

应用科研履历数据的情报研究进展

■ 卫炯圻^{1,2,3} 鲁晶晶^{1,2} 谭宗颖¹

¹中国科学院文献情报中心 北京 100190 ²中国科学院大学 北京 100049

³中国科学院心理研究所 北京 100101

摘要: [目的/意义] 由于科研履历提供了其他数据源所没有的丰富而且独特的信息,情报学研究者正越来越意识到它的重要性,并借助科研履历数据在多个领域开展了成果丰硕的各类研究。科研履历正逐步被看作是情报分析与研究的重要数据来源。[方法/过程] 采用宏微观和横纵向的两维分析方法,对近20年以科研履历数据作为主要数据源的学术研究文献进行调研与梳理。[结果/结论] 提出履历数据至少可以被应用到科研人员的职业成长、人才流动与科研合作、研究群体的特征分析以及科研项目、政策、建制的评估等4个领域的情报研究中。专家库带来的大量结构化履历数据以及以履历数据为枢纽形成的数据链,将推动情报学研究向纵深化发展。

关键词: 科研履历 情报研究 CV数据 职业成长 人才流动 项目评估

分类号: G250

DOI:10.13266/j.issn.0252-3116.2015.07.019

科研人员的个人职业履历(curriculum vitae, CV)越来越被当代的科技情报研究者看作是一种重要的数据来源。尤其是最近20年, CV数据被作为主要数据源用于科技人才的职业发展与职业特征、人才流动、科研合作以及科学项目和建制的评价等多个研究领域,表现出独特的数据价值。本文将通过对以 CV数据为主要数据源的各类研究进行综述,对 CV数据在多个领域的应用进行分析,并结合当前研究的主要问题,提出未来 CV数据的可能应用前景。

1 CV数据的优点

CV数据本身具有诸多优点。首先, CV数据的组织形式都是围绕个人的。所有的 CV数据可以确定地指向某一个个体,而文献计量数据中对个人数据的精确计量总会因为人名的模糊和多义而困难重重。其次, CV数据非常普遍,信息量大且种类丰富,可以形成一个人的全景数据^[1]。几乎每个科研人员都有自己的 CV,因为 CV对每一个职业研究人员都非常重要,其用途广泛,可用于记录个人科研成果、宣传自己的科研活动、申请基金和表彰、获得职业认可、寻求新的职位等。CV的普遍性使其在分析某些文献计量数据不能很好

覆盖的学科领域或研究产出时显得非常有用。B. Lepori 和 C. Probst 描绘瑞士传播学的领域地图时,选择了 CV数据^[2],便是因为 Web of Science 的数据无法有效覆盖传播学中非英语的和期刊论文之外的研究产出,这会遗漏大部分需要被纳入的研究成果。这时,研究者履历便成为了比标准的文献计量数据更为有效的数据源。CV数据的丰富性表现为它记载有年龄、性别、职业、专业方向、职位、成果、教育经历、专业经历、项目经历、教学活动、迁移、学术交流活动等各类信息,而且这些信息从其他数据源较难获得。CV数据的丰富性还表现在它一般按时间顺序来组织,可以形成一份较完备的纵向研究数据。对于关心时间序列关系和长时程效应的研究者来说,这是非常宝贵的。并且 CV日益成为一种公开可获取的数据资料,一项对 2 154 名研究人员的调查显示,有 61% 的受调查者有自己公开的 CV网页^[3],这大大增加了 CV数据的可获得性。同时,这还带来 CV的第三个优点,即它的非侵入性——无需打扰研究对象或获得研究对象的认可就可以获得所需数据,减少了因此造成的数据偏差,也降低了数据收集的时间成本和经济成本。

作者简介: 卫炯圻(ORCID:0000-0003-1228-3960),博士研究生;鲁晶晶(ORCID:0000-0003-2199-692X),博士研究生;谭宗颖(ORCID:0000-0003-3945-7174),研究员,通讯作者, E-mail: tanzy@mail.las.ac.cn。

收稿日期: 2015-02-15 **修回日期:** 2015-03-02 **本文起止页码:** 140-146 **本文责任编辑:** 王传清

2 CV 数据研究简史

国内的研究者通常援引国外文献,将 1996 年美国佐治亚大学的 RVM(research value mapping)项目看作是 CV 研究的“拓荒实验”与“先河”^[4-5]。RVM 项目由美国能源部和国家自然科学基金会资助,主要目的是在大学与研究机构的职能不再限于纯粹学术活动的背景下,采用新的社会学与经济学方法对公共资金资助的研究项目进行价值评估^[6]。但在 Web of Science 和 Google Scholar 上,以“curriculum vitae”和 CV 作为主题词进行粗检,可以找到很多在 1996 年之前就利用 CV 数据开展的研究工作。甚至,有些工作已经做得非常成熟。因此,不能将 RVM 项目看作是拓荒形式的工作。比如,1964 年,P. B. Price 等人就广泛地搜集职业医师的各种资料,其中就包括个人简历,用来评价医师的工作表现^[7]。1974 年,M. W. Rossiter 仅利用简历数据,就勾勒出美国 1920 年前男女科学家在职业生涯模式上的差异,并发现,科学界普遍存在的性别歧视对女性科学家来说,成为了很难克服的职业发展屏障^[8]。同样是针对科研群体的研究,1989 年,巴西的研究者利用巴西生物化学家对其他同行科学家 CV 所展示出的工作成就进行分数评定,以研究科学界的社会分层结构^[9]。

利用 CV 数据进行项目评估,则在 1988 年就已有了之。加拿大某大学开设了一个新型医学学士项目,参加该项目的学生在完成了两年的正式医学学习后,将在专人的指导下,从事一定形式的短期医学研究。对参与该项目的毕业生和控制组毕业生(没有参加该项目的)的 CV 数据进行对照,J. M. Gerrard 等人发现,即便是短期科研经历也会影响大学生的职业选择和人生发展,参与新项目的学生毕业后更多地选择了学术研究工作,并且其研究表现也好于控制组^[10]。1992 年,S. Bonzi 的研究工作已经非常接近当今研究者对 CV 数据的常用处理方法。他对某一所大学的 411 名研究人员的简历进行编码与分析,揭示出不同职称、性别和学科之间的差异,以及在职业生涯中各个时期的科研产出情况。该研究依据 CV 数据对样本人群进行分组,并进行组间比较。结果显示,产出与职称、学科和性别这 3 个因素都密切相关,但性别上的产出差异可能更多地是由职称和学科差异造成的。该研究还发现,职业发展早期的产出数量与后期的产出数量有 0.67 的显著相关,但随着产出数量的增加,篇均被引次数却逐步减少,早期与后期被引次数的相关只有 0.3^[11]。

虽然本文不太认可将 RVM 项目说成是拓荒性质的工作,但是依旧非常肯定 RVM 项目在 CV 数据的推广和应用上做出了前所未有的突出贡献。正是 RVM 项目组成员发表了一系列研究论文,才引发了情报研究者对 CV 数据的广泛关注,才真正带动了对 CV 数据的重视和应用。

3 CV 数据研究概述

受到 RVM 项目的推动,近 20 年间在图书情报领域,利用 CV 数据开展的研究工作非常多,本文从两个维度——宏微观维度和横纵向维度,将这些研究进行组织与分类,提出 CV 数据可用于解决的 4 类科学问题:科研人员的职业成长研究,人才流动与科研合作研究,科研群体的特征分析研究和科研项目、政策、建制的评估研究,如表 1 所示:

表 1 CV 数据的主要应用研究领域

维度	横向	纵向
宏观	科研群体的特征分析	科研项目、政策、建制的评估
微观	人才流动与科研合作	科研人员的职业成长

3.1 第一类研究:科研人员的职业成长

对微观个体进行纵向研究时,往往会用 CV 数据对科学家的个人成长和职业发展进行分析,以了解影响科学家职业生涯的各种要素,特别是某种经历对随后一段时间内个人科研活动的影响。通常,CV 里会记录一个研究者的各种个人经历(职业的和非职业的),这是其他数据源所不具备的独特资料,成为职业成长研究的重要数据来源。作为 RVM 项目的一部分工作,M. W. Lin 和 B. Bozeman 结合 CV 数据和问卷数据,分析了有全职的产业工作经历的科学家与始终在高校和科研机构供职的科学家在职业发展上有何不同。采用回归分析的统计方法,在控制了性别、年龄、职称与合作等因素后,他们发现有产业经历的科学家的论文总数量较少,但会培养更多的学生。而且,对于职称较低的年轻科学家和女性科学家而言,产业工作经验反倒会有助于增加年均产出数量^[12]。类似的研究还有 M. Gaughan 以美国国立卫生研究所(NIH)的 76 份 CV 数据为研究样本,尝试用 Cox 模型进行存活分析,来比较不同性质的研究组织中科学家学术生涯的差异^[13]。

CV 数据还可以用于研究科学训练中的不同经历对职业成长所构成的持久影响。比如,田瑞强等人对 ESI 中 233 名高被引的华人研究者的履历进行编码,研究了高层次人才的职业成长。采用医学上常用的存活分析,发现在中国获得博士学位和在外国获得博士学

位的研究人员在职业发展初期会形成两种不同的成长模式。同时,博士毕业后的职业选择也会对个人的学术成长有显著的影响,选择工作在研究所加速了职业成长,而博士后的经历却延缓了职业的成长^[14]。但博士后经历对职业成长的影响会因科学家所在国家的不同而有所差异。M. Gaughan 和 S. Robin 也采用 CV 数据对法国和美国各 400 余名物理学家和生物学家进行了比较,发现博士后研究经历对两国科学家的职业发展的影响不同。在美国,博士后研究经历并没有延缓获得永久研究职位的时间,而在法国,拥有一个博士后职位在一定程度上阻挠或延迟了其学术生涯的发展^[15]。

3.2 第二类研究:人才流动与科研合作

对微观层面的个人进行横向研究时,研究者通常利用 CV 所记录的教育经历、专业经历和访学经历来研究人才流动、科学知识的传播和科研合作等问题,用 CV 数据研究人才流动尤其热门。使用文献计量数据中的署名数据也可以研究人才流动^[16-17]。但署名数据远不如 CV 数据准确可靠,因为署名变化与人才流动之间并不存在必然的动态联系。因此,结合文献计量数据和 CV 数据来研究人才流动对科研活动的影响,日趋成为人才流动研究的一种常用数据方法。C. Cañibano 等人的研究便利用这样的数据组合,发现人才流动(体现为流动次数,比如短期居留次数、长期居留次数,呆过的研究中心和国家的数量等)与获得国际研究经费有一定的正相关性,但却与论文数量无关。特别是在分子生物学领域,Probit 概率分析和结构方程分析的结果显示,人才流动与科研产出之间的相关甚至是负的^[18]。不过,人才流动与科研活动之间的关系并没有这么简单明了。D. D. Filipp 等人以西班牙一所大学的教职员工 CV 数据和文献计量数据相结合,发现有国际间流动经历的研究者有更好的科研表现,比如更高的平均期刊影响因子、更高的篇均被引频次、更低的零被引论文比例和更高的国际合作论文比例,并且对所有职称的研究人员均是如此^[19]。但是,该研究仅仅将人才流动划分“有流动经历”和“没有流动经历”两类,且均为同一所大学的教职工,研究尚不深入,取样也存在偏差。尚不能准确认定是人才的流动导致了科研表现的优异,两者之间的关系有待更进一步的研究。

另有一些研究则利用 CV 数据关注国际间的短暂人才流动。C. Cañibano 等人将短暂流动界定为居留时间在 1 周到 2 年内的人才流动,通过对 10 349 名研究

者的 CV 进行分析,研究发现,越年轻的学者,有过短暂流动经历的可能性越大,有过流动经历的人越多。流动率也存在学科差异。在社会科学和人文科学中,人才流动更为频繁。研究揭示出的人才流动模式上的差异提示决策者,制定一套标准能对如此多样的流动都起到积极的激励作用的科技政策可能是不现实的^[20]。“人才回流”也是人才流动研究中的一个重要议题。K. Jonkers 和 R. Tijssen 利用 CV 数据分析中国生命科学领域的留学归国人员,并结合了案例研究,显示出归国科学家依旧保持了与原留学国科研人员的合作关系,这种归国流动有助于形成国际科学合作,实现流出国和流入国的双赢^[21]。但该研究取样数量少,只有 76 个人,回归分析的数据甚至只有 10 人。更为重要的是,与海外经历有关的变量与科研产出的关系并不明显,更为明显的相关是从事研究活动的时间,或获得博士学位后的时间与研究产出的关系。这种相关更多地可以归结为研究者的研究阅历或研究水平的提高,或者社会与人力资本的逐步积累所致,并不能完全将其解释为海外经历或流动对科研表现的影响。

还有一些研究者利用 CV 数据对不同研究群体的流动特征进行描绘。田瑞强等人对 ESI 中 233 名高被引的华人研究者的 CV 进行编码后,采用社会网络分析方法对他们攻读博士学位期间、博士后期间和工作期间这 3 个阶段所形成的流动网络进行描绘,发现高被引华人研究者的主流流动模式是一个以美国为绝对中心,并在英国、中国、新加坡和瑞士之间集中流动的形态^[22]。但考虑到该研究中,高被引的华人大多数在美国拿到博士学位,读博士后并工作,这样的结果实在并不让人意外。徐孝娟等人^[23]采用 CV 数据方法勾勒了国际计算机信息系统领域中研究人员的合作和人才流动情况。该研究的主要数据源依旧为 WoS 数据,仅部分地使用 CV 数据以作佐证。仅对该研究所定义的 302 名核心作者中发文排名前 10 的 10 份简历进行分析,并以 10 份数据就认为核心作者主要为教授,研究重点领域为生物化学信息技术、电子工程和化学信息学等,不能不说有些草率。在 U. Sandström 对瑞典 326 名高级医学科学家的流动特征进行描绘的研究中,也存在一些方法上的问题^[24]。她创造性地使用了聚类分析方法,将 326 名医学科学家分成了 4 组:高流动科学家(mobility)、低流动科学家(immobility)、杰出科学家(excellence)和企业型科学家(entrepreneur),其后的研究结论均依据这个分组而展开。但是,分组完全依据统计结果得出,而非来自理论或先验的命题。

高流动科学家不仅流动次数多,多数也在国外获得博士学位;杰出科学家不仅产出多,年龄也大。可见,分类的依据并不单纯,所以依此得出的结论无法给出可靠的解释。虽然使用了复杂的统计方法,但结论的可靠性可能还不如上文所提到的 S. Bonzi 在 1992 年的研究^[11]。用 CV 数据分析人才国际流动的研究还有 Y. Yamashit 和 D. Yoshinaga 的工作^[25]。他们对人工智能领域 2004-2006 年间的 392 名高被引论文的作者和 320 名零被引论文的作者的简历进行了分析,发现这两类人存在不同的国际间流动模式,高被引论文的作者往往由其他国家,特别是中国和印度向美国流动,而零被引论文的作者基本上是没有流动的本国学者。作者还试图分析高被引论文的作者群体和零被引论文作者群体的平均年龄以及归国人员对两类论文的贡献。但由于在选择群体的时候,作者错误地认为高被引论文的作者就是高水平作者,零被引论文的作者就是低水平作者,以单篇论文的被引情况来划分作者,导致在此基础上的任何结论都站不住脚,有些结果甚至无法解释,比如该研究发现,高被引作者的学术年龄为 15.7 年,而零被引论文作者群的平均学术年龄却达到 18.1 年。

3.3 第三类研究:科研群体的特征分析

对群体的横向研究中, CV 数据被用来勾勒某一群体,如杰出科学家、女性科学家或某个学科科研群体的特征。我国研究者尤其关心高层次科技人才这个群体。比如,用长江学者、百人计划研究员和杰出青年基金获得者的 CV 数据来分析他们的主要特征,通过描述统计,研究者发现,在这些高层次科技人才中,女性较少,上述 3 类人才中,女性科学家的比例均为个位数,长江学者中女性入选者甚至只有 4%,而同期的《中国科技人力资源发展研究报告》显示,女性科学家的比例有 37%。这两个比例数据形成了悬殊的差异。高层次科技人才的其他特征还有:入选年龄呈逐年上升趋势,群体成员几乎都有海外科研经历,尤其是美国的研究经历等^[26]。这些群体特征的描绘对制定科技人才政策有非常重要的参考价值。类似的研究还有贾佳等人的“中国高被引科学家基本特征研究——以 2014 汤森路透高被引科学家为例”^[27]、鲍雪莹等人的“基于履历信息的国际科技人才特征分析——以近十年诺贝尔物理、化学、生理或医学奖得主为例”^[28]等等。虽然不同的研究对高层次科学家的界定不同,但基本上都属于一类研究。

男性科学家和女性科学家是两类特征非常明显的

科研群体。M. Sabatier 等人的工作关注男性科学家和女性科学家在获得个人职业成就(在该研究中主要指晋升)过程中的差异^[29]。CV 简历在这里就体现出了明显的丰富性,它不仅仅提供了论文列表,还包括研究项目、咨询活动、教学活动和管理活动等各种信息。结果显示,对两性科学家的差别对待是持续一贯的,并且对女性科学家采用了与男性不同的晋升标准。女性科学家需要在各种研究活动,比如教学、研究、管理等过程中,都展示出更多的职业投入才有可能获得晋升。还有一些学者实验性地用 CV 数据来勾勒学术共同体的认知、社会与建制性的复杂结构。比如, B. Lepori 等人对瑞士传播学专家的 76 份 CV 研究发现,瑞士传播学专家的受教育背景主要来自 13 个学科,而出乎意料的是,有传播学教育背景的人只有 32%。继续结合成果列表和其他数据源,从建制/社会的维度和认知/主题的维度,便可以画出该领域的结构地图,区分出 5 个研究集群和他们相应的研究主题,并可以显示瑞士传播学与比邻国家的相互关系^[2]。相比于以上描述群体特征和领域地图的研究来说, E. Leahey 的研究对 CV 数据的使用则更为深入一些。他没有停留在对群体的 CV 数据进行简单编码和描述的层面上,而是结合文献计量数据,构造出新的概念——专业化程度(extent of research specialization),用以表示一名研究者对某一个研究领域的专注程度,并利用结构方程模型分析发现,女性科学家与男性科学家的产出差异是由她们在专业化程度上比男性科学家较差所致^[30]。

3.4 第四类研究:科研项目、政策、建制的评估

当 CV 数据用于群体的纵向研究时,可以评估某一个或某一类科研项目、科技政策和建制的效果。广受关注的 RVM 项目,就是设计利用 CV 数据,并结合问卷数据来评估美国科研资助制度和研究组织形式的变化对科学活动的影响^[11]。作为该项目的研究成果之一, M. Ganghan 和 B. Bozeman 用 CV 信息来比较两种研究资助形式对科研活动的影响:一个是申请资助模式;另一个是研究中心资助模式。对 1 041 名获得博士学位的科学家和工程师的 CV 数据进行分析,确定两种不同的资助模式至少可以对科学家的应用产出与学术产出的模式构成一定的影响。研究中心资助的情况下,科学家更有可能获得更多的应用项目资助。M. Ganghan 和 B. Bozeman 不单单对两种资助模式下的科学家简历进行了比照,还试图使用多元回归分析的统计方法来研究资助模式与论文产出之间的因果联系^[31]。C. Cañibano 等人则尝试利用 CV 数据来评估基

金申请中同行评价制度的效能^[32]。为了验证 CV 信息是否可以用于同行评议,他随机选取了 3 个学科(物理与空间科学、分析生物学和哲学文献学)的 266 个样本,其中有 36 人最终获得了基金资助。对这些研究者的 CV 数据,包括性别、年龄、获得博士学位的时间、流动和研究成果等进行编码,采用判别分析和逻辑回归分析两种统计方法研究发现,同行评议的过程主要看重申请者的论文数量,这表示该系统的评议主要依据的还是研究的杰出程度,而不是其他,比如性别、年龄等,也就是说,并不存在同行评议过程中的性别歧视。

4 CV 数据的缺陷

没有什么数据是完美的, CV 数据也是一样。国内的研究者对此已经有过一些分析。田瑞强等人指出, CV 数据的 4 个主要缺陷为:可获得性差、数据异质性强、信息缺省较多和编码困难^[4]。对此,周建中等人也持有类似的观点^[5]。不过,在这 4 个主要缺陷中,可获得性、异质性和编码困难问题都是相对的,将其看作是缺点还是优点,取决于参照的对象和对 CV 数据的基本态度。对于无法访问 Web of Sciences 文摘索引数据库的人来说,网络公开的 CV 数据比 Web of Sciences 提供的文献计量数据要更容易获取。J. S. Dietz 等人的论文就表示, CV 数据“在一定程度上是标准化的,且相对容易获取”^[1]。至于编码问题,在大多数质性研究中,对素材的编码,例如访谈文本和现场观察笔记等,远比 CV 编码要复杂得多。质性研究创制的众多编码方法、程序与工具都可被 CV 数据编码借鉴。因此,真正值得注意的是信息缺省带来的系统性偏差问题。并非所有人的科研履历都可以获得,这使得研究取样无法随机,研究结论也只适用于公布了研究履历的科学家,这会造成一定的系统偏差。而科研人员又可能不会公布他们所有的信息,只会选择性地列出部分研究成果,这又会造成一定的系统偏差。因此,对选用 CV 数据的研究结果的解释要谨慎。只有了解这些偏差,才能对研究结论有更恰当的理解。不过,这些系统偏差是否就否定了 CV 数据的价值呢?答案是否定的。系统偏差同样存在于实验数据、调查数据和 Web of Sciences 的文献计量数据中,但是这不影响我们广泛使用它们作为文献计量和情报研究的基础数据。

5 CV 研究展望

目前,无论是在国内还是国外,在采用 CV 数据作为主要数据源的研究中,样本数据量均比较小,取样往

往局限于单个国家的单个机构或单个项目的数据(多为方便取样),且数据被多次重复使用。于是,不少研究结论存在局限性,或仅能视之为利用 CV 数据进行情报研究的一种方法学上的尝试。数据难以大量获取,成为制约开展大规模研究的瓶颈。建立专家学者库一定程度上有助于解决这个问题。一些国家,比如西班牙和巴西等,已经着手建立本国的学者专家库^[17,33],一些机构也建立起自己的专家库,比如香港大学机构知识库的专家页(<http://hub.hku.hk/>)和中国科学院的 iAuthor 平台(<http://iauthor.cn/>)等。对专家学者库的结构化 CV 数据,进行数据收割和二次加工都比较容易,可以形成较大规模的数据集,这将有助于 CV 数据相关研究的开展,增加研究数量,拓展研究命题。不过,目前多数专家库依旧比较侧重对公开的知识成果进行登记、展示与传播,缺少对专家个人元数据的规范和描述,信息量还不够大。

当前研究中存在的另外一个问题是 CV 数据缺少与其他数据的深度关联。对 CV 数据的研究与分析在理论上代表了“产量范式”向“能力范式”的转移^[1],从关注研究活动的定量产出,转变为强调评估长期的知识产出能力^[22],但在实际研究中,常见的数据关联模式是结合 CV 数据与文献计量数据,依旧将研究人员视为一组研究成果的合集,就像将一本刊物看作是其所有刊文的合集一样。这种研究取向缺乏对人的能动性的考虑,采取的依旧是产量评价的范式。其结果是,连相关关系所反映的实质内容也难以说清,更别说去了解复杂的因果关系了。举例来说, C. Gaughan 等试图发现两种资助模式的差别,但是通过对 CV 数据的回归分析,了解到两类资助模式所资助的科学家的科研活动是不同的,却无法准确地将这种差异解释为由资助模式的差别所致^[18]。

因此,笔者建议将 CV 数据看作是研究人员在职业生涯中个人资本逐步积累过程的一个记录,以时间为主要维度组织起来的 CV 数据可以作为枢纽,与其他类型的过程数据关联起来,从而构成一条条数据链,来支撑更为深层次的研究发现。以“高科技人才流动”研究为例,在宏观和中观层面, CV 数据可以关联机构的经费投入、整体科研实力、研究声望、建制形态、组织模型、文化环境等数据;在微观层面, CV 数据可以关联起研究人员对流入机构的态度及其变化、流动动机、对环境的适应与调节行为等数据,这样通过 CV 数据结合实际科研活动的深入观察和调查,才能将研究活动还原为能用性的个体在特定物理环境、社会环境、文

化环境和制度环境下的知识探索过程。

参考文献:

- [1] Dietz J S, Chompalov I, Bozeman B, et al. Using the curriculum vita to study the career paths of scientists and engineers: An exploratory assessment[J]. *Scientometrics*, 2000, 49(3): 419-442.
- [2] Lepori B, Probst C. Using curricula vitae for mapping scientific fields: A small-scale experience for Swiss communication sciences [J]. *Research Evaluation*, 2009, 18(2): 125-134.
- [3] Kousha K, Thelwall M. Disseminating research with Web CV hyperlinks[J]. *Journal of the Association for Information Science and Technology*, 2014, 65(8): 1615-1626.
- [4] 田瑞强,姚长青,袁军鹏,等.基于科研履历的科技人才流动研究进展[J]. *图书与情报*, 2013(5): 119-125.
- [5] 周建中,肖小溪.科技人才政策研究中应用 CV 方法的综述与启示[J]. *科学学与科学技术管理*, 2011, 32(2): 151-156.
- [6] Dietz J S, Bozeman B. Academic careers, patents, and productivity: Industry experience as scientific and technical human capital [J]. *Research Policy*, 2005, 34(3): 349-367.
- [7] Price P B, Taylor C W, Richards Jr J M, et al. Measurement of physician performance [J]. *Academic Medicine*, 1964, 39(2): 203-211.
- [8] Rossiter M W. Women scientists in America before 1920[J]. *American Scientist*, 1974, 62(3): 312-323.
- [9] De Meis L, Longo P H, Falcão E. The learning process in science: A study among Brazilian biochemists[J]. *Biochemical Education*, 1989, 17(3): 127-132.
- [10] Gerrard J M, Fish I, Tate R, et al. Evaluation of the careers of graduates of the University of Manitoba's BSc (Medicine) program [J]. *Canadian Medical Association Journal*, 1988, 139(11): 1063-1068.
- [11] Bonzi S. Trends in research productivity among senior faculty[J]. *Information Processing & Management*, 1992, 28(1): 111-120.
- [12] Lin Minwei, Bozeman B. Researchers' industry experience and productivity in university-industry research centers: A scientific and technical human capital explanation[J]. *Journal of Technology Transfer*, 2006, 31(2): 269-290.
- [13] Gaughan M. Using the curriculum vitae for policy research: An evaluation of National Institutes of Health center and training support on career trajectories [J]. *Research Evaluation*, 2009, 18(2): 117-124.
- [14] 田瑞强,姚长青,袁军鹏,等.基于履历信息的海外华人高层次人才成长研究:生存风险视角[J]. *中国软科学*, 2013(10): 59-67
- [15] Gaughan M, Robin S. National science training policy and early scientific careers in France and the United States [J]. *Research Policy*, 2004, 33(4): 569-581.
- [16] Laudel G. Studying the brain drain: Can bibliometric methods help? [J]. *Scientometrics*, 2003, 57(2): 215-237.
- [17] Halevi G, Moed H. International scientific migration and collaboration patterns following a bibliometrics line of investigation[J]. arXiv preprint arXiv:1212.5194
- [18] Cañibano C, Otamendi J, Andújar I. Measuring and assessing researcher mobility from CV analysis: The case of the Ramón y Cajal programme in Spain[J]. *Research Evaluation*, 2008, 17(1): 17-31.
- [19] Filippo D D, Casado E S, Gómez I. Quantitative and qualitative approaches to the study of mobility and scientific performance: A case study of a Spanish university [J]. *Research Evaluation*, 2009, 18(3): 191-200.
- [20] Cañibano C, Otamendi F, Solís F. International temporary mobility of researchers: A cross-discipline study[J]. *Scientometrics*, 2011, 89(2): 653-675.
- [21] Jonkers K, Tijssen R. Chinese researchers returning home: Impacts of international mobility on research collaboration and scientific productivity[J]. *Scientometrics*, 2008, 77(2): 309-333.
- [22] 田瑞强,姚长青,潘云涛,等.基于履历数据的海外华人高层次科技人才流动研究: 社会网络分析视角[J]. *图书情报工作*, 2014, 58(19): 92-99.
- [23] 徐孝娟,朱庆华,彭希羨,等.基于社会网络及履历分析的国际科技人才合作模式研究——以信息系统领域为例[J]. *现代情报*, 2014, 34(9): 16-23.
- [24] Sandström U. Combining curriculum vitae and bibliometric analysis: mobility, gender and research performance [J]. *Research Evaluation*, 2009, 18(2): 135-142.
- [25] Yamashita Y, Yoshinaga D. Influence of researchers' international mobilities on publication: A comparison of highly cited and uncited papers [J]. *Scientometrics*, 2014, 101(2): 1475-1489.
- [26] 牛翥,周建中.基于 CV 分析方法对中国高层次科技人才的特征研究[J]. *北京科技大学学报: 社会科学版*, 2012, 28(2): 96-102.
- [27] 贾佳,潘云涛,马峥.中国高被引科学家基本特征研究——以 2014 汤森路透高被引科学家为例[J]. *科技与出版*, 2014, (12): 149-151.
- [28] 鲍雪莹,陈贡,刘木林.基于履历信息的国际科技人才特征分析——以近十年诺贝尔物理、化学、生理或医学奖得主为例 [J]. *现代情报*, 2014, 34(9): 4-9.
- [29] Sabatier M, Carrere M, Mangematin V. Profiles of academic activities and careers: Does gender matter? An analysis based on French life scientists' CVs [J]. *Journal of Technology Transfer*, 2006, 31(3): 311-324.
- [30] Leahey E. Gender differences in productivity: Research specialization as a missing link [J]. *Gender and Society*, 2006, 20(6): 754-780.
- [31] Gaughan M, Bozeman B. Using curriculum vitae to compare some impacts of NSF research grants with research center funding [J]. *Research Evaluation*, 2002, 11(1): 17-26.
- [32] Cañibano C, Otamendi J, Andújar I. An assessment of selection processes among candidates for public research grants: The case of

the Ramón y Cajal Programme in Spain[J]. Research Evaluation, 2009, 18(2):153-161.

- [33] D' Onofrio M. The public CV database of Argentine researchers and the "CV-minimum" Latin-American model of standardization of CV information for R&D evaluation and policy-making[J]. Research Evaluation, 2009, 18(2):95-103.

作者贡献说明:

卫垆圻:选题确定、资料收集、初稿撰写与修订;

鲁晶晶:资料收集、初稿审订;

谭宗颖:选题确定、初稿审订和最终定稿。

Research Progress on Scientists' Curriculum Vitae Data

Wei Tongqi^{1,2,3} Lu Jingjing^{1,2} Tan Zongying¹

¹National Science Library, Chinese Academy of Sciences, Beijing 100190

²University of Chinese Academy of Sciences, Beijing, 100049

³Institute of Psychology, Chinese Academy of Sciences, Beijing 100101

Abstract: [**Purpose/significance**] The researchers are realizing the significance of scientists' curriculum vitae (CV) for its richness and uniqueness. The information researchers and analysts were able to make many scientific findings in various domains by exploring the scientist' CV data. CV is becoming as the important data source of information research and analysis. [**Method/process**] The current study reviewed the academic studies using CV data as their main materials during past two decades. [**Result/conclusion**] We found that there are at least four research domains suitable for CV analysis: scientist' career growth, scientific mobility and cooperation, the characteristic analysis of research group, and the evolution of program, policy or institution. Furthermore, CV could act as a data hub linking more various datasets, and improve the information research and analysis.

Keywords: researcher CV information analysis CV dataset career growth talent mobility program evolution

(上接第 139 页)

- [26] Mckinsey. Periscope [EB/OL]. [2014-12-26]. <http://solutions.mckinsey.com/periscope>.

- [27] Mckinsey. Clickfox [EB/OL]. [2015-01-23]. <http://solutions.mckinsey.com/index/IE8/solutions/clickfox.html>.

- [28] Mckinsey. Finalta [EB/OL]. [2015-01-23]. <http://www.finalta.eu/home/home.asp>.

- [29] Mckinsey. spotlight [EB/OL]. [2015-01-25]. <http://solutions.mckinsey.com/index/IE8/solutions/spotlight.html>.

作者贡献说明:

张军:负责研究选题、框架与方法设计及论文撰写;

周磊:参与智库调研与论文撰写;

慕慧鸽:参与智库调研与论文撰写。

Development and Trends of Quantitative Research Methods on

International Authoritative Think-tanks

Zhang Jun¹ Zhou Lei¹ Mu Huige^{1,2}

¹ Wuhan Documentaion and Information Center of the Chinese Academy of Sciences, Wuhan 430071

² University of Chinese Academy of Sciences, Beijing 100049

Abstract: [**Purpose/significance**] Think-tanks influence government decision-making through high level and high quality research and consulting activities. Innovative quantitative methodology can provide reliable basis for think-tank's research and enhance authority and credibility. Exploring the characteristics and trends of quantitative research methods used by international authoritative think-tanks, can provide some reference for Chinese think-tanks. [**Method/process**] By investigating quantitative research methods used in selected Think-tanks' research reports in recent years, the authors summarized these methods' characteristics and trends, then proposed some suggestions to Chinese think-tanks in strategic and practice aspect. [**Result/conclusion**] The results shows that modern think-tank emphasis comprehensive field view, interdisciplinary research, automated and intelligent tools and full chain procedure.

Keywords: think-tank quantitative analysis methods quantitative analysis tools