(LREC) – COLING – ACL 2018

Udo Hahn

Statistics

	LREC 2018	COLING 2018	ACL 2018
Submissions	1102	1.017 (129 withdrawn!)	1.621 (L: 1045 | S: 576)
Accepted papers	718	331	L: 256/1018 | S: 125/526 = 381
Acceptance ratio	**65.0%** (NC: „inclusive“)	**37,3%**	**24.7%**
# reviewers	1.273	1.029	1.610
# participants	~1200	~700	~1500
Ranking	B	A – [→ B+]	A+
Our contributions	Buechel & Hahn (main)	Buechel & Hahn (main)	ECONLP Workshop (org)
	Lohr et al. (main)	Hellrich et al. (demo)	ECONLP Workshop: Buechel et al. (short)

COLING 2018 – Santa Fé, New Mexico, USA

84.000 Einwohner, 2200m Höhe (!)

COLING 2018 - Tutorials

- *NLP for Conversations: Sentiment, Summarization, and Group Dynamics – Gabriel Murray, Giuseppe Carenini and Shafiq Joty*
- *Deep Bayesian Learning and Understanding – Jen-Tzung Chien*
- *Deep Learning for Dialogue Systems – Yun-Nung Chen, Asli Celikyilmaz and Dilek Hakkani-Tur*
- Practical Parsing for Downstream Applications – Daniel Dakota and Sandra Kübler
- Frame Semantics across Languages: Towards a Multilingual FrameNet – Collin Baker, Michael Ellsworth, Miriam Petruck and Swabha Swayamdipta
- Data-Driven Text Simplification – Sanja Štajner and Horacio Saggion

COLING 2018 – Major Themes

- **Sentiment + Emotion Analysis (I-IV), Humor/Sarcasm/Rumor (I-II)**
- **Embeddings (I-III), Distributional Semantics**
- Machine Learning (I-II)
- Parsing (I-II), Generation (I-II)
- Coreference, Discourse (I-II)
- Low-Resource Languages
- Ethics
- Historical Linguistics (L. change)
- **Named Entity Recognition (I-II), Information Extraction (I-II)**
- **Dialog Systems (I-II), QA (I-II)**
- **Machine Translation (I-III)**
- Summarization
- Multimodality

COLING 2018 - Keynotes

- James Pustejovsky: Visualizing Meaning – Modeling Communication through Multimodal Simulations
  - Human-computer/human robot interaction
  - Contextual / situated grounding (nice VoxWorld video, blocks world: „grasp a cup“)
  - VoxML: modeling language for constructing contextualized 3D virtual realization
  - Multiple logics combined
- Hannah Rohde: Why are you telling me this? Relevance and informativity in language processing
  - Inference of relevance and processing of (un)informative information (from a human/cognitive) perspective
  - Discourse coherence relations: implicit (inferable) vs. explicit relations – psycholinguistic experiments (insertion of appropriate adverbials into blank text, given discourse context); self-paced reading time experiments to determine inference for relevance and redundancy for (lack of) informativeness, e.g., color relative to pieces of clothing and fruits
- Min-Yen Kan: Research – fast and slow
  - Acceleration of science: arXiv, GitHub /Jupyter Notebook, Shared Tasks
  - Research fast: works but we don’t know why; research fast+slow: it works and we think we know why and we’ll advocate for it
- Fabiola Henri: Investigating a discriminative approach to creolization
  - Standard claim: creole language (simplifications of European languages) do not have morphology – yet, they have one (though simpler) [25/47 slides shown]

COLING 2018 – Interesting Stuff & Observations

- Andrey Kutuzov, Lilja Øvrelid, Terrence Szymanski, Erik Velldal: Diachronic word embeddings and semantic shifts: a survey
- Andrew Moore and Paul Rayson: Bringing replication and reproduction together with generalisability in NLP: Three reproduction studies for Target Dependent Sentiment Analysis
  - Huge differences between original publication and replication studies when domains and genres are changed
- INCEption annotation tool (TU Darmstadt)

COLING 2018 – Interesting Stuff & Observations

- Zied Bouraoui, Shoaib Jameel, Steven Schockaert: Relation Induction in Word Embeddings Revisited
  - Cosine-based word similarity gives counter-intuitive results: (horses,horse) vs. (berlin, germany)
  - Solution combines Bayesian linear regression similarity between s and t and high probability of s-t, s and t being in a semantic relation
- Alan Akbik, Duncan Blythe, Roland Vollgraf: Contextual String Embeddings for Sequence Labeling
  - Zalando WE framework (flair) increases effectiveness (WE and system are being distributed)
  - Problems: Word ambiguity solved by contextualizing embeddings; fixed vocabulary -> meaningful embeddings for any type of word → propose contextual string embeddings (character-based)!
  - Contextualization is realized by combining embeddings from left2right and right2left: outperforms word-level embeddings; but concatenate string embeddings with standard word embeddings! Very good results for NE tasks (English, German) – move from word-level to char-level language modelling
- Jie Yang, Shuailong Liang, Yue Zhang: Design Challenges and Misconceptions in Neural Sequence Labeling
  - Nice survey of problems of reproducibility of sequence labeling tasks (POS, NER, chunking) → common framework for comparison (hyperparameters, data sets, etc.): NCRF++ (GitHub)

COLING 2018 – Workshop on Trolling, Aggression and Cyberbullying

- Keynote Talk – Rada Mihalcea: What Hides Behind Online Identity (and Anonymity)
  - Cyber-self often different from real self: live 24-7: addictive behavior, fear of missing out, perfect (eating disorders, low self-esteem), part of unnaturally large groups (isolation, negative relations), different self (bullying, prediation) → identity deception (pretending to be someone you are not)
  - Linguistic style different when engaging in identity deception: age, binary gender
  - New identity deception dataset (AMT): open ended questions (600 individuals) related to portray one of 4 fake identities (18/65 aged : m/f) using real and fake identity (imagine you are 18/65:f/m) – can we recognize instances of identity deception? (e.g., female claiming to be male) : result: 85.8 (using SVM) on all identities
  - Finding gender deceivers (predicting gender deception) : 86% men lying (being females) are harder to identify than women pretending to be men
  - Finding age deceivers (predicting age deception): 82.7% (old people are better to deceive than young ones)
  - How to spot a liar (related to gender and age)? LIWC, n-grams, word embeddings
  - Lying about others (fake news) not about themselves: celebrity and political domains (ngrams, punctuation, LIWC, readability, syntax as features) – ngrams, readability but particularly LIWC are performing well
  - Human vs. automatic fake news detection: 74-76% (system, 70-80% (humans)
  - More than just language … video clips from 121real trials → gesture annotation

COLING 2018 – Workshop on Trolling, Aggression and Cyberbullying

- RiTUAL-UH at TRAC 2018 Shared Task: Aggression Identification: Niloofar Safi Samghabadi, Deepthi Mave, Sudipta Kar and Thamar Solorio
  - 130 registered, 30 submitted, 25 papers
  - Annotated dataset of 2 languages: Hindi (1200 tweets) & English (1269 tweets): 3 classes –overly aggressive, covertly aggressive, non-aggressive, posts from Facebook pages (political groups and news)
  - Methods: LSTM, RNN (2nd best used SVM with char and word-level n-grams) – danger of overfitting when surprise dataset had to be analyzed (particularly for Hindi)
  - Summary: NN approaches not necessarily better than standard ML; data augmentation (translation, pseudolabelling) using other resources is effective; careful preprocessing pays off
  - Semeval 2019: task 6 – HateEval and OffensEval (started now), training at end of 2018, testing begin of 2019

COLING 2018 – LaTech Workshop

- Keynote Talk by Ted Underwood: Measurement and Human Perspective
  - Perspectival knowledge – crucial for qualitative vs. quantitative approaches to (digital) humanities
  - Bayesian approaches can take this into account

COLING 2018 – German Impact

- TU Darmstadt [6]
  - Ilia Kuznetsov and Iryna Gurevych - From Text to Lexicon: Bridging the Gap between Word Embeddings and Lexical Resources
  - Steffen Eger, Johannes Daxenberger, Christian Stab and Iryna Gurevych – Cross-lingual Argumentation Mining: Machine Translation (and a bit of Projection) is All You Need!
  - Erik-Lân Do Dinh, Steffen Eger and Iryna Gurevych – Killing Four Birds with Two Stones: Multi-Task Learning for Non-Literal Language Detection
  - Andreas Hanselowski, Avinesh PVS, Benjamin Schiller, Felix Caspelherr, Debanjan Chaudhuri, Christian M. Meyer and Iryna Gurevych – A Retrospective Analysis of the Fake News Challenge Stance-Detection Task
  - Lisa Beinborn, Teresa Botschen and Iryna Gurevych - Multimodal Grounding for Language Processing
  - Daniil Sorokin and Iryna Gurevych – Modeling Semantics with Gated Graph Neural Networks for Knowledge Base Question Answering
  - Jan-Christoph Klie, Michael Bugert, Beto Boullosa, Richard Eckart de Castilho & Iryna Gurevych – The INCEpTION Platform: Machine-Assisted and Knowledge-Oriented Interactive Annotation [demo]
- U Stuttgart [4]
  - Jeremy Barnes, Roman Klinger and Sabine Schulte im Walde – Projecting Embeddings for Domain Adaption: Joint Modeling of Sentiment Analysis in Diverse Domains
  - Evgeny Kim and Roman Klinger – Who Feels What and Why? Annotation of a Literature Corpus with Semantic Roles of Emotions
  - Laura Ana Maria Bostan and Roman Klinger - An Analysis of Annotated Corpora for Emotion Classification in Text
  - Ina Roesiger, Arndt Riester and Jonas Kuhn – Bridging resolution: Task definition, corpus resources and rule-based experiments
  - Markus Gärtner, Sven Mayer, Valentin Schwind, Eric Hämmerle, Emine Turcan, Florin Rheinwald, Gustav Murawski, Lars Lischke & Jonas Kuhn – NLATool: an Application for Enhanced Deep Text Understanding [demo]
- Leibniz ScienceCampus, Heidelberg/Mannheim [2,5]
  - Ines Rehbein and Josef Ruppenhofer – Sprucing up the trees – Error detection in treebanks
  - Josef Ruppenhofer, Michael Wiegand, Rebecca Wilm and Katja Markert – Distinguishing affixoid formations from compounds
  - Marc Schulder, Michael Wiegand and Josef Ruppenhofer – Automatically Creating a Lexicon of Verbal Polarity Shifters: Mono- and Crosslingual Methods for German

COLING 2018 – German Impact

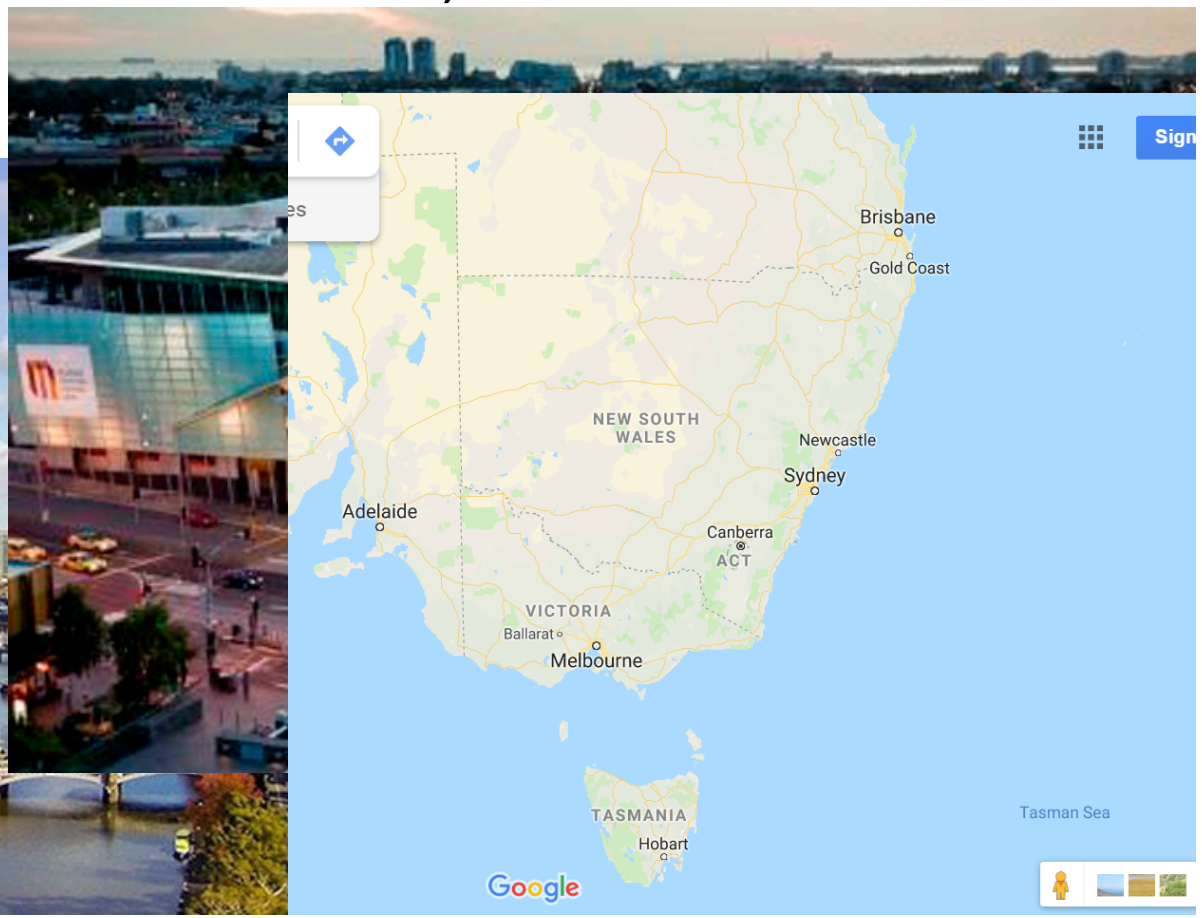
- U Tübingen [2]
  - Zarah Weiß and Detmar Meurers – Modeling the Readability of German Targeting Adults and Children: An empirically broad analysis and its cross-corpus validation
  - Johannes Dellert – Combining Information-Weighted Sequence Alignment and Sound Correspondence Models for Improved Cognate Detection
- U Hamburg [2]
  - Seid Muhie Yimam and Chris Biemann – Par4Sim: Adaptive Paraphrasing for Text Simplification
  - Arne Köhn – Incremental Natural Language Processing: Challenges, Strategies, and Evaluation
- U Passau [2]
  - Matthias Cetto, Christina Niklaus, André Freitas and Siegfried Handschuh – Graphene: Semantically-Linked Propositions in Open Information Extraction
  - Christina Niklaus, Matthias Cetto, André Freitas and Siegfried Handschuh – A Survey on Open Information Extraction
  - Matthias Cetto, Christina Niklaus, André Freitas & Siegfried Handschuh – Graphene: a Context-Preserving Open Information Extraction System [demo]
- BU Weimar [2]
  - Martin Potthast, Tim Gollub, Kristof Komlossy, Sebastian Schuster, Matti Wiegmann, Erika Patricia Garces Fernandez, Matthias Hagen and Benno Stein – Crowdsourcing a Large Corpus of Clickbait on Twitter
  - Henning Wachsmuth, Manfred Stede, Roxanne El Baff, Khalid Al Khatib, Maria Skeppstedt and Benno Stein – Argumentation Synthesis following Rhetorical Strategies
- FSU Jena [1,5]
  - Sven Buechel and Udo Hahn – Emotion Representation Mapping for Automatic Lexicon Construction (Mostly) Performs on Human Level
  - Johannes Hellrich, Sven Buechel & Udo Hahn – JeSemE: Interleaving Semantics and Emotions in a Web Service for the Exploration of Language Change Phenomena [demo]

COLING 2018 – German Impact

- U Duisburg-Essen
  - Sebastian Dungs, Ahmet Aker, Norbert Fuhr and Kalina Bontcheva – Can Rumour Stance Alone Predict Veracity?
- U München
  - Wenpeng Yin, Yadollah Yaghoobzadeh and Hinrich Schütze – Recurrent One-Hop Predictions for Reasoning over Knowledge Graphs
- U des Saarlandes, Spoken Language Systems
  - Marc Schulder, Michael Wiegand and Josef Ruppenhofer – Automatically Creating a Lexicon of Verbal Polarity Shifters: Mono- and Crosslingual Methods for German
- University of Mannheim
  - Sanja ˇStajner and Ioana Hulpus – Automatic Assessment of Conceptual Text Complexity Using Knowledge Graphs
- TU Dresden
  - Kilian Gebhardt – Generic refinement of expressive grammar formalisms with an application to discontinuous constituent parsing
- FU Berlin, Department of Literary Studies
  - Timo Baumann, Hussein Hussein and Burkhard Meyer-Sickendiek – Style Detection for Free Verse Poetry from Text and Speech
- [Fach-]Hochschule Hannover
  - Jean Charbonnier and Christian Wartena – Using Word Embeddings for Unsupervised Acronym Disambiguation
- Zalando Research, Berlin
  - Alan Akbik, Duncan Blythe and Roland Vollgraf – Contextual String Embeddings for Sequence Labeling
- U DÜsseldorf
  - Andreas van Cranenburgh – Active DOP: an Active Learning Constituency Treebank Annotation Tool [demo]
- HITS Heidelberg
  - Mark-Christoph Müller | Michael Strube – Transparent, Efficient, and Robust Word Embedding Access with WOMBAT [demo]
- KIT Karlsruhe
  - Florian Dessloch, Thanh-Le Ha, Markus Müller, Jan Niehues, Thai Son Nguyen, Ngoc-Quan Pham, Elizabeth Salesky, Matthias Sperber, Sebastian Stüker, Thomas Zenkel & Alexander Waibel – KIT Lecture Translator: Multilingual Speech Translation with One-Shot Learning [demo]
- JvGU Frankfurt/Main
  - Daniel Baumartz | Tolga Uslu | Alexander Mehler – LTV: Labeled Topic Vector [demo]

ACL 2018 – Melbourne, Australia

5M Einwohner (Larger M., 31m Höhe)



ACL 2018

- Huge growth in membership (3000++)
- Asian & Pacific Chapter of the ACL founded (AACL)
- Call for papers for a special issue of Computational Linguistics on "Computational approaches in historical linguistics after the quantitative turn" – typological (phylogenetic) approaches sought after (MPI)

ACL 2018 - Tutorials

- *100 Things You Always Wanted to Know about Semantics & Pragmatics But Were Afraid to Ask* – Emily M. Bender
- *Neural Approaches to Conversational AI* – Jianfeng Gao,Michel Galley, and Lihong Li
- *Variational Inference and Deep Generative Models* – Wilker Aziz and Philip Schulz
- *Connecting Language and Vision to Actions* – Peter Anderson, Abhishek Das, and QiWu
- Beyond Multiword Expressions: Processing Idioms and Metaphors – Valia Kordoni
- *Neural Semantic Parsing* – Luke Zettlemoyer,Matt Gardner, Pradeep Dasigi, Srinivasan Iyer, and Alane Suhr
- *Deep Reinforcement Learning for NLP* – William YangWang, Jiwei Li, and Xiaodong He
- *Multi-lingual Entity Discovery and Linking* – Avirup Sil, Heng Ji, Dan Roth, and Silviu-Petru Cucerzan

ACL 2018 – Major Themes

- **Machine Learning (I-III)**
- **Semantics (I-II), Word Semantics (I-II)**
- **Parsing (I-II), Semantic Parsing (I-III), Generation (I-II)**
- **Discourse (I-II)**
- Inference & Reasoning
- Resources, Annotation
- Language & Document Models

- **Machine Translation (I-II), Multilinguality (I-II)**
- **Information Extraction (I-III), Text Mining, Argument Mining**
- **Question Answering (I-III), Dialog Systems & Discourse (I-III)**
- **Sentiment (I-II), Social Media**
- Summarization (I-II)
- Evaluation
- Vision
- Multimodality
- Information Retrieval
- Psycholinguistics & Cognitive Modeling

ACL 2018 - Keynotes

- Anton van den Hengel: Deep Neural Network, and some things they're not very good at
  - 110 ML persons @ Australian Institute for Machine Learning
  - Vision: benchmark driven DL – what if problem, data or information changes
  - Visual Question Answering (VQA) – given pictures , ask questions about their contents, get answers (training data: images, questions, answers) requires explicitly coded knowledge (e.g., DBPedia)
  - Solving the learning problem: Neural Turing Machine (computes reasons for answers) by coupling image recognition and NL – attention as reasoning (attention = associative chaining?) – no fixed ontology (Visual Question Answering-CVVP17)
- Lifetime Achievement Award: Mark Steedman
  - History of Combinatory Categorial Grammar (CCG) – why is language combinatory? – from Lambda calculus to statistical CCG ; CCG in the Age of DL
  - Semantics matters more ! which semantic representation – unsolved problem?

ACL 2018 - Keynotes

- Carolyn Penstein Rosé : Who is the Bridge Between the What and the How ?
  - Distinguish the what of language, namely its propositional content, and the how of language, or its form, style, or framing.
  - What bridges between these realms are social processes that motivate the linguistic choices that result in specific realizations of propositional content situated within social interactions, designed to achieve social goals.
  - this talk probes into a specific quality of discussion referred to as *Transactivity*, the extent to which a contribution articulates the reasoning of the speaker, that of an interlocutor, and the relation between them (associated with solidarity, influence, expertise transfer, and learning)

ACL 2018 – Interesting Stuff & Observations

- Significance testing & NLP
  - The Hitchhiker's Guide to Testing Statistical Significance in Natural Language Processing: Dror, Baumer, Shlomov, and Reichart
    - Only 1/5 of NLP papers use correct statistical testing – survey of testing problems
    - Decision procedure for selection f proper tests: If (distribution is known): select parametric test; IF dataset is small (..) …
  - Replicability Analysis for Natural Language Processing: Testing Significance with Multiple Datasets: Dror, Baumer, Bogomolov, and Reichart [TACL]
    - Look at type I/II errors and overall type error
    - Rather than showing huge tables better specify k (Fisher or Bonferroni) and give p-values for these k cases
- Small to huge data set generalizations
  - Predicting accuracy on large datasets from smaller pilot data: Johnson, Anderson, Dras, and Steedman
  - More data = better accuracy, higher quality produces better accuracy, but how much data are needed (engineer question) → statistical power analysis (Cohen, 1992) : how large must a data set be to guarantee statistical significance
  - Idea: extrapolate performance from small pilot data to predict performance on much larger data – provide 9 extrapolation methods (Power law, inverse square-root, biased power law, etc.) on 8 text corpora using FastText classifier; hyper-parameter estimation for each test configuration!

ACL 2018 – Interesting Stuff & Observations

- Scalability of named entity types
  - 10k types
  - Learning NE types hierarchies
  - *Murty, Verga, Vilnis, Radovanovic, and McCallum: Hierarchical Losses and New Resources for Fine-grained Entity Typing and Linking*
  - *Choi, Levy, Choi, and Zettlemoyer: Ultra-Fine Entity Typing* [medical appl.]
- Subjective embeddings
  - Searching for the X-Factor: Exploring corpus subjectivity for word embeddings: Tkachenko, Chia, and Lauw
  - Wikipedia (neutral, no bias) vs. Amazon reviews (subjective, opinionated) → objective vs. subjective embeddings;
  - 3 classifications tasks: sentiment, subjectivity and topic classification – similar performance of objective/subjective embeddings on objective tasks but subjective embeddings better for subjective genres → more information in subjective embeddings → sentiment vectors (SentiVec) = Word2Vec + Lexical Resource

ACL 2018 – Interesting Stuff & Observations

- Reading comprehension (machine reading)
  - The NarrativeQA Reading Comprehension Challenge: Koˇciský, Schwarz, Blunsom, Dyer, Hermann, Melis, and Grefenstette
  - Narrative – construct answers that cannot be read off from the text (requires reasoning and context), taken from books and movie scripts
  - Why, how, list questions answered
- Language understanding in the large
  - Whodunnit? Crime Drama as a Case for Natural Language Understanding: Frermann, Cohen, and Lapata
  - Whodunnit – CSI (Crime Scene Investigation) ran for 15 seasons (337 episodes → lots of data), each episode 40-64 minutes. Video + audio + textual data
  - Whodunit phrased as a sequence labeling problem → LSTM detective

ACL 2018 – BioNLP Workshop

- 55 participants
- New resources:
  - AllenNLP @ github (Allen AI)
  - CLEW – Clinical Language Engineering Workbench (Jon Griffith) @ github
- Sub-word information in pre-trained biomedical word representations: evaluation and hyper-parameter optimization: Dieter Galea, Ivan Laponogov, and Kirill Veselkov
  - Word2vec vs. fastText comparison ; optimizing hyper-parameters can get performance improvements comparable to latest architectures – look at characters!
- Invited Presentation: "A Corpus with Multi-Level Annotations of Patients, Interventions and Outcomes to Support Language Processing for Medical Literature"– Ben Nye
  - Support for evidence-based medicine, improve accessibility of medical literature : 5k abstracts from PubMed, annotation by crowdworkers (filtering very important!); questions decomposed into PICO components (Participant/Problem, Intervention, Comparator, Outcomes) → increasingly more detailed info is annotated

ACL 2018 – EcoNLP Workshop

- ~40 participants
- Adobe, Zalando, French investment company
- Causality Analysis of Twitter Sentiments and StockMarket Returns – Narges Tabari, Piyusha Biswas, Bhanu Praneeth, Armin Seyeditabari, Mirsad Hadzikadic, andWlodek Zadrozny
  - Causality analysis of Twitter based on Granger causality
  - Crowdworkers judge sentiment of tweets (-2 to +2, neg/pos) related to companies → 2000 negs/8000 pos tweets, 9000 neutrals
  - used pos/neg dictionary (Loughran)
  - Random Forest and SVM classifiers: 80% F
  - Test whether *senti (score) causes stock market return* or stock return causes sentiment (score)

ACL 2018 – German Impact

- Co-Chair: Iryna Gurevych, TU Darmstadt, Germany
- U des Saarlandes (CI, MPI)
  - Jonas Groschwitz, Matthias Lindemann, Meaghan Fowlie, Mark Johnson & Alexander Koller – AMR dependency parsing with a typed semantic algebra [L]
  - Dominic Seyler, Tatiana Dembelova, Luciano Del Corro, Johannes Hoffart & Gerhard Weikum – A Study of the Importance of External Knowledge in the Named Entity Recognition Task [S]
  - Stefan Grünewald, Sophie Henning & Alexander Koller – Generalized chart constraints for efficient PCFG and TAG parsing [S]
  - Prabal Agarwal, Jannik Strötgen, Luciano Del Corro, Johannes Hoffart & Gerhard Weikum – diaNED: Time-Aware Named Entity Disambiguation for Diachronic Corpora [S]
- LMU München
  - Nina Poerner, Hinrich Schütze & Benjamin Roth – Evaluating neural network explanation methods using hybrid documents and morphosyntactic agreement [L]
  - Viktor Hangya, Fabienne Braune, Alexander Fraser & Hinrich Schütze – Two Methods for Domain Adaptation of Bilingual Tasks: Delightfully Simple and Broadly Applicable [L]
  - Philipp Dufter, Mengjie Zhao, Martin Schmitt, Alexander Fraser & Hinrich Schütze – Embedding Learning Through Multilingual Concept Induction [L]
  - Wenpeng Yin, Dan Roth & Hinrich Schütze – End-Task Oriented Textual Entailment via Deep Explorations of Inter-Sentence Interactions [S]
- U Heidelberg
  - Todor Mihaylov & Anette Frank – Knowledgeable Reader: Enhancing Cloze-Style Reading Comprehension with External Commonsense Knowledge [L]
  - Julia Kreutzer, Joshua Uyheng & Stefan Riezler – Reliability and Learnability of Human Bandit Feedback for Sequence-to-Sequence Reinforcement Learning [L]
  - Carolin Lawrence & Stefan Riezler – Improving a Neural Semantic Parser by Counterfactual Learning from Human Bandit Feedback [L]

ACL 2018 – German Impact

- U Stuttgart
  - Jeremy Barnes | Roman Klinger | Sabine Schulte im Walde – Bilingual Sentiment Embeddings: Joint Projection of Sentiment Across Languages [L]
  - Martin Riedl & Sebastian Padó – A Named Entity Recognition Shootout for German [S]
- BU Weimar
  - Henning Wachsmuth, Shahbaz Syed & Benno Stein – Retrieval of the Best Counterargument without Prior Topic Knowledge [L]
  - Khalid Al Khatib, Henning Wachsmuth, Kevin Lang, Jakob Herpel, Matthias Hagen & Benno Stein – Modeling Deliberative Argumentation Strategies on Wikipedia [L]
- RWTH Aachen
  - Weiyue Wang, Derui Zhu, Tamer Alkhouli, Zixuan Gan & Hermann Ney – Neural Hidden Markov Model for Machine Translation [S]
  - Albert Zeyer, Tamer Alkhouli & Hermann Ney – RETURNN as a Generic Flexible Neural Toolkit with Application to Translation and Speech Recognition [demo]
- U Mannheim + U Hamburg
  - Dmitry Ustalov, Alexander Panchenko, Andrey Kutuzov, Chris Biemann & Simone Paolo Ponzetto – Unsupervised Semantic Frame Induction using Triclustering [S]
- Facebook Research
  - Holger Schwenk – Filtering and Mining Parallel Data in a Joint Multilingual Space [S]
- U Bielefeld
  - Matthias Hartung, Hendrik ter Horst, Frank Grimm, Tim Diekmann, Roman Klinger & Philipp Cimiano – SANTO: A Web-based Annotation Tool for Ontology-driven Slot Filling [demo]

2018 Germany‘s Top Performers

	ACL 2018	COLING 2018	Σ
U d. Saarlandes	4	1	5
LMU München	3,5	1	4,5
U Stuttgart	2	4,5	6,5
TU Darmstadt	–	6,5	6,5
BU Weimar	2	2	4
U Heidelberg	3	–	3
RWTH Aachen	1,5	–	1,5
U Hamburg	0,5	2	2,5
LeibnizCampus MA/HD	–	2,5	2,5
U Passau	–	2,5	2,5
U Tübingen	–	2	2
U Mannheim	0,5	1	1,5
FSU Jena	–	1,5	1,5

	ACL 2018	COLING 2018	Σ
U Bielefeld	1	–	1
Facebook Res.	1	–	1
U Duisburg-Essen	–	1	1
TU Dresden	–	1	1
FU Berlin	–	1	1
FH Hannover	–	1	1
Zalando	–	1	1
U Düsseldorf	–	0,5	0,5
HITS Heidelberg	–	0,5	0,5
KIT Karlsruhe	–	0,5	0,5
JXvG Frankfurt/Main	–	0,5	0,5

Over-all Summary

- Language understanding is back on the scene (after more than 30 years!) – this includes **reasoning** tasks, e.g., for machine reading, vision-based QA
- **Neural** Networks
  - Various models are investigated (GAN, attention-based → context)
  - Incorporating non-textial resources (lexicons/terminologies, knowledge graphs)
  - Transfer learning, multi-task learning
  - interpretability
- Methodology underlying **statistical** analysis is criticially reflected