

论文关联数据相关工作说明

1 论文关联数据共享的意义

论文关联数据是指通过基础研究、应用研究、实验开放等产生的用于支撑学术论文发表的数据，以及通过观测检测、考察调查、检验检测等方法取得并用于形成论文图表、支撑论文研究结论的原始数据及其衍生数据。论文关联数据的共享，是论文研究成果和结论的重要佐证，提升研究可验证性和透明度；数据的共享，提升科技期刊论文的受关注度和影响力；数据的重复使用，节省人力物力，让科学家聚焦于创新研究；论文数据作为一种新型学术成果，将促进学科领域间的创新合作和人才培养，推进开放式科学实践。

2 论文关联数据的类型及要求

作者用于直接支撑论文结论的数据应当共享，本刊鼓励作者共享；作者为开展论文课题研究而产生的且反映在论文中的数据，或作者为开展论文课题研究而进行重复使用或分析的数据鼓励共享；作者为开展论文课题研究从试验或观察中得到的原始的、未加工的且未反映在论文中的数据自愿共享。涉及科研伦理、敏感信息、保密信息或共享数据将损害第三方合法权益等情况的数据不宜共享。对于不宜共享的数据，若已进行不可逆的脱敏处理或因其他合法合理的原因可以共享的，数据作者提供脱敏声明或其他证明文件后可以共享。

本刊鼓励论文作者提交论文关联数据进行公共保存，保障论文结论可验证、可重用，提高论文可信度，维护论文作者学术声誉。提交论文关联数据的论文将获得优先评审、优先录用、优先出版的资格，后续本刊若有优秀论文评选活动将优先考虑已提交论文关联数据的论文。

3 数据标准和格式

论文关联数据的共享不限制作者提交的数据格式。但从数据可重用、可长期访问等方面考虑，请作者优先参考推荐格式列表。

若所选数据文件格式尚不在下表中，本刊推荐作者从以下3个方面自行评估数据文件的可重用和长期可持续访问情况：

- (1) 采用的文件格式是领域内的常用格式；
- (2) 采用的文件格式有开放的规范或标准；
- (3) 采用的文件格式独立于特定的软件、开发人员或供应商。

若无法完全符合上述要求，也请尽可能满足以上三点原则。

论文关联数据推荐格式列表

文件类型	推荐格式
纯文本文件	Unicode text(.txt)
标签语言文件	XML(.xml) HTML(.html) Related files: .css, .xslt, .js, .es
文本文档	PDF/A(.pdf)
RDF	RDF/XML (.rdf) Trig (.trig) Turtle (.ttl) NTriple (.nt) JSON-LD
电子数据表	CSV(.csv) ODS(.ods)
数据库文件	SQL(.sql) SIARD (.siard) DB tables (.csv)
统计数据	SPSS Portable (.por) STATA (.dta) DDI (.xml) Data and setup (.csv +.txt) R
光栅图片	JPEG (.jpg, .jpeg) TIFF (.tif, .tiff) PNG (.png) JPEG 2000 (.jp2) DICOM (.dcm)
矢量图片	SVG (.svg)
音频文件	BWF (.bwf) MXF (.mxf) Matroska (.mka) FLAC (.flac)

4 数据使用许可协议

作者可选择的数据使用许可协议有以下8种国际通用许可协议：CC0、CC-BY 4.0、CC BY-SA 4.0、CC BY-NC 4.0、CC BY-NC-SA 4.0、CC BY-ND 4.0、CC BY-NC-ND 4.0（以上协议的开放程度依次递减）和ODbL（用于数据库类型的数据）。数据使用者须根据作者选择的数据使用许可协议对数据进行使用。

为了更好地传播和共享论文关联数据，本刊推荐作者使用CC-BY 4.0许可协议。

5 数据共享的方式

5.1 数据共享

数据共享的方式有两种：直接共享和有条件共享，作者可根据实际需要二选一。

1) 直接共享

作者提交的论文关联数据，一旦通过评审，将即刻发布，即元数据和数据文件开放给公众访问获取。

2) 有条件共享

数据保护期后获取：作者提交的论文关联数据可设置保护期。保护期内，公众仅能访问数据的元数据，而无法下载获取数据文件。保护期后，数据自动转为开放获取状态，公众皆可访问获取其元数据和数据文件。

依申请获取：作者提交的论文关联数据因特殊原因不宜开放共享，数据使用者须向作者提出数据访问申请，得到作者授权或同意后，才能获取数据文件。对于不宜开放共享的数据，作者应当向本刊提供相关情况的书面说明并提供相关证明材料。

5.2 数据存储库

本刊鼓励论文关联数据在数据存储库中提交并共享。本刊对数据存储库的遴选标准和推荐列表，如下：

5.2.1 遴选标准

数据存储库应遵循FAIR共享原则，保障科学数据的可发现性、可访问性、互操作性和可重用性。具体包括以下内容和要求：

- 1) 数据存储库应为科学数据提供可发现的数据标识；
- 2) 数据存储到存储库，应提供可供编辑部、审稿人审核使用的链接和公开共享的永久访问地址；
- 3) 数据元数据和文件应支持互操作，提供开放的元数据收割接口；
- 4) 数据存储库应支持数据的重用，提供通用的标准元数据和传播分发路径，为数据重用提供支持保障。

5.2.2 推荐列表

科学数据银行 Science Data Bank (ScienceDB)

6 数据审核

本数据存储平台对作者提交的论文关联数据进行形式审核，主要审核其是否已上传数据、上传的数据是否可访问、是否已提交数据可用性声明（如要求提供）等内容。

作者应在录用阶段完成论文关联数据的提交。

7 数据可用性声明

《数据可用性声明》是对论文关联数据是否可获取以及具体获取方式的说明文件或文字，其主要包括论文关联数据的存储方式、访问链接等内容。《数据可用性声明》的使用有助于提高研究透明度和可重复性，对论文结论的重复验证、数据重用、科研诚信等均具有重要意义。

《数据可用性声明》应当包括以下内容：数据的存储方式及访问链接、数据的唯一标识符（如有）、打开或使用数据的软件或工具名称。若论文关联数据为不宜共享的数据，应当明确说明并给出访问数据的具体条件及方式。作者应将《数据可用性声明》附于论文正文之后、参考文献之前。

《数据可用性声明》模板如下，供作者参考：

1. 一般模板：本篇论文的关联数据（DOI/CSTR）可在（数据库名称）数据库（永久网页链接）中访问获取。
2. 需特殊软件工具打开的数据的模板：本篇论文的关联数据（DOI/CSTR）可在（数据库名称）数据库（永久网页链接）中访问获取，该论文关联数据的打开软件为（软件名称）。
3. 不宜共享的数据的模板：本篇论文的关联数据为不宜共享的数据，可依据合理理由从作者处获取，作者联系方式：

8 数据的引用

8.1 数据引用的意义

论文关联数据作为作者的智慧劳动成果应当得到尊重和认可，规范的数据引用能够提升科学数据的影响力，同时为引用该数据的研究提供佐证。

8.2 数据引用的格式

本刊推荐的数据引用格式参照《信息技术科学数据引用》(GB/T 35294—2017)。

9 帮助和支持

9.1 科学数据银行Science Data Bank (ScienceDB) 数据提交流程

- 1) 平台注册与登录: 作者可选择数据存储库平台账户注册登录, 也可选择存储库平台支持的第三方用户认证登录。
- 2) 数据提交与修改: 作者通过数据提交入口, 创建一个新的数据, 此时数据即被分配唯一标识(CSTR和DOI)和私有访问链接。作者在此流程中填写元数据信息、上传数据文件、选择开放共享方式等信息, 此时作者可以暂存草稿并进行信息更改, 直至提交。数据提交后, 若数据审核方反馈了修改意见, 用户则需要对数据进行修改, 直至数据通过审核。数据发布前, 数据仅可通过私有访问链接进行编辑部审稿等内部传播。
- 3) 数据发布与传播: 数据审核通过并发布后, 数据唯一标识生效, 数据可在平台、传播索引平台被检索到。
- 4) 数据更新: 数据一旦发表, 如需追加文件或进行数据更新操作, 作者需重新发起数据新版本的创建, 并对更新信息进行提交。与旧版本一样, 用户更新信息需要审核后才能发布。

9.2 联系ScienceDB

联系人: 中国科学院计算机网络信息中心 姜璐璐 李宗闻

电 话: 010-58812538

邮 箱: sciencedb@cnic.cn

地 址: 北京市海淀区东升南路二号院中科院信息化大厦