Robin Kokot

Leuven, Belgium ✉ robin.edu.hr@gmail.com 📞 +32 47 808 70 43 ⭘ rokokot

About me

I am a graduate student in Artificial Intelligence at KU Leuven with expertise in multilingual NLP, language model interpretability, and formal methods. My research focuses on probing large language models for linguistic generalization, developing neuro-symbolic methods for low resource machine translation, and implementing advanced computational techniques for scalable linguistic reasoning systems.

Education

Sep 2024 – Jun 2025
KU Leuven, Faculty of Engineering Science
M.Sc. in Artificial Intelligence (Speech & Language Technology)

- Thesis: *Formal Probing of Multilingual Models for Interrogative Semantics*
- **Coursework:** Advanced NLP, Machine Translation, Bayesian Machine Learning

Sep 2022 – Sep 2024
KU Leuven, Institute of Philosophy
M.A. in Philosophy (Cum Laude)

- Focus: Formal logic, epistemic justification applied to model uncertainty quantification

Research Projects

Jan 2024 – Present **Linguistic Probing of Glot500**
KU Leuven NLP Lab

- Designed a diagnostic evaluation framework to assess linguistic generalization in multilingual LLMs
- Developed a dataset for question-type classification accuracy (polar vs. wh-questions) and linguistic complexity metric scoring
- Developed and implemented cross-lingual transfer studies
- **Tools Used:** PyTorch, Hugging Face Transformers, XLM-RoBERTa

Sep 2023 – Present **AAC Pictogram-to-Text Translation**
Independent Research

- Modeled pictogram sequences as a constrained seq2seq task using linear logic programming to improve compositional generalization
- Designed and implemented custom decoding algorithms, evaluated algorithms across traditional MT metrics
- **Tools Used:** OpenNMT(eole), Beam Search with Semantic Constraint Satisfaction algorithm, multimodal encodings, vision transformers

Technical Skills

Programming: Python (PyTorch, TensorFlow, Keras), SQL, Bash, LaTeX

ML/NLP: Hugging Face Transformers, scikit-learn, pandas, seaborn, NLTK, spaCy, AllenNLP

Specialized: Fine-Tuning LLMs, RAG pipelines, Model Quantization, Linguistic Probing, Semantic Schema Induction

HPC: Linux for HPC, PBS scripting, Bash Automation for Cluster Environments
Data Science and Information Visualization: Pandas, Numpy, Matplotlib, Seaborn, Python DB, SQLalchemy

Formal Methods: Bayesian Evidence Theory, Linear Logic Programming, Inquisitive Semantics

Languages

Native: Croatian — **Fluent:** English — **Basic:** German, French