MetaStudio

Service Overview

Issue   01
Date   2025-12-17

Huawei Cloud Computing Technologies Co., Ltd.

Contents

1 What Is MetaStudio?

The MetaStudio solution offers a wide range of services, such as virtual avatar video production, livestreaming, and intelligent interaction, breathing new life into content creation across industries.

Scenarios

Education: virtual avatar courseware creation and virtual teachers

Government and public services: virtual media personalities, presentations by virtual avatars, and conferences with virtual avatars

Healthcare: intelligent healthcare customer service and virtual doctor training

Finance: virtual customer service personnel, virtual brand ambassadors, and virtual avatar training video production

Culture and tourism: virtual customer service personnel, virtual guides, and virtual docents

Broadcasting and TV: virtual variety show hosts and virtual news anchors

Interactive entertainment and e-commerce: virtual avatar livestreaming, virtual avatar short video production, and virtual streamers of live commerce

2 Advantages

One-Stop Content Generation on the User-Friendly MetaStudio Console

You need to register with MetaStudio, upload an authorization letter, and record and upload an audio/video for model training to generate a custom avatar image and voice. Select elements such as backgrounds, voices, and avatars, and implement capabilities such as video production and livestreaming and intelligent interaction based on intelligent text or audio control.

Huawei Cloud Pangu Virtual Human Model for New Digital Marketing Techniques Across Industries

MetaStudio leverages Huawei's global cloud infrastructure (compute-network convergence and ultra-low latency) and diverse computing power (CPUs/GPUs/NPUs), along with the Huawei Cloud Pangu virtual human model, to simplify the training and generation of virtual avatars, objects, and spaces. These virtual avatars are ideal for video production, livestreaming, intelligent interaction, and virtual media personalities.

Fast Iteration of Virtual Avatars to Stay Relevant

- Accurate lip sync: accuracy > 95%, capable of speaking many languages, and strong generalization
- Natural facial expression: AI eye contact correction for continuous eye contact and enhanced facial realism
- Realistic images: intelligent relighting for seamless integration with backgrounds
- Modeling in complex scenarios: walking and motion training for smoother movements of virtual avatars
- Multi-language generalization: able to speak more than 20 languages after one-off mother tongue-based training

90%+ Less Expensive than Humans

- One video shooting and training session creates a virtual avatar that is always ready to use. This avatar remains consistent and stable no matter the time, length, or amount of use. Lip and motion sync is so high that you cannot tell the difference between the virtual avatar and its human counterpart.

- The freedom of motion of a virtual avatar depends on the video used for training. Virtual avatars are less expensive and more efficient and stable than real humans.

More Realistic and Immersive Livestreaming

- Virtual streamers interact with objects, stream with real-world backgrounds, reply to on-screen comments, and allow takeover by a real person anytime. This makes the experience more engaging and immersive.

- Matrix management and advanced policy configuration make virtual avatar livestreaming more flexible.

Real-Time Intelligent Interaction for Enhanced Customer Satisfaction and Brand Engagement

You can quickly integrate open web SDKs and APIs into third-party service systems to complete tasks such as training and Q&A by AI virtual employees.

- Interaction at a latency of 1.5 seconds

  Low-latency interaction streamlines communication between users and virtual avatars.

- Capabilities open for integration

  - Device web SDK for intelligent interaction

  - APIs for avatar control

  - APIs of virtual human brain

Open Virtual Human Capabilities and Multiple Collaboration Models

- After completing real-name authentication, you can create content on the MetaStudio console.

- You can integrate virtual avatar APIs and SDKs into your own platforms.

- We partner with professional livestreaming services to develop solutions tailored to your needs. We also provide expert support for delivering and running virtual avatar livestreams.

Elevated Product Competitiveness

- More efficient virtual avatar production

  The platform supports simultaneous training of thousands of models to batch generate virtual avatars.

- Premium quality

  30%+ improvement in model quality

- End-to-end (E2E) service capabilities

- This solution can seamlessly integrate with other AIGC products to deliver E2E capabilities, ranging from virtual avatar video generation to editing.

Adoption by Many Industries and Scenarios

MetaStudio virtual avatars are qualified for many roles, such as brand ambassador, variety show host, live commerce streamer, news anchor, trainer, culture and tourism guide, customer service personnel, and marketing promoter.

Global Presence

MetaStudio, available in Huawei Cloud regions in the Chinese mainland, Singapore, and Ireland, empowers customers to accelerate their global business growth.

3 Functions

MetaStudio mainly provides the functions below. You can check if a function is available in a region on the console.

☷ NOTE

Image Modeling

Shoots a video of a human to generate a virtual avatar.

The virtual avatar can apply to video production, livestreaming, and intelligent interaction.

Note:

- You cannot download or export a created virtual avatar model to your local device.
- Virtual avatar models are not general ones and are, therefore, incompatible with third-party services. Models generated on MetaStudio can be used only within MetaStudio.
- Virtual avatars do not support clothes or face swap.
- If a training video used for image modeling includes choreography, you can see on the card of the generated virtual avatar, meaning you can add actions to this virtual avatar.

Voice Modeling

Uses a recorded human voice to generate a voice model.

The voice model can be used to dub virtual avatars in video production, livestreaming, and intelligent interaction.

Note:

- You cannot download or export a created voice model to your local device.
- Voice models are not general ones and are, therefore, incompatible with third-party services. Models generated on MetaStudio can be used only within MetaStudio.

Video Production

Uses a preset or custom virtual avatar image and voice to generate audio/video content.

The generated videos are applicable in a wide range of scenarios, such as teaching and training.

Livestreaming

Uses a preset or custom virtual avatar image and voice for livestreaming.

You can livestream on a platform by:

pushing RTMP streams from Huawei Cloud MetaStudio to a third-party livestreaming platform. You need to obtain an ingest URL from the platform. If the ingest URL is not available, you can start livestreaming through window capture.

Intelligent Interaction

Performs interactive Q&A between users and virtual avatars equipped with a third-party brain. Interactive virtual avatars are qualified for many roles, including shopping guide, culture and tourism guide, and customer service personnel.

> ⚠️ **CAUTION**
>
> The answers are given by the integrated third-party large language model (LLM) or knowledge base.

Asset Management

- You can upload models, PowerPoint files, animations, materials, videos, scenes, images, and music from a local device.
- Voice and model assets can be transferred to other tenants.
  This function is not available yet. To use it, **submit a service ticket**.
- Assets can be deleted.

4 Constraints

4.1 Service Usage

Before using MetaStudio, you need to learn about the **service constraints**.

Table 4-1 Service constraints

Constraint	Description
Region	MetaStudio is now available in the CN North-Beijing4 and AP-Singapore regions.
Preset image usage	The preset virtual avatars on the MetaStudio console are valid for three years (till the middle of 2027). The notice of bringing a preset image offline will be released on the live network three months in advance.
Violations	Do not publish non-compliant content on MetaStudio. For details, see **What Content Is Not Allowed on MetaStudio?**

4.2 APIs

Constraints on APIs

See **Table 4-2**.

Table 4-2 Constraints on APIs

Item	Constraint
Timbre	Before selecting a timbre by calling livestreaming or intelligent interaction APIs, check whether the timbre is available on the live room editing or dialog setting page of the MetaStudio console. The API calling will be successful only when the timbre is available on the page.
Intelligent interaction	The following three intelligent interaction API groups apply only to iFLYTEK AIUI and iFLYTEK Spark Cognitive Model: ● Virtual Avatar Knowledge Base Skill Management ● Virtual Avatar Knowledge Base Intent Management ● Virtual Avatar Knowledge Base Question Management

API Request Throttling

MetaStudio sets a limit on the number of API calls to prevent service interruption caused by repeated API calls in a short period of time.

> **NOTICE**
>
> If the error message **APIGW.0308** is returned when an API is called, the API request throttling is triggered. You need to adjust the API calling interval by referring to the following table. Alternatively, wait for a period of time and then send API requests.
>
> For details about error codes starting with **APIGW**, see **Error Codes**.

Table 4-3 API request throttling

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
Digital Asset Management	Creating an Asset	10 times per second	5 times per second	5 times per second
	Querying Assets	20 times per second	10 times per second	10 times per second
	Updating an Asset	10 times per second	5 times per second	5 times per second
	Deleting an Asset	10 times per second	5 times per second	5 times per second
	Querying Asset Details	20 times per second	10 times per second	10 times per second

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
	Restoring a Deleted Asset	10 times per second	5 times per second	5 times per second
	Batch Asset Operations	Once per second	Once per second	Once per second
	Querying Asset Summary	20 times per second	10 times per second	10 times per second
	Querying Asset Copy Information	20 times per second	10 times per second	10 times per second
	Copying an Asset	4 times per second	3 times per second	3 times per second
File Management	Creating a File and Obtaining the URL for File Upload	10 times per second	5 times per second	5 times per second
	Deleting a File	10 times per second	5 times per second	5 times per second
	Confirming the Upload of a File	10 times per second	5 times per second	5 times per second
	Creating a Large File	10 times per second	5 times per second	5 times per second
Video Script Management	Creating a Video Script	10 times per second	5 times per second	5 times per second
	Querying Video Scripts	20 times per second	10 times per second	10 times per second
	Querying Video Script Details	20 times per second	10 times per second	10 times per second
	Updating a Video Script	10 times per second	5 times per second	5 times per second
	Deleting a Video Script	10 times per second	5 times per second	5 times per second
	Copying a Video Script	10 times per second	5 times per second	5 times per second
Video Production Management	Querying Video Production Tasks	20 times per second	10 times per second	10 times per second

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
Virtual Avatar Video Production Management	Creating a Virtual Avatar Video Production Task	5 times per second	3 times per second	3 times per second
	Querying Virtual Avatar Video Production Task Details	20 times per second	10 times per second	10 times per second
	Canceling a Virtual Avatar Video Production Task in the Waiting Status	5 times per second	3 times per second	3 times per second
Virtual Avatar Video Subtitle File Management	Creating a Virtual Avatar Video Subtitle File	5 times per second	3 times per second	3 times per second
	Querying Virtual Avatar Video Subtitle File Details	20 times per second	10 times per second	10 times per second
Intelligent Live Room Management	Creating an Intelligent Live Room	5 times per second	3 times per second	3 times per second
	Querying Intelligent Live Rooms	20 times per second	10 times per second	10 times per second
	Querying Intelligent Livestreaming Script Details	20 times per second	10 times per second	10 times per second
	Updating Intelligent Live Room Information	5 times per second	3 times per second	3 times per second
	Deleting an Intelligent Live Room	4 times per second	3 times per second	3 times per second
	Creating an Interaction Rule Library for Live Rooms	4 times per second	3 times per second	3 times per second

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
	Querying Interaction Rule Libraries for Live Rooms	20 times per second	10 times per second	10 times per second
	Updating an Interaction Rule Library for Live Rooms	4 times per second	3 times per second	3 times per second
	Deleting an Interaction Rule Library for Live Rooms	4 times per second	3 times per second	3 times per second
Livestreaming Task Management	Starting a Virtual Avatar Intelligent Livestreaming Task	4 times per second	3 times per second	3 times per second
	Querying Livestreaming Tasks of a Live Room	20 times per second	10 times per second	10 times per second
	Querying Virtual Avatar Intelligent Livestreaming Task Details	20 times per second	10 times per second	10 times per second
	Ending a Virtual Avatar Intelligent Livestreaming Task	4 times per second	3 times per second	3 times per second
	Controlling Virtual Avatar Intelligent Livestreaming	4 times per second	3 times per second	3 times per second
	Querying All Virtual Avatar Livestreaming Tasks of a Tenant	10 times per second	5 times per second	5 times per second
	Reporting Livestreaming Events	4 times per second	3 times per second	3 times per second
Live Product Management	Creating a Product	4 times per second	3 times per second	3 times per second

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
	Querying Products	20 times per second	10 times per second	10 times per second
	Updating a Product	4 times per second	3 times per second	3 times per second
	Querying Product Details	20 times per second	10 times per second	10 times per second
	Deleting a Product	4 times per second	3 times per second	3 times per second
	Configuring a Product Asset Portfolio	4 times per second	3 times per second	3 times per second
Third-Party Live Platform Management	Creating a Third-Party Live Platform	Once per second	Once per second	Once per second
	Updating Third-Party Live Platform Information	Once per second	Once per second	Once per second
	Deleting Third-Party Live Platform Information	Once per second	Once per second	Once per second
	Querying Live Platforms	20 times per second	10 times per second	10 times per second
	Querying Third-Party Live Platforms	20 times per second	10 times per second	10 times per second
	Querying Products on a Third-Party Live Platform	5 times per second	3 times per second	3 times per second
Virtual Avatar Dialog Management	Creating an Interactive Dialog	5 times per second	3 times per second	-
	Querying Interactive Dialogs	20 times per second	10 times per second	10 times per second
	Querying Interactive Dialog Details	20 times per second	10 times per second	10 times per second
	Updating an Interactive Dialog	5 times per second	3 times per second	-

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
	Deleting an Interactive Dialog	5 times per second	3 times per second	-
Virtual Avatar Dialog Task Management	Creating a Dialog URL	5 times per second	3 times per second	-
	Starting a Virtual Avatar Intelligent Interaction Task	5 times per second	3 times per second	-
	Ending a Virtual Avatar Intelligent Interaction Task	5 times per second	3 times per second	-
	Querying Virtual Avatar Intelligent Interaction Tasks	20 times per second	10 times per second	10 times per second
Virtual Avatar Application Management	Creating an Application	5 times per second	3 times per second	-
	Querying Applications	20 times per second	10 times per second	10 times per second
	Querying Application Details	20 times per second	10 times per second	10 times per second
	Modifying an Application	5 times per second	10 times per second	-
	Deleting an Application	5 times per second	3 times per second	-
Virtual Avatar Authentication Code Management	Creating a One-Off Authentication Code	20 times per second	10 times per second	-
Virtual Avatar Activation Code Management	Creating an Activation Code	5 times per second	3 times per second	-
	Querying Activation Codes	20 times per second	10 times per second	10 times per second
	Querying Activation Code Details	20 times per second	10 times per second	10 times per second
	Modifying an Activation Code	5 times per second	3 times per second	-

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
	Resetting an Activation Code	5 times per second	3 times per second	-
	Deleting an Activation Code	5 times per second	3 times per second	-
Virtual Avatar Knowledge Base Skill Management	Creating a Knowledge Base Skill	4 times per second	3 times per second	3 times per second
	Querying Knowledge Base Skills	20 times per second	10 times per second	10 times per second
	Modifying a Knowledge Base Skill	4 times per second	3 times per second	3 times per second
	Querying Knowledge Base Skill Details	20 times per second	10 times per second	10 times per second
	Deleting a Knowledge Base Skill	4 times per second	3 times per second	3 times per second
	Exporting a Knowledge Base Skill	Once per second	Once per second	Once per second
Virtual Avatar Knowledge Base Intent Management	Creating a Knowledge Base Intent	4 times per second	3 times per second	3 times per second
	Querying Knowledge Base Intents	20 times per second	10 times per second	10 times per second
	Querying Knowledge Base Intent Details	20 times per second	10 times per second	10 times per second
	Modifying a Knowledge Base Intent	4 times per second	3 times per second	3 times per second
	Deleting a Knowledge Base Intent	4 times per second	3 times per second	3 times per second

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
	Creating a Knowledge Base Intent and Question	4 times per second	3 times per second	3 times per second
Virtual Avatar Knowledge Base Question Management	Creating a Knowledge Base Question	4 times per second	3 times per second	3 times per second
	Querying Knowledge Base Questions	20 times per second	10 times per second	10 times per second
	Querying Knowledge Base Question Details	20 times per second	10 times per second	10 times per second
	Modifying a Knowledge Base Question	4 times per second	3 times per second	3 times per second
	Deleting a Knowledge Base Question	4 times per second	3 times per second	3 times per second
	Batch Creating Knowledge Base Questions	Once per second	Once per second	Once per second
	Batch Modifying Knowledge Base Questions	Once per second	Once per second	Once per second
Virtual Avatar Greeting Management	Creating a Greeting	4 times per second	3 times per second	3 times per second
	Querying Greetings	20 times per second	10 times per second	10 times per second
	Querying Greeting Details	20 times per second	10 times per second	10 times per second
	Modifying a Greeting	4 times per second	3 times per second	3 times per second
	Deleting a Greeting	4 times per second	3 times per second	3 times per second

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
	Determining Whether to Configure a Greeting	4 times per second	3 times per second	3 times per second
	Querying Whether Greetings Are Enabled	20 times per second	10 times per second	10 times per second
Virtual Avatar Frequently Asked Question Management	Creating a Frequently Asked Question	4 times per second	3 times per second	3 times per second
	Querying Frequently Asked Questions	20 times per second	10 times per second	10 times per second
	Querying Frequently Asked Question Details	20 times per second	10 times per second	10 times per second
	Modifying a Frequently Asked Question	4 times per second	3 times per second	3 times per second
	Deleting a Frequently Asked Question	4 times per second	3 times per second	3 times per second
Virtual Avatar Keyword Record Management	Creating a Keyword Record	4 times per second	3 times per second	3 times per second
	Querying Keyword Records	20 times per second	10 times per second	10 times per second
	Querying Keyword Record Details	20 times per second	10 times per second	10 times per second
	Modifying a Keyword Record	4 times per second	3 times per second	3 times per second
	Deleting a Keyword Record	4 times per second	3 times per second	3 times per second
	Determining Whether to Configure a Keyword	4 times per second	3 times per second	3 times per second

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
	Querying Whether Keywords Are Enabled	20 times per second	10 times per second	10 times per second
Virtual Avatar Agency Management	Querying Agencies	20 times per second	10 times per second	10 times per second
	Creating an Agency	Once per second	Once per second	-
	Deleting an Agency	Once per second	Once per second	-
Image Modeling Management	Creating a Virtual Avatar Model Training Task	10 times per second	3 times per second	3 times per second
	Querying Virtual Avatar Model Training Tasks	20 times per second	10 times per second	10 times per second
	Querying Virtual Avatar Model Training Task Details	20 times per second	10 times per second	10 times per second
	Deleting a Virtual Avatar Model Training Task	10 times per second	3 times per second	3 times per second
	Updating a Virtual Avatar Model Training Task	10 times per second	3 times per second	3 times per second
	Executing a Virtual Avatar Model Training Task as a Tenant	10 times per second	3 times per second	3 times per second
Voice Modeling Task Management	Querying Personalized Audio Duration	20 times per second	10 times per second	10 times per second
	Creating a Voice Training Task (Basic Edition)	10 times per second	3 times per second	3 times per second
	Creating a Voice Training Task (Advanced Edition)	10 times per second	3 times per second	3 times per second

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
	Creating a Voice Training Task (Premium Edition)	10 times per second	3 times per second	3 times per second
	Querying Voice Training Tasks	20 times per second	10 times per second	10 times per second
	Submitting a Voice Training Task	10 times per second	3 times per second	3 times per second
	Querying Voice Training Task Details	20 times per second	10 times per second	10 times per second
	Deleting a Voice Training Task	10 times per second	3 times per second	3 times per second
	Querying Task Operation Logs	20 times per second	10 times per second	10 times per second
	Obtaining the URL for Uploading a Voice File	20 times per second	10 times per second	10 times per second
	Obtaining the Review Result of a Voice Training Task	20 times per second	10 times per second	10 times per second
	Setting the Task Batch	10 times per second	3 times per second	3 times per second
	Confirming the Online Recording Result	10 times per second	3 times per second	3 times per second
	Obtaining the Confirmed Online Recording Result	20 times per second	10 times per second	10 times per second
TTS Management	Creating a TTS Preview Task	5 times per second	3 times per second	3 times per second
	Obtaining a TTS Preview File	20 times per second	10 times per second	10 times per second
	Creating an Asynchronous TTS Task	5 times per second	3 times per second	3 times per second

API Category	API	Single Tenant (CN North-Beijing4)	Single Tenant (CN East-Shanghai1)	Single Tenant (AP-Singapore)
	Obtaining an Asynchronous TTS Task	20 times per second	10 times per second	10 times per second
	Configuring a Tenant-Level Custom Pronunciation for TTS	10 times per second	3 times per second	3 times per second
	Obtaining a Tenant-Level Custom Pronunciation for TTS	10 times per second	3 times per second	3 times per second
	Deleting a Tenant-Level Custom Pronunciation for TTS	10 times per second	3 times per second	3 times per second
	Modifying a Tenant-Level Custom Pronunciation for TTS	10 times per second	3 times per second	3 times per second

Table 4-4 API request throttling

API Category	API	Single Tenant	System-based
Tenant Management	Viewing Tenant Resource Usage	10 times per second	20 times per second
	Viewing Tenant Resources	10 times per second	20 times per second
	Collecting Statistics on Resource Quantity in a Time Period	10 times per second	20 times per second
Subscription Management	Subscribing to a MetaStudio Product	Once per second	10 times per second

4.3 Image Modeling

See **Table 4-5**.

Table 4-5 Constraints on image modeling

Constraint	Description
User	Only enterprise users can customize virtual avatar images on MetaStudio.
Usage	- You cannot download or export a created virtual avatar model to your local device. - Virtual avatar models are not general ones and are, therefore, incompatible with third-party services. Models generated on MetaStudio can be used only within MetaStudio. - Virtual avatars do not support clothes or face swap. - If a training video used for image modeling includes choreography, you can see on the card of the generated virtual avatar, meaning you can add actions to this virtual avatar.

4.4 Voice Modeling

See **Table 4-6**.

Table 4-6 Constraints on voice modeling

Constraint	Description
User	Only enterprise users can customize virtual avatar voices on MetaStudio.
Usage	- You cannot download or export a created voice model to your local device. - Voice models are not general ones and are, therefore, incompatible with third-party services. Models generated on MetaStudio can be used only within MetaStudio.
DupDub	You can use a preset third-party voice of DupDub after purchasing it. There is no constraint on the use of the voice.

4.5 Video Production

See **Table 4-7**.

Table 4-7 Constraints on virtual avatar video production

Video Production Setting	Constraint
Video script	- A video script can contain a maximum of 100 scenes. - A video script can contain only one virtual avatar image model and one voice model.
Text control	- A maximum of 10,000 characters are allowed in each scene. - A maximum of 100,000 characters are allowed in all scenes. - If a single scene contains Speech Synthesis Markup Language (SSML) tags, the text size must be less than 128 KB.
Speech control	Audio files uploaded in a single scene cannot be larger than 100 MB.
Video format	You can insert MP4, M4V, MKV, MOV, FLV, 3GP, WMV, AVI, and WebM videos. Note: On the video production page, if an FLV, 3GP, WMV, or AVI video is inserted, you cannot preview the video and can see only a preview image of the video. This is due to browser incompatibility. However, after video compositing, the inserted video can play seamlessly with the main video. Requirements on an inserted video: - Resolution ≤ 1080p - Frame rate ≤ 30 FPS - AV1, VP8, VP9, H.264, or H.265 encoding - Only WebM videos support VP8 and VP9 encoding. - If a WebM video uses AV1 encoding, the video cannot be composited. - Only the Chrome browser supports AV1 and H.265 encoding. - Video size < 1 GB - A maximum of two video overlays Precautions of video upload: - An uploaded video overlay cannot be modified. You can modify it only on your local device and then upload the new one to the console. - The aspect ratio of a video overlay is locked. You can adjust its width and height only on your local device and then import the new one. - A video overlay cannot exceed 30 minutes.

Video Production Setting	Constraint
Audio format	You can insert MP3, M4A, and WAV audios. Requirements on an inserted audio: ● Only mono audios can be extracted. ● The audio size should be less than 500 MB.
Image format	You can insert PNG, JPG, or JPEG images. Requirements on an inserted image: ● Resolution ≤ 1080p ● Image size < 500 MB
Subtitling	Punctuation marks (such as ,.:;!?...) in subtitles will be automatically removed. ● If the punctuation mark to be automatically removed is in the middle of a piece of text, a space is added after the punctuation mark is removed. ● If the punctuation mark to be automatically removed is at the end of a paragraph, it will be removed directly. Some punctuation marks (""()·~--) cannot be automatically removed.
Video production task	● Retention duration: Historical tasks can be retained for six months. After that, the tasks are not displayed on the **Video Production** page of **Task Center** on the console. ● Queue limit: A maximum of 20 tasks can be queued at the same time. ● Concurrency limit: A maximum of 20 videos can be composited at the same time.

Video Production Setting	Constraint
PowerPoint file	Constraints: • You can upload a PowerPoint file of no more than 100 slides. • If the font of the PowerPoint file cannot be displayed on the MetaStudio console, convert the PowerPoint file to a PDF file before importing it. For details, see **How Do I Convert a PowerPoint File to a PDF File?** • A PowerPoint file containing special effects such as animation cannot be displayed on the MetaStudio console. You are advised to export the PowerPoint file as a video and then import it. See **How Do I Present the Animation of a PowerPoint File?** • If a PowerPoint file fails the parsing and cannot be uploaded, rectify the fault by referring to **What If a PowerPoint File Cannot Be Uploaded Due to a Parsing Failure?**

4.6 Livestreaming

Constraints on virtual avatar livestreaming:

- **Constraints on Live Room Settings**
- **Ending a Livestreaming Task**
- **Live Video Playback**
- **Live Content Review**

Constraints on Live Room Settings

See **Table 4-8**.

Table 4-8 Constraints on live room settings

Live Room Setting	Constraint
Text control	• All paragraphs in each scene must contain more than 200 characters. • There must be fewer than 2,000 characters in each paragraph (fewer than 3,000 if SSML tags are contained). • All paragraphs in a live room must contain fewer than 300,000 characters if SSML tags are contained.

Live Room Setting	Constraint
Speech control	Audio files uploaded (longer than 1 minute in total) in a single scene cannot be larger than 100 MB.
Video format	You can insert MP4, M4V, MKV, MOV, FLV, 3GP, WMV, and AVI videos. Note: On the live room page, if an FLV, 3GP, WMV, or AVI video is inserted, you cannot preview the video and can see only a preview image of the video. This is due to browser incompatibility. However, during a livestream, the inserted video can play. Requirements on an inserted video: ● Resolution ≤ 1080p ● Frame rate ≤ 30 FPS ● AV1, H.264, or H.265 encoding Only the Chrome browser supports AV1 and H.265 encoding. ● Video sampling format: YUV420 ● Video size < 1 GB ● A maximum of three video overlays, including the videos inserted during the interaction setting Precautions of video upload: ● An uploaded video overlay cannot be modified. You can modify it only on your local device and then upload the new one to the console. ● The aspect ratio of a video overlay is locked. You can adjust its width and height only on your local device and then import the new one. ● Transparent videos can be overlaid only in video production, but not on a live video. An uploaded MOV video file with alpha channel (that is, a transparent video) will be transcoded into a WebM video file and stored in the MetaStudio asset library. The livestreaming service allows importing a transparent MOV video file, but the imported MOV video file will be transcoded into a WebM one. WebM videos cannot be overlaid on a live video.
Audio format	You can insert MP3, M4A, and WAV audios. Requirements on an inserted audio: ● Only mono audios can be extracted. ● The audio size should be less than 500 MB.

Live Room Setting	Constraint
Image format	You can insert PNG, JPG, JPEG, or BMP images. Requirements on an inserted image: • Resolution ≤ 1080p • Image size < 500 MB
Subtitling	Punctuation marks (such as ,.:;!?...) in subtitles will be automatically removed. • If the punctuation mark to be automatically removed is in the middle of a piece of text, a space is added after the punctuation mark is removed. • If the punctuation mark to be automatically removed is at the end of a paragraph, it will be removed directly. Some punctuation marks (""()·~--) cannot be automatically removed.
Transparent background	Allowed only for video production and intelligent interaction

Ending a Livestreaming Task

See **Table 4-9**.

Table 4-9 Ending a livestreaming task

Scenario	Operation
Ending a livestreaming task as a tenant	A tenant can end a livestreaming task: • on the livestreaming page of the MetaStudio console; • alternatively, by calling the API for **ending a virtual avatar livestreaming task**.
Automatically ending a livestreaming task	To protect tenants' interests, a livestreaming task will automatically end when: • there is no RTMP stream push by the MetaStudio console and the playback window has been closed for about two to three minutes; • APIs are called for livestreaming, there is no RTMP stream push, and viewers have left the RTC room for about two to three minutes.
Freezing a livestreaming task	If the live content involves compliance risks, the service O&M personnel directly freeze the livestreaming task on the O&M page.

Live Video Playback

See **Table 4-10**.

Table 4-10 Live video playback

Live Video Playback	Operation
Controlled by a tenant	MetaStudio allows controlling livestreams so that streamers can capture a complete live video.
	Before starting livestreaming, deselect **Automatic livestreaming after loading** on the console (that is, set **play_policy.auto_play_script** to **FALSE**). A livestream will not start until the tenant's instruction of starting livestreaming is sent (for API calling, the **play_start** instruction is sent through the **RTSA COMMAND** instruction). Before the instruction is sent, the virtual streamer remains silent.

Live Content Review

See **Table 4-11**.

Table 4-11 Live content review

Content Review	Operation
Live room content review	Content review on MetaStudio:
	• By default, each configuration item of a live room, such as the live room name, description, text, audio, video, and image, will be automatically reviewed to eliminate risks of violation and avoid generating illegal live content.
	• If a tenant has used another content risk control system and is committed to content risk control, the tenant can **submit a service ticket** to disable content review on MetaStudio.
	• If complaints about live content are received, MetaStudio will freeze the live rooms and livestreaming tasks that violate the regulations in accordance with applicable laws and regulations. When necessary, all resources of the tenant will be frozen.

4.7 Intelligent Interaction

Table 4-12 lists the minimum hardware configuration requirements for intelligent interaction.

Table 4-12 Hardware device requirements

Hardware	Minimum Configuration
CPU	i5 or later
Memory	More than 8 GB
Disk	More than 128 GB
OS	Windows or Android
Microphone	Voice pickup within 1 meter
Speaker	Required
Touchscreen	Required
Bandwidth	More than 4 Mbit/s

4.8 Asset Management

See the following sections:

- **Asset Storage Billing**
- **Asset Creation Failure**
- **Asset Freezing**

Asset Storage Billing

See **Table 4-13**.

Table 4-13 Asset storage billing

Asset Storage	Billing Description
Free storage space	MetaStudio provides 5 GB asset storage space for free for each tenant to upload audios, videos, and images. If the tenant needs more storage space, they need to buy a monthly asset management package.
Customized image model storage	The storage space occupied by the timbre models of a customized image model on MetaStudio is not included in the tenant's asset storage billing.
Customized video asset storage	The storage space used by a customized video asset on MetaStudio will not be charged in the first six months. The billing starts after six months.
Asset storage billing	If an asset is not completely deleted, is frozen or in the recycle bin, billing of its storage space will continue.

Asset Creation Failure

During asset creation, the asset status will be automatically set to **Failed** when:

- The main asset file is not uploaded within 24 hours.
- Review of the asset title, description, and main file content failed.
- The main asset file does not match the created file information.

Asset Freezing

See **Table 4-14**.

Table 4-14 Asset freezing

Scenario	Operation
Freezing tenant assets	MetaStudio is jointly and severally liable for the risks of a tenant's asset content. If complaints about asset content are received, MetaStudio will freeze non-compliant assets in accordance with applicable laws and regulations. When necessary, all resources of the tenant will be frozen. Frozen assets cannot be used as materials for content generation. Tenants can delete frozen assets or apply for unfreezing assets.
Requesting for unfreezing an asset	You can submit a request for unfreezing an asset. After the asset is unfrozen, its asset status will become the one before it was unfrozen. If an asset in the recycle bin is unfrozen, its asset status will become the one before it was deleted.

5 Concepts

Virtual Human

A virtual human exists in a digital space and has a realistic, human-like appearance and behavioral characteristics, as well as certain intelligence and emotional capabilities that enable it to interact with people and express itself. A virtual human is also known as a virtual image or virtual avatar. The core technologies behind virtual humans include computer vision (CV), computer graphics, motion capture and control, image rendering, and AI.

A virtual support personality is a hyper-realistic virtual avatar built on composited images based on deep neural networks. It has the following characteristics:

- 2D model, which is generated after training a human video
- No facial expression & skeletal data
- AI-controlled only
- Specified facial expressions and motions

Script, Image Layout, and Settings

Video compositing is required for video production and livestreaming. The following uses livestreaming as an example.

- Video materials: as shown in area 1 in the figure, including templates, roles, backgrounds, overlays, videos, products, and text. You can use the preset materials or customize materials.
- Image layout: as shown in area 2 in the figure. You can insert roles, backgrounds, overlays, text boxes, and videos into the image layout, and adjust the size and layer position of each material. All settings are visible in the image layout.
- Livestream settings: as shown in area 3 in the figure
- Script: Script design, including material preparations, layout design, and livestream settings, is required for video compositing to produce videos and livestream.

Figure 5-1 Page for editing a live room
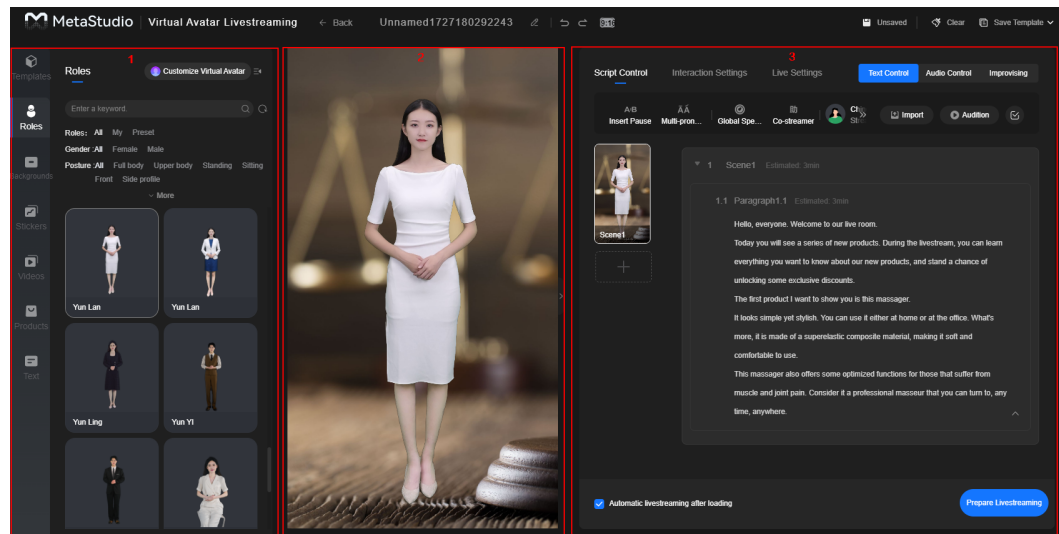
Role

Virtual avatar model

Action

Action of the virtual avatar model

6 Billing Overview

You can choose the yearly/monthly, pay-per-use, or one-off payment billing mode.

Billing Items

The billing items of MetaStudio include virtual avatars.

For details, see **Table 6-1**.

On the MetaStudio console, click **Buy Resource** on the **Overview** page to purchase required resources. For details, see **Purchase Guide**.

Table 6-1 Billing items

Category	Billing Item	Description	Billing Mode
Virtual avatar	Image modeling	You are charged based on the number of created virtual avatar images. Each time a virtual avatar image is created, one quota will be deducted from your purchased resource package. **CAUTION** If you delete the image asset created using a purchased resource package with auto-renewal enabled, you need to cancel auto-renewal first. A message will be displayed in the warning dialog box when you delete an image asset. Perform operations as prompted. Assume that you purchase a one-year image modeling resource package: ● You can create only one image asset within the validity period (one year). ● After one year, the purchased resource package will be frozen. If you do not renew after the package is frozen, the image asset created using the package will also be frozen and unavailable. ● If the grace period and retention period expire after the one-year validity period, the billed resources will be deleted, along with the image asset created using these resources.	Yearly

Cate gory	Billing Item	Description	Billing Mode
	Voice modeling	You are charged based on the number of created voices. Each time a virtual avatar voice is created, one quota will be deducted from your purchased resource package. **CAUTION** If you delete the voice asset created using a purchased resource package with auto-renewal enabled, you need to cancel auto-renewal first. A message will be displayed in the warning dialog box when you delete a voice asset. Perform operations as prompted. Assume that you purchase a one-year voice modeling resource package: ● You can create only one voice asset within the validity period (one year). ● After one year, the purchased resource package will be frozen. If you do not renew after the package is frozen, the voice asset created using the package will also be frozen and unavailable. ● If the grace period and retention period expire after the one-year validity period, the billed resources will be deleted, along with the voice asset created using these resources.	Yearly

Cate gory	Billing Item	Description	Billing Mode
	Video production	You are charged based on the duration of a virtual avatar video. To enable pay-per-use billing, go to the **My Resources** area of the MetaStudio console. Click **Enable Pay-per-Use** in the **Virtual avatar video production** card. In the **Information** dialog box displayed, agree to the agreement. Click **OK**. Note: ● The duration of a composited video is expressed in second. The actual charging will convert the duration from seconds to minutes. The value is accurate to two decimal places and the last digit is rounded off. ● Even if a composited video is frozen due to inappropriate content, the video production fees are still charged.	Pay-per-use
	Livestream ing	The subscription entry is not open. You can use virtual avatar livestreaming only when: ● you have **submitted a service ticket** for enabling virtual avatar livestreaming; ● the administrator has allocated livestreaming resources to you. Note: If the number of concurrent livestreaming tasks exceeds the allocated quota, the excess tasks cannot be executed due to insufficient resources.	Allocated by the administrator

Cate gory	Billing Item	Description	Billing Mode
	Intelligent interaction	You can purchase multiple dialogs. Each dialog is free of charge within the validity period of the yearly/ monthly subscription. Note: • The actions of a virtual avatar for intelligent interaction are generated after training based on the video of its human counterpart. The virtual avatar (1080p) features precise lip sync and seamless fusion with the background. • Intelligent interaction supports multiple knowledge bases (in both Chinese and English) to meet different needs. • Equipped with trained avatar and voice models, MetaStudio's interactive virtual avatars can use knowledge bases to interact with users in real time. • The more dialogs you purchase, the more interactive dialogs can be performed with a virtual avatar simultaneously. • Resources will become unavailable after the validity period expires, so we recommend enabling auto- renewal. • If the number of concurrent intelligent interaction tasks exceeds the allocated quota, the excess tasks cannot be executed due to insufficient resources.	Yearly/Monthly

Billing Modes

You can choose either the pay-per-use or yearly billing mode.

- Pay-per-use: You pay as you go and just pay for what you use. In each billing cycle, the fee is calculated by multiplying the actual usage of each resource by its unit price. Then, this fee is taken from your account balance.

- Yearly/Monthly: a prepaid billing mode. You pay in advance for a subscription term, and in exchange, you get a discounted rate. The longer the subscription
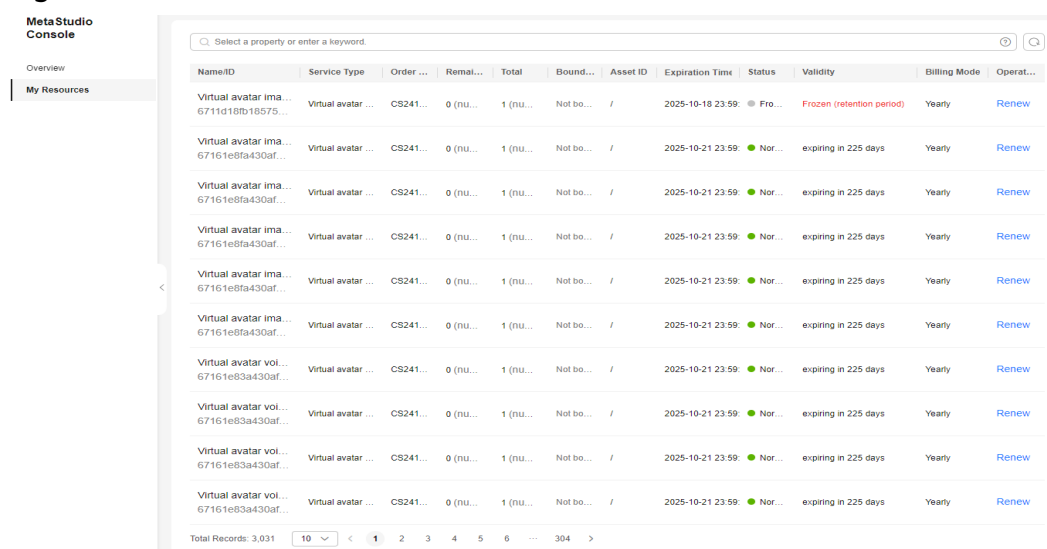
term, the bigger the discount. This mode is a good option for long-term, stable services.

Resource Validity Period

Package billing starts from the effective time after the order is paid. You can use the resources in the package within the validity period. With pay-per-use billing, you are charged based on the actual resource usage and there is no validity period limit.

You can view the expiration time of the resource package in **Overview** > **My Resources** of the MetaStudio console, as shown in **Figure 6-1**.

Figure 6-1 Resource list

Renewal

The yearly billing mode allows for renewal before the quota is used up. To renew the subscription, go to the **Renewals** page. If the renewal fails, the fee will be deducted from the account balance.

Unsubscription

- Yearly packages for virtual avatar image modeling and voice modeling cannot be unsubscribed.
- On the service overview page, if pay-per-use billing is enabled, you can disable it.
  - Enabling pay-per-use billing: Click **Enable Pay-per-Use** in the resource card. In the dialog box displayed, agree to the agreement. Click **OK**.
  - Disabling pay-per-use billing: Click **Disable** in the resource card. In the dialog box displayed, click **OK**.

Expiration and Arrears

- Arrears: Prepaid package resources do not expire. With pay-per-use billing, fees are deducted based on each billing item in its billing cycle. Insufficient

account balance will cause fee deduction of the previous billing cycle to fail, making your account in arrears. If your account is in arrears, your cloud service resources will enter a grace period and probably a retention period. For more information, see **Resource Suspension and Release**.

You can view the arrears information of your account and **pay the arrears** in **Billing Center**.

Complete the renewal during the retention period. If the service is in arrears and enters the retention period, the service will be unfrozen after the renewal.

- Service expiration: Pay-per-use resources do not expire. If a prepaid resource package expires, the remaining resources will be cleared. Since then, you will be billed for subsequent resource usage on a pay-per-use basis.

7 Permissions Management

If you need to grant different permissions to employees in your organization to access your purchased MetaStudio resources, Identity and Access Management (IAM) is a good choice. IAM provides identity authentication, fine-grained permissions management, and access control, securing the access to your cloud resources.

With IAM, you can create IAM users and grant them permissions to access only specific resources. For example, you can create IAM users for some software developers in your organization to allow them to use MetaStudio resources but not to delete resources.

If your account does not need individual IAM users for permissions management, skip this section.

IAM is a free service. You pay only for the resources in your account. For more information about IAM, see **IAM Service Overview**.

MetaStudio does not support enterprise projects.

MetaStudio Permissions

By default, new IAM users do not have permissions granted. You need to add a user to one or more groups, and attach permissions policies or roles to these groups. Users inherit permissions from the groups to which they are added and can perform specified operations on cloud services based on the permissions.

MetaStudio is a project-level service deployed and accessed in specific physical regions. To grant MetaStudio permissions to a user group, specify the scope as region-specific projects and select projects (such as **ap-southeast-2**). If you select **All projects**, the permissions will take effect for user groups in all region-specific projects. When accessing MetaStudio, the users need to switch to a region where they have been granted permissions to use this service.

You can grant users permissions by using roles and policies.

- Roles: A type of coarse-grained authorization mechanism that defines permissions related to user responsibilities. This mechanism provides only a limited number of service-level roles for authorization. When using roles to grant permissions, you also need to assign other roles on which the permissions depend to take effect. However, roles are not an ideal choice for fine-grained authorization and secure access control.

- Policies: A fine-grained authorization strategy that defines permissions required to perform operations on specific cloud resources under certain conditions. This mechanism allows for more flexible policy-based authorization, meeting requirements for secure access control.

  See **Table 7-1**.

Table 7-1 System-defined policies of MetaStudio

Policy Name	Description	Type	Policy Content
MetaStudio FullAccess	Administrator permissions of MetaStudio. This policy contains all permissions of MetaStudio.	System-defined policy	**MetaStudio FullAccess Policy**
MetaStudio ReadOnlyAccess	Read-only permissions of MetaStudio. Users granted these permissions can only view MetaStudio data.	System-defined policy	**MetaStudio ReadOnlyAccess Policy**

Helpful Links

- **IAM Service Overview**
- **Creating a User and Granting MetaStudio Permissions**

MetaStudio FullAccess Policy

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

MetaStudio ReadOnlyAccess Policy

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