How to Write a Great Thesis?
Best (and worst) practices from choosing a topic to handing in

Roman Haas

With material from Dr. Elmar Juergens

In close cooperation with the Academic Advisors at TUM Informatics
Questions?

• Please join our Tweedback sessions
  – Bachelor’s / Master’s Thesis  
    17.06. 11:00 am  
  – Guided Research  
    17.06. 2:30 pm  

• For details, please refer to moodle
2011 – 2017

2017 – now

Research project “SOFIE”
Agenda

1. Why?
2. What’s important?
3. Choosing a topic
4. Doing the work
5. Presentation
• Slides
• Detailed Essays
• FAQ
Agenda

1. Why?
2. What’s important?
3. Choosing a topic
4. Doing the work
5. Presentation
Refactoring Suggestions for Long Methods
Using Network Analysis for Recommendation of Central Software Classes
Daniela Steidl, 2012,
Intl. Working Conference on Reverse Engineering
Kingston, Kanada

Revealing Missing Bug-Fixes in Code Clones in Large-Scale Code Bases
Martin Pöhlmann, 2013,
Intl. Workshop on Software Quality and Maintainability
Genova, Italien

Deriving Extract Method Refactoring Suggestions for Long Methods
Roman Haas, 2016,
Software Quality Days
Wien, Österreich

Ticket Coverage: Putting Test Coverage into Context
Jakob Rott, 2017,
Workshop on Emerging Trends in Software Metrics
Buenos Aires, Argentinien
Agenda

1. Why?
2. What’s most important?
3. Choosing a topic
4. Doing the work
5. Presentation
Automatic Categorization of Open-Source Software

Jonathan Batchu

Bachelor's Thesis in Computer Science

Discovery of missed reuse opportunities

Tobias Volke
Thesis Projects

• Bachelor
  – CS, Games: 15 ECTS, 4 months writing time
  – Information Systems: 25 ECTS, 5 months writing time including additional project

• Master
  – For all study courses: 30 ECTS and 6 months writing time
Roles

Author

Advisor

Supervisor
Roles

Author
- Creates solution
- Implements code
- Writes text
- Presents work

Advisor
- Invents topic
- Feedback for work
- Evaluation

Supervisor
- Feedback topic
- Formal Evaluation
Agenda

1. Why?
2. What’s important?
3. Choosing a topic
4. Doing the work
5. Presentation
Topics interesting for Author

Topics interesting for Advisor
Sources

- Courses of studies
- Chair webpages
- (Old) topic descriptions
- List of topics at in.tum.de:
  
  http://intranet.in.tum.de/pages/14qh3nhl9lzn/Thesisankuendigungen

Ask potential advisors! They are looking forward to your initial application!
Thesis Application is an Application!

You should address two important points:
• Motivation: why are you interested?
• Strengths: why should you work on the topic?

Optionally
• CV, motivation letter
• Feedback: TUM Career Center
What if I don’t have a Topic in Mind?

- Ask potential advisors for ideas

- As an advisor, I do not expect
  - Students to come up with thesis topics
  - Students to apply only for documented topics
Hallo Herr Juergens,


Haben Sie noch Themen für eine Bachelorarbeit aus den Bereichen Software Engineering/ Testing/ Analyse? Falls ja, könnten wir uns mal auf einen Kaffee treffen, um uns darüber genauer zu unterhalten?

Danke für Ihre Antwort!

Viele Grüße

Roman Haas
How to Find a Topic in Times of COVID-19

• Still the same: contact potential advisors

• Send a strong application mail and convince your potential advisor that he should meet you virtually to discuss thesis topics

• For additional hints, see https://www.in.tum.de/fuer-studierende/coronavirus/
Author interest

Advisor interest

Location factors of environment
<table>
<thead>
<tr>
<th>TUM</th>
<th>Industry</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Contact to research (PhD?)</td>
<td>• Contact to company (Job?)</td>
</tr>
<tr>
<td>• Easier to find advisor</td>
<td>• Easier to find topic relevant in practice</td>
</tr>
<tr>
<td>• Scientific publication more probable</td>
<td>• You might get paid</td>
</tr>
<tr>
<td>• More visionary topics</td>
<td>• Publication more difficult (Exception: Case study)</td>
</tr>
<tr>
<td>• Experience in Supervision</td>
<td>• Synchronization of advisors requires more effort</td>
</tr>
<tr>
<td>⇒ <strong>Insight into scientific work</strong></td>
<td>⇒ <strong>Insight into Company</strong></td>
</tr>
</tbody>
</table>

**Special Case**

Spin-Offs/Research institutes: sometimes best of both worlds
Forschungsarbeiten @CQSE

- Agenda: How to do a BA/GR/MA @CQSE
  - Analysis Implementation
  - Evaluation/ Case Study
  - Advisors
  - Pitch of current topics

- Remote Event in times of COVID-19

- Would like to be invited?

http://cqse.eu/feedback-tum-talk
When does a topic suck?

If you cannot tell, whether a solution for it is good, or not.

Most important factors:

- Is there a clear problem statement?
- Is there a way to evaluate alternative solutions?

Why important?

- Helps you choose between alternatives
- Helps you convince your advisor
- Helps your advisor to convince his professor (your supervisor)
Unterstützung von Sprachentwicklung durch Visualisierung

Images taken from slides from Ludwig
Refactoring Suggestions for Long Methods

Which candidate would you use more likely for an extract method refactoring?
- TOP 1
- TOP 2/3
- Random

Would you use the selected candidate for an extract method refactoring?
- Yes
- Yes, with slight modification
- No
Topic-Antipatterns

- Search my Literature
- Implementation only
- Choose my Tool

- Lack of objective evaluation criteria
- Hard to guide your own work
- Publication difficult
Schedule

- **Internal Thesis**
  - 1-2 Months
  - 4-6 Months
  - 1-2 Months

- **External Thesis**
  - 3-6 Months

- **Spin-Offs/Research Institutes**
  - 2-3 Months
Formal Registration

My Advice: Register immediately.
When to Register?

• My advice: Register immediately!

• Still, there is no risk: you can withdraw without any consequences within the first third of your writing time.
Further Administrative Details

• Registration Forms: CS, Games, Information Systems (W-Info)
• Duration 6 Months
• Extension
  – Because of Sick Leave: No problem
  – Because of other reasons: Potentially complex. Try to address using scoping
Agenda

1. Why?
2. What’s important?
3. Choosing a topic
4. Doing the work
5. Presentation
What characterizes scientific work?

German: Wissenschaftliches Arbeiten: schafft Wissen.
⇒ Scientific work creates knowledge

• Separate opinion / hypothesis from empirical results
• Separate your results from previous work

Goal: Enable reader to trace (or in some cases reproduce) your results.
Author

Advisor
Author Responsibilities

- Time management
- Request feedback
- Communicate problems
- Make decisions

To make advisor happy, report in each meeting
- Status last meeting
- What happened since then
- What you plan to do next, which problems you see
Advisor Responsibilities

- Topic definition and clarification
- Scoping
- Giving Feedback

Not Responsibility

- Micro-Management & Time management
- Take away decisions
- Proof-read complete work
Meet how often?

- Regular meeting
- Meeting on demand
- Feedback for outline
Parts

1. Introduction
2. Fundamentals
3. Related Work
4. Solution Approach
5. Evaluation
6. Future Work
7. Conclusion
Parts

1. Introduction
2. Fundamentals
3. Related Work
4. Solution Approach
5. Evaluation
6. Future Work
7. Conclusion

> 2/3 of thesis
Thesis Architecture

The outline is the architecture of your thesis. It decomposes your document into components (called chapters) with dependencies between them (called references). As for software, the architecture plays a crucial role for the success of your project.

Since text is hard to refactor (much harder than source code), it is tedious manual work to fix an outline that does not work properly later. Minimize this risk by 1) using a standard architecture and 2) early validation of a prototype (through supervisor feedback).

Standard Architecture

A suitable software architecture allows the most important operations to be performed with high performance, even if this increases the cost of other operations. A search engine architecture, for example, is optimized for performance of search queries. Other text operations, say spell checking, are slow or not supported at all.

A suitable thesis outline allows a readers to answer the most frequent questions with high performance (i.e. without having to read the whole document). Which problem does the thesis solve? Did the approach work? What are its results and limitations? Which section should I read for details? Less important questions, say in which order the ideas occurred to the author, may not be answerable at all.
Tools

• LaTeX
• Version Control System (e.g., gitlab.lrz.de)
• JabRef

• Template (e.g.): https://github.com/fwalch/tum-thesis-latex

Formal requirements:

• DE: http://www.in.tum.de/fuer-studierende/pruefungen-und-formalitaeten/abschlussarbeit.html
<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00</td>
<td>Coffee Break, 10:00</td>
</tr>
<tr>
<td>12:00</td>
<td>Lunch, 12:00</td>
</tr>
<tr>
<td>15:00</td>
<td>Coffee Break, 15:00</td>
</tr>
</tbody>
</table>
Agenda

1. Why?
2. What’s important?
3. Choosing a topic
4. Doing the work
5. Presentation
How to Draft Your Presentation

When I listen to a thesis presentation, I need to get the big picture before I care about the details. Until I have understood the problem statement, for example, I do not care how an algorithm works or how its average-case amortized runtime complexity beats existing solutions. Not even if it is presented with a nifty visualization.

In other words, if the big picture of a presentation—its structure—is messed up, no amount of clever visuals can save it. The first priority is thus to get the presentation structure right.

What do I mean by presentation structure? The presentation’s composition of of its constituent parts. How much time you spend on which part. Which one you focus on. Which thesis content you shorten, or leave out entirely. Whether you include an example, presentation. Which running example. And so on.

Creating the structure is about making choices. Much like software architecture decisions, these choices are difficult to change later on. The goal of creating the presentation structure is thus to deliberately make and evaluate these choices early on, when they are still cheap to change.

When I prepared my first presentations, I always started by creating slides right away. This is a fundamentally bad idea, as it distracts from the structure.

As soon as I write my first letter on a slide, it has a font, a size, a color and a position. If I draw objects to form an info graphic, they have relative positions, alignment, proximity, contrast. All these properties force me to make choices. Or accept default settings, which are often ugly. Since ugly slides are irritating, they demand fixing. This ties up my time with tweaking. However, none of the visual properties of slides matter for presentation structure.

The key to focus on structure is to avoid details altogether. Use a medium that does not allow details. Thus shy slide-ware (Powerpoint and such) when drafting presentation structure.

My favourite drafting medium are whiteboard marker, post-its and a wall [1]. Since I am pretty clumsy with the whiteboard marker, and since the post-its are small, I simply cannot do details.

I use the process described by Nancy Duarte in her book Resonate. To make it tangible, I use examples from a presentation I did on presentation design [2].

Brainstorm
Choosing the topic also means choosing an advisor. This is obvious. It also means choosing the amount of time the advisor has. This is less obvious. Let me explain.
How to Rehearse Your Thesis Presentation

A new set of presentation slides is like a program that has probably contains bugs. It reduces the pain for all stakeholders, if you test it to discover (and fix) its bugs before exhibiting yours, if you test it to discover (and fix) its bugs before exhibiting yours.

My test process for presentations has three steps. The first is a test of a program. The second one is in front of a test audience done by representative users. The third one is with your group.

Step One: Alone

The first rehearsal step is to give the presentation to an individual who is not in your group. This is critical to understand where the material is strong and where it needs to be improved. It also helps you to determine whether your delivery style is effective.

Step Two: Test Audience

The second rehearsal step is to give the presentation to a test audience.

Two to four persons make the perfect test audience size. A single person can miss too many problems or be too subjective. More than four add little value but complicate the rehearsal, since everybody wants their say. Personally, I prefer three test audience members.

The test audience comments are the more helpful, the more similar the test audience is to the audience of the final presentation. Computer science students are thus preferable over parents over grandparents. However, take what you can get. A rehearsal in front of your grandparents is still far better than no rehearsal in front of an audience.

I use this process for test presentations with an audience:

- Hand out pen and paper, including a printout of the slides. It is easiest to note down visual slice problems on the slides themselves. It also allows you to collect the notes afterwards.

- Plan at least three times the amount of time of the presentation for the entire meeting (e.g. 60min rehearsal meeting to test a 20min presentation).

http://thesisguide.org/2015/03/04/how-to-draft-your-presentation/
After handing in

• Celebrate!
• Make presentation appointment early
• Communicate your time constraints early to advisor
Avoiding Typical Pitfalls (not only) at TUM

This post is from Roman Haas (he was advised by Elmar in his Bachelor's Thesis).

It focuses on more or less typical problems that appeared to him and his friends during our theses and how to avoid them. They are described by antipatterns, i.e. there is always a description of a problem and a possible solution for it. The problems are sorted by the moment when you should pay attention to them: the first ones may appear at the beginning of your work, the latter ones appear during your work or at the end.

Rating Criteria Surprise

Problem: You are not happy with the final grade of your thesis because you spent too much time on things that are not (so) relevant for the grade.
Solution: Ask your supervisor at the beginning of the work which criteria will be used to rate the thesis. Does the code that you wrote for the thesis influence the grade? What about the presentation at the end?

Wasting Time on Thesis Template

Problem: At the end of your writing time you realize that there are formal errors like formatting issues and content of the cover page in your thesis.
Solution: Look for a thesis template at the very beginning of your work, e.g. ask your friends which template they used. Do also check at the beginning whether all formal
http://cqse.eu/feedback-tum-talk
TUM provides the platform.

It is your responsibility, to create your environment to write a successful thesis.
Thanks!

haas@cqse.eu

@r2h293

More information available at https://thesisguide.org