How to do a Structured Literature Review in computer science

Technical Report · May 2015

CITATIONS
10

READS
5,699

1 author:

Anders Kofod-Petersen
Alexandra Instituttet
82 PUBLICATIONS **660** CITATIONS

SEE PROFILE

How to do a Structured Literature Review in computer science

Anders Kofod-Petersen

Version 0.2

October 8, 2014

Contents

1 Introduction

How to write a reference list

Doing a *systematic literature review* is a formal way of synthesising the information available from available primary studies relevant to a set of research questions. The use of systematic literature reviews have traditionally been widespread primarily in medicine (e.g. the well known *Cochrane reviews* [1]). Unfortunately it has been used to a much lesser extend in computer science (for an example of how to do reviews in software engineering see: [2]). Systematic literature reviews stand apart from, in computer science the more traditional unsystematic surveys by using a strict methodological framework with a set of well defined steps carried out in accordance with a predefined protocol.

Using a systematic literature review is in no way a guarantee of finding all relevant literature in a given area. However, there are several advantages in using it: A systematic literature review can map out existing solutions before a researcher attempts to tackle an area; it helps researchers in avoiding bias in their work; publishing these reviews also benefits the community by allowing others to avoid duplicating the effort; it allows researchers to identify gaps of knowledge; and it highlights the areas where additional research is required.

If a systematic literature review is conducted thoroughly it fulfils the advantages described above and thereby gains scientific value.

This documents attempts to give a short introduction to how to conduct a structured literature review within computer science. The examples used are taken from [3].

2 Structure of a systematic literature review

A systematic review has three main phases: *i*) planning, *ii*) conducting and *iii*) reporting. Each of these phases are divided into several steps.

The first phase involves planning the review and can be broken down into these five steps:

1. Identification of the need for a review
2. Commissioning a review
3. Specifying the research question(s)
4. Developing a review protocol
5. Evaluating the review protocol

This second phase is the actually review of the literature. It consists of five steps:

1. Identification of research
2. Selection of primary studies
3. Study quality assessment
4. Data extraction and monitoring
5. Data synthesis

The last phase deals with how to disseminate the newly acquired knowledge. It consists of three steps:

1. Specifying dissemination strategy
2. Formatting the main report
3. Evaluating the report

3 Performing a structured literature review

3.1 Planning the review

For the purpose of this document we can assume that a need has already been identified (step 1) and that a review has been commissioned (step 2). This description will cover steps 3 and 4 in the planning phase; as step 5 has been included in step 4.

Step 3: Specifying the research question(s)

Attempting a literature review is obviously closely coupled with some specific area and/or problem. Thus, it is assumed that a specific problem ($\mathcal{P}$) is tackled using some specific constraints, methods and/or approaches ($\mathcal{C}$) to develop a system, application or algorithm ($\mathcal{S}$). In computer science we would typically like to know what existing solutions are available, how they compare, what the strength of the evidence is and what implications these solutions have. Writing these points down gives us the following research questions:

RQ1 What are the existing solutions to $\mathcal{P}$?

RQ2 How does the different solutions found by addressing RQ1 compare to each other with respect to $\mathcal{C}$?

RQ3 What is the strength of the evidence in support of the different solutions?

RQ4 What implications will these findings have when creating $\mathcal{S}$?

Step 4: Developing a review protocol

The review protocol is very important as it defines exactly how each step is to be carried out; thus, the work is reproducible. It can be beneficial to create an initial protocol and review the upcoming step whenever a step is concluded. Doing this iteratively then covers the fifth step *evaluating the review protocol*. Appendix **??** gives an example of a complete review protocol.

3.2 Conducting the review

With the developed protocol in hand it is now possible to conduct the review. This phase contains five steps: Identification of research in the literature, selecting the primary studies deemed relevant, evaluate the corpus with respect to the chosen quality parameters, extract the relevant data, and synthesise the data.

Step 1: Identification of research

The goal of this step is to retrieve all the literature relevant to the defined research questions. To do this a search strategy must be defined. This strategy should specify *which sources* to be searched and *how to* search them. The list of sources to be searched will traditionally contain the relevant on-line digital libraries as well as a set of journals and conferences relevant to the area[1] The following list contains the most obvious general computer science archives: ACM digital library, IEEE Xplore, ISI web of knowledge, ScienceDirect, CiteSeer, SpringerLink and Wiley Inter Science. For domain specific sources it is worth contacting experts in the domain who will normally know which are the best.

Once the list is complete the specific search terms can be defined as well as the procedure for searching the sources.

The search strings are formed by grouping *key terms*into groups. Each group contains terms that are either synonyms, different forms of the same word, or terms that have similar or related semantic meaning within the domain. Table 1 exemplifies this approach. The terms

[1]Obviously off-line sources should be searched as well. However, at least in computer science most relevant sources are on-line.

chosen should by closely related to the first research question (what are the existing solutions to $\mathcal{P}$?).

Table 1: Search terms

	Group 1	Group 2	Group 3	Group 4
Term 1	$Synonym_1$	$Synonym_2$	$Synonym_3$	$Synonym_4$
Term 2	$Synonym_1$	$Synonym_2$	$Synonym_3$	$Synonym_4$
Term 3		$Synonym_2$	$Synonym_3$	
Term 4			$Synonym_3$	
Term 5			$Synonym_3$	

Each of the, in this case four groups can be designed to retrieve different sets of the relevant literature. The primary goal is to find the literature that is the intersection of the sets (see Figure 1).

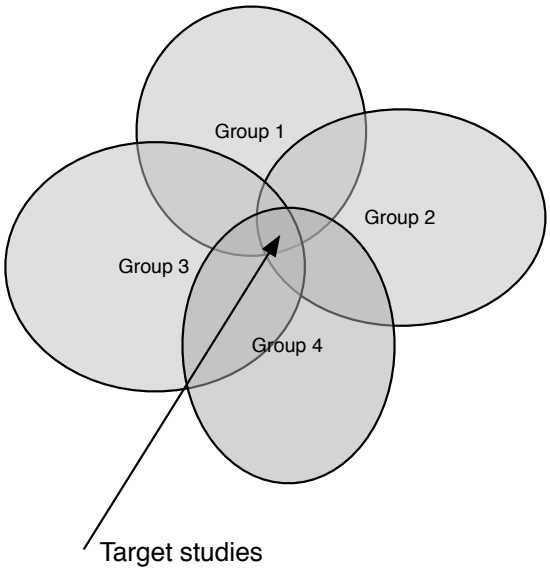

Figure 1: Relevant studies

Implementing this search strategy can be achieved by applying the AND ($\wedge$) and OR ($\vee$), where the OR operator can used within the groups and the AND operator between the groups. Following the example in Table 1 the following search string will capture the structure:

$$([G1, T1] \vee [G1, T2]) \wedge ([G2, T1] \vee [G2, T2] \vee [G2, T3]) \wedge ([G3, T1] \vee [G3, T2] \vee [G3, T3] \vee [G3, T4] \vee [G3, T5]) \wedge ([G4, T1] \vee [G4, T2])$$

The set of papers constructed by applying this search strategy is now ready to go though the selection process.

Step 2: Selection of primary studies

Applying the search described above will most likely return a number of articles far larger than manageable. The relevant articles should now be selected. The protocol should described exactly which criteria should be applied in this selection process. However, some points can be regarded as general and used as removal criteria:

1. Duplicates (keep the highest ranking source),
2. The same study published in different sources (keep the highest ranking source),
3. Studies published before a certain date (or even after).

Applying this selection now leaves us with a set of relevant studies that can now be filtered with respect to quality.

Step 3: Study quality assessment

The purpose of this step is to filter away studies that are not thematically relevant to the area chosen. The protocol should define exactly which inclusion (IC) and quality criteria (QC) are employed (Table 2 gives some examples of criteria).

Table 2: Inclusion and quality criteria

Criteria identification	Criteria
IC 1	The study's main concern is $\mathcal{P}$
IC 2	The study is a primary study presenting empirical results
IC 4	The study focuses on $\mathcal{C}$
IC 5	The study describes an $\mathcal{S}$
QC 1	There is a clear statement of the aim of the research
QC 2	The study is put into context of other studies and research

The criteria can be divided into: primary, secondary and quality screening criteria. In the example described in Table 2 IC 1 and 2 would be the primary; IC 3 and 4 the secondary; and QC 1 and 2 the quality criteria. The criteria can now be applied in a three stage process:

1. Abstract inclusion criteria screening,
2. Full text inclusion criteria screening,
3. Full text quality screening.

Each step should be thoroughly documented as part of the final protocol. Once the set of studies have gone through this process it is (most likely) further reduced and can now go though the next step of detailed quality assessment.

The final quality assessment is done to answer the third research question (*What is the strength of the evidence in support of the different solutions?*). To do this further quality criteria, supplementing QC 1 and 2 in Table 2 should be developed. Examples of this could be (QC 1 and 2 is duplicated as questions for completeness):

QC 1 Is there is a clear statement of the aim of the research?

QC 2 Is the study is put into context of other studies and research?

QC 3 Are system or algorithmic design decisions justified?

QC 4 Is the test data set reproducible?

QC 5 Is the study algorithm reproducible?

QC 6 Is the experimental procedure throughly explained and reproducible?

QC 7 Is it clearly stated in the study which other algorithms the study's algorithm(s) have been compared with?

QC 8 Are the performance metrics used in the study explained and justified?

QC 9 Are the test results thoroughly analysed?

QC 10 Does the test evidence support the findings presented?

Each of the studies under considerations should be classified according to these 10 quality criteria. The protocol should clearly specify the granularity of the score, e.g. yes (1 point), partly ($\frac{1}{2}$ point) or no (0 point). The protocol should further specify the threshold for studies to be accepted and if it is acceptable to have, e.g. zero points in certain criteria.

All the studies have been classified and a suitable set of worthy studies has been selected. Now the data from each study can now be extracted.

Step 4: Data extraction and monitoring

Step 5: Data synthesis

References

[1] J. P. T. Higgins and S. Green, editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011].* The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

[2] B. A. Kitchenham. Guidelines for performing systematic literature reviews in software engineering version 2.3. Technical Report EBSE-2007-01, Keele University and University of Durham, 2007.

[3] Terje Nesbakken Lillegraven and Arnt Christian Wolden. Design of a bayesian recommender system for tourists presenting a solution to the cold-start user problem. Master's thesis, Department of Computer and Information Science, NTNU, 2010.