DOI 编码：10.3969/j.issn.1672-884x.2018.07.018

组织人力资源大数据研究框架与文献述评

刘善仕 孙 博 葛淳棉 彭秋萍 周怀康
（华南理工大学工商管理学院）

摘要：基于数据类型和研究问题层次构建了人力资源大数据的研究框架，并从内容数据、关系数据以及用户行为数据 3 个角度对国内外相关文献进行了回顾。研究结果表明，基于大数据的人力资源管理研究目前尚处于初级发展阶段，相关研究主要集中于宏观层面人力资本管理、人力资源流动、政策评价、组织间网络等主题，对微观层面人力资源管理涉及的招聘、培训、绩效管理等主题的研究还不够深入。鉴于此，对未来可以基于人力资源大数据进行研究的方向进行了展望，以进一步推进国内基于人力资源大数据的研究。

关键词：人力资源大数据；在线简历；社交网络；电子化人力资源管理；网络招聘

中图法分类号：C93 **文献标志码**：A **文章编号**：1672-884X(2018)07-1098-09

The Research Framework of Human Resources Big Data and a Literature Review

LIU Shanshi SUN Bo GE Chunmian PENG Qiuping ZHOU Huaikang
(South China University of Technology, Guangzhou, China)

Abstract: This paper constructs a research framework of Human Resources Big Data based on different data types and levels of research topics, and further reviews the existing literature from the perspectives of content data, relational data and user behavior data. The results demonstrate that the study of human resource management using big data is still in the primary development stage. The relevant literature mainly focuses on topics at the macro level, such as talent flow, policy evaluation, inter-organizational network, etc. However, research on recruitment, training and performance management at the micro level is not as comprehensive. In view of this, we have given some research directions on human resources with big data.

Key words: human resources big data; online resume; social networks; e-hrm; e-recruiting

随着电子化人力资源管理（E-HRM）、网络职业招聘平台、社交网络应用、在线劳动力市场等技术在组织人力资源管理中应用程度的不断加深，以及大数据技术的日臻成熟，组织收集、存储和使用与人力资源管理相关的海量数据将变得更加便捷和系统化[1]。这些以结构化、半结构化和非结构化形式存在的数据的量化也将成为可能[2]。在此背景下，可用于宏、微观人力资源管理研究的人力资源大数据应运而生。同以往抽样调查数据相比，人力资源大数据具备样本量大、实时性、动态性、价值性等特征[3]，更能真实、深入地反映个体与组织的活动和状态[4]，并最大程度控制量化研究中的测量偏误[5]和统计偏误[6]，以及探索更为细致的研究问题[7]，使得人力资源管理研究迎来新的机遇。

但需要指出的是，作为一个新兴研究主题，学术界目前尚未厘清人力资源大数据的概念、研究范围、数据来源与类型以及适用的理论。鉴于此，在回顾国内外现有相关文献的基础上，本研究首先讨论和总结了人力资源大数据的内涵、数据来源和研究框架；其次，按照提出的人力资源大数据研究框架，将现有基于大数据的人力资源管理研究分为基于内容、关系和用户行为 3 个方面，并进一步探究了不同类型数据下宏、微观人力资源管理的研究主题及其对应的理论基础；最后，指出了人力资源大数据存在

收稿日期：2017-12-26

基金项目：国家自然科学基金资助项目(71503084)；中央高校基本科研业务费资助项目(2017BQ049)；广东省普通高校哲学社会科学重点实验室资助项目(2015WSYS001)；广州市哲学社会科学十三五规划资助课题(2017GZZK10)

的问题以及潜在的研究方向。

1　人力资源大数据的概念、数据来源与研究框架

1.1　人力资源大数据的概念

得益于大数据在社会管理、商务管理以及企业管理等领域的广泛应用，人们对大数据的认识越来越全面[8]。学术界和实务界从不同的视角界定了大数据的概念，包括大数据的特征[9]、价值[10]、来源[9]和处理方法[11]等。尽管如此，目前尚无一个得到学界广泛认同的大数据定义。但有关大数据概念的一个共识是，大数据是对“海量数据”和“大规模数据”等概念的继承和发展，并在数据体量、复杂性和产生速度等3个方面远远超出了传统的数据形态[8]。当前，有以下比较具有代表性的大数据定义：美国国家科学基金会[9]认为，大数据是指基于仪器、互联网交易、电子邮件、视频等数据源所生成的大规模、多样的、复杂的、纵向的和/或分布式的数据集。

与此同时，虽然基于大数据的人力资源管理研究也越来越多，但当前学术界和实务界并没有关于人力资源大数据的正式定义。从数据来源和研究层次看，相关研究更多的是基于大数据对特定的人力资源管理问题进行研究。一方面，这些研究运用区别于传统抽样的大数据方法来获得大规模反映个体行为、关系和状态的数据；另一方面，相关研究基于这些数据来研究微观层面组织内部人力资源管理涉及的招聘、培训和绩效管理等活动，以及宏观层面组织、区域的人力资本管理等问题。例如，KLUEMPER等[12]基于社交应用的内容数据对个体的个性进行评估，用于识别与组织匹配的应聘者；GUO等[13]基于内部工作视频监控数据研究了员工危险行为的干预以及培训方案的优化；刘善仕等[14]基于在线简历的职业生涯数据，从社会网络的视角研究了企业间人才流动对其创新绩效的影响。

由此，基于大数据的核心概念和代表性定义，本研究认为，人力资源大数据是指在信息技术和互联网技术发展的背景中产生的，可反映组织及其个体的行为、关系或状态的，并能够用于宏、微观层面人力资源管理研究的海量数据集。

1.2　人力资源大数据的来源与数据类型

基于对文献的分析和实践的观察，本研究总结了当前人力资源大数据的五大主要来源。具体包括：①网络职业招聘平台。网络职业招聘平台的不断涌现，使得以往现场招聘所涉及的活动逐渐转移到在线平台以电子形式呈现，并带来了个人简历、招聘单位需求信息以及雇佣双方匹配信息等3个方面的数据。②社交网络应用。社交网络作为基于用户关系的在线内容创造与传播平台，是人们用来分享意见、见解、经验和观点的工具，其可以提供基于内容的数据、基于用户互动的关系数据以及基于用户的使用行为数据。③在线劳动力市场。线下劳动力市场向线上发展，带来了与工作任务和工作者相关的信息。④在线知识社区、搜索引擎。在线知识问答社区以及网络搜索引擎是个体在工作和生活中遇到问题寻求解答的在线空间，这些来源可以为科学研究提供与员工及其求教相关的信息。⑤人力资源管理信息系统。人力资源管理信息系统的主要功能是系统收集、管理和分析以组织内部为主的人力资源信息，为组织提供涉及招聘、挑选、培训、绩效管理等在人力资源管理过程中产生的数据。

总体而言，以上五大数据源所产生的数据可以按照数据类型分为基于内容的数据、基于关系的数据和基于用户行为的数据3类。表1按照这种数据分类方式将五大数据源所产生的数据进行了细分，并指出了其可能主要适用的研究层次。

表1　人力资源大数据的来源与数据结构

类别	代表性应用	基于内容的数据	基于关系的数据	基于用户行为的数据	研究层次
网络职业招聘平台	领英、前程无忧、猎聘、企业门户网站等	个体所在地、年龄、学历、婚育和健康状况等重要背景信息；个人自评、求职意向等职业生涯历史信息	基于共同背景的关系（如同乡、校友、同事等）；基于职业流动的关系	简历投递、招聘信息浏览和推送回应等	组织内、组织间、区域
社交网络应用	脸书、推特、微博、微信等	文字、图片、视频	用户之间的互动关系	用户发表内容的形式、频率、长短等	组织内、组织间、区域
在线劳动力市场	MTurk、UpWork、猪八戒网、阿里众包等	工作要求、劳动力供需、雇佣匹配	基于共同任务、共同雇佣的关系	工作申请行为、浏览行为等	区域、组织内
在线知识社区、搜索引擎	Quora、知乎、果壳等；谷歌、雅虎、百度等	个人信息、提问内容、回答内容	基于主题的关系	基于搜索行为的数据、求教频率等	组织内
人力资源管理信息系统	Oracle、Workday、SAP、金蝶、用友等	基于工作内容的数据；招募、挑选、培训、绩效等活动产生的数据	竞争合作关系、轮岗数据等	工作场所操作行为、工作过程等	组织内

注：资料来源为本文作者整理。

1.3 人力资源大数据的研究框架

作为大数据的子类之一，人力资源大数据同样面临是否需要理论，以及理论与数据如何驱动科学研究的问题。在总结学术界关于大数据理论基础的必要性以及研究范式的讨论的基础上，本研究提出人力资源大数据的研究框架。

1.3.1 理论基础

学术界对基于大数据的科学研究和商业决策是否需要理论支撑的问题，存在两种不同的观点。理论挑战派指出，大数据时代需要注重相关关系，而可以忽略甚至是取代因果关系。例如，ANDERSON[15]认为，足够多的数据其自身就可以呈现一定的规律，因而相关关系可以取代因果关系，这时科学甚至可以没有模型、理论，或者实际上可以完全不依赖任何机制性解释而发展。理论拥护派则表示，大数据时代的到来为科学研究和商业决策提供了前所未有的机遇，而这个时代比任何时候都更加需要理论以及理论的创新，用以指导基于大数据的研究，避免大数据研究的泛滥。例如，王天思[16]认为，因果关系是人类理性行为与活动的基本依据，人类理性本身不可能否定因果关系。大数据所凸显的相关关系的确从实践层面实质性地推进了对传统因果概念的深入反思，但如何从相关关系中推断出因果关系才是大数据研究的真正问题所在，失去与因果关联的相关关系很容易陷入虚假泥淖。

1.3.2 演绎导向与归纳导向

演绎导向的研究和归纳导向的研究都是理解和解释现象的主流研究方法[17]。从理论开始的研究被认为是演绎导向的假设检验研究；从观察开始的研究则被认为是归纳导向的建设理论研究[18]。然而，归纳导向的建设理论研究在组织科学中的发展一直较慢，其主要原因是受制于数据的规模和质量，即归纳驱动的研究很难从相对较少的观测值中获得强有力的结论[3]。大数据时代将有可能改变这种局面。虽然大数据在本质上仍然属于样本数据，但是其数量和质量较以往有大幅提升趋近总体的趋势，是对现实世界近似真实的描绘。

大数据技术包括对数据的收集、管理、分析和应用，是研究的工具和素材[2]。其本身既不具有归纳性，也不具有演绎性[3]。由此，研究者需要根据研究目的来使用大数据。归纳导向的建设理论研究从观察现象到形成理论和命题，大数据分析师和理论研究者通常是使用数据中所观察到的模式来指导未来的决策。这种数据驱动的大数据分析与传统研究中的归纳推理有一定的相似之处。演绎导向的假设检验研究则是先寻找可能的理论基础进行逻辑演绎推导，然后收集精细的大数据进行验证。实际上，大数据为未来实证研究提供了一种兼顾演绎和归纳的可能，即研究者可以先基于高质量、大规模的数据做归纳分析，寻找变量间可能存在的关系，然后再使用演绎方法予以论证并实证。

1.3.3 研究框架

基于数据类型和研究问题的层次，本研究提出了由人力资源大数据研究议题和相关理论基础两个层次组成的人力资源大数据研究框架（见图1）。基于该框架，本研究将现有基于人力资源大数据的实证研究归为6类：①个体评估、培训需求识别、绩效管理（微观层面基于内

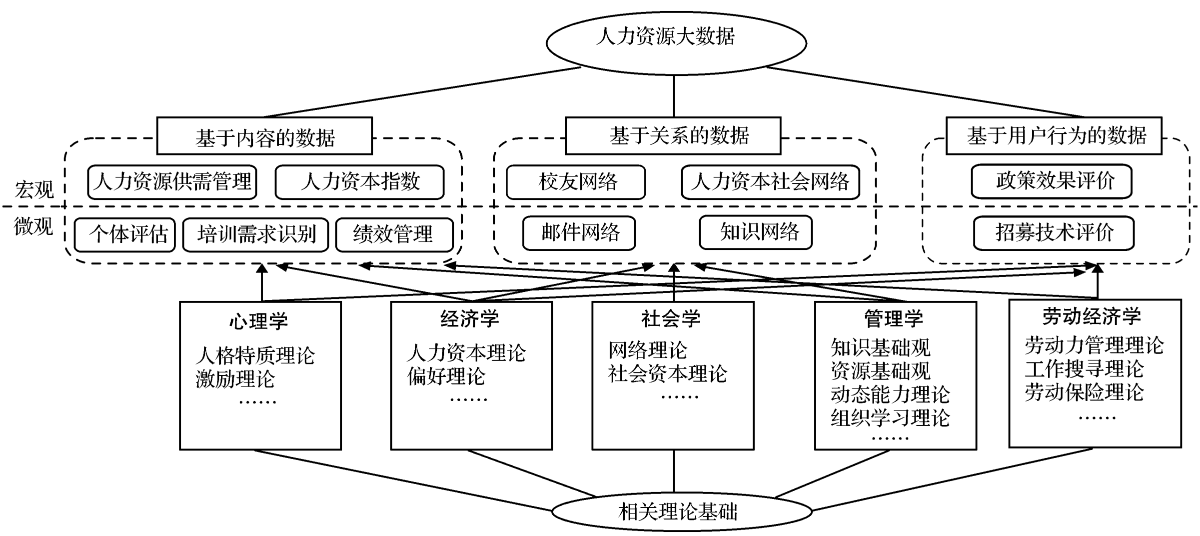

图1 人力资源大数据研究框架图

容数据的研究)；②人力资源供需管理、人力资本指数(宏观层面基于内容数据的研究)；③邮件网络、知识网络(微观层面基于关系数据的研究)；④人力资本社会网络、校友网络(宏观层面基于关系数据的研究)；⑤招募技术评价(微观层面基于用户行为数据的研究)；⑥政策效果评价(宏观层面基于用户行为数据的研究)。此外,现有研究的理论主要来源于心理学、经济学、社会学、管理学等学科领域。

2 人力资源大数据研究现状

2.1 基于内容数据的研究

从内容上看,人力资源大数据提供了个体和组织在使用相关技术过程中留下的信息,包括文本、图像以及视频等形式的数据。例如,表1中,由基于内容的数据栏可知,个体使用社交网络发表的个人评论,在网络职业招聘平台上投递的简历,用人单位发布的职业要求,使用电子化人力资源管理系统产生的培训、绩效信息等。现有基于内容数据的实证研究,在微观层面主要关注个体评估、培训需求识别、绩效管理等主题,主要理论是人格特质理论、激励理论等;在宏观层面关注人力资源供需管理、人力资本指数等主题,主要理论是人力资本理论、劳动力管理理论等。

2.1.1 个体评估

因为社会赞许和个体欺瞒的存在,以往基于问卷的自我报告式个性测试、能力测试量表具有有限的信度和效度,不能很好地反映应聘者的真实情况[19]。基于社交网络内容的个体评估可以较好地解决评估的真实性问题[4]。这是因为,一方面,与个体交往的熟人可以监督个体所发布的信息真伪;另一方面,社交网络平台包含了大量诸如个人想法、图像、社会活动等信息,可以在某种程度上反映个性、认知能力、自我效能、人与组织匹配程度等个体特征和状态[12,20]。例如,具有自恋型人格的个体可能会在社交网络上发布更多的社交活动、自我激励的文字和图片[21]。

尽管基于社交网络内容的个体评估可能成为组织挑选员工的新方法,但现有的实证研究还比较匮乏,并且出现了较大的分歧。比较有影响的研究如KLUEMPER等[12]检验了基于社交网络内容的个体评估。该研究结果表明,基于社交网络历史信息的个性评估与基于自我评价、第三方评价的个性评估具有较高的一致性,并能够稳定预测个体的工作能力和绩效。相反,VAN IDDEKINGE等[5]认为,基于社交网络内容的评估并非是一个好的员工挑选方法。该研究从挑选的目的出发,分别比较了基于社交网络的评估与基于问卷的评估对应聘者加入组织后的工作行为和绩效的预测,发现后者具有更好的预测力。

2.1.2 培训需求识别

制定符合个体需求的培训方案将有利于个体能力的提升和职业生涯的发展,但传统的组织培训方案制定流程对个体需求的挖掘还不够。一般而言,个体的培训主要包括组织公共培训和具有针对性的个体职业培训,这些培训多是基于组织需求和个体申报而制定的。事实上,个体对自己的实际培训需求和职业生涯发展是难以完全了解的,组织所提供的公共培训也不都能照顾到个体的个性化需求。由此,以个体的实际需求为导向,企业可以通过大数据技术挖掘个体潜在的需求和参照以往成功经验两种方式制定个体的培训方案。

个体在知识社区的求助、社交网络信息、浏览器搜索行为、工作场所操作行为等均可以作为需求识别的信息源。组织可根据这些信息来确定个体的培训需求,也可以依据人口统计学特征、教育经历以及绩效记录等因素进行个体聚类分析,即找到与个体属于同一类别的、已获得职业成功的个体,参照其培训经历制定个体的培训方案[3]。现有研究已经尝试分析基于工作场所行为以及海量职业生涯分析的培训需求识别。例如,GUO等[13]利用视频分析技术,基于安全行为分析方法观察和分析建筑工作现场的不安全行为,并制定相关培训方案。

2.1.3 绩效管理

电子化绩效管理系统、智能手机、电脑等电子设备的应用也给组织的绩效管理研究带来了新机会。这些新技术可以为组织完整地记录员工的日常工作行为、工作过程、绩效以及绩效改进等信息,为深入、系统地研究员工绩效的形成与提高提供了坚实的数据基础。当前相关的研究主要集中在绩效过程管理,即利用各种电子设备记录员工的日常工作过程,并利用数据处理技术对这些视频、音频、文本等信息进行结构化处理,形成可量化的数据,以探索绩效形成的过程。例如,GUO等[13]采用大数据技术对海量的非结构化视频信息进行处理以研究绩效过程,为今后基于非结构化视频数据的研究提供了参照。

2.1.4 人力资源供需管理

在劳动力市场人力资源的供需研究中，存在偏重劳动力供给而对劳动力需求研究不足的情况，这在一定程度上与缺乏有效数据来源有关[22]。以往基于问卷的劳动力供需调查常常受到数据收集成本的限制，难以获得大规模、高质量的数据集。而在线职业网站的数据成本则相对比较固定，并且随着平台规模和影响力的增大，其能收集到更大数量和更高质量的数据[7]。基于在线职业网站的数据，研究者可以从供给、需求以及供需动态匹配等3个角度研究人力资源供需管理。

首先，研究者可以基于个体简历的职业生涯数据还原特定范围内人力资源的存量、结构与流动情况，以反映劳动力市场供给。例如，GE等[6]基于在线简历的职业经历描绘了企业间的人才流动，并研究了IT人才、科学家等知识工作者的职业流动所带来的知识溢出效应。其次，通过汇聚特定范围内的组织的招聘信息，可以一定程度上还原劳动力市场的需求状况与特点。例如，SHEN等[23]基于“智联招聘”2008～2010年期间超过105万条的招聘数据，研究了中国企业在招聘中是否存在性别偏好的特点。最后，基于劳动力市场的供需动态可以研究劳动力供需的动态匹配。例如，HORTON等[24]利用实验方法，研究了工作报酬区间的动态生成。该研究设定雇主在发布工作时必须先给出工作的报酬，然后将这些信息推荐给潜在的雇员，询问雇员是否接受，再基于大量匹配数据的排序，以获得该份工作的劳动报酬区间。

2.1.5 人力资本指数

人力资本指数在某种程度上可以反映一个区域的人才存量、结构和动态变化，可以有效预测区域经济发展、创新活力等重要绩效指标。受限于数据可得性，以往主流的人力资本指数测度方法侧重使用年龄、受教育年限等因素来构造指数，致使研究者难以全面、深入、动态地研究人力资本的存量、结构和变化，以及因此带来的对经济后果的影响。简历中的工作经历、社会关系、年龄、教育经历等成分为更好地测量人力资本指数提供了契机。

2.2 基于关系数据的研究

从关系上看，人力资源大数据为研究组织内外部关系提供了丰富的数据基础，包括共同背景、共同行为等数据。现有基于关系数据的实证研究，在微观层面主要是从内部邮件网络来研究内部工作网络的结构及其与领导绩效的关系，以及知识社区用户网络；在宏观层面主要是基于简历的职业流动数据和背景信息来构建企业层面的人才流动网络、校友网，并从个体中心网络视角研究网络结构及其对企业经济后果的影响。基于关系数据的实证研究其主要的理论基础有网络理论、社会资本理论、人力资本理论等。

2.2.1 邮件网络与知识网络

相比于组织正式层级结构，非正式社会网络是更加自发的协调机制和更有效率的信息与知识传播渠道。非正式网络不仅有利于调和组织内部不同部门员工之间的冲突，以更高效的方式解决问题，还可以通过加快信息扩散和知识分享而有利于组织绩效。信息通信技术在工作领域的应用，使得组织内部的非正式网络更容易形成，同时也为科学研究提供了大量构建内部网络关系的数据。例如，TASHIRO等[25]基于一家国际化制造企业中层经理的邮件往来关系构建了企业内部邮件网络，并从社会网络结构的视角研究了个体在网络中的位置与其领导力绩效的关系。该研究发现，占据中介中心位置和点入中心位置越多的经理，其领导力绩效越好。这说明具有影响力的领导其更容易被信任，而处于沟通网络的中心位置。此外，也有基于公共知识社区、论坛等在线平台的用户关系、发帖关系构建的多层次关系网络。例如，唐洪婷等[26]基于小米科技MIUI(米柚)社区用户的发帖内容及互动关系构建了多层次用户知识网络，并研究了其在大众协同创新中的作用。

2.2.2 人力资本社会网络与校友网络

随着职业社交网络的出现和数据收集技术的进步，基于在线简历数据的组织间社会网络研究将有可能成为未来人力资源管理在宏观层次的重要方向。特别是最近10年里，一批学者基于个体简历的职业生涯数据和背景数据，从社会网络视角构建了人才流动网络、校友网络等组织间关系网络，并取得较好的研究成果[27,28]。

这些研究认为，企业可以通过人才流动或校友关系获取和控制其他企业的信息、知识等人力资本资源以及社会资本，而企业在关系网络中所占据的位置一定程度表征了组织所获得资源的质量和数量，这些资源将进一步影响到企业的行为和绩效[14]。中心度和结构洞是常用的网络位置指标，中心度表示组织在整体网络中处于核心位置的程度，表征组织在整个网络中获取资源的能力，一定程度衡量了资源来源

的广泛程度和资源的丰富程度；而结构洞表示的是焦点企业与两个不直接相连的企业有直接联结的关系模式，表示组织获得异质性资源的能力，一定程度上表征了资源的差异程度。例如，在人才流动网络研究中，DOKKO 等[27]的研究分析了企业在人才流动网络中的中心度位置对企业制定行业标准的影响；SHIPILOV 等[28]从中心度的视角分别研究了流入中心度和流出中心度对企业创新绩效的影响，以及组织地位和地理距离对这种关系的调节作用；刘善仕等[14]整合人力资本理论和社会网络理论，提出了人力资本社会网络的概念，并采用大规模在线简历数据构建了企业间人才流动网络，从中心度和结构洞的视角研究不同类型人才流动对企业创新的影响。此外，在校友网络研究中，申宇等[29]基于我国上市公司高管简历的教育背景信息构建了企业校友网络，并研究了校友网络对基金业绩的影响。

2.3 基于用户行为数据的研究

从用户行为上看，个体和组织在使用相关技术过程中会留下一系列的操作信息，包括点击、浏览、行为频率、行为时长等数据。例如，表1中，由基于用户行为的数据栏可知，个体使用社交网络发表内容的频率、长短，在网络职业招聘平台上投递简历的次数、浏览招聘信息的停留区域、时长等。现有基于用户行为数据的实证研究，在微观层面主要是从个体申请工作行为的角度来评价新技术用于企业招募时的效果，其主要理论是用户偏好理论；在宏观层面从工作申请以及工作搜索行为来评估相关人才政策效果，其主要理论是劳动保险理论、工作搜寻理论等。

2.3.1 招募技术评价

高效的招募工作有利于组织提高备选人才池中人才的质量和多样性。随着科学技术的进步及其在人力资源管理领域中的应用，越来越多的组织采用基于网页、职业社交网络等技术的招募[30]。据统计，美国大约90%的大型组织正在采用至少一种技术发布招聘广告[31]。这些新技术在企业员工招募管理中的应用效果研究日益引起关注。

鉴于此，现有研究从员工申请工作行为的角度研究了基于网页、社交媒体等新型招募技术的效果。首先，相关研究认为，一些招募技术所展现的特征可以更好地展示组织的形象、与工作相关的信息，从而影响到应聘者对组织的态度和工作的兴趣，吸引更多的应聘者提交工作申请[1]。例如，DINEEN 等[32]指出，基于 Web 2.0 的网络招募较 Web 1.0 具有强烈的交互特征，个体在浏览招募信息的过程中，可以与后台工作人员进行互动，从而获得更好的使用体验和信息反馈，帮助应聘者对人岗匹配、个体需求与能力匹配、个体与组织匹配进行更加深入的评估，有利于吸引更多高质量的应聘者浏览招募信息，以及促使其做出工作申请行为。也有研究从优化技术的角度研究了招募方法的改进。例如，HORTON[33]利用实验的方法，通过改变 Odesk（在线工作社区）工作推荐功能的算法，研究不同算法对雇主招聘成功率的影响。

2.3.2 政策效果评价

人才的存量、结构和流动是人力资源管理在区域层次重点研究的问题。以往基于抽样调查、人口普查的人才管理和人才政策制定，存在时效性不佳、成本高等特点，使得政策制定部门难以实时掌握人才资源的供需情况并制定有针对性的人才管理策略。随着大数据收集、处理技术的日臻成熟，通过对人才流动等个体活动记录进行量化处理，可以得到实时反映人力资本资源转移情况的数据，从而有助于相关主体实时、精准地掌握人才供需情况，制定高效、适用的人才管理策略。人才政策评价是当前基于用户行为数据在区域人才管理研究中的主要关注点。例如，MARINESCU[30]基于凯业必达网站的工作申请和职位空缺的数据，研究了在美国金融危机期间，政府延长失业补助期限对工作申请和职位空缺的影响，以评估失业补助期限延长政策的有效性。

3 人力资源大数据存在的问题

3.1 人力资源大数据的效度问题

人力资源大数据的效度问题是人力资源大数据研究最容易受到质疑的问题之一[20]。从现有文献看，质疑主要源于以下3个方面：①测量内容没有充分地反映目标构念。持慎用基于社交网络来挑选员工的学者认为，大多数社交网络并非是为挑选员工而设计的应用，其主要功能偏重社会交往，其数据内容大多数与工作无关。这就造成基于社交媒体历史信息的个性与能力测试可能存在信息不全的问题，导致基于社交网络的评估的效度较低，使得评价结果出现偏差[5,34]。②数据量大与能否准确地反映关键概念不存在必然的关系。虽然，大量、丰富的数据有助于降低统计偏差，但海量、多来源的数据并不意味着其可以充分地反映目标构念。社

交网络尽管带来了海量内容信息，但是海量信息的价值存在边际递减的规律，数据量越大并非一定能提高预测的精准度。③测量体系难以结构化或标准化，导致测量体系缺乏对比性[20]。不同的数据来源其内容和格式可能存在较大程度的差别，这就导致很难对基于不同来源的评估做出对比。例如，同为社交网络应用，领英偏重职业社交，而脸书偏重非职业社交，这就导致这两个平台在内容和数据格式上存在一定程度的差别[34]。由此，在个性测试中，针对同一个对象基于领英的评估和基于脸书的评估可能存在较大的差别。

然而，即便是人力资源大数据存在这些潜在的效度问题，但并非完全不能应用于研究和实践。首先，为避免信息不全所造成的决策偏误，组织可以基于已有信息来验证应聘者所提供的信息的真实性，以及基于消极信息做出拒绝录用的决策。其次，组织仍然可以基于不同平台的内容特征做出相应的个体评估。例如，利用非职业社交网络的信息进行个体的个性评估，采用职业社交网络的信息进行工作能力评估[35]。最后，基于社交网络内容的员工评价也可以做到一定程度的结构化。例如，收集性别、民族和身体状况等信息来表征相应的构念，利用关键事件法来定义个体的某些状态和行为[20]。

3.2 人力资源大数据的收集和来源问题

人力资源大数据在收集过程中不依赖于随机抽样，因此，无法确定抽样数据的总体情况，这将导致最终研究结果可能是建立在有偏差的数据集之上[7]。与问卷抽样调查不同，基于大数据技术收集的劳动力市场信息，很难得知数据来源的背景特征。相比人口现状调查数据，研究者对人力资本大数据样本的背景信息可能知之甚少；同时，匿名情况下个体所提供的信息和展现的行为可能与真实情况相去甚远。在无法对数据的总体情况进行描述的情况下，为降低抽样偏误，可以将收集的在线数据与基本统计数据进行匹配和比对。例如，劳动信息市场中的工资、教育、年龄、性别等在线数据可以与人口普查或其他补充来源的人口调查数据进行匹配和比对。

3.3 人力资源大数据的伦理问题

人力资源大数据也具有大数据普遍存在的隐私问题，特别是基于社交网络的应用与研究，其隐私保护与研究边界的权衡问题尤其突出。首先，基于网络的数据收集不同于问卷调查，不需要得到被访者的许可，隐私侵犯可能伤害到消费者和用户的利益。其次，基于社交网络应用信息的招聘和挑选可能更容易产生个体歧视，进而导致不公正的招聘与挑选。基于姓名、图片、个人心情、评论等信息可以判断个体的种族、性别、年龄等敏感信息。基于这类敏感信息的招聘和挑选决策使得歧视更加隐秘。由此，随着大数据的发展，相应的政策和规范需要配套发展起来。对于隐私问题的处理，一方面，需要界定大数据责任，包括大数据预测技术应用中道德选择和责任承担问题，建立互联网背景下的大数据收集和使用监管体制（如保护隐私的责任由个体转到数据使用者）；另一方面，在数据使用过程中，从技术层面尽量最大化地隐藏用户隐私（如通过数据的模糊处理、匿名处理进行个体隐私保护）。

4 未来研究趋势

人力资源大数据时代的到来为人力资源管理研究提供了全新的数据基础，使得以往因缺乏精细数据而难以深入研究的问题获得新生。但是现有基于人力资源大数据的研究尚处于初级发展阶段，这种特点主要表现在：①相关文献主要集中于宏观层面人力资本管理、人力资源流动、政策评价、组织间网络等主题，对微观层面人力资源管理涉及的招聘、培训、绩效管理等主题的研究还不多。②为数不多的微观层面研究仍属于探索性研究。例如，比较新技术与传统方法的优劣、基于内容数据的个体评估的准确性讨论、基于工作监控视频数据的简单分析等。这种状况可能与当前大数据处理技术尚难处理非结构化数据、组织内部人力资源大数据比较敏感而难于获取，以及研究者尚未明确人力资源大数据情境下的研究范式等现状有关。即便如此，对未来进行有益的探索也是很有必要和具有价值的。

本研究从数据类型和研究问题层次的角度构建了一个可扩展性研究框架。有关人力资源大数据的研究主题及其相关理论基础还有待延伸，未来基于人力资源大数据的研究在宏观、微观层面具体可从以下6点开展进一步探索。

4.1 微观层面

（1）*关注精准招聘* 一方面，通过精准招聘在目标人群中寻找与组织、岗位匹配度高的个体，实现招募信息的精准投放；另一方面，制定个性化的招募广告吸引个体提交工作申请，探索在提高应聘者数量的同时提升应聘者质量。

例如,基于大数据的素质模型识别、个体招聘信息偏好研究等。

(2)**研究基于社交应用的个体评估**　探索基于不同平台的个体评估可以获得何种指标,如何根据不同的平台构建结构化的测度体系。基于关键事件和文本分析的个体评估,不同的事件和字段具有何种意义,可以表征何种构念。例如,基于职业社交平台的工作评估,基于非职业社交平台的个体个性评估。

(3)**探索基于聚类的培训方案制定与个体职业生涯模式识别**　即研究个体的职业生涯培训经历、个体特征与职业生涯成功之间的关系,建立三者之间的函数关系,构建智能培训推荐系统。利用机器学习,通过不断地为该系统输入成功的案例优化其算法,以提升其推荐精准度。此外,通过对海量的职业生涯数据进行聚类分析,识别具有代表性的职业生涯模式,然后对每种职业生涯模式所涵盖的个体的背景信息进行分析,获取具有共性的个体特征,以此建立职业生涯智能推荐系统。

(4)**探索基于人力资源大数据的绩效考核与过程管理**　即分析不同岗位高绩效人才的基本信息、业绩数据、工作时间利用度、工作效率、工作生活习惯、社会行为等信息,描绘高绩效人才所具备的标准和特征,以此作为相关员工的绩效考核依据。此外,基于绩效大数据的实证研究还可以探索更深入的绩效产生过程。例如,在绩效主义盛行的背景下,许多企业为追求漂亮的绩效结果数据,而实行了诸如即时绩效管理等严苛的工作流程管理方法,这些方法也有可能导致员工的激进工作行为,并带来不良的组织影响。

4.2　宏观层面

(1)**丰富和深化基于关系数据的研究**　现有基于简历数据的组织社会网络研究呈现以下两个特点:①在数据上偏重以职业生涯经历构建关系网络,对简历数据中其他重要的背景信息(如教育背景、籍贯、同事等)研究不足,未来可以深入探索基于背景信息构建的组织校友网络、同乡网络、同事网络;②现有组织间网络集中从个体中心网络视角研究特定网络的影响后果,未来可以从二元关系、个体中心网络以及整体网络视角来研究组织间社会网络的形成、影响后果及作用机制,甚至是多层次网络的动态演化。

(2)**研究人才政策评价**　当前,各级政府都十分重视人才引进工作,出台了各具特色的人才吸引政策,但在这些政策中哪些因素会对个体的流动产生重要影响还未有确切定论。未来可以基于简历数据描绘人才的多层面流动,并以人才流动来衡量政策因素的效果及其整体政策效能。

参　考　文　献

[1] STONE D L, DEADRICK D L, LUKASZEWSKI K M, et al. The Influence of Technology on the Future of Human Resource Management [J]. Human Resource Management Review, 2015, 25(2): 216-231.

[2] GROVER P, KAR A K. Big Data Analytics: A Review on Theoretical Contributions and Tools Used in Literature [J]. Global Journal of Flexible Systems Management, 2017, 18(2): 1-27.

[3] MCABEE S T, LANDIS R S, BUIKE M I. Inductive Reasoning: The Promise of Big Data [J]. Human Resource Management Review, 2017, 27(2): 277-290

[4] BACK M D, STOPFER J M, VAZIRE S, et al. Facebook Profiles Reflect Actual Personality, Not Self-Idealization [J]. Psychological Science, 2010, 21(3): 372-374

[5] VAN IDDEKINGE C H, LANIVICH S E, ROTH P L, et al. Social Media for Selection? Validity and Adverse Impact Potential of a Facebook-Based Assessment [J]. Journal of Management, 2016, 42(7): 1 811-1 835.

[6] GE C, HUANG K, PNG I P L. Engineer/Scientist Careers: Patents, Online Profiles, and Misclassification Bias [J]. Strategic Management Journal, 2016, 37(1): 232-253.

[7] HORTON J J, TAMBE P. Labor Economists Get Their Microscope: Big Data and Labor Market Analysis[J]. Big Data, 2015, 3(3): 130-137.

[8] 冯芷艳,郭迅华,曾大军,等. 大数据背景下商务管理研究若干前沿课题 [J]. 管理科学学报, 2013, 16(1): 1-9.

[9] NSF. Core Techniques and Technologies for Advancing Big Data Science & Engineering (BIGDATA) [EB/OL]. (2012-03-30)[2017-12-26]. http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504767, 2012.

[10] GARTNER. Big Data[EB/OL]. [2017-12-20]. http://www.gartner.com/it-glossary/big-data/, 2013.

[11] WIKIPEDIA. Big Data[EB/OL]. [2017-12-26]. http://en.wikipedia.org/wiki/Big_data, 2013.

[12] KLUEMPER D H, ROSEN P A. Future Employment Selection Methods: Evaluating Social Networking Web Sites [J]. Journal of Managerial Psychol-

ogy,2009,24(6):67-580.

[13] GUO S Y,DING L Y,LUO H B,et al. A Big Data Based Platform of Workers' Behavior:Observations from the Field [J]. Accident Analysis and Prevention,2016,93(2):299-309.

[14] 刘善仕,孙博,葛淳棉,等. 人力资本社会网络与企业创新——基于在线简历数据的实证研究 [J]. 管理世界,2017(7):88-98.

[15] ANDERSON C. The End of Theory:The Data Deluge Makes the Scientific Method Obsolete [J]. Wired Magazine,2008,16(7):16-07.

[16] 王天思. 大数据中的因果关系及其哲学内涵 [J]. 中国社会科学,2016(5):22-42.

[17] MANTERE S,KETOKIVI M. Reasoning in Organization Science [J]. Academy of Management Review,2013, 38(1):70-89.

[18] 陈晓萍,徐淑英,樊景立. 组织与管理研究的实证方法 [M]. 第2版. 北京:北京大学出版社,2012.

[19] MORGESON F P, CAMPION M A, DIPBOYE R L. Reconsidering the Use of Personality Tests in Personnel Selection Contexts [J]. Personnel Psychology,2007,60(3):683-729.

[20] BROWN V R,VAUGHN E D. The Writing on the Facebook Wall:The Use of Social Networking Sites in Hiring Decisions [J]. Journal of Business and Psychology,2011,26(2):219-225.

[21] BUFFARDI L E, CAMPBELL W K. Narcissism and Social Networking Web Sites [J]. Personality and Social Psychology Bulletin,2008,34(10):1 303-1 314.

[22] PELLIZZARI M. Employers' Search and the Efficiency of Matching [J]. British Journal of Industrial Relations,2011,49(1):25-53.

[23] SHEN K,KUHN P. Do Chinese Employers Avoid Hiring Overqualified Workers? Evidence from an Internet Job Board [M]. POLACHEK S, TATSIRAMOS K . In Labor Market Issues in China. West Yorkshire: Emerald Group Publishing Limited, 2013:1-30.

[24] HORTON J J,JOHARI R . At What Quality and What Price? Eliciting Buyer Preferences as a Market Design Problem[C]//Proceedings of the Sixteenth ACM Conference on Economics and Computation, 2015(6):507.

[25] TASHIRO H,LAU A,MORI J,et al. E-Mail Networks and Leadership Performance [J]. Journal of the Association for Information Science and Technology,2012,63(3):600-606.

[26] 唐洪婷,李志宏,秦睿. 基于超网络的大众协同创新社区用户知识模型研究[J]. 管理学报, 2017,14(6):859-867.

[27] DOKKO G, ROSENKOPF L. Social Capital for Hire? Mobility of Technical Professionals and Firm Influence in Wireless Standards Committees [J]. Organization Science,2010,21(3):677-695.

[28] SHIPILOV A,GODART F C,CLEMENT J. Which Boundaries? How Mobility Networks Across Countries and Status Groups Affect the Creative Performance of Organizations [J]. Strategic Management Journal,2017,38(6):1232-1252.

[29] 申宇,赵静梅,何欣. 校友关系网络、基金投资业绩与"小圈子"效应[J]. 经济学(季刊),2015(4):403-428.

[30] MARINESCU I. The General Equilibrium Impacts of Unemployment Insurance: Evidence from a Large Online Job Board [J]. Journal of Public Economics, 2017,150(6):14-29.

[31] MACKELDEN L. How Do Big Recruiters Recruit Online [J]. Retrieved March, 2013,30:2014.

[32] DINEEN B R,NOE R A. Effects of Customization on Application Decisions and Applicant Pool Characteristics in a Web-Based Recruitment Context [J]. Journal of Applied Psychology, 2009, 94 (1): 224-234.

[33] HORTON J J. The Effects of Algorithmic Labor Market Recommendations: Evidence from a Field Experiment [J]. Journal of Labor Economics,2017, 35(2):345-385.

[34] ROTH P L,BOBKO P,VAN IDDEKINGE C H,et al. Social Media in Employee Selection Related Decisions: A Research Agenda for Uncharted Territory [J]. Journal of Management,2016,42(1):269-298.

[35] ROULIN N, BANGERTER A. Social Networking Websites in Personnel Selection [J]. Journal of Personnel Psychology,2013,12(3):143-151.

(编辑　郭恺)

通讯作者: 葛淳棉(1986～),男,广东汕头人。华南理工大学(广州市　510641)工商管理学院副教授、博士研究生导师。研究方向为战略管理、信息经济学、大数据分析等。E-mail:bmgecm@scut.edu.cn