

# Systematic Literature Reviews



**Idaho State  
University**

Computer  
Science

**Isaac Griffith**

CS 6620 - Empirical Software Engineering  
Department of Computer Science  
Idaho State University

**ROAR**

# Outline

At the end of today's lecture you should be able to

- Know and understand the steps to conduct a SLR
- Understand various analysis approaches used for an SLR
- Structure and produce a review report
- Know why and when to conduct a review

# Introduction

## Systematic Literature Reviews (SLRs)

- Conducted to “identify, analyze, and interpret all available evidence related to a specific research question”
- Aims to give a complete, comprehensive, and valid picture of existing evidence
- This must be done in a scientific and rigorous way, using the following 3-step process:
  - ① Planning the review
  - ② Conducting the review
  - ③ Reporting the review

# Planning the Review

Planning an SLR requires the following actions:

- Identification of the need for a review
- Specifying the review questions
- Developing a review protocol

# Need for a Review

The need for an SLR is typically one or more of the following:

- Aiming to understand the state-of-the art in a research area
- A desire to use empirical evidence in decision-making



# Review Questions

- The area of review and the questions set the focus.
- To develop appropriate questions use the PICO(C) method:
  - **P: Population** - the population in which the evidence is collected, i.e., which group of people, programs or businesses are of interest for the review?
  - **I: Intervention** - the intervention applied in the empirical study, i.e., which technology, tool or procedure is under study?
  - **C: Comparison** - the comparison to which the intervention is compared, i.e. how is the control treatment defined?
  - **O: Outcomes** - the outcomes of the experiment should not only be statistically significant, but also be significant from a practical point of view.
  - **C: Context** - the context of the study must be defined, which is an extended view of the population, including whether it is conducted in academia or industry, in which industry segment, and also the incentives for the subjects.
- Additionally the experiment designs to include should be noted in the research question.



# Review Protocol

- Defines the review procedures and acts as a log
- The protocol should contain:
  - Background and rationale
  - Research questions
  - Search strategy for primary studies
  - Study selection criteria
  - Study selection procedures
  - Study quality assessment checklists and procedures
  - Data extraction strategy
  - Synthesis of the extracted data
  - Dissemination strategy
  - Project timetable



# Review Protocol

- Ensures consistency
- Ensures validity
- You should also conduct a pre-study to
  - scope the research questions
- Be open to modifying questions during protocol development



# Conducting the Review

A review is conducted by executing the following tasks:

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

# Research Identification

This step focuses on the search strategies that will be used to find primary studies, which include:

- Primary search involves two approaches
  - Manual Search
  - Database Search
- Snowball sampling
- Grey Literature



# Manual Search

- Using knowledge of the research area, a manual search can be conducted in the following locations:
  - Prominent Author's webpages
  - References of existing literature reviews
  - Known Conferences for the research area
  - Known Journals for the research area
- The goal is to obtain a representative sample of the papers from which snowball sampling can then be used.
  - The tradeoff between manual and database search is
    - Less false positives up front with manual search than with database search
    - More rounds of snowball sampling required for manual search
    - Higher chance of missing key primary studies with manual search if the original sample was not representative enough

# Database Search

If you do not have enough knowledge to conduct a manual search, a database search will be necessary. This proceeds as follows:

- Develop your search string
- Database selection
- Search execution

Note that there is always a tradeoff between:

- finding all relevant primary studies
- being overwhelmed with false positives to be removed manually



# Search String

- Use the research questions to create sets of keywords
- You should have one set per item of your PICO(C) question criteria
- For each item in each set, identify applicable synonyms for each keyword and add to the set.
- Build a Boolean search string from the selected keywords
  - **OR** together keywords within a set
  - surround sets with parentheses
  - **AND** together the sets
- If necessary, refine the expression to improve your results.
- You should analyze the sensitivity of the results of your search to refine the search string.



# Database Selection

- A single database is usually not enough to identify your primary studies.
- Using multiple databases will inevitably result in duplicates, but a greater sample of the population as well.
- For CS/SE Studies you should use at least the following:
  - IEEExplore
  - ACM Digital Library
- You should use 2 or more general indexes such as:
  - INSPEC/Compendex
  - Web of Science



# Database Selection

- Additionally, you can use the following indexes
  - ScienceDirect
  - SpringerLink
- **Do NOT use Google Scholar** since the results are not replicable
- Realize: you cannot find all primary studies for a given topic, and the what is found is simply a sample



# Search Execution

- Keep a detailed record of the search findings.
  - Number found
  - Included studies
  - Excluded studies
  - Duplicate studies



# Snowball Sampling

- Snowball sampling is simply a manual search for primary studies based on a selection of identified primary studies.
- There are two types of snowball sampling: Forward and Backward

## Backward Snowballing

- Search the references of a primary study for new primary studies
- Can be done either by:
  - Searching actual paper reference section, or
  - Searching through listed references on online paper listing
- This process continues as new studies are added until no new studies are added

## Forward Snowballing

- Search the items citing a primary study for new primary studies
- Can be done either by:
  - Using Google scholar cited-by listings
  - Using “Citations” section at paper listing

# Primary Study Selection

- Primary studies are selected based on a set of well-defined inclusion/exclusion criteria
  - These need to be defined before the search is conducted to reduce bias
- To reduce threats to validity it is wise to have more than one reviewer
  - Because selection is based on researcher judgment, you need to address this
  - After all researchers have made their assessments
    - Measure the inter-rater agreement with Cohen's Kappa
    - Use the Think-Aloud technique to attempt to come to a consensus
    - Measure the inter-rater agreement again
    - Report all of this in the final report



# Primary Study Selection

- Studies are selected based on researcher judgment and the application of inclusion/exclusion criteria
- Reducing the search results to the set of primary studies should be done in an iterative fashion
  - Start by removing those which can easily be excluded from their title or abstract alone
  - Next expand to those studies which can be excluded by their introduction and conclusions
  - The remaining studies should be thoroughly reviewed by a full text reading
- Remove any papers with more than one version (conference and journal) keeping only the most recent version



# Publication Bias

- There is a bias associated with published primary studies
  - They typically focus on positive results and discount negative results
- To overcome this bias you should consider gray literature
  - Technical Reports
  - Dissertations and Theses
  - Rejected Publications
  - Works in progress

# Study Quality Assessment

- Primary study quality is important as it
  - Can be used to analyze cause of contradicting results
  - Can be used in weighting the value of evidence from primary studies
- We measure quality through the use of checklists, several of which have been published
- Note that the quality of the study not the reporting are to be evaluated



# Data Extraction

- After primary studies are selected, data is extracted.
- A form for data collection should be developed from research questions
  - Expedites data collection
  - Increases reliability of data collection
- You should conduct a small data extraction using your forms on a subset of studies to validate the form

# Data Synthesis

There are several approaches to synthesize the data from a literature review

- Meta-analysis
- Descriptive synthesis
- Qualitative approaches for inhomogenous and mixed-method studies

Independent of approach used, a sensitivity analysis should take place

- Analyzes whether results are consistent across different subsets



# Meta-Analysis

- Most advanced form of data synthesis
- Assumptions
  - primary studies are homogenous
  - or cause of in-homogeneity is known
- Compares **effect sizes** and **p** values to assess synthesized outcome
- Primarily applicable to replication studies





# Meta-Analysis

- Studies to be included in a meta-analysis must:
  - Be of the same type, for example, formal experiments
  - Have the same test hypothesis
  - Have the same measures of the treatment and effect constructs
  - Report the same explanatory factors
- Process
  - ① Decide which studies to include in the meta-analysis
  - ② Extract the effect size from the primary study report, or estimate if there is no effect size published
  - ③ Combine the effect sizes from the primary studies to estimate and test the combined effect.



# Meta-Analysis

- In addition to the procedures noted
  - Include an analysis of **publication bias**
    - Funnel plots - where observed effect sizes are plotted against measure of study size (inverse of variance or other dispersion measure)
- Effect size (i.e. difference between mean values) for each study
  - normalize between studies by dividing by combined standard deviation
- Statistical evaluation
  - ① Determine homogeneity by evaluating heterogeneity using either the **Q test** or **Likelihood Ratio test**
  - ② Homogenous studies use a fixed-effects model
  - ③ Inhomogenous studies use a random effects model



# Descriptive Synthesis

- tabulates data from the primary studies to shed light on the research questions
- tabulated data should contain the following, at a minimum:
  - Sample size for each intervention
  - Estimates of effect size for each intervention (with standard errors)
  - Difference between mean values for each intervention
  - Confidence interval for the differences
  - Units used for measuring the effect
- Forest plot can visualize
  - means of difference between treatments of each study
  - variance of difference between treatments of each study



# Qualitative Approaches

- **Thematic Analysis** - aims at identifying, analyzing and reporting patterns or themes in the primary studies. At a minimum, it organizes and presents the data in rich detail, and interprets various aspects of the topic under study.
- **Narrative Synthesis** - tells the story originating from the primary evidence.
- **Comparative Analysis** - aims at analyzing complex causal connections. It uses Boolean logic to explain relations between cause and effect in the primary studies.



# Qualitative Approaches

- **Case Survey** - Aggregates existing research by applying a survey instrument of specific questions to each primary study. The data from the survey is quantitative and is aggregated using statistical methods.
- **Meta-ethnography** - translates studies into one another, and synthesizes the translations into concepts beyond the individual studies. The data is the interpretations and explanations in the primary studies.
- **Meta-Analysis** - can use statistical methods to integrate quantitative data from several cases

# Reporting the Review

The results of your research should be published where your intended audience can review it

## Practitioner Oriented

- ① Practitioner Journals and Magazines
- ② Press Releases to popular or specialist press
- ③ Short summary leaflets
- ④ Posters
- ⑤ Websites
- ⑥ Direct communication

## Academia

- ① Academic Journals
- ② Academic Conferences
  - If publication constraints prevent providing necessary detail for replication, write an accompanying Tech Report which can be posted online.



# Related Readings

- Chapter 4 of the book
- Kitchenham and Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering", version 2.3 - on moodle.



# A final word

- SLR and SMS are extremely useful studies
  - Help new researchers gain entry into a field of research
  - Develops the background section of a thesis or dissertation
  - Serve as the first key publication of your graduate degree
  - Help you identify current research challenges





**Are there any questions?**