DOI:10. 19834/j. cnki. yjsjy2011. 2026. 03. 15　　CSTR:32289. 14. yjsjy2011. 2026. 03. 15

多模态视角下博士研究生就业质量智能评价体系及实践进路

王　霆,林申琦

(中国政法大学 商学院,北京 100088)

摘　要: 在推进高水平科技自立自强与产业结构升级的背景下,博士研究生作为国家战略人才力量的重要组成部分,其就业质量直接关系到教育链、人才链与创新链的协同效能。本研究聚焦博士研究生就业区域失衡、职业错位与能力错配的三重困境,探讨运用多模态数据融合技术的价值意蕴和实施逻辑,构建涵盖就业求职、初次就业与职业发展的全周期就业质量智能评价体系。运用 LDA 主题模型、德尔菲法与层次分析法,系统提炼博士研究生就业质量评价指标,融合结构化数据、文本数据与行为数据,运用多模态大模型算法,实现动态价值引领、个体能力画像与政策治理协同等智能评价结果应用。本研究提出建设三级联动多模态数据仓、开发双维度智能诊断工具、实施分阶段精准施策等实践进路,推动博士研究生从"学术单通道"向"多元高质量"转型。

关键词: 博士研究生;就业质量;多模态数据;智能评价;人才培养

中图分类号: F015;G643　　**文献标识码:** A　　**文章编号:** 2095-1663(2026)03-0118-10

《国民经济和社会发展第十五个五年规划纲要》在"一体推进教育科技人才发展"部分明确指出"深化教育科技人才一体改革,强化规划衔接、政策协同、资源统筹、评价联动,促进科技自主创新和人才自主培养良性互动。"博士研究生作为国家高水平科技自立自强的战略储备人才队伍,其就业质量已不仅是教育议题,更是关乎国家教育链、人才链与创新链深度耦合的治理命题。然而,随着全球教育结构调整、学术科研单位岗位饱和、科技革命及产业变革的加速演进,博士研究生的就业也面临着诸多挑战,如毕业生总量增加、就业结构失衡、就业质量不高、就业风险增加等。基于此,本文通过厘清博士研究生就业质量的现状与困境,在多模态视角下探讨博士生就业质量智能评价体系及实践进路,为数智语境下算法治理与政策模拟提供可靠的数据支撑,推动博士生能力结构与国家创新需求精准对接。

一、博士研究生就业质量的现状与困境

就业质量研究始于 20 世纪初,"科学管理之父"弗雷德里克·温斯洛·泰勒(Frederick Winslow Taylor)从企业的视角提出"就业质量"的概念。国际劳工组织提出的"体面劳动"(Decent Work)和欧盟理事会提出的"工作质量"(Quality in Job)与就业质量有共同的关注点:三者不仅关注劳动者的工作稳定性、薪资待遇、工作环境、职业发展等客观因

收稿日期: 2026-01-20

作者简介: 王霆(1976—),男,江苏南京人,中国政法大学商学院教授,博士生导师,中国政法大学人力资源开发与管理研究中心主任,中国劳动学会劳动科学教育专委会副主任委员。

林申琦(1999—),女,山东日照人,中国政法大学商学院博士研究生。

基金项目: 国家自然科学基金面上项目"基于数据挖掘的我国大学生高质量就业研究:评价体系、影响因素和实现路径"(71874205);中国政法大学科研创新年度规划(中央高校基本科研业务费专项资金资助)项目"新质生产力对地区就业韧性的影响效应及机制研究"(25KYGH001);中国政法大学钱端升杰出学者支持计划资助项目

素，还涉及个人满意度等主观感受。高校毕业生就业质量是对毕业生满足社会潜在需求情况的反馈。整体来看，博士研究生群体就业质量呈现以下特征。

（一）就业区域分布不均

当前博士毕业生就业呈现出明显的区域集聚与流动失衡特征。东部地区凭借雄厚的经济基础、丰富的科研资源与优渥的人才政策，对博士毕业生形成强大的“虹吸效应”。数据显示，东部博士毕业生本地留存率超过八成，而中部地区最低[1]，部分省份博士授权点增长滞后、科研平台不足，博士“产地”人才供给与“户籍地”岗位需求错位，毕业生即使有意返乡，也难以获得匹配岗位。

西部地区虽博士点扩张迅速，2013—2019 年博士点由 532 个增至 667 个，增幅为 25%，高于东部的 19%[2]，但科研经费、产业平台、岗位层次仍相对薄弱，导致“产出”远超“吸纳”。部分西部籍博士择业预期偏高，就业岗位以高校等学术部门为主，选择企业就业者较少[3]，导致宁可待业或延长博士后阶段，也不愿下沉至地方高校或企业，进一步推高待业率[4]。区域间发展不均衡、资源配置差异与社交网络依赖，成为博士毕业生区域流动失衡的重要结构性原因。

（二）职业选择非学术化

博士毕业生职业路径正逐步从“学术唯一”向“多元就业”转变。我国顶尖研究型大学的博士毕业生进入学术机构工作的比例已不足 50%，选择到企业、党政机关等其他非学术领域就业的博士毕业生比例正在不断上升[5]，博士学位与学术职业的解耦已成为一种趋势。在传统教育体制中，博士教育的基本目标是培养从事学术职业的科研人员。尽管当前博士研究生职业选择仍以高校科研岗为主，企业、事业单位为辅，但是非学术化趋势愈发明显，北京大学（49.1%）、清华大学（44.1%）这两所顶尖研究型大学，此类情况占比已接近一半[6]。就学科而言，经济学、法学、管理学的博士毕业生更倾向于选择行政单位就业，理学、农学博士毕业生倾向于选择其他事业单位就业，除此之外，理学、工学博士毕业生还倾向于民营企业和三资企业就业[7]。

这一现象既源于博士个体内在动机的转变，也受到外部就业结构变化的推动。一方面，部分博士在读期间逐渐意识到自身科研兴趣不足，倾向于寻求更稳定或更具社会影响力的岗位；另一方面，学术岗位趋于饱和，高校编制收紧、晋升压力大、考核机制严苛，削弱了学术职业的吸引力[8]。总体来看，博士择业“非学术化”既是个体理性选择的结果[9]，也是学术劳动力市场供需失衡的反映。

（三）就业能力有待提升

博士生的就业能力是决定其能否顺利就业的重要因素，主要来源于其在研究生阶段所获得的高层次人力资本[10]。尽管当前博士研究生就业率维持在较高水平，但是“高学历—低适配”的现象日益凸显，原因在于博士研究生的就业能力与劳动力市场的岗位需求发生结构性错配。

博士研究生在求职过程中，呈现学术能力单维突出、其他方面能力较弱的特征，博士生培养已偏离就业市场的需求[11]。在博士教育期间，博士培养长期遵循“专业—科研—论文”单线研究模式，必修或选修课程也侧重培养研究能力、分析能力、写作能力和专业知识[12]，导致博士研究生在沟通协作、项目管理、商业敏锐度等横向技能上严重缺失。面对学术岗位有限的现状，博士研究生就业方向转型迫在眉睫，仅具备较高科研能力的博士研究生在复杂的劳动力市场中竞争力相对较弱，甚至无法发挥其学历优势。Roy 等学者将韧性、团队合作、管理等软能力纳入博士研究生就业能力测量指标框架[13]。

综上所述，博士研究生就业质量面临“区域失衡—职业错位—能力错配”三重困境，既有结构性制度约束，也暴露出培养体系与市场需求脱节的问题。构建科学、动态、精准的就业质量智能评价体系，已成为推动博士研究生教育链与人才链协同升级的迫切需求。通过文献回顾发现[14-15]，现有相关研究得出的结论主要源于深度访谈或问卷调查两类数据，此数据收集方式受地区和个体人口统计学特征的限制，往往缺乏一定的代表性与全面性，难以科学、准确、动态地评价博士研究生就业质量。针对此困境，本文探究运用多模态数据融合技术的价值意蕴和实施逻辑，构建博士研究生就业质量多模态智能评价体系并提出实践进路，为我国培养高层次创新型人才提供扎实的研究基础。

二、博士研究生就业质量多模态智能评价体系的价值意蕴

多模态大模型是一类基于深度学习架构（如 Transformer）的大规模预训练模型，能够联合建模文本、图像、音频等多种模态数据，实现跨模态理解与生成任务。在教育学领域，多模态大模型已经应用到思想政治教育、在线学习、认知能力测评等多个场景，本文尝试在多模态视角下构建博士研究生就

业质量智能评价体系,推动人才培养从传统模式向深度融合、智能化方向积极转型,为高等教育领域提供全新的技术支撑和实践进路。

(一)实现动态价值引领,适配多元化就业生态

传统静态评价体系难以捕捉快速变迁的就业市场对博士人才的价值定义。多模态评价体系则依托动态知识图谱技术,实时抓取并分析国家科技政策、产业报告、招聘数据等海量信息,精准描绘不同职业赛道对博士能力要求的变化轨迹,从而实现对博士职业价值观的动态引导。例如,当系统智能识别到"人工智能伦理治理"或"碳中和前沿技术"等新兴交叉领域的人才需求在快速增长时,可即时生成行业前沿发展报告与人才需求能力清单,通过就业指导平台向相关学科博士生及其导师进行精准推送,使博士生培养活动能够紧扣社会需求变化,动态调整博士生的知识结构和职业预期。这种市场人才需求信号驱动下的动态反馈机制,打破高校博士生人才培养的静态学术价值导向,构建了一个博士生就业供需智能匹配的职业价值导航系统。它不仅精准地化解了博士生的职业认同焦虑,更从源头上引导高层次人才流向国家社会最需要的领域,实现个人价值与社会发展需求的动态统一。

(二)刻画个体能力画像,提供个性化发展导航

博士研究生就业质量多模态评价体系的核心突破在于能对每位博士研究生的可迁移能力进行动态、精准、智能的画像。它不再依赖单一的论文指标,而是通过动态采集博士生在学期间的课程项目、科研实践、社会服务、竞赛活动等多模态行为数据,利用智能模型(如自然语言处理、社会网络分析)进行评估。系统能够精准识别出博士研究生在跨学科团队协作、复杂问题建模或科研成果转化等方面的优势与短板,并生成一份动态更新的个人能力发展报告。在此基础上,多模态评价体系能智能地为每位博士生"量身定制"发展建议,例如向一位展现出强沟通能力但与产业界接触较少的工科博士生,精准推荐参与企业开放课题或创新创业实训营。高度个性化的、基于动态数据的发展导航,可使博士研究生的能力培养从"大水漫灌"转变为"精准滴灌",为其应对高度不确定性的未来职业市场奠定坚实且全面的能力基础。

(三)打造智慧治理平台,优化宏观政策效能

在宏观层面,多模态评价体系作为一个动态感知和智能分析的"仪表盘",为国家和区域的博士研究生人才治理提供了精准施策支持。传统的人才政策制定往往依赖于抽样调查、宏观统计年鉴和有限的专家经验,存在数据滞后、维度单一、难以进行政策预评估等局限。多模态评价体系则通过构建一个互联互通的数据共享平台,动态整合个人职业发展数据、高校培养数据、劳动力市场数据、产业创新数据,利用多模态数据融合技术,精准揭示博士人才的跨区域流动规律、产业集群效应、创新能力与区域经济发展的关联强度,乃至特定政策干预对博士就业选择与职业发展的实际影响。例如,地方政府在区域规划设立新研发中心时,可通过该平台精准查询到全国相关领域博士的存量、分布与流动意愿,从而制定出更具针对性的人才引进政策。这标志着博士研究生就业工作从传统的经验化、碎片化管理,迈向了一个基于全息数据、具备动态反馈和科学决策能力的智慧治理新时代,极大提升了国家高层次人才资源的配置效率。

三、博士研究生就业质量多模态智能评价体系的实施逻辑

本文探讨如何基于多模态数据融合技术智能评价博士研究生就业质量,在文献研究基础上,通过LDA主题模型文本分析法和德尔菲专家决策法的主客观分析,构建博士研究生就业质量评价指标体系,运用层次分析法确定就业质量各评价维度之间的权重。借助反映各指标的多模态数据融合技术,对博士生就业质量开展智能评价,并运用大数据挖掘就业质量影响因素,整体实施逻辑如图1所示。

(一)构建博士研究生就业质量评价指标体系

在系统梳理国内外就业质量的经典著作与相关文献基础上,本文认为博士研究生作为高层次研究人才的重要储备力量,需要动态考察其就业质量,覆盖就业求职、初次就业、就业后职业发展各阶段。不仅要关注博士生就业时的供需匹配情况,也要关注其就业后动态发展情况;不仅要关注博士生薪酬福利、职业发展、成果产出等方面的就业客观收益,还要关注其工作满意度、职业成就感、工作压力感知等就业主观感受。因此,本文从就业求职阶段、初次就业阶段、职业发展阶段三个时段采集博士研究生就业质量评价指标,形成整体研究框架(见表1的第1列)。

根据科学性、全面性、独立性等指标选择原则,在上述博士生就业质量评价框架下,本文运用隐含狄利克雷分布(Latent Dirichlet Allocation,LDA)模型开展文本聚类分析,从供给与需求两方面提炼

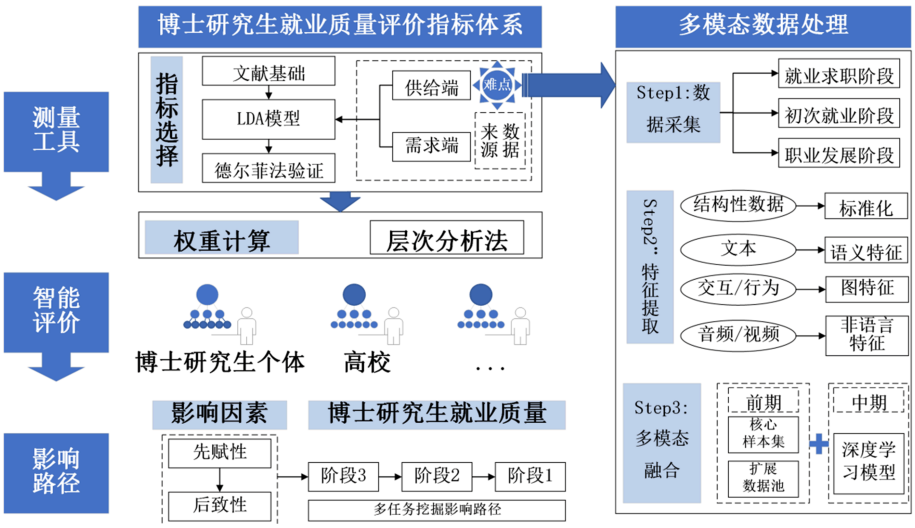

图 1　博士研究生就业质量多模态评价体系的实施逻辑

测量指标，并采用德尔菲背对背决策法对指标进行筛选验证。其中，LDA 模型是 Blei 等提出的一个三层概率生成模型。作为一种对文本主题信息进行建模的方法，它能够挖掘文字背后的语义关联，使得文本处理更加智能化[16-17]。从需求方来看，博士研究生就业去向主要是科研院所、高等院校和高科技企业，本文选择代表性用人单位进行走访调研，围绕博士研究生的人才引进渠道、招聘现状、职业规划等展开深入访谈并做好详细文字记录。从供给方来看，博士研究生经常在微博、小红书等社交软件发表评价文字[18-20]，本文通过 Python 爬取关于就业的相关评价文字，与访谈记录相整合，清洗并删除无意义的评价数据，再运用 LDA 模型开展主题聚类，查找数据中的潜在变量或隐藏结构，与前期文献梳理指标进行汇总融合。在此基础上，本文选择就业领域资深学者背对背访谈，确定最终博士研究生就业质量评价指标（见表 1 第 2 列），并采用层次分析法计算各维度和测量指标的权重。

（二）博士研究生就业质量评价多模态数据处理

在多模态研究中，数据处理是关键，其核心技术难点在于数字、文本、音频、视频、图像等异构数据的对齐、融合与表征。本文通过数据采集与模态界定、数据清洗与特征提取、数据融合处理等步骤，实现博士研究生就业质量的多模态智能评价。

首先，数据采集与模态界定。本文采用多源融合的数据采集策略，在确定表 1 博士生就业质量评价指标体系后，明确各类数据的采集来源和模态属性（见表 1 第 3、第 4 列）。主要数据来源包括：一是公共数据源，如高校年度就业质量报告、教育部门就业数据库等；二是网络数据源，如在线招聘网站的求职简历、招聘职位描述、浏览互动次数、学术成果数据库及知乎、小红书、微博等社交媒体数据等；三是调查数据源，通过问卷调查和深度访谈获取的博士研究生结构化数据和主观感知数据。本研究针对不同数据进行模态界定，包括结构性数据、文本数据、交互/行为数据、主观感知数据等。

其次，数据清洗与特征提取。将不同模态的原始数据转化为机器可读、可分析的特征向量，对结构化数据通过归一化、标准化处理，同时处理缺失值；对文本数据去噪处理，建立有关博士生就业的语料库，使用词嵌入模型以及预训练语言模型获取深度语义特征；对交互/行为数据构建网络图，计算节点中心性、密度等图特征；对于音频/视频等数据通过文本转译，并提取语速、音调、面部表情等非语言特征。当然，上述数据清洗处理过程要通过数据脱敏流程，运用匿名化处理技术，符合数据安全与伦理合规要求。

再次，多模态数据融合。在处理多模态数据时可以采用前期融合、中期融合和后期融合三种策略。考虑到博士生就业质量评价数据综合了宏观与微观、静态与动态、客观与主观等多维数据，而同一个体样本难以获取如此多的数据，这是不现实也没有必要的。多模态数据融合追求的是获取足够多的样本和更为丰富的维度，使得在群体层面构建的评价模型是稳健可靠的。因此，本文将数据分为“核心样本集”和“扩展数据池”。核心样本集通过招募五年内毕业、覆盖不同代表性高校和学科的博士生志愿者，在告知数据用途和隐私保护措施基础上，志愿者签署知情同意书，这部分数据样本量不求大但要求全维度、高质量，尽可能获取全部多模态数据作为模型训练与验证的“基准真相”。在此基础上，通过一些非敏感的统一标识符来连接不同来源的数据，比

如利用“毕业院校+学科专业+毕业年份”数据,对匿名化的招聘平台求职群体和高校就业去向群体进行群体级别的匹配,从而形成样本量大但每个个体数据维度并不完整的扩展数据池。再通过中期融合策略,运用深度学习模型(比如 Transformer、跨模态注意力机制)完成数据处理,实现模型在中间层自动学习不同模态特征的交互关系,完成智能化动态数据融合。

表 1 博士研究生就业质量评价指标体系及多模态数据结构

时间节点	测量指标		数据来源	数据模态
就业求职阶段(就业前)	投递简历次数		博士生问卷调研	结构化数据
	参加面试次数			
	获取 offer 数量			
	工作环境		求职简历 招聘职位描述	文本数据
	能力要求			
	职责描述			
	求职行为偏好		招聘网络 社交媒体	交互/行为数据
	与雇主互动热度			
初次就业阶段(入职起始)	薪酬待遇		博士生问卷调研	结构化数据
	专业匹配			
	岗位类型(科研、教学、管理等)			
	单位性质(政府、事业单位、企业、国际组织等)			
	工作合同类型(长聘、预聘、临时等)			
职业发展阶段(就业后五年内)	主观感受	工作满意度	深度访谈录音 视频自述	音频/视频等主观感知数据
		工作稳定性		
		职业成就感		
		工作压力感知		
	客观成果	科研成果数量	论文、专利、项目等 科研成果数据库	结构化数据
		科研成果等级		结构化数据
		科研合作关系		科研成果统计图片等 交互数据

(三)多模态数据智能评价博士研究生各层级就业质量

通过多模态数据融合,运用深度学习算法可以开展博士研究生就业质量多层级智能评价,满足博士生个体、学科、高校以及地区的不同应用需要。比如,在博士生个体层面,生成个性化就业质量评估报告,包括就业质量水平以及与同类群体的比较分析,为其职业发展提供参考;在学科和高校层面,形成博士生就业质量综合评估动态分析报告,从就业质量视角建立专业人才发展预警机制,为教育改革提供数据决策支持;在地区层面,测算就业质量发展指数,开展长期持续动态追踪研究,分析博士生就业质量和人才培养发展趋势。

(四)梳理博士研究生就业质量的微观影响因素及其测量指标

接下来,需进一步分析影响博士研究生就业质量的各类因素,探讨这些因素对博士研究生就业质量的影响机制,为模型训练提供原始特征值及其数据来源。博士研究生就业质量的影响因素包括经济增长、产业结构、高等教育投入等宏观因素,以及人力资本、社会资本等微观因素。对于博士研究生而言,微观因素对其就业质量产生更为直接的影响,因

此本文主要考察博士生个体微观因素的影响。本文通过访谈法，深入了解博士研究生的就业选择、就业能力与职业规划，剖析其就业质量特征的成因。本文从就业供需双方选择合适对象，最终选择 20 名具有五年工作经验的博士、10 名科研院所的负责人与 10 名企业负责人进行实地访谈，从人力资本、社会资本、人口统计学角度剖析就业质量的影响因素（见表 2）。

表 2　博士研究生就业质量的影响因素及其测量指标

影响因素	维度	测量指标
人力资本	附着性人力资本	毕业高校声誉
		学科评估等级
	主体性人力资本	综合成绩排名
		在校科研成果
		参与社会实践
社会资本	先赋性社会资本	父亲教育水平
		母亲教育水平
		家庭年收入
		父亲的职业
		母亲的职业
	后致性社会资本	社会网络规模
		社会网络跨度
		社会网络密度
人口统计学变量	性别、年龄、生源地等	

人力资本是体现在人身上的技能和生产知识的存量，能够在就业中被用人单位认可，企业也愿意为之付酬，这些要素往往是博士生通过接受教育投资和参与社会实践而形成的[21]。对于博士研究生而言，其人力资本的构成因素主要包括毕业高校声誉、学科评估等级等附着性人力资本，以及个人的综合成绩排名、在校科研成果、参与社会实践等主体性人力资本。

社会资本是社会关系嵌入资源的总和，它可以对生活在特定社交网络中的人们产生优势和预期回报。对于博士研究生而言，影响他们职业选择的社会资本因素可以分为两个维度：先赋性社会资本和后致性社会资本。先赋性社会资本往往采用家庭社会经济地位（Social Economic Status，SES）来测量，一般包括父母的受教育程度、家庭收入以及父母职业等变量[22]。而后致性社会资本是指个体经过后天社会网络交往所获得的积累性社会资本，往往采用社会网络分析模型中的网络规模、网络跨度与网络密度三个维度来测量[23]。此外，性别、年龄、政治面貌、生源地、专业等人口统计学变量对于就业质量也会产生影响。相关影响因素的数据主要来源于博士毕业生志愿者填写的调研问卷，且采集数据均为结构化数据，可与前述“核心样本集”数据相匹配。

上述影响博士研究生就业质量的微观因素可以分为先赋性因素和后致性因素。先赋性因素包括性别、年龄、学前户籍和先赋性社会资本，这些是博士研究生自身不能改变的因素。后致性因素包括后致性社会资本和人力资本，这些因素是可以由博士研究生自己积累的（见图 2）。初始地位比较高的人，无论该地位是先赋的还是后致的，往往可以获得更多的社会资源，得到更好的工作[24]。

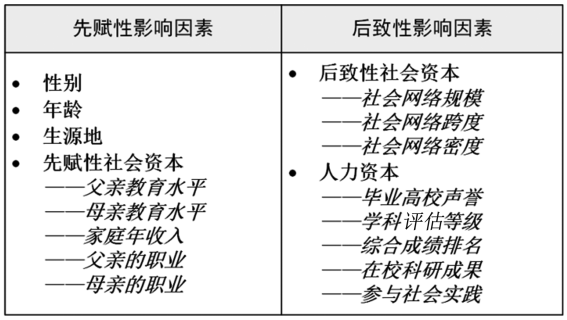

图 2　博士研究生就业质量微观影响因素分类

（五）多任务数据挖掘探索博士研究生就业质量影响路径

本文采用多任务数据挖掘方法，揭示影响博士研究生就业质量的关键因素以及各组变量之间复杂的内在影响机制，探索实现博士生高质量就业的有效路径，为优化博士生人才培养政策提供实证依据。运用多任务数据挖掘方法来探索各微观因素对博士生就业质量的影响机制，首先需要构建研究问题的系统模型（见图 3）。本研究将影响博士生就业质量的微观因素分成先赋性因素（Group I）、后致性因素（Group II），将博士研究生就业质量按照动态时点分为就业求职阶段（Group III）、初次就业阶段（Group IV）和职业发展阶段（Group V）。为了有效探索各原始特征值之间的内在关系，本研究将探索性研究分解为四个数据挖掘任务，即任务 A、任务 B、任务 C 和任务 D。

任务 A 是通过数据挖掘来探索博士生初次就业质量对其后期职业发展的影响。该任务旨在揭示博士生就业起点与职业长远发展的动态关联，追踪初次就业质量多维指标对其职业成长速度、学术影

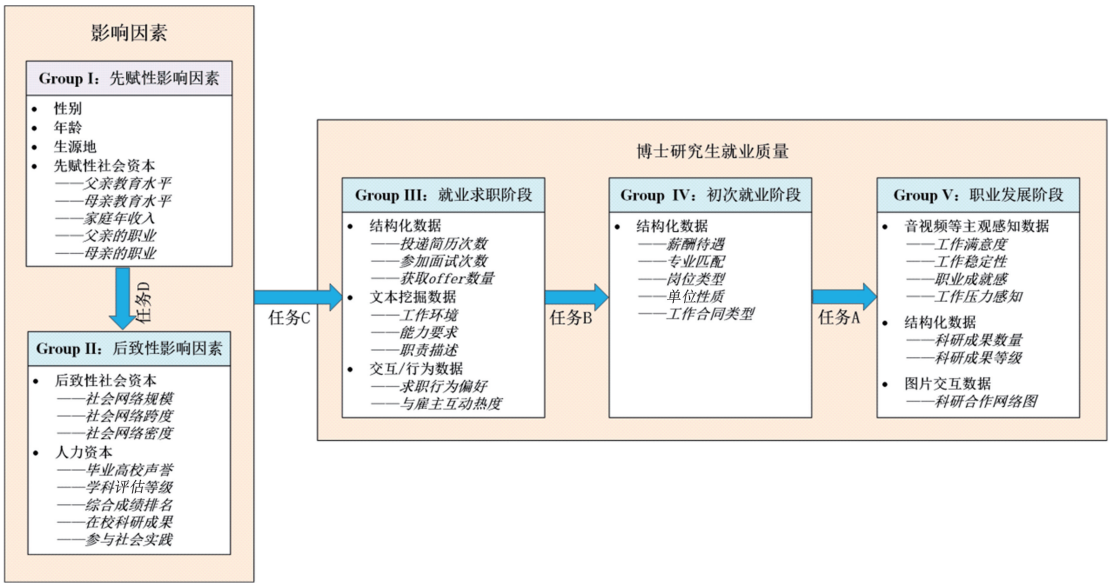

图 3 多任务数据挖掘探索博士研究生就业质量影响路径

响力及职业稳定性的滞后效应,旨在突破传统截面研究的局限,为博士生就业指导的前置干预策略提供实证依据。

任务 B 是通过数据挖掘来探索博士生就业求职经历对其初次就业质量的影响。该任务聚焦求职过程中的关键行为变量(如信息搜寻强度、职业决策策略)对就业结果的影响机制,运用序列模式挖掘技术,识别求职行为的最优路径,旨在揭示博士生人力资本向职业资本的转化过程,从人才市场供需匹配视角重塑博士生就业能力培养体系。

任务 C 是通过数据挖掘来探索博士生先赋性因素和后致性因素对其就业求职经历的影响。该任务通过构建家庭背景—教育经历—求职行为的因果链模型,量化分析个体先赋性因素与后致性因素对求职过程的差异化影响及交互效应,旨在识别促进就业公平的关键变量,为构建补偿性支持政策、阻断代际劣势传递提供数据支撑。

任务 D 是通过数据挖掘来探索博士生先赋性因素对后致性因素的影响。该任务深度挖掘家庭社会经济状况等先赋性变量如何通过影响博士生科研产出、社会资本积累、人力资本建构等后致性因素,形成隐蔽性的优势累积机制,旨在揭示教育过程中可能出现的不平等传递路径,批判性审视博士生培养体系的公平性。

本研究采用多任务数据挖掘方法,揭示各原始特征值及其分组数据对博士生各阶段就业质量的影响机制,从而基于大数据提炼博士研究生人才成长路径,推动高层次研究人才的高质量培养。

四、博士研究生就业质量多模态智能评价体系的实践进路

前文系统构建了涵盖就业求职、初次就业、职业发展三阶段的博士研究生就业质量多模态智能评价体系。该体系突破传统单一数据源局限,创造性地融合结构化数据、文本数据、交互/行为数据及主观感知数据,形成“核心样本集”与“扩展数据池”协同的异构数据融合架构。根据前述智能评价体系的实施逻辑,本研究构建“数据采集—智能诊断—精准施策”三位一体的实践进路,将多模态评价指标转化为可落地、可监测、可迭代的治理工具,实现从静态评估到动态治理的范式跃迁。

(一)建设三级联动多模态数据仓,搭建博士生就业质量数据共享平台

数据采集是多模态博士生就业质量评价体系的基础与前提。构建系统化、标准化、安全可控的数据采集与共享机制,能够打破传统评价中数据来源单一、信息孤岛等问题,实现多源异构数据的深度融合与持续供给,为后续智能诊断与精准施策提供坚实的数据基础,是推动博士生就业评价从经验判断走向数据驱动的关键支撑。

一是国家整合多源数据。国家层面建设“博士生就业质量多模态数据总仓”,整合教育部全国研究生教育质量信息平台的结构化数据、主流招聘平台的文本与行为数据、学术成果数据库的科研产出数据,以及微博、知乎等社交媒体的主观感知数据。总仓采用“湖仓一体”(一种融合数据湖与数据仓库的

现代数据架构）架构，对结构化数据实行归一化处理，对文本数据构建就业领域专用语料库并运用Word2Vec模型提取语义特征，对行为数据构建动态网络图计算节点中心性，对音视频数据通过语音识别与情感计算提取语速、音调、面部表情等非语言特征。所有数据经差分隐私技术处理，确保个体信息无法追溯，仅保留群体分析所需的统计特征。

二是区域聚焦属地需求。区域层面建设省级数据分仓，聚焦属地产业需求，实时抓取本地龙头企业、专精特新“小巨人”企业的技术岗位变动、薪酬曲线及招聘文本，通过自然语言处理（Natural Language Processing，NLP）技术分析职位描述的能力要求，运用跨模态注意力机制将企业需求数据与高校培养数据进行对齐，回传总仓完成供需校准。例如，长三角分仓可重点抓取集成电路、生物医药等战略性新兴产业的岗位数据，成渝分仓则聚焦电子信息、航空航天等领域的招聘需求。

三是高校采集行为数据。高校层面建设校级数据节点，负责采集博士生从入学到毕业的全培养周期行为数据，包括课程项目参与度、实验室打卡记录、社会实践时长、竞赛活动轨迹等，每周定期将数据汇总推送至上级仓。为保护博士生隐私权，所有数据采集须遵循“知情同意、最小必要、目的限制”三原则，由博士生签署电子授权书，明确告知其数据用途与保存期限，并提供“一键退出”机制。配套出台《博士生就业质量数据安全管理条例》，规范数据使用流程，所有数据操作须经隐私安全评估后方可进行，并同步建立操作日志的区块链存证机制，确保全程可追溯、不可篡改。

（二）开发双维度智能诊断工具，构建博士生就业质量动态预警系统

智能诊断工具是将多模态评价指标转化为可操作、可视化、可预警的管理工具，实现对博士生就业质量的早识别、早预警与早干预。通过面向个体与群体的双维度诊断，不仅能够为博士生提供个性化发展指导，也能为学科、高校与区域层面的政策调整提供数据支持，推动就业工作从事后统计向事前预警和事中干预转型。

一是开发个体层面的应用工具。面向博士生个体，开发“就业质量智能诊断”微信小程序，整合前文验证的多模态评价指标，生成三种动态应用工具：首先是能力画像报告，融合结构化数据（学业成绩、科研产出）与文本数据（学术论文关键词、简历文本），运用LDA模型与预训练语言模型提取深度语义特征，识别博士生“跨学科协作”“复杂问题建模”“科研成果转化”等维度的显性优势与隐性短板，并以雷达图形式直观呈现。其次是就业准备度指数，实时计算社会资本规模、简历匹配精度等指标的加权得分，生成0—100分的综合指数。当指数低于40分时，红灯预警，系统自动锁定其就业指导中心账号，强制要求预约求职行为矫正咨询，并推送产业导师优先匹配名额与心理援助热线。最后是职业轨迹预测，运用时序模型分析博士生行为数据与往届毕业生轨迹的相似度，预测其未来六个月的就业风险等级与职业发展瓶颈，提供“学术－产业”双通道的发展建议。所有诊断报告和应用工具均遵循“数据可用不可见”原则，原始数据加密存储于高校私有云，仅向学生本人展示脱敏后的分析结果，导师需经学生授权后，方可查阅其数据的摘要信息，以保障学生的数据知情权与自主决定权。

二是建立群体层面的动态系统。面向群体层面，建立学科、高校、区域三级就业质量动态预警系统。首先在学科层面，每月生成《博士生就业质量学科健康度报告》，重点监测专业匹配度、区域留存率、职业成长斜率等核心指标。当某学科专业匹配度连续三届低于全国均值10%时，系统会提示红色预警，推送微专业增设建议方案。例如，若某校生物学博士专业匹配度持续偏低，系统通过招聘文本数据发现大量生物信息学岗位需求未被满足，则建议该学科增设生物信息学微专业，课程体系包括机器学习在生物学的应用、基因组数据分析等模块，由数据仓智能匹配行业专家授课。其次在高校层面，每学期形成《博士生就业质量校际对标分析报告》，从培养－需求匹配度、就业－成长转化率等维度与同类院校横向比较。若某校毕业生区域留存率低于平均水平15%，系统启动地域适应力提升计划，推送与地方政府、龙头企业共建产业联合培养基地的具体方案，包括基地数量、合作模式、经费预算等细节。最后在区域层面，打造博士人才供需驾驶舱，实时呈现三大核心功能：流动监测功能动态展示毕业生跨区域流动轨迹，当监测到某区域人才净流出率单月超过5%时，自动推送人才政策紧急评估提醒和调整建议；产业匹配功能实时计算各产业集群的博士人才饱和度与创新能力关联指数，为地方政府区域发展规划提供人才储备充足度评级；政策模拟功能通过输入拟调整的科研经费投入、住房补贴力度等政策参数，系统模拟该政策对三年后博士就业选择与职业发展的影响曲线，避免地方间人才补贴陷入

恶性竞争的困局。为恪守数据伦理，系统仅提供群体级分析，个体数据在聚合时自动脱敏；建立数据使用审计委员会，定期审查算法公平性，防止对特定地区、学科、人群产生系统性偏见。该系统将前文静态指标体系转化为动态治理仪表盘，使管理者能够一屏观全局、一键调政策。

（三）实施分阶段精准施策，提升博士生就业质量智慧治理能力

精准施策是基于多模态智能评价与诊断结果，将数据洞察转化为博士生就业质量智慧治理行动的关键环节。根据前述多任务数据挖掘方法，探索博士研究生就业质量的关键因素和有效路径，实施分阶段、差异化、系统性的干预措施，有效回应博士研究生就业质量提升的复杂诉求，实现从数据驱动诊断到政策智慧生成的闭环治理。

一是建立动态跟踪与补偿机制，破解职业发展起点依赖困境。教育部门应联合人社部门，依托多模态数据仓，为每位博士毕业生建立为期 5－10 年的职业发展电子档案，持续追踪其岗位流动、薪酬增长、成果产出与主观满意度。系统需自动识别高风险起点群体（如初始岗位严重不匹配、起薪显著低于同类均值），并触发补偿性支持政策。例如，对入职非优势区域基层科研单位的博士，自动匹配“学术青椒托举计划”，提供额外的科研启动经费、访问学者名额及跨校合作网络接入权限，以弥补其初始平台劣势。

二是设计最优求职行为的精准干预方案，提升职业搜寻效率。各高校就业指导中心应升级为智能就业能力发展中心，系统通过分析学生的简历投递模式、面试表现状况、雇主互动数据等，为学生生成个性化的求职行为优化清单。例如，对存在“海投”但面试转化率低的学生，系统可推送针对性的模拟面试训练模块；对信息搜寻效率低的学生，则智能推荐其尚未触达但匹配度高的隐形招聘渠道与产业导师。通过靶向矫正求职行为，缩短人力资本向职业资本转化的周期。

三是构建博士生差异化的就业支持体系，促进就业机会实质公平。政策设计需从“一刀切”转向“梯度支持”。首先，面向家庭社会经济地位较低的博士生，设立专项基金，不仅为其提供经济资助，更匹配学术导师与职业导师双轨指导，并资助其参与国内外高端学术会议与产业实训，以快速弥补在社会资本与学术视野等方面存在的差距。其次，在招聘环节，鼓励重点用人单位在初审环节引入盲审机制，聚焦后致性能力指标；在政策考核中，对积极聘用多样化背景博士生的用人单位给予倾斜支持。通过这些措施，减小先赋性因素的直接约束，增强后天努力的回报率。

四是实施促进机会均等的公平支持计划，保障培养过程内涵公平。高校需对内部资源配置与评价机制进行“公平性审计”。关键举措包括：推行“研究资源公平获取计划”，确保学生享有重大科研项目参与、先进实验设备使用的机会，依托预约系统与成果署名规范，向所有博士生尤其是无强学术背景的学生开放资源；改革奖学金评定体系，降低单一导师评价权重，引入团队匿名评审与成果盲审制度，并设立“跨学科创新”“成果转化”等专项奖励，拓宽优秀博士生评价的标准。此外，建立导师指导质量监测系统，特别关注导师对不同背景学生的指导投入与支持差异。

参考文献：

[1] 许丹东，翟月，陈洪捷. 博士毕业生区域流动就业的特征与影响因素：基于东、中、西部地区的分析[J]. 学位与研究生教育，2023(12)：18-26.

[2] 郭丛斌，方晨晨，何菲. 博士研究生教育对经济增长的影响：基于博士学位授权点区域差异的视角[J]. 教育研究，2022，43(5)：124-138.

[3] 许丹东，沈文钦，陈洪捷. 西部地区博士毕业生去哪儿就业：博士毕业生的区域流动及其影响因素[J]. 中国高教研究，2021(8)：34-40.

[4] 罗洪川，高玉建，向体燕，等. 西部地区学术型博士毕业生去向及就业因素研究：基于 2015-2020 年博士研究生毕业数据的实证分析[J]. 研究生教育研究，2023(1)：72-81，97.

[5] 樊秀娣，沈晶晶. 从单一到多元：博士生职业发展多样化的现实需求及教育支持策略研究[J]. 高校教育管理，2024(1)：81-91，112.

[6] 胡德鑫，师悦. 中国研究型大学博士职业选择趋向与治理路径研究[J]. 研究生教育研究，2020(6)：53-61.

[7] 罗洪川，向体燕，高玉建，等. 我国博士毕业生去向及就业特征分析：基于 2015-2020 年博士毕业生数据的分析[J]. 学位与研究生教育，2022(1)：53-62.

[8] 陈庆，杨颉，石雪怡. 博士生就业路径多元化的推进举措探析：基于部分 OECD 国家实践的分析[J]. 研究生教育研究，2025(3)：92-100.

[9] 黄维海，李树岳. “理想奔现”与“现实优化”：博士生就业多元化源流分析[J]. 中国高教研究，2024(12)：70-76.

[10] 陈纯槿，郅庭瑾. 博士毕业生就业能力的影响机理及发展路径[J]. 研究生教育研究，2024(4)：98-105.

[11] 卿石松. 博士生就业问题调查及对策分析[J]. 学位与研究生教育，2017(1)：43-49.

[12] 卿石松. 人文社科博士毕业生就业能力供需匹配分析[J]. 华东师范大学学报(教育科学版)，2020(12)：85-98.
[13] Roy D，Jiménez López M D，García Álvarez M E. Hires-PhD：A Transversal Skills Framework for Diversifying PhD Employability[J]. Humanities & Social Sciences Communications，2025，12(1)：1-16.
[14] 王佳雨，沈文钦，翟月. 博士生就业信息获取及其对就业意向的影响：基于认知信息处理理论的质性分析[J]. 研究生教育研究，2025(5)：60-68.
[15] 陈纯槿. 海外留学如何影响博士毕业生的学术职业选择？——来自全球博士生项目中国调查的证据[J]. 上海交通大学学报(哲学社会科学版)，2024(5)：79-91.
[16] Blei D M，Ng A Y，Jordan M I. Latent Dirichlet Allocation[J]. Journal of Machine Learning Research，2003，3：993-1022.
[17] Wei X，Croft W B. LDA-based Document Models for Ad-hoc Retrieval[C]. In：Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval，2006：178-185.
[18] 周全，汤书昆. 博士生视角下和谐导学关系构建关键要素探析：基于学生微博自我表露数据的研究[J]. 研究生教育研究，2016(4)：33-40.
[19] 乔慧慧. 规训与抵抗：博士生学术认同的形成机制研究[J/OL]. 重庆高教研究：1-15[2025-12-23]. http://kns. cnki. net/kcms/detail/50. 1028. G4. 20251223. 1529. 002. html.
[20] 李彩玉，曲铁华. 有谁能懂我的痛苦：社交媒体中博士生负性情绪的叙事研究[J]. 重庆高教研究，2025(4)：70-80.
[21] 梁超，于鑫凡，张荣杰. 家庭借贷约束、人力资本形成与长期发展：基于大学生生源地助学贷款的研究[J]. 管理世界，2025(8)：109-134.
[22] 崔晓娟，朱绘杰. 人力资本、社会资本对大学生基层就业意愿的影响研究：基于西部 57 所高校的调查分析[J]. 高教探索，2025(2)：119-128.
[23] Lin Nan. Building a Network Theory of Social Capital[J]. Connections，1999，22(1)：28-51.
[24] 王霆. 大学生高质量就业的影响机制研究：人力资本与社会资本的视角[J]. 高教探索，2020(2)：108-114.

Intelligent System and Practical Approach for Evaluating Employment Quality of Doctoral Graduates from the Multimodal Perspective

WANG Ting，LIN Shenqi

(*Business School，China University of Political Science and Law，Beijing* 100088，*China*)

Abstract：Against the backdrop of advancing high-level technological self-reliance and industrial upgrading，doctoral graduates represent a vital component of the national strategic talent reserve. Their employment quality directly influences the synergistic performance of the education chain，talent chain，and innovation chain. This study addresses the critical triple challenges in doctoral employment：regional imbalance，occupational misalignment，and skills mismatch. It explores the value implication and implementation logic by applying multimodal data fusion technology to construct a full-cycle intelligent evaluation system encompassing job hunting，initial employment，and career development. By integrating the Latent Dirichlet Allocation (LDA) topic model，the Delphi method，and the Analytic Hierarchy Process (AHP)，this study systematically extracts key evaluation indicators for doctoral employment quality. It synthesizes structured data，textual data，and behavioral data，and employs multimodal large model algorithms to generate applicable solutions from intelligent evaluation result such as dynamic value-based guidance，individual's competency profiling，and coordinated policy governance. This study also proposes practical pathways including constructing a hierarchical three-tier collaborative multimodal data warehouse，developing a two-dimensional intelligent diagnostic tool，and implementing phased，targeted intervention strategies. These approaches are designed to drive the transition of doctoral graduates from a "singular academic track" to "diverse，high-quality career pathways".

Keywords：doctoral graduates；employment quality；multimodal data；intelligent evaluation；talent cultivation