

开放环境下农业科技信息资源开放共享 ——以 FAO 的实践为例

曹雨晴¹, 冯东昕², 张学福^{1,3}, 孙巍^{1,3}, 罗婷婷¹, 鲜国建^{1,3*}

(1. 中国农业科学院农业信息研究所, 北京 100081; 2. Food and Agriculture Organization of the United Nations (FAO), Rome 00153; 3. 农业农村部农业大数据重点实验室, 北京 100081)

摘要: [目的 / 意义] 开放科学运动的兴起极大地促进了学术传播以及科研成果利用。开放共享作为推进开放科学运动的重要手段, 也是推动全球农业科技发展的必要措施。通过梳理开放环境下联合国粮农组织 (Food and Agriculture Organization of the United Nations, FAO) 面向全球农业科技信息资源的开放共享实践, 有助于为推动中国实现农业科技信息资源的开放共享、促进科研成果的转化及利用提供借鉴。[方法 / 过程] 从率先构建并积极践行农业信息管理标准 AIMS、持续维护与多渠道开放共享多语种叙词表 AGROVOC, 以及精心建设与运营国际农业科技信息检索系统 AGRIS 等多维度, 梳理 FAO 为推进国际农业科技信息资源开放共享的实践进展。[结果 / 结论] 基于 FAO 为推动国际农业科技信息资源开放共享得出的实践经验, 从构建规范化管理体系、提高资源开放共享能力以及促进国际合作等方面对中国农业科技信息资源开放共享的建设工作提出展望。

关键词: 开放科学; 开放共享; FAO; AGRIS; AGROVOC; 农业科技信息

中图分类号: G250

文献标识码: A

文章编号: 1002-1248 (2020) 12-0050-09

引用本文: 曹雨晴, 冯东昕, 张学福, 等. 开放环境下农业科技信息资源开放共享——以 FAO 的实践为例[J]. 农业图书情报学报, 2020, 32(12): 50-58.

收稿日期: 2020-10-22

基金项目: 中国农业科学院创新工程重点任务“FAO- 中国农科院区域创新平台及‘一带一路’能力建设”和“农业科技大数据融合计算关键技术”(CAAS-ASTIP-2016-A11)

作者简介: 曹雨晴, 女, 硕士研究生, 研究方向为知识图谱构建与应用。冯东昕, 女, 博士, 研究员, 研究方向为植物分子病理学、农业国际合作。张学福, 男, 研究员, 博士生导师, 研究方向为农业信息管理和科技情报研究。孙巍, 女, 博士, 研究员, 硕士生导师, 研究方向为农业信息管理和科技情报研究。罗婷婷, 女, 硕士, 助理馆员, 研究方向为知识组织、大数据融汇治理、信息管理与信息系统

***通信作者:** 鲜国建 (ORCID: 0000-0003-4332-1958), 男, 博士, 研究员, 硕士生导师, 研究方向为大数据融汇治理、知识组织、知识图谱。E-mail: xianguojian@caas.cn

Opening up and Sharing Agricultural Scientific and Technological Information Resources: Case Study of the Practice of FAO

CAO Yuqing¹, FENG Dongxin², ZHANG Xuefu^{1,3}, SUN Wei^{1,3}, LUO Tingting¹, XIAN Guojian^{1,3*}

(1. Agricultural Information Institute of CAAS, Beijing 100081; 2. Food and Agriculture Organization of the United Nations (FAO), Rome 00153; 3. Key Laboratory of Agricultural Big Data, Ministry of Agriculture and Rural Affairs, Beijing 100081)

Abstract: [Purpose/Significance] The global rise of the Open Science Movement has greatly promoted the academic dissemination and the uptakes of scientific research results. As an important means of advancing the Movement of Open Science, openness and sharing is key to promoting the development of global agricultural science and technology. By reviewing the practice of the Food and Agriculture Organization of the United Nations (FAO) for opening up and sharing global agricultural scientific and technological information resources under an open environment, it provides reference for promoting the opening-up and sharing of agricultural scientific and technological information resources and the conversion and utilization of scientific research results in China. [Method/Process] Through a multi-dimensional review of the practice progress of FAO in promoting the opening-up and sharing of international agricultural scientific and technological information resources, the main steps are to take the lead in the construction and active implementation of Agricultural Information Management Standards (AIMS) and multi-channel openness and sharing of multilingual thesaurus AGROVOC, and elaboration on the construction and operation of the International System for Agricultural Science and Technology (AGRIS). [Results/Conclusions] Based on the experience gained by FAO to promote the opening-up and sharing of international agricultural scientific and technological information resources, we summarize some useful suggestions for improving the openness and sharing of agricultural scientific and technological information resources in China from the aspects of the establishment of a standardized management system, improvement of the capability development and promotion of international cooperation.

Keywords: open science; openness and sharing; FAO; AGRIS; AGROVOC; agricultural sci-tech information

1 引 言

当前, 开放科学运动在世界范围内迅速发展, 以开放性和连通性为基础的开放科学为科学研究的整个生命周期 (包括计划、执行、获取、评估等) 带来了巨大的社会文化变革和技术变革。早在 2013 年欧洲委员会副主席尼利·克洛斯便指出: “我们已经步入开放科学时代”^[1]。迈克尔·尼尔森^[2]最早提出开放科学, 并将其定义为不同形式的科学数据在科学发现初期实现一定程度的开放共享。科学技术的迅猛发展推动了数

据密集型科研范式的兴起与发展, 所产生的海量科学数据使科学数据的开放共享变得十分重要。开放数据工具、开放获取平台、开放同行评审以及开放公众参与活动的兴起都代表了开放科学正在深深影响着科研领域, 科技信息资源共享有利于形成更加丰富的数据环境, 以方便研究人员更加轻松快捷地发布、发现、访问和使用数据。

开放科学运动的全球兴起极大地促进了科研学术成果的国际间传播和利用。例如由欧盟主导的欧洲开放科学云 (European Open Science Cloud, EOSC)^[3]、OpenAIRE 项目^[4]并制定 “FAIR” 规范^[5], 均致力于促

进跨学科和跨地域的数据驱动研究成果能够更方便地可发现、可访问、可交互和可复用,为一线欧洲研究人员提供了无比优越的开放获取信息资源服务。作为农业领域的联合国专门机构,FAO 一直致力于对农业信息和知识的开发、维护和拓展。在农业科技信息资源开发利用方面,FAO 紧随国际开放科学运动的发展,充分借助以语义网技术为代表的新兴信息技术,不断促进农业科技信息资源的开放共享。

2 FAO 推进农业科技信息资源开放共享概况

FAO 是联合国各成员国间讨论如何共同努力解决粮食和农业问题的国际组织,其目标是保证世界上所有人的粮食安全,确保人们有规律地获得足够的高质量食品,过上积极健康的生活。截至 2020 年,已有 75 年历史的 FAO 共有 194 个成员国,一个成员组织和两个准成员,在全世界 130 多个国家,正在来自中国的屈冬玉总干事的有力领导下,为创建“Zero Hungry 零饥饿”世界、实现“可持续发展目标 (SDG)”积极开展工作。而积极推动国际农业科技信息资源的开放共享,将有助于农业科技知识和研究成果在全球范围内的传播与转化应用,也是实现 FAO 可持续发展目标的重要举措。

长期以来,FAO 通过构建农业领域多语种叙词表 AGROVOC 并对农业科技信息进行规范化标引组织,建设国际农业科技信息检索系统 (International System for Agricultural Science and Technology, AGRIS),以促进全球科研机构及相关人员对农业科技信息资源的高效检索。随着信息技术的不断发展,针对海量异构农业科技信息资源,FAO 率先提出构建农业信息管理标准平台 (Agricultural Information Management Standards (AIMS) Portal,为全球科研机构及相关人员提供最佳获取农业信息标准、技术和实践的相关信息。

此外,为满足全球科研机构及相关人员,尤其是发展中国家的科研机构及相关人员对科技文献信息日益增加的需求,FAO 还做出了一系列举措促进全球农业科技信息资源的开放共享。例如:2002 年,FAO

联合国际主要期刊出版商以及包含世界卫生组织 (WHO) 在内的相关国际性组织,共同合作搭建了 Research4Life 知识共享平台^[6],旨在为成员国的科研机构提供免费或低费用的在线资源访问,以缩小发达国家与发展中国家之间的知识鸿沟。其中由 FAO 主导的全球农业在线获取平台 (Access to Global Online Research in Agriculture, AGORA) 于 2003 年正式启动,目前已为全球 9 000 多家科研机构及研究人员提供包括农业、食品、营养、动物科学及农业相关的生物环境科学和社会科学等领域的主要期刊资源。2015 年,FAO 依托开放数据平台 F1000Research 门户网站开展农业发展开放知识项目 (Open Knowledge in Agricultural Development, OKAD),以进一步促进农业科技领域信息资源的开放共享,尤其是针对原始数据集的开放共享。2016 年,FAO 针对现有的农业领域知识组织体系构建了丰富且全面的农业领域本体知识库 AgroPortal^[7],为农业科技信息资源的开放共享奠定基石。同时,基于已有的农业信息管理标准平台 (AIMS Portal) 和国际农业科学技术情报系统 AGRIS,FAO 与国际开放知识库同盟 (Confederation of Open Access Repositories, COAR) 以及研究数据同盟 (Research Data Alliance, RDA) 共同合作开发并推广了农业科技领域的语义产品及规范,使农业科技信息领域资源的开放获取进展又步入了一个新的阶段^[8]。2019 年,FAO 又联合全球农业和营养开放数据协会 (Global Open Data for Agriculture and Nutrition, GODAN)、农业农村技术中心 (Technical Centre for Agricultural and Rural Cooperation, CTA) 等,共同开展线上课程“农业数据管理、共享和农业发展服务”,以进一步提高农业及相关领域研究人员对农业开放数据的认识,同时促进与各种开放数据管理项目建立合作关系^[9]。

3 FAO 推进农业科技信息资源开放共享的重要举措

FAO 为推进农业科技信息资源的开放共享的实践进展,从率先构建并积极践行农业信息管理标准

AIMS、持续维护与多途径开放共享多语种词 AGROVOC 以及精心建设与运营国际农业科技信息检索系统 AGRIS 等方面开展工作。其总体思路如图 1 所示。具体包括：基于统一的农业信息管理标准对海量资源进行规范化处理，生成大规模的结构化书目型元数据库；同时通过对农业领域多语种叙词表 AGROVOC 的不断扩展和完善，实现对海量信息资源的高效组织索引，以促进农业科技信息资源的开放共享。

3.1 率先构建并积极践行农业信息管理标准

3.1.1 AIMS 简介

FAO 作为国际性组织，坚持标准先行。农业信息管理标准^[10]（Agriculture Information Management Standards, AIMS）旨在为全球农业领域的学者和机构提供优质的农业科技信息资源，同时推进对研究成果的开放共享。AIMS 同时也作为农业领域的资源聚合器，集合了来自不同信息源的海量农业科技信息资源，方便用户利用语义网技术和农业信息标准，将相关领域的合作伙伴的资源开放共享实践应用到农业领域中。此外，AIMS 是全球农业研究与发展信息共享体系^[11]（Coherence in Information for Agricultural Research for Development, CIARD）的重要组成部分，为全球农业领域的学者和机构提供最佳获取农业信息标准、技术和实践的相关信息，为实现农业信息服务资源互操作，

推动全球开放农业知识共享提供了强有力支持。

3.1.2 AIMS 指导推动的开放共享服务

AIMS 为促进对农业信息资源的开放共享，通过提供一系列管理农业信息资源的工具和标准去实现。主要包括以下几种形式。

（1）促进农业领域研究成果开放获取（Open Access）。AIMS 为促进农业信息资源的开放获取，提供了一整套可促进研究成果可发现和复用的方法论，通过收集农业领域有效的参考数据源、研究农业领域主要机构的开放获取案例以及主要机构的开放获取授权以促进农业领域的开放获取。

（2）构建专业词表支撑资源索引和词表服务（Vocabulary）。AIMS 构建并完善农业领域的专业词表以支撑内容和信息资源索引，并且通过 VEST 目录提供更多类型的词表服务，如规范文档、分类体系、概念图、受控词表、本体等。

（3）制定元数据标准促进资源获取与共享（Metadata）。AIMS 通过提供一整套的通用元数据属性、鼓励使用规范数据、受控词汇表与句法编码标准，以促进农业信息资源的获取与共享。主要包括了 M2B 标准^[12]（the Meaningful Bibliographic Metadata）、VEST 目录和农业元数据元素集（Agricultural Metadata Element Set, AgMES）等，为农业信息的管理、获取以及利用提供帮助。

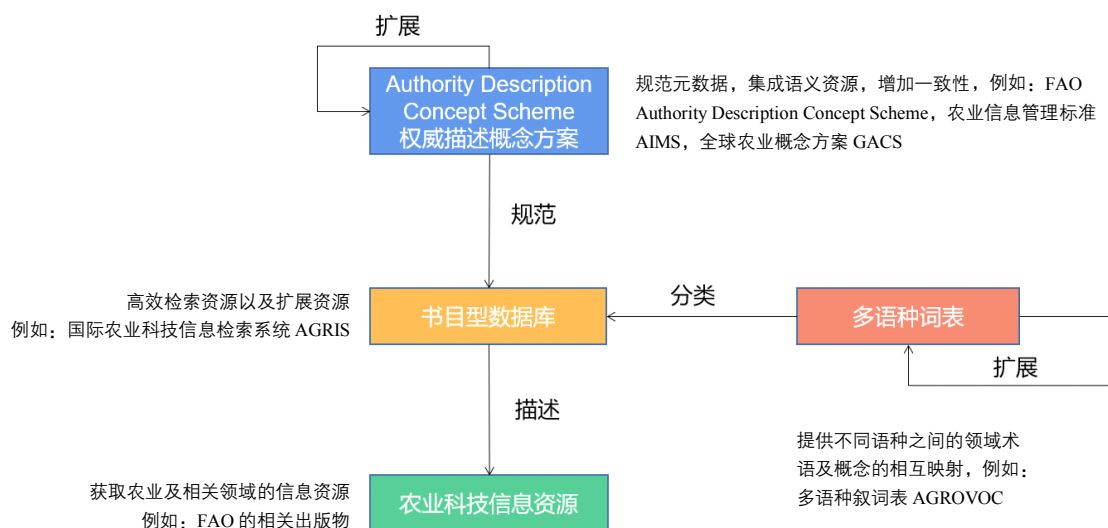


图 1 FAO 推动农业科技信息资源开放共享的总体思路

* 注：翻译自 https://www.w3.org/2005/Incubator/1ld/wiki/File:FAO_Relationships_between_Use_Cases.png

(4) 研发信息管理工作促进互操作 (Information Management Tools)。AIMS 为了丰富元数据的内容, 研制了相关信息管理工具以匹配元数据格式, 保证符合 OAI-PMH 格式的元数据可被充分利用。例如农业海洋资源管理系统 AgriOceanDSpace、农业元数据生成系统 AgriMetaMaker 和农业文档管理系统 AgriDrupal 都促进了描述农业信息资源相关元数据的互操作和复用。

(5) 关联数据 (Linked Data)。为了保证语义网技术的充分利用, AIMS 专门面向关联数据提供了相关标准和知识服务工具, 例如将 AGROVOC 以关联数据的形式提供在线服务。

(6) 推动开放关联书目数据 (Linked Open Data)。AIMS 为了进一步促进农业信息资源的开放获取, 近年来组织编制了关联书目数据指南 LODE-BD, 实现对不同开放知识库中的书目数据进行编码处理, 并形成关联开放数据可用的数目数据。

此外, AIMS 还提供网络社区管理服务: 发布相关进展新闻、组织开展网络研讨会、构建有关元数据、提供信息工具和叙词表的信息目录、收集并分享实践案例以及定期访问 AIMS 社区的相关研究人员以帮助用户更好地通过 AIMS 获取农业科技信息资源, 并促进全球农业研究者之间的良性互动。AIMS 社区旨在让全球农业领域研究者针对农业信息管理的标准、相关工具^[13] (例如 AgriOcean、DSpace、AgriDrupal 等) 以及方法论进行探讨以促进农业信息资源的开放共享。

3.2 持续维护与多途径开放共享多语种叙词表

3.2.1 AGROVOC 简介

AGROVOC 是一部多语种农业叙词表, 1980 年由 FAO 和欧盟委员会联合开发, 涵盖了农业、林业、渔业、食物安全及其他相关学科领域。AGROVOC 最初作为农业领域的核心词汇表被广泛使用。2003 年起, AGROVOC 基于互联网技术由传统农业核心叙词表面向本体重新开发, 成为了农业领域的权威知识库, 使得农业领域的相关概念和专业术语通过 AGROVOC 实现了一定程度上的语义关联。2011 年起, AGROVOC 又开始以关联数据的形式为农业领域在被广泛应用。

相较于传统的基于词的索引, 以关联数据形式发布的 AGROVOC 是基于 URI 的索引, 可与其他标准词汇表通过映射链接相互关联, 从而实现高效的互操作。最新发布的 2020 年 10 月版本的 AGROVOC 词表包含 37 831 个概念和 781 393 个术语, 涵盖了 40 种语言, 基于 RDF/SKOS-XL 格式与超过 20 种国际大型词表在农业领域建立语义关联。

3.2.2 AGROVOC 的扩展与应用

AGROVOC 主要以基于 SKOS 格式的概念体系和开放关联数据集两种形式面向用户提供访问和获取服务: 用户可登录 FAO 门户网站利用 SKOSMOS 浏览器直接浏览查询, 也可以通过访问 SPARQL 端点查询获取, 还可以直接开放下载不同格式的 AGROVOC 等。

为了进一步推动农业科技数据更加符合“FAIR”规范, AGROVOC 通过 VocBench3^[14]平台进行协作编辑管理, 以维护叙词表内容、本体模型、代码列表和权威资源等重要资源, 同时提供历史编辑记录、安全验证、出版流程、基于用户角色访问权限的多用户管理功能。VocBench 最初是为促进 AGROVOC 的协助编辑而被开发使用, 自从 2019 年与 ART 小组建立合作, 逐渐发展为基于 SKOS 的开发协作环境, 并逐步实现针对 AGROVOC 的细分领域概念体系的管理, 从而促进细分专业领域资源的进一步开放共享。目前已成功实现与土地治理、法律与政策、土壤科学、渔业及水产科学等细分专业领域的合作。

此外, FAO 高度重视与全球农业信息管理机构以及领域专家的长期合作。目前 FAO 与包括中国农业科学院在内的专业机构合作构建了 AGROVOC 编辑团队合作网络, 来自 25 个国家的 30 多个组织机构的专家奉献他们的专业知识, 根据农业科技的最新进展, 通过添加新的概念与词汇, 添加和更新术语, 丰富了 AGROVOC 的概念和术语。例如 AGROVOC 针对最新当前需求在最新的版本中增加了有关 COVID-19 的相关术语内容^[15]。2018 年以来, 已先后在荷兰、德国和在线视频 (2020 年因新冠肺炎疫情改为在线) 召开了三届编辑工作会议, 来自中国农业科学院 (CAAS)、国际应用生物科学中心 (CABI)、美国国家农业图书

馆 (NAL)、德国农业技术与建设理事会 (KTBL)、法国农业科学院 (INRA) 等组织机构的专家一起, 围绕 AGROVOC 的标准规范、技术平台和应用推广等方面进行广泛交流和学术研讨。

作为开放关联数据集发布的 AGROVOC, FAO 团队正在进一步对其内容进行逐步扩展, 以推动其在更多相关领域的利用。AGROVOC 通过与中国农业科学叙词表 (CAT) 在类的 18 个外部词表进行相互映射以扩大使用范围 (<http://aims.fao.org/vest-registry/vocabularies/agrovoc>), 以便 FAO 构建全球性信息资源网络为实现进一步资源开放共享贡献力量, 为更广泛的用户群体和机构提供高效信息服务。随着 AGROVOC 的不断完善和发展, 对 AGROVOC 的利用也趋向多样化, 尤其是将 AGROVOC 广泛应用于专业领域图书馆以及机构知识库中, 以实现文献资源内容的索引从而有利于更深层次的文本挖掘。例如在 2015 年启动的“全球农业概念方案 (Global Agricultural Concept Scheme, GACS)”^[16], 旨在基于 AGROVOC 开放关联数据集, 建立与国际应用生物科学中心叙词表 (CABI)、美国国家农业图书馆叙词表 (NALT) 核心概念集的映射, 充当连接面向用户的叙词表和语义上更精确、更专业的领域本体的中心枢纽, 从而与其他相关领域数据集建立精准连接, 以提高农业科技信息资源的可发现性和语义互操作性。

3.3 精心建设与运营国际农业科技信息检索系统

3.3.1 AGRIS 概况

国际农业科技信息检索系统 (AGRIS) 是由 FAO 联合 150 多个国家、地区以及国际性组织共同创建的面向农业领域科技文献的多语种元数据库。AGRIS 服务最初提出始于 1974 年, 旨在创建一个可获取全球农业信息资源的大型农业书目型数据库。随着语义网技术的不断发展以及开放共享观念的不断深入, FAO 对 AGRIS 进行重新设计架构, 使其成为基于关联开放数据的集成系统, 不仅为广大农业领域研究者提供信息资源的访问和获取, 并且提供语义知识服务以扩展现

有知识。目前, AGRIS 共收录了 1 200 万余条记录, 资源类型包括期刊论文、专著、图书章节以及技术报告和会议论文等灰色文献, 其中 300 万余条记录有全文获取链接, 大多数记录基于 AGROVOC 词表进行了标引。2019 年, 新增收录了 1 500 余个多来源的开放数据集。此外, AGRIS 还可通过 Google Scholar 进行索引和访问, 以扩展全球访问。为了尽可能实现农业科技信息资源的开放共享, 在用户遵循相关使用政策的前提下, AGRIS 授予用户对核心数据的完全开放访问, 同时利用各种有效方式 (例如电子邮件、定期会议等) 邀请用户群体对 AGRIS 的未来发展提出意见, 以实现更进一步的农业信息资源的开放共享。AGRIS 每月为全球超过 40 万的农业研究工作者及相关人员提供农业科技信息资源。

3.3.2 AGRIS 特色服务功能

AGRIS 作为 FAO 推动农业科技信息资源开放共享实践进展中的重要一环, 旨在提供高质量的特色信息服务, 因此 AGRIS 具备以下特点。

(1) 提供多语种检索服务。AGRIS 作为面向国际的农业科技信息检索系统, 多语言检索为提供高质量信息服务奠定基石。基于多语种叙词表 AGROVOC 的 AGRIS, 可将用户关键词 (不经过分词处理) 直接与 AGROVOC 中的各语种优选标签和候选标签进行字符匹配, 系统将会把匹配上的叙词概念或术语的所有语种标签字符串重组后, 与 Solr 索引进行匹配实现多语种检索, 以保证用户查询的精准度, 同时还可提供关键词的多语种专业术语翻译。此外, 自 2014 年以来, FAO 与中国农业科学院建立合作, 利用 AGROVOC 与 CAT 的映射关系, 将大量的中文元数据引入 AGRIS 数据库, 大大丰富了数据库的知识基础并提供了多语种检索中文资源的支持。

(2) 对异构元数据的集成处理。AGRIS 作为大型书目型数据库的检索高效性还依赖于对异构元数据的集成处理: 从海量数据源中抽取的异构元数据记录通过 AGRIS 输入应用程序直接生成 RDF 格式记录作为 AGRIS 内部模型; 再应用必需的过滤器进行一定的过滤 (例如针对期刊、发布者和数据提供者相关信息的

实体消歧和内容清洗)；接着利用 SemaGrow 搜索器^[17]和关键词提取器 AgroTagger 生成基于 AGROVOC 的 URI，然后由 RDF 管理器负责将三元组添加到 AGRIS 三元组存储，并根据 Apache Solr 的要求为它们建立索引；最后基于 Mashup 页面，将所有记录链接到外部信息源，并使用 Google 自定义 API 接口获取全文。图 2 为 AGRIS 针对异构元数据的集成处理流程^[18]。

(3) 构建了语义链接网络。为保证农业领域研究机构及相关人员对全球农业科技信息资源的开放获取，AGRIS 一直致力于提供有效的农业语义链接服务。自 2013 年起，AGRIS 将数据资源转换为开放关联数据 LOD，利用 URI 标识实现对数据资源的访问和复用。此外，当用户进行相关检索时，AGRIS 引用 Mashup 技术来丰富用户的查询，提供对更多丰富有效的信息源的访问接口以提升用户的信息获取质量。同时用户可利用 Mashup 页面（图 3）所显示的相关记录对外部网络服务进行异步查询以获取更加丰富的扩展信息。AGRIS 中的语义链接主要依赖于两种机制：一种是 URI 链接机制，即通过 URI 来标识和定位目标资源，用户通过访问 SPARQL 端点直接查询数据；另一种是 RDF 链接机制，即利用开放关联数据集中的实体和概念之间的 RDF 链接机制构建开放关联数据网络，以实现语义层面的数据共享。

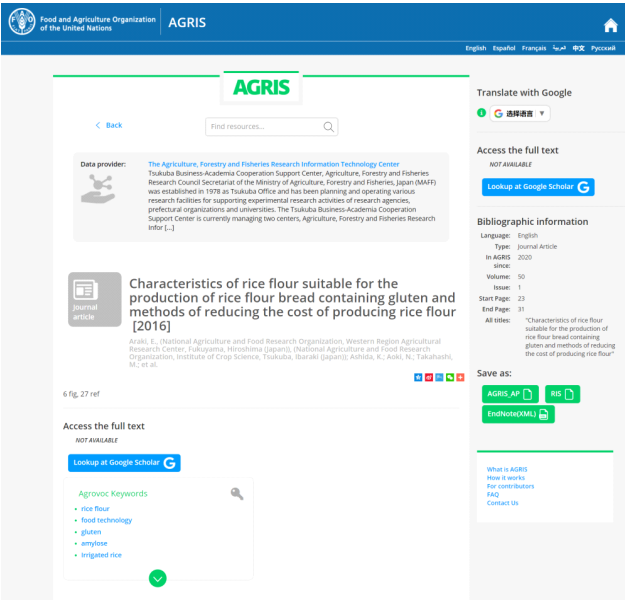


图 2 AGRIS 的 Mashup 界面

(4) 基于用户驱动的信息服务。基于开放共享的理念，AGRIS 的信息源和信息服务的发展是基于用户驱动的，即依赖于用户所提供的各种反馈：例如填写 AGRIS 的反馈表单；接受 AGRIS 团队的阶段性调查；参加 AIMS 网络研讨会等。目前，随着 AGRIS 的不断升级，其向用户提供了越来越多的高价值语义知识服务和不断拓新的庞大信息源；同时，积极的用户反馈又反过来促进了基于 AGRIS 的农业科技信息资源的开放共享。

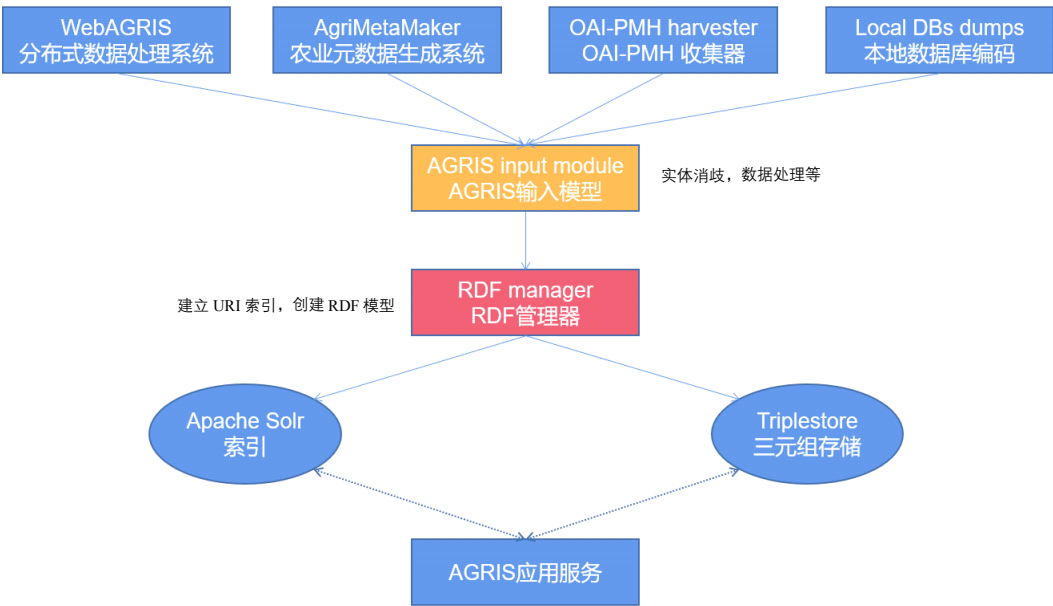


图 3 AGRIS 面向异构元数据的集成处理流程

4 对推进中国农业科技信息资源开放共享的启示

作为国际化组织, FAO 高度重视信息资源的共建、开放、共享, 实现了科技信息服务对组织战略目标实现的支撑和促进作用。无论是为推进全球农业科技信息资源规范管理而制定的农业信息管理标准 AIMS, 对多语种叙词表 AGROVOC 的长期维护和多途径开放共享, 以及对国际农业科技情报系统 AGRIS 的精心运营, 还是为了缩小发达国家与发展中国家之间的资源差距、实现全球农业科技信息资源开放共享而做出的一系列有效实践及其成果, 都对欧洲乃至全球范围内开放科学以及开放获取运动带来极大的积极影响, 同时也对中国研究农业科技信息资源开放共享提供鲜活的案例借鉴。

尽管目前国内各农业科技信息机构已经开始加强合作, 以实现农业科技信息资源的共享与集成。例如, 由中国农业科学院农业信息研究所联合多家单位共同建设的国家农业科学数据共享中心和农业专业知识服务系统, 旨在有效盘活、挖掘、抢救和保存中国农业科技大数据资源^[9]。但与发达国家相比, 中国农业开放科学研究和信息资源开放共享尚属起步阶段, 仍存在诸多不足, 主要表现为: 资源建设相对缺乏整体性和系统性; 资源整合和挖掘深度不足以提供知识增值服务; 资源版权制约资源开放共享进程。

因此, 针对上述问题, 结合 FAO 推进农业科技信息资源开放共享的实践, 作者提出“十四五”时期, 中国可从以下几方面加强创新与应用服务。

(1) 完善农业科技信息资源的管理标准体系, 加强对农业科技信息资源的规范化管理。当前国内农业科技信息资源虽然规模庞大, 但是并未效仿 FAO 建立高效一致的管理标准体系, 导致资源建设缺乏一定的整体性和系统性, 对实现农业科技信息资源开放共享带来了一定的局限性。为此, 可基于 FAO 所推崇的广域合作的数据联盟机制, 按照统一标准和规范将零散的农业科技信息资源进行集中注册, 构建统一的信息

资源门户网站负责集成各类资源, 为用户提供一站式数据检索和定位服务, 并由分布式的数据服务系统提供数据获取服务^[20], 构建一个集成农业领域的各项信息资源和技术工具的标准化农业信息管理体系, 以保证用户可以通过该体系高效获取农业领域的一系列信息标准、技术和应用实践, 同时推动用户参与农业科技信息资源开放共享的积极性。

(2) 增强农业科技信息资源的语义关联与互操作性, 提高资源的开放共享能力和水平。中国目前现有的农业科技信息资源数据库多数是以相互独立的形式存在, 农业科技信息资源间的语义关联性和互操作性较差, 从而阻碍了进一步的农业科技信息资源开放与共享。因此, 国内需积极践行可发现、可获取、可互操作和可重用的 (FAIR) 国际准则, 加强知识组织体系、关联数据、知识图谱等语义知识组织技术创新应用, 增强农业科技信息资源的语义关联与互操作性, 例如对中国农业科学叙词表 (CAT) 进行不断完善和维护以实现 CAT 的多途径开放共享, 同时参考借鉴 FAO 推动农业科技信息资源开放共享实践的总体思路、工作流程、协同编辑等方面, 将新兴技术与包含丰富语义关系的知识组织体系进行高度融合, 最终提高中国农业科技信息资源的开放共享能力和水平。

(3) 加强与国际农业信息相关机构的合作, 提升中国农业科技的国际影响力。随着中国纵深推进乡村振兴、“藏粮于技”等重大战略, 加快实现由农业大国向农业强国转变, 并在国际事务中发挥主导引领作用, 这对中国农业科技信息资源开发利用提出了新需求和新要求。因此, 下一步更加需要加强与 FAO 等国际农业组织机构的合作交流, 进一步深化农业科技信息资源开放利用。例如基于前期与 FAO 在叙词表建设与映射、文献数据共享等基础上, 组建联合研究中心或实验室, 将大数据、人工智能等技术与农业学科领域交叉融合, 构建区域性或全球性农业科技大数据采集与治理平台, 充分整合全球农业科技文献、科研数据、海外数据、专业领域特色资源等, 建设全球农业大数据与信息服务平台和开放型知识交流传播体系, 支撑引领中国农业科研创新向数据密集型科学发现第

四范式和开放科学转型, 为提升中国农业科技的国际影响力发挥更大作用。同时, 在推进中国农业科技知识发现服务时, 可参考 FAO 的资源管理模式, 加强对国际社区可开放获取的农业科技信息资源的分析、收集及整合, 加快实现面向全球范围的农业科技信息资源知识发现服务。

参考文献:

- [1] KROES N. Opening up scientific data[EB/OL]. [2020-09-15]. https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_13_236.
- [2] NIELSEN M. An informal definition of OpenScience[EB/OL]. [2020-09-15]. <http://openscience.org/an-informal-definition-of-openscience/>.
- [3] MONS B, NEYLON C, VELTEROP J, et al. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud[J]. Information services & use, 2017, 37(1): 49-56.
- [4] RETTBERG N, SCHMIDT B. OpenAIRE - Building a collaborative Open Access infrastructure for European researchers[J]. Liber quarterly the journal of European research libraries, 2012, 22(3): 160-175.
- [5] QIN J, COLL I S, ZENG M L, et al. Technical and policy underpinnings of FAIR data principles[J]. Proceedings of the association for information science and technology, 2019, 56(1): 569-571.
- [6] BARTOL T. Information literacy and international capacity development initiatives in life sciences: AGORA, OARE, HINARI, ARDI (Research4Life-R4L)[C]//European conference on information literacy, Springer, Cham, 2013: 338-344.
- [7] JONQUET C, TOULET A, ARNAUD E, et al. AgroPortal: A vocabulary and ontology repository for agronomy[J]. Computers and electronics in agriculture, 2018, 144: 126-143.
- [8] 顾立平, 赵昆华, 段美珍, 等. 开放获取跟踪扫描 2016 第 05 期[J]. 开放获取跟踪扫描报告, 2016(5).
- [9] SHONHE L. Sharing open data in agriculture: a learning curve for developing countries[M]. Open access implications for sustainable social, political, and economic development, IGI global: 244-266.
- [10] MALAPELA T, SUBIRATS I, DISTER S, et al. A reference Space for information management standards, tools and methodologies: the AIMS portal[C]// Ifla Wlic, 2013.
- [11] PESCE V. The role of the CIARD RING in the building of the CIARD framework for data and information sharing: now and in the future[J]. Annals of the New York academy of Sciences, 2011, 95(95): 969-974.
- [12] SUBIRATS I, ZENG M. Meaningful bibliographic metadata (M2B): Recommendations of a set of metadata properties and encoding vocabularies.[Rome]: Agricultural information management standards (AIMS), food and agriculture organization (FAO) of the United Nations[J]. Retrieved may, 2012, 8: 2012.
- [13] SUBIRATS I, MALAPELA T, DISTER S, et al. Reorienting open repositories to the challenges of the semantic web: Experiences from FAO's contribution to the resource processing and discovery cycle in repositories in the agricultural domain [C]//Research conference on metadata and semantic research, Springer, Berlin, Heidelberg, 2012: 158-167.
- [14] STELLATO A, FIORELLI M, TURBATI A, et al. VocBench 3: A collaborative semantic web editor for ontologies, thesauri and lexicons[J]. Semantic web, 2020(1): 1-27.
- [15] The aims team. Watch: AGROVOC presents work on COVID-19 in the multilingual thesaurus during NKOS 2020 workshop [EB/OL]. [2020-09-15]. <http://aims.fao.org/activity/blog/watch-agrovoc-presents-work-covid-19-multilingual-thesaurus-during-nkos-2020-workshop>.
- [16] BAKER T, WHITEHEAD B, MUSKER R, et al. Global agricultural concept space: Lightweight semantics for pragmatic interoperability[J]. NPJ science of food, 2019, 3(1): 1-8.
- [17] CHARALAMBIDIS A, TROUMPOUKIS A, KONSTANTOPOULOS S. SemaGrow: Optimizing federated SPARQL queries[C]//Proceedings of the 11th international conference on semantic systems, 2015: 121-128.
- [18] CELLI F, KEIZER J, JAKUES Y, et al. Discovering, Indexing and Interlinking Information Resources[J]. F1000research, 2015, 4: 432.
- [19] 赵瑞雪, 赵华, 朱亮. 国内外农业科学大数据建设与共享进展[J]. 农业大数据学报, 2019, 1(1): 24-37.
- [20] 李成赞, 张丽丽, 侯艳飞, 等. 科学大数据开放共享: 模式与机制[J]. 情报理论与实践, 2017, 40(11): 45.