

Student guide for reports and theses at MLAI

Abstract

This document 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	
1.1 Conventions Throughout this Document	
2 Scientific writing in general	
2.1 Language	
2.2 Formalia	
2.3 Notation and conventions	
3 Visual Presentation	
3.1 Text	
3.2 Non-Text elements	
4 Structure of a thesis	
5 Scopes and scales	
5.1 Theses	
5.2 Reports	
6 A Possible Timeline for your Thesis	
7 Thesis Proposal / Thesis Exposé	
8 Helpful Resources	
A Checklist	
B List of required Latex packages	

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
✗ the experiments were conducted by X →
✓ X conducted the experiments
- ✗ Sentences spanning multiple lines, especially those that contain many nested sentences, multiple arguments, many examples and enumerations should be avoided as to not unnecessarily fatigue the reader such that concentrating on the content is as easy at the end of the document as it was in the beginning. → ✓ Avoid sentences that span multiple lines. Convoluted sentence construction is especially tiresome in longer documents.
It is okay to write long sentences initially just to 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.
- Choose 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)
 - the pagecount starts with the Introduction (first chapter) and ends with the Conclusion (last chapter).
 - The Bibliography is not part of the page count and you can have as many References as you like, without cutting into your page budget.
- 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).
✗ ... by now is a well established fact. [42] → ✓ ... by now is a well established fact [42]. ✗ ... by now is a well established fact. (Adams, 1979) → ✓ ... by now is a well established fact (Adams, 1979).
- Exception are "loud references", that is if the sentence is broken without using the reference explicitly: "Horváth et al. (2020) have shown ...". Note that (Author, Year) citations differ in this case:
✗ (Welke, 2020) uses references incorrectly. → ✓ Welke (2021) shows that people can learn from past mistakes.
(Note the different placement of brackets)
- 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: Most are associated with a scientific publication. If they are distributed via a website this is also the place where you can usually find information on how to cite them properly.
- Different L^AT_EX-packages are available to handle citations and references.
 - natbib allows (Author, Year), as well as [num] styles. It has dedicated commands \citep{} for "silent" and \citet{} for "loud" citations.

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 L^AT_EX.

3.1 Text

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

3.2 Non-Text elements

- L^AT_EX: 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
- Each axis on each plots needs to be labeled
- 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**!

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

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 L^AT_EX-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

- *Context*: Start with a general introduction of the the research field and steer towards your specific topic.
- *Problem statement*: Roughly outline the problem that this work is concerned with. Formalize the task and mention what the goal is.
- *Relevance*: Explain why this problem is relevant (i.e. motivate using e.g. applications) and why solving it isn't trivial.
- *Method*: Discuss how the problem is approached and finally solved. List necessary 'ingredients' that are part of the method.
- *Related work*: Very shortly mention related work concerned with the same or a closely related problem.

- *Contributions*: Summarize all main contributions and results of your work.
- *Thesis structure*: Briefly outline the structure of the remaining sections.
- Further hints:
 - It is advisable to include an easy to understand example, ideally using a visualization.
 - After reading the introduction, the reader should already have a rough idea about all major contents of your work.
 - More detailed information on writing introductions is provided in e.g. [5].

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
- 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

- | Deadline | Ongoing Activities | Work Packages | Tasks | Mile Stones |
|---|--------------------------|---|---|---|
| One month after topic selection | discussing | Topic Selection | Find an interesting topic and a rough research question. Discuss with your supervisor | Agreement with your supervisor on topic |
| One month before the official submission deadline | writing; reading | Research Question and Conceptualization | Formulate a precise research question; Decide which methods to use; Decide how to evaluate your methods; Get feedback from your supervisor; | Accepted Proposal |
| Three months before submission deadline | writing; reading; coding | Literature Review and Digestion | Get to know the literature; Read and take notes; Write a first draft of the related work section; | Draft of the related work section |
| Two months before submission deadline | writing; reading; coding | Implementation and Debugging | Prepare toy datasets and workflow; Implement and debug your algorithms and system; Optionally find and test libraries; | A working implementation of your idea |
| One month before submission deadline | writing; reading | Empirical Evaluation | Run experiments; Evaluate results; Visualize results; Draw conclusions; Run more experiments and repeat | Draft of the experimental section |
| One month before the official submission deadline | writing; reading | Drafting the Final Manuscript | Sketch section, subsection, paragraph contents; Write paragraphs; Check order; Finally look for the big picture; | Draft version sent to supervisor and several friends&family |
| | writing; re-reading | Review and Rewrite | Revise content; Revise language; Revise formally; | Final manuscript |
| | writing; rehearsing | Preparation of Presentation | Prepare slides; Rehearse talk; | Presentation slides; A 30min talk |
| One Week before the official submission deadline | worrying | Presentation | Hand in Thesis and accompanying materials; If not submitting online, send final pdf version to your supervisor | Thesis submitted; Presentation given |

Project Group

- ## Seminar Report

- # Lab Report

- ## 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 Thesis Proposal / Thesis Exposé

Before you formally register your thesis at the examination office we expect you to have a written exposé ready. This piece of writing has to be agreed upon by yourself and your supervisor. You can see it as an agreement on the scope of your thesis. It should be 1–3 pages and contain

- your detailed research question(s),
- an initial (brief) review of relevant related work,
- a short overview of methods and experiments that you want to apply to answer your question(s), and

Figure 1: A possible timeline of your thesis.

- an outline of the structure of your thesis.

Your exposé can ideally serve as the basis of your actual thesis.

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](https://doi.org/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).
- [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](https://doi.org/10.1007/978-1-4471-6639-9). URL: <https://www.springer.com/gp/book/9781447166382> (visited on 08/17/2020).

A Checklist

Scientific Writing		ToDo	Done
Work is in present tense (except for Related Work and parts of Conclusion)			
Active voice			
No contractions in the work			
No sentences spanning multiple-lines			
The text is consistent, British or American English			
The text is spell-checked			
The length of the work matches the requirements			
The citation style is consistent and correct			
The bibliography is correct and complete			
Every result which is not from this particular document is referenced			
There is no use ambiguous words			
The definitions are formally given			
The notation is consistent			
Visual Presentation		ToDo	Done
There is no headline after a headline			
The font is readable with font size 10-12pt			
All floats have a caption, reference and standalone explanation			
All plot axes have labels			
There is no screenshot of a table			
All graphics are vector graphics or tizpictures			
Structure		ToDo	Done
The work is structured properly, according to Section 4			

B List of required Latex packages

In this sections you can find a list of the packages used for the different templates.

- Preamble:
 - polyglossia
 - algorithm
 - algpseudocode
 - enumitem
 - csquotes
 - metalogo
 - fancyvrb
 - varioref
 - hyperref
 - cleveref
 - biblatex
- Report and Thesis:
 - geometry
 - translations
 - xcolor
 - graphicx
 - booktabs
 - mathtools
 - ntheorem
 - iftex
 - inputenc
 - fontenc
 - textcomp
 - tgpagella
 - fontspec
- Thesis:
 - fmtcount
 - emptypage
 - subcaption
 - caption