

Research & Innovation

– Scientific Methodology–

Building the Literature Review

Olivier Boissier

Mines Saint-Etienne <http://www.emse.fr/~boissier>



CPS2 M1 — Fall 2020

Outline

Identify

Reading

Telling

Writing

Structured Literature Review (SLR)

Some steps

Demonstrate that you totally know the field

Literature review is a springboard for describing your idea

- ▶ Identify relevant papers and/or school/institute seminars
- ▶ Read the papers in varying detail
- ▶ Keep notes on each paper/seminar
- ▶ Weave these into a story
- ▶ Write your report

Source: Informatics Research Review, Stratis Viglas, University of Edinburgh

Outline

Identify

Reading

Telling

Writing

Structured Literature Review (SLR)

How to identify papers to read

- ▶ Search for relevant work
 - ▶ On the internet
 - ▶ In the library
 - ▶ Use previous related state of the art reports
- ▶ Follow-up the citations in the papers you read (the past)
- ▶ See who cited papers you read (the future)
- ▶ There are library and online resources (cf. **See Course on Bibliography**)
 - ▶ Citeseer and ISI Web of Knowledge
 - ▶ Google Scholar
- ▶ Define a thesaurus, a set of keywords related to problem, hypothesis

Source: Informatics Research Review, Stratis Viglas, University of Edinburgh

Outline

Identify

Reading

Telling

Writing

Structured Literature Review (SLR)

Reading/listening to a purpose

Always have questions in mind when reading or listening

- ▶ What are the aims and objectives of the work?
- ▶ What was achieved?
- ▶ What claims are being made?
- ▶ Is the supporting evidence convincing?
- ▶ Where will my (potential) project fit in?

Source: Informatics Research Review, Stratis Viglas, University of Edinburgh

Reading to different depths

- ▶ Some work is central to your concerns, some less so
- ▶ Need to vary reading depth
 - ▶ Some need only skim
 - ▶ Some read in depth
 - ▶ Some in between
- ▶ Could be 20+ papers in total, but only 3 or 4 in depth
 - ▶ But you need to cite everything

Source: Informatics Research Review, Stratis Viglas, University of Edinburgh

How to skim

- ▶ Read title, abstract, introduction, conclusion, bibliography, key sections
- ▶ Identify main contribution of paper
- ▶ How does it relate to other work?
- ▶ Identify key questions to be addressed and hunt for answers

Source: Informatics Research Review, Stratis Viglas, University of Edinburgh

- ▶ Make several passes over the paper
 - ▶ Start by skimming
 - ▶ Then read in increasing detail
- ▶ Apply techniques to your own examples
- ▶ Try explaining the ideas to a friend
- ▶ Apply paper reviewing techniques (see course on Research evaluation)

Source: Informatics Research Review, Stratis Viglas, University of Edinburgh

Outline

Identify

Reading

Telling

Writing

Structured Literature Review (SLR)

Telling a story

- ▶ Literature survey is part of motivation
- ▶ How did this field develop?
- ▶ How did it start?
- ▶ What are the rival approaches?
- ▶ How do pieces of work relate?
- ▶ Where are we now?
- ▶ What remains to be done?
- ▶ What are the hot topics?

Source: Informatics Research Review, Stratis Viglas, University of Edinburgh

Outline

Identify

Reading

Telling

Writing

Structured Literature Review (SLR)

Writing Literature review

- ▶ Introduction: identify and motivate topic
- ▶ Main body
 - ▶ Summarise each piece of work
 - ▶ Give critical analysis
 - ▶ Compare and contrast (Use comparison table along some well chosen criteria)
- ▶ Conclusion
 - ▶ What is state of the field?
 - ▶ Where next?
- ▶ Bibliography: list all (and only) papers cited

Source: Informatics Research Review, Stratis Viglas, University of Edinburgh

Technical writing

- ▶ When I read a technical paper, I want to know exactly what is going on
- ▶ It's not a mystery novel; there is no plot, only facts and (maybe) opinion
- ▶ Do not try to write flamboyantly; it confuses the reader
- ▶ Close your Thesaurus; NOW!
- ▶ Use terminology; it's there for a reason
- ▶ Read it three times a day, with every meal

Source: Informatics Research Review, Stratis Viglas, University of Edinburgh

Avoid plagiarism

- ▶ Quotations must be acknowledged
- ▶ Including close paraphrase
- ▶ Use quote marks and cite source
- ▶ Best practice in research
- ▶ Plagiarism carries serious penalties

Source: Informatics Research Review, Stratis Viglas, University of Edinburgh

Writing References

- ▶ List all and only papers cited
- ▶ example: [Hacker 2000] Hacker, A., “A model of free will”, Journal of Computational Theology, pp 1-42, Vanity Press, 2000.
- ▶ There are several styles and publication types
- ▶ LaTeX supports several
- ▶ Provide the following: author(s) name(s), paper title, journal/book/conference title, year, pagination, volume/number, editor(s), publisher

Source: Informatics Research Review, Stratis Viglas, University of Edinburgh

Outline

Identify

Reading

Telling

Writing

Structured Literature Review (SLR)

Planning the review

Conducting the review

Introduction

Structured Literature Review (SLR)

What is it? a formal way of synthesizing the available information from primary studies relevant to a set of research questions

- ▶ use of a strict methodological framework
- ▶ use of well defined steps
- ▶ use of a predefined protocol

Notice that using a systematic literature review is in no way a guarantee of finding all relevant literature in a given area, however it can:

- ▶ map out existing solutions before attempting to tackle an area
- ▶ help researchers in avoiding bias in their work, in identifying gaps in their knowledge
- ▶ benefit the community by avoiding duplicate efforts
- ▶ highlight areas where additional research is required

Source of the following slides: Guidelines for performing Systematic Literature Reviews in SE, Kitchenham et al., 2007

Advantages and disadvantages

Structured Literature Review (SLR)

- ▶ Advantages are:
 - ▶ The well-defined methodology makes it less likely that the results of the literature are biased, although it does not protect against publication bias in the primary studies.
 - ▶ They can provide information about the effects of some phenomenon across a wide range of settings and empirical methods. If studies give consistent results, systematic reviews provide evidence that the phenomenon is robust and transferable. If the studies give inconsistent results, sources of variation can be studied.
 - ▶ In the case of quantitative studies, it is possible to combine data using meta-analytic techniques. This increases the likelihood of detecting real effects that individual smaller studies are unable to detect.
- ▶ The major disadvantage of systematic literature reviews is that they require considerably more effort than traditional literature reviews. In addition, increased power for meta-analysis can also be a disadvantage, since it is possible to detect small biases as well as true effects.

Features

Structured Literature Review (SLR)

A SLR differs of a conventional expert literature review by:

- ▶ a *review protocol* that specifies the research question being addressed and the methods that will be used to perform the review.
- ▶ a *defined search strategy* that aims to detect as much of the relevant literature as possible.
- ▶ the search strategy is *documented* so that readers can assess their rigour and the completeness and repeatability of the process (bearing in mind that searches of digital libraries are almost impossible to replicate).
- ▶ *explicit inclusion and exclusion criteria* to assess each potential primary study.
- ▶ *specification* of the information to be obtained from each primary study including *quality criteria* by which to evaluate each primary study.
- ▶ being a prerequisite for quantitative meta-analysis.

SLR process

Structured Literature Review (SLR)

1. Planning the review
 - 1.1 Identification of the need for a review
 - 1.2 (Commissioning a review)
 - 1.3 Specifying the research question(s)
 - 1.4 Developing a review protocol
 - 1.5 (Evaluating the review protocol)
2. Conducting the review
 - 2.1 Identification of research
 - 2.2 Selection of primary studies
 - 2.3 Study quality assessment
 - 2.4 Data extraction and monitoring
 - 2.5 Data synthesis
3. Reporting the review
 - 3.1 Specifying dissemination strategy
 - 3.2 Formatting the main report
 - 3.3 Evaluating the report

Specifying the research questions

Structured Literature Review (SLR)– Planning the review

Specifying the research questions is the most important part of any systematic review. The review questions drive the entire systematic review methodology:

- ▶ The search process must identify primary studies that address the research questions.
- ▶ The data extraction process must extract the data items needed to answer the questions.
- ▶ The data analysis process must synthesise the data in such a way that the questions can be answered.

The critical issue is to ask the right question. It is usually one that:

- ▶ is meaningful and important to practitioners as well as researchers
- ▶ will lead either to changes or to increased confidence
- ▶ will identify discrepancies between commonly held beliefs and reality
- ▶ could be based on PICOC (Population (e.g. users, applications), Intervention (e.g. methodology, tool, technology), Comparison, Outcome, Context) criteria

Specifying the research questions (2)

Structured Literature Review (SLR)– Planning the review

Finding keywords:

- ▶ Analyze the research questions and extract the first keywords;
- ▶ Use papers from the area of interest to find new words and synonyms of the words already found;
- ▶ Define the PICOC to better delineate the scope/aims of the systematic review (optional).

Defining search string:

- ▶ The basic rule is to separate the keywords.
- ▶ For each separated word, find its synonyms and concatenate them with the OR connector.
- ▶ After the definition of the groups of words with their synonyms, concatenate them with AND to end the string.

example: (pragmatic OR pragmatics OR pragmatism) AND (interoperability OR interoperate OR interoperable OR interoperation OR similarity OR integrate OR integration) AND (solution OR method OR technique OR model OR tool OR framework OR architecture OR infrastructure OR approach) AND (computational OR system OR application OR software)

Specifying the research questions (3)

Structured Literature Review (SLR)– Planning the review

1. Defining search engines
 - ▶ Depends on the area of systematic review.
 - ▶ Considers the main electronic search engines that include the researches produced in the area of Computer Science
 - ▶ See course on bibliography
2. String refinement by verifying if the returned papers appear to be relevant. If it is not the case, refine the search string
3. Search string execution by adapting its syntax to the search engines
4. Download and store search results in BIB format and by selecting citations and abstract information

The review protocol

Structured Literature Review (SLR)– Planning the review

- ▶ Background. The rationale for the survey.
- ▶ The research questions that the review is intended to answer.
- ▶ The strategy that will be used to search for primary studies including search terms and resources to be searched. Resources include digital libraries, specific journals, and conference proceedings. An initial mapping study can help determine an appropriate strategy.
- ▶ Study selection criteria. Study selection criteria are used to determine which studies are included in, or excluded from, a systematic review. It is usually helpful to pilot the selection criteria on a subset of primary studies.
- ▶ Study selection procedures. The protocol should describe how the selection criteria will be applied e.g. how many assessors will evaluate each prospective primary study, and how disagreements among assessors will be resolved.
- ▶ Study quality assessment checklists and procedures. The researchers should develop quality checklists to assess the individual studies. The purpose of the quality assessment will guide the development of checklists.
- ▶ Data extraction strategy. This defines how the information required from each primary study will be obtained. If the data require manipulation or assumptions and inferences to be made, the protocol should specify an appropriate validation process.
- ▶ Synthesis of the extracted data. This defines the synthesis strategy. This should clarify whether or not a formal meta-analysis is intended and if so what techniques will be used.
- ▶ Dissemination strategy (if not already included in a commissioning document).
- ▶ Project timetable. This should define the review schedule.

Identification of research

Structured Literature Review (SLR)– Conducting the review

SLR aims at finding as many primary studies relating to the search process by using an unbiased search strategy

Search strategies are usually iterative and benefit from:

- ▶ Preliminary searches aimed at both identifying existing systematic reviews and assessing the volume of potentially relevant studies.
- ▶ Trial searches using various terms derived from the research question.
- ▶ Checking trial research strings against lists of already known primary studies.
- ▶ Consultations with experts in the field

Identification of research (2)

Structured Literature Review (SLR)– Conducting the review

Publication bias (positive results are more likely to be published than negative results) can lead to systematic bias in SLR. Special efforts are made to address this problem covered by the strategies and including:

- ▶ Scanning the grey literature
- ▶ Scanning conference proceedings
- ▶ Contacting experts and researchers working in the area and asking them if they know of any unpublished results.

Documenting the search so that the process is transparent and replicable (as far as possible):

- ▶ The review must be documented in sufficient detail for readers to be able to assess the thoroughness of the search.
- ▶ The search should be documented as it occurs and changes noted and justified.
- ▶ The unfiltered search results should be saved and retained for possible reanalysis.

Identification of research (3)

Structured Literature Review (SLR)– Conducting the review

Data Source	Documentation
Digital Library	Name of database Search strategy for the database Date of search Years covered by search
Journal Hand Searches	Name of journal Years searched Any issues not searched
Conference proceedings	Title of proceedings Name of conference (if different) Title translation (if necessary) Journal name (if published as part of a journal)
Efforts to identify unpublished studies	Research groups and researchers contacted (Names and contact details) Research web sites searched (Date and URL)
Other sources	Date searched/contacted URL Any specific conditions pertaining to the search

Selection of primary studies

Structured Literature Review (SLR)– Conducting the review

Study selection criteria:

- ▶ are intended to identify those primary studies that provide direct evidence about the research question.
- ▶ should be decided during the protocol definition in order to reduce the likelihood of bias, although they may be refined during the search process.
- ▶ may be distinguished in inclusion and exclusion criteria, based on the research question.
- ▶ should be piloted to ensure that they can be reliably interpreted and that they classify studies correctly.

Further selection criteria can be:

- ▶ Language, Journal, Authors
- ▶ Setting
- ▶ Participants or subjects
- ▶ Research Design
- ▶ Sampling method
- ▶ Date of publication.

Study quality assessment

Structured Literature Review (SLR)– Conducting the review

- ▶ To provide still more detailed inclusion/exclusion criteria.
- ▶ To investigate whether quality differences provide an explanation for differences in study results.
- ▶ As a means of weighting the importance of individual studies when results are being synthesised.
- ▶ To guide the interpretation of findings and determine the strength of inferences.
- ▶ To guide recommendations for further research.

An initial difficulty is that there is no agreed definition of study quality. It is suggested that quality relates to the extent to which the study minimises bias and maximises internal and external validity.

Data extraction and monitoring

Structured Literature Review (SLR)– Conducting the review

Data extraction forms:

- ▶ must be designed to collect all the information needed to address the review questions and the study quality criteria.
- ▶ If the quality criteria are to be used to identify inclusion/exclusion criteria, they require separate forms (since the information must be collected prior to the main data extraction exercise).
- ▶ If the quality criteria are to be used as part of the data analysis, the quality criteria and the review data can be included in the same form.
- ▶ In most cases, data extraction will define a set of numerical values that should be extracted for each study (e.g. number of subjects, treatment effect, confidence intervals, etc.).
- ▶ Numerical data are important for any attempt to summarise the results of a set of primary studies and are a prerequisite for meta-analysis (i.e. statistical techniques aimed at integrating the results of the primary studies).

Data extraction and monitoring (2)

Structured Literature Review (SLR)– Conducting the review

Data extraction procedures:

- ▶ Whenever feasible, data extraction should be performed independently by two or more researchers.
- ▶ Data from the researchers must be compared and disagreements resolved either by consensus among researchers or arbitration by an additional independent researcher.
- ▶ Uncertainties about any primary sources for which agreement cannot be reached should be investigated as part of any sensitivity analyses.
- ▶ A separate form must be used to mark and correct errors or disagreements.
- ▶ If several researchers each review different primary studies because time or resource constraints prevent all primary papers being assessed by at least two researchers, it is important to employ some method of checking that researchers extract data in a consistent manner.
- ▶ For single researchers such as PhD students, other checking techniques must be used including for instance supervisors, or use of rest-retest process

Data synthesis

Structured Literature Review (SLR)– Conducting the review

Descriptive narrative data synthesis:

- ▶ Extracted information about the studies (i.e. intervention, population, context, sample sizes, outcomes, study quality) should be tabulated in a manner consistent with the review question.
- ▶ Tables should be structured to highlight similarities and differences between study outcomes.
- ▶ It is important to identify whether results from studies are consistent with one another (i.e. homogeneous) or inconsistent (e.g. heterogeneous).
- ▶ Results may be tabulated to display the impact of potential sources of heterogeneity, e.g. study type, study quality, and sample size.

Quantitative data synthesis should be presented in tabular form including:

- ▶ Sample size for each intervention.
- ▶ Estimates effect size for each intervention with standard errors for each effect.
- ▶ Difference between the mean values for each intervention, and the confidence interval for the difference.
- ▶ Units used for measuring the effect.

References

- ▶ A TUTORIAL and source of multiple resources:
<https://guides.lib.monash.edu/researching-for-your-literature-review/>
- ▶ Systematic Literature Review:
 - ▶ A Tutorial:
<https://guides.lib.monash.edu/systematic-review>
 - ▶ Guidelines for performing Systematic Literature Reviews in SE, Kitchenham et al., 2007
 - ▶ Systematic Literature Review in Computer Science-A Practical Guide, F Weidt, R Silva - Relatorios Tecnicos do DCC/UFJF, 2016