

TABLE OF CONTENTS

DECLARATION.....	2
TABLE OF CONTENTS.....	3
LIST OF TABLES.....	5
LIST OF FIGURES.....	6
<i>ABSTRACT</i>	8
CHAPTER 1 INTRODUCTION.....	9
1.1 Research Background.....	9
1.2 Research Problem.....	10
1.3 Research Scope.....	10
1.4 Research Objective.....	10
1.5 Research Advantage.....	11
CHAPTER 2 LITERATURE REVIEW.....	12
CHAPTER 3 FUNDAMENTAL THEORY.....	18
3.1 Money Laundering.....	18
3.2 Data Model.....	20
3.2.1 Relational Data Model.....	21
3.2.2 Graph Data Model.....	22
3.3 Multi-Model Data Management.....	23
CHAPTER 4 RESEARCH METHODOLOGY.....	25
4.1 Research Overview.....	25
4.2 Research Dataset and Schema.....	26
4.2.1 Relational Model Schema.....	27
4.2.2 Graph Model Schema.....	28
4.3 Data Architecture.....	29
4.4 Research Evaluation and Analysis.....	34
CHAPTER 5 IMPLEMENTATION.....	32
5.1 Implementation Tools.....	32
5.3 Cloud Initialization.....	39
5.3.1 Environment Configuration.....	40
5.3.4 Graph Spanner Schema Building.....	44
5.4 Account Workflow Construction.....	44
5.5 Transaction Workflow Construction.....	47
5.6 Evaluation Testing.....	58
5.6.1 Graph Model Evaluation.....	58
CHAPTER 6 RESULT AND DISCUSSION.....	63

6.1 Result.....	63
6.2 Evaluation on Query Depth.....	65
6.3 Evaluation on Data Volume.....	66
6.4 Discussion.....	67
CHAPTER 7 CONCLUSION AND RECOMMENDATION.....	70
7.1 Conclusion.....	70
7.2 Suggestion.....	71
REFERENCES.....	72
APPENDIX.....	76

LIST OF TABLES

Table 2.1. Summary of previous literature.....	16
Table 3.1. Relational model vs. graph model.....	21
Table 4.1. Dataset summary.....	27
Table 4.2. Dataframe library performance (Mozzillo et al., 2023).....	30
Table 4.3. Transaction incremental schema.....	32
Table 4.4. Query benchmarks.....	34
Table 4.5. Benchmark scenario.....	35
Table 5.1. Implementation device.....	32

LIST OF FIGURES

Figure 3.1. Money laundering stages.....	19
Figure 3.2. AML typologies (Altman et al., 2023).....	20
Figure 3.3. Relational data model (Brown, 2024).....	21
Figure 3.4. Graph data model (Pugsley, 2023).....	22
Figure 3.5. Comparison of polyglot persistence and multi-model database (Schmitt, 2022).....	24
Figure 4.1. Flowchart of the research framework.....	25
Figure 4.2. Dataset sample.....	27
Figure 4.3. Relational model diagram.....	28
Figure 4.4. Graph model diagram.....	29
Figure 4.5. Data architecture.....	29
Figure 4.6. Acquisition and pre-processing flow.....	31
Figure 4.7. Data orchestration schema.....	32
Figure 4.8. Data quality checking schema.....	34
Figure 4.9. Task execution scenario.....	36
Figure 5.1. Environment initialization.....	34
Figure 5.2. Raw sample of account data.....	35
Figure 5.3. Account construction.....	35
Figure 5.4. Raw sample of transaction and pattern data.....	36
Figure 5.5. Initial construction of transaction.....	36
Figure 5.6. Transaction pattern extraction.....	37
Figure 5.7. Transaction column transformation.....	38
Figure 5.8. Final construction of transaction.....	39
Figure 5.9. Pattern construction.....	39
Figure 5.10. Service enablement.....	40
Figure 5.11. Cloud storage set up.....	40
Figure 5.12. BigQuery schema construction.....	42
Figure 5.13. Spanner schema construction.....	44
Figure 5.14. Graph schema construction.....	44
Figure 5.15. Load account to BQ.....	45
Figure 5.16. Account JSON schema.....	46
Figure 5.17. Load account to spanner.....	47
Figure 5.18. Raw ingestion initialization.....	48
Figure 5.19. BQ Raw ingestion job.....	48

Figure 5.20. BQ delta ingestion job.....	49
Figure 5.21. Delta ingestion procedure.....	52
Figure 5.22. Full scale ingestion procedure.....	53
Figure 5.23. Full scale ingestion job.....	54
Figure 5.24. Data replication initialization.....	55
Figure 5.25. Transaction and pattern JSON schema.....	56
Figure 5.26. Load to spanner function.....	57
Figure 5.27. Data replication to spanner.....	57
Figure 5.28. Graph model queries.....	59
Figure 5.29. Relational model queries.....	62
Figure 6.1. Elapsed time distribution.....	64
Figure 6.2. CPU time distribution.....	64
Figure 6.3. Standard deviation distribution.....	65