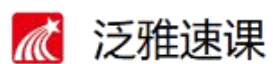


数字人速课 使用手册  
综合简版



超星集团  
2025 年 02 月

## 目 录

一、 平台使用前准备内容 .....	3
1. 数字人采样 .....	3
2. 数据筛选及处理 .....	6
3. PPT 及讲稿整理 .....	6
二、 平台使用流程说明 .....	9
数字人( <a href="http://chaoxing.com">chaoxing.com</a> ) .....	9
泛雅速课( <a href="http://chaoxing.com">chaoxing.com</a> ) .....	9
1. 用户登录 .....	9
2. 数字人定制 .....	9
3. 创建速课 .....	12

超星数字人是基于人工智能技术和计算机图像学技术，利用高清视频渲染、深度语言学习等技术为教师创建定制的 2D 高仿真虚拟数字人或计算机模拟虚拟形象。超星数字人联合解决方案以创新教学服务模式为着力点，深度优化服务人员工作流程，实现时间与空间的双重压缩，有效节省人力与物力成本，达成整体效率的显著提升，为教育服务领域的提质增效开辟全新路径。

仿真数字人以 10 分钟实拍为基准，利用人工智能技术和自然语言处理能力，将知识资源分享以数字人的方式代替原始人工录制，再通过合成素材、场景、讲稿等素材，快速生成优质视频并通过线下场馆或是线上融媒体进行传播，丰富学校资源库建设，为教学提供更多丰富知识存储。

本文将按照使用顺序，进行功能介绍及注意事项说明。

## 一、平台使用前准备内容

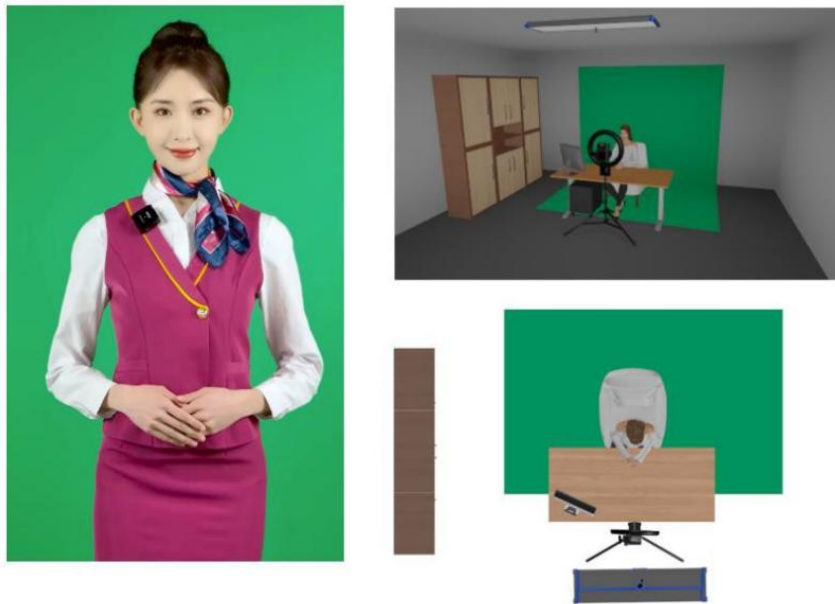
### 1. 数字人采样

#### (1) 影棚相关

**背景：**可选择实景或者绿幕背景，实景与绿幕拍摄，均支持抠除背景，后续使用数字人可替换为其他背景。实景拍摄需要保持头发边缘整齐，背景整洁，与着装及发色形成明显反差，建议纯色或浅色，背景要静态不能动态。

**光线：**可以选择通用光或者立体光，通用光难度低，对灯光设备要求低；立体光需要有多个补拍设备，所拍人物形象更立体。

**环境：**确保安静、无杂音，拍摄时除了读稿，不得有其他声音。



**录制设备：**需具备 4k 拍摄能力，具备 10bit422 录制能力，具备帧内编码录制能力的机器更佳，高解像力镜头，高速存储卡，稳定的三脚架等辅助设备及隔音箱。



## (2) 形象声音采集

原始录制素材通常包括三部分：

- a) 驱动视频：2-3 分钟有手势动作的视频；
  - b) 训练视频：12 分钟视频，开头结尾保留 30s 静默数据，中间动作自然流畅，视频画面连续；
  - c) 训练音频：12 分钟音频，建议在专业录音棚录制，避免混响，杂音；
- \*驱动视频和训练视频中被采集人物形象、衣着、造型、站位尽量保持一致。

## 声音采集环节注意事项——

**录音环境：**请在安静小房间录制，关闭门窗、空调等易产生噪声的声源。不宜在会议室、卫生间、空旷的房间录制(会产生混响/回声)。

**录音设备：**建议有线耳机或麦克风，不可使用蓝牙耳机，不要距离太远。

**录音姿势：**麦克风/有线耳机等收音设备和口部相距 1 拳距离。

**音频格式：**离线提供音频的客户注意，不要保存 .mp3、采样率不要低于 24000。

## 声音训练数据规范——

针对离线录音，小样本 pttts 不能少于 80 句，标准定制不低于 10 分钟。

**质量：**采样率不低于 24kHz，32k/48k 等均可，不支持类似 16k、8k，位深不低 16bits。

**格式：**wav 格式，不支持其他 mp3、pcm、aac 等。

**时长：**有效音频 10min 以上，可考虑多录，效果更佳，20 分钟定制精品小样本。

**长度：**不支持一长段音频训练，必须切分为单句（句子长度按照正常说话语句停顿）。

### (3) 讲述人须知

**形象妆造：**建议穿有垂感、有腰身的衣服；避免穿条纹、细格、与背景颜色相近的衣服，适合哑光妆面，构图上建议竖屏正面拍摄、推荐近景半身；建议渗透仪边缘不卡边、不出画；头部保持在画面中间 1/3 处，不要太近或者太远；站姿坐姿均可，处于自然舒展的状态效果更好。

## 拍摄前准备——

请需被采集对象准备 1800 字左右（约十分钟）的讲稿，讲稿内容尽可能丰富，包含概念、理论和示例，能充分展示被采集对象的讲解状态为佳。

## 拍摄要点——

**读稿流程：**【保持静默 15S】 --> 【每读一句稿，停顿 2-4S】 --> 【读完稿子，结束】

**发声：**张嘴即发声，避免嘴巴张开却没有声音。

**情绪：**情绪饱满，保持微笑读稿。

**停顿：**每句话读完，嘴部闭合时，自然闭嘴即可，避免抿嘴、露牙。

**动作：**核心是自然，保持自然说话时做动作的状态即可。

### 录音要点——

- 1) 吐字清晰，自然流畅，语速不要忽快忽慢、音量不要过大过小。
- 2) 录音过程中，适当阅稿及换气，不要太急促，显得不自然。
- 3) 不要晃动身体和椅子，容易产生杂音。

## 2. 数据筛选及处理

### (1) 筛选驱动视频【用于形象和动作训练】

对驱动视频素材中的手势动作、脸部表情、身体晃动进行筛选评估，筛选5分钟以上连续动作片段进行处理，确保衔接点顺畅，无明显拼接。确保无误后将视频压缩在2G以内输出保存。

训练视频剪辑完成后检查：

- a) 视频中被采集人物没有出框部位；
- b) 背景声音没有混响；
- c) 没有除被采集人物外的其他声音；
- d) 人物没有大幅移动，导致侧脸等画面而影响面部数据采集；
- e) 数据最后一帧无黑屏现象，确保一直有画面。

### (2) 剪辑训练音频【用于声音训练】

确保训练音频中无其他杂音，无混响。确保无误后进行降噪处理。剪辑选择20秒以内的有效数据，语句停顿自然，发音清楚。

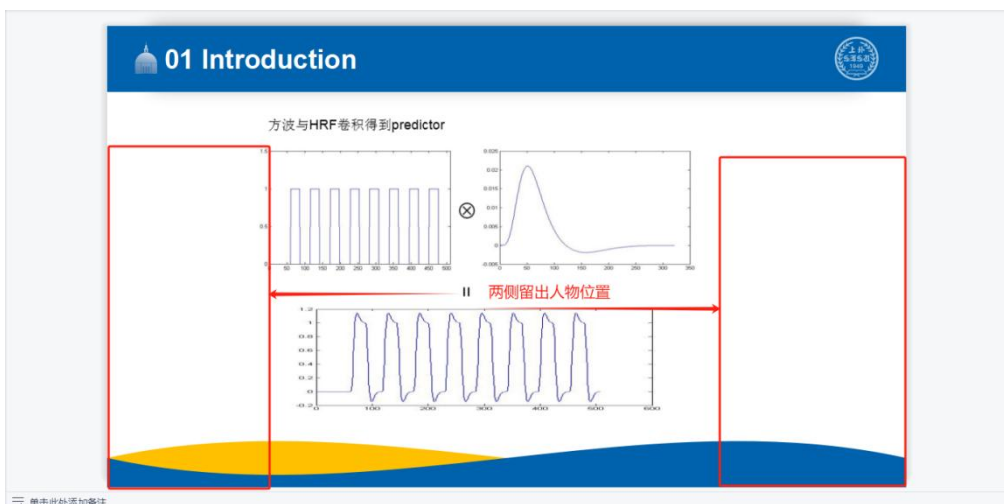
### (3) 再次检查，上传训练

再次检查训练视频和训练音频，检查无误后上传平台训练。

## 3. PPT 及讲稿整理

### (1) PPT 大小及比例

- a) 建议 PPT 大小【<200MB】，文件越大解析时长越久；
- b) 文件默认制作及输出比例为【16:9】，其余非适配尺寸有画幅拉伸；
- c) 如果需要出现人像，可以集中空出 PPT 左侧或右侧【1/4-1/3】的位置。

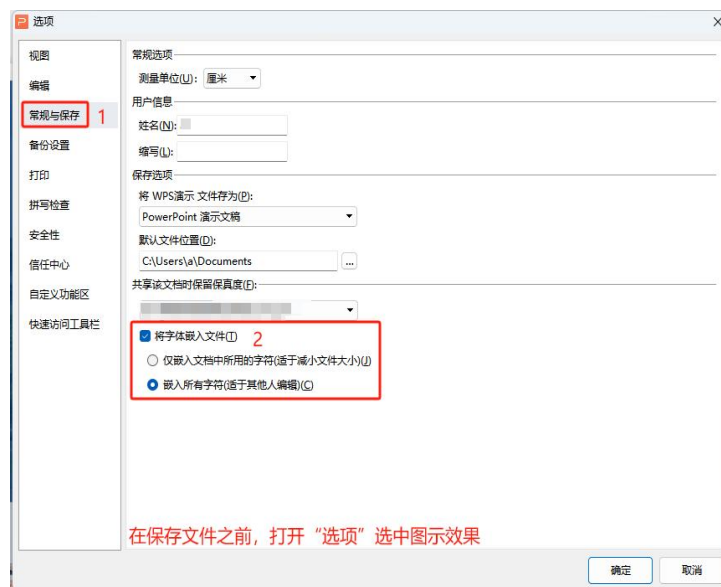


## (2) 切页及内页动画

PPT 内页动画均支持解析至时间轴内，可编辑查看，【拖动时间轴顺序及长短，可自由调整动画时长及效果】。

## (3) 特殊字体等

文件中的字体形式多样，为避免合成视频中存在识别误差，请在保存 PPT 时进行如下操作：



考虑到版权字体等使用情况，建议用户 PPT 文件内容在版权使用范围内，减少艺术字等的使用。

## (4) 音视频富媒体文件

音视频解析后默认保留原始时长，为后期合成效果自然，应选择【播放-自动】，详细位置处理可以拖动时间轴，避免被数字人讲话段落交叉重叠。



### (5) 讲稿及备注

讲稿内知识点表述应当口语化，过于书面的报告方式会以 100%还原的效果合成语音朗读，在段落间应调整好字符标点等，以使语音输出自然。

中英文混合的段落，涉及到字母缩写等需要手动编辑停顿，其余常规英文词组可正常发音。

涉及到特殊公式，或其他理工学科的特殊字符，建议编辑文稿中选择替换发音，以使语音输出准确。

The image shows a presentation slide titled '教育部公布的18个“人工智能+高等教育”应用场景典型案例' (18 Typical Cases of 'AI+Higher Education' Application Scenarios Announced by the Ministry of Education). The slide contains a table with 18 cases. The 10th case, '东南大学 大学物理课程智慧AI助教系统' (Southeast University University Physics Course Smart AI Tutor System), is highlighted with a red box. The 16th case, '西安电子科技大学 打造AI赋能身份新模式, 启动教学质量提升新引擎' (Xi'an Jiaotong University Creating a New Mode of AI-empowered Identity, Starting a New Engine for Quality Improvement), is also highlighted with a red box. Below the table, there is a red-bordered text box containing the following text:

在今年四月份，首批18个应用场景典型案例也进行了公示，其中教育东南大学AI助教系统和西安电子科技大学AI赋能智慧督导两个案例。这两个案例都是超星协助学校进行建设和申报的。那么，AI赋能下可以实现智慧教学，将会对我们的教学带来很大改变。对于老师来讲，可以为老师减负，让老师从原本繁重的教学或答疑工作中解脱出来，同时借助AI手段实现更加精准的教学，可以更好地了解学生的学情，对学生的一些学习成果进行更好的分析，更好的实现因材施教。从学生的角度来看，也可以实现更加个性化、自适应的学习模式，能够随时随地进行学习，向老师或AI来提问，同时也可以虚拟知识图谱、智慧课程来实现跨课程、跨专业的学习，更好地实现复合型人才培养和拔尖人才的培养。

准备ppt每页对应文稿

## 二、平台使用流程说明

[数字人\(https://ai-man.chaoxing.com/front/figureView\)](https://ai-man.chaoxing.com/front/figureView)

超星数字人:可在此定制教师形象及声音,将符合条件音视频素材上传后即可生成专属 AI 数字人,并将其与用在授课视频制作/讲座/宣传/读书分享等多个播报场景。

[泛雅速课\(https://qk-course.chaoxing.com/front/workList\)](https://qk-course.chaoxing.com/front/workList)

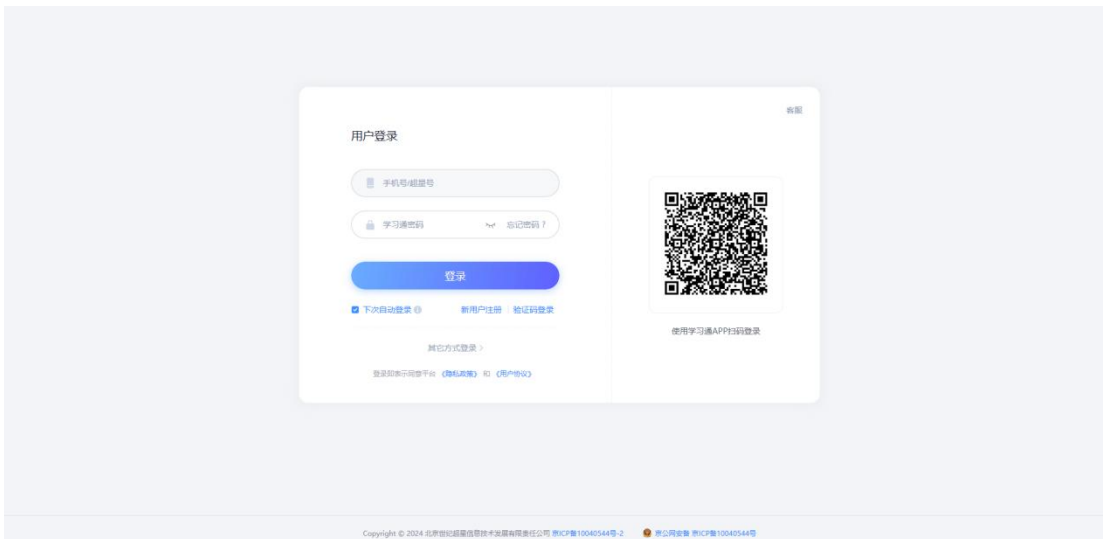
泛雅速课:PPT+数字人=课程视频。可利用平台提供的标准形象及用户定制形象,以及上传的 PPT 及备注文稿,快速高效合成授课内容。

平台使用前,需综合统计用户信息(姓名、账户 ID、所属单位、学工号等基本条目),统一登记授权。

### 1. 用户登录

进入以上网站后使用学习通账号密码、APP 扫码、手机验证码登录(未注册用户可以进行新用户注册)

登陆完成后进入使用页面(未授权无法使用)



### 2. 数字人定制

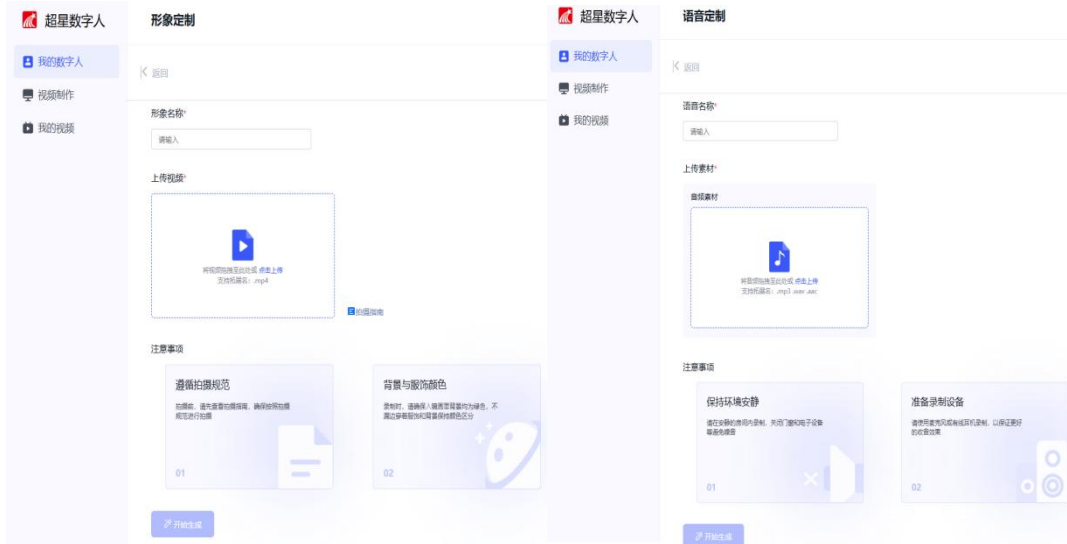
将上述所采集筛选后的数据(视频和音频)导入平台,在平台上自动训练,完成训练后即可生成指定人物形象,为后续视频制作所使用。

#### (1) 形象定制

填写形象名称并确保您的拍摄数据是绿幕等背景,保证您的数据为 5 分钟以

上有有效数据，点击提交，等待数据上传至 100%且在平台显示生成中，方可退出页面。一般定制时长约 18 小时，克隆完成后可在视频制作页面体验效果。

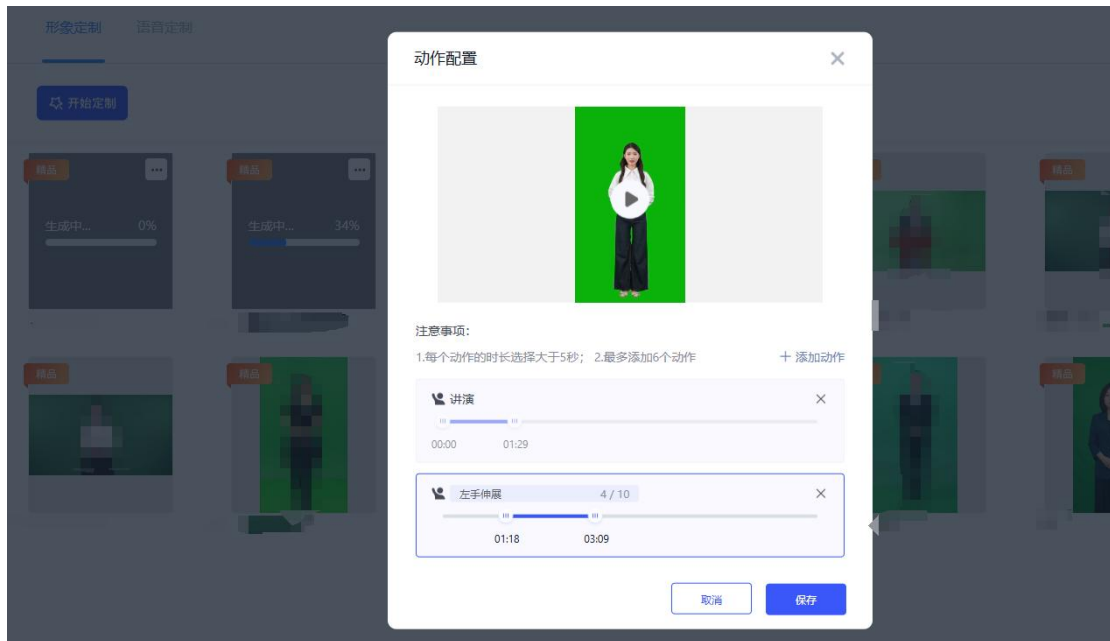
0%的状态即为排队中。用户默认可定制自身有效数据 1 个。



## (2) 动作配置

已经生成好的视频，可以在此将记录好的动作节点进行动作配置，每个形象可以定制 6 个动作片段，每个动作的时长选择需要大于 5 秒。

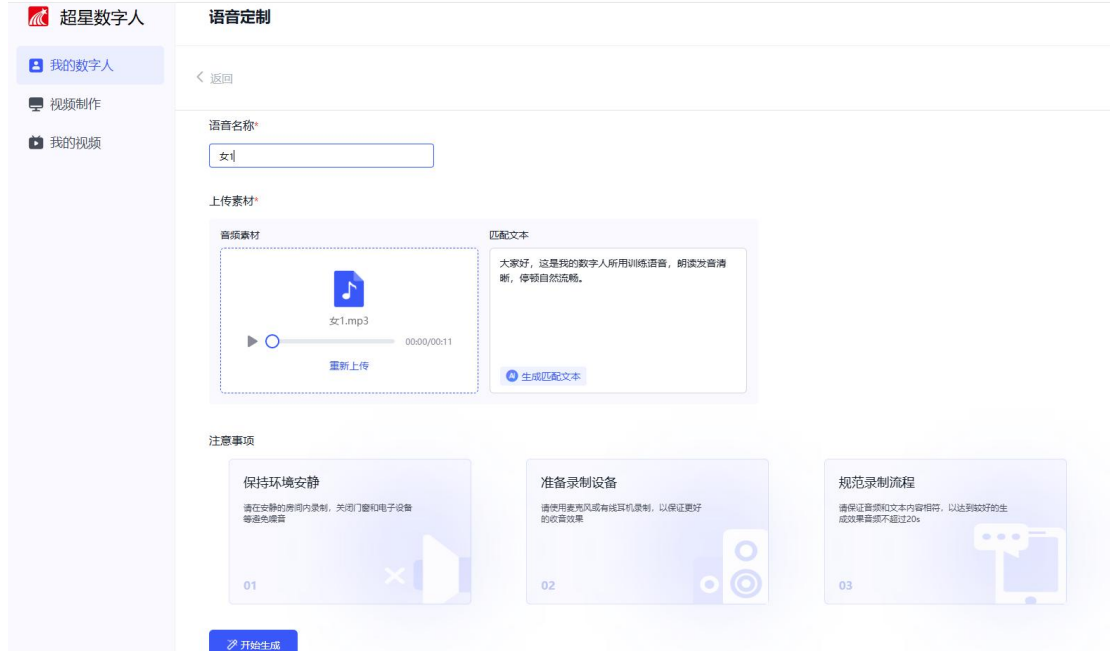
配置好的动作将会同步在 泛雅速课 的形象选择中，以便制作不同场景的视频文件。



## (3) 声音定制

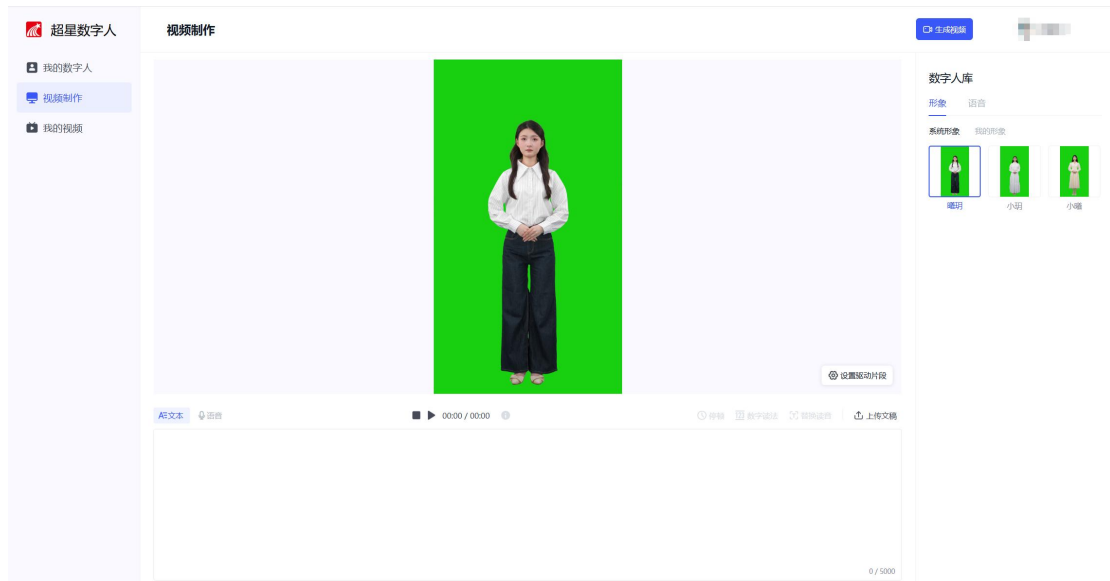
音频不超过 20s，请保证音频和文本内容相符，以达到较好的生成效果，上传后三分钟即可悬浮鼠标点击试听效果。

支持形象和声音绑定。

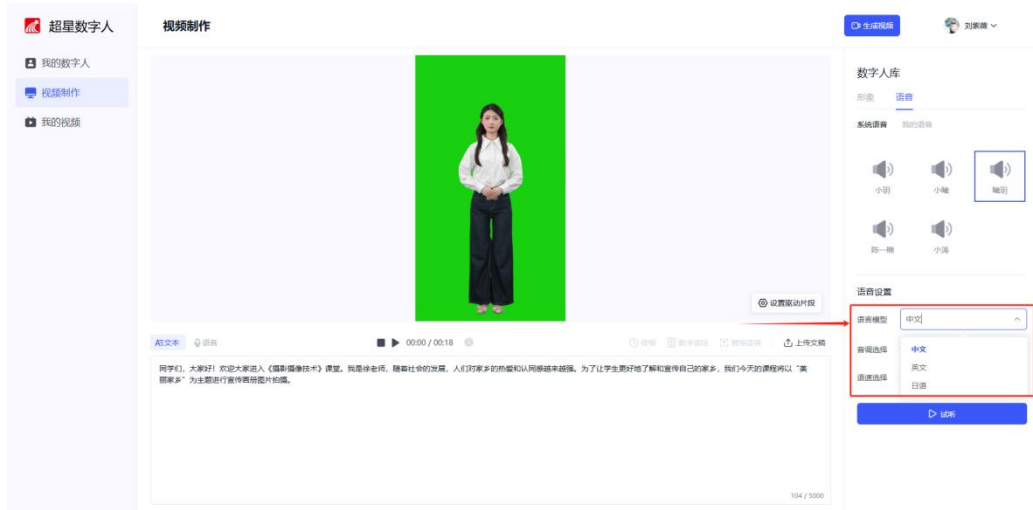


#### (4) 视频制作

选中制定教师形象及语音，将准备好的文稿输入其中，并执行增加停顿、替换发音、数字读法标记等常规编辑，及时试听，确保文稿输出准确，完成编辑后，点击生成视频，即可获得所需绿幕背景文件。生成好的视频文件在“我的视频”中查看。



语音模型功能不支持自动翻译，为提高匹配文稿的对应语音进行选择。  
文件支持重命名、删除及重新编辑操作。

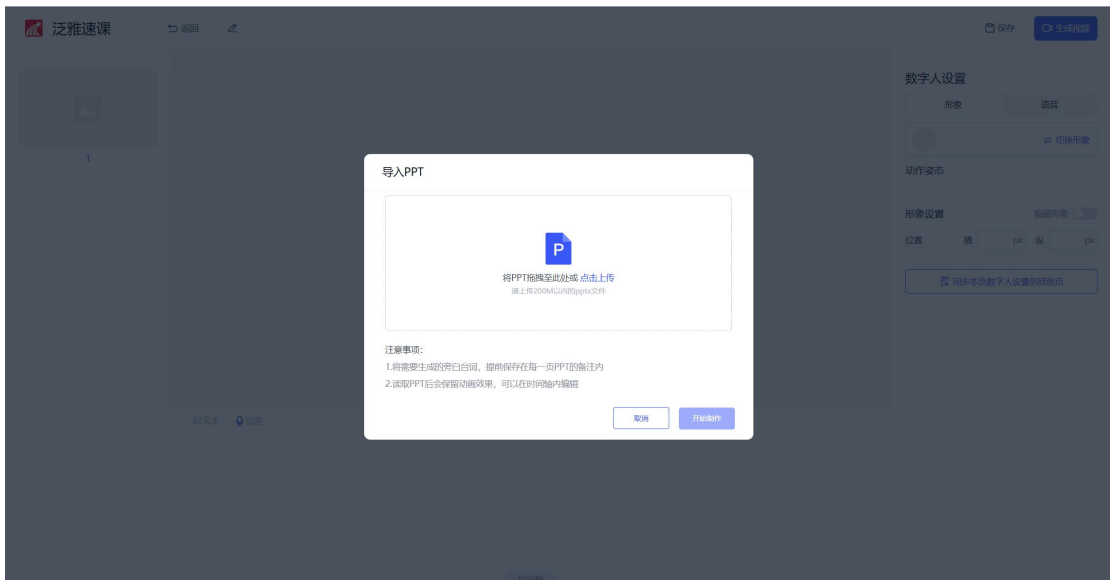


其中，重新编辑功能支持打开原有项目，并在重新保存生成后依旧保留原有项目，新生成内容不会覆盖原有文本内容，方便用户多种选择，在原有文本基础上做新改动。



### 3. 创建速课

进入速课网站并完成登录后，点击右上角【创建速课】按钮导入 PPT。



导入并解析准备好的 PPT 文件，解析后保留备注文稿及切页动画，进入工作台编辑页面。

用户预览左侧文件页面及详情页备注文字，确保解析数据完整，后进行数字人设置。

## (1) 数字人设置

### ① 形象选择

选中所需要展示的数字人形象，并在页面中调节数字人形象的大小位置，将所需要的数字人形象配置的动作添加到对应页面。

### ② 声音选择

声音与形象并不绑定，需要单独选中当前老师的语音数据，并试听文本内容，根据文件需要调整语音语调和语速，根据输入文本信息，选择对应语音模型。

### ③ 形象设置

形象设置页，显示当前数字人的位置坐标，并支持【同步】全选或者选中指定页面，进行数字人设置同步。



## (2) 文本与时间轴编辑

在工作台页面中，会自动导入 ppt 内的备注文字作为旁白文本，自动解析 ppt 内设置的动画效果，也可以二次编辑旁白文本。

### ① 文本合成语音

针对文本效果执行以下操作：停顿/替换发音/数字读法。

试听当页备注文稿转语音合成(TTS)输出，并【选择时长后添加停顿】，使语音输出自然;针对发音异常，【选中指定文本(单字)并替换成同音字】，试听

并应用。

为保留用户原音，支持在线录音或者用音频文件驱动生成数字人视频。



单次选中文本字数上限为“400 字”，超出不予加载;加载耗时受选中文本字数影响;除单句试听外，还提供当前页全文本试听。

根据需要自主开关【字幕】。

点击左侧页面【预览播放】按钮，可生成含数字人(无口型)的整体播放效果。



## ②调整时间轴动画

点击时间轴上的播放按钮，可加载当前页面动画效果，进行预览。

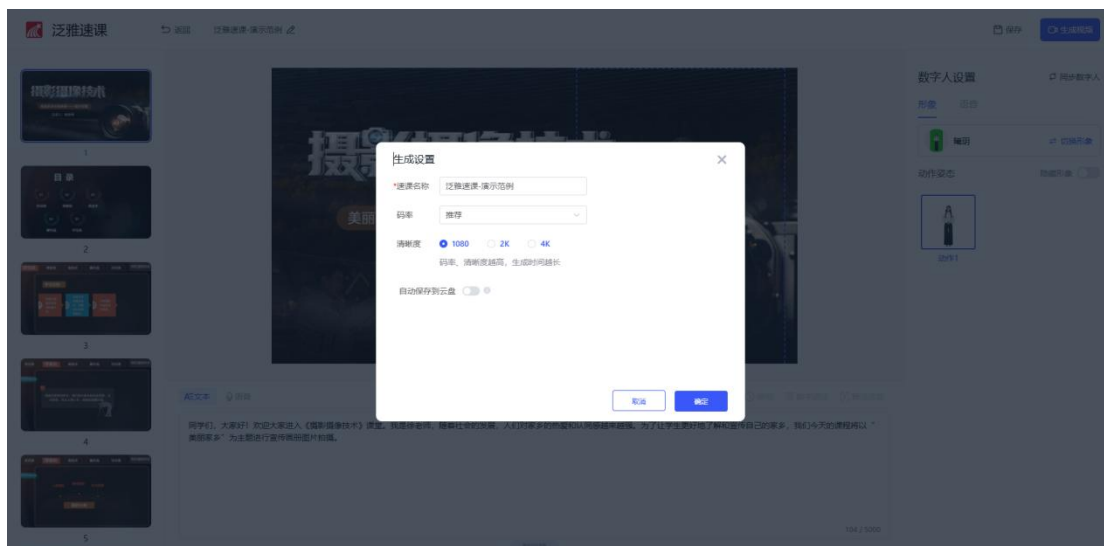
## (3) 保存及导出

建议:进行每一步配置后都及时【保存】，保存后点击生成。

编辑完成后，点击工作台右上角【保存】按钮保存编辑状态，下次打开项目时会展示最近一次保存的配置效果。

a) 生成中允许再次编辑，可选择是否更新保存配置，是否重新生成(重新生成原有进度清零)；

b) 文件生成用时，受【文件内备注文稿字数及动画数量】影响，可在项目管理页/数据详情页查看生成进度。



c) 生成后导出文件, 默认视频规格为【1080P、推荐码率】, 可以自行选择更高清晰度数据, 主要调整为数字人的分辨率, 格式为 MP4 视频格式/ZIP 压缩文件可根据需要导出字幕等文件, 也支持直接对接到学习通云盘。

d) 点击确定后, 会根据选择的下载内容通过浏览器下载到本地。