

•编者按•

生物多样性数据论文：规范化和多样化

张健^{1*}，黄晓磊^{2*}，周玉荣³，马克平⁴

1. 华东师范大学生态与环境科学学院，浙江天童森林生态系统国家野外科学观测研究站，上海 200241；2. 闽台作物有害生物生态防控国家重点实验室，福建农林大学植物保护学院，福州 350002；3. 中国科学院植物研究所《生物多样性》编辑部，北京 100093；4. 中国科学院植物研究所植被与环境变化国家重点实验室，北京 100093

张健，黄晓磊，周玉荣，马克平 (2021) 生物多样性数据论文：规范化和多样化. 生物多样性, 29, 1147–1148. doi: 10.17520/biods.2021369.
Zhang J, Huang XL, Zhou YR, Ma KP (2021) Biodiversity data papers: Standardizing and diversifying data sharing. Biodiversity Science, 29, 1147–1148. doi: 10.17520/biods.2021369.

Biodiversity data papers: Standardizing and diversifying data sharing

Jian Zhang^{1*}，Xiaolei Huang^{2*}，Yurong Zhou³，Keping Ma⁴

1 Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, School of Ecological and Environmental Sciences, East China Normal University, Shanghai 200241

2 State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops, College of Plant Protection, Fujian Agriculture and Forestry University, Fuzhou 350002

3 The Editorial Office of Biodiversity Science, Institute of Botany, Chinese Academy of Sciences, Beijing 100093

4 State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093

2021年10月，联合国《生物多样性公约》第15次缔约方大会(COP15)将在我国昆明举办。会议将审议通过2020后全球生物多样性框架，确定2030年全球生物多样性保护目标。与之密切相关的联合国可持续发展目标(Sustainable Development Goals, SDGs)、生物多样性和生态系统服务政府间科学-政策平台(IPBES)和政府间气候变化专门委员会(IPCC)评估报告等，也对当前和未来的生物多样性保护设定了短期和长期目标。这些战略目标的实现，无一不需要海量的、高质量的、可发现、可访问、可互操作和可再用的生物多样性大数据作为支撑。

数据论文是数据共享的主要载体之一。它以描述科学数据为核心目的，而不是报道研究结果或研究假说。数据论文基于结构简单的、易于理解的形式分享数据，从而让更多的研究人员和公众了解、使用并正确引用数据(Chavan & Penev, 2011)。科研活动产生的数据在开放共享过程中应遵循可发现(Findability)、可访问(Accessibility)、可互操作(Interoperability)和可再用(Reusability)的FAIR原则

(Jacobsen et al, 2019; 宋佳等, 2021)，数据论文发表也应该遵守这个基本原则。生物多样性数据共享和发表不仅是提升生物多样性研究深度和广度以及开展整合性研究的需要，也是科学制定生物多样性保护政策的重要基础(黄晓磊和乔格侠, 2014)。近年来，在生物多样性研究领域，数据论文和数据期刊也逐渐流行起来。据不完全统计，目前可发表生物多样性数据论文的期刊不少于33种，详见GBIF网站列表(<https://www.gbif.org/data-papers>)。

《生物多样性》期刊一直在积极推动中国生物多样性数据共享工作，2011年设立了旨在共享物种信息的“生物编目”栏目，2014年设立了“数据论文”栏目，并通过组织生物多样性信息学、数据共享专辑/专题等形式(马克平, 2014; 严岳鸿等, 2016)发表了多篇关于生物多样性数据共享的讨论文章和数据论文(如: 黄晓磊和乔格侠, 2014; 李俊洁和黄晓磊, 2016; 张健, 2017)。基于前期的探索和经验，我们逐渐意识到当前我国生物多样性数据共享领域存在的一些问题，比如有些作者未能很好地遵循

* 共同通讯作者 Co-authors for correspondence. E-mail: jzhang@des.ecnu.edu.cn; huangxl@fafu.edu.cn

FAIR原则, 共享数据的多样性程度还不够高。因此, 在2020年编委会讨论的基础上, 在多位编委的共同努力下, 我们组织了本期数据论文专题, 希望能为将来的生物多样性数据共享和作者投稿提供参考。


本专题共收录7篇数据论文, 涉及生物功能性状数据、物种多样性和地理分布数据、生态遥感数据、生物多样性监测数据等不同类型。生物功能性状数据对于理解物种演化和生态系统功能等具有重要价值, 王彦平等(2021)报道了包含1,445种中国鸟类的生活史和生态学特征数据集, 佟一杰等(2021)报道了包含1,302种锹甲科成虫(该科全球99%的已知种)的前胸背板和鞘翅外轮廓形态这些典型连续性形态性状的数据集。生物多样性格局和生物区系演化等方面的研究离不开物种多样性和地理分布数据的积累, 李俊洁等(2021)共享了102科2,090属7,822种中国半翅目昆虫的地理分布信息, 程洁婕等(2021)共享了720种(包括114变种、45亚种和2变型)中国野生杜鹃花属(*Rhododendron*)植物的分类信息和地理分布数据。近年来, 无人机、红外相机等新技术越来越成为生物多样性监测的重要手段。张昭臣等(2021)报道了4个中国亚热带森林动态监测样地的无人机可见光遥感影像数据, 可以为森林生物多样性和林冠生态学等研究提供重要支撑。李晟等(2021)报道了四川老河沟保护地2011–2015年野生动物红外相机监测数据集, 李盼


盼等(2021)共享了基于3种采集方法在广西花坪样地不同地点所收集的563个形态种8,914头甲虫标本照片数据集。这些数据论文均采用了相同的格式, 相应的数据集也均保存至公共数据库并提供了下载链接, 实现了免费公开共享。


数据格式的标准化和数据论文的规范化是实现数据共享FAIR原则的重要基础, 共享数据类型的多样化则有助于从整体上促进生物多样性科学的发展。以此专题中的论文为范例, 《生物多样性》期刊对今后的数据论文投稿在稿件格式、数据量、数据存储等方面提出了更严格的规范化要求。为鼓励更多类型生物多样性数据的快速共享, 编辑部也将通过提高数据论文稿件的审理与出版速度、减免版面费等方式予以支持。具体可参考期刊网站的投稿指南和模板。我们期望越来越多的研究者能够认识到数据共享对于生物多样性科学研究和保护的重要意义, 通过标准化和多样化的数据共享来促进生物多样性研究水平的提高和保护成效的提升。

文中引用的参考文献见附录1 (<http://www.biodiversity-science.net/fileup/PDF/2021-369.pdf>)。

ORCID

张健  <https://orcid.org/0000-0003-0589-6267>

黄晓磊  <https://orcid.org/0000-0002-6839-9922>

马克平  <https://orcid.org/0000-0001-9112-5340>

(责任编辑: 周玉荣)

参考文献

- Chavan V, Penev L (2011) The data paper: A mechanism to incentivize data publishing in biodiversity science. *BMC Bioinformatics*, 12(Suppl. 15), S2.
- Cheng JJ, Li MJ, Yuan TH, Huang H, Yang GL, Bai XX (2021) A dataset on wild *Rhododendron* and geographical distribution information in China. *Biodiversity Science*, 29, 1175–1180. (in Chinese with English abstract) [程洁婕, 李美君, 袁桃花, 黄红, 杨桂丽, 白新祥 (2021) 中国野生杜鹃花属植物名录与地理分布数据集. 生物多样性, 29, 1175–1180.]
- Huang XL, Qiao GX (2014) Sharing and publishing of biodiversity data: Recent trends and future suggestions. *Biodiversity Science*, 22, 293–301. (in Chinese with English abstract) [黄晓磊, 乔格侠 (2014) 生物多样性数据共享和发表: 进展和建议. 生物多样性, 22, 293–301.]
- Jacobsen A, de Miranda Azevedo R, Juty N, Batista D, Coles S, Cornet R, Courtot M, Crosas M, Dumontier M, Evelo CT, Goble C, Guizzardi G, Hansen KK, Hasnain A, Hettne K, Heringa J, Hooft RWW, Imming M, Jeffery KG, Kaliyaperumal R, Kersloot MG, Kirkpatrick CR, Kuhn T, LaBastida I, Magagna B, McQuilton P, Meyers N, Montesanti A, van Reisen M, Rocca-Serra P, Pergi R, Sansone SA, da Silva Santos LOB, Schneider J, Strawn G, Thompson M, Waagmeester A, Weigel T, Wilkinson MD, Willighagen EL, Wittenburg P, Roos M, Mons B, Schultes E (2020) FAIR principles: Interpretations and implementation considerations. *Data Intelligence*, 2, 10–29.
- Li JJ, Liu HH, Wu YX, Zeng LD, Huang XL (2021) A dataset on the diversity and geographical distributions of hemipteran insects in China. *Biodiversity Science*, 29, 1154–1158. (in Chinese with English abstract) [李俊洁, 刘欢欢, 吴杨雪, 曾凌达, 黄晓磊 (2021) 中国半翅目昆虫多样性和地理分布数据集. 生物多样性, 29, 1154–1158.]
- Li JJ, Huang XL (2016) Analysis of publication trends of biodiversity data papers. *Biodiversity Science*, 24, 1317–1324. (in Chinese with English abstract) [李俊洁, 黄晓磊 (2016) 生物多样性数据论文发表趋势分析. 生物多样性, 24, 1317–1324.]
- Li PP, Tong YJ, Cao HY, Rong GS, Qin SQ, Yang XK, Wang GQ, Bai M (2021) A photographic dataset of the beetle specimens from a SITE100 standardized sampling area in Huaping, Guangxi. *Biodiversity Science*, 29, 1165–1169. (in Chinese with English abstract) [李盼盼, 佟一杰, 曹浩宇, 容国森, 覃诗晴, 杨星科, 王国全, 白明 (2021) 广西花坪SITE100样地甲虫标本照片数据集. 生物多样性, 29, 1165–1169.]
- Li S, Wang DJ, Chen XH, Bu HL, Liu XG, Jin T (2021) The wildlife camera-trapping dataset of Laohegou Protected Area, Sichuan Province (2011–2015). *Biodiversity Science*, 29, 1170–1174. (in Chinese with English abstract) [李晟, 王大军, 陈祥辉, 卜红亮, 刘小庚, 靳彤 (2021) 四川老河沟保护地2011–2015年野生动物红外相机监测数据集. 生物多样性, 29, 1170–1174.]
- Ma KP (2014) Rapid development of biodiversity informatics in China. *Biodiversity Science*, 22, 251–252. (in Chinese) [马克平 (2014) 生物多样性信息学在中国快速发展. 生物多样性, 22, 251–252.]
- Song J, Wen LM, Li Y (2021) Scientific data sharing FAIR data principles: Background, content and practice. *Information and Documentation Services*, 42(1), 57–68. (in Chinese with English abstract) [宋佳, 温亮明, 李洋 (2021) 科学数据共享FAIR原则: 背景、内容及实践. 情报资料工作, 42(1), 57–68.]
- Tong YJ, Zhang MN, Josh JS, Wan X, Yang XK, Bai M (2021) A geometric morphometric dataset of stag beetles. *Biodiversity Science*, 29, 1159–1164. (in Chinese with English abstract) [佟一杰, 张萌娜, Josh Jenkins Shaw, 万霞, 杨星科, 白明 (2021) 全球锹甲的几何形态学数据集. 生物多样性, 29, 1159–1164.]
- Wang YP, Song YF, Zhong YX, Chen CW, Zhao YH, Zeng D, Wu YR, Ding P (2021) A dataset on the life-history and ecological traits of Chinese birds. *Biodiversity Science*, 29, 1149–1153. (in Chinese with English abstract) [王彦平, 宋云枫, 钟雨茜, 陈传武, 赵郁豪, 曾頔, 吴亦如, 丁平 (2021) 中国鸟类的生活史和生态学特征数据集. 生物多样性, 29, 1149–1153.]
- Yan YH, Huang XL, Ma KP (2016) Sharing biodiversity data through academic publishing. *Biodiversity Science*, 24, 1315–1316. (in Chinese) [严岳鸿, 黄晓磊, 马克平 (2016) 通过发表实现生物多样性数据共享. 生物多样性, 24, 1315–1316.]
- Zhang J (2017) Biodiversity science and macroecology in the era of big data. *Biodiversity Science*, 25, 355–363. (in Chinese with English abstract) [张健 (2017) 大数据时代的生物多样性科学与宏生态学. 生物多样性, 25, 355–363.]
- Zhang ZC, Hu JB, Yang QS, Lian JY, Li BH, Wang XH, Ye WH, Zhang J (2021) An unmanned aerial vehicle visible light remote sensing image dataset of four forest dynamic plots in subtropical China. *Biodiversity Science*, 29, 1181–1185. (in Chinese with English abstract) [张昭臣, 胡健波, 杨庆松, 练据愉, 李步杭, 王希华, 叶万辉, 张健 (2021) 中国亚热带4个森林动态监测样地无人机可见光遥感影像数据集. 生物多样性, 29, 1181–1185.]