



中國計算機學會
CHINA COMPUTER FEDERATION



中国中文信息学会
Chinese Information Processing Society of China

第十届语言与智能高峰论坛 The 10th Language and Intelligence Summit

2025.08.10

会议手册



中国·乌鲁木齐



组委会名单

- **程序委员会主席**

毛先领 北京理工大学

任昭春 莱顿大学

- **组织委员会主席**

杨雅婷 中国科学院新疆理化技术研究所

王昊奋 同济大学

- **青年科学报告主席**

林成华 曼彻斯特大学

许嘉蓉 复旦大学

- **竞赛论坛主席**

曹云波 中国人民解放军军事科学院

纪焘 复旦大学

吴华 百度

- **宣传主席**

覃立波 中南大学

赵森栋 哈尔滨工业大学



第十届语言与智能高峰论坛日程

八月十日

(火焰山厅)

特邀报告

时间	内容	主讲人	主持人
09:45-10:00	开幕式		杨雅婷
10:00-10:10	吾守尔院士捐赠仪式		
10:10-11:10	基于具身场景认知的视觉导航	蒋树强	毛先领
11:10-12:10	飞桨框架 3.0 加速大模型技术创新	胡晓光	
12:10-12:25	茶歇		
12:25-13:25	生成式语音与音乐处理技术	谢磊	任昭春
13:25-13:55	Klear:快手语言大模型的技术探索和应用	张富峥	
13:55-15:10	午餐休息		
15:10-16:10	跨越 AI 技术周期:大模型的演进路径与落地实践	黄非	

青年科学家报告

时间	内容	主讲人	主持人
16:10-16:40	Beyond One-Size-Fits-All: Inversion Learning for Highly Effective NLG Evaluation Prompts	林成华	许嘉蓉
16:40-17:10	智能体的革命:聚焦核心要素与能力演进	雷文强	
17:10-17:25	茶歇		
17:25-17:55	语音大模型技术及应用进展	孔常青	林成华
17:55-18:25	情智兼备的高表现语音对话研究	赵洲	

竞赛论坛(百度)

时间	内容	主讲人	主持人
18:25-18:40	LIC2025 命题立意分享	张叔夏:百度 AI 技术生态数据生态运营负责人	赵森栋
18:40-18:50	赛题&数据集分享*3	林肯: MyTVSuper Limited 原创内容主管	
18:50-18:55	颁奖仪式		
18:55-19:25	冠军选手分享*3		
19:25-19:35	闭幕式		王昊奋

蒋树强

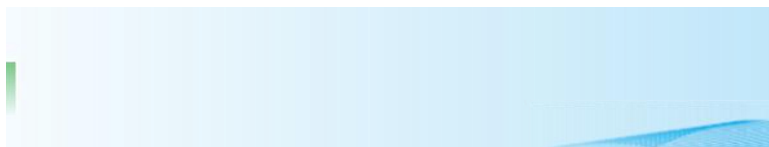


报告题目：基于具身场景认知的视觉导航

摘要：具身智能是真实物理世界中人工智能的重要表现形态，具身导航是指智能体根据任务目标，感知与理解周围环境并执行移动动作完成任务，这是具身智能系统与真实世界交互的关键技术之一。生理学与认知科学研究表明，场景认知在导航任务中起着至关重要的作用，智能体不仅通过当前观察来感知环境，还能通过过往经验和记忆对未观测的场景进行推测，如何将场景认知与导航任务相结合，为智能体提供预想与决策能力，是一项值得研究的重要问题。本报告将首先介绍具身智能与具身导航的研究背景，并汇报基于具身场景认知的视觉导航研究进展，包括基于场景知识图的物体导航、基于网格记忆地图的视觉语言导航、基于神经辐射表征前瞻探索的视觉语言导航等技术，最后介绍具身导航从虚拟到真实环境的适配并给出演示。

简介：蒋树强，中国科学院大学特聘教授，博士生导师，先后担任期刊《IEEE TMM》、《ACM ToMM》、《IEEE Multimedia》、《计算机研究与发展》、《JCST》、《CAD 学报》编委，中国人工智能学会具身智能专委会主任、中国计算机学会多媒体专委会副主任、中国自动化学会网络计算专委会副主任、ACM SIGMM 中国分会副主席，研究方向是多媒体内容分析和具身智能技术。主持承担科技创新 2030-“新一代人工智能”重大项目、国家自然科学基金青年基金 A 类(杰青)、B 类(优青)、重点等项目 20 余项，发表论文 200 余篇，获授权专利 20 余项，多项技术应用到实际系统中，先后获省部级或学会奖励 5 项。

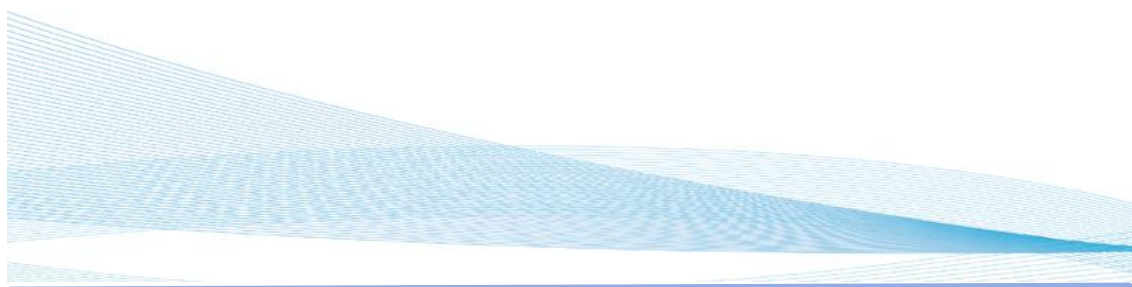
胡晓光



报告题目：飞桨框架 3.0 加速大模型技术创新

摘要：深度学习框架作为连接底层算力芯片与上层模型应用的核心基础软件，正加速大模型等技术创新的步伐。2025 年 4 月，飞桨框架正式发布 3.0 版本；2025 年 6 月，基于飞桨研制的文心 4.5 系列模型正式开源。本次报告将阐释飞桨框架 3.0 的设计思想；着重展示其核心特性：动静统一自动并行、大模型训推一体、科学计算高阶微分、神经网络编译器、异构多芯适配；并以文心 4.5 系列模型为例，探讨如何使用飞桨框架 3.0 以加速大模型等前沿领域技术创新与产业应用。

简介：胡晓光，正高级工程师，现任百度深度学习技术平台部杰出架构师，主要研究方向包括自然语言处理、深度学习框架、科学智能等。设计了飞桨 2.0 全新 API 体系，打造飞桨动静统一、训推一体特色；突破飞桨框架 3.0 自动并行、神经网络编译器、高阶自动微分等技术，研制飞桨科学计算工具组件赛桨，实现微分方程高效求解，助力科学探索。研究成果曾获国家科技进步二等奖、中国电子学会科技进步一等奖、北京市科技进步一等奖等。



谢磊



报告题目：生成式语音与音乐处理技术

摘要：本次报告将和大家一起探讨生成式语音与音乐处理技术前沿技术。首先交流基于语言模型的语音合成与增强技术，特别关注语音合成的风格化和灵活可控。而后探讨说唱生成和音乐生成，包括以歌词与伴奏为输入的高质量说唱合成技术和基于扩散模型的全曲生成技术。随后介绍一个开源的语音理解大模型的构建过程，以及在此基础上如何构建的基于语音理解的共情语音对话模型。最后对未来研究方向和趋势进行展望。

简介：谢磊，西北工业大学教授/博导，音频语音与语言处理实验室(ASLP@NPU)负责人。曾在比利时布鲁塞尔自由大学、香港城市大学和香港中文大学工作，获得教育部新世纪优秀人才支持计划，陕西省青年科技新星、全球前2%顶尖科学家（斯坦福大学&Elsevier）、华为云 AI 名师等荣誉。在音频语音相关领域发表论文 400 余篇，Google Scholar 引用 1.4 万次。多次获得学术会议最佳论文奖和国际评测冠军。众多研究成果在产业界落地应用。担任国际语音通信协会 ISCA 中文口语语言处理兴趣组(SIG-CSLP)副主席、IEEE 语音和语言技术委员会（IEEE SLTC）委员、中国计算机学会语音听觉与对话专委会常务委员、IEEE/ACM Transactions on Audio, Speech and Language Processing 和 IEEE Signal Processing Letters 高级领域编委（SAE）等。

张富峥



报告题目：Klear：快手语言大模型的技术探索和应用

摘要：Klear 是快手公司从 0 到 1 独立自主研发的一系列语言大模型及 Agent 系统，包括模型的预训练、对齐、推理能力增强以及 Agent 系统建设等。全面专业的评测结果表明，Klear LLM 模型在同等尺寸下达到了行业的先进水平。在技术探索上，我们在合成数据、形式化数学证明、Diffusion LLM 等前沿方向也有创新性的成果。此外，我们正在搭建 Deep Research Agent，基于端到端创新性的 RL 算法，在相关评测中效果有了显著的提升。

简介：张富峥博士，长期从事大模型、自然语言处理(NLP)、知识图谱(Knowledge Graph)和搜索推荐等 AI 领域的工作，在 AI 技术研发创新、产品应用落地、技术团队管理和跨团队协作等方面均拥有丰富的工作经验。目前在快手公司担任 General Manager，语言大模型中心负责人，负责 Klear 系列大模型的技术建设及相关业务应用。在此之前，富峥于 2015-2018 期间在 MSRA 担任研究员，负责知识图谱和推荐系统相关的研究以及在小冰、Bing 新闻等业务场景的应用，于 2018-2021 期间在美团担任技术总监，负责知识图谱、预训练和商品搜索等方向的技术研发和业务应用。此外，富峥在 ACL、KDD、WWW、EMNLP 等相关领域的顶级会议/顶级期刊上发表论文 70 余篇，谷歌引用 15000+，长期担任相关学术活动的评委。连续多年入选了斯坦福大学发布的“全球前 2% 顶尖科学家榜单”和“人工智能全球最具影响力学者榜单 AI 2000”。

黄非

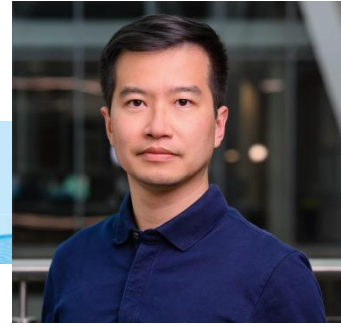


报告题目：跨越 AI 技术周期：大模型的演进路径与落地实践

摘要：本次演讲将以“技术周期”为线索，系统回顾人工智能特别是大模型的发展演化，分析当前主流大语言模型的技术路线、能力边界与产业落地挑战。我们将探讨如何从“感知智能”迈向“认知智能”，以及在算力、数据、安全、成本等现实约束下，大模型如何实现可持续、可控、可用的路径选择。演讲还将介绍阿里巴巴在大模型领域的最新进展，包括开源的“千问”模型体系（Qwen）、多模态融合、社区生态建设与产业应用实践，展示我们如何在开放共建的理念下，推动大模型从“实验室智能”走向“产业级智能”。

简介：黄非博士，长期从事自然语言处理和人工智能方向研究开发。目前担任阿里巴巴通义自然语言智能实验室负责人兼阿里云副总裁。他带领团队研发通义自然语言大模型体系，在机器阅读理解(MRC)，图文问答（VQA）和中文理解（CLUE）等任务上实现首次超越人类结果。以通义千问大模型为基础，赋能软件开发，角色对话，办公效率，智能司法等行业合作伙伴。他在人工智能顶级会议和期刊发表文章 300+篇，国际国内专利 100+项，谷歌引用 20000+，H-index 65，曾担任 ACL，TACL 等学术会议领域主席和期刊编辑等。

Chenhua Lin

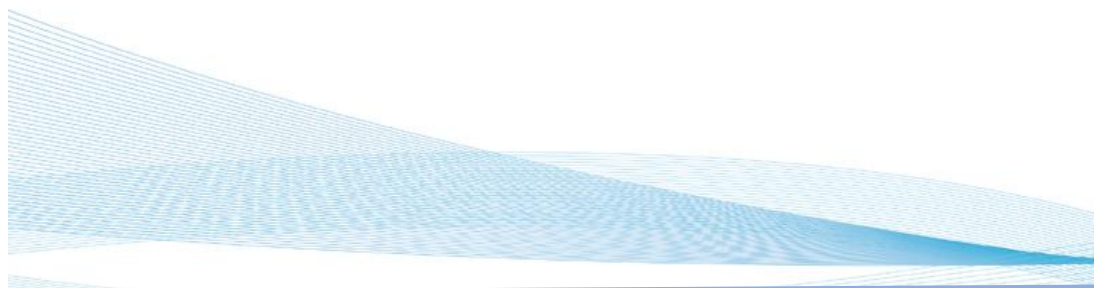


Title: Beyond One-Size-Fits-All: Inversion

Learning for Highly Effective NLG Evaluation Prompts

Abstract: Evaluating natural language generation (NLG) systems is inherently challenging. While human evaluation remains the gold standard, it is difficult to scale and often suffers from inconsistencies and demographic biases. LLM-based evaluation offers a scalable alternative but is highly sensitive to prompt design, where small variations can lead to significant discrepancies. In this talk, I will introduce an inversion learning method that learns effective reverse mappings from model outputs back to their input instructions, enabling the automatic generation of highly effective, model-specific evaluation prompts. This method is simple, requires only a single evaluation sample, and eliminates the need for manual prompt engineering, thereby improving both the efficiency and robustness of LLM-based evaluation.

Bio: Chenghua Lin is a Full Professor and Chair in Natural Language Processing in the Department of Computer Science at The University of Manchester. His research lies at the intersection of machine learning and natural language processing, with a focus on language generation, multimodal LLMs, and evaluation methods. He currently serves as Chair of the ACL SIGGEN Board, a member of the IEEE Speech and Language Processing Technical Committee, and Associate Editor for Computer Speech and Language. He has received several prizes and awards for his research and academic leadership, including the CIKM Test-of-Time Award, the INLG Best Paper Runner-up Award, and an Honourable Mention for the Scottish Informatics and Computer Science Alliance (SICSA) Supervisor of the Year Award. He has also held numerous program and chairing roles for *ACL conferences, including Documentation Chair for ACL'25, Publication Chair for ACL'23, Workshop Chair for AACL-IJCNLP'22, Program Chair for INLG'19, and Senior Area Chair for EMNLP'20, ACL'22-'23, EACL'23, NAACL'25, and AACL'25.



雷文强



题目：智能体的革命：聚焦核心要素与能力演进

摘要：当前，大模型智能体技术正在经历一场深刻的范式变革。这一变革以三大核心要素为支撑：信息组织的统一化架构、系统设计的模块化范式以及行为能力的自主化机制。基于这些基础性突破，智能体系统实现了向高阶能力的演化：流程自动化效能的显著提升；交互式信息获取与智能决策的协同优化；知识体系的动态积累与诱导式演进。这种多维度的能力跃迁形成了自我强化的正向发展循环。此次技术革命不仅从根本上拓展了智能体的能力边界，同时凸显出关键性技术挑战与未来演进路径。同时，它也为人机协同范式创造了全新的可能性空间，将推动智能系统与应用生态的范式升级。

简介：雷文强，四川大学计算机学院院长助理、教授、博士生导师，国家级青年人才，博士毕业于新加坡国立大学。从事基于自然语言处理、信息检索、人机交互系统研究。先后主持国家重点研发计划课题、国家自然科学基金面上项目等，并以第一作者或者通讯作者发表中国计算机学会 A 类长文（CCF-A）数十篇，多篇一作/通讯文章两年内引用过百，获 ACM MM2020 最佳论文奖，ACL 2024 领域主席奖。多次在国际顶级会议，比如 ACL，SIGIR 上做 Tutorial。先后担任各大顶级国际会议比如 ACL，KDD，AAAI，IJCAI，WSDM，EMNLP 等（高级）程序委员会委员，并担任新加坡全国自然语言处理会议 SSNLP2021 的程序委员会主席、以及担任重要期刊 ACM Trans. on Web 的客座编委。

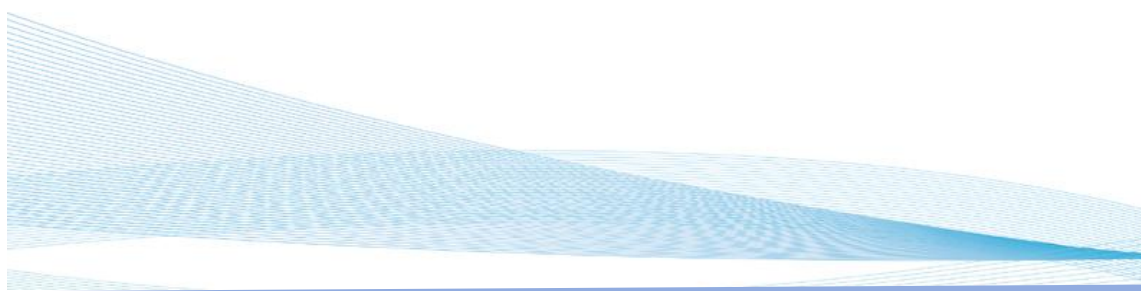
孔常青



题目：语音大模型技术及应用进展

摘要：在国内外认知大模型发展迅速的背景下，语音大模型的研发也逐渐成为业界主流；报告首先对国内外语音大模型的发展趋势进行解读，然后对讯飞在语音合成、语音识别、语音交互以及语音同传等各个领域的语音大模型研发情况进行展开介绍，最后对相关技术在智能汽车座舱、翻译机、透明屏、办公本等典型场景的落地应用成效进行介绍。

简介：2015年毕业于中国科学技术大学，长期从事智能语音相关技术的研发和产品化工作，主要研究方向包括语音识别、语音唤醒以及语音大模型等方向。先后带队参与语音识别算法优化、多语种语音识别系统研发、语音大模型研发等公司级重点项目，有丰富的核心技术研究、产品化和平台化的工作经验，并推动了多项成果在产品中落地，申请发明专利 63 项，产生了良好的经济和社会效益。



赵洲



题目：情智兼备的高表现语音对话研究

摘要：语音对话系统的出现为大语言模型提供了一个全新的交互入口,使其从“指尖的工具”进化为“身边的伙伴”。但大语言模型在接入语音模态后,智商显著下降,且在理解用户情绪、进行共情交流等“情商”维度上仍显不足,导致交互体验生硬、缺乏人性化。本报告聚焦于“情智兼备的高表现语音对话研究”这一前沿课题,旨在打破“智商”与“情感”的壁垒,探索构建一个全新的对话系统框架。该框架将融合上下文深度理解推理与多模态情感计算。我们的核心目标是让系统不仅能高效解决用户的实际问题(高智商),更能敏锐感知其语气、情感与潜在意图,并给予富有同理心和个性化的回应(高情商)。通过此项研究,我们期望推动语音对话从“功能型工具”向“伙伴型伴侣”的转变,打造出更自然、更可信、更具吸引力的交互体验,为下一代人机交互树立新的标杆。

简介：赵洲,浙江大学计算机学院教授、博士生导师、国家青年人才、科技部2030重点研发计划青年科学家,主持国家自然科学基金联合基金重点、浙江省杰青等项目。主要研究方向为自然语言理解、计算机视觉计算和生成式模型,在国际期刊TPAMI和会议NeurIPS, ICML, ICLR, CVPR等上发表100余篇论文,谷歌学术引用2万+,相关技术被应用于微软、字节、Stability AI、华为等公司,获2022年度教育部科技进步一等奖、2023年度上海市科学技术一等奖、2024年度人工智能学会科技进步一等奖、2021年度中国电子学会科技进步一等奖、2022年浙江大学竺可桢学院十佳专业导师,连续4年(2021-2024)入选“全球前2%顶尖科学家榜单”,连续2年(2023-2024)入选“中国高被引学者”,指导的多位学生获华为天才少年奖和浙江大学竺可桢奖学金。

Host



Platinum Sponsors



Golden Sponsors

