MTAT.03.270

Seminar on Enterprise Software

(3 ECTS)

Dietmar Pfahl

dietmar.pfahl@ut.ee

Schedule – sessions in red are mandatory

- 09 February - Introductory Session 1 (Dietmar Pfahl)
- 16 February - Introductory Session 2 (Dietmar Pfahl)
- 23 February & 02 March & 09 March - Individual Consultation (optional & on request by student only - first come first serve principle) - Online or face-to-face
- 16 March – Student Presentations 1 (work in pairs)
- 23 March – Student Presentations 2 (work in pairs)
- 30 March – Student Presentations 3 (work in pairs)
- 06 April – Student Presentations 4 (work in pairs)
- 13 April – Student Presentations 5 (work in pairs)
- 20 April – Student Presentations 6 (work in pairs)
- 27 April – Student Presentations 7 (work in pairs)
- May 04 & 11 – Consultation for final reports (optional & on request by student(s) only - first come first serve principle)

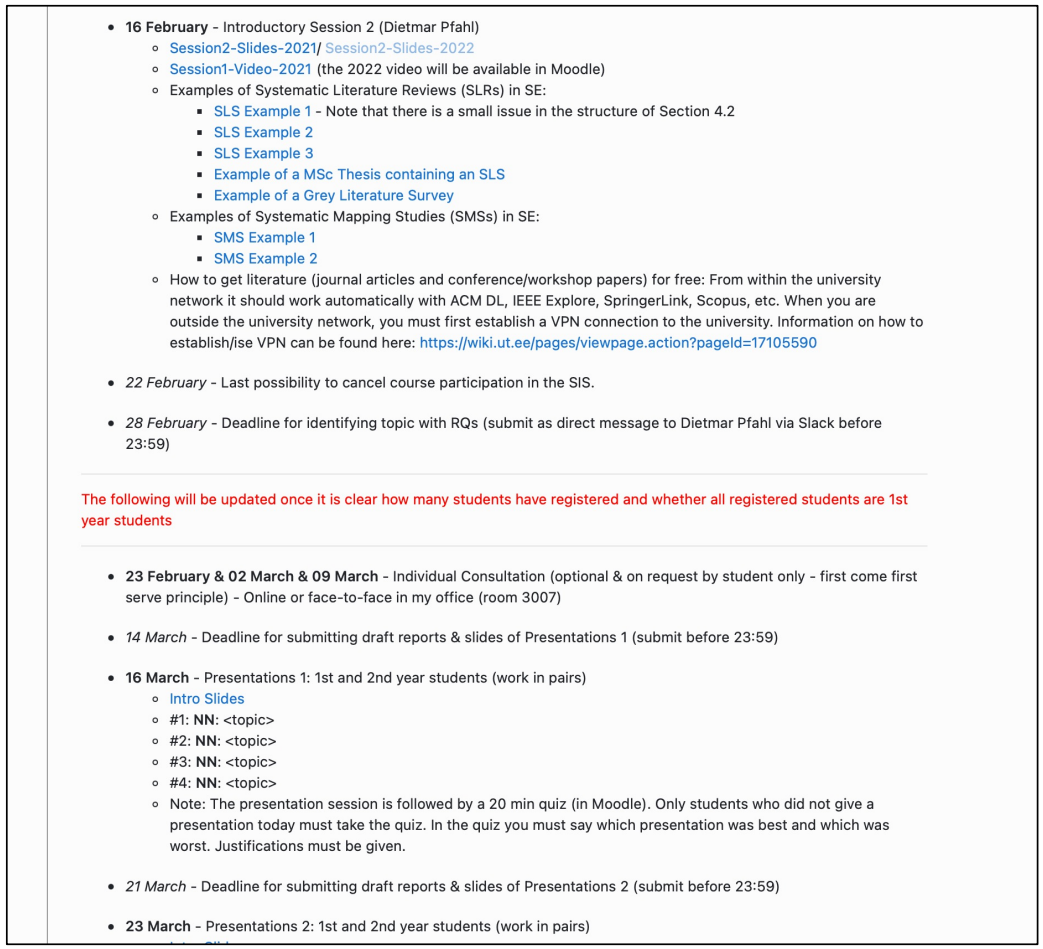
- **16 February** - Introductory Session 2 (Dietmar Pfahl)
  - Session2-Slides-2021/ Session2-Slides-2022
  - Session1-Video-2021 (the 2022 video will be available in Moodle)
  - Examples of Systematic Literature Reviews (SLRs) in SE:
    - SLS Example 1 - Note that there is a small issue in the structure of Section 4.2
    - SLS Example 2
    - SLS Example 3
    - Example of a MSc Thesis containing an SLS
    - Example of a Grey Literature Survey
  - Examples of Systematic Mapping Studies (SMSs) in SE:
    - SMS Example 1
    - SMS Example 2
  - How to get literature (journal articles and conference/workshop papers) for free: From within the university network it should work automatically with ACM DL, IEEE Explore, SpringerLink, Scopus, etc. When you are outside the university network, you must first establish a VPN connection to the university. Information on how to establish/ise VPN can be found here: https://wiki.ut.ee/pages/viewpage.action?pageId=17105590
- *22 February* - Last possibility to cancel course participation in the SIS.
- *28 February* - Deadline for identifying topic with RQs (submit as direct message to Dietmar Pfahl via Slack before 23:59)

The following will be updated once it is clear how many students have registered and whether all registered students are 1st year students

- **23 February & 02 March & 09 March** - Individual Consultation (optional & on request by student only - first come first serve principle) - Online or face-to-face in my office (room 3007)
- *14 March* - Deadline for submitting draft reports & slides of Presentations 1 (submit before 23:59)
- **16 March** - Presentations 1: 1st and 2nd year students (work in pairs)
  - Intro Slides
  - #1: **NN:** <topic>
  - #2: **NN:** <topic>
  - #3: **NN:** <topic>
  - #4: **NN:** <topic>
  - Note: The presentation session is followed by a 20 min quiz (in Moodle). Only students who did not give a presentation today must take the quiz. In the quiz you must say which presentation was best and which was worst. Justifications must be given.
- *21 March* - Deadline for submitting draft reports & slides of Presentations 2 (submit before 23:59)
- **23 March** - Presentations 2: 1st and 2nd year students (work in pairs)

Schedule – sessions in red are mandatory

- 09 February - Introductory Session 1 (Dietmar Pfahl)
- 16 February - Introductory Session 2 (Dietmar Pfahl)
- 23 February & 02 March & 09 March - Individual Consultation (optional & on request by student only - first come first serve principle) - Online or face-to-face
- ~~16 March – Student Presentations 1 (work in pairs)~~ ???
- 23 March – Student Presentations 2 (work in pairs)
- 30 March – Student Presentations 3 (work in pairs)
- 06 April – Student Presentations 4 (work in pairs)
- 13 April – Student Presentations 5 (work in pairs)
- 20 April – Student Presentations 6 (work in pairs)
- 27 April – Student Presentations 7 (work in pairs)
- May 04 & 11 – Consultation for final reports (optional & on request by student(s) only - first come first serve principle)



- **16 February** - Introductory Session 2 (Dietmar Pfahl)
  - Session2-Slides-2021/ Session2-Slides-2022
  - Session1-Video-2021 (the 2022 video will be available in Moodle)
  - Examples of Systematic Literature Reviews (SLRs) in SE:
    - SLS Example 1 - Note that there is a small issue in the structure of Section 4.2
    - SLS Example 2
    - SLS Example 3
    - Example of a MSc Thesis containing an SLS
    - Example of a Grey Literature Survey
  - Examples of Systematic Mapping Studies (SMSs) in SE:
    - SMS Example 1
    - SMS Example 2
  - How to get literature (journal articles and conference/workshop papers) for free: From within the university network it should work automatically with ACM DL, IEEE Explore, SpringerLink, Scopus, etc. When you are outside the university network, you must first establish a VPN connection to the university. Information on how to establish/ise VPN can be found here: https://wiki.ut.ee/pages/viewpage.action?pageId=17105590
- *22 February* - Last possibility to cancel course participation in the SIS.
- *28 February* - Deadline for identifying topic with RQs (submit as direct message to Dietmar Pfahl via Slack before 23:59)

The following will be updated once it is clear how many students have registered and whether all registered students are 1st year students

- **23 February & 02 March & 09 March** - Individual Consultation (optional & on request by student only - first come first serve principle) - Online or face-to-face in my office (room 3007)
- *14 March* - Deadline for submitting draft reports & slides of Presentations 1 (submit before 23:59)
- **16 March** - Presentations 1: 1st and 2nd year students (work in pairs)
  - Intro Slides
  - #1: **NN:** <topic>
  - #2: **NN:** <topic>
  - #3: **NN:** <topic>
  - #4: **NN:** <topic>
  - Note: The presentation session is followed by a 20 min quiz (in Moodle). Only students who did not give a presentation today must take the quiz. In the quiz you must say which presentation was best and which was worst. Justifications must be given.
- *21 March* - Deadline for submitting draft reports & slides of Presentations 2 (submit before 23:59)
- **23 March** - Presentations 2: 1st and 2nd year students (work in pairs)

Tasks

For reports and presentations:
1st and 2nd year students work in pairs
Feedback must be given individually

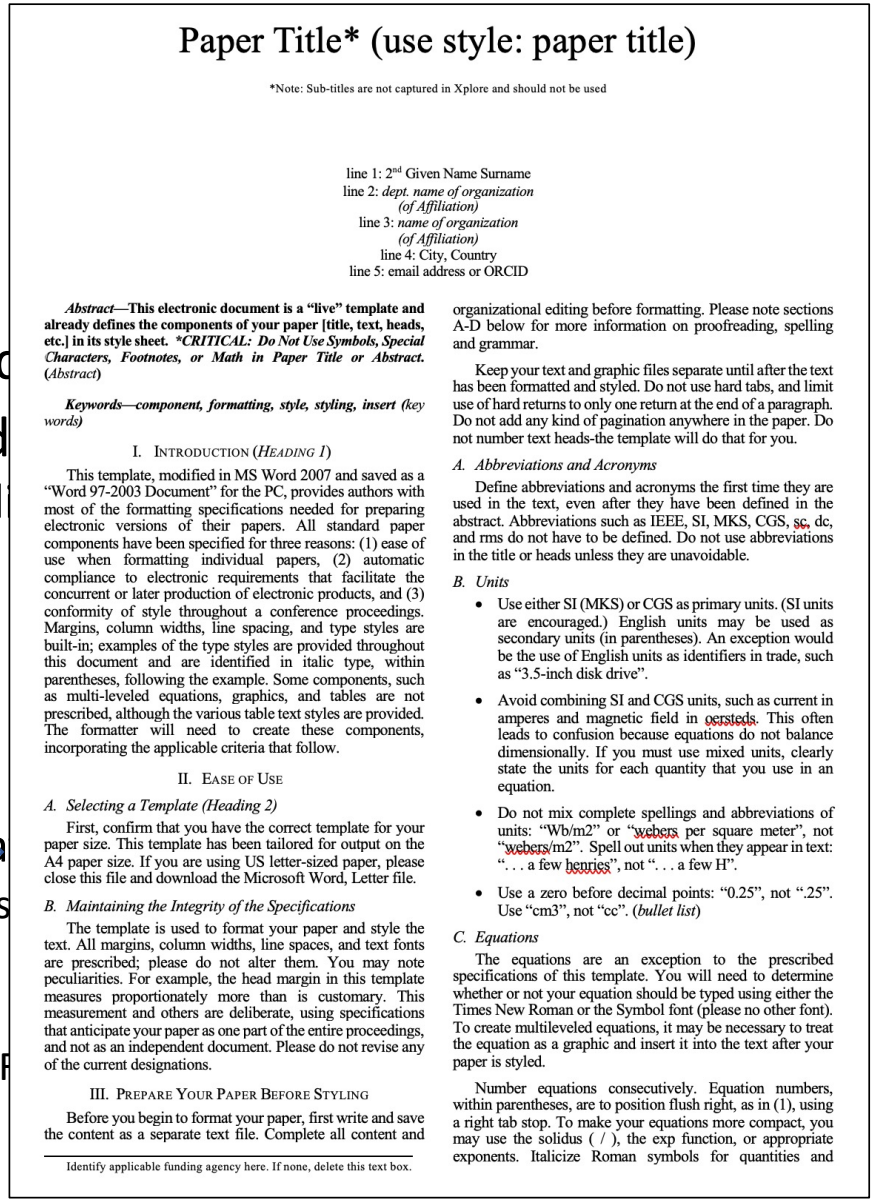
- Draft Report + Presentation Slides (submit 2 days before presentation)
- Feedback on presentations (via Quiz in Moodle at end of session)
  - Must give feedback at least for 5 sessions (excluding the one where you give a presentation)
- Final Report

- Submission formats:
  - Report: IEEE template (~4 pages) – take A4 format
    - Template is here: https://www.ieee.org/conferences/publishing/templates.html
    - Submit PDF via course wiki
  - Presentation slides: must fit in 12 min
    - Submit PDF (and Powerpoint if animations make PDF unreadable) via course wiki

Tasks

- Draft Report + Presentation Slides (submit 2 d
- Feedback on presentations (via Quiz in Mood
  - Must give feedback at least for 5 sessions (exclud presentation)
- Final Report

- Submission formats:
  - Report: IEEE template (~4 pages)
    - Template is here: https://www.ieee.org/conferences
    - Submit PDF via course wiki
  - Presentation slides: must fit in 12 min
    - Submit PDF (and Powerpoint if animations make PDF



Paper Title* (use style: paper title)

*Note: Sub-titles are not captured in Xplore and should not be used

line 1: 2nd Given Name Surname
line 2: *dept. name of organization (of Affiliation)*
line 3: *name of organization (of Affiliation)*
line 4: City, Country
line 5: email address or ORCID

Abstract*—This electronic document is a "live" template and already defines the components of your paper [title, text, heads, etc.] in its style sheet. **CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract. (*Abstract*)

Keywords—component, formatting, style, styling, insert (*key words*)

I. Introduction (*Heading 1*)

This template, modified in MS Word 2007 and saved as a "Word 97-2003 Document" for the PC, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. All standard paper components have been specified for three reasons: (1) ease of use when formatting individual papers, (2) automatic compliance to electronic requirements that facilitate the concurrent or later production of electronic products, and (3) conformity of style throughout a conference proceedings. Margins, column widths, line spacing, and type styles are built-in; examples of the type styles are provided throughout this document and are identified in italic type, within parentheses, following the example. Some components, such as multi-leveled equations, graphics, and tables are not prescribed, although the various table text styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow.

II. Ease of Use

A. Selecting a Template (Heading 2)

First, confirm that you have the correct template for your paper size. This template has been tailored for output on the A4 paper size. If you are using US letter-sized paper, please close this file and download the Microsoft Word, Letter file.

B. Maintaining the Integrity of the Specifications

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

III. Prepare Your Paper Before Styling

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

B. Units

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as "3.5-inch disk drive".
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
- Do not mix complete spellings and abbreviations of units: "Wb/m2" or "webers per square meter", not "webers/m2". Spell out units when they appear in text: ". . . a few henries", not ". . . a few H".
- Use a zero before decimal points: "0.25", not ".25". Use "cm3", not "cc". (*bullet list*)

C. Equations

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and

Identify applicable funding agency here. If none, delete this text box.

Seminar on Enterprise Software 2021/22 spring

Manage this course

Edit page

- Home
- Sessions & Deadlines
- Deliverables
  - Submission
- List of Topics
- Message Board (Slack)

Edit sidebar

Deliverables

The following deliverables must be submitted:

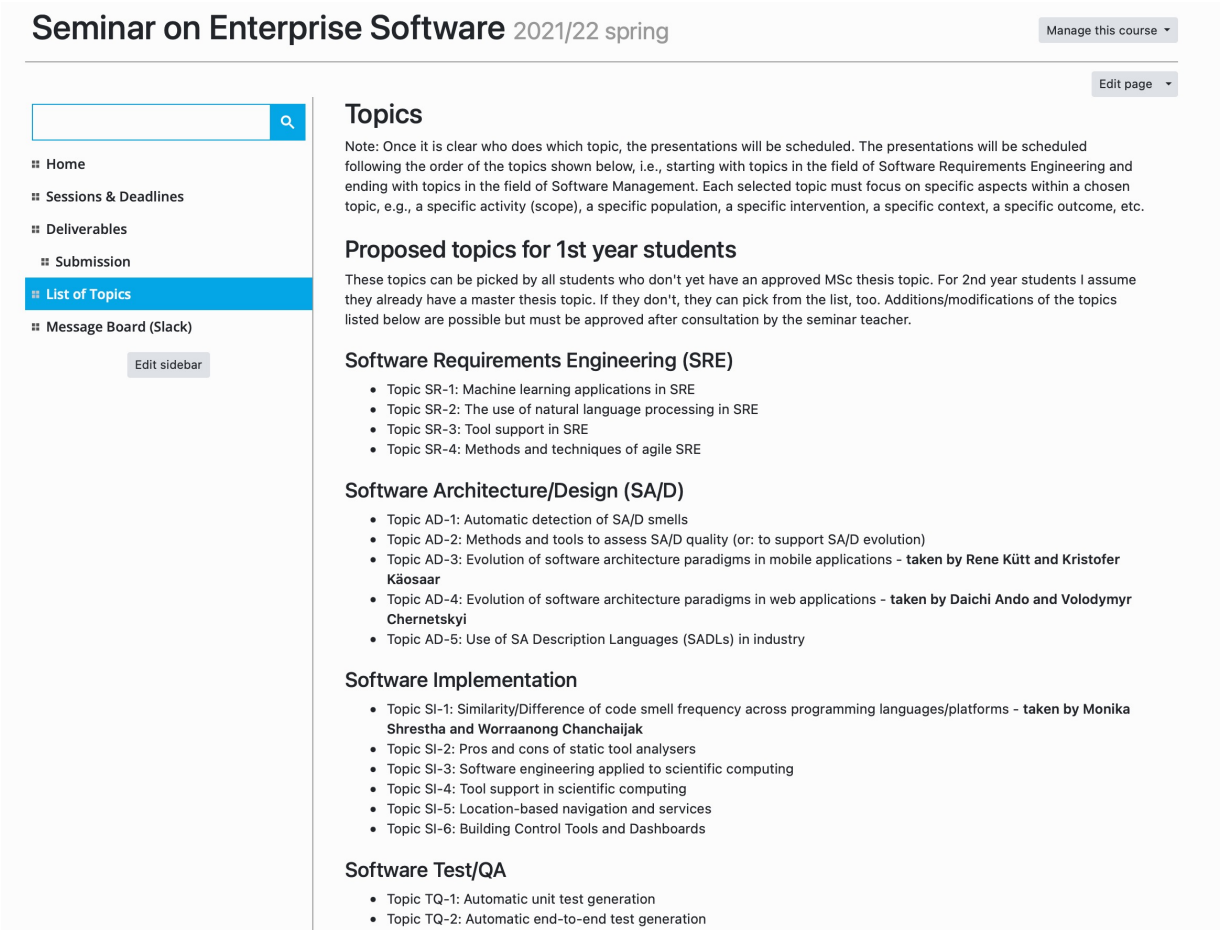
- Draft Report + Presentation Slides
  - For the report use the IEEE 2-column template (Word, LaTex, or Overleaf) for conference papers (A4 format): https://www.ieee.org/conferences/publishing/templates.html
  - Good example (posted with permission from Liisa Sakermann and Rain Hallikas):
    - Draft Report
    - Data Extraction Spreadsheet (filled)
    - Presentation Slides
  - Another good example (posted with permission from Janna-Liisa Leemets & Richard Õnnis):
    - Draft Report
    - Data Extraction Spreadsheet (filled)
    - Presentation Slides
- Presentation (during seminar session - 12 min)
- Final Report
  - For the report use the IEEE 2-column template (Word, LaTex, or Overleaf) for conference papers (A4 format): https://www.ieee.org/conferences/publishing/templates.html
- Presentation assessments (of all presentations on days where a student does not present himself/herself)

In order to not fail the course, all deliverables must be submitted on time (no extensions are granted). Regarding presentation assessments: Only one presentation day may be missed. Presentation assessments must be submitted via a quiz in Moodle.

Assessment criteria for deliverables and minimum conditions for passing the course successfully will be given in the lecture.

Topics + Research Questions (RQs)

- Pick from list on course wiki
- Refine according to interest
- Own proposal possible but must be approved
- Deadline: Feb 28 (incl. 4 RQs)



Seminar on Enterprise Software 2021/22 spring

Manage this course

Edit page

Home
Sessions & Deadlines
Deliverables
Submission
List of Topics
Message Board (Slack)
Edit sidebar

Topics

Note: Once it is clear who does which topic, the presentations will be scheduled. The presentations will be scheduled following the order of the topics shown below, i.e., starting with topics in the field of Software Requirements Engineering and ending with topics in the field of Software Management. Each selected topic must focus on specific aspects within a chosen topic, e.g., a specific activity (scope), a specific population, a specific intervention, a specific context, a specific outcome, etc.

Proposed topics for 1st year students

These topics can be picked by all students who don't yet have an approved MSc thesis topic. For 2nd year students I assume they already have a master thesis topic. If they don't, they can pick from the list, too. Additions/modifications of the topics listed below are possible but must be approved after consultation by the seminar teacher.

Software Requirements Engineering (SRE)

- Topic SR-1: Machine learning applications in SRE
- Topic SR-2: The use of natural language processing in SRE
- Topic SR-3: Tool support in SRE
- Topic SR-4: Methods and techniques of agile SRE

Software Architecture/Design (SA/D)

- Topic AD-1: Automatic detection of SA/D smells
- Topic AD-2: Methods and tools to assess SA/D quality (or: to support SA/D evolution)
- Topic AD-3: Evolution of software architecture paradigms in mobile applications - **taken by Rene Kütt and Kristofer Käosaar**
- Topic AD-4: Evolution of software architecture paradigms in web applications - **taken by Daichi Ando and Volodymyr Chernetskyi**
- Topic AD-5: Use of SA Description Languages (SADLs) in industry

Software Implementation

- Topic SI-1: Similarity/Difference of code smell frequency across programming languages/platforms - **taken by Monika Shrestha and Worraanong Chanchaijak**
- Topic SI-2: Pros and cons of static tool analysers
- Topic SI-3: Software engineering applied to scientific computing
- Topic SI-4: Tool support in scientific computing
- Topic SI-5: Location-based navigation and services
- Topic SI-6: Building Control Tools and Dashboards

Software Test/QA

- Topic TQ-1: Automatic unit test generation
- Topic TQ-2: Automatic end-to-end test generation

MSc Thesis Template

Guidelines Document
+
Thesis Template

- Link: https://cs.ut.ee/en/content/thesis-deadlines-and-guidelines
- The typical structure of a thesis consists of the following components:
  - Title page
  - Information sheet
  - Table of Contents
  - Introduction
  - Terms and notions (optional)
  - Chapters
  - Summary
  - References
  - Appendices (if relevant)
  - License

Introduction	Why do it? (Context / Motivation)
	What to achieve? (Goals)
Related Literature	What exists already?
Methodology	How to achieve goals?
Results	What was achieved?
Discussion	How well were goals achieved?
	What are the limitations?
Conclusions	Summary of achievements
	What next?

Methods -> Literature

- Systematic Literature Survey:
  - B.A. Kitchenham, S. Charters, Guidelines for Performing Systematic Literature Reviews in Software Engineering Technical Report EBSE-2007-01, 2007.
- *Mapping Study:*
  - *K. Petersen, R. Feldt, M. Shahid, M. Mattsson, Systematic Mapping Studies in Software Engineering. EASE 2008, 2008.*
- More materials on research methods in SE
  - See course wiki (Sessions & Deadlines):

https://courses.cs.ut.ee/2022/enterprise-seminar1/spring/Main/Seminars

Today's Program

- SLR and SMS examples
- Guidelines on how to write an SMS
- Tips & Tricks for report writing
- Tips & Tricks for presenting

- Examples of Systematic Literature Reviews (SLRs) in SE:
  - SLS Example 1 - Note that there is a small issue in the structure of Section 4.2
  - SLS Example 2
  - SLS Example 3
  - Example of a MSc Thesis containing an SLS
- Examples of Systematic Mapping Studies (SMSs) in SE:
  - SMS Example 1
  - SMS Example 2

Example SLR https://courses.cs.ut.ee/2022/enterprise-seminar1/spring/Main/Seminars

Software Development Analytics in Practice:
A Systematic Literature Review

João Caldeira[a], Fernando Brito e Abreu[a], Jorge Cardoso[b,c], Toacy Oliveira[d]

[a]*Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR-IUL, Lisboa, Portugal*
[b]*CISUC, Dept. of Informatics Engineering*
University of Coimbra, Portugal
[c]*Huawei Munich Research Center, Germany*
[d]*Systems Engineering and Computer Science Department*
Federal University of Rio de Janeiro, RJ, Brazil

Abstract

Context: Software Development Analytics is a research area concerned with providing insights to improve product deliveries and processes. Many types of studies, data sources and mining methods have been used for that purpose.

Objective: This systematic literature review aims at providing an aggregate view of the relevant studies on Software Development Analytics in the past decade (2010-2019), with an emphasis on its application in practical settings.

Method: Definition and execution of a search string upon several digital libraries, followed by a quality assessment criteria to identify the most relevant papers. On those, we extracted a set of characteristics (study type, data source, study perspective, development life-cycle activities covered, stakeholders, mining methods, and analytics scope) and classified their impact against a taxonomy.

Results: Source code repositories, experimental case studies, and developers are the most common data sources, study types, and stakeholders, respectively. Product and project managers are also often present, but less than expected. Mining methods are evolving rapidly and that is reflected in the long list identified. Descriptive statistics are the most usual method followed by correlation

Email addresses: jcppc@iscte-iul.pt (João Caldeira), fba@iscte-iul.pt (Fernando Brito e Abreu), jcardoso@dei.uc.pt (Jorge Cardoso), toacy@cos.ufrj.br (Toacy Oliveira)

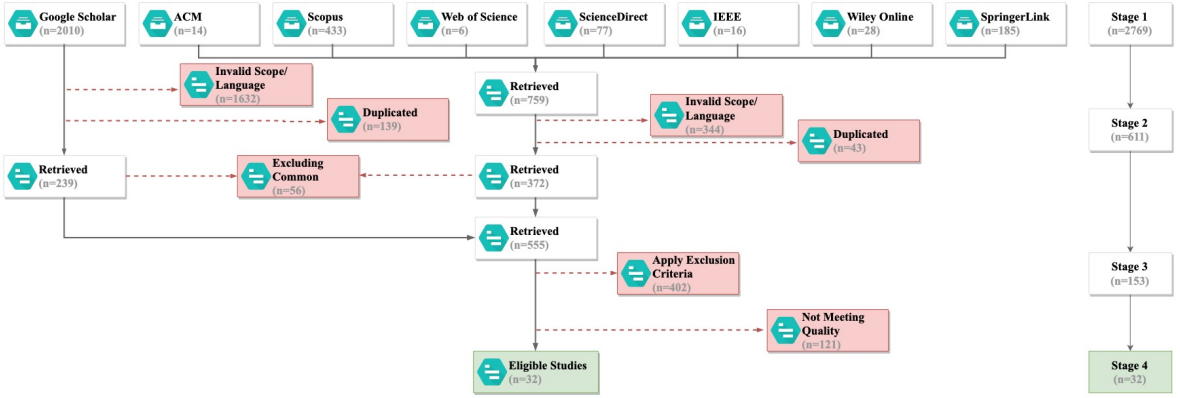


Figure 1: Study Selection Process Stages

- 8 data sources
- 8 research questions
- 6 exclusion criteria
- 3 inclusion criteria
- 13 quality criteria
- From 2769 primary studies in stage 1 down to 32 studies in stage 4 (those were analyzed in detail)

Example SLR https://courses.cs.ut.ee/2022/enterprise-seminar1/spring/Main/Seminars

IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 37, NO. 2, MARCH/APRIL 2011 283

Systematic Review and Aggregation of Empirical Studies on Elicitation Techniques

Oscar Dieste, *Member*, *IEEE*, and Natalia Juristo, *Senior Member*, *IEEE*

Abstract—We have located the results of empirical studies on elicitation techniques and aggregated these results to gather empirically grounded evidence. Our chosen surveying methodology was systematic review, whereas we used an adaptation of comparative analysis for aggregation because meta-analysis techniques could not be applied. The review identified 564 publications from the SCOPUS, IEEEXPLORE, and ACM DL databases, as well as Google. We selected and extracted data from 26 of those publications. The selected publications contain 30 empirical studies. These studies were designed to test 43 elicitation techniques and 50 different response variables. We got 100 separate results from the experiments. The aggregation generated 17 pieces of knowledge about the interviewing, laddering, sorting, and protocol analysis elicitation techniques. We provide a set of guidelines based on the gathered pieces of knowledge.

Index Terms—Elicitation methods, performance measures, experimentation, systematic literature review.

1 Introduction

Elicitation is the process by means of which a software analyst gathers information about the problem domain.

The analyst uses a series of analyst-user interaction mechanisms, called elicitation techniques, to acquire information. A very wide range of elicitation techniques have been proposed: interviews [27], protocol analysis [28], laddering [29], work groups [30], etc. Apart from these commonly known techniques, there are a great many variations [11], [31], [32], a lot of underused techniques [13], [33], as well as combinations of techniques [34], [35].

Some experiences suggest that elicitation techniques are more or less equivalent for simple and well-defined problems [36], [37]. For real problems though, a number of studies suggest that elicitation techniques are not interchangeable, and there are far-reaching differences with regard to what type of knowledge each technique can uncover [1], [4], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47]. Other aspects, like quantity of information or elicitation efficiency, are features that might distinguish one elicitation technique from another [6], [15], [24].

A number of authors have suggested the need to research the differences between elicitation techniques [15], [38], [39], [48]. Unfortunately, there are few either theoretical or empirical comparative studies analyzing the potential of one technique compared with another [15], [49], [50], [51]. Some studies have reviewed elicitation techniques [38], [51], [52], [53], [54], [55] in an attempt at organization and categorization. However, as they are separate studies and no effort has been made to combine their results, they are underexploited.

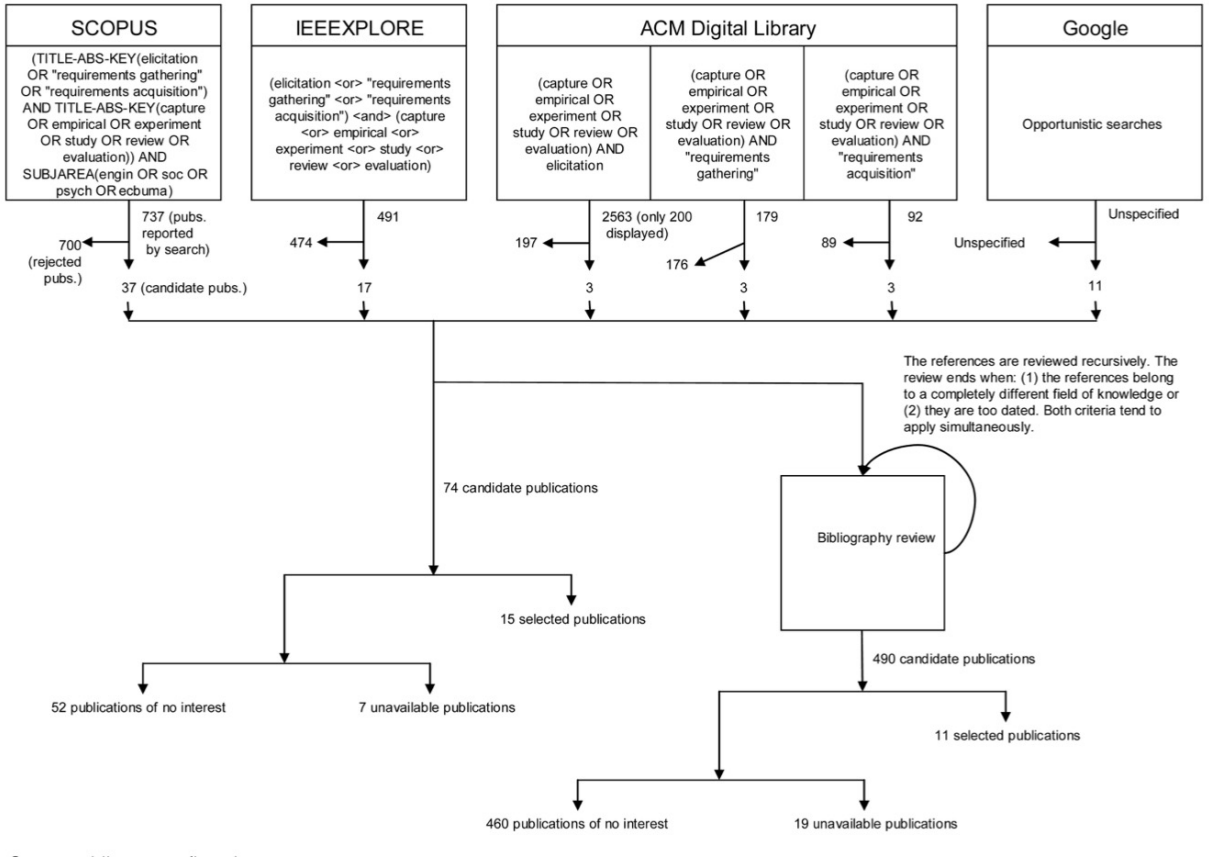
Some experts have proposed criteria about when to use certain elicitation techniques [56], [57], [58], [59], [60], [61]. These criteria are supported by the expert's personal experience or by theoretical concerns, but not by evidence. However, any aspect distinguishing one elicitation technique from another is hard to address theoretically. This is because the applicable theory provided by cognitive psychology [62] is not yet mature enough to be able to theoretically predict how effective elicitation techniques are.

The aim of this paper is to generate knowledge about the applicability of elicitation techniques by aggregating the results of existing empirical studies. The procedure recommended by [63] is used for review. A review protocol was put together as recommended in [63]. To build the protocol, we first ran a nonprotocoled pilot study. However, we had to make changes to [63] on several occasions because the proposed procedural steps were not always suitable for aggregating the identified empirical studies. The most marked differences are in the results aggregation procedure. We had to develop our own procedure, based on [64], [65], for results aggregation as the meta-analysis techniques suggested by [63] can only be applied under certain circumstances that do not generally hold in experimental SE. For this reason, this paper both describes the results of and details the tasks carried out in each review process activity, highlighting the deviations from [63]. For a very early review of this research, see [66].

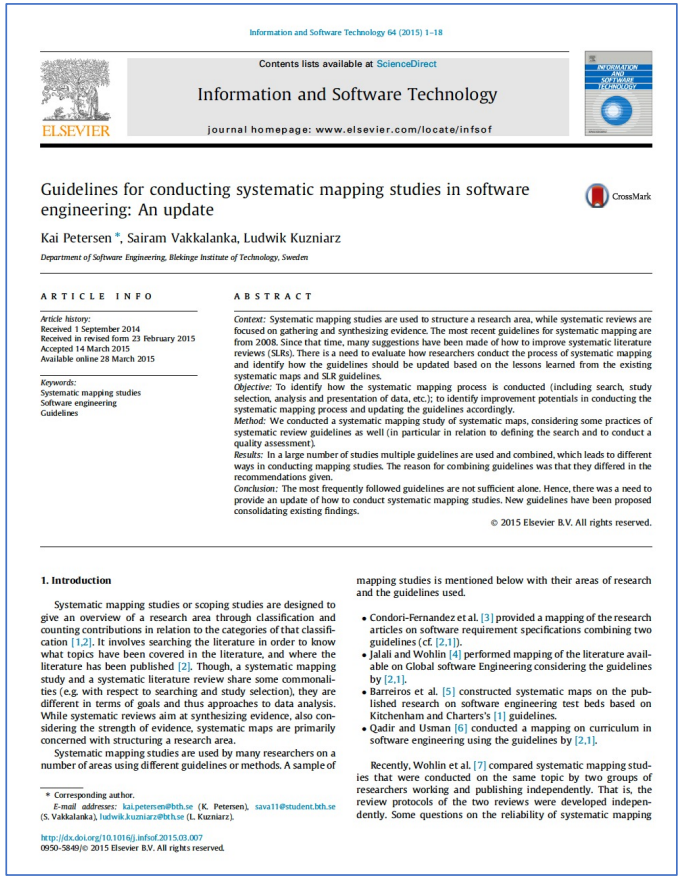
The paper is structured as follows: Section 2 describes the selected studies. Section 3 shows the empirical results extracted from the selected studies. Section 4 describes the aggregation procedure and Section 5 explains the mechanism for assessing the quality of the empirical studies and the sensitivity analysis run. Section 6 presents the aggregations. Section 7 describes a sensitivity analysis performed on the aggregations. Finally, Section 8 gives an overview of the new knowledge and Section 9 discusses how reliable this knowledge is.

• *The authors are with the Facultad de Informática, Universidad Politécnica de Madrid, Campus de Montegancedo, 28660 Boadilla del Monte, Spain. E-mail: {odieste, natalia}@fi.upm.es.*

Manuscript received 1 May 2009; revised 24 Sept. 2009; accepted 9 Dec. 2009; published online 11 Feb. 2010.
Recommended for acceptance by B. Nuseibeh.
For information on obtaining reprints of this article, please send e-mail to: tse@computer.org, and reference IEEECS Log Number TSE-2009-03-0052.
Digital Object Identifier no. 10.1109/TSE.2010.33.



Example SMS https://courses.cs.ut.ee/2022/enterprise-seminar1/spring/Main/Seminars



Information and Software Technology 64 (2015) 1–18

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Guidelines for conducting systematic mapping studies in software engineering: An update

Kai Petersen *, Sairam Vakkalanka, Ludwik Kuzniarz

Department of Software Engineering, Blekinge Institute of Technology, Sweden

ARTICLE INFO

Article history:
Received 1 September 2014
Received in revised form 23 February 2015
Accepted 14 March 2015
Available online 28 March 2015

Keywords:
Systematic mapping studies
Software engineering
Guidelines

ABSTRACT

Context: Systematic mapping studies are used to structure a research area, while systematic reviews are focused on gathering and synthesizing evidence. The most recent guidelines for systematic mapping are from 2008. Since that time, many suggestions have been made of how to improve systematic literature reviews (SLRs). There is a need to evaluate how researchers conduct the process of systematic mapping and identify how the guidelines should be updated based on the lessons learned from the existing systematic maps and SLR guidelines.

Objective: To identify how the systematic mapping process is conducted (including search, study selection, analysis and presentation of data, etc.); to identify improvement potentials in conducting the systematic mapping process and updating the guidelines accordingly.

Method: We conducted a systematic mapping study of systematic maps, considering some practices of systematic review guidelines as well (in particular in relation to defining the search and to conduct a quality assessment).

Results: In a large number of studies multiple guidelines are used and combined, which leads to different ways in conducting mapping studies. The reason for combining guidelines was that they differed in the recommendations given.

Conclusion: The most frequently followed guidelines are not sufficient alone. Hence, there was a need to provide an update of how to conduct systematic mapping studies. New guidelines have been proposed consolidating existing findings.

1. Introduction

Systematic mapping studies or scoping studies are designed to give an overview of a research area through classification and counting contributions in relation to the categories of that classification [1,2]. It involves searching the literature in order to know what topics have been covered in the literature, and where the literature has been published [2]. Though, a systematic mapping study and a systematic literature review share some commonalities (e.g. with respect to searching and study selection), they are different in terms of goals and thus approaches to data analysis. While systematic reviews aim at synthesizing evidence, also considering the strength of evidence, systematic maps are primarily concerned with structuring a research area.

Systematic mapping studies are used by many researchers on a number of areas using different guidelines or methods. A sample of mapping studies is mentioned below with their areas of research and the guidelines used.

- Condori-Fernandez et al. [3] provided a mapping of the research articles on software requirement specifications combining two guidelines (cf. [2,1]).
- Jalali and Wohlin [4] performed mapping of the literature available on Global software Engineering considering the guidelines by [2,1].
- Barreiros et al. [5] constructed systematic maps on the published research on software engineering test beds based on Kitchenham and Charters's [1] guidelines.
- Qadir and Usman [6] conducted a mapping on curriculum in software engineering using the guidelines by [2,1].

Recently, Wohlin et al. [7] compared systematic mapping studies that were conducted on the same topic by two groups of researchers working and publishing independently. That is, the review protocols of the two reviews were developed independently. Some questions on the reliability of systematic mapping

* Corresponding author.
E-mail addresses: kai.petersen@bth.se (K. Petersen), sava11@student.bth.se (S. Vakkalanka), ludwik.kuzniarz@bth.se (L. Kuzniarz).

http://dx.doi.org/10.1016/j.infsof.2015.03.007
0950-5849/

A mapping study of mapping studies ….

We will use this in the following to illustrate the steps of a mapping study

Today's Program

- SLR and SMS examples
- Guidelines on how to write an SMS
- Tips & Tricks for report writing
- Tips & Tricks for presenting

Homework (not graded):
Read the papers on SLS posted on course wiki, i.e.

Kai Petersen, Robert Feldt, Shahid Mujtaba, and Michael Mattsson (2008) Systematic mapping studies in software engineering. In Proceedings of the 12th international conference on Evaluation and Assessment in Software Engineering (EASE'08). BCS Learning & Development Ltd., Swindon, GBR, 68–77. [https://www.scienceopen.com/hosted-document?doi=10.14236/ewic/EASE2008.8]

Kai Petersen, Sairam Vakkalanka, Ludwik Kuzniarz (2015) Guidelines for conducting systematic mapping studies in software engineering: An update. Information and Software Technology, Volume 64, Pages 1-18. [https://doi.org/10.1016/j.infsof.2015.03.007]

SMS – Main Steps

1. Research Questions
2. Data Search (Database, Manual, Snowballing)
3. Study Selection and Quality Assessment
4. Data Extraction
5. Analysis and Classification

Information and Software Technology 64 (2015) 1–18

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Guidelines for conducting systematic mapping studies in software engineering: An update

Kai Petersen *, Sairam Vakkalanka, Ludwik Kuzniarz

Department of Software Engineering, Blekinge Institute of Technology, Sweden

ARTICLE INFO

Article history:
Received 1 September 2014
Received in revised form 23 February 2015
Accepted 14 March 2015
Available online 28 March 2015

Keywords:
Systematic mapping studies
Software engineering
Guidelines

ABSTRACT

Context: Systematic mapping studies are used to structure a research area, while systematic reviews are focused on gathering and synthesizing evidence. The most recent guidelines for systematic mapping are from 2008. Since that time, many suggestions have been made of how to improve systematic literature reviews (SLRs). There is a need to evaluate how researchers conduct the process of systematic mapping and identify how the guidelines should be updated based on the lessons learned from the existing systematic maps and SLR guidelines.

Objective: To identify how the systematic mapping process is conducted (including search, study selection, analysis and presentation of data, etc.); to identify improvement potentials in conducting the systematic mapping process and updating the guidelines accordingly.

Method: We conducted a systematic mapping study of systematic maps, considering some practices of systematic review guidelines as well (in particular in relation to defining the search and to conduct a quality assessment).

Results: In a large number of studies multiple guidelines are used and combined, which leads to different ways in conducting mapping studies. The reason for combining guidelines was that they differed in the recommendations given.

Conclusion: The most frequently followed guidelines are not sufficient alone. Hence, there was a need to provide an update of how to conduct systematic mapping studies. New guidelines have been proposed consolidating existing findings.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Systematic mapping studies or scoping studies are designed to give an overview of a research area through classification and counting contributions in relation to the categories of that classification [1,2]. It involves searching the literature in order to know what topics have been covered in the literature, and where the literature has been published [2]. Though, a systematic mapping study and a systematic literature review share some commonalities (e.g. with respect to searching and study selection), they are different in terms of goals and thus approaches to data analysis. While systematic reviews aim at synthesizing evidence, also considering the strength of evidence, systematic maps are primarily concerned with structuring a research area.

Systematic mapping studies are used by many researchers on a number of areas using different guidelines or methods. A sample of mapping studies is mentioned below with their areas of research and the guidelines used.

- Condori-Fernandez et al. [3] provided a mapping of the research articles on software requirement specifications combining two guidelines (cf. [2,1]).
- Jalali and Wohlin [4] performed mapping of the literature available on Global software Engineering considering the guidelines by [2,1].
- Barreiros et al. [5] constructed systematic maps on the published research on software engineering test beds based on Kitchenham and Charters's [1] guidelines.
- Qadir and Usman [6] conducted a mapping on curriculum in software engineering using the guidelines by [2,1].

Recently, Wohlin et al. [7] compared systematic mapping studies that were conducted on the same topic by two groups of researchers working and publishing independently. That is, the review protocols of the two reviews were developed independently. Some questions on the reliability of systematic mapping

* Corresponding author.
E-mail addresses: kai.petersen@bth.se (K. Petersen), sava11@student.bth.se (S. Vakkalanka), ludwik.kuzniarz@bth.se (L. Kuzniarz).

http://dx.doi.org/10.1016/j.infsof.2015.03.007
0950-5849/© 2015 Elsevier B.V. All rights reserved.

Based on:

Kai Petersen, Sairam Vakkalanka, Ludwik Kuzniarz (2015) Guidelines for conducting systematic mapping studies in software engineering: An update. Information and Software Technology, Volume 64, Pages 1-18. [https://doi.org/10.1016/j.infsof.2015.03.007]

Method
(Section 3 in the paper)

Information and Software Technology 64 (2015) 1–18

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Guidelines for conducting systematic mapping studies in software engineering: An update

Kai Petersen*, Sairam Vakkalanka, Ludwik Kuzniarz

Department of Software Engineering, Blekinge Institute of Technology, Sweden

ARTICLE INFO

Article history:
Received 1 September 2014
Revised and accepted 12 February 2015
Accepted 14 March 2015
Available online 28 March 2015

Keywords:
Systematic mapping studies
Evidence engineering
Guidelines

ABSTRACT

Context: Systematic mapping studies are used to structure a research area, while systematic reviews are focused on gathering and synthesizing evidence. The most recent guidelines for systematic mapping are from 2008. Since then researchers conducting mapping studies have gained more experience, leading to updated guidelines and improved methods for conducting systematic mapping studies.
Objective: To identify how the systematic mapping process is conducted (including search, study selection, analysis and presentation of the results); to propose improvement potentials to conducting the systematic mapping process and updating the guidelines accordingly.
Method: We conducted a systematic mapping study of systematic maps, considering some practices of systematic review guidelines as well (in particular in relation to defining the search and to conduct a quality assessment).
Results: In a large number of studies multiple guidelines are used and combined, which leads to different ways in conducting mapping studies. The reason for combining guidelines was that they differed in the aspects they addressed.
Conclusion: The most frequently followed guidelines are not sufficient alone. Hence, there was a need to provide an update of how to conduct systematic mapping studies. New guidelines have been proposed consolidating existing findings.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Systematic mapping studies or scoping studies are designed to give an overview of a research area through classification and counting contributions in relation to the categories of that classification [1]. These categories can, for example, be focuses of research, what topics have been covered in the literature, and where the literature has been published [2]. Though, a systematic mapping study and a systematic literature review share some commonalities (e.g. with respect to searching and study selection), they are different in terms of their goals, and consequently the methods. While systematic reviews aim at synthesizing evidence, also considering the strength of evidence, systematic maps are primarily concerned with structuring a research area.

Systematic mapping studies are used by many researchers on a number of areas using different guidelines or methods. A sample of mapping studies is mentioned below with their areas of research and the guidelines used.

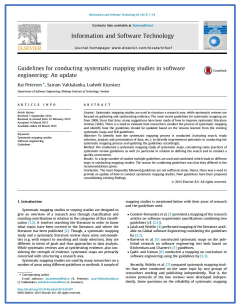
- Condori-Fernandez et al. [3] provided a mapping of the research articles on software requirement specifications combining two guidelines [4,5].
- Jalali and Wohlin [6] performed mapping of the literature available on Global software Engineering considering the guidelines by [2,7].
- Barreiros et al. [8] constrained systematic maps on the publication trends in software testing. Their work was both based on Kitchenham and Charters [4,9] guidelines.
- Qadir and Usman [9] conducted a mapping on curriculum in software engineering using the guidelines by [2,7].

Recently, Wohlin et al. [7] compared systematic mapping studies that were conducted on the same topic by two groups of researchers working and publishing independently. That is, the review protocols of the two reviews were developed independently. Some questions on the reliability of systematic mapping

* Corresponding author.
E-mail addresses: kai.petersen@bth.se (K. Petersen), sairam.vakkalanka@bth.se (S. Vakkalanka), ludwik.kuzniarz@bth.se (L. Kuzniarz).

http://dx.doi.org/10.1016/j.infsof.2015.03.007
0950-5849/© 2015 Elsevier B.V. All rights reserved.

SMS-Step 1: Research Questions



- Overall goal: …to determine how systematic mapping processes have been executed in software engineering.

- Research Questions (RQs):
  - RQ1: Which guidelines are followed to conduct the systematic mapping studies in software engineering?
  - RQ2: Which software engineering topics are covered?
  - RQ3: Where and when were mapping studies published?
  - RQ4: How was the systematic mapping process performed?

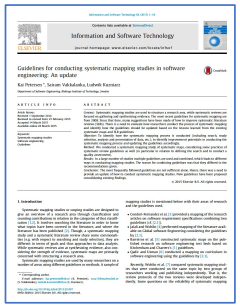
    This includes, for example:

    – Identification of studies (search, inclusion and exclusion)

    – Categorization and Classification schemes and processes

    – Visualization of results

SMS-Step 1: Research Questions



- Overall goal: …to determine how systematic mapping processes have been executed in software engineering.

- Research Question
  - RQ1: Which guid
    software engine
  - RQ2: Which soft
  - RQ3: Where and
  - RQ4: How was th

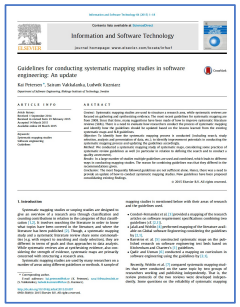
    This inc
    – Identi
    – Categ
    – Visua

Important to remember:

In your Master's Thesis, the RQs for the SMS (used to find related work) are usually not exactly the same as the RQs of your thesis topic.

They are only the same if your thesis itself is an SMS (or SLR).

SMS-Step 1: Research Questions



- Overall goal: …to determine how systematic mapping processes have been executed in software engineering.

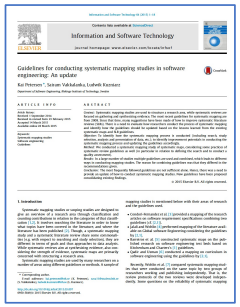
- Research Question
  - RQ1: Which guid
    software engine
  - RQ2: Which soft
  - RQ3: Where and
  - RQ4: How was th

    This inc
    – Identi
    – Categ
    – Visua

How to find RQs for SMS?

Typically, RQs for SMS examine the extent, range and nature of a research activity of interest. In the context of SE, ”research activity” may refer to a class of SE methods (processes), teqchniques, and tools.

SMS-Step 1: Research Questions



- Overall goal: …to determine how systematic mapping processes have been executed in software engineering.

- Research Questions (RQs):
  - RQ1: Which guidelines are followed to conduct the systematic mapping studies in software engineering?
  - RQ2: Which software engineering topics are covered?
  - RQ3: Where and when were mapping studies published?
  - RQ4: How was the systematic mapping process performed?

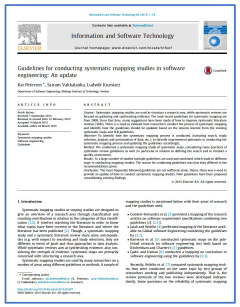
    This includes, for example:

    – Identification of studies (search, inclusion and exclusion)

    – Categorization and Classification schemes and processes

    – Visualization of results

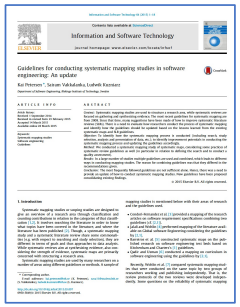
SMS-Step 2: Data Search (Database, Manual, Snowballing)



Used PICO (Population, Intervention, Comparison and Outcomes) to identify keywords to be used in search strings

- Population: may refer to specific software engineering role, category of software engineer, an application area or an industry group. Here: the population consists of systematic mapping/scoping studies.
- Intervention: refers to a software methodology, tool, technology, or procedure. Here: no clear intervention to be investigated (in addition to SMS)
- Comparison: Here we compare the different processes of conducting the maps by means of identifying the different strategies that have been used by them. Though, no empirical comparison is made, at this stage alternative strategies are identified.
- Outcomes: Here: No measurable outcome is considered, as we do not focus on empirical studies evaluating systematic mapping
- Context: Software engineering – Note: The authors call it “scope”

SMS-Step 2: Data Search (**Database**, Manual, Snowballing)

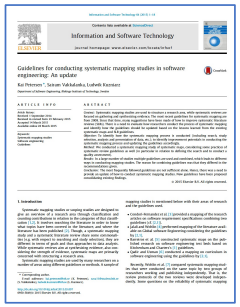


Used PICOC (Population, Intervention, Comparison, Outcomes, and Context) to identify keywords to be used in search strings

- Population
- Intervention
- Comparison
- Outcomes
- Context

  (Note:
  'Context' is not used in the example paper)

SMS-Step 2: Data Search (**Database**, Manual, Snowballing)



Used PICOC (Population, Intervention, Comparison, Outcomes, and Context) to identify keywords to be used in search strings

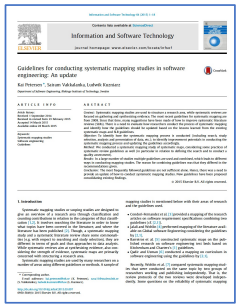
- Population
- Intervention
- Comparison
- Outcomes
- Context

  (Note: 'Context' is not used in the example paper)

In software engineering experiments, the populations might be any of the following:

- A specific software engineering role e.g. testers, managers.
- A category of software engineer, e.g. a novice or experienced engineer.
- An application area e.g. IT systems, command and control systems.
- An industry group such as Telecommunications companies, or Small IT companies.

SMS-Step 2: Data Search (**Database**, Manual, Snowballing)



Used PICOC (Population, Intervention, Comparison, Outcomes, and Context) to identify keywords to be used in search strings

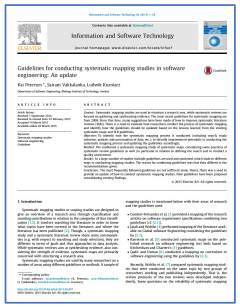
- Population
- Intervention
- Comparison
- Outcomes
- Context

(Note:
'Context' is not used in the example paper)

The intervention is the software methodology/tool/technology/procedure that addresses a specific issue

Examples:
technologies to perform specific tasks such as requirements specification, system testing, or software cost estimation

SMS-Step 2: Data Search (**Database**, Manual, Snowballing)



Used PICOC (Population, Intervention, Comparison, Outcomes, and Context) to identify keywords to be used in search strings

- Population
- Intervention
- Comparison
- Outcomes
- Context

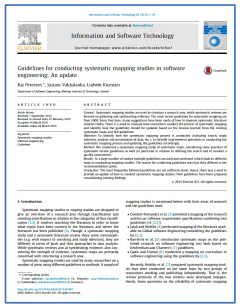
(Note: 'Context' is not used in the example paper)

This is the software engineering methodology/tool/technology/procedure with which the intervention is being compared

- When the comparison technology is the conventional or commonly-used technology, it is often referred to as the "control" treatment
- The control situation must be adequately described; "not using the intervention" is inadequate as a description of the control treatment

If training is required, beware of confounding factors (i.e., training effect versus effect of the technique)

SMS-Step 2: Data Search (**Database**, Manual, Snowballing)



Used PICOC (Population, Intervention, Comparison, Outcomes, and Context) to identify keywords to be used in search strings

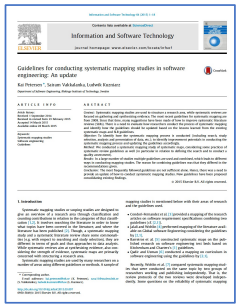
- Population
- Intervention
- Comparison
- Outcomes
- Context

(Note: 'Context' is not used in the example paper)

Outcomes should relate to factors of importance to practitioners such as improved reliability, reduced production costs, reduced time to market, …

- All relevant outcomes must be specified; in some cases we require interventions that improve some aspect of software production without affecting another, e.g., improved reliability with no increase in cost
- A problem is the use of surrogate measures (e.g., defects found during system testing as a surrogate for quality, or coupling measures for design quality; using surrogate measures may be misleading and conclusions based on such studies may be less robust

SMS-Step 2: Data Search (**Database**, Manual, Snowballing)



Used PICOC (Population, Intervention, Comparison, Outcomes, and Context) to identify keywords to be used in search strings

- Population
- Intervention
- Comparison
- Outcomes
- Context

  (Note: 'Context' is not used in the example paper)

This is the context in which the comparison takes place (e.g. academia or industry), the participants taking part in the study (e.g. practitioners, academics, consultants, students), and the tasks being performed (e.g. small scale, large scale)

- Note: Many software experiments take place in academia using student participants and small scale tasks; such experiments might not be representative of what might occur with practitioners working in industry

Exercise – PICOC

- Population?
- Intervention?
- Comparison?
- Outcome?

What are the PICOC criteria in the following RQs?

- RQ 1: 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?

- RQ 2: 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?

Note: "study" refers to the primary study reported in the literature that is reviewed.

Exercise - PICOC

- Population?
- Intervention?
- Comparison?
- Outcome?

What are the PICOC criteria in the following RQs?

- RQ 1: 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?

Population

Intervention

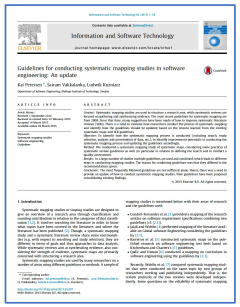
Comparison

- RQ 2: 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?

Outcome

Note: "study" refers to the primary study reported in the literature that is reviewed.

SMS-Step 2: Data Search (Database, Manual, Snowballing)

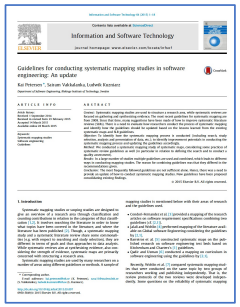


In the example SMS, based on the RQs and PICO(C), the identified keywords are *guidelines*, *systematic mapping* and *software engineering* which were grouped into sets and their synonyms were considered to formulate the search string.

- Set 1: Scoping the search for software engineering, i.e. ‘‘software engineering’’.
- Set 2: Search terms directly related to the intervention, e.g. ‘‘systematic mapping’’, and ‘‘systematic maps’’.
- Set 3: Search terms related to the process of classification and categorization, e.g. “methods”, “tools”, “classification”, “framework”.

Each set will be connected with ‘AND’ in the search string …

SMS-Step 2: Data Search
(Database, Manual, Snowballing)



Data sources

+

Search Strings

Database	Search
IEEE Xplore	(("Systematic mapping" or "systematic map" or "systematic mapping study" or "systematic mapping studies" or "systematic maps") AND (("Methods" or "framework" or "model" or "practice") OR ("tools" OR "tool" OR "techniques") OR ("categorization" OR "classification" OR "grouping") OR ("guidelines" OR "rules")) AND "software engineering")
ACM DL	(("Systematic mapping" OR "systematic map" OR "systematic mapping study" OR "systematic mapping studies") AND software engineering)
Scopus	("Systematic mapping" OR "systematic map" OR "systematic mapping study" OR "systematic mapping studies" OR "systematic maps") AND ("software engineering") OR (model OR method OR approach OR tools OR tool OR techniques OR framework OR practice OR classification OR categorization OR process OR guidelines OR rules OR strategy OR way)
Inspec / Compendix	((((Systematic mapping OR systematic map OR systematic mapping study OR systematic mapping studies OR systematic maps) WN All fields) AND (((Methods OR framework OR model OR practice) OR (tools OR tool OR techniques) OR (categorization OR classification OR grouping) OR (guidelines OR rules)) WN All fields)) AND ((software engineering) WN All fields))

Number of studies per database:

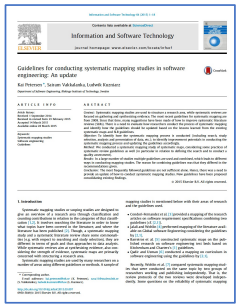
→ 5610

→ 360

→ 1215

→ 567

SMS-Step 2: Data Search
(Database, Manual, Snowballing)



Data sources

+

Search Strings

Database	Search
IEEE Xplore	((“Systematic mapping” or “systematic map” or “systematic mapping study” or “systematic mapping studies” or “systematic maps”) AND ((“Methods” or “framework” or “model” or “practice”) OR (“tools” OR “tool” OR “techniques”) OR (“categorization” OR “classification” OR “grouping”) OR (“guidelines” OR “rules”)) AND “software engineering”)
ACM DL	((“Systematic mapping” OR “systematic map” OR ystematic mapping g)
	natic map” OR “systematic apping studies” OR e engineering”) OR (model OR tool OR techniques OR cation OR categorization ; OR strategy OR way)
	natic map OR systematic pping studies OR ND (((Methods OR OR (tools OR tool OR classification OR grouping) OR (guidelines OR rules)) WN All fields)) AND ((software engineering) WN All fields))

Software Engineering Digital Libraries

- **IEEExplore**
- **ACM Digital library**
- **SpringerLink**
- **Google scholar** (scholar.google.com)
- **Citeseer library** (citeseer.ist.psu.edu)
- **Inspec** (www.iee.org/Publish/INSPEC/)
- **ScienceDirect** (www.sciencedirect.com)
- **EI Compendex** (www.engineeringvillage2.org/Controller/Servlet/AthensService)

Number of studies per database:

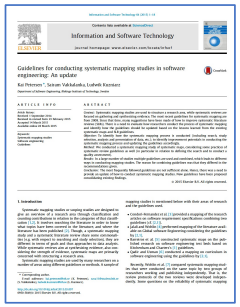
→ 5610

→ 360

→ 1215

→ 567

SMS-Step 2: Data Search
(Database, Manual, Snowballing)



Data sources

+

Search Strings

Dedicated SE sources:
- IEEE Transactions on Software Engineering (J)
- ACM Transactions on Software Engineering Methodology (J)
- Empirical Software Engineering (J)
- Information and Software Technology (J)
- Journal of Software: Evolution and Process (J)
- Journal of Systems and Software (J)
- International Conference on Software Engineering (C)
- Evaluation and Assessment in Software Engineering (C)

Etc.

+ Gray literature (blogs, PhD and MSc theses, etc.)

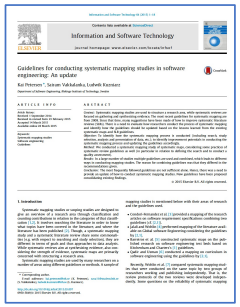
Software Engi
- **IEEExplore**
- **ACM Digital library**
- **SpringerLink**
- **Google scholar** (scholar.g
- **Citeseer library** (citeseer.ist.psu.edu)
- **Inspec** (www.iee.org/Publish/INSPEC/)
- **ScienceDirect** (www.sciencedirect.com)
- **EI Compendex** (www.engineeringvillage2.org/Controller/Servlet/AthensService)

; OR strategy OR way)

ıatic map OR systematic
ɔping studies OR
ND (((Methods OR
OR (tools OR tool OR
classification OR
grouping) OR (guidelines OR rules)) WN All fields)) AND
((software engineering) WN All fields))

567

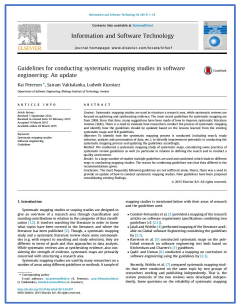
SMS-Step 3: Study Selection & Quality Assessment



Inclusion Criteria (applied to titles and abstracts):

- Studies present the research method and result of a systematic mapping study.
- Studies are in the field of software engineering.
- Studies were published online in the time frame 2004 (first systematic review and evidence-based software engineering guidelines were published) to 2012 (the study was conducted during 2013).

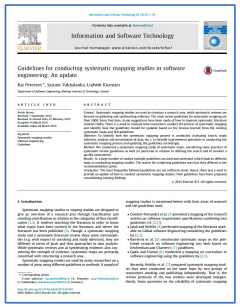
SMS-Step 3: Study Selection & Quality Assessment



Exclusion Criteria:

- Studies presenting summaries of conferences/editorials or guidelines/templates for conducting mapping studies.
- Studies presenting non-peer reviewed material.
- Studies not presented in English.
- Studies not accessible in full-text.
- Books and gray literature.
- Studies that are duplicates of other studies.

SMS-Step 3: Study Selection & Quality Assessment



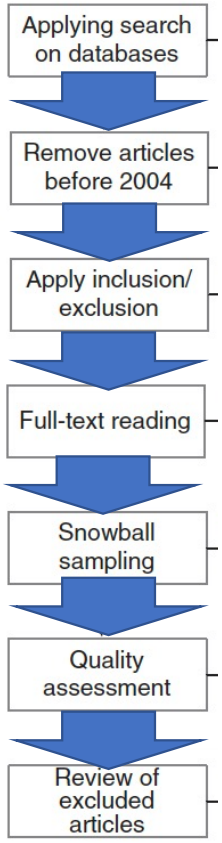
Quality Criteria applied to identified primary studies:

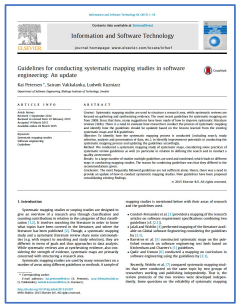
- Is the motivation for conducting systematic mapping clearly stated?
- Is the process of conducting systematic mapping clearly defined (study identification, data extraction, classification)?
- Is there any empirical evidence for the defined mapping process?
  - This question is concerned with whether the results of the mapping study are presented. That is, studies focusing on evaluating mapping studies without presenting their results are excluded.

Note: many more quality checks on the selected publications could be made, i.e., focusing on all kinds of biases

SMS-Step 3: Study Selection & Quality Assessment

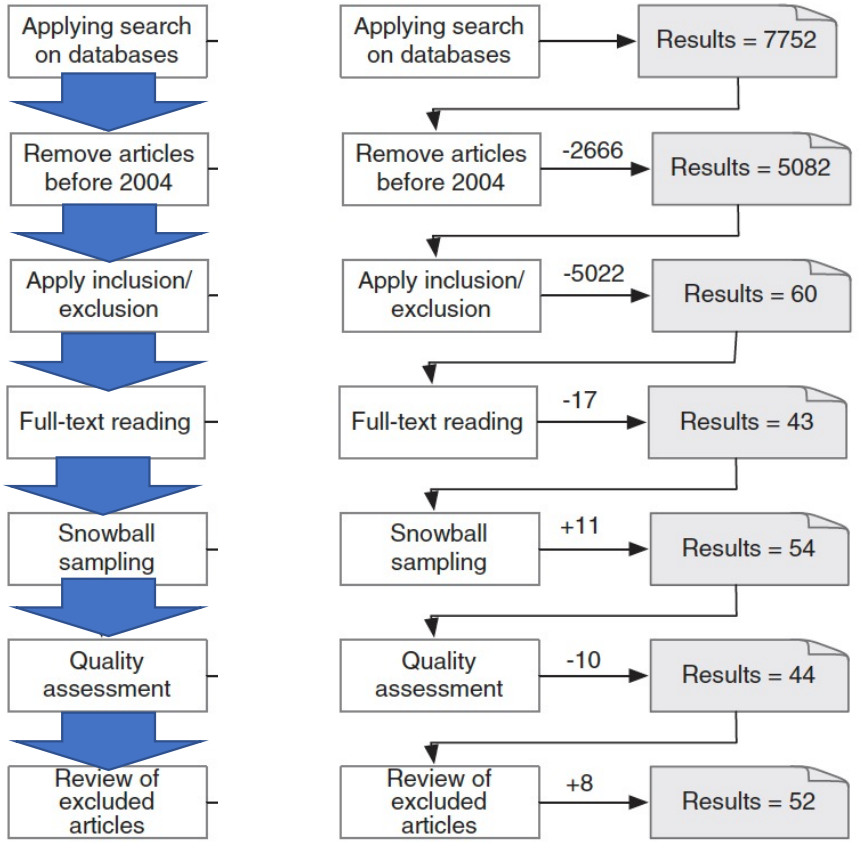
- Results of the identification process (incl. quality assessment)

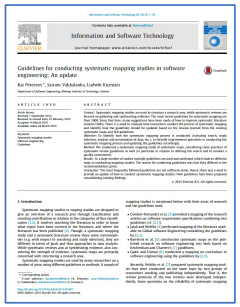




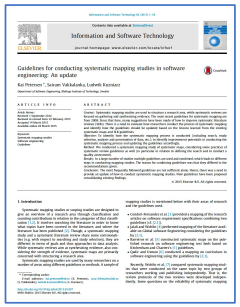
SMS-Step 3: Study Selection & Quality Assessment

- Results of the identification process (incl. quality assessment)



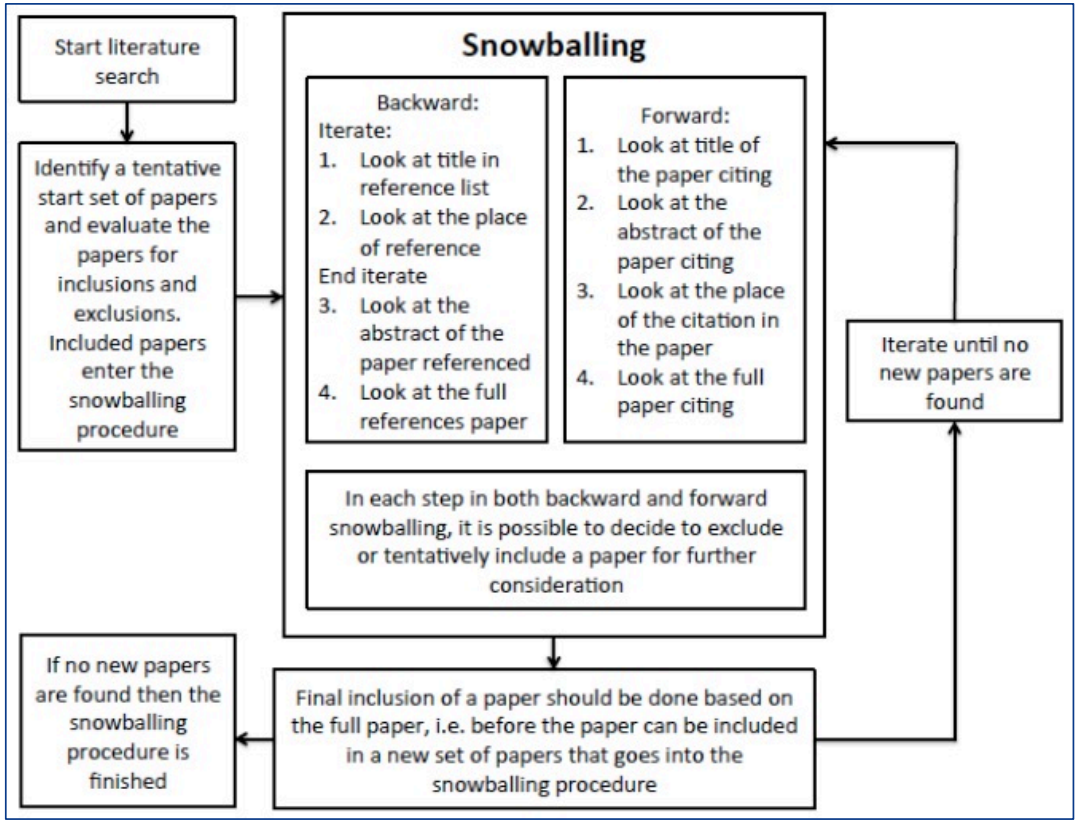


SMS-Step 3: Study Selection & Quality Assessment

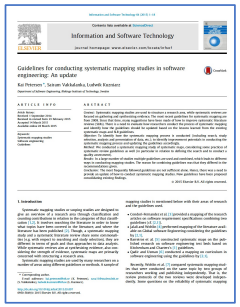


Snowballing:

- Backward
- Forward



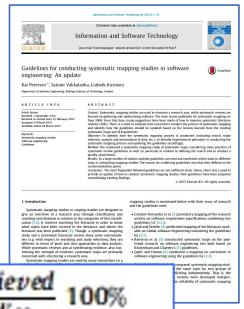
SMS-Step 3: Study Selection & Quality Assessment



Quality assessment of the identification process:

- If there are several persons doing the SMS have them review the results of each step
  - Applies to: formation and application of search strings, application of inclusion/exclusion criteria, data extraction and analysis, etc.
- Check for sensitivity & precision of search results

SMS-Step 3: Study Selection & Quality Assessment

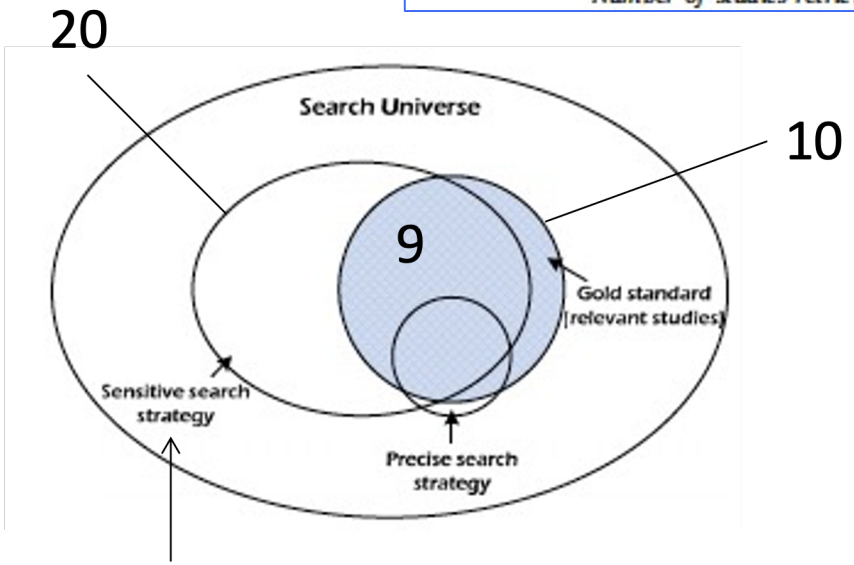


$$Sensitivity = \frac{Number\ of\ relevant\ studies\ retrieved}{Total\ number\ of\ relevant\ studies}\ 100\%$$

$$Precision = \frac{Number\ of\ relevant\ studies\ retrieved}{Number\ of\ studies\ retrieved}\ 100\%$$

Check Sensitivity & Precision

- Assume there are 10 relevant studies (= gold standard)
- Assume you have automatically detected (using search string) 20 studies
- After analysis, 9 of the 20 selected studies are in the gold standard (i.e., they are relevant)

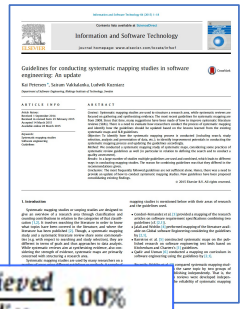


Sensitivity = 9/10 = 90%

Precision = 9/20 = 45%

This looks like a sensitive search strategy.

SMS-Step 3: Study Selection & Quality Assessment

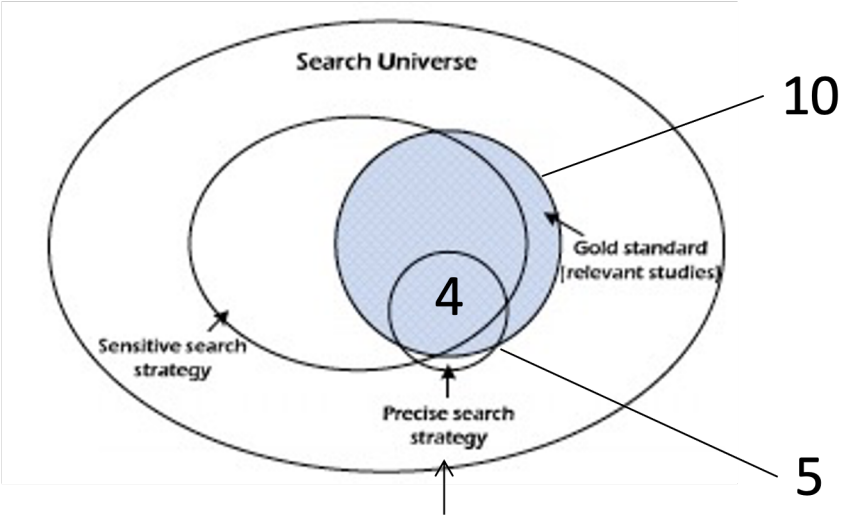


$$Sensitivity = \frac{Number\ of\ relevant\ studies\ retrieved}{Total\ number\ of\ relevant\ studies}\ 100\%$$

$$Precision = \frac{Number\ of\ relevant\ studies\ retrieved}{Number\ of\ studies\ retrieved}\ 100\%$$

Check Sensitivity & Precision

- Assume there are 10 relevant studies (= gold standard)
- Assume you have automatically detected (using search string) 5 studies
- After analysis, 4 of the 5 selected studies are in the gold standard (i.e., they are relevant)

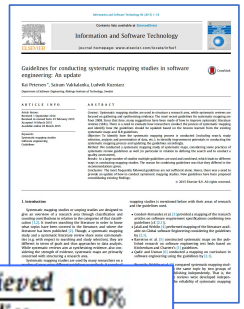


Sensitivity = 4/10 = 40%

Precision = 4/5 = 80%

This looks like a precise search strategy.

SMS-Step 3: Study Selection & Quality Assessment



$$Sensitivity = \frac{Number\ of\ relevant\ studies\ retrieved}{Total\ number\ of\ relevant\ studies} 100\%$$

$$Precision = \frac{Number\ of\ relevant\ studies\ retrieved}{Number\ of\ studies\ retrieved} 100\%$$

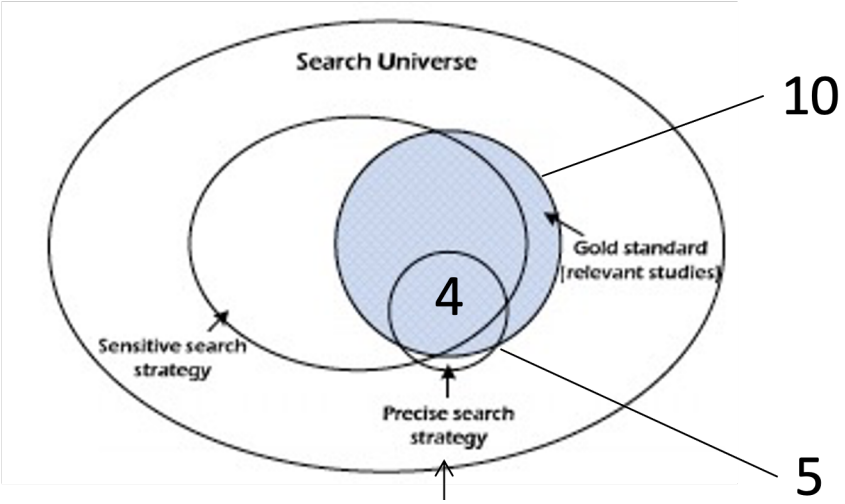
Problem:

You don't know the true gold standard

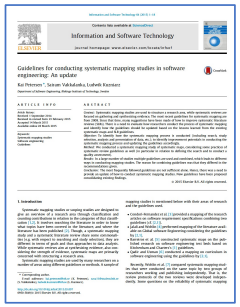
=>

Solution:

Use pseudo-gold standard (= set of known relevant studies found manually)



SMS-Step 4: Data Extraction

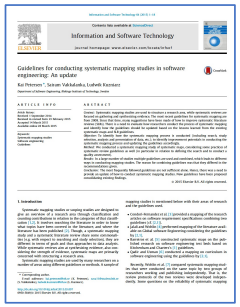


- Data Extraction Items:

SWEBOK =
Software Engineering
Body of Knowledge

Data item	Value	RQ
General		
Study ID	Integer	
Article Title	Name of the article	
Author Name	Set of Names of the authors	
Year of Publication	Calendar year	RQ3
Area in SE	Knowledge areas in SWEBOK	RQ2
Venue	Name of publication venue	RQ3
Process		
Guidelines	Which guidelines were adopted	RQ1
Search strategy	What search strategy is followed, and how were studies selected	RQ4
Search type	Manual or automated or both	RQ4
Classification schemes	How were articles classified	RQ4
Visualization type	What visualization types were used in order to present the data in a pictorial manner	RQ4

SMS-Step 4: Data Extraction



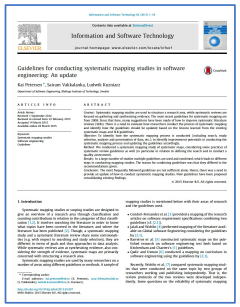
- Data Extraction Sheet:

RQ1: Which guidelines are followed to conduct the systematic mapping studies in software engineering?
RQ2: Which software engineering topics are covered?
RQ3: Where and when were mapping studies published?
RQ4: How was the systematic mapping process performed?

Study ID	Art. Title	Auth. Names	RQ1: Guidelines	RQ2: SE Area	RQ3: Pub. Venue	RQ3: Pub. Year	RQ4: Search Strategy	RQ4: manual / automatic	RQ4: Classif. Strategy	RQ4: Visual. Type
…										
…										

For the purpose of this seminar: 8 to 10 papers are sufficient (no need to be complete)

SMS-Step 5: Analysis and Classification



Information and Software Technology

Guidelines for conducting systematic mapping studies in software engineering: An update

Kai Petersen *, Sairam Vakkalanka, Ludwik Kuzniarz

Results Section

Results
(Section 4 in the paper)

Information and Software Technology 64 (2015) 1–18

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Guidelines for conducting systematic mapping studies in software engineering: An update

Kai Petersen *, Sairam Vakkalanka, Ludwik Kuzniarz

Department of Software Engineering, Blekinge Institute of Technology, Sweden

ARTICLE INFO

Article history:
Received 1 September 2014
Revised 11 December 2014, 31 February 2015
Accepted 13 March 2015
Available online 28 March 2015

Keywords:
Systematic mapping studies
Software engineering
Guidelines

ABSTRACT

Context: Systematic mapping studies are used to structure a research area, while systematic reviews are focused on gathering and synthesizing evidence. The most recent guidelines for systematic mapping are from 2008. Since that time, many suggestions have been made of how to improve systematic literature reviews (SLRs). There is a need to evaluate how researchers conduct the process of systematic mapping and update the guidelines based on the lessons learned from the existing systematic maps and SLR guidelines.

Objective: To identify how the systematic mapping process is conducted (including search, study selection, analysis and presentation of the results); to propose improvement potentials to conducting the systematic mapping process and updating the guidelines accordingly.

Method: We conducted a systematic mapping study of systematic maps, considering some practices of systematic review guidelines as well (in particular in relation to defining the search and to conduct a quality assessment).

Results: A large number of studies used/cite guidelines are used and considered, which leads to different ways in conducting mapping studies. The reasons for combining guidelines was that they differed in the recommendations given.

Conclusion: The existing frequently followed guidelines are not sufficient alone. Hence, there was a need to provide an update of how to conduct systematic mapping studies. New guidelines have been proposed consolidating existing findings.

1. Introduction

Systematic mapping studies or scoping studies are designed to give an overview of a research area through classification and counting contributions in relation to the categories of that classification [1]. These categories can, for example, be used to describe what topics have been covered in the literature, and where the literature has been published [2]. Though, a systematic mapping study and a systematic literature review share some commonalities (e.g. with respect to searching and study selection), they are different in their goal, and consequently approaches to data analysis. While systematic reviews aim at synthesizing evidence, also considering the strength of evidence, systematic maps are primarily concerned with structuring a research area.

Systematic mapping studies are used by many researchers on a number of areas using different guidelines or methods. A sample of mapping studies is mentioned below with their areas of research and the guidelines used.

- Condori-Fernandez et al. [3] provided a mapping of the research articles on software requirement specifications combining two guidelines [4,5].
- Jalali and Wohlin [6] performed mapping of the literature available on Global software Engineering considering the guidelines by [2,5].
- Barreiros et al. [9] constructed systematic maps on the publication trends in software engineering and built bases on Kitchenham and Charters [4] guidelines.
- Qadir and Usman [8] conducted a mapping on curriculum in software engineering using the guidelines by [2,5].

Recently, Wohlin et al. [7] compared systematic mapping studies that were conducted on the same topic by two groups of researchers working and publishing independently. That is, the review protocols of the two reviews were developed independently. Some questions on the reliability of systematic mapping

* Corresponding author.
E-mail addresses: kai.petersen@bth.se (K. Petersen), sairam.vakkalanka@bth.se (S. Vakkalanka), ludwik.kuzniarz@bth.se (L. Kuzniarz).

http://dx.doi.org/10.1016/j.infsof.2015.03.007
0950-5849/

SMS: Results for RQ1 (Guidelines)

- Review process
  - Guidelines
  - ~~Study identification~~
  - ~~Quality evaluation~~
  - ~~Data extraction and classification~~
  - ~~Visualization~~
  - ~~Validity threats~~

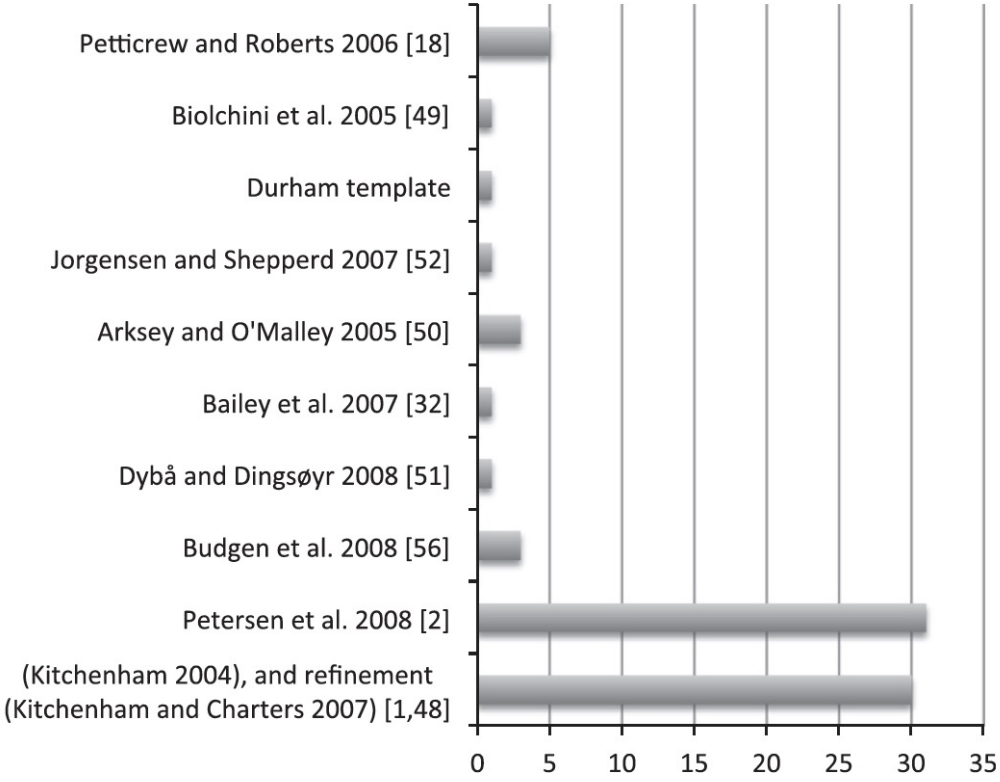


Fig. 5. Guidelines.

SMS: Results for RQ2 (Topics)

- Topics

All categories covered …
… but gaps in Education and SW configuration management

SWEBOK categories

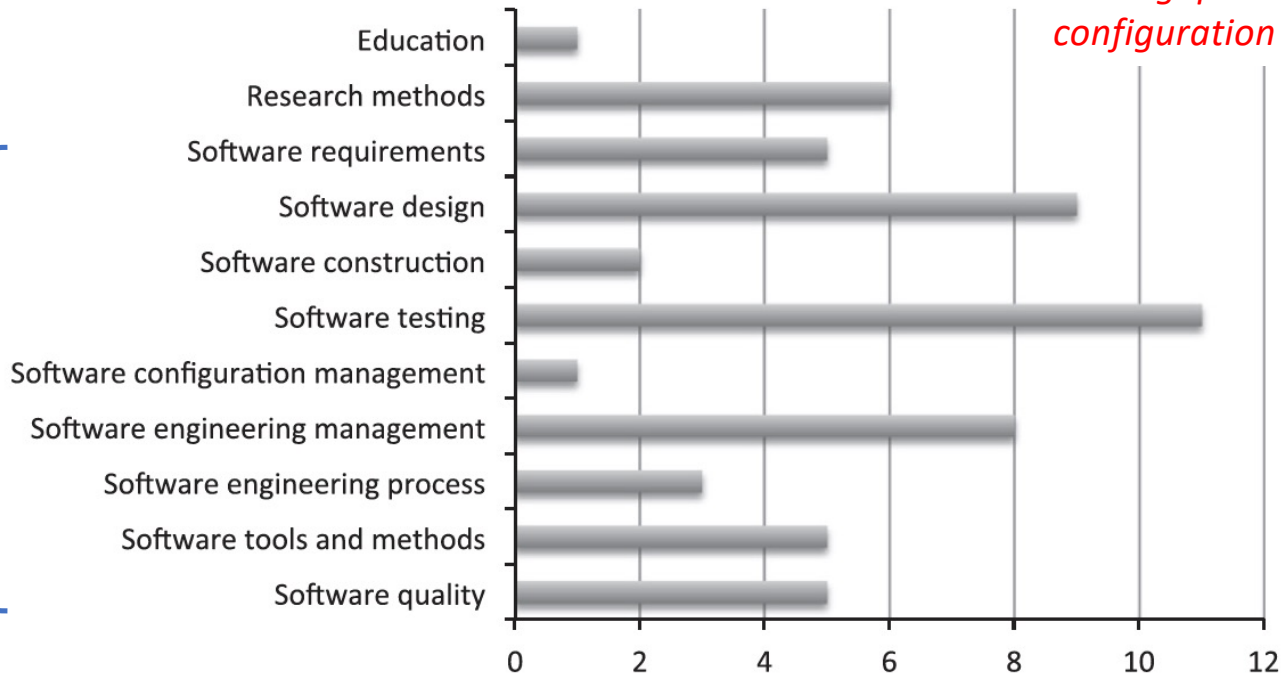


Fig. 3. Topics covered in systematic mapping studies.

SMS: Results for RQ3 (Where & When)

- Frequency of publications (2007-2012)

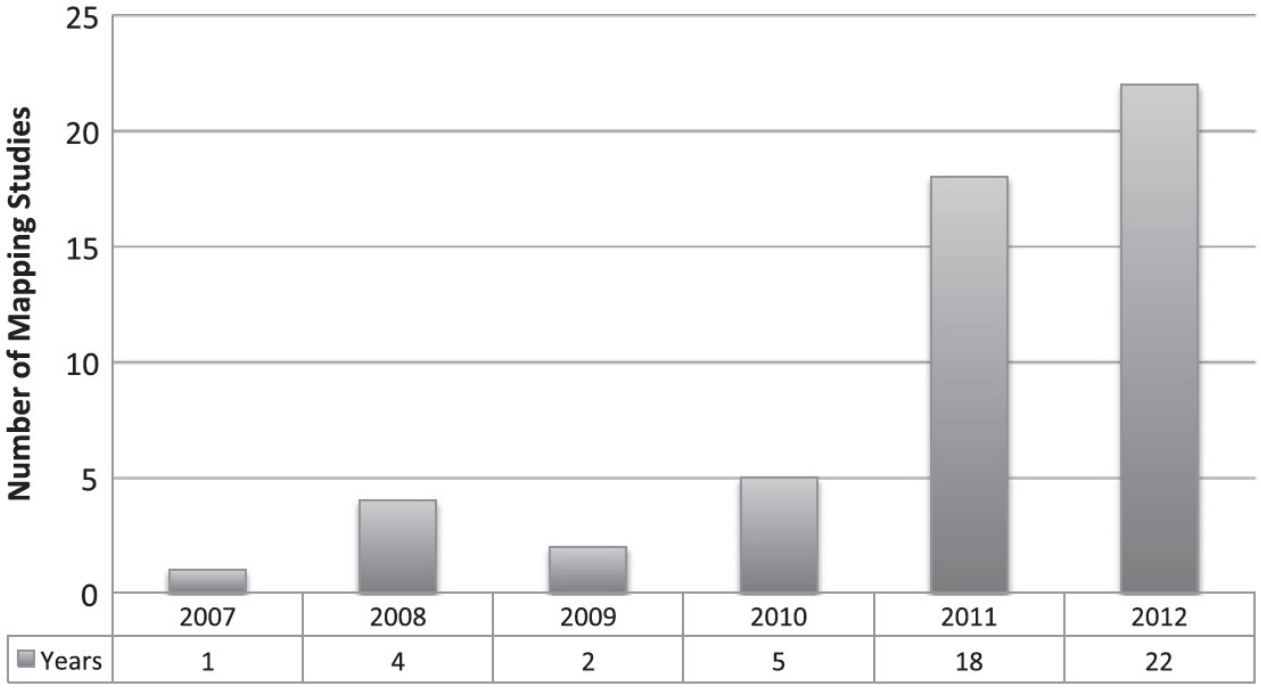


Fig. 2. Publications per year.

SMS: Results for RQ3 (Where and When)

- Venues (Type? Frequency?)

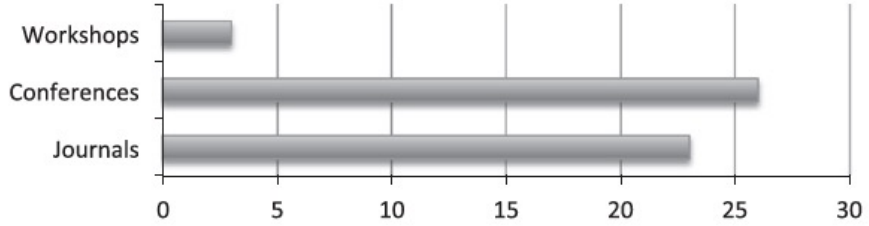


Fig. 4. Venue types.

Table 4
Targeted venues.

Rank	Venue	Studies
1	Information & Software Technology (IST)	[23,33,12,34–36,13,37,22,20,38,30]
2	Evaluation and Assessment in Software Engineering (EASE)	[39–46]
3	Empirical Software Engineering and Measurement (ESEM)	[32,5,3,47]

SMS: Results for RQ4

- Review process
  - ~~Guidelines~~
  - Study identification
  - Quality evaluation
  - Data extraction and classification
  - Visualization
  - Validity threats

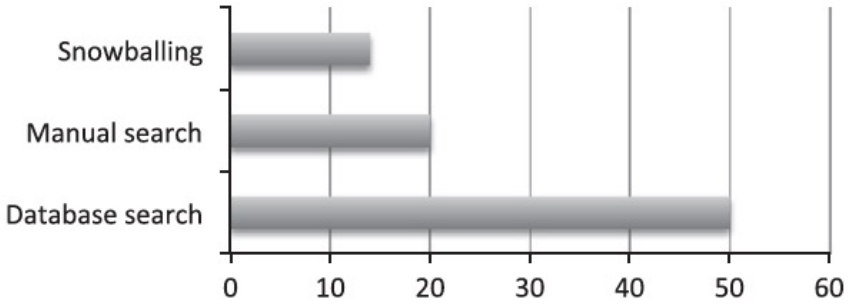


Fig. 6. Choosing search strategies.

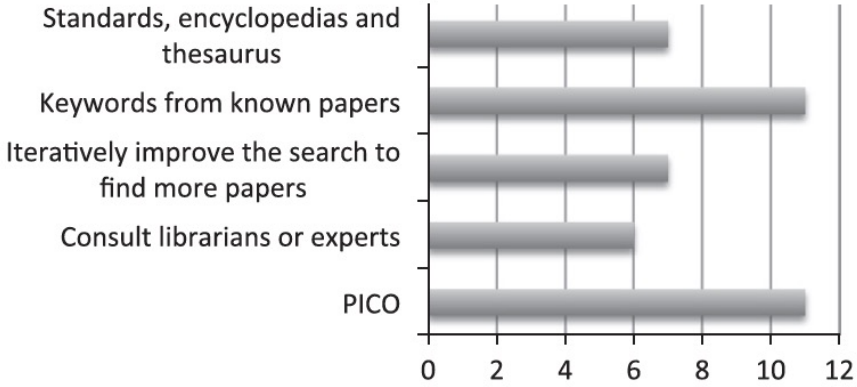


Fig. 7. Developing the search strategy.

SMS: Results for RQ4

- Review process
  - ~~Guidelines~~
  - Study identification
  - Quality evaluation
  - Data extraction and classification
  - Visualization
  - Validity threats

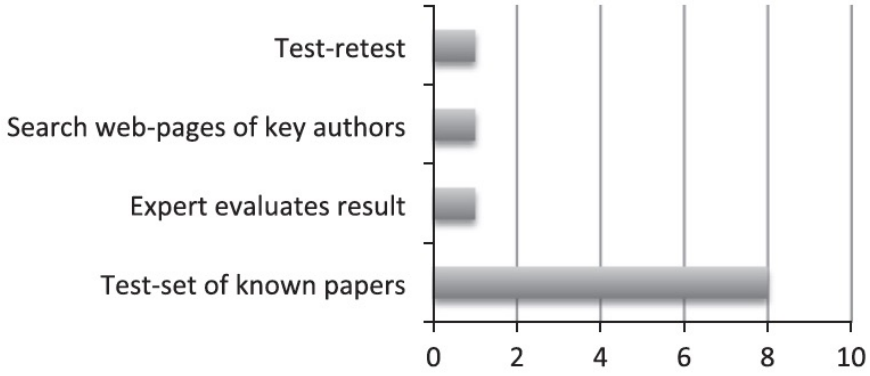


Fig. 8. Evaluating the search strategy.

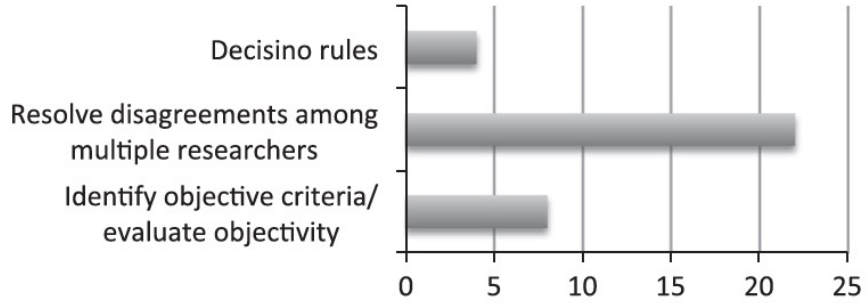


Fig. 9. Inclusion and exclusion.

SMS: Results for RQ4

- Review process
  - ~~Guidelines~~
  - Study identification
  - Quality evaluation
  - Data extraction and classification
  - Visualization
  - Validity threats

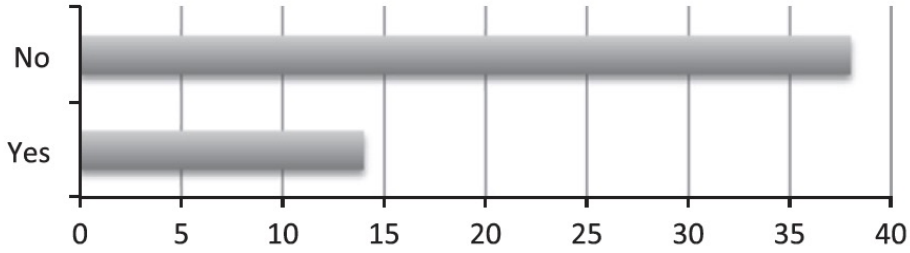


Fig. 10. Quality assessment.

14 out of 52 did not evaluate the quality of the identified studies.

SMS: Results for RQ4

- Review process
  - ~~Guidelines~~
  - Study identification
  - Quality evaluation
  - Data extraction and classification
  - Visualization
  - Validity threats

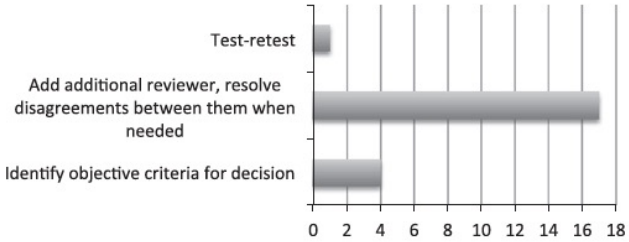


Fig. 11. Data extraction and classification process.

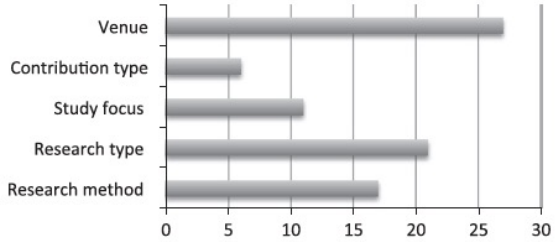


Fig. 12. Topic-independent classification.

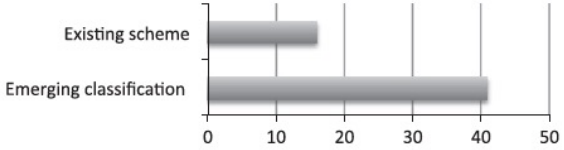


Fig. 13. Topic-related classification.

SMS: Results for RQ4

- Review process
  - ~~Guidelines~~
  - Study identification
  - Quality evaluation
  - Data extraction and classification
  - Visualization
  - Validity threats

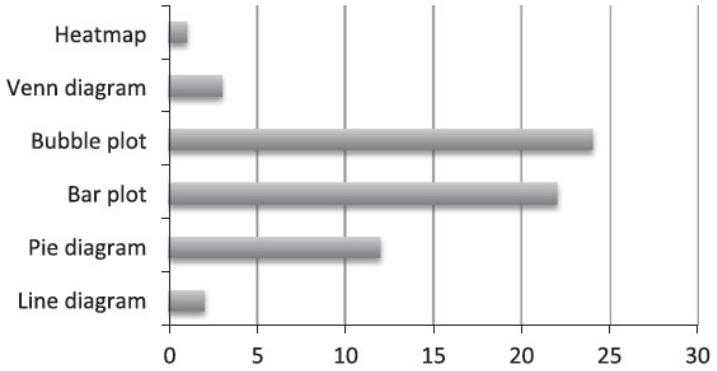


Fig. 14. Visualizations.

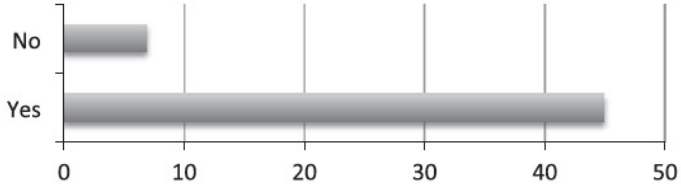


Fig. 15. Discussion of validity threats.

Based on the results, improvements to the process of conducting SMS are suggested regarding:

(1) Planning

(2) Conducting

(3) Reporting

(see Section 5 in the paper)

Today's Program

- SLR and SMS examples
- Guidelines on how to write an SMS
- Tips & Tricks for report writing
- Tips & Tricks for presenting

Hints & Tips – Technical Writing

Importance of good writing

- It helps your readers to:
  - distinguish between more important and less important material
  - quickly find specific information
  - grasp the flow of argument, the assumptions (incl. previous/related work), and the context
- Remember: Often you don’t have control over the recipients of your document

Hints & Tips – Technical Writing

Style issues:

- No colloquial expressions or slang!
- Appropriate usage of terminology
- Consistent wording (one concept – one expression)
- Short sentences
- Normally: result-oriented (not process-oriented)
- Introduce abbreviations before first usage
- Neutral style instead of “we …”, “our …”, etc.
- After each headline there should be some text …

Hints & Tips – Technical Writing

Figures and tables:

- Have a caption (incl. brief description and numbering)

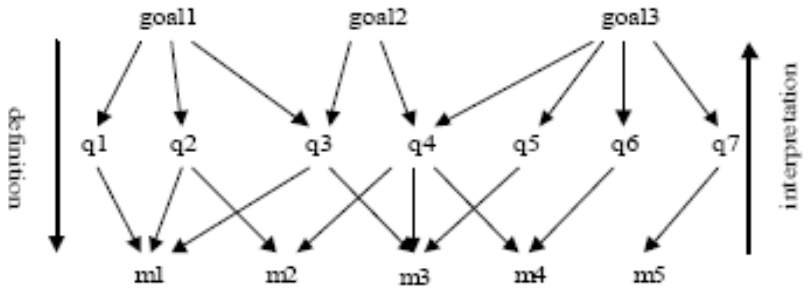


Fig. 1. A sample GQM.

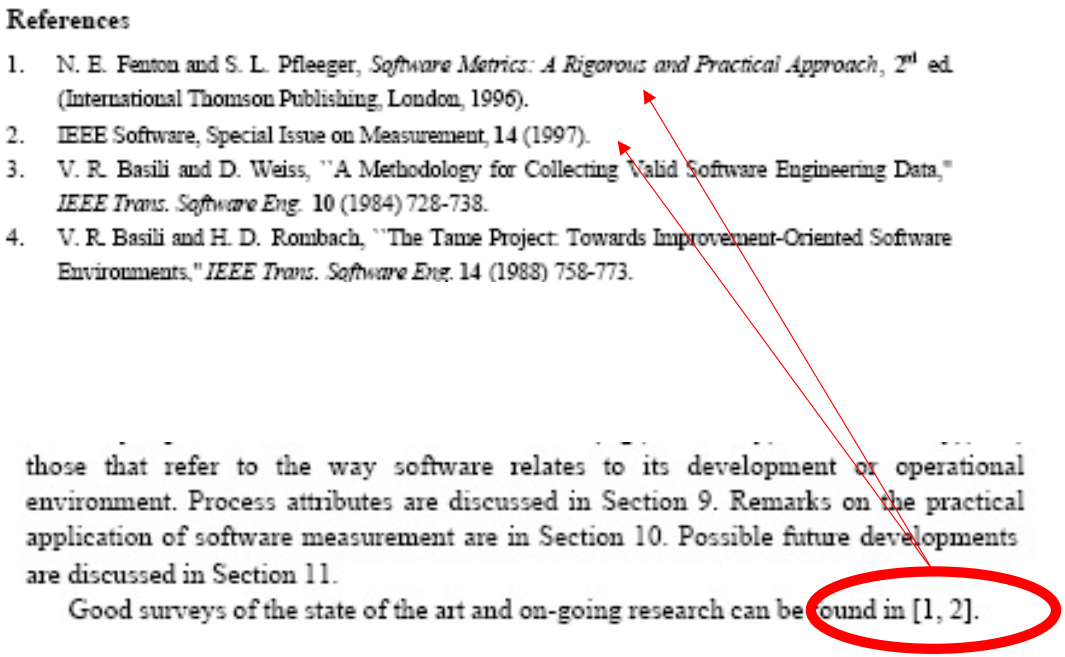
- Are referenced from within the related text

density? Each question is then refined into measures that can be collected on the field. For instance, defect density may be defined as the ratio of the number of defects found to the number of lines of code. Figure 1 shows this top-down refinement of goals into measures. Figure 1 also shows that several measurement goals may be pursued at the same time, and questions and measures may be reused across goals, thus decreasing the effort for adding further goals to an existing set of goals, questions, and measures.

Hints & Tips – Technical Writing

References:

- All literature included in the list of references must be referenced (at least once)



References

1. N. E. Fenton and S. L. Pfleeger, *Software Metrics: A Rigorous and Practical Approach*, 2nd ed. (International Thomson Publishing, London, 1996).
2. IEEE Software, Special Issue on Measurement, 14 (1997).
3. V. R. Basili and D. Weiss, "A Methodology for Collecting Valid Software Engineering Data," *IEEE Trans. Software Eng.* 10 (1984) 728-738.
4. V. R. Basili and H. D. Rombach, "The Tame Project: Towards Improvement-Oriented Software Environments," *IEEE Trans. Software Eng.* 14 (1988) 758-773.

those that refer to the way software relates to its development or operational environment. Process attributes are discussed in Section 9. Remarks on the practical application of software measurement are in Section 10. Possible future developments are discussed in Section 11.

Good surveys of the state of the art and on-going research can be found in [1, 2].

How bibliography needs to be organized? References should appear as …

References should appear as …

- **Book with one author**:
  - Author, A. A. (2005). Title of work. Location/City, State: Publisher.
- **Book with two authors**:
  - Author, A. A., & Author, B. B. (2005). Title of work. Location/City, State: Publisher.
- **Book with more than two authors**:
  - Author, A. A., Author, B. B., & Author, C. C. (2005). Title of work. Location/City, State: Publisher.
- **Journal article**:
  - Sawyer, S., & Tapia, A. (2005). The sociotechnical nature of mobile computing work: Evidence from a study of policing in the United States. International Journal of Technology and Human Interaction, 1(3), 1-14.
- **A publication in press**:
  - Junho, S. (in press). Roadmap for e-commerce standardization in Korea. International Journal of IT Standards and Standardization Research.

How bibliography needs to be organized? References should appear as …

References should appear as …

- **Journal article**:
  - Sawyer, S., & Tapia, A. (2005). The sociotechnical nature of mobile computing work: Evidence from a study of policing in the United States. International Journal of Technology and Human Interaction, 1(3), 1-14.
- **A publication in press**:
  - Junho, S. (in press). Roadmap for e-commerce standardization in Korea. International Journal of IT Standards and Standardization Research.
- **Report from a university**:
  - Broadhurst, R. G., & Maller, R. A. (1991). Sex offending and recidivism (Tech. Rep. No. 3). Nedlands, Western Australia: University of Western Australia, Crime Research Centre.
- **Published proceedings**:
  - Deci, E. L., & Ryan, R. M. (1991). A motivational approach to self: Integration in personality. In R. Dienstbier (Ed.), Nebraska Symposium on Motivation: Vol. 38. Perspectives on motivation (pp. 237-288). Lincoln: University of Nebraska Press.

How bibliography needs to be organized? References should appear as …

References should appear as …

- **Unpublished doctoral dissertation or master's thesis**:
  - Wilfley, D. (1989). Interpersonal analyses of bulimia: Normal-weight and obese. Unpublished doctoral dissertation, University of Missouri, Columbia.
- **A presented paper**:
  - Lanktree, C., & Briere, J. (1991, January). Early data on the Trauma Symptom Checklist for Children (TSC-C). Paper presented at the meeting of the American Professional Society on the Abuse of Children, San Diego, CA.
- **Web site**:
  - VandenBos, G., Knapp, S., & Doe, J. (2001). Role of reference elements in the selection of resources by psychology undergraduates. Journal of Bibliographic Research, 5, 117- 123. Retrieved October 13, 2001, from http://jbr.org/articles.html

In-Text Citations

- In-text citations should appear with author surname followed by publication year in parentheses

  (Brown, 2002)

- Citing several references in-text:

  In most organizations, data resources are considered to be a major resource (Brown, 2002; Krall & Johnson, 2005; Smith, 2001).

  Brown (2002) states that the value of data is recognized by most organizations

  “In most organizations, data resources are considered to be a major organization asset” (Smith, 2001, pp. 35-36) and must be carefully monitored by the senior management.

  Brown (2002) states that “the value of data is realized by most organizations” (p. 45).

- If you have organized the citations with number in brackets:

  In most organizations, data resources are considered to be a major resource [15, 30, 84].

Tips for Writing a Thesis

How theses get written: some cool tips …

- ***http://www.cs.toronto.edu/~sme/presentations/thesiswriting.pdf***

Today's Program

- SLR and SMS examples
- Guidelines on how to write an SMS
- Tips & Tricks for report writing
- Tips & Tricks for presenting

Hints & Tips – Presentation

- What to present:
  - Topic and Goals (RQs)
    - Also say what you find interesting about the RQs
  - Method (SMS Steps)
    - Data Search (Database, Manual, Snowballing)
    - Study Selection and Quality Assessment
    - Data Extraction
    - Analysis and Classification
  - Results (Answers to RQs)
  - Your learning experience
    - What was difficult?
    - What was most time-consuming?
    - … ?

Information and Software Technology 64 (2015) 1–18

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Guidelines for conducting systematic mapping studies in software engineering: An update

Kai Petersen *, Sairam Vakkalanka, Ludwik Kuzniarz

Department of Software Engineering, Blekinge Institute of Technology, Sweden

ARTICLE INFO

Article history:
Received 1 September 2014
Received in revised form 23 February 2015
Accepted 14 March 2015
Available online 28 March 2015

Keywords:
Systematic mapping studies
Software engineering
Guidelines

ABSTRACT

Context: Systematic mapping studies are used to structure a research area, while systematic reviews are focused on gathering and synthesizing evidence. The most recent guidelines for systematic mapping are from 2008. Since that time, many suggestions have been made of how to improve systematic literature reviews (SLRs). There is a need to evaluate how researchers conduct the process of systematic mapping and identify how the guidelines should be updated based on the lessons learned from the existing systematic maps and SLR guidelines.

Objective: To identify how the systematic mapping process is conducted (including search, study selection, analysis and presentation of data, etc.); to identify improvement potentials in conducting the systematic mapping process and updating the guidelines accordingly.

Method: We conducted a systematic mapping study of systematic maps, considering some practices of systematic review guidelines as well (in particular in relation to defining the search and to conduct a quality assessment).

Results: In a large number of studies multiple guidelines are used and combined, which leads to different ways in conducting mapping studies. The reason for combining guidelines was that they differed in the recommendations given.

Conclusion: The most frequently followed guidelines are not sufficient alone. Hence, there was a need to provide an update of how to conduct systematic mapping studies. New guidelines have been proposed consolidating existing findings.

1. Introduction

Systematic mapping studies or scoping studies are designed to give an overview of a research area through classification and counting contributions in relation to the categories of that classification [1,2]. It involves searching the literature in order to know what topics have been covered in the literature, and where the literature has been published [2]. Though, a systematic mapping study and a systematic literature review share some commonalities (e.g. with respect to searching and study selection), they are different in terms of goals and thus approaches to data analysis. While systematic reviews aim at synthesizing evidence, also considering the strength of evidence, systematic maps are primarily concerned with structuring a research area.

Systematic mapping studies are used by many researchers on a number of areas using different guidelines or methods. A sample of mapping studies is mentioned below with their areas of research and the guidelines used.

- Condori-Fernandez et al. [3] provided a mapping of the research articles on software requirement specifications combining two guidelines (cf. [2,1]).
- Jalali and Wohlin [4] performed mapping of the literature available on Global software Engineering considering the guidelines by [2,1].
- Barreiros et al. [5] constructed systematic maps on the published research on software engineering test beds based on Kitchenham and Charters's [1] guidelines.
- Qadir and Usman [6] conducted a mapping on curriculum in software engineering using the guidelines by [2,1].

Recently, Wohlin et al. [7] compared systematic mapping studies that were conducted on the same topic by two groups of researchers working and publishing independently. That is, the review protocols of the two reviews were developed independently. Some questions on the reliability of systematic mapping

* Corresponding author.
E-mail addresses: kai.petersen@bth.se (K. Petersen), sava11@student.bth.se (S. Vakkalanka), ludwik.kuzniarz@bth.se (L. Kuzniarz).

http://dx.doi.org/10.1016/j.infsof.2015.03.007
0950-5849/

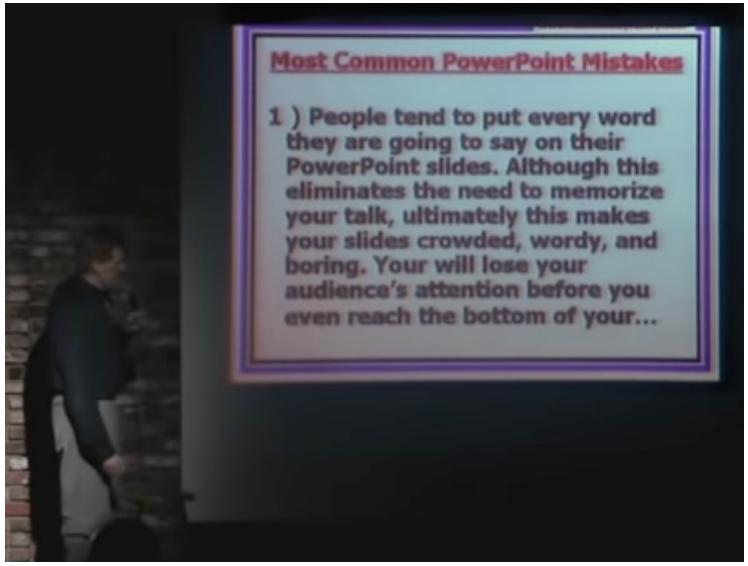
Hints & Tips – Presentation

- How to present:
  - Use the time – but don’t exceed (12 min presentation + 3 min discussion)
  - Balance content (per slide and across whole presentation)
  - Speak clearly (not too fast / not too low)
  - Let the slides support your talk – don’t use slides to write down what you say
  - Use visuals/diagrams/tables where appropriate

Use these bullet points to rank the presenters and justify your ranking!

How **not** to do a presentation …

- Don McMillan: Life After Death by PowerPoint
  - http://www.youtube.com/watch?v=WGiePuNFXwY
  - http://www.youtube.com/watch?v=zDvm1PVtgWo
  - http://www.youtube.com/watch?v=lpvgfmEU2Ck



Next Steps

- 22 February - Last possibility to cancel course participation in the SIS.
- 23 February & 02 March & 09 March - Individual Consultation (optional & on request by student only - first come first serve principle) - Online or face-to-face in my office (room 3007)
- 28 February - Deadline for identifying topic with RQs
- 14 March - Deadline for submitting draft reports & slides of Presentations 1 (submit before 23:59)
- 16 March - Presentations 1: 1st and 2nd year students (work in pairs) …