

OKSANA DEREZA

@ oksana.dereza@gmail.com

+353899404168

Gaillimh, Éire

github.com/ancatmara



Seanteach an Sagairt, Príomhsráid, An Cheathrú Rua, Co. na Gaillimhe, H91 FXR7

TAITHÍ OIBRE

Cúntóir Teagaisc

Ollscoil na Gaillimhe

Meán Fómhair 2022 – Nollaig 2022 Gaillimh

- Saotharlanna “Bunchúrsa Phróiseáil Teanga Nádúrtha” a mhúineadh
- Bileoga oibre agus tascanna a chruthú agus a ghrádú

Cúntóir Taighde

An Coláiste Ollscoile, Baile Átha Cliath

January 2021 – Bealtaine 2021 Ar líne

- Taighde a dhéanamh ar mhodhanna daonnachtaí digiteacha agus ar bhogearraí le haghaidh staidéir dlí meánaoiseach
- Múnla sonraí nasctha a fhréamhshamlú ar réimsí dlí na Meánaoise
- Cabhrú le togra deontais a scríobh don **tionscadal FLEXI** (bronnadh deontais IRC agus ERC dó)

Meantóir

Skillfactory

Meán Fómhair 2019 – Eanáir 2020 Ar líne

- Meantóireacht ar mhic léinn an chúrsa ‘Eolaíocht Sonraí Réadúil’
- Ranganna teagaisc bolscaireachta ar Python a chruthú

Léachtóir

Ollsoil Náisiúnta Taighde ‘Ard-Scoil Eacnamaíochta’

Meán Fómhair 2017 – Lúnasa 2019 Moscó

- Python le haghaidh Próiseála Teanga Nádúrtha
- Bunchúrsa Python (do mhic léinn na daonnachtaí)
- Inniúlacht Digiteach (do mhic léinn na daonnachtaí)
- Meaisínfhoghlaím le haghaidh Próiseála Teanga Nádúrtha

Léachtóir

Stát-Ollscoil Mhoscó

Eanáir 2018 – Iúil 2019 Moscó

- Bunchúrsa Gaeilge
- Léitheoireacht ar Litríocht Nua-Ghaeilge
- An Nua-Ghaeilge Luath

Anailísí sonraí & Ríomhtheangeolaí

Kaspersky Lab

Feabhra 2015 – Iúil 2018 Moscó

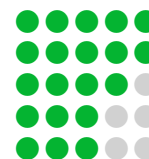
- Measúnú agus ceartú ar thorthaí uathrangaithe leathanach gréasáin
- Tascanna PTN: anailís tuairimí, rangú aiseolais, rangú ceisteanna gréasáin, scagadh ábhair, eastóscadh eochairfhocail

SCILEANNA TEICNIÚLA

Python Próiseáil Teanga Nádúrtha
LaTeX Unix Meaisínfhoghlaím
Eolaíocht Sonraí

SCILEANNA TEANGA

Rúisis – cainteoir dúchais
Béarla – C2
Gaeilge
Fraincis
Danmhairgis



OIDEACHAS

Iarrthóir PhD

Insight, Ionad na hAnailísíochta Sonraí, Ollscoil na Gaillimhe

2019 – anois Éire

Taighdeoir ar **thionscadal CARDAMOM** (“Comparative Deep Models for Minority and Historical Languages”)

Máistreacht sa Ríomhtheangeolaíocht agus Teoiric Teanga

Ollsoil Náisiúnta Taighde ‘Ard-Scoil Eacnamaíochta’

2015 – 2017 An Rúis

Tráchtas: “Anailís mhoirfeolaíoch ar shonraí Sean-Ghaeilge le líonraí néaracha”

Céim speisialtóra san Fhileolaíocht

Stát-Ollscoil Mhoscó

2010 – 2015 An Rúis

Speisialtóir = BA + MA
Tráchtas: “Séimeantaic cáilíochtaí fisiciúla sna teangacha Gaelacha: ‘trom’ agus ‘éadrom’”.

Mac léinn malairte

Coláiste na Tríonóide BÁC, Roinn na Gaeilge

2014 Éire

Mac léinn malairte

Ollscoil Chóbanhávan, Roinn na Léinn & Teangeolaíochta Lochlannach

2012 An Danmhairg

An Dr. John P. McCrae

@ john.mccrae@nuigalway.ie

An tAonad um Shonraí Teangeolaíochta
Insight, Ionad na hAnailísíochta Sonraí
Pairc Ghnó na Gaillimhe, An Daingean
Gaillimh, H91 AEX4

An Dr. Fangzhe Qiu

@ fangzhe.qiu@ucd.ie

Scoil na Gaeilge, an Léinn Cheiltigh agus
an Bhéaloidis, Áras Newman
An Coláiste Ollscoile, Baile Átha Cliath
Belfield, BÁC 4

An Dr. Victor Bayda

+353872153671

Comhchoiste Ghaeltacht Uíbh Ráthaigh
Ceann Eich, Bóthar Coiredín, Cathair
Saidhbhín, Co. Chiarraí, V23 R894

COMHHDHÁLA

- The semantics of Proto-Celtic **k^wenno*: history and typology. // *First European Symposium in Celtic Studies*. Trier, 2013
- Semantic differences, meaning and distribution of the word *ceann* in Irish and Scottish Gaelic. // *Rannasachadh 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.
- 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.
- A Lexicon-Based Lemmatizer for Old Irish (poster). // *AINL FRUCT: Artificial Intelligence and Natural Language Conference*. St. 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.
- 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.
- Automatic Dating of Medieval Irish Texts with Cluster Analysis // *International Medieval Congress*. Leeds, 2019.
- Machine-Assisted Dating of Medieval Irish Texts with Cluster Analysis // *The XVIth International Congress Of Celtic Studies*. Bangor, 2019.

FOILSEACHÁIN

- 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.
- 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
- Lemmatization for Ancient Languages: Rules or Neural Networks? // *Series "Communications in Computer and Information Science"*, Springer, 2018
- Lemmatisation for Under-Resourced Languages with Sequence-to-Sequence Learning: A Case of Early Irish // *EPiC Series in Language and Linguistics*, 2019
- A Multilingual Evaluation Dataset for Monolingual Word Sense Alignment // (with S. Ahmadi, J. McCrae, S. Nimb, F. Khan, M. Monachini, B. Pedersen, T. Declerck, T. Wissik, A. Bellandi, I. Pisani, T. Troelsgard, S. Olsen, S. Krek, V. Lipp, T. Váradi, L. Simon, A. Gyorffy, C. Tiberius, T. Schoonheim, Y. Ben Moshe, M. Rudich, R. Abu Ahmad, D. Lonke, K. Kovalenko, M. Langemets, J. Kallas, T. Fransen, D. Cillessen, D. Lindemann, M. Alonso, A. de Castro Salgado, J. L. Sancho, R. Ureña-Ruiz, J. Porta Zamorano, K. Simov, P. Osenova, Z. Kancheva, I. Radev, R. Stankovic, A. Perdih, D. Gabrovsek) // *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC)*, 2020