EACL 2021

The 16th Conference of the European Chapter of the Association for Computational Linguistics

Proceedings of the System Demonstrations

April 19 - 23, 2021

©2021 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-954085-05-3

Introduction

Welcome to the proceedings of the EACL System Demonstration Sessions. This volume contains the papers of the system demonstrations presented at the 16th conference of the European Chapter of the Association for Computational Linguistics (EACL) on April 19th - April 23rd, 2021.

The EACL 2021 demonstrations track welcomed submissions ranging from early research prototypes to mature production-ready systems. We received 56 submissions this year, of which 39 were selected for inclusion in the program after having been reviewed by two members of the program committee. This year, the EACL demo track received a very large amount of high-quality research papers. We therefore decided to feature an inclusive programme where all excellent publications were chosen to be presented at the meeting. Following current practice, we further asked the programme committee to pay special attention to issues related to ethics. We would like to warmly thank the members of the program committee for their timely help in reviewing the submissions.

Lastly, we thank the many authors that submitted their work to the demonstrations track. Because this year the EACL conference is completely virtual, the demonstration paper talks are therefore pre-recorded and authors are invited to present their work with posters during the two live demo sessions through gather.town.

Dimitra Gkatzia and Djamé Seddah
EACL 2021 Demonstration Track Chairs

Organizing Committee

- **System Demonstration Chairs:**
  Dimitra Gkatzia, Djamé Seddah

- **Program Committee:**
  Omri Abend, Nikolaos Aletras, Shlomo Argamon, Rachel Bawden, Timothée Bernard, Marie Candito, Xavier Carreras, Alberto Cetoli, Jhih-Jie Chen, Manuel R. Ciosici, Miruna-Adriana Clinciu, Raphael Cohen, Michael Desmond, Sourav Dutta, Catherine Finegan-Dollak, Dimitra Gkatzia, Johannes Heinecke, Daniel Hershcovich, David M. Howcroft, Ganesh Jawahar, Giannis Karamanolakis, Rabeeh Karimi Mahabadi, Andrew Kirby, Arne Köhn, Philippe Laban, Dong-Ho Lee, Manling Li, Xuezhe Ma, Prodromos Malakasiotis, Yuval Marton, Ivan Vladimir Meza Ruiz, Khalil Mrini, Pedro Javier Ortiz Suárez, Xutan Peng, Benoît Sagot, Sashank Santhanam, Sebastin Santy, Thomas Scialom, Djamé Seddah, Somayajulu Sripada, Carl Strathearn, Simon Wells, Guillaume Wisniewski, Runxin Xu, Hamada Zahera, Yi Zhang

Table of Contents

Using and comparing Rhetorical Structure Theory parsers with rst-workbench
Arne Neumann 1

SF-QA: Simple and Fair Evaluation Library for Open-domain Question Answering
Xiaopeng Lu, Kyusong Lee and Tiancheng Zhao 7

Finite-state script normalization and processing utilities: The Nisaba Brahmic library
Cibu Johny, Lawrence Wolf-Sonkin, Alexander Gutkin and Brian Roark 14

CovRelex: A COVID-19 Retrieval System with Relation Extraction
Vu Tran, Van-Hien Tran, Phuong Nguyen, Chau Nguyen, Ken Satoh, Yuji Matsumoto and Minh Nguyen 24

MATILDA - Multi-AnnoTator multi-language InteractiveLight-weight Dialogue Annotator
Davide Cucurnia, Nikolai Rozanov, Irene Sucameli, Augusto Ciuffoletti and Maria Simi 32

AnswerQuest: A System for Generating Question-Answer Items from Multi-Paragraph Documents
Melissa Roemmele, Deep Sidhpura, Steve DeNeefe and Ling Tsou 40

T-NER: An All-Round Python Library for Transformer-based Named Entity Recognition
Asahi Ushio and Jose Camacho-Collados 53

Forum 4.0: An Open-Source User Comment Analysis Framework
Marlo Haering, Jakob Smedegaard Andersen, Chris Biemann, Wiebke Loosen, Benjamin Milde, Tim Pietz, Christian Stöcker, Gregor Wiedemann, Olaf Zukunft and Walid Maalej 63

SLTEV: Comprehensive Evaluation of Spoken Language Translation
Ebrahim Ansari, Ondřej Bojar, Barry Haddow and Mohammad Mahmoudi 71

Trankit: A Light-Weight Transformer-based Toolkit for Multilingual Natural Language Processing
Minh Van Nguyen, Viet Dac Lai, Amir Pouran Ben Veyseh and Thien Huu Nguyen 80

DebIE: A Platform for Implicit and Explicit Debiasing of Word Embedding Spaces
Niklas Friedrich, Anne Lauscher, Simone Paolo Ponzetto and Goran Glavaš 91

A Dashboard for Mitigating the COVID-19 Misinfodemic
Zhengyuan Zhu, Kevin Meng, Josue Caraballo, Israa Jaradat, Xiao Shi, Zeyu Zhang, Farahnaz Akrami, Haojin Liao, Fatma Arslan, Damian Jimenez, Mohanmmed Samiul Saeef, Paras Pathak and Chengkai Li 99

EasyTurk: A User-Friendly Interface for High-Quality Linguistic Annotation with Amazon Mechanical Turk
Lorenzo Bocchi, Valentino Frasnelli and Alessio Palmero Aprosio 106

ASAD: Arabic Social media Analytics and unDerstanding
Sabit Hassan, Hamdy Mubarak, Ahmed Abdelali and Kareem Darwish 113

COCO-EX: A Tool for Linking Concepts from Texts to ConceptNet
Maria Becker, Katharina Korfhage and Anette Frank 119

A description and demonstration of SAFAR framework
Karim Bouzoubaa, Younes Jaafar, Driss Namly, Ridouane Tachicart, Rachida Tajmout, hakima khamar, hamid jaafar, Lhoussain aouragh and Abdellah Yousfi 127

InterpreT: An Interactive Visualization Tool for Interpreting Transformers
Vasudev Lal, Arden Ma, Estelle Aflalo, Phillip Howard, Ana Simoes, Daniel Korat, Oren Pereg, Gadi Singer and Moshe Wasserblat 135

Representing ELMo embeddings as two-dimensional text online
Andrey Kutuzov and Elizaveta Kuzmenko 143

LOME: Large Ontology Multilingual Extraction
Patrick Xia, Guanghui Qin, Siddharth Vashishtha, Yunmo Chen, Tongfei Chen, Chandler May, Craig Harman, Kyle Rawlins, Aaron Steven White and Benjamin Van Durme 149

MadDog: A Web-based System for Acronym Identification and Disambiguation
Amir Pouran Ben Veyseh, Franck Dernoncourt, Walter Chang and Thien Huu Nguyen 160

Graph Matching and Graph Rewriting: GREW tools for corpus exploration, maintenance and conversion
Bruno Guillaume 168

Massive Choice, Ample Tasks (MaChAmp): A Toolkit for Multi-task Learning in NLP
Rob van der Goot, Ahmet Üstün, Alan Ramponi, Ibrahim Sharaf and Barbara Plank 176

SCoT: Sense Clustering over Time: a tool for the analysis of lexical change
Christian Haase, Saba Anwar, Seid Muhie Yimam, Alexander Friedrich and Chris Biemann . . . 198

GCM: A Toolkit for Generating Synthetic Code-mixed Text
Mohd Sanad Zaki Rizvi, Anirudh Srinivasan, Tanuja Ganu, Monojit Choudhury and Sunayana Sitaram 205

T2NER: Transformers based Transfer Learning Framework for Named Entity Recognition
Saadullah Amin and Guenter Neumann 212

European Language Grid: A Joint Platform for the European Language Technology Community
Georg Rehm, Stelios Piperidis, Kalina Bontcheva, Jan Hajic, Victoria Arranz, Andrejs Vasiljevs, Gerhard Backfried, Jose Manuel Gomez-Perez, Ulrich Germann, Rémi Calizzano, Nils Feldhus, Stefanie Hegele, Florian Kintzel, Katrin Marheinecke, Julian Moreno-Schneider, Dimitris Galanis, Penny Labropoulou, Miltos Deligiannis, Katerina Gkirtzou, Athanasia Kolovou, Dimitris Gkoumas, Leon Voukoutis, Ian Roberts, Jana Hamrlova, Dusan Varis, Lukas Kacena, Khalid Choukri, Valérie Mapelli, Mickaël Rigault, Julija Melnika, Miro Janosik, Katja Prinz, Andres Garcia-Silva, Cristian Berrio, Ondrej Klejch and Steve Renals 221

A New Surprise Measure for Extracting Interesting Relationships between Persons
Hidetaka Kamigaito, Jingun Kwon, Young-In Song and Manabu Okumura 231

Paladin: an annotation tool based on active and proactive learning
Minh-Quoc Nghiem, Paul Baylis and Sophia Ananiadou 238

Story Centaur: Large Language Model Few Shot Learning as a Creative Writing Tool
Ben Swanson, Kory Mathewson, Ben Pietrzak, Sherol Chen and Monica Dinalescu 244

FrameForm: An Open-source Annotation Interface for FrameNet
Büşra Marşan and Olcay Taner Yıldız 257

OCTIS: Comparing and Optimizing Topic models is Simple!
Silvia Terragni, Elisabetta Fersini, Bruno Giovanni Galuzzi, Pietro Tropeano and ANTONIO CANDELIERI 263

ELITR Multilingual Live Subtitling: Demo and Strategy
Ondřej Bojar, Dominik Macháček, Sangeet Sagar, Otakar Smrž, Jonáš Kratochvíl, Peter Polák, Ebrahim Ansari, Mohammad Mahmoudi, Rishu Kumar, Dario Franceschini, Chiara Canton, Ivan Simonini, Thai-Son Nguyen, Felix Schneider, Sebastian Stüker, Alex Waibel, Barry Haddow, Rico Sennrich and Philip Williams 271

Breaking Writer's Block: Low-cost Fine-tuning of Natural Language Generation Models
Alexandre Duval, Thomas Lamson, Gaël de Léséleuc de Kérouara and Matthias Gallé 278

OPUS-CAT: Desktop NMT with CAT integration and local fine-tuning
Tommi Nieminen 288

Domain Expert Platform for Goal-Oriented Dialog Collection
Didzis Goško, Arturs Znotins, Inguna Skadina, Normunds Gruzitis and Gunta Nešpore-Bērzkalne 295

Which is Better for Deep Learning: Python or MATLAB? Answering Comparative Questions in Natural Language
Viktoriia Chekalina, Alexander Bondarenko, Chris Biemann, Meriem Beloucif, Varvara Logacheva and Alexander Panchenko 302

PunKtuator: A Multilingual Punctuation Restoration System for Spoken and Written Text
Varnith Chordia 312

Conversational Agent for Daily Living Assessment Coaching Demo
Raymond Finzel, Aditya Gaydhani, Sheena Dufresne, Maria Gini and Serguei Pakhomov 321

HULK: An Energy Efficiency Benchmark Platform for Responsible Natural Language Processing
Xiyou Zhou, Zhiyu Chen, Xiaoyong Jin and William Yang Wang 329

Conference Program

Using and comparing Rhetorical Structure Theory parsers with rst-workbench
Arne Neumann

SF-QA: Simple and Fair Evaluation Library for Open-domain Question Answering
Xiaopeng Lu, Kyusong Lee and Tiancheng Zhao

Finite-state script normalization and processing utilities: The Nisaba Brahmic library
Cibu Johny, Lawrence Wolf-Sonkin, Alexander Gutkin and Brian Roark

CovRelex: A COVID-19 Retrieval System with Relation Extraction
Vu Tran, Van-Hien Tran, Phuong Nguyen, Chau Nguyen, Ken Satoh, Yuji Matsumoto and Minh Nguyen

MATILDA - Multi-AnnoTator multi-language InteractiveLight-weight Dialogue Annotator
Davide Cucurnia, Nikolai Rozanov, Irene Sucameli, Augusto Ciuffoletti and Maria Simi

AnswerQuest: A System for Generating Question-Answer Items from Multi-Paragraph Documents
Melissa Roemmele, Deep Sidhpura, Steve DeNeefe and Ling Tsou

T-NER: An All-Round Python Library for Transformer-based Named Entity Recognition
Asahi Ushio and Jose Camacho-Collados

Forum 4.0: An Open-Source User Comment Analysis Framework
Marlo Haering, Jakob Smedegaard Andersen, Chris Biemann, Wiebke Loosen, Benjamin Milde, Tim Pietz, Christian Stöcker, Gregor Wiedemann, Olaf Zukunft and Walid Maalej

SLTEV: Comprehensive Evaluation of Spoken Language Translation
Ebrahim Ansari, Ondřej Bojar, Barry Haddow and Mohammad Mahmoudi

Trankit: A Light-Weight Transformer-based Toolkit for Multilingual Natural Language Processing
Minh Van Nguyen, Viet Dac Lai, Amir Pouran Ben Veyseh and Thien Huu Nguyen

DebIE: A Platform for Implicit and Explicit Debiasing of Word Embedding Spaces
Niklas Friedrich, Anne Lauscher, Simone Paolo Ponzetto and Goran Glavaš

A Dashboard for Mitigating the COVID-19 Misinfodemic
Zhengyuan Zhu, Kevin Meng, Josue Caraballo, Israa Jaradat, Xiao Shi, Zeyu Zhang, Farahnaz Akrami, Haojin Liao, Fatma Arslan, Damian Jimenez, Mohammed Samiul Saeef, Paras Pathak and Chengkai Li

No Day Set (continued)

EasyTurk: A User-Friendly Interface for High-Quality Linguistic Annotation with Amazon Mechanical Turk
Lorenzo Bocchi, Valentino Frasnelli and Alessio Palmero Aprosio

ASAD: Arabic Social media Analytics and unDerstanding
Sabit Hassan, Hamdy Mubarak, Ahmed Abdelali and Kareem Darwish

COCO-EX: A Tool for Linking Concepts from Texts to ConceptNet
Maria Becker, Katharina Korfhage and Anette Frank

A description and demonstration of SAFAR framework
Karim Bouzoubaa, Younes Jaafar, Driss Namly, Ridouane Tachicart, Rachida Tajmout, hakima khamar, hamid jaafar, Lhoussain aouragh and Abdellah Yousfi

InterpreT: An Interactive Visualization Tool for Interpreting Transformers
Vasudev Lal, Arden Ma, Estelle Aflalo, Phillip Howard, Ana Simoes, Daniel Korat, Oren Pereg, Gadi Singer and Moshe Wasserblat

Representing ELMo embeddings as two-dimensional text online
Andrey Kutuzov and Elizaveta Kuzmenko

LOME: Large Ontology Multilingual Extraction
Patrick Xia, Guanghui Qin, Siddharth Vashishtha, Yunmo Chen, Tongfei Chen, Chandler May, Craig Harman, Kyle Rawlins, Aaron Steven White and Benjamin Van Durme

MadDog: A Web-based System for Acronym Identification and Disambiguation
Amir Pouran Ben Veyseh, Franck Dernoncourt, Walter Chang and Thien Huu Nguyen

Graph Matching and Graph Rewriting: GREW tools for corpus exploration, maintenance and conversion
Bruno Guillaume

Massive Choice, Ample Tasks (MaChAmp): A Toolkit for Multi-task Learning in NLP
Rob van der Goot, Ahmet Üstün, Alan Ramponi, Ibrahim Sharaf and Barbara Plank

SCoT: Sense Clustering over Time: a tool for the analysis of lexical change
Christian Haase, Saba Anwar, Seid Muhie Yimam, Alexander Friedrich and Chris Biemann

GCM: A Toolkit for Generating Synthetic Code-mixed Text
Mohd Sanad Zaki Rizvi, Anirudh Srinivasan, Tanuja Ganu, Monojit Choudhury and Sunayana Sitaram

No Day Set (continued)

T2NER: Transformers based Transfer Learning Framework for Named Entity Recognition
Saadullah Amin and Guenter Neumann

European Language Grid: A Joint Platform for the European Language Technology Community
Georg Rehm, Stelios Piperidis, Kalina Bontcheva, Jan Hajic, Victoria Arranz, Andrejs Vasiljevs, Gerhard Backfried, Jose Manuel Gomez-Perez, Ulrich Germann, Rémi Calizzano, Nils Feldhus, Stefanie Hegele, Florian Kintzel, Katrin Marheinecke, Julian Moreno-Schneider, Dimitris Galanis, Penny Labropoulou, Miltos Deligiannis, Katerina Gkirtzou, Athanasia Kolovou, Dimitris Gkoumas, Leon Voukoutis, Ian Roberts, Jana Hamrlova, Dusan Varis, Lukas Kacena, Khalid Choukri, Valérie Mapelli, Mickaël Rigault, Julija Melnika, Miro Janosik, Katja Prinz, Andres Garcia-Silva, Cristian Berrio, Ondrej Klejch and Steve Renals

A New Surprise Measure for Extracting Interesting Relationships between Persons
Hidetaka Kamigaito, Jingun Kwon, Young-In Song and Manabu Okumura

Paladin: an annotation tool based on active and proactive learning
Minh-Quoc Nghiem, Paul Baylis and Sophia Ananiadou

Story Centaur: Large Language Model Few Shot Learning as a Creative Writing Tool
Ben Swanson, Kory Mathewson, Ben Pietrzak, Sherol Chen and Monica Dinalescu

FrameForm: An Open-source Annotation Interface for FrameNet
Büşra Marşan and Olcay Taner Yıldız

OCTIS: Comparing and Optimizing Topic models is Simple!
Silvia Terragni, Elisabetta Fersini, Bruno Giovanni Galuzzi, Pietro Tropeano and ANTONIO CANDELIERI

ELITR Multilingual Live Subtitling: Demo and Strategy
Ondřej Bojar, Dominik Macháček, Sangeet Sagar, Otakar Smrž, Jonáš Kratochvíl, Peter Polák, Ebrahim Ansari, Mohammad Mahmoudi, Rishu Kumar, Dario Franceschini, Chiara Canton, Ivan Simonini, Thai-Son Nguyen, Felix Schneider, Sebastian Stüker, Alex Waibel, Barry Haddow, Rico Sennrich and Philip Williams

Breaking Writer's Block: Low-cost Fine-tuning of Natural Language Generation Models
Alexandre Duval, Thomas Lamson, Gaël de Léséleuc de Kérouara and Matthias Gallé

OPUS-CAT: Desktop NMT with CAT integration and local fine-tuning
Tommi Nieminen

Domain Expert Platform for Goal-Oriented Dialog Collection
Didzis Goško, Arturs Znotins, Inguna Skadina, Normunds Gruzitis and Gunta Nešpore-Bērzkalne

Which is Better for Deep Learning: Python or MATLAB? Answering Comparative Questions in Natural Language
Viktoriia Chekalina, Alexander Bondarenko, Chris Biemann, Meriem Beloucif, Varvara Logacheva and Alexander Panchenko

No Day Set (continued)

PunKtuator: A Multilingual Punctuation Restoration System for Spoken and Written Text
Varnith Chordia

Conversational Agent for Daily Living Assessment Coaching Demo
Raymond Finzel, Aditya Gaydhani, Sheena Dufresne, Maria Gini and Serguei Pakhomov

HULK: An Energy Efficiency Benchmark Platform for Responsible Natural Language Processing
Xiyou Zhou, Zhiyu Chen, Xiaoyong Jin and William Yang Wang