大数据分析与应用技术国家工程实验室

NATIONAL ENGINEERING LABORATORY FOR BIG DATA ANALYSIS AND APPLICATIONS

大数据分析与应用技术
国家工程实验室
NATIONAL ENGINEERING LABORATORY FOR
BIG DATA ANALYSIS AND APPLICATIONS

大数据分析与应用技术
国家工程实验室
NATIONAL ENGINEERING LABORATORY FOR
BIG DATA ANALYSIS AND APPLICATIONS

目录 CONTENTS

主任致辞 ····· 02

实验室概况 ····· 05

发展历程 ····· 07

实验室组织架构 ····· 09

常务理事单位 ····· 11

常务理事会 ····· 13

技术委员会 ····· 14

实验室主任、副主任 ····· 16

分中心介绍与科研成果 ····· 17

实验室核心科研团队 ····· 18

共性技术研发平台 ····· 42

数据处理及统计分析中心 ····· 42

机器学习中心 ····· 46

数据管理与分析中心 ····· 52

可视分析中心 ····· 55

知识集成和智能决策中心 ····· 61

行业应用示范中心 ····· 68

大数据高性能分析处理中心 ····· 68

智慧交通中心 ····· 72

多媒体大数据分析中心 ····· 74

网信智能中心 ····· 77

大数据分析技术创新中心 ····· 81

国家工程实验室研究基地 ····· 87

北京大学重庆大数据研究院 ····· 88

北京大学长沙计算与数字经济研究院 ····· 110

北京大学人文社科数智化研究基地 ····· 136

主任致辞

为践行国家大数据战略，由北京大学牵头，联合中国科学院数学与系统科学研究院、中山大学、北京奇虎科技有限公司、北京嘀嘀无限科技发展有限公司、中国信息安全研究院等多家单位，共同申请建设大数据分析与应用技术国家工程实验室，并于 2017 年 2 月获得了国家发展和改革委员会的正式批复。

大数据分析与应用技术国家工程实验室的核心定位是扎根中国大地，瞄准国家的重大需求、特别是针对具有中国特色的问题进行发力和突破。同时，实验室还将以企业和市场的需求为导向，以提高产业自主创新能力为目标，力争在我国大数据分析与应用领域形成一批重大的标志性成果。

大数据分析与应用技术国家工程实验室将对科研与人才培养的体制机制进行积极探索和创新，实验室将与北京大学大数据科学研究中心、实验室共建单位等机构深度协作，整合各方资源，发挥各自优势，努力实现机构共建、人才联聘、成果共享的运行新机制。

大数据分析与应用技术国家工程实验室培养当前非常紧缺的大数据人才，致力于将以往分学科、分阶段的“串联式”培养方式升级为多学科交叉、理论与实践并行的“并联式”培养方式，联合北京大学数学科学学院、信息科学技术学院等院系，聚合数学、计算机、工程应用等多方面的资深专家，为大数据领域的优秀拔尖人才营造出良好的成长环境，提高人才培养的效率，提升人才利用大数据相关理论与实践经验解决具体问题的实战能力。

大数据分析与应用技术国家工程实验室的发展离不开主管单位、合作伙伴、诸位专家及广大客户的大力支持。在此，感谢大家对我们工作的无私指导和帮助，欢迎大家对实验室的发展提出宝贵意见和建议！

国家工程实验室主任：张平 院士

新世纪初，大数据作为新事物进入我国时，与信息安全、移动互联网、物联网以及后来的区块链等国内有发展基础的新技术相比，我们在其发展应对上应该说是比较仓促和不充分的，包括数据的分布式存储、处理和分析应用等很多方面的技术储备都非常薄弱，人才短板更是大问题。直到“国家大数据战略”提出后的 2016 年，国内的大数据发展也还是相对稚嫩，产业认知、数据基础、技术和产品储备、政策和法律法规环境等各方面均未形成体系，产业生态仍不完备。

在这样的背景下，2016 年下半年，我们以提升国家大数据发展能力为使命，联合兄弟学校、科研院所和合作企业等，共同申请了大数据分析与应用技术国家工程实验室，并于 2017 年初获批。国家工程实验室的核心定位就是通过构筑一套整合产、学、研多方面能力资源的创新体系，协同开展大数据应用技术创新和工程化落地实践，帮助我国解决好大数据规模化、产业化发展中的关键一环，推动我国大数据的整体技术进步和产业发展。

经过几年沉淀，国家工程实验室已在数据预处理、数据模型 / 算法、交互式可视化数据分析及知识图谱等方面，为社会贡献了很多高质量的大数据技术创新成果，并在大数据统计学习、文本大数据分析、图像视频大数据分析、时空大数据分析等诸多方向开展了卓有成效的工程落地实践。

时至今日，我国在大数据政策和法律法规环境构建、学科建设与人才培养、技术和产品创新，以及数据积累和基础设施建设等各方面都取得了很大进展。在新的发展形势下，国家工程实验室将以深厚的技术创新能力和工程落地实践能力为基础，持续发挥桥梁纽带作用，面向各行业领域典型大数据应用场景的实际需求，加强与业界同行的合作，携手为我国大数据产业发展做出努力。

国家工程实验室主任：郑纬民 院士

实验室概况

大数据分析与应用技术国家工程实验室由国家发展和改革委员会批复建设，由北京大学牵头，联合中国科学院数学与系统科学研究院、中山大学、北京奇虎科技有限公司、北京嘀嘀无限科技发展有限公司、中国信息安全研究院等多家单位共建，在国家发改委高技术司、教育部科技司等主管部门的指导下，于 2017 年 3 月 21 日在北京大学正式揭牌。实验室的主要建设内容是在现有研发和试验条件基础上，建设大数据分析系统开发、可视化展示、测试与评估、重大应用示范与系统集成研发平台。

实验室的主要任务：针对我国大数据分析技术和综合应用能力弱、缺乏面向应用的原创性分析技术等问题，围绕提升大数据综合分析能力与智能决策水平的迫切需求，建设大数据分析与应用技术研究平台，支撑开展大数据预处理与质量控制、大数据统计模型与分析算法、分布式大数据挖掘系统、交互式大数据可视分析、基于大数据与机器学习的新一代智能决策系统等技术的研发和工程化。通过建立大数据通用分析技术研发与应用试验平台，形成国内一流的科研环境，主动承担国家和行业重大科研项目，在大数据统计学习、文本大数据分析、图像视频大数据分析、时空大数据分析、图数据与知识图谱分析等重点研究方向取得一批关键技术成果并成功转化，构建大数据分析应用领域自主知识产权与标准体系，形成可持续的产学研协同创新机制，为推动我国大数据分析与应用技术的进步和产业发展提供技术支撑。

实验室紧密围绕大数据分析与应用技术发展的需要，开展相关产业关键技术攻关、重要技术标准制订，集聚与培养产业急需的技术创新人才。探索和实施适合具体行业特点的创新方案，着力解决行业发展的重大技术问题，提高产业自主创新能力，建立该领域发展趋势与重大问题的研究机制，积极完成国家有关部门委托的科研课题和行业技术咨询工作，积极提供行业发展研究情况和重要进展，更好地为国家和行业发展服务。

实验室着眼核心定位和远期目标，在承担国家重大任务、大数据科学研究、技术转移与辐射、人才培养等方面取得了显著成绩。

2020 年 12 月 5 日，大数据分析与应用技术国家工程实验室以优秀的成绩通过了国家发改委委托教育部组织的验收。

▲ 验收会议与会专家与实验室部分成员合影

发展历程

2017 年 2 月

国家发改委批复由北京大学作为牵头单位，联合相关单位筹建大数据分析与应用技术国家工程实验室。

2017 年 3 月

大数据分析与应用技术国家工程实验室揭牌仪式暨大数据分析与应用技术创新论坛。

2018 年 5 月

成立大数据分析与应用技术国家工程实验室第一届理事会，并召开第一届理事会第一次会议。

2020 年 10 月

大数据分析与应用技术国家工程实验室联合相关单位共同成立数字生态协同创新平台，首次发布数字生态指数。

2020 年 11 月

北京大学批复大数据分析与应用技术国家工程实验室在北京大学校内按照实体研究机构运行。

2020 年 11 月

北京大学与重庆高新区管理委员会签订合作协议，共建北京大学重庆大数据研究院。大数据分析与应用技术国家工程实验室作为校内承建二级单位负责落地实施。

2020 年 12 月

大数据分析与应用技术国家工程实验室以优秀的成绩通过国家发改委委托教育部组织的验收。

2022 年 1 月

北京大学、长沙市人民政府和长沙高新区管理委员会签订合作协议，共建北京大学长沙计算与数字经济研究院。大数据分析与应用技术国家工程实验室作为校内承建二级单位负责落地实施。

2022 年 4 月

牵头筹建北京大学人文社科数智化研究基地。

2023 年 7 月

大数据分析与应用技术国家工程实验室组建新一届技术委员会，并召开 2023 年会议。

2024 年 8 月

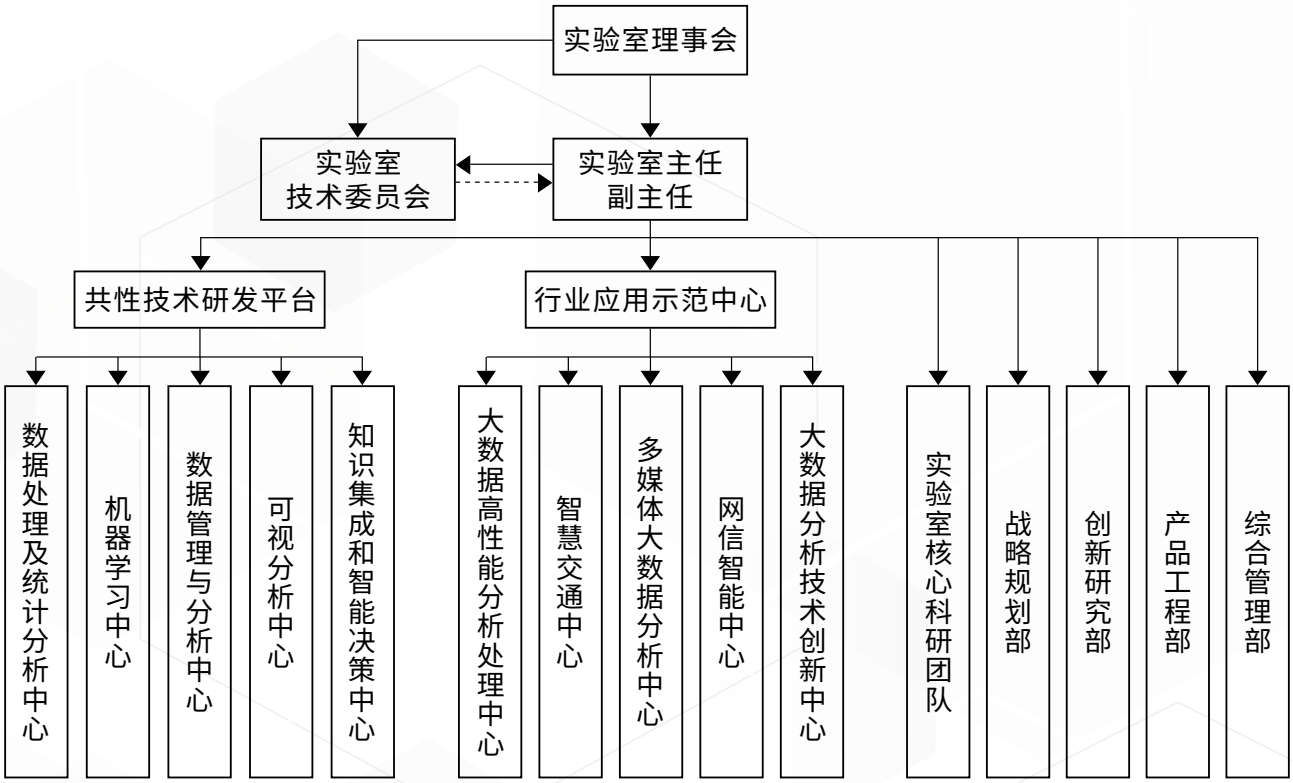
依托北京大学重庆大数据研究院孵化的企业北太振寰（重庆）科技有限公司，顺利完成首轮千万级融资，以破亿估值圆满收官。

2025 年 3 月

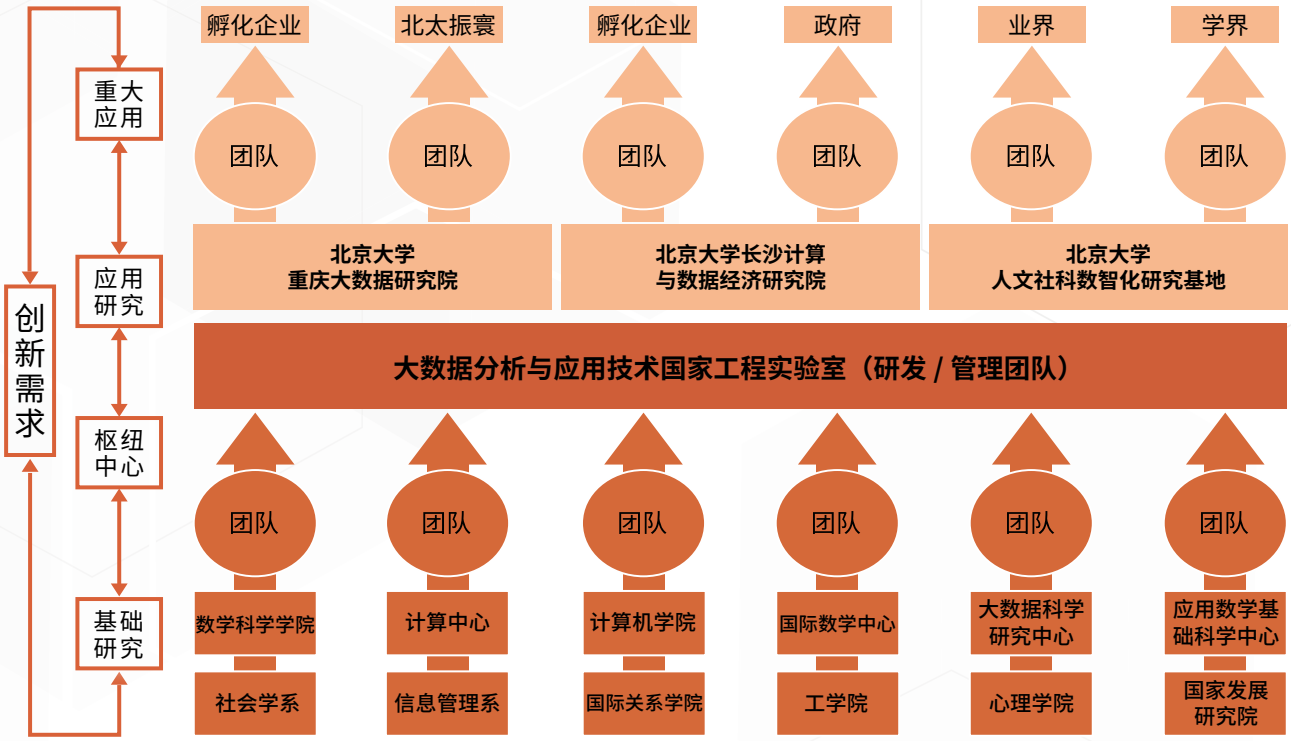
截至 2025 年 3 月，依托地方研究院孵化公司 16 家，其中 5 家获得第一笔融资。

实验室
组织架构

实验室制订了理事会章程，聘任了实验室主任团队，成立了技术委员会，建立了健全的实验室组织架构和管理制度。截至 2025 年，实验室组建了 10 个分中心，并设立四个部门，战略规划部、创新研究部、产品工程部和综合管理部。



◆大数据学术创新链



常务理事单位

◆北京大学

北京大学是我国著名的研究型、综合型大学，数学、统计、信息科学等多个学科方向参与国家工程实验室建设。北京大学在大数据分析与应用方面具有综合学科优势与坚实研究基础。北京大学拥有齐全的学科，大数据学科体系完整，成功申报了全国首批数据科学与大数据技术专业，拥有数据科学的本科、硕士与博士培养全过程，并在交叉学科平台建设上成立了大数据科学研究中心。北京大学拥有“机器感知与智能教育部重点实验室”、“数字视频编解码技术国家工程实验室”等相关研究创新平台，还与多个企业进行了广泛合作，建立了长效的产学研合作关系，为相关部门提供大数据分析技术支撑和服务，并积累了大量的数据资源和经验。北京大学积极探索建立大数据学术创新链，以服务国家战略需求为导向，通过聚合与优化资源，将校内外研究机构与合作单位作为关键节点，协同创新，加强创新链和产业链的对接，形成完整科研创新闭环，以实现取得标志性应用成果、服务社会经济发展和国家战略需求的目标。

◆中国科学院数学与系统科学研究院

中国科学院数学与系统科学研究院成立于 1998 年 12 月，是由中国科学院数学研究所、应用数学研究所、系统科学研究所及计算数学与科学工程计算研究所 4 个研究所整合而成，是一个综合性的国立学术研究机构，覆盖了数学与系统科学的主要研究方向。

中国科学院数学与系统科学研究院历年来获得科技重要奖励 500 余项，包括首届国家最高科学技术奖、5 项国家自然科学奖一等奖、1 项国家科学技术进步奖特等奖、1 项有“东方诺贝尔奖”之称的邵逸夫数学科学奖、27（或者 29）项国家自然科学奖二等奖以及 50 余项国际学术奖励和荣誉。不少重要的应用成果受到中央领导和实际部门的充分肯定，显示了数学和系统科学理论在国民经济主战场的作用。研究院科研队伍实力雄厚，现有职工 338 余人，其中科研人员 248 人，包括中国科学院院士 16 人，中国工程院院士 1 人，第三世界科学院院士 6 人，中国科学院“百人计划”入选者和国家杰出青年基金获得者约 80 人。中国科学院数学与系统科学研究院历来重视人才培养，有 15 个硕士点、12 个博士点（二级学科），至今培养硕士与博士研究生 3000 多人（其中，硕士研究生 1588 名，博士研究生 1876 名），博士后 232 名。研究院下属的 4 个研究所是我国最早被批准具有硕士、博士学位授予权的单位之一，是首批国家批准的博士后流动站之一。

◆北京奇虎科技有限公司

北京奇虎科技有限公司（以下简称 360 公司）经过不断努力和持续创新，已成为综合实力较强的网络安全企业。360 公司终端安全软件在国内拥有 6 亿 PC 用户、9 亿智能手机用户，在国外用户数也已经超过 3 亿。360 公司成立以来，以保护国家网络安全为使命，以创新为发展原动力，在多个领域取得突破性成果，自主研发的核心技术达到世界领先水平。360 公司总部设立在北京，拥有办公场所面积近 7 万平米，同时公司在上海、广州、深圳、雄安新区、天津、成都、西安、武汉、珠海、大连等地设有分支机构，为科技研发人员提供了良好的研发条件以及业务运行环境。360 公司员工约 1 万人，其中超过 50% 为产品及技术研发人员。

◆ 中山大学

中山大学学科门类覆盖面广，是教育部直属高校中学科门类最齐全的学校之一，在大数据分析方面具有坚实的研究基础。中山大学计算机学院成立于 2015 年 6 月，由学校整合原信息科学与技术学院计算机科学系、软件学院、超级计算学院以及数学与计算科学学院计算数学方向（部分人员）等计算机相关专业的优势资源组建而成。学院拥有强大的计算平台“天河二号”，以国家超级计算广州中心（以下简称广州超算中心）“天河二号”超级计算机为平台，筹建了“大数据科学研究中心”。由此，学院形成“一院两中心”的发展格局，与广州超算中心和大数据科学研究中心实现一体化运行，互相支撑、共同发展。同时，学院围绕超级计算的行业应用和科研方向组建了一支高水平的技术团队，支撑广州超算中心的应用推广和服务拓展。学院坚持面向学术前沿、坚持面向国家重大战略需求、坚持面向国家和区域经济社会发展，贯彻“理工结合，学科交叉”的发展战略，按照学校“大平台、大团队、大项目”的部署，推动学术和科研的发展；围绕“德才兼备、领袖气质、家国情怀”的人才培养目标，致力于培养具有国际视野的应用型、复合型、创新型的高素质计算机人才和计算与数据分析处理人才。

◆ 北京嘀嘀无限科技发展有限公司

北京嘀嘀无限科技发展有限公司于 2013 年成立，是全球领先的一站式移动出行平台，为 5.5 亿用户提供出租车、快车、专车、豪华车、公交、小巴、代驾、企业级、共享单车、共享电单车、共享汽车等全面的出行和运输服务，为超过 3000 万车主及司机获得灵活赚取收入的机会。公司运用大数据驱动的人工智能技术，以共享经济实践响应中国互联网创新战略，与不同社群及行业伙伴协作互补，通过技术创新解决全球交通、环保和就业挑战。公司一直非常重视研发投入，拥有一支全球顶尖的研发队伍，多个研究方向或产品曾获国家科技计划项目支持，技术成果多次获评国际领先水平，拥有多项核心技术和专利群。

◆ 中国信息安全研究院

中国信息安全研究院是中国电子信息产业集团的全资二级单位，是中国电子加快打造国家网信产业核心力量和组织平台的产业总体支撑单位，致力于成为我国网信领域具有战略支撑作用的一流智库，主要从事战略发展、产业研究、数字科技、网信测评、网络安全、高新电子等业务，拥有武器装备科研生产单位保密资格认证（一级），参与支撑工信部、网信办、密码局等国家相关部委工作。中国信安位于北京未来科学城，是国家双创示范基地，建有大数据分析与应用技术国家工程实验室网信智能中心和国家保密科技测评中心、电子信息产业系统测评实验室等多个科研创新平台，是中国电子网络安全和信息化产业基地业主单位，建筑规模近 30 万平米，已孵化培育中国电子智能制造和工业互联网等新兴产业。

常务理事会

王　杰

国家工程实验室理事长

北京大学数学科学学院教授

北京大学原副校长

朱立锋

国家工程实验室常务理事

中国电子信息产业集团总经理助理数字办常务副主任

中国信息安全研究院董事长

中电工业互联网有限公司董事长

教授级高工

谢　冰

国家工程实验室常务理事

教授，北京大学科学研究部部长

胡振泉

国家工程实验室常务理事

奇虎 360 副总裁

杨　毅

国家工程实验室常务理事

滴滴出行智能平台事业群总经理

何伟玲

国家工程实验室常务理事

中山大学科学研究院基地处副处长（主持工作）

广东省医学会精准医学与分子诊断学分学会副主任委员

戴彧虹

国家工程实验室常务理事

中国科学院数学与系统科学研究院副院长

IFORS 会士

亚太运筹学会联合会主席

技术委员会（按姓氏拼音排序）

袁亚湘

国家工程实验室技术委员会主任
中国科协副主席
中国科学院院士
发展中国家科学院院士
巴西科学院通讯院士
国际工业与应用数学联合会主席

陈松蹊

中国科学院院士
北京大学讲席教授
北京大学统计科学中心联席创始主任

鄂维南

中国科学院院士
北京大学数学科学学院讲席教授
国家工程实验室联席主任

高　文

中国工程院院士
计算机专家
北京大学教授
鹏城实验室主任

管晓宏

中国科学院院士
系统工程学家
西安交通大学教授

廖湘科

中国工程院院士
计算机系统软件专家
国防科技大学教授

林毅夫

北京大学博雅讲席教授
北京大学国家发展研究院名誉院长

罗智泉

大数据分析与应用科学家
加拿大皇家科学院院士
中国工程院外籍院士
香港中文大学（深圳）副校长、校长讲座教授
深圳市大数据研究院院长

马志明

中国科学院院士
第三世界科学院院士
中国科学院应用数学研究所研究员
南开大学统计与数据科学学院院长

钱德沛

中国科学院院士
北京航空航天大学教授
中国计算机学会会士
曾任国家工程实验室副主任

汤　涛

中国科学院院士
欧洲科学院院士
发展中国家科学院院士
计算数学家
中国数学会副理事长
广州南方学院校长

唐立新

中国工程院院士
东北大学副校长
工业智能与系统优化国家级前沿科学中心主任及首席科学家
智能工业数据解析与优化教育部重点实验室主任

徐文伟

华为科学家咨询委员会主任

徐宗本

中国科学院院士
数学家
信号与信息处理专家
西安交通大学教授

张平文

中国科学院院士
武汉大学校长、党委副书记
北京大学博雅讲席教授
国家工程实验室主任

郑志明

中国科学院院士
信息处理专家
北京航空航天大学教授

实验室主任、副主任

张平文

国家工程实验室主任
中国科学院院士
武汉大学校长、党委副书记
北京大学博雅讲席教授

鄂维南

国家工程实验室联席主任
中国科学院院士
北京大学数学科学学院讲席教授

袁晓如

国家工程实验室常务副主任
北京大学智能学院研究员

郑伟诗

国家工程实验室副主任
中山大学计算机学院副院长

殷宇辉

国家工程实验室副主任
奇虎 360 副总裁

傅毅明

国家工程实验室副主任
中国信息安全研究院院长助理
正高级工程师

吴国斌

国家工程实验室副主任
滴滴出行科技生态与发展部总监

刘　歆

国家工程实验室副主任
中国科学院数学与系统科学研究院计算数学与科学工程计算研究所副所长

夏壁灿

国家工程实验室（北京大学校内）副主任
北京大学数学科学学院教授

杨　超

国家工程实验室（北京大学校内）副主任
北京大学数学科学学院教授

分中心介绍与科研成果

实验室核心科研团队

◆ 团队简介

大数据分析与应用技术国家工程实验室组建优秀的科研团队，承接实验室国家级重要科研项目，服务国家和社会的重大需求。

◆ 团队成员介绍

张平文

中国科学院院士、发展中国家科学院院士。历任北京大学校长助理，副校长，党委常委、副校长等职务，现任武汉大学校长、党委副书记；北京大学博雅讲席教授，大数据分析与应用技术国家工程实验室主任，北京大学重庆大数据研究院首席科学家，北京大学长沙计算与数字经济研究院首席科学家；中国工业与应用数学学会理事长；国家自然科学基金委“科学计算与机器学习”基础科学中心项目负责人。主要从事复杂流体的数学理论和计算方法，大数据分析与应用等方面研究，发表论文百余篇。曾获国家自然科学奖二等奖、教育部高等学校自然科学奖一等奖、何梁何利基金科学与技术进步奖、冯康科学计算奖等奖项。2011 年受邀在第七届国际工业与应用数学大会上作一小时报告，2014 年受邀在美国工业与应用数学年会上作大会报告，2018 年受邀在国际数学家大会作 45 分钟邀请报告，2020 年当选美国工业与应用数学学会会士。担任学术期刊 CSIAM Transactions on Applied Mathematics 主编，Applied Mathematics and Mechanics 等十余种国内外期刊编委。

鄂维南

中国科学院院士，美国数学学会会士，美国工业与应用数学学会会士，英国物理学会会士。北京大学讲席教授、大数据分析与应用技术国家工程实验室联席主任、北京大学国际机器学习研究中心主任、北京科学智能研究院院长、北京大数据研究院院长。1982 年毕业于中国科学技术大学数学系，1985 年获中国科学院计算中心硕士学位，1989 年获美国加州大学洛杉矶分校博士学位。主要从事计算数学、应用数学及其在力学、物理、化学和工程等领域中的应用等方面的研究。曾获首届美国青年科学家与工程师总统奖，冯康科学计算奖，国际工业与应用数学协会颁发的 Collatz 奖，ICIAM 麦克斯韦奖，美国工业与应用数学学会 R. E. Kleinman 奖，美国工业与应用数学学会卡门 (Theodore von Karman) 奖，SIAM 和 ETH Zürich 联合授予的 Peter Henrici 奖及 ACM 戈登·贝尔奖（ACM Gordon Bell Prize），第五届北京市华人华侨“京华奖”。2022 年国际数学家大会 1 小时报告人。

陈松蹊

中国科学院院士，北京大学长沙计算与数字经济研究院首席科学家。现任中国概率统计学会理事长、中国统计学会副会长、美国科学促进会会士、美国统计学会会士、数理统计研究所会士。曾任北京大学光华管理学院商务统计与经济计量系主任、北京大学统计科学中心联席创始主任。曾任统计学顶级期刊《The Annals of Statistics》和美国统计学会会刊《JASA》副主编，目前任《Environmentrics》和《中国科学 (数学)》副主编。目前已发表论文 120 余篇，Web of Science H-index 33。曾获 Iowa State University 教员杰出研究奖、第七届厉以宁科研奖、2017 年教育部自然科学一等奖。研究成果得到了国内外同行的高度认可，入选斯坦福大学 2020-2022 榜单。

钱德沛

中国科学院院士，北京航空航天大学教授，中国计算机学会会士。2018 年至 2021 年期间，担任大数据分析与应用技术国家工程实验室副主任。长期从事高性能计算机体系结构与实现技术、网络计算、分布并行系统和计算机网络等方面的研究，曾主持过国家重大项目、国家自然科学基金重点项目等研发工作，取得了重要的科技成果。发表学术论文 300 余篇，获得国家科技进步奖二等奖两项、北京市科学技术一等奖和二等奖各一项、山东省科学技术一等奖一项、国家教委科技进步二等奖三项。

袁晓如

北京大学智能学院研究员、长聘副教授、博士生导师，北京大学教材建设博雅特聘教授，大数据分析与应用技术国家工程实验室常务副主任，机器感知与智能教育部重点实验室副主任。研究方向包括复杂流场数据可视化，高维 / 时空数据，交通、社会媒体数据的分析，可视化的快速构建方法。高动态范围可视化的工作获得 2005 年 IEEE VIS 大会最佳应用论文奖。2013 年来指导实验室团队 7 次在 IEEE VAST 可视化分析挑战赛中获奖。数十次担任 IEEE VIS，EuroVis, IEEE PacificVis 等国际可视化会议程序委员会委员，2017 年 IEEE VIS 大会论文共同主席 (SciVis)，创建中国可视化与可视分析 (ChinaVis) 大会。中国计算机学会理事，杰出会员，杰出讲者。中国计算机学会大数据专家委员会委员，人机交互专委会常务委员和计算机辅助设计与图形学专委会委员。中国图象图形学学会理事、可视化与可视分析专业委员会主任。

夏壁灿

北京大学数学科学学院教授，大数据分析与应用技术国家工程实验室（校内机构）副主任。主要研究方向计算机数学、自动推理和程序验证。在实代数、实几何自动推理方面取得了若干研究成果，并编制了能自动发现不等式型定理的软件 DISCOVERER。

杨　超

北京大学大数据分析与应用技术国家工程实验室（校内机构）副主任，北京大学长沙计算与数字经济研究院院长，北京大学数学科学学院长聘教授，教育部长江学者特聘教授。长期从事与超大规模并行计算相关的模型、算法、软件和应用研究，主持国家自然科学基金重点项目、重大研究计划集成项目、科技创新 2030 重大项目等国家级科研项目 10 余项，发表学术论文 100 余篇，研究成果先后获 2016 年国际计算机协会“戈登·贝尔”奖、2017 年中国科学院杰出科技成就奖、2017 年 CCF-IEEE CS 青年科学家奖、2018 年茅以升北京青年科技奖、2020 年首届王选杰出青年学者奖等国内外重要奖项。目前担任 National Science Review 编委、SIAM Journal on Scientific Computing 编委，中国工业与应用数学学会高性能计算与数学软件专业委员会副主任，中国新一代人工智能产业技术创新战略联盟 AI 指令集与开发接口标准专题组组长，IEEE C/DC 2941 标准工作组副主席等职。

胡　俊

北京大学数学科学学院党委书记、教授，北京大学重庆大数据研究院院长。兼任 Adv Appl. Math. Mech. 执行主编、多个期刊的编委、北京计算数学学会理事长、中国数学会常务理事、中国大坝工程学会大坝数值模拟专委会副主任委员、重庆市工业软件应用发展协会副会长、北大 - 华为数学联合实验室主任。主要从事非标准有限元方法的研究，特别是弹性力学问题、线性化 Einstein-Bianchi 方程组及相关问题的非标准有限元方法的构造与数值分析的研究，解决了弹性力学问题混合有限元方法的构造这个长期悬而未决的公开问题，首次构造了线性化 Einstein-Bianchi 方程组保结构的稳定有限元方法。曾获教育部自然科学奖一等奖、国家杰出青年科学基金、冯康科学计算奖、中国数学会计算数学分会首届青年创新奖。

邱泽奇

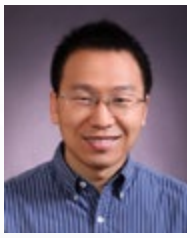
北京大学中国社会与发展研究中心主任，社会学系教授、博士生导师，数字治理研究中心主任，教育部“长江学者奖励计划”特聘教授，北京大学博雅特聘教授。曾先后担任北京大学中国社会科学调查中心创办主任（2006-2011）、重庆大学公共管理学院院长（2012-2016）。2007 年入选教育部新世纪优秀人才支持计划，享受国务院政府特殊津贴。主要研究领域为数字社会与治理、技术应用与社会变迁、组织社会学、社会调查与研究方法等。主持和参与国家社会科学基金、国家自然科学基金、教育部人文社会科学重点研究基地课题等国家级项目十余项，在国内外期刊上发表学术论文百余篇。

段慧玲

中国科学院院士，国家杰出青年基金获得者，长江学者特聘教授，国家万人计划科技创新领军人才，美国机械工程师学会会士（ASME Fellow）。现任北京大学博雅讲席教授、工学院（本科生学院）院长，中国力学学会常务理事，中国力学学会女科技工作者专业委员会主任，中国复合材料学会常务理事，国际理论与应用力学联盟 IUTAM Symposia Panel for Solid Mechanics 委员，固体力学顶尖期刊 JMPS 编委。主要从事界面力学，流固耦合力学研究，在国际主流期刊发表论文 230 余篇，主持基金委重大项目、装备重大基础研究项目等，成果入选十三五国家自然科学基金资助项目优秀成果选编。作为第一完成人获国家自然科学二等奖，教育部自然科学一等奖和北京市教学成果一等奖；并获全国创新争先奖、中国青年科技奖、中国青年女科学家奖、美国机械工程师学会 The Sia Nemat-Nasser Early Career 奖、全国三八红旗手、宝钢教育基金会优秀教师特等奖、德国洪堡研究奖、北京市优秀研究生指导教师、第十届“徐芝纶力学奖”等多项奖励。2023 年，段慧玲教授新当选中国科学院院士（数理学部）。

李　若

北京大学博雅特聘教授、北京大学数学科学学院副院长、北京大学重庆大数据研究院副院长、北京大学大数据分析与应用技术国家工程实验室核心科研成员。获聘教育部长江特聘教授，曾任中国数学会计算数学分会理事、副理事长，北京计算数学学会理事长、监事长；曾获教育部高校自然科学奖一等奖（第三完成人）、科技部中青年科技创新领军人才、冯康科学计算奖、第九届国际工业与应用数学大会报告人、吴文俊应用数学奖、CSIAM 会士等荣誉；曾主持和参与国家杰出青年基金、科技部 973 项目等。主要研究方向为偏微分方程数值解和计算流体力学，发表科研论文 100 余篇，学术成就主要在动理学方程的模型约化理论、约化模型的数值方法方面取得突破性进展，解决了存在六十多年的 Grad 矩方程组的双曲性缺失问题。



李铁军

北京大学数学科学学院教授，北京大学数学科学学院信息与计算科学系主任，北京大学大数据科学研究中心副主任。国家自然科学基金委优秀青年及杰出青年基金获得者，入选教育部“新世纪优秀人才”支持计划。研究领域为随机模型及算法、数据科学的理论及应用。在化学反应随机动力学稀有事件、能量景观、tau-leaping 方法的分析；复杂网络的模型约化；复杂流体多尺度理论及数值分析方面做出了重要成果。

卢　朓

北京大学数学科学学院副教授、博士生导师。研究方向包括电磁场和电磁波、量子力学和半导体载流子输运等相关的建模、算法和模拟。2003 年获中国计算数学学会青年优秀论文一等奖，2017 年获得北京大学教学优秀奖（研究生部分），2018 年获得北京大学宝洁奖教金。迄今累计发表杂志论文三十余篇。

陈　钟

北京大学信息科学技术学院教授、区块链研究中心主任，金融信息化研究中心主任，网络与信息安全实验室主任，北京大学软件与微电子学院创始院长。主要研究领域为面向领域的软件工程、网络与信息安全、区块链技术等，曾主持和参加完成国家科技重大专项、国家 863、国家自然科学基金以及国家发改委、国家保密局、人民银行、工信部、公安部、北京市等部委级科技专项等二十余项，曾获国家科技进步二等奖两项、国家级教学成果一等奖一项，部委级科技成果奖励多项。陈钟教授为中国计算机学会会士、常务理事，担任中国软件行业协会副理事长，中国开源软件推进联盟副理事长，全国区块链与分布式记账技术标准化技术委员会委员，全国信息安全、信息保密标准化技术委员会委员，科技部十四五规划“区块链”专家组成员等。

宋　洁

北京大学先进制造与机器人学院党委委员、书记，教育部长江学者特聘教授，北京大学长沙计算与数字经济研究院副院长。主要研究领域为在线学习、仿真优化及在系统工程领域的应用，发表学术论文 60 余篇。承担国家自然科学基金创新群体、重大、重点等多项课题，主持国家发改委、国家粮食局、国家电网等单位重点研究课题。国际电气和电子工程师协会高级会员（IEEE Senior Member），担任 IEEE 国际自动化管理技术委员会主席（首任亚洲学者），同时入选 IEEE 杰出报告人。曾先后入选教育部青年长江学者，长江学者特聘教授。2022 年获系统科学与系统工程青年科技奖。研究获得日内瓦国际发明金奖，IEEE 协会期刊最佳论文奖，IISE 国际工业与系统工程协会最佳论文荣誉奖，INFORMS 最佳论文奖，工程管理教指委“最佳指导导师奖”，北京市高等教育教学成果一等奖，多次获得北京大学教学优秀奖和教学改革优秀奖。

董　彬

北京大学北京国际数学研究中心长聘教授，北京大学国际机器学习研究中心副主任，北京大学国家生物医学成像科学中心研究员，北京大学长沙计算与数字经济研究院副院长。2003 年本科毕业于北京大学数学科学学院，2005 年在新加坡国立大学数学系获得硕士学位，2009 年在美国加州大学洛杉矶分校数学系获得博士学位。主要研究领域为科学计算、机器学习及其在计算成像和数据分析中的应用。在国际重要学术期刊和会议上发表论文 90 余篇，现任期刊 Inverse Problems and Imaging 编委，CSIAM Transactions on Applied Mathematics、Journal of Computational Mathematics、Journal of Machine Learning 副主编。2014 年获得求是杰出青年学者奖，2015 年入选中组部“千人计划”青年项目，2020 年入选中组部“万人计划”领军人才，2022 年受邀在世界数学家大会（ICM）作 45 分钟报告学家大会 (ICM) 作 45 分钟报告，2023 年入选新基石研究员项目同年获得王选杰出青年学者奖。

张志华

北京大学数学科学学院统计学教授。主要从事机器学习、应用统计和数值计算的交叉学科的研究。研究范围涉及理论、模型、算法和应用等多个层面，研究方向主要包括机器学习与模式识别、统计建模与计算、数值代数与优化、分布式计算构架等。

刘云淮

北京大学博雅特聘教授、博导，ACM 中国理事会副主席，北京大数据研究院副院长，国家杰出青年，中组部“万人计划”青年拔尖人才，国家重点研发计划项目负责人。刘云淮博士 2008 年毕业于香港科技大学计算机科学和工程专业。他的研究方向主要在智慧城市、移动计算、物联网、群智感知等方向。发表论文 100 余篇，两次获得 IEEE ICDCS, IEEE SANER 国际会议最佳论文。

黄　卓

北京大学国家发展研究院副院长、教授，BiMBA 商学院院长，北京大学南南合作与发展学院副校长，北京大学数字金融研究中心常务副主任，北京大学长沙计算与数字经济研究院副院长。斯坦福大学经济学博士，曾获得斯坦福大学经济系“最佳博士生候选人论文奖”。主要研究领域为数字经济与数字金融、金融计量、金融工程、大数据分析。主持多项国家自然科学基金和国家高端智库课题，研究成果发表在国内外一流期刊如 Journal of Econometrics 和《经济研究》《管理世界》等。2014 年获得应用计量经济学领域国际权威期刊 Journal of Applied Econometrics 的“Richard Stone 最佳论文奖”，2015 年获得第七届高等学校科学研究优秀成果奖 (人文社会科学) 论文类二等奖，2020 年获得北京大学教学卓越奖。目前还担任中国金融四十人论坛特邀研究员、中国衍生品青年论坛联席秘书长、数字经济开放研究计划首任秘书长。

文再文

北京大学北京国际数学研究中心长聘教授，北京大学工学院工业工程与管理系主任。主要研究最优化算法与理论及其在机器学习和人工智能中的应用，先后获得国家自然科学基金委优秀青年科学基金、中组部青年拔尖人才计划、中国青年科技奖、国家万人计划科技创新领军人才和北京市杰出青年中关村奖。主持国家自然科学基金重点项目等多个课题，发表学术论文 70 余篇，与合作者一起开发有多重网格优化软件包 MGLS，稀疏优化软件包 FPCAS，低秩矩阵恢复软件包 LMaFit，保正交约束软件包 ARNT，半定规划软件包 SSNSDP 和大规模特征值分解软件包 Arrabit。2021 年由高教出版社出版教材《最优化 : 建模、算法与理论》和《最优化计算方法》，得到了北京大学、清华大学、复旦大学和中国科技大学等超百所高校采用，截至 2022 年 4 月累计印刷 4 次 1.7 万余本。

徐克付

国家工程实验室特聘研究员。从事网络空间安全领域的技术研究与工程开发十余年，作为项目负责人完成了国家自然科学基金、国家 863 重点项目、国家发改委重大专项、中国科学院战略性先导专项、华为研究院委托研发项目等十余项项目的研究工作；在国内外学术期刊及国际会议上发表论文 50 余篇，申请发明专利十余项，长期参与国家网络安全保障工作，多次获中央部委表彰，科研成果《全周期云数据安全管控体系及应用支撑平台》获教育部科技进步一等奖。

研究方向：网络空间安全、数据安全与数据流通

王剑晓

北京大学博雅青年学者，国家工程实验室副研究员，国家自然科学基金优秀青年项目获得者，国家重点研发计划青年科学家；美国斯坦福大学访问学者，清华大学工学博士、工学和经济学学士。从事数据驱动优化决策理论方法及其在低碳能源系统韧性、储能与氢能技术领域的应用研究。作为第一和通讯作者，在 Nature Sustainability、Nature Communications、Engineering、IEEE Transactions 等期刊和会议发表论文 100 余篇；第一作者编写 Springer 英文专著 1 部，参编 IEEE Press 英文专著 1 部、中文专著 4 部；曾获日内瓦国际发明金奖、中国电工技术学会科技进步二等奖，入选国家工信部优秀研究成果，得到中国科技部及工信部致谢，获《人民日报》报道。担任 Cell 旗下期刊 The Innovation 青年编委、IEEE Transactions on Industry Applications 编委、IEEE PES 中国常务理事、IEEE IAS 北京大学指导教师等，入选科技部第六次国家技术预测专家、教育部全国研究生教育评估监测专家、第九批北京市优秀青年人才、福布斯中国 U30 精英榜、全国仿真创新应用优秀指导教师、第七届中国科协青年人才托举工程、北京市“千人进千企”专项行动产业特派员、吴文俊人工智能优秀青年、中国科协高端科技创新智库青年等。

张　一

国家工程实验室特聘副研究员，香港大学博士，北京大学学士。发表 21 篇学术论文，出版一本学术著作，以主持或主要负责人身份参与 29 项科研基金和项目。曾（兼）任北京大数据研究院经济大数据研究中心常务副主任，北京大学重庆大数据研究院数字化转型中心副主任，新华社媒体融合生产技术与系统国家重点实验室青年科学家兼首席科学家助理，香港大学太古学者协会会长，香港大学研究生会会长。

研究方向：大数据分析与应用、数字经济、数字化转型

王　娟

国家工程实验室特聘副研究员。管理学博士，北京大学博士后，美国印第安纳大学凯利商学院访问学者。相关论文发表在 Asia Pacific Journal of Management, Chinese Management Studies 等 SSCI 期刊，以及《电子政务》、《西安交通大学学报（社科版）》、《管理评论》、《社会科学家》等 CSSCI 期刊。

研究方向：数字经济、数字化转型、国际商务管理

崇　滨

国家工程实验室特聘副研究员。北京大学理学博士，清华大学博士后。在机器学习，数字经济等领域发表十余篇 SCI/SSCI 等核心论文。曾在匹兹堡大学，MIT 斯隆管理学院，台湾大学，哈佛大学交流访问。主持和参与 973 计划，国家自然科学基金重大、重点和面上项目，省部级课题。先后参与企业数字化转型、全球数字人才和数字技能发展、数据要素化治理等报告和书稿撰写，获相关部门采纳和领导批示。

研究方向：数据要素化、大数据专利分析、数字化转型

王冉冉

国家工程实验室助理研究员。北京大学企业大数据研究中心研究员，北京大学重庆大数据研究院产业金融研究中心执行主任，统计学博士。曾担任北京大学中国企业创新创业调查执行负责人；参与中国小微经营者调查，并负责编制小微经营者信心指数；在 China Economic Review，China Economic Journal，产业经济评论，IEEE Access，International Conference for Machine Learning and Cybernetics 等经济学和机器学习相关期刊上发表多篇论文。

研究方向：金融风险管理、小微企业研究、机器学习算法

张　敏

国家工程实验室助理研究员、计算数学博士。主持国家自然科学青年基金，参与研发国产通用型科学计算软件北太天元，主要负责研发工具箱、多域动态系统仿真工具等。获第十七届钟家庆数学奖、入选北京市科协 2024-2026 年度青年人才托举工程。曾为北京大学博雅博士后、美国堪萨斯大学数学系访问学者，曾主持博士后科学基金面上资助，多次参与国防基础科研核科学挑战专题项目等。在计算数学高水平期刊 SIAM J. Sci. Comput., J. Comput. Phys., J. Sci. Comput. 等上发表论文 10 余篇。

研究方向 : 偏微分方程数值解、科学计算软件

刘浩洋

国家工程实验室助理研究员，计算数学博士。曾为北京大学博雅博士后；参与北太天元数值计算通用软件研发，担任软件内核架构师，产品研发中心技术负责人，并获得多部国内专利授权；获得 CSIAM 2023 青年人才托举项目资助；在 SIAM Journal on Scientific Computing, SIAM Journal on Optimization 等国际知期刊发表多篇学术论文；参与编写《最优化 : 建模、算法与理论》、《最优化计算方法》学术教材。

研究方向：最优化算法与理论、科学计算软件

侯　琛

国家工程实验室助理研究员。2019 年 1 月毕业于清华大学，获工学博士学位，2019 年 2 月至 2021 年 3 月在腾讯公司任高级工程师，2021 年 4 月至 2023 年 9 月在中国农业大学任副教授，以第一作者在 IEEE Trans. ASE, SMC-S, SC 和 VLSI 等期刊上发表论文 20 余篇，以第一发明人获得国家发明专利授权 60 余项，香港同步授权 30 余项，同步申请美国专利 4 项，主持国家自然科学基金青年科学基金 1 项，获得 2023 年度 IEEE Transactions on Automation Science and Engineering 最佳论文奖（第一作者）和 2019 年清华大学优秀博士学位论文奖。

研究方向：复杂网络化系统与智能数据系统的运行、控制与优化

金鹏展

国家工程实验室助理研究员。2016 年本科毕业于中国科学技术大学数学系，2021 年博士毕业于中国科学院数学与系统科学研究院，2021 年至 2024 年为北京大学数学科学学院博雅博士后。曾两度赴美国布朗大学应用数学系从事合作研究。曾获中国博士后科学基金面上一等资助。在 SIAM 系列，Nature Mach. Intell.，IEEE Trans. Neural Networks，Neural Networks 等期刊发表论文 10 余篇。

研究方向：机器学习与科学计算

张　博

国家工程实验室特聘副研究员、工学博士，曾为北京大学博士后。在 Comput.-Aided Civil Infrastruct. Eng., Swarm Evol. Comput., IEEE Trans. Intell. Transp. Syst., J. Comput. Phys. 等期刊会议上发表学术论文 10 余篇，已获授权多项发明专利。目前或曾经主持科研项目包括：国家自然科学基金青年项目、博士后科学基金面上项目、北大 - 华为数学联合实验室项目。

研究方向：科学计算、机器学习

陈德良

北京大学长沙计算与数字经济研究院副院长，中国国际技术转移经理人。长期从事环保、能源、工业等领域的大数据平台建设与顶层设计，作为核心研究人员参与科技部、国家发改委等多项科研课题，在各类期刊上发表学术论文数十余篇；参与建设多个科技创新平台和推动设立多个校企联合研发平台。

王　涛

北京大学重庆大数据研究院副院长，中关村数智人工智能产业联盟副秘书长（兼），重庆数字经济研究中心特聘专家，高级技术经纪人。曾任北京大学大数据分析与应用技术国家工程实验室郑州中心副主任、北京大数据研究院博雅大数据学院执行院长等。

黄文彬

北京大学信息管理系主任助理、长聘副教授、博导、研究员，北京大学信息化与信息管理研究中心主任、北京大学信息管理系大数据管理与应用教研室主任，北京大学重庆大数据研究院副院长。曾获教育部国家级教学成果奖二等奖、北京市高等教育教学成果奖二等奖、黄廷芳 / 信和青年杰出学者奖、北京大学教学优秀奖等荣誉。致力于基于数据科学、多媒体技术、人工智能等技术面向人文社会科学 (教育、情报、出版、金融、科学学) 的应用研究。主持、参与国家社会科学重点项目、面上项目、以及中宣部、文旅部、网信办委托项目等，参撰著作 3 部，发表论文 50 余篇。

王立威

北京大学智能学院教授。主要研究领域：长期从事机器学习研究。目前主要致力于机器学习基础理论，隐私算法的设计与分析，以及医疗影像诊断算法与系统的开发。王立威教授在 NeurIPS, ICML, TPAMI 等国际顶级期刊和会议上发表论文 150 余篇。关于深度学习优化理论工作曾入选 ICML19 十大最具影响力论文。入选 AI’s 10 to Watch，首位获得该奖项的亚洲学者。获得首届国家自然科学基金优秀青年基金。带领团队获得首届天池 AI 医疗大赛冠军。

步　一

北京大学信息管理系助理教授，国家工程实验室联聘助理教授。北京大学管理学学士，美国印第安纳大学理学硕士、信息学博士。从事信息计量学和科学评价相关研究，具体研究方向包括人文社会科学大数据研究、复杂网络与系统、科学评价、科技情报等。在 Journal of the Association for Information Science and Technology 等 SSCI/SCI 期刊上发表论文数十篇，担任 Nature Humanities and Social Sciences Communications 等期刊编委。入选第七届中国科协青年人才托举工程，担任国际信息科学技术学会科技情报专业委员会主席。

孟　凡

北京大学信息管理系助理教授，国家工程实验室联聘助理教授。2012 年获北京大学学士学位，2017 年获中国科学院大学博士学位。主要研究兴趣包括机器学习、深度学习及其在图书情报、金融等领域的应用。相关成果发表在 IEEE TNNLS、IEEE TCYB、IEEE TITS、Neural Computing and Applications、Knowledge-Based Systems、ICDM 等高水平期刊会议上。

黄　得

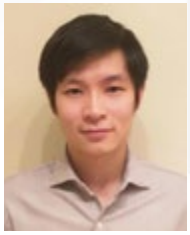
北京大学数学科学学院研究员 / 助理教授，国家工程实验室联聘助理教授。主要研究方向为偏微分方程分析与计算、流体方程分析、随机矩阵理论。曾获 2022 年海外优青项目。

张　成

北京大学数学科学学院概率统计系助理教授，博士生导师，国家工程实验室联聘研究员 / 助理教授。研究兴趣主要在统计推断、机器学习和计算生物学的交叉领域。等变标准化流变分贝叶斯演化推断的工作受邀在 2020 年 NeurIPS 大会作口头报告。在 Journal of The American Statistical Association, Bayesian Analysis, Statistics and Computing, NeurIPS, ICML, ICLR 等统计学著名期刊以及机器学习顶级会议发表多篇论文。

张瑞勋

北京大学数学科学学院金融数学系研究员、国家工程实验室联聘研究员、博雅青年学者、博士生导师，入选海外高层次人才计划青年项目，曾获 S&P Global Academic ESG Research Award 国际研究奖。2011 年获北京大学数学与应用数学、经济学（双学位）学士学位，2015 年获 MIT 应用数学博士学位。主要研究兴趣包括金融统计、金融科技，可持续投资，数据科学，以及金融和人工智能交叉领域。



梅文俊

北京大学工学院力学与工程科学系助理教授、研究员。国家工程实验室联聘助理教授。主要研究方向为社会网络动力学的建模与分析、网络博弈与群体智能。近 5 年在 Nature Communications、IEEE Transactions on Automatic Control、Automatica 等国际顶级期刊与会议发表论文十余篇。参与国家自然科学基金重点项目 1 项。担任 SSCI 国际期刊 Journal of Mathematical Sociology 编委。

何冠楠

北京大学工学院助理教授、国家工程实验室联聘助理教授。国家海外高层次人才计划青年项目专家。2014 年获清华大学学士学位，2019 年获美国卡耐基梅隆大学博士学位，2019-2021 于美国麻省理工学院从事博士后研究，MIT Energy Fellow。主要研究领域为能源系统优化与数据分析、能源经济与政策，尤其是电化学与氢储能系统的复杂决策。曾在 Nature Energy, Joule，Energy & Environmental Science 等著名期刊与会议发表论文二十余篇。

刘　洋

北京大学王选计算机研究所助理教授、研究员，博士生导师，国家工程实验室联聘助理教授。博士毕业于英国剑桥大学计算机科学专业，曾任英国牛津大学博士后研究员。研究方向为跨媒体智能，即运用人工智能算法处理跨媒体信息分析与智能融合的相关问题，具体围绕多媒体信息表征，跨媒体对齐与智能融合，面向开放动态环境的跨媒体分析三个方面展开。研究成果多次发表在计算机视觉、机器学习顶级会议上（CVPR, ICCV, AAAI 等），获国际专利授权，四项国际竞赛获奖。

吴　磊

北京大学数学科学学院信息与计算科学系助理教授，国家工程实验室联聘助理教授。2012 年获南开大学学士学位，2018 年获北京大学博士学位。2018-2021 年在普林斯顿大学应用与计算数学系从事博士后研究。2021 年 12 月入职北京大学。研究方向为深度学习的数学理论。

袁　坤

北京大学国际机器研究中心助理教授，国家工程实验室联聘助理教授。2011 年在西安电子科大获得学士学位，2014 年在中国科大获得硕士学位，2019 年在美国加州大学洛杉矶分校获得博士学位，2019 年至 2022 年在阿里巴巴达摩院美国西雅图研究中心任高级算法专家。主要从事大规模分布式机器学习与最优化方法的研究。 在 2018 年获得 IEEE 信号处理协会青年作者最佳论文奖。主要研究方向：优化理论与算法、分布式机器学习、分布式信号处理。

邱韫哲

北京大学信息管理系助理教授，国家工程实验室联聘助理教授。2013 年毕业于清华大学工业工程系，2016 年获得北京大学工业工程与管理系硕士学位，2022 年博士毕业于美国圣路易斯华盛顿大学奥林商学院供应链管理专业。在 MSOM、IISE Transaction、IEEE TASE、IJPR 等国际期刊发表学术论文与会议论文十余篇。研究方向包括供应链管理与金融、大数据管理与应用、可持续供应链、仿真优化等。

唐聿劼

北京大学工学院工业工程与管理系助理教授，国家工程实验室联聘助理教授。2013 年获清华大学电子信息科学与技术专业学士学位，2019 年获美国加州理工学院博士学位，2019 年至 2022 年于哈佛大学工程与应用科学学院任博士后研究员，2022 年入职北京大学。主要研究方向为分布式优化、强化学习理论及其在能源系统优化中的应用。

尤鹏程

北京大学工学院工业工程与管理系助理教授，研究员，博士生导师，国家工程实验室联聘助理教授。主要研究方向包括优化控制、强化学习、市场机制等在低碳电力能源系统、能源经济中的应用。于浙江大学获得学士和博士学位，曾任美国约翰霍普金斯大学博士后研究员和讲师，也曾在美国加州理工学院和美国西北太平洋国家实验室从事过科研工作，入选微软亚洲研究院铸星学者计划。

周沛劼

北京大学前沿交叉学科研究院国际机器学习中心研究员、博士生导师、助理教授，国家工程实验室联聘助理教授。研究领域为应用数学，主要科研兴趣为数据驱动的动力学建模与计算，研究成果发表在 Nature Communications, Physical Review X, Molecular Systems Biology, Nature Genetics, Nature Machine Intelligence, Nucleic Acids Research, Briefings in Bioinformatics 等应用数学、数据科学和计算系统生物学的交叉学科期刊。曾于 2019 年获得北京大学优秀博士论文奖。

黄　晶

国家工程实验室算法专家
研究方向：人工智能在互联网、金融、零售行业的应用

赵　杰

国家工程实验室高级算法工程师
研究方向：图像处理、机器学习、深度学习

李　楠

国家工程实验室算法工程师
研究方向：自然语言处理、文本挖掘

李新海

国家工程实验室大数据应用与算法研发高级工程师
研究方向：数据安全、实时计算、大数据应用

董豆豆

国家工程实验室软件工程师
研究方向：应用统计、可靠性分析

李　镇

国家工程实验室算法工程师
研究方向：运筹优化、机器学习、数据分析

芦佩峰

国家工程实验室软件工程师
研究方向：医疗大数据统计分析与应用开发

◆ 战略规划部

邓　冰

国家工程实验室战略规划部部长
博士
高级工程师

王新民

国家工程实验室党支部书记
北京大学长沙计算与数字经济研究院秘书长兼院长助理
高级项目经理
高级工程师
高级信息系统项目管理师
管理咨询师

孙　震

国家工程实验室项目经理

◆创新研究部

杨　超

国家工程实验室副主任兼创新研究部部长

北京大学数学科学学院教授

徐克付

国家工程实验室特聘研究员

王剑晓

国家工程实验室副研究员

张　一

国家工程实验室特聘副研究员

王　娟

国家工程实验室特聘副研究员

崇　滨

国家工程实验室特聘副研究员

王冉冉

国家工程实验室助理研究员

张　敏

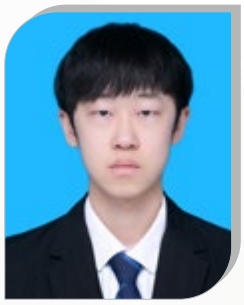
国家工程实验室助理研究员

刘浩洋

国家工程实验室助理研究员

侯　琛

国家工程实验室助理研究员



金鹏展

国家工程实验室助理研究员

张　博

国家工程实验室助理研究员

• 博士后

杨明远

余　倩

陈齐浩

徐　成

◆产品工程部

袁晓如

国家工程实验室常务副主任兼产品工程部部长

北京大学智能学院研究员

黄　晶

国家工程实验室产品工程部副部长、算法专家

黄永贵

国家工程实验室高级工程师

赵　杰

国家工程实验室高级算法工程师

李　楠

国家工程实验室算法工程师

李新海

国家工程实验室大数据应用与算法研发工程师

董豆豆

国家工程实验室软件工程师

李　镇

国家工程实验室算法工程师

芦佩峰

国家工程实验室软件工程师

◆综合管理部

董　盼

国家工程实验室综合管理部部长

黎　娜

国家工程实验室高级宣传专员

钱　玲

国家工程实验室科研助理

文丽君

国家工程实验室财务助理

王晓鹏

国家工程实验室科研助理

李思辰

国家工程实验室人力资源发展专员

◆实验室代表性科研成果

•数字生态指数　负责人：张平文

数字中国建设是引领中国迈向经济强国的重要引擎。“十四五”规划明确指出，要营造良好的数字生态，助力数字政府、数字经济和数字社会等数字中国各项建设。实验室联合全国多方成立数字生态协同创新平台，共同开展数字生态指数及相关理论研究，为营造良好数字生态提供科学评估依据和实践抓手工具。数字生态指数大会旨在发布数字生态指数及相关专题研究成果，并邀请国内顶级研究学者、企业高管、智库专家等共聚一堂，就数字经济、数字社会、数字政府等前沿问题进行思想碰撞、实践探讨和格局研判。首届大会于 2020 年 10 月在北京召开，第二届大会于 2021 年 10 月在重庆召开，第三届大会于 2022 年 11 月 26 日在线上召开。“首届数字生态与治理论坛暨数字生态指数 2023 发布会”于 9 月 22 日、23 日在长沙召开。新华社每年对发布会进行系列报道，相关成果在政府部门、研究机构以及智库媒体获得广泛关注。

▲ 数字生态指数（2020）发布会，北京

▲ 数字生态指数 2020

▲ 数字生态指数（2021）发布会，重庆

▲ 数字生态指数 2021

▲ 数字生态指数（2022）发布会，线上

▲ 数字生态指数 2022

▲ 数字生态指数（2023）发布会，长沙

▲ 数字生态指数 2023

▲ 数字生态指数 (2024) 发布会，武汉

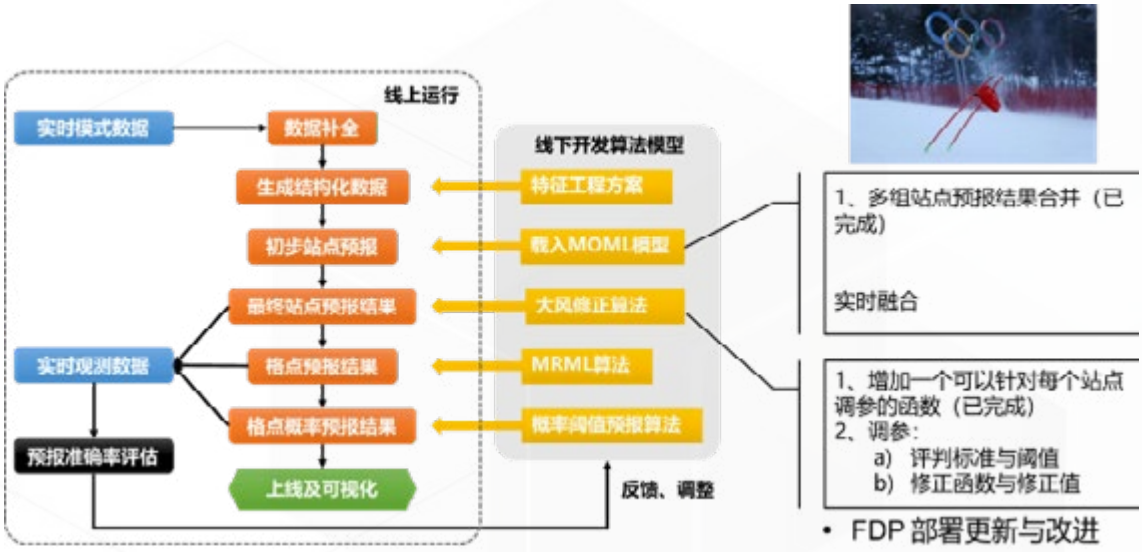
▲ 数字生态指数 2024

数字生态指数提炼了四型联动的中国省级数字生态现状——全面领先型、赶超壮大型、发展成长型、蓄势突破型；呈现了中国城市级数字生态发展格局，并总结了三型驱动的城市数字生态发展模式——能力驱动型、应用驱动型、基础驱动型。此外，国际数字生态指数对中美欧等世界 40 多个主要国家进行发展水平和驱动模式等研究，对国际数字治理格局进行深度研判。相关数据及研究报告在官方网站 digiteco.com.cn。

为进一步推动我国数字生态建设，研究小组代表实验室申请加入了全国信息化和工业化融合管理标准化技术委员会下属的数字经济标准工作组和数字化转型标准工作组（下图所示）。未来，打造理想数字生态的路径不是唯一的，不同地区数字生态发展水平和结构差异要求各地寻求因地制宜、切合实际的差异化发展战略。

• 智慧冬奥 2022 天气预报示范　负责人：张平文

为进一步加强北京冬奥气象预报工作，满足冬奥会“百米级、分钟级”精细化天气预报需求，中国气象局组织实施了“智慧冬奥 2022 天气预报示范计划（SMART2022-FTP）”。该计划征集国内优秀高分辨率区域数值天气模式和人工智能等客观预报技术、方法和系统，通过在 2021 年北京冬奥测试赛和 2022 年北京冬奥会正式赛期开展天气预报示范的方式，为冬奥气象预报服务提供强有力科技支撑。该计划旨在充分集成并应用近年来气象预报现代化的发展成果，以冬奥预报需求为导向，坚持研发、应用、评估等全链条衔接贯通，加快气象服务的科技创新，由北京大学张平文院士领衔的气象大数据实验室团队应邀参加，其精细化预报产品在 2021 年冬奥测试赛服务期间表现优异，入选 2022 冬奥气象服务站点天气预报产品八项之一。



• 科学战疫

a. 基于轨迹大数据的疫情风险评估与密接筛查

联合攻关，精细化疫情风险评估和疑似病例预测研究。研发轨迹大数据与物理平均场理论的 HiRES 风险地图时空模型、HiRES-p 个人流行病感染风险客观评分模型，以及基于该模型的密接个体筛查算法。提供宏观风险度量，计算时间、空间、个体三类微观风险；数据保留全部空间维度和个体维度，对时间适当降维，信息留存最大化。

合作单位致信感谢：“在四月黑龙江防控中发挥了重要作用，黑龙江省当地政府相关部门给予了高度肯定。在支撑吉林省疫情防控工作中也取得了阶段性成果。”

b. 多源大数据疫情防控研判系统

·PESA-Global：实现对国际疫情发展态势做出预测和研判，每天上报一次《疫情预测日报》，每周上报一次《疫情防控政策研判报告》。

·精准预测获国家有关部委的“点赞”：“体现了团队卓越的科研创新能力和坚实的工程实践能力，也体现出北京大学围绕国家重大战略需求开展科学研究和社会服务的精神。”

c. 国内外疫情可视化

国内外疫情可视化：2020 年 1 月下旬开展相关工作，1 月 30 日首次发布，与国内外志愿者协同数据整理、设计。应用于中央广播电视总台（CCTV）视听新媒体中心，系统每天输出全国各省市新冠疫情数据可视化效果 , 支持视听 新媒体中心数据可视化系列短视频《独家“视”角》的制作；应用于腾讯微信疫情数据动态可视化。（分别于 2020 年 2 月 22 日发布视频和 2 月 7 日上线）。

d. 新冠病毒潜伏期估计及传染动力学模型

综合利用流行病学数据、病毒基因数据和交通流量数据等各种信息，从数学模型和统计模型相结合的角度对本次新型冠状病毒疫情的发生和发展过程进行深入研究，力图阐明新 型冠状病毒的疾病流行特点和规律，为防控新型冠状病毒疫 情提供以数据为基础的定量化决策建议”。

e. 新冠知识图谱数据和访问平台

OpenKG 开放的新冠知识图谱数据包括八大类共 17 个知识图谱，分别为新冠科研图谱、新冠百科图谱、新冠健康图 谱、新冠临床图谱、新冠英雄图谱、新冠热点事件图谱、新冠流行病学图谱和新冠物资图谱，均提供访问接口和详细的使用手册。

f. 城市交通活力恢复指数

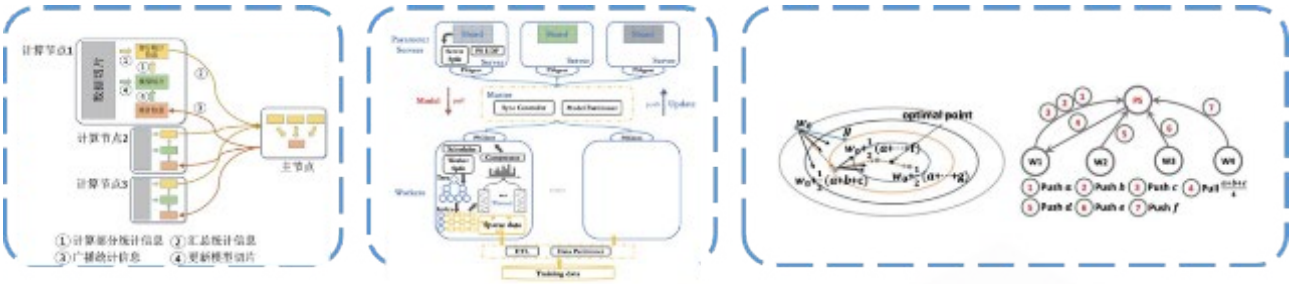
实验室联合常务理事单位滴滴出行，中国电子信息产业发展研究院（赛迪研究院），基于滴滴平台城市交通轨迹、道路拥堵数据和通勤订单数据所 计算的城市交通活力恢复指数可对相关政策研判及研究提供支撑。共同发布城市交通活力恢复指数排行榜。此次发布的城市交通活力恢复指数，是结合城市交通轨迹、道路拥堵数据及滴滴平台通勤数据综合分析得出。指数越高，说明城市交通活动量更积极。

• 基于参数服务器的机器学习系统 Angel 负责人：崔斌

随着数据规模爆发式增长，数据特征的维度往往能够达到十亿甚至百亿的规模。而在训练高维机器学习模型的过程中，各个计算节点需要互相通信以完成分布式的训练。这个过程往往意味着巨大的存储和通信开销，远远超过了单个机器的能力。

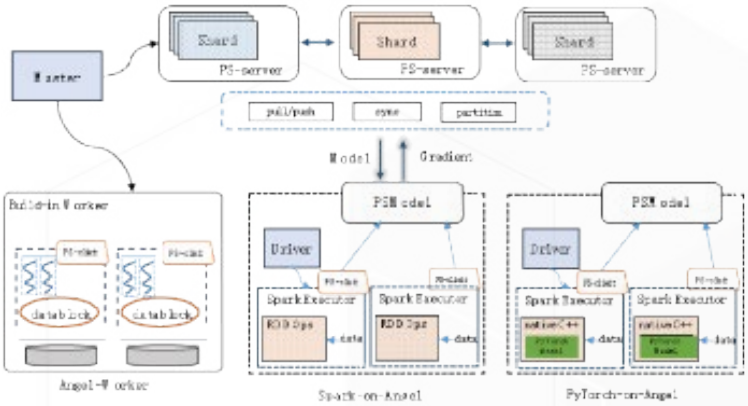
通过分布式训练系统架构，机器学习算法并行方案设计，梯度通信以及一致性同步协议优化，探索更高效的分布式拓展方案。

广泛应用于工业界互联网公司的广告精准推荐、用户兴趣挖掘等场景，在开源网站 GitHub 上获得来自国内外开发者的上千颗 Star；相关工作得到国家重点研发计划支持。

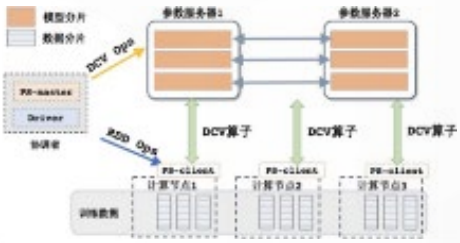


列式划分的分布式训练　　基于特征并行的GBDT算法设计　　异构感知的一致性同步协议优化

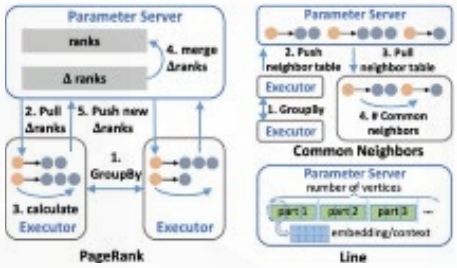
通过结合参数服务器和机器学习算法二者的特点，开发了很多高效、易用的机器学习算法，并与大数据处理引擎 Spark、深度学习系统 PyTorch 构建良好的生态，广泛适用于不同场景的机器学习应用。



▲ Angel 系统架构示意图



▲ Spark-on-Angel 中的模型并行



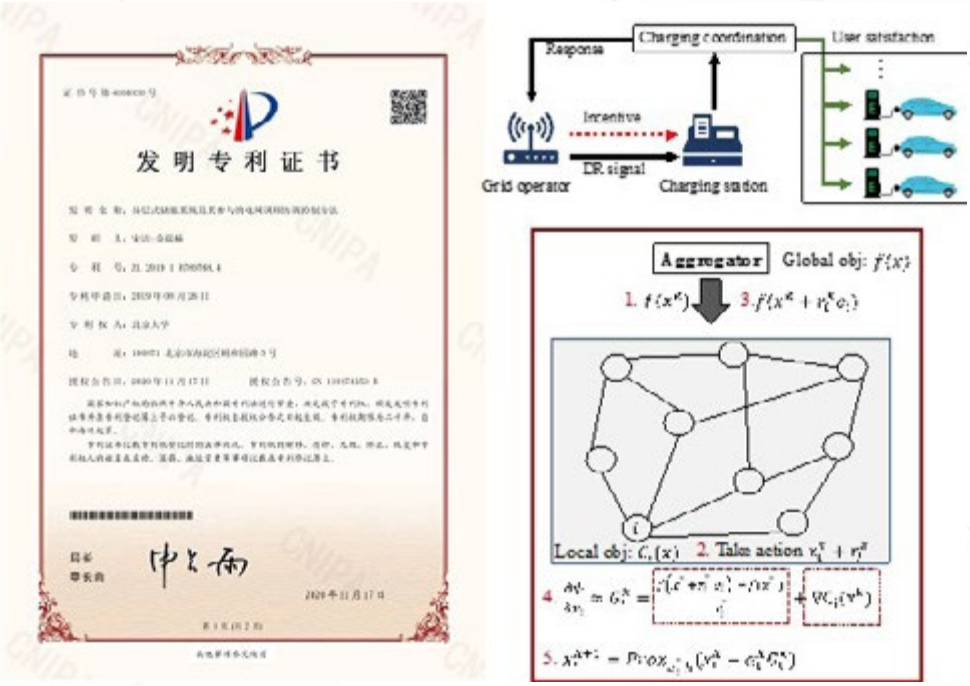
▲ PyTorch-on-Angel 中的图算法优化

• 机理与数据驱动的能源系统建模优化 负责人：宋洁

分布式架构是未来能源系统的重要特征之一。在“双碳”目标背景下，分布式能源快速发展，能源生产逐步向集中式与分散式并重转变，系统模式由单一主体为主逐步向与海量异构分布式资源并行发展。发电侧与用电侧接入大量分布式可再生能源，需求侧管理也需要协调大量分布式主体。然而不同于集中式的系统优化，分布式系统的规划与控制需要考虑网络拓扑结构、节点通信约束、隐私保护以及求解空间过大等一系列问题。

为此，针对分布式储能系统规划的求解空间过大以及网络潮流非线性约束等问题，基于负荷数据，仿真优化各节点容量配置；针对分布式能源系统的运行优化，采用改进一致性控制算法协调多主体、实现用电成本优化；对于充电站通过协调异质、可变规模的电动汽车充电参与需求侧响应的问题，设计基于数据驱动和深度强化学习算法的分布式充电决策架构，共享单主体的强化学习算法与参数以提高训练与决策效率；对于分布式的台区日冻结量安全预警管理，首先利用基于国网上海电力公司的配用电大数据平台开发的台区画像辅助日冻结量预测算法给出台区预警等级，构建最小化激励成本与风险成本的优化问题，对台区进行需求侧响应并提供个性化激励，保证电力系统安全稳定运行并降低风险管理成本约 75%；针对分布式需求响应中的隐私保护问题，考虑设计完全避免分布式主体与外界进行数据交换的分布式优化算法，分布式主体仅通过接收全局信息进行梯度估计和决策，理论分析表明一定条件下算法能够在避免数据交换的同时得到最优解。

相关研究成果已发表在 IEEE Transactions 等国际顶级期刊，获得发明专利授权 1 项，获得 IEEE 会议 Best Student Paper Award，并获得 2020 中国工业工程与精益管理创新大赛特等奖和 2022 日内瓦国际发明特别展金奖。为分布式能源系统建模与优化提供理论基础与算法设计，助力我国未来低碳能源系统下分布式能源系统快速发展。

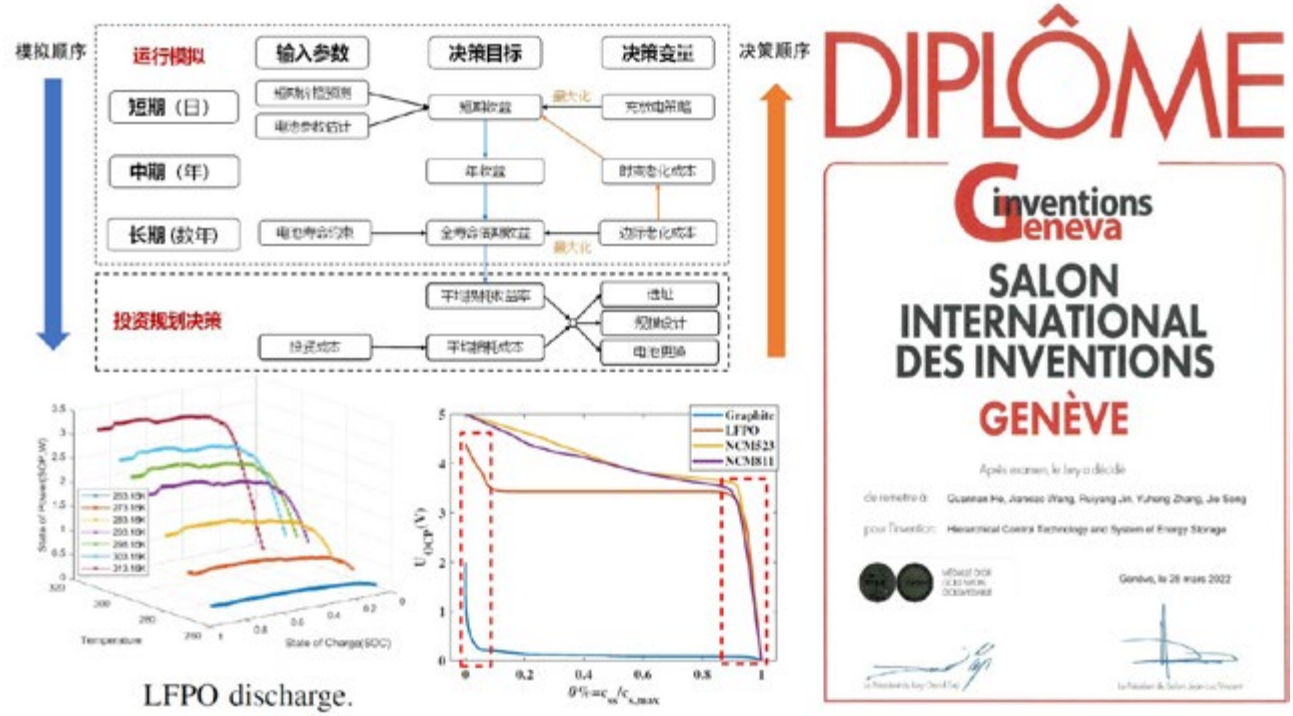


▲ 分布式能源系统研究成果

电化学储能是实现能源与交通系统电气化、低碳化的关键技术。电化学储能的老化是不可避免的复杂非线性过程，严重影响电化学储能的经济性与安全性。过去的电化学储能优化决策与经济评估缺乏对电化学储能老化特性的精确建模及老化成本的严谨定义，存在一系列不合理假设，影响了决策的最优性和评估的精确性。

为此，研究内嵌多物理场一致性的电解制氢与电池储能建模与控制方法，提出表征电化学多物理场参量分布特性的端口机理模型，基于数据驱动的方法精准辨识多单体多物理场参量分布不一致性，为电解制氢及储能的高安全、长寿命优化控制提供基础策略。进而，提出了电化学储能的跨期综合优化决策框架，涵盖短期调度与控制决策和长期运行规划和维护决策，形成多时间尺度耦合决策模型，基于美国加州电网的算例分析表明，与现有方法相比，所提出的决策框架可以实现 30% 以上的经济效益提升。

相关研究成果发表于 Nature Energy、Joule、IEEE Transactions 等国际顶级期刊，获得 2022 年日内瓦国际发明展金奖，为高比例新能源接入下新型电力系统的安全经济运行提供创新性解决路径，为我国健康有序发展氢能和储能技术产业、实现碳中和目标提供高质量政策建议。



▲ 储能建模与跨期决策研究成果

诸如光伏、风电、径流式水电的随机能源强烈依赖天气条件，已有发电功率预测研究关注于提升机器学习方法本身的泛化能力，将数值气象直接作为输入，强非线性对学习器性能提出严格要求。然而，忽视了风电、光伏发电与数值气象之间的物理规律。

为此，提出解析模型与数据融合驱动的新能源预测模型，利用发电功率与数值气象解析模型刻画的关键特征作为输入，提升数据驱动方法的学习性能。相比国际已有的新能源预测模拟方法，所开展研究能够实现新能源随机发电资源的精细预测模拟。基于 24 小时前、分钟级颗粒度、3km×3km 空间网格化的数值天气预报，由高空云图推演计算地球表面风速、直射 / 散射 / 反射光强、环境温度等数值气象特征。

相关研究成果发表于 Nature Communications、IEEE Transactions 等国际顶级期刊，获得由 Elsevier 和国际太阳能协会全球两年举办一次的“可再生转型十佳奖”、全国发明展金奖、IEEE 会议 Best Paper Award 等，为我国风电和光伏技术产业的快速发展提供基础理论工具，助力提升高比例新能源电力系统的可观性。

▲ 新能源预测及应用研究成果

• 粮食应急保障体系布局与优化运作　负责人：宋洁

传统燃气、电力、粮食等供应链调度依赖集中控制的技术策略与通讯方式，然而随着供应链各环节分散主体的大规模接入，已有研究普遍提出云 - 边计算协同的分布式并行理论方法，充分利用边缘侧的计算能力，规避集中主站的通讯负担和计算压力。已有分散决策文献选取定步长或与迭代次数相关函数变步长更新拉氏乘子，选取不当甚至导致振荡和发散。

为此，提出基于动态乘子优化的分布式计算方法，研究基于神经网络学习的初值选取算法，以数据驱动强化均衡趋优过程的收敛性。在粮食安全领域，提出数据驱动的中国粮食应急保障技术，构建对应急因素敏感的粮食应急安全评价体系，划分粮食安全区域，建立基于时空网络的具有鲁棒性的动态应急随机规划模型，服务于粮食宏观调控应急的快速反应和科学决策，进一步保障我国粮食安全。在能源安全领域，提出可再生能源、氢能、热能与电力系统多轮驱动的综合能源耦合供应体系，研究支撑高比例甚至 100% 新能源电力系统稳定运行的综合能源优化配置技术，形成异质能源共享互补的能量流分析方法。

相关研究成果发表于 Energy Environmental Science、IEEE Transactions 等国际顶级期刊，服务于国家粮食和物资储备局应急保障体系，纳入十四五粮食安全规划，得到国家粮食与物资储备局的批示与致谢，为我国粮食安全与能源安全提出政策建议，带来重大经济效益和社会效益。

▲ 粮食应急保障体系研究成果

数据处理及统计分析中心

◆ 中心简介

数据处理及统计分析中心由北京大学周晓华教授团队负责，承担大数据预处理与质量控制技术的研发应用，办公地址位于北京大学北京国际数学研究中心。

该中心组建了 17 人的核心科研团队，在大数据处理与控制技术、自然语言处理、图像处理、因果推断、重大疫情发生和发展规律的数学统计和建模等方向上取得了关键性突破，并成功应用于阿尔茨海默病 (AD) 患者和正常人的大脑连接研究、非依从情况下行为干预效果的因果推断等。

◆ 中心团队成员介绍

周晓华教授

北京大学讲席教授，北京大学生物统计系主任，北京大学国家药品医疗器械监管科学研究院副院长，北京大学重庆大数据研究院副院长、研究院智慧中医药研究中心主任，西部（重庆）科学城高新区“金凤凰” 杰出人才 , 美国科学促进会（AAAS）会士，美国统计学会 (ASA) 会士，国际数理统计研究院 (IMS) 会士。曾荣获国际贝叶斯统计科学学会 Mitchell 奖、中国产学研合作创新奖、世界中医药学会联合会中医药国际贡献奖 - 科技进步二等奖、国家基金委海外杰青、教育部高层次文教专家、教育部海外名师、2022 十大重庆科技创新年度人物等奖项和荣誉。研究方向主要集中于医学诊断学的统计方法，因果推断方法，医疗大数据建模，新兴传染病的数学和统计建模，中医药研究的统计方法，发表 SCI 学术论文 290 余篇。

耿　直

北京大学数学科学学院教授，北京大学数理统计研究所所长

研究方向：因果推断、数理统计、生物医学统计、因果网络、贝叶斯网络、图模型

房祥忠

北京大学数学科学学院教授

研究方向：生存分析、可靠性、纵向数据、基尼系数、空间统计等

邓明华

北京大学数学科学学院教授及概率统计系副系主任，北京大学定量生物学中心教授

研究方向：生物信息学

艾明要

北京大学数学科学学院教授，统计学教研室主任

研究方向：试验设计与分析、计算机试验、高维数据分析、应用统计

葛　颢

北京国际数学研究中心长聘副教授
研究方向：随机过程与物理、化学、生物的交叉

林　伟

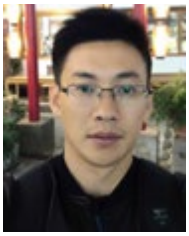
北京大学数学科学学院研究员，国家第十一批“青年千人计划”入选者
研究方向：高维统计、大数据分析、因果推断、生存分析

席瑞斌

北京大学数学科学学院研究员
研究方向：统计理论和算法、计算基因组学和生物信息学

贾金柱

北京大学公共卫生学院研究员
研究方向：高维统计推断、机器学习和因果推断



杨　伟

北京大数据研究院中医大数据中心助理主任，现为中国中医科学院中医临床基础医学研究所数据分析小组负责人

黄新霆

北京大学访问学者
研究方向：中医智能辩证论治、医疗大数据、医疗信息标准

候　艳

北京大学助理教授，北京大学公共卫生学院生物统计系副研究员

张云俊

北京大学公共卫生学院生物统计系副教授

梁宝生

北京大学公共卫生学院生物统计系助理教授

尤　翀

北京大学北京国际数学研究中心副研究员
研究方向：变分贝叶斯、混合模型、模型选择、传染病模型

占　翔

北京大学副教授
研究方向：生物统计与生物信息学、统计遗传、基于希尔伯特再生核空间的统计建模、高维分子组学数据分析与统计推断

胡涛波

北京大学人民医院乳腺外科医师、武汉大学工业科学研究院兼职教授
研究方向：乳腺癌外科手术和精准诊疗、临床数据分析和统计建模、医学生物信息学

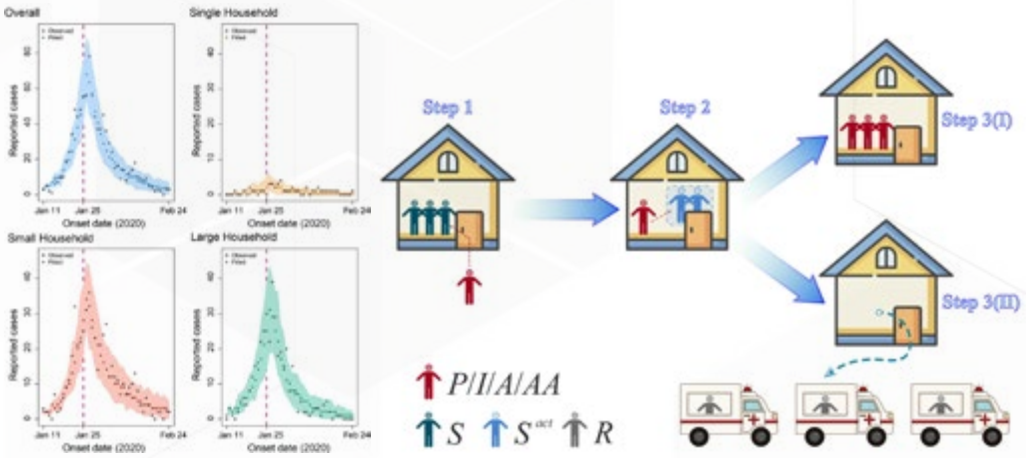
◆科研成果

• 重大疫情发生与发展规律的数学和统计建模　负责人：周晓华

人类与传染病的斗争从古至今一直都在进行。从古老的黑死病到近代的艾滋病，从甲型流感到非典，可以说人类社会的发展史就是一部与传染疾病作斗争的历史，而每一次的传染病疫情都会给全球的发展造成巨大的冲击。尤其是 2020 年伊始，突如其来的新型冠状病毒（COVID-19）疫情在全球范围内持续肆掠，如何科学、高效防控成为了亟待解决的问题。北京大学周晓华团队综合利用流行病学数据、病毒基因数据、交通流量数据等各种信息，从统计模型和数学模型相结合的角度对本次新冠疫情的发生和发展过程进行深入研究，聚焦于病毒特征分析、疫情现状分析、干预措施评估和人群易感性分析等关键问题。

具体来说，项目团队（1）构建了符合于新冠疫情特征的动力学传播模型，解释了新冠独特的传播动态，揭示了家庭感染在新冠疫情传播中的作用，充分考虑到潜伏期的传染性、无症状感染者和医学追踪以及不同时间点防控措施对于疫情的影响，克服传统动力学模型的不足；（2）利用新的数学方法对病毒的潜伏期、隔代时间、再生数等特征进行估计，不仅减少了偏倚，并充分利用丰富的现成数据，提高了病毒特征估计的准确性；（3）证明了新冠的传输异质性相对较高，存在超级传播事件的概率较高，并首次定义超级传播者是感染 6 个及以上二代病例的感染者，从而确定了天津某传播链中存在超级传播者；（4）鉴定了“冷冻”病毒是导致 2020 年 6 月北京新发地疫情和 2021 年 5 月安徽 - 辽宁疫情中主要原因；（5）提出动态调整新冠病毒疫苗的分配以减少我国死亡率、住院率等；（6）通过数学模型对我国清零防疫相关政策提出建议并获得高度关注，揭示了在 2021 年夏秋的病毒变异特性和人群免疫水平下，一旦我国采取了上述参考国家的防疫政策，即便在高度乐观保守的估计下，我国每日新增确诊病例将达到数十万例，每日新增重症也将接近万例。

周晓华团队的相关工作不仅在国际顶刊上进行发表（例如 Science Advances，Nature Communications 等），还产生了较大的社会影响，其中仅对针对我国清零防疫相关政策的工作就有 370 多家重要媒体参与报道本项工作，国内媒体有人民日报、人民政协网、光明网、央视网、南方都市报、观察者网、凤凰网澎湃新闻、今日头条等；国际媒体有英国路透社、美国有线电视新闻网、英国每日邮报、英国世界新闻报、美国彭博社、美国新闻周刊、美国新闻和世界报道等等。周晓华团队为新冠疫情流行病学调查、疫情监测、防疫政策等作出重要贡献。



▲ 周晓华团队研发的重大疫情发生和发展规律的数学统计和建模示例图

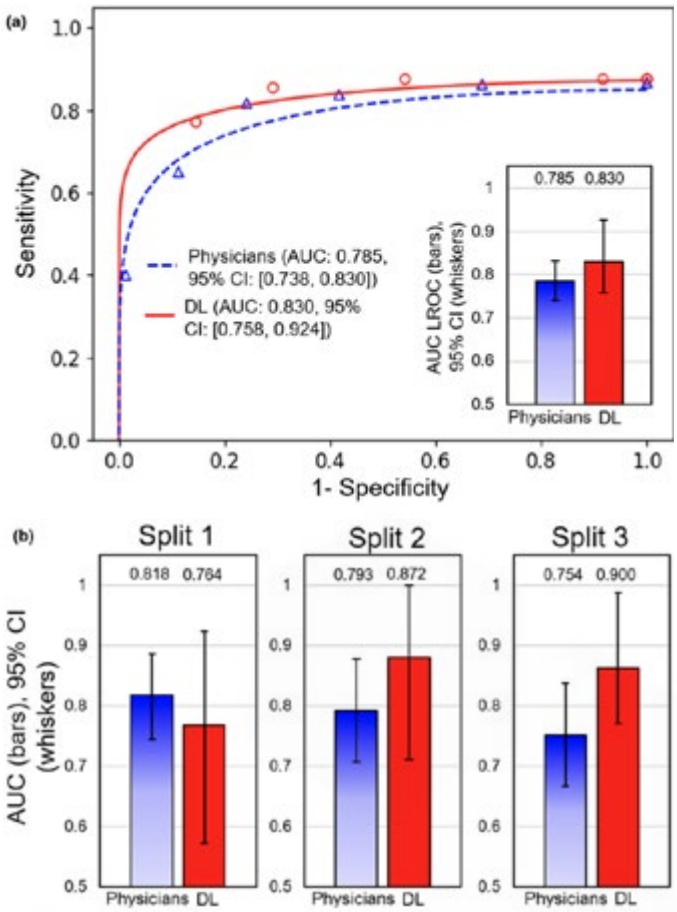
已完成疫情数据共享、预警和可视化平台的开发，建成包括数据采集系统、国内数据监控平台、国际数据监控平台和手机端数据监控平台。

获得相应软件著作权证书 3 项；

软件名称	时间	依托单位
疫情数据收集系统	2020.10	北京大学
新冠肺炎全国疫情趋势分析系统 APP	2020.10	北京大学
新冠肺炎疫情监控分析平台	2020.10	北京大学

• 诊断医学　负责人：周晓华

诊断医学统计方法主要研究诊断测验准确率的评估和推断，正确评估医学诊断方法的准确性、有效性对医药学、影像学研究的发展起着至关重要的作用。本项研究主要包括三方面内容。第一方面的研究集中于影像学中包含多个病变和病变位置信息的诊断的准确度研究。针对包含位置信息的自由响应接受者操作特性试验数据，本研究通过 FROC 类曲线和相应的曲线下面积等指标来评判其诊断准确性，并研究了相应的参数估计方法、非参数估计方法、曲线以及曲线下面积的回归方法，对影像学的位置准确率的评估方法的发展做出了重要的贡献。第二方面的研究集中于伴随诊断的统计方法。在伴随诊断试剂的随机临床试验因现实因素和伦理问题不能正常开展时，本研究提出了新的实验设计方法和相应的统计分析方法，通过结合外部数据和已有的实验数据，对目标伴随诊断试剂阴性、阳性组的药效进行估计，填补了伴随诊断实验统计分析方法的空白。第三方面的研究聚焦于测序方法准确性的研究。在以往的研究中，通过测序检验基因突变往往缺乏金标准，测序检验基因突变的准确性也没有得到合理的评估。本研究利用贝叶斯方法研究了测序方法检验基因突变的准确率，为基因测序领域的准确性评价提供了重要的参考。本项目有多篇已发表和在投的论文，成果丰富。



▲ 诊断医学统计方法对医药学、影像学研究的发展起着至关重要的作用

机器学习中心

◆ 中心简介

机器学习中心由北京大学张志华教授团队负责，承担大数据分析支撑理论与算法的研发应用。

机器学习是数据科学和人工智能的理论基础和核心算法，实验室的任务是研究机器学习的机理，包括但不限于深度学习和强化学习，并且运用机理去解决现实问题，发展机器学习理论，研究新的机器学习模型和原型，设计高效的算法以及实现大规模的算法软件库 / 平台。

该中心的主要任务包括：

1、发展理解深度学习机制的理论；
2、发展高效的优化多层神经网络的算法；
3、针对不同类型问题的网络结构设计；
4、贝叶斯模型和产生式博弈模型；
5、强化学习及其应用；
6、在自然语言处理、精准医学、量化经济等领域的应用。

◆ 中心团队成员介绍

董　彬

机器学习中心主任，北京大学北京国际数学研究中心长聘教授，北京大学国际机器学习研究中心副主任，北京大学国家生物医学成像科学中心研究员

鄂维南

中国科学院院士，北京大学数学科学学院讲席教授
研究方向：计算数学、应用数学及其在力学、物理、化学和工程等领域中的应用等

张志华

北京大学数学科学学院统计学教授
研究方向：统计学、机器学习、自然语言处理

邵嗣烘

北京大学数学科学学院计算数学副教授
研究方向：量子化学和量子物理中的算法、微分方程数值解、图谱理论和谱聚类算法

文再文

北京国际数学研究中心长聘副教授，北京大学工学院工业工程与管理系主任
研究方向：最优化问题理论与算法及其在数据科学和人工智能中的应用

张　立

北京大学大数据科学研究中心助理研究员
研究方向：生物医疗图像处理，机器学习

赵　杰

国家工程实验室算法工程师
研究方向：图像处理、机器学习、深度学习

袁　坤

北京大学国际机器学习研究中心助理教授
研究方向：优化理论与算法、分布式机器学习、分布式信号处理

周沛劼

北京大学前沿交叉学科研究院国际机器学习研究中心助理教授

研究方向：计算系统生物学

张文涛

北京大学国际机器学习研究中心助理教授、研究员、博士生导师

研究方向：数据驱动的机器学习、图机器学习、机器学习系统和 AI4Science

◆科研成果

• 统计学、数学和计算机科学的交叉领域研究　负责人：张志华

张志华团队主要从事统计学、数学和计算机科学的交叉领域研究。具体地，试图应用统计推理、运筹优化和随机分析等来研究机器学习和数据科学的机理和算法，并且运用它们去解决实际问题。近年来在随机梯度算法、深度神经网络、联邦学习和强化学习等理论分析方面取得了一系列的成果：

1）提出了一个统一框架来分析二阶方法的局部和全局收敛性，引入了 Nesterov 加速去改善随机二阶算法的收敛性能，由此提出了加速的近似牛顿算法。给出了求解一类优化问题和非线性方程组的经典拟牛顿方法的显性超线性收敛率的理论结果。

2）在深度学习理论方面的工作主要涉及神经网络的几何景观和优化轨迹。具体地，证明了超宽稀疏神经网络依旧具备良好的几何景观，提供了线性神经网络在分类和回归两种常见问题设置下的轨迹收敛分析。

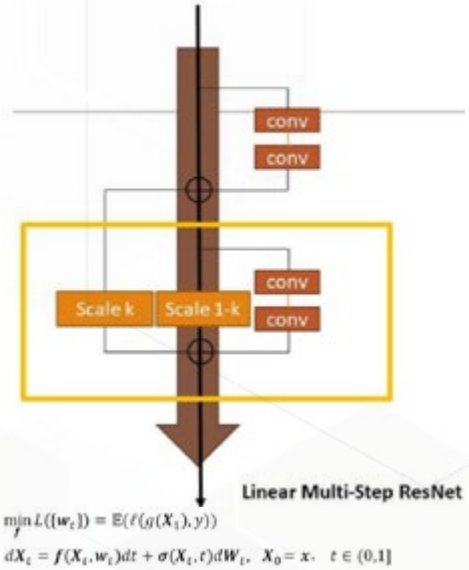
3）首次建立了联邦学习（FL）在数据非独立情况的收敛性结果，理论上揭示了通讯有效性和收敛率之间的权衡；发展了函数极限中心定理，由此首次建立了 FL 中的 Local SGD 方法的统计估计和推断理论。我们发展起来的分析工具独立于 Local SGD 问题本身，它针对一类归纳定义的随机过程给出了依概率意义下的一致收敛速度的精确刻画，该工具可作为经验过程 Donsker 理论中等度连续性质的推广，帮助我们证明这类更复杂的随机过程的 Donsker 不变原理。

4）系统地讨论了强化学习的稀疏策略问题，给出了构造稀疏策略的一类充分必要条件；建立了鲁棒马尔可夫过程的样本复杂度和渐进统计理论；给出了著名的 Q-learning 算法的渐近收敛性和半参数统计理论结果。即将由人民邮电出版社出版《深度强化学习》一书。开源了大规模并行的强化学习算法软件包，针对强化学习中环境模拟速度慢、GPU 擅长批量处理样本、强化学习模型的参数更新及推断可异步执行等特点，作出了创新性的设计与改进。

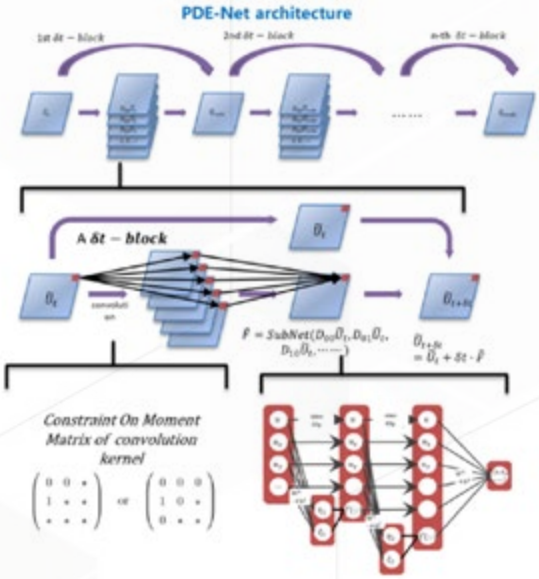
这些成果有助于我们理解机器学习和数据科学的前沿基础性问题。它们发表在机器学习、人工智能、运筹优化领域顶级期刊或会议上，如期刊 Journal of Machine Learning Research （JMLR）、Artificial Intelligence、Mathematical Programming 和会议 COLT、NeurIPS、ICML、ICLR、AISTATS 等。表明团队具有高度的国际影响力。

• 计算数学方法与深度学习的结合　负责人：董彬

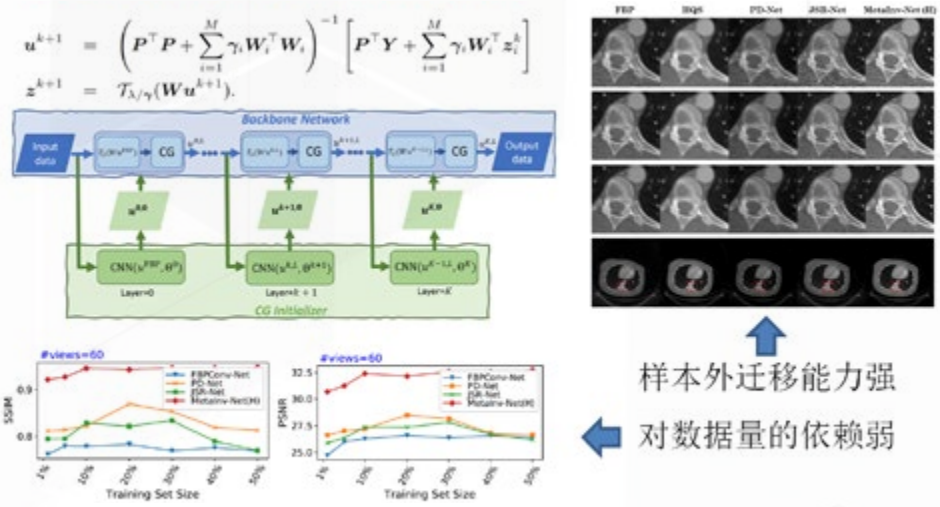
董彬团队尝试计算数学方法与深度学习相结合，建立了数值微分方程和深层网络构架的联系，并针对图像科学中的重要问题，设计机理与数据融合的新算法。该项目一方面用微分方程数值格式来系统的指导深度学习中的深层网络构架设计，这也正是深度学习领域最核心的问题之一，而团队提出的是一个全新的切入点，并在大规模图像识别数据集上验证了该方法的有效性（图 1）；另一方面，利用深度学习端到端的训练思想和深层网络的表达能力，结合偏微分方程有限差分方法，提出一个如何从海量数据中学习未知 PDE 模型并同时能做精确预测的算法框“PDE-Net”。PDE-Net 能够在保证预测能力的同时也保证模型的可解释性，作为反问题求解器以及数据驱动的建模方案初见成效（图 2）；更进一步，团队继续秉承机理与数据融合的思路，提出融合深度学习、强化学习、最优控制和 PDE 的图像重建模型（图 3）、图像分割模型（图 4）、PDE 求解算法和加速对抗训练算法（图 5），这些模型和算法也同时具备了优良的性能和较好的可解释性。



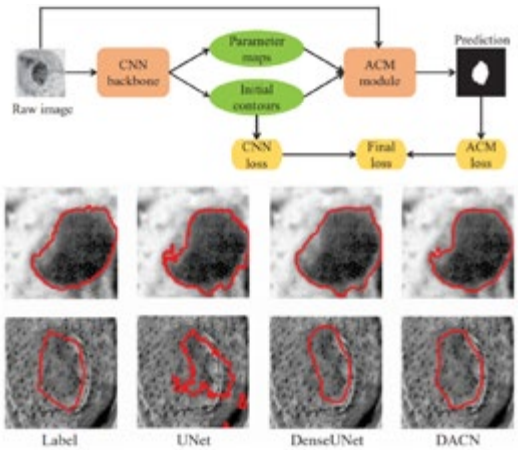
▲ 图 1 数值 ODE 算法诱导新型深层网络构架
Y. Lu et al., ICML 2018.



▲ 图 2 数值 PDE 算法诱导深层网络构架
Z. Long et al., ICML 2018.
Z. Long, Y. Lu and B. Dong, JCP 2019.



▲ 图 3 机理与数据融合的 CT 图像重建模型
H.M. Zhang et al., IEEE TMI 2021.

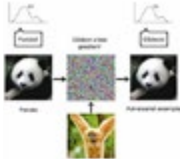


▲ 图 4 神经网络推断 PDE 模型初始化和超参数
M. Zhang, B. Dong and Q. Li, MICCAI 2020.

$$\min_{\theta} \max_{\|\eta\|_\infty \le \epsilon} J(\theta,\eta) := \frac{1}{N}\sum_{i=1}^{N} \ell_i(x_{i,T}) + \frac{1}{N}\sum_{i=1}^{N}\sum_{t=0}^{T-1} R_t(x_{i,t};\theta_t)$$

$$\text{subject to } x_{i,1} = f_0(x_{i,0}+\eta_i,\theta_0), i=1,2,\cdots,N$$

$$x_{i,t+1} = f_t(x_{i,t},\theta_t), t=1,2,\cdots,T-1$$



对抗训练的最优控制刻画

Training Methods	Clean Data	PGD-20 Attack	Training Time (mins)
Natural train	95.03%	0.00%	233
PGD-3[1,2]	90.07%	39.18%	1134
PGD-5[1,2]	89.65%	43.85%	1574
PGD-10[1,2]	87.30%	47.04%	2713
Free-8[1,3]	86.29%	47.00%	667
YOPO-3-5 (Ours)	87.27%	43.04%	299
YOPO-5-3 (Ours)	86.70%	47.98%	476

[1] Code from https://github.com/ashafahi/free_adv_train.
Table 3: Results of Wide ResNet34 for CIFAR10.

2-5 倍训练加速

Training Methods	Clean Data	PGD-20 Attack	CW Attack	Training Time (mins)
TRADES-10[1,2]	86.14%	44.50%	58.40%	633
TRADES-YOPO-3-4 (Ours)	87.82%	46.13%	59.48%	259
TRADES-YOPO-2-5 (Ours)	88.15%	42.48%	59.25%	218

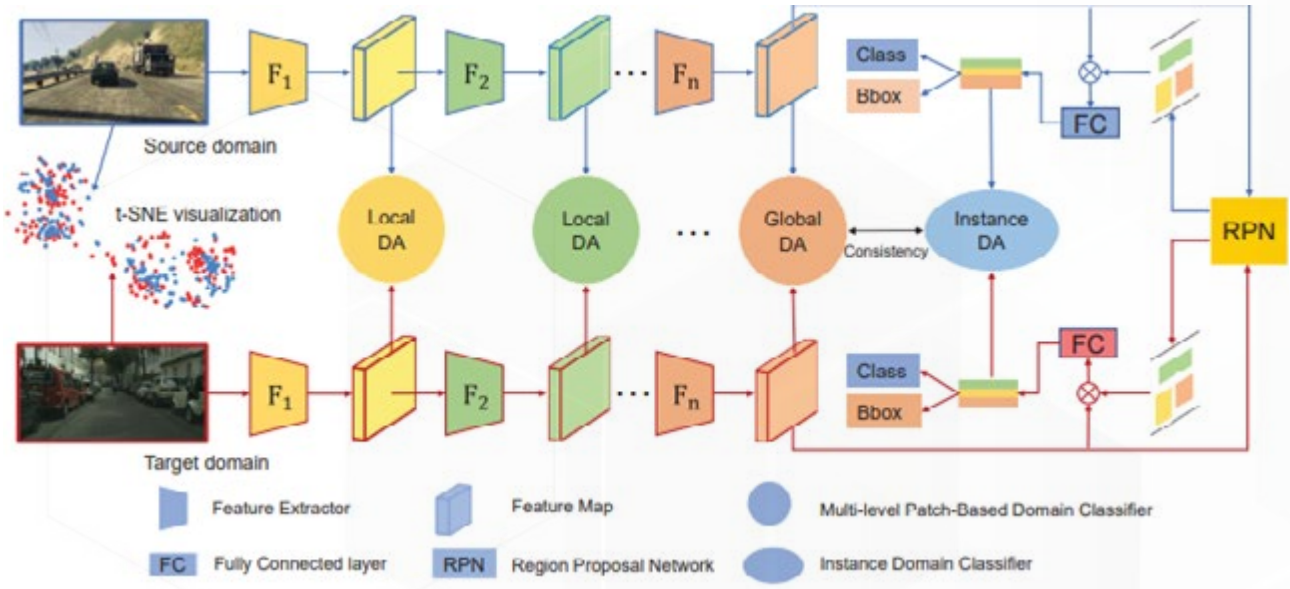
Table 4: Results of training PreAct-Res18 for CIFAR10 with TRADES objective

▲ 图 5 最优控制理论诱导快速对抗训练算法
D. Zhang et al., NeurIPS 2019.

• 多模态领域知识迁移方法研究　负责人：张立

相关论文 1：《Multi-level Domain Adaptive learning for Cross-Domain Detection》, Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops, 2019。

内容简介：本工作希望解决不同领域图像中因为光照、风格、尺度和外观的变化所引起的图像分析模型性能变化，使得在某一模态图像训练得到的模型能在其他模态图像上也能具有较高的性能。以前的工作使用对抗性训练来对齐领域偏移中的全局特征并实现图像信息传输。然而，这些方法不能有效地匹配局部特征的分布，导致跨域对象检测的改进有限。为了解决这个问题，我们提出了一个多级自适应模型来同时对齐局部级特征和全局级特征的分布。我们通过多个实验评估我们的方法，包括不利天气适应、合成数据适应和跨相机适应。在大多数对象类别中，所提出的方法相对于最先进的技术实现了卓越的性能，这证明了我们方法的有效性和稳健性，相关工作获得 ICCV 2019 TaskCV workshop 最佳论文。

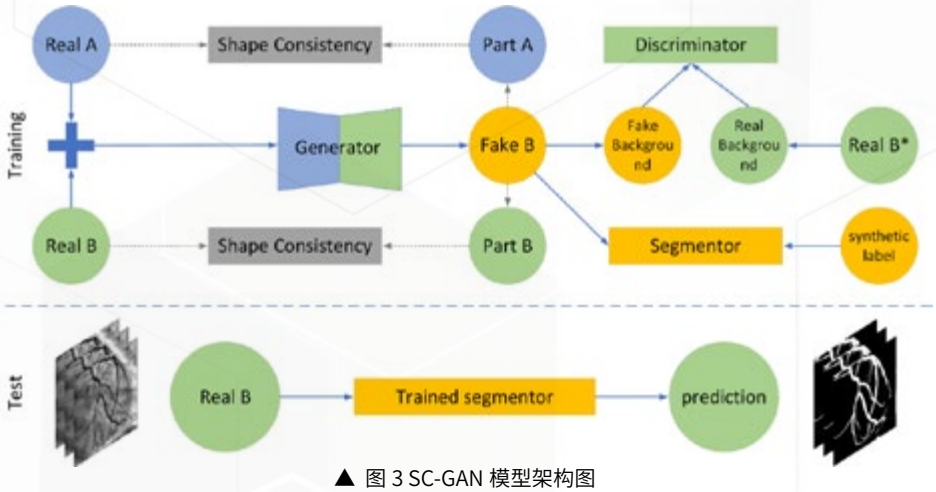


▲ 图 1 网络架构图

▲ 图 2 从 Cityscapes 到带雾 Cityscapes 数据集的适应性定性结果

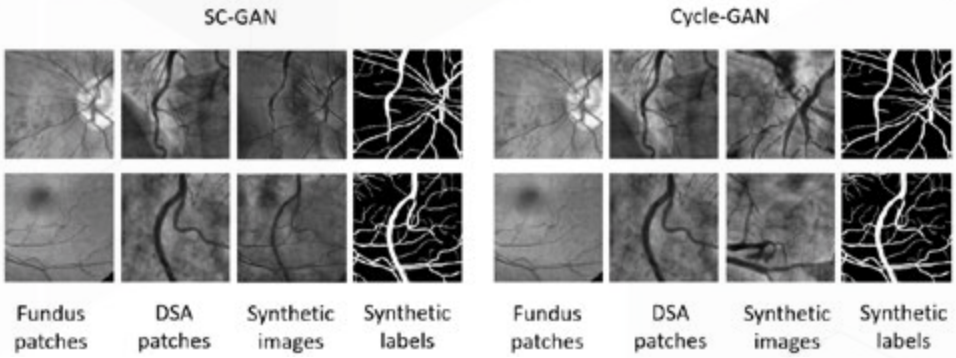
相关论文 2：《Annotation-Free Cardiac Vessel Segmentation via Knowledge Transfer from Retinal Images》, International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2019。

内容简介：在医疗图像分析领域，大部分领域知识迁移都发生在同一器官或对象的不同模态数据之间，不同器官数据之间的知识迁移少之又少，本工作是第一篇将眼底血管和心脏血管的图像信息加以结合，提出了一种无需标注的基于知识迁移的形状一致生成对抗网络模型（SC-GAN），该模型可使用来自公开标注的眼底数据集中的知识来分割冠状动脉。所提出的网络以端到端的方式进行训练，并生成和分割合成图像，这些图像可保持冠状动脉造影的背景并保留视网膜血管和冠状动脉的血管结构。



▲ 图 3 SC-GAN 模型架构图

下面是我们的 SC-GAN 方法一些实验结果。首先是图像合成的结果。跟 SC-GAN 相比，我们的方法合成的图像背景更加逼真，而且合成的血管也能够很好的与合成的标签相对应。

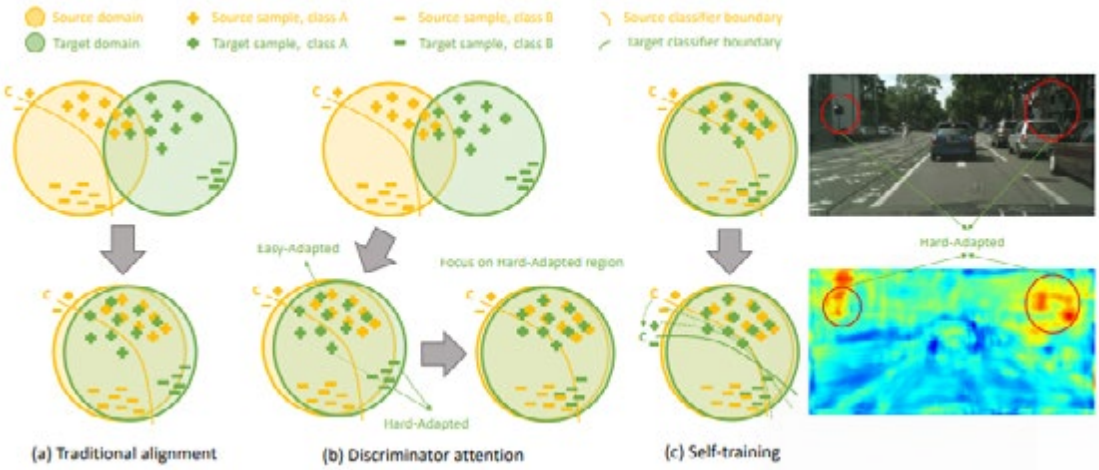


▲ 图 4 SC-GAN 模型合成图像示例

该成果被提名 MICCAI 2019 青年科学家奖。大会收录论文 538 篇，其中仅有 14 篇获得该提名。获得该提名的还包括麻省理工学院、牛津大学、约翰霍普金斯大学、美国梅奥诊所等团队。

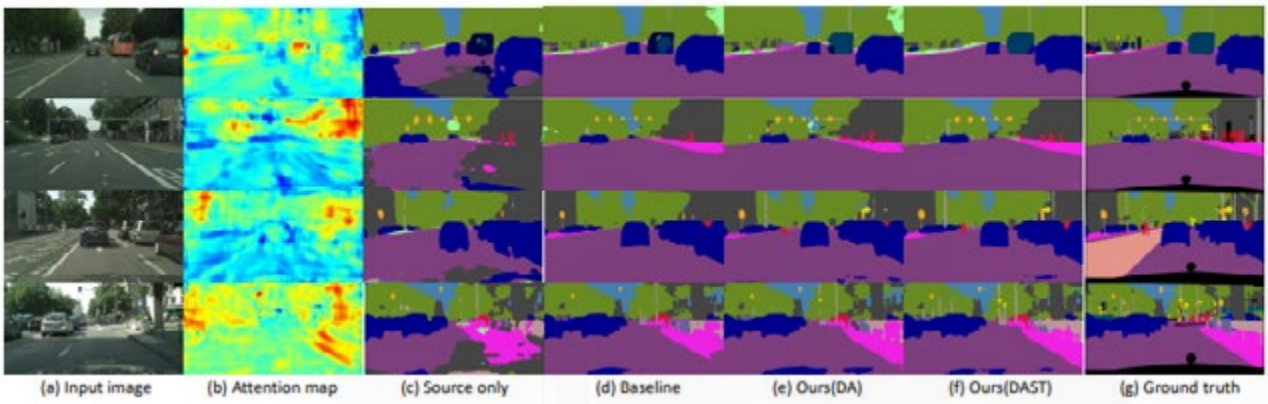
相关论文 3：《DAST: Unsupervised Domain Adaptation in Semantic Segmentation Based on Discriminator Attention and Self-Training》, The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI), 2021。

内容简介：本工作首次将无监督领域自适应在图像分割上的作用提升到了像素级别，这将提高未标记现实世界数据语义分割的性能。在本文中，我们提出了一种使用鉴别器注意力和自我训练策略来减少领域偏移的新方法。判别器注意力策略包含一个两阶段的对抗学习过程，如图 5 所示。它明确区分对齐良好（领域不变）和对齐不良（领域特定）的特征，然后引导模型专注于后者。



▲ 图 5 对抗学习方法

自训练策略自适应地改进模型对目标域的决策边界，这隐含地促进了域不变特征的提取。通过结合这两种策略，我们提出了一种有效且直观的无监督域适应方法用于语义分割，结合了鉴别器注意和自我训练的策略。实验结果和定性示例证明我们的方法在基准数据集上优于以前的最先进方法。



▲ 图 6 分割结果和注意力图

研究贡献：医疗图像分析任务往往具有多模态、少标注等特点，研究多模态领域知识迁移方法，有利于将不同模态的信息融合在一起，对比单一模态的分析可能具有更好的表现。本课题成果一正是针对医疗图像处理的这一特点，不仅实现了跨模态之间的知识迁移，还实现了跨器官之间的信息传递，具有较好的科研先进性，并且具有较好的现实意义。

数据管理与分析中心

◆ 中心简介

数据管理与分析中心由北京大学崔斌教授团队负责，承担大数据挖掘技术的研发应用。

在大数据挖掘技术方面，该中心主要关注文本、视频、时空等典型数据类型，以数据管理和分析相关的系统、算法、应用为主要研究方向，开展大规模并行数据挖掘技术的研发，建设软硬件系统平台，推动计算机相关学科的发展和产学研协同创新。

该中心的主要任务包括：

1、针对典型非结构化数据分析，重点研发大规模文本数据挖掘和自然语言理解的基础技术、面向海量视频图像数据的语义分析技术和时空数据分析技术等。

2、研制大规模并行数据管理平台与挖掘算法库，研制包括图数据库等新型数据库系统，并将平台和系统实际应用到航天等领域。

3、研制大规模分布式机器学习系统，支持高维深度学习模型，支持 CPU-GPU 等异构大数据处理平台。

4.、开展网络相关大数据技术研发，包括社交网络、内容分发网路、移动无线网络等不同场景的数据分析挖掘，大规模并行和分布式算法的推理和优化。

5、结合文理大数据科研需求，逐步构建可持续发展的包含数据共享、基础软硬件和应用软件为一体的大数据共享与服务平台，推进前沿交叉研究。

◆ 中心团队成员介绍

崔　斌

数据管与分析中心主任，北京大学计算机学院教授，网络与信息系统研究所所长
研究方向：数据库系统、大数据管理与分析

宋令阳

北京大学电子学院博雅特聘教授
研究方向：无线通信网络、信号处理和机器学习

王腾蛟

北京大学计算机学院教授
研究方向：大数据管理、互联网数据分析、数据仓库与数据挖掘

高　军

北京大学计算机学院教授
研究方向：分布式数据管理、图数据分析和深度学习

边凯归

北京大学计算机学院副教授，网络与信息系统研究所副所长
研究方向：移动计算、物联网、社交网络

杨　智

北京大学计算机学院网络与信息系统研究所副研究员
研究方向：大数据分析与处理研究

杨　仝

北京大学计算机学院新体制副教授，博士生导师
研究方向：大数据流处理、网络测量、机器学习

陈　薇

北京大学计算机学院网络与信息系统研究所，副研究员，中国软件行业协会数据与智能分会副秘书长
研究方向：大数据管理

闫宏飞

北京大学计算机学院副教授
研究方向：计算金融

黄　群

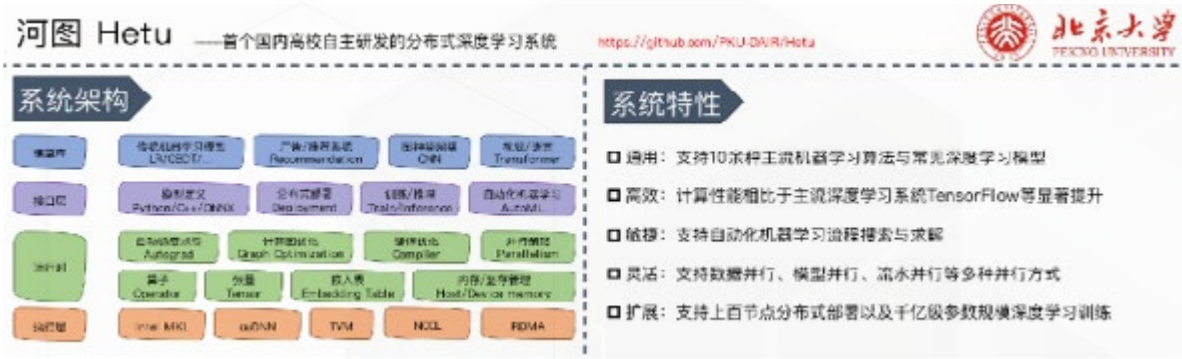
北京大学计算机学院网络与信息系统研究所，研究员，博士生导师
研究方向：计算机网络 / 分布式系统

陈立军

北京大学计算机学院副教授
研究方向：数据库与信息系统

◆科研成果

• 分布式深度学习系统河图　项目负责人：崔斌



河图（Hetu）是一个兼顾创新性和可用性的分布式深度学习系统，是首个由国内高校自主研发的分布式深度学习系统。该系统与现有相关基线系统完全解耦，并具有多项先进特性：

·通用：采用标准的数据流图方式定义深度学习模型，底层实现了丰富的 CPU 以及 GPU 算子库，支持大量常见深度学习模型的计算；另外，该系统还支持 10 余种主流机器学习算法的高效求解，如：LR、GBDT、LDA 等。

·高效：采用多项技术创新，使得其计算性能相比于主流深度学习系统 TensorFlow 等显著提升，在常见的 DNN、CNN、RNN 基准测试上取得了至少 30% 的性能优势。

·敏捷：支持自动化机器学习流程搜索与求解，包括：特征工程、模型选择、超参数调优等，使机器学习模型开发过程更容易。

·灵活：支持参数服务器和全局规约通信架构，支持数据并行、模型并行以及流水线并行等多种并行训练方式，针对硬件环境，通过多种策略组合提供最优分布式部署方案。

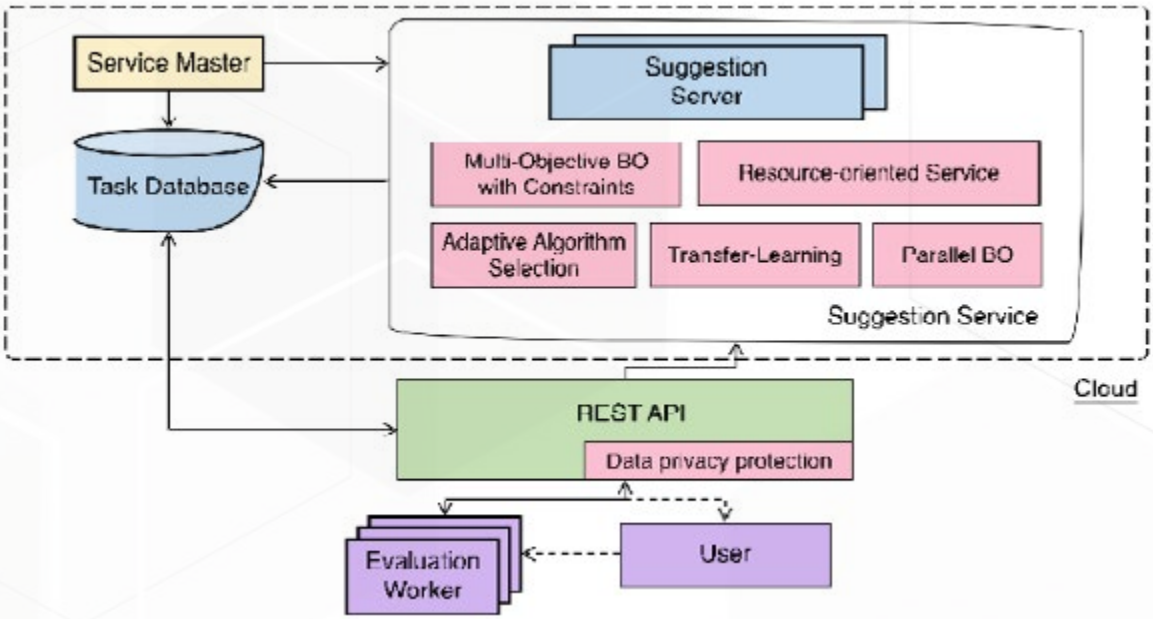
·扩展：在硬件上，通过显存优化和通信优化技术，支持上百节点的分布式部署；在模型上，支持千亿级参数规模的深度学习模型训练，已经在 Criteo、Open Graph Benchmark 等多个大规模基准数据集上进行了测试。

本项目由国家重点研发计划支持，相关成果已经发表了多篇 CCF-A 类高水平学术论文，目前项目已经开源在：https://github.com/PKU-DAIR/Hetu。

• OpenBox: 高效的通用黑盒优化系统　项目负责人：崔斌

近年来，人工智能与机器学习备受人们关注，越来越多的人使用机器学习模型解决实际问题，如数据分析与预测、人脸识别、商品推荐等。在应用机器学习模型的过程中，模型超参数的选择对模型性能有着重要影响，超参数优化问题成为了机器学习的重要挑战之一。超参数优化是典型的黑盒优化问题，对于优化目标不存在具体表达式或梯度信息，且验证代价较大，其目标是在有限的验证次数内，尽快找到全局最优点。

OpenBox 是我们针对黑盒优化设计的一套开源系统，以贝叶斯优化为基础，高效求解黑盒优化问题。其使用场景广泛，不仅支持传统的单目标黑盒优化 (如超参数优化)，还支持多目标优化、带约束条件优化、多种参数类型、 迁移学习、分布式并行验证、多精度优化等。OpenBox 不仅支持本地安装与优化调用，还提供在线优化服务，用 户可通过网页可视化监控并管理优化过程，也可以部署私有优化服务。 架构图如下：



OpenBox 使用 Python 编写，开源项目地址为 : https://github.com/PKU-DAIR/open-box。

可视分析中心

◆中心简介

可视分析中心由北京大学袁晓如研究员团队负责，承担大数据可视分析技术的研发应用，该中心针对大数据分析与应用的挑战，发展大数据可视化与可视分析的核心技术，形成完整的系列方法和工具集合；培养一流大数据可视化人才，推动我国大数据产业发展。

该中心的主要任务包括：

1、大规模图数据的表示、布局、交互技术，并实现超长时变图数据的可视化技术；

2、沉浸式大数据可视化分析环境技术，并研究基于大屏幕的多人合作交互技术；

3、针对高维大数据的可视分析技术等。

可视分析是对大数据进行分析的一种有效手段，日益受到重视。可视分析技术的目标是使数据分析过程透明化。它结合了可视化、人机交互和自动分析技术。在一个典型的可视分析流程中，自动分析的结果通过可视化展示给用户，用户通过人机交互技术评价、修改和改进自动分析模型，从而得到新的自动分析结果。通过这种方式，可视分析技术将人的经验智慧与机器的运算能力紧密地结合在一起。这其中，由人来定义分析任务和识别复杂的模式，由机器来存储和分析大量的数据。分析结果的可视化则成为人与机器协作的桥梁。

◆中心团队成员介绍

袁晓如

国家工程实验室常务副主任
兼可视分析中心主任
北京大学智能学院研究员
研究方向：可视化，可视分析，人机交互

汪国平

北京大学计算机学院教授
研究方向：计算机图形学，人机交互

刘　瑜

地球与空间科学学院副院长
研究方向：地理信息科学，时空大数据分析方法和应用

曾　钢

北京大学智能学院研究员
研究方向：计算机视觉和图形学，三维场景建模，基于图像或交互的建模和渲染，自动图像分割

裴玉茹

北京大学智能学院副教授
研究方向：计算机视觉，图像处理，数字几何处理

周　亮

北京大学健康医疗大数据国家研究院助理教授、副研究员
研究方向：科学可视化，信息可视化，视觉感知，健康数据可视化研究

周秉锋

北京大学王选计算机研究所图形图像研究室主任、研究员
研究方向：机器人运动学图形模拟、彩色图像处理及系统、数字图像半色调技术、数字几何处理、基于图像的建模与绘制、计算摄影学、虚拟现实

李　胜

北京大学计算机学院副教授
研究方向：计算机图形学、虚拟现实 / 增强现实、智能数字媒体技术

陈立翰

北京大学心理与认知科学学院副教授、脑与认知科学系副系主任
研究方向：为多感觉通道知觉、注意和认知神经科学

陈文广

北京大学信息管理系信息系统研究室主任、副教授
研究方向：包括数据管理与数据分析、智能人机交互、知识组织

陈毅松

北京大学计算机学院副教授
研究方向：计算机视觉、计算机图形学、图像和视频处理、模式识别、机器学习、数据挖掘和智能系统

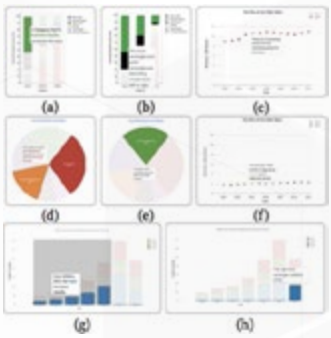
冯　洁

北京大学王选计算机研究所工程师
研究方向：基于图像的三维建模和真实感绘制、基于图像的表面材质建模、三维立体显示技术、光场显示技术、彩色图像处理、虚拟现实

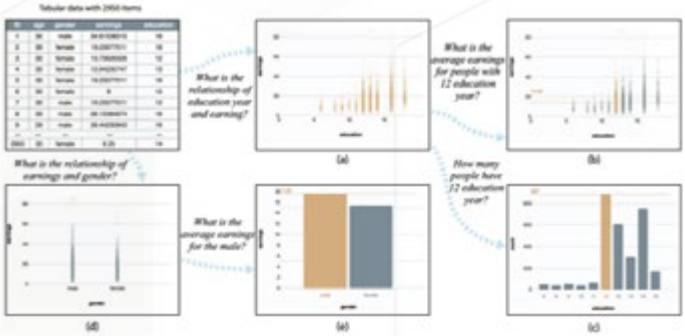
◆科研成果

•智能可视化语法、组件与系统　项目负责人：袁晓如

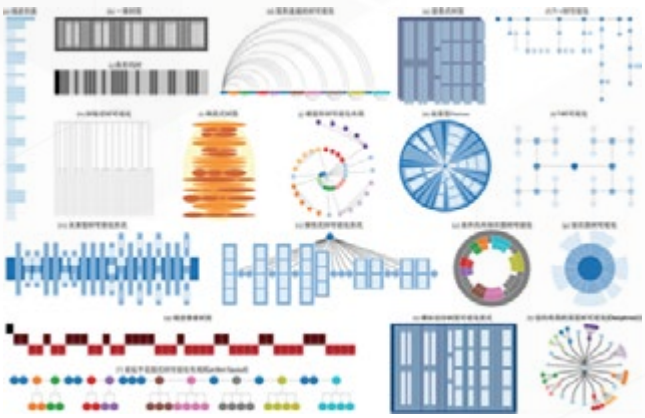
研发从数据到构型和交互反馈的全链条智能可视化体系。包括支持层次数据、多组分数据等的可视化领域描述性语法，设计开发可视化交互构建系统，支持用户无代码编写构建可视化形式。支持基于深度学习模型的自然语言与可视化交互，支持可视化自动问答、可视化描述和标题的智能生成，支持自动生成动画内容提升可视化与自然语言描述的理解。相关论文发表在 IEEE VIS, ACM CHI 和 IEEE PacificVis 等国际人机交互和可视化会议上。



▲ 针对图表和自然语言自动生成高亮动画效果



▲ 针对表格数据自动对自然语言生成可视化



▲ 基于 GoTree 描述性语法创建的树可视化形式部分集合

• 面向领域的可视化应用　项目负责人：袁晓如

新型冠状病毒肺炎疫情可视分析： 在 2020 年新冠疫情期间，针对疫情的发展变化，开展了有针对性的可视化工作，制作了全国各地的疫情变化晴雨表、各类疫情地图、疫情信息图表、相关社交媒体可视分析等可视化工具。其中，疫情变化晴雨表在 2020 年 2 月 7 日在腾讯微信医疗健康板块首次上线，应用于腾讯微信疫情数据动态可视化。相关工作还应用于中央广播电视总台（CCTV）视听新媒体中心，系统每天输出全国各省市新冠疫情数据可视化效果。

▲ 腾讯微信疫情数据动态可视化

面向人文领域的可视化工具集：在实验室团队长期可视化方法研究的积累基础上，针对社科人文的需要，从表格多属性、网络、时空等多角度全方位研发构建了面向人文领域的可视化工具集。当前工具集已经包括历史人物时间线、空间迁移地图、历史名人轨迹、网络关系地图以及标注众包等工具系统，在与文学、历史、考古、管理等领域合作中实践应用。构建的可视化项目成果在中国可视化与可视分析古籍创意竞赛、IEEE 数据可视分析竞赛等赛事中获奖，多次在国内国际艺术展中获邀参展。

▲ 人物迁居流向地图

▲ 人物迁居轨迹地图

▲ 历史名人时间线

▲ 众包标注系统

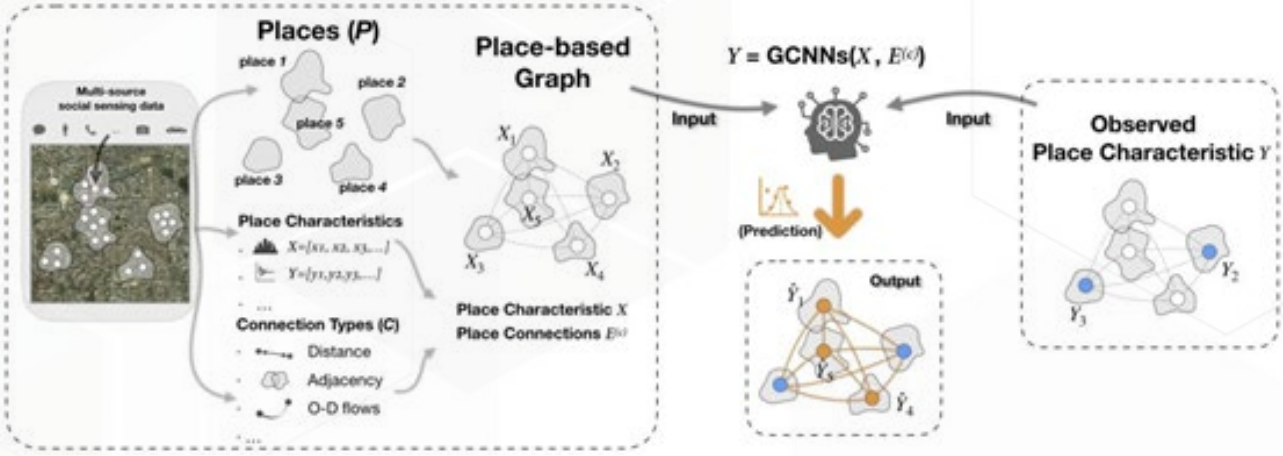
• 大规模空间场景建模与分析　项目负责人：汪国平、刘瑜、曾刚、周秉锋

基于分布式架构的大规模三维重建：使用基于图像和视频的三维重建技术进行针对 3D 模型的精细重建以及最新的分布式架构针对大规模城市场景的快速三维重建，自主研制了二三维重建系统 i23D，系统包括从数据采集到场景处理组织存储显示的全流程技术体系。系统工具被用于多个博物文物和智慧城市的三维场景快速重建中，并提供在线的三维场景重建服务。

▲ 航拍图像的大规模三维重建模型

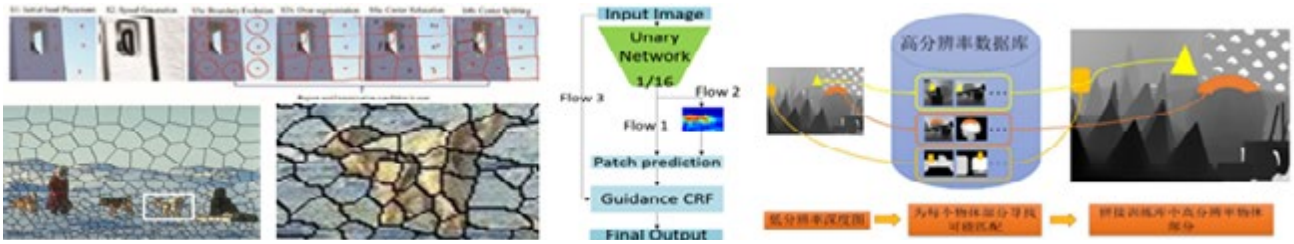
▲ 文物重建的三维模型

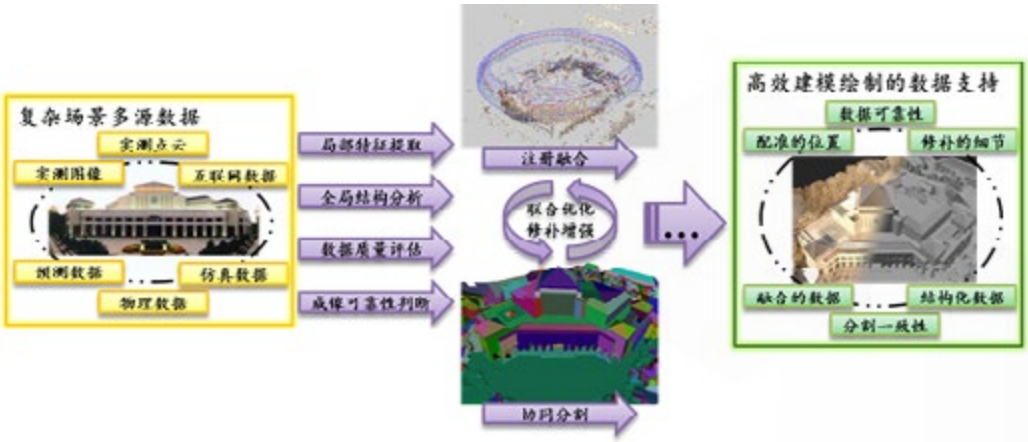
图卷积神经网络（GCNN）应用于不规则的地理空间分布的建模和推测：提出场所图 (placebased graph) 对复杂的场所属性和多源的场所连接进行形式化表达，通过训练图卷积神经网络 (graph convolutional neural network, GCNN) 实现地理空间回归，充分挖掘地理语境中场所属性的可预测性知识。 本研究为不规则的地理空间分布提供了一种全新的神经网络建模思路，可以用于辅助评估空间模式，衡量城市 社会经济发展与人类活动的关系，进而指导城市规划和居民生活。 论文于 2020 年 1 月正式发表于国际地理学旗舰刊物 Annals of the American Association of Geographers。



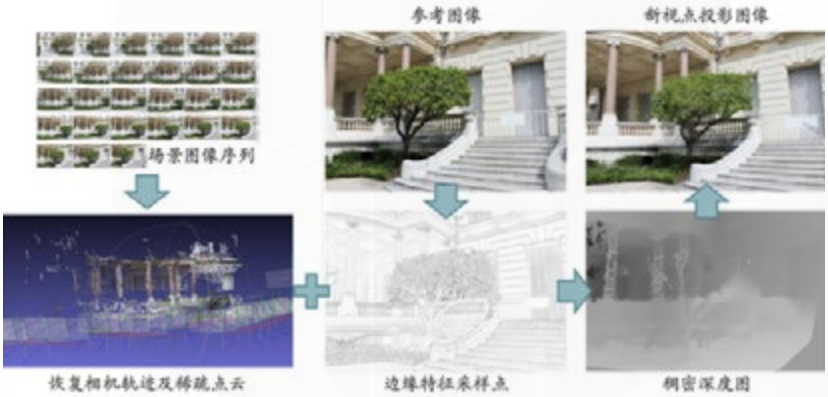
▲ 构造场所图以及使用图卷积神经网络推测未知场所属性的示意图

基于结构分析的场景建模：针对场景建模和绘制在鲁棒性、精度和计算效率方面的问题，从加强输入数据的类型、规模和质量入手，研究来自实测和社会的多源异构场景数据的结构化分析方法，引入仿真、预测、物理数据，挖掘在布局、构成、形变和互动等方面更多的结构约束，探索多源数据协同分割和语义识别方法。该尝试在统一的理论框架下表达数据融合、分割识别和几何重建的研究目标，有助于对问题形成更全面和完善的理解，并在理论上保障算法结果的最优性，通过多源数据协同分割识别与几何模型的联合优化提高结构分析和场景重建的一致性，更好地支持数据驱动智能化场景建模和绘制。





基于图像的三维场景高效建模与绘制：以通用相机获取的场景参考图像序列为输入，结合重要性采样等方法，充分利用参考图像中的颜色、纹理、边缘等特征，将稀疏的深度信息在图像平面上进行传播，从而快速、准确地恢复出稠密深度图像。在此基础上实现了在任意新视点下快速生成照片真实感新视图的方法。该方法仅需利用普通相机拍摄参考图像，对输入设备和数据具有更少的限制，应用范围也更广泛。

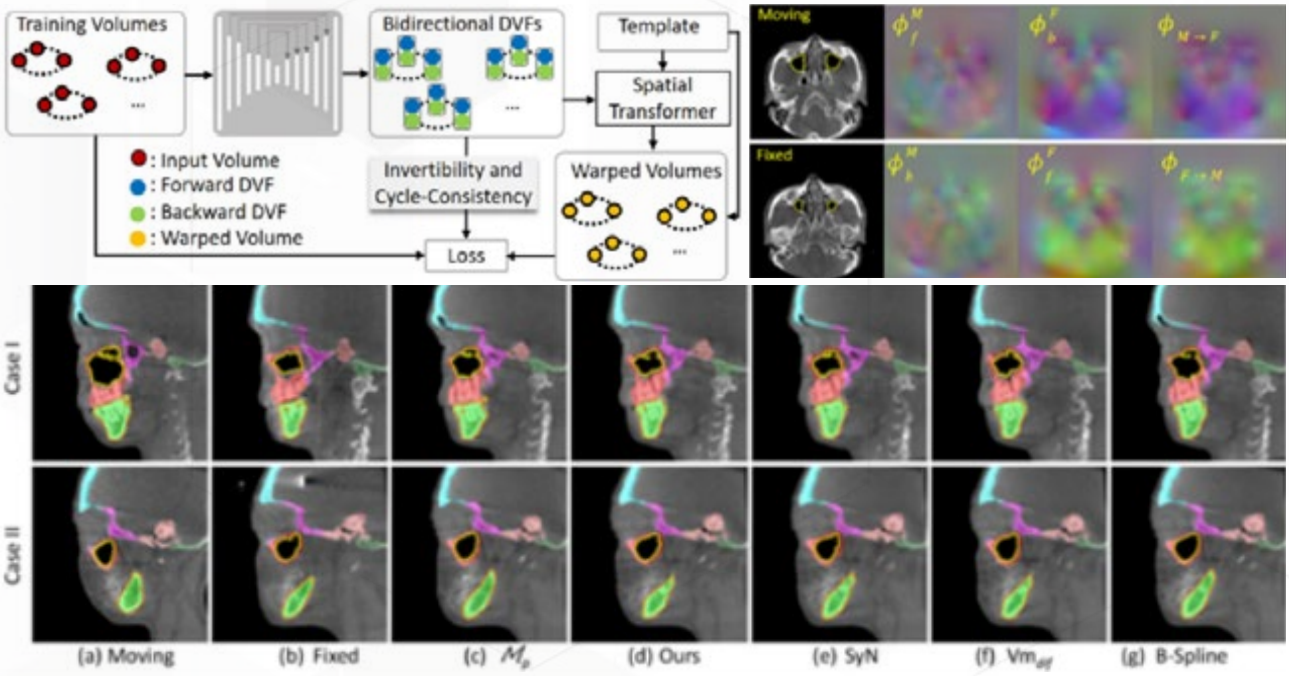


▲ 从参考图像序列进行三维场景新视点视图绘制流程

• 医疗大数据分析与可视化　项目负责人：裴玉茹、周亮

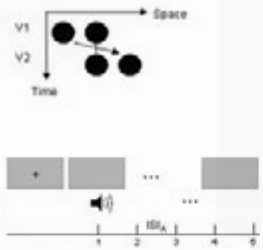
三维图像稠密体素对应与配准：引入全卷积神经网络预测输入图像和共享的模板图像之间的正向和反向变形场，利用任意的成对或者三元组图像，保证三维图像变形闭环一致对应。

数据驱动空间填充曲线脑部核磁影像集合数据可视化：提出数据驱动的空间填充曲线在尽量保留数据原始维度中的特征情况下降至一维，相比传统空间填充曲线（例如 Peano-Hilbert 曲线），减少降维中丢失的原始数据特征，使人通过对比一维线图或箱型图就能进行定量的集合数据可视化。



▲ 数据驱动空间填充曲线脑部核磁影像集合数据可视化

• “部分贝叶斯整合”理论模型　项目负责人：陈立翰



▲ 研究的刺激序列示意图

课题组构建了感觉通道信息权重动态变化的“部分贝叶斯整合”理论模型，挑战了此前众多研究所确立的“全”或“无”的整合规则，并首次揭示了人们利用“几何平均”的跨通道知觉组织规律。该研究为日常生活中、项目所经历的复杂的视听交互等场景，奠定了崭新的跨通道“平均加工”知觉组织原则。研究针对实验心理学领域中的跨感觉通道信息的加工机制和计算机理，提供了全新的视角和解释，补充了普通心理学和实验心理学关于多感觉通道的“知觉”组织理论。研究发表在 Journal of Experimental Psychology: General（2018）上。

知识集成和智能决策中心

◆中心简介

知识集成和智能决策中心由北京大学王选计算机研究所邹磊教授牵头，浙江大学、武汉大学、奇虎 360、滴滴出行等共同组成，承担大数据智能知识管理与决策支持技术的研发应用。

该中心的目标是开展基于大数据与机器学习的新一代智能决策系统等技术的研发和工程化。中心将从以下方面开展科研和应用落地工作：

（1）知识集成和知识图谱平台技术；

（2）学科领域知识图谱构建及平台应用；

（3）知识图谱和智能决策的工业应用。

面向大数据的知识发现、管理和决策支持服务平台是大数据分析流程中关键的一环。其中包括四个方面的主要内容：一是面向多源异构数据源的知识图谱构建和融合，将不同的数据源的数据进行有效的整合，形成面向特定领域或者开放领域的完备的知识图谱；二是海量知识数据的存储和查询等数据管理问题，一个面向海量知识数据的高效的知识图谱数据管理系统是提供知识服务和支撑上层决策的基础；三是智能知识检索和分析，为用户提供一种基于可视化技术的、交互式的知识数据访问和分析的平台接口，方便普通用户使用知识图谱；四是在上述知识图谱的构建和分析的基础上，针对复杂的问题提出定性定量综合集成建模体系，从而有效支持相关决策。

◆中心团队成员介绍

邹　磊

知识集成和智能决策中心主任，北京大学王选计算机研究所教授

研究方向：知识图谱、图数据库、大数据分析系统

杨晓光

国家工程实验室副主任兼知识发现与智能决策研究中心主任，中国科学院数学与系统科学研究院研究员，中国科学院系统科学研究所副所长

研究方向：金融风险管理，宏观经济分析，复杂网络，博弈论等

刘　颖

中国科学院大学经济与管理学院副教授

研究方向：商业数据分析、互联网经济等

汤　帜

北京大学王选计算机研究所所长、研究员

研究方向：文档分析与理解、信息检索

冯岩松

北京大学王选计算机研究所副教授
研究方向：自然语言处理、信息抽取

孙　栩

北京大学信息科学技术学院研究员
研究方向：自然语言处理、机器学习、深度学习

陈华钧

浙江大学计算机科学与技术学院教授
研究方向：知识图谱、自然语言处理、大数据系统

洪　亮

武汉大学信息管理学院教授
研究方向：大数据分析、知识图谱、大数据安全与保密、人工智能

冷大炜

奇虎 360 资深算法专家
研究方向：计算机视觉、深度学习、多模态预训练模型

李　奘

滴滴出行首席算法工程师
研究方向：数据挖掘、知识图谱、增长算法

熊　超

滴滴出行专家算法工程师、智能客服算法负责人
研究方向：人机交互、多轮对话、智能客服、知识图谱

◆科研成果

• 图数据库系统 gStore

核心产品 gStore 是由研发团队研发的面向 RDF 知识图谱的开源图数据库系统。邹磊教授及其团队在国家自然科学基金重点项目，科技部重点研发计划，军委科技委国防创新特区项目的持续资助下，在注重学术创新的同时，坚持自主研发和工程系统研发，历经十余年的努力，已经形成了面向百亿三元组规模的高效图数据库系统，代码完全自主可控，拥有“论文 - 软件著作权 - 专利”完整自主知识产权体系，并于 2018 年 2 月获得教育部自然科学二等奖（“大规模图结构数据管理”）。gStore 系统的开源版本近一年以来被下载超过一万次。目前 gStore 系统目前已经完成国产自主可控软硬件适配工作，获得的适配证明表明 gStore 支持麒麟、UOS 操作系统，支持飞腾 FT-1500, FT2000+、鲲鹏、海光、兆芯申威等国产处理器。2021 年 6 月，gStore 图数据库通过中国信通院图数据库基础能力专项评测。2022 年 8 月，发布 gStore 1.0 版本。

项目主页：http://www.gstore.cn

GitHub 开源主页：https://github.com/pkumod/gStore/

▲ gStore Workbench

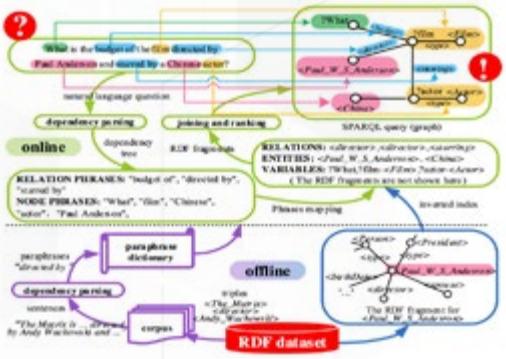
▲ gStore 商业版本应用到多个企事业单位的核心知识图谱业务

▲ gStore 适配证明及产品测评证书

• 自然语言问答系统 gAnswer

gAnswer 是面向知识图谱的自然语言问答系统，采用“语义查询图”来表达自然语言问答的语义，并利用子图匹配的方法来解决自然语言理解中的消歧问题，同时查询知识库得到答案。发表包括 SIGMOD、TKDE 和 EMNLP 在内多篇学术论文；申请了多项相关专利。gAnswer 系统整体开源，应用到知识图谱问答系统项目中。其评测成绩优异，在欧盟举办的 QALD 9 的面向知识图谱的自然语言问答系统的评测中排名第一。

GitHub 开源主页：https://github.com/pkumod/gAnswer



▲ gAnswer 系统框架图

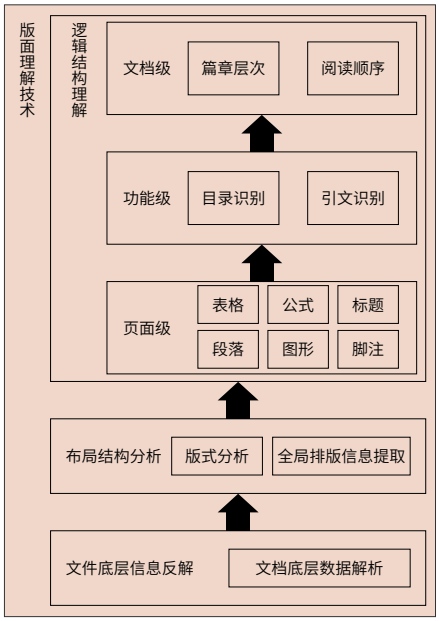
▲ gAnswer Workbench

• 知识图谱构建系统 gBuilder

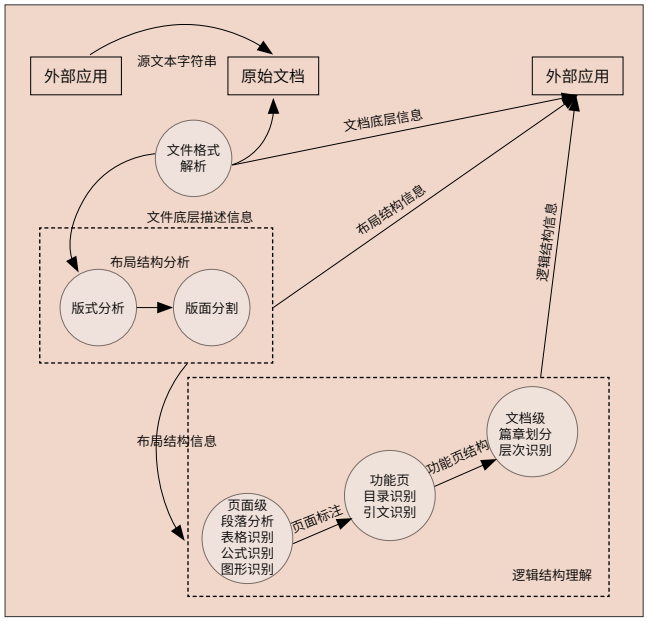
gBuilder 是知识图谱自动化构建平台，解决知识图谱生命周期中一个基础且重难点问题即如何自动化构建知识图谱。通过结合 NLP 技术、机器学习、人工智能、知识图谱、图数据库等众多技术，打造的一个针对结构化数据和非结构化数据的知识图谱自动化构建平台，实现数据向知识的转化。 提供 schema 可视化设计，通过拖拽的方式进行 schema 设计，无需编写代码。通过映射方式直连关系型数据库，结构化数据无损转化为 RDF 三元组数据。非结构化数据主要通过深度学习模型流水线构造的方式来对非结构化数据的实体、属性关系进行学习与识别，实现数据和知识的自动化转化。

• 版面理解技术

版面理解技术应用于知识图谱的知识提取阶段，协助完成从非结构化知识库中提取出实体、关系、属性等知识要素，属于知识提取的基础支撑技术。版面理解技术基于深度学习，解决从结构检索到语义理解的难题，实现主题级和语义级的检索。利用深度学习网络优良的序列处理能力学习数据元之间的嵌套与层次关系，以得到更优的结构识别结果。版面理解技术通过技术成果转化已经成功地应用到若干知名企业的核心产品当中，如方正阿帕比的 Apabi Maker 和方正电子公司的方正慧云中学生题库搜索与推荐系统，明显提升了产品的功能效果和产品在市场上的竞争力。

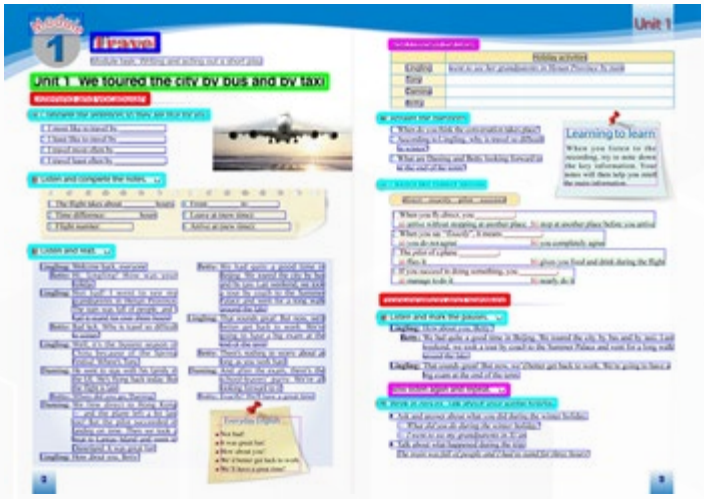


▲ 版面理解技术模块架构图



▲ 版面理解技术业务流程图

外语教学与研究出版社有限责任公司的出版产品主要为外语教材，需要给教学机构配套开发教学辅助系统 / 学生学习系统和点读功能，以往系统中出版内容需要在付印 PDF 文件中逐一拷贝、点读产品需要逐句圈选链接音视频文件。为了解决流程低效率和高错误率的问题，版面理解技术产品需要根据外研社英语基础教材和英语高等教材进行版面理解开发，正确识别出版物各级标题、正文（按句）并输出页面信息、结构块标签和位置信息。



▲ 版面理解技术信息提取结果

在该项目中，通过使用版面理解技术，多页教材在以分钟为单位的时间里即可完成版面识别、内容块标签、位置信息标注工作，联合麦片编辑系统或外研社基础数字开发系统即形成不同区域的热区并按区拾取页面内容，很大程度地减少了编辑人员手工操作的工作量、提高了工作效率和工作准确度。

• PKUBase

在知识获取方面研发了大规模开放域中文结构化知识库 。PKUBase 通过一系列自然语言处理手段及机器学习方法，从半结构化和非结构化可信文本，例如维基百科、百度百科、Freebase 等资源中构建的海量中文结构化知识资源，其中包含标准中文类别体系框架、近千万级中文实体、超过 6000 万中文知识条目。PKUBase 是以 RDF（资源描述框架）三元组标准数据形式进行组织和管理的，支持常见的标准图数据管理引擎。用户可以通过编写满足 W3C 的 SPARQL 1.0 标准的结构化查询语句对 PKUBase 进行知识检索。

• OpenKG

国家工程实验室积极参与和支持了中文领域开放知识图谱 OpenKG 的工作，典型的包括为 OpenKG 提高基于 gStore 的图谱查询引擎，开放并开源了 gAnswer 等开源工具等，得到广泛的关注。

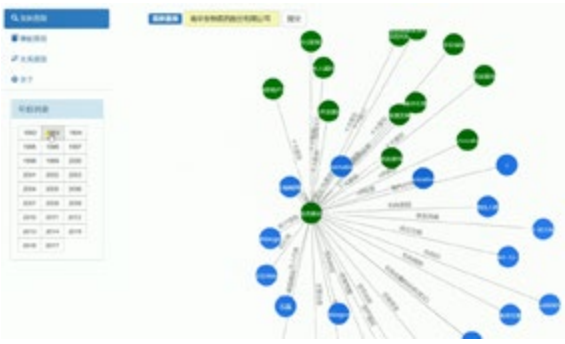
OpenKG（http://openkg.cn/）是由中国中文信息学会语言与知识计算专业委员会倡导的开放知识图谱项目，其资源库包含来自学校和公司的开放图谱数据及工具集，它的发起致力于促进中文知识图谱的开放与互联以及促进知识图谱和语义技术的普及和广泛应用。该项目在宣传开放之于 KG 价值及理念的同时，也在不断搜集、整理和聚集国内外主要的开放知识库，并利用自身的资源推动开放域的链接知识图谱的建设。当前 OpenKG 项目的工作可分为 OpenBase 众包平台搭建，kg4ai 知识图谱的构建，cnschema 标准，大规模百科知识对齐等四条线。

以 OpenBase 为例，它是一个基于众包的知识构建与数据集成平台，通过人工和自动化相结合的方式解决图谱构建、标注和迭代的难题。用户可以将结构化的三元组数据上传到平台，使得具有领域知识的其他用户可以对三元组进行审核，最终由验收者进行验收，确保图谱数据更准确。任意用户可浏览已上传的图谱数据。

• 知融平台

由武汉大学、北京大学与深圳证券信息有限公司合作开发，基于 gStore 数据库研发了面向资本市场大规模图数据管理关键技术，融合了全量的工商注册企业数据以及商业银行股权结构数据，对十亿级别三元组的大规模知识图谱实现了准确、高效、鲁棒的查询与管理，实现了对资本市场的实体、实体关系和实体属性的知识关联查询，多层股权结构的发现与揭示，以及交叉持股，资本系等蕴含潜在金融风险的股权结构的准确发现和深度分析。

▲ 知融体系架构中的图数据库系统“gStore”

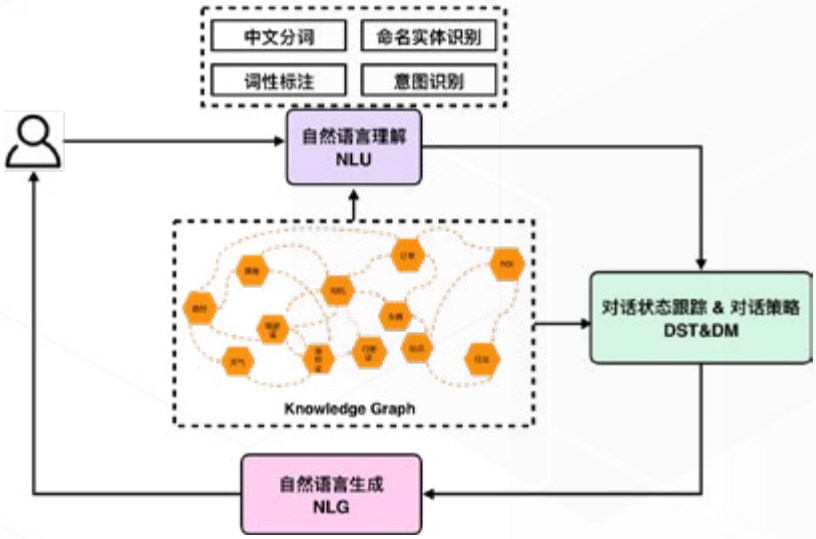


▲ 多层股权查询示例

• 滴滴多轮对话机器人

多轮对话机器人是滴滴智能客服系统的重要组成部分，通过多轮次的人机对话不断地获取司机 / 乘客的意图和相关业务信息，从而快速、有效地解决司机 / 乘客的问题。和大多数任务型对话系统不同的是，滴滴多轮对话机器人引入了知识图谱（KG）辅助识别用户意图和维护对话状态，借助图谱强大的知识表达能力和知识推理能力，能高效准确地解答用户的问题。目前滴滴智能客服可以有效降低客服的人工进线量，并更好地帮助人工客服提高效率。

多轮对话机器人主要应用于滴滴的智能客服业务，当用户（司机 / 乘客）进线咨询相关业务问题时，意图理解模块首先会对用户的问题进行关键信息抽取，并基于知识图谱的知识推理能力精准获取用户意图；在对话模块，通过基于知识图谱的分析推理找到用户问题的答案；最后基于自然语言生成模块生成相应的回答文本返回给用户，解答用户的问题。



大数据高性能分析处理中心

◆ 中心简介

大数据高性能分析处理中心由中山大学负责承建，由王甲海教授牵头负责。

该中心研究基于天河二号的大数据系统支撑平台与产业化应用。随着大数据相关的技术和产业不断发展，数据量越来越大，应用也越来越广泛，需要有超大规模软硬件作为基础来对大数据分析与管理进行支撑。而目前国内的大数据相关应用，包括在公共管理、制造领域、电力领域、电子商务领域、智能交通领域、金融领域、科学领域、教育领域、健康医疗领域、电信领域、文化领域等，这些大数据应用在超大规模数据的处理上普遍专注的是针对领域中的特定需求的算法和技术的研发，缺乏系统软件层面的支撑。这大大制约了大数据技术的发展。基于天河二号的大数据系统支撑平台的建设与产业化可以很好地满足这方面的需求，提供可用于不同领域的系统软件研发环境和平台，将为各领域的大数据技术发展提供直接的支撑和推动作用，促进区域与国家的大数据产业的快速发展，支撑战略性新兴产业的培育和发展。

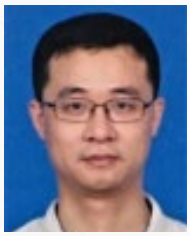
◆ 中心团队成员介绍

中心负责人 王甲海

毕业于富山大学（日本），工学博士，中山大学计算机学院教授 / 博士生导师，大数据与计算智能研究所所长，大数据分析与应用技术国家工程实验室大数据高性能分析处理中心主任，研究兴趣包括人工智能（智能优化与学习），大数据分析与挖掘。提出了多种先进智能算法，有效求解复杂优化问题，推动智能优化算法在智慧物流等领域的实际应用；对深度学习有深入研究，并应用于社交网络分析和自然语言处理，研制了舆情分析等实用算法和系统；利用数据驱动的深度学习方法，研究面向城市感知的群智涌现机理与计算方法，在智慧城市中得到应用验证。在 IEEE Transactions，AAAI 等国际知名期刊和会议上发表系列文章，主持 / 完成 4 个国家自然科学基金和多个省部级基金，主持科技部“新一代人工智能”重大项目课题 1 项，主持完成军队装备预研项目 1 项。参与国家自然科学基金联合基金重点项目 1 项。 2011 年获广州市首届珠江科技新星人才专项资助。IEEE 高级会员，CCF 杰出会员。

郑伟诗

国家工程实验室副主任、中山大学计算机学院副院长



江　颖

中山大学计算机学院教授

研究方向：医学影像、积分方程数值算法

赖韩江

中山大学计算机学院副教授

研究方向：大数据分析、信息检索

刘　军

浪潮集团 AI&HPC 总经理

研究方向：人工智能与高性能计算

刘玉葆

中山大学计算机学院教授
研究方向：数据库与数据挖掘

钱德沛

中国科学院院士，北京航空航天大学教授，中国计算机学会会士，曾任国家工程实验室副主任
研究方向：高性能计算机体系结构与实现技术、并行与分布计算、计算机网络管理与性能测量等

权小军

中山大学计算机学院教授
研究方向：数据挖掘、自然语言处理

饶洋辉

中山大学计算机学院副教授
研究方向：情感计算、主题建模、事件检测

苏勤亮

中山大学计算机学院副教授
研究方向：机器学习、贝叶斯分析、自然语言处理

吴　迪

中山大学计算机学院教授、副院长
研究方向：新型网络计算系统、大数据分析与应用等

吴维刚

中山大学计算机学院教授、副院长
研究方向：云计算与边缘计算、大数据与高性能计算、深度学习、区块链

张海樟

中山大学数学学院教授
研究方向：机器学习、稀疏逼近、应用分析理论及算法

陈志广

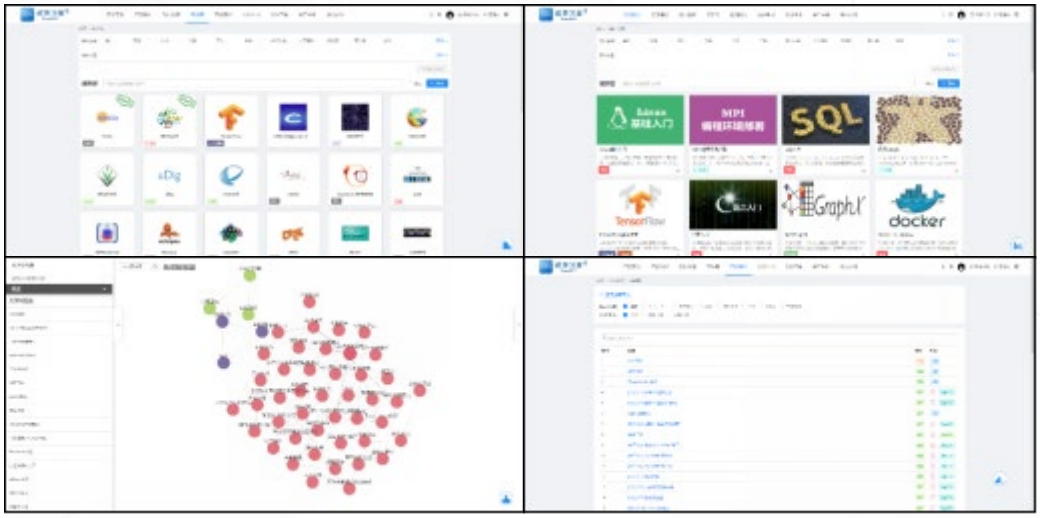
中山大学计算机学院副教授
研究方向：大数据存储与处理、并行与分布式计算

◆科研成果

• 超算教育实践平台“超算习堂”

由中山大学牵头，清华大学、北京大学、中国科技大学等多所国内一流高校参与共同合作研发了国内规模最大的超算教育实践平台“超算习堂”，获得了国家重点研发计划重点专项支持，多家超算中心为平台提供机时资源与技术支持。该平台致力于提升广大学生在超算领域上的科学素养与实践能力，平台涵盖课程教育、在线编程、在线实训、案例展示等诸多内容，包括大量编程实战题、领域经典案例以及编程竞赛平台的进阶挑战。

超算习堂是一个国家级开放的超算教育实践平台，提供超算实践教育课程、软件、案例、题库等多种不同类型的教育资源。为满足用户的不同教育需求，构建了在线课程、在线实训、交互协作平台、竞赛平台、软件库等多种核心功能模块。平台最终部署在国家高性能计算环境中，提供了多层次、多样化的超算教育实践服务，让学生在实践过程中更为深刻地理解整个过程，熟练掌握运用并行计算思维求解问题的方法，该平台为我国超算教育人才培养提供了有力保障。



平台支持十多所国内知名大学 30 余门高性能相关课程的实践练习，为用户提供了数万次并行程序评测；教育实践平台为一系列超算大赛（如 SC、ASC、PAC、RDMA 等）的参赛选手提供培训和机时支持；所培训的参赛学生获得一系列重要奖项（包括 SC 总分冠军、ASC ePrize 计算挑战奖、RDMA 一等奖、PAC 优化金奖等）。

• 面向生物多组学分析和药物设计领域的高性能计算应用平台

国家超级计算广州中心面向生物多组学分析和药物设计领域行业应用需求，研制和建设简易实用、大规模高效、稳定运行的高性能计算应用平台。突破领域应用跨平台超算计算架构、多模态大数据综合管理与分析处理、大规模跨尺度关联融合、高通量任务高效计算等关键技术，开发和集成主流数据库、应用软件、且具有可扩展能力，提供应用所需的数据资源、软件资源、计算资源、简易交互界面等全方位服务功能，有效支撑包括高等院校、科研机构、行业用户的大规模多组学数据分析、药物筛选和设计等应用计算需求，实现疾病辅助诊断、药物靶标发现、新型药物设计等领域业务应用，加快国产超算在生物医药应用领域的推广和普及，提高我国超算应用水平。

生物平台的总体目标是基于多个重要生物组学数据库的镜像，对这些数据建立统一的数据标准；构建标准组学数据分析模块，将不同层次多组学数据整合在一起，实现数据的融合分析；基于现有的知识库，结合人工智能算法、对这些生物组学数据进行深度挖掘和注释。

本平台实现了对不同层次数据的知识库进行整合，统一实体和关系命名，建立知识图谱。整合的数据库包括基因突变（COSMIC、Clinvar、HGMD 等）、转录组（eQTL、Roadmap）、表观遗传（EOCODE、FANTOM、 Roadmap Epigenomics）、非编码 RNA(NONCODE 等)、药物基因组学（LINCS、PharmGKB、TTD）、蛋白质结构（PDB）和功能（GO、Uniprot、Reactome、KEGG）, 蛋白质与蛋白质 (StringDB)、小分子（BindingDB、ChEMBL、Drugbank）的相互作用，以及基因与疾病（CTD、 OMIM、SIDER）关系等。

平台整合了各类公共癌症组学数据集。前期研究中项目已经下载了 TCGA 中 33 种癌症的基因组、转录组、DNA 甲基化以及病理影像组数据，并获得了 TCGA 下载完整序列数据的完全授权。接下来，将整合 GEO 中其它分散的相关癌症数据集合，也将进一步申请获得 ICGC 及 UK Biobank 的授权。联合军事科学院布置了国际蛋白质组学共享平台 iProX 的镜像，对外提供数据和计算服务。

通过与医院科室合作，整合医院的癌症组学数据资源。前期，项目组与中山大学附属第一医院团队合作，获得了 768 名宫颈癌患者的全基因组测序数据和病理、CT 等临床影像组学数据；与中山附二医院团队组合作，获得了 200 例膀胱癌患者的临床影像组、癌组织和癌旁组织的外显子组与转录组数据。未来将继续利用中山大学十家顶级附属医院的优势，通过合作整合更多的癌症组学数据。

在数据分析处理方面，针对基因组的驱动，项目组将根据数据库中定义的基因功能区域、基因、基因模块、基因通路划分，结合知识图谱中的基因突变对应的致病预测分值，对每个模块中出现突变的分值进行加和，通过比较癌症和千人基因组、亚洲十万人基因组、以及自有的组学数据，发现新的癌症驱动模式。针对转录组等其它组学数据，包括转录组（含 miRNA、lncRNA）、单细胞测序、蛋白质组、药物基因组等组学类型，根据知识图谱建立分子间的关系图，利用网络分析或图卷积学习挖掘关键分子或网络模块。

除此之外，围绕着生物数据分析流程平台建设，项目组在生物大分子预测、组学数据分析、多组学数据关联和融合、数据挖掘算法、知识图谱构建和应用，以及云平台建设等方面已经奠定了重要的研究基础，为平台构建奠定了技术基础，相应的研究和技术。

智慧交通中心

◆ 中心简介

智慧交通中心由北京嘀嘀无限科技发展有限公司负责承建。该中心依托全球最大的共享出行平台，致力推动智慧交通相关科学研究合作，促进人才培养和学术、产业交流，整合大数据资源、平台资源和技术能力，增强交通大数据和云平台赋能智慧交通产业实践的落地工作，实现业务问题解决与科技研发的相互驱动。

该中心的主要任务包括：

1、继续开发和优化不同场景下的交通大数据收集、存储、分析和呈现系统，推动“滴禹”综合交通大数据平台进一步优化迭代；

2、探索交通大数据赋能城市交通，在信号灯配时优化、拥堵研判、车载视频应用、城市交通安全治理等方面开展广泛的行业应用实践；

3、以实际业务催生学术研究和技术研发，开发新产品、新应用、创新方法等，在国内外期刊或会议上发表学术论文、讨论，申请专利等；

4、以季度为单位制作和发布我国主要城市交通运行报告，真实反映城市交通情况，在宏观监测、城市治理和规划建设等多方面做出贡献；

5、积极联系实验室及其他分中心，协同合作，共创共赢。

该中心将重点利用出行大数据，建设综合交通信息系统，探索建设包括共享出行在内的一站式的出行解决方案与大数据平台，提高多样化的交通方式协同工作效率。

◆ 中心团队成员介绍

杨　毅

国家工程实验室常务理事，滴滴出行智能平台事业群总经理
研究方向：智慧交通、大数据、人工智能

吴国斌

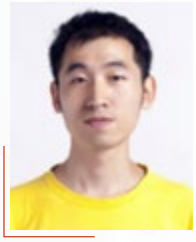
国家工程实验室副主任，滴滴出行科技生态与发展部总监
研究方向：人工智能、大数据

王　琤

滴滴杰出工程师
研究方向：机器学习、图像识别、人机交互

孙伟力

智慧交通中心主任，滴滴 AIoT 高级产品专家
研究方向：信号优化、视频识别、AIoT 技术



张凌宇

智慧交通中心副主任，滴滴出行 AI Labs 高级研究员
研究方向：人工智能、大数据、机器学习、数据挖掘

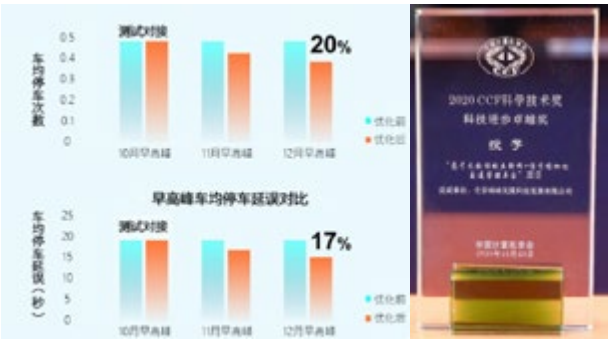
梁　舰

智慧交通中心副主任，滴滴智慧交通高级技术专家
研究方向：智慧交通、人工智能

◆科研成果

•信号优化实现“评估 - 诊断 - 优化”的闭环迭代技术

滴滴基于自身大数据和技术优势，自主研发了互联网动态信号控制优化系统，能根据车辆轨迹数据对交通状态进行测量，通过相应的算法进行分析，将分析的结果反馈给信号灯，协调各个路口红绿灯的时间，实现对大规模区域内信号系统的问题进行实时评估诊断、系统性和动态地调整优化，实现“评估 - 诊断 - 优化”的闭环迭代，有效帮助城市缓解拥堵。基于大数据的互联网 + 信号精细化交通管理平台，成果获得 2020CCF 科学技术奖科技进步卓越奖 。



•基于桔视的 AIoT 智能驾驶感知平台

滴滴基于自主研发的桔视智能车载设备，集成了车外行车记录、行程中车内安全录像和安全求助等功能，并构建了一套内外联动的智能化驾驶安全系统。不仅能对车外的前向碰撞、前车急刹、车道偏离等场景及时预警，还能通过对内的驾驶安全检测系统，对司机的开车分心、玩手机等不良行为进行提醒，在司机进入疲劳状态前，通过语音提醒司机注意行车通风、及时休息，避免疲劳驾驶，让出行更安全。

目前这一驾驶安全解决方案已逐步从网约车拓展至货运业务，持续帮助提升货运司机服务水平和驾驶安全，保护司乘双方权益。

在吴文俊人工智能科学技术奖十周年颁奖盛典暨 2020 中国人工智能产业年会上，滴滴的“基于桔视的 AIoT 智能驾驶感知平台”获评第十届吴文俊人工智能科技进步奖企业技术创新工程项目。

•智慧交通大脑形成“一云四中枢”的闭环管理体系

智慧交通大脑通过结合城市交通信息、基础路网信息、交通设施信息建设了“一云四中枢”的闭环管理体系，即交通云和感知、分析、决策、评估四个中枢。感知中枢可以从交通云获取实时信息与历史规律，主动识别异常交通状态、交通事故信息等；分析中枢负责实时推演和预测交通运行状态，为不良交通状况找到“病因”；决策中枢则针对各类交通事件的研判结果，智能生成多套优化预案；评估中枢负责对各预案的预期效果进行评估，确定最优方案，并且可以追踪方案实施效果，智能识别方案问题，最终将结果反馈至基础云平台。

滴滴为济南交通大脑建设的大数据平台作为交通大脑的基础设施，负责了数据从采集，处理，存储到数据服务的全链路建设。济南交警支队通过发挥“济南交通大脑”中“车研网、渣土车通行备案、动态违法预警打击”三个功能模块作用，通过卫星定位轨迹追踪、电子围栏、图片二次识别以及人脸识别等技术手段，实现对违法渣土车精准打击，获“山东公安科学技术进步一等奖”。

多媒体大数据分析中心

◆ 中心简介

多媒体大数据分析中心由北京奇虎科技有限公司负责承建，由 360 人工智能研究院负责。

针对多媒体数据的处理问题，中心将研究对多源异构数据进行合理且有效的综合分析，通过卷积神经网络、循环神经网络、生成对抗网络等多种方法进行建模。多媒体数据处理的计算复杂度较高，在大数据处理场景下这个问题尤为突出，为此，中心将重点研发在资源受限场景下的加速技术，如模型量化，模型结构自优化等。

该中心的主要任务包括：

1、发展多媒体处理技术，包括但不限于：图像分类、图像语义分析、图像快速索引、人物理解、场景理解、视频分析等。

2、研究资源受限场景下的模型优化，重点关注如何在保证精度的前提下进一步提升模型效率，推动工业化部署。

3、进一步加强与相关实验室、中心的合作，保持技术的先进性。

4、将相关技术应用于大规模生产环境，在实际中检验并优化相关技术级理论。

◆ 中心团队成员介绍

殷宇辉
国家工程实验室副主任
奇虎 360 副总裁

李　娜
奇虎 360 政企安全副总裁

冷大炜
奇虎 360 资深算法专家
研究方向：计算机视觉、深度学习、多模态预训练模型

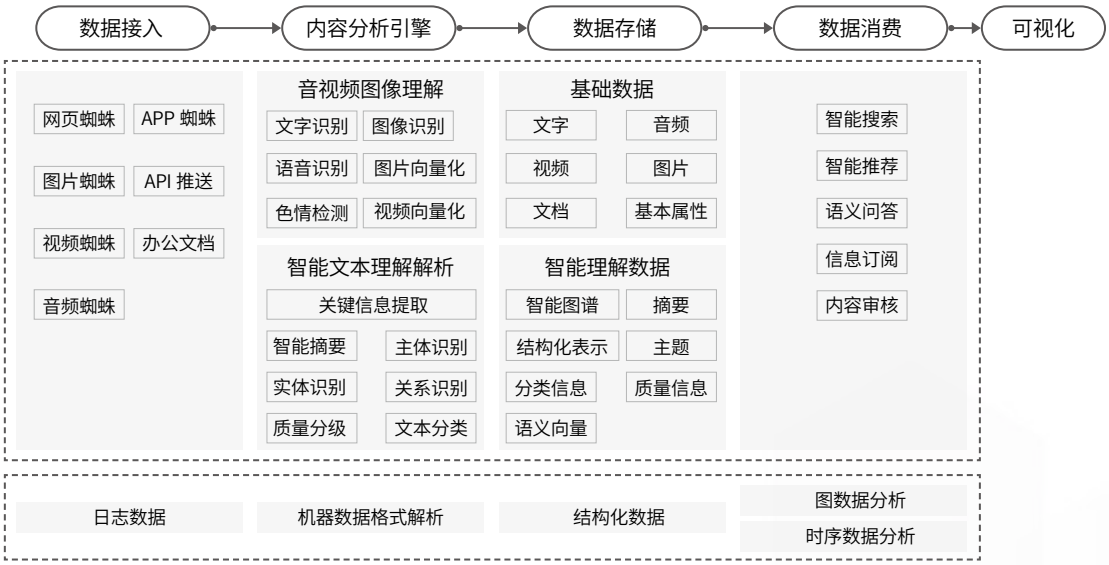
张　锋
奇虎 360 资深算法专家
研究方向：NLP、深度学习、知识图谱

刘焕勇
奇虎 360 算法专家
研究方向：知识图谱、知识表示、深度学习

◆科研成果

• 智能内容服务平台

智能内容服务平台是底层内容统一存储和统一管理的基础平台，提供 API、连接器等方式服务于各种应用场景。提供 AI 驱动的非结构化数据处理工作流，提供内容理解、信息提取、知识生成、智能搜索为核心的基础能力为业务赋能。



• 可视化大数据分析与技术应用

主要工作

·研发人形和人脸检测识别，跌倒和攀爬等危险行为检测、车辆和路标检测等技术，应用在 360 智能门铃和摄像头、360 行车记录仪和 360 家庭安全大脑产品中。

·研发 2D 地图构建、3D 环境建模、自动驾驶运动引擎等技术，应用在 360 扫地机器人中。

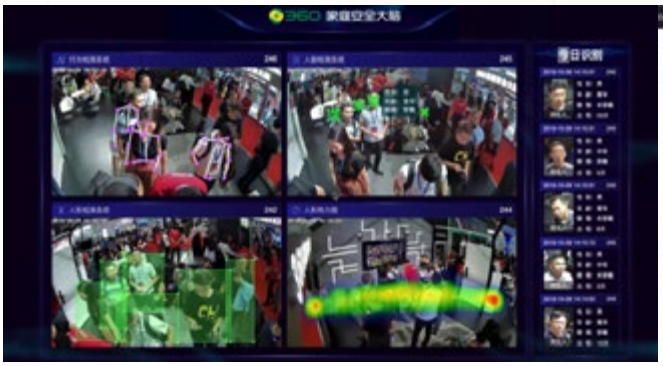
·研发语音识别、语义理解、对话管理、决策和推荐算法等技术，应用在智能外呼机器人系统、智能门禁对话系统、360 儿童手表、360 音箱、360 行车记录仪等产品中。

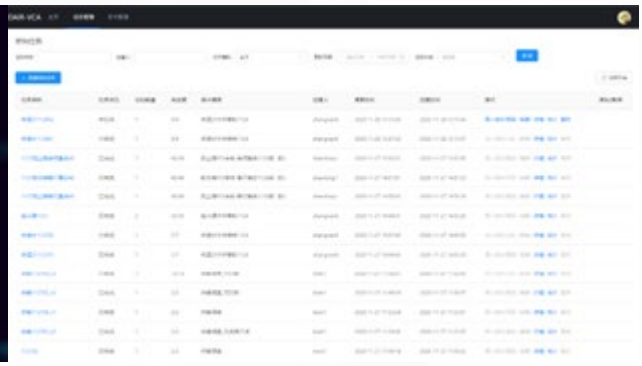
典型案例

·远华实业：智能安全监管系统 - 云端管理

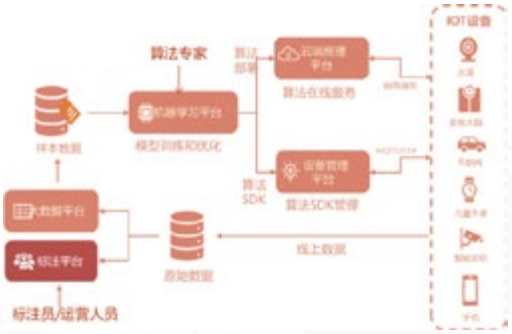
·联想：Lenovo X1 扫地机器人海外产品

·朝阳区政府：疫情防控智能外呼助手





• 端云融合的 AIOT 主动学习平台



建立包含：数据获取、数据标注、分布式模型训练、嵌入式端侧模型移植、模型部署、OTA 升级等组件的 AIOT 主动学习平台

优势

·采用端云分层计算方式，可适应低端设备，可节约云端成本；

·相对传统的人工训练 SDK 模式，理想状态下，在无需算法工程师参与的情况下，系统在人工标注 / 运营的支持下，可自主学习，自我进化；

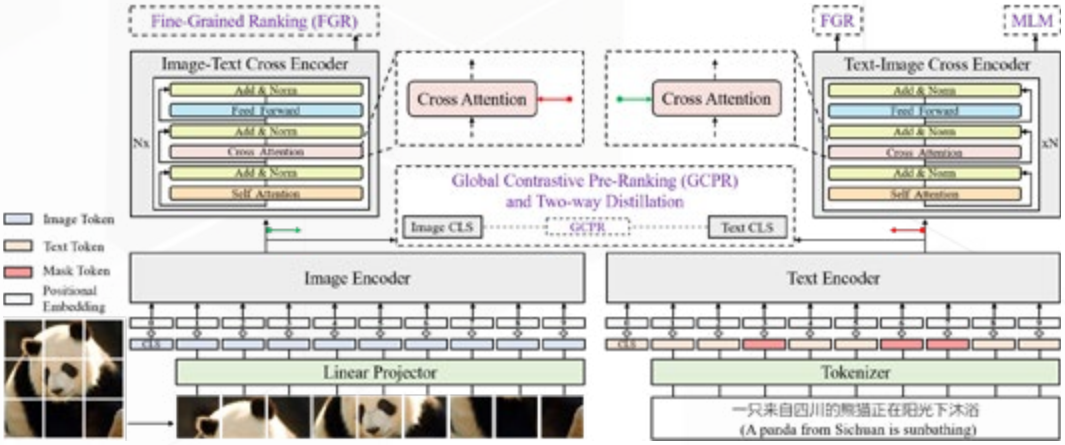
·场景数据挖掘和积累，将业内领先的 IOT 设备保有量转化为数据和场景的壁垒，可打造比 AI 创业公司更加先进的场景化 AI 技术；

·主动模型训练，无干预或弱干预，对于已有任务可有效节约维护人力，对于新增 AI 需求，具有快速迁移能力。

• 知识图谱应用平台

▲ 一站式全流程的知识图谱构建与应用平台产品

• 多模态模型应用



工作重点

奇虎 360 自研跨模态图文模型，并积累中文基准图文数据集 ZERO。视觉语言预训练框架 R2D2 用于大规模跨模态学习，基于所提出的 Zero-Corpus 数据集进行预训练，并在多个下游任务上进行测试，R2D2 取得了多项超越 SOTA 的结果，实现了中文数据集上图文跨模态算法效果的突破，可为多种下游任务如图文搜索、多模态推荐、图像生成等任务提供新的方法及框架。

网信智能中心

◆中心简介

网信智能中心（原“智能制造与智能服务中心”）由中国信息安全研究院负责承建，致力打造“中电安全数脑、数字经济智库、数字人才基地”三位一体的数字经济安全发展新型基础设施。

中心主要任务包括：

1. 中电安全数脑：基于中国电子 PKS 信创生态构筑“数字底座”，以全球自主编码、可信区块链为核心打造“数字身份”，形成包括多领域“数字场景”和全流程“数字资产”的融合基础设施。

2. 数字经济智库：为各部委和地方政府提供咨询服务，开展企业数字化转型、数字身份相关技术标准研制，与众多合作伙伴发布行业报告和智库产品。

3. 数字人才培养：获批国家双创示范基地和北京市高精尖产业技能提升培训机构，面向政企管理人员、技术研发人员等开展数字经济、网络安全、数字化转型等相关培训。

4. 自主安全计算产业链大数据平台：平台用数字化平台赋能“链长制”，构建产业链数字治理平台，绘制重点产业链知识图谱，以产业链布局创新链，以创新链发展产业链，支撑基于产业链的精准招商、项目部署、产业布局、精准施策等科学高效决策，促进产业高质量发展。

5. 企业联合创新实验室：中心与科技企业成立联合创新实验室，为企业数字化转型提供需求分析、总体规划与技术研发、成果转化等专项服务，建设大数据安全应用示范工程。

◆中心团队成员介绍

傅毅明

国家工程实验室副主任，网信智能中心主任，中国信息安全研究院院长助理，数字经济所所长，正高级工程师
研究方向：数字科技与信息安全等

吴余龙

网信智能中心副主任，武汉力龙信息科技董事长
研究方向：智慧城市

张　翔

网信智能中心副主任，交通运输部城市公共交通智能化化行业重点实验室副主任
研究方向：数据治理、智慧城市

陈　宁

网信智能中心副主任，中国社科院城市发展研究会研究员
研究方向：可信数字身份、区块链、数字经济

颜　阳

网信智能中心副主任 国家信息中心深圳大数据研究院学术委员会委员
研究方向：大数据、数字经济、区块链

耿　科

网信智能中心副主任，北京家音顺达数据技术有限公司技术总监
研究方向：数据治理、数据安全

陈彦华

中国信息安全研究院市场总监
研究方向：数字身份、数字经济

王旭阳

中国信息安全研究院研究总监
研究方向：宏观经济、产业经济、大数据分析与应用

王　鲸

中国信息安全研究院项目主管
研究方向：大数据规划、数据治理

冯　祥

中国信息安全研究院技术主管
研究方向：数字身份、区块链

张隆裕

中国信息安全研究院研究员
研究方向：数字身份、区块链

王隆泰

中国信息安全研究院市场主管
研究方向：数字经济、数据治理

王瀚璋

网信智能中心工程师
研究方向：数字身份、数据要素

刁虹钰

中国信息安全研究院数据分析师
研究方向：数量经济

越子桐

中国信息安全研究院运营主管助理
研究方向：数字经济

◆科研成果

中电安全数脑

• 中电安全数脑核心技术——码链融合数字身份

数字身份体系作为推动社会治理数字化转型的关键基础设施，以网络可信身份、全球 MA 统一标识代码、北斗网格码等为工具，融合人、物、时间空间等关键要素，通过多码融合与“码 + 链”优势互补，提升标识的科学性与上链数据的可信性，构建互信、互联、互通的全球网络应用生态，激活数据要素流通交易。

• 中电安全数脑技术架构

中电安全数脑基于中国电子 PKS 信创生态构筑“数字基座”，叠加数字身份、数字场景、数字价值构成数字大脑新框架。以全球统一标识代码、可信区块链为核心打造“码链融合”数字身份，形成包括多领域“数字场景”和全流程“数字价值”的融合基础设施。数字场景由网信中心组织协同产业生态伙伴共同开发与示范推广，包含但不限于教育、交通、金融、能源等应用领域，为数字资产建设提供场景验证支持。数字价值是数字经济发展过程中产生的全新资产形态，是数字经济发展的重要载体。

▲ 加快构建数字化安全发展新体系

重点工程应用示范

• 网络安全生态数脑

网络安全生态数脑智库系统以打造知识大数据驱动的网络安全生态智库为目标，面向相关网络安全主管部门和行业机构，提供网络安全生态指数、知识服务和精准情报的智能中枢平台。 网络安全生态数脑涵盖 9 大类数据资源、37 个基础数据库、3500 余名网络安全行业专家信息等，涉及 500 万知识大数据。

• 自主安全计算产业链大数据平台

平台用数字化平台赋能“链长制”，基于海量多源数据分析能力与专家经验知识，构建产业链数字治理平台，绘制重点产业链知识图谱，精准识别产业链发展的难点、断点、堵点，系统把握产业链的核心关键技术、环节、卡点，以产业链布局创新链，以创新链发展产业链，支撑基于产业链的精准招商、项目部署、产业布局、精准施策等科学高效决策，统筹考虑项目审批、要素保障、生态环保、技术创新等工作，为企业提供常态化“链式服务”。通过自主培育、招商引资、招才引智，优化产业布局，实现强链、补链、展链，促进产业高质量发展。

• 信创资产管理平台

信创资产管理平台总结提炼业务运营和生产运营先进新兴技术实践，运用新技术平台、码链融合数字身份、MA 全球统一编码、物联网大数据监测，把信创资产的“身份证”贯穿资产到全生命周期，建立统一资产台账，摸清信创资产底数，实现信创资产的可视可溯与一体化安全服务。平台构建全领域、智能化、自动化、人性化的资产设备管理平台，解决现阶段设备资产管理的痛点问题。

大数据分析技术创新中心

◆中心简介

大数据分析技术创新中心是大数据分析与应用技术国家工程实验室和北京大数据研究院联合共建的产业化合作与服务的重要载体与平台。中心以城市治理与决策的数字化与智能化转型赋能为目标，主要研究方向：一、智慧城市建设和运营模式、机制体制、标准规范体系研究；二、人工智能在城市治理决策中的应用；三、数据运营的模式、规范研究，数据运营的核心技术攻关。

◆中心团队成员介绍

程　超

大数据分析技术创新中心主任，北京大数据研究院院长助理。

主要研究方向是智能化治理、智慧城市应用落地、数据要素运营等，中国大数据产业发展指数总负责人，具有丰富大数据和智慧城市项目咨询和实施经验。先后参与负责北京市海淀区、东城区、经济开发区、上海化工区、天津滨海、江苏盐城、云南玉溪、贵州黔西南等地大数据和智慧城市顶层设计咨询和项目落地。

阴法超

大数据分析技术创新中心核心团队成员。

原百度集团解决方案总监、明略科技集团副总裁兼市场战略咨询中心总经理、太极计算机股份公司事业部副总经理；具有丰富的企业管理咨询、信息化规划设计经验，在人工智能、大数据等领域具有多年的实战经验。

欧高炎

大数据分析技术创新中心核心团队成员。

北京大数据研究院研究员，北京大学计算机博士、计算数学博士后，大数据教育联盟秘书长，北京大数据协会副会长，中国计算机学会数据库专委会委员，阿里云 MVP(最有价值专家)，北京市大数据人才培训示范基地培训负责人。

何海东

大数据分析技术创新中心核心团队成员。

主要研究方向是大数据、人工智能等前沿技术在生产应用领域的产品设计和研发。具有 20 多年的 IT 行业工作经验，先后就职于北大方正、华创高科等企业，担任研发负责人等角色。先后荣获教育部科技进步一等奖、国家科技进步二等奖、方正集团突出贡献奖等。

张　翔

大数据分析技术创新中心特聘专家，北京北控智慧城市科技发展有限公司技术总监。

主要研究方向为清结算系统、数据治理、大数据应用、跨行业数据融合。曾参与建设全球最大的一卡通系统，并负责系统清算、运营、大数据部分；负责京津冀互联互通大数据项目的架构、建设及运营。多项市政交通行业垂直的大数据应用实践填补国内空白。

顾敏洁

大数据分析技术创新中心核心团队成员，北京大数据研究院上海分中心主任。

长期从事企业信息化、数字化、智能化咨询及实施服务，曾任 CIBER（NYSE:CBR）中国区副总。创立了多家公司，先后帮助了数百家企业及机构客户实现数字化、智能化转型升级。

葛忠海

大数据分析技术创新中心核心团队成员，北京大数据研究院院长助理。

主要研究方向是智慧城市、大数据产业等发展规划研究。曾就职于首都信息发展股份有限公司、北京市政务网络管理中心、北京软件和信息服务交易所有限公司等单位，有多年政府部门信息化工作经验和科技园区工作经验，在主机技术、网络技术、信息安全技术、信息化项目全流程管理技术等方面经验丰富。

金雅昭

大数据分析技术创新中心核心团队成员。

长期以来致力于将大数据和人工智能技术与应用场景相结合，通过综合性解决方案促进技术的产业落地，参与筹建由大成基金、360、雪球等十余家各行业龙头企业组成的产业大数据俱乐部。带领团队帮助中国人民银行、上海市公安局、北京市质监局、上海烟草集团等众多各行业标杆客户实现数字化与智能化应用升级，灯塔项目曾获总书记参观与首肯，拥有丰富的理论与实战经验。

冷　莹

大数据分析技术创新中心核心团队成员。

主要研究方向是大数据、智慧城市的规划和咨询等。曾参与科技部重点研发计划、北京经济技术开发区大数据发展规划研究、王府井智慧商街建设研究、大数据小镇发展模式研究、北京市东城区大数据平台和领导驾驶舱建设顶层设计等项目，具有十余年信息化、大数据和智慧城市设计咨询经验。

陈　杨

大数据分析技术创新中心核心团队成员。

主要研究方向为政务数据治理。曾参与多个行业领域的大数据咨询规划相关的工作，包括了基于知识图谱的智慧城市中间层项目、全国煤炭交易中心大数据与指数项目、基于大数据技术的困难群体精准预测分析项目、东城区大数据工作咨询服务项目、街区更新之王府井智慧商街建设研究项目等，在数据治理和大数据项目落地方面具有丰富经验。

贾雅洁

大数据分析技术创新中心核心团队成员。

主要研究方向是大数据、智慧城市的规划和咨询等。参与北京市大数据综合发展指数编制、北京市东城区政务大数据建设顶层设计、北京市东城区“十四五”时期信息化发展规划编制、北京市人工智能算法安全研究等项目，在智慧城市、数字经济顶层设计规划方面具有丰富经验。

◆科研成果

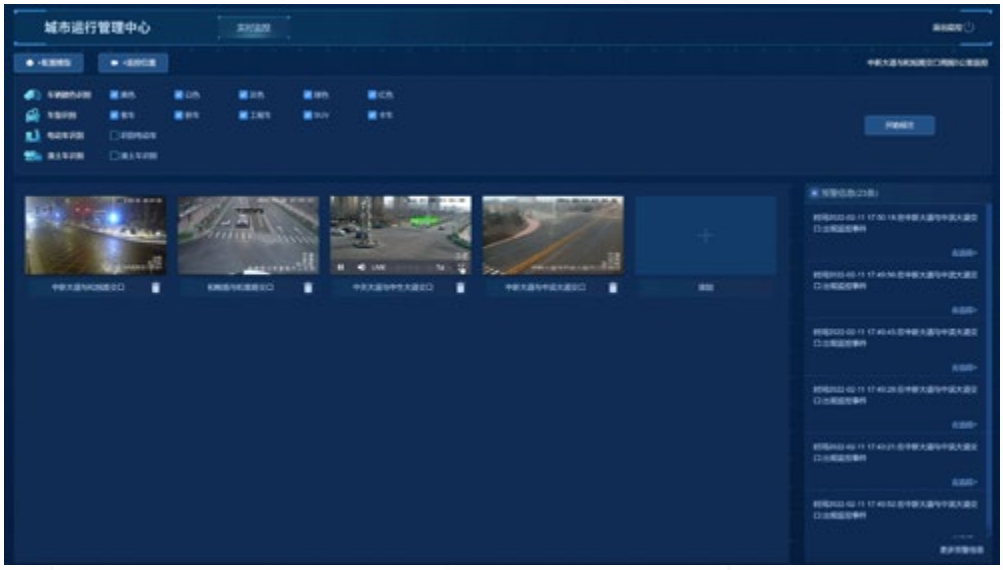
• 城市智能模型操作系统

随着智慧城市的建设与发展，异构数据极速增长，其中非结构化数据蕴含着丰富的城市运行信息，但管理和利用仍然困难；智慧城市 AI 应用整体投入成本高，各环节开发都需要大量人员与时间，重复建设问题严重；AI 模型受不同厂商和技术栈的影响互不相通，无法有效整合，面对复杂场景缺乏协同，且 AI 模型建成后运维优化容易被厂商“绑架”，模型孤岛的问题逐渐显露。为此，团队研发城市智能模型操作系统，实现了模型的可用、可管、可控。系统内预置丰富的城市 AI 模型，覆盖不同的城市应用场景，支持快速调用以及二次开发，降低智能化应用建设门槛。围绕图像、文本、视频、音频等多模态的数据，提供自动标注、人工标注等多种形式的数据标注途径，满足不同模型开发任务的需求，为模型训练提供数据基础，并针对不同用户，提供零代码、低代码、开发者等多样的模型开发模式，适配多种开发框架，满足模型训练、测试、推理、调优等各阶段的需要。对完成开发的 AI 模型提供审核、上线、评估、监测、更新等全流程的模型管理功能，保障模型质量、择优落地应用、一体化解决智慧城市领域模型管理问题。

▲ 城市智能模型操作系统功能架构

▲ 城市智能模型操作系统界面

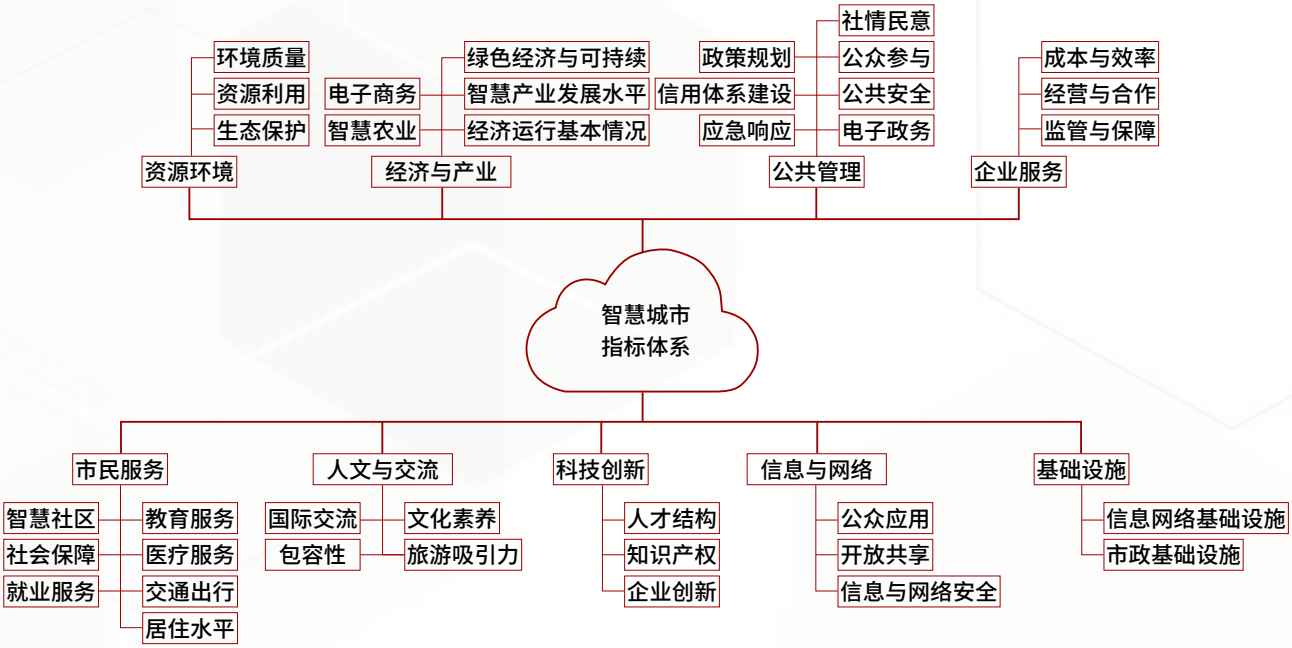
城市智能模型操作系统将通过新开发 AI 模型或整合城市已有 AI 模型组建城市 AI 基础模型库，将 AI 模型实时加载、配置、应用到任意数据源，有效降低模型的工程化成本，提升模型落地效率，目前已在某区实现相应功能，通过城市智能模型操作系统管理城市 AI 模型，并将任意厂商开发的模型加载、配置到普通摄像头中，利用异构数据开展城市管理。



▲ 城市智能模型操作系统应用界面

• 城市综合运行监测（城市体检）指标体系

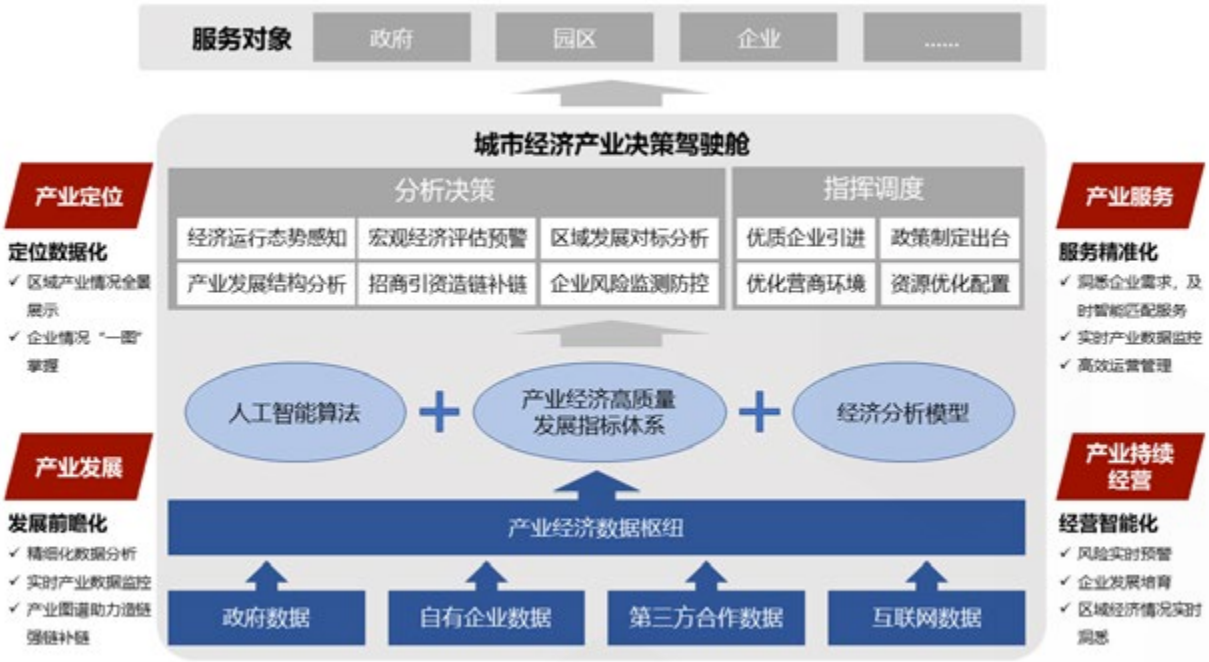
目前国内外的智慧城市指标体系大都是基于宏观数据统计综合评价理论方法的传统套路。随着大数据时代海量实时数据源不断更新和数据分析方法不断优化，中心基于大数据时代的城市运行时空数据，利用文本挖掘、知识图谱、特征状态分析和大数据可视化等分析技术，建立多尺度、多时段、多类型动态监测数据指标库，并构建一整套实时的城市综合运行监测指标体系，并结合在一些地方试点的相关经验，形成了整套解决方案，包括城市级的指标体系、600 多个全部可以量化计算的微观三级指标和一系列的指标体系应用案例，实现所有指标可量化可计算，实现基于特征状态参数对某个区域的城市综合运行情况进行监测、预警和分析。



▲ 指标体系框架

• 经济产业决策驾驶舱

城市经济产业决策驾驶舱以建设集产业经济动态感知、分析决策、综合评价、企业服务为一体的城市经济产业中枢为目标，汇聚区域政府数据、北京大数据研究院自有企业数据、第三方合作数据及互联网数据形成产业经济数据枢纽，构建经济产业高质量发展指标体系，结合机器学习算法，建立产业经济分析模型，从宏观、产业、企业三个层面对城市经济运行指标进行实时动态监测，异常及时预警，为经济管理部门配置要素资源和制订产业政策提供有效的参考依据，辅助领导进行科学决策，最终建成信息展示全面、业务监测实时、运行决策有力、机制运转高效的城市经济产业决策驾驶舱，使之成为透视全区经济运行的“慧眼”和调控城市经济运行的“智脑”。



▲ 城市经济产业决策驾驶舱功能架构

·宏观经济态势预测预警

综合应用政务数据和社会数据，选取对宏观经济变化产生影响的微观指标，通过人工智能技术建立预测模型，预测宏观经济指标的变化趋势，客观反映城市经济发展现状，对宏观经济变化和异常情况做到早发现、早预警。并回溯引起变化的指标项，剖析异常原因，使管理者提前掌握经济变化情况，为制定经济决策提供参考，克服传统统计方法的滞后性等问题。

·产业经济高质量发展

产业经济以城市主导产业为核心，聚焦园区经济、载体经济、楼宇经济等新业态，对产业结构、产业链生态、企业定位与分工等进行深入分析，梳理产业链环节分布、上下游关系和细分领域，多角度绘制产业全景图并进行可视化分析，全方位展示产业发展现状。明确地区的差异化定位和产业协同，推动经济资源、基础设施与行业政策的联动协同，让人、资金、产业、企业等要素在一张总体蓝图的指引下实现有效运转。

·企业管理服务精准高效

围绕城市企业“定、找、引、育、服、投”全生命周期进行管理和服务，基于企业数据，对企业发展状况进行评估和分析，对符合地区主导产业的企业招商引资，对城市内具有潜力的企业进行政策扶持，推动与城市经济发展不适配的企业转型升级。同时基于大数据技术解决企业管理服务中的运营效率低、服务感知差、管理体系散等问题，提升政府、园区的运营能力与服务能力。

▲ 经济产业决策平台

国家工程实验室研究基地

北京大学重庆大数据研究院

◆研究院简介

• 建设背景

2020 年 8 月 18 日，北京大学校长郝平（时任）率团赴重庆，拜会重庆市委书记陈敏尔（时任），重庆市委副书记、市长唐良智（时任），双方签署市校战略合作协议。

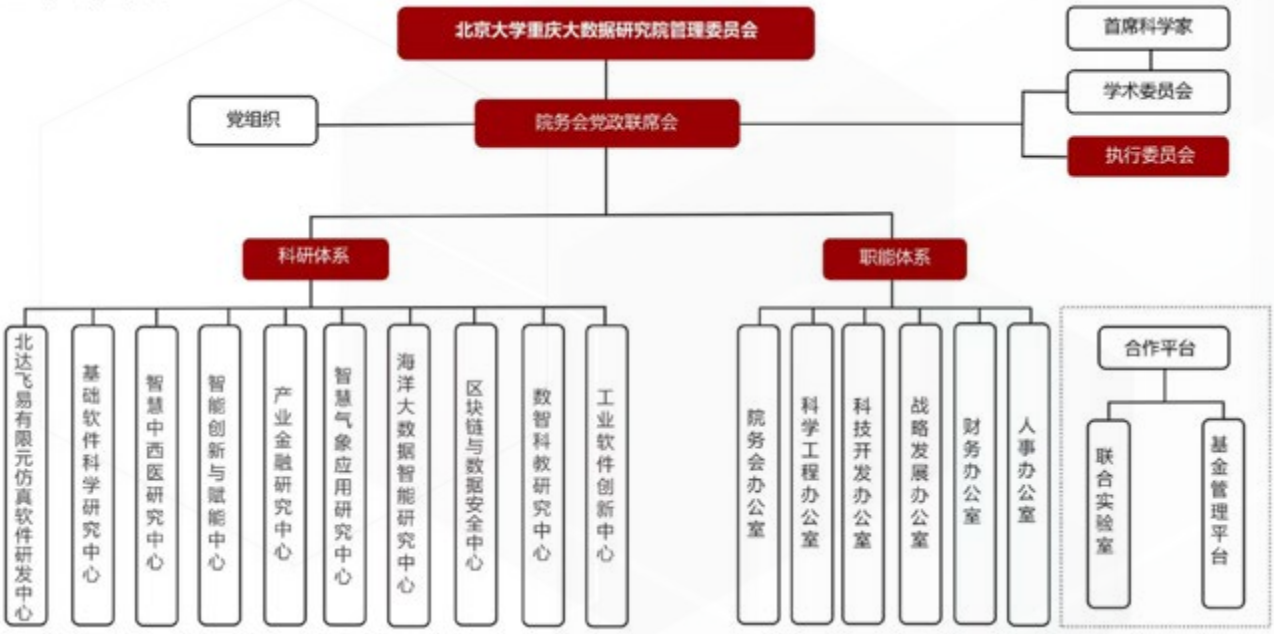
2020 年 11 月 21 日，重庆高新技术产业开发区管理委员会与北京大学正式签订合作协议。双方在西部（重庆）科学城共建北京大学重庆大数据研究院。

• 研究院简介

北京大学重庆大数据研究院成立于 2021 年，是在重庆市政府指导下，由重庆高新技术产业开发区管理委员会和北京大学共同举办的具有独立法人资格的重庆市属事业单位。

研究院扎根西部（重庆）科学城，紧密服务于国家战略需求和重庆发展需要，组建超 200 人的高水平科研队伍，聚焦大数据智能化和数字化转型的共性关键技术研发和成果转移转化，创新“高校 + 校内组织机构 + 异地科研机构”三位一体的科技成果转化模式，已初步打通数学学科从基础研究到落地应用的全链条，形成了以基础软件为“根茎”，工业软件、行业应用软件为“枝干”的软件群。目前研究院正以科技企业为载体、以创新产品为抓手、以市场化发展为目标，联合产学研用金打造创新链、产业链、资金链、人才链融合发展的创新生态，高水平服务现代化新重庆建设。

◆ 组织架构



◆首席科学家

张平文

中国科学院院士、发展中国家科学院院士。历任北京大学校长助理，副校长，党委常委、副校长等职务，现任武汉大学校长、党委副书记；北京大学博雅讲席教授，大数据分析与应用技术国家工程实验室主任，北京大学重庆大数据研究院首席科学家；中国工业与应用数学学会理事长；国家自然科学基金委“科学计算与机器学习”基础科学中心项目负责人。主要从事复杂流体的数学理论和计算方法，大数据分析与应用等方面研究，发表论文百余篇。曾获国家自然科学奖二等奖、教育部高等学校自然科学奖一等奖、何梁何利基金科学与技术进步奖、冯康科学计算奖等奖项。2011 年受邀在第七届国际工业与应用数学大会上作一小时报告，2014 年受邀在美国工业与应用数学年会上作大会报告，2018 年受邀在国际数学家大会作 45 分钟邀请报告，2020 年当选美国工业与应用数学学会会士。担任学术期刊 CSIAM Transactions on Applied Mathematics 主编，Applied Mathematics and Mechanics 等十余种国内外期刊编委。

◆院领导

胡　俊　院长

北京大学数学科学学院党委书记、教授，北京大学重庆大数据研究院院长。兼任 Adv Appl. Math. Mech. 执行主编、多个期刊的编委、北京计算数学学会理事长、中国数学会常务理事、中国大坝工程学会大坝数值模拟专委会副主任委员、重庆市工业软件应用发展协会副会长、北大 - 华为数学联合实验室主任。主要从事非标准有限元方法的研究，特别是弹性力学问题、线性化 Einstein-Bianchi 方程组及相关问题的非标准有限元方法的构造与数值分析的研究，解决了弹性力学问题混合有限元方法的构造这个长期悬而未决的公开问题，首次构造了线性化 Einstein-Bianchi 方程组保结构的稳定有限元方法。曾获教育部自然科学奖一等奖、国家杰出青年科学基金、冯康科学计算奖、中国数学会计算数学分会首届青年创新奖。

周晓华　副院长

北京大学讲席教授，北京大学生物统计系主任，北京大学国家药品医疗器械监管科学研究院副院长，北京大学重庆大数据研究院副院长、研究院智慧中医药研究中心主任，西部（重庆）科学城高新区“金凤凰” 杰出人才 , 美国科学促进会（AAAS）会士，美国统计学会 (ASA) 会士，国际数理统计研究院 (IMS) 会士。曾荣获国际贝叶斯统计科学学会 Mitchell 奖、中国产学研合作创新奖、世界中医药学会联合会中医药国际贡献奖 - 科技进步二等奖、国家基金委海外杰青、教育部高层次文教专家、教育部海外名师、2022 十大重庆科技创新年度人物等奖项和荣誉。研究方向主要集中于医学诊断学的统计方法，因果推断方法，医疗大数据建模，新兴传染病的数学和统计建模，中医药研究的统计方法，发表 SCI 学术论文 290 余篇。

李　若　副院长

北京大学博雅特聘教授、北京大学数学科学学院副院长、北京大学重庆大数据研究院副院长、北京大学大数据分析与应用技术国家工程实验室核心科研成员。获聘教育部长江特聘教授，曾任中国数学会计算数学分会理事、副理事长，北京计算数学学会理事长、监事长；曾获教育部高校自然科学奖一等奖（第三完成人）、科技部中青年科技创新领军人才、冯康科学计算奖、第九届国际工业与应用数学大会报告人、吴文俊应用数学奖、CSIAM 会士等荣誉；曾主持和参与国家杰出青年基金、科技部 973 项目等。主要研究方向为偏微分方程数值解和计算流体力学，发表科研论文 100 余篇，学术成就主要在动理学方程的模型约化理论、约化模型的数值方法方面取得突破性进展，解决了存在六十多年的 Grad 矩方程组的双曲性缺失问题。

王　涛　副院长

北京大学重庆大数据研究院副院长，中关村数智人工智能产业联盟副秘书长（兼），重庆数字经济研究中心特聘专家，高级技术经纪人。曾任北京大学大数据分析与应用技术国家工程实验室郑州中心副主任、北京大数据研究院博雅大数据学院执行院长等。

黄文彬　副院长

北京大学信息管理系主任助理、长聘副教授、博导、研究员，北京大学信息化与信息管理研究中心主任、北京大学信息管理系大数据管理与应用教研室主任，北京大学重庆大数据研究院副院长。曾获教育部国家级教学成果奖二等奖、北京市高等教育教学成果奖二等奖、黄廷芳 / 信和青年杰出学者奖、北京大学教学优秀奖等荣誉。致力于基于数据科学、多媒体技术、人工智能等技术面向人文社会科学 (教育、情报、出版、金融、科学学) 的应用研究。主持、参与国家社会科学重点项目、面上项目、以及中宣部、文旅部、网信办委托项目等，参撰著作 3 部，发表论文 50 余篇。

◆科研体系

自 2021 年落地西部 (重庆) 科学城以来，研究院始终聚焦国家战略需求和重庆发展需要，致力于攻克大数据和数字化转型的核心科学技术问题和创新应用瓶颈，构建了一个由多个研究中心组成的科研体系。各研究中心由该领域杰出学者挂帅，组建高水平团队，持续开展关键核心技术攻关。至今，已推出原创成果 20 余项，形成了以基础软件为“根茎”，工业软件、行业应用软件为“枝干”的软件群。随着这些产品的推广应用，科技创新正转化为实际生产力，持续不断地为现代化新重庆建设注入新动能。

- 北达飞易有限元仿真软件研发中心
- 基础软件科学研究中心
- 智慧中西医研究中心
- 智能创新与赋能中心
- 产业金融研究中心
- 智慧气象应用研究中心
- 海洋大数据智能研究中心
- 区块链与数据安全中心
- 数智科教研究中心

……

• 北达飞易有限元仿真软件研发中心

BeFEM 北达飞易

北达飞易有限元仿真软件研发中心是在北京大学数学科学学院、北京大学大数据分析与应用技术国家工程实验室的支持下，由北京大学重庆大数据研究院于 2022 年 4 月在重庆高新区成立。中心聚焦 CAE 底层数学问题攻关及其产品化，以“打造新一代算法引擎，为国产 CAE 崛起筑基”为使命，研发 CAE 软件国产化引擎，为实现我国工业仿真软件自主研发、核心技术安全可控贡献北大力量。

团队负责人

胡俊　主任

北京大学数学科学学院党委书记、教授

曾获教育部自然科学奖一等奖、国家杰出青年科学基金、冯康科学计算奖、中国数学会计算数学分会首届青年创新奖等荣誉

研究方向：非标准有限元方法

吴金彪　执行主任

北京大学数学科学学院副教授

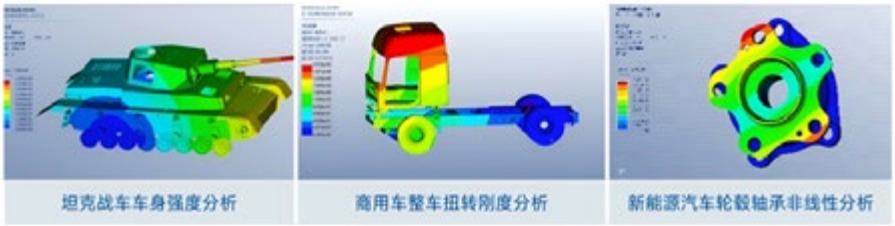
研究方向：多重网格方法

科研成果

北达飞易新一代 CAE 算法引擎：北达飞易依托北京大学数学科学学院胡俊教授课题组在有限元领域的原始创新能力，成功将团队奠基性成果弹性力学问题混合有限元算法、原创性成果 AI 混合迭代法和偏微分方程直接元法等一系列创新成果转化落地，构建起新一代算法引擎，为国产 CAE 软件的崛起奠定了坚实基础。

▲ 产品架构图

北达飞易有限元仿真分析软件：基于新一代 CAE 算法引擎开发的具有完全知识产权的北达飞易有限元仿真分析软件，可有效支撑大规模自由度模型的快速求解，实现对结构刚强度、疲劳、非线性、动力学等问题的高精度仿真，满足新时代先进制造和基础设施行业对工业仿真软件的现实需求。



• 基础软件科学研究中心

基础软件科学研究中心是北京大学重庆大数据研究院研究数学关键核心基础软件与工业软件底层技术的重点中心，下设数值计算实验室、系统仿真与设计实验室、复杂系统智能分析与应用实验室。中心开展科学计算软件、工业基础软件等核心技术研究与产品孵化，从国产科学计算底层出发，探索并突破关键核心技术，研发相关核心数学基础软件与工业基础软件，在瞄准替换国外软件的发展过程中，实现国产自主可控，助力国家科技战略力量。

团队负责人

李若　中心主任

北京大学重庆大数据研究院副院长
北京大学数学科学学院副院长、博雅特聘教授、博士生导师

卢朓　中心执行主任

北京大学重庆大数据研究院基础软件科学研究中心执行主任
北京大学数学科学学院副教授、博士生导师

科研成果

1 北太天元科学计算与系统仿真软件

北太天元科学计算与系统仿真软件将应用数学的科研成果转化为实际生产力，是面向科学研究与工程技术开发的基础性生产工具，汇集沉淀数学、物理、计算机和工程技术等多方面的知识，作为国内工业研发设计类软件的计算底层，为各专业软件提供计算内核，赋能核心技术研发，推进实现数字化智能研发与制造的目标。北太天元科学计算与系统仿真软件由四部分组成：北太天元科学计算软件、北太真元系统仿真软件、工具箱、插件。

▲ 北太天元产品架构图

·北太天元科学计算软件

北太天元科学计算软件面向科学计算，为各领域科学家与工程师提供一体化的科学计算平台，具备强大的 M 语言集成开发环境与底层数学函数库，支持数值计算、数据分析与可视化、算法与应用程序开发等工作，内置了丰富的工具箱，如控制系统工具箱、图像处理工具箱等，满足不同领域需求，并通过 SDK 与 API 接口，扩展支持各学科计算场景。北太天元内核作为根技术，具备解释器机制与词法语法体系，是软件的计算引擎，并配备了核心函数库、词法语法分析、程序调试器与开发者工具（SDK）。

·北太真元系统仿真软件

北太真元系统仿真软件面向工程计算场景，是基于北太天元内核能力构建的系统建模与仿真软件。北太真元提供了完善的模块化、图形化的建模仿真环境，内置丰富模块库及求解器，可实现对连续 / 离散 / 混合时间系统的精准模拟，满足各种类仿真需求。在模型构建上，北太真元基于模型的系统工程设计方法论，能够有效解决航空航天、电力电子、控制工程等领域中的跨学科仿真和优化问题。

·工具箱

工具箱是软件的扩展模块，用于提供额外的函数、功能、工具，针对不同领域的科学计算和编程需求，提供针对性的领域解决方案。工具箱分为通用工具箱和专业工具箱。通用工具箱作为基础工具集，适用于广泛的专业和行业，为用户提供一系列常用的数学、数据分析和编程工具。专业工具箱针对特定专业，旨在协助专门领域的研究人员和工程人员解决复杂的专业性问题。

·插件

使用北太天元提供的开发者工具（SDK）开发的软件组件，为北太天元扩展函数、功能或特性。

• 智慧中西医研究中心

智慧中西医研究中心下设中医药大数据实验室、中医药真实世界研究实验室、智慧医疗实验室，依托北京大学数学科学院、北京大学生物统计系、北京国际数学研究中心，联合国内多家知名医院、医药院校及科研机构、大型医药企业，由生物统计、医学、人工智能、大数据等多专业人员组成多学科研究团队，研发自然语言处理、基于 AI 的多标签证候预测，证候量表评价，精准医学和诊断医学、中医疗效评价分析等技术。

研究方向聚焦于中西医重大疾病早筛与干预、中西医药真实世界研究、中医证候诊断与疗效评价、中医辨证论治智能化、医疗服务与药品监管、脑科学、生物基因等多个领域，并通过技术的成果转化打造面向多种应用场景、智能一体化的核心产品。研究中心致力于推动医疗健康服务模式创新，助力名老中医经验传承和医学早筛，提升基层医药服务能力和技术水平，加速医药健康产品研发创新，打造支撑医药健康传承创新发展的数智化底座平台。

团队负责人

周晓华

北京大学讲席教授、北京大学生物统计系主任、北京大学国家药品医疗器械监管科学研究院副院长、北京大学重庆大数据研究院副院长、研究院智慧中西医研究中心主任、西部（重庆）科学城高新区“金凤凰” 杰出人才、美国科学促进会（AAAS）会士、美国统计学会 (ASA) 会士、国际数理统计研究院 (IMS) 会士。

黄新霆

北京大学大数据分析与应用国家工程实验室数据处理与统计分析中心成员、世界中医药学会联合会中医药大数据产业分会秘书长、重庆大模型联盟副理事长、中国中医药信息研究会人工智能分会常务理事、西部（重庆）科学城高新区“金凤凰” 领军人才、国家发改委 2023 年重点支持项目“中医辨证论治智能辅助系统开发与应用项目”技术负责人。

科研成果

1 智慧中医药一体化服务平台

智慧中医药一体化服务平台打通中医药大数据从平台建设、采集治理、分析挖掘到智能应用的完整闭环，实现以大数据为基础、中医理论为指导、具备自我学习能力的中医辨证论治智能辅助系统和中西医结合知识库系统赋能基层中医，以面向个人的中医治未病系统和中医智能预诊和随访系统赋能中医大健康管理，以名老中医经验数智化传承系统赋能名老中医经验传承，以中医证候诊断与疗效评价平台，解决中医药证候诊断与疗效评价的标准化难题，进而为中医药传承创新发展提供数智化支撑。

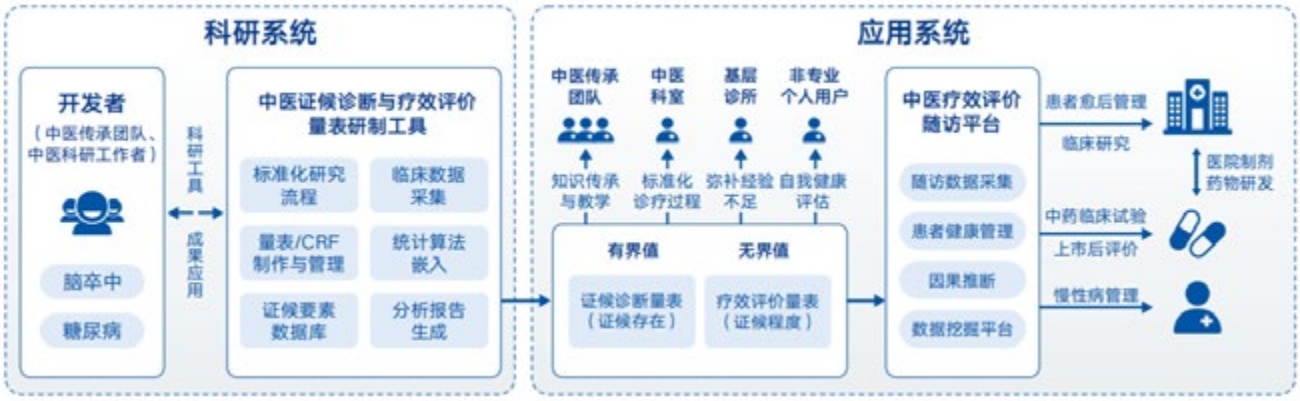
中医辨证论治智能辅助系统开发与应用项目已成功入选国家发改委生物技术与信息技术融合应用工程重点方向支持项目名单。

该项目结合中医望闻问切等诊断需求以及针灸、艾灸等治疗需求，基于中医学方法论，研发集成式中医辨证论治智能辅助系统，挖掘经典医籍和名医经验，建立多维辨证辅助算法和方证适配模型，实现中医治法、中药处方智能推荐。结合中医医疗机构智慧医院建设，搭建医养康中医信息化服务云平台，为服务对象建立中医个人健康档案，提升重点人群中医健康管理水平。建设周期 2-3 年。建成后，将在全国 30 家以上医疗机构推动中医智能辅助诊疗系统落地示范应用。

科研成果

2 中医证候诊断与疗效评价平台

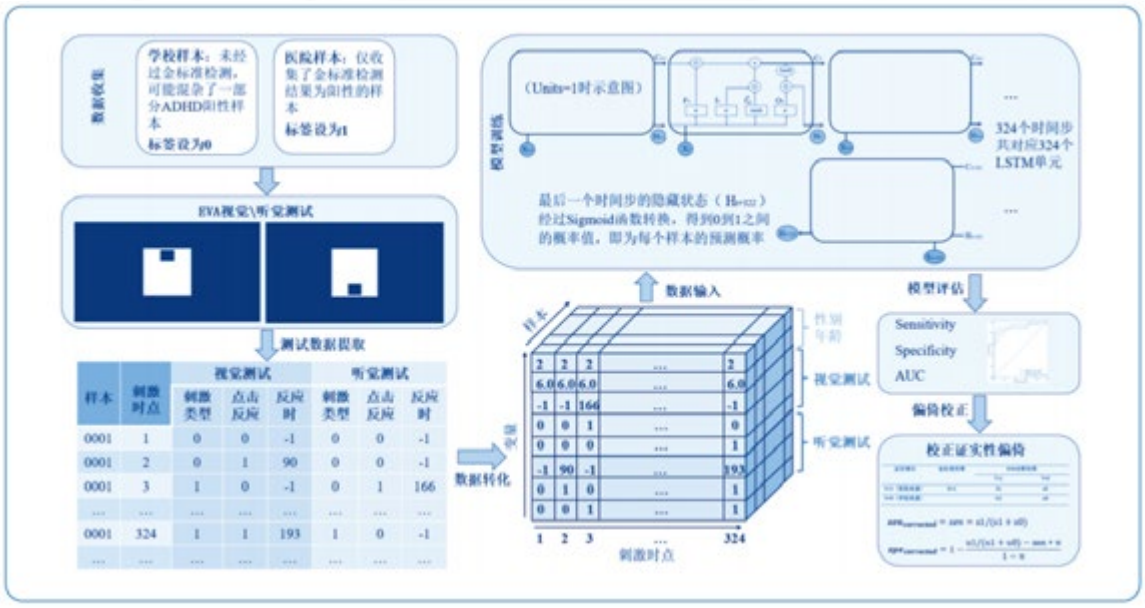
中医证候诊断与疗效评价平台是国内首个中医诊断量表全流程研制平台。平台以中医证候诊断量表研制技术规范为理论依据，以证候要素库、专家库、算法库为支撑，结合数据结构化操作、经验智能化转化等技术，首创性提出了在无“金标准”的情况下量表的考评方法、中医证候疗效评估的因果推断统计方法，全面助力中医药研究团队、中医院、中药药企等用户开展中医证候诊断量表研制与疗效评价。



▲ 系统架构

3 ADHD 辅助诊断系统 EVA

ADHD 辅助诊断系统 EVA 利用校正证实性偏倚的统计方法，解决数据收集过程中较为常见的金标准缺失问题，并通过将深度学习算法与 EVA 测试数据有机结合，提升模型的灵敏度和特异度。同时，通过解决 TOVA 现存的诸多问题，EVA 能更准确地评估中国儿童的注意力水平，达到早发现早治疗的目的，实现“筛未病，测预后”，为学生心理健康提供更为全面和深入的保障。



▲ EVA 建模流程图

• 智能创新与赋能中心

智能创新与赋能中心紧密围绕国家及地区的科技战略部署，顺应科技发展的前沿趋势，以数字化转型与人工智能为研究领域，以大模型、大数据等尖端技术为基石，以务实、有效推动政府及企业等主体的数字化、智能化转型为目标，提供创新服务与智库支持。中心下设数字化转型促进中心和政务智能问答数据科学实验室，致力于打造基于大模型的全流程产品及服务体系，包括战略规划咨询、数据治理咨询、大模型应用解决方案、培训与宣贯全流程四个板块。

团队负责人

战略委员会

张平文　主任委员
中国科学院院士
发展中国家科学院院士
北京大学博雅讲席教授

周　剑　主任委员
点亮智库秘书长
国信院院长
中信联副理事长兼秘书长

主任团队

邱泽奇　主任
北京大学中国社会与发展研究中心主任
北京大学数字治理研究中心主任
北京大学教授、博士生导师，长江学者特聘教授

黄　晶　副主任
北京大学大数据分析与应用技术国家工程实验室产品与工程部副部长
算法专家

核心职能

1 技术产业和政策研究 ·把握转型方向
开展技术产业相关发展趋势研判，推动产业转型，助力产业发展，打通国家和重庆战略政策落地的最后一公里。

2 平台服务落地与推广 ·降低转型难度
搭建重庆市数字化转型公共服务平台，整合市场化平台服务资源，以精准需求挖掘和精准服务画像为基础提升供需对接能力。

3 场景和解决方案开发 ·优化转型供给
以用户需求为导向，整合跨产业、跨专业的创新资源，围绕重庆重点行业开展端到端的业务场景开发，并提出解决方案，打造并推广示范样板。

4 数字人才培养和培训 ·激发转型活力
构建产学研用专家库，围绕重庆产业特点构建知识体系，开发数字化人才培养和认证体系，培育数字化人才。

5 学术交流和产业合作 ·构建转型生态
开展多层次宣贯与产业交流活动，协同发展的合作机制和商业模式。

产品服务

中心致力于打造围绕大模型的全流程产品及服务体系，包括战略规划咨询、数据治理咨询、大模型应用解决方案、培训与宣贯全流程四个板块。

战略规划咨询

制定长期战略、关注多维视角、确保方向准确，使 AI 技术赋能业务长期发展

培训与宣贯

解决理解、监管、掌握、运用的问题

数据治理咨询

基于未来人工智能模型训练的需要，开展数据治理咨询

大模型最后一公里应用解决方案

解决通用大模型成本高、性能差、无法保障数据安全的问题

战略规划咨询

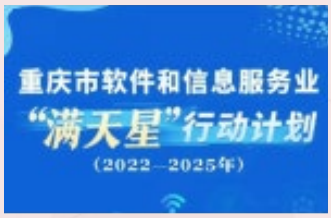
以政策研判和行业分析为输入，以目标愿景为方向，以业务场景优化为切入，以大模型应用场景设计为抓手，通过数字资源优化强治理、挖价值，通过人才建设定机制、强管理、保执行。

- 识别转型核心需求及潜在瓶颈；
- 构建数字化转型顶层设计和人工智能应用场景规划；
- 确立明确实施步骤，确保业务、流程和数字技术深度结合。

▲ 战略规划咨询架构

典型项目

政府咨询服务



重庆市软信产业发展
“满天星” 计划

数字生态指数
成渝数字生态研究

数字生态指数
成渝数字生态研究

企业咨询服务

城投公司数字化
转型顶层规划

金融担保公司数字化
转型顶层规划

广电传媒公司
数据治理咨询

大模型应用解决方案

梳理不同领域专业知识与业务逻辑，充分挖掘、匹配大模型能力，开展大模型数据治理，训练行业大模型 / 微调大模型，解决应用场景的最后一公里。

“源启小政”政务智能问答产品

“源启小政”政务智能问答系统是中心基于大语言模型开发的产品，具备强大的多轮实时交互问答能力。其功能涵盖了从政务咨询到政策匹配的多个场景，包括政务咨询问答、政务信息查询、办事流程指引政策精准匹配等。相较于传统的问答产品，意图理解更精准、内容生成更优质、依据溯源更准确、提问引导更高效。

▲ “源启小政”产品架构

• 产业金融研究中心

产业金融研究中心下设智慧金融实验室和新型智慧城市实验室。中心依托人工智能、大数据技术和先进算法，面向金融机构和政府，研发产业链智能分析平台和各类金融科技产品，构建“政府 - 产业 - 金融机构”生态圈，以强化金融机构产业分析和风险防控能力，助力地方政府“加快建设现代化产业体系”。

团队负责人

中心主任

鄂维南

中国科学院院士，美国数学学会会士，美国工业与应用数学学会会士。北京大学讲席教授、大数据分析与应用技术国家工程实验室联席主任、北京大学国际机器学习研究中心主任、北京科学智能研究院院长。

中心执行主任

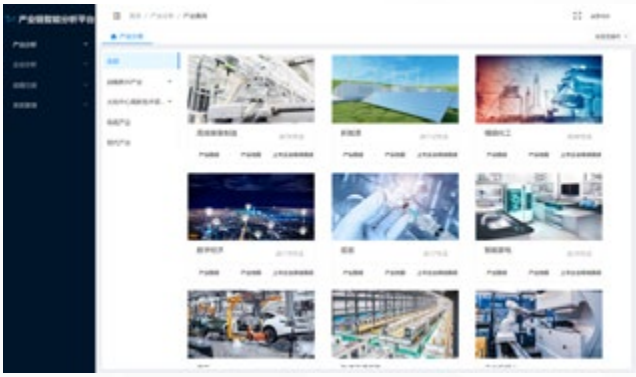
王冉冉

北京大学大数据分析与应用技术国家工程实验室助理研究员

研究领域为企业研究、金融风险管理、以及机器学习算法

科研成果

智城链 - 产业链智能分析平台是一款集成了产业智能分析、监测与产业数字招商的一体化平台。平台中内置了超过 100 条现代化产业的产业图谱、企业图谱、智能分析模型、BI 报表等工具，可帮助用户快速摸清本地产业发展情况并与全国各地区进行对比分析，快速对目标企业进行线上调研，并快速制定产业招商名单。通过大数据手段，数智赋能产业政策制定、产业发展监测、强链补链招商、产业金融创新、产业风险传导等产业相关场景。



▲ 智城链 - 产业链智能分析平台界面

• 智慧气象应用研究中心

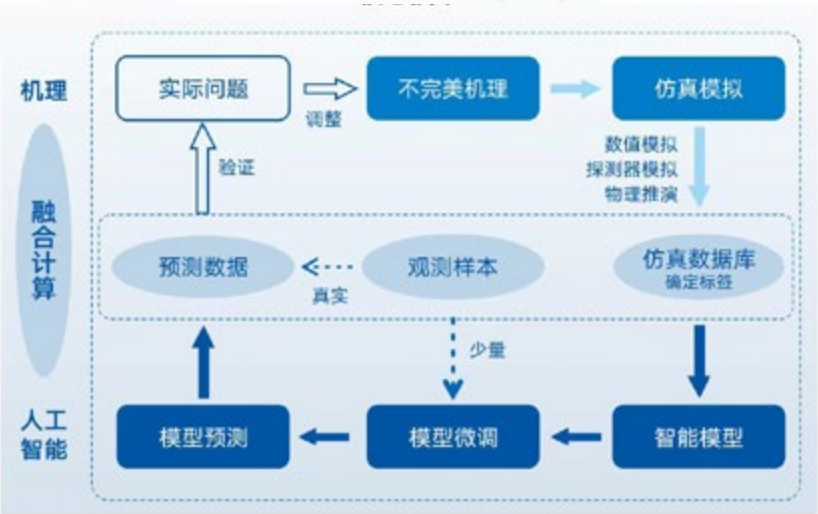
智慧气象应用研究中心下设智能会商与人工智能天气预报实验室、智能遥感监测实验室。中心以“服务国家重大战略需求”和“做出有世界影响的成果”为愿景，坚持北京大学创新链、气象部门业务链和地方产业链“三链融合”，着力大跨度学科交叉融合，以“AI 赋能气象”科技创新为动力，深入推进人工智能、大数据等新一代信息技术与监测、预报和服务等全气象产业链条有机融合，开展颠覆性和前沿引领性研究，形成一批技术领先、应用广泛的智能气象服务产品。中心致力于为气象以及遥感、能源、交通、农业、国防、环保、保险等衍生领域用户提供气象遥感大数据处理、模式智能优化、精细化数值仿真、风能太阳能资源评估以及“双碳”监测分析等前沿技术服务及智慧解决方案，助力我国气象服务产业智能化升级。

团队负责人

张平文

中国科学院院士、发展中国家科学院院士。历任北京大学校长助理，副校长，党委常委、副校长等职务，现任武汉大学校长、党委副书记；北京大学博雅讲席教授，大数据分析与应用技术国家工程实验室主任；中国工业与应用数学学会理事长；国家自然科学基金委“科学计算与机器学习”基础科学中心项目负责人。

核心技术

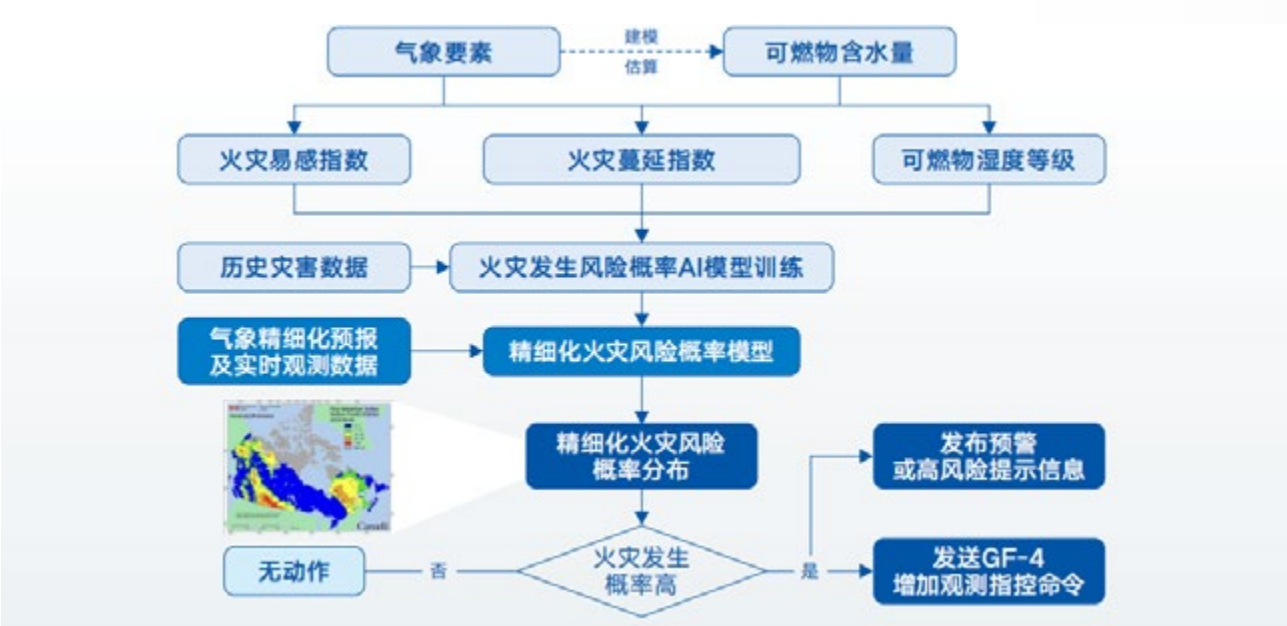


主要产品

科研成果

中心基于气象遥感大数据处理，人工智能大模型方法和多源数据融合算法的研究，培育孵化了两个产品：面向新型电力系统的电力气象预报大模型，基于数据增强的森林火险智能监测预警系统。电力气象预报大模型可以对电力气象关键要素实现高分辨率的准确预报，为电网新能源出力预测、中长期电力电量平衡提供气象预报的支持。森林火险智能监测预警系统可以提供森林火灾全生命周期的初期识别、发展预测、风险管理、灾害决策、事后处理的技术服务与支持，帮助用户提高自然灾害监测精度和预警有效性，提升应对灾害的能力。

▲面向新型电力系统电力气象预报大模型



▲ 森林火险概率预报系统

• 海洋大数据智能研究中心

海洋大数据智能研究中心是由北京大学国际关系学院、北京大学海洋战略研究中心、北京大学海洋研究院等单位的专家学者组成的多学科研究团队，下设开源大数据智能决策实验室、多源异构数据融合实验室。中心充分发挥社科研究经验与大数据智能技术结合的领先优势，运用 AI+ 海洋大数据，深入开展数据采集、多源异构数据融合、智能决策等关键技术研发和智慧海洋的应用领域探索。凭借团队在海洋安全、数据融合和智能决策等方面的深厚积累，中心将为广大用户提供集数据、算法和系统于一体的解决方案，助力相关行业海域态势感知能力的提升。

团队负责人

中心主任

胡　波

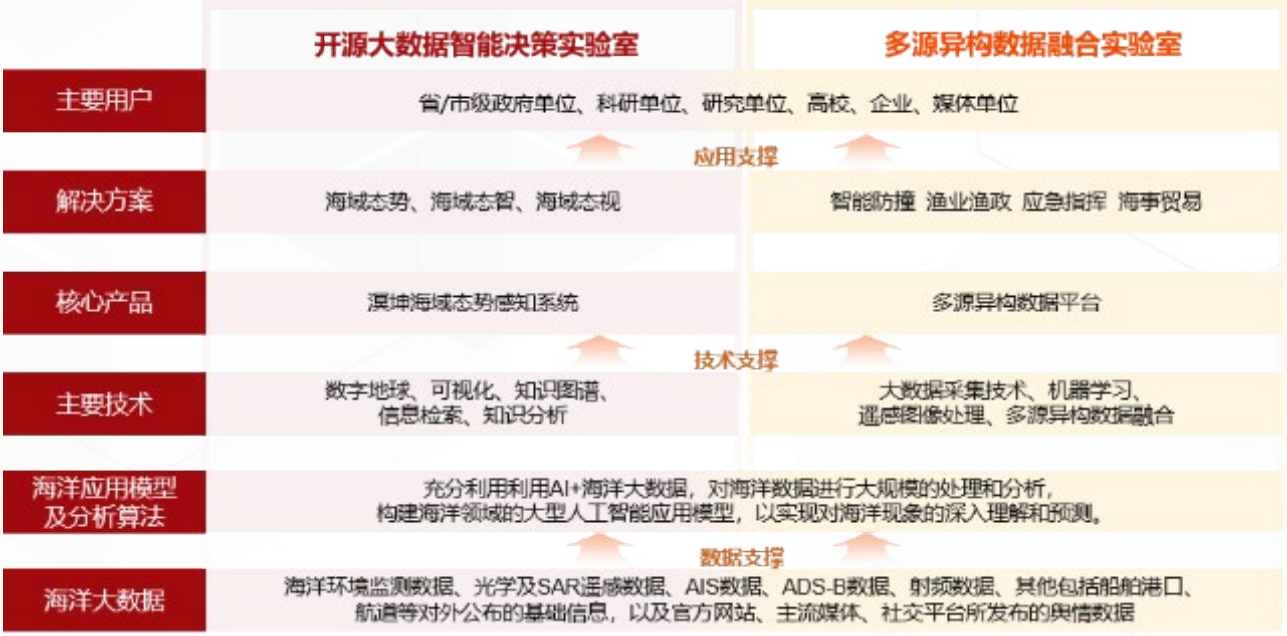
北京大学国际关系学院研究员、北京大学海洋战略研究中心主任，兼任“南海战略态势感知计划”主任，长期从事海洋战略与政策等方面的研究，主要关注领域为海权战略、社科研究中的数据赋能、国际安全及美国军事。

中心副主任

王继民

北京大学信息管理系教授，北京大学海洋研究院中心成员，《文献与数据学报》副主编。研究方向：机器学习、Web 挖掘、科学评价、信息可视化等。曾主持国家“核高基”重大科技专项子课题、国家社科基金重点项目、以及国家发改委、教育部、北京市科委等科研课题 30 余项。

中心架构

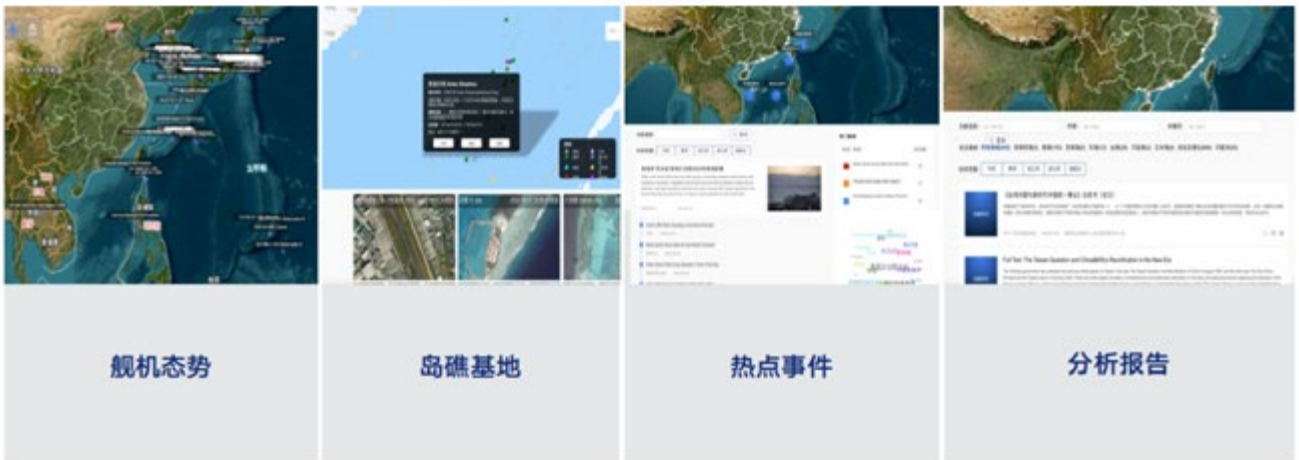


科研成果

1 “溟坤海域态势感知系统”

“溟坤海域态势感知系统”基于地理空间信息系统，融合了 AIS、ADS-B、商业遥感、无线电和新闻舆情等多源异构数据，在数字地球上叠加数据层、虚拟对象等覆盖地形，并结合海域研究经验以符合业务和行业逻辑的方式创建逼真的可视化场景，全面展现海上态势信息，为用户提供多维、全景、动态数据系统。在战略态势领域填补国内空白，态势数据 + 互联网新闻信息数据的融合处于国内外探索前列。

▲ “溟坤海域态势感知系统”界面



▲ 核心功能

• 区块链与数据安全中心

区块链与数据安全中心面向数据要素市场化配置开展区块链与数据安全技术的研究和应用，基于区块链、大数据与隐私计算技术打造数据安全共享与交易平台，推动区块链在数据确权保护、数据安全治理、数据可信共享和数据要素交易领域的技术创新和应用推广。

目前研究主要分为三部分：面向大数据应用的数据安全防护体系研究、基于区块链的政务数据共享平台研究、数据要素市场化配置创新模式与支撑技术研究。

团队负责人

陈　钟

北京大学计算机学院教授、区块链研究中心主任，金融信息化研究中心主任，网络与信息安全实验室主任，北京大学软件与微电子学院创始院长。同时兼任中国计算机学会会士、常务理事，担任中国软件行业协会副理事长，中国开源软件推进联盟副理事长，全国区块链与分布式记账技术标准化技术委员会委员，全国信息安全、信息保密标准化技术委员会委员，科技部十四五规划“区块链”专家组成员等。主要研究领域为面向领域的软件工程、网络与信息安全、区块链技术等，曾主持和参加完成国家科技重大专项、国家 863、国家自然科学基金以及国家发改委、国家保密局、人民银行、工信部、公安部、北京市等部委级科技专项二十余项，曾获国家科技进步二等奖两项、国家级教学成果一等奖一项，部委级科技成果奖励多项。

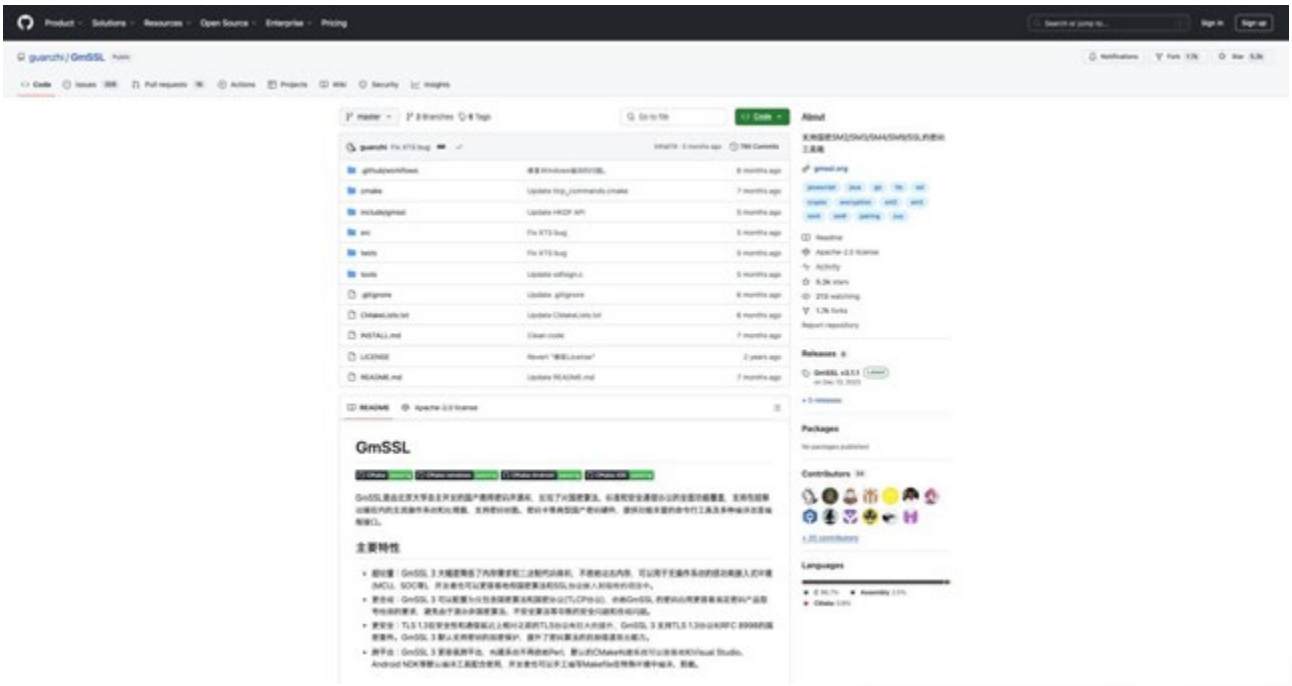
科研成果

1 RegChain

RegChain 是内置 RegLang 智能合约执行引擎、GmSSL 国密算法等自研核心技术的内生安全、高效扩展、自主可控的国产基础联盟链平台。

2 GmSSL

GmSSL 自主开发的国产商用密码开源库，实现了对国密算法、标准和安全通信协议的全面功能覆盖。



• 数智科教研究中心

数智科教研究中心旨在以科教融合协同发展为目标，以科教大数据为基础，以智能算法为依托，以精准画像分析为导向，构建科教大数据管理与分析平台，开展监测评估、统筹资源、政策制定、转化 / 就业服务等科学技术管理和高等教育管理实际问题的研究与实践，促进“大教育”的科研、“大科学”的教育，推动我国科学技术管理和高等教育管理数智化建设，助力科教兴国战略实施。

团队负责人

黄文彬

北京大学信息管理系主任助理、长聘副教授、博导、研究员，北京大学信息化与信息管理研究中心主任、北京大学信息管理系大数据管理与应用教研室主任，北京大学重庆大数据研究院副院长。曾获教育部国家级教学成果奖二等奖、北京市高等教育教学成果奖二等奖、黄廷芳 / 信和青年杰出学者奖、北京大学教学优秀奖等荣誉。致力于基于数据科学、多媒体技术、人工智能等技术面向人文社会科学 (教育、情报、出版、金融、科学学) 的应用研究。主持、参与国家社会科学重点项目、面上项目、以及中宣部、文旅部、网信办委托项目等，参撰著作 3 部，发表论文 50 余篇。

科研成果

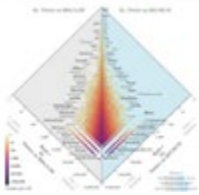
科研项目：近三年来主持与科教管理与评价、大数据分析与应用等相关的国家级项目 3 项，省部级项目 4 项，企事业单位委托项目 10+ 项，校级项目 5 项。

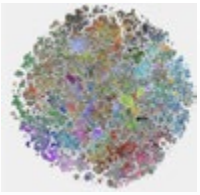
科学论文：近三年来在数据科学、信息资源管理、科学技术管理、高等教育管理、计算机科学与技术等领域发表中英文期刊和会议论文 60+ 篇。

平台系统：近三年来形成与科技信息和教育信息大数据展示平台 2 个。

基于科教大数据的研究生培养画像研究

以研究生培养过程中关于科研项目管理、课程管理、校园生活等多方面的数据为前提，采用数字画像技术通过指标构建、画像标签等手段对研究生、教师及学院三个层级进行多维度的呈现。





◆联合实验室

与重庆嘉发实业（集团）有限公司共建大数据智能化应用联合实验室：围绕企业数字化转型顶层设计、智慧运维、节能环保、精准营销、智慧中医等领域开展合作研究，解决传统企业数字化转型的核心科学问题及创新应用瓶颈。

与博雅正链（北京）科技有限公司共建区块链创新应用联合实验室：探索大数据技术与区块链技术在数字生态中的深度融合与创新，推动产业链与创新链真正有效结合，共同赋能数字生态发展。

与北京无疆脑智科技有限公司共建脑科学大数据联合实验室：围绕注意力缺陷多动障碍、3-6岁学龄前儿童神经发育水平障碍、严重精神障碍、自闭症等研究方向，打造脑科学和精神健康数字化解决方案。

与中电信翼金科技有限公司、北源创智（重庆）科技有限公司共建人工智能研发创新联合实验室：聚焦大模型及人工智能技术在金融科技领域的深度应用，研发智慧营销大模型、金融行业洞察智囊以及数字化 CFO 等创新金融 AI 产品，引领金融科技未来发展。

◆联系方式

地址：重庆高新区科学谷 10 号楼
电话：023-65257000
网址：http://cqbdri.pku.edu.cn

公众号二维码

北京大学长沙计算与数字经济研究院

◆ 研究院简介

• 建设背景

北京大学长沙计算与数字经济研究院是北京大学、长沙市人民政府和湖南湘江新区（长沙高新区）管理委员会共建的长沙市属事业单位。研究院以建设高水平新型研发机构为目标，围绕先进计算、大数据、数字经济与数字化转型等领域的核心科学技术问题和重大应用难题，广泛集聚全球顶尖人才团队，开展基础研究、应用研究和政策研究，培养创新人才，孵化科技企业，构建具有国际影响力、引领国内先进计算和数字经济发展的人才聚集平台、前沿研究平台和成果转化平台。

人才聚集平台

前沿研究平台

成果转化平台

◆建设历程

2021.5.21

湖南省委常委、长沙市委书记吴桂英率团访问北京大学，同日，北京大学与长沙市人民政府签署深化全面合作协议。

北京大学与长沙市人民政府签署深化全面合作协议

2022.1.14

北京大学、长沙市人民政府、长沙高新区管委会在长沙举行共建北京大学计算与数字经济研究院合作协议签约仪式。

共建北京大学计算与数字经济研究院合作协议签约仪式

2022.3.29

研究院完成登记注册。

2022.6.9

研究院第一届理事会第一次会议在长沙顺利召开，研究院各项工作正式启动。

研究院第一届理事会第一次会议与会人员合影

2023.9.1

中国科学院院士、北京大学党委常委、副校长张锦等领导为研究院揭牌。

北京大学长沙计算与数字经济研究院揭牌仪式

2024.4.7

研究院第一届管理委员会（原理事会）第二次会议在长沙顺利召开。

2024.9.20

北京大学党委书记郝平率团访问湖南推动校地合作，实地调研研究院。

北京大学党委书记郝平实地调研研究院

2024.12.10

湖南省党政代表团访问北京大学，双方签署战略合作协议，研究院被列为湖南省和北京大学共建的重大创新平台。

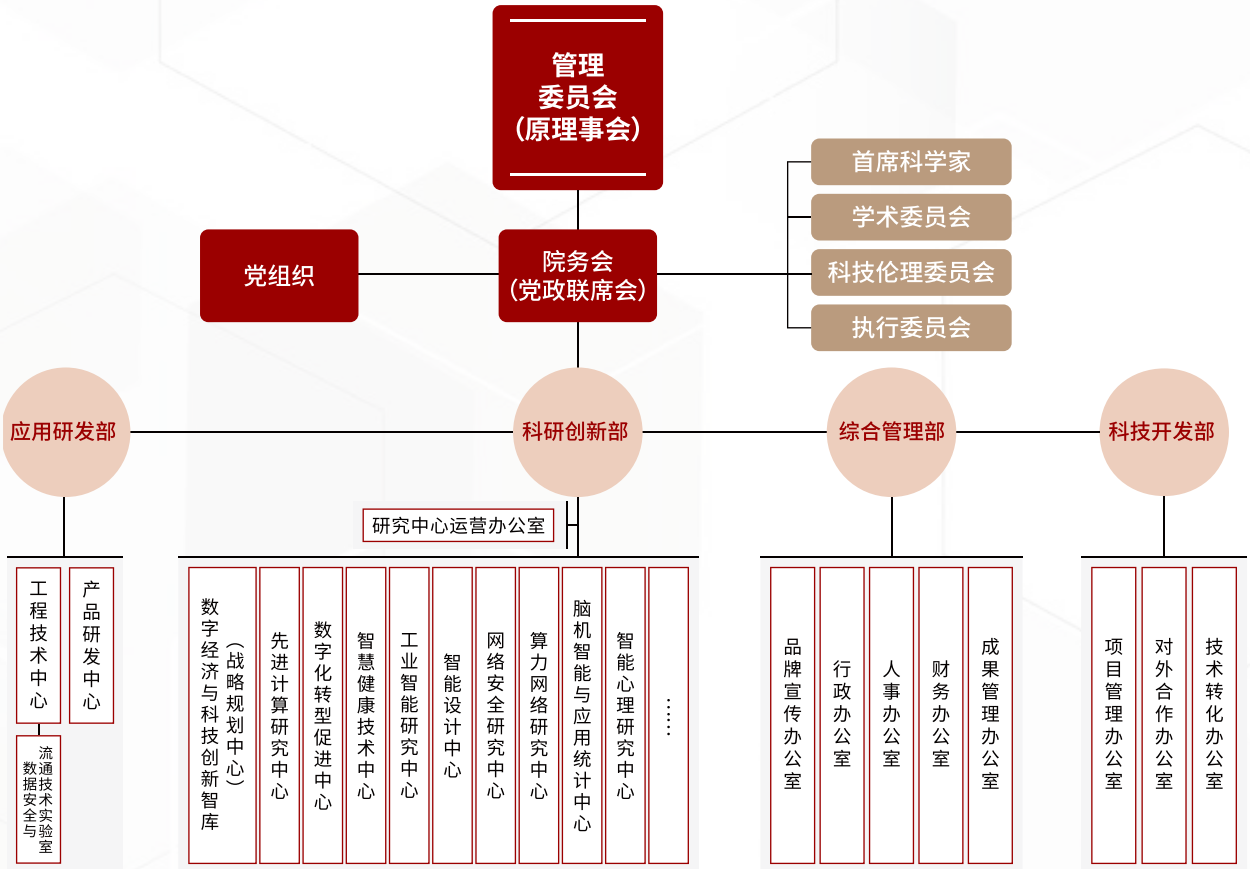
湖南省人民政府与北京大学签署战略合作协议

◆建设计划

北京大学长沙计算与数字经济研究院近期的建设与发展将充分发挥自身在人才储备和科研能力方面的优势，全面调动各方资源，结合湖南省、长沙市本地的实际需求，围绕体制机制创新、人才团队建设、应用研发落地三方面分阶段有序开展。

- 体制机制创新
- 人才团队建设
- 应用研发落地

◆组织架构



◆首席科学家

研究院聘请中国科学院院士、发展中国家科学院院士张平文教授和中国科学院院士、中国概率统计学会理事长陈松蹊教授共同担任研究院首席科学家。

张平文

中国科学院院士、发展中国家科学院院士。历任北京大学校长助理，副校长，党委常委、副校长等职务，现任武汉大学校长、党委副书记；北京大学博雅讲席教授，大数据分析与应用技术国家工程实验室主任，北京大学长沙计算与数字经济研究院首席科学家；中国工业与应用数学学会理事长；国家自然科学基金委“科学计算与机器学习”基础科学中心项目负责人。主要从事复杂流体的数学理论和计算方法，大数据分析与应用等方面研究，发表论文百余篇。曾获国家自然科学奖二等奖、教育部高等学校自然科学奖一等奖、何梁何利基金科学与技术进步奖、冯康科学计算奖等奖项。2011 年受邀在第七届国际工业与应用数学大会上作一小时报告，2014 年受邀在美国工业与应用数学年会上作大会报告，2018 年受邀在国际数学家大会作 45 分钟邀请报告，2020 年当选美国工业与应用数学学会会士。担任学术期刊 CSIAM Transactions on Applied Mathematics 主编，Applied Mathematics and Mechanics 等十余种国内外期刊编委。

陈松蹊

中国科学院院士，北京大学长沙计算与数字经济研究院首席科学家。现任中国概率统计学会理事长、中国统计学会副会长、美国科学促进会会士、美国统计学会会士、数理统计研究所会士。曾任北京大学光华管理学院商务统计与经济计量系主任、北京大学统计科学中心联席创始主任。曾任统计学顶级期刊《The Annals of Statistics》和美国统计学会会刊《JASA》副主编，目前任《Environmentrics》和《中国科学 (数学)》副主编。目前已发表论文 120 余篇，Web of Science H-index 33。曾获 Iowa State University 教员杰出研究奖、第七届厉以宁科研奖、2017 年教育部自然科学一等奖。研究成果得到了国内外同行的高度认可，入选斯坦福大学 2020-2022 榜单。

◆院领导

杨超

院长

- 教育部长江学者特聘教授
- 北京大学数学科学学院教授
- 北京大学大数据分析与应用技术国家工程实验室副主任

宋洁

副院长

- 教育部长江学者特聘教授
- 北京大学先进制造与机器人学院党委委员、书记

董彬

副院长

- 北京大学北京国际数学研究中心教授
- 北京大学国际机器学习研究中心副主任
- 北京大学国家生物医学成像科学中心研究员

黄卓

副院长

- 北京大学国家发展研究院副院长、教授
- BiMBA 商学院院长
- 北京大学南南合作与发展学院副院长
- 北京大学数字金融研究中心常务副主任

陈德良

副院长

- 中国国际技术转移经理人
- 长沙市C类人才(省市级领军人才)
- 中共长沙市科技新型社会组织行业委员会委员
- 长沙市科协第十三届委员会常务委员

王新民

秘书长兼院长助理

- 北京大学大数据分析与应用技术国家工程实验室党支部书记
- 高级工程师
- 高级信息系统项目管理师
- 管理咨询师

◆研究成果

研究院打造了一个智库和多个研究中心的科研创新体系。智库团队将为先进计算、数字经济与科技创新等产业的高质量发展提供政策建议和智力支持；各研究中心将积极培育科研力量，围绕相关领域的核心科学技术问题和重大应用难题，蓄力攻克“卡脖子”技术，加快人才链、创新链、产业链的融合对接，为长沙市、湖南省乃至中部地区计算产业和数字经济的发展提供新动能。

• 数字经济与科技创新智库

依托北京大学的学术影响力和国家高端智库平台，聚焦数字经济与高质量发展、科技创新、数字金融和科技金融、先进计算和前沿技术产业发展等重点领域，提供智库研究、政策建议和战略咨询服务，打造地方政府、产业界和学术界之间的对话交流平台，立足湖南、辐射中部、面向全国，为中央和地方政府提供战略性、前瞻性的智力支撑，致力于成为在数字经济和科技创新领域国内领先和中部地区最有影响力的智库机构。

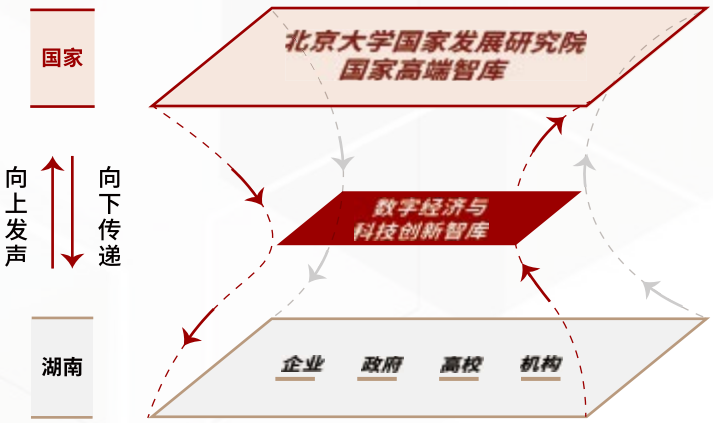
智库主任

黄卓 副院长

北京大学国家发展研究院副院长、教授
BiMBA 商学院院长
北京大学南南合作与发展学院副院长
北京大学数字金融研究中心常务副主任

斯坦福大学经济学博士，曾获得斯坦福大学经济系“最佳博士生候选人论文奖”。主要研究领域为数字经济与数字金融、金融计量、金融工程、大数据分析。主持多项国家自然科学基金和国家高端智库课题，研究成果发表在国内外一流期刊如 Journal of Econometrics 和《经济研究》《管理世界》等。2014 年获得应用计量经济学领域国际权威期刊 Journal of Applied Econometrics 的“Richard Stone 最佳论文奖”，2015 年获得第七届高等学校科学研究优秀成果奖 (人文社会科学) 论文类二等奖，2020 年获得北京大学教学卓越奖。目前还担任中国金融四十人论坛特邀研究员、中国衍生品青年论坛联席秘书长、数字经济开放研究计划首任秘书长。

建设中国特色新型智库



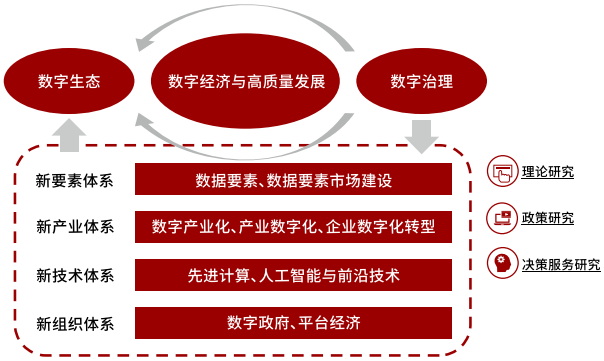
成果落地

- 专题研讨会
- 高端系列讲座
- 智库研究要报
- 重大课题研究报告

研究方向

1 数字经济与高质量发展

围绕数字经济与高质量发展，开展基于数字产业化与产业数字化、先进计算与前沿技术产业、企业数字化转型、数据要素市场建设、数字生态与数字治理、数字政府、平台经济等相关领域的政策研究与科学问题探讨，形成以理论指导实践、实践丰富理论的良性互动，为地区数字经济与高质量发展提供决策支撑作用。

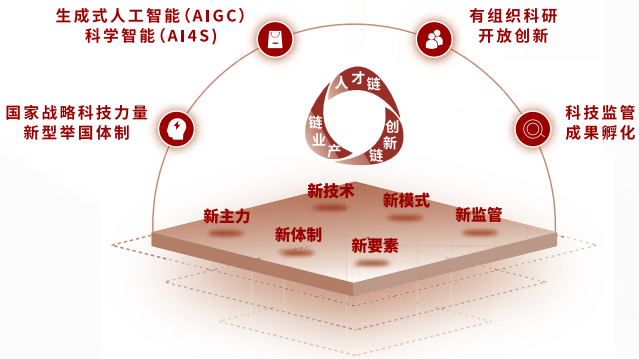


2 数字金融与科技金融

聚焦金融改革、发展和监管面临的突出问题，围绕科技金融、绿色金融、普惠金融、养老金融、数字金融等方向开展研究，引导金融更好地为重大战略、重点领域和薄弱环节提供优质服务，全面提升金融服务实体经济质效，为加快金融强国建设提供有力支撑。

3 科技创新与新质生产力

针对数字经济时代下科技创新的新特征、新趋势、新挑战，聚焦战略科技力量、创新体制机制、创新组织模式、创新资源利用、创新人才培养以及科研成果转化等方面进行理论前沿问题研究，并持续探索新时代科技创新的科学内涵，为科技强国建设、湖南科技强省建设提供战略支撑，为新质生产力跃迁升级提供战术指引，以催化新动能新优势的塑造与凝聚。



著作成果（部分）

新质生产力

平台经济通识

数字金融的力量：为实体经济赋能

The Digital Financial Revolution in China

• 先进计算研究中心

中心以先进计算技术推动社会经济发展为使命，旨在先进计算领域攻坚克难，研发适配国产硬件的关键技术和应用软件，构建新型国产计算应用生态，将领先算力高效转化为解决科学与工程问题的能力，满足国家和地区的重大战略需求和产业需求，支撑经济社会高质量发展。

中心主任

杨超 院长

教育部长江学者特聘教授
北京大学数学科学学院教授
北京大学大数据分析与应用技术国家工程实验室副主任

长期从事与超大规模并行计算相关的模型、算法、软件和应用研究，主持国家自然科学基金重点项目、重大研究计划集成项目、科技创新2030重大项目等国家级科研项目10余项，发表学术论文100余篇，研究成果先后获2016年国际计算机协会“戈登·贝尔”奖、2017年中国科学院杰出科技成就奖、2017年CCF-IEEE CS青年科学家奖、2018年茅以升北京青年科技奖、2020年首届王选杰出青年学者奖等国内外重要奖项。目前担任National Science Review编委、SIAM Journal on Scientific Computing编委，中国工业与应用数学学会高性能计算与数学软件专业委员会副主任，中国新一代人工智能产业技术创新战略联盟AI指令集与开发接口标准专题组组长，IEEE C/DC 2941标准工作组副主席等职。

研究方向

1 跨平台高性能多物理仿真技术

针对流体、结构、传热等复杂多物理场景中所面临的大时空跨度、强非线性和超大规模计算等关键技术挑战，面向国际主流商用处理器和国产新型众核处理器，开展并行求解算法的核心技术攻关，研发新一代工业仿真设计CAE计算引擎，提升智能制造、能源双碳、医疗健康、智慧城市和航空航天等领域的工业仿真设计能力，赋能“战略性新兴产业”和“未来产业”，助力我国高水平科技自立自强。

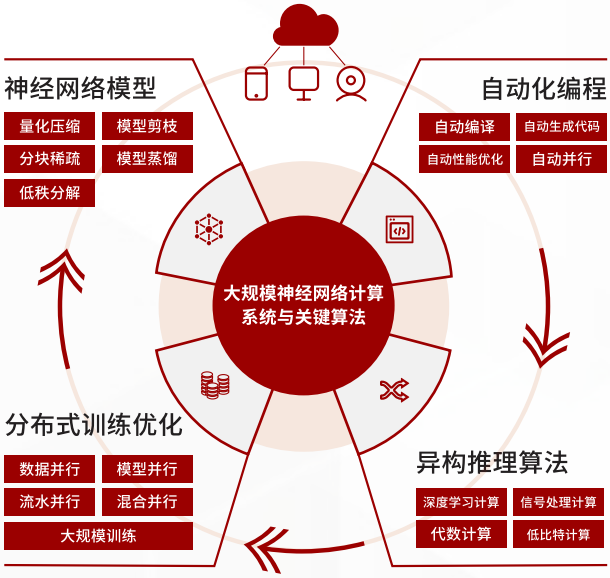
研究方向

2 高性能计算库与接口标准化技术

针对科学计算、大数据、人工智能、无线通信等领域所涉及的基础数学、线性代数、张量代数等各类基础性计算问题，面向CPU、GPU、NPU等主流硬件平台，屏蔽底层硬件细节，制定科学、统一的算子与算法接口标准，并在此基础上研发高性能计算库，推动行业技术共同进步，为国家重点科研、应用领域提供有力的技术支撑。

3 大规模神经网络计算系统与关键算法

针对大规模神经网络模型计算效率低、能源消耗大等问题，结合云、边、端异构计算系统在体系结构、网络连接等方面的不同特性，研发计算、访存、通信效率高度优化的神经网络模型、分布式训练算法和异构推理算法等，并结合国产计算芯片的特点，通过自动编译、代码生成等手段降低编程门槛，提升运行效率，推动国产人工智能算力的普及应用。



• 数字化转型促进中心

中心积极响应“十四五”规划和2035年远景目标纲要，致力于企业数字化转型和传统产业升级的国家战略，通过对典型行业的数字化转型研究，深入理解数据要素的内涵，探索数字经济及相关领域产业应用的实践路径，建设成为国内外一流的数字化转型咨询机构和产业孵化基地，助推中国经济的高质量发展。

中心主任

宋洁 副院长

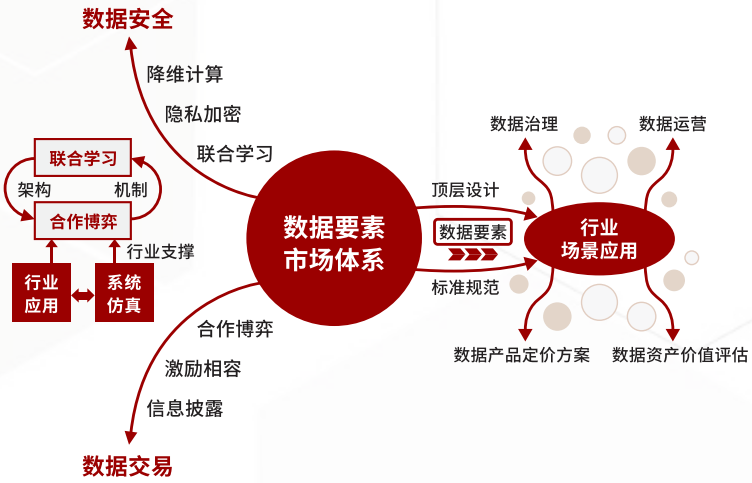
教育部长江学者特聘教授
北京大学先进制造与机器人学院党委委员、书记

主要研究领域为在线学习、仿真优化及在系统工程领域的应用，承担国家自然科学基金创新群体、重大、重点等多项课题，主持国家发改委、国家粮食局、国家电网等单位重点研究课题。国际电气和电子工程师协会高级会员，担任 IEEE 国际自动化管理技术委员会主席 (首任亚洲学者)，同时入选 IEEE 杰出报告人。曾先后入选教育部青年长江学者，长江学者特聘教授。研究获得日内瓦国际发明金奖，中国电工技术学会科技进步二等奖，IEEE-TASE 期刊最佳论文奖，IISE 国际工业与系统工程协会最佳论文荣誉奖，工程管理教指委“最佳指导导师奖”，北京市高等教育教学成果一等奖，多次获得北京大学教学优秀奖和教学改革优秀奖。

研究方向

1 数据要素市场体系

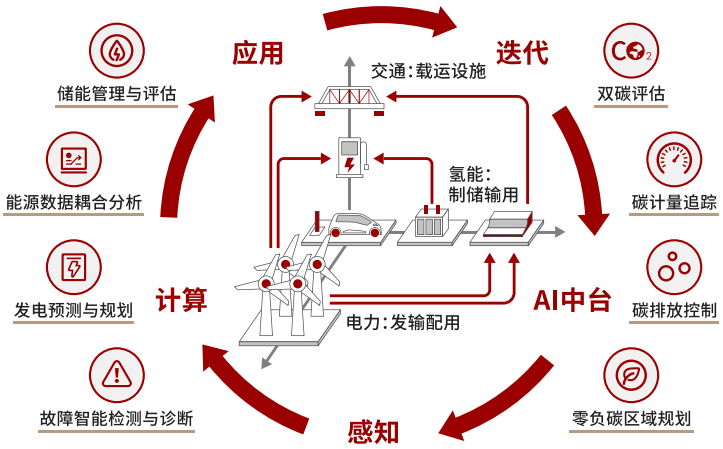
围绕数据要素市场，开展数据要素市场配置、数据交易、数据定价机制等理论与政策研究，应用大数据分析、人工智能算法等手段，深度挖掘数据资源价值，提升数据资源利用效率和社会价值，推动数据交易中心或平台建设，为数据资产化、商品化提供有效载体。充分发挥数据要素在经济发展中的催化作用，推动经济社会全面数字化转型和高质量发展。



研究方向

2 能源大数据与双碳应用

聚焦“数字-能源-经济”协同发展的理论和应用研究，充分利用人工智能、大数据、云计算和物联网等先进技术，开展对全国资源禀赋、产业结构及能源结构的全方位分析，实现对能源、环境、交通、工业生产等复杂过程的精准建模，研究低碳转型路径，为政府制定双碳政策提供理论依据，为能源企业低碳转型提供决策依据及技术支持。



3 人工智能技术赋能企业数字化转型

面向本土产业发展和企业数字化转型需求，研究数字化转型最新理论与先进企业成功案例，提炼转型路径，挖掘大数据和人工智能等前沿技术在数字化转型各阶段的应用场景，自主研发基于大模型的智能化综合服务平台DataAgent，进一步整合生态资源，提供包含评级诊断、顶层规划、数据治理、技术咨询、平台开发及落地运维的全生命周期陪伴式服务，助力传统企业向数智企业转型升级。

• 智慧健康技术中心

中心聚焦医疗健康领域的前沿问题，以临床需求为牵引，凝练、总结、归纳关键科学和数学问题，发挥北京大学数学、机器学习和跨学科交叉的优势，开展面向智慧医疗的前沿算法研究和软件系统的开发，联合国内各大医院，针对性地解决临床上的“卡脖子”问题，研究多模态生物医学数据前沿算法，研发精准诊断的智能辅助系统。

主要成员

中心主任

董彬 副院长

北京大学北京国际数学研究中心教授
北京大学国际机器学习研究中心副主任
北京大学国家生物医学成像科学中心研究员

2003 年本科毕业于北京大学数学科学学院，2005 年在新加坡国立大学数学系获得硕士学位，2009 年在美国加州大学洛杉矶分校数学系获得博士学位。博士毕业后曾在美国加州大学圣迭戈分校数学系任访问助理教授，2011 — 2014 年在美国亚利桑那大学数学系任助理教授，2014 年底入职北京大学。主要研究领域为科学计算、机器学习及其在计算成像和数据分析中的应用。2014 年获得求是杰出青年学者奖，2022 年受邀在世界数学家大会（ICM）做 45 分钟报告，2023 年入选新基石研究员项目。

医学顾问

沈琳

北京大学肿瘤医院主任医师、教授
北京学者、北京市突出贡献专家

唐磊

北京大学肿瘤医院主任医师、教授
医学影像科副主任

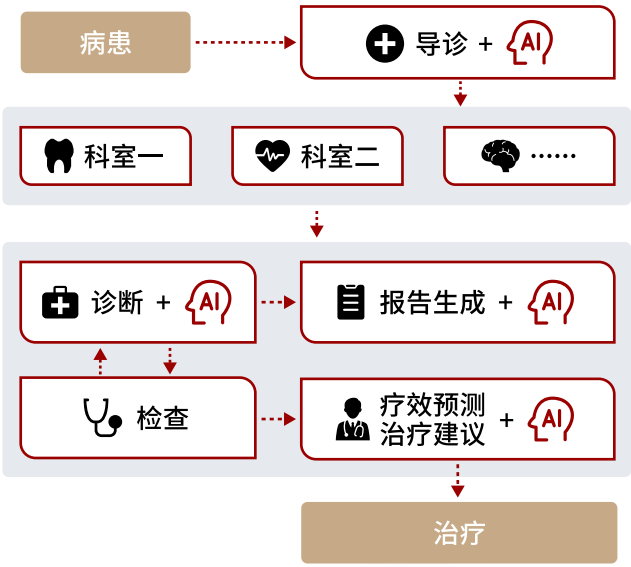
宋伟涛

中南大学湘雅医院副主任医师
眼科中心党支部书记、副主任

研究方向

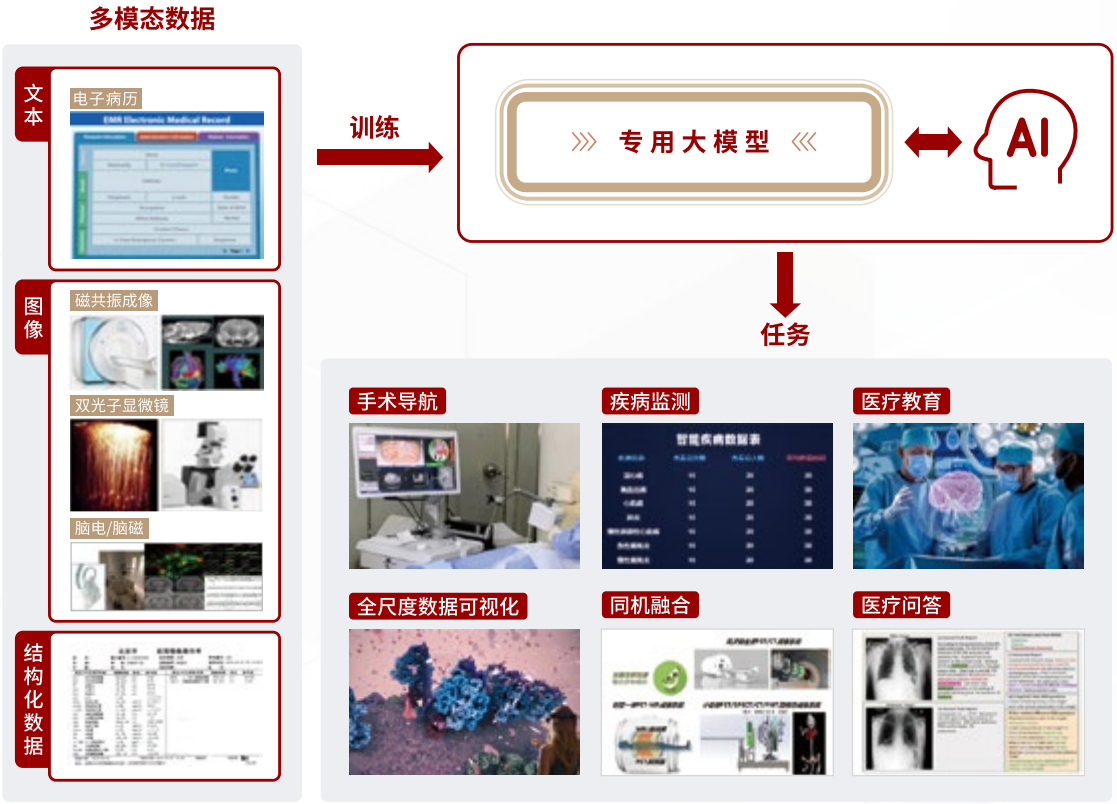
1 面向医疗领域的垂类大语言模型

聚焦医疗资源短缺问题，以医疗大语言模型为核心，研发与推广智慧医疗助手，辅助医生诊疗，同时为患者提供更高效、更精确的治疗建议，实现全流程、全周期的智慧诊疗。



2 多模态机器学习前沿算法

集中在多模态生物医学数据分析等研究，促进机器学习前沿算法在计算机视觉、自然语言处理和生物医学等领域的应用。



• 工业智能研究中心

中心坚持以服务国家战略、国计民生及企业数智化转型为使命，面向智能制造、供应链、交通运输、消费零售、芯片研发等领域的智能决策需求，搭建基于运筹优化算法及人工智能关键技术的智能决策产品“智解 - 智算 - 智问”，赋能产业数智化转型升级与发展。

中心主任

文再文

教育部长江学者特聘教授
北京大学北京国际数学研究中心教授
北京大学工学院工业工程与管理系主任

主要研究最优化算法与理论及其在机器学习和人工智能中的应用，先后获得国家自然科学基金委优秀青年科学基金、中组部青年拔尖人才计划、中国青年科技奖、国家万人计划科技创新领军人才和北京市杰出青年中关村奖。主持国家自然科学基金重点项目等多个课题，发表学术论文 70 余篇，与合作者一起开发有多重网格优化软件包 MGLS，稀疏优化软件包 FPC_AS，低秩矩阵恢复软件包 LMaFit，保正交约束软件包 ARNT，半定规划软件包 SSNSDP 和大规模特征值分解软件包 Arrabit。2021 年由高教出版社出版教材《最优化：建模、算法与理论》和《最优化计算方法》，得到了北京大学、清华大学、复旦大学和中国科技大学等超百所高校采用，截至 2023 年 5 月累计印刷 5 次 2 万余本。

研究方向

聚集三大核心产品的工业智能决策平台：

1 智解：智能决策计算引擎

研发具有自主知识产权的高性能、高扩展性的数学规划求解器，内置丰富的约束条件和目标函数，支持用户自定义复杂的业务场景及决策问题，满足不同场景下大规模复杂问题的优化求解，为上层业务模型求解提供计算引擎。

研究方向

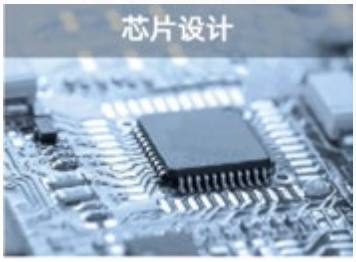
2 智算：智能决策大脑

面向智能制造、数字供应链、能源电力、交通运输、金融等领域的复杂决策场景，通过问题分析、建模、求解、决策，帮助企业解决大规模复杂的决策问题，赋能管理优化和流程优化，从而达到企业价值和效率的指数级提升。

3 智问：智能决策大模型

探索研发基于大模型与运筹优化技术结合的智能决策轻量化应用，通过与用户交互，动态理解用户的意图，选择或自动生成业务问题模型及优化求解算法，实现业务场景与人工智能算法应用的无缝衔接，让每个人都能作出智能的决策。

CNC精密加工智能排产APS



EDA工具链建模、优化与仿真技术

物流配送多式联运优化技术

• 智能设计中心

中心聚焦开展智能设计领域的战略咨询和技术研发，重点研究智能设计对长沙市、湖南省乃至全国计算与数字经济与社会经济发展环境等的关系问题，探索实现智能化目标的技术路径，将智能设计中心建设成为长沙智能设计领域综合服务平台。智能设计中心具备服务企业技术需求、科技成果转移转化、产业项目孵化、政府决策咨询、公共技术支撑五大功能。

中心主任

袁晓如

北京大学智能学院研究员
北京大学教材建设博雅特聘教授
大数据分析与应用技术国家工程实验室常务副主任

研究方向包括复杂流场数据可视化，高维/时空数据，交通、社会媒体数据的分析，可视化的快速构建方法。高动态范围可视化的工作获得2005年IEEE VIS大会最佳应用论文奖。2013年来指导实验室团队7次在IEEE VAST可视化分析挑战赛中获奖。数十次担任IEEE VIS, EuroVis,IEEE PacificVis等国际可视化会议程序委员会委员，2017年IEEE VIS大会论文共同主席（SciVis），创建中国可视化与可视分析（ChinaVis）大会。中国计算机学会理事，杰出会员，杰出讲者。中国计算机学会大数据专家委员会委员，人机交互专委会常务委员和计算机辅助设计与图形学专委会委员。中国图象图形学学会理事、可视化与可视分析专业委员会主任。

研究方向

• 网络安全研究中心

中心旨在探索构建新形势高对抗下的网络安全新体系，为国家安全贡献力量，以建设成为业内前沿技术的策源地、顶端人才的聚集地，服务国家重大战略需求，并以科技创新为核心动力促进网络安全产业发展。

主要成员

中心主任

张蓓

原北京大学计算中心主任
CERNET专家委员会委员

长期从事计算机网络、信息安全、计算机应用等领域的研究和建设工作，承担国家重大科研项目 50 余项，发表学术论文 100 余篇，获得国家科技进步二等奖 1 项，部级科技进步一等奖 1 项、二等奖 2 项。自 2001 年起担任 CERNET 专家委员会委员、北京大学计算中心副主任，2008 年起担任北京大学计算中心主任、网络中心主任，负责 CERNET 北大主节点、CERNET2 北大核心节点、北京大学校园网、智慧校园、网络安全、高性能计算、大数据分析、数据中心等领域的建设和管理工作，视野开阔，经验丰富，成效显著。

中心副主任

周昌令

- 北京大学高级工程师
- 北京大学计算中心网络安全室主任

研究领域

网络安全、攻防对抗、网络流量分析、网络管理、机器学习等

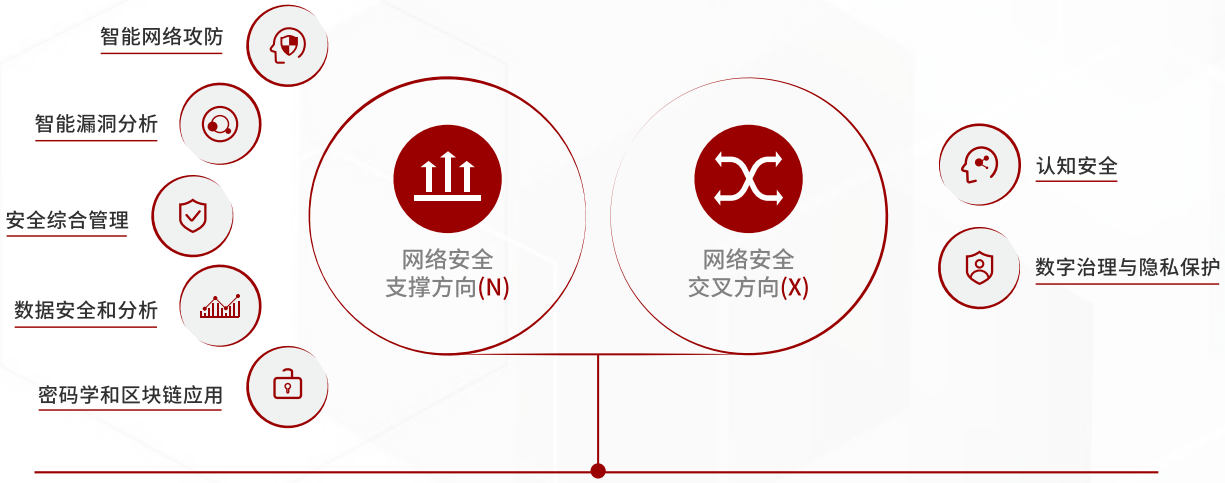
文伟平

- 北京大学教授
- 北京大学软件安全研究小组首席专家

研究领域

系统与软件安全、大数据与云安全、区块链安全、智能计算安全等

研究方向



按照“网络安全支撑N+网络安全交叉X”两个主要方向来建设组织科研业务开展

• 算力网络研究中心

中心旨在通过突破算力网络领域的核心关键技术，建立标准化的算力中心管理范式，打通各算力中心间管理、使用、结算壁垒，实现算力资源跨域统一调配，建立一体化算力网络大市场，进而优化算力资源市场化配置，激活算力网络市场生态，满足国家新型基础设施建设需求，推动我国数字经济高质量发展。

中心主任

樊春

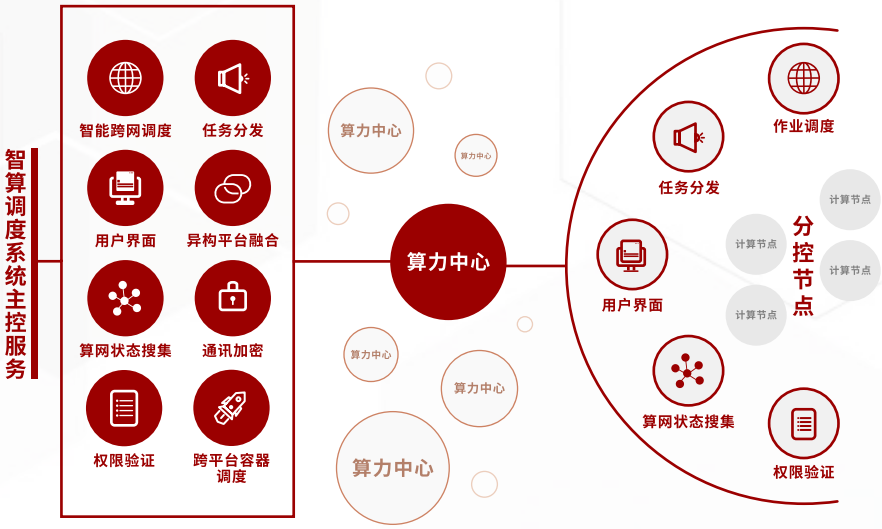
北京大学计算中心系统管理室主任、正高级工程师
北京大学高性能计算平台主任工程师
中国计算机学会高性能计算专业委员会执行委员

研究工作主要涉及高性能计算、人工智能算子、计算流体力学和纳米材料计算等领域。担任中国工业与应用数学学会高性能计算专业委员会委员，开放科学计算联盟副理事长。主持建设北京大学校级公共平台多套高性能计算集群，领导建设了算力中心一体化运维管理系统，形成了一套标准化的算力中心管理范式，并在多所高校和科研院所集群中推广应用。主持和参与了科技创新 2030- 新一代人工智能重大项目、国家重点研发计划、863 计划、发改委重点专项、“十三五”国家重大科技基础设施、广东省重点领域研发计划等多个项目。作为北京大学超算队指导老师，带领学生获得第十届、第十一届 ASC 世界大学生超算竞赛冠军，SC20 国际大学生超算竞赛第二名，SC23 国际大学生超算竞赛第二名、LINPACK 基准性能测试最高性能奖等优秀成绩。

研究方向

1 算力网络融合技术

中心针对算力资源碎片化的现状，聚焦算力网络融合技术，围绕算力中心标准化管理、算力资源跨域融合、算网市场建设等问题开展研究。拟通过可视化交互式门户平台的使用模式、开源中立管理平台的算力资源接入模式、连接算力资源供需两端的交易模式，解决算力网络市场中使用、接入、交易三个关键问题。建成算力网络统一大市场，激活算力网络市场生态，促进算力网络经济发展。



研究方向

2 算力网络资源调度技术

中心计划研究算力网络中的资源智能调度技术。面向算力网络中的超算资源、AI智算资源和云计算资源等多种算力资源，研制能够支持跨域互联、可伸缩、高性能、高可靠、高可用、易部署的智能调度系统，实现对百万核心以上规模算力资源的调度能力。通过多维度统一智能调度，打通算力孤岛，提高算力资源综合使用率，促进数据要素流通应用，推进我国新型基础设施高质量发展。

• 脑机智能与应用统计中心

中心旨在推动脑科学和应用统计两个领域的科研成果转化和落地。脑科学领域聚焦于非侵入式脑机接口技术、分析算法和采集设备的研发和推广，以癫痫的人工智能辅助诊断为出发点，积极探索痴呆、抑郁、睡眠等领域的关键需求和解决方案。应用统计领域旨在打通从数据到数据赋能之间的分析瓶颈，挖掘数据中隐含的生产力，重点关注粮食安全、能源安全、低碳环保和先进制造行业的数据科学服务，打造典型案例助力数字中国建设。

中心主任

陈松蹊

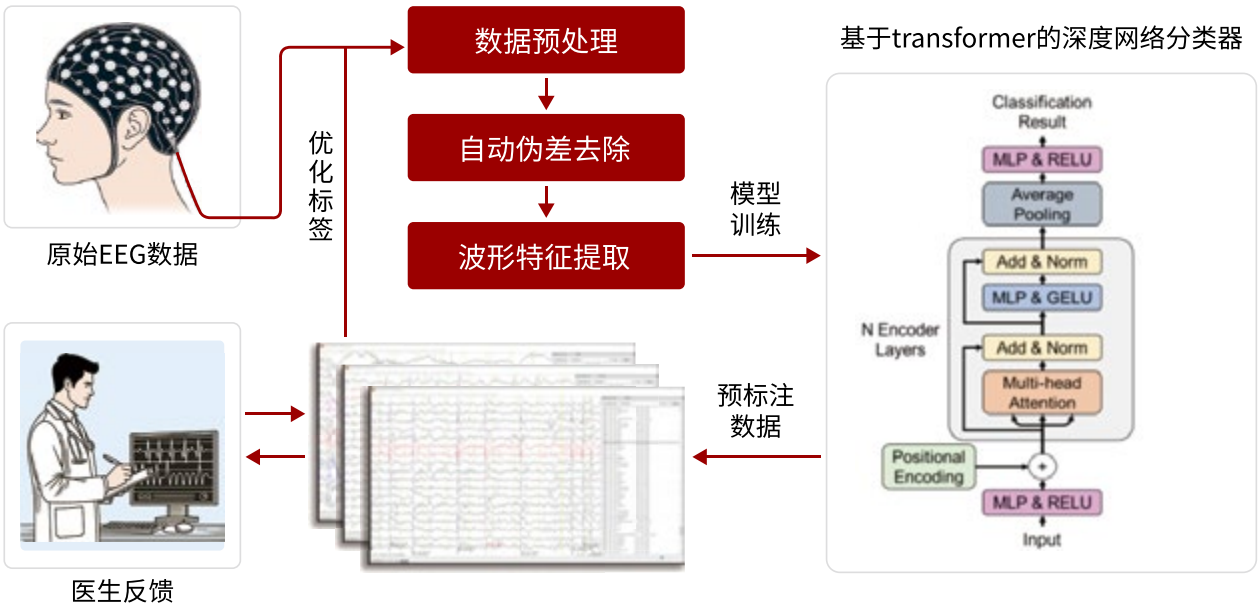
中国科学院院士
中国概率统计学会理事长
北京大学长沙计算与数字经济研究院首席科学家

中国科学院院士，北京大学长沙计算与数字经济研究院首席科学家。现任中国概率统计学会理事长、中国统计学会副会长、美国科学促进会会士、美国统计学会会士、数理统计研究所会士。曾任北京大学光华管理学院商务统计与经济计量系主任、北京大学统计科学中心联席创始主任。曾任统计学顶级期刊《The Annals of Statistics》和美国统计学会会刊《JASA》副主编，目前任《Environmentrics》和《中国科学(数学)》副主编。目前已发表论文 120 余篇，Web of Science H-index 33。曾获 Iowa State University 教员杰出研究奖、第七届厉以宁科研奖、2017 年教育部自然科学一等奖。研究成果得到了国内外同行的高度认可，入选斯坦福大学 2020-2022 榜单。

研究方向

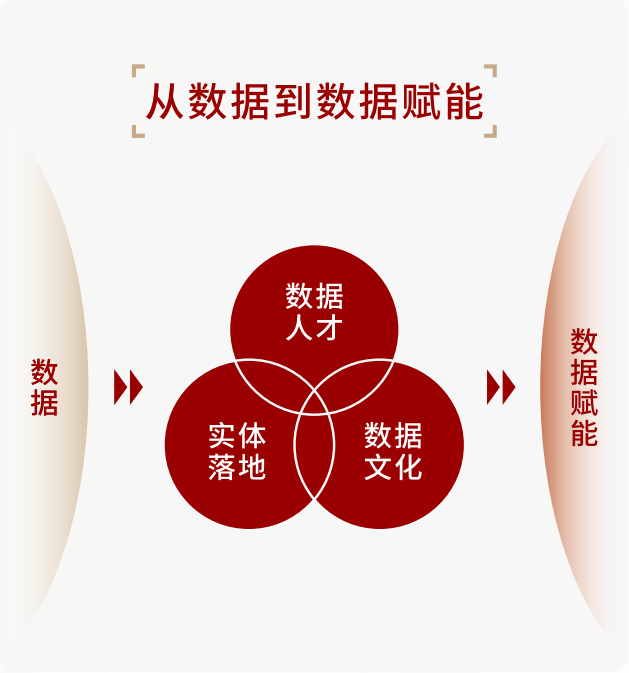
1 非侵入式脑机接口智能分析技术

脑电图(EEG)、脑磁图(EMG)在脑部疾病特别是癫痫诊断中起到核心作用。针对癫痫诊疗全流程中所涉及的癫痫发作时段识别、发作预测、发作间期放电检测、病灶溯源定位和病种分析等各类医疗需求，研发完整的AI智能诊疗方案，并在此基础上提升诊疗效率，推动优质医疗资源下沉，提高脑机接口技术在患者中的可触达性。



2 数据赋能与统计分析

依据最新的统计学方法论，将数据分析与企业销售、管理、供应链深度融合，帮助实体企业实现数据的价值，并在数据科学、数据分析层面助力企业数据文化建设。围绕从数据采集、数据分析到管理决策的全流程服务，培养专业的统计师和数据分析师，加强与大数据、人工智能学科群的融合与创新。



• 智能心理研究中心

当前我国面临严峻的心理健康挑战，专业人才短缺、资源分布不均及标准缺失等问题突出。智能心理研究中心充分利用 AI 技术优势，精准对焦，为优化我国心理健康服务体系提供有效方案，加速推动体系现代化与标准化进程。我们致力于提升国民整体心理健康水平，构建符合中国国情的特色心理健康服务体系，以科技赋能，守护每一颗心灵。

中心主任

臧寅垠

北京大学心理与认知科学学院
研究员
中国心理学会注册督导师
英国心理学会特许心理学家、
副会士

研究领域为临床心理学、创伤后应激障碍（PTSD）及焦虑相关障碍、儿童青少年抑郁焦虑等心理疾病的临床循证干预研究。在中、英、美三国主导或参与了多个临床心理学研究项目，包括于我国地震灾区开展的对成人及儿童创伤后应激障碍的治疗的临床研究，以及多个大型多中心临床干预试验等。已发表 SCI/SSCI 收录英文期刊论文 40 余篇，并承担多个国家级科研项目。

研究方向

1 多模态心理评测与人机互融干预

融合多种数据源，利用跨模态深度学习模型，全面分析和理解个体的心理状态，实现高度精确的评测。基于动态评估结果，设计和实施个性化的心理干预方案。充分利用AI的自适应性、互动性和灵活性，满足不同人群的特异性心理服务需求。通过大规模数据分析和临床试验验证干预措施的有效性和安全性，为心理健康服务提供强有力的循证支持，推动心理健康服务的精准化和智能化发展。

研究方向

2 情感计算与情感智能

结合心理学理论与先进的AI模型，通过深度学习和神经网络分析，提升AI理解和解读复杂人类情感状态的能力。增强AI情感识别的准确性，提升其共情能力，加强情感交互表现。研究并探索如何在医疗、教育、工程等多种场景中有效应用情感智能技术。通过不断的学习和优化，推动人类与AI在情感层面的深度融合，带来更加智能和人性化的未来。

AI在心理健康状况中的应用

压力水平

问题识别技巧

问题解决能力

思维能力

心理健康状况检测

逻辑推理

基于AI技术的应用

数据收集和结构化

特征提取与表征

精神疾病的分类

实时监控和干预

检测潜在生物标志物

• 数据安全与流通技术实验室

数据已成为重要的国家战略资源和推动经济发展质量变革、效率变革、动力变革的新型生产要素，实验室以安全促发展，积极构建大数据安全技术保障体系，研发数据安全与数据要素流通关键技术、重要产品与核心平台，建设数据安全流通基础设施，参与行业及国家标准规范制定，促进数据要素有序流通、释放数据价值、护航数字经济发展。

实验室负责人

徐克付

北京大学大数据分析与应用技术国家工程实验室特聘研究员
中国人工智能学会智慧能源专业委员会委员
中国医疗保健国际交流促进会健康大数据与数字化分会委员

从事大数据、人工智能领域、网络安全领域的技术研究与工程研发十余年；作为项目负责人完成了国家重点研发计划、国家发改委重大专项、中国科学院战略性先导专项等十余项项目的研究及开发工作；长期支持国家部委大型工程项目的建设及国家网络安全保障工作，在国家网络安全重大专项中表现优异，连续三年获中央领导表彰；在国内外学术期刊及国际会议上发表论文 50 余篇、申请发明专利十余项、参与十余项行业及国家标准制定、《国家网络安全审查战略研究》入选 2015 乌镇世界互联网大会十大重要成果，《全周期云数据安全管控与应用支撑平台》获教育部科技进步一等奖。

研究方向

1 大数据生命周期安全防护与管控技术

随着人工智能、云计算、移动互联网和物联网等技术的融合发展，传统的基于边界安全域和基于已知特征库的网络安全防护方式已经无法有效应对大数据环境下新的安全威胁。针对数据生命周期的安全防护问题，综合利用数据源验证、访问控制、敏感数据识别、自动化数据分类分级、数字资产管理、流转管控、追踪溯源与态势感知等技术，建立数据生命周期安全纵深的防御体系；借助大数据分析、人工智能等技术，实现自动化威胁识别、风险阻断和攻击溯源，实现从被动防御到主动检测的转变，提升大数据安全防御水平，保障企业、政府核心数据资产与国家关键信息基础设施安全。

隐私安全：去标识化 | 匿名化 | 数据脱敏 | 数据失真

数据安全：分级分类 | 数据隔离 | 追踪溯源 | 质量管理 | 数据加密 | 数据销毁 | 元数据管理 | 数据防泄露

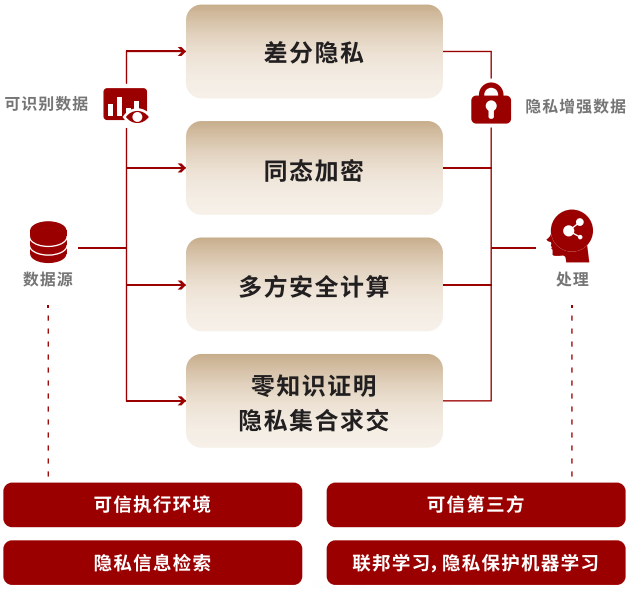
平台安全：接口鉴权 | 访问控制 | 身份认证 | 配置管理 | 传输安全 | 备份恢复 | 访问鉴权 | 安全审计

大数据基础软件：hadoop | hive | HBASE | Spark | STORM | kafka

研究方向

2 面向数据要素流通的隐私计算技术

隐私计算技术是在保护数据本身不对外泄露的前提下，多个参与方通过协同对自有数据处理、联合建模运算、分析输出结果、挖掘数据价值的一类信息技术。针对数据流通中的个人信息保护、权益分配、数据安全保障、追溯审计等问题，研究通过隐私计算技术解决数据有效流通和价值释放问题，在实现数据价值挖掘的同时，真正做到数据“可用不可见”；通过对多方安全计算技术、联邦学习技术、可信执行环境技术的研究，研发隐私计算技术平台，为政务、金融、医疗等领域个人隐私保护、数据合作共享、数据联合建模、公共数据开放等数据共享应用场景提供支撑，赋能数据安全保护及数据要素价值释放。



3 数据跨主体、跨域、跨行业可信流通技术

加快建设数据流通基础设施对推动数据要素市场全面发展至关重要，针对数据流通过程中面临的困难挑战，聚焦基础设施的总体功能架构、网络结构、关键技术、应用场景、规范保障等内容展开研究工作。针对数据跨主体、跨域、跨行业流通与应用需求，基于“隐私计算+区块链”技术，研发自主知识产权、安全、可信、可控、可追溯的数据流通产品、平台及服务，支撑大规模数据、众多数据方和复杂共享流通应用场；研究推动隐私计算跨平台的互联互通，构建产业级、国家级的数据流通基础设施，实现跨企业、跨行业、跨区域数据要素市场化流通，推动建设隐私计算与数据流通产业联盟与产业生态。

◆联系方式

地址：湖南省长沙市岳麓区尖山湖路2号
邮箱：contact@icode.pku.edu.cn
网站：https://icode.pku.edu.cn

公众号二维码

北京大学人文社科数智化研究基地

◆基地简介

北京大学人文社科数智化研究基地获批成立于 2022 年 4 月。基地旨在将数字化、智能化手段融入传统人文社会科学的研究，塑造新范式，提供新视角、新方法和新工具，以战略性、前瞻性和系统性思路开展有组织的研究，推动人文社会科学在数字时代的发展与传播，提升数字中国的关键能力，形成核心竞争力。基地将建设成为学科建设的助推器、战略智库的数据池和文理交融的汇合地，促进学术交流，服务国家战略，为国家决策提供科学支撑，推动数字中国建设。这是数字时代的深谋远见。基地由北京大学大数据分析与应用技术国家工程实验室牵头，汇聚北京大学相关领域团队、虚实体机构入驻，目前已筹建入驻 6 个实验室：数字生态与数字治理实验室；区域与国别数据库实验室；国际组织数据库实验室；古籍数字化实验室；语言学实验室；图像学实验室。

▲ 北京大学人文社科数智化研究基地效果图

▲ 学科交叉创新平台

◆实验室介绍

• 数字生态与治理实验室

数字生态与治理课题组是北京大学大数据分析与应用技术国家工程实验室的智库研究团队，旨在开拓探索“数字中国建设”战略实施的理论基础，深化剖析“数据基础制度”基本问题，重点部署“数字化转型”工程任务，从而更好服务大数据技术赋能国家治理体系和治理能力现代化。课题组由实验室主任张平文院士指导，实验室核心成员邱泽奇、宋洁两位教授主持，成员除了实验室相关研究员与博士后外，还广泛联结了社会学系、工学院、国发院、国际关系学院、政府管理学院、经济学院等众多北大教授团队。

实验室负责人

张平文

武汉大学校长、党委副书记
北京大学博雅讲席教授
大数据分析与应用技术国家工程实验室主任

邱泽奇

北京大学中国社会与发展研究中心主任
社会学系教授、博士生导师
数字治理研究中心主任
教育部“长江学者奖励计划”特聘教授
北京大学博雅特聘教授

宋洁

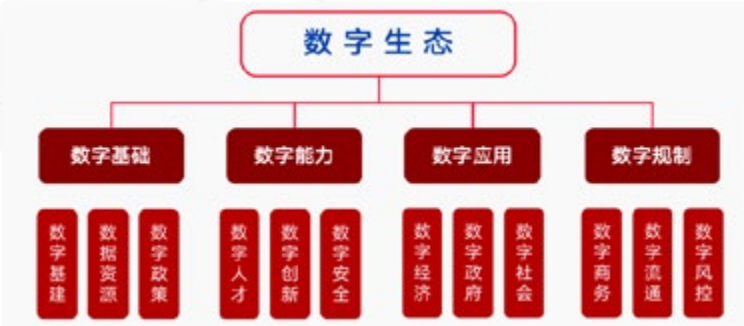
北京大学先进制造与机器人学院党委委员、书记
教育部长江学者特聘教授
北京大学长沙计算与数字经济研究院副院长

科研成果

近几年来，课题组承担了国家发改委、中央网信办、国家自然科学基金、国家社会学科基金、中国科学院学部等委托的系列重大课题，支撑了中共中央、国务院《数字中国建设整体布局规划》文件起草、国家发改委数字化转型伙伴行动倡议、数据要素流通与治理的机制与政策研究等重大任务；培养了一批博士后、博士生、硕士生等科研人才，发表了一批具有影响力的期刊论文与学术专著；参与了诸多地方政府与行业企业的数字化转型工程项目，赋能了我国政府、企业和社会数字化发展过程。

“数字生态”是课题组开创并持续深化研究的新理论与新视角。我们将数字生态定义为数字主体互连互动形成的多层次多结构多形态且彼此依存、合作和冲突的动态系统，数据是数字生态的核心要素，政府、企业、社会是数字生态的行动主体，数字化转型是数字生态的发展过程。

“数字生态指数”是评估全球和国内各地区数字经济发展的科学工具，包括数字基础、数字能力、数字应用、数字规制四个一级指标，12 个二级指标，测量指标来自“数字生态协同创新平台”合作单位研制的相关指数。



“数字生态协同创新平台”是实验室为数字生态理论与指数研究搭建的共创联合体，广泛联合了清华大学等高等院校，中国经济信息社等产业服务机构，国家电网等企业机构，以及滴滴、阿里云等互联网公司，目前包括全国 30 多个指数团队。

2020-2024 年间，数字生态指数已连续 5 年发布，对 31 个省（自治区、直辖市）、360 多个地级市以及 150 多个国家和地区进行了追踪评估。

研究成果一览：

·指数报告：

《数字生态指数 2020》、《数字生态指数 2021》、《数字生态指数 2022》、《数字生态指数 2023》

·指数网站：

http://digiteco.com.cn

·学术论文：

《中国科学院院刊》3 篇、《电子政务》3 篇和《西安交通大学学报（社会科学版）》2 篇

·学术著作：

《数据要素的特征：信息、权属、价值、安全、交易》、《数字生态与治理（第一辑）》

·党建读物：

《数字经济干部读本》、《数字经济学习问答》

·标准起草：

《数据交易服务指南》、《数据交易安全评估指南》、《数据资产登记指南》、《数据资产质量评估》、《数字化转型 成熟度模型》、《数字化转型 成熟度评估指南》、《基于码链融合的数字身份应用框架》等

·实验室参加的标准组织：

全国信息化和工业化融合管理标准化技术委员会数字化转型标准工作组（TC573WG16）和数字经济标准工作组（TC573/WG9）；

北京市数字经济标委会数据交易工作组；中国电子工业标准化技术协会；中信联—数字化转型成熟度评估工作委员会。

• 区域与国别数据库实验室

区域与国别数据库实验室旨在推进学科深度交叉融合，打造以数据算法和人工智能为底层架构的“综合区域国别研究”集群以及“Global Mind 全球实时开放数据智能平台”，赋能社会科学量化研究，推进区域国别学一级学科建设。实验室秉持“学以致用，用以强学”原则，大力开展智库建设，实现学术研究和应用研究的高效转化，服务国家战略。

实验室负责人

张平文

武汉大学校长、党委副书记
北京大学博雅讲席教授
大数据分析与应用技术国家工程实验室主任

王腾蛟

北京大学计算机学院教授、博士生导师
国家教育部“长江学者”特岗学者
北京大学计算社会科学研究中心主任

翟崑

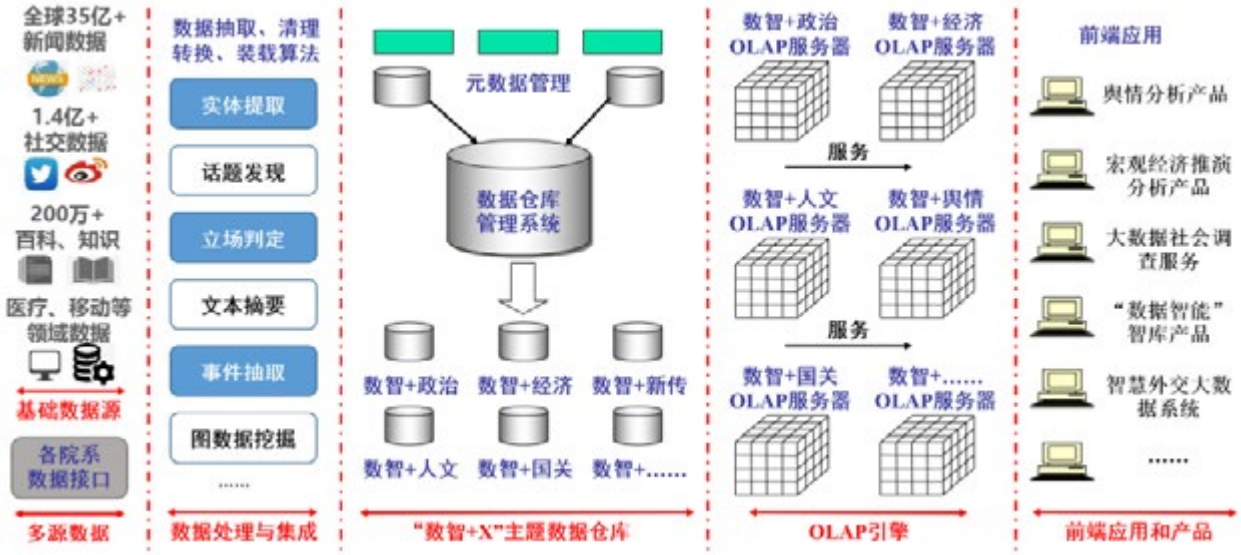
北京大学国际关系学院教授，博士生导师
北京大学区域与国别研究院副院长
北京大学海洋研究院全球互联互通研究中心主任
北京大学东盟国家研究中心主任
全球湄公河研究中心（中国中心）副主任
北京大学国际战略研究院特约研究员

科研成果

· Global Mind 全球实时开放数据智能服务平台　项目负责人：王腾蛟

Global Mind 以智能数据分析为导向，通过构建多主题数据仓库，实现多源异构数据的统一建模、负载均衡和访问机制，为海量数据统一存储和管理提供支持。该系统现已集成海量全球开放数据，实现“计算 +X”多主题数据仓库和多维分析引擎，为开展多领域定制化数据分析和智能决策提供了有效的数据管理支撑。

系统实时采集并提取全球主要国家和地区的新闻媒体、社交媒体等数据，累计数据量已超过 40 亿条。该系统为研究人员提供了宝贵的数据资源，促进了跨学科研究的开展。系统能够实时主动捕捉数据的结构、模式和内容的变化，通过主动学习构建预测模型。系统以“主动感知”为核心，构建了主动感知的数据提取、主动感知的数据管理和主动感知的数据分析创新技术体系。帮助企业和机构高效利用信息资产，驱动业务增长和创新。



▲ 全球实时开放数据智能服务平台架构

▲ 全球实时开放数据采集总控平台

· 全球智多星系统　项目负责人：翟崑

北京大学大数据分析与应用技术国家工程实验室、国际关系学院、区域国别研究院、北京大学长沙计算与数字经济研究院共同开发智多星系统。全球每个国家、地区都有自身独特的运转系统。该项目依托国际政治、国际关系、区域国别学科基础，开展对不同国家、地区的复杂系统研究。通过跟踪全球主要智库和国际组织的动态，构建动态系统研究中全维度监测区域与国别的方法，将特定国家、区域置于全球动态演进的框架下进行观察和分析，并推动形成相关理论，建立动态监测中国与世界关系的评估框架与指标体系。相关理论成果已在北京大学大数据分析与应用技术国家工程实验室及北京大学长沙计算与数字经济研究院的支持下完成了产品研发及在线部署，在该学科的理论研究和人才培养等方面发挥了较好的促进作用。

▲ 全球智多星系统

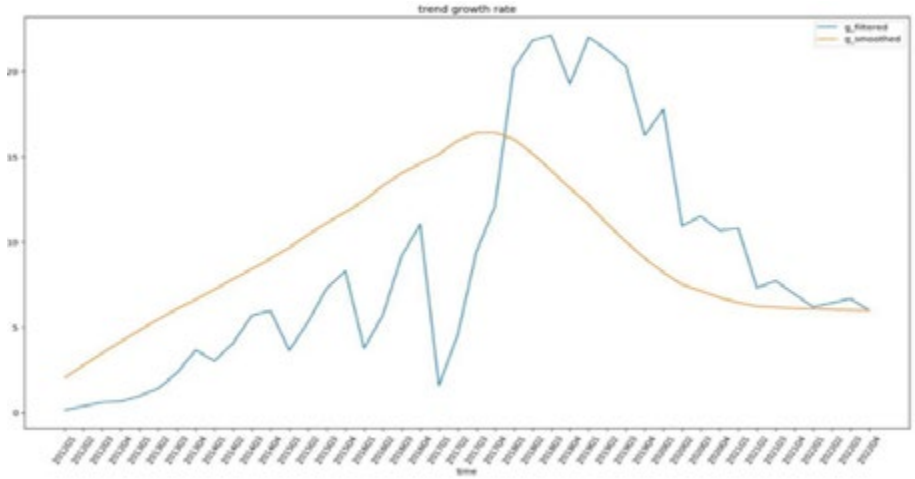
· 东南亚数据集搜索系统　项目负责人：王继民

东南亚对中国来说有很强的地理战略意义，随着数据驱动研究范式的发展，与东南亚进行友好合作、政策互利、人文研究等越来越离不开相关数据研究的支持。近些年来，东南亚国家在开放政府数据门户、开放研究数据方面成果颇多。本项目对区域与国别研究相关数据平台、东南亚国家政府数据平台、研究数据平台开展，为相关研究提供了数据导航，促进数据的应用和数据价值的再发挥。

国家	平台	语种	数据量
印度尼西亚	印度尼西亚政府开放数据平台	印尼语	45476
马来西亚	马来西亚政府开放数据平台	马来语	12330
越南	越南政府开放数据平台	越南语	9533
泰国	泰国政府开放数据平台	泰语	5718
柬埔寨	东南亚开放发展数据平台	英语	3877
老挝	东南亚开放发展数据平台	英语	2715
缅甸	东南亚开放发展数据平台	英语	2490
新加坡	新加坡政府开放数据平台	英语	1891
菲律宾	菲律宾统计协会-开放统计数据	英语	1791
泰国	东南亚开放发展数据平台	英语	1410
文莱	文莱政府开放数据平台	英语	604
老挝	国家统计局-统计数据库	老挝语	162
东帝汶	人道主义数据中心	英语	74
东南亚11国家	Figshare	英语	7013
东南亚11国家	Harvard Dataverse	英语	811
东南亚10国家（除柬埔寨）	Repositori Ilmiah Nasional (RIN)	印尼语、英语	673
东南亚10国家（除文莱）	Research Data Australia	英语	622
东南亚11国家	UK Data Service	英语	121
东南亚10国家（除东帝汶）	ICPSR、UNC Dataverse等27个平台	英语	525

· 印度经济潜在增速估计与演变路径　项目负责人：王熙

潜在经济增长率作为宏观经济学中的一个基础性概念，不但对于一国中长期经济发展规划制定和公共政策制定都具有重要意义，清晰估算重要他国 GDP 潜在经济增速也有益于我国客观科学的指定经济竞争策略。本项目系统性收集我国以及印度各类宏观经济数据，开发了动态隐因子算法估计了我国与印度的长期经济增速，开创性将动态因子模型与机器学习算法结合发展出了一套基于宏观大数据进行即时预测框架。



• 国际组织数据库实验室

实验室负责人

张平文

武汉大学校长、党委副书记
北京大学博雅讲席教授
大数据分析与应用技术国家工程实验室主任

张海滨

北京大学国际关系学院副院长
国际组织与国际公共政策系主任、教授、博士生导师
北京大学碳中和研究院副院长
北京大学全球健康发展研究院副院长
北京大学国际组织研究中心主任

科研成果

· 国际领导人画像数据库　项目负责人：张海滨　技术负责人：黄晶

数据库通过对联合国系统、政府间新型组织及重点半官方组织和民间组织的基本信息、评估类信息的整理形成数据基础，发布有国际影响力的研究成果及智库报告。

截至目前：国际组织领导人画像库新增到 2 万领导人，其中有 2000 个领导人可以找到个人信息，并针对这些领导人完成基本信息、教育信息、工作信息的提取工作；国际组织数据库展示网站二期于 2024 年 3 月上线，二期新增国际组织领导人数据库介绍信息和数据展示分析区，用来展示国际组织领导人的职业、年龄分布和社交网络。

网站地址：iodb.pku.edu.cn

• 溟坤海域态势感知系统实验室（依托北京大学重庆大数据研究院建设）

实验室负责人

胡波

北京大学国际关系学院研究员
北京大学海洋战略研究中心主任
“南海战略态势感知计划”主任
北京大学重庆大数据研究院海洋大数据智能研究中心主任

科研成果

· **“溟坤海域态势感知系统”(MK-MDA) 项目负责人：胡波**

“溟坤海域态势感知系统”(MK-MDA) 是依托北京大学大数据分析与应用技术国家工程实验室、北京大学重庆大数据研究院的重点研发项目，基于地理空间信息系统，融合 AIS、ADS-B、商业遥感、电磁频谱和新闻舆情等多源异构数据，为用户从全局视角整体洞悉海上军事、政治、经济、环境等态势，为国家安全与国际安全等领域的研究提供数据赋能和实证支撑。产品功能填补国内海域态势感知数融合领域的空白，态势数据和热点舆情数据的融合技术达到全球领先水平。目前已实现产业转化，孵化公司溟坤数科（重庆）智能科技有限公司。

▲ 溟坤海域态势感知系统

微信公众号二维码

地址：北京大学静园六院
电话：010-62767117
网址：https://bda.pku.edu.cn
公共邮箱：gdsjsys@pku.edu.cn

大数据分析与应用技术
国家工程实验室

NATIONAL ENGINEERING LABORATORY FOR
BIG DATA ANALYSIS AND APPLICATIONS

2025 年 7 月