

# 大模型审核接口文档（端到端）

## 目录

- **使用须知**
  - API服务域名
  - 公共头
  - 签名认证访问
- **大模型文本审核API**
  - 1.1 提交大模型文本审核任务
    - 1.1.1 请求说明
    - 1.1.2 请求体说明
    - 1.1.3 响应体说明
  - 1.2 查询大模型文本审核任务
    - 1.2.1 请求说明
    - 1.2.2 请求体说明
    - 1.2.3 响应体说明
- **大模型图像审核API**
  - 2.1 提交大模型文本审核任务
    - 2.1.1 请求说明
    - 2.1.2 请求体说明
    - 2.1.3 响应体说明
  - 2.2 查询大模型文本审核任务
    - 2.2.1 请求说明
    - 2.2.2 请求体说明
    - 2.2.3 响应体说明

## 使用须知

### API服务域名

全局服务，域名为：

1	Region	Endpoint	Protocol
2	全局	sinan.cloud.baidu.com	HTTP and HTTPS

## 公共头

### 公共请求头

1	公共头域	描述	是否必要
2	host	请求host。	是
3	authorization	包含 Access Key 与请求签名	是
4	x-bce-date	计算签名请求头	是
5	content-type	application/json; charset=utf-8	是
6	content-length	上传 Object 的大小，以字节为单位	否

API要求使用JSON格式的结构体来描述一个请求的具体内容。

### 公共响应头

1	公共头部	描述
2	Content-Type	application/json; charset=utf-8。
3	x-bce-request-id	后端生成，并自动设置到响应头域中。

API服务采用JSON格式的消息体作为响应返回的格式。

## 签名认证访问

百度智能云采用统一的API鉴权认证机制，通过使用Access Key Id / Secret Access Key加密的方法来验证某个请求的发送者身份，详情请见[鉴权认证机制](#)。Access Key Id（AK）用于标示用户，Secret Access Key（SK）是用户用于加密认证字符串和百度智能云用来验证认证字符串的密钥，其中SK必须保密，只有用户和百度智能云知道。

当百度智能云接收到用户的请求后，系统将使用相同的SK和同样的认证机制生成认证字符串，并与用户请求中包含的认证字符串进行比对。如果认证字符串相同，系统认为用户拥有指定的操作权限，并执行相关操作；如果认证字符串不同，系统将忽略该操作并返回错误码。

## 1. 大模型文本审核API

### 1.1 提交大模型文本审核任务

1.1.1 请求说明

</> Bash

```
1 POST /v1/text_censor/llm/submit
2 host: sinan.cloud.baidu.com
3 authorization: <bce-authorization-string>
4 x-bce-date: <x-bce-date>
5 content-type: application/json
```

1.1.2 请求体说明

1	字段名称	字段类型	必要性	字段描述
2	contents	List	必须	审核文本内容，单条文本内容不超过2000字
3	scene	String	可选	使用场景，「昵称」、「聊天」、「社区」，默认「昵称」
4	strategy	String	可选	准召策略，「highPrecision」、「highRecall」，默认「highRecall」

</> 请求示例 JSON

```
1 curl -L 'https://sinan.cloud.baidu.com/v1/text_censor/llm/submit' \
2 -H 'authorization: bce-auth-v1/xxxx/2025-08-25T10:32:13Z/7200/host;x-bce-date/xxxx' \
3 -H 'x-bce-date: 2025-08-25T11:31:53Z' \
4 -H 'Content-Type: application/json' \
5 -d '{
6     "contents": ["hhhh", "祖国大陆"],
7     "scene": "社区",
8     "strategy": "highPrecision"
9 }'
```

1.1.3 响应体说明

1	字段名称	字段类型	字段描述
2	taskId	String	提交任务Id
	status	String	分析状态， <i>PROVISIONING</i> , // 0 <i>PROCESSING</i> , // 1

3			<i>FINISHED</i> , // 2
	createTime	Date	创建时间
	scene	String	使用场景
4	strategy	String	准召策略

</> 响应示例JSON

```
1 {
2   "taskId": "rhuhu8zttht07n7k87g0",
3   "status": "PROVISIONING",
4   "scene": "社区",
5   "strategy": "highPrecision",
6   "createTime": "2025-08-29T04:33:03Z"
7 }
```

## 1.2 查询大模型文本审核任务

### 1.2.1 请求说明

</> Bash

```
1 POST /v1/text_censor/llm/pull
2 host: sinan.cloud.baidu.com
3 authorization: <bce-authorization-string>
4 x-bce-date: <x-bce-date>
5 content-type: application/json
```

### 1.2.2 请求体说明

1	字段名称	字段类型	必要性	字段描述
2	taskId	String	必须	任务Id

</> 请求示例JSON

```
1 curl -L 'https://sinan.cloud.baidu.com/v1/text_censor/llm/pull' \
2 -H 'authorization: bce-auth-v1/xxxx/2025-08-25T10:32:13Z/7200/host;x-bce-date/xxxx' \
3 -H 'x-bce-date: 2025-08-25T11:31:53Z' \
4 -H 'Content-Type: application/json' \
5 -d '{
```

```
6     "taskId": "rhuhu8zttht07n7k87g0"
7   }'
```

1.2.3 响应体说明

1	字段名称	字段类型	字段描述
2	taskId	String	提交任务时返回的taskId
3	status	String	分析状态 <i>PROVISIONING</i> , // 0 <i>PROCESSING</i> , // 1 <i>FINISHED</i> , // 2
4	createTime	Date	创建时间
5	finishTime	Date	分析完成时间
6	scene	String	使用场景
7	strategy	String	准召策略
8	results	List	审核文本内容
9	• subtaskId	String	子任务Id
10	• text	String	文本内容
11	• result	Object	单条文本审核结果
12	◦ conclusion	String	「违规」、「正常」、「不确定」、「审核失败」四选一
13	◦ details	List	违规点，仅当“结论=违规”时输出
14	▪ ret	Object	单个违规点结果
15	• label	String	违规类型涉黄、涉政

</> 响应示例 - 运行中

JSON

```
1 {
2   "taskId": "rhuhu8zttht07n7k87g0",
3   "status": "PROCESSING",
4   "scene": "社区",
5   "strategy": "highPrecision",
6   "results": [
7     {
8       "text": "祖国大陆"
9     }
10  ],
```

```
11     "createTime": "2025-08-29T04:35:13Z"
12 }
```

</> 响应示例 - 分析完成

JSON

```
1 {
2     "taskId": "rhuhu8zttht07n7k87g0",
3     "status": "FINISHED",
4     "scene": "社区",
5     "strategy": "highPrecision",
6     "results": [
7         {
8             "subTaskId": "rhuhua8w4nyv277efsu",
9             "text": "祖国大陆",
10            "result": {
11                "conclusion": "pass"
12            }
13        }
14    ],
15    "createTime": "2025-08-29T04:35:13Z",
16    "finishTime": "2025-08-29T04:35:29Z"
17 }
```

## 2. 大模型图像审核API

### 2.1 提交大模型文本审核任务

#### 2.1.1 请求说明

</>

Bash

```
1 POST /v1/img_censor/llm/submit
2 host: sinan.cloud.baidu.com
3 authorization: <bce-authorization-string>
4 x-bce-date: <x-bce-date>
5 content-type: application/json
```

#### 2.1.2 请求体说明

1	字段名称	字段类型	必要性	字段描述
---	------	------	-----	------

2	urls	List<String>	必须	审核图像内容，图像url列表 建议使用：jpg、png、webp、bmp格式 文件大小：10M内
3	scene	String	可选	使用场景，「头像」、「社区」，默认「头像」
4	strategy	String	可选	准召策略，「highPrecision」、 「highRecall」，默认 「highRecall」

</> 请求示例JSON

```
1 curl -L 'https://sinan.cloud.baidu.com/v1/img_censor/llm/submit' \  
2 -H 'authorization: bce-auth-v1/xxxx/2025-08-25T10:32:13Z/7200/host;x-bce-  
   date/xxxx' \  
3 -H 'x-bce-date: 2025-08-25T11:31:53Z' \  
4 -H 'Content-Type: application/json' \  
5 -d '{  
6     "urls": ["http://bj.bcebos.com/xxxx.jpeg"],  
7     "scene": "社区",  
8     "strategy": "highPrecision"  
9 }'
```

2.1.3 响应体说明

1	字段名称	字段类型	字段描述
2	taskId	String	提交任务Id
3	status	String	分析状态 <i>PROVISIONING</i> , // 0 <i>PROCESSING</i> , // 1 <i>FINISHED</i> , // 2
4	createTime	Date	创建时间
5	scene	String	使用场景
6	strategy	String	准召策略

</> 响应示例JSON

```
1 {  
2     "taskId": "rhuhsihdzvqstc85p4j",  
3     "status": "PROVISIONING",
```

```
4     "scene": "社区",
5     "strategy": "highPrecision",
6     "createTime": "2025-08-29T04:33:03Z"
7 }
```

## 2.2 查询大模型文本审核任务

### 2.2.1 请求说明

</> Bash

1 POST /v1/img\_censor/llm/pull
2 host: sinan.cloud.baidu.com
3 authorization: <bce-authorization-string>
4 x-bce-date: <x-bce-date>
5 content-type: application/json

### 2.2.2 请求体说明

1	字段名称	字段类型	必要性	字段描述
2	taskId	String	必须	任务Id

</> 请求示例 JSON

1 curl -L 'https://sinan.cloud.baidu.com/v1/img\_censor/llm/pull' \
2 -H 'authorization: bce-auth-v1/xxxx/2025-08-25T10:32:13Z/7200/host;x-bce-
 date/xxxx' \
3 -H 'x-bce-date: 2025-08-25T11:31:53Z' \
4 -H 'Content-Type: application/json' \
5 -d '{
6 "taskId": "rhttxpakag33ekwa3rx"
7 }'

### 2.2.3 响应体说明

1	字段名称	字段类型	字段描述
2	taskId	String	提交任务时返回的taskId
	status	String	分析状态 <i>PROVISIONING</i> , // 0



3			<i>PROCESSING</i> , // 1
			<i>FINISHED</i> , // 2
	createTime	Date	创建时间
4	finishTime	Date	分析完成时间
	scene	String	使用场景
	strategy	String	准召策略
5	results	List	图像审核结果
	• subtaskId	String	子任务Id
	• url	String	原始请求地址
6	• result	Object	单条图像审核结果
	◦ conclusion	String	「违规」、「正常」、「不确定」、「审核失败」四选一
	◦ data	List	违规点，仅当“结论=违规”时输出
7	▪ ret	Object	单个违规点结果
	• label	String	违规类型涉黄、涉政

</> 响应示例 - pass

JSON

```
1 {
2   "taskId": "rhttxpakag33ekwa3rx",
3   "status": "FINISHED",
4   "scene": "社区",
5   "strategy": "highPrecision",
6   "results": [
7     {
8       "subTaskId": "rhttxqcmwyg43b3yieu",
9       "url": "http://bj.bcebos.com/xxxx.jpeg",
10      "result": {
11        "conclusion": "pass"
12      }
13    }
14  ],
15  "createTime": "2025-08-28T11:43:00Z",
16  "finishTime": "2025-08-28T11:43:20Z"
17 }
```