

# 重要开放农业科学数据资源建设现状综述

柴苗岭<sup>1,3</sup>, 黄琳<sup>2</sup>, 任运月<sup>2</sup>

(1. 中国科学院成都文献情报中心, 成都 610041; 2. 四川大学公共管理学院图书馆学系, 成都 610065;  
3. 中国科学院大学经济与管理学院图书情报与档案管理系, 北京 100049)

**摘要:** [目的/意义] 系统梳理了重要开放农业科学数据资源现状与特征, 为“数字政府”管理使用科学数据提供借鉴与参考。[方法/过程] 确定农业科学数据的概念和内涵, 从资源类型及数据格式、学科范围、知识组织方法和服务模式 4 个维度进行调研和分析。具体从政府、农业科研机构与大学、农业图书情报机构网络调研入手, 选择重要的农业科学数据资源网站/平台进行检索, 筛选 165 篇中文文献, 167 篇英文文献作为补充参考来源。[结果/结论] 建议基于农业科技宏观和中观管理需求, 综合发挥科研机构和大学、图书情报机构、政府和企业的资源优势 and 学科优势, 采集和集成多类型、多学科的异构异源的科学数据资源, 加强数据内部的关联和融合, 为数字政府转型提供支持。

**关键词:** 农业; 科学数据; 综述; 开放数据; 知识组织; 服务模式

**中图分类号:** G250      **文献标识码:** A      **文章编号:** 1002-1248 (2020) 10-0025-10

**引用本文:** 柴苗岭, 黄琳, 任运月. 重要开放农业科学数据资源建设现状综述[J]. 农业图书情报学报, 2020, 32(10): 25-34.

## A Review of Construction of Major Agricultural Open Scientific Data Resources

CHAI Miaoling<sup>1,3</sup>, HUANG Lin<sup>2</sup>, REN Yunyue<sup>2</sup>

(1. Chengdu Library and Information Center, Chinese Academy of Sciences, Chengdu 610041; 2. Sichuan University, School of Public Administration, Department of Library Science, Chengdu 610065; 3. University of Chinese Academy of Sciences, School of Economics and Management, Department of Library Information and Archives Management, Beijing 100049)

**Abstract:** [Purpose/Significance] This paper systematically reviews the current situation and characteristics of major agricultural open scientific data resources, so as to provide references for a "digital government" to manage and use

收稿日期: 2020-09-08

基金项目: 四川省科技计划国际合作项目“面向知识服务的农业科学数据和科技文献关联融汇服务系统研究与示范”(2017HH0094)

作者简介: 柴苗岭 (OCRID: 0000-0002-5047-2167), 博士, 副研究馆员, 中国科学院成都文献情报中心, 研究方向为资源建设、知识组织、知识产权情报研究。黄琳 (OCRID: 0000-0002-2807-3944), 硕士, 四川大学公共管理学院图书馆学, 研究方向为信息服务研究。任云月 (ORCID: 0000-0002-2662-268X), 硕士, 四川大学公共管理学院图书馆学, 研究方向为信息管理方法与应用

scientific data. [Method/Process] The concept and connotation of agricultural scientific data were determined, and the investigation and analysis were carried out from four dimensions: resource type and data format, the scope of a subject field, the method of knowledge organization and the service mode. Based on a survey of government agents, agricultural research institutions, universities and agricultural library and information institutions, this paper investigated important agricultural science data websites and platforms constructed by those organizations, and selected 165 Chinese papers and 167 English papers as supplementary references. [Results/Conclusions] Based on the needs of macro and meso management of agricultural science and technology, it is suggested to give full play to the role of scientific research institutions, universities, libraries and information institutions, governments and enterprises in terms of their advantages in resources and specific subject fields, collect and integrate multi-type and multi-discipline heterogeneous scientific data resources, strengthen the internal association and integration of data, and provide support for the transformation of a digital government.

**Keywords:** agricultural; scientific data; review; open data; knowledge organization; service model

## 1 引言

科学数据在解读上有所不同,但总体是围绕科学活动产生的数据进行定义。如张莉等认为科学数据是从事农业科技活动产生的原始性和基础性数据,以及按照不同需求系统加工后的数据集和相关信息,既包括农业及相关部门长期积累的大规模调查、观测、探测和试验等工作所获得的海量科学数据,也包括广大农业科研人员在研究工作中所产生的大量科学数据<sup>[1]</sup>。2018年3月发布的《国务院办公厅关于印发科学数据管理办法的通知(国办发〔2018〕17号)》认为“科学数据主要包括在自然科学、工程技术科学等领域,通过基础研究、应用研究、试验开发等产生的数据,以及通过观测监测、考察调查、检验检测等方式取得并用于科学研究活动的原始数据及其衍生数据”。

随着大数据技术的日趋成熟,开放获取运动的推广和发展,在农业领域,农业科学数据作为战略性、基础性的科技资源越来越受到重视。利用农业科学数据提供知识发现,指导农业生产发展、经营监管以及促进科技创新成为农业科技管理的重要需求。另外,当下中国数字政府建设及管理正处于关键期,而“数字政府”的共同机理是,借用现代信息技术手段搭建

基础信息平台,通过全方位收集、整理、分析、存储和输出信息,确保政府在全面、迅速、及时、准确掌握信息的基础上进行科学决策和高效管理<sup>[2]</sup>。由此可见农业科学数据的管理也同样是数字政府建设管理不可或缺的部分。在这一环境下,了解和学习各国开放农业科学数据的应用和复用将为中国农业科技管理提供借鉴和参考。其次,围绕国内外政府及研究机构建设的开放农业科学数据资源平台,研究其特征和功能,对利用图书情报机构的资源优势 and 学科优势,参与到科学数据管理服务资源平台建设中也有重要的参考意义。

## 2 国内外研究现状及调研方法

### 2.1 国内研究现状

自2001年底中国国家科学数据共享工程启动第一个试点——气象科学数据共享试点以来,资源环境、农业、人口与健康、基础与前沿等领域科学数据建设与共享工作纷纷展开<sup>[3]</sup>。政府主导的系统化科学数据平台建设始于2004—2005年<sup>[4]</sup>。关于农业科学数据的研究也随之出现并增多。第一,农业科学数据资源共享研究。从数据生命周期来说,其研究内容涵盖了资源整合<sup>[5]</sup>、数据存储<sup>[6]</sup>、数据监管<sup>[7]</sup>、数据汇交与管理<sup>[8]</sup>、

数据开放与出版<sup>[9]</sup>、数据重用<sup>[10]</sup>等环节。研究内容包括农业科学数据共享管理机制<sup>[11,12]</sup>、共享服务<sup>[13]</sup>、共享战略<sup>[1]</sup>、共享体系建设<sup>[14]</sup>、系统建设<sup>[15]</sup>、用户需求<sup>[16]</sup>、共享政策<sup>[17]</sup>、共享技术<sup>[18]</sup>以及区域性农业科学数据共享研究(如,新疆<sup>[19,20]</sup>、辽宁<sup>[21]</sup>等)。第二,农业科学数据资源建设研究。资源是科学数据共享管理的逻辑起点<sup>[22]</sup>。农业数据资源整合作为农业数据开放共享的基础亦受到研究者关注。杨从科<sup>[23]</sup>以数据分类、数据整合、数据政策等数据资源建设手段和内容为重点研究农业资源建设平台构成的技术路线和基本要素。赵瑞雪<sup>[24]</sup>、马海龙<sup>[25]</sup>等则重点关注了国家农业科学数据共享中心建设情况,介绍了该中心的系统架构和功能、数据库网站内容、数据共享途径以及用户服务模式等。曾小红等<sup>[26]</sup>关注热带农业科学数据的数据资源、标准化建设及实施。刘海清等<sup>[27]</sup>聚焦中国贵州省、广东省和海南省的农业科学数据资源平台的建设情况。赵瑞雪等<sup>[2]</sup>从政策制定及执行、数据资源建设模式与资源整合、数据开放与出版等角度对国内外农业科学大数据建设和共享进展进行了总结分析。

## 2.2 国外研究现状

国外对于农业科学数据的研究与实践均早于国内。早在1960年,美国便成立国家大气研究中心进行地球科学数据的建模、收藏和保存工作<sup>[28]</sup>。同时,相关信息政策对科学数据建设有较大影响,如美国《信息公开法》、欧盟《数据库法律保护指令》以及英国《数据保护法》等<sup>[29]</sup>。现今,美国已建立了较为全面的农业科学数据资源系统和农业计算机网络系统,欧盟、德国、韩国等国家亦致力于农业科学数据开放共享<sup>[30]</sup>。农业科学数据资源建设技术完善是国外研究涉及的重要内容,包括元数据体系结构的建立<sup>[31]</sup>、开放信息系统模型框架和农业资源平台关键的软硬件技术研究<sup>[32]</sup>、利用商业智能集成农业数据<sup>[33]</sup>、基于物联网开发数据平台<sup>[34]</sup>等研究。此外,从数据管理层面出发的数据互操作<sup>[35]</sup>、跨学科数据管理<sup>[36]</sup>、数据管理人员及合作<sup>[37]</sup>、跨部门数据共享<sup>[38]</sup>研究及农业数据开放的安全性和隐私问题<sup>[39]</sup>亦受到国外学者的关注。

总的来说,国内外聚焦于农业科学数据资源建设的研究呈现出由基础架构和体系框架设计到跨学科跨部门数据共享和资源建设实践情况的研究趋势。虽然涵盖较广,但主要是从政策战略、体系框架、技术路线等较为宏观角度进行资源整合建设情况的考察,缺少对资源建设平台本身数据组织和服务方式等微观层面的考察。

## 2.3 研究方法和步骤

本文以重要农业开放科学数据资源为目标进行调研和分析。重要开放农业科学数据资源主要指重要农业国家的政府机构、科研机构和大学、图书情报机构建立的科学数据平台中的资源,并且能够以开放形式获取和分析。具体研究路径有3个步骤。

(1) 确定重要农业国家和开放农业科学数据资源来源。首先从农业综合实力国家通识排名靠前的政府机构、农业科研机构、图书情报机构入手,扫描文献、新闻,选择拟调研的国家和国际农业组织,初步确定重要开放农业科学数据的范围。然后从国家级的农业管理部门调研入手,采用机构网站链接扩展机构范围。重点调研国家和国际组织为:中国、美国、加拿大、澳大利亚、德国、法国、英国、日本、新西兰、印度、联合国粮食及农业组织、国际农业及生物学中心、世界土壤信息中心等,机构类型覆盖了国际组织和政府机构、图书情报机构、科学研究及大学(表1)。

(2) 采用网络调研方法,对机构自建的科学数据系统和科学数据基础设施进行调研,并按照机构的重要程度、资源的相似度、网络链接与重要机构相关度排序,筛选出重要农业科学数据资源。

(3) 采用文献调研法,用关键词“科学数据 OR 农业科技信息基础设施 OR 农业科学大数据 OR 农业知识服务 AND 机构名”“Research Data OR Science Data AND 机构名”“Agriculture + Data Infrastructure”分别在CNKI、Web of Science、EI Village下检索,共检索到相关中文文献174篇、英文文献256篇,筛选后中文文献共165篇,英文文献167篇。以筛选后的文献资源作为补充本文数据资源信息的重要参考。

表 1 重要农业开放科学数据调研国家及机构示例

机构类型	国家/组织	机构名称
科研机构	中国	中国农业科学院
		中国科学院
	美国	美国冷泉港实验室
	英国	国际农业及生物科学中心
	荷兰	世界土壤信息中心
	法国	法国农业科学与环境研究院
	日本	NARO <sup>[40]</sup> 日本国立农业研究组织
图书情报机构	中国	中国农业科学院农业信息研究所
	美国	美国国家农业图书馆
	美国	美国生物科学信息服务社
政府管理部门/ 国际组织	国际组织	联合国粮食及农业组织
国际组织		联合国环境规划署
	中国	中国农业部
	美国	美国农业部
	法国	法国农业部
	日本	日本农林水产省
	德国	德国农业部
	澳大利亚	澳大利亚农业部
	印度	National Informatics Centre (NIC), Ministry of Electronics & Information Technology, Government of India

(4) 通过调研现有国内外研究成果，找出目前研究的不足，然后从这些不足之处对国内外具有代表性的农业科学数据资源网站/平台进行深入调查研究，并为中国农业科学数据资源建设提出切实建议。

表 2 农业科学数据资源类型特征

数据类型	定义	农业科学数据资源保存机构	成果载体
科学研究活动数据	包括原生数据和衍生数据。其中原生数据指不依赖于现有数据而产生的数据，观测监测、考察调查、检验检测数据。衍生数据指原生数据被记录、存储后，经过算法加工、计算、聚合而成的系统的、可读取、有使用价值的、可以产生知识产权的数据	科研机构、大学、中心、实验室、观测站等	数据集、数据库、期刊、专著、专利、报告等
研究数据	为了获得关于现象和可观察事实的基本原理的新知识(揭示客观事物的本质、运动规律，获得新发现、新学说)而进行的实验性或理论性研究，它不以任何专门或特定的应用或使用为目的	科研机构、大学、中心、实验室、观测站、图书情报机构	数据集、数据库、地图、期刊、专著、会议、报告、专利、项目等
应用数据	为获得新知识而进行的创造性研究，主要针对某一特定的目的或目标	农业管理机构、农业图书情报机构	期刊、专著、专利、项目、获奖成果、计划、规划、战略等
试验数据	利用从基础研究、应用研究和实际经验所获得的现有知识，为产生新的产品、材料和装置，建立新的工艺、系统和服务，以及对已产生和建立的上述各项作实质性的改进而进行的系统性工作	农业管理机构、科研机构及大学、企业	数据集、数据库、专利、期刊、会议、报告、项目等

### 3 资源建设现状与分析

#### 3.1 资源类型及数据格式特征

本文按照国务院科学数据管理办法<sup>[41]</sup>，将资源划分为科学研究活动数据、基础研究、应用研究、试验活动的科学数据 4 类。将 4 类数据参照国家统计局和杨立新等对原生数据和衍生数据的定义<sup>[42]</sup>，从数据类型出发，对数据来源、数据载体进行对比分析(表 2)。观察到数据来源机构主要是科研院所和大学(包括中心、实验室、观测站等)，图书情报机构、政府管理部门和企业。数据载体主要有数据集、科学论文、专著、发明专利、新产品、新工艺、项目、报告、政策、规划、战略、获奖成果等。

从调研情况来看，农业科学数据形式以文本、数值、图像、视频、语音、文字为主，常见的数据和数据集格式有 csv、geojson、html、kml、xml、txt、wfs、wms、pdf、dwg、shp、zip、FileGDB、GPKG、MapInfo File、TIFF 等。参见表 3 列举的数据形式和格式。

#### 3.2 资源学科范围

科学数据资源在不同国家、平台和机构中的分类有所区别。产生这种情况的主要原因有 3 点。一是数据产出机构不同，存在认知差异。二是学科划分粒度

表3 重要农业科学数据资源组织方法、学科范围、数据形式及格式

资源名称	机构类型	学科范围	数据载体	数据形式	数据格式
国家(中国)农业科学数据中心	科研机构	作物科学、动物科学、热作科学、渔业科学、草地与草业科学、农业资源与环境科学、农业区划科学、植物保护科学、农业微生物科学、食品营养与加工科学、农业农村经济科学、农业科技基础	数据集、数据库	文本、数值、图像、视频、语音等	csv、xlsx、xls、zip等
中国科学院科学数据云	科研机构	农业、土壤、水利、水土保持	数据集、数据库	数值型、文本型、图像、视频等	Rpb、tiff、xml、jpg、rpb、txt、docx、xml等
美国农业数据共享 Ag Data Commons <sup>[44]</sup>	政府机构	农业经济学、生物能源、动物与牲畜、食物与营养、基因组学与遗传学、农业生态系统与环境、植物与农作物、农产品	数据集、数据库	文本型、图像、数值等	Html、csv、pdf、xls、txt、bin、tar、zip、docx、png、gz、jpg、pptx、api、geotiff、xml、accdb、ascii、tgz等
加拿大农业图书馆	图书情报机构	农业与食品科学: 农业经济学、农学、作物科学、畜牧学、植物病理学、植物学、真菌学、食品科学技术、农业病虫害、昆虫学、乳业科学、土壤科学、兽医学	书籍、政府文件、会议录、地图、期刊论文、杂志文章、网络资源、同行评审、报纸、微缩胶卷、音视频记录等	文本、音频、图像等	
英国发现开放数据 (Find open data)	政府机构	农业主题设置在环境主题下, 主要包含的学科有化学品、气候变化和能源、商业捕鱼、渔业和船舶、能源基础设施、环境许可证、粮食和农业、海洋、污染和环境质量、河流维护、洪水和海岸侵蚀、农村和农村、废物和再循环、水工业; 野生动物、动物、生物多样性和生态系统	数据集、数据库	文本、数值、图像、视频、语音等	csv、geojson、html、kml、xml、txt、wfs、wms、zip等
新西兰政府开放数据网站-农林渔 <sup>[45]</sup>	政府机构	农业、林业、渔业	数据集、数据库	文本、数值、图像、视频、语音等	html、tiff、csv、kml、pdf、dwg、shp、FileGDB、GPKG、MapInfo File等
Gramene 比较植物基因组资源	科研机构	该项目具有用于植物、真菌、无脊椎动物后生物、细菌和原生生物基因组的特定Web 门户。旨在提供分类学参考点, 给出可以理解基因的进化背景, 以及涵盖所有主要的非脊椎动物实验生物体、农业重要物种、病原体 and 载体	物种全基因组、野生水稻的部分基因组序列、基因的遗传和物理图谱、表达序列标签位点和数量性状位点、表型特征和编译的描述等	数值、图像	bed、csv、tsv、gtf、gff、gff3、fasta、rtf等
世界土壤信息中心	科研机构	土壤及相关的气候、地质、地貌、植被、土地利用和土地适宜性等地理信息。	文献、国家报告、书籍和地图	文本、图像、数值、视频	Pdf、xml等
联合国粮食及农业组织 AGRIS	政府机构	涵盖农业、林业、渔业、食品等领域的相关分类, 包含 37 000 多个概念和 750 000 多个术语, 覆盖 38 个语种	提供粮食及农业领域的参考书、期刊、专著、图书、数据和灰色文献(即未发表的科学和技术报告、论文、学位论文和会议论文)	文本、图像、数值	xml、txt、pdf、zip、csv等

粗细不同，综合资源的学科要求大而全，数据格式丰富；专题资源的学科要求小而精，或者数据格式相对单一。三是服务方向不同，有面向基础研究、应用和试验开发之分。

农业科学及农业科学数据在不同的国家、不同的机构划分有一定的差异。将各国、各机构产出的农业科学数据（表3）可以划分为农业科学和工程技术<sup>[43]</sup>。基于《中国科学院图书分类法》的划分，农业科学主要包括农业总论、农业经济，农业基础科学，农业工程、农田水利，农业机械化、农业机械及农具，农艺学，植物学，农作物，园艺，林业、林业科学，畜牧、兽医、蚕蜂、水产；工程技术主要包括食品、物流运输及智慧农业等相关学科。

### 3.3 知识组织方法

农业科学数据组织的环境与传统资源组织方法有所不同。主要体现在资源来源、组织主体、组织对象的格式和形式的多样化，结果呈现方式上更多采用动态、可视化满足用户喜好。组织目标则从用户需求和资源内容出发，以科学数据的知识序化为主，兼顾用户的参与和交互。

选择什么样的科学数据知识组织方法，科学数据加工到什么程度受到多重原因影响。这是由于全球信息化技术发展不均衡，同时也受信息安全的影响和信息组织技术能力的限制。从网站、数据库等科学数据资源获取和知识发现角度来看主要采用了基于内容层面的标记、元数据、关联数据、主题词、本体、知识图谱、融合法等知识组织方法。论文选择了有代表性的 AGRIS、中国科学院数据云等，从方法特征、学科范围、数据形式的角度进行分析，发现采用的知识组织技术与科学数据的复杂性、语义控制程度、用户需求多样性成正相关关系，而且在网络环境下，本体技术是较为常用的组织方法。

### 3.4 资源服务模式

基于不同的服务群体，农业科学数据资源平台利用其数据资源提供各式服务。重要资源平台资源服务内容（表4）从数据管理来看，覆盖了数据产生、存储、加工、发布、利用的整个科学数据生命周期；从数据范围来看，原始资料、二手资料以及元数据都包含其中，本机构数据和其他相关机构数据亦集成于此；从面向对象来看，用户范围广、人员层次多样，包括

表4 重要资源平台资源服务内容

国家	资源平台	资源服务内容
中国	农业科技信息资源共建共享平台	文献传递；资源共享、科技查新、定题检索；交流培训；特色服务（研究所科研知识环境；领域知识服务系统、学科团队信息环境、农业科研系统馆藏资源集成揭示系统、农业专家学者系统）
	重点农产品市场信息平台	提供农业统计资料、重点品种信息、市场热点、农产品价格排行、重点农产品批发价格、200 指数、期货价格、进出口贸易及统计资料
美国	Ag Data Commons	提供数据目录和存储库功能。主要服务有：科学元数据的创建和管理；公众通过美国农业部资助的研究和公共 API 目录获取数据；用于长期归档的数据存储库（根据需要支持 DOI 服务）
加拿大	Open Government- Agriculture 国家土壤数据库（NSDB）	数据、信息的上传、访问与下载，开放数据分析，政府咨询以及申请访问受约束未公开的政府信息 提供包含土壤、景观，以及所有加拿大的气候数据的计算机可读文件的下载
意大利	FAOSTAT	为 245 多个国家以及地区提供免费的粮食和农业数据，其涵盖了所有粮农组织（FAO）区域各组自 1961 年以来至最近可获取的所有数据
荷兰	The Ministry of Agriculture, Nature and Food Quality	来自农业、自然和食品质量部的 67 个文件的浏览、检索与下载
澳大利亚	DAWE Data Catalogue	查找和获取信息、制作地图、发布和描述资源、信息分析
	NCI Data Catalogue	超级计算、数据服务、数据收集管理、虚拟研究环境、数据存储、可视化、多维数据集、HPC 优化
新西兰	Stats NZ	信息中心服务、综合数据基础设施（IDI）服务、纵向业务数据库服务、数据实验室访问服务、大学教职员及学生服务、定制数据服务

了农户、科研人员、政府用户、企业用户和其他社会公众。

根据上表内容并结合前文相关调研, 可将农业科学数据资源平台现有的资源服务模式做如下总结。

(1) 资源驱动型服务模式。该服务模式以资源为中心, 强调资源的单向提供。平台将来自本机构或全国乃至全球的相关数据集成起来, 用户仅能对平台提供的数据进行访问。例如荷兰农业、自然和食品质量部允许用户浏览、检索与下载其开放文件; 中国的“重点农产品市场信息平台”提供农业统计资料、重点品种信息等农产品市场信息的检索、浏览与下载; 加拿大国家土壤数据库集成了联邦和各省实地调研收集的有关加拿大土壤、景观以及气候的数据, 并提供下载; 粮农组织统计数据库提供粮农组织 (FAO) 1961 年以来至最近可获取的所有数据。

(2) 交互开发型服务模式。该服务模式重视用户参与, 强调平台与用户的互动。平台允许用户创建个人账户, 用户不仅能够检索、浏览以及下载平台提供的资源, 也能按照一定规则添加、组织资源, 建立个人资源库。如美国农业部开放数据网站则提供个人账户功能, 允许其资助的研究者提交、创建、组织以及管理相应数据集, 并通过科学的元数据创建和管理以实现美国农业部资助的研究数据系统和数据产品可查找、可访问、可互操作和重复利用; 澳大利亚 DAWE Data Catalogue 平台提供强大的元数据编辑器, 支持多类型文件的上传和在线描述。

(3) 面向需求型服务模式。该服务模式以用户为中心, 通过细分用户群体, 提供个性化定制服务。如新西兰统计局的“Stats NZ”平台, 将用户细分为研究人员和大学教职工, 提供集成数据基础架构、纵向业务数据库、数据实验室访问功能以及个性化的定制数据服务; 中国“农业科技信息资源共建共享平台”面

向研究机构、学科团队提供特色研究服务。

(4) 智能导向型服务模式。该服务模式注重现代信息技术的应用, 提供相应技术服务。大数据时代, 要实现数据利用的最大化, 对数据的整合与分析必不可少。加拿大的 Open Government-Agriculture 澳大利亚的 NCI Data Catalogue 和 DAWE Data Catalogue 等都提供相应的数据分析和可视化功能。在该模式下, 平台与用户不是简单的资源与需求的对接, 更要充分通过智能技术挖掘用户潜在需求, 提供智能化的资源服务。

### 3.5 机构类型数据资源的比较

根据以上对国内外重要农业科学数据资源的调查和梳理, 可发现科研机构和大学、图书情报机构、政府和企业三者 in 农业科学数据建设上各有优势。尤其是在数据载体和数据类型方面, 优势差异明显, 具体对比如表 5 所示。

由于不同机构的性质及目的差异, 其在建设农业科学数据资源时侧重有所不同, 如政府机构/国际组织更具有集成优势, 数据更为庞大和宏观, 科研机构和大学则在研究型数据上更具优势, 资源主题类型也更丰富多样, 图书情报机构则更侧重于科技文献数据资源建设, 提供更具指导性的数据资源。

## 4 结论及建议

在政府数字化转型战略需求下, 作为国计民生的重要领域, 农业在科学数据的使用中占有重要地位, “数据是数字化转型的核心, 也是政府转型战略的一部分”<sup>[49]</sup>。建议政府机构、科研院所和大学、图书情报机构基于自身的资源优势、学科优势, 加强沟通和合作, 重点从 4 个方面增强对农业科学数据的开发和利用, 积极参与数字政府建设。

表 5 机构类型资源建设优势对比

机构类型	数据载体	数据类型
政府机构/国际组织	专利、项目、获奖成果、计划、规划、战略等	应用数据
科研机构及大学	数据集、数据库、地图	科学活动数据、基础研究数据、试验数据
图书情报机构	期刊、专著、会议、报告、专利、期刊、专著	科技文献数据

## 4.1 资源类型及学科范围

在资源覆盖范围上, 农业科学数据资源基本形成数据形态广覆盖、数据范围多领域的布局。科学数据资源建设的主要倡议者和建设者是国际组织、科技管理部门、图书情报机构、科研机构 and 高校。政府数字管理倾向于使用带有统计意义的数据。

建议从情报利用的角度, 加强多元异构资源的采集和集成, 并挖掘用于支持农业管理的新的科学数据类型, 不局限数据格式和类型。如体现农业种植面积规模变化的产量数据、体现畜禽养殖规模变化的生产数据。利于相关管理部门从宏观变化中掌握规律。

## 4.2 资源组织方法

就知识组织方式来说, 呈现出组织个性化、信息检索单一化的特征。第一, 在信息分类方式上, 典型农业科学数据资源系统信息分类多关注于农业数据本身, 而缺乏对用户的关注。随着数据共享的不断发展, 用户群体不断呈现出多样化的特点, 除科研人员外, 农业科学数据的使用还包括政府用户、企业用户、农户以及其他社会公众等<sup>[1]</sup>。第二, 在信息检索方面, 无论是检索输入还是检索结果的处理, 典型农业科学数据资源系统提供的方式关联性不足, 对用户的检索效率和检索效果有一定影响。有利的条件是, 调研中的农业科学数据资源系统在满足一定条件下, 能够免费浏览和下载数据, 数据开放性较好。

建议在农业管理数据生命周期基础上, 对异构异源的科学数据进行有效整合, 强调建立在管理和应用上的资源组织逻辑和数据关联, 例如科学数据在政府宏观和中观管理的功能, 为数字政府建设提供参考。其次, 建立基于语义的科学数据内部融合关联研究。如科学文献与基础研究、试验数据、原始数据的关联, 有效复用科学数据和本体, 加强数据循环。

## 4.3 资源服务内容

通过对农业科学数据资源平台的调查, 发现资源驱动服务模式发展已相对成熟, 各平台资源数量颇具

规模。在该种模式下, 集成化、精准服务是其发展的重要趋向, 相关部门、机构、项目团队的资源通过统一的描述和组织方式进行聚合, 实现数据共建共享。美国、加拿大等的政府开放数据网站即是如此, 用户能够在该一站式平台查找到农业核心领域、农业相关领域乃至整个社会经济领域的信息。此外, 面向用户、智能导向是资源平台服务模式建设的重点内容。随着平台资源建设的不断发展, 平台资源量增加、用户增多, 依靠智能技术组织、分析、定制和共享数据以满足用户多样化、差异化信息需求成为平台发展重点。最后, 对于农业科学数据服务平台来说, 多种服务模式并没有割裂, 而应该实现多种服务模式的综合。提供资源驱动、面向用户、数据集成、智能分析的综合服务。

此外, 在科学数据使用和建设过程中, 还需要增强数据安全意识。通过科学数据开放地理区域、使用范围和使用频率有效管理。

## 4.4 资源管理要务

提高数据治理能力是数字政府建设及管理的重要内容, 从整体来看, 通过数据汇集形成政务大数据资源池, 其中就包括人口、自然资源、空间地理等公共基础数据库, 而农业科学数据管理则是极为重要的数据治理领域。要实现农业科学数据有效的资源管理, 政策和技术保障二者缺一不可。首先要构建整体的数据治理技术架构, 保障农业科学数据治理的协调一致, 明晰数据采集、数据处理、数据共享、数据再利用的流程, 发挥现代技术优势; 其次, 有关农业科学数据治理方面, 政策法规的制定、组织机构职责的明确、数据管理效果的评估等都是数字政府建设及管理的重要方向。

### 参考文献:

- [1] 张莉. 中国农业科学数据共享发展研究[D]. 北京: 中国农业科学院, 2006.
- [2] 梁木生. 略论“数字政府”运行的技术规制[J]. 中国行政管理, 2001(6): 20-21.

- [3] 科学数据共享工程[EB/OL]. [2020-09-08]. [http://www.most.gov.cn/ztl/kjzg60/kjzg60hhcj/kjzg60jcyj/200909/t20090911\\_72832.htm](http://www.most.gov.cn/ztl/kjzg60/kjzg60hhcj/kjzg60jcyj/200909/t20090911_72832.htm).
- [4] 田野. 区块链环境下的异构政府数据语义共享机制研究[J]. 农业图书情报学报, 2020, 32(1): 12-22.
- [5] 贺玲玉. 基于关联数据的农业信息资源整合研究[D]. 武汉: 华中师范大学, 2014.
- [6] 王剑, 黄朝光. 海量农业科学数据存储体系架构与方法研究[J]. 广东农业科学, 2015, 42(2): 152-156.
- [7] 陆丽娜, 王萍, 于啸. 农业科学数据监管平台构建研究[J]. 图书情报工作, 2017, 61(10): 68-73.
- [8] 赵瑞雪. 农业科学数据共享中数据汇交与管理研究[J]. 科技管理研究, 2009, 29(8): 284-286.
- [9] 赵华, 王健. 科学数据出版现状及对中国农业科学数据出版的启示[J]. 农业展望, 2016, 12(8): 53-57.
- [10] 彭秀媛, 王枫, 周国民. 面向重用的农业科学数据共享模式研究[J]. 农业经济, 2019(1): 87-89.
- [11] 张莉. 论农业科学数据的共享管理[J]. 农业图书情报学刊, 2006(12): 36-39.
- [12] 高峰, 王剑. 基于合作博弈的农业科学数据平台共享管理机制研究[J]. 技术与创新管理, 2013, 34(5): 457-460.
- [13] 张小红, 李思经. 农业科学数据共享服务研究[J]. 中国科技论坛, 2006(5): 127-130.
- [14] 张莉. 农业科学数据共享体系建设分析——以国家农业科学数据中心建设为例[J]. 情报探索, 2007(10): 66-68.
- [15] 卓文飞. 论我国农业科技信息集成服务系统的构建[J]. 中国科技论坛, 2010(3): 134-138.
- [16] 王华丽, 张磊磊, 王新哲, 等. 新疆农业科学数据共享用户需求调查分析[J]. 新疆农业科技, 2014(5): 1-3.
- [17] 赵华, 王剑. 科学数据共享政策及其对中国农业科学数据共享的启示[J]. 农业展望, 2015, 11(11): 70-74.
- [18] 彭秀媛, 王枫, 周国民. 农业科学数据共享技术系统研究[J]. 园艺与种苗, 2019, 39(2): 53-57.
- [19] 王华丽, 张磊磊, 王新哲. 基于协同理论的新疆农业科学数据共享研究[J]. 农业现代化研究, 2015, 36(4): 708-715.
- [20] 刘莹, 王华丽. 新疆农业科学数据共享保障机制构建[J]. 黑龙江农业科学, 2016(6): 118-122.
- [21] 彭秀媛, 王枫, 周国民. 辽宁省农业科学数据共享情况调查与分析[J]. 农业经济, 2017(1): 59-61.
- [22] 科学数据共享调研组. 科学数据共享工程的总体框架[J]. 中国基础科学, 2003(1): 65-70.
- [23] 杨从科. 中国农业科学数据资源建设研究[D]. 北京: 中国农业科学院, 2007.
- [24] 赵瑞雪, 李娇, 张洁, 等. 多场景农业专业知识服务系统构建研究[J]. 农业图书情报学报, 2020, 32(1): 4-11.
- [25] 马海龙, 陈佑启, 邹金秋. “国家农业科学数据共享中心农业区划数据分中心”系统设计与建设[J]. 中国农业资源与区划, 2010, 31(1): 75-80.
- [26] 曾小红, 王强. 热带农业科学数据标准化建设及对策研究[J]. 热带农业科学, 2011, 31(3): 54-57, 65.
- [27] 刘海清, 卢琨, 刘晓珂, 等. 农业科学大数据资源平台在热区的实践思考[J]. 农业大数据学报, 2019, 1(4): 58-64.
- [28] 柏永青, 杨雅萍, 孙九林. 国内外科学数据管理办法研究进展[J]. 农业大数据学报, 2019, 1(3): 5-20, 4.
- [29] 张莉. 中国农业科学数据共享发展研究[D]. 北京: 中国农业科学院, 2006.
- [30] 陈婷婷, 官波, 沈祥成, 等. 大数据时代开放式农业信息知识库构建研究[J]. 东北农业科学, 2018, 43(5): 60-64.
- [31] SGOUROPOULOU C, KOUTOUMANOS A. Building metadata architectures: A case for e-research[J]. International journal of metadata semantics & ontologies, 2014, 9(4): 275-288.
- [32] ZHANG M, LI Z, LI F. Discussion on key technologies in forestry fundamental scientific information cloud service platform[C]//Green computing & communications. IEEE, 2013.
- [33] WISNUBHADRA I, ADITHAMA S P, BAHARIN S S K, et al. Agriculture spatiotemporal business intelligence using open data integration [C]//2019 international seminar on research of information technology and intelligent systems (ISRITI), 2019.
- [34] AYDIN S, AYDIN M N. Semantic and syntactic interoperability for agricultural Open-Data platforms in the context of IoT using Crop-Specific trait ontologies[J]. Applied sciences, 2020, 10(13): 1-27.
- [35] ENAYAT, RAJABI, SALVADOR, et al. A linked and open dataset from a network of learning repositories on organic agriculture[J]. British journal of educational technology, 2015.
- [36] MARTIN C, CADIOU C, EMMANUELLE J-O. Data management:

- New tools, new organization, and new skills in a French research institute[J]. 2017.
- [38] COUTU S, BECKER-RESHEF I, WHITCRAFT A K, et al. Food security: Underpin with public and private data sharing [J]. *Nature*, 2020, 578(7796): 515-515.
- [39] FERRAG M A, SHU L, YANG X, et al. Security and privacy for green IoT-Based agriculture: Review, blockchain solutions, and challenges[J]. *IEEE access*, 2020: 99.
- [40] 農研機構|国立研究開発法人農業・食品産業技術総合研究機構[EB/OL]. [2020-07-24]. <http://www.naro.affrc.go.jp/>.
- [41] 科学技术[EB/OL]. [2020-08-04]. [http://www.stats.gov.cn/tjsj/zbjss/201912/t20191202\\_1713041.html](http://www.stats.gov.cn/tjsj/zbjss/201912/t20191202_1713041.html).
- [42] 杨立新. 衍生数据是数据专有权的客体[N]. *中国社会科学报*, 2016-07-13(005).
- [43] 中国科学院图书馆图书分类法第三版[D]. 北京: 科学出版社, 1994.
- [44] Ag data commons|providing central access to USDA's open research data[EB/OL]. [2020-07-24]. <https://data.nal.usda.gov/>.
- [45] Agriculture, forestry and fisheries-data gov. Nz-discover and use data [EB/OL]. [2020-07-29]. <https://catalogue.data.govt.nz/group/3a53bc45-ab5a-478d-b613-c6b31ba0857c?organization=ministry-for-primary-industries>.
- [46] About-data gov. UK[EB/OL]. [2020-07-29]. <https://data.gov.uk/about>.