

目录

目录	1
产品概述	3
名词解释	3
行业场景与资源	3
数据维护	3
数据审批	3
数据发现	3
产品版本	3
产品优势	4
统一元数据管理	4
强大的数据血缘分析	4
快速数据识别与发现	4
企业级数据安全合规	4
产品功能	4
库表管理	4
数据血缘管理	4
数据热度分析	4
数据权限管理	4
数据质量监控	4
数据地图	4
新建数据源	4
数据库创建及发布	5
新建数据库	5
发布到测试	5
发布到生产	5
查看发布申请	5
发布申请撤回	5
发布审批	5
数据表创建及发布	6
新建数据表	6
表发布到测试	6
表发布到生产	6
查看发布申请	6
发布申请撤回	6
发布审批	6
标签管理	7
新建一级标签	7
导入标签	7
业务元数据管理	7
分类管理	7
元模型管理	7
元数据维护	7
数据关联配置	8
数据搜索	8
数据目录	8
数据分布	8
数据血缘关系	9
数据质量	9
规则配置	9
质量查询	9

质量告警	9
数据权限管理	9
权限申请	10
申请记录	10
撤销申请	10
权限审批	10
数据授权	10
常见问题	11

产品概述

金山云智·数据管理是一种全方位的企业数据治理服务，提供云端传统关系型数据库、Nosql数据库、大数据平台数据仓库的统一管理，提供数据地图、数据质量、数据血缘分析能力。通过统一元数据管理解决金融、政府等行业数据孤岛、数据质量、数据合规使用问题。助力客户掌握数据全景，打通采数、识数、取数、用数全流程，实现数据管理无死角。

名词解释

- **元数据** 描述数据及其环境的数据，包括技术元数据、业务元数据、操作元数据。
- **业务元数据** 具体一定业务含义或高度抽象的元数据，如业务术语、数据字典。
- **技术元数据** 描述数据层的元数据，分为逻辑元数据和物理元数据，如库表逻辑结构和物理结构。
- **元数据托管** 非托管的数据源数据管理只提供数据源链接信息；托管的数据源支持元数据同步、库表创建和库表权限管理。
- **元数据同步** 从数据源同步库表结构信息到统一元数据服务。
- **设计态** 数据管理遵从先设计后发布的理念，设计态的数据跟任何环境没有关系，可随时修改。设计完的库表可发布到物理环境。设计完的业务数据发布后别人才可见。
- **密级** 数据具有密级属性，分为1、2、3、4、5档，1为最高密级，5为最低密级。数据密级跟用户密级对应，当用户密级不够时无法查看相关数据。
- **业务元模型** 类似一张表，定义了业务数据的结构和属性，如数据标准清单是一种业务元模型，包含了标准编号、标准名称、标准编码等字段及字段属性。

行业场景与资源

数据维护

- 开发人员维护数据表结构、创建技术检核质量规则；
- 业务人员维护业务数据，创建业务检核质量规则；
- 有权限的用户可对具体的表、字段、业务数据打标签。

数据审批

管理人员进行元数据变更审批，数据权限审批，质量规则发布审批。

数据发现

- 所有用户可通过数据目录发现自己感兴趣的数据，查看具体数据库表详情，业务数据详情，并申请数据权限；
- 所有用户可通过数据搜索发现自己感兴趣的数据，查看具体数据库表详情，业务数据详情，并申请数据权限；

产品版本

数据管理将按版本进行开通，分标准版、专业版和企业版，用户在购买大数据云的时候默认开通标准版。各版本预计包含的功能模块如下：

一级菜单	二级菜单	标准版	专业版	企业版
数据源管理		√	√	√
元数据管理	库表管理	√	√	√
	标签管理	√	√	√
	业务数据	√	√	√
	数据关联	√	√	√
	数据权限			
数据权限	数据授权	√	√	√
	我的申请	√	√	√
	我的审批	√	√	√
发布管理	我的申请	√	√	√
	我的审批	√	√	√
数据质量	规则配置			√
	质量查询			√

	质量告警		✓
数据血缘		✓	✓
数据地图	数据搜索	✓	✓
	数据目录	✓	✓
	数据分布	✓	✓

产品优势

统一元数据管理

- 多种异构数据源元数据同步与托管
- 在线库表模型设计与多源发布
- 灵活的业务元模型定义，统一数据标准

强大的数据血缘分析

- 多级血缘查询，快速定位问题
- 通过上下游关系，深入理解数据
- 评估数据变更影响分析

快速数据识别与发现

- 数据资产一键检索定位
- 技术、业务、标签多角度构建数据资产目录
- 标签、热度、质量指标快速定位有用数据

企业级数据安全合规

- 严格的数据权限申请与审批流程
- 严格的元数据变更管理流程
- 逻辑层、物理层双层数据鉴权机制
- 数据密级管理与脱敏处理

产品功能

库表管理

支持创建、删除、修改数据库，支持创建、修改、删除表等操作。支持库表模型设计，库表可由设计态发布到测试环境和生产环境，支持一个库/表同时发布到多个数据源。

数据血缘管理

从数据采集/同步，到数据ETL、统计分析、结果查询等数据产生的全链路对数据血缘进行分析，便于了解数据流链路情况。

数据热度分析

根据统一SQL查询服务使用情况统计各个库表的查询使用情况分布；并结合数据生产冷热情况，分析最终数据利用情况。

数据权限管理

支持管理关系型数据库、ES、Hbase、kafka、图数据库、对象存储等数据源的数据权限，支持数据权限的在线申请和审批。

数据质量监控

支持批量设置数据质量检核规则，对数据处理流程全链路各个环节进行数据质量的监控管理。支持质量检核告警记录查询与管理。

数据地图

支持库表基本信息、分区、存储量分布等各类统计信息查看。支持以技术目录、业务目录、标签目录等多种方式查看元数据信息。支持按关键字全文检索所有类型元数据信息，支持按技术目录、业务目录、标签目录等多种高级方式筛选元数据。

新建数据源

1. 登录大数据云平台，点击【数据管理】，进入【数据源管理】页面
2. 点击【新增数据源】按钮
3. 点击选择需要添加的数据源，目前支持Oracle、MySQL、Hive、HBase、Kafka、Elasticsearch、Redis、FTP、HTTP、ArangoDB等多种数据源。
4. 填写配置（以Hive为例，点击选择Hive，填写新增hive数据源对话框中的配置）。
5. 点击【连通性测试】按钮，测试通过后点击完成。

【补充说明】

如果要连接云上RDS数据源，需在RDS配置大数据云IP白名单。

大数据云外网IP：120.92.11.137

大数据云内网IP：10.69.66.222，10.69.66.149（两个都需配上白名单）

数据库创建及发布

新建数据库

数据源创建完成后，进入【库表管理】页面。

1. 点击【库表管理】按钮，点击【新建数据库】按钮。
2. 填写新建数据库对话框中的配置。
3. 点击【确认】按钮，提示创建成功。

发布到测试

1. 新建数据库完成后，点击【发布到测试】。
2. 选择数据源，以hive数据源为例，点击【确认】按钮提示发布成功。

发布到生产

1. 管理员审批通过后，在【库表管理】中【测试环境】下【我设计的库表】中显示。
2. 点击【发布到生产】
3. 点击【确认】按钮，提示发布成功。
4. 管理员审批通过后数据库将在【库表管理】中【生产环境】【我设计的库表】下显示。

查看发布申请

1. 点击【发布申请】按钮可以查看发布的申请号。
2. 点击申请号可以查看发布明细
3. 点击【已结束的申请】查看已结束的申请记录

发布申请撤回

对于已经提交的发布到测试或生产的申请可以进行撤回

1. 点击【发布申请】
2. 点击选择测试或生产环境查看相应申请。
3. 点击【撤回】，输入撤回原因，点击【确认】。

发布审批

1. 管理员进入数据管理页面，点击【发布审批】。

2. 选择【测试环境】或【生产环境】查看环境下的待审批内容。
3. 待我审批模块，查看并审批其项目下待审批记录，包括表申请
4. 点击申请号查看详细信息，点击【查看】可以查看流程状态。
5. 点击通过按钮，弹出通过申请框，输入审批意见点击确认按钮提示审批成功。
6. 点击驳回按钮，弹出驳回申请框，输入审批意见点击确认按钮提示审批成功。
7. 点击我已审批菜单，查看已经处理过的申请记录，包括表。

数据表创建及发布

新建数据表

1. 数据库创建完成并审批通过后，点击库表管理中数据表选项卡
2. 点击新建数据表按钮
3. 填写创建数据表配置信息
4. 单击下一步按钮
5. 填写新建表页面中填写字段设置。
6. 单击下一步按钮
7. 进行分区设置
8. 点击完成按钮。提示创建成功，并且列表自动刷新展示刚刚新建的表。

表发布到测试

1. 新建数据表完成后，在列表点击发布到测试按钮。
2. 选择数据源，数据库，点击确认按钮。提示发布成功。

表发布到生产

1. 管理员审批通过后，在【库表管理】中【测试环境】下【我设计的库】表点击【显示表】显示找到要发布的数据表。
2. 点击发布到生产按钮。
3. 点击确认按钮，提示发布成功。管理员审批通过后，在【库表管理】中【生产环境】下【我设计的库表】点击【显示表】显示。

查看发布申请

1. 点击发布申请按钮可以查看发布的申请号。
2. 点击申请号可以查看发布明细
3. 点击【已结束的申请】查看已结束的申请记录

发布申请撤回

对于已经提交的发布到测试或生产的申请可以进行撤回。

1. 点击【发布申请】
2. 点击选择测试或生产环境查看相应申请。
3. 点击【撤回】，输入撤回原因，点击【确认】。

发布审批

1. 管理员进入数据管理页面，点击【发布审批】。

2. 选择【测试环境】或【生产环境】查看环境下的待审批内容。
3. 待我审批模块，查看并审批其项目下待审批记录，包括表申请
4. 点击申请号查看详细信息，点击【查看】可以查看流程状态。
5. 点击通过按钮，弹出通过申请框，输入审批意见点击确认按钮提示审批成功。
6. 点击驳回按钮，弹出驳回申请框，输入审批意见点击确认按钮提示审批成功。
7. 点击我已审批菜单，查看已经处理过的申请记录，包括表。

标签管理

新建一级标签

1. 点击选择【标签管理】，点击【新建一级标签】。
2. 填写一级标签名称，点击【确认】。说明：一级标签名称仅支持英文、数字、下划线，切不能以数字或下划线开头，最大100字符。

导入标签

1. 点击【导入标签】。
2. 点击下载【标签导入模板.xlsx】，按照要求填写各级目录名称，点击文件夹图标在您的计算机上浏览找到填写好的模板进行导入，点击【确认】。

业务元数据管理

分类管理

业务元数据的分类，如数据标准、数据字典、业务术语等，由全局数据管理员维护。启用的业务元数据分类全员可见，未启用的只有全局数据管理员可见。可视化业务包是默认生成的，主要用于可视化BI。

1. 左侧菜单选中业务元数据，默认进入分类管理tab页。点击【新增分类】按钮弹出设置框，输入分类名称和描述，点击【确认】即可。
2. 新增的业务分类需要启用才能生效，点击操作列的【启用】按钮启用确认提示，点击【确定】即可。当前启用状态的分类不支持修改和删除。

元模型管理

元模型隶属于业务分类，如数据标准分类下可以包含多个标准元模型。元模型管理提供元模型的新建、发布、下线、修改、授权、更改责任人、删除功能。由数据资产管理进行元模型的管理和维护。

1. 新建元模型支持自定义模板和导入模板两种方式。自定义模板需要用户手动逐个填入模板字段及相关定义。导入模板支持下线模板文件，将模板字段定义批量写入模板文件，然后上传到平台。

自定义模板类似定义一张表，需要定义该元模型包含的字段内容及相关属性。此处的字段属性跟建表时类似，此处不再赘述。需要说明的是“是否展示字段”，其含义是：元模型的默认展示字段，不展示的字段将会被隐藏，搜索结果不展现此类字段的内容，主要为了防止定义的字段太多，无法展示，且展示意义不大的问题。导入模板方式提供模板下载和上传，可以批量设置元模型字段。

2. 新建的元模型为待发布状态，可点击操作列下的【发布】按钮进行发布。发布元模型需要全局数据资产管理审批。元模型的发布申请记录可以在发布申请列表看到。审批通过的元模型全员可见。发布的元模型不用后可以下线，下线的元模型其他人不可见，只有元模型创建者和上级管理员可见。
3. 发布后的元模型支持修改，修改后元模型状态由已发布变为待发布更新。点击【发布】按钮将走变更发布流程。
4. 元模型可以授权给其他人，授权的用户才能向该元模型里添加业务数据。
5. 当人员离职或变更时支持修改元模型的责任人。将元模型转交给其他数据资产管理维护。
6. 下线的元模型支持删除。

元数据维护

定义了业务元模型下一步就可以添加业务数据。切换到元数据维护tab页，左侧目录树会显示发布的业务分类和元模型。选中具体的元模型右侧区域显示元模型定义的展示字段。

1. 点击【新增元数据】可单条新增业务数据记录。点击【导入】可向该元模型批量导入业务数据。只有授权的用户方可添加业务元数据。
2. 新增的业务数据需要发布才能生效，可点击具体记录操作列下的【发布】按钮单条发布，或选中每条记录前面的复选框，点击上方【发布】按钮进行批量发布。业务元数据发布需要元模型创建人审批。发布后的业务数据支持下线。
3. 若发布的业务数据需要变更，可直接进行修改，修改后的业务数据需要重新发布。
4. 已下线的业务数据支持删除
5. 点击操作列下的【历史】按钮可查看该条业务数据的历史变更记录。
6. 已发布的业务数据支持打标签，点击操作列下的【打标】按钮弹出打标界面。
7. 元数据维护页面默认展示元模型的展示字段（新建元模型时定义），要看其他字段内容可点击最右侧的设置图标，调整要展示的字段。

数据关联配置

数据管理支持技术元数据管理和业务元数据管理，以及技术元数据与业务元数据、业务元数据与业务元数据之间的关联关系配置。关联关系配置需要先定义关系，然后基于定义的关系配置具体的元数据关联。关系定义了关系类别（目前提供业务关联业务、业务关联技术两种关系类型）以及源实体和目标实体。业务的实体为具体的元模型，技术的实体和表或字段。

1. 左侧菜单切换到元数据关联页面，默认进入关系管理tab页。点击【新增关系】按钮可新建关系。
2. 新建的关系需要启用才能生效，生效的关系通过【停用】可以失效，即关系不可用。停用状态的关系方可删除。
3. 关系定义了之后可以进行关系配置。切换tab页到关系配置，系统默认会选中一个关系，下面展示该关系定义的源实体。可通过下拉列表切换关系类别和关系名称，下面的内容会相应联动变化。
4. 点击源实体的任意一行可设置该行内容与目标实体的关联关系。弹出的目标实体页面支持按已发布的关联和未发布的关联查看。选中源实体记录要关联的目标实体记录，系统会自动保存。
5. 关闭目标实体弹框，点击右侧【发布】按钮，即可发布刚才建立的关联关系。发布也需要走审批流程，关联关系的发布申请由项目数据管理员进行审批。

数据搜索

1. 点击【数据搜索】，在屏幕上放选择搜索内容所属的元数据空间（测试环境或生产环境），输入关键字即可进行简单搜索
2. 如需提高搜索效率，点击【展开高级选项】进行筛选
3. 如果选择【按技术元数据】，您可以通过限定数据库类型找到相应的数据库缩小搜索范围，如果选择【按数据标签】，您可以按照【标签管理】处设置的标签进行查找，在屏幕下方选择结果的展现形式（全部，表级或列级）。

数据目录

您可以通过【数据目录】功能按照技术元数据、数据标签、业务元数据查看数据库表信息和业务数据信息。

1. 点击【数据目录】，在页面上方选择数据空间。
2. 在屏幕右侧选择技术元数据或数据标签查看相应目录，点击下箭头展开下一级，点击数据库名称查看基本信息。
3. 点击【表信息】可以查看该数据库中各数据表的基本信息。
4. 点击左侧目录树的表，可以查看该数据表的详细信息，如基本信息、字段信息、数据预览、权限拓扑等。
5. 点击字段信息中具体某个字段后面的【详情】按钮，可以查看该字段的详细信息

数据分布

您可以通过【数据分布】查看项目数、数据量、数据库 / 表情况。同时也可以通过搜索某个项目查看该项目的数据情况。

1. 点击【数据地图】>【数据分布】，在页面上方选择数据空间。
2. 在搜索框搜索想要了解的项目名称，或者点击关系图中的某个项目，进入该项目的数据分布。
3. 继续在搜索框搜索数据库，或者在关系图中点击某个数据库，进入该数据库的数据分布情况。
4. 继续在搜索框搜索数据表，或者在关系图中点击某个数据表，进入该数据表的数据情况。

数据血缘关系

专业版和企业版支持数据血缘关系查询功能。您可以通过【数据搜索】，查看数据的血缘关系。

1. 点击【数据地图】>【数据搜索】，在搜索框里输入您要查找的数据。
2. 点击您需要查看的数据，进入详情页面，点击【数据血缘】Tab，选择上游及下游层数，查看数据血缘关系。
3. 上下滑动鼠标，缩放血缘关系图。鼠标经过某个节点，显示该节点的信息。鼠标点击某个节点，跳转到相应的节点详情页面。

数据质量

数据质量模块当前支持对Oracle、Hive、MySQL等关系型数据库进行质量检核。提供质量规则配置、质量任务查询、质量告警管理等功能。

规则配置

当前质量规则配置提供五种业务检核算子，只有企业版支持配置业务检核。目前支持取值范围检查、取值约束配置、数据格式检查、非空检查、唯一性检查等。

操作说明如下：

1. 新建规则 业务人员、开发人员可以新建规则。菜单切换到数据质量下面的规则配置，点击【新增规则】按钮弹出创建规则配置框。
2. 规则试运行 当前版本发布的业务检核规则支持试运行，用生产的几条数据进行试跑，可以批量试运行。试运行成功点击【试运行成功！点击查看结果】链接可以跳转到质量查询页面进行查看。
3. 规则发布 新建的规则发布后才能生效，点击发布后，确认即可发布成功。
4. 规则下线 不用的规则可以下线。点击下限，确认后即可下线该规则。
5. 规则修改、删除 未发布的规则支持修改和删除；已发布的规则支持修改阈值和告警信息。
6. 规则查看 展开左侧的目录树，可以根据表或字段查看该表下所有字段或当前字段配置的所有规则。

质量查询

当前版本质量规则在数据管理的质量模块发布之后需要到数据集成的去使用规则。数据集成的同步、加工任务会根据需求运行相关的检核规则，并生成质量检核任务。成功运行的质量检核任务会同步到数据管理的质量查询页面，该页面按照任务维度展示检核结果，可查看具体某次检核任务的质量检核结果。

1. 质量任务可以按照任务状态、数据源类型、数据源名称、所属项目、任务类型、业务时间、执行时间进行过滤，按具体的表或任务ID进行搜索。
2. 定位到的质量任务支持点击操作列的【详情】按钮查看详情。
3. 详情里展示了该任务跑的具体有哪些规则，以及规则相关的一些配置信息和每条规则在本次任务检核中检测到的异常数据，点击异常数据列的具体数字可查看具体的异常数据。
4. 每个任务的每条规则支持查看历史检核结果。点击操作列【历史结果】按钮弹出历史结果框。

质量告警

质量检核任务产生的告警信息统一在质量告警页面展现。可通过勾选【只看失败告警】查看发送失败的告警记录。也可通过告警级别、告警ID、规则名称、告警接收人搜索具体的告警信息。

数据权限管理

数据权限管理主要包括两方面：

- 租户或者数据管理员直接给项目或者用户添加表或者列级别的白名单
- 其他角色的人员申请项目或者个人的权限，数据管理员审批通过、驳回或者通过后回收权限。

权限申请

1. 除数据管理员角色以外的其他角色的人员（包括：项目管理员，开发人员，运维人员，业务人员等）进入数据管理模块，点击【元数据管理】下面的【数据目录】
2. 选中数据源（以hive为例），点击表名称，数据空间展示表详情。
3. 点击【权限申请】按钮，弹出权限申请的页签，选择申请对象，选择权限，选择申请对象，选择申请项目，点击【确认】按钮，权限申请成功

说明：

1. 申请类型为列权限时，申请对象只能为个人；申请类型为表权限时，申请对象可以为项目，个人或者项目和个人。
2. kafka类型数据源，没有列权限，只有表权限，不需要选择申请类型；kafka类型数据源的权限选择只有两个选项：insert和select。
3. 申请个人和项目的表权限，会同时生成两条申请记录，审批记录也是两条。

申请记录

1. 进入数据管理模块【数据权限管理】->【申请记录】查看提交的权限申请。
2. 点击【申请号】查看记录详情。
3. 点击【查看】查看流程状态。
4. 点击【已结束的申请】查看已结束的申请记录。

撤销申请

登录账户，点击选择【数据权限管理】->【申请记录】->【进行中的申请】，找到想要撤回的申请号，点击右侧【撤回】。

权限审批

1. 使用数据管理员账号登录进入数据管理模块，点击【数据权限管理】，展开下拉列表中点击【权限审批】
2. 根据申请号选择提交的申请，点击【流程】可查看提交申请的用户，备注中显示提交原因。
3. 点击【通过】（或者驳回）审批该申请。

说明：通过申请，如果申请的时候选择的是项目，该表的项目白名单会新增一条数据，项目管理员对该数据表有对应的权限；如果选择的是个人，用户白名单新增一条数据，申请人对该表有对应的权限。

数据授权

使用数据管理员或者租户的账号登录，进入数据管理模块，点击【数据权限管理】，展开的下拉框中点击【数据授权】，数据空间显示数据授权列表。

1. 白名单设置

(1) 选中需要授权的表，点击【白名单设置】按钮，弹出白名单设置标签，选择授权类型，选择权限，选择授权对象，授权项目和个人都支持多选，点击【确认】

说明：授权类型为表权限时，授权对象可以为项目，个人或者项目和个人；授权对象为列权限时，授权对象只可以为个人；个人和项目均支持多选和搜索。

(2) 授权成功后，点击【查看】按钮，可以看到授权的项目白名单和个人白名单。

说明：查看项目白名单和用户白名单，点击【新增白名单】按钮，可以给表增加白名单，操作类似白名单设置。

2. 公开设置

(1) 进入【数据权限管理】->【数据授权】，找到表名点击【公开设置】。

(2) 选择授权类型和公开权限，点击【确认】。

常见问题

1. 什么是数据管理？

数据管理是一种全方位的企业数据治理服务，提供云端传统关系型数据库、Nosql数据库、大数据平台数据仓库的统一管理，提供数据地图、数据质量、数据血缘分析能力，助力企业全面管理数据资产。

2. 为什么需要数据管理？

- 如果您的数据存在多个零散、异构的数据存储系统里，如oracle、HBase、Kafka、ES等，想要对所有数据的元数据及权限进行统一管理，您可能需要数据管理。
- 如果您现在的数据系统封闭，让用户发现数据、获取数据的成本很高，您可能需要数据管理。
- 如果您现在的数据质量不高，缺乏相关标准和编码规范，想要快速提升数据质量，进行数据溯源与影响分析，您可能需要数据管理。

3. 数据管理提供哪些功能？

数据管理产品最大的亮点在于提供统一的元数据管理。无论是关系型数据库、NoSql数据库、图数据库、消息队列还是对象存储，均能进行元数据映射管理，并托管数据权限。同时支持数据标准、数据字典、业务术语等业务元模型管理与关联关系配置，批量设置数据标签。在全域元数据的基础上数据管理提供数据地图、数据血缘等高级应用，并提供技术、业务多种数据质量检核规则，全面提升数据管控、数据发现、数据共享效率，保障数据质量提升和数据安全合规使用。

4. 如何开通数据管理？

目前仅支持线下开通，请联系商务人员，会为您定制解决方案

5. 数据管理不同版本有什么区别，适用场景是什么？

- 数据管理标准版功能较少，配置的资源相对较少，支持的并发量较小，适合小规模团队使用，适用于侧重数据开发，对数据管理无特别要求的场景。
- 专业版提供数据标签、数据目录、数据搜索、数据血缘等更多功能，能够快速定位数据，以及发现数据之间的关系。配置的物理资源相对较多，支持的并发数较多，适合于中等规模数据量开发管理场景。
- 企业版相对专业版来说提供业务元数据、元数据关联、数据地图、数据质量等更多高级功能，能够更好的满足企业数据治理的场景，全方位的了解企业数据资产。配置的物理资源是最高的，支持的并发数也最多，适合大型企业复杂的数据治理场景。

对于一些特大企业，如果以上版本配置的资源满足不了需求，可进行定制化资源配置，提供更多的服务资源。