

Nikos Papasasantopoulos



work experience

senior data scientist

Priceline, Booking Holdings

Oct 2022 – Jul 2024

Data Science Team, Priceline Partner Solutions

remote, UK

- development, design, and productionisation of ML solutions powering B2B systems of Priceline Partner Solutions.
- replaced legacy NLP application with modern, more explainable, tagger solution: increased mean accuracy to 97%, reduced latency by 50%, increased trust; the app is now the default for the whole inventory in production.
- prototype, design, and deployment of a computer vision pipeline, enabling multiple B2B vision applications.
- work in automation of training and evaluation pipelines for ML applications; automated evaluation process for NLP applications, reducing time to deployment of trained models by 24h.
- regular prototypes for ML applications, including chatbots, inventory item deduplication, pricing restriction language simplification: frequent communication and presentations to stakeholders and deployment of live demos.

research scientist

Huawei Technologies UK R&D

2019 – 2022

Knowledge Graph Team

Edinburgh, UK

- research on Natural Language Processing applications over Knowledge Graphs.
- results disseminated in technical reports, presentations, research papers, and patent applications.
- work on various stages of the research cycle: literature review, brainstorming, prototype implementation, experiments, scalability optimisations, and related engineering problems.
- introduced the team to development best practices for research, including testing, code reviews, and CI/CD.

research engineer

Institute for Language & Speech Processing

2014 – 2016

Athena Research Centre

Athens, Greece

- NLP for search engine customisation and multilingual support: developed filters for Apache Solr and Elasticsearch.
- implemented/enhanced search functionality in several web applications, resulting in improved search quality.
- work on research projects on entity linking, question answering, and information extraction.

education

- **PhD in NLP and Machine Learning**, University of Edinburgh, UK. **2016 – 2019**
thesis: *Multi-view Representation Learning for Natural Language Processing Applications*.
- **MEng in Electrical and Computer Engineering**, National Technical University of Athens, Greece. **2006 – 2012**
thesis on unsupervised entity and relation extraction from text in Greek language.

technologies and languages

- languages: python, SQL, past experience with Java.
- ML: pytorch, tensorflow, keras, numpy, scipy, scikit-learn, huggingface, mlflow, sagemaker, langchain.
- web: flask, FastAPI, Django.
- cloud: AWS, kubernetes.
- dev: git, CI/CD pipelines.

other activities

- **peer review**

AAACL (2020, 2022, 2023), ACL (2019, 2021, 2023), EACL (2021, 2023), EMNLP (2018, 2019, 2020, 2021, 2022, 2023), IJCAI (2020), LREC (2022), NAACL (2021), Semantic Web Journal (2023).

- **teaching**

School of Informatics, University of Edinburgh, 2017 – 2019

tutor and marker in “Informatics Research Review”, marker in “Processing Formal and Natural Languages”, teaching assistant in “Topics in Natural Language Processing”, tutor in “Accelerated NLP”.

School of Electrical and Computer Engineering, National Technical University of Athens, 2007 – 2008

lab assistant, “Introduction to Computer Programming”.

private tutor, 2007 – 2016

professional high school private tutor, preparing students for Greek national university entrance exams.

publications

WenYu Huang, Mirella Lapata, Pavlos Vougiouklis, Nikos Papasrantopoulos, and Jeff Pan. *Retrieval Augmented Generation with Rich Answer Encoding*. IJCNPP-AAACL 2023.

Pavlos Vougiouklis, Nikos Papasrantopoulos, Danna Zheng, David Tuckey, Chenxin Diao, Zhili Shen, and Jeff Pan. *FASTRAT: Fast and Efficient Cross-lingual Text-to-SQL Semantic Parsing*. IJCNPP-AAACL 2023.

Tianyi Li, WenYu Huang, Nikos Papasrantopoulos, Pavlos Vougiouklis, and Jeff Z. Pan. *Task-specific Pre-training and Prompt Decomposition for Knowledge Graph Population with Language Models*. arXiv, abs/2208.12539, 2022.

Georgios Stoilos, Nikos Papasrantopoulos, Pavlos Vougiouklis, and Patrik Banský. *Type Linking for Query Understanding and Semantic Search*. KDD Applied Data Science Track, 2022.

Nikos Papasrantopoulos and Shay B. Cohen. *Narration Generation for Cartoon Videos*. arXiv, abs/2101.06803, 2021.

Nikos Papasrantopoulos, Lea Frermann, Mirella Lapata, and Shay B. Cohen. *Partners in crime: Multi-view Sequential Inference for Movie Understanding*. EMNLP-IJCNLP, 2019.

Shashi Narayan, Ronald Cardenas, Nikos Papasrantopoulos, Shay B. Cohen, Mirella Lapata, Jiangsheng Yu, and Yi Chang. *Document Modeling with External Attention for Sentence Extraction*. ACL, 2018.

Nikos Papasrantopoulos, Helen Jiang, and Shay B. Cohen. *Canonical Correlation Inference for Mapping Abstract Scenes to Text*. AAAI, 2018.

Shashi Narayan, Nikos Papasrantopoulos, Mirella Lapata, and Shay B. Cohen. *Neural Extractive Summarization with Side Information*. arXiv, abs/1704.04530, 2017.

Renars Liepins et al. *The SUMMA platform prototype*. EACL, Software Demonstrations, 2017.

Konstantina Papanikolaou, Haris Papageorgiou, Nikos Papasrantopoulos, Theoni Stathopoulou, and George Papastefanatos. *“Just the facts” with PALOMAR: Detecting Protest Events in Media Outlets and Twitter*. AAAI ICWSM, 2016.

Nikos Papasrantopoulos, Nick Poulakis, and Christina Anagnostopoulou. *A Computational Analysis Study of Children’s Songs from Different Countries*. ESCOM, 2015.

Georgios Stratiogiannis, Nikos Papasrantopoulos, Maria Pontiki, Giorgos Siolas, Dimitris Galanis, Andreas Stafylopatis, and Haris Papageorgiou. *NTUA ILSP in the TAC KBP 2014 Entity Linking Challenge*. 2014.

patents

Nikos Papasrantopoulos, Pavlos Vougiouklis, David Tuckey, Danna Zheng, and Jeff Pan. *Devices and Methods for Generating an SQL Query based on Natural Language Query*. 2024.

Nikos Papasrantopoulos, Pavlos Vougiouklis, Patrik Banský, and Jeff Pan. *Devices and Methods for Selecting Answer Entities of a Knowledge Base*. 2022.

Georgios Stoilos, Nikos Papasrantopoulos, Pavlos Vougiouklis, Patrik Banský, and Yantao Jia. *Apparatus and Method for Type Matching of a Text Sample*. 2021.