

Master Thesis Important Deadlines:

1. Milestone Thesis Design | João Deadline: **10 February** | UvA Deadline: **19 February**

Should include:

- **Introduction**
 - Problem statement, context/field, scientific relevance/research gap and research questions
- **Related Work**
 - Grounding your work in the literature, coining state-of-the-art research and hypotheses
- **Methodology**
 - Describing and justifying methods, data, reliability/validity and plan for the [evaluation](#)
- **Risk Assessment**
 - Describing the risks and presenting a Plan B for each of them
- **Project Plan**
 - Describing a timeline via a Gantt chart or table with achievements per week

2. Milestone Exploratory Data Analysis | João Deadline: **10 March** | UvA Deadline: **19 March**

You want to know your data at several levels:

- Corpus
- Variables by themselves
- Interaction between variables

Be sure to upload the structure of your dataset. List all the features. Please find one extra dataset, next to that set provided by your internship (or university project). The best option is open source with published results and scores. In this way, you know what your model should aim for. Even if your model performs poorly on your internship data but is somewhat in line with the outside data set, you can get the master thesis to a pass. You will then need an error analysis with some cunning hypotheses on the causes of this strange mismatch. This extra dataset gives you a piece of mind, and if something goes wrong, a direction to look for reasons.

3. Milestone Draft Methodology Section | Send to João: **7 April** | UvA Deadline: **16 April**

Why did you choose the current methodology over other plausible alternatives? Does the research design probably measure the effect and is it credible in terms of reliability and validity? How did you go about solving or making progress on the research problem? In

particular, you pay attention to that new little thing, that change, that great idea, that you add to the existing method.

Use most space for your addition. You can make a separate section on the experimental setup (or incorporate it in the methodology section). The function of the experimental setup is that the reader can replicate your experiments. You give *all* the setting you used in your experiments. The description of the data is also incorporated in this section. You can use insightful graphs from your EDA (if relevant in your case).

4. **Milestone Draft Results Section** | Send to João **12 May** | UvA Deadline: **21 May**

How are the outcomes related to your research question in the form of a table or graphic? Structure the section in such a way that the reader only has to look at these two things (and can safely skip all else): the question and the table/figure with (elaborate) caption. It is important that tables/figures are unambivalent. There is much need for a perfect caption, perfect labels and smart design of table or figure. The reader should be able to use 100% of his or her brain to understand the outcomes, not to try to figure out what was meant. Please note that you explicitly answer your research questions in the conclusion, not in the results, given the limitations you qualify in the discussion.

Most good computer science papers conclude that something is so many percent faster, cheaper, smaller, significant or otherwise better than something else. Avoid vagueness, handwaving results in terms as "very small". If you must be vague, you are only given licence to do so when you can talk about orders of magnitude of improvement. There is a tension in that you should not provide numbers that can easily be misinterpreted, but on the other hand you do have room for all the caveats.

5. **Milestone Draft Thesis** | Send to João **9 June** | UvA Deadline: **18 June**

At this point, it is important to pay special attention to the discussion section. You go into an explanation why you did (not) find what you expected beforehand. You should critically reflect on concepts like generalizability, reliability and validity. Formulate options for further research that follow from your research. Be sure that the limitations are reflected in the conclusion as a qualification.

The function of the discussion section is to compare your results with the *state of the art* and to reflect upon the results and limitations of the study. You can start with a section in which you compare the results of your research with previous studies. Be sure to make use of

concrete results. Make use of references. If you did not find what was expected, it is important that you go deeper into possible reasons for this. Are there comparable studies in which no results were found? What would be the mostly likely scenarios for not finding the expected results? It is important that you are able to justify that fixable problems with the research set-up are not the most likely reason for not finding the expected results. It will help to formulate a theory for your results.

The limitations of the study should be noted. Preferably, they are related to key concepts like reproducibility, scalability, generalizability, reliability and validity. Make sure that if the issues are serious that you make clear how you take the limitations into account in the conclusion. You can already hint to future work to which you come back in the conclusion section. Consider whether there are alternative conclusions that are consistent with the results that have been presented. It is helpful if you already discuss what the value of your research is considering previous research. Does it fill the research gap? This gives a bridge to the conclusion.

6. **Final Thesis Deadline: 30 June**