EMNLP 2021

The Third Workshop on New Frontiers in Summarization (NewSum 2021)

Workshop Proceedings

November 10, 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 978-1-955917-04-9

Message from the Workshop Chairs

Developing intelligent systems which can produce concise, fluent, and accurate summaries has been a long-standing goal in natural language processing. The aim of this workshop is to provide a research forum for cross-fertilization of ideas towards this goal. We seek to bring together researchers from a diverse range of fields (e.g., summarization, language generation, cognitive and psycholinguistics) for discussions on key issues related to automatic summarization. This includes discussion on novel paradigms/frameworks, multilingual and cross-lingual setups, shared tasks of interest, information integration and presentation, new evaluation protocols, applied research and applications, and possible future research foci. The workshop aims to pave the way towards building a cohesive research community, accelerating knowledge diffusion, developing new tools, datasets, and resources that are in line with the needs of academia, industry, and government. This is the third edition of the workshop, following our previous workshops at EMNLP 2017 and EMNLP 2019. The workshop received 18 long paper submissions, of which 9 were accepted, and 14 short paper submissions, of which 7 were accepted. This resulted in an overall acceptance rate of 50%. We are pleased to have three invited speakers at this year's workshop: Asli Celikyilmaz (Facebook AI Research), Sebastian Gehrmann (Google Research), and Shashi Narayan (Google Research). Together, they cover a broad spectrum of work in summarization and adjacent areas. We would like to thank our invited speakers, as well as our programme committee members and workshop participants.

Giuseppe Carenini, Jackie Chi Kit Cheung, Yue Dong, Fei Liu, Lu Wang

Organizing Committee

Giuseppe Carenini (University of British Columbia, Canada)
Jackie Chi Kit Cheung (McGill University, Canada)
Yue Dong (McGill University & MILA, Canada)
Fei Liu (University of Central Florida, USA)
Lu Wang (University of Michigan, USA)

Invited Speakers

Shashi Narayan (Google)
Asli Celikyilmaz (Facebook AI Research)
Sebastian Gehrmann (Google)

Technical Committee

Enamul Hoque (York University)
Jiacheng Xu (The University of Texas at Austin)
Rui Zhang (Penn State University)
Hou Pong Chan (University of Macau)
Yuntian Deng (Harvard University)
Kristjan Arumae (Amazon)
Xiaojun Wan (Peking University)
Chris Kedzie (Rasa Technologies Inc.)
Naoaki Okazaki (Tokyo Institute of Technology)
Manabu Okumura (Tokyo Institute of Technology)
Yang Liu (Microsoft)
Tadashi Nomoto (National Institute of Japanese Literature)
Linzi Xing (University of British Columbia)
Ari Rappoport (Hebrew University)
Felice Dell'Orletta (Istituto di Linguistica Computazionale "A. Zampolli" (CNR), Pisa, Italy)
Margot Mieskes (University of Applied Sciences Darmstadt, Germany)
Rodrigo Souza Wilkens (University of Essex)
Maxime Peyrard (EPFL)
Benoit Favre (Aix-Marseille University LIS/CNRS)
Tobias Falke (Amazon)
Thiago Alexandre Salgueiro Pardo (University of São Paulo)
Jessica Ouyang (University of Texas at Dallas)
Wencan Luo (Google)
Florian Boudin (Université de Nantes - France)
Juan-Manuel Torres-Moreno (LIA Avignon Université)
Michael Elhadad (Ben Gurion University)
Esaú Villatoro Tello (Universidad Autónoma Metropolitana Unidad Cuajimalpa, México)
Yuning Mao (University of Illinois at Urbana-Champaign)
Wen Xiao (University of British Columbia)
Xinyu Hua (Northeastern University)
Patrick Huber (University of British Columbia)
Abram Handler (University of Colorado)
Wojciech Kryściński (Salesforce Research)

Alexander Fabbri (Yale University)
Greg Durrett (UT Austin)
Yang Gao (Royal Holloway, University of London, UK)
Ramakanth Pasunuru (UNC Chapel Hill)
Ido Dagan (Bar-Ilan University)

Table of Contents

Sentence-level Planning for Especially Abstractive Summarization
Andreas Marfurt and James Henderson 1

Template-aware Attention Model for Earnings Call Report Generation
Yangchen Huang, Prashant K. Dhingra and Seyed Danial Mohseni Taheri 15

Knowledge and Keywords Augmented Abstractive Sentence Summarization
Shuo Guan 25

Rewards with Negative Examples for Reinforced Topic-Focused Abstractive Summarization
Khalil Mrini, Can Liu and Markus Dreyer 33

A Novel Wikipedia based Dataset for Monolingual and Cross-Lingual Summarization
Mehwish Fatima and Michael Strube 39

Evaluation of Summarization Systems across Gender, Age, and Race
Anna Jørgensen and Anders Søgaard 51

Evaluation of Abstractive Summarisation Models with Machine Translation in Deliberative Processes
Miguel Arana-Catania, Rob Procter, Yulan He and Maria Liakata 57

Capturing Speaker Incorrectness: Speaker-Focused Post-Correction for Abstractive Dialogue Summarization
Dongyub Lee, Jungwoo Lim, Taesun Whang, chanhee lee, Seungwoo Cho, Mingun Park and Heuiseok Lim 65

Measuring Similarity of Opinion-bearing Sentences
Wenyi Tay, Xiuzhen Zhang, Stephen Wan and Sarvnaz Karimi 74

EASE: Extractive-Abstractive Summarization End-to-End using the Information Bottleneck Principle
Haoran Li, Arash Einolghozati, Srinivasan Iyer, Bhargavi Paranjape, Yashar Mehdad, Sonal Gupta and Marjan Ghazvininejad 85

Context or No Context? A preliminary exploration of human-in-the-loop approach for Incremental Temporal Summarization in meetings
Nicole Beckage, Shachi H Kumar, Saurav Sahay and Ramesh Manuvinakurike 96

Are We Summarizing the Right Way? A Survey of Dialogue Summarization Data Sets
Don Tuggener, Margot Mieskes, Jan Deriu and Mark Cieliebak 107

Modeling Endorsement for Multi-Document Abstractive Summarization
Logan Lebanoff, Bingqing Wang, Zhe Feng and Fei Liu 119

SUBSUME: A Dataset for Subjective Summary Extraction from Wikipedia Documents
Nishant Yadav, Matteo Brucato, Anna Fariha, Oscar Youngquist, Julian Killingback, Alexandra Meliou and Peter Haas 131

TLDR9+: A Large Scale Resource for Extreme Summarization of Social Media Posts
Sajad Sotudeh, Hanieh Deilamsalehy, Franck Dernoncourt and Nazli Goharian 142

A New Dataset and Efficient Baselines for Document-level Text Simplification in German
Annette Rios, Nicolas Spring, Tannon Kew, Marek Kostrzewa, Andreas Säuberli, Mathias Müller and Sarah Ebling 152

Conference Program

9:00–10:30 **Morning Session I**

9:00–9:10 *Open remarks*
NewSum Organizers

9:10–10:00 *Keynote I*
Sashi Narayan (Google)

10:00–10:10 *Sentence-level Planning for Especially Abstractive Summarization*
Andreas Marfurt and James Henderson

10:10–10:20 *Template-aware Attention Model for Earnings Call Report Generation*
Yangchen Huang, Prashant K. Dhingra and Seyed Danial Mohseni Taheri

10:20–10:25 *Knowledge and Keywords Augmented Abstractive Sentence Summarization*
Shuo Guan

10:25–10:30 *Rewards with Negative Examples for Reinforced Topic-Focused Abstractive Summarization*
Khalil Mrini, Can Liu and Markus Dreyer

10:30–11:00 ***Coffee break I***

11:00–12:00 **Morning session II**

11:00–11:50 *Keynote II*
Sebastian Gehrmann (Google)

11:50–12:00 *A Novel Wikipedia based Dataset for Monolingual and Cross-Lingual Summarization*
Mehwish Fatima and Michael Strube

12:00–13:00 ***Lunch break***

November 10 (continued)

13:00–14:30 Afternoon session I

13:00–13:50 *Keynote III*
Asli Celikyilmaz (Facebook AI Research)

13:50–13:55 *Evaluation of Summarization Systems across Gender, Age, and Race*
Anna Jørgensen and Anders Søgaard

13:55–14:00 *Evaluation of Abstractive Summarisation Models with Machine Translation in Deliberative Processes*
Miguel Arana-Catania, Rob Procter, Yulan He and Maria Liakata

14:00–14:10 *Capturing Speaker Incorrectness: Speaker-Focused Post-Correction for Abstractive Dialogue Summarization*
Dongyub Lee, Jungwoo Lim, Taesun Whang, chanhee lee, Seungwoo Cho, Mingun Park and Heuiseok Lim

14:10–14:20 *Measuring Similarity of Opinion-bearing Sentences*
Wenyi Tay, Xiuzhen Zhang, Stephen Wan and Sarvnaz Karimi

14:20–14:30 *EASE: Extractive-Abstractive Summarization End-to-End using the Information Bottleneck Principle*
Haoran Li, Arash Einolghozati, Srinivasan Iyer, Bhargavi Paranjape, Yashar Mehdad, Sonal Gupta and Marjan Ghazvininejad

14:30–15:00 ***Coffee break***

15:00–16:15 Afternoon session II

15:00–15:10 *Context or No Context? A preliminary exploration of human-in-the-loop approach for Incremental Temporal Summarization in meetings*
Nicole Beckage, Shachi H Kumar, Saurav Sahay and Ramesh Manuvinakurike

15:10–15:20 *Are We Summarizing the Right Way? A Survey of Dialogue Summarization Data Sets*
Don Tuggener, Margot Mieskes, Jan Deriu and Mark Cieliebak

15:20–15:30 *Modeling Endorsement for Multi-Document Abstractive Summarization*
Logan Lebanoff, Bingqing Wang, Zhe Feng and Fei Liu

November 10 (continued)

15:30–15:35 *SUBSUME: A Dataset for Subjective Summary Extraction from Wikipedia Documents*
Nishant Yadav, Matteo Brucato, Anna Fariha, Oscar Youngquist, Julian Killingback, Alexandra Meliou and Peter Haas

15:35–16:15 ***EMNLP finding papers - Summarization***

16:15–16:45 ***Coffee break***

16:45–18:00 **Afternoon session III**

16:45–16:55 *TLDR9+: A Large Scale Resource for Extreme Summarization of Social Media Posts*
Sajad Sotudeh, Hanieh Deilamsalehy, Franck Dernoncourt and Nazli Goharian

16:55–17:00 *A New Dataset and Efficient Baselines for Document-level Text Simplification in German*
Annette Rios, Nicolas Spring, Tannon Kew, Marek Kostrzewa, Andreas Säuberli, Mathias Müller and Sarah Ebling

17:00–18:00 ***Mentoring Program***