What Makes a Good Master's Thesis? A Checklist Built from My Feedback Emails*

Roman Klinger

2026-11-02

Over the years I have written many emails to students whose theses I supervise: comments on proposals, on drafts, on slides for the final talk. At some point I noticed that I kept writing the same things. In 2021 I collected the most frequent points in a short thesis checklist on this website, and I have been sending students the link ever since.

For this post I wanted to go one step further and find out what I actually tell students. I found mails I sent to students and then anonymized them (removed header, and also name mentions from the text). Then, I used an LLM to collect the feedback. This left 374 anonymized emails to students (I did send many more, but the anonymization process did fail sometimes). The LLM collected the feedback from these texts, grouped it by thesis chapter, and merged it with my existing checklist. What you read below is the result.

If you only take one thing away, let it be this: a thesis answers a research question. It is not a report on a piece of software you built, and it is not a diary of what you did in which order. Almost everything below follows from that.

Before you start: the proposal

The comment I have written most often in all these years is some version of: this is a research question, not an engineering task. The proposal is where you fix that, so it pays to take it seriously.

- **Ask a research question, don't build a tool.** The goal is new knowledge. Software is only the means you need to answer the question.
- **Make your research questions testable.** Phrase them as hypotheses or comparisons ("A outperforms B on C"). "Is it possible to…?" is trivially answered with "yes" and can never be falsified.
- **Keep them few and distinct.** One clear main question is better than a long list. Number them, and if you compare, name both sides.
- **Always compare something.** Plan at least a baseline and an improved model that differ only in the factor you study. Reporting one F score doesn't really tell you something if you don't know if that's good or bad.

*Published as blog post at https://www.romanklinger.de/blog/2026-11-02-thesis-checklist/

- **Goals say what, the work plan says how.** A goals section can be five lines long. Details such as Python libraries and or doing grid search do not belong there.
- **Write the work plan so that someone else could carry it out.** Imagine handing it to a person with your training who has to deliver exactly what you want. What information would they be missing? Say what each step should achieve, and treat specific tools and models as options. If you commit to one tool, you fail at the first problem.
- **Consider to use the usual shape.** Data, (annotation), trivial baseline, advanced model per research question, evaluation, analysis, writing. You can setup your thesis differently, but you should have reason why you do.
- **Milestones are deadlines.** Give concrete dates, not durations or "March to April". One milestone every two to four weeks works well.
- **Plan the writing.** Expect about two months. Write in parallel with your experiments. You might figure out things while writing that you still want to do in an experiment.
- **Check for show-stoppers early.** Can you get the data, are you allowed to crawl it, does an annotated corpus exist? Name a fallback.
- **Mark optional goals and start with the safe core.** Optional goals that you do not reach do not count against you.
- **Don't hope, and don't try.** A negative answer to a research question is as valuable as a positive one.
- **The proposal is preparation, not thesis time.** It typically takes about a month, including the reading, and several rounds of feedback before you register.

Introduction

Introduction and conclusion are the frame of your thesis. Everything in between exists to answer the questions you pose at the start.

- **Introduce and motivate the topic clearly.**
- **Motivate from a research gap, not from the method or from yourself.** "LDA exists, so let's test it" is not a motivation, and neither is personal interest. Make the reader feel the pain that makes your research necessary.
- **Problem before solution.** First explain why the phenomenon matters, then let the need for your method follow.
- **State your research questions or hypotheses explicitly.**
- **Show a concrete example early,** ideally on the first page. Examples are always good.
- **Back up claims with citations, also in the introduction.** "It is well known that…": known by whom?
- **Make your contribution explicit,** compared to the closest prior work.
- **Define central terms and the exact task** before you use them, and say what is out of scope.
- **Stay high-level.** Detailed related work, metrics and implementation come later.
- **Frame it around the scientific problem,** not (in the first place) the company or project that triggered it.

Background and related work

Related work is not a list of papers. It shows the state of the art as a whole and where your own work starts.

- **Another student of your program should understand the thesis without reading anything else.** A thesis is not a paper: explain every method you use, with formal definitions.
- **Say why each background topic is there,** and cut what you never use later.
- **Describe related methods well enough that your choice of method is understandable,** including the alternatives you did not pick.
- **Describe all the work that is your starting point.**
- **Compare and group, don't list.** Use thematic subsections. For each group, say what the papers share, how they differ, and how your work relates to them.
- **Be comprehensive in your core area.** One paper per topic looks anecdotal, and gaps suggest missing knowledge.
- **Cover the phenomenon in its home discipline,** for instance psychological theories of emotion and how they are measured, not only the NLP methods.
- **Give comparable sections similar depth.** If you use SVMs and neural networks, both deserve the same attention.
- **Search systematically.** Use the ACL Anthology and Google Scholar, and follow both what a good paper cites and who cites it.

Methods and data

The methods chapter describes what you built, on the conceptual level, and precisely enough that someone could rebuild it in ten years.

- **Separate concept from implementation.** Don't mix code snippets with mathematical explanations. That you wrote a script or stored things in JSON is irrelevant – the concrete file format is probably an important, but arbitrary choice. You could also use XML without loss of generality. If you want to write about implementation, add a section "Implementation Details".
- **Make clear what is yours and what you reused.** Don't hide your own model development under "Experiments", or it looks as if you only ran existing tools.
- **Start with an overview figure of the whole system,** then explain each component.
- **Describe methods formally where you can,** with examples next to the equations.
- **Motivate every design choice.** Otherwise the features fall from the sky.
- **Describe your data completely.** Sources, collection dates, splits, and a statistics table (number of documents, average length, labels per class).
- **Describe the annotation completely.** Who annotated, which guidelines, which training, how labels were aggregated, the agreement, the payment, and a screenshot of the interface.
- **Keep the setup out of the methods.** Datasets, experimental settings and user studies belong in the experiments chapter. Imagine the methods to contribute to a system that can be of use later. The experimental setting section explains how you make use of the methods to answer the research questions.

- **Probably you want to call it “Methods”, not “Methodology”.** It is not the same thing.
- **Give corpora, models and experiments short names** that the reader can remember. “Small model” and “big model” are easier to understand than M1 and M2.

Experiments and results

Organize your results around the research questions, and never leave the reader alone with a table.

- **Design the experiments to answer the research questions, and explain each one.**
- **An experiment is a comparison.** You change one factor and observe the effect. Running one model on one dataset only gives you numbers.
- **Structure the chapter by research question:** question, setup, result, explicit answer. Open with a paragraph that tells the reader where the chapter is going.
- **Prepare tables so that the reader sees how they answer a research question.** One main comparison table is better than many small ones. Details go to the appendix.
- **Talk the reader through every table and figure:** what you want to find out, how the table is organized, what you see and where you see it.
- **A reader should be able to repeat all experiments** with the information in the thesis.
- **Always include trivial baselines** (random, majority class), and always say what an improvement is relative to.
- **Also pick strong baseline**, if available.
- **Select models on validation data.** Hyperparameter tuning is not a result.
- **Analyze why something works or fails.** Count error classes and show concrete examples. An ablation study is not an error analysis.
- **“Significant” requires a significance test.** Don’t over-interpret tiny differences. Better avoid using the term if you didn’t do a statistical test.
- **Close to the deadline, stop experimenting and write.** Less work, described properly, is better than more work described badly.

Discussion, conclusion and abstract

Reading only your introduction and your conclusion should give me all your findings.

- **Answer the research questions from the introduction.**
- **Summarize the main take-home messages.**
- **Mark speculation as speculation.** “This might be because…” is fine. Keep facts and interpretation apart.
- **Discuss limitations and alternative explanations that are specific to your work,** not generic ones.
- **Future work names the research questions your findings raised,** not things you would have liked to do.

- **Keep the quality up to the last page.** Language and argumentation often get weaker in the final chapters.
- **The abstract mirrors the thesis:** problem, goal or research question, method, results, conclusion. Be concrete, don't tease ("the reader will see…"), and leave out citations.

Citations and bibliography

Every statement in your thesis is either intuitively clear to every reader or backed by a reference.

- **Cite all work you used.** Never leave the impression that you came up with something yourself when you didn't.
- **Wherever someone's work shaped your text, say so.** Copying sentences or figures is plagiarism.
- **Cite peer-reviewed work.** Avoid blog posts and Wikipedia as much as possible.
- **Cite arXiv papers only if no published version exists,** for instance in the ACL Anthology.
- **Make every entry complete,** so that a reader can easily find the source.
- **Make entries consistent:** first names always or never abbreviated, venue names as in the official publication, volumes, numbers and pages where available.
- **Don't trust BibTeX from Google Scholar.** In my experience it is more often wrong than right. Use the ACL Anthology and read the warnings of BibTeX or biber. Many publishers of journals also fail to deliver correct bib files. AI anyway often fails to do so.
- **Check that the paper you cite actually says what you claim it says.**
- **Use `\citet` and `\citep` correctly.** Write "Klinger [11] showed", not "[11] showed", and put several citations into one bracket.
- **Use BibTeX from the proposal on.**
- **Check whether you may reuse a figure.** If not, redraw it and cite the original.

Writing and style

A thesis is a scientific text, not a diary of what you did in which order.

- **Don't write a progress report.** "I did this, observed that, and then I did…" does not belong in a thesis. Research exists without a time dimension.
- **Take the reader by the hand.** Keep a red thread. In every section, the reader should know which research question it serves.
- **Be precise and concrete.** Avoid "some", "many" and "etc.". Say how something works, not only that it exists.
- **Use terms precisely and consistently.** It is emotion analysis, not emotional analysis. Avoid buzzwords like "AI", and LLMs don't "think".
- **Introduce every abbreviation,** and avoid abbreviations unless you use them very often.
- **Keep a neutral scientific register.** No contractions, no colloquialisms, no pathos, no "naturally", no self-praise.
- **Commit.** Don't "try" to do something. The system does it or it does not.

- **Plan before you write.** Section goal, list of points, order, then sentences. One idea per paragraph, and check every transition.
- **Prefer active voice.** “We” means you and the reader together.
- **Capitalize only proper nouns.** The existance of a common acronym does not make “natural language processing” a proper noun. “Table 5” and “Section 3” are.
- **Proofread.** Let the text rest for a few days and read it on paper.
- **Read Higham’s Handbook of Writing for the Mathematical Sciences,** in particular Sections 1, 2, 3 and 5.
- **Write the text yourself.** Text generated by a language model shows and breaks the flow. Declare any use of such tools.

Figures, tables and math

A figure or table should be understandable from itself and its caption.

- **Captions explain what is shown,** including the metric, abbreviations, and what bold face means.
- **Don’t force placement with `[h]`.** Floats belong at the top of a page and are referenced in the text.
- **Use vector graphics that are readable without zooming.** No screenshots, no squashed pixel images.
- **Use booktabs and no vertical lines.** Align digits, two decimals are often enough, and no engineering notation (write 0.000034 or 3.4×10^{-5}, not 3.4e-05).
- **Format numbers sensibly.**
- **Set multi-letter function names upright** (`\operatorname{softmax}`, `\textrm{precision}`). Otherwise LaTeX typesets the product s·o·f·t·m·a·x.
- **Equations are part of the sentence,** including punctuation. Define every symbol and index.

Structure and formalities

Readers expect the standard structure. Deviate only if you have a reason.

- **Use the conventional outline:** Introduction, Background and Related Work, (Corpus), Methods, Experiments and Results, Discussion and Conclusion.
- **Put each thing where it belongs.** No results in the methods chapter, no new methods in the results chapter.
- **Avoid singleton subsections,** and don’t write more than two or three pages without a subheading.
- **Headings tell the reader what to expect.** Use names a reader understands, not invented labels.
- **Complete the title page:** title, study program, supervisors and examiners with their correct roles and titles, dates.
- **Keep your LaTeX clean:** blank lines between paragraphs (not `\\`), proper quotation marks and dashes, no unresolved “??”, chapters starting on right-hand pages, URLs printed in full.
- **Length is a signal.** There is no minimum, but a master’s thesis far below about 80 pages usually misses something.

Working with your supervisor

Feedback right before the deadline is useless, because there is no time left to act on it.

- **Send a full draft at least two weeks before submission.** I need about a week to read it (a careful read takes four to eight hours), and you need about a week to work in the comments.
- **Send one complete PDF** with title page, table of contents and bibliography, not fragments, and tell me where to focus.
- **My comments are examples.** Apply them everywhere they fit. If I don't comment on a section, that does not mean I like it as is. A comment somewhere else might also apply here.
- **Address every comment** and mark what you changed.
- **Bring concrete questions and your own ideas** to our meetings, not only problems.
- **It is your thesis.** You may disagree with a suggestion, and that is absolutely fine (actually: appreciated) but let's talk about it.
- **Contact the examination office early** about formalities and deadlines. You can also contact our secretary for help.

The final talk

The talk follows the same logic as the thesis: problem first, then questions, then answers.

- **Open with a motivating example,** not with an agenda. Show the outline only once the goal is clear.
- **Present results question by question.** Each slide states the question it answers.
- **Less text, more pictures.** Keep the font size constant and readable, and build complex figures step by step.
- **Number your slides,** so that people can ask about slide 12.
- **Plan about one slide per minute.** Details go to backup slides. Skip the basics everybody in the room knows.
- **End with take-home messages,** not with a references slide. Supervisors and reviewers belong on the title slide.
- **Send me your slides one to two weeks ahead,** so that there is time for a round of feedback.
- **Come to the final talks of other students.** We expect you to attend them; it is also the best way to see what works in a talk. If you don't receive the invitations, ask me to add you to the mailing list.

How my feedback changed over the years

Going through almost two decades of emails was also a look in the mirror. My feedback did not change direction, but its focus moved from "write clearly" to "ask a real question".

- **Research questions moved to the centre.** Research questions were part of my feedback from the start, but over the years they became the point I make most often: a thesis answers a testable question and is not an engineering project.
- **Planning became a process.** Around 2018, proposals got a fixed structure with dated milestones and several feedback rounds. In 2021, the comments I kept repeating became the checklist on this website.
- **Generative AI is the one new topic.** It only shows up in the last few years: declare any use, generated text is noticeable, and language models don't "think".

Where I have not always been consistent

I did not find a case where I flatly contradicted myself, but there are places where I told different students slightly different things. Here is what I actually mean.

1. **How detailed should a proposal be?** I ask for a work plan so concrete that someone else could carry it out, and I also tell students not to over-detail it with preprocessing steps or premature model choices. Both are true: be concrete about what each step must achieve and how you will decide, and leave open what you cannot know yet.
2. **Hypotheses.** I ask for explicit hypotheses, but I also object when the introduction already presents the expected answer as a fact. A hypothesis is a claim you test, not a result you announce.
3. **Length.** I say that length does not matter, and I also say that below about 80 pages something is usually missing. Length is not a goal, but a short thesis is a warning sign.

Things I keep writing

Some sentences show up in my emails again and again. If you are one of my students, you have probably read at least one of them.

- "We do research; development is just needed to answer the research question."
- "Make the reader feel the pain."
- "Don't leave the reader alone with the interpretation."
- "Take the reader by the hand."
- "Imagine somebody else should do the work for you."
- "Milestones have a deadline, not a starting date."
- "An experiment is a comparison, not a model run."
- "A 'no' is just as good a result as a 'yes'."
- "Numbers are only informative in comparison to others."
- "Sometimes less is more; you don't have to show everything."
- "Abbreviations help the writer, but very seldom the reader."
- "Examples, examples, examples."

I hope this helps you write a thesis that you are proud of.