图片内容安全

操作指南

腾讯云

文档目录

操作指南
敏感词排行榜查询指引

最近更新时间：2024-01-11 10:14:21

功能说明

用户可以使用敏感词排行榜功能对天御内容安全机审模型的审核结果中出现的**敏感关键词命中数据**进行查询。
本文将介绍如何查看敏感关键词排行榜。

操作步骤

1、登录 内容安全控制台 ，在左侧导航栏中，单击**数据统计** > **排行榜**。

> **说明：**
> 子账号使用敏感词排行榜查询如未被主账号授予查询权限，需要按照 子账号授权指引 完成ListUsers权限授权后，才可查询敏感词数据。

2、根据业务需求，筛选字段应用名、场景名、内容类型（图片）、识别类型、处理建议、账号、时间范围、敏感关键词、即可查看统计数据。

> **说明：**
> 时间范围字段筛选时长按照天进行筛选，筛选周期为1-30天。

3、根据业务需求筛选对应字段后即可查看纠错统计数据。

> **说明：**
> 1、敏感词排行可支持显示排名前1000的敏感词；
> 2、敏感词数据支持导出，点击下载按钮即可。

字段	说明
排行名次	按照敏感关键词命中次数正序排名
关键词	模型识别到的违规关键词
处置建议	模型识别到的违规关键词处理建议，包含违规、疑似
识别类型	敏感关键词命中违规的类型
命中次数	敏感关键词对应筛选条件下命中总次数
命中词库类型	命中词库包含系统默认库和自定义词库
操作	1、命中详情：点击“命中详情”可前往-控制台-机审明细-查看对应命中的敏感关键词的汇总明细； 2、关键词详情： • 当关键词命中词库为系统默认库时，“关键词详情”将置灰，不支持查看系统默认词库明细； • 当关键词命中词库为自定义词库时，点击“关键词详情”可查看当前命中词库，前往关键词命中详情页面；

4、命中关键词管理（可选）

当关键词命中词库为自定义词库时，点击“关键词详情”可查看当前命中词库，前往关键词命中详情页面，并可以进行以下操作：

管理：点击“管理”按钮页面将跳转至控制台-名单管理-关键词管理页面，管理该关键词所属词库；

删除：点击”删除“，该命中关键词将会从命中的自定义词库名单中删除；

批量删除：如果关键词同时命中多个自定义词库，勾选对应词库后点击“批量删除”，该关键词将从所选自定义词库中删除；

纠错操作指引
纠错统计

最近更新时间：2023-10-26 09:17:41

功能说明

用户可以使用 纠错功能 对天御内容安全机审模型的审核结果进行修改，以满足业务需求。
本文将介绍如何查看已完成纠错的案例的数据统计。

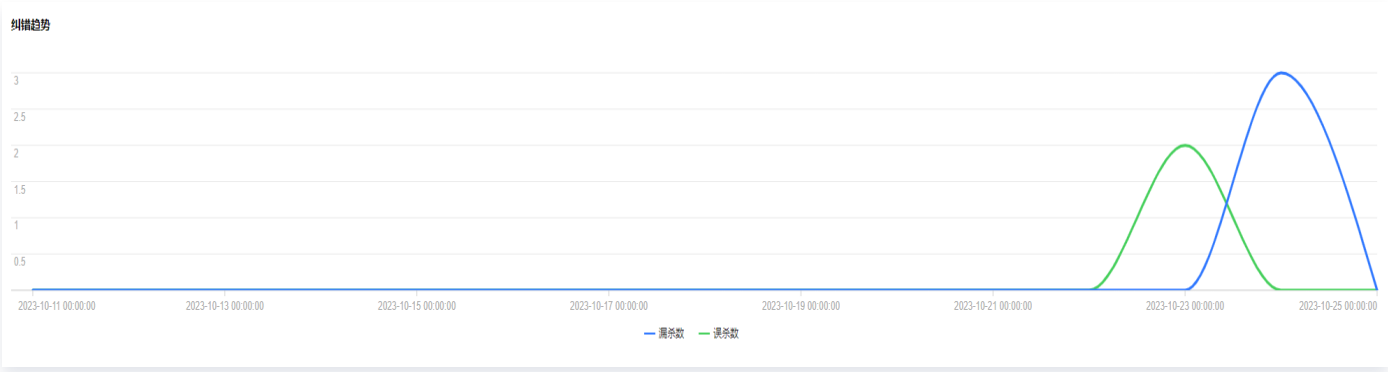
操作步骤

1. 登录 内容安全控制台，在左侧导航栏中，单击**数据统计 > 纠错统计**。

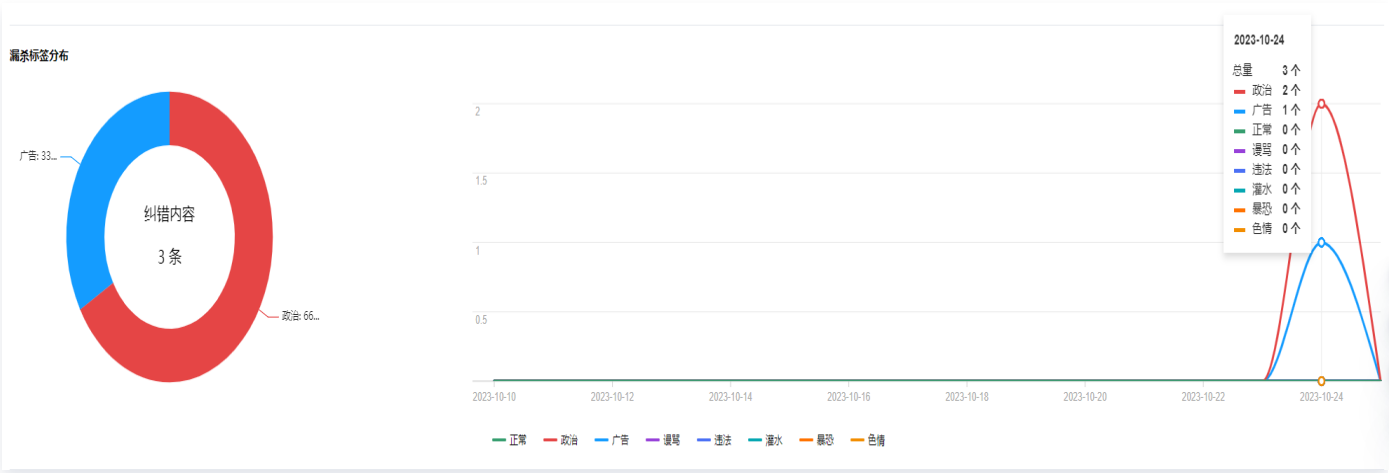
2. 根据业务需求，筛选字段应用、场景、服务类型（图片）、统计周期后即可查看统计数据。

3. 根据业务需求筛选对应字段后即可查看纠错统计数据。
   - **纠错总量**：进行了纠错操作的案例总量统计。
   - **漏杀个数**：机审模型审核识别为“正常”，用户纠正其为风险标签的数据。
   - **误杀个数**：机审识别为“风险标签”,客户纠正为“正常”或“其他风险标签”。

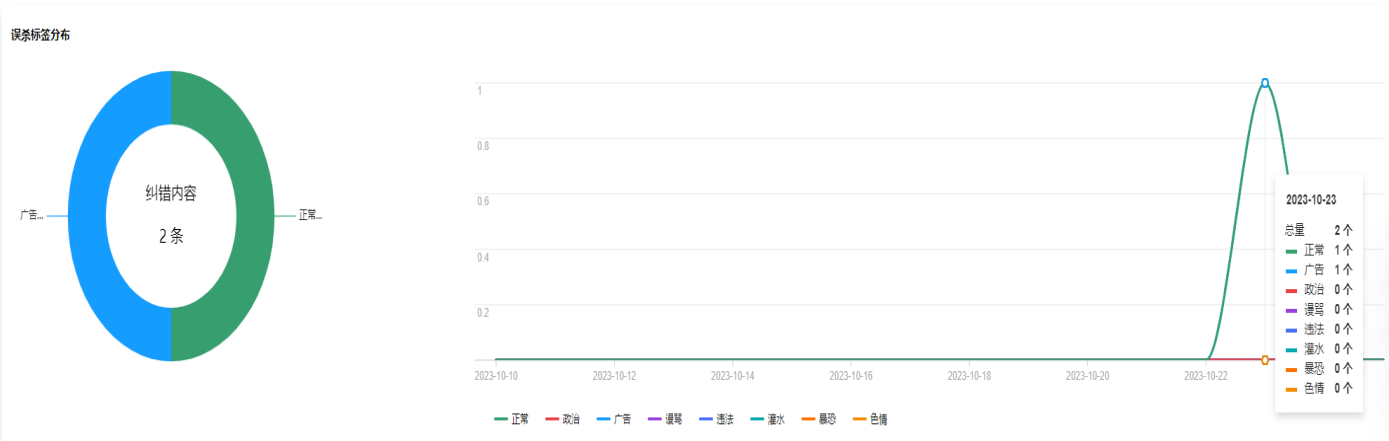
- **纠错趋势：**漏杀数和误杀数在不同时间节点的变化趋势。



- **漏杀标签分布：**不同的漏杀标签数量占比以及不同时间节点变化趋势。



- **误杀标签分布：**不同的误杀标签数量占比以及不同时间节点变化趋势。



审核结果纠错

最近更新时间：2023-12-13 16:07:41

功能说明

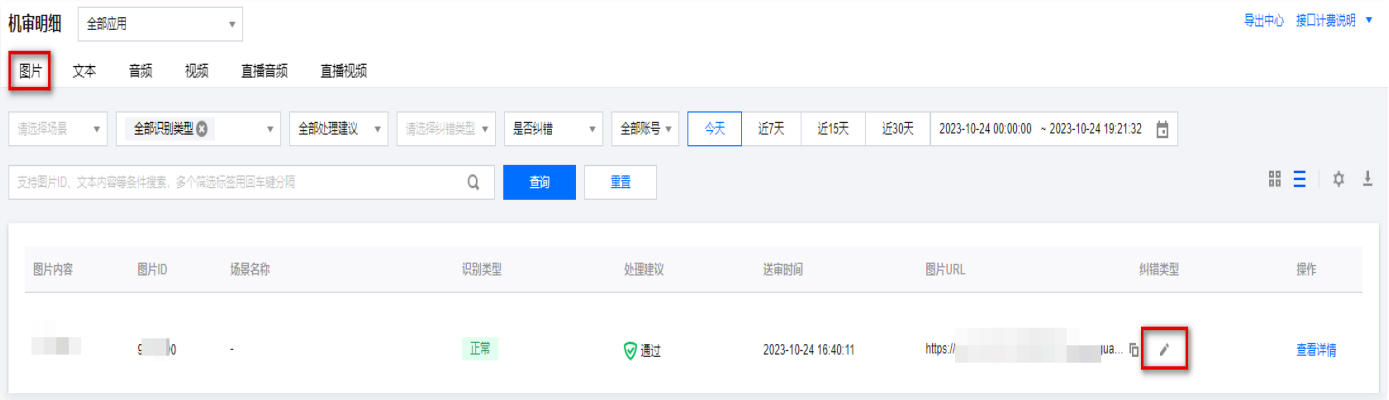
用户可以使用纠错功能对天御内容安全机审模型的审核结果进行修改，以满足业务需求。
本文将介绍如何进行审核结果纠错操作。

操作步骤

1. 登录 内容安全控制台 ，在左侧导航栏中，单击**明细查询** > **机审明细**。

> **说明：**
> 子账号使用审核结果纠错如未被主账号授予权限，需要按照 子账号授权指引 完成ListUsers权限及CMS策略权限授权后，才可使用审核结果纠错功能。

2. 在机审明细的图片页面，找到需要纠错案例，单击**纠错类型**列的 。



3. 在纠错选项页面，查看当前图片机审类型，并选择您认为正确的纠错类型，进行纠错操作。

> **注意：**
> - 纠错操作仅限一次，纠错成功后“纠错类型”将会显示为您所选对应标签。
> - 纠错操作包含：纠错结果回调、加入纠错名单。两类操作可全选或二选一，**不可不选**。

参数名称	参数说明
纠错结果回调	首次进行纠错操作，需配置回调地址，后续进行纠错操作将展示上次填写的回调地址。更多详情请参见 纠错回调数据结构 。 注意：回调地址可支持修改。
加入纠错名单	加入纠错名单后，md5值完全相同的数据，再次送审时将会按照纠错后的标签输出(加入名单后预计10分钟内生效)。

4. 填写完成，单击**确定**完成纠错。

纠错回调数据结构

```
{
  "RequestId": "123",   // 机审请求ID
  "DataId": "456",      // 机审数据ID
  "ContentType": "image", // 送审文件类型：text或image
  "Content": "http://xxx.xxx.xxx",    // 送审原始内容
  "Suggestion": "Block",   // 处理建议
  "Label": "Porn"       // 纠错标签
}
```

纠错名单

最近更新时间：2023-11-20 16:35:22

功能说明

用户可以使用 纠错功能 对天御内容安全机审模型的审核结果进行修改，以满足业务需求。
本文中将介绍如何查看已加入纠错名单的案例。

操作步骤

1. 登录 内容安全控制台，在左侧导航栏中，单击**名单管理** > **纠错名单**。
2. 在纠错名单的图片页面，支持查看已加入纠错名单的图片案例。

> **说明：**
> 可按照账号维度筛选并查看不同账号所操作的纠错案例。

字段名称	字段说明
图片内容	已完成纠错操作的图片内容
图片 ID	已完成纠错操作的图片请求ID
生效类型	仅支持全局生效，全局生效指：纠错结果将在生效至所有应用场景
识别类型	机审模型原始审核结果标签
纠错类型	业务方所选择的结果标签
操作	删除，业务方可将已加入纠错操作的案例进行删除。

	如进行了删除操作，送审相同md5图片，机审模型不再按照纠错结果输出，将按照机审模型原始审核结果输出。