

# Sebastian Ruder

runder.io • sebastian@runder.io

DeepMind c/o Google, 6 St Pancras Square, Kings Cross, N2C 4AG • London, UK

---

## 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

- **DeepMind** **London, United Kingdom**  
*Research Scientist* 03/2019 – Present
  - **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. Cyprien de Masson d'Autume, **Sebastian Ruder**, Lingpeng Kong, Dani Yogatama. Episodic Memory in Life-long Language Learning. To be published in *Proceedings of NeurIPS 2019*.
4. Julian Eisenschlos\*, **Sebastian Ruder\***, Piotr Czapla\*, Marcin Kadras\*, Sylvain Gugger and Jeremy Howard. MultiFiT: Efficient Multi-lingual Language Model Fine-tuning. In *Proceedings of EMNLP 2019*.

5. Paula Czarnowska, **Sebastian Ruder**, Edouard Grave, Ryan Cotterell and Ann Copestake (2019). Don't Forget the Long Tail! A Comprehensive Analysis of Morphological Generalization in Bilingual Lexicon Induction. In *Proceedings of EMNLP 2019*.
6. 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.
7. **Sebastian Ruder**, Joachim Bingel, Isabelle Augenstein, Anders Søgaard (2019). Latent Multi-task Architecture Learning. In *Proceedings of AAAI 2019*, Honolulu, Hawaii.
8. 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.
9. **Sebastian Ruder\***, Ryan Cotterell\*, Yova Kementchedjhieva, Anders Søgaard (2019). A Discriminative Latent-Variable Model for Bilingual Lexicon Induction. In *Proceedings of EMNLP 2018*, Brussels, Belgium.
10. Yova Kementchedjhieva, **Sebastian Ruder**, Ryan Cotterell, Anders Søgaard (2018). Generalizing Procrustes Analysis for Better Bilingual Dictionary Induction. In *Proceedings of CoNLL 2018*, Brussels, Belgium.
11. **Sebastian Ruder**, Barbara Plank (2018). Strong Baselines for Neural Semi-supervised Learning under Domain Shift. In *Proceedings of ACL 2018*, Melbourne, Australia.
12. Jeremy Howard\*, **Sebastian Ruder\*** (2018). Universal Language Model Fine-tuning for Text Classification. In *Proceedings of ACL 2018*, Melbourne, Australia.
13. Anders Søgaard, **Sebastian Ruder**, Ivan Vulić (2018). On the Limitations of Unsupervised Bilingual Dictionary Induction. In *Proceedings of ACL 2018*, Melbourne, Australia.
14. **Sebastian Ruder**, John Glover, Afshin Mehrabani, Parsa Ghaffari (2018). 360° Stance Detection. In *Proceedings of NAACL-HLT 2018: System Demonstrations*, New Orleans, US.
15. 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.
16. **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.
17. **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

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

## Tutorials

19. **Sebastian Ruder**, Anders Søgaard, Ivan Vulić (2019). Unsupervised Cross-Lingual Representation Learning. In *Proceedings of ACL 2019: Tutorial Abstracts*, pages 31–38.
20. **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

21. 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.

---

\*Equal contribution.

22. **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.
23. Ian D. Wood and **Sebastian Ruder** (2016). Emoji as emotion tags for tweets. In *Emotion and Sentiment Analysis Workshop, LREC*, Portorož, Slovenia.
24. **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.
25. **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.

## Preprints

26. Chris Hokamp, **Sebastian Ruder**, John Glover (2018). Off-the-Shelf Unsupervised NMT. arXiv preprint arXiv:1811.02278.
27. **Sebastian Ruder** (2017). An Overview of Multi-Task Learning in Deep Neural Networks. arXiv preprint arXiv:1706.05098.
28. **Sebastian Ruder**, Parsa Ghaffari, John G. Breslin (2017). Data Selection Strategies for Multi-Domain Sentiment Analysis. arXiv preprint arXiv:1702.02426.
29. **Sebastian Ruder**, Parsa Ghaffari, John G. Breslin (2017). Knowledge Adaptation: Teaching to Adapt. arXiv preprint arXiv:1702.02052.
30. **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 <i>Cusanuswerk</i> , 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 <i>Bundeswettbewerb Fremdsprachen</i> , national foreign languages competition | 2007 – 2008       |
| • First and second prizes <i>Landeswettbewerb Mathematik</i> , 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 2019<sup>1</sup>; Publication chair at CoNLL 2019; Publicity chair at EMNLP 2019

---

<sup>1</sup><https://www.eurnlp.org/>

- **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 2018<sup>2</sup>, 2019
- **Meetups:** NLP Dublin Meetup<sup>3</sup>

## Reviewing

- **Journals:** Transactions on Audio, Speech and Language Processing; Artificial Intelligence; IEEE Computational Intelligence Magazine; Natural Language Engineering
- **Conferences:** ACL 2018, 2019; ICML 2019; NeurIPS 2019; NAACL-HLT 2019; ICLR 2019; EMNLP 2018, 2019; CoNLL 2018
- **Workshops:** RepL4NLP 2018, DeepLo 2018, SemEval-2016 Task 5

---

## Talks

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

---

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

<sup>3</sup><https://www.meetup.com/NLP-Dublin/>