The 10th CCF International Conference on Natural Language Processing and Chinese Computing

NLPCC 2021 Handbook

13 October – 17 October, 2021
Qingdao, China

Preface by the Program Committee Co-Chairs

Welcome to NLPCC 2021, the tenth CCF International Conference on Natural Language Processing and Chinese Computing. Following the success of previous conferences held in Beijing (2012), Chongqing (2013), Shenzhen (2014),Nanchang (2015), Kunming (2016), Dalian (2017), Hohhot (2018), Dunhuang(2019), and Zhengzhou (2020), this year's NLPCC is held at Qingdao, a beautifulcoastal city in East China. As a premier international conference on Natural Lan-guage Processing and Chinese Computing, organized by the CCF-NLP (Technical Committee of Natural Language Processing, China Computer Federation,formerly known as Technical Committee of Chinese Information, China Computer Federation), NLPCC 2021 serves as an important forum for researchersand practitioners from academia, industry, and government to share their ideas,research results and experiences, and to promote their research and technicalinnovations in the fields.

The fields of natural language processing (NLP) and Chinese computing (CC) have boomed in recent years. Following NLPCC's tradition, we welcomed submissions in ten areas for the main conference: Fundamentals of NLP; Machine Translation and Multilinguality; Machine Learning for NLP; Informtion Extraction and Knowledge Graph; Summarization and Generation; Question Answering; Dialogue Systems; Social Media and Sentiment Analysis; NLP Applications and Text Mining; Multimodality and Explainability. On the submission deadline, we are trilled to recieve a record number of 446 valid submissions to the main conference.

After a rigid review process, out of 446 submissions (some of which were withdrawn or rejected without review due to format issues or policy violations),104 papers were finally accepted to appear in the main conference, where 89 are written in English and 15 in Chinese, resulting in an acceptance rate of 23.3%. Among them, 72 submissions were accepted as oral papers and 32 as poster papers. Specifically, ten papers were nominated by our area chairs for the best paper award. An independent best paper award committee was formed to select the best papers from the shortlist. This proceedings include only the accepted English papers; the Chinese papers will appear in the ACTA Scientiarum Naturalium Universitatis Pekinensis. In addition to the main proceedings, 3 papers were accepted to the student workshop, 22 papers were accepted to the Evaluation workshop, and 2 papers were accepted to the Explainable AI (XAI) workshop.

We are honored to have four internationally renowned keynote speakers —Rada Mihalcea (University of Michigan), Yanchao Bi (Beijing Normal University), Sebastian Riedel (University College London and Facebook AI Research)and Graham Neubig (Carnegie Mellon University) —share their findings on recent research progress and achievements in natural language processing.

We would also like to thank all the people who have contributed to NLPCC2021.First of all, we would like to thank our 20 area chairs for their hard work for recruiting reviewers, monitoring the review and discussion processes, and carefully rating and recommending submissions. We would like to thank all 432 reviewers for their time and efforts to review the submissions. We are very grateful to Tim Baldwin, Chin-Yew Lin, Kang Liu, Deyi Xiong and Yue Zhang for their participation in the best paper committee. We are also grateful for the help and support from the general chairs, Tim Baldwin and Jie Tang, and from the organization committee chairs, Zhumin Chen, Pengjie Ren, and Xiaojun Wan. Special thanks go to Yu Hong and Ruifang He, the publication chairs, for their great help. We greatly

appreciate all your help!

Finally, we would like to thank all the authors who submitted their work to NLPCC 2021, and thank our sponsors for their contributions to the conference.Without your support, we cannot have such a strong conference program.

We are looking forward to seeing you at NLPCC 2021 in Qingdao and hope you enjoy the conference!

Lu Wang, University of Michigan, Ann Arbor
Yansong Feng, Peking University
NLPCC 2021, Program Co-Chairs

Organizing Committee

Organizer
China Computer Federation (CCF)

Host
Shandong University

Publishers
Lecture Notes on Artificial Intelligence (LNAI), Springer Verlag
ACTA Scientiarum Naturalium Universitatis Pekinensis

General Chairs

Tim Baldwin	University of Melbourne
Jie Tang	Tsinghua University

Program Committee Chairs

Lu Wang	University of Michigan
Yansong Feng	Peking University

Student Workshop Chairs

Fang Kong	Soochow University
Meishan Zhang	Tianjin University

Evaluation Chairs

Jiajun Zhang	Institute of Automation, Chinese Academy of Sciences
Yunbo Cao	Tencent

Tutorial Chairs

Minlie Huang	Tsinghua University
Zhongyu Wei	Fudan University

Publication Chairs

Yu Hong	Soochow University
Ruifang He	Tianjin University

Journal Coordinator

Yunfang Wu	Peking University

Conference Handbook Chair

Xiaomeng Song	Shandong University

Sponsorship Chairs

Dongyan Zhao	Peking University
Zhaochun Ren	Shandong University

Publicity Chairs

Wei Jia	Beijing Academy of Artificial Intelligence (BAAI)
Jing Li	Hong Kong Polytechnic University

Organization Committee Chairs

Zhumin Chen	Shandong University
Pengjie Ren	Shandong University
Xiaojun Wan	Peking University

Area Chairs

Fundamentals of NLP

Liang Huang	Oregon State University
Kewei Tu	ShanghaiTech University

Machine Translation and Multilinguality

Derek Wong	University of Macau
Boxing Chen	Alibaba Group

Machine Learning for NLP

Yangfeng Ji	University of Virginia
Lingpeng Kong	The University of Hong Kong

Information Extraction and Knowledge Graph

Lifu Huang	Virginia Tech
Shizhu He	Chinese Academy of Sciences

Summarization and Generation

Rui Zhang	Penn State University
Jin-ge Yao	Microsoft Research

Question Answering

Huan Sun	The Ohio State University
Yiming Cui	Harbin Institute of Technology

Dialogue Systems

Jessy Li	The University of Texas at Austin
Wei Wu	Meituan

Social Media and Sentiment Analysis

Fei Liu	University of Central Florida
Duyu Tang	Tencent AI Lab

NLP Applications and Text Mining

Wenpeng Yin	Salesforce Research
Zhunchen Luo	PLA Academy of Military Science

Multimodality and Explainability

Xin Wang	UC Santa Cruz
Zhongyu Wei	Fudan University

Treasurer

Yajing Zhang	Soochow University
Xueying Zhang	Peking University

Webmaster

Hui Liu	Peking University

Program Committee

Program Committee Chairs

Lu Wang	University of Michigan
Yansong Feng	Peking University

Program Committee Members

Ayana	Inner Mongolia University of Finance and Economics
Bo An	Institute of Software, Chinese Academy of Sciences
Xiang Ao	Institute of Computing Technology, Chinese Academy of Sciences
Jiaqi Bai	Beihang University
Guangsheng Bao	Westlake University
Junwei Bao	JD AI Research
Qiming Bao	The University of Auckland
Paul Buitelaar	Data Science Institute, National University of Ireland Galway
Xuanting Cai	Facebook
Xiangrui Cai	Nankai University
Yi Cai	South China University of Technology
Yixin Cao	Nanyang Technological University
Yuan Cao	Google Brain
Shuyang Cao	Univerisity of Michigan
Pengfei Cao	Institute of Automation, Chinese Academy of Sciences
Ziqiang Cao	Soochow University
Hailong Cao	Harbin Institute of Technology
Zhangming Chan	Peking University
Zhumin Chen	Shandong University
Sihao Chen	University of Pennsylvania
Xiuying Chen	Peking University
Jiajun Chen	Nanjing University
Junkun Chen	Oregon State University
Sanxing Chen	University of Virginia
Hongshen Chen	JD.com
Zhipeng Chen	iFLYTEK Research
Jie Chen	Ahu University
Qingcai Chen	Harbin Institute of Technology,Shenzhen
Lei Chen	School of Data Science, Fudan University
Wenliang Chen	Soochow University
Shuang Chen	Harbin Institute of Technology
Xinyu Chen	Soochow University
Ruijun Chen	Yunnan University
Bo Chen	iscas.ac.cn
Yulong Chen	Zhejiang University, Westlake University
Huadong Chen	ByteDance
Pei Chen	Texas A&M University
Lu Chen	Shanghai Jiao Tong University
Guanyi Chen	Utrecht University

Yidong Chen	Department of Artificial Intelligence, School of Informatics, Xiamen University
Liwei Chen	Peking University
Xinchi Chen	Amazon AWS
Hanjie Chen	University of Virginia
Wei Chen	School of Data Science, Fudan University
Hannah Chen	University of Virginia
Yubo Chen	Institute of Automation, Chinese Academy of Sciences
Kehai Chen	National Institute of Information and Communications Technology
Mingda Chen	Toyota Technological Institute at Chicago
Jiangjie Chen	Fudan University
Jifan Chen	UT Austin
Jiaao Chen	Georgia Institute of Technology
Jiayang Cheng	Fudan University
Xin Cheng	Peking University
Qikai Cheng	Wuhan University
Zewen Chi	Beijing Institute of Technology
Zewei Chu	Google
Chenhui Chu	Kyoto University
James Cross	Facebook
Peng Cui	Harbin Institute of Technology
Yiming Cui	Harbin Institute of Technology
mao cunli	kunming university of science and technology
Junqi Dai	Fudan University
Xinyu Dai	National Key Laboratory for Novel Software Technology,Nanjing University
Xiang Deng	The Ohio State University
Xiao Ding	Harbin Institute of Technology
Chenchen	Ding NICT
Zhicheng Dou	Renmin University of China
Dejing Dou	University of Oregon
Ziyi Dou	UCLA
Rotem Dror	University of Pennsylvania
Wanyu Du	University of Virginia
Jinhua Du	Investments AI, AIG
Xinya Du	Cornell University
Junwen Duan	Central South Universit
Chaoqun Duan	Harbin Institute of Technology
Xiangyu Duan	Soochow University
Philipp Dufter	Center for Information and Language Processing, LMU Munich
Kevin Duh	Johns Hopkins University
Ismail El Maarouf	Fortia Financial Solutions
Alexander Fabbri	Yale University
Chuang Fan	Harbin Institute of Technology,Shenzhen
Zhihao Fan	Fudan University
Zhiyuan Fang	Arizona State University
Xiachong Feng	Harbin Institute of Technology
Zhangyin Feng	Harbin Institute of Technology
Xiaocheng Feng	Harbin Institute of Technology,SCIR lab
Shi Feng	Northeastern University
Jiazhan Feng	Peking University
Xingyu Fu	Upenn

Yao Fu	The University of Edinburgh
Guohong Fu	Soochow University
Zihao Fu	The Chinese University of Hong Kong
Yi Fung	University of Illinois at Urbana-Champaign
Jun Gao	Harbin Institute of Technology,Shenzhen
Shen Gao	Peking University China
Yang Gao	Royal Holloway, University of London
Yifan Gao	The Chinese University of Hong Kong
Michaela Geierhos	Universität der Bundeswehr München
Ruiying Geng	Alibaba Group
Erfan Ghadery	KU Leuven
Heng Gong	Harbin Institute of Technology
Yeyun Gong	Microsoft Research Asia
Jiachen Gu	University of Science and Technology of China
Yu Gu	The Ohio State University
Jian Guan	Tsinghua University
Yuhang Guo	Beijing Institute of Technology
Daya Guo	Sun Yat-Sen University
Qipeng Guo	Fudan University
Junliang Guo	University of Science and Technology of China
Jiale Han	Beijing University of Posts and Telecommunications
Fangqiu Han	facebook Inc
Xudong Han	The university of Melbourne
Lifeng Han	Dublin City University
Wenjuan Han	Beijing Institute for General Artificial Intelligence
Tianyong Hao	School of Computer Science,South China Normal University
Shizhu He	Institute of Automation, Chinese Academy of Sciences
Zhongjun He	Baidu, Inc.
Ziwei He	Shanghai Jiao Tong University
Ben He	University of Chinese Academy of Sciences
Ruifang He	Tianjin University
Xudong Hong	Saarland University / MPI Informatics
Lei Hou	Tsinghua University
Yacheng Hsu	The University of Hong Kong Taiwan
Huang Hu	STCA NLP Group, Microsoft
Zhe Hu	Baidu
Zikun Hu	National University of Singapore
Minghao Hu	PLA Academy of Military Science
Baotian Hu	Harbin Institue of Technology,Shenzhen
Wenpeng Hu	School of Mathematical Sciences,Peking University
Zechuan Hu	ShanghaiTech University
Xinyu Hua	Northeastern University
Zhen Huang	National University of Defense Technology
Jiangping Huang	Chongqing University of Posts and Telecommunications
Shujian Huang	National Key Laboratory for Novel Software Technology,Nanjing University
Qingbao Huang	Guangxi University;South China University of Technology
Zhongqiang Huang	Alibaba Group
Yongjie Huang	Sun Yat-sen University
Junjie Huang	Beihang University
Yi-Ting Huang	Academia Sinica Taiwan
Julia Ive	Imperial College London

Lei Ji	MSRA
Menglin Jia	Cornell University
Zixia Jia	ShanghaiTech University
Ping Jian	Beijing Institute of Technology
Xuhui Jiang	Institute of Computing Technology , Chinese Academy of Sciences
Yong Jiang	Alibaba DAMO Academy
Xin Jiang	Huawei Noah's Ark Lab
Ming Jiang	University of Illinois at Urbana-Champaign
Yichen Jiang	University of North Carolina at Chapel Hill
Wenbin Jiang	Baidu Inc.
Peng Jin	Leshan Normal University
Xisen Jin	University of Southern California
Zhijing Jin	Max Planck Institute for Intelligent Systems
Lin Jun	alibaba
Zixuan Ke	University of Illinois at Chicago
Dongjin Kim	KAIST
Fang Kong	Soochow University
Bernhard Kratzwald	ETH Zurich
Tuan Lai	University of Illinois at Urbana-Champaign
Viet Lai	University of Oregon
Yuxuan Lai	Peking University
Man Lan	East China Normal University
Chongshou Li	Southwest Jiaotong Univiersity
Maoxi Li	School of Computer Information Engineering, Jiangxi Normal University
Bin Li	Nanjing Normal University
Mingzhe Li	Peking University
Hongzheng Li	School of Foreign Languages, Beijing Institute of Technology
Chenliang Li	Wuhan University
Zhongyang Li	Harbin Institute of Technology
Manling Li	UIUC
Dongfang Li	Harbin Institute of Technology,Shenzhen
Xinyi Li	National University of Defense Technology
Zejun Li	Fudan University
Peifeng Li	Soochow University
Irene Li	Yale University
Jian Li	Huawei Noah's Ark Lab
Piji Li	Tencent AI Lab
Xin Li	Alibaba Group
Lishuang Li	Dalian University of Technology
Fei Li	Wuhan University
Weikang Li	Peking University
Yachao Li	Soochow University
Liangyou Li	Noah's Ark Lab, Huawei Technologies
Qintong Li	Shandong University
Xinhang Li	Tsinghua University
Linjie Li	Microsoft
Yanran Li	The Hong Kong Polytechnic University
Jing Li	Department of Computing, The Hong Kong Polytechnic University
Junhui Li	Soochow University, Suzhou

Zuchao Li	Shanghai Jiao Tong University
Lin Li	Qinghai Normal University
Zhenghua Li	Soochow University
Ruizhe Li	University of Sheffield
Xilai Li	Amazon
Peng Li	Institute of Information Engineering, CAS
Liunian Harold Li	UCLA
Binyang Li	University of International Relations
Xiaolong Li	SAS Institute Inc.
Zhixu Li	Soochow University
Zujie Liang	Sun Yat-Sen University
Paul Pu Liang	Carnegie Mellon University
Ying Lin	Apple
Junyang Lin	Alibaba Group
Zhaojiang Lin	The Hong Kong University of Science and Technology
Zi Lin	Google
Zhouhan Lin	Shanghai Jiao Tong University
Ye Liu	National University of Singapore
Chang Liu	Peking University
Ling Liu	University of Colorado Boulder
Qun Liu	Chongqing University of Posts and Telecommunications
Tianyu Liu	Peking University
Shujie Liu	Microsoft Research Asia, Beijing,China
Qun Liu	Huawei Noah's Ark Lab
Dayiheng Liu	Alibaba DAMO Academy
Qingbin Liu	National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, University of Chinese Academy of Sciences
Yongbin Liu	School of Computer Science, University of South China
Xuebo Liu	University of Macau
Jian Liu	Beijing Jiaotong University
Xianggen Liu	Tsinghua University
Jiangming Liu	The University of Edinburgh
Lemao Liu	Tencent AI Lab
Xiao Liu	Beijing Institute of Technology
Qian Liu	Beihang University
Yunfei Long	University of Essex
Xin Lu	Harbin Institute of Technology
Hengtong Lu	Beijing University of Posts and Telecommunications
Yaojie Lu	Institute of Software, Chinese Academy of Sciences
Yinglong Ma	North China Electric Power University
Yun Ma	Peking University
Xianling Mao	Beijing Institute of Technology
Tao Meng	UCLA
Haitao Mi	Ant Group US
Tao Mingxu	Peking University
Xiangyang Mou	Rensselaer Polytechnic Institute
Preslav Nakov	Qatar Computing Research Institute, HBKU
Feng Nie	sysu.edu.cn
Qiang Ning	Amazon

Yawen Ouyang	National Key Laboratory for Novel Software Technology,Nanjing University
Vardaan Pahuja	The Ohio State University
Huijie Pan	The University of Hong Kong
Jiaxin Pei	University of Michigan
Xutan Peng	The University of Sheffield
Baolin Peng	Microsoft Research
Longhua Qian	Soochow University
Tao Qian	Hubei University of Science and Technology
Libo Qin	Harbin Institute of Technology
Vardaan Pahuja	The Ohio State University
Huijie Pan	The University of Hong Kong
Jiaxin Pei	University of Michigan
Xutan Peng	The University of Sheffield
Baolin Peng	Microsoft Research
Longhua Qian	Soochow University
Tao Qian	Hubei University of Science and Technology
Libo Qin	Harbin Institute of Technology
Yanxia Qin	School of Computer Science and Technology, Donghua University
Liang Qiu	University of California, Los Angeles
Zhaochun Ren	Shandong University
Shuo Ren	Beihang University
Pengjie Ren	Shandong University
Martin Schmitt	Center for Information and Language Processing, LMU Munich
Stephanie Schoch	University of Virginia
Lei Sha	University of Oxford
Zhihong Shao	Tsinghua University
Lei Shen	Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences
Haoyue Shi	Toyota Technological Institute at Chicago
Weijia Shi	ucla.edu
Weiyan Shi	Columbia University
Lei Shu	Amazon AWS AI
Chenglei Si	University of Maryland, College Park
Jyotika Singh	ICX Media, Inc.
Yiping Song	National University of Defense Technology
Kaiqiang Song	University of Central Florida
Yixuan Su	University of Cambridge
Elior Sulem	University of Pennsylvania
Simeng Sun	University of Massachusetts Amherst
Huan Sun	The Ohio State University
Kai Sun	Cornell University
Chengjie Sun	Harbin Institute of Technology
Chuanqi Tan	Alibaba Group
Minghuan Tan	Singapore Management University
Zhixing Tan	Tsinghua University
Danie Tang	National Engineering Research Center for Big Data Technology and System, Service Computing Technology and System Lab,Huazhong University of Science and Technology
Jintao Tang	National University of Defense Technology

Duyu Tang	Tencent
Buzhou Tang	Harbin Institute of Technology Shenzhen Graduate School
Ruixuan Tang	University of Virginia
Chongyang Tao	Microsoft Corporation
Zhiyang Teng	Westlake University
Zhiliang Tian	Hong Kong University of Science and Technology
Zhoujin Tian	Beihang University
Lifu Tu	Toyota Technological Institute at Chicago
Yu Wan	NLP2CT Lab, University of Macau
Liran Wang	Beihang University
Lingzhi Wang	The Chinese University of Hong Kong
Di Wang	Woobo Inc
Wei Wang	Shenzhen International Graduate School, Tsinghua University
Ruize Wang	Academy for Engineering and Technology, Fudan University
Xing Wang	Tencent
Tao Wang	Department of Biostatistics & Health Informatics, King's College London
Jingjing Wang	Soochow University
Le Wang	Chinese Academy of Military Science
Hongwei Wang	University of Illinois Urbana-Champaign
Hongling Wang	Soochow University
Shaonan Wang	National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences
Sijia Wang	Virginia Tech
Xun Wang	Umass Amherst
Zhen Wang	The Ohio State University
Rui Wang	Shanghai Jiao Tong University
Siyuan Wang	Fudan University
Yaqiang Wang	Chengdu University of Information Technology
Dingmin Wang	University of Oxford
Zijian Wang	AWS AI Labs
Qingyun Wang	University of Illinois at Urbana-Champaign
Ke Wang	Peking University
Kexiang Wang	Peking University
Pancheng Wang	National University of Defence Technology
Ge Wang	ShanghaiTech University
Hong Wang	University of California Santa Barbara
Zhuoyu Wei	Yuanfudao Inc.
Zhongyu Wei	School of Data Science, Fudan University
Haoyang Wen	Carnegie Mellon University
Shuangzhi Wu	Tencent
Sixing Wu	School of Electronics Engineering and Computer Science, Peking University
Junshuang Wu	Beihang University
Yu Wu	Microsoft Research Asia
Lianwei Wu	School of Software Engineering,Xi'an Jiaotong University
Chiensheng Wu	Salesforce
Congying Xia	University of Illinois at Chicago
Rui Xia	Nanjing University of Science and Technology
Wen Xiao	University of British Columbia

Tong Xiao	Northeastern University
Yuqiang Xie	Institute of Information Engineering, Chinese Academy of Sciences
Deyi Xiong	Tianjin University
Hao Xiong	Alibaba
Can Xu	STCA NLP Group, Microsoft
Ruijian Xu	Peking University
Silei Xu	Stanford University
Jingjing Xu	ByteDance AI Lab
Tong Xu	University of Science and Technology of China
Yiheng Xu	Microsoft Research Asia
Yan Xu	Hong Kong University of Science and Technology
Jinan Xu	Beijing Jiaotong University
Yadollah Yaghoobzadeh	University of Tehran
Yuanmeng Yan	Beijing University of Posts and Telecommunications
Lingyong Yan	Baidu Inc.
Songlin Yang	ShanghaiTech University
Liang Yang	Dalian University of Technology
Muyun Yang	Harbin Institute of Technology
Yaqin Yang	Paypal United States
Jingxuan Yang	Beijing University of Posts and Telecommunications
Qiang Yang	KAUST
Ziqing Yang	iFLYTEK Research
Haoran Yang	The Chinese University of Hong Kong
Baosong Yang	Alibaba Damo Academy, Alibaba Inc.
Cheng Yang	Beijing University of Posts and Telecommunications
Zixiaofan Yang	Columbia University
Zhiwei Yang	College of Computer Science and Technology, Jilin University
Zhewei Yao	UC Berkeley
Jianmin Yao	Soochow University
Yiqun Yao	University of Michigan
Wenlin Yao	Tencent AI Lab
Zhe Ye	Microsoft
Xiaoyuan Yi	Tsinghua University
Pengcheng Yin	Carnegie Mellon University
Da Yin	University of California, Los Angeles (UCLA)
Zhiwei Yu	Microsoft Research Asia
Adams Yu	Google Brain United States
Tiezheng Yu	The Hong Kong University of Science and Technology
Heng Yu	Alibaba
Pengfei Yu	Department of Computer Science,University of Illinois at Urbana-Champaign
Dian Yu	Tencent AI Lab
Xiaodong Yu	University of Pennsylvania
Chunyuan Yuan	Institute of Information Engineering,Chinese Academy of Sciences
Zheng Yuan	University of Cambridge
Xiang Yue	The Ohio State University
Zhiyuan Zeng	Beijing University of Posts and Telecommunications
Qi Zeng	University of Illinois at Urbana-Champaign

Shuang (Sophie) Zhai	University of Oklahoma
Danqing Zhang	Amazon
Jiajun Zhang	Institute of Automation Chinese Academy of Sciences
Zixuan Zhang	University of Illinois Urbana-Champaign
Han Zhang	Vispek
Peng Zhang	Tianjin University
Meishan Zhang	Tianjin University, China
Zhifei Zhang	Tongji University
Wenxuan Zhang	The Chinese University of Hong Kong
Hongming Zhang	HKUST
Biao Zhang	University of Edinburgh
Qi Zhang	Fudan University
Liwen Zhang	ShanghaiTech University
Chengzhi Zhang	Nanjing University of Science and Technology
Iris Zhang	Cornell University
Zhirui Zhang	Alibaba DAMO Academy
Shuaicheng Zhang	Virginia Polytechnic Institute and State University
Tongtao Zhang	Siemens Corporate Technology
Dongxu Zhang	University of Massachusetts,Amherst
Xuanyu Zhang	Du Xiaoman Financial
Fan Zhang	Tianjin University, China
Hao Zhang	Agency for Science, Technology and Research
Zhuosheng Zhang	Shanghai Jiao Tong University
Mengjie Zhao	Center for Information and Language Processing, LMU Munich
Yufan Zhao	Microsoft
Yang Zhao	Institute of Automation, Chinese Academy of Sciences
Xiang Zhao	National University of Defense Technology
Zhenjie Zhao	Nanjing University of Information Science and Technology
Sanqiang Zhao	University of Pittsburgh, Department of Informatics and Networked Systems, School of Computing and Information
Lulu Zhao	Beijing University of Posts and Telecommunications
Kai Zhao	Google LLC
Jie Zhao	Amazon
Renjie Zheng	Baidu Research
Chen Zheng	Michigan State University
Ming Zhong	University of Illinois at Urbana-Champaign
Wanjun Zhong	Sun Yat-Sen University
Junsheng Zhou	Nanjing Normal University
Wenxuan Zhou	University of Southern California
Guangyou Zhou	School of Computer Science, Central China Normal University
Qingyu Zhou	Tencent
Ben Zhou	University of Pennsylvania
Xian Zhou	Information Research Center of Military Science
Jie Zhu	Soochow University
Muhua Zhu	Tencent Inc.
Junnan Zhu	Institute of Automation, Chinese Academy of Sciences
Shuaichen Chang	The Ohio State University
Mengxing Dong	Soochow University

Yanqing He	Institute of Scientific and Technical Information of China
Yu Hong	Soochow University
Zhiling Jin	Soochow University
Shasha Li	College of Computer, National University of Defense Technology
Zichao Li	McGill University
Xiao Li	University of Reading
Xing Shi	DiDi Research America Inc.
Linfeng Song	Tencent AI Lab
Wei Song	Capital Normal University
Jianyou Wang	University of California, San Diego
Changxing Wu	East China Jiaotong University
Kai Yang	Peking University
Tao Yu	Yale University
Yazhou Zhang	Zhengzhou University of Light Industry
Shi Zong	Nanjing University
Wangchunshu Zhou	Beihang Univeristy

Organizers and Supports

Organizer

Hosts

Publishers

Special Technical Sponsor

Diamond Sponsors

Platinum Sponsors

Golden Sponsors

Sliver Sponsors

Contents

1 Organization **1**

2 Conference Venues **3**
Venue . . . 3
Main Conference & Workshop . . . 4

3 Conference Agenda **6**
Tutorial (Online) . . . 6
Workshop . . . 7
Main Conference & Workshop . . . 8

4 Tutorials : Oct, 13 – 14 **10**
Schedule . . . 10
Tutorial 1 . . . 11
Tutorial 2 . . . 12
Tutorial 3 . . . 13
Tutorial 4 . . . 14
Tutorial 5 . . . 15
Tutorial 6 . . . 16
Tutorial 7 . . . 17
Tutorial 8 . . . 18

5 3 Workshops : Oct, 15 **20**
MMNLP Workshop: The Second Workshop on Multimodal Natural Language Processing . . 23
Baidu Workshop:Frontier Progresses and Applications of Pre-trained Models 28
XAI Workshop: Align with Changes: A Practice of NLP from Explanation to Adaption . . . 34

6 Main Conference & 2 Workshops : Oct, 16 **40**
Keynote Talk 1: Rada Mihalcea 41
Keynote Talk 2: Yanchao Bi 42
Industrial Talk 1: Junlan Feng 43
Industrial Talk 2: Min Chu 45
Oral Paper Sessions 46
Posters Session 50
Student Workshop 54
Evaluation Workshop 58

7 Main Conference : Oct, 17 **59**
Keynote Talk 3: Sebastian Riedel 60
Keynote Talk 4: Graham Neubig 61
Industrial Talk 3: Songfang Huang 62
Industrial Talk 4: Yang Wu 63
Oral Paper Sessions 64

Index **70**

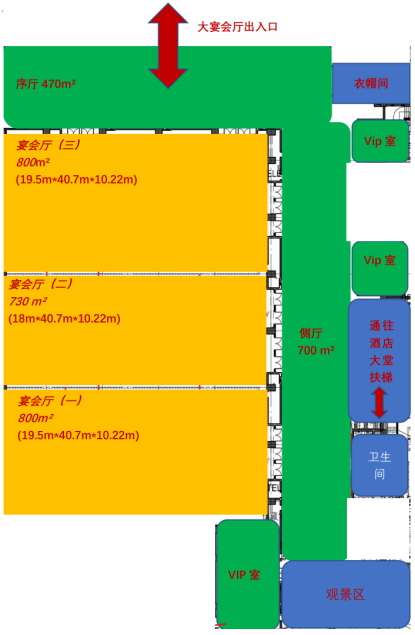
1 Organization

NLPCC (CCF International Conference on Natural Language Processing & Chinese Computing) is the annual conference of CCF TCNLP (Technical Committee of Natural Language Processing, China Computer Federation). As a leading conference on the field of NLP & Chinese Computing of CCF, NLPCC is the premier forum for the NLP researchers and practitioners from academia, industry, and government in China and Pacific Asia area to share their ideas, research results and experiences, which will highly promote the research and technical innovation in these fields domestically and internationally.

2 Conference Venues

Venue

Haitian Hotel QINGDAO, 青岛海天大酒店
1st Floor



Main Conference & Workshop

Haitian Hotel, Qingdao
G Floor Diagram

3 Conference Agenda

Tutorial (Online)

Oct. 13

08:55-09:00	Opening Remarks	Tutorial Chairs
09:00-10:30	Lecture 1: Natural Language Question Answering over Knowledge Graph	Lei Zou
10:45-12:15	Lecture 2: Learning towards Knowledge-driven Dialogue System between Human and Computer	Rui Yan
14:00-15:30	Lecture 3: Group-Theoretic Self-Supervised Representation Learning	Hanwang Zhang
15:45-17:15	Lecture 4: OOD Sample and New Class under Weakly Supervised Scenario	Yu-Feng Li

Oct. 14

09:00-10:30	Lecture 5: Will Non-Autoregressive Translation Dominate Neural Machine Translation?	Yang Feng
10:45-12:15	Lecture 6: Self-supervised Reinforcement Learning	Jianye Hao
14:00-15:30	Lecture 7: Generative Adversarial Nets Meet Reinforcement Learning	Weinan Zhang
15:45-17:15	Lecture 8: Robustness in Natural Language Processing	Qi Zhang & Tao Gui

Workshop

Oct. 15 3 Workshops		
09:00-22:00	NLPCC 2021 Onsite Registration	Lobby
14:00-17:00	MMNLP Workshop *Chair: Youzheng Wu*	Qindao Function Room 1, G floor
14:00-17:00	Baidu Workshop *Chair: Hua Wu, Yu Sun*	Qindao Function Room 2, G floor
14:00-17:00	XAI Workshop *Chair:Yao Meng*	Meeting Room 08 & 09, G floor
15:30-16:00	Break	G floor public area
17:00-17:30	Grand Data Award	Meeting Room 08 & 09, G floor
18:00-21:00	Dinner	Haitian Grand Ballroom 2(Buffet), 1st floor
19:00 -21:00	TCNLP Annual Meeting (Only TCNLP Members)	Qindao Function Room 1, G floor

Main Conference & Workshop

Oct. 16 Main Conference + 2 Workshops

09:00-12:00	NLPCC 2021 Onsite Registration	Lobby
08:30-9:00	NLPCC 2021 Opening Ceremony	Haitian Grand Ballroom 1, 1st floor
09:00-10:00	Keynote Talk 1, by Rada Mihalcea	Haitian Grand Ballroom 1, 1st floor
10:00-10:15	Industrial Talk 1, by Junlan Feng	Haitian Grand Ballroom 1, 1st floor
10:15-10:45	Break & Group Photo	Outdoor plaza
10:45-11:45	Keynote Talk 2, by Yanchao Bi	Haitian Grand Ballroom 1, 1st floor
11:45-12:00	Industrial Talk 2, by Min Chu	Haitian Grand Ballroom 1, 1st floor
12:00-13:30	Lunch	Haitian Grand Ballroom 2, 1st floor
13:30-15:30	Posters Session 1	G floor public area
	Oral 1: Fundementals of NLP	Haitian Grand Ballroom 1, 1st floor
	Oral 2: Information Extraction	Meeting room 06,07, G floor
	Student Workshop	Meeting room 08,09, G floor
15:30-16:00	Break	
16:00-18:00	Posters Session 2	G floor public area
	Oral 3: Machine Translation and Multilinguality	Haitian Grand Ballroom 1, 1st floor
	Oral 4: Machine Learning	Meeting room 06,07, G floor
	Evaluation Workshop	Meeting room 08,09, G floor
18:30-21:00	Dinner	Haitian Grand Ballroom 2(Banquet), 1st floor

Oct. 17 Main conference

09:00-10:00	Keynote Talk 3, by Sebastian Riedel	Haitian Grand Ballroom 1(Online), 1st floor
10:00-10:15	Industrial Talk 3, by Songfang Huang	Haitian Grand Ballroom 1(Online), 1st floor
10:15-10:45	Break	
10:45-11:45	Keynote Talk 4, by Graham Neubig	Haitian Grand Ballroom 1(Online), 1st floor
11:45-12:00	Industrial Talk 4, by Yang Wu	Haitian Grand Ballroom 1, 1st floor
12:00-13:30	Lunch	Haitian Grand Ballroom 2(Buffet), 1st floor
13:30-15:00	Oral 5: Summarization and Generation	Haitian Grand Ballroom 1, 1st floor
	Oral 6: Dialogue System	Qindao Function Room 1, G floor
	Oral 7: Multimodality	Meeting room 06,07, G floor
	Oral 8: Sentiment Analysis	Meeting room 08,09, G floor
15:00-15:15	Break	
15:15-16:45	Oral 9: Text Mining	Haitian Grand Ballroom 1, 1st floor
	Oral 10: NLP Applications 1	Qindao Function Room 1, G floor
	Oral 11: Question Answering and Knowlege Graph	Meeting room 06,07, G floor
	Oral 12: NLP Applications 2	Meeting room 08,09, G floor
17:00-17:10	QA for Keynote Talk 3	Haitian Grand Ballroom 1, 1st floor
17:10-17:40	Award & Closing Ceremony	Haitian Grand Ballroom 1, 1st floor
18:00-21:00	Dinner	Haitian Grand Ballroom 2(Buffet), 1st floor

4

Tutorials : Oct, 13 – 14

Schedule

Wednesday, 13 Oct (Online)

08:55-09:00	Opening Remarks	Tutorial Chairs
09:00-10:30	Lecture 1: Natural Language Question Answering over Knowledge Graph	Lei Zou
10:45-12:15	Lecture 2: Learning towards Knowledge-driven Dialogue System between Human and Computer	Rui Yan
14:00-15:30	Lecture 3: Group-Theoretic Self-Supervised Representation Learning	Hanwang Zhang
15:45-17:15	Lecture 4: OOD Sample and New Class under Weakly Supervised Scenario	Yu-Feng Li

Thursday, 14 Oct (Online)

09:00-10:30	Lecture 5: Will Non-Autoregressive Translation Dominate Neural Machine Translation?	Yang Feng
10:45-12:15	Lecture 6: Self-supervised Reinforcement Learning	Jianye Hao
14:00-15:30	Lecture 7: Generative Adversarial Nets Meet Reinforcement Learning	Weinan Zhang
15:45-17:15	Lecture 8: Robustness in Natural Language Processing	Qi Zhang & Tao Gui

Tutorial 1

Natural Language Question Answering over Knowledge Graph

Lei Zou

Wednesday, 13 Oct, 09:00 – 10:30

Abstract: As more and more structured data become available on the web, the question of how end users can access this body of knowledge becomes of crucial importance. As a de facto standard of a knowledge base, RDF repository is a collection of triples. Although SPARQL is a standard way to access RDF data, it remains tedious and difficult for end users because of the complexity of the SPARQL syntax and the RDF schema. An ideal system should allow end users to profit from the expressive power of Semantic Web standards (such as RDF and SPARQLs) while at the same time hiding their complexity behind an intuitive and easy-to-use interface. In this talk, I first review two categories of existing methods on natural language question answering (Q/A) over RDF knowledge graph — one is IR (Information Retrieval)-based and the other one is called semantic parsing method. Besides, the game changer — pre-trained language model have made great progress in NLP, including KBQA. I will also briefly review its development and impact on KBQA task. Then, I will talk about our KBQA system (gAnswer), which is based on graph matching-based technique. Our method constructs semantic query graph to represent semantics of natural language question and answer, and then utilize subgraph matching technique to address the disambiguation issue in natural language understanding and query the answer from KB. gAnswer has achieved excellent evaluation results on multiple benchmarks, and won the championship in QALD-9 knowledge graph QA system competition (hosted by HOBBIT). We share relevant source codes at Github.

Lecturer: Dr. Lei Zou is a professor in WangXuan Institute of Computer Technology of Peking University, the assistant to the director of Center for Data Science of Peking University, and a faculty member of National Engineering Laboratory for Big Data Analysis Technology and Application (PKU), China. Prof. Lei Zou's recent research interests include graph databases, knowledge graph, particularly in graph-based RDF data management, natural language question answering over knowledge graph and graph analytics and machine learning and big data systems. He has published more than 50 papers, including more than 30 papers published in reputed journals and major international conferences, such as SIGMOD, VLDB, ICDE, TODS, TKDE, VLDB Journal, AAAI, EMNLP. Lei Zou's research is supported by NSFC-Young Excellent Talent Project and National Key Research and Development Program of China. Furthermore, Prof. Zou obtained Newton Advanced Fellowships of UK Royal Society.

Tutorial 2

Learning towards Knowledge-driven Dialogue System between Human and Computer

Rui Yan

Wednesday, 13 Oct, 10:45 – 12:15

Abstract: Recently, intelligent dialogue techniques between human and computer have been improved significantly. A series of industrial applications have been developed along with the development of dialogue systems, such as Microsoft XiaoIce, AliMe from Alibaba, and Amazon Echo, etc. Intelligent human-computer dialogue systems have provided a huge space for future applications which may be beyond our imagination. However, we are still facing with a bottleneck of further progress of dialogue systems: the lack of knowledge usually leads to inadequate and awkward conversations with low intelligence. In this tutorial, we will review the current techniques to incorporate knowledge into the human-computer dialogue systems in order to make the conversations more natural and more human-like.

Lecturer: Dr. Rui Yan is an associate professor with tenure at Gaoling School of Artificial Intelligence, Renmin University of China. He was a tenure-track assistant professor at Wangxuan Institute of Computer Technology, Peking University, and a senior researcher at Baidu Inc. He was selected as a young scientist at Beijing Academy of Artificial Intelligence (BAAI) and a startrack young fellow of Microsoft Research Asia (MSRA). Till now he has published more than 100 highly peer-reviewed publications with more than 5,000 citations. He regularly served as an area chair/senior PC member for top-tier international conferences. He has been invited to give tutorial talks for these conferences as well.

Tutorial 3

Group-Theoretic Self-Supervised Representation Learning

Hanwang Zhang

Wednesday, 13 Oct, 14:00 – 15:30

Abstract: A good visual representation is an inference map from observations (images) to features (vectors) that faithfully reflects the structure and transformations of the underlying generative factors (semantics), who are invariant to environmental changes. In this paper, we formulate the notion of "good" representation from a group-theoretic view using Higgins' definition of disentangled representation, and show that existing Self-Supervised Learning (SSL) can only learn augmentation-related features such as lighting and view shifts, leaving the rest of high-level semantics entangled. To break the limitation, we propose an iterative SSL method: Iterative Partition-based Invariant Risk Minimization (IP-IRM), which successfully grounds the abstract group actions into a concrete SSL optimization. At each iteration, IP-IRM first partitions the training samples into subsets. In particular, the partition reflects an entangled semantic group action. Then, it leverages IRM to learn subset-invariant sample similarities, where the invariance guarantees to disentangle the corresponding semantic. We prove that IP-IRM converges with a full-semantic disentangled representation, and show its effectiveness on various feature disentanglement and SSL benchmarks.

Lecturer: Dr. Hanwang Zhang is an Assistant Professor at Nanyang Technological University's School of Computer Science and Engineering. His research interests include Computer Vision, Natural Language Processing, Causal Inference, and their combinations. His work has received numerous awards including the IEEE AI's-10-To-Watch 2020, TMM Prize Paper Award 2020, Alibaba Innovative Research Award 2019, ACM ToMM Best Paper Award 2018, Nanyang Assistant Professorship 2018, ACM SIGIR Best Paper Honourable Mention Award 2016, and ACM MM Best Student Paper Award 2012. Hanwang and his team work actively in causal inference for connecting vision and language. For example, their scene graph detection benchmark won the IEEE CVPR Best Paper Finalist 2019 and their visual dialog agent won the 1st place in Visual Dialog Challenge 2019 and 2nd place in 2018/2020.

Tutorial 4

OOD Sample and New Class under Weakly Supervised Scenario

Yu-Feng Li

Wednesday, 13 Oct, 15:45 – 17:15

Abstract: Machine learning algorithms tend to fail when the training and test data contain examples from unknown distribution, e.g., out-of-distribution (OOD) sample or sample from new class, which becomes one major challenge to deploy machine learning models in real-world tasks. Previous studies mainly focused on supervised or unsupervised scenarios, while the efforts on weakly supervised scenarios remain to be limited. In this talk, we briefly introduce some recent research on weakly supervised learning suffering from samples from unknown distribution. Firstly, we present some efforts on robust weakly supevised learning, and then some more accurate semi-supervised and label noise learning algorithm affected by OOD samples, respectively. Then, we present two attempts on detecting examples from new class under streaming data. Experimental results verify the superiority of our proposed approaches, and reveal the possible research direction in the future.

Lecturer: Dr. Yu-Feng Li is an associate professor of the National Key Laboratory for Novel Software Technology, Nanjing University. He received the BSc and PhD degrees in computer science from Nanjing University, China, in 2006 and 2013, respectively. His research interests include semi-supervised learning, weakly supervised learning, and optimization. He has published more than 50 papers in top-tier journals and conference proceedings. He serves as associate editor of the Machine Learning, Neural Network, etc. He serves as program co-chair of IEEE Bigcomp 2020, CCML 2021, MLA 2020, journal track co-chair of ACML 2021, workshop co-chair of ACML 2018, tutorial co-chair of ACML 2019, etc, and area chair/senior pc member of ICML, IJCAI, AAAI, ACML, PAKDD, etc.

Tutorial 5

Will Non-Autoregressive Translation Dominate Neural Machine Translation?

Yang Feng

Thursday, 14 Oct, 09:00 – 10:30

Abstract: Non-autoregressive translation (NAT) has shown its superiority in translation speed and is drawing more attention recently, but it also faces some problems, among which multimodality is the most prominent. Multimodality is the inherent problem of NMT and is further aggravated in NAT due to lacking sequential modeling. In this tutorial, we will review the development of NAT modeling, explore the problems of NAT and give our solution to multimodality.

Lectuer: Dr.Yang Feng is a Professor in Institute of Computing Technology, Chinese Academy of Sciences (ICT/CAS). She was picked into "New Baixing Talent Introduction Program" of ICT/CAS in 2017 and was selected as one of the outstanding members of CCAI in 2019 (70 people in total). Her research focuses on machine translation and dialogue and her work on machine translation has won Best Long Paper Award of ACL 2019. She has served as (Senior) Area Co-chair for top-tier conferences of NLP (e.g., ACL/EMNLP) and an Action Editor of ACL rolling review. She is a recipient of Technical Committee on NLP "Young Scientist".

Tutorial 6

Self-supervised Reinforcement Learning

Jianye Hao

Thursday, 14 Oct, 10:45 – 12:15

Abstract: Reinforcement learning has achieved great success in recent years but still suffers from sample efficiency and scalability issues, which significantly limits the wide application of RL to real application scenarios. One of major bottleneck of RL is the limited representation power in terms of both environments and policies. In this talk, I will introduce how to leverage self-supervised techniques to increase the representation power of RL from different aspects such as state, policy, action, and finally boost the learning efficiency and scalability across different scenarios and tasks.

Lecturer: Dr. Jianye Hao is Associate Professor at Tianjin University and Director of Huawei Noah's Ark Decision-making and Reasoning Lab. His research area focuses on reinforcement learning and multiagent systems. Dr. Hao has published over 100 peer-reviewed papers in top conferences and journals and won a number of best paper awards such as ASE2019, DAI2019 and CoRL2020. The research of his team has been successfully applied in various domains such as Game AI, E-commerce recommendation, network optimization, supply chain optimization and so on.

Tutorial 7

Generative Adversarial Nets Meet Reinforcement Learning

Weinan Zhang

Thursday, 14 Oct, 14:00 – 15:30

Abstract: Generative adversarial nets (GANs) have been widely studied during the recent development of deep learning and unsupervised learning. With an adversarial training mechanism, GAN manages to train a generative model to fit the underlying unknown real data distribution under the guidance of the discriminative model estimating whether a data instance is real or generated. Such a framework is originally proposed for fitting continuous data distribution such as images, thus it is not straightforward to be directly applied to information retrieval scenarios where the data is mostly discrete, such as IDs, text and graphs. In this talk, I will focus on discussing the GAN techniques and the variants on discrete data modeling with reinforcement learning techniques. Specifically, I will start the talk with the fundamentals of GAN framework and its extension on discrete data generation via reinforcement learning. Then I will discuss about the connection between GAN and imitation learning. Finally, I will discuss GAN on discrete data generation tasks, including IRGAN, SeqGAN, GraphGAN, CommunityGAN etc, and some advanced applications.

Lecturer: Dr. Weinan Zhang is now an associate professor at Shanghai Jiao Tong University. His research interests include reinforcement learning, deep learning and data science with various real-world applications of recommender systems, search engines, text mining and generation, game AI etc. He has published over 100 research papers on international conferences and journals and has been serving as a (senior) PC member at ICML, NeurIPS, ICLR, KDD, AAAI, IJCAI, SIGIR etc. and a reviewer at JMLR, TOIS, TKDE, TIST etc. He was granted as ACM Rising Star Award 2017 and Alibaba DAMO Young Scholar Award 2018. His research won the Best Paper Honorable Mention Award at SIGIR 2017, the Best Paper Award at DLP-KDD Workshop 2019 and the Best System Paper Award at CoRL 2020. Weinan earned his Ph.D. from the Computer Science Department of University College London in 2016 and B.E. from ACM Class of Shanghai Jiao Tong University in 2011.

Tutorial 8

Robustness in Natural Language Processing

Qi Zhang, Tao Gui

Thursday, 14 Oct, 15:45 – 17:15

Abstract: The recent breakthroughs in deep learning theory and technology provide strong support for the wide application of NLP technology, such as question answering systems, information extraction, and machine translation. A large number of models have emerged, of which the performances surpass that of humans when the training and test data are independent and identically distributed (i.i.d.). However, recent research had reported that highly accurate deep neural networks (DNN) can be vulnerable to carefully crafted adversarial examples distribution shift, data transformation, and shortcut learning. Recently, researchers have begun to explore ways to detect robustness prior to model deployment. In this talk, we will give some brief introduction about recent studies on robustness in NLP.

Lecturer: Dr. Qi Zhang is a professor in the school of computer science at the Fudan University. His Dr. degree in Computer Science was received from Fudan Univerisity, in 2009. Since April 2009, he has been with the School of Computer Science, Fudan University. His research focuses on natural language processing , with special emphasis on roubustness, information extraction, sentiment analysis, question answering, and so on. He has published more than 100 papers on top-tier conferences in related research fields, including ACL, EMNLP, ICML, ICLR, NIPS, IJCAI, AAAI and so on. His recent research won the WSDM Best Paper Runner Up Award, CoLING Area Chair Feavour Favorites Award. He is a recipient IBM faculty Award and ACM Shanghai Rising Star Award.

Lecturer: Dr. Tao Gui is an pre-tenured associate professor in Institute of Modern Languages and Linguistics at Fudan University. He also got his Phd at Fudan University in 2021. His research focuses on natural language processing, information extraction robust model, and model interpretability. He has published more than 20 papers on top-tier conferences in related research fields, including ACL, EMNLP, SIGIR, IJCAI, AAAI and so on.

5

3 Workshops : Oct, 15

	Session	Chair	Venue
09:00-22:00	NLPCC 2021 Onsite Registration		Lobby, 1st Floor
14:00-17:00	MMNLP Workshop	Youzheng Wu	Qindao Function Room 1, G Floor
	Baidu Workshop	Hua Wu,Yu Sun	Qindao Function Room 2, G Floor
	XAI Workshop	Yao Meng	Meeting Room 08&09, G Floor
15:30 -16:00	Break		G floor public area
17:00 -17:30	Grand Data Award		Meeting Room 08&09, G Floor
18:00 -21:00	Dinner		Haitian Grand Ballroom 2 (Buffet), 1st Floor
19:00 -21:00	TCNLP Annual Meeting (Only TCNLP Members)		Qindao Function Room 1, G Floor

MMNLP Workshop: The Second Workshop on Multimodal Natural Language Processing

14:00-17:00	Chair: Youzheng Wu
14:00-14:05	Opening
14:05-14:40	Multimodal Emotion Recognition *Qin Jin*
14:40-15:15	Multimodal Neural Machine Translation *Jinsong Su*
15:15-15:50	Vision to Language: from Independency, Interaction, to Symbiosis *Ting Yao*
15:50-16:25	Multimodal Information Extraction for Social Media Posts *Jianfei Yu*
16:25-17:00	Speech and NLP Research at Huawei Noah’s Ark Lab *Yichun Yin*

Baidu Workshop：Frontier Progresses and Applications of Pre-trained Models

14:00-17:00	Chair: Hua Wu, Yu Sun
14:00-14:10	Opening Speech：Distinguished Architect of Baidu, Leader of Baidu ERNIE platform *Yu Sun*
14:10-14:40	Keynote Speech 1：Postdoc at the Language Technologies Institute of Carnegie Mellon University *Pengfei Liu*
14:40-14:45	Q & A
14:45-15:15	Keynote Speech 2：Core member of AliceMind, Machine Intelligence laboratory of Alibaba Damo Academy *Fuli Luo*
15:15-15:20	Q & A
15:20-15:50	Keynote Speech 3：Staff Software Engineer of Baidu NLP department *Shuohuan Wang*
15:50-15:55	Q & A
15:55-16:25	Keynote Speech 4：Postdoc at Institute of Automation, Chinese Academy of Sciences *Lingxiang Wu*
16:25-16:30	Q & A
16:30-17:00	Adjourn/Networking

14:00-17:00	Chair:Yao Meng
14:00-14:10	Welcome
14:10-14:40	Keynote Speech: Construction of event semantics and the explanation of the process and the results *Endong Xun*
14:40-14:45	Q&A
14:45-15:15	Keynote Speech: Towards Interpretable and Robust Natural language Predictive Model *Baotian Hu*
15:15-15:20	Q&A
15:20-15:30	Coffee Break
15:30-15:50	Oral Session：Steffen et al. ”Detecting Covariate Drift with Explanations” from DFKI.
15:50-16:10	Oral Session：Prateek et al. ”A Data-Centric Approach Towards Deducing Bias in Artificial Intelligence Systems for Textual Contexts” from SAP.
16:10-16:25	Industrial Speeches: Intelligent Document Processing and Knowledge Graph Construction *Yunwen Chen*
16:25-16:40	Industrial Speeches:The Application of the Semantic Recognition in Enterprise Search Engine *Xi Zhang*
16:40-16:55	Industrial Speeches:Practice of Multilingual Knowledge Alignment *Cai Chao*

MMNLP Workshop: The Second Workshop on Multimodal Natural Language Processing

Oct. 15
14:00-17:00, Qindao Function Room 1 , G Floor

Multimodal Emotion Recognition

Qin Jin

14:05-14:40

Automatic emotion recognition is an indispensable ability of intelligent human-computer interaction systems. The behavior signals of human emotion expression are multimodal, including voice, facial expression, body language, bio-signals etc. Therefore, our research focuses on robust emotion recognition via fusing multimodalities. This talk will present our recent works on multimodal emotion recognition from different perspectives.

Dr. Qin Jin is a full professor in School of Information at Renmin University of China (RUC), where she leads the AI • M3 lab. She received her Ph.D. degree in 2007 at Carnegie Mellon University. Before joining RUC in 2013, she was a research faculty (2007-2012) and a research scientist (2012) at Carnegie Mellon University and IBM China Research Lab. Her research interests are in intelligent multimedia computing and human computer interaction. Her team's recent works on video understanding and multimodal affective analysis have won various awards in international challenge evaluations, including CVPR ActivityNet Dense Video Captioning challenge, NIST TrecVID VTT evaluation, ACM Multimedia Audio-Visual Emotion Challenge etc.

Multi-modal Neural Machine Translation

Jinsong Su

14:40-15:15

With the rapid development of deep learning, researches on neural machine translation (NMT) have shifted from traditional purely textual NMT to multi-modal NMT. In this report, I will first give the task definition of multi-modal NMT, summarize the typical studies in this aspect, and finally look forward to the future work.

Dr. Su Jinsong is a Professor at School of Informatics, Xiamen University. He received the Ph.D. degree in Chinese Academy of Sciences in 2011. His research interests mainly include natural language processing, neural machine translation and text generation. He has published more than 70 CCF-A/B papers, won Hanwang Youth Innovation Excellence Award and is supported by Fujian outstanding youth fund.

Vision to Language: from Independency, Interaction, to Symbiosis

Ting Yao

15:15-15:50

Vision and Language are two fundamental capabilities of human intelligence. Humans routinely perform tasks through the interactions between vision and language, supporting the uniquely human capacity to talk about what they see. That motivates us researchers to expand the horizons of such cross-modal analysis. In particular, vision to language is probably one of the hottest topics in the past five years, with a significant growth in both volume of publications and extensive applications. In this talk, we look into the problem of vision to language, from three different perspectives: 1) Independency –aim for a thorough image/video understanding for language generation; 2) Interaction –explore the (1st, 2nd, ⋯) interaction across vision and language inputs; 3) Symbiosis –learn a universal encoder-decoder structure for vision-language tasks. Moreover, we will also discuss the real-world deployments or services of vision to language.

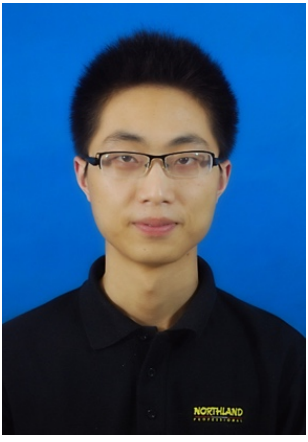
Dr. Ting Yao is currently a Principal Researcher in Vision and Multimedia Lab at JD AI Research, Beijing, China. His research interests include video understanding, vision and language, and deep learning. Prior to joining JD.com, he was a Researcher with Microsoft Research Asia, Beijing, China. Ting is the lead architect of a few top-performing multimedia analytic systems in international benchmark competitions such as ActivityNet Large Scale Activity Recognition Challenge 2019-2016, Visual Domain Adaptation Challenge 2019-2017, and COCO Image Captioning Challenge. He is also the lead organizer of MSR Video to Language Challenge in ACM Multimedia 2017 & 2016, and built MSR-VTT, a large-scale video to text dataset that is widely used worldwide. His works have led to many awards, including ACM SIGMM Outstanding Ph.D. Thesis Award 2015, ACM SIGMM Rising Star Award 2019, and IEEE TCMC Rising Star Award 2019. He is an Associate Editor of IEEE Trans. on Multimedia.

Multimodal Information Extraction for Social Media Posts

Jianfei Yu

15:50-16:25

Recent years have witnessed the explosive growth of multimodal user-generated contents on social media platforms such as Facebook, Instagram, Twitter and Snapchat. The analysis of social media posts now has to take into consideration of not only text but also other modalities of data such as images and videos. In this talk, I will first share some recent progress in multimodal information extraction on social media posts, and then report our recent work on multimodal named entity recognition and entity-level sentiment analysis. Finally, I will discuss several limitations of current studies, and point out some promising directions in the future.



Jianfei Yu is an Associate Professor at School of Computer Science and Engineering, Nanjing University of Science and Technology (NJUST). He obtained his Ph.D. from Singapore Management University (SMU) in 2018, and has worked at SMU as a Research Scientist before joining NJUST. His research focuses on deep learning and transfer learning for many natural language processing tasks, including information extraction, sentiment analysis, question answering, social media analytics. He has served on the Virtual Infrastructure Committee of ACL 2021, and given a half-day tutorial about Fine-Grained Opinion Mining in IJCAI 2019.

Speech and NLP Research at Huawei Noah's Ark Lab

Yichun Yin

16:25-17:00

The Noah's Ark Lab is the AI research center for Huawei Technologies. Its mission is to make significant contributions to both the company and society by innovating in AI. In this talk, I will introduce some important research work of Noah's Ark lab in Speech and NLP. The topics include speech processing, pretrained language models, multimodal processing, etc.

Dr. Yichun Yin is a researcher at Huawei Noah's Ark Lab. He received PhD from the School of EECS at Peking University in 2018, where he was advised by Ming Zhang. Before that, he was an undergraduate student at University of Science and Technology Beijing.

Baidu Workshop:Frontier Progresses and Applications of Pre-trained Models

Oct. 15
14:00-17:00, Qindao Function Room 2, G Floor

Introduction: Since the pre-trained language model BERT had been proposed in 2018, pre-trained models have achieved state-of-the-art results in various Natural Language Processing (NLP) tasks. Particularly, the GPT-3 model with 175 billion parameters shows its strong task-agnostic zero-shot/few-shot learning capabilities. Recently some researchers proposed the concept of the "Foundation Model" based on this efficient paradigm of "pre-training + fine-tuning", believing that the pre-trained models will profoundly change all aspects of human society.
Grasping the opportunity of rapid development of the pre-trained model, this workshop invites distinguished scholars and experts to give keynote speeches and share their cutting-edge progresses and practices in pre-trained models. The latest advances of pre-trained models in simple terms will be presented. The speakers will also show practical examples of how to apply the pre-trained model in real applications to the audience. Finally, future direction and work in the area of pre-trained model will be deeply discussed as well.

Distinguished Architect of Baidu, Leader of Baidu ERNIE platform

Yu Sun

14:00-14:10

Yu Sun,is a Distinguished Architect of Baidu and the leader of Baidu ERNIE platform. His main research topics cover natural language understanding, dialogue system, deep learning and information retrieval. He leads the research and development of Baidu language understanding technology and platform Wenxin (ERNIE)—key technology of Baidu search engine and other global leading technologies. Yu’ s works are widely used in search engine, smart speaker, map and other products, which significantly improved the user experience effecting hundreds of millions of users. He has won more than 10 worldwide champions including SemEval, the world’s largest semantic evaluation, and published dozens of papers at ACL, NAACL, AAAI, and obtained more than 70 patents. He has also pocketed SAIL (the highest award) of the World Artificial Intelligence Conference (WAIC), excellent scientific and technological achievement award of Chinese Association for Artificial Intelligence (CAAI) as well as First Prize of scientific and technological progress award of Chinese Institute of Electronics (CIE).

The Fourth Paradigm of Modern Natural Language Processing Techniques

Pengfei Liu

14:10-14:40

In this talk, Pengfei will summarize and analyze four paradigms in the development of statistical natural language processing techniques, and argue that prompt-based learning is a promising new paradigm that may represent another major change in the way we look at NLP. Then he will try to organize the current state of knowledge in this rapidly developing field by providing an overview and formal definition of prompting methods and highlight some potential challenges to motivate future works.

Pengfei Liu, is a postdoc at the Language Technologies Institute of Carnegie Mellon University and serves as a co-lecturer in the CMU Natural Language Processing course. His research topics currently focus on information extraction, text generation, and NLP system evaluation. He serves as area chairs of NAACL, EMNLP, NeurIPS, ACL Rolling Review, etc., and wins the best demo paper award in ACL 2021. He also won a CAAI outstanding doctoral dissertation, Baidu scholarship, and so on.

AliceMind: Alibaba's Collection of Encoder-decoders from Machine Intelligence of Damo Lab

Fuli Luo

14:45-15:15

AliceMind, aka. ALIbaba's Collection of Encoder-decoders from MinD (Machine intelligence of Damo) is a leading and systematic framework for pre-trained language models. AliceMind includes language understanding model StructBERT (ICLR 2020), generative language model PALM (EMNLP 2020), cross-lingual language model VECO (ACL 2021), cross-modal language model StructVBERT (CVPR 2021 VQA Challenge Champion), structural language model StructuralLM (ACL 2021). Recently, AliceMind released PLUG, a large-scale Chinese language understanding and generation model with 21 billion parameters, and opened an API for testing. The AliceMind team has published more than 20 papers. At present, the pre-trained models within the AliceMind are leading the leaderboard of major international benchmarks, including GLUE, CLUE, XTREME, VQA Challenge, DocVQA, MS MARCO, etc. AliceMind currently has the core technical competitiveness, being one of the large-scale pre-trained language model frameworks that cover technology, platform, application, ecosystem, and so on.

Fuli Luo, master of Peking University. Her research focuses on natural language understanding and generation, including language model pre-training, text style transfer, sentiment controllable text generation, and word sense disambiguation. Currently, she works on language model pre-training, model compression, fine-tuning. She is now working at Alibaba and is the core member of AliceMind, a leading framework for pre-trained language models created by the Machine Intelligence laboratory of Alibaba Damo Academy. She has published more than 20 research papers, and most of them are published in top-tier conferences and journals. She has served as the Program Committee Member of various international conferences including ACL, EMNLP, AAAI, IJCAI, etc.

The Algorithm, Platform and Application of ERNIE

Shuohuan Wang

15:20-16:10

This talk is focusing on the algorithm, platform and application of ERNIE, which is the pre-training technology of Baidu NLP. ERNIE was made a series of breakthroughs in various fields such as language understanding, language generation, cross-lingual and cross-modal understanding. ERNIE also played an important role in a series of products such as Baidu search, Baidu App and Xiaodu. In July 2021, ERNIE 3.0, large-scale knowledge enhanced pre-training for language understanding and generation, has achieved SoTA Results in 54 Chinese NLP tasks, crowned 1st place on SuperGLUE leaderboard.

Shuohuan Wang, is a Staff Software Engineer of Baidu NLP department. He received Master degree from Peking University in 2015 and Bachelor degree from Beihang University in 2012. His primary work is on pre-training in NLP and the work has achieved SoTA results in several global benchmarks such as GLUE、SuperGLUE、XTREME and SemEval. The related papers have been published on AAAI, ACL, EMNLP etc. He has more than 70 patents filed or granted.

Exploration and practice of multimodal pretrained models

Lingxiang Wu

16:15-16:45

Pre-training model has attracted intensive attention since Google proposed BERT in 2018. With "big data + large modal", the unsupervised training paradigm has demonstrated impressive performance on several downstream tasks. Zidongtaichu, proposed by Institute of Automation, Chinese Academy of Sciences, consists of three single-modal base models and a series of cross-modal encoders/decoders. In the single modal, object-level training and efficient transformer structure are explored in vision tasks. In the vision-language pretraining, knowledge is embedded to achieve better semantic understanding and generalization capability. In the three-modal works, the first pre-trained model that connects three modalities of text, vision, and audio in a unified framework is proposed. By token-, modality-, and sample- level pretext tasks, it can effectively adapt to a series of cross-modal understanding and generation tasks given single-, two-, or three-modal inputs. Zidongtaichu can support practical applications, including a cross-modal chatbot, generalized representation, image-to-audio generation, image generation etc. Intensive computing power enables large models and incubates various applications. Multi-modal pretraining is a keypoint towards the more general artificial intelligence, and it is expected to become a platform-level technology in the future. In this talk, I'll introduce the Zidongtaichu in details and provide an expectation of multi-modal pretraining in the future development.

Lingxiang Wu, is a Postdoc in Institute of Automation, Chinese Academy of Sciences, supervised by Jinqiao Wang. Receiving her Ph.D. from University of Technology, Sydney, she primarily works on vision language pretraining in multi-modal scenarios. Her works have been published on TMM, TCSVT, PR, etc. Before joining the Institute of Automation, Chinese Academy of Sciences, she worked as a data scientist for Woolworths Group Limited, and a senior RD in Baidu subsequently.

XAI Workshop: Align with Changes: A Practice of NLP from Explanation to Adaption

Oct. 15
14:00-17:00, Meeting Room 08 & 09, G Floor

Construction of event semantics and the explanation of the process and the results

Endong Xun

14:10-15:20

The report mainly introduces the construction of event semantics based on the method of knowledge, which is used to extract knowledge and establish the graphical meaning of knowledge. In the computing process, it attempts to identify the structures of the event and the entity and the semantic relationship between the two. Using symbols and parameters, the combination of the calculations ensures the explanation of the process and the results.

Prof. Endong Xun Doctoral Advisor, Dean of the School of Information Science, Beijing Language and Culture University, and dean of the Institute of Linguistics Intelligence. He received a Master's and Ph.D. from the Harbin Institute of Technology. He worked successively at the Microsoft Asia Research Institute and at the Hong Kong University of Science and Technology. His research areas are natural language processing and language education technology. In recent years, he has presided over a number of the National 863 Program and National Natural Science Foundation of China, undertook a number of cooperation projects between ministries and companies, and published dozens of articles. The main content of his research is semantic computing in Chinese and building language resources. The BCC corpus developed in charge is the most important on-line corpus of China and the preferred corpus for language research and development. He put forward the parataxis graph theory and construction method for Chinese semantic analysis, lead the team to unveil the major research tasks of artificial intelligence of the Ministry of Industry and Information Technology, and carry out the application of Chinese semantic understanding in many fields.

Towards Interpretable and Robust Natural language Predictive Model

Baotian Hu

14:10-15:20

Neural predictive models have been successfully applied in many natural language processing (NLP) tasks. However, the success has come at the expense of transparency in the model's decision process. Recently, the interpretability of deep models has attracted great attention, which has motivated efforts to interpret how these black-box models work. Previous work has shown that neural networks are easy to be attacked, which naturally brings about the question of whether the interpretability will be affected by the attack. Intuitively, a crucial property that interpretability should satisfy its robustness to local perturbations of the input. Though some recent studies have focused on linking interpretability and adversarial robustness, there is no explicit statement about existing models have both two properties. In this talk, I will present some findings on the interpretability of natural language predictive models: (a) the combination of explicit and latent domain knowledge representations have been useful for performance improvements and interpretability in machine reading comprehension. (b) making predictions can benefit from generating the explanation in the decision process. (c) robust models lead to better-extracted rationales.

Dr. Baotian Hu is an Assistant Professor in School of Computer Science and Technology of Harbin Institute of Technology (Shenzhen); His research interests lie in natural language generation, information extraction, and medical information processing, etc. He received doctorate degree in computer science and technology from Harbin Institute of Technology in 2016. In recent years, he has published more than 30 papers on top academic journals and conferences in the field (such as NeurIPS, ACL, EMNLP, AAAI, JAMIA, TNNLS, etc.) which has been cited more than 1800. Due to his outstanding contributions in the field of Chinese information processing, he has won the Excellent Doctoral Dissertation Nomination Award of the Chinese Information Society of China in 2018.

Intelligent Document Processing and Knowledge Graph Construction

Yunwen Chen

16:10-16:25

Intelligent document processing utilizes artificial intelligence technologies such as natural language processing (NLP), optical character recognition (OCR), computer vision (CV), deep learning (DL) ,and machine learning (ML) to classify, categorize and extract relevant information，and then verifies the extracted data, which enhances the ability to extract and understand the transformation of information from structured and unstructured data. Meanwhile, it supports extracting and organizing structured and unstructured documents and data in various scenarios of industries at any time, enabling enterprises to handle documents procedurally, automatically, and intelligently. Knowledge graph technology, based on the mulit-source heterogeneous data fusion from diverse industries，constructs a knowledge middle platform. Intelligent document processing,knowledge graph and robotic process automation are now widely applied to various scenarios in finance, manufacturing, communication, legal, audit, media, banking, government and other text intensive industries. In the securities industry, OCR ,RPA,natural language processing and other technologies are used to automatically analyze financial reports, unify report formats, balance trial calculations, cross-form check, information accuracy and integrity checks, report processing and information filling can also be customized to meet business requirements. In the manufacturing industry, the knowledge graph is applied to the quality management system to realize the mapping of knowledge and experience in failure documents and information.

Dr. Yunwen Chen is the founder and CEO of DataGrand. Alongside his PhD in computer science from Fudan University, he is a renowned computing expert with membership in the Association for Computing Machinery (ACM), the Institute of Electrical and Electronics Engineers (IEEE), and senior membership in the China Computer Federation (CCF). He was named one of Shanghai Pudong’s Hundred Talent’s Program and member of the national committee of CPPCC, and was awarded the Shanghai Excellent Doctoral Thesis Award. He is the owner of more than 30 national invention patents, has participated in and won the international data mining competition multiple times, and translated the book ”Algorithms of the Intelligent Web” into Chinese. Yunwen previously served as the chief data officer of Shanda Literature, the senior director of Tencent Literature, and the core technology research and development engineer of Baidu. His areas of expertise are machine learning, natural language processing, search engine and recommendation –fields in which he has extensive research and engineering experience.

The Application of the Semantic Recognition in Enterprise Search Engine

Xi Zhang

16:25-16:40

In the talk, I’ ll be trying to explain the differences between the traditional keyword search and the semantic search, and how the NLP technology is applied in the data ingestion, the index building and the ranking on the search results. Apart from the technology utilizing, I’ ll also share an example to show how is this state-of-the-art search engine capable of helping the enterprise to leverage more interesting search scenarios.

Xi Zhang is now the head of AI Strategic Group in Laiye with the master degree of IT, specialized in data science, from University of New South Wales, Australia. He is the member of ACS(Australian Computer Society) and EA(Engineers Australia) since 2016. He and his team have been serving over a hundred enterprise customers of various industries with the AI solutions of ASR, NLU, TTS, CV and OCR, focusing on the application of the NLP technologies ranging from conversational AI, text extraction, document classification to semantic search, machine reading comprehension and insights engine. He is also the GM of the Laiye Conversational AI Platform, prospecting for productizing the NLP technologies to serve as many customers as possible in the worldwide digital transformation trend.

Practice of Multilingual Knowledge Alignment

Cai Chao

16:40-16:55

In the task of building realistic knowledge graph, the knowledge we want to mine is often hidden in the text of various languages, the more influential the object entity is，the more languages it refers. To build a powerful globalized cross-cultural knowledge graph, we need to extract and incorporate clues from multilingual text. This report focuses on three knowledge alignment methods, based on graph structure, based on entity attributes, and paragraph-based approaches, which apply different contexts of knowledge in text and graph to address the needs of different scenarios.

Dr. Cai Chao is Engineer, Director of Big Data Research Center of Global Tone Communication Technology Co., Ltd. He mainly works in the field of science and technology big data, knowledge graph, language semantics, etc., published more than 10 SCI, EI search papers, applied for 9 invention patents, won various types of technical achievements award more than 10 items. He led a team to build a large-scale economic-technology knowledge graph, which integrated global scientific and technological data, industrial data and enterprise data. The work connected investment and entrepreneurship, S & T research and industrialization, enterprise cultivation and listing and other links. Based on the knowledge graph，some software such as Science and Technology Evaluation System, Urban Science and Technology Innovation Service System was developed, which achieved good economic and social benefits.

6

Main Conference & 2 Workshops : Oct, 16

	Session		Venue
09:00-12:00	NLPCC 2021 Onsite Registration		Lobby
08:30-09:00	NLPCC 2021 Opening Ceremony		Haitian Grand Ballroom 1, 1st floor
09:00-10:00	Keynote Talk 1, by Rada Mihalcea		
10:00-10:15	Industrial Talk 1, by Junlan Feng		
10:15-10:45	Break & Group Photo		Outdoor plaza
10:45-11:45	Keynote Talk 2, by Yanchao Bi		Haitian Grand Ballroom 1, 1st floor
11:45-12:00	Industrial Talk 2, by Min Chu		
12:00-13:30	Lunch		Haitian Grand Ballroom 2, 1st floor
13:30-15:30	Sessions & Workshop	Chair	Venue
	Posters Session 1	-	G floor public area
	Oral 1: Fundementals of NLP	Jiajun Zhang	Haitian grand Ballroom 1, 1st floor
	Oral 2: Information Extraction	Ruifeng Xu	Meeting room 06,07, G floor
	Student Workshop	Fang Kong	Meeting room 08,09, G floor
15:30-16:00	Break		
16:00-18:00	Sessions & Workshop	Chair	Venue
	Posters Session 2	-	G floor public area
	Oral 3: Machine Translation and Multilinguality	Hao Zhou	Haitian Gand Ballroom 1, 1st floor
	Oral 4: Machine Learning	Wayne Xin Zhao	Meeting room 06,07, G floor
	Evaluation Workshop	Jiajun Zhang	Meeting room 08,09, G floor
18:00-21:00	Dinner		Haitian Grand Ballroom 2, 1st floor

Keynote Talk 1

Oct. 16 , Chair: *Tim Baldwin*
09:00-10:00, Haitian Grand Ballroom 1, 1st floor

Towards People-centric Word Representations

Rada Mihalcea

Most of the work to date in natural language processing has relied on general purpose language representations, obtained from training one model on very large text collections. While this approach is effective for those people whose language style is well represented in the data, it quickly fails when it comes to the language spoken by those whose style diverges from the majority. In this talk, I will describe research that points to the value of people-centric language representations, and show that we can effectively use information about the people behind the words to build better natural language processing models.

Rada Mihalcea is the Janice M. Jenkins Collegiate Professor of Computer Science and Engineering at the University of Michigan and the Director of the Michigan Artificial Intelligence Lab. Her research interests are in computational linguistics, with a focus on lexical semantics, multilingual natural language processing, and computational social sciences. She serves or has served on the editorial boards of the Journals of Computational Linguistics, Language Resources and Evaluations, Natural Language Engineering, Journal of Artificial Intelligence Research, IEEE Transactions on Affective Computing, and Transactions of the Association for Computational Linguistics. She was a program co-chair for EMNLP 2009 and ACL 2011, and a general chair for NAACL 2015 and *SEM 2019. She currently serves as ACL President. She is the recipient of a Presidential Early Career Award for Scientists and Engineers awarded by President Obama (2009), an ACM Fellow (2019) and a AAAI Fellow (2021). In 2013, she was made an honorary citizen of her hometown of Cluj-Napoca, Romania.

Keynote Talk 2

Oct. 16 , Chair: *Yue Zhang*
10:45-11:45, Haitian Grand Ballroom 1,1st floor

Semantic representation in the human brain

Yanchao Bi

Human brain stores tremendous amount of knowledge about this world, which is the foundation of language, thought, and reasoning. What's the neural codes of semantic knowledge representation? Is the knowledge "roses are red" simply the memory trace of perceiving the color of roses, stored in the brain circuits within color-sensitive neurons? What about knowledge that is not directly perceived by senses, such as "freedom" or "rationality"? What's their relationship with machine-based knowledge representation? I will present some work from my lab that addresses this issue using cognitive, neuroimaging, and neuropsychological methods with healthy subjects, individuals with sensory deprivation (blind and deaf) or with brain damage. The findings point to a highly distributed system incorporating two different types of information coding –one based on distributed sensory experiences (embodied) and one based on language (symbolic).

Yanchao Bi is a ChangJiang professor in IDG/McGovern Institute for Brain Research and the State Key Laboratory of Cognitive Neuroscience and Learning, at Beijing Normal University. She received her PhD from the Department of Psychology, Harvard University in 2006. Her current work focuses on the study of functional and neural architecture associated with semantic memory and language, using cognitive, neuropsychological and neuroimaging methods. She serves on the board of "Society of Neurobiology of Language", the editorial board of Journals Elife, Cognition, Cognitive Neuropsychology, and Neurobiology of Language. She has won various awards, scholarships or recognitions such as "The National Science Fund for Distinguished Young Scholars", "The National Science Fund for Excellent Young Scholars", Sackler scholar of psychophysiology, Fulbright scholar, and "rising star" in the Observer by the American psychological association.

Industrial Talk 1

Oct. 16, Chair: *DongYan Zhao*
10:00-10:15, Haitian Grand Ballroom 1, 1st floor

Beyond Open Source: Jiutian BiSheng Platform

Junlan Feng
China Mobile

Open-source software(OSS) has made fundamental contribution to nearly every issue in AI, though it is largely absent from open discussions. There are thousands of implementations of machine learning and deep learning algorithms/frameworks . Examples are Scikit-learn, R , Tensorflow, PyTorch, PaddlePaddle, Spark, Hadoop, etc. They are not simply some of the tools, they are the best AI/BigData tools. What is possible beyond Open Source? We believe Open Platform, Open Development, Open Education are part of the answer. Jiutian BiSHeng Platform(JBP) is designed for students, researchers and teachers. It provides workable instances for popular open source projects, particularly AI projects and saves developers' time from setting it up, sorting out mismatches between OS versions and open source projects, fixing known and unknown bugs commonly seen in open source software. It provides tools and computing resources to support teachers to teach in a digital coding room, where the teacher and the students in the same course can share source codes and co-development environment. It provides students an open platform to conveniently learn and research. This talk will explain how JBP works.

Junlan Feng Vice Chairman of the China Artificial Intelligence Industry Alliance, Chief Scientist at China Mobile，General Manager of AI and Intelligent Operation R&D Center，Board Chair of Linux Foundation Network.Dr. Feng received her Ph.D. on Speech Recognition from Chinese Academy of Sciences, and joined AT&T Labs Research in 2001, as a principal researcher on Speech recognition, language understanding and data mining until 2013. Dr. Feng has led R&D of China Mobile on artificial intelligence and intelligent operation since September 2013. She is an IEEE senior member. She had served as an IEEE speech and language committee member and an IEEE industry committee member. She is a frequent reviewer and organizer for major data mining, speech, and natural language international conferences and journals. Dr. Feng has over 100 professional publications and co-authored a book. She holds 51 issued U.S and international patents and 3 issued Chinese patents. She has 46 pending patent applications.Dr. Feng has won more than 20 domestic and foreign R&D awards, including AT&T CTO Award in 2009. "JiuTian", the AI platform developed by her team won the single product gold award of China Mobile in 2019, the second prize of scientific and technological achievements of China Institute of Communications, the third prize of scientific and technological progress of China

Institute of Electronics, and the 2018 Innovation Project Award of Deep Integration of Artificial Intelligence and Real Economy by the Ministry of Industry and Information Technology.

Industrial Talk 2

Oct. 16, Chair: DongYan Zhao
11:45-12:00, Haitian Grand Ballroom 1, 1st floor

Scenario Driven Innovation-Modeling and Improving the Sales Process by Discourse Analysis and Training-bot

Min Chu
AISpeech

Businesses that sale high value products such as cars, apartments or houses often spend a lot in advertising through various channels and the cost for win a customer keeps increasing. However, the sales management team normally has limited method to monitor the in-house sales process, which is the last and most important stage for closing the deal. They don't know whether the salesperson is friendly and professional enough, or whether they have overcommitments. This presentation introduces our innovative solution for such scenarios, which digitalizes the sales process, constructs a three-tier sales model, discovers pieces of outstanding examples, builds up training bots with these examples and setups evaluation matrix for the sales process. Many unseen challenges emerge when we go deeper in the frontier, some are solved or partially solved, more are waiting for deep exploring.

Dr. Chu joined AISpeech as a vice present in 2017. She built up AIspeech Beijing R&D center from scratch. The R&D center focused on developing key technologies in knowledge-driven spoken dialogue systems and exploring new business and application opportunities, and later incubated the Business Unit for Intelligent Services.
Before joining AISpeech, Dr. Chu spent 8 years with Alibaba, leading various R&D efforts in big data, search and speech technologies and products. She initiated the iDST speech interaction team which supports the speech needs within Alibaba and Alipay (Yun OS, Alipay, Taobao, DingDing etc).
Before Alibaba, Dr. Chu worked with MSRA for about 10 years. Her main research interests are in ASR, TTS, NLP, machine learning, big data etc. She has published 100+ academic papers and applied 50+ patents.

Oral Paper Sessions

Oral Paper Session 1: Fundementals of NLP (Haitian Grand Ballroom 1, 1st floor)
Session Chair: *Jiajun Zhang*

13:30-13:45	Coreference Resolution: Are the eliminated spans totally worthless? *Xin Tan, Longyin Zhang and Guodong Zhou*
13:45-14:00	Chinese Macro Discourse Parsing on Dependency Graph Convolutional Network *Yaxin Fan, Feng Jiang, Xiaomin Chu, Peifeng Li and Qiaoming Zhu*
14:00-14:15	CTRD: A Chinese Theme-Rheme Discourse Dataset *Biao Fu, Yiqi Tong, Dawei Tian, Yidong Chen, Xiaodong Shi and Ming Zhu*
14:15-14:30	Multi-Level Cohesion Information Modeling for Better Written and Dialogue Discourse Parsing *Jinfeng Wang, Longyin Zhang and Fang Kong*
14:30-14:40	Break
14:40-14:55	ProPC: A Dataset for In-domain and Cross-Domain Proposition Classification Tasks *Mengyang Hu, Pengyuan Liu and Lin Bo*
14:55-15:10	Predicting Categorial Sememe for English-Chinese Word Pairs via Representations in Explainable Sememe Space *Baoju Liu, Lei Hou, Juanzi Li and Xin Lv*
15:10-15:25	汉语近义词辨析知识库构建研究 *Juan Li*
15:25-15:40	大规模中文具体度词典的构建及具体度推理技术初探 *Zhipeng Xie and Ran Bi*

Oral Paper Session 2: Information Extraction (Meeting room 06,07, G floor)
Session Chair: *Ruifeng Xu*

13:30-13:45	Employing Multi-granularity Features to Extract Entity Relation in Dialogue *Qiqi Wang and Peifeng Li*
13:45-14:00	Entity-Aware Relation Representation Learning for Open Relation Extraction *Zihao Liu, Yan Zhang, Huizhen Wang and Jingbo Zhu*
14:00-14:15	HAIN: Hierarchical Aggregation and Inference Network for Document-Level Relation Extraction *Nan Hu, TaoLin Zhang, ShuangJi Yang, Wei Nong and XiaoFeng He*
14:15-14:30	ReMERT: Relational Memory-based Extraction for Relational Triples *Chongshuai Zhao, Xudong Dai, Lin Feng and Peng Liu*
14:30-14:40	Break
14:40-14:55	Employing Sentence Compression to improve Event Coreference Resolution *Xinyu Chen, Sheng Xu, Peifeng Li and Qiaoming Zhu*
14:55-15:10	BRCEA: Bootstrapping Relation-aware Cross-lingual Entity Alignment *Yujing Zhang, Feng Zhou and X.Y. Li*
15:10-15:25	Recognition of Nested Entity with Dependency Information *Yu Xia and Fang Kong*
15:25-15:40	Incorporate Lexicon into Self-training: A Distantly Supervised Chinese Medical NER *Zhen Gan, Zhucong Li, Baoli Zhang, Yubo Chen, Jing Wan, Kang Liu, Jun Zhao, Yafei Shi and Shengping Liu*

Oral Paper Session 3: Machine Translation and Multilinguality (Haitian Grand Ballroom 1, 1st floor)
Session Chair: *Hao Zhou*

16:00-16:15	Contrastive Learning for Machine Translation Quality Estimation *Hui Huang, Hui Di, Jian Liu, Yufeng Chen, Kazushige Ouchi and Jinan Xu*
16:15-16:30	Improving Non-Autoregressive Machine Translation with Soft-Masking *Shuheng Wang, Shumin Shi and Heyan Huang*
16:30-16:45	Adaptive Transformer for Multilingual Neural Machine Translation *Junpeng Liu, Kaiyu Huang, Jiuyi Li, Huan Liu and Degen Huang*
16:45-17:00	Guwen-UNILM: Machine Translation Between Ancient and Modern Chinese Based on Pre-Trained Models *Zinong Yang, KeJia Chen and Jingqiang Chen*
17:00-17:10	Break
17:10-17:25	Learning to Select Relevant Knowledge for Neural Machine Translation *Jian Yang, Juncheng Wan, Shuming Ma, Haoyang Huang, Dongdong Zhang, Yong Yu, Zhoujun Li and Furu Wei*
17:25-17:40	A 融合小句对齐知识的汉英神经机器翻译 *Guoyi Miao, Mingtong Liu, Yufeng Chen, Jinan Xu, Yujie Zhang and Wenhe Feng*
17:40-17:55	融合双通道音节特征的藏文 La 格例句自动分类模型 *Mabao Ban, Rangjia Cai and Mazhaxi ZGUO*

Oral Paper Session 4: Machine Learning (Meeting room 06,07, G floor)
Session Chair: *Wayne Xin Zhao*

16:00-16:15	RAST: A Reward Augmented Model for Fine-Grained Sentiment Transfer *Xiaoxuan Hu, Hengtong Zhang, Wayne Xin Zhao, Yaliang Li, Jing Gao and JiRong Wen*
16:15-16:30	Sentence-State LSTMs for Sequence-to-Sequence Learning *Xuefeng Bai, Yafu Li, Zhirui Zhang, Mingzhou Xu, Boxing Chen, Weihua Luo, Derek F. Wong and Yue Zhang*
16:30-16:45	Exploring Generalization Ability of PretrainedLanguage Models on Arithmetic and Logical Reasoning *Cunxiang Wang, Boyuan Zheng, Yuchen Niu and Yue Zhang*
16:45-17:00	AutoNLU: Architecture Search for Sentence and Cross-sentence Attention Modeling with Re-designed Search Space *Wei Zhu*
17:00-17:10	Break
17:10-17:25	AutoTrans: Automating Transformer Design via Reinforced Architecture Search *Wei Zhu, Xiaoling Wang, Yuan Ni and Guotong Xie*
17:25-17:40	Accelerating Pretrained Language Model Inference Using Weighted Ensemble Self-Distillation *Jun Kong, Jin Wang and Xuejie Zhang*
17:40-17:55	A Word-level Method for Generating Adversarial Examples using Whole-sentence Information *Yufei Liu, Dongmei Zhang, Chunhua Wu and Wei Liu*

Posters Session

Posters Session 1 (G floor public area)

P1-1	SynXLM-R: Syntax-enhanced XLM-R in Translation Quality Estimation *Bin Ni, Xiaolei Lu and Yiqi Tong*
P1-2	Advertisement Extraction from Content Marketing Articles via Segment-aware Sentence Classification *Xiaoming Fan and Chenxu Wang*
P1-3	基于层次化分类模型的电力文本分类方法 *Xiaona Chen, Pengfei Gao and Yinglong Ma*
P1-4	Memetic Federated Learning for Biomedical Natural Language Processing *Xinya Zhou, Conghui Tan, Di Jiang, Bosen Zhang, Si Li, Yajing Xu, Qian Xu and Sheng Gao*
P1-5	基于跨语种预训练语言模型 XLM-R 的神经机器翻译方法 *Qian Wang, Maoxi Li, Shuixiu Wu and Mingwen WANG*
P1-6	FABERT: A Feature Aggregation BERT-Based Model for Document Reranking *Xiaozhi Zhu, Leung-Pun Wong, Lap-Kei Lee, Hai Liu and Tianyong Hao*
P1-7	Capturing Global Informativeness in Open Domain Keyphrase Extraction *Si Sun, Zhenghao Liu, Chenyan Xiong, Zhiyuan Liu and Jie Bao*
P1-8	Event Argument Extraction via a Distance-Sensitive Graph Convolutional Network *Lu Dai, Bang Wang, Wei Xiang and Yijun Mo*
P1-9	Exploit Vague Relation: An Augmented Temporal Relation Corpus and Evaluation *Liang Wang, Sheng Xu, Peifeng Li and Qiaoming Zhu*
P1-10	Searching Effective Transformer for Seq2Seq Keyphrase Generation *Yige Xu, Yichao Luo, Yicheng Zou, Zhengyan Li, Qi Zhang, Xipeng Qiu and Xuanjing Huang*
P1-11	基于异质信息网络元路径的药物-靶标相互作用预测模型 *Yiming Liao, Chunping Ouyang, Yongbin Liu and Fuyu Hu*
P1-12	Generating Relevant, Correct and Fluent Answers in Natural Answer Generation *Yongjie Huang, Meng Yang and Yang Ni*
P1-13	Prerequisite Learning with Pre-trained Language and Graph Embedding Models *Bangqi Li, Boci Peng, Yifeng Shao and Zhichun Wang*

P1-14 GeoCQA: A Large-scale Geography-Domain Chinese Question Answering Dataset from Examination
Zhen Cui, Bin Wang and Jiangzhou Ju

P1-15 面向问答领域的数据增强方法
Jiajie Ding, Kang Xiao, Heng Ye, Xiabing Zhou and Min Zhang

P1-16 融合角色指代的多方对话关系抽取方法研究
Yang Xu, jiang yuru, yuyao zhang and He Weikai

P1-17 Enhancing Model Robustness via Lexical Distilling
Wentao Qin and Dongyan Zhao

P1-18 Multi-stage Multi-modal Pre-training for Video Representation
Chunquan Chen, Lujia Bao, Weikang Li, Xiaoshuai Chen, Xinghai Sun and Chao Qi

P1-19 Nested Causality Extraction on Traffic Accident Texts as Question Answering
Gongxue Zhou, Weifeng Ma, Yifei Gong, Liudi Wang, Yaru Li and Yulai Zhang

P1-20 Task 1 - Argumentative Text Understanding for AI Debater (AIDebater)
Li Yu Ming, Xia Mao Jin and Wang Yi Dong

P1-21 Two Stage Learning for Argument Pairs Extraction
Shuheng Wang, Zimo Yin, Wei Zhang, Derong Zheng and Xuan Li

P1-22 Context-Aware and Data-Augmented Transformer for Interactive Argument Pair Identification
Yuanling Geng, Shuqun Li, Fan Zhang, Shaowu Zhang, Liang Yang and Hongfei Lin

P1-23 ARGUABLY @ AI Debater-NLPCC 2021 Task 3: Argument Pair Extraction from Peer Review and Rebuttals
Guneet Singh Kohli, Prabsimran Kaur, Muskaan Singh, Tirthankar Ghosal and Prashant Rana

P1-24 Sentence Rewriting for Fine-Tuned Model Based on Dictionary: Taking the Track 1 of NLPCC 2021 Argumentative Text Understanding for AI Debater as an Example
Pan He, Yan Wang and Yanru Zhang

P1-25 Knowledge Enhanced transformers System for Claim Stance Classification
Xiangyang Li, Zheng Li and Sujian Li

Posters Session 2 (G floor public area)

P2-1	BERT-based Meta-learning Approach with Looking Back for Sentiment Analysis of Literary Book Reviews *Hui Bao, Kai He, Xuemeng Yin, Xuanyu Li, Xinrui Bao, Haichuan Zhang, Jialun Wu and Zeyu Gao*
P2-2	Generating Informative Dialogue Responses with Keywords-Guided Networks *Heng-Da Xu, Xian-Ling Mao, Zewen Chi, Fanshu Sun, Jingjing Zhu and Heyan Huang*
P2-3	ISWR: an Implicit Sentiment Words Recognition Model Based on Sentiment Propagation *Qizhi Li, Xianyong Li, Yajun Du and Xiaoliang Chen*
P2-4	Towards unifying the explainability evaluation methods for NLP *Diana Lucaci and Diana Inkpen*
P2-5	引入图像信息的多模态复述生成模型 *Chao Ma, Zhang Wan, YuJie Zhang, Jinan Xu and Yufeng Chen*
P2-6	Zero-Shot Deployment for Cross-Lingual Dialogue System *Lu Xiang, Yang Zhao, Junnan Zhu, Yu Zhou and Chengqing Zong*
P2-7	基于语境相似度的中文分词一致性检验研究 *Wei Liu, Kaiyu Huang, Hao Yu and Degen Huang*
P2-8	Background Semantic Information Improves Verbal Metaphor Identification *Wentao Qin and Dongyan Zhao*
P2-9	An Aspect-Centralized Graph Convolutional Network for Aspect-based Sentiment Classification *Weixiang Zhao, Yanyan Zhao, Xin Lu and Bing Qin*
P2-10	MultiWOZ 2.3: A multi-domain task-oriented dialogue dataset enhanced with annotation corrections and co-reference annotation *Ting Han, Ximing Liu, Ryuichi Takanobu, Yixin Lian, Chong Hsuan Huang, Dazhen Wan, Wei Peng and Minlie Huang*
P2-11	Variational Autoencoder with Interactive Attention for Affective Text Generation *Ruijun Chen, Jin Wang and Xuejie Zhang*
P2-12	Syntax and Coherence - The Effect on Automatic Argument Quality Assessment *Xichen Sun, WenHan Chao and Zhunchen Luo*
P2-13	EmoDialoGPT: Enhancing DialoGPT with Emotion *Yuxiang Jia, Shuai Cao, Changyong Niu, Yutuan Ma, Hongying Zan, Rui Chao and weicong zhang*
P2-14	ExperienceGen 1.0: A Text Generation Challenge which Requires Deduction and Induction Ability *Hu Zhang, pengyuan liu, Dong Yu and sanle zhang*

P2-15	CUSTOM: Aspect-Oriented Product Summarization for E-Commerce *Jiahui Liang, Junwei Bao, Yifan Wang, Youzheng Wu, Xiaodong He and Bowen Zhou*
P2-16	融合上下文的残差门卷积实体抽取 *Fenglong Su*
P2-17	Enhanced Few-shot Learning with Multiple-Pattern-Exploiting Training *jiali Zeng, Yufan Jiang, Shuangzhi Wu and Mu Li*
P2-18	MSDF: A General Open-Domain Multi-Skill Dialog Framework *Yu Zhao, Xinshuo Hu, Yunxin Li, Baotian Hu, Dongfang Li, Sichao Chen and Xiaolong Wang*
P2-19	RoKGDS: A Robust Knowledge Grounded Dialog System *Jun Zhang,Yuxiang Sun, Yushi Zhang,Weijie Xu , JiaHao Ying, Yan Yang, Man Lan, Meirong Ma, Hao Yuan and Jianchao Zhu*
P2-20	TKB²ert: Two-stage Knowledge Infused Behavioral Fine-tuned BERT *Jiahao Chen, Zheyu He, Yujin Zhu and liang Xu*
P2-21	A Unified Information Extraction System Based on Role Recognition and Combination *Yadong Zhang and Man Lan*
P2-22	A Simple but Effective System for Multi-format Information Extraction *Feiliang Ren, Yaduo Liu, Longhui Zhang, Xiaofeng Zhao and Shujuan Yin*
P2-23	The Solution of Xiaomi AI Lab to the 2021 Language and Intelligence Challenge: Multi-Format Information Extraction Task *wen dai, xinyu hua, rongrong lv, ruipeng bo and shuai chen*
P2-24	A Unified Platform for Information Extraction with Two-stage Process *Chongshuai Zhao, Dongjie Guo, Xudong Dai, Chengmin Gu, Lingling Fa and Peng Liu*
P2-25	The Overview of 2021 Language and Intelligence Challenge *Shangwen Lv, Hongyu Li, Zhengyu Niu, Quan Wang, Jing Liu, Hua Wu*

Student Workshop

Chair: Fang Kong
Meeting room 08,09, G floor

Technology and Application of Conversational AI

Shansong Yang (Hisense Visual Technology Co., Ltd)
13:30-14:00

Abstract: In this talk, we present the challenges, solutions, and trends of natural language processing in conversational interaction from industrial application perspective. Although deep learning has made significant progress, in practical applications, the conversational agents still face the challenges of language expression diversity, ambiguity, and scenario transfer. Therefore, a comprehensive use of natural language processing, knowledge graphs, multi-modal modeling is an important way to realize conversational intelligence.

Dr. Shansong Yang is the technical director of Hisense Visual Technology Co., Ltd. He received his Ph.D. from Zhejiang University in 2016. His research interests include natural language understanding, knowledge graph and conversational AI. Currently he is leading a team whose focus is research and development of natural language processing, text-to-speech, automatic speech recognition and their industrial applications.

NLP Technologies Applied on Language and Beyond

Zhaoyan Ming (Zhejiang University City College)
14:00-14:30

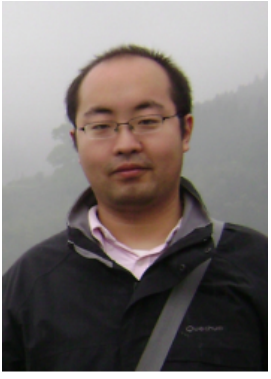
Abstract: The linguistic research community has seen tremendous advances in NLP technologies in recent years. At the same time, the applications of NLP on a diverse type of data are receiving significant attention and success. These include data in the form of languages and other types of sequential data, such as time series and spatial series. In this talk, we will show some typical NLP applications in multi-disciplinary research. We hope to inspire more young researchers to take up NLP research and expand its horizon to broad spaces.

Dr. Zhaoyan Ming is a researcher at College of Computing and Computer Science, Zhejiang University City College (ZUCC). She received her Ph.D. degree in 2011 at National University of Singapore (NUS). Before joining ZUCC, she led the wellness project at NUS-Tsinghua-Southampton Center of Extreme Search. Her research interests are in natural language processing, multimedia analysis, and their applications in health and life science. She has published 30+papers in top conferences and premier journals, including SIGIR, EMNLP, TKDE, TOIS, Multimedia Modeling, Biomedical Signal Processing and Control, etc., and two granted international patents licensed to industry.

Sentiment Forecasting in Dialogue

Zhongqing Wang (Soochow University)
14:30-15:00

Abstract: Sentiment forecasting in dialogue aims to predict the polarity of next utterance to come, and can help speakers revise their utterances in sentimental utterances generation. However, the polarity of next utterance is normally hard to predict, due to the lack of content of next utterance (yet to come). In this study, we propose a Neural Sentiment Forecasting (NSF) model to address inherent challenges. The motivations behind are that, since next utterance is yet to come, it would be helpful if we can simulate next utterance from the context. Moreover, the polarity of next utterance can be much influenced by the context, it is necessary to consider the influence of the context for next utterance. Therefore, we employ a neural simulation model to simulate the next utterance based on the context (previous utterances encountered). Moreover, we employ a sequence influence model to learn both pair-wise and seq-wise influence.



Zhongqing Wang is Associate Professor at Soochow University. He obtained the Ph.D degree from Soochow University supervised by Prof. Guodong Zhou. His research interests lie in the area of Information Extraction, Sentiment Analysis, and Social Computing. He has published more than 20 papers in top-tier Conference and journals, including IJCAI, ACL, EMNLP, COLING, and TASLP. His has severed on the program committees/reviewer boards of several prestigious international journals and conferences, including TASLP, AAAI, IJCAI, ACL etc.

Recent Advances in Unified NER

Fei Li (Wuhan University)
15:00-15:30

Abstract: Named entity recognition (NER) is a fundamental task of natural language processing (NLP), which is usually treated as a sequence labeling problem. Unfortunately, entities are not always regular text spans and some irregular entities may be overlapped or nested with each other, or consist of several discontinuous spans. Overlapped and discontinuous entities bring new challenges and opportunities for NER, since traditional sequence labeling models are not applicable. In this talk, I will introduce some recent advances for unified NER, where flat, overlapped and discontinuous entities can be recognized in one model. First, I will introduce span-based methods for unified NER. Such method first builds a span graph and then entities can be decoded out from this graph. Second, I will introduce seq2seq-based methods for unified NER, where an input sentence is first encoded by an encoder and then entities are generated by a decoder. Both methods achieve competitive results on the widely-used datasets for unified NER. These models help us promote the development of NER in NLP.

Fei Li is working as the associate researcher and master supervisor in the School of Cyber Science and Engineering of Wuhan University. He is the member of Youth Work Committee of Chinese Information Processing Society of China, the member of Youth Work Committee of Artificial Intelligence Association of China. His research interests include information extraction, text classification, deep learning, and data mining in biomedical fields. He has published more than 30 papers in the field of natural language processing, artificial intelligence and other high-level journals and conferences (such as ACL, AAAI, IJCAI, EMNLP, TASLP, IPM, etc.). In recent years, he has presided one humanities and social science research project of the Ministry of Education, one Natural Science Foundation of Hubei Province, one independent scientific research project of Wuhan University, and participated in several National Natural Science Foundation projects. He also serves as a reviewer of ACL, EMNLP, AAAI and other international conferences and journals.

Evaluation Workshop

Evaluation Workshop Session 1:AIDebater (Meeting Room 08,09, G floor)
Chair: *Jiajun Zhang*

16:00-16:10	Overview of Argumentative Text Understanding for AI Debater Challenge *Jian Yuan, Liying Cheng, Ruidan He, Yinzi Li, Lidong Bing, Zhongyu Wei, Qin Liu, Chenhui Shen, Changlong Sun, Luo Si, Changjian Jiang and Xuanjing Huang*
16:10-16:20	Distant Finetuning with Discourse Relations for Stance Classification *Lifeng Jin, Kun Xu, Linfeng Song and Dong Yu*
16:20-16:30	ACE: A Context-Enhanced Model for Interactive Argument Pair Identification *Yi Wu and Pengyuan Liu*
16:30-16:40	A Hierarchical Sequence Labeling Model for Argument Pair Extraction *Jingyi Sun, Qinglin Zhu, Jianzhu Bao, Jipeng Wu, Caihua Yang, Wang Rui and Ruifeng Xu*
16:40-16:50	Break

Evaluation Workshop Session 2:FewCLUE (Meeting Room 08,09, G floor)
Chair: *Jiajun Zhang*

16:50-17:00	Few-shot Learning for Chinese NLP tasks *Liang Xu, Xiaojing Lu, Chenyang Yuan, Xuanwei Zhang, Yuan hu, Huilin Xu, Guoao Wei, Xiang Pan, Junyi Li, Jianlin Su, Zhenyu Yang, Renfen Hu and Hai Hu*
17:00-17:10	When Few-shot Learning Meets Large-scale Knowledge-enhanced Pre-training: Alibaba at FewCLUE *Ziyun Xu, Chengyu Wang, Peng Li, Yang Li, Ming Wang, Boyu Hou, Minghui Qiu, Chengguang Tang and Jun Huang*
17:10-17:20	Break

Evaluation Workshop Session 3:AutoIE 2 (Meeting Room 08,09, G floor)
Chair: *Jiajun Zhang*

17:20-17:30	Overview of the NLPCC 2021 Shared Task: AutoIE2 *Weigang Guo, Xuefeng Yang, Xingyu Bai, Taiqiang Wu, Weijie Liu, Zhe Zhao, Qi Ju and Yujiu Yang*
17:30-17:40	BIT-Event at NLPCC-2021 Task 3: Subevent Identification via Adversarial Training *Xiao Liu, Ge Shi, Bo Wang, Changsen Yuan, Heyan Huang, Chong Feng and Lifang Wu*
17:40-18:00	Break

7

Main Conference : Oct, 17

	Session		Venue
09:00-10:00	Keynote Talk 3, by Sebastian Riedel		Haitian Grand Ballroom 1(Online), 1st floor
10:00-10:15	Industrial Talk 3, by Songfang Huang		Haitian Grand Ballroom 1(Online), 1st floor
10:15-10:45	Break		
10:45-11:45	Keynote Talk 4, by Graham Neubig		Haitian Grand Ballroom 1(Online), 1st floor
11:45-12:00	Industrial Talk4, by Yang Wu		Haitian Grand Ballroom 1(Online), 1st floor
12:00-13:30	Lunch		Haitian Grand Ballroom 2(Buffet), 1st floor
13:30-15:00	Oral Paper Sessions 5-8	Chair	Venue
	Summarization and Generation	Zhaochun Ren	Haitian Grand Ballroom 1, 1st floor
	Dialogue System	Pengjie Ren	Qindao Function Room 1, G floor
	Multimodality	Zhongyu Wei	Meeting room 06,07, G floor
	Sentiment Analysis	Liang Yang	Meeting room 08,09, G floor
15:00-15:15	Break		
15:15-16:45	Oral Paper Sessions 9-12	Chair	Venue
	Text Mining	Tianyong Hao	Haitian Grand Ballroom 1, 1st floor
	NLP Applications 1	Tong Xiao	Qindao Function Room 1, G floor
	Question Answering and Knowlege Graph	Minghao Hu	Meeting room 06,07, G floor
	NLP Applications 2	Wenpeng Lu	Meeting room 08,09, G floor
17:00-17:10	QA for Keynote Talk 3		Haitian Grand Ballroom 1, 1st floor
17:10-17:40	Award & Closing Ceremony		Haitian Grand Ballroom 1, 1st floor
18:00-21:00	Dinner		Haitian Grand Ballroom 2(Buffet), 1st floor

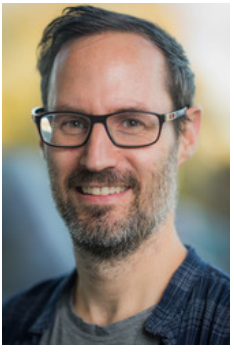
Keynote Talk 3

Oct. 17 , Chair: *Yansong Feng*
09:00-10:00, Haitian Grand Ballroom 1, 1st floor

Parametric vs Nonparametric Knowledge, and what we can learn from Knowledge Bases

Sebastian Riedel

Traditionally, AI and Machine Learning communities have considered knowledge from the perspective of discrete vs continuous representations, knowledge bases (KBs) vs dense vectors or logic vs algebra. While these are important dichotomies, in this talk I will argue that we should put more focus on another: parametric vs non-parametric modelling. Roughly, in the former a fixed set of parameters is used, in the latter parameters grow with data. I will explain recent advances in knowledge intensive NLP from this perspective, show the benefit of hybrid approaches, and discuss KBs as non-parametric approaches with relatively crude assumptions about what future information needs will be. By replacing these assumptions with a learnt model, we show that such "modern KBs" are a very attractive alternative or complement to current approaches.



Sebastian Riedel is a researcher at Facebook AI research, professor in Natural Language Processing and Machine Learning at the University College London (UCL) and an Allen Distinguished Investigator. He works in the intersection of Natural Language Processing and Machine Learning, and focuses on teaching machines how to read and reason. He was educated in Hamburg-Harburg (Dipl. Ing) and Edinburgh (MSc., PhD), and worked at the University of Massachusetts Amherst and Tokyo University before joining UCL.

Keynote Talk 4

Oct. 17, Chair: *Lu Wang*
10:45-11:45, Haitian Grand Ballroom 1, 1st floor

How Can We Know What and When Language Models Know?

Graham Neubig

One recent remarkable finding in natural language processing is that by training a model to simply predict words in a sentence, language models can learn a significant amount of world knowledge. In this presentation, I will discuss a new paradigm of ”prompting”, which attempts to elicit knowledge from language models by presenting a textual prompt such as ”CMU is located in ___” and asking the language model to fill in the answer. I will first give an outline of this paradigm as a whole, then present research regarding two questions. First: how can we most effectively elicit this knowledge from language models in this way, and what are the connections to other methods for parameter-efficient training of neural NLP models? Second: how can we best know when these predictions are accurate, and when they are no better than a random guess?

Prof. Graham Neubig is an associate professor at the Language Technologies Institute of Carnegie Mellon University. His work focuses on natural language processing, specifically multi-lingual models that work in many different languages, and natural language interfaces that allow humans to communicate with computers in their own language. Much of this work relies on machine learning, and he is also active in developing methods and algorithms for machine learning over natural language data. He publishes regularly in the top venues in natural language processing, machine learning, and speech, and his work has won awards at EMNLP 2016, EACL 2017, NAACL 2019, and ACL 2021.

Industrial Talk 3

Oct. 17 , Chair: *ZhaoChun Ren*
10:00-10:15, Online

AliceMind –Technologies, Platform, and Applications

Songfang Huang
Alibaba

In this talk, we present AliceMind, Alibaba's efforts on large scale pre-trained language models. We will introduce our technology innovations on generative, multilingual, and multimodal pre-trained language models. In particular, technologies on how to build and serve extra-large pre-trained language models will be mentioned. We will also talk about AliceMind Platform, and illustrate several business application scenarios of AliceMind in digital economy.

Dr. Songfang Huang is senior staff algorithm engineer of Language Technologies Lab, Alibaba DAMO Academy. He leads R&D on large scale pre-trained language models and its applications. His team develops AliceMind, Alibaba's Collection of Encoder-decoders from Machine Intelligence of DAMO, which has been widely used internally and externally. Prior to Alibaba, he was affiliated with IBM Research (Watson and China) working on speech and language processing. He obtained his PhD from The University of Edinburgh.

Industrial Talk 4

Oct. 17 , Chair: *ZhaoChun Ren*
11:45-12:00, Haitian Grand Ballroom 1, 1st floor

Applications of Vision and/or Language Technologies in Video Content Production

Yang Wu
Tencent ARC Lab

Video is becoming the major media in content production, distribution, and consumption in our daily lives. It has the advantanges of being able to include visual, audio, and textual information in the same media, in line with our natural experiences. While video distribution and consumption are growing explosively, video production, especially high-quality content generation, still faces great challenges due to the requirement of complex or even professional skills. This talk shares the latest progresses in Tencent ARC lab on building the fundamental components for video content production, especially those using cutting-edge vision and/or language technologies.

Yang Wu received a BS degree and a Ph.D degree from Xi'an Jiaotong University in 2004 and 2010, respectively. He is currently an expert researcher with Tencent ARC Lab. From Jul. 2019 to May 2021, he was a program-specific senior lecturer with Kyoto University. He was an assistant professor of the NAIST International Collaborative Laboratory for Robotics Vision from Dec.2014 to Jun. 2019. Before that, he was a program specific researcher with Kyoto University. His research is in the fields of computer vision, multimedia, and machine learning, with special interests in real applications. He has published more than 80 quality papers, including a AutoML best paper, two best student papers (BMVC and ICPR), works of three championships of global open challenges, and many oral presentations at major international conferences (ICCV, ECCV, BMVC, etc.).

Oral Paper Sessions

Oral Paper Session 5: Summarization and Generation (Haitian Grand Ballroom 1, 1st floor)
Session Chair: *Zhaochun Ren*

13:30-13:45	Explore Coarse-grained Structures for Syntactically Controllable Paraphrase Generation *Erguang Yang, Mingtong Liu, Deyi Xiong, YUJIE ZHANG, Yao Meng, Changjian Hu, Jinan Xu and Yufeng Chen*
13:45-14:00	Chinese Poetry Generation with Metrical Constraints *Yingfeng Luo, changliang li, Canan Huang, Chen Xu, Xin Zeng, Binghao Wei, Tong Xiao and Jingbo Zhu*
14:00-14:15	CNewSum: A Large-scale Chinese News Summarization Dataset with Human-annotated Adequacy and Deducibility Level *Danqing Wang, Jiaze Chen, Xianze Wu, Hao Zhou and Lei Li*
14:15-14:30	Diversified Paraphrase Generation with Commonsense Knowledge Graph *Xinyao Shen, Jiangjie Chen and Yanghua Xiao*
14:30-14:45	Question Generation from Code Snippets and Programming Error Messages *Bolun Yao, Wei Chen, Yeyun Gong, Bartuer Zhou, Jin Xie, Zhongyu Wei, Biao Cheng and Nan Duan*
14:45-15:00	Extractive Summarization of Chinese Judgment Documents via Sentence Embedding and Memory Network *Yan Gao, Zhengtao Liu, Juan Li and Jin Tang*

Oral Paper Session 6: Dialogue System (Qindao Function Room, G floor)
Session Chair: *Pengjie Ren*

13:30-13:45	Multi-Intent Attention and Top-k Network with Interactive Framework for Joint Multiple Intent Detection and Slot Filling *Xu Jia, Jiaxin Pan, Youliang Yuan and Min Peng*
13:45-14:00	Enhancing Long-Distance Dialogue History Modeling for Better Dialogue Ellipsis and Coreference Resolution *Zixin Ni and Fang Kong*
14:00-14:15	Few-Shot NLU with Vector Projection Distance and Abstract Triangular CRF *Su Zhu, Lu Chen, Ruisheng Cao, Zhi Chen, Qingliang Miao and Kai Yu*
14:15-14:30	Cross-domain Slot Filling with Distinct Slot Entity and Type Prediction *Shudong Liu, Peijie Huang, Zhanbiao Zhu, Hualin Zhang and Jianying Tan*
14:30-14:45	Knowledge-Grounded Dialogue with Reward-Driven Knowledge Selection *Shilei Liu, Xiaofeng Zhao, Bochao Li and Feiliang Ren*
14:45-15:00	Exploiting Explicit and Inferred Implicit Personas for Multi-turn Dialogue Generation *Ruifang Wang, Ruifang He, Longbiao Wang, Yuke Si, Huanyu Liu, Haocheng Wang and Jianwu Dang*

Oral Paper Session 7:Multimodality(Meeting room 06,07, G floor)
Session Chair: *Zhouyu Wei*

13:30-13:45	An Object-Extensible Training Framework for Image Captioning *Yike Wu, Ying Zhang and Xiaojie Yuan*
13:45-14:00	XGPT: Cross-modal Generative Pre-Training for Image Captioning *Qiaolin Xia, Haoyang Huang, Nan Duan, Dongdong Zhang, Lei Ji, Zhifang Sui, Edward Cui, Taroon Bharti and Ming Zhou*
14:00-14:15	Relation-aware Multi-hop Reasoning for Visual Dialog *Yao Zhao, Lu Chen and Kai Yu*
14:15-14:30	MUMOR: A Multimodal Dataset for Humor Detection in Conversations *Jiaming Wu, Hongfei LIN, Liang Yang and Bo Xu*
14:30-14:45	Skeleton-Based Sign Language Recognition with Attention-enhanced Graph Convolutional Networks *Wuyan Liang and Xiaolong Xu*
14:45-15:00	Multi-Modal Sarcasm Detection Based on Contrastive Attention Mechanism *Xiaoqiang Zhang, Ying Chen and Guangyuan Li*

Oral Paper Session 8: Sentiment Analysis (Meeting room 08,09, G floor)
Session Chair: *Liang Yang*

13:30-13:45	Locate and Combine: A Two-Stage Framework for Aspect-Category Sentiment Analysis *Yang Wu, Zhenyu Zhang, Yanyan Zhao and Bing Qin*
13:45-14:00	Highway-Based Local Graph Convolutional Network For Aspect Based Sentiment Analysis *Shiguan Pang, Zehao Yan, Yun Xue, Weihao Huang, Bixia Tang and Anan Dai*
14:00-14:15	Dual Adversarial Network Based on BERT for Cross-domain Sentiment Classification *Shaokang Zhang, Xu Bai, Lei Jiang and Huailiang Peng*
14:15-14:30	Aspect-Sentiment-Multiple-Opinion Triplet Extraction *Fang Wang, Yuncong Li, Sheng-hua Zhong, Cunxiang Yin and Yancheng He*
14:30-14:45	Semantic Enhanced Dual-channel Graph Communication Network for Aspect-based Sentiment Analysis *Zehao Yan, Shiguan Pang and Yun Xue*
14:45-15:00	基于多通道压缩双线性池化的情感原因句子对提取模型 *Jin Huang, Shi Xu, Ercong Cai, Zhijie Wu, Meimei Guo and Jia Zhu*

Oral Paper Session 9: Text Mining (Haitian Grand Ballroom 1, 1st floor)
Session Chair: *Tianyong Hao*

15:15-15:30	A Residual Dynamic Graph Convolutional Network for Multi-label Text Classification *Bingquan Wang, Jie Liu, Shaowei Chen, Xiao Ling, Shanpeng Wang, Wenzheng Zhang, Liyi Chen and Jiaxin Zhang*
15:30-15:45	BERT-KG：A Short Text Classification Model Based on Knowledge Graph and Deep Semantics *Yuyanzhen Zhong, zhiyang zhang, Weiqi Zhang and Juyi Zhu*
15:45-16:00	A Simple Baseline for Cross-domain Few-shot Text Classification *Chen Zhang and Dawei Song*
16:00-16:15	Leveraging Lexical Common-Sense Knowledge for Boosting Bayesian Modeling *Yashen Wang*
16:15-16:30	A Aggregating inter-viewpoint relationships of user's review for accurate recommendation *Xingchen He, Xuling Zheng, Yidong Chen and Guocheng Zhang*

Oral Paper Session 10: NLP Applications 1(Qindao Function Room 1, G floor)
Session Chair: *Tong Xiao*

15:15-15:30	Sentence Ordering by Context-enhanced Pairwise Comparison *Haowei Du, Jizhi Tang and Dongyan Zhao*
15:30-15:45	Uncertainty-aware Self-paced Learning for Grammatical Error Correction *Kai Dang, Jiaying Xie, Jie Liu and Shaowei Chen*
15:45-16:00	结合自监督学习的多任务文本语义匹配方法研究 *Yuan Chen and Xinying Qiu*
16:00-16:15	Shared Component Cross Punctuation Clauses Recognition in Chinese *Xiang Liu, Ruifang Han, Shuxin Li, Yujiao Han, Mingming Zhang, Zhilin Zhao and zhiyong luo*
16:15-16:30	Pre-trained Language models for Tagalog with Multi-source data *Shengyi Jiang, Yingwen Fu, Xiaotian Lin and Nankai Lin*

Oral Paper Session 11: Question Answering and Knowlege Graph (Meeting room 06,07, G floor)
Session Chair: *Minghao Hu*

15:15-15:30	ThinkTwice: A Two-Stage Method for Long-Text Machine Reading Comprehension *Mengxing Dong, Bowei Zou, Jin Qian, Rongtao Huang and Yu Hong*
15:30-15:45	EviDR: Evidence-Emphasized Discrete Reasoning for Reasoning Machine Reading Comprehension *Yongwei Zhou, Junwei Bao, Haipeng Sun, Jiahui Liang, Youzheng Wu, Xiaodong He, Bowen Zhou and Tiejun Zhao*
15:45-16:00	Attention Based Reinforcement Learning with Reward Shaping for Knowledge Graph Reasoning *Sheng Wang, Xiaoying Chen and Shengwu Xiong*
16:00-16:15	基于多跳路径的知识库补全 *Yinmiao Wang and Zhimin Han*

Oral Paper Session 12: NLP Applications 2 (Meeting room 08,09, G floor)
Session Chair: *Wenpeng Lu*

15:15-15:30	Syntax and Sentiment Enhanced BERT for Earliest Rumor Detection *Xin Miao, Dongning Rao and Zhihua Jiang*
15:30-15:45	Emotion Classification with Explicit and Implicit Syntactic Information *Nan Chen, Qingrong Xia, Xiabing Zhou, Wenliang Chen and Min Zhang*
15:45-16:00	Metaphor Recognition and Analysis via Data Augmentation *Liang Yang, Jingjie Zeng, Shuqun Li, Zhexu Shen, Yansong Sun and Hongfei Lin*
16:00-16:15	A Dual-Attention Neural Network for Pun Location and Using Pun-Gloss Pairs for Interpretation *Shen Liu, Meirong Ma, Hao Yuan, Jianchao Zhu, Yuanbin Wu and Man Lan*

Index

Cai,Ercong, 67
Gao,Jing , 50
Huang,Hui, 49
jiang,yuru, 52
Kong,Fang, 47
Li,Juanzi , 47
Li,X.Y, 48
Liu,Baoju, 47
Liu,Huan, 49
Neubig,Graham, 62
Nong,Wei , 48
Wang,Liang, 51
Wang,Lu, 62
Wang,Qiqi, 48
Wang,Quan, 54
Wang,Zhichun, 51
Wu,Youzheng, 7
Zhang,Ying, 66

Bai,Xu, 67
Bai,Xuefeng , 50
Baldwin,Tim, 42
Ban,Mabao , 49
Bao,Hui, 53
Bao,Jie, 51
Bao,Junwei, 54, 69
Bao,Lujia, 52
Bao,Xinrui, 53
Bi,Ran , 47
Bi,Yanchao, 43
Bo,Lin , 47
Bo,Ruipeng, 54

Cai,Rangjia , 49
Cao,Ruisheng, 66
Cao,Shuai, 53
Cao,Xuezhi, 59
Cao,Yunbo, 59
Chao,Cai, 23, 39
Chao,Rui, 53
Chao,WenHan, 53
Chen,Boxing , 50
Chen,Chunquan, 52
Chen,Jiahao , 54
Chen,Jiangjie, 65
Chen,Jiaze, 65
Chen,Jingqiang , 49
Chen,KeJia , 49

Chen,Lei, 59
Chen,Liyi, 68
Chen,Lu, 66
Chen,Nan, 69
Chen,Ruijun, 53
Chen,Shaowei, 68
Chen,Shuai, 54
Chen,Sichao, 54
Chen,Wei, 65
Chen,Wenliang, 69
Chen,Xiaoliang, 53
Chen,Xiaona, 51
Chen,Xiaoshuai, 52
Chen,Xiaoying, 69
Chen,Xinyu , 48
Chen,Yidong, 68
Chen,Yidong , 47
Chen,Ying, 66
Chen,Yuan, 68
Chen,Yubo , 48
Chen,Yufeng, 53, 65
Chen,Yufeng , 49
Chen,Yunian, 59
Chen,Yunwen, 23, 37
Chen,Zhi, 66
Cheng,Biao, 65
Cheng,Chen, 59
Chi,Zewen, 53
Chu,Min, 46
Chu,Xiaomin , 47
Cui,Zhen, 52

Dai,Anan, 67
Dai,Wen, 54
Dai,Xudong, 54
Dai,Xudong , 48
Dang,Jianwu, 66
Dang,Kai, 68
Di,Hui , 49
Ding,Jiajie, 52
Dong,Mengxing, 69
Du,Haowei, 68
Du,Yajun, 53
Duan,Nan, 65, 66

Edward,Cui, 66

Fa,Lingling, 54
Fan,Xiaoming, 51
Fan,Yaxin, 47
Feng,Junlan, 44
Feng,Lin , 48
Feng,Wenhe , 49
Feng,Yang, 15
Feng,Yansong, 61
Fu,Biao , 47
Fu,Yingwen, 68

Gan,Zhen , 48
Gao,Pengfei, 51
Gao,Sheng, 51
Gao,Yan, 65
Gao,Zeyu, 53
Geng,Yuanling, 52
Ghosal,Tirthankar, 52
Gong,Yeyun, 65
Gong,Yifei, 52
Gu,Chengmin, 54
Gui,Tao, 18
Guo,Dongjie, 54
Guo,Meimei, 67

Han,Ruifang, 68
Han,Ting, 53
Han,Yujiao, 68
Hao,Jianye, 16
Hao,Tianyong, 51, 68
He,Kai, 53
He,Pan, 52
He,Ruifang, 66
He,Weikai, 52
He,Xiaodong, 54, 69
He,XiaoFeng , 48
He,Xingchen, 68
He,Yancheng, 67
He,Zheyu , 54
Hong,Yu, 69
Hou,Lei , 47
Hu,Baotian, 23, 36, 54
Hu,Changjian, 65
Hu,Fuyu, 51
Hu,Hai, 59
Hu,Mengyang , 47

Hu,Minghao, 69
Hu,Nan, 48
Hu,Xiaoxuan , 50
Hu,Xinshuo, 54
Hua,Xinyu, 54
Huang,Canan, 65
Huang,Chong Hsuan, 53
Huang,Degen, 53
Huang,Degen , 49
Huang,Haoyang, 66
Huang,Haoyang , 49
Huang,Heyan, 53
Huang,Heyan , 49
Huang,Jin, 67
Huang,Kaiyu, 53
Huang,Kaiyu , 49
Huang,Minlie, 53
Huang,Peijie, 66
Huang,Rongtao, 69
Huang,Songfang, 63
Huang,Weihao, 67
Huang,Xuanjing, 51
Huang,Yongjie, 51

Inkpen,Diana, 53

Jia,Xu, 66
Jia,Yuxiang, 53
Jiang,Di, 51
Jiang,Feng , 47
Jiang,Lei, 67
Jiang,Shengyi, 50, 68
Jiang,Yufan, 54
Jiang,Zhihua, 69
Jie,Lu, 51
Jie,Zhanming, 59
Jin,Qin, 21, 24
Ju,Jiangzhou, 52

Kaur,Prabsimran, 52
Kohli,Guneet Singh, 52
Kong,Fang, 66
Kong,Fang , 48
Kong,Jun , 50

Lan,Man, 54, 69
Lan,Man , 54
Lee,Lap-Kei, 51
Lei,Ji, 66
Li,Bochao, 66
li,changliang, 65
Li,Dongfang, 54
Li,Guangyuan, 66
Li,Hongyu , 54
Li,Jiuyi , 49
Li,Juan, 65
Li,Juan , 47
Li,Junyi, 59
Li,Lei, 65
Li,Lu, 59
Li,Maoxi, 51
Li,Minglei, 59
Li,Mu, 54
Li,Peifeng, 51
Li,Peifeng , 47, 48
Li,Qizhi, 53
Li,Shuqun, 52, 69
Li,Shuxin, 68
Li,Si, 51
Li,Sujian, 52
Li,Weikang, 52
Li,Xiangyang, 52
Li,Xianyong, 53
Li,Xuan, 52
Li,Xuanyu, 53
Li,Yafu , 50
Li,Yaliang , 50
Li,Yaru, 52
Li,Yu Ming, 52
Li,Yu-Feng, 14
Li,Yuncong, 67
Li,Yunxin, 54
Li,Zheng, 52
Li,Zhengyan, 51
Li,Zhoujun , 49
Li,Zhucong , 48
Lian,Yixin, 53
Liang,Jiahui, 54, 69
Liang,Wuyan, 66
Liao,Yiming, 51
LIN,Hongfei, 66
Lin,Hongfei, 52, 59, 69

Lin,Nankai, 68
Lin,Xiaotian, 68
Ling,Xiao, 68
Liu, Jing , 54
Liu, Shilei, 66
Liu,Hai, 51
Liu,Huanyu, 66
Liu,Jian , 49
Liu,Jie, 68
Liu,Junpeng , 49
Liu,Kang, 48
Liu,Mingtong, 65
Liu,Mingtong , 49
Liu,Peng, 54
Liu,Peng , 48
Liu,Pengfei, 22, 31
liu,pengyuan, 53
Liu,Pengyuan , 47
Liu,Qun, 55, 59
Liu,Shen, 69
Liu,Shengping , 48
Liu,Shudong, 66
Liu,Wei, 53
Liu,Wei , 50
Liu,Xiang, 68
Liu,Ximing, 53
Liu,Yaduo, 54
Liu,Yongbin, 51
Liu,Yufei , 50
Liu,Yunfeng, 59
Liu,Zhenghao, 51
Liu,Zhengtao, 65
Liu,Zhiyuan, 51
Liu,Zihao , 48
Lu,Wenpeng, 69
Lu,Xin, 53
Lucaci,Diana, 53
Luo,Fuli, 22, 32
Luo,Weihua , 50
Luo,Yichao, 51
Luo,Yingfeng, 65
Luo,zhiyong, 68
Luo,Zhunchen, 53
Lv,Rongrong , 54
Lv,Shangwen, 54
Lv,Xin , 47
Ma,Chao, 53
Ma,Meirong, 69
Ma,Meirong , 54
Ma,Shuming , 49
Ma,Weifeng, 52
Ma,Yinglong, 51
Ma,Yutuan, 53
Mao,Jiaxin, 57
Mao,Xian-Ling, 53
Meng,Yao, 65
Miao,Guoyi , 49
Miao,Qingliang, 66
Miao,Xin, 69
Mihalcea,Rada, 42
Min,Zhang, 52
Ming,Zhou, 66
Mo,Yijun, 51

Ni,Bin, 51
Ni,Yang, 51
Ni,Yuan , 50
Ni,Zixin, 66
Niu,Changyong, 53
Niu,Yuchen , 50
Niu,Zhengyu , 54

Ouchi,Kazushige , 49
Ouyang,Chunping, 51

Pan,Jiaxin, 66
Pang,Shiguan, 67
Peng,Boci, 51
Peng,Huailiang, 67
Peng,Min, 66
Peng,Tao, 59
Peng,Wei, 53

Qi,Chao, 52
Qian,Jin, 69
Qin,Bing, 53, 67
Qin,Wentao, 52, 53
Qiu,Xinying, 68
Qiu,Xipeng, 51

Rana,Prashant, 52
Rao,Dongning, 69

Ren,Feiliang, 54, 66
Ren,Pengjie, 66
Ren,ZhaoChun, 64
Ren,Zhaochun, 65
Riedel,Sebastian, 61

Shao,Yifeng, 51
Shen,Xinyao, 65
Shen,Zhexu, 69
Shi,Shumin , 49
Shi,Xiaodong , 47
Shi,Yafei , 48
Si,Yuke, 66
Song,Dawei, 68
Su,Fenglong, 54
Su,Jinsong, 21, 25
Sun,Fanshu, 53
Sun,Haipeng, 69
Sun,Jian, 59
Sun,Si, 51
Sun,Xichen, 53
Sun,Xinghai, 52
Sun,Yansong, 69
Sun,Yu, 7, 22, 30
Sun,Yuxiang, 54

Takanobu,Ryuichi, 53
Tan,Conghui, 51
Tan,Jianying, 66
Tan,Xin, 47
Tang,Bixia, 67
Tang,Jin, 65
Tang,Jizhi, 68
Taroon,Bharti, 66
Tian,Dawei , 47
Tong,Yiqi , 47

Wan,Dazhen, 53
Wan,Jing , 48
Wan,Juncheng , 49
Wan,Zhang, 53
Wang,Bang, 51
Wang,Bin, 52
Wang,Bingquan, 68
Wang,Chenxu, 51
Wang,Cunxiang , 50
Wang,Danqing, 65
Wang,Fang, 67
Wang,Haocheng, 66
Wang,Huizhen , 48
Wang,Jin, 53
Wang,Jin , 50
Wang,Jinfeng , 47
Wang,Kai, 59
Wang,Liudi, 52
Wang,Longbiao, 66
WANG,Mingwen, 51
Wang,Qian, 51
Wang,Ruifang, 66, 67
Wang,Shanpeng, 68
Wang,Sheng, 69
Wang,Shuheng, 52
Wang,Shuheng , 49
Wang,Shuohuan, 22, 33
Wang,Xiaoling , 50
Wang,Xiaolong, 54
Wang,Yan, 52
Wang,Yashen, 68
Wang,Yi Dong, 52
Wang,Yifan, 54
Wei,Binghao, 65
Wei,Furu , 49
Wei,Zhongyu, 65
Wei,Zhouyu, 66
Wen,JiRong , 50
Wong,Derek F. , 50
Wong,Leung-Pun, 51
Wu,Benhong, 59
Wu,Chunhua , 50
Wu,Fangzhao, 58
Wu,Hua , 7, 54
Wu,Jialun, 53
Wu,Jiaming, 66
Wu,Lingxiang, 22, 34
Wu,Shuangzhi, 54
Wu,Shuixiu, 51
Wu,Xianze, 65
Wu,Yang, 64, 67
Wu,Yike, 66
Wu,Youzheng, 54, 69
Wu,Yuanbin, 69
Wu,Zhijie , 67

Xia ,Mao Jin, 52
Xia,Qiaolin, 66
Xia,Qingrong, 69
Xia,Yu , 48
Xiang,Lu, 53
Xiang,Wei, 51
Xiao,Kang, 52
Xiao,Tong, 65, 68
Xiao,Yanghua, 65
Xiaolei,Lu, 51
Xie,Guotong , 50
Xie,Jiaying, 68
Xie,Jin, 65
Xie,Rui, 59
Xie,Zhipeng , 47
Xinya,Zhou, 51
Xiong,Chenyan, 51
Xiong,Deyi, 65
Xiong,Shengwu, 69
Xu, Weijie, 54
Xu,Bo, 66
Xu,Chen, 65
Xu,Heng-Da, 53
Xu,Huimin, 59
Xu,Jinan, 53, 65
Xu,Jinan , 49
Xu,Jinan , 49
Xu,Liang, 54, 59
Xu,Mingzhou , 50
Xu,Qian, 51
Xu,Ruifeng, 48, 59
Xu,Sheng, 51
Xu,Sheng , 48
Xu,Shi, 67
Xu,Xiaolong , 66
Xu,Yajing, 51
Xu,Yang, 52
Xu,Yige, 51
Xue,Yun, 67
Xun,Endong, 23, 35

Yan,Rui, 12
Yan,Zehao, 67
Yang,Erguang, 65
Yang,Jian , 49
Yang,Liang, 52, 59, 66, 67, 69
Yang,Meng, 51
Yang,Min, 59
Yang,ShuangJi , 48
Yang,Xuefeng, 59
Yang,Yan, 54
Yang,Zinong , 49
Yao,Bolun, 65
Yao,Jin-Ge, 55
Yao,Ting, 21, 26
Ye,Heng, 52
Yin,Cunxiang, 67
Yin,Shujuan, 54
Yin,Xuemeng, 53
Yin,Yichun, 21, 28, 59
Yin,Zimo, 52
Ying,JiaHao , 54
Yinmiao Wang, 69
Yiqi,Tong, 51
Yu,Dong, 53
Yu,Hao, 53
Yu,Jianfei, 21, 27
Yu,Jintao, 59
Yu,Kai, 66
Yu,Yong , 49
Yuan,Hao, 69
Yuan,Hao , 54
Yuan,Xiaojie, 66
Yuan,Youliang, 66

Zan,Hongying, 53
Zeng,jiali, 54
Zeng,Jingjie, 69
Zeng,Xin, 65
ZGUO,Mazhaxi , 49
Zhang,Baoli , 48
Zhang,Bosen, 51
Zhang,Chen, 68
Zhang,Dongdong, 66
Zhang,Dongdong , 49
Zhang,Dongmei , 50
Zhang,Fan, 52
Zhang,Guocheng, 68
Zhang,Haichuan, 53
Zhang,Hanwang, 13
Zhang,Hengtong , 50
Zhang,Hu, 53

Zhang,Hualin, 66
Zhang,Jiajun, 47
Zhang,Jiaxin, 68
Zhang,Jing, 56
Zhang,Jinhui, 59
Zhang,Jun, 54
Zhang,Longhui, 54
Zhang,Longyin, 47
Zhang,Longyin , 47
Zhang,Min, 69
Zhang,Mingming, 68
Zhang,Qi, 18, 51
zhang,sanle, 53
Zhang,Shaokang, 67
Zhang,Shaowu, 52
Zhang,TaoLin , 48
Zhang,Wei, 52
zhang,weicong, 53
Zhang,Weinan, 17
Zhang,Weiqi, 68
Zhang,Wenzheng, 68
Zhang,Xi, 23, 38
Zhang,Xiaoqiang, 66
Zhang,Xuanwei, 59
Zhang,Xuejie, 53
Zhang,Xuejie , 50
Zhang,Yadong, 54
Zhang,Yan , 48
Zhang,Yanru, 52
Zhang,Yue, 43
Zhang,Yue , 50
Zhang,Yue , 50
ZHANG,YUJIE, 65
Zhang,Yujie, 53
Zhang,Yujie , 49
Zhang,Yujing , 48
Zhang,Yulai, 52
Zhang,Yushi , 54
Zhang,Yuyang, 59
zhang,yuyao, 52
Zhang,Zhenyu, 67
Zhang,Zhirui , 50
zhang,zhiyang, 68
Zhao, Xiaofeng, 66
Zhao,Chongshuai, 54
Zhao,Chongshuai , 48
Zhao,Dongyan, 52, 53, 68
Zhao,Jun, 48
Zhao,Tiejun, 69
Zhao,Wayne Xin , 50
Zhao,Weixiang, 53
Zhao,Xiaofeng, 54
Zhao,Xin,Wayne, 55
Zhao,Yang, 53
Zhao,Yanyan, 53, 67
Zhao,Yao, 66
Zhao,Yu, 54
Zhao,Zhilin, 68
Zheng,Boyuan , 50
Zheng,Derong , 52
Zheng,Xuling, 68
Zhifang,Sui, 66
Zhimin Han, 69
Zhong,Sheng-hua, 67
Zhong,Yuyanzhen, 68
Zhou,Bartuer, 65
Zhou,Bowen, 69
Zhou,Bowen , 54
Zhou,Feng , 48
Zhou,Fengqing, 59
Zhou,Gongxue, 52
Zhou,Guodong, 47
Zhou,Hao, 49, 65
Zhou,Xiabing, 52, 69
Zhou,Yongwei, 69
Zhou,Yu, 53
Zhu,Jia, 67
Zhu,Jianchao, 54, 69
Zhu,Jingbo, 65
Zhu,Jingbo , 48
Zhu,Jingjing, 53
Zhu,Junnan, 53
Zhu,Juyi, 68
Zhu,Ming , 47
Zhu,Qiaoming, 47, 51
Zhu,Qiaoming , 48
Zhu,Su, 66
Zhu,Wei , 50
Zhu,Xiaozhi, 51
Zhu,Yujin , 54
Zhu,Zhanbiao, 66
Zong,Chengqing, 53

Zou,Bowei, 69
Zou,Lei, 11
Zou,Yicheng, 51