

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/>
Videlectures: http://videlectures.net/isabelle_augenstein/

RESEARCH INTERESTS

Natural Language Processing • Machine Learning • Weakly Supervised Learning • Fact Checking • Explainability • Multilingual Learning • Question Answering

RESEARCH EXPERIENCE

- since 01/2021 **Head of the Natural Language Processing Section** at the Department of Computer Science, University of Copenhagen
- since 09/2020 **Deputy Head of Research** at CheckStep Ltd. (part-time)
- since 02/2020 **Associate Professor** in Natural Language Processing and Machine Learning at Department of Computer Science, University of Copenhagen
- 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
2016 Sheffield
- 06/2010 – **Research Assistant (Part-Time)** at AIFB, Karlsruhe Institute of
09/2012 Technology
- 03 – 04/2010 **Research Intern** at AIFB, Karlsruhe Institute of Technology

RESEARCH PROJECTS

- 04/2021 – Learning to Explain Attitudes on Social Media (EXPANSE), funding
03/2025 scheme: DFF (Independent Research Fund Denmark) Sapere Aude, *PI*
- 02/2021 – Typologically informed dependency parsing, funder: Swedish Research
01/2025 Council, *Host, postdoctoral fellow: Miryam de Lhoneux*
- 09/2021 – Privacy Black & White, funding scheme: University of Copenhagen Da-
08/2023 ta+, *PI*
- 09/2021 – Gendered Democratic Participation on Social Media, funding scheme:
08/2023 University of Copenhagen Data+, *co-PI*
- 09/2021 – Error and Accountability in Digitised Public Decision Making, funding
08/2023 scheme: University of Copenhagen Data+, *co-PI*
- 01/2020 – Structured Multilinguality for Natural Language Processing, funder:
12/2023 Swedish Research Council, *co-PI*
- 01/2020 – Detecting Gender-Biased Language on Social Media, funding scheme:
12/2023 DFF Project 1, *PI, PHD student on project: Karolina Stańczak*
- 01/2020 – Conceptualizing and Measuring Well-Being Using Statistical Semantics,
12/2023 funder: Swedish Research Council, *Host, postdoctoral fellow: Oscar Kjell*
- 09/2019 – Multi3Generation: Multi-task, Multilingual, Multi-modal Language Ge-
08/2023 neration, funding programme: COST (European Cooperation in Science
and Technology), *coordinator of 23-partner European project*
- 10/2019 – Data Enrichment to Advance Scientific Information Extraction, funding
09/2022 programme: Horizon 2020 Marie Skłodowska-Curie COFUND, *Co-PI,
PhD student on project: Dustin Wright*
- 07/2019 – Improving the democratic debate in a time of fake news,
06/2020 funder: Innovation Fund Denmark, *PI*
- 09/2019 – Transfer Learning of Textual Representations in the Life Science Domain,
08/2022 funder: Innovation Fund Denmark, Base Life Science, *PI, PhD student
on project: Andreas Nugaard Holm*
- 05/2019 – Subjective Question Answering, funder: Megagon Labs, *PI, postdocs on
04/2020 project: Johannes Bjerva, Sagnik Ray Choudhury*

- 05/2019 – Advanced Language Models for Writing Assistants (ALMA),
04/2022 funder: Innovation Fund Denmark, Ordbogen A/S,
Co-PI, PhD student on project: Andrea Lekkas
- 03/2019 – Tailoring knowledge for machine reading comprehension in both rich and
02/2022 scarce in resource languages and domains, funding scheme: Horizon 2020
Marie Skłodowska-Curie COFUND, *Co-PI, PhD student on project: Pepa
Atanasova*
- 10/2018 – Accessibility Increased through Real-life Multi-lingual Artificial Intelli-
01/2019 gence (AIRMAIL),
funder: Ministry of Higher Education and Science, Denmark, *PI*
- 03/2018 – Multilingual question answering for customer service bots, funder: Bot-
02/2021 XO, *Co-PI, PhD student on project: Ana Valeria Gonzalez*
- 02/2018 – Real Time Big Data Product Analysis – Product Management System in
07/2018 International Markets (ReProsis), funder: Eurostars, *PI from 02/2018,*
research assistant on project: Sune Andreas Dybro Debel
- 08/2017 – Automated Fact Checking of Claims Online,
11/2017 funder: Ministry of Higher 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: <i>1.9</i>
2008 – 2011	Study of Computational Linguistics and Psychology (B.A.) Heidelberg University <i>Thesis topic:</i> LODifier - Generating RDF from Natural Language <i>Supervisors:</i> Prof. Sebastian Padó, Prof. Sebastian Rudolph

SCHOOL EDUCATION

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

OTHER EDUCATIONAL ACTIVITIES

01/2020 12/2020	- Pre-leadership programme for female researchers, University of Copenhagen, Denmark
02/2019 06/2019	- PhD Student Supervision Course, University of Copenhagen, Denmark
01/2018 01/2019	- University Pedagogy Course, University of Copenhagen, Denmark
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

Google Scholar Bibliometrics: publications: 94; citations: 2220; h-index: 23; i10-index: 42

SELECTED PUBLICATIONS

- Pepa Atanasova, Jakob Grue Simonsen, Christina Lioma, **Isabelle Augenstein**. *Generating Fact Checking Explanations*. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL 2020).
- Alexander Hoyle, Lawrence Wolf-Sonkin, Hanna Wallach, **Isabelle Augenstein**, Ryan Cotterell. *Unsupervised Discovery of Gendered Language through Latent-Variable Modeling*. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL HLT 2019).
- Johannes Bjerva, Yova Kementchedjheva, Ryan Cotterell, **Isabelle Augenstein**. *A Probabilistic Generative Model of Linguistic Typology*. Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2019).
- **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).
- **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.

PAPERS IN CONFERENCE PROCEEDINGS

- Clara Meister, Stefan Lazov, **Isabelle Augenstein**, Ryan Cotterell. *Is Sparse Attention more Interpretable?* Proceedings of the Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021), August 2021.
- Dustin Wright, **Isabelle Augenstein**. *CiteWorth: Cite-Worthiness Detection for Improved Scientific Document Understanding*. Findings of the Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021), August 2021.
- Wojciech Ostrowski, Arnav Arora, Pepa Atanasova, **Isabelle Augenstein**. *Multi-Hop Fact Checking of Political Claims*. Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI 2021), August 2021.
- Wei Zhao, Steffen Eger, Johannes Bjerva, **Isabelle Augenstein** (2021). *Inducing Language-Agnostic Multilingual Representations*. Proceedings of 10th Joint Conference on Lexical and Computational Semantics (*SEM 2021), August 2021.

- Johannes Bjerva, **Isabelle Augenstein** (2021). *Does Typological Blinding Impede Cross-Lingual Sharing?*. Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2021), April 2021.
- Pepa Atanasova, Jakob Grue Simonsen, Christina Lioma, **Isabelle Augenstein** (2020). *A Diagnostic Study of Explainability Techniques for Text Classification*. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020), November 2020.
- Pepa Atanasova, Dustin Wright, **Isabelle Augenstein** (2020). *Generating Label Cohesive and Well-Formed Adversarial Claims*. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020), November 2020.
- Dustin Wright, **Isabelle Augenstein** (2020). *Transformer Based Multi-Source Domain Adaptation*. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020), November 2020.
- Johannes Bjerva, Nikita Bhutani, Behzad Golshan, Wang-Chiew Tan, **Isabelle Augenstein** (2020). *SubjQA: A Dataset for Subjectivity and Review Comprehension*. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020), November 2020.
- Farhad Nooralahzadeh, Giannis Bekoulis, Johannes Bjerva, **Isabelle Augenstein** (2020). *Zero-Shot Cross-Lingual Transfer with Meta Learning*. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020), November 2020.
- Dustin Wright, **Isabelle Augenstein** (2020). *Claim Check-Worthiness Detection as Positive Unlabelled Learning*. Findings of the Association for Computational Linguistics: EMNLP 2020, November 2020.
- Anna Rogers, **Isabelle Augenstein** (2020). *What Can We Do to Improve Peer Review in NLP?* Findings of the Association for Computational Linguistics: EMNLP 2020, November 2020.
- Nils Rethmeier, Vageesh Kumar Saxena, **Isabelle Augenstein** (2020). *TX-Ray: Quantifying and Explaining Model-Knowledge Transfer in (Un-)Supervised NLP*. In Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI 2020).
- Pepa Atanasova, Jakob Grue Simonsen, Christina Lioma, **Isabelle Augenstein**. *Generating Fact Checking Explanations*. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL 2020).
- Pranav A, **Isabelle Augenstein**. *2kenize: Tying Subword Sequences for Chinese Script Conversion*. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL 2020).
- Johannes Bjerva, Wouter Kouw, **Isabelle Augenstein**. *Back to the Future – Sequential Alignment of Text Representations*. Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI 2020).
- **Isabelle Augenstein**, Christina Lioma, Dongsheng Wang, Lucas Chaves Lima, Casper Hansen, Christian Hansen, Jakob Grue Simonsen. *MultiFC: A Real-World Multi-Domain Dataset for Evidence-Based Fact Checking of Claims*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2019).

- Johannes Bjerva, Yova Kementchedjheva, Ryan Cotterell, **Isabelle Augenstein**. *Uncovering Probabilistic Implications in Typological Knowledge Bases*. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL HLT 2019).
- Alexander Hoyle, Lawrence Wolf-Sonkin, Hanna Wallach, **Isabelle Augenstein**, Ryan Cotterell. *Unsupervised Discovery of Gendered Language through Latent-Variable Modeling*. Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL HLT 2019).
- Johannes Bjerva, Yova Kementchedjheva, Ryan Cotterell, **Isabelle Augenstein**. *A Probabilistic Generative Model of Linguistic Typology*. Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2019).
- Alexander Hoyle, Lawrence Wolf-Sonkin, Ryan Cotterell, Hanna Wallach, **Isabelle Augenstein**. *Combining Disparate Sentiment Lexica with a Multi-View Variational Autoencoder*. Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2019).
- Mareike Hartmann, Tallulah Jansen, **Isabelle Augenstein**, Anders Søgaard. *Issue Framing in Online Discussion Fora*. Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2019).
- Sebastian Ruder, Joachim Bingel, **Isabelle Augenstein**, Anders Søgaard (2019). *Latent Multi-Task Architecture Learning*. Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI 2019), January 2019.
- Yova Kementchedjheva, Johannes Bjerva, **Isabelle Augenstein**. *Copenhagen at ConLL–SIGMORPHON 2018: Multilingual Inflection in Context with Explicit Morphosyntactic Decoding*. Proceedings of ConLL–SIGMORPHON 2018.
- 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).
- 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).
- 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).

- Johannes Bjerva, **Isabelle Augenstein**. *From Phonology to Syntax: 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).
- **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 Civravegna (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

- **Isabelle Augenstein**. *Determining the Credibility of Science Communication*. In Proceedings of the Second Workshop on Scholarly Document Processing (SDP at NAACL 2021), June 2021.
- Johannes Bjerva, Elizabeth Salesky, Sabrina J. Mielke, Aditi Chaudhary, Giuseppe G. A. Celano, Edoardo M. Ponti, Ekaterina Vylomova, Ryan Cotterell, **Isabelle Augenstein**. *SIGTYP 2020 Shared Task: Prediction of Typological Features*. Proceedings of TyP-NLP: The Second Workshop on Typology for Polyglot NLP (TyP-NLP at EMNLP 2020), November 2020.
- Lukas Muttenthaler, **Isabelle Augenstein**, Johannes Bjerva. *Unsupervised Evaluation for Question Answering with Transformers*. Proceedings of the 3rd Workshop on Analyzing and interpreting neural networks for NLP (BlackboxNLP at EMNLP 2020), November 2020.
- Alok Debnath, Nikhil Pinnaparaju, Manish Shrivastava, Vasudeva Varma, **Isabelle Augenstein**. *Semantic Textual Similarity of Sentences with Emojis*. Proceedings of the 8th International Workshop on Natural Language Processing for Social Media (SocialNLP at TheWebConf 2020), April 2020.
- Mareike Hartmann, Yevgeniy Golovchenko, **Isabelle Augenstein**. *Mapping (Dis-)Information Flow about the MH17 Plane Crash*. Proceedings of the 2019 Workshop on NLP4IF: censorship, disinformation, and propaganda (NLP4IF at EMNLP 2019), November 2019.
- Ana Valeria Gonzalez, **Isabelle Augenstein**, Anders Søgaard. *Retrieval-Based Goal-Oriented Dialogue Generation*. Proceedings of the 3rd Conversational AI Workshop at NeurIPS 2019, December 2019
- Ana Valeria Gonzalez, **Isabelle Augenstein**, Anders Søgaard. *Retrieval-Based Goal-Oriented Dialogue Generation*. Proceedings of the 3rd Conversational AI Workshop at NeurIPS 2019, December 2019
- Joachim Bingel, Victor Petrén Bach Hansen, Ana Valeria Gonzalez, Paweł Budzianowski, **Isabelle Augenstein**, Anders Søgaard (2019). *Domain Transfer in Dialogue Systems without Turn-Level Supervision*. Proceedings of the 3rd Conversational AI Workshop (ConvAI at NeurIPS 2019).

- Johannes Bjerva, Katharina Kann, **Isabelle Augenstein** (2019). *Transductive Auxiliary Task Self-Training for Neural Multi-Task Models*. Proceedings of the 2nd Workshop on Deep Learning for Low Resource Natural Language Processing (DeepLo at EMNLP 2019).
- Mostafa Abdou, Cezar Sas, Rahul Aralikkatte, **Isabelle Augenstein**, Anders Søgaard (2019). *X-WikiRE: A Large, Multilingual Resource for Relation Extraction as Machine Comprehension*. Proceedings of the 2nd Workshop on Deep Learning for Low Resource Natural Language Processing (DeepLo at EMNLP 2019).
- 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).
- 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

- Luna De Bruyne, Pepa Atanasova, **Isabelle Augenstein** (2021). *Joint Emotion Label Space Modelling for Affect Lexica*. Computer Speech & Language, to appear.
- Johannes Bjerva, Robert Östling, Maria Han Veiga, Jörg Tiedemann, **Isabelle Augenstein**. *What do Language Representations Really Represent?*. Computational Linguistics, Vol. 45, No. 2, June 2019.
- 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

- Karolina Stańczak, Sagnik Ray Choudhury, Tiago Pimentel, Ryan Cotterell, *Isabelle Augenstein*. Quantifying Gender Bias Towards Politicians in Cross-Lingual Language Models. CoRR, abs/2104.07505, April 2021.

- **Isabelle Augenstein**, Ivan Habernal (2021). *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Tutorial Abstracts (EACL 2021)*, April 2021.
- Momchil Hardalov, Arnav Arora, Preslav Nakov, **Isabelle Augenstein** (2021). *Cross-Domain Label-Adaptive Stance Detection*. CoRR, abs/2104.07467, April 2021.
- Sheikh Muhammad Sarwar, Dimitrina Zlatkova, Momchil Hardalov, Yoan Dinkov, **Isabelle Augenstein**, **Preslav Nakov** (2021). *A Neighbourhood Framework for Resource-Less Content Flagging*. CoRR, abs/2103.17055, March 2021.
- Momchil Hardalov, Arnav Arora, Preslav Nakov, **Isabelle Augenstein** (2021). *A Survey on Stance Detection for Mis- and Disinformation Identification*. CoRR, abs/2103.00242, February 2021.
- Preslav Nakov, Vibha Nayak, Kyle Dent, Ameya Bhatawdekar, Sheikh Muhammad Sarwar, Momchil Hardalov, Yoan Dinkov, Dimitrina Zlatkova, Guillaume Bouchard, **Isabelle Augenstein** (2021). *Detecting Abusive Language on Online Platforms: A Critical Analysis*. CoRR, abs/2103.00153, February 2021.
- Nils Rethmeier, **Isabelle Augenstein** (2021). *A Primer on Contrastive Pretraining in Language Processing: Methods, Lessons Learned and Perspectives*. CoRR, abs/2102.12982, February 2021.
- Tamar Solorio, Mahsa Shafaei, Christos Smallis, **Isabelle Augenstein**, Margaret Mitchell, Ingrid Stapf, Ioannis Kakadiaris (2021). *White Paper - Creating a Repository of Objectionable Online Content: Addressing Undesirable Biases and Ethical Considerations*. OpenReview Preprint, February 2021.
- Andreas Nugaard Holm, Barbara Plank, Dustin Wright, **Isabelle Augenstein** (2020). *Longitudinal Citation Prediction using Temporal Graph Neural Networks*. CoRR, abs/2012.05742, December 2020.
- Andrea Lekkas, Peter Schneider-Kamp, **Isabelle Augenstein** (2020). *Multi-Sense Language Modelling*. CoRR, abs/2012.05776, December 2020.
- Nils Rethmeier, **Isabelle Augenstein** (2020). *Data-Efficient Pretraining via Contrastive Self-Supervision*. CoRR, abs/2010.01061, October 2020.
- Liesbeth Allein, **Isabelle Augenstein**, Marie-Francine Moens (2020). *Time-Aware Evidence Ranking for Fact-Checking*. CoRR, abs/2008.09112, September 2020.
- Zeerak Waseem, Smarika Lulz, Joachim Bingel, **Isabelle Augenstein** (2020). *Disembodied Machine Learning: On the Illusion of Objectivity in NLP*. OpenReview Preprint, June 2020.
- **Isabelle Augenstein**, Spandana Gella, Sebastian Ruder, Katharina Kann, Burcu Can, Alexis Conneau, Johannes Welbl, Xian Ren and Marek Rei. *Proceedings of The Fourth Workshop on Representation Learning for NLP*. (Repl4NLP at ACL 2019).
- **Isabelle Augenstein**, Sameer Singh, Andrew McCallum. *Proceedings of The First Conference on Automatic Knowledge Base Construction*. (AKBC 2019).

- **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.
- 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

- 1st place in ConLL-SIGMORPHON 2018 task on Morphological Inflection in Context
- 3rd place out of 50 scored systems in 2017 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 AND PANELS

- *Determining the Credibility of Science Communication*. Second Workshop on Scholarly Document Processing at NAACL 2021, June 2021. *Invited talk, online*

- *Careers in NLP* Panel at NAACL 2021, June 2021. *Online panel*
- *What should we be teaching?* Panel at Teaching NLP workshop at NAACL 2021, June 2021. *Online panel*
- *How to pursue a career in academia/industry* Panel at Mexican NLP Summer School 2021, June 2021. *Online panel*
- *Evaluating Automatically Generated Fact Checking Explanations*. Distinguished Speakers: Oxford Women in Computer Science, May 2021. *Invited talk, online*
- *On the different stages of adaptation in domain adaptation*. Adapt-NLP: The Second Workshop on Domain Adaptation for NLP at EACL 2021, April 2021. *Invited talk, online*
- *Toward Explainable Fact Checking*. University of Copenhagen; DIKU Business Club webinar on ‘Fighting Harmful Content on the Web’, March 2021. *Invited talk, online*
- *Toward Explainable Fact Checking*. University of Cambridge; Language Technology Lab (LTL) seminar, January 2020. *Invited talk, online*
- *Toward Explainable Fact Checking*. Law, Economics, and Data Science Group at ETH Zurich, December 2020. *Invited talk, online*
- *What can ML do for knowledge graphs – and what can graphs do for ML?* Knowledge Connexions panel at Knowledge Graph Conference, December 2020. *Online panel*
- *Explainable Fact Checking as Structured Prediction*. Structured Prediction for NLP workshop at EMNLP 2020, November 2020. *Invited talk, online*
- *Towards Explainable Fact Checking*. ACL FEVER Workshop Invited Talk, July 2020. *Invited talk, online*
- *Towards Explainable Fact Checking*. University of Cambridge, NLIP Seminar Series, June 2020. *Invited talk, online*
- *Towards Explainable Fact Checking*. Google Research & Google News Seminar, June 2020. *Invited talk, online*
- *What can typological knowledge bases and language representations tell us about linguistic properties?*. University of Zurich, Department of Informatics Colloquium, March 2020. *Invited talk, all expenses paid*
- *Tracking False Information Online*. TU Dresden, Computer Science Colloquium, February 2020. *Invited talk, all expenses paid*
- *Gender Bias in Language*. UCPH task force on unconscious bias in recruitment and retention, December 2019.

- *Tracking False Information Online*. IT University Copenhagen, NLP Seminar Series, November 2019.
- *MultiFC: A Real-World Multi-Domain Dataset for Evidence-Based Fact Checking of Claims.*. Copenhagen Business School, Seminar on Fact Checking, November 2019.
- *Tracking False Information Online*. 5th Workshop on Noisy User-generated Text (WNUT) at EMNLP-IJCNLP 2019, Hong Kong, November 2019.
- *What can typological knowledge bases and language representations tell us about linguistic properties?*. Typology for Polyglot NLP workshop at ACL, Florence, Italy, August 2019.
- *Natural Language Understanding with Limited Training Data*. Institute of Informatics, Hamburg, Germany, April 2019. *Invited talk, all expenses paid*
- *Natural Language Understanding with Limited Training Data*. QUARTZ ITN workshop, Copenhagen, Denmark, April 2019. *Invited talk*
- *Natural Language Understanding with Limited Training Data*. NLP Meetup Series, Uppsala, Sweden, March 2019. *Invited talk, all expenses paid*
- *Natural Language Understanding with Limited Training Data*. University of Copenhagen SCIENCE AI Centre, Copenhagen, Denmark, March 2019. *Invited talk*
- *Natural Language Understanding with Limited Training Data*. Heidelberg Institute for Theoretical Studies, Heidelberg, Germany, January 2019. *Invited talk, all expenses paid*
- *Learning with limited labelled data in NLP: multi-task learning and beyond*. Johns Hopkins University, Center for Language and Speech Processing, Baltimore, Maryland, USA, June 2018. *Invited talk, all expenses paid*
- *Learning with limited labelled data in NLP: multi-task learning and beyond*. New York University, Machine Learning for Language group, New York City, New York, USA, June 2018. *Invited talk*
- *Learning with limited labelled data in NLP: multi-task learning and beyond*. 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, Copenhagen, Denmark, 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.

ORGANISING COMMITTEES

2021	3rd Conference for Truth and Trust Online (TTO 2021), Program Co-Chair
2021	16th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2021), Tutorial Co-Chair and Session Chair
2021	1st Workshop on Natural Language Processing by the European Laboratory for Learning and Intelligent Systems (ELLIS NLP 2021), Session Chair
2020	24th Conference on Empirical Methods in Natural Language Processing (EMNLP 2020), Diversity & Inclusion Co-Chair
2020	13th NIST Text Analysis Conference 2020, Knowledge Base Population track (TAC KBP), Scientific Advisory Board Member
2020	5th Workshop on Representation Learning for NLP (Repl4NLP) at ACL 2020, Advisory Board Member
2020	4th WiNLP Workshop on Widening NLP at ACL 2020, Advisory Board Member
2019	4th Workshop on Representation Learning for NLP (Repl4NLP) at ACL 2019, Primary Organiser
2019	3rd WiNLP Workshop on Widening NLP at ACL 2019, Advisory Board Member
2019	2nd Workshop on Learning with Limited Labeled Data: Weak Supervision and Beyond (LLD 2019) at ICLR 2019, Organiser
2019	1st Conference on Automated Knowledge Base Construction (AKBC 2019), Program Co-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
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

AREA CHAIRING

2021	2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021), Senior Area Chair for Machine Learning
2021	59th Annual Meeting of the Association for Computational Linguistics (ACL 2021), Area Chair for Special Theme Track “NLP for the Social Good”
2021	18th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2021), Area Chair for Machine Learning for NLP: Classification and Structured Prediction Models
2020	58th Annual Meeting of the Association for Computational Linguistics (ACL 2020), Senior Area Chair for Machine Learning
2020	2nd Conference on Automated Knowledge Base Construction (AKBC 2020), Area Chair
2019	22rd Conference on Empirical Methods in Natural Language Processing (EMNLP 2019), Area Chair for Machine Learning
2019	57th Annual Meeting of the Association for Computational Linguistics (ACL 2019), Area Chair for Information Extraction and Text Mining
2019	17th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2019), Area Chair for Sentiment Analysis

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
2014	2nd University of Sheffield Engineering Symposium (USES 2014), Artificial Intelligence & Virtual Reality Track Chair

JOURNAL REVIEWING

since 05/2021	Action editor of “ACL Rolling Review”
since 07/2020	Editorial board member of the journal “Computer Speech & Language”
2020-2021	Editorial board member of the Northern European Journal of Language Technology (NEJLT)
since 2016	Transactions of the Association for Computational Linguistics (TACL), ad-hoc reviewer since 2016, standing reviewer since 2020
since 2016	Computational Linguistics (CL), ad-hoc reviewer
2020	KÄ¼nstliche Intelligenz - Special Issue on NLP and Semantics (KI), ad-hoc reviewer
2016	Information Processing & Management (IP&M), ad-hoc reviewer

PROGRAMME COMMITTEES

- Programme Committee Member for Conferences: ICLR 2019, NeurIPS 2019, NeurIPS 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 ACL 2020, FEVER at EMNLP 2019, 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)

GRANT PROPOSAL REVIEWING

2021	University of Copenhagen, Data+ programme
2021	University of Cambridge, Trinity College Junior Research Fellowships programme
2020-2021	European Language Grid (ELG), a Horizon 2020 research and innovation programme, https://www.european-language-grid.eu/
2020	EDGE Fellowship Programme, a H2020 MSCA Cofund action, https://edge-research.eu/
2020-2021	ELITE-S Fellowship Programme, a H2020 MSCA Cofund action, http://elite-fellowship.eu/
2019	Riksbankens Jubileumsfond (RJ), https://www.rj.se/

OTHER

- 18th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2021), Best Paper Awards Committee Member
- 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2019), Best Paper Awards Committee Member
- President of the newly created ACL Special Interest Group on Representation Learning (SIGREP), October 2018 - September 2022
- Co-creator and maintainer of the BIG Directory of members of underrepresented groups and supporters in Natural Language Processing, since July 2018
- Co-organiser of the Copenhagen NLP meetup, since July 2017

PROFESSIONAL MEMBERSHIPS

since 2021	Young Royal Danish Academy of Sciences and Letters (DUA), member
since 2020	European Lab for Learning and Intelligent Systems (ELLIS), Scholar, founding faculty of ELLIS CPH uni, founding faculty of ELLIS NLP program
since 2020	ACL ACL Special Interest Group on Typology (SIGTYP), member
since 2018	ACL Special Interest Group on Representation Learning (SIGREP), president and co-founder

since 2015 Association for Computational Linguistics, member

UNIVERSITY INTERNAL SERVICES

since 2021 Head of the Natural Language Processing section at the Department of Computer Science, University of Copenhagen

2020 Machine Learning section responsible for teaching coordination at the Department of Computer Science, University of Copenhagen

since 2017 On panels for recruitment and selection of University of Copenhagen Computer Science staff

2019 On committee for streamlining the tenure-track process at the University of Copenhagen

2017 On committee for communications at the University of Copenhagen

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

01/2021 Advanced Language Processing School (ALPS) 2021, teacher, tutorial on Explainability in NLP

07/2020 10th Lisbon Machine Learning School (LxMLS) 2020, invited talk on Explainability in NLP

2020 Guest lecturer at Department of Computer Science, University of Copenhagen
Module: Data Science. Designed lecture materials.

2019 Co-lecturer at Department of Computer Science, University of Copenhagen
Module: Advanced Topics in Natural Language Processing. Co-developed a new course and designed new lecture materials.

since 2019 Co-lecturer at Department of Computer Science, University of Copenhagen
Module: Web Science. Re-designed lecture materials.

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

- Miryam de Lhoneux. *Typologically informed dependency parsing*. UCPH Postdoctoral Researcher, advisor, since February 2021.
- Shailza Jolly. *Generating fact-checking explanations in low-resource settings*. UCPH Visiting PhD Student, advisor, January - May 2021.
- Sagnik Ray Choudhury. *Explainable Question Answering*. UCPH Postdoctoral Researcher, advisor, since July 2020.
- Karolina Stańczak. *Unsupervised Cross-Lingual Detection of Gender Biases*. UCPH PhD Student, advisor, since June 2020.
- Oscar Kjell. *Conceptualizing and Measuring Well-Being Using Statistical Semantics*. UCPH Postdoctoral Researcher, advisor, since January 2020.
- Dustin Wright. *Data Enrichment to Advance Scientific Information Extraction*. UCPH PhD Student, advisor, since October 2019.
- Andreas Nugaard Holm. *Transfer Learning of Textual Representations in the Life Science Domain*. UCPH PhD Student, advisor, since September 2019.
- Nils Rethmeier. *Interpretability of Black-Box Representation Learning Models*. UCPH PhD Student, advisor, since August 2019.

- Pepa Atanasova. *Explainability for Fact Checking*. UCPH PhD Student, advisor, since March 2019.
- Wei Zhao. *Understanding cross-lingual language representations*. UCPH Visiting PhD Student, advisor, January - March 2021.
- Liesbeth Allein. *Time-Aware Evidence Ranking for Fact-Checking*. UCPH Visiting PhD Student, advisor, February - March 2021.
- Farhad Nooralahzadeh. *Domain Adaptation for Information Extraction*. UCPH Visiting PhD Student, advisor, May - July 2019.
- Luna De Bruyne. *Weakly Supervised Emotion Recognition*. UCPH Visiting PhD Student, advisor, April - June 2019.
- Giannis Bekoulis. *Joint Information Extraction*. UCPH Visiting PhD Student, advisor, March - May 2019.
- Andrea Lekkas. *Advanced Language Models for Writing Assistants*. SDU PhD Student, co-advisor, May 2019 - March 2021.
- Johannes Bjerva. *Low-Resource Learning in Natural Language Processing*. UCPH Postdoctoral Researcher, advisor, July 2017 - August 2020.
- Ana Valeria Gonzalez. *Multilingual question answering for customer service bots*. UCPH PhD Student, co-advisor, thesis submitted in February 2021.
- Yova Kementchedjhieva. *Morphological Segmentation with Distant Supervision*. UCPH PhD Student, co-advisor, thesis submitted in March 2021.
- Mareike Hartmann. *Detecting Disinformation with Multilingual Stance Detection*. UCPH PhD Student, co-advisor from September 2017 onwards, thesis submitted in October 2019.
- 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, February 2018 - July 2018.
- Georgios Spithourakis. *Numeracy of Language Models: Joint Modelling of Words and Numbers*. PhD in Computer Science, thesis submitted in July 2018, 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, thesis submitted in October 2019, co-advised with Maria Liakata at the University of Warwick via the Alan Turing Institute.

SUPERVISION OF MASTER'S AND BACHELOR'S STUDENTS

- Wojciech Ostrowski. *Generating Fact-Checking Explanations*. UCPH MSc in Computer Science Dissertation, co-supervised by Pepa Atanasova, submitted in 2020.
- Lukas Muttenthaler. *Subjective Question Answering*. UCPH MSc in IT & Cognition Dissertation, co-supervised by Johannes Bjerva, submitted in 2020.

- Marcus Hansen. *Text Mining and Processing on Patients Experiences on Psilocybin*. UCPH BSc in Computer Science Dissertation, co-supervised by Melanie Ganz-Benjaminsen, submitted in 2020.
- Jimmie Jin, Asger Thorleif Knudsen and Sylvester Lee. *Stance Detection of Climate Change Tweets*. UCPH BSc in Computer Science Dissertation, co-supervised by Dustin Wright, submitted in 2020.
- Sebastian Hansen, Bjørn Møller and William Hansen. *Graph Convolutions on Co-Author Networks*. UCPH BSc in Computer Science Dissertation, co-supervised by Andreas Holm, submitted in 2020.
- Mihai Popovici. *Stance Detection in Scientific Reviews*. UCPH MSc in Computer Science Dissertation, co-supervised by Rahul Aralikkatte, submitted in 2019.
- Yi He. *Stance Detection in Scientific Reviews*. UCPH MSc in Computer Science Research Project and Dissertation, co-supervised by Rahul Aralikkatte, submitted in 2019.
- Haining Tong. *Zero-Shot Relation Extraction via Description Learning*. UCPH MSc in Computer Science Research Project and Dissertation, co-supervised by Matt Lamm, submitted in 2019.
- Zhong Xuan. *Zero-Shot Relation Extraction using Graph Neural Networks*. UCPH MSc in Computer Science Research Project, co-supervised by Raghavendra Selvan, submitted in 2019.
- Raluca Alexandra Fetic. *Efficient encoding of large documents for stance detection*. UCPH MSc in Computer Science Dissertation, submitted in 2019.
- Peter Spliid. *A Framework for Typological Evaluation of Language Representations*. UCPH MSc in Computer Science Dissertation, submitted in 2018.
- Magnus Alexander Johansen. *Learning to ask questions about products*. UCPH MSc in Computer Science Dissertation, submitted in 2018.
- Andrea Lekkas. *Learning to ask questions about products*. UCPH MSc in Computer Science Dissertation, submitted in 2018.
- Leendert Bastian van Doorn. *Variational encoding of location for dialect-conditional language generation*. UCPH MSc in Computer Science Research Project, co-supervised by Johannes Bjerva, submitted 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, submitted in 2018.
- Ke Zhai. *Deep Learning of Linguistic Features for Scandinavian Languages*. UCPH MSc in Computer Science Research Project, co-supervised by Johannes Bjerva, submitted in 2018.
- Thomas Nyegaard-Signorla and Casper Veistrup Helms. *Modelling affect in tweets*. UCPH BSc in Computer Science Dissertation, joint project, co-supervised by Johannes Bjerva, submitted in 2018.
- Rogan Inglis. *Machine Reading for Scientific Publications Using Generative Regularisation*. UCL MSc in Machine Learning Dissertation, co-supervised by 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

2021 – 2025	DDF (Independent Research Fund Denmark) Sapere Aude Research Leader Grant for project ‘Learning to Explain Attitudes on Social Media (EXPANSE)’, PI. DKK 7,950,000.
2021 – 2024	University of Copenhagen Data+ Grant for project ‘Privacy Black & White’, PI. DKK 2,000,000.
2021 – 2024	University of Copenhagen Data+ Grant for project ‘Gendered Democratic Participation on Social Media’, co-PI. DKK 915,000.
2021 – 2024	University of Copenhagen Data+ Grant for project ‘Error and Accountability in Digitised Public Decision Making’, co-PI. DKK 1,730,000.
2021 – 2024	Swedish Research Council International Postdoc Grant for project ‘Typologically informed dependency parsing’, host.
2020 – 2023	Swedish Research Council International Postdoc Grant for project ‘Conceptualizing and Measuring Well-Being Using Statistical Semantics’, host.
2020 – 2024	Swedish Research Council Grant for project ‘Structured Multilinguality for Natural Language Processing’, co-PI. SEK 4,000,000.
2020 – 2023	Independent Research Fund Denmark Grant for project ‘Detecting Gender-Biased Language on Social Media’, PI. DKK 2,875,582.
2019 – 2022	COST Action from the European Commission, networking and training funding for 23-partner project ‘Multi3Generation: Multi-task, Multilingual, Multi-modal Language Generation’, project coordinator. EUR 520,000.
2019 – 2022	H2020-MSCA-COFUND Grant from the European Commission, partial PhD fellowship for one PhD student for project ‘Data Enrichment to Advance Scientific Information Extraction’, co-PI. DKK 940,000.

- 2019 – 2022 **Innovation Fund Denmark Grant** for project ‘Transfer Learning of Textual Representations in the Life Science Domain’, PI. Funding for one industrial PhD with Base Life Science. DKK 972,000
- 2019 – 2020 **Innovation Fund Denmark Grant** for project ‘Improving the democratic debate in a time of fake news’, PI. Funding for a consultancy on an InnoBooster project with ULobby. DKK 400,000.
- 2019 **Megagon Labs Research Award** for project ‘Subjective Question Answering’, PI. USD 106,162.
- 2018 **H2020-MSCA-COFUND Grant** from the European Commission, partial PhD fellowship for one PhD student for project ‘Tailoring knowledge for machine reading comprehension in both rich and scarce in resource languages and domains’, co-PI. DKK 940,000.
- 2018 **Innovation Fund Denmark Grant** for project ‘Advanced Language Models for Writing Assistants (ALMA)’, co-PI. Funding for one industrial PhD with Ordbogen A/S. DKK 972,000.
- 2018 **EUopSTART Grant** from Ministry of Higher Education and Science, Denmark for project ‘Accessibility Increased through Real-life Multilingual Artificial Intelligence (AIRMAIL)’, PI. DKK 75,000.
- 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
- 2018 **NVIDIA GPU Grant** for project ‘Understanding Online Discourse with Conversational Stance Detection’, September 2018, PI. 1 NVIDIA TITAN V GPU
- 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

- 2019 **DIKU Dissemination Award 2019**, DKK 20,000 cash prize
- 2019 **Megagon Labs Research Award** for project ‘Subjective Question Answering’, PI. USD 106,162.
- 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/MLO/>
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

EXTERNAL HIRING COMMITTEES

2021 University of Oslo, Department of Informatics,
Assessment committee for Associate Professorship

PHD COMMITTEES

2021-2022 Aditi Chaudhary, *Automated Extraction of Grammatical Specifications for Under-Resourced Languages*,
PhD thesis committee at Carnegie Mellon University, USA

2021 Michael Schlichtkrull, *Incorporating Structure into Neural Models for Language Processing*,
University of Amsterdam, the Netherlands

2021 Simon Hellemann Flachs, *Computational Grammatical Error Correction: Bridging the Gap from Academia to Industry*,
University of Copenhagen, Denmark

2019 Luiza Sayfullina, *Machine Learning methods for classification of unstructured data*,
Aalto University, Finland

2019 Giannis Bekoulis, *Neural Approaches to Sequence Labeling for Information Extraction*,
University of Ghent, Belgium

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

MEDIA

2020 “Forskere mener, de blev fyret af Google, fordi de påpegede racisme. Nu går flere i vrede”. Politiken article discussing Google’s firing of prominent Ethical AI researchers, 26 February 2021, <https://tinyurl.com/468kvp7z>

2020 “Nyt studie peger på flere udfordringer med peer review”. Science Report article discussing our EMNLP Findings paper, 2 December 2020, <https://sciencereport.dk/ny-viden/nyt-studie-peg-er-paa-flere-udfordringer-med-peer-review/>

- 2020 “How Can We Improve Peer Review in NLP?”. Ms Coffee bean YouTube video discussing our EMNLP Findings paper, 24 November 2020, <https://www.youtube.com/watch?v=ttDkpIq-sD4&feature=youtu.be>
- 2020 “How Can We Improve Peer Review in NLP?”. The Gradient article discussing our EMNLP Findings paper, 14 November 2020, <https://thegradient.pub/how-can-we-improve-peer-review-in-nlp/>
- 2019 Gender Stereotypes And Sexist Language In Literature. **75+ international news outlets** reporting the findings of our ALC 2019 paper ‘Unsupervised Discovery of Gendered Language through Latent-Variable Modeling’. Link to the Forbes article: <https://www.forbes.com/sites/jessedamiani/2019/08/30/massive-machine-learning-study-demonstrates-gender-stereotyping-and-sexist-language-in-literature>
- 2017 “Organizing the SemEval task on scientific information extraction”. Allen Institute for Artificial Intelligence podcast ‘NLP Highlights’, 27 October 2017, <https://soundcloud.com/nlp-highlights/39-organizing-the-semeval-task-on-scientific-information-extraction-with-isabelle-augenstein>
- 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, 25 June 2021