Pontus Stenetorp

University College London
Gower Street
London WC1E 6BT, UK
📞 +44 (0)20 3549 5395
🌐 pontus.stenetorp.se

Education

PhD, *University of Tokyo*, Computer Science. **2010–2013**
"On the Applicability and Representation of Lexical Prior Knowledge"

MSc Eng, *Royal Institute of Technology (KTH)*, Computer Science. **2005–2010**
"Automated Extraction of Swedish Neologisms using a Temporally Annotated Corpus"

Work experience

Senior Research Associate, *University College London*, London, United Kingdom. **2017–**
Multi-step inference using deep memory-based models.

Research Associate, *University College London*, London, United Kingdom. **2015–2017**
Models learning from a minimal amount of human supervision.

Postdoctoral Research Fellow, *University of Tokyo*, Tokyo, Japan. **2013–2015**
Joint learning of latent categories, structures, and their representations.

Special Project Research Associate, *National Institute of Informatics*, Tokyo, Japan. **2013**
Vector composition and representation learning for dependency parsing.

Teaching Assistant, *Royal Institute of Technology (KTH)*, Stockholm, Sweden. **2006–2009**
Weekly lectures for up to 15 students for several classes.

Quality-Assurance Engineer, *LucidEra Inc.*, San Mateo, California, USA. **2008**
Performance testing and test automation.

Software Developer, *Norrorts Konferens HB*, Djursholm, Sweden. **2006–2007**
Software development for conference activities.

Achievements

Funding

Grant, *Centre for Doctoral Training in Data Intensive Science PhD Studentship Grant, Co-PI, GBP 80,000.* **2017–2021**

Fellowship, *Japan Society for the Promotion of Science (JSPS), JPY 8,688,000.* **2013–2015**

Grant, *KAKENHI #13F03041, PI, JPY 2,300,000.* **2013–2015**

Scholarship, *Monbukagakusho (MEXT), JPY 1,740,000.* **2012–2013**

Stipend, *University of Tokyo Overseas Visit Support, JPY 220,000.* **2012**

Scholarship, *Japan Student Services Organization (JASSO), JPY 780,000.* **2011–2012**

Stipend, *Royal Swedish Academy of Sciences (KVA), SEK 100,000.* **2011**

Stipend, *Royal Swedish Academy of Sciences (KVA), SEK 100,000.* **2010**

Stipend, *Lars Hierta Memorial Foundation, SEK 35,000.* **2010**

Stipend, *Ebba Danelius Foundation, SEK 120,000.* **2009**

Stipend, *Adolf Lindgren Foundation, SEK 16,000.* **2008**

Awards

Outstanding paper, *15th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2017).*	**2017**
Best presentation, *JSPS London Research Promotion Conference.*	**2016**
Best proposal, *Facebook AI Research (FAIR), USD 1,250.*	**2016**
Nomination, *Gerd Eriksson Teaching Award, KTH.*	**2007**
Nomination, *Gerd Eriksson Teaching Award, KTH.*	**2006**

Services

Organising committee

2017: Conference on Empirical Methods in Natural Language Processing (EMNLP)

2016: *Programme Chair*, JuliaCon 2016 (274 attendees)

2015: Biomedical Linked Annotation Hackathon (BLAH)

2013: BioNLP Shared Task (22 groups participating)

2011: BioNLP Shared Task (24 groups participating)

Programme committee

2017: BioNLP, EACL, EMNLP, Ethics in NLP, Games4NLP, IJCNLP, JuliaCon, WiNLP, and RepEval

2016: AKBC, COLING, EMNLP, IJCAI, JuliaCon, and SCIDOCA

2015: EMNLP and JuliaCon

2013: BioNLP and IJCNLP

2012: BioNLP and SMBM

Departmental services

2015–: *Hiring committee member*, University College London (3 hires)

2015–: *PhD student screening committee member*, University College London

Other

2015–: *Lead Organiser*, South England Natural Language Processing (SE NLP) Meetup

2016: *Organiser*, Google Summer of Code (GSoC) for the Julia programming language

2013: *Local Organiser*, The 5th International Symposium on Languages in Biology and Medicine

Teaching

Classes

Lecturer, *University College London*, London, UK.	**2015–**

Research Methods (COMP3095) and Research Group Project (COMP3096). Weekly project-based classes.

Guest lecturer, *University College London*, London, UK.	**2015–**

Statistical Natural Language Processing (COMPM083/COMPGI19). Guest lecturer for deep learning portion of course.

Teaching Assistant, *Royal Institute of Technology (KTH)*, Stockholm, Sweden.   **2005–2008**
Algorithms, Data Structures and Complexity (DD1352), Introduction to Computer Science (DD1339), Programming Paradigms (DD1361), and Programming Techniques (DD1310). Lab assistance, homework grading, and flipped classroom lectures with up to 15 students.

Supervision

Co-supervisor: *University College London*, Tom Crossland, *PhD*, 2021 (expected).

Co-supervisor: *University College London*, Johannes Welbl, *PhD*, 2018 (expected).

Supervisor: *University College London*, Sudhanshu Kasewa, *MSc*, 2017 (expected).

Supervisor: *University College London*, Jonas Vetterle, *MSc*, 2017 (expected).

Supervisor: *University College London*, Iuliana-Elena Parasca, *MEng*, 2017, now at Google Zürich.

Supervisor: *University College London*, Aleksandar Rusinov, *MEng*, 2017, now at Morgan Stanley.

Co-supervisor: *Tohoku University*, Shimaoka Sonse, *MSc*, 2017, now at Recruit Holdings Co, Ltd.

Supervisor: *University College London*, Jonathan Godwin, *MSc*, 2016, now at Bloomsbury AI.

Supervisor: *University College London*, Mark Neumann, *MSc*, 2016, now at the Allen Institute for Artificial Intelligence.

Invited talks

Machines that read and reason, *Language, Knowledge and People in Perspective Workshop*, Leiden, Netherlands.   **2017**
Invited talk, paid travel and accommodation.

Machine Learning Communities Day, *Seldon/Google*, London, UK.   **2017**
Invited participant.

Learning to Generate Training Data, *University of Cambridge*, Cambridge, UK.   **2017**
Invited talk, paid travel.

Weakly Supervised Machine Reading, *CREST International Symposium on Big Data Application*, Tokyo, Japan.   **2017**
Invited talk, with Isabelle Augenstein.

Weakly Supervised Machine Reading, *Nara Institute of Science and Technology*, Ikoma, Japan.   **2017**
Invited talk, with Isabelle Augenstein.

Learning to Generate Textual Data, *National Institute of Advanced Industrial Science and Technology*, Tokyo, Japan.   **2017**
Invited talk.

Learning to Generate Textual Data, *Tohoku University*, Sendai, Japan.   **2016**
Invited talk, paid travel and accommodation.

Learning to Generate Textual Data, *National Institute of Informatics*, Tokyo, Japan.   **2016**
Invited talk.

Machines that read the scientific literature, *Japan Society for the Promotion of Science (JSPS) London Japan-UK Research Promotion Conference*, Embassy of Japan, London, UK.   **2016**
Invited talk, best presentation award.

Natural Language Processing with Julia, *Manchester Julia Workshop*, University of Manchester, Manchester, UK.   **2016**

Plenary talk, paid travel and accommodation.

Natural Language Processing: History, motivations, and what the near future may hold for intelligent machines, *Waseda University*, Tokyo, Japan. Invited talk.	**2016**
Learning to Generate Textual Data, *AI4Exam Workshop*, National Institute of Informatics, Tokyo, Japan. Invited talk, paid travel and accommodation.	**2016**
Defining Words with Words, *Kyoto University*, Kyoto, Japan. Invited talk.	**2016**
Defining Words with Words, *Osaka University*, Osaka, Japan. Invited talk.	**2016**
Computers that Read: Uncovering the Structure of Language with Deep Learning, *Artificial Intelligence Research (AIR) society*, Osaka University, Osaka, Japan. Invited talk, with Jason Naradowsky.	**2016**
Defining Words with Words, *Tohoku University*, Sendai, Japan. Invited talk, paid travel and accommodation.	**2016**
Defining Words with Words, *University of Tokyo*, Tokyo, Japan. Invited talk.	**2016**
Leveraging Monolingual Data for Crosslingual Compositional Word Representations, *Alan Turing Institute Deep Learning Workshop*, Edinburgh, Scotland. Invited poster.	**2015**
Natural Language Processing, *Google NLP PhD Summit*, Zürich, Switzerland. Invited participant, paid travel and accommodation.	**2015**
Suitably naming a child with multiple nationalities using Julia, *JuliaCon*, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. Invited talk.	**2015**
Unified Annotation Tooling, *National Science Foundation-Sponsored Workshop*, Sunny Isles Beach, Florida, USA. Invited participant, paid travel and accommodation.	**2015**
Computers and Language, *Yamanashi Prefectural Hikawa High School*, Yamanashi, Japan. Invited talk, paid travel.	**2014**
A Brief JuliaCon Report and Natural Language Processing with Julia, *Julia-Tokyo Meetup*, Tokyo, Japan. Invited talk.	**2014**
Natural Language Processing with Julia, *JuliaCon*, University of Chicago, Chicago, Illinois, USA. Invited talk.	**2014**
A Year and a Few Months Later: Results and Musings on the State of NLP and BioNLP, *Microsoft Research*, Seattle, Washington, USA. Invited talk.	**2012**
Assisting Gene Mention Tagging by Utilising Largely Unsupervisedly Induced Semantic Class Taggers, *Tech Talks 'n Mix*, Google, Tokyo, Japan. Invited poster.	**2010**

Commissions of Trust

Chairman, *KTH Japan Alumni Chapter*, Tokyo, Japan.   **2013–2015**

Chairman of the Education Committee, *Computer Science Chapter, Student Union*   **2008**
at the Royal Institute of Technology, Stockholm, Sweden.

Member of the Board, *Computer Science Chapter, Student Union at the Royal*   **2006–2007**
Institute of Technology, Stockholm, Sweden.

Municipality Chairman, *Piratpartiet (Pirate Party)*, Stockholm, Sweden.   **2006**

Certificates

Test of English for International Communication (TOEIC), *990/990*.   **2013**

Test of English as a Foreign Language (TOEFL), *120/120*.   **2009**

Certificate of Advanced English (CAE), *Grade A, University of Cambridge*.   **2004**

Language skills

Swedish: Native

English: Fluent

Japanese: Intermediate

German: Elementary

Open source software and contributions

Creator: *brat*, a widely adopted text annotation tool.

Contributor: *Julia*, modern open-source programming language for technical computing.

A complete list of public software repositories can be found at: https://github.com/ninjin

Publications

32. Sonse Shimaoka, Pontus Stenetorp, Kentaro Inui, and Sebastian Riedel. **Neural Architectures for Fine-grained Entity Type Classification** *In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers, Valencia, Spain, April 2017.* (**Outstanding paper award**, **acceptance rate: 27%**)

31. Galen Han, Mark Menezes, Leah Halaseh, Jędrzej Stuczyński, Daniele Menara, Sebastian Riedel, and Pontus Stenetorp. **Towards a Word Sheriff 2.0: Lessons learnt and the road ahead** *Games4NLP Symposium, Valencia, Spain, April 2017.*

30. Jonathan Godwin, Pontus Stenetorp, and Sebastian Riedel. **Deep Semi-Supervised Learning with Linguistically Motivated Sequence Labeling Task Hierarchies** *arXiv, December 2016.*

29. Neumann Mark, Pontus Stenetorp, and Sebastian Riedel. **Learning to Reason with Adaptive Computation** *Interpretable Machine Learning for Complex Systems at the 2016 Conference on Neural Information Processing Systems (NIPS), Barcelona, Spain, December 2016.*

28. Guillaume Bouchard, Pontus Stenetorp, and Sebastian Riedel. **Learning to Generate Textual Data** *In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, Texas, November 2016.* (**Acceptance rate: 26%**)

27. Iuliana-Elena Parasca, Andreas Lukas Rauter, Jack Roper, Aleksandar Rusinov, Guillaume Bouchard, Sebastian Riedel, and Pontus Stenetorp. **Defining Words with Words: Beyond the Distributional Hypothesis** *In Proceedings of the 1st Workshop on Evaluating Vector-Space Representations for NLP, Berlin, Germany, August 2016.* (**Best proposal award**)

26. Sonse Shimaoka, Pontus Stenetorp, Kentaro Inui, and Sebastian Riedel. **An Attentive Neural Architecture for Fine-grained Entity Type Classification** *In Proceedings of the 5th Workshop on Automated Knowledge Base Construction, San Diego, California, USA, June 2016.* (**Citations: 6**)

25. Jin-Dong Kim, Kevin Cohen, Nigel Collier, Zhiyong Lu, and Pontus Stenetorp. **Introduction to the Biomedical Linked Annotation Hackathon (BLAH) 2015 Symposium** *In BioMed Central Proceedings, August 2015.*

24. Kazuma Hashimoto, Pontus Stenetorp, Makoto Miwa, and Yoshimasa Tsuruoka. **Task-Oriented Learning of Word Embeddings for Semantic Relation Classification** *In Proceedings of the Nineteenth Conference on Computational Natural Language Learning, Beijing, China, July 2015.* (**Acceptance rate: 30%**, **citations: 12**)

23. Sampo Pyysalo, Jorge Campos, Juan Miguel Cejuela, Filip Ginter, Kai Hakala, Chen Li, Pontus Stenetorp, and Lars Juhl Jensen. **Sharing annotations better: RESTful Open Annotation** *In Proceedings of ACL-IJCNLP 2015 System Demonstrations, Beijing, China, July 2015.* (**Citations: 7**)

22. Hubert Soyer, Goran Topić, Pontus Stenetorp, and Akiko Aizawa. **CroVeWA: Crosslingual Vector-Based Writing Assistance** *In Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, Denver, Colorado, USA, June 2015.*

21. Hubert Soyer, Pontus Stenetorp, and Akiko Aizawa. **Leveraging Monolingual Data for Crosslingual Compositional Word Representations** *In Proceedings of the 3rd International Conference on Learning Representations, San Diego, California, USA, May 2015.* (**Acceptance rate: 20%**, **citations: 16**)

20. Kazuma Hashimoto, Pontus Stenetorp, Makoto Miwa, and Yoshimasa Tsuruoka. **Jointly Learning Word Representations and Composition Functions Using Predicate-Argument Structures** *In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, October 2014.* (**Acceptance rate: 27%**, **citations: 15**)

19. Pontus Stenetorp, Sampo Pyysalo, Sophia Ananiadou, and Jun'ichi Tsujii. **Generalising Semantic Category Disambiguation with Large Lexical Resources for Fun and Profit** *Journal of Biomedical Semantics, June 2014.*

18. Pontus Stenetorp. **Transition-based Dependency Parsing Using Recursive Neural Networks** *Deep Learning Workshop at the 2013 Conference on Neural Information Processing Systems (NIPS), Lake Tahoe, Nevada, USA, December 2013.* (**Citations: 14**)

17. Pontus Stenetorp, Wiktoria Golik, Thierry Hamon, Donald C. Comeau, Rezarta Islamaj Doğan, Haibin Liu, and W. John Wilbur. **BioNLP Shared Task 2013: Supporting Resources** *In Proceedings of the BioNLP Shared Task 2013 Workshop, Sofia, Bulgaria, August 2013.*

16. Pontus Stenetorp. **On the Applicability and Representation of Lexical Prior Knowledge** *Doctor of Philosophy (PhD) Thesis, University of Tokyo, Tokyo, Japan, March 2013.*

15. Pontus Stenetorp, Sampo Pyysalo, Goran Topić, Sophia Ananiadou, and Akiko Aizawa. **Normalisation with the brat rapid annotation tool** *In Proceedings of the 5th International Symposium on Semantic Mining in Biomedicine, Zürich, Switzerland, September 2012.*

14. Pontus Stenetorp, Hubert Soyer, Sampo Pyysalo, Sophia Ananiadou, and Takashi Chikayama. **Size (and Domain) Matters: Evaluating Semantic Word Space Representations for Biomedical Text** *In Proceedings of the 5th International Symposium on Semantic Mining in Biomedicine, Zürich, Switzerland, September 2012.* (**Citations: 10**)

13. Pontus Stenetorp, Sampo Pyysalo, Tomoko Ohta, Sophia Ananiadou, and Jun'ichi Tsujii. **Bridging the Gap Between Scope-based and Event-based Negation/Speculation Annotations: A Bridge Not Too Far** *In Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics, Jeju, Republic of Korea, July 2012.*

12. Panot Chaimongkol, Pontus Stenetorp, and Aizawa Aizawa. **Utilising Bilingual Lexical Resources for Technical Term Extraction** *In Proceedings of The 26th Annual Conference of the Japanese Society for Artificial Intelligence, Yamaguchi, Japan, June 2012.*

11. Sampo Pyysalo, Pontus Stenetorp, Tomoko Ohta, Jin-Dong Kim, and Sophia Ananiadou. **New Resources and Perspectives for Biomedical Event Extraction** *In BioNLP: Proceedings of the 2012 Workshop on Biomedical Natural Language Processing, Montréal, Canada, June 2012.*

10. Pontus Stenetorp, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun'ichi Tsujii. **brat: a Web-based Tool for NLP-Assisted Text Annotation** *In Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics, Avignon, France, April 2012.* (**Citations: 330**)

9. Pontus Stenetorp, Goran Topić, Sampo Pyysalo, Sophia Ananiadou, and Jun'ichi Tsujii. **stav: text annotation visualiser** *In Proceedings of NLP2012, Hiroshima, Japan, March 2012.*

8. Pontus Stenetorp, Sampo Pyysalo, Sophia Ananiadou, and Jun'ichi Tsujii. **Almost Total Recall: Semantic Category Disambiguation Using Large Lexical Resources and Approximate String Matching** *In Proceedings of the Fourth International Symposium on Languages in Biology and Medicine, Singapore, Singapore, December 2011.*

7. Pontus Stenetorp, Sampo Pyysalo, Sophia Ananiadou, and Jun'ichi Tsujii. **Investigating Approaches to Semantic Category Disambiguation Using Large Lexical Resources and Approximate String Matching** *Information Processing Society of Japan Special Interest Groups Notes, Ishigaki, Japan, November 2011.*

6. Pontus Stenetorp, Goran Topić, Sampo Pyysalo, Tomoko Ohta, Jin-Dong Kim, and Jun'ichi Tsujii. **BioNLP Shared Task 2011: Supporting Resources** *In Proceedings of BioNLP Shared Task 2011 Workshop, Portland, Oregon, USA, June 2011.* (**Citations: 36**)

5. Pontus Stenetorp, Sampo Pyysalo, and Jun'ichi Tsujii. **SimSem: Fast Approximate String Matching in Relation to Semantic Category Disambiguation** *In Proceedings of BioNLP 2011 Workshop, Portland, Oregon, USA, June 2011.* (**Citations: 5**)

4. Pontus Stenetorp. **Automated Extraction of Swedish Neologisms using a Temporally Annotated Corpus** *Master of Science in Engineering (MSc Eng) Thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, March 2010.*

3. Pontus Stenetorp, and Jason Jingshi Li. **Evaluating Cases in Legal Disputes as Rival Theories** *New Frontiers in Artificial Intelligence, March 2010.* (10.1007/978-3-642-14888-0_6)

2. Jason Jingshi Li, and Pontus Stenetorp. **Evaluating cases in legal disputes as rival theories** *In Proceedings of the Third International Workshop on Juris-informatics (JURISIN 2009), Tokyo, Japan, November 2009.*

1. Han-Cheol Cho, Do-Gil Lee, Jung-Tae Lee, Pontus Stenetorp, Jun'ichi Tsujii, and Hae-Chang Rim. **A Novel Word Segmentation Approach for Written Languages with Word Boundary Markers** *In Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, Suntec, Singapore, August 2009.* (**Acceptance rate: 26%**)

Citations are counted as unique publications citing the work according to Google Scholar. Only citation counts greater than or equal to 5 are shown.