



自然语言处理团队

Natural Language Processing Group

data layout evaluation
understanding interaction
type linguistics processed
computer machine text human word
design intelligence science tag process automatic
summarization & intelligence
[ɒ̩ˈtɒmətɪk]
[ˈæɪnəlɪzɪs]
[ˈnætʃrəl]
[ˈleɪnɪŋ]
[ˈstætɪstɪkəl]
[ˈnjuːz]
[ˈkɪwɔːrdz]
[ˈɪnpuːt]
[ˈtɪpɒɡrəfi]



■ 团队介绍

北京交通大学自然语言处理研究组（语言智能与大数据处理研究所）成立于2009年8月，现已发展为具有国际先进国内领先的从事自然语言处理研究的科研团体之一。研究室现拥有教授1人，副教授2人、讲师2人、博士生、硕士生等50余人组成的研究队伍。研究组以自然语言理解为核心技术，在重视包括分词、句法分析、语义分析、语言模型等基础研究的同时，注重机器翻译、信息抽取、大数据处理、语音识别与合成、对话系统、文本情感计算、智能机器人、智能人机接口和资源构建、多语种多模态大模型及其应用、认知智能、自然语言生成、机器写作等多方面的应用研发。



迄今为止，研究组已培养毕业博士和硕士100余人。曾先后与中国科学院计算技术研究所、中国科学院自动化研究所、日本情报通信研究机构（NICT）等国际知名企业及研究机构合作研发，与美国、日本、加拿大、欧洲、新加坡、香港、台湾等多个国家或地区建立了合作关系。



研究组在AI赋能轨道交通方向有丰富的学科交叉前沿研究经验，在自然语言处理、机器翻译和大语言模型等方面积累了丰富的模型研究与系统研发经验，主持相关领域的国家自然科学基金项目、国家重点研发计划、863计划专项课题、科技部中日合作项目和科技部中越合作项目等若干项国家项目和国际合作项目，取得了一系列重要的研究成果。

研究组本着“团结、合作、创新、发展”的实验室文化，以精益求精的务实风格和勇于探索的求真精神，精诚合作，开拓创新。今后，我们将以自然语言理解为核心，以提高社会生产力为目标，努力开拓创新，将研究成果在机器翻译、文本大数据处理（云计算）、人机对话系统、语音语言智能控制等应用领域进行成果转化。



导师介绍



徐金安

教授，博士生导师

北京交通大学语言智能与大数据处理研究所主任 1992年毕业于北方交通大学，获学士学位；同年被分配到铁道部第四勘测设计院上海分院工作，1997年任工程师职务。1999年至2006年留学日本，分别于2003年和2006年获得日本国北海道大学大学院工学研究科电子情报专攻自然语言处理专业工学硕士和工学博士学位；2006年至2009年就职于日本电气株式会社中央研究院研究员。2009年归国，任教于北京交通大学计算机学院。CCF杰出会员，中国计算机学会自然语言处理专委会等委员、中文信息学会专委会委员、人工智能学会多语信息处理专委会副秘书长。

研究方向：机器翻译、自然语言处理、人机交互、大数据处理、文本情感分析、对话系统、知识图谱、多语种多模态大模型及其应用等。

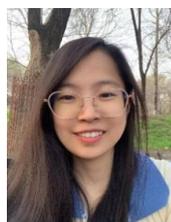


陈钰枫

副教授，博士生导师

1999年至2003年于北京交通大学获得学士学位；2003年至2008年于中科院自动化所获得博士学位；2008年至2014年任职于中科院自动化所模式识别国家重点实验室，2014年至今任教于北京交通大学。

研究方向：多语言文本处理、机器翻译、知识图谱、对话系统、多语种多模态大模型等。



韩文娟

高聘副教授，硕士生导师

中国科学院大学博士，加州大学洛杉矶分校访问学者，新加坡国立大学(NUS)研究员。EMNLP 2022研讨会(UM-IoS Workshop)主席，ACM SIGAI CHINA 新星奖，铁科院审稿顾问，北京通用人工智能研究院专家顾问。在国内外公认的本学科权威学术刊物上发表40余篇。研究方向为多模态大模型，及在综合交通等安全关键领域的应用。

研究方向：自然语言处理、多模态信息处理、认知智能、大模型等、AIGC。



刘健

讲师，硕士生导师

2009年至2013年于东北大学获得学士学位；2013年至2015年于东北大学获得硕士学位；2016年至2020年于中国科学院自动化研究所获得博士学位；2020年至今任教于北京交通大学计算机学院。

研究方向：自然语言处理、知识图谱、信息抽取、多语言信息抽取。



黄锴宇

讲师，硕士生导师

大连理工大学计算机和日语双学士、博士，清华大学智能产业研究院博士后，合作导师为刘洋教授。多年来从事机器翻译、预训练模型与自然语言处理等相关领域的研究。围绕多语言自然语言处理、对话建模、迁移学习及持续学习研究，目前已发表40余篇高水平国际会议/期刊论文。曾获2023年ACL杰出论文奖、优秀博士论文等荣誉。

研究方向：多语言信息处理，对话问答，地理科学智能，信息检索。



■ 研究内容

自然语言处理：包括语音识别、自然语言理解、生成、人机交互及所涉及的中间阶段，是人工智能和计算机科学的交叉学科。自然语言处理涉及词法，句法分析等基础任务，机器翻译，自动问答等核心技术，并将其应用于推荐系统，搜索引擎等领域。



随着深度学习的发展，自然语言处理具有诸多研究热点，正处于黄金发展阶段。尤其是 2022 年 ChatGPT、GPT4.0 等多语种多模态大规模预训练模型的诞生，把 NLP 技术的发展上升到了通用人工智能和认知智能的新阶段！

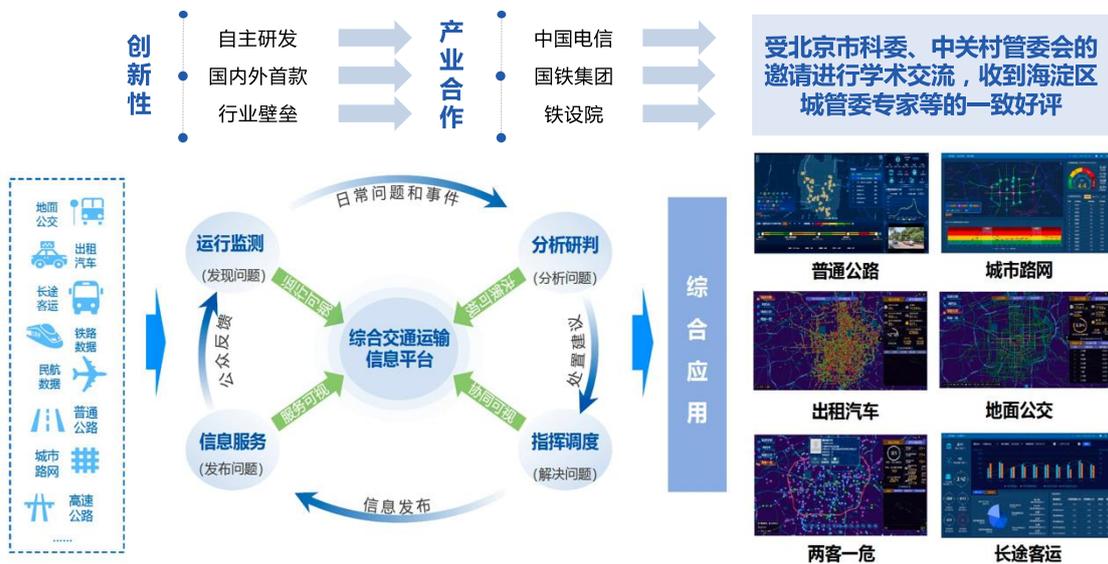
■ 研究成果

本研究组的代表性成果是在日汉机器翻译研究方面，开创性地提出了融合格依存文法的日汉依存树到串统计机器翻译理论并构建了翻译系统，并且我们还自主研发了端到端神经网络机器翻译系统，在国际国内的机器翻译评测任务中，先后夺得 WAT2016 夺得 3 个第一、CCMT2020 夺得 4 个第一、WMT2020 和 WMT2021 分别夺得英汉第一、CCMT2021 夺得 5 个第一名。



研究成果

依托“交通大数据与人工智能”教育部重点实验室、联合中国计算机学会智慧交通分会，发布了国内外首款交通大模型 TransGPT·致远。将城市中与交通相关的人、车、路、环境数据汇集到一个平台，准确、实时、精准的掌握交通态势，主动发现交通异常、分析原因、参与调度。



研究组累计翻译出版译著 4 部、发表学术论文 200 余篇，其中包括 EMNLP、ACL、AAAI、IJCAI、COLING、TALLIP、TPAMI、TASLP、AI 等国际顶会和期刊 50 余篇，申报国家级发明专利 50 余项。连续 11 年校级优秀硕士毕业论文（赖思羽等）。其中包括 2021 年发表顶会论文 17 篇、2022 年 21 篇，2023 年 ACL 录用 7 篇，投稿 EMNLP2023 累计 19 篇，2024 年 ACL 录用 5 篇。申报发明专利 50 余项。

研究成果



在研及完成的科研项目主要包括：科技部科技创新 2030 重大专项课题、国家重点研发计划项目子任务、国家自然科学基金面上项目、北京市自然科学基金、科技部中日国际合作项目，科技部中越合作项目等重要科研项目 30 余项。



■ 学长风采



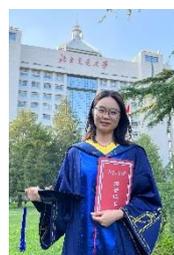
刘直进

2021 年毕业，以大咖 OFFER 就职于腾讯微信模式识别研究组，学硕在读期间共发表 ACL、AAAI、EMNLP 等顶级会议 3 篇，夺得 WMT2020 机器翻译评测英汉方向第一名。



梁云龙

2020 级计算机科学与技术专业博士研究生，中共党员，师从徐金安教授。曾荣获 2 次国家奖学金和 1 次知行奖学金提名奖。科学研究方面能力突出，专注学科前沿，在导师的指导下，致力于攻克对话机器翻译课题中译文不流畅、译文不能充分传达对话特性等难题，已发表 16 篇国际顶级期刊或会议论文（均为一作或共一，其中 CCF A: 10 篇），Google 论文引用次数 792；曾荣获 WMT 国际机器翻译评测对话翻译 2 个世界冠军、国家级人工智能类竞赛亚军以及企业级亚军等成绩。



赖思羽

2020 级学硕，曾获国家奖学金、北京市优秀毕业生和校级优秀毕业论文。在校期间一作发表 nlp 领域顶会论文 3 篇，专利和软著各一项。目前就职中国人寿研发中心，从事算法相关工作。



梁晨

21 级硕士，硕士期间从事信息抽取的研究，参与文心一言推理能力建设，共发表包含 ACL、EMNLP 以及 SCI 1 区期刊在内的 4 篇论文，获得北京交通大学硕士优秀学位论文，现已入职百度。



■ 学长风采



吴培昊
2015 年毕业
字节跳动 AI lab



刘江鸣
2015 年毕业
爱丁堡大学读博



明芳
2017 年毕业
铁科院



李少童
2018 年毕业
腾讯



刘方旭
2019 年毕业
腾讯

更多去向

毕业年份	去向单位
2023 年	百度、字节跳动、创新工场、中国交通银行上海分行
2022 年	百度、字节跳动、中国工商银行、京东研究院
2021 年	腾讯微信、字节跳动、中国农业银行北京分行、创新工场
2020 年	腾讯 AI 研究院、腾讯 CSIG 智能平台部、字节跳动 AI lab、国家开发银行上海分行、中国银行北京分行、中国工商银行上海分行
2019 年	腾讯 CSIG 智能平台部、中国移动通信有限公司研究院、字节跳动 AI lab、360 算法中心、航天信息股份有限公司、北京日报社技术部、京东 AI 平台与研发部、中国银行国际金融研修院
2018 年	腾讯 AI 研究院、搜狗 AI 研究部、苏州移动研究院、航天信息股份有限公司、国家电网、中国银行
2017 年	华为诺亚方舟实验室、建信基金、北京市专利局、中国铁道科学研究院、北京市建筑设计研究院有限公司、中国银行



■ 团队瞬间



■ 欢迎优秀的你加入



嗨! 🤝 欢迎加入北交 NLP 团队!

BJTU-NLP 欢迎你 (中文)

BJTU-NLP ལོ་ 10 ལོ་ལྷན་སྐྱེས་ལྷན་པུ་སྐྱེས་ (蒙文)

BJTU-NLP ལྷན་སྐྱེས་ལྷན་པུ་སྐྱེས་ (藏文)

سنزنى



群聊: 2024年-保研夏令营-NLP
团队群



该二维码7天内(7月22日前)有效, 重新进入将更新





北京交通大学 计算机科学与技术学院
BEIJING JIAOTONG UNIVERSITY School of Computer Science & Technology

知行

北京交通大学自然语言处理团队制作

Made by BJTU-NLP

2024年7月