

**INTEGRATION OF SMALL LANGUAGE MODELS (SLMs)
FOR ACCOUNTING AUTOMATION IN AN ERP**

By

Theodora Andrea Lydia

12102013

BACHELOR'S DEGREE

in

INFORMATION TECHNOLOGY

FACULTY OF ENGINEERING AND INFORMATION TECHNOLOGY



SWISS GERMAN UNIVERSITY

The Prominence Tower

Jalan Jalur Sutera Barat No. 15, Alam Sutera

Tangerang, Banten 15143 - Indonesia

Revision after the Thesis Defense on 8 July 2025

STATEMENT BY THE AUTHOR

I hereby declare that this submission is my own work and to the best of my knowledge, it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at any educational institution, except where due acknowledgement is made in the thesis.



Theodora Andrea Lydia
Student

22 July 2025
Date

Approved by:



Dr. Ir. Heru Purnomo Ipung, M.Eng.
Thesis Advisor

22 July 2025
Date



Leonardo Bahar, S.TI., B.I.T.
Thesis Co-Advisor

22 July 2025
Date

Dr. Maulahikmah Galinium, S.Kom, M.Sc
Dean

22 July 2025
Date

ABSTRACT

INTEGRATION OF SMALL LANGUAGE MODELS FOR ACCOUNTING AUTOMATION IN AN ERP

By

Theodora Andrea Lydia

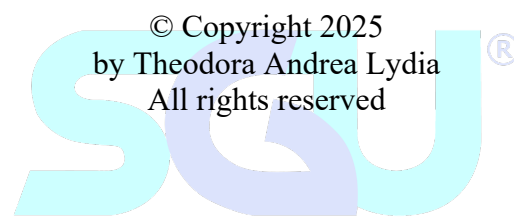
Dr. Ir. Heru Purnomo Ipung, M.Eng, Advisor

Leonardo Bahar, S.TI., B.I.T., Co-Advisor

SWISS GERMAN UNIVERSITY

This thesis explores the application of Small Language Models (SLMs) for accounting automation, namely in invoice data extraction and account reconciliation. The purpose of this thesis is to evaluate the feasibility of using SLMs for these tasks and to implement a prototype tool that is integrated into an ERP system. This thesis follows the design science research methodology to develop and evaluate the SLM-based tools. This thesis compares several different SLMs with different prompts to find the most reliable and accurate model and prompt. Testing is conducted with synthesized exploratory data. Results show that certain SLMs are able to achieve these functionalities with decent accuracy and inference time. This suggests that SLMs are a viable and efficient solution for accounting automation in ERP systems.

Keywords: Small Language Models, Account Reconciliation, Invoice Data Extraction, Low-Resource AI Applications, Financial Automation



ACKNOWLEDGMENTS

I would like to express my gratitude to my advisors in Indonesia and Germany for their direction and guidance, my internship company, Codedoc, for supporting my thesis, other lecturers for their input and support, and the friends and family who have been cheering me on in this journey.



TABLE OF CONTENTS

	Page
STATEMENT BY THE AUTHOR.....	2
ABSTRACT.....	3
ACKNOWLEDGMENTS.....	5
TABLE OF CONTENTS.....	6
LIST OF FIGURES.....	9
LIST OF TABLES.....	10
CHAPTER 1 – INTRODUCTION.....	11
1.1 Background.....	11
1.2 Research Problems.....	13
1.3 Research Objectives.....	13
1.4 Significance of Study.....	13
1.5 Research Questions.....	13
1.6 Hypothesis.....	13
1.7 Research Scope.....	13
1.8 Research Limitations.....	14
CHAPTER 2 – LITERATURE REVIEW.....	15
2.1 Theoretical Perspective.....	15
2.1.1 Account Reconciliation.....	15
2.1.2 Large Language Models (LLMs) and Small Language Models (SLMs).....	15
2.1.3 Vision Language Models.....	16
2.1.4 Prompt Engineering.....	17
2.1.5 RAG.....	17
2.1.6 ERP.....	18
2.2 Related Works	
2.2.1 Chain-of-Thought Prompting Elicits Reasoning in Large Language Models (Wei et al., 2022).....	18
2.2.2 Entity Matching using Large Language Models (Peeters et al., 2024).....	18
2.2.3 Leveraging Large Language Models for Entity Matching (Huang et al., 2024).....	19

2.2.4	Optimizing Accounting Processes Through the Integration for AI in Data Entry (Dr. Frederick et al., 2024).....	19
2.2.5	Automation of Accounting Process: Impact of Artificial Intelligence (Dr. Chukwuani & Dr. Egiyi, 2020).....	19
CHAPTER 3 – RESEARCH METHODOLOGY.....		20
3.1.	Invoice Data Extraction Tool Development.....	21
3.1.1	Model Selection.....	21
3.1.2	Prompt Engineering.....	21
3.1.3	Preprocessing.....	21
3.1.4	Postprocessing.....	21
3.2	Account Reconciliation Tool Development.....	22
3.2.1	Preliminary Testing, Model Selection, and Prompt Engineering.....	22
3.2.2	Tool Mechanism.....	22
3.2.3	Preprocessing.....	23
3.2.4	Postprocessing.....	23
3.3	Software Development.....	23
3.3.1	System Architecture.....	23
3.3.2	Deployment Environment.....	24
3.3.3	Security Considerations.....	24
3.3.4	Unit Testing.....	24
3.3.5	System Integration Testing.....	24
3.4	Evaluation.....	24
3.4.1	Dataset.....	24
3.4.2	Evaluation Metrics.....	26
CHAPTER 4 – RESULTS AND DISCUSSIONS.....		28
4.1	Invoice Data Extraction.....	28
4.1.1	Model Comparison.....	28
4.1.2	Prompt Design and Parameter Selection.....	30
4.1.3	Quantitative Testing.....	31
4.2	Account Reconciliation.....	32
4.2.1	Preliminary Testing.....	32

4.2.2	Model Selection.....	33
4.2.3	Prompt Design.....	33
4.2.4	Quantitative Testing.....	33
4.3	System Integration.....	35
CHAPTER 5 – CONCLUSION AND RECOMMENDATIONS.....		38
REFERENCES.....		40
APPENDICES.....		43
CURRICULUM VITAE.....		48



LIST OF FIGURES

Figures

1. Research overview.....	20
2. Invoice data extraction flowchart.....	20
3. Pseudocode of account reconciliation RAG.....	22
4. Account reconciliation flowchart using SLM and RAG.....	23
5. Example of simple invoice.....	25
6. Example of medium invoice.....	25
7. Example of hard invoice.....	26
8. Confusion matrix of 100 transactions.....	34
9. System Architecture Diagram.....	35
10. Account reconciliation feature.....	36
11. Account reconciliation result.....	36
12. Account reconciliation result.....	37
13. Invoice extraction tool.....	37



LIST OF TABLES

Table

1. Model performance comparison for invoice extraction.....	28
2. Selected model performance for 100 invoices.....	31
3. SLM testing result.....	34
4. SLM + RAG testing result.....	35

