Systematic Reviews with ReLiS

Prof. Eugene Syriani
Outline

1. Background on systematic reviews

2. Tool support: ReLiS

3. Challenges of conducting systematic reviews

4. Discussion: systematic review RL4SE
Secondary research

• **Narrative review**
  – Overview of current knowledge in a topic
  – *Informal survey*

• **Systematic review**
  – More detailed and rigorous
  – Review a research question
  – Reproducible methodology for searching papers and analyzing results

• **Meta-analysis**
  – Systematic review with statistical analysis to compare papers and derive new interpretations or new findings
Systematic review (SR)

(Kitchenham, 2004)

Research strategy used to extract new knowledge from existing studies

- Collect empirical proofs from existing studies corresponding to eligibility criteria to answer a specific research question

- Systematic $\Rightarrow$ methodological, repeatable, low bias
Systematic reviews in software engineering (SE)

- **Systematic Literature Review (SLR)** *(Kitchenham, 2007)*
  - Collects, synthesizes, and analyzes results of primary studies
  - Studies them in detail to answer precise questions empirically
  - Quality of primary studies is important

- **Systematic Mapping Studies (SMS)** *(Petersen, 2015)*
  - A.k.a. Scoping review
  - Classifies primary studies into categories to demonstrate research trends in a domain
  - Overview of a research domain and on quantity of available research
  - Analyzes a larger corpus than SLR
Process of Systematic Literature Reviews

Phase 1: Plan Review
1. Specify Research Questions
2. Develop Review Protocol
3. Validate Review Protocol

Phase 2: Conduct Review
4. Identify Relevant Research
5. Select Primary Studies
6. Assess Study Quality
7. Extract Required Data
8. Synthesize Data

Phase 3: Document Review
9. Write Review Report
10. Validate Report

(Brereton et al., 2007)
Process of Systematic Mapping Studies

(Petersen et al., 2015)
Systematic review process

Iterative process!

Planning ➔ Conducting ➔ Reporting

Protocol ➔ Activities ➔ Report

➢ Study identification
➢ Study selection
➢ Quality assessment
➢ Data extraction
➢ Synthesis

Refine ➔ Document
## Planning

### Need to perform SR
- Scientific, temporal reasons
- Number of primary studies available
- New review or update

### Research questions
- Scope
- Objectives
- PICOTS format
- Time interval

### Paper identification strategy
- Data sources
- Keywords
- Automatic search
- Manual search
- Snowballing
- Duplicate resolution
- Pilot studies

### Study selection strategy
- Inclusion, exclusion criteria
- Phases: meta-data, entire text
- Number and role of reviewers
- Conflict resolution strategy

### Evaluation of study quality
- Q&A check list
- Quality threshold
- Bias risk

### Data extraction and synthesis strategy
- Classification scheme
- Data extraction form
- Quantitative and qualitative analysis techniques of the data

### Dissemination strategy
- Report of protocol
- Publication of results
- Appendix for bibliography, extracted data, statistical analysis
Conducting

- **Identify candidate studies**
  - Execute search
  - Collect meta-information and complete text
  - Clean corpus, eliminate duplicates, correct references

- **Study selection**
  - Multiple phases
  - Distribution of screening among reviewers
  - Conflict detection
  - Conflict resolution
  - Validation by expert

- **Quality assessment**
  - Identify control list of studies
  - Evaluate each study by answering to precise questions
  - Decide to include/exclude study with low score

- **Data extraction and synthesis**
  - Fill data extraction form per retained study
  - Collaboration among reviewers
  - Conflict detection and resolution
  - Validation by expert
  - Synthesize to statistically analyze data
Reporting

Publish corpus
- References of included studies

Publish extracted data
- Tabular format
- Explanation of codes

Publish protocol report
- Who, Where, How
- Pilot studies
- Manual corrections

Publish results
- As a scientific contribution
- Good motivation
- Related work on topic and on methodology
- Interpret findings
PRISMA flow diagram
PRISMA flow diagram (simplified)
Notable standards

- **PRISMA** checklist [https://prisma.shinyapps.io/checklist/](https://prisma.shinyapps.io/checklist/)
- **PICOTS** to formulate scope
- Methodology for **systematic literature reviews** ([Kitchenham et al., 2007](#))
- Methodology for **systematic mapping studies** ([Petersen et al., 2015](#))
- Guide to **snowballing** ([Wohlin, 2014](#))
- Systematic review **updates** ([Mendes et al., 2020](#))
- **Experimentation** in software engineering ([Wohlin et al., 2012](#))
Problem in conducting SR

Complex process

• A lot of activities
• Iterative process
• Collaboration

Time consuming
Error prone

SR tools

• Automate activities
• Generative approach
• On-the-fly installation
• Web based

Reduce time
Improve SR quality
Tool support for SR

• Reference management & text-mining
  – EndNote, RefWorks, Zotero, JabRef, BiBler

• Specific to software engineering
  – StArt, Parsifal, SESRA, SLuRP, SRLTOOL, ReLiS

• Used in other disciplines
  – Eppi Reviewer, Convidence, DistillerSR, RevMan

Cannot modify configuration of SR procedure on-the-fly
Cannot pilot reviews
ReLiS

The team

• Brice Bigendako’s M.Sc. thesis (main initial developer)
  – [https://papyrus.bib.umontreal.ca/xmlui/handle/1866/20187](https://papyrus.bib.umontreal.ca/xmlui/handle/1866/20187)

• Mitacs GlobalLink interns (improving functionalities)
  – Naveen Kumar, Purboshi Das, Gauransh Kumar

• Student projects
  – Mahamat Youssouf Issa (test automation), Lucas Hornung (report generation)

• Research assistant: Naima Essadi (integration)

• Supervision: Prof. Eugene Syriani
ReLiS

Flexible tool for conducting SR collaboratively and iteratively

Planning → Conducting → Reporting

DSL for configuration
Generated web application
Synthesize data
SR workflow in ReLiS
SR workflow in ReLiS
SR workflow in ReLiS

- Manual entry
- CSV
- BibTeX
- EndNote
SR workflow in ReLiS
SR workflow in ReLiS
SR workflow in ReLiS
SR workflow in ReLiS
SR workflow in ReLiS
ReLiS in action

In this video we will show how to perform a systematic review following this process:

https://youtu.be/EFkMGfwCBsM

https://youtu.be/U5zOmk2vWy8
Feature coverage evaluation using (Al-Zubidy et al., 2017)

- 4th place on high level requirement: 55%
- 2nd place with no Integrated search + Text mining: 76%
- 1st place on all requirements: 58%
ReLiS tool availability

http://relis.iro.umontreal.ca

https://github.com/geodes-sms/relis
Tools to use with ReLiS

Bibliography curation

- **BiBler**: standard formatting for BibTeX
- **EndNote**: collects abstracts + keywords, efficient duplicate detection, collaboration
- **JabRef**, **Zotero**: grouping, extract directly from web

Analysis of extracted data

- **Excel**: cleaning data, clustering
- **R**, **SPSS**: statistical methods, graphs

Reporting

- **LaTeX/MiKTeX**
- **Overleaf**, **TeXStudio**
Tips using ReLiS

Setup and import

• **Assign roles appropriately**: administrator, project manager, validator, reviewer

• You may need to **query the database** for more advanced manipulations

• Typical **workflow to import papers**:
  1. Download references from online databases in EndNote format (ensure abstract is available)
  2. Create groups for each source
  3. Eliminate duplicates
  4. Export them to BiBler Export format
  5. Import them into BiBler, clean, and save
  6. Import them into ReLiS
Tips using ReLiS

Screening

• Beware that all statistics are updated continuously
  – If you add new papers in a screening phase, all statistics are going to be updated
  – If you resolve conflicts they will not appear in the statistics and kappa anymore

➢ Take screenshots of the statistics page at key moments

• Project managers can reassign papers to other reviewers if needed
Tips using ReLiS

Data extraction

• Define your **classification schema before** producing data in the project
  – Every category should be properly defined
  ➢ If you significantly change the classification scheme after screening, create a new project with classification only and import the included papers from the other project

• Use **DynamicList** fields as much as possible for better statistics and more flexibility

• **Maintain a log** of important milestones in a separate document, e.g.
  – Date and number of imported papers from a source, search string
  – Date of each activity and screenshot of their statistics
Challenges encountered in Systematic Reviews

Lessons learned from applying SLRs in SE
(Brereton et al., 2007) + mine
Challenges encountered in SR

1. Specify research questions

Expect to **revise RQs during protocol** development, as your understanding of the problem increases

Not while conducting the review!

A **pre-review mapping study** may help in scoping RQs

Include an RQ about **publication trends** and an RQ about **tool support** when appropriate
Challenges encountered in SR

2. Develop review protocol

- All reviewers must **collaborate** to develop the review protocol

- **Piloting the protocol** is essential to find mistakes in data collection, may indicate to change the methodology to address the RQs
Challenges encountered in SR

3. Validate review protocol

Specific validation process separate from the protocol

Ideally, external reviewers should undertake this validation process
Challenges encountered in SR

4. Identify relevant research

Explore **alternative search strategies** to achieve different sort of search completion criteria

Search many **different sources**: no single source indexes all studies in SE
ACM DL, IEEE Xplore, SpringerLink, Web of Science, Google Scholar, Scopus

Some excluded papers (especially secondary studies) can make up your **related work**
Challenges encountered in SR

5. Select primary studies

The standard of SE *abstracts* is too poor to rely on when selecting primary studies: also review the *conclusions*

Make sure all reviewers understand the *inclusion* and *exclusion criteria*: the scope must be well-defined

Decide when to resort to *snowballing*: missed articles, enlarge corpus
Challenges encountered in SR

6. Assess study quality

Some disciplines require a QA step, but in SE it depends:
Sometimes for SLR, not for SMS

Decide how QA is used in data aggregation and analysis
Challenges encountered in SR

7. Extract required data

For large corpus, have 2 reviewers for each paper: one **data extractor** and one **data checker**

Make sure **everyone understands** the protocol and data extraction process

Identify **clusters** of papers for repeated/similar studies
Challenges encountered in SR

8. Synthesize data

Systematic reviews in SE are likely to be qualitative in nature

Tabulating data is a useful aggregation, yet explain how the aggregated data actually answers the RQs

Explain or interpret unexpected observed phenomena from the results
Challenges encountered in SR

9+10. Write review report + Validate report

Keep a detailed **record of decisions** made throughout the review process

To validate it, publish a **technical report**, have it reviewed then write a scientific paper about the findings

Report **threats to validity** and how they have been mitigated

**Read** other systematic reviews
Examples of systematic reviews in SE


Discussion
Discussion

Systematic review on The use of reinforcement learning in SE

RQ1: What characteristics of SE problems RL has been used for?
   – What SE problem? In what SE activity? How is the RL method presented?

RQ2: What characteristics of RL methods have been used in SE?
   – What type of RL? What RL algorithm? How are state/policy/value function/reward signal represented? To what extent is the environment modeled? How are the results of the RL method presented/explainable?

RQ3: What RL tools are used and how they are integrated in the process?
   – Tool popularity in RL and SE communities? Documentation? Appreciation by SE?

RQ4: What are the typical challenges and limitations of using RL in SE?

RQ5: What are the publication trends in RL in SE?