

# Antonios Minas KRASAKIS

BIRTH DATE: 27 February 1991  
PUBLICATIONS: [Google Scholar](#)



[krasakis.com](#)  
[in/krasakis](#)



[github.com/littlewine](#)  
[amkrasakis@gmail.com](mailto:amkrasakis@gmail.com)

## PROFESSIONAL EXPERIENCE

MAY 2025 - Nov 2025	<b>Microsoft Research, Research Scientist Intern</b> Conducted research on Small Language Models (SLMs) for low-resource languages and dialects. Developed synthetic data generation pipelines, scalable methods for building parallel corpora of low-resource dialects, and dialect identification models.
SEPT 2022 - MAR 2023	<b>Amazon, Applied Scientist Intern</b> Worked on Amazon product search with a focus on multilingual retrieval, experimentally evaluating Learned Sparse Retrieval (LSR) within the product search pipeline and scoping multilingual extensions for new and emerging markets.
DEC 2018 - AUG 2019	<b>BOL.com, Data Scientist</b> Led the design and implementation of a content-based recommendation system for related products, including system architecture, data collection and annotation, and online A/B testing.
APR 2018 - JULY 2018	<b>Elsevier, Data Scientist Intern</b> Developed a meta-learning framework for weakly supervised biomedical information extraction from scientific articles, leading to a research publication.
OCT 2014 - JUL 2015	<b>Intelen Services Ltd., Data Scientist Intern</b> Performed data analytics and applied machine learning on energy consumption time-series data for forecasting and customer segmentation.

## EDUCATION

SEPT 2019 - MAY 2025	<b>Ph.D. in Computer Science, University of Amsterdam</b> Thesis: <a href="#">Improving Passage Retrieval for Conversational Search</a>
SEPT 2017 - SEPT 2018	<b>M.Sc. Data Science, University of Amsterdam</b> Cum Laude   GPA: 8,4/10,0
SEPT 2008 - JUL 2015	<b>Dipl. Electrical &amp; Computer Engineering, Aristotle University of Thessaloniki</b>

## AREAS OF INTEREST

- INFORMATION RETRIEVAL (IR)
- RETRIEVAL AUGMENTED GENERATION
- LARGE LANGUAGE MODELS (LLMs)
- CONVERSATIONAL SEARCH

## PUBLICATIONS

CIKM 2025	<a href="#">Constructing Set-Compositional &amp; Negated Representations for First-Stage Ranking</a> A.M. Krasakis, A. Yates, E. Kanoulas
ArXiv (2024)	<a href="#">Corpus-Informed Retrieval Augmented Generation of Clarifying Questions</a> A.M. Krasakis, A. Yates, E. Kanoulas
TOIS 2023	<a href="#">Contextualizing and Expanding Conversational Queries without Supervision</a> A.M. Krasakis, A. Yates, E. Kanoulas
SIGIR 2022	<a href="#">Zero-shot Query Contextualization for Conversational Search</a> A.M. Krasakis, A. Yates, E. Kanoulas

ICTIR 2020 | [Analysing the Effect of Clarifying Questions on Document Ranking in Conversational Search](#)  
A.M. Krasakis, M. Aliannejadi, N. Voskarides, E. Kanoulas

AKBC 2019 | [Semi-supervised Ensemble Learning with Weak Supervision for Biomedical Relationship Extraction](#)  
A.M. Krasakis, E. Kanoulas, G. Tsatsaronis

## SHARED TASK PARTICIPATION

---

Summer '23 | TREC-iKAT:  
Combining LLMs & retrieval for creating personalized conversational assistants. Benchmarking LLMs in Retrieve-then-Generate (RAG) & Generate-then-Retrieve pipelines.  
 *Our submission achieved the best results in this track! [paper]*

Summer '21 | TREC-CAsT:  
Introduced a novel approach to perform zero-shot Conversational Search, using a token-level Dense Retriever. [\[paper\]](#)

## LANGUAGES

---

ENGLISH: Full professional proficiency [C1 Level]

GERMAN: Limited working proficiency [B2 Level]

GREEK: Mother-tongue