

基于科技信息大数据的科技人才科研综合能力评价及应用研究^{*}

王运红 潘云涛 赵筱媛

(中国科学技术信息研究所 北京 100038)

〔摘要〕 在对科技人才科研综合能力分析的基础上,构建科技人才科研综合能力评价模型,设计基于科技信息大数据的评价指标体系,包括基本素养、科研产出影响力、科研管理能力、学术潜力和学术地位 5 个维度,研究其在人才引进工作中的应用。

〔关键词〕 多维评价模型;科研综合能力;评价指标

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2017.12.002

Evaluation on Comprehensive Scientific Research Capability of Technology Talents and Application Study Based on Big Data of Scientific and Technological Information WANG Yun-hong, PAN Yun-tao, ZHAO Xiao-yuan, Institute of Scientific and Technology Information of China, Beijing 100038, China

〔Abstract〕 Based on the analysis of comprehensive scientific research capability of technology talents, the paper constructs the evaluation model of comprehensive scientific research capability of technology talents, designs the evaluating indicator system based on big data of scientific and technological information, including the 5 dimensions like basic literacy, influence of scientific research output, scientific research management capability, academic potential and academic status, and studies on its application in talent introduction.

〔Keywords〕 The multi-dimensional evaluation model; Comprehensive scientific research ability; Evaluation index

1 引言

科学、合理的科技人才评价体系是人才工作的

“指挥棒”,不仅影响着科技人才的质量和水平,同时影响科技人才引进、管理、培养和使用等。中共中央印发的《关于深化人才发展体制机制改革的意见》(以下简称“意见”)指出,当前人才评价领域中存在评价结果难以反映人才整体研究能力和水平等问题。目前人才评价存在的问题主要表现为:一是人才评价的导向简单化,过于强调论文、课题等显性指标,忽视了人才研究综合能力、学术潜力等;二是人才评价标准单一化,对不同领域、门类和层级的人才都使用简单的量化指标衡量;三是评价标准针对性不强,没有建立分层、分类的科学评价体系。针对当前我国人才评价中的突出问题,建立一套能够反映科技人才的科研综合能力,可以实

〔收稿日期〕 2017-11-24

〔作者简介〕 王运红,高级工程师,发表论文 11 篇。

〔基金项目〕 科技创新战略研究专项项目“基于全球视野的杰出科技人才资源发现、培养和使用机制研究”(项目编号:ZLY201628);国家科技支撑计划课题“面向科技情报分析的信息服务平台研发与应用示范”(项目编号:2015BAH25F02)。

际操作应用,指导人才工作的评价体系具有重要价值。

在科技人才评价模型和指标的研究方面,研究人员多数提出品德、知识、能力、业绩作为衡量人才的主要指标,但是细化指标时,很多指标不具备操作性,如能力中有沟通能力、思维能力、预见能力,业绩有经济效益、社会效益;模型构建和指标设计分层不清晰,指标值不便于获取和比较,因此在对科技人才的科研综合能力评价上无法深入和实现^[1-4]。本研究在对科技人才科研综合能力进行分析的基础上,构建科技人才的科研综合能力评价模型,设计基于科技信息大数据的科研综合能力评价指标体系,研究科研综合能力评价模型在人才引进、培养、发现等工作中的应用。本研究设计的科研综合能力评价指标有基本素养、科研产出影响力、科研管理能力、学术潜力和学术地位5个维度,设计了3级指标体系,定量指标以科技信息大数据为基础获得客观数据,定性指标辅助并设计量化值以便比较,进而全面、动态、科学地反映科技人才的科研综合能力。

2 科研综合能力及科技信息内涵

2.1 科技人才科研综合能力

能力是直接影响活动效率,使活动顺利完成的个性心理特征。能力又分一般能力和特殊能力,一般能力是在很多基本活动中表现出来的能力,适用于广泛的活动范围。特殊能力是表现在某些专业活动中的能力,只适用于某种狭窄的活动范围。科研能力就是在科研活动中表现出来的特殊能力,广义来讲是个很宽泛的概念,指一个人在其所从事的专业中以科学的思维和适当的方法对未知领域进行科学探索的能力,反映其发现问题、认识问题和解决问题的能力,综合体现一个人专业知识的深度和广度。科技人才的科研能力有很多方面,从解决问题的角度,从问题发现和问题解决能力、资料搜集和处理能力、逻辑思维能力、动手能力、创新能力等角度进行研究。从能力本身结构角度,将科研人才的能力分为一般能力、特殊能力和创造能力^[5]。本

研究中科技人才的科研综合能力不仅是泛指的能力,而是从科技人才从事的科研活动以及产生的科研成果和学术活动来定义其科研综合能力,包括对基本素养的定义、科研产出成果水平、科研管理能力、同行评议的学术地位以及学术活动中表现出来的学术潜力,可以较全面地评价科技人才的研究基础、研究能力、研究水平和发展潜力,为科技人才评价、培养、引进等工作提供客观依据。

2.2 科技信息

信息的定义非常广泛,科技信息的定义则在信息的定义基础上,限定在科学技术领域,在科学研究与技术开发、应用过程中产生的各种信息资源,以及科技活动过程中所需要的各种类型的信息,当这些信息汇集时就形成科技信息资源^[6]。科技信息是一种战略资源,反映科研活动的全过程,从科研设想、科研实施、成果发布和成果转化利用的各个环节都有科技信息的产生,科技信息资源都发挥了重要的价值,同时在科研活动的过程中也不断有新的科技信息资源产生^[7]。国家科技发展中长期规划中的“科技信息”是指“人类社会科技活动所产生的基本科学技术数据、资料,以及面向不同需求加工整理形成的各种科学数据产品和各种载体的科技图书、期刊、报告、论文、专利等科技文献”^[8]。本研究中的科技信息大数据不仅数据量大,而且信息种类和来源渠道多。从更广义的角度认识科技信息资源,认为科技信息资源是指从事科技活动的人力、物力、财力以及组织、管理、信息等硬、软件要素进行信息化后的资源总称,包含科技人力信息资源、科技物力信息资源、科技财力信息资源、科技制度和组织信息资源、科技成果产出信息等。本研究的评价模型以科技信息大数据为来源,主要指科技人才基本信息以及在科技活动中产生和形成的科研成果产出信息资源,包括科技人才基本信息、科技文献信息和科技项目信息和科技机构信息等。其中科技文献信息资源范围最广,包括图书、期刊论文、会议论文、学位论文、科技报告、科技成果、专利和标准等。围绕科技人才的这些数据反映科技人才不同于其他人才的个性特征,本研究称之为

为科技人才的全景信息数据，见图 1。

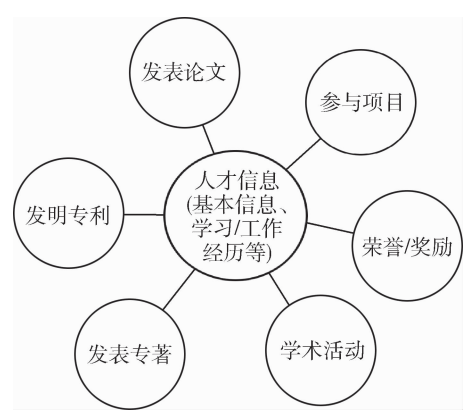


图 1 科技人才全景信息数据

3 科研综合能力评价模型

目前国内外科技人才评价的基本原理主要包括胜任力模型理论、个体差异原理、思维的模糊性原理、事物的可测性原理及测评一致原理等，把行为科学、心理科学、管理学、计量学和人才测度理论等作为进行科技人才评价的理论基础”^[9]。虽然各种评价模型和指标体系从不同的角度对科技人才的评价进行了有益探索，目前将综合评价模型与指标关联，定性指标与定量指标结合，落实到实际操作层面，支撑实际工作评价模型的并不多见，而兼顾考察人才学术潜力的可量化评价更少。本研究构建的科技人才科研综合能力模型包括基本素养、科研产出影响力、科研管理能力、学术潜力和学术地位 5 个要素，每个要素都有测度的系列指标，面向不同类型的科技人才，构建出科技人才的科研综合能力评价模型，见图 2。基本素养评价的是科技人才的道德信誉，主要考察科技人才的材料真实性，履历是否造假或隐瞒，科研产出是否造假、虚报。科研产出影响力主要从学术论文、发明专利等产出数量和质量来考评，通过论文和专利的被引情况、转化情况考察科研产出在同行中的影响力，进而判断

其学术水平。科研管理能力从其参与或者主持的科研项目数量、连续性、承担的角色，分析判断其科研项目的实施能力。学术潜力从科研产出的趋势、研究延续性、学术会议参加情况并加以年龄段分析，可看出其学术活跃度和未来学术能力发展情况。学术地位主要评价科技人才在其研究领域的学术地位和认可度，主要数据来源是获得的科技奖励、荣誉称号和学术任职情况。

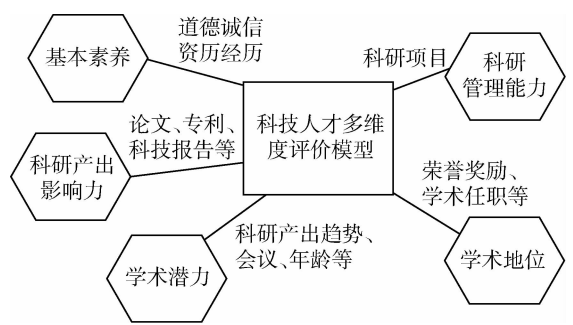


图 2 科技人才科研综合能力评价模型

4 科研综合能力评价指标体系

4.1 指标设计

从文献调研的人才评价指标体系结构设计来看，基本都沿用基本素质、学术水平和成果产出的评价体系结构，但是指标体系的设计必须考虑实际应用需求，将指标体系使用起来，为实际工作提供参考和指导。本研究设计的评价指标体系，见表 1。其面向人才评价工作的实际应用，把非量化指标如行为特征、个人素质、创新能力难以测度的指标去掉，或用便于获取和考量的指标；规避量化指标过分强调科研产出数量的问题，更加注重科研产出的质量，同时增加对科技人才的学术潜力、学术地位的考察。本指标针对不同类型科技人才，面向不同的应用目标和对象以及人才发展的不同阶段，在应用中灵活调整指标的权重，更便于实际操作。

表 1 科研综合能力评价指标

1 级指标	2 级指标	3 级指标	说明
基本素养	道德诚信	履历真实性	F – 有造假
		成果真实性	T – 未检出诚信问题
			U – 不确定（包括有举报未核实）
	资历经历	学历学位	学历学位分级并赋值
		职称	职称分级并赋值
		工作单位	对工作单位按照机构排名分级赋值
科研产出影响力	论文产出影响力	论文数量	论文数量分区赋值
		高被引论文比例	第 1 作者或通讯作者的高被引文章占总高被引文章的比例
		论文相对影响力	论文的被引频次与该领域论文的篇均被引频次的比值分区赋值
	专利产出影响力	发明专利数量	专利分不同来源的数量分区赋值
		发明专利相对影响力	专利被引频次与本技术领域篇均被引频次比值分区赋值
科研管理能力	项目实践	项目级别与数量	项目分级参与数量分区赋值
		项目管理	不同项目，不同角色进行权重赋值
		团队规模	团队数量、层次分级
学术潜力	论文产出趋势	国际、国内、会议论文	选取近 10 年论文发表情况做趋势分析，反映论文产出的上升、下降、平稳趋势
	专利产出趋势	国际、国内专利	选取近 10 年发明专利情况做趋势分析，反映专利产出的上升、下降、平稳趋势。不同专利做不同权重赋值
学术地位	学术交流	学术会议	参加会议情况，会议发言或者论文发表
	荣誉奖励	科技荣誉数量	针对不同科技荣誉分级赋值
		科技奖励数量	不同级别奖励分级赋值
	学术任职	科研团体中职务	对职务分级赋值
		国内外期刊编委	对期刊分级赋值
	同行评议	同行认可度	定性指标，专家邀请

注：为科研产出影响力 3 级指标数据便于获取和计算，在科研产出影响力指标中，简化 3 级指标，只保留论文和专利影响力，科技报告、标准、著作、科技成果等产出的影响力暂不作为评价指标。科技报告、标准、著作和科技成果的影响力数据获取方式和评价方法，将作为进一步研究的内容之一。

4.2 系统构建流程

基于科技信息大数据的科研综合能力评价指标系统流程，见图 3。基础数据采集是在海量信息中对科技信息数据的采集、清洗、分类等工作，将在科技人才评价中用到的科技人才工作经历、学习经历、诚信信息、荣誉奖励等基本信息和科研成果产出信息，包括论文、专利以及参与项目情况等数据全部采集入库。指标计算是通过专家调查法得到不同类型科技人才的指标权重，经大数据分析、科学计量学、

聚类算法和趋势分析，对每个 3 级指标进行计算赋值；定量指标和定性指标计算获得不同的赋值后，对指标进行拟合，最后得出综合评价结果。



图 3 科研综合能力评价指标系统构建流程

5 科研综合能力评价在引才工作中的应用示例

5.1 概述

本研究已经为某机构引才进行实际评价，以生

命科学领域的两个被评人为例，介绍科技人才科研综合能力评价指标系统的应用对引才工作的支撑作用。被评价人的基本情况，见表 2。

表 2 被评价人基本情况

评价指标	被评人 A	被评人 B
学位	Doctor（博士）	Doctor（博士）
毕业院校	国内 985（本科）、新加坡国立大学（硕士）、 哈佛大学（博士）	国内 985（本科）、北京大学（硕士）、北京大学 （硕士）、加州大学交流学者
研究领域	Life Sciences and Medicine 生命科学和药物	Life Sciences and Medicine 生命科学和药物
研究方向	Cell Biology（细胞生物学）、Biochemistry（化学 生物学）、Molecular Biology（分子生物学）	Cell Biology（细胞生物学）、Biochemistry（化学 生物学）
工作机构	机构 A（QS 排名前 10）	机构 B（QS 排名 25 – 30）
出生年份	1985 年	1965 年
机构任职	Research Officer（研究人员）	Professor（教授）

5.2 评价目标

为某机构的人才计划引进具有国际工作经验的青年学科带头人，要求具有博士以上学位，被国际同行高度认可，具有较大影响力的学术产出成果，良好的学术水平和发展潜力，带领本研究机构在 5 年内具备国际领先的学术水平，产出具有国际影响力的成果，引导产业的发展。针对不同的评价目标，可以对指标体系中指标的权重、定性指标评价标准、定量指标计算和赋值情况进行调整。

5.3 评价指标赋值

5.3.1 概述 本评价指标是定性指标和定量指标的融合，定量指标的数据可获得性、科学性和客观性才是能够在实际工作中应用的基础。几个重要指标的赋值说明如下。

5.3.2 基本素养 （1）道德诚信从科研诚信（包括成果是否抄袭、造假等）、品质道德（是否有品质道德的问题，包括伪造履历等）两个角度衡量，取值及含义：F – 有造假和诚信问题；T – 未检出诚信问题。两者必须同时为 T，诚信度取值 T，否则为 F。该项指标在本示例中是 1 票否决指标，如

检测出道德诚信有问题，后面的评价指标再高，不再考虑被评人的入选资质。（2）资历经历对被评价人的学历学位、专业职称、工作单位的影响力等角度测量，可以化定性指标为定量指标，考察被评价人作为引进人才的基本条件和素养是否达标。

5.3.3 科研产出影响力 科学计量学在科研评价实践中的应用越来越广，在传统评价指标如论文量、论文总被引频次、论文篇均被引频次、期刊影响因子及其相关指标等基础上，基于引证分析的学术影响力评价指标（h 指数、g 指数、A 指数及 R 指数等）不断趋于合理、完善^[10]。现有指标均存在着不同程度的缺陷，如无法同时兼顾论文数量和质量、忽视定性评价的作用等，因此国内外的学者对科学计量学指标在人才评价中的应用进行不同程度的改进^[11]。本指标从数量、质量、相对影响力 3 个角度对人才进行评价，指标更针对人才本身，以及产出的质量考量。（1）论文产出影响力。论文反映科技人才的基础研究能力，论文产出的影响力反映论文的质量。本示例中论文来源是 SCI 核心合集、EI、CPCI 的论文数量之和。高被引论文比例是选择该作者发表所有论文里被引次数最高 1 篇论文为代表作，本研究选取基本科学指标（Essential Science

Indicators, ESI) 中各领域的高被引论文为基准来评估该论文在领域内的影响力。百分位数值是某个研究领域里一部分论文被引次数的最低被引次数。百分位数值越小, 最低被引次数值越高。百分位数值是用来评估论文影响力的基准。如位于前 1% 的论文最低被引次数为 15, 那么这个领域同年发表论文的被引次数至少要高于 15 次才归属于前 1%。ESI 对高被引论文的百分位分区为: 0.01%、0.10%、1.00%、10.00%、20.00% 和 50.00%, 影响力依次降低。论文相对影响力是论文影响力用 1 篇论文的被引频次与该领域总体论文的平均被引频次的比值表示。公式可以表示为:

相对影响力 = $\frac{1 \text{ 篇论文的被引频次}}{\text{该领域论文的篇均被引频次}}$

(2) 专利产出影响力。专利考察一个人的技术创新和研发能力。本示例的专利数据来源于中国国家知识产权局的专利数据库和德温特专利数据库。作为引才工作主要考察国际发明专利、专利被引和转化情况。

5.3.4 科研管理能力 (1) 项目实践。将项目类别分为省部级项目、国家级项目、国际合作项目和其他项目, 按照参与的级别对项目分类赋权重。(2) 项目管理。将承担角色分为负责人、项目骨干、参与人 3 个级别。评价同时与项目类别对应考量。

5.3.5 学术潜力 (1) 论文产出趋势。将 SCI、

EI、CPCI 近 10 年论文的数量按年度分布做出趋势图, 按照“上升”、“平稳”、“下降”给予不同的分值。(2) 专利产出趋势。将中国发明专利和德温特专利库中的国际专利数量按年度分布做出趋势图, 按照“上升”、“平稳”、“下降”给予不同的分值。(3) 学术交流从会议论文看参加学术会议情况, 评价其学术活跃度。本示例从国内会议、国际会议两个角度衡量学术交流的活跃性。

5.3.6 学术地位 (1) 荣誉奖励。将荣誉和奖励分为 6 个类别, 分别是学协会科技奖励/荣誉、国内地市级奖励/荣誉、国内省部级、国家级、国际级及顶尖国际奖励/荣誉。根据不同的奖励和荣誉, 结合考虑评价目标中对被评人的资质要求, 对不同的荣誉奖励分级赋值。本评价指标系统中, 如图灵奖、诺贝尔物理奖等顶尖国际奖励/荣誉获得者 1 票入选。(2) 学术任职。将学术任职分为普通学术职务、重要学术职务, 不同职务可以分区赋值。(3) 同行评议。本指标是定性指标, 为补充参考, 通过两两比较后得出同行评议的模糊值输出。

5.4 评价结果综述示例

通过对两位参评人的各项指标计算和综合分析后, 得出综合评述结果, 提供给引才机构, 从实际工作需求出发来决定入选人。被评价人评价结果, 见表 3。

表 3 被评价人评价结果

评价指标	被评人 A	被评人 B
基本素养	道德诚信没有问题; 学术研究基础非常好	科研诚信没有问题, 学术研究基础非常好
科研产出影响力	只有论文产出, 近 5 年论文产出数量较多, 有作为第 1 作者或通讯作者的高被引论文	论文近 5 年论文产出数量较少, 无作为第 1 作者或通讯作者的高被引论文, 近 5 年论文产出影响力总分值在中游。有国内专利 5 项, 专利无被引, 无转化
科研管理能力	参与项目较多, 项目骨干和参与人, 未找到项目负责人资料	科研项目较多, 多作为项目负责人和骨干参与项目, 科研管理能力较强
学术潜力	学术论文和会议论文上升趋势, 积极参与学术会议并发表文章, 在近 3 年发文量和质量提升很快, 年龄在 35 岁以下, 具有很好的研究能力和学术潜力	学术论文呈平稳趋势, 近 5 年产出下降; 专利近 5 年无产出; 积极参与学术会议并担任职务
学术地位	从其为本领域国际一流期刊担任编委和评审的经历看, 具有较高的国际同行认同度	主要为国内期刊担任编委和评审, 具有较高的国内同行认同度

6 结语

国内外的学者对科技人才的评价指标及其应用进行多角度的研究和探索,但是科技人才评价研究主要的问题是评价指标过于宽泛、指标值难以获取、定量指标过于侧重数量而忽视质量等问题,由此造成法科技人才评价在实际应用中难以操作的缺陷。本文从实际工作需求出发,针对性地对科技人才评价指标进行 5 个维度设计,细化指标,从科技信息大数据可获得的角度进行实证研究,指标体系较全面地涵盖了定性和定量指标,一定程度上弥补现有指标研究缺乏科学性和普适性、不能实际操作的不足。后续研究将对指标体系继续完善,针对不同应用场景,合理确定指标权重,建立更加完整合理的评价模型,为更多的人才管理提供支撑。

参考文献

- 1 文魁,谭永生.试论我国人才评价指标体系的构建[J].首都经济贸易大学学报,2005,7(2):5-8.
- 2 鲁捷,侯健.科技人才评估指标体系探讨[J].中国人力资源,2005,(1):48-51.
- 3 吴欣.高层次创新型科技人才评价指标体系研究[J].信息资源管理学报,2014,(3):107-112.

- 4 刘永,王继娜,周惠来,等.科技人才研究综述[J].河南图书馆学刊,2007,27(3):41-44.
- 5 魏晓通,宋丹.研究型大学教师科研能力结构分析[J].辽宁行政教育学院学报,2014,31(3):30-34.
- 6 赵伟,赵奎涛,王运红,等.科技信息资源共享与服务价值传递分析[J].科技进步与对策,2009,26(15):8-9.
- 7 徐建武,张秀梅,程煜华.我国科技信息资源消费行为研究[J].数字图书馆论坛,2014,(12):36-37.
- 8 中国科学技术信息研究所.国家创新能力基础设施建设规划研究报告[R].2007.
- 9 李思宏,罗瑾琰,张波.科技人才评价维度与方法进展[J].科学管理研究,2007,25(2):76-79.
- 10 杜建,张玢.学术影响力评价指标之间的相关性分析——基于医学领域某一细分学科的视角[J].评价与管理,2010,8(4):68-76.
- 11 李阳,许培扬.国内外科技人才学术影响力定量评价方法研究[J].医学信息学杂志,2012,33(8):38-43.
- 12 朱郑州,苏渭珍,王亚沙.我国科技人才评价的问题研究[J].科技管理研究,2011,31(15):132-135.
- 13 李光红,杨晨.高层次人才评价指标体系研究[J].科技进步与对策,2007,24(4):186-189.
- 14 田瑞强,袁军鹏,朱梦皎,等.高层次科技人才分层引进及评价研究[J].全球科技经济瞭望,2015,30(12):1-6.

2018 年《医学信息学杂志》征订启事

《医学信息学杂志》是国内医学信息领域创刊最早的医学信息学方面的国家级期刊。主管：国家卫生和计划生育委员会；主办：中国医学科学院；承办：中国医学科学院医学信息研究所。中国科技核心期刊（中国科技论文统计源期刊），RCCSE 中国核心学术期刊（武汉大学中国科学评价研究中心，Research Center for Chinese Science Evaluation），美国《化学文摘》、《乌利希期刊指南》及 WHO 西太区医学索引（WPRIM）收录，并收录于国内 3 大数据库。主要栏目：专论，医学信息技术，医学信息研究，医学信息组织与利用，医学信息教育，动态等。读者对象：医学信息领域专家学者、管理者、实践者，高等院校相关专业的师生及广大医教研人员。

2018 年《医学信息学杂志》国内外公开发行，每册定价：15 元（月刊），全年 180 元。邮发代号：2-664，全国各地邮局均可订阅。也可到编辑部订购：北京市朝阳区雅宝路 3 号（100020）医科院信息所《医学信息学杂志》编辑部；电话：010-52328673，52328674，52328671。

《医学信息学杂志》编辑部