Seminar from 2013/2014

Patricia Lago
VU University Amsterdam
Think: do I really need to print this?
Study material and Useful resources

- Study material all available on [www.cs.vu.nl/~patricia > Teaching](http://www.cs.vu.nl/~patricia)
Contents

• Module organization (changed from academic year 2013/2014)

• Essentials of literature study design

• Resources and Materials
What will you learn

• How to formulate a research question

• How to translate it into a sound search query

• How to identify the right on-line literature search engines

• How to perform a motivated selection of the literature for further analysis
Overview of module organization

- Identify topic of interest
- Contact potential tutor
- Agree on the (draft) research question, or at least the research problem to be addressed

This lecture
(under supervision of your tutor)

- Design your literature study
- Running the study
- Prepare the final report
- Present study results to the research group of your tutor
Grading criteria

• quality of study design;

• rationale for literature selection;

• quality of results (scientific quality of the analysis, discussion of the findings, reflection in the drawn conclusions, clear answer to main research question);

• quality of final report (clarity, organization);

• correctness and completeness of references and citations;

• final presentation to the research group where the student has performed the literature study;
Contents

• Module organization (changed from academic year 2013/2014)
• Essentials of literature study design
• Resources and Materials
Essentials of literature study design

Planning the review
- Identify the need for a review
- Customize the review protocol

Conducting the review
- Select primary studies
- Data extraction
- Data analysis

Reporting the review
- Write final report
- Prepare presentation

Adapted from [Applying Systematic Reviews to Diverse Study Types An Experience Report, 2007]
Two types of systematic studies

Planning the review
Conducting the review
Reporting the review

More suitable for a time-boxed literature study
• Systematic search
• Systematic analysis

Long and effort-intensive
Identify the need for a literature review

Example

- Cloud-based software is assumed to be more energy-efficient and less expensive. (Such properties are becoming Service Level Agreements/Objectives.)

- Is it true?

- Are there “software architecture types” that we can reuse to engineer energy-efficient software for cloud provisioning?
Customize the review protocol

- RQ → keywords
- Query
- I/E criteria
- Selection process (from title to abstract to contents); selection of a subset of the primary studies
- Data extraction
- Data analysis
Customize the review protocol /2

1. RQ → keywords
   • RQ: what kind of software architectures for cloud service provisioning allow to define service level objectives regarding energy efficiency?

2. Five keywords
   • Add synonyms

3. RQ → Query string
   • “software architecture” AND
   • cloud AND
   • service AND
   • “(energy OR power) efficiency” AND
   • (SLA OR SLO OR “service level“)
## Inclusion/Exclusion criteria

<table>
<thead>
<tr>
<th><strong>Criterion</strong></th>
<th><strong>Rationale</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I1</strong> A study that directly proposes software architectures, architectural styles or strategies that are generally reusable or applicable, or indirectly proposes them from a service provisioning perspective.</td>
<td>We want to identify how software architectures affect energy efficiency, thus we need articles proposing software architectures that are generally reusable or applicable, or indirectly proposing them from a service provisioning perspective.</td>
</tr>
<tr>
<td><strong>I2</strong> A study that addresses energy efficiency as a quality attribute.</td>
<td>We want to investigate whether energy efficiency is considered, either implicitly or explicitly, by providers or experts, as a quality attribute for cloud services.</td>
</tr>
<tr>
<td><strong>I3</strong> A study that is developed by either of academics and practitioners.</td>
<td>Both academic and industrial solutions are relevant to this study.</td>
</tr>
<tr>
<td><strong>I4</strong> A study that is published in software engineering/cloud computing field.</td>
<td>Software engineering is our reference field, but cloud computing research can provide us an insight on what trends are set in terms of software architectures for cloud.</td>
</tr>
<tr>
<td><strong>I5</strong> A study that is peer-reviewed.</td>
<td>A peer-reviewed paper guarantees a certain level of quality and contains reasonable amount of content.</td>
</tr>
<tr>
<td><strong>I6</strong> A study that is written in English.</td>
<td>For feasibility reasons papers written in other languages than English are excluded.</td>
</tr>
<tr>
<td><strong>E1</strong> A study that does not propose software solutions for energy efficiency.</td>
<td>Traditionally, energy efficiency has been regarded as an hardware issue. We want to drive past this assumption and address the software impact of power consumption.</td>
</tr>
<tr>
<td><strong>E2</strong> A study that does not imply any type of service provisioning.</td>
<td>We are not interested in solutions that generally increase the energy efficiency of a datacenter, without having in mind how to provide an energy-efficient service to a customer.</td>
</tr>
<tr>
<td><strong>E3</strong> A study that does not consider</td>
<td>We are not interested in studies that consider energy effi-</td>
</tr>
</tbody>
</table>
Best practices – Selection process

• Each primary study must fulfil all Inclusion Criteria and none of the Exclusion Criteria

• The query (hence the RQ) must often be “scoped down” to limit the number of primary studies

  • Keep the RQ relevant but not too broad

• A typical systematic literature review can cover between 30 and 150 primary studies

  • Target about 15-20 primary studies, find a good reason for their selection

  • E.g.: primary studies cover two types of works, surveys and specific technical contributions. I select the first to give a broader overview of the field, and keep track of the second (in my excel sheet – see next slides)
Conducting the review

BEST PRACTICES – SEARCH ENGINE

- Google scholar
- ACM Digital Library (or ACM Digital Guide)
- IEEE Explore
- Springer-link
- Elsevier Scopus
- ISI Web of Knowledge
• An Excel sheet can be a valuable tool to document your protocol

<table>
<thead>
<tr>
<th>ID</th>
<th>Yes/no</th>
<th>Title of the article</th>
<th>Author(s) of the article</th>
<th>Publication year</th>
<th>Inclusion criteria, mark with X if met</th>
<th>Exclusion criteria, mark with X if met</th>
<th>Mark X if only the title is read</th>
<th>Mark X if the abstract is read</th>
<th>Mark X if the full text is read</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>53</td>
<td>no</td>
<td>Cloud Computing Models</td>
<td>Geraldi, Eugenie</td>
<td>2015</td>
<td>X X X X X</td>
<td>X X X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>no</td>
<td>The cloud adoption toolkit: Addressing the challenges of cloud</td>
<td>Greenwood, David</td>
<td>2016</td>
<td>X X X X X</td>
<td>X X X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>X</td>
<td>Energy saving techniques in datacenters</td>
<td>Grodenic, I and</td>
<td>2012</td>
<td>X X X X X</td>
<td>X X X</td>
<td>O</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
BEST PRACTICES - EXCEL

- Data Columns: ID, Title, Author, Pub. Year, Notes
- Allow you to perform simple demographic analysis (e.g. publication trend graph)
- Filtering/Sorting capabilities

Figure 1: Number of primary studies selected per year.
Figure 2: Distribution of primary studies according to contribution type.
BEST PRACTICES - EXCEL

● Inclusion/Exclusion criteria columns

● Keep track of which study you selected/discarded and why!
Best Practices - Excel

- **Selected** column

- Once you made your final decision on a study

- You can filter out discarded studies (neat!)
BEST PRACTICES - EXCEL

● **Data Extraction** sheet

<table>
<thead>
<tr>
<th>Year</th>
<th>Type</th>
<th>Notes</th>
<th>Energy Efficiency</th>
<th>Elements</th>
<th>Stakeholders involved</th>
<th>Validation of the solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>Master Thesis</td>
<td>Energy Monitoring + Self-Adaptation – Static classification based on hardware specifications</td>
<td>Energy-aware composition and provisioning mechanism for the VEs on the IT side the “greenest” resources available that fulfill the application requirements will be selected consolidating their allocation on the most efficient hardware, and shutting down or putting on stand-by the unused resources to minimize consumption. The selection is proposed to consider indicators for general classes of energy-efficient resources to give priority to the usage of certain VEs over others. the suggested way to proceed for selecting the VEs in the choice of IT resources, is to assign such indicators (Energy Saving Indicators (ESIs)) to prioritize IT VEs provided by providers certified for different EE characteristics including the usage of new generation energy-efficient hardware components.</td>
<td>Logical Infrastructure Composition Layer (LICL), Enhanced Network Control Plane (ENCP), Service Middleware Layer (SML), Providers, Virtual Requests Management module, the Data Infrastructure Management module and the Service and Resource Management module, also referred as Virtual IT Manager (VITM). Operators, Customers; contains the Green Traffic Engineering Network Operators; Database (GreenTE-DB) module</td>
<td>Absent</td>
<td></td>
</tr>
</tbody>
</table>

● You can embed your extraction form in the same Excel file linking it to the primary studies in the first sheet
Data analysis

• Most difficult yet most important part

• Ask yourself e.g.:
  • What do I learn? What do I observe? Similarities and differences?

• Given your RQ,
  • E.g. are the architecture solutions proposed in the primary studies of certain types?

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Techniques</th>
<th>Description</th>
<th>Reference studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Static classification</td>
<td>Energy classification of software based upon the power consumption specifications of the hardware components</td>
<td>(Alessandro, 2010)(Curtis, 2008)(Götz et al., 2011)</td>
</tr>
</tbody>
</table>
Contents

• Module organization (changed from academic year 2013/2014)

• Essentials of literature study design

• Resources and Materials
Resources and materials: guidelines deliverables

• The **final report** should include:
  • study design; overview of selected literature; analysis of the literature; discussion and conclusions

• The **final presentation** should include:
  • (overview of) research questions, study design, study execution;
  • (in depth) discussion of analysis results
Resources and materials: downloads

• Tools

• Templates

• Further reading

Planning the review
- Identify the need for a review
- Customize the review protocol

Conducting the review
- Select primary studies
- Data extraction
- Data analysis

Reporting the review
- Write final report
- Prepare presentation
If you don’t know where you're going, any road will do…

Now you do know!