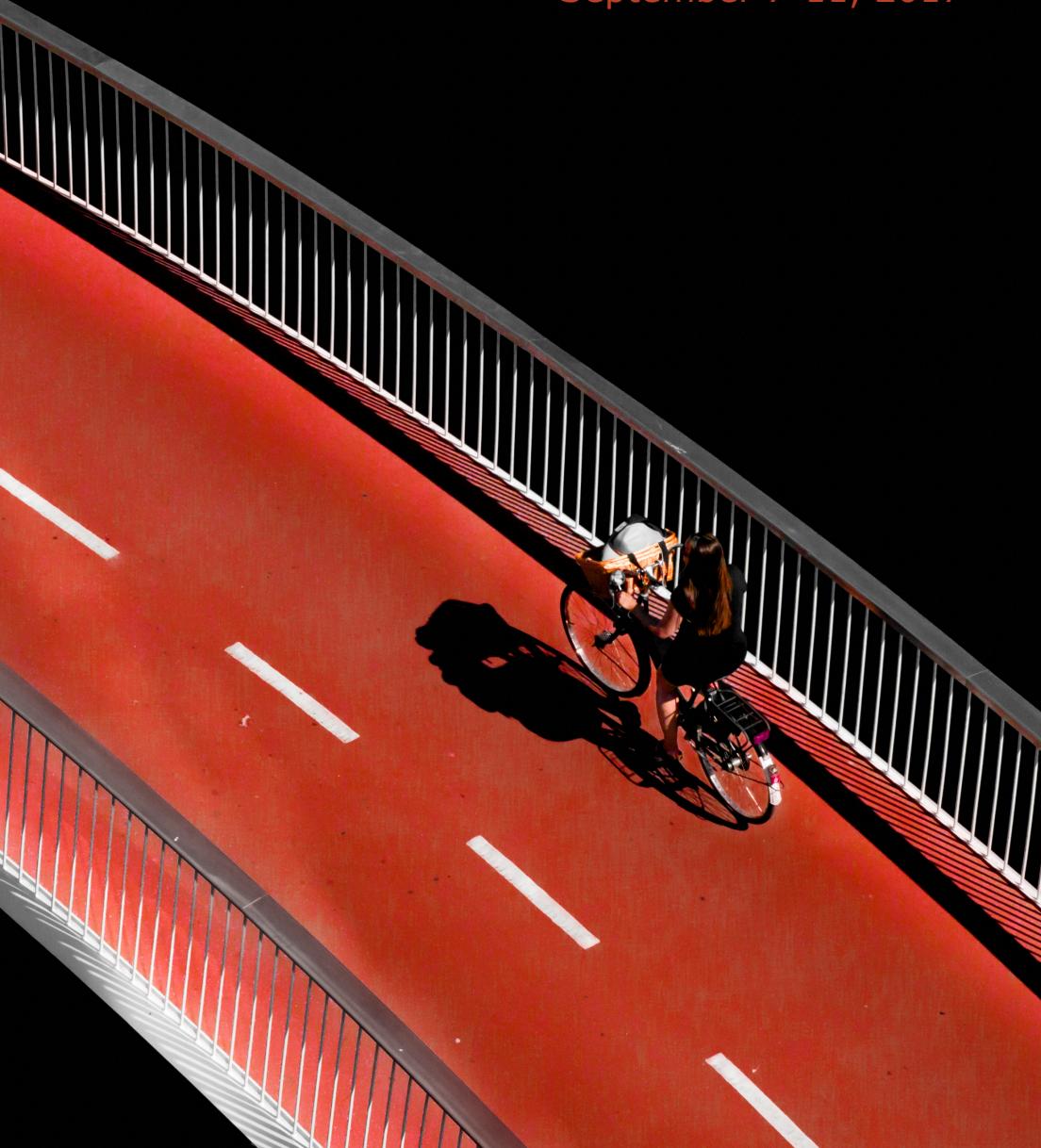


emnlp 2017

Conference Handbook

Copenhagen, Denmark
September 7-11, 2017



Cover design by Aurelia Bunescu

Cover photo by Matthew James

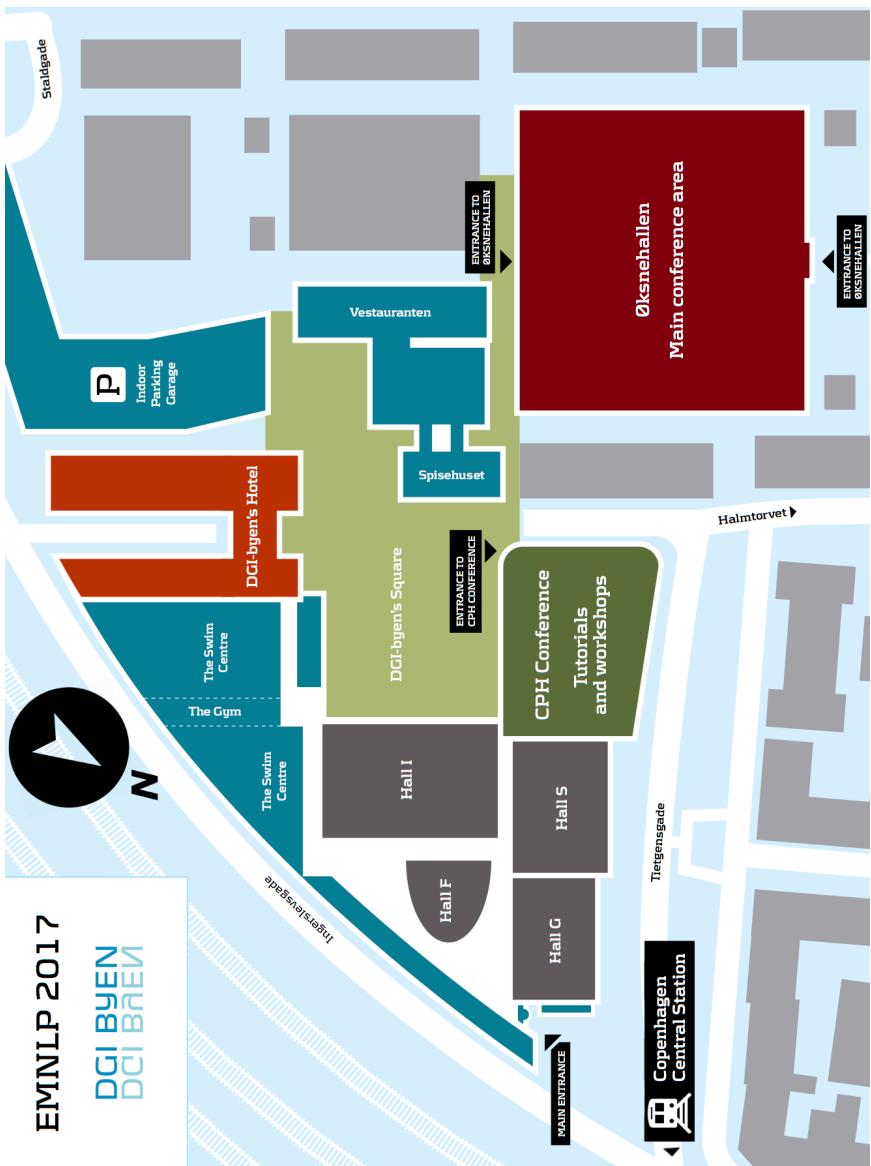
Deep gratitude to Matt Post for invaluable advice with creating this handbook

Handbook assembled by Joachim Bingel

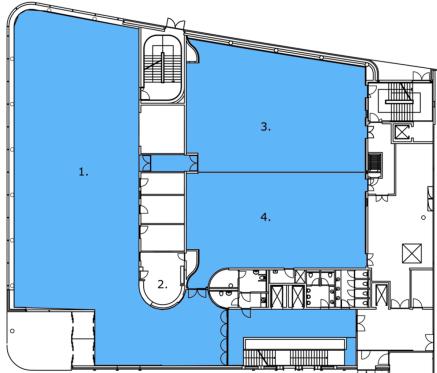
Printing by Frederiksbergs Bogtrykkeri A/S

EMNLP 2017

DGI BYEN
DCI BYEN

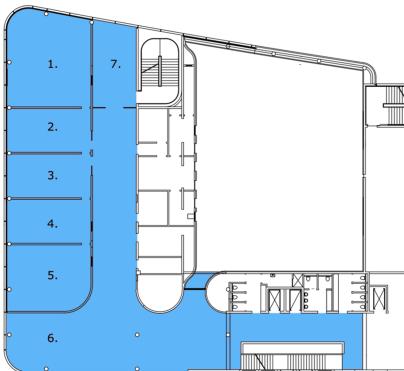


CPH Conference Floorplan (Workshops and Tutorials)



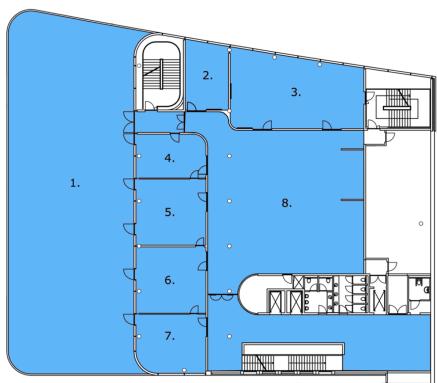
CPH CONFERENCE: GROUND FLOOR

1. Lobby 3. Sankt Hans Torv
2. Reception 4. Nørrebro Runddel



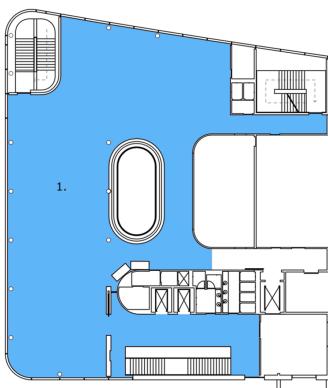
CPH CONFERENCE: 1st FLOOR

1. Kastrup Lufthavn 5. Amager Strandpark
2. Christianshavn 6. Break/ Lounge area
3. Islands Brygge 7. Break/ Lounge area
4. Christiania



CPH CONFERENCE: 2nd FLOOR

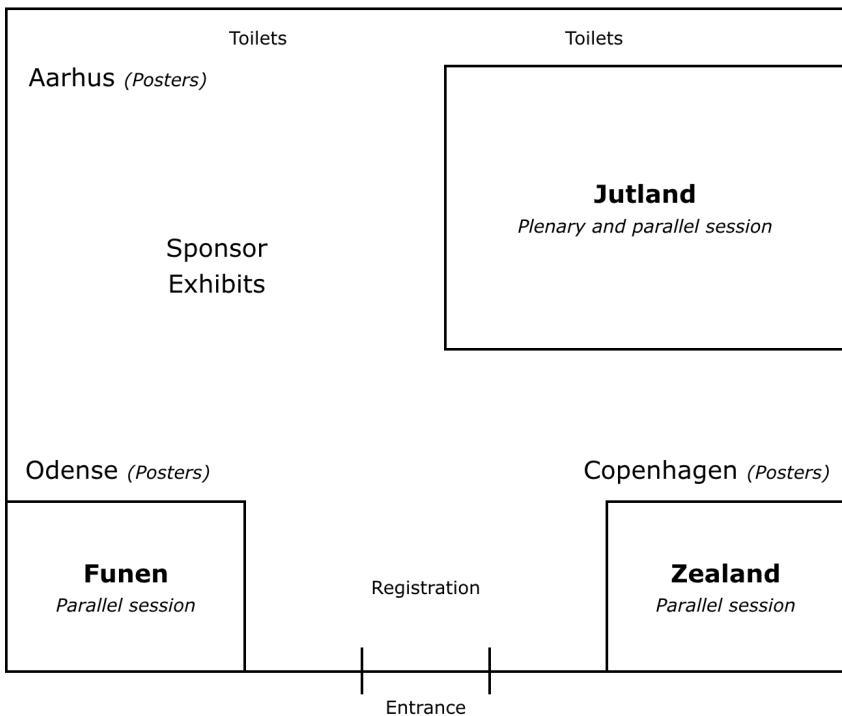
1. Roof terrace 5. Vesterbro Torv
2. Istedgade 6. Enghave Plads
3. Hovedbanen 7. Købyen
4. Tivoli 8. Break / Lounge area



CPH CONFERENCE: 3rd FLOOR

1. Østerbro

Øksnehallen Floorplan
(Main Conference)



Contents

Table of Contents	i
1 Conference Information	1
Message from the General Chair	1
Message from the Program Committee Co-Chairs	3
Organizing Committee	5
Program Committee	6
Venue Info	8
Meal Info	8
Welcome Reception	8
2 Tutorials: Thursday, September 7	11
T1: Acquisition, Representation and Usage of Conceptual Hierarchies	12
T2: Computational Sarcasm	13
T3: Graph-based Text Representations: Boosting Text Mining, NLP and Information Retrieval with Graphs	15
T4: Semantic Role Labeling	17
3 Tutorials: Friday, September 8	19
T5: Memory Augmented Neural Networks for Natural Language Processing	20
T6: A Unified Framework for Structured Prediction: From Theory to Practice	22
T7: Cross-Lingual Word Representations: Induction and Evaluation	24
4 Workshops: Thursday–Friday, September 7–8	25
W1: Second Conference on Machine Translation (WMT)	26
W2: Subword and Character LEvel Models in NLP	32
W3: Natural Language Processing meets Journalism	34
W4: 3rd Workshop on Noisy User-generated Text	36
W5: 2nd Workshop on Structured Prediction for Natural Language Processing	38
W6: New Frontiers in Summarization	39
W7: Workshop on Speech-Centric Natural Language Processing	41
W8: 3rd Workshop on Discourse in Machine Translation	43
W9: Workshop on Stylistic Variation	45
W10: 12th Workshop on Innovative Use of NLP for Building Educational Applications	47

W11: 4th Workshop on Argument Mining	51
W12: 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis	53
W13: 2nd Workshop on Evaluating Vector Space Representations for NLP	56
W14: Building Linguistically Generalizable NLP Systems	58
5 Main Conference: Saturday, September 9	61
Demo Session	61
Invited Talk: Dan Jurafsky	65
6 Main Conference: Sunday, September 10	67
Invited Talk: Sharon Goldwater	68
Social Event	69
7 Main Conference: Monday, September 11	71
Invited Talk: Nando de Freitas	72
SIGDAT Business Meeting	73
8 Anti-harassment policy	75
Author Index	77
9 Local Guide	87
General	87
Sightseeing	88
Shops	90
Events	91
Day-Trips	91
Other Things to Do	92
Dance the Night Away	92
Parks	92
Eating & Drinking	92
Restaurants Close to the Conference Venue	93
Culinary Highlights in the City	95
Cafes and Bars	97
Health	98

Conference Information

Message from the General Chair

Thank you so much for joining us in Copenhagen! Welcome to a cosmopolitan city of fantastic restaurants, lovely seascapes, rich history, and lots and lots of cyclists!

We have an exciting program lined up for you, with three Invited talks, fifteen workshops, seven tutorials, nine TACL presentations, 322 reviewed papers presented as both oral talks and posters, and twenty-one demos. I am especially grateful to our Program Chairs, Rebecca Hwa and Sebastian Riedel, who did a fantastic job managing a backbreaking 1,500 paper submissions (1466 reviewed papers). This involved 51 Area chairs and 980 reviewers. We tried some new things this year (never conducive to a smooth process) including a more careful handling of the COIs that result from Area Chair submissions, and the addition of a meta-review step to encourage more thoughtful reviewing. We are soliciting feedback on the meta-review process, from both reviewers and authors. Despite the additional time involvement, many of the Area Chairs embraced this new approach, and would like to repeat it. However, there are clearly a few dissenters, since Rebecca and Sebastian ended up writing around 200 meta-reviews themselves at the last minute! We are also trying to raise the visibility and status of the poster sessions by integrating them as parallel sessions alongside oral talks, with poster session chairs. This is in response to the survey results from EMNLP 2015 that indicated a decided preference for smaller, more frequent poster sessions during the day rather than evening mega-sessions. Finally, Rebecca and Sebastian are bringing you three outstanding invited speakers, Dan Jurafsky, Sharon Goldwater, and Nando de Freitas. No program chairs ever worked harder to bring you a superb set of presentations in an attendee friendly setting.

I am also very grateful to Victoria Fossum and Karl Moritz Hermann, our Workshop Chairs, who put together a terrific slate of fifteen workshops, and paid meticulous attention to ensuring that each workshop could hold exactly the poster sessions, invited talks and special events that it required. Our tutorial chairs, Alexandra Birch and Nathan Schneider, also outdid themselves, providing especially tempting tutorial offerings. Matt Post deserves to be singled out, for being an Advisor to our conscientious and successful Handbook Chair, Joachim Bingel, as well as becoming a welcome last minute addition to our excellent team of Demo Chairs, Lucia Specia and Michael Paul. Thanks are due to our Website Chair, Anders Johannsen, who responded promptly and deftly to all of our requests, and to our Student Volunteer and Student Sponsorship Chairs, Zeljko Agic and Yonatan Bisk, who brought you the helpful and energetic volunteers who keep things running smoothly.

Last but not least, many thanks to your hosts, our Local Arrangements Chairs, Dirk Hovy and Anders Søgaard and their team. Their concern has been increasing the enjoyment of

Message from the General Chair

your experience, and to that end they proposed a stunning venue, put together an amazing reception and Social Event, chose your conference bags, issued all the invitation letters for visas, helped create all the signs, etc., etc., etc. Dan Hardt, our Sponsorship chair, working with Anders and Dirk, raised an unusual amount of local sponsorships, all to defray the cost of the Social Event.

As always, we are extremely indebted to our generous sponsors. Our platinum sponsors are Google, Amazon, Baidu, Apple, Facebook, Bloomberg and Siteimprove. Gold sponsors include IBM Research, Microsoft, ebay, SAP, textkernel, Maluuba, Recruit Institute of Technology, and Deloitte. Silver sponsors are Nuance, Oracle, Sogou, Huawei, Duolingo, CVTE, and Wizkids. Grammarly and Yandex are our Bronze sponsors.

Finally, many, many thanks to our Area Chairs, our reviewers, and our authors, whose outstanding research is being showcased here for your delectation. *Nyd det mens det varer!*

Best Regards,
Martha Palmer
EMNLP 2017 General Chair

Message from the Program Committee Co-Chairs

Welcome to the 2017 Conference on Empirical Methods in Natural Language Processing! This is an exciting year; we have received a new record-high in the number of submissions: 1,509 papers. After discounting early withdraws, duplicates, and other invalid submissions, we sent out 1,418 submissions (836 long papers, 582 short papers) to be reviewed by the program committee. Ultimately, 216 long papers (25.8% acceptance rate) and 107 short papers (18.4% acceptance rate) have been accepted for presentation, making a total of 323 papers and an overall acceptance rate of 22.8%.

This year's technical program consists of three invited talks and 113 oral presentations and 219 poster presentations for the 323 long and short accepted papers as well as nine papers accepted to the Transactions of the Association for Computational Linguistics. To accommodate all the presentations in a compressed timeframe, we opted to have plenary sessions for the invited talks and the winners of the Best Paper Awards, while allotting three parallel oral sessions and thematically related poster sessions for all other presentations. We chose to have concurrent poster and oral sessions for several reasons. First, this is the preferred model of the majority (51.6%) of participants who filled out the EMNLP 2015 post-conference survey. Second, this allows us to spread out the poster presentations across three days in smaller thematically related clusters. Finally, this maximises the number of acceptances for the high quality submissions we received; by having more poster sessions, we are able to maintain the acceptance rates at the previous year's level despite an increase in submissions by 40%.

It would not have been possible to properly handle such a large number of submissions without the generous voluntary help from all the members of the program committee, which consists of 980 reviewers overseen by 51 area chairs. We continued last year's experiment of defining twelve relatively broad topic areas and assigning multiple area chairs to facilitate consistent ranking of larger sets of papers. Most technical program decisions, from the selection of papers to the modes of presentation to the choice of outstanding papers, are primarily made in a bottom-up fashion: reviewers assessed and scored papers, made recommendations for oral vs poster decisions, and marked papers suitable for best paper awards; area chairs ensured the quality of assessments, encouraged discussions and assembled opinions into their own recommendations; finally, we construct the technical program, considering the recommendations from the area chairs while taking into account venue constraints and balance across areas. A new experimental feature of this year's EMNLP reviewing process is the "meta review," in which the area chairs briefly summarize the major discussions between the reviewers to give authors a more transparent view of the process.

Per EMNLP tradition, awards are given to outstanding papers in three categories: Best Long Paper, Best Short Paper, and Best Resource Paper. The selection process is bottom-up: based on the reviewers and area chairs' recommendations, we nominated four papers for each category; we invited expert members to form a Best Papers committee for each category; each committee reviews the candidates and select the winners. The awarded papers will be presented at a special plenary session on the last day of the conference.

We are extremely grateful that three amazing speakers have agreed to give invited talks at EMNLP. Nando de Freitas (Google Deepmind) will discuss simulated physical environments, and whether language would benefit from the development of such environments, and could contribute toward improving such environments and agents within them. Sharon Goldwater (University of Edinburgh) will describe work on developing unsupervised speech technology for those of the world's 7,000 or so languages not spoken in large rich countries. Dan Jurafsky (Stanford University) will talk about processing the language of policing to automatically measure linguistic aspects of the interaction from discourse factors like conversational structure to social factors like respect.

The conference would not have been possible without the support of various people inside and outside of the committee. In particular, we would like to thank:

- Martha Palmer, whose encouragement and advice as the general chair has been invaluable every step of the way;

Message from the Program Committee Co-Chairs

- Chris Callison-Burch, who has given us excellent advice and support in his capacity as the SIGDAT Secretary;
- Priscilla Rasmussen, who always has the right answers;
- Xavier Carreras and Kevin Duh, who generously shared their experiences as the chairs of EMNLP 2016;
- Anders Johannsen, who is lightning fast with website updates;
- Our 51 area chairs: David Bamman, Mohit Bansal, Roberto Basili, Chris Biemann, Jordan Boyd-Graber, Marine Carpuat, Joyce Chai, David Chiang, Jinho Choi, Jennifer Chu-Carroll, Trevor Cohn, Cristian Danescu-Niculescu-Mizil, Dipanjan Das, Hal Daume, Mona Diab, Mark Dredze, Jacob Eisenstein, Sanja Fidler, Alona Fyshe, Dan Gildea, Ed Grefenstette, Hannaneh Hajishirzi, Julia Hockenmaier, Kentaro Inui, Jing Jiang, Philipp Koehn, Mamoru Komachi, Anna Korhonen, Tom Kwiatkowski, Gina Levow, Bing Liu, Nitin Madnani, Mausam, Rada Mihalcea, Marie-Francine Moens, Saif M. Mohammad, Mari Ostendorf, Sameer Pradhan, Alexander Rush, Anoop Sarkar, William Schuler, Hinrich Schütze, Sameer Singh, Thamar Solorio, Vivek Srikanth, Amanda Stent, Tomek Strzalkowski, Mihai Surdeanu, Andreas Vlachos, Scott Wenthau Yih, Zhang Yue;
- The best papers award committee members: Chris Brew, Mike Collins, Kevin Duh, Adam Lopez, Ani Nenkova, Bonnie Webber, Luke Zettlemoyer;
- Preethi Raghavan and Siddharth Patwardhan, the publications co-chairs and Joachim Bingel, the conference handbook chair;
- Dirk Hovy and Anders Søgaard, the local arrangements co-chairs;
- Rich Gerber and Paolo Gai at SoftConf.

Finally, we'd like to thank SIGDAT for the opportunity to serve as Program Co-Chairs of EMNLP 2017. It is an honor and a rewarding learning experience. We hope you will be as inspired by the technical program as we are.

EMNLP 2017 Program Co-Chairs
Rebecca Hwa, University of Pittsburgh
Sebastian Riedel, University College London

Organizing Committee

TODO double-check with Homepage

General Chair

Martha Palmer, University of Colorado

Local Arrangements Chair

Priscilla Rasmussen, ACL Business Manager

Program Committee Co-chairs

Rebecca Hwa, University of Pittsburgh

Sebastian Riedel, University College London

Local Arrangements Co-chairs

Dirk Hovy, University of Copenhagen

Anders Søgaard, University of Copenhagen

Local Sponsorship Chair

Daniel Hardt, Copenhagen Business School

Workshop Co-chairs

Victoria Fossum, Google

Karl Moritz Hermann, DeepMind

Tutorial Co-chairs

Alexandra Birch, Edinburgh University

Nathan Schneider, Georgetown University

Demos Co-chairs

Michael Paul, University of Colorado

Lucia Specia, University of Sheffield

Publications Sr Chair

Siddharth Patwardhan, Apple

Publications Jr Chair

Preethi Raghavan, TJ Watson Lab, IBM

Publicity Chair

Isabelle Augenstein, University College London

Web Chair

Anders Johannsen, Apple

Conference Handbook Chair

Joachim Bingel, University of Copenhagen

Conference Handbook Advisor

Matt Post, Johns Hopkins University

Handbook Proofreader

Pontus Stenetorp, University College London

Student Scholarship Co-chair and Student Volunteer Coordinator

Željko Agić, IT University of Copenhagen

Yonatan Bisk, University of Southern California, ISI

SIGDAT Liason

Chris Callison-Burch, University of Pennsylvania

Program Committee

Program Committee Co-chairs

Rebecca Hwa, University of Pittsburgh
Sebastian Riedel, University College London

Area Chairs

Information Extraction, Information Retrieval, and Question Answering

Mihai Surdeanu (QA), University of Arizona
Jing Jiang, Singapore Management University
Hinrich Schütze, LMU Munich
Sameer Singh, UC Irvine
Scott Wen-tau Yih, MSR
Tomek Strzalkowski, SUNY Albany

Language and Vision

Sanja Fidler, University of Toronto
Hannaneh Hajishirzi, University of Washington

Linguistic Theories and Psycholinguistics

William Schuler, The Ohio State University

Machine Learning

Mohit Bansal, UNC Chapel Hill
Jordan Boyd-Graber, University of Colorado
Trevor Cohn, University of Melbourne
Hal Daumé, University of Maryland
Alona Fyshe, University of Victoria
Anoop Sarkar, Simon Fraser University

Machine Translation and Multilinguality

Marine Carpuat, University of Maryland
David Chiang, University of Notre Dame
Mona Diab, George Washington University
Dan Gildea, University of Rochester
Philipp Koehn, Johns Hopkins University

Segmentation, Tagging, and Parsing

Jinho Choi, Emory University
Julia Hockenmaier, University of Illinois at Urbana-Champaign
Alexander Rush, Harvard University
Zhang Yue, Singapore University of Technology and Design

Semantics

Roberto Basili, University of Roma, Tor Vergata
Chris Biemann, University of Hamburg
Ed Grefenstette, DeepMind
Tom Kwiatkowski, Google
Sameer Pradhan, cemantix.org and Boulder Learning, Inc
Vivek Srikumar, University of Utah

Sentiment Analysis and Opinion Mining

Bing Liu, University of Illinois at Chicago
Rada Mihalcea, University of Michigan
Saif M. Mohammad, National Research Council Canada

Social Media and Computational Social Science

David Bamman, University of California, Berkeley
Cristian Danescu-Niculescu-Mizil, Cornell University
Mark Dredze, Johns Hopkins University
Jacob Eisenstein, Georgia Tech

Spoken Language Processing

Mari Ostendorf, University of Washington

Summarization, Generation, Discourse, Dialogue

Joyce Chai, Michigan State University

Jennifer Chu-Carroll, Elemental Cognition

Kentaro Inui, Tohoku University

Gina Levow, University of Washington

Amanda Stent, Bloomberg LP

Text Mining and NLP Applications

Dipanjan Das, Google

Mamoru Komachi, Tokyo Metropolitan University

Anna Korhonen, University of Cambridge

Nitin Madnani, Educational Testing Service (ETS)

Marie-Francine Moens, KU Leuven

Thamar Solorio, University of Houston

Andreas Vlachos, University of Sheffield

Venue Info

Please note that EMNLP takes place in *two different locations*. All tutorials and workshops are located in **CPH Conference**, while the main conference takes place in **Øksnehallen**. There are maps with room locations, as well as an area map, in the front matter.

Meal Info

The following meals are provided as part of your registration fee:

- During the coffee breaks in the mornings and afternoons, **snacks** are provided.
- There will also be **snacks** at the welcome reception.
- A **full dinner** with various options is provided as part of the social event.

Welcome Reception

Thursday, September 7, 2017, 19:00 – 22:00

CPH Conference, Room *Østerbro*

Catch up with your colleagues at the **Welcome Reception!** It will be held following tutorials on Thursday evening.

Thursday, September 7, 2017

TODO full overview for all days

2

Tutorials: Thursday, September 7

Overview

7:30–6:00	Registration	<i>Foyer</i>
7:30–9:00	Breakfast	<i>TODO NOT AVAILABLE</i>
9:00–12:30	Morning Tutorials	
	Acquisition, Representation and Usage of Conceptual Hierarchies	<i>TODO</i>
	<i>Marius Pasca</i>	
	Computational Sarcasm	<i>TODO</i>
	<i>Pushpak Bhattacharyya and Aditya Joshi</i>	
10:30–11:00	Coffee break	
12:30–2:00	Lunch break	
2:00–5:30	Afternoon Tutorials	
	Graph-based Text Representations: Boosting Text Mining, NLP and Information Retrieval with Graphs	<i>TODO</i>
	<i>Fragkiskos D. Malliaros and Michalis Vazirgiannis</i>	
	Semantic Role Labeling	<i>TODO</i>
	<i>Diego Marcheggiani, Michael Roth, Ivan Titov, and Benjamin Van Durme</i>	
3:30–4:00	Coffee break	
6:00–9:00	Welcome Reception	<i>CPH Conference, Room Østerbro</i>

Tutorial 1

Acquisition, Representation and Usage of Conceptual Hierarchies

Marius Pasca

Thursday, September 7, 2017, 9:00–12:30pm

TODO

Through subsumption and instantiation, individual instances (“artificial intelligence”, “the spotted pig”) otherwise spanning a wide range of domains can be brought together and organized under conceptual hierarchies. The hierarchies connect more specific concepts (“computer science subfields”, “gastropubs”) to more general concepts (“academic disciplines”, “restaurants”) through IsA relations. Explicit or implicit properties applicable to, and defining, more general concepts are inherited by their more specific concepts, down to the instances connected to the lower parts of the hierarchies. Subsumption represents a crisp, universally-applicable principle towards consistently representing IsA relations in any knowledge resource. Yet knowledge resources often exhibit significant differences in their scope, representation choices and intended usage, to cause significant differences in their expected usage and impact on various tasks.

This tutorial examines the theoretical foundations of subsumption, and its practical embodiment through IsA relations compiled manually or extracted automatically. It addresses IsA relations from their formal definition; through practical choices made in their representation within the larger and more widely-used of the available knowledge resources; to their automatic acquisition from document repositories, as opposed to their manual compilation by human contributors; to their impact in text analysis and information retrieval. As search engines move away from returning a set of links and closer to returning results that more directly answer queries, IsA relations play an increasingly important role towards a better understanding of documents and queries. The tutorial teaches the audience about definitions, assumptions and practical choices related to modeling and representing IsA relations in existing, human-compiled resources of instances, concepts and resulting conceptual hierarchies; methods for automatically extracting sets of instances within unlabeled or labeled concepts, where the concepts may be considered as a flat set or organized hierarchically; and applications of IsA relations in information retrieval.

Marius Pasca is a research scientist at Google. Current research interests include factual information extraction from unstructured text within documents and queries and its applications to Web search.

Tutorial 2

Computational Sarcasm

Pushpak Bhattacharyya and Aditya Joshi

Thursday, September 7, 2017, 9:00–12:30pm

TODO

Sarcasm is a form of verbal irony that is intended to express contempt or ridicule. Motivated by challenges posed by sarcastic text to sentiment analysis, computational approaches to sarcasm have witnessed a growing interest at NLP forums in the past decade. Computational sarcasm refers to automatic approaches pertaining to sarcasm. The tutorial will provide a bird's-eye view of the research in computational sarcasm for text, while focusing on significant milestones.

The tutorial begins with linguistic theories of sarcasm, with a focus on incongruity: a useful notion that underlies sarcasm and other forms of figurative language. Since the most significant work in computational sarcasm is sarcasm detection: predicting whether a given piece of text is sarcastic or not, sarcasm detection forms the focus hereafter. We begin our discussion on sarcasm detection with datasets, touching on strategies, challenges and nature of datasets. Then, we describe algorithms for sarcasm detection: rule-based (where a specific evidence of sarcasm is utilised as a rule), statistical classifier-based (where features are designed for a statistical classifier), a topic model-based technique, and deep learning-based algorithms for sarcasm detection. In case of each of these algorithms, we refer to our work on sarcasm detection and share our learnings. Since information beyond the text to be classified,

Prof. Pushpak Bhattacharyya is the current President of ACL (2016-17). He is the Director of IIT Patna and Vijay and Sita Vashee Chair Professor in IIT Bombay, Computer Science and Engineering Department. He was educated in IIT Kharagpur (B.Tech), IIT Kanpur (M.Tech) and IIT Bombay (PhD). He has been visiting scholar and faculty in MIT, Stanford, UT Houston and University Joseph Fourier (France). Prof. Bhattacharyya's research areas are Natural Language Processing, Machine Learning and AI. He has guided more than 250 students (PhD, masters and Bachelors), has published more than 250 research papers and led government and industry projects of international and national importance. A significant contribution of his is Multilingual Lexical Knowledge Bases and Projection. Author of the text book 'Machine Translation' Prof. Bhattacharyya is loved by his students for his inspiring teaching and mentorship. He is a Fellow of National Academy of Engineering and recipient of Patwardhan Award of IIT Bombay and VNMM award of IIT Roorkey- both for technology development, and faculty grants of IBM, Microsoft, Yahoo and United Nations.

Aditya Joshi is a PhD student at IITB-Monash Research Academy, a joint PhD programme between Indian Institute of Technology Bombay, India and Monash University, Australia, since January 2013. His PhD advisors are Pushpak Bhattacharyya (IITB) and Mark Carman (Monash). His primary research focus is computational sarcasm where he has explored different ways in which incongruity can be captured in order to detect and generate sarcasm. In addition, he has worked on innovative applications of NLP such as sentiment analysis for Indian languages, drunk-texting prediction, news headline translation, political issue extraction, etc.

contextual information is useful for sarcasm detection, we then describe approaches that use such information through conversational context or author-specific context.

We then follow it by novel areas in computational sarcasm such as sarcasm generation, sarcasm v/s irony classification, etc. We then summarise the tutorial and describe future directions based on errors reported in past work. The tutorial will end with a demonstration of our work on sarcasm detection.

This tutorial will be of interest to researchers investigating computational sarcasm and related areas such as computational humour, figurative language understanding, emotion and sentiment sentiment analysis, etc. The tutorial is motivated by our continually evolving survey paper of sarcasm detection, that is available on arXiv at: Joshi, Aditya, Pushpak Bhattacharyya, and Mark James Carman. "Automatic Sarcasm Detection: A Survey." arXiv preprint arXiv:1602.03426 (2016).

Tutorial 3

Graph-based Text Representations: Boosting Text Mining, NLP and Information Retrieval with Graphs

Fragkiskos D. Malliaros and Michalis Vazirgiannis

Thursday, September 7, 2017, 9:00–12:30pm

TODO

Graphs or networks have been widely used as modeling tools in Natural Language Processing (NLP), Text Mining (TM) and Information Retrieval (IR). Traditionally, the unigram bag-of-words representation is applied; that way, a document is represented as a multi-set of its terms, disregarding dependencies between the terms. Although several variants and extensions of this modeling approach have been proposed (e.g., the n-gram model), the main weakness comes from the underlying term independence assumption. The order of the terms within a document is completely disregarded and any relationship between terms is not taken into account in the final task (e.g., text categorization). Nevertheless, as the heterogeneity of text collections is increasing (especially with respect to document length and vocabulary), the research community has started exploring different document representations aiming to capture more fine-grained contexts of co-occurrence between different terms, challenging the well-established unigram bag-of-words model. To this direction, graphs constitute a well-developed model that has been adopted for text representation. The goal of this tutorial is to offer a comprehensive presentation of recent methods that rely on graph-based text representations to deal with various tasks in NLP and IR. We will describe basic as well as novel graph theoretic concepts and we will examine how they can be applied in a wide

Fragkiskos D. Malliaros is currently a data science postdoctoral scholar in the Department of Computer Science and Engineering at UC San Diego and member of the Artificial Intelligence group. Right before that, he was a postdoctoral researcher in Ecole Polytechnique, France from where he also received his Ph.D. degree in 2015. He obtained his Diploma and his M.Sc. degree from the University of Patras, Greece in 2009 and 2011 respectively. He is the recipient of the 2012 Google European Doctoral Fellowship in Graph Mining and the 2015 Thesis Prize by Ecole Polytechnique. During the summer of 2014, he was a research intern at the Palo Alto Research Center (PARC), working on anomaly detection in social networks. His research interests span the broad area of data science, with focus on graph mining, social network analysis, applied machine learning and natural language processing.

Michalis Vazirgiannis is a Professor in Ecole Polytechnique, France and the leader of the Data Science and Mining (DaSciM) team. He holds a degree in Physics, a M.Sc. in Robotics, both from University of Athens, Greece, and a M.Sc. in Knowledge Based Systems from Heriot-Watt University (Edinburgh, UK). He acquired his Ph.D. degree from the Dept. of Informatics, University of Athens. He has worked as a researcher in different places: NTUA, GMD-IPSI (currently Fraunhofer-IPSI), Germany Fern-Universitaet Hagen, in project VERSO (later GEMO) in INRIA/Paris, in IBM India Research Laboratory and in MPI fur Informatik (Saarbruecken, Germany). He held a Marie Curie Intra-European fellow in area of P2P Web Search, hosted by INRIA FUTURS, Paris. His current research interests are on graph mining, text mining and recommendation algorithms. He is chairing the “AXA Data Science” chair in Ecole Polytechnique and has collaborations with the industry including Google and Airbus.

range of text-related application domains. All the material associated to the tutorial will be available at: http://fragkiskosm.github.io/projects/graph_text_tutorial

Tutorial 4

Semantic Role Labeling

Diego Marcheggiani, Michael Roth, Ivan Titov, and Benjamin Van Durme

Thursday, September 7, 2017, 2:00–5:30pm

TODO

This tutorial describes semantic role labelling (SRL), the task of mapping text to shallow semantic representations of eventualities and their participants. The tutorial introduces the SRL task and discusses recent research directions related to the task. The audience of this tutorial will learn about the linguistic background and motivation for semantic roles, and also about a range of computational models for this task, from early approaches to the current state-of-the-art. We will further discuss recently proposed variations to the traditional SRL task, including topics such as semantic proto-role labeling.

We also cover techniques for reducing required annotation effort, such as methods exploiting unlabeled corpora (semi-supervised and unsupervised techniques), model adaptation across languages and domains, and methods for crowdsourcing semantic role annotation (e.g., question-answer driven SRL). Methods based on different machine learning paradigms, including neural networks, generative Bayesian models, graph-based algorithms and bootstrapping style techniques.

Beyond sentence-level SRL, we discuss work that involves semantic roles in discourse. In particular, we cover data sets and models related to the task of identifying implicit roles and linking them to discourse antecedents. We introduce different approaches to this task from

Diego Marcheggiani is a postdoctoral researcher at the University of Amsterdam. He graduated with a Ph.D. in Computer Science from the University of Venice and during this period he worked at the ISTI-CNR in Italy as a researcher. His research focus ranges from relation extraction to semantic role labeling and frame-semantic parsing. He is interested in supervised and unsupervised learning approaches in the scope of tensor factorization models and neural networks.

Michael Roth is a postdoctoral researcher and DFG research fellow at Saarland University and University of Illinois at Urbana-Champaign, respectively. He graduated with a Ph.D. in Computational Linguistics from Heidelberg University in 2013. His research focus lies on computational models of language that can facilitate automatic text understanding beyond the sentence level. Recent work includes neural-network based approaches to semantic role labeling and discourse-level frame-semantic parsing. His models are the current state-of-the-art on the CoNLL-2009 and FrameNet 1.5 data sets.

Ivan Titov is an Associate Professor at the University of Amsterdam. He is the recipient of an ERC Starting Grant, a personal Vidi Grant from the Dutch NSF (NWO) and a Google Focused Research Award. Ivan is an action editor for the Journal of Machine Learning Research (JMLR) and Transactions of ACL (TACL), as well as an editorial board member of the Journal of Artificial Intelligence Research (JAIR). His interests are in probabilistic modeling of language, primarily in semantics and syntax as well as in multilingual NLP and semi-supervised learning for NLP.

Benjamin Van Durme is an Assistant Professor at the Johns Hopkins University in Computer Science, with a courtesy appointment in Cognitive Science, and the lead of the Natural Language Understanding group at the Human Language Technology Center of Excellence. (HLTCOE). His research is broadly focused on discovering and extracting knowledge from language, exploring topics such as low resource, multilingual information extraction; scalable, streaming algorithms for processing large collections; and semantic analysis at various levels of complexity.

the literature, including models based on coreference resolution, centering, and selectional preferences. We also review how new insights gained through them can be useful for the traditional SRL task.

Tutorials: Friday, September 8

Overview

7:30–6:00	Registration	<i>Foyer</i>
2:00–5:30	Afternoon Tutorials	
	Memory Augmented Neural Networks for Natural Language Processing	
	<i>TODO</i>	
	<i>Caglar Gulcehre and Sarath Chandar</i>	
	A Unified Framework for Structured Prediction: From Theory to Practice	
	<i>TODO</i>	
	<i>Wei Lu</i>	
	Cross-Lingual Word Representations: Induction and Evaluation	<i>todo</i>
	<i>Manaal Faruqui, Anders Søgaard, and Ivan Vulic</i>	
3:30–4:00	Coffee break	
6:00–9:00	Welcome Reception	<i>CPH Conference, Room Østerbro</i>

Tutorial 5

Memory Augmented Neural Networks for Natural Language Processing

Caglar Gulcehre and Sarath Chandar

Friday, September 8, 2017, 2:00–5:30pm

TODO

Designing of general-purpose learning algorithms is a long-standing goal of artificial intelligence. A general purpose AI agent should be able to have a memory that it can store and retrieve information from. Despite the success of deep learning in particular with the introduction of LSTMs and GRUs to this area, there are still a set of complex tasks that can be challenging for conventional neural networks. Those tasks often require a neural network to be equipped with an explicit, external memory in which a larger, potentially unbounded, set of facts need to be stored. They include but are not limited to, reasoning, planning, episodic question-answering and learning compact algorithms. Recently two promising approaches based on neural networks to this type of tasks have been proposed: Memory Networks and Neural Turing Machines.

In this tutorial, we will give an overview of this new paradigm of “neural networks with memory”. We will present a unified architecture for Memory Augmented Neural Networks (MANN) and discuss the ways in which one can address the external memory and hence read/write from it. Then we will introduce Neural Turing Machines and Memory Networks as specific instantiations of this general architecture. In the second half of the tutorial, we will focus on recent advances in MANN which focus on the following questions: How can we read/write from an extremely large memory in a scalable way? How can we design

Caglar Gulcehre is currently a research scientist at Deepmind. He finished his PhD in University of Montreal under the supervision of Yoshua Bengio. His work mainly focuses on applications of neural networks, in particular recurrent architectures such as GRU and LSTMs on NLP and sequence to sequence learning tasks. His research also investigates different optimization approaches and architectures which are easier to optimize for neural networks. His recent research focuses on building neural network models that have external memory structures. He has done research internships at IBM Watson Research Center, Google Deep Mind. He was a PC Member at ECML and IJCAI 2016 Deep Reinforcement Learning Workshop. Prior to joining MILA as a PhD student, he finished his master degree in Middle East Technical University in Cognitive Science department.

Sarath Chandar is currently a PhD student in University of Montreal under the supervision of Yoshua Bengio and Hugo Larochelle. His work mainly focuses on Deep Learning for complex NLP tasks like question answering and dialog systems. He also investigates scalable training procedure and memory access mechanisms for memory network architectures. In the past, he has worked on multilingual representation learning and transfer learning across multiple languages. His research interests includes Machine Learning, Natural Language Processing, Deep Learning, and Reinforcement Learning. Before joining University of Montreal, he was a Research Scholar in IBM Research India for a year. He has previously given a tutorial on "Multilingual Multimodal Language Processing using Neural Networks" at NAACL 2016.

efficient non-linear addressing schemes? How can we do efficient reasoning using large scale memory and an episodic memory? The answer to any one of these questions introduces a variant of MANN. We will conclude the tutorial with several open challenges in MANN and its applications to NLP.

We will introduce several applications of MANN in NLP throughout the tutorial. Few examples include language modeling, question answering, visual question answering, and dialogue systems.

Tutorial 6

A Unified Framework for Structured Prediction: From Theory to Practice

Wei Lu

Friday, September 8, 2017, 2:00–5:30pm

TODO

Structured prediction is one of the most important topics in various fields, including machine learning, computer vision, natural language processing (NLP) and bioinformatics. In this tutorial, we present a novel framework that unifies various structured prediction models.

The hidden Markov model (HMM) and the probabilistic context-free grammars (PCFGs) are two classic generative models used for predicting outputs with linear-chain and tree structures, respectively. As HMM’s discriminative counterpart, the linear-chain conditional random fields (CRFs) (Lafferty et al., 2001) model was later proposed. Such a model was shown to yield good performance on standard NLP tasks such as information extraction. Several extensions to such a model were then proposed afterward, including the semi-Markov CRFs (Sarawagi and Cohen, 2004), tree CRFs (Cohn and Blunsom, 2005), as well as discriminative parsing models and their latent variable variants (Petrov and Klein, 2007). On the other hand, utilizing a slightly different loss function, one could arrive at the structured support vector machines (Tschantzidis et al., 2004) and its latent variable variant (Yu and Joachims, 2009) as well. Furthermore, new models that integrate neural networks and graphical models, such as neural CRFs (Do et al., 2010) were also proposed.

In this tutorial, we will be discussing how such a wide spectrum of existing structured prediction models can all be implemented under a unified framework that involves some basic building blocks. Based on such a framework, we show how some seemingly complicated structured prediction models such as a semantic parsing model (Lu et al., 2008; Lu, 2014) can be implemented conveniently and quickly. Furthermore, we also show that the framework can be used to solve certain structured prediction problems that otherwise cannot be easily handled by conventional structured prediction models. Specifically, we show how to use such a framework to construct models that are capable of predicting non-conventional structures, such as overlapping structures (Lu and Roth, 2015; Muis and Lu, 2016a). We will also discuss how to make use of the framework to build other related models such as topic mod-

Wei Lu is an Assistant Professor at the Singapore University of Technology and Design (SUTD), directing the StatNLP research group. He received his Ph.D. from the National University of Singapore (NUS) in 2009. He visited CSAIL, Massachusetts Institute of Technology (MIT) in 2007–2008, and worked as a postdoctoral research associate at the University of Illinois at Urbana-Champaign in 2011–2013. His research interests include developing mathematical models and machine learning algorithms for solving natural language processing problems. He is particularly interested in semantic processing (in a broad sense). His papers appeared at venues such as ACL, EMNLP, NAACL, AAAI, and CIKM. He served as a program committee member for conferences such as ACL, EMNLP, NAACL, EACL, AAAI, IJCAI and NIPS, and is currently a member of the standing reviewer team for TACL. He served as an area co-chair for ACL 2016 and received the best paper award at EMNLP 2011.

els and highlight its potential applications in some recent popular tasks (e.g., AMR parsing (Flanigan et al., 2014)).

The framework has been extensively used by our research group for developing various structured prediction models, including models for information extraction (Lu and Roth, 2015; Muis and Lu, 2016a; Jie et al., 2017), noun phrase chunking (Muis and Lu, 2016b), semantic parsing (Lu, 2015; Susanto and Lu, 2017), and sentiment analysis (Li and Lu, 2017). It is our hope that this tutorial will be helpful for many natural language processing researchers who are interested in designing their own structured prediction models rapidly. We also hope this tutorial allows researchers to strengthen their understandings on the connections between various structured prediction models, and that the open release of the framework will bring value to the NLP research community and enhance its overall productivity.

Tutorial 7

Cross-Lingual Word Representations: Induction and Evaluation

Manaal Faruqui, Anders Søgaard, and Ivan Vulić

Friday, September 8, 2017, 2:00–5:30pm

todo

In recent past, NLP as a field has seen tremendous utility of distributional word vector representations as features in downstream tasks. The fact that these word vectors can be trained on unlabeled monolingual corpora of a language makes them an inexpensive resource in NLP. With the increasing use of monolingual word vectors, there is a need for word vectors that can be used as efficiently across multiple languages as monolingually. Therefore, learning bilingual and multilingual word embeddings/vectors is currently an important research topic. These vectors offer an elegant and language-pair independent way to represent content across different languages.

This tutorial aims to bring NLP researchers up to speed with the current techniques in cross-lingual word representation learning. We will first discuss how to induce cross-lingual word representations (covering both bilingual and multilingual ones) from various data types and resources (e.g., parallel data, comparable data, non-aligned monolingual data in different languages, dictionaries and thesauri, or, even, images, eye-tracking data). We will then discuss how to evaluate such representations, intrinsically and extrinsically. We will introduce researchers to state-of-the-art methods for constructing cross-lingual word representations and discuss their applicability in a broad range of downstream NLP applications.

We will deliver a detailed survey of the current methods, discuss best training and evaluation practices and use-cases, and provide links to publicly available implementations, datasets, and pre-trained models.

Manaal Faruqui is a research scientist at Google NYC currently working on industrial-scale NLP problems. Manaal received his PhD in the Language Technologies Institute at Carnegie Mellon University. He has worked on problems in the areas of representation learning, distributional semantics and multilingual learning. He has won one of the best paper awards at NAACL 2015. He organized the workshop on cross-lingual and multilingual models in NLP at NAACL 2016.

Anders Søgaard is a full professor of Computer Science (NLP and Machine Learning) at the University of Copenhagen. Anders is interested in transfer learning and has worked on semi-supervised learning, domain adaptation, and cross-language adaptation of NLP models. He is particularly interested in transferring models to very low-resource languages. He holds an ERC Starting Grant, as well as several grants from national research councils and private research foundations. He has won three best paper awards at major ACL conferences. He gave a tutorial on domain adaptation at COLING 2014.

Ivan Vulić is a research associate at the University of Cambridge. He received his PhD summa cum laude at KU Leuven in 2014. Ivan is interested in representation learning, distributional and multi-modal semantics in monolingual and multilingual contexts, and transfer learning for enabling cross-lingual NLP applications. His work has been published in top-tier ACL and IR conferences. He gave a tutorial on multilingual topic models at ECIR 2013 and WSDM 2014, and organized a Vision & Language workshop at EMNLP 2015.

4

Workshops

Note: all workshops are located in **CPH Conference**. Please see the map with room locations in the front matter.

Thursday–Friday

TODO	Second Conference on Machine Translation (WMT)	p.26
------	--	------

Thursday

TODO	Subword and Character LEvel Models in NLP	p.32
TODO	Natural Language Processing meets Journalism	p.34
TODO	3rd Workshop on Noisy User-generated Text	p.36
TODO	2nd Workshop on Structured Prediction for Natural Language Processing	p.38
TODO	New Frontiers in Summarization	p.39
TODO	Workshop on Speech-Centric Natural Language Processing	p.41

Friday

TODO	3rd Workshop on Discourse in Machine Translation	p.43
TODO	Workshop on Stylistic Variation	p.45
TODO	12th Workshop on Innovative Use of NLP for Building Educational Applications	p.47
TODO	Workshop on Argument Mining	p.51
TODO	8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis	p.53
TODO	2nd Workshop on Evaluating Vector Space Representations for NLP	p.56
TODO	Building Linguistically Generalizable NLP Systems	p.58

Workshop 1: Second Conference on Machine Translation (WMT)

Organizers: *Philipp Koehn and Barry Haddow*

Venue: TODO

Thursday, September 7, 2016

8:45–9:00 **Opening Remarks**

9:00–10:30 **Session 1: Shared Tasks Overview Presentations I**

9:00–9:40 **Shared Task: News Translation**

- Findings of the 2017 Conference on Machine Translation (WMT17)
Ondřej Bojar, Rajen Chatterjee, Christian Federmann, Yvette Graham, Barry Haddow, Matthias Huck, Philipp Koehn, Varvara Logacheva, Christof Monz, Matteo Negri, Matt Post, Raphael Rubino, Lucia Specia, and Marco Turchi

9:40–10:10 **Shared Task: Multimodal Translation**

- Findings of the Second Shared Task on Multimodal Machine Translation and Multilingual Image Description
Desmond Elliott, Stella Frank, Loïc Barrault, Fethi Bougares, and Lucia Specia

10:10–10:30 **Shared Task: Biomedical Translation**

- Findings of the WMT 2017 Biomedical Translation Shared Task
Antonio Jimeno Yepes, Aurelie Neveu, Mariana Neves, Karin Verspoor, Ondřej Bojar, Arthur Boyer, Cristian Grozea, Barry Haddow, Madeleine Kittner, Yvonne Lichtblau, Pavel Pecina, Roland Roller, Rudolf Rosa, Amy Siu, Philippe Thomas, and Saskia Trescher

10:30–11:00 **Coffee Break**

11:00–12:30 **Session 2: Shared Tasks Poster Session I**

11:00–12:30 **Shared Task: News Translation**

- CUNI submission in WMT17: Chimera goes neural
Roman Sudarikov, David Mareček, Tom Kocmi, Dusan Varis, and Ondřej Bojar
- LIMSIWMT'17
Franck Burlot, Pooyan Safari, Matthieu Labeau, Alexandre Allauzen, and François Yvon
- SYSTRAN Purely Neural MT Engines for WMT2017
Yongchao Deng, Jungi Kim, Guillaume Klein, Catherine KOBUS, Natalia Segal, Christophe Servan, Bo Wang, Dakun Zhang, Josep Crego, and Jean Senellart
- FBK's Participation to the English-to-German News Translation Task of WMT 2017
Mattia Antonino Di Gangi, Nicola Bertoldi, and Marcello Federico
- The JHU Machine Translation Systems for WMT 2017
Shuoyang Ding, Huda Khayrallah, Philipp Koehn, Matt Post, Gaurav Kumar, and Kevin Duh

- The TALP-UPC Neural Machine Translation System for German/Finnish-English Using the Inverse Direction Model in Rescoring
Carlos Escolano, Marta R. Costa-jussà, and José A. R. Fonollosa
- LIUM Machine Translation Systems for WMT17 News Translation Task
Mercedes García-Martínez, Ozan Caglayan, Walid Aransa, Adrien Bardet, Fethi Bougares, and Loïc Barrault
- Extending hybrid word-character neural machine translation with multi-task learning of morphological analysis
Stig-Arne Grönroos, Sami Virpioja, and Mikko Kurimo
- The AFRL-MITLL WMT17 Systems: Old, New, Borrowed, BLEU
Jeremy Gwinnup, Timothy Anderson, Grant Erdmann, Katherine Young, Michael Kazi, Elizabeth Salesky, Brian Thompson, and Jonathan Taylor
- University of Rochester WMT 2017 NMT System Submission
Chester Holtz, Chuyang Ke, and Daniel Gildea
- LMU Munich’s Neural Machine Translation Systems for News Articles and Health Information Texts
Matthias Huck, Fabienne Braune, and Alexander Fraser
- Rule-based Machine translation from English to Finnish
Arovi Hurskainen and Jörg Tiedemann
- NRC Machine Translation System for WMT 2017
Chi-ku Lo, Boxing Chen, Colin Cherry, George Foster, Samuel Larkin, Darlene Stewart, and Roland Kuhn
- The Helsinki Neural Machine Translation System
Robert Östling, Yves Scherrer, Jörg Tiedemann, Gongbo Tang, and Tommi Nieminen
- The QT21 Combined Machine Translation System for English to Latvian
Jan-Thorsten Peter, Hermann Ney, Ondřej Bojar, Ngoc-Quan Pham, Jan Niehues, Alex Waibel, Franck Burlot, François Yvon, Mārcis Pinnis, Valters Šics, Joost Bastings, Miguel Rios, Wilker Aziz, Philip Williams, Frédéric Blain, and Lucia Specia
- The RWTH Aachen University English-German and German-English Machine Translation System for WMT 2017
Jan-Thorsten Peter, Andreas Guta, Tamer Alkhouri, Parnia Bahar, Jan Rosendahl, Nick Rossenbach, Miguel Graça, and Hermann Ney
- The Karlsruhe Institute of Technology Systems for the News Translation Task in WMT 2017
Ngoc-Quan Pham, Jan Niehues, Thanh-Le Ha, Eunah Cho, Matthias Sperber, and Alexander Waibel
- Tilde’s Machine Translation Systems for WMT 2017
Mārcis Pinnis, Rihards Krišlauks, Tomis Miks, Daiga Deksne, and Valters Šics
- C-3MA: Tartu-Riga-Zurich Translation Systems for WMT17
Matiss Rikters, Chantal Amrhein, Maksym Del, and Mark Fishel
- The University of Edinburgh’s Neural MT Systems for WMT17
Rico Sennrich, Alexandra Birch, Anna Currey, Ulrich Germann, Barry Haddow, Kenneth Heafield, Antonio Valerio Miceli Barone, and Philip Williams
- XMU Neural Machine Translation Systems for WMT 17
Zhixing Tan, Boli Wang, Jimming Hu, and Yidong Chen
- The JAIST Machine Translation Systems for WMT 17
Long Trieu, Trung-Tin Pham, and Le-Minh Nguyen
- Sogou Neural Machine Translation Systems for WMT17
Yuguang Wang, Shanbo Cheng, Liyang Jiang, Jiajun Yang, Wei Chen, Muze Li, Lin Shi, Yanfeng Wang, and Hongtao Yang

- PJIIT's systems for WMT 2017 Conference
Krzysztof Wolk and Krzysztof Marasek
- Hunter MT: A Course for Young Researchers in WMT17
Jia Xu, Yi Zong Kuang, Shondell Baijoo, Jacob Hyun Lee, Uman Shahzad, Mir Ahmed, Meredith Lancaster, and Chris Carlan
- CASICT-DCU Neural Machine Translation Systems for WMT17
Jinchao Zhang, Peerachet Porkaew, Jiawei Hu, Qiuye Zhao, and Qun Liu

11:00–12:30 Shared Task: Multi-Modal Translation

- LIUM-CVC Submissions for WMT17 Multimodal Translation Task
Ozan Caglayan, Walid Aransa, Adrien Bardet, Mercedes García-Martínez, Fethi Bougares, Loïc Barrault, Marc Masana, Luis Herranz, and Joost van de Weijer
- DCU System Report on the WMT 2017 Multi-modal Machine Translation Task
Iacer Calixto, Koel Dutta Chowdhury, and Qun Liu
- The AFRL-OSU WMT17 Multimodal Translation System: An Image Processing Approach
Jeremy Gwinnup, John Duselis, Michael Hutt, James Davis, and Joshua Sandwick
- CUNI System for the WMT17 Multimodal Translation Task
Jindřich Helcl and Jindřich Libovický
- Generating Image Descriptions using Multilingual Data
Alan Jaffe
- OSU Multimodal Machine Translation System Report
Mingbo Ma, Dapeng Li, Kai Zhao, and Liang Huang
- Sheffield MultiMT: Using Object Posterior Predictions for Multimodal Machine Translation
Pranava Swaroop Madhyastha, Josiah Wang, and Lucia Specia
- NICT-NAIST System for WMT17 Multimodal Translation Task
Jingyi Zhang, Masao Utiyama, Eiichiro Sumita, Graham Neubig, and Satoshi Nakamura

11:00–12:30 Shared Task: Biomedical Translation

- Automatic Threshold Detection for Data Selection in Machine Translation
Mirela-Stefania Duma and Wolfgang Menzel

12:30–14:00 Lunch

14:00–15:30 Session 3: Invited Talk

14:00–15:30 Holger Schwenk (Facebook): Multilingual Representations and Applications in NLP

15:30–16:00 Coffee Break

16:00–17:30 Session 4: Research Papers on Lexicon and Morphology

16:00–16:15 Sense-Aware Statistical Machine Translation using Adaptive Context-Dependent Clustering *Xiao Pu, Nikolaos Pappas, and Andrei Popescu-Belis*

16:15–16:30 Improving Word Sense Disambiguation in Neural Machine Translation with Sense Embeddings *Annette Rios Gonzales, Laura Mascarell, and Rico Sennrich*

16:30–16:45 Word Representations in Factored Neural Machine Translation *Franck Burlot, Mercedes García-Martínez, Loïc Barrault, Fethi Bougares, and François Yvon*

- 16:45–17:00 Modeling Target-Side Inflection in Neural Machine Translation
Aleš Tamchyna, Marion Weller-Di Marco, and Alexander Fraser
- 17:00–17:15 Evaluating the morphological competence of Machine Translation Systems
Franck Burlot and François Yvon
- 17:15–17:30 Target-side Word Segmentation Strategies for Neural Machine Translation
Matthias Huck, Simon Riess, and Alexander Fraser

Friday, September 8, 2017

- 9:00–10:30 **Session 5: Shared Tasks Overview Presentations II**
- 9:00–9:20 **Shared Task: Quality Estimation**
- 9:20–9:40 **Shared Task: Metrics**
- Results of the WMT17 Metrics Shared Task
Ondřej Bojar, Yvette Graham, and Amir Kamran
- 9:40–10:00 **Shared Task: Automatic Post-Editing**
- 10:00–10:15 **Shared Task: Bandit Learning**
- A Shared Task on Bandit Learning for Machine Translation
Artem Sokolov, Julia Kreutzer, Kellen Sunderland, Pavel Danchenko, Witold Szymaniak, Hagen Fürstenau, and Stefan Riezler
- 10:15–10:30 **Shared Task: Neural Training**
- Results of the WMT17 Neural MT Training Task
Ondřej Bojar, Jindřich Helcl, Tom Kocmi, Jindřich Libovický, and Tomáš Musil
- 10:30–11:00 **Coffee Break**
- 11:00–12:30 **Session 6: Shared Tasks Poster Session II**
- 11:00–12:30 **Shared Task: Quality Estimation**
- Sentence-level quality estimation by predicting HTER as a multi-component metric
Eleftherios Avramidis
 - Predicting Translation Performance with Referential Translation Machines
Ergun Biçici
 - Bilexical Embeddings for Quality Estimation
Frédéric Blain, Carolina Scarton, and Lucia Specia
 - Improving Machine Translation Quality Estimation with Neural Network Features
Zhiming Chen, Yiming Tan, Chenlin Zhang, Qingyu Xiang, Lilin Zhang, Maoxi Li, and Mingwen WANG
 - UHH Submission to the WMT17 Quality Estimation Shared Task
Melania Duma and Wolfgang Menzel

- Predictor-Estimator using Multilevel Task Learning with Stack Propagation for Neural Quality Estimation
Hyun Kim, Jong-Hyeok Lee, and Seung-Hoon Na
- Unbabel's Participation in the WMT17 Translation Quality Estimation Shared Task
André F. T. Martins, Fabio Kepler, and Jose Monteiro
- Feature-Enriched Character-Level Convolutions for Text Regression
Gustavo Paetzold and Lucia Specia

11:00–12:30 Shared Task: Metrics

- UHH Submission to the WMT17 Metrics Shared Task
Melanie Duma and Wolfgang Menzel
- MEANT 2.0: Accurate semantic MT evaluation for any output language
Chi-ku Lo
- Blend: a Novel Combined MT Metric Based on Direct Assessment — CASICT-DCU submission to WMT17 Metrics Task
Qingsong Ma, Yvette Graham, Shugen Wang, and Qun Liu
- CUNI Experiments for WMT17 Metrics Task
David Mareček, Ondřej Bojar, Ondřej Hübisch, Rudolf Rosa, and Dusan Varis
- chrF++: words helping character n-grams
Maja Popović
- bleu2vec: the Painfully Familiar Metric on Continuous Vector Space Steroids
Andre Tättar and Mark Fishel

11:00–12:30 Shared Task: Automatic Post-Editing

- LIG-CRISTAL Submission for the WMT 2017 Automatic Post-Editing Task
Alexandre Berard, Laurent Besacier, and Olivier Pietquin
- Multi-source Neural Automatic Post-Editing: FBK's participation in the WMT 2017 APE shared task
Rajen Chatterjee, M. Amin Farajian, Matteo Negri, Marco Turchi, Ankit Srivastava, and Santanu Pal
- The AMU-UEdin Submission to the WMT 2017 Shared Task on Automatic Post-Editing
Marcin Junczys-Dowmunt and Marcin Junczys-Dowmunt
- Ensembling Factored Neural Machine Translation Models for Automatic Post-Editing and Quality Estimation
Chris Hokamp
- Neural Post-Editing Based on Quality Estimation
Yiming Tan, Zhiming Chen, Liu Huang, Lilin Zhang, Maoxi Li, and Mingwen Wang
- CUNI System for WMT17 Automatic Post-Editing Task
Dusan Varis and Ondřej Bojar

11:00–12:30 Shared Task: Bandit Learning

- The UMD Neural Machine Translation Systems at WMT17 Bandit Learning Task
Amr Sharaf, Shi Feng, Khanh Nguyen, Kiante Brantley, and Hal Daumé III
- LIMSI Submission for WMT'17 Shared Task on Bandit Learning
Guillaume Wisniewski

11:00–12:30 Shared Task: Neural Training

- Variable Mini-Batch Sizing and Pre-Trained Embeddings
Mostafa Abdou, Vladan Gloncak, and Ondřej Bojar

- The AFRL WMT17 Neural Machine Translation Training Task Submission
Jeremy Gwinnup, Grant Erdmann, and Katherine Young

12:30–14:00 **Lunch**

14:00–15:15 Session 7: Research Papers on Syntax and Deep Models

- 14:00–14:15 Predicting Target Language CCG Supertags Improves Neural Machine Translation

Maria Nadejde, Siva Reddy, Rico Sennrich, Tomasz Dwojak, Marcin Junczys-Dowmunt, Philipp Koehn, and Alexandra Birch

- 14:15–14:30 Exploiting Linguistic Resources for Neural Machine Translation Using Multi-task Learning

Jan Niehues and Eunah Cho

- 14:30–14:45 Tree as a Pivot: Syntactic Matching Methods in Pivot Translation
Akiva Miura, Graham Neubig, Katsuhiro Sudoh, and Satoshi Nakamura

- 14:45–15:00 Deep architectures for Neural Machine Translation

Antonio Valerio Miceli Barone, Jindřich Helcl, Rico Sennrich, Barry Haddow, and Alexandra Birch

- 15:00–15:15 Biasing Attention-Based Recurrent Neural Networks Using External Alignment Information

Tamer Alkhouri and Hermann Ney

15:15–16:00 **Coffee Break**

16:00–17:15 Session 8: Research Papers on Domain Adaptation and External Data

- 16:00–16:15 Effective Domain Mixing for Neural Machine Translation

Denny Britz, Quoc Le, and Reid Pryzant

- 16:15–16:30 Multi-Domain Neural Machine Translation through Unsupervised Adaptation

M. Amin Farajian, Marco Turchi, Matteo Negri, and Marcello Federico

- 16:30–16:45 Adapting Neural Machine Translation with Parallel Synthetic Data
Mara Chinea-Rios, Álvaro Peris, and Francisco Casacuberta

- 16:45–17:00 Copied Monolingual Data Improves Low-Resource Neural Machine Translation

Anna Currey, Antonio Valerio Miceli Barone, and Kenneth Heafield

- 17:00–17:15 Guiding Neural Machine Translation Decoding with External Knowledge

Rajen Chatterjee, Matteo Negri, Marco Turchi, Marcello Federico, Lucia Specia, and Frédéric Blain

Workshop 2: Subword and Character LLevel Models in NLP

Organizers: *Manaal Faruqui, Hinrich Schütze, Isabel Trancoso, and Yadollah Yaghoobzadeh*

Venue: TODO

Thursday, September 7, 2017

09:00–09:10 **Opening Remarks (Manaal Faruqui)**

09:10–09:50 **Subword-level Information in NLP using Neural Networks
(Tomas Mikolov)**

09:50–10:30 **Chewing the Fat about Mincing Words (Invited Talk by Noah Smith)**

10:30–11:00 **Coffee break**

11:00–11:40 **Neural WFSTs: Tutorial Talk**

11:40–12:10 **Best paper presentations**

12:10–14:00 **Poster session & Lunch break**

- Character and Subword-Based Word Representation for Neural Language Modeling Prediction
Matthieu Labeau and Alexandre Alauzen
- Learning variable length units for SMT between related languages via Byte Pair Encoding
Anoop Kunchukuttan and Pushpak Bhattacharyya
- Character Based Pattern Mining for Neology Detection
Gaël Lejeune and Emmanuel Cartier
- Automated Word Stress Detection in Russian
Maria Ponomareva, Kirill Milintsevich, Ekaterina Chernyak, and Anatoly Starostin
- A Syllable-based Technique for Word Embeddings of Korean Words
Sanghyuk Choi, Taeuk Kim, Jinseok Seol, and Sang-goo Lee
- Supersense Tagging with a Combination of Character, Subword, and Word-level Representations
Youhyun Shin and Sang-goo Lee
- Weakly supervised learning of allomorphy
Miikka Silfverberg and Mans Hulden
- Character-based recurrent neural networks for morphological relational reasoning
Olof Mogren and Richard Johansson
- Glyph-aware Embedding of Chinese Characters
Falcon Dai and Zheng Cai
- Exploring Cross-Lingual Transfer of Morphological Knowledge In Sequence-to-Sequence Models
Huiming Jin and Katharina Kann
- Language Generation with Recurrent Generative Adversarial Networks without Pre-training
Ofir Press, Amir Bar, Ben Bogin, Jonathan Berant, and Lior Wolf

- EXTENDED ABSTRACT Align and Copy: Hard Attention Models for Morphological Inflection Generation
Tatyana Ruzsics, Peter Makarov, and Simon Clematide
- (EXTENDED ABSTRACT) Natural Language Generation through Character-Based RNNs with Finite-State Prior Knowledge
Raghav Goyal, Marc Dymetman, and Eric Gaussier
- Patterns versus Characters in Subword-aware Neural Language Modeling
Zhenisbek Assylbekov and Rustem Takhanov

14:00–14:40 **Fully Character Level Neural Machine Translation**

14:40–15:50 **Poster session & Coffee break**

- Unlabeled Data for Morphological Generation With Character-Based Sequence-to-Sequence Models
Katharina Kann and Hinrich Schütze
- Vowel and Consonant Classification through Spectral Decomposition
Patricia Thaine and Gerald Penn
- Syllable-level Neural Language Model for Agglutinative Language
Seunghak Yu, Nilesh Kulkarni, Haejun Lee, and Jihie Kim
- Character-based Bidirectional LSTM-CRF with words and characters for Japanese Named Entity Recognition
Shotaro Misawa, Motoki Taniguchi, Yasuhide Miura, and Tomoko Okuma
- Word Representation Models for Morphologically Rich Languages in Neural Machine Translation
Ekatерина Vylomova, Trevor Cohn, Xuanli He, and Gholaemreza Haffari
- Spell-Checking based on Syllabification and Character-level Graphs for a Peruvian Agglutinative Language
Carlo Alva and Arturo Oncevay
- What do we need to know about an unknown word when parsing German
Bich-Ngoc Do, Ines Rehbein, and Anette Frank
- A General-Purpose Tagger with Convolutional Neural Networks
Xiang Yu, Agnieszka Falenska, and Ngoc Thang Vu
- Reconstruction of Word Embeddings from Sub-Word Parameters
Karl Stratos
- Inflection Generation for Spanish Verbs using Supervised Learning
Cristina Barros, Dimitra Giatzia, and Elena Lloret
- Neural Paraphrase Identification of Questions with Noisy Pretraining
Gaurav Singh Tomar, Thyago Duque, Oscar Täckström, Jakob Uszkoreit, and Dipanjan Das
- Sub-character Neural Language Modelling in Japanese
Viet Nguyen, Julian Brooke, and Timothy Baldwin
- Byte-based Neural Machine Translation
Marta R. Costa-jussà, Carlos Escolano, and José A. R. Fonollosa
- Improving Opinion-Target Extraction with Character-Level Word Embeddings
Soufian Jebara and Philipp Cimiano

15:50–16:30 **Acoustic Word Embeddings**

16:30–17:30 **Panel discussion**

17:30–17:45 **Closing remarks (Hinrich Schuetze)**

Workshop 3: Natural Language Processing meets Journalism

Organizers: *Octavian Popescu and Carlo Strapparava*

Venue: TODO

- Predicting News Values from Headline Text and Emotions
Maria Pia di Buono, Jan Šnajder, Bojana Dalbelo Basic, Goran Glavaš, Martin Tutek, and Natasa Milic-Frayling
- Predicting User Views in Online News
Daniel Hardt and Owen Rambow
- Tracking Bias in News Sources Using Social Media: the Russia-Ukraine Maidan Crisis of 2013–2014
Peter Potash, Alexey Romanov, Mikhail Gronas, Anna Rumshisky, and Mikhail Gronas
- What to Write? A topic recommender for journalists
Alessandro Cucchiarelli, Christian Morbidoni, Giovanni Stilo, and Paola Velardi
- Comparing Attitudes to Climate Change in the Media using sentiment analysis based on Latent Dirichlet Allocation
Ye Jiang, Xingyi Song, Jackie Harrison, Shaun Quegan, and Diana Maynard
- Language-based Construction of Explorable News Graphs for Journalists
Rémi Bois, Guillaume Gravier, Eric Jamet, Emmanuel Morin, Pascale Sébillot, and Maxime Robert
- Storyteller: Visual Analytics of Perspectives on Rich Text Interpretations
Maarten van Meersbergen, Piek Vossen, Janneke van der Zwaan, Antske Fokkens, Willem van Hage, Inger Leemans, and Isa Maks
- Analyzing the Revision Logs of a Japanese Newspaper for Article Quality Assessment
Hideaki Tamori, Yuta Hitomi, Naoaki Okazaki, and Kentaro Inui
- Improved Abusive Comment Moderation with User Embeddings
John Pavlopoulos, Prodromos Malakasiotis, Juli Bakagianni, and Ion Androutsopoulos
- Incongruent Headlines: Yet Another Way to Mislead Your Readers
Sophie Chesney, Maria Liakata, Massimo Poesio, and Matthew Purver
- Unsupervised Event Clustering and Aggregation from Newswire and Web Articles
Swen Ribeiro, Olivier Ferret, and Xavier Tannier
- Semantic Storytelling, Cross-lingual Event Detection and other Semantic Services for a Newsroom Content Curation Dashboard
Julian Moreno-Schneider, Ankit Srivastava, Peter Bourgonje, David Wabnitz, and Georg Rehm
- Deception Detection in News Reports in the Russian Language: Lexics and Discourse
Dina Pisarevskaya
- Fake news stance detection using stacked ensemble of classifiers
James Thorne, Mingjie Chen, Giorgos Myrianthous, Jiashu Pu, Xiaoxuan Wang, and Andreas Vlachos
- From Clickbait to Fake News Detection: An Approach based on Detecting the Stance of Headlines to Articles
Peter Bourgonje, Julian Moreno Schneider, and Georg Rehm

- 'Fighting' or 'Conflict'? An Approach to Revealing Concepts of Terms in Political Discourse
Linyuan Tang and Kyo Kageura
- A News Chain Evaluation Methodology along with a Lattice-based Approach for News Chain Construction
Mustafa Toprak, Özer Özkahraman, and Selma Tekir
- Using New York Times Picks to Identify Constructive Comments
Varada Kolhatkar and Maite Taboada
- An NLP Analysis of Exaggerated Claims in Science News
YINGYA LI, Jieke Zhang, and Bei Yu

Workshop 4: 3rd Workshop on Noisy User-generated Text

Organizers: *Kai-Wei Chang, Ming-Wei Chang, Vivek Srikumar, and Alexander M. Rush*

Venue: TODO

September 7

9:00–9:05 **Opening**

9:05–9:50 **Invited Talk: Common Sense Knowledge as an Emergent Property of Neural Conversational Models (Bill Dolan)**

9:50–10:35 **Oral Session I**

9:50–10:05 Boundary-based MWE segmentation with text partitioning
Jake Williams

10:05–10:20 Towards the Understanding of Gaming Audiences by Modeling Twitch Emotes
Francesco Barbieri, Luis Espinosa Anke, Miguel Ballesteros, Juan Soler, and Horacio Saggion

10:20–10:35 Churn Identification in Microblogs using Convolutional Neural Networks with Structured Logical Knowledge
Mourad Gridach, Hatem Haddad, and Hala Mulk

10:35–11:00 **Coffee Break**

11:00–12:30 **Oral Session II**

11:00–11:15 To normalize, or not to normalize: The impact of normalization on Part-of-Speech tagging
Rob van der Goot, Barbara Plank, and Malvina Nissim

11:15–11:30 Constructing an Alias List for Named Entities during an Event
Anietie Andy, Mark Dredze, Mugizi Rwebangira, and Chris Callison-Burch

11:30–11:45 Incorporating Metadata into Content-Based User Embeddings
Linzi Xing and Michael J. Paul

11:45–12:00 Simple Queries as Distant Labels for Predicting Gender on Twitter
Chris Emmery, Grzegorz Chrupala, and Walter Daelemans

12:00–12:15 A Dataset and Classifier for Recognizing Social Media English
Su Lin Blodgett, Johnny Wei, and Brendan O'Connor

12:15–12:30 Evaluating hypotheses in geolocation on a very large sample of Twitter
Bahar Salehi and Anders Søgaard

12:30–14:00 **Lunch**

14:00–14:45 **Invited Talk: Tweets in Finance (Miles Osborne)**

14:45–14:55 **Lightning Talks**

- The Effect of Error Rate in Artificially Generated Data for Automatic Preposition and Determiner Correction
Fraser Bowen, Jon Dehdari, and Josef Van Genabith
- An Entity Resolution Approach to Isolate Instances of Human Trafficking Online
Chirag Nagpal, Kyle Miller, Benedikt Boecking, and Artur Dubrawski

- Noisy Uyghur Text Normalization
Osman Tursun and Ruket Cakici
- Crowdsourcing Multiple Choice Science Questions
Johannes Welbl, Nelson F. Liu, and Matt Gardner
- A Text Normalisation System for Non-Standard English Words
Emma Flint, Elliot Ford, Olivia Thomas, Andrew Caines, and Paula Buttery
- Huntsville, hospitals, and hockey teams: Names can reveal your location
Bahar Salehi, Dirk Hovy, Eduard Hovy, and Anders Søgaard
- Improving Document Clustering by Removing Unnatural Language
Myungha Jang, Jinho D. Choi, and James Allan
- Lithium NLP: A System for Rich Information Extraction from Noisy User Generated Text on Social Media
Preeti Bhargava, Nemanja Spasojevic, and Guoning Hu

14:55–15:30 **Shared Task Session**

14:55–15:10 Results of the WNUT2017 Shared Task on Novel and Emerging Entity Recognition
Leon Derczynski, Eric Nichols, Marieke van Erp, and Nut Limsopatham

15:10–15:20 A Multi-task Approach for Named Entity Recognition in Social Media Data
Gustavo Aguilar, Suraj Maharjan, Adrian Pastor López Monroy, and Thamar Solorio

15:20–15:30 Distributed Representation, LDA Topic Modelling and Deep Learning for Emerging Named Entity Recognition from Social Media
Patrick Jansson and Shuhua Liu

- Multi-channel BiLSTM-CRF Model for Emerging Named Entity Recognition in Social Media
Bill Y. Lin, Frank Xu, Zhiyi Luo, and Kenny Zhu
- Transfer Learning and Sentence Level Features for Named Entity Recognition on Tweets
Pius von Däniken and Mark Cieliebak
- Context-Sensitive Recognition for Emerging and Rare Entities
Jake Williams and Giovanni Santia
- A Feature-based Ensemble Approach to Recognition of Emerging and Rare Named Entities
Utpal Kumar Sikdar and Björn Gambäck

15:30–16:30 **Poster Session**

16:30–17:15 **Invited Talk: Modeling Language as a Social Construct (Dirk Hovy)**

17:15–17:30 **Closing and Best Paper Awards**

Workshop 5: 2nd Workshop on Structured Prediction for Natural Language Processing

Organizers: *Leon Derczynski, Wei Xu, Alan Ritter, and Timothy Baldwin*

Venue: TODO

Thursday, September 7, 2016

9:00–10:30 **Section 1**

9:00–9:15 **Welcome (Organizers)**

9:15–10:00 **Invited Talk**

10:00–10:30 Dependency Parsing with Dilated Iterated Graph CNNs
Emma Strubell and Andrew McCallum

10:30–11:00 **Coffee Break**

11:00–12:15 **Section 2**

11:00–11:45 **Invited Talk**

11:45–12:15 **Poster Madness**

12:15–14:00 **Lunch**

14:00–15:30 **Section 3**

14:00–14:45 **Poster Session**

- Entity Identification as Multitasking
Karl Stratos
- Towards Neural Machine Translation with Latent Tree Attention
James Bradbury and Richard Socher
- Structured Prediction via Learning to Search under Bandit Feedback
Amr Sharaf and Hal Daumé III
- Syntax Aware LSTM model for Semantic Role Labeling
Feng Qian, Lei Sha, Baobao Chang, LuChen Liu, and Ming Zhang
- Spatial Language Understanding with Multimodal Graphs using Declarative Learning based Programming
Parisa Kordjamshidi, Taher Rahgooy, and Umar Manzoor
- Boosting Information Extraction Systems with Character-level Neural Networks and Free Noisy Supervision
Philipp Meerkamp and Zhengyi Zhou

14:45–15:30 **Invited Talk**

15:30–16:00 **Coffee Break**

16:00–17:30 **Section 4**

16:00–16:45 **Invited Talk**

16:45–17:15 Piecewise Latent Variables for Neural Variational Text Processing
Iulian Vlad Serban, Alexander Ororbia II, Joelle Pineau, and Aaron Courville

17:15–17:30 **Closing**

Workshop 6: New Frontiers in Summarization

Organizers: *Lu Wang, Jackie Chi Kit Cheung, Giuseppe Carenini, and Fei Liu*

Venue: TODO

08:45–10:30 Morning Session 1

08:45–08:50 Opening Remarks

08:50–09:50 Invited Talk (Andreas Kerren)

09:50–10:10 Video Highlights Detection and Summarization with Lag-Calibration
based on Concept-Emotion Mapping of Crowdsourced Time-Sync
Comments

Qing Ping and Chaomei Chen

10:10–10:30 Multimedia Summary Generation from Online Conversations: Current
Approaches and Future Directions
Enamul Hoque and Giuseppe Carenini

10:30–11:00 Break

11:00–12:30 Morning Session 2

11:00–12:00 Invited Talk (Katja Filippova)

12:00–12:15 Low-Resource Neural Headline Generation
Ottokar Tilk and Tanel Alumäe

12:15–12:30 Towards Improving Abstractive Summarization via Entailment
Generation
Ramakanth Pasunuru, Han Guo, and Mohit Bansal

12:30–14:00 Lunch

14:00–15:30 Poster Session

- Coarse-to-Fine Attention Models for Document Summarization
Jeffrey Ling and Alexander Rush
- Automatic Community Creation for Abstractive Spoken Conversations
Summarization
Karan Singla, Evgeny Stepanov, Ali Orkan Bayer, Giuseppe Carenini, and Giuseppe Riccardi
- Combining Graph Degeneracy and Submodularity for Unsupervised
Extractive Summarization
Antoine Tixier, Polykarpos Meladianos, and Michalis Vazirgiannis
- TL;DR: Mining Reddit to Learn Automatic Summarization
Michael Völske, Martin Potthast, Shahbaz Syed, and Benno Stein
- Topic Model Stability for Hierarchical Summarization
John Miller and Kathleen McCoy
- Learning to Score System Summaries for Better Content Selection
Evaluation.
Maxime Peyrard, Teresa Botschen, and Iryna Gurevych
- Revisiting the Centroid-based Method: A Strong Baseline for
Multi-Document Summarization
Demian Gholipour Ghalandari

- Reader-Aware Multi-Document Summarization: An Enhanced Model
and The First Dataset
Piji Li, Lidong Bing, and Wai Lam
- A Pilot Study of Domain Adaptation Effect for Neural Abstractive
Summarization
Xinyu Hua and Lu Wang

15:30–17:15 **Afternoon Session**

15:30–16:30 **Invited Talk (Ani Nenkova)**

17:10–17:15 **Closing Remarks**

Workshop 7: Workshop on Speech-Centric Natural Language Processing

Organizers: *Nicholas Ruiz and Srinivas Bangalore*

Venue: TODO

September 7, 2017

8:50–9:00 Opening Remarks (**Nicholas Ruiz and Srinivas Bangalore**)

9:00–10:00 Invited Talk

10:00–10:30 Session I

- Functions of Silences towards Information Flow in Spoken Conversation
Shammur Absar Chowdhury, Evgeny Stepanov, Morena Danieli, and Giuseppe Riccardi

10:30–11:00 Coffee Break

11:00–12:30 Session II

- Encoding Word Confusion Networks with Recurrent Neural Networks for Dialog State Tracking
Glorianna Jagfeld and Ngoc Thang Vu
- Analyzing Human and Machine Performance In Resolving Ambiguous Spoken Sentences
Hussein Ghaly and Michael Mandel
- Parsing transcripts of speech
Andrew Caines, Michael McCarthy, and Paula Buttery
- Enriching ASR Lattices with POS Tags for Dependency Parsing
Moritz Stiefel and Ngoc Thang Vu

12:30–14:00 Lunch

14:00–15:30 Session III

- End-to-End Information Extraction without Token-Level Supervision
Rasmus Berg Palm, Dirk Hovy, Florian Laws, and Ole Winther
- Spoken Term Discovery for Language Documentation using Translations
Antonios Anastasopoulos, Sameer Bansal, David Chiang, Sharon Goldwater, and Adam Lopez
- Amharic-English Speech Translation in Tourism Domain
Michael Melese, Laurent Besacier, and Million Meshesha
- Speech- and Text-driven Features for Automated Scoring of English Speaking Tasks
Anastassia Loukina, Nitin Madnani, and Aoife Cahill

15:30–16:00 Coffee Break / Poster Discussion

16:00–16:25 Session IV

- Improving coreference resolution with automatically predicted prosodic information
Ina Roesiger, Sabrina Stehwien, Arndt Riester, and Ngoc Thang Vu

One-day Workshops

16:25–17:50 **Round-table: Issues in Speech-centric NLP**

17:50–18:00 **Closing**

Workshop 8: 3rd Workshop on Discourse in Machine Translation

Organizers: *Bonnie Webber, Andrei Popescu-Belis, and Jörg Tiedemann*

Venue: TODO

Friday, September 8, 2017

09:00–10:30 Session 1

09:00–09:10 Introduction

09:10–09:40 Findings of the 2017 DiscoMT Shared Task on Cross-lingual Pronoun Prediction

Sharid Loáiciga, Sara Stymne, Preslav Nakov, Christian Hardmeier, Jörg Tiedemann, Mauro Cettolo, and Yannick Versley

09:40–10:10 Validation of an Automatic Metric for the Accuracy of Pronoun Translation (APT)

Lesly Miculicich Werlen and Andrei Popescu-Belis

10:10–10:30 Poster Boaster

10:30–11:00 Coffee Break

11:00–12:30 Session 2a: Regular Track Posters

- Using a Graph-based Coherence Model in Document-Level Machine Translation

Leo Born, Mohsen Mesgar, and Michael Strube

- Treatment of Markup in Statistical Machine Translation

Mathias Müller

11:00–12:30 Session 2b: Shared Task Posters

- A BiLSTM-based System for Cross-lingual Pronoun Prediction

Sara Stymne, Sharid Loáiciga, and Fabienne Cap

- Neural Machine Translation for Cross-Lingual Pronoun Prediction

Sébastien Jean, Stanislas Lautry, Orhan Firat, and Kyunghyun Cho

- Predicting Pronouns with a Convolutional Network and an N-gram Model

Christian Hardmeier

- Cross-Lingual Pronoun Prediction with Deep Recurrent Neural Networks v2.0

Juhani Luotolahti, Jenna Kanerva, and Filip Ginter

11:00–12:30 Session 2c: Posters Related to Oral Presentations

- Combining the output of two coreference resolution systems for two source languages to improve annotation projection

Yulia Grishina

- Discovery of Discourse-Related Language Contrasts through Alignment Discrepancies in English-German Translation

Ekaterina Lapshinova-Koltunski and Christian Hardmeier

- Neural Machine Translation with Extended Context

Jörg Tiedemann and Yves Scherrer

- Translating Implicit Discourse Connectives Based on Cross-lingual Annotation and Alignment
Hongzheng Li, Philippe Langlais, and Yaohong Jin

12:30–14:00 **Lunch Break**

14:00–15:30 **Session 3**

- 14:00–14:30 Neural Machine Translation with Extended Context
Jörg Tiedemann and Yves Scherrer

- 14:30–14:50 Discovery of Discourse-Related Language Contrasts through Alignment Discrepancies in English-German Translation
Ekaterina Lapshinova-Koltunski and Christian Hardmeier

- 14:50–15:10 Translating Implicit Discourse Connectives Based on Cross-lingual Annotation and Alignment
Hongzheng Li, Philippe Langlais, and Yaohong Jin

- 15:10–15:30 Combining the output of two coreference resolution systems for two source languages to improve annotation projection
Yulia Grishina

15:30–16:00 **Coffee Break**

16:00–17:30 **Session 4**

- 16:00–16:30 Lexical Chains meet Word Embeddings in Document-level Statistical Machine Translation
Laura Mascarell

- 16:30–16:50 On Integrating Discourse in Machine Translation
Karin Sim Smith

16:50–17:30 **Final Discussion and Conclusion**

Workshop 9: Workshop on Stylistic Variation

Organizers: Julian Brooke, Thamar Solorio, and Moshe Koppel

Venue: TODO

Friday September 8th, 2017

9:00–9:10 **Opening remarks (Julian Brooke, Thamar Solorio, and Moshe Koppel)**

9:10–10:00 **Invited Talk: Style Analysis for Practical Semantic Interpretation of Text (Ani Nenkova)**

10:00–10:30 From Shakespeare to Twitter: What are Language Styles all about?
Wei Xu

10:30–11:00 **Coffee Break**

11:00–12:30 **Technical Papers I**

11:00–11:30 Shakespearizing Modern Language Using Copy-Enriched Sequence to Sequence Models
Harsh Jhamtani, Varun Gangal, Eduard Hovy, and Eric Nyberg

11:30–12:00 Discovering Stylistic Variations in Distributional Vector Space Models via Lexical Paraphrases
Xing Niu and Marine Carpuat

12:00–12:30 Harvesting Creative Templates for Generating Stylistically Varied Restaurant Reviews
Shereen Oraby, Sheideh Homayon, and Marilyn Walker

12:30–14:00 **Lunch**

14:00–14:50 **Invited Talk: Problems in Personality Profiling (Walter Daelemans)**

14:50–15:30 **Poster Session**

- Is writing style predictive of scientific fraud?
Chloé Braud and Anders Søgaard
- "Deep" Learning : Detecting Metaphoricity in Adjective-Noun Pairs
Yuri Bizzoni, Stergios Chatzikyriakidis, and Mehdi Ghanimifard
- Authorship Attribution with Convolutional Neural Networks and POS-Eliding
Julian Hitschler, Esther van den Berg, and Ines Rehbein
- Topic and audience effects on distinctively Scottish vocabulary usage in Twitter data
Philippa Shoemark, James Kirby, and Sharon Goldwater
- Differences in type-token ratio and part-of-speech frequencies in male and female Russian written texts
Tatiana Litvinova, Pavel Seredin, Olga Litvinova, and Olga Zagorovskaya
- Modeling Communicative Purpose with Functional Style: Corpus and Features for German Genre and Register Analysis
Thomas Haider and Alexis Palmer
- Stylistic Variation in Television Dialogue for Natural Language Generation
Grace Lin and Marilyn Walker

One-day Workshops

15:30–16:00 **Coffee Break**

16:00–17:30 **Technical Papers II**

16:00–16:30 Controlling Linguistic Style Aspects in Neural Language Generation
Jessica Ficler and Yoav Goldberg

16:30–17:00 Approximating Style by N-gram-based Annotation
Melanie Andresen and Heike Zinsmeister

17:00–17:30 Assessing the Stylistic Properties of Neurally Generated Text in
Authorship Attribution
*Enrique Manjavacas, Jeroen De Gussem, Walter Daelemans, and
Mike Kestemont*

17:30–17:35 **Closing Remarks (Julian Brooke, Thamar Solorio, and Moshe Koppel)**

Workshop 10: 12th Workshop on Innovative Use of NLP for Building Educational Applications

Organizers: *Joel Tetreault, Jill Burstein, Claudia Leacock, and Helen Yannakoudakis*

Venue: TODO

Friday, September 8, 2017

08:45–09:00 Load Oral Presentations

09:00–10:30 Session 1

09:00–09:15 Opening Remarks

09:15–09:40 Question Difficulty – How to Estimate Without Norming, How to Use for Automated Grading
Ulrike Pado

09:40–10:05 Combining CNNs and Pattern Matching for Question Interpretation in a Virtual Patient Dialogue System
Lifeng Jin, Michael White, Evan Jaffe, Laura Zimmerman, and Douglas Danforth

10:05–10:30 Continuous fluency tracking and the challenges of varying text complexity
Beata Beigman Klebanov, Anastassia Loukina, John Sabatini, and Tenaha O'Reilly

10:30–11:00 Break

11:00–12:35 Session 2

11:00–11:25 Auxiliary Objectives for Neural Error Detection Models
Marek Rei and Helen Yannakoudakis

11:25–11:50 Linked Data for Language-Learning Applications
Robyn Loughnane, Kate McCurdy, Peter Kolb, and Stefan Sealent

11:50–12:10 Predicting Specificity in Classroom Discussion
Luca Lugini and Diane Litman

12:10–12:35 A Report on the 2017 Native Language Identification Shared Task
Sherzain Malmasi, Keelan Eganini, Aoife Cahill, Joel Tetreault, Robert Pugh, Christopher Hamill, Diane Napolitano, and Yao Qian

12:35–14:00 Lunch

14:00–15:30 Poster Session

14:00–14:45 Poster Session A

- Evaluation of Automatically Generated Pronoun Reference Questions
Arief Yudha Satria and Takenobu Tokunaga
 - Predicting Audience's Laughter During Presentations Using Convolutional Neural Network
Lei Chen and Chong Min Lee
 - Collecting fluency corrections for spoken learner English
Andrew Caines, Emma Flint, and Paula Buttery
-

- Exploring Relationships Between Writing & Broader Outcomes With Automated Writing Evaluation
Jill Burstein, Dan McCaffrey, Beata Beigman Klebanov, and Guangming Ling
- An Investigation into the Pedagogical Features of Documents
Emily Sheng, Prem Natarajan, Jonathan Gordon, and Gully Burns
- Combining Multiple Corpora for Readability Assessment for People with Cognitive Disabilities
Victoria Yaneva, Constantin Orasan, Richard Evans, and Omid Rohanian
- Automatic Extraction of High-Quality Example Sentences for Word Learning Using a Determinantal Point Process
Arseny Tolmachev and Sadao Kurohashi
- Distractor Generation for Chinese Fill-in-the-blank Items
Shu Jiang and John Lee
- An Error-Oriented Approach to Word Embedding Pre-Training
Youmna Farag, Marek Rei, and Ted Briscoe
- Investigating neural architectures for short answer scoring
Brian Riordan, Andrea Horbach, Aoife Cahill, Torsten Zesch, and Chong Min Lee
- Human and Automated CEFR-based Grading of Short Answers
Anaïs Tack, Thomas François, Sophie Roekhaut, and Cédric Fairon
- GEC into the future: Where are we going and how do we get there?
Keisuke Sakaguchi, Courtney Napolis, and Joel Tetreault
- Detecting Off-topic Responses to Visual Prompts
Marek Rei
- Combining Textual and Speech Features in the NLI Task Using State-of-the-Art Machine Learning Techniques
Pavel Ircing, Jan Svec, Zbynek Zajic, Barbora Hladka, and Martin Holub
- Native Language Identification Using a Mixture of Character and Word N-grams
Elham Mohammadi, Hadi Veisi, and Hessam Amini
- Ensemble Methods for Native Language Identification
Sophia Chan, Maryam Honari Jahromi, Benjamin Benetti, Aazim Lakhani, and Alona Fyshe
- Can string kernels pass the test of time in Native Language Identification?
Radu Tudor Ionescu and Marius Popescu
- Neural Networks and Spelling Features for Native Language Identification
Johannes Bjerva, Gintare Grigonyte, Robert Östling, and Barbara Plank
- A study of N-gram and Embedding Representations for Native Language Identification
Sowmya Vajjala and Sagnik Banerjee
- A Shallow Neural Network for Native Language Identification with Character N-grams
Yunita Sari, Muhammad Rifqi Fatchurrahman, and Meisyarah Dwiaستuti
- Fewer features perform well at Native Language Identification task
Taraka Rama and Çağrı Çöltekin

14:45–15:30 Poster Session B

- Structured Generation of Technical Reading Lists
Jonathan Gordon, Stephen Aguilar, Emily Sheng, and Gully Burns
- Effects of Lexical Properties on Viewing Time per Word in Autistic and Neurotypical Readers
Sanja Štajner, Victoria Yaneva, Ruslan Mitkov, and Simone Paolo Ponzetto

- Transparent text quality assessment with convolutional neural networks
Robert Östling and Gintare Grigonyte
- Artificial Error Generation with Machine Translation and Syntactic Patterns
Marek Rei, Mariano Felice, Zheng Yuan, and Ted Briscoe
- Modelling semantic acquisition in second language learning
Ekaterina Kochmar and Ekaterina Shutova
- Multiple Choice Question Generation Utilizing An Ontology
Katherine Stasaki and Marti A. Hearst
- Simplifying metaphorical language for young readers: A corpus study on news text
Magdalena Wolska and Yulia Clausen
- Language Based Mapping of Science Assessment Items to Skills
Farah Nadeem and Mari Ostendorf
- Connecting the Dots: Towards Human-Level Grammatical Error Correction
Shamil Chollampatt and Hwee Tou Ng
- Question Generation for Language Learning: From ensuring texts are read to supporting learning
Maria Chinkina and Detmar Meurers
- Systematically Adapting Machine Translation for Grammatical Error Correction
Courtney Napoles and Chris Callison-Burch
- Fine-grained essay scoring of a complex writing task for native speakers
Andrea Horbach, Dirk Scholten-Akoun, Yuning Ding, and Torsten Zesch
- Exploring Optimal Voting in Native Language Identification
Cyril Goutte and Serge Léger
- CIC-FBK Approach to Native Language Identification
Ilia Markov, Lingzhen Chen, Carlo Strapparava, and Grigori Sidorov
- The Power of Character N-grams in Native Language Identification
Artur Kulmizev, Bo Blankers, Johannes Bjerva, Malvina Nissim, Gertjan van Noord, Barbara Plank, and Martijn Wieling
- Classifier Stacking for Native Language Identification
Wen Li and Liang Zou
- Native Language Identification on Text and Speech
Marcos Zampieri, Alina Maria Ciobanu, and Liviu P. Dinu
- Native Language Identification using Phonetic Algorithms
Charese Smiley and Sandra Kübler
- A deep-learning based native-language classification by using a latent semantic analysis for the NLI Shared Task 2017
Yoo Rhee Oh, Hyung-Bae Jeon, Hwa Jeon Song, Yun-Kyung Lee, Jeon-Gue Park, and Yun-Keun Lee
- Fusion of Simple Models for Native Language Identification
Fabio Kepler, Ramón Astudillo, and Alberto Abad
- Stacked Sentence-Document Classifier Approach for Improving Native Language Identification
Andrea Cimino and Felice Dell'Orletta

15:30–16:00 **Break**

16:00–17:30 **Session 3**

16:00–16:25 Using Gaze to Predict Text Readability
Ana Valeria Gonzalez-Garduño and Anders Søgaard

One-day Workshops

16:25–16:50 Annotating Orthographic Target Hypotheses in a German L1 Learner Corpus
Ronja Laarmann-Quante, Katrin Ortmann, Anna Ehlert, Maurice Vogel, and Stefanie Dipper

16:50–17:15 A Large Scale Quantitative Exploration of Modeling Strategies for Content Scoring
Nitin Madnani, Anastassia Loukina, and Aoife Cahill

17:15–17:30 **Closing Remarks**

Workshop 11: 4th Workshop on Argument Mining

Organizers: *Ivan Habernal, Iryna Gurevych, Kevin Ashley, Claire Cardie, Nancy Green, Diane Litman, Georgios Petasis, Chris Reed, Noam Slonim, and Vern Walker*

Venue: TODO

Friday, September 8, 2017

8:50–9:50 Welcome session

8:50–9:00 Welcome (Workshop Chairs)

9:00–9:50 Invited talk (Christian Kock, Dept. of Media, Cognition and Communication, University of Copenhagen)

9:50–10:30 Paper session I

9:50–10:10 200K+ Crowdsourced Political Arguments for a New Chilean Constitution
Constanza Fierro, Claudio Fuentes, Jorge Pérez, and Mauricio Quezada

10:10–10:30 Analyzing the Semantic Types of Claims and Premises in an Online Persuasive Forum
Christopher Hidey, Elena Musi, Alyssa Hwang, Smaranda Muresan, and Kathy McKeown

10:30–11:00 Coffee break

11:00–12:30 Paper session II

11:00–11:20 Annotation of argument structure in Japanese legal documents
Hiroaki Yamada, Simone Teufel, and Takenobu Tokunaga

11:20–11:40 Improving Claim Stance Classification with Lexical Knowledge Expansion and Context Utilization
Roy Bar-Haim, Lilach Edelstein, Charles Jochim, and Noam Slonim

11:40–12:00 Mining Argumentative Structure from Natural Language text using Automatically Generated Premise-Conclusion Topic Models
John Lawrence and Chris Reed

12:00–12:20 Building an Argument Search Engine for the Web
Henning Wachsmuth, Martin Potthast, Khalid Al Khatib, Yamen Ajjour, Jana Puschmann, Jiani Qu, Jonas Dorsch, Viorel Morari, Janek Bevendorff, and Benno Stein

12:30–14:30 Lunch break

14:30–15:30 Poster session

14:30–15:30 Argument Relation Classification Using a Joint Inference Model
Yufang Hou and Charles Jochim

14:30–15:30 Projection of Argumentative Corpora from Source to Target Languages
Ahmet Aker and Huangpan Zhang

14:30–15:30 Manual Identification of Arguments with Implicit Conclusions Using Semantic Rules for Argument Mining
Nancy Green

14:30–15:30 Unsupervised corpus—wide claim detection
Ran Levy, Shai Gretz, Benjamin Sznajder, Shay Hummel, Ranit Aharonov, and Noam Slonim

One-day Workshops

14:30–15:30 Using Question-Answering Techniques to Implement a Knowledge-Driven Argument Mining Approach
Patrick Saint-Dizier

14:30–15:30 What works and what does not: Classifier and feature analysis for argument mining
Ahmet Aker, Alfred Sliwa, Yuan Ma, Ruishen Lui, Niravkumar Borad, Seyedeh Ziyaei, and Mina Ghobadi

15:30–16:00 **Coffee break**

16:00–17:00 **Paper session III**

16:00–16:20 Unsupervised Detection of Argumentative Units through Topic Modeling Techniques
Alfio Ferrara, Stefano Montanelli, and Georgios Petasis

16:20–16:40 Using Complex Argumentative Interactions to Reconstruct the Argumentative Structure of Large-Scale Debates
John Lawrence and Chris Reed

16:40–17:00 Unit Segmentation of Argumentative Texts
Yamen Ajjour, Wei-Fan Chen, Johannes Kiesel, Henning Wachsmuth, and Benno Stein

17:00–17:30 **Wrap-up discussion**

Workshop 12: 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis

Organizers: *Alexandra Balahur, Saif Mohammad, and Erik van der Goot*

Venue: TODO

Friday, 8th September 2017

08:30–08:40 Opening Remarks

08:40–10:30 Session 1: Irony, stance and negotiating interpersonal meaning

08:40–09:15 Detecting Sarcasm Using Different Forms Of Incongruity
Aditya Joshi

09:15–09:40 Assessing State-of-the-Art Sentiment Models on State-of-the-Art
Sentiment Datasets
Jeremy Barnes, Roman Klinger, and Sabine Schulte im Walde

09:40–10:05 Annotation, Modelling and Analysis of Fine-Grained Emotions on a
Stance and Sentiment Detection Corpus
*Hendrik Schuff, Jeremy Barnes, Julian Möhme, Sebastian Padó, and
Roman Klinger*

10:05–10:30 Ranking Right-Wing Extremist Social Media Profiles by Similarity to
Democratic and Extremist Groups
Matthias Hartung, Roman Klinger, Franziska Schmidtke, and Lars Vogel

10:30–11:00 Coffee Break

11:00–12:30 Session 2: Emotion Intensity Task

11:00–11:40 WASSA-2017 Shared Task on Emotion Intensity
Saif Mohammad and Felipe Bravo-Marquez

11:40–12:05 IMS at EmoInt-2017: Emotion Intensity Prediction with Affective Norms,
Automatically Extended Resources and Deep Learning
Maximilian Köper, Evgeny Kim, and Roman Klinger

12:05–12:30 Prayas at EmoInt 2017: An Ensemble of Deep Neural Architectures for
Emotion Intensity Prediction in Tweets
Prayas Jain, Pranav Goel, Devang Kulshreshtha, and Kaushal Kumar Shukla

12:30–14:00 Lunch Break

14:00–15:30 Session 3: Sentiment, stance and emotion

14:00–14:35 Latest News in Computational Argumentation: Surfing on the Deep
Learning Wave, Scuba Diving in the Abyss of Fundamental Questions
Iryna Gurevych

14:35–15:00 Towards Syntactic Iberian Polarity Classification
David Vilares, Marcos García, Miguel A. Alonso, and Carlos Gómez-Rodríguez

15:00–15:15 Toward Stance Classification Based on Claim Microstructures
Filip Boltuzic and Jan Šnajder

15:15–15:30 Linguistic Reflexes of Well-Being and Happiness in Echo
Jiaqi Wu, Marilyn Walker, Pranav Anand, and Steve Whittaker

15:30–16:00 Coffee Break

16:00–17:15 **Session 4: Preferences and values as determiners of sentiment and emotion**

16:00–16:35 Forecasting Consumer Spending from Purchase Intentions Expressed on Social Media
Viktor Pekar and Jane Binner

16:25–16:50 Mining fine-grained opinions on closed captions of YouTube videos with an attention-RNN
Edison Marrese-Taylor, Jorge Balazs, and Yutaka Matsuo

16:50–17:15 Understanding human values and their emotional effect
Alexandra Balahur

17:15–17:25 **Break**

17:25–18:25 **Session 5: Posters (Main Workshop and Emotion Intensity Task)**

- Did you ever read about Frogs drinking Coffee? Investigating the Compositionality of Multi-Emoji Expressions
Rebeca Padilla López and Fabienne Cap
- Investigating Redundancy in Emoji Use: Study on a Twitter Based Corpus
Giulia Donato and Patrizia Paggio
- Modeling Temporal Progression of Emotional Status in Mental Health Forum: A Recurrent Neural Net Approach
Kishaloy Halder, Lahari Poddar, and Min-Yen Kan
- Towards an integrated pipeline for aspect-based sentiment analysis in various domains
Orphee De Clercq, Els Lefever, Gilles Jacobs, Tijl Carrels, and Veronique Hoste
- Building a SentiWordNet for Odia
Gaurav Mohanty, Abishek Kannan, and Radhika Mamidi
- Lexicon Integrated CNN Models with Attention for Sentiment Analysis
Bonggun Shin, Timothy Lee, and Jinho D. Choi
- Explaining Recurrent Neural Network Predictions in Sentiment Analysis
Leila Arras, Grégoire Montavon, Klaus-Robert Müller, and Wojciech Samek
- GradAscent at EmoInt-2017: Character and Word Level Recurrent Neural Network Models for Tweet Emotion Intensity Detection
Egor Lakomkin, Chandrakant Bothe, and Stefan Wermter
- NUIG at EmoInt-2017: BiLSTM and SVR Ensemble to Detect Emotion Intensity
Vladimir Andryushchkin, Ian Wood, and James O' Neill
- Unsupervised Aspect Term Extraction with B-LSTM & CRF using Automatically Labelled Datasets
Athanasiou Giannakopoulos, Claudiu Musat, Andreea Hossmann, and Michael Baeriswyl
- PLN-PUCRS at EmoInt-2017: Psycholinguistic features for emotion intensity prediction in tweets
Henrique Santos and Renata Vieira
- Textmining at EmoInt-2017: A Deep Learning Approach to Sentiment Intensity Scoring of English Tweets
Hardik Meisher, Rupsa Saha, Priyanka Sinha, and Lipika Dey
- YNU-HPCC at EmoInt-2017: Using a CNN-LSTM Model for Sentiment Intensity Prediction
You Zhang, Hang Yuan, Jin Wang, and Xuejie Zhang
- Seernet at EmoInt-2017: Tweet Emotion Intensity Estimator
Venkatesh Duppada and Sushant Hiray

- IITP at EmoInt-2017: Measuring Intensity of Emotions using Sentence Embeddings and Optimized Features
Md Shad Akhtar, Palaash Sawant, Asif Ekbal, Jyoti Pawar, and Pushpak Bhattacharyya
- NSEmo at EmoInt-2017: An Ensemble to Predict Emotion Intensity in Tweets
Sreekanth Madisetty and Maunendra Sankar Desarkar
- Tecnolengua Lingmotif at EmoInt-2017: A lexicon-based approach
Antonio Moreno-Ortiz
- EmoAtt at EmoInt-2017: Inner attention sentence embedding for Emotion Intensity
Edison Marrese-Taylor and Yutaka Matsuo
- YZU-NLP at EmoInt-2017: Determining Emotion Intensity Using a Bi-directional LSTM-CNN Model
Yuanye He, Liang-Chih Yu, K. Robert Lai, and Weiyi Liu
- DMGroup at EmoInt-2017: Emotion Intensity Using Ensemble Method
Song Jiang and Xiaotian Han
- UWat-Emote at EmoInt-2017: Emotion Intensity Detection using Affect Clues, Sentiment Polarity and Word Embeddings
Vineet John and Olga Vechtomova
- LIPN-UAM at EmoInt-2017: Combination of Lexicon-based features and Sentence-level Vector Representations for Emotion Intensity Determination
Davide Buscaldi and Belem Priego
- deepCybErNet at EmoInt-2017: Deep Emotion Intensities in Tweets
Vinayakumar R and Prabaharan Poornachandran

18:25–18:30 **Closing remarks**

Workshop 13: 2nd Workshop on Evaluating Vector Space Representations for NLP

Organizers: *Samuel Bowman, Yoav Goldberg, Felix Hill, Angeliki Lazaridou, Omer Levy, Roi Reichart, and Anders Søgaard*

Venue: TODO

Friday 8 September 2017

09:00 Opening Remarks

09:20–09:55 Shared task report

- The RepEval 2017 Shared Task: Multi-Genre Natural Language Inference with Sentence Representations
Nikita Nangia, Adina Williams, Angeliki Lazaridou, and Samuel Bowman

09:55–10:30 Yejin Choi (University of Washington)

10:30–11:00 Coffee Break (set up posters)

11:00–11:35 Jakob Uszkoreit (Google Research)

11:35–12:10 Kyunghyun Cho (New York University)

12:10–12:30 Few Minutes Madness (Evaluation Proposals)

- Traversal-Free Word Vector Evaluation in Analogy Space
Xiaoyin Che, Nico Ring, Willi Raschkowski, Haojin Yang, and Christoph Meinel
- Hypothesis Testing based Intrinsic Evaluation of Word Embeddings
Nishant Gurnani
- Evaluation of word embeddings against cognitive processes: primed reaction times in lexical decision and naming tasks
Jeremy Auguste, Arnaud Rey, and Benoit Favre
- Playing with Embeddings : Evaluating embeddings for Robot Language Learning through MUD Games
Anmol Gulati and Kumar Krishna Agrawal
- Recognizing Textual Entailment in Twitter Using Word Embeddings
Octavia-Maria řulea

12:30–14:00 Lunch (somewhere together if pos)

14:00–14:30 Contributed Talks (shared task systems)

14:00–14:15 Recurrent Neural Network-Based Sentence Encoder with Gated Attention for Natural Language Inference *Qian Chen, Xiaodan Zhu, Zhen-Hua Ling, Si Wei, Hui Jiang, and Diana Inkpen*

14:15–14:30 Shortcut-Stacked Sentence Encoders for Multi-Domain Inference *Yixin Nie and Mohit Bansal*

14:30–15:30 Posters and discussion

- Character-level Intra Attention Network for Natural Language Inference
Han Yang, Marta R. Costa-jussà, and José A. R. Fonollosa
-

- Refining Raw Sentence Representations for Textual Entailment Recognition via Attention
Jorge Balazs, Edison Marrese-Taylor, Pablo Loyola, and Yutaka Matsuo
- LCT-MALTA's Submission to RepEval 2017 Shared Task
Hoa Vu

15:30–16:00 **Working Coffee Break**

16:00–17:30 **Presentation of Findings and Panel Discussion**

Workshop 14: Building Linguistically Generalizable NLP Systems

Organizers: *Emily M. Bender, Hal Daumé III, Allyson Ettinger, and Sudha Rao*

Venue: TODO

Friday, September 8, 2017

09:00–09:15 **Welcome Note**

- Towards Linguistically Generalizable NLP Systems: A Workshop and Shared Task
Allyson Ettinger, Sudha Rao, Hal Daumé III, and Emily M. Bender

09:15–10:00 **Invited Talk (Aurelie Herbelot)**

10:00–12:10 **Session 1: Research Contribution Papers**

10:00–10:25 Analysing Errors of Open Information Extraction Systems
Rudolf Schneider, Tom Oberhauser, Tobias Klatt, Felix A. Gers, and Alexander Löser

10:30–11:00 **Coffee Break**

11:00–11:45 **Invited Talk (Grzegorz Chrupała)**

11:45–12:10 Massively Multilingual Neural Grapheme-to-Phoneme Conversion
Ben Peters, Jon Dehdari, and Josef van Genabith

12:10–12:30 "Build It Break It, Language Edition" Shared Task Overview

12:30–14:00 **Lunch Break**

14:00–14:45 **Invited Talk (Martha Palmer)**

14:45–15:35 **Session 2: Shared Task Description Papers**

14:45–15:10 BIBI System Description: Building with CNNs and Breaking with Deep Reinforcement Learning
Yitong Li, Trevor Cohn, and Timothy Baldwin

15:10–15:35 Breaking NLP: Using Morphosyntax, Semantics, Pragmatics and World Knowledge to Fool Sentiment Analysis Systems
Taylor Mahler, Willy Cheung, Micha Elsner, David King, Marie-Catherine de Marneffe, Cory Shaing, Symon Stevens-Guille, and Michael White

15:35–16:00 **Coffee Break**

16:00–17:15 **Poster Session**

- An Adaptable Lexical Simplification Architecture for Major Ibero-Romance Languages
Daniel Ferrés, Horacio Saggion, and Xavier Gómez Guinovart
 - Cross-genre Document Retrieval: Matching between Conversational and Formal Writings
Tomasz Jurczyk and Jinho D. Choi
 - ACTSA: Annotated Corpus for Telugu Sentiment Analysis
Sandeep Sricharan Mukku and Radhika Mamidi
-

- Strawman: An Ensemble of Deep Bag-of-Ngrams for Sentiment Analysis
Kyunghyun Cho
- Breaking Sentiment Analysis of Movie Reviews
Ieva Staliūnaitė and Ben Bonfil

17:15–17:30 **Closing Remarks**

Main Conference: Saturday, September 9

Demo Session

Time: 10:30–12:10

Location: Foyer

The NLTK FrameNet API: Designing for Discoverability with a Rich Linguistic Resource

Nathan Schneider and Chuck Wooters

A new Python API, integrated within the NLTK suite, offers access to the FrameNet 1.7 lexical database. The lexicon (structured in terms of frames) as well as annotated sentences can be processed programmatically, or browsed with human-readable displays via the interactive Python prompt.

Argotario: Computational Argumentation Meets Serious Games

Ivan Habernal, Raffael Hannemann, Christian Pollak, Christopher Klamim, Patrick Pauli, and Iryna Gurevych

An important skill in critical thinking and argumentation is the ability to spot and recognize fallacies. Fallacious arguments, omnipresent in argumentative discourse, can be deceptive, manipulative, or simply leading to ‘wrong moves’ in a discussion. Despite their importance, argumentation scholars and NLP researchers with focus on argumentation quality have not yet investigated fallacies empirically. The nonexistence of resources dealing with fallacious argumentation calls for scalable approaches to data acquisition and annotation, for which the serious games methodology offers an appealing, yet unexplored, alternative. We present Argotario, a serious game that deals with fallacies in everyday argumentation. Argotario is a multilingual, open-source, platform-independent application with strong educational aspects, accessible at www.argotario.net.

An Analysis and Visualization Tool for Case Study Learning of Linguistic Concepts

Cecilia Ovesdotter Alm, Benjamin Meyers, and Emily Prud'hommeaux

We present an educational tool that integrates computational linguistics resources for use in non-technical undergraduate language science courses. By using the tool in conjunction with evidence-driven pedagogical case studies, we strive to provide opportunities for students to gain an understanding of linguistic concepts and analysis through the lens of realistic problems in feasible ways. Case studies tend to be used in legal, business, and health education contexts, but less in the teaching and learning of linguistics. The approach introduced also has potential to encourage students across training backgrounds to continue on to computational language analysis coursework.

GraphDocExplore: A Framework for the Experimental Comparison of Graph-based Document Exploration Techniques

Tobias Falke and Iryna Gurevych

Graphs have long been proposed as a tool to browse and navigate in a collection of documents in or-

der to support exploratory search. Many techniques to automatically extract different types of graphs, showing for example entities or concepts and different relationships between them, have been suggested. While experimental evidence that they are indeed helpful exists for some of them, it is largely unknown which type of graph is most helpful for a specific exploratory task. However, carrying out experimental comparisons with human subjects is challenging and time-consuming. Towards this end, we present the *GraphDocExplore* framework. It provides an intuitive web interface for graph-based document exploration that is optimized for experimental user studies. Through a generic graph interface, different methods to extract graphs from text can be plugged into the system. Hence, they can be compared at minimal implementation effort in an environment that ensures controlled comparisons. The system is publicly available under an open-source license.

SGNMT – A Flexible NMT Decoding Platform for Quick Prototyping of New Models and Search Strategies

Felix Stahlberg, Eva Hasler, Danielle Saunders, and Bill Byrne

This paper introduces SGNMT, our experimental platform for machine translation research. SGNMT provides a generic interface to neural and symbolic scoring modules (predictors) with left-to-right semantic such as translation models like NMT, language models, translation lattices, n-best lists or other kinds of scores and constraints. Predictors can be combined with other predictors to form complex decoding tasks. SGNMT implements a number of search strategies for traversing the space spanned by the predictors which are appropriate for different predictor constellations. Adding new predictors or decoding strategies is particularly easy, making it a very efficient tool for prototyping new research ideas. SGNMT is actively being used by students in the MPhil program in Machine Learning, Speech and Language Technology at the University of Cambridge for course work and theses, as well as for most of the research work in our group.

StruAP: A Tool for Bundling Linguistic Trees through Structure-based Abstract Pattern

Kohsuke Yanai, Misa Sato, Toshihiko Yanase, Kenzo Kurotsuchi, Yuta Koreeda, and Yoshiaki Niwa

We present a tool for developing tree structure patterns that makes it easy to define the relations among textual phrases and create a search index for these newly defined relations. By using the proposed tool, users develop tree structure patterns through abstracting syntax trees. The tool features (1) intuitive pattern syntax, (2) unique functions such as recursive call of patterns and use of lexicon dictionaries, and (3) whole workflow support for relation development and validation. We report the current implementation of the tool and its effectiveness.

KnowYourNyms? A Game of Semantic Relationships

Ross Mechanic, Dean Fulgoni, Hannah Cutler, Sneha Rajana, Zheyuan Liu, Bradley Jackson, Anne Cocos, Chris Callison-Burch, and Marianna Apidianaki

Semantic relation knowledge is crucial for natural language understanding. We introduce “KnowYourNyms?”, a web-based game for learning semantic relations. While providing users with an engaging experience, the application collects large amounts of data that can be used to improve semantic relation classifiers. The data also broadly informs us of how people perceive the relationships between words, providing useful insights for research in psychology and linguistics.

The Projector: An Interactive Annotation Projection Visualization Tool

Alan Akbik and Roland Vollgraf

Previous works proposed annotation projection in parallel corpora to inexpensively generate treebanks or propbanks for new languages. In this approach, linguistic annotation is automatically transferred from a resource-rich source language (SL) to translations in a target language (TL). However, annotation projection may be adversely affected by translational divergences between specific language pairs. For this reason, previous work often required careful qualitative analysis of projectability of specific annotation in order to define strategies to address quality and coverage issues. In this demonstration, we present THE PROJECTOR, an interactive GUI designed to assist researchers in such analysis: it allows users to execute and visually inspect annotation projection in a range of different settings. We give an overview of the GUI, discuss use cases and illustrate how the tool can facilitate discussions with the research community.

Interactive Visualization for Linguistic Structure

Aaron Sarnat, Vidur Joshi, Cristian Petrescu-Prahova, Alvaro Herrasti, Brandon Stilson, and Mark Hopkins

We provide a visualization library and web interface for interactively exploring a parse tree or a forest of parses. The library is not tied to any particular linguistic representation, but provides a general-purpose API for the interactive exploration of hierarchical linguistic structure. To facilitate rapid understanding of a complex structure, the API offers several important features, including expand/collapse functionality, positional and color cues, explicit visual support for sequential structure, and dynamic highlighting to convey node-to-text correspondence.

DLATK: Differential Language Analysis ToolKit

H. Andrew Schwartz, Salvatore Giorgi, Maarten Sap, Patrick Crutchley, Lyle Ungar, and Johannes Eichstaedt

We present Differential Language Analysis Toolkit (DLATK), an open-source python package and command-line tool developed for conducting social-scientific language analyses. While DLATK provides standard NLP pipeline steps such as tokenization or SVM-classification, its novel strengths lie in analyses useful for psychological, health, and social science: (1) incorporation of extra-linguistic structured information, (2) specified levels and units of analysis (e.g. document, user, community), (3) statistical metrics for continuous outcomes, and (4) robust, proven, and accurate pipelines for social-scientific prediction problems. DLATK integrates multiple popular packages (SKLearn, Mallet), enables interactive usage (Jupyter Notebooks), and generally follows object oriented principles to make it easy to tie in additional libraries or storage technologies.

QUINT: Interpretable Question Answering over Knowledge Bases

Abdalghani Abujabal, Rishiraj Saha Roy, Mohamed Yahya, and Gerhard Weikum

We present QUINT, a live system for question answering over knowledge bases. QUINT automatically learns role-aligned utterance-query templates from user questions paired with their answers. When QUINT answers a question, it visualizes the complete derivation sequence from the natural language utterance to the final answer. The derivation provides an explanation of how the syntactic structure of the question was used to derive the structure of a SPARQL query, and how the phrases in the question were used to instantiate different parts of the query. When an answer seems unsatisfactory, the derivation provides valuable insights towards reformulating the question.

Function Assistant: A Tool for NL Querying of APIs

Kyle Richardson and Jonas Kuhn

In this paper, we describe Function Assistant, a lightweight Python-based toolkit for querying and exploring source code repositories using natural language. The toolkit is designed to help end-users of a target API quickly find information about functions through high-level natural language queries, or descriptions. For a given text query and background API, the tool finds candidate functions by performing a translation from the text to known representations in the API using the semantic parsing approach of (Richardson and Kuhn, 2017). Translations are automatically learned from example text-code pairs in example APIs. The toolkit includes features for building translation pipelines and query engines for arbitrary source code projects. To explore this last feature, we perform new experiments on 27 well-known Python projects hosted on Github.

MoodSwipe: A Soft Keyboard that Suggests MessageBased on User-Specified Emotions

Chieh-Yang Huang, Tristan Labetoulle, Ting-Hao Huang, Yi-Pei Chen, Hung-Chen Chen, Vallari Srivastava, and Lun-Wei Ku

We present MoodSwipe, a soft keyboard that suggests text messages given the user-specified emotions utilizing the real dialog data. The aim of MoodSwipe is to create a convenient user interface to enjoy the technology of emotion classification and text suggestion, and at the same time to collect labeled data automatically for developing more advanced technologies. While users select the MoodSwipe keyboard, they can type as usual but sense the emotion conveyed by their text and receive suggestions for their message as a benefit. In MoodSwipe, the detected emotions serve as the medium for suggested texts, where viewing the latter is the incentive to correcting the former. We conduct several experiments to show the superiority of the emotion classification models trained on the dialog data, and further to verify good emotion cues are important context for text suggestion.

ParlAI: A Dialog Research Software Platform

Alexander Miller, Will Feng, Dhruv Batra, Antoine Bordes, Adam Fisch, Jiasen Lu, Devi Parikh, and Jason Weston

We introduce ParlAI (pronounced “par-lay”), an open-source software platform for dialog research im-

plemented in Python, available at <http://parl.ai>. Its goal is to provide a unified framework for sharing, training and testing dialog models; integration of Amazon Mechanical Turk for data collection, human evaluation, and online/reinforcement learning; and a repository of machine learning models for comparing with others' models, and improving upon existing architectures. Over 20 tasks are supported in the first release, including popular datasets such as SQuAD, bAbI tasks, MCTest, WikiQA, QACNN, QADailyMail, CBT, bAbI Dialog, Ubuntu, OpenSubtitles and VQA. Several models are integrated, including neural models such as memory networks, seq2seq and attentive LSTMs.

HeidelPlace: An Extensible Framework for Geoparsing

Ludwig Richter, Johanna Geiß, Andreas Spitz, and Michael Gertz

Geographic information extraction from textual data sources, called geoparsing, is a key task in text processing and central to subsequent spatial analysis approaches. Several geoparsers are available that support this task, each with its own (often limited or specialized) gazetteer and its own approaches to toponym detection and resolution. In this demonstration paper, we present HeidelPlace, an extensible framework in support of geoparsing. Key features of HeidelPlace include a generic gazetteer model that supports the integration of place information from different knowledge bases, and a pipeline approach that enables an effective combination of diverse modules tailored to specific geoparsing tasks. This makes HeidelPlace a valuable tool for testing and evaluating different gazetteer sources and geoparsing methods. In the demonstration, we show how to set up a geoparsing workflow with HeidelPlace and how it can be used to compare and consolidate the output of different geoparsing approaches.

Unsupervised, Knowledge-Free, and Interpretable Word Sense Disambiguation

Alexander Panchenko, Fide Marten, Eugen Ruppert, Stefano Faralli, Dmitry Ustalov, Simone Paolo Ponzetto, and Chris Biemann

Interpretability of a predictive model is a powerful feature that gains the trust of users in the correctness of the predictions. In word sense disambiguation (WSD), knowledge-based systems tend to be much more interpretable than knowledge-free counterparts as they rely on the wealth of manually-encoded elements representing word senses, such as hypernyms, usage examples, and images. We present a WSD system that bridges the gap between these two so far disconnected groups of methods. Namely, our system, providing access to several state-of-the-art WSD models, aims to be interpretable as a knowledge-based system while it remains completely unsupervised and knowledge-free. The presented tool features a Web interface for all-word disambiguation of texts that makes the sense predictions human readable by providing interpretable word sense inventories, sense representations, and disambiguation results. We provide a public API, enabling seamless integration.

NeuroNER: an easy-to-use program for named-entity recognition based on neural networks

Franck Dernoncourt, Ji Young Lee, and Peter Szolovits

Named-entity recognition (NER) aims at identifying entities of interest in a text. Artificial neural networks (ANNs) have recently been shown to outperform existing NER systems. However, ANNs remain challenging to use for non-expert users. In this paper, we present NeuroNER, an easy-to-use named-entity recognition tool based on ANNs. Users can annotate entities using a graphical web-based user interface (BRAT): the annotations are then used to train an ANN, which in turn predict entities' locations and categories in new texts. NeuroNER makes this annotation-training-prediction flow smooth and accessible to anyone.

SupWSD: A Flexible Toolkit for Supervised Word Sense Disambiguation

Simone Papandrea, Alessandro Raganato, and Claudio Delli Bovi

In this demonstration we present SupWSD, a Java API for supervised Word Sense Disambiguation (WSD). This toolkit includes the implementation of a state-of-the-art supervised WSD system, together with a Natural Language Processing pipeline for preprocessing and feature extraction. Our aim is to provide an easy-to-use tool for the research community, designed to be modular, fast and scalable for training and testing on large datasets. The source code of SupWSD is available at <http://github.com/SI3P/SupWSD>.

Interactive Abstractive Summarization for Event News Tweets

Ori Shapira, Hadar Ronen, Meni Adler, Yael Amsterdamer, Judit Bar-Ilan, and Ido Dagan

We present a novel interactive summarization system that is based on abstractive summarization, derived from a recent consolidated knowledge representation for multiple texts. We incorporate a couple of inter-

action mechanisms, providing a bullet-style summary while allowing to attain the most important information first and interactively drill down to more specific details. A usability study of our implementation, for event news tweets, suggests the utility of our approach for text exploration.

LangPro: Natural Language Theorem Prover*Lasha Abzianidze*

LangPro is an automated theorem prover for natural language. Given a set of premises and a hypothesis, it is able to prove semantic relations between them. The prover is based on a version of analytic tableau method specially designed for natural logic. The proof procedure operates on logical forms that preserve linguistic expressions to a large extent. %This property makes the logical forms easily obtainable from syntactic trees. %, in particular, Combinatory Categorial Grammar derivation trees. The nature of proofs is deductive and transparent. On the FraCaS and SICK textual entailment datasets, the prover achieves high results comparable to state-of-the-art.

Interactive Visualization and Manipulation of Attention-based Neural Machine Translation*Jaesong Lee, Joong-Hwi Shin, and Jun-Seok Kim*

While neural machine translation (NMT) provides high-quality translation, it is still hard to interpret and analyze its behavior. We present an interactive interface for visualizing and intervening behavior of NMT, specifically concentrating on the behavior of beam search mechanism and attention component. The tool (1) visualizes search tree and attention and (2) provides interface to adjust search tree and attention weight (manually or automatically) at real-time. We show the tool gives various methods to understand NMT.

Invited Talk: Dan Jurafsky

“Does This Vehicle Belong to You”? Processing the Language of Policing for Improving Police-Community Relations

Saturday, September 9, 2017, 9:00–10:10am

Jutland

Abstract: Police body-cameras have the potential to play an important role in understanding and improving police-community relations. In this talk I describe a series of studies conducted by our large interdisciplinary team at Stanford that use speech and natural language processing on body-camera recordings to model the interactions between police officers and community members in traffic stops. We use text and speech features to automatically measure linguistic aspects of the interaction, from discourse factors like conversational structure to social factors like respect. I describe the differences we find in the language directed toward black versus white community members, and offer suggestions for how these findings can be used to help improve the fraught relations between police officers and the communities they serve.

Biography: Dan Jurafsky is Professor and Chair of Linguistics and Professor of Computer Science, at Stanford University. His research has focused on the extraction of meaning, intention, and affect from text and speech, on the processing of Chinese, and on applying natural language processing to the cognitive and social sciences. Dan's deep interest in NLP education led him to co-write with Jim Martin the widely-used textbook "Speech and Language

Demo Session

Processing" (whose 3rd edition is in (slow) progress) and co-teach with Chris Manning the first massive open online class on natural language processing. Dan was the recipient of the 2002 MacArthur Fellowship and is a 2015 James Beard Award Nominee for his book, "The Language of Food: A Linguist Reads the Menu".

6

Main Conference: Sunday, September 10

Invited Talk: Sharon Goldwater

Towards more universal language technology: unsupervised learning from speech

Sunday, September 10, 2017, 9:00–10:10am

Jutland

Abstract: Speech and language processing has advanced enormously in the last decade, with successful applications in machine translation, voice-activated search, and even language-enabled personal assistants. Yet these systems typically still rely on learning from very large quantities of human-annotated data. These resource-intensive methods mean that effective technology is available for only a tiny fraction of the world's 7000 or so languages, mainly those spoken in large rich countries.

This talk describes our recent work on developing *unsupervised* speech technology, where transcripts and pronunciation dictionaries are not used. The work is inspired by considering both how young infants may begin to acquire the sounds and words of their language, and how we might develop systems to help linguists analyze and document endangered languages. I will first present work on learning from speech audio alone, where the system must learn to segment the speech stream into word tokens and cluster repeated instances of the same word together to learn a lexicon of vocabulary items. The approach combines Bayesian and neural network methods to address learning at the word and sub-word levels.

Biography: Sharon Goldwater is a Reader at the University of Edinburgh's School of Informatics, where she is a member of the Institute for Language, Cognition and Computation. She received her PhD in 2007 from Brown University and spent two years as a postdoctoral researcher at Stanford University before moving to Edinburgh. Her research interests include unsupervised learning for speech and language processing, computer modelling of language acquisition in children, and computational studies of language use. Dr. Goldwater co-chaired the 2014 Conference of the European Chapter of the Association for Computational Linguistics and is Chair-Elect of EACL. She has served on the editorial boards of the Transactions of the Association for Computational Linguistics, the Computational Linguistics journal, and OPEN MIND: Advances in Cognitive Science (a new open-access journal). In 2016, she received the Roger Needham Award from the British Computer Society, awarded for "distinguished research contribution in computer science by a UK-based researcher who has completed up to 10 years of post-doctoral research."

Social Event

Sunday, September 10, 2017, 18:00–22:00

Øksnehallen Courtyard

The EMNLP 2017 social event will take place in the court yard just outside the venue, directly after the last session of Sunday. Some social events are designed to give attendees a chance to see a popular tourist attraction, but this year, the social event is pure fun and relaxation. All attendees are invited to grab free food at our food trucks, listen to great music, or just chill in the lounge areas. Of course the social event will include surprises, also, but - hey, they're surprises.

7

Main Conference: Monday, September 11

Invited Talk: Nando de Freitas

Physical simulation, learning and language

Monday, September 11, 2017, 9:00–10:10am

Jutland

Abstract: Simulated physical environments, with common physical laws, objects and agents with bodies, provide us with consistency to facilitate transfer and continual learning. In such environments, research topics such as learning to experiment, learning to learn and emergent communication can be easily explored. Given the relevance of these topics to language, it is natural to ask ourselves whether research in language would benefit from the development of such environments, and whether language can contribute toward improving such environments and agents within them. This talk will provide an overview of some of these environments, discuss learning to learn and its potential relevance to language, and present some deep reinforcement learning agents that capitalize on formal language instructions to develop disentangled interpretable representations that allow them to generalize to a wide variety of zero-shot semantic tasks. The talk will pose more questions than answers in the hope of stimulating discussion.

Biography: I was born in Zimbabwe, with malaria. I was a refugee from the war in Mozambique and thanks to my parents getting in debt to buy me a passport from a corrupt official, I grew up in Portugal without water and electricity, before the EU got there, and without my parents who were busy making money to pay their debt. At 8, I joined my parents in Venezuela and began school in the hood; see City of God. I moved to South Africa after high-school and sold beer illegally in black-townships for a living until 1991. Apartheid was the worst thing I ever experienced. I did my BSc in electrical engineering and MSc in control at the University of the Witwatersrand, where I strived to be the best student to prove to racists that anyone can do it. I did my PhD on Bayesian methods for neural networks at Trinity College, Cambridge University. I did a postdoc in Artificial Intelligence at UC Berkeley. I became a Full Professor at the University of British Columbia, before joining the University of Oxford in 2013. I quit Oxford in 2017 to join DeepMind full-time, where I lead the Machine Learning team. I aim to solve intelligence so that future generations have a better life. I have been a Senior Fellow of the Canadian Institute for Advanced Research for a long time. Some of my recent awards, mostly thanks to my collaborators, include: Best Paper Award at the International Conference on Machine Learning (2016), Best Paper Award at the International Conference on Learning Representations (2016), Winner of round 5 of the Yelp Dataset Challenge (2015), Distinguished Paper Award at the International Joint Conference on Artificial Intelligence (2013), Charles A. McDowell Award for Excellence in Research (2012), and Mathematics of Information Technology and Complex Systems Young Researcher Award (2010).

SIGDAT Business Meeting

Date: Monday, September 11, 2017

Time: 12:40–13:40 PM

Venue: Jutland

Chair: Pascale Fung

All attendees are encouraged to participate in the business meeting.

8

Anti-harassment policy

The open exchange of ideas, the freedom of thought and expression, and respectful scientific debate are central to the aims and goals of the ACL. These require a community and an environment that recognizes the inherent worth of every person and group, that fosters dignity, understanding, and mutual respect, and that embraces diversity. For these reasons, ACL is dedicated to providing a harassment-free experience for all the members, as well as participants at our events and in our programs.

Harassment and hostile behavior are unwelcome at any ACL conference, associated event, or in ACL-affiliated on-line discussions. This includes: speech or behavior that intimidates, creates discomfort, or interferes with a person's participation or opportunity for participation in a conference or an event. We aim for ACL-related activities to be an environment where harassment in any form does not happen, including but not limited to: harassment based on race, gender, religion, age, color, appearance, national origin, ancestry, disability, sexual orientation, or gender identity. Harassment includes degrading verbal comments, deliberate intimidation, stalking, harassing photography or recording, inappropriate physical contact, and unwelcome sexual attention. The policy is not intended to inhibit challenging scientific debate, but rather to promote it through ensuring that all are welcome to participate in shared spirit of scientific inquiry.

It is the responsibility of the community as a whole to promote an inclusive and positive environment for our scholarly activities. In addition, anyone who experiences harassment or hostile behavior may contact any current member of the ACL Executive Committee or contact Priscilla Rasmussen, who is usually available at the registration desk during ACL conferences. Members of the executive committee will be instructed to keep any such contact in strict confidence, and those who approach the committee will be consulted before any actions are taken.

Approved by ACL Executive Committee in 2016.

The policy is also available from ACL's main page.

Index

- Abad, Alberto, 49
Abdou, Mostafa, 30
Abujabal, Abdalghani, 63
Abzianidze, Lasha, 65
Adler, Meni, 64
Agrawal, Kumar Krishna, 56
Aguilar, Gustavo, 37
Aguilar, Stephen, 48
Aharonov, Ranit, 51
Ahmed, Mir, 28
Ajjour, Yamen, 51, 52
Akbik, Alan, 62
Aker, Ahmet, 51, 52
Akhtar, Md Shad, 55
Al Khatib, Khalid, 51
Alkhouli, Tamer, 27, 31
Allan, James, 37
Allauzen, Alexandre, 26, 32
Alonso, Miguel A., 53
Alumäe, Tanel, 39
Alva, Carlo, 33
Amini, Hessam, 48
Amrhein, Chantal, 27
Amsterdamer, Yael, 64
Anand, Pranav, 53
Anastasopoulos, Antonios, 41
Anderson, Timothy, 27
Andresen, Melanie, 46
Androutsopoulos, Ion, 34
Andryushchkin, Vladimir, 54
Andy, Anietie, 36
Apidianaki, Marianna, 62
Aransa, Walid, 27, 28
Arras, Leila, 54
Ashley, Kevin, 51
Assylbekov, Zhenisbek, 33
Astudillo, Ramón, 49
Auguste, Jeremy, 56
Avramidis, Eleftherios, 29
Aziz, Wilker, 27
Baeriswyl, Michael, 54
Bahar, Parnia, 27
Bajoo, Shondell, 28
Bakagianni, Juli, 34
Balahur, Alexandra, 53, 54
Balazs, Jorge, 54, 57
Baldwin, Timothy, 33, 38, 58
Ballesteros, Miguel, 36
Banerjee, Sagnik, 48
Bangalore, Srinivas, 41
Bansal, Mohit, 39, 56
Bansal, Sameer, 41
Bar, Amir, 32
Bar-Haim, Roy, 51
Bar-Ilan, Judit, 64
Barbieri, Francesco, 36
Bardet, Adrien, 27, 28
Barnes, Jeremy, 53
Barrault, Loïc, 26–28
Barrett, Maria, 87
Barros, Cristina, 33
Bastings, Joost, 27
Batra, Dhruv, 63
Bayer, Ali Orkan, 39
Beigman Klebanov, Beata, 47, 48
Bender, Emily M., 58
Benetti, Benjamin, 48
Berant, Jonathan, 32
Berard, Alexandre, 30
Berg, Esther van den, 45
Bertoldi, Nicola, 26

- Besacier, Laurent, 30, 41
Beverendorff, Janek, 51
Bhargava, Preeti, 37
Bhattacharyya, Pushpak, 11, 13, 32, 55
Bicici, Ergun, 29
Biemann, Chris, 64
Bing, Lidong, 40
Bingel, Joachim, 2, 87
Birner, Jane, 54
Birch, Alexandra, 27, 31
Bizzoni, Yuri, 45
Bjerva, Johannes, 48, 49
Blair, Frédéric, 27, 29, 31
Blankers, Bo, 49
Blodgett, Su Lin, 36
Boecking, Benedikt, 36
Bogin, Ben, 32
Bois, Rémi, 34
Bojar, Ondrej, 26
Bojar, Ondřej, 26, 27, 29, 30
Boltuzic, Filip, 53
Bonfil, Ben, 59
Borad, Niravkumar, 52
Bordes, Antoine, 63
Born, Leo, 43
Bothe, Chandrakant, 54
Botschen, Teresa, 39
Bougares, Fethi, 26–28
Bourgonje, Peter, 34
Bowen, Fraser, 36
Bowman, Samuel, 56
Boyer, Arthur, 26
Bradbury, James, 38
Brantley, Kianté, 30
Braud, Chloé, 45
Braune, Fabienne, 27
Bravo-Marquez, Felipe, 53
Briscoe, Ted, 48, 49
Britz, Denny, 31
Brooke, Julian, 33, 45
Bunescu, Aurelia, 2
Buono, Maria Pia di, 34
Burlot, Franck, 26–29
Burns, Gully, 48
Burstein, Jill, 47, 48
Buscaldi, Davide, 55
Buttery, Paula, 37, 41, 47
Byrne, Bill, 62
- Çöltekin, Çağrı, 48
Şulea, Octavia-Maria, 56
Caglayan, Ozan, 27, 28
Cahill, Aoife, 41, 47, 48, 50
Cai, Zheng, 32
Caines, Andrew, 37, 41, 47
Cakici, Ruket, 37
Calixto, Iacer, 28
Callison-Burch, Chris, 36, 49, 62
- Cap, Fabienne, 43, 54
Cardie, Claire, 51
Carenini, Giuseppe, 39
Carlan, Chris, 28
Carpels, Tijl, 54
Carpuat, Marine, 45
Cartier, Emmanuel, 32
Casacuberta, Francisco, 31
Cettolo, Mauro, 43
Chan, Sophia, 48
Chandar, Sarath, 19, 20
Chang, Baobao, 38
Chang, Kai-Wei, 36
Chang, Ming-Wei, 36
Chatterjee, Rajen, 26, 30, 31
Chatzikyriakidis, Stergios, 45
Che, Xiaoyin, 56
Chen, Boxing, 27
Chen, Chaomei, 39
Chen, Hung-Chen, 63
Chen, Lei, 47
Chen, Lingzhen, 49
Chen, Mingjie, 34
Chen, Qian, 56
Chen, Wei-Fan, 52
Chen, Wei, 27
Chen, Yi-Pei, 63
Chen, Yidong, 27
Chen, Zhiming, 29, 30
Cheng, Shanbo, 27
Chernyak, Ekaterina, 32
Cherry, Colin, 27
Chesney, Sophie, 34
Cheung, Jackie Chi Kit, 39
Cheung, Willy, 58
Chiang, David, 41
Chinea-Rios, Mara, 31
Chinkina, Maria, 49
Cho, Eunah, 27, 31
Cho, Kyunghyun, 43, 59
Choi, Jinho D., 37, 54, 58
Choi, Sanghyuk, 32
Chollampatt, Shamil, 49
Chowdhury, Shammur Absar, 41
Chrupała, Grzegorz, 36
Cieliebak, Mark, 37
Cimiano, Philipp, 33
Cimino, Andrea, 49
Ciobanu, Alina Maria, 49
Clausen, Yulia, 49
Clematide, Simon, 33
Cocos, Anne, 62
Cohn, Trevor, 33, 58
Costa-jussà, Marta R., 27, 33, 56
Courville, Aaron, 38
Crego, Josep, 26
Crutchley, Patrick, 63
Cucchiarelli, Alessandro, 34

- Currey, Anna, 27, 31
Cutler, Hannah, 62
Daelemans, Walter, 36, 46
Dagan, Ido, 64
Dai, Falcon, 32
Dalbelo Basic, Bojana, 34
Danchenko, Pavel, 29
Danforth, Douglas, 47
Danieli, Morena, 41
Däniken, Pius von, 37
Das, Dipanjan, 33
Daumé III, Hal, 30, 38, 58
Davis, James, 28
De Clercq, Orphee, 54
de Freitas, Nando, 72
De Gussem, Jeroen, 46
Dehdari, Jon, 36, 58
Deksne, Daiga, 27
Del, Maksym, 27
Delli Bovi, Claudio, 64
Dell'Orletta, Felice, 49
Deng, Yongchao, 26
Derczynski, Leon, 37, 38
Dernoncourt, Franck, 64
Desarkar, Maunendra Sankar, 55
Dey, Lipika, 54
Di Gangi, Mattia Antonino, 26
Ding, Shuoyang, 26
Ding, Yuning, 49
Dinu, Liviu P., 49
Dipper, Stefanie, 50
Do, Bich-Ngoc, 33
Donato, Giulia, 54
Dorsch, Jonas, 51
Dredze, Mark, 36
Dubrawski, Artur, 36
Duh, Kevin, 26
Duma, Melania, 29, 30
Duma, Mirela-Stefania, 28
Duppada, Venkatesh, 54
Duque, Thyago, 33
Duselis, John, 28
Dutta Chowdhury, Koel, 28
Dwiastuti, Meisyarah, 48
Dwojak, Tomasz, 31
Dymetman, Marc, 33
Edelstein, Lilach, 51
Ehlert, Anna, 50
Eichstaedt, Johannes, 63
Ekbal, Asif, 55
Elliott, Desmond, 26
Elsner, Micha, 58
Emmery, Chris, 36
Erdmann, Grant, 27, 31
Erp, Marieke van, 37
Escolano, Carlos, 27, 33
Espinosa Anke, Luis, 36
Ettinger, Allyson, 58
Evanini, Keelan, 47
Evans, Richard, 48
Fairon, Cédrick, 48
Falenska, Agnieszka, 33
Falke, Tobias, 61
Farag, Youmna, 48
Farajian, M. Amin, 30, 31
Faralli, Stefano, 64
Faruqui, Manaal, 19, 24, 32
Favre, Benoit, 56
Federico, Marcello, 26, 31
Federmann, Christian, 26
Felice, Mariano, 49
Feng, Shi, 30
Feng, Will, 63
Ferrara, Alfio, 52
Ferrés, Daniel, 58
Ferret, Olivier, 34
Ficler, Jessica, 46
Fierro, Constanza, 51
Firat, Orhan, 43
Fisch, Adam, 63
Fishel, Mark, 27, 30
Flint, Emma, 37, 47
Fokkens, Antske, 34
Fonollosa, José A. R., 27, 33, 56
Ford, Elliot, 37
Foster, George, 27
François, Thomas, 48
Frank, Anette, 33
Frank, Stella, 26
Fraser, Alexander, 27, 29
Fuentes, Claudio, 51
Fulgoni, Dean, 62
Fung, Pascale, 73
Fürstenau, Hagen, 29
Fyshe, Alona, 48
Gambäck, Björn, 37
Gangal, Varun, 45
García, Marcos, 53
García-Martínez, Mercedes, 27, 28
Gardner, Matt, 37
Gaussier, Eric, 33
Geiß, Johanna, 64
Genabith, Josef van, 58
Germann, Ulrich, 27
Gers, Felix A., 58
Gertz, Michael, 64
Ghaly, Hussein, 41
Ghanimifard, Mehdi, 45
Ghobadi, Mina, 52
Gholipour Ghalandari, Demian, 39
Giannakopoulos, Athanasios, 54
Gildea, Daniel, 27

- Ginter, Filip, 43
Giorgi, Salvatore, 63
Gkatzia, Dimitra, 33
Glavaš, Goran, 34
Gloncak, Vladan, 30
Goel, Pranav, 53
Goldberg, Yoav, 46, 56
Goldwater, Sharon, 41, 45, 68
Gómez Guinovart, Xavier, 58
Gómez-Rodríguez, Carlos, 53
Gonzalez-Garduño, Ana Valeria, 49
Goot, Erik van der, 53
Goot, Rob van der, 36
Gordon, Jonathan, 48
Goutte, Cyril, 49
Goyal, Raghav, 33
Graça, Miguel, 27
Graham, Yvette, 26, 29, 30
Gravier, Guillaume, 34
Green, Nancy, 51
Gretz, Shai, 51
Gridach, Mourad, 36
Grigonyte, Gintare, 48, 49
Grishina, Yulia, 43, 44
Gronas, Mikhail, 34
Grönroos, Stig-Arne, 27
Grozea, Cristian, 26
Gulati, Anmol, 56
Gulcehre, Caglar, 19, 20
Guo, Han, 39
Gurevych, Iryna, 39, 51, 53, 61
Gurnani, Nishant, 56
Guta, Andreas, 27
Gwinnup, Jeremy, 27, 28, 31
- Ha, Thanh-Le, 27
Habernal, Ivan, 51, 61
Haddad, Hatem, 36
Haddow, Barry, 26, 27, 31
Haffari, Gholamreza, 33
Hage, Willem van, 34
Haider, Thomas, 45
Halder, Kishaloy, 54
Hamill, Christopher, 47
Han, Xiaotian, 55
Hannemann, Raffael, 61
Hardmeier, Christian, 43, 44
Hardt, Daniel, 34
Harrison, Jackie, 34
Hartmann, Mareike, 87
Hartung, Matthias, 53
Hasler, Eva, 62
He, Xuanli, 33
He, Yuanye, 55
Heafield, Kenneth, 27, 31
Hearst, Marti A., 49
Helcl, Jindřich, 28, 29, 31
Herranz, Luis, 28
- Herrasti, Alvaro, 62
Hidey, Christopher, 51
Hill, Felix, 56
Hiray, Sushant, 54
Hitomi, Yuta, 34
Hitschler, Julian, 45
Hladka, Barbora, 48
Hokamp, Chris, 30
Holtz, Chester, 27
Holub, Martin, 48
Homayon, Sheideh, 45
Honari Jahromi, Maryam, 48
Hopkins, Mark, 62
Hoque, Enamul, 39
Horbach, Andrea, 48, 49
Hossmann, Andreea, 54
Hoste, Veronique, 54
Hou, Yufang, 51
Hovy, Dirk, 37, 41, 87
Hovy, Eduard, 37, 45
Hu, Guoning, 37
Hu, Jiawei, 28
Hu, Jinming, 27
Hua, Xinyu, 40
Huang, Chieh-Yang, 63
Huang, Liang, 28
Huang, Liu, 30
Huang, Ting-Hao, 63
Hübsch, Ondřej, 30
Huck, Matthias, 26, 27, 29
Hulden, Mans, 32
Hummel, Shay, 51
Hurskainen, Arvi, 27
Hutt, Michael, 28
Hwa, Rebecca, 4
Hwang, Alyssa, 51
- Inkpen, Diana, 56
Inui, Kentaro, 34
Ionescu, Radu Tudor, 48
Ircing, Pavel, 48
- Jackson, Bradley, 62
Jacobs, Gilles, 54
Jaffe, Alan, 28
Jaffe, Evan, 47
Jagfeld, Glorianna, 41
Jain, Prayas, 53
James, Matthew, 2
Jamet, Eric, 34
Jang, Myungha, 37
Jansson, Patrick, 37
Jean, Sébastien, 43
Jebbara, Soufian, 33
Jeon, Hyung-Bae, 49
Jhamtani, Harsh, 45
Jiang, Hui, 56
Jiang, Liyang, 27

- Jiang, Shu, 48
Jiang, Song, 55
Jiang, Ye, 34
Jimeno Yepes, Antonio, 26
Jin, Huiming, 32
Jin, Lifeng, 47
Jin, Yaohong, 44
Jochim, Charles, 51
Johansson, Richard, 32
John, Vineet, 55
Joshi, Aditya, 11, 13, 53
Joshi, Vidur, 62
Junczys-Dowmunt, Marcin, 30, 31
Jurafsky, Dan, 65
Jurczyk, Tomasz, 58
- Kageura, Kyo, 35
Kamran, Amir, 29
Kan, Min-Yen, 54
Kanerva, Jenna, 43
Kann, Katharina, 32, 33
Kannan, Abishek, 54
Kazi, Michaeel, 27
Ke, Chuyang, 27
Kepler, Fabio, 30, 49
Kestemont, Mike, 46
Khayrallah, Huda, 26
Kiesel, Johannes, 52
Kim, Evgeny, 53
Kim, Hyun, 30
Kim, Jihie, 33
Kim, Jun-Seok, 65
Kim, Jungi, 26
Kim, Taeuk, 32
King, David, 58
Kirby, James, 45
Kittner, Madeleine, 26
Klamm, Christopher, 61
Klatt, Tobias, 58
Klein, Guillaume, 26
Klinger, Roman, 53
KOBUS, Catherine, 26
Kochmar, Ekaterina, 49
Kocmi, Tom, 26, 29
Koehn, Philipp, 26, 31
Kolb, Peter, 47
Kolhatkar, Varada, 35
Köper, Maximilian, 53
Koppel, Moshe, 45
Kordjamsjidi, Parisa, 38
Koreeda, Yuta, 62
Kreutzer, Julia, 29
Krišlauks, Rihards, 27
Ku, Lun-Wei, 63
Kuang, Yi Zong, 28
Kübler, Sandra, 49
Kuhn, Jonas, 63
Kuhn, Roland, 27
- Kulkarni, Nilesh, 33
Kulmizev, Artur, 49
Kulshreshtha, Devang, 53
Kumar, Gaurav, 26
Kunchukuttan, Anoop, 32
Kurimo, Mikko, 27
Kurohashi, Sadao, 48
Kurotsuchi, Kenzo, 62
- Laarmann-Quante, Ronja, 50
Labeau, Matthieu, 26, 32
Labetoulle, Tristan, 63
Lai, K. Robert, 55
Lakhani, Aazim, 48
Lakomkin, Egor, 54
Lam, Wai, 40
Lancaster, Meredith, 28
Langlais, Philippe, 44
Lapshinova-Koltunski, Ekaterina, 43, 44
Larkin, Samuel, 27
Lauly, Stanislas, 43
Lawrence, John, 51, 52
Laws, Florian, 41
Lazaridou, Angeliki, 56
Le, Quoc, 31
Leacock, Claudia, 47
Lee, Chong Min, 47, 48
Lee, Haejun, 33
Lee, Jacob Hyun, 28
Lee, Jaesong, 65
Lee, Ji Young, 64
Lee, John, 48
Lee, Jong-Hyeok, 30
Lee, Sang-goo, 32
Lee, Timothy, 54
Lee, Yun-Keun, 49
Lee, Yun-Kyung, 49
Leemans, Inger, 34
Lefever, Els, 54
Léger, Serge, 49
Lejetine, Gaël, 32
Levy, Omer, 56
Levy, Ran, 51
Li, Dapeng, 28
Li, Hongzheng, 44
Li, Maoxi, 29, 30
Li, Muze, 27
Li, Piji, 40
Li, Wen, 49
Li, YINGYA, 35
Li, Yitong, 58
Liakata, Maria, 34
Libovický, Jindřich, 28, 29
Lichtblau, Yvonne, 26
Limsopatham, Nut, 37
Lin, Bill Y., 37
Lin, Grace, 45
Ling, Guangming, 48

- Ling, Jeffrey, 39
Ling, Zhen-Hua, 56
Litman, Diane, 47, 51
Litvinova, Olga, 45
Litvinova, Tatiana, 45
Liu, Fei, 39
Liu, LuChen, 38
Liu, Nelson F., 37
Liu, Qun, 28, 30
Liu, Shuhua, 37
Liu, Weiyi, 55
Liu, Zheyuan, 62
Lloret, Elena, 33
Lo, Chi-ku, 27, 30
Loaiciga, Sharid, 43
Logacheva, Varvara, 26
López Monroy, Adrian Pastor, 37
Lopez, Adam, 41
Löser, Alexander, 58
Loughnane, Robyn, 47
Loukina, Anastassia, 41, 47, 50
Loyola, Pablo, 57
Lu, Jiasen, 63
Lu, Wei, 19, 22
Lugini, Luca, 47
Lui, Ruishen, 52
Luo, Zhiyi, 37
Luotolahti, Juhani, 43
- Ma, Mingbo, 28
Ma, Qingsong, 30
Ma, Yuan, 52
Madhyastha, Pranava Swaroop, 28
Madisetty, Sreekanth, 55
Madnani, Nitin, 41, 50
Maharjan, Suraj, 37
Mahler, Taylor, 58
Makarov, Peter, 33
Maks, Isa, 34
Malakasiotis, Prodromos, 34
Malliaros, Fraggiskos D., 11, 15
Malmasi, Shervin, 47
Mamidi, Radhika, 54, 58
Mandel, Michael, 41
Manjavacas, Enrique, 46
Manzoor, Umar, 38
Marasek, Krzysztof, 28
Marcheggiani, Diego, 11, 17
Mareček, David, 26, 30
Markov, Ilia, 49
Marneffe, Marie-Catherine de, 58
Marrese-Taylor, Edison, 54, 55, 57
Marten, Fide, 64
Martins, André F. T., 30
Masana, Marc, 28
Mascarell, Laura, 28, 44
Matsuo, Yutaka, 54, 55, 57
Maynard, Diana, 34
- McCaffrey, Dan, 48
McCallum, Andrew, 38
McCarthy, Michael, 41
McCoy, Kathleen, 39
McCurdy, Kate, 47
McKeown, Kathy, 51
Mechanic, Ross, 62
Meerkamp, Philipp, 38
Meersbergen, Maarten van, 34
Meinel, Christoph, 56
Meisher, Hardik, 54
Meladianos, Polykarpou, 39
Melese, Michael, 41
Menzel, Wolfgang, 28–30
Mesgar, Mohsen, 43
Meshesha, Million, 41
Meurers, Detmar, 49
Meyers, Benjamin, 61
Miceli Barone, Antonio Valerio, 27, 31
Miculicich Werlen, Lesly, 43
Miks, Toms, 27
Milic-Frayling, Natasa, 34
Milintsevich, Kirill, 32
Miller, Alexander, 63
Miller, John, 39
Miller, Kyle, 36
Misawa, Shotaro, 33
Mitkov, Ruslan, 48
Miura, Akiva, 31
Miura, Yasuhide, 33
Mogren, Olof, 32
Mohammad, Saif, 53
Mohammadi, Elham, 48
Mohanty, Gaurav, 54
Mohme, Julian, 53
Montanelli, Stefano, 52
Montavon, Grégoire, 54
Monteiro, Jose, 30
Monz, Christof, 26
Morari, Viorel, 51
Morbidi, Christian, 34
Moreno Schneider, Julian, 34
Moreno-Ortiz, Antonio, 55
Moreno-Schneider, Julian, 34
Morin, Emmanuel, 34
Mukku, Sandeep Sriraman, 58
Mulki, Hala, 36
Müller, Klaus-Robert, 54
Müller, Mathias, 43
Muresan, Smaranda, 51
Musat, Claudiu, 54
Musi, Elena, 51
Musil, Tomáš, 29
Myrianthous, Giorgos, 34
- Na, Seung-Hoon, 30
Nadeem, Farah, 49
Nadejde, Maria, 31

- Nagpal, Chirag, 36
Nakamura, Satoshi, 28, 31
Nakov, Preslav, 43
Nangia, Nikita, 56
Napoles, Courtney, 48, 49
Napolitano, Diane, 47
Natarajan, Prem, 48
Negri, Matteo, 26, 30, 31
Neubig, Graham, 28, 31
Neveol, Aurelie, 26
Neves, Mariana, 26
Ney, Hermann, 27, 31
Ng, Hwee Tou, 49
Nguyen, Khanh, 30
Nguyen, Le-Minh, 27
Nguyen, Viet, 33
Nichols, Eric, 37
Nie, Yixin, 56
Niehues, Jan, 27, 31
Nieminen, Tommi, 27
Nissim, Malvina, 36, 49
Niu, Xing, 45
Niwa, Yoshiki, 62
Noord, Gertjan van, 49
Nyberg, Eric, 45
- O'Neill, James, 54
Oberhauser, Tom, 58
O'Connor, Brendan, 36
Oh, Yoo Rhee, 49
Ohkuma, Tomoko, 33
Okazaki, Naoaki, 34
Oncevay, Arturo, 33
Oraby, Shereen, 45
Orasan, Constantin, 48
O'Reilly, Tenaha, 47
Ororbia II, Alexander, 38
Ortmann, Katrin, 50
Ostendorf, Mari, 49
Östling, Robert, 27, 48, 49
Ovesdotter Alm, Cecilia, 61
Özkahraman, Özer, 35
- Padilla López, Rebeca, 54
Padó, Sebastian, 53
Pado, Ulrike, 47
Paetzold, Gustavo, 30
Paggio, Patrizia, 54
Pal, Santanu, 30
Palm, Rasmus Berg, 41
Palmer, Alexis, 45
Palmer, Martha, 2
Panchenko, Alexander, 64
Papandrea, Simone, 64
Pappas, Nikolaos, 28
Parikh, Devi, 63
Park, Jeon-Gue, 49
Pasca, Marius, 11, 12
- Pasunuru, Ramakanth, 39
Paul, Michael J., 36
Pauli, Patrick, 61
Pavlopoulos, John, 34
Pawar, Jyoti, 55
Pecina, Pavel, 26
Pekar, Viktor, 54
Penn, Gerald, 33
Pérez, Jorge, 51
Peris, Álvaro, 31
Petasis, Georgios, 51, 52
Peter, Jan-Thorsten, 27
Peters, Ben, 58
Petrescu-Prahova, Cristian, 62
Peyrard, Maxime, 39
Pham, Ngoc-Quan, 27
Pham, Trung-Tin, 27
Pietquin, Olivier, 30
Pineau, Joelle, 38
Ping, Qing, 39
Pinnis, Márvis, 27
Pisarevkaya, Dina, 34
Plank, Barbara, 36, 48, 49
Poddar, Lahari, 54
Poesio, Massimo, 34
Pollak, Christian, 61
Ponomareva, Maria, 32
Ponzetto, Simone Paolo, 48, 64
Poornachandran, Prabaharan, 55
Popescu, Marius, 48
Popescu, Octavian, 34
Popescu-Belis, Andrei, 28, 43
Popović, Maja, 30
Porkaew, Peerachet, 28
Post, Matt, 2, 26
Potash, Peter, 34
Potthast, Martin, 39, 51
Press, Ofir, 32
Priego, Belem, 55
Prud'hommeaux, Emily, 61
Pryzant, Reid, 31
Pu, Jiashu, 34
Pu, Xiao, 28
Pugh, Robert, 47
Purver, Matthew, 34
Puschmann, Jana, 51
- Qian, Feng, 38
Qian, Yao, 47
Qu, Jiani, 51
Quegan, Shaun, 34
Quezada, Mauricio, 51
- R, Vinayakumar, 55
Raganato, Alessandro, 64
Rahgooy, Taher, 38
Rajana, Sneha, 62

- Rama, Taraka, 48
Rambow, Owen, 34
Rao, Sudha, 58
Raschkowski, Willi, 56
Reddy, Siva, 31
Reed, Chris, 51, 52
Rehbein, Ines, 33, 45
Rehm, Georg, 34
Rei, Marek, 47–49
Reichart, Roi, 56
Rey, Arnaud, 56
Ribeiro, Swen, 34
Riccardi, Giuseppe, 39, 41
Richardson, Kyle, 63
Richter, Ludwig, 64
Riedel, Sebastian, 4
Riess, Simon, 29
Riester, Arndt, 41
Riezler, Stefan, 29
Rifqi Fatchurrahman, Muhammad, 48
Rikters, Matiss, 27
Ring, Nico, 56
Riordan, Brian, 48
Rios Gonzales, Annette, 28
Rios, Miguel, 27
Ritter, Alan, 38
Robert, Maxime, 34
Roekhaut, Sophie, 48
Roesiger, Ina, 41
Rohanian, Omid, 48
Roller, Roland, 26
Romanov, Alexey, 34
Ronen, Hadar, 64
Rosa, Rudolf, 26, 30
Rosendahl, Jan, 27
Rossenbach, Nick, 27
Roth, Michael, 11, 17
Rubino, Raphael, 26
Ruiz, Nicholas, 41
Rumshisky, Anna, 34
Rupperl, Eugen, 64
Rush, Alexander M., 36
Rush, Alexander, 39
Ruzsics, Tatyana, 33
Rwebangira, Mugizi, 36
- Sabatini, John, 47
Safari, Pooyan, 26
Saggion, Horacio, 36, 58
Saha Roy, Rishiraj, 63
Saha, Rupsa, 54
Saint-Dizier, Patrick, 52
Sakaguchi, Keisuke, 48
Salehi, Bahar, 36, 37
Salesky, Elizabeth, 27
Samek, Wojciech, 54
Sandwick, Joshua, 28
Santia, Giovanni, 37
- Santos, Henrique, 54
Sap, Maarten, 63
Sari, Yunita, 48
Sarnat, Aaron, 62
Sato, Misa, 62
Satria, Arief Yudha, 47
Saunders, Danielle, 62
Sawant, Palaash, 55
Scarton, Carolina, 29
Scherrer, Yves, 27, 43, 44
Schmidtke, Franziska, 53
Schneider, Nathan, 61
Schneider, Rudolf, 58
Scholten-Akoun, Dirk, 49
Schuff, Hendrik, 53
Schulte im Walde, Sabine, 53
Schütze, Hinrich, 32, 33
Schwartz, H. Andrew, 63
Sébillot, Pascale, 34
Segal, Natalia, 26
Selent, Stefan, 47
Senellart, Jean, 26
Sennrich, Rico, 27, 28, 31
Seol, Jinseok, 32
Serban, Iulian Vlad, 38
Seredin, Pavel, 45
Servan, Christophe, 26
Sha, Lei, 38
Shahzad, Uman, 28
Shain, Cory, 58
Shapira, Ori, 64
Sharaf, Amr, 30, 38
Sheng, Emily, 48
Shi, Lin, 27
Shin, Bonggun, 54
Shin, Joong-Hwi, 65
Shin, Youhyun, 32
Shoemark, Philippa, 45
Shukla, Kaushal Kumar, 53
Shutova, Ekaterina, 49
Sics, Valters, 27
Sidorov, Grigori, 49
Sikdar, Utpal Kumar, 37
Silfverberg, Miikka, 32
Sim Smith, Karin, 44
Singla, Karan, 39
Sinha, Priyanka, 54
Siu, Amy, 26
Sliwa, Alfred, 52
Slonim, Noam, 51
Smiley, Charese, 49
Socher, Richard, 38
Søgaard, Anders, 19, 24, 36, 37, 45, 49, 56
Sokolov, Artem, 29
Soler, Juan, 36
Solorio, Thamar, 37, 45
Song, Hwa Jeon, 49
Song, Xingyi, 34

- Spasojevic, Nemanja, 37
Specia, Lucia, 26–31
Sperber, Matthias, 27
Spitz, Andreas, 64
Srikumar, Vivek, 36
Srivastava, Ankit, 30, 34
Srivastava, Vallari, 63
Stahlberg, Felix, 62
Staliūnaitė, Ieva, 59
Starostin, Anatoly, 32
Stasaski, Katherine, 49
Stehwien, Sabrina, 41
Stein, Benno, 39, 51, 52
Stepanov, Evgeny, 39, 41
Stevens-Guille, Symon, 58
Stewart, Darlene, 27
Stiefel, Moritz, 41
Stilo, Giovanni, 34
Stilson, Brandon, 62
Strapparava, Carlo, 34, 49
Stratos, Karl, 33, 38
Strube, Michael, 43
Strubell, Emma, 38
Stymne, Sara, 43
Sudarikov, Roman, 26
Sudoh, Katsuhito, 31
Sumita, Eiichro, 28
Sunderland, Kellen, 29
Svec, Jan, 48
Syed, Shahbaz, 39
Sznajder, Benjamin, 51
Szolovits, Peter, 64
Szymaniak, Witold, 29
- Taboada, Maite, 35
Tack, Anaïs, 48
Täckström, Oscar, 33
Takhanov, Rustem, 33
Tamchyna, Aleš, 29
Tamori, Hideaki, 34
Tan, Yiming, 29, 30
Tan, Zhixing, 27
Tang, Gongbo, 27
Tang, Linyuan, 35
Taniguchi, Motoki, 33
Tannier, Xavier, 34
Tättar, Andre, 30
Taylor, Jonathan, 27
Tekir, Selma, 35
Tetreault, Joel, 47, 48
Teufel, Simone, 51
Thaine, Patricia, 33
Thomas, Olivia, 37
Thomas, Philippe, 26
Thompson, Brian, 27
Thorne, James, 34
Tiedemann, Jörg, 27, 43, 44
Tilk, Ottokar, 39
- Titov, Ivan, 11, 17
Tixer, Antoine, 39
Tokunaga, Takenobu, 47, 51
Tolmachev, Arseny, 48
Tomar, Gaurav Singh, 33
Toprak, Mustafa, 35
Trancoso, Isabel, 32
Trescher, Saskia, 26
Trieu, Long, 27
Turchi, Marco, 26, 30, 31
Tursun, Osman, 37
Tutek, Martin, 34
- Ungar, Lyle, 63
Ustalov, Dmitry, 64
Uszkoreit, Jakob, 33
Utiyama, Masao, 28
- Šics, Valters, 27
Šnajder, Jan, 34, 53
Štajner, Sanja, 48
Vajjala, Sowmya, 48
Van Durme, Benjamin, 11, 17
Van Genabith, Josef, 36
Varis, Dusan, 26, 30
Vazirgiannis, Michalis, 11, 15, 39
Vechtomova, Olga, 55
Veisi, Hadi, 48
Velardi, Paola, 34
Versley, Yannick, 43
Verspoor, Karin, 26
Vieira, Renata, 54
Vilares, David, 53
Virpioja, Sami, 27
Vlachos, Andreas, 34
Vogel, Lars, 53
Vogel, Maurice, 50
Vollgraf, Roland, 62
Völske, Michael, 39
Vossen, Piek, 34
Vu, Hoa, 57
Vu, Ngoc Thang, 33, 41
Vulić, Ivan, 19, 24
Vylomova, Ekaterina, 33
- Wabnitz, David, 34
Wachsmuth, Henning, 51, 52
Waibel, Alexander, 27
Waibel, Alex, 27
Walker, Marilyn, 45, 53
Walker, Vern, 51
Wang, Bo, 26
Wang, Boli, 27
Wang, Jin, 54
Wang, Josiah, 28
Wang, Lu, 39, 40
WANG, Mingwen, 29

- Wang, Mingwen, 30
Wang, Shugen, 30
Wang, Xiaoxuan, 34
Wang, Yanfeng, 27
Wang, Yuguang, 27
Webber, Bonnie, 43
Wei, Johnny, 36
Wei, Si, 56
Weijer, Joost van de, 28
Weikum, Gerhard, 63
Welbl, Johannes, 37
Weller-Di Marco, Marion, 29
Wermter, Stefan, 54
Weston, Jason, 63
White, Michael, 47, 58
Whittaker, Steve, 53
Wieling, Martijn, 49
Williams, Adina, 56
Williams, Jake, 36, 37
Williams, Philip, 27
Winther, Ole, 41
Wisniewski, Guillaume, 30
Wolf, Lior, 32
Wolk, Krzysztof, 28
Wolska, Magdalena, 49
Wood, Ian, 54
Wooters, Chuck, 61
Wu, Jiaqi, 53
- Xiang, Qingyu, 29
Xing, Linzi, 36
Xu, Frank, 37
Xu, Jia, 28
Xu, Wei, 38, 45
- Yaghoobzadeh, Yadollah, 32
Yahya, Mohamed, 63
Yamada, Hiroaki, 51
Yanai, Kohsuke, 62
Yanase, Toshihiko, 62
Yaneva, Victoria, 48
Yang, Han, 56
Yang, Haojin, 56
Yang, Hongtao, 27
Yang, Jiajun, 27
Yannakoudakis, Helen, 47
Young, Katherine, 27, 31
Yu, Bei, 35
Yu, Liang-Chih, 55
Yu, Seunghak, 33
Yu, Xiang, 33
Yuan, Hang, 54
Yuan, Zheng, 49
Yvon, François, 26–29
- Zagorovskaya, Olga, 45
Zajic, Zbynek, 48
Zampieri, Marcos, 49

Local Guide

*This guide was written by Maria Barrett, Joachim Bingel, Mareike Hartmann, Dirk Hovy.
For the most up-to-date version, please visit <http://www.emnlp2017.net>*

General

Welcome to Copenhagen!

To **get around** the city, best buy a multi-day pass or **rent bikes** at a local bike shop. There are also city bikes at many locations throughout the city. When biking, clearly indicate where you are going. Do not stop without indicating so (by raising a hand as if to greet someone). Do not swerve and change lanes. Any of these things will bring out the inner Viking in otherwise mild-mannered Danish cyclists.

Public transport is excellent and gets you everywhere. You certainly don't need a taxi between almost anywhere in the city and the airport. But note that there are three zones to get from the airport to central Copenhagen (and regular ticket controls!). Otherwise two zones will get you around central Copenhagen. Find connections between addresses in Denmark at <http://journeyplanner.dk>. If you stay longer and travel more than 10 times, you might want to get a **Rejsekort** (reloadable travel card valid throughout all of Denmark on all modes of transportation), which you can buy at vending machines placed at every metro station. Note that the card itself costs 80 DKK (non-refundable), and you'll need to top it up before your first ride.

For getting a sense of the **city's geography**, it is helpful to realize that there are three major boroughs surrounding the Inner City in the West, North, and East, which are aptly called *Vesterbro* (where the venue is located), *Nørrebro*, and *Østerbro*. Crammed between *Vesterbro* and *Nørrebro* lies *Frederiksberg*. In the north, the Inner City is walled off by a stretch of artificial lakes, *Søerne*, and in the south, on the other side of the canal, there's the island of *Amager*.

You can pay with **credit card** pretty much anywhere, although some places (usually smaller ones) require a Danish card.

The **country code** is +45.

Sightseeing

Kødbyen is a lively place, and you can spend the entire conference within the walls of the old meatpacking district, trying new restaurants, listening to upcoming bands, and drinking hipster coffee with the locals. However, DGI Byen – the part of Kødbyen, in which Øksnehallen and Cph Conference are located – contains a number of other facilities. This includes indoor swimming and gym, that are free for participants. Check out <http://www.dgi-byen.com/> for more information. The following sightseeing suggestions all take you out of Kødbyen.

Carlsberg Glyptotek

Nice collection of statues and paintings. The courtyard has a cafe where you can sit among palm trees. Free entrance on Tuesdays.

<http://www.glyptoteket.com>

Dantes Plads 7 1556 Copenhagen K	Tue, Wed, Fri, Sat, Sun 11:00–18:00 Thu 11:00–22:00
-------------------------------------	--

Christiania

Mostly known for weed and dead hippie dreams. Actually has nice jazz concerts and other events (often in 'Operaen', not to be confused with the big opera house). The old parts along the ramparts are full of nice, self-made houses. Nice vegetarian restaurant, Morgensteden.

<http://www.visitcopenhagen.com/copenhagen/culture/alternative-christiania>

Main entrance is on Prinsessegade, but there are several other ways into the area.

Danish Design Museum

Next to Kastellet in an old hospital. A series of several rooms arranged around a courtyard, each presenting one decade or artist of Danish design (lamps, chairs, etc.). Nice cafe, and a museum shop with many of the things you just saw.

<http://www.designmuseum.dk/en>

Bredgade 68 1260 Copenhagen K	Tue, Fri, Thu, Sat, Sun 10:00–18:00 Wed 10:00–21:00
----------------------------------	--

Den Sorte Diamant

The Royal Library. Modern addition to the old library. Great architecture. Changing exhibitions in the basement, a free exhibition of rare books and old manuscripts in a pop-art decorated room upstairs. Also features changing photography exhibitions.

<http://www.kb.dk/en/>

Søren Kierkegaards Pl. 1 1221 Copenhagen K	Mon–Fri 08:00–21:00 Sat 09:00–19:00
---	--

Get a GoBoat

Get an electric-powered boat for up to eight people and set to sea (well, don't actually set to sea: the canal). You can explore Copenhagen's waterways on your own while enjoying some drinks and snacks that you bring yourself or pick up, e.g. at Papirøen.

<http://goboard.dk/en/>

Islands Brygge 10
2300 Copenhagen S

Harbour bus

Cheapskate boat sightseeing. Pay a regular bus ticket or show your multi-day pass and get a sea view of Copenhagen. Stops e.g. by Den Sorte Diamant, Nyhavn, The Opera House and Langelinje.

<https://www.google.dk/maps/place/Havnebussen/>

Kastellet

The old castle at the northern end of the city. Shaped like a star, and fun to walk on. Next to the Little Mermaid, which you can happily skip, but if you must, you can see it from here. Was built to defend the city, but captured before it was finished.

National Museum

One of the best Viking collections in the world.

<http://en.natmus.dk/museums/the-national-museum-of-denmark/>

Ny Vestergade 10
1471 Copenhagen K

Rundetårn

Round tower with a spiraling pathway inside, built so that the king could ride up on a horse to the observatory. Nice view over the city, and you can run down the spiral and shout "wheee!"

<http://www.rundetaarn.dk/en/>

Købmagergade 52A
1150 Copenhagen K

Torvehallerne

Two modern market halls, filled with food vendors and food-related items. Some overpriced stuff, but also a lot of specialized shops with things that are otherwise hard to find. The smørrebrød shop is becoming a new in-spot.

<http://torvehallernekbh.dk>

Frederiksborggade 21 Mon–Wed 10:00–19:00 Sat 10:00–18:00
1360 Copenhagen K Thu–Fri 10:00–20:00 Sun 11:00–17:00

The Little Mermaid

Yeah... Don't be too disappointed.

Langelinie
2100 Copenhagen Ø

Statens Museum for Kunst

Decent art museum. Has a bar on the first Friday of the month.

<http://www.smk.dk/en/>

Sølvgade 48-50	Tue–Sun 11:00–17:00
1307 Copenhagen K	Wed 11:00–20:00

Ørestad

If you like new architecture, take the metro and visit Ørestad. It is a new area of Copenhagen with award-winning houses. There is a mall directly in front of Ørestad station (Field's), or take the metro out to Vestamager station and walk up the Stallet house.

<http://www.visitcopenhagen.com/copenhagen/architecture/architectural-orestad>

Jægersborggade

This street in Nørrebro is a hipster's paradise. Very close to Assistens Kirkegård, it is home to a number of cafés/bars, galleries, craft and (second-hand) clothing shops, as well as combinations thereof (SneakersCoffee, BeerVinyl, ...). Great for finding gifts to take home.

<http://jaegersborggade.com/wpAB/en/shops/>

Shops

B&W Hallerne. *Second hand furniture in a giant, factory hall, by the water, in industrial Copenhagen. Only open every other weekend. Bring cash.* Refshalevej 171, 1432 Copenhagen. <https://www.facebook.com/BW-LOPPEMARKED-171235476283625/>

Faraos Cigarer. *Cartoon shop – or three, actually, but all next-door – close to Rundetårn.* Skindergade 27, 1159 Copenhagen K. <https://www.faraos.dk>

Sort kaffe & Vinyl. *Hipster record shop with good coffee.* Skydebanevænge 4 1709 Copenhagen V. <https://www.facebook.com/sortkaffeogvinyl/>

København K. *Second-hand clothing in a street packed with fashion stores.* Studiestræde 30, 1455 Copenhagen K. <http://koebenhavnk.com>

Dance and Music Venues

Global CPH. *World music.* Nørre Allé 7, 2200 Copenhagen N. <http://globalcph.dk/english/>

Danshallerne. *Dance and theatre.* Bohrsgade 19, 1799 Copenhagen V. <http://www.dansehallerne.dk/en/>

La Fontaine. *Intimate, 100-capacity veteran of the Scandinavian jazz scene staging nightly jam sessions.* Kompanistræde 11, 1208 Copenhagen K. <http://lafontaine.dk>

Montmartre. *Classic jazz joint in central Copenhagen.* Store Regnegade 19A, 1110 Copenhagen K. <http://www.jazzhusmontmartre.dk>

Mojo. *Blues, jazz & folk from Scandinavian & international artists in a compact club open until 5am.* Løngangstræde 21C, 1468 Copenhagen K. <https://mojo.dk>

Pumpehuset. *Rock, pop, hip hop, world music – lots of different music styles, but generally nice, national and international bookings.* Studiestræde 52, 1554 Copenhagen V. <http://pumpehuset.dk/>

Royal Theatre. *See ballet at the old stage, a play in the new theatre house by the sea or opera in the Opera House.* <https://kglteater.dk/en/>

Vega. *Rock and pop scene not too far from the venue.* Enghavevej 40, 1674 Copenhagen V. <http://vega.dk>

Events

Moreover, if your conference schedule permits, here's a list of events happening in Copenhagen and the wider region, during EMNLP:

Cirque du Soleil. Malmö Arena, September 6–10. Hyllie Stationstorg 2, 215 32 Malmö, Sweden. <https://www.cirquedesoleil.com/sweden/malmo/shows>

Copenhagen World Music. Runs September 6–10 at different venues in Copenhagen. <http://cpwhorld.dk>

Luisi leads Danish composer Carl Nielsen's 5th Symphony in Koncerthuset. September 7. Koncerthuset, Emil Holms Kanal 20, 2300 Copenhagen S. <http://drkoncerthuset.dk/>

Day-Trips

Arken

Modern art museum south of the city, in an unassuming suburb. Nice changing exhibitions and a modern collection.

<http://uk.arken.dk>

Skovvej 100
2635 Ishøj

Louisiana

Modern art museum on the coast north of the city. The buildings are integrated into a park overlooking the cliffs towards Sweden. Great exhibits and a world-renowned collection. Sneak in picnic stuff and eat on the lawns.

<https://en.louisiana.dk>

G1 Strandvej 13
3050 Humlebæk

Roskilde Viking Ship Museum

30 min by train from the main station, in a lovely small town lies this museum, which contains five well-preserved Viking ships that were sunk there to form a barrier. See the 1:1 viking ship reconstructions and try to sail one.

<http://www.vikingeskibsmuseet.dk/en/>

Vindeboder 12
4000 Roskilde

Other Things to Do

If you're less into classical sightseeing but want to live a day as a Copenhagener would (big claim, we know...), the following suggestions might be interesting. Obviously, the best way to get around is by bike.

Dance the Night Away

If you're into electronic music, you'll probably enjoy Copenhagen's old Meat Packing District (*Kødbyen*), where all the following bars are located. *Jolene* and *Bakken* are small and sweaty, and true to their origin and location. They look a bit run-down (the area used to be all slaughterhouses and meat auction halls). There's also KB18 with more deep house/techno, but don't go there before 1 am. *Mesteren & Lærlingen* plays hip hop/soul/reggae.

Parks

Copenhagen has a number of very nice parks that invite you to hang out or do sports. The biggest, *Fælledparken*, is right next to the Department of Computer Science and has a little lake, otherwise it's a big green meadow with lots of runners and football players. A hidden gem with a super beautiful (albeit artificially created) lake, or rather trench, is *Østre Anlæg*, right behind the National Art Gallery. You can have a barbecue there if you bring your own coal, stationary grills are provided. *Assistens Kirkegård* in Nørrebro is a very beautiful graveyard. People from other places might find this strange, but it's a favourite pastime among Copenhageners to hang out here. *Nørrebroparken* is wonderful in springtime, go there to join young Copenhageners for a beer or a frisbee game (disclaimer: don't literally try to join them, they are Danes. They will panic). Just outside of Copenhagen near Gentofte station, is *Bernstorff's Park* which – aside from usual park stuff and a castle – holds Pometet which is a very broad range of Danish fruit trees as well as a smaller selection of exotic fruit trees. Originally for the royal family, but nowadays free for visitors to sample. During weekends in spring, summer and early autumn you can have afternoon tea in Queen Louise' teahouse or in the her rose garden.

Eating & Drinking

Unless you come from Norway, everything will be more expensive than in your home country. It's best not to convert the prices and just take them as-is. A coffee/latte costs 20–50 DKK, a beer almost everywhere 50 DKK. A main course in a restaurant is typically 150–250 DKK.

The following sections list a selection of restaurants along with the price range and their distance from the conference venue.

Price range indicator:

- \$ Main course < 150 DKK
- \$\$ Main course 150–250 DKK
- \$\$\$ Main course > 250 DKK

Restaurants Close to the Conference Venue

The following restaurants are in very close proximity to the conference venue, most of them in *Kødbyen*, the old meat packing district, which is now a center of Copenhagen nightlife and home to numerous restaurants offering foods from around the globe. The following are all great picks, but there are plenty more for you to discover.

Fiskebaren

0.4 km \$\$\$

Allegedly one of the 10 best fish restaurants in Europe. Very nice dishes, but expensive. Go for the starters and medium dishes and share. Make sure to check out the column-shaped aquarium.

<http://fiskebaren.dk>

Flæsketorvet 100 1711 Copenhagen V	Mon–Thu 17:30–00:00 Sat 11:30–02:00 Fri 17:30–02:00 Sun 11:30–00:00
---------------------------------------	--

Fleisch

0.3 km \$\$

Restaurant in Kødbyen with its own butcher counter. Have some meat-based open-face sandwiches there, or grab a few of their beer sausages to go and enjoy them outside.

<http://www.fleisch.dk>

Slagterboderne 7 1716 Copenhagen V	Tue–Thu 11:30–00:00 Sun 11:30–00:00 Fri–Sat 11:30–01:00
---------------------------------------	--

Hija de Sanchez

0.3 km \$

Ex-Noma chef decided to leave the big business and make everyday tacos. Get 3 for 100DKK.

<http://www.hijadesanchez.dk>

Slagterboderne 8 1716 Copenhagen V	Mon–Fri 17:30–00:00 Sun 17:30–22:00
---------------------------------------	--

Isted Grill

0.7 km \$

Not a restaurant, but a hole-in-the-wall late-night snack food place. Get the flækesteg sandwich (grilled pork roast with pickles and red cabbage) after a night of drinking. Cheap and tasty.

Istedgade 92 1650 Copenhagen V	Sun–Thu 12:00–00:00 Fri–Sat 12:00–02:00
-----------------------------------	--

Madklubben

0.7 km \$

Local Guide

Several locations, closest on Vesterbrogade. Nice "home-made" food, reasonable prices. Very good for groups (with a reservation). Vegetarian options.

<http://madklubben.dk/en/>

Vesterbrogade 62
1620 Copenhagen V

Magasasa Dim Sum & Cocktails

0.5 km \$

The name says it all. Raw interior. Reasonably priced.

<http://magasasa.dk/dim-sum-cocktails/?lang=en>

Flæsketorvet 54–56
1711 Copenhagen V

Mother

0.4 km \$

Great pizza!

<http://mother.dk>

Høkerboderne 9
1712 Copenhagen V

Nose2Tail

0.6 km \$\$_

As cozy as a former slaughterhouse basement can get (the one in the Kødbyen basement is the best of their locations). Serves daily changing meat, fish, and innard dishes. Freshly made cracklings (Danish delicacy made from pork skin), served with bacon mayonnaise – because fat! Chase with a Fernet Branca. Very good beer and friendly staff with awesome leather aprons.

<http://nose2tail.dk/>

Kødboderne 9
1711 Copenhagen
Tue–Thu 18:00–00:00
Fri–Sat 18:00–01:00

PatePate

0.2 km \$

Small plates, international cuisine. Great for sharing. Medium price range. Also nice to start your evening in Kødbyen. Vegetarian options.

<http://www.patepate.dk/>

Slagterboderne 1
1716 Copenhagen V

Mon–Wed	09:00–00:00	Fri	09:00–01:00
Thu	09:00–01:00	Sat	11:30–01:00

Restaurant Cofoco

0.3 km \$\$_

Part of the Cooking for Copenhagen group. Small delicious plates. Slightly upscale, but reasonable.

<http://cofoco.dk/en/>

Abel Cathrines Gade 7
1654 Copenhagen V

Tommi's Burger Joint 0.4 km \$

Tasty minimalistic burgers! Part of a small international chain originating from Iceland.

<https://www.burgerjoint.dk/>

Høkerboderne 21-23 Thu–Sun 11:00–22:00
1712 Copenhagen V Mon–Wed 11:00–21:00

WEDOFOOD 0.2 km \$

Homemade salads. Vegetarian options.

<http://wedofood.dk>

Halmstorvet 21 Mon–Sat 10:00–21:00
1700 Copenhagen V Sun 10:00–20:00

Warpigs 0.3 km \$\$

One word: pork. Great craft beer.

<http://warpigs.dk>

Flæsketorvet 25 Mon–Thu 11:30–00:00
1711 Copenhagen V Fri–Sat 11:00–02:00

Culinary Highlights in the City

Brør 1.3 km \$\$\$

High-end dining, comes with reasonably priced wine pairings. More expensive, but worth it.

<http://www.restaurantbror.dk>

Sankt Peders Stræde 24A Wed–Sun 17:30–00:00
1453 Copenhagen K

Geist 2.3 km \$\$

Super-minimalist nordic cooking. The menu describes exactly what you get ("Carrots, braised in orange juice, with ginger"), but whatever it is, it's cooked to perfection! You order several small plates, each reasonably priced, but it adds up. You can watch the busy, but eerily quiet kitchen, lit only by candles. Fancy cocktail bar, too, albeit a bit short on classics.

<http://restaurantgeist.dk>

Kongens Nytorv 8 All days 12:00–15:00 and 17:30–01:00
1050 Copenhagen K

Gran Torino 2.3 km \$

Part of the Madklubben group. Set meals (from 200 DKK) includes pasta, pizza and extra nice Tiramisu.

Local Guide

<http://madklubben.dk/gran-torino/>

Sortedam Dossering 5
2200 Copenhagen N

Mon–Fri 17:30–00:00
Sun 17:30–22:00

Höst

1.6 km \$\$\$

Get a fixed price menu with lots of interesting culinary and visual effects (clams on burning juniper bushes). Pricey, but worth the money. Throw in an extra few hundred for the wine pairing, and you won't be disappointed.

<http://hostvakst.dk/host/restaurant/?lang=en>

Nørre Farimagsgade 41
1364 Copenhagen K

All days 17:30–00:00

Morgenstedet

3.3 km \$

Vegetarian Restaurant in the heart of Christiania. Has an improvised canteen feel to it, but very tasty food.

Fabriksområdet 134
1440 Copenhagen K

Tue–Sun 12:00–21:00

Papirøen

3.3 km \$\$

A collection of street food vendors in one old factory building, ranging from duck-fat fries to Asian noodle salads. And plenty of drinks. If the weather is nice, you can sit in deck chairs and watch the boats on the canal. Cheap to mid-range depending on the vendor.

<http://copenhagenstreetfood.dk/en/>

Trangravsvej 14, hal 7/8
1436 Copenhagen K

Mon–Thu 12:00–21:00

Fri–Sun 12:00–21:00

Ramen to Biiru

1.4 km \$

Japanese ramen meets Danish craft beer.

<http://ramentobiiru.dk/vesterbro/>

Enghavevej 58
1674 Copenhagen V

Mon–Thu 12:00–22:00 Sun 12:00–21:00

Fri–Sat 12:00–23:00

SimpleRAW

1.7 km \$

Rawfood. Vegan and vegetarian.

<https://www.simpleraw.dk>

Gråbrødre Torv 9
1154 Copenhagen K

Mon–Sat 10:00–22:00

Sun 10:00–20:00

Cafés and Bars

Bang og Jensen**0.9 km**

At the far end of Istedgade, both a cafe and a bar, and extremely hyggelig.

<http://www.bangogjensen.dk>

Istedgade 130

Mon-Fri 07:30-02:00 Sun 10:00-00:00

1650 Copenhagen V

Sat 10:00-02:00

Bastard**1.3 km**

Board game café.

<http://bastardcafe.dk>

Rådhusstræde 13

Mon–Thu, Sun 12:00–00:00

1466 Copenhagen K

Fri–Sat 12:00–02:00

Kaffe**0.7 km**

On Istedgade, at the intersection with Skydebaneegade, and easy to miss. Full of wooden trinkets and good coffee. Very small!

<http://kaffeistedgade.dk>

Istedgade 90

Mon–Fri 08:00–22:00

1650 Copenhagen V

Sat–Sun 09:00–22:00

Library Bar**0.6 km**

Next to the train station, the bar of the Plaza Hotel. Plush leather sofas, dark wood panels, and on some nights: live piano and songs. Mixed crowd, go in a suit or with hiking clothes.

<https://ligula.se/en/the-library-bar/>

Bernstorffsgade 4

Mon–Thu 16:00–00:00

1577 Copenhagen V

Fri–Sat 16:00–01:00

Mikkeller**0.4 km**

Slightly overhyped micro-brewery. Some good beers, some misses. If you ever wondered what chocolate-liquorice blueberry porter tastes like: here you might find out. Excellent sausages to go with the beer.

<http://mikkeller.dk/location/mikkeller-bar-viktoriagade-copenhagen/>

Viktoriagade 8

Sun–Wed 13:00–01:00 Sat 12:00–02:00

1655 Copenhagen V

Thu–Fri 13:00–02:00

Paludan**2.3 km**

Sit in an antique book shop and sip a beer. Or a coffee. Very decent snack food (try the charcuterie board). Also a good place to work.

<https://www.paludan-cafe.dk/home-eng>

Local Guide

Fiolstræde 10
1171 Copenhagen K

Mon–Fri 09:00–22:00 Sun 10:00–22:00
Sat 10:00–22:00

Risteriet

0.2 km

Coffee place close to Kødbyen.

<http://www.risteriet.dk/risteriet-halmtorvet/>

Helgolandsgade 21
1700 Copenhagen V

Mon–Fri 07:30–18:00 Sun 09:00–18:00
Sat 09:00–18:00

Sort kaffe og vinyl

0.6 km

Another small coffee place, around the corner from Kaffe. You can also buy old vinyl disks while sipping your latte.

<https://facebook.com/sortkaffeogvinyl/>

Skydebanegade 4
1709 Copenhagen

Mon–Fri 08:00–19:00 Sun 09:00–18:00
Sat 09:00–19:00

Health

Pharmacies are few and far between. If you need one, there is a 24h pharmacy next to the train station on Vesterbrogade 6, 1620 Copenhagen V, 0.6 km from the venue.

For **24h medical advice**, call 1813. They can also help you see a doctor outside regular opening hours.

In **emergency situations**, call 112.



We're a company of pioneers.

It's our job to make bold bets, and we get our energy from inventing on behalf of customers. **Success is measured against the possible, not the probable.** For today's pioneers, that's exactly why there's no place on Earth they'd rather build than Amazon.

Apply today.

We are hiring applied scientists and engineers around the world.

TEAM SPOTLIGHTS

Email your resume to
EMNLP2017@amazon.com
and learn more at
www.amazon.jobs/emnlp

Core Machine Learning
Amazon Music
Alexa
Personalization Sciences
AWS Deep Learning
Amazon Lex
Amazon Polly
Amazon Search/A9
Smart Home
Amazon Go
Advertising Technology
Alexa Shopping
Retail Systems Machine Learning
Customer Service Technology
Lab 126



Θ Siteimprove

—
One of Denmark's Top 10 Best IT Workplaces*

In one of the world's happiest countries

With some of the world's best restaurants

**And as Carlsberg would say –
probably the best colleagues in the world!**

We have a small but dedicated
department committed to being the
best in Natural Language Processing –
are you our next team member?

*2015, 2016 Great Place to Work® Institute

**Stop by our
booth to find
out more!**

content/ads/platinum/baidu.pdf

content/ads/platinum/facebook.pdf

Data Science at Bloomberg

Bloomberg's core product, the Terminal, is a must-have for the most influential people in finance. Bloomberg scientists and engineers develop finance-oriented NLP applications for NER & NED, question answering, sentiment analysis, market impact indicators, social media analysis, topic clustering and classification, recommendation systems and predictive models of market behavior. Our customers rely on this information to make swift financial decisions.

We are hiring in NLP, IR, ML and Data Science

We publish in leading academic venues and contribute to open source

We fund academic research through the Bloomberg Data Science Research Grants program

We welcome faculty visitors (summer and sabbatical)

We offer internships and postdocs

Find out more:

techatbloomberg.com/nlp
github.com/bloomberg

©2017 Bloomberg L.P. S809518567 0817



Bloomberg

Ask us about
our work in:

Data Science Machine Learning and Artificial Intelligence



Machine Translation



Signal Processing



Language Understanding



Information Extraction



Machine Learning



Computer Vision



Behavior Modeling



Data Analytics



Trust & Safety



Economics



Dialog Modeling



Deep Learning



Evolutionary Algorithms

... and so much more!

eBay Inc.
labs.ebay.com



ADVANCE YOUR CAREER



CONNECT WITH US



STAY INFORMED



BOOST YOUR RESEARCH

100110010111
110001101101
100000110011
1111011101
1011011100

Microsoft

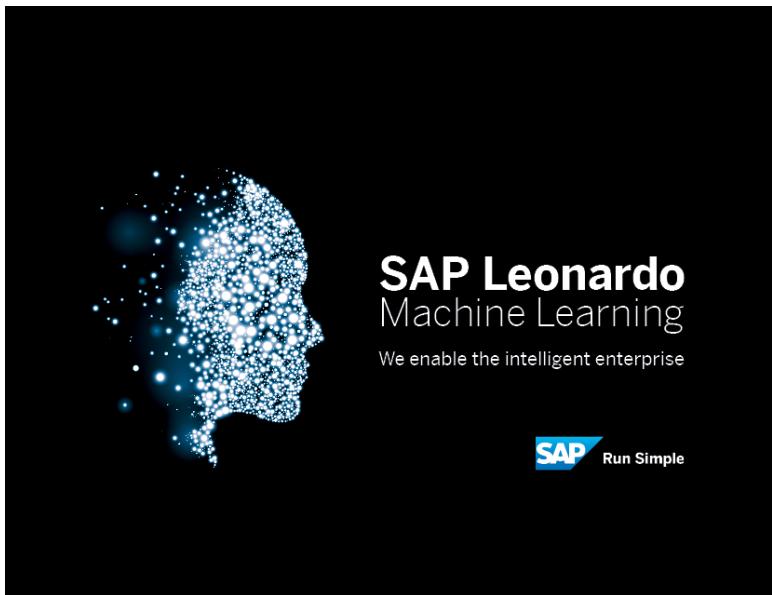
Together, we can achieve more

At Microsoft Research, we're inventing the future of computing. We relentlessly push the boundaries of technology, actively collaborate with world-class researchers, and passionately support the next generation of scientists.

Engage with us: Microsoft.com/research

Microsoft Research

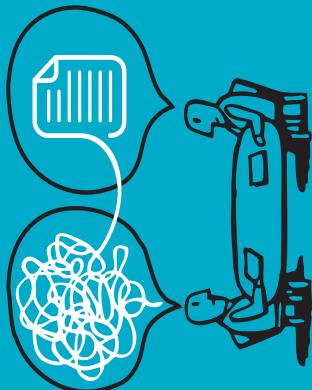
content/ads/gold/ibm.png



textkernel

Innovation, Artificial Intelligence and Machine Learning for matching jobs and people

- Information extraction and retrieval
- Recommendation systems
- Deep Learning and Big Data
- Week long hackathon
- Influence on products



(↑ Our Talent Team rotated this on purpose)

We're hiring!

Visit textkernel.careers or email us at jobs@textkernel.com

content/ads/gold/maluuba.png

Enriching lifestyles with Information Technology

Recruit Institute of Technology (RIT) is the technology hub and research lab for Recruit Holdings, a company that provides over 200 online services in the areas of human resources, travel, housing, restaurants and many other areas in which people make daily lifestyle decisions.

We conduct research in several areas, including natural language processing, machine learning, artificial intelligence, data management, and data integration. We collaborate with universities and publish in top-notch conferences.



Example Projects:

NLP for search engines: RIT is developing natural language processing techniques to improve the understanding of queries that users pose to vertical search engines and the textual content these engines store through conversational interfaces.

Chatbots: we are developing tools for easily building chatbots in vertical domains, including how to pose questions and extract answers and finding natural conversation flows.

Joybot: we are fielding our NLP technology in Joybot, a conversational agent that helps users set goals for increasing their happiness.

444 Castro St Suite 900, Mountain View, CA 94041, USA | www.recruit.ai

content/ads/gold/deloitte.png

#1

Business Software

Complete. Modern. Integrated.

ORACLE®

oracle.com/about
or call 1.800.ORACLE.1

Copyright © 2017, Oracle and/or its affiliates. All rights reserved.



The **Noah's Ark Lab** is a research lab of **Huawei Technologies**, located in Hong Kong and Shenzhen. The mission of the lab is to make significant contributions to both the company and society by innovating in data mining, artificial intelligence, and related fields. Founded in 2012, the lab has now grown to be a research organization with many significant achievements. We welcome talented researchers and engineers to join us to realize their dreams.

research.nuance.com

Inventing a more
human conversation
with technology

To learn more about Nuance's R&D team of talented scientists, linguists and engineers pioneering human-machine intelligence and communication, visit research.nuance.com

 NUANCE

<content/ads/silver/sogou.png>

Shape the future of language learning.

Duolingo scientists and engineers are reinventing how 200 million+ people learn languages around the world.

We're looking for creative ML and NLP researchers to join our innovative, interdisciplinary, and award-winning team!

duolingo.com/jobs

duolingo



MAXHUB

— Efficient Conference Platform

In this diverse and open era MAXHUB breaks the limitation of previous product design logic, develop all-in-one solution of HD display, touch writing, wireless presenting instead single feature, also compatible with different teleconference software& hardware solution and many office applications, aiming to improve meeting efficiency of financial institutions, technology industries, real estate companies, manufacturing industries, consulting services and administrative organizations.



- | | |
|--|------------------------------------|
| Whiteboard fluent writing | Wireless Mirroring |
| Dual pen writing | Voice Assistant |
| scan to take and share meeting minutes | Dual camera |
| Email meeting records | 4K / 1080P high definition display |
| Annotate in any input channel | High performance modular design |

For more information, <http://www.maxhub.vip>

* Different types have different functions.



content/ads/silver/trustpilot.png

content/ads/silver/wizkids.png

Your AI Startup Starts Here.

NextAI is a startup accelerator located in Toronto, Canada. NextAI is for entrepreneurs, researchers and scientists looking to launch AI-enabled ventures. Come build your company in one of the global hotspots for AI research and commercialization.



Come build your
AI company in
Canada!

- Up to \$200K in funding
- Workspace in Toronto
- Award-winning technical & business faculty
- Access to corporate mentorship
- World leading technology platforms & scientists
- Startup Visa designated entity

Applications open on September 20th, 2017

For more information and to apply, visit: nextai.com

NEXT AI

content/ads/bronze/snapchat.png

G grammarly



grammarly.com/jobs

- We help the world's 2B English speakers make their communication clear, effective, and error-free
- We are growing fast, with 10M+ active users in Chrome and volumes of data
- We use a variety of techniques including deep learning to push the boundaries of writing assistance

content/ads/bronze/yandex.png

The EMNLP organizers gratefully acknowledge the support from the following sponsors.

Platinum



Gold



Machine Intelligence for People and Jobs



A Microsoft company



Silver



Bronze

