

DIEGO COLLARANA

Developing Neuro-Symbolic AI systems that work!

@ diego.collarana.vargas@fit.fraunhofer.de
Germany

www.fit.fraunhofer.de/kellm

+49-17635347746

diego-collarana

Zöllnerstrasse 7, 01307 Dresden

0000-0002-2583-0778



EXPERIENCE

Head of Knowledge-Enhanced Foundational Models

Fraunhofer FIT — 7 Team members

Jul. 2023 – Ongoing

Sankt Agustin, DE

- Develop interfaces between KGs and LLMs, e.g., GraphRAG
- Strategies for acquiring AI research and industry projects
- Design neuro-symbolic AI agents for data analytics
- Develop Graph Foundation Models for enterprises

Team Lead Knowledge Graphs for Conversational AI

Fraunhofer IAIS — 5 team members

Oct. 2019 – Jun. 2023

Dresden, DE

- Data preparation and quality for pretraining LLMs
- KGs development for conversational apps
- Develop advance KG question answering systems

Senior Research Engineer and Project Leader

Fraunhofer IAIS

Apr. 2016 – Sep. 2019

Sankt Agustin, DE

- Ideate new applications for KG Embeddings and GNNs
- Implementing KGs for Industry 4.0

Innovation Department Chief

MedSpazio Srl — 5 team members

Mar. 2014 – Feb. 2015

Cochabamba, BO

- Designing innovative mobile healthcare apps with KGs

Software Engineer and Research Assistant

DFKi and Ontology Engineering Group

Feb. 2011 – Dec. 2013

Kaiserslautern (DE) and Madrid (ES)

- Implement REST APIs for accessing KGs for connecting a VR app

PROJECTS

Jupiter AI Factory – #JAIF

European High-Performance Computing Joint Undertaking

2025 – 2028

#JAIF is a central pillar of Europe's AI strategy. Through #JAIF, we aim to support businesses training next-generation AI foundational models using Europe's first exascale-class supercomputer. Our goal

EDUCATION

Ph.D. in Artificial Intelligence

University of Bonn

Sep. 2015 – Feb. 2019

Thesis: Strategies and Techniques for Federated Semantic Knowledge Retrieval and Integration (**Magna Cum Laude, 1.0**)

M.Sc. in Software Engineering

Technische Universität Kaiserslautern (DE) and Universidad Politecnica de Madrid (ES)

Sep. 2011 – Dec. 2013

Thesis: Requirements Engineering and Design of Gamified Mobile Apps (**Score 1.3**)

B.Sc. in Computer Systems

Universidad Catolica Boliviana

Feb. 1999 – Feb. 2004

TOP PUBLICATIONS

Books and Standards

- G. Rehm, D. Collarana, L. Hennig, et al., *Knowledge Graphs for Language Models and Language Models for Knowledge Graphs – Hybrid Applications of Symbolic and Sub-symbolic AI* (DIN SPEC 91526), 2025-05. Berlin: DIN Media GmbH, 2025, p. 66. DOI: 10.31030/3610747.
- D. Collarana, *Strategies and Techniques for Federated Semantic Knowledge Retrieval and Integration* (Studies on the Semantic Web). Amsterdam, Netherlands: IOS Press, 2019, vol. 44, ISBN: 978-1-64368-046-0. DOI: 10.3233/SSW200002-MONO.

Journal Articles

- A. Rivas, D. Collarana, M. Torrente, and M. Vidal, "A neuro-symbolic system over knowledge graphs for link prediction," *Semantic Web*, vol. 15, no. 4, pp. 1307–1331, 2024.

Conference Proceedings

is to assist German start-ups and SMEs in developing robust, secure, and privacy-compliant AI applications. We leverage our expertise in managing complex knowledge graphs, aided by GenAI.

Next Generation of Dataspaces Initiative – #NGDI
Fraunhofer Program for Int. Cooperation and Networking

📅 2024 – 2027

This lighthouse project is developing innovative and interoperable data spaces for the European data economy. Specifically, we explore the roles of Knowledge Graphs, Large Language Models, and Agentic AI in automating interoperability among heterogeneous dataspace by leveraging the FAIR principles.

OpenGPT-X
Bundesministerium für Wirtschaft und Klimaschutz

📅 2022 – 2025

We developed Teuken, the first large-scale multilingual LLM specifically designed to perform competitively in various EU languages. This project advanced AI language models in Germany, promoting technological sovereignty. Additionally, we investigated the integration of LLMs architectures with KGs, resulting in models that can produce factually accurate text.

The SPEAKER Platform
Federal Ministry for Economic Affairs and Energy

📅 2020 – 2023

Knowledge is essential for conversational assistants. As the leader of the SPEAKER data connector component, we develop strategies for continuous collection and enhancement of language data, enabling high-quality, unbiased annotated data for training NLU modules, even in data-scarce scenarios. Our Smart Data Connector (SDC) transforms diverse enterprise data into actionable insights, enabling conversational assistants to comprehend various business domains and deliver accurate responses.

MOST PROUD OF

- 🏆 Acquired Fundings worth +10 Mio Euro (2020-2024)
Co-PI in the acquisition of #JAIF, #WestAI, #NGDI and more
- 🏆 Best Publication Award (2023)
L3S Research Center, Germany
- 🏆 Google AI Faculty Award (2020)
For supporting excellent research
- 🏆 Best Paper Award (2017)
17th Conference of Web Engineering

STRENGTHS

Hard-working

Eye for detail

Motivator & Leader

Customer Engagement

Funding Acquisition

RDF/OWL Stack

KG Construction

SPARQL

KGQA

GNN

Python

Fine-Tuning

Agentic AI

Hugging Face

LangChain

- Y. Peng, D. Collarana, E. Jin, and S. Decker, “Thgfm: Dual-branch temporal heterogeneous graph foundation model,” in *International Semantic Web Conference*, Springer, 2026.
- Y. Mou, L. Liu, S. K. Sowe, D. Collarana, and S. Decker, “Leveraging llms few-shot learning to improve instruction-driven knowledge graph construction,” in *International Conference on Very Large Data Bases, VLDB 2024, Workshops, VLDB.org*, 2024.
- J. Soruco, D. Collarana, A. Both, and R. Usbeck, “QALD-9-ES: A spanish dataset for question answering systems,” in *International Conference on Semantic Systems Semantics*, IOS Press, 2023.
- A. Sadeghi, D. Collarana, D. Graux, and J. Lehmann, “Embedding knowledge graphs attentive to positional and centrality qualities,” in *European Conference ECML PKDD. Research Track*, Springer, 2021.
- D. Collarana, M. Galkin, C. Lange, S. Scerri, S. Auer, and M. Vidal, “Synthesizing knowledge graphs from web sources with the MINTE framework,” in *International Semantic Web Conference*, Springer, 2018.
- D. Collarana, T. Heuss, J. Lehmann, et al., “A question answering system on regulatory documents,” in *Legal Knowledge and Information Systems - JURIX*, IOS Press, 2018.

LANGUAGES

English	●●●●●
Spanish	●●●●●
German	●●●●●
French	●●●●●

REFEREES

Prof. Dr. Sören Auer
@ Director of Leibniz Information Centre for Science and Technology
✉ auer@tib.eu

Prof. Dr. Maria-Esther Vidal
@ Head of Scientific Data Management Research Group
✉ maria.vidal@tib.eu

Prof. Dr. Jens Lehmann
@ Principal Scientist at Amazon
✉ jlehmann@amazon.com