

教育部工程研究中心年度报告

(2025年1月——2025年12月)

工程中心名称：数据库与商务智能

所属技术领域：信息与电子工程

工程中心主任：陈红

工程中心联系人/联系电话：张晓莹/15210984900

依托单位名称：中国人民大学

2026 年 5 月 11 日填报

一、技术攻关与创新情况

工程研究中心总体目标和定位是依托中国人民大学，围绕国家大数据战略需求，在数据库与商务智能领域探索大数据管理与分析的关键新兴科技，研发具有重大影响力的新一代数据库技术与系统、大数据分析技术与平台、大数据方法与智慧治理，大力发展我国信息系统升级产业，服务国民经济主战场。同时承担国家和地方重大科研攻关任务，推动数据库与商务智能相关领域的科研创新，为国家培养数据库、大数据等领域的高端创新型领军人才。2025年新增国家级科研项目14项，省部级项目7项，新获批纵向科研经费900余万元。在数据库、商务智能、大数据等领域取得一些创新成果，以第一作者或通讯作者身份在国内外重要期刊或会议上发表高水平学术论文130余篇，出版著作教材5部，新增专利授权23项。

中心研发围绕以下三方面进行：

（1）新一代数据库技术与系统。探索一系列面向复杂需求的新形态数据库技术，包括跨域分布式数据库、异构硬件数据库、云原生数据库、智能数据库、全同态压缩技术等；与国产数据库企业联合研制一系列有影响力的系统和工具。

（2）大数据分析技术与平台。研究云原生大数据分析技术与平台、数据要素可流通与分析技术、智能数据分析、元数据管理、隐私计算、区块链等。

（3）大数据方法与智慧治理。具体包括智慧金融科技、智慧医疗、智慧城市、智慧养老等交叉方向。

代表性成果：

1. 融合AI技术的新一代融合数据库产品

面对AI大模型时代带来的复杂应用场景，以“打造融合AI技术的新一代融合数据库产品”为长期目标，围绕“四个一体化”推进技术攻关：（1）金仓融合数据库（KESV9 2025版）：集成“四个一体化能力”，在实现对异构数据库高度兼容的基础上，内置向量等多模型计算引擎。在性能上，通过并行处理、虚拟事务堆栈、多版本并发控制优化等技术，在关键业务场景中达到2万TPS的吞吐量，在电网营销平台中超越Oracle性能，达到4万TPS。（2）金仓企业级统一管控平台（KEMCC）：通过可视化、智能化手段，解决了国产数据库大规模部署运维的难题。KEMCC能够纳管金仓全栈产品，支持从安装部署到监控诊断、批量升级、参数修改的全生命周期管理。如在头部汽车制造企业，成功支撑了超过1000个数据库实例的统一管控，大幅提升了运维效率。（3）金仓数据库一体机（云数据库AI版）：作为金仓“交付即智能”的代表性产品，深度融合金仓数据库、超融合平台、高性能硬件及AI大模型，深度嵌入AI能力，以“性能、智能、安全”为核心，搭载“的卢运维智能体”“赤兔加速引擎”，专设“绝影安全模块”，提供开箱即用、成本可控的数据库私有云解决方案。（4）企业级智能海量数据集成平台（KFS Ultra）：基于全新研发的“掣电融合数据复制引擎”，采用Serverless架构，实现了资源弹性伸缩和高可用性。可支持超100种数据源，覆盖关系型、非结构化、向量、时序等多种类型，独创“零停机数据校验与修复”功能，可在业务不停机的情况下，精准校验海量数据，解决数据的一致性难题。

2. OLAP数据库一体机：面向国产新硬件的高性能数据库一体机

系统

由中国人民大学联合中国移动信息技术有限公司研发的OLAP数据库一体机系统，解决当前国产数据库一体机在高性能数据库分析方面的空白。系统存储层基于中国人民大学数据库团队自主研发的Pixels列式存储引擎，计算层基于中国移动研发的梧桐数据库查询引擎。在计算层，Pixels针对国产CPU和国产NVMe SSD、ZNS SSD进行软硬件协同优化，重点解决目前大数据分析系统中存在的内存多次拷贝问题、缺少异步I/O优化问题、和数据压缩后解压速度慢问题，实现了从NVMe SSD到CPU内存一次性数据传输后零数据拷贝、基于io-uring的高效异步I/O框架、以及可以快速解压的轻量级数据编码算法。以上优化使得数据库一体机存储层在NVMe全闪存的存储层上数据访问性能相比目前主流开源列存储引擎系统1-2倍，并且随NVMe SSD数量增加呈现更好的扩展性。目前基于Pixels的数据库一体机存储引擎以及和梧桐数据库查询引擎集成，在中国移动的青海等多个省公司落地应用。

3. 自主数据科学智能体 DeepAnalyze

DeepAnalyze围绕“如何使人工智能系统自主完成复杂数据科学任务”智能发展领域关键问题，面向传统数据科学工具依赖固定流程、难以处理开放环境下复杂任务的瓶颈，提出了课程式代理型训练方法，模拟人类学习路径，按照“从单一能力到复合能力”的渐进方式，在真实环境中持续提升大模型的数据理解、推理规划、工具调用与任务执行能力；提出了面向数据科学任务的轨迹合成框架，自动化构建超过50万条数据科学推理与环境交互数据，为大模型在复杂搜索空间中学习正确决策路径提供高质量训练支撑。

二、成果转化与行业贡献

（一）总体情况

2025年，依托工程研究中心的前沿科技成果，电科金仓加速技术向产业应用的转移转化，实现多项核心产品的迭代。金仓融合型数据库（KESV9 2025版）、企业级智能海量数据集成平台（KFS Ultra）、企业级统一管控平台（KMCC）、数据库一体机（云数据库AI版）等产品，通过独立实施或联合渠道伙伴、生态伙伴，完成工程技术成果产品化、市场化的转化。2025年，公司签约额保持超30%的增长，结构持续优化：其中，党政行业收入占比显著下降，医疗、交通、金融、能源、运营商等关键行业领域收入占比全面提升，取得关键应用领域销售套数第一，医疗、交通行业双第一的成绩，累计装机部署超过100万套。

2025年，金仓的技术成果转化主要涵盖“四个一体化”：多语法体系一体化兼容、多应用场景一体化承载、多模数据一体化处理、AI融合的开发运维一体化管理。此外，还实现了基于虚拟事务堆栈的高性能自治事务技术、基于共用表达式编译执行的高性能PL/SQL技术、基于固定缓冲池的索引热点页面冲突性能优化、基于活动快照间隔的精确旧数据清理技术、WAL并行回放技术等多项关键核心技术的转化。目前，金仓已形成技术驱动加市场驱动的双驱动正向循环机制，并通过专利、软著、生态共建、社区建设等方式强化知识产权与人才保障。

（二）工程化案例

（1）中汇亿达-Comestar资金交易管理系统

项目基本情况：ComStar是中国外汇交易中心旗下子公司，致力于为金融机构FICC业务（固定收益、外汇与大宗商品）提供跨市场、全流程、一站式解决方案，业务覆盖交易下单、智能交易、资金业务管理、风险管理、对客业务、资管等领域，服务全国300余家金融机构的日常交易管理。该项目系统涉及多个复杂功能模块，对数据库的高可靠、高安全、高性能提出了极高的要求。

针对ComStar资金交易管理系统（覆盖近300家金融机构，含二十余类专业解决方案）的复杂业务场景，项目采用多数据协议一体化兼容框架实现对Oracle的100%兼容，通过单进程多事务技术将自治性能提升5-6倍，结合过程语言共享缓存、SIMD高精度数值计算、物理代价优化及KFS异构数据同步双轨并行迁移等关键技术，遵循“深度调研—反向适配—全链路测试—双轨并行—平滑上线”路径，采用“KFS+KDOTS”技术组合拳，以KFS准实时同步机制实现业务“零感知”迁移，迁移时间较业内平均缩短50%以上，并部署主备KES架构满足高可用金融级容灾标准。系统上线后多月来持续稳定运行，日终批次业务耗时节省约20%，损益报表分析等核心场景查询响应速度提升至原系统近三倍，验证了国产数据库支撑金融核心交易系统的硬核实力，为金融行业核心系统信创化提供了可复制、可推广的实践标杆。项目入选2025“金信通”金融科技创新应用典型案例。

（2）一汽集团：全业务、全流程信创替代

一汽集团为国有特大型汽车集团，该信创替代项目覆盖生产运营、车联网、供应链等20余类核心业务，横跨IT基础、大制造、大营销、大数据等多个领域，对数据库的高并发、高可用及复杂业务

支撑能力提出了极高要求。

针对该项目特点，金仓采用多数据协议一体化兼容框架，实现对Oracle、MySQL的100%语法兼容与性能级平替，依托KFS异构数据同步软件实现不停机平滑迁移，结合共享存储集群与读写分离架构保障高并发场景下的系统稳定性和可用性。同时，针对汽车行业特定场景（如TSP-TBOX智能联网系统），联合定制相关功能组件，填补了该领域的空白。项目遵循“深度调研—方案制定—阶梯迁移—双轨并行—平滑上线”的实施路径，按照“办公管理系统→生产运营系统→核心业务系统”的顺序分阶段推进，累计部署2400+套数据库，适配上线应用系统300+个，覆盖集团20余类业务。此外，项目产出《关系型数据库国产化替换迁移技术标准》（T/BJCSA07-2025），形成了“标准制定—工具迁移—定制开发—人才赋能”的全流程可复制方法论。该项目验证了国产数据库在高并发、复杂业务场景下的成熟度，为汽车行业及重点领域国产化替代与数智化改造提供了可推广的实践标杆。

（三）行业服务情况

2025年，公司通过多元化行业服务深度赋能合作伙伴与用户，持续构建开放共赢的产业生态。在生态共建方面，启动“常青藤计划”，授权ISV在长达5年内免费使用金仓数据库进行开发与测试，大幅降低合作成本与门槛，推动产品在能源、金融、运营商、交通、医疗等关键领域落地，同时积极适配第三方生态工具，并持续推进与华为云、阿里云等主流云平台的融合，通过一体机、云数据库等形式丰富交付方式。

在技术咨询与培训方面，人才发展中心全年累计培训近18,000人，举办各类培训329场，同比增长18%，内容覆盖产品使用、性能优化、数据迁移、运维管理等，并通过出版专业书籍，如2025年出版《金仓数据库KingbaseES性能优化》，向行业分享最佳实践。

在行业协会与联盟活动方面，金仓连续当选信创工委WG2数据库组组长单位，积极参与并主导行业标准制定；2025年，金仓与信通院联合发布首个《融合型数据库技术研究报告》，引领行业发展。

在品牌与国际拓展方面，金仓积极参加国内各类重要行业展会，并首次走出国门参加新加坡行业展会及香港“一带一路”高峰论坛，持续巩固扩大品牌影响力。

2025年，金仓通过生态共建、知识赋能、行业引领与品牌建设，持续推动国产数据库的产业化应用与产业链认知提升，为关键信息基础设施自主可控作出了重要贡献。

三、学科发展与人才培养

（一）支撑学科发展情况

数据库、大数据技术是计算机科学与技术学科的重要分支，工程研究中心聚焦数据库、大数据系统的研发，在科技创新、成果转化、人才培养、队伍建设等方面为计算机科学与技术一级学科发展建设提供了重要支撑。主要表现在以下方面：

第一，促进了依托学科的科研创新和成果转化，提升了依托学科承接国家级、省部级项目的能力。2025年，中心承担了国家重点

研发计划课题、国家自然科学基金重点项目等国家级科研项目。

第二，为创新型人才培养提供了重要实践基地，优化了创新型工程学科人才培养条件和环境。工程中心向本科生、研究生提供课程所需的实践设备、实验空间、工程指导等，并选拔优秀学生进入工程中心进行工程项目锻炼。2025年，牵头建设数据与人工智能研发实验中心、中国人民大学社会发展创新工程中心（智慧治理），与头部企业合作，成立中国人民大学-小米大模型联合重点实验室、中国人民大学-中电科金仓数据库系统创新联合实验室。

第三，工程中心全面支持计算机学科拔尖人才培养，支持强基班、图灵实验班的数据库、数据分析、数据智能相关课程开设，包括《数据科学导论》、《数据库系统概论荣誉课程》、《机器学习》等。中心同时为强基班、图灵班人才培养过程中的课程实践、专业实习和毕业设计提供环境与题目，在实际工程实践和科研项目中锻炼拔尖创新人才的创新能力与动手能力。

第四，工程中心教师继续牵头建设教育部“101计划”数据库系统虚拟教研室和教育部—“智能基座”数据库课程虚拟教研室，召开2025年虚拟教研室年会、举办“101计划”数据库系统课程导教班、担任CCF计算机课程改革导教班（CCD2025）“数据库系统概论”课程助教教师。数据库课程虚拟教研室建设成果显著，“‘线下激活线上、实体赋能虚拟’：通过年会切实推进协同育人和产教融汇”入选2025年教育部典型教研方法。

（二）人才培养情况

2025年，培养博士硕士毕业生130人，在读研究生441人。涌现

出一批优秀学生，例如，博士生周昆入选2025年“CCF博士学位论文激励计划”；博士生官佳薇、杜盼、许骞入选2025年中国科协青年科技人才培育工程博士生专项计划；博士生官佳薇获批国家自然科学基金青年学生基础研究项目（博士研究生）；博士生葛加可入选CCF信息系统专委会优秀博士学位论文激励计划；硕士生刘家真入选2024年度中国电子学会硕士学位论文激励计划。

（三）研究队伍建设情况

工程中心一直重视人才队伍建设，特别关注青年骨干教师的成长，培养了一批优秀青年教师。2025年，工程中心引进科研人员4名，分别是王文轩讲师、傅哲讲师、张帆进讲师、张绍磊讲师。教师具体情况如下：

王文轩，讲师、青年英才、玉兰学者，入选国家级青年人才项目。香港中文大学博士，曾在香港科技大学担任博士后研究员。研究方向包括：大模型应用与大模型安全。在ACL、ICML、ICLR、CVPR等国内外学术期刊和会议上发表论文30余篇，获ACM SIGSOFT杰出论文奖、EMNLP 2025最佳论文提名奖。

傅哲，讲师，获得北卡罗来纳大学夏洛特分校计算与信息系统博士学位。研究方向为商务智能与智慧决策系统，在SIGIR、EMNLP、POM等国际顶级计算科学与管理科学会议与期刊上发表论文十余篇，其中一作/通讯论文6篇。获PACIS2024最佳论文奖、Essam El-Kwae师生杰出研究奖等奖励。

张帆进，讲师，青年英才。清华大学博士&博士后，师从唐杰教授。研究方向包括知识图谱、数据挖掘、大语言模型。在KDD、

TKDE、WWW等CCF-A类会议期刊发表论文十余篇，谷歌被引3000余次。主导发布多个版本亿级开放学术图谱OAG。曾获国家科学技术进步二等奖和中国人工智能学会年度优秀科技成果奖。

张绍磊，讲师。他于中国科学院计算技术研究所获得博士学位。研究方向涵盖大语言模型、智能体系统、数据科学。相关研究成果在NeurIPS、ACL、ICLR等国际人工智能与自然语言处理会议发表论文30余篇，谷歌学术引用1300余次。开源的多语言与多模态大模型项目在GitHub社区累计获得10000+星标，模型下载量万余次。相关自研算法在谷歌、华为、百度联合举办的国际实时翻译比赛AutoSimTrans2021 流式输入赛道获得第一名，主导开发的 百聆大模型被国际开放基准委员会评为“Top 100 Open-source Achievements”。

工程中心在科研项目、国际交流、仪器设备等方面为科研人员提供支持，帮助科研人员快速成长。例如，新教师根据自身研究方向选择加入研究团队，由经验丰富的老教师指导其开展科研工作；为教师赴国外科研院所短期学术交流、参加国际高水平会议提供经费支持；充分利用校级高性能计算集群和“人大行云”大数据平台，为科学研究提供了较好的工作环境。

四、开放与运行管理

（一）主管部门、依托单位支持情况

中国人民大学作为工程中心的依托单位，在以下几方面支持中心建设和发展：

在学科建设方面，中国人民大学是我国最早开展数据库技术教

学研究和数据库管理系统开发的单位之一，起步于1978年，由老一辈计算机科学家萨师煊教授等人所创立。中国人民大学始终站在学科前沿，跟踪国际先进技术，研究了关系数据库技术、并行数据库技术、数据仓库与数据挖掘技术、新硬件环境下的数据库技术等，先后建立了“数据工程与知识工程”教育部重点实验室、“大数据管理与分析方法研究”北京重点实验室、“新一代智能搜索与推荐”教育部工程研究中心、“大模型与智慧治理”北京市重点实验室、这些都为工程中心的研发创造了有利条件。

在人才引进和团队建设方面，大幅度提高引进待遇，吸引海内外高层次人才。不定期举办德科大讲堂、信息讲坛、信息系统双周论坛，邀请海内外知名学者交流研讨，加强国内外合作。支持中心固定人员出国访学，并积极争取国家、省部以及学校的各类留学基金和海外留学名额。鼓励中心人员参加国际交流和国际合作，提供高水平国际会议参会经费支持，资助高水平论文发表并提供后续配套研究经费，不断提升团队整体水平和国际影响力。

在研究生培养方面，在中国人民大学任职的中心固定人员均在招收和培养计算机科学与技术学科的博士和硕士研究生。设置荣誉辅修项目，依托国家重大重点科技任务，由中心固定人员担任学术导师，选拔优秀本科生加入科研团队。

（二）仪器设备开放共享情况

工程研究中心配备完善、先进的仪器设备，为中心进行各项科学研究提供了保障。云计算平台、智算平台、新硬件大数据管理与分析实验平台等为学院师生提供云计算、大数据、深度学习、多媒

体数据分析等计算及存储服务，支撑大数据、云计算、机器学习、人工智能等方面的众多研究课题，为各类重点重大项目提供重要计算服务资源保障。

（三）学风建设情况

工程中心一直将学风建设作为人才培养的核心环节常抓不懈，在制度建设、学生管理、师德师风建设、课程督导等方面持续深化学风引导，营造公平公正、健康向上的师生学习和科研环境。主要措施围绕以下几点：（1）持续强化师德师风建设，充分发挥师德师风建设与监督委员会作用，扎实开展教师师德考核，切实将师德师风建设摆在与教学科研同等重要的位置，在岗位聘任、评奖评优、绩效考核等工作中严格落实师德师风“一票否决制”。坚持正面榜样引领与反面案例警示教育并重，定期召开师德师风专题会议，做好警示教育，选树宣传师德榜样典型。组织撰写萨师煊、陈余年等“大先生”先进事迹文稿，2025年，推出教师专访13篇。密切关注教师思想动态，定期召开座谈会，及时掌握教师思想状况与生活情况，把苗头性、倾向性问题化解在萌芽状态。（2）强化课程督导，夯实课堂教学主阵地。本年度，学院持续完善课程督导机制，学院领导班子成员、资深教授深入课堂教学一线开展常态化督导。督导组采取随机听课、重点巡查相结合的方式，重点检查教师教学规范、学生出勤率等情况，课后及时与任课教师、学生代表交流反馈，形成“督导—反馈—改进—复查”的闭环管理，有效推动了优良学风的形成与巩固。（3）从严从实做好学业警示工作，建立学业困难学生台账，定期展开学生学业情况谈心谈话，向任课教师了解

学生课业进展，教务、学工、班主任多方发力，建立一对一帮扶小组，全力支持学生克服学业困难；（4）全面落实警示教育工作，从新生入学教育开始，在每学期开学、期中、期末等重要节点，组织专题会议或主题班会向全体学生强调考试诚信与学术诚信问题，加强法律法规、校规校纪学习，严禁考试作弊、学术不端等行为，全流程、常态化对学生开展警示宣传工作，为营造良好学风守住底线；（5）持续开展“信手相连”学业辅导活动，由高年级本科生与研究生面向低年级本科生对难度较大的数学课、专业课开展辅导与答疑，并面向全校学生开展Python、大语言模型、OpenClaw等热点课程与讲座，发挥信息学院专业优势，突出“传帮带”作用，传承信院学子“敢为人先，科技报国”的精神品质。

（四）技术委员会工作情况

2025年12月16日，召开2025年度技术委员会会议。工程中心主任陈红教授向技术委员会做工作汇报，全面介绍了中心在队伍建设、前沿研究成果、深化校企合作、学科建设与人才培养等方面取得的进展。技术委员会对工程研究中心的下一步建设与发展展开了热烈、深入的讨论，对工程研究中心的建设提出了很多宝贵的建议。

五、下一年度工作计划

在技术研发方面，立足数据库与商务智能领域前沿，继续围绕新一代数据库技术与系统、大数据分析技术与平台、大数据方法与政府决策这一研发目标，形成“基础研究-技术攻关-应用验证”全链条创新体系。依托国家级、省部级各类重大重点科研项目，通过

持续不断的技术创新，掌握“卡脖子”关键技术，打破技术壁垒，服务国家战略需求，形成具有国际影响力的亮点成果

在成果转化方面，充分发挥现有联合实验室的平台优势，继续与国产数据库企业建立多种形式的产学研协作，构建“需求牵引-技术供给-成果转化”闭环生态。整合优化科技资源配置，立足国家战略需求，共同承担国家重大科技任务，解决数据库、商务智能、大数据领域国家“卡脖子”技术难题。继续拓展政产学研用融合通道，理顺科研成果转化的机制，在关键核心技术突破、科技成果转化、高水平工程技术人才培养等方面发挥重要示范作用。

在人才培养方面，深化培养体制和培养模式的改革，以国家重大项目和重大应用工程为载体培养工程型创新人才；鼓励高年级本科生和研究生积极参与科研项目，用真问题驱动学生工程能力的培养。

在团队建设方面，坚持“引育并举”，打造一支结构合理、创新能力突出的高水平人才队伍。继续加强宣传，多种渠道吸引国内外高层次人才，提高引进待遇。鼓励国际合作，与国内外科研机构建立长久深入的合作。此外，不断完善科研人员考核制度与奖励机制，加强队伍梯度建设。

在制度优化方面，积极争取依托单位中国人民大学在经费、人才引进等方面的政策支持；不断完善中心内部管理制度，规范人、财、物的管理，提高仪器设备的高效共享。

六、问题与建议

无

七、审核意见

工程中心负责人审核意见：

本年报的数据及内容真实有效，信息填报属实

工程研究中心主任：

年 月 日

依托单位审核意见：

依托单位：

（单位公章）

年 月 日

八、年度运行情况统计表

研究方向	研究方向1	大数据分析技术与平台		学术带头人		陈红
	研究方向2	新一代数据库技术与系统		学术带头人		杜小勇
	研究方向3	大数据方法与智慧治理		学术带头人		李翠平
	研究方向4			学术带头人		
工程中心面积	4359.0 m ²			当年新增面积		0.0 m ²
固定人员	51 人			流动人员		0 人
获奖情况	国家级科技奖励	一等奖	0项	二等奖	0项	
	省、部级科技奖励	一等奖	0项	二等奖	0项	
当年项目到账总经费	2267.94万元	纵向经费	989.32万元	横向经费	1278.62万元	
当年知识产权与成果转化	专利等知识产权持有情况	有效专利	168项	其他知识产权	104项	
	参与标准与规范制定情况	国际/国家标准	0项	行业/地方标准	0项	
	以转让方式转化科技成果	合同项数	0项	其中专利转让	0项	
		合同金额	0.0万元	其中专利转让	0万元	
		当年到账金额	0.0万元	其中专利转让	0.0万元	
	以许可方式转化科技成果	合同项数	0项	其中专利许可	0项	
		合同金额	0.0万元	其中专利许可	0.0万元	
		当年到账金额	0.0万元	其中专利许可	0.0万元	

		以作价投资方式 转化科技成果		合同项数	0项	其中专利作价	0项
				作价金额	0.0万元	其中专利作价	0.0万元
		产学研合作情况		技术开发、咨询 、服务项目合同 数	24项	技术开发、咨询 、服务项目合同 金额	1289.8万 元
当年服务情况		技术咨询		329次		培训服务	18000人次
学科发 展与人才 培养	依托学科 (据实增删)	学科1	数据库	学科2		学科3	
	研究生 培养	在读博士	209人		在读硕士		278人
		当年毕业博士	25人		当年毕业硕士		104人
	学科建设 (当年情况)	承担本 科课程	7168学时	承担研究生 课程	1840学时	大专院校 教材	6部
研究队 伍建设	科技人才	教授	25人	副教授	19人	讲师	7人
	访问学者	国内		0人	国外	0人	
	博士后	本年度进站博士后		0人	本年度出站博士后		0人