Using Large Language Models to Convert Documents to Knowledge Graphs to Check for Completeness and Consistency

by Michael Wacey

B.S. in Computer Science, June 1985, Washington University in St. Louis M.S. in Computer Science, December 1990, Drexel University

A Praxis submitted to

The Faculty of
The School of Engineering and Applied Science
of The George Washington University
in partial satisfaction of the requirements
for the degree of Doctor of Engineering

March 8, 2026

Praxis directed by

Professor M. Elbasheer Professorial Lecturer in Engineering and Applied Science The School of Engineering and Applied Science of The George Washington University certifies that [Your Name] has passed the Final Examination for the degree of Doctor of Engineering as of December 18, 2025. This is the final and approved form of the Praxis.

Using Large Language Models to Convert Documents to Knowledge Graphs to Check for Completeness and Consistency

Michael Wacey

Praxis Research Committee:

Professor M. Elbasheer, Professorial Lecturer in Engineering and Applied Science, Praxis Director

XYZ, Professor of Online Engineering Programs, Chair

XYZ, Professorial Lecturer of Engineering and Applied Science, Committee Member © Copyright 2025 by Michael Wacey All rights reserved

Dedication

Acknowledgments

Abstract

Using Large Language Models to Convert Documents to Knowledge Graphs to Check for Completeness and Consistency

Table of Contents

| Dedication | iv |
|---|---|
| Acknowledgments | v |
| Abstract | vi |
| List of Figures | ix |
| List of Figures | ix |
| List of Tables | x |
| List of Tables | x |
| Glossary of Terms | хi |
| Chapter 1: Introduction 1.1 Background and Research Motivaton 1.1.1 Document Consistency and Completeness 1.1.2 Attention Model 1.1.3 Large Language Models 1.1.4 Knowledge Graphs 1.1.5 Municipal Laws in Pennsylvania 1.1.6 Research Motivation 1.2 Problem Statement 1.3 Thesis Statement 1.4 Research Objectives 1.5 Research Questions 1.6 Research Hypotheses 1.7 Research Scope and Limitations 1.7.1 Research Scope 1.7.2 Research Limitations 1.8 Praxis Organization | 1 1 1 2 2 2 3 4 4 4 5 5 5 6 6 |
| Chapter 2: Literature Review 2.1 Introduction | 8 |
| 2.2 Some Section | 8 |
| 2.2.1 Subsection 1 | 10 |
| 2.3 Another Section | 11 |
| 2.3.1 Subsection | 12 |
| 2.3.2 Subsection | 12 13 |

| Chapter 3: Methodology | 15 |
|---|----|
| 3.1 Introduction | 15 |
| 3.2 Some Subsection | 15 |
| 3.3 Some Subsection | 16 |
| 3.3.1 Some Subsection | 16 |
| 3.3.2 Some Subsection | 17 |
| 3.4 Conclusion | 17 |
| Chapter 4: Results and Analysis | 19 |
| 4.1 Introduction | 19 |
| 4.2 Some Subsection | 19 |
| 4.3 Some Subsection | 20 |
| 4.4 Some Subsection | 21 |
| 4.5 Some Subsection | 23 |
| 4.5.1 Subsection | 24 |
| 4.6 Some Section | 25 |
| 4.6.1 Conclusion | 26 |
| Chapter 5: Discussion and Conclusions | 28 |
| 5.1 Conclusion | 28 |
| 5.2 Contribution to the Body of Knowledge | 28 |
| 5.3 Recommendations for Future Research | 29 |

List of Figures

| 2.1 | Histogram of XYZ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 9 |
|-----|------------------|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|---|
|-----|------------------|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|---|

List of Tables

| 2.1 | My Table About Something | 12 |
|-----|---|----|
| 4.1 | Title of the table every first letter capitalized | 21 |
| 4.2 | Test-2: Transformer vs. AutoTAB Performance Metrics | 24 |

Glossary of Terms

AI Artificial Intelligence

ANN Artificial Neural Network

API Application Programming Interface

GAT Graph Attention Networks

GCNN Graph Convolutional Neural Network

GNN Graph Neural Network

ML Machine Learning

NER Named Entity Recognition

SHAP SHapley Additive exPlanations

SVM Support Vector Machine

TN True Negative

TP True Positive

UBB User-Based Batching

UBS User-Based Sequencing

Chapter 1: Introduction

1.1 Background and Research Motivaton

Ensuring document quality involves verifying completeness, consistency, and correctness (Zowghi & Gervasi, 2003). While evaluating correctness often necessitates access to knowledge external to the document and understanding the document's intent, completeness and consistency can be assessed within the document itself. This research focuses on developing automated methods using Large Language Models (LLMs) to address the latter two aspects. The specific focus is on converting a large document into a knowledge graph that can be used in future research to check the consistency and completeness of a document.

1.1.1 Document Consistency and Completeness

Much effort has been put into ensuring structural consistency of documents, see (Laban et al., 2021) or focused on a specific domain, see (Tröls et al., 2022). In the literature, the term coherence is often used for consistency, (Shen et al., 2021).

1.1.2 Attention Model

In 2017 a team at Google proposed a neural network model based solely on the Attention Mechanism, (Vaswani et al., 2017). This model looks at all words within a prompt to see the impact that they have on each other. By paying attention to everything, it ensures that the model fully understands the meaning of what is being considered. While the scope has

been increased, Attention models must limit the window that they consider due to limited space and the need for rapid performance.

1.1.3 Large Language Models

All LLMs are derived from the Attention Model and have a window of information that they can look at. As documents grow larger, they exceed this window and therefore cannot be processed completely by an LLM. However, checking a document does not require paying attention to the entire document. Instead, the LLM can be used to identify entities of interest throughout the document. Then, instead of paying attention to the entire document, it can focus on the entities that are of a similar type. This combines the best of both attention and focus.

1.1.4 Knowledge Graphs

A knowledge graph is a compact and human-readable way to represent information compared to a neural network as used in LLM. It is a directed acyclic graph with nodes used to represent things of importance, entities, and edges used to represent relationships between nodes. The nodes can have attributes to provide richer detail. For example, a car node could have a color attribute to represent the car's color. One useful way to interpret a knowledge graph is as a group of frames; see (Minsky, 1974).

1.1.5 Municipal Laws in Pennsylvania

There are over 1,200 townships of the second class in Pennsylvania. Each township must create and maintain its own set of laws. These laws cover Police, Fire Companies, Zoning, Subdivision and Land Development among many other items. These documents end up being thousands of pages long.

Each time a new law is added it can add one or more inconsistencies or completeness issues. Lawyers and law compilers focus on checking for this but it is complicated and they often miss things.

It is complicated because the same things are discussed in multiple areas of the document. For example, a set of laws may define the various boundaries for plots of land early in the document. This could include minimum frontage, setback from neighbors, and many other items. Each set of boundaries would be for a specific zoning area. At the end of the document the zoning areas are defined. A simple inconsistency would be defining plot boundaries for a zoning area that is never defined. There are many other such interactions between sections of municipal laws that need to be kept synchronized.

1.1.6 Research Motivation

Despite extensive research on analyzing small documents or specific document sections, there is a significant gap in addressing the challenges of complete, large-scale document analysis. The need for automated consistency and completeness checks is critical in various industries. Currently, these tasks are often performed manually, requiring substantial time and resources while still potentially yielding suboptimal results. This research aims to bridge this gap by developing an effective and efficient automated solution.

Within the scope of this research, local regulations of townships in Pennsylvania go through a time consuming and complicated process to get published. After the governing body enacts a law, it is sent to an organization to compile it into existing laws of the township. This is a manual and intensive process to determine if any of the existing laws are affected by the new law. Even with this, there are many cases of new laws that make a set of existing laws incomplete or inconsistent.

1.2 Problem Statement

Municipal laws in Pennsylvania Townships, authored by multiple people over time, develop inconsistencies and are incomplete (D. Curley, Easttown Supervisor, personal communication, September 16, 2024; A. Rau, Esq., Easttown Solicitor, personal communication, September 20, 2024; J. Sanders, personal communication, October 25, 2024), leading to annual revenue losses of hundreds of thousands of dollars. (M. Wacey, Easttown Supervisor, personal communication, September 23, 2024).

1.3 Thesis Statement

An LLM-based tool to convert a document into an attributed knowledge graph can be used to check for consistency and completeness will allow municipal lawyers to create consistent and complete law documents which prevent costly disputes and reduce revenue losses.

1.4 Research Objectives

The primary objective of this research is to develop a tool capable of automatically processing documents of any size into a coherent set of entities in a knowledge graph. This tool will leverage advanced techniques to analyze document content, identify potential entities, and provide access to the knowledge graph.

The created knowledge graph will be analyzed to determine whether it is appropriate to check the document for inconsistencies and incompleteness.

This will include introducing issues in the source documents and then highlighting how easy they are to observe in the knowledge graph.

1.5 Research Questions

To achieve the research objectives, the following research questions will be considered.

- **RQ1:** Can an LLM be used to convert a large document into a knowledge graph?
- **RQ2:** Can an LLM be used to process multiple knowledge graphs into a typed cluster of knowledge graphs.
- **RQ3:** Can a typed cluster of knowledge graphs be used to check the source document for consistency and completeness?

1.6 Research Hypotheses

Research will be conducted to test the following hypotheses.

- **H1:** An LLM can be used to convert a large document into a knowledge graph.
- **H2:** An LLM can be used to process multiple knowledge graphs into a typed cluster of knowledge graphs.
- **H3:** A typed cluster of knowledge graphs can be used to check the source document for consistency and completeness.

1.7 Research Scope and Limitations

1.7.1 Research Scope

This research uses Pennsylvania township laws as a case study to train and test the developed tool. These laws, available in the public domain in both PDF and Word formats, are chosen due to their extensive length, multiple authors, and rigorous manual checks for consistency and completeness. Although the focus is on these legal documents, the developed approach is intended to be generalizable to various document types.

For this effort, the development will focus on the creation of Knowledge Graphs that represent the document or documents under review. There will be an analysis to ensure that the knowledge graphs created are suitable to check the document for consistency and completeness. However, that development will be left for future work.

1.7.2 Research Limitations

Potential limitations might include computational constraints, challenges in handling specific document formats or language complexities, and the need for further refinement of neural network models. This research will not perform the actual test for consistency and completeness.

1.8 Praxis Organization

The remainder of this research is organized as follows.

- **Chapter 2 Literature Review:** This chapter provides a comprehensive review of the relevant literature on:
 - The creation of knowledge graphs from documents by LLMs.
 - The processing of multiple knowledge graphs into a combined knowledge graph by LLMs.
 - The ability of knowledge graphs to be used to represent the original document in ensuring consistency and completeness.

- The process of creating and maintaining local laws in Pennsylvania.
- Background on checking documents for consistency and completeness.
- **Chapter 3 Methodology:** This chapter details the statistical and machine learning methodologies employed in this research, including data pre-processing, model selection, training, and evaluation.
- **Chapter 4 Results:** This chapter presents and analyzes the results of the data analysis, addressing each research question and hypothesis. It also evaluates the performance of the proposed methodology and tool.
- Chapter 5 Discussion and Conclusion: This chapter concludes the
 research with a discussion of key findings, contributions to the field,
 recommendations for practical applications and potential avenues for
 future research.

Chapter 2: Literature Review

2.1 Introduction

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

2.2 Some Section

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel

leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

This is how you insert an image by referencing it like this: Figure 2.1 shows XYZ.

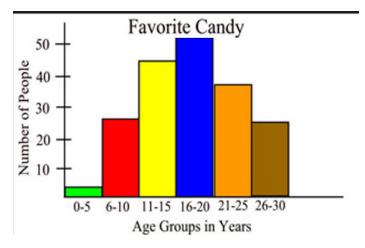


Figure 2.1: Histogram of XYZ

2.2.1 Subsection 1

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

This is how to enter an equation and reference it. Equation 2.1 and 2.1 show how XYZ is implemented.

$$TE_{(pos,2i)} = sin(pos/23^{2i/Lm})$$
 (2.1)

$$KN_{(pos,2i+1)} = cos(pos/453^{2i/Lm})$$
 (2.2)

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer

sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

2.3 Another Section

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

This is the correct format for a table, including the caption and proper alignment (centered and italicized).

Table 2.1: My Table About Something

| Reference | Technique |
|----------------|----------------|
| something here | Technique here |

2.3.1 Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

2.3.2 Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a,

magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

2.3.2.1 Sub subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

2.4 Summary

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida

Chapter 3: Methodology

3.1 Introduction

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

3.2 Some Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel

leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

3.3 Some Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

3.3.1 Some Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer

sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

3.3.2 Some Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

3.4 Conclusion

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna

fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

At the end of Chapter 3, you must restate your research questions and hypotheses exactly as they are. The format should follow the structure shown below. Please Ensure all text coloring is removed.:

RQ1:

RQ2:

RQ3:

H1:

H2:

H3:

Example:

This methodological approach addresses the research questions and hypotheses outlined in Chapter 1 and repeated below:

RQ1: You will repeat your research question1 here?

RQ2: You will repeat your research question2 here?

RQ3: RQ2: You will repeat your research question3 here?

H1: RQ2: You will repeat your research Hyothesis1 here.

H2: You will repeat your research Hyothesis2 here.

H3: You will repeat your research Hyothesis2 here.

Chapter 4: Results and Analysis

4.1 Introduction

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

4.2 Some Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel

leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

4.3 Some Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

This is how you build and format the table. The caption is centered and italized and in the top. The table number is bolded. Please remove the red coloring i use to illutrate this.

Table 4.1 presents an overview of xyz and zkm and xyz xyz. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit,vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris

Table 4.1: Title of the table every first letter capitalized

| Factor1 | Test 1 | Test 2 |
|----------------|--------|--------|
| Something here | 123 | 123 |
| Something here | 123 | 123 |
| Something here | 123 | 1123 |
| Something here | 16 | 123 |
| Something here | 123 | 123 |
| Something here | 123 | 123 |
| | | |

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

4.4 Some Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida

mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

This is the correct way to add and format an equation. Ensure the formatting is consistent, and remove any red coloring used for illustration.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum..

This is the correct way to add and format an equation. Ensure the

formatting is consistent, and remove any red coloring used for illustration.

$$Precision = \frac{TP}{TP + FP} \tag{4.2}$$

4.5 Some Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

This is the correct way to build and format a table. The caption should be centered, italicized, and placed at the top. The table number should be bolded. In this example, the table is long, so ensure the table font matches the font of your document. It is not allowed to significantly reduce the table font size just to make it fit. In this case, I used a slightly smaller font, depending on the situation—only if necessary. However, try to avoid having too many columns to maintain readability. Make sure to remove the red coloring used for illustration.

Table 4.2 depicts the xxxx.

Table 4.2: Test-2: Transformer vs. AutoTAB Performance Metrics

| Model | Method | A | P | R | F1 | AUC | FNR | FPR |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|
| | Sub1 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 |
| Mod1 | Sub2 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 |
| | Sub3 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 |
| | Sun1 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 |
| Mod2 | Sub2 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 |
| | Sub3 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 | 0.0123 |

4.5.0.1 Sub Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

4.5.1 Subsection

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida

mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

4.6 Some Section

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a,

magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

4.6.1 Conclusion

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

At the end of Chapter 3, you must restate your research questions and hypotheses exactly as they are. The format should follow the structure shown below. Please Ensure all text coloring is removed.:

RQ1:

| RQ2: |
|----------|
| RQ3: |
| H1: |
| H2: |
| Н3: |
| Example: |

The results from this Chapter address the research questions and hypotheses outlined in Chapter 1 and repeated below:

RQ1: How do Transformer encoders compare to Autoencoders in terms of accuracy, precision, and recall when detecting malicious insider threats?

RQ1: You will repeat your research question1 here?

RQ2: You will repeat your research question2 here?

RQ3: RQ2: You will repeat your research question3 here?

H1: RQ2: You will repeat your research Hyothesis1 here.

H2: You will repeat your research Hyothesis2 here.

H3: You will repeat your research Hyothesis2 here.

Chapter 5: Discussion and Conclusions

5.1 Conclusion

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

5.2 Contribution to the Body of Knowledge

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel

leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

5.3 Recommendations for Future Research

References

- Laban, P., Dai, L., Bandarkar, L., & Hearst, M. A. (2021). Can transformer models measure coherence in text? re-thinking the shuffle test. *arXiv* (Cornell University). https://doi.org/10.48550/arxiv.2107.03448
- Minsky, M. (1974, June). A framework for representing knowledge. http://hdl.handle.net/1721.1/6089
- Shen, A., Mistica, M., Salehi, B., Li, H., Baldwin, T., & Qi, J. (2021). Evaluating document coherence modeling. *Transactions of the Association for Computational Linguistics*, *9*, 621–640. https://doi.org/10.1162/tacl_a_00388
- Tröls, M. A., Marchezan, L., Mashkoor, A., & Egyed, A. (2022). Instant and global consistency checking during collaborative engineering. Software and systems modeling, 21(6), 2489–2515. https://doi.org/10.1007/s10270-022-00984-4
- Vaswani, A., Shazeer, N., Brain, G., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & Kaiser, Ł. (2017). Attention is all you need. Advances in Neural Information Processing Systems.
- Zowghi, D., & Gervasi, V. (2003). On the interplay between consistency, completeness, and correctness in requirements evolution. *Information and Software Technology*, 45(14), 993–1009. https://doi.org/10.1016/S0950-5849(03)00100-9