

音频内容安全 快速入门



腾讯云

【 版权声明 】

©2013–2025 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

文档目录

快速入门

音频内容安全服务

音频 AI 生成识别服务

快速入门

音频内容安全服务

最近更新时间：2025-09-15 16:04:42

步骤一：登录注册

注册腾讯云账号，并完成 [实名认证](#)，详情请参见 [账号注册教程](#)。

❗ 说明

如果您已拥有腾讯云账号，可略过此步骤。

步骤二：创建子账号（可选）

在 [注册腾讯云](#) 账号时，所生成的账号为主账号，拥有该主账号下所有云资源的管理权限。如果您需要其他用户来协助您一起管理账号下的云资源，可以通过使用访问管理功能创建子账号，并为其绑定权限策略。该功能可以创建一个或多个子账号提供给团队成员，详情请参见 [新建子用户](#)。

子账号刚创建时，默认是没有任何权限的，需要主账号或管理员为其绑定策略，才能使子账号拥有某些云资源的操作权限。您可以为子账号配置内容安全 CAM 授权，详情请参见 [CAM 授权指引](#)。

步骤三：开通服务

前提条件

- 已获取腾讯云主账号或子账号。
- 登录账号已拥有音频内容安全服务的读写权限。
- 账户余额需大于0元，如账户余额小于0元，请完成 [账户充值](#) 后前往开通服务。

试用说明

- **新用户免费试用权益：**新用户首次开通服务后，可获得10小时的免费试用包（不含“音频 AI 生成识别”能力），有效期为15天。

❗ 说明：

首次开通领取试用包仅为获取测试权益，非正式服务开通。

- **后续计费与服务开通规则：**
 - 若试用包到期未用完、提前用尽，需 [购买正式服务](#) 或开通后付费服务，否则无法正常调用接口。
 - 后付费模式将按照“点播 - 音频内容安全”的按量后付费功能结算，结算方式为日结，单价标准为1元 / 小时。

常见问题提示:

- 老用户（曾领取过试用包）接入时，若未开通正式服务，会提示“服务未开通”。
 - 新用户领取试用包后，若15天有效期内未使用或未开通正式服务，试用包过期后接入也会提示“服务未开通”。
- **解决方案:** 若遇到“服务未开通”提示，需 [购买正式服务](#) 或开通按量后付费功能，完成后即可正常调通接口。

内容安全

概览

数据统计

内容审核明细查询

名单管理

应用管理

服务体验

生成识别

资源包管理

云产品内容审核

资源包管理

图片 文本 音频 视频 直播音频 直播视频

① 1. 新开通服务的用户，可获得识别10小时音频的免费试用额度，有效期为15天

2. 试用套餐包过期或消耗完毕后，请购买正式套餐包，否则将默认开启后付费模式

3. 如不再继续使用产品服务，停止调用服务接口，则不会产生对应的服务费用

4. 如需了解费用详情，请前往腾讯云费用中心[腾讯云费用中心](#)

购买套餐包 开通后付费

付费模式: 预付费 套餐类型: 全部套餐

资源包类型	来源	套餐状态
T-Sec-腾讯安全天御-音频内容安全（AMS）/试用包	赠送	

共 1 页

操作步骤

1. 登录 [内容安全控制台](#)，单击立即开通。

腾讯云 控制台

支持通过实例ID、IP、名称等搜索资源 快捷操作 帮助文档 备案 工具 各版支持 试用 费用

内容安全

概览

数据统计

明细查询

名单管理

应用管理

服务体验

套餐包管理

云产品内容审核

您好，欢迎使用内容安全应用管理

支持在内容安全控制台中创建多个应用，以应用维度查询审核数据用量、审核情况等，同时支持在应用下添加多个业务场景，为不同的业务场景配置具体的审核服务。

① 1. 首次使用内容安全服务将赠送试用套餐包，套餐包详情及使用情况可在套餐包管理中查看

2. 每个审核服务没有可用套餐包或未开通后付费时，无法在场景中添加服务，购买套餐包或开通后付费后，方可正常配置审核服务

立即开通

2. 选择音频内容安全服务，单击确定完成开通。

审核服务管理



- 1、请勾选应用涉及的全部审核服务类型，后续可在各应用场景管理中使用，若未勾选则不可用
- 2、若未购买服务时，勾选将开启后付费模式
- 3、当勾选的内容包含音频内容安全、视频内容安全、视频流内容安全、音频流内容安全之一的时候，天御将会为您创建COS桶存储对应的送审内容，为了保证审核服务不出错，请勿删除对应的COS桶

- ☐ 文本内容安全
 ☐ 图片内容安全
 ☒ 音频内容安全
- ☐ 视频内容安全
 ☐ 视频流内容安全
 ☐ 音频流内容安全

▶ 附加功能

确定

取消

3. 在开通音频内容安全服务时，音频内容安全需要您为内容安全创建服务角色，并授权调用其他云服务接口，单击**同意授权**，完成身份校验后即可完成COS资源授权，正常使用或测试音频内容安全服务。

服务授权



执行本服务相关操作时将用到其他云服务功能。

需要您为 **内容安全** 创建服务角色，并授权调用其他云服务的接口。相关信息如下：

角色名称	VM_QCSRole (服务角色)
角色描述	当前角色为内容安全（VM）服务角色，该角色将在已关联策略的权限范围内访问您的其他云服务资源。
权限策略	预设策略 QcloudAccessForVMRole ⓘ

同意授权

取消

步骤四：配置策略

建议您使用配置任务策略，可根据业务需求配置识别策略，用于个性化服务体验。

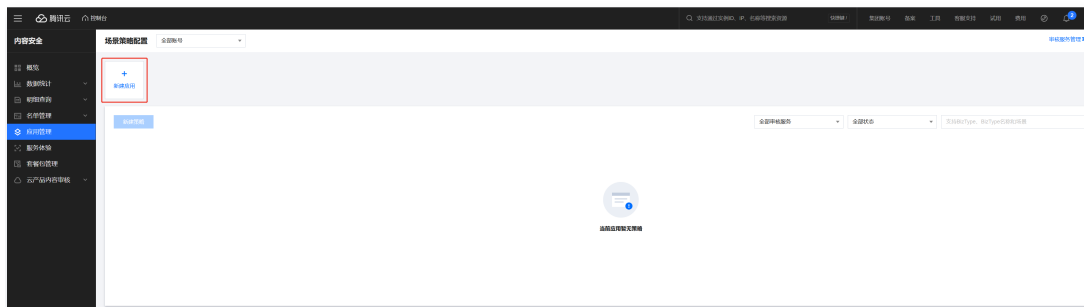


说明

- 腾讯云内容安全服务已预设有默认策略，如使用默认策略，可略过此步骤。
- 默认策略为天御多行业模型沉淀的策略配置，适用于大部分的内容安全需求。
- 如您想要配置更符合业务场景的策略，可按照以下步骤创建自定义策略实现个性化服务。

操作步骤

1. 登录 [内容安全控制台](#)，在概览页单击应用管理 > 新建应用，新建需要配置音频内容安全管理的应用。



2. 在应用管理页面，单击新建应用，输入应用名称，单击保存完成应用创建。

新建应用

* 应用名称

测试_音频

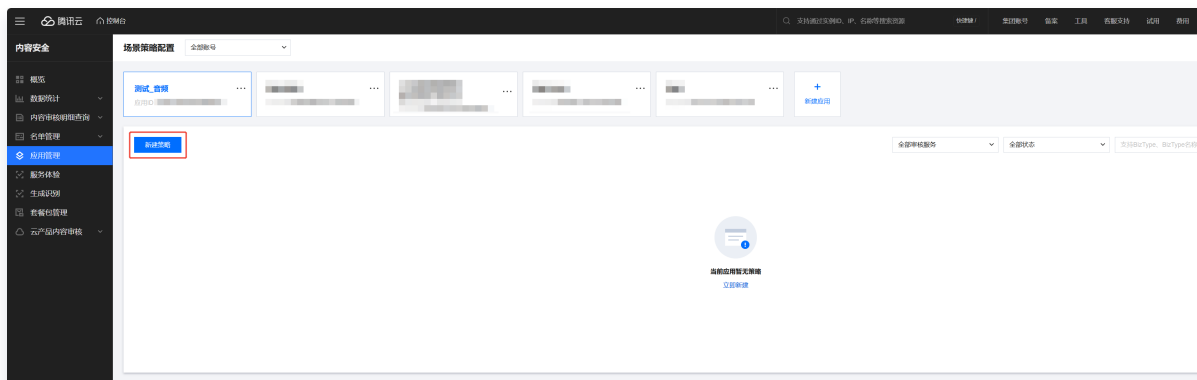
应用说明

非必填项

保存

取消

3. 在应用管理页面，单击新建策略。



4. 在新建策略页面，选择音频内容安全，输入Biztype、Biztype名称、场景，单击下一步进入策略模板配置。

新建策略

×

1 策略基础信息

>

2 策略模板配置

* 管理审核服务

☐ 文本内容安全

☐ 图片内容安全

☒ 音频内容安全

☐ 视频内容安全

☐ 视频流内容安全

☐ 音频流内容安全

管理审核服务

* Biztype

text_ams

BizType 是策略集。一个业务场景要对应到一个BizType。
例:biztype='nickname_input', biztype名称='昵称审核', 场景='个人主页'。

* Biztype名称

测试-音频

场景

非必填项，用于对BizType的分类

下一步

取消

参数名称	描述
场景名称	场景的文字描述，可使用中文、英文、数字及下划线组合，长度不超过20个字符
关联审核服务	选择场景需要关联的审核服务

5. 在策略模板配置页面，根据具体业务选择对应的行业、业务类型、业务场景，并单击确定引用模板保存，即可完成策略创建。

策略模板

×

ⓘ

• 点击“确认引用模板”后，策略配置页面会刷新成对应的标签和词库模板配置；

• 更多审核策略模板，请联系您熟悉的架构师/售后/运营人员/提工单进行配置；

已选策略：工具 / 私域会议 / im消息

系统推荐模板

复用已有策略

行业	业务类型	业务场景
社交	语聊房	动态广场
通用		
工具		

确认引用模板

取消

音频内容安全策略模板介绍		
行业	业务类型	业务场景
社交	语聊房	动态广场
工具	私域会议	im 消息
通用	通用	通用

ⓘ 说明：

• 单击确认引用模板后，策略配置页面会刷新成对应的标签和词库模板配置。

• 如未找到业务分类，请选择通用模板。

• 如需更多审核策略模板，请联系您熟悉的架构师、售后、运营人员或 提交工单 进行配置。

6. 策略创建完成，在策略详情页面会显示当前系统使用的策略，如需自定义更改，单击**策略配置**，进入策略详情，选择您想开启的识别策略。

点播音频审核策略详情

场景名称：测试-音频

×

1 策略基础配置

>

2 关联名单库

策略基础信息

* 策略名称

测试-音频

* bizType名称 ①

text_ams1

识别策略配置

策略模板

①

• 请基于业务需求选择需要的标签分类，选中对应分类代表需要进行识别打击该内容

• 三级标签是模型用于过滤，系统命中后会返回二级标签

音频识别

音频文本识别

☒

0

☐

0

恢复默认

标签名称	标签说明	标签级别	检测状态
No data			

点播音频文件截帧配置

音频切片时长

15

30

45

60

-

30

+

审核结果回调配置

☐ 违规回调 ①

☐ 全量回调 ①

下一步

取消

参数名称	描述
识别策略配置	根据业务需求，选择是否需要识别不同类型的识别内容。

点播音频文件截帧配置	音频切片时长：设置音频切片的时间长度，可选15s、30s、60s，也可自行填写设置，默认为30s。 审核结果回调配置： - 违规回调：仅返回违规审核结果。 - 全量回调：返回全部审核结果。
关联自定义库	选择是否需要使用自定义词库用于内容识别，如暂无自定义库，可跳过，或保存当前策略后前往 步骤五：配置自定义词库 。

7. 完成识别策略配置后，单击下一步，选择是否需要关联名单库用于内容识别，如暂无关联名单库，可跳过，或保存当前策略后前往。

点播音频审核策略详情

场景名称：测试-音频

策略基础配置

关联名单库

关联自定义词库

选择词库

请输入关键词

词库名称

处理类型

通过
(Pass)

疑似
(Review)

违规
(Block)

已选择 (0)

词库名称

处理类型

上一步

确定

取消

8. 完成后单击确认，完成策略配置。

版权所有：腾讯云计算（北京）有限责任公司

第11 共21页

步骤五：配置词库（可选）

前提条件

- 已获取腾讯云主账号或子账号。
- 登录账号已开通音频内容安全服务。

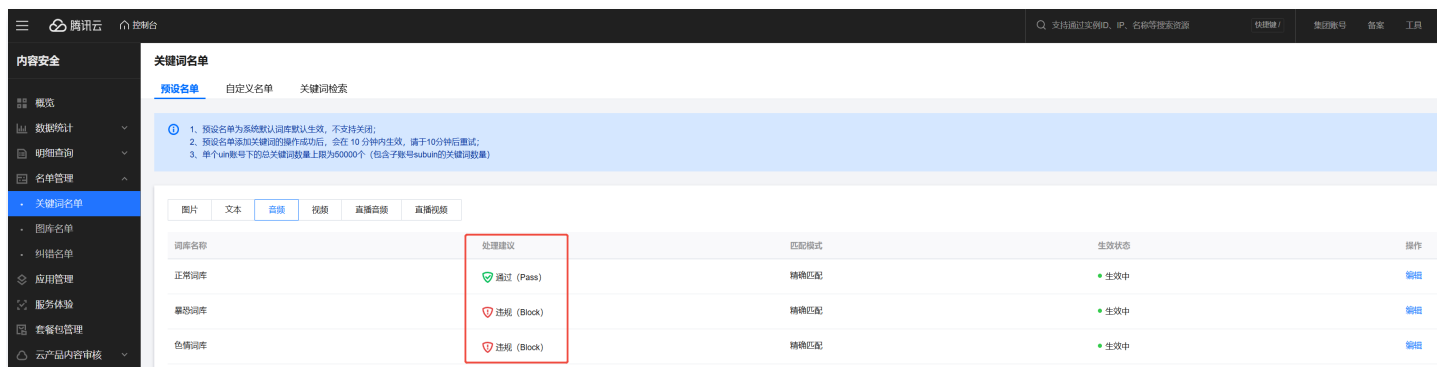
预设词库

说明：
如无需添加预设词库关键词，可略过此步骤。

1. 登录 [内容安全控制台](#)，在左侧导航栏中，单击**名单管理** > **关键词名单**。
2. 在预设名单页面，产品类型选择**音频**。



3. 在预设名单页面，支持查看系统预设的词汇，包括：正常词库、违法词库、广告词库等。



说明：

预设词库的不同颜色代表不同的屏蔽逻辑，绿色代表通过，红色代表违规。

4. 在预设名单页面，单击**编辑**，可以添加关键词。

❗ **说明：**

预设词库默认开启，不支持关闭、删除操作。

5. 在编辑词库页面，单击**添加关键词**，输入关键词，单击**确定**，即可保存关键词至当前词库。

← 管理词库

通过

正常

完成添加的关键词10分钟内生效

匹配模式 精确匹配 状态 ● 生效中

添加关键词

批量删除

☐ 关键词

样本ID

机审标签

关键词类型

添加关键词

×

关键词类型 ☒ 普通词 ☐ 组合词 ☐ 英文

添加关键词 *

多个以回车换行分隔，一次提交不超过2000个

关键词备注

请输入关键词的备注说明，最多可输入100个字

0 / 100

1、每个关键词以换行来确定，单个关键词长度20个汉字以内

2、支持以复制粘贴的方式批量导入，单次最多添加关键词2000个，词与词之间用换行符分隔

3、单个uin账号下的总关键词数量上限为50000个（包含子账号subuin的关键词数量）

4、如需添加组合词，选择关键词类型：组合词，使用&符号对两个或者多个词进行连接，添加格式为“关键词1&关键词2...”，其中关键词只支持纯中文，这些关键词同时出现时则会命中

确定

取消

参数名称	描述
关键词	- 每个关键词以换行来确定，单个关键词长度20 个汉字以内。 - 支持以复制粘贴的方式批量导入，单次最多添加关键词2000个，词与词之间用换行符分隔。 - 单个词库的关键词添加个数上限为10000个。 - 单个 uin 账号下的总关键词数量上限为50000个（包含子账号 subuin 的关键词数量）。 - 如需添加组合同，选择关键词类型：组合同，使用&符号对两个或者多个词进行连接，添加格式为“关键词1&关键词2...”，其中关键词只支持纯中文，这些关键词同时出现时则会命中。

自定义词库

自定义词库用于添加个性化的识别词库，识别图片上是否含有违规文本。

❗ 说明

如无需配置自定义词库，可略过此步骤。

- 登录 [内容安全控制台](#)，在左侧导航栏中，选择**名单管理> 关键词名单 > 自定义名单**。
- 在自定义名单页面，单击**新建词库**，填写文本库名称及按业务需求添加关键词。

新建关键词库

词库名称

请输入关键词库名称

可使用中文、英文、数字及下划线组合，长度32个字符内

识别结果

☒ 违规 (Block) ☐ 疑似 (Review) ☐ 放过 (Pass)

匹配模式

☒ 模糊匹配 ☐ 精确匹配

匹配模式选择后，不可变更。匹配模式的逻辑分别为：
精确匹配：对输入的文本和关键词做直接的一致性匹配，例如：配置类关键词“香蕉”，则“我想吃香蕉”会命中，但“我想吃香x蕉”则不命中
模糊匹配：用于检测可能发生变体的词，当前支持拆分字、形似字、音似字、简繁体、大小写、大写数字六种形式。

确定

取消

参数名称	描述
文本库名称	词库的文字描述，可使用中文、英文、数字及下划线组合，长度不超过32个字符

处理建议	可选择违规或疑似 - 违规：确认为不良信息。 - 疑似：可能为不良信息，需要人工识别。 - 放过：业务方认为的正常信息。
匹配模式	可选择精确匹配或模糊匹配 - 精确匹配：对输入文本进行匹配，匹配对象需完全一致 - 模糊匹配：用于检测可能发生变体的词，当前支持拆分字、形似字、音似字、简繁体、大小写、大写数字等形式的相似词进行匹配

3. 单击**确定**，即可创建自定义词库。
4. 在自定义词库页面下方的列表中，将显示刚创建的词库。在创建词库弹窗，填写自定义库的相关信息，结合业务需求配置自定义库。

腾讯云

控制台

关键词名单

预设名单

自定义名单

关键词检索

1. 对于名单内容的操作成功后，会在 10 分钟内生效

2. 单个un编号下的总关键词数量上限为50000个（包含子编号subun的关键词数量）

新建词库

输入词库名

词库名称	识别结果 Y	匹配模式 Y	库信息修改时间	操作
测试2	通过 (Pass)	模糊匹配	2025-03-26 11:51:15	关联场景 管理 删除 导出
测试1	疑似 (Review)	模糊匹配	2025-03-26 11:51:08	关联场景 管理 删除 导出
测试	违规 (Block)	模糊匹配	2025-03-26 11:50:19	关联场景 管理 删除 导出

!

说明：

自定义词库的不同颜色代表不同的屏蔽逻辑，红色代表违规，橙色代表疑似，绿色代表正常。

5. 在自定义词库页面，单击**管理**，选择添加关键词，单击**确定**，即可保存关键词至当前词库。

添加关键词

×

机审标签 * 正常

关键词类型 ☒ 普通词 ☐ 组词 ☐ 英文

添加关键词

输入框添加

文件上传

输入框一次最多添加2000个词。若超过该数量，则请使用文件上传的方式。

请输入需要添加的关键词，每个关键词以换行来确定，单个关键词长度20个汉字以内，添加多个关键词以回车换行分隔，一次提交不超过2000个

关键词备注

请输入关键词的备注说明，最多可输入100个字

0 / 100

确定

取消

参数名称	描述
识别类型	识别模型对应的违规类型。
关键词	- 每个关键词以换行来确定，单个关键词长度20 个汉字以内。 - 支持以复制粘贴的方式批量导入，单次最多添加关键词2000个，词与词之间用换行符分隔。 - 单个词库的关键词添加个数上限为10000个。 - 单个 uin 账号下的总关键词数量上限为50000个（包含子账号 subuin 的关键词数量）。 - 如需添加组词，选择关键词类型：组词，使用&符号对两个或者多个词进行连接，添加格式为“关键词1&关键词2...”，其中关键词只支持纯中文，这些关键词同时出现时会命中。

ⓘ 说明：

自定义词库配置完成后，可以在 [步骤四：配置策略](#) 时将自定义词库与策略关联使用。

步骤六：体验服务

完成以上步骤后，您可以选择创建识别策略，通过立即识别音频，体验音频内容安全服务。

前提条件

- 已获取腾讯云主账号或子账号。
- 登录账号已开通音频内容安全服务。

操作步骤

1. 登录 [内容安全控制台](#)，在左侧导航栏中，选择[服务体验](#)> [音频](#)。
2. 在服务体验页面，选择所需的应用策略，通过如下任意一种方式指定识别音频。
 - 从本地上传：单击[点击上传](#)，选择需要上传的本地文件。
 - 通过 URL 上传：在输入框内填入识别内容 URL 或 COS 存储地址。

❗ 说明：

其中音频文件需满足以下要求：

- 音频文件大小：50 MB以内。
- 音频文件时长：3 分钟以内。
- 音频码率类型：128 Kbps – 256 Kbps。
- 音频文件支持格式：wav、mp3、aac、flac、amr、3gp、m4a、wma、ogg、ape。



3. 单击**立即审核**，待系统识别完成后在页面下方会显示该音频的识别结果。

步骤七：接入服务

如果您需要接入音频内容安全服务，需通过调用 API 接口完成接入，详情请参考 [调用方式](#)。

❗ 说明

在调用 API 接口之前，您需要获取腾讯云 API 访问密钥。腾讯云通过 SecretId 和 SecretKey 对开发者的身份和权限进行校验，您可以参考以下步骤获取腾讯云 API 访问密钥：

进入 [云 API 密钥管理](#) 页面，在左侧导航栏选择**访问密钥** > **API 密钥管理**，单击**新建密钥**创建密钥，保存 SecretId 和 SecretKey 供后续 API 调用传参。

音频 AI 生成识别服务

最近更新时间：2025-06-03 17:48:42

服务概述

音频内容安全提供了对 AI 生成内容的识别能力，本文将为您介绍[页面体验](#)和 [API 调用](#)两种音频生成识别服务的使用方式。

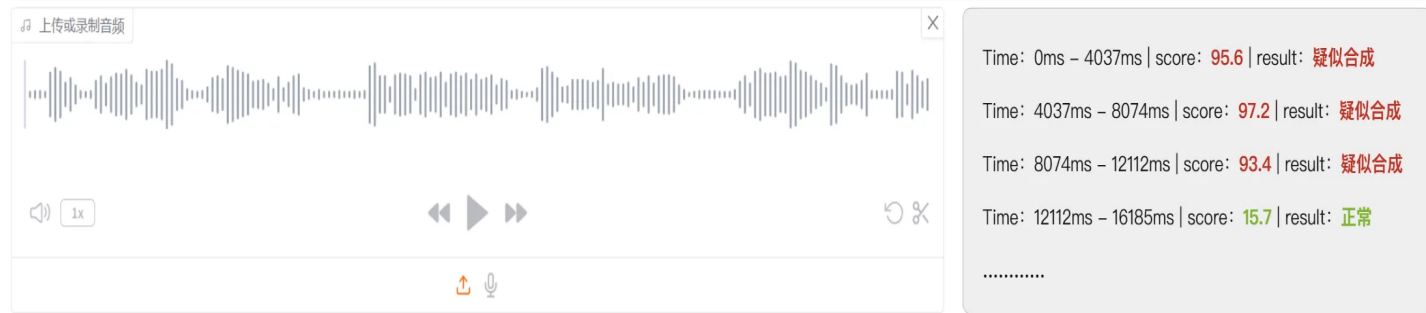
快捷入口：

- 页面体验：[内容安全控制台-生成识别](#)
- API 调用：[音频内容安全接入文档](#)

能力介绍

大模型技术的快速发展，给内容创作领域带来了更多的可能性，但同时也降低了创作门槛，带来了虚假信息、生成非法内容等问题。本产品能识别音频是否为 AI 工具生成，覆盖 Fastspeech 、 VITS 、 VALL-E 、 ChatTTS 、 CosyVoice 等数十种主流音频生成场景，在实际应用中识别准召率高。

逻辑上，系统会将音频按照每4秒为一小段进行截取，然后对截取的每小段鉴别是否生成识别，如下图所示：



获取到每小段的识别结果后，系统会整体来看生成段落在整体视频段落的占比。我们的建议值为：

- 合成段数占比<10%，判断为非合成音频。
- 合成段数占比 10% - 30%，判断为疑似少量合成音频，建议人工二次校验。
- 合成段数占比 >30%，判断为音频整体生成痕迹较重，大概率为合成音频。

服务对接

说明：

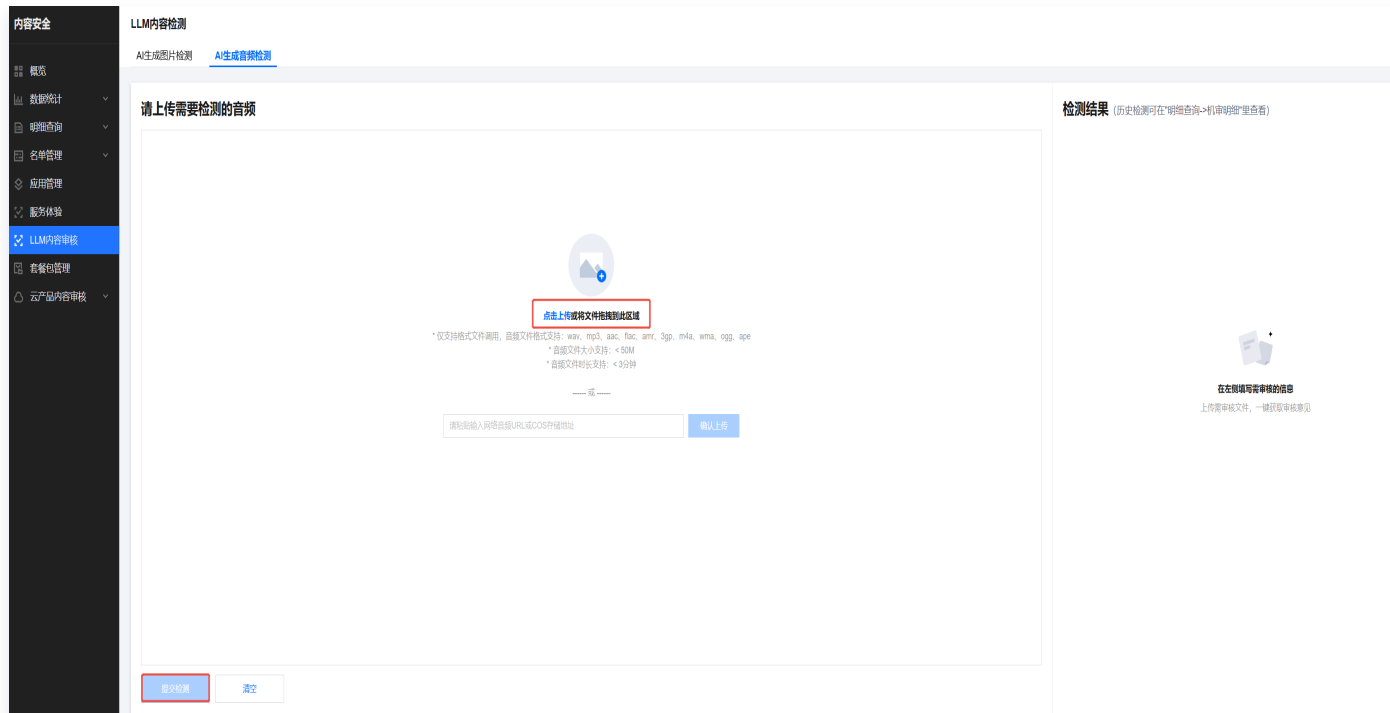
本服务的使用，需要您先登录腾讯云账号。若您未注册腾讯云账号，可参考 [账号注册教程](#) 先行注册。

方式一 通过控制台页面使用

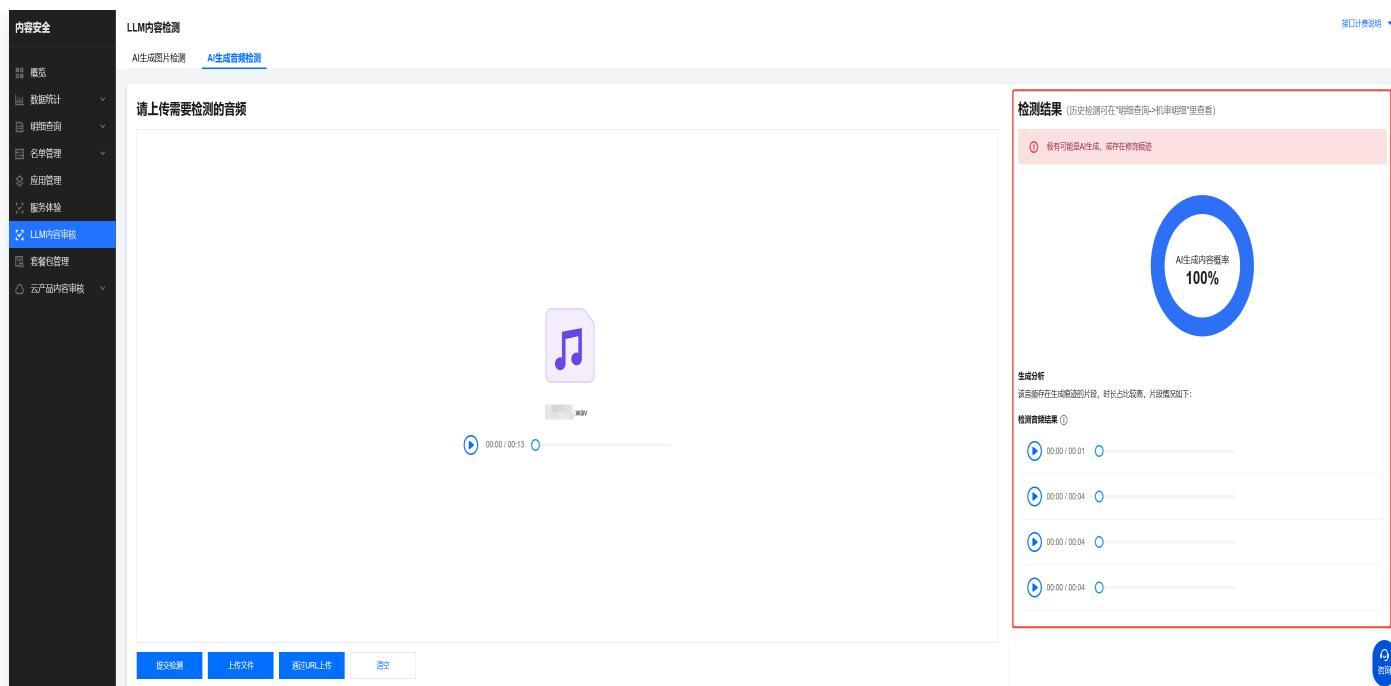
1. 登录 [内容安全控制台](#)，单击 **LLM 内容审核**，跳转后单击**初始化配置**，立即体验。



2. 在指定区域单击**上传文件**或将文件拖拽到此区域。



3. 上传文件或将文件拖拽到此区域后，单击**提交检测**，即可查看**检测结果**。



方式二：通过 API 接口调用

若您需要和您的业务系统做对接，则可使用 API 接口的方式来调用我们的服务。

- 开通服务：**登录 [内容安全控制台](#)，单击 **LLM 内容审核**，跳转后单击**初始化配置**，立即体验。
- 获取接口参数：**单击“应用管理”，即可见如下图所示的“AI 生成检测配套策略”，其中 Biztype 字段对应的值，即接口调用的必备传值。
- 系统对接：**请参见 [音频内容安全接入文档](#) 来完成系统接口的对接，其中，接口参数中 Biztype 的传值，即为上一步您在“应用管理”页面查看的 Biztype 值，Type 的传值为AUDIO_AIGC。

