ACL 2024

62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024)

Proceedings of the Conference

August 11-16, 2024

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-098-1

Introduction

Welcome to this year's ACL tutorial session, a highlight of our conference. We are thrilled to have you with us!

The ACL tutorial session aims to provide attendees with a thorough introduction to key topics in our fast-evolving research field, delivered by expert researchers. This year, as in recent years, the process of calling for, submitting, reviewing, and selecting tutorials was a collaborative effort across multiple conferences: EACL, NAACL, ACL, and EMNLP.

We assembled a review committee consisting of the tutorial chairs from EACL (Sharid Loaiciga, Mohsen Mesgar), NAACL (Rui Zhang, Nathan Schneider, Snigdha Chaturvedi), and the interim EMNLP tutorial chair (Isabelle Augenstein). Each tutorial proposal was meticulously reviewed by a panel of three reviewers, who assessed them based on criteria such as clarity, preparedness, novelty, timeliness, instructors' experience, potential audience, open access to teaching materials, and diversity (including multilingualism, gender, age, and geolocation). Out of 27 submissions, 6 were selected for presentation at ACL.

We would like to thank the tutorial authors for their commitment, dedicated collaboration and flexibility while organizing the conference. Finally, our thanks go to the conference organizers for effective collaboration, and in particular to the general chair Claire Gardent.

Enjoy the session!

Warm regards,
ACL 2024 Tutorial Co-chairs
Luis Chiruzzo
Hung-yi Lee
Leonardo F. R. Ribeiro

Organizing Committee

General Chair

Claire Gardent, CNRS and Université de Lorraine

Program Chairs

Lun-Wei Ku, Academia Sinica
Andre Martins, Instituto Superior Técnico / Instituto de Telecomunicações / Unbabel
Vivek Srikumar, University of Utah

Local Organization

Thepchai Supnithi, NECTEC and AIAT
Prachya Bookwan, NECTEC and AIAT
Thanaruk Theeramunkong, SIIT and AIAT

Tutorial Chairs

Luis Chiruzzo, Universidad de la República
Hung-yi Lee, National Taiwan University
Leonardo Ribeiro, Amazon Alexa Seattle

Program Committee

iv

Program Chairs

Lun-Wei Ku, Academia Sinica
Andre Martins, Instituto Superior Técnico / Instituto de Telecomunicações / Unbabel
Vivek Srikumar, University of Utah

Table of Contents

Program

Sunday, August 11, 2024

09:00 - 12:30	*Tutorial 1 - Computational Linguistics for Brain Encoding and Decoding: Principles, Practices and Beyond*
09:00 - 12:30	*Tutorial 2 - Automatic and Human-AI Interactive Text Generation (with a focus on Text Simplification and Revision)*
09:00 - 12:30	*Tutorial 3 - Vulnerabilities of Large Language Models to Adversarial Attacks*
14:00 - 17:30	*Tutorial 4 - Computational Expressivity of Neural Language Models*
14:00 - 17:30	*Tutorial 5 - Watermarking for Large Language Models*
14:00 - 17:30	*Tutorial 6 - Presentation Matters: How to Communicate Science in the NLP Venues and in the Wild?*

Computational Linguistics for Brain Encoding and Decoding: Principles, Practices and Beyond

Jingyuan Sun **Shaonan Wang** **Zijiao Chen** **Jixing Li** **Marie-Francine Moens**

Computational linguistics (CL) has witnessed tremendous advancements in recent years, with models such as large language models demonstrating exceptional performance in various natural language processing tasks. These advancements highlight their potential to help understand brain language processing, especially through the lens of brain encoding and decoding. Brain encoding involves the mapping of linguistic stimuli to brain activity, while brain decoding is the process of reconstructing linguistic stimuli from observed brain activities. CL models that excel at capturing and manipulating linguistic features are crucial for mapping linguistic stimuli to brain activities and vice versa. Brain encoding and decoding have vast applications, from enhancing human-computer interaction to developing assistive technologies for individuals with communication impairments. This tutorial will focus on elucidating how computational linguistics can facilitate brain encoding and decoding. We will delve into the principles and practices of using computational linguistics methods for brain encoding and decoding. We will also discuss the challenges and future directions of brain encoding and decoding. Through this tutorial, we aim to provide a comprehensive and informative overview of the intersection between computational linguistics and cognitive neuroscience, inspiring future research in this exciting and rapidly evolving field.

Jingyuan Sun, Postdoc researcher in the Department of Computer Science, KU Leuven, Belgium
email: jingyuan.sun@kuleuven.be
website: https://www.kuleuven.be/wieiswie/en/person/00155742
Jingyuan Sun has published papers in artificial intelligence and natural language processing journals and conferences such as TNNLS, Scientific Data, AAAI, IJCAI, EMNLP, COLING, ECAI, etc. He also serves as a (senior) PC member for these above conferences. He is a reviewer of TPAMI.

Shaonan Wang, Associate Professor in the State Key Laboratory of Multimodal Artificial Intelligence System, Institute of Automation, Chinese Academy of Sciences
email: shaonan.wang@nlpr.ia.ac.cn
website: https://wangshaonan.github.io/
Shaonan Wang has contributed papers to natural language processing journals and conferences such as Information Sciences, TNNLS, ACL, EMNLP, AAAI, and IJCAI. Her work also spans neurolinguistics, with publications in Scientific Data, Brain and language, Cognitive, Affective, and Behavioral Neuroscience, and more. She serves as an editor for Neurobiology of Language and TALLIP and holds the role of curriculum chair at Neuromatch Academy.

Jixing Li, Assistant Professor at in the Department of Linguistics and Translation at the City University of Hong Kong
email: jixingli@cityu.edu.hk
website: https://jixing-li.github.io/
Jixing Li is an Assistant Professor at in the Department of Linguistics and Translation at the City University of Hong Kong. Her research combines NLP models and neuroimaging methods to examine syntactic and semantic analyses in the brain. Her work has been published in the Journal of Neuroscience, Brain and Language, Annual Review of Linguistics, etc. She serves as an ad-hoc reviewer for top journals in the field of neurolinguistics and is on the editorial board of Communications Psychology.

Zijiao Chen, PhD candidate in the Multimodal Neuroimaging in Neuropsychiatric Disorders Laboratory at the National University of Singapore
email: zijiao.chen@u.nus.edu
Zijiao Chen is a PhD candidate in the Multimodal

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 5: Tutorial Abstracts), pages 1–2
August 11-16, 2024 ©2024 Association for Computational Linguistics

Neuroimaging in Neuropsychiatric Disorders Laboratory at the National University of Singapore. Her research primarily centers on representation learning within neuroimaging data and brain decoding. She has published papers in artificial intelligence and neuroimage conferences and journals such as CVPR, OHBM, NCAA, and AD.

Marie-Francine Moens, Full Professor in the Department of Computer Science, KU Leuven
email: sien.moens@cs.kuleuven.be
website: https://people.cs.kuleuven.be/~sien.moens/
Marie-Francine Moens is a Full Professor in the Department of Computer Science, KU Leuven. She is a fellow of the European Laboratory for Learning and Intelligent Systems (ELLIS). She is an associate editor of the journal IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI). She holds the ERC Advanced Grant CALCULUS (2018-2024) granted by the European Research Council. She used to be the chair of the European Chapter of the Association for Computational Linguistics (EACL) and was a member of the executive board of the Association for Computational Linguistics (ACL). From 2012 to 2016 she was the coordinator of the MUSE project financed by Future and Emerging Technologies (FET) - Open of the European Commission.

Automatic and Human-AI Interactive Text Generation
(with a focus on Text Simplification and Revision)

Yao Dou **Philippe Laban** **Claire Gardent** **Wei Xu**

In this tutorial, we focus on text-to-text generation, a class of natural language generation (NLG) tasks, that takes a piece of text as input and then generates a revision that is improved according to some specific criteria (e.g., readability or linguistic styles), while largely retaining the original meaning and the length of the text. This includes many useful applications, such as text simplification, paraphrase generation, style transfer, etc. In contrast to text summarization and open-ended text completion (e.g., story), the text-to-text generation tasks we discuss in this tutorial are more constrained in terms of semantic consistency and targeted language styles. This level of control makes these tasks ideal testbeds for studying the ability of models to generate text that is both semantically adequate and stylistically appropriate. Moreover, these tasks are interesting from a technical standpoint, as they require complex combinations of lexical and syntactical transformations, stylistic control, and adherence to factual knowledge, – all at once. With a special focus on text simplification and revision, this tutorial aims to provide an overview of the state-of-the-art natural language generation research from four major aspects – Data, Models, Human-AI Collaboration, and Evaluation – and to discuss and showcase a few significant and recent advances: (1) the use of non-retrogressive approaches; (2) the shift from fine-tuning to prompting with large language models; (3) the development of new learnable metric and fine-grained human evaluation framework; (4) a growing body of studies and datasets on non-English languages; (5) the rise of HCI+NLP+Accessibility interdisciplinary research to create real-world writing assistant systems.

Yao Dou, Ph.D. student in the College of Computing at the Georgia Institute of Technology
email: `douy@gatech.edu`
website: `https://yao-dou.github.io/`
Yao Dou is a Ph.D. student in the College of Computing at the Georgia Institute of Technology, advaised by Prof. Wei Xu. His research interests lie in natural language processing and machine learning. His recent work focuses on text simplification, text evaluation, and social media privacy.

Philippe Laban, Research Scientist at Salesforce Research
email: `plaban@salesforce.com`
website: `https://tingofurro.github.io/`
Philippe Laban is a Research Scientist at Salesforce Research. His research is at the intersection of NLP and HCI, focusing on several tasks within text generation, including text simplification and summarization. He received his Ph.D. in Computer Science from UC Berkeley in 2021. His thesis is titled "Unsupervised Text Generation and its Application to News Interfaces". His recent work has focused on expanding the scope of text simplification to the paragraph and document-level and evaluating textediting interfaces. He publishes in both *ACL and HCI conferences, including work on interactive user interface design for NLP applications.

Claire Gardent, Senior Research Scientist at the French National Center for Scientific Research (CNRS)
email: `claire.gardent@loria.fr`
website: `https://members.loria.fr/CGardent/`
Claire Gardent is a Senior Research Scientist at the French National Center for Scientific Research (CNRS), based at the LORIA Computer Science research unit in Nancy, France. In 2022, she was selected as an ACL Fellow and was awarded the CNRS Silver Medal. She works in the field of NLP with a particular interest in Natural Language Generation. In 2017, she launched the WebNLG challenge, a shared task where the goal is to generate text from Knowledge Base fragments. She

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 5: Tutorial Abstracts), pages 3–4
August 11-16, 2024 ©2024 Association for Computational Linguistics

has proposed neural models for simplification and summarization; for the generation of long-form documents such as multi-document summaries and Wikipedia articles; for multilingual generation from Abstract Meaning Representations and for response generation in dialog. She currently heads the AI XNLG Chair on multi-lingual, multi-source NLG and the CNRS LIFT Research Network on Computational, Formal and Field Linguistics

Wei Xu, Assistant Professor in the College of Computing at the Georgia Institute of Technology
email: `wei.xu@cc.gatech.edu`
website: `https://cocoxu.github.io/`
Wei Xu is an Assistant Professor in the College of Computing at the Georgia Institute of Technology. Her recent research focuses on text generation (including data construction, controllable model, human and automatic evaluation), stylistics, analyzing and evaluating large language models (including multilingual capability, cross-lingual transfer learning, cultural bias, and cost efficiency). She is a recipient of the NSF CAREER Award, CrowdFlower AI for Everyone Award, best paper award from COLING 2018, and honorable mention from ACL 2023. She is an NAACL executive board member and regularly serves as a (senior) area chair for *ACL conferences. She frequently gives invited talks at universities and companies. She has given tutorials on "NLP for Social Media and Text Analysis" and has organized multiple workshops, including WNUT, GEM, and TSAR.

Computational Expressivity of Neural Language Models

Alexandra Butoi **Ryan Cotterell** **Anej Svete**

Language models (LMs) are currently at the forefront of NLP research due to their remarkable versatility across diverse tasks. However, a large gap exists between their observed capabilities and the explanations proposed by established formal machinery. To motivate a better theoretical characterization of LMs' abilities and limitations, this tutorial aims to provide a comprehensive introduction to a specific framework for formal analysis of modern LMs using tools from formal language theory (FLT). We present how tools from FLT can be useful in understanding the inner workings and predicting the capabilities of modern neural LM architectures. We will cover recent results using FLT to make precise and practically relevant statements about LMs based on recurrent neural networks and transformers by relating them to formal devices such as finite-state automata, Turing machines, and analog circuits. Altogether, the results covered in this tutorial will allow us to make precise statements and explanations about the observed as well as predicted behaviors of LMs, as well as provide theoretically motivated suggestions on the aspects of the architectures that could be improved.

Alexandra Butoi, Ph.D. Candidate at ETH Zürich.
email: alexandra.butoi@inf.ethz.ch
website: https://rycolab.io/authors/alexandra/
Her current interests include formalisms for mildly context-sensitive languages and parsing.

Ryan Cotterell, Assistant professor at ETH Zürich in the Institute for Machine Learning.
email: ryan.cotterell@inf.ethz.ch
website: https://rycolab.io/authors/ryan/
His research focuses on a wide range of topics, including information-theoretic linguistics, parsing, computational typology and morphology, and bias and fairness in NLP systems.

Anej Svete, Ph.D. Candidate at ETH AI Centre at ETH Zürich.
email: anej.svete@inf.ethz.ch
website: https://anejsvete.github.io/
His main research interests lie at the intersection of formal language theory and LMs, where he is working on improving our understanding of the formal properties of modern architectures.

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 5: Tutorial Abstracts), pages 5–5
August 11-16, 2024 ©2024 Association for Computational Linguistics

Computational Linguistics for Brain Encoding and Decoding: Principles, Practices and Beyond

Jingyuan Sun **Shaonan Wang** **Zijiao Chen** **Jixing Li** **Marie-Francine Moens**

Computational linguistics (CL) has witnessed tremendous advancements in recent years, with models such as large language models demonstrating exceptional performance in various natural language processing tasks. These advancements highlight their potential to help understand brain language processing, especially through the lens of brain encoding and decoding. Brain encoding involves the mapping of linguistic stimuli to brain activity, while brain decoding is the process of reconstructing linguistic stimuli from observed brain activities. CL models that excel at capturing and manipulating linguistic features are crucial for mapping linguistic stimuli to brain activities and vice versa. Brain encoding and decoding have vast applications, from enhancing human-computer interaction to developing assistive technologies for individuals with communication impairments. This tutorial will focus on elucidating how computational linguistics can facilitate brain encoding and decoding. We will delve into the principles and practices of using computational linguistics methods for brain encoding and decoding. We will also discuss the challenges and future directions of brain encoding and decoding. Through this tutorial, we aim to provide a comprehensive and informative overview of the intersection between computational linguistics and cognitive neuroscience, inspiring future research in this exciting and rapidly evolving field.

Jingyuan Sun, Postdoc researcher in the Department of Computer Science, KU Leuven, Belgium
email: jingyuan.sun@kuleuven.be
website: https://www.kuleuven.be/wieiswie/en/person/00155742

Jingyuan Sun has published papers in artificial intelligence and natural language processing journals and conferences such as TNNLS, Scientific Data, AAAI, IJCAI, EMNLP, COLING, ECAI, etc. He also serves as a (senior) PC member for these above conferences. He is a reviewer of TPAMI.

Shaonan Wang, Associate Professor in the State Key Laboratory of Multimodal Artificial Intelligence System, Institute of Automation, Chinese Academy of Sciences
email: shaonan.wang@nlpr.ia.ac.cn
website: https://wangshaonan.github.io/

Shaonan Wang has contributed papers to natural language processing journals and conferences such as Information Sciences, TNNLS, ACL, EMNLP, AAAI, and IJCAI. Her work also spans neurolinguistics, with publications in Scientific Data, Brain and language, Cognitive, Affective, and Behavioral Neuroscience, and more. She serves as an editor for Neurobiology of Language and TALLIP and holds the role of curriculum chair at Neuromatch Academy.

Jixing Li, Assistant Professor at in the Department of Linguistics and Translation at the City University of Hong Kong
email: jixingli@cityu.edu.hk
website: https://jixing-li.github.io/

Jixing Li is an Assistant Professor at in the Department of Linguistics and Translation at the City University of Hong Kong. Her research combines NLP models and neuroimaging methods to examine syntactic and semantic analyses in the brain. Her work has been published in the Journal of Neuroscience, Brain and Language, Annual Review of Linguistics, etc. She serves as an ad-hoc reviewer for top journals in the field of neurolinguistics and is on the editorial board of Communications Psychology.

Zijiao Chen, PhD candidate in the Multimodal Neuroimaging in Neuropsychiatric Disorders Laboratory at the National University of Singapore
email: zijiao.chen@u.nus.edu

Zijiao Chen is a PhD candidate in the Multimodal

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 5: Tutorial Abstracts), pages 6–7
August 11-16, 2024 ©2024 Association for Computational Linguistics

Neuroimaging in Neuropsychiatric Disorders Laboratory at the National University of Singapore. Her research primarily centers on representation learning within neuroimaging data and brain decoding. She has published papers in artificial intelligence and neuroimage conferences and journals such as CVPR, OHBM, NCAA, and AD.

Marie-Francine Moens, Full Professor in the Department of Computer Science, KU Leuven
email: `sien.moens@cs.kuleuven.be`
website: `https://people.cs.kuleuven.be/~sien.moens/`
Marie-Francine Moens is a Full Professor in the Department of Computer Science, KU Leuven. She is a fellow of the European Laboratory for Learning and Intelligent Systems (ELLIS). She is an associate editor of the journal IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI). She holds the ERC Advanced Grant CALCULUS (2018-2024) granted by the European Research Council. She used to be the chair of the European Chapter of the Association for Computational Linguistics (EACL) and was a member of the executive board of the Association for Computational Linguistics (ACL). From 2012 to 2016 she was the coordinator of the MUSE project financed by Future and Emerging Technologies (FET) - Open of the European Commission.

Computational Linguistics for Brain Encoding and Decoding: Principles, Practices and Beyond

Jingyuan Sun **Shaonan Wang** **Zijiao Chen** **Jixing Li** **Marie-Francine Moens**

Computational linguistics (CL) has witnessed tremendous advancements in recent years, with models such as large language models demonstrating exceptional performance in various natural language processing tasks. These advancements highlight their potential to help understand brain language processing, especially through the lens of brain encoding and decoding. Brain encoding involves the mapping of linguistic stimuli to brain activity, while brain decoding is the process of reconstructing linguistic stimuli from observed brain activities. CL models that excel at capturing and manipulating linguistic features are crucial for mapping linguistic stimuli to brain activities and vice versa. Brain encoding and decoding have vast applications, from enhancing human-computer interaction to developing assistive technologies for individuals with communication impairments. This tutorial will focus on elucidating how computational linguistics can facilitate brain encoding and decoding. We will delve into the principles and practices of using computational linguistics methods for brain encoding and decoding. We will also discuss the challenges and future directions of brain encoding and decoding. Through this tutorial, we aim to provide a comprehensive and informative overview of the intersection between computational linguistics and cognitive neuroscience, inspiring future research in this exciting and rapidly evolving field.

Jingyuan Sun, Postdoc researcher in the Department of Computer Science, KU Leuven, Belgium
email: jingyuan.sun@kuleuven.be
website: https://www.kuleuven.be/wieiswie/en/person/00155742

Jingyuan Sun has published papers in artificial intelligence and natural language processing journals and conferences such as TNNLS, Scientific Data, AAAI, IJCAI, EMNLP, COLING, ECAI, etc. He also serves as a (senior) PC member for these above conferences. He is a reviewer of TPAMI.

Shaonan Wang, Associate Professor in the State Key Laboratory of Multimodal Artificial Intelligence System, Institute of Automation, Chinese Academy of Sciences
email: shaonan.wang@nlpr.ia.ac.cn
website: https://wangshaonan.github.io/

Shaonan Wang has contributed papers to natural language processing journals and conferences such as Information Sciences, TNNLS, ACL, EMNLP, AAAI, and IJCAI. Her work also spans neurolinguistics, with publications in Scientific Data, Brain and language, Cognitive, Affective, and Behavioral Neuroscience, and more. She serves as an editor for Neurobiology of Language and TALLIP and holds the role of curriculum chair at Neuromatch Academy.

Jixing Li, Assistant Professor at in the Department of Linguistics and Translation at the City University of Hong Kong
email: jixingli@cityu.edu.hk
website: https://jixing-li.github.io/

Jixing Li is an Assistant Professor at in the Department of Linguistics and Translation at the City University of Hong Kong. Her research combines NLP models and neuroimaging methods to examine syntactic and semantic analyses in the brain. Her work has been published in the Journal of Neuroscience, Brain and Language, Annual Review of Linguistics, etc. She serves as an ad-hoc reviewer for top journals in the field of neurolinguistics and is on the editorial board of Communications Psychology.

Zijiao Chen, PhD candidate in the Multimodal Neuroimaging in Neuropsychiatric Disorders Laboratory at the National University of Singapore
email: zijiao.chen@u.nus.edu

Zijiao Chen is a PhD candidate in the Multimodal

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 5: Tutorial Abstracts), pages 8–9
August 11-16, 2024 ©2024 Association for Computational Linguistics

Neuroimaging in Neuropsychiatric Disorders Laboratory at the National University of Singapore. Her research primarily centers on representation learning within neuroimaging data and brain decoding. She has published papers in artificial intelligence and neuroimage conferences and journals such as CVPR, OHBM, NCAA, and AD.

Marie-Francine Moens, Full Professor in the Department of Computer Science, KU Leuven
email: sien.moens@cs.kuleuven.be
website: https://people.cs.kuleuven.be/~sien.moens/
Marie-Francine Moens is a Full Professor in the Department of Computer Science, KU Leuven. She is a fellow of the European Laboratory for Learning and Intelligent Systems (ELLIS). She is an associate editor of the journal IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI). She holds the ERC Advanced Grant CALCULUS (2018-2024) granted by the European Research Council. She used to be the chair of the European Chapter of the Association for Computational Linguistics (EACL) and was a member of the executive board of the Association for Computational Linguistics (ACL). From 2012 to 2016 she was the coordinator of the MUSE project financed by Future and Emerging Technologies (FET) - Open of the European Commission.

Watermarking for Large Language Models

Xuandong Zhao **Yu-Xiang Wang** **Lei Li**

As AI-generated text increasingly resembles human-written content, the ability to detect machine-generated text becomes crucial in both the computational linguistics and machine learning communities. In this tutorial, we aim to provide an in-depth exploration of text watermarking, a subfield of linguistic steganography with the goal of embedding a hidden message (the watermark) within a text passage. We will introduce the fundamentals of text watermarking, discuss the main challenges in identifying AI-generated text, and delve into the current watermarking methods, assessing their strengths and weaknesses. Moreover, we will explore other possible applications of text watermarking and discuss future directions for this field. Each section will be supplemented with examples and key takeaways.

Xuandong Zhao, Ph.D. Candidate in Computer Science at the University of California, Santa Barbara.
email: `xuandongzhao@cs.ucsb.edu`
website: `https://xuandongzhao.github.io`
He is advised by Professors Yu-Xiang Wang and Lei Li. His research combines machine learning and natural language processing to build ethical and trustworthy generative AI systems. He has published several papers in leading NLP/ML conferences such as ACL, NAACL, EMNLP, ICML, AISTATS, and UAI. In addition to his research, Xuandong has actively participated in the academic community, serving on the Program Committee for ACL/EMNLP 2022/2023, ICML 2023, NeurIPS 2023, and ICLR 2024. His professional experience extends beyond academia, having completed internships at industry giants including Microsoft, Google, and Alibaba. Xuandong is the recipient of the prestigious Chancellor's Fellowship from UC Santa Barbara. He obtained his Bachelor of Science degree in Computer Science from Zhejiang University in 2019.

Yu-Xiang Wang, Faculty Member of the Computer Science Department at UCSB.
email: `yuxiangw@cs.ucsb.edu`
website: `https://sites.cs.ucsb.edu/yuxiangw`
Prior to joining UCSB, he was a scientist with Amazon AI in Palo Alto. Even before that he was with the Machine Learning Department at Carnegie Mellon University and had the pleasure of being jointly advised by Stephen Fienberg, Alex Smola, Ryan Tibshirani and Jing Lei. Over the years Yu-Xiang has worked on a diverse set of problems in the broad area of statistical machine learning, e.g., trend filtering, differential privacy, subspace clustering, large-scale learning / optimization, bandits / reinforcement learning, just to name a few. His most recent quests include making differential privacy practical and developing a statistical foundation for off-policy reinforcement learning.

Lei Li, Assistant Professor in the Language Technology Institute at Carnegie Mellon University
email: `leili@cs.cmu.edu`
website: `https://lileicc.github.io`
His research interest lies in natural language processing, machine translation, and AI-powered drug discovery. He received his B.S. from Shanghai Jiao Tong University and Ph.D. from Carnegie Mellon University. His dissertation work on fast algorithms for mining co-evolving time series was awarded ACM KDD best dissertation (runner up). His recent work on AI writer Xiaomingbot received 2nd-class award of Wu Wen-tsün AI prize in 2017. He is a recipient of ACL 2021 best paper award, CCF Young Elite award in 2019, and CCF distinguished speaker in 2017. His team won first places for five language translation directions in WMT 2020 and the best in corpus filtering challenge. Previously, he worked at

ByteDance as the founding director of AI Lab. He has served organizers and area chair/senior PC for multiple conferences including KDD, ACL, EMNLP, ICML, ICLR, NeurIPS, AAAI, IJCAI, and CIKM. He has started ByteDance's machine translation system, VolcTrans and many of his algorithms have been deployed in production.

Author Index