

英国环境领域数据发布平台发展现状与内容剖析*

姜颖

(对外经济贸易大学图书馆, 北京 100029)

摘要: 自然科学数据经过仪器观测、实验、仿真等方式产生, 具有重要参考、重用和研究价值。环境领域自然科学数据获取成本巨大且具有唯一性, 其共享对科研结论验证、数据再利用和节约成本意义明显。以英国环境领域数据发布平台为调研对象, 从平台概况、建设目标、数据来源及规模、服务方式及管理政策多方面介绍建设现状, 从基本模块、检索服务、浏览与导航服务、咨询/帮助服务角度分析平台内容, 归纳平台特点, 以期为我国平台建设提供借鉴。

关键词: 自然科学数据; 环境科学; 数据服务; 数据管理

中图分类号: G252.1

DOI: 10.3772/j.issn.1673-2286.2018.02.009

1 科学数据平台是推动资源共享的基础

美国国家科学基金会 (National Science Foundation, United States, NSF) 在《面向21世纪探索的网络基础设施》报告中指出“科技研究和教育越来越朝向数据密集型转变, 即以数字技术、仪器仪表和各种网络来收集、生成、共享和分析数据, 并进行扩散。每天通过实验、观察和模拟进行生产、获取、分析、整合和存储的数据高达吉字节 (G-bytes)”^[1]。在数据密集型科研环境下, 数据或事实是科研工作的根本, 在缺乏数据佐证或事实依据的情况下, 人们无法进行科研成果的验证, 科学结果与论点无法成立^[2]。

科学数据是科研人员在科学研究过程中通过科学实验、实际调查等方式产生和获得的数据资料^[3], 伴随科研活动的整个生命周期, 真实记录科学研究的全过程, 对科学研究成果具有直接的支撑和佐证作用^[4]。自然科学是研究自然界的物质形态、结构、性质和运动规律的科学, 包括数学、物理学、化学、生物学、天文学等基础科学和医学、农学、气象学、材料学等应用科学^[5], 自然科学数据大多经过仪器观测、实验、探测、仿真等

方式产生, 其创建与获取大多需特定的研究对象、实验试剂、仪器设备、实验环境和专人引导等因素支持。一些自然科学领域的环境科学数据, 如外太空的观察数据、深海的地质数据和大规模核辐射对环境危害的数据等, 不仅获取成本巨大, 且再获取性较低^[6]。除国家保密因素外, 环境科学数据的共享对于环境领域科研结论的重复验证、数据的再利用和投入成本的节约意义明显; 而环境科学数据的发布平台是实现和推动自然环境科学数据共享的基础与前提^[7]。

英国自然科学研究水平居于世界前列, 自然科学数据管理和服务工作亦起步较早, 已有多项自然科学数据发布平台和管理项目问世, 内容涉及地球与环境科学、天文学、生物学、医学等多个领域^[8]。自2004年1月OECD成员国签署《开放获取公共资助研究数据的宣言》, 英国自然科学领域数据发布平台的建设随即提上日程。作为主要研究资助机构的英国七大研究理事会: 自然环境研究理事会 (Natural Environment Research Council, NERC)、艺术与人文研究理事会、生物技术与生物科学研究理事会、工程与自然科学研究理事会、经济与社会研究理事会、医学研究理事会、科

* 本研究得到对外经济贸易大学2015年度校级科研课题“学术机构知识库支持科研及优化策略研究” (编号: 15YB06) 资助。

学与技术设施研究理事会,成为科学数据发布平台建设的主要机构。NERC作为首轮加入OECD国家科学数据共享计划的公共机构^[9],目前在数据政策制定和数据共享平台建设方面,均较成熟和先进。由其建设的5个环境科学数据共享平台:国家地球科学数据中心(National Geophysical Data Center, NGDC)、环境信息中心(Environmental Information Data Centre, EIDC)、海洋数据中心(British Oceanographic Data Center, BODC)、环境信息分析中心(Centre for Environmental Data Analysis, CEDA)和英国极地数据中心(UK Polar Data Centre, PDC),发布的数据涉及陆地、淡水、海洋、冰川、气候、极地等多个自然环境领域,内容涵盖性强,其建设较成熟且极具代表性,是环境科学领域甚至自然科学领域科学数据发布平台建设的典范。

本文选取上述5个平台作为调研对象,调研时间为2017年10月26日~11月25日。采用网络调研法和文献内容分析法在调研各平台建设现状的基础上分析其发布内容,并对其建设特点进行总结,以期为我国环境科

学领域甚至整个自然科学领域科学数据平台建设提供参考。

2 英国自然环境科学数据发布平台建设现状

2.1 平台概况

如表1所示,5个平台管理和发布的科学数据数量众多,且分属不同的环境科学领域。其中,NGDC主要发布地球科学领域的数据和信息,EIDC侧重于陆地和淡水科学领域,BODC发布生物学、化学、物理学和地球物理领域的海洋环境数据,CEDA主要提供气候、太阳能和地球观测的卫星数据,PDC负责发布极地数据。5个平台均受到NERC的资助,并得到各自领域国家重点数据监管和调查机构的数据支持。如NGDC得到英国地质调查局(British Geological Survey, BGS)的数据支持,EIDC得到英国生态水文中心(Center for Ecology & Hydrology, CEH)的数据支持。

表1 5个平台基本情况统计

平台名称	数据范围	平台概况	建设目标
NGDC	地球科学数据	由BGS主持、NERC资助搭建,管理地球科学领域超过400个数据集	收集、保存和管理地球科学领域有价值的数据和信息,并将其广泛提供给全球用户
EIDC	陆地和淡水环境资源数据	由CEH主持、NERC资助搭建,负责管理和发布与陆地和淡水科学有关的国家重要数据集	广泛收集、长期保存和管理环境科学信息数据,为研究人员和研究组织提供长期的数据管理计划,保证数据的长期保存和公共传播
BODC	海洋环境数据	受NERC资助,存储管理和发布有关海洋环境数据,维护和开发国家海洋学数据库	大量收集和整理可公开获取的生物学、化学、物理学领域的海洋数据,并将其广泛提供给全球用户
CEDA	气候、地球观测、太阳能系统数据	NERC资助搭建,前身为科学技术设施研究理事会Rutherford Appleto实验室,存档和管理各种形式的气候、太阳能系统以及各种地球观测数据	长期收集、归档保存并管理气候、太阳能和地球观测数据,并将其广泛提供给全球用户
PDC	极地数据	NERC资助搭建,实地调查、收集、管理极地数据等信息	长期致力于确保在极地地区采集的环境数据的公共访问和广泛提供

2.2 平台建设目标

5个平台的建设目标既具有一致性,又具有特殊性。一致性体现在都是为收集、保存和管理各领域有价值的科学数据,方便公众对科学数据的发现、获取和使用,促进自然科学数据共享管理的进一步发展,提高科学研究的效率和社会价值;特殊性体现在不同类型

的环境数据平台现阶段目标侧重点有所区别。表1中NGDC负责国家地球科学领域科学数据的收集、保存和管理,主要目标是将收集的地球科学数据广泛提供给全球用户,PDC则长期致力于确保在极地地区采集的环境数据的公共获取和广泛提供,这两个平台的建设目标在于保证数据的全球共享;而EIDC广泛收集、长期保存和管理环境科学信息数据,为研究人员和研

究组织提供长期的数据管理计划、保证数据的长期保存,其主要目标为数据的长期管理与保存。

2.3 平台数据来源及规模

平台的数据来源主要有两种,一是自行研究产生的数据,二是收集其他机构或个人的科学数据^[10]。自行研究产生的数据指研究机构将日常工作中得到的数据建成结构化的数据集,如PDC发布的数据全部都是数据中心的自有数据,即极地数据中心在极地地区利用仪器、卫星等工具实地采集的环境数据。而其他4个平台的数据来源既包含自行研究数据也包含收集的其他项目或人员的数据。如EIDC既要管理和发布CEH以及NERC环境科学领域研究产生的数据,也要承担其他研究项目和人员的数据管理委托。

5个平台的数据量和类型均各具规模。根据平台的统计数据,CEDA包含5 000多个数据集;NGDC拥有地球科学领域400多种形式的数据集,包括环境监测数据,在线电子数据库(如岩石收藏数据库),钻孔岩芯、岩石、矿物和化石数据的纸质数字资料收藏,照片扫描本和地图数据等形式;BODC拥有海洋科学领域超过2.2万个测量变量的数据集,数据类别包含声学、测深和地形、海平面、声像学、盐度、水柱化学等。

2.4 平台服务方式

自然科学数据发布平台是负责收集、处理、加工信息的机构,也是对科学数据实施管理的机构^[11]。环境科学数据发布平台的服务内容包含科学数据的收集、处理、加工和管理,服务方式包括开放数据服务、数据管理服务、数据检索服务、数据存储服务、数据咨询服务和数据分析服务(见表2)。

表2 5个平台的数据服务方式

平台名称	服务方式					
	开放数据	数据管理	数据检索	数据存储	数据咨询	数据分析
NGDC	√	√	√	√	√	-
EIDC	√	√	√	√	√	-
BODC	√	√	√	√	√	-
CEDA	√	√	√	√	√	√
PDC	√	√	√	-	√	√

注:“√”表示存在此项服务方式。

其中,开放数据、数据管理、数据检索和数据咨询是5个平台均可提供的服务方式。

(1) 开放数据服务。开放数据即面向公众的免费数据服务,以开放数据目录的形式向公众提供科学数据列表及内容。5个平台均采用部分数据开放的模式。如NGDC将开放数据服务命名为“Open Geoscience”,遵循开放政府许可^[12]针对公众提供的免费数据服务;CEDA将平台数据分为公共数据(即开放数据服务)、注册用户可访问数据、有限制数据和永久限制数据。

(2) 数据管理服务。数据管理包括数据选择、组织、描述、保存、访问、使用和评估等工作过程和实践操作,涵盖指定标准、元数据创建和文献链接等活动^[13]。EIDC和CEDA均向用户提供科学数据管理计划,EIDC指出数据管理计划是确保项目成果价值实现的重要步骤,EIDC配备专门的数据管理专员,在项目之初予以指导,并在之后的工作中将管理不断调整和细化。

(3) 数据检索服务。作为常规服务项目,5个平台均根据自身数据特点向用户提供多种检索途径和方法(如主题组合检索、专业检索式检索以及地图检索等),帮助用户定位目标数据。

(4) 数据存储服务。平台除对自行研究产生的数据进行存储外,还可提供将其他机构或个人所提交数据的长期保存服务。除PDC外,其他4个数据平台均对外提供数据存储服务。

(5) 数据咨询服务。以疑问解答为目的的数据咨询服务是发布平台的必备服务,NGDC和BODC甚至专设咨询服务岗位。

(6) 数据分析服务是面向用户的高级数据服务形式,CEDA和PDC均提供专门的数据深层分析服务,可根据用户个人需求定制数据。

2.5 平台管理政策

5个平台的数据管理政策大致可分为数据质量监管政策、数据存储政策和数据传播政策。如表3所示,数据质量监管政策包括制定数据管理计划、对数据管理执行情况的监督和数据管理指导与服务,数据存储政策包括数据标准、元数据标准、数据保存和数据范围,数据传播政策包括数据共享、数据引用、数据版权、数据安全和数据访问。

(1) 数据质量监管政策。主要针对其他研究机构和个人提交的数据,因此除PDC外(PDC只存储自行

表3 5个平台的管理政策

平台名称	数据质量监管政策			数据存储政策				数据传播政策				
	数据管理计划	数据管理执行情况的监督	数据管理指导与服务	数据标准	元数据标准	数据保存	数据范围	数据共享	数据引用	数据版权	数据安全	数据访问
NGDC	-	-	√	√	√	√	√	√	√	√	-	√
EIDC	√	√	√	√	√	√	√	√	√	√	√	√
BODC	-	-	√	√	√	√	√	√	√	√	√	√
CEDA	√	-	√	√	√	√	√	√	√	√	-	√
PDC	-	-	-	√	√	√	√	√	√	√	√	√

注：“√”表示存在此项政策内容。

研究产生的数据),其他4个平台均制定了相关数据质量监管政策,对数据管理提供指导与服务。EIDC和CEDA均要求数据提交者提供数据管理计划,计划通常包括数据采集、整理、分析、存储等,是保障数据质量的重要途径^[4]。此外,EIDC还对数据管理执行情况提供监督。

(2) 数据存储政策。数据存储可分为两大类,一类是平台自有数据的存储,另一类是其他机构或个人的数据存储。PDC仅支持本单位实地勘测和研究整理的数据存储,其他4个平台既支持本单位也支持其他机构或个人的相关领域数据存储。5个平台均对其平台数据标准、元数据标准和数据范围作出明确说明。如EIDC数据标准包括数据格式,数据是否被用在同行评审出版物,数据是否符合收录主题,数据是否可重复利用,数据范围不受地理位置和时间的限制,数据类型包含表格、图像、数据库、软件、模型等多种形式。同时,EIDC提供专门的元数据标准指导,指明其元数据标准高于英国位置计划“the UK Location Programme”^[15]的界定标准;BODC数据存储总则对数据提交机制、数据文件格式和文件命名、参数说明、附带元数据等均作出明确界定,并针对其包含的11种元数据(如电流表数据、海平面数据、水样数据、浮标数据等)存储给予专门说明。

(3) 数据传播政策。包括数据共享、数据引用、数据版权、数据安全和数据访问政策,5个平台均对数据传播政策进行说明,其中BODC、EIDC和PDC对数据安全作了规定。BODC规定数据所有权不受存储管理的影响,即数据所有权归属于数据发起人或与研究机构保持一致;EIDC要求项目数据由管理专员负责数据的定期备份,以防止软硬件故障、病毒感染或数据丢失。备份程序取决于当地情况、数据的感知价值和风险水平。如果数据包含个人信息,应注意创建副本的最小数

量;PDC在其隐私保护政策中明确说明PDC会第一时间保护用户安全和隐私,绝对不会出售、出租或与第三方共享用户个人信息,并从安全的角度指导用户进行浏览器、IP、服务器等的设置。

3 数据发布平台内容分析

5个环境科学数据发布平台已建设得较为成熟,网站模块齐备,内容丰富。下面从基本模块、检索服务、浏览与导航服务、咨询/帮助服务四个角度对平台内容进行分析,如表4所示。

3.1 基本模块

5个平台在基本模块的设置上特点鲜明,可归纳为共有模块和个性模块两类。其中数据检索和数据引用为5个平台网站共有的模块设置。通过检索服务,用户可直接利用数据关键词的简单检索或高级检索功能直接定位所需数字资源。科学数据引用具有体现已有工作价值、显示已有数据价值、保证数据的可用性等重要意义^[16]。5个平台引用模块包含被引数据目录(标题、作者和DOI)、引用数据集、数据引用的过程描述等内容。除PDC以外的其他4个平台均提供数据存储服务。在存储模块中,各平台均发布数据存储指南,如《NGDC数据存储指南》^[17]对存储数据的范围、格式、原则等问题进行详细说明,按照存储步骤为用户提供网站存储导引,或通过提供存储数据模板指导用户完成数据存储。

此外,不同平台也按照自身特色设置了个性化模块。如NGDC包含数据管理和开放数据模块,数据管理模块对平台数据管理给予指导,开放数据模块是平台

表4 5个平台网站内容分析

平台名称	基本模块	检索服务	浏览与导航服务	咨询/帮助服务
NGDC	数据检索、数据存储、数据引用、数据管理	数据目录检索：数据关键词简单检索。 地图检索：利用地图中地理位置，检索此位置相关数据信息	浏览：关键词浏览（按字母顺序A~Z）、数据集浏览（按字母顺序A~Z）、按关键词类别浏览元数据、按地理位置浏览数据集、按主题浏览以及目的浏览数据	邮件咨询、人工咨询服务岗（图形与数据显示咨询、数据请求和传递咨询）
EIDC	数据检索、数据存储、数据引用、数据支持	数据目录检索：数据关键词简单检索、高级检索（利用不同选项关键词和逻辑算符及其他位置算符等构造专业的检索式）、地图检索	资源导航	邮件咨询、现场人工咨询、网站帮助页面（视频与文本教程、FAQs）
BODC	数据检索、数据存储、数据引用	数据目录检索：关键词简单检索、组合检索（时间、数据类型等组合选项）	资源导航、项目导航	邮件咨询、电话咨询、传真咨询、现场人工咨询
CEDA	数据检索、数据存储、数据引用、数据分析（数据模型、定制数据）	数据目录检索：简单检索、高级检索。 元数据检索：关键词简单检索。 分类检索：项目、仪器、平台、计算	浏览：按照数据集、元数据、项目、仪器、平台、计算等进行浏览	邮件咨询、电话咨询、传真咨询、帮助文档
PDC	数据检索、数据管理、发现元数据、数据引用和出版	数据目录检索：简单检索。 多类型数据库系统检索：简单检索、组合检索。 元数据检索：关键词简单检索和高级检索	按主题浏览数据，研究项目导航	邮件咨询、电话咨询、传真咨询

面向公众免费开放数据的集合；EIDC专门设立数据支持模块对平台的科学数据管理计划及其监督机制给予详细指导和说明。

3.2 检索服务

检索服务是用户定位目标数据的直接手段。数据目录检索是5个自然数据平台均可提供的检索服务。在数据目录中，利用数据关键词进行搜索，以确定目标数据。除目录检索外，由于平台数据类型、数据开放形式等因素不同，各平台可提供的其他检索方式相差较大，有些平台仅提供简单检索，有些平台则提供多种复杂检索方式（多主题检索、检索式检索等）。

NGDC的目录检索只有一个简单检索的检索框，检索精准度不高；而EIDC提供不同选项（主题、机构、作者、日期等）和逻辑算符（与、或、非）及其他位置算符等高级检索方式；BODC的检索界面包含时间范围、数据类型、参数选择、项目、平台、仪器、水深、仪器深度、

国家、文件格式等多个选项，供读者组合检索；PDC将数据搭建为不同的数据系统，如发现元数据系统、南极数据系统、大气数据系统、极光边界图像数据库和极地航空地球物理数据门户等，分别提供检索服务。

除数据检索外，NGDC、CEDA、PDC还提供发现元数据的简单或高级检索服务。对于检索结果，5个平台均要求访问者注册为系统用户方可下载数据，并且存在一部分受限和须付费购买的数据，用户可通过有效身份认证对限制数据提出申请，申请通过后才可访问数据。

3.3 浏览与导航服务

数据浏览和资源导航功能是用户快速了解平台资源和服务的有效途径。5个平台结合自身数据类型、数据开放形式等因素，提供不同程度的资源浏览和导航服务。如NGDC提供数据关键词字母顺序浏览、数据集字母顺序浏览、元数据关键词归属类别浏览、数据集地

理位置浏览、主题（如3D建模、土地利用规划、地震、能源等）浏览以及目的（施工、挖掘、引流、农业、栖息地分析等）浏览；PDC将南极数据系统的数据按照冰雪、臭氧、气候、地图和地名等主题排列，用户可按主题浏览数据。

3.4 咨询/帮助服务

咨询/帮助服务是5个数据服务平台共有的服务项目，能够及时解答用户在获取数据、存储数据和使用数据过程中遇到的问题。目前5个平台提供的咨询服务多样，包含邮箱咨询、电话咨询、人工现场答疑、网页FAQs、各类入门和使用视频教程、用户手册等。如NGDC在数据部门下设专职人工咨询服务岗位，专门解答用户在图形与数据显示与数据请求传递等方面的问题，并在每一个具体服务页面下方提供不同咨询邮箱；EIDC在提供邮箱咨询的同时，配以专门的网站帮助页面，页面中包含不同问题的回答集锦、存储数据的使用视频、数据教程下载、扩展功能介绍和用户注册流程等。

4 数据发布平台建设内容总结及启示

5个英国自然科学数据发布平台是国际上相对成熟的建设范例，无论是服务方式、管理政策还是网站内容配置方面均具有借鉴价值，可以为我国科学数据平台建设提供经验参考。

4.1 数据来源与经费来源渠道稳定且可持续

平台的数据来源渠道主要有政府部门或基金会的资助、高校或研究机构的资助以及私营部门的捐赠^[18]。相比而言，前两种资助渠道的稳定性和可持续性较高。本文调研的数据平台均受NERC资助。很多受到政府部门、基金会、科研机构等资助的数据平台都是基于该基金会或科研机构所资助的研究项目数据搭建起来的^[19]。如作为科研资助机构，NERC于2011年发布科学数据政策^[20]，要求受NERC资助的科研项目，其科学数据必须得到长期保存，在数据中心的说明页面明确规定NERC资助的自然科学类项目的数据必须存储到由其资助搭建的相关类别的自然数据发布平台中^[21]；同时，数据发布平台的数据支持单位也是数据来源的重要保障，如BGS和CEH均是平台的数据支持机构，数

据来源可靠且稳定。

4.2 服务内容丰富多样

平台数据服务的内容不仅代表平台当前服务现状，也代表平台今后的服务发展方向。经过不断的动态调整和完善，5个平台可提供的数据服务内容已涵盖数据生命周期的多个环节，不仅包括数据存储咨询、数据发现支持、数据标准指南等咨询型服务，还包含开放数据、数据处理、数据搜索（数据定制）、数据分析和数据管理等技术型服务。同时，5个平台还根据自身特点，提供特色服务。如CEDA考虑到用户对不同层次数据需求，为注册用户提供深层次数据检索、分析和定制服务；EIDC在提供帮助用户制定科学数据管理计划服务的同时，还对计划的实施提供监督和指导。

科学数据的管理政策涉及数据存储、数据质量监管和数据传播等多个方面，5个平台科学数据管理政策制定得较为齐备，从数据存储标准、元数据标准、数据类型范围到数据管理计划和监督，以及数据共享、发布和出版，均作了详细的政策文件导引。如EIDC的《EIDC数据管理计划指导手册》^[22]和《EIDC元数据指导手册》^[23]，BODC的《BODC数据存储总则》^[24]，以及CEDA的《CEDA数据存储步骤》^[25]等，有效地保证了发布平台各环节工作的开展。

4.3 具备明确的权利责任机制

数据发布平台涉及数据的收集、保存、管理、发布等一系列环节，明确的权利责任机制将有助于各环节的实施和完善。明确权利责任机制不仅包括对用户和机构的权利责任限制和说明，也包括数据服务机构对其内部职位权利责任的部署。

对用户和机构的权利责任限制和说明如BODC要求用户存储数据前，同意其权利义务条款，具体包括：①确认本人是数据发起人或获得数据发起人的许可，同意BODC数据存储条款；②数据所有权归属与数据发起人或研究机构保持一致；③通过授予BODC许可以存储数据，保证数据永久存储于BODC，并满足BODC数据政策规定的数据使用和重复使用条款；④授予BODC将数据格式转换为其他便于保存和访问的格式。对其内部职位权利责任的部署如NGDC设置专门的数据服务部门，提供数据服务经理岗位，并详细制定

其工作内容,包括管理BGS提供和传递的数据项目、提供对BGS数据合作伙伴的技术解决方案、策划有助于BGS数据合作伙伴的商业类型和开放获取类型的系列活动、探索提高数据访问能力的方法、研究提供数据的多种创新方法(如Web服务和智能手机服务等);同时,设置知识产权与授权经理、数据显示与图形设计工程师、数据处理和传递咨询专员,并对相关岗位权利与义务作出明确界定。

4.4 平台网站内容布局合理且界面友好

作为自然科学数据发布平台,数据的浏览和导航、数据检索和存储模块是核心部分。经过不断试验调整,5个平台网站均采用简明实用的内容布局方式,即在主页中将关键功能作为单独模块,便于用户第一时间定位需求和服务。同时根据不同数据领域特点,配置新闻、热点项目和咨询服务模块。数据浏览和导航模块是向用户展现平台资源的直接方式,平台提供多种浏览和导航方式供用户选择,如按照字母字顺、元数据类别、数据类别、地理位置、数据采集时间等;检索模块由于受到简单关键词检索的限制,提供专业检索和组合检索等高级检索方式;存储模块提供步骤引导指导用户依照流程完成存储。整体而言,5个平台网站界面用户友好程度较高。

5 结语

英国环境领域自然科学数据发布平台的建设经验表明,完善的数据管理和服务不仅可以使历史数据得以归档、长期保存和共享利用,也可以推动新增数据的业务化管理和在线发布,极大地推动科学数据的有效管理和利用。由于我国还处于科学数据管理服务的起步阶段,在未来的发展中会遇到很多的困难与挑战,如缺少相应的政策与经费支持、数据保存与共享格式难以统一、缺乏科学数据长期保存技术与机制等。英国环境领域科学数据服务在这些方面已经作出大胆尝试,为我国相关机构在开展服务方面提供参考和借鉴。因此,我国在开展科学数据共享服务时,可根据我国国情充分了解不同领域存储和共享的需求和特点,并合理借鉴国外的成熟经验,以便更好地开展此项服务。

参考文献

- [1] NSF. Cyberinfrastructure Vision for 21st Century Discovery [EB/OL]. [2017-10-13]. <https://www.nsf.gov/pubs/2007/nsf0728/>.
- [2] Goporis Conference 2013 on Non-Textual Information Strategy and Innovation Beyond Text [EB/OL]. [2017-07-28]. <http://www.nontextualinformation2013.de/index.php/programme>.
- [3] HYOUNGJOO P, DIETMAR W. An examination of research data sharing and reuse: implications for data citation practice [J]. *Scientometrics*, 2017, 111: 443-461.
- [4] ROBERT B, SURESH K S. Implementation of data citations and persistent identifies at the OPNL DAAC [J]. *Ecological Informatics*, 2016 (33): 10-16.
- [5] Wikipedia. Natural science [EB/OL]. [2017-10-28]. https://en.wikipedia.org/wiki/Natural_science.
- [6] GORLIZ O, STAAB S. Federated data management and query optimization for linked open data [J]. *New Directions in Web Data Management*, 2011, 331: 109-137.
- [7] SILVIO P, ALEXANDER D, TANYA G. Setting our bibliographic references free: towards open citation data [J]. *Journal of Documentation*, 2015, 71 (2): 253-277.
- [8] Symposium On International Scientific Data Sharing [EB/OL]. [2017-10-29]. <https://www.cni.org/news/symposium-on-international-scientific-data-sharing-april-18-19-washington-dc>.
- [9] 顾立平. 科学数据开放获取的政策研究 [M]. 北京: 科学技术文献出版社, 2016: 108.
- [10] 司莉, 邢文明. 科学数据管理与共享的理论与实践 [M]. 武汉: 武汉出版社, 2017: 21.
- [11] 邱春艳. 欧盟科学数据开放获取实践及启示 [J]. *情报理论与实践*, 2016, 39 (11): 138-144.
- [12] Open Government Licence for public sector information [EB/OL]. [2018-01-01]. http://www.centralbedfordshire.gov.uk/Images/open-government-licence_tcm3-11743.pdf.
- [13] Guidelines for Responsible Data Management in Scientific Research [EB/OL]. [2018-01-01]. <https://ori.hhs.gov/images/ddblock/data.pdf>.
- [14] Scientific Data Management in the Coming Decade [EB/OL]. [2018-01-01]. <http://www.productmanualguide.com/newpdf/scientific-data-management-in-the-coming-decade.pdf>.
- [15] Metadata Guidelines2 [EB/OL]. [2017-11-13]. <http://www.agi.org.uk/storage/standards/uk-gemini/MetadataGuidelines2.pdf>.

- [16] MIKE T, PAUL W. Regression for citation data: an evaluation of different methods [J]. *Journal of Informetrics*, 2014 (8): 963-971.
- [17] British Geological Survey. Good data deposit guidelines [EB/OL]. [2017-11-11]. <http://www.bgs.ac.uk/services/ngdc/goodData.html>.
- [18] 国家科技基础条件平台中心. 国家科学数据资源发展报告 2016 [M]. 北京: 科学技术文献出版社, 2016: 45.
- [19] 黄国彬, 屈亚杰. 英国科研资助机构的科学数据共享政策调研 [J]. *图书馆论坛*, 2017 (5): 124-132.
- [20] NERC. Data centres [EB/OL]. [2017-11-13]. <http://www.nerc.ac.uk/research/sites/data/policy2011.asp>.
- [21] NERC Data Policy [EB/OL]. [2017-10-30]. <http://www.nerc.ac.uk/research/sites/data/policy/data-policy/>.
- [22] EIDC Data Management Guidance & Template [EB/OL]. [2017-11-11]. <http://eidc.ceh.ac.uk/support/dataManagementGuidance>.
- [23] EIDC. Metadata Guidance [EB/OL]. [2017-11-11]. <http://eidc.ceh.ac.uk/deposit/metadata-guidance>.
- [24] British Oceanographic Data Centre. Submitting data to BODC [EB/OL]. [2017-11-11]. https://www.bodc.ac.uk/submit_data/submission_guidelines/.
- [25] Centre for Environmental Data Analysis. Steps to archiving data with CEDA [EB/OL]. [2017-11-12]. <http://help.ceda.ac.uk/article/138-steps-to-archiving-data-with-ceda>.

作者简介

姜颖, 女, 1984年生, 馆员, 硕士, 研究方向: 科学数据管理, E-mail: uibeijing@126.com。

Development Status and Content Analysis of Publishing Platform of Environmental Natural Science Data in UK

JIANG Ying

(Library of University of International Business and Economics, Beijing 100029, China)

Abstract: Most of the natural science data were produced by instrumental observation, experiment, exploration and simulation, and which has great value in scientific research. The environmental science data such as outer space observation data and deep geological data has great meaning in data sharing, data validation and science research conclusion of reuse as its huge cost and uniqueness. The paper selects 5 more mature science data publishing platform as research object. By investigating related websites and literatures, it first introduced the present situation of the construction of platform, including platform overview, construction goal, resource scale, service mode and management policy; then analyzed the content of the platform's web site from the basic module, retrieval service, navigation services, consulting services. Finally, it summarized the construction features of the five data publishing platforms and provided reference and inspiration for the science data platform construction in China.

Keywords: Natural Science Data; Environmental Science; Data Service; Data Management

(收稿日期: 2018-01-05)

《数字图书馆论坛》在2016年度 复印报刊资料转载指数排名中喜获佳绩

由中国人民大学人文社会科学学术成果评价研究中心联合书报资料中心研制的2016年度复印报刊资料转载指数排名于2017年3月28日正式发布。

在“图书馆、情报与档案管理学科期刊”全文转载排名中,《数字图书馆论坛》转载率位列第15名,综合指数位列第20名。

该排名根据人大复印报刊资料近100种学术系列期刊在2015年度转载的学术论文数据,从转载量、转载率、综合指数三个维度对中国人文社科期刊和教学科研机构进行统计形成。