中国科技资源导刊 ISSN 1674-1544 CHINA SCIENCE & TECHNOLOGY RESOURCES REVIEW
2010 年 9 月 第 42 卷第 5 期 37-42 ISSN 1674-1544 Vol.42 No.5 37-42，Sept. 2010

初探山西省科技计划项目科学数据的汇交管理

刘 军 范文虎
（山西省科学技术情报研究所，山西太原 030001）

摘 要：科学数据是科技计划项目成果的重要内容。科技计划项目数据汇交对于推进山西省科学数据整合共享、改善科技创新环境、提高科学数据利用率和计划项目管理水平具有重要意义。文章结合山西省科学数据共享平台的建设，通过制订数据汇交管理办法、建立数据汇交管理依托单位、将数据汇交融于科技计划项目管理流程、项目分类汇交和汇交数据的更新等 5 个方面，提出山西省科技计划项目科学数据汇交管理工作的思路，最后对汇交材料的格式进行了设计，供读者参考。

关键词：科技计划项目；科学数据；数据汇交；数据共享；科技管理
中图分类号：G352 **文献标识码：**A **DOI：**10.3772/j.issn.1674-1544.2010.05.007

Study on the Management of Scientific Data Convergence of the Projects under Scientific and Technological Plan of Shanxi Province

Liu Jun, Fan Wenhu
(Shanxi Science & Technology Information Institute , Taiyuan 030001)

Abstract: Scientific data are the important contents of the achievements of the projects under sci-tech plan. The convergence of the data of the projects under sci-tech plan is significant for promoting the integration and sharing of the scientific data in Shanxi Province, improving the environment of sci-tech innovation and raising the utilization rate of scientific data and the management level of the projects under sci-tech plan. For the construction of scientific data sharing platform of Shanxi Province, this paper puts forward some ideas on the management of scientific data convergence of the projects under sci-Tech Plan of Shanxi Province from five aspects, i.e., working out the measures for the management of data convergence, establishing the supporting institutions for the management of data convergence, integrating data convergence into the management process of the projects under sci-tech plan, carrying out the classified convergence of data and updating the convergent data. The paper also designs the format of convergence materials for reference.

Keywords: sci-tech plan, scientific data, data convergence, data sharing, sci-tech management

1 引 言

科学数据包括行业部门长期采集和管理的科学数据和各类科技计划项目产生的研究型科学数据[1]。山西省的科技计划有科技产业化环境建设、科技攻关、基础研究、国际科技合作、科技创新、软科学研究、基础条件平台等各类科技计

第一作者简介：刘军（1968- ），男，副研究员，研究方向：科技文献共享与服务平台和科学数据共享平台建设、科技情报技术。

收稿日期：2010 年 5 月 13 日。

划，在这些计划的项目研究和实施过程中，会产生大量的科学数据，包括原始性观测数据、探测数据、试验数据、实验数据、调查数据、考察数据、统计数据。这些数据具有本区域的时间和空间特性，是评价计划项目执行情况的基本依据之一。然而，针对这些数据的汇交管理，由于缺乏明确具体的管理制度，缺乏统一的数据保存和管理，大部分科技计划项目的科学数据不能汇交上来，仍然保存在基层单位或个人手中，导致数据的毁坏或濒临丢失，无法实现科学数据共享与服务，失去其科学价值和使用价值的局面。

山西省科技计划项目的数据汇交工作是实现全省科学数据共享、改善全省科技创新环境的基础工作，通过全省科技计划项目的数据汇交，可以实现对全省科学数据的统一保存管理，避免科学数据的分散保存、丢失和重复建设，充分提高科学数据的利用效率。同时，山西省科学数据也是对我国科学数据的重要补充和完善，对于下一步参与全国科学数据共享具有重要的意义。

2　数据汇交管理的思路

早在 2003 年，在国家科学数据共享工程中就颁布了《国家科技计划项目科学数据汇交暂行办法（草案）》。2008 年，国家重点基础研究发展计划（973）资源环境领域项目数据汇交工作正式启动。该项工作依托资源与环境信息系统国家重点实验室建立数据汇交中心，开展项目数据的接收、保存、管理和服务工作，取得了阶段性的进展，为我国科技计划项目的科学数据汇交工作起到示范作用。现结合山西省各类科技计划管理办法，提出山西省科技计划项目科学数据汇交管理工作的思路。

2.1　制订相应的科学数据汇交办法和制度

参照《国家科技计划项目科学数据汇交暂行办法（草案）》制订《山西省科技计划项目科学数据汇交办法》来规范每一个环节的责任义务，制订相应的汇交计划、汇交义务人、汇交数据的审核、汇交数据的保管、汇交工作的管理及汇交奖惩[2]等内容，使科学数据汇交有法可依，有章可循，作为科学数据汇交工作顺利进行的保障。

2.2　增加数据汇交管理工作依托部门

当前，在山西省科技计划项目管理中，从项目开始的计划指南发布到最后项目验收形成的归档材料大都由各计划项目归口管理的处室独自管理，项目归档的材料由处室分散保存，尤其是项目中产生和加工的科学数据，只是纸件或文字说明形式，真正的实体数据处室并没有要求上报和统一保存。因此，在下一步的科学数据汇交管理工作中，科技管理部门可以增加数据汇交管理依托部门来对全省的科学数据进行统一的管理。当前山西省科学数据共享平台建设项目由全省多家基础数据持有单位联合共建，整合的内容以参建单位现有的数据资源为主。因此，科技管理部门应当借助全省科技基础条件平台建设计划为契机，以科学数据共享平台建设为突破口，以科学数据共享平台建设牵头单位为依托，成立科学数据共享中心（简称数据中心），负责全省科学数据汇交管理，负责历史或今后科技计划项目产生的实体数据的统一保存和管理，全面开展这项工作。

2.3　将数据汇交工作纳入科技计划项目管理流程

目前，山西省科技厅系统的各类科技计划项目归口于不同的处室管理，由于各项目具有不同的特点，因此其管理流程也具有差异性，但归结起来，各项目的管理经历计划指南发布、项目申报、项目评议审查、项目立项、项目签定任务书（合同书）、项目实施、项目中期汇报（检查）、项目验收、资料归档等 9 个阶段。在这 9 个阶段中，涉及数据汇交管理的主要有申报和验收 2 个阶段，因此，可以通过对项目申报和项目验收阶段的重新部署，将数据汇交工作纳入到项目管理中，作为计划项目管理的内容，进行数据汇交的管理工作。具体的管理流程如图 1 所示。

2.3.1　项目申报阶段

（1）计划管理部门在原有的项目申报书中增加数据汇交计划[3]部分。

（2）项目承担单位向数据中心提交数据汇交计划进行审核。

（3）数据中心在审核过程中，应结合项目的情况，详细了解项目的具体研究内容，并通过本省及全国范围的科学数据共享平台，对汇交计划

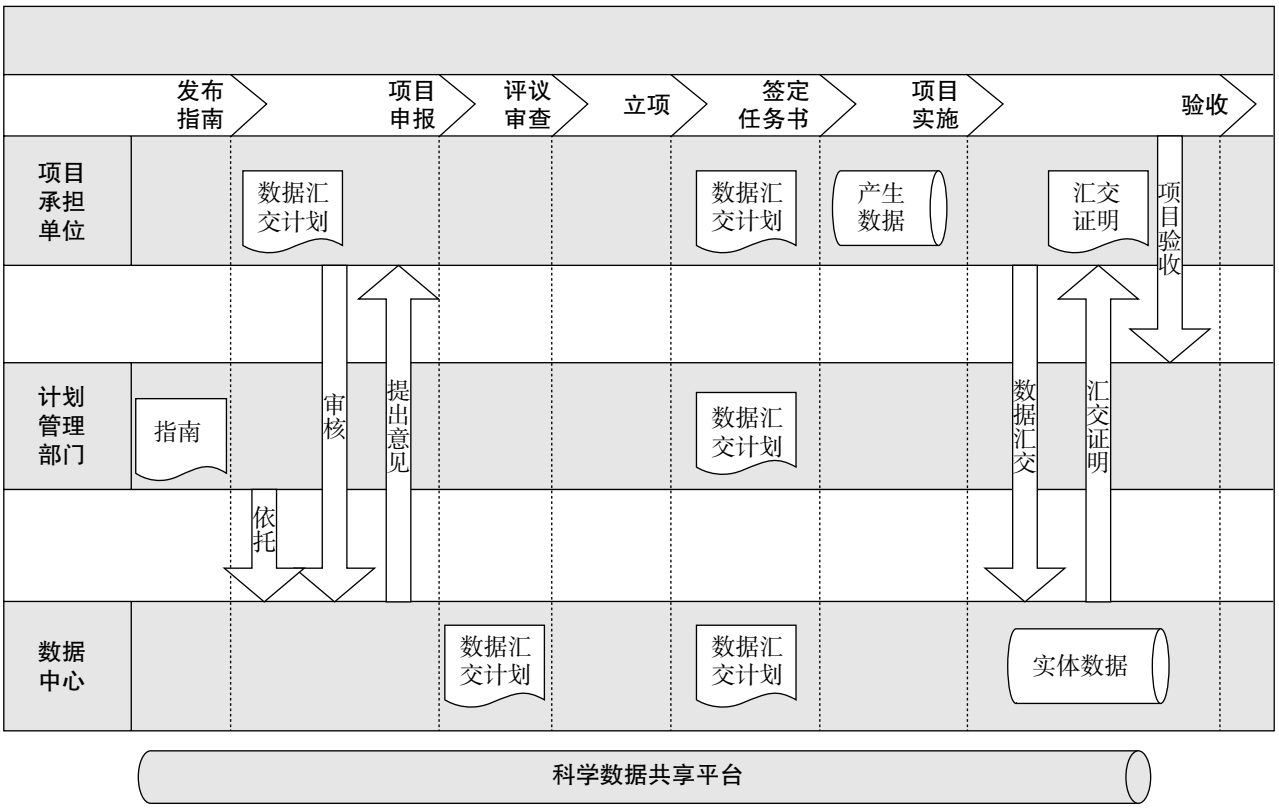

图 1　数据汇交管理工作流程

中的科学数据建设内容进行查询，对汇交计划进行审核，防止汇交计划的不全、重复、少报等现象，最后给予审核意见。审核通过的计划，项目承担单位会同其他申报材料报送计划管理部门进行项目的评议和审查。审核不通过的计划，数据中心提出修订意见，等待项目承担单位进行修订。

2.3.2　项目评议审查阶段

在计划管理部门组织的项目评议审查过程中，对于没有通过数据中心审核的数据汇交计划，将视为申报材料不完整对待。

2.3.3　签定任务书阶段

对立项的项目三方（项目承担单位、科技管理部门、数据中心）进行数据汇交任务书的签订，任务书中对于科学数据加工的考核指标根据汇交计划的内容确定，并明确项目承担单位在汇交科学数据中的各项法律责任。

2.3.4　项目实施阶段

在项目实施阶段，项目承担单位应根据数据汇交任务书的要求，对项目中产生和加工的科学数据进行整理，保证数据的真实性、完整性和质量要求。

2.3.5　项目验收阶段

（1）项目承担单位在项目验收的前期向数据中心提出数据汇交请求，并按数据中心制订的标准规范要求进行数据汇交。

（2）数据中心提供基于网络的数据汇交管理软件，并负责对所汇交的数据（元数据和实体数据）进行质量审核和检查。经审核合格后，开具数据汇交证明，作为项目承担单位的验收材料。对于汇交审核不合格的数据，不开具数据汇交证明。

（3）计划管理部门要结合数据中心开具的汇交证明材料对项目进行验收。

2.3.6　数据的保存和共享阶段

（1）所有汇交的数据统一由数据中心分类整理保存。

（2）按照各汇交数据的共享范围，通过山西省科学数据共享平台，向社会开放，实现共享。

2.4　项目的分类汇交

数据汇交工作要按照各计划项目的实施情况进行分类汇交，这里的项目分类是指已结题项目、未结题项目和新项目。对于各计划中已经结

题和未结题项目可以作为省科技计划条件平台计划中的新项目开展数据汇交，完成实体数据的整理和加工。对于新启动的项目在数据汇交管理办法颁布并实施后，按照各科技计划项目管理办法要求进行数据汇交。

2.5　汇交数据的更新

对于通过数据汇交保存到数据中心的各类数据，涉及需要持续更新的数据集，项目承担单位虽然已经完结项目，但只要后期数据集有更新，就有义务将增量数据及时汇交到数据中心，以保证数据的连续性。对于项目承担单位或数据提供单位，数据中心可根据情况支付一定的数据加工费用。

3　数据汇交相关材料的参考格式设计

3.1　数据汇交计划

作为项目申报材料的一部分，计划应明确在项目实施过程中，是否具有科学数据的产生、具体产生的数据集名称和共享范围、计划汇交的期限、汇交义务人及其通讯方式等信息。数据汇交计划参考格式见表 1。

3.2　数据汇交任务书

数据汇交任务书是项目任务书的一部分，是项目承担单位、数据中心、计划管理部门三方签订的合同。任务书中除应明确项目承担单位数据汇交的

表 1　数据汇交计划参考格式

<table>
<tr><td colspan="4">名　　称</td><td colspan="2">填写内容</td></tr>
<tr><td colspan="4">1. 本项目是否有科学数据产生：</td><td colspan="2">□是（请填写以下内容）　□ 否</td></tr>
<tr><td colspan="4">2. 科学数据产生的方式：</td><td colspan="2">□已有数据的加工整理　□项目中新采集或产生</td></tr>
<tr><td colspan="6">3. 本项目产生的科学数据集为：</td></tr>
<tr><td>序号</td><td colspan="3">名　　称</td><td colspan="2">共享范围</td></tr>
<tr><td>1</td><td colspan="3"></td><td colspan="2">□ 公开　□ 非公开</td></tr>
<tr><td>2</td><td colspan="3"></td><td colspan="2">□ 公开　□ 非公开</td></tr>
<tr><td>3</td><td colspan="3"></td><td colspan="2">□ 公开　□ 非公开</td></tr>
<tr><td colspan="6">4、以上数据集计划汇交的期限为：　年　月　日——　年　月　日</td></tr>
<tr><td rowspan="2">5、汇交义务人</td><td>姓　名</td><td></td><td>单位名称</td><td colspan="2"></td></tr>
<tr><td>电　话</td><td></td><td>手　机</td><td colspan="2"></td></tr>
<tr><td colspan="6">6、汇交义务人签字：　日期：</td></tr>
<tr><td colspan="6">7、数据中心审核意见：　签字：　日期：</td></tr>
</table>

表 2　数据汇交任务书参考格式

<table>
<tr><td colspan="6">1. 本项目完成的科学数据集为：</td></tr>
<tr><td>序号</td><td colspan="3">名　　称</td><td colspan="2">共享范围</td></tr>
<tr><td>1</td><td colspan="3"></td><td colspan="2">□公开　□非公开</td></tr>
<tr><td>2</td><td colspan="3"></td><td colspan="2">□公开　□非公开</td></tr>
<tr><td>3</td><td colspan="3"></td><td colspan="2">□公开　□非公开</td></tr>
<tr><td colspan="6">2. 相关的条款：（参考内容见后）</td></tr>
<tr><td colspan="6">3. 汇交义务人：

签字：　日期</td></tr>
<tr><td colspan="6">4. 各方意见</td></tr>
<tr><td colspan="2">项目承担单位：

签字（盖章）：　日期：</td><td colspan="2">数据中心：

签字（盖章）：　日期：</td><td colspan="2">计划管理处室：

签字（盖章）：　日期：</td></tr>
</table>

其中，相关的条款是项目承担单位、数据中心、计划管理部门三方责任和义务的说明。

数据集任务内容外，还要指定数据汇交义务人和明确三方的权利和义务，最后经各方签字盖章生效。此任务书可在原有任务书格式上进行扩展，加入数据汇交任务的内容，其参考格式见表 2。

对于项目承担单位要明确以下权利和义务：（1）项目承担单位必须选定汇交义务人，汇交义务人对汇交的科学数据有发表权、署名权、修改权、保护科学数据完整权、使用权等。（2）汇交义务人必须按照数据中心的技术规定、程序、期限保质保量完成本任务书所列的科学数据集的汇交。（3）汇交义务人对汇交的科学数据的真实性负责并承担有关法律责任。（4）汇交义务人在汇交数据前应当通过所在单位的审核。

对于数据中心要明确以下权利和义务：（1）数据中心作为山西省科技计划项目科学数据汇交管理依托部门，负有妥善保管科学数据的义务。（2）数据中心有义务敦促汇交义务人及时汇交科学数据。（3）数据中心不得非法披露非公开的科学数据或封锁公开的科学数据。（4）数据中心对项目承担单位提出的数据汇交请求，应在 5 ~ 7 个工作日内给予答复并完成数据汇交工作。（5）数据中心自收到汇交数据之日起 30 个工作日内，在山西省科学数据共享平台网站上予以公开。

对于计划管理部门要明确以下权利和义务：（1）计划管理部门负责科学数据汇交的管理与监督工作。（2）计划管理部门对于在数据汇交过程中提供虚假数据情节严重的项目承担单位，有权取消该单位再次承担计划项目的资格。

3.3　数据汇交核心元数据参考格式

核心元数据是项目承担单位向数据中心汇交的基础数据，元数据反映了承担单位已经完成和加工的实体数据，是用户了解和查询实体数据的首要信息，是数据汇交工作规范化管理的基础。其参考格式见表 3。

表 3　数据汇交核心元数据参考格式

序号	元素名称	定　　义
1	数据集名称	数据集的中英文提名
2	数据集关键词	描述数据集内容的词语或短语
3	数据集描述	对数据集内容的文本描述
4	来源的项目编号	支持数据集产生和加工的计划项目编号
5	数据集类型	主要指关系型数据库和文件系统，也可以将图像、音频、视频、软件等视为数据集。
6	数据量	数据集所包含数据量的说明（记录数和存储量）
7	数据集来源	对其他资源的参照，当前数据资源部分或全部源自这些参照资源
8	数据集提供者	提供数据资源的机构，即存储数据库所在的单位
9	数据集贡献者	对数据集内容创建做出贡献的个人或组织的名称
10	数据集更新频率	数据集在多长时间更新一次
11	数据集时间	与此数据集的创建和修改事件相关的时间
12	数据集语种	指数据集内容的语言种类
13	数据集字符集	数据集依存数据库管理系统所使用的字符集
14	数据集共享范围	指数据集共享的范围，是公开还是非公开
15	URL	数据集提供网络服务的链接地址
16	关联数据集	与当前数据集相关的其他数据集
17	数据集范围	数据集内容所涉及的分类、时间和空间范围
18	数据集负责单位	数据集分发人（组织）的联系信息
19	元数据创建日期	元数据的创建时间
20	元数据修改日期	元数据的最近修改时间
21	元数据联系信息	元数据的联系信息，在此指平台服务单位
22	元数据标准	元数据所参考的标准

其中，1–18由项目承担单位填写，19–22由数据中心补充填写。

注：元数据标准参考中国科学院科学数据库核心元数据标准V1.1版本[4]。

3.4　数据汇交接收证明

数据汇交接收证明是数据中心按照承担单位数据汇交任务书的内容，对元数据及实体数据集进行汇交，并对数据质量进行审核通过后，为项目承担单位开具的证明材料，证明材料客观地反映了数据汇交实体数据集的内容。其参考格式见表 4。

表 4　数据汇交接收证明参考格式

数据中心对XXX单位完成的省XXX计划下的XXXXXX项目的如下科学数据集内容进行了数据汇交和接收。特此证明。 数据中心（盖章）：　　日期				
1、项目汇交的科学数据集为：				
序号	名称	共享范围	记录数	备注
1		□公开　□非公开		
2		□公开　□非公开		
3		□公开　□非公开		

4　结　语

科学资源已成为提高自主创新能力的基础条件和有力支撑[5]。科技资源的科学数据的汇交工作是逐步优化山西省科技计划项目管理的重要工作内容，是实现全省科学数据共享、改善全省科技创新环境的重要手段之一。通过全省科技计划项目的数据汇交，实现对全省科学数据的统一保存管理，避免科学数据的分散保存、丢失和重复建设，能够充分提高科学数据的利用效率和科技计划项目管理水平，对于本省科学数据共享平台的持续发展及参与全国范围的科学数据共享具有重要作用。同时，数据汇交工作的进行不仅需要管理办法的约束，更在于管理办法的认知和执行力度，只有这样才能保障科学数据共享平台建设的长效进行。本文通过借鉴国家有关部门数据汇交工作的经验，提出山西省科技计划项目科学数据汇交管理工作思路，许多细节还需成熟和深酌。在此只愿能抛砖引玉，以促进山西省科技计划项目科学数据汇交工作的尽早开展。

参考文献

[1] Lu Peng,Miao Liangtian,Li Zhixiong,et al. Investigation on Scientific Data Sharing Situation and Its Analysis[J]. Earthquake,2007, 27(3):125−130.
〔路鹏，苗良田，李志雄，等．“973”计划资源环境领域数据汇交进展与数据分析 [J]. 地震 ,2007,27(3): 125−130.〕

[2] Lu Peng,Miao Liangtian,Mo Jihong,et al. Speculation of Law System in the National Science and Technology Plan of Scientific Data Remit[J]. Recent Developments in World Seismology,2007(10): 28−31.
〔路鹏，苗良田，莫纪宏，等．国家科技计划项目科学数据汇交的法律制度建设构想 [J]. 国际地震动态 ,2007(10): 28−31.〕

[3] 科学数据共享工程门户网站．国家科技计划项目科学数据汇交暂行办法（草案）[EB/OL]. [2010−02−05]. http://www.sciencedata.cn/fagui.php.

[4] 中国科学院科学数据库网站．科学数据库核心元数据标准 [EB/OL].[2007−10−23]. http://www1.csdb.cn/prohtml/0.projects.standard/pages/0014.html.

[5] Xu Guanhua. Strengthening the Research on S&T Resources and Promoting Its Sharing[J]. China Science & Technology Review, 2008(3): 3−5.(in Chinese)
〔徐冠华．加强科技资源研究　促进科技资源共享 [J]. 中国科技资源导刊，2008(3): 3−5.〕

（上接第 5 页）

Journals[J].Wuhan University Journal: Humanity Sciences, 2009, 62(5): 635−638.（in Chinese）
〔程郁缀，刘曙光．防止学术不端和期刊的责任 [J]. 武汉大学学报：人文科学版，2009, 62(5): 635−638.〕

[12] Ioannidis，John P A. Why Most Published Research Findings are False[EB/OL]. [2010−07−29]. http://www.plosmedicine.org/article/info%3Adoi%2F12.1371%2Fjournal.pmed.0020124.

[13] Hu Liangping.The Main Elements Affecting the Quality of Bio−Medical and Clinical Research Projects[J].Chinese Journal of Pharmaceutical Technology Economics & Management, 2010(4): 72−74.（in Chinese）
〔胡良平．影响生物医学和临床科研课题质量的要因 [J]. 中国医药技术经济与管理，2010(4): 72−74.〕

[14] Chang Simin.The Bad Behaviors of Academic Morality in Sci−Tech Papers and Editorial Control[J].Journal of Henan Agricultural University, 2004, 38(3): 346−349.（in Chinese）
〔常思敏．科技论文中的不良学术道德行为及编辑控制．河南农业大学学报，2004,38(3): 346−349.〕