yet published [13], and will hereafter respectively called Study 1 (S1), Study 2 (S2), and Study 3 (S3). We selected S1, S2 and S3 for the following reasons: (1) they were conducted and double-checked by reviewers with experience in conducting secondary studies; (2) the research topic focus of S1, S2 and S3 represents a very specific area within SE with a small number of studies; therefore, we believed this would be a more controlled context to test our strategies; (3) they have not been recently updated; and (4) the list of included studies were available.

The forward snowballing technique was evaluated using the measures Recall and Precision [5]. Recall is the ability of a search technique to obtain all relevant studies. Since we cannot guarantee that the results from a secondary study include ALL the relevant studies, true recall (also called sensitivity) cannot be calculated. An alternative is to calculate the Relative Recall. Therefore, we considered the overall existing set of relevant studies as the sum of unique relevant studies identified in our update efforts by both sources: (i) searching using Google Scholar as electronic database; and (ii) searching using IEEE Xplore. In the present study, we calculated relative recall (RR) as:

$RR\_GoogleScholar = \frac{a}{b}$ where:
a = set of relevant studies retrieved by Google Scholar;
b = set of relevant studies retrieved by both sources.

$RR\_IEEE\ Xplore = \frac{a}{b}$ where:
a = set of relevant studies retrieved by IEEE Xplore;
b = set of relevant studies retrieved by both sources.

The precision of a source is the amount of relevant studies amongst the studies retrieved by the source, i.e., the ability to detect no or few irrelevant studies. In the present study, we calculated precision (P) as:

$P\_GoogleScholar = \frac{a}{b}$ where:
a = set of relevant studies retrieved by Google Scholar;
b = set of all studies retrieved by Google Scholar.

$P\_IEEE\ Xplore = \frac{a}{b}$ where:
a = set of relevant studies retrieved by IEEE Xplore;
b = set of all studies retrieved by IEEE Xplore.

2.2 Conduction and Preliminary Results

This section details the steps and results of our multiples conductions.

Our first step was to obtain the list of relevant studies to be included in three secondary studies selected by us, and to use it as our seed set (starting set). The seed set of S1 [3] was composed of 19 studies. We also considered S1 itself to compose the seed set, as recommended by Felizardo et al. [6]. 16 studies formed the seed set of S2 [4]. Similarly to S1, S2 also composed its seed set. S3 [13] had a seed set containing 17 studies. Unlike S1 and S2, S3 did not compose its seed set, because it has not been published yet.

We have updated three studies (S1, S2 and S3) and each study was updated twice, i.e., the citations were extracted with the help of two different electronic databases, Google Scholar and IEEE Xplore. Results of these six updates are summarized in Table 2. For example, the update of Study S1 using Google Scholar was executed in three rounds (see Table 2, line 2, column 3). During the first round 393 studies were identified, being, 111 duplicated, 36 included and 357 excluded. In the second round, through the 36 studies included in the last round a total of 295 new studies were obtained being, 113 duplicated, nine included, 173 excluded and eight studies have not been cited. Finally, in the last round, 61 studies were retrieved being, 36 duplicated, zero included, 25 excluded and five studies have not been cited.

Table 2 shows an overview of studies analyzed, included and lost (studies not cited). Considering the six updates we analyzed, the three combinations are detailed as follows.

Combination 1: *Searching using Google Scholar as electronic database to obtain new evidence during secondary studies updating using forward snowballing.*

Updating S1 using forward snowballing and Google Scholar involved the reading of 749 studies, out of which 45 were included (see Figure 1 – Update 1). With respect to S2, using Google Scholar, 355 candidate studies were analyzed and 5.8% (28 – see Figure 1, Update 3) were included. Using Google Scholar to update S3, 441 studies were read and 25 included. In general, the use of generic

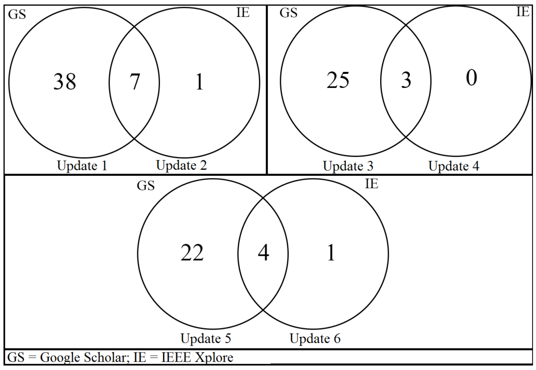
Table 2: Results of updates

Metrics	Update 1 – S1/GS	Update 2 – S1/IE	Update 3 – S2/GS	Update 4 – S2/IE	Update 5 – S3/GS	Update 6 – S3/IE
Rounds (R)	3 (R1, R2, R3)	3 (R1, R2, R3)	3 (R1, R2, R3)	2 (R1, R2)	3 (R1, R2, R3)	2 (R1, R2)
Studies	749 (393+295+61)	109 (101+8+0)	355 (239+113+3)	12 (10+2)	411 (380+59+2)	66 (63+3)
Included	45 (36+9+0)	8 (7+1+0)	28 (24+4+0)	3 (3+0)	25 (24+1+0)	5 (5+0)
Excluded	464 (246+173+25)	67 (62+5+0)	244 (163+80+1)	6 (5+1)	272 (158+45+2)	29 (28+1)
Duplicated	260 (111+113+36)	34 (32+2+0)	82 (52+29+1)	3 (2+1)	210 (197+13+0)	32 (30+2)
Not cited	13 (0+8+5)	4 (1+2+1)	9 (0+8+1)	5 (3+2)	12 (2+10+0)	6 (3+3)
GS = Google Scholar; IE = IEEE Xplore						

Table 3: Updates: Relative Recall and Precision

Update – Study	Relative Recall	Precision
Update 1 – S1/GS	45 ÷ 46 = 97%	45 ÷ 749 = 6%
Update 2 – S1/IE	8 ÷ 46 = 17%	8 ÷ 109 = 7%
Update 3 – S2/GS	28 ÷ 28 = 100%	28 ÷ 355 = 7%
Update 4 – S2/IE	3 ÷ 28 = 10%	3 ÷ 12 = 25%
Update 5 – S3/GS	25 ÷ 26 = 96%	25 ÷ 441 = 5%
Update 6 – S3/IE	5 ÷ 26 = 19%	5 ÷ 66 = 7%
GS = Google Scholar; IE = IEEE Xplore		

Figure 1: Intersection of Results



databases leads to reading of a large number of studies, not necessarily relevant. For example, in Update 5 – S3, Google Scholar, 441 studies were read and only 26 were selected (5% of precision).

As shown in Table 3, RR of Google Scholar ranged from 96% to 100% (S1, Update 1, RR: 97%, S2, Update 3, RR: 100%, and S3, Update 5, RR: 96%), showing comprehensiveness in searching for evidence.

Combination 2: *Searching using IEEE Xplore as electronic database to obtain new evidence during secondary studies updating using forward snowballing.*

In order to update S1 with IEEE Xplore, 109 studies were read, 101 excluded, 8 included (see Figure 1 – Update 2) and 4 were not found (lost), not allowing identification of new evidence. Using IEEE Xplore only 14.55% of candidate studies were identified and 17.77% of the evidence was included. Although only eight studies were included, IEEE Xplore revealed one relevant study, which was not identified using Google Scholar. With respect to S2, using IEEE Xplore, 12 candidate studies were analyzed and 10.7% (3 of 28) were included. Using IEEE Xplore to update S3, 66 were read (14.96%) and five were included (see Figure 1 – Update 6).

Six studies were not located in this electronic database. However, one of the included studies was identified exclusively by IEEE Xplore. One plausible explanation is that although Google Scholar indexes IEEE Xplore, there may be a delay in making the papers available in generic databases. In addition, the Google Scholar coverage and its sources are unclear.

Considering P, IEEE Xplore and Google Scholar presented similar results (variation of 5% to 7%) – see Table 3. One exception was Update 4 (P = 25%), which out of the 12 studies identified by IEEE Xplore, only three were included.

RR of IEEE Xplore (see Table 3) ranged from 10% to 19%. RR of the three updates using IEEE Xplore (S1, Update 2, RR: 17%, S2, Update 4, RR: 10% and S3, Update 6, RR: 19%) was lower than the RR of the other three updates using Google Scholar (S1, Update 1, RR: 97%, S2, Update 3, RR: 100%, and S3, Update 5, RR: 96%).

Combination 3: *Searching using both, IEEE Xplore and Google Scholar as complementary electronic databases to obtain new evidence during secondary studies updating using forward snowballing.*

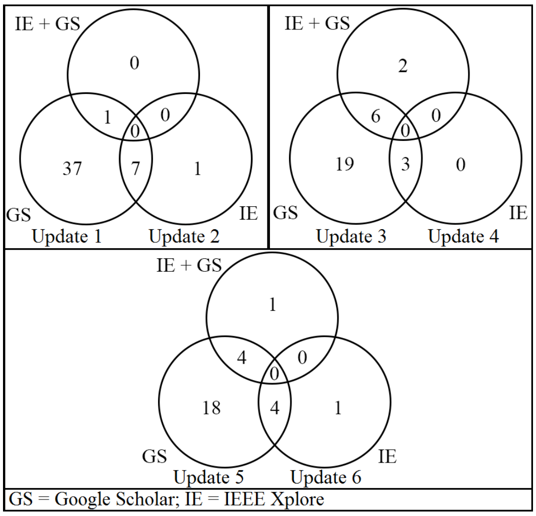
IEEE Xplore precision would be increased if there were no lost evidence. For example, during Update 2, S1, which composed its seed set, was not found in IEEE Xplore because it was indexed in another database and four studies were not identified. Update 4 lost nine studies while Update 6 lost six. In this context, we re-executed these three updates considering the lost studies as new seed sets and Google Scholar as search source.

The results of these re-executions are demonstrated in Tables 4, 5 and Figure 2. For example, the re-execution of Update 2 was performed in two rounds (see Table 4, line 3, column 2). In the first round, three studies were found being, one included, two excluded and 7 studies have not been cited by any other study (zero citations). During the second round, the single study included in the previous round was not cited by no other study.

Using Google Scholar to update S1, 45 studies were included (P = 6%). Eight relevant studies were detected using IEEE Xplore (P = 7%). Finally, adding Google Scholar to locate studies lost by IEEE Xplore, one more study would have been found (P would increase from 7% to 8%). Regarding S2, 28 studies were included by Google Scholar (P = 7%) and 3 by IEEE Xplore (P = 25%). Google Scholar as a complementary database to IEEE Xplore (re-execution of Update 4) ensured the detection of two new studies (P = 9%). Five studies (P = 7%) were included using IEEE Xplore to update S3, one of which had not been detected by Google Scholar, which included 26 studies

(P = 5%). During the re-execution of Update 6 (Precision = 9%), a new study, detected exclusively during re-execution, was revealed.

Figure 2: Re-executions: Intersection of Results



We observed that advantages in adding a complementary database to locate non-indexed studies by a specific database include: (i) precision increaser – it was observed in two of the three studies analyzed; (ii) discovery of new studies – in two re-executions new studies (not previously detected) were discovered.

2.3 Discussions

Felizardo et al. [6] and Wohlin [14] have updated, using forward snowballing, the SLR initially conducted by Kitchenham et al. [9], which was published in 2007. Felizardo et al. [6] performed four iterations looking at the citations that were extracted with the help of electronic databases, such as IEEE Xplore and ACM Digital Library. Forward snowballing achieved 7.5% as P and 92.8% as RR. Additional studies were identified, however, one study was missed (it was not retrieved). They concluded that forward snowballing reduces effort of updating SLRs, when comparing with search strings, but exposes the risk of missing relevant studies.

Similar to Felizardo et al. [6], re-executions of Updates 2, 4 and 6 also adopted multiple electronic databases. Comparing the results, precisions were similar, ranging from 7.5% to 9 %. However, re-executions RRs (19%, 36% and 35%) were lower than the RR reached by Felizardo et al. [6] (92.8%). One explanation is that re-executions seed set did not contain all studies included in the previous version of the SLR and the SLR itself. We reinforce Felizardo et al. [6] recommendation that the seed set could be formed by the primary studies included in a previous version of an SLR and the SLR itself.

Wohlin [14] used the citations provided by Google Scholar to identify possible additional studies. Only one iteration was executed. The author concluded that snowballing found all studies and some additional ones. Comparing results from Updates 1, 3 and 5, which used Google Scholar to obtain the citations, with Wohlin results, it is possible to confirm that Google Scholar revealed new studies, however, in our study, unlike Wohlin's study, Google Scholar did not identify all studies; two were not found.

One issue of using a single database is that none of them is able to return all relevant studies in the context of a secondary study. Choosing only one database can result in loss of evidence. Combining databases would prevent the loss of evidence.

Some lessons learned from our preliminary results are:

(1) relevant studies are contained in both large databases (e.g., Google Scholar, Scopus, etc.) and smaller specific datasets focused on specific subject areas (e.g, IEEE Xplore). An interesting aspect is that the scope of SLR might lead to select one database because a specific topic of interest might be covered better on it. This issue will be investigated by us;
(2) When using a specific database (e.g., IEEE Xplore) there is a risk in missing a lot of relevant studies;
(3) The use of a generic database (e.g. Google Scholar) is sufficient to discover the majority of the studies, however, there is no guarantee that all relevant studies will be detected;
(4) Google Scholar presents a high number of citations for a given study. A feature that may contribute to the high number of citations found by Google Scholar is that it includes duplicates. For example, a citation published in two different sources, such as preprint and journal paper, will be counted as two citations. We also noted that Google Scholar sometimes includes non-scholarly citations, e.g., editorial notes; and
(5) The citations storage is a tedious and time-consuming task. Despite this, none of the electronic databases investigated by us have provided support to automate it. A potential future work for the academic community is to automate the citation storage (download of citations) for snowballing application.

The main limitation of this research is that our analysis was based on applying forward snowballing to update secondary studies only using two electronic databases. However, we selected two different types of databases, Google Scholar and IEEE Xplore. Our results have provided some valuable insight and indication as to how much specific and generic electronic databases can support forward snowballing application to update secondary studies. Thus, as a first-cut assessment, we believe our research met its goal. Other threats to validity are described below.

- Internal validity. It is important to highlight that we strictly followed the guidelines for conducting snowballing suggested by Wohlin [14].
- Construct validity. In our multiple conductions, the study inclusion criteria can be considered a potential confusing factors; however, the studies were independently selected by three experienced researchers. Discussion meetings were also held where the judgments (inclusion or exclusion of a study) were reviewed by other two reviewers. This step was carried out as part of each round.
- Reliability. The citations were extracted with the help of electronic databases such as Google Scholar and IEEE Xplore.
- External validity. The citations were extracted from two different databases (Google Scholar – generic database and IEEE Xplore – specific database). Other databases, e.g., Scopus, ACM, among others, were not analyzed. However, the

Table 4: Results of re-execution

Metrics	Re-execution Update 2 – S1	Re-execution Update 4 – S2	Re-execution Update 6 – S3
Seed set	8	9	5
Rounds (R)	2 (R1, R2)	2 (R1, R2)	3 (R1, R2, R3)
Studies	3 (R1:3+R2:0)	109 (R1:67+R2:42)	35 (R1:32+R2:2+R3:1)
Included	1(R1:1+R2:0)	8 (R1:8+R2:0)	5 (R1:4+R2:1+R3:0)
Excluded	2 (R1:2+R2:0)	99 (R1:57+R2:42)	25 (R1:25+R2:0+R3:0)
Duplicated	0 (R1:0+R2:0)	2 (R1:2+R2:0)	5 (R1:3+R2:1+R3:1)
Not cited	7 (R1:6+R2:1)	6 (R1:3+R2:3)	5 (R1:2+R2:2+R3:1)
New studies	0	2	1
The re-executions were performed using both Google Scholar and IEEE Xplore			

Table 5: Final Results

Combination 1 – Google Scholar	**Combination 2 – IEEE Xplore**	**Combination 3 – IEEE Xplore + Google Scholar**
Update 1 – S1; Included=45; RR=97%; P=6%	**Update 2 – S1**; Included=(7+1*)=8; RR=17%; P=7%	**Re-execution of Update 2 – S1**; Included=(7+1*)+1=9; RR=19%; P=8%
Update 3 – S2; Included=28; RR=100%; P=7%	**Update 4 – S2**; Included=3; RR=10%; P=25%	**Re-executation of Update 4 – S2**; Included=3+(6+2**)=11; RR=36%; P=9%
Update 5 – S3; Included=26; RR=96%; P=5%	**Update 6 – S3**; Included=(4+1*)=5; RR=18%; P=7%	**Re-executation of Update 6 – S3**; Included=(4+1*)+(4+1**)=10; RR=35%; P=9%
* *New study identify exclusively by IEEE Xplore*		
** *New study identify exclusively during the update re-execution*		

study findings are not generalizable and replications on other sources to reinforce our preliminary indications are required. A relevant future work would be evaluate multiple databases in opposite to Google Scholar to find out if its worthy using Google Scholar or a mix of databases (which one?). Moreover, other databases, such as, Scopus or Compendex, maybe provide higher precision and recall rates. We will continue the evaluation of other combinations by conducting more updates conduction.

3 CONCLUSIONS

The emergent contributions of our work to the body of knowledge in SLR are the following: (i) evaluation of two different electronic databases to assist applying forward snowballing to update secondary studies; and (ii) empirical evidence regarding the effects of the use of different electronic databases to support employing forward snowballing in updating secondary studies context.

ACKNOWLEDGMENTS

We wish to express our gratitude to Dr. Willian M. Watanabe for his support and valuable insights and suggestions to our research.

The authors thank the financial support received from CNPq (project 401033/2016-3), Brazil.

REFERENCES

[1] D. Badampudi, C. Wohlin, and K. Petersen. 2015. Experiences from using snowballing and database searches in systematic literature studies. In 19^{th} *Int. Conference on Evaluation and Assessment in Software Engineering (EASE'15)*. 17:1–17:10.

[2] M. Chen, A. Babar, and H. Zhang. 2010. Towards an Evidence-based Understanding of Electronic Data Sources. In 14^{th} *Int. Conference on Evaluation and Assessment in Software Engineering (EASE'10)*. 135–138.

[3] E.F. de Souza, R.A. Falbo, and N.L. Vijaykumara. 2013. Ontologies in software testing: A Systematic Literature Review. In *CEUR Workshop Proceedings*. 71–82.

[4] E.F. de Souza, R.A. Falbo, and N.L. Vijaykumara. 2015. Knowledge management initiatives in software testing: A mapping study. In *Information and Software Technology*, Vol. 57. 378–391.

[5] K. Dickersin, R. Scherer, and C. Lefebvre. 1994. Systematic reviews: Identifying relevant studies for systematic reviews. *British Medical Journal* 309, 6964 (1994), 1286–1291.

[6] K.R. Felizardo, E. Mendes, M. Kalinowski, E.F. Souza, and N.L. Vijaykumar. 2016. Using Forward Snowballing to update Systematic Reviews in Software Engineering. In 10^{th} *Int. Symposium on Empirical Software Engineering and Measurement (ESEM'16)*. 1–6.

[7] B. Kitchenham and P. Brereton. 2013. A Systematic Review of Systematic Review Process Research in Software Engineering. *Information and Software Technology* 55, 12 (2013), 2049–2075.

[8] B.A. Kitchenham and S. Charters. 2007. *Guidelines for performing Systematic Literature Reviews in Software Engineering*. Technical Report EBSE 2007-001. Keele University and Durham University, UK.

[9] B.A. Kitchenham, E. Mendes, and G.H. Travassos. 2007. Cross versus Within-Company Cost Estimation Studies: A Systematic Review. *IEEE Transactions on Software Engineering* 33, 5 (2007), 316–329.

[10] D. Moher, A. Tsertsvadze, A. Tricco, M. Eccles, J. Grimshaw, M. Sampson, and N. Barrowman. 2008. When and how to update systematic reviews. *Cochrane Database Syst Rev* 1 (2008), MR000023.

[11] E. Mourão, M. Kalinowski, L. Murta, E. Mendes, and C. Wohlin. 2017. Investiganting the use of a hybrid search strategy for systematic review. In 11^{st} *Int. Symposium on Empirical Software Engineering and Measurement (ESEM'17)*.

[12] P. Singh and K. Singh. 2017. Exploring Automatic Search in Digital Libraries: A Caution Guide for Systematic Reviewers. In 21^{st} *Int. Conference on Evaluation and Assessment in Software Engineering (EASE'17)*. 236–241.

[13] Willian M. Watanabe. 2017. Towards Cross-Browser Incompatibilities Detection: a Systematic Literature Review. In *Under Review*.

[14] Claes Wohlin. 2016. Second-generation Systematic Literature Studies Using Snowballing. In 20^{th} *Int. Conference on Evaluation and Assessment in Software Engineering (EASE'16)*. 15:1–15:6.