

Nikitas Theodoropoulos

✉ nikitastheodorop@gmail.com

🌐 [nikitas-theo](https://nikitas-theo.github.io)

🔗 nikitas-theo.github.io

EDUCATION

National Technical University of Athens (NTUA)

BSc & MSc in Electrical and Computer Engineering (5-year joint degree; 300 ECTS)

Sep 2015 - Nov 2024

Athens, Greece

• GPA: 8.83/10, Major: Computer Science

• Final thesis project: Leveraging Synthetic Story Data for Sample-Efficient Language Pre-training.

PUBLICATIONS

1. **BabyBabelLM: A Multilingual Benchmark of Developmentally Plausible Training Data.** Jaap Jumulet, ..., Nikitas Theodoropoulos, ..., Leshem Choshen (26 authors, listed alphabetically). 2025. Under Review. [Arxiv](#) [🔗](#)
2. **BERTtime Stories: Investigating The Role of Synthetic Story Data in Language Pre-training.** Nikitas Theodoropoulos, Giorgos Filandrianos, Vassilis Lyberatos, Maria Lymperaïou, and Giorgos Stamou. *The 2nd BabyLM Challenge at CoNLL*. 2024. [ACL](#) [🔗](#)
3. **From {Solution Synthesis} to {Student Attempt} Synthesis for Block-Based Visual Programming Tasks.** Adish Singla and Nikitas Theodoropoulos. *International Conference on Educational Data Mining (EDM)*. 2022. [Arxiv](#) [🔗](#)

EXPERIENCE

Independent Researcher at BabyBabelLM

BabyBabelLM Project [🔗](#)

April 2025 - Oct 2025

Remote participation

• Created a multilingual benchmark of developmentally plausible data for 45 languages.

• Collaborated with an international team, collected data for 3 languages, worked on data processing and validation.

MSc Thesis Researcher at AILS of NTUA

Artificial Intelligence and Learning Systems Lab (AILS) at NTUA [🔗](#)

May 2024 - Nov 2024

Athens, Greece

• Investigated efficient and cognitively-inspired language modeling, advised by Prof. [Giorgos Stamou](#).

• Trained models with human-like data constraints of at most 100 million words.

• Work published at the 2nd BabyLM Challenge at CoNLL, 2024.

Research Engineer at the department of Informatics at NKUA

Department of Informatics at the National Kapodistrian University of Athens (NKUA) [🔗](#)

Sep 2023 - Apr 2024

Athens, Greece

• Investigated multi-agent RL methods for minimizing carbon emissions and energy usage in computer networks.

• Participated in writing industry and EU-funded research proposals.

Research Internship (Visiting Scholar, Fellowship) at MPI-SWS

Machine Teaching group at MPI-SWS [🔗](#)

Feb 2021 - Jul 2022

Saarbrücken, Germany

• Research topics: AI for education, program synthesis, reinforcement learning, advised by Prof. [Adish Singla](#).

• Designed a benchmark for predicting student behavior at block-based visual programming. Published at EDM, 2022.

Undergraduate Researcher at SLP of NTUA

Speech and Language Processing group (SLP) at NTUA [🔗](#)

Feb 2020 - Feb 2021

Athens, Greece

• Investigated use of brain-derived word representations (fMRI) for improving NLP task performance advised by Prof. [Alexandros Potamianos](#).

• Problem: find an unsupervised dictionary between two languages using brain-derived word embeddings.

CMMRS Pre-doctoral Research School

The Cornell, Maryland, Max Planck Pre-doctoral Research School (CMMRS) [🔗](#)

Aug 3-9, 2020

Saarbrücken, Germany (virtually)

• Selected for a competitive one-week program, learned about state-of-the-art research and prepared for grad school.

Teaching Experience

School of Electrical & Computer Engineering (ECE) at NTUA [🔗](#)

Athens, Greece

• TA for university courses, *Intro to Programming* and *Programming Techniques*.

2016-2017

• Gave algorithm courses to prepare the national team for the Junior Balkan Olympiad of Informatics.

2017

ACHIEVEMENTS

IEEEExtreme 14.0 coding contest, 2nd place in Greece, top 2% worldwide

2021

IEEEExtreme 13.0 coding contest, 2nd place in Greece, top 4% worldwide

2020

Award by EFG Eurobank, distinction in the Greek Nationwide University Entrance Examination

2015

SKILLS

Technical skills: Python, PyTorch, Scikit-Learn, NumPy, Pandas, Matplotlib, C, C++, Rust, ~~TeX~~TeX, Unix/Linux, Bash, Haskell, SQL, MatLab, Assembly, basic server/system administrator skills.

Languages: Greek (native), English (fluent, C2), French (conversational), German (basic).