

面向数字产业需求的大数据岗位胜任力模型构建研究

孙 洁^{1*}, 刘 馨², 闫欣宜¹

¹北京联合大学商务学院, 北京

²北京联合大学应用文理学院, 北京

收稿日期: 2026年9月3日; 录用日期: 2026年10月16日; 发布日期: 2026年10月27日

摘 要

数字经济高质量发展对大数据人才提出结构性新要求。本文结合文本分析方法, 对数据分析师、数据产品经理等企业核心岗位展开系统调研, 基于岗位描述的词频分析与语义归类, 构建了包含4个一级指标及12个二级指标的岗位胜任力模型。研究发现: 数字产业对人才的需求正从“学科匹配”转向“能力匹配”, 其中“数据要素化”能力成为区分传统数据人才与新型数据人才的关键标识。基于此, 本文提出大数据管理与应用专业在培养目标精准化、实践教学体系重构、产教融合机制创新等方面的优化路径, 为应用型高校的大数据人才培养提供决策参考。

关键词

数字产业, 岗位胜任力, 大数据, 人才培养

Research on Constructing a Competency Model for Big Data Positions to Meet the Needs of the Digital Industry

Jie Sun^{1*}, Xin Liu², Xinyi Yan¹

¹School of Business, Beijing Union University, Beijing

²College of Applied Arts and Sciences, Beijing Union University, Beijing

Received: September 3, 2026; accepted: October 16, 2026; published: October 27, 2026

Abstract

The high-quality development of the digital economy has created new, structural demands for big

*通讯作者。

文章引用: 孙洁, 刘馨, 闫欣宜. 面向数字产业需求的大数据岗位胜任力模型构建研究[J]. 创新教育研究, 2026, 14(10): 589-595. DOI: 10.12677/ces.2026.1410824

data talent. Employing text analysis methods, this study conducts a systematic survey of core corporate roles—such as data analysts and data product managers—and constructs a competency model comprising four primary indicators and twelve secondary indicators based on word frequency analysis and semantic categorization of job descriptions. The findings reveal that the digital industry's demand for talent is shifting from “disciplinary alignment” to “competency alignment”, with the ability to handle “data as a factor of production” serving as the key differentiator between traditional and new-generation data professionals. Based on these insights, the paper proposes optimization pathways for “Big Data Management and Application” majors—focusing on refining training objectives, restructuring practical teaching systems, and innovating industry-education integration mechanisms—thereby providing a decision-making reference for talent cultivation in applied universities.

Keywords

Digital Industry, Job Competency, Big Data, Talent Cultivation

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

数字经济已成为我国经济发展的重要引擎，其核心产业的基础性、支柱性作用日益显著。在此背景下，北京市充分发挥数字经济先行优势，产业规模与增速位居全国前列，已成为驱动高质量发展的重要力量。为强化数字人才对数字经济发展的支撑，北京市多部门联合印发《北京市加快数字人才培育支撑数字经济发展实施方案（2024-2026 年）》提出：“探索数字领域新工科研究与实践，加强高等院校数字领域相关学科专业建设，加大交叉学科人才培养力度，建设一批数字领域‘一流专业’‘一流课程’，提升首都急需紧缺领域卓越工程师自主培养能力。”¹然而，数字经济快速发展也对数字人才供给提出了更高要求。中国信息通信研究院发布的《数字经济就业影响研究报告》显示，2020 年我国数字化人才缺口接近 1100 万²，企业的跨领域复合型人才需求快速上升。作为 2018 年首批开设的面向数据人才培养的专业，大数据管理与应用专业肩负着为数字产业输送复合型人才的重要使命，但现有人才培养模式在对接数字产业特色需求方面仍显不足，对数据要素化能力培养以及真实产业场景嵌入的关注亟待加强[1] [2]。

因此，本研究以数字产业为调研对象，选取北京市数字企业作为主要样本来源，聚焦数据分析师、数据产品经理等核心岗位，通过文献研究、半结构化访谈与文本挖掘，系统分析岗位能力要素，构建岗位胜任力模型，并提出人才培养优化建议。

2. 文献综述

随着数字经济发展，政策以数字产业化和产业数字化为核心，着力培养战略科学家、领军人才、技术人才和技能人才等四类人才队伍。在宏观政策指引下，学界对数字人才需求的研究逐渐从宏观政策解读转向微观岗位能力挖掘。任心雨等(2025)聚焦长三角旅游业数字人才需求，提出将其能力模型划分为运营型、市场型和技术型三类[3]。叶仁道等(2024)运用 LDA 主题模型和 TF-IDF 方法分析新兴产业招聘数据，发现大数据岗位更重视数据分析、SQL 查询和 Python 编程等技能组合[4]。潘维超等(2025)基于 K-

¹https://www.beijing.gov.cn/zhengce/zhengcefagui/202407/t20240712_3743944.html

²https://www.ndrc.gov.cn/fggz/jyysr/jysrsbxf/202112/t20211216_1308107.html

means 聚类将数字人才需求划分为三类,构建了以数字知识技能为核心的胜任力框架[5]。刘彩娜等(2025)进一步指出,数据类岗位对机器学习等专业技能和软件技能有着多维度的刚性要求[6]。“数据要素化”是指将原始数据资源转化为具有明确经济价值的生产要素的过程,已有研究为理解数字人才的市场需求提供了量化依据,但其分析对数据要素市场化改革背景下“数据要素化”能力的关注仍显不足。

对岗位能力的系统刻画离不开胜任力模型的理论支撑。McClelland(1973)最早提出胜任力概念,将其界定为能够区分特定岗位绩效水平的个体特征[7]。此后发展的冰山模型与洋葱模型将胜任力划分为显性素质与隐性素质。本研究所构建的四维模型,正是对冰山模型在数字产业情境下的具体化:技术基础能力与管理决策能力偏向显性,可通过系统教学训练获得;而偏隐性部分的场景应用能力与数据要素化能力需在真实产业浸润中养成。然而,既有胜任力模型多从通用管理情境中抽象而来,对数据要素市场化这一独特制度情境下新型能力结构的解释力有限。

文本分析与挖掘方法能够从海量非结构化文本中自动提取潜在主题与知识结构,被广泛应用于政策评估、职业需求识别与用户行为画像等领域。叶提芳等(2026)运用 BERTopic 模型对 878 份数据要素政策进行主题挖掘,识别出场景应用、数据治理与市场培育三大核心领域[8];聂奕轩等(2026)采用 LDA 主题模型与文本挖掘揭示了低年资急诊护士心理调适需求的六维结构[9];邹锬等(2026)立足产业化视角设计了大模型文本挖掘系统[10];王德庄等(2024)则基于招聘网站文本分析,构建了图书情报硕士数据类岗位胜任力模型[11]。现有研究针对大数据管理与应用专业人才胜任力模型的系统性文本挖掘研究较为缺乏,对数据要素市场化配置改革背景下新型数字人才能力结构的探讨尤为稀缺。

作为新兴交叉学科,大数据管理与应用专业的培养模式正处于快速探索期。邓胜利等(2024)的调查报告显示,该专业学位教育在课程体系与产业需求的对接上仍存在明显差距[1]。余海燕等(2023)、刘彩娜等(2025)分别从伦理治理与模块化课程角度探索了培养模式优化路径[2][6]。童杰(2024)采用 K-means 算法建立了大数据岗位工程实践能力评价模型,为能力量化评估提供了方法参考[12]。孙洁等(2025)基于岗位需求视角,指出课程内容与产业真实场景的脱节是制约人才培养质量的关键瓶颈[13]。叶光辉等(2023)分析了高校相关专业的联合建设路径[14]。但总体而言,高校人才培养方案与产业真实场景之间的适配短板依然突出。

综上所述,现有文献为本研究提供了理论及方法基础,但仍存在以下局限:其一,大数据岗位需求分析多聚焦通用技术技能,对“数据要素化”能力的识别不足;其二,胜任力模型研究针对大数据管理与应用专业的系统性研究相对匮乏;其三,产教融合研究对数字产业特色场景的回应不够深入。鉴于此,本文以北京数字产业相关企业作为调研样本,聚焦核心岗位,通过文献研究、实地访谈与文本挖掘方法,系统构建岗位胜任力模型。

3. 岗位胜任力模型构建

(1) 产业需求与岗位特征

北京市在金融科技、智慧城市、数字医疗等领域具有显著优势。北京国际大数据交易所作为全国首个基于数据要素登记确权的数据交易平台,其设立为数据要素市场化配置提供了制度与空间载体,推动数据要素在金融服务、城市治理等场景的深度应用,催生了对“懂合规、懂资产、懂场景”的复合型数据人才的巨大需求。

从北京市数字产业相关企业的招聘数据来看,大数据相关岗位类型多样化与能力要求呈现多元化特征。数据分析师是需求量最大的岗位,不仅要掌握数据可视化等技术工具,更需要具备业务理解能力。数据产品经理贯通数据全生命周期,协调技术、业务、法务等多方资源。数据运营类岗位侧重用户增长、私域运营等业务场景。数据开发与治理类岗位聚焦数据仓库建设与合规管理。量化分析类岗位在金融科

技领域需求突出,涉及量化模型构建与策略回测。上述岗位类型共同构成了北京数字产业大数据人才的需求图谱,也为进一步识别大数据岗位的共性能力要求提供了基础。因此,本文将研究重点聚焦于不同岗位类型背后的共性能力要求,归纳形成大数据岗位的共性胜任力模型。

调研发现,当前北京数字产业人才供需存在结构性矛盾:一是层次错配,具备“数据要素化”思维的复合型中层人才严重短缺;二是能力错配,高校培养侧重技术工具操作,对真实产业场景嵌入不足,企业反映应届生实战经验欠缺;三是制度错配,校企协同产教融合尚未形成深度协同机制。模型构建正是基于上述现实需求。

(2) 研究设计

本研究采用“政策-需求-实践”三角驱动框架,综合运用文献研究与招聘大数据文本挖掘方法。在文献梳理基础上,确定以数据分析师、数据产品经理为核心岗位,以数据治理工程师、量化分析师等为补充岗位,采集北京地区大数据相关岗位的描述文本。

数据采集覆盖市面主流招聘网站,采集原始数据 300 余条,经清洗剔除重复发布、信息不完整等低质量样本后,获得有效样本 188 条。清洗后的数据完整包含“岗位职责”与“任职要求”两个核心字段,涵盖互联网、电商、金融科技、政务等多个产业场景。

岗位类型分布方面,数据分析师类岗位占比最高,其次是数据运营、数据产品经理以及数据治理工程师、量化分析师等岗位,样本覆盖快手、字节跳动、科大讯飞等头部互联网与科技企业,以及中国联通等国有与政务数据服务企业,企业规模从 50 人以下初创企业到万人以上大型集团均有涉及。岗位类型与企业的多元分布,为胜任力模型构建提供了较为全面的数据基础。

在文本分析方法上,首先对岗位职责与任职要求分别构建语料库,运用 Python 的 jieba 分词库进行中文分词,以哈工大停用词表过滤无实际意义词汇,并加载自定义词典以提高分词准确性。自定义词典主要收录了“数据资产”“数据合规”“数据挖掘”等专业术语,以确保分词准确性。随后采用词频统计方法提取高频关键词,分别统计岗位职责与任职要求两个子语料库中的词频分布,以发现招聘文本中反复出现的能力关键词。高频关键词的选取阈值为频次 ≥ 28 次,取词频分布前 15% 分位数作为筛选阈值,低于该阈值的关键词因语义支撑不足不予纳入指标体系。关键词的语义归类由两位行业专家独立完成,依据关键词的共现语境与能力属性进行标注。编码完成后,计算编码者间信度 Kappa 系数为 0.84,达到良好一致水平(>0.80),对于分歧条目经第三位专家仲裁后达成共识,确保了归类结果的客观性。

(3) 基于文本挖掘的胜任力模型构建

对有效文本进行词频统计后,本研究结合关键词语义指向、岗位任务及能力属性,对高频关键词进行归类与整合。

基于关键词属性划分一级维度,将高频关键词按其指涉对象与功能属性,将其划分为四个一级维度,分别命名为:技术基础能力(A)、管理决策能力(B)、场景应用能力(C)、数据要素化能力(D)。其中,归类遵循以下原则:编程语言、数据处理工具及算法等方法类关键词,描述岗位所需的技术操作能力,归入技术能力维度;业务理解、项目管理、团队协作等关键词体现岗位在业务支持、组织协同的要求,归入管理决策维度;电商、金融、智慧城市等关键词体现大数据技术在不同产业场景中的应用需求,归入场景应用维度;数据资产、数据合规、数据安全及数据产品类关键词,描述岗位在数据要素化方向上的专业能力,归入数据要素化维度。

基于能力层次与岗位职能特征对二级指标进一步划分,并结合高频关键词及其语义关联,对各级维度内部能力要求进行细分。技术基础能力维度按“工具应用-方法分析-算法建模”的递进逻辑划分为三个层次:数据处理与工具应用(A1)对应基础工具应用层,侧重大数据岗位人员利用 SQL、Python、Spark 等技术工具开展数据分析工作的基础能力。统计分析 with 数据建模(A2)对应方法分析层,侧重运用统计分

析、数据分析及建模方法的理解与应用。数据挖掘与算法应用(A3)对应算法建模层,侧重模型构建、算法应用以及从数据中发现潜在规律的能力。

管理决策能力维度按职能逻辑划分为三个层次:商业分析与决策支持(B1)对应业务层,侧重业务需求理解与策略制定能力;项目协作与沟通管理(B2)对应组织层,侧重项目推进与团队协作能力;数据质量与治理管理(B3)对应治理层,侧重数据标准制定、数据质量和数据治理流程等方面的管理能力。

场景应用能力维度按产业场景类型划分为三个方向,金融科技应用(C1)对应金融行业的数据分析与智能决策场景;智慧城市应用(C2)对应政务、城市治理及公共数据应用场景;数字商业运营场景应用(C3)对应电商、用户增长和数字化运营等商业场景。

数据要素化能力在本研究中特指大数据岗位从业者所具备的贯穿数据要素价值实现的综合能力,即实现数据从资源向资产转化的系统性能力。按照“资产识别-安全合规-产品转化”的逻辑划分为三个层次:数据资产识别与评估(D1)侧重对数据资源识别、质量判断和价值评估的能力;数据合规与安全管理(D2)侧重数据开发利用过程中防范安全与合规风险能力;数据产品设计与运营(D3)侧重数据产品的设计、应用与持续运营能力。

由此形成包含 4 个一级指标、12 个二级指标的大数据岗位胜任力模型指标体系,结果如表 1 所示。

Table 1. Indicator system for the competency model of big data positions based on text mining

表 1. 基于文本挖掘的大数据岗位胜任力模型指标体系

一级指标	二级指标	关键词(频次)		
A. 技术基础能力	A1 数据处理与工具应用	SQL (107)	Python (82)	Spark (50)
	A2 统计分析与数据建模	统计学(79)	数据分析(43)	建模(30)
	A3 数据挖掘与算法应用	模型(157)	算法(64)	数据挖掘(33)
B. 管理决策能力	B1 商业分析与决策支持	业务(449)	策略(82)	决策(59)
	B2 项目协同与沟通管理	团队(211)	项目(165)	沟通(128)
	B3 数据质量与治理管理	标准(79)	数据治理(41)	质量管控(36)
C. 场景应用能力	C1 金融科技应用	量化(157)	风控(82)	金融(46)
	C2 智慧城市应用	安全(50)	治理(41)	智慧(28)
	C3 数字商业运营应用	运营(150)	用户(126)	电商(31)
D. 数据要素化能力	D1 数据资产识别与评估	质量(49)	数据资产(47)	价值(41)
	D2 数据合规与安全管理	数据安全(44)	数据合规(36)	保护(34)
	D3 数据产品设计与运营	产品(260)	隐私(52)	计算(30)

表中数据显示,四个能力维度在词频统计中均获得支撑。技术基础能力维度中,SQL、Python、模型等关键词呈高频分布,表明数据处理工具的操作能力与算法建模能力是大数据岗位普遍的技术要求,其中“模型”为全部技术类关键词中频次最高者,反映出建模能力在岗位需求中的关注度较高。

管理决策能力维度中,“业务”一词为全部关键词中频次最高者,表明业务理解与业务需求相关要求在招聘文本中具有较高的关注度。团队、项目、沟通等关键词的高频出现,体现出大数据相关工作需要与业务部门、技术团队及其他职能部门开展协同,岗位人员除了技术能力外,还需兼备技术与管理的复合能力。

场景应用能力维度中,运营与用户类关键词频次最高,与互联网平台企业对数据分析岗位的大规模招聘需求直接相关;涉及金融、电商、智慧城市等垂直场景的关键词,验证了场景应用能力维度的多产

业覆盖特征。

在数据要素化能力维度中,产品、隐私、数据安全等关键词均达到可观测频次,表明数据价值开发及规范利用等相关能力已经进入大数据岗位的能力要求。“产品”作为出现频次较高的关键词,与数据产品设计、开发及运营等岗位任务有较强关联,反映出数据产品设计与运营能力已成为数据要素化方向的显性需求。

(4) 模型特征分析

“数据要素化能力”在词频统计中呈现突出地位,该维度关键词在文本中已达到可观测频次,表明在数据要素市场化配置改革持续推进的背景下,企业对数据要素化能力的关注已逐步在招聘需求中体现。未来的数据人才必须理解数据不仅是分析对象,更是生产要素和资产,要具备将数据转化为价值的能力。

“场景应用能力”下设三个二级指标,词频分布呈现梯度特征:运营、用户等互联网场景词频最高,金融、电商等场景居中,智慧城市等垂直场景相对偏低,反映了北京数字产业多领域协同发展的现实格局。脱离产业场景的数据分析难以产生实际价值,高校人才培养必须强化场景嵌入,使学生能够运用数据思维解决复杂产业问题。

四维能力结构呈现“冰山模型”特征。词频结果显示,技术基础能力与管理决策能力维度的关键词分布集中且频次高,属于“冰山上”的显性素质,可通过课程教学与技能训练获得;场景应用能力与数据要素化能力维度的关键词分布较为分散且频次相对较低,属于“冰山下”的隐性素质,需要在真实项目实战中通过长期浸润与反思内化形成。这提示在人才培养中,既要重视显性技能的标准化训练,更要通过产教融合机制为隐性素质的养成提供真实情境支撑。

招聘平台岗位数据验证显示:技术基础能力对应的 Python、SQL 等技能关键词覆盖度最高;场景应用能力对应的行业知识要求在垂直领域岗位中出现频率较高;数据要素化能力相关的治理、合规、资产等关键词虽当前出现频率相对较低,但呈快速增长趋势,侧面印证了该维度属于面向未来的核心能力。

4. 结论与建议

(1) 研究结论

首先,数字产业对数据人才的需求正从“技术工具操作型”向“数据要素化复合型”转变,“数据资产识别”“合规流通”等数据要素化能力已成为区分新型数字人才与传统数据人才的核心标识。

其次,当前高校人才培养在培养目标、实践内容、协同机制等方面与产业实际需求存在结构性偏差,亟须通过深化产教融合改革弥合这一鸿沟。

最后,本研究构建了包含技术基础能力、管理决策能力、场景应用能力、数据要素化能力四个维度的岗位胜任力模型,为大数据管理与应用专业培养方案优化提供了可操作的框架。

(2) 对策建议

一是精准化设置培养目标。将区域产业战略需求转化为专业培养目标与学习成果标准,将“数据要素化”思维与实践能力作为核心毕业要求,设置能力达成指标与评价量规。

二是重构实践课程体系。构建“基础核心层 + 场景嵌入层 + 前沿交叉层”三层模块化实践教学课程体系。其中,场景嵌入层应依托北京本地标杆场景,开发脱敏实战案例库。

三是创新产教融合机制创新。将企业真实业务项目改造为课堂实训课题,为学生配备“学术 + 产业”双导师。建立制度性通道,破解脱敏数据使用、知识产权归属、指导教师工作量认定等关键制度障碍。

四是改革教学评价方式。构建基于学习分析技术的过程性评价体系,制定学生数字素养能力评价标准,明确“数据要素化”思维与实践能力的观测指标与评价量规,建立学生自评、同伴互评、导师评价、企业反馈的多主体参与机制。

(3) 研究局限与展望

本研究存在以下局限：一是文本挖掘数据主要来自北京地区主流招聘平台，对国内其他城市及中小企业覆盖面不足，可能影响结论的普适性。二是模型构建基于词频统计，尚未结合企业实际绩效数据进行实证检验，存在一定主观偏差。三是样本集中于金融科技、智慧城市等优势领域，对数字文创、工业互联网等新兴领域的覆盖有待拓展。

未来将扩大招聘平台与岗位类型覆盖范围，开展基于本模型的毕业生追踪调查以检验模型效度，并基于动态招聘数据实现胜任力模型的持续迭代优化。

基金项目

北京联合大学教育教学研究与改革重点项目“面向北京数字产业需求的大数据管理与应用专业产教融合实践教学体系重构”(JJ2026Z002)。

参考文献

- [1] 邓胜利, 汪璠, 夏苏迪. 大数据管理与应用本科专业学位教育发展状况调查报告(2023) [J]. 图书情报知识, 2024, 41(1): 69-79+101.
- [2] 余海燕, 刘进, 卢安文, 等. 大数据管理与应用专业能力与伦理治理协同培养模式探析[J]. 科学咨询(科技·管理), 2023(1): 4-9.
- [3] 任心雨, 龚晨愉, 方宇萱, 等. 数字化驱动旅游业转型: 长三角地区数字经济人才需求与能力模型研究[J]. 可持续发展, 2025, 15(5): 61-70.
- [4] 叶仁道, 李文倩, 卢毅. 新工科下大数据人才岗位需求与任职能力分析——以新兴产业为例[J]. 杭州电子科技大学学报(社会科学版), 2024, 20(2): 60-71.
- [5] 潘维超, 郇文庆, 刘兴波, 等. 基于招聘大数据与机器学习的数字人才需求与能力模型构建[J]. 软件导刊, 2025, 24(4): 173-183.
- [6] 刘彩娜, 路健, 范增民. 大数据管理与应用专业模块化课程体系建设与实践[J]. 无线互联科技, 2025, 22(10): 119-123.
- [7] McClelland, D.C. (1973) Testing for Competence Rather than for "Intelligence". *American Psychologist*, **28**, 1-14.
- [8] 叶提芳, 张志刚, 张耀峰, 等. 数据要素政策与企业数字化转型: 基于文本挖掘与效应评估的实证研究[J/OL]. 数理统计与管理, 1-19. <https://doi.org/10.13860/j.cnki.sltj.20260721-001>, 2026-09-02.
- [9] 聂奕轩, 葛冲, 谢砚辞, 等. 基于社交媒体文本挖掘低年资急诊护士心理调适需求[J]. 护理学杂志, 2026, 41(15): 89-94.
- [10] 邹锬, 余雪, 刘艳, 等. 产业化视角下大模型文本数据挖掘系统研究[J]. 高科技与产业化, 2026, 32(7): 42-44.
- [11] 王德庄, 林萌萌. 图书情报硕士数据类岗位胜任力模型构建研究——基于企业招聘网站的数据分析[J]. 情报探索, 2024(6): 1-9.
- [12] 童杰. 一种基于大数据岗位的工程实践能力评价模型[J]. 深圳信息职业技术学院学报, 2024, 22(1): 15-23.
- [13] 孙洁, 刘静, 常馨予. 基于岗位需求的专业实践教学体系构建研究[J]. 教育进展, 2025, 15(6): 62-67.
- [14] 叶光辉, 曹高辉, 夏立新. 我国高校信息管理学院系大数据管理与应用专业联合建设路径分析[J]. 情报学报, 2023, 42(2): 231-240.