

“应用驱动创新”培养“硬核” 数据科技创新人才的探索与实践

金澈清** 钱卫宁 周烜 周傲英

华东师范大学数据科学与工程学院, 上海 200062

摘要 当前高校在培养数据科技创新人才时, 正面临三重严峻挑战: 学科知识体系滞后于技术迭代, 课程设置常落后于行业前沿; 学生创新思维深度不足, 难以将知识转化为解决复杂问题的能力; 教学与真实产业场景脱节, 导致所学难为所用。针对这一困境, 华东师范大学数据科学与工程学院以“硬核”为内核, 在十多年探索中构建出一套卓有成效的人才培养范式。该模式包括适应数据科技创新人才培养的知识体系、基于“应用驱动创新”理念的校企联动人才培养方法、和以实验室为载体的产教研融合人才培养方法。经过十多年实践探索, 该模式展现出显著的培养效果: 重新梳理了相关专业的教学体系并出版了系列教材, 建设了一支高水平、产研融合的师资队伍, 培养了一批受到就业市场欢迎的优秀研究生, 以水杉在线为载体进行了大规模教育科技工程实践。该模式可以为相关学科人才培养提供借鉴。

关键字 应用驱动创新, 产教融合, 数据科技, 研究生培养

Explorations and Practices in Cultivating High-Caliber Talent in Data Science and Technology Innovation through “Needs drive Innovation” Philosophy

Cheqing Jin Weining Qian Xuan Zhou Aoying Zhou

School of Data Science and Engineering
East China Normal University, Shanghai, 200062, China

Abstract—Currently, universities face three formidable challenges in cultivating data science and technology innovation talent: (i) disciplinary knowledge systems lag behind technological iteration, with curricula often trailing industry frontiers; (ii) students lack sufficient depth in innovative thinking, struggling to translate knowledge into solving complex real-world problems; (iii) teaching remains disconnected from genuine industry scenarios, resulting in the knowledge acquired in classrooms often proving inadequate in practical settings. To address these challenges, the School of Data Science and Engineering at East China Normal University has developed an effective talent cultivation paradigm centered on advanced core technology training through over a decade of exploration. This model encompasses: a knowledge system adapted for nurturing data science innovation talent, an industry-academia collaborative training approach based on the principle of “needs drive innovation”, and a laboratory-based methodology integrating production, teaching, and research. After over a decade of practical exploration, this model has demonstrated remarkable training outcomes: (i) restructuring pedagogical systems for relevant disciplines and publishing a series of textbooks; (ii) building a high-caliber, industry-integrated faculty team; (iii) cultivating cohorts of outstanding postgraduates highly sought after in the job market; (iv) implementing large-scale EdTech engineering practices via the Shuishan Online platform. This paradigm offers valuable insights for talent cultivation in related academic fields.

Keywords—Needs drive innovation, industry-education integration, data technology, graduate education

1 引言

数据, 是现实世界在网络信息空间的投影, 隐含着丰富的信息和规律。它和土地、劳动、资本和科技一样, 是一种生产要素, 是建设数字中国的基础, 是驱动数字化转型和数字经济的一种新能源动力(Power)。基于数据, 已经并且还将继续出现一大批新兴应用。而近年来人工智能技术和应用的爆发式发展, 如同电机(Motor)之于电(Electric Power),

极大地发挥出数据能(Data Power)的潜力。在新兴信息技术应用大发展、国际科技竞争加剧的背后, 亟需发展硬核科技, 解决卡脖子问题。硬核科技是应用、系统和理论的组合体。有成功的应用, 有成体系的技术, 再加上在前两者基础上抽象出来的概念和理论, 三者缺一不可, 方可称为“硬核”。培养“硬核”数据科技创新人才面临如下三个挑战。

如何重构体现数据学科持续发展的知识体系。数据科技发展迅速, 体现在知识体系具有学科特征。但是现有计算机类培养体系没有与时俱进, 凸显数据的

* 通讯作者: 金澈清 csteic3@163.com。

重要性,其知识体系无法适应数据时代的发展趋势。

如何培养“造车人”,而非“驾驶员”。驾驶员只需要了解如何使用工具、操作软件;而驾驶员则需要了解系统的底层思维,解决卡脖子问题,并且将成果开源,转化成实际生产力,服务社会需求和国家战略。

如何以应用倒驱,激发学生掌握硬核科技愿景。现有人才培养方式并未与应用场景深度融合,未能针对“真实问题”开展科学研究,其成果局限性显著,难以实现复合式、高层次数据专业人才的培育和输送。

当前,一些学者已经在探索如何培养数据科技创新人才。周傲英等人指出,数据是驱动新时代发展的新能源,数据科学与工程大数据时代的新兴交叉学科[1-2]。王红等人提出,需要依靠产教融合理念来培养应用型人才[3]。金澈清等人提出,要推进产教融合以打通数据学科研究生培养的“最后一公里”[4]。毛嘉莉等人设计了校企融合的非全日制专业硕士定制化

本文以培育“硬核”数据科技创新人才为核心目标,以“应用驱动创新”理念为导向,探索如何深度对接企业真实需求开展科研创新,设计相应的人才培养模式,推动科研成果转化成生产力,解决产教融合“两张皮”问题,培育出既懂理论又懂实践、既能创新又能落地的、面向未来的复合型数据人才,真正满足国家和社会对高质量数据人才的需求。

2 “应用驱动创新”数据人才培养模式

为应对前面提到的三大挑战难题,提出了五个解决方案,包括培养前置与本硕博贯通相结合的培养方式、重构的知识体系梳理与课程体系、实验室为载体的产教研融合人才培养、产研融合的师资团队建设、支撑工程实践的教育科技等,如图1所示。

(1) 培养前置与本硕博贯通相结合的培养方式

进入研究生段学业之前,引导学生通过丰富的在线学习平台进行预习,入学后的课程先修要求对接前

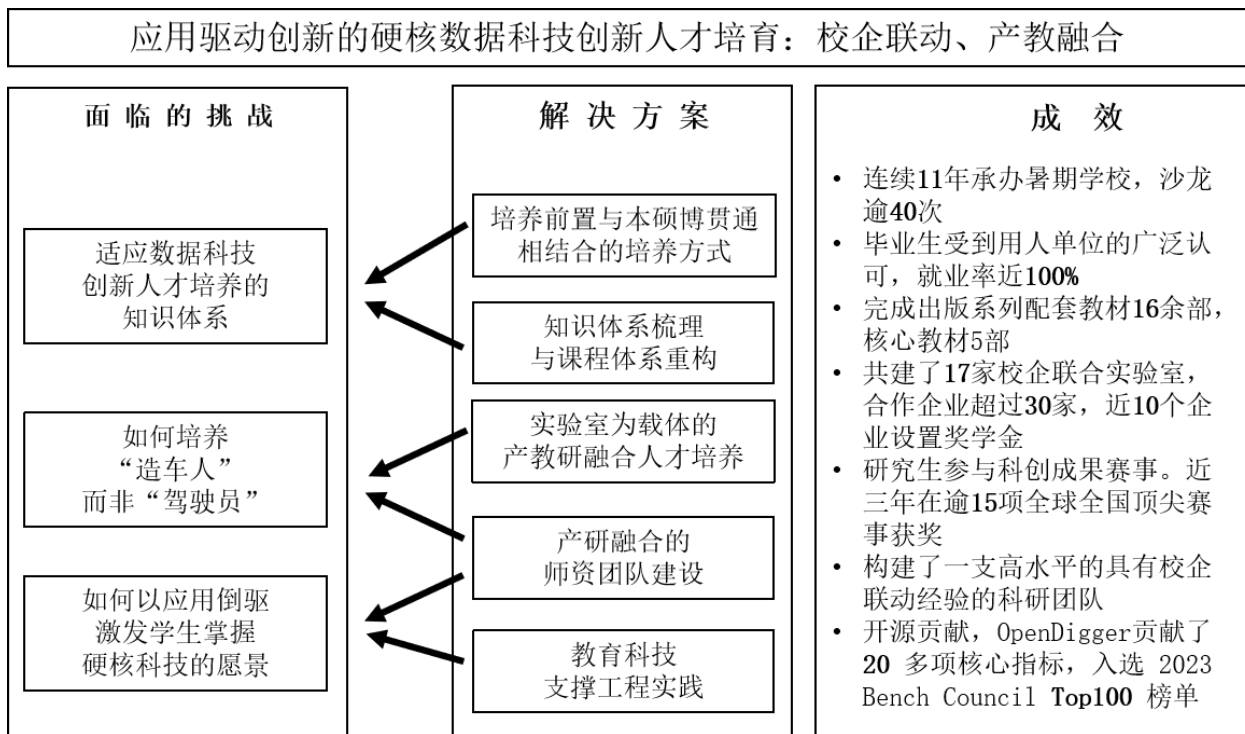


图1 硬核数据科技创新人才培育模式

人才培养模式[5]。何倩等人将产教融合思想引入到计算机网络课程建设中来,并进行了实践[6]。金劲彪等人研究如何在产教融合背景下对产业教师的队伍建设进行制度保障[7]。郭忠文等探讨了“产、教、学、研”四位一体研究生产教融合培养模式研究[8]。修佳鹏等针对特色化示范性软件学院的产教融合实践创新做了有益的探索[9]。总体上来说,系统性地探索与总结“硬核”数据创新人才培养模式的工作还不多见。

置预习;对于本校直升的学生,推行落地“本硕贯通”,引导本科高年级学生修读研究生课程。本硕博全学段研究生须参加暑期的“前沿知识讲堂”+“实践课堂”全流程,暑期学校作为课程学习的协同形式,是为“硬核”素养的学业水平考查环节。

在非全日制研究生的招生与录取过程中,引入了“双录”管理模式。该类研究生通常属于定向培养生源,同时具备企业员工和高校学生的双重身份。由于其课

程学习多集中于晚间和周末等时段，学习时长相对有限，较全日制学生存在明显差距。因此，如何有效保障并提升非全日制研究生的培养质量，成为一项重要课题。“双录”模式作为应对该挑战的有益探索，涵盖以下三个层面：其一，该模式推动了校企合作关系的进一步深化——学校与企业依托既往良好合作基础，共同制定培养方案、合作开发课程，致力于定向领域专业人才的系统培育；其二，在招生录取阶段，学生须先后通过企业的入职面试和学校的入学复试，方可进入联合培养程序，体现“双主体”协同、订单式培养的特征；其三，校企双方形成理念共识与责任共担，真正实现“我中有你，你中有我”的深度融合与联合育人。

（2）知识体系的梳理与课程体系的重构

将专业选修课划分成三个模块：数据系统类、智能算法类和应用交叉类，赋予标签，供师生根据学业规划来选择。教师依据研究成果反向“植入”艰深的工程原理，结合来自于合作企业的原创案例，以综合性、开放性的案例思考为牵引，激发学生深入掌握理论，深刻理解问题的意义和挑战，培养研究生高质量甚至创造性地解决实际问题的能力。

图2是新的课程知识体系。将整个课程分为多个大类。面向系统的课程板块，包括数据库系统实现、高性能计算与并行计算、大规模数据处理系统、区块链与分享型数据库、系统优化、数据库内核实践、当代数据库系统原理与实现、数据存储系统与技术、分布式事务处理等。面向数据基础板块，包括数据科学与工程理论基础、统计推断。在人工智能板块，包括机器学习、深度学习与强化学习、计算机视觉、深度学习、统计与机器学习等。除此之外，开设数据中台、计算教育学、专业英语、工程伦理、开源软件开发与社区治理、数据科学与工程研究方法、数据管理与分析前沿、软件开发管理与实践等应用型课程。以上课程“演进式”+“多维度”地构建学生的知识体系。图3则详细介绍了各门课程的核心技术内容。

在教学体系中，还需要重视实践教学与第二课堂。实践环节有助于增强学生的实践创新能力。在案例教学过程中，学生深刻体会到真实案例的重要价值，转变了以往将“实践简单等同于毕业实习”的片面认知，并在参与项目实践的过程中进一步端正了学习态度。联合实验室作为一类自主建设类基地，具备较高自主性和可持续性，学院与教师可借助持续稳定的项目，为研究生提供高度契合的研究方法训练与工程实践指导，帮助他们在真实业务场景中顺利完成从学生到初级职业者的角色过渡。第二课堂则是在长期校企合作背景下逐步形成的行业实验室模式，致力于应对现实行业中的具体问题。基于行业实验室所获得的研究成

果，教师将其中涉及的复杂工程原理进行反向整合，开发出原创教学案例并引入课堂教学。通过设计综合性和开放性的案例思考题，引导学生以小组讨论、辩论等形式深入剖析复杂工程项目，最终由教师进行总结与点拨，使学生不仅扎实掌握理论知识，更能深刻理解工程问题的实质与挑战，进而在实践中实现举一反三，提升高质量乃至创造性解决实际问题的能力。例如，由学院科研团队开发的在线教育平台（水杉平台）、区块链数据库系统（StarChain）、分布式数据库系统（Cedar）等，均已作为典型教学案例应用于研究生教学之中。

（3）实验室为载体的产教研融合人才培养

基于实验室导师制度管理和培养研究生。实验室承担明确的科研任务，包括校企联合实验室建设、企业合作项目以及面向国家重大需求的科研项目等。实验室导师直接负责实验室的管理和学生的培养。实验室导师与学术导师以及导师小组的其他成员一起负责把握学生培养的质量，深度参与博士生的培养方案制定、资格考试、开题、中期考核、预答辩等培养环节。研究生在联合实验室中针对应用的痛点，寻找和抽象问题，开展研究工作，并以应用为检验成果有效性最重要的标准。



图2 课程体系



图3 各门课程的核心内容

在这种模式下,校企双方充分发挥各自优势,进行深度融合与创新。课程内容不再局限于传统的理论知识,而是大量融入企业真实项目案例和实际应用场景,让研究生在学习过程中能够紧密接触行业动态,了解企业实际需求,培养他们解决实际问题的能力。共建联合实验室为研究生提供实践创新平台的重要支撑。实验室导师与企业导师协同互补,有效弥合了研究生培养中理论与实践相互脱节的短板。师生以研发项目为核心展开紧密协作,双导师分别从实践应用与学术研究两个维度对学生进行指导,从而全面提升其综合能力。

(4) 产研融合的师资队伍团队建设

学院构建相匹配的师资队伍,注重发挥团队中的教学名师的引领作用。调整并重组团队师资,盘活存量;增加教师资源供给,做优增量。首先,制定教师能力提升计划,自主培养具有工程实践思维和能力的教师。通过教师深度介入校企对接过程,深入了解业务场景、难题,并主导技术创新过程,积累了丰富的产学研合作经验,有效提升对实际问题的敏锐度、形成从理论知识到解决方案的育人模式,同时保障项目顺利实施、保障项目实践作为培养学生的基本属性。其

次,有意识地引进具有深厚企业工作背景的专家担任全职教师,这些“双师型”的教师,其行、企业工作经历将极大地丰富高校人才培养的维度、方法和方式,与其他教师形成互补。最后,“弹性引进”,即高校和学院吸收企业界专家型研发人员和管理者深度介入高校人才培养,包括共同研发课程并授课,开设前沿讲座和组织技术研讨等,指导学生项目和论文等。依托专业博士培养项目来建设授课专家库,企业专家纳入课堂讲课、暑期学校交流、研究生学术沙龙等,引入行业前沿,使得教学内容实时匹配产业发展。对于学生培养的成效需要有客观评价依据传统的以课程分数、发表论文作为单一评价维度的方式难以全面、真实反映学生的学习效果和学业水平,需要有破有立地建立多维、多元的评价效果。

学院着力打造结构匹配、能力突出的师资队伍,突出教学名师的示范引领作用,通过优化现有师资配置激活存量资源,并积极拓展引进渠道、提升新增师资质量。具体举措包括:首先,多举措提升与发展教师的能力,自主培育具备工程实践思维与能力的师资。通过鼓励教师深入参与校企合作,熟悉实际业务场景与核心问题,主导技术研发与创新,积累产学研协同经验,从而增强对现实问题的洞察力,构建从理论到

实践的育人机制,并确保项目实践在培养学生过程中的核心地位。其次,主动引进具有丰富企业工作经验的专业人才担任全职教师。此类教师能够将其行业经历融入教学,拓宽人才培养的途径和方式,与现有师资形成有效互补。最后,创新“弹性引进”机制,积极引入企业专家型研发与管理人才,深度参与课程共建、授课讲座、技术研讨及学生项目与论文指导。依托专业博士培养项目,建设企业专家授课资源库,将其纳入常规课堂教学、暑期学校、学术沙龙等环节,使教学内容紧密对接产业前沿。在学生培养成效评估方面,需构建更加科学多元的评价体系。传统仅依赖课程分数和论文发表的评价方式难以全面客观反映学生的真实能力与学习效果,亟待突破单一维度,建立多角度、多形式的综合评价机制。

(5) 教育科技支撑的工程实践

积极建设包括MOOCs、在线知识库与学习平台在内的线上虚拟教研体系。随着各领域数据量的迅猛增长,近年来涌现出诸如Hadoop、Storm、Spark等开源系统,显著降低了开发人员的学习与使用门槛。当前,我国已进入新的发展阶段,正迈向创新型国家前列的关键时期,破解重点领域“卡脖子”问题尤为迫切。这也对人才培养提出了更高要求——不仅应会熟练运用这些开源系统,更须能够深入掌握底层核心技术,具备开发具有自主知识产权的开源系统的能力。因此,在培养目标上,应突出塑造兼具“数据科学家”与“系统架构师”能力的人才,强调对学生原始创新能力和实践创造力的培养,引导其胸怀国之大者,走上推动行业转型升级与实施替代工程并重的发展道路,真正助力解决制约国家发展的核心技术难题。

作为集“教、学、练、测、评、创”全链路上工程实训平台,自主研发的在线学习和评测系统“水杉在线”,除了将教学资源上线,支持个性化培养模式,还创建数据驱动的工程能力过程性评价,自动记录用户的学习行为。平台收集的数据信息用于促进研究生教育的优化,比如因材施教,教学相长。

3 创新点

本人才培养模式的创新点主要体现在三方面。

(1) 本硕博一体化螺旋式上升的硬核科技人才培养体系

培育模式引导教师(低学段)和导师(高学段)深入应用,深刻理解目标理念,然后用过“本硕博”贯通的人才模式设计,避免了传统分段式教育中可能出现的知识断层,使学生能够在更长的周期内扎实地掌握学科知识,形成专业技能。通过将前沿课程/第二课堂与实践应用紧密结合,低年级学生接触到最新的

研究成果,培养创新思维和前沿视野;而高年级学生则能够将所学知识直接应用于实践,参与科研项目或企业合作,推动科研成果的落地与推广。

(2) 科技支撑的工程实践能力培养方法

首先,教学模式致力于创造真实的需求环境,将学生置于挑战中,从而激发其科技精神,激励他们去探索亟待解决的“真问题”,不断推动科技的边界。其次,开发在线支持平台,其本身既是科技创新赋能教育的真实场景,也是“应用驱动创新”的实践样板。将人工智能、区块链技术深度融入教学评价,破解过程性评价规模化难题,精准把握学生的学习状况,从而优化教学策略。再次,推行开源教育,以开源思想来指导学生的工程实践,鼓励学生能够使用开源软件,为开源社区贡献代码,并且产生了开源评估方案。

(3) “需求-培养-成果产出-应用转化”闭环式工程研究人才培养方式

技术创新的起点和终点都不在技术,而在应用需求。以产业需求为导向,赋能行业创新动力。依托校企联合实验室,设计多维度的培育方案,实现了多层次研究生对接多层次应用场景的闭环。此外,研究生教育资源共享,实现资源的配置闭环。遴选优秀博士进入国家专项,遴选部分全日制硕士依托国家专项、交叉平台等培育。面向国民经济领域,通过“有组织科研”培养其解决实际问题;与合作企业协同培育非全日制研究生,帮助在职学生深入了解工程实践中的解决方案,同时,学生反馈行业领域中实时业务场景进展,为教学提供丰富的素材;专业博士的培育中,提供跨学科交流的机会,学生则将行业中的技术需求和最新动态带入高校,增强科研的实用性和针对性。

4 实践成效

自从2013年华东师范大学成立国内首个数据科学与工程研究院(后于2016年改为数据科学与工程学院),团队就开始探索创新人才培养模式,经过10余年的研讨探索、实践完善、优化升级,形成了如下成效。

(1) 在培养前置与本硕博贯通培养方面

连续12年承办暑期学校。从2014年开始连续举办12届以“数据科学与工程”为主题的暑期学校,邀请合作企业和高校教师作为讲师,帮助研究生了解行业最新技术和市场趋势,拓宽视野;企业提前了解研究生的能力和潜力,为招聘优秀人才提供渠道;暑期学校开放对社会学员的注册,促进双向交流。以开放的态度向社会共享资源,促进产学研的良性互动。

毕业生受到用人单位的广泛认可。研究生就业情况优异,学生曾在国际算法大赛、CCF中国数字金融

赛、OceanBase 数据库大赛、华为软件精英挑战赛、机器人开发者大赛等赛事中获得一等奖等奖项。5 年来毕业生平均毕业去向落实率接近 100%，约 40% 的同学签约世界 500 强企业，涵盖华为、腾讯、阿里巴巴、小红书、美团、百度、快手等互联网头部公司，涉及互联网、电子信息、装备制造、金融等行业。优秀校友中，不乏入选江苏省 333 科技人才工程等殊荣，获得国家级、省部级科技进步奖等荣誉者。

表 1 暑期学校开设情况

年份	主题
2014	OceanBase系统：开源与应用
2015	数据科学与工程：大数据时代的新兴交叉学科
2016	面向科学发现的大数据管理
2017	面向新经济的数据管理系统
2018	面向异构体系结构的数据管理
2019	数据驱动的计算教育学
2020	数据中台：互联网时代的数据库
2021	数据驱动的系统优化
2022	面向科学和工程的人工智能
2023	智能时代的数据系统
2024	区块链技术及其创新应用
2025	智慧在线学习的本质与实践

2025 届研究生实现了 100% 就业，具体如图 4 所示，其实六成以上学生在互联网/科技大厂就业。

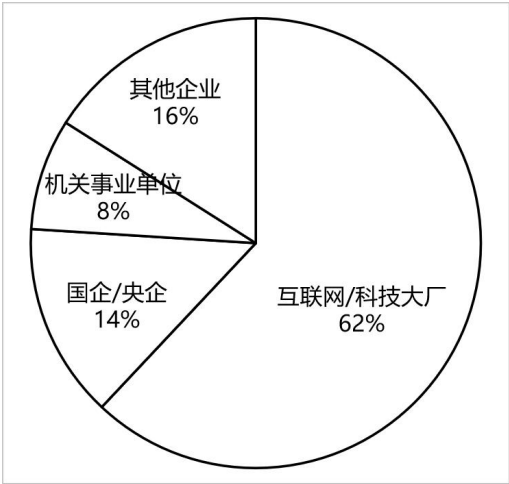


图 4 2025 届硕士研究生就业分布图

(2) 在知识体系梳理与课程体系重构方面

完成出版系列配套教材。鉴于数据学科发展迅速，缺乏相关适配的教材。项目组从课程讲义起步，经过多年整合、重构、迭代和完善，出版了 5 部相关教材。其中，《数据科学与工程算法》《数据科学与工程数学基础》作为数据学科的基础课程，从算法和数据角度打下坚实专业知识基础；《区块链导论：原理、技术与应用》《分布式计算系统》《云计算系统》面向专业选修课程，实现了国际国内最新研究成果与基础知识的体系化建设，展现了极高的学术价值和教学实用性。教材不仅得到了教育界专家的高度评价，还获得了出版领域权威机构——高等教育出版社的认可，进一步印证了该教材的卓越水平与广泛影响力。部分教材入选了战略新兴领域“十四五”高等教育教材名录，为数据研究生培育提供了一部兼具理论深度与实践指导意义的优质教学资源。

表 2 已出版的系列教材

年份	作者	标题	出版社
2021	高明，胡卉芪	数据科学与工程算法基础	高等教育出版社
2022	徐辰	分布式计算系统	高等教育出版社
2023	王伟，陆雪松，蒲鹏，张琰彬	云计算系统	高等教育出版社
2023	张召，金澈清，田继鑫，周傲英	区块链导论：原理、技术与应用	高等教育出版社
2024	黄定江	数据科学与工程数学基础	华东师范大学出版社

(3) 实验室为载体的产教研融合人才培养模式

共同建设了 17 家校企联合实验室。通过与各垂直领域的领先企业建立起广泛而深入的合作关系，组织科研力量解决企业中存在的真实需求。先后与蚂蚁集团、宝武集团欧冶云商等建立了 17 家校企联合实验室。研究生以这些联合实验室为实训平台，针对真实场景下的问题开展研究，研究成果迅速落地，解决实际问题。此外，有 10 家企业在数据学院设置了奖学金（奖教金）。

研究生参与的科创成果释放创新活力。通过高质量的校企合作，集中资源优势对国家重大需求进行科研攻关、实现重大原始创新突破，解决“卡脖子”问题的基础理论和关键技术难题。科研产出在各方面已取得了丰富的成果，从 2017 年开始，已经牵头获得 1 项国家科技进步二等奖、4 项省部级科学技术一等奖、1 项一级学会科技进步一等奖。

（4）产研融合的师资团队建设

构建了一支高水平的、具有校企联动经验的科研团队。采用两种方式构建教学团队。一方面，积极从产业界引进具有工程实践经验的专家，以解决复杂工程问题，目前团队中有9人具有企业工作经验。其次，积极通过工程项目和联合实验室培养在职教师，使其增加解决工程问题的能力。最后，通过双聘方式引进企业专家。

（5）教育科技支撑的工程实践

开源贡献：核心项目 OpenDigger 是一款面向开源生态的数据挖掘与信息服务系统，已在开源贡献评估、复杂网络分析、协作关系建模等前沿领域发表了超过80篇学术论文，推动了开源治理理论与实践的双向发展。OpenDigger 也荣获 2023 年度上海开源创新卓越成果特等奖，并成功入选 2023 Bench Council Top100 榜单。

水杉在线：作为教育部首批公布的18个“人工智能+高等教育”应用场景典型案例，团队自主研发的“水杉在线”，作为新一代大规模个性化在线智慧学习平台以数据驱动为核心，集成了“教、学、练、测、评、创”等多个环节，形成了一个综合性的在线虚拟学习社区，突破地域和场景的限制，辐射500多所高校10万余名学生。

5 结束语

数据已经成为驱动经济社会发展的新能源，但是

如何培养“硬核”数据科技创新人才仍然还在探索之中。本文在“应用驱动创新”理念驱动之下，设计了“硬核”数据科技创新人才的培养模式，已经在华东师范大学进行了十多年的探索，并且辐射到多所高校，在实践中收到了良好的成效。

参考文献

- [1] 周傲英, 钱卫宁, 王长波. 数据科学与工程: 大数据时代的新兴交叉学科[J]. 大数据, 2015, 1(2): 2015022. 引用页码.
- [2] 奚春雁. 加快推进计算机科学教育 做数据科学教育的探路人: 华东师范大学副校长周傲英教授专访[J]. 计算机教育, 2020(1): 2-8.
- [3] 王红, 苗强, 李慧. 基于产教融合的应用型人才培养研究与实践[J]. 计算机教育, 2021(1): 101-105.
- [4] 金澈清, 钱卫宁, 陈佳, 周傲英. 产教融合打通数据学科研究生培养的“最后一公里”[J]. 计算机教育, 2021(10): 85-88.
- [5] 毛嘉莉, 周傲英, 金澈清, 钱卫宁. 校企融合的非全日制专业硕士定制化人才培养模式[J]. 中国计算机学会通讯, 2024, 20(3): 69-73.
- [6] 何倩, 廖冰洁, 覃匡宇, 陈俊彦, 董庆贺. 基于产教融合的计算机网络课程思政建设与实践[J]. 计算机教育, 2024(3): 111-114, 119.
- [7] 金劲彪, 孙丽珍. 产教融合背景下产业教师队伍建设的制度保障[J]. 高等工程教育研究, 2024(1): 42-47.
- [8] 郭忠文, 将若冰. “产、教、学、研”四位一体研究生产教融合培养模式研究[J]. 计算机技术与教育学报, 2023, 11(2): 18-22.
- [9] 修佳鹏, 王莉, 陈虹雨, 周海伟, 邝坚. 产教融合的特色化示范性软件学院实践创新体系探索与实践[J]. 计算机技术与教育学报, 2023, 11(2): 75-80.