Tiberiu Boros

Date of birth: 02/10/1985 | **Nationality:** Romanian | **Phone number:**

(+40) 745310081 (Mobile) | **Email address:** tiberiu44@gmail.com | **Website:**

https://www.linkedin.com/in/tiberiu-boros-68913854 | **Address:** Romania (Home)

europass

WORK EXPERIENCE

2017 – CURRENT
MACHINE LEARNING ENGINEER ADOBE SYSTEMS ROMANIA

Research and development centred on implementing and integrating machine learning models in applications.
1. machine learning for interpreting application logs
2. active development and maintenance of an open-source NLP framework: https://github.com/adobe/NLP-Cube
3. two accepted patents and one pending
4. independent contribution related to machine learning on various on-going projects

9 – 2018 Bucharest, Romania
SEARCH SCIENTIST II RESEARCH INSTITUTE FOR ARTIFICIAL INTELLIGENCE, ROMANIAN ACADEMY

Research activities centered on:
1. Spoken language processing: Text-to-speech Synthesis using data-driven approaches (Concatenative Unit Selection, Statistical Parametric Speech Synthesis and Neural generative Speech Synthesis)
2. Natural language processing: (a) low-level text processing such as Tokenization, Part-of-Speech Tagging, Lemmatization, Syllabification, Letter-to-Sound conversion, Lexical Stress Detection; (b) high-level processing: text classification, document retrieval, dialogue systems, machine translation, parsing
3. Interesting projects and results: large tagset labeling with feed forward neural networks, offline query alteration (Microsoft Speller Challenge) based on Google 1TB corpus, development of a complete processing framework for text-to-speech synthesis, small-data-driven text classification and parameter extraction for dialogue systems using distributed word representations and convolutional neural networks, continuous handwriting recognition for old-style English using Bidirectional Long Short-Term Memory Networks, background learning for adaptive motion detection and object tracking

2009 – 2016
SOFTWARE DEVELOPER EXPERT ONE RESEARCH

oftware development related to document indexing solution:
1. automatic document classification
2. OCR/ICR recogniton and custom-solutions
3. continuous handwriting recognition

01/08/2005 – 01/03/2009 Pitesti, Romania
SOFTWARE DEVELOPER S.C. IWAVE SOLUTIONS S.R.L.

At that time, iWave was a small company in which job attributes were not very well defined. As such, my roles within the company were:
1. System administration for a small ISP cluster: ~500 clients with PPTP Internet and hosted website+emails.
2. Developing custom-tailored software in: C++, PhP and JAVA
3. Development of custom firmware for ARM-based routers: the main goal of the project was to enable ChilliSpot in ZCom Wireless Routers

EDUCATION AND TRAINING

01/11/2009 – 01/11/2013 Bucharest, Romania
PHD IN COMPUTER SCIENCE Research Institute for Artificial Intelligence "Mihai Draganescu", Romanian Academy

The title of my thesis is "Contributions to the modeling and implementation of Text-to-Speech Synthesis systems. Case study: Romanian Language"
During the preparation of my thesis I validated my research by publishing original contributions in international conferences (ACL, LREC, RANLP, IEEE SpeD and many others) and by participating with success in a number of international challenges: Microsoft Speller Challenge (4th place and a Microsoft Research Award), Blizzard Challenge 2013 and IWSLT.
I graduated in 2013 with "summa cum laude".

Level in EQF EQF level 8

01/10/2005 – 01/08/2009 Pitesti, Romania
BACHELORS DEGREE IN COMPUTER SCIENCE Faculty of Electronics, Computers and Telecomunications, University of Pitesti

Level in EQF EQF level 6

LANGUAGE SKILLS

Mother tongue(s): **ROMANIAN**
Other language(s):

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken production	Spoken interaction	
ENGLISH	C2	C2	C1	C1	C2

Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user

DIGITAL SKILLS

Python ┊ Machine Learning ┊ Pytorch ┊ Natural Language Processing ┊ Text to Speech

PUBLICATIONS

Publications (selected)

2024
- Boroş, T., Chivereanu, R., Dumitrescu, S., & Purcaru, O. (2024, May). Fine-Tuning and Retrieval Augmented Generation for Question Answering Using Affordable Large Language Models. In *Proceedings of the Third Ukrainian Natural Language Processing Workshop (UNLP)@ LREC-COLING 2024* (pp. 75-82).

2023
- Boros, T., Dumitrescu, S. D., Mironica, I., & Chivereanu, R. (2023). Generative Adversarial Training for Text-to-Speech Synthesis Based on Raw Phonetic Input and Explicit Prosody Modelling. Speech Synthesis Workshop

2022
- Boros, T. et al. Machine Learning and Feature Engineering for Detecting Living off the Land Attacks. In *IoTBDS*.

2021
- Boros, T. et al. A Principled Approach to Enriching Security-related Data for Running Processes through Statistics and Natural Language Processing. In *IoTBDS*.

2019
- Boroş, T. et al. *Tripod: Learning Latent Representations for Sequences*. In ConsiLR

2018
- Boroş, T. et al. "NLP-Cube: End-to-end raw text processing with neural networks." *In CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*.
- Boroş, T. et al. "GBD-NER at PARSEME Shared Task 2018: Multi-Word Expression Detection Using Bidirectional Long-Short-Term Memory Networks and Graph-Based Decoding." *In LAW-MWE-CxG-2018*.
- Dumitrescu, S.D., and Boros, T. "Attention-free encoder decoder for morphological processing.". *In CoNLL SIGMORPHON*.

2017
- Dumitrescu, S.D., Boroş, T., and Tufiş, D. "RACAI's Natural Language Processing pipeline for Universal Dependencies." *In CoNLL UD Shared Task.*
- Boros, T. et al. "Fast and accurate decision trees for natural language processing tasks.". In *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP.*
- Boroş, T et al. "A data-driven approach to verbal multiword expression detection. PARSEME Shared Task system description paper." *In MWE.*

2015
- Boroş, T. et al. Robust deep-learning models for text-to-speech synthesis support on embedded devices. In MEDES'15.

2014
- Boroş, T. et a. RACAI GEC – A hybrid approach to Grammatical Error Correction. In Proceedings of CoNLL. Baltimore, USA
- Dumitrescu, S.D., Boroş, T. and Ion, R. Crowd-sourced, automatic speech-corpora collection – building the Romanian Anonymous Speech Corpus. In CCURL2014.
- Boroş, T., Stan, A., Watts, O. and Dumitrescu, S.D. RSS-TOBI - A Prosodically Enhanced Romanian Speech Corpus. In LREC.

2013
- Dumitrescu, S.D. and Boroş, T.. A unified corpora-based approach to Diacritic Restoration and Word Casing. In LTC 2013
- Boroş, T. et al. Large tagset labeling using Feed Forward Neural Networks. Case study on Romanian Language. In ACL.
- Boroş, T. et al. Improving the RACAI Neural Network MSD Tagger. In Engineering Applications of Neural Networks. Springer, vol. 383
- Boroş, T. Contributions to the modeling and implementation of Text-To-Speech Synthesis System. Case Study: Romanian Language. PhD thesis, Research Institute for Artificial Intelligence "Mihai Drăgănescu", Romanian Academy

HONOURS AND AWARDS

Honours and awards

PhD in Computer Science "summa cum laude"
Microsoft Research Award
Romanian Academy recognition award

ORGANISATIONAL SKILLS

Organisational skills

I have had previous experience as a team leader at iWave Solutions and also I have a solid experience as a WP-leader in multiple EU and government-funded projects at RACAI.

COMMUNICATION AND INTERPERSONAL SKILLS

Communication and interpersonal skills

I have very good communication skills developed by representing iWave Solutions in customer relationships and also by sustaining presentations in multiple International Conferences

JOB-RELATED SKILLS

Job-related skills

Experience with Machine Learning algorithms, Natural and Spoken Language Processing, Deep Learning architectures
Experience with multiple programming languages such as: Python, C++, C#, Java.
Previous experience with Pytorch, DyNET, Tensorflow, libRNN, AForge and Accord.NET
Some basic knowledge in: HTML, Javascript, JQuery, Bootstrap

OPEN SOURCE CONTRIBUTIONS

Open source contributions

NLP-Cube - multilingual raw text processing framework using deep-learning
- GIT: https://github.com/adobe/NLP-Cube

libLOL - security tool for detecting previously unseen living off the land commands
  • GIT: https://github.com/adobe/libLOL

TTS-Cube - end-to-end speech synthesis system using neural networks
  • GIT: https://github.com/tiberiu44/TTS-Cube
  • Web (audio samples): https://tiberiu44.github.io/TTS-Cube/

DyNET (individual contributions) - deep learning framework in C++ with JAVA and Python bindings - contributed to implementation of sparse LSTMs
  • GIT: https://github.com/clab/dynet