Christos Theodoropoulos, PhD

Senior Data Scientist at EarlyTracks
Brussels, Belgium, 1000
WWW: https://christos42.github.io/
christostheodoropoulos@outlook.com

Professional Statement - Summary

Purpose-driven Senior Data Scientist improving the way electronic health records are managed. Results-driven PhD researcher at KU Leuven specializing in Natural Language Processing and Knowledge Extraction focused on the biomedical domain, leveraging an extensive background as a Research Scientist Intern at IBM Research, where I contributed to advancements in Person-Centric Knowledge Graphs, filed a patent, and pioneered an open-source research project. Accomplished Machine Learning Engineer and Researcher with the I-SENSE Group, leading a pivotal role in a Computer Vision project centered on eye tracking and object detection. Proven expertise as a Data Scientist at a dynamic tech startup in life sciences, specializing in signal processing, feature selection, and signal reconstruction. A fast learner with a keen appetite for exploring and evolving, I thrive in challenging environments and desire to engage in purposeful projects. My personal and collaborative growth mindset positions me as an adaptable professional, ready to contribute to innovative solutions and tackle complex challenges.

Professional and Research Experience

Senior Data Scientist Jan. 2025 – Present
EarlyTracks *Brussels, Belgium, Hybrid*
- Apply multi-lingual Natural Language Processing solutions to process and analyze diverse healthcare data.
- Implement Named Entity Recognition (NER) and Named Entity Linking (NEL) approaches to extract critical medical information and link it to medical terminologies (e.g. SNOMED CT).
- Develop Relation Extraction models to identify relationships within the data, improving the semantic understanding of medical records.
- Leverage Large Language Models (LLMs) for deep learning-based understanding and processing of healthcare content.
- Collaborate closely with cross-functional teams to design scalable solutions that are not only technically robust but also meet the real-world needs of healthcare professionals.

Doctoral Researcher Sep. 2020 – Dec. 2024
LIIR, KU Leuven *Leuven, Belgium*
- Research in Information Extraction, Knowledge Discovery, and Entity-Centric Knowledge Graphs focused on the biomedical domain.
- Developed an end-to-end open-sourced paradigm enhancing Knowledge Discovery for diseases from biomedical publications.
- Implemented a framework using the Contrastive Learning paradigm, incorporating knowledge graphs and unstructured text.
- Taught exercises and prepared capstone project in Information Retrieval and Search Engines course.
- Supervised two research-oriented master theses for the advanced master in Artificial Intelligence.
- Participated in OxML summer schools on health and bio AI and representation learning and generative AI, organized by AI for Global Goals in collaboration with CIFAR and the University of Oxford.

Research Scientist Intern Jul. 2022 – Nov. 2022
IBM Research *Dublin, Ireland*
- Researched Person-Centric Knowledge Graph Extraction from Electronic Health Records using HSPO ontology.
- Applied Graph Neural Network (GNN) training for solving Intensive Care Unit readmission prediction.
- Experimented with embedding learning for heterogeneous graphs utilizing various graph structures.
- Abstractly developed the framework to be applicable to other downstream predictive tasks (e.g. mortality prediction).
- Heavily contributed to IBM's open-source project (HSPO ontology).
- Published two research papers as the first author and filed a patent as the primary inventor.

Machine - Deep Learning Engineer, Researcher Mar. 2019 – Sep. 2019
I-SENSE Group, ICCS, NTUA *Athens, Greece*
- Worked on an internal project related to Eye-Tracking and Object-Detection systems.
- Implemented advanced Deep Learning (DL) architectures for Gaze Localization.
- Developed process for Generative Adversarial Networks quality evaluation using pre-trained DL models and Random Forest.
- Performed data pre-processing and manipulation for the unification of image datasets related to driving conditions.

Data Scientist Mar. 2018 – Jul. 2018
Feel Therapeutics *Athens, Greece*
- Developed proprietary algorithms for signal processing of Heart Rate Variability (signal reconstruction and quality estimation).
- Created procedure for data retrieval from cloud storage space (AWS Athena).

Education

KU Leuven Sept. 2020 – Mar. 2025
Doctor of Philosophy - PhD, Computer Science *Leuven, Belgium*
- Research Areas: Natural Language Processing, Information Extraction, Deep Learning, Knowledge Graphs, Knowledge Discovery
- Advisor: Prof. Marie-Francine Moens

KU Leuven Sept. 2019 – Aug. 2020
Advanced MSc, Artificial Intelligence *Leuven, Belgium*
- Grade: 81.17%, Magna Cum Laude - Great Distinction
- Dissertation: "Automatic artifact removal of resting-state fMRI with Deep Neural Networks", Advisor: Prof. Sabine Van Huffel

National and Technical University of Athens Oct. 2012 – Oct. 2018
Integrated BSc & MSc, Electrical & Computer Engineering *Athens, Greece*
- Grade: 81.8%, Great Distinction
- Dissertation: "Deep Learning Techniques for Emotion Recognition through Facial Expressions", Advisor: Prof. Stafylopatis

Selected Publications and Patents

Christos Theodoropoulos, James Henderson, Andrei C. Coman, Marie-Francine Moens. **Enhancing Biomedical Knowledge Discovery for Diseases: An Open-Source Framework Applied on Rett Syndrome and Alzheimer's Disease**. IEEE Access 12. 2024. 180652–180673. (Journal)

Christos Theodoropoulos, Natalia Mulligan, Joao H Bettencourt-Silva. **Evaluating the Predictive Features of Person-Centric Knowledge Graph Embeddings: Unfolding Ablation Studies**. In Proceedings of the 34th Medical Informatics Europe Conference. 2024. (MIE '24). (oral presentation)

Christos Theodoropoulos, Natalia Mulligan, Thaddeus Stappenbeck, Joao H Bettencourt-Silva. **Representation Learning for Person or Entity-centric Knowledge Graphs: An Application in Healthcare.** In Proceedings of the 12th Knowledge Capture Conference. 2023. (K-CAP '23). Association for Computing Machinery. (oral presentation)

Christos Theodoropoulos, Natalia Mulligan, Joao H Bettencourt-Silva, Marco Luca Sbodio. **Method/System to rank and improve usage of relations in learning Multi-Relational Graphs using Graph Neural Networks.** U.S. Patent. 2023. (filed)

Christos Theodoropoulos, Marie-Francine Moens. **An Information Extraction Study: Take In Mind the Tokenization!** In Proceedings of the 13th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT). 2023. (oral presentation)

Christos Theodoropoulos, James Henderson, Andrei C. Coman, Marie-Francine Moens. **Imposing Relation Structure in Language-Model Embeddings Using Contrastive Learning.** In Proceedings of the 25th Conference on Computational Natural Language Learning. 2021. (oral presentation)

Christos Theodoropoulos, Christos Chatzichristos, and Sabine Van Huffel. **Automatic artifact removal of resting-state fMRI with Deep Neural Networks.** In Proceedings of the 29th European Signal Processing Conference (EUSIPCO). IEEE, 2021. (oral presentation)

Selected Projects

Person-Centric Knowledge Graph Extraction
- Developed a data pre-processing and analysis pipeline to extract Person-Centric Knowledge Graphs in Resource Description Framework format using HSPO ontology schema and MIMIC-III dataset.
- Implemented a downstream task-agnostic transformation process to create different graph versions, experiment with the level of heterogeneity, and extract knowledge graphs in a GNN-friendly format.
- Trained Graph Convolution/Attention Networks to solve the ICU readmission prediction task using an insufficiently large dataset (appr. 2.5K training records).

Automatic Denoising of Resting-State fMRI
- Designed and developed a Deep Learning framework using 3D CNN and LSTM layers for artifact removal of resting-state fMRI.
- Handled spatial and temporal information by implementing a weighted voting schema.
- Achieved comparable to state-of-the-art performance.

Emotion Recognition
- Manipulated video data annotated with emotion markers using Deep Learning and Transfer Learning techniques for emotion recognition through facial expressions.
- Approached the problem as classification and regression based on the emotional space (valence-arousal).

Honors and Awards

Scholarship for PhD studies, FWO fellowship
KU Leuven *Sep. 2020*

Scholarship for MSc studies
Bodossaki Foundation and Eugenides Foundation *Sep. 2019*

Scholarship for BSc and MSc studies
Bakopouleio Foundation *2012-2017*

Technical, Soft Skills and Talks

Programming: Python (PyTorch, PyTorch-Geometric, Scikit-learn, Huggingface, Spacy, Numpy, ONNX, Streamlit, vLLM, llama-cpp), Matlab, C

Agile Development, Versioning, and Project Management: Git, DVC, Poetry, Hydra, JIRA, Scrum

Cloud and DevOps: Azure, Docker, fastAPI

Conference/Other Talks: DTAI Seminar '24 - KU Leuven, MIE '24, Keynote talk on Knowledge Discovery - Idiap visit '24, EUSFLAT '23, Show & Tell - IBM Research '22, CoNLL '21, EUSIPCO '21

Job Shadowing: Data/AI Scientist Role in B12 Consulting '24, Business Day in Upstream '18, Business Day in Oracle '18

Languages: English (fluent), Greek (native), French (elementary)

Soft skills: Learnability, Communication, Collaboration, Critical Thinking, Problem-Solving, Time Management, Adaptability, Mentorship, Attention to Detail, Active listening, Self-motivation