Second Call for Papers

The 14th Workshop on Innovative Use of
NLP for Building Educational Applications (BEA14)
Florence, Italy; August 2, 2019
(co-located with ACL)

https://sig-edu.org/bea/current

Submission Deadline: Friday, April 26, 2019, 11:59pm EST

WORKSHOP DESCRIPTION

The BEA Workshop is a leading venue for NLP innovation in the context of educational applications. It is one of the largest one-day workshops in the ACL community with over 80 attendees in the past several years. The growing interest in educational applications and a diverse community of researchers involved resulted in the creation of the [Special Interest Group in Educational Applications (SIGEDU)](https://sig-edu.org/bea/current) in 2017, which currently has 144 members.

The workshop's continuing growth highlights the alignment between societal needs and technological advances. NLP capabilities can now support an array of learning domains, including writing, speaking, reading, science, and mathematics, as well as the related intra-personal (e.g., self-confidence) and inter-personal (e.g., peer collaboration) skills. Within these areas, the community continues to develop and deploy innovative NLP approaches for use in educational settings. In the writing and speech domains, automated writing evaluation (AWE) and speech scoring applications, respectively, are commercially deployed in high-stakes assessment and in instructional contexts (e.g., Massive Open Online Courses (MOOCs) and K-12 classrooms). Commercially-deployed plagiarism detection is also commonly used in both K-12 and higher education settings. For writing, the focus is on innovations that support writing tasks requiring source use, argumentative discourse, and factual content accuracy. For speech, there is an interest in advancing automated scoring to include the evaluation of discourse and content features in responses to spoken assessments. General advances in speech technology have promoted a renewed interest in spoken dialog and multimodal systems for instruction and assessment, for instance, for workplace interviews and simulated teaching environments. The explosive growth of mobile applications for game-based and simulation-based applications is another area where NLP has begun to play a large role, especially for language learning.

NLP for educational applications has gained visibility outside of the NLP community. First, the Hewlett Foundation reached out to the public and private sectors and sponsored two competitions: one for automated essay scoring, and the other for scoring of short

response items. The motivation driving these competitions was to engage the larger scientific community in this enterprise. Learning @ Scale is a relatively new venue for NLP research in education. MOOCs now incorporate AWE systems to manage *several thousand* assignments that may be received during a single MOOC course. MOOCs for Refugees have more recently emerged in response to the current social situations. Courses include language learning, and we can imagine that AWE and other NLP capabilities could support coursework. Another breakthrough for educational applications within the CL community is the presence of a number of shared-task competitions over the past several years – including three shared tasks on grammatical error detection and correction. NLP/Education shared tasks have seen new areas of research, such as the Automated Evaluation of Scientific Writing at BEA 11, Native Language Identification at BEA 12, Second Language Acquisition Modelling at BEA 13, and Complex Word Identification at BEA 13. These competitions increased the visibility of, and interest in, our field.

The 14th BEA workshop will have oral presentation sessions and a large poster session in order to maximize the amount of original work presented. We expect that the workshop will continue to highlight novel technologies and opportunities for educational NLP in English as well as other languages. The workshop will solicit both full papers and short papers for either oral or poster presentation. We will solicit papers that incorporate NLP methods, including, but not limited to: automated scoring of open-ended textual and spoken responses; game-based instruction and assessment; educational data mining; intelligent tutoring; peer review; grammatical error detection and correction; learner cognition; spoken dialog; multimodal applications; tools for teachers and test developers; and use of corpora. Specific topics include:

Automated scoring/evaluation for written student responses (across multiple genres)
- Content analysis for scoring/assessment
- Detection and correction of grammatical and other types of errors
- Argumentation, discourse, sentiment, stylistic analysis, & non-literal language

Intelligent Tutoring (IT), Collaborative Learning Environments
- Educational Data Mining: Collection of user log data from educational applications
- Game-based learning
- Multimodal communication (including dialog systems) between students and computers

Learner cognition
- Assessment of learners' language and cognitive skill levels
- Systems that detect and adapt to learners' cognitive or emotional states
- Tools for learners with special needs

Use of corpora in educational tools
- Data mining of learner and other corpora for tool building
- Annotation standards and schemas / annotator agreement

Tools and applications for classroom teachers and/or test developers
- NLP tools for second and foreign language learners
- Semantic-based access to instructional materials to identify appropriate texts
- Tools that automatically generate test questions

SHARED TASK: Grammatical Error Correction

The BEA14 workshop will also host a shared task on Grammatical Error Correction. Grammatical error correction (GEC) is the task of automatically correcting grammatical errors in text; e.g. [I follows his advices → I followed his advice]. It can be used to not only help language learners improve their writing skills, but also alert native speakers to accidental mistakes or typos.

GEC gained significant attention in the HOO and CoNLL shared tasks between 2011 and 2014 (Dale and Kilgarriff, 2011; Dale et al., 2012; Ng et al., 2013; Ng et al., 2014), but has since become much more difficult to evaluate given a lack of standardised experimental settings. In particular, recent systems have been trained, tuned and tested on different combinations of corpora using different metrics (Yannakoudakis et al., 2017; Chollampatt and Ng, 2018a; Ge et al., 2018; Junczys-Dowmunt et al., 2018). One of the main aims of this proposal is hence to once again provide a platform where different approaches can be evaluated under the same test conditions.

Another significant problem facing the field is that system performance is still primarily benchmarked against the CoNLL-2014 test set, even though this 5-year-old dataset only represents a very narrow domain of first year, South-East Asian undergraduates in Singapore. This means systems have increasingly overfit to a very specific type of English and so do not generalise well to other domains. Our proposal hence introduces a new dataset that represents a much more diverse cross-section of English language domains.

More information can be found on the task webpage:
https://www.cl.cam.ac.uk/research/nl/bea2019st/

The organizing committee for the shared task consists of: Christopher Bryant (University of Cambridge), Mariano Felice (University of Cambridge) and Ted Briscoe (University of Cambridge).

SUBMISSION INFORMATION

We will be using the ACL Submission Guidelines for the BEA14 Workshop this year. Long papers may consist of up to eight (8) pages of content, plus unlimited references; final versions of long papers will be given one additional page of content (up to 9 pages) so that reviewers' comments can be taken into account. Short papers may consist of up to four (4)

pages of content, plus unlimited references. Upon acceptance, short papers will be given five (5) content pages in the proceedings. Authors are encouraged to use this additional page to address reviewers' comments in their final versions.

Papers which describe systems are also invited to give a demo of their system. If you would like to present a demo in addition to presenting the paper, please make sure to select either "full paper + demo" or "short paper + demo" under "Submission Category" in the START submission page.

Previously published papers cannot be accepted. The submissions will be reviewed by the program committee. As reviewing will be blind, please ensure that papers are anonymous. Self-references that reveal the author's identity, e.g., "We previously showed (Smith, 1991) ...", should be avoided. Instead, use citations such as "Smith previously showed (Smith, 1991) ...".

We have also included conflict of interest in the submission form. You should mark all potential reviewers who have been authors on the paper, are from the same research group or institution, or who have seen versions of this paper or discussed it with you.

We will be using the START conference system to manage submissions:
https://www.softconf.com/acl2019/bea/

IMPORTANT DATES

Submission Deadline: Friday, April 26, 2019, 11:59pm EST
Notification of Acceptance: Friday, May 24, 2019
Camera-ready Papers Due: Monday, June 3, 2019
Workshop: Friday, August 2, 2019

ORGANIZING COMMITTEE

- Helen Yannakoudakis, University of Cambridge (primary contact)
- Ekaterina Kochmar, University of Cambridge
- Claudia Leacock, Grammarly
- Nitin Madnani, Educational Testing Service
- Ildikó Pilán, Develop Diverse
- Torsten Zesch, University of Duisburg-Essen

Workshop contact email address: bea.nlp.workshop@gmail.com

PROGRAM COMMITTEE

- Tazin Afrin, University of Pittsburgh
- David Alfter, University of Gothenburg
- Dimitrios Alikaniotis, Grammarly
- Rajendra Banjade, Audible Inc. (an Amazon company)
- Timo Baumann, Carnegie Mellon University
- Lee Becker, Pearson
- Beata Beigman Klebanov, ETS
- Kay Berkling, Cooperation State university Karlsruhe Germany
- Suma Bhat, University of Illinois at Urbana-Champaign
- Sameer Bhatnagar, Polytechnique Montreal
- Joachim Bingel, University of Copenhagen
- Karim Bouzoubaa, Mohammed V University in Rabat
- Chris Brew, Facebook
- Ted Briscoe, University of Cambridge
- Julian Brooke, University of British Columbia
- Dominique Brunato, Institute for Computational Linguistics, ILC-CNR, Pisa, Italy
- Christopher Bryant, University of Cambridge
- Paula Buttery, University of Cambridge
- Andrew Caines, University of Cambridge
- Mei-Hua Chen, Department of Foreign Languages and Literature, Tunghai University
- Martin Chodorow, ETS & City University of New York
- Shamil Chollampatt, National University of Singapore
- Mark Core, University of Southern California
- Vidas Daudaravicius, UAB VTEX
- Kordula De Kuthy, University of Tübingen
- Carrie Demmans Epp, University of Alberta
- Yo Ehara, Faculty of Informatics, Shizuoka Institute of Science and Technology
- Keelan Evanini, Educational Testing Service
- Mariano Felice, University of Cambridge
- Michael Flor, Educational Testing Service
- Thomas François, Université catholique de Louvain
- Michael Gamon, Microsoft Research
- Dipesh Gautam, The University of Memphis
- Sian Gooding, University of Cambridge
- Jonathan Gordon, Vassar College
- Cyril Goutte, National Research Council Canada
- Iryna Gurevych, UKP Lab, TU Darmstadt

- Na-Rae Han, University of Pittsburgh
- Jiangang Hao, Educational Testing Service
- Homa Hashemi, Microsoft
- Trude Heift, Simon Fraser University
- Derrick Higgins, American Family Insurance
- Heiko Holz, LEAD Graduate School & Research Network at the University of Tuebingen
- Andrea Horbach, University Duisburg-Essen
- Chung-Chi Huang, Frostburg State University
- Yi-Ting Huang, Academia Sinica
- Radu Tudor Ionescu, University of Bucharest
- Lifeng Jin, The Ohio State University
- Pamela Jordan, University of Pittsburgh
- Taraka Kasicheyanula, University of Oslo
- Elma Kerz, RWTH Aachen
- Fazel Keshtkar, St. John's University
- Mamoru Komachi, Tokyo Metropolitan University
- Lun-Wei Ku, Academia Sinica
- Ji-Ung Lee, UKP Lab, Technische Universität Darmstadt
- John Lee, City University of Hong Kong
- Lung-Hao Lee, National Central University
- Ben Leong, Educational Testing Service
- James Lester, North Carolina State University
- Chen Liang, Facebook
- Diane Litman, University of pittsburgh
- Yang Liu, Laix
- Peter Ljunglöf, University of Gothenburg
- Anastassia Loukina, ETS
- Xiaofei Lu, Pennsylvania State University
- Luca Lugini, University of Pittsburgh
- Nabin Maharjan, University of Memphis
- Jean Maillard, University of Cambridge
- Shervin Malmasi, Harvard Medical School
- Montse Maritxalar, University of the Basque Country
- Ditty Mathew, IIT Madras
- Julie Medero, Harvey Mudd College
- Beata Megyesi, Uppsala University
- Detmar Meurers, Universität Tübingen
- Elham Mohammadi, CLaC Laboratory, Concordia University
- Maria Moritz, University of Goettingen

- William Murray, Pearson
- Courtney Napoles, Grammarly
- Diane Napolitano, Educational Testing Service
- Hwee Tou Ng, National University of Singapore
- Huy Nguyen, LingoChamp
- Rodney Nielsen, University of North Texas
- Nobal Niraula, Boeing Research and Technology
- Yoo Rhee Oh, Electronics and Telecommunications Research Institute (ETRI)
- Constantin Orasan, University of Wolverhampton
- Ulrike Pado, HFT Stuttgart
- Alexis Palmer, University of North Texas
- Martí Quixal, Universitat Oberta de Catalunya
- Zahra Rahimi, University of Pittsburgh
- Lakshmi Ramachandran, Amazon Search
- Hanumant Redkar, Indian Institute of Technology Bombay
- Marek Rei, University of Cambridge
- Robert Reynolds, Brigham Young University
- Brian Riordan, ETS
- Kat Robb, University of Leeds
- Andrew Rosenberg, Google
- Mark Rosenstein, Pearson
- Alla Rozovskaya, City University of New York
- C. Anton Rytting, University of Maryland
- Keisuke Sakaguchi, Allen Institute for Artificial Intelligence
- Allen Schmaltz, Harvard University
- Mat Schulze, San Diego State University
- Burr Settles, Duolingo
- Serge Sharoff, University of Leeds
- Swapna Somasundaran, ETS
- Richard Sproat, Google
- Helmer Strik, Centre for Language and Speech Technology (CLST), Centre for Language Studies (CLS), Radboud University Nijmegen
- Jan Švec, NTIS, University of West Bohemia
- Anaïs Tack, UCLouvain & KU Leuven
- Yuen-Hsien Tseng, National Taiwan Normal University
- Giulia Venturi, Institute for Computational Linguistics "A. Zampolli", Italy
- Aline Villavicencio, Federal University of Rio Grande do Sul (Brazil) and University of Essex (UK)
- Carl Vogel, Trinity College Dublin
- Elena Volodina, University of Gothenburg

- Shuting Wang, Facebook Inc
- Michael White, The Ohio State University
- Michael Wojatzki, LTL, University of Duisburg-Essen
- Magdalena Wolska, Eberhard Karls Universität Tübingen
- Huichao Xue, LinkedIn
- Victoria Yaneva, National Board of Medical Examiners / University of Wolverhampton
- Zheng Yuan, University of Cambridge
- Marcos Zampieri, University of Wolverhampton (UK)
- Klaus Zechner, ETS
- Fan Zhang, Google
- Haoran Zhang, University of Pittsburgh
- Ramon Ziai, University of Tübingen