

腾讯云智能数智人 数智人 API 接入模式



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

数智人 API 接入模式

数智人 API 接入模式概述

数智人 aPaas 接口调用方式

会话交互接入 API 文档

会话交互接入方案概述

云渲染会话交互接入 API 文档

云渲染会话交互服务 API 概述

会话管理

创建会话

使用数智人平台项目创建会话

使用形象资产创建会话

查询会话状态

开启会话

关闭会话

查询会话列表

查询 uin 下的会话列表

查询数智人项目下的会话列表

查询个人资产形象下的会话列表

数智人驱动指令

驱动指令说明

WebSocket 长连接驱动方式

创建长连接通道

文本驱动指令（WebSocket 长连接）

流式文本驱动指令

音频驱动指令

心跳指令

长连接下行消息

HTTP 短连接驱动方式

文本驱动指令（HTTP 短连接）

其他接口

实时更新数智人画面参数

获取进入 TRTC 房间密钥

交互数智人实践教程

附录

交互数智人 API 接入常见问题

视频流嵌入信息说明

错误码

端渲染会话交互接入 API 文档

端渲染会话交互接入概述

创建长连接通道

端渲染驱动接口

直播数智人服务 API 文档

直播数智人服务 API 概述

创建直播间

查询直播间信息

直播接管

查询直播间列表

关闭直播间

直播回调通知

2D真人形象、声音定制 API

形象、声音定制 API 概述

形象定制

接口调用逻辑示意图

获取上传临时令牌

上传素材到腾讯云 COS

定制接口

进度查询接口

效果确认接口

人脸检测任务创建

人脸检测任务状态查询

声音复刻（基础版）

接口调用逻辑示意图

音质检测任务创建接口

音质检测任务状态查询接口

获取上传临时令牌

上传素材到腾讯云 COS

定制接口

进度查询接口

效果确认接口

声音复刻（极速版）

接口调用逻辑示意图

获取声音训练文本

获取上传临时令牌

上传素材到腾讯云 COS

提交音质检测任务

查询音质检测任务进度

定制接口

进度查询接口

结果码说明

接口对接常见问题

视频生成服务 API 文档

视频生成服务 API 概述

音频制作接口

视频制作接口-基础版

视频制作接口-进阶版

视频制作接口-视频免训练

视频制作接口-照片免训练

音视频制作进度查询接口

附录

附录一：结果码字典

附录二：回调请求体格式

接口对接常见问题

个人资产管理 API 文档

个人资产管理 API 概述

查询小样本形象列表

查询服务资产信息

查询声音资产信息

查询形象资产信息-查询主播

查询形象资产信息-查询主播下的所有形象

查询形象支持的动作列表

附录1-服务资产类型

附录2-情感风格

附录3-数智人类型

附录4-语种列表

接口对接常见问题

配置端 iPaaS 接入

产品介绍

基础版接入指引

数智人 API 接入模式

数智人 API 接入模式概述

最近更新时间：2025-05-06 21:19:12

数智人支持以 API 形式接入的服务有以下几类：

会话交互：借助大模型可以实现跟数智人实时进行对话的交互场景，主要用于导览、知识问答等场景。区分云渲染和端渲染。

直播：以虚拟人形象构建长时间的直播，支持长文本的提前录入。

形象及声音定制：通过 API 发起形象或声音的定制请求，会消耗已购买的配额。

视频生成：通过 API 形式生成指定的视频（含音频），会消耗时长包。

个人资产管理：通过 API 形式查询、分配已购买的资源。

iPaaS 嵌入服务：帮助用户快速将数智人平台的 UI 交互集成到第三方业务系统或平台系统里。

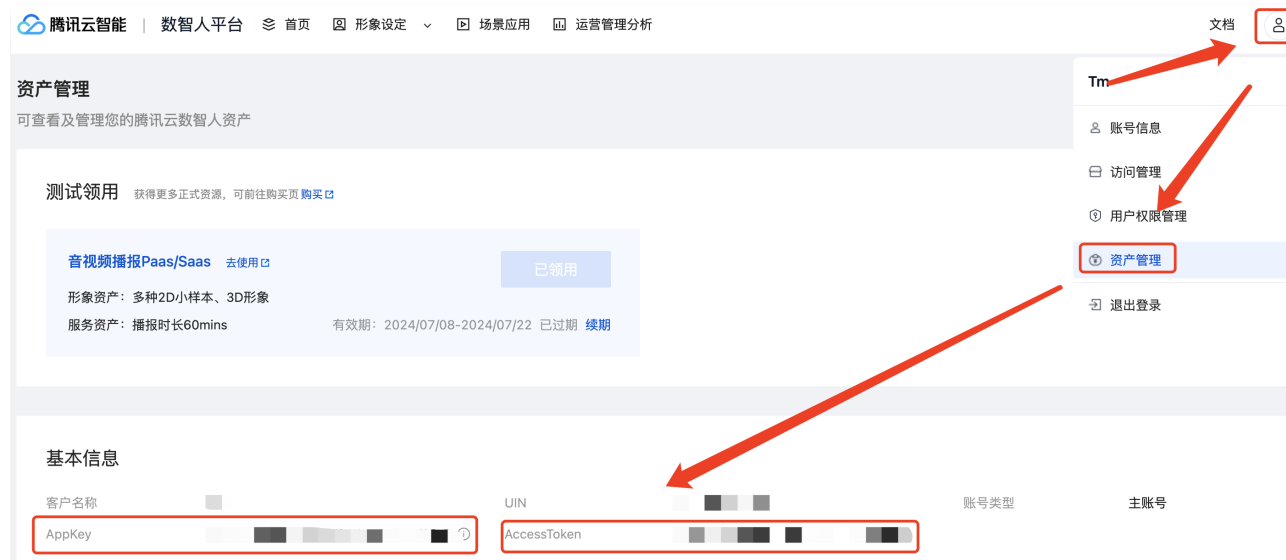
数智人 aPaas 接口调用方式

最近更新时间：2025-04-01 16:13:33

本文主要提供 生成腾讯云智能数智人 aPaas 平台接口调用URL 的方式，以及常见编程语言 demo，语言类型包括 Golang、Java、C++、Python、JavaScript。

一、必要参数获取

appkey、accesstoken请在 [数智人平台资产管理页面](#) 获取：



二、公共参数

在调用 aPaas 任一接口时，都需在 URL 中以 QueryString 形式携带下面的等公共参数。

参数	类型	必须	说明
appkey	string	是	根据各接口文档提供的方式获取到的 appkey。
timestamp	string	是	请求时间戳，单位：秒。时间戳需要和当前时间的差异不过能超过五分钟，否则会鉴权失败。
requestid	string	否	只有 建立会话长链接 接口需要携带该参数，其他接口不需要携带。获取方式请参考对应接口文档。
signature	string	是	请求签名（参考 签名方法 ）。

携带参数示例：

示例1：https协议，注意每个参数的值替换成实际的值

```
https://gw.tv.s.qq.com/v2/ivh/sessionmanager/sessionmanagerservice/createsession?appkey={{appkey}}&timestamp={{timestamp}}&signature={{signature}}
```

示例2：建立会话长链接

```
wss://gw.tv.s.qq.com/v2/ws/ivh/interactdriver/interactdriverservice/commandchannel?appkey={{appkey}}&requestid={{requestid}}&timestamp={{timestamp}}&signature={{signature}}
```

⚠ 注意：

[建立会话长链接](#) 接口的 url 会有 appkey、requestid、timestamp 三个参数，注意都要拼接。

三、签名方法

请求参数签名步骤如下：

1. signature 签名规则如下，这里以下为例说明(示例为伪代码仅供参考)：

```
appkey = example_appkey
accesstoken = example_accesstoken
域名路由 = https://api.example.com/v2/ivh/example_uri
```

2. 根据对应接口文档，对 query 参数中除 signature 之外的所有需要的参数按字典序进行排序，作为签名原文，用于生成签名。示例中仅使用 appkey 和 timestamp 两个参数（各个接口需要的参数请在相关文档中确认），排序拼接后的字符串示例为：

```
appkey=example_appkey&timestamp=1717639699
```

3. 对签名原文使用 accesstoken 进行 HmacSha256 加密，之后再 base64 编码：

```
hashBytes = HmacSha256("appkey=example_appkey&timestamp=1717639699", "example_accesstoken")
signature = Base64Encode(hashBytes)
```

4. 得到 signature 签名值为：

```
aCNWYzZdplxWVo+JsqzZc9+J9XrwWWITfX3eQpsLVno=
```

5. 将 signature 值进行 urlencode（必须进行 URL 编码，否则将导致鉴权失败）后拼接得到最终请求 URL 为：

```
https://api.example.com/v2/ivh/example_uri?
appkey=example_appkey&timestamp=1717639699&signature=aCNWYzZdplxWVo%2BJsqzZc9%2BJ9XrwWWITfX3eQpsLVno%
3D
```

四、签名生成 Demo

用户只需要将访问对应接口需要的除 signature 以外所有公共参数填入 demo 中相应位置，即可生成请求签名并拼接成访问接口的完整 URL。

Golang

```
package main

import (
    "crypto/hmac"
    "crypto/sha256"
    "encoding/base64"
    "fmt"
    "net/url"
    "sort"
    "strconv"
    "time"
)

func GenSignature(signingContent string, accessToken string) string {

    // 计算 HMAC-SHA256 值
    h := hmac.New(sha256.New, []byte(accessToken))
    h.Write([]byte(signingContent))

    // 将 HMAC-SHA256 值进行 Base64 编码
    hashInBase64 := base64.StdEncoding.EncodeToString(h.Sum(nil))
}
```

```
// URL encode
encodeSign := url.QueryEscape(hashInBase64)

// 拼接签名
signature := "&signature=" + encodeSign

return signature
}

func GenReqURL(parameter map[string]string, accessToken string, baseUrl string) string {
    // 按字典序拼接待计算签名的字符串
    signingContent := ""
    pathKey := make([]string, 0, len(parameter))
    for k := range parameter {
        pathKey = append(pathKey, k)
    }
    sort.Strings(pathKey)
    for _, k := range pathKey {
        if signingContent != "" {
            signingContent += "&"
        }
        signingContent += k + "=" + parameter[k]
    }

    // 计算签名
    signature := GenSignature(signingContent, accessToken)

    // 拼接访问接口的完整URL
    return baseUrl + "?" + signingContent + signature
}

func main() {
    baseUrl := "https://api.example.com/v2/ivh/example_uri"
    accesstoken := "example_accesstoken"
    wssUrl := "wss://api.example.com/v2/ws/ivh/example_uri"
    // 示例一(访问需要提供 appkey 和 timestamp 参数的接口):
    // 用户按需填入生成签名所需的公共参数
    parameter := map[string]string{
        "appkey": "example_appkey",
        // 用户提供的时间戳应保证单位为秒且与当前时间相差不得超过五分钟
        // 示例时间戳:
        // "timestamp": "1717639699",
        // 推荐使用下面的语句生成当前时间戳:
        "timestamp": strconv.FormatInt(time.Now().Unix(), 10),
    }
    url1 := GenReqURL(parameter, accesstoken, baseUrl)
    // 使用示例时间戳输出应当如下:
    // Example 1:https://api.example.com/v2/ivh/example_uri?
    appkey=example_appkey&timestamp=1717639699&signature=aCNWYzZdplxWVo%2BJsqzZc9%2BJ9XrwWWITfX3eQpsLVn
    o%3D
    fmt.Println("Example 1:" + url1)

    // 示例二(访问需要提供 appkey, requestID 和 timestamp 参数的接口):
    // 用户按需填入生成签名所需的公共参数
    parameter = map[string]string{
        "appkey": "example_appkey",
        "requestid": "example_requestid",
        // 用户提供的时间戳应保证单位为秒且与当前时间相差不得超过五分钟
        // 示例时间戳:
    }
```

```
// "timestamp": "1717639699",
// 用户若自己提供时间戳应保证单位为秒且与当前时间相差不得超过五分钟
"timestamp": strconv.FormatInt(time.Now().Unix(), 10),
}
url2 := GenReqURL(parameter, accesstoken, wssUrl)
// 使用示例时间戳输出应当如下:
// Example 2:wss://api.example.com/v2/ws/ivh/example_uri?
appkey=example_appkey&requestid=example_requestid&timestamp=1717639699&signature=QVenICk0VHtHGYZKXM
6IC%2BW1CjZC1joSr%2Fx0gfKKYT4%3D
fmt.Println("Example 2:" + url2)
}
```

Java

```
import java.util.*;
import java.util.stream.Collectors;
import java.time.Instant;
import java.net.URLEncoder;
import java.nio.charset.StandardCharsets;
import javax.crypto.Mac;
import javax.crypto.spec.SecretKeySpec;
import java.util.Base64;
import java.security.InvalidKeyException;
import java.security.NoSuchAlgorithmException;
import java.io.UnsupportedEncodingException;

public class Presigned {

    public static String GenSignature(String signingContent, String accessToken) {
        try {
            // 计算 HMAC-SHA256 值
            Mac sha256_HMAC = Mac.getInstance("HmacSHA256");
            SecretKeySpec secret_key = new
            SecretKeySpec(accessToken.getBytes(StandardCharsets.UTF_8), "HmacSHA256");
            sha256_HMAC.init(secret_key);

            String hashInBase64 =
            Base64.getEncoder().encodeToString(sha256_HMAC.doFinal(signingContent.getBytes(StandardCharsets.UTF_8)));

            // URL encode
            String encodeSign = URLEncoder.encode(hashInBase64, StandardCharsets.UTF_8.toString());

            // 拼接签名
            String signature = "&signature=" + encodeSign;

            return signature;
        } catch (NoSuchAlgorithmException | InvalidKeyException | UnsupportedEncodingException e) {
            e.printStackTrace();
            return null;
        }
    }

    public static String GenReqURL(Map<String, String> parameter, String accessToken, String
    baseUrl) {
```

```
// 按字典序拼接待计算签名的字符串
String signingContent = parameter.entrySet().stream()
    .sorted(Map.Entry.comparingByKey())
    .map(entry -> entry.getKey() + "=" + entry.getValue())
    .collect(Collectors.joining("&"));

// 计算签名
String signature = GenSignature(signingContent, accessToken);

// 拼接访问接口的完整URL
return baseUrl + "?" + signingContent + signature;
}

public static void main(String[] args) {
    String baseUrl = "https://api.example.com/v2/ivh/example_uri";
    String accesstoken = "example_accesstoken";
    String wssUrl = "wss://api.example.com/v2/ws/ivh/example_uri";

    // 示例一(访问需要提供 appkey 和 timestamp 参数的接口):
    // 用户按需填入生成签名所需的公共参数
    Map<String, String> parameter = new TreeMap<>();
    parameter.put("appkey", "example_appkey");
    // 用户提供的的时间戳应保证单位为秒且与当前时间相差不得超过五分钟
    // 示例时间戳:
    // parameter.put("timestamp", "1717639699");
    // 推荐使用下面的语句生成当前时间戳:
    parameter.put("timestamp", String.valueOf(Instant.now().getEpochSecond()));
    String url = GenReqURL(parameter, accesstoken, baseUrl);
    // 使用示例时间戳输出应当如下:
    // Example 1:https://api.example.com/v2/ivh/example_uri?
    appkey=example_appkey&timestamp=1717639699&signature=aCNWYzZdplxWVo%2BJsqzZc9%2BJ9XrwWWITfX3eQpsLVn
    o%3D

    System.out.println("Example 1:" + url);

    // 示例二(访问需要提供 appkey, requestId 和 timestamp 参数的接口):
    parameter.clear();
    parameter.put("appkey", "example_appkey");
    parameter.put("requestid", "example_requestid");
    // 用户提供的的时间戳应保证单位为秒且与当前时间相差不得超过五分钟
    // 示例时间戳:
    // parameter.put("timestamp", "1717639699");
    // 推荐使用下面的语句生成当前时间戳:
    parameter.put("timestamp", String.valueOf(Instant.now().getEpochSecond()));
    url = GenReqURL(parameter, accesstoken, wssUrl);
    // 使用示例时间戳输出应当如下:
    // Example 2:wss://api.example.com/v2/ws/ivh/example_uri?
    appkey=example_appkey&requestid=example_requestid&timestamp=1717639699&signature=QVenICk0VHtHGYZKXM
    6IC%2BW1CjZC1joSr%2Fx0gfkKYT4%3D

    System.out.println("Example 2:" + url);
}
}
```

C++

```
#include<iostream>
#include <map>
#include<string>
#include<vector>
#include<algorithm>
#include <ctime>
#include <iomanip>
#include <sstream>
#include<openssl/hmac.h>
#include<openssl/sha.h>

using namespace std;

static const string base64_chars =
    "ABCDEFGHIJKLMNOPQRSTUVWXYZ"
    "abcdefghijklmnopqrstuvwxyz"
    "0123456789+/";

string url_encode(const string &value) {
    ostringstream escaped;
    escaped.fill('0');
    escaped<< hex;

    for (size_t i = 0; i< value.length(); ++i) {
        char c = value[i];
        if (isalnum(c) || c == '-' || c == '_' || c == '.' || c == '~') {
            escaped << c;
        } else {
            escaped<< uppercase;
            escaped << '%'<< setw(2)<< int((unsigned char) c);
            escaped<< nouppercase;
        }
    }

    return escaped.str();
}

string base64_encode(unsigned char const* bytes_to_encode, unsigned int in_len) {
    string ret;
    int i = 0;
    int j = 0;
    unsigned char char_array_3[3];
    unsigned char char_array_4[4];
    while (in_len--) {
        char_array_3[i++] = *(bytes_to_encode++);
        if (i == 3) {
            char_array_4[0] = (char_array_3[0] & 0xfc) >> 2;
            char_array_4[1] = ((char_array_3[0] & 0x03) << 4) + ((char_array_3[1] & 0xf0) >> 4);
            char_array_4[2] = ((char_array_3[1] & 0x0f) << 2) + ((char_array_3[2] & 0xc0) >> 6);
            char_array_4[3] = char_array_3[2] & 0x3f;

```



```
        for(i = 0; (i < 4) ; i++)
            ret += base64_chars[char_array_4[i]];
        i = 0;
    }
}
if (i)
{
    for(j = i; j < 3; j++)
        char_array_3[j] = '\\0';
    char_array_4[0] = (char_array_3[0] & 0xfc) >> 2;
    char_array_4[1] = ((char_array_3[0] & 0x03) << 4) + ((char_array_3[1] & 0xf0) >> 4);
    char_array_4[2] = ((char_array_3[1] & 0x0f) << 2) + ((char_array_3[2] & 0xc0) >> 6);
    char_array_4[3] = char_array_3[2] & 0x3f;
    for (j = 0; (j < i + 1); j++)
        ret += base64_chars[char_array_4[j]];
    while((i++ < 3))
        ret += '=';
}
return ret;
}

string GenSignature(const string& signingContent, const string& accessToken) {

    // 计算 HMAC-SHA256 值
    unsigned char hash[EVP_MAX_MD_SIZE];
    unsigned int hashLength;
    HMAC(EVP_sha256(), accessToken.c_str(), accessToken.length(), (unsigned char *)
signingContent.c_str(), signingContent.length(), hash, &hashLength);

    // 将 HMAC-SHA256 值进行 Base64 编码
    string hashInBase64 = base64_encode(hash, hashLength);

    // URL encode
    string encodeSign = url_encode(hashInBase64);

    // 拼接签名
    string signature = "&signature=" + encodeSign;

    return signature;
}

string GenReqURL(const map<string, string>& parameter, const string& accessToken, const string&
baseURL) {
    // 按字典序拼接待计算签名的字符串
    string signingContent;
    vector<string> pathKey;
    for (const auto& p : parameter) {
        pathKey.push_back(p.first);
    }
    sort(pathKey.begin(), pathKey.end());
```

```
for (const auto& k : pathKey) {
    if (!signingContent.empty()) {
        signingContent += "&";
    }
    signingContent += k + "=" + parameter.at(k);
}

// 计算签名
string signature = GenSignature(signingContent, accessToken);

// 拼接访问接口的完整URL
return baseUrl + "?" + signingContent + signature;
}

// 编译时需要链接到OpenSSL库，可以使用下面的命令：
// g++ presigned.cpp -o presigned -lcrypto

int main() {
    string baseUrl = "https://api.example.com/v2/ivh/example_uri";
    string accesstoken = "example_accesstoken";
    string wssUrl = "wss://api.example.com/v2/ws/ivh/example_uri";

    // 示例一(访问需要提供 appkey 和 timestamp 参数的接口):
    // 用户按需填入生成签名所需的公共参数
    map<string, string> parameter1 = {
        {"appkey", "example_appkey"},
        // 用户提供的时间戳应保证单位为秒且与当前时间相差不得超过五分钟
        // 示例时间戳:
        // {"timestamp", "1717639699"},
        // 推荐使用下面的语句生成当前时间戳:
        {"timestamp", to_string(time(NULL))}
    };
    string url = GenReqURL(parameter1, accesstoken, baseUrl);
    // 使用示例时间戳输出应当如下:
    // Example 1:https://api.example.com/v2/ivh/example_uri?
    appkey=example_appkey&timestamp=1717639699&signature=aCNWYzZdplxWVo%2BJsqzZc9%2BJ9XrwwWITfX3eQpsLVn
    o%3D
    cout << "Example 1:"<< url<< endl;

    // 示例二(访问需要提供 appkey, requestID 和 timestamp 参数的接口):
    // 用户按需填入生成签名所需的公共参数
    map<string, string> parameter2 = {
        {"appkey", "example_appkey"},
        {"requestid", "example_requestid"},
        // 用户提供的时间戳应保证单位为秒且与当前时间相差不得超过五分钟
        // 示例时间戳:
        // {"timestamp", "1717639699"},
        // 推荐使用下面的语句生成当前时间戳:
        {"timestamp", to_string(time(NULL))}
    };
    url = GenReqURL(parameter2, accesstoken, wssUrl);
    // 使用示例时间戳输出应当如下:
```

```
// Example 2:wss://api.example.com/v2/ws/ivh/example_uri?
appkey=example_appkey&requestid=example_requestid&timestamp=1717639699&signature=QVenICk0VHtHGYZKXM
6IC%2BW1CjZC1joSr%2Fx0gfKKYT4%3D
cout << "Example 2:"<< url<< endl;

return 0;
}
```

Python

```
import hmac
import hashlib
import time
import base64
from urllib.parse import quote

# 用户可以在函数内部生成时间戳，只需要传入appkey和accessToken即可获取访问接口所需的公共参数和签名
def GenSignature(signing_content, access_token):

    # 计算 HMAC-SHA256 值
    h = hmac.new(access_token.encode(), signing_content.encode(), hashlib.sha256)

    # 将 HMAC-SHA256 值进行 Base64 编码
    hash_in_base64 = base64.b64encode(h.digest()).decode()

    # URL encode
    encode_sign = quote(hash_in_base64)

    # 拼接签名
    signature = f"&signature={encode_sign}"

    return signature

def GenReqURL(parameter, access_token, base_url):
    # 按字典序拼接待计算签名的字符串
    signing_content = '&'.join(f'{k}={parameter[k]}' for k in sorted(parameter.keys()))

    # 计算签名
    signature = GenSignature(signing_content, access_token)

    # 拼接访问接口的完整 URL
    return f'{base_url}?{signing_content}{signature}'
```

```
def main():
    base_url = 'https://api.example.com/v2/ivh/example_uri'
    access_token = 'example_accesstoken'
    wss_url = 'wss://api.example.com/v2/ws/ivh/example_uri'

    # 示例一 (访问需要提供 appkey 和 timestamp 参数的接口):
    # 用户按需填入生成签名所需的公共参数
    parameter1 = {
        'appkey': 'example_appkey',
        # 用户提供的时间戳应保证单位为秒且与当前时间相差不得超过五分钟
        # 示例时间戳:
        # 'timestamp': '1717639699'
        # 推荐使用下面的语句生成当前时间戳:
        'timestamp': int(time.time()) # 使用当前时间戳 (单位: 秒)
    }
    url1 = GenReqURL(parameter1, access_token, base_url)
    # 使用示例时间戳输出应当如下:
    # Example 1:https://api.example.com/v2/ivh/example_uri?
    appkey=example_appkey&timestamp=1717639699&signature=aCNWYzZdplxWVo%2BJsqzZc9%2BJ9XrwWWITfX3eQpsLVno%3D
    print('Example 1:', url1)

    # 示例二 (访问需要提供 appkey, requestID 和 timestamp 参数的接口):
    # 用户按需填入生成签名所需的公共参数
    parameter2 = {
        'appkey': 'example_appkey',
        'requestid': 'example_requestid',
        # 用户提供的时间戳应保证单位为秒且与当前时间相差不得超过五分钟
        # 示例时间戳:
        # 'timestamp': '1717639699'
        # 推荐使用下面的语句生成当前时间戳:
        'timestamp': int(time.time()) # 使用当前时间戳 (单位: 秒)
    }
    url2 = GenReqURL(parameter2, access_token, wss_url)
    # 使用示例时间戳输出应当如下:
    # Example 2:wss://api.example.com/v2/ws/ivh/example_uri?
    appkey=example_appkey&requestid=example_requestid&timestamp=1717639699&signature=QVenICK0VHtHGYZKXM6IC%2BW1CjZC1joSr%2Fx0gfKKYT4%3D
    print('Example 2:', url2)

if __name__ == '__main__':
    main()
```

JavaScript

```
const crypto = require('crypto');
```

```
// 用户可以在函数内部生成时间戳，只需要传入appkey和accessToken即可获取访问接口所需的公共参数和签名
function GenSignature(signingContent, accessToken) {
  // 计算 HMAC-SHA256 值
  const hmac = crypto.createHmac('sha256', accessToken);
  hmac.update(signingContent);

  // 将 HMAC-SHA256 值进行 Base64 编码
  const hashInBase64 = hmac.digest('base64');

  // URL encode
  const encodeSign = encodeURIComponent(hashInBase64);

  // 拼接签名
  const signature = `&signature=${encodeURIComponent(encodeSign)}`;

  return signature;
}

function GenReqURL(parameter, accessToken, baseUrl) {
  // 按字典序拼接待计算签名的字符串
  let signingContent = '';
  const pathKey = Object.keys(parameter).sort();
  for (const k of pathKey) {
    if (signingContent !== '') {
      signingContent += '&';
    }
    signingContent += `${k}=${parameter[k]}`;
  }

  // 计算签名
  const signature = GenSignature(signingContent, accessToken);

  // 拼接访问接口的完整URL
  return `${baseUrl}?${signingContent}${signature}`;
}

function main() {
  const baseUrl = 'https://api.example.com/v2/ivh/example_uri';
  const accesstoken = 'example_accesstoken';
  const wssUrl = 'wss://api.example.com/v2/ws/ivh/example_uri';

  // 示例一 (访问需要提供 appkey 和 timestamp 参数的接口):
  // 用户按需填入生成签名所需的公共参数
  const parameter1 = {
    appkey: 'example_appkey',
    // 用户提供的时间戳应保证单位为秒且与当前时间相差不得超过五分钟
    // 示例时间戳:
    // timestamp: '1717639699',
    // 推荐使用下面的语句生成当前时间戳:
    timestamp: Math.floor(Date.now() / 1000), // 使用当前时间戳 (单位: 秒)
  };
}
```

```
};
const url1 = GenReqURL(parameter1, accesstoken, baseUrl);
// 使用示例时间戳输出应当如下:
// Example 1:https://api.example.com/v2/ivh/example_uri?
appkey=example_appkey&timestamp=1717639699&signature=aCNWYzZdplxWVo%2BJsqzZc9%2BJ9XrwWWITfX3eQpsLVn
o%3D
console.log('Example 1:', url1);

// 示例二 (访问需要提供 appkey, requestID 和 timestamp 参数的接口):
const parameter2 = {
  appkey: 'example_appkey',
  requestid: 'example_requestid',
  // 用户提供的时间戳应保证单位为秒且与当前时间相差不得超过五分钟
  // 示例时间戳:
  // timestamp: '1717639699',
  // 推荐使用下面的语句生成当前时间戳:
  timestamp: Math.floor(Date.now() / 1000), // 使用当前时间戳 (单位: 秒)
};
const url2 = GenReqURL(parameter2, accesstoken, wssUrl);
// 使用示例时间戳输出应当如下:
// Example 2:wss://api.example.com/v2/ws/ivh/example_uri?
appkey=example_appkey&requestid=example_requestid&timestamp=1717639699&signature=QVenICK0VHtHGYZKXM
6IC%2BW1CjZC1joSr%2Fx0gfKKYT4%3D
console.log('Example 2:', url2);
}

main();
```

五、接口调用注意事项

发送请求调用数智人相关接口时，请求体都要求有 header 和 payload 参数，请注意不要把这里的 header 理解为请求头，它也是请求体的一个参数。

以下为请求体的正确写法示例：

```
{
  "Header": {

  },
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "ParamKey1": "paramValue1",
    "ParamKey2": "paramValue2"
  }
}
```

以下为请求体的错误写法示例：

```
{
  "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
  "ParamKey1": "paramValue1",
  "ParamKey2": "paramValue2"
}
```

会话交互接入 API 文档

会话交互接入方案概述

最近更新时间：2025-05-06 21:42:22

数智人整体会话交互接入方案分为两种形态，云渲染和端渲染两种场景。

- **云渲染**：将数字人的模型、动作、场景等数据传输至远程服务器集群，通过云端高性能计算资源（如 GPU 集群）完成渲染计算，最终将生成的音视频流（或图像序列）实时回传至用户终端（如手机、电脑、VR 设备等）的技术方案。其核心是“计算在云端，显示在终端”。
- **端渲染**：端渲染是将数字人的模型、资源及渲染引擎直接部署在用户本地设备（如 PC、手机、智能硬件），通过设备自身的 CPU/GPU 完成渲染计算，并在本地实时显示的技术方案。其核心是“计算与显示均在终端”。

为了更好的服务用户，针对我们提供的工具平台，我们还提供了 [配置端 iPaaS 接入](#)，方便有平台集成能力的厂商直接集成我们的工具平台，以便给用户提供更便捷的数字人操作体验。详情可以跳转相关链接进行查看：

- [云渲染会话交互接入 API 文档](#)
- [端渲染会话交互接入 API 文档](#)
- [配置端 iPaaS 接入](#)

云渲染会话交互接入 API 文档

云渲染会话交互服务 API 概述

最近更新时间：2025-05-06 21:42:22

本文档主要描述 腾讯云智能数智人 交互aPaas 平台的开放接口协议，以及阐述接口的调用流程。

调用前须知

调用交互数智人服务接口前请确保您进行了形象采购以及购买了交互服务，详情请查看购买指南中的 [购买指引](#)。

云渲染会话交互接入方案有两种对接形式，包括 API 接入方式和 [SDK 接入方式](#)。云渲染场景的数字人形象需要通过云端引擎来驱动，所以必须购买的服务包括：**在有效期内的形象、会话互动并发**，如下图所示：

- **形象的来源有三种：**
 - 租赁我们已有的形象，该模式下需要在购买页购买“形象租赁”中选择形象进行下单，默认有效期1年。1年后需重新购买。
 - 找我们定制形象，该模式下需要联系我们，进行专门的形象定制。会产生三笔费用：
 - 定制费用，联系我们获取；
 - 引擎费用，在购买页下单“云端驱动引擎”；
 - 定制形象的续期费用，形象1年后会过期，需要在过期前下单“定制形象续期”对定制形象进行续费。
 - 客户自己提供形象，该模式下需要联系我们，进行形象的对接，形式上也属于定制形象，但没有定制费用。会产生两笔费用：
 - 引擎费用，在购买页下单“云端驱动引擎”；
 - 形象续期费用，在过期前下单“定制形象续期”对形象进行续费。
- **会话交互：**该交互用于构建多路实时会话，例如用户需要同时保持2路不相关的实时会话，则需要购买2路“会话互动并发数”。

云智能数智人

返回产品详情

2D形象

3D形象

声音定制

更多服务

购买须知

计费概述

购买须知

使用须知

退款说明

选择配置

形象类型

3D写实

3D半写实

3D卡通

形象类型说明

采购类型

形象租赁

定制形象续期

云端驱动引擎

选择角色

瑞光

智皓

小微

瑞晓3

形象可用时长

1个月

2个月

3个月

6个月

1年

2年

3年

资源包（服务资产）

会话互动并发数（云渲染场景）

0

路

音视频播报开发

0

路

音频播报合成小时包

0

个

视频播报合成小时包

0

个

端渲染会话驱动次数包

0

个

服务合约

我已阅读并同意《产品服务协议》和《服务等级协议》

接口调用环境

正式环境域名：
https://gw.tv.s.qq.com
wss://gw.tv.s.qq.com

接口调用方式

调用每个接口时需要在接口的 query 中传入公共参数及签名。参数及签名方式说明参考以下链接：
[数智人 aPaas 接口调用方式](#)

交互数智人 API 概览

功能类别	API 标题		接口	说明
会话管理	创建会话		POST /v2/ivh/sessionmanager /sessionmanagerservice/createsession	使用数智人平台项目创建会话，获取数智人画面的流播放地址。
		使用数智人平台项目创建会话		

		使用形象资产创建会话	POST /v2/ivh/sessionmanager/sessionmanagerservice/createsessionbyasset	使用客户有权限的形象资产直接创建会话，获取数智人画面的流播放地址。用于未在数智人平台创建项目的情况。
	查询会话状态		POST /v2/ivh/sessionmanager/sessionmanagerservice/statssession	查询指定会话的当前状态。
	开启会话		POST /v2/ivh/sessionmanager/sessionmanagerservice/startsession	会话就绪之后，必须调用此口才可以进行驱动。
	关闭会话		POST /v2/ivh/sessionmanager/sessionmanagerservice/closesession	关闭会话，停止数智人推流，释放并发。
	查询会话列表	查询 uin 下的会话列表	POST /v2/ivh/sessionmanager/sessionmanagerservice/listsessionofuin	用于查询某个uin账号所有进行中的会话列表。
		查询数智人项目下的会话列表	POST /v2/ivh/sessionmanager/sessionmanagerservice/listsessionofprojectid	用于查询数智人项目下的所有进行中的会话列表。
		查询个人资产形象下的会话列表	POST /v2/ivh/sessionmanager/sessionmanagerservice/listsessionofassetvk	用于查询个人资产形象下的所有进行中的会话列表。
数智人指令驱动				
❗ 说明： 指令驱动 适用场景 参考 驱动 指令说明		创建长连接通道	WSS /v2/ws/ivh/interactdriver/interactdriverservice/commandchannel	建立 websocket 长连接通道，上行发送驱动指令、接收下行消息。
		文本驱动指令 (WebSocket)		建立长连接通道后，可以通过 WebSocket 长连接发送文本驱动数智人。
		流式文本驱动指令 (WebSocket)		建立长连接通道后，可以通过 websocket 长连接发送流式文本驱动数智人。
		音频驱动指令 (WebSocket)		建立长连接通道后，可以通过 websocket 长连接发送音频驱动数智人。
		心跳指令 (WebSocket)		在超过10分钟后如果云端未收到任何驱动指令会自动断开连接并释放并发。创建长通道后，通过发送心跳指令，可对会话进行续期。发送间隔大于30s，小于10分钟
直接驱动(HTTPS)		文本驱动指令 (HTTP 短连接)	POST /v2/ivh/interactdriver/interactdriverservice/command	HTTP方式发送文本驱动指令 ❗ 说明：

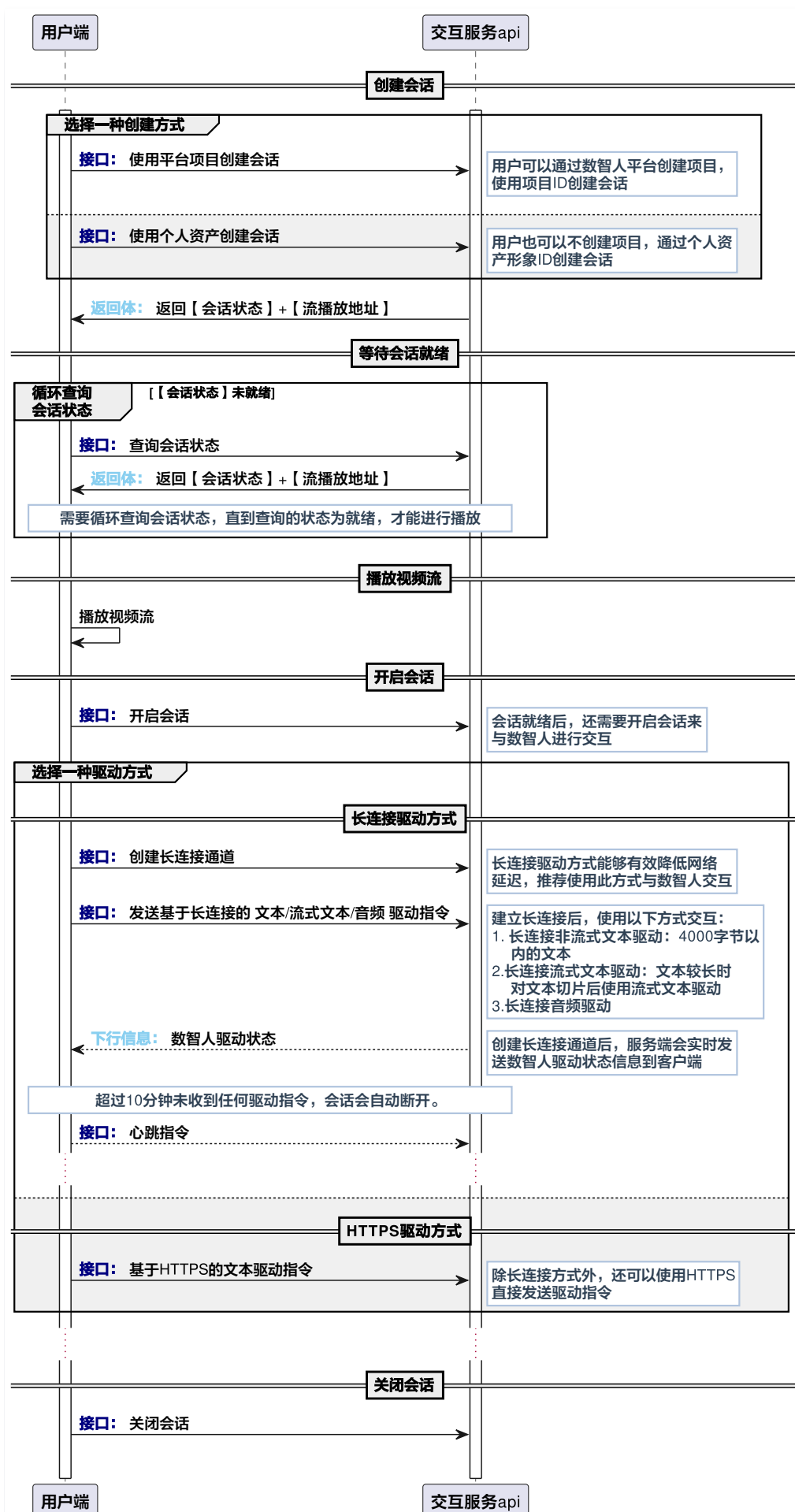
				不用建立长连接通道，常用于试用、一句话驱动等场景。
其他接口	实时更新数智人画面参数	POST /v2/ivh/sessionmanager/sessionmanagerservice/updatesessionconfig	用于实时更新视频流画面中的人物大小、位置、背景图。此接口不再建议使用，后续不再继续迭代新特性。	
	获取进入 TRTC 房间密钥	POST /v2/ivh/sessionmanager/sessionmanagerservice/gettrtcsign	当创建直播会话时选择了 trtc 协议，除数字人之外的其他用户若需加入房间使用此接口获取该用户的入房密钥。	

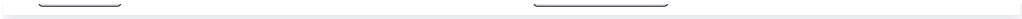
接口接入流程

接入指引

- 通过下列步骤获取交互数智人服务：
- 1. 购买相关产品及服务：[购买链接](#)；
 - 2. 创建会话：使用下列两种方式之一与云端创建会话，获取数智人视频流[会话状态](#)与[流播放地址](#)：
 - [使用数智人平台项目创建会话](#)
 - [使用形象资产创建会话](#)
 - 3. 等待流就绪：当[会话状态](#)为准备中，代表数智人模型正在加载中，时长一般不会超过两分钟。通过调用[查询会话状态](#)接口循环查询，直到[会话状态](#)为进行中。[会话状态](#)状态码说明见[查询会话状态](#)；
 - 4. 播放视频流；
 - 5. 开启会话：视频流就绪后，需要调用[开启会话](#)接口，使数智人可以通过指令进行驱动；
 - 6. 发送驱动指令：各种驱动指令说明及要求请参考[驱动指令说明](#)；
 - 7. 关闭会话：交互完成后请[关闭会话](#)。

接入流程





会话管理

创建会话

使用数智人平台项目创建会话

最近更新时间：2025-07-03 16:23:52

接口说明

使用数智人平台项目创建会话，获取数智人画面的流播放地址。

调用协议

- HTTPS + JSON
- POST /v2/ivh/sessionmanager/sessionmanagerservice/createsession
- Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识，长度为32的 UUID。
SessionId	String	否	会话唯一标识，公有云环境不允许传入由云端分配，私有化环境可传入。
VirtualmanProjectId	String	是	数智人项目 ID。 
UserId	String	是	用户的唯一标识，由调用方自己维护，以相同的 UserId 创建新流，会导致上一个该 UserId 流关闭。 ⚠ 注意： 当使用 TRTC 协议时，该参数为数智人进房用户，调用方自身进房用户不能与该用户相同，如果相同会将数智人踢出房间，导致断流。
Protocol	String	是	视频流协议，当前支持参数：RTMP、TRTC、WebRTC，不同之处请参见 交互数智人 API 接入常见问题 。
DriverType	int	是	数智人驱动方式。 1：纯文本驱动。 3：音频驱动（原声），此模式下支持发送音频和文本两种驱动模式。
ProtocolOption	ProtocolOption	否	协议自定义参数。 • Protocol 选择 TRTC 则需要传入该参数下的 TRTC 相关字段。 • 若用户需要自定义云直播推流地址，则不论哪种 Protocol，都需要传入该参数下的 CssCustomPushUrl 字段。
ExtraInfo	ExtraInfo	否	建流扩展参数。

StreamMaxInterval	int	否	流式文本非子句模式或音频驱动时数据上送最大间隔时长，范围为2000 – 6000，默认为2000，单位为 ms。
-------------------	-----	---	--

ProtocolOption

名称	类型	必选	描述
TrtcUseExternalApp	boolean	否	是否使用外部 TRTC AppId，如果不使用，将使用数智人平台统一的 TRTC AppId。 ⚠ 注意： 数智人平台统一的 TRTC AppId，仅用于调试阶段，投产时，由用户自行在腾讯云申请 TRTC AppId。
TrtcAppId	String	否	TRTC AppId（使用外部 TRTC AppId 时必填），获取方式请参见 实时音视频-应用概览 。
TrtcRoomId	int	否	TRTC 数字房间号（使用外部 TRTC AppId 时 TrtcRoomId 和 StrTrtcRoomId 必填一个，如果两个都填写，优先使用数字房间号）。
TrtcStrRoomId	String	否	TRTC 字符串房间号（使用外部 TRTC AppId 时 TrtcRoomId 和 StrTrtcRoomId 必填一个，如果两个都填写，优先使用数字房间号）。
TrtcAutoGenRoomIdType	int	否	当 TrtcRoomId 和 TrtcStrRoomId 都不填，系统自动生成的房间号类型（前提是使用数智人平台统一的 TRTC AppId），0：数字类型 1：字符串类型，默认数字类型。
TrtcUserSig	String	否	TRTC 数智人用户签名（使用外部 TRTC AppId 时必填，参数 UserId 为 TRTC 数智人用户，数智人需要使用该签名才能加入房间），计算方式请参见 实时音视频-用户鉴权 。
TrtcPrivateMapKey	String	否	TRTC 数智人用户权限密钥（使用外部 TRTC AppId 时必填）。
CssCustomPushUrl	String	否	自定义云直播推流地址，推流协议固定使用 RTMP，使用 RTMP 推流给云直播，可以使用云直播支持的各类拉流协议进行拉流播放。 - 要求的推流地址格式为： <code>rtmp://xyz.com/cssAppName/streamid?txSecret={0}&txTime={1}</code> - txSecret 和 txTime 计算方式参考： <code>https://cloud.tencent.com/document/product/267/32720</code>

ExtralInfo

名称	类型	必选	描述
AlphaChannelEnable	bool	否	是否开启原始视频拼接 alpha 输出拼接视频流。 ⓘ 说明： 假设原始输出宽度为 iW，则实际视频流宽度为：iW + pW + iW - pW 为中间间隔(pad)宽度：$(16 - iW \% 16) \% 16$ - alpha 视频宽度起始点：iW + pW - alpha 视频宽度：iW ⚠ 注意： 2D 小样本（高精）形象暂不支持该特性。

返回参数

名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识。
SessionId	String	是	会话的唯一标识。
SessionStatus	int	是	会话状态： 1：进行中（已就绪）：命中缓存直接得到播流地址。 3：准备中（未就绪）：未命中缓存需等待模型加载，需要使用 查询会话状态 接口轮询流状态，直到流状态变为1为止。一般不超过2分钟。
PlayStreamAddr	String	否	播放地址。 ❗ 说明： 如果通过 CssCustomPushUrl 参数指定了自定义云直播推流地址，则不返回此字段。

请求示例（RTMP 协议）

```
{
  "Header": {
  },
  "Payload": {
    "ReqId": "f2612aa810014e8997f95bda97917268",
    "VirtualmanProjectId": "1b2c9f1c268a4edfbf6c6274da31b890",
    "UserId": "virtualhuman",
    "Protocol": "rtmp",
    "DriverType": 1
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "Message": "",
    "RequestID": "3c5a8d999d8e1fbc6c574595a1b31d5d"
  },
  "Payload": {
    "ReqId": "f2612aa810014e8997f95bda97917268",
    "SessionId": "m318552187863054171",
    "SessionStatus": 1,
    "PlayStreamAddr": "rtmp://live.qq.com/live/m789"
  }
}
```

请求示例（TRTC 协议，使用数智人平台统一的 TRTC AppId）

```
{
  "Header": {
  },
  "Payload": {
    "ReqId": "f2612aa810014e8997f95bda97917268",
```



```
    "VirtualmanProjectId": "1b2c9f1c268a4edfbf6c6274da31b890",
    "UserId": "virtualhuman",
    "Protocol": "trtc",
    "DriverType": 1,
    "ProtocolOption": {
        "TrtcUseExternalApp": false,
        "TrtcRoomId": 12345
    }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "6cf04886b38e5855a86b64e945068ca6",
    "SessionID": "gz6240ef0317504134763113857",
    "DialogID": "",
    "Code": 0,
    "Message": ""
  },
  "Payload": {
    "SessionId": "m587049736339457633",
    "PlayStreamAddr": "trtc://12345?
userId=virtualhuman&appId=1400695865&userSig=eJwszcsOgjAQheF3mbUhA5SiTVwYWXmLWN2wwzDgeCG1bQxqfHcTYPt-
yTlFOG50QJ1hS6DSCHHSF66o9VwzWVDgyXnEGAZy1a00hitQoUCUs2Qqk1G4AQVPfc7Wju52a*vF6toKXXhZ7N*dzj*nC2K2LEVeHAz
tmvk46flBoMI0QRHGIpVDffXnUYDw*wcAAP--sGoxjA__&privateMapKey=eJxEjl9LwzAUR7-
LfZZxs6XNFvAhpNbinD5Yuudo7kYYzpg-
tlb87kJb8Pd4Dhx*P9A*vqxo8C4QSLFGvJmIs3RN7uQogIREMSFuYFbRXoz3zoJkHLHcFduyWIw7gwTdVqHdXfuHZkSdjxTu3rz55HF
MWH-U*3PWHm02*vk73i7J5N4JJBMFcrbhoppjRe8wkKkKUae9*Nt1L9gT1tRXVpjmLfmb07DOP-S*5rer1eIfz*BQAA---
uwULB&virtualManUserId=123450005_ivh_anchor",
    "SessionStatus": 3,
    "ReqId": "ac8ffc41970a0da06bb87ce3471282be"
  }
}
```

请求示例（TRTC 协议，使用外部 TRTC AppId）

```
{
  "Header": {
  },
  "Payload": {
    "ReqId": "f2612aa810014e8997f95bda97917268",
    "VirtualmanProjectId": "1b2c9f1c268a4edfbf6c6274da31b890",
    "UserId": "virtualhuman",
    "Protocol": "trtc",
    "DriverType": 1,
    "ProtocolOption": {
        "TrtcUseExternalApp": true,
        "TrtcAppId": "123322865",
        "TrtcRoomId": 12345,
        "TrtcUserSig":
        "eJyrVgrxCdYrSy1SslIy0jNQ0gHzM1NS80oy0zLBwiWpXSBsZQqeKU7MSCgswUJStDEwMDM0tTCzNTiExqRUFmUSpQ3NTU1MjAwAAi
        WpKZCxIzN7EAQnMLE6gpmelAk11989Iik4r8s-xMHQtyAzOcC51LKwPyS3LcAhyznd2jHMNCDSLDC7XDwz1t1WoBf4UyHg__",
        "TrtcPrivateMapKey": "123123"
    }
  }
}
```

```
}  
}
```

返回示例

```
{  
  "Header": {  
    "RequestID": "2d46357a566aa584b75f92f295aa5d50",  
    "SessionID": "gza802cc9317504133607701099",  
    "DialogID": "",  
    "Code": 0,  
    "Message": ""  
  },  
  "Payload": {  
    "SessionId": "m586856211152897989",  
    "SessionStatus": 3,  
    "ReqId": "ad262ea663c84768591a086d6c2e96fd"  
  }  
}
```

使用形象资产创建会话

最近更新时间：2025-07-03 16:23:52

接口说明

可使用客户有权限的形象资产直接创建会话，获取数智人画面的流播放地址。

调用协议

- HTTPS + JSON
- POST /v2/ivh/sessionmanager/sessionmanagerservice/createsessionbyasset
- Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识，长度为32的 UUID。
SessionId	String	否	会话唯一标识。公有云环境不允许传入由云端分配，私有化环境可传入。
AssetVirtualmanKey	String	是	形象资产 ID。从数智人平台的资产管理中获取，如图： 
UserId	String	是	用户的唯一标识，由调用方自己维护，以相同的 UserId 创建新流，会导致上一个该 UserId 流关闭。 ⚠ 注意： 当使用 TRTC 协议时，该参数为数智人进房用户，调用方自身进房用户不能与该用户相同，如果相同会将数智人踢出房间，导致断流。
Protocol	String	是	视频流协议，当前支持参数：RTMP、TRTC、WebRTC，不同之处请参见 交互数智人 API 接入常见问题 。
DriverType	int	是	数智人驱动方式。 1：纯文本驱动。 3：音频驱动（原声），此模式下支持发送音频和文本两种驱动模式。
ProtocolOption	ProtocolOption	否	协议自定义参数，同 使用数智人平台项目创建会话 参数 ProtocolOption。Protocol 选择 TRTC 则需要传入该参数下的 TRTC 相关字段；若用户需要自定义云直播推流地址，则不论哪种 Protocol，都需要传入该参数下的 CssCustomPushUrl 字段。
SpeechParam	SpeechParam	否	音色相关参数。
ExtralInfo	ExtralInfo	否	建流扩展参数，同 使用数智人平台项目创建会话 参数 ExtralInfo。
Backgrounds	Arrays of [VideoLayer]	否	背景，背景和前景总数不能超过5。 📌 说明： - 对于未抠除拍摄背景的主播，此参数无法生效。 - 对于抠除拍摄背景的主播，背景生效优先顺序为：

			2.1 此参数设置的背景。 2.2 绿幕。
Foregrounds	Arrays of [VideoLayer]	否	前景，背景和前景总数不能超过5
StreamMaxInterval	int	否	流式文本非子句模式或音频驱动时数据上送最大间隔时长，范围为2000 – 6000，默认为2000，单位为 ms。

ProtocolOption

名称	类型	必选	描述
TrtcUseExternalApp	boolean	否	是否使用外部 TRTC AppId，如果不使用，将使用数智人平台统一的 TRTC AppId。 ⚠ 注意： 数智人平台统一的 TRTC AppId，仅用于调试阶段，投产时，由用户自行在腾讯云申请 TRTC AppId。
TrtcAppId	String	否	TRTC AppId（使用外部 TRTC AppId 时必填），获取方式请参见 实时音视频-应用概览 。
TrtcRoomId	int	否	TRTC 数字房间号（使用外部 TRTC AppId 时 TrtcRoomId 和 StrTrtcRoomId 必填一个，如果两个都填写，优先使用数字房间号）。
TrtcStrRoomId	String	否	TRTC 字符串房间号（使用外部 TRTC AppId 时 TrtcRoomId 和 StrTrtcRoomId 必填一个，如果两个都填写，优先使用数字房间号）。
TrtcAutoGenRoomIdType	int	否	当 TrtcRoomId 和 TrtcStrRoomId 都不填，系统自动生成的房间号类型（前提是使用数智人平台统一的 TRTC AppId），0：数字类型 1：字符串类型，默认数字类型。
TrtcUserSig	String	否	TRTC 数智人用户签名（使用外部 TRTC AppId 时必填，参数 UserId 为 TRTC 数智人用户，数智人需要使用该签名才能加入房间），计算方式请参见 实时音视频-用户鉴权 。
TrtcPrivateMapKey	String	否	TRTC 数智人用户权限密钥（使用外部 TRTC AppId 时必填）。
CssCustomPushUrl	String	否	自定义云直播推流地址，推流协议固定使用 RTMP，使用 RTMP 推流给云直播，可以使用云直播支持的各类拉流协议进行拉流播放。 - 要求的推流地址格式为： <code>rtmp://xyz.com/cssAppName/streamid?txSecret={0}&txTime={1}</code> - txSecret 和 txTime 计算方式参考： <code>https://cloud.tencent.com/document/product/267/32720</code>

SpeechParam

名称	类型	必选	描述
Speed	float	否	语速（1.0为正常语速，范围[0.5–1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快，DriverType 为音频驱动类型时，语速控制不生效）
TimbreKey	String	否	音色 key，默认使用形象自有音色。
Volume	int	否	音量大小，范围[0, 10]，对应音量大小。默认为0，代表正常音量，值越大音量越高。
TimbreLanguage	String	否	音色语种，可选语种参考个人资产管理 API 分页查询音色列表 接口，多语种音色在合成时必须选择对应语种。

VideoLayer

参数名称	类型	必选	描述
URL	String	是	图层资源地址。 - 图片格式：支持 jpg、jpeg、png、mjpeg、gif。 - 图片大小：建议不超过2M。
X	int	是	资源放置的左上 X 坐标。
Y	int	是	资源放置的左上 Y 坐标。
Width	int	是	资源输出的宽度。
Height	int	是	资源输出的高度。

返回参数

名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识。
SessionId	String	是	会话的唯一标识。
SessionStatus	int	是	会话状态： 1：进行中（已就绪），命中缓存直接得到播流地址。 3：准备中（未就绪），未命中缓存需等待模型加载，需要使用 查询会话状态 接口轮询流状态，直到流状态变为1为止。一般不超过2分钟。
PlayStreamAddr	String	否	播放地址。 说明： 如果通过 <code>CssCustomPushUrl</code> 参数指定了自定义云直播推流地址，则不返回此字段。

请求示例（RTMP 协议）

```
{
  "Header": {
  },
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "AssetVirtualmanKey": "s34d23f4fd5hvdbsqfd5be5g8c5fg7h1",
    "DriverType": 1,
    "UserId": "virtualhuman",
    "Protocol": "rtmp"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "f87643e2e0bb5af25127bc4c474caf2a",
    "SessionID": "gz6240ef0317515200540972438",
    "DialogID": "",
    "Code": 0,
    "Message": ""
  },
}
```

```

    "Payload": {
      "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
      "SessionId": "m318552187863054171",
      "SessionStatus": 1,
      "PlayStreamAddr": "rtmp://liveplay.ivh.qq.com/live/m318552187863054171"
    }
  }
}

```

请求示例（TRTC 协议，使用数智人平台统一的 TRTC AppId）

```

{
  "Header": {

  },
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "AssetVirtualmanKey": "s34d23f4fd5hvdbsqfd5be5g8c5fg7h1",
    "DriverType": 1,
    "UserId": "virtualhuman",
    "Protocol": "trtc"
    "ProtocolOption": {
      "TrtcUseExternalApp": false,
      "TrtcRoomId": 12345
    }
  }
}

```

返回示例

```

{
  "Header": {
    "RequestID": "5c586477ffaf5d53d0ecdb2bc022c446",
    "SessionID": "gz6240ef0317504138822405697",
    "DialogID": "",
    "Code": 0,
    "Message": ""
  },
  "Payload": {
    "SessionId": "m5877303041064972016",
    "PlayStreamAddr": "trtc://12345?userId=virtualhuman&appId=1400695865&userSig=eJwszk8LgjAYx-H38pxD5nJuDrp4icAKW0F1G7jsyT*pbVpE7z1Qr98P-Ph94Zg7wb7AxITglZjAUzU1u8oelAQo*ddbq8u0rXMPkrK3TTYAbSDwgJIyZCNgvmIMGps7EfqoQPKneHbFFkml7j91MKFae83a2jm08f4YmVm9U8abEyIH3OSOAvhaBT7ccH1CPw*wcAAP--P9Mz0Q__&privateMapKey=eJxEj1FrwjAURv-LfR5ym1pbA3sIiDpmmc6qz8kS06up69JEN8f**0A79j2eAwe*b6gW64H5bMkb4DlDfR0uYUaE-GA4cz*RClq2MjT3D3nT7KtiUNPBkijSZZMcp6QxY45CFVYSnZ4W2*ujxV7vV6Ocrnta7K2c6inR7Y0i4m44*iwsc*GagxwJM8w2GSFgW709gZr*IeOAghSt24Lz3f1mpWJ2o3jZv02r2shCgn7n27yaT431-2fLvABgg-vwEAAP--Q*VH*g__&virtualManUserId=12345_ivh_anchor",
    "SessionStatus": 3,
    "ReqId": "c81f28f73ddeec64b0738e0a316e65d6"
  }
}

```

请求示例（TRTC 协议，使用外部 TRTC AppId）

```

{
  "Header": {

```

```

    },
    "Payload": {
      "ReqId": "f2612aa810014e8997f95bda97917268",
      "AssetVirtualmanKey": "0fee10a003ea4c80a36c5b11ea8082f8",
      "UserId": "virtualhuman",
      "Protocol": "trtc",
      "DriverType": 1,
      "ProtocolOption": {
        "TrtcUseExternalApp": true,
        "TrtcAppId": "123322865",
        "TrtcRoomId": 12345,
        "TrtcUserSig":
"eJyrVgrxCdYrSy1SslIy0jNQ0gHzM1NS80oy0zLBwiWpxSUGBsZQqeKU7MSCgswUJStDEwMDM0tTNTiExqRUFmUSpQ3NTU1MjAwAAiWpKZCxIzN7EAQnMLE6gpmelAk11989Iik4r8s-xMHQtyAzOcC51LKwPyS3LcAhyznd2jHMNCDSLDC7XDwz1t1WoBf4UyHg__",
        "TrtcPrivateMapKey": "123123"
      }
    }
  }
}

```

返回示例

```

{
  "Header": {
    "RequestID": "2d46357a566aa584b75f92f295aa5d50",
    "SessionID": "gza802cc9317504133607701099",
    "DialogID": "",
    "Code": 0,
    "Message": ""
  },
  "Payload": {
    "SessionId": "m586856211152897989",
    "SessionStatus": 3,
    "ReqId": "ad262ea663c84768591a086d6c2e96fd"
  }
}

```

查询会话状态

最近更新时间：2024-11-29 16:28:52

接口说明

查询指定会话的当前状态。

调用协议

HTTPS + JSON
POST /v2/ivh/sessionmanager/sessionmanagerservice/statssession
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	必选	类型	描述
ReqId	是	String	单次请求唯一标识，长度为32的 uuid。
SessionId	是	String	会话唯一标识。

返回参数

参数名称	必选	类型	描述
ReqId	是	String	单次请求唯一标识。
SessionStatus	是	int	会话状态。1：进行中(已就绪)，2：已关闭，3：准备中(未就绪)，4：创建会话失败
PlayStreamAddr	否	string	流播放地址。
SpeakStatus	否	string	数智人状态： - Initial：初始状态。 - WaitingTextStart：等待文本播报开始。 - TextStart：文本播报中。 - WaitingTextOver：等待文本播报结束。 - TextOver：文本播报结束。 - WaitingAudioStart：等待音频开始播报。 - AudioStart：音频开始播报。 - WaitingAudioOver：等待音频播报结束。 - AudioOver：音频播报结束。 - Error：驱动错误（与 TextOver 和 AudioOver 含义一样表示一次驱动的终态），仅表示最近一次驱动失败，不影响继续发送驱动命令。
IsSessionStarted	是	bool	是否已经开启会话，只有开启状态，才可以发送驱动指令。 ⚠ 注意： 1. 返回 false 时，需要调用开启会话接口进行开启。 2. 一旦会话被开启过，不论当前会话状态如何，该字段都会返回 true。
ErrorCode	是	int	当前状态对应的错误码，0为正常。错误码列表参考 错误码
ErrorMessage	否	string	当前状态对应的错误信息。

请求示例

```
{
```



```
"Header": {},
"Payload": {
  "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
  "SessionId": "m318552187863054171"
}
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "Message": "",
    "RequestId": "s34d23f4fd5hvdbsqfd5be5g8c5fg7h1"
  },
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "SessionStatus": 1,
    "PlayStreamAddr": "rtmp://live.qq.com/live/m789",
    "SpeakStatus": "TextOver",
    "IsSessionStarted": true,
    "ErrorCode": 0,
    "ErrorMessage": ""
  }
}
```

开启会话

最近更新时间：2024-11-11 21:58:52

接口说明

[查询会话状态](#)直到流就绪之后，调用此接口开启会话，之后才可以发送驱动指令。

调用协议

HTTPS + JSON
POST /v2/ivh/sessionmanager/sessionmanagerservice/startsession
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识，长度为32的uuid。
SessionId	String	是	会话唯一标识。

返回参数

名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识。

请求示例

```
{
  "Header": {},
  "Payload": {
    "SessionId": "m318552187863054171",
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "Message": "",
    "RequestID": "aeeac7e4cf1b3db317c90db91acb6267"
  },
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf"
  }
}
```

关闭会话

最近更新时间：2024-11-29 16:28:52

接口说明

关闭直播会话，停止数智人推流，释放并发。

调用协议

HTTPS + JSON
POST /v2/ivh/sessionmanager/sessionmanagerservice/closesession
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识，长度为32的 uuid。
SessionId	String	是	会话唯一标识。

返回参数

名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识。

请求示例

```
{
  "Header": {},
  "Payload": {
    "SessionId": "m318552187863054171",
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "Message": "",
    "RequestID": "m123adfafvbadsafd"
  },
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf"
  }
}
```

查询会话列表

查询 uin 下的会话列表

最近更新时间：2024-11-29 16:28:52

接口说明

用于查询某个 uin 账号所有进行中的会话。

调用协议

HTTPS + JSON
POST /v2/ivh/sessionmanager/sessionmanagerservice/listsessionofuin
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识，长度为32的 uuid。

返回参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识。
Sessions	Session	是	会话列表数组。

Session

名称	类型	必选	描述
UserId	String	是	用户的唯一标识。
SessionId	String	是	会话的唯一标识。
SessionStatus	int	是	会话状态。1：进行中(已就绪)，2：已关闭，3：准备中(未就绪)，4：建流失败
PlayStreamAddr	String	否	流播放地址，建流成功后返回。
DriverType	String	是	驱动类型。1：文本驱动；3：音频驱动(原声)。
IsSessionStarted	bool	是	是否已经开启会话，只有开启状态，才可以发送驱动指令。
VirtualmanType	String	是	数智人类型名称。
CreateType	int	是	创建会话的方式。0：使用数智人项目ID创建会话；1：使用形象资产 ID 创建会话。
VirtualmanKey	string	否	数智人 VirtualmanKey，形象唯一标识。
AssetVirtualmanKey	string	否	形象资产 ID。
VirtualmanProjectId	string	否	数智人项目 ID。

请求示例

```
{
  "Header": {},
}
```

```
{
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "Message": "",
    "RequestID": "3c5a8d999d8e1fbc6c574595a1b31d5d"
  },
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "Sessions": [
      {
        "UserId": "virtualhuman",
        "SessionId": "m318552187863054171",
        "SessionStatus": 1,
        "PlayStreamAddr": "rtmp://live.qq.com/live/m789",
        "DriverType": 1,
        "IsSessionStarted": true,
        "VirtualmanType": "small_sample_2d",
        "CreateType": 0,
        "VirtualmanKey": "5c298g1ba5b34b4a8fb0806fe8fb3e6a",
        "AssetVirtualmanKey": "",
        "VirtualmanProjectId": "5c298g1ba5b34b4a8fb0806fe8fb3e6a"
      }
    ]
  }
}
```

查询数智人项目下的会话列表

最近更新时间：2024-11-29 16:28:52

接口说明

用于查询数智人项目下的所有进行中的会话列表。

调用协议

HTTPS + JSON
POST /v2/ivh/sessionmanager/sessionmanagerservice/listsessionofprojectid
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识，长度为32的 uuid。
VirtualmanProjectId	String	是	数智人项目 ID。获取方式如图： 

返回参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识。
Sessions	Session	是	会话列表数组

Session

名称	类型	必选	描述
UserId	String	是	用户的唯一标识。
SessionId	String	是	会话的唯一标识。
SessionStatus	int	是	会话状态。1：进行中(已就绪)，2：已关闭，3：准备中(未就绪)，4：建流失败
PlayStreamAddr	String	否	流播放地址，建流成功后返回
DriverType	String	是	数智人类型。1：文本驱动；3：语音驱动(原声)。
IsSessionStarted	bool	是	是否已经开启会话，只有开启状态，才可以发送驱动指令。
VirtualmanType	String	是	数智人类型名称。
CreateType	int	是	创建会话的方式。0：使用数智人项目ID创建会话；1:使用形象资产ID创建会话。
VirtualmanKey	string	是	数智人 VirtualmanKey，形象唯一标识。

AssetVirtualmanKey	string	否	形象资产 ID。
VirtualmanProjectId	string	是	数智人项目 ID。

请求示例

```
{
  "Header": {},
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "VirtualmanProjectId": "5c298g1ba5b34b4a8fb0806fe8fb3e6a"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "Message": "",
    "RequestId": "3c5a8d999d8e1fbc6c574595a1b31d5d"
  },
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "Sessions": [
      {
        "UserId": "virtualhuman",
        "SessionId": "m318552187863054171",
        "SessionStatus": 1,
        "PlayStreamAddr": "rtmp://live.qq.com/live/m789",
        "DriverType": 1,
        "IsSessionStarted": true,
        "VirtualmanType": "small_sample_2d",
        "CreateType": 0,
        "VirtualmanKey": "5c298g1ba5b34b4a8fb0806fe8fb3e6a",
        "AssetVirtualmanKey": "",
        "VirtualmanProjectId": "5c298g1ba5b34b4a8fb0806fe8fb3e6a"
      }
    ]
  }
}
```

查询个人资产形象下的会话列表

最近更新时间：2024-11-29 16:28:52

接口说明

用于查询个人资产形象下的所有进行中的会话列表。

调用协议

HTTPS + JSON
POST /v2/ivh/sessionmanager/sessionmanagerservice/listsessionofassetvk
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识，长度为32的 uuid。
AssetVirtualmanKey	String	是	个人资产形象 ID。从数智人平台的资产管理中获取，如图： 

返回参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识。
Sessions	Session	是	会话列表数组。

Session

名称	类型	必选	描述
UserId	String	是	用户的唯一标识。
SessionId	String	是	会话的唯一标识。
SessionStatus	int	是	状态。1：进行中(已就绪)，2：已关闭，3：准备中(未就绪)，4：建流失败
PlayStreamAddr	String	否	流播放地址，建流成功后返回。
DriverType	String	是	驱动类型。1：文本驱动；3：语音驱动(原声)。
IsSessionStarted	bool	是	是否已经开启会话，只有开启状态，才可以发送驱动指令。
VirtualmanType	String	是	数智人类型名称。
CreateType	int	是	创建会话的方式。0：使用数智人项目ID创建会话；1：使用形象资产 ID 创建会话。
VirtualmanKey	string	否	数智人 VirtualmanKey，形象唯一标识。
AssetVirtualmanKey	string	否	形象资产 ID。

VirtualmanProjectId	string	否	数智人项目 ID。
---------------------	--------	---	-----------

请求示例

```
{
  "Header": {},
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "AssetVirtualmanKey": "7c298g1ba5b34b4a8fb0806fe8fbge6a"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "Message": "",
    "RequestId": "3c5a8d999d8e1fbc6c574595a1b31d5d"
  },
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "Sessions": [
      {
        "UserId": "virtualhuman",
        "SessionId": "m318552187863054171",
        "SessionStatus": 1,
        "PlayStreamAddr": "rtmp://live.qq.com/live/m789",
        "DriverType": 1,
        "IsSessionStarted": true,
        "VirtualmanType": "small_sample_2d",
        "CreateType": 1,
        "VirtualmanKey": "",
        "AssetVirtualmanKey": "7c298g1ba5b34b4a8fb0806fe8fbge6a",
        "VirtualmanProjectId": ""
      }
    ]
  }
}
```

数智人驱动指令

驱动指令说明

最近更新时间：2025-05-22 17:37:52

指令说明

驱动方式	指令类型	驱动指令		适用场景
长连接驱动方式 (WebSocket)	音频指令	音频驱动指令		使用音频驱动数智人。 参考 音频驱动场景实践教程
	文本指令	流式文本驱动指令	子句模式	1. 适用于直播场景； 2. 可支持插播功能。 推荐直接使用直播数智人服务，参考 直播数智人服务 API 文档
			非子句模式	适用于用户自行接入大模型，用大模型的流式输出驱动数智人的场景。 参考 用户自行接入大模型实践教程
		文本驱动指令（WebSocket 长连接）		1. 适用于数智人播报的场景； 2. 适用于在数智人平台绑定了对话服务的场景。 说明： 数智人绑定对话服务请参考 数智人项目的创建与管理 参考 数智人绑定对话服务实践教程
	心跳指令	心跳指令		用于长连接的保活，推荐每1~2分钟发送一次。
直接驱动方式 (HTTPS)	文本指令	文本驱动指令（HTTP 短连接）		1. 适用于短连接版数智人播报的场景； 2. 常用于试用、一句话驱动等场景。

- 说明：**
指令类型：
 - 文本指令：**发送文本指令是发送播报文本给云端，由云端合成语音进而驱动数字人讲话的指令；
 - 音频指令：**发送音频指令是发送语音流式分片给云端，由云端进行原声或变声驱动数字人讲话的指令；
 - 心跳指令：**发送心跳指令是为了维系长连接和会话不断开，如果客户端不发送心跳且无有效上下行数据，云端超过3分钟主动断开长连接，10分钟主动关闭会话断流。
- 名词解释：**
 - 子句与非子句：**子句是构成复杂段落的基本单位，能够表达一个完整的陈述，如“腾讯云智能数智人采用语音交互、虚拟形象模型生成等多项 AI 技术”；非子句是指不构成完整句子结构的片段，如“智”、“能”、“数智”、“人”。
 - 插播：**当已经发送子句A1、A2、A3、A4，数智人服务端正在处理A2时，此时发送插播子句X，数智人服务端处理完A2后，会优先处理X，最终的播报顺序为A1、A2、X、A3、A4。
 - 播报与对话服务：**播报是单向的，用户发送什么内容，数智人就说什么内容；对话是双向的，用户发送问题后，数智人根据数智人平台绑定的对话服务来说出答案。
 - 对话服务：**目前对话服务支持腾讯行业大模型对话以及云小微对话。
 - 腾讯行业大模型：**[腾讯云智能体开发平台](#)（Tencent Cloud Agent Development Platform），是面向企业客户及合作伙伴的，基于大语言模型的知识应用构建平台，结合企业专属数据，提供知识问答等应用范式，更快更高效地完成大模型应用的构建，推动大语言模型在企业服务场景的应用落地。
 - 云小微对话：**即使用 [文本机器人](#) 服务，支持主动询问式多轮对话，结合智能学习能力，实现高准确率人机对话。

驱动方式与指令类型的对应关系

新建流时的 DriverType 参数和指令类型的对应关系如下：

- DriverType=1：文本驱动 支持发送文本指令（SEND_TEXT）、心跳指令（SEND_HEARTBEAT）
- DriverType=3：音频驱动 支持发送文本指令（SEND_TEXT）、发送音频指令（SEND_AUDIO）、心跳指令（SEND_HEARTBEAT）

当DriverType=3时，文本与音频的发送顺序如下：

1. 当文本正在播报，即下行消息中数智人状态SpeakStatus=TextStart（文本播报中）时，如果要发送音频指令，需要发送空文本指令的打断指令（Interrupt设置为true），直至返回下行消息中 SpeakStatus=TextOver（文本播报结束）才可以发送音频。
2. 当文本正在播报时，如果要播报新文本，直接发送新文本的文本指令，当前正在播报的文本被打断完成会收到 TextOver 的下行消息，新文本播报开始会收到 TextStart，播报结束会收到 TextOver 。
3. 当音频驱动进行中，如果要中断音频驱动发送文本指令驱动，需要发送音频指令（IsFinal 设置为 true）结束语音驱动，直至收到 AudioOver 事件之后才可以发文本或者音频。

WebSocket 长连接驱动方式

创建长连接通道

最近更新时间：2024-08-29 15:12:21

接口说明

建立 websocket 长连接通道，上行发送驱动指令、接收下行消息。

调用协议

WebSocket + JSON

WSS /v2/ws/ivh/interactdriver/interactdriverservice/commandchannel

Header Content-Type: application/json;charset=utf-8

建连方式（重要）

websocket url：参考 [数智人 aPaas 接口调用方式](#) 接口调用方式组织 url，建连后可通过连接对象发送驱动指令来控制数智人。

⚠ 注意：

- 建立长链接通道需要使用公共参数requestid来发起连接请求。
- 需额外注意的是 requestid 参数填写会话 id（SessionId）的值。例如 SessionId=m318552187863054171，那么 url 如下：

```
wss://domain/v2/ws/ivh/interactdriver/interactdriverservice/commandchannel?  
appkey=xxx&requestid=m318552187863054171&timestamp=xxx&signature=xxx
```

- websocket 可能会因为网络抖动、云端服务更新等操作产生断连，当检测到断开时只要 session 状态没有关闭则需要重新建连。
- 请在调用 [开启会话](#) 接口后，再建立 websocket 长链接，不然连接不成功。一个会话只允许创建一个长链接通道。

文本驱动指令（WebSocket 长连接）

最近更新时间：2025-05-26 10:37:02

接口说明

创建[长连接通道](#)后，可以通过 websocket 长连接发送文本驱动数智人。

 **注意：**
发送文本间隔必须大于1s。

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次驱动的唯一标识, 32位的 uuid(不包含 ‘-’)。  注意： 每次发送都新生成一个 ReqId。
SessionId	String	是	会话唯一标识。
Command	String	是	SEND_TEXT：发送文本。
Data	Data	是	数据对象

Data（发送文本）

名称	类型	必选	描述
Text	String	否	要播报的文本内容，缺省值：字符串最长4000字节。
Interrupt	Boolean	否	打断标记，会打断当前数智人的播报。  注意： 需要等到 websocket 下行消息返回 TextStart 标记后才可发送打断请求。
ChatCommand	String	否	对话指令类型，当数智人项目绑定客服对话或 腾讯云智能体开发平台 时生效。 NotUseChat：客服对话或智能体开发平台配置的问答知识库不生效，数智人播报发送的内容。 传空或不传：客服对话或智能体开发平台配置的问答知识库生效，数智人播报问答知识库中的答案内容。
ChatRoundId	String	否	对话轮次 id，推荐使用 uuid。缺省值："" 使用腾讯云智能体开发平台对话服务的数智人项目需要传入此字段，用于区分多轮对话。
CloudAiVisitorBizId	String（64）	否	腾讯云智能体开发平台的访客ID。
VideoSeiInfo	String	否	需要在视频流 SEI 中携带的信息，字符串内容为 json 格式。发送后会在视频流播报本次响应时在 SEI 中携带。当前仅支持 2D 形象，视频流中信息说明参见： 视频流嵌入信息说明 。
SmartActionEnabled	bool	否	是否开启智能动作，缺省值：false。为 true 并且输入的文本或者话术增强后的文本没有动作标签则会生成智能动作。

请求示例

```
{
  "Header": {},
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "SessionId": "m318552187863054171",
    "Command": "SEND_TEXT",
    "Data": {
      "Text": "您好，我是人工合成的数智人"
    }
  }
}
```

流式文本驱动指令

最近更新时间：2025-07-03 16:23:52

接口说明

创建长连接通道 后，可以通过 WebSocket 长连接发送流式文本驱动数智人。

流式文本驱动指令分为子句模式与非子句模式，区别如下：

1. **服务端是否会重新组句**：非子句模式下服务端会重新组句，子句模式下服务端不会重新组句。例如非子句模式下，依次发送流式文本数据包：“您”、“好”、“我是”、“数”、“智”、“人”，那么服务端收到后就会重新组句为“您好，我是数智人”；子句模式下则不会重新组句，所以需要发送一句连贯的话语例如“您好，我是数智人”。
2. **是否支持插播**：非子句模式不支持插播，子句模式支持插播能力，设置字段 IsInsertSentence=true 即可。
3. **是否会补 final**：非子句模式下，若在指定时间间隔（默认2秒，可以在创建会话时通过参数 StreamMaxInterval 进行设置）内未发送流式文本驱动指令，那么服务端就会自动补充 final 数据包，建议在一段文本内容结束后及时发送 final 数据包避免出现数智人卡顿的情况。

	服务端是否会重新组句	服务端是否会补 final	是否支持插播
子句模式	×	×	✓
非子句模式	✓	✓	×

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次驱动的唯一标识。 每一段流式文本指定一个32位 UUID 值(不包含 ‘-’)。 ⚠ 注意： 每一段流式文本序列指定一个 ReqId，而不是每一次请求指定一个。 值得注意的是，在发送 final 包时仍然使用的是指定的这一个 ReqId，不需要再生成一个。
SessionId	String	是	会话唯一标识。
Command	String	是	SEND_STREAMTEXT，发送流式文本。
Data	Data	是	数据对象

Data

名称	类型	必选	描述
Text	string	是	流式文本内容，只需要发送增量的文本。每个片包字符串长度限制2000字节。 ⚠ 注意： 子句模式发送要求： 1. 输入文本应为完整的句子序列，例如以，。！？等标点符号分割的文本。 2. 单个子句长度应大于10个汉字或单词。 非子句模式发送要求： 1. 流式非子句不支持 Text 中带 SSML 标签。 2. 发送速度不能过慢，过慢会导致出现非预期的 SpeakStatus，推荐发送速度与大模型返回数据的速度一致。
Seq	int	是	流式文本片包序号，序号必须从1开始。

IsFinal	bool	否	用于标记本次流式文本驱动是否结束，默认值 false。 ⚠ 注意： 1. 发送 final 包的时机： 当流式文本驱动指令的文本内容 Text 都已发送完毕后，需要再发送一个 IsFinal=true 的驱动指令（这里简称 final 包，Text 字段可带有文本内容也可以是空串），来结束当次流式文本驱动，使数字人回到静默状态。 2. 没有及时发送 final 包会出现的问题： 对于流式文本驱动指令，包含子句模式与非子句模式，都可能会出现数智人播报期间卡顿的问题。
Interrupt	bool	否	强制打断标记，缺省值：false。传入 true 并且 Text 为空，表示打断流式文本。 ⚠ 注意： 对于如何发送强制打断的驱动指令（Interrupt=true），在 Data 中只需要填充 Seq 以及 Interrupt 的数据，其他字段置空处理即可，注意不可省去 Seq（设置为非0即可），不然会打断无效，参考示例三。
SmartActionEnabled	bool	否	是否开启智能动作，缺省值：false。为 true 并且输入的文本或者话术增强后的文本没有动作标签则会生成智能动作。
IsSentence	bool	否	是否是子句模式，缺省值：false。为 true 服务端不会做重新组句。
IsInsertSentence	bool	否	是否是插入的子句，缺省值：false。为 true 并且是子句模式则表示当前分片需要插播。
VideoSeiInfo	string	否	需要在视频流 SEI 中携带的信息，字符串内容为 JSON 格式。发送后会在视频流播报本次响应时在 SEI 中携带。当前仅支持 2D 形象，视频流中信息说明请参见 视频流嵌入信息说明。

请求示例

示例一：正常的流式文本片包

```
{
  "Header": {},
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "SessionId": "m318552187863054171",
    "Command": "SEND_STREAMTEXT",
    "Data": {
      "Text": "流式文本内容",
      "Seq": 1,
      "IsFinal": false
    }
  }
}
```

示例二：流式文本结束包

```
{
  "Header": {},
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "SessionId": "m318552187863054171",
```



```

    "Command": "SEND_STREAMTEXT",
    "Data": {
      "Text": "流式文本结束了",
      "Seq": 60,
      "IsFinal": true
    }
  }
}

```

示例三：打断正在播报的非子句模式或子句模式的流式文本

```

{
  "Header": {},
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "SessionId": "m318552187863054171",
    "Command": "SEND_STREAMTEXT",
    "Data": {
      "Seq": 40,
      "Interrupt": true
    }
  }
}

```

示例四：正常的流式文本子句模式片包

```

{
  "Header": {},
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "SessionId": "m318552187863054171",
    "Command": "SEND_STREAMTEXT",
    "Data": {
      "Text": "流式文本子句模式内容",
      "Seq": 1,
      "IsFinal": false,
      "IsSentence": true,
      "IsInsertSentence": false//如果是需要插播的子句，则设置该值为true
    }
  }
}

```

音频驱动指令

最近更新时间：2025-07-03 16:23:52

接口说明

创建长连接通道 后，可以通过 WebSocket 长连接发送音频驱动数智人。
发送音频时请特别留意请求参数下方的发送注意事项，不然会导致发送不成功。

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次驱动的唯一标识。 每一段音频指定一个32位 UUID 值(不包含 ‘-’)。 ⚠ 注意： 每一段流式音频序列指定一个 ReqId，而不是单独一次音频分片请求指定一个。 值得注意的是，在发送 final 包时仍然使用的是指定的这一个 ReqId，不需要再生成一个。
SessionId	String	是	会话唯一标识。
Command	String	是	SEND_AUDIO，发送音频。
Data	Data	是	数据对象。

Data

名称	类型	必选	描述
Audio	string	是	音频原始数据的 byte 数组，经 Base64 编码后的字符串。只支持： - 格式：PCM - 采样率：16kHz - 采样位深：16bits - 声道：单声道 可以使用 FFmpeg 工具对音频格式进行调整 FFmpeg。更多请参见 FFmpeg 转化音频格式说明 。
Seq	int	是	音频片包序号，序号必须从1开始。
IsFinal	bool	否	用于标记本次音频驱动是否结束，默认值 false。 ⚠ 注意： 当数据包发送完毕后，必须再发送一个 IsFinal=true 的空数据包（Audio 字段填充空串）结束当次音频驱动使数字人回到静默状态。 没有及时发送 final 包会出现的问题： - 可能会出现数智人播报期间卡顿的问题。 - 后台在指定时间间隔（默认2秒，可以在创建会话时通过参数 StreamMaxInterval 进行设置）内既没收到音频包又没收到 IsFinal=true 的尾包，则后台会自动补一个 IsFinal=true 的尾包，使数智人进入静默状态。
VideoSeiInfo	string	否	需要在视频流 SEI 中携带的信息，字符串内容为 JSON 格式。发送后会在视频流播报本次响应时在 SEI 中携带。当前仅支持 2D 形象，视频流中信息说明请参见 视频流嵌入信息说明 。

⚠ 注意:

1. 如果是麦克风实时收音发送，每录制160ms（5120B）的数据即可发出，中间不需要间隔等待；如果是读取离线音频文件发送，片包大小为160ms（5120B），片包发送间隔为120ms。
2. 最后一个片包大小按实际发送（需小于160ms）。
3. 发送音频实时率要介于[0.75,1]，小于0.75会触发限速，大于1会导致画面卡顿。例如160ms音频片包大小，发送间隔不能低于120ms，不能高于160ms。

请求示例

```
{
  "Header": {},
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "SessionId": "m318552187863054171",
    "Command": "SEND_AUDIO",
    "Data": {
      "Audio": "声音二进制数据编码Base64后的值",
      "Seq": 1,
      "IsFinal": false
    }
  }
}
```

心跳指令

最近更新时间：2025-02-20 10:11:22

接口说明

在超过10分钟后如果云端未收到任何驱动指令会自动断开连接并释放并发。[创建长连接通道](#)后，通过发送心跳指令，可对会话进行续期。

 **注意：**
发送间隔大于30s，小于1分钟。

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识，长度为32的 uuid。
SessionId	String	是	会话唯一标识。
Command	String	是	SEND_HEARTBEAT，发送心跳。
Data	Data	是	数据对象

Data

名称	类型	必选	描述
Text	string	是	固定填写 PING

请求示例

```
{
  "Header": {},
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "SessionId": "m318552187863054171",
    "Command": "SEND_HEARTBEAT",
    "Data": {
      "Text": "PING",
    }
  }
}
```

长连接下行消息

最近更新时间：2025-05-22 17:37:52

创建长连接通道后，数智人服务端将会实时将数智人驱动状态信息返回给客户端。

 **注意：**

当 websocket 重连时，云端会主动推送近3次驱动请求触发的播报状态，以防止客户端在网络抖动导致重连过程中丢失对应的驱动反馈事件。

名称	类型	必选	描述
Type	int	是	下行的消息类型。小样本形象其他返回类型不用关注。 3：播报状态。 4：大模型返回内容。 5：大模型思考过程内容。 9：驱动失败。
SessionId	String	是	视频流会话的唯一标识
ReqId	String	是	一次驱动的唯一标识
Seq	int	是	流式子句的唯一序号
SpeakStatus	String	否	Initial：初始状态。 WaitingTextStart：等待文本播报开始。 TextStart：文本播报中。 WaitingTextOver：等待文本播报结束。 TextOver：文本播报结束。 WaitingAudioStart：等待音频开始播报。 AudioStart：音频开始播报。 WaitingAudioOver：等待音频播报结束。 AudioOver：音频播报结束。 Error：驱动失败（与 TextOver 和 AudioOver 含义一样表示一次驱动的终态），仅表示最近一次驱动失败，不影响继续发送驱动命令。 SentenceNext：流式的子句模式返回该状态，客户端收到该状态表示可以发送下一条子句。 SentenceStart：流式的子句模式返回该状态，客户端收到该状态表示当前子句开始播报。 SentenceOver：流式的子句模式返回该状态，客户端收到该状态表示当前子句播报结束。
FinalType	int	否	当 SpeakStatus 返回 AudioOver 时有值，标记音频播报结束的 Final 包来源： 1-客户输入； 2-服务端超时自动补包
Text	String	否	当 Type 为2、4时有此字段，提问语。
TextPro	String	否	当 Type 为2、4时有此字段，播报文本，带 ssml 标签。
TextDisplay	String	否	当 Type 为2、4时有此字段，播报文本，用于客户端展示。
ContentType	String	否	当 Type 为2、4时有此字段，为2时表示流式非子句输入的内容经过解析后返回的内容，为4时表示大模型返回的内容： 0：未知 1：普通字符串 2：有序列表 3：无序列表 4：图片链接 5：HTTP 链接 6：表格 8：标题 9：SSML
TtsSupport	bool	否	当 Type 为4时有此字段，大模型返回内容是否需要经过 tts 播报。

Final	bool	否	当 Type 为4时有此字段，大模型返回内容是否是最后一个子句。
IsHighLight	bool	否	当 Type 为4时有此字段，大模型返回内容是否需要端上做高亮处理。
Uninterrupt	bool	否	当 Type 为2、4时有此字段，当前播报句是否可打断： true: 不可打断 false: 可打断
Muted	bool	否	当 Type 为2、4时有此字段，当前播报文本是否关闭收音。
InteractionType	String	否	当 Type 为2时有此字段，可用于配置平台自定义，常见类型有：弹窗、图片等。
InteractionContent	String	否	当 Type 为2时有此字段，用于下方弹窗、图片等非文本类的特殊消息。
ErrorCode	int	是	错误码，值为0正常，非0代表错误，例如请求体不符合要求、驱动指令的时机不符合协议约定、请求速率过快等。参考7.错误码列表
ErrorMessage	String	否	错误描述
CloudAiExtra	String	否	腾讯云智能体开发平台返回的扩展字段。返回示例： { "record_id": "b24e6505-be78-4ce9-ad9e-7075dd227994", "references": [{ "id": "1780434945714421760", "type": 2, "url": "https://oaqbot.qidian.qq.com/#/preview/md?id=1780227052790087680", "name": "客户会", "doc_id": "11296" }] } 点赞点踩接口参考 评价消息 参考来源具体字段参考： MsgRecordReference

HTTP 短连接驱动方式

文本驱动指令（HTTP 短连接）

最近更新时间：2025-06-17 16:10:51

接口说明

通过 HTTP 直连的方式发送文本驱动数智人。

 **注意：**
发送文本间隔必须大于1s。

调用协议

HTTPS + JSON
POST /v2/ivh/interactdriver/interactdriverservice/command
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识。
SessionId	String	是	会话唯一标识。
Command	String	是	SEND_TEXT：发送文本。
Data	Data	是	数据对象

Data（发送文本）

名称	类型	必选	描述
Text	string	否	要播报的文本内容，缺省值：""，字符串最长4000。
Interrupt	boolean	否	强制打断标记，缺省值：false。如果不使用强制打断，在数智人平台中该文本被设置为不可打断时，则会收到不可打断的反馈提示。如果使用强制打断，系统会忽略数智人平台中的不可打断设置，直接执行强制打断。
ChatCommand	string	否	对话指令类型，当数智人项目绑定客服对话 bot 时生效。 - NotUseChat：客服 bot 配置的问答知识库不生效，数智人播报发送的内容。 - 传空或不传：客服 bot 配置的问答知识库生效，数智人播报问答知识库中的答案内容。
VideoSeilInfo	string	否	需要在视频流 SEI 中携带的信息，字符串内容为 json 格式。发送后会在视频流播报本次响应时在 SEI 中携带。当前仅支持 2D 形象，视频流中信息说明请参见 视频流嵌入信息说明

返回参数

无需关注返回内容，仅关注返回错误码。

请求示例

```
//发送文本
{
  "Header": {},
  "Payload": {
    "ReqId": "a936625150a66be74afd8f4e5e82b785",
    "SessionId": "m318552187863054171",
```

```

        "Command": "SEND_TEXT",
        "Data": {
            "Text": "您好呀"
        }
    }
}
//绑定客服对话后，不使用对话，播报发送的内容
{
    "Header": {},
    "Payload": {
        "ReqId": "a936625150a66be74afd8f4e5e82b785",
        "SessionId": "m318552187863054171",
        "Command": "SEND_TEXT",
        "Data": {
            "Text": "您好呀",
            "ChatCommand": "NotUseChat" //使用对话的数智人项目建的文本驱动的流，本次驱动纯文本驱动不走对话
        }
    }
}
//强制打断
{
    "Header": {},
    "Payload": {
        "ReqId": "a936625150a66be74afd8f4e5e82b785",
        "SessionId": "m318552187863054171",
        "Command": "SEND_TEXT",
        "Data": {
            "Interrupt": true
        }
    }
}
}

```

返回示例

```

{
    "Header": {
        "Code": 0,
        "Message": "",
        "RequestID": "m123adfafvbadsafd",
    }
}

```


其他接口

实时更新数智人画面参数

最近更新时间：2024-08-22 11:13:31

接口说明

用于实时更新视频流画面中的人物大小、位置、背景图。

说明：
此接口不再建议使用，后续不再继续迭代新特性。当需要设置数智人背景、人物大小时的推荐使用方法：设置绿幕背景推流，自行使用流媒体软件抠绿叠加前后景、缩放人物大小。

注意：
此接口请求频次限制：单一session每5秒可请求一次。

调用协议

HTTPS + JSON
POST /v2/ivh/sessionmanager/sessionmanagerservice/updatesessionconfig
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识。
SessionId	String	是	会话id。
SpaceParam	SpaceParam	否	空间信息，2D小样本支持此参数。
ImageParam	ImageParam	否	形象信息，3D形象支持此参数。

SpaceParam

名称	类型	必选	描述
BackgroundUrl	String	否	背景 URL 地址（小于1M大小，最好与输出分辨率对齐，否则会被拉伸填充）要求：jpg、jpeg、png（其他格式不支持），不传使用上次背景图。
ImagePosition	float	是	人物在背景图中的X位置，人物图片中轴居中X为0，取值范围为[-0.5,0.5]，不可移出背景图范围，线性值。
ImagePositionVertical	float	是	人物在背景图中的Y位置，人物图片贴底Y为0，取值范围为[-0.5,0.25]，向下可以移出背景图，线性值。
ImageZoom	float	是	人物相对于原始素材图的缩放比例，由缩放后的大小和原始人物素材图计算得到的缩放比例。计算公式：缩放后人物图片的宽度/(预览区域背景宽度/目标分辨率宽度)/人物原始素材的宽度。建议取值范围[0.25,2]。

ImageParam

名称	类型	必选	描述
ImageVerticalOffset	float	否	3D形象上半身垂直方向转动角度。

ImageHorizontalOffset	float	否	3D形象上半身水平方向转动角度。
ImageExtraParam	float	否	3D形象变色参数，具体参数可通过资产管理的 查询主播下的所有形象 接口查询。

返回参数

名称	类型	必选	描述
ReqId	String	是	单次请求唯一标识。

请求示例

```
{
  "Header": {},
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "SessionId": "m123adfafvbadsafd",
    "ImageParam": {
      "ImageVerticalOffset": 10,
      "ImageHorizontalOffset": 10,
      "ImageExtraParam": "{\\"JacketColor\\":{\\"colorValue\\":10491928},\\"DressColor\\":{\\"colorValue\\":10491928},\\"shoeColor\\": {\\"colorValue\\": 10491928}}"
    },
    "SpaceParam": {
      "BackgroundUrl": "https://virtualhuman-cos-prod-1251316161.cos.ap-nanjing.myqcloud.com/ai-resource/1518804312073965568",
      "ImagePosition": -0.01,
      "ImagePositionVertical": -0.01,
      "ImageZoom": 1
    }
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "Message": "",
    "RequestID": "123",
  },
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
  }
}
```

获取进入 TRTC 房间密钥

最近更新时间：2025-06-13 14:37:22

接口说明

当创建会话时选择了 TRTC 协议，除数字人之外的其他用户若需加入房间使用此接口获取该用户的入房密钥。

 **注意：**
仅适用于使用数智人内部 TRTC 账号的情况。

调用协议

HTTPS + JSON
POST /v2/ivh/sessionmanager/sessionmanagerservice/gettrtcsign
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	必选	类型	描述
ReqId	是	String	单次请求唯一标识。
SessionId	是	string	会话 id。
RoomId	否	int	int 类型房间 id，RoomId 和 StrRoomId 必填一个，都填时以 int 类型房间号为准。
StrRoomId	否	string	string 类型房间号 id，RoomId 和 StrRoomId 必填一个，都填时以 int 类型房间号为准。
UserId	是	string	加入房间的用户 id。

返回参数

参数名称	必选	类型	描述
ReqId	是	String	单次请求唯一标识。
UserSig	是	string	房间密钥。
PrivateMapKey	是	string	权限密钥。

请求示例

```
{
  "Header": {},
  "Payload": {
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "SessionId": "8da33dd4a662ad5b",
    "RoomId": 1200111,
    "UserId": "avcvdafasfds"
  }
}
```

返回示例

```
{
```

```
"Header": {
  "Code": 0,
  "Message": "",
  "RequestID": "s7aa08das3dd4a662ad5b4508c5b77cg"
},
"Payload": {
  "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
  "UserSig": "vafdasfda",
  "PrivateMapKey": "2ad5be508c"
}
}
```

交互数智人实践教程

最近更新时间：2024-12-12 17:27:52

交互数智人实践教程概述

交互数智人 apaas 接口的调用建议参考以下流程：

1. **初始化 API 调用对象**：设置 API 调用的基础信息，包括 appkey、access token 以及 HTTP 和 WebSocket 的域名。
2. **创建视频流**：使用形象资产 ID 或数智人平台项目 ID 创建视频流，指定视频流协议和驱动类型。
3. **检测会话状态**：循环检查会话状态，直到会话状态变为已就绪（SessionStatus=1）。
4. **开启会话**：调用 API 以开启会话。
5. **创建长链接通道**：建立长链接通道，以便进行持续的通信。启动两个线程，一个用于定期获取下行信息，另一个用于定期发送心跳指令，以保持连接活跃。
6. **发送驱动指令**：根据不同的场景（参见本文档：[1. 数智人绑定对话服务场景](#)、[2. 直播场景](#)、[3. 用户自行接入大模型场景](#)、[4. 音频驱动场景](#)），发送相应的驱动指令，详细的适用场景与驱动指令的关系参考 [驱动指令说明](#)。
7. **关闭长连接通道和会话**：根据不同场景的情况关闭长连接通道和会话。

其中步骤需要根据不同的场景来选择不同的驱动指令，其他步骤的逻辑基本一致。以下是该流程的代码 Demo（目前只提供 go 语言版本）：

```
Go

package main

import (
    "encoding/base64"
    "fmt"
    "io"
    "log"
    "os"
    "strings"
    "time"

    uuid "github.com/satori/go.uuid"
    "k8s.io/apimachinery/pkg/util/rand"
)

func main() {
    //第一步:初始化api调用对象
    var api = NewVHApi(
        "your appkey", //your appkey
        "your accesstoken", //your accesstoken
        "https://gw.tvs.qq.com", //http domain
        "wss://gw.tvs.qq.com") //wss domain
    //第二步:使用形象资产ID或数智人项目ID创建视频流(使用形象资产ID可以指定音色相关参数,对话场景这些也没有明确;使用数智人项目ID里面包含了音色信息以及对话服务配置)
    //这里使用数智人项目ID创建视频流,以webrtc协议为例
    createSessionRsp, err := api.CreateSession(CreateSessionReq{
        UserId: "your user id",
        VirtualmanProjectId: "your virtualman project id",
        Protocol: "webrtc", //可选:webrtc、rtmp、trtc
        DriverType: 1, //drivetype=1代表支持文本驱动指令;drivetype=3代表支持音频和文本驱动指令
    })
    if err != nil {
        log.Fatalf("err:%v", err)
    }
    fmt.Println("获取流播放地址", createSessionRsp.PlayStreamAddr)
    for {
        var userInput string
```

```
    fmt.Println("是否成功复制地址 (输入'y'确定): ")
    fmt.Scanln(&userInput)
    if strings.ToLower(userInput) == "y" {
        break
    }
}

//第三步:循环检测会话状态直到SessionStatus=1 (已就绪)
sessionStatus := createSessionRsp.SessionStatus
for sessionStatus != 1 {
    time.Sleep(time.Second * 1)
    statSessionRsp, err := api.StatSession()
    if err != nil {
        log.Fatalf("err:%v", err)
    }
    sessionStatus = statSessionRsp.SessionStatus
}

//第四步:开启会话
_, err = api.StartSession()
if err != nil {
    log.Fatalf("err:%v", err)
}

//第五步:创建长链接通道
err = api.CommandChannel()
if err != nil {
    log.Fatalf("err:%v", err)
}

// 开启协程每过3s获取下行信息
go func() {
    for {
        time.Sleep(time.Second * 3)
        fmt.Println("\n-----开始调用长连接下行信息-----")
        api.ReadWssMessage()
    }
}()

//开启协程每过1分钟发送心跳指令
go func() {
    for {
        time.Sleep(time.Minute * 1)
        fmt.Println("\n-----开始调用长连接发送心跳指令-----")
        api.SendHeartBeatCommand()
    }
}()

//第六步:发送驱动指令
//根据不同的场景选择不同的函数发送驱动指令, 如大模型对话, 调用以下函数 (更多场景请见下文)
largeModelDialogue()

//第七步:关闭长连接通道和会话
//90s后关闭长连接通道 (根据自己的场景来选择断连时间)
time.Sleep(time.Second * 90)
api.CloseSession()
}
```

数智人绑定对话服务实践教程

数智人绑定的对话服务目前主要有如下两种配置：

1. 云小微对话（客服机器人对话）
2. 腾讯行业大模型对话

❗ 说明：

请到[数智人的项目创建与管理](#)绑定数智人对话服务，只能使用[数智人平台项目创建会话](#)方式。

对话服务推荐采用[文本驱动指令（WebSocket 长连接）](#)，以下是这两种配置的代码示例：

Go

```
// 客服机器人对话
func robotDialogue() {
    fmt.Println("\n-----云小微对话（客服机器人对话）-----")
    err := api.SendTextDriveWssCommand(TextDriveWssCommand{
        Data: &TextWssData{
            Text:      "微信简介", //根据客服机器人问答知识库回答
            ChatCommand: "",      //使用客服对话or大模型对话
        },
    })
    if err != nil {
        log.Fatalf("err:%v", err)
    }
}
```

Go

```
// 腾讯行业大模型对话
func largeModelDialogue() {
    fmt.Println("\n-----腾讯行业大模型对话-----")
    err := api.SendTextDriveWssCommand(TextDriveWssCommand{
        Data: &TextWssData{
            Text:      "你是谁", //可以在模型管理里面自定义配置问答
            ChatCommand: "",    //使用客服对话or大模型对话
            ChatRoundId: strings.Replace(uuid.NewV4().String(), "-", "", -1)[:32], //使用腾讯行业大模型对话服务的数智人项目需要传入此字段，用于区分多轮对话
        },
    })
    if err != nil {
        log.Fatalf("err:%v", err)
    }
}
```

直播场景实践教程

直播场景推荐使用直播数智人服务，实践教程可直接参考[直播数智人服务 API 文档](#)。

用户自行接入大模型实践教程

用户自行接入大模型：指用户自行接入了 ChatGPT、腾讯混元等大模型，输出通常为流式文本。

该场景推荐使用[流式文本驱动指令](#)中的非子句模式(isSentence=false)，非子句模式下服务端会对客户端发送的文本重新组句，能很好的配合大模型的流式输出，用户接入大模型的场景完整 Demo 参考[ASR/ChatGPT/数智人 整合 demo](#)。

以下是该场景对应的代码示例：

Go

```
// 用户自行接入大模型来驱动数智人的场景
func SendNotSentenceStreamText() {
    //模拟大模型的返回结果
    rspText := "我是一个数智人，我在测试发送流式文本非子句模式。"
    //模拟大模型生成流式文本
    streamChan := streamResponse(rspText, time.Millisecond*500)
    reqId := strings.Replace(uuid.NewV4().String(), "-", "", -1)[:32]
    seq := 1
    for char := range streamChan {
        api.SendStreamTextDriveWssCommand(StreamTextDriveWssCommand{
            ReqId: reqId,
            Data: &StreamTextWssData{
                Text: char,
                Seq: seq,
            },
        })
        seq++
    }
    // 发送final数据包
    api.SendStreamTextDriveWssCommand(StreamTextDriveWssCommand{
        ReqId: reqId,
        Data: &StreamTextWssData{
            Seq: 1,
            IsFinal: true,
        },
    })
}

func streamResponse(text string, delay time.Duration) <-chan string {
    ch := make(chan string)
    go func() {
        for _, char := range text {
            //模拟大模型流式返回延时
            time.Sleep(delay)
            ch <- string(char)
        }
        close(ch)
    }()
    return ch
}
```

音频驱动场景实践教程

用户可以使用音频文件来驱动数智人，以下是该场景对应的代码示例：

Go

```
func SendAudio() {
    /*
        音频格式：格式-PCM，采样率-16kHz，采样位深-16bits，声道-单声道。
        1.如果是麦克风实时收音发送，每录制160ms（5120B）的数据即可发出，中间不需要间隔等待；如果是读取离线音频文件
        发送，片包大小为160ms（5120B），片包发送间隔为120ms。
        2.最后一个片包大小按实际发送（需小于160ms）。
    */
}
```


3. 当数据包发送完毕后，必须再发送一个IsFinal=true的空数据包（Audio字段填空串）结束当次音频驱动使数字人回到静默状态。

4. 发送音频实时率要介于[0.75,1]，小于0.75会触发限速，大于1会导致画面卡顿。比如160ms音频片包大小，发送间隔不能低于120ms，不能高于160ms。

```
*/
chunkSize := 5120 //块大小-5120B
sleepBase := 120 //间隔ms
seq := 1
reqId := strings.Replace(uuid.NewV4().String(), "-", "", -1)[:32]
//读取音频二进制文件
voiceFile, err := os.Open("your_audio_file.pcm")
if err != nil {
    log.Fatalf("err:%v", err)
}
defer voiceFile.Close()
buffer := make([]byte, chunkSize)
for {
    n, err := voiceFile.Read(buffer)
    if err != nil && err != io.EOF {
        fmt.Println("Error reading file:", err)
        break
    }
    if n == 0 { // 文件读取完毕
        break
    }
    err = api.SendVoiceDriveCommand(VoiceDriveCommand{
        ReqId: reqId,
        Data: &VoiceData{
            Audio:  base64.StdEncoding.EncodeToString(buffer), //声音二进制数据编码Base64后的值。
            Seq:    seq,
            IsFinal: false,
        },
    })
    if err != nil {
        log.Fatalf("err:%v", err)
    }
    seq++ // 递增序列号
    // 等待120ms到160ms的随机时间
    sleepDuration := time.Duration(sleepBase+rand.Intn(41)) * time.Millisecond
    time.Sleep(sleepDuration)
}
// 发送final数据包
api.SendVoiceDriveCommand(VoiceDriveCommand{
    ReqId: reqId,
    Data: &VoiceData{
        Seq:    seq,
        IsFinal: true,
    },
})
}
```

完整 Demo 列表

Demo 内容	Demo 概述	语言	最新版本文件名	Demo 下载链接
ASR/ChatGPT/数智人整合 demo	本 Demo 展示了如何通过 ASR/ChatGPT，与数智人进行交互。	python	asr_chatgpt_vh_demo_20241212.zip	点击下载

附录

交互数智人 API 接入常见问题

最近更新时间：2024-12-30 11:31:42

三种视频建流协议有什么区别？

视频建流 Protocol 支持参数：rtmp、trtc、webrtc，这三者有什么区别？

1. RTMP：

- RTMP 是 Adobe 公司开发的一种基于 TCP 的流媒体传输协议，主要用于实时音视频流的发布和播放。
- 它是一种成熟的、广泛使用的直播协议，特别是在 Flash 播放器中。
- RTMP 协议在延迟和稳定性方面表现良好，但需要专用服务器支持，且基于 TCP，可能不适用于所有网络环境。
- 推荐端上使用 [播放器 SDK](#) 播放器播放视频流。

2. TRTC：

- TRTC 是腾讯云提供的实时通信服务，它支持多种传输协议（包括 UDP），以适应不同的网络环境和性能要求。
- TRTC 提供了丰富的 SDK 和服务端支持，使得开发者可以更容易地集成实时音视频通信功能。
- 作为一种云服务，TRTC 还提供了诸如房间管理、用户管理等附加功能，产品文档参考 [实时音视频](#)。

3. WebRTC：

- WebRTC 是一种支持实时音视频通信的开放标准，由 W3C 和 IETF 共同制定。
- WebRTC 基于 P2P（点对点）技术，可以在不依赖中心服务器的情况下建立直接连接，从而降低延迟并提高通信效率。
- WebRTC 支持多种传输协议（如 SCTP、DTLS 等），并且可以在不同的网络环境下工作。
- WebRTC 的主要优势在于其开放性和兼容性，可以在现代浏览器中无需插件即可运行。
- 推荐端上使用 [播放器 SDK](#) 播放器播放视频流。

总结：

RTMP 是一种成熟的直播协议，适用于需要稳定性和低延迟的传统直播场景。

TRTC 是腾讯云提供的实时通信服务，提供了丰富的 SDK 和服务端支持，适用于需要快速集成和额外功能的场景。

WebRTC 是一种开放标准的实时通信技术，适用于需要低延迟、高效率和开放性的现代 Web 应用。

SEI 是什么？

SEI（Supplemental Enhancement Information）信息是视频编码中的一种辅助增强信息，用于提供一些额外的、非必要的信息，以帮助解码器更好地处理视频流或提供一些附加信息给用户。SEI 信息可以包含各种类型的数据，如图像处理指示、版权信息、定时信息等。

- 在采用 H.264 协议的 RTMP 或 WebRTC 建流中增加 SEI 信息，可以使得视频流携带更多的自定义数据或特定于应用的信息。这些信息对于标准的视频播放可能不是必需的，但对于特定的应用场景（如互动直播、广告插入等）来说可能非常有用。
- 举个例子，假设在一个直播应用中，主播希望在直播过程中向观众发送一些实时信息，例如当前的时间、直播间的互动提示或者广告信息。通过将 SEI 信息嵌入到 H.264 编码的视频流中，主播可以方便地将这些信息与视频内容一起传输给观众。观众的设备在解码视频时能够解析出这些 SEI 信息，并根据需要进行处理或展示给用户。这样，直播应用就能够在不影响视频流畅度的情况下，向用户提供更多的实时信息和互动功能。

如何使用 ffmpeg 转换合格的 pcm 音频格式？

FFmpeg 是一个非常强大的开源多媒体处理工具，它可以处理视频、音频和其他多媒体文件。

可以从官网进行下载：<https://ffmpeg.org/>

假设您有一个音频文件 input.wav，您想将其转换为 PCM 格式的文件 output.pcm。可以采用如下命令：

```
ffmpeg -i input.wav -f s16le -ar 16000 -ac 1 -c:a pcm_s16le output.pcm
-ar 16000：设置采样率为16kHz
-f s16le：设置采样位深为16bits
-ac 1：设置为单声道
-c:a pcm_s16le：指定音频编解码器为16位PCM
```

为了确保格式转化成功，可以使用 ffprobe 查看音频信息，具体命令如下：

```
ffprobe -i output.pcm
```

```
Input #0, pcm_s16le, from 'output.pcm':
```

```
Duration: 00:00:30.00, bitrate: 1411 kb/s
  Stream #0:0: Audio: pcm_s16le ([1][0][0][0] / 0x0001), 16000 Hz, mono, s16, 1411 kb/s
```

关键信息如下：

16000 Hz, mono, s16

16000 Hz 代表了采样率为16kHz

mono 代表为单声道

s16 代表采样位深为16bits

视频流嵌入信息说明

最近更新时间：2025-04-16 12:06:53

SEI (Supplemental Enhancement Information) 是视频编码中的一种辅助增强信息，用于提供一些额外的、非必要的的数据，以帮助解码器更好地处理视频流或提供一些附加信息给用户。SEI信息可以包含各种类型的数据，如图像处理指示、版权信息、定时信息等。

1. 数智人交互 aPaas 中采用 H264 协议的 RTMP、WebRTC 和 TRTC 建流可以增加 SEI 信息。
2. 嵌入流的 SEI 信息包含用户请求携带的 SEI 数据信息和驱动的状态信息，采用 JSON 格式封装。
3. SEI 数据包格式如下：uuid+data，前16个字节为 uuid，第17个字节开始为 JSON 格式的包体数据。

```
{
  "sei_data": {

  },
  "status_data": {
    "STATUS": "AudioStart"
  }
}
```

❗ 说明：

- sei_data 值为用户上传的 VideoSeiInfo 原样返回。
- status_data 值为服务端返回信息的扩展字段，前端支持 STATUS 为当前驱动的状态，支持 AudioStart、AudioOver、SentenceStart、SentenceOver、TextStart、TextOver（第一个文本的 TextStart 会覆盖 SentenceStart）。
- sei_data 和 status_data 的key 都保持有，如果没有对应值，则对应 value 为空字符串，也就是说即便 sei_data 和 status_data 对应的数据为空，也需要保证存在这个字段。

错误码

最近更新时间：2024-08-22 11:13:31

错误码说明

如果返回结果的Header中 Code 值非0，则表示调用 API 接口失败。例如：

```
{
  "Header": {
    "Code": -100000,
    "Message": "系统内部错误",
    "RequestID": "gz1ab1a40c17222461134894879",
  }
}
```

错误码列表

代码	含义	措施
-101000	系统内部错误	重试或联系客服定位错误
-101001	数智人服务算力资源不足	重试或联系客服定位错误
-101002	下载图片失败	重试或联系客服定位错误
-111001	更新流参数失败	重试或联系客服定位错误
-111002	连接异常断开	重试或联系客服定位错误
-111003	数智人模型加载过程出现错误	重试或联系客服定位错误
-111004	数智人服务内部建流超时	重试或联系客服定位错误
-111005	文本或音频驱动数智人时出现错误	请阅读异常断开返回描述，如果是系统错误类型，联系客服定位错误
100001	请求参数错误	字段格式、长度、类型、非空约束等错误，查看协议文档，对照字段传输要求
100005	未授权操作	请按授权内容请求相关接口或资源
100008	超过并发限制	请检查已用并发数和并发数上限
100012	请求过于频繁	请按协议要求频率发送
100014	并发配额不足	请检查并发配额是否充足或是否在有效期内
100015	数智人形象已过期	请检查有效期
110013	会话已关闭	请确认是否在驱动已关闭的流，如果异常关流，请使用 statsession 查看 code 和 message 定位异常关流原因
110014	会话未就绪	请确认流创建完成 status 为1且使用了 startsession 开启了会话。
110015	因发送驱动指令时机不对导致请求未被处理	请参考对应接口文档要求发送指令。
110016	对未开启的session发起驱动指令	请确认流创建完成 status 为1且使用了 startsession 开启了会话。
110018	会话不存在	请检查会话状态是否已关闭或者会话未创建。
110021	trtc 服务欠费或可用时长不足	请检查 trtc 服务是否欠费。
110022	同 userid 建流会话被关闭	请检查建流时传输的 userid。

端渲染会话交互接入 API 文档

端渲染会话交互接入概述

最近更新时间：2025-04-16 12:06:53

本文档主要描述 腾讯云智能数智人 端渲染驱动 aPaas 的开放接口协议，适用于3D和2D端渲染交互场景。

接口调用环境

正式环境：`https://gw.tvs.qq.com`

注意事项

调用该接口前请确保下单购买了下图红框中的产品：



接口调用方式

调用每个接口时需要在接口的 query 中传入公共参数及签名。参数及签名方式说明参考以下链接：

[数智人 aPaas 接口调用方式](#)

接口接入流程

接入指引：

第1步，开通数智人平台账号，开通方式参考 [购买指南](#)；

第2步，登录平台 [数智人交互平台](#)。创建数智人项目，得到项目的 appkey、accesstoken、virtualmankey，通过 [创建长连接通道](#) 接口建立 websocket 连接；

第3步，通过长连接通道发送驱动请求，并获取数智人驱动下行消息。请求、返回内容见 [端渲染驱动接口](#)；

第4步，如果驱动结束，根据DriverType驱动类型，发送结束标记或断开连接。

创建长连接通道

最近更新时间：2024-08-09 19:25:01

建立 websocket 长连接通道，上行发送驱动指令、接收下行消息。

调用协议

WebSocket + JSON

WSS /v2/ws/ivh/interactdriver/interactdriverservice/driverengine

Header Content-Type: application/json;charset=utf-8

建连方式

websocket url: 参考 [数智人 aPaas 接口调用方式](#) 接口调用方式组织 url。

端渲染驱动接口

最近更新时间：2025-04-18 10:44:02

创建[长连接通道](#)后，可以通过 websocket 长连接发送文本获取驱动数据。

请求参数

参数名称	类型	必选	描述
ReqId	string	是	单次请求唯一标识，长度为32的 uuid
StreamId	string	是	会话 ID，用于区分多轮会话，长度为32的 uuid(不包含 ‘-’)
VirtualmanProjectId	string	是	数智人项目ID，可在数智人项目中获取
InputText	string	否	请求文本内容。当 DriverType 为 TEXT 时，不可为空
SpeechParam	SpeechParam	否	定义输出音频的详细参数
DriverType	string	是	驱动类型 1. TEXT: 文本驱动； 2. CHAT: 文本对话驱动 3. STREAM_TEXT: 流式文本驱动；
ChatCommand	string	否	对话指令，默认值 CHATTING 1. CHATTING: 对话 2. START_CHAT: 开启对话 3. STOP_CHAT: 中断对话
InputTextType	string	否	InputText的类型，默认为MARKDOWN 1. MARKDOWN: markdown 格式，包含纯文本，可支持流式 2. SSML: ssml 标准格式，不支持流式
Seq	int	否	流式文本分片的ID
IsFinal	bool	否	流式文本分片的结束标记（每一段流式文本结束必须传入结束标记）

SpeechParam

参数名称	类型	必选	描述
TimbreKey	string	否	音色值，默认使用数智人项目配置音色。可使用的音色列表，可通过 分页查询音色列表 获取音色取值
Speed	float	否	语速: 1.0为正常语速，范围[0.5–2.0]，值为0.5时播报语速最慢，值为2.0时播报语速最快。未指定时默认取数智人项目配置的语速值
Volume	int	否	音量大小，范围[-10, 10]，对应音量大小。默认为0，代表正常音量，值越大音量越高
EmotionCategory	string	否	控制合成音频的情感，仅支持多情感音色使用，可选值参考 分页查询音色列表
EmotionIntensity	int	否	控制合成音频情感程度，取值范围为 [50,200]，只有 EmotionCategory 不为空时生效
SmartActionEnabled	bool	否	是否开启智能动作，默认不开启
SubtitleType	int	否	字幕返回的模式，按字级别还是词级别。默认按字 0: 按字； 1: 按词；

TimbreLanguage	string	否	音色语种，可选语种参考个人资产管理 API 分页查询音色列表 接口，多语种音色在合成时必须选择对应语种。
----------------	--------	---	--

长连接下行消息

参数名称	类型	必选	描述
ReqId	string	是	单次请求 ID，和入参一致
StreamId	string	是	会话 ID，用于区分多轮会话，和入参一致
DriverRspType	string	是	响应类型 1. REPLY: 返回 ReplyRsp，对应会话信息 2. SPEECH: 返回 SpeechRsp，对应音频信息
ReplyRsp	ReplyRsp	否	会话响应，当 DriverRsp 为REPLY 时返回
SpeechRsp	SpeechRsp	否	音频信息相应，当 DriverRsp为SPEECH 时返回
ErrorCode	int	是	错误码
ErrorMessage	string	是	错误信息

ReplyRsp

参数名称	类型	必选	描述
ReplyType	string	是	回复语类型。 1. cloudAiGpt: 腾讯云大模型对话 2. yunxiaowei: 云小微客服对话 3. cloudAiWaiting: 首包超时等待话术 4. cloudAiTimeOut: 超时未响应话术，会话结束 5. sensitive: 输入文本或回复语中包含敏感内容时，返回的固定话术 6. input: InputText为纯文本或流式文本时输入的内容 7. enhanceText: 未配置对话服务时，匹配到话术管理的内容 8. cloudAiThought: 腾讯云大模型思考过程内容
ReplyPro	string	否	播报内容，包含ssml标签
ReplyDisplay	string	是	展示内容，包含富文本标签
InteractionType	string	否	特殊消息类型
InteractionContent	string	否	特殊消息内容，用于下发弹窗、图片等非文本类的特殊消息
Uninterrupt	bool	是	当前播报内容是否可打断
Muted	bool	是	当前播报内容，是否关闭收音
SeqNo	int	是	子句序号，当 ReplyType为cloudAiGp t时，正常回复语序号从1开始，其余固定话术从0开始
ContentType	int	是	回复语内容类型 0: 未知 1: 普通字符串 2: 有序列表 3: 无序列表 4: 图片链接 5: http链接 6: 表格 8: 标题

			9: SSML
TtsSupport	bool	是	当前子句是否播报
IsFinal	bool	是	是否为最后一句
IsHighLight	bool	是	是否需要高亮展示

SpeechRsp

参数名称	类型	必选	描述
Audio	string	是	base64 编码的 pcm 音频数据
ThDim	int	是	口型维度
ThFeat	Array of float	是	口型数据
Phn	Array of [PhnInfo]	是	音素信息
Word	Array of [WordInfo]	是	分词信息
Final	bool	是	整句结束标识
SentenceFinal	bool	是	流式子句结束标识
Sampling	int	是	采样率
Action	Array of [Action]	是	动作信息
Subtitle	Array of [SubtitleInfo]	是	字幕信息
RealThType	string	是	口型参数
Expression	Array of [Expression]	是	表情信息
SeqNo	int	是	子句序号
SentenceStart	bool	是	子句开始
ThFeatFinal	bool	是	口型结束标识

PhnInfo

参数名称	类型	必选	描述
Phn	string	是	音素
Start	string	是	起始时间，单位为0.1us，该数值/10000为ms
End	string	是	结束时间，单位为0.1us，该数值/10000为ms

WordInfo

参数名称	类型	必选	描述
Phn	string	是	音素
Word	string	是	对应单词

Action

参数名称	类型	必选	描述
Pos	string	是	动作名称
Start	string	是	起始时间点，单位为0.1us，该数值/10000为ms

SubtitleInfo

参数名称	类型	必选	描述
Word	string	是	对应单词
Start	string	是	起始时间点，单位为0.1us，该数值/10000为ms
End	string	是	结束时间点，单位为0.1us，该数值/10000为ms
PosStart	string	是	文本中的起始 unicode 位置，注意为左闭右开形式[PosStart,PosEnd)
PosEnd	string	是	文本中的结束 unicode 位置，注意为左闭右开形式[PosStart,PosEnd)

Expression

参数名称	类型	必选	描述
Name	string	是	表情名称
Start	string	是	起始时间点，单位为0.1us，该数值/10000为ms
End	string	是	结束时间点，单位为0.1us，该数值/10000为ms
Loc	string	是	表情在文本中的 unicode 位置
Flag	string	是	B: 该段文本包含表情起始; I: 该段文本属于表情中间的一部分 E: 该段文本表情结束 S: 该段文本包含表情的起始、结束

请求示例

```
{
  "Header": {},
  "Payload": {
    "VirtualmanProjectId": "253b2a182d694a60bed82635b18025a2",
    "InputText": "在人工智能产业中，哪些领域的AI发展基础条件表现较优？",
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "StreamId": "92597c35-3a99-415e-9bae-3124771b7749",
    "DriverType": "TEXT",
    "SpeechParam": {
      "TimbreKey": ""
    }
  }
}
```

返回示例

```
//DriverRspType为REPLY
{
  "Header": {
    "RequestID": "fe0e4c13f2a34cb69b2475d8483f28de",
  }
}
```

```

    "SessionID": "gza802cc9317231084402578413",
    "DialogID": "",
    "Code": 0,
    "Message": ""
  },
  "Payload": {
    "DriverRspType": "REPLY",
    "ErrorCode": 0,
    "ErrorMessage": "",
    "ReplyRsp": {
      "ContentType": 1,
      "InteractionContent": "",
      "InteractionType": "",
      "IsFinal": true,
      "IsHighLight": true,
      "Muted": false,
      "ReplyDisplay": "哪些领域的AI发展基",
      "ReplyPro": "\u003cspeak\u003e哪些领域的AI发展基础条件表现较优? \u003c/speak\u003e",
      "ReplyType": "input",
      "SeqNo": 2,
      "TtsSupport": true,
      "UninterrId": "fe0e4c13f2a34cb69b2475d8483f28de",
      "SpeechRsp": {
        "Action": [],
        "Audio": "",
        "Expression": [],
        "Final": false,
        "Phn": [],
        "RealThType": "",
        "Sampling": 0,
        "SentenceFinal": false,
        "SentenceStart": false,
        "SeqNo": 0,
        "Subtitle": [],
        "ThDim": 0,
        "ThFeat": [],
        "ThFeatFinal": false,
        "Word": []
      },
    },
    "StreamId": "92597c35-3a99-415e-9bae-3124771b7749"
  }
}

//DriverRspType为SPEECH
{
  "Header": {
    "RequestID": "fe0e4c13f2a34cb69b2475d8483f28de",
    "SessionID": "gza802cc9317231084402578413",
    "DialogID": "",
    "Code": 0,
    "Message": ""
  },
  "Payload": {
    "DriverRspType": "SPEECH",
    "ErrorCode": 0,
    "ErrorMessage": "",
    "ReplyRsp": {
      "ContentType": 0,

```

```
"InteractionContent": "",
"InteractionType": "",
"IsFinal": false,
"IsHighLight": false,
"Muted": false,
"ReplyDisplay": "",
"ReplyPro": "",
"ReplyType": "",
"SeqNo": 0,
"TtsSupport": false,
"Uninterrupt": false
},
"ReqId": "fe0e4c13f2a34cb69b2475d8483f28de",
"SpeechRsp": {
  "Action": [],
  "Audio": "", //内容过长，不展示
  "Expression": [],
  "Final": false,
  "Phn": [
    {
      "End": "200000",
      "Phn": "sil0",
      "Start": "0"
    },
    {
      "End": "1100000",
      "Phn": "z4",
      "Start": "200000"
    },
    {
      "End": "2700000",
      "Phn": "ai4",
      "Start": "1100000"
    },
    {
      "End": "3800000",
      "Phn": "r2",
      "Start": "2700000"
    },
    {
      "End": "5100000",
      "Phn": "en2",
      "Start": "3800000"
    },
    {
      "End": "5800000",
      "Phn": "g1",
      "Start": "5100000"
    },
    {
      "End": "7300000",
      "Phn": "ong1",
      "Start": "5800000"
    },
    {
      "End": "8100000",
      "Phn": "zh4",
      "Start": "7300000"
    }
  ],
}
```

```

    {
      "End": "9000000",
      "Phn": "iii4",
      "Start": "8100000"
    },
    {
      "End": "9800000",
      "Phn": "n2",
      "Start": "9000000"
    },
    {
      "End": "11300000",
      "Phn": "eng2",
      "Start": "9800000"
    },
    {
      "End": "12800000",
      "Phn": "ch3",
      "Start": "11300000"
    },
    {
      "End": "14000000",
      "Phn": "an3",
      "Start": "12800000"
    },
    {
      "End": "16000000",
      "Phn": "ie4",
      "Start": "14000000"
    },
    {
      "End": "17100000",
      "Phn": "zh1",
      "Start": "16000000"
    },
    {
      "End": "19200000",
      "Phn": "ong1",
      "Start": "17100000"
    },
    {
      "End": "24200000",
      "Phn": "sil0",
      "Start": "19200000"
    }
  ],
  "RealThType": "3D_standard",
  "Sampling": 24000,
  "SentenceFinal": false,
  "SentenceStart": true,
  "SeqNo": 1,
  "Subtitle": [
    {
      "End": "2700000",
      "PosEnd": "1",
      "PosStart": "0",
      "Start": "200000",
      "Word": "在"
    }
  ],

```

```

    {
      "End": "5100000",
      "PosEnd": "2",
      "PosStart": "1",
      "Start": "2700000",
      "Word": "人"
    },
    {
      "End": "7300000",
      "PosEnd": "3",
      "PosStart": "2",
      "Start": "5100000",
      "Word": "工"
    },
    {
      "End": "9000000",
      "PosEnd": "4",
      "PosStart": "3",
      "Start": "7300000",
      "Word": "智"
    },
    {
      "End": "11300000",
      "PosEnd": "5",
      "PosStart": "4",
      "Start": "9000000",
      "Word": "能"
    },
    {
      "End": "14000000",
      "PosEnd": "6",
      "PosStart": "5",
      "Start": "11300000",
      "Word": "产"
    },
    {
      "End": "16000000",
      "PosEnd": "7",
      "PosStart": "6",
      "Start": "14000000",
      "Word": "业"
    },
    {
      "End": "19200000",
      "PosEnd": "9",
      "PosStart": "7",
      "Sta": "16000000",
      "Word": "中, "
    }
  ],
  "ThDim": 52,
  "ThFeat": [], //内容过长, 不展示
  "ThFeatFinal": false,
  "Word": [
    {
      "Phn": "z-ai4",
      "Word": "在"
    },
    {

```

```
    "Phn": "r-en2|g-ng1",
    "Word": "人工"
  },
  {
    "Phn": "zh-iii4|n-eng2",
    "Word": "智能"
  },
  {
    "Phn": "ch-an3|ie4",
    "Word": "产业"
  },
  {
    "Phn": "zh-ong1",
    "Word": "中"
  }
]
},
"StreamId": "925999-415e-9bae-3124771b7749"
}
```


直播数智人服务 API 文档

直播数智人服务 API 概述

最近更新时间：2024-11-13 14:51:52

本文档主要描述 腾讯云智能数智人 直播aPaas平台的开放接口协议，以及阐述接口的调用流程。

调用前须知

直播数智人服务是基于交互数智人服务来实现的，所需购买产品同交互数智人服务。
调用直播数智人服务接口前请确保您进行了形象采购以及购买了交互服务，详情请查看购买指南中的 [购买指引](#)。

接口调用环境

正式环境：`https://gw.tv.s.qq.com`

接口调用方式

调用每个接口时需要在接口的 query 中传入公共参数及签名。参数及签名方式说明参考以下链接：
[数智人 aPaas 接口调用方式](#)

API 概览

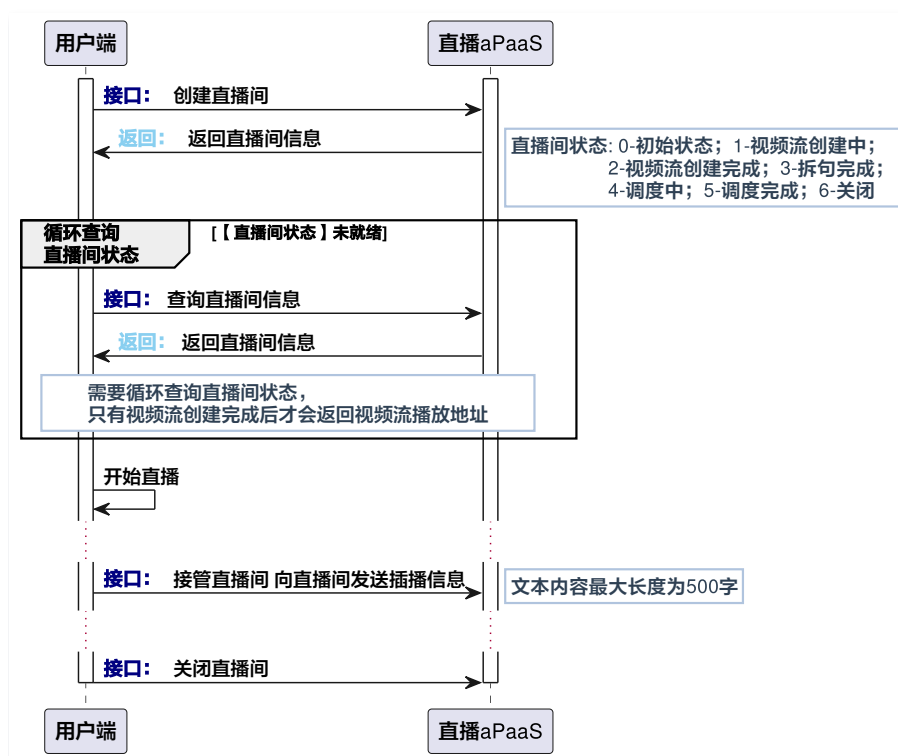
API 标题	接口	说明
创建直播间	POST /v2/ivh/liveroom/liveroomservice/openliveroom	创建直播间
查询直播间信息	POST /v2/ivh/liveroom/liveroomservice/getliveroom	查询直播间信息
查询直播间列表	POST /v2/ivh/liveroom/liveroomservice/listliveroomofappkey	查询直播间列表，返回使用某 appkey 创建的未关闭的直播间
直播接管	POST /v2/ivh/liveroom/liveroomservice/takeover	向直播间发送插播信息
关闭直播间	POST /v2/ivh/liveroom/liveroomservice/closetliveroom	关闭直播间

接口接入流程

接入指引

- 通过下列步骤使用直播数智人服务：
1. 购买相关产品及服务：[购买链接](#)；
 2. 创建直播间：调用 [创建直播间](#) 接口即可创建；
 3. 等待直播间就绪：通过调用 [查询直播间信息](#) 接口循环查询，直到直播间状态为视频流创建完成（Status=2）；
 4. 开始直播：直播过程中可通过调用 [直播接管](#) 接口来进行插播；
 5. 关闭直播：直播完成请关闭 [关闭直播间](#)。

接入流程



代码示例

```
Go

package main

import (
    "fmt"
    "log"
    "time"
)

func main() {
    // 第一步: 初始化api调用对象
    var api = NewVHApi(
        "your_appkey",           //appkey
        "your_accesstoken",      //accesstoken
        "https://gw.tvs.qq.com") //httpOrigin

    // 第二步: 开启直播间
    openLiveroomRsp, err := api.OpenLiveroom(OpenLiveroomReq{
        CycleTimes: 5, //循环次数
        //CallbackUrl: "https://callback.example.com/callback_example_url",
        VideoStreamReq: &VideoStreamReq{
            VirtualmanProjectId: "your_virtualman_project_id", //数智人项目ID
            Protocol:          "webrtc",                     //视频流协议
        },
        Scripts: []*ScriptReq{
            {
                Content: "直播文本内容:腾讯云智能数智人(Tencent Cloud AI Digital Human,TCADH)采用语音交互、虚拟形象模型生成等多项 AI 技术,实现唇形语音同步和表情动作拟人等效果,广泛应用于虚拟形象播报和实时语音交互两大场景。虚拟形象播报根据文本内容快速合成音视频文件,落地于媒体、教育、会展服务等场景,实时语音交互支持即时在线对话,可赋能智能客服、语音助理等场景。",
                Backgrounds: []*VideoLayer{
```

```
        {Url: "your_background_image_url"},
    },
},
},
}))
if err != nil {
    log.Fatalf("err:%v", err)
}
liveRoomId := openLiveroomRsp.LiveRoomId

fmt.Println("获取开启直播间接口中返回的VideoStreamPlayUrl:", openLiveroomRsp.VideoStreamPlayUrl)

// 查询直播间
for openLiveroomRsp.Status != 2 {
    openLiveroomRsp, _ = api.GetLiveroom(liveRoomId)
}

// 接管直播演示, 插播并可补充背景图片
time.Sleep(30 * time.Second)
api.TakeoverLiveroom(TakeoverReq{
    LiveRoomId: liveRoomId,
    Content:    "插播测试: 文本内容, 最大长度500字",
})

time.Sleep(120 * time.Second)
// 第三步: 关闭直播间
api.CloseLiveroom(liveRoomId)
}
```

创建直播间

最近更新时间：2025-05-22 17:37:52

接口说明

创建直播间。

调用协议

HTTPS + JSON
POST /v2/ivh/liveroom/liveroomservice/openliveroom
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	string	是	请求 Id，长度为32的 uuid
VideoStreamReq	VideoStreamReq	是	创建会话的参数
CycleTimes	int	是	循环次数（不支持无限循环），取值范围：(0,500)
Scripts	Arrays of [ScriptReq]	是	直播间脚本的参数，当前数组只支持一条数据
CallbackUrl	string	否	回调通知的 url。回调协议参考 直播回调通知 ⚠ 注意： 回调签名使用的 signKey 通过线下方式提供。

VideoStreamReq参数

参数名称	类型	必选	描述
VirtualmanProjectId	string	否	数智人平台项目 ID，与 AssetVirtualmanKey 二选一。获取方式如下 
AssetVirtualmanKey	string	否	资产形象 ID，与 VirtualmanProjectId 二选一。获取方式如下 
Protocol	string	是	视频流协议，当前支持参数：rtmp、trtc、webrtc
SpeechParam	SpeechParam	否	音色相关参数，使用 AssetVirtualmanKey 建流时，可以选择指定这个参数

AnchorParam	AnchorParam	否	数智人相关参数，使用 AssetVirtualmanKey 建流时，可以选择指定这个参数来调整位置、大小
ExtraInfo	ExtraInfo	否	创建会话扩展参数

SpeechParam参数

参数名称	类型	必选	描述
Speed	float	否	语速（1.0为正常语速，范围[0.5–1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快）
TimbreKey	string	否	音色key，可通过 查询声音资产信息 接口获得
Volume	int	否	音量大小，范围[0, 10]，对应音量大小。默认为0，代表正常音量，值越大音量越高

AnchorParam参数

参数名称	类型	必选	描述
HorizontalPosition	float	否	定义主播横向位置（0为中间位置，推荐范围为 [-0.5, 0.5]，-0.5为最左，0.5为最右）
VerticalPosition	float	否	定义主播纵向位置（0为中间位置，推荐范围 [-0.5, 0.5]，-0.5为向上，0.5为向下）
Scale	float	否	缩放比例，（1是默认大小，推荐范围(0,10]）

ExtraInfo参数

名称	类型	必选	描述
AlphaChannelEnable	bool	否	是否开启原始视频拼接 alpha 输出拼接视频流。 说明： 假设原始输出宽度为iW，则实际视频流宽度为：iW + pW + iW pW 为中间间隔(pad)宽度：(16 - iW%16) % 16 alpha 视频宽度起始点：iW + pW alpha 视频宽度：iW

ScriptReq参数

参数名称	类型	必选	描述
Content	string	是	文本内容，上限40万个字符（字符数按 unicode 字符数计算）
Backgrounds	Arrays of VideoLayer	否	背景，背景和前景总数不能超过5 说明： 1. 对于未抠除拍摄背景的主播，此参数无法生效。 2. 对于抠除拍摄背景的主播，背景生效优先顺序为： 2.1 此参数设置的背景； 2.2 数智人平台项目配置的背景（使用数智人平台项目场景）； 2.3 绿幕。
Foregrounds	Arrays of VideoLayer	否	前景，背景和前景总数不能超过5

VideoLayer参数

参数名称	类型	必选	描述
Url	string	是	图层资源地址 图片格式：支持jpg、jpeg、png、mjpeg、gif 图片大小：建议不超过2M
X	int	是	资源放置的左上x坐标
Y	int	是	资源放置的左上y坐标
Width	int	是	资源输出的宽度
Height	int	是	资源输出的高度

返回参数

参数名称	类型	必须	描述
ReqId	string	是	单次请求 Id，和入参 ReqId 一致
LiveRoomId	string	是	直播间 id
Status	int	是	直播间状态，0-初始状态；1-视频流创建中；2-视频流创建完成；3-拆句完成；4-调度中；5-调度完成；6-关闭
VideoStreamPlayUrl	string	否	视频流播放地址（视频流创建完成才会返回，即Status=2）

请求示例

```
{
  "Header": {},
  "Payload": {
    "ReqId": "3429db2406c04a5e94a3f63478e8778f",
    "VideoStreamReq": {
      "VirtualmanProjectId": "12b24hub97f447f28ec17c26cc17c985",
      "Protocol": "webrtc"
    },
    "CycleTimes": 10,
    "Scripts": [{
      "Content": "直播文本内容：Tencent Cloud AI Digital Human，TCADH）采用语音交互、虚拟形象模型生成等多项 AI 技术，实现唇形语音同步和表情动作拟人等效果，广泛应用于虚拟形象播报和实时语音交互两大场景。虚拟形象播报根据文本内容快速合成音视频文件，落地于媒体、教育、会展服务等场景，实时语音交互支持即时在线对话，可赋能智能客服、语音助理等场景。"
    }]
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "3c5a8d999d8e1fbc6c574595a1b31d5d",
    "SessionID": "gz92439a3317234609262925106",
    "DialogID": "",
    "Code": 0,
    "Message": ""
  },
  "Payload": {
    "LiveRoomId": "lm16406256654745613158",
  }
}
```

```
"Status": 1,  
"ReqId": "3429db2406c04a5e94a3f63478e8778f"  
}  
}
```

查询直播间信息

最近更新时间：2024-11-01 17:24:22

接口说明

创建直播间后，可查询直播间信息。

调用协议

HTTPS + JSON
POST /v2/ivh/liveroom/liveroomservice/getliveroom
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	string	是	请求Id，长度为32的uuid
LiveRoomId	string	是	直播间id

返回参数

参数名称	类型	必选	描述
ReqId	string	是	单次请求Id，和入参一致
LiveRoomId	string	是	直播间id
Status	int	是	直播间状态，0-初始状态；1-视频流创建中；2-视频流创建完成；3-拆句完成；4-调度中；5-调度完成；6-关闭
VideoStreamPlayUrl	string	否	视频流播放地址（视频流创建完成才会返回）

请求示例

```
{
  "Header": {},
  "Payload": {
    "LiveRoomId": "1m16441358952693772460",
    "ReqId": "89ba8c61e9854b62902db33b8daf48c1"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "aeeac7e4cf1b3db317c90db91acb6267",
    "SessionID": "gz92439a3317234630252833536",
    "DialogID": "",
    "Code": 0,
    "Message": ""
  },
  "Payload": {
    "LiveRoomId": "1m16441358952693772460",
    "VideoStreamPlayUrl": "webrtc://xxx.com/live/1m16441358952693772460?min_delay_ms=100",
    "Status": 2,
  }
}
```



```
"ReqId": "89ba8c61e9854b62902db33b8daf48c1"  
}  
}
```

直播接管

最近更新时间：2024-11-01 19:34:42

接口说明

向直播间发送插播信息。

调用协议

HTTPS + JSON
POST /v2/ivh/liveroom/liveroomservice/takeover
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	string	是	请求Id，长度为32的uuid
LiveRoomId	string	是	直播间id
Content	string	是	文本内容，最大长度500字。

返回参数

参数名称	类型	必选	描述
ReqId	string	是	单次请求Id，和入参一致

请求示例

```
{
  "Header": {},
  "Payload": {
    "ReqId": "369ded5b8aaa4e6ab942d3f58baf7523",
    "LiveRoomId": "1m16441358952693772460",
    "Content": "我在做插播测试：文本内容，最大长度500字"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "a58f83539f222119d247f2e0f61994f7",
    "SessionID": "gza802cc9317234630253261542",
    "DialogID": "",
    "Code": 0,
    "Message": ""
  },
  "Payload": {
    "ReqId": "369ded5b8aaa4e6ab942d3f58baf7523"
  }
}
```

查询直播间列表

最近更新时间：2024-11-01 17:21:42

接口说明

查询直播间列表，查询使用当前appkey创建的未关闭的直播间。

调用协议

HTTPS + JSON
POST /v2/ivh/liveroom/liveroomservice/listliveroomofappkey
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	string	是	请求Id，长度为32的uuid
PageIndex	int	是	页码，从1开始
PageSize	int	是	每页条数，大于0，小于等于1000

返回参数

参数名称	类型	必选	描述
ReqId	string	是	单次请求Id，和入参一致
LiveRooms	Arrays of [LiveRoom]	是	直播间数组，无数据时不返回此字段
Total	int	是	总条数

LiveRoom参数

参数名称	类型	必选	描述
LiveRoomId	string	是	直播间id
Status	int	是	直播间状态，0-初始状态；1-视频流创建中；2-视频流创建完成；3-拆句完成；4-调度中；5-调度完成；6-关闭
VideoStreamPlayUrl	string	否	视频流播放地址（视频流创建完成才会返回，即Status=2）

请求示例

```
{
  "Header": {},
  "Payload": {
    "ReqId": "49eb1ed343a248acaf14c63b15d93697",
    "PageIndex": 1,
    "PageSize": 10
  }
}
```

返回示例

```
{
```

```
"Header": {
  "RequestID": "21add85761b654bdaf714be092ed36e6",
  "SessionID": "gza802cc9317234630181331439",
  "DialogID": "",
  "Code": 0,
  "Message": ""
},
"Payload": {
  "ReqId": "49eb1ed343a248acaf14c63b15d93697",
  "LiveRooms": [{
    "LiveRoomId": "lm16430281628057611272",
    "VideoStreamPlayUrl": "webrtc://xxx.com/live/lm16430281628057611272?min_delay_ms=100",
    "Status": 4
  }],
  "Total": 1
}
```

关闭直播间

最近更新时间：2024-11-01 17:21:42

接口说明

关闭直播间。

调用协议

HTTPS + JSON
POST /v2/ivh/liveroom/liveroomservice/closetiveroom
Header Content-Type: application/json;charset=utf-8

请求参数

参数名称	类型	必选	描述
ReqId	string	是	请求Id，长度为32的uuid
LiveRoomId	string	是	直播间id

返回参数

参数名称	类型	必选	描述
ReqId	string	是	单次请求Id，和入参一致

请求示例

```
{
  "Header": {},
  "Payload": {
    "LiveRoomId": "lm16441358952693772460",
    "ReqId": "158eac2f0e2740819efc15c4ffbfab30"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "68bb87d6f1fc5c8cab6c66a001fb5571",
    "SessionID": "gza802cc9317234630255161543",
    "DialogID": "",
    "Code": 0,
    "Message": ""
  },
  "Payload": {
    "ReqId": "158eac2f0e2740819efc15c4ffbfab30"
  }
}
```

直播回调通知

最近更新时间：2024-11-01 17:21:42

接口说明

用于在创建直播间后回调直播信息给客户端。

调用协议

请求方式：HTTPS-POST
数据格式：JSON格式请求（"Content-Type":"application/json"）

请求构建

公共参数

参数名称	类型	必选	描述
reqid	string	是	请求Id，长度为32的uuid
timestamp	string	是	请求时间戳，单位：秒。
signature	string	是	请求签名（参考签名方法）。

请求构建步骤

1. signature 签名规则如下（类似于 [数智人 aPaas 接口调用方式](#)，只是公共参数有变化），假设signKey和callbackUrl如下：

```
signKey = example_signkey
callbackUrl = https://callback.example.com/callback_example_url
```

 **注意：**
签名使用的 signKey 通过线下方式提供。

2. 根据请求参数，对除 signature 之外的所有需要的参数按字典序进行排序，作为签名原文，用于生成签名。使用 reqid 和 timestamp 两个参数，排序拼接后的字符串示例为：

```
reqid=f6d83764e3854dc88bed0f87067e5b7f&timestamp=1719316714
```

3. 对签名原文使用 signKey 进行 HmacSha256 加密，之后再进行 base64 编码：

```
hashBytes = HmacSha256("reqid=f6d83764e3854dc88bed0f87067e5b7f&timestamp=1719316714","example_signkey")
signature = Base64Encode(hashBytes)
```

4. 得到 signature 签名值为：

```
ZdFlMA2f+8e1/LNKiWBhB64zF4KVXhDvLHWhh3NwU=
```

5. 将 signature 值进行 urlencode（必须进行 URL 编码，否则将导致鉴权失败）后拼接得到最终请求 URL 为：

```
https://callback.example.com/callback_example_url?
reqid=f6d83764e3854dc88bed0f87067e5b7f&timestamp=1719316714&signature=ZdFlMA2f%2B8e1%2FLNKiWBhB64zF4KVK
XHdVbLHWhh3NwU%3D
```

返回参数

参数名称	类型	必选	描述
ReqId	string	是	请求Id，长度为32的uuid
MsgType	string	是	消息类型，标识本消息是依据什么场景发送的
MsgData	string	是	消息数据，以json格式发送。不同消息类型的结构不一样。

MsgData：直播间状态通知（MsgType=LiveRoomStatus）

参数名称	类型	必选	描述
LiveRoomId	string	是	直播间状态
Status	int	是	直播间状态，0-初始状态；1-视频流创建中；2-视频流创建完成；3-拆句完成；4-调度中；5-调度完成；6-关闭
VideoStreamPlayUrl	string	否	视频流播放地址（视频流创建完成才会返回）

MsgData：直播间脚本播报状态通知（MsgType=LiveRoomScriptStatus）

参数名称	类型	必选	描述
LiveRoomId	string	是	直播间状态
Seq	int	是	脚本拆句后的子句id
Text	string	是	脚本拆句后的子句内容
SpeakStatus	string	是	播报状态，SentenceStart-子句播报开始；SentenceOver-子句播报结束
Timestamp	string	是	时间戳，精度到秒

2D真人形象、声音定制 API

形象、声音定制 API 概述

最近更新时间：2024-11-13 15:25:52

本文档主要描述 腾讯云智能2D数字人 aPaas 平台的开放接口协议。

一、接口调用环境

正式环境：`https://gw.tvs.qq.com`

二、接口调用方式

[数智人 aPaas 接口调用方式](#)

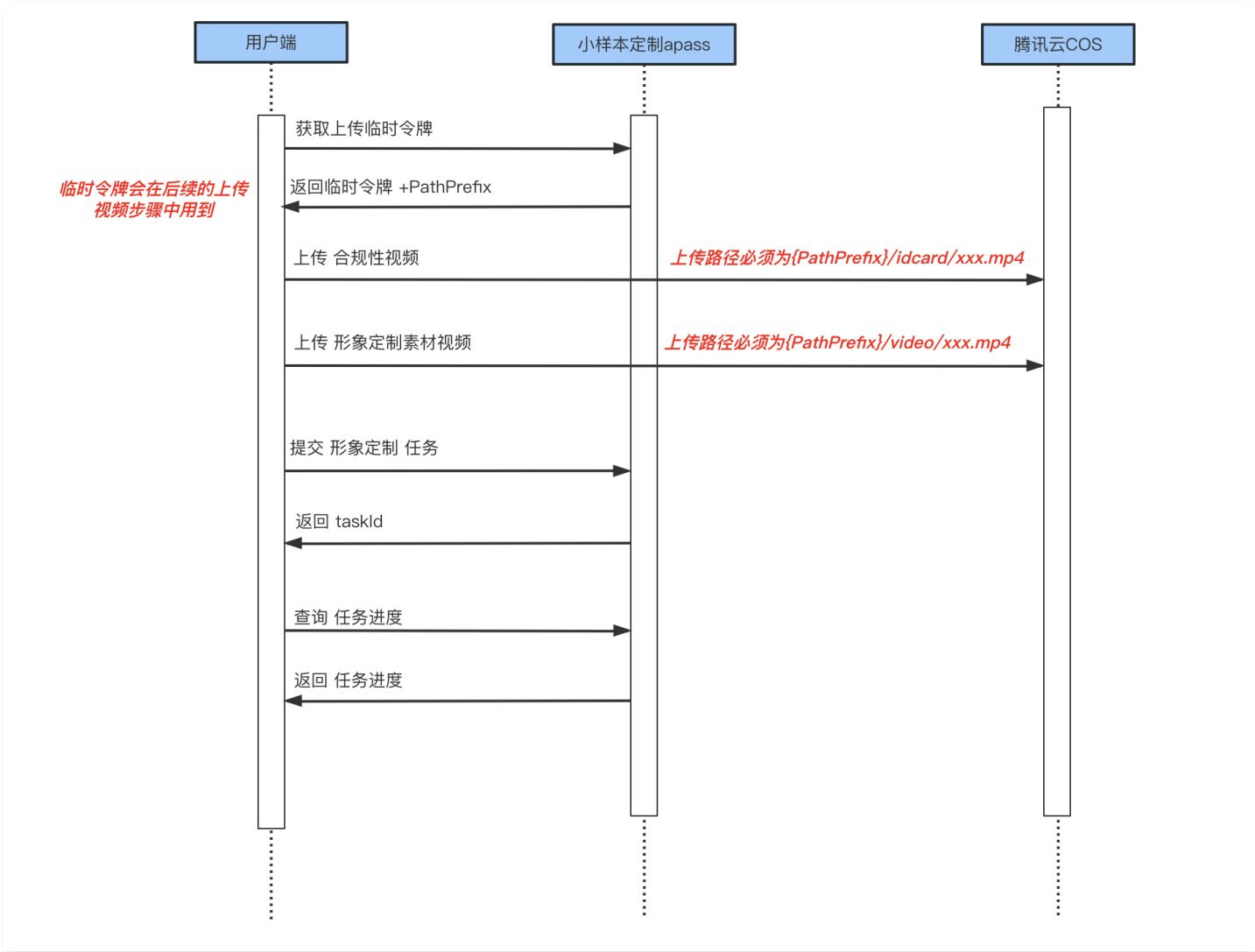
⚠ 注意：

如果通过腾讯云计费直接采购。appkey 和token请到 [数智人平台资产管理页面](#) 查询。

形象定制

接口调用逻辑示意图

最近更新时间：2025-03-14 16:06:02



1. 初始说明

- 用户端需先获取临时令牌，该令牌将在后续上传视频时使用。

2. 获取临时令牌

- 用户端发起“获取上传临时令牌”请求。
- 服务端返回包含临时令牌 + PathPrefix的响应，用于后续文件上传的鉴权和路径定位。

3. 视频上传阶段

- 用户端依次上传两类视频：
 - 合规性视频（用于内容审核），上传路径必须为 {PathPrefix}/idcard/xxx.mp4。
 - 形象定制素材视频（用于生成定制形象），上传路径必须为 {PathPrefix}/video/xxx.mp4。

4. 任务提交

- 用户端提交“形象定制”任务。
- 服务端返回 taskId（任务唯一标识），供后续查询进度使用。

5. 任务进度查询

- 用户端通过 taskId 定期查询任务进度。

- 服务端返回当前任务进度状态（如处理中、已完成等）。

获取上传临时令牌

最近更新时间：2024-11-18 17:24:52

在调用 [定制接口](#) 前，需要将素材文件上传到指定的云端存储。上传时，首先通过该接口获取上传临时令牌和指定的上传路径，再通过 [上传素材到腾讯云 COS](#) 接口上传对应的文件到云端存储。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/customservice/getuploadcredentials
Header Content-Type: application/json;charset=utf-8

请求参数

无

返回参数

参数	类型	必须	说明
Credentials	object	是	临时密钥信息
Credentials.Token	string	是	临时密钥令牌
Credentials.TmpSecretId	string	是	临时证书密钥 ID
Credentials.TmpSecretKey	object	是	临时证书密钥 Key
StartTime	int	是	临时证书有效起始时间，返回 Unix 时间戳，精确到秒
ExpiredTime	int	是	临时证书有效截止时间，返回 Unix 时间戳，精确到秒
PathPrefix	string	是	上传到 COS 上的路径前缀

返回示例

```
{
  "Header": {
    "RequestID": "c18ae4e49d89e408231c5b3ceea23c",
    "SessionID": "gzfa9105ca17310669340273673",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "Credentials": {
      "TmpSecretId": "AKID*****ib",
      "TmpSecretKey": "DVki*****k0=",
      "Token": "8oz1l*****49esQlGO940LX2tacXiCv4YS2qdBsiQ"
    },
    "ExpiredTime": 1731074134,
    "PathPrefix": "https://cos-test.ap-guangzhou.myqcloud.com/customer-pipeline/123/4506c29b-2904-45e2-b7d7-3052fe2493c2/",
    "StartTime": 1731066934
  }
}
```

最近更新时间: 2025-01-07 16:55:22

对接腾讯云 COS 的 SDK，调用其中的 putObject 方法将素材上传，上传时需要用到获取到的上传地址和上传凭证。

腾讯云 COS 的 SDK文档：[对象存储-SDK 概览](#)。

java 上传示例可以参考：[对象存储-上传对象](#)。

其他语言示例可以参考：[对象存储-SDK 概览](#)。

调用 SDK 中的上传对象方法时，需要用到临时密钥和存储桶相关信息，这些信息都可以通过 [获取上传临时令牌](#) 接口获取，下文举例说明。

临时密钥相关信息见下图:

```

"Payload": {
  "Credentials": {
    "TmpSecretId": "AKID59fPFGv2n-1ygDE9J",
    "TmpSecretKey": "Kg5hYnF+Ado0a3X1s2zEhgA",
    "Token": "FGSQNKc3Q90WCLeZ5jn2pdHo3cGEuPKab8e01dfcb983fbc8719373ed702600fw9B4mLcxks10FXxJAW-j1ldnREtKIZ8JIfam6Jo9iRn7I3IvvcHCsaMzC23lryFFcGHmE35KwV1y3XlRqYvsRB8Jj4CcDZ0vXVW7Rz7MAqr6Eb5xUyST1234LxjcupBGj5OW7Ij3HhNGMzK270414jgIsVvocC2_H6kWqILcT5QK56Nhm1c1ZnmGCLYQe0Hsrt-W2uoP_9y03SFLKvLcRnSowURVBumJqP0SIis_5FKK4q6W0JdqXPeGI7XxNBSzYzmcmjKjIiHnJbxQ6fHgVPGj5lEuG0ITGdXU5Pv6B6iqxwrfBBBuAaBvGhcd4gPIE0mKZ5fhcgCpXyNtIPLUxxTiM73GnQMhBcpj8PJtCso_c1GffuSCWnahTki-IsCH_vUXZWCYtI8lJcQeKtWbN1VCG4190CRz8NeusGNZiqzK6Jxklkin-6eGLCMTj3Y05D6NMROvtXzMLN20FS-yBj7QcWdhTe_FD7d-rliozkwjNZNzmmss9HKsYF43ggyXdm1guUW87EEPai33oZ6Q_zG7Z6jJl0BYcQkTW4gcXnj58PAugt9Uwfx8T_dT18DjJlRv-HD11J9Q0lF5RHuNyw09iQWcpXNh95d1kwhluzY5XH7H95ebm365v2p5Af5N13mqwhdAkREaCBFVMG9cMaarGyUvbaZecBfB6s0Adwz04500Ikdy7GZi5UwE7_Y_QkUqFEECj82RuEYheUCb1N20vXVW7Rz7MAqr6Eb5xUyST1234LxjcupBGj5OW7Ij3HhNGMzK270414jgIsVvocC2_H6kWqILcT5QK56Nhm1c1ZnmGCLYQe0Hsrt-W2uoP_9y03SFLKvLcRnSowURVBumJqP0SIis_5FKK4q6W0JdqXPeGI7XxNBSzYzmcmjKjIiHnJbxQ6fHgVPGj5lEuG0ITGdXU5Pv6B6iqxwrfBBBuAaBvGhcd4gPIE0mKZ5fhcgCpXyNtIPLUxxTiM73GnQMhBcpj8PJtCso_c1GffuSCWnahTki-IsCH_vUXZWCYtI8lJcQeKtWbN1VCG4190CRz8NeusGNZiqzK6Jxklkin-6eGLCMTj3Y05D6NMROvtXzMLN20FS-yBj7QcWdhTe_FD7d-rliozkwjNZNzmmss9HKsYF43ggyXdm1guUW87EEPai33oZ6Q_zG7Z6jJl0BYcQkTW4gcXnj58PAugt9Uwfx8T_dT18DjJlRv-HD11J9Q0lF5RHuNyw09iQWcpXNh95d1kwhluzY5XH7H95ebm365v2p5Af5N13mqwhdAkREaCBFVMG9cMaarGyUvbaZecBfB6s0Adwz04500Ikdy7GZi5UwE7_Y_QkUqFEECj82RuEYheUCb1N20vXVW7Rz7MAqr6Eb5xUyST1234LxjcupBGj5OW7Ij3HhNGMzK270414jgIsVvocC2_H6kWqILcT5QK56Nhm1c1ZnmGCLYQe0Hsrt-W2uoP_9y03SFLKvLcRnSowURVBumJqP0SIis_5FKK4q6W0JdqXPeGI7XxNBSzYzmcmjKjIiHnJbxQ6fHgVPGj5lEuG0ITGdXU5Pv6B6iqxwrfBBBuAaBvGhcd4gPIE0mKZ5fhcgCpXyNtIPLUxxTiM73GnQMhBcpj8PJtCso_c1GffuSCWnahTki-IsCH_vUXZWCYtI8lJcQeKtWbN1VCG4190CRz8NeusGNZiqzK6Jxklkin-6eGLCMTj3Y05D6NMROvtXzMLN20FS-yBj7QcWdhTe_FD7d-rliozkwjNZNzmmss9HKsYF43ggyXdm1guUW87EEPai33oZ6Q_zG7Z6jJl0BYcQkTW4gcXnj58PAugt9Uwfx8T_dT18DjJlRv-HD11J9Q0lF5RHuNyw09iQWcpXNh95d1kwhluzY5XH7H95ebm365v2p5Af5N13mqwhdAkREaCBFVMG9cMaarGyUvbaZecBfB6s0Adwz04500Ikdy7GZi5UwE7_Y_QkUqFEECj82RuEYheUCb1N20vXVW7Rz7MAqr6Eb5xUyST1234LxjcupBGj5OW7Ij3HhNGMzK270414jgIsVvocC2_H6kWqILcT5QK56Nhm1c1ZnmGCLYQe0Hsrt-W2uoP_9y03SFLKvLcRnSowURVBumJqP0SIis_5FKK4q6W0JdqXPeGI7XxNBSzYzmcmjKjIiHnJbxQ6fHgVPGj5lEuG0ITGdXU5Pv6B6iqxwrfBBBuAaBvGhcd4gPIE0mKZ5fhcgCpXyNtIPLUxxTiM73GnQMhBcpj8PJtCso_c1GffuSCWnahTki-IsCH_vUXZWCYtI8lJcQeKtWbN1VCG4190CRz8NeusGNZiqzK6Jxklkin-6eGLCMTj3Y05D6NMROvtXzMLN20FS-yBj7QcWdhTe_FD7d-rliozkwjNZNzmmss9HKsYF43ggyXdm1guUW87EEPai33oZ6Q_zG7Z6jJl0BYcQkTW4gcXnj58PAugt9Uwfx8T_dT18DjJlRv-HD11J9Q0lF5RHuNyw09iQWcpXNh95d1kwhluzY5XH7H95ebm365v2p5Af5N13mqwhdAkREaCBFVMG9cMaarGyUvbaZecBfB6s0Adwz04500Ikdy7GZi5UwE7_Y_QkUqFEECj82RuEYheUCb1N20vXVW7Rz7MAqr6Eb5xUyST1234LxjcupBGj5OW7Ij3HhNGMzK270414jgIsVvocC2_H6kWqILcT5QK56Nhm1c1ZnmGCLYQe0Hsrt-W2uoP_9y03SFLKvLcRnSowURVBumJqP0SIis_5FKK4q6W0JdqXPeGI7XxNBSzYzmcmjKjIiHnJbxQ6fHgVPGj5lEuG0ITGdXU5Pv6B6iqxwrfBBBuAaBvGhcd4gPIE0mKZ5fhcgCpXyNtIPLUxxTiM73GnQMhBcpj8PJtCso_c1GffuSCWnahTki-IsCH_vUXZWCYtI8lJcQeKtWbN1VCG4190CRz8NeusGNZiqzK6Jxklkin-6eGLCMTj3Y05D6NMROvtXzMLN20FS-yBj7QcWdhTe_FD7d-rliozkwjNZNzmmss9HKsYF43ggyXdm1guUW87EEPai33oZ6Q_zG7Z6jJl0BYcQkTW4gcXnj58PAugt9Uwfx8T_dT18DjJlRv-HD11J9Q0lF5RHuNyw09iQWcpXNh95d1kwhluzY5XH7H95ebm365v2p5Af5N13mqwhdAkREaCBFVMG9cMaarGyUvbaZecBfB6s0Adwz04500Ikdy7GZi5UwE7_Y_QkUqFEECj82RuEYheUCb1N20vXVW7Rz7MAqr6Eb5xUyST1234LxjcupBGj5OW7Ij3HhNGMzK270414jgIsVvocC2_H6kWqILcT5QK56Nhm1c1ZnmGCLYQe0Hsrt-W2uoP_9y03SFLKvLcRnSowURVBumJqP0SIis_5FKK4q6W0JdqXPeGI7XxNBSzYzmcmjKjIiHnJbxQ6fHgVPGj5lEuG0ITGdXU5Pv6B6iqxwrfBBBuAaBvGhcd4gPIE0mKZ5fhcgCpXyNtIPLUxxTiM73GnQMhBcpj8PJtCso_c1GffuSCWnahTki-IsCH_vUXZWCYtI8lJcQeKtWbN1VCG4190CRz8NeusGNZiqzK6Jxklkin-6eGLCMTj3Y05D6NMROvtXzMLN20FS-yBj7QcWdhTe_FD7d-rliozkwjNZNzmmss9HKsYF43ggyXdm1guUW87EEPai33oZ6Q_zG7Z6jJl0BYcQkTW4gcXnj58PAugt9Uwfx8T_dT18DjJlRv-HD11J9Q0lF5RHuNyw09iQWcpXNh95d1kwhluzY5XH7H95ebm365v2p5Af5N13mqwhdAkREaCBFVMG9cMaarGyUvbaZecBfB6s0Adwz04500Ikdy7GZi5UwE7_Y_QkUqFEECj82RuEYheUCb1N20vXVW7Rz7MAqr6Eb5xUyST1234LxjcupBGj5OW7Ij3HhNGMzK270414jgIsVvocC2_H6kWqILcT5QK56Nhm1c1ZnmGCLYQe0Hsrt-W2uoP_9y03SFLKvLcRnSowURVBumJqP0SIis_5FKK4q6W0JdqXPeGI7XxNBSzYzmcmjKjIiHnJbxQ6fHgVPGj5lEuG0ITGdXU5Pv6B6iqxwrfBBBuAaBvGhcd4gPIE0mKZ5fhcgCpXyNtIPLUxxTiM73GnQMhBcpj8PJtCso_c1GffuSCWnahTki-IsCH_vUXZWCYtI8lJcQeKtWbN1VCG4190CRz8NeusGNZiqzK6Jxklkin-6eGLCMTj3Y05D6NMROvtXzMLN20FS-yBj7QcWdhTe_FD7d-rliozkwjNZNzmmss9HKsYF43ggyXdm1guUW87EEPai33oZ6Q_zG7Z6jJl0BYcQkTW4gcXnj58PAugt9Uwfx8T_dT18DjJlRv-HD11J9Q0lF5RHuNyw09iQWcpXNh95d1kwhluzY5XH7H95ebm365v2p5Af5N13mqwhdAkREaCBFVMG9cMaarGyUvbaZecBfB6s0Adwz04500Ikdy7GZi5UwE7_Y_QkUqFEECj82RuEYheUCb1N20vXVW7Rz7MAqr6Eb5xUyST1234LxjcupBGj5OW7Ij3HhNGMzK270414jgIsVvocC2_H6kWqILcT5QK56Nhm1c1ZnmGCLYQe0Hsrt-W2uoP_9y03SFLKvLcRnSowURVBumJqP0SIis_5FKK4q6W0JdqXPeGI7XxNBSzYzmcmjKjIiHnJbxQ6fHgVPGj5lEuG0ITGdXU5Pv6B6iqxwrfBBBuAaBvGhcd4gPIE0mKZ5fhcgCpXyNtIPLUxxTiM73GnQMhBcpj8PJtCso_c1GffuSCWnahTki-IsCH_vUXZWCYtI8lJcQeKtWbN1VCG4190CRz8NeusGNZiqzK6Jxklkin-6eGLCMTj3Y05D6NMROvtXzMLN20FS-yBj7QcWdhTe_FD7d-rliozkwjNZNzmmss9HKsYF43ggyXdm1guUW87EEPai33oZ6Q_zG7Z6jJl0BYcQkTW4gcXnj58PAugt9Uwfx8T_dT18DjJlRv-HD11J9Q0lF5RHuNyw09iQWcpXNh95d1kwhluzY5XH7H95ebm365v2p5Af5N13mqwhdAkREa
```

存储桶相关信息需要从 PathPrefix 中解析见下图:

[illegible]

- bucket名称: 解析出来的值为 virtualhuman-cos-prod-1251316161。
- region名称: 解析出来的值为 accelerate。
- key(bucket 下的存储路径): 解析出来的值为 customer-pipeline/352/59c6f6cd-8f92-45f9-8352-64ca8fb92eaa/。

注意：

解析出来的值为 key 的前半部分，还需要根据上传素材类型来拼接后半部分，如上传的是 pdf 授权书素材，则需要拼接上 idcard/xxx.pdf，如下图所示，具体后半部分值的定义请参考 [定制接口](#) 中的相关参数说明。

腾讯云智能数智人 购买指南 数智人形象介绍 数智人平台操作指引 服务端 API 接入 数智人 aPaas 接口调用方式 2D真人形象、声音定制 API 概述 形象定制 接口调用逻辑示意图 获取上传临时令牌 上传素材到腾讯云 COS 定制接口 进度查询接口 效果确认接口 人脸检测任务创建 人脸检测任务状态查询 声音复刻（基础版） 接口调用逻辑示意图 音质检测任务创建接口 音质检测任务状态查询接口 获取上传临时令牌 上传素材到腾讯云 COS 定制接口 进度查询接口 效果确认接口 声音复刻（极速版）	IdentityWrittenCosUrl	string	否	pdf 格式授权书 URL，要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名 <code>[customer-pipeline/{数字}]/[uuid]/idcard/b.pdf</code> - 该格式主要适用于“书面版授权书”，以完整清晰的扫描件形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityCosUrl 二选一或二者都填	本页目录： 调用协议 请求参数 返回参数 请求示例 返回示例
	MaterialCosUrl	string	否	形象定制素材 url 地址要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源url地址后增加 video 路径，例如 <code>[customer-pipeline/{数字}]/[uuid]/video/c.mp4</code> - 视频大小，不超过5GB，4K不超过10G - 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 - 视频分辨率：1080P、4K（3840*2160），如定制高精版，则必须为4K - 视频长宽比：16:9 (9:16) - 视频帧率：不低于25fps，不超过60fps - 视频格式：mp4, mov 备注：形象定制时 video 路径必传 声音定制素材 url 地址要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如 <code>[customer-pipeline/{数字}]/[uuid]/audio/c.zip</code> - 压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色；压缩时不要新建文件夹，直接选中所有 wav 压缩 - 对于单个 zip 包里的音频注意事项： ①音频数量：每个 zip 包里可以包含1条或多条 wav 格式音频，总量不超过10条 ②音频大小：每个 zip 包里的音频大小总和不超过1G ③音频格式：每条音频需为 wav 格式，其他格式的音频请先转码为 wav 格式后再压缩成 zip 包 ④音频采样率：采样率大于等于24k即可，建议24k、36k	

定制接口

最近更新时间：2025-06-09 16:25:22

使 该接 提交定制的请求，通过 [进度查询接口](#) 查询定制的各个阶段以及各阶段的相关信息。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/customservice/make
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
AnchorName	string	是	主播名称： 1. 该名称主要用于标识该定制形象/声音，可以根据实际情况自定义 2. 命名参考：如果该主播定制的数量只有1个，可以直接以主播名称命名，例如“张三”；也可加上衣服名称来更好的标识，例如“张三蓝西服”等 3. 不超过50个字符，不少于2个字符，只允许有汉字、字母、数字、下划线、横线 4. 不能重名
MakeType	string	是	定制类别： IMAGE：2D小样本-标准版（专属口型）形象定制 IMAGE_GENERAL：2D小样本-标准版（通用口型）形象定制 IMAGE_4K：2D小样本-高精版（专属口型）形象定制 IMAGE_PHOTO：2D小样本（照片数智人）形象定制 ⚠ 注意：选择 IMAGE_PHOTO 类型时，强烈建议预先执行 人脸检测任务创建 以提高成功率 VOICE：声音复刻（基础版） ZERO_SHOT_VOICE：声音复刻（极速版）
IdentityCosUrl	string	否	视频格式授权书 URL，要求： 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/a.mp4 2. 该格式主要适用于“口述版授权书”，也可将“ 书面版授权书 ”拍成完整清晰的视频形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityWrittenCosUrl 二选一或二者都填 3. 大小要求：不超过5G
IdentityWrittenCosUrl	string	否	pdf 格式授权书 URL，要求： 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/b.pdf 2. 该格式主要适用于“ 书面版授权书 ”，以完整清晰的扫描件形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityCosUrl 二选一或二者都填 3. 大小要求：不超过10M
MaterialCosUrl	string	否	形象定制素材 url 地址要求： 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源url地址后增加 video 路径，例如/customer-pipeline/{数字}/{uuid}/video/c.mp4 2. 视频大小，不超过5GB，4K不超过10G 3. 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 4. 视频分辨率：1080P、4K（3840*2160），如定制高精版，则必须为4K 5. 视频长宽比：16:9 (9:16)

			6. 视频帧率：不低于25fps，不超过60fps 7. 视频格式：mp4, mov 备注：形象定制时 video 路径必传 声音定制素材 url 地址要求： 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如/customer-pipeline/{数字}/{uuid}/audio/c.zip 2. 压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色；压缩时不要新建文件夹，直接选中所有 wav 压缩 3. 对于单个 zip 包里的音频注意事项： ① 音频数量：每个 zip 包里可以包含1条或多条 wav 格式音频，总量不超过10条 ② 音频大小：每个 zip 包里的音频大小总和不超过1G ③ 音频格式：每条音频需为 wav 格式，其他格式的音频请先转码为 wav 格式后再压缩成 zip 包 ④ 音频采样率：采样率大于等于24k即可，建议24k、36k ⑤ 音频命名：名称不要有空格或者特殊字符，后缀名需要为小写 “.wav” 备注：除 ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型必填 照片数智人定制素材 url 地址要求： 1. url 地址为通过上传素材到腾讯云 COS上传到指定路径的资源 url 地址后增加 photo 路径，例如/customer-pipeline/{数字}/{uuid}/photo/example.png 2. 图片名称：不少于2个字符，只允许有汉字、字母、数字、下划、横线；图片格式：支持jpg、jpeg、png、webp；图片大小：不超过16M；图片宽高比：支持 1:1 9:16 16:9 4:3 3. 人物清晰正面照，脸部居中，表情自然，嘴部闭合 备注：照片定制时 photo 路径必传
SexType	string	是	性别： MALE：男 FEMALE：女
Notes	string	否	定制备注信息，100字以内
TextDriver	string	否	用于生成驱动 demo 的文本内容，字符长度在4-1000之间（包括 ssmll 标签,单个中文算一个字符）
VoiceDriverCosFile	string	否	生成驱动 demo 的音频文件路径，要求： 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如:/customer-pipline/{数字}/{uuid}/audio/example.wav 2. 音频文件大小不超过10M，格式支持wav、mp3、wma、m4a、aac
Audiold	string	否	ZERO_SHOT_VOICE 定制类型必填:填写 查询音质检测任务进度 通过后返回的 Audiold
TimbreInfo	TimbreTemplate	否	形象定制时若需要指定音色，或尝试用视频里的声音快速复刻，则填充该字段。
IsHaveBackground	bool	否	形象定制类型，训练出来的形象是否带原始背景。默认“否”，即不带原始背景，在应用过程中可以按需更换背景 true：保留拍摄背景 false：抠除拍摄背景（支持更换背景，仅支持绿幕视频和白墙视频）
RemoveBackground	int	否	当 IsHaveBackground=false 时，进一步填充该参数。默认值是1。 1：抠除拍摄背景-绿幕视频 2：抠除拍摄背景-白幕视频
AlphaCosUrl	string	否	自定义 Alpha 通道视频地址，视频上传路径参照 形象定制素材 路径
ZeroshortLang	string	否	ZERO_SHOT_VOICE 定制类型的语言类型，chi-中文普通话，eng-英文，can-粤语，默认chi
PhotoVersion	int	否	默认0 0：快速版。在10分钟内即可完成数智人定制，只支持头部轻微晃动，面部特征稳定。建议执行 人脸检测任务 (CheckerVersion=1)

			1: 优享版。在1个小时内可完成数智人定制，头部和身体都支持轻微晃动。建议执行 人脸检测任务 (CheckerVersion=2)。
--	--	--	--

TimbreTemplate 信息参数

参数	类型	必须	说明
TimbreCode	string	否	制作文本驱动 demo 时指定的音色 code
Speed	string	否	制作文本驱动 demo 时指定音色的语速,默认1.0。语速（范围[0.5–1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快）
IsWithTimbre	bool	否	尝试用视频里的声音快速复刻，额外生成 demo（由于视频里的音质质量不能保证，可能会存在快速复刻失败的情况）。该 demo 不影响其他 demo 的正常合成。

返回参数

参数	类型	必须	说明
TaskId	string	是	制作的任务 ID，携带 TaskId 访问 进度查询接口 ，可获得制作进度和制作结果

请求示例

```
{
  "Header": {},
  "Payload": {
    "AnchorName": "马老师",
    "MakeType": "IMAGE_GENERAL",
    "IdentityCosUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/f30e4d52-00ba-4e36-8a26-26c348bf786e/idcard/identity.mp4",
    "MaterialCosUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/78675441-08bf-43f2-93db-f3936068bce7/video/test.mp4",
    "SexType": "FEMALE",
    "RemoveBackground": 2
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "c18aeee4e49d89e408231c5b3ceea23c",
    "SessionID": "gzfa9105ca17310669340273673",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "TaskId": "1212_29934_w"
  }
}
```


进度查询接口

最近更新时间：2025-03-07 12:14:53

通过TaskId来查询任务的制作进度和结果。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/customservice/getprogress
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
TaskId	string	是	定制接口 返回的 TaskId

返回参数

参数	类型	必须	说明
Status	string	是	任务制作状态： MAKING：进行中 SUCCESS：制作成功 FAIL：制作失败
FailMessage	string	否	当 Status: FAIL 时，返回制作失败原因，便于排查问题
Stage	string	否	当前定制阶段（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageId	string	否	定制阶段 Id（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageV2	string	否	（自2024.1.1之后发起的进度查询请求，以当前字段为准） 当 Status: MAKING 时，返回任务进行中的具体阶段： SUBMIT: 素材提交 AUDIT: 人工审核 PROCESS: 预处理（仅形象定制有） TRAIN: 智能训练 CONFIRM: 效果确认（照片定制无此阶段） END: 完成
StatusV2	string	否	当 StageV2 处于某个阶段时，该阶段的具体状态： WAIT_PROCESS: 待处理 PROCESSING: 处理中 SUCCESS: 成功 FAIL: 失败
StageInfo	string	否	处于某个阶段的额外信息，JSON 格式的字符串，不同阶段的 JSON 格式类型见下述表格，目前只有效果确认阶段（StageV2 为 CONFIRM）有值。
AssetList	Array of [AssetInfo]	否	返回定制的形象或音色信息，目前支持返回的定制类型有： IMAGE: 2D小样本-标准版（专属口型）形象定制 IMAGE_GENERAL: 2D小样本-标准版（通用口型）形象定制 IMAGE_4K: 2D小样本-高精版（专属口型）形象定制 IMAGE_PHOTO: 2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE: 声音复刻（极速版） VOICE: 声音复刻（基础版）

不同阶段 StageInfo 额外信息见下表：

StageV2	类型	描述	示例
CONFIRM: 效果确认阶段	string of [Customer ConfirmCo ntent]	当阶段处于效果确认阶段（StageV2: CONFIRM），且状态处于待处理中时（StatusV2: WAIT_PROCESS），返回待确认效果 Demo 视频，客户可以根据返回的 Demo 视频通过调用 效果确认接口 对 Demo 视频效果进行认同/驳回操作	<pre>{\\"TrainResult\\": [\\"url1\\",\\"url2\\"],\\"RejectReason\\":\\"\\",\\"TrainRecord\\":[{\\"StartTime\\":\\"2023-11-01 23:00:17\\",\\"Record\\":[\\"url1\\",\\"url2\\"]}]}</pre>

CustomerConfirmContent

参数	类型	必须	说明
TrainResult	Array of [string]	是	供客户进行效果确认的 Demo 视频 url 地址
RejectReason	string	是	数智人侧不接受重训的理由
TrainRecord	Array of [TrainRecord]	否	历史训练的 Demo 视频清单

TrainRecord

参数	类型	必须	说明
StartTime	string	是	历史训练时间
Record	Array of [string]	是	历史训练的 Demo 视频 url 地址

AssetInfo

参数	类型	必须	说明
VirtualmanTypeCode	string	否	数智人类型 Code，参考 附录3-数智人类型 ，形象定制返回
VirtualmanKey	string	否	数智人 virtualmankey，形象定制返回
Resolution	string	否	数智人分辨率，形象定制返回
OriginZoom	string	否	缩放系数，保留三位小数，形象定制返回
VirtualmanResourceId	int32	否	数智人资源 ID，形象定制返回
TimbreKey	string	否	音色key，声音复刻返回
TimbreSample	string	否	音色demo链接，声音复刻返回
MakeType	string	是	定制类型： IMAGE_PHOTO：2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE：声音复刻（极速版）

请求示例

```
{  
  "Header": {},  
}
```

```
"Payload": {  
  "TaskId": "1212_29934_w"  
}  
}
```

返回示例

```
{  
  "Header": {  
    "RequestID": "gz1a74c25f17030396653416949",  
    "SessionID": "gz1a74c25f17030396653416950",  
    "DialogID": "",  
    "Code": 0,  
    "Message": "ok"  
  },  
  "Payload": {  
    "StageInfo": "{\\\"TrainResult\\\":[\\\"url1\\\",\\\"url2\\\"],\\\"RejectReason\\\":\\\"\\\",\\\"TrainRecord\\\":  
[\\\"StartTime\\\":\\\"2023-11-01 23:00:17\\\",\\\"Record\\\":[\\\"url1\\\",\\\"url2\\\"]]}\"",  
    "Status": "MAKING",  
    "StatusV2": "WAIT_PROCESS",  
    "FailMessage": "",  
    "StageV2": "CONFIRM"  
  }  
}
```

效果确认接口

最近更新时间：2024-11-13 15:22:12

当通过进度查询接口查到处于待确认效果阶段时（Status:MAKING，StageV2:CONFIRM，StatusV2:WAIT_PROCESS），对 demo 视频进行效果确认。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/customservice/customerconfirm
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
TaskId	string	是	定制接口 返回的 TaskId
Operate	string	是	操作： AGREE：认同 REJECT：驳回
Reason	string	否	驳回原因，当 Operate:REJECT 时必填，最多300字

返回参数

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "1212_29934_w",
    "Operate": "AGREE"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "gz1a74c25f17030396653416949",
    "SessionID": "gz1a74c25f17030396653416950",
    "DialogID": "",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {}
}
```

人脸检测任务创建

最近更新时间：2024-12-25 16:47:53

传入 RequestId，图片的 url 和检测版本，创建人脸检测任务，返回 TaskId。使用 TaskId 可以进行进度查询。通过预先的人脸检测，可以提高照片数智人的定制成功率。

调用协议

HTTPS + JSON
POST /v2/ivh/facedetectserver/facedetectservice/facedetect
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
RequestId	string	是	本次请求的唯一标识ID 示例值：569ae497-f266-4565-93bd-93aaf822524e
ImageUrl	string	是	要进行检测的图片的 url
CheckerVersion	int	否	检测版本，支持项：[1,2]。不传默认为1。 值为1时： - 检查图片的比例，目前支持[9:16, 16:9, 1:1, 4:3] - 检查结果只包含人脸框信息和角度 值为2时： - 不检查图片比例 - 检查结果包含遮挡比例，建议裁剪的区域坐标等信息。根据建议裁剪的区域对原图片进行裁剪，然后再进行数智人制作效果会更好。

返回参数

参数	类型	必须	说明
TaskId	string	是	携带TaskId访问 人脸检测任务状态查询 接口，可获得检测进度和检测结果

请求示例

```
{
  "Header": {},
  "Payload": {
    "RequestId": "569ae497-f266-4565-93bd-93aaf822524e",
    "ImageUrl": "https://virtualhuman-cos-prod-1251316161.cos.ap-gangzhou.myqcloud.com/customer-pipeline/704/e60fe7c8-ec69-44cf-bc5e-c4c754614e82/photo/48922.png",
    "CheckerVersion": 2
  }
}
```

返回示例

```
{
  "Header": {
    "RequestId": "gz1a74c25f17030396653416949",
    "SessionID": "gz1a74c25f17030396653416950",
    "DialogID": "",
    "Code": 0,
  }
}
```

```
    "Message": "ok"
  },
  "Payload": {
    "TaskId": "569ae497-f266-4565-93bd-93aaf822524e"
  }
}
```

人脸检测任务状态查询

最近更新时间：2024-11-13 15:25:52

通过 TaskId 来查询人脸检测任务的进度和结果。

调用协议

HTTPS + JSON
POST /v2/ivh/facedetectserver/facedetectservice/getfacedetectstate
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
TaskId	string	是	人脸检测任务创建 接口返回的TaskId

返回参数

参数	类型	必须	说明
TaskId	string	是	任务ID
TaskState	string	是	任务状态： - WAITING: 任务等待中 - PROCESSING: 任务执行中 - SUCCESS: 任务执行成功 - FAILED: 任务执行失败
ErrorCode	int	否	任务执行失败时有值，错误码，可参考附录
ErrorMsg	string	否	任务执行失败时有值，错误信息
IsPassDetect	bool	否	任务执行成功时有值，是否通过人脸检测
Faces	Array of FaceInfo	否	任务执行成功时有值，检测到的人脸框

FaceInfo

参数	类型	必须	说明
Box	BoxInfo	否	检测到的人脸框
Rotation	float	否	人脸的角度
Score	float	否	人脸置信度 *
AreaRatio	float	否	人脸区域占比 *
FaceRadio	float	否	脸部遮挡比例 0~1，越大遮挡越大 *
FeatureRadio	float	否	五官遮挡比例 0~1，越大遮挡越大 *
Yaw	float	否	yaw角度，-90~90 表明侧脸程度，越大姿态越大 *
Pitch	float	否	pitch角度，-90~90 表明俯仰程度，越大姿态越大 *
MaxCrop	BoxInfo	否	建议的最大裁剪区域 *

MinCrop	BoxInfo	否	建议的最小裁剪区域 *
---------	---------	---	-------------

 **说明：**
表格中带 * 的参数只有创建时 CheckerVersion=2 的任务才会有检测结果。

BoxInfo

参数	类型	必须	说明
Left	int	否	左上角的x坐标点
Top	int	否	左上角的y坐标点
Width	int	否	宽度
Height	int	否	高度

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "569ae497-f266-4565-93bd-93aaf822524e"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "gz1a74c25f17030396653416949",
    "SessionID": "gz1a74c25f17030396653416950",
    "DialogID": "",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "ErrorCode": 0,
    "IsPassDetect": true,
    "Faces": [
      {
        "Yaw": 11.440315246582,
        "Pitch": -3.8738098144531,
        "MaxCrop": {
          "Top": 0,
          "Height": 512,
          "Width": 512,
          "Left": 0
        },
        "MinCrop": {
          "Top": 0,
          "Height": 512,
          "Width": 512,
          "Left": 0
        },
        "Box": {
          "Top": 103,
```



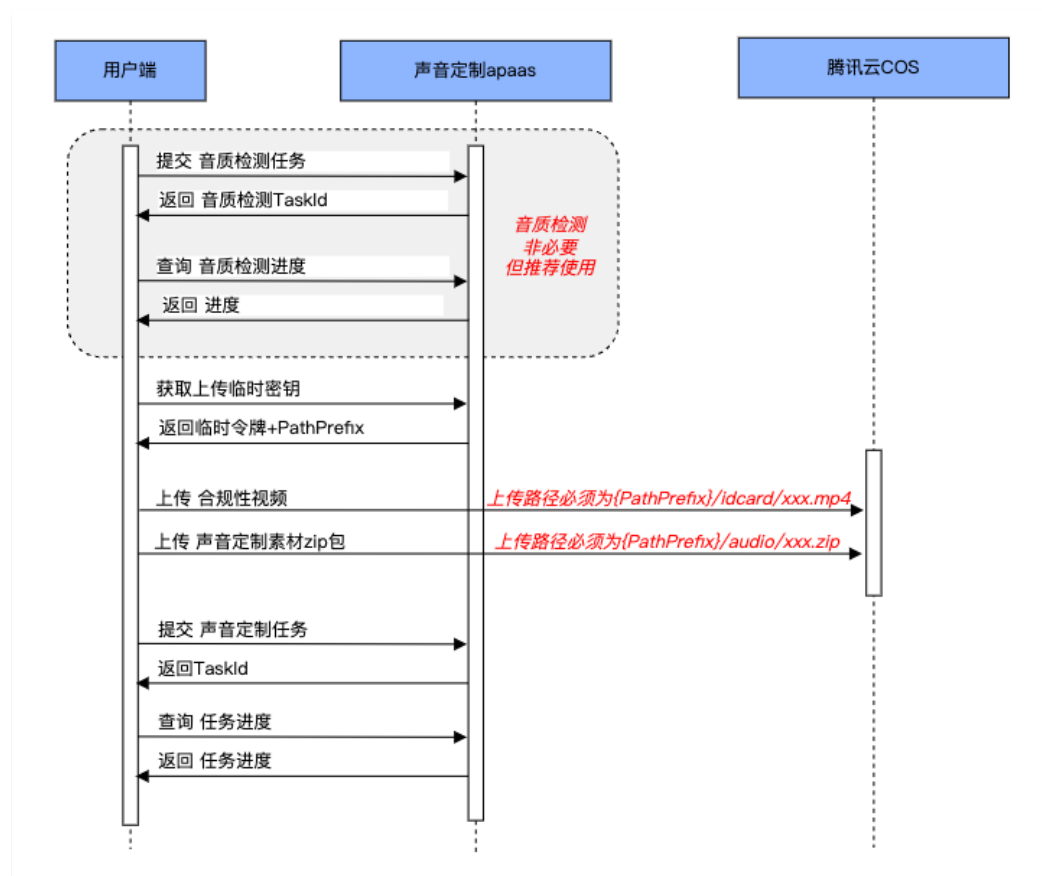
```
        "Height": 406,
        "Width": 320,
        "Left": 85
    },
    "Rotation": -7,
    "Score": 1,
    "AreaRatio": 0.49560546875,
    "FaceRadio": 0,
    "FeatureRadio": 0
}
],
"TaskState": "SUCCESS",
"ErrorMsg": "",
"TaskId": "569ae497-f266-4565-93bd-93aaf822524e"
}
```

附录一：错误码字典

代码	含义	处理方式
100001	请求参数错误	修正请求入参
100002	参数取值错误	修正请求入参
100003	图片格式不支持	选择新的图片
100004	图片尺寸不支持	调整图片尺寸重新检测
100005	图片比例不支持	调整图片比例重新检测
200001	内部错误	联系管理员
200002	文件下载失败	联系管理员
200003	任务不存在	联系管理员
300001	图片中没有检测到人脸	更换新的图片
300002	图片中检测到多个人脸	更换新的图片
300003	人脸框不满足要求，可能过大或者过小	更换新的图片

声音复刻（基础版） 接口调用逻辑示意图

最近更新時間：2025-03-14 16:06:02



1. 音质检测阶段 (可选)

- 用户提交音质检测任务，系统返回一个音质检测 TaskId。
- 用户可选择性查询任务进度，系统返回当前进度。

2. 文件上传准备

- 用户需先获取上传临时密钥，系统返回临时令牌 + PathPrefix。

3. 文件上传要求

- 合规性视频：必须上传至路径 `(PathPrefix)/idcard/xxx.mp4`。
- 声音定制素材压缩包：必须上传至路径 `(PathPrefix)/audio/xxx.zip`。

4. 声音定制任务提交

- 用户提交声音定制任务，系统返回声音定制 TaskId。
- 用户可查询任务进度，系统返回当前状态（如处理中、完成等）。

音质检测任务创建接口

最近更新时间：2025-06-09 16:25:22

可以通过该接口提前检测声音质量，提升定制效率。

支持的音质指标：

- 信噪比指数：大于等于25合格
- 混响指数：大于等于30合格
- 截幅指数：小于等于10合格
- 有效时长：大于等于10分钟合格
- 文本准确率：（当传入“音频参考文本”时该指标有效）
 - 中文普通话：大于等于85%合格
 - 英文：大于等于50%合格
 - 粤语：大于等于70%合格

调用协议

HTTPS + JSON

POST /v2/ivh/algotaskserver/algotaskservice/audioevaluation

Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
AudioUrl	string	是	待检测音频的 URL，音频格式目前支持 wav、mp3、m4a、aac、wma 和 asf 格式，建议文件大小不超过20M
ReferenceText	string	否	音频参考文本，可选。不填表示不需要“文本准确率”指标
Language	string	否	语言类型。chi-中文普通话，eng-英文，can-粤语。默认chi

返回参数

参数	类型	必须	说明
TaskId	string	是	已创建音质检测任务的 id

请求示例

```
{
  "Header": {},
  "Payload": {
    "AudioUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/63232275-a349-4f27-8e33-aa4ac791094c/voice_record/hello.wav",
    "ReferenceText": "蓝蓝的天空银河里，有只小白船，船上有棵桂花树，白兔在游玩。渡过那条银河水，走向云彩国，走过那个云彩国，飘向明天",
    "Language": "chi"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "1e2c786806e6a59d8c2c70bdbd94f70e",
  }
}
```

```
"SessionID": "gza802cc9317328691207007711",  
"DialogID": "",  
"Code": 0,  
"Message": "ok"  
},  
"Payload": {  
  "TaskId": "837a583863b745e59391456943ffc22f"  
}  
}
```

音质检测任务状态查询接口

最近更新时间：2024-11-29 11:26:32

通过 TaskId 来查询音质检测任务的状态和结果。

调用协议

HTTPS + JSON
POST /v2/ivh/algotaskserver/algotaskservice/getaudioevaluationstate
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
TaskId	string	是	音质检测任务创建接口 返回的 TaskId

返回参数

参数	类型	必须	说明
TaskState	string	是	任务状态： CREATED：已创建 QUEUING：排队中 PROCESSING：处理中 SUCCESS：任务完成 FAIL：任务失败
Result	object	是	音质检测任务结果
Result.Snr	object	是	音频的信噪比指标
Result.Snr.Value	float	是	音频的信噪比指标测量值
Result.Snr.QualifiedRange	Array of [float]	是	音频的信噪比指标合格范围 [lower_limit,upper_limit]
Result.Snr.Description	string	是	音频的信噪比指标说明
Result.C50	object	是	音频的混响指标
Result.C50.Value	float	是	音频的混响指标测量值
Result.C50.QualifiedRange	Array of [float]	是	音频的混响指标合格范围 [lower_limit,upper_limit]
Result.C50.Description	string	是	音频的混响指标说明
Result.Clipping	object	是	音频的截幅指标
Result.Clipping.Value	float	是	音频的截幅指标测量值
Result.Clipping.QualifiedRange	Array of [float]	是	音频的截幅指标合格范围 [lower_limit,upper_limit]
Result.Clipping.Description	string	是	音频的截幅指标说明
Result.Duration	object	是	音频的时长指标
Result.Duration.Value	float	是	音频的时长
Result.Duration.QualifiedRange	Array of [float]	是	音频的时长合格范围 [lower_limit,upper_limit]
Result.Duration.Descriptio	string	是	音频的时长指标说明

n			
Result.Wer	object	是	音频的文本准确率指标
Result.Wer.Value	float	是	文本准确率，仅“参考文本”不为空时有效
Result.Wer.QualifiedRange	Array of [float]	是	文本准确率合格范围 [lower_limit,upper_limit]
Result.Wer.Description	string	是	文本准确率说明
Result.AlignDetails	array of [TextAlign]	是	文本对齐结果
Result.Sentences	array of [Sentences]	是	短句截取时间
ErrorMsg	string	是	任务失败提示信息

TextAlign

参数	类型	必须	说明
Detail	Array of [string]	是	文本每个字符对齐结果，如["好","好"]/["好","<eps>"], 其中"<eps>"为占位符

Sentences

参数	类型	必须	说明
StartMs	int	是	句子开始毫秒数
EndMs	int	是	句子结束毫秒数
Sentence	string	是	句子内容

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "63232275-a349-4f27-8e33-aa4ac791094c"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "Message": "",
    "RequestID": "28499a-cf6d-4f7b-b7fa-4a5a54662b38"
  },
  "Payload": {
    "Status": "SUCCESS",
    "Result": {
      "Snr": {
        "Value": 72.67048,
        "QualifiedRange": [
          30,
```

```
    2147483600
  ],
  "Description": "音频的信噪比, 越大越好"
},
"C50": {
  "Value": 57.082405,
  "QualifiedRange": [
    30,
    2147483600
  ],
  "Description": "混响指标, 越大越好"
},
"Clipping": {
  "QualifiedRange": [
    -2147483600,
    0
  ],
  "Description": "截幅指数: 小于等于10合格",
  "Value": 0
},
"Duration": {
  "Value": 24.21,
  "QualifiedRange": [
    600,
    2147483600
  ],
  "Description": "有效时长: 单位秒, 大于等于10分钟合格"
},
"Wer": {
  "Value": 0.91358024,
  "QualifiedRange": [
    0.7,
    1
  ],
  "Description": "文本准确率: 大于等于70%合格"
},
"Sentences": [
  {
    "StartMs": 340,
    "EndMs": 4510,
    "Sentence": "激光祛斑是利用激光的光热解原理, "
  },
  {
    "StartMs": 4510,
    "EndMs": 6040,
    "Sentence": "激光照射皮肤后, "
  },
  {
    "StartMs": 6040,
    "EndMs": 8640,
    "Sentence": "会被皮肤病变部位的色素吸收, "
  },
  {
    "StartMs": 8900,
    "EndMs": 12660,
    "Sentence": "而激光产生的高能量就会将这些色素击碎, "
  },
  {
    "StartMs": 12890,
```

```

        "EndMs": 17155,
        "Sentence": "色素碎屑会随人体的各项循环排除体外，"
    },
    {
        "StartMs": 17155,
        "EndMs": 19050,
        "Sentence": "从而达到祛斑的效果。"
    }
],
"AlignDetails": [
    {
        "Detail": [
            "<eps>",
            "大"
        ]
    },
    {
        "Detail": [
            "家",
            "家"
        ]
    },
    {
        "Detail": [
            "<eps>",
            "好"
        ]
    }
]
},
"ErrorMsg": ""
}
}

```


获取上传临时令牌

最近更新时间：2024-11-18 17:24:52

在调用 [定制接口](#) 前，需要将素材文件上传到指定的云端存储。上传时，首先通过该接口获取上传临时令牌和指定的上传路径，再通过 [上传素材到腾讯云 COS](#) 接口上传对应的文件到云端存储。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/customservice/getuploadcredentials
Header Content-Type: application/json;charset=utf-8

请求参数

无

返回参数

参数	类型	必须	说明
Credentials	object	是	临时密钥信息
Credentials.Token	string	是	临时密钥令牌
Credentials.TmpSecretId	string	是	临时证书密钥 ID
Credentials.TmpSecretKey	object	是	临时证书密钥 Key
StartTime	int	是	临时证书有效起始时间，返回 Unix 时间戳，精确到秒
ExpiredTime	int	是	临时证书有效截止时间，返回 Unix 时间戳，精确到秒
PathPrefix	string	是	上传到 COS 上的路径前缀

返回示例

```
{
  "Header": {
    "RequestID": "c18ae4e49d89e408231c5b3ceea23c",
    "SessionID": "gzfa9105ca17310669340273673",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "Credentials": {
      "TmpSecretId": "AKID*****ib",
      "TmpSecretKey": "DVki*****k0=",
      "Token": "8oz1l*****49esQlGO940LX2tacXiCv4YS2qdBsiQ"
    },
    "ExpiredTime": 1731074134,
    "PathPrefix": "https://cos-test.ap-guangzhou.myqcloud.com/customer-pipeline/123/4506c29b-2904-45e2-b7d7-3052fe2493c2/",
    "StartTime": 1731066934
  }
}
```

最近更新时间: 2025-01-07 16:55:22

对接腾讯云 COS 的 SDK，调用其中的 putObject 方法将素材上传，上传时需要用到获取到的上传地址和上传凭证。

腾讯云 COS 的 SDK文档：[对象存储-SDK 概览](#)。

java 上传示例可以参考：[对象存储-上传对象](#)。

其他语言示例可以参考：[对象存储-SDK 概览](#)。

⚠ 注意:

调用 SDK 中的上传对象方法时，需要用到临时密钥和存储桶相关信息，这些信息都可以通过 [获取上传临时令牌](#) 接口获取，下文举例说明。

临时密钥相关信息见下图:

```

"Payload": {
  "Credentials": {
    "TmpSecretId": "AKID59fPpGv2n-1ygDE91...MsMmaC21SRpG9V2WpA8xPsW000EM8E",
    "TmpSecretKey": "Kg5hYNF+AdoAa3Xl2sZehgA...nU=",
    "Token": "fG5qNkc3Q90WCLe25jn2pdHo3cGEuPKab8e01dfcb983fbc8719373ed702600fw9B4mmLcxks10FXxJAW-j1dnREtKIz8Ifam6Jo9rIn73IIvvcHcsaMZC31ryfGHCmS35Z0vXVW7r2MAqr6EbX5uY5T1234LxjcupBGjSOW7Ij3HHNgMZK270414jgIsVvocC2_H6kWq1LT5QK56Nhmic1ZNmgCLYQe0Hsrt-W2uoP_9y03SFLKvlrCnSowURVBumJqP0SI...nBS5zyzmckjKjIhngBxQ6EGhVPgJ51EuGOitGXuPsV96BQ1xzqwrBBbuAaBvGhcd4gPIE0mZK5fhcgCcPyXntIPlUxxTiM73nGnQMBhpcj8PJtCso_clGffuSCwnahTKI-IsCh_V419D9CzBneuSGN2iqxR6jxkLkin-6egLCMtj3Y0506NMROvtzMLN20FS-yb7jqCwDhTe_FD7d-zLiozkjWNZNmss59HKsYF43qgyXDm1guUw87EEPaI33oZ6Q_zG7Z6jL0yOv8T_dT18DjJlRv-HD1139Q01FR5HuNyw09iQWcpXnKh95d1kWhlurY5XH7H95ebm365v2PfA5N13mqWdhdAkREaCBFWMG9qCmaARgyUvbvzaECBfB6s0Adwz04500IKdy7GzjS1wUE7VZ0vXVW7r2MAqr6EbX5uY5T1234LxjcupBGjSOW7Ij3HHNgMZK270414jgIsVvocC2_H6kWq1LT5QK56Nhmic1ZNmgCLYQe0Hsrt-W2uoP_9y03SFLKvlrCnSowURVBumJqP0SI...JHUBmic1U09xrouplQsKv0f2XQ2v7D6d5pZhkzK216j784dE19JwgHL6XHdFV3ZRIUWEdN8ojxpvBuNRTD2EzRVs6PpJlV15gHayiy_yT4xk34ijI01VMRC0UpucJ6Kcu-zmx5Fp...baziWwkdH-UF5BCqinQp0p0Qnmwm3r1QTWE84n9HHu4LTyFA0vLdNqdxA6tnCCVwA6MYmFLTakf9x541sQWUv26HF118wLnF0zKIPQz9DUCXxz1IoC55TKK-qx0rI03EPQn0i_w20r9MYML5mnS2GnxuK61w-cS1xzR6uv8BPoX3WVs1BK1bz7a-2fds4Xb0gMrEdwuto5KN76y69gYYAKalrVr8nDmT8Phc1pzEHgdrv6KHuUmFR6Y-nd6n__yE12FNndm_MbvzrRx0Qn...JHUBmic1U09xrouplQsKv0f2XQ2v7D6d5pZhkzK216j784dE19JwgHL6XHdFV3ZRIUWEdN8ojxpvBuNRTD2EzRVs6PpJlV15gHayiy_yT4xk34ijI01VMRC0UpucJ6Kcu-zmx5Fp...baziWwkdH-UF5BCqinQp0p0Qnmwm3r1QTWE84n9HHu4LTyFA0vLdNqdxA6tnCCVwA6MYmFLTakf9x541sQWUv26HF118wLnF0zKIPQz9DUCXxz1IoC55TKK-qx0rI03EPQn0i_w20r9MYML5mnS2GnxuK61w-cS1xzR6uv8BPoX3WVs1BK1bz7a-2fds4Xb0gMrEdwuto5KN76y69gYYAKalrVr8nDmT8Phc1pzEHgdrv6KHuUmFR6Y-nd6n__yE12FNndm_MbvzrRx0Qn...Q9kr-m91Uo_nrByfm2Rbz1houAv28GyW3JMj-HBE2V113h6dUvYGCeAqPnT2S85CDLz1CD3tY7I"
  },
  "ExpiredTime": "1736240204",
  "PathPrefix": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipe-...dd-4ed9-81c6-748a8a10dc1a",
  "StartTime": "1736233004"
}

```

存储桶相关信息需要从 PathPrefix 中解析见下图:

```

"Payload": {
  "Credentials": {
    "TmpSecretId": "AKID6mp_7lbof_ow5_DTYz[REDACTED]Qbz1MlvTfFjxqbg1FrN9",
    "TmpSecretKey": "9pkrCgY6duNwhJ4/[REDACTED]l87PGk=",
    "Token":
      "fGSqNkc3Q90WCLeZ5jn2pdHu3cGEuPaK23738bd8cb693d875f4e2ead9f6ff446w9B4mmLcxks10FXXJAW-jtudiJnfPdPht0RvKEC3PshPrv2nrXPLjwTst8A-mGxVxkpM-YIE-
      kF9izYreejnxz26uCoCL61AXy39v6CdCHXoWkdKwZyc1LWio2QVMrshS0uKr26H770MmvDjH9vK19I9V8K4K7mJaE3GDisk5tJpAqFtNd5tSRGPsIErdgHuf2jUeF9wvnuV1yy
      KgOhj7GwzwWkw4j2jtr1B6fbNzYfTxPhvijLwDC50Dg613pgiZqwtu7GtoUxdFOXWQNYlaHiiz1NRr25EL0t_07tAsdL6Ae7uj-3lbykQbU88e5JHSRvs4p0GvE58ILUNw5Ihpt
      Pmc2z29mXtSvwwkx2Y-KiQBYHsF81PaEuMhK4eU60303xdjd5_h6N4C0u6QFet3PUY77Kp1t7L4K0670JqERriKEgBypyzYqvVhAd4YZ0-TABrhymYPUVukdbmzW4Idim8tzuU16f
      Afq132ux4IX-NrieAGMRifks1q1XfRfHo2i1QealyUubddGoyVvK6y5eIjiko5SMMoDrPBpZyoZJwaLHt12mbm-hDKHvcgBe-bMe7qs7ze6NiyqwrECDCrVmdTA8j0EVPKGuY8r
      HnhmQvc8W77hBuQvgcCPxR-X_xAUd-6ydv5vGi_8mLcd07h35Nvk0TrXyNuMTmqjTEboisVC74jP68pQMZNy--UULZ9C_tI6Y938xwf5xPB70MzrMTP7IEInMPqsDzGWWglR-U-JB
      [REDACTED]
      B0d6Qf3XN2PCL1q1gkJsU8jZBLQ4AkC5XVwfEbSwh4TcLn500yqp6ks7cEXBZeAhAoWRJ_APBDNTI72-XnlYdgBbcjY6G3xXuDRbBGUs8QrCi_3mMYNMMyNq4f1bAnR1sdAnJfF
      EFwPKyLD4cs4bEbbatbMt-1u37ovXD4EUcNyrR6YZ1jNws27n8gAL0t9FDEb-aHICXaYyGc5gEXLGHfb16gJsomjwZaOfpXtI4RMPBP_6D0f6RGEOIF_0r5AwH_9_QsGwIXuW18r
      Q8Cf0vWU-zk0Q2F4x9CpQwUdsN118FIYC3I3sYTViS5gIh2P_-L0dkrwjplz5-yRglbDjh5baljXMLGqhipeohMsYxQSHDfgkAm1WFwqpp4vKt-0S-1kuYv94BqTRkM8M2_OvoRN2
      iI657HD13z4dGncSvQ_ R495R1NK471Cne-NDoqRMbCvFYX-mgDS7Fht2t2FsvkxueYX0iRmnpA_475qQdpYn1_PNYConJ4YISICpX3XUNeoFRTJc2miwSENL7vH0H0kM02hN1k
      K51s0d130zesLRkCbD0sAZyJd_5e1pEF7SaulMdm7FMkwq3th0CwacrlY062DNcWEQBBM-00uwr7Cp3d7BdhakLAXn-RAO-C12vgMw-wx1LBQ57ixN7yJE28ukCsZ68Qc14_ Q0UsGGuY8r
      1AGhjpbCN_7XzH-WAPGfgrGXJWlDT-zkHJYRPjMbB19zx2WCifCycChtbz27XREQCN6BwmnXvtqApjKAdr9ZQtTk0ANdvm53zz1jxd0HttAkJL9YCPp0iN_2YjHI5t4e7PMc_Its
      E18iC0lqs57wtRErFQspFgtAWR3dmjgTYTHehPIP_K42k5WA2t6g303iz_3_zamGTZngv4uDPTVYr-QpfqAQMQDPaRL132QTkariiMa8F8d10DhjuvBDPzIWzA6D85tsTJMHJ
      3KLJUiqfAgwVcnHwHJmrdmJo_AVyZ9yUaqNqH0G0AIEUtYlzh35emcD6vb1V2ZKci3gJgrTSg"
  }
},
"ExpiredTime": 1736241459,
"PathPrefix": "https://virtualhuman-cos-prod-1251316163.cos.accelerate.myqcloud.com/customer-pipeline/352/59c6f6cd-8f92-45f9-8352-64ca8fb92eaa/",
"StartTime": 1736234259

```

- bucket名称: 解析出来的值为 virtualhuman-cos-prod-1251316161。
- region名称: 解析出来的值为 accelerate。
- key(bucket 下的存储路径): 解析出来的值为 customer-pipeline/352/59c6f6cd-8f92-45f9-8352-64ca8fb92eaa/。

注意：

解析出来的值为 key 的前半部分，还需要根据上传素材类型来拼接后半部分，如上传的是 pdf 授权书素材，则需要拼接上 idcard/xxx.pdf，如下图所示，具体后半部分值的定义请参考 [定制接口](#) 中的相关参数说明。

腾讯云智能数智人 购买指南 数智人形象介绍 数智人平台操作指引 服务端 API 接入 数智人 aPaas 接口调用方式 2D真人形象、声音定制 API 概述 形象定制 接口调用逻辑示意图 获取上传临时令牌 上传素材到腾讯云 COS 定制接口 进度查询接口 效果确认接口 人脸检测任务创建 人脸检测任务状态查询 声音复刻（基础版） 接口调用逻辑示意图 音质检测任务创建接口 音质检测任务状态查询接口 获取上传临时令牌 上传素材到腾讯云 COS 定制接口 进度查询接口 效果确认接口 声音复刻（极速版）	IdentityWrittenCosUrl	string	否	pdf 格式授权书 URL，要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名 <code>[customer-pipeline/{数字}]/[uuid]/idcard/b.pdf</code> - 该格式主要适用于“书面版授权书”，以完整清晰的扫描件形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityCosUrl 二选一或二者都填	调用协议 请求参数 返回参数 请求示例 返回示例
	MaterialCosUrl	string	否	形象定制素材 url 地址要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源url地址后增加 video 路径，例如 <code>[customer-pipeline/{数字}]/[uuid]/video/c.mp4</code> - 视频大小，不超过5GB，4K不超过10G - 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 - 视频分辨率：1080P、4K（3840*2160），如定制高精版，则必须为4K - 视频长宽比：16:9 (9:16) - 视频帧率：不低于25fps，不超过60fps - 视频格式：mp4, mov 备注：形象定制时 video 路径必传 声音定制素材 url 地址要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如 <code>[customer-pipeline/{数字}]/[uuid]/audio/c.zip</code> - 压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色；压缩时不要新建文件夹，直接选中所有 wav 压缩 - 对于单个 zip 包里的音频注意事项： ①音频数量：每个 zip 包里可以包含1条或多条 wav 格式音频，总量不超过10条 ②音频大小：每个 zip 包里的音频大小总和不超过1G ③音频格式：每条音频需为 wav 格式，其他格式的音频请先转码为 wav 格式后再压缩成 zip 包 ④音频采样率：采样率大于等于24k即可，建议24k、36k	

定制接口

最近更新时间：2024-12-25 17:12:02

使 该接 提交定制的请求，通过 [进度查询接口](#) 查询定制的各个阶段以及各阶段的相关信息。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/customservice/make
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
AnchorName	string	是	主播名称： 1. 该名称主要用于标识该定制形象/声音，可以根据实际情况自定义 2. 命名参考：如果该主播定制的数量只有1个，可以直接以主播名称命名，例如“张三”；也可加上衣服名称来更好的标识，例如“张三蓝西服”等 3. 不超过50个字符，不少于2个字符，只允许有汉字、字母、数字、下划线、横线 4. 不能重名
MakeType	string	是	定制类别： IMAGE：2D小样本-标准版（专属口型）形象定制 IMAGE_GENERAL：2D小样本-标准版（通用口型）形象定制 IMAGE_4K：2D小样本-高精版（专属口型）形象定制 IMAGE_PHOTO：2D小样本（照片数智人）形象定制 ⚠ 注意：选择 IMAGE_PHOTO 类型时，强烈建议预先执行 人脸检测任务创建 以提高成功率 VOICE：声音复刻（基础版） ZERO_SHOT_VOICE：声音复刻（极速版）
IdentityCosUrl	string	否	视频格式授权书 URL，要求： 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/a.mp4 2. 该格式主要适用于“口述版授权书”，也可将“ 书面版授权书 ”拍成完整清晰的视频形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityWrittenCosUrl 二选一或二者都填
IdentityWrittenCosUrl	string	否	pdf 格式授权书 URL，要求： 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/b.pdf 2. 该格式主要适用于“ 书面版授权书 ”，以完整清晰的扫描件形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityCosUrl 二选一或二者都填
MaterialCosUrl	string	否	形象定制素材 url 地址要求： 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源url地址后增加 video 路径，例如/customer-pipeline/{数字}/{uuid}/video/c.mp4 2. 视频大小，不超过5GB，4K不超过10G 3. 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 4. 视频分辨率：1080P、4K（3840*2160），如定制高精版，则必须为4K 5. 视频长宽比：16:9 (9:16) 6. 视频帧率：不低于25fps，不超过60fps 7. 视频格式：mp4, mov

			声音定制素材 url 地址要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如/customer-pipeline/{数字}/{uuid}/audio/c.zip - 压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色；压缩时不要新建文件夹，直接选中所有 wav 压缩 - 对于单个 zip 包里的音频注意事项： ①音频数量：每个 zip 包里可以包含1条或多条 wav 格式音频，总量不超过10条 ②音频大小：每个 zip 包里的音频大小总和不超过1G ③音频格式：每条音频需为 wav 格式，其他格式的音频请先转码为 wav 格式后再压缩成 zip 包 ④音频采样率：采样率大于等于24k即可，建议24k、36k ⑤音频命名：名称不要有空格或者特殊字符，后缀名需要为小写 “.wav” 备注：除 ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型必填 照片数智人定制素材 url 地址要求： - url地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 photo 路径，例如/customer-pipeline/{数字}/{uuid}/photo/example.png - 图片名称：不少于2个字符，只允许有汉字、字母、数字、下划、横线；图片格式：支持jpg、jpeg、png、webp；图片大小：不超过16M；图片宽高比：支持 1:1 9:16 16:9 4:3 - 人物清晰正面照，脸部居中，表情自然，嘴部闭合
SexType	string	是	性别： MALE：男 FEMALE：女
Notes	string	否	定制备注信息，100字以内
TextDriver	string	否	用于生成驱动 demo 的文本内容，字符长度在4-1000之间（包括 ssml 标签,单个中文算一个字符）
VoiceDriverCosFile	string	否	生成驱动 demo 的音频文件路径，要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如:/customer-pipeline/{数字}/{uuid}/audio/example.wav - 音频文件大小不超过10M，格式支持wav、mp3、wma、m4a、aac
Audiold	string	否	ZERO_SHOT_VOICE 定制类型必填:填写 查询音质检测任务进度 通过后返回的 Audiold
TimbreInfo	TimbreTemplate	否	形象定制时若需要指定音色，或尝试用视频里的声音快速复刻，则填充该字段。
IsHaveBackground	bool	否	形象定制类型，训练出来的形象是否带原始背景。默认“否”，即不带原始背景，在应用过程中可以按需更换背景 true：保留拍摄背景 false：抠除拍摄背景（支持更换背景，仅支持绿幕视频和白墙视频）
RemoveBackground	int	否	当 IsHaveBackground=false 时，进一步填充该参数。默认值是1。 1：抠除拍摄背景-绿幕视频 2：抠除拍摄背景-白幕视频
AlphaCosUrl	string	否	自定义 Alpha 通道视频地址，视频上传路径参照 形象定制素材 路径
ZeroshortLang	string	否	ZERO_SHOT_VOICE 定制类型的语言类型，chi-中文普通话，eng-英文，can-粤语，默认chi
PhotoVersion	int	否	默认0 0：快速版。在10分钟内即可完成数智人定制，只支持头部轻微晃动，面部特征稳定。建议执行 人脸检测任务 (CheckerVersion=1) 1：优享版。在1个小时内可完成数智人定制，头部和身体都支持轻微晃动。建议执行 人脸检测任务 (CheckerVersion=2)。

TimbreTemplate 信息参数

参数	类型	必须	说明
TimbreCode	string	否	制作文本驱动 demo 时指定的音色 code
Speed	string	否	制作文本驱动 demo 时指定音色的语速,默认1.0。语速（范围[0.5–1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快）
IsWithTimbre	bool	否	尝试用视频里的声音快速复刻，额外生成 demo（由于视频里的音质质量不能保证，可能会存在快速复刻失败的情况）。该 demo 不影响其他 demo 的正常合成。

返回参数

参数	类型	必须	说明
TaskId	string	是	制作的任务 ID，携带 TaskId 访问 进度查询接口 ，可获得制作进度和制作结果

请求示例

```
{
  "Header": {},
  "Payload": {
    "AnchorName": "张三",
    "MakeType": "VOICE",
    "IdentityCosUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/f30e4d52-00ba-4e36-8a26-26c348bf786e/idcard/identity.mp4",
    "SexType": "FEMALE",
    "MaterialCosUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/110/77ef6290-0991-4e45-9451-0ef8587985a3/audio/张三.zip"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "c18aeee4e49d89e408231c5b3ceea23c",
    "SessionID": "gzfa9105ca17310669340273673",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "TaskId": "1212_29934_w"
  }
}
```

进度查询接口

最近更新时间：2025-03-07 12:14:53

通过TaskId来查询任务的制作进度和结果。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/customservice/getprogress
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
TaskId	string	是	定制接口 返回的 TaskId

返回参数

参数	类型	必须	说明
Status	string	是	任务制作状态： MAKING：进行中 SUCCESS：制作成功 FAIL：制作失败
FailMessage	string	否	当 Status: FAIL 时，返回制作失败原因，便于排查问题
Stage	string	否	当前定制阶段（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageId	string	否	定制阶段 Id（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageV2	string	否	（自2024.1.1之后发起的进度查询请求，以当前字段为准） 当 Status: MAKING 时，返回任务进行中的具体阶段： SUBMIT: 素材提交 AUDIT: 人工审核 PROCESS: 预处理（仅形象定制有） TRAIN: 智能训练 CONFIRM: 效果确认（照片定制无此阶段） END: 完成
StatusV2	string	否	当 StageV2 处于某个阶段时，该阶段的具体状态： WAIT_PROCESS: 待处理 PROCESSING: 处理中 SUCCESS: 成功 FAIL: 失败
StageInfo	string	否	处于某个阶段的额外信息，JSON 格式的字符串，不同阶段的 JSON 格式类型见下述表格，目前只有效果确认阶段（StageV2 为 CONFIRM）有值。
AssetList	Array of [AssetInfo]	否	返回定制的形象或音色信息，目前支持返回的定制类型有： IMAGE: 2D小样本-标准版（专属口型）形象定制 IMAGE_GENERAL: 2D小样本-标准版（通用口型）形象定制 IMAGE_4K: 2D小样本-高精版（专属口型）形象定制 IMAGE_PHOTO: 2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE: 声音复刻（极速版） VOICE: 声音复刻（基础版）

不同阶段 StageInfo 额外信息见下表：

StageV2	类型	描述	示例
CONFIRM: 效果确认阶段	string of [Customer ConfirmCo ntent]	当阶段处于效果确认阶段（StageV2: CONFIRM），且状态处于待处理中时（StatusV2: WAIT_PROCESS），返回待确认效果 Demo 视频，客户可以根据返回的 Demo 视频通过调用 效果确认接口 对 Demo 视频效果进行认同/驳回操作	<pre>{\"TrainResult\": [\"url1\", \"url2\"], \"RejectReason\": \"\", \"TrainRecord\": [{\"StartTime\": \"2023-11-01 23:00:17\", \"Record\": [\"url1\", \"url2\"]}]}</pre>

CustomerConfirmContent

参数	类型	必须	说明
TrainResult	Array of [string]	是	供客户进行效果确认的 Demo 视频 url 地址
RejectReason	string	是	数智人侧不接受重训的理由
TrainRecord	Array of [TrainRecord]	否	历史训练的 Demo 视频清单

TrainRecord

参数	类型	必须	说明
StartTime	string	是	历史训练时间
Record	Array of [string]	是	历史训练的 Demo 视频 url 地址

AssetInfo

参数	类型	必须	说明
VirtualmanTypeCode	string	否	数智人类型 Code，参考 附录3-数智人类型 ，形象定制返回
VirtualmanKey	string	否	数智人 virtualmankey，形象定制返回
Resolution	string	否	数智人分辨率，形象定制返回
OriginZoom	string	否	缩放系数，保留三位小数，形象定制返回
VirtualmanResourceId	int32	否	数智人资源 ID，形象定制返回
TimbreKey	string	否	音色key，声音复刻返回
TimbreSample	string	否	音色demo链接，声音复刻返回
MakeType	string	是	定制类型： IMAGE_PHOTO：2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE：声音复刻（极速版）

请求示例

```
{  
  "Header": {},  
}
```



```
"Payload": {  
  "TaskId": "1212_29934_w"  
}  
}
```

返回示例

```
{  
  "Header": {  
    "RequestID": "gz1a74c25f17030396653416949",  
    "SessionID": "gz1a74c25f17030396653416950",  
    "DialogID": "",  
    "Code": 0,  
    "Message": "ok"  
  },  
  "Payload": {  
    "StageInfo": "{\\\"TrainResult\\\":[\\\"url1\\\",\\\"url2\\\"],\\\"RejectReason\\\":\\\"\\\",\\\"TrainRecord\\\":  
[\\\"StartTime\\\":\\\"2023-11-01 23:00:17\\\",\\\"Record\\\":[\\\"url1\\\",\\\"url2\\\"]]}\"",  
    "Status": "MAKING",  
    "StatusV2": "WAIT_PROCESS",  
    "FailMessage": "",  
    "StageV2": "CONFIRM"  
  }  
}
```

效果确认接口

最近更新时间：2024-11-13 15:22:12

当通过进度查询接口查到处于待确认效果阶段时（Status:MAKING，StageV2:CONFIRM，StatusV2:WAIT_PROCESS），对 demo 视频进行效果确认。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/customservice/customerconfirm
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
TaskId	string	是	定制接口 返回的 TaskId
Operate	string	是	操作： AGREE：认同 REJECT：驳回
Reason	string	否	驳回原因，当 Operate:REJECT 时必填，最多300字

返回参数

请求示例

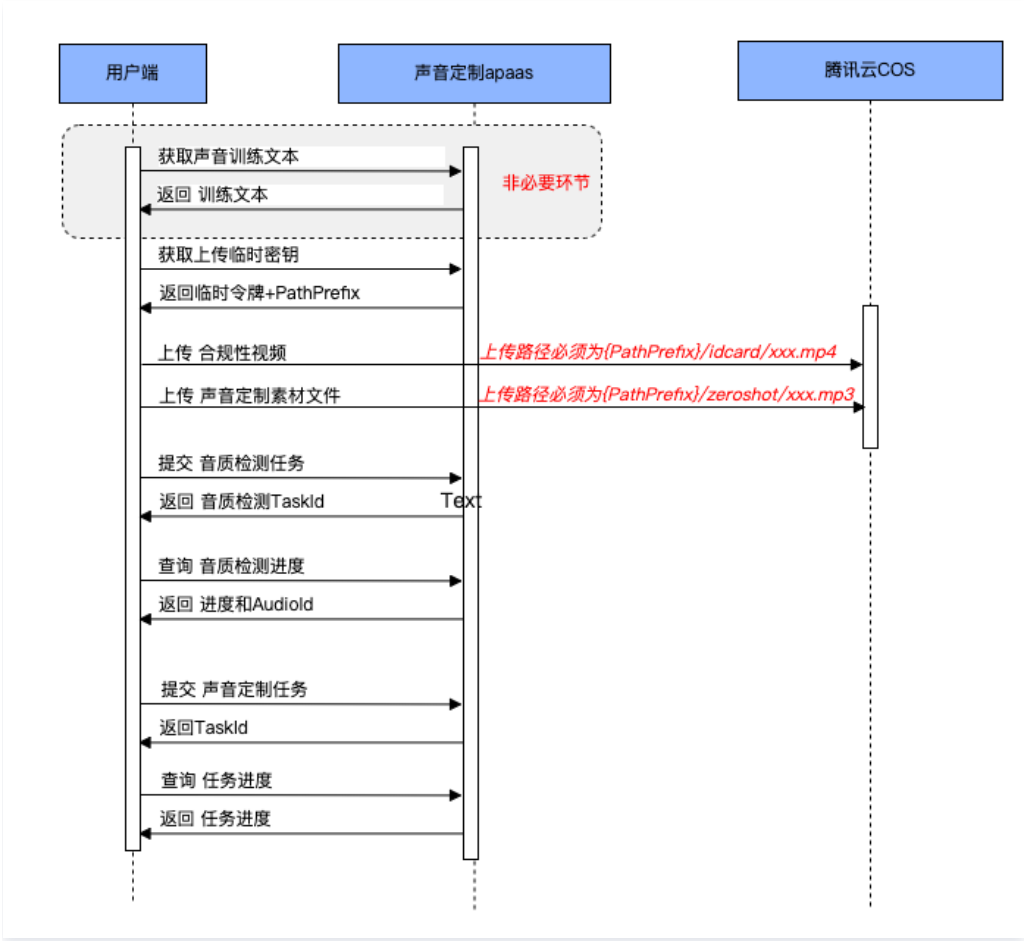
```
{
  "Header": {},
  "Payload": {
    "TaskId": "1212_29934_w",
    "Operate": "AGREE"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "gz1a74c25f17030396653416949",
    "SessionID": "gz1a74c25f17030396653416950",
    "DialogID": "",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {}
}
```

声音复刻（极速版） 接口调用逻辑示意图

最近更新时间：2025-03-14 16:06:02



- 1. 初始化阶段（可选）
 - 用户可获取声音训练文本，系统返回训练文本。
- 2. 文件上传准备
 - 用户获取上传临时密钥，系统返回临时令牌 + PathPrefix。
- 3. 文件上传规则
 - 合规性视频：必须上传至路径 {PathPrefix}/idcard/xxx.mp4 。
 - 声音定制素材文件：必须上传至路径 {PathPrefix}/zeroshot/xxx.mp3 。
- 4. 音质检测任务
 - 用户提交音质检测任务，系统返回音质检测 TaskId。
 - 用户可查询检测进度，系统返回进度和 Audioid。
- 5. 声音定制任务
 - 用户提交声音定制任务，系统返回 TaskId。
 - 用户可查询任务进度，系统返回任务进度。

获取声音训练文本

最近更新时间：2024-11-29 17:15:42

接口说明

调用该接口获取需要朗读的文档内容，然后使用该文本内容诵读成音频文件当作声音制作素材。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/zeroshotservice/gettext
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
Language	String	否	获取对应语言的文本：chi-中文普通话，eng-英文，can-粤语

返回参数

参数	类型	必须	说明
Texts	Array of [Text]	是	可选择的文本列表

Text

参数	类型	必须	说明
TextId	string	是	文本ID
Text	string	是	文本

请求示例

```
{
  "Header": {},
  "Payload": {
    "Language": "chi"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "DialogID": "",
    "Message": "",
    "SessionID": "gz92439a3317328659145933517",
    "RequestID": "9e2bf1f4dafe4250409132f78defcaea"
  },
  "Payload": {
    "Texts": [
      {
        "TextId": "ec28499a-cf6d-4f7b-b7fa-4a5a54662b38",
        "Text": "本次降价的300多款产品 其中肉类零食占比较多一些"
```

```
    }  
  ]  
}  

```

获取上传临时令牌

最近更新时间：2024-11-18 17:24:52

在调用 [定制接口](#) 前，需要将素材文件上传到指定的云端存储。上传时，首先通过该接口获取上传临时令牌和指定的上传路径，再通过 [上传素材到腾讯云 COS](#) 接口上传对应的文件到云端存储。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/customservice/getuploadcredentials
Header Content-Type: application/json;charset=utf-8

请求参数

无

返回参数

参数	类型	必须	说明
Credentials	object	是	临时密钥信息
Credentials.Token	string	是	临时密钥令牌
Credentials.TmpSecretId	string	是	临时证书密钥 ID
Credentials.TmpSecretKey	object	是	临时证书密钥 Key
StartTime	int	是	临时证书有效起始时间，返回 Unix 时间戳，精确到秒
ExpiredTime	int	是	临时证书有效截止时间，返回 Unix 时间戳，精确到秒
PathPrefix	string	是	上传到 COS 上的路径前缀

返回示例

```
{
  "Header": {
    "RequestID": "c18ae4e49d89e408231c5b3ceea23c",
    "SessionID": "gzfa9105ca17310669340273673",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "Credentials": {
      "TmpSecretId": "AKID*****ib",
      "TmpSecretKey": "DVki*****k0=",
      "Token": "8oz1l*****49esQlGO940LX2tacXiCv4YS2qdBsiQ"
    },
    "ExpiredTime": 1731074134,
    "PathPrefix": "https://cos-test.ap-guangzhou.myqcloud.com/customer-pipeline/123/4506c29b-2904-45e2-b7d7-3052fe2493c2/",
    "StartTime": 1731066934
  }
}
```

上传素材到腾讯云 COS

最近更新时间: 2025-01-07 16:55:22

对接腾讯云 COS 的 SDK，调用其中的 `putObject` 方法将素材上传，上传时需要用到获取到的上传地址和上传凭证。

腾讯云 COS 的 SDK文档：[对象存储-SDK 概览](#)。

java 上传示例可以参考：[对象存储-上传对象](#)。

其他语言示例可以参考：[对象存储-SDK 概览](#)。

⚠ 注意:

调用 SDK 中的上传对象方法时，需要用到临时密钥和存储桶相关信息，这些信息都可以通过 [获取上传临时令牌](#) 接口获取，下文举例说明。

临时密钥相关信息见下图:

```

"Payload": {
  "Credentials": {
    "TmpSecretId": "AKID59fFPGv2n-1ygDE9l",
    "TmpSecretKey": "Kg5hYnF+Ado0a3X1s2zEhgA",
    "Token": "fGSQnKc3Q90WCLeZ5jn2pdHo3cGEUpKAb8e01dfcb983fbc8719373ed702600fw9B4mmLcxks10FXJXAW-j1dnREtKIZ8jIfam6Jo9rIn73IIvvhcSaMZC231ryFFcGHmE35KvW1y3X1RQYvRB8j4CccDZovXVW7rZ7MAqr6EbX5uY5T1234LxjcupBgjSOW7Ij3HhNGMKZk270414jgIsVvocC2_H6KwQilc150K56Nhm1c2LnmGCLYQe0Hsrt-Wu2op_9y03SFLKvlcrnSowURVBumJqP0Siis_5FKK4q6w0JdqXPeGI7XxNBsZyzcmjkmJiIhnjBxQ6iGhVPgJ5lEuG0itGxuPsV96BQixqwrFBbBuAaBvGhcd4gPIE0mZk5fhcgCcPyXntIPlUxxTmI73GnQMhbcjp8PJtCso_clGffuSvCwnahmTKi-IsCh_VUXZWCYTl8jCqeKtWbN1VCG419DCRzBneusGNZiqxR6qkxLkin-6egLcMTj3Y05D6mNRvotXzMLN20FS-ybJ7qCwdHte_Fd7d-rLiozjkWNZnmssms9HKsYF43qgyXDM1guUW87EEPai330z6_zG7Z6jJl0bYcQkTW4gcXnj58PAugt9UwIxa8T_d187djJlRv-HD1139001F5RHuNyw09iQwcpXnKh95d1kwhlurY5Xh7H95ebm365v2pfa5N13mqwdhkdAKReaCBFVGm9qCmaaRgyUvzbaeCEBf6s0Adwz04500Ikdy76ZjSiUwE7v_QkUqFEECj82RuEYheUcb1nZovXVW7rZ7MAqr6EbX5uY5T1234LxjcupBgjSOW7Ij3HhNGMKZk270414jgIsVvocC2_H6KwQilc150K56Nhm1c2LnmGCLYQe0Hsrt-Wu2op_9y03SFLKvlcrnSowURVBumJqP0Siis_5FKK4q6w0JdqXPeGI7XxNBsZyzcmjkmJiIhnjBxQ6iGhVPgJ5lEuG0itGxuPsV96BQixqwrFBbBuAaBvGhcd4gPIE0mZk5fhcgCcPyXntIPlUxxTmI73GnQMhbcjp8PJtCso_clGffuSvCwnahmTKi-IsCh_VUXZWCYTl8jCqeKtWbN1VCG419DCRzBneusGNZiqxR6qkxLkin-6egLcMTj3Y05D6mNRvotXzMLN20FS-ybJ7qCwdHte_Fd7d-rLiozjkWNZnmssms9HKsYF43qgyXDM1guUW87EEPai330z6_zG7Z6jJl0bYcQkTW4gcXnj58PAugt9UwIxa8T_d187djJlRv-HD1139001F5RHuNyw09iQwcpXnKh95d1kwhlurY5Xh7H95ebm365v2pfa5N13mqwdhkdAKReaCBFVGm9qCmaaRgyUvzbaeCEBf6s0Adwz04500Ikdy76ZjSiUwE7v_QkUqFEECj82RuEYheUcb1nZovXVW7rZ7MAqr6EbX5uY5T1234LxjcupBgjSOW7Ij3HhNGMKZk270414jgIsVvocC2_H6KwQilc150K56Nhm1c2LnmGCLYQe0Hsrt-Wu2op_9y03SFLKvlcrnSowURVBumJqP0Siis_5FKK4q6w0JdqXPeGI7XxNBsZyzcmjkmJiIhnjBxQ6iGhVPgJ5lEuG0itGxuPsV96BQixqwrFBbBuAaBvGhcd4gPIE0mZk5fhcgCcPyXntIPlUxxTmI73GnQMhbcjp8PJtCso_clGffuSvCwnahmTKi-IsCh_VUXZWCYTl8jCqeKtWbN1VCG419DCRzBneusGNZiqxR6qkxLkin-6egLcMTj3Y05D6mNRvotXzMLN20FS-ybJ7qCwdHte_Fd7d-rLiozjkWNZnmssms9HKsYF43qgyXDM1guUW87EEPai330z6_zG7Z6jJl0bYcQkTW4gcXnj58PAugt9UwIxa8T_d187djJlRv-HD1139001F5RHuNyw09iQwcpXnKh95d1kwhlurY5Xh7H95ebm365v2pfa5N13mqwdhkdAKReaCBFVGm9qCmaaRgyUvzbaeCEBf6s0Adwz04500Ikdy76ZjSiUwE7v_QkUqFEECj82RuEYheUcb1nZovXVW7rZ7MAqr6EbX5uY5T1234LxjcupBgjSOW7Ij3HhNGMKZk270414jgIsVvocC2_H6KwQilc150K56Nhm1c2LnmGCLYQe0Hsrt-Wu2op_9y03SFLKvlcrnSowURVBumJqP0Siis_5FKK4q6w0JdqXPeGI7XxNBsZyzcmjkmJiIhnjBxQ6iGhVPgJ5lEuG0itGxuPsV96BQixqwrFBbBuAaBvGhcd4gPIE0mZk5fhcgCcPyXntIPlUxxTmI73GnQMhbcjp8PJtCso_clGffuSvCwnahmTKi-IsCh_VUXZWCYTl8jCqeKtWbN1VCG419DCRzBneusGNZiqxR6qkxLkin-6egLcMTj3Y05D6mNRvotXzMLN20FS-ybJ7qCwdHte_Fd7d-rLiozjkWNZnmssms9HKsYF43qgyXDM1guUW87EEPai330z6_zG7Z6jJl0bYcQkTW4gcXnj58PAugt9UwIxa8T_d187djJlRv-HD1139001F5RHuNyw09iQwcpXnKh95d1kwhlurY5Xh7H95ebm365v2pfa5N13mqwdhkdAKReaCBFVGm9qCmaaRgyUvzbaeCEBf6s0Adwz04500Ikdy76ZjSiUwE7v_QkUqFEECj82RuEYheUcb1nZovXVW7rZ7MAqr6EbX5uY5T1234LxjcupBgjSOW7Ij3HhNGMKZk270414jgIsVvocC2_H6KwQilc150K56Nhm1c2LnmGCLYQe0Hsrt-Wu2op_9y03SFLKvlcrnSowURVBumJqP0Siis_5FKK4q6w0JdqXPeGI7XxNBsZyzcmjkmJiIhnjBxQ6iGhVPgJ5lEuG0itGxuPsV96BQixqwrFBbBuAaBvGhcd4gPIE0mZk5fhcgCcPyXntIPlUxxTmI73GnQMhbcjp8PJtCso_clGffuSvCwnahmTKi-IsCh_VUXZWCYTl8jCqeKtWbN1VCG419DCRzBneusGNZiqxR6qkxLkin-6egLcMTj3Y05D6mNRvotXzMLN20FS-ybJ7qCwdHte_Fd7d-rLiozjkWNZnmssms9HKsYF43qgyXDM1guUW87EEPai330z6_zG7Z6jJl0bYcQkTW4gcXnj58PAugt9UwIxa8T_d187djJlRv-HD1139001F5RHuNyw09iQwcpXnKh95d1kwhlurY5Xh7H95ebm365v2pfa5N13mqwdhkdAKReaCBFVGm9qCmaaRgyUvzbaeCEBf6s0Adwz04500Ikdy76ZjSiUwE7v_QkUqFEECj82RuEYheUcb1nZovXVW7rZ7MAqr6EbX5uY5T1234LxjcupBgjSOW7Ij3HhNGMKZk270414jgIsVvocC2_H6KwQilc150K56Nhm1c2LnmGCLYQe0Hsrt-Wu2op_9y03SFLKvlcrnSowURVBumJqP0Siis_5FKK4q6w0JdqXPeGI7XxNBsZyzcmjkmJiIhnjBxQ6iGhVPgJ5lEuG0itGxuPsV96BQixqwrFBbBuAaBvGhcd4gPIE0mZk5fhcgCcPyXntIPlUxxTmI73GnQMhbcjp8PJtCso_clGffuSvCwnahmTKi-IsCh_VUXZWCYTl8jCqeKtWbN1VCG419DCRzBneusGNZiqxR6qkxLkin-6egLcMTj3Y05D6mNRvotXzMLN20FS-ybJ7qCwdHte_Fd7d-rLiozjkWNZnmssms9HKsYF43qgyXDM1guUW87EEPai330z6_zG7Z6jJl0bYcQkTW4gcXnj58PAugt9UwIxa8T_d187djJlRv-HD1139001F5RHuNyw09iQwcpXnKh95d1kwhlurY5Xh7H95ebm365v2pfa5N13mqwdhkdAKReaCBFVGm9qCmaaRgyUvzbaeCEBf6s0Adwz04500Ikdy76ZjSiUwE7v_QkUqFEECj82RuEYheUcb1nZovXVW7rZ7MAqr6EbX5uY5T1234LxjcupBgjSOW7Ij3HhNGMKZk270414jgIsVvocC2_H6KwQilc150K56Nhm1c2LnmGCLYQe0Hsrt-Wu2op_9y03SFLKvlcrnSowURVBumJqP0Siis_5FKK4q6w0JdqXPeGI7XxNBsZyzcmjkmJiIhnjBxQ6iGhVPgJ5lEuG0itGxuPsV96BQixqwrFBbBuAaBvGhcd4gPIE0mZk5fhcgCcPyXntIPlUxxTmI73GnQMhbcjp8PJtCso_clGffuSvCwnahmTKi-IsCh_VUXZWCYTl8jCqeKtWbN1VCG419DCRzBneusGNZiqxR6qkxLkin-6egLcMTj3Y05D6mNRvotXzMLN20FS-ybJ7qCwdHte_Fd7d-rLiozjkWNZnmssms9HKsYF43qgyXDM1guUW87EEPai330z6_zG7Z6jJl0bYcQkTW4gcXnj58PAugt9UwIxa8T_d187djJlRv-HD1139001F5RHuNyw09iQwcpXnKh95d1kwhlurY5Xh7H95ebm365v2pfa5N13mqwdhkdAKReaCBFVGm9qCmaaRgyUvzbaeCEBf6s0Adwz04500Ikdy76ZjSiUwE7v_QkUqFEECj82RuEYheUcb1nZovXVW7rZ7MAqr6EbX5uY5T1234LxjcupBgjSOW7Ij3HhNGMKZk270414jgIsVvocC2_H6KwQilc150K56Nhm1c2LnmGCLYQe0Hsrt-Wu2op_9y03SFLKvlcrnSowURVBumJqP0Siis_5FKK4q6w0JdqXPeGI7XxNBsZyzcmjkmJiIhnjBxQ6iGhVPgJ5lEuG0itGxuPsV96BQixqwrFBbBuAaBvGhcd4gPIE0mZk5fhcgCcPyXntIPlUxxTmI73GnQMhbcjp8PJtCso_clGffuSvCwnahmTKi-IsCh_VUXZWCYTl8jCqeKtWbN1VCG419DCRzBneusGNZiqxR6qkxLkin-6egLcMTj3Y05D6mNRvotXzMLN20FS-ybJ7qCwdHte_Fd7d-rLiozjkWNZnmssms9HKsYF43qgyXDM1guUW87EEPai330z6_zG7Z6jJl0bYcQkTW4gcXnj58PAugt9UwIxa8T_d187djJlRv-HD1139001F5RHuNyw09iQwcpXnKh95d1kwhlurY5Xh7H95ebm365v2pfa5N13mqwdhkdAKReaCBFVGm9qCmaaRgyUvzbaeCEBf6s0Adwz04500Ikdy76Z
```

存储桶相关信息需要从 PathPrefix 中解析见下图:

[illegible]

- bucket名称: 解析出来的值为 virtualhuman-cos-prod-1251316161。
- region名称: 解析出来的值为 accelerate。
- key(bucket 下的存储路径): 解析出来的值为 customer-pipeline/352/59c6f6cd-8f92-45f9-8352-64ca8fb92eaa/。

⚠ 注意:

解析出来的值为 key 的前半部分，还需要根据上传素材类型来拼接后半部分，如上传的是 pdf 授权书素材，则需要拼接上 idcard/xxx.pdf，如下图所示，具体后半部分值的定义请参考 [定制接口](#) 中的相关参数说明。

腾讯云智能数智人 购买指南 数智人形象介绍 数智人平台操作指引 服务端 API 接入 数智人 aPaas 接口调用方式 2D真人形象、声音定制 API 概述 形象定制 接口调用逻辑示意图 获取上传临时令牌 上传素材到腾讯云 COS 定制接口 进度查询接口 效果确认接口 人脸检测任务创建 人脸检测任务状态查询 声音复刻（基础版） 接口调用逻辑示意图 音质检测任务创建接口 音质检测任务状态查询接口 获取上传临时令牌 上传素材到腾讯云 COS 定制接口 进度查询接口 效果确认接口 声音复刻（极速版）	IdentityWrittenCosUrl	string	否	pdf 格式授权书 URL，要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名 <code>[customer-pipeline/{数字}]/[uuid]/idcard/b.pdf</code> - 该格式主要适用于“书面版授权书”，以完整清晰的扫描件形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityCosUrl 二选一或二者都填	本页目录： 调用协议 请求参数 返回参数 请求示例 返回示例
	MaterialCosUrl	string	否	形象定制素材 url 地址要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源url地址后增加 video 路径，例如 <code>[customer-pipeline/{数字}]/[uuid]/video/c.mp4</code> - 视频大小，不超过5GB，4K不超过10G - 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 - 视频分辨率：1080P、4K（3840*2160），如定制高精版，则必须为4K - 视频长宽比：16:9 (9:16) - 视频帧率：不低于25fps，不超过60fps - 视频格式：mp4, mov 备注：形象定制时 video 路径必传 声音定制素材 url 地址要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如 <code>[customer-pipeline/{数字}]/[uuid]/audio/c.zip</code> - 压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色；压缩时不要新建文件夹，直接选中所有 wav 压缩 - 对于单个 zip 包里的音频注意事项： ①音频数量：每个 zip 包里可以包含1条或多条 wav 格式音频，总量不超过10条 ②音频大小：每个 zip 包里的音频大小总和不超过1G ③音频格式：每条音频需为 wav 格式，其他格式的音频请先转码为 wav 格式后再压缩成 zip 包 ④音频采样率：采样率大于等于24k即可，建议24k、36k	

提交音质检测任务

最近更新时间：2024-11-29 11:26:32

朗读 [获取声音训练文本](#) 内容，录制好音频文件上传到指定 cos 路径后，需要进行音频质量检测，检测通过的音频会返回 Audioid，使用该 Audioid 调用定制接口进行一句话声音定制。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/zeroshotservice/detectaudioquality
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
TextId	string	否	获取声音训练文本 接口返回的 TextId
ReferenceText	string	否	音频参考文本，ReferenceText 和 TextId 为两种不同方式获取声音训练文本，可以都不填或二选一，能获取更加准确音质检测结果
AudioUrl	string	是	语音文件 url 地址，要求： 1. url 地址为通过4.1上传到指定路径的资源 url 地址后增加 zeroshot 路径，例如 /customer-pipline/{数字}/{uuid}/zeroshot/xxx.mp3 2. 建议上传音频时长为10-30S，单个文件不超过20M 3. 音频格式支持 wav、mp3、m4a、wma、asf，采样率需大于等于16K；对于压缩格式，建议码率大于等于128kbps 4. 音频命名为2-50个字符，只允许有汉字、字母、数字、下划线、横线

返回参数

参数	类型	必须	说明
TaskId	string	是	制作的任务 ID，携带 TaskId 访问 查询音质检测任务进度 ，可获得检测进度和检测结果

请求示例

```
{
  "Header": {},
  "Payload": {
    "AudioUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipline/1322/c9d435f3-9dfb-42c8-9b39-c0117d3cd35c/zeroshot/90450181730994661.mp3"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "DialogID": "",
    "Message": "ok",
    "RequestID": "c9d435f3-9dfb-42c8-9b39-c0117d3cd35c"
  },
  "Payload": {
    "TaskId": "6da6f935-de23-4337-b753-64cca3627362"
  }
}
```

```
}  
}
```

查询音质检测任务进度

最近更新时间：2024-11-13 15:22:12

通过 TaskId 来查询音质检测的进度和结果。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/zeroshotservice/getdetectprogress
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
TaskId	string	是	提交音质检测任务 返回的 TaskId

返回参数

参数	类型	必须	说明
AudioId	string	是	音频ID
DetectionCode	int	是	检测code： 0：表示当前语音通过 -1：表示检测失败，需要重试 -2：表示语音检测不通过，提示用户再重新录制一下（通常漏读，错读，或多读） -3：表示语音中噪声较大，不通过
DetectionMsg	string	否	检测提示信息
TaskState	string	否	检测任务状态： PROCESSING：处理中 SUCCESS：任务处理完成 FAIL：任务处理失败

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "c9d435f3-9dfb-42c8-9b39-c0117d3cd35c"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "gz1a74c25f17030396653416949",
    "SessionID": "gz1a74c25f17030396653416950",
    "DialogID": "",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "TaskState": "SUCCESS",
    "AudioId": "17xsjw8-9dfb-42c8-9b39-sjdue972hds",
  }
}
```

```
"DetectionCode": 0,  
"DetectionMsg": "success"  
}  
}
```

定制接口

最近更新时间：2024-11-29 11:26:32

使 该接 提交定制的请求，通过 [进度查询接口](#) 查询定制的各个阶段以及各阶段的相关信息。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/customservice/make
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
AnchorName	string	是	主播名称： 1. 该名称主要用于标识该定制形象/声音，可以根据实际情况自定义 2. 命名参考：如果该主播定制的数量只有1个，可以直接以主播名称命名，例如“张三”；也可加上衣服名称来更好的标识，例如“张三蓝西服”等 3. 不超过50个字符，不少于2个字符，只允许有汉字、字母、数字、下划线、横线 4. 不能重名
MakeType	string	是	定制类别： IMAGE：2D小样本-标准版（专属口型）形象定制 IMAGE_GENERAL：2D小样本-标准版（通用口型）形象定制 IMAGE_4K：2D小样本-高精版（专属口型）形象定制 IMAGE_PHOTO：2D小样本（照片数智人）形象定制 ⚠ 注意：选择 IMAGE_PHOTO 类型时，强烈建议预先执行 人脸检测任务创建 以提高成功率 VOICE：声音复刻（基础版） ZERO_SHOT_VOICE：声音复刻（极速版）
IdentityCosUrl	string	否	视频格式授权书 URL，要求： 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/a.mp4 2. 该格式主要适用于“口述版授权书”，也可将“ 书面版授权书 ”拍成完整清晰的视频形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityWrittenCosUrl 二选一或二者都填
IdentityWrittenCosUrl	string	否	pdf 格式授权书 URL，要求： 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/b.pdf 2. 该格式主要适用于“ 书面版授权书 ”，以完整清晰的扫描件形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityCosUrl 二选一或二者都填
MaterialCosUrl	string	否	形象定制素材 url 地址要求： 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源url地址后增加 video 路径，例如/customer-pipeline/{数字}/{uuid}/video/c.mp4 2. 视频大小，不超过5GB，4K不超过10G 3. 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 4. 视频分辨率：1080P、4K（3840*2160），如定制高精版，则必须为4K 5. 视频长宽比：16:9 (9:16) 6. 视频帧率：不低于25fps，不超过60fps 7. 视频格式：mp4, mov

			声音定制素材 url 地址要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如/customer-pipeline/{数字}/{uuid}/audio/c.zip - 压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色；压缩时不要新建文件夹，直接选中所有 wav 压缩 - 对于单个 zip 包里的音频注意事项： ①音频数量：每个 zip 包里可以包含1条或多条 wav 格式音频，总量不超过10条 ②音频大小：每个 zip 包里的音频大小总和不超过1G ③音频格式：每条音频需为 wav 格式，其他格式的音频请先转码为 wav 格式后再压缩成 zip 包 ④音频采样率：采样率大于等于24k即可，建议24k、36k ⑤音频命名：名称不要有空格或者特殊字符，后缀名需要为小写 “.wav” 备注：除 ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型必填 照片数智人定制素材 url 地址要求： - url地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 photo 路径，例如/customer-pipeline/{数字}/{uuid}/photo/example.png - 图片名称：不少于2个字符，只允许有汉字、字母、数字、下划、横线；图片格式：支持jpg、jpeg、png、webp；图片大小：不超过16M；图片宽高比：支持 1:1 9:16 16:9 4:3 - 人物清晰正面照，脸部居中，表情自然，嘴部闭合
SexType	string	是	性别： MALE：男 FEMALE：女
Notes	string	否	定制备注信息，100字以内
TextDriver	string	否	用于生成驱动 demo 的文本内容，字符长度在4-1000之间（包括 ssml 标签,单个中文算一个字符）
VoiceDriverCosFile	string	否	生成驱动 demo 的音频文件路径，要求： - url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如:/customer-pipeline/{数字}/{uuid}/audio/example.wav - 音频文件大小不超过10M，格式支持wav、mp3、wma、m4a、aac
Audiold	string	否	ZERO_SHOT_VOICE 定制类型必填:填写 查询音质检测任务进度 通过后返回的 Audiold
TimbreInfo	TimbreTemplate	否	形象定制时若需要指定音色，或尝试用视频里的声音快速复刻，则填充该字段。
IsHaveBackground	bool	否	形象定制类型，训练出来的形象是否带原始背景。默认“否”，即不带原始背景，在应用过程中可以按需更换背景 true：保留拍摄背景 false：抠除拍摄背景（支持更换背景，仅支持绿幕视频和白墙视频）
RemoveBackground	int	否	当 IsHaveBackground=false 时，进一步填充该参数。默认值是1。 1：抠除拍摄背景-绿幕视频 2：抠除拍摄背景-白幕视频
AlphaCosUrl	string	否	自定义 Alpha 通道视频地址，视频上传路径参照 形象定制素材 路径
ZeroshortLang	string	否	ZERO_SHOT_VOICE 定制类型的语言类型，chi-中文普通话，eng-英文，can-粤语，默认chi
PhotoVersion	int	否	默认0 0：快速版。在10分钟内即可完成数智人定制，只支持头部轻微晃动，面部特征稳定。建议执行 人脸检测任务 (CheckerVersion=1) 1：优享版。在1个小时内可完成数智人定制，头部和身体都支持轻微晃动。建议执行 人脸检测任务 (CheckerVersion=2)。

TimbreTemplate 信息参数

参数	类型	必须	说明
TimbreCode	string	否	制作文本驱动 demo 时指定的音色 code
Speed	string	否	制作文本驱动 demo 时指定音色的语速,默认1.0。语速（范围[0.5–1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快）
IsWithTimbre	bool	否	尝试用视频里的声音快速复刻，额外生成 demo（由于视频里的音质质量不能保证，可能会存在快速复刻失败的情况）。该 demo 不影响其他 demo 的正常合成。

返回参数

参数	类型	必须	说明
TaskId	string	是	制作的任务 ID，携带 TaskId 访问 进度查询接口 ，可获得制作进度和制作结果

请求示例

```
{
  "Header": {},
  "Payload": {
    "AnchorName": "马老师",
    "MakeType": "ZERO_SHOT_VOICE",
    "AudioId": "audiof1d85a5c766e47c690ff61cbf43d5364",
    "IdentityCosUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/f30e4d52-00ba-4e36-8a26-26c348bf786e/idcard/identity.mp4",
    "MaterialCosUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/78675441-08bf-43f2-93db-f3936068bce7/video/test.mp4",
    "SexType": "FEMALE"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "c18aeee4e49d89e408231c5b3ceea23c",
    "SessionID": "gzfa9105ca17310669340273673",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "TaskId": "1212_29934_w"
  }
}
```

进度查询接口

最近更新时间：2025-03-07 12:14:53

通过TaskId来查询任务的制作进度和结果。

调用协议

HTTPS + JSON
POST /v2/ivh/assetmanager/customservice/getprogress
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
TaskId	string	是	定制接口 返回的 TaskId

返回参数

参数	类型	必须	说明
Status	string	是	任务制作状态： MAKING：进行中 SUCCESS：制作成功 FAIL：制作失败
FailMessage	string	否	当 Status: FAIL 时，返回制作失败原因，便于排查问题
Stage	string	否	当前定制阶段（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageId	string	否	定制阶段 Id（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageV2	string	否	（自2024.1.1之后发起的进度查询请求，以当前字段为准） 当 Status: MAKING 时，返回任务进行中的具体阶段： SUBMIT: 素材提交 AUDIT: 人工审核 PROCESS: 预处理（仅形象定制有） TRAIN: 智能训练 CONFIRM: 效果确认（照片定制无此阶段） END: 完成
StatusV2	string	否	当 StageV2 处于某个阶段时，该阶段的具体状态： WAIT_PROCESS: 待处理 PROCESSING: 处理中 SUCCESS: 成功 FAIL: 失败
StageInfo	string	否	处于某个阶段的额外信息，JSON 格式的字符串，不同阶段的 JSON 格式类型见下述表格，目前只有效果确认阶段（StageV2 为 CONFIRM）有值。
AssetList	Array of [AssetInfo]	否	返回定制的形象或音色信息，目前支持返回的定制类型有： IMAGE: 2D小样本-标准版（专属口型）形象定制 IMAGE_GENERAL: 2D小样本-标准版（通用口型）形象定制 IMAGE_4K: 2D小样本-高精版（专属口型）形象定制 IMAGE_PHOTO: 2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE: 声音复刻（极速版） VOICE: 声音复刻（基础版）

不同阶段 StageInfo 额外信息见下表：

StageV2	类型	描述	示例
CONFIRM: 效果确认阶段	string of [Customer ConfirmCo ntent]	当阶段处于效果确认阶段（StageV2: CONFIRM），且状态处于待处理中时（StatusV2: WAIT_PROCESS），返回待确认效果 Demo 视频，客户可以根据返回的 Demo 视频通过调用 效果确认接口 对 Demo 视频效果进行认同/驳回操作	<pre>{\\"TrainResult\\": [\\\"url1\\\",\\\"url2\\\"],\\"RejectReason\\":\\\"\\\",\\"TrainRecord\\":[{\\"StartTime\\":\\"2023-11-01 23:00:17\\\",\\"Record\\":[\\\"url1\\\",\\\"url2\\\"]}]}</pre>

CustomerConfirmContent

参数	类型	必须	说明
TrainResult	Array of [string]	是	供客户进行效果确认的 Demo 视频 url 地址
RejectReason	string	是	数智人侧不接受重训的理由
TrainRecord	Array of [TrainRecord]	否	历史训练的 Demo 视频清单

TrainRecord

参数	类型	必须	说明
StartTime	string	是	历史训练时间
Record	Array of [string]	是	历史训练的 Demo 视频 url 地址

AssetInfo

参数	类型	必须	说明
VirtualmanTypeCode	string	否	数智人类型 Code，参考 附录3-数智人类型 ，形象定制返回
VirtualmanKey	string	否	数智人 virtualmankey，形象定制返回
Resolution	string	否	数智人分辨率，形象定制返回
OriginZoom	string	否	缩放系数，保留三位小数，形象定制返回
VirtualmanResourceId	int32	否	数智人资源 ID，形象定制返回
TimbreKey	string	否	音色key，声音复刻返回
TimbreSample	string	否	音色demo链接，声音复刻返回
MakeType	string	是	定制类型： IMAGE_PHOTO：2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE：声音复刻（极速版）

请求示例

```
{  
  "Header": {},  
}
```

```
"Payload": {  
  "TaskId": "1212_29934_w"  
}  
}
```

返回示例

```
{  
  "Header": {  
    "RequestID": "gz1a74c25f17030396653416949",  
    "SessionID": "gz1a74c25f17030396653416950",  
    "DialogID": "",  
    "Code": 0,  
    "Message": "ok"  
  },  
  "Payload": {  
    "StageInfo": "{\\\"TrainResult\\\":[\\\"url1\\\",\\\"url2\\\"],\\\"RejectReason\\\":\\\"\\\",\\\"TrainRecord\\\":  
[\\\"StartTime\\\":\\\"2023-11-01 23:00:17\\\",\\\"Record\\\":[\\\"url1\\\",\\\"url2\\\"]]}\"",  
    "Status": "MAKING",  
    "StatusV2": "WAIT_PROCESS",  
    "FailMessage": "",  
    "StageV2": "CONFIRM"  
  }  
}
```

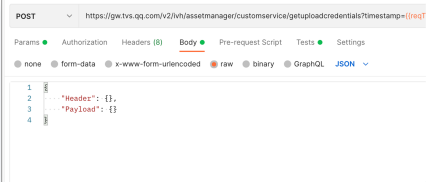
结果码说明

最近更新时间：2024-08-23 10:39:03

代码	含义	措施
0	成功	无
100000	系统内部错误	重试或联系管理员定位错误
100001/100003	请求参数错误（包括参数格式、类型等错误，缺少参数错误，必传参数没填）	对照 API 协议核查参数格式、类型
100002	参数取值错误	对照 API 协议核查参数列表
100005	未授权操作	检查对请求的会话是否有操作权限
110006	记录不存在	检查 taskid 是否正确
100008	超过配额限制	分配的配额已经用完

接口对接常见问题

最近更新时间：2024-08-23 17:05:43

序号	问题描述	解决方式	备注
1	调用接口时出现鉴权不通过	1. 确认签名时参数按照字典序排列了 2. 确认时间戳参数为当前时间 3. 参考右侧的示例参数，按照自己的签名步骤计算一下，如果计算结果与示例一致则表示签名计算步骤无误。	参数： appkey = e38267c0e86411ebb02aed82acb0ed99 accesstoken = f68f2d10ae9e4604b76fb05cf46bccec timestamp=1646636485 签名结果： BfWuaC9kmaicCggXc693uK%2BsZQ8q e88O4HVQNTdwZuo%3D
2	调用接口出现"miss Content Length"错误	1. 请确认请求的 body 是否加了 header 参数和 payload 参数，不加的话会报错。如右图所示。 2. 参数的类型要严格按照接口文档中填充。 3. 检查 url 是否拼写正确	
3	UnauthorizedOperation: 未授权操作: 未给该 appkey 创建项目	未给客户开通小样本形象定制接口的使用权限，请联系腾讯相关人员开通	
4	视频素材是否需要上传到客户的腾讯云 COS 上	不需要。通过临时令牌上传会上传到数智人内部的 COS 上，不需要客户额外开通腾讯云 COS	
5	调用 cos 的 sdk 时的 Bucket、Region、Key 如何获取	可以从 prefix 中解析出来，如右图所示	

视频生成服务 API 文档

视频生成服务 API 概述

最近更新时间：2024-11-13 14:38:22

本文档主要描述 腾讯云智能数智人 视频生成aPaas 平台的开放接口协议。

一、接口调用环境

正式环境：<https://gw.tvs.qq.com>

二、权限申请及开通

视频生成 aPaas 接口使用权限可以通过 [资产领用](#) 页面领用，领用后即可开通15天试用权限。

三、接口调用方式

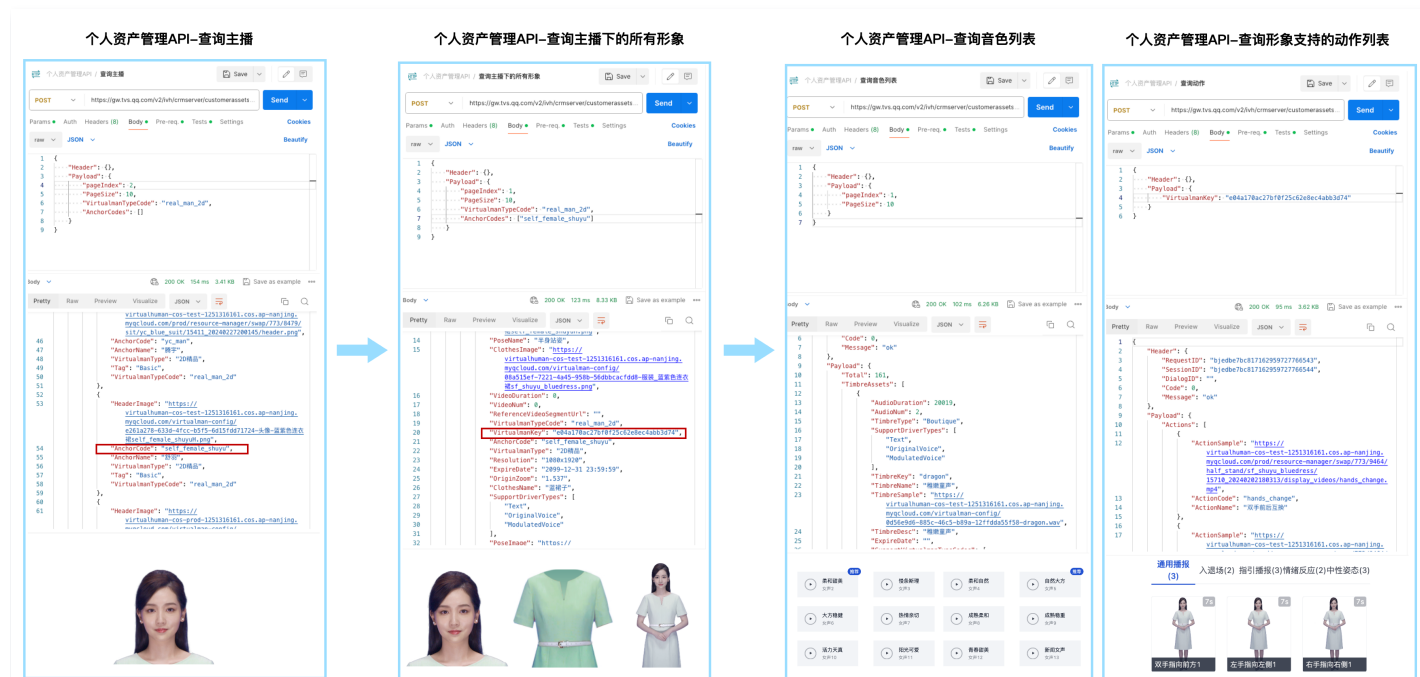
数智人 aPaas 接口调用方式

⚠ 注意：

如果通过腾讯云计费直接采购。appkey 将只能在选购的形象包内切换形象，不同形象包不通用。

四、接入流程示意图

资源查询逻辑示意图



音频制作接口

最近更新时间：2024-11-29 17:11:52

接口说明

对输入文本进行试听，需要试听的音色可以通过接口 [查询声音资产信息](#) 查询。

调用协议

HTTPS + JSON
POST /v2/ivh/videomaker/broadcastservice/tts
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
TimbreKey	string	否	音色key，当VirtualmanKey为空时TimbreKey不能为空
VirtualmanKey	string	否	定义播报的角色、服装、姿态、分辨率等信息，参数为枚举值，当TimbreKey为空时VirtualmanKey不能为空，默认选取匹配该形象的第一个音色制作音频。
InputSsml	string	是	需要播报的文本内容，支持ssml标签，上限2万字（字数按unicode字符数计算）
Speed	float	是	语速（1.0为正常语速，范围[0.5-1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快）
AudioStorageS3Url	string	否	可传入含鉴权s3协议存储url，音频成品会上传至该url
SampleRate	int	否	采样率，支持24000（24k）和16000（16k），默认24000(24k)
Codec	string	否	音频格式，支持mp3和wav，默认mp3
SentenceMaxWords	int	否	分句的字数上限，字数范围 [0,999]，传0或不传时默认值30
SentenceDisplayPunctuation	string	否	分句里需要展示的标点符号，其中，传入特殊字符"0"表示标点符号都不展示，特殊符号"1"（默认取值）表示标点符号全都展示，也可自定义传入需要展示的标点符号
SentenceSplitPunctuation	string	否	需要断句的标点符号，默认取值。；？！……！？
Volume	int	否	音量大小，范围[0, 10]，对应音量大小。默认为0，代表正常音量，值越大音量越高。 ❗ 说明： TimbreKey 在 male_1-20、female_1-23（即男声1-20、女声1-23）间的音色不支持音量调节。
EmotionCategory	string	否	控制合成音频的情感，仅支持多情感音色使用，可选值参考个人资产管理API 4.5音色列表接口。
EmotionIntensity	int	否	控制合成音频情感程度，取值范围为 [50,200]，只有 EmotionCategory 不为空时生效。
TimbreLanguage	string	否	音色语种，可选语种参考个人资产管理 API 分页查询音色列表 接口，多语种音色在合成时必须选择对应语种。

返回参数

参数	类型	必须	说明
TaskId	string	是	音频制作的任务 ID，携带 TaskId 访问<音视频制作进度查询接口>，可获得该视频的制作进度和下载地址。

请求示例

```
{
  "Header": {},
  "Payload": {
    "VirtualmanKey": "k3dda2d3b56f4ec89778f2e032c2e294",
    "InputSsm1": "你好虚拟主播",
    "Speed": 1
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "DialogID": "",
    "Message": "",
    "RequestID": "fde854eaa981c7f2f7285d1c7eca335b",
    "SessionID": "gzb7dec22117297528294581119"
  },
  "Payload": {
    "TaskId": "81883d47c6154edf8e276531f09227b6"
  }
}
```

视频制作接口-基础版

最近更新时间：2025-06-09 21:04:41

接口说明

使用 ssml 文本和数智人进行视频制作，通过 [音视频制作进度查询接口](#) 最终返回成品视频和字幕文件。

 **注意：**
不支持定义主播位置等未剪辑高级参数，如需使用需要切换到 [视频制作接口-进阶版](#) 接口。

调用协议

HTTPS + JSON
POST /v2/ivh/videomaker/broadcastservice/videomake
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
VirtualmanKey	string	是	定义播报的角色、服装、姿态、分辨率等信息，参数为枚举值。  说明： 可以通过 查询形象资产信息-查询主播下的所有形象 接口查询
InputSsml	string	是	播报的文本内容，支持ssml标签，支持的标签类型参照 数智人 SSML 标记语言规范 ，标签写法参照示例，内容不能换行，符号需转义。上限2万字（字数按 unicode 字符数计算）。DriverType 为空、或 Text 时，该字段必填  说明： 当 Format 为 TransparentMov 时，上限为2千字。
SpeechParam	object	是	定义音频的详细参数
SpeechParam.Speed	float	是	语速（1.0为正常语速，范围[0.5-1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快，DriverType为音频驱动类型时，语速控制不生效）
SpeechParam.TimbreKey	string	否	音色key，默认使用形象自有音色
SpeechParam.Volume	int	否	音量大小，范围[0，10]，对应音量大小。默认为0，代表正常音量，值越大音量越高。  说明： TimbreKey 在 male_1-20、female_1-23（即男声1-20、女声1-23）间的音色不支持音量调节。
SpeechParam.EmotionCategory	string	否	控制合成音频的情感，仅支持多情感音色使用，可选值参考个人资产管理 API 分页查询音色列表 接口。
SpeechParam.EmotionIntensity	int	否	控制合成音频情感程度，取值范围为 [50,200]，只有 EmotionCategory 不为空时生效。
SpeechParam.TimbreLanguage	string	否	音色语种，可选语种参考个人资产管理 API 分页查询音色列表 接口，多语种音色在合成时必须选择对应语种。
VideoParam	object	否	定义合成视频的详细参数，

VideoParam.Format	string	否	视频输出格式，默认 TransparentWebm TransparentWebm：带 Alpha 通道的 webm 格式视频 TransparentMov：带 Alpha 通道的 mov 格式视频 Mp4：绿幕 mp4 格式视频；非抠像形象为原始素材背景
CallbackUrl	string	否	当用户增加回调 url 时，将把视频制作结果以固定格式发送 post 请求到该 url 地址，固定格式见 附录二：回调请求体格式 ，需注意： 1、限制 CallbackUrl 长度小于1000 2、只发送一次请求，无论是哪种问题导致的请求失败，都不会再进行发送。
DriverType	string	否	驱动类型，默认Text 1. Text：文本驱动，要求 InputSsmI 字段必填 2. OriginalVoice：原声音频驱动，要求 InputAudioUrl 字段必填 3. ModulatedVoice：变声音频驱动，可通过 Speech.TimbreKey 指定音色，未填写时使用主播默认音色
InputAudioUrl	string	否	驱动数智人的音频 url，当 DriverType 为 OriginalVoice、ModulatedVoice 时，该字段必填。 音频格式要求： 1. 小样本形象时长不超过60分钟，不少于0.5秒；非小样本形象时长不超过10分钟，不少于0.5秒 2. 支持格式：wav、mp3、wma、m4a、aac 3. 建议文件大小：不大于100M 说明： 当 Format 为 TransparentMov 时，时长不超过6分钟，不少于0.5秒；
VideoStorageS3Url	string	否	可传入含鉴权 s3 协议存储 url，视频成品会上传至该 url
SubtitleStorageS3Url	string	否	可传入含鉴权 s3 协议存储 url，字幕成品会上传至该 url
ConcurrencyType	string	否	视频制作任务使用的并发类型，默认采用优先使用独享，其次使用共享的策略 1. Exclusive：独享并发，若无独享并发，任务提交失败 2. Shared: 共享并发

返回参数

参数	类型	必须	说明
TaskId	string	是	视频制作的任务 ID，携带 TaskId 访问 音视频制作进度查询接口 ，可获得制作进度和制作结果

请求示例

文本驱动

```
{
  "Header": {},
  "Payload": {
    "VirtualmanKey": "k3dda2d3b56f4ec89778f2e032c2e294",
    "InputSsmI": "你好，我是虚拟<phoneme alphabet=\"py\" ph=\"fu4\">主</phoneme>播",
    "SpeechParam": {
      "Speed": 1.0
    },
    "VideoParam": {
      "Format": "TransparentWebm"
    },
    "DriverType": "Text"
  }
}
```

```

    }
  }
}

```

音频驱动

```

{
  "Header": {},
  "Payload": {
    "VirtualmanKey": "k3dda2d3b56f4ec89778f2e032c2e294",
    "InputAudioUrl": "https://videomaker-
resources.ivh.qq.com/5da17352aa1e4cba9cada7bd20571295.mp3",
    "SpeechParam": {
      "Speed": 1.0
    },
    "VideoParam": {
      "Format": "TransparentWebm"
    },
    "DriverType": "OriginalVoice"
  }
}

```

返回示例

```

{
  "Header": {
    "Code": 0,
    "DialogID": "",
    "Message": "",
    "RequestID": "fde854eaa981c7f2f7285d1c7eca335b",
    "SessionID": "gzb7dec22117297528294581119"
  },
  "Payload": {
    "TaskId": "81883d47c6154edf8e276531f09227b6"
  }
}

```

视频制作接口-进阶版

最近更新时间：2025-06-09 21:04:41

接口说明

使用 ssmI 文本和数智人进行视频制作，通过 [音视频制作进度查询接口](#) 最终返回成品视频和字幕文件。在原有接口基础上、进阶版接口新增了部分资源参数，扩展了微剪辑能力，支持能力如下表：

LOGO	支持调整 logo 位置（自定义X轴Y轴）
	支持调整 logo 大小（缩放比例）
指定视频背景	在MP4视频文件中，可以指定背景图。
主播	横向位置调节
	主播大小调整
	主播角度调整
	⚠ 注意： 仅支持3D形象
片头/片尾视频	可指定远程地址添加视频片头片尾
内嵌字幕	打开、关闭内嵌字幕

调用协议

HTTPS + JSON
POST /v2/ivh/videomaker/broadcastservice/videomakeadvanced
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
VirtualmanKey	string	是	定义播报的角色、服装、姿态、分辨率等信息，参数为枚举值。 ⓘ 说明： 可以通过 查询形象资产信息-查询主播下的所有形象 接口查询
InputSsmI	string	否	播报的文本内容，支持ssmI标签，支持的标签类型参照 数智人 SSML 标记语言规范 ，标签写法参照示例，内容不能换行，符号需转义。上限2万字（字数按unicode字符数计算）。DriverType为空、或Text时，该字段必填 ⓘ 说明： 当 Format 为 TransparentMov 时，上限为2千字。
SpeechParam	object	是	定义音频的详细参数
SpeechParam.Speed	float	是	语速（1.0为正常语速，范围[0.5-1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快，DriverType为音频驱动类型时，语速控制不生效）
SpeechParam.TimbreKey	string	否	音色key，默认使用形象自有音色

SpeechParam.Volume	int	否	音量大小，范围[0, 10]，对应音量大小。默认为0，代表正常音量，值越大音量越高。 说明： TimbreKey 在 male_1-20、female_1-23（即男声1-20、女声1-23）间的音色不支持音量调节。
SpeechParam.EmotionCategory	string	否	控制合成音频的情感，仅支持多情感音色使用，可选值参考个人资产管理API 分页查询音色列表 接口。
SpeechParam.EmotionIntensity	int	否	控制合成音频情感程度，取值范围为 [50,200]，只有 EmotionCategory 不为空时生效。
SpeechParam.TimbreLanguage	string	否	音色语种，可选语种参考个人资产管理 API 分页查询音色列表 接口，多语种音色在合成时必须选择对应语种。
VideoParam	object	否	定义合成视频的详细参数，
VideoParam.Format	string	否	视频输出格式，默认值：Mp4 TransparentWebm: 带 Alpha 通道的 webm 格式视频，支持部分微剪辑能力（主播参数支持） TransparentMov: 带 Alpha 通道的 mov 格式视频，支持部分微剪辑能力（主播参数支持） Mp4: 支持微剪辑能力的 mp4 格式视频，抠像形象默认为绿幕背景，可通过 BackgroundFileUrl 参数设置背景；非抠像形象为原始素材背景。
VideoParam.BackgroundImageUrl	string	否	视频背景图片/视频下载路径，支持jpg、png、mp4格式，图片/视频分辨率需要与视频分辨率相同，不传默认为绿幕视频。文件大小限制500M。
VideoParam.VideoHeadFileUrl	string	否	片头视频，支持mp4格式，分辨率需要与视频分辨率相同，文件大小限制500M
VideoParam.VideoTailFileUrl	string	否	片尾视频，支持mp4格式，分辨率需要与视频分辨率相同，文件大小限制500M
VideoParam.ShowSubtitles	boolean	否	是否在视频内显示字幕，默认不展示，展示会大大增加视频制作时间
VideoParam.SubtitlesParam	object	否	定义视频中字幕如何显示的相关参数
VideoParam.SubtitlesParam.MaxWords	int	否	字幕每页展示的字数上限，字数范围 [0,999]；默认值为0，默认展示规则：在分辨率80%宽度范围内展示，超出时分页。
VideoParam.SubtitlesParam.DisplayPunctuation	string	否	字幕需要展示的标点符号，其中，传入特殊字符"0"表示标点符号都不展示，特殊符号"1"（默认取值）表示标点符号全都展示，也可自定义传入需要展示的标点符号
VideoParam.SubtitlesParam.SplitPunctuation	string	否	字幕需要分页的标点符号，默认取值：；？！………！？
VideoParam.LogoParams	Array of [LogoParam]	否	定义视频中logo相关参数
VideoParam.SmartActionEnabled	bool	否	是否开启智能动作，默认不开启。生效条件：DriverType=Text、且InputSsml中不携带动作标签
VideoParam.AnchorParam	object	否	定义视频中主播相关参数

VideoParam.AnchorParam.HorizontalPosition	float	否	定义主播横向位置（0为中间位置）
VideoParam.AnchorParam.VerticalPosition	float	否	定义主播纵向位置（0为中间位置）
VideoParam.AnchorParam.Scale	float	否	主播大小（1是默认大小，范围(0,15]）
VideoParam.AnchorParam.Angle	int	否	主播角度（默认0度，范围[0,360]），目前仅3D形象支持。
VideoParam.AnchorParam.AnchorExtraParam	string	否	主播额外可配置参数，目前仅包含3D服装变色参数。不同主播可配置的参数不同，详情见接口 查询形象资产信息-查询主播下的所有形象 的 SupportAnchorExtraParam 参数；需要将该参数组织为 json 字符串格式，请参考请求示例组织 json
VideoParam.SmallSampleParam	object	否	定义小样本相关特殊参数，当为非小样本形象时，此参数不生效
VideoParam.SmallSampleParam.MakeType	string	否	定义小样本制作类型： Default：默认配置，会从随机起始帧开始制作，StartIdx 和 EndIdx 均不生效 Custom：指定视频片段，可以填写 StartIdx 和 EndIdx 来选择片段，默认生成的视频用该“指定的视频片段”进行正倒放循环 Circle：首尾帧归位，可以填写 StartIdx 来指定视频开始帧数(EndIdx 不生效)，注意：该类型只用于声音驱动
VideoParam.SmallSampleParam.StartIdx	int	否	起始帧号，当MakeType为Custom或Circle时生效，当填写的时候，默认生成的视频从该位置开始；不填的时候，默认选取开始帧号，即第0帧
VideoParam.SmallSampleParam.EndIdx	int	否	结束帧号，当MakeType为Custom时生效，当填写的时候，默认结束的视频到该位置结束；不填的时候，默认选取该“选中的全部视频片段”的结束帧号
CallbackUrl	string	否	当用户增加回调url时，将把视频制作结果以固定格式发送 post 请求到该 url 地址，固定格式见 附录二：回调请求体格式 。 需注意： 1. 限制 CallbackUrl 长度小于1000 2. 只发送一次请求，无论是哪种问题导致的请求失败，都不会再进行发送。
DriverType	string	否	驱动类型，默认Text 1. Text：文本驱动，要求 InputSsmI 字段必填 2. OriginalVoice：原声音频驱动，要求 InputAudioUrl 字段必填 3. ModulatedVoice：变声音频驱动，可通过 Speech.TimbreKey 指定音色，未填写时使用主播默认音色
InputAudioUrl	string	否	驱动数智人的音频 url，当 DriverType 为 OriginalVoice、ModulatedVoice 时，该字段必填。音频格式要求： 1. 小样本形象时长不超过60分钟，不少于0.5秒；非小样本形象时长不超过10分钟，不少于0.5秒 2. 支持格式：wav、mp3、wma、m4a、aac 3. 建议文件大小：不大于100M ① 说明： 当 Format 为 TransparentMov 时，时长不超过6分钟，不少于0.5秒；
VideoStorageS3Url	string	否	可传入含鉴权s3协议存储url，视频成品会上传至该url
SubtitleStorageS3Url	string	否	可传入含鉴权s3协议存储url，字幕成品会上传至该url
ConcurrencyType	string	否	视频制作任务使用的并发类型，默认采用优先使用独享，其次使用共享的策略 1. Exclusive：独享并发，若无独享并发，任务提交失败 2. Shared：共享并发

LogoParam

参数	类型	必须	说明
LogoFileUrl	string	否	logo 图片文件下载路径，支持jpg、png格式
PositionX	int	否	logo 图片左上角的X坐标（坐标范围和视频分辨率相关）
PositionY	int	否	logo 图片左上角的Y坐标（坐标范围和视频分辨率相关）
Scale	float	否	logo 图片缩放比例（1.0为原始图片大小，范围(0,15]）

VideoParam.Anchor.AnchorExtraParam 示例

 **说明：**
colorValue 转换规则：十六进制Hex色值，去掉#后转为十进制，e.g：#FAFAFA =>16448250

```
{
  "shirtColor": {
    "colorValue": 10491928
  },
  "clotheColor": {
    "colorValue": 10491928
  },
  "shoeColor": {
    "colorValue": 10491928
  }
}
```

返回参数

参数	类型	必须	说明
TaskId	string	是	视频制作的任务ID，携带TaskId访问 音视频制作进度查询接口 ，可获得制作进度和制作结果

请求示例

```
{
  "Header": {},
  "Payload": {
    "VirtualmanKey": "k3dda2d3b56f4ec89778f2e032c2e294",
    "InputSsm1": "你好，我是虚拟<phoneme alphabet=\"py\" ph=\"fu4\">主</phoneme>播",
    "SpeechParam": {
      "Speed": 1
    },
    "VideoParam": {
      "Format": "Mp4",
      "BackgroundFileUrl": "http://virtualhuman-cos-test-1251316161.cos.ap-nanjing.myqcloud.com/bg.png",
      "VideoHeadFileUrl": "http://virtualhuman-cos-test-1251316161.cos.ap-nanjing.myqcloud.com/video_head.mp4",
      "VideoTailFileUrl": "http://virtualhuman-cos-test-1251316161.cos.ap-nanjing.myqcloud.com/video_tail.mp4",
      "ShowSubtitles": true,
      "LogoParams": [
        {

```

```

        "LogoFileUrl": "http://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/logo.png",
        "PositionX": 1561,
        "PositionY": 751,
        "Scale": 1.0
    }
],
    "AnchorParam": {
        "HorizontalPosition": 0.5,
        "Scale": 1.0,
        "AnchorExtraParam": "{\\"hair\\":{\\"boolValue\\":true},\\"Jacket\\":
{\\"boolValue\\":true},\\"Shoes\\":{\\"boolValue\\":true},\\"Skirt\\":{\\"boolValue\\":true},\\"hairColor\\":
{\\"colorValue\\":12759296},\\"JacketColor\\":{\\"colorValue\\":7976228},\\"ShoesColor\\":
{\\"colorValue\\":16777215},\\"SkirtColor\\":{\\"colorValue\\":16777215}}}"
    }
}
}
}

```

返回示例

```

{
    "Header": {
        "Code": 0,
        "DialogID": "",
        "Message": "",
        "RequestID": "fde854eaa981c7f2f7285d1c7eca335b",
        "SessionID": "gz7dec22117297528294581119"
    },
    "Payload": {
        "TaskId": "81883d47c6154edf8e276531f09227b6"
    }
}

```

视频制作接口-视频免训练

最近更新时间：2025-06-25 20:02:42

接口说明

无需训练，即可基于一段真人视频素材，通过输入文本或音频，生成人物说话口型与输入内容相匹配的新视频。
通过 [音视频制作进度查询接口](#) 最终返回成品视频。未上传自定义存储 URL 的音视频资源只保留7天，请尽快下载。

调用协议

HTTPS + JSON
POST /v2/ivh/videomaker/broadcastservice/videomakenotrain
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
RefVideoUrl	string	是	模板视频 URL，该字段必填。 1. 支持封装格式：mp4、mov、avi ⚠ 注意： 视频本身要求编码格式为 H.264 格式，若不是，可通过下述方法转换。请参见 编码格式转换。 2. 文件大小：不超过5G。 3. 视频分辨率：单边像素要求在360 – 4096之间。 4. 视频时长：支持1 – 600秒，建议时长在10秒到120秒（当 ConcurrencyType 为 Exclusive 时，支持20分钟视频时长）。 5. 请保证文件的下载速度，否则会下载失败或导致视频制作实时率下降。 对于输入视频素材的规范性要求： 1. 人脸画面：要求真人出镜（如果是卡通人物，需要人物五官和真人比例相近），画面中的人脸说话时，建议正对镜头，水平转动不超过45度，俯仰不超过15度；人脸尽量不遮挡，面部光线稳定。 2. 说话音频：对音频无限制。 对于输出视频的说明： 1. 输出格式：mp4 2. 输出分辨率：最大支持输出单边4096，即 4K 视频。 2.1 当 ConcurrencyType 为 Shared 时，支持最大输出2560*1440。当输入视频分辨率 <=2560*1440时，输出分辨率和输入视频一致。 2.2 当 ConcurrencyType 为 Exclusive 时，支持最大输出4096*4096。当输入视频分辨率 <=4096*4096时，输出分辨率和输入视频一致。（4K 视频是 2K 视频合成实时率的2倍）
DriverType	string	是	驱动类型，该字段必填。 1. Text：文本驱动，要求 InputSsml 字段必填。 2. OriginalVoice：原声音频驱动，要求 InputAudioUrl 字段必填。
IdentityWrittenUrl	string	否	PDF 格式授权书，格式为 PDF，文件大小小于10M。
IdentityVideoUrl	string	否	视频格式授权书，格式为 mp4，文件大小小于5G。

InputAudioUrl	string	否	驱动数智人的音频 URL，当 DriverType 为 OriginalVoice 时，该字段必填。 - 支持格式：wav、mp3、wma、m4a、aac、ogg - 时长不超过10分钟，不少于1秒（当 ConcurrencyType 为 Exclusive 时，支持20分钟音频时长） - 大小要求：不超过100MB - 请保证文件的下载速度，否则会下载失败或导致视频制作实时率下降。
InputSsml	string	否	播报文本内容，支持 SSML 标签，支持的标签类型参照 数智人 SSML 标记语言规范，标签写法参照示例，内容不能换行，符号需转义。上限2000字，不少于4个字（字数按 unicode 字符数计算）。DriverType为空、或 Text 时，该字段必填。
SpeechParam	object	否	定义音频的详细参数。DriverType 为 Text 时，该字段必填。
SpeechParam.Speed	float	否	语速（1.0为正常语速，范围[0.5–1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快，DriverType 为音频驱动类型时，语速控制不生效），DriverType 为 Text 时，该字段必填。
SpeechParam.TimbreKey	string	否	音色 key，DriverType 为Text时，该字段必填。
SpeechParam.Volume	int	否	音量大小，范围[0, 10]，对应音量大小。默认为0，代表正常音量，值越大音量越高。 说明： TimbreKey 在 male_1-20、female_1-23（即男声1-20、女声1-23）间的音色不支持音量调节。
SpeechParam.EmotionCategory	string	否	控制合成音频的情感，仅支持多情感音色使用，可选值参考个人资产管理 API 分页查询音色列表 接口。
SpeechParam.EmotionIntensity	int	否	控制合成音频情感程度，取值范围为 [50,200]，只有 EmotionCategory 不为空时生效。
SpeechParam.TimbreLanguage	string	否	音色语种，可选语种参考个人资产管理 API 分页查询音色列表 接口，多语种音色在合成时必须选择对应语种。
ConcurrencyType	string	否	视频制作任务使用的资源类型。 - Exclusive：使用并发调用，不扣除小时包，需要购买并发，如果没有购买，任务提交失败。 - Shared：调用会扣除小时包，需要购买小时包，如果没有购买，任务提交失败。 - 不填：如果购买了并发或者并发和小时包都购买则为“Exclusive”，如果没有购买并发但购买了小时包则为“Shared”，如果都没有购买，任务提交失败。
VideoLoop	int	否	当音频时长大于视频时长时，生成的视频会对齐音频时长。以下是两种对齐方式： 0：反向拼接、1：正向拼接。默认 0。
VideoParametersConsistent	int	否	视频输出标准（码率与帧率）对齐选项，默认0。 0：不强制对齐输入视频格式。 1：对齐输入视频格式。 参数说明： 一、选择“0”时 - 码率： - 默认采用 CRF=17 的编码方式，编码器会根据视频内容复杂度动态调整输出码率，确保质量稳定。 - 帧率： - 默认输出帧率为 25fps。 二、选择“1”时 - 码率处理规则： - 输入码率 < 1000kbps → 输出码率调整为 1000kbps。 - 输入码率 > 9000kbps → 输出码率调整为 9000kbps。 - 输入码率在 [1000kbps, 9000kbps] 范围内 → 输出码率与输入一致。 - 帧率处理规则：

			○ 输入帧率 < 15fps → 输出帧率调整为 15fps。 ○ 输入帧率 > 60fps → 输出帧率调整为 60fps。 ○ 输入帧率在 [15fps, 60fps] 范围内 → 输出帧率与输入一致。
CallbackUrl	string	否	当用户增加回调 URL 时，将把视频制作结果以固定格式发送 POST 请求到该 URL 地址，固定格式见 附录二：回调请求体格式 ，需注意： 1. 限制 CallbackUrl 长度小于1000。 2. 只发送一次请求，无论是哪种问题导致的请求失败，都不会再进行发送。
VideoParam	object	否	定义合成视频的详细参数。
VideoParam.RefPhotoUrl	string	否	用户上传的人脸参考图 URL。 输入视频中包含多张人脸时，VideoMakeNoTrain 视频生成 API 仅能选择一张人脸作为目标替换其口型，实现口型与音频相匹配的效果。此参数用来指定以哪个人物作为目标。 若不输入人脸参考图，默认将选择视频中第一个有人脸的画面中，人脸占比最大的人物为目标。 图像文件要求： 1. 文件大小：文件≤10MB 2. 图像大小：单边分辨率要求在[192~4096] 3. 格式支持 jpg、jpeg、png、bmp、webp 4. 内容：需包含一张清晰的人物正脸，且为视频中出现的角色。
VideoParam.DisableOcclusionDetect	int	否	遮挡检测开关。默认0。 0：开启遮挡检测。开启后，当嘴部有遮挡时，不会驱动口型。 1：关闭遮挡检测。关闭后，当嘴部有遮挡时，也会驱动口型。
VideoParam.EnableFakeTeeth	int	否	假牙处理开关。默认0。 0：关闭假牙处理。关闭后，最终合成的视频里的牙齿，会学习原视频里的牙齿特征。 1：开启假牙处理。开启后，最终合成的视频里的牙齿会重新生成，不会参考原视频的牙齿特征。
VideoParam.MakeType	string	否	视频定制类型。 • Default：默认配置，视频片段默认会从原视频第0帧开始，StartTime 和 EndTime 均不生效 • Custom：指定视频片段，可以填写 StartTime 和 EndTime 来选择片段（需要大于5s），默认生成的视频用该“指定的视频片段”进行正倒放循环。 • Circle：首尾帧归位，可以填写 StartTime 来指定视频开始时间（EndTime 不生效），并且此时 VideoLoop 参数不生效。
VideoParam.StartTime	float	否	起始时间，单位秒（有效数字为小数点后3位），当 MakeType 为 Custom 或 Circle 时生效，当填写的时候，默认生成的视频从该位置开始；不填的时候，默认选取视频开始时间。
VideoParam.EndTime	float	否	结束时间，单位秒（有效数字为小数点后3位），当 MakeType 为 Custom 时生效，当填写的时候，默认结束的视频到该位置结束；不填的时候，默认选取该“选中的全部视频片段”的结束时间
VideoParam.DisableIntervals	Array of [DisableInterval]	否	定义视频中不驱动时间片段列表（目前最大支持5个时间片段），列表需按照时间顺序递增，否则任务提交失败。 例如：[[1,2],[3,4],[5,6]] 表明原视频的1s~2s、3s~4s、5s~6s不做口型驱动。

DisableInterval

参数	类型	必须	说明
StartTime	float	否	不驱动片段起始时间，单位秒，（有效数字为小数点后3位）
EndTime	float	否	不驱动片段结束时间，单位秒，（有效数字为小数点后3位）

返回参数

参数	类型	必须	说明
----	----	----	----

TaskId	string	是	视频制作的任务 ID，携带 TaskId 访问 音视频制作进度查询接口 ，可获得制作进度和制作结果
--------	--------	---	---

请求示例

文本驱动

```
{
  "Header": {},
  "Payload": {
    "RefVideoUrl": "http://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/ref_video.mp4",
    "DriverType": "Text",
    "InputSsm1": "你好，我是虚拟<phoneme alphabet=\\"py\\" ph=\\"fu4\\">主</phoneme>播",
    "SpeechParam": {
      "TimbreKey": "female_1",
      "Volume": 1,
      "Speed": 1.0
    }
  }
}
```

音频驱动

```
{
  "Header": {},
  "Payload": {
    "RefVideoUrl": "http://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/ref_video.mp4",
    "DriverType": "OriginalVoice",
    "InputAudioUrl": "http://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/audio.mp3"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "DialogID": "",
    "Message": "",
    "RequestID": "fde854eaa981c7f2f7285d1c7eca335b",
    "SessionID": "gzb7dec22117297528294581119"
  },
  "Payload": {
    "TaskId": "81883d47c6154edf8e276531f09227b6"
  }
}
```

编码格式转换小教程

该教程专用于将视频转成 H.264 编码格式，主要原理为：通过 FFmpeg 实现转换。

第一步：安装 FFmpeg

- Windows 系统：

从 FFmpeg 官网下载 Windows 版本的压缩包。

- macOS 系统:

可以使用 Homebrew 进行安装, 命令为 `brew install ffmpeg`。

- Linux 系统:

- Ubuntu/Debian 系统使用命令 `sudo apt-get install ffmpeg`。
- Fedora 系统使用命令 `sudo dnf install ffmpeg`。

第二步：转换

1. 准备好两个路径：转换前视频路径和转换后视频路径，然后在“终端”中输入以下指令。

```
ffmpeg -i 转换前视频路径 -c:v libx264 -preset medium -crf 23 -c:a aac -b:a 128k 转换后视频路径
```

例如：

- 转换前视频路径为：“/Users/xxxx/Downloads/picture/video/视频名.mov”
- 转换后视频路径为：“/Users/xxxx/Downloads/picture/video/转换后视频名.mp4”

2. 则终端输入指令为：

```
ffmpeg -i "/Users/xxxx/Downloads/picture/video/视频名.mov" -c:v libx264 -preset medium -crf 23 -c:a aac -b:a 128k "/Users/xxxx/Downloads/picture/video/转换后视频名.mp4"
```

⚠ 注意：

Windows 系统需要通过 `cd` 切换目录到压缩包 .exe 所在的文件夹后，再执行指令。

第三步：验证

验证是否为 H.264：在上一步指定路径下获取已转换视频后，可借助系统自带播放器或第三方 VLC 等软件检查。

视频制作接口-照片免训练

最近更新时间：2025-06-25 20:02:42

接口说明

无需训练，即可基于一个真人照片素材，通过输入文本或音频，生成人物说话口型与输入内容相匹配的新视频。
通过 [音视频制作进度查询接口](#) 最终返回成品视频。**未上传自定义存储 URL 的音视频资源只保留7天，请尽快下载。**

调用协议

HTTPS + JSON
POST /v2/ivh/videomaker/broadcastservice/phototovideonotrain
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
RefPhotoUrl	string	是	模板照片，格式支持 jpg、jpeg、png、bmp、webp。 1. 文件大小必须在10M内。 2. 图片单边分辨率要求在[192~4096]之间。 3. 图片宽高比在1:2到2:1范围内。 4. 需要有真人或类真人的卡通人脸，避免上传无人脸、人脸不完整、不清晰、偏转角度过大、嘴部被遮挡的图片。
DriverType	string	是	驱动类型，该字段必填。 1. Text：文本驱动，要求 InputSsml 字段必填。 2. OriginalVoice：原声音频驱动，要求 InputAudioUrl 字段必填。
InputAudioUrl	string	否	驱动数智人的音频 URL，当 DriverType 为 OriginalVoice 时，该字段必填。 音频格式要求： 1、时长在 [2, 60] 秒。 2、支持格式：wav、mp3、wma、m4a、aac、ogg。 3、文件大小：不超过20M。
InputSsml	string	否	播报文本内容，支持 SSML 标签，支持的标签类型参照 数智人 SSML 标记语言规范 ，标签写法参照示例，内容不能换行，符号需转义。上限300字（文本驱动底层会转换为音频，如果时长超过60秒时任务会制作失败），不少于4个字（字数按 unicode 字符数计算）。DriverType 为空、或 Text 时，该字段必填。
SpeechParam	object	否	定义音频的详细参数。DriverType 为 Text 时，该字段必填。
SpeechParam.Speed	float	否	语速（1.0为正常语速，范围[0.5~1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快，DriverType 为音频驱动类型时，语速控制不生效），DriverType 为 Text 时，该字段必填。
SpeechParam.TimbreKey	string	否	音色 Key，DriverType 为 Text 时，该字段必填。
SpeechParam.Volume	int	否	音量大小，范围[0, 10]，对应音量大小。默认为0，代表正常音量，值越大音量越高。 说明： TimbreKey 在 male_1-20、female_1-23（即男声1-20、女声1-23）间的音色不支持音量调节。
SpeechParam.EmotionCategory	string	否	控制合成音频的情感，仅支持多情感音色使用，可选值参考个人资产管理 API 分页查询音色列表 接口。
SpeechParam.EmotionIntensity	int	否	控制合成音频情感程度，取值范围为 [50,200]，只有 EmotionCategory 不为空时生效。

SpeechParam.TimbreLanguage	string	否	音色语种，可选语种参考个人资产管理 API 分页查询音色列表 接口，多语种音色在合成时必须选择对应语种。
ConcurrencyType	string	否	视频制作任务使用的资源类型。 1. Exclusive：使用并发调用，不扣除小时包，需要购买并发，如果没有购买，任务提交失败。 2. Shared：调用会扣除小时包，需要购买小时包，如果没有购买，任务提交失败。 3. 不填：如果购买了并发或者并发和小时包都购买则为“Exclusive”，如果没有购买并发但购买了小时包则为“Shared”，如果都没有购买，任务提交失败。
CallbackUrl	string	否	当用户增加回调 URL 时，将把视频制作结果以固定格式发送 POST 请求到该 URL 地址，固定格式见 附录二：回调请求体格式 ，需注意： 1. 限制 CallbackUrl 长度小于1000。 2. 只发送一次请求，无论是哪种问题导致的请求失败，都不会再进行发送。
VideoParam	object	否	定义输出视频的相关参数，不填时取各字段默认值。
VideoParam.EmotionLevel	int	否	输出视频的人物情绪强度：可选强度：1，2，3；默认值是2。越大音频控制强度越强，但可能导致不自然。

返回参数

参数	类型	必须	说明
TaskId	string	是	视频制作的任务 ID，携带 TaskId 访问 音视频制作进度查询接口 ，可获得制作进度和制作结果

请求示例

文本驱动

```
{
  "Header": {},
  "Payload": {
    "RefPhotoUrl": "http://virtualhuman-cos-test-1251316161.cos.ap-nanjing.myqcloud.com/ref_photo.jpg",
    "DriverType": "Text",
    "InputSsml": "你好，我是虚拟<phoneme alphabet=\"py\" ph=\"fu4\">主</phoneme>播",
    "SpeechParam": {
      "TimbreKey": "female_1",
      "Volume": 1,
      "Speed": 1.0
    }
  }
}
```

音频驱动

```
{
  "Header": {},
  "Payload": {
    "RefPhotoUrl": "http://virtualhuman-cos-test-1251316161.cos.ap-nanjing.myqcloud.com/ref_photo.jpg",
    "DriverType": "OriginalVoice",
    "InputAudioUrl": "http://virtualhuman-cos-test-1251316161.cos.ap-nanjing.myqcloud.com/audio.mp3"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "DialogID": "",
    "Message": "",
    "RequestID": "fde854eaa981c7f2f7285d1c7eca335b",
    "SessionID": "gzb7dec22117297528294581119"
  },
  "Payload": {
    "TaskId": "81883d47c6154edf8e276531f09227b6"
  }
}
```

音视频制作进度查询接口

最近更新时间：2025-06-23 19:20:32

接口说明

通过 TaskId 来查询任务的制作进度和结果，当返回值里的 Progress 字段值为100时，即可通过 MediaUrl 获取最终音视频的下载地址，如果是视频制作也会返回 SubtitlesUrl 字段获取最终 SRT 字幕地址。

 **注意：**

制作完成的视频资源默认存储在数智人内部账号的 COS 下，只保留7天，可以通过在制作接口中的 VideoStorageS3Url 字段指定第三方存储 URL，这样制作完的视频会上传到指定地址。

调用协议

HTTPS + JSON
POST /v2/ivh/videomaker/broadcastservice/getprogress
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
TaskId	String	是	任务 ID，在 音频制作接口 、 视频制作接口-基础版 、 视频制作接口-进阶版 接口中返回。

返回参数

参数	类型	必须	说明
Progress	int	是	制作进度，-1~100，-1代表生成失败，100代表生成成功（预留字段，目前不具备参考意义）
MediaUrl	String	是	音视频结果地址
SubtitlesUrl	String	否	当制作视频时，返回视频对应的 SRT 字幕地址
Status	String	是	制作状态 "COMMIT": 已提交需要排队 "MAKING": 制作中 "SUCCESS": 制作成功 "FAIL": 制作失败
ArrayCount	int	是	Status 为 "COMMIT" 状态时在该任务之前排队的任务数量
FailCode	int	否	制作失败返回的错误码， 对应附录一中的“音视频制作结果错误码”
FailMessage	String	否	制作失败返回的失败原因，便于排查问题
TextTimestampResult	Array of [Sentence]	否	该字段返回 TTS 试听任务的文本时间戳信息
Duration	int	是	视频时长，单位 ms

Sentence

参数	类型	必须	说明
Sentence	String	是	拆句得出的句子
Words	Array of [Word]	是	句子中每个字的信息

Word

参数	类型	必须	说明
Word	String	是	句子中的一个字
StartTimestamp	long	是	该字的时间起点，该数值/10000为 ms，相当于单位为 0.1us
EndTimestamp	long	是	该字的时间尾点，该数值/10000为 ms，相当于单位为 0.1us

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "81883d47c6154edf8e276531f09227b6"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "Message": "",
    "RequestID": "fde854eaa981c7f2f7285d1c7eca335b",
    "SessionID": "gzb7dec22117297528294581119"
  },
  "Payload": {
    "Progress": 100,
    "MediaUrl": "https://videomaker-resources.ivh.qq.com/81883d47c6154edf8e276531f09227b6.mp3",
    "SubtitlesUrl": "",
    "ArrayCount": 0,
    "FailMessage": "",
    "Duration": 11810,
    "FailCode": 0,
    "TextTimestampResult": [
      {
        "Sentence": "你好虚拟主播。",
        "Words": [
          {
            "Word": "你",
            "EndTimestamp": 6100000,
            "StartTimestamp": 4500000
          },
          {
            "Word": "好",
            "EndTimestamp": 8200000,
            "StartTimestamp": 6100000
          },
          {
            "Word": "虚",
            "EndTimestamp": 9700000,
            "StartTimestamp": 8200000
          },
          {
            "Word": "拟",
            "EndTimestamp": 11100000,
            "StartTimestamp": 9700000
          }
        ]
      }
    ]
  }
}
```

```
    },  
    {  
      "Word": "主",  
      "EndTimestamp": 12900000,  
      "StartTimestamp": 11100000  
    },  
    {  
      "Word": "播。",  
      "EndTimestamp": 16000000,  
      "StartTimestamp": 12900000  
    }  
  ]  
}  
]  
}
```

附录

附录一： 结果码字典

最近更新时间：2025-05-14 17:43:22

一、接口错误码

代码	含义	措施
0	成功	无
-101000	系统内部错误	重试或联系管理员定位错误
-101001	算力不足	重试或联系管理员定位错误
-101002	下载图片失败	确认提供的图片 url 是否可正常访问、下载
-101003	请求S3失败	重试或联系管理员定位错误
100001/100003	请求参数错误（包括参数格式、类型等错误，缺少参数错误，必传参数没填）	对照 API 协议核查参数格式、类型
100002	参数取值错误	对照 API 协议核查参数列表
100005	未授权操作	检查对请求的会话是否有操作权限
100008	超过配额限制	会话超过并发限制
100009	资源不存在	当前权限范围是否越权
110005	记录已经存在	过快的重复请求
110006	记录不存在	请检查请求的会话是否存在
110007	virtualmankey 已过期	检查 key 所属主账号的有效期限
100013	未找到有效的小时包或小时包余量不足	申请或加购对应小时包
100015	数智人形象已经过期	申请形象延期
100016	未找到形象资产	确认使用的 VirtualmanKey 是否有效
801000	调用远程服务异常	重试或联系管理员定位错误
801002	调用远程服务，返回状态码异常	重试或联系管理员定位错误
801004	操作频繁，请稍后重试	避免同一时间提交多个任务
801005	文本安全校验未通过	请检查 ssml 是否有敏感信息

二、音视频制作结果错误码

代码	含义	措施
0	成功	无
801000	系统内部错误	重试或联系管理员定位错误
801001	数据库操作错误	重试或联系管理员定位错误

801002	调用远程服务错误	重试或联系管理员定位错误
801003	任务制作超过重试次数	重试或联系管理员定位错误
801004	任务制作过程发生崩溃	重试或联系管理员定位错误
801006	操作S3失败	重试或联系管理员定位错误
801007	任务执行超时	重试或联系管理员定位错误
801008	下载文件失败	检查提供的文件资源是否可正常访问、下载
801009	参数错误	检查 FailMessage 中提示的参数是否符合接口要求
801010	获取 url 资源失败	检查提供的 url 资源是否可正常访问、下载
801012	任务参数解析失败	重试或联系管理员定位错误
801014	下载字幕文件失败/下载生成视频失败	重试或联系管理员定位错误
801015	上传文件到指定的S3失败	检查提供的 VideoStorageS3Url、SubtitleStorageS3Url 是否可用
801401	音频制作，解析任务参数失败	重试或联系管理员定位错误
801402	音频制作，创建本地文件失败	重试或联系管理员定位错误
801403	音频制作，接收音频数据失败	重试或联系管理员定位错误
801404	音频制作，接收音频数据，发生业务异常	重试或联系管理员定位错误
801405	音频制作，mp3 文件转换失败或合成背景音乐失败	重试或联系管理员定位错误
801501	视频制作，解析参数错误	重试或联系管理员定位错误
801503	视频制作，调用远程服务失败	重试或联系管理员定位错误
801504	视频制作，创建拼接文件失败	重试或联系管理员定位错误
801506	视频制作，制作 logo/字幕或者拼接视频失败	重试或联系管理员定位错误
801507	视频制作，传入图片等资源分辨率转换失败	重试或联系管理员定位错误
801508	视频制作，获取小样本指定视频片段帧数失败	重试或联系管理员定位错误
801509	视频制作，检查小样本输入帧数不符合规定	检查 StartIdx、EndIdx 是否符合接口定义
801510	视频制作，音频驱动文件时长校验失败	检查音频驱动文件、重试或联系管理员定位错误
802000	免训练视频制作，照片审核失败	检查上传的照片内容是否涉及敏感信息
802001	免训练视频制作，照片人脸检测失败	检查上传的照片中是否仅有单一人脸
802002	免训练视频制作，预处理失败	重试或联系管理员定位错误
802003	免训练视频制作，SDK 处理错误	请检查输入视频或照片中是否存在人脸丢失情况
802004	免训练视频制作，回调错误	重试或联系管理员定位错误

附录二： 回调请求体格式

最近更新时间：2024-11-13 14:38:22

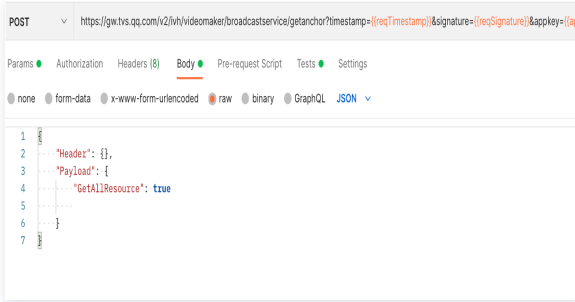
参数	类型	必须	说明
Progress	int	是	制作进度，-1代表生成失败，100代表生成成功（预留字段，目前不具备参考意义）
MediaUrl	string	是	音视频结果地址
SubtitlesUrl	string	否	当制作视频时，返回视频对应的SRT字幕地址
Status	string	是	制作状态 "SUCCESS": 制作成功 "FAIL": 制作失败
FailMessage	string	否	制作失败返回的失败原因，便于排查问题
TaskId	string	是	请求视频制作返回的 TaskId

请求示例

```
{
  "Payload": {
    "Progress": 100,
    "MediaUrl": "https://videomaker-resources.ivh.qq.com/broadcast/Basic/video/25215772ab424ad38858f6b692426989.mp4",
    "SubtitlesUrl": "https://videomaker-resources.ivh.qq.com/broadcast/Basic/subtitle/25215772ab424ad38858f6b692426989.srt",
    "Status": "SUCCESS",
    "FailMessage": "",
    "TaskId": "25215772ab424ad38858f6b692426989"
  }
}
```

接口对接常见问题

最近更新时间：2024-04-26 17:01:52

问题描述	解决方式	备注
调用接口时出现鉴权不通过	1. 确认签名时参数按照字典序排列了 2. 确认时间戳参数为当前时间 3. 参考右侧的示例参数，按照自己的签名步骤计算一下，如果计算结果与示例一致则表示签名计算步骤无误。	参数： appkey = e38267c0e86411ebb02aed82acb0ed99 accesstoken = f68f2d10ae9e4604b76fb05cf46bcce timestamp=1646636485 签名结果： BfWuaC9kmaicCggXc693uK%2BsZQ8qe88O4HVQwZuo%3D
调用接口出现参数错误	1. 请确认请求的 body 是否加了 header 参数，不加的话会报错。如右图所示。 2. 参数的类型要严格按照接口文档中填充。	
查询出的形象数量不对或查询不到形象	如果开通的是小样本形象需要参考 分页查询小样本形象列表接口 查询。 如果开通的是2d精品形象需要参考 客户资源查询主播接口 查询。 ❗ 说明： 官网下单播报小时包后，开通的形象既有2d小样本，又有2d精品，需要参考两套接口分别查询相关形象。	
是否支持生成透	支持，可以生成 webm 格式的带透明通道的视频	

明背景的视频		
为什么webm透明背景的视频播放出来是绿幕的	说明客户使用的播报器不支持 webm 格式，可以在 Chrome 中播放看看，如果出来的视频是黑底的，说明视频本身没问题。	
形象位置参数调整后不生效	1. 确认调用的接口是基础版还是进阶版，只有进阶版接口支持。 2. 确认视频输出格式是否为 GreenScreenMp4，设置为该值时不支持调整位置。	

主播位置计算方法



参数1: HorizontalPosition: 主播横向往位置 (默认值是0, 负数为向左调整, 正数为向右调整)
定义: 屏幕中轴线距形象中轴线距离/ 整个屏幕宽度;
公式: $HorizontalPosition = (x + anchorWidth / 2 - containerWidth / 2) / containerWidth$

参数2: VerticalPosition: 主播纵向位置 (默认值是0, 负数为向上调整, 正数为向下调整)
定义: 形象底部距屏幕底部距离 / 两倍形象高度
公式: $VerticalPosition = (anchorHeight + y - wrapperHeight) / 2 / anchorHeight$

参数3: Scale: 主播大小 (默认值是: 1; 数值越小, 主播越小)
缩放规则: 在左下顶点保持不动, 按照形象宽度、形象高度等比例缩放。

相关参数的定义:
①anchorHeight: 即为叠加了缩放系数的形象高度, 即"原始图片高度" * originZoom * Scale
②anchorWidth: 即为叠加了缩放系数的形象宽度, 即"原始图片宽度" * originZoom * Scale
③原始图片高度、宽度: 即《个人资产管理api接口协议》→4.1分页查询小样本形象列表接口中, PoselImage的高度和宽度。
④originZoom: 即默认的缩放系数, 为了在不同的分辨率下人物默认有最佳展示效果, 赋予了一个默认的缩放系数, 可以在《个人资产管理api接口协议》→4.1分页查询小样本形象列表接口 查询。

如何计算数智人初始大小

数智人查询接口中会返回OriginZoom(原始缩放系数)和PoselImage(数智人姿态图片), OriginZoom 乘以 PoselImage的宽高即为数智人初始大小。

文案中插入动作不生效

动作前后需要预留一定的文字, 每个动作对预留文字都不一样, 建议前后预留10个字以上。

数智人是否支持调整语速, 韵律, 多音字

可以, 参考接口文档中的ssml标签 [附录二: SSML 规范](#) 部分。

指定了外部的S3存储，数智人视频未上传上去	请检查S3的权限设置是否允许外部写入，或者参数中的S3链接是否带入了鉴权信息。	
没有字幕文件	只有文本驱动才有字幕文件，声音驱动没有	
超过配额限制：提交音频制作数量超过限制：1	开通限时默认只开通1路音频制作并发，即同时只能有一个音频任务处于制作中，必须等上一个制作完成后才能提交下一个任务。可以通过购买音频并发解决。	
超过配额限制：提交视频	1. 针对每种形象类型，每个用户默认支持同时处理5个视频制作任务，由主子账号共享。这5个任务使用的并发资源所有客户共享，可能会出现任务排队的情况。 2. 可增购播报独享并发，例如购买了1路播报独享并发，则一共支持同时处理6个视频制作任务。增购的并发由该用户专享。 该问题解决方案： 1. 用户增加任务排队策略； 2. 购买独享并发。	

制作数量超过限制：
n

个人资产管理 API 文档

个人资产管理 API 概述

最近更新时间：2024-11-11 15:12:11

本文档主要描述客户资产管理接口，客户可以通过该套接口查询到客户当前开通的数智人相关资产，包括**形象资产**、**声音资产**和**服务资产**，服务资产包括定制配额、播报小时报、播报并发以及交互并发等。

接口调用环境

正式环境：`https://gw.tvs.qq.com`

接口调用方式

具体内容请参考 [数智人 aPaas 接口调用方式](#)

查询小样本形象列表

最近更新时间：2024-12-16 10:00:24

接口说明

查询客户拥有的小样本形象列表。

调用协议

HTTPS + JSON
POST /v2/ivh/crmserver/customerassetsservice/describessmallsampleimage
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
PageIndex	int	是	当前页
PageSize	int	是	页面大小（最大不超过100）
AnchorCodes	Array of string	否	AnchorCode 筛选条件
VirtualmanKey	string	否	数智人 VirtualmanKey

返回参数

参数	类型	必须	说明
Virtualmans	Array of [SmallSampleImage]	是	小样本形象列表
Total	int	是	总数

SmallSampleImage

参数	类型	必须	说明
VirtualmanKey	string	是	数智人 VirtualmanKey，形象唯一标识
AnchorName	string	是	数智人主播名称
AnchorCode	string	是	数智人主播 Code
ClothesName	string	是	数智人服装
PoseName	string	是	数智人姿态
Resolution	string	是	数智人分辨率
HeaderImage	string	是	数智人头像图片 URL
PoselImage	string	是	数智人姿态图片 URL
ClothesImage	string	是	数智人服装图片 URL
SupportDriverTypes	Array of [string]	是	数智人支持的驱动类型 1. Text: 文本驱动 2. OriginalVoice: 原声音频驱动 3. ModulatedVoice: 变声音频驱动
EffectiveDate	string	是	资产生效时间，格式：yyyy-MM-dd HH:mm:ss

EffectiveTime	int	是	资产生效时间戳，单位毫秒
ExpireDate	string	是	资产到期时间，格式：yyyy-MM-dd HH:mm:ss
ExpireTime	int	是	资产到期时间戳，单位毫秒
VideoDuration	int	是	数智人累计制作视频时长，单位为毫秒
VideoNum	int	是	数智人累计制作视频个数
OriginZoom	float	是	数智人初始缩放系数
OriginZoomStr	string型浮点数	是	数智人初始缩放系数，值精确到小数点后3位
ReferenceVideoSegmentUrl	string	是	小样本选用视频片段 URL

请求示例

```
{
  "Header": {},
  "Payload": {
    "PageIndex": 1,
    "PageSize": 10,
    "AnchorCodes": ["wl_pink_skirt_suit_stand"],
    "VirtualmanKey": "d3bf659f364243ceacbe026710bdc25f"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "6dbf3272a613d3a5cafa69f8bda562ef",
    "SessionID": "gze0f59a6c17311577017347309",
    "DialogID": "",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "Total": 1,
    "Virtualmans": [
      {
        "AnchorCode": "wl_pink_skirt_suit_stand",
        "AnchorName": "微澜-粉短裙套装-全身站姿",
        "ClothesImage": "https://virtualhuman-cos-test-1251316161.cos.ap-nanjing.myqcloud.com/virtualman-config/00e92768-3d23-40d7-aa39-94df36a44a38-Mask_group粉短裙套装.png",
        "ClothesName": "粉短裙套装",
        "EffectiveDate": "2023-10-31 14:49:44",
        "EffectiveTime": 1698734984000,
        "ExpireDate": "2023-11-14 14:49:44",
        "ExpireTime": 1699944584000,
        "HeaderImage": "https://virtualhuman-cos-prod-1251316161.cos.ap-nanjing.myqcloud.com/virtualman-config/c0948a75-b4fd-461e-88e6-c15ef7c3b8f7-weilan_xyb_fenfuanqun.png",
        "OriginZoom": 0.332,
        "OriginZoomStr": "0.332",
        "PoseImage": "https://virtualhuman-cos-prod-1251316161.cos.ap-nanjing.myqcloud.com/virtualman-config/e75180c0-393b-4fec-9895-98717ddd3db5-weilan_xyb_fenfuanqun.png",
        "PoseName": "全身站姿",
      }
    ]
  }
}
```

```
        "ReferenceVideoSegmentUrl": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/prod/resource-manager/small/57/159/4873/model_6859_20230812184617/pv.mp4",  
        "Resolution": "1280x720",  
        "SupportDriverTypes": [  
            "Text",  
            "OriginalVoice",  
            "ModulatedVoice"  
        ],  
        "VideoDuration": 120000,  
        "VideoNum": 2,  
        "VirtualmanKey": "d3bf659f364243ceacbe026710bdc25f"  
    }  
]  
}
```

查询服务资产信息

最近更新时间：2024-11-11 15:12:11

接口说明

查询客户拥有的服务资产信息，包括播报小时包、播报并发、交互并发、形象定制配额、声音复刻定制配额和音色小时包等。

调用协议

HTTPS + JSON
POST /v2/ivh/crmserver/customerassetsservice/getserviceasset
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
ServiceType	string	否	服务资产类型，具体类型见 附录1 ，不传默认查询所有类型
PageIndex	int	是	当前页
PageSize	int	是	页面大小（最大不超过100）

返回参数

参数	类型	必须	说明
ServiceAssetInfos	Array of [ServiceAssetInfo]	是	服务资产列表
Total	int	是	总数

ServiceAssetInfo

参数	类型	必须	说明
ServiceType	string	是	服务资产类型
ExpireDate	string	是	服务资产到期时间，格式：yyyy-MM-dd HH:mm:ss
UniversalAssetInfoDetail	UniversalAssetInfoDetail	是	定制配额类资产详情
UniversalConcurrencyInfoDetail	UniversalConcurrencyInfoDetail	是	并发类资产详情
UniversalHourInfoDetail	UniversalHourInfoDetail	是	小时包类资产详情
InteractConcurrencyInfoDetail	[InteractConcurrencyInfoDetail]	否	2D小样本（专属）交互并发资产详情
BroadcastHourInfoDetail	[BroadcastHourInfoDetail]	否	2D小样本（专属）播报小时包资产详情
BroadcastConcurrencyInfoDetail	[BroadcastConcurrencyInfoDetail]	否	2D小样本（专属）播报并发数资产详情
ImageAssetInfoDetail	[ImageAssetInfoDetail]	否	2D小样本（专属）形象定制配额资产详情
SmallSampleGeneral2DImageAssetInfoDetail	[SmallSampleGeneral2DImageAssetInfoDetail]	否	2D小样本（通用）形象定制配额资产详情
SmallSampleGeneral2DInter	[SmallSampleGeneral2DInter	否	2D小样本（通用）交互并发资产详情

ractConcurrencyInfoDetail	actConcurrencyInfoDetail]		
SmallSampleGeneral2DBroadcastHourInfoDetail	[SmallSampleGeneral2DBroadcastHourInfoDetail]	否	2D小样本（通用）播报小时包资产详情
SmallSampleGeneral2DBroadcastConcurrencyInfoDetail	[SmallSampleGeneral2DBroadcastConcurrencyInfoDetail]	否	2D小样本（通用）播报并发数资产详情
RealMan2DInteractConcurrencyInfoDetail	[RealMan2DInteractConcurrencyInfoDetail]	否	2D精品交互并发数资产详情
RealMan2DBroadcastHourInfoDetail	[RealMan2DBroadcastHourInfoDetail]	否	2D精品播报小时包资产详情
RealMan2DBroadcastConcurrencyInfoDetail	[RealMan2DBroadcastConcurrencyInfoDetail]	否	2D精品播报并发数资产详情
RealMan3DInteractConcurrencyInfoDetail	[RealMan3DInteractConcurrencyInfoDetail]	否	3D写实交互并发资产详情
RealMan3DBroadcastHourInfoDetail	[RealMan3DBroadcastHourInfoDetail]	否	3D写实播报小时包资产详情
RealMan3DBroadcastConcurrencyInfoDetail	[RealMan3DBroadcastConcurrencyInfoDetail]	否	3D写实播报并发数资产详情
TimbreAssetInfoDetail	[TimbreAssetInfoDetail]	否	声音复刻配额资产详情
TimbreHourInfoDetail	[TimbreHourInfoDetail]	否	音色小时包资产详情

UniversalAssetInfoDetail

参数	类型	必须	说明
CustomizationNum	int	是	定制数量
RemainNum	int	是	剩余数量

UniversalConcurrencyInfoDetail

参数	类型	必须	说明
Concurrency	int	是	并发数

UniversalHourInfoDetail

参数	类型	必须	说明
OrderMilliseconds	int	是	播报小时包时长，单位为毫秒
RemainMilliseconds	int	是	播报小时包剩余时长，单位为毫秒

InteractConcurrencyInfoDetail

参数	类型	必须	说明
Concurrency	int	是	并发数

BroadcastHourInfoDetail

参数	类型	必须	说明
OrderMilliseconds	int	是	播报小时包时长，单位为毫秒

RemainMilliseconds	int	是	播报小时包剩余时长，单位为毫秒
--------------------	-----	---	-----------------

BroadcastConcurrencyInfoDetail

参数	类型	必须	说明
Concurrency	int	是	并发数

ImageAssetInfoDetail

参数	类型	必须	说明
CustomizationNum	int	是	形象定制数量
RemainNum	int	是	形象定制剩余数量

SmallSampleGeneral2DImageAssetInfoDetail

参数	类型	必须	说明
CustomizationNum	int	是	定制数量
RemainNum	int	是	剩余数量

SmallSampleGeneral2DInteractConcurrencyInfoDetail

参数	类型	必须	说明
Concurrency	int	是	并发数

SmallSampleGeneral2DBroadcastHourInfoDetail

参数	类型	必须	说明
OrderMilliseconds	int	是	播报小时包时长，单位为毫秒
RemainMilliseconds	int	是	播报小时包剩余时长，单位为毫秒

SmallSampleGeneral2DBroadcastConcurrencyInfoDetail

参数	类型	必须	说明
Concurrency	int	是	并发数

RealMan2DInteractConcurrencyInfoDetail

参数	类型	必须	说明
Concurrency	int	是	并发数

RealMan2DBroadcastHourInfoDetail

参数	类型	必须	说明
OrderMilliseconds	int	是	播报小时包时长，单位为毫秒
RemainMilliseconds	int	是	播报小时包剩余时长，单位为毫秒

RealMan2DBroadcastConcurrencyInfoDetail

参数	类型	必须	说明
OrderMilliseconds	int	是	播报小时包时长，单位为毫秒

RemainMilliseconds	int	是	播报小时包剩余时长，单位为毫秒
--------------------	-----	---	-----------------

RealMan3DInteractConcurrencyInfoDetail

参数	类型	必须	说明
Concurrency	int	是	并发数

RealMan3DBroadcastHourInfoDetail

参数	类型	必须	说明
OrderMilliseconds	int	是	播报小时包时长，单位为毫秒
RemainMilliseconds	int	是	播报小时包剩余时长，单位为毫秒

RealMan3DBroadcastConcurrencyInfoDetail

参数	类型	必须	说明
Concurrency	int	是	并发数

TimbreAssetInfoDetail

参数	类型	必须	说明
CustomizationNum	int	是	声音复刻数量
RemainNum	int	是	声音复刻剩余数量

TimbreHourInfoDetail

参数	类型	必须	说明
OrderMilliseconds	int	是	音色小时包时长，单位为毫秒
RemainMilliseconds	int	是	音色小时包剩余时长，单位为毫秒

请求示例

```
{
  "Header": {},
  "Payload": {
    "ServiceType": "",
    "PageIndex": 1,
    "PageSize": 10
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "890efbe5efdc8f75a75a47a8a19eeac",
    "SessionID": "gz92439a3317313012788482967",
    "DialogID": "",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
```

```
"ServiceAssetInfos": [  
  {  
    "BroadcastHourInfoDetail": {  
      "OrderMilliseconds": 999999999,  
      "RemainMilliseconds": 881318391  
    },  
    "ExpireDate": "2024-04-05 11:31:07",  
    "ServiceType": "BroadcastHour",  
    "UniversalHourInfoDetail": {  
      "OrderMilliseconds": 999999999,  
      "RemainMilliseconds": 881318391  
    }  
  },  
  {  
    "ExpireDate": "2024-04-11 14:59:46",  
    "ImageAssetInfoDetail": {  
      "CustomizationNum": 5,  
      "RemainNum": 0  
    },  
    "ServiceType": "ImageAsset",  
    "UniversalAssetInfoDetail": {  
      "CustomizationNum": 5,  
      "RemainNum": 0  
    }  
  }  
],  
"Total": 2  
}
```

查询声音资产信息

最近更新时间：2025-05-14 14:08:02

接口说明

查询客户拥有的声音资产信息，包括复刻声音和公共声音。

调用协议

HTTPS + JSON
POST /v2/ivh/crmserver/customerassetsservice/describetimbre
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
PageIndex	int	是	当前页
PageSize	int	是	页面大小（最大不超过100）
TimbreKey	string	否	数智人音色 code，支持精确匹配
TimbreName	string	否	数智人音色名称，支持模糊匹配
PurchaseType	string	否	购买类型，Lease:租赁，Custom:定制
TimbreKeyList	Array of string	否	数智人音色 code 数组，当前参数与 TimbreKey 同时存在时，优先取当前字段

返回参数

参数	类型	必须	说明
TimbreAssets	Array of [TimbreAssetDetail]	是	声音资产列表
Total	int	是	总数

TimbreAssetDetail

参数	类型	必须	说明
TimbreKey	string	是	数智人音色 Code
TimbreName	string	是	数智人音色名称
TimbreSample	string	是	数智人音色试听样例 URL
TimbreDesc	string	是	数智人音色描述
SupportVirtualmanTypeCodes	Array of [string]	是	音色支持的数智人类型 Code，参考 附录3-数智人类型
SupportDriverTypes	Array of [string]	是	音色支持的驱动类型 1. Text: 文本驱动 2. OriginalVoice: 原声音频驱动 3. ModulatedVoice: 变声音频驱动
ExpireDate	string	否	资产到期时间，为空表示长期有效 格式：yyyy-MM-dd HH:mm:ss

AudioDuration	int	是	音频累计调用时长，单位为毫秒
AudioNum	int	是	音频累计调用数量
EmotionalCategories	Array of [string]	否	音色支持的情感风格列表，仅多情感音色才有数据，含义见 附录2-情感风格
TimbreGender	int	是	音色性别，1:男，2:女
PurchaseType	string	是	购买类型，Lease:租赁，Custom:定制
TimbreType	string	是	音色类型 Boutique: 精品 SampleBasic: 小样本基础版 SampleFast: 小样本极速版
TimbreLanguages	Array of [string]	否	音色支持的语种列表，仅部分音色支持多语种，含义见 附录4-语种列表

请求示例

```
{
  "Header": {},
  "Payload": {
    "PageIndex": 1,
    "PageSize": 10
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "9ab72535a33d42efea712fb19aef0971",
    "SessionID": "gze0f59a6c17313015362017196",
    "DialogID": "",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "TimbreAssets": [
      {
        "AudioDuration": 221587,
        "AudioNum": 11,
        "EmotionalCategories": [
          "6",
          "0",
          "1",
          "2",
          "3",
          "4",
          "5",
          "7",
          "8",
          "9"
        ],
        "ExpireDate": "2025-09-01 18:00:00",
        "PurchaseType": "Lease",
        "SupportDriverTypes": [
```

```

        "Text",
        "OriginalVoice",
        "ModulatedVoice"
    ],
    "SupportVirtualmanTypeCodes": [
        "real_man_2d",
        "small_sample_2d",
        "real_man_3d",
        "half_realistic_3d",
        "cartoon_3d",
        "small_sample_general_2d",
        "small_sample_4k_2d",
        "small_sample_photo_2d"
    ],
    "TimbreDesc": "温柔磁性",
    "TimbreGender": 2,
    "TimbreKey": "female_1",
    "TimbreLanguages": [],
    "TimbreName": "女声1",
    "TimbreSample": "https://virtualhuman-cos-prod-1251316161.cos.ap-
nanjing.myqcloud.com/virtualman-config/6bf3915e-fcc2-49b7-8fe7-650e1cbf6633-
10001_2_33_2D_demo_nv1.wav",
    "TimbreType": "Boutique"
},
{
    "AudioDuration": 9389,
    "AudioNum": 2,
    "EmotionalCategories": [],
    "ExpireDate": "",
    "PurchaseType": "Lease",
    "SupportDriverTypes": [
        "Text",
        "OriginalVoice"
    ],
    "SupportVirtualmanTypeCodes": [
        "real_man_2d",
        "small_sample_2d",
        "real_man_3d",
        "half_realistic_3d",
        "cartoon_3d",
        "small_sample_general_2d",
        "small_sample_4k_2d",
        "small_sample_photo_2d"
    ],
    "TimbreDesc": "对话风格",
    "TimbreGender": 1,
    "TimbreKey": "male_11",
    "TimbreLanguages": [],
    "TimbreName": "男声11",
    "TimbreSample": "https://virtualhuman-cos-prod-1251316161.cos.ap-
nanjing.myqcloud.com/virtualman-config/84fcec66-885e-4b9d-a66e-7fe1fab3394-
0c9a40260f20447ab772e2e02d46e1ae.wav",
    "TimbreType": "Boutique"
}
],
"Total": 2
}
}

```

查询形象资产信息-查询主播

最近更新时间：2025-03-25 17:57:02

接口说明

查询客户拥有的形象资产信息，查询主播列表。

调用协议

HTTPS + JSON
POST /v2/ivh/crmserver/customerassetsservice/getanchor
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
VirtualmanTypeCode	string	是	数智人类型 Code，参考 附录3-数智人类型
AnchorCodes	Array of [string]	否	数智人主播 Code
PurchaseType	string	否	购买类型，Lease:租赁，Custom:定制
PageIndex	int	是	当前页
PageSize	int	是	页面大小（最大不超过100）

返回参数

参数	类型	必须	说明
Virtualmans	Array of [Virtualman]	是	主播列表
Total	int	是	总数

Virtualman

参数	类型	必须	说明
AnchorName	string	是	数智人主播名称
AnchorCode	string	是	数智人主播 Code
VirtualmanType	string	是	数智人类型名称
VirtualmanTypeCode	string	是	数智人类型 Code
PurchaseType	string	是	购买类型，Lease:租赁，Custom:定制
HeaderImage	string	是	数智人头像 URL
Tag	string	是	数智人模型tag，分三类： 1. Basic 2. Standard 3. Advanced 目前 tag 仅影响数智人横纵向位置功能

请求示例

```
{
  "Header": {},
  "Payload": {
    "VirtualmanTypeCode": "real_man_3d",
    "AnchorCodes": [],
  }
}
```



```
    "PageIndex": 1,  
    "PageSize": 10  
  }  
}
```

返回示例

```
{  
  "Header": {  
    "RequestID": "4ff85e77c7c5067cb9a79802879db144",  
    "SessionID": "gze0f59a6c17313022465941770",  
    "DialogID": "",  
    "Code": 0,  
    "Message": "ok"  
  },  
  "Payload": {  
    "Virtualmans": [  
      {  
        "HeaderImage": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/resource-manager/3D_Cartoon/xiaowei/202306091551/header.png",  
        "VirtualmanTypeCode": "real_man_3d",  
        "VirtualmanType": "3D写实",  
        "Tag": "Basic",  
        "PurchaseType": "Custom",  
        "AnchorCode": "xiaowei",  
        "AnchorName": "小微"  
      }  
    ],  
    "Total": 1  
  }  
}
```

查询形象资产信息-查询主播下的所有形象

最近更新时间：2025-03-25 17:57:02

接口说明

查询客户拥有的形象资产信息，查询某主播下的形象列表。

调用协议

HTTPS + JSON
POST /v2/ivh/crmserver/customerassetsservice/getimagebyanchor
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
VirtualmanTypeCode	string	是	数智人类型 Code，参考 附录3-数智人类型
AnchorCodes	Array of [string]	否	数智人主播 Code
PageIndex	int	是	当前页
PageSize	int	是	页面大小（最大不超过100）
VirtualmanKey	string	否	数智人 VirtualmanKey
PurchaseType	string	否	购买类型，Lease:租赁，Custom:定制

返回参数

参数	类型	必须	说明
Virtualmans	Array of [ImageAssetDetail]	是	该主播下的形象列表 ❗ 说明： 当对应 VirtualmanTypeCode 下未根据接口参数查询到形象时，形象列表返回为空。
Total	int	是	总数

ImageAssetDetail

参数	类型	必须	说明
AnchorName	string	是	数智人主播名称
AnchorCode	string	是	数智人主播 Code
VirtualmanType	string	是	数智人类型
VirtualmanTypeCode	string	是	数智人类型 Code
VirtualmanKey	string	是	数智人Virtualmankey
ClothesName	string	是	数智人服装名称
PoseName	string	是	数智人姿态
Resolution	string	是	数智人分辨率

HeaderImage	string	是	数智人头像图片 URL
PoselImage	string	是	数智人姿态图片 URL
ClothesImage	string	是	数智人服装图片 URL
SupportDriverTypes	Array of [string]	是	数智人支持的驱动类型 Text: 文本驱动 OriginalVoice: 原声音频驱动 ModulatedVoice: 变声音频驱动
SupportAnchorExtraParams	Array of [SupportAnchorExtraParam]	是	数智人可配置形象参数说明，目前仅包含3D服装变色参数
VideoDuration	int64	是	数智人累计调用时长，单位为毫秒
VideoNum	int64	是	数智人累计视频个数
ExpireDate	string	是	资产到期时间，格式：yyyy-MM-dd HH:mm:ss
OriginZoom	string	是	缩放系数，保留三位小数
ReferenceVideoSegmentUrl	string	是	选用视频片段
PurchaseType	string	是	购买类型，Lease:租赁，Custom:定制

SupportAnchorExtraParam

参数	类型	必须	说明
Key	string	是	配置项名称
Description	string	是	配置项描述
Type	int	是	配置项的取值类型。1：字符串，2：数值，3：布尔
Value	string[]	是	当前配置项，可配置的字段列表

请求示例

```
{
  "Header": {},
  "Payload": {
    "VirtualmanTypeCode": "small_sample_2d",
    "AnchorCodes": [
      "wl_pink_skirt_suit_stand"
    ],
    "PageIndex": 1,
    "PageSize": 10
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "612258c65d2da9401710b06e16e0bed1",
    "SessionID": "gzfa9105ca17313024090292042",
  }
}
```

```
"DialogID": "",
"Code": 0,
"Message": "ok"
},
"Payload": {
  "Total": 2,
  "Virtualmans": [
    {
      "AnchorCode": "wl_pink_skirt_suit_stand",
      "AnchorName": "微澜-粉短裙套装-全身站姿",
      "ClothesImage": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/virtualman-config/00e92768-3d23-40d7-aa39-94df36a44a38-Mask group粉短裙套装.png",
      "ClothesName": "粉短裙套装",
      "ExpireDate": "2024-03-13 21:37:02",
      "HeaderImage": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/prod/resource-manager/small/57/159/4873/model_11131_20240429141943/header.png",
      "OriginZoom": "1.000",
      "PoseImage": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/prod/resource-
manager/small/57/159/4873/model_11131_20240429141943/static_rgba.png",
      "PoseName": "全身站姿",
      "ReferenceVideoSegmentUrl": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/prod/resource-manager/small/57/159/4873/model_11131_20240429141943/pv.mp4",
      "Resolution": "1080x1920",
      "SupportAnchorExtraParams": [],
      "SupportDriverTypes": [
        "Text",
        "OriginalVoice",
        "ModulatedVoice"
      ],
      "VideoDuration": 0,
      "VideoNum": 0,
      "VirtualmanKey": "2cc66ae2950e4968880676a055013234",
      "VirtualmanType": "2D小样本(专属)",
      "VirtualmanTypeCode": "small_sample_2d",
      "PurchaseType": "Custom"
    },
    {
      "AnchorCode": "wl_pink_skirt_suit_stand",
      "AnchorName": "微澜-粉短裙套装-全身站姿",
      "ClothesImage": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/virtualman-config/00e92768-3d23-40d7-aa39-94df36a44a38-Mask group粉短裙套装.png",
      "ClothesName": "粉短裙套装",
      "ExpireDate": "2024-03-13 21:37:02",
      "HeaderImage": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/prod/resource-manager/small/57/159/4873/model_11131_20240429141943/header.png",
      "OriginZoom": "0.571",
      "PoseImage": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/prod/resource-
manager/small/57/159/4873/model_11131_20240429141943/static_rgba.png",
      "PoseName": "全身站姿",
      "ReferenceVideoSegmentUrl": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/prod/resource-manager/small/57/159/4873/model_11131_20240429141943/pv.mp4",
      "Resolution": "1920x1080",
      "SupportAnchorExtraParams": [],
      "SupportDriverTypes": [
        "Text",
        "OriginalVoice",
        "ModulatedVoice"
      ]
    }
  ]
}
```

```
    ],  
    "VideoDuration": 0,  
    "VideoNum": 0,  
    "VirtualmanKey": "fe880d9385304e688f9a8e469370fc90",  
    "VirtualmanType": "2D小样本(专属)",  
    "VirtualmanTypeCode": "small_sample_2d",  
    "PurchaseType": "Custom"  
  }  
]  
}  
}
```

查询形象支持的动作列表

最近更新时间：2024-11-19 11:30:22

接口说明

根据数智人VirtualmanKey，查询形象支持的动作列表。

 **注意：**
该接口只支持查询3d形象的动作。2d小样本形象不支持动作，调用该接口返回为空。

调用协议

HTTPS + JSON
POST /v2/ivh/crmserver/customerassetsservice/getaction
Header Content-Type: application/json;charset=utf-8

请求参数

参数	类型	必须	说明
VirtualmanKey	string	是	已分配的数智人形象 Virtualmankey

返回参数

参数	类型	必须	说明
Actions	Array of [Action]	否	动作列表  说明： 当该形象无支持动作、或 VirtualmanKey 不存在时，动作列表返回为空。

Action

参数	类型	必须	说明
ActionName	string	是	数智人动作名称
ActionCode	string	是	数智人动作 Code，对应 ssml 动作标签
ActionSample	string	是	数智人动作样例 URL

请求示例

```
{
  "Header": {},
  "Payload": {
    "VirtualmanKey": "8af9c7a960ab44c9bfe8a48deae6a250"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "28c713ede08f42afc058e4c57d7ecdc0",
  }
}
```

```
"SessionID": "gzfa9105ca17313027784787766",
"DialogID": "",
"Code": 0,
"Message": "ok"
},
"Payload": {
  "Actions": [
    {
      "ActionCode": "2hands_forward1",
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/2hands_forward1.mp4",
      "ActionName": "双手指向前方1"
    },
    {
      "ActionCode": "admission",
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/admission.mp4",
      "ActionName": "入场"
    },
    {
      "ActionCode": "appearance",
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/appearance.mp4",
      "ActionName": "出场"
    },
    {
      "ActionCode": "bow",
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/bow.mp4",
      "ActionName": "弯腰鞠躬"
    },
    {
      "ActionCode": "waving_hand1",
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/waving_hand1.mp4",
      "ActionName": "单手挥手"
    },
    {
      "ActionCode": "welcome",
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/welcome.mp4",
      "ActionName": "欢迎"
    },
    {
      "ActionCode": "left_across_waist",
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/resource-
manager/3d_virtualman/meta_yoyo/2023/display_videos/left_across_waist.mp4",
      "ActionName": "左手体侧滑动"
    },
    {
      "ActionCode": "left_side1",
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/left_side1.mp4",
      "ActionName": "左手指向左侧1"
    },
    {
      "ActionCode": "left_up2",
```

```
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/left_up2.mp4",  
      "ActionName": "左手指向左上2"  
    },  
    {  
      "ActionCode": "right_side1",  
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/right_side1.mp4",  
      "ActionName": "右手指向右侧1"  
    },  
    {  
      "ActionCode": "right_slide_down",  
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/resource-  
manager/3d_virtualman/meta_yoyo/2023/display_videos/right_slide_down.mp4",  
      "ActionName": "右手体侧向下滑动"  
    },  
    {  
      "ActionCode": "right_up2",  
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/right_up2.mp4",  
      "ActionName": "右手指向右上2"  
    },  
    {  
      "ActionCode": "applaud",  
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/applaud.mp4",  
      "ActionName": "鼓掌"  
    },  
    {  
      "ActionCode": "compliment_state3",  
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/resource-  
manager/3d_virtualman/meta_yoyo/2023/display_videos/compliment_state3.mp4",  
      "ActionName": "双手点赞"  
    },  
    {  
      "ActionCode": "ok2",  
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/ok2.mp4",  
      "ActionName": "体侧比OK"  
    },  
    {  
      "ActionCode": "show_love3",  
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/show_love3.mp4",  
      "ActionName": "单手比心"  
    },  
    {  
      "ActionCode": "support",  
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/support.mp4",  
      "ActionName": "加油"  
    },  
    {  
      "ActionCode": "turn_around",  
      "ActionSample": "https://virtualhuman-cos-test-1251316161.cos.ap-  
nanjing.myqcloud.com/resource-manager/3d_virtualman/meta_yoyo/2023/display_videos/turn_around.mp4",  
      "ActionName": "转圈"
```



```
    }  
  ]  
}  

```

附录1—服务资产类型

最近更新时间：2025-04-08 14:05:52

参数	说明
ImageAsset	2D小样本（专属）形象定制配额
InteractConcurrency	2D小样本（专属）交互并发
BroadcastHour	2D小样本（专属）播报小时包
BroadcastConcurrency	2D小样本（专属）播报并发
SmallSampleGeneral2DImageAsset	2D小样本（通用）形象定制配额
SmallSampleGeneral2DInteractConcurrency	2D小样本（通用）交互并发
SmallSampleGeneral2DBroadcastHour	2D小样本（通用）播报小时包
SmallSampleGeneral2DBroadcastConcurrency	2D小样本（通用）播报并发
SmallSample4K2DImageAsset	2D小样本（高精）形象定制配额
SmallSample4K2DInteractConcurrency	2D小样本（高精）交互并发
SmallSample4K2DBroadcastHour	2D小样本（高精）播报小时包
SmallSample4K2DBroadcastConcurrency	2D小样本（高精）播报并发
SmallSamplePhoto2DImageAsset	2D小样本（照片）形象定制配额
SmallSamplePhoto2DInteractConcurrency	2D小样本（照片）交互并发
SmallSamplePhoto2DBroadcastHour	2D小样本（照片）播报小时包
SmallSamplePhoto2DBroadcastConcurrency	2D小样本（照片）播报并发
NoTrainVideoBroadcastHour	2D免训练播报小时包（视频素材）
NoTrainVideoBroadcastConcurrency	2D免训练播报并发（视频素材）
RealMan2DInteractConcurrency	2D精品交互并发
RealMan2DBroadcastHour	2D精品播报小时包
RealMan2DBroadcastConcurrency	2D精品播报并发
RealMan3DInteractConcurrency	3D写实交互并发
RealMan3DBroadcastHour	3D写实播报小时包
RealMan3DBroadcastConcurrency	3D写实播报并发
HalfRealistic3DInteractConcurrency	3D半写实交互并发
HalfRealistic3DBroadcastHour	3D半写实播报小时包
HalfRealistic3DBroadcastConcurrency	3D半写实播报并发
Cartoon3DInteractConcurrency	3D卡通交互并发
Cartoon3DBroadcastHour	3D卡通播报小时包
Cartoon3DBroadcastConcurrency	3D卡通播报并发
TimbreAsset	声音复刻（基础版）定制配额

ZeroShotTimbreAsset	声音复刻（极速版）定制配额
TimbreHour	音色小时包
SmallSampleImageRenewalAsset	2D形象-专属口型在架服务配额
SmallSampleGeneralImageRenewalAsset	2D形象-通用口型在架服务配额
RealMan3DImageRenewalAsset	3D形象在架服务配额
TimbreRenewalAsset	音色定制在架服务配额

附录2-情感风格

最近更新时间：2024-08-23 10:02:15

参数	说明
0	高兴
1	新闻
2	客服
3	恐惧
4	诗歌
5	故事
6	中性
7	悲伤
8	广播
9	生气
tc_amaze	震惊
tc_angry	生气
tc_aojiao	傲娇
tc_call	客服
tc_disgusted	厌恶
tc_exciting	兴奋
tc_fear	恐惧
tc_happy	高兴
tc_jieshuo	解说
tc_neutral	中性
tc_news	新闻
tc_peaceful	平静
tc_poetry	诗歌
tc_radio	广播
tc_sad	悲伤
tc_sajiao	撒娇
tc_story	故事

附录3-数智人类型

最近更新时间：2024-08-23 18:01:11

参数	说明
real_man_2d	2D精品
real_man_3d	3D写实
half_realistic_3d	3D半写实
cartoon_3d	3D卡通
small_sample_2d	2D小样本（专属）
small_sample_general_2d	2D小样本（通用）
small_sample_4k_2d	2D小样本（高精）
small_sample_photo_2d	2D小样本（照片）

附录4–语种列表

最近更新时间：2025-03-12 17:03:53

参数	说明
am-ET	阿姆哈拉语(埃塞俄比亚)
ar-EG	阿拉伯语(埃及)
ar-SA	阿拉伯语(沙特阿拉伯)
bg-BG	保加利亚语(保加利亚)
bs-BA	波斯尼亚语(波斯尼亚和黑塞哥维那)
ca-ES	加泰罗尼亚语
cs-CZ	捷克语(捷克)
da-DK	丹麦语(丹麦)
de-AT	德语(奥地利)
de-CH	德语(瑞士)
de-DE	德语(德国)
el-GR	希腊语(希腊)
en-AU	英语(澳大利亚)
en-GB	英语(英国)
en-IE	英语(爱尔兰)
en-IN	英语(印度)
en-US	英语(美国)
es-ES	西班牙语(西班牙)
es-MX	西班牙语(墨西哥)
et-EE	爱沙尼亚语(爱沙尼亚)
eu-ES	巴斯克语
fi-FI	芬兰语(芬兰)
fil-PH	菲律宾语(菲律宾)
fr-BE	法语(比利时)
fr-CH	法语(瑞士)
fr-FR	法语(法国)
gl-ES	加利西亚语
he-IL	希伯来语(以色列)
hr-HR	克罗地亚语(克罗地亚)
hu-HU	匈牙利语(匈牙利)
id-ID	印度尼西亚语(印度尼西亚)

it-IT	意大利语(意大利)
ja-JP	日语(日本)
jv-ID	爪哇语(拉丁语、印度尼西亚)
ko-KR	韩语(韩国)
mk-MK	马其顿语(北马其顿)
ms-MY	马来语(马来西亚)
nb-NO	书面挪威语(挪威)
ne-NP	尼泊尔语(尼泊尔)
nl-BE	荷兰语(比利时)
nl-NL	荷兰语(荷兰)
pl-PL	波兰语(波兰)
ps-AF	普什图语(阿富汗)
pt-BR	葡萄牙语(巴西)
pt-PT	葡萄牙语(葡萄牙)
ro-RO	罗马尼亚语(罗马尼亚)
sk-SK	斯洛伐克语(斯洛伐克)
sl-SI	斯洛文尼亚语(斯洛文尼亚)
sv-SE	瑞典语(瑞典)
sw-KE	斯瓦希里语(肯尼亚)
th-TH	泰语(泰国)
tr-TR	土耳其语(土耳其)
uk-UA	乌克兰语(乌克兰)
vi-VN	越南语(越南)
yue-CN	中文(粤语，简体)
zh-CN	中文(普通话，简体)
zh-HK	中文(粤语，繁体)
zh-TW	中文(台湾普通话，繁体)
zu-ZA	祖鲁语(南非)

接口对接常见问题

最近更新时间：2024-08-23 10:29:43

问题描述	解决方式	备注
个人资产接口是否支持查询已制作视频记录明细	暂不支持，当前支持查询总视频个数和总耗时小时数。	-

配置端 iPaaS 接入

产品介绍

最近更新时间：2025-05-26 10:37:02

适用场景

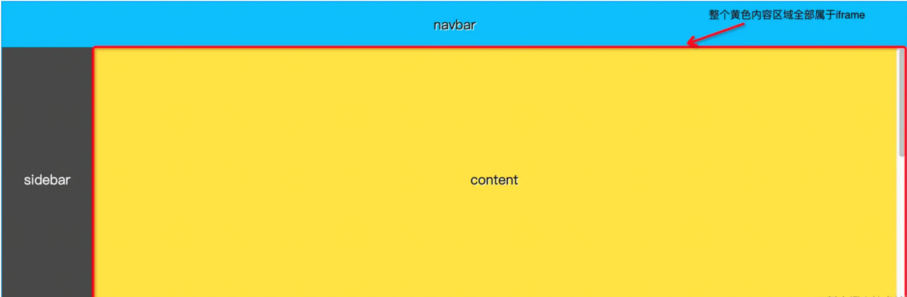
iPaaS（Integration Platform as a Service）的模式，集成配置端（数智人平台）页面并进行主题色、品牌 logo 的配置后，即可在任一业务系统内直接享受到与原生数智人平台一致的云服务。

适合接入方

- 业务诉求：希望集成数智人平台核心配置能力、快速搭建业务系统的客户：
 - 如，软件服务渠道商。通过 iPaaS 集成数智人平台模式，降低各级集成商接入的门槛，扶持更多伙伴快速起量；
 - 如，需要基于客户内部业务平台去调用数智人服务的 ISV 集成方。客户无法接受独立系统或单点登录，期望深度集成融合到自己的管理系统中使用，则可通过 iPaaS 集成数智人平台模式，快速在客户已有系统中新增数智人服务模块。
- 研发能力：有前端集成能力。

支持特性

被集成的数智人平台页面可直接满足以下特性：

序号	特性	具体描述
1	新页面可零/低成本嵌入	嵌入框架搭建好后，后续新增页面可以快速嵌入，无需引发增量研发工作
2	页面内更新无感知	客户接入只需要验证 url 跳转，每次页面内功能/UI 升级时 测试不需要全量回归验证，页面内功能升级可无感知
3	iframe 方式嵌入	对指定的页面范围，能提供除了菜单导航外页面内容的 iframe 嵌入方式 
4	iframe 自适应	可支持传入嵌入范围的宽和高，iframe 能自适应展示，且尽量保证： ①弹窗居中； ②嵌入部分不出现滚动条； ③浮层弹窗等能在最上层弹出
5	跨域接入	可支持客户跨域接入，无论客户平台使用什么网关和账号体系： ① 客户账号租户：支持客户通过映射表或类似方式，将自己的账号体系转换为腾讯云 UIN，能以不同租户的身份来访问平台嵌入页面； ② 主子账号：支持客户维护腾讯云父子账号租户关系，若客户指定了主子账号关系，能实现在嵌入页面内操作主子账号的资产分配功能
6	每客户唯一 license+短链	即A客户的 license 仅限A 且仅限A唯一对接嵌入，如果A将接入链接给到B，B客户无法接入或者接入后A无法接入
7	页面前进后退+跳转	支持通过浏览器内的 go back 返回前序页面； 支持通过客户传入的 postMessage 跳转到指定页面；
8	主题色级别的换肤	可支持一个主题对应一个配置文件，可通过客户写入配置文件的方式来实现主题风格的变化

9	logo 等图片资源的配置	可支持一个主题对应一个配置文件，可通过客户写入配置文件的方式来实现 logo 图片的变化
10	嵌入页面的配置	可指定嵌入页面范围

接入流程

1. 与商务经理沟通下单 iPaaS 服务（5万元/年）；
2. 由交付架构师支持接入。提供您的业务回调信息到交付架构师处，进行接入授权；
3. 参考调试工具确认接入页面范围和配置项；
4. 参考接入规范完成接入（一般情况“天”级别即可完成）。

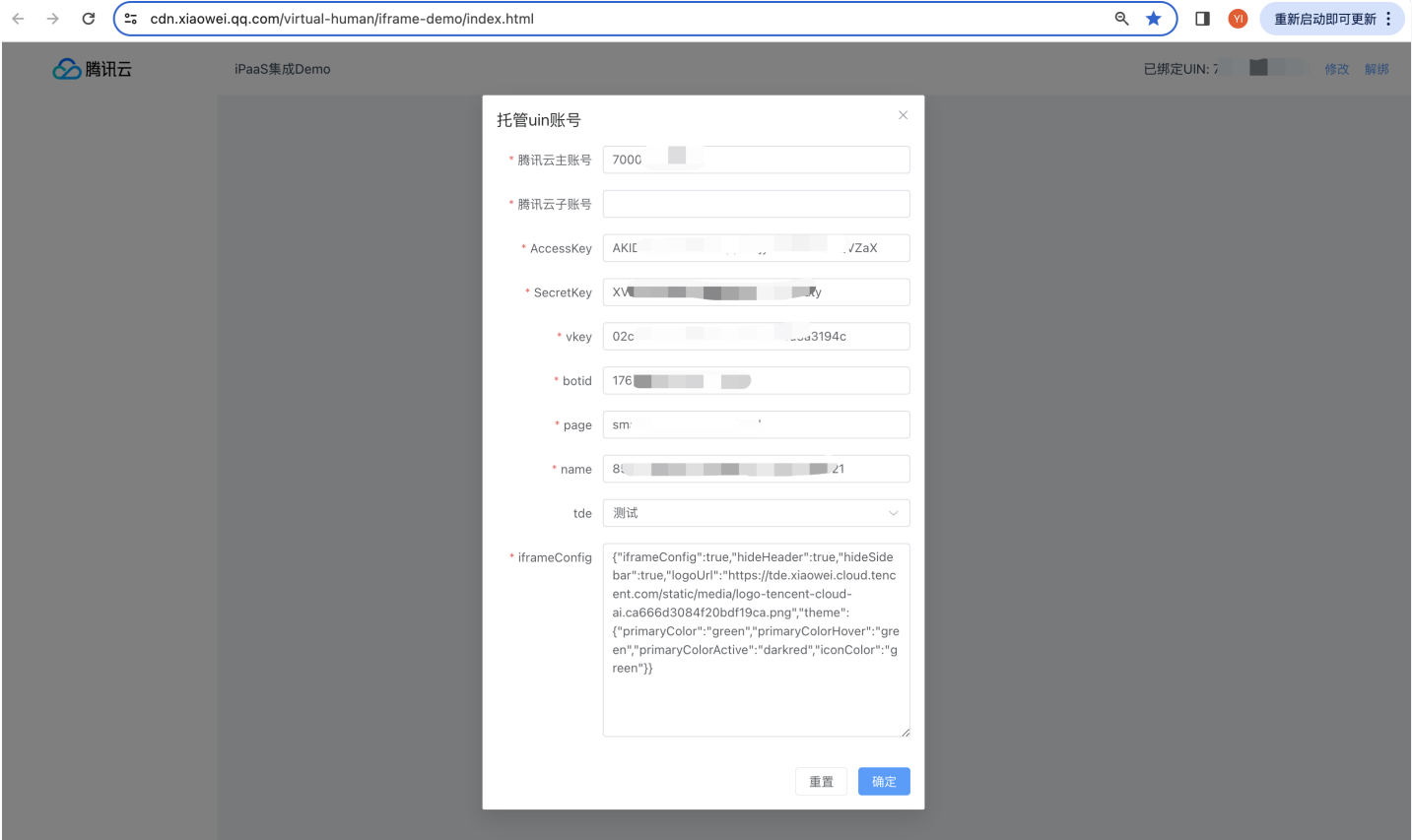
调试工具

为集成方提供了集成 Demo 参考，可直观理解各子页面集成后的展示效果，以及更改主题色、logo配置后的展示效果。

!

Demo示例:

<https://cdn.xiaowei.qq.com/virtual-human/iframe-demo/index.html>



参数说明：

vkey 数智人项目的 vkey，botid 为客户托管账号绑定的腾讯云智能体开发平台机器人 id，page 播报页面类型，name 播报数智人 name 参数。

腾讯云

iPaaS集成Demo

已绑定UIN: 700000446505

修改

解绑

音视频播报列表

音视频制作页面-编辑

音视频制作页面-新建

会话互动列表

交互大模型页面

运营管理

资产管理

< 音视频制作

测试

全身站姿 / 个人服饰

女声7

播报内容

造型

声音

输出设置

文本驱动

语音驱动

音色

全部

定制

通用

多情感

客服

资讯

对话

朗读

特色

外语

方言

御姐

云羽 (旧)

叮当智能屏音色

叮当

温柔磁性

new女声1new

柔和甜美

new女声2new

慢条斯理

new女声3new

自然大方

new女声5new

大方稳健

new女声6new

热情亲切

new女声7new

萝莉

云萝

成熟女声

云叶

温柔磁性

女声1

柔和甜美

女声2

慢条斯理

女声3

柔和自然

女声4

自然大方

女声5

大方稳健

女声6

热情亲切

女声7

成熟柔和

女声8

成熟稳重

女声9

活力天真

new女声10new

0.5倍

0.75倍

1.0倍

1.25倍

1.5倍

生成音频

生成视频

最近更新时间: 2025-05-26 20:04:21

1. 云账号托管能力建设:

- ```
%7B%22version%22%3A%222.0%22%2C%22statement%22%3A%5B%7B%22effect%22%3A%22allow%22%2C%22action%22%3A%5B%22ivh%3A%*%22%5D%2C%22resource%22%3A%5B%22%*%22%5D%7D%5D%7D
```





子账号 aksk 获取路径截图

- 第三步：集成方按照接口规范提供一个公网 api，作为数智人校验 token（即上一步中提到的 token）的回调接口。接口规范详见：[token 校验回调接口规范](#)，可邮件发给我们，也可通过商务同学传递。**提示：**merchantID（即客户标识）在未获取的情况下，可以先自定义个值，等数智人方提供后，替换即可。

**说明：**  
第一步和第三方操作，可发邮件至：[txc\\_avatar@tencent.com](mailto:txc_avatar@tencent.com)（主题注明“ipaas 接入申请”，邮件内容主要提供两个信息：“1、UIN：xxxxxx，2、token 校验回调接口：<https://xxxxx>”），我们会在邮件回复中提供“客户标识”值。

- 第四步：参照 [产品介绍](#) 中的 demo 来构建数据的“信息录入交互页面”。数据信息包括：UIN、vkey、botid、page、name等等。

**注意：**  
demo 中的 AccessKey 和 SecretKey 为永久密钥，可根据实际情况调整为不需要。

2. 数智人平台菜单框架开发及集成：

开发者通过 iframe 集成，使用下面各个菜单页的 URL 在客户业务系统中嵌入数智人平台页面。

菜单各个页面的拼接规则为：**主域名+路由地址**

主域名为：<https://xiaowei.cloud.tencent.com>，路由地址详见下文中的表格清单。

示例：

“音视频播报列表”页面，其路由地址为：`/ivh/#/application?iframe=true&from=xxx&token=xxx`

则该页面的访问地址URL：`https://xiaowei.cloud.tencent.com/ivh/#/application?iframe=true&from=xxx&token=xxx`

统一参数：

from 指客户标识，由数智人侧分配，详见 [第一步：准备工作](#)；

token 由集成商生成，背后对应客户的 uin、aksk，当访问页面时数智人侧会回调接口通过 token 换取 uin 和 aksk，详见上文。

目前数智人平台支持以下菜单页面的嵌入（持续更新中），完整平台功能范围见 [数智人平台操作指引](#)。

| 菜单页面名称     | 菜单页面路由地址                                                                                           | 备注                                                     |
|------------|----------------------------------------------------------------------------------------------------|--------------------------------------------------------|
| 音视频播报列表    | <code>/ivh/#/application?iframe=true&amp;from=xxx&amp;token=xxx</code>                             | —                                                      |
| 音视频制作页面-编辑 | <code>/ivh/#/bvh_page?<br/>page=xxx&amp;name=xxx&amp;iframe=true&amp;from=xxx&amp;token=xxx</code> | page 为制作视频所选用的数智人类型，详见下文的数智人类型；<br>name 为播报视频id，详见下文图示 |

|            |                                                                                                                             |                                                                    |
|------------|-----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------|
| 音视频制作页面-新建 | <code>/ivh/#/bvh_page? page=xxx&amp;iframe=true&amp;from=xxx&amp;token=xxx</code>                                           | page 为制作视频所选用的数智人类型，详见下文的数智人类型                                     |
| 会话互动列表     | <code>/ivh/#/application?iframe=true&amp;from=xxx&amp;token=xxx</code>                                                      | -                                                                  |
| 交互大模型页面    | <code>/ivh/#/lke/robot-manage/knowledge/qa-source? vkey=xxx&amp;botid=xxx&amp;iframe=true&amp;from=xxx&amp;token=xxx</code> | vkey 为数智人项目的 virtualmanProjectId，botid 为客户托管账号绑定的为 智能体开发平台 机器人 id; |
| 运营管理       | <code>/ivh/#/analysis?iframe=true&amp;from=xxx&amp;token=xxx</code>                                                         | -                                                                  |
| 资产管理       | <code>/ivh/#/asserts_management? iframe=true&amp;from=xxx&amp;token=xxx</code>                                              | -                                                                  |

数智人类型代号

| 数智人类型                   |           |
|-------------------------|-----------|
| real_man_2d             | 2D精品      |
| small_sample_2d         | 2D小样本（专属） |
| half_realistic_3d       | 3D半写实     |
| cartoon_3d              | 3D卡通      |
| real_man_3d             | 3D写实      |
| small_sample_general_2d | 2D小样本（通用） |
| small_sample_photo_2d   | 2D小样本（照片） |
| noTrain                 | 免训练视频     |

腾讯云智能 | 数智人平台 首页 形象设定 场景应用 运营管理分析 文档

智人场景应用

新建或编辑数智人场景应用项目

鼠标放在此处，点击“复制视频ID”后，获取视频ID

会话互动 音视频播报

新建音视频播报项目

4月28日2007 哟哟3  
2025/04/28 20:07

4月28日1957 哟哟3  
2025/04/28 19:57

4月17日1539 哟哟3  
2025/04/17 15:39

复制视频ID

4月17日1511 哟哟3  
2025/04/17 15:11

4月17日1456 哟哟3  
2025/04/17 14:56

4月17日1151 叮当  
2025/04/17 11:51

4月17日1143 艾云  
2025/04/17 11:43

4月17日1133 瑶光  
2025/04/17 11:33

4月17日1130 小微  
2025/04/17 11:30

草稿

草稿

草稿

获取播报视频id



### 3. token 回调检查接口:

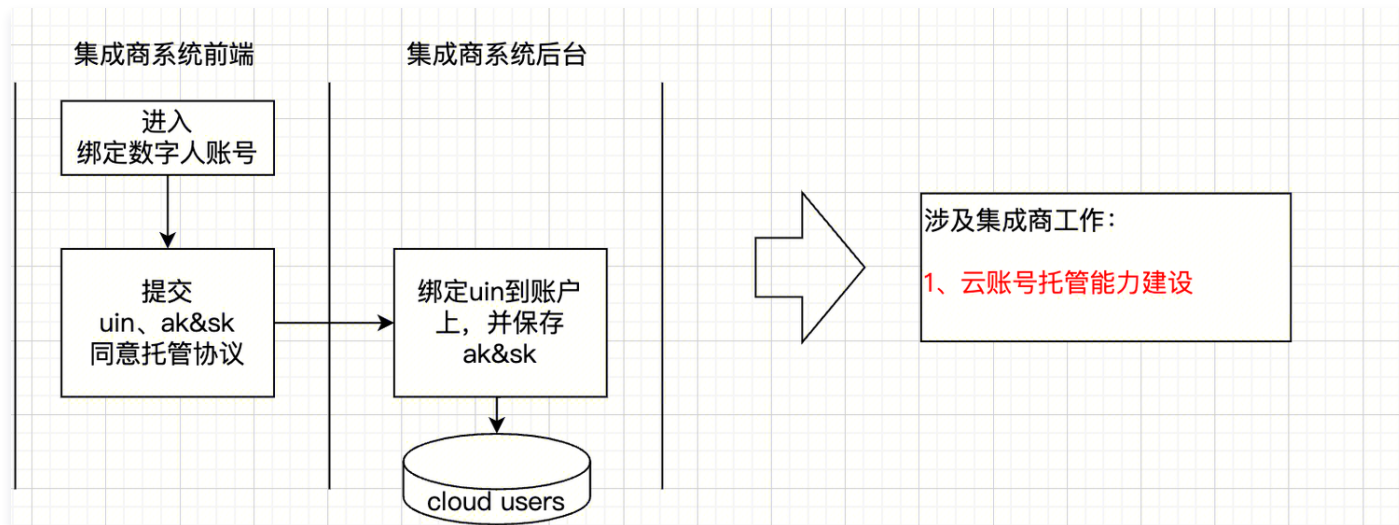
数字人进行 token 回调, 返回客户托管的 aksk 信息;

当数智人页面打开时, 会拿到 url 上的带参 token, 回调客户系统检查 token 合法有效性, 并希望客户返回该 token 对应的客户已托管的 uin、ak、sk 信息, 以便后续能访问内嵌的数智人平台页面xxxx。

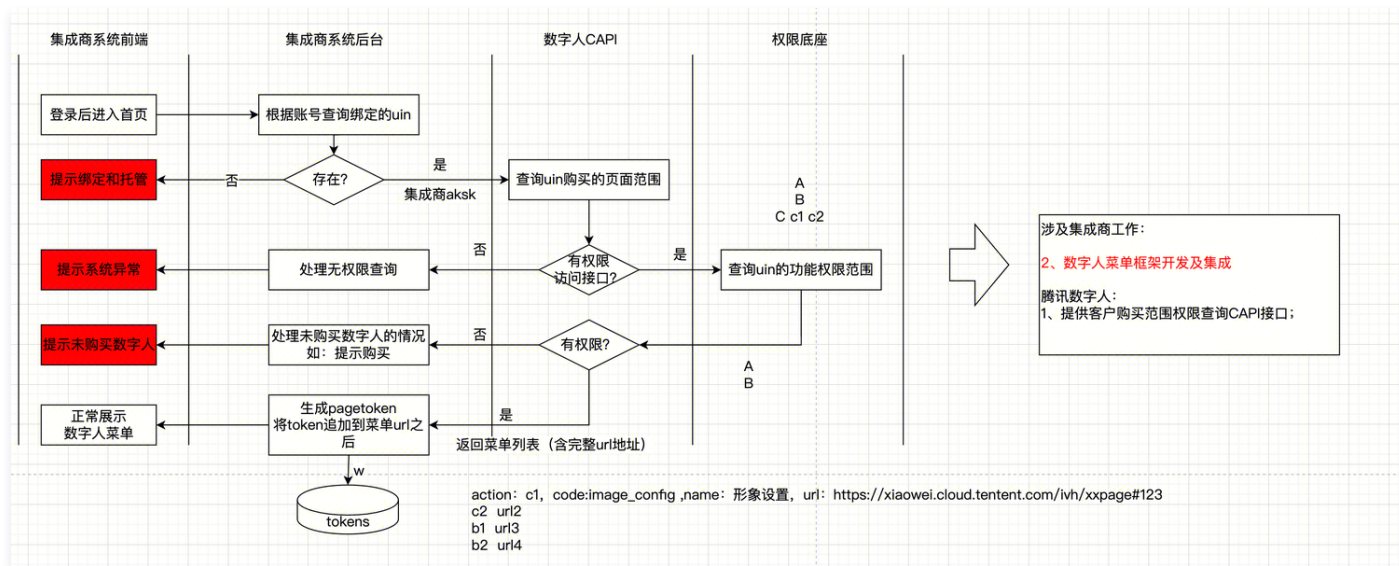
若上文第一部分的第四步已经完成, 此处可忽略。

## 二、集成过程说明

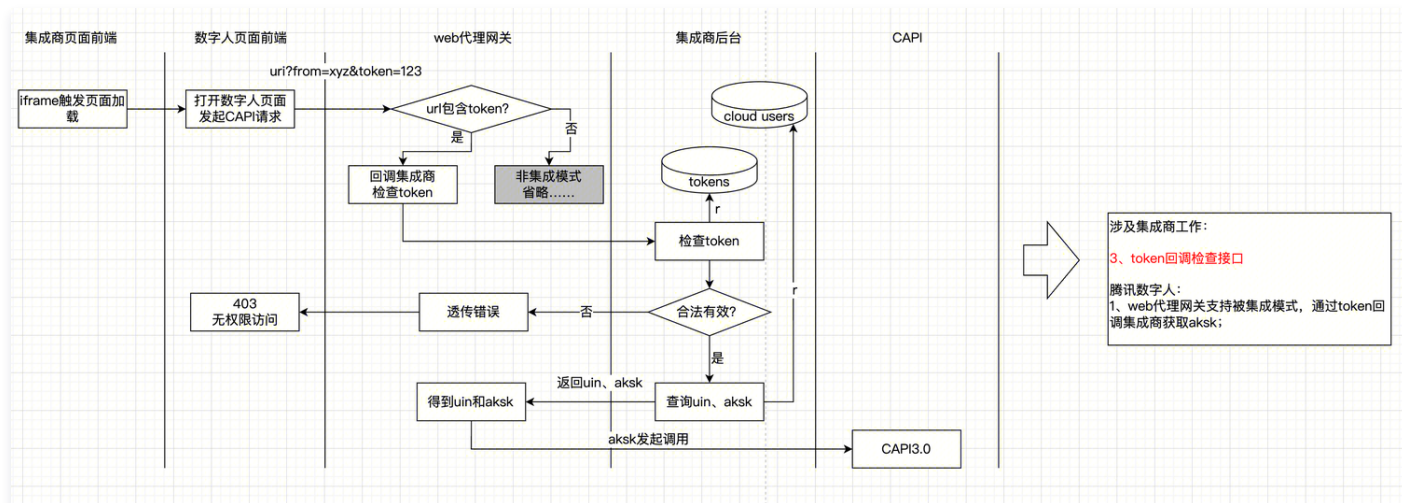
### 1. 客户 uin、aksk 托管。



### 2. 集成商系统主框架加载【流程图简化, 只区分腾讯 vs 客户的鉴权和查询】



### 3. 集成商系统内打开数字人页面



## 三、token 校验回调接口规范

协议: https

请求method: POST

请求content-type: application/json

请求body:

```

{
 "token": "token", // 集成商生成, 在页面嵌入时在url中追加的token参数。
}

```

返回结构要求

```

{
 "request_id": "request_id", // 唯一请求id
 "code": 0,
 "message": "ok",
 "merchantID": "merchantID", // 集成商id, 腾讯分配, 在页面嵌入时在url中追加的from参数。
 "merchantUin": "merchantUin", // 托管的客户uin
 "MerchantSubAccountUin": "MerchantSubAccountUin", // 托管的客户子账号uin
 "sessionToken": {
 "secretId": "secretId", // 秘钥id, 建议使用临时秘钥返回
 "secretKey": "secretKey", // 秘钥key, 建议使用临时秘钥返回
 }
}

```

接口开发完成后, 提交给厂商的产品运营人员, 通过上文中的第四步所开发的“信息录入交互页面”进行相关配置。

## 四、页面样式规范

### 可配置样式范围

针对 iframe 嵌入的页面, 目前已支持一键更换主题色、logo 配置的功能。

### 配置方式

具体能力详见下面的配置内容:

```

{
 "iframeConfig": true, // 表示当前是iframe潜入, 并且支持换肤能力
}

```



```
"hideHeader": true, // 隐藏数智人平台的头部内容
"hideSidebar": true, // 隐藏数智人平台的左侧菜单内容
"logoUrl": "/logo.png", // 更换数智人头部的logo信息
// 页面相关的tab可以通过权限控制隐藏显示。
// 主题色配置
"theme": {
 "primaryColor": "green",
 "primaryColorHover": "green",
 "primaryColorActive": "darkred",
 "iconColor": "green"
 //其他antd支持的主题样式配置。
}
}
```