

科技基础资源调查专项特征分析及其成效评价模型构建

王 祎¹ 王卷乐^{2,3} 郝丽娜^{2,4} 李 凯^{2,5}

(1. 国家科技基础条件平台中心, 北京 100862; 2. 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, 北京 100101; 3. 江苏省地理信息资源开发与利用协同创新中心, 江苏南京 210023; 4. 江苏海洋大学海洋技术与测绘学院, 江苏连云港 222005; 5. 中国矿业大学(北京)地球科学与测绘工程学院, 北京 100083)

摘要: 国家科技基础资源调查专项建立起了一条从本底资料收集到信息汇聚共享的重要科技资源供给通道, 成为我国基础科学研究的重要“源头活水”。但是科技基础资源涉及的学科领域多样且资源调查数据生命周期较长, 因此准确把握其数据流动特征、有针对性地加强科技基础资源调查的评价和管理已成为科技基础资源调查专项有力实施和提高成效的关键。在分析科技基础资源调查专项定位和特征的基础上, 挖掘科技基础资料采集、加工整理、汇交保存到共享应用各阶段对科技管理措施的要求, 采用定性评价与定量评价相结合方法, 构建科技基础资源调查专项评价模型, 为科技基础资源调查专项的科学管理和提升成效提供支撑。

关键词: 基础研究; 科技基础性工作; 资源调查; 科学数据共享; 成果评价模型

DOI: 10.3772/j.issn.1674-1544.2022.05.001

CSTR: 15994.14.issn.1674.1544.2022.05.001

中图分类号: G307

文献标识码: A

Special Feature Analysis of Science and Technology Fundamental Resources Investigation Program and Its Effectiveness Evaluation Model Construction

WANG Yi¹, WANG Juanle^{2,3}, HAO Lina^{2,4}, LI Kai^{2,5}

(1. National Science and Technology Infrastructure Center, Beijing 100862; 2. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101; 3. Jiangsu Province Geography Collaborative Innovation Center for Information Resource Development and Utilization, Nanjing 210023; 4. School of Marine Technology and Geomatics, Jiangsu Ocean University, Lianyungang 222005; 5. School of Geoscience & Surveying Engineering, China University of Mining & Technology (Beijing), Beijing 100083)

Abstract: The National Science and Technology Fundamental Resources Investigation Program has established an important channel for the flow of scientific and technical resources from the collection of physical background data to the convergence and sharing of information, and has become an important “source of living water” for basic research in China. Due to the diverse subject areas and the long life cycle of the flow of

作者简介: 王祎(1975—), 男, 国家科技基础条件平台中心副研究员, 研究方向为科技资源开放服务管理; 王卷乐(1976—), 男, 中国科学院地理科学与资源研究所研究员, 博士生导师, 研究方向为科学数据共享、地理信息系统与遥感应用(通信作者); 郝丽娜(1995—), 女, 中国科学院地理科学与资源研究所硕士生, 研究方向为数字城市空间信息采集与处理; 李凯(1998—), 男, 中国科学院地理科学与资源研究所硕士生, 研究方向为大地测量学与测量工程。

基金项目: 国家科技基础条件平台专项课题“科技基础资源调查专项成效分析及评价模型研究”(2022WT10)。

收稿日期: 2022年6月27日。

science and technology fundamental resources, it is a key task to grasp the characteristics of their data flow and to enhance the evaluation and management of this program in a targeted manner. This study analyses the positioning of the Science & Technology Fundamental Resources Investigation Program, analyzes its flow characteristics from the collection, processing and collation, remittance and preservation to the sharing and application of scientific and technical information, mining deeper into the requirements for scientific and technical management measures at each stage, and on the basis of this whole chain analysis, the evaluation model of the special science and technology basic resources survey is constructed by combining qualitative evaluation and quantitative evaluation, which is expected to provide support for the scientific management and enhance the effectiveness of the special science and technology basic resources survey.

Keywords: basic research, science and technology fundamental work, resource investigation, scientific data sharing, effectiveness evaluation model

0 引言

基础研究是认识自然现象,揭示自然规律,获取新知识、新理论和新方法的研究活动,是科技创新的源头活水^[1]。科技基础性工作是对基本科学数据、资料和相关信息进行系统考察、采集、鉴定,并进行评价和综合分析,以探求基本规律,推动这些科学资料的流动与使用的工作^[2],是科技工作中的重要组成部分。《国家“十一五”基础研究发展规划》将前沿基础科学研究分为3类^[3]:一是好奇心驱动的自由探索,二是国家任务的定向性基础研究,三是对基础科学数据、资料等系统地进行采集、分析、综合等基础性工作。因此,科技基础性工作是国家基础研究的三大类型之一。

作为科技创新的重要源泉,科技基础性工作长期受到国际社会的广泛关注。欧美等国家的探险家长期持续对南北极和青藏高原地区开展多学科的综合科学考察,获得了一手资料。美国国家水生资源调查项目(NARS)通过调查和统计,评估全国沿海水域、湖泊和水库、河流和溪流以及湿地的质量。加拿大启动了2016—2021年北极区域环境研究计划(ARES),目标是研究未来海上石油和天然气活动对北极地区的潜在环境和社会经济影响。俄罗斯高度重视地理考察工作,该国地理学会会长通常由国防部长兼任。在典籍志书编研方面,2011年由美国密苏里植物园等

牵头,开始编写世界在线植物志;丹麦正在牵头编写泰国植物志。在工程材料领域,美国通过材料自然环境腐蚀试验积累了大量材料环境腐蚀数据,20世纪50年代将试验研究范围扩展至全球。20世纪60年代,日本开始本土系统的材料腐蚀数据收集,为其高品质钢铁工业和防腐蚀新技术发展奠定科学基础。在人口健康领域,美国于1956年通过《全国健康调查法》授权开展全国健康调查,且于1990年国会通过《全国营养监测和相关调查法案》。纵观国外科技基础资源调查的现状,主要发达国家对资源的收集具有前瞻性和计划性,而且科技资源收集的范围早已从本土延伸到境外,收集的领域和类型也无所不及,丰富的科技资源收集和积累助力其占领基础研究科技发展的制高点。我国也长期重视科技基础性工作,并通过科技项目等方式给予持续支持,还设立了“科技基础资源调查”专项,连续多年开展科技基础资源调查工作。科技基础资源调查专项支持的重点集中在考察调查与资料整编两大部分内容。具体可再分为综合科学考察、专项科学调查和科技资料整编。综合科学考察是对特定区域或主体开展跨圈层要素的综合性资料调查、采集、整编和科学规律分析,并形成结论。专项科学调查是对特定区域或主体,开展专题要素的资料调查、采集、整编和科学规律分析。科技资料整编是围绕某项科学主题深度加工成统一表达,形成典籍、志书、图集、报告等成果。科技基础

资源调查专项实施以来,支持获取了大量具有重要科学价值、长时间序列的原始数据和资料,保存了为数众多的样品和标本,编制了一系列图集图鉴和重要志书,为推动学科发展、科技进步和经济社会发展发挥了重要的基础性作用^[4]。

尽管我国科技基础性工作取得了较大进展和成效,但与发达国家相比尚存很大的差距,还不能满足国内经济、社会和科技发展对科技基础资源不断增长的急迫需求。面对新冠肺炎疫情对经济社会的持续影响、国际地区冲突不断加剧、全球资源供应链中断风险加大等现状,原创性科技成果和自主知识产权的关键技术成为国家创新发展的重要基础。与此同时,在我国倡议的新的人类命运共同体正处于构建之中,对科技基础资源的收集、挖掘和利用提出了更高的要求。在这一新的需求牵引下,科技基础资源调查专项要主动适应,谋求新的发展,这就需要深入思考其特征和定位及其在实施中的管理要求,有针对性地分类加强其成果评价模式。

1 科技基础资源调查专项定位和特点分析

美国唐纳德·司托克斯^[5]基于巴斯德微生物研究的案例,将科学研究依据是否具有基本认识和实际应用两个特点划分成“象限四重图”。第I象限为由好奇心驱动的基础研究,称为波尔象限。第II象限为由应用引发的基础研究,如巴斯德对发酵的研究解决了甜菜制酒的现实问题,被称为巴斯德象限。第III象限代表追求实际应用的生活工作技能训练与整理试验,被称为爱迪生象限。第IV象限表示既没有基本认识,也没有明确的实际应用目的,而是对一些特定现象或专题的研究,被称为皮特森象限。

科技基础资源调查是面向科学目标和国家战略需求开展的,对自然本底数据和材料的获取、对已有数据和材料的整理与编研等科技基础性工作,其是“以实用为目的”和“以求知为目的”的两大基础科学研究的重要基础资源支撑。科技基础资源调查专项主要从事标准化、重复性、周期性的调查工作,且调查人员要经过系统的培训

和训练,因此最接近于技能训练与整理实验(第IV象限)。在实际调查过程中,存在面向天文、空间、物理等基础学科领域的调查,也存在面向国家生态文明、“一带一路”建设、区域高质量发展、健康中国等国家需求的调查,以及面向交通建设、材料腐蚀等工程领域的调查,因此与纯基础研究(第I象限)、应用引发的基础研究(第II象限)以及纯工程应用研究(第III象限)都有交集。但总体应用出口方向,集中在应用引发的基础研究领域,即第II象限(巴斯德象限),如图1所示。

通过科技基础资源调查与科学研究项目和行业调查的对比(图2)可知,科技基础资源调查专项具有基础性、长期性和综合性特点。

(1) 基础性。科技基础资源调查专项围绕自然本底信息和基础科技资源,开展第一手资料的获取和已有历史资料的整编,面向科技界和全社会开放共享,为学科发展、科技创新和行业领域应用提供基础支撑。

(2) 长期性。科技基础资源调查对自然界各要素开展持续性野外观测、周期性区域本底更新以及长时序历史科学数据和资料整编,为揭示地球系统长期演化规律和区域可持续发展提供基础支撑。

(3) 综合性。科技基础资源调查专项项目覆盖范围非常广,涉及地理、地质、地球物理、生态、生物、农林、环境、大气、海洋、工程材料、健康等多个学科领域。单一的行业调查数据往往只局限在业务工作范围之内,不能掌握全貌。科技基础资源调查克服地理单元之间的行业部门管理界限,促进跨行业部门的科学数据采集和系统化整编。

2 科技基础资源调查专项数据流动过程分析

2.1 资源类型划分

科技基础资源调查专项所产生的数据资源价值代表性体现在数据库。数据库可分为资源型数据库、研究型数据库与参考型数据库^[6]。这3类数据资源特点也反映了科技基础资源调查项目的

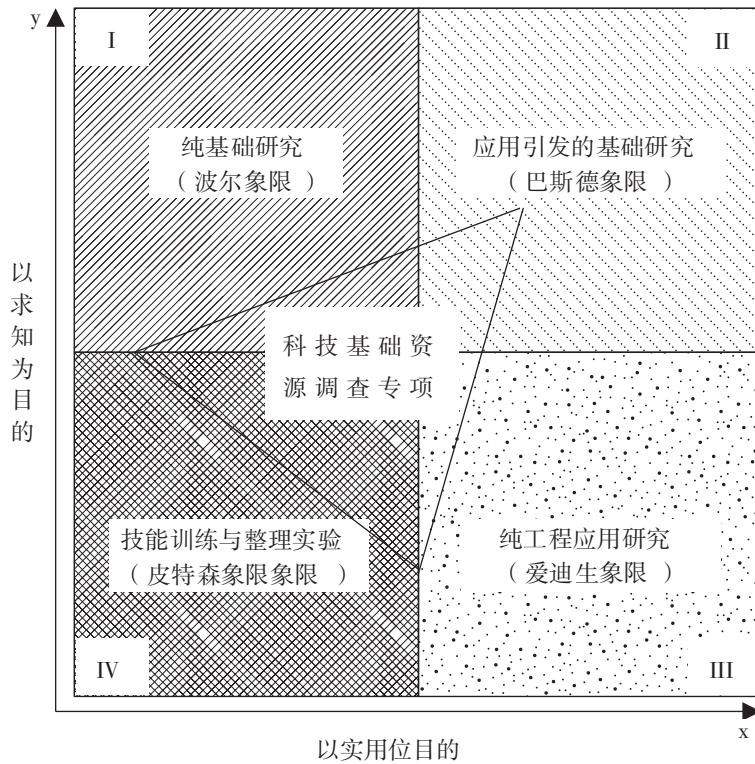


图1 科技基础资源专项科学研究在巴斯德象限中定位

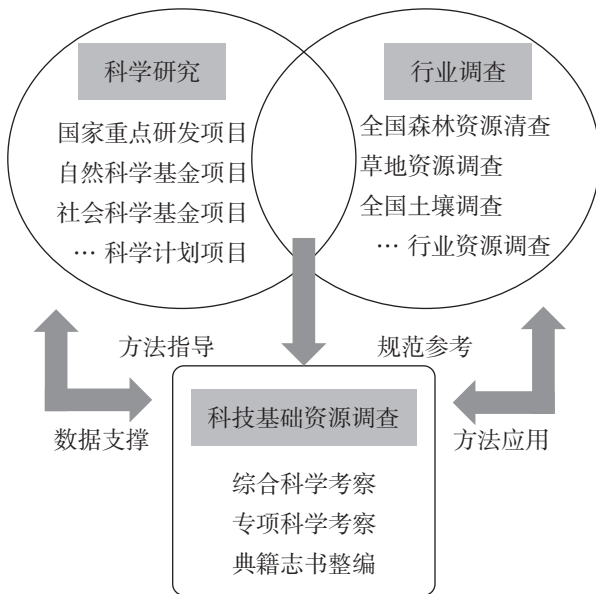


图2 科技基础资源调查与科学研究项目和行业调查的关系

差异，分别对应为资源增长型、应用导向型和标准比对型。

资源型数据库是指那些数据资源规模大、数据时空序列完整且覆盖面大，但缺乏特别针对性的应用对象，而是面向普遍性的应用群体和领

域。如美国国家航空航天局（NASA）的对地观测领域的12个分布式数据中心^[7]，均是资源型数据库。对应于此的科技基础资源调查专项项目则称为资源增长型的项目，即通过调查或资料整编大量增长数据资源。针对资料空白区，则通过数据资源采集来填补空白。针对周期性调查区域，则通过资源增长来不断增加本区域的资源数量，提高数据的长时间序列和数据可比性。多数科技基础资源调查专项项目具有此类特点，如冰川、冻土、湿地、湖泊、沙漠等周期性调查和资料积累等。

研究型数据库是指通过某个项目研究所产生的数据，这些数据一般只经过有限的处理，通常只为特定的应用目标服务^[8]。在科技基础资源调查专项中存在这类以特定的应用目标和研究群体的项目，即应用导向类的项目。如针对特定生态脆弱地区的黄土高原生态系统与环境变化考察项目、针对西北农业缺水的西北干旱地区农业经济用水量调查专项、针对南水北调的水环境变化的南水北调（中线）水源地生物群落环境调查专

项等。

参考型数据库旨在为大范围应用提供标准的数据支持，因此要求数据更为权威和规范，能够成为数据对比的基线。典型的参考型数据库包括蛋白质数据库、美国国立卫生研究院的基因序列数据库 GenBank 等，其他同类数据能否发表要先和这里的数据进行比对。相对应地，科技基础资源调查专项也有立典、立志和标准规范类的项目。如中国地层立典剖面及若干断代全球界线层型项目、中国动物志、中国孢子植物志和中国植物志的编研项目、中国自然疫源性疾病流行病学图集编研项目等。

据不完全统计，科技基础资源调查专项产出研究型数据库、资源型数据库与参考型数据库的占比分别为 26%、32% 和 33%。这些数据资源在经过早期分散阶段采集和加工，后期成果发表和确权后，逐渐完成资源标识和价值确定（图 3）。只有具有完整的科技资源标识和价值权益归属的数据，才能具有科学数据开放、传播、流动和增效的能力。

2.2 资源流动过程

科技基础资源调查的任务是获取各种科技资源。这些资源通过科技管理的疏导，汇聚到各类科学数据中心和资源库，形成全社会的科技基础资源供给。科技基础资源调查的数据内容覆盖

领域广，如覆盖了农业、林业、生物、地理、地质、海洋、大气、健康、工程等诸多领域。这些多样化的科技基础资源就像“毛细血管”一样收集整理起来汇聚到国家科技基础条件平台这个“心脏”中，然后输送到各个行业领域“组织”里，为整个科技创新和经济社会发展的“身体”提供源源不断的数据能量。

科技基础资源调查的数据流动过程分为科技基础资源采集、处理分析以及数据汇交（图 4）。科技基础资源数据来源于地球系统和生物圈（含人类活动）各要素，经过数据处理和分析后，得到规范化和结构化的数据体。进而对数据实体进行汇编、集成、统计，形成数据库、标准规范、咨询报告、期刊文献等不同的成果形式。由于不同项目面向的问题特点不同，科技基础资源调查专项项目的成果总体上可以分为以资源增长为特征的、以解决应用需求导向为特征的和以标准比对参考型为特征的不同类型。这些特点不同的科技基础资源均汇交到相应的国家科学数据中心或资源库^[9]，并对外提供共享服务。

2.3 资源变化特征

科技基础资源的数据流动体现出以下 3 个方面的特点：一是从无到有，即从无数据、暗数据到有数据和开放数据；二是从分散、碎片和多样化的数据向集中、规范和统一的数据资源转变；

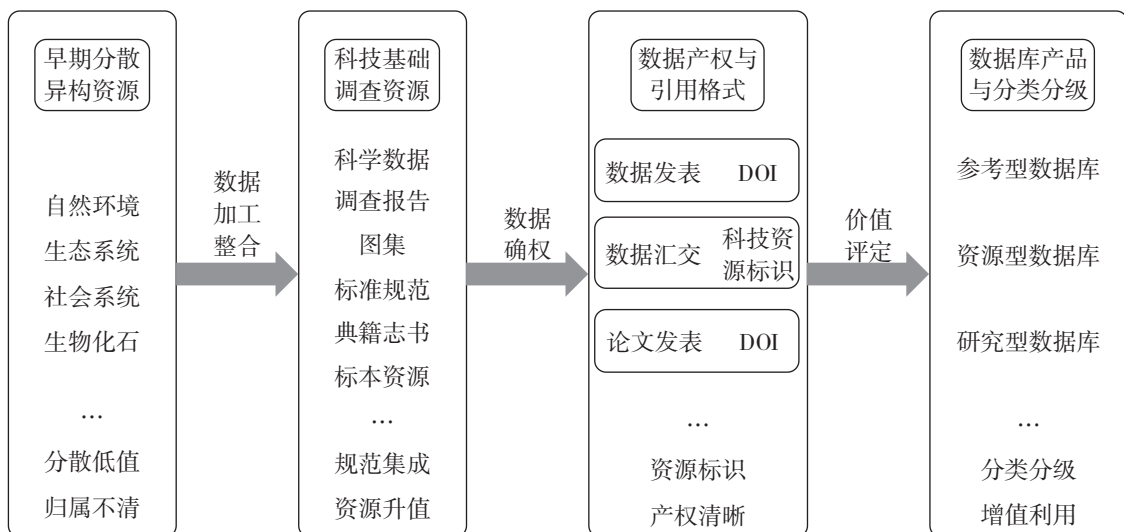


图 3 科技基础资源调查数据流转

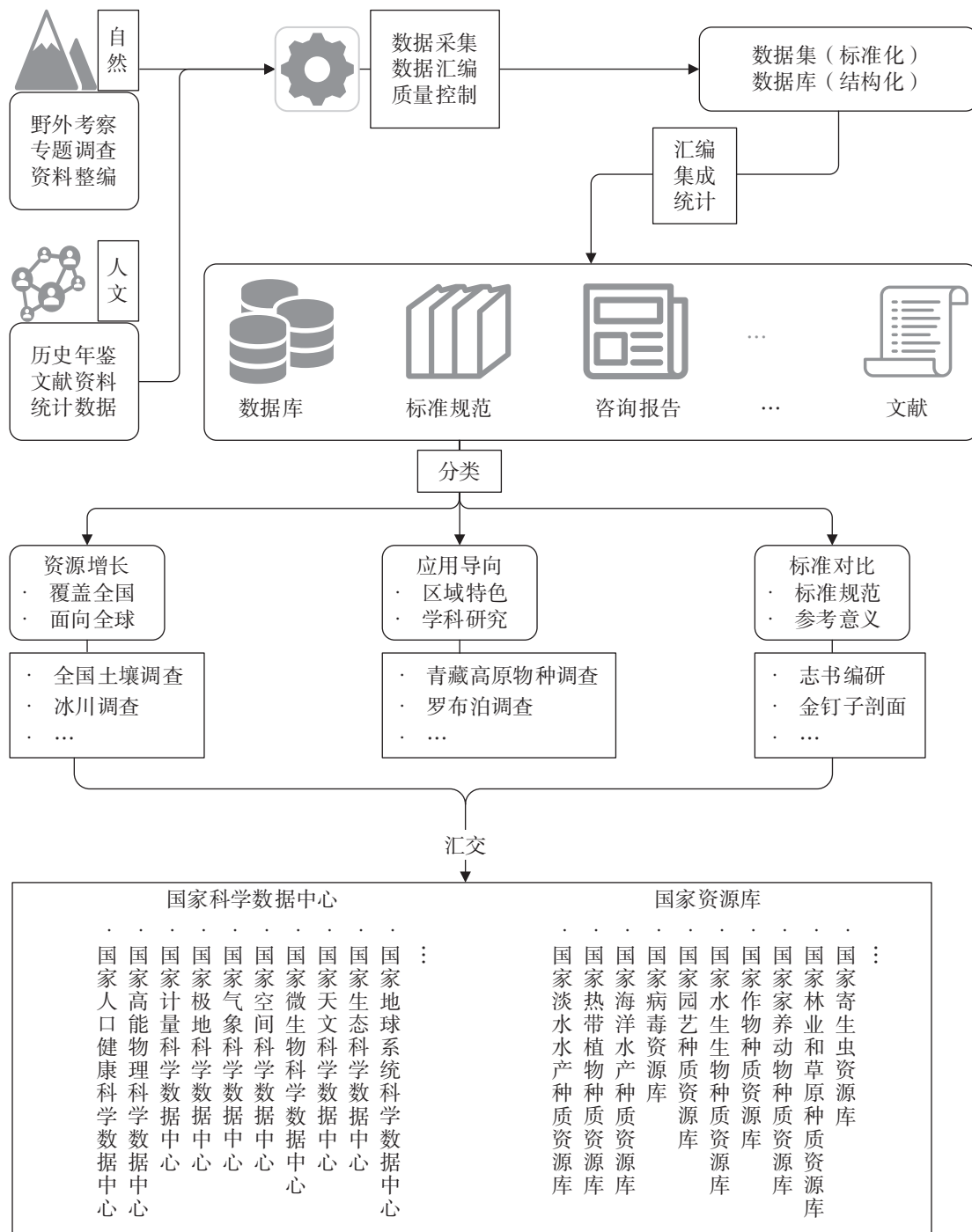


图4 科技基础资源调查数据流动形式

三是从低值到增值的价值增加过程，数据资源的价值和生产力得以提高。

(1) 从无到有。这主要体现在资源空白的地区。如在新中国成立初期，我国很多地区缺少资料、家底不清，呈现大量的空白区。这个期间对西藏、华南地区、黄河中游、黑龙江流域、新

疆、青藏柴达木盆地盐湖、黄河及长江流域土壤调查、华南和云南橡胶为主的热带生物资源、西北地区治沙、青甘地区以及西部地区南水北调、内蒙古和宁夏等地区进行考察^[10]。在这个过程中，产生大量的野外考察第一手数据，通过整编积累成为重要的数据集、考察报告和地图资源，

如编制了各区域考察报告、全国 1:100 万土地资源图和 1:100 万草场资源图等。

(2) 从多样到统一。科技基础资源数据源具有多源、多学科、多类型等多样化特征。以志书为例，任何一部志书编修的数据来源都是多源的，通常涉及国家或地方机构和考察团队等^[11]。只有通过以上多来源资料的汇编、整理和加工后，才能形成有科学依据、可以长期参考和标准比对的基础资料。因此，典籍、志书、图集的编研始终是科技基础资源调查专项的标志性成果特色。

(3) 从低值到高值。科技基础资源调查所产生的数据是具有价值属性的。随着数字经济的发展，数据成为有巨大价值潜力的资源，各行各业通过对数据资源的收集、整理、汇编，加快挖掘出更多有效信息服务于生活与应用^[12]。将原始分散的数据进行专业的整理、加工、集成，挖掘数据的价值、增加数据的可利用性、提取数据的有效信息、提升数据要素的价值，成为当前的重要的应用范式。因此，科技基础资源采集加工的过程也是数据增值的过程。通常数据准确度和整合

度越高，其价值越大^[13]。科技基础资源数据价值随数据流动的变化见图 5。对应于不同类型的资源，其价值曲线有一定的差异。标准比对型数据短期内增长不高，但最终价值最大。应用导向型数据是为具体应用而产生的，因此短期内价值提升最快，但长期增值相对不大。资源增长型数据价值居于中间，其随着数据量的不断增大，也具有较大的价值。

3 科技基础资源调查专项成效评价模型的构建

3.1 总体管理环节

适宜的成果评价手段不仅利于科技基础资源调查专项提高效益，而且使得项目管理过程高效且有章可循。在科技基础资源调查专项实施过程中，伴随着数据的流动，其在不同阶段的管理和评价内容是不同的（图 6）。在数据采集阶段，数据离散程度较高，需要适应多源数据的采集和获取；在数据处理阶段，要对不同来源数据进行规范化，强调标准化和数据质量控制；在数据汇交阶段，数据按统一规范进行汇交，并制备相应的

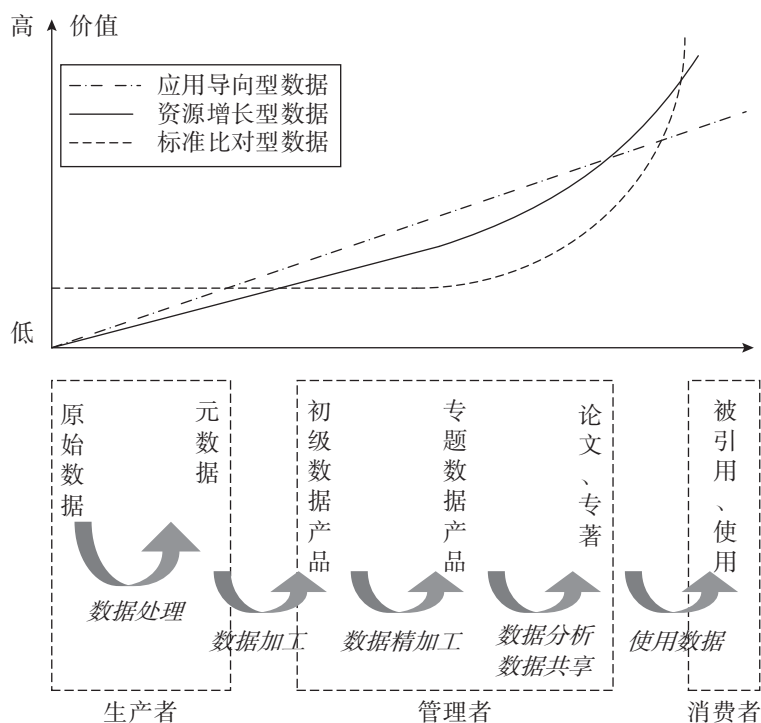


图 5 科技基础资源调查数据价值变化

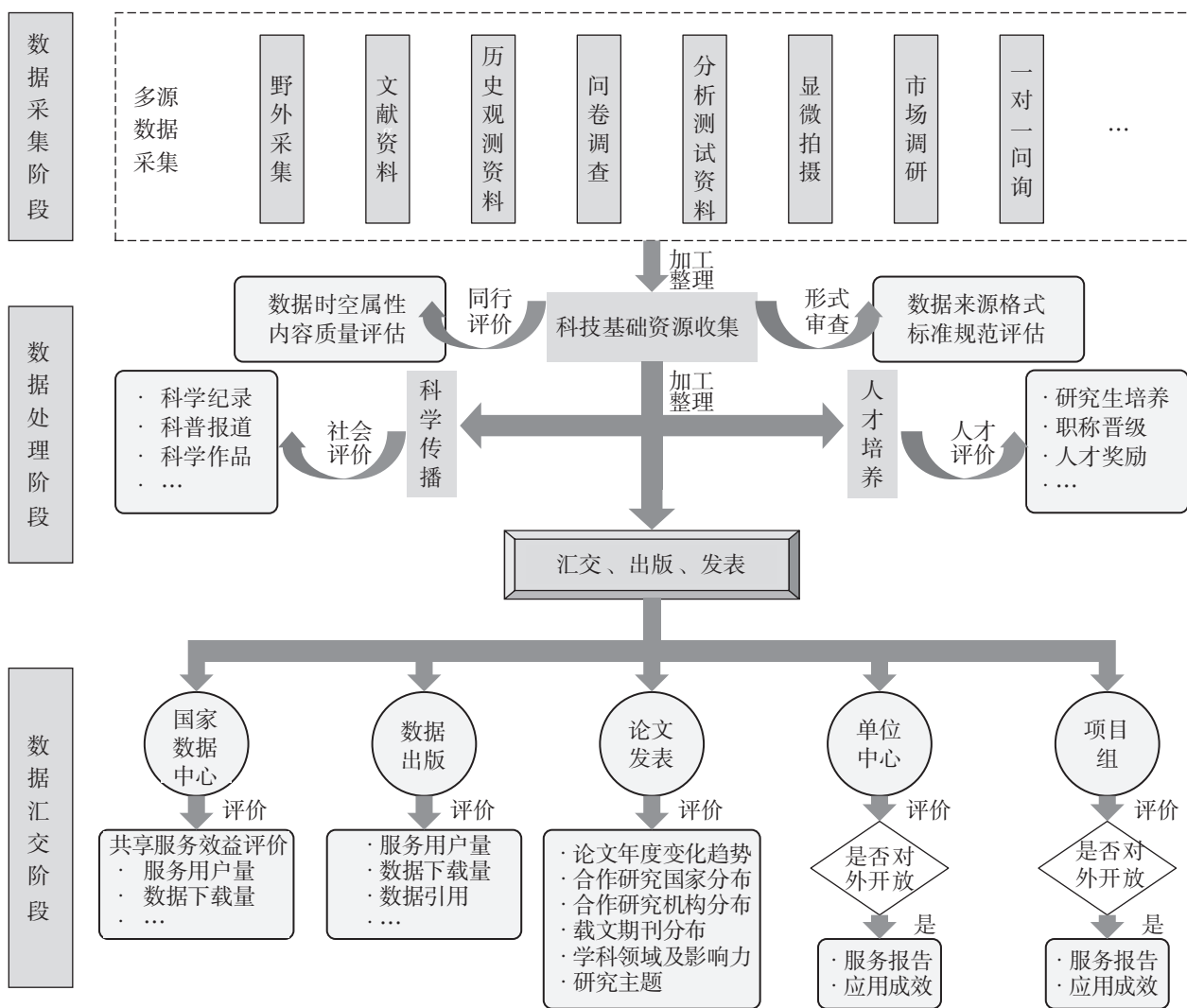


图6 数据流动过程及其评价方法

数据文档和元数据说明，以实现最大程度的开放共享和应用。对于科技基础资源成效的评价主要包括以下 11 个方面。

(1) 数据数量评价，主要针对提交数据的数量是否达到任务书中汇交量指标，以及在服务系统中的数据公开共享率进行评价。

(2) 数据质量评价，主要通过同行评价的手段，对项目汇交数据内容的时间跨度、空间范围、数据价值、数据不确定性程度等进行评价。在数据汇交的过程中，尤其要重视数据质量的评价，并在课题汇交、项目汇交的层级管理过程中充分体现数据质量的检查和溯源。

(3) 数据规范化评价，主要是审查数据的格式和内容是否符合相应的标准规范，以及在开放

服务中的资源描述和内容规范化程度。

(4) 社会影响评价，主要体现在科学数据的传播及其影响力，通过计量分析的方法对相关科学记录、科学作品、科普报道等的社会影响力进行评价。

(5) 人才培养评价，主要对其项目培养的获得学位的硕士生数量、博士生数量、获得职称晋级的人数以及获得的人才荣誉称号等情况进行评价。

(6) 论文成果评价，主要是对科技基础资源调查成果的学术影响力等进行评价。如通过计量项目产生的中英文论文数量、合作研究国家、合作机构、载文期刊、学科领域及其影响力、研究主题进行评价。

(7) 标志性成果评价, 图集志书典籍专著等标志性成果评价包括对成果的数量、质量、领域覆盖面、支撑成效、引用率、影响力等进行评价。

(8) 成果奖励评价, 主要对科技基础资源调查所获奖项的类型、等级、数量等, 以及对成果的价值和影响力进行评估。

(9) 技术先进性评价, 考虑科技基础资源调查专项在信息化、大数据、物联网、云计算等技术支持下的应用情况, 鼓励采用数字化和智能技术对专项赋能。

(10) 数据共享评价, 主要对出版和发布数据集的数量, 以及数据在国家科学数据中心的用户访问量、下载量、数据引用量进行评价, 并对数据应用的反馈情况进行收集和评价。

(11) 机构共享评价, 主要对数据生产依托单位或机构的数据共享情况、成果宣传情况等进行评价。

3.2 分类评价

科技基础资源调查专项项目评价的类别可分为资源增长型、应用导向型和标准比对型。同一科技基础资源调查项目可能交叉分属不同的类型, 如某些项目既属于资源增长型, 又具有特殊的时代紧迫需求特点而属于应用导向型。不同类的资源需要差异化的管理方式, 对不同科技基础资源调查成果建立管理模型。

(1) 资源增长型。资源增长型以扩大数据量为基本特点, 要求面向全球、覆盖全国, 形成本底资料。资源增长型项目评价的重点也是其量, 包括时间、空间、要素, 覆盖面、覆盖的深度, 与同类已有数据相比的弥补度和稀缺性等。目的是填补空白和加强厚度, 即通过填补空白以求完整, 通过加强厚度以求数据可比提高精度。此类资源的评价对应管理举措上, 可以从数据获取手段、数据处理方法、数据库建设以及在填补空白方面等进行评价。综上所述, 其评价要素集中在填补空白、数据时空完整性以及数据先进性等方面。

(2) 应用导向型。应用导向型是结合区域或

学科发展需求, 要求针对性的目标清晰, 短期内的成果支撑服务的作用明显, 且具有区域或学科的代表性、科学性以及可推广性。如落实中央的重要战略部署、针对国家关切或区域发展获取全球或区域数据资源、针对地方紧迫的应用需求而开展的相关调查等。其对应管理层面, 重点关注其应用成效如何, 如是否形成上报中央的咨询建议批示、是否有提供地方应用的证明、是否有重大的用户使用证明等。综上所述, 其评价要素集中在目标针对性、应用证明和重要应用案例。

(3) 标准比对型。标准比对型要求资料精准、具有参考性和标准性, 以利于长期比对, 因而要求成果具有科学性、准确性和先进性。其对应管理层面, 可以对数据标准化、数据先进性以及在学界的引用情况等进行评价。此类项目的应用价值是引用率, 即通过对比其在应用中被引用和对比的次数, 来评估数据资源的价值。综上所述, 其评价要素集中在成果的规模和数量、成果引用次数、国内外影响力等。

3.3 评价指标体系

国家科技基础资源调查专项项目成效评价通常由数据汇交评价和综合绩效评价两部分组成。但这两部分内容相互独立, 并未形成一体, 且后者主要是通过同行评审, 缺少充分的定量指标支持。然而, 这并不是评价中最主要的问题, 其最大的难点在于以下 3 个方面: 一是现有的评价采用统一的评价标准, 未对项目特点进行分类, 这会导致不同特点和分类的项目采用一个标准尚不能充分发挥基础资源调查专项的特点和优势; 二是尚未兼顾科技基础资源调查专项影响面大、数据流动周期长和成果类型多样的特点, 也未考虑其数据论文产出、机构影响力、成果滞后效应等方面; 三是尚未形成定量与定性相结合的评价指标体系, 未来可以对各项目进行包括数据汇交和综合绩效评价在内的定量评估(作为后台数据), 与同行会议评审的定性评价共同形成完整的综合评估模式。

(1) 科技基础资源成果评价指标体系

整体评估指标包括通用指标和特色指标两

个部分。通用的指标体系包括资源汇交、资源应用、资源成果和社会影响4个方面,占分值的80%。特色指标根据其所属的资源增长型、应用导向型、标准比对型进行专项针对性打分,占分

值的20%。通用指标和特色指标如前文所述,并对其打分,分值见表1。专用指标可针对3类不同类别的项目分别设置。

(2) 指标权重确定

表1 科技基础资源成果评价指标体系

指标类型	一级指标	二级指标	详细描述	分值	评分标准
通用指标 (必选)	资源汇交评价(25%)	科技资源数量 T_{11}	科技资源汇交的数量,以及能够在开放服务体系中可以检索发现的数量	8~10	数据量远超任务书规定的数量,且按时汇交
				6~7	数据量略超过任务书规定的数量,且按时汇交
				≤5	数据量等于或未达到规定的数量,且按时汇交
		科技资源质量 T_{12}	科技资源的同行评价情况,以及在开放服务体系中用户的质量反馈情况	8~10	数据审核检查错误率小于3%,且数据覆盖空间范围完整,时间序列长。
				6~7	数据审核检查错误率小于5%,且数据覆盖空间范围较完整,时间序列较长
				≤5	数据审核检查错误率小于10%,数据覆盖空间和时间范围不完整
		科技资源规范性 T_{13}	科技资源遵循的标准规范情况,以及在开放服务中的资源描述和内容规范化程度	4~5	元数据文档、数据相关描述齐全,且数据本身符合相应规范标准
				2~3	元数据文档、数据相关描述、数据本身基本符合规范标准
				≤1	元数据文档、数据相关描述、数据本身未达到规范标准
	资源应用评价(15%)	数据共享情况 T_{21}	通过国家科学数据中心服务绩效体系,获取有关科技资源的共享利用情况,主要包括下载次数、下载量、应用证明等	8~10	数据对外开放共享,数据访问量大于3000次/年,下载量大于200次/年,且有应用证明(或用户反馈)
				4~7	数据对外开放共享,数据访问量大于500次/年,下载量大于100次/年,或有应用证明(或用户反馈)
				≤5	数据协议开放共享或暂不共享或数据访问量低于500次/年
		机构共享情况 T_{22}	通过本机构或项目组持续运营的应用服务系统中,获取其科技资源的共享利用情况	4~5	单位或机构有报道宣传,数据提供开放下载,且运行服务良好,下载量大
				2~3	单位或机构有报道宣传,且数据提供下载
				≤1	单位或机构无报道宣传,亦不提供下载
	资源成果评价(25%)	典籍志书图集等代表性成果 T_{31}	科技资源相关的成果进一步形成图集、志书、专著出版的情况	8~10	典籍志书图集等出版成果多于考核指标要求,获得相关奖项或出版基金资助
				6~7	按时完成考核指标要求的出版成果
				≤5	典籍志书图集等出版小于考核指标
		论文/数据发表 T_{32}	科技资源相关的论文发表和数据出版情况,并借助文献计量分析其科学影响力	4~5	数据论文发表多,汇交科技论文远超指标,且学术影响力大
				2~3	数据论文有发表,汇交科技论文较超指标,具有一定学术影响力
				≤1	汇交科技论文低于指标要求
		奖励情况 T_{33}	科技资源获得的各类奖励情况	4~5	获得国家级奖项
				2~3	获得省部级及以上奖项
				≤1	获得其他奖项或者无
技术先进性 T_{34}		在信息化、大数据、物联网、云计算等方面应用情况	4~5	广泛使用先进信息技术,或有显著技术创新	
			2~3	较为广泛使用先进信息技术	
			≤1	较少或未使用先进信息技术	

(续表)

指标类型	一级指标	二级指标	详细描述	分值	评分标准
通用指标 (必选)	社会影响评价(15%)	科学传播 T ₄₁	科技资源制作成为科普作品的情况, 以及其在社会上的传播和影响力	8~10	在国家级媒体或核心以上期刊传播科普作品或科学报道
				6~7	在省部级媒体或其他期刊传播科普作或科学报道
				≤5	在其他媒体报道或者未进行报道
		人才培养 T ₄₂	科技资源相关的人才培养情况	4~5	培养有院士、长江、杰青、优青等人才及大量研究生
				2~3	培养专业骨干人才及大量研究生
				≤1	培养部分专业人才及少量研究生
特色指标 (单选)	资源增长型 特色评价 (20%)	填补空白情况 T ₅₁	科技基础资源调查专项的数据增量情况	6~7	数据直接填补空白
				4~5	数据补充填补空白
				≤3	数据无填补空白
		数据时空完整性 T ₅₂	科技基础资源调查专项的时空序列完整情况	6~7	数据时空序列长且完整
				4~5	数据时空序列完整
				≤3	数据时空序列不完整
		数据先进性 T ₅₃	科技基础资源调查专项的数字化赋能能力	=6	先进技术赋能显著, 且有应用
				4~5	先进技术赋能, 但仅尝试
				≤3	采用传统非新技术
	应用导向型 特色评价 (20%)	目标针对性 T ₅₁	科技基础资源调查专项的目标需求	6~7	直接面向中央部署需求
				4~5	面向地方领域需求
				≤3	目标针对性不明确
		应用证明 T ₅₂	科技基础资源调查专项的应用案例支撑情况	6~7	有国家领导人批示
				4~5	有地方或领域应用证明
				≤3	无相关证明
		重要应用案例 T ₅₃	科技基础资源调查专项的数字化赋能能力	=6	有重要应用案例且影响大
				4~5	有应用案例
				≤3	无应用案例
	标准比对型 特色评价 (20%)	成果的规模和数量 T ₅₁	权威和参考型的成果发表的数量和质量	6~7	成果数量大且质量高
				4~5	成果数量和质量符合要求
				≤3	成果数量和质量未达标
		引用次数 T ₅₂	权威和参考型的成果被引用的次数	6~7	引用次数高
				4~5	有引用
				≤3	无引用
国内外影响力 T ₅₃		权威和参考型的成果的国内外影响	=6	成果具有国际和国内影响	
			4~5	成果具有国内影响	
			≤3	成果无国内影响	

成果评价指标权重见表2。在确定权重之前, 选择专家分别对表2的通用指标和特色指标进行重要性打分。

以表2中的资源汇交数量T11为例, 标准化后的 $Y = \{y_{11}, y_{21}, \dots, y_{n1}\}$ 如下:

$$y_{11} = [x_{11} - \min(x_{11}, x_{21}, \dots, x_{n1})] / [\max(x_{11}, x_{21}, \dots, x_{n1}) - \min(x_{11}, x_{21}, \dots, x_{n1})]$$

$$y_{21} = [x_{21} - \min(x_{11}, x_{21}, \dots, x_{n1})] / [\max(x_{11}, x_{21}, \dots, x_{n1}) - \min(x_{11}, x_{21}, \dots, x_{n1})]$$

.....

$$y_{n1} = [x_{n1} - \min(x_{11}, x_{21}, \dots, x_{n1})] / [\max(x_{11}, x_{21}, \dots, x_{n1}) - \min(x_{11}, x_{21}, \dots, x_{n1})]$$

各指标信息熵的计算如下:

$$E_j = -\frac{1}{\ln n} \sum_{i=1}^n p_{ij} \ln p_{ij} \quad (1)$$

$$p_{ij} = y_{ij} / \sum_{i=1}^n y_{ij} \quad (2)$$

根据标准化数据Y逐次利用式(2)、式(1)计算 P_{ij}, E_j 得到10类指标的信息熵。

各指标权重的计算公式确定为:

表2 科技基础资源成果评价指标确权表

指标类型	指标	X_1	X_2	...	X_n	
通用指标	科技资源数量 T_{11}	x_{11}	x_{21}	...	x_{n1}	
	科技资源质量 T_{12}	x_{12}	x_{22}	...	x_{n2}	
	科技资源规范性 T_{13}	x_{13}	x_{23}	...	x_{n3}	
	数据共享情况 T_{21}	x_{14}	x_{24}	...	x_{n4}	
	机构共享情况 T_{22}	x_{15}	x_{25}	...	x_{n5}	
	典籍志书图集代表性成果 T_{31}	x_{16}	x_{26}	...	x_{n6}	
	论文/数据发表 T_{32}	x_{17}	x_{27}	...	x_{n7}	
	奖励情况 T_{33}	x_{18}	x_{28}	...	x_{n8}	
	技术先进性 T_{34}	x_{19}	x_{29}	...	x_{n9}	
	科学传播 T_{41}	x_{110}	x_{210}	...	x_{n10}	
	人才培养 T_{42}	x_{111}	x_{211}	...	x_{n11}	
特色指标	资源增长型	填补空白情况 T_{51}	x_{112}	x_{212}	...	x_{n12}
		科技资源质量 T_{52}	x_{113}	x_{213}	...	x_{n13}
		科技资源规范性 T_{53}	x_{114}	x_{214}	...	x_{n14}
	应用导向型	目标针对性 T_{51}	x_{115}	x_{215}	...	x_{n15}
		科技资源质量 T_{52}	x_{116}	x_{216}	...	x_{n16}
		科技资源规范性 T_{53}	x_{117}	x_{217}	...	x_{n17}
	标准比对型	成果的规模和数量 T_{51}	x_{118}	x_{218}	...	x_{n18}
		科技资源质量 T_{52}	x_{119}	x_{219}	...	x_{n19}
		科技资源规范性 T_{53}	x_{120}	x_{220}	...	x_{n20}

$$W_i = \frac{1 - E_i}{10 - \sum E_i} \quad i = 1, 2, \dots, 10 \quad (3)$$

再将信息熵E转化为权重W，得到10类数据指标的权重。

(3) 定量评价

$$Z = \sum_{i=1}^{10} P_{ii} W_i s_i \quad (4)$$

邀请m位专家 (P_1, P_2, \dots, P_m) 针对表1对科技基础资源调查专项成果进行综合评价。 P_{ii} ($i = 1, 2, \dots, m$) 为各指标项的中值/均值, s_i 为各指标在表1中的满分分值, 按式(4), 计算资源科技资源成果的评价量化值Z。Z越大表示该科技基础资源调查专项成果价值越高。

$$S = Z + M \quad (5)$$

最终的项目百分制总得分S由式(5)计算得到。式中,Z为该科技基础资源调查专项成果通用指标价值;M为该科技基础资源调查专项成果特色指标价值,针对表2中的3类特色专项指标进行专家评审打分,确权和打分方式与Z类似。

(4) 综合评定

采用“定量评价+定性评价”结合法,式

(1)~式(3)为定量评价环节,该专项项目得分可记录在项目负责人的科研诚信系统中。以定量评价为基础,对科技基础资源专项项目开展定性评价,定量评价的结果可以作为定性评价的支撑。

4 结语与建议

国家创新驱动发展对原始创新提出了更大的需求,原始创新对科技基础资源的依赖程度越来越高。面向我国“十四五”规划和2021—2035年国家中长期科技发展规划,我国在生物多样性保护、粮食安全、区域高质量发展、海洋强国、生态文明建设、健康中国、“一带一路”建设等方面开展了一系列重大部署,这都对科技基础资源收集、挖掘、利用等提出了新的需求。在新形势下,科技基础资源调查面临着更多的机遇与挑战。结合对科技基础资源调查专项管理和评价模型的分析,提出以下建议。

(1) 建立科技基础资源调查专项总体调控模式。建立科技基础资源调查总体跟踪专家组,及

时跟踪和发现专项实施中的问题,并能够通过会议协调机制及时收集各方面建议。利用信息技术建立科技基础资源调查实施沙盘系统,对已完成的大气、海洋、地理、地质、生物、农业、健康、工程等各领域项目进行归纳和展示,在数据目录、空间平面和三维可视角度进行查看和分析。以上模式的作用:一是利于掌握全或布局,了解是否存在空白区;二是查看布局与目标的差异度,以及与当前关注的可持续发展目标领域的差异;三是加强薄弱区或需要对快速变化区域进行周期性的调查安排等。

(2) 建立科技基础资源调查专项项目分类管理的评价机制。针对资源增长型、应用导向型、标准比对型 3 类科技基础资源调查专项项目的不同特点,建立其定量与定性评估的记录系统,以对应跟踪其项目执行效果。把综合评价结果登记进入相关负责人、单位机构的诚信系统,为相关负责人及其研究单位的未来项目立项提供一定的诚信记录依据。

(3) 兼顾科学数据长尾效应,充分借助于文献计量分析方法,发掘科技基础资源调查成果影响力。持续分析国家科技基础资源调查专项论文产出和全球相关科研态势,形成科研脉动报告。结合成果追踪评估和文献计量分析等手段,追踪其结题 2 年内的成果,进而完成全面评估,进一步挖掘科学数据的潜在价值和应用成效。后评估方式可以使数据使用效益和论文成效效益的量化跟踪形成闭环监督,并促进成果的宣传和影响力扩大。

(4) 加大技术创新力度,鼓励科技基础资源调查专项充分利用各类大数据技术。提倡科学调查和资料整编中的新装备、新技术、新方法应用。随着信息时代的发展,传统的科技基础数据资料有了越来越多的快速增长渠道。如生态观测正从传统的、不连续的样方地面观测向“天—空—地”多尺度、多要素、多过程的综合观测转变,由单一的生态学研究人员参与向全民参与的“公民科学”模式转变^[14]。以往的传统野外调查数据增加了手机移动数据和社交媒体大数据,具

有了大数据特点的科技基础资源调查专项和数据驱动的科学研究的结合更为紧密。

致谢:感谢科技基础资源调查专项领域专家的指导和课题组成员的调研支持。

参考文献

- [1] 叶玉江. 持之以恒加强基础研究 夯实科技自立自强根基[J]. 中国科学院院刊, 2022, 37(5): 589-595.
- [2] 中央级科研院所科技基础性工作专项[EB/OL]. (2006-10-21) [2022-06-26]. http://www.most.gov.cn/ztlz/swkjhh/swkjhh/200610/t20061021_36390.html.
- [3] 科学技术部. 关于印发《国家“十一五”基础研究发展规划》的通知; 国科发计字〔2006〕436号[EB/OL]. (2006-10-23) [2022-06-26]. http://www.most.gov.cn/xxgk/xinxifenlei/fdzdgnr/gjkjgh/200811/t20081129_65771.html.
- [4] 王祎, 王卷乐, 王晓洁, 等. 我国科技基础资源调查专项实施布局分析与启示[J]. 中国科技资源导刊, 2021, 53(3): 33-43.
- [5] STOKES D E. Pasteur's quadrant: basic science and technological innovation[M]. Washington: Brookings Institution Press, 2011: 734-736.
- [6] FOUNDATION N S. Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century[R]. Alexandria, VA: National Science Foundation, 2005: 8.
- [7] 王卷乐, 王明明, 石蕾, 等. 科学数据管理态势及其对我国地球科学领域的启示[J]. 地球科学进展, 2019, 34(3): 306-315.
- [8] 王卷乐. 地球系统科学数据集成共享研究: 标准视角[M]. 北京: 气象出版社, 2015: 14.
- [9] 白燕, 杨雅萍, 王祎. 科技基础性工作专项资源环境领域项目数据汇交进展与分析[J]. 中国科技资源导刊, 2020, 52(5): 52-62.
- [10] 孙鸿烈, 成升魁, 封志明. 60年来的资源科学: 从自然资源综合考察到资源科学综合研究[J]. 自然资源学报, 2010, 25(9): 1414-1423.
- [11] 张凤雨. 志书数据问题探讨[J]. 中国地方志, 2015(3): 21-27, 63.
- [12] 王伟玲, 刘梦冉. 激发数据流通活力 释放数据潜在价值[J]. 网信军民融合, 2021(4): 17-19.
- [13] 熊巧琴, 汤珂. 数据要素的界权、交易和定价研究进展[J]. 经济学动态, 2021(2): 143-158.
- [14] 于贵瑞, 何洪林, 周玉科. 大数据背景下的生态系统观测与研究[J]. 中国科学院院刊, 2018, 33(8): 832-837.