

# Sebastian Ruder

sebastianruder.com • ruders.sebastian@gmail.com  
+353 89 237 9192 • 9 New Bride Street • Dublin, Ireland

---

## Experience

- **AYLIEN**  
*Research Scientist*

**Dublin, Ireland**  
*10/2015 – Present*

  - Developed new aspect-based sentiment analysis (ABSA) endpoint<sup>1</sup> from scratch; created data collection and training pipeline; implemented semi-supervised convolutional neural networks.
  - Developed multi-lingual models for sentiment analysis on-par with state-of-the-art<sup>2</sup>.
  - My current work focuses on developing and improving systems for sentiment analysis with a focus on aspects and entities and adapting systems to new domains and languages. For the latter objective, I also work on developing new general-purpose cross-lingual embeddings for words and documents.
- **INSIGHT Centre for Data Analytics**  
*Research Assistant*

**Galway, Ireland**  
*10/2015 – Present*

  - As a member of the INSIGHT Centre, I collaborate with other researchers on emotion detection, sentiment analysis, and review analysis.
- **IBM**  
*Extreme Blue Intern, Watson*

**Munich, Germany**  
*08/2015 – 09/2015*

  - Design and implementation of text analysis ML components applied to customer data of leading German insurance company *Versicherungskammer Bayern*; automatically identifies structural semantics and sentiment of incoming e-mails, e.g. complaints and classifies email based on reason for complaint.
  - Pitched project to audience at European Expo and was chosen as one of eight teams to pitch to IBM customers; project was referred to as a "lighthouse project for Watson in Europe" by jury members.
  - Project was awarded Digital Thought Leadership award in leading contest of German insurance industry by leading German newspaper *Süddeutsche Zeitung* and Google<sup>3</sup> and covered by *Süddeutsche Zeitung*<sup>4</sup>.
- **Microsoft**  
*Linguistic Engineering Intern, Xbox*

**Dublin, Ireland**  
*02/2015 – 06/2015*

  - Contributed to developing an ML system for analyzing linguistic complexity of strings in C# for localization prioritization during testing; performed feature analysis and framed problem as anomaly detection.
  - Created proof of concept and implemented terminology validation algorithm using morphology generation.
  - Evangelized customer sentiment analysis efforts, drove cross-team collaboration, and provided insights to stakeholders.
- **The OpenCog Foundation**  
*Google Summer of Code Intern*

**opencog.org**  
*Summer 2014*

  - Enabled system to make common-sense inferences using deductive reasoning, e.g. *All men are mortal. Socrates is a man. → Socrates is mortal.*
  - Applied inference using probabilistic logic networks on the output of a relationship extractor.
  - Documented and extended Python code for temporal inference.
- **Lingenio GmbH**  
*Software Engineering Intern*

**Heidelberg, Germany**  
*Spring 2014*

  - Created a converter from TBX to Lingenio native format and vice versa.
  - Integrated TBX term bases in Dictionary Server; created localized web service using Jinja2, Flask-Babel, and lighttpd.
- **SAP**  
*Working Student, Development University*

**Walldorf, Germany**  
*02/2013 – 02/2014*

---

<sup>1</sup><https://developer.aylien.com/text-api-demo?tab=absa>

<sup>2</sup><https://developer.aylien.com/text-api-demo?tab=sentiment>

<sup>3</sup><https://www.sv-veranstaltungen.de/site/fachbereiche/versicherungs-leuchtturm>

<sup>4</sup><http://www.sueddeutsche.de/wirtschaft/kuenstliche-intelligenz-aerger-fuer-watson-1.2772927>

- Created content for internal programming and Design Thinking courses.
- Automated reporting processes, e.g. reduced expenditure of work for monthly training report by > 75%, i.e. from 8 hours to 2 hour using Excel / VBA scripts.

- **TEMIS**

*Freelancing Developer*

**Heidelberg, Germany**

02/2013 – 10/2013

- Created a cosine metric-based word sense disambiguation system leveraging tens of GB of text extracted from Wikipedia and DBpedia in Python; achieved performance comparable to the state-of-the-art.

## Education

- **National University of Ireland**

*College of Engineering and Informatics, Ph.D. Natural Language Processing*

**Galway, Ireland**

10/2015 – Present

- Research in domain adaptation and transfer learning of ML and deep learning techniques with applications to cross-lingual text classification, particularly sentiment analysis, personalization, and dialogue modeling.

- **Ruprecht-Karls-Universität Heidelberg**

*Institute of Computational Linguistics, B.A. Computational Linguistics, English Linguistics*

**Heidelberg, Germany**

10/2012 – 09/2015

- Final grade: 1.0 (German scale), i.e. GPA 4.0; thesis: *Construction and Analysis of an Emotion Proposition Store*
- Relevant courses: Statistics, Algorithms and Data Structures, Machine Learning, Formal Syntax & Semantics
- Relevant online courses: Machine Learning (Stanford), AI (MIT), Intro to Algorithms (Berkeley), Deep Learning for NLP (Stanford), Deep Learning (Oxford)

- **Trinity College**

*School of Computer Science and Statistics, Computer Science and Language*

**Dublin, Ireland**

09/2014 – 01/2015

- Semester abroad
- relevant courses: AI, Fuzzy Logic, High-Tech Entrepreneurship

## Awards

- Scholarship of the Irish Research Council 10/2015 – Present
- Cusanuswerk scholarship of the German state 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

## Languages and Technologies

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

**Technologies:** SciPy, NumPy, Keras, TensorFlow, 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

## Publications

- Sebastian Ruder, Parsa Ghaffari, John G. Breslin (2016). Towards a continuous modeling of natural language domains. In *Uphill Battles in Language Processing, EMNLP*, Austin, Texas, US.
- Sebastian Ruder, Parsa Ghaffari, John G. Breslin (2016). A Hierarchical Model of Reviews for Aspect-based Sentiment Analysis. In *EMNLP*, Austin, Texas, US.
- Ian D. Wood and Sebastian Ruder (2016). Emoji as emotion tags for tweets. In *Emotion and Sentiment Analysis Workshop, LREC*, Portorož, Slovenia.

- Sebastian Ruder, Peiman Barnaghi, John G. Breslin (2016). Analysis and Applications of a Novel Corpus of Influencers on Twitter. In *Twitter for Research Conference*, Galway, Ireland.
  - 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.
  - 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.
- 

## Talks

- INSIGHT SIG NLP meetup, August 2016: A Hierarchical Model of Reviews for Aspect-based Sentiment Analysis
- NLP Dublin meetup, August 2016: Softmax Approximations for Learning Word Embeddings and Language Modelling