

# 国家科学数据中心 可持续发展机制研究

黄鼎成<sup>1</sup> 李晓波<sup>2</sup> 李集明<sup>3</sup> 徐 枫<sup>4</sup> 王卷乐<sup>5</sup>

(1. 中国科学院地质与地球物理研究所, 北京 100029;

2. 国土资源部信息中心, 北京 100812; 3. 国家气象信息中心, 北京 100081;

4. 国家信息中心, 北京 100045; 5. 中国科学院地理科学与资源研究所, 北京 100101)

**摘 要:** 本文从理论分析入手, 结合科学数据共享工程实施经验, 以基地建设的思路探讨了基于可持续性的科学数据管理机构概念模型; 讨论了从国家公共财政持续稳定的投入保障、投入责任与绩效管理、政策导向与法律制度规范以及以荣誉机制、回收服务成本和一线研究参与制度为内容的协调与适应管理等方面, 建立科学数据中心的可持续发展机制。

**关键词:** 数据共享; 科学数据中心; 可持续发展; 机制

**中图分类号:** TP3 **文献标识码:** A **DOI:** 10.3772/j.issn.1674-1544.2008.05.001

## Study on Sustainable Development Mechanism of National Scientific Data Center

Huang Dingcheng<sup>1</sup>, Li Xiaobo<sup>2</sup>, Li Jiming<sup>3</sup>, Xu Feng<sup>4</sup>, Wang Juanle<sup>5</sup>

(1. Institute of Geology and Geophysics, CAS, Beijing 100029;

2. Information Center of Ministry of Land and Resources P. R. C, Beijing 100812;

3. Information Center of China Meteorological Administration, Beijing 100081;

4. State Information Center of China, Beijing 100045;

5. Institute of Geographical Sciences and Natural Resources Research, CAS, Beijing 100101)

**Abstract:** Based on the practice experiences of Scientific Data Sharing Program (SDSP) in China in the last 6 years, the paper brings forward the sustainable concept model for scientific data management organizations through theory analysis. The sustainable development mechanism of national scientific data center is designed under the framework of sustainable concept model according to the idea of long term base construction for data sharing. This mechanism is composed of 4 parts, i. e. funding mechanism, evaluation management mechanism, policy guide and bylaw, standard mechanism, corresponding & adaptation mechanism. The paper also analyzed some details in corresponding & adaptation mechanism, such as honor mechanism, data publication strategy, scientists' participant mechanism and so on.

**Keywords:** data sharing, scientific data center, sustainable development, mechanism

第一作者简介: 黄鼎成(1941-), 男, 中国科学院地质与地球物理研究所研究员, 研究方向是工程地质与岩土工程。

收稿日期: 2008年9月10日。

国家科技基础条件平台及其科学数据共享工程(以下简称“工程”)的建设将不断地推进科学数据信息资源为科技创新、政府决策、企业发展和社会公众服务,是一项持久性的工作。它必然要通过从数据管理到用户的有效途径,实现对科学数据管理与共享服务的质量、效率、效果的最优化,推进数据资源在低成本下强势流动。国家级科学数据中心是由不同领域科学数据中心组成的国家科学数据管理与共享服务体系。这项“工程”建设的管理,旨在不断提高其对科技创新的基础支撑能力,提高共享服务的质量、效率和效果,建立健全可持续发展机制,切实为中国的科技进步与创新提供高效的科学数据、信息支持。

## 1 可持续性的制约因素

科学数据是贯穿科技创新活动全过程的最活跃因素。因此,各个数据中心都是在整体科技活动中运行,是在社会中运行的一个开放系统,其正常秩序的建立,取决于彼此关联与相互作用的五大要素之间的平衡与协调,如图1所示。

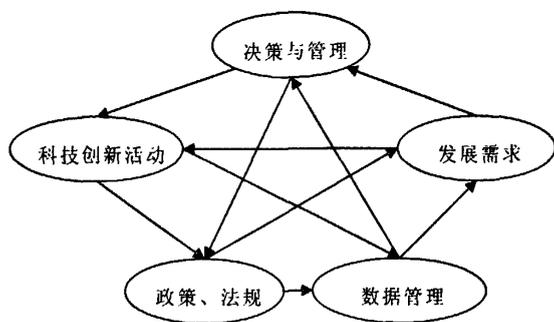


图1 科学数据管理与共享服务系统的关联要素及其相互作用

(1)发展需求。是指国家经济建设、社会发展和安全保障对科学和技术发展的需求,这一强劲驱动力最具灵性,是科技活动的源动力。它导致了国家不同时期的科技发展目标与各类科技计划的产生,决定了不同时期科技攻坚的方向与重点。

(2)决策与管理。即国家发展需求所决定的科技发展目标及其发展规划和决策者的意志,一方面表现为对某一时期科技发展的方向与重点的抉择(诸如,2005年“国家中长期科学和技术发展规划纲要”的出台和各类科技计划的调整);另一方面为了保证目标的实现,必然发动一切组织力量,包括不断增强的国家公共财政对科技的投入,利用政策导向和法律制度规范来调整和规范包括科学数据管理与共享的各种行为在内的科技活动。

(3)科技创新活动。即出成果,出人才。这不仅需要国家在人、财、物上的高强度投入,而且直接推进了与之相关的政策、法规的建设,和对包括“工程”所属的各个数据中心在内的各类科技基础条件平台与建设与发展的导向与制约(即科学数据中心建设与发展的科技创新需求导向)。

(4)政策、法规。是科技活动的政策导向以及科技活动所产生的各种社会关系和利益关系的调整与约束,是维护科技活动正常秩序,满足国家发展需求的保障。在这里强调政策的导向性与法规的约束性相结合的管理,推进科学数据共享,调整在共享过程中当事各方之间以权利和义务为内容的各种社会关系问题,规范数据管理与共享行为。

(5)科学数据管理。是以科学数据中心为具体操作平台开展的科学数据管理与服务,通过其具体数据内容服务实现对科学研究的支撑。它的运行积累又可能为管理者提供对相关发展计划的修改和资源配置决策的依据。

由以上简要分析可以看出,在对科学数据中心建设与运行的管理中,需求是导向。这就意味着,一个科学数据中心的存在及其可持续性,最根本的是它能否达到国家规划目标,它的作为能否满足科技进步与创新以及经济增长、社会发展和国家安全对相关领域科学数据、信息的需求。“大家都有用,大家都敢用”是它的立身之本,科技界的公允才有它存在与发展的空间。要能为决策者提供一个更为全面的价值理解,就能赢得国家公共财政稳定而持续的投入。而政策、法规体系则是对科学数据管理与共享行为的导向与规范。

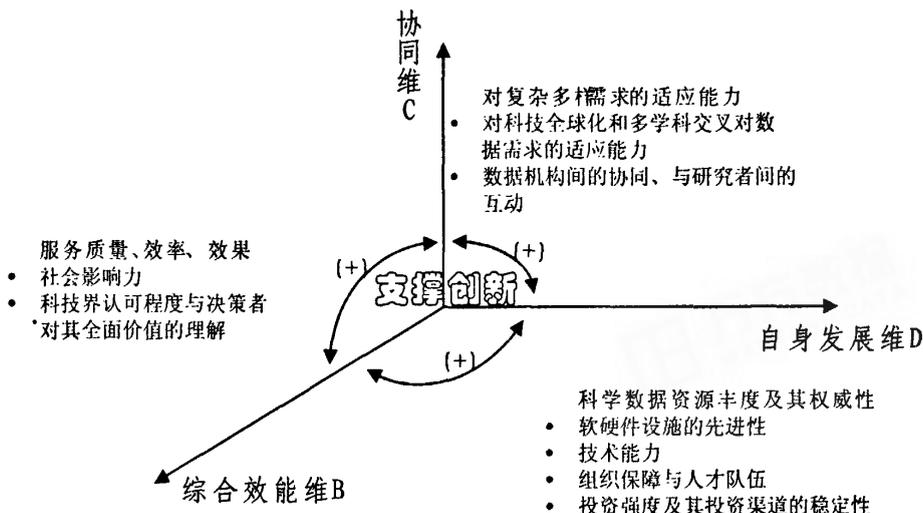


图2 基于可持续性的科学数据管理机构概念模型  
(据“科技基础设施对科技创新能力的支撑研究课题组”2007年研究报告修改)

## 2 管理模式

如前所述，作为国家科学数据管理与共享服务体系的每一个数据中心都是在整体科技活动中运行，它是围绕科技创新——出成果出人才的目标运转的。从本质上讲，每一个科学数据中心输出的结果，决定于内部对输入配置的运转能力和水平。也就是说，我们必须从整体科技活动的视角去考察数据中心与诸因素彼此关联、相互作用的协同程度，考察每一数据中心内部的协调及其对外部输入各种配置的协调运转能力与水平。科学数据中心的有效运行，必须维持其可持续性、稳定性和协调性，不断提高其自身发展的能力，协调和适应来自社会各方面科技创新活动的发展变化与不断提高的、多样的需求，从而有效地支撑与促进科技创新活动。

图2是一个科学数据中心基于可持续性概念的三维模型。是围绕支撑科技创新这个终极目标，以发展为主线，考察它为科技创新活动提供科学数据、信息资源的能力，及其与科技创新这个复杂过程协同演进的能力，包括其自身潜在的、现实的和决定性的作用力与影响力。研究认为，科学数据管理机构对科技创新支撑的成效，取决于其竞争力与功能为基础的服务和数据管理

队伍的技术水平。主要表现为面向全社会的、有效的科学数据内容服务，使研究者对科学数据、信息的需求在其中能够得到满足。对科学数据中心的数据管理与共享服务的可持续性可以从自身发展能力、协同能力和综合效能3个方面进行分析。如果它们之间能够实现正面资助或最大限度地减少负反馈，或者以最短的时间消除无序，这样的数据中心将处于稳定、协调状态，具备可持续发展的条件。

在图2中：

(1)自身发展维D，更多地体现了一个数据中心的权威性数据资源的丰度与来源的稳定性，信息技术与设施的先进性，组织体制及其队伍的技术水平与服务质量、政策执行和协调的管理机制能力及其文化氛围。实际上，自身的发展潜力或能力就是它对科技创新支撑的基本能力。

(2)协同维C，体现的是机构内部诸要素彼此关联程度及其对外部环境变化的适应与协同程度。诸如：与其他科学数据管理机构之间互联互通，协同工作，特别是数据信息资源互为整合互为利用的态状；能否适应科技创新全球化和多学科交叉融合趋势，而开展高效的数据内容服务；不断开发适应性的软件工具、模型以及与研究者互动的网络协同工作环境；能否坚持为全社会服务及其对不同用户的无歧视、平等的原则；能否

在国家安全和知识产权保护的原则下,最大限度实现资源共享。归纳起来就是适应不断变化的发展态势,满足复杂多样的需求及其高质、高效的服务。

(3)综合效能维B,则是每一科学数据中心的系统产出,是最终答卷。它通过与外部实现数据资源高效共享,维持一个良好的数据资源管理与共享秩序,达到在整个科技活动大系统中的协调平衡,通过其服务质量、效率和效果全面实现它对科技创新持续支撑的功能。

显然,3个维度之间是互为促进和彼此约束的,如:自身发展将强化其协同能力,推进共享服务水平,提高应对由于外部环境变化而可能产生风险的能力;自身发展的高潜能和对外部变化的高度适应与高质量的服务,必然使整个系统对科技创新活动产生高效支撑;因其科学数据资源的高效使用,并产生明显的科技、社会、经济效益(或效果),以其对科技创新强有力的支撑,而赢得全社会的支持,又进一步增强系统的自身发展及其与外部的协同能力。这是一个正向的良性循环,它有利于规划目标的实现,其发展是可持续的。目前数据共享工程可持续发展的关键有两个,一是各个行业数据系统的运行,是由各个行业自己支持的。二是共享,是由数据共享工程支持的,主要是支持各个行业将行业数据转化成可以共享的科学数据。且不谈行业对数据共享的阻力问题,即使行业没有阻力,也需要共享工程的支持。这是主要的。文中的3个方面没有阐述准确。当然数据的需求和共享效果也是持续发展的因素。

### 3 可持续性建设发展机制

为了保障“工程”建设的决策管理层和执行主体之间、“工程”与创新研究主体之间的良性互动,实现对科技创新产生强有力的基础支撑,建立健全可持续性机制至关重要。这并非纯技术问题,更重要的是管理问题。诸如:国家公共财政投入是实现这种支撑的必要条件,而投入责任与绩效管理机制又是投入效益最大化的保证;维护科学数据管理与共享秩序,需要政策和法律制度的

保障;充分发挥致力于创新研究的专家群体的作用,才能保证数据资源建设的正确方向;遵循统一的技术标准与考量指标,才能确保科学、完整、有序的管理与共享服务;形成区别于社会其他数据、信息工作的特色和符合时代发展的管理体制等。

#### 3.1 国家公共财政持续稳定的投入保障机制

目前,建设“工程”的重要性、必要性及其发挥的作用已有了广泛的共识。现在需要做的是以“工程”实施的实际效果与社会影响力,来赢得科技界的认可,并为投资者提供一个更为全面的价值理解,从而建立一种持续稳定的投入保障与激励机制。

实践表明,作为投资主渠道的国家公共财政持续稳定投入的方式,是与“工程”建设发展阶段有关,如图3所示。

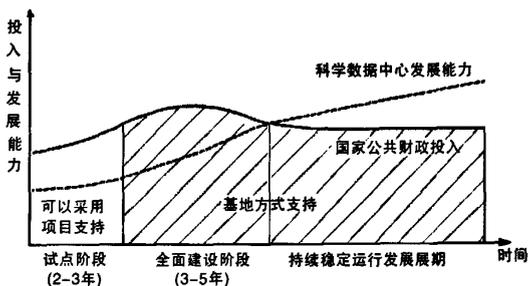


图3 科学数据中心建设、运行的国家公共财政投入与中心发展能力示意图

试点阶段一般为2~3年,国家公共财政资助可以采用项目形式管理。通过试点具备作为国家科学数据中心建设的条件,根据“工程”规划的建设标准,经专家论证进入全面建设阶段。建设阶段一般期限为3~5年,执行边建设边服务的方针,国家公共财政是以基地建设的思路进行适当规模的投入。建设任务完成后,对科学数据中心逐个进行认证评估,达到建设标准并通过认证评估者后,正式冠名为国家科学数据中心,进入持续运行阶段。持续运行阶段是持久性的,国家公共财政是从一个国家级基地角度,每年给予足额的、稳定的运行经费,主管部门根据国家科技

计划实施的需求, 下达与经费相宜的数据管理与服务任务; 数据中心更要在运行中不断提升自身发展能力和对科技创新的支撑能力, 推进科学数据集成化管理与共享服务工作的持续发展。对那些工程、技术领域的科学数据中心, 通过两个阶段的建设, 其数据管理与服务工作将实现良性循环, 有可能逐渐步入市场机制的运行模式。在运行发展期的国家公共财政投入更多地体现为授权协议与委托任务的投入。

### 3.2 投入责任与绩效管理机制

“工程”建设的思路是基地建设, 采用了以国家公共财政投入为主渠道的多元参与共建的事业型运行模式。因此“工程”及每一个科学数据中心必然要承担投入责任, 接受社会公众的监督。因此, 必须建立一种与国家公共持续稳定的投入保障机制相配套的投入责任与绩效管理机制, 以保证“工程”的效益最大化及其可持续发展。

综合分析认为, 数据的管理与共享服务绩效是考量一个数据中心规划目标的实现和投入的效益(效果)的落脚点, 是对全社会多样化需求的适应与满足程度及其可持续性的直接体现, 也是优化资金配置的依据。所以说, 国家对每一个科学数据中心的有效管理, 可以通过规划、执行、资源配置和绩效评估相互促进、互为制约的投入责任与绩效管理机制来实现, 如图 4 所示。

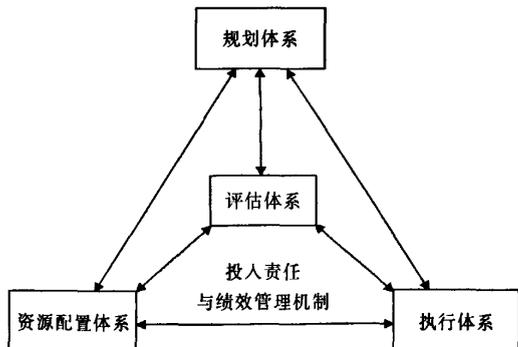


图 4 科学数据管理机构绩效评估在其管理体制中的位置

由图 4 可以看出, 体现以需求为导向的、代表国家决策意志的规划是评估的依据, 而评估的结果是为“工程”管理者、投资者提供更为全面的价值理解, 既为规划完善提供反馈, 又为资源优

化配置提供依据, 还为科学数据中心提供可持续发展的方向。通过绩效评估而优化资源配置, 实际上就是进一步明确投资者和接受投资的数据中心的责任, 降低公众的成本。在政府更好地实现为公众服务职能的同时, 使每一个数据中心的工作都能接受投资者和社会公众的监督, 而不断提升其为科技创新服务的质量、效率和效果以绩效来管理或决定投入, 投入又决定持续, 这是普通的规律。评估科学数据的绩效需要确立以满足需求的数据内容服务为导向, 建立认证评估基准、绩效评估办法与指标体系, 形成一个完善的投入责任与绩效管理机制。

### 3.3 政策导向与法律制度规范机制

“工程”在科技创新体系的建设中发挥支撑作用, 需要有一套相对完备的政策法规体系的保障, 需要尽快制定并颁布实施有关科学数据管理与共享服务的政策性文件和部门规章制度。从“工程”建设及其可持续发展角度, 在协调管理与政策法规体系制定过程中, 特别需要解决(调整)的问题是:

(1) 适当调整国家公共财政对科技投入的增长机制和投资方向(结构)。要确立科学数据共享工程的法律地位; 明确以基地思路加强这类公共平台的建设, 并以国家公共财政持续稳定的投入, 确保其可持续发展。

(2) 建立科学数据、信息等科技资源管理、共享机制。既要明确利用国家科技投入和物质条件开展的科技活动产出(固定资产、知识产权及其科学数据、信息等)的归属, 又要明释这些产出的使用、转移和管理规则。既要协调科学数据、信息在社会公有物品和市场商品之间的关系; 又要界定知识产权保护范围和社会公共功能范围。既要确定和调整共享过程中当事各方之间以权利和义务为内容的各种社会关系, 又要确立科学数据信息发布策略、办法的规则和数据共享技术的法律地位。必须指出, 数据共享最大的问题是我国没有数据共享的上位法, 缺乏法律的支持。可是, 知识产权法、保密法等都制约了数据的共享。因此, 在目前的数据共享工程建设过程中最重要的是促进立法、建立协调管理机制等。

### 3.4 协调与适应管理机制

所谓“协调与适应管理”机制,是与国家公共财政持续稳定的投入保障机制相配套、相辅相成的,以荣誉机制、回收服务成本发布策略和创新研究者参与制度为主要内容的一种管理思路。其一,由“工程”建设运行经费投入后所产生的数据库、数据分析产品和相应的软件工具,统一实施登记制度、引用制度和用户评价反馈制度;其二,各科学数据中心在基于科学数据资源的不同内容和用户的不同使用目的而采用分类分级的发布策略的基础上,可以实施服务成本收费(不包含生产成本)和定制产品的协议收费制度;其三,遵循“开放、联合”方针,建立一线研究者的参与制度。

(1) 实施登记、引用和用户评价反馈制度就是对各数据中心建设与运行所产生的各类数据(库)产品,实施统一的登记,分别赋予数据标识编码,由“工程”门户网站定期发布。借鉴科技文献引用制度,制定关于科学数据引用制度。同时,每一个科学数据中心都必须建立用户评价意见档案,形成活跃的与用户群之间的联系与合作关系,藉以提高数据质量和服务质量。显然,一个数据中心登记的数据库、数据产品和软件越多,被引用的数据量及其被引用的范围越大,引用频次越高,用户参与作用越大,那么其工作成效就越大,这样对其绩效评估也就可以有量化的表述,达到公平和透明,增强了各科学数据中心的积极性与荣誉感。

(2) 分类分级的数据发布策略服务成本回收是指科学数据中心在根据数据资源的不同内容和不同用途而制定符合实际的分类分级发布策略的基础上,建立不同的发布方式和方法。公开与共用是原则,不公开与限制是例外。基于我国当前所积累的绝大多数资源是由政府投入所获得,纳税人有权要求共享。而共享并不能简单地归结为无偿服务(指无任何收费),是要根据数据产品形态、数据产品权属与涉密和需求形态的综合分析,以求不同的发布方式。按服务范围可以分为公开发布、安全性有限发布和知识产权有限发布;按发布酬报可分为无偿服务、服务成本回

收服务和低成本(非赢利)有偿服务(或协议收费)。也就是说,以多样的发布方式方法,坚持可以收取服务成本费用,那些定制的专业性强的数据产品可以采用协议收费(低于生产成本)。这样,不会因为单纯强调无偿服务(指无任何收费)而可能对那些正在兴起的数据、信息内容服务产业的发展构成影响;并尽可能减少那些无任何委托协议机构,轻易聚积“工程”无偿提供的数据、信息资源,构成他们“寻租”的资本;同时,也使“工程”从职人员从中体味到自身的服务价值。显然,收费标准及其收入的使用规范都需要进一步研究分析,决不能违背为科技创新活动提供基础性支撑和公益服务、使从职人员从中体味到自身的服务价值的初衷。

(3) 基于“工程”的基本功能,强化一线研究者参与机制是为多学科领域一线研究者提供多样化的、能够满足大规模计算的数据内容服务。有一线研究者广泛参与,可以把握数据产品的科学内涵及其质量,从而造就“工程”的权威性数据分析产品(库)。一要强化“工程”专家委员会和各数据中心专家委员会的作用;二要聚集一线研究者参与需求分析、高端数据产品生产、软件工具开发以及对数据产品质量的把握;三要高度重视用户反馈意见;四要开展跟踪服务,建立广泛的数据联盟与合作伙伴等。

### 参考文献

- [1] 国家科技基础条件平台建设实施方案研究课题组. 国家科技基础条件平台建设实施方案研究[R]. 北京: 科学技术文献出版社, 2006: 200-249.
- [2] 黄鼎成, 郭增艳. 科学数据共享管理研究[M]. 北京: 中国科学技术出版社, 2002: 34-84.
- [3] 李新男. 关于国家科技基础条件建设的若干思考[J]. 中国科技资源导刊, 2008, 40(1): 6-12.
- [4] 国家信息工程技术研究中心、国家信息安全基础设施研究中心. 电子政务总体设计与技术实现[M]. 北京: 电子工业出版社, 2003.
- [5] 黄鼎成, 李晓波, 莫纪宏. 科学数据共享法规体系建设的若干思考[J]. 中国基础科学, 2003(6): 38-43.
- [6] 姚长青, 等. DOI在精品科技基础全文数据库中的应用[J]. 中国科技资源导刊, 2008, 40(1): 13-15.