Heidelberg University
Institute of Computer Science
Data Science Group

Master's Thesis

Here Comes the Title

Name:	Name of the author
Matriculation Number:	1234 5678 9
Supervisor:	Prof. Dr. Michael Gertz
Date of Submission:	November 27, 2026

Declaration of Independent Work

SURNAME, First name:
Student ID No.:

I hereby declare that I independently prepared the [Bachelor's][Master's] thesis submitted by me on [DATE] as a PDF file (readable and searchable using standard PDF reader software), entitled

[TITLE]

supervised by Prof. Dr. Michael Gertz, without any unauthorized assistance from third parties. I have used no sources or aids other than those specified and have clearly identified all passages taken verbatim or in substance from published or unpublished works, including those from digital or AI-based sources.

I confirm that any use of AI tools was discussed in advance with the supervisor of the thesis and that the agreed-upon rules were observed.

I assume full responsibility for the scientific quality and content of this thesis, the methodology chosen and the process used to prepare it, as well as for the literature cited.

I agree that my thesis may be checked as part of university procedures for plagiarism as well as for the potential misuse of automated text and code generation.

Place, Date	Signature

Disclosure of AI Tools

In accordance with the Declaration of Independent Work, the following table summarizes all AI tools used during the preparation of this thesis and the specific tasks for which they were employed. I reviewed and verified all code, experimental outputs, literature interpretations, and written text, and I take full personal responsibility for them.

AI Tool	**Purpose and Scope of Use**
Example: Gemini	Used to translate English abstract into Zusammenfassung.
Example: Claude Code	Assisted in writing implementation and evaluation code for packages X, Y, and Z faster and keeping track of experimental result files.

Zusammenfassung

Die Zusammenfassung muss auf Deutsch **und** auf Englisch geschrieben werden. Die Zusammenfassung sollte zwischen einer halben und einer ganzen Seite lang sein. Sie soll den Kontext der Arbeit, die Problemstellung, die Zielsetzung und die entwickelten Methoden sowie Erkenntnisse bzw. Ergebnisse übersichtlich und verständlich beschreiben.

Abstract

The abstract has to be given in German **and** English. It should be between half a page and one page in length. It should cover in a readable and comprehensive style the context of the thesis, the problem setting, the objectives, and the methods developed in this thesis as well as key insights and results.

Contents

1 Introduction **1**
- 1.1 Motivation . . . 1
- 1.2 Objectives and Contributions . . . 1
- 1.3 Structure of the Thesis . . . 1

2 Background and Related Work **2**
- 2.1 (Example) Question Answering . . . 2
- 2.2 (Example) Conversational AI . . . 2
- 2.3 (Example) Retrieval Augmented Generation . . . 2
- 2.4 Related Work . . . 2

3 Conceptual Framework **3**
- 3.1 Overview . . . 3
- 3.2 (Example) Document Model . . . 3
- 3.3 (Example) Reasoning Path Construction . . . 3
- 3.4 (Example) And Another Section . . . 4

4 Experimental Evaluation **5**
- 4.1 Data . . . 5
- 4.2 Implementation . . . 5
- 4.3 Metrics . . . 5
- 4.4 Evaluation Results . . . 5

5 Conclusions and Future Work **6**

1 Introduction

This chapter provides an overview of the work. The section on motivation, in particular, should be written in a way that is easy to understand for the general reader. Use cases and examples demonstrating the problem to be solved are always helpful. The introduction should also include important references.

1.1 Motivation

What is this thesis about? Give examples of the problem(s) at hand. Illustrations are usually helpful for a better understanding (images say more than 1000 words). Why is the topic important and relevant? In what context and application settings or domains? Give some key references.

1.2 Objectives and Contributions

This section describe the challenges to be addressed in this thesis and the problem statement or scientific hypothesis, and in particular the objectives of the work. These typically include the development of a conceptual model (methods, algorithms etc.) and an experimental evaluation of the model and conceptual framework, respectively.

1.3 Structure of the Thesis

This section is usually quite short and essentially only describes the structure of the rest of the thesis. For example: In Chapter 2, we give an overview of the foundations of the work as well as related work. In Chapter 3, we then present . . . . . etc.

2 Background and Related Work

The first few sections in this chapter introduce the basics for the work. These can be, for example (based on what topics in the Data Science Group are covered), the basic concepts and terminology underlying information extraction, knowledge graphs, question answering, transformer models, and the like. The introductory paragraph of this chapter gives a brief outline of the chapter, again with references to the respective sections. For example, in Section 2.1 we describe the key components of question answering (QA) methods. Conversational AI frameworks, which often build on QA, are covered in Section 2.2.

2.1 (Example) Question Answering

Such sections typically include a lot of references.

2.2 (Example) Conversational AI

Such sections typically include a lot of references.

2.3 (Example) Retrieval Augmented Generation

Such sections typically include a lot of references.

2.4 Related Work

Typically, the last section of this chapter then discusses related works, ideally in some kind of structured way, i.e., not just a laundry list of papers. References are provided in the following format: This was first shown in the work of Spitz and Gertz [Spitz and Gertz, 2016] .... Further details are covered in Newman's book on networks [Newman, 2010] ....

3 Conceptual Framework

This chapter usually builds the main part of the thesis. Before the first section, a brief overview in the form of a few sentences with references to subsequent sections should be given. Of course, one has to find a proper title for this chapter that fits the topic of the thesis.

Example: In the following Section 3.1, an overview of the requirements of the conceptual framework to be developed are detailed. Section 3.2 then introduces the underlying document model. ...

3.1 Overview

About 2-4 pages that describe the requirements, objectives, and methods the the methods, algorithms etc. developed in the following sections. Often, figures of conceptual pipelines or architectures help to see how methods and components interact with each other and what flow of information and data there is. Note that the description in this section should be more ”technical” or ”formal” than in the introduction (Chapter 1, since the reader is now familiar with the basics and related work. Also – at least for a thesis in the Data Science group – this chapter is independent of any implementation but can refer to this; the same holds for specific datasets.

3.2 (Example) Document Model

In this and the following sections, the contributions of the work are first motivated, formally described in a clear manner (often mathematically, i.e., with definitions, etc.), and – where necessary and useful – illustrated with examples. Note again that the descriptions in this chapter are mostly independent of any specific implementation and datasets; these will be detailed in the following chapter. In this section section, one typically formally introduces some base models, e.g., for documents or knowledge graphs.

3.3 (Example) Reasoning Path Construction

This and the following sections detail the solutions at a conceptual (abstract) level. Follow the rule: motivate and put in context, formalize (mathematically), explain and give examples. Explain alternative approaches (scientific discourse) and why you prefer one approach over the other.

3.4 (Example) And Another Section

And more good stuff . . .

4 Experimental Evaluation

The structure of this chapter heavily depends on the topic and its depth of experimental coverage. The chapter typically starts with a section describing the data used for the evaluation, including proper statistics of the data, references to sources etc. Show examples!

As for the previous two chapters, also this chapter starts with a paragraph outlining the structure of the chapter. The following sections show a typical structure of this chapter.

4.1 Data

Give a good overview of the properties of the data (statistics) used for the experiments, where the data is coming from, how it has been processed etc. Show some examples, such as table entries, text fragments, QA-pairs ....

4.2 Implementation

Provide a figure of the system architecture or data processing pipeline you have implemented. State what specific packages and system settings you have used. Do not provide any source code!

4.3 Metrics

What metrics have you used to evaluate the methods detailed in Chapter 3? Why did you pick these metrics?

4.4 Evaluation Results

Start with the objectives of the evaluation. What baseline has been used for the evaluation? Then describe the evaluation results in a detailed and structured way. Emphasize key insights and results. Show cases where the proposed methods work well and cases where they did not. In case you use large language models (LLMs) as part of the evaluation, provide details about the LLMs (source, version, parameter settings).

5 Conclusions and Future Work

Here, the most important results and findings of the work are summarized once again (not simply a repetition of the structure of the previous chapters!), which new concepts, methods, and tools have been developed, which problems can now be solved (more efficiently), and an outlook on further work is given (e.g., what you would do if you had another 6 months to work on the project).

Bibliography

[Newman, 2010] Newman, M. (2010). *Networks: an introduction*. Oxford University Press.

[Spitz and Gertz, 2016] Spitz, A. and Gertz, M. (2016). Terms over LOAD: leveraging named entities for cross-document extraction and summarization of events. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2016, Pisa, Italy, July 17-21, 2016*, pages 503–512.