Student guide for reports and theses at MLAI

Abstract

This documents collects some guidelines for writing theses, seminar reports, and lab reports within the MLAI group at the University of Bonn. Generally, you should ensure that the reader always has to know what a section (subsection, paragraph) is about and how it is relevant to the overall context (**red thread**), that you explain everything in a way such that a colleague of yours would understand it (**target audience**), that you define everything properly and uniquely and do not contradict yourself (**inner logic**), that you correctly reference everything that is not your own work (**no plagiarism**), and that you motivate and explain all choices you made and point out strengths and flaws of your own work (**critical self-reflection**).

Contents

1 Initial Remarks

Writing is a skill that needs to be refined over time. We hope this document will give you a head start, not only by providing best practice advice, but by directing your attention to certain things that will enable you to better reflect on the quality of your document.

Disclaimer This document is meant as a guide to give you an idea of our expectations regarding written presentation of your results. It can not and does not constitute an exhaustive list of our formal requirements. Thus this document does not provide a base for arguments over grades.

1.1 Conventions Throughout this Document

This document collects hints and best practices for scientific writing in general and the various text formats that students have to hand in at our group in particular. Whenever it is necessary to specify different parameters for different text types, we use the following acronyms:

BA Bachelor's thesis

MA Master's thesis

PG Projektgruppenbericht

SR Seminar report

LR Lab report

2 Scientific writing in general

This section contains best practices, tips and tricks on good formal, scientific writing. Presenting your work in an accessible way is as much part of the scientific process as developing methods and conducting experiments. A well structured text reflects sound methodical thinking and makes it easier for the reader to concentrate on the content of your work. [5]

2.1 Language

- Style
  - Use present tense (exception: Related Work and parts of Conclusion)
  - No contractions: don't → do not, hasn't → has not
  - Use active voice "X conducted the experiments" not "the experiments were conducted by X"
  - Avoid long sentences spanning multiple lines (it is okay to write long sentences initially so you have all your thoughts collected in the document. When subsequently improving this first draft, break down those long sentences into shorter ones so the reader does not get lost.)
  - Chose between British and American English and stay consistent within it
  - Use spell checkers
- Instead of "I" use "We", as in you, the author, walks the reader through your thinking process

2.2 Formalia

- The document has to be a **PDF**.
- Length
  - PG, SR, LR: 8-10 pages per person required
  - BA: required are 25-50 pages (we recommend 40)
  - MA: required are 25-100 pages (we recommend 50)
  - Counting starts with the Introduction (first chapter) and ends with the Conclusion (last chapter)
- Citations
  - Citations style **either (Author, year) or [num]**. We recommend the first but chose what fits your style of writing best as long as you stay consistent and don't mix both citation styles.
  - Citations are placed at the **end** of the sentence, before the dot ("silent" reference). Exception: "Horváth et al. (2020) have shown …" ("loud" reference).
  - You need to mark **everything** with a reference that is not part of the current contribution or the current result (i.e. your own work) of this particular document. In particular, referencing your own older work (e.g., a seminar report from last semester) requires a citation. This can include figures (also recreated ones), tables, algorithms, lemmata, results, arguments etc.
- Bibliography
  - Not all types of sources (web pages, proceedings, books) come with the same set of meta data (author, year, edition, date of access, link, chapter, pages) neither is necessarily all the information relevant. Present the information that is necessary for the reader to find the cited source (and perhaps the relevant section) but avoid overloading the entries with redundant or unnecessary information.
  - Scientific context/ related work: journal article > conference > workshop paper
  - Basics: textbooks
  - Misc: web pages
  - Datasets: available on websites but most of them are associated with a publication

2.3 Notation and conventions

- Adhere to field specific conventions regarding vocabulary and technical terms, variables and acronyms.
- Avoid using ambiguous words (normal, canonical …) except when properly introduced beforehand
- Definitions should be as formal as possible, a prose explanation of something is good style but only complementary.
- Different sources can use the same variable with different meanings. When citing you can adapt those definitions as to avoid ambiguous use for the variables within your document. Stay **consistent** within your document!
- See Section 8, esp. [2, pp. 1-8]

3 Visual Presentation

In addition to a well structured content a tidy and consistent visual presentation greatly benefits readability. We highly suggest you use LaTeX.

3.1 Text

- No headline after a headline
- Use a readable font
- Font size: 10–12pt

3.2 Non-Text elements

- LaTeX: **Floats** [1]
  - $\exists \Rightarrow$ use them
  - Floats wrap around nearly everything that is non-text. They enumerate themselves automatically, can be referenced by name set in \label{…} via \ref{…} and provide a place for a \caption{…}
  - Each float needs to be referenced and explained in the text
  - Each float has a caption that gives a standalone explanation
  - Place floats close to where they are referenced to in the text
- On a plot each axis needs to have a label
- When citing a table from another document, do not use a screen shot bitmap, but instead rebuild the table
- Never use image files such as .png / .jpeg when avoidable. Use **vector graphics**!

4 Structure of a thesis

Most theses (and articles, for that matter), are similarly structured. A similar structure helps your more seasoned readers (think: examiners) to focus on the content (i.e., what you did). It is good to adhere to the standard, everything that distracts the reader from the content that you want to present should be avoided.

Your content and results should be creative, not your layout. We provide a LaTeX-template for theses and reports on our homepage. You may use it as you see fit and change it according to your liking for your thesis.

1. Title page
   - Title
   - Name
   - Date
   - Supervisor
   - Examiner(s)
   - *optional*: Logo of University of Bonn and the MLAI group
2. Abstract
   - BA/MA: max 1 page
   - Reports: max $\frac{1}{4}$ page
   - content: what, why, how, results
3. Acknowledgments/ inspirational quote (optional)
4. Eigenständigkeitserklärung
5. Table of content
6. Lists of Figures, Tables, Abbreviations and Variables (or after bibliography; BA: optional, MA: probably helpful)
7. Introduction
   - cf. Section 8, esp. [5]
   - motivation
   - gist of related work
   - research questions in prose
   - outline and structure of thesis
8. Related Work
9. Preliminaries
   - Formally introduce definitions, notation and concepts required to understand the remainder of the document
   - Target audience: explain everything with as much detail such that a fellow student of yours would be able to understand the document completely. If something was new to you before the thesis: explain it.
   - You can (and should) give illustrative examples using pictures/ figures to aid comprehension
10. Main section
    - This can span multiple chapters
    - start with a formal definition of your problem/ research question(s)
    - describe and motivate your solution; take your time, describe preliminary experiments and trials, visualize examples

[1] https://en.wikibooks.org/wiki/LaTeX/Floats,_Figures_and_Captions

- how will you solve the problem
- why did you chose to do something one way over the other (if two ways are basically equivalent, state this fact and say that you simply preferred doing it this way)

11. Experiments
    Before conduction
    - Describe datasets
    - Setup
    - What metrics will you report within one experiment (e.g. error or loss) and across experiments (e.g. standard deviation on results)
    - Baselines/ frames of reference (e.g. other algorithms)
    - (Briefly) describe all hyper-parameters (of all the methods you use) you had to chose and explain your choice
    - Explain how your solution will solve the research problem

    After conduction
    - Present the results, if possible visualize them
    - complete, extensive and objective description of results (this may feel very dry and weird to write at first)
    - interpret the results
    - point out expected/ unexpected results
    - determine whether the results constitute a success regarding the research questions
12. Conclusion and Future Work
    - Summary of thesis
    - Repeat research question(s)
    - What was your contribution to solve it
    - Result of the interpretation
    - What could be the next steps or further avenues of research from here on?
13. Bibliography
    - See Bibliography in 2.2
14. Appendix
    - Additional Figures
    - lengthy/ detailed proofs

5 Scopes and scales

5.1 Theses

Bachelor Thesis

- Independent research of literature and interpretation of results.
- Appropriate presentation as a scientific document
- Citing mostly text books and only few (proper) papers is okay

Master Thesis

- Work autonomously
- Considerable length
- Solve a significant research problem
- Presentation of the research problem and solution within the context of current research
- Reflecting knowledge on the current state of the art
- Working with primary sources: Papers and the most recent research
- Adhere to scientific principles
- Demonstrate ability to creatively apply knowledge on the problem at hand

5.2 Reports

Project Group

- See BA in 5.1 but less content and less literature research
- Work together as a group

Seminar Report

- Summarize and discuss scientific papers autonomously
- Summarize content within the context of the seminar
- Discuss, analyze and compare content with fellow students and try to see the overall direction of research

Lab Report

- MA in 5.1 but less content

6 A Possible Timeline for your Thesis

It is your task to structure your own work. Part of writing a thesis is showing us that you are able to work independently on a scientific topic. Hence, we will not force you to do anything or to adhere to any structure that we might feel suitable for a thesis. However, Figure 1 provides a generic timeline for your thesis that you might consider as a starting point for your own time table.

7 Supervision

When writing a thesis or a report, you will have a supervisor who is a member of our group.

8 Helpful Resources

Many smart people have written on writing. Here are a few references that extend (and probably partially contradict) this compact document. Feel free to have a look!

- "How to Read a Paper" by Keshav [1]
- *Writing for Computer Science* by Zobel [5]
- *Mathematical writing* by Knuth, Larrabee, and Roberts [2]
- *How To Speak by Patrick Winston* by Winston [4]
- *The elements of style* by Strunk and White [3]

References

[1] S. Keshav. "How to Read a Paper". In: *SIGCOMM Comput. Commun. Rev.* 37.3 (July 2007), pp. 83–84. ISSN: 0146-4833. DOI: 10.1145/1273445.1273458. URL: https://doi.org/10.1145/1273445.1273458.

[2] Donald E. Knuth, Tracy L. Larrabee, and Paul M Roberts. *Mathematical writing*. Vol. 14. Mathematical Association of America, Washington, DC, 1989.

[3] William Strunk and E. B. White. *The elements of style*. 4th edition. Pearson, 2000.

[4] Patrick Winston. *How To Speak by Patrick Winston*. MIT OpenCourseWare. 2019. URL: https://www.youtube.com/watch?v=Unzc731iCUY (visited on 09/24/2020).

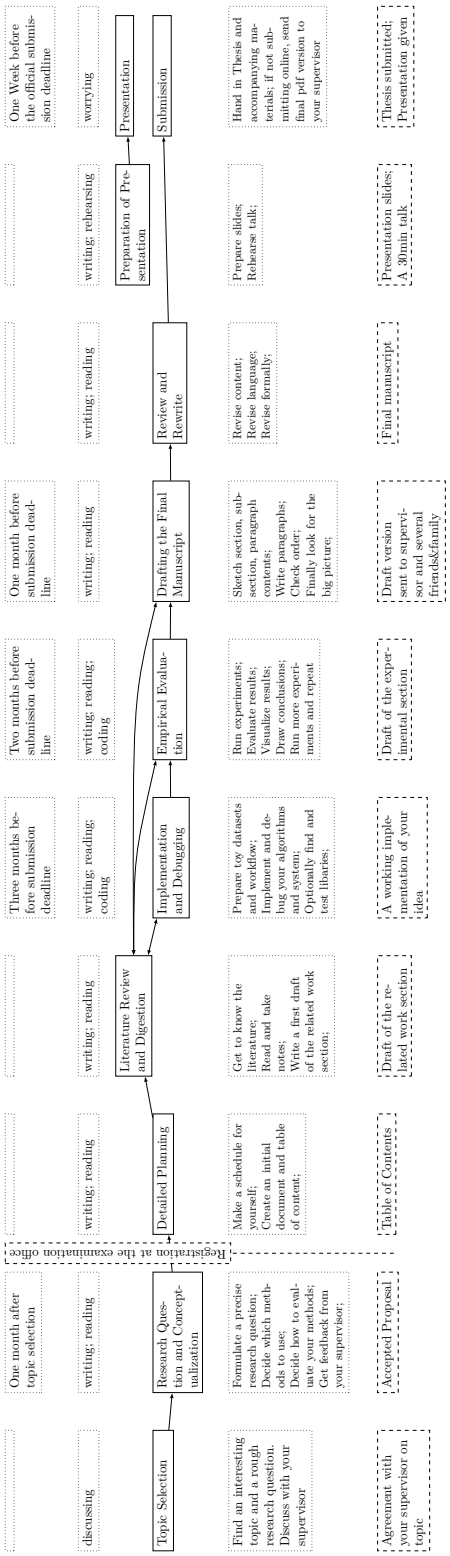

Figure 1: A possible timeline of your thesis.

[5] Justin Zobel. *Writing for Computer Science*. 3rd ed. London: Springer-Verlag, 2014. ISBN: 978-1-4471-6638-2. DOI: 10.1007/978-1-4471-6639-9. URL: https://www.springer.com/gp/book/9781447166382 (visited on 08/17/2020).