

DB32

江苏省地方标准

DB32/T 4977—2024

农业农村大数据平台数据共享技术规范

Technical specification of data sharing for agricultural and rural
big data platform

2024-12-27 发布

2025-01-27 实施

江苏省市场监督管理局 发布
中国标准出版社 出版

目 次

前言Ⅲ

1 范围1

2 规范性引用文件1

3 术语和定义1

4 缩略语1

5 体系架构2

6 功能架构2

 6.1 概述2

 6.2 数据资源采集管理3

 6.3 数据资源治理管理3

 6.4 数据资源目录管理3

 6.5 共享交换管理4

 6.6 共享交换订阅4

7 共享业务架构4

 7.1 概述4

 7.2 数据资源提供4

 7.3 数据资源使用6

 7.4 数据资源共享7

 7.5 其他需求对接8

8 安全要求8

 8.1 日志要求8

 8.2 网络安全要求8

 8.3 数据安全要求9

 8.4 账号管理要求9

附录A(规范性) 江苏省农业农村大数据平台使用方申请及审批表10

附录B(规范性) 用户责任书11

参考文献12

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由江苏省农业农村厅提出并归口。

本标准起草单位：江苏省互联网农业发展中心、中国农业大学、北京中农信达信息技术有限公司、北京佳格天地科技有限公司、农芯(南京)智慧农业研究院有限公司、江苏省智慧农业研究会。

本标准主要起草人：吴昊、尚芬芬、魏祥帅、黄慧宇、李道亮、刘烨、顾竹、陈天恩、徐茂、赵然、王聪、杨李君、刘鹏、王绪琪、张帅、张强、吴众望、彭欣、杨锐、袁泽宸、陈雯、魏晓华、王振智、李震、王坦、张盼、田梦玲、甄华、鲍新宇、杨文治、刘雪、苏傲。

农业农村大数据平台数据共享技术规范

1 范围

本文件规定了农业农村大数据平台(以下简称“大数据平台”)数据共享的体系架构、功能架构、共享业务架构和网络安全要求。

本文件适用于江苏省各级农业农村部门间数据共享。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中,注日期的引用文件,仅该日期对应的版本适用于本文件;不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

GB/T 22239 信息安全技术 网络安全等级保护基本要求

NY/T 3501—2019 农业数据共享技术规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

农业农村大数据 agricultural and rural big data

具有农业农村地域性、季节性、多样性、周期性等特征的来源广泛、类型多样、结构复杂、数据量巨大、具有潜在价值的数据集合。

3.2

农业农村大数据平台 agricultural and rural big data platform

整合多渠道农业农村数据,引入数据挖掘展现技术,以专业分析为导向,面向服务对象提供数据查询、在线分析、共享交流等应用服务的系统环境。

3.3

数据资源目录 catalogue of agricultural big data resources

与组织业务活动相适应的,便于检索的多层次分类、分级的数据资源编排形式。

[来源:GB/T 42450—2023]

3.4

无条件共享数据 unconditional sharing data

可提供给各级农业农村部门共享使用的农业农村数据。

3.5

有条件共享数据 conditional sharing data

可提供给省内指定农业农村部门共享使用或者经审批后方可提供的农业农村数据。

4 缩略语

下列缩略语适用于本文件。

API:应用程序编程接口(Application Programming Interface)

JSON:轻量级的数据交换格式(JavaScript Object Notation)

MySQL:一种开源的关系型数据库管理系统(My Structured Query Language)

RESTful:表述性状态传递(Representational State Transfer)

5 体系架构

通过大数据平台完成省内主要农业农村数据资源汇聚,建立并管控省内数据传输通道。省内各级农业农村部门向大数据平台提供或从平台获取数据资源,实现省内农业农村数据资源交换与共享,并通过大数据平台与农业农村部平台、省政务数据共享交换平台实现对接。大数据平台共享体系架构如图 1 中虚线范围所示。

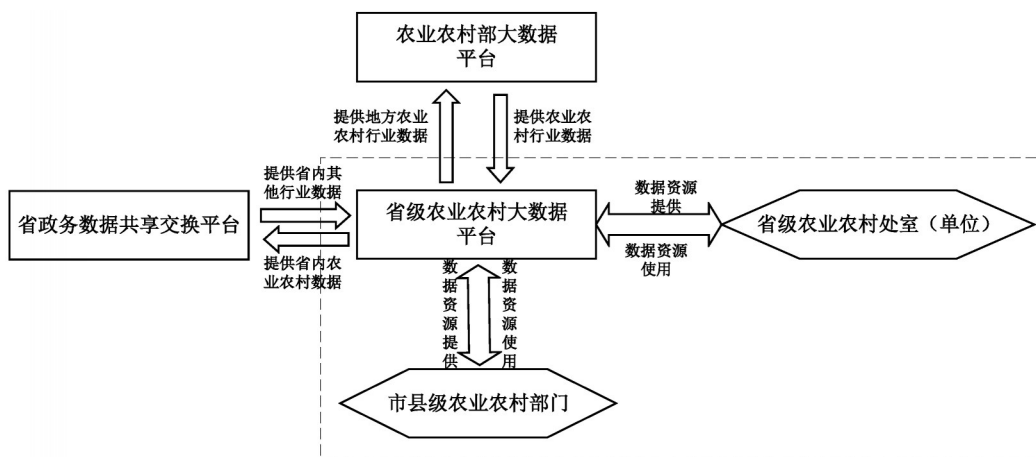


图 1 大数据平台共享体系架构

6 功能架构

6.1 概述

大数据平台功能架构包含数据资源采集管理、数据资源治理管理、数据资源目录管理、共享交换管理和共享交换订阅五大类功能,功能架构如图 2 所示。

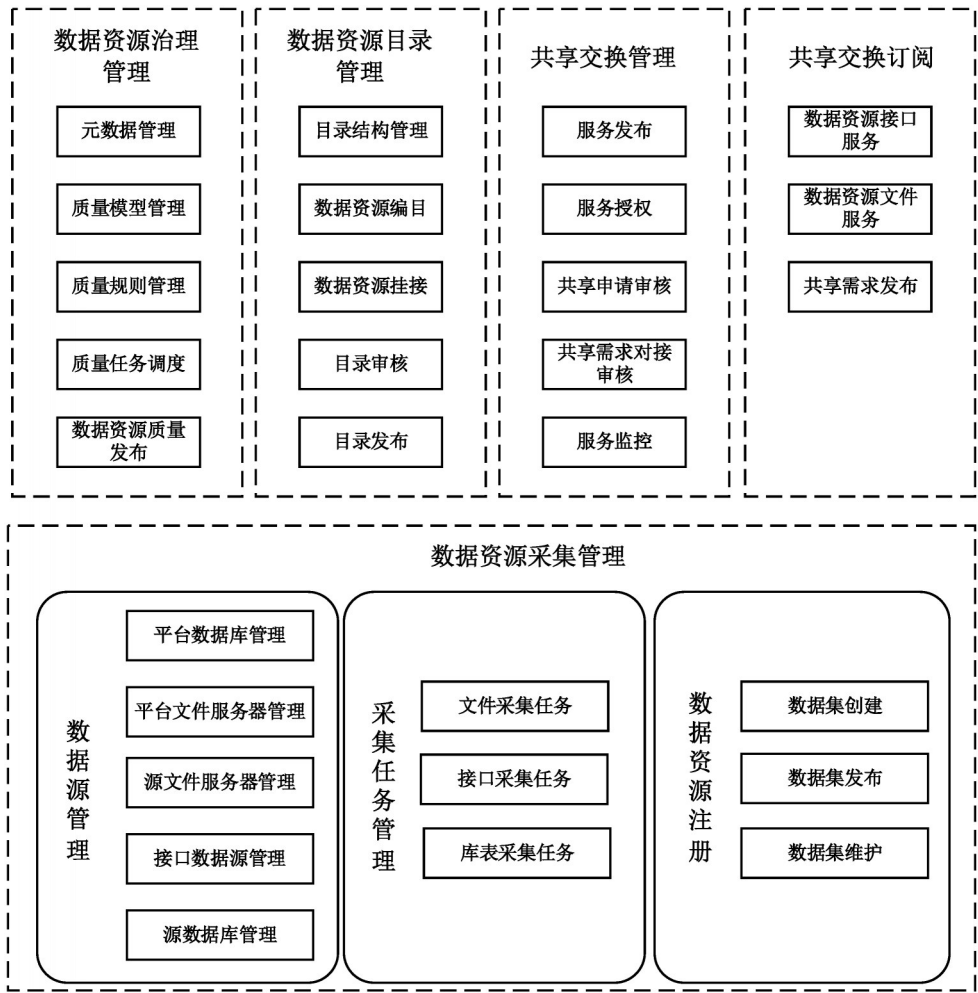


图2 功能架构

6.2 数据资源采集管理

数据资源提供方使用数据资源采集管理进行数据对接。该模块支持数据源管理、采集任务管理和数据资源注册管理,提供文件采集、接口采集、库表采集等数据资源提供方式,实现源数据资源向平台数据资源汇聚、集中存储,并能基于平台资源快速完成数据资源注册,包括数据集创建、发布、维护等。

6.3 数据资源治理管理

数据资源提供方、平台管理方使用数据资源治理管理对原始数据资源进行数据资源质量稽核。数据资源治理应支持元数据管理、质量模型管理、质量规则管理、质量任务调度、数据资源质量发布等,为促进数据资源质量的持续改进提供技术支持。原始数据资源本身的质量管理和清洗工作由数据资源提供方负责。

6.4 数据资源目录管理

数据资源目录管理应支持目录结构的灵活配置管理,并能方便、快捷地进行数据资源目录编制、数据资源挂接、审核和发布。目录管理应提供灵活的查询功能,便于目录定位查找,同时为了保证目录内容及实时更新,应提供对发布的目录进行维护的功能。

6.5 共享交换管理

数据资源申请方可以在共享交换管理模块进行数据资源申请操作。该模块支持服务发布与授权、共享申请审核、共享需求对接审核、服务监控等功能。

6.6 共享交换订阅

数据资源申请方可以在共享交换订阅管理模块对周期性数据资源进行订阅操作。共享交换订阅作为面向平台所有用户的共享门户,提供统一的数据资源服务目录以及数据资源服务查询、获取功能,应支持用户发布、跟踪新数据资源共享需求。

7 共享业务架构

7.1 概述

大数据平台共享业务架构由数据资源提供方、数据资源使用方、平台管理方三种角色和数据资源提供、数据资源使用、数据资源共享与其他需求对接四个业务过程组成。其中,其他需求对接是指大数据平台已有数据资源服务无法满足的数据共享需求时采用的共享办法。大数据平台共享业务架构符合江苏省公共数据管理办法。

7.2 数据资源提供

数据资源提供过程主要包括目录编制与审核、数据资源注册与挂接、数据资源审核、服务发布。数据资源提供方完成目录编制、审核和数据资源注册、挂接后,由平台管理方进行数据资源审核,审核通过后的数据资源将以服务形式发布。

7.2.1 数据资源提供方要求

数据资源提供方在提供数据资源时应满足以下要求:

- 数据资源提供方应在大数据平台上编制数据资源目录,并注册相应数据资源;
- 数据资源注册的信息应与目录信息保持一致,接口方式注册需事先由数据资源提供方组织并通过接口调用测试;数据资源注册时,应逐一确定数据资源内容、基本信息、注册形式、更新方式等,根据不同数据资源提供方式,提供数据资源链接或文件,并将数据资源与目录进行关联;
- 提供的数据资源有变更的,数据资源提供方应及时更新、维护数据资源目录和数据资源,确保所提供的数据资源与本部门掌握数据资源一致;
- 数据资源提供方应确保信息的真实、可靠、完整;
- 数据资源提供方对资源的管理要求应符合 NY/T 3501—2019 中 4.2.1 的规定。

7.2.2 数据资源目录管理要求

数据资源目录管理要求如下:

- 应以十大板块(种植业、畜牧业、渔业、农机化、农田建设、质量安全、农村经济、乡村发展、科教与信息、政务服务)为基础,建立统一的数据资源目录版本;
- 目录发布前应进行合规性检测,包括检测目录格式、编制范围等;
- 建立目录更新机制,定期开展数据资源目录的梳理和一致性检查。

7.2.3 数据资源提供规则

数据资源提供规则分为以下两种:

- a) 增量提供:一次性将所有历史数据提供给平台,并周期性将数据资源提供给平台;
- b) 全量提供:周期性将全量数据资源提供给平台。

数据资源提供方应优先采用增量方式提供数据资源,当数据资源不具备增量标识且人为无法指定增量标识时宜采用全量方式提供数据资源。

7.2.4 数据资源授权要求

申请无条件共享数据资源的,数据资源使用方能自动获得授权。申请有条件共享数据资源的,由数据资源提供方通过平台授权使用,大数据平台应设置授权答复期限。

7.2.5 数据资源提供方式

7.2.5.1 数据文件

以数据文件的形式提供数据资源,文件提供方式应满足以下要求:

- a) 结构化数据资源应支持按需提供和周期性提供;
- b) 结构化数据资源应支持 wps、doc、xml、txt、docx、html、ppt、et、xls、xlsx、csv 等文件格式;
- c) 结构化数据资源应支持文件服务器在线传输和计算机本地文件直接上传;
- d) 结构化数据资源和非结构化数据资源应满足表 1 中的要求。

表 1 数据文件提供技术规范

序号	属性		规则说明
1	文件名称		[数据名称].[文件序号].[文件创建日期:YYYYMMDD].[后缀名]
2	文件内容	结构化类型	1) 列为序列;第 1 行为数据元中文名称; 2) 始数据内容包含 html 标签,需去除; 3) 重复值不超过数据总量的 5%; 4) 相应文件格式为:xml、csv、xls、xlsx
3		非结构化类型	若为行列式类型的表单数据,需转换为结构化数据后提供
4	文件大小		单个文件大小不超过 2 G

7.2.5.2 数据接口

以数据接口方式提供数据资源,数据接口提供方式应满足以下要求:

- a) 应支持按需提供和实时提供;
- b) 应支持任意数据库类型的数据存储;
- c) 应支持根据共享数据资源业务内容定制接口开发;
- d) 应满足表 2 中的要求。

表 2 数据接口提供技术规范

序号	属性	规则说明
1	接口地址	1) 应采用共享交换数据接口专用二级域名; 2) 应采用 RESTful 设计规范设计接口地址,定义好资源版本号、资源英文名称、直接访问接口地址默认返回值等
2	通信协议	应支持 http、https 协议

表 2 数据接口提供技术规范（续）

序号	属性		规则说明
3	接口参数	基本参数	应定义数据条数获取限制、身份认证、数据分页、数据排序方式、数据返回方式(JSON 或 xml)等基本参数及说明
		业务参数	应定义业务查询参数及说明
4	返回结果		应定义返回结果的数据内容、说明及返回格式,返回格式包括但不限于 JSON、xml 等格式
5	错误代码		应定义返回错误时的错误代码和错误原因
6	接口性能		1) 返回结果时间应不超过 3 s; 2) 应支持至少 300 以上的访问并发量
7	接口所在网络区域		接口所在网络区域为政务外网

7.2.5.3 交换库

数据资源提供方以交换库形式提供数据资源,将共享数据同步至该交换库。交换库由平台管理方通过大数据平台创建,交换库提供方式应满足以下要求:

- a) 应支持按需提供和周期性提供;
- b) 应确保交换库数据与业务数据的一致性;
- c) 应支持关系型数据库类型交换库;
- d) 应满足表 3 的要求。

表 3 库方式数据提供技术规范

序号	属性	规则说明
1	数据库名称	用户申请后平台应支持自动创建生成
2	数据表信息	应包含字段名称、字段类型、字段长度及字段注释
3	数据质量	1) 包含空值行数不超过总行数的 2%; 2) 应去除重复值; 3) 应与业务系统保持同步更新

7.2.6 需求响应要求

数据资源使用方向数据资源提供方发出数据需求申请后,数据资源提供方在规定日期内向数据资源使用方做出响应,若数据资源提供方同意需求申请,则应通过大数据平台向数据资源使用方提供。如果数据资源提供方不同意提供需求数据,应向数据资源使用方和江苏省农业农村大数据资源主管部门说明原因。

7.3 数据资源使用

数据资源使用过程主要是数据资源服务的调用。数据资源使用方作为数据资源共享的需求方,通过平台提供的接口、数据库视图、数据文件订阅、联机查询等服务方式,获取所需数据资源。

7.3.1 数据资源使用方要求

数据资源使用方在使用数据资源时应满足以下要求:

- a) 数据资源使用方应严格遵守国家相关法律法规,保证共享数据资源仅用于本部门相关业务工作需要,不应直接或以改变数据资源形式等方式提供给第三方,也不应用于或变相用于其他目的;
- b) 属于有条件共享类的数据资源,应严格按照供需双方约定的方式和适用范围使用数据资源;
- c) 数据资源提供方和平台管理方在数据资源使用方测试联调期间应给与必要的技术支持;
- d) 数据资源使用方应根据数据资源使用情况,向平台反馈数据资源使用情况,包括数据资源满足需求情况、数据资源质量情况及数据资源使用效果;
- e) 数据资源使用方对数据资源的管理要求应符合 NY/T 3501—2019 中 4.2.3 的规定。

7.3.2 数据资源使用方式

7.3.2.1 数据文件订阅

以数据文件订阅方式使用数据资源,应符合以下要求:

- a) 使用方式应支持周期性获取和按需获取;
- b) 周期性获取应依托交换通道技术实现文件的传输,获取同一批次的资源应包括索引文件和数据文件;
- c) 按需获取应支持直接下载文件。

7.3.2.2 数据接口服务

以数据接口服务方式获取或验证数据资源,应符合以下要求:

- a) 应包含认证接口和业务接口,认证接口用于验证调用者身份的合法性,业务接口用于提供数据资源服务,认证成功后可调用数据资源服务接口;
- b) 认证接口信息应包含服务地址、接口路径、接口类型、接口请求方式、接口编码、内容类型;
- c) 业务接口信息应包含接口路径、接口类型、接口请求方式、接口编码、内容编码类型。

7.4 数据资源共享

7.4.1 概述

数据资源共享由数据资源提供方、数据资源使用方和平台管理方共同完成,数据资源共享过程包括数据资源申请、数据资源授权和共享监测分析。

7.4.2 职责分工

- a) 数据资源使用方应在线提交数据资源使用申请,并明确使用范围、应用场景、申请依据等;
- b) 数据资源提供方负责审核数据资源使用方提交的有条件共享类数据资源的申请,对同意共享的数据资源进行线上授权,对不同意共享的数据资源应说明理由;
- c) 平台管理方对数据资源共享全流程进行监测分析。

7.4.3 共享数据变更要求

共享数据变更要求包括:

- a) 数据资源提供方变更或下架数据资源时,应提前和已订阅的数据资源使用方协商,并填写数据资源变更及下架原因上报江苏省农业农村大数据资源主管部门;非因法定事由,数据资源提供方不应在共享中的数据资源提出变更;
- b) 以文件方式共享的数据资源变更时,平台应保留数据资源原始版本,并将变更后的新版本命名为最新版本编号发布至平台;

- c) 以数据服务接口方式共享的数据资源一旦数据结构发生变更,应和数据资源使用方协商沟通。

7.5 其他需求对接

其他需求对接由数据资源提供方、数据资源使用方和平台管理方共同完成,过程包括共享需求提出、需求审核与反馈、需求跟踪。数据资源提供方、数据资源使用方和平台管理方在数据资源共享过程中的职责分工如下:

- a) 数据资源使用方应在线发布数据资源使用需求,并明确使用范围、应用场景、申请依据、数据格式、数据结构、数据资源更新频率、数据资源使用方式等主要信息;
- b) 数据资源提供方负责评估、审核数据资源使用方提交的共享需求申请,对同意共享的数据资源进行数据资源提供、线上授权,对不同意共享的数据资源应说明理由;数据资源提供方对数据资源使用方的数据资源调用记录进行监控,以便对数据资源异常调用方进行相关处理;
- c) 平台管理方负责跟踪共享需求对接情况。

8 安全要求

8.1 日志要求

平台应对服务器、安全系统、应用支撑以及业务系统的日志信息进行统一管理,实现全面监控和实时告警,记录的日志包含但不限于登录日志、操作日志、服务调用日志、服务监控日志,要求如下:

- a) 登录日志信息,包括登录名、姓名、登录时间等;
- b) 操作日志信息,包括操作人、操作时间、操作事件、操作记录、事件类型等;
- c) 服务调用日志信息,包括服务名称、调用入、调用时间、调用结果等;
- d) 服务监控日志信息,包括服务名称、服务状态、监控时间等;
- e) 平台应保证平台日志无法删除、修改或被覆盖;
- f) 平台宜提供对日志进行统计、查询、分析及生成报表的功能。

8.2 网络安全要求

8.2.1 通则

大数据平台应部署在政务云的政务外网区域。所有涉及需要跨网的环节,都通过政务云的安全跨网交换平台进行。

8.2.2 强化身份认证和访问控制

确保只有经过授权的用户能够访问共享数据。使用多因素身份验证技术,提高身份验证的安全性。同时,实施访问权限管理,确保用户只能访问其被授权的数据。

8.2.3 强化 API 安全

确保 API 的安全配置和管理对 API 进行身份验证、授权和访问控制。定期审计和评估 API 的安全性,修复潜在的漏洞和风险。

8.2.4 持续漏洞管理

进行持续的漏洞扫描和漏洞管理,及时修补发现的安全漏洞。采用自动化的漏洞管理工具,确保及时应用安全补丁和更新。

8.3 数据安全要求

8.3.1 数据加密

应采用密码技术保证个人信息、重要数据、关键业务数据等在传输过程中的保密性,必要时应对数据进行加密存储。

8.3.2 数据脱敏与防泄漏

应对个人信息、重要数据、关键业务数据等进行脱敏处理,防止数据泄露。应通过有效技术手段防止数据被恶意爬取。

8.3.3 共享安全

共享安全要求包括:

- a) 在数据资源共享过程中,提供和使用数据资源的信息系统应按照 GB/T 22239 的要求开展等级保护定级;
- b) 大数据平台应建立数据资源接入和使用安全规范,明确资源使用方式、授权机制和使用的权限范围等,建立规范的数据资源共享审核流程,明确各方职责分工和安全责任;
- c) 大数据平台应制定数据服务接口调用安全控制策略,提供对资源服务的安全限制和安全控制措施,如身份鉴别、授权策略、访问控制机制、签名、时间戳、安全协议等,对数据服务接口调用的参数进行限制或过滤,设置异常触发告警机制,并对数据接口调用情况进行定期审计。

8.4 账号管理要求

账号管理要求包括:

- a) 各级农业农村部门向江苏省农业农村大数据资源主管部门申请开通数据资源管理员账号时,应如实填写大数据平台用户申请表(应符合附录 A 的规定)并签订用户责任书(应符合附录 B 的规定);
- b) 各部门数据资源管理员应严格审核本部门下注册的数据资源提供员账号和数据资源使用员账号;
- c) 各部门数据资源管理员账号信息变更前,应将账号信息变更需求按程序报送至江苏省农业农村大数据资源主管部门审核;
- d) 各部门数据资源管理员遗失账号或密码,应书面提交申请报送到江苏省农业农村大数据资源主管部门核准找回账号或密码。

附 录 A
(规范性)

江苏省农业农村大数据平台使用方申请及审批表

江苏省农业农村大数据平台使用方申请及审批表如表 A.1、表 A.2 所示。

表 A.1 江苏省农业农村大数据平台使用方申请表

申请日期：	
申请单位	
申请人	
手机	
有效期限	
申请资源名称	
访问 IP	
数据申请范围	
申请用途	
申请时间	

表 A.2 江苏省农业农村大数据平台使用方审批表

编号：	审批日期：
经办人	
审批人	
经办意见	
审核意见	
备注	

附 录 B
(规范性)
用户责任书

用户应认真阅读用户责任书,自觉遵守其各项规定。

- 1) 用户在使用共享交换平台时,应遵守国家的有关法律、法规和行政规章制度。
- 2) 用户在使用共享交换平台时,应严格遵守《江苏省政务数据资源管理办法》。
- 3) 用户只能把从共享交换平台获取的共享信息用于本单位履行职责需要,不应用于其他任何目的。
- 4) 未经提供单位的同意,用户不应擅自向社会发布和公开所获取的共享信息。
- 5) 用户登录共享交换平台所使用的账号,仅限于用户个人使用,不应交由他人使用。用户应妥善保管账号。用户原因造成该账号泄露的,一切后果由用户承担。
- 6) 当用户的账号在使用过程中出现异常情况时,用户应积极配合交换平台管理员解决。
- 7) 用户违规使用共享交换平台涉及商业秘密或个人隐私的共享信息,或造成商业秘密或个人隐私泄露的,按国家有关法律法规的规定处理,构成犯罪的,依法追究刑事责任。

经办人(签字):

申请单位(盖章)

日期: 年 月 日

参 考 文 献

- [1] GB/T 37722—2019 信息技术 大数据存储与处理系统功能要求
 - [2] GB/T 38673—2020 信息技术 大数据 大数据系统基本要求
 - [3] GB/T 41187 农业物联网应用服务
 - [4] GB/T 42528—2023 时空大数据技术规范
 - [5] DB37/T 4413.1—2021 生态环境数据共享技术规范 第1部分:城市空气
 - [6] DB52/T 1123—2016 政府数据 数据分类分级指南
 - [7] 国家数据共享交换平台(政务外网)省级平台接入指南(试行)
 - [8] 政务信息资源目录编制指南(试行)(发改高技[2017]1272号)
 - [9] 江苏省公共数据管理办法(江苏省人民政府令 第148号)
-