

OKSANA DEREZA

@ oksana.dereza@gmail.com

+79032960371

Galway, Ireland

github.com/ancatmara

ORCID: 0000-0003-2896-6037

ResearcherID: Q-6140-2017

SPIN-RSCI: 4494-8088



EXPERIENCE

Mentor

Skillfactory

September 2019 – January 2020 Online

- Helping students with the 'Real Data Science' course
- Providing individual feedback on homeworks and overall performance
- Creating promo tutorials

Course author & Lecturer

National Research University 'Higher School of Economics'

September 2017 – August 2019 Moscow

- Python for Natural Language Processing
- An Introduction to Python (for students in humanities)
- Digital Literacy (for students in humanities)
- Machine Learning for Natural Language Processing
- Functional and Cognitive Models in Linguistics

Course author & Lecturer

Lomonosov Moscow State University

January 2018 – July 2019 Moscow

- Beginner Modern Irish
- Readings in Modern Irish Literature
- Early Modern Irish

Data analyst & Computational linguist

Kaspersky Lab

February 2015 – July 2018 Moscow

- Web content analysis
- Natural language processing

Translator

Clever Media Group

February 2015 – February 2017 Moscow

Translating French and English educational books for children into Russian

CERTIFICATES

- Certificate of Proficiency in English (2014)
- Introduction to Machine Learning (2016)
- Introduction to \LaTeX (2016)

TECHNICAL SKILLS

Python

LaTeX

R

Unix

NLProc

Machine Learning

Data Science

LANGUAGES

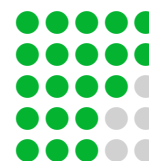
Russian

English

Irish

French

Danish



EDUCATION

PhD student

Insight Centre of Data Analytics, NUI Galway

September 2019 – now

Part of the CARDAMOM project
("Comparative Deep Models for Minority and Historical Languages")

MA in Computational Linguistics and Language Theory

NRU HSE

September 2015 – June 2017

Thesis title: "Morphological analysis of Old Irish data with neural networks"

Degree of Specialist in Philology. Teacher of the Irish Language and Foreign Literature

Lomonosov Moscow State University

September 2010 – June 2015

Specialist = BA + MA

Thesis title: "Semantics of physical qualities in Goidelic: 'heavy' and 'light'".

Student exchange

Trinity College Dublin, School of Irish

September 2014 – December 2014

Student exchange

University of Copenhagen, Department for Scandinavian Studies and Linguistics

February 2012 – June 2012

CONFERENCES

- The semantics of Proto-Celtic **k^wenno*: history and typology. // *First European Symposium in Celtic Studies*. Trier, 2013
- Семантика рефлексов протокельтской основы **k^wenno*: история и типология. // *Сектор анатолийских и кельтских языков РАН*. Москва, 2014
- Семантика протокельтского **k^wenno*: история и типология. // *XXI Международная научная конференция студентов, аспирантов и молодых ученых «Ломоносов»*. Москва, 2014
- Semantic differences, meaning and distribution of the word *ceann* in Irish and Scottish Gaelic. // *Rannsachadh na Gàidhlig*. Edinburgh, 2014.
- The polysemy of TRWM 'heavy' in Middle Welsh (with Dr. Elena Parina). // *Language Diversity in Wales*. Aberystwyth, 2014.
- The semantics of **k^wenno*- reflexes: history and typology (poster). // *Indian Summer School in Typology, HSE*. Moscow, 2014.
- Polysemous adjectives in Goidelic languages: 'heavy' and 'light'. // *The 7th International Colloquium of Societas Celto-Slavica*. Bangor, 2014.
- Корпусные исследования семантики и сочетаемости слов на материале гойдельских языков. // *ConCorT Junior: Корпусные технологии и компьютерные методы в современной гуманитарной науке*. Москва, 2015.
- A complex approach to spellchecking and autocorrection for Russian (with Alyona Fenogenova, Anna Marakasova & Dmitri Kayutenko). // *XXII International conference on Computational Linguistics and Intellectual Technologies 'DIALOGUE'*. Moscow, 2016.
- Automatic morphological analysis for Russian: a comparative study (poster, with Alyona Fenogenova & Dmitri Kayutenko). // *XXII International conference on Computational Linguistics and Intellectual Technologies 'DIALOGUE'*. Moscow, 2016.
- A homonymy-oriented metric for morphological analyzers (poster, with Alyona Fenogenova & Dmitri Kayutenko). // *Computational Linguistics and Language Science Workshop*. Moscow, 2016.
- Building a dictionary-based lemmatizer for Old Irish. // *The 2nd Celtic Language Technology Workshop, JEP-TALN-RECITAL*. Paris, 2016.
- Verb-noun collocation and government model extraction from large corpora (poster, with Vladislav Tushkanov). // *The 3^d International Workshop on Concept Discovery in Unstructured Data*. Moscow, 2016.
- Intratextual relations in the Ulster cycle. // *The 8th International Colloquium of Societas Celto-Slavica*. Heidelberg, 2016.
- Создание сети героев Уладского цикла древнеирландских саг на основе неаннотированного корпуса. // *ConCorT: Корпусные технологии, Digital Humanities и современное знание*. Нижний Новгород, 2016.
- Создание визуального словаря коллокаций на основе размеченного корпуса (с Владиславом Тушкановым). // *ConCorT: Корпусные технологии, Digital Humanities и современное знание*. Нижний Новгород, 2016.
- A Lexicon-Based Lemmatizer for Old Irish (poster). // *AINL FRUCT: Artificial Intelligence and Natural Language Conference*. Saint-Petersburg, 2016.
- Lemmatizing Old Irish: Lexicons, Rules and Neural Networks. // *Spring Symposium on Computational and Corpus-based Linguistics (organized by Chronologicon Hibernicum)*. Maynooth, 2017.
- Lemmatisation for under-resourced languages with sequence-to-sequence learning: A case of Early Irish // *The 3d Kolmogorov Workshop "Computational linguistics and language science"*. Moscow, 2018.

PUBLICATIONS

- Paired adjectives in the Middle Welsh *Ystorya Bown o Hamtwn*. // *Dafydd Johnston, Elena Parina and Maxim Fomin (eds.) 'Ynllaweniawn, ynllawniaith': Proceedings of the 6th International Colloquium of Societas Celto-Slavica. Studia Celto-Slavica 7*. Aberystwyth: University of Wales Centre for Advanced Welsh and Celtic Studies, 2015, p. 281 – 282.
- Стилистика средневековых переводов: ирландская и валлийская версии жести о Беде из Амтона. // *Проконьев, А. Ю. (ред). Проблемы истории и культуры средневекового общества: материалы XXXIII всероссийской конференции студентов, аспирантов и молодых ученых «Курбатовские чтения» (26–29 ноября 2013 года)*. Санкт-Петербург: Свое издательство, 2015, с. 396 – 402.
- Королевская власть в Уэльсе в XI–XII вв.: между героическим прошлым и туманным будущим. // *Левин, Ф. (ред.) Структуры власти в кельтских обществах. VOX MEDII AEVI №1 (12)*, Санкт-Петербург, 2015, с. 15 – 23
- «Основы знаний об Ирландии» Джеффри Китинга и король с говорящим именем. // *Левин, Ф. (ред.) Структуры власти в кельтских обществах. VOX MEDII AEVI №1 (12)*, Санкт-Петербург, 2015, с. 84 – 92
- Корпусные исследования семантики качественных прилагательных на материале гойдельских языков. // *Focal, fíos agus foghlaim: Тезисы международной конференции, посвященной отечественной школе кельтологии и юбилею д.ф.н., профессора Татьяны Андреевны Михайловой*. Москва, 2016, с. 21 – 23
- Семантика качественных прилагательных в гойдельских языках: 'тяжелый' и 'легкий'. // *Сборник научно-исследовательских работ студентов-победителей Конкурса НИРС 2015 г., М. : Изд. дом Высшей школы экономики, 2016, с. 210 – 225*
- Building a Dictionary-Based Lemmatizer for Old Irish // *Actes de la conférence conjointe JEP-TALN-RECITAL 2016, volume 6: CLTW*. Paris, 2016, p. 12 – 17
- Physical Qualities in Goidelic: A Corpus Study of Polysemy and Collocability. // *Proceedings of the Seventh International Colloquium of Societas Celto-Slavica, held at Bangor, Wales, 4 – 7 September 2014*, 2018, p. 85 – 99

CONFERENCES

- A Searchable Database of Early Irish Texts. // *The 9th International Colloquium of Societas Celto-Slavica*. Quimper, 2018.
- A searchable database of Early Irish texts // *32nd Irish Conference of Medievalists*. Cork, 2018.
- Lemmatization for ancient languages: rules or neural networks? // *Artificial Intelligence and Natural Language Conference (AINL)*. St. Petersburg, 2018.

PUBLICATIONS

- Lemmatisation for under-resourced languages with sequence-to-sequence learning: A case of Early Irish // *EPiC Series in Language and Linguistics*, 2019
- Lemmatization for Ancient Languages: Rules or Neural Networks? // *Series "Communications in Computer and Information Science"*, Springer, 2018