

Sebastian Ruder

ruder.io • sebastian@ruder.io
Google, Tucholskystraße 2, 10117 Berlin, Germany

Education

- **National University of Ireland** **Galway, Ireland**
College of Engineering and Informatics, Ph.D. Natural Language Processing 10/2015 – 02/2019
 - **University of Copenhagen** **Copenhagen, Denmark**
Research visit, Natural Language Processing Group, Computer Science Department 04/2017 – 06/2017
 - **Trinity College** **Dublin, Ireland**
Semester abroad, School of Computer Science and Statistics, Computer Science and Language 09/2014 – 01/2015
 - **Ruprecht-Karls-Universität Heidelberg** **Heidelberg, Germany**
Institute of Computational Linguistics, B.A. Computational Linguistics, English Linguistics 10/2012 – 09/2015
-

Experience

- **Google** **Germany, Berlin**
Research Scientist 11/2021 – Present
 - **DeepMind** **London, United Kingdom**
Research Scientist 03/2019 – 10/2021
 - **AYLIEN** **Dublin, Ireland**
Research Scientist 10/2015 – 02/2019
 - **IBM** **Munich, Germany**
Extreme Blue Intern, Watson 08/2015 – 09/2015
 - **Microsoft** **Dublin, Ireland**
Linguistic Engineering Intern, Xbox 02/2015 – 06/2015
 - **The OpenCog Foundation** **opencog.org**
Google Summer of Code Intern Summer 2014
 - **Lingenio GmbH** **Heidelberg, Germany**
Software Engineering Intern Spring 2014
 - **SAP** **Walldorf, Germany**
Working Student, Development University 02/2013 – 02/2014
 - **TEMIS** **Heidelberg, Germany**
Freelancing Developer 02/2013 – 10/2013
-

Books

1. Anders Søgaard, Ivan Vulić, **Sebastian Ruder**, Manaal Faruqui (2019). Cross-lingual Word Embeddings. *Morgan & Claypool Publishers*.

PhD thesis

2. **Sebastian Ruder** (2019). Neural Transfer Learning for Natural Language Processing. *National University of Ireland, Galway*.

Refereed conference publications

3. Alexis Conneau, Ankur Bapna, Yu Zhang, Min Ma, Patrick von Platen, Anton Lozhkov, Colin Cherry, Ye Jia, Clara Rivera, Mihir Kale, Daan Van Esch, Vera Axelrod, Simran Khanuja, Jonathan H Clark, Orhan Firat, **Sebastian Ruder**, Jason Riesa, Melvin Johnson (2022). XTREME-S: Evaluating Cross-lingual Speech Representations. In *Proceedings of Interspeech 2022*.
4. Xinyi Wang, **Sebastian Ruder**, Graham Neubig (2022). Expanding Pretrained Models to Thousands More Languages via Lexicon-based Adaptation. In *Proceedings of ACL 2022*.
5. Alham Fikri Aji, Genta Indra Winata, Fajri Koto, Samuel Cahyawijaya, Ade Romadhony, Rahmad Mahendra, Kemal Kurniawan, David Moeljadi, Radityo Eko Prasojo, Timothy Baldwin, Jey Han Lau, **Sebastian Ruder** (2022). One Country, 700+ Languages: NLP Challenges for Underrepresented Languages and Dialects in Indonesia. In *Proceedings of ACL 2022*.
6. Michael T€nzer, **Sebastian Ruder**, Marek Rei (2022). Memorisation versus Generalisation in Pre-trained Language Models. In *Proceedings of ACL 2022*.
7. **Sebastian Ruder***, Ivan Vulić*, Anders S€gaard* (2022). Square One Bias in NLP: Towards a Multi-Dimensional Exploration of the Research Manifold. In *Findings of ACL 2022*.
8. Daan van Esch, Tamar Lucassen, **Sebastian Ruder**, Isaac Caswell, Clara E. Rivera (2022). Writing system and speaker metadata for 2,800+ language varieties. In *Proceedings of LREC 2022*.
9. Shamsuddeen Hassan Muhammad, David Ifeoluwa Adelani, **Sebastian Ruder**, Ibrahim Said Ahmad, Idris Abdulmumin, Bello Shehu Bello, Monojit Choudhury, Chris Chinenye Emezue, Anuoluwapo Aremu, Saheed Abdul, Pavel Brazdil (2022). Naijasenti: A Nigerian Twitter Sentiment Corpus for Multilingual Sentiment Analysis. In *Proceedings of LREC 2022*.
10. Yi Tay*, Vinh Q. Tran*, **Sebastian Ruder**, Jai Gupta, Hyung Won Chung, Dara Bahri, Zhen Qin, Simon Baumgartner, Cong Yu, Donald Metzler (2021). Charformer: Fast Character Transformers via Gradient-based Subword Tokenization. In *Proceedings of ICLR 2022*.
11. Vamsi Aribandi, Yi Tay, Tal Schuster, Jinfeng Rao, Huaixiu Steven Zheng, Sanket Vaibhav Mehta, Honglei Zhuang, Vinh Q Tran, Dara Bahri, Jianmo Ni, Jai Gupta, Kai Hui, **Sebastian Ruder**, Donald Metzler (2022). ExT5: Towards Extreme Multi-Task Scaling for Transfer Learning. In *Proceedings of ICLR 2022*.
12. Rabeeh Karimi Mahabadi, James Henderson, **Sebastian Ruder** (2021). Compacter: Efficient Low-Rank Hypercomplex Adapter Layers. In *Proceedings of NeurIPS 2021*.
13. Angeliki Lazaridou, Adhi Kuncoro, Elena Gribovskaya, Devang Agrawal, Adam Liska, Tayfun Terzi, Mai Gimenez, Cyprien de Masson d’Autume, Tomas Kocisky, **Sebastian Ruder**, Dani Yogatama, Kris Cao, Susannah Young, Phil Blunsom (2021). Mind the Gap: Assessing Temporal Generalization in Neural Language Models. In *Proceedings of NeurIPS 2021*.
14. Stefan Daniel Dumitrescu, Petru Rebeja, Beata Lorincz, Mihaela Gaman, Andrei Avram, Mihai Ilie, Andrei Pruteanu, Adriana Stan, Lorena Rosia, Cristina Iacobescu, Luciana Morogan, George Dima, Gabriel Marchidan, Traian Rebedea, Madalina Chitez, Dani Yogatama, **Sebastian Ruder**, Radu Tudor Ionescu, Razvan Pascanu, Viorica Patraucean (2021). LiRo: Benchmark and leaderboard for Romanian language tasks. In *Proceedings of NeurIPS 2021 Datasets and Benchmarks Track*.
15. **Sebastian Ruder**, Noah Constant, Jan Botha, Aditya Siddhant, Orhan Firat, Jinlan Fu, Pengfei Liu, Junjie Hu, Graham Neubig, Melvin Johnson (2021). XTREME-R: Towards More Challenging and Nuanced Multilingual Evaluation. In *Proceedings of EMNLP 2021*.
16. Jonas Pfeiffer, Ivan Vulić, Iryna Gurevych, **Sebastian Ruder** (2021). UNKS Everywhere: Adapting Multilingual Language Models to New Scripts. In *Proceedings of EMNLP 2021*.
17. Samuel Cahyawijaya, Genta Indra Winata, Bryan Wilie, Karissa Vincentio, Xiaohong Li, Adhiguna Kuncoro, **Sebastian Ruder**, Zhi Yuan Lim, Syafri Bahar, Masayu Leylia Khodra, Ayu Purwarianti, Pascale Fung (2021). IndoNLG: Benchmark and Resources for Evaluating Indonesian Natural Language Generation. In *Proceedings of EMNLP 2021*.
18. Alan Ansell, Edoardo Maria Ponti, Jonas Pfeiffer, **Sebastian Ruder**, Goran Glavaš, Ivan Vulić, Anna Korhonen (2021). MAD-G: Multilingual Adapter Generation for Efficient Cross-Lingual Transfer. In *Findings of EMNLP 2021*.

19. Xinyi Wang, Yulia Tsvetkov, **Sebastian Ruder**, Graham Neubig (2021). Efficient Test Time Adapter Ensembling for Low-resource Language Varieties. In *Findings of EMNLP 2021*.
20. Rabeeh Karimi Mahabadi, **Sebastian Ruder**, Mostafa Dehghani, James Henderson (2021). Parameter-efficient Multi-task Fine-tuning for Transformers via Shared Hypernetworks. In *Proceedings of ACL 2021*.
21. Phillip Rust, Jonas Pfeiffer, Ivan Vulić, **Sebastian Ruder**, Iryna Gurevych (2021). How Good is Your Tokenizer? On the Monolingual Performance of Multilingual Language Models. In *Proceedings of ACL 2021*.
22. Xinyi Wang, **Sebastian Ruder**, Graham Neubig (2021). Multi-view Subword Regularization. In *Proceedings of NAACL 2021*.
23. Hyung Won Chung, Thibault Févry, Henry Tsai, Melvin Johnson, **Sebastian Ruder** (2021). Rethinking embedding coupling in pre-trained language models. In *Proceedings of ICLR 2021*.
24. Yi Tay, Mostafa Dehghani, Samira Abnar, Yikang Shen, Dara Bahri, Philip Pham, Jinfeng Rao, Liu Yang, **Sebastian Ruder**, Donald Metzler (2021). Long Range Arena: A Benchmark for Efficient Transformers. In *Proceedings of ICLR 2021*.
25. Nicolas Garneau, Mareike Hartmann, Anders Sandholm, **Sebastian Ruder**, Ivan Vulić, Anders Søgaard (2021). Analogy Training Multilingual Encoders. In *Proceedings of AAAI 2021*.
26. Paula Czarowska, **Sebastian Ruder**, Ryan Cotterell, Ann Copestake (2020). Morphologically Aware Word-Level Translation. In *Proceedings of COLING 2020*.
27. Jonas Pfeiffer, Andreas Rücklé, Clifton Poth, Aishwarya Kamath, Ivan Vulić, **Sebastian Ruder**, Kyunghyun Cho, Iryna Gurevych (2020). AdapterHub: A Framework for Adapting Transformers. In *Proceedings of EMNLP 2020: System Demonstrations*.
28. Jonas Pfeiffer, Ivan Vulić, Iryna Gurevych, **Sebastian Ruder** (2020). MAD-X: An Adapter-based Framework for Multi-task Cross-lingual Transfer. In *Proceedings of EMNLP 2020*.
29. Marcin Kardas, Piotr Czapla, Pontus Stenetorp, **Sebastian Ruder**, Sebastian Riedel, Ross Taylor, Robert Stojnic (2020). AxCell: Automatic Extraction of Results from Machine Learning Papers. In *Proceedings of EMNLP 2020*.
30. Ivan Vulić*, **Sebastian Ruder***, Anders Søgaard* (2020). Are All Good Word Vector Spaces Isomorphic? In *Proceedings of EMNLP 2020*.
31. Junjie Hu*, **Sebastian Ruder***, Aditya Siddhant, Graham Neubig, Orhan Firat, Melvin Johnson (2020). XTREME: A Massively Multilingual Multi-task Benchmark for Evaluating Cross-lingual Generalization. In *Proceedings of ICML 2020*.
32. Mikel Artetxe*, **Sebastian Ruder***, Dani Yogatama, Gorka Labaka, Eneko Agirre (2020). A Call for More Rigor in Unsupervised Cross-lingual Learning. In *Proceedings of ACL 2020*.
33. Mikel Artetxe, **Sebastian Ruder**, Dani Yogatama (2020). On the Cross-lingual Transferability of Monolingual Representations. In *Proceedings of ACL 2020*.
34. Cyprien de Masson d'Autume, **Sebastian Ruder**, Lingpeng Kong, Dani Yogatama (2019). Episodic Memory in Lifelong Language Learning. In *Proceedings of NeurIPS 2019*.
35. Julian Eisenschlos*, **Sebastian Ruder***, Piotr Czapla*, Marcin Kadras*, Sylvain Gugger, Jeremy Howard (2019). MultiFiT: Efficient Multi-lingual Language Model Fine-tuning. In *Proceedings of EMNLP 2019*.
36. Paula Czarowska, **Sebastian Ruder**, Edouard Grave, Ryan Cotterell, Ann Copestake (2019). Don't Forget the Long Tail! A Comprehensive Analysis of Morphological Generalization in Bilingual Lexicon Induction. In *Proceedings of EMNLP 2019*.
37. Goran Glavaš, Robert Litschko, **Sebastian Ruder**, Ivan Vulić (2019). How to (Properly) Evaluate Cross-Lingual Word Embeddings: On Strong Baselines, Comparative Analyses, and Some Misconceptions. In *Proceedings of ACL 2019, Florence, Italy*.
38. **Sebastian Ruder**, Joachim Bingel, Isabelle Augenstein, Anders Søgaard (2019). Latent Multi-task Architecture Learning. In *Proceedings of AAAI 2019, Honolulu, Hawaii*.
39. Victor Sanh, Thomas Wolf, **Sebastian Ruder** (2019). A Hierarchical Multi-task Approach for Learning Embeddings from Semantic Tasks. In *Proceedings of AAAI 2019, Honolulu, Hawaii*.
40. **Sebastian Ruder***, Ryan Cotterell*, Yova Kementchedjheva, Anders Søgaard (2019). A Discriminative Latent-Variable Model for Bilingual Lexicon Induction. In *Proceedings of EMNLP 2018, Brussels, Belgium*.

41. Yova Kementchedjheva, **Sebastian Ruder**, Ryan Cotterell, Anders Søgaard (2018). Generalizing Procrustes Analysis for Better Bilingual Dictionary Induction. In *Proceedings of CoNLL 2018*, Brussels, Belgium.
42. **Sebastian Ruder**, Barbara Plank (2018). Strong Baselines for Neural Semi-supervised Learning under Domain Shift. In *Proceedings of ACL 2018*, Melbourne, Australia.
43. Jeremy Howard*, **Sebastian Ruder*** (2018). Universal Language Model Fine-tuning for Text Classification. In *Proceedings of ACL 2018*, Melbourne, Australia.
44. Anders Søgaard, **Sebastian Ruder**, Ivan Vulić (2018). On the Limitations of Unsupervised Bilingual Dictionary Induction. In *Proceedings of ACL 2018*, Melbourne, Australia.
45. **Sebastian Ruder**, John Glover, Afshin Mehrabani, Parsa Ghaffari (2018). 360° Stance Detection. In *Proceedings of NAACL-HLT 2018: System Demonstrations*, New Orleans, US.
46. Isabelle Augenstein*, **Sebastian Ruder***, Anders Søgaard (2018). Multi-task Learning of Pairwise Sequence Classification Tasks Over Disparate Label Spaces. In *Proceedings of NAACL-HLT 2018*, New Orleans, US.
47. **Sebastian Ruder**, Barbara Plank (2017). Learning to select data for transfer learning with Bayesian Optimization. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Copenhagen, Denmark.
48. **Sebastian Ruder**, Parsa Ghaffari, John G. Breslin (2016). A Hierarchical Model of Reviews for Aspect-based Sentiment Analysis. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 999–1005, Austin, Texas, US.

Journal publications

49. **Sebastian Ruder**, Ivan Vulić, Anders Søgaard (2019). A Survey Of Cross-lingual Word Embedding Models. *Journal of Artificial Intelligence Research*.

Tutorials

50. **Sebastian Ruder**, Avirup Sil (2021). Multi-Domain Multilingual Question Answering. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: Tutorial Abstracts*, pages 17–21.
51. **Sebastian Ruder**, Anders Søgaard, Ivan Vulić (2019). Unsupervised Cross-Lingual Representation Learning. In *Proceedings of ACL 2019: Tutorial Abstracts*, pages 31–38.
52. **Sebastian Ruder**, Matthew E. Peters, Swabha Swayamdipta, Thomas Wolf (2019). Transfer Learning in Natural Language Processing. In *Proceedings of the NAACL 2019: Tutorials*, pages 15–18.

Refereed workshop publications

53. Aida Nematzadeh*, **Sebastian Ruder***, Dani Yogatama* (2020). On Memory in Human and Artificial Language Processing Systems. In *Proceedings of the Bridging AI and Cognitive Science Workshop at ICLR 2020*.
54. Matthew Peters*, **Sebastian Ruder***, Noah A. Smith (2019). To Tune or Not to Tune? Adapting Pretrained Representations to Diverse Tasks. In *Proceedings of RepL4NLP at ACL 2019*, Florence, Italy.
55. **Sebastian Ruder**, Parsa Ghaffari, John G. Breslin (2016). Towards a continuous modeling of natural language domains. In *Proceedings of EMNLP 2016 Workshop on Uphill Battles in Language Processing: Scaling Early Achievements to Robust Methods*, pages 53–57, Austin, Texas, US.
56. Ian D. Wood and **Sebastian Ruder** (2016). Emoji as emotion tags for tweets. In *Emotion and Sentiment Analysis Workshop, LREC*, Portorož, Slovenia.
57. **Sebastian Ruder**, Parsa Ghaffari, John G. Breslin (2016). INSIGHT-1 at SemEval-2016 Task 4: Convolutional Neural Networks for Sentiment Classification and Quantification. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016)*, San Diego, US.
58. **Sebastian Ruder**, Parsa Ghaffari, John G. Breslin (2016). INSIGHT-1 at SemEval-2016 Task 5: Convolutional Neural Networks for Multilingual Aspect-based Sentiment Analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016)*, San Diego, US.

*Equal contribution.

Preprints

59. **Sebastian Ruder** (2016). An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747.
-

Certificates and awards

- CogX Outstanding Contribution in AI – PhD thesis 06/2019
 - Google Developer Expert – Machine Learning 12/2017 – 02/2019
 - Scholarship of the Irish Research Council 10/2015 – 02/2019
 - Scholarship of the *Cusanuswerk*, one of the 13 German sponsorship organizations 04/2014 – 09/2015
 - Microsoft Certified Professional (Programming in C#) 06/2015
 - Best Delegate award in various Model United Nations conferences 11/2012 – 01/2014
 - Second and third prizes *Bundeswettbewerb Fremdsprachen*, national foreign languages competition 2007 – 2008
 - First and second prizes *Landeswettbewerb Mathematik*, state mathematics competition 2006 – 2008
-

Languages and Technologies

Programming Languages: Python, Java, C#, R, C, \LaTeX , Prolog, JavaScript, SPARQL

Technologies: SciPy, NumPy, Keras, TensorFlow, DyNet, scikit-learn, NLTK, CoreNLP, MALLET, Weka, UNIX, Git

Natural Languages: Fluent in German and English, advanced in French and Spanish, beginner in Portuguese and Latin

Open Source Contributions: The OpenCog Foundation

Service and professional activities

Organizing

- **Conferences:** Program chair at EurNLP 2019¹
- **Tutorials:** Transfer Learning in NLP at NAACL-HLT 2019; Unsupervised Cross-lingual Representation Learning at ACL 2019
- **Workshops:** 4th Workshop on Representation Learning for NLP (RepL4NLP) at ACL 2019
- **Sessions:** NLP Session at Deep Learning Indaba 2018², 2019
- **Meetups:** NLP Dublin Meetup³

Conference chairing

- **Senior area chair:** COLING 2020; ACL 2021
 - **Area chair:** ACL 2020; EMNLP 2020; EACL 2021
 - **Publication chair:** CoNLL 2019
 - **Publicity chair:** EMNLP 2019
-

¹<https://www.eurnlp.org/>

²<http://www.deeplearningindaba.com/schedule-2018.html>

³<https://www.meetup.com/NLP-Dublin/>

Reviewing

- **Journals:** Transactions of the Association for Computational Linguistics (TACL); Transactions on Audio, Speech and Language Processing; Artificial Intelligence; IEEE Computational Intelligence Magazine; Natural Language Engineering; Journal of Artificial Intelligence Research; Transactions on Pattern Analysis and Machine Intelligence
 - **Conferences:** ACL Rolling Review; ACL 2018, 2019, 2022; ICML 2019, 2020; NeurIPS 2019, 2020; NAACL-HLT 2019; ICLR 2019, 2020; EMNLP 2018, 2019, 2020, 2021, 2022; CoNLL 2018; AACL 2020
 - **Workshops:** RepL4NLP 2018, DeepLo 2018, SemEval-2016 Task 5, ACL SRW 2020
-

Talks

Talks are available at <http://ruder.io/talks/>.