CHRIS SAMARINAS

College of Information and Computer Sciences, University of Massachusetts Amherst

140 Governors Drive, Amherst, Massachusetts 01003-9264

chris.samarinas@gmail.com - algoprog.com

SUMMARY

Machine Learning researcher with focus on applications in Natural Language Processing (Information Retrieval, Information Extraction, Question Answering, Knowledge Graphs, Recommendation Systems). Solved challenging problems during my work at 3 AI start-ups and 3 universities. PhD student in Information Retrieval at UMass Amherst.

EDUCATION

University of Massachusetts Amherst, Amherst, USA   September 2021 – 2026 (expected)
M.Sc./PhD in Computer Science

- Research in Information Retrieval 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 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 – Present
Research Assistant

- Research Assistant at the Center for Intelligent Information Retrieval, supervised by Hamed Zamani. Research in *conversational search* for the Alexa Prize Taskbot Challenge, funded by Amazon.

Adobe Research, Remote (or San Jose, CA), USA   June 2022 – August 2022
Incoming Summer Research Intern

- I'm going to do research in *contextual advertising* (TBD) with Xiang Chen from the Big Data Experience Lab.

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   October 2018 – January 2020
Research Engineer

- 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   June 2018 – January 2019
Research Engineer

- 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.

PUBLICATIONS

Improving Evidence Retrieval for Automated Explainable Fact-Checking
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*, **25%** acceptance rate
[code] [demo 1, demo 2] [news] [blog post]

WAMBy: An Information Retrieval Approach to Web-Based Question Answering
Chris Samarinas, Grigorios Tsoumakas, long paper at *SETN 2018*, **33%** acceptance rate

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, *EMB IEEE Student Chapter AUTH, StudentGuru*   2018
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.

Hackathon Thessaloniki, Honor, *Municipality of Thessaloniki*   2014
awarded for developing CityReport in 24 hours, an online web application for instant issues reporting.

Great Moment for Education Scholarship, Award, *Eurobank Ergasias*   2013
scholarship for achieving the highest grade at school at the national university admission examination.

National Mathematical Olympiad, First Prize, *Hellenic Mathematical Society*   2009
awarded for competing in the final stage "Archimedes", top 50 out of over 20,000 participants.

GRANTS

Bodossaki Foundation PhD scholarship, *Bodossaki Foundation*   2021
$60,000 scholarship for my 5-year PhD studies in Information Retrieval.

Tensorflow Research Cloud Program, *Google*   2020
Granted free access to TPU pods for 6 months for our project on large-scale latent retrieval.

Innovate UK Smart Grant, *UK Research and Innovation,* **1%** acceptance rate   2019
£400,000 for our work at Plural AI, building a knowledge engine for the finance industry.

Horizon 2020 Grant, *EU Funding for Research and Innovation,* **16%** acceptance rate   2019
£50,000 for our research at Plural AI on entity extraction and semantic search.

REFERENCES

- **Wynne Hsu,** Professor of CS at *National University of Singapore*, Director of *IDS*, co-founder of *EyRIS*

- **Jaron Collis**, co-founder and former CTO at *Plural AI*

- **Grigorios Tsoumakas**, Associate Professor of Machine Learning and Knowledge Discovery at *Aristotle University of Thessaloniki*, Greece, co-founder of *Medoid AI*