

**How to choose: Survey?
Systematic Mapping Review?
Systematic Literature Review?
Snowballing?**



Maíra Marques Samary

mmarques@dcc.uchile.cl

Agenda



- What it is?
 - Survey
 - Systematic Reviews
 - Snowballing
- How to do it?
 - Systematic Reviews
 - Research Questions
 - Database
 - Strings
 - Inclusion/Exclusion Criteria
 - Final Part

What it is?



Survey

- ◉ Definition – “act of surveying an area...”

How to do it?

- ◉ Find a seminal paper of the specific topic you choose and look for papers that this paper cited or papers that cited that paper. How to end the cycle of looking for papers, you never know...
- ◉ Validity? None – is not possible to reproduce
- ◉ Is it possible to publish it? Nowadays, hardly possible

What it is?



Systematic Reviews

- Uses rigorous, systematic, and transparent methods to minimize bias.
- Transparent means that the decisions are clearly documented.
- To identify existing evidence on something, summarize the evidence, and limit bias.
- To identify any gaps in current research in order to suggest areas for further investigation.
- To provide a framework/background in order to appropriately position new research activities.

**Systematic
Mapping
Review**

**Systematic
Literature
Review**

What it is?



Systematic Mapping Review

- ◉ Goal - provide an overview of a research area, and identify the amount, the type of research and results available
- ◉ Map the frequencies of publication over time to see trends
- ◉ Identify forums and relevant authors in which research in the area has been published
- ◉ Use – when you are beginning and are not sure about the area you want to do your research
- ◉ Advantages – it can be published, and may be a chapter of your thesis
- ◉ Effort - High

What it is?



Systematic Literature Review

- ◉ Goal - provide a status of the research being done on a specific research area, and identify the efforts being done to answer the open questions of this research area
- ◉ Identify authors that are working or have worked with the topic and at what point their research is published
- ◉ Use – when you already know your research area
- ◉ Advantages – if done properly can be published in a journal, and it will definitely be a thesis chapter
- ◉ Effort – VERY HIGH

What it is?



Snowballing

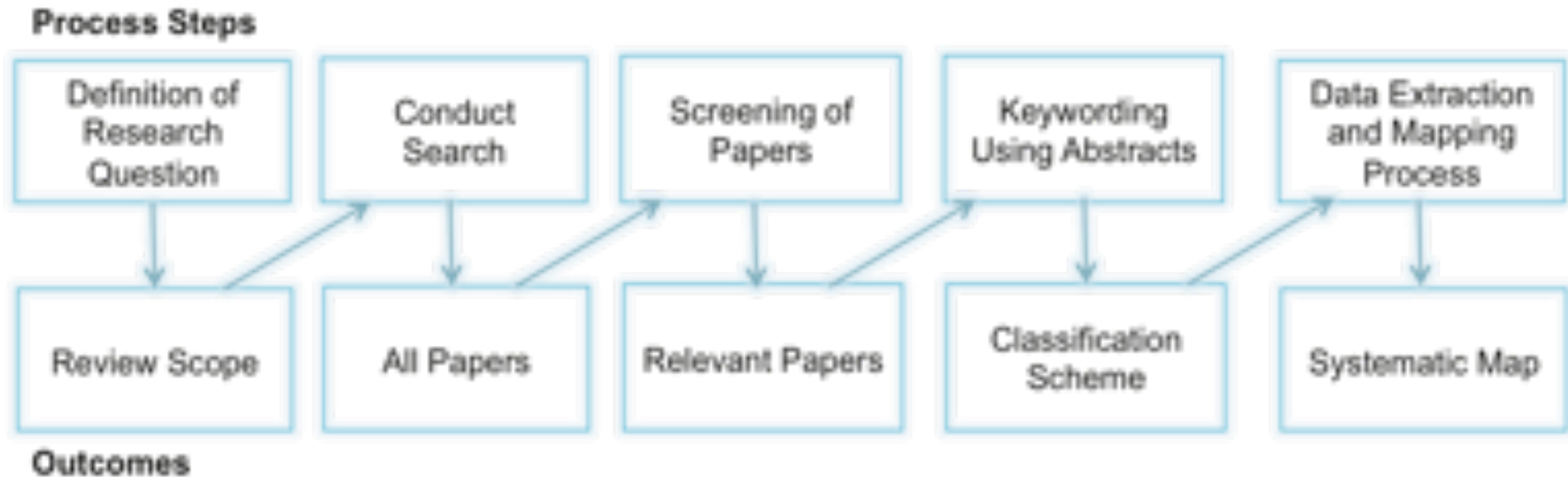
- ◉ It is a continuous, recursive process of gathering, searching, scanning and aggregating references
- ◉ It is being used as a way of search string validation – it is not mandatory
- ◉ Select seminal papers of your area, verify if these papers and that some papers that refers them are in your list of papers being analyzed.



How to do it?

- Systematic Mapping Review
- Systematic Literature Review

How to do it?



Petersen, F., R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic Mapping Studies in Software Engineering". Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering, (2008), pp. 71-80.

How to do it?



Systematic Mapping Review Questions:

- General:
 - RQ1: Which journals include papers on “area”?
 - RQ2: What are the most investigated “area” topics and how have these changed over time?
 - RQ3: What are the most frequently applied research methods, and in what study context?
- More Specific:
 - RQ1: What areas in “****” are addressed and how many articles cover the different areas?
 - RQ2: What types of papers are published in the area and in particular what type of evaluation and novelty do they constitute?

How to do it?



Systematic Literature Review Questions

- Questions are specific to the area
- An example:
 - What approaches have been reported regarding the analysis and support of PL evolution?
 - What PL artifacts can evolve?
 - How can PL artifacts evolve?
 - What PL evolution profiles are supported?
 - What techniques and methods are used to support PL evolution?
 - What is the maturity level of these approaches?

How to do it?



Databases

- Most relevant ones:
 - IEEEExplore
 - ACM Digital library
 - Citeseer library (citeseer.ist.psu.edu)
 - Inspec (www.iee.org/Publish/INSPEC/)
 - ScienceDirect (www.sciencedirect.com)
 - EI Compendex
 - Scopus
 - SpringerLink
 - Elsevier
- Tip: do not use all of them, choose 5 or 6

How to do it?



Strings

- The search strings are derived from the research questions.
- First: identify keywords
 - product line or product family
 - system family or software family or product family
 - evolution
 - erosion
 - refactoring or placement or continuous
- Create and test search string
 - (product line or product family) and (system family or software family or product family) and evolution and erosion and (refactoring or placement or continuous)

Now, the easy part is finished,
lets see the hard part

How to do it?



Choose a paper database management:

- ◉ Most used ones:
 - Papers
 - EndNote
 - Mendeley
 - RefDB
 - Zotero
 - JabRef

How to do it?



Some Databases Problems:

- ◉ Depending on the database there is no automatic root recognition (you have to use the plural of every word)
- ◉ Depending on the database you have a limitation of how many words you can use on your string (IEEE 15 words)
- ◉ Some databases offer advanced or expert search (ACM, IEEE) and others you do not have anything, not even a precedence rule

How to do it?



Databases Problems:

- ◉ Export Problems
 - IEEEExplore - exports in csv file
 - ACM Digital library – exports one by one in bib or csv

My solution:

Crawler

(ej. ACM – 30 seconds between each paper)

- ◉ Converter from csv to bib – minimum data

How to do it?



- ◉ Search strings – a lot of papers
- ◉ Merge of all the papers data
- ◉ 10% - 20% - repeated papers

How to do it?



- Inclusion criteria:
 - English, peer-reviewed articles in journals, conferences and workshops, published until Mar. 2015
 - Focus on product lines
 - Discuss some aspect of PL evolution
 - Approach and its validation are reasonably present
- Exclusion criteria:
 - No discussion of PL evolution
 - The paper only discusses evolution of code artifacts
 - The paper does not include validation of the approach
 - The paper is not in English
 - The paper is not peer reviewed
 - The paper is not available online
 - The paper is a survey or SLR, it does not present the outcomes of some specific research work

How to do it?



Now the fun begins!!!!

- ◉ Read all the titles and abstracts and decide if the paper will be included or excluded from your work
- ◉ More than one researcher
- ◉ It is time consuming and boring
- ◉ When there is no agreement between the inclusion/exclusion of some papers a third person may be asked to decide
- ◉ At the end you will probably have 10% - 20% of the initial papers.

How to do it?



Final part:

- ◉ Read all the papers (full reading)
- ◉ Evaluate if each one of the papers attend the inclusion/exclusion criteria
- ◉ If attend the include criteria, gather the information you need to answer your research questions
- ◉ If there is one reader, if there are doubts in the inclusion/exclusion criteria, the advisor should help decide
- ◉ If there is more than one reader, an agreement is needed between the readers on the inclusion/exclusion of each paper

How to do it?



- ◉ If you survived all this work, you just have to write your systematic literature review
- ◉ If you are performing a mapping, you just have to analyze and plot the data (normally a bubble plot) - and I'm still looking for an easy way to do it, if someone find out, please tell me!!!!
- ◉ All the data you have, how many papers on each phase, why important papers were not considered, you have to record and report it

