

● 季淑娟, 董月玲, 王晓丽 (北京科技大学 图书馆, 北京 100083)

基于文献计量方法的学科评价研究

摘 要: 基于文献计量方法, 运用系统工程思想, 研究学科评价系统的构成, 通过对系统的输入变量、指标要素、资源环境和输出结果的分析, 结合对学科评价流程的解析, 探讨了学科评价过程中应关注的诸如学科分类、数据源选取以及评价指标设计等影响因素, 为今后有效地开展学科评价工作提供借鉴和参考。

关键词: 文献计量方法; 系统分析; 学科评价

Abstract: Based on the bibliometrical method and using the idea of system engineering, this article studies the structure of the disciplinary evaluation system. By analyzing the input variables, evaluation indicators, resources environment and output results of the system, and in combination with the analysis of the disciplinary evaluation process, the article discusses the factors which should be paid attention to during the disciplinary evaluation process such as subject classification, data source selection and evaluation indicator design in an attempt to provide reference for future disciplinary evaluation work.

Keywords: bibliometrical method; system analysis; disciplinary evaluation

学科评价方法主要有定性评价和定量评价。基于文献计量的学科评价, 是针对某一学科群体, 以文献体系和文献计量特征为研究对象, 采用数学、统计学等计量方法, 研究学科文献体系的分布结构、数量关系、变化规律, 进而探讨学科发展和建设过程中的某些结构、特征和规律。其主要是针对文献群实体自身运动变化规律的研究, 借助文献计量提供结构性强的描述数据, 对学科领域的技术与发展加以推断, 从而为学科的学术产出、学科影响力, 学科的发展潜力与创新能力进行评价分析^[1]。科学合理的学科评价能针对学科研究和发展跟踪世界研究的最新进展, 发现研究热点趋势, 激发新的研究思路, 寻找研究合作伙伴; 同时也使科研人员客观地了解自身学术水平和影响, 准确评估自身科研成就, 为申报科研项目、申请国家自然科学基金提供信息, 为科学研究、奖励、管理提供依据。对管理决策部门正确地评价科研学术活动, 合理地制定激励措施, 促进科学发展和社会进步, 都有一定的参考价值。因此, 运用系统分析的思想, 基于文献计量的方法, 研究学科评价过程的影响因素, 为今后有效地开展学科学术群体构成、实例分析和学术声誉等评价工作提供借鉴和参考。

1 学科评价系统的构成与分析

基于文献计量的学科评价系统主要由明确的学科评价目标、限定的学科属性特征、表征学科学术水平的文献计量特征指标、获取和计算评价指标的数据来源和分析工具

构成, 如图1所示。

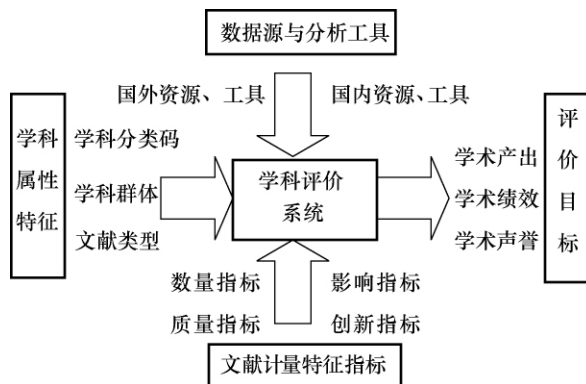


图1 学科评价系统的构成

学科属性特征是系统的输入变量, 文献计量特征指标是系统的要素, 数据来源与分析工具属系统的资源环境, 评价目标是系统的输出结果。通过对评价系统目标分析, 明确组成系统的基本要素, 寻求系统可用的资源环境, 准确地判定评价标准, 提出可行的评价方案和流程, 有效开展评价工作, 取得满足需求的评价结果。

1.1 评价系统的目标分析

系统分析首要是限定问题、确定目标。学科评价的目标可以是对学科领域的研究热点与趋势的跟踪分析, 也可以是学科研究群体的学术声誉与影响。针对文献产出的数量、质量及在一定范围内的学术影响, 通过选择计量指标, 运用合理的计算方法, 借助有效的计量工具, 对各种

文献计量学指标计算并进行定量分析。从多尺度的评价角度分析,学科评价可以分为3种不同尺度,宏观大尺度评价是在世界范围内评价某一学科的状态,是以国家为单元的学科领域;中观尺度评价是以研究机构为单元的机构之间的学科排名,如中国重点大学的学科排名;微观尺度是以某个机构的学科团队或个人为评价单元,评价这个团队或个人在某一学科中所处的学术地位和影响。因此,开展学科评价工作之初,首先要明确评价尺度和评价目标。

1.2 学科属性特征与文献分类

学科评价的基本单元是学科。关于学科的概念和内涵并没有统一、确切的定义。可以认为学科是通过持续地对某一科学领域或某一问题的研究形成相对稳定的学术研究领域,经科学分类而形成完整的、相对独立的知识体系。其应有明确、长远的学术发展方向,有自己的理论基础,且构成合理的理论体系。基于文献计量学进行学科评价,首先应明确学科的分类标准,在一致的标准下才可能进行比较评价。《中华人民共和国学科分类与代码国家标准》(GB/T 13745—2009,简称《学科分类与代码》)是中华人民共和国关于学科分类的国家标准。该标准规定了学科的分类与代码,分为自然科学、农业科学、医药科学、工程与技术科学、人文与社会科学5个门类,62个一级学科或学科群,676个二级学科或学科群,2382个三级学科^[2]。文献的学科分类标准与国家的学科分类标准不同,且不同国家对于文献的学科分类也不同,我国文献学科分类体系采用国家推广使用的《中国图书馆分类法》^[3]。其分为马克思主义、列宁主义、毛泽东思想、哲学、社会科学、自然科学、综合性图书5个部类,22个大类。我国的文献评价机构和文献检索工具多采用《中国图书馆分类法》为学科分类基础。为了能采集到分学科的文献相关统计数据,许多评价体系也采用《中国图书馆分类法》的学科体系划分学科。

国外的评价系统比较成熟,但国外的文献学科划分与国内文献学科分类不同。如Essential Science Indicators (ESI)是汤姆森路透汇集和分析ISI Web of Science所收录的学术文献和其所引用的参考文献建立的分析型数据库,受到了国内外学者的特别关注和认可,被广泛应用于各类学术评价中,其所收录的期刊被分为工学、生命科学、社会科学、理学、农学、医学等22个学科。其提供的文献学科分类是自行规定的,与我国的学科划分和图书馆分类法对文献的学科分类有很大不同。因此,在学科评价之初,应明确学科的界定以及文献分类。

1.3 文献计量特征指标

文献的发表、出版与被利用情况从一个侧面反映了学科群体的研究实力和水平,代表着学科的发展速度和活跃

程度,体现了学科研究成果的受关注程度和学术声誉。文献特征的技术交流等情况,为定量开展学科评价工作提供依据。

1) 学术产出。在学科发展过程中会形成许多有价值的文献,诸如图书、期刊、报纸、论文集、专利、档案、网络文献等。以文献为单元,通过定性和定量分析寻求学科发展的一些脉络。某一学科的学术产出数量大,说明该学科学术领域比较活跃,学科的受关注度高。如果是某个机构的某一学科的学术产出数量大,说明该学科在该机构占有一定的优势。学术产出质量的定量评价通常要借助一些参照对象,比如:以参考刊载刊物质量来评价论文质量,或以是否被公认国际知名数据库、检索工具收录等来评价。常用特征指标有:国内论文发表数量、国际论文发表数量、核心期刊论文发表数量、人均发表论文数量、被知名检索工具收录的数量等。

2) 学术影响。学术影响代表了学科的学术竞争力,学术影响可看作是某一学科的研究者或机构等对于该学科发展的贡献。文献的学术影响通常体现在同行关注与引用数量上,在一定程度上体现了学科的学术价值、学术水平和影响力。有研究认为作者发表论文的被引频次与作者的学科影响力成正相关关系。有基于期刊的特征指标,如影响因子、被引次数、核心期刊的发文数;也有基于分析作者和单篇文章的特征指标,如被引数量、篇均被引次数、h指数、被引速率、高被引论文数等,现在有些评价也考虑了文献的网络传播和开放存取等学术评价指标。

3) 学术创新。学术创新是从创新的角度考察学科学术活动的自主创新能力和创新水平,它是学科建设的关键。专利本身就具有创新性,是科技创新的具体体现,是转化为生产力最宝贵的知识财富之一。科技成果可以很好地反映科研工作的转化能力和效率,也是学术创造力的一部分。获得的专利数量和成果数量越多,说明该学科的学术创造力越强。常用特征指标:公开的专利数量、科技成果数、国内外专利授权数、获得的科技奖励数、科研成果转化数等指标。

4) 学术交流。学术交流是推动高校学科发展的不竭动力,高层次、高水平的学术交流活动,可以引导教师立足于学科的前沿,促进学科的交叉与融合,营造良好的科研氛围。因此,学术交流能力反映了学科发展过程中参与国际、国内交流的活跃程度,以及与国内外其他机构的合作交流情况,是学科学术水平评价的重要内容。常用特征指标:会议发表论文数和与其他机构合作的论文数。

1.4 系统分析平台与工具

文献计量的方法常常应用统计分析工具来研究作者、出版物和文献利用之间的关系与模式,过去的研究已经提

出了一些评价指标且已被广泛应用。有些公司也提供一些数据库分析工具用以直接提取分析指标。选择和使用先进的科技评价系统和工具,通过科学合理的指标数据提取和大规模、大尺度的数据分析,借助多维度、多侧面的关联计算以及形象生动的可视化展示,及时获取相关领域的学科评价数据,是开展学科评价的关键环节,准确、及时获得有效的评价数据是开展学科评价的基础。目前,常用的国外分析与数据来源系统有汤姆森路透的 ISI Web of Science, 期刊引用报告数据库 (Journal Citation Reports, JCR), 基本科学指标数据库 (ESI), CPCI (Conference Proceedings Citation Index, 原 ISTEP) 会议论文引文索引, 德温特专利创新索引数据库 (DII), 爱思唯尔 (Elsevier) 出版社研发的 Scopus 文摘引文数据库, 工程索引 (EI) 数据库, 等等。

常用的国内数据分析与来源系统有: 中国科学引文数据库 (Chinese Science Citation Database, CSCD)。它是教育部学位中心学科评估的指定数据库, 其收录我国中英文科技核心期刊和优秀期刊千余种, 已积累从 1989 年到现在的论文记录 300 万条, 引文记录近 1 700 万条^[4]。中文社会科学引文索引数据库 (Chinese Social Sciences Citation Index, CSSCI) 用于检索中文人文社会科学领域的论文收录和被引用情况, 至 2009 年已积累来源文献近 100 余万篇, 引文文献 600 余万篇^[5]。由于我国的期刊出版物较多, 因此, 国内对论文的评价也参考论文所刊载的期刊是否为核心期刊。核心期刊是指较为集中地刊载原创性论文的学术性期刊, 其刊载论文的学术水平较高, 并能反映某一学科的最新研究成果及前沿学术与技术研究动态和发展趋势, 且论文的密度大, 文摘率、引用率、利用率等相对较高, 在本专业领域具有较高的影响, 能够体现学科性、集中性和代表性。《中文核心期刊要目总览》(2008 年版) 研究评出的核心期刊有 1 983 种^[6]。

2 学科评价的流程解析

2.1 学科评价的流程

学科评价流程如图 2 所示。

2.2 评价流程解析

1) 确定评价目标是评价工作的首要任务。在科学界定学科范围和评价尺度目标的基础上, 应明确评价的目的是寻求学科教学与研究的前沿热点, 还是为了比较学校间的学科发展水平, 或是要探究发现学科发展的规律; 评价的服务对象是为管理决策者提供参考, 还是供科研人员了解, 或是供学科研究人员借鉴。

2) 分析学科的特点和文献的关系。由于各个学科都有自己的特点, 尤其在文献的数量、利用与评价指标等方

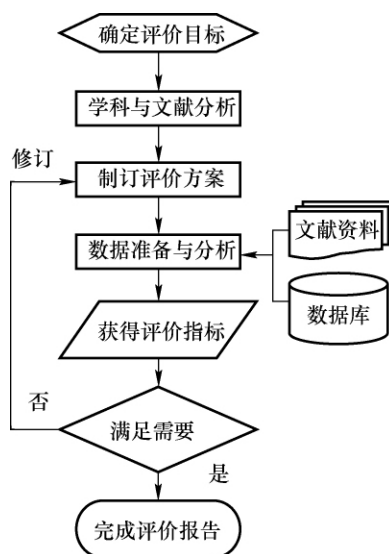


图2 评价系统的流程

面差异很大, 在研究过程中还要分别对待。如对于理学的物理学、化学等研究, 工学的计算机与材料工程等研究和对于图书情报学、心理学、社会学等研究有很多不同。文献类型的筛选、国内外资源的选择、学术研究成果表现形式等不同学科也会有差异。在明确学科分类的基础上, 通过学科查找对应的学科文献分类, 定位学科文献的范围和尺度, 提出参与评价的机构限定, 更进一步是学科群体的人员限定及文献类型的因素确定, 明确系统的输入变量。

3) 制订评价方案是整个流程中至关重要的一步。其包括评价内容、评价的具体方法和工作程序、评价指标设计与定义等。指标的获取主要依赖计算指标所用原始数据的可获取性和一些专用的数据库工具。在这个过程中应注重方案的可行性, 尤其是评价指标的可获取性、工具、方法及可行性, 确保评价工作的有效进行。最可行的方案也许不一定是最佳方案, 但它应该是在评价约束条件内, 根据评价标准筛选出的最现实可行的方案。

4) 学科评价对象主要体现在对学科科研成果的评价、对学科学术出版物的评价、对学科队伍的评价、对学科科研机构的评价。因此, 应从评价研究的实际出发, 根据数据指标特性选取合适的数据库。如评价某一学科在世界范围内的影响与竞争情况, 可以利用 ESI 数据库系统, 其提供进入全球某学科领域 1% 的机构的学术相关指标, 有针对性地对分析国际科技文献, 从而了解一些著名人物、机构、国家和学术期刊在某一学科领域的发展和影响, 获取相应的评价指标, 如某时间段内的相关机构发表的学科论文数量、论文的被引次数、论文的平均被引次数、高被引文章的相关指标。如果需要获取某学科的期刊学术评价指标, 国际上比较著名的有 JCR 数据库, 可以获取不同学科期刊的影响因子、发文量、学科中哪些国内期刊被 SCI 收

录,等等。国内的期刊评价数据可以取自 CSCD、CSSCI 或中国学术期刊电子杂志社每年出版的《中国学术期刊影响因子年报》等。针对一些专业性较强的学科,可以选择一些专业类型的数据库参与评价。如生命科学文摘索引数据库 (BIOSIS Previews), 生物医学文摘数据库 (MEDLINE), 化学文摘数据库 (SciFinder Scholar)。

5) 定量评价结果最终以直观的计量数据表现。如表现学科学术产出的数量指标,表现学科学术影响力的影响指标,表现学科发展潜力和创新能力的成果指标。基于文献评价的计量指标设计、数据准备、数据分析以及指标的计算方法对评价结果影响很大,因此,必须正确地理解指标,解读相关数据的含义,才能准确地查询与整合数据。可以借助数据挖掘技术或统计分析方法进行数据的分类、估值、相关性分析、聚集等工作,根据相应的评价指标得到的结果可通过其他途径进行实证研究,如通过问卷调查、专家咨询、与其他结论来对比等定性的方法来判定,如果定性定量结果相吻合或有合理、科学的解释,则评价工作达到目标。如果偏离较大或有异常,则应调整评价方案,修改约束条件或评价标准,甚至调整限定问题,重新进行评价工作。

6) 评价工作的最终表现形式应以评价报告的方式体现。评价报告目前并没有规范的内容与格式,但其基本内容应包含:学科简介、评价目标、评价目的与内容、评价流程、计量指标、数据来源或计量工具、检索数据、对比分析以及评价结果,并应给出明确的结论,必要时应附加原始检索数据。

3 学科评价过程中应考虑的影响因素

3.1 学科与文献分类

表1 10种国外期刊在不同系统中的学科分类

序号	刊名	ISSN	出版 国	不同系统的学科分类		
				Elsevier	中图 法	SCI
1	Acta Materialia	1359-6454	英国	材料科学	TF	材料类
2	Scripta Materialia	1359-6462	英国	材料科学	TF	材料类
3	Intermetallics	0966-9795	英国	材料科学	TF	材料类
4	Journal of Alloys and Compounds	0925-8388	瑞士	材料科学	TF	材料类
5	Synthetic Metals	0379-6779	瑞士	材料科学	TF	材料类
6	Applied Surface Science	0169-4332	荷兰	材料科学	TF	材料类
7	Hydrometallurgy	0304-386X	荷兰	化学工程	TF	冶金类
8	Materials Characterization	1044-5803	美国	材料科学	TF	材料类
9	Powder technology	0032-5910	瑞士	材料科学	TG	化学工程
10	NDT & E International	0963-8695	英国	工程学	TG	材料类

表1所示为10种国外期刊在3个不同数据系统中的学科分类情况^[7]。可以看出,3个不同体系对同一种期刊的分类可能存在不同的划分,而且,不同学科文献的利用与引用行为也可能存在差异,评价数据的参考指标变化较大,因此,在开展学科评价研究过程中,应具体学科具体分析,正确理解学科领域的研究内容,准确定位学科范畴,避免不同学科之间的定量分析比较。

3.2 数据源的选取

1) 权威性和知名度。数据的权威、可信是开展学科评价选择数据来源时比较关注的问题。国内外常用的有汤姆森路透基于SCI, SSCI, A&HCL开发的ESI计量分析数据库和爱思唯尔出版社研发的Scopus学术评价系统。2007年上海交通大学公布的“世界五大学术领域排名”中的论文高被引人数、期刊论文发表篇数、发表在前20%期刊比例等指标就是取自ESI, SCI, SSCI和JCR等数据库。武汉大学中国科学评价研究中心2007年公布的世界一流大学及学科竞争力评价结果,其原始数据全部取自在全世界有着极其广泛影响的ESI数据库和DII专利库,从而保证了数据的权威性和可信度^[8]。

2) 数据的深度和广度。不同数据库收录文献的标准和限制不同,故文献的数量也不同,由此产生的统计分析数据差异很大。ESI数据库收录全世界11 000多种期刊的1 000万篇文章。Scopus收录来自全球5 000多家出版社的近18 000多种期刊,拥有4 100万条记录,其中不仅包含同行评议期刊,还包含部分开放存取期刊和纯电子期刊等^[9]。由于数据库的内容覆盖和学科侧重不同,使得用户可以根据需求选择数据来源。全球著名的英国《时报高等教育增刊》的“世界大学排行榜”2007年公布的评价数据中的论文被引用率(Citations/Faculty)的数值来源,由原来的Thomson Scientific公司的ESI改为Scopus数据库中2001—2006年期间的数据。因为,“与ESI比较,Scopus数据库有较多工程与医疗健康学术论文,也收录更多亚太地区文章。但也有人认为,ESI收录了更多的人文学论文,而且有更多资料收集分析经验”^[10]。因此,选择数据来源要考虑评价所应覆盖的数据深度和广度。

3) 关注多元化的网络数据。随着互联网的发展,网络逐渐成为全球性的学术资源,网络上资源更加丰富,是学术交流与沟通的重要工具,借助网络学术传播能增加学术的能见度和影响力。例如,Google Scholar的特点是提供引文链接。它的引文分析功能有独到之处,不仅能广泛反映学术期刊之间的相互引用关系,还能反映学术期刊、书籍、预印本库、机构库、各类非联机出版物等多种类型文献之间的相互引用关系,因此其在文献相互引用中所能覆盖的范围要远远大于现有的引文数据库。目前,西班牙网

络计量研究中心的网络实验室正在建立全球大学的资料库,借助 Google, Google Scholar, Yahoo, Live Search, Exalead 等网络搜索引擎进行大学排名。其评价指标包含:规模——网站搜寻的各大学网页的页数;能见度——各大学网页对内外被链接数、被浏览数目等;学术档案——不同类型如 ppt, pdf, doc 等文档总数量;学术论文——通过网站所搜寻到的学术文章、报告与相关其他学术研究等的数量。可见,网络资源在今后学术评价中的作用不容忽视,网络数据的引入将会使评价的内容更加多元化。

3.3 评价指标的科学设计

目前基于文献计量的评价指标一般采用多指标综合评价,科学规划与选取指标、合理控制指标数量相当重要。既要考虑覆盖评价学科内容,也应考虑数据的科学性和可行性,在全面反映评价学科内容的基础上,避免出现因指标过多造成的指标同质现象和数据量大不易获取或处理难度大等问题,简化评价工作。

1) 指标的分类。选取指标时首先要考虑其代表性和评价功能,各种指标按照学科评价的内容分成若干类,使相似指标尽可能归为一类。除人为定性划分外,也可采用聚类分析的方法,或通过因子分析从多个评价指标中选择少数几个综合指标的降维多元统计方法。

2) 指标的相关性。由于指标间常常存在着不同程度的相关性,比如,用于评价学科影响力的评价指标有他引频次、被引频次、被引论文率、未被引论文率,等等。为避免相似或类同指标的重复使用,选取指标时应考虑选取低相关度指标。可以采用主成分分析法和因子分析法进行指标的相关性分析,提高评价的科学性。

3) 指标的权重。评价学科各个指标在整个评价系统中的价值高低、相对重要程度以及所占比例的大小通常量化地用“权重系数”代表。权重是综合评价中的一个重要指标,合理地分配权重是多因素、多水平量化评估过程的关键因素。因此,权重的构成是否合理,直接影响到评估结果的科学性。确定权重的方法有很多,如专家咨询法(Delphi 法)、专家排序法、层次分析法、秩和比法(RSR)、相关系数法等,但各种方法均有其优缺点。

3.4 数据核准与清洗

数据分析的结果依赖于数据质量,评价结果的准确性与可靠性在很大程度上取决于所用的评价数据。受作者在文章中用词、姓名、机构名称等规范标引的影响,同时,也有数据库本身数据标引与检索方式的问题,不同数据库的检索结果会出现不同程度的误差或错误,需要在提取评价数据时给予关注。以在 SCI 数据库中检索机构论文收录情况为例,由于该数据库没有提供“机构名称”检索途径,一般通过“地址”字段进行检索,由于机构地址写

法的多样性,检索中需要格外注意,以北京科技大学为例,检索结果有“UNIV SCI & TECHNOL BEIJING”,“BEIJING UNIV SCI & TECHNOL”,“UNIV SCI & TECHNOL, BEIJING”,“UNIV SCI & TECHNOL, DEPT ... BEIJING”等多种记录形式,包括国家名称变更、不规范标注等现象^[11],个人学术成果检索时同样存在着标引的多样性、重名作者等问题。上述现象说明在目前没有更理想的检索方法时,检索结果必须经过仔细核准与校对,检测数据中存在的错误、缺失、异常、不一致或不完备等问题,剔除错误、不准确或无分析价值的数据,填补或删除缺失数据,合并重复数据,确保评价数据的准确性。

4 结束语

学科评价是一项科学性极强的工作,基于文献计量的方法开展学科定性评价正处于实践研究阶段,尚无标准的流程、规范的指标和公认的报告格式,但此项工作的开展有助于为学科前沿领域的研究、影响与发展提供一种科学的评价方法。本文以系统分析的思想研究学科评价的整体构成与工作流程等问题,希望对今后科学、规范、有效地开展学科评价工作提供有益的借鉴。□

参考文献

- [1] 邱均平,王曰芬.文献计量内容分析法[M].北京:国家图书馆出版社,2008:263.
- [2] GB/T 13745—2009 学科分类与代码[M].北京:中国标准出版社,2009.
- [3] 中国图书馆分类法编辑委员会.中国图书馆分类法简本[M].4版.北京:北京图书馆出版社,2000.
- [4] 中国科学文献服务系统[EB/OL]. [2011-03-17]. <http://sdb.csdli.ac.cn>.
- [5] 南京大学中国社会科学研究评价中心[EB/OL]. [2011-03-31]. <http://essci.nju.edu.cn>.
- [6] 朱强,戴龙基,蔡蓉华.中文核心期刊要目总览[M].北京:北京大学出版社,2008.
- [7] 戴龙基,蔡蓉华.国外科学技术核心期刊总览[M].北京:北京大学出版社,2004.
- [8] 邱均平,赵蓉英,马瑞敏.世界一流大学及学科竞争力评价的意义、理念与实践[J].科技进步与对策,2007,24(5):138-142.
- [9] <http://www.info.sciverse.com/scopus/about>.
- [10] 侯永淇.全球与各国大学排名研究[M].北京:高等教育出版社,2007.
- [11] 董琳.学科评价之文献计量数据准备[J].情报理论与实践,2010,33(6):49-52.

作者简介:季淑娟,女,1963年生,博士,研究馆员。

董月玲,女,1976年生,硕士,馆员。

王晓丽,女,1974年生,硕士,馆员。

收稿日期:2011-06-09