Metodologi Penelitian

5. *Systematic Literatur Review (SLR)*

Husni

husni@trunojoyo.ac.id
http://husni.trunojoyo.ac.id

Course Outline

1. Pengantar Penelitian

2. Tahapan Penelitian

3. Literature Review

4. Penulisan Ilmiah dan Publikasi Penelitian

5. Systematic Literature Review (SLR)

6. Pembimbingan dan Presentasi Penelitian

5. Systematic Literature Review (SLR)

5.1 Pengantar SLR

5.2 Tahapan Planning

5.3 Tahapan Conducting

5.4 Tahapan Reporting

5.1 Pengantar SLR

Literature Review

- Literature Review is a critical and in depth evaluation of previous research (Shuttleworth, 2009) (*https://explorable.com/what-is-a-literature-review*)

- A summary and synopsis of a particular area of research, allowing anybody reading the paper to establish the reasons for pursuing a particular research

- A good Literature Review evaluates quality and findings of previous research

Why doing Literature Review?

- To establish connection or relationship between existing knowledge and the problem to be solved

- To refine the research problem

- To identify significance of research

- To define research question

Literature Review

- This literature review aims to identify and analyze the state-of-the-art research in the software defect prediction field

- Type of Literature Review:
    1. **Traditional Review**
    2. **Systematic Literature Review or Systematic Review**
    3. **Systematic Mapping Study (Scoping Study)**
    4. **Tertiary Study**

- SLR is now well established review method in the field of software engineering

(Kitchenham & Charters, Guidelines in performing Systematic Literature Reviews in Software Engineering, EBSE Technical Report version 2.3, 2007)

1. Traditional Review

- Provides an overview of the research findings on particular topics
- Advantages: produce insightful, valid syntheses of the research literature if conducted by the expert
- Disadvantages: vulnerable to unintentional and intentional bias in the selection, interpretation and organization of content
- Examples:
  - Liao et al., Intrusion Detection System: A Comprehensive Review, Journal of Network and Computer Applications, 36(2013)
  - Galar et al., A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), Vol. 42, No. 4, July 2012
  - Cagatay Catal, Software fault prediction: A literature review and current trends, Expert Systems with Applications 38 (2011)

2. Systematic Mapping Study

- Suitable for a very broad topic
- Identify clusters of evidence (making classification)
- Direct the focus of future SLRs
- To identify areas for future primary studies
- Examples:
  - Neto et al., A systematic mapping study of software product lines testing, Information and Software Technology Vol. 53, Issue 5, May 2011
  - Elberzhager et al., Reducing test effort: A systematic mapping study on existing approaches, Information and Software Technology 54 (2012)

3. Systematic Literature Review (SLR)

- The purpose of a systematic literature reviews is to provide as complete a list as possible of all the published studies relating to a particular subject area

- A process of identifying, assessing, and interpreting all available research evidence, to provide answers for a particular research question

- A form of secondary study that uses a well-defined methodology

- SLRs are well established in other disciplines, particularly medicine. They integrate an individual clinical expertise and facilitate access to the outcomes of the research

(Kitchenham & Charters, Guidelines in performing Systematic Literature Reviews in Software Engineering, EBSE Technical Report version 2.3, 2007)

4. Tertiary study

- Is a SLR of SLRs
- To answer a more wider question
- Uses the same method as in SLR
- Potentially less resource intensive
- Examples:
  - Kitchenham et al., Systematic literature reviews in software engineering – A tertiary study, Information and Software Technology 52 (2010)
  - Cruzes et al., Research synthesis in software engineering: A tertiary study, Information and Software Technology 53 (2011)

Tahapan SLR

1. Formulate the review's research question
2. Develop the review's protocol

PLANNING

1. Identify the relevant literature
2. Perform selection of primary studies
3. Perform data extraction
4. Assess studies' quality
5. Conduct synthesis of evidence

CONDUCTING

Write up the SLR report/paper

REPORTING

5.2 Tahapan Planning

1. Formulate the review's research question
2. Develop the review's protocol

Rumusan Masalah: *Research Question* (RQ)

- Is the most important part in any SLR

- Is not necessarily the same as questions addressed in your research

- Is used to guide the search process

- Is used to guide the extraction process

- Data analysis (synthesis of evidence) is expected to answer your SLR's RQ

Formulation of RQ

- Features of good question:
  - The RQ is meaningful and important to practitioners and researchers.
  - The RQ will lead to changes in current software engineering practice or to increase confidence in the value of current practice
  - The RQ will identify discrepancies between commonly held beliefs and the reality
- RQ can be derived primarily based on researcher's interest
  - An SLR for PhD thesis should identify existing basis for the research work and where it fits in the current body of knowledge

Formulation of RQ

The formulation of RQs about effectiveness of a treatment should focus on 5 elements known as PICOC:

- **Population (P)** - the target group for the investigation (e.g. people, software etc.)

- **Intervention (I)** - specifies the investigation aspects or issues of interest to the researchers

- **Comparison (C)**– aspect of the investigation with which the intervention is being compared to

- **Outcomes (O)**– the effect of the intervention

- **Context (C)**– the setting or environment of the investigation

(Petticrew et al., Systematic Reviews in the Social Sciences: A Practical Guide, Blackwell Publishing, 2006)

Example of PICOC (Kitchenham et al., 2007)

Kitchenham et al., A Systematic Review of Cross- vs. Within-Company Cost Estimation Studies, IEEE Transactions on Software Engineering, 33 (5), 2007

Population:	Software or web project
Intervention:	Cross-company project effort estimation model
Comparison:	Single-company project effort estimation model
Outcomes:	Prediction or estimate accuracy
Context:	None

Example of PICOC (Salleh et al., 2011)

Salleh et al., Empirical studies of pair programming for CS/SE teaching in higher education: A systematic literature review. IEEE Transactions on Software Engineering, 37(4), 2011

Population:	CS/SE students in higher education
Intervention:	Pair programming
Comparison:	N/A
Outcomes:	Pair Programming's effectiveness
Context:	Review(s) of all empirical studies of PP within the domain of CS/SE in higher education

Example of RQs

Kitchenham et al., A Systematic Review of Cross- vs. Within-Company Cost Estimation Studies, IEEE Transactions on Software Engineering, 33 (5), 2007

- RQ1: What evidence is there that cross-company estimation models are not significantly different from within-company estimation models for predicting effort for software/Web projects?
- RQ2: What characteristics of the study data sets and the data analysis methods used in the study affect the outcome of within- and cross-company effort estimation accuracy studies?
- RQ3: Which experimental procedure is most appropriate for studies comparing within- and cross-company estimation models?

Example of RQs

Davis et al., Effectiveness of Requirements Elicitation Techniques: Empirical Results Derived from a Systematic Review, 14th IEEE Requirements Engineering Conference, 2006

- RQ: What elicitation technique is most efficient in a particular setting?

Example of RQs

Radjenovic et al., "Software fault prediction metrics: A systematic literature review"

- RQ1: Which software metrics for fault prediction exist in literature?

- RQ2: What data sets are used for evaluating metrics?

SLR Protocol

- A plan that specifies the basic review procedures (method)

- Components of a protocol:

  1. Background
  2. Research Questions
  3. Search terms
  4. Selection criteria
  5. Quality checklist and procedures
  6. Data extraction strategy
  7. Data synthesis strategy

5.3 Tahapan Conducting

1. Identify the relevant literature
2. Perform selection of primary studies
3. Perform data extraction
4. Assess studies' quality
5. Conduct synthesis of evidence

1. Identifying Relevant Literature

- Involves a comprehensive and exhaustive searching of studies to be included in the review

- Define a search strategy

- Search strategies are usually iterative and benefit from:
  - Preliminary searches (to identify existing review and volume of studies)
  - Trial searches (combination of terms from RQ)
  - Check the search results against list of known studies
  - Consult the experts in the field

Common Approach to Construct Search String

- Derive major terms used in the review questions based on the PICOC

- List the keywords mentioned in the article

- Search for synonyms and alternative words

- Use the boolean OR to incorporate alternative synonyms

- Use the boolean AND to link major terms

Common Approach to Construct Search String

- Derive major terms used in the review questions based on the PICOC

- List the keywords mentioned in the article

- Search for synonyms and alternative words

- Use the boolean OR to incorporate alternative synonyms

- Use the boolean AND to link major terms

E.g. Search String

Salleh et al. (2011)

- The complete search term initially used :

    (student* OR undergraduate*) AND (pair programming OR pair-programming) AND ((experiment* OR measurement OR evaluation OR assessment) AND (effective* OR efficient OR successful)

- A very limited number of results retrieved when using the complete string, thus a much simpler string was derived.

- Subject librarian suggested to revise the search string:

    "pair programming" OR "pair-programming"

E.g. Search String

- Kitchenham et al. (2007) used their structured questions to construct search strings for use with electronic databases:
  - *Population*: software OR application OR product OR Web OR WWW OR Internet OR World-Wide Web OR project OR development
  - *Intervention*: cross company OR cross organisation OR cross organization OR multiple-organizational OR multiple-organisational model OR modeling OR modelling effort OR cost OR resource estimation OR prediction OR assessment
  - *Contrast*: within-organisation OR within-organization OR within-organizational OR within-organisational OR single company OR single organisation
  - *Outcome*: Accuracy OR Mean Magnitude Relative Error
- The search strings were constructed by linking the four OR lists using the Boolean AND

Sources of Evidence

- Digital libraries

- Reference lists from relevant primary studies and review articles

- Journals (including company journals such as the IBM Journal of Research and Development), grey literature (i.e. technical reports, work in progress)

- Conference proceedings

- Research registers

- The Internet (google)

- Direct contact specific researcher(s)

E.g. Sources of Evidence

Salleh et al. (2011)

- Online databases used:
  - ACM Digital Library, Current Contents, EBSCOhost, IEEExplore, ISI Web of Science, INSPEC, ProQuest, Sage Full text Collection, ScienceDirect, SpringerLink, Scopus
- Other search engines used: Google scholar, Citeseer, Agile Alliance.
- Some databases were selected based on previous studies we were aware of.

E.g. Sources of Evidence

Kitchenham et al. (2007)

- The search strings were used on 6 digital libraries:
  - INSPEC , El Compendex, Science Direct, Web of Science, IEEExplore, ACM Digital library

- Search specific journals and conf. proceedings:
  - Empirical Software Engineering (J)
  - Information and Software Technology (J)
  - Software Process Improvement and Practice (J)
  - Management Science (J)
  - International Software Metrics Symposium (C)
  - International Conference on Software Engineering (C)

- Manual search:
  - Evaluation and Assessment in Software Engineering (C)

- Check references of each relevant article

- Contact researchers

Managing Bibliography

- Use relevant Bibliographic package to manage large number of references

- E.g. Mendeley, EndNote, Zotero, JabRef Reference Manager etc.

Documenting the Search

- The process of conducting SLR must be transparent and replicable

- The review should be documented in sufficient detail

- The search should be documented and changes noted

- Unfiltered search results should be saved for possible reanalysis

Data Source	Documentation
Digital Library	Name of Database, Search strategy, Date of search, years covered by search
Journal Hand Searches	Name of journal, Years searched
Conference proceedings	Title of proceedings/Name of conference, Journal name (if published as part of a journal)

2. Selection of Studies

- Primary studies need to be assessed for their actual relevance

- Set the criteria for including or excluding studies (decided earlier during protocol development, can be refined later)

- Inclusion & exclusion criteria should be based on RQ

- Selection process should be piloted

- Study selection is a multistage process

2. Selection of Studies

- Primary studies need to be assessed for their actual relevance

- Set the criteria for including or excluding studies (decided earlier during protocol development, can be refined later)

- Inclusion & exclusion criteria should be based on RQ

- Selection process should be piloted

- Study selection is a multistage process

E.g. Selection of Studies

Kitchenham et al. (2007) used the following **inclusion** criteria:

- Any study that compared predictions of cross-company models with within-company models based on analysis of single company project data.

They used the following **exclusion** criteria:

- Studies where projects were only collected from a small number of different sources (e.g. 2 or 3 companies)
- Studies where models derived from a within-company data set were compared with predictions from a general cost estimation model.

E.g. Selection of Studies

Salleh et al. (2011)

- **Inclusion criteria**:
  - to include any empirical studies of PP that involved higher education students as the population of interest.
- **Exclusion criteria**:
  - Papers presenting unsubstantiated claims made by the author(s), for which no evidence was available.
  - Papers about Agile/XP describing development practices other than PP, such as test-first programming, refactoring etc.
  - Papers that only described tools (software or hardware) that could support the PP practice.
  - Papers not written in English.
  - Papers involving students but outside higher education

3. Assessing Studies' Quality

- To provide more detailed Inclusion/Exclusion criteria
- To check 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 synthesized
- To guide the interpretation of findings and determine the strength of inferences
- To guide recommendations for further research

Assessing Studies' Quality

- Quality relates to the extent to which the study minimizes bias and maximizes internal and external validity (Khan et al. 2001)

- Quality Concepts Definition (Kitchenham & Charter, 2007)

Terms	Synonyms	Definition
Bias	Systematic error	tendency to produce results that depart systematically from the 'true' results. Unbiased results are internally valid
Internal Validity	Validity	The extent to which the design and conduct of the study are likely to prevent systematic error. Internal validity is a prerequisite for external validity
External Validity	Generalizability, Applicability	The extent to which the effects observed in the study are applicable outside of the study

Assessing Studies' Quality

- Assessing quality of studies:
    - Methodology or design of the study
    - Analysis of studies' findings
- Quality checklist or instrument need to be designed to facilitate quality assessment
- Most quality checklists include questions aimed at assessing the extent to which articles have addressed bias and validity

E.g. Study Quality Assessment - Salleh et al. (2011)

Item	Answer
1. Was the article referred? [30]	Yes/No
2. Were the aim(s) of the study clearly stated? [16], [67]	Yes/No/Partially
3. Were the study participants or observational units adequately described? For example, students' programming experience, year of study etc. [44], [68]	Yes/No/Partially
4. Were the data collections carried out very well? For example, discussion of procedures used for collection, and how the study setting may have influenced the data collected [44], [48], [67], [68]	Yes/No/Partially
5. Were potential confounders adequately controlled for in the analysis? 67]	Yes/No/Partially
6. Were the approach to and formulation of the analysis well conveyed? For example, description of the form of the original data, rationale for choice of method/tool/package [48], [67], [68]	Yes/No/Partially
7. Were the findings credible? For example, the study was methodologically explained so that we can trust the findings; findings/conclusions are resonant with other knowledge and experience [48], [44], [68]	Yes/No/Partially

E.g. Study Quality Assessment

Kitchenham et al. (2007) constructed a quality questionnaire based on 5 issues affecting the quality of the study:

1. Is the data analysis process appropriate?
2. Did studies carry out a sensitivity or residual analysis?
3. Were accuracy statistics based on the raw data scale?
4. How good was the study comparison method?
5. The size of the within-company data set (e.g < 10 projects considered poor quality)

4. Data Extraction

- Involve reading the full text article
- Data extracted from primary studies should be recorded using *data extraction form*
- The form should be designed and piloted when the protocol is defined
- Collect all the information that can be used to answer the RQ and the study's quality criteria
- Both quality checklist and review data can be included in the same form
- In case of duplicates publications (reporting the same data), refer the most complete one
- For validation, a set of papers should be reviewed by 2 or more researchers. Compare results and resolve any conflicts

5. Synthesis of Evidence

- Involves collating and summarizing the results of the included primary studies

- Key objectives of data synthesis (Cruzes & Dyba, 2011):

  - to analyze and evaluate multiple studies

  - to select appropriate methods for integrating or providing new interpretive explanations about them

- Synthesis can be:

  - Descriptive (narrative/non-quantitative)

  - Quantitative (e.g. meta-analysis)

(Cruzes et al., Research Synthesis in Software Engineering: A tertiary study, Information and Software Technology, 53(5), 2011)

Descriptive Synthesis (Narrative)

"An approach to the synthesis of findings from multiple studies that relies primarily on the use of words and text to summarize and explain the findings of the synthesis. It adopts a textual approach to the process of synthesis to 'tell the story' of the findings from the included studies." (Popay et al. 2006)

- Use tables to tabulate information extracted from included studies (e.g. population, number of included studies, study quality etc.)

- Tables should be structured to highlight similarity or differences of study outcomes

- Were the findings consistent (homogeneous) or inconsistent?

Quantitative Synthesis (Meta-Analysis)

- Meta-analysis can be used to aggregate results or to pool data from different studies

- The outcome of a meta-analysis is an average effect size with an indication of how variable that effect size is between studies

- Meta-analysis involves three main steps:
    - 1. Decide which studies to be included in the meta-analysis
    - 2. Estimate an effect size for each individual study
    - 3. Combine the effect sizes from the individual studies to estimate and test the combined effect

- Results of the meta-analysis can be presented in a forest plot

5.4 Tahapan Reporting

Write up the SLR report/paper

Reporting SLR results in Journals

- Some journals and conferences include a specific topic on SLR:
    - Information & Software Technology has an editor specializing in systematic reviews
    - Journal of Systems and Software
    - Expert Systems with Applications
    - IEEE Transactions on Software Engineering
    - International Symposium on Empirical Software Engineering & Measurement (ESEM)
    - International Conference on Evaluation & Assessment in Software Engineering (EASE)
    - International Workshop on Evidential Assessment of Software Technologies (EAST)

Reporting Structure

- ## Introduction
  - General introduction about the research. State the purpose of the review. Emphasize the reason(s) why the RQ is important. State the significance of the review work and how the project contributes to the body of knowledge of the field.

- ## Main Body
  - Review method – briefly describe steps taken to conduct the review
  - Results – findings from the review
  - Discussion – implication of review for research & practice

- ## Conclusions

Referensi

- Abbott, M., & McKinney, J. (2013). Understanding and Applying Research Design. John Wiley & Sons, Inc.

- Berndtsson, M., Hansson, J., & Olsson, B. (2008). Thesis Projects: a Guide for Students in Computer Science and Information Systems (2nd ed.). London: Springer-Verlag

- Blaxter, L., Hughes, C., & Tight, M. (2006). How to Research (3rd ed.). Open University Press

- Blessing, L. T. M., & Chakrabarti, A. (2009). DRM, a Design Research Methodology. Springer-Verlag London

- Cohen, L., Manion, L., & Morrison, K. (2005). Research Methods in Education (5th ed.). Taylor & Francis Group

- Dawson, C. W. (2009). Projects in Computing and Information Systems A Student's Guide (2nd ed.). Pearson Education Limited

- Jonker, J., & Pennink, B. (2010). The Essence of Research Methodology. Springer-Verlag Berlin Heidelberg

- Lichtfouse, E. (2013). Scientific Writing for Impact Factor Journals. Nova Science Publishers, Inc.

5.4 Referensi

- Kothari, C. (2004). Research Methodology: Methods and Techniques. New Age International

- Might, M. (2010). The Illustrated Guide to a Ph.D. Matt.might.net. Retrieved from *http://matt.might.net/articles/phd-school-in-pictures/*

- Marczyk, G., DeMatteo, D., & Fertinger, D. (2005). Essentials of Research Design and Methodology. John Wiley & Sons, Inc.

- Rea, L. M., & Parker, R. A. (2014). Designing and Conducting Survey Research: A Comprehensive Guide (4th ed.). John Wiley & Sons, Inc.

- Runeson, P., Host, M., Rainer, A., & Regnell, B. (2012). Case Study Research in Software Engineering: Guidelines and Examples. John Wiley & Sons, Inc.

- Sahu, P. K. (2013). Research Methodology: A Guide for Researchers In Agricultural Science, Social Science and Other Related Fields. Springer

- Veit, R., Gould, C., & Gould, K. (2013). Writing, Reading, and Research (9th ed.). Cengage Learning