

# 服务市场需求的大数据人才培养模式研究

王婉婧, 尹心萍\*, 翟羽佳

天津师范大学管理学院, 天津

收稿日期: 2023年5月25日; 录用日期: 2023年6月15日; 发布日期: 2023年6月27日

## 摘要

研究立足于当下高校大数据人才培养与市场人才需求现状, 明确人才培养的缺失与偏差, 为加快大数据人才供需两侧的有效对接提供支撑。采用社会网络分析与实体抽取方法, 对大数据专业培养方案和大数据相关岗位招聘公告中体现出的人才能力特征进行分析与匹配, 发现大数据人才培养与市场需求之间存在一定偏差, 提出聚焦国家战略需要, 以市场需求为导向, 以学科特质为内核, 以技术工具为抓手, 高校、企业、国家协同培育的人才培养模式。

## 关键词

人才培养, 市场需求, 大数据

# Research on the Big Data Talent Training Model for Serving Market Demand

Wanjing Wang, Xinping Yin\*, Yujia Zhai

School of Management, Tianjin Normal University, Tianjin

Received: May 25<sup>th</sup>, 2023; accepted: Jun. 15<sup>th</sup>, 2023; published: Jun. 27<sup>th</sup>, 2023

## Abstract

Based on the current situation of big data talent cultivation and market demand in universities, this study clarifies the deficiencies and deviations in talent cultivation, providing support for accelerating the effective connection between the supply and demand of big data talents. By using social network analysis and entity extraction methods, the talent capability characteristics reflected in the big data professional training plan and big data related job recruitment announcements were analyzed and matched, and found that there is a certain deviation between big data talent cultivation and market demand. This paper proposes a talent cultivation model focusing on national strategic needs, oriented by market demand, with disciplinary traits as the core and

\*通讯作者。

technological tools as the grip, and cultivated by universities, enterprises and the state in collaboration.

## Keywords

Talent Cultivation, Market Demand, Big Data

Copyright © 2023 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

大数据作为新一轮科技革命和产业变革的重要驱动力量,已成为一种新的战略资源,是重塑国家竞争优势的新机遇。近年来,我国大数据产业规模高速增长,呈现良好发展态势[1]。但大数据人才队伍建设进程依然滞后于大数据技术与产业发展,人才教育培养目标 and 人才市场需求之间形成了一定的偏差[2]。以市场需求为导向探究大数据人才培养模式,对于有效提升大数据人才供需双方的能力融合程度,加速大数据高质量人才培养有着重大意义。

近年来,关于大数据人才培养的研究主要集中在教育学界与信息资源管理学界。其中教育学界主要以产教融合、技术赋能、协同创新等时代关键词为主题,在新工科及新商科背景下,对大数据人才培养机制与路径进行深度的理论研究。信息资源管理学界主要以不同研究主体为离群点,从跨学科视角对大数据人才培养策略提出调整建议,如陈沫等[3]对国内外大数据相关专业的培养目标和课程设置模式进行调研,总结学科教育发展特色,设计情报学取向的大数据专业人才培养计划;王晰巍等[4]以国内外18所典型高校为样本,对大数据管理与应用专业的理论课程体系和实践教学环节进行对比分析,提出产学研协同为导向的新文科大数据管理与应用专业本科人才培养模式。

大数据人才市场需求方面的相关研究主要集中在两个方面。一是选取不同招聘来源中的招聘数据来分析挖掘数据人才需求侧对于人才多层面的要求与关注重点,以反映当前大数据人才社会需求现状与趋势,国外如 Fang 等[5]通过收集网站招聘信息对数据分析师的技能需求进行分析,国内如袁毅等[6]通过招聘综合网站进行招聘文本挖掘,桑秀丽等[7]通过企业官网和事业单位的招聘数据探讨培养目标与市场需求的差距。二是侧重于在研究方法上突破新的技术方法将冗杂的招聘信息转化为人才需求特征。如 Cui X 等[8]引入了 BiLSTM 和 Attention 机制来提高实体识别的效果,提出了用来提取互联网招聘信息岗位描述中的技能实体的 BiLSTM-Attention-CRF 模型;易新河等[9]利用插入在 Bi-LSTM 层和 CRF 层之间的域自适应层(CDTL-PSE)来实现从各个源域到目标域的跨域迁移学习,有效提高了专业技能词的识别能力。

通过对高校人才培养及市场人才需求研究成果的梳理,发现在大数据人才培养路径完善的探究中,鲜少有研究将其与市场需求的融合破困作为突破点。因此,本研究立足于当下高校大数据人才培养与市场人才需求现状,明确大数据人才培养的缺失与失衡,探索以市场需求为导向的人才培养模式。

## 2. 数据获取与处理

本研究选取“大数据管理与应用”专业与“数据科学与大数据技术”专业的高校培养方案作为人才供给侧数据来源,共采集培养方案262份。筛选在重要内容上有详细阐述的培养方案作为文本数据,最终总计192份。选取智联招聘网作为人才需求侧的数据来源,在智联招聘官网以“大数据”作为关键词,限定条件“学历要求”为“本科”,“地区”限定为“北京/上海/广州”,包括“岗位名称”“招聘要求”

等主要数据字段，共获取数据 1748 条，经去重和文本质量筛选后得到有效招聘数据 1569 条。

本研究主要探究大数据人才在专业知识与技能方面的能力，包括理论知识、经验知识、专业技能等。首先，对培养方案与招聘公告中的专业知识与技能实体进行人工提取，同时构建能力同义词词典，将文本中出现的内容不一致但表达意义相同的实体匹配替换。最终共明确培养方案能力实体 164 个，招聘公告能力实体 215 个。

### 3. 研究结果分析

本研究采用社会网络分析法，分别绘制人才培养与市场需求的能力度中心性共现网络(见图 1、图 2)，网络中的圆形节点表示知识技能培养实体，节点的大小表示实体出现的频次高低，节点颜色的深浅表示实体中心性的程度，网络中的边表示两个实体之间存在共现关系。

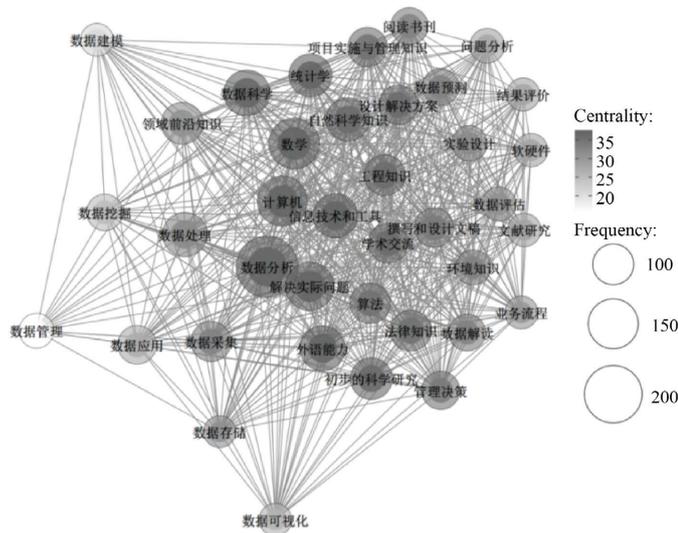


Figure 1. Talent cultivation capability degree centrality coexistence network  
图 1. 人才培养能力度中心性共现网络

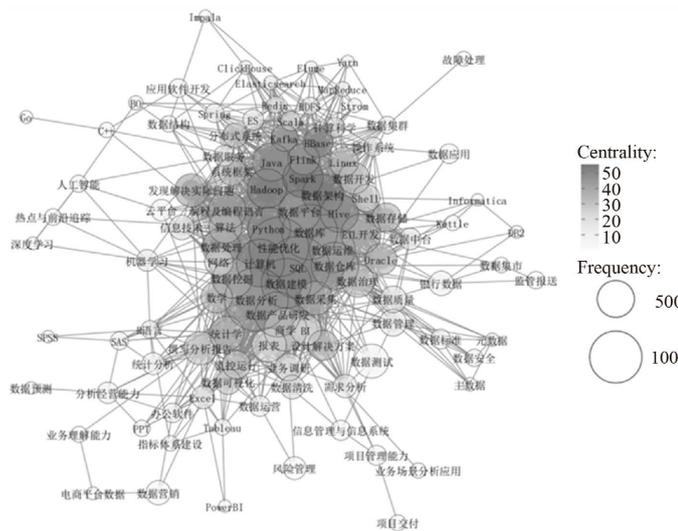


Figure 2. Market demand capability degree centrality coexistence network  
图 2. 市场需求能力度中心性共现网络

计算大数据人才培养与市场需求能力实体的点度中心度，并降序排列，如表 1 所示。

**Table 1.** Talent cultivation capability degree centrality and market demand capability degree centrality

**表 1.** 人才培养能力点度中心度与市场需求能力点度中心度

序号	人才培养知识技能实体	度	序号	市场需求知识技能实体	度
1	数据分析	874	1	SQL	2800
2	计算机	577	2	性能优化	1854
3	数学	561	3	数据库	1737
4	解决实际问题	553	4	Spark	1708
5	外语能力	472	5	数据分析	1552
6	信息技术和工具	462	6	Hive	1493
7	数据处理	454	7	Python	1424
8	数据科学	438	8	数据建模	1254
9	自然科学知识	432	9	Hadoop	1160
10	统计学	421	10	数据仓库	1101
11	初步的科学研究	394	11	HBase	959
12	法律知识	391	12	Flink	920
13	算法	386	13	商学	919
14	工程知识	374	14	编程及编程语言	908
15	设计解决方案	372	15	Java	891
16	数据采集	370	16	数据运维	848
17	领域前沿知识	367	17	数据架构	843
18	撰写和设计文稿	363	18	ETL 开发	830
19	学术交流	358	19	计算机	780
20	数据应用	342	20	Oracle	771

从图 1 和表 1 可以发现，当前高校将计算机、数学、统计学等基础课程作为培养根基，既培养大数据人才“信息技术和工具”“数据处理”等大数据相关专业最核心的专业理论知识与学科技能，也包含“解决实际问题”的实践技能以及“工程知识”等通识知识。从度的数值梯度来看，数据分析能力是大数据人才的培养核心。从图 2 及表 1 可以发现，大数据人才的能力需求以主流数据库及相关工具、大数据平台及生态组件、编程及编程语言为主，其中 SQL、Java、Python、Spark 为核心技术工具，数据分析、性能优化、计算机为核心能力。最需要掌握的知识包括“性能优化”“数据分析”“数据建模”“数据仓库”等大数据专业知识以及“商学”“计算机”等学科背景知识。

以度作为核心匹配标准，本研究将供需双方的能力实体进行比对(见表 2)。从需求侧映射供给侧可以发现，仅有“数据分析”完整的出现在人才培养实体列表中。考虑到文本差异性，进一步将“SQL”“Spark”“Hive”等需求实体抽象后，可与培养实体中的“信息技术和工具”进行匹配。将供给侧反向映射需求侧，可以看出各高校将人才培养重心放在基础学科知识及专业理论知识上，而就业市场已经要求人才在任职前应具备运用所学大数据技术工具解决实际工作问题的能力。而导致人才供需匹配度较低的深层次原因的是高校培养与市场需求所处的人才环节不同，而不同环节对应的目标不同，高校以培养出可为国

**Table 2.** Talent cultivation and market demand capability entity contrast  
**表 2.** 人才培养与市场需求能力实体比对

序号	人才培养知识技能实体	度	序号	市场需求知识技能实体	度
1	数据分析	874	1	SQL	2800
2	计算机	577	2	性能优化	1854
3	数学	561	3	数据库	1737
4	解决实际问题	553	4	Spark	1708
5	外语能力	472	5	数据分析	1552
6	信息技术和工具	462	6	Hive	1493
7	数据处理	454	7	Python	1424
8	数据科学	438	8	数据建模	1254
9	自然科学知识	432	9	Hadoop	1160
10	统计学	421	10	数据仓库	1101

家为社会所用的大数据人才为宗旨，是人才培养的过程环节；企业则以选拔具备就业潜质的大数据可发展性人才，在实际工作中引领其在国家需求和社会需求中发挥最大作用为导向，是人才培养的结果环节和人才应用的发展环节。因此，在大数据专业知识结构如此庞大[10]的情况下，单方面要求高校在人才供需对接上致力的效益甚微，从国家角度和企业角度都应合力并助力人才培养过程，校企国三方融合培养模式才是大数据人才供需对接的有效手段。

#### 4. 结论

本研究以高校培养方案和网络招聘公告为双向抓手，一方面探究高校培养方案中大数据人才培养的能力素质特征，另一方面挖掘社会需求中大数据人才任职的能力素质特征，并进一步将供需两侧进行匹配分析，明确人才培养生命周期的缺失与不足。基于研究结果与分析，本研究提出以下培养模式。

(一) 聚焦国家战略需要，以市场需求为导向，释放学科建设新动能。

近年来，中央及地方出台各类大数据政策文件支持、引导大数据人才培养，高校应加强面向重大战略的统筹和部署，更新细化培养方案，及时调整人才培养方向。同时，也应实时跟进社会人才需求发展现状及变化趋势，以培养为国家所用、为社会所用的高质量大数据人才为行动准则，把市场需求中现存的核心技能作为重心，新兴技术作为个性拓展明确人才培养定位与毕业能力要求。

(二) 以学科特质为内核，以技术工具为抓手，形成“大数据+”课程培养体系。

根据研究结果分析，当前大数据人才培养单位明确将计算机、数学、统计学、外语等基础课程作为培养根基，将数据分析能力作为培养核心，并拓展其技术性、实践性、创新型。因此在整个知识技能培养体系中，应格外关注数据分析能力、数学基础、大数据技术工具应用等学科核心知识技能，尤其在供需匹配中产生偏差的运用技术工具、编程语言、数据库等解决实际问题方面提高重视度，把握这些关键能力有助于大数据人才加速拓展形成完整的知识技能体系，提高人才培养速度与质量，促进人才供需有效对接。同时，由于其学科交叉性，大数据人才除需掌握与本专业相关的知识技能外，还应进一步拓宽知识视野，例如高校可以增加社会学、语言文学、量子物理等跨领域课程供学生选择，形成“大数据+”的课程培养体系。

(三) 构建与企业、国家协同培育机制

人才供需对接的深度融合需高校、企业、国家三方相互合作、并驾齐驱，以解决大数据人才就业不

适应性及岗位人才缺口等问题。其中,高校应以社会需求为导向,以学科特性为内核,以技术工具为抓手,做好高质量人才的供给者;企业需要推进人才生态发展,深度参与人才培养生命周期环节中,做好人才培养的接替者,渗入到人才培养的过程环节,充当数据支持者的角色[11],为高校人才培养提供大量真实数据,帮助教育教学实践性。同时提供企业优秀员工作为辅助师资,下沉到课堂中为接替大数据人才培养做好准备;国家可助力构建学校、企业、国家协同培育机制,做好大数据人才发展的引领者。通过制定地方大数据产业发展计划和扶持政策,协调当地院校和本土企业,促成校企联合培养模式的具体实施,例如牵头或出资建设大数据人才培养项目,推进跨高校、跨学科、跨领域的大数据产业发展。

## 基金项目

天津师范大学教学改革项目“新文科背景下数据科学课程教学新模式的构建与实践”(项目编号:JGYB01220066)。

## 参考文献

- [1] 中国信息通信研究院. 大数据白皮书(2022) [EB/OL]. [http://www.caict.ac.cn/kxyj/qwfb/bps/202301/t20230104\\_413644.htm](http://www.caict.ac.cn/kxyj/qwfb/bps/202301/t20230104_413644.htm), 2022-01-04.
- [2] 帆软数据应用研究院. 中国数据人才供需关系与培养报告[EB/OL]. <https://research.fanruan.com/research/special-topic/data-talent-report-2020/>, 2020-12-30.
- [3] 陈沫, 李广建, 陈聪聪. 情报学取向的“数据科学与大数据技术”专业人才培养[J]. 图书情报工作, 2019, 63(12): 5-11.
- [4] 王晰巍, 李玥琪, 贾若男, 等. 新文科背景下大数据管理与应用专业人才培养模式[J]. 图书情报工作, 2021, 65(17): 45-56.
- [5] Fang, F. and Zhou, Y. (2021) A Study on Recruitment of Data Analyst Based on Text Mining and Visualization Technology. *Journal of Physics: Conference Series*, **1952**, 042017. <https://doi.org/10.1088/1742-6596/1952/4/042017>
- [6] 袁毅, 陶鑫琪, 李瑾萱, 刘娅娴, 汪晓芸, 景香玉. 基于招聘文本实体挖掘的人才供需分析——以人工智能领域为例[J]. 图书情报工作, 2022, 66(14): 101-118.
- [7] 桑秀丽, 黄天一, 刘继云, 解瑞玺. 工程管理类专业学位硕士培养目标对产业需求的适应性研究[J]. 科技和产业, 2021, 21(4): 130-138.
- [8] Cui, X., Dai, F., Sun, C., et al. (2021) BiLSTM-Attention-CRF Model for Entity Extraction in Internet Recruitment Data. *Procedia Computer Science*, **183**, 706-712. <https://doi.org/10.1016/j.procs.2021.02.118>
- [9] 易新河, 杨鹏, 文益民. 中文招聘文档中专业技能词抽取的跨域迁移学习[J]. 数据分析与知识发现, 2022, 6(Z1): 274-288.
- [10] 崔晓龙, 张敏, 张磊, 边胜琴, 郭茜. 新工科背景下应用型大数据人才培养课程群研究与建设[J]. 实验技术与管理, 2021, 38(2): 213-218.
- [11] 杨玲. 高校人才培养与人才需求深度融合破困研究[J]. 河北广播电视大学学报, 2022, 27(1): 81-84.