

Curriculum vitae et studiorum of Natalia Kuzminykh born on May 17, 1994

EDUCATION

Title: PhD Computing

Released by: School of Engineering, Computing and Mathematics, University of Plymouth (UK)

Project: Embodying LLMs for Enhanced Environmental Interaction and Knowledge Application

Supervision: Dr. Stefano Nolfi, Dr. Amir Aly, Dr. Nicola Milano

Period of activity from: 01/10/2024 to Present

Title: Linguistics and Cognitive studies (Laurea Magistrale) | 110/110

Date: 13/06/2022

Released by: Department of Social, Political and Cognitive Studies, Università degli Studi di Siena (Italy)

Dissertation: Assessment of Complexity Metric in Russian Agreement Configurations

Supervision: Prof. Dr. Christiano Chesi & Prof. Dr. Luigi Rizzi

Period of activity from: 10/2018 to 06/2022

Title: BA in Linguistics. Interpreting and Translation Studies curriculum

Date: 05/07/2018

Released by: Institute of Language and Literature, Udmurt State University (Izhevsk, Russia)

Dissertation: Conveying the linguistic components of the promotional business films in translation

Supervision: Prof. Olga Kuzyaeva

Overseas mobility: Part of the educational program (14 c.f.) was mastered in the University of Helsinki (Finland)

Period of activity from: 09/2014 to 07/2018

OTHER CERTIFIED COURSES

Title: Artificial Intelligence Professional Program

Date: 19/04/2023

Released by: Stanford Online

Courses: XCS224N: NLP with Deep Learning, XCS221: Artificial Intelligence Principles and Techniques, XCS224U: Natural Language Understanding

Period of activity from: 09/2022 to 03/2023

Title: Experimental Methods in Language Acquisition Research XVI 2020

Date: 22/04/2020

Released by: Utrecht Institute of Linguistics (Netherlands)

Courses: ERP, Automatic Speech Recognition, PRAAT, CHILDES, Statistics with R etc.

Period of activity from: 15/04/2020 to 17/04/2020

Title: The PhD course SPRÅK8861: Methods of Linguistics. "Psycholinguistics Methods: Language Research from Conception to Publication"

Date: 30/10/2019

Released by: Norwegian University of Science and Technology (Norway)

Courses: Data Science, Introduction to R, Scientific Writing

Period of activity from: 14/10/2019 to 18/10/2019

RESEARCH PROJECTS

Title: Graduate Fellowship: *"Embodying large language models – Future AI Research"*

Supervisor: prof. Aldo Gangemi

Institution: *Institute of Cognitive Sciences and Technologies, CNR, Catania (Italy)*

Period of activity from: 01/02/2024 to -01/02/2025

Title: Research internship: *"Improving Complexity and Acceptability Prediction with Multi-task Learning on Self-Supervised Annotations"*

Supervisor: Research Director Christiano Chesi

Institution: *NETS laboratory on Neurolinguistics, Computational Linguistics and Theoretical Syntax, Scuola Superiore Studi Pavia IUSS, Pavia (Italy)*

Period of activity from: 15/03/2021 to 31/05/2021

Title: Research internship | Erasmus+ | *"Effect of Age and Language Phonological Complexity on Children's Vocalizations"*

Supervisor: Research Director Alejandrina Cristia

Institution: *Laboratoire de Sciences Cognitives et de Psycholinguistique, Ecole Normale Supérieure (ENS), EHESS, CNRS, PSL University, Paris, (France)*

Period of activity from: 01/03/2020 to 31/08/2020

ACADEMIC EXPERIENCE

Title: ESL Instructor

Course taught: English Language Course for Short-cycle Tertiary Education Program students

Institution: College of Law and Finance, Izhevsk, (Russia)

Period of activity from: 09/2017 to 04/2018

INDUSTRY EXPERIENCE

Title: Generative AI Engineer

Services Provided: - Backend development of custom AI solutions using GPT frameworks and executing end-to-end implementations in Python, Java; - Integration of prompt engineering techniques like Zero-shot, Few-shot Prompting, and Retrieval Augmented Generation; - Fine-tuning the performance and quality of generative AI models and assistants using metrics such as BLEU, ROUGE, perplexity, etc.; - Models optimization using best practices such as version control, unit testing, code review, etc.; - Deployment via RESTful APIs, Flask, Docker.

Organization: *Freelance (with VAT number) | Remote from Italy*

Period of activity from: 06/2023 to Present

Title: Conversational AI Engineer

Services Provided: - Enhanced multilingual AI system performance by updating data and developing a CoNLL-U to XML conversion script; - Conducted Intent Recognition research, analysing studies and benchmarking against solutions like RASA and Dialogflow; - Led evaluation of the in-house model's performance across various metrics, (incl. precision, recall, F1-score, and accuracy) on a custom dataset for thorough assessment and validation.

Organization: *QuestIT (Siena, Italy)*

Period of activity from: 10/2022 to 05/2023

CONFERENCE TALKS

Title: "What day-long audio-recordings can tell us about Children's Vocal development" [English]

Conference: Conference Linguistic Representations and Language Processing

Institution: University of Münster, Münster (Germany)

Period of activity: March 2021

Title: "Conveying the linguistic components of the promotional business films in translation" [Russian]

Conference: XLVI Students' Scientific Conference

Institution: Udmurt State University, Izhevsk, (Russia)

Period of activity: April 2018

PUBLICATIONS

Kuzminykh N. & Winder P. (2023) Part 5: How to Monitor a Large Language Model. Technical Report, <https://winder.ai/part-5-monitor-large-language-model/> [English]

Kuzminykh N. & Winder P. (2023) Part 4: How to Deploy a ChatGPT Model or LLM. Technical Report, <https://winder.ai/part-4-deploy-chatgpt-model-llm/> [English]

Kuzminykh N. & Winder P. (2023) Part 3: Training Custom ChatGPT and Large Language Models. Technical Report, <https://winder.ai/part-3-training-custom-chatgpt-large-language-models/> [English]

Kuzminykh N. & Winder P. (2023) Part 2: An Overview of LLM Development & Training ChatGPT. Technical Report, <https://winder.ai/part-2-overview-llm-development-training-chatgpt/> [English]

Kuzminykh N. & Winder P. (2023) Part 1: Introduction to Large Language Models and ChatGPT. Technical Report, <https://winder.ai/introduction-large-language-models-chatgpt/> [English]

Kuzminykh N. (2022). Assessment of Complexity Metric in Russian Agreement Configurations. Master Thesis, University of Siena, Italy [English]

Kuzminykh N. (2021). Effect of Age and Language Phonological Complexity on Children's Vocalizations. Poster, SRCO Biennial Meeting, Society for Research in Child Development, Minneapolis (USA) [English]

Kuzminykh N. (2018). Conveying the linguistic components of the promotional business films in translation. BA Thesis, Udmurt State University, (Russia) [Russian]

AWARDS & HONORS

2022 – 2023	<i>IKIGAI Training Grant – Monte dei Paschi Siena Foundation</i>
2018 – 2021	<i>Regional Scholarship – Tuscany Regional Agency (Italy)</i>
2020	<i>Erasmus+</i>
2017	<i>Scholarship for Overseas mobility</i>

ADDITIONAL INFORMATION

Languages	Russian (Native), English (Proficient Level), Italian (Intermediate Level), German (Elementary Level), French (Pre-Intermediate Level)
Research Skills	<p>Experience in end-to-end development of neural networks for NLP tasks (incl. LSTMs, RNNs, LLMs) such as word prediction, summarization, and NER.</p> <p>Skilled in semantic analysis using word vectorization methods (Word2Vec, SVD, GloVe).</p> <p>Experience in employing contextual LMs (Transformers, BERT, GPT) for semantic and syntactic analysis.</p> <p>Experience in utilizing dependency parsing to uncover word relationships within sentences.</p> <p>Experience in building and fine-tuning language and translation models for advanced word prediction.</p> <p>Competent in information retrieval using both traditional and neural network approaches.</p> <p>Experience in designing Web Experiments (Ibex Farm, Qualtrics)</p> <p>Experience in Linguistic Annotation (ELAN software & ACLEW DAS annotation template)</p> <p>Knowledge of Eye-Tracking Data Collection & Analysis (OpenSesame software)</p> <p>Experience in working with both Textual (CHILDES, WordSmith & NLTK library) and Audio (ELAN, Praat) Corpora</p> <p>Project Management, Strong focus on achieving goals, Planning skills, organizational skills, precision, good interpersonal skills and proactivity</p>
Computational Skills	<p>Coding languages: R, Python (incl. PyTorch & TensorFlow), C & Bash</p> <p>Experience with Typescript, React Native, SQL</p> <p>Experience in Research Reproducibility (via GitHub and OSF)</p> <p>Knowledge of Collaborative Tools (GitHub, Google Colab)</p> <p>Knowledge of DevOps tools (Docker)</p> <p>Knowledge of Cloud computing tools: Azure, Google Cloud</p>

The applicant is aware of the criminal sanctions to which she is liable for making false statements