Frankfurt University of Applied Sciences

Faculty of Computer Science and Engineering

Conversational Copilot for Enterprise Architects

A Retrieval-Augmented Approach

Thesis to Obtain the Academic Degree

Master of Science (M.Sc.)

Submitted by

Hendrik Gruber

Matriculation Number: 1458240

Advisor : Prof. Dr. Jürgen Jung

Co-Advisor : Dr. Rainer Schlör

DECLARATION

I hereby assure that I wrote the present work independently and that I did not use any other sources than those given in the bibliography.

All passages that are taken verbatim or correspondingly from published or not yet published sources are marked as such.

The drawings or images in this work were created by myself or provided with a corresponding source reference.

This work has not been submitted to any other examination authority in the same or a similar form.

Frankfurt, 28. February, 2026

Hendrik Gruber

ABSTRACT

Lorem ipsum ...

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AI	Artificial Intelligence
EA	Enterprise Architect / Architecture
EAM	Enterprise Architecture Management
GenAI	Generative Artificial Intelligence
RAG	Retrieval-Augmented Generation

USAGE OF GENERATIVE AI

OpenAI GPT-5 (OpenAI, 2025) was used in order to find sources and summarize them.

Part I

THESIS

1

INTRODUCTION

1.1 MOTIVATION AND THESIS QUESTION

In the field of

2 BACKGROUND INFORMATION AND TERMINOLOGY

This chapter introduces ...

2.1 TERMINOLOGY

Go over the basics of what EAM, Chatbots, etc. are

3 CURRENT STATE OF THE ART

Briefly explain why a literature analysis is important. Define the scope (what fields you looked at, which databases, what keywords). Define the research method and how you narrowed it down from x sources to y sources.i

3.1 ENTERPRISE ARCHITECTURE MANAGEMENT

theories, digital twin efforts, EA tool landscapeStandards or frameworks (e.g., TOGAF, ArchiMate, IATA ONE Record, LeanIX). Theoretical foundations (auch auf prozessmanagment eingehen, wie der aktuelle Prozess aussieht, wenn die Landschaft geändert werden soll) Current tools and methods Research prototypes in EA

Authors Jung and Fraunholz 2021 [5] lay foundational work from which many EAM concepts can be derived.

3.2 LARGE LANGUAGE MODELS, CONVERSATIONAL AGENTS, AND RETRIEVAL-AUGEMENTED GENERATION

strengths, hallucination issues, graph-RAG enhancements Theoretical foundations Current tools and methods

This paper covers how ai tools are more scalable than manual expertise analysis of things. The source is highly relevant. Look at the summary in notebookLM. 05.10.25 [4]

This 2025 paper has ideas on how changing knowledge-graphs (e.g. through updates) can be handled [6]. It looks at temporal data and how to handle it. This might be relevant since addressing how a changing application landscape can be handled will probably be a challenge.

This paper gives an overview on how to control the dialog sequence and also notes 4 types of dialog options for chatbots in the related works section: [7].

This paper [10] covers how a chatbot can support in task-planning and output generation. Might be helpful in understanding how my chatbot can tell the EA how to conduct changes in the application landscape.

This paper [1] states how proactive dialogue systems work and can be improved. It goes into detail on 3 types of dialogue systems: clarification, target-guided, and non-collaborative dialogues. All 3 of these have a certain relevance for the EAM Chatbot.

3.3 COMPARABLE PROJECTS AND PROTOTYPES

Proof-of-concepts, research prototypes, industry whitepapers, GitHub projects.

Tools like ChatEA, LeanIX AI features, or Microsoft Copilot integrations in architecture/governance.

A prototypical graph-based RAG approach for text-summarization has been created by Microsoft: y[2]. The accompanying paper is here: [3]

3.4 EVALUATIONS AND LIMITATIONS

Studies analyzing strengths/weaknesses of RAG, embedding quality, hallucination mitigation.

Papers about user interaction with EA tools, chatbot evaluation frameworks, usability challenges

This paper [9] gives a standardized method and framework for evaluating coversational AI agents.

This paper [8] proposes a benchmark for open-ended multi-turn conversational agents. I think this paper focuses more on evaluating agents and comparing their results, but maybe i can copy their evaluation methods and benchmarks?

4

METHODOLOGY

5

EXPERIMENTS

6

EVALUATION OF THE RESULTS AND CHALLENGES

7

CONCLUSION AND FUTURE WORK

Todo

Part II

APPENDIX

BIBLIOGRAPHY

[1] Yang Deng, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, and Tat-Seng Chua. *Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration*. 2023. arXiv: 2305.13626 [cs.CL]. URL: https://arxiv.org/abs/2305.13626.

[2] Darren Edge, Ha Trinh, Newman Cheng, Joshua Bradley, Alex Chao, Apurva Mody, Steven Truitt, Dasha Metropolitansky, Robert Osazuwa Ness, and Jonathan Larson. "From Local to Global: A Graph RAG Approach to Query-Focused Summarization." 2024. URL: https://www.microsoft.com/en-us/research/publication/from-local-to-global-a-graph-rag-approach-to-query-focused-summarization/.

[3] Darren Edge, Ha Trinh, Newman Cheng, Joshua Bradley, Alex Chao, Apurva Mody, Steven Truitt, Dasha Metropolitansky, Robert Osazuwa Ness, and Jonathan Larson. *From Local to Global: A Graph RAG Approach to Query-Focused Summarization*. 2025. arXiv: 2404.16130 [cs.CL]. URL: https://arxiv.org/abs/2404.16130.

[4] Jiawei Gu, Xuhui Jiang, Zhichao Shi, Hexiang Tan, Xuehao Zhai, Chengjin Xu, Wei Li, Yinghan Shen, Shengjie Ma, Honghao Liu, et al. "A survey on llm-as-a-judge." In: *arXiv preprint arXiv:2411.15594* (2024).

[5] Jürgen Jung and Bardo Fraunholz. *Masterclass Enterprise Architecture Management*. Springer, 2021.

[6] Dong Li, Yichen Niu, Ying Ai, Xiang Zou, Biqing Qi, and Jianxing Liu. *T-GRAG: A Dynamic GraphRAG Framework for Resolving Temporal Conflicts and Redundancy in Knowledge Retrieval*. 2025. arXiv: 2508.01680 [cs.AI]. URL: https://arxiv.org/abs/2508.01680.

[7] Zhigen Li et al. *ChatSOP: An SOP-Guided MCTS Planning Framework for Controllable LLM Dialogue Agents*. 2025. arXiv: 2407.03884 [cs.CL]. URL: https://arxiv.org/abs/2407.03884.

[8] Xiao Liu, Hao Yu, Hanchen Zhang, Yifan Xu, Xuanyu Lei, Hanyu Lai, Yu Gu, Hangliang Ding, Kaiwen Men, Kejuan Yang, et al. "Agentbench: Evaluating llms as agents." In: *arXiv preprint arXiv:2308.03688* (2023).

[9] Anna Wolters, Arnold Arz von Straussenburg, and Dennis M. Riehle. "Evaluation Framework for Large Language Model-based Conversational Agents." In: July 2024.

[10] Wenshuo Zhai, Jinzhi Liao, Ziyang Chen, Bolun Su, and Xiang Zhao. "A Survey of Task Planning with Large Language Models." In: *Intelligent Computing* (2025). DOI: 10.34133/icomputing.0124.