

农业科学数据共享初探

刘 茜

(中国农业科学院农业信息研究所, 北京 100081)

摘要: 作为重要的科学资产和公共信息资源, 科学数据正在爆炸式的增长。因此, 整合科学数据, 完善科学数据管理, 实现科学数据共享具有重要的意义。通过对国家农业数据共享中心平台进行分析, 提出促进科学数据共享的政策, 认为科学数据的共享离不开国家相关法律法规的制定和科研资助机构相关政策的扶持。应充分重视科学数据共享的战略地位, 解决科学数据产权保护的问题, 提升科研人员的共享意识, 培养科学数据管理的专业人才, 探索科学数据共享的增值模式。

关键词: 农业; 科学数据; 科学数据共享; 科学数据管理; 知识产权

中图分类号: G350

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2016.04.007

Preliminary Study to Promote the Sharing of Agricultural Scientific Data

LIU Qian

(Agricultural Information Institute, CAAS, Beijing 100081)

Abstract: As an important scientific asset and public information resources, scientific data is explosive growth. Therefore, it is important to integrate scientific data, improve the scientific data management, and achieve scientific data sharing. The paper analyzed the AgriData, we found the measures to promote the sharing of scientific data under the background of big data. We found that scientific data sharing is inseparable from the development of national laws and regulations, and research funding agencies' policy. We should pay full attention to the strategic position of sharing scientific data, solve the problem of protection of property rights of scientific data, enhance the shared awareness of researchers, train professionals of scientific data management, and explore value-added model of scientific data sharing.

Keywords: agriculture, scientific data, scientific data sharing, scientific data management, intellectual property

1 引言

作为重要的科学资产和公共信息资源, 农业科学数据正在爆炸式地增长, 科研人员和广大农民对农业科学数据的需求日益迫切。我国是农业资源大国。长期以来, 积累了非常丰富的科学数

据资源, 但大部分没有进行有策略的管理, 无法进行有效的共享, 从而造成了巨大的数据资源浪费。相比其他学科, 农业科学数据的研究也相对较少。以中国知网为例, 我国近5年有关科学数据的研究论文有2600多篇, 但与农业科学数据相关的研究论文只有100多篇。因此, 开展农业

作者简介: 刘茜(1986—), 女, 中国农业科学院农业信息研究所助理研究员, 研究方向: 科学数据管理与共享。

基金项目: 中国农业科学院农业信息研究所2016年度科技创新工程项目“农业大数据挖掘与分析关键技术及平台构建”(CAAS-ASTIP-2016-AII); 国家社会科学基金项目“科学数据用户相关性标准与使用模式实证研究”(14BTQ056)。

收稿时间: 2016年4月5日。

科学数据共享研究, 加快打破农业科学数据使用屏障, 是国家发展战略的必然需要, 对增强我国农业整体科技水平、提高农业科技创新能力、促进经济发展具有重要意义。

科学研究越来越依赖于科学大数据的使用, 联合国教科文组织界定了三种开放存取的信息资源, 即科学出版物、教育资源、科学数据。由此可见, 科学数据位列其中^[1]。2002年, 我国启动科学数据共享工程, 希望建设面向全国的科学数据共享服务体系。在这个背景下, 2005年国家农业科学数据共享中心平台开始建设。经过10年的建设与发展, 取得了显著的成效。本文将国家农业科学数据共享中心平台为例, 整合农业科学数据, 完善农业科学数据管理, 对实现农业科学数据共享进行初步探讨。

2 国家农业数据共享中心平台概况

由中国农业科学院农业信息研究所主持, 中国农业科学院部分专业研究所、中国水产科学研

究院、中国热带农业科学院等单位参加的国家农业科学数据共享中心(以下简称“中心”)于2005年正式启动建设, 2009年通过科技部和财政部的开放共享评议, 2011年通过科技部、财政部联合评审, 成为首批认定的23个国家级科技平台之一。该中心的主页如图1所示, 其向全国科研院所、大专院校、政府部门、企业和社会公众提供农业科学数据共享服务, 为农业部重大专项、农业各级各类科技计划(项目/课题)提供了必要的数据支撑。中心通过门户网站www.agridata.cn, 向社会提供24小时×365天的不间断在线数据共享服务, 并开通了在线客服, 提供在线技术咨询、线下数据共享服务等。

2.1 科学数据资源建设

农业科学数据类型多样, 专业门类多跨度大, 并且相对分散, 整合困难很大, 因此, 中心采用学科分类的资源整合策略, 建立了作物科学、动物科学和动物医学、农业科技基础、渔业与水产科学、热带作物科学、草地与草业科

AgriData
国家农业科学数据共享中心

全部 [v] 搜索 高级搜索

发布数据 发布需求

资源总量: 3217.95GB
覆盖学科: 12大类
总下载量: 698.27GB
总访问量: 1448632次
科研团体: 1302个
专家学者: 36名
中外联盟: 18个
注册用户: 25590人

新增数据	更多...	新增需求	更多...
中国粮食生产情况数据库	2014-06-27	求农资农具方面的销量数据	2016-04-30
中国畜产品生产数据库	2014-06-27	求农业科技统计的相关数据	2016-04-24
社会主义新农村信息化案例数据库	2014-06-27	全国葡萄品种种植面积数据极其...	2016-04-16
中国县域农业产业结构数据库	2014-06-27	农业成本需求	2016-04-10
世界农畜产品贸易量排名数据库	2014-06-27	我想要 作物资源图像相关的数据	2016-03-02
世界农畜产品生产量排名数据库	2014-06-27	求作物物种分布数据	2016-02-21
植物基因组数据库	2014-06-27	我想要 中国地方猪品种数据	2016-02-11
粮食作物功能基因数据库	2014-06-27	沿海11省市2004-201...	2016-01-23

图1 国家农业数据共享中心主页

学、农业资源与环境科学、农业区划科学、农业微生物科学、农业生物技术与生物安全、食品工程与农业质量标准、农业信息与科技发展等 12 大类的学科资源整合框架。目前，资源总量达到 3217.95GB。已整合的农业科学数据资源几乎涵盖了农业各个学科领域。其中，重点学科资源整合总量占国内总量的 85% 以上，其余学科的资源整合量占国内总量的 60% 以上。这些数据包括历史珍贵资料、实地调查和科研数据等，内容真实可靠，具有较大的科研价值。

2.2 科学数据共享标准规范

根据农业行业科学数据资源整合与共享的实际需要^[2]，中心制订了农业科学数据共享管理办法、农业科学数据检查与质量控制管理办法、农业科学数据汇交管理办法等 4 个农业科学数据共享管理办法，数据制作、数据组织、数据管理和共享的标准规范 10 余个，以及农业专业领域的标准规范 50 余个。这些标准规范是保证中心工作顺利推进的基础和保障。中心科学数据共享标准规范具体分类如图 2 所示。

(1) 框架体系：框架体系定义了一系列的标准规范，目的是为了说明各个标准的应用范围、作用和相互关系，并建立了参考模型和术语等标准。

(2) 数据管理：包含一系列的数据标准，例如：数据加工流程、数据分类标准、数据采集、数据组织、质量控制和数据维护等。

(3) 数据制作：包含一系列的数据加工流程

的操作标准和规范，例如：元数据标准、数据标引规范、数据著录标准和数据表示规范等。

(4) 数据服务：包含农业科学数据发布、交换和共享方面的相关标准，例如：农业科学数据的转换格式和方法、互操作方法和规则、用户认证、数据库性能监督和改进等。

2.3 科学数据资源整合模式

面对农业大数据资源量大面广、类型多样、分散性强等特点，农业科学数据共享中心组建了由“数据主中心—数据分中心—数据节点”三个层次组成的数据整合工作体系^[3]，加强数据整合过程中各个环节的组织与管理。

在具体的操作层面上，明确数据管理者、拥有者、使用者的责任和义务，建立合理科学的数据整合汇交机制。第一，按照农业学科的相应特点和数据共享的需要，制定了农业科学数据汇交管理办法，将汇交责任和义务进行明确说明；第二，按一定机制将数据所有者（单位或个人）组织起来，如服务协议方式、共建共享方式、项目合同方式等，形成农业科学大数据的汇交责任关系，进一步明确数据汇交人的义务、数据种类、数据格式、数据质量等内容；第三，按照中心统一的标准规范，数据所有者应对科学数据进行规范化的整理和加工。通过农业科学数据汇交管理系统，提交数据和文档，包括数据源、元数据、数据字典和数据使用说明等；第四，对于上交的数据库，中心进行数据审核和验收，并进行二次标引和数据融合，最终进行网络化组织和发布。

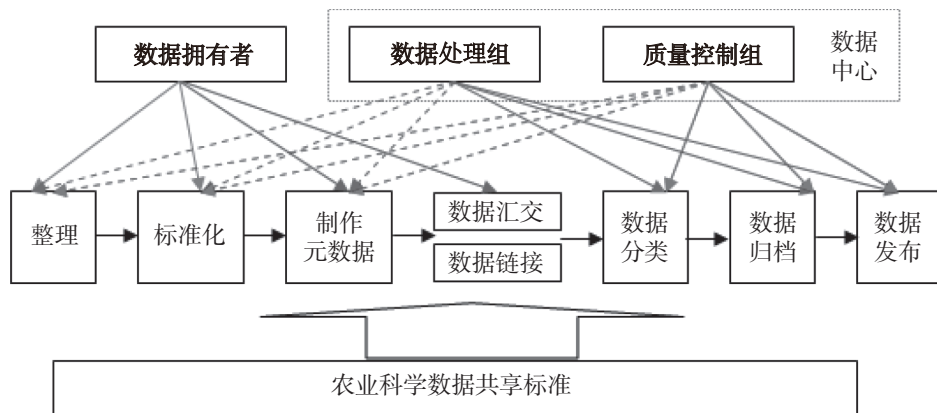


图 2 国家农业科学数据共享中心数据处理业务流程

2.4 国家农业数据共享中心服务成效

(1) 科技支撑

国家科技计划项目是农业科学数据的主要生产源头,是农业科学数据共享中心数据整合与数据服务的重点。为此,农业科学数据共享中心借鉴“学科馆员”的理念,建立了与重大和重点农业科研项目的对接机制,要求农业科学数据共享中心的每个参加单位都要与本学科领域的科研项目实现有效对接。通过跟踪国家重大科技计划,如“973”计划、“863”计划等,提供项目论证、立项、执行等各环节的数据支撑服务。这不仅有助于提高项目的科学性、执行的高效性和产出成果的创新性,而且可以直接吸收项目研究产生的最新科学数据,通过共享让更多的科研人员受益。

(2) 社会效益

农业科学数据共享中心积极为政府决策提供支持,在促进我国国民经济发展中发挥了应有的作用,取得了良好的社会效益。在已开展的水稻、玉米、小麦、大豆、油菜、棉花、柑橘、苹果、生猪、奶牛等10个产业技术体系建设中,农业科学数据共享中心为水稻、玉米、小麦、大豆、油菜、棉花、柑橘等产业研究建立了专题数据集,为相关产业体系的建设提供了重要的数据支撑。在国家农业科学数据中心和重点实验室体系之间建立快速信息通道,为全国农业科技创新活动提供数据支撑,取得了良好的社会效益。

3 我国农业科学数据共享存在的问题

经过对国家农业数据共享中心平台的科学数据资源建设、科学数据共享规范和科学数据整合模式的分析,以及对中心多年管理经验的总结,并结合合作方和用户反馈的意见,提出如下几点目前我国农业科学数据共享面临的问题。

(1) 农业科学数据的采集分散、数字化程度较低

目前,我国农业科学数据的采集、管理和维护主要是通过政府投资完成,数据资源分布分散于国家科研机构、农业高等院校和农业企业之

中。在现有的管理体制下,这些分散的农业科学数据资源归部门所有,数据的积累、管理和应用相对滞后。特别是有些数据资料长期分散在部门、单位或个人手中,不但难以充分发挥作用,而且还将面临散失的风险。同时,我国农业科学数据的数字化程度较低,许多长期积累的历史资料、古籍等珍贵历史数据尚未进行规模化的数字化处理。

(2) 农业科学数据共享的积极性不高

虽然国家农业数据共享中心改变了过去农业科学数据管理集中度较低的问题,整合了全国的优秀农业科学数据资源,但负责相应学科的单位或部门对科学数据上传和更新的积极性并不高。这是由于农业科研人员的科学数据共享意识还比较淡薄,同时科学数据共享对科学数据的知识产权保护存在缺失,这就大大降低了科研人员共享科学数据的积极性。

(3) 农业科学数据共享服务形式单一

目前,国家农业科学数据共享中心主要向用户提供农业科学数据的查询和下载,对平台中的科学数据进行分析的服务相对欠缺。对于平台中的农业科学数据进行数据挖掘、数据分析和数据可视化处理不够,没有向用户提供及时有效的数据分析结果,这就增加了不同用户对同一数据的重复性分析,从而加大了工作强度,降低了科学研究的效率。

4 促进农业科学数据共享的对策建议

4.1 加强农业科学数据共享法律法规体系建设

农业科学数据共享需要国家制定相关法律法规予以许可和在一定程度上限制,对于国家或地方政府支持的非营利机构,只要所开发的农业科学数据不危及国家安全和核心产权利益,就应该向社会开放,但需要说明的是,这些科学数据不能用于以营利为目的的商业活动。同时,为尊重科研人员的科研劳动,使用这些科学数据时应加以标注。对于由企业、单位自筹或个人开发的农业科学数据,一般应采取科学数据共享收费的方式,以保障他们的利益,提高其数据共享的积

极性。

但目前在我国还没有专门针对科学数据共享的相关法律。作为科技领域的基本法律,《中华人民共和国促进科技成果转化法》和《中华人民共和国科学技术进步法》并没有对科学数据共享做出具有针对性的规定,其法律原则只能对科学数据共享活动进行指导^[4]。而欧美发达国家的政府部门都已制定了针对科学数据管理和共享的法律法规,确立了明确的科学数据的管理规范。因此,制定科学数据共享的法律法规,是科学数据共享的必要要求,是保障科研人员利益、提升科学数据共享积极性的重要一环。只有加强科学数据共享法律法规体系建设,国家农业数据共享才能有法可依,才能有效保障农业科研人员的利益,提升农业科研人员共享科学数据的积极性。

4.2 加强农业科学数据管理,培养农业科研人员的科学数据共享意识

农业领域内的科研机构、数据中心和科学团体等公共部门是产生和拥有农业科学数据的重要主体,但由于学科划分、部门利益的驱使,以及农业科研人员共享意识的淡薄,农业科学数据共享存在诸多屏障。在整个科学数据共享的体系中,科研人员既是提供科学数据的主体,也是科学数据使用的主体。在科研过程中,只有提升科研人员的科学数据共享意识,科研人员才能认真进行科学数据管理,进而实现高质量的科学数据共享。

当然,促进农业科学数据共享,单纯依靠农业科研人员共享意识的提升是远远不够的。所以,科学数据的共享与完善的科学数据管理是密不可分的,制定完备的科学数据管理政策是科学数据共享的保障。从国外发达国家的科学数据共享经验来看,美国、英国和澳大利亚等国的科研资助机构和高校都要求科研项目申请者,在提交项目申请时,附带相应的科学数据管理计划,目的是能够有效管理和保存科研过程中产生的科学数据^[5]。而我国的科研资助机构和高校的科研项目申请恰恰没有相关的规定,使得我国科研人员缺乏保存和管理科学数据的意识,科研过程中产

生的大量实验数据没有进行有效的管理,致使科学数据无法共享,其价值无法发挥,也不利于科研成果的继承和积累,严重影响了我国科研投入的效益。

因此,借鉴发达国家的经验,我国的农业科研资助管理机构应该要求项目申请者在提交项目申请的同时提交相应的农业科学数据管理计划,并将其纳入项目申请和执行情况的考核指标之中,使得农业科学数据能够得到有效管理,进而促使农业科研人员进行农业科学数据共享。

4.3 探索农业科学数据增值利用模式

农业科学数据共享不但可以服务于农业科学研究,也可以服务于企业创新和社会人才培养。通过推进拥有农业科学数据的不同主体进行农业科学数据共享,探索跨学科、跨领域、跨部门的科学数据整合和共享,进而达到农业科学数据增值利用的目的。

首先,国家和政府部门应主动开放农业科学数据共享。在大数据时代,国家和政府部门掌握大量的农业科学数据,应该主动对这些农业科学数据进行开放共享^[6],积极开放农业科学数据二次开发的标准应用接口,鼓励单位、企业和个人充分利用政府开发的农业数据资源,使农业科学数据创造出全新的、更大的价值。

其次,农业科学数据是具有巨大增值潜力的数据资源。我国国家公共部门的科学数据主要集中于地理、交通、国土、气象、农业、公共卫生等领域,而欧盟的《欧洲公共信息部门开发潜力》研究报告显示,科学数据的增值领域主要集中在地理、交通和气象领域^[7]。因此,在满足国家对农业科学数据需求、坚持公益性原则的同时,应大力发展农业科学数据的增值利用,探索市场化运行的服务机制,实现农业科学数据增值利用的可持续发展。对此,我国可以借鉴美国、欧盟等发达国家的成功经验,鼓励企业或个人参与增值开发,满足多样性的社会需求。对以营利为目的的企业或个人,其开发和增值利用农业科学数据应探索有偿服务的模式,实现商业化运

(下转第56页)

- [7] 赵勇. 中小企业专利信息利用中的障碍研究[J]. 现代情报, 2004(11): 194-197.
- [8] 王根, 周斌. 地方专利信息服务平台建设研究[J]. 情报探索, 2014(12): 60-64.
- [9] 郭强. 基于区域产业集群的专利信息服务平台建设研究[J]. 科技进步与对策, 2013(10): 59-62.
- [10] 何铁宝. 企业专利信息分析与专利战略工作[J]. 企业科技与发展 2009(20): 16-17.
- [11] 刘启强, 夏立新. 试析我国企业专利信息咨询服务[J]. 中国信息导报, 2007(1): 45-47.
- [12] 周洁. 专利信息利用对我国企业发展的作用研究[J]. 江苏科技信息, 2015(8): 11-13.
- [13] 林映州. 企业专利信息运用的研究与实践[J]. 广东科技, 2008(1): 27-28.

(上接第45页)

营, 形成农业科学数据产业, 开发农业科学数据产品。

4.4 培养农业科学数据管理的专业人才

农业科学数据共享是一项长期的工作。在大数据时代的今天, 面对海量的农业科学数据迫切需要建立专业的农业数据管理人才队伍。为了提高科学数据管理人才的专业水平, 一些国外的图书情报教育机构已经推出了科学数据管理的相关培训, 并在大学设立了相关课程。美国已经有 54 所大学开设了数据管理的相关课程, 共 202 个教育计划, 涉及 476 门课程^[8]。

目前, 我国图书情报教育机构还没有设立科学数据管理的相关专业。面对新形势的挑战, 高校和研究机构的图书馆应该把握科学数据共享的发展方向, 适应科学数据管理发展的新需求。调整高校图书情报专业的课程设置和人才培养方案, 增加科学数据管理的专业方向, 培养专业的农业科学数据管理人才。

5 结语

本文以农业科学数据共享中心为例, 分析了当前农业科学数据共享存在的问题, 提出了促进农业科学数据共享的对策。即, 要加强科学数据共享法律法规体系建设, 加大科研资助机构相关政策的扶持力度; 要重视科学数据共享的战略地位, 加强科学数据管理, 培养科研人员的科学数据共享意识; 要积极探索科学数据增值利用模

式; 要努力培养科学数据管理的专业人才。通过提升我国农业科学数据共享的水平, 发展农业科学数据共享的新模式, 形成我国农业科学数据共享的长效可持续发展的运行机制, 从而增强我国整体科技水平, 提高科技创新能力, 促进经济整体发展。虽然农业科学数据共享中心的建设, 在我国农业科学数据的管理与共享方面取得了一定的成就, 但今后需要做的工作和努力还有很多, 例如科学数据分析、科学大数据挖掘等问题都需要进一步思考并加以解决。因此, 今后对于农业科学数据共享的研究将会是农业科研领域的重要议题, 贯穿于整个农业科技活动的始终。

参考文献

- [1] 黄如花, 王斌, 周志峰. 促进我国科学数据共享的对策[J]. 图书馆, 2014(3): 7-13.
- [2] 白如江, 冷伏海. “大数据”时代科学数据整合研究[J]. 情报理论与实践, 2014(1): 94-99.
- [3] 杨友清, 陈雅. 科学大数据共享研究: 基于国际科学数据服务平台[J]. 新世纪图书馆, 2014(3): 24-28.
- [4] 王晴. 论科学数据开放共享的运行模式、保障机制及优化策略[J]. 国家图书馆学刊, 2014(1): 3-9.
- [5] 司莉, 邢文明. 国外科学数据管理与共享政策调查及对我国的启示[J]. 情报资料工作, 2013(1): 61-66.
- [6] 孟小峰, 慈祥. 大数据管理: 概念、技术与挑战[J]. 计算机研究与发展, 2013(1): 146-169.
- [7] Europe's public sector information[EB/OL]. [2014-01-01]. http://ec.europa.eu/information_society/policy/psi/docs/pdfs/brochure/psi_brochure_en.pdf.
- [8] 黄如花, 邱春艳. 国外科学数据共享研究综述[J]. 情报资料工作, 2013(4): 24-30.