

# CURRICULUM VITAE

## ISABELLE AUGENSTEIN

### PERSONAL DETAILS

---

Isabelle Augenstein  
Universitetsparken 5  
2100 Copenhagen, Denmark  
Email: augenstein@di.ku.dk  
WWW: <http://isabelleaugenstein.github.io/>  
Twitter: @IAugenstein  
Github: <https://github.com/isabelleaugenstein>

LinkedIn: <http://uk.linkedin.com/pub/isabelle-augenstein/7a/36b/824/>  
Google Scholar: <http://scholar.google.de/citations?user=DjJp0dcAAAAJ>  
Slideshare: <http://www.slideshare.net/isabelleaugenstein/>  
Videolectures: [http://videolectures.net/isabelle\\_augenstein/](http://videolectures.net/isabelle_augenstein/)

### RESEARCH INTERESTS

---

Statistical Natural Language Processing • Machine Learning • Information Extraction • Knowledge Base Population • Weakly Supervised Learning • Structured Prediction • Representation Learning

### RESEARCH EXPERIENCE

---

- since 07/2017    **Tenure-Track Assistant Professor** in Natural Language Processing and Machine Learning at Department of Computer Science, University of Copenhagen
- 03/2016    -    **Postdoctoral Research Associate** in Machine Reading Group at Department of Computer Science, University College London
- 06/2017
- 10/2015    -    **Research Associate** in Natural Language Processing Group at Department of Computer Science, The University of Sheffield
- 02/2016
- 04 – 09/2015    **Research Associate (Part-Time)** in Natural Language Processing Group at Department of Computer Science, The University of Sheffield
- 10/2012    -    **PhD Student** at Department of Computer Science, The University of Sheffield
- 2016
- 06/2010    -    **Research Assistant (Part-Time)** at AIFB, Karlsruhe Institute of Technology
- 09/2012
- 03 – 04/2010    **Research Intern** at AIFB, Karlsruhe Institute of Technology

## RESEARCH PROJECTS

---

- 07/2017 – Real Time Big Data Product Analysis – Product Management System  
07/2017 in International Markets (ReProsis), funder: Eurostars  
*PI from 02/2017, research assistant on project:  
Sune Andreas Dybro Debel*
- 08/2017 – Automated Fact Checking of Claims Online, funder: Ministry of Higher  
09/2017 Education and Science, Denmark  
*PI*
- 09/2017 – Low-Resource Learning in Natural Language Processing, funder: Depart-  
02/2019 ment of Computer Science, University of Copenhagen  
*PI, postdoc on project: Johannes Bjerva*
- 03/2016 – Construction of Method and Algorithm Knowledge Graphs, funded by  
06/2017 Elsevier  
*Researcher*
- 04/2015 – PHEME (Computing Veracity Across Media, Languages, and Social Net-  
02/2016 works), EU FP7 ICT  
*Researcher*
- 10/2012 – LODIE (Linked Open Data for Information Extraction), EP/J019488/1  
11/2015 *PhD student associated with project*
- 10/2012 – Web Relation Extraction with Distant Supervision  
12/2015 *PhD topic*
- 2010 – 2012 LODifier - Generating Linked Data from Unstructured Text  
*Research topic at AIFB, BA thesis topic*

## ACADEMIC EDUCATION

---

- 10/2012 – Study of Computer Science (PhD)  
2016 **University of Sheffield**  
*Thesis topic: Web Information Extraction using Linked Data  
Supervisors: Prof. Fabio Ciravegna, Dr. Diana Maynard  
PhD thesis submitted: 12/2015, PhD viva: 3/2016, PhD awarded:  
09/2016*
- 2011 – 2012 Study of Computational Linguistics (M.A.)  
**Heidelberg University**
- 08/2011 B.A. Computational Linguistics and Psychology  
Final mark: *1.9*
- 2008 – 2011 Study of Computational Linguistics and Psychology (B.A.)  
**Heidelberg University**  
*Thesis topic: LODifier - Generating RDF from Natural Language  
Supervisors: Prof. Sebastian Padó, Prof. Sebastian Rudolph*

## SCHOOL EDUCATION

---

July 2008	Abitur Final mark: 1.5
2005 – 2008	Heinrich-Wieland-Schule, Technisches Gymnasium, Pforzheim
1999 – 2005	Theodor-Heuss-Gymnasium, Mühlacker
1995 – 1999	Ulrich-von-Dürrenz-Schule, Mühlacker

## OTHER EDUCATIONAL ACTIVITIES

---

07/2015	8th GATE Text Mining Summer Training Course, Sheffield, UK
01 - 05/2015	Machine Learning, Coursera, Tutor: Prof. Andrew Ng
07/2013	10th Summer School on Ontology Engineering and the Semantic Web (SSSW 2013), Cercedilla, Spain

## PUBLICATIONS

---

### SELECTED PUBLICATIONS

- **Isabelle Augenstein**, Sebastian Ruder, Anders Søgaard. *Multi-task Learning of Pairwise Sequence Classification Tasks Over Disparate Label Spaces*. Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2018), to appear.
- Johannes Bjerva, **Isabelle Augenstein**. *Unsupervised Linguistic Typology at Different Levels with Language Embeddings*. Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2018), to appear.
- **Isabelle Augenstein**, Tim Rocktäschel, Andreas Vlachos, Kalina Bontcheva. *Twitter Stance Detection with Bidirectional Conditional Encoding*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2016), November 2016.
- **Isabelle Augenstein**, Anders Søgaard. *Multi-task learning of keyphrase boundary classification*. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), Short Papers, August 2017.
- Ben Eisner, Tim Rocktäschel, **Isabelle Augenstein**, Matko Bošnjak, Sebastian Riedel. *emoji2vec: Learning Emoji Representations from their Description*. Proceedings of the International Workshop on Natural Language Processing for Social Media at EMNLP 2016, November 2016. **Best paper award**.

## PAPERS IN CONFERENCE PROCEEDINGS

- Yova Kementchedjhieva, Johannes Bjerva, **Isabelle Augenstein**. *Copenhagen at ConLL–SIGMORPHON 2018: Multilingual Inflection in Context with Explicit Morphosyntactic Decoding*. Proceedings of ConLL–SIGMORPHON 2018, to appear.
- Ana Valeria Gonzalez-Garduño, **Isabelle Augenstein**, Anders Søgaard. *A strong baseline for question relevancy ranking*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2018), to appear.
- Miryam de Lhoneux, Johannes Bjerva, **Isabelle Augenstein**, Anders Søgaard. *Parameter sharing between dependency parsers for related languages*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2018), to appear.
- Dirk Weissenborn, Pasquale Minervini, Tim Dettmers, **Isabelle Augenstein**, Johannes Welbl, Tim Rocktäschel, Matko Bošnjak, Jeff Mitchell, Thomas Demeester, Pontus Stenetorp, Sebastian Riedel. *Jack the Reader – A Machine Reading Framework*. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL 2018), System Demonstrations.
- **Isabelle Augenstein**, Sebastian Ruder, Anders Søgaard. *Multi-task Learning of Pairwise Sequence Classification Tasks Over Disparate Label Spaces*. Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2018), to appear.
- Johannes Bjerva, **Isabelle Augenstein**. *Unsupervised Linguistic Typology at Different Levels with Language Embeddings*. Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2018), to appear.
- **Isabelle Augenstein**, Anders Søgaard. *Multi-task learning of keyphrase boundary classification*. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), Short Papers, August 2017.
- Ed Collins, **Isabelle Augenstein**, Sebastian Riedel. *A Supervised Approach to Extractive Summarisation of Scientific Papers*. Proceedings of SIGNLL Conference on Computational Natural Language Learning (CoNLL 2017), July 2017.
- **Isabelle Augenstein**, Tim Rocktäschel, Andreas Vlachos, Kalina Bontcheva. *Twitter Stance Detection with Bidirectional Conditional Encoding*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2016), November 2016.
- George Spithourakis, **Isabelle Augenstein**, Sebastian Riedel. *Numerically Grounded Language Models for Semantic Error Correction*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2016), November 2016.
- Piroska Lendvai, **Isabelle Augenstein**, Kalina Bontcheva, Thierry Declerck. *Monolingual Social Media Datasets for Detecting Contradiction and Entailment*. Language Resources and Evaluation Conference (LREC), May 2016.

- **Isabelle Augenstein**, Andreas Vlachos, Diana Maynard (2015). *Extracting Relations between Non-Standard Entities using Distant Supervision and Imitation Learning*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2015), September 2015.
- **Isabelle Augenstein**, Diana Maynard, Fabio Ciravegna (2014). *Relation Extraction from the Web using Distant Supervision*. Proceedings of the 19th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2014), November 2014. **Nominated for best research paper.**
- **Isabelle Augenstein** (2014). *Joint Information Extraction from the Web using Linked Data*. Proceedings of the 12th International Semantic Web Conference (ISWC 2014), October 2014.
- Ziqi Zhang, Anna Lisa Gentile, Eva Blomqvist, **Isabelle Augenstein**, Fabio Ciravegna (2013). *Statistical Knowledge Patterns: Identifying Synonymous Relations in Large Linked Datasets*. Proceedings of the 12th International Semantic Web Conference (ISWC 2013), October 2013.
- Ziqi Zhang, Anna Lisa Gentile, **Isabelle Augenstein**, Eva Blomqvist, Fabio Ciravegna (2013). *Mining Equivalent Relations from Linked Data*. Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: Short Papers (ACL 2013), August 2013.
- Anna Lisa Gentile, Ziqi Zhang, **Isabelle Augenstein**, Fabio Ciravegna (2013). *Unsupervised Wrapper Induction using Linked Data*. Proceedings of the 8th International Conference on Knowledge Capture (K-CAP 2013), June 2013.
- **Isabelle Augenstein**, Anna Lisa Gentile, Barry Norton, Ziqi Zhang, Fabio Ciravegna (2013). *Mapping Keywords to Linked Data Resources for Automatic Query Expansion*. Proceedings of the 10th Extended Semantic Web Conference (ESWC 2013), May 2013. **Best Workshop Paper of the 2nd International Workshop on Knowledge Discovery and Data Mining Meets Linked Open Data (KNOW@LOD 2013).**
- **Isabelle Augenstein**, Sebastian Padó and Sebastian Rudolph (2012). *LODifier: Generating Linked Data from Unstructured Text*. Proceedings of the 9th Extended Semantic Web Conference (ESWC 2012), May 2012.

#### PAPERS IN WORKSHOP PROCEEDINGS

- Anders Søgaard, Miryam de Lhoneux, **Isabelle Augenstein**. *Nightmare at test time: How punctuation prevents parsers from generalizing*. Proceedings of the Analyzing and interpreting neural networks for NLP (BlackboxNLP at EMNLP 2018), to appear.
- Katharina Kann, Johannes Bjerva, **Isabelle Augenstein**, Barbara Plank and Anders Søgaard. *Character-level Supervision for Low-resource POS Tagging*. Proceedings of the 1st Workshop on Deep Learning Approaches for Low Resource Natural Language Processing (DeepLo at ACL 2018)

- Thomas Nyegaard-Signori, Casper Veistrup Helms, Johannes Bjerva, **Isabelle Augenstein**. KU-MTL at SemEval-2018 Task 1: Multi-task Identification of Affect in Tweets. Proceedings of International Workshop on Semantic Evaluation (SemEval), June 2018.
- Johannes Bjerva, **Isabelle Augenstein**. Tracking Typological Traits of Uralic Languages in Distributed Language Representations. Proceedings of Fourth International Workshop on Computational Linguistics for Uralic Languages (IWCLUL), January 2018.
- **Isabelle Augenstein**, Mrinal Das, Sebastian Riedel, Lakshmi Nair Vikraman, Andrew McCallum. SemEval 2017 Task 10: ScienceIE - Extracting Keyphrases and Relations from Scientific Publications. Proceedings of International Workshop on Semantic Evaluation (SemEval), August 2017.
- Elena Kochkina, Maria Liakata, Arkaitz Zubiaga, **Isabelle Augenstein**. *Turing at SemEval-2017 Task 8: Sequential approach to Rumour Stance Classification with branch-LSTM*. Proceedings of International Workshop on Semantic Evaluation (SemEval at ACL), August 2017. **Highest scoring system for SemEval-2017 Task 8 Subtask A, best paper award.**
- Elena Kochkina, Maria Liakata, **Isabelle Augenstein**. *Sequential Approach to Rumour Stance Classification*. Proceedings of the ACL Workshop on Women and Underrepresented Minorities in Natural Language Processing, July 2017.
- Ben Eisner, Tim Rocktäschel, **Isabelle Augenstein**, Matko Bosnak, Sebastian Riedel. *emoji2vec: Learning Emoji Representations from their Description*. Proceedings of the International Workshop on Natural Language Processing for Social Media at EMNLP 2016, November 2016. **Best paper award.**
- **Isabelle Augenstein**, Andreas Vlachos and Kalina Bontcheva. *USFD at SemEval-2016 Task 6: Any-Target Stance Detection on Twitter with Autoencoders*. Proceedings of the International Workshop on Semantic Evaluation (SemEval 2016), June 2016.
- Leon Derczynski, **Isabelle Augenstein**, Kalina Bontcheva (2015). *USFD: Twitter NER with Drift Compensation and Linked Data*. Proceedings of the ACL Workshop on Noisy User-generated Text (W-NUT 2015), July 2015.
- **Isabelle Augenstein** (2014). *Seed Selection for Distantly Supervised Web-Based Relation Extraction*. Proceedings of the COLING Workshop on Semantic Web and Information Extraction (SWAIE 2014), August 2014.
- Eva Blomqvist, Ziqi Zhang, Anna Lisa Gentile, **Isabelle Augenstein**, Fabio Ciravegna (2013). *Statistical Knowledge Patterns for Characterising Linked Data*. Proceedings of the 4th Workshop on Ontology and Semantic Web Patterns (WOP 2013), October 2013.

## JOURNAL ARTICLES

- Arkaitz Zubiaga, Elena Kochkina, Maria Liakata, Rob Procter, Michal Lukasik, Kalina Bontcheva, Trevor Cohn, **Isabelle Augenstein**. Discourse-Aware Rumour Stance Classification in Social Media Using Sequential Classifiers. *Information Processing and Management*, Volume 54, Issue 2, March 2018, Pages 273-290.

- **Isabelle Augenstein**, Leon Derczynski, Kalina Bontcheva. Generalisation in Named Entity Recognition: A Quantitative Analysis. *Computer Speech & Language*, February 2017.
- Ziqi Zhang, Anna Lisa Gentile, **Isabelle Augenstein**, Eva Blomqvist, Fabio Ciravegna (2016). *An Unsupervised Data-driven Method to Discover Equivalent Relations in Large Linked Datasets*. *Semantic Web Journal*, Volume 8, Number 2 / 2017.
- **Isabelle Augenstein**, Diana Maynard, Fabio Ciravegna (2016). *Distantly Supervised Web Relation Extraction for Knowledge Base Population*. *Semantic Web Journal*, Volume 7, Number 4 / 2016.

## BOOKS

- Diana Maynard, Kalina Bontcheva, **Isabelle Augenstein**. *Natural Language Processing for the Semantic Web*. Morgan & Claypool - Synthesis Lectures on the Semantic Web, December 2016.

## OTHER

- **Isabelle Augenstein**, Kris Cao, He He, Felix Hill, Spandana Gella, Jamie Kiros, Hongyuan Mei, Dipendra Misra. *Proceedings of The Third Workshop on Representation Learning for NLP*. (Repl4NLP at ACL 2018).
- Benjamin Riedel, **Isabelle Augenstein**, Georgios Spithourakis, Sebastian Riedel (2017). *A simple but tough-to-beat baseline for the Fake News Challenge stance detection task*. *CoRR*, abs/1707.03264, July 2017.
- Sebastian Ruder, Joachim Bingel, **Isabelle Augenstein**, Anders Søgaard (2017). *Sluice networks: Learning what to share between loosely related tasks*. *CoRR*, abs/1705.08142, May 2017.
- Piroska Lendvai, **Isabelle Augenstein**, Dominic Rout, Kalina Bontcheva, Thierry Declerck. *Algorithms for Detecting Disputed Information: Final Version*. FP7-ICT Collaborative Project ICT-2013-611233 PHEME Deliverable D4.2.2 (WP4), June 2016.
- Ziqi Zhang, Anna Lisa Gentile, **Isabelle Augenstein** (2014). *Linked Data as Background Knowledge for Information Extraction on the Web*. *ACM SIGWEB Newsletter*, July 2014.

## DISSERTATIONS

- **Isabelle Augenstein** (2016). *Web Information Extraction using Distant Supervision*. PhD theses: University of Sheffield, Department of Computer Science, UK, July 2016.
- **Isabelle Augenstein** (2011). *LODifier - Generating RDF from Natural Language*. Heidelberg University, Department of Computational Linguistics, July 2011.

## PARTICIPATION IN SHARED TASKS

---

- 3rd place out of 50 scored systems in Fake News Challenge stance detection task, USD 400 cash prize
- 1st place in SemEval 2017 Task 8 “Determining rumour veracity and support for rumours” SubTask A (stance detection)
- 6th place in SemEval 2016 Task 6 “Detecting Stance in Tweets” SubTask B (weakly supervised task)
- 2nd place in ACL 2015 W-NUT Twitter Named Entity Recognition no-types challenge

## TALKS

---

- *Multi-task Learning of Pairwise Sequence Classification Tasks*. Johns Hopkins University, Center for Language and Speech Processing, Baltimore, Maryland, USA, June 2018. *Invited talk, all expenses paid*
- *Multi-task Learning of Pairwise Sequence Classification Tasks*. New York University, Machine Learning for Language group, New York City, New York, USA, June 2018. *Invited talk*
- *Multi-task Learning of Pairwise Sequence Classification Tasks*. University of Ghent, Information Extraction and Retrieval group, Ghent, Belgium, May 2018. *Invited talk, all expenses paid*
- *Learning with limited labelled data in NLP: multi-task learning and beyond*. University of Edinburgh, ILCC Seminar Series, Edinburgh, UK, May 2018. *Invited talk, all expenses paid*
- *Learning to Read for Automated Fact Checking*. Ukrainian AI/NLP Meetup Series, Kiev, Ukraine, November 2017. *Invited talk, all expenses paid*
- *Learning what to share in multi-task learning*. Grammarly, Kiev, Ukraine, November 2017. *Invited talk, all expenses paid*
- *Learning to Read for Automated Fact Checking*. Computer Science Department Lecture Series, Aarhus University, Aarhus, Denmark, November 2017. *Invited talk, paid for travel and meals*
- *Learning to Read for Automated Fact Checking*. AU Datalab Lecture Series, Aarhus University, Aarhus, Denmark, November 2017. *Invited talk, paid for travel and meals*



- *Machine Reading Using Neural Machines*. University of Linköping, Linköping, Sweden, November 2017. *Invited talk, paid for travel, accommodation and meals*
- *NLP beyond the sentence level*. Structured Prediction for NLP workshop at EMNLP, July 2017.
- *Machine Reading Using Neural Machines*. Microsoft Research, Bellevue, Washington, USA, July 2017. *Invited talk, paid for flights, accommodation and meals*
- *Learning to Read for Automated Fact Checking*. Google and DeepMind, London, UK, April 2017. *Invited talk*.
- *Weakly supervised Machine Reading for Science Publications*. National Institute for Informatics (NII), Tokyo, Japan, January 2017. *Invited talk and all expenses paid visit.*
- *Weakly supervised Machine Reading for Science Publications*. University of Tokyo Department of Computer Science, Tokyo, Japan, January 2017. *Invited talk and all expenses paid visit.*
- *Weakly supervised Machine Reading for Science Publications*. NAIST Computational Linguistics Lab, Ikoma, Japan, January 2017. *Invited talk and all expenses paid visit.*
- *Machine Reading*. Big Data Application Symposium, NAIST Big Data Project, Ikoma, Japan, January 2017. *Invited talk, all expenses paid 1-week research visit and additional stipend.*
- *Information Extraction meets Question Answering*. Answering Questions using Web Data (WDAqua) Marie Skłodowska-Curie Innovative Training Network (ITN) R&D Week, Athens, Greece, December 2016. *Invited talk and all expenses paid 3-day visit.*
- *Learning to Detect Stance and Represent Emojis*, University of Cambridge, Cambridge, UK. Cambridge NLIP Seminar Series, November 2016. *Invited talk and paid travel.*
- *Learning to Detect Stance and Represent Emojis*, South England NLP Meetup, London, UK. *Invited talk.*
- *Weakly Supervised Machine Reading*, University of Sheffield, Sheffield, UK. University of Sheffield NLP Seminar Series, October 2016. *Invited talk.*
- *Weakly Supervised Machine Reading*, University of Copenhagen, Copenhagen, DK. University of Copenhagen Computer Science department, October 2016. *Invited talk and all expenses paid 1-week research visit.*
- *Distant Supervision with Imitation Learning*, Heriot-Watt University, Edinburgh, UK. University of Edinburgh Computer Science Seminar Series, November 2015. *Invited talk and paid travel.*
- *Distant Supervision with Imitation Learning*, University of Sheffield, Sheffield, UK. University of Sheffield NLP Seminar Series, November 2015.
- *Web Information Extraction using Distant Supervision*, University of Sheffield, Sheffield, UK. University of Sheffield NLP Seminar Series, January 2015.

## LANGUAGE SKILLS

---

- German: Native
- English: Fluent
- Danish, Swedish: Moderate
- French, Spanish: Basic

## PROGRAMMING AND SCRIPTING LANGUAGES

---

- Python, Java: Very Good
- LaTeX, HTML: Good
- C++, Bash, XML, Javascript: Moderate
- C, Assembler, Visual Basic, PHP, R: Basic

## IT EXPERIENCE

---

- Natural Language Processing Software: GATE, StanfordNLP, OpenNLP, NLTK, WordNet
- Machine Learning Software: scikit-learn, TensorFlow, DyNet, Weka, Mallet, CRF-Suite
- Numerical Analysis Software: Octave, R, SPSS, gnuplot
- Other Software: Tomcat, Hadoop, Eclipse, PyCharm, Lucene, Amazon Web Services
- Operating systems: Linux, Mac OS, Windows
- Databases: MySQL, Jena, GraphDB (OWLIM)

## ORGANISING COMMITTEES

---

- |      |  |
|------|--|
| 2018 | 22nd Conference on Empirical Methods in Natural Language Processing (EMNLP 2018), Area Chair for Information Extraction and Question Answering |
|------|--|

2018	27th International Conference on Computational Linguistics (Coling 2018), Area Chair
2018	3rd Workshop on Representation Learning for NLP (Repl4NLP) at ACL 2018, Primary Organiser
2018	13th International Workshop on Semantic Evaluation (SemEval 2019), Steering Committee member
2018	2nd WiNLP Workshop on Widening NLP at ACL 2018, Advisory Board Member
2017	1st Workshop on Learning with Limited Labeled Data: Weak Supervision and Beyond (LLD 2017) at NIPS 2017, Organiser
2017	1st Workshop on Deep Structured Prediction (DeepStruct 2017) at ICML 2017, Organiser
2017	21st Conference on Empirical Methods in Natural Language Processing (EMNLP 2017), Publicity Chair
2017	1st Women and Underrepresented Minorities in Natural Language Processing Workshop (WiNLP 2017) at ACL 2017, Organiser
2017	11th Workshop on Semantic Evaluation (SemEval 2017) at ACL 2017, Primary Organiser of Task 10, ScienceIE: Extracting Keyphrases and Relations from Scientific Publications, Session Chair
2015	5th European Semantic Web Summer School (ESWCSS 2015), Community Manager
2014	2nd University of Sheffield Engineering Symposium (USES 2014), Artificial Intelligence & Virtual Reality Track Chair
2013 - 2014	President of USES, the University of Sheffield Postgraduate Society
2008 - 2012	Member of the student council at the Department of Computational Linguistics, Heidelberg University

## PROGRAMME COMMITTEES & REVIEWING

---

- Reviewing for Journals: Computational Linguistics journal (CL) 2017-2018, Transactions of the Association for Computational Linguistics (TACL) 2016, Information Processing & Management (IP&M) 2016
- Programme Committee Member for Conferences: NIPS 2018, KI 2018, ACL 2018, NAACL 2018, EMNLP 2017, NoDaLiDa 2017, EACL 2017 (Information Extraction Track), COLING 2016 (Applications Track), EMNLP 2016 (Machine Learning Track), EMNLP 2015 (Information Extraction Track), ISWC 2015 (Research Track), SEMAPRO 2015, ESWC 2015 (Poster & Demo Track), WWW 2015 (Research Track), ISWC 2014 (Research Track, In Use Track), SEMAPRO 2014

- Programme Committee Member for Workshops: FEVER at EMNLP 2018, W-NUT at EMNLP 2018, DeepLo at ACL 2018, SRW at ACL 2018, WiNLP at NAACL 2018, Repl4NLP at EMNLP 2017, Ethics in NLP at ACL 2017, SemEval at ACL 2017, SemEval at NAACL 2016, KET at WWW 2015, AHA!-Workshop at COLING 2014, LD4IE at ISWC 2014, LD4IE at ISWC 2013, NLP&LOD 2013 at RANLP 2013
- Subreviewer for Conferences: ACL 2015, ESWC 2014 (Poster & Demo Track)

## UNIVERSITY INTERNAL SERVICES

---

since 2017	On committee for communications at the University of Copenhagen
since 2017	On panel for recruitment and selection of University of Copenhagen Computer Science staff
since 2017	Co-organiser of Copenhagen NLP meetup
2017	Co-organiser of South England NLP meetup
2016 - 2017	On panel for recruitment and selection of UCL Computer Science staff
2016 - 2017	Publicity manager for UCL Machine Reading group
2016 - 2017	Marking of MSc Machine Learning dissertations

## TEACHING EXPERIENCE

---

### LECTURES & TUTORIALS

2019	Co-lecturer at Department of Computer Science, University of Copenhagen <i>Module:</i> Web Science
since 2018	Course responsible and lecturer at Department of Computer Science, University of Copenhagen <i>Module:</i> Natural Language Processing. Designed and taught a completely new course.
2018	Course responsible and lecturer at Department of Computer Science, University of Copenhagen <i>Module:</i> Web Science
08/2017	Deep Learning for Data Analysis Summer School 2017, Copenhagen Summer University, Lecture on Applications of RNNs in NLP.
2016	Guest Lecturer at Department of Computer Science, University College London <i>Module:</i> Statistical Natural Language Processing (Dr. Sebastian Riedel).
09/2015	Bridges Summer School 2016, Tutor, Tutorial on Practical Machine Learning for Social Media Analysis

09/2015	ESWC Summer School 2015, Tutor, Hands-On session on NLP for the Semantic Web
09/2014	ESWC Summer School 2014, Tutor, Hands-On session on Linked Data for NLP with GATE, Secondary tutor for two student projects
2013 – 2014	Demonstrator at Department of Computer Science, University of Sheffield <i>Modules:</i> Web Technologies, The Intelligent Web (Prof. Fabio Ciravegna)

#### MENTORING OF POSTDOCS AND PHD STUDENTS

- Johannes Bjerva. *Low-Resource Learning in Natural Language Processing*. UCPH Postdoctoral Researcher, since July 2017.
- Ana Valeria Gonzalez. *Multilingual question answering for customer service bots*. UCPH PhD Student, co-advisor, since March 2018.
- Simon Flachs. *Grammatical Error Correction*. UCPH PhD Student, co-advisor, since March 2018.
- Yova Kementchedjhieva. *Morphological Segmentation with Distant Supervision*. UCPH PhD Student, co-advisor, since December 2017.
- Mareike Hartmann. *Detecting Disinformation with Multilingual Stance Detection*. UCPH PhD Student, co-advisor, since September 2017.
- Miryam de Lhoneux. *Robust Multilingual Parsing in Low-Resource Settings*. UCPH Research Intern, co-advisor, January - April 2018.
- Sune Andreas Dybro Debel. *Real Time Big Data Product Analysis*. UCPH Research Assistant, since October 2017.
- George Spithourakis. *Multimodal distributional semantics for understanding natural language*. PhD in Computer Science Dissertation, expected submission in 2017, co-advised with Sebastian Riedel, Steffen Petersen at University College London.
- Elena Kochkina. *Modelling the conversational structure of tweets for rumour detection with deep learning*. PhD in Computer Science Dissertation, expected submission in 2018, co-advised with Maria Liakata at the University of Warwick via the Alan Turing Institute.
- Mentor at Women in Natural Language Processing lunch at ACL 2016, sponsored by Microsoft

#### SUPERVISION OF MASTER'S AND BACHELOR'S STUDENTS

- Peter Spliid. *A Framework for Typological Evaluation of Language Representations*. UCPH MSc in Computer Science Dissertation, to submit in 2018.
- Magnus Alexander Johansen. *Learning to ask questions about products*. UCPH MSc in Computer Science Dissertation, to submit in 2018.

- Andrea Lekkas. *Learning to ask questions about products*. UCPH MSc in Computer Science Dissertation, to submit in 2018.
- Raluca Alexandra Fetic. *Efficient encoding of large documents for stance detection*. UCPH MSc in Computer Science Dissertation, to submit in 2018.
- Nina Alef. *Learning to Predict Personality Traits from Social Media*. UCPH MSc in Computer Science Dissertation, to submit in 2018.
- Leendert Bastian van Doorn. *Variational encoding of location for dialect-conditional language generation*. UCPH MSc in Computer Science Research Project, co-supervised with Johannes Bjerva, to submit in 2018.
- Christoffer Trysøe, Andreas Borgstad, Xuwen Zhang. *Modelling customer behaviour from clickstream data*. UCPH MSc in Computer Science Research Project, co-supervised with Fabian Gieseke, to submit in 2018.
- Ke Zhai. *Deep Learning of Linguistic Features for Scandinavian Languages*. UCPH MSc in Computer Science Research Project, co-supervised with Johannes Bjerva, submitted in 2018.
- Thomas Nyegaard-Signorita and Casper Veistrup Helms. *Modelling affect in tweets*. UCPH BSc in Computer Science Dissertation, joint project, co-supervised with Johannes Bjerva, submitted in 2018.
- Rogan Inglis. *Machine Reading for Scientific Publications Using Generative Regularisation*. UCL MSc in Machine Learning Dissertation, co-advised with Pasquale Minervini, submitted in 2017.
- Benjamin Riedel. *Simple yet powerful: A system for news article stance detection*. UCL MSc in Machine Learning Dissertation, submitted in 2017, co-supervised with Sebastian Riedel, George Spithourakis.
- Danish Sheikh. *Extracting Keyphrases and Relations from Scientific Publications*. UCL MEng in Computer Science Dissertation, degree awarded in 2017.
- Edward Collins. *Generating Automatic Summaries of Scientific Papers*. UCL MEng in Computer Science Dissertation, degree awarded in 2017.
- Ben Eisner. *emoji2vec: Learning Emoji Representations from their Description*. UCL BSc Research Intern, 2016, co-supervised with Tim Rocktäschel.
- Rupert Chaplin. *End-to-end Differentiable and Interpretable Deep Learning for Natural Language Programming*. UCL MSc in Machine Learning Dissertation, degree awarded in 2016, co-supervised with Tim Rocktäschel.
- Dhruv Ghulati. *Distant Supervision and Cost Sensitive Classification for Weakly Supervised Claim Detection*. UCL MSc in Computer Science Dissertation, degree awarded in 2016, co-supervised with Sebastian Riedel, George Spithourakis.

## SCHOLARSHIPS, GRANTS

---

- 2018            **CCC Grant** for project ‘Natural Language Processing Seminar Series’, PI. Funding for organising a seminar series. DKK 34,500.
- 2018 – 2021    **BotXO PhD Fellowship Grant** for project ‘Multilingual question answering for customer service bots’, co-PI. Salary for a full 3-year PhD fellow, plus travel expenses
- 2018            **Eurostars Grant** for project ‘Real Time Big Data Product Analysis – Product Management System in International Markets (ReProsis)’, PI from February 2018. EUR 235,662 UCPH, EUR 1,313,362 total
- 2017            **EUopSTART Grant** from Ministry of Higher Education and Science, Denmark for project ‘Automated Fact Checking of Claims Online’, PI. DKK 50,000.
- 2017            **NVIDIA GPU Grant** for project ‘Semi-supervised multi-task learning for heterogenous label spaces’, September 2017, PI. 1 NVIDIA TITAN Xp GPU
- 2017            Discipline Specific Workshop grant from Computing Research Association (**CRA-DSW grant**) for organising a Women and Underrepresented Minorities workshop (WiNLP), to be held in the United States in 2018, co-PI. <http://cra.org/cra-w/discipline-specific-mentoring-workshops-dsw/>  
Award amount: USD 15,000
- 2017            **National Science Foundation (NSF) grant** for organising WiNLP 2017, held in Canada in 2017, author. Award amount: USD 20,000
- 2012            **Department of Computer Science, University of Sheffield PhD Studentship**  
Stipend for a full 3-year research scholarship, plus tuition fees

## AWARDS, RESEARCH VISITS

---

- 2018            **1st place in CoNLL-SIGMORPHON shared task** on universal morphological reinflection, Task 2 on morphological inflection in context
- 2017            **DIKU Dissemination Award 2017**, DKK 20,000 cash prize
- 2017            **3rd place in Fake News Challenge stance detection task** out of 50 scored systems, USD 400 cash prize
- 2017            Invited to 3-day **Microsoft Research Faculty Summit** in Bellevue, Washington, USA, paid for flights, accommodation and meals
- 2017            1st place in SemEval RumourEval stance detection task, and **Outstanding paper** award for the task paper at the International Workshop on Semantic Evaluation at ACL 2017

- 2016 Invited to 1-week **Research Visit** in Japan (NAIST, NII, University of Tokyo) by NAIST Big Data Project, paid for flights, accommodation, meals, and additional stipend
- 2016 Invited to 1-week **Research Visit** in Computer Science Department, University of Copenhagen, paid for flights, accommodation and meals
- 2016 **Best workshop paper** of the 4th International Workshop on Natural Language Processing for Social Media at EMNLP 2016 (SocialNLP)
- 2015 Invited to 4-day **Google NLP PhD Summit** in Zurich, paid for flights, accommodation and meals
- 2014 **Nominated for best research paper award** at EKAW 2014
- 2014 **ISWC 2014 student travel grant**  
€800 for registration and accommodation
- 2013 **Best workshop paper** of the 2nd International Workshop on Knowledge Discovery and Data Mining Meets Linked Open Data (KNOW@LOD)
- 2008 **Literary Society Scheffelbund, Scheffel prize**  
<http://web1.karlsruhe.de/Kultur/ML0/>  
Prize for the best exam results in German obtained by school leavers at grammar schools in the states of Baden-Württemberg  
€80, 5-year membership in Scheffelbund and book subscription

#### PHD COMMITTEES

---

- 2018 Lucas Sterckx, *Knowledge Base Population from Text and Graphs*, University of Ghent, Belgium

#### MEDIA

---

- 2017 “The new faces of data science”. Elsevier Connect article featuring a profile on my research, 17 November 2017, <https://www.elsevier.com/connect/the-new-faces-of-data-science>
- 2017 “L’IA au service de la recherche scientifique”. L’atelier BNP Paribas article describing our research project at Elsevier and how it can aid peer review, April 2017, <https://atelier.bnpparibas/prospective/article/l-ia-service-recherche-scientifique>
- 2017 “If AI Can Fix Peer Review in Science, AI Can Do Anything”. Wired article describing our research project at Elsevier and how it can aid peer review, also describing perspectives from other prominent AI researchers, 21 February 2017, <https://www.wired.com/2017/02/ai-can-solve-peer-review-ai-can-solve-anything/>



2017

Our emoji2vec paper which won the best paper award at SocialNLP at EMNLP 2016, described in Denny Britz's AI Newsletter (<https://www.getrevue.co/profile/wildml/issues/the-wild-week-in-ai-29-ai-partnerships-ec2-gpu-instances-google-s-neural-machine-translation-model-new-datasets-31604>), on a tumblr blog (<http://procedural-generation.tumblr.com/>) and receiving a significant amount of mentions on Twitter (<https://twitter.com/search?f=tweets&vertical=default&q=emoji2vec>), October 2016

## REFERENCES

---

On request

Copenhagen, 12 October 2018