ATHANASIOS (NASSOS) KATSAMANIS

Cofounder, CTO, and Research Director in Speech and Conversational AI

nkatsam@athenarc.gr | Mistra 4, 18757, Keratsini, Greece | cvsp.cs.ntua.gr/~nassos | nassosoassos
nkatsam | nkatsam

EXPERIENCE

CTO & Co-founder

Auxilis AI

Mar 2025 – ongoing | Athens, Greece

- Working on R&D activities related to the company's main product, i.e., a conversational AI agent for healthcare targeting UK primary care.

Research Director (on leave)

Athena R.C., ILSP

Feb 2019 – ongoing | Athens, Greece

- Doing research in speech analysis, affective computing, conversational AI, multimodal speech processing, multimodal interfaces, human behavior modeling.
- Principal investigator in a national research project on developing technologies for the therapy of speech (PLan-V), a European project on virtual assistants for healthcare (DataTools4Heart) and projects with industrial partners (Cosmote, NBG) and startups (Behavioral Signals, soniphi).
- Leading the R&D group working on automated spoken dialog systems.
- Teaching two graduate classes at the University of Athens on speech recognition, synthesis, and dialogue systems. Supervisor of 12 MSc/MEng thesis students.
- Elected member of the Scientific Council and deputy director (since 2021) of the Institute for Language and Speech Processing (ILSP)

Head of Technology

Behavioral Signals

Dec 2016 – Feb 2025 | Los Angeles, CA

- Leading the design and implementation of R&D activities related to the company's main product, namely an engine for emotion recognition from speech and a deepfake speech detection system.

Cofounder and CTO

beenotes

Jan 2015 – Nov 2016 | Athens, Greece

- Led the design and implementation of the company's R&D plan and incorporated automatic speech recognition and synthesis technologies into the company's products. Developed an automated dialog system as an application for mobile phones.

Research Associate

Athena R.C. and National Technical University of Athens

Mar 2013 – Jul 2016 | Athens, Greece

- Did research and supervised graduate students on various research themes, including keyword detection in speech, distant speech recognition in smart homes, and emotion recognition from speech.

MOST PROUD OF

Works citing my research
I have more than 80 publications in international peer-reviewed journals and conferences with more than 3000 citations in total

Participating in Research Projects
I have participated in more than 15 national and international research projects in Greece, Europe and in the US, as a principal investigator (2), research associate (8), or research assistant (5)

Scholarships and awards
I have been awarded 3 scholarships (Onassis Public Foundation, Greek National Scholarship Organization) for my studies and research and 2 awards for exceptional performance during my studies

Open Source Software
I have developed sailalign, an open-source software for speech-text alignment, that has been used by more than 100 research groups worldwide.

Social Impact
I have been one of the main contributors to Theano, the Greek dialogue system for COVID-19. Theano presented users with COVID-19 statistics and facts and informs them about the best health practices as well as the latest COVID-19 related guidelines.

EDUCATION

PhD, Electrical and Computer Engineering

National Technical University of Athens

Dec 2003 – Jun 2009

Dipl. Eng., Electrical and Computer Engineering

National Technical University of Athens

Sep 1998 – Oct 2003

Postdoctoral Research Associate

Univ. of Southern California

Oct 2009 – Apr 2012 Los Angeles, CA

- Did research and supervised graduate students in various research projects related to speech technologies, virtual agents, and emotion recognition.

PUBLICATION HIGHLIGHTS

Indicative Works I have Co-authored

- Paraskevopoulos, G. et al. (2024). "Sample-Efficient Unsupervised Domain Adaptation of Speech Recognition Systems: A Case Study for Modern Greek". In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 32, pp. 286–299. DOI: 10.1109/TASLP.2023.3328280.
- Sgouropoulos, D. et al. (2024). "Emotion-Aware Speech Popularity Prediction: A Use-Case on TED Talks". In: *2024 12th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 300–308. DOI: 10.1109/ACII63134.2024.00039.
- Voukoutis, L. et al. (2024). *Meltemi: The first open Large Language Model for Greek*. arXiv: 2407.20743 [cs.CL]. URL: https://arxiv.org/abs/2407.20743.
- Chatzoudis, G. et al. (2022). "Zero-Shot Cross-lingual Aphasia Detection using Automatic Speech Recognition". In: *Interspeech*.
- Ciaperoni, M. et al. (2022). "SIEVE: A Space-Efficient Algorithm for Viterbi Decoding". In: *Proceedings of the 2022 International Conference on Management of Data*, pp. 1136–1145.
- Ventoura, N. et al. (Aug. 2021). "Theano: A Greek-speaking conversational agent for COVID-19". In: *Proceedings of the 1st Workshop on NLP for Positive Impact*. Online: Association for Computational Linguistics, pp. 36–46.
- Chatziagapi, A. et al. (2019). "Data Augmentation Using GANs for Speech Emotion Recognition." In: *Interspeech*, pp. 171–175.
- Tsiami, A. et al. (2019). "A behaviorally inspired fusion approach for computational audiovisual saliency modeling". In: *Signal Processing: Image Communication* 76, pp. 186–200.
- Filntisis, P. P. et al. (2017). "Video-realistic expressive audio-visual speech synthesis for the Greek language". In: *Speech Communication* 95, pp. 137–152.
- Katsamanis, A., V. Pitsikalis, et al. (2017). "Multimodal gesture recognition". In: *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations-Volume 1*, pp. 449–487.
- Rodomagoulakis, I. et al. (2017). "Room-localized spoken command recognition in multi-room, multi-microphone environments". In: *Computer Speech & Language* 46, pp. 419–443.
- Gibson, J. et al. (2015). "Multiple instance learning for behavioral coding". In: *IEEE Transactions on Affective Computing* 8.1, pp. 81–94.
- Lee, C.-C. et al. (2014). "Computing vocal entrainment: A signal-derived PCA-based quantification scheme with application to affect analysis in married couple interactions". In: *Computer Speech & Language* 28.2, pp. 518–539.
- Narayanan, S. et al. (2014). "Real-time magnetic resonance imaging and electromagnetic articulography database for speech production research (TC)". in: *The Journal of the Acoustical Society of America* 136.3, pp. 1307–1311.

REFEREES

Prof. Petros Maragos
@ National Technical University of Athens
maragos@cs.ntua.gr

Prof. Shrikanth Narayanan
@ Univ. of Southern California
shri@ee.usc.edu

LIFE PHILOSOPHY

"If you have built castles in the air, your work need not be lost; that is where they should be. Now put the foundations under them." (H.D. Thoreau)

- Black, M. P. et al. (2013). “Toward automating a human behavioral coding system for married couples’ interactions using speech acoustic features”. In: *Speech communication* 55.1, pp. 1–21.
- Metallinou, A., A. Katsamanis, and S. Narayanan (2013). “Tracking continuous emotional trends of participants during affective dyadic interactions using body language and speech information”. In: *Image and Vision Computing* 31.2, pp. 137–152.
- Metallinou, A., M. Wollmer, et al. (2012). “Context-sensitive learning for enhanced audiovisual emotion classification”. In: *IEEE Transactions on Affective Computing* 3.2, pp. 184–198.
- Katsamanis, A., M. Black, et al. (2011). “SailAlign: Robust long speech-text alignment”. In: *Proc. of workshop on new tools and methods for very-large scale phonetics research*.
- Katsamanis, A., G. Papandreou, and P. Maragos (2009). “Face active appearance modeling and speech acoustic information to recover articulation”. In: *IEEE Transactions on Audio, Speech, and Language Processing* 17.3, pp. 411–422.
- Papandreou, G. et al. (2009). “Adaptive multimodal fusion by uncertainty compensation with application to audiovisual speech recognition”. In: *IEEE Transactions on Audio, Speech, and Language Processing* 17.3, pp. 423–435.