



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU



科技學院
Faculdade de Ciências e Tecnologia
Faculty of Science and Technology



Natural Language Processing & Portuguese –
Chinese Machine Translation Laboratory

自然語言處理與中葡機器翻譯實驗室



優創譯
科技有限公司
UTran-i Technology LTD.

The 2025 Workshop on Multilinguality in the Age of Large Language Models

Celebrating the **15th Anniversary** of the **NLP²CT Laboratory**

From Words to Worlds:
Groundbreaking Innovation in Machine Translation

Meeting Handbook

*Natural Language Processing & Portuguese – Chinese
Machine Translation Laboratory*

July 16, 2025 (Wed) | 9:30 AM - 5:00 PM

Venue: E11 - G015

Sponsors (In Alphabetical Order)



NLP²CT Laboratory
Official Account



Registration Form



Meeting Agenda

Time	Topic	Speakers	Venus
9:30am – 9:50am	Opening	University of Macau Derek F. Wong	E11-G015
Group Photo			
9:50am – 10:40am	Multilinguality in Language and Speech	Johns Hopkins University Philipp Koehn	E11-G015
Tea Break (Poster Session)			
11:00am – 11:50am	Compositional Generalization and LLM for Machine Translation	Westlake University Yue Zhang	
11:50am – 12:10am	Scaling Machine Translation to Long Documents with LLM Agents	Harbin Institute of Technology, Shenzhen Xuebo Liu	E11-G015
12:10am – 12:45am	Poster Session		
Lunch			
2:30pm – 2:45pm	Towards a Personal Foundation Agent	AutoArk Xiaodong Zeng	
2:45pm – 3:00pm	The Road from Natural Language Processing to Language Agent	Alibaba International Longyue Wang	
3:00pm – 3:15pm	From Words to New Worlds	Xunxu Intelligence Yikai Zhou	E11-G015
3:15pm – 3:30pm	Building a High-Quality Data Foundation for AI Large Models Based on DeepData Engine	DeepFuture & DeepTranx Yuchu Lin	
3:30pm – 3:45pm	Efficient Language Modeling for Neural Machine Translation: From NMT to LLMs	ByteDance Jianhui Pang	
Tea Break (Poster Session)			
4:00pm – 4:55pm	Lessons from Academia and Industry: Multilingual AI in the Age of LLM	Panel Discussion	
4:55pm – 5:00pm	Closing	University of Macau Derek F. Wong	E11 Learning Commons



Host & Speakers

Derek F. Wong



Derek F. Wong is an Associate Professor in the Department of Computer and Information Science at the University of Macau, where he leads the Natural Language Processing & Portuguese-Chinese Machine Translation Laboratory (NLP2CT Lab). His research focuses primarily on neural machine translation and natural language processing. Dr. Wong's contributions have been recognized with prestigious awards, including the Second-Class Prize in the Science and Technology Progress Award (2012) and the Second-Class Prize in the Technological Invention Award (2022), both conferred by the Macao Science and Technology Award. He serves as an Associate Editor for ACM TALLIP and IEEE/ACM TASLP, as well as an Action Editor for TACL. Additionally, Dr. Wong has held roles in major conferences, including Program Chair, Senior Area Chair, Area Chair, and Program Committee (PC) Member for leading events such as ACL, EMNLP, ICML, ICLR, NeurIPS, AAAI, IJCAI, MT Summit, CWMT, and NLPCC.

“

Welcome Speech & Review of Groundbreaking Innovations in Machine Translation

Derek F. Wong is an Associate Professor in the Department of Computer and Information Science at the University of Macau, where he leads the Natural Language Processing & Portuguese-Chinese Machine Translation Laboratory (NLP2CT Lab). His research focuses primarily on neural machine translation and natural language processing. Dr. Wong's contributions have been recognized with prestigious awards, including the Second-Class Prize in the Science and Technology Progress Award (2012) and the Second-Class Prize in the Technological Invention Award (2022), both conferred by the Macao Science and Technology Award. He serves as an Associate Editor for ACM TALLIP and IEEE/ACM TASLP, as well as an Action Editor for TACL. Additionally, Dr. Wong has held roles in major conferences, including Program Chair, Senior Area Chair, Area Chair, and Program Committee (PC) Member for leading events such as ACL, EMNLP, ICML, ICLR, NeurIPS, AAAI, IJCAI, MT Summit, CWMT, and NLPCC.

”



Philipp Koehn



Philipp Koehn is Professor of Computer Science at Johns Hopkins University. He received his PhD from the University of Southern California and previously worked as postdoctoral scholar at the Massachusetts Institute of Technology and professor at the University of Edinburgh. He also worked in various positions in industry, such as Meta, Systran, and Omnisien Technology. He is best known for his contributions to machine translation but also more broadly to language and speech processing. He published two textbooks in the field; led the development of Moses, the dominant statistical machine translation open-source toolkit; and published over 200 research papers. He received the Award of Honor from the International Association for Machine Translation and was named a Fellow by the Association for Computational Linguistics.

“

Multilinguality in Language and Speech

Since the field of machine translation adopted a deep learning approach which led to the development of the Transformer models, the various subfields in language and speech processing have come closer together. This talk will present recent research on speech translation and improving how language models can handle many languages beyond their current focus on English.

”

Yue Zhang

Yue Zhang is a tenured Professor at Westlake University (<https://frccchang.github.io>). His research interests include fundamental NLP and its machine learning algorithms, and his recent research focuses on LLM reasoning and AI scientist. His major contributions to the field include machine learning algorithms for structured prediction (e.g., parsing and IE), neural NLP models (i.e., lattice and graph LSTM), and generalization for NLP/LM (e.g., OOD and logical reasoning). He co-authored the Cambridge University Press book “Natural Language Processing -- a Machine Learning Perspective” and served as a PC co-chair for CCL 2020 and EMNLP 2022, test-of-time award committee co-chairs for ACL 2024 and 2025, action editor for TACL, and associate editor for TASLP, TALLIP, TBD, and CSL. He won the best paper awards of IALP 2017 and COLING 2018, best paper honorable mention of SemEval 2020, and best paper nomination for ACL 2018 and ACL 2023.



“

Compositional Generalization and LLM for Machine Translation

Neural MT has witnessed significant advances recently, yet they still face challenges. In this talk, I will share some observations on the generalization issues of MT, and specifically compositional generalization. Quantitative results will be given, and some method to address generalization issues will be discussed. In addition, I will share some comparisons between LLM translation and professional translators, which reveal similarities and difference. Finally, I will discuss the collaboration between LLM translators and human users.

”



Xuebo Liu



Xuebo Liu is an Associate Professor at the School of Computer Science, Harbin Institute of Technology, Shenzhen. In 2021, he earned his Ph.D. from the Faculty of Science and Technology at the University of Macau. His research focuses on natural language processing, Data-Centric AI, domain-specific large language models, and machine translation. He has published over 40 papers in top-tier conferences such as ACL, EMNLP, ICLR, and NeurIPS, with more than 30 as the first or corresponding author. He serves as (Senior) Area Chair for ACL 2022/2024, EMNLP 2022-2025, and NAACL 2024-2025.

“

Scaling Machine Translation to Long Documents with LLM Agents

Despite their remarkable progress in machine translation, large language models (LLMs) still struggle with the inherent challenges of translating lengthy documents like novels. The primary obstacles—*intrinsic token limits and translation omissions*—critically undermine terminological consistency and narrative coherence. This talk presents our agent-based framework for addressing these deficiencies. We introduce a methodology that leverages Monte Carlo Tree Search (MCTS) to generate synthetic, domain-specific long-chain reasoning data. Furthermore, our approach offers valuable insights into adapting foundational LLMs to specialized tasks, using machine translation as a representative example of a complex domain.

”

Xiaodong Zeng

Xiaodong Zeng is a seasoned leader in the technology industry with over a decade of experience in AI, IoT, and product innovation. His journey spans significant roles including founding CEO at AutoArk Technology, and CTO at Ant Group where he spearheaded AI innovations. He has a robust academic background in Natural Language Processing and Machine Translation from the University of Macau, which complements his extensive professional tenure. Notably, Xiaodong has led his teams through critical growth phases at major platforms like Alipay, significantly enhancing user engagement and operational efficiency through innovative algorithmic strategies. His pioneering work in AI and Green Technology has been recognized globally, making him a recipient of prestigious awards like the MIT Technology Review's Innovators Under 35.



“

Towards a Personal Foundation Agent

Two "foray areas" for AI Agent commercialization: Vertical domains and personal domains: Why build a "foundational" personal agent? What are its three key elements? The pursuit of optimal interaction: Our journey in developing an end-to-end multimodal model like GPT-4o.

”

Longyue Wang



Longyue Wang is currently a Senior Algorithm Expert at AI Business, Alibaba International. He received his M.Sc. degree from the University of Macau in 2014 and his Ph.D. from Dublin City University in 2018. His research interests include large language models, multimodality, natural language processing, and machine translation. He has published more than 80 papers in leading AI journals and conferences, such as Nature Communications, IEEE TPAMI, ICLR and ACL. He has received the Best Thesis Award from the European Association for Machine Translation (EAMT) and a Best Paper Award Nomination at ACM MM 2024.

The Road from Natural Language Processing to Language Agent

Language is not only a medium for communication but has also become a unified interface that empowers artificial intelligence (AI) with perception, action, and collaboration. The advent of large language models (LLMs) has shifted our focus from “AI research for natural language processing (NLP)” to “General AI research based on NLP.” This talk will briefly review the evolution of NLP, introduce how LLMs expand the capabilities of language agents—including autonomous decision-making, planning, and executing complex tasks—and look ahead to their applications across various domains as well as future challenges.

Yikai Zhou

Yikai Zhou holds a Master's degree from the NLP2CT Laboratory at the University of Macau and is currently pursuing a Ph.D. in Artificial Intelligence at the Macau University of Science and Technology (MUST). As an NLP researcher, Yikai Zhou has published papers in prestigious academic conferences such as ACL and EMNLP. His primary research focus lies in the bionic self-awareness and self-evolution of intelligent agents. Yikai Zhou is also a contributor to the drafting of several industry standards in the field of artificial intelligence in Shenzhen. In the industry, Yikai Zhou founded Sequence Logic, a company dedicated to intelligent technology. He also serves as a senior AI consultant for UBTECH Robotics, a leading humanoid robotics company. Yikai Zhou is committed to applying intelligent agents in commercial practices and integrating them with embodied intelligence.



From Words to New Worlds

Reconnect with old friends and explore the transformative power of artificial intelligence in this engaging workshop. We'll discuss the real-world impacts of current AI technologies on industry, sharing practical insights and experiences that shape our understanding of its potential. Together, we'll delve into predictions for the future, examining how AI will continue to reshape our world.

Yuchu Lin

Dr. Lin Yuchu, Ph.D. in Artificial Intelligence, is the Founder and CEO of DeepFuture & DeepTranx, dedicated to building an AI ecosystem. With a robust research foundation, he has participated in the Parallel Corpus Construction project at the University of Macau's NLP Laboratory and Big Data Mining initiatives at the DACC Laboratory. Regarding intellectual property, Dr. Lin holds 13 authorized invention patents (including 5 data patents, 3 algorithm patents, and 5 model patents) and 32 software copyrights. The companies he founded have obtained authoritative certifications including High-Tech Enterprise, Specialized, Refined, Differential, and Innovative SME, Gazelle Cultivation Enterprise. Notably, Dr. Lin has established a closed-loop "Data-Model-Application" ecosystem that advances digital economy cooperation between Macao and Portuguese-speaking countries, and this ecosystem putting energy into cross-border data flow.

“

Building a High-Quality Data Foundation for AI Large Models Based on DeepData Engine

We pioneer multi-modal data mining and multiple LLM platforms for Portuguese-speaking markets, the minor Belt & Road languages, and major linguistic domains. Cultivating data assets with deep expertise, Driving innovation as China's premier data provider - the Scale.AI of CHINA.

”

Jianhui Pang

Jianhui Pang is a Natural Language Processing Engineer specializing in Machine Translation (MT) and Efficient Large Language Models (Efficient-LLM). He holds a PhD from the University of Macau, where he pioneered modular architectures for traditional neural machine translation systems and developed an anchor-based training methodology for efficient inference models, achieving dramatic KV cache compression rates of up to 99%. His research has been published in premier NLP conferences and journals, including ACL, COLING, Computational Linguistics (CL), and Transactions of the Association for Computational Linguistics (TACL). Jianhui has gained significant research experience through internships at Alibaba DAMO Academy, Tencent AI Lab, and Alibaba's Tongyi Lab. His current research interests focus on data synthesis techniques and advancing the efficiency of large language models. He has recently held a position as an Algorithm Engineer at ByteDance.



“

Efficient Language Modeling for Neural Machine Translation: From NMT to LLMs

This talk presents efficient techniques for machine translation before and after the rise of large language models (LLMs). We first review traditional neural machine translation methods such as back-translation, data rejuvenation, and modular model designs aimed at improving data and model efficiency. Then, we highlight how LLMs bring new opportunities through supervised fine-tuning with small datasets, in-context learning, LoRA for efficient adaptation, and multilingual capabilities. We also touch on recent advances like KV cache compression and future directions including multimodal and real-time translation tasks.

”



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU



科技學院
Faculdade de Ciências e Tecnologia
Faculty of Science and Technology



Natural Language Processing & Portuguese –
Chinese Machine Translation Laboratory

自然語言處理與中葡機器翻譯實驗室



優創譯
科技有限公司
UTran-i Technology LTD.

Sponsors



AutoArk
无界方舟

AutoArk



深绎未来
DEEPFUTURE

深译科技
DEEPTANX

DeepFuture Technology
(Guangdong Hengqin) Co. Ltd.



新译科技
NewTranx Tech

NewTranx Information
Technology (Shenzhen) Co. Ltd.



寻序智能

SequenceLogic
(Shenzhen) Co., Ltd.