UNIVERSITÄT BAYREUTH

– P&E Master's Programme –
Chair of Philosophy, Computer Science & Artificial Intelligence

Automating the Modelling of Transformative Artificial Intelligence Risks

—

"An Epistemic Framework for Leveraging Frontier AI Systems to Upscale Conditional Policy Assessments in Bayesian Networks on a Narrow Path towards Existencial Safety"

—

A thesis submitted at the Department of Philosophy

for the degree of *Master of Arts in Philosophy & Economics*

Author:

Valentin Jakob Meyer
Valentin.meyer@uni-bayreuth.de
Matriculation Number: 1828610
Tel.: +49 (1573) 4512494
Pielmühler Straße 15
52066 Lappersdorf

Supervisor:

Dr. Timo Speith

Word Count:
30.000

Source / Identifier:
Document URL

26th of May 2025

Preface

This is a Quarto book.

To learn more about Quarto books visit `https://quarto.org/docs/books`.

Abstract

Outline(s)

Frontmatter

Prefatory Apparatus: Illustrations and Terminology — Quick References

List of Tables

List of Graphics & Figures

List of Abbreviations

esp. especially
f., ff. following
incl. including
p., pp. page(s)
MAD Mutually Assured Destruction

Glossary

Chapter 1

Introduction

10% of Grade:

• introduces and motivates the core question or problem • provides context for discussion (places issue within a larger debate or sphere of relevance) • states precise thesis or position the author will argue for • provides roadmap indicating structure and key content points of the essay

~ 14% of text ~ 4200 words

• introduces and motivates the core question or problem

1.1 Motivation: Problem Statement

1.2 Motivation: Research Question

• provides context for discussion (places issue within a larger debate or sphere of relevance)

1.3 Scope: Aim & Context of the Research

1.4 Significance of the Research: Theory of Change

• states precise thesis or position the author will argue for

1.5 Thesis Statement & Position: (Aim of the Paper)

• provides roadmap indicating structure and key content points of the essay

1.6 Overview: Structure & Approach of the Paper (Roadmap — Theory of Change)

1.7 Table of Contents

Chapter 2

Context

20% of Grade:

• demonstrates understanding of all relevant core concepts • explains why the question/thesis/problem is relevant in student's own words (supported by quotations) • situates it within the debate/course material • reconstructs selected arguments and identifies relevant assumptions • describes additional relevant material that has been consulted and integrates it with the course material as well as the research question/thesis/problem

~ 29% of text ~ 8700 words

1. successively (chunk my chunk) introduce concepts/ideas — and 2. ground each with existing literature

Chapter 3

AMTAIR

20% of Grade:

• provides critical or constructive evaluation of positions introduced • develops strong (plausible) argument in support of author's own position/thesis • argument draws on relevant course material • claim/argument demonstrates understanding of the course materials incl. key arguments and core concepts within the debate • claim/argument is original or insightful, possibly even presents an original contribution to the debate

~ 29% of text ~ 8700 words

Chapter 4

Discussion

10% of Grade:

• discusses a specific objection to student's own argument • provides a convincing reply that bolsters or refines the main argument • relates to or extends beyond materials/arguments covered in class

~ 14% of text ~ 4200 words

Chapter 5

Conclusion

10% of Grade:

• summarizes thesis and line of argument • outlines possible implications • notes outstanding issues / limitations of discussion • points to avenues for further research • overall conclusion is in line with introduction

~ 14% of text ~ 4200 words

References

Appendix A

Appendices

Appendix B

Appendix A

Appendix C

Appendix B

Appendix D

Appendix C

Appendix E

Appendix D

TestText

Appendix F

Affidavit

Appendix G

testtext