

WORK EXPERIENCE

Pre-Doctoral Researcher – Computer Vision

IBM Research

May 2025 – Present Zurich

- Conducting research on multi-modal machine learning for document understanding within the Docling Team.
- Focusing on Vision-Language Models (VLMs) for document parsing and chemical structure extraction.
- Supervised by Dr. Peter W. J. Staar in the AI for Knowledge group.

Master Thesis – NLP

IBM Research

Oct 2024 – Apr 2025 Zurich

- End-to-End Evaluation of Chunking Methods for Retrieval Augmented Generation Systems.

Working Student – NLP

Reeeliance

May 2024 – Sep 2024 Remote (Hamburg)

- R&D in NLP and Data Integration. (Focus: Agentic RAG Systems)

Working Student – Robotics

DLR (German Aerospace Center)

Dec 2021 – Sep 2022 Weßling

- Construction and supervision of robotic systems and testing.
- Department of Mechatronic Systems.

RESEARCH & PUBLICATIONS

CVPR 2026 Main Paper – Computer Vision

First Author

June 2026

- MarkushGrapher2: End-to-end Multimodal Recognition of Chemical Structures (to be published at CVPR 2026). (Paper | Code | Models | Datasets | Website)
- Introduces an end-to-end multimodal model for recognizing complex chemical structures from document images (including text). The model is able to extract standard molecular structures as well as multimodal markush structures, containing text descriptions, representing entire families of molecules.
- Application: Advancing large scale patent analysis and drug discovery.

Seminar Project – NLP

TU Munich (SEBIS Chair)

Apr 2023 – Jul 2023

- Investigated Graph Representations for NLP: researching ways to augment LLMs using domain knowledge stored in Knowledge Graphs.

Research Project – Computer Vision

TU Munich

Apr 2023 – Jul 2023

- Developed a pipeline for 6D Pose Estimation for Unseen Categories: trained a deep neural network to estimate globally-informed correspondences between partial scene and model point clouds based on functional mapping, enabling 6D pose estimation for unseen object categories.

Bachelor Thesis – Computer Vision

TU Munich

May 2021 – Nov 2021

- Object Detection and Pose Estimation in Control Cabinet Construction using CAD Matching. Conducted as part of the RoMaFo Project – Robotic System and Machine Vision for the assembly of deformable objects for customized production.

ABOUT ME

I am a Ph.D. student at ETH Zürich in the Computer Vision Lab led by Prof. Dr. Ender Konukoglu, and a Pre-Doctoral Researcher at IBM Research – Zurich with Dr. Peter W. J. Staar in the AI for Knowledge group.

My research focuses on multi-modal machine learning for document understanding. I work within the Docling Team, focusing on Vision-Language Models (VLMs) for document parsing and chemical structure extraction.

I hold a M.Sc. in Robotics, Cognition, Intelligence from TU Munich. During My Masters, I gathered experience in (3D) Computer Vision, NLP and Robotics. My most recent publication is MarkushGrapher-2: End-to-end Multimodal Recognition of Chemical Structures (to be published at CVPR 2026).

EDUCATION

PhD, Computer Vision Lab

ETH Zurich

May 2025 – Present Zurich

Computer Vision Lab, Prof. Dr. Ender Konukoglu

Exchange Semester – NLP

Keio University

Sep 2023 – Mar 2024 Tokyo

M.Sc., Robotics, Cognition, Intelligence

TU Munich

Apr 2022 – Apr 2025 Munich

Exchange Semester – Computer Vision

National Taiwan University

Feb 2020 – Jul 2020 Taipei

B.Sc., Engineering Science

TU Munich

Oct 2017 – Nov 2021 Munich

LANGUAGES

- English (Fluent) German (Fluent)
- Spanish (Advanced)

SKILLS

- Computer Vision NLP Deep Learning VLMs
- Knowledge Graphs Python PyTorch
- Document Understanding
- Chemical Structure Recognition
- 3D Computer Vision Robotics RAG Systems