

ALIBABA CLOUD

阿里云

智能媒体管理
产品简介

文档版本：20211229

 阿里云

法律声明

阿里云提醒您阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
Courier字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <i>Instance_ID</i>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

1.什么是智能媒体管理	05
2.产品优势	07
3.产品架构	10
4.应用场景	12

1.什么是智能媒体管理

阿里云智能媒体管理（Intelligent Media Management，简称IMM），场景化封装数据智能分析管理，为云上文档、图片、视频数据，提供一站式数据处理、分析、检索等管控体验。

产品概述

智能媒体管理针对不同行业的业务场景封装整合完整的处理能力，提供文档的格式转换及预览，图片的内容识别、人脸检测、二维码检测、人脸搜索等功能，适合媒资管理、智能网盘、社交应用、图库图床等开发者使用。智能媒体管理可以结合对象存储（OSS）、表格存储（Tablestore）为文档管理、图片社交分析等领域提供实用的场景化一站式解决方案。

功能特性

 **说明** 如需体验相应功能，请通过控制台操作。具体操作，请参见[场景体验](#)。

● 文档转换和预览

将文档相关的格式转换和预览整合，快速实现文档的智能管理能力。

功能	说明
格式转换	将PPTX、PPT、XLS、DOC、PDF、HTML、HTM等48种文档格式转换为JPG、PNG、PDF、TXT和VECTOR向量格式。更多信息，请参见 文档格式转换 。
文档预览	根据实际需要选择合适的方式进行文档预览。 <ul style="list-style-type: none">文档预览V1：将输入文档转换为VECTOR向量格式后，通过对接智能媒体管理提供的前端渲染引擎，实现更易用、功能更强大、定制化的文档预览效果。更多信息，请参见文档预览V1。文档预览V2：获取文档的预览地址和AccessToken后，无需指定iframe元素，JS文件会自动在自定义元素下生成iframe，并通过JS文件设置AccessToken即可快捷的实现文档预览。更多信息，请参见文档预览V2。

● 图片智能检测

将内容识别、人脸检测等AI功能整合，快速实现图片的智能管理能力。

功能	说明
内容识别	识别图片中场景、物体、事件等信息，获取到标签的元数据信息。更多信息，请参见 内容识别 。
人脸检测	检测图片中的人脸以及人的年龄、性别、心情等，获取到人脸的元数据信息。更多信息，请参见 人脸检测 。
二维码检测	检测图中的二维码以及二维码中存储的内容，详情请参见 二维码识别 。
商标检测	检测图片中的户外用品、奢侈品牌、数码3C、汽车、互联网企业、体育用品等领域企业的商标。更多信息，请参见 商标检测 。
人体检测	检测图片中的人体区域和置信度。更多信息，请参见 人体检测 。

功能	说明
人脸搜索	搜索与指定图片最相似的前N张图片，结果按相似度降序排列。更多信息，请参见 人脸搜索 。
人脸对比	比较两张图片中分别最大的两个人脸的相似度。更多信息，请参见 人脸对比 。
图片盲水印	为图片添加图片或文字类型的盲水印。盲水印添加后，在图片中不能直接看到该水印，但是可以通过使用智能媒体管理的解析图片盲水印功能恢复图中隐藏的水印。更多信息，请参见 图片盲水印 。

应用场景

智能媒体管理可应用于文档管理、图片社交分析、家庭设备数据存储等场景。更多信息，请参见[应用场景](#)。

使用产品

- 智能媒体管理控制台

使用Web服务页面方便您管理智能媒体管理，您可以登录[智能媒体管理控制台](#)操作项目，以及体验不同功能的效果。更多信息，请参见[快速入门](#)。

- 智能媒体管理SDK

使用SDK方便您灵活使用智能媒体管理，SDK基于服务API实现，且提供和服务API同样的能力。更多信息，请参见[SDK参考](#)和[API手册](#)。

- OSS控制台

为OSS Bucket绑定智能媒体管理后，通过OSS控制台可以使用智能媒体管理的功能，例如文档预览、人脸识别等。更多信息，请参见[快速开始](#)。

产品定价

请参见[计费说明](#)。

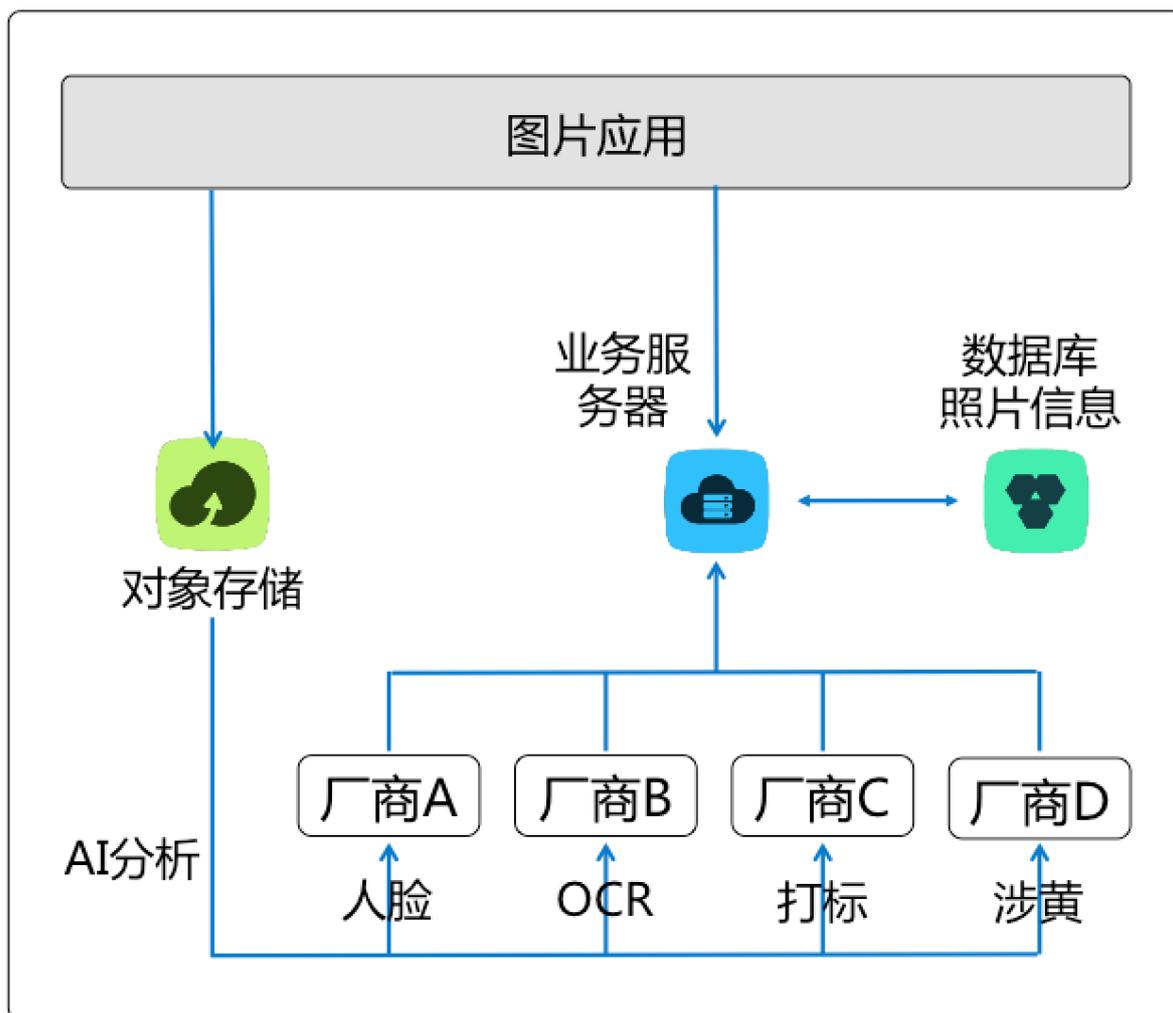
2. 产品优势

基于对象存储OSS中的文档、图片、视频等多媒体数据，智能媒体管理典型的应用场景有图片应用、网盘等。

场景痛点分析

- 图片应用

对于图片应用，首先上传图片、视频到对象存储中，随着业务的不断发展和法规政策的监管需求，会增加鉴黄检测、标签检测、人脸检测、OCR识别等AI分析功能，如下图所示。



通常，图片应用会选择采用业务服务器安装不同厂家的AI分析能力，通过这些功能得到关键的元数据信息，然后保存到数据库中，从而更好的支撑元数据检索，支撑业务和监管的需求。采用此方案存在如下问题：

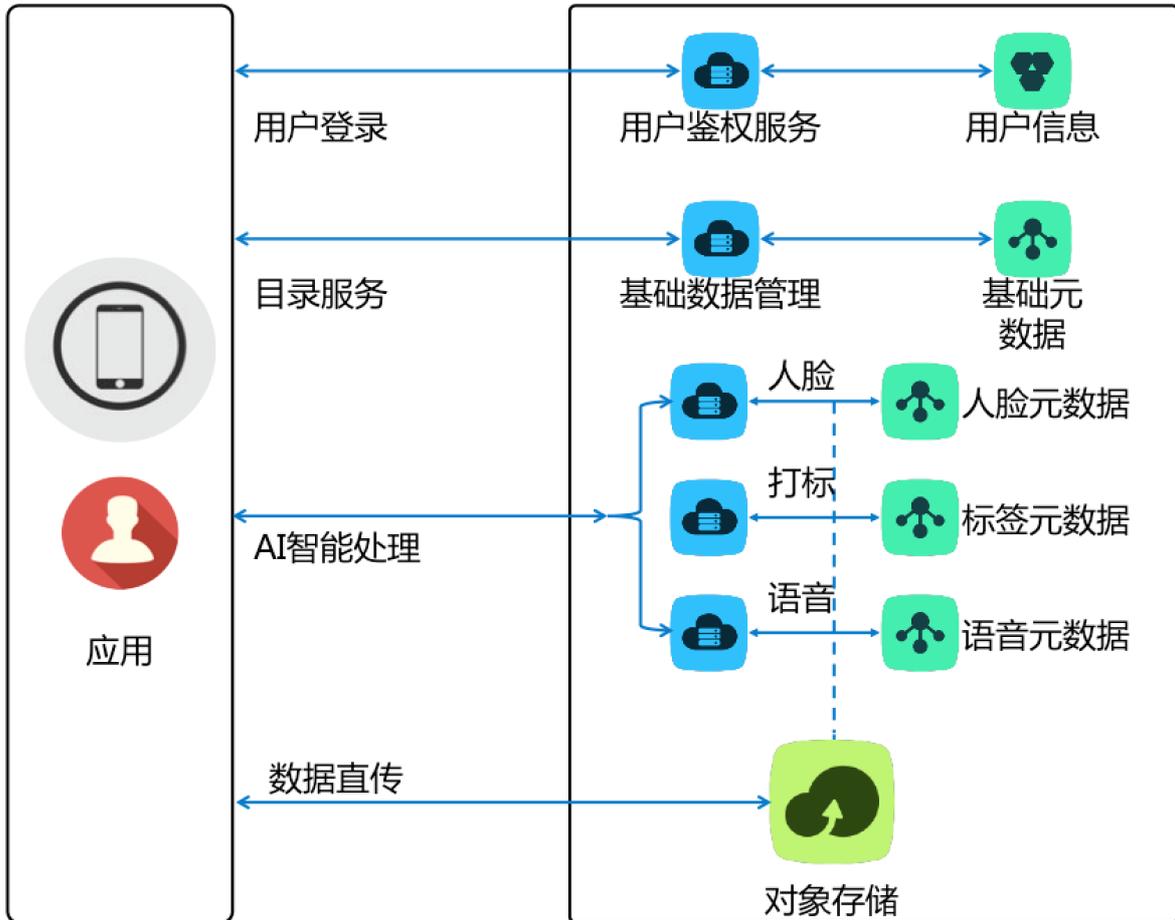
- 接口不统一
因为有多厂家的选型，需要考虑不同厂家接口的兼容性。
- 资源浪费
同一张图片会被多次读取，甚至是传输到外部网络，浪费网络带宽。

○ 无存量数据的低成本批量处理方案

采用厂商的同步处理价格高昂，需要提供存量数据的低成本批量处理方案，接受异步接口返回检测结果。例如针对存量的OSS Bucket中的所有图片，进行打标处理。

● 网盘

对于网盘应用，通常需要用户登录、目录服务、数据直传OSS、AI智能处理等功能。通过后端使用服务器和数据库，其架构如下图所示。



为了支撑网盘的数据管理，通常需要提供各类元数据管理，特别是在AI智能处理场景下，需要定义相关的存储格式，处理数据库的异常，带来了较大的开发难度。采用此方案存在如下问题：

○ 元数据表设计难度大

针对不同的元数据，需要分类设计各类表结构存储，存在技术门槛。

○ 多维度元数据管理有挑战

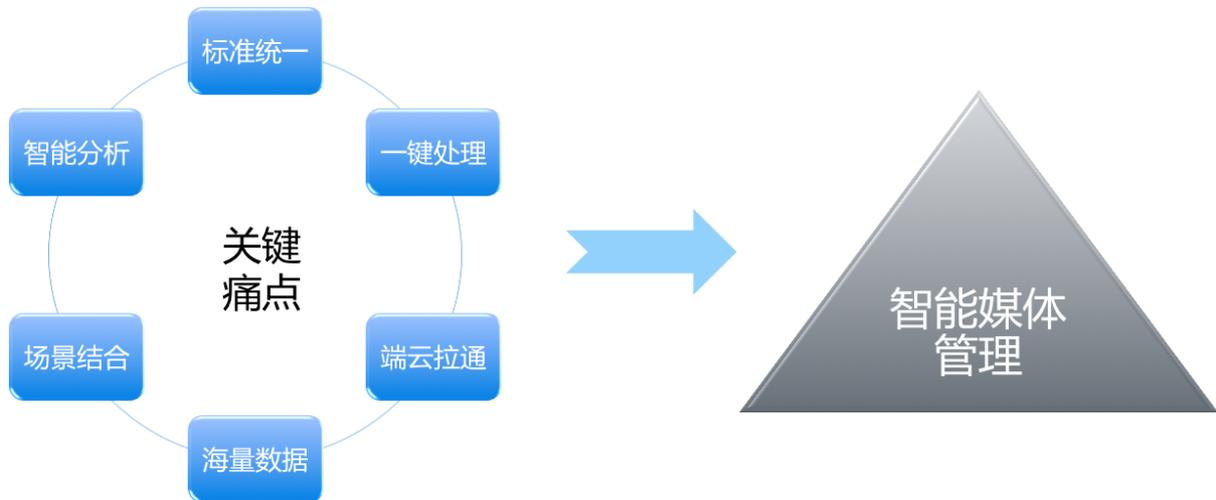
需要组合多种元数据，进行关联查询处理，存在设计挑战。

○ 维护元数据的一致性难题

解决元数据在异常场景下的恢复处理，是系统级难题。

智能媒体管理产品优势

智能媒体管理围绕海量数据、端云拉通、标准统一、智能分析、场景结合、一键处理6个关键点进行设计，提供基于场景的AI智能处理方案，如下图所示。



通过针对性的架构设计，智能媒体管理具有如下优势：

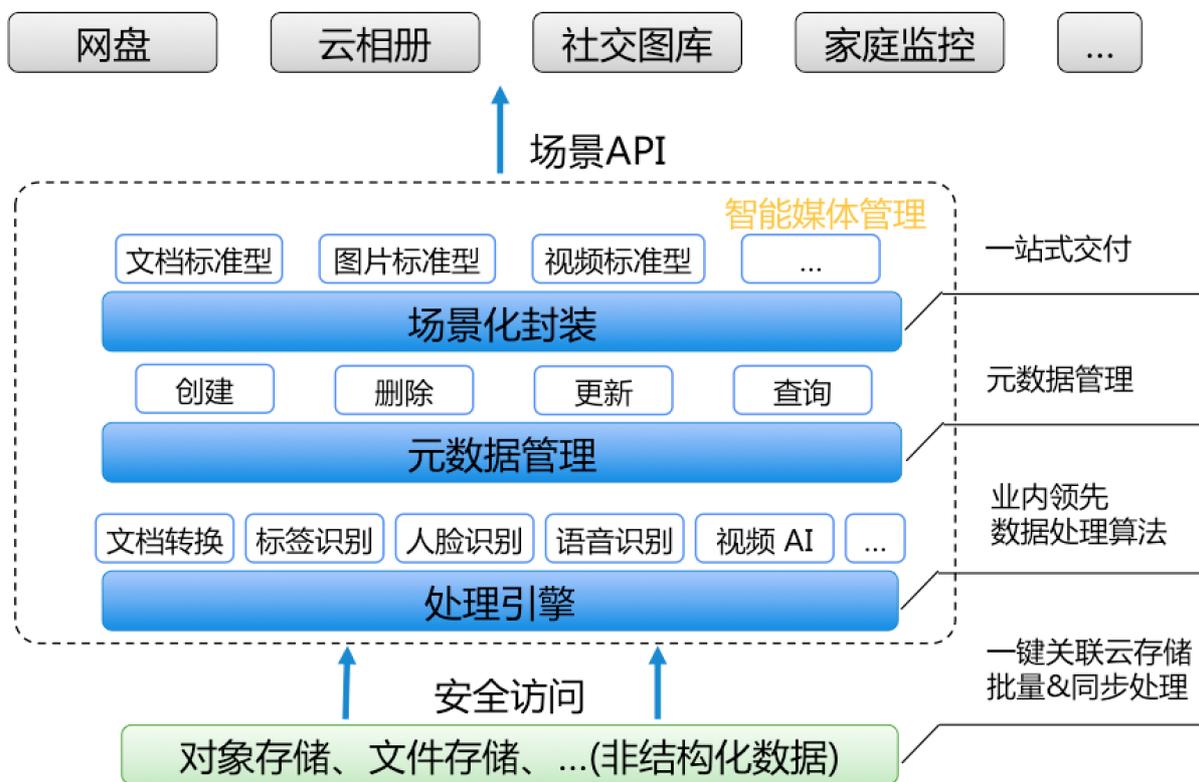
- 存储数据无缝贴合
与对象存储OSS直接关联，自动处理云上数据。
- 丰富数据处理
结合业界领先的识别、处理能力，为应用处理提供丰富的功能支持。
- 简化运维
提供Serverless化服务，无需关心业务运维。
- 场景化一站式解决方案
面向场景的构建快捷的元数据管理，快速实现应用。

3.产品架构

智能媒体管理围绕海量数据、端云拉通、标准统一、智能分析、场景结合、一键处理六个关键点采用分层架构进行设计，分层架构包含处理引擎、元数据管理、场景化封装三层。

智能媒体管理采用分层架构进行设计，分层架构包含处理引擎、元数据管理、场景化封装三层，且存在依赖的上下文，如下图所示。

- 对下依赖对象存储、文件存储等阿里云存储服务，通过安全的机制访问阿里云存储中的非结构化数据（例如图片、视频），提取有价值的信息。
- 对上基于场景理解进行封装，支撑网盘、云相册、社交图库、家庭监控等图片和视频应用场景，为应用提供新的价值。



处理引擎层

基于阿里云存储提供就近构建计算框架，该框架支持批量异步处理、准实时同步处理，在一键关联阿里云存储（例如指定OSS Bucket的目录前缀、指定OSS Bucket的某个对象）后，实现快速的自动数据处理，通过整合业界领先的数据处理算法，目前处理引擎提供如下功能：

- 文档格式转换

支持将包括OFFICE在内的48种格式的文档转换为JPG、PNG、PDF、TXT、VECTOR 5种格式，可用于网盘文档浏览等场景。
- 内容识别

识别图片中场景、物体、事件等信息，实现图片的自动打标，可用于图片内容审核、图片检索等场景。
- 人脸检测

检测图片中的人脸以及人的年龄、性别、心情等，可用于相册分类等场景。
- 二维码检测

检测图片中的二维码以及二维码中存储的内容，可以判断图片中是否含有二维码信息，输出二维码包含的信息，可用于图像内容审核等场景。

- 商标检测

检测图片中的商标及商标名称，可以判断图片中是否含有商标信息，输出商标名称，可用于图片内容审核等场景。

- 人体检测

检测图片中的人体区域和置信度，可用于异常行为检测等场景。

- 人脸搜索

搜索与指定图片最相似的前N张图片，结果按相似度降序排列，可用于会员管理、相册分类、目标人员搜索等场景。

- 人脸对比

比较两张图片中分别最大的两个人脸的相似度，可用于身份识别验证等场景。

- 图片盲水印

为图片添加图片或文字类型的盲水印。盲水印添加后，在图片中不能直接看到该水印，但是可以通过使用智能媒体管理的解析图片盲水印功能恢复图中隐藏的水印，可用于图片版权追溯等场景。

元数据管理层

基于处理引擎提供的功能，通过对场景的深入理解和梳理，智能媒体管理封装了场景的元数据设计，对外提供场景的元数据访问接口，简化场景应用的设计难度、无需关注元数据索引数据库的运维工作，目前支持的元数据索引如下：

- 人脸聚类索引

构建元数据集合，然后调用人脸分组的索引接口分析图片，将得到的元数据加入到该元数据集合中，从而可以得到该集合中相似的人脸。通过该索引，可以快速的支撑网盘的人脸相册、家庭监控的陌生人检测、新零售的顾客管理等场景。

- 标签分组索引

构建元数据集合，然后调用标签分组的索引接口分析图片，将得到的元数据加入到该元数据集合中，从而可以根据标签搜索图片。通过该索引，可以快速的支撑网盘的场景相册、家庭监控的宠物跟踪、低俗图片等标签的搜索。

场景封装层

通过阿里云对场景的支撑，把处理引擎层和元数据管理层的功能进行包装，并按照资源包方式提供出来，从而简化使用，方便应用快速的接入，实现AI和场景的紧密结合，目前支持的场景实例如下：

- 文档标准型

将文档相关的格式转换和预览整合，快速实现文档的智能管理能力。

- 图片标准型

将内容识别、人脸检测等AI功能整合，快速实现图片的智能管理能力。

4.应用场景

智能媒体管理为云存储上的文档、图片、视频等非结构化数据提供智能分析处理能力，并通过提取关键元数据构建索引，实现非结构化数据的查询能力，可应用于文档管理、图片社交分析、家庭设备数据存储等场景。

文档管理场景

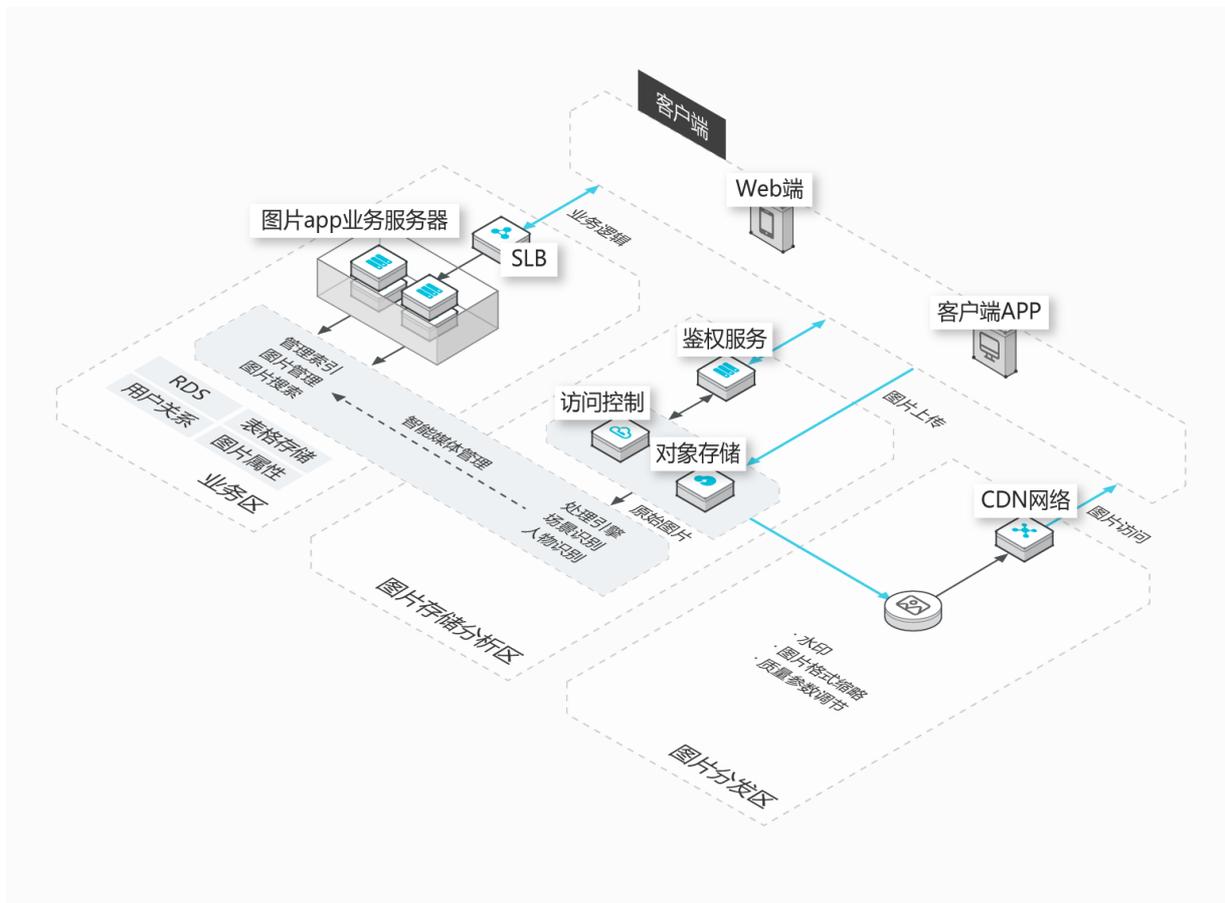
在网盘、邮箱、文档管理等应用中，使用智能媒体管理提供的文档标准型项目，可以快速实现如下场景：

- 文档预览

使用格式转换功能，将48种常见文档格式转换为JPEG、PNG、PDF、TXT、VECTOR 5种目标类型，然后基于格式转换的结果，结合前端渲染引擎实现在PC和移动设备上的文档预览。
- 全文检索

提取DOC文档中的文字并按页输出，基于文字提取的结果，按文档对应页构建全文索引，实现按页的全文检索。

图片社交分析场景



在图片社交、电商网站、图库等应用中，使用智能媒体管理提供的图片标准型项目，可以快速实现如下场景：

- 图片内容审核

使用内容识别功能，有效识别图片中的低俗和违禁内容，例如色情、暴力恐怖、违法违规等，满足监管部门对内容日益严格的监管要求，规避运营风险。

使用二维码检测功能，判断图片中是否含有二维码以及输出二维码的内容。

使用商标检测功能，判断图片中是否含有商标以及输出商标名称。

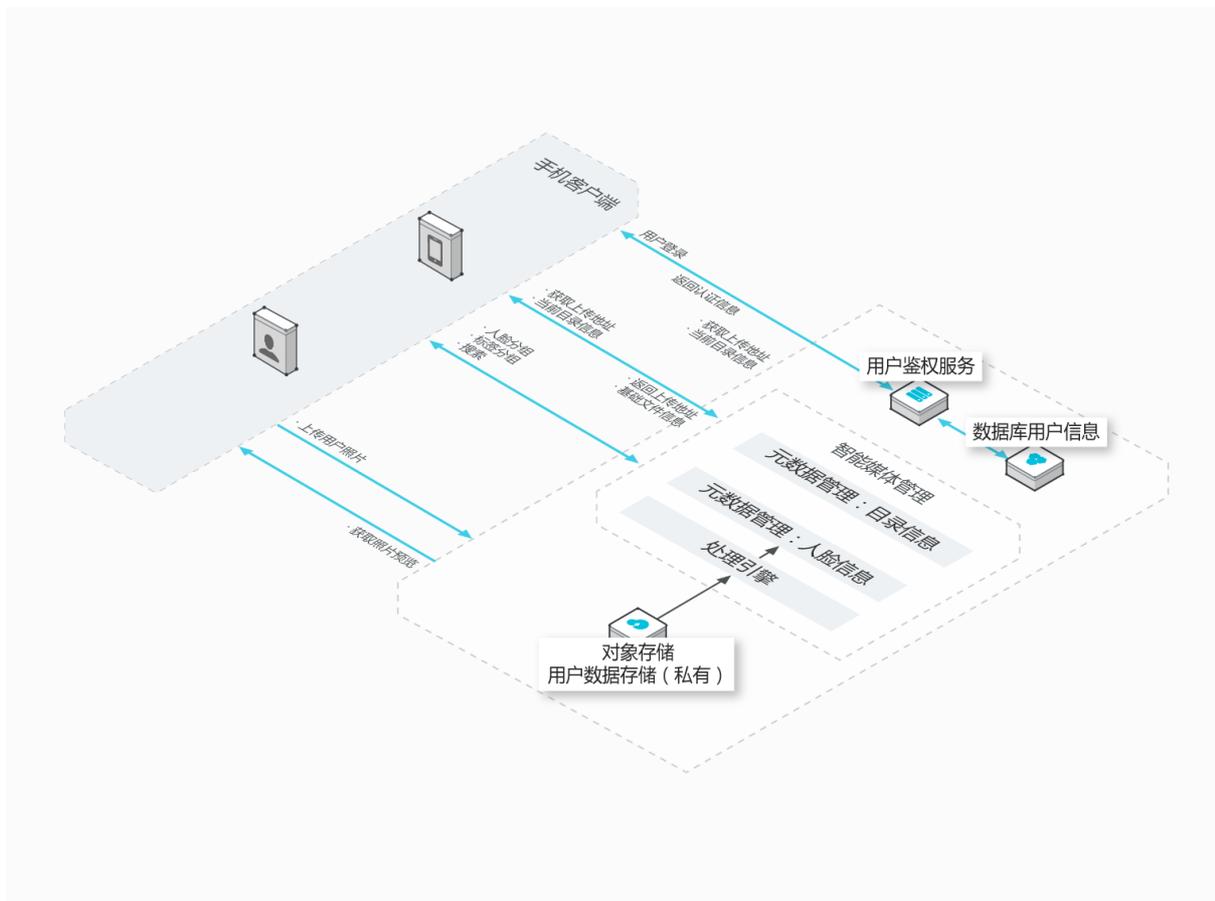
• 图片分类和检索

使用内容识别功能，为图片打标分组后，通过搜索关键字在图库、素材网站、网盘、智能相册等应用中搜索匹配的图片。

• 图片版权追溯

使用图片盲水印功能，为图片添加图片或文字类型的盲水印。盲水印添加后，在图片中不能直接看到该水印，但是可以通过使用智能媒体管理的解析图片盲水印功能恢复图中隐藏的水印。

家庭设备数据存储场景



家庭设备（例如摄像头）上传家庭图片、监控视频到OSS后，智能媒体管理分析OSS中保存的多个家庭设备上传图片、视频的人脸信息，得到人脸检测、分组等元数据信息，利用元数据信息可以方便的实现人脸相册、陌生人检测等功能。

同时，还能够将人脸的元数据信息，反向同步回各个家庭设备，您可以通过家庭设备即使在网络OFFLINE时也能够享受智能媒体管理的智能人脸AI能力。

通过此方式，实现了多设备的关联，端和云的互动，让智能媒体管理的AI能力更加的普惠，从而带来如下的优势：

• 多设备智能管理

通过多设备的图片、视频数据处理，可以在云上实现组合分析和智能管理。

• 端和云的智能协作

用极少的带宽得到图片、视频的AI元数据，实现设备端快速享受云端的AI智能。

在家庭监控、智能相册等应用中，使用智能媒体管理提供的图片标准型项目，可以快速实现如下场景：

- 人脸分类和检索

使用人脸搜索功能，在图库中搜索与指定人脸最相似的前N张图片，结果按相似度降序排列。

- 身份识别验证

使用人脸对比功能，比较两张图片中分别最大的两个人脸的相似度，可以检测陌生人或验证两个人是否为同一个人。