

数字内容生产线

产品介绍

文档版本 01
发布日期 2025-01-14



版权所有 © 华为云计算技术有限公司 2025。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为云计算技术有限公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为云计算技术有限公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为云计算技术有限公司

地址：贵州省贵安新区黔中大道交兴功路华为云数据中心 邮编：550029

网址：<https://www.huaweicloud.com/>

目录

1 什么是数字内容生产线.....	1
2 功能特性.....	2
3 产品优势.....	4
4 使用限制.....	6
4.1 区域限制.....	6
4.2 预置模特使用限制.....	6
4.3 API 使用限制.....	6
4.4 形象制作限制.....	15
4.5 声音制作限制.....	15
4.6 视频制作限制.....	16
4.7 视频直播限制.....	18
4.8 Flexus 分身数字人限制.....	21
4.9 照片数字人限制.....	22
4.10 智能交互限制.....	22
4.11 资产管理说明.....	23
4.12 SP 用户限制.....	25
5 常用概念.....	26
6 计费说明.....	28
7 权限管理.....	35
8 算法备案公示.....	37

1 什么是数字内容生产线

数字内容生产线（MetaStudio）解决方案，提供数字人视频与直播、智能交互等服务，使能千行百业场景应用。

应用场景

教育：数字人课件制作、虚拟老师教学。

政府及公共事业：数字人代言、数字人宣讲、数字人会议等。

医疗健康：智能医疗客服、虚拟医生培训等。

金融：数字人智能客服、数字人企业代言、数字人培训视频制作等。

文旅：智能客服、数字人导览、数字人讲解等。

广电传媒：虚拟综艺主持人、虚拟新闻主播等。

互娱电商：数字人文娱直播、数字人短视频制作、电商直播等。

2 功能特性

数字内容生产线Metastudio通过AI学习真人形象和声音，来生成数字人模型，相应功能为形象制作、声音制作和Flexus分身数字人。再将模型用于音视频内容的创作，相应功能为视频制作、视频直播和智能交互，主要是生成画面内容。也可以直接通过静态的人脸照片生成视频，对应功能为照片数字人。

功能详情，如表2-1所示。

表 2-1 功能特性

功能名称	功能描述
形象制作	<p>拍摄真人视频，建模生成数字人形象。</p> <p>数字人形象可用于生成音视频内容，应用于视频制作、视频直播和智能交互。</p> <p>须知</p> <ul style="list-style-type: none">制作好的数字人模型不支持下载或导出至本地。数字人模型不是通用模型，不与第三方服务兼容。用户在MetaStudio制作的模型，仅支持在本服务内使用。
Flexus分身数字人	<p>分身数字人Flexus版针对中小企业客户，提供1080P的实景数字人、1分钟语料声音克隆能力和120分钟1080P视频制作能力，帮助客户通过手机拍摄即可轻松制作数字人。Flexus分身数字人不仅简化制作流程，还显著降低技术门槛，且购买简单，为中小企业提供超简单、超高性价比数字人。</p> <p>须知</p> <p>Flexus版视频制作和标准版视频制作的资源相互独立，各自计费。</p>
声音制作	<p>录制真人声音，建模生成声音模型。</p> <p>声音模型可为数字人配音，生成音视频内容，应用于视频制作、视频直播和智能交互，或生成数字人名片和照片数字人。</p> <p>须知</p> <ul style="list-style-type: none">制作好的声音模型不支持下载或导出至本地。声音模型不是通用模型，不与第三方服务兼容。用户在MetaStudio制作的模型，仅支持在本服务内使用。

功能名称	功能描述
视频制作	使用系统预置或自定义生成的数字人形象、声音，制作音视频内容。制作的视频可应用于教学、培训等各类场景中。
视频直播	使用系统预置或自定义的数字人形象、声音，制作音视频直播内容。可通过下述方式，在直播平台开播。 RTMP直推方式，即华为云MetaStudio直接把画面传输给第三方直播平台，需从直播平台获取推流地址。如果直播平台无法提供推流地址，可通过窗口捕获方式开播。
智能交互	对接第三方大脑，通过数字人交互的方式与用户进行问答。场景应用广泛，包括智能导购、文旅导览、智能问答、营业厅客服等。 注意 数字人回答的内容，是由对接的第三方大模型或知识库返回的答案决定的。
资产管理	<ul style="list-style-type: none">支持从本地上传模型、PPT、动画、素材、视频、场景、图片、音乐等资产。支持转移声音和模型资产给其他租户使用。此功能暂未对外开放，如需使用，需提交工单申请。支持删除资产。
数字人名片	支持为分身数字人制作形象名片。通过名片可以查看数字人形象和说话效果。
照片数字人	支持通过静态的人脸照片生成视频。
SP用户	MetaStudio支持SP模式，此模式对大客户开放，用于分销场景。SP用户支持购买大颗粒度资源包，并分配资源给关联租户使用。

3 产品优势

控制台自助服务，一站式快速生成所需内容

用户申请账号并上传相关授权，拍摄训练所需音视频素材，上传进行模型训练，生成自定义形象和声音。选择背景、声音、模特等内容，基于文本或语音智能驱动，实现视频制作、视频直播、智能交互等能力。

华为云盘古数字人大模型，赋能千行百业数字化营销新模式

MetaStudio服务依托华为云基础设施、海量算力（CPU/GPU/NPU）、全球一张网（算网融合、超低时延），通过华为云盘古数字人大模型，训练生成数字人、数字物、数字空间，实现数字人视频制作、直播、智能交互和数字人IP代言等场景。

数字人升级迭代快，提升产品竞争力

- 口型精准：口型匹配度>95%，多语种无缝切换，泛化能力强。
- 表情自然：AI眼神纠正，持续眼神交流，细节更丰富。
- 图像真实：智能重打光，背景融合更真实。
- 多种复杂场景建模：支持人物走动、侧身等训练，肢体动作更自然。
- 多语言泛化：一次母语训练，多语种支持；支持20+语种，覆盖主流语言。

相对真人成本下降 90%以上

- 一次拍摄与训练后，可无限次使用；无时间约束、无时长限制、无容量限制；统一质量，稳定性高；口型、动作匹配度高，无限接近真实。
- 仅分身数字人动作自由度受限于采集视频的动作范围，综合对比数字人比真人成本低、效率高，且质量稳定。

数字人直播更真实、更沉浸

- 支持持物直播、实景直播、弹幕智能互动、真人随时接管，让数字人直播更真实、更沉浸。
- 支持数字人直播矩阵式管理，高级策略配置管理直播更灵活。

实时智能交互，提升客户满意度和品牌形象

开放的Web SDK及API可快速集成到第三方业务系统，实现AI数字员工问答、服务、陪练等场景任务。

- 1.5秒超低时延交互
实时驱动极低延时，保障用户真实体验。
- 多模型数字人大脑
 - 华为云智能问答机器人CBS
 - 华为云盘古大模型
 - 科大讯飞星火交互大模型及AIUI
- 能力开放可集成
 - 智能交互端侧Web SDK
 - 形象驱动API
 - 数字人大脑API

数字人能力开放，多种合作模式灵活选择

- 实名安全认证后，可使用控制台在线自助完成内容制作。
- 集成数字人API、SDK至企业自建平台。
- 联合专业直播服务伙伴，定制场景化解决方案，实现数字人直播系统交付和运营专业支持。

帮助客户提升产品竞争力

- 大幅提升数字人生产效率
满足批量生成数字人诉求，平台支持几百到上千的并发模型训练任务。
- 大幅提升数字人生产质量
模型质量提升30%以上，数字人名片迅速引流新阵地直播产品，业务推广无后顾之忧。
- 解决产品端到端服务能力
- 可接入其他AIGC产品，打造AIGC数字人视频生成到剪辑端到端的解决方案能力。

多行业多场景广泛应用

MetaStudio数字人已广泛应用于企业代言、综艺主持、电商带货、新闻播报、培训讲师、文旅导览、营业厅客服、市场营销等场景。

安全稳定可靠

中国站、国际站（新加坡）、欧洲站（爱尔兰）公有云均已部署，支撑客户全球业务拓展。

4 使用限制

在使用数字内容生产线服务前，您需要了解本服务的使用限制信息。

4.1 区域限制

数字内容生产线服务目前只支持“华北-北京四”和“亚太-新加坡”区域。

4.2 预置模特使用限制

MetaStudio控制台预置模特的有效使用期为3年，在2027年的年中前用户可以放心使用。具体下线时间请关注MetaStudio服务，会提前3个月在现网开放下线公告。

4.3 API 使用限制

使用 API 调用声音的限制说明

用户使用视频直播或智能交互接口调用音色时，需要先确认能否在MetaStudio控制台的视频直播间编辑界面或智能交互对话设置页面，选到当前音色。

如果界面正常展示且可以选择此音色使用，说明接口可以正常调用此音色，否则接口无法调用。

API 流控

数字内容生产线服务对服务端API设置了调用次数限制，避免出现短时间内重复调用API，服务中断的情况。

须知

如果调用API时，返回“APIGW.0308”错误信息，是触发了API流控限制。需要参考下述表格中的流控值，调整调用API的间隔时间，控制请求的发送速度；或者等待一段时间后，再发送API请求。

“APIGW”开头的错误码列表，详见[错误码说明](#)。

表 4-1 API 流控限制

接口分类	接口名称	单租户接口流控 (华北-北京四)	单租户接口流控 (华东-上海一)	单租户接口流控 (亚太-新加坡)
数字资产管理	创建资产	10次/秒	5次/秒	5次/秒
	查询资产列表	20次/秒	10次/秒	10次/秒
	更新资产	10次/秒	5次/秒	5次/秒
	删除资产	10次/秒	5次/秒	5次/秒
	查询资产详情	20次/秒	10次/秒	10次/秒
	恢复被删除的资产	10次/秒	5次/秒	5次/秒
	批量资产操作	1次/秒	1次/秒	1次/秒
	查询资产概要	20次/秒	10次/秒	10次/秒
	查询资产复制信息	20次/秒	10次/秒	10次/秒
	复制资产	4次/秒	3次/秒	3次/秒
文件管理	创建文件并获取上传URL	10次/秒	5次/秒	5次/秒
	删除文件	10次/秒	5次/秒	5次/秒
	确认文件已上传	10次/秒	5次/秒	5次/秒
	创建大文件	10次/秒	5次/秒	5次/秒
视频制作剧本管理	创建视频制作剧本	10次/秒	5次/秒	5次/秒
	查询视频制作剧本列表	20次/秒	10次/秒	10次/秒
	查询视频制作剧本详情	20次/秒	10次/秒	10次/秒
	更新视频制作剧本	10次/秒	5次/秒	5次/秒
	删除视频制作剧本	10次/秒	5次/秒	5次/秒
	复制视频制作剧本	10次/秒	5次/秒	5次/秒
数字人视频制作管理	查询视频制作任务列表	20次/秒	10次/秒	10次/秒
分身数字人视频制作管理	创建分身数字人视频制作任务	5次/秒	3次/秒	3次/秒
	查询分身数字人视频制作任务详情	20次/秒	10次/秒	10次/秒

接口分类	接口名称	单租户接口流控 (华北-北京四)	单租户接口流控 (华东-上海一)	单租户接口流控 (亚太-新加坡)
	取消等待中的分身数字人视频制作任务	5次/秒	3次/秒	3次/秒
数字人名片制作管理	创建数字人名片制作	5次/秒	3次/秒	3次/秒
	查询数字人名片制作任务列表	20次/秒	10次/秒	10次/秒
	查询数字人名片制作任务详情	20次/秒	10次/秒	10次/秒
	更新数字人名片制作	5次/秒	3次/秒	3次/秒
	删除数字人名片制作任务	5次/秒	3次/秒	3次/秒
照片数字人视频制作管理	创建照片分身数字人视频制作任务	5次/秒	3次/秒	3次/秒
	查询照片分身数字人视频制作任务详情	20次/秒	10次/秒	10次/秒
	取消等待中的照片分身数字人视频制作任务	5次/秒	3次/秒	3次/秒
	创建照片检测任务	5次/秒	3次/秒	3次/秒
	查询照片检测任务详情	20次/秒	10次/秒	10次/秒
分身数字人字幕文件生成管理	创建分身数字人视频字幕文件	5次/秒	3次/秒	3次/秒
	查询分身数字人视频字幕文件详情	20次/秒	10次/秒	10次/秒
智能直播间管理	创建智能直播间	5次/秒	3次/秒	3次/秒
	查询智能直播间列表	20次/秒	10次/秒	10次/秒
	查询智能直播剧本详情	20次/秒	10次/秒	10次/秒
	更新智能直播间信息	5次/秒	3次/秒	3次/秒

接口分类	接口名称	单租户接口流控 (华北-北京四)	单租户接口流控 (华东-上海一)	单租户接口流控 (亚太-新加坡)
	删除智能直播间	4次/秒	3次/秒	3次/秒
	创建智能直播间互动规则库	4次/秒	3次/秒	3次/秒
	查询智能直播间互动规则库列表	20次/秒	10次/秒	10次/秒
	更新智能直播间互动规则库	4次/秒	3次/秒	3次/秒
	删除智能直播间互动规则库	4次/秒	3次/秒	3次/秒
直播任务管理	启动数字人智能直播任务	4次/秒	3次/秒	3次/秒
	查询某个智能直播间下直播任务列表	20次/秒	10次/秒	10次/秒
	查询数字人智能直播任务详情	20次/秒	10次/秒	10次/秒
	结束数字人智能直播任务	4次/秒	3次/秒	3次/秒
	控制数字人直播过程	4次/秒	3次/秒	3次/秒
	查询租户所有数字人直播任务列表	10次/秒	5次/秒	5次/秒
	上报直播间事件	4次/秒	3次/秒	3次/秒
直播商品管理	创建商品	4次/秒	3次/秒	3次/秒
	查询商品列表	20次/秒	10次/秒	10次/秒
	更新商品	4次/秒	3次/秒	3次/秒
	查询商品详情	20次/秒	10次/秒	10次/秒
	删除商品	4次/秒	3次/秒	3次/秒
	商品资产组合配置	4次/秒	3次/秒	3次/秒
第三方直播平台管理	创建第三方直播平台	1次/秒	1次/秒	1次/秒
	更新第三方直播平台信息	1次/秒	1次/秒	1次/秒

接口分类	接口名称	单租户接口流控 (华北-北京四)	单租户接口流控 (华东-上海一)	单租户接口流控 (亚太-新加坡)
	删除第三方直播平台信息	1次/秒	1次/秒	1次/秒
	查询直播平台列表	20次/秒	10次/秒	10次/秒
	查询第三方直播平台信息	20次/秒	10次/秒	10次/秒
	查询第三方直播平台商品列表	5次/秒	3次/秒	3次/秒
智能交互数字人对话管理	创建智能交互对话	5次/秒	3次/秒	-
	查询智能交互对话列表	20次/秒	10次/秒	10次/秒
	查询智能交互对话详情	20次/秒	10次/秒	10次/秒
	更新智能交互对话信息	5次/秒	3次/秒	-
	删除智能交互对话	5次/秒	3次/秒	-
智能交互数字人对话任务管理	创建对话链接	5次/秒	3次/秒	-
	启动数字人智能交互任务	5次/秒	3次/秒	-
	结束数字人智能交互任务	5次/秒	3次/秒	-
	查询数字人智能交互任务	20次/秒	10次/秒	10次/秒
智能交互数字人应用管理	创建应用	5次/秒	3次/秒	-
	查询应用列表	20次/秒	10次/秒	10次/秒
	查询应用详情	20次/秒	10次/秒	10次/秒
	修改应用	5次/秒	10次/秒	-
	删除应用	5次/秒	3次/秒	-
智能交互数字人鉴权码管理	创建一次性鉴权码	20次/秒	10次/秒	-
智能交互数字人激活码管理	创建激活码	5次/秒	3次/秒	-
	查询激活码列表	20次/秒	10次/秒	10次/秒
	查询激活码详情	20次/秒	10次/秒	10次/秒

接口分类	接口名称	单租户接口流控 (华北-北京四)	单租户接口流控 (华东-上海一)	单租户接口流控 (亚太-新加坡)
	修改激活码	5次/秒	3次/秒	-
	重置激活码	5次/秒	3次/秒	-
	删除激活码	5次/秒	3次/秒	-
智能交互数字人知识库技能管理	创建知识库技能	4次/秒	3次/秒	3次/秒
	查询知识库技能列表	20次/秒	10次/秒	10次/秒
	修改知识库技能	4次/秒	3次/秒	3次/秒
	查询知识库技能详情	20次/秒	10次/秒	10次/秒
	删除知识库技能	4次/秒	3次/秒	3次/秒
	导出知识库技能	1次/秒	1次/秒	1次/秒
智能交互数字人知识库意图管理	创建知识库意图	4次/秒	3次/秒	3次/秒
	查询知识库意图列表	20次/秒	10次/秒	10次/秒
	查询知识库意图详情	20次/秒	10次/秒	10次/秒
	修改知识库意图	4次/秒	3次/秒	3次/秒
	删除知识库意图	4次/秒	3次/秒	3次/秒
	创建知识库意图和问法	4次/秒	3次/秒	3次/秒
智能交互数字人知识库问法管理	创建知识库问法	4次/秒	3次/秒	3次/秒
	查询知识库问法列表	20次/秒	10次/秒	10次/秒
	查询知识库问法详情	20次/秒	10次/秒	10次/秒
	修改知识库问法	4次/秒	3次/秒	3次/秒
	删除知识库问法	4次/秒	3次/秒	3次/秒
	批量创建知识库问法	1次/秒	1次/秒	1次/秒
	批量修改知识库问法	1次/秒	1次/秒	1次/秒

接口分类	接口名称	单租户接口流控 (华北-北京四)	单租户接口流控 (华东-上海一)	单租户接口流控 (亚太-新加坡)
智能交互数字人欢迎词管理	创建欢迎词	4次/秒	3次/秒	3次/秒
	查询欢迎词列表	20次/秒	10次/秒	10次/秒
	查询欢迎词详情	20次/秒	10次/秒	10次/秒
	修改欢迎词	4次/秒	3次/秒	3次/秒
	删除欢迎词	4次/秒	3次/秒	3次/秒
	修改欢迎词功能开关	4次/秒	3次/秒	3次/秒
	查询欢迎词功能开关	20次/秒	10次/秒	10次/秒
智能交互数字人热点问题管理	创建热点问题	4次/秒	3次/秒	3次/秒
	查询热点问题列表	20次/秒	10次/秒	10次/秒
	查询热点问题详情	20次/秒	10次/秒	10次/秒
	修改热点问题	4次/秒	3次/秒	3次/秒
	删除热点问题	4次/秒	3次/秒	3次/秒
智能交互数字人热词记录管理	创建热词记录	4次/秒	3次/秒	3次/秒
	查询热词记录列表	20次/秒	10次/秒	10次/秒
	查询配置热词记录详情	20次/秒	10次/秒	10次/秒
	修改热词记录	4次/秒	3次/秒	3次/秒
	删除热词记录	4次/秒	3次/秒	3次/秒
	修改热词功能开关	4次/秒	3次/秒	3次/秒
	查询热词功能开关	20次/秒	10次/秒	10次/秒
智能交互数字人委托管理	查询委托	20次/秒	10次/秒	10次/秒
	创建委托	1次/秒	1次/秒	-
	删除委托	1次/秒	1次/秒	-
分身形象制作管理	创建分身数字人模型训练任务	10次/秒	3次/秒	3次/秒
	查询分身数字人模型训练任务列表	20次/秒	10次/秒	10次/秒

接口分类	接口名称	单租户接口流控 (华北-北京四)	单租户接口流控 (华东-上海一)	单租户接口流控 (亚太-新加坡)
	查询分身数字人模型训练任务详情	20次/秒	10次/秒	10次/秒
	删除分身数字人模型训练任务	10次/秒	3次/秒	3次/秒
	更新分身数字人模型训练任务	10次/秒	3次/秒	3次/秒
	租户执行分身数字人模型训练任务命令	10次/秒	3次/秒	3次/秒
声音制作任务管理	查询用户配置的个性化音频时长	20次/秒	10次/秒	10次/秒
	创建基础版语音训练任务	10次/秒	3次/秒	3次/秒
	创建进阶版语音训练任务	10次/秒	3次/秒	3次/秒
	创建高级版语音训练任务	10次/秒	3次/秒	3次/秒
	查询语音训练任务列表	20次/秒	10次/秒	10次/秒
	提交语音训练任务	10次/秒	3次/秒	3次/秒
	查询语音训练任务详情	20次/秒	10次/秒	10次/秒
	删除语音训练任务	10次/秒	3次/秒	3次/秒
	查询任务操作日志	20次/秒	10次/秒	10次/秒
	获取语音文件上传地址	20次/秒	10次/秒	10次/秒
	获取语音训练任务审核结果	20次/秒	10次/秒	10次/秒
	设置任务批次	10次/秒	3次/秒	3次/秒
	确认在线录音结果	10次/秒	3次/秒	3次/秒
	获取在线录音确认结果	20次/秒	10次/秒	10次/秒
语音合成管理	创建TTS试听任务	5次/秒	3次/秒	3次/秒
	获取TTS试听文件	20次/秒	10次/秒	10次/秒

接口分类	接口名称	单租户接口流控 (华北-北京四)	单租户接口流控 (华东-上海一)	单租户接口流控 (亚太-新加坡)
	创建TTS异步任务	5次/秒	3次/秒	3次/秒
	获取TTS异步任务	20次/秒	10次/秒	10次/秒
	设置TTS租户级自定义读法配置	10次/秒	3次/秒	3次/秒
	获取TTS租户级自定义读法配置	10次/秒	3次/秒	3次/秒
	删除TTS租户级自定义读法配置	10次/秒	3次/秒	3次/秒
	修改TTS租户级自定义读法配置	10次/秒	3次/秒	3次/秒
3D数字人风格管理	查询数字人风格列表	20次/秒	10次/秒	-
3D数字人风格化照片建模	创建照片建模任务	2次/秒	2次/秒	-
	照片建模任务列表查询	20次/秒	10次/秒	-
	照片建模任务详情查询	20次/秒	10次/秒	-
	基于图片URL创建照片建模任务	2次/秒	2次/秒	-
3D数字人语音驱动	创建语音驱动任务	2次/秒	2次/秒	-
	获取语音驱动任务列表	20次/秒	10次/秒	-
	创建语音驱动表情动画任务	2次/秒	2次/秒	-
	获取语音驱动表情数据	20次/秒	10次/秒	-
	获取语音驱动数据	20次/秒	10次/秒	-
3D数字人视频驱动	创建视频驱动任务	2次/秒	2次/秒	-
	查询视频驱动任务列表	20次/秒	10次/秒	-
	查询视频驱动任务详情	20次/秒	10次/秒	-
	停止视频驱动任务	2次/秒	2次/秒	-

接口分类	接口名称	单租户接口流控 (华北-北京四)	单租户接口流控 (华东-上海一)	单租户接口流控 (亚太-新加坡)
	控制数字人驱动	2次/秒	2次/秒	-

表 4-2 API 流控限制

接口分类	接口名称	单租户接口流控	系统流控
租户管理	查看租户资源用量信息	10次/秒	20次/秒
	查看租户资源列表	10次/秒	20次/秒
	统计时间段内资源数量	10次/秒	20次/秒

4.4 形象制作限制

形象制作限制说明，如表4-3所示。

表 4-3 形象制作限制

限制项	说明
用户限制	MetaStudio不支持个人用户定制形象，仅支持企业用户。 定制分身形象时，从平台下载的 授权书 中有明确说明，定制分身形象需授权给公司使用。
使用限制	<ul style="list-style-type: none"> 制作好的数字人模型不支持下载或导出至本地。 数字人模型不是通用模型，不与第三方服务兼容。用户在MetaStudio制作的模型，仅支持在本服务内使用。

4.5 声音制作限制

声音制作限制说明，如表4-4所示。

表 4-4 声音制作限制

限制项	说明
用户限制	MetaStudio不支持个人用户定制声音，仅支持企业用户。 定制声音时，从平台下载的 授权书 中有明确说明，定制音色需授权给公司使用。
使用限制	<ul style="list-style-type: none"> 制作好的声音模型不支持下载或导出至本地。 声音模型不是通用模型，不与第三方服务兼容。用户在MetaStudio制作的模型，仅支持在本服务内使用。



4.6 视频制作限制

分身数字人视频制作限制说明，如[表4-5](#)所示。

表 4-5 视频制作限制

视频制作配置	限制
视频制作剧本	<ul style="list-style-type: none"> 一个视频制作剧本的场景数量不能超过50个。 一个视频制作剧本，仅支持使用一个数字人模型和一个音色模型。
文本驱动	<ul style="list-style-type: none"> 每个单场景输入文字的数量不超过1万字。 所有场景输入文字的数量不超过10万字。 每个单场景如果包含语音合成标记语言SSML标签，文字数量需小于128K。
音频驱动	每个单场景上传音频的大小不超过100M。

视频制作配置	限制
视频格式	<p>插入的视频素材支持的格式为：MP4、M4V、MKV、MOV、FLV、3GP、WMV、AVI和WEBM。</p> <p>说明 在视频制作界面，如果添加视频，且视频格式为FLV、3GP、WMV、AVI时，无法展示视频播放效果，仅能图片展示。这是浏览器不兼容导致的，但是不影响视频合成效果。视频合成后，这些插入的视频可以正常播放和展示。</p> <p>插入的视频素材，要求如下所示：</p> <ul style="list-style-type: none"> ● 分辨率最大支持1080P。 ● 帧率小于等于30fps。 ● 编码为AV1、VP8、VP9、H264或H265格式。 <ul style="list-style-type: none"> - VP8、VP9编码仅WEBM视频支持。 - WEBM视频如果为AV1编码格式，不支持合成视频。 - AV1、H265编码格式目前支持Chrome浏览器，不支持Edge浏览器。 ● 视频素材总大小小于1GB。 ● 同时叠加的视频总数量最多支持2个。
音频格式	<p>音频支持的格式为：MP3、M4A、WAV。</p> <p>音频要求如下所示：</p> <ul style="list-style-type: none"> ● 仅提取单声道音频。 ● 音频素材总大小小于500MB。
图片格式	<p>图片支持的格式为：PNG、JPG、JPEG。</p> <p>图片要求如下所示：</p> <ul style="list-style-type: none"> ● 分辨率最大支持1080P。 ● 图片素材总大小小于500MB。
字幕	<p>字幕中正常断句的符号，如中文符号、，。：；！？…等；英文符号,;!?...等，系统会自动去除。</p> <ul style="list-style-type: none"> ● 如果自动去除的符号在一段文字中间，自动去除后，会再增加一个空格。 ● 如果自动去除的符号在段落末尾，仅直接去除。 <p>有些符号不会自动去除，如中文符号“”《》（）·~——等；英文符号""《》（）·~--等。</p>

视频制作配置	限制
Flexus分身数字人	<p>Flexus分身数字人左上角会有标识。未带有标识的数字人均为标准版形象。</p> <p>需注意Flexus视频制作和标准版视频制作分开计费，从各自的套餐包中扣除费用。</p> <ul style="list-style-type: none"> Flexus视频制作：必须使用Flexus形象，声音使用不受限制。 <ul style="list-style-type: none"> 按视频制作时长计费，计费项为“Flexus分身数字人视频制作”，详见计费说明。 系统声音和Flexus声音免费使用。第三方声音的计费方式，详见购买出门问问套餐包。 标准版视频制作：必须使用标准版形象，声音不支持使用Flexus声音。 <ul style="list-style-type: none"> 按视频制作时长计费，计费项为“分身数字人视频制作”，详见计费说明。 系统声音免费使用。第三方声音的计费方式，详见购买出门问问套餐包。
视频制作任务	<p>视频制作历史任务保留6个月。满6个月后控制台“任务中心”菜单的“视频制作”界面，不再展示此任务。</p>
PPT	<p>操作限制，如下所示：</p> <ul style="list-style-type: none"> PPT文件不支持超过50页。 如果PPT的字体不支持在MetaStudio控制台展示，可能会导致视频制作时展示的PPT页面排版与实际排版不一致，建议将PPT转存为PDF文件后再导入。 操作详情请参考PPT如何转为PDF文件？。 如果PPT中有动画、动效，是不支持在MetaStudio控制台展示的，建议将PPT导出为视频后再导入。 操作详情请参考如何在MetaStudio控制台呈现PPT的动画效果？。

4.7 视频直播限制

分身数字人直播限制，如下所示：

- [直播间配置约束](#)
- [直播任务退出说明](#)
- [直播播放说明](#)
- [直播内容审核说明](#)

直播间配置约束

直播间配置限制，如[表4-6](#)所示。

表 4-6 直播间限制

直播间配置	限制
文本驱动	<ul style="list-style-type: none"> • 每个单场景中所有段落输入文字的总量需大于200字。 • 每个段落的文本字符数量需小于2000，如果包含语音合成标记语言SSML标签，需小于3000。 • 整个直播间所有段落的文本总字符数，如果包含语音合成标记语言SSML标签时，需小于30万。
音频驱动	每个单场景上传音频的时长需大于1分钟，大小不超过100M。
视频格式	<p>视频支持的格式为：MP4、M4V、MKV、MOV、FLV、3GP、WMV和AVI。</p> <p>说明 在直播间制作界面，如果添加视频，且视频格式为FLV、3GP、WMV、AVI时，无法展示视频播放效果，仅能图片展示。这是浏览器不兼容导致的，但是不影响直播效果。直播过程中，这些插入的视频可以正常播放和展示。</p> <p>视频要求如下所示：</p> <ul style="list-style-type: none"> • 分辨率最大支持1080P。 • 帧率小于等于30fps。 • 编码为AV1、H264或H265格式。 AV1、H265编码格式目前支持Chrome浏览器，不支持Edge浏览器。 • 视频采样格式为YUV420。 • 视频素材总大小小于1GB。 • 同时叠加的视频总数量为3，包含互动时配置的视频。
音频格式	<p>音频支持的格式为：MP3、M4A、WAV。</p> <p>音频要求如下所示：</p> <ul style="list-style-type: none"> • 仅提取单声道音频。 • 音频素材总大小小于500MB。
图片格式	<p>图片支持的格式为：PNG、JPG、JPEG。</p> <p>图片要求如下所示：</p> <ul style="list-style-type: none"> • 分辨率最大支持1080P。 • 图片素材总大小小于500MB。
字幕	<p>字幕中正常断句的符号，如中文符号、，。：；！？…等；英文符号,;!?...等，系统会自动去除。</p> <ul style="list-style-type: none"> • 如果自动去除的符号在一段文字中间，自动去除后，会再增加一个空格。 • 如果自动去除的符号在段落末尾，仅直接去除。 <p>有些符号不会自动去除，如中文符号“”《》（）~——等；英文符号""《》()~--等。</p>
透明背景	视频直播不支持透明背景设置，仅视频制作和智能交互支持。

直播任务退出说明

直播任务退出说明，如表4-7所示。

表 4-7 直播任务退出说明

直播任务退出现场	操作说明
租户自行退出直播任务	<p>租户可通过下述方式，退出直播任务。</p> <ul style="list-style-type: none"> 在MetaStudio控制台开播页面，停止直播任务。 调用结束数字人智能直播任务接口，停止直播任务。
服务自动关停直播任务	<p>下述场景下，生成的直播视频流没有观看者，服务为保障租户利益，会自动关停直播任务。</p> <ul style="list-style-type: none"> 在MetaStudio控制台开播，且没有配置RTMP推流时，用户关闭了视频播放窗口一段时间以后，约2~3分钟。 调用API开播，且没有配置RTMP推流时，客户端退出RTC房间一段时间以后，约2~3分钟。
特殊退出场景	<p>如果直播内容涉及安全合规风险，服务运维人员，会通过运维界面，直接冻结直播任务。</p>

直播播放说明

直播播放说明，如表4-8所示。

表 4-8 直播播放说明

直播播放配置	操作说明
租户控制直播播放的配置	<p>为保障主播可以截取到完整的直播视频，MetaStudio直播服务提供了控制直播播放的配置。</p> <p>启动直播前，Console控制台界面，需要去勾选“加载完自动直播”（即配置项“play_policy.auto_play_script”设置为“FALSE”）。启动直播任务后，会等待租户的开始直播指令（如果接口调用的方式，会通过RTSA COMMAND指令发送play_start指令），才会启动直播，否则数字人一直保持静默状态。</p>

直播内容审核说明

直播内容审核说明，如表4-9所示。



表 4-9 直播间内容审核说明

内容审核机制	操作说明
直播间内容审核	<p>MetaStudio服务内容审核说明，如下所示：</p> <ul style="list-style-type: none"> 默认会对直播间的各个配置项，如直播间名称、描述、文本、音频、视频、图片等内容进行机器审核。以排除内容违规风险，避免生成不合法的直播内容。 如果租户已对接其他内容风控系统，且承诺内容风控自行负责，可提交工单申请关闭MetaStudio服务内容审核功能。 如果服务接收到直播内容投诉，按照国家相关部门要求，会冻结违规的直播间及直播任务。情节严重的，还会冻结租户的全部资源。

4.8 Flexus 分身数字人限制

Flexus分身数字人限制说明，如[表4-10](#)所示。

表 4-10 声音制作限制

限制项	说明
Flexus形象和声音使用限制	Flexus分身数字人目前仅支持视频制作，不支持视频直播和智能交互。
视频制作限制	<p>Flexus数字人采用实景拍摄，视频制作时不支持切换背景。</p> <p>Flexus分身数字人左上角会有标识。未带有标识的数字人均为标准版形象。</p> <p>需注意Flexus视频制作和标准版视频制作分开计费，从各自的套餐包中扣除费用。</p> <ul style="list-style-type: none"> Flexus视频制作：必须使用Flexus形象，声音使用不受限制。 <ul style="list-style-type: none"> 按视频制作时长计费，计费项为“Flexus分身数字人视频制作”，详见计费说明。 系统声音和Flexus声音免费使用。第三方声音的计费方式，详见购买出门问问套餐包。 标准版视频制作：必须使用标准版形象，声音不支持使用Flexus声音。 <ul style="list-style-type: none"> 按视频制作时长计费，计费项为“分身数字人视频制作”，详见计费说明。 系统声音免费使用。第三方声音的计费方式，详见购买出门问问套餐包。

4.9 照片数字人限制

照片数字人视频制作限制说明，如表4-11所示。

表 4-11 照片数字人视频制作限制

照片数字人视频制作配置	限制
人像照片	<p>制作照片数字人的人像照片拍摄要求，如下所示：</p> <ul style="list-style-type: none"> 真人正面照，不支持动漫卡通形象；面部占比50%以上，侧脸俯仰角度不超过15度。 五官清晰无遮挡，目视前方。 良好的亮度，确保面部细节。 照片内不能出现多张人脸。 <p>人像照片格式要求，如下所示：</p> <ul style="list-style-type: none"> 支持的格式为：PNG、JPG、JPEG、WEBP。 大小不超过16MB。 分辨率最大支持1080P，如果超出，会自动缩放为1080P，且需保证分辨率大于等于80*80。
音频	<p>音频录制要求，如下所示：</p> <ul style="list-style-type: none"> 单声道音频。 清晰人声无背景音乐与杂音。 <p>音频格式要求，如下所示：</p> <ul style="list-style-type: none"> 支持的格式为：MP3、WAV。 大小不超过20MB，时长不超过5分钟。
文本驱动	<p>文本驱动方式的要求如下所示：</p> <ul style="list-style-type: none"> 文本输入长度不超过2000字。 通过文本合成的音频时长不能超过5分钟。 <p>在照片数字人制作页面，选择“文本驱动”方式驱动照片说话时，需要输入文本并选择音色，再单击“试听”预估音频时长。如果超过5分钟，会导致照片数字人制作失败。</p>

4.10 智能交互限制

智能交互对硬件设备的最低配置要求，如表4-12所示。

表 4-12 硬件设备要求

硬盘配置	最低配置
CPU	i5及以上版本。

硬盘配置	最低配置
内存	8G以上。
硬盘	128G及以上。
操作系统	Windows或Android系统。
麦克风	需要有麦克风，且要保障1米以内的拾音效果。
扬声器	需要有扬声器。
触摸屏	需要有触摸屏。
带宽	4M以上。

由于浏览器的安全策略限制，仅支持通过“<https://域名>”方式访问，否则无法正常加载运行智能交互WebSDK。

智能交互界面对浏览器的要求，如表4-13所示。

表 4-13 浏览器适配详情

操作系统类型	浏览器类型	浏览器版本
Windows	Chrome浏览器	91+
	Edge浏览器	80+
Android	移动版Chrome浏览器	91+
iOS	移动版Safari浏览器	17.3+
Mac	Safari浏览器	15.6+

表 4-14 透明背景支持情况

操作系统类型	浏览器类型	浏览器版本
Windows	Chrome浏览器	94+
	Edge浏览器	94+
Android	移动版Chrome浏览器	94+
iOS	移动版Safari浏览器	17.3+
Mac	Safari浏览器	16.4+

4.11 资产管理说明

资产管理说明，如下所示：

- [资产存储计费说明](#)
- [资产创建失败说明](#)
- [资产冻结说明](#)

资产存储计费说明

资产存储计费说明，如[表4-15](#)所示。

表 4-15 资产存储计费说明

资产存储	计费说明
免费存储空间	MetaStudio服务免费赠送每个租户5GB资产存储空间，用于上传各类音视频、图片等素材。如果租户需要更多的存储空间，请购买单独的资产管理包月产品。
定制形象模型存储	基于MetaStudio服务定制生成的形象模型，其所包含的音色模型所占存储空间，暂不计入租户资产存储计费中。
定制视频资产存储	基于MetaStudio服务定制生成的视频资产，6个月以内不计入资产存储空间计费；满6个月以后，参与计费。
资产存储计费说明	资产只要没有被彻底删除，均需要进行租户的资产存储空间计费。即使资产处于冻结状态，或在回收站时，也需参与计费。

资产创建失败说明

如下创建资产时的异常场景，会导致资产被自动置为失败状态。

- 超过24小时未上传资产主文件。
- 资产标题、描述和主文件内容审核失败。
- 服务校验资产主文件与用户创建的文件信息不匹配。

资产冻结说明

资产冻结说明，如[表4-16](#)所示。

表 4-16 资产冻结说明

场景	操作说明
服务冻结租户资产	MetaStudio服务对租户资产内容的风险有一定连带责任。如果服务接收到资产内容投诉，按照国家相关部门要求，会冻结违规资产。情节严重的，还会冻结租户的全部资源。 被冻结的资产不能用作内容生成的素材。租户可以删除被冻结的资产或者申请解冻资产。

场景	操作说明
用户申请解冻资产	用户可以通过申诉的方式，申请解冻资产，解冻后资产会恢复至解冻前的状态。 如果处于回收站中的资产被冻结，解冻后，其状态会恢复至被删除前的状态。

4.12 SP 用户限制

表 4-17 限制说明

限制项	说明
开放范围	SP用户仅对大客户开放申请和使用。
功能范围	<ul style="list-style-type: none">SP用户仅支持购买大颗粒度资源包，并分配资源给关联租户使用。SP用户不支持访问MetaStudio控制台，进行控制台操作。

5 常用概念

数字人

数字人是一种以数字形式存在于数字空间中的虚拟人物，它具有拟人或真人的外貌、行为特点，并具备一定的智能和情感，可以进行交互和表达。数字人也可以被称之为虚拟形象、数字虚拟人、虚拟数字人等。数字人的核心技术主要包括计算机视觉、计算机图形学、动作捕捉和驱动、图像渲染和人工智能等。

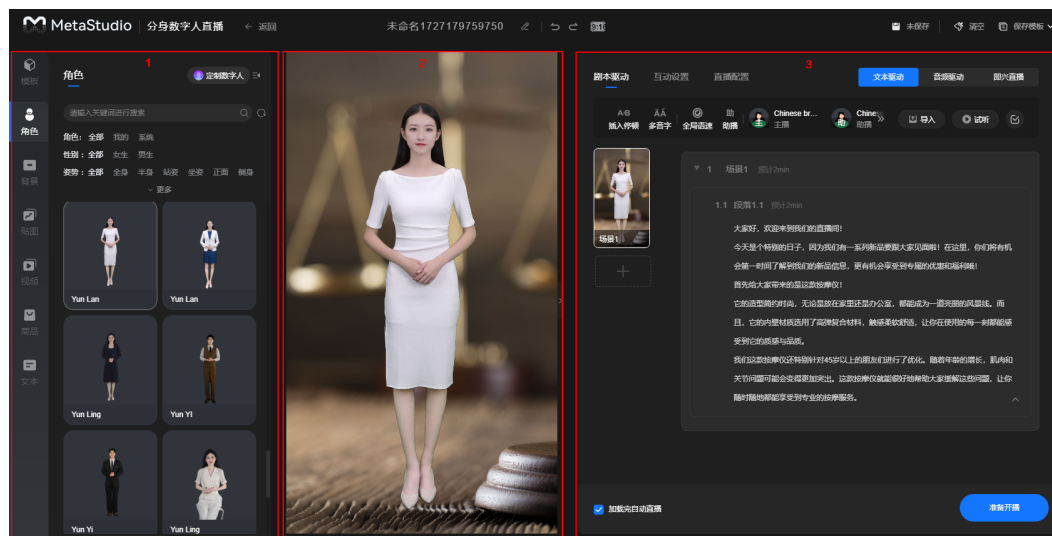
- 服务型数字人：利用深度神经网络进行图像合成、高度拟真的虚拟人。
具备如下特点：
 - 2D模型，通过拍摄真人视频训练生成
 - 无表情&骨骼数据
 - 只能由AI驱动
 - 使用既定表情&动作
- IP型数字人：由计算机图形学技术创造出来的与人类形象高度接近的数字化形象。
具备如下特点：
 - 3D模型，通过艺术设计生成
 - 有表情、骨骼数据
 - 可由真人动作驱动
 - 能呈现自定义表情&动作

剧本、画布和设置

视频制作、视频直播功能，都需要进行视频创作，下面以视频直播为例，进行说明。

- 视频素材：如图中1区域所示，包括模板、角色、背景、贴图、视频、商品、文本等素材。支持使用预置的，也支持自定义生成。
- 画面布局：简称“画布”，如图中2区域所示。可以在画布中插入人物形象、背景、贴图、文本框、视频等内容，支持调整各项直播素材的大小和图层位置。所有的设置效果均在画布中直观可见。
- 开播设置：简称“设置”，如图中3区域所示，包括视频直播的各类设置。
- 剧本设计：简称“剧本”，视频制作和视频直播的全流程制作过程，统称为剧本设计，包含素材准备、画布设计、开播设置等操作，最终合成视频。

图 5-1 直播间编辑界面



角色

数字人模型。

动作

数字人模型展示的动作。

捏脸

用户可根据自己的想法自由捏造风格化虚拟形象。

风格化素材

建模/捏脸时采用的五官、服饰、装饰等素材，比如发型、脸型、装饰、眉毛、鼻子等。

6 计费说明

数字内容生产线支持按需计费、包年/包月（包周期）、一次性计费等计费模式，以满足不同场景下的用户需求。

计费项

数字内容生产线计费项包含分身数字人、3D数字人和资产管理费用。详细的价格信息，详见[价格计算器](#)。

所需费用按照不同的计费方式产生，具体详情如[表6-1](#)所示。

须知

用户可以在MetaStudio控制台服务概览页面，单击“购买资源”，购买所需资源。如需体验新用户优惠活动，可以查看[MetaStudio数字人专场](#)。

表 6-1 计费项说明

类别	计费项	计费说明	计费方式
分身数字人	Flexus分身数字人	按Flexus数字人生成个数和声音克隆个数计费。 须知 每提交一个Flexus分身数字人定制任务，会同时生成一个Flexus形象和一个Flexus声音，对应消耗一个形象资源和一个声音资源。	<ul style="list-style-type: none">• 按需预付费套餐包• 一次性计费

类别	计费项	计费说明	计费方式
	Flexus分身数字人视频制作	<p>按Flexus数字人视频制作时长计费。</p> <p>视频制作计费说明，如下所示：</p> <ul style="list-style-type: none"> 不同计费方式的抵扣顺序为：一次性计费 > 按需预付费套餐包 > 按需计费。 如果合成视频的费用使用一次性计费，那么额度用尽后，需重新购买使用。 如果合成视频的费用使用按需预付费套餐包抵扣，额度使用完后，转按需付费。 <p>须知</p> <ul style="list-style-type: none"> Flexus数字人视频制作时，必须使用flexus形象；声音可以使用flexus声音或标准版声音。标准版视频制作，不支持使用flexus形象或声音。 Flexus数字人与标准版数字人的视频制作资源不共享，分开计费。 	<ul style="list-style-type: none"> 按需计费 按需预付费套餐包 一次性计费
	分身数字人形象制作	<p>按分身数字人形象制作的个数计费。</p> <p>每成功制作1个分身数字人形象，从所购预付费资源包中扣除1个。</p>	一次性计费
	分身数字人视频制作	<p>按分身数字人视频制作时长计费。</p> <p>视频制作的时长有限制，不能大于所购预付费资源包中的时长规格。</p> <p>视频制作计费说明，如下所示：</p> <ul style="list-style-type: none"> 不同计费方式的抵扣顺序为：一次性计费 > 按需预付费套餐包 > 按需计费。 如果合成视频的费用使用一次性计费，那么额度用尽后，需重新购买使用。 如果合成视频的费用使用按需预付费套餐包抵扣，额度使用完后，转按需付费。 <p>须知</p> <ul style="list-style-type: none"> 平台合成的视频按秒记录时长，套餐包以分钟记录时长。所以合成视频的总时长会从秒转化为分钟，保留小数点后两位，末位四舍五入。 视频制作生成的视频即使因内容违规被冻结，依然会收取制作费用。 一次性计费模式下的19.9元“分身数字人视频制作60分钟体验包”，是未购买过MetaStudio资源的新用户专享福利，仅支持用来视频制作，不支持训练形象。 	<ul style="list-style-type: none"> 按需计费 按需预付费套餐包 一次性计费

类别	计费项	计费说明	计费方式
	声音制作	<p>按声音制作的个数计费。</p> <p>包括如下类型：</p> <ul style="list-style-type: none"> ● 基础版：20条语料音频 ● 进阶版：100条语料音频 ● 高级版：1小时以上的长音频 WAV文件 <p>每成功制作1个声音，从所购预付费资源包中扣除1个。</p>	一次性计费
	分身数字人视频直播	<p>按数字人各路直播的总时长计费。</p> <ul style="list-style-type: none"> ● 包年/包月有效期内，自动走包年/包月套餐。如果超出包年/包月的路数，转按需计费。 ● 超出包年/包月有效期，转按需计费。 <p>须知 路数就是并发数，是视频直播同时执行的任务数量。需要几路并发，就买几路。如果已开通按需计费，当并发任务数超出已购买路数时，超出的路数自动走按需结算，费用从当前账户余额中扣除。</p>	<ul style="list-style-type: none"> ● 按需计费 ● 包年/包月

类别	计费项	计费说明	计费方式
	分身数字人智能交互	<p>支持购买多路，每路按时长计费。</p> <p>须知：</p> <ul style="list-style-type: none"> 智能交互中数字人分身的动作是根据数字分身拍摄形象训练生成的，唇形和文字声音匹配，数字分身和背景融合自然，清晰度1080P。 智能交互根据客户需要，后台提供多种知识库配置，支持配置中英双语。 智能交互基于已有训练生成的数字分身形象模型和声音模型，结合知识库，可与用户进行实时对话。 用户购买路数越多，同时进行数字人交互对话的数量则越多。 资源超过有效期后将无法使用，建议开通自动续费功能。 路数就是并发数，是智能交互同时执行的任务数量。需要几路并发，就买几路。如果已开通按需计费，当并发任务数超出已购买路数时，超出的路数自动走按需结算，费用从当前账户余额中扣除。 	包年/包月
	照片数字人视频制作	<p>按照片数字人视频制作时长计费。</p> <p>计费时优先使用套餐包中的额度，该额度使用完后，转按需付费。</p> <p>须知</p> <ul style="list-style-type: none"> 平台合成的视频按秒记录时长，套餐包以分钟记录时长。所以合成视频的总时长会从秒转化为分钟，保留小数点后两位，末位四舍五入。 当前资源默认置灰，即不支持购买。如需购买，仅支持企业用户联系产品经理或者提交工单开通购买，个人用户不支持。 	<ul style="list-style-type: none"> 按需计费 按需预付费套餐包

类别	计费项	计费说明	计费方式
	分身数字人名片制作	<p>按照数字人名片制作的个数计费。</p> <p>每成功制作1个数字人名片，从所购预付费资源包中扣除1个。</p> <p>计费时优先使用套餐包中的额度，该额度使用完后，转按需付费。</p> <p>分身数字人名片的含义：客户根据其单张正面照（头肩像），调用相关服务，进行形象复刻生成分身数字人名片，该数字人名片形式是介绍名片或公司信息的一段视频。</p> <p>须知 当前资源默认置灰，即不支持购买。如需购买，仅支持企业用户联系产品经理或者提交工单开通购买，个人用户不支持。</p>	<ul style="list-style-type: none"> ● 按需计费 ● 按需预付费套餐包
3D数字人	照片建模	<p>按照3D风格化基础模型建模，调用API接口的次数计费。</p> <p>每成功调用一次3D风格化基础模型建模的API，从所购预付费资源包中扣除1次。</p> <p>计费时优先使用套餐包中的额度，该额度使用完后，转按需付费。</p> <p>照片建模的含义：根据单张照片自动生成3D风格化基础模型。</p>	<ul style="list-style-type: none"> ● 按需计费 ● 按需预付费套餐包
	语音驱动	<p>按照输入文本或语音驱动数字人口型、表情等动作的时长进行计费。</p> <p>计费时优先使用套餐包中的额度，该额度使用完后，转按需付费。</p> <p>语音驱动的含义：根据输入文本或语音驱动数字人口型、表情等动作。</p>	<ul style="list-style-type: none"> ● 按需计费 ● 按需预付费套餐包
	视觉驱动	<p>按摄像头采集真人动作驱动各路数字人动作的总时长进行计费。</p> <ul style="list-style-type: none"> ● 包年/包月有效期内，自动走包年/包月套餐。 如果超出包年/包月的路数，转按需计费。 ● 超出包年/包月有效期，转按需计费。 <p>视觉驱动的含义：通过摄像头采集真人运动数据，驱动数字人运动。</p>	<ul style="list-style-type: none"> ● 按需计费 ● 包年/包月

类别	计费项	计费说明	计费方式
资产管理	资产管理	<p>按资产管理的规格（资产存储空间）进行计费。</p> <p>需注意：</p> <ul style="list-style-type: none">首次购买或创建资产时，免费赠送5G存储空间。下述场景，使用资产空间不会计费：<ul style="list-style-type: none">用户上传的分身数字人模型、3D数字人模型和声音。平台生成的分身数字人模型、3D数字人模型和声音。平台合成的视频资产6个月内，默认不计费。6个月以后，开始计费。 <p>资产管理的含义：提供数字资产存储、简单分类、搜索等基础管理能力。</p>	包年/包月

计费模式

数字内容生产线提供按需计费、包年/包月、一次性计费三种计费模式，以满足不同场景下的用户需求。

- 按需计费模式：包含按需和按需预付费套餐包两种计费方式。
 - 按需计费：一种后付费模式，即先使用再付费。每个计费周期分别统计各资源的实际使用量，并根据使用量*单价的方式计算出费用，然后从账户余额中扣除对应金额。
 - 按需预付费套餐包：您也可以通过按需预付费套餐包的方式，提前购买使用额度，获取更多的优惠。套餐包需要在有效期内使用，且超出套餐包的部分自动转按需计费。
- 包年/包月计费模式：一种预付费模式，即先付费再使用，按照订单的购买周期进行结算。购买周期越长，享受的折扣越大。一般适用于需求量长期稳定的成熟业务。
- 一次性计费：一次性预付费套餐包，由MetaStudio服务进行结算。

资源有效期

套餐包从订单付款成功后的生效时间开始计费，请在有效期内使用套餐包中的资源。按需付费则按照资源的实际使用情况进行计费，无有效期的限制。

资源包的“到期时间”可以在MetaStudio控制台的服务概览页面查看，如图6-1所示。

图 6-1 资源列表

名称/ID	产品类型	订单号	剩余量	总量	到期时间	状态	计费模式	操作
风格化视频建模	照片建模		20 (次)	20 (次)	-		一次性	
视觉驱动时长	视觉驱动		9.48 (分钟)	10 (分钟)	-		一次性	
照片数字人视频制作	照片数字人视频制作		1,267.87 (...)	6,000 (分钟)	-		一次性	
语音驱动	语音驱动		3.95 (分钟)	5 (分钟)	-		一次性	
分身数字人形象制作高级版	分身数字人形象制作		3 (个)	3 (个)	-		一次性	
分身数字人制作百万字声...	声音合成		100 (万字)	100 (万字)	-		一次性	
分身数字人智能交互 20240306T15:00:149250...	分身数字人智能交互		3,000 (路)	3,000 (路)	2025-03-06 23:59:59 58天后到期	● 正常	包周期	
声音制作高级版 20240607T15:25:025650...	声音制作		777 (个)	872 (个)	2025-06-07 23:59:59 151天后到期	● 正常	一次性	
分身数字人名片制作 20241012T14:15:072090...	分身数字人名片制作		808 (个)	856 (个)	2025-10-12 23:59:59 278天后到期	● 正常	一次性	
分身数字人视频制作 20241012T14:15:054320...	分身数字人视频制作		5,569.75 (...)	8,112.82 (...)	2025-10-22 23:59:59 288天后到期	● 正常	一次性	

续费

数字内容生产线续费说明，如下所示：

- 如果使用按需预付费套餐包或包年/包月计费方式，支持在用完前续费。如需续费，请进入[续费管理](#)页面，完成续费操作。如果未能续费成功，那么产生的费用会从账户余额中扣除。
- 如果使用一次性计费，不支持续费，需要在MetaStudio控制台再次购买相同的资源包。

退订

- 包年/包月套餐包支持退订。
- 按需预付费套餐包、一次性计费不支持退订。
- 在服务概览页面，如果通过下述方式开通了按需计费，不支持关闭。
单击资源卡片中的“去开通”，在弹出的“开通按需计费服务提示”对话框中，勾选同意协议。单击“确定”，开通按需计费。

到期与欠费

- 欠费：预付费资源包没有欠费概念，按需付费是根据每个计费项在对应的计费周期进行扣费的，若余额不足，则无法对上一个计费周期进行扣费，从而导致数字内容生产线服务欠费。当您的账户欠费后，华为云将根据您的客户等级来定义云服务资源宽限期、保留期时长，更多欠费停服处理规则请参见[保留期](#)。
您可以在[费用中心](#)查看账户的欠费信息，并参照[欠费还款](#)核销欠款。
请注意在保留期完成续费操作。如果服务欠费后进入保留期，会在续费后解冻，服务恢复正常使用。
- 服务到期：按需付费没有到期概念。预付费资源包到期后，若资源包中仍有额度将一次性清零，产生的费用将从余额中扣除。

7 权限管理

如果您需要对购买的MetaStudio资源，给企业中的员工设置不同的访问权限，以达到不同员工之间的权限隔离，您可以使用统一身份认证服务（Identity and Access Management，简称IAM）进行精细的权限管理。该服务提供用户身份认证、权限分配、访问控制等功能，可以帮助您安全地控制资源地访问。

通过IAM，您可以在账号中给员工创建IAM用户，并授权控制他们对资源的访问范围。例如您的员工中有负责软件开发的人员，您希望他们拥有MetaStudio的使用权限，但是不希望他们拥有删除MetaStudio等高危操作的权限，那么您可以使用IAM为开发人员创建用户，通过授予仅能使用MetaStudio，但是不允许删除MetaStudio的权限策略，控制他们对MetaStudio资源的使用范围。

如果账号已经能满足您的要求，不需要创建独立的IAM用户进行权限管理，您可以跳过本章节，不影响您使用MetaStudio服务的其他功能。

IAM是华为云提供权限管理的基础服务，无需付费即可使用，您只需要为您账号中的资源进行付费。关于IAM的详细介绍，请参见《[IAM产品介绍](#)》。

MetaStudio不支持企业项目。

MetaStudio 权限

默认情况下，新建的IAM用户没有任何权限，您需要将其加入用户组，并给用户组授予策略或角色，才能使得用户组中的用户获得对应的权限，这一过程称为授权。授权后，用户就可以基于被授予的权限对云服务进行操作。

MetaStudio部署时通过物理区域划分，为项目级服务。授权时，“作用范围”需要选择“区域级项目”，然后在指定区域（如华北-北京四）对应的项目（cn-north-4）中设置相关权限，并且该权限仅对此项目生效；如果在“所有项目”中设置权限，则该权限在所有区域项目中都生效。访问MetaStudio时，需要先切换至授权区域。

权限根据授权精细程度分为角色和策略。

- **角色：** IAM最初提供的一种根据用户的工作职能定义权限的粗粒度授权机制。该机制以服务为粒度，提供有限的服务相关角色用于授权。由于各服务之间存在业务依赖关系，因此给用户授予角色时，可能需要一并授予依赖的其他角色，才能正确完成业务。角色并不能满足用户对精细化授权的要求，无法完全达到企业对权限最小化的安全管控要求。
- **策略：** IAM最新提供的一种细粒度授权的能力，可以精确到具体服务的操作、资源以及请求条件等。基于策略的授权是一种更加灵活的授权方式，能够满足企业对权限最小化的安全管控要求。

如表7-1所示，包括了MetaStudio的所有系统策略。

表 7-1 MetaStudio 系统权限

策略名称	描述	策略类别	策略内容
MetaStudio FullAccess	MetaStudio管理员权限，拥有该权限的用户可以拥有MetaStudio支持的全部权限。	系统策略	MetaStudio FullAccess策略内容
MetaStudio ReadOnlyAccess	MetaStudio只读权限，拥有该权限的用户仅能查看MetaStudio数据。	系统策略	MetaStudio ReadOnlyAccess策略内容

相关链接

- [IAM产品介绍](#)
- 创建用户组、用户并授予MetaStudio权限请参考：[创建用户并授权使用MetaStudio](#)

MetaStudio FullAccess 策略内容

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Action": [
        "metastudio:*"
      ],
      "Effect": "Allow"
    }
  ]
}
```

MetaStudio ReadOnlyAccess 策略内容

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Action": [
        "metastudio:*:show*"
      ],
      "Effect": "Allow"
    }
  ]
}
```

8 算法备案公示

下述内容为MetaStudio服务提供的算法备案信息、基本原理、运行机制和目的意图等内容，以保障用户的知情权，方便用户更好的选择和使用MetaStudio服务。

华为云 MetaStudio 分身数字人驱动算法

表 8-1 分身数字人驱动算法

算法项	描述
算法名称	华为云MetaStudio分身数字人驱动算法
备案编号	网信算备520111252474601240045号
算法基本原理	<p>分身数字人驱动算法是指通过深度学习生成数字人驱动模型，模型生成后，输入音频来合成数字人视频的一种技术。</p> <p>其基本情况包括：</p> <ul style="list-style-type: none">● 输入数据：真人视频、音频。● 算法原理：通过深度学习算法来学习真人视频，生成驱动该真人形象的数字人模型。通过该模型输入音频，合成数字人视频。● 输出结果：数字人视频。● 应用领域：分身数字人驱动算法可以应用于真人视频自动生成，包括新闻播报，课件制作等场景，以取代真人视频拍摄，提升视频内容生产效率。
算法运行机制	<ol style="list-style-type: none">1. 选择一段真人视频。2. 输入真人视频，经过平台专家安全审核通过，且用户授权使用后，由训练人员选取真人视频中符合要求的视频进行预处理。预处理完成后，进行深度学习训练，生成该真人形象的数字人驱动模型。3. 推理阶段输入一段音频。4. 音频输入至数字人驱动模型后，经过模型推理生成数字人形象播报视频。5. 数字人视频通过审核后返回给用户。

算法项	描述
算法应用场景	分身数字人驱动算法可以应用于真人视频自动生成，包括新闻播报、课件制作等场景，以取代真人视频拍摄，提升视频内容生产效率。
算法目的意图	可以使用授权过的真人视频，在预训练模型基础上，生成真人数字人驱动模型。该模型可基于音频生成口型匹配的数字人视频，实现真人视频自动生成，包括新闻播报、课件制作等场景，以取代真人视频拍摄，提升视频内容生产效率。

华为云 MetaStudio 分身数字人声音制作算法

表 8-2 声音制作算法

算法项	描述
算法名称	华为云MetaStudio分身数字人声音制作算法
备案编号	网信算备520111252474601240079号
算法基本原理	分身数字人声音制作算法是指使用深度学习算法生成数字人声音模型，再使用该模型通过输入文字生成数字人语音的一种技术。 其基本情况包括： <ul style="list-style-type: none"> ● 输入数据：真人语音音频。 ● 算法原理：通过深度学习算法，学习真人语音音频生成数字人声音模型，通过该模型，输入文本生成数字人语音。 ● 输出结果：接近真人音色的数字人语音。 ● 应用领域：分身数字人声音制作算法可以应用于数字人视频合成时的语音配音。在新闻播报、课件制作等场景模拟真人配音，提升数字内容生产效率。
算法运行机制	训练阶段： <ol style="list-style-type: none"> 1. 用户上传一段真人语音音频及授权书作为输入。 2. 音频经过人工安全审核和授权认证后，由训练人员标注用于训练的音频数据，使用深度学习算法训练生成数字人声音模型。 推理阶段： <ol style="list-style-type: none"> 1. 用户上传一段文本作为输入文本内容，由系统自动审核。 2. 输入文本使用数字人声音模型推理生成数字人语音。
算法应用场景	分身数字人声音制作算法可以应用于数字人语音合成。在新闻播报、课件制作等场景模拟真人配音，提升数字内容生产效率。

算法项	描述
算法目的意图	可以通过有授权的真人语音音频，在预训练模型基础上微调生成数字人声音模型，该模型可用于基于文本合成类似真人音色的数字人语音。

华为云 MetaStudio 数字人照片建模算法

表 8-3 照片建模算法

算法项	描述
算法名称	华为云MetaStudio数字人照片建模算法
备案编号	网信算备520111252474601230033号
算法基本原理	<p>数字人照片建模算法是指使用深度学习算法将已授权的人像照片信息转换为数字人3D模型的一种技术。</p> <p>其基本情况包括：</p> <ul style="list-style-type: none"> ● 输入数据：有授权的人像照片。 ● 算法原理：使用深度学习算法，将人像照片转换为数字人3D模型。 ● 输出结果：数字人3D模型。 ● 应用领域：数字人照片建模算法可应用于影视制作、仿真形象生成、虚拟现实等领域，能加速数字人3D模型产业的生产效率。
算法运行机制	<ol style="list-style-type: none"> 1. 选择一张已授权的人像照片作为输入。 2. 输入的人像照片经过安全过滤，判断是否通过安全筛选，若不通过则不进行数据生成和结果返回操作。 3. 将人像照片输入至算法模型中，通过特征提取、人脸重建、素材特征匹配等进行人头重建，生成3D人头模型。 4. 生成的人头模型通过配上默认身体素材配件，形成完整的3D数字人模型，并返回结果。
算法应用场景	数字人照片建模算法可应用于影视制作、仿真形象生成、虚拟现实等领域，能加速数字人模型产业的生产效率。
算法目的意图	<p>使用已授权的人像照片生成3D数字人模型，应用于如下场景：</p> <ul style="list-style-type: none"> ● 用于影视动漫制作角色初始模型的快速生成。 ● 用于虚拟现实Avatar形象快速生成。 ● 其他需要数字人3D形象的领域，快速提升形象生成效率。

华为云 MetaStudio 数字人视觉驱动算法

表 8-4 视觉驱动算法

算法项	描述
算法名称	华为云MetaStudio数字人视觉驱动算法
备案编号	网信算备520111252474601240053号
算法基本原理	<p>数字人视觉驱动算法是指使用神经网络，将视频中的人脸表情和人体姿态，转换为表情基系数及数字人骨骼驱动数据的一种技术。</p> <p>其基本情况包括：</p> <ul style="list-style-type: none">● 输入数据：单人表演视频。● 算法原理：通过深度学习算法识别人脸表情和人体姿态，转换为表情基系数及数字人骨骼驱动数据。● 输出结果：表情基系数、数字人骨骼驱动数据。● 应用领域：应用于影视制作、虚拟人姿态控制等场景，加速影视制作的动画生成效率，提升虚拟人控制体验。
算法运行机制	<ol style="list-style-type: none">1. 输入为单人表演视频。2. 通过视频抽帧得到单张图片。经过安全过滤，判断是否通过安全筛选，若不通过则不进行数据生成和结果返回操作。3. 将视频图片输入至算法模型中，将视频图像分割为面部、手部和身体三个区域。4. 使用深度学习算法，识别面部区域转化为面部表情，识别手部区域转化为手部骨骼驱动数据，识别身体转化为人体骨骼驱动数据。5. 对算法输出系数进行平滑处理及异常数据过滤，返回结果。
算法应用场景	数字人视觉驱动算法可用于影视制作、虚拟人姿态控制等场景，加速影视制作的动画生成效率，提升虚拟人控制体验。
算法目的意图	<ul style="list-style-type: none">● 视频驱动数字人动作，实现影视制作动画快速生成。● 虚拟现实Avatar形象姿态控制。● 其他需要通过表演者驱动数字人动作表情的领域，以提升形象动画生产效率。

华为云 MetaStudio 数字人语音驱动算法

表 8-5 语音驱动算法

算法项	描述
算法名称	华为云MetaStudio数字人语音驱动算法
备案编号	网信算备520111252474601240061号
算法基本原理	<p>数字人语音驱动算法是指使用深度学习将语音转换成3D数字人表情和肢体驱动数据的一种技术。</p> <p>其基本情况包括：</p> <ul style="list-style-type: none">● 输入数据：语音音频数据。● 算法原理：通过深度学习算法，提取语音音频中的特征，并转化为表情驱动的表情基系数。● 输出结果：表情基系数。● 应用领域：应用于3D数字人文本和语音驱动场景，包括：短视频制作、直播和交互等。在特定场景中，可替代人快速生成视频内容，提升内容生成的效率。
算法运行机制	<ol style="list-style-type: none">1. 训练阶段输入预先获取的高质量语音及其表情基系数，通过学习获得语音特征与表情基系数的关系。2. 用户使用时，输入音频流或音频文件。3. 音频经过安全审核后，进入下一步操作，否则不返回结果。4. 对音频特征提取后，获取音频特征，再通过算法将音频特征转换为表情基系数。5. 返回结果数据。
算法应用场景	数字人语音驱动算法可用于短视频制作、直播、交互等场景。在特定场景中，可替代人快速生成视频内容，以提升内容生成的效率。
算法目的意图	通过学习语音与表情基系数的关系，实现使用语音生成视频的能力。在使用数据人形象生成视频的场景，包括短视频制作、直播、智能交互等，可快速生成不同台词的视频内容。