

Chris Samarin

College of Information and Computer Sciences, University of Massachusetts Amherst
140 Governors Drive, Amherst, Massachusetts 01003-9264

chris.samarinas@gmail.com - algotprog.com

SUMMARY

I'm a third year PhD student in Information Retrieval at UMass Amherst. My recent work is on conversational information seeking, representation learning and ranking for information retrieval. In the past, I have solved challenging NLP problems (contextual recommendations, conceptual search, relation extraction, fact-checking) at 3 AI start-ups (adtech, fintech, biotech) and in academia (NUS). I have also participated in the Alexa Prize twice (2022, 2023). My long-term goal is to help overcome the limitations of machine learning models (efficiency, robustness, limited training resources, grounding), and improve multimodal content organization & discovery.

EDUCATION

University of Massachusetts Amherst, Amherst, USA

September 2021 – 2026 (expected)

M.Sc./PhD in Computer Science

- Research in conversational information seeking at the [Center for Intelligent Information Retrieval](#) (CIIR)
- Supervisor: [Hamed Zamani](#)

Aristotle University of Thessaloniki, Thessaloniki, Greece

October 2013 – Jun 2017

Bachelor of Science (B.Sc.) in Computer Science

- GPA: **9.67**/10.00, ranked **1**/120, highest GPA since **1992** out of 2,500+ CS graduates, top **0.04%**
- Ranked **1**/120 in the national university entrance exam in my department and top **0.2%** nationwide.
- Thesis: *Mining answers from the web using information retrieval techniques and machine learning*

WORK EXPERIENCE

University of Massachusetts Amherst, Amherst, USA

September 2021 – 2026 (expected)

Research Assistant

- Research Assistant at the [Center for Intelligent Information Retrieval](#), supervised by [Hamed Zamani](#).
- Research in *conversational information seeking* (conversational IR, clarifying questions).
- Team Lead for the [Alexa Prize Taskbot Challenge 2 \(2022\)](#)
- Team Member for the [Alexa Prize Taskbot Challenge \(2021\)](#)

Adobe Research, Remote, USA

June 2022 – December 2022

Research Intern

- Research in *multi-step query reformulation for dense retrieval* at the Real-time Experiences and Algorithms Lab (REAL) with [Xiang Chen](#) and [Saayan Mitra](#).

Ro5, Remote

February 2021 – August 2021

Machine Learning Consultant (Knowledge Graph)

- **Ro5** is a biotechnology company that accelerates drug discovery, development, repurposing, and clinical trials.
- Responsible for the automated construction of a *biomedical knowledge graph* from research papers.
- Worked on *biomedical entity extraction*, *supervised entity resolution* and *relation classification*.

National University of Singapore, Singapore

February 2020 – April 2021

Research Assistant

- Research on *latent retrieval for large-scale fact-checking and question answering* with prof. Wynne Hsu, president of [IDS](#) and prof. Mong Li Lee, president of [CTIC](#). (IEEE ICTAI 2020)
- Developed a framework for large-scale semantic search that returns text snippets related to questions or statements and deployed it in a real-world semantic search engine for COVID-19, launched by the president of NUS.
- Research on *multi-stage evidence retrieval for large-scale fact-checking*. (NAACL 2021) and on *explainable fake news detection based on content and social context*.

Plural AI, London, United Kingdom
Research Engineer

October 2018 – January 2020

- **Plural AI** is a fintech start-up that builds a *knowledge engine* for the finance industry. The company focuses on extracting content from corporate websites and financial statements to enrich a financial *knowledge graph*, and support question answering on complex financial queries.
- Developed a novel framework for *concept extraction* and *saliency scoring* from corporate websites. Owned the whole machine learning pipeline, from the *annotation tool development* and website crawling, to the productionization and deployment.
- Developed a smart tool for efficient curation of an industrial *category graph*.
- Developed a scalable semantic search service of companies using a novel *topic modeling* technique.

Donaco LTD, London, United Kingdom
Research Engineer

June 2018 – January 2019

- **Donaco** is an online advertising start-up that tries to revolutionize online donations.
- Designed a system for contextual recommendation of charity campaigns in news articles.

Freelance Software Engineer (early stuff)

2009 – November 2018

- My first YouTube utility app got over **1M** downloads and was **acquired** by a small company in Poland.
- Developed a metasearch engine for a client in Germany.
- Developed a location-based web search interface for a client in Canada.
- Developed and sold a script that bypassed Google's Custom Search API limitations and offered an interface for 'instant search', which among others was also used by a [company](#) in Turkey that got **\$2.5M** government funding.
- I was a short-term contractor for the development of the new version of the *historical* [HotBot](#) metasearch engine.

RECENT & ONGOING RESEARCH

A Synthetic Data Generation Framework for End-To-End Retrieval-Augmented Task-Oriented Dialogue

[Chris Samarinas](#), Pracha Promthaw, Atharva Nijasure, Hansi Zeng, Julian Killingback, Hamed Zamani
under submission to COLM 2024

ProCIS: A Benchmark for Proactive Retrieval in Conversations

[Chris Samarinas](#), Hamed Zamani, *under review, SIGIR 2024*

Multi-Modal Augmentation for Large Language Models with Applications to Task-Oriented Dialogues

[Chris Samarinas](#), Pracha Promthaw, Rohan Lekhwani, Sheshera Mysore, Sung Ming Huang, Atharva Nijasure, Hansi Zeng, Hamed Zamani, proceedings of the *Alexa Prize Taskbot Challenge 2022*

Revisiting Open Domain Query Facet Extraction and Generation - [\[code\]](#)

[Chris Samarinas](#), Arkin Dharawat, Hamed Zamani, long paper at *ICTIR 2022*

MarunaBot: An Extensible Retrieval-Focused Framework for Task-Oriented Dialogues

Amit Hattimare, Arkin Dharawat, Yelman Khan, Yen-Chieh Lien, [Chris Samarinas](#), George Wei, Yulin Yang, Hamed Zamani, proceedings of the *Alexa Prize Taskbot Challenge 2021*

Improving Evidence Retrieval for Automated Explainable Fact-Checking - [\[code\]](#)

[Chris Samarinas](#), Wynne Hsu, Mong Li Lee, system demonstration at *NAACL 2021*

Latent Retrieval for Large-Scale Fact-Checking and Question Answering with NLI Training

[Chris Samarinas](#), Wynne Hsu, Mong Li Lee, long paper at *IEEE ICTAI 2020*
[\[code\]](#) [\[news\]](#) [\[blog post\]](#)

SKILLS

Programming languages

Python, Java, C, C++, Prolog

Web development

JavaScript, HTML5, PHP, MySQL, MongoDB

Machine Learning

Tensorflow, PyTorch

Others

Kubernetes, neo4j, GCP, Latex, Android development

AWARDS

Bioengineering Hackathon, 1st place , <i>EMB IEEE Student Chapter AUTH, StudentGuru</i> 1st place for developing DropSense in 48 hours, a platform that monitors and connects elderly with their relatives and paid carers using Wear OS smartwatches.	2018
Hackathon Thessaloniki, Honor , <i>Municipality of Thessaloniki</i> awarded for developing CityReport in 24 hours, an online web application for instant issues reporting	2014
Great Moment for Education Scholarship, Award , <i>Eurobank Ergasias</i> scholarship for achieving the highest grade at school at the national university admission examination	2013
National Mathematical Olympiad, First Prize , <i>Hellenic Mathematical Society</i> awarded for competing in the final stage "Archimedes", top 50 out of over 20,000 participants	2009

GRANTS

Bodossaki Foundation PhD scholarship , <i>Bodossaki Foundation</i> \$60,000 scholarship for my 5-year PhD studies in Information Retrieval	2021
Tensorflow Research Cloud Program , <i>Google</i> Granted free access to TPU pods for 6 months for our project on large-scale latent retrieval	2020
Innovate UK Smart Grant , <i>UK Research and Innovation</i> , 1% acceptance rate £400,000 for our work at Plural AI, building a knowledge engine for the finance industry	2019
Horizon 2020 Grant , <i>EU Funding for Research and Innovation</i> , 16% acceptance rate £50,000 for our research at Plural AI on entity extraction and semantic search	2019

SERVICE & MENTORING

Reviewer:

- ACM Transactions on Information Systems (TOIS)

Mentor:

- 2022 Early Research Scholars Program (ERSP) at UMass
- PhD Applicant Support Program (PASP) at UMass
- Supervised 3 master's and 1 undergraduate students on research projects