

科学数据出版成果著录规范化研究

刘 闯^{1*}, 马军花¹, Paul F. Uhler², 石瑞香¹

1. 中国科学院地理科学与资源研究所, 北京 100101, 中国;

2. Consultant on Information Policy and Management, 纽约 12723, 美国

摘 要: 近些年来, 科学数据出版发展迅速, 出版的数据成果越积越多。这些数据成果在被应用过程中如何被标注和著录成为科学数据应用过程中重要的问题之一。科学数据著录坚持开放性原则符合国际组织一致倡导的科学数据开放和共享原则。本文列举了目前科学数据应用过程中采用论著文本中单独段落说明数据来源、在致谢中说明数据来源、在作者分工中包含部分数据来源信息等方式著录的案例, 提出了将科学数据分为五种类型, 数据著录分为两种类型的学术观点。五种数据类型分别为: 正式出版的数据、世界数据系统(WDS)公开发布的数据、国家级数据中心发布的数据、个人或团队发布的数据和内部使用的数据。两种著录方式分别为参考文献著录方式和脚注著录方式。其中, 将有出版资质单位正式出版公开发行的数据、世界数据系统公开发布的数据、国家级数据中心公开发布的数据采用参考文献著录方式; 个人或团队发布的数据和内部使用的数据采用脚注著录方式。本文同时对实体数据、数据记录、数据论文的引用著录格式进行了探讨, 以示例的方式说明了不同数据类型所采用的不同标注和著录的格式。作者认为在科学数据出版和公开发表的方式与渠道多样化的情况下, 广泛应用、规范引用和标准化著录这些数据成果对推动科学数据开放共享、提高数据质量、尊重数据原创者知识产权, 避免“用而不引”、“引而不标”、“标不规范”的现象、更好地促进科研大数据的快速发展均具有重要的实践意义。

关键词: 科学数据; 数据类型; 著录方式; 参考文献; 脚注; 规范化

DOI: 10.3974/geodp.2018.02.01

1 问题的提出: 科学数据出版成果是否需要规范化著录?

科学数据是科学研究基础性资源^[1], 是科技活动重要产出成果, 也是科学发现必不可少的表达和佐证。科学数据出版发展迅速, 出版的数据成果越积越多^[2]。这些科学数据成果在被应用过程中是否也需要标注和著录? 如果需要, 那么如何规范科学数据出版成果的著录?

1.1 目前数据应用过程中的标引表达方式

目前数据应用过程中的标引主要有三种表达方式: 论(著)文中单设小节中说明数据来源; 致谢中说明数据来源; 作者分工中涉及数据来源信息。

收稿日期: 2018-05-10; 修订日期: 2018-06-01; 出版日期: 2018-06-25

*通讯作者: 刘闯 L-3684-2016, 中国科学院地理科学与资源研究所, lchuang@igsrr.ac.cn

作者 ID: 刘闯 L-3684-2016, 马军花 S-9830-2017, Uhler, P. F. G-6246-2016, 石瑞香 L-4389-2016

引用格式: 刘闯, 马军花, Uhler, P. F.等. 科学数据出版成果著录规范化研究[J]. 全球变化数据学报, 2018, 2(2): 123-134. DOI: 10.3974/geodp.2018.02.01.

1.1.1 论(著)文中单设小节文字说明数据来源

目前数据应用表达方式非常普遍的一种做法是在论(著)文中单设小节文字说明数据来源(示例1、示例2)。

示例1^[3]:

“数据来源:经济园区空间规模已16个城市5期建设用地TM遥感数据为依据,5期数据分别为1990、1995、2000、2005、2010,其来源于中国科学院地理科学与资源研究所。经济园区的社会经济数据来源于1989-2015年《开发区年鉴》以及各经济园区官网数据,城市社会经济数据来源于1985-2015年《中国城市统计年鉴》以及格式统计年鉴。”

示例2^[4]:

“本文的基础数据主要包括:①2015年雄安三县土地利用与覆被,空间分辨率为30m,源自中国科学院资源环境科学数据中心(<http://www.resdc.cn>),主要用于生境质量分析;②ASTER GDEM数据,空间分辨率为30m,源自地理空间数据云(<http://www.gscloud.cn/>),用于土壤保持服务中地形因子的测算;③2015年气象站点数据,包括气温和降水数据,来自中国气象数据共享网(<http://data.cma.cn/>),用于评估产水和土壤保持服务;④社会经济数据,来源于雄安三县统计年鉴,用于计算粮食产量;⑤Open Street Map路网数据(<https://www.openstreetmap.org/>)和2015年8月的VIIRS/DNB夜间灯光数据(<https://ncc.nesdis.noaa.gov/VIIRS/>),分别用于评估生境质量和修正基本阻力面。”

1.1.2 致谢中说明数据来源

数据应用另一种比较普遍的表达方式是在致谢中说明数据来源(示例3、示例4)。

示例3^[5]:

“Acknowledgements

We would like to thank Drs. FAN Qiang, LI Wei-le and HUANG Yan for providing various datasets and advices throughout this research.”

示例4^[6]:

“Acknowledgements

The authors would like to express appreciation to the Pakistan Meteorological Department (PMD) and the Capital Development Authority (CDA), Islamabad, for providing valuable data to conduct this research.”

1.1.3 作者分工中说明数据来源

有些数据应用者将应用的数据在作者分工中加以说明。

示例5^[7]

“Contributions

W. W., Y. H., D. L., and X. G. designed the study. W. W. wrote the paper. D. L. Y. M., Y. H., P. X., H. D., and X. G. reviewed the paper and contributed to the data analysis. Y. Y., Z. H., and Y. M. helped examine and improving the dataset. Y. Y. and Z. H. collected part of the first version of the 2014 sub-dataset. P. X. and H. D. contributed to the collection of the 1960s and 2005 sub-datasets. X. G. helped to provide the 2014 GF-1 data.”

从上述示例中可以看出,目前在科学数据应用标引方式非常不规范,不仅标注位置不规范,在著录内容上也各不同。

1.2 数据引文规范化发展态势

随着科学数据开放和共享程度的提高,认可科学家在科学数据领域对科学做出的贡献和保护数据知识产权,规范性引用科学数据的呼声在科技界越来越强烈。为了全面推动科学数据引文规范化,国际科学技术数据委员会(CODATA)与国际科学技术信息委员会(ICSTI)于2010年联合成立数据引文标准和实践任务组(CODATA-ICSTI Data Citation

Standards and Practices)。该任务组的成立和领衔组织召开的十几次国际学术讨论会的实践活动,在世界范围内全面开启了科学数据引文规范化运动。

2014年是国际科学数据出版年,其代表性的标志是《Scientific Data》(Springer Nature出版集团)、《Geoscience Data》(英国皇家气象学会)两个期刊和“全球变化变化科学研究数据出版系统”(中国科学院地理科学与资源研究所、中国地理学会)于2014年正式上线,这项不约而同的数据出版举措宣告了世界范围内的科学数据出版开始了。虽然三家数据出版单位采取的出版方式不同,但是,实体数据与数据论文关联出版的本质是一致的,即数据论文均采用学术期刊的方式出版,实体数据采取一体化或分布式方式全部开放,网络可以获取。科学数据出版成果,特别是实体数据出版成果,作为独立的学术成果在国际学术界已成共识。

由于科学数据纳入出版领域,科学数据规范化标引问题刻不容缓地提上出版日程。为了解决这个问题,2014年,美国FORCE 11即“研究交流与电子学者的未来”组织(the Future Of Research Communications and E-scholarship)提出数据引文原则(Data Citation Principles)。还有一些国家和国际组织,例如,英国数字解译中心(UK Digital Curation Center)、科学研究数据全球联盟(RDA)数据引文工作组等均积极推动数据引文规范化的工作。在FORCE 11数据引文原则中规定了8项基本原则,其核心精神是要求数据应用者将引用的数据纳入参考文献著录系统,如同信息与文献中参考文献著录规范化一样,其基本出发点是引文要达到不仅人可以理解,同时计算机可以互操作。

我国于2015年发布了中华人民共和国国家标准GB/T 7714—2015《信息与文献 参考文献著录规则》^[8],2017年发布了中华人民共和国国家标准GB/T 35294—2017《信息技术 科学数据引用》^[9]。然而,在科学数据出版和科学数据出版成果引文如何实施国际共识的数据引文原则,科学数据数字化出版成果引文规范如何实施?这个问题成为目前科学数据、科技信息和数字化出版交叉领域亟待解决的问题。

2 问题的深入:科学数据标引是否需要坚持开放性原则?

国际科学技术数据委员会(CODATA)、国际科学理事会(ISC)、世界数据系统(WDS)、科学研究数据全球联盟(RDA)等国际数据学术组织以及国际地球观测组织(GEO)、世界气象组织(WMO)等与数据密切相关的政府间组织均制定了数据开放的数据共享原则。联合国于2017年在南非举办的首次世界数据峰会论坛(WDF)通过的“开普敦宣言”和“行动计划”中明确指出:及时开放数据和保障数据质量是数据的核心问题^[10]。因此,在科学数据标引中提倡数据开放是国际数据界业已达成共识的科学数据开放与共享基本原则实施的重要方面。

由欧洲倡导“FAIR数据原则(Findable, Accessible, Inter-operational, Re-useful)”^[11]、世界经济合作组织(OECD)“数据共享原则”、由中国主导的国际科学数据委员会(CODATA)发展中国家任务组通过的“发展中国家科学数据共享原则”^[12]中也都分别强调了数据开放原则和数据的计算机互操作原则。很显然,数据不开放的情况下,数据间的互操作、智能化数据分析等都难以实现。从数据互操作的发展方向来分析,科学数据标引倡导和鼓励数据的开放性也非常重要。

科学数据的应用与学术论著的应用有所不同的是绝大多数学术论著的引用是学术观点、科学方法、科学发现或科学结论,而科学数据的应用很少是引用学术观点或科学发现结论,而是直接应用数据成果。多数情况下,这些科学数据可以成为不同学术观点或科学发现的共用数据。例如:IPCC 预测未来气候变暖模拟结果时,使用相同数据集,由于不同模型计算,得出不同的预测结果^[13]。在这个过程中,更重要的是科学数据不是作为学术观点的参考,而是被直接应用。它的质量如何直接影响了在此基础上产生新的数据集的质量。因此,数据质量可靠性检查或处理是数据出版和数据再应用的必要步骤。数据不开放的话,数据集的同行评议无法进行。当然,保护数据集原作者的知识产权,即数据记录的查重(“全球变化科学研究数据出版系统”的数据记录查重标准是10%原则^[14]),也无法进行。

在这种情况下,科学数据标引坚持开放性原则对实施国际共识的科学数据共享原则和促进科学数据共享具有重要意义。

那么是否一些内部使用的数据、一些未经过同行专家评审或未公开出版的数据就不需要标引了?我们的观点是需要标注和著录,但是,需要与公开出版的数据在标引和著录方式上有所区别。

3 解决问题思路:五种数据类型和两种标引方式的划分

科学数据与学术论著不同之处在于它更加复杂。首先对引用的数据进行分类,然后根据不同类型采取不同标引方式。

3.1 五种数据类型的划分

为了引文规范化,根据目前国内外数据开放的具体情况,特别是实体数据开放的情况,数据分为五种类型,即:正式出版的数据、世界数据系统(WDS)公开发布的数据、国家级数据中心发布的数据、个人或团队发布的数据和内部使用的数据。

3.1.1 正式出版的数据

正式出版的数据是由有出版资质的单位出版的数据。这类数据考核的关键点是数据正式出版,出版者一般包括有出版资质的出版社或学术期刊编辑部。出版方式包括数据集(实体数据)出版、数据论文出版、数据集包含在学术论文或学术专著中一并出版等不同的出版方式。

3.1.2 世界数据系统(WDS)公开发布并承担长期保藏和开放责任的数据

世界数据系统(WDS)是国际科学理事会(ICSU)于2008年批准成立的跨领域以实体数据永久保藏和提供可信赖的共享服务为宗旨的国际学术组织。世界数据系统(WDS)正式成员(Regular Member)资格需要国际科学理事会授权组织的评审专家组按照WDS正式成员批准程序进行资格评审、授权和定期检查。获得WDS正式成员的数据中心(仓储中心)开放的数据是被学术界认可的、可信赖的数据。根据世界数据系统(WDS)对正式成员的管理规定,对于数据标引来说,具有两个可信赖的核心点:其一,数据得到永久保藏和计算机网络开放共享服务,因此,在数据标引后,可以做到计算机网络可以查询,检验;其二,数据质量可以信赖。截至2018年3月,世界数据系统共有72个正式成员和7个正在审核中的正式成员(表1)。

表 1 国际科学理事会（ICSU）世界数据系统（WDS）批准的 WDS 正式成员

名称	国家	批准时间	主办机构
Incorporated Research Institutions for Seismology (IRIS), Data Services	USA	13-12-2011	National Science Foundation (NSF), USA
WDC-Geoinformatics and Sustainable Development	Ukraine	13-12-2011	National Technical University of Ukraine
ISRIC-WDC Soils	Netherlands	13-12-2011	ISRIC-World Soil Information
WDC for Climate	Germany	13-12-2011	German Climate Computing Center
WDC-Meteorology, Asheville	USA	13-12-2011	NOAA National Centers for Environmental Information (NCEI)
Center de Donnees astronomiques de Strasbourg (CDS)	France	13-12-2011	University of Strasbourg / CNRS
World Glacier Monitoring Service, Zurich	Switzerland	13-12-2011	University of Zurich
Australian Antarctic Data Center	Australia	13-12-2011	Australian Government, Antarctic Division
Chinese Astronomical Data Center	China	13-12-2011	Chinese National Astronomical Observatories, MOST/CAS
WDC-Renewable Resources and Environment	China	13-12-2011	Institute of Geographic Sciences and Natural Resources Research, CAS
Flanders Marine Institute, Data Center	Belgium	13-12-2011	Flanders Marine Institute
World Data Service for Oceanography	USA	13-12-2011	NOAA National Centers for Environmental Information (NCEI)
International Earth Rotation and Reference Systems	Germany	13-12-2011	Federal Agency for Cartography and Geodesy, Germany
Fish Database of Taiwan (Academia Sinica, Taiwan)	Taipei, China	13-12-2011	Academia Sinica Center for Digital Culture
WDC-Oceanography, Tianjin	China	13-12-2011	National Marine Data and Information Service
World Data Service for Geophysics	USA	13-12-2011	NOAA National Centers for Environmental Information (NCEI)
PANGAEA-Data Publisher for Earth & Environmental Science	Germany	13-12-2011	Center for Marine Environmental Sciences, University of Bremen; and Alfred Wegener Institute for Polar and Marine Research
WDC-Solar-Terrestrial Physics, Moscow	Russian Federation	13-12-2011	Geophysical Center of the Russian Academy of Sciences
WDC-Sunspot Index and Long-term Solar Observations (SILSO)	Belgium	24-02-2012	Solar Influences Data analysis Center-Royal Observatory of Belgium
WDC-Oceanography, Obninsk	Russian Federation	24-02-2012	All-Russian Research Institute of Hydrometeorological Information
WDC-Remote Sensing of the Atmosphere	Germany	24-02-2012	German Aerospace Center (DLR)
WDC-Geomagnetism, Copenhagen	Denmark	24-02-2012	National Space Institute
International Service of Geomagnetic Indices	France	24-02-2012	EOST, Université de Strasbourg/CNRS, France
WDC-Geomagnetism, Edinburgh	United Kingdom	24-02-2012	British Geological Survey
WDC for Solid Earth Physics, Moscow	Russian Federation	24-02-2012	Geophysical Center of the Russian Academy of Sciences,
WDC-Meteorology, Obninsk	Russian Federation	24-02-2012	All-Russian Research Institute of Hydrometeorological Information
WDC-Solar Activity / BASS2000	France	24-02-2012	LESIA, Observatoire de Paris
WDC-Geomagnetism, Kyoto	Japan	17-04-2012	Data Analysis Center for Geomagnetism and Space Magnetism, Kyoto University

续表

名称	国家	批准时间	主办机构
Interdisciplinary Earth Data Alliance	USA	26-06-2012	Lamont-Doherty Earth Observatory of Columbia University.
WDC-Space Weather, Australia	Australia	13-07-2012	The Bureau of Meteorology, Australian Government
NSIDC DAAC	USA	13-07-2012	University of Colorado
Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC)	USA	20-08-2012	Oak Ridge National Laboratory (ORNL)
World Stress Map Project	German	20-08-2012	German Research Center for Geosciences
WDC-CU	USA	04-10-2012	National Snow and Ice Data Center, University of Colorado
WDC-Ionosphere and Space Weather	Japan	22-11-2012	National Institute of Information and Communications Technology (NICT)
Ukrainian Geospatial Data Center	Ukraine	26-11-2012	Space Research Institute of Ukraine
Data Center for Geography, Moscow	Russian Federation	30-11-2012	Lomonosov Moscow State University, Faculty of Geography
WDC-Earth Resources Observation and Science (EROS) Center	USA	29-01-2013	USGS Earth Resources Observation and Science (EROS) Center
The Language Archive	Netherlands	29-01-2013	Max Planck Institute for Psycholinguistics
World Data Service for Paleoclimatology	USA	29-01-2013	NOAA National Centers for Environmental Information (NCEI)
DataFirst	South Africa	29-01-2013	University of Cape Town
WFCC-MIRCEN World Data Center for Microorganisms	China	30-01-2013	World Federation for Culture Collections (WFCC)
Goddard Earth Sciences Data and Information Services Center (GES DISC)	USA	23-03-2013	NASA/GSFC
Crustal Dynamics Data Information System (CDDIS)	USA	27-03-2013	NASA GSFC
Chinese Space Science Data Center	China	28-05-2013	Center for Space Science and Applied Research, CAS
Cold and Arid Regions Science Data Center at Lanzhou (CARD)	China	03-06-2013	Cold and Arid Regions Environmental and Engineering Research Institute, CAS
Global Hydrology Resource Center (GHRC)	USA	03-06-2013	NASA
Italian Center for Astronomical Archive-IA2	Italy	03-06-2013	INAF
Inter-university Consortium for Political and Social Research (ICPSR)	USA	03-06-2013	University of Michigan
Atmospheric Science Data Center (Distributed Active Archive Center)	USA	26-06-2013	NASA
WDC-Geomagnetism, Mumbai	India	28-03-2014	Indian Institute of Geomagnetism, New Panvel, India
Canadian Astronomy Data Center/ Canadian Virtual Observatory	Canada	15-04-2014	Canadian Astronomy Data Center/ Canadian Virtual Observatory
Alaska Satellite Facility	USA	19-05-2014	Geophysical Institute, Fairbanks
Ocean Networks Canada	Canada	14-07-2014	University of Victoria
Environment Climate Data Sweden	Sweden	13-08-2014	Swedish National Data Service
Socioeconomic Data and Applications Center (SEDAC)	USA	20-04-2015	Center for International Earth Science Information Network, Columbia University
UNAVCO, Inc.	USA	20-04-2015	UNAVCO, Inc.
Land Processes Distributed Active Archive Center	USA	26-06-2015	USGS

续表

名称	国家	批准时间	主办机构
Permanent Service for Mean Sea Level (PSMSL)	United Kingdom	02-10-2015	National Oceanography Center
DANS	Netherlands	06-11-2015	KNAW
Research Institute for Sustainable Humanosphere, Kyoto University	Japan	17-02-2016	Research Institute for Sustainable Humanosphere, Kyoto University
The Cambridge Crystallographic Data Center	United Kingdom	29-07-2016	The Cambridge Crystallographic Data Center
Global Biodiversity Information Facility	Denmark	29-08-2016	Global Biodiversity Information Facility
WDC for Geophysics, Beijing	China	09-09-2016	Institute of Geology and Geophysics, Chinese Academy of Sciences
Global Change Research Data Publishing and Repository	China	23-09-2016	Institute of Geographical Sciences and Natural Resources Research, CAS, Geographical Society of China
Canadian Cryospheric Information Network/Polar Data Catalogue	Canada	11-10-2016	Canadian Space Agency, Meteorological Service of Canada, Natural Resources Canada
Swedish National Data Service	Sweden	14-12-2016	Swedish Research Council
Center for Astronomical Data of the Institute of Astronomy of the Russian Academy of Sciences	Russian Federation	09-02-2017	Institute of Astronomy of the Russian Academy of Sciences
Ocean Biology Data Active Archive Center (OB.DAAC)	USA	26-05-2017	National Aeronautics and Space Administration
Neotoma Paleoecological Database	USA	01-09-2017	The Neotoma Community
ImmPort Repository	USA	07-02-2018	DAIT, NIAID, NIH
Worldwide Protein Data Bank (wwPDB)	USA	19-02-2018	Worldwide Protein Data Bank (wwPDB)

注：来自 WDS 网站：<https://www.icsu-wds.org/community/membership/regular-members>。

3.1.3 国家级科学数据中心发布的数据

国家级科学数据中心是指由国家政府主管部门认证批准的、承担科学数据永久保藏和提供开放共享服务任务的单位。一般情况下，国家级科学数据中心保藏的数据分为数据中心公开发布，全社会共享数据；数据中心公开发布元数据，一部分人（通常为资格审核或内部使用人员）共享实体数据；内部发布、内部共享数据三种情况。国家级科学数据中心发布的数据是指由国家级数据中心公开发布、通过计算机网络可以获取到的数据，包括实体数据。对于数据引文规范化来说，后两种情况，可纳入到内部使用的数据类型处理。由于国家主管部门对国家级数据中心发布的数据有规范的管理办法，包括数据质量、数据产权、数据安全等方面。所以，国家级科学数据中心发布、全社会共享的数据原则上也具有下述两个特点：其一，数据得到永久保藏和计算机网络开放共享服务；其二，数据质量可以信赖。

3.1.4 个人或团队通过计算机网络发布的数据

有一些科学家个人或科研团队直接将个人或科研团队研发的数据通过计算机网络发布，这些科学数据对数据传播和共享具有时效性强、专业性强等特点。但是，科学家个人或科研团队不能承诺对数据永久保藏和长期提供共享服务，不能提供同行专家评审通过的

证明。因此,这些数据很难保证多年后通过计算机网络可以被查询到。

3.1.5 内部使用的数据

由于各种原因或某种考虑,一些数据在一定时期内不能在计算机网络上公开向全社会开放共享,但是,这些数据可以在这个时期在一定范围内或对通过资格审核后的用户范围内共享。这部分数据的标引与公开出版的数据标引方式应该有所不同。

3.2 两种著录方式

根据数据是否开放、是否得到永久保藏和是否经过同行专家评审而数据质量可靠的基本原则,上述五种数据类型可采取两种不同的标引方式,其一是采取参考文献的著录方式,其二是采取脚注的标引方式。

3.2.1 参考文献方式著录

纳入参考文献著录方式的数据原则上是作者应用或主要参考的数据,数据评审专家可以此作为评审重要参考和比对的数据。它的基本特点是数据开放,数据得到永久保藏,数据通过同行专家评审。根据目前国内外数据管理的情况,如下其一者可采取参考文献著录方式:

(1) 有出版资质的单位公开出版的数据

经过国内外有关单位批准、具备数据出版资质的出版单位公开出版的数据。这类数据主要以三种形式出版,包括:

- 1) 有出版资质的出版社出版的数字化专著及实体数据;
- 2) 以专门出版实体数据为核心业务的电子期刊出版实体数据;
- 3) 以出版学术论文为核心业务,同时,将论文关联实体数据作为附件出版实体数据。

其中,学术期刊应该具有国际标准连续出版物号(ISSN号)^①或国内统一连续出版物号(CN号)。

(2) 世界数据系统(WDS)正式成员公开发布的数据

根据国际科学理事会(ISC)世界数据系统(WDS)目前发布的信息,世界数据系统正式成员列于表1。这些数据中心公开发布的数据的应用都可以纳入到参考文献目录中。

(3) 国家级数据中心公开发布的数据

目前,世界各国都在响应联合国大数据白皮书的号召,重视国家级数据中心的建设。很多国家建设了国家级开放数据、开放科学共享平台和共享网。例如:美国国家航空航天局(NASA)数据中心群、中国国家科技基础条件平台、南非国家开放科学共享网等就是典型的案例。有一些国家的数据中心保藏的数据分为两部分,其一是公开发布的数据,另一部分是内部使用、不公开发布的数据。纳入到可在参考文献著录的部分是公开发布的数据。

3.2.2 脚注方式著录

个人或团队通过计算机网络发布的数据和内部使用(不公开发布)的数据的著录采取脚注式方式。由于前者通常没有经过数据出版过程中必要的同行专家评审或缺乏数据长期保藏、网络共享的保障,因此,采用脚注的方式著录有助于数据用户在应用过程中增加对该数据应用前的数据质量检查和数据预处理程序和数据下载风险的认知。对于内部使用的数据采取脚注方式著录主要是告知用户这些数据的特殊使用政策和特殊的评审程序。

^① ISSN 中国国家中心,北京中关村南大街33号国家图书馆,100080.

4 列入参考文献的数据著录格式

列入参考文献的数据著录格式分为以下四种情况: 引用数据集标注位置与著录格式, 引用数据记录标注位置与著录格式, 引用数据论文标注位置与著录格式, 引用论著中作为附件数据文件的标注位置与著录格式。

4.1 引用数据集标注位置与著录格式

4.1.1 引用数据集标注位置

如果在文章或著作中引用已经正式出版的数据集(实体数据), 需要在文中或著作中论及到该数据集的位置用上标中括号标出著录顺序号。

示例 6^[15]:

“基于 ABPM 模型的全球海洋初级生产力遥感监测 9 km 分辨率月度数据集(2003–2012)^[8]的名称、作者、地理区域、数据年代、时间分辨率、空间分辨率、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表。”

4.1.2 引用数据集著录格式

应用他人公开出版或被世界数据系统(WDS)正式成员或国家级数据中心(数据共享平台)公开发布的数据集需要在参考文献栏目中通过中括号顺序号方式著录应用的数据集(实体数据), 著录格式包括析出数据集主要责任者(作者). 析出数据集题名[文献类型标识/文献载体标识]. 连续出版物题名: 其他题名信息, 年. 数字对象唯一标识符。

示例 7^[16]:

隋翠娟, 张占海, 吴辉碇等. 北极海冰季节统计数据(1979–2012) [DB/OL]. 全球变化科学研究数据出版系统, 2016. DOI: 10.3974/geodb.2016.03.05.V1.

4.2 引用数据记录标注位置和著录格式

“全球变化科学研究数据出版系统”数据共享政策中关于“摘取数据部分记录创作新数据的作者需要遵循 10% 引用原则”, 即从本数据集中摘取的数据记录少于新数据集总记录量的 10%, 同时需要在摘取的数据记录标注数据来源”的规定, 引用数据记录需要标注。通常需要采取二种著录兼有方式, 即不仅要在参考文献中按照引用数据集标注位置和著录格式标引, 而且还要在实体数据中以备注的方式标出引用数据记录的来源。

4.2.1 引用数据记录标注位置和备注格式

引用数据记录需要以备注的格式在实体数据的适当位置(如果实体数据是数据表, 通常备注放到表格下面)以标出数据引来源。

如果在数据集中引用其他数据成果的数据记录, 需要在该数据记录的位置上标星号*标出备注的位置, 并在表格末尾添加备注内容。

示例 8^[17]:

在《中国地形起伏度公里网格数据集》(DOI: 10.3974/geodb.2018.03.16.V1)中, 有一列数据记录(“中国分省地形起伏度统计表”中)是中国分省行政区划编码, 这时就需要在该栏表头右上方加*标注(见图 1)。在表格末尾添加备注。

表1 中国分省地形起伏度统计表			
Table 1 Statistics of the Relief Degree of Land Surface (Rdls) in provinces of China			
Province Name	省名	CODE*	Rdls
		行政编码*	地形起伏度
Beijing	北京市	110000	0.6553
Tianjin	天津市	120000	0.0247
Hebei Province	河北省	130000	0.6519
.....
Hong Kong Special Administrative Region	香港特别行政区	810000	0.8006
Macao Special Administrative Region	澳门特别行政区	820000	0.1532
*备注：行政编码引自文献：中华人民共和国民政部. 中华人民共和国行政区划简册2014. 中国地图出版社. 2014.			
*Note: The code was cited from the reference: Ministry of Civil Affairs of the People's Republic of China. People's Republic of China administrative division 2014. SinoMaps Press. 2014.			

图 1 数据记录中中国分省行政区划编码引用示例

4.3 引用数据论文（学术期刊连续出版物中的析出文献）标注位置和著录格式

4.3.1 引用数据论文（学术期刊连续出版物中的析出文献）标注位置

如果在文章或著作中引用已经正式出版的数据论文(学术期刊连续出版物中的析出文献), 需要在文中或著作中论及到该数据论文的位置用上标中括号 ([]) 标出著录顺序号 (示例 9)。

示例9^[18]：引用数据论文标注位置

“妥善应对全球变化，离不开科学研究的支撑，全球变化因而成为地球科学最活跃的研究领域之一。国际全球变化研究始于 20 世纪 80 年代，由多个国际科学研究计划引领，当前主要以 2012 年国际科学理事会(ICSU)、国际社会科学理事会(ISSC)等多家国际组织联合发起的“未来地球计划(Future Earth)”(2014–2023)^[2]主导”。

4.3.2 引用数据论文（学术期刊连续出版物中的析出文献）著录格式

引用数据论文著录格式包括析出文献主要责任者(作者). 析出文献题名 [文献类型标识/文献载体标识] . 连续出版物题名: 其他题名信息, 年, 卷(期): (页码). 数字对象唯一标识符.

示例 10^[19]:

林丽娜, 王颖杰, 张洁等. 中国南极科学考察 CTD 数据集[J]. 全球变化数据学报, 2017, 2(3): 157–164. DOI: 10.3974/geodp.2017.02.04

4.4 引用论著中作为支持科学结论的数据文件的标注位置与著录格式

越来越多的学术期刊要求作者在发表论文章的同时要把关联的数据一并发表，这些数据是作为支持信息发表。这一部分数据引用的标注位置与该论文的标注位置相同，所不同的是著录格式有所区别。引用数据论文著录格式包括析出文献主要责任者（作者）. 析出文献题名 [文献类型标识/文献载体标识]. 连续出版物题名: 其他题名信息, 年, 卷(期): (页码). 数字对象唯一标识符. 数据所在文中的部位, 数据格式: 数据文件名. 文件扩展名(数据量), 数据获取和访问路径 (示例 11)。

示例 11^[20]:

Zhang, Q. Q., Ying, G. G., Pan, C. G., *et al.* Comprehensive evaluation of antibiotics emission and fate in the river basins of China: source analysis, multimedia modeling, and linkage to bacterial resistance [J]. *Environ. Sci. Technol.*, 2015, 49 (11): 6772–6782. DOI: 10.1021/acs.est.5b00729. Supporting Information, MS Excel: es5b00729_si_001.xlsx (2.43 MB), <https://pubs.acs.org/doi/10.1021/acs.est.5b00729>.

5 采用脚注方式的数据标引格式

个人或团队通过计算机网络发布的数据和内部使用（不公开发布）的数据的标引采取

脚注方式著录。脚注格式的序号采取由计算机自动生成圈码①上标的方式。在 PDF 数据格式的文本中，脚注出现在本页文底，以分割线将其与正文隔开（示例 12，图 2）。

示例 12（PDF）格式^[21]：

表 3 “三生用地”的提取方案		
用地类型	用地定义	提取依据
水源涵养用地	重要的江水源和水源补给的分布区域	生态功能区划 ^① ，中国自然保护区区划 ^[17] ，土地利用图
土壤保持用地	主要的水土流失和治理区域	生态功能区划 ^① ，全国水土保持规划 ^[18] ，土地利用图
.....
农村生活用地	包括土地利用中农村居民点的分布范围	土地利用图
注：引自张红旗等 ^[16] ，资源科学，2015。		
<div>① 中华人民共和国环境保护部，中国科学院. 全国生态功能区划. 2008.</div> <div>② 中华人民共和国国家林业局. 全国防沙治沙规划（2011-2020 年）. 2013.</div> <div>③ 中国 1：400 万水系矢量图（中国科学院资源环境科学数据中心）.</div> <div>④ 中华人民共和国国家林业局，中华人民共和国财政部. 国家级公益林区划界定办法. 2009.</div> <div>⑤ WCMC 自然保护区（http://protectedplanet.net/）.</div>		

图 2 正文中的脚注标注

示例 13（HTML）格式^[21]：

在 HTML 格式中，全部脚注按照顺序号集中汇集到全文的最后部分，以分割线将其与正文隔开（示例 13，图 3）。

<div>① 中华人民共和国环境保护部，中国科学院. 全国生态功能区划. 2008.</div> <div>② 中华人民共和国国家林业局. 全国防沙治沙规划（2011-2020 年）. 2013.</div> <div>③ 中国 1：400 万水系矢量图（中国科学院资源环境科学数据中心）.</div> <div>④ 中华人民共和国国家林业局，中华人民共和国财政部. 国家级公益林区划界定办法. 2009.</div> <div>⑤ WCMC 自然保护区（http://protectedplanet.net/）.</div>

图 3 HTML 格式尾标注

6 结束语

随着大数据时代的到来，科学数据出版成果越来越多。在科学数据出版和公开发表的方式和渠道多样化的情况下，广泛应用、规范引用和标准化著录这些数据成果对推动科学数据开放共享、提高数据质量、尊重数据原创者知识产权，避免“用而不引”、“引而不标”、“标不规范”现象，更好地促进科研大数据的快速发展均具有重要的实践意义。

参考文献

[1] 中华人民共和国国务院办公厅. 关于印发科学数据管理办法的通知[OL]. 2018. http://www.gov.cn/gongbao/content/2018/content_5283177.htm.

[2] 刘闯. 论全球变化科学研究数据出版[J]. 地理学报, 2014, 69(增刊): 03-11. DOI: 11821/DLXB2014s000.

[3] 孙占秀, 栾维新, 马瑜等. 中国沿海不同区位经济园区空间扩张特征研究[J]. 自然资源学报, 2018,

- 33(2): 262–274. DOI: 10.11849/zrzyxb.20170052.
- [4] 彭建, 李慧雷, 刘焱序等. 雄安新区生态安全格局识别与优化策略[J]. 地理学报, 2018, 73(4): 701–710. DOI: 10.11821/dlxb201804009.
- [5] Meng, Q. K., Miao, F., Zhen, J., *et al.* Impact of earthquake-induced landslide on the habitat suitability of giant panda in Wolong, China [J]. *Journal of Mountain Science*, 2016, 13(10): 1789–1805. DOI: 10.1007/s11629-015-3734-0.
- [6] Abbas, T., Nabi, G., Boota, W. M., *et al.* Uncertainty analysis of runoff and sedimentation in a forested watershed using sequential uncertainty fitting method [J]. *Science in Cold and Arid Regions*, 2016, 8(4): 297–310. DOI: 10.3724/SP.J.1226.2016.00297.
- [7] Wan, W., Long, D., Hong, Y., *et al.* A lake data set for the Tibetan Plateau from the 1960s, 2005, and 2014 [J]. *Scientific Data*, 2016, 3: 160039. DOI: 10.1038/sdata.2016.39.
- [8] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. 信息与文献 参考文献著录规则: GB/T 7714—2015 [S]. 北京: 中国标准出版社, 2015: 5.
- [9] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. 信息技术 科学数据引用: GB/T 35294—2017 [S]. 北京: 中国标准出版社, 2017: 12.
- [10] High-level Group for Partnership, Coordination and Capacity-Building for statistics for the 2030 Agenda for Sustainable Development (HLG-PCCB). Cape Town Global Action Plan for Sustainable Development Data [OL]. 2017. <https://undataforum.org/WorldDataForum/wp-content/uploads/2017/01/Cape-Town-Action-Plan-For-Data-Jan2017.pdf>.
- [11] FORCE 11. The FAIR Data Principles [OL]. <https://www.force11.org/group/fairgroup/fairprinciples>.
- [12] 国际科学数据委员会发展中国家数据保藏与任务组. 发展中国家科学数据共享原则(内罗毕数据共享原则) [J]. 全球变化数据学报, 2017, 1(1): 12–15. DOI: 10.3974/geodp.2017.01.03.
- [13] IPCC. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 2014: 151.
- [14] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策 [OL]. DOI: 10.3974/dp.policy.2014.05 (2017 年更新).
- [15] 陶醉, 周翔, 马胜. 基于 ABPM 模型的全球海洋初级生产力遥感监测 9 km 分辨率月度数据集 (2003–2012)[J]. 全球变化数据学报, 2017, 1(2): 149–156. DOI: 10.3974/geodp.2017.02.03.
- [16] 隋翠娟, 张占海, 吴辉碇等. 北极海冰季节统计数据集(1979–2012) [DB/OL]. 全球变化科学研究数据出版系统, 2016. DOI: 10.3974/geodb.2016.03.05.V1.
- [17] 游珍, 封志明, 杨艳昭. 中国地形起伏度公里网格数据集[DB/OL]. 全球变化科学研究数据出版系统, 2018. DOI: 10.3974/geodb.2018.03.16.V1.
- [18] 徐冠华, 葛全胜, 方修琦等. 中国全球变化国家基础研究计划(2011–2020) [J]. 全球变化数据学报, 2017, 1(1): 16–26. DOI: 10.3974/geodp.2017.01.04.
- [19] 林丽娜, 王颖杰, 张洁等. 中国南极科学考察 CTD 数据集[J]. 全球变化数据学报, 2017, 2(3): 157–164. DOI: 10.3974/geodp.2017.02.04.
- [20] Zhang, Q. Q., Ying, G. G., Pan, C. G., *et al.* Comprehensive evaluation of antibiotics emission and fate in the river basins of China: Source analysis, multimedia modeling, and linkage to bacterial resistance [J]. *Environmental Science & Technology*, 2015, 49(11): 6772–6782. DOI: 10.1021/acs.est.5b00729.
- [21] 张红旗, 许尔琪, 朱会义. 中国生态、居住与工业土地利用空间数据集[J]. 全球变化数据学报, 2017, 1(4): 419–425. DOI: 10.3974/geodp.2017.04.06.