

## Education

### PhD student in Computer Science, Instituto Superior Técnico

Computer Science and Engineering Department; supervisor: Professor [André F. T. Martins](#)

Collaborative project [MAIA](#) (Multilingual AI Agent Assistants) in the Carnegie Mellon Portugal Program partnership involving Instituto de Telecomunicações, INESC-ID, Carnegie Mellon University and Unbabel.

*Research interests: uncertainty estimation, probabilistic deep learning, nlp evaluation and interpretability*

Lisbon, Portugal

2020 – present

### MSc in Data Science, HSE University

Computer Science Department; supervisor: Professor [Ekaterina Artemova](#)

Moscow, Russia

2018 – 2020

### BSc in Applied Mathematics and Computer Science, HSE University

Computer Science Department; supervisor: Professor [Ekaterina Artemova](#)

Moscow, Russia

2014 – 2018

## Experience

### Instituto Superior Técnico

PhD Researcher

- Uncertainty estimation for machine translation evaluation.

Lisbon, Portugal

Sep 2020 – present

### HSE University

Research Internship (Part-time)

- Question-answering for Russian language.

Moscow, Russia

Oct 2019 – Aug 2020

### Unbabel

Applied AI Research Summer Internship

- Language model-based data filtering and selection for machine translation.

Lisbon, Portugal

Jul 2019 – Oct 2019

### Digital Decisions

Data Scientist (Python)

- Hierarchical classification and routing of the customers messages.
- Named entity recognition for a speech assistant.

Moscow, Russia

Feb 2019 – Mar 2020

### Sberbank

Data Science Internship (Python)

- Time series analysis and prediction. Forecasting demand for cash at ATMs.

Moscow, Russia

Apr 2018 – Jun 2018

### Diginetica

Junior Data Scientist/Engineer (Python)

- Spell checker tool for search queries aimed to suggest corrections for misspelled words and offer synonyms.
- Snippets from product reviews: extraction of ultra-concise summaries of opinions over products.

Moscow, Russia

Jun 2017 – Dec 2017

### Yandex

Python Software Engineering Internship

- Analysis of the personalities in the flow of search queries.
- Marking and classification of the various object definitions, ranking them in order of relevance for users.

Moscow, Russia

Summer 2015

## Publications

\*equal contribution

### Uncertainty-Aware Machine Translation Evaluation

Taisiya Glushkova, Chrysoula Zerva, Ricardo Rei, André F. T. Martins

Presented at the *Sixth Conference on Machine Translation (WMT21) at EMNLP 2021*  
*EMNLP Findings 2021*

[[ACL Anthology](#), slides]

### IST-Unbabel 2021 Submission for the Quality Estimation Shared Task

Chrysoula Zerva, Daan van Stigt, Ricardo Rei, Ana C. Farinha, Pedro G. Ramos, José G. C. de Souza, Taisiya Glushkova, Miguel Vera, Fabio Kepler, André F. T. Martins

*EMNLP WMT 2021*

[[wmt21](#)]

## Are References Really Needed? Unbabel-IST 2021 Submission for the Metrics Shared Task

[wmt21]

Ricardo Rei, Ana C. Farinha, Chrysoula Zerva, Daan van Stigt, Craig Stewart, Pedro G. Ramos,

Taisiya Glushkova, André F. T. Martins, Alon Lavie

EMNLP WMT 2021

## DaNetQA: a yes/no Question Answering Dataset for the Russian Language

[arXiv]

Taisiya Glushkova\*, Alexey Machnev\*, Alena Fenogenova, Tatiana Shavrina, Ekaterina Artemova, Dmitry I. Ignatov

AIST 2020

## Char-RNN and Active Learning for Hashtag Segmentation

[arXiv]

Taisiya Glushkova, Ekaterina Artemova

CICLing 2019, published in Springer LNCS

## Skills

Programming languages:	Python, R, C++, HTML, CSS, $\LaTeX$
Tools:	Jupyter Notebook, Google Colab, git, bash
Main libraries:	numpy, pandas, matplotlib, scipy, scikit-learn, nltk, gensim, fasttext, keras, pytorch, allennlp, deeppavlov, transformers
Languages:	English(Advanced), Russian(Native), German(A2), Portuguese(A1)

## Teaching

### Instituto Superior Técnico

Teaching Assistant for Deep Learning course

Lisbon, Portugal  
Nov 2021 – present

### HSE University

Teaching Assistant for NLP course

Teaching Assistant for Python course

Moscow, Russia  
Sep 2018 – Dec 2019  
Sep 2016 – Sep 2018

## University Projects

<b>Graduation Thesis (MSc)</b> , DaNetQA: a dataset for binary question answering in Russian	2019 – 2020
<b>Coursework</b> , Detecting Uncertain Information in Natural Language Texts	2018 – 2019
<b>Graduation Thesis (BSc)</b> , Character Level Models for Hashtag Segmentation	2017 – 2018
<b>Coursework</b> , Collecting Dictionaries for Sentiment Analysis of Reviews	2016 – 2017
<b>Coursework</b> , Extracting Timestamps from News in English Language	2015 – 2016
<b>Real Time Tracker</b> , Desktop App for managing tasks and projects in real time (analogue of Trello) Front end: HTML, CSS, jQuery, AngularJS framework Back end: Python, Django, SQLite	2015 – 2016
<b>Redis-server</b> , Redis is an open source, in-memory data structure store, used as database, cache and message broker. My goal was to implement the analogue by using C++ and boost library and to practise writing Unit-tests.	2015

## Extracurricular Activities

Participated in LxMLS (8th Lisbon Machine Learning School)	Summer 2018
Freshmen student advisor: helped students to choose courseloads and adjust to university life	2015 – 2017
Organized university-wide events such as the Annual Open House	2015
Hobbies: watercolor, hiking	