The 4th New Frontiers in Summarization (with LLMs) Workshop

Welcome to join SigSUMM: www.sigsumm.org

Our Objective

- Fostering exchange of ideas related to automatic summarization
- Bringing together experts from various disciplines
- Discussing the latest updates and key issues

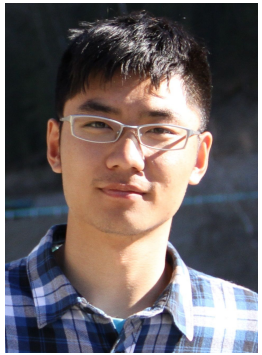
Summarization: Current status and Challenges

- LLMs show an advanced ability on automatic summarization
- LLMs have also been widely applied for evaluation
- Challenges:
  - Hallucination problems
  - Reliability of LLM evaluator
  - Long-document summarization
  - Multi-lingual / cross-lingual summarization
  - Customized summarization
  - Real-time summarization
  - ....

Keynote speakers

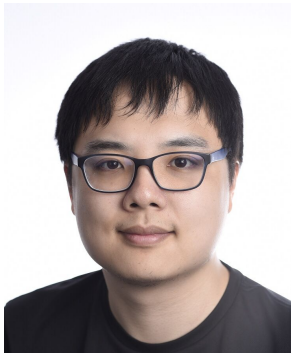
Kathleen McKeown
Columbia University

Jackie Cheung
McGill University



Rui Zhang
Penn State University

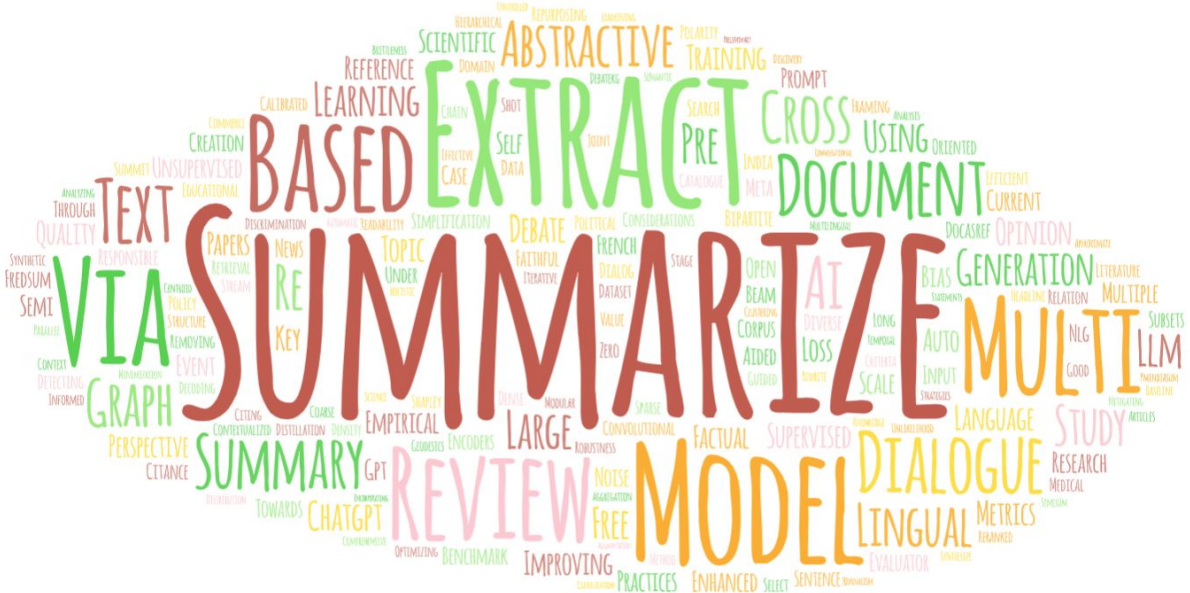
Iz Beltagy
Allen Institute for AI



Chenguang Zhu
Zoom

Statistics of the papers

- Number of submissions: 31, accepted papers: 13, acceptance rate: 42%
- Number of Findings papers to present: 25



Workshop Papers

- Is ChatGPT a Good NLG Evaluator? A Preliminary Study
- Zero-Shot Cross-Lingual Summarization via Large Language Models
- SimCSum: Joint Learning of Simplification and Cross-lingual Summarization for Cross-lingual Science Journalism
- Extract, Select and Rewrite: A Modular Sentence Summarization Method
- Summarization-based Data Augmentation for Document Classification
- In-context Learning of Large Language Models for Controlled Dialogue Summarization: A Holistic Benchmark and Empirical Analysis
- From Sparse to Dense: GPT-4 Summarization with Chain of Density Prompting
- Generating Extractive and Abstractive Summaries in Parallel from Scientific Articles Incorporating Citing Statements
- Supervising the Centroid Baseline for Extractive Multi-Document Summarization
- DebateKG – Automatic Policy Debate Case Creation with Semantic Knowledge Graphs
- Unsupervised Opinion Summarization Using Approximate Geodesics
- Analyzing Multi-Sentence Aggregation in Abstractive Summarization via the Shapley Value
- Improving Multi-Stage Long Document Summarization with Enhanced Coarse Summarizer

Findings Papers

- Bipartite Graph Pre-training for Unsupervised Extractive Summarization with Graph Convolutional Auto-Encoders
- Can you Summarize my learnings? Towards Perspective-based Educational Dialogue Summarization
- Citance-Contextualized Summarization of Scientific Papers
- DocAsRef: An Empirical Study on Repurposing Reference-based Summary Quality Metrics as Reference-free Metrics
- Enhancing abstractiveness of summarization models through calibrated distillation
- Extractive Summarization via ChatGPT for Faithful Summary Generation
- Factual Relation Discrimination for Factuality-oriented Abstractive Summarization
- FREDSum: A Dialogue Summarization Corpus for French Political Debates
- Hierarchical Catalogue Generation for Literature Review: A Benchmark
- Improving the Robustness of Summarization Models by Detecting and Removing Input Noise
- Large-Scale and Multi-Perspective Opinion Summarization with Diverse Review Subsets
- LLM aided semi-supervision for efficient Extractive Dialog Summarization
- Medical Text Simplification: Optimizing for Readability with Unlikelihood Training and Reranked Beam Search Decoding
- Mitigating Framing Bias with Polarity Minimization Loss
- Open Domain Multi-document Summarization: A Comprehensive Study of Model Brittleness under Retrieval
- PMIndiaSum: Multilingual and Cross-lingual Headline Summarization for Languages in India
- Re-Examining Summarization Evaluation across Multiple Quality Criteria
- Responsible AI Considerations in Text Summarization Research: A Review of Current Practices
- Summarizing Multiple Documents with Conversational Structure for Meta-Review Generation
- SummIt: Iterative Text Summarization via ChatGPT
- Synthesize, if you do not have: Effective Synthetic Dataset Creation Strategies for Self-Supervised Opinion Summarization in E-commerce
- Topic-Informed Dialogue Summarization using Topic Distribution and Prompt-based Modeling
- Unsupervised Opinion Summarization Using Approximate Geodesics
- Using LLM for Improving Key Event Discovery: Temporal-Guided News Stream Clustering with Event Summaries

Schedule

Time	Session Chair	Event & Details
08:50 - 09:00	Wen Xiao	Opening Remarks
09:00 - 09:45	Lei Yu	**Keynote I - Kathleen McKeown (Columbia University)** *Addressing Large Language Models that Lie: Case Studies in Summarization*
09:45 - 10:30	Lei Yu	**Keynote II - Jackie Cheung (McGill University)** *Open Problems in Automatic Summarization*
10:30 - 11:00	-	Coffee Break
11:00 - 11:45	Wen Xiao	**Keynote III - Rui Zhang (Penn State University)** *Are Large Language Models Fair Summarizers?*
11:45 - 12:30	Wen Xiao	**Keynote IV - Iz Beltagy (Allen Institute for AI)** *The Quest for Open Language Models*
12:30 - 14:00	-	Lunch Break
14:00 - 14:45	Wen Xiao	**Keynote V - Chenguang Zhu (Zoom)** *Facing the Challenges and Opportunities of LLMs*
14:45 - 15:30	Wen Xiao	**Lightning Talks** (Workshop papers and Findings papers)
15:30 - 16:00	-	Coffee Break
16:00 - 17:30	Wen Xiao	**Poster Session** (In-person/Virtual: Gathertown) (Workshop papers and Findings papers)

Program Committee

Manabu Okumura (Tokyo Institute of Technology)
Ido Dagan (Bar-Ilan University)
Ming Zhong (UIUC)
Kristjan Arumae (Qualtrics)
Pengcheng He (Microsoft Research)
Naoaki Okazaki (Tokyo Institute of Technology)
Zhe Hu (Baidu Inc)
Wojciech Kryscinski (Salesforce Research)
Haopeng Zhang (University of California Davis)
Hou Pong Chan (University of Macau)
Yang Liu (Microsoft)
Kaiqiang Song (Tencent AI Lab)
Juan-Manuel Torres-Moreno (LIA Avignon Université)
Jing Jiang (Singapore Management University)
Ziqiang Cao (Soochow University)
Margot Mieskes (University of Applied Sciences, Darmstadt)
Felice Dell'Orletta (Istituto di Linguistica Computazionale «A. Zampolli», CNR, Pisa, Italy)
Xinnuo Xu (University of Edinburgh)
Richard Evans (University of Wolverhampton)
Esau Villatoro-Tello (Idiap Research Institute)
Susana Bautista (Universidad Francisco de Vitoria)
Tobias Falke (Amazon Alexa)
Kellie Webster (Google)
Giulia Venturi (Institute for Computational Linguistics "A. Zampolli" (ILC-CNR)
Jessica Ouyang (University of Texas at Dallas)
Wencan Luo (Google)
Rui Zhang (Penn State University)
Linzi Xing (University of British Columbia)
Jiacheng Xu (Salesforce AI Research)
Tadashi Nomoto (National Institute of Japanese Literature)
Chao Zhao (UNC Chapel Hill)
Ori Shapira (Amazon)
Patrick Huber (UBC)
Florian Boudin (Nantes Université)
Xinyu Hua (Bloomberg)
Elena Lloret (University of Alicante, Spain)
Alexander Fabbri (Salesforce AI Research)
Tanya Goyal (UT Austin)
Yuntian Deng (Harvard University)
Maxime Peyrard (EPFL)
Arpit Sood (Meta)
Niyathi Allu (University of California, Riverside)
Priyanshu Sharma (University of California, Riverside)

Organizers

Yue Dong
University of California, Riverside, USA

Wen Xiao
Microsoft Azure AI, Canada

Wang Lu
University of Michigan, USA

Fei Liu
Emory University, USA

Giuseppe Carenini
University of British Columbia, Canada

Thanks!
Now let’s dive into the main sessions!