Research Methods in Cybersecurity

Raimundas Matulevičius
rma@ut.ee
Introductory Seminar

• How to write 2-page problem statement?
• How to perform literature review?
• **How bibliography needs to be organised?**
• Few more quick tips for presentation
How bibliography needs to be organised?

References should appear as

• **Book with one author:**
  • Author, A. A. (2005). Title of work. Location/City, State: Publisher.

• **Book with two authors:**

• **Book with more than two authors:**

• **Journal article:**

• **A publication in press:**
How bibliography needs to be organised?
References should appear as

• **Edited book:**

• **Chapter in an edited book:**

• **Report from a university:**

• **Published proceedings:**
How bibliography needs to be organised?

References should appear as

- **Unpublished doctoral dissertation or master’s thesis:**

- **A presented paper:**

- **Web site:**
How bibliography needs to be organised?

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).
How bibliography needs to be organised?

In-Text Citations

• If you have organised the citations with number in brackets

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

STEP 1: MAKE A PLAN.

STEP 2: SET A DEADLINE.
Thesis plan

**OK, I need a plan for writing my thesis.**

**Step 1: Make a plan.**

**Step 2: Set a deadline.**

**Step 3: Freak out.**
Structure of the Thesis

• Introduction
• Background / State of the Art
• Contribution
• Validation
• Discussion and Conclusion
Introduction

• Introduction
  • Motivation
  • Scope
  • Research problem
    • Research question
  • Contribution
  • Structure description

• Background / State of the Art
• Contribution
• Validation
• Discussion
• Conclusion
Background / State of the Art

• Introduction

• **Background / State of the Art**
  • Survey of existing technology
  • Survey of existing related work

• Contribution
• Validation
• Discussion
• Conclusion


http://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf
Contribution

- Introduction
- Background / State of the Art
- Contribution
  - Research method
  -...
- Validation
- Discussion
- Conclusion
Validation

• Introduction
• Background / State of the Art
• Contribution

• Validation
  • Theoretical validation
  • Empirical validation
  • Experimental validation

• Discussion
• Conclusion


Proof-of-concept
e.g., Prototype
Discussion

• Introduction
• Background / State of the Art
• Contribution
• Validation

• Discussion
  • Limitations
  • Related work
    • Comparison of your results (method, framework, calculations, findings) to similar studies
    • Differences, advantages and limitations

• Conclusion
Conclusion

• Introduction
• Background / State of the Art
• Contribution
• Validation
• Discussion

• Conclusion
  • Not **ONLY** summary !!!
  • Summarising the thesis - 1 paragraph, 5 lines at most.
  • Concluding remarks: Contribution to theory - Contribution to practice
  • Generalisation
  • Future work
• **Introduction**
  - Motivation
  - Scope
  - Research problem
    - Research question
  - Contribution
  - Structure description

• **Background / State of the Art**
  - Survey of existing technology
  - Survey of existing related work

• **Contribution**
  - Research method
  - …

• **Validation**
  - Theoretical validation
  - Empirical validation
  - Experimental validation

• **Discussion**
  - Limitations
  - Related work
    - Comparison of your results (method, framework, calculations, findings) to similar studies
    - Differences, advantages and limitations

• **Conclusion**
  - Not ONLY summary !!!
  - Summarising the thesis
    - 1 paragraph, 5 lines at most.
  - Concluding remarks: Contribution to theory - Contribution to practice
  - Generalisation
  - Future work
Writing Good Software Engineering Research Papers

by Mary Shaw


presented by Raimundas Matulevičius
University of Tartu, Estonia
A good research paper or thesis should answer a number of questions
A good research paper or thesis should answer a number of questions

• **What, precisely, was your contribution?**
  • What question did you answer?
  • Why should the reader care?
  • What larger question does this address?

• **What is your new result?**
  • What new knowledge have you contributed that the reader can use elsewhere?
  • What previous work (yours or someone else's) do you build on? What do you provide a superior alternative to?
  • How is your result different from and better than this prior work?
  • What, precisely and in detail, is your new result?

• **Why should the reader believe your result?**
  • What standard should be used to evaluate your claim?
  • What concrete evidence shows that your result satisfies your claim?
A good research paper or thesis should answer a number of questions

- **What, precisely, was your contribution?**
  - What question did you answer?
  - Why should the reader care?
  - What larger question does this address?

- **What is your new result?**
  - What new knowledge have you contributed that the reader can use elsewhere?
  - What previous work (yours or someone else's) do you build on? What do you provide a superior alternative to?
  - How is your result different from and better than this prior work?
  - What, precisely and in detail, is your new result?

- **Why should the reader believe your result?**
  - Why should the reader care?
  - What standard should be used to evaluate your claim?
  - What concrete evidence shows that your result satisfies your claim?

If you answer these questions clearly,

- You will probably communicate your result well
- Your result represents an interesting, sound, and significant contribution
- You will have a good chance of getting thesis accepted for publication in a conference or journal
- You will have a good chance of publishing the paper (or few) in a prestigious conference or journal
What, precisely, was your contribution?
What kinds of questions do software engineers investigate?

<table>
<thead>
<tr>
<th>Type of question</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method or means of development</td>
<td>How can we do/create/modify/evolve (or automate doing) X?</td>
</tr>
<tr>
<td></td>
<td>What is a better way to do/create/modify/evolve X?</td>
</tr>
<tr>
<td>Method for analysis or evaluation</td>
<td>How can I evaluate the quality/correctness of X?</td>
</tr>
<tr>
<td></td>
<td>How do I choose between X and Y?</td>
</tr>
<tr>
<td>Design, evaluation, or analysis of a particular instance</td>
<td>How good is Y? What is property X of artifact/method Y?</td>
</tr>
<tr>
<td></td>
<td>What is a (better) design, implementation, maintenance, or adaptation for application X?</td>
</tr>
<tr>
<td></td>
<td>How does X compare to Y?</td>
</tr>
<tr>
<td></td>
<td>What is the current state of X / practice of Y?</td>
</tr>
<tr>
<td>Generalization or characterization</td>
<td>Given X, what will Y (necessarily) be?</td>
</tr>
<tr>
<td></td>
<td>What, exactly, do we mean by X? What are its important characteristics?</td>
</tr>
<tr>
<td></td>
<td>What is a good formal/empirical model for X?</td>
</tr>
<tr>
<td></td>
<td>What are the varieties of X, how are they related?</td>
</tr>
<tr>
<td>Feasibility study or exploration</td>
<td>Does X even exist, and if so what is it like?</td>
</tr>
<tr>
<td></td>
<td>Is it possible to accomplish X at all?</td>
</tr>
</tbody>
</table>

What do thesis committees look for?

• A clear statement of the specific problem you solved

• The question about software development you answered

• An explanation of how the answer will help solve an important software engineering problem
What is your new result?
What kinds of results do you produce?

<table>
<thead>
<tr>
<th>Type of result</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedure or technique</td>
<td>New or better way to do some task, such as design, implementation, maintenance, measurement, evaluation, selection from alternatives; includes techniques for implementation, representation, management, and analysis; a technique should be operational—not advice or guidelines, but a procedure</td>
</tr>
<tr>
<td>Qualitative or descriptive model</td>
<td>Structure or taxonomy for a problem area; architectural style, framework, or design pattern; non-formal domain analysis, well-grounded checklists, well-argued informal generalizations, guidance for integrating other results, well-organized interesting observations</td>
</tr>
<tr>
<td>Empirical model</td>
<td>Empirical predictive model based on observed data</td>
</tr>
<tr>
<td>Analytic model</td>
<td>Structural model that permits formal analysis or automatic manipulation</td>
</tr>
<tr>
<td>Tool or notation</td>
<td>Implemented tool that embodies a technique; formal language to support a technique or model (should have a calculus, semantics, or other basis for computing or doing inference)</td>
</tr>
<tr>
<td>Specific solution, prototype, answer, or judgment</td>
<td>Solution to application problem that shows application of SE principles – may be design, prototype, or full implementation; careful analysis of a system or its development, result of a specific analysis, evaluation, or comparison</td>
</tr>
<tr>
<td>Report</td>
<td>Interesting observations, rules of thumb, but not sufficiently general or systematic to rise to the level of a descriptive model.</td>
</tr>
</tbody>
</table>
What do thesis committees look for?

• Interesting, novel, exciting results that

**Significantly enhance the ability**

• to develop and maintain software
• to know the quality of the software
• to recognize general principles about software
• to analyze properties of software

Questions that committee asks about your thesis

What, precisely, do you claim to contribute?

- If your result ought to work on large systems
  - explain why you believe it scales

- If you claim your method is "automatic"
  - should not require human intervention
    - automatic when it's operating but requires manual assistance to configure
    - automatic except for certain cases how often the exceptions occur

- If you claim your result is "distributed"
  - should not have a single central controller or server
    - If it does, what part is distributed
    - what part is not

- If you're proposing a new notation for an old problem
  - why your notation is clearly superior to the old one

- If your thesis is an "experience report"
  - use of a previously-reported tool or technique in a practical software project
    - what idea the reader can take away from the thesis to use in other settings

<table>
<thead>
<tr>
<th>Level</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awful</td>
<td>• I completely and generally solved … (unless you actually did!)</td>
</tr>
<tr>
<td>Bad</td>
<td>• I worked on galumphing. (or studied, investigated, sought, explored)</td>
</tr>
<tr>
<td>Poor</td>
<td>• I worked on improving galumphing. (or contributed to, participated in, helped with)</td>
</tr>
<tr>
<td>Good</td>
<td>• I showed the feasibility of composing blitzing with flitzing.</td>
</tr>
<tr>
<td></td>
<td>• I significantly improved the accuracy of the standard detector.</td>
</tr>
<tr>
<td></td>
<td>(or proved, demonstrated, created, established, found, developed)</td>
</tr>
<tr>
<td>Better</td>
<td>• I automated the production of flitz tables from specifications.</td>
</tr>
<tr>
<td></td>
<td>• With a novel application of the blivet transform, I achieved a 10% increase in speed and a 15% improvement in coverage over the standard method.</td>
</tr>
<tr>
<td></td>
<td>▼</td>
</tr>
<tr>
<td>-------</td>
<td>----</td>
</tr>
<tr>
<td>Bad</td>
<td>▼</td>
</tr>
<tr>
<td>Poor</td>
<td>▼</td>
</tr>
<tr>
<td>Good</td>
<td>▲</td>
</tr>
</tbody>
</table>
| Better| ▲  | • I automated the production of flitz tables from specifications.  
     |     | • With a novel application of the blivet transform, I achieved a 10% increase in speed and a 15% improvement in coverage over the standard method. |

"Try not. *Do*, or *do* not. There is no *try.*" -- Yoda
Questions that committee asks about your thesis

What, precisely, is the result?

- **If you introduce a new model**
  - be clear about its power
    - empirical data, formal semantics, mathematical principles

- **If you introduce a new metric**
  - define it precisely

- **If you introduce a new architectural style, design pattern, or similar design element,**
  - treat it as if it were a new generalization or model

- **If your contribution is principally the synthesis or integration of other results or components**
  - be clear about why the synthesis is itself a contribution
Questions that committee asks about your thesis

What, precisely, is the result?

• If your paper is chiefly a report on experience applying research results to a practical problem
  • say what the reader can learn from the experience

• If a tool plays a featured role in your paper
  • what is the role of the tool?

• If a system implementation plays a featured role in your paper
  • what is the role of the implementation?
    • architecture, implementation technique, capability, performance, system is a result itself

Why should the reader believe your result?
<table>
<thead>
<tr>
<th>Type of validation</th>
<th>Examples</th>
</tr>
</thead>
</table>
| **Analysis**       | I have analyzed my result and find it satisfactory through rigorous analysis, e.g. …  
For a formal model  
For an empirical model  
For a controlled experiment  
    rigorous derivation and proof  
    data on use in controlled situation  
    carefully designed experiment with statistically significant results |
| **Evaluation**     | Given the stated criteria, my result...  
For a descriptive model  
For a qualitative model  
For an empirical model  
    adequately describes phenomena of interest ...  
    accounts for the phenomena of interest...  
    is able to predict ... because ..., or  
    generates results that fit actual data ...  
Includes feasibility studies, pilot projects |
| **Experience**     | My result has been used on real examples by someone other than me, and the evidence of its correctness/usefulness/effectiveness is ...  
For a qualitative model  
For an empirical model or tool  
For a notation or technique  
    narrative  
    data, usually statistical, on practice  
    comparison of systems in actual use |
| **Example**        | Here’s an example of how it works on  
For a technique or procedure  
    a "slice of life" example based on a real system ...  
    a system that I have been developing ...  
    a toy example, perhaps motivated by reality  
The "slice of life" example is most likely to be convincing, especially if accompanied by an explanation of why the simplified example retains the essence of the problem being solved. Toy or textbook examples often fail to provide persuasive validation, (except for standard examples used as model problems by the field). |
| **Persuasion**     | I thought hard about this, and I believe passionately that ...  
For a technique  
For a system  
For a model  
    if you do it the following way, then ...  
    a system constructed like this would ...  
    this example shows how my idea works  
Validation purely by persuasion is rarely sufficient for a research paper. Note, though, that if the original question was about feasibility, a working system, even without analysis, can suffice |
| **Blatant assertion** | No serious attempt to evaluate result. This is highly unlikely to be acceptable |
What do committees look for?

Why should the reader believe your result?

• If you claim to improve on prior art
  • compare your result objectively to the prior art

• If you used an analysis technique
  • follow the rules of that analysis technique

• If the technique is not a common one
  • explain the technique and standards of proof, and be clear about your adherence to the technique
What do committees look for?

Why should the reader believe your result?

• If you offer practical experience as evidence for your result
  • establish the effect your research has

• If you performed a controlled experiment
  • explain the experimental design

• If you performed an empirical study
  • explain what you measured, how you analyzed it, and what you concluded

• If you use a small example for explaining the result
  • provide additional evidence of its practical use and scalability
How do you combine the elements into a research strategy?
1. Formulate the research question
2. Prepare the structure of your thesis
3. Write abstract
4. Investigate needed technology
5. Review related work
6. Develop contribution
7. Validate contribution
8. Write conclusion and introduction
9. Write introduction
1. Formulate the research question
2. Prepare the structure of your thesis
3. Write abstract
4. Investigate needed technology
5. Review related work
6. Develop contribution
7. Validate contribution
8. Write conclusion and introduction
9. Write introduction
1. Formulate the research question
2. Prepare the structure of your thesis
3. Write abstract
4. Investigate needed technology
5. Review related work
6. Develop contribution
7. Validate contribution
8. Write conclusion and introduction
9. Write introduction

Read/review everything carefully

Re-write to improve !!!