

新文科计算社会科学平台使用说明

四、数据资源使用手册

登录地址为：<http://dataverse.openaiplan.com:8080/>

1. 数据资源的核心概念

研究数据（Research data）与数据文件（Data file）

研究数据指科研人员通过观察、探测、实验、建模、仿真和调查等形式收集的、用于验证研究发现的、能够直接或间接产出学术成果的完整事实材料及其数字化产品，例如：文档、数据文件、问卷、模型、算法、软件或代码、图片、音视频资料等。开放研究数据平台只保存数字化的研究数据产品，主要形态为包含数据信息的电子文件，即数据文件。数据文件包括通常意义上含有一条或多条数据记录的电子表格或数据库文件，也包括相关的源代码、流程图、说明文档等文件。

数据集（Dataset）

数据集是开放研究数据平台中的基本数据单位之一，必须从属于一个数据空间。数据集可以包含一个或多个数据文件，可由一人或多人管理，并面向可控的用户群（不超过所属数据空间的用户群）开放研究数据的使用权。一个数据集通常对应一个具体的研究项目，或大型研究项目的一个阶段。

在开放研究数据平台的浏览与搜索结果中，数据集带有蓝色边框，图标为一张写满字的纸片。

2. 使用数据资源

了解平台的基本构成和用户界面，联系注册平台账号并完善用户信息，为借助开放研究数据平台提升您的研究工作做好第一步准备。

浏览和搜索

浏览：

开放研究数据平台的首页下方展示着一系列“特色数据空间”。点击其中任何一个，即进入该数据空间的信息视图。在此可阅读该空间的说明介绍文字，浏览（或进一步搜索）空间内的数据集。可以使用排序和翻页功能，迅速浏览数据空间内的所有数据集。在数据空间和数据集清单的左侧，平台提供了多种筛选工具，可以限定是否显示数据空间、是否显示数据集、是否显示文件（默认未选中），以及根据分类、单位、发布时间、作者、学科、关键词、提交日期等元数

据项目进行快速筛选。数据空间、数据集、文件等可能同一列浏览与搜索结果清单中出现，但不难区分：数据空间为橙色边框，抽象元宇宙图标；数据集为蓝色边框，写字纸片图标；数据文件为灰色边框，图标多种多样。

简单搜索：

简单搜索方式会在数据空间、数据集、数据文件的**所有**元数据字段范围内展开搜索，并返回相关的搜索结果。有两个位置可以找到简单搜索的输入框：

(1) 在任何页面的顶端，点击放大镜图标展开的输入框，都可以做简单搜索。

(2) 在任何搜索结果的上方，都可以在输入框中输入搜索词，随时再次进行简单搜索。

高级搜索：

执行过一次简单搜索后，在搜索结果页面，搜索框的右侧会出现“高级搜索”的链接。点击后进入高级搜索界面。高级搜索界面提供在数据空间、数据集或文件的个别元数据字段中根据指定的搜索词执行的搜索功能。字段之间的逻辑关系为“与”，即返回的每一条搜索结果都应匹配所有填写了内容的搜索输入框。

查看各级信息页面

从浏览或搜索结果中点击任意数据空间，在打开的信息页面中可以找到有关该数据空间的说明文字，以及所包含的密级数据空间和数据集的列表。

从浏览或搜索结果中点击任意数据集，在打开的信息页面中可以找到有关该数据集的描述、学科、关键词等信息，以及所包含的数据文件列表、使用许可和版本沿革。

引用和下载

引用数据集：

在查看数据集的信息页面时，数据集标题下方浅蓝色背景的一段文字为该数据集的引文格式，包含多种元数据信息，如作者、发布年份、标题、DOI、数据发布平台版本号、UNF 数字指纹。

需要引用时，请按照您应遵守的学术写作规范重新编排上述信息的顺序和格式。

下载引用文件：

在查看数据集的信息页面时，浅蓝色背景的引文右侧有一个“下载引用”按钮。点击后，下拉菜单提供 EndNote XML 和 RIS 格式两种常见的参


考文献管理软件引文著录格式。请选择自己适用的格式下载。

部分采用专用格式的数据文件，点击  按钮后，会出现下拉菜单，包含一个子菜单，其中也提供 RIS 格式和 EndNote XML 两种引文著录



下载数据文件

数据集的信息页面下方，可以找到数据文件的列表。

其中部分文件右侧可见  按钮。其中有些采用专用格式的数据文件可能提供多种可下载的变异格式，然则点击“下载”按钮后会出现下拉菜单，如下图所示。

原始文件格式 (SPSS SAV)

制表符分割文件

RDATA格式

可变元

Data Subset

开始下载前，系统还会提示访问条款。如果您理解并愿意接受条款，点击“接受”按钮后即可开始下载该文件。有部分文件可能见不到“下载”按钮，而附有红色锁具图标，代表文件受限。受限的文件可能只允许开放研究数据平台的注册或认证用户下载，而您还没登录开放研究数据平台。受限的文件可能还需要向管理员申请访问许可后才能下载，而您暂未申请，或申请未获批准。

如果您在搜索和浏览时选择了显示数据文件的筛选项，结果列表中的某些数据文件的右侧可以见到一个砖红色的锁形图标，这表示该文件的使用权是“受限”的，即便注册并登录到开放研究数据平台，也需要申请访问许可并获批后才能使用。

【提示】说明文档类的外围数据文件，往往无需登录即可下载使用。然而，无论是否登录，您都应当遵守平台用户使用协议，以及您在下载之前点击接受的访问条款，避免滥用任何数据文件。

【文件校验】 系统显示每一个数据文件的 MD5 校验值，帮助用户检验下载文件的完整度。对于外围数据文件来说，这也是辨别修订版本的依据之一。下载文件后，Windows 用户可使用微软提供的 File Checksum Integrity Verifier 小工具检查文件的完整性，Linux/Unix 用户则可使用 md5sum 等命令，苹果电脑 Mac OS X 用户可使用 Terminal 终端下的 md5 命令进行校验。

完善账户信息

开放研究数据平台可供匿名访客浏览并利用部分完全开放的数据，但如需利用其他数据或发布自己的数据，就必须登录到平台中。这需要与平台管理员联系以创建登录账号。

【安全提示】 如果您在公用电脑上登录开放研究数据平台，离开前请务必退出登录状态，即点击页面右上方自己的姓名，再点击下拉菜单中的“退出”命令。

申请数据权限

如果您已登录到开放研究数据平台，在数据文件列表中却找不到所需文件的下载按钮，表明您尚未加入管理员设置的用户组，不具备特定数据的使用权。

作为数据的利用者，您可以主动向数据管理者申请加入用户组，获得数据的使用权。

(1) 请尝试在数据文件列表的上方点击“申请加入用户组下载文件”文字之下的“申请”按钮。

(2) 如果申请顺利提交，页面中会浮现提示，表明您已经成功发送了申请。

(3) 数据的管理者也会接到通知，然后在后台管理页面审核您的申请。

如果审核通过，管理者批准您使用数据，您的邮箱将收到通知邮件。然后就可以回到先前不能下载的数据文件那里，点击新出现的“下载”按钮了。

如果您的申请被管理者拒绝，您也会收到相应的通知邮件。

真实而全面的账户信息有助于申请获得批准，但数据管理者有权决定批准或不批准您的申请。开放研究数据平台不干涉数据管理者的自由决定权。


有的数据空间或数据集对申请者的账户信息做了特殊要求。

可能要求填报您所在的机构、组织、单位的公务电子邮箱，而非个人邮箱（或将个人邮箱填入“其他邮箱”一栏）。

如果您无法提供机构邮箱，有的数据空间也接受通过电子邮件向管理者发送工作证件的电子复本的审核途径，详见相应数据空间的说明信息。

在线分析浏览数据

部分数据文件采用了平台系统可识别的专用格式，如 SPSS 统计数据格式等。平台不仅为这类格式提供下载数据文件单独的引文著录文件和多种变异数据文件格式，而且提供便捷的在线分析浏览功能。

凡是符合规格的数据文件，如果您已经申请取得了访问权，那么在浏览或搜索界面（文件列表）中，不仅可以见到“下载”按钮，还能见到一个  浏览按钮，点击后同样提示访问条款，您理解并接受后才会打开在线分析浏览的界面。