

# 虚拟数字人

- 可分为卡通、拟人、写实、超写实等风格

Calvin Tang

2021年12月

179209347@qq.com

## 虚拟数字人概述

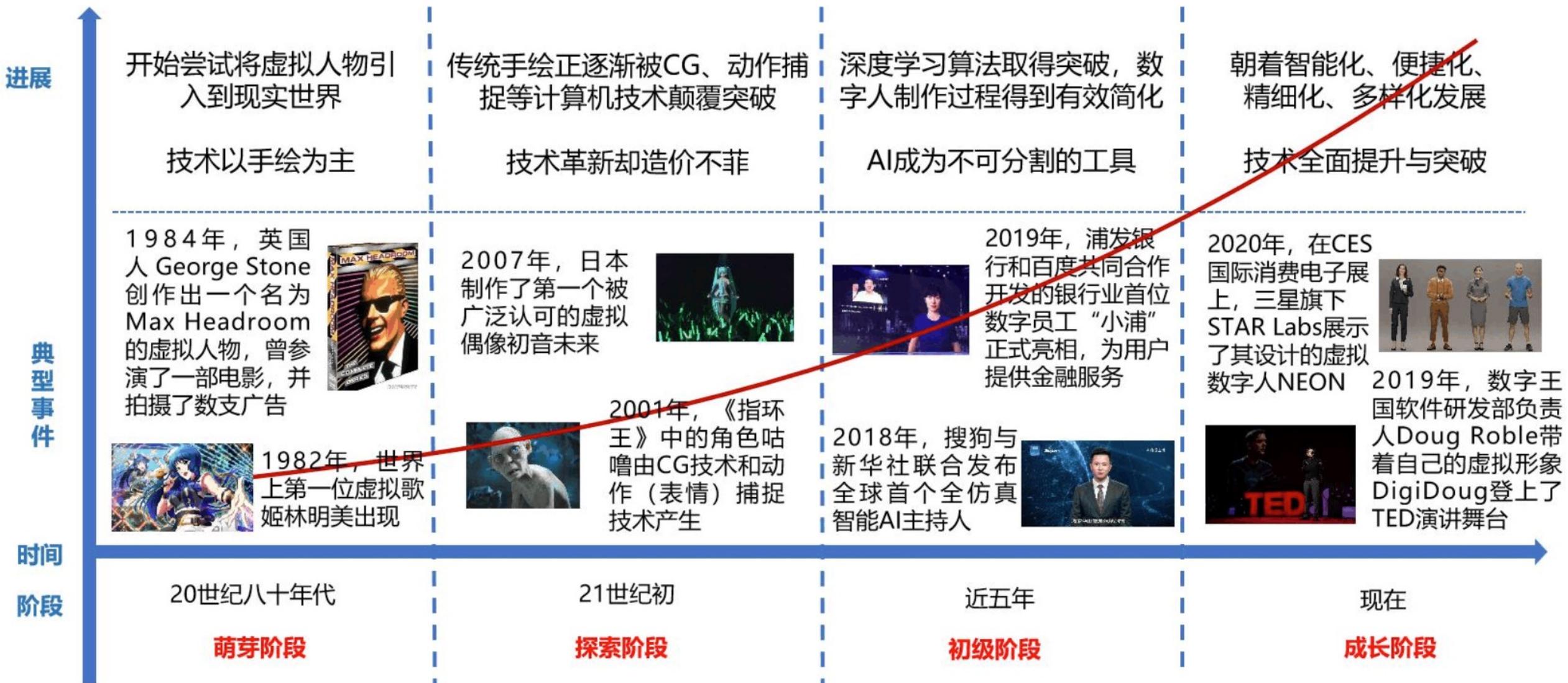
虚拟数字人具备以下三方面特征:

- 拥有**人的外观**，具有特定的相貌、性别和性格等人物特征；
- 拥有**人的行为**，具有用语言、面部表情和肢体动作表达的能力；
- 拥有**人的“思想”**，具有识别外界环境、并能与人交流互动的能力。

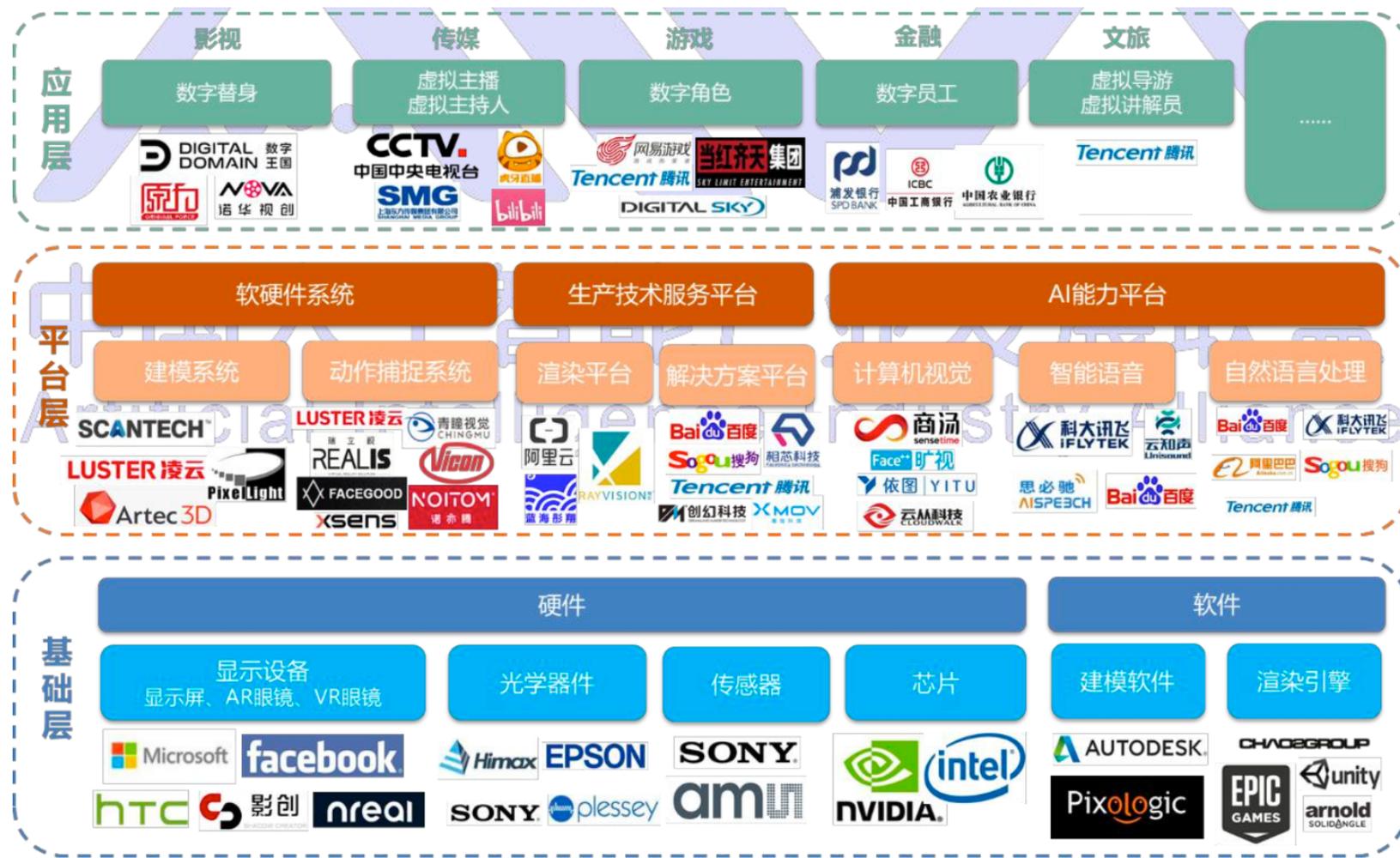


智能顾问

# 虚拟数字人发展历程



# 虚拟数字人产业视图



**应用层**是指虚拟数字人技术结合应用场景领域，切入各类，形成行业应用解决方案，赋能行业领域。

**平台层**包括软硬件系统、生产技术服务平台、AI能力平台，为虚拟数字人的制作及开发提供技术能力。

**基础层**为虚拟数字人提供基础软硬件支撑，硬件包括显示设备、光学器件、传感器、芯片等，基础软件包括建模软件、渲染引擎。

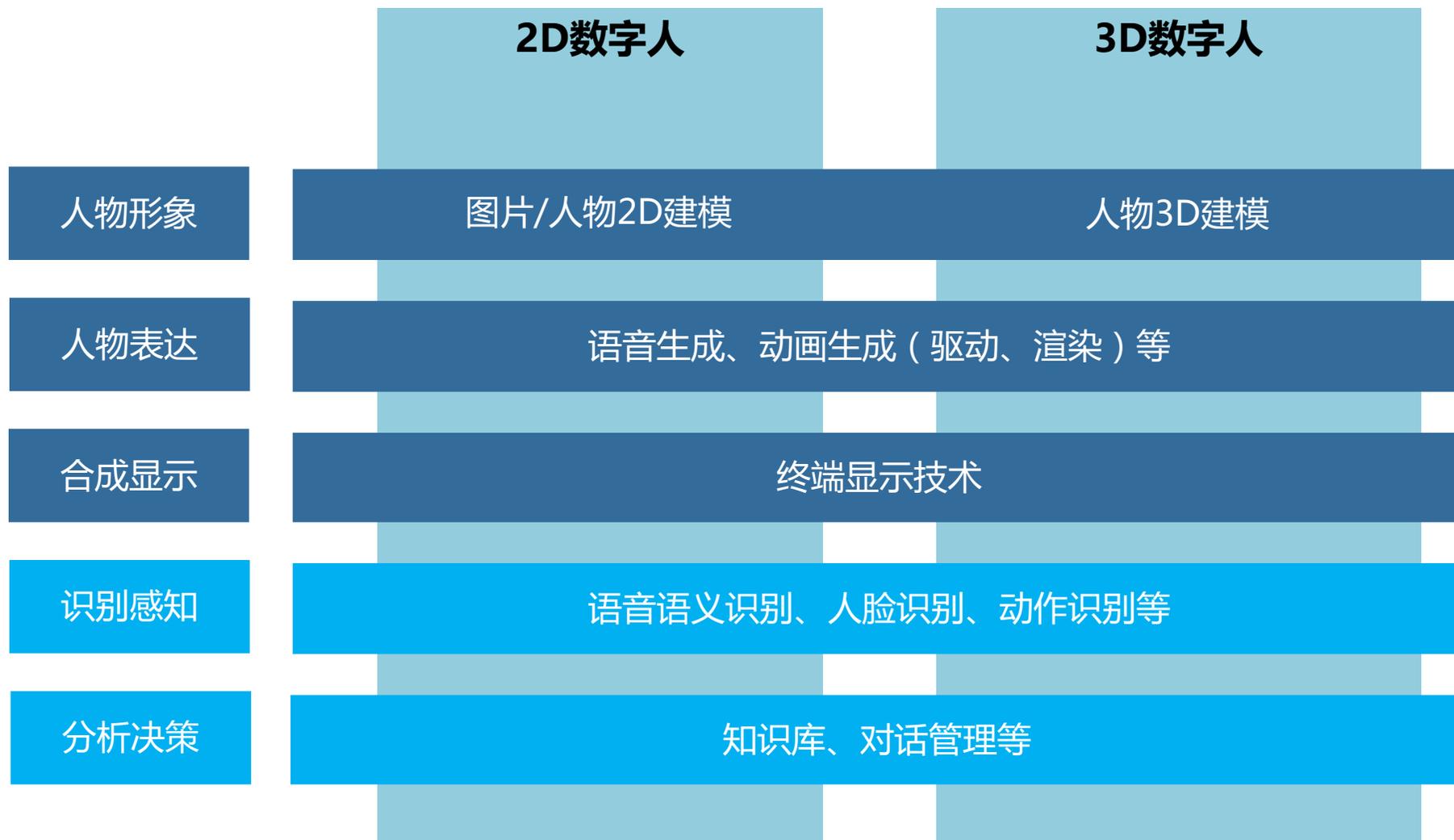
## 虚拟数字人系统构成

虚拟数字人系统一般情况下由人物形象、语音生成、动画生成、音视频合成显示、交互等5个模块构成：

- **人物形象**根据人物图形资源的维度，可分为 2D 和 3D 两大类，从外形上又可分为**卡通、拟人、写实、超写实**等风格;
- **语音生成**模块和**动画生成**模块可分别基于文本生成对应的人物语音以及与之相匹配的人 物动画;
- **音视频合成显示**模块将语音和动画合成视频，再显示给用户。
- **交互**模块使数字人具备交互功能，即通过语音语义识别等智能技术识别用户的意图，并根据用户当前意图决定数字人后续的语音和动作，驱动人物开启下一轮交互。

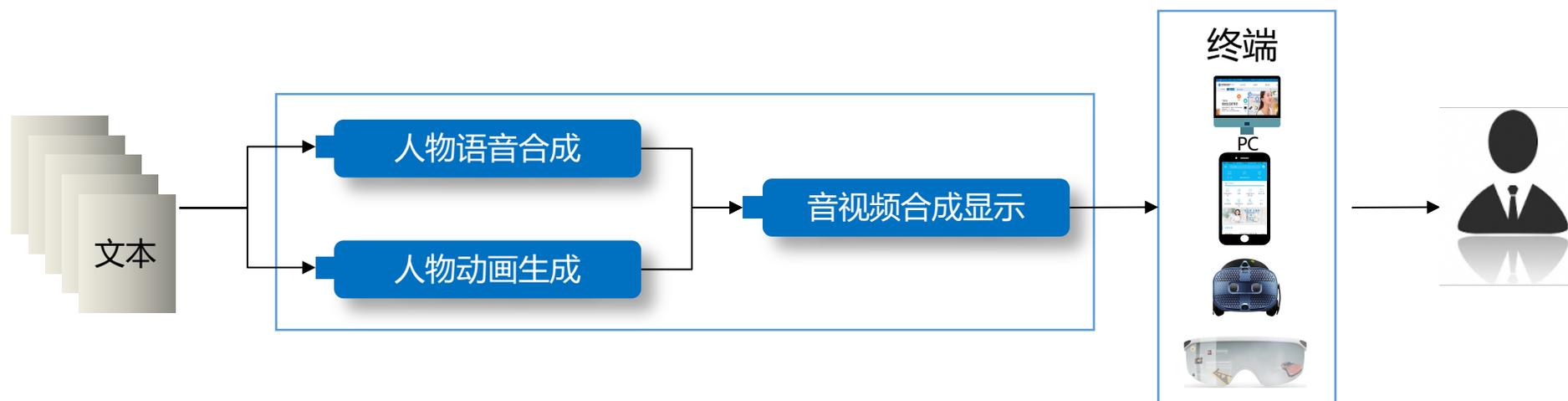


# 虚拟数字人系统构成



## 交互模块

交互模块为扩展项，根据其有无，可将数字人分为交互型数字人和非交互型数字人。

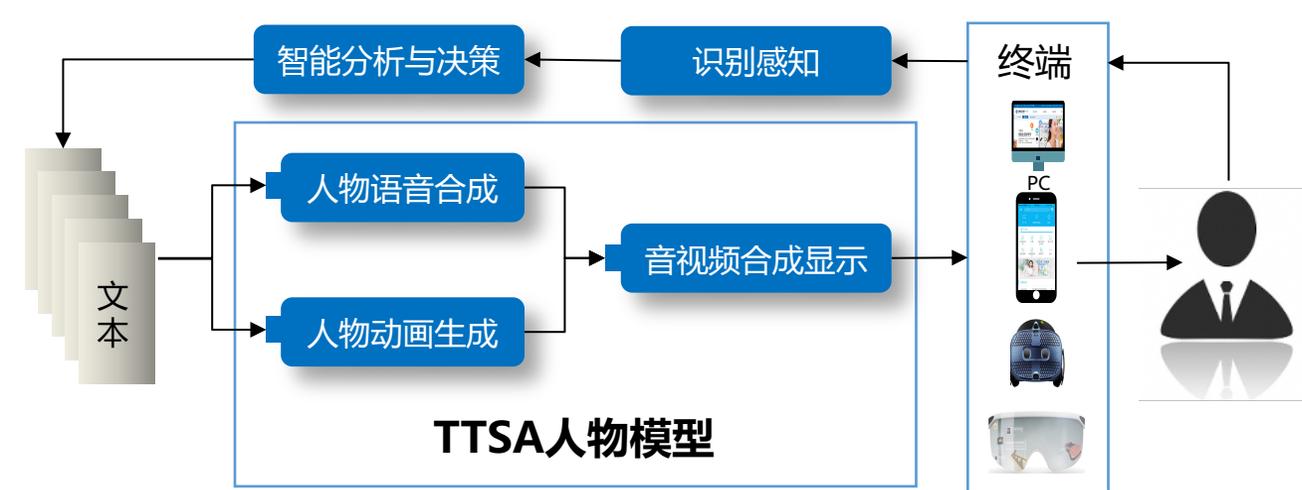


非交互型数字人系统运作流程

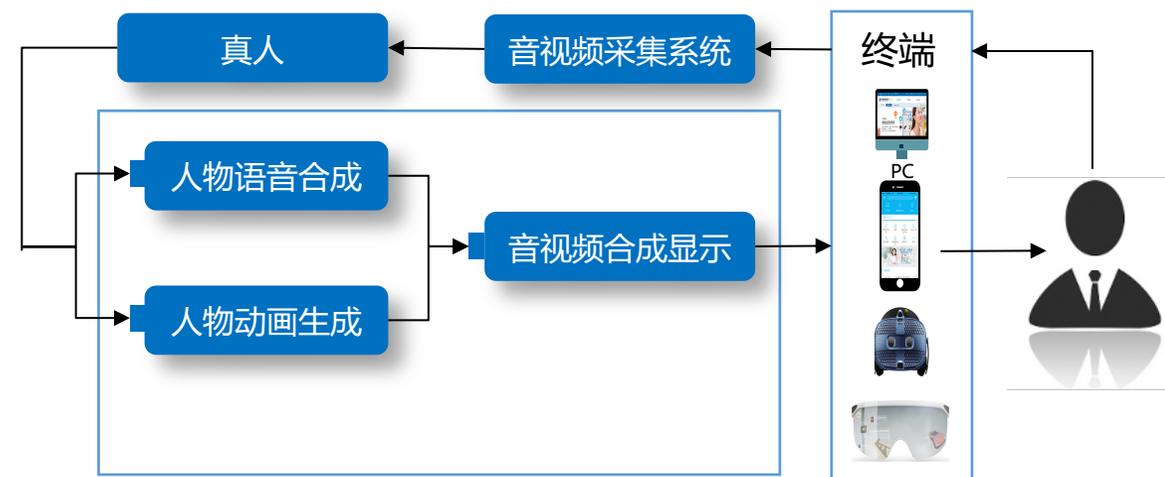
## 交互模块

交互型数字人根据驱动方式的不同可分为**智能驱动型**和**真人驱动型**。

- 智能驱动型数字人**可通过智能系统自动读取并解析识别外界输入信息，根据解析结果决策数字人后续的输出文本，然后驱动人物模型生成相应的语音与动作来使数字人跟用户互动。智能驱动型数字人可通过文本驱动生成语音和对应动画，业内将此模型称为TTSA(Text To Speech & Animation)人物模型。
- 真人驱动型数字人**则是通过真人来驱动数字人，主要原理是真人根据视频采集系统传来的用户视频，与用户实时语音，同时通过动作捕捉采集系统将真人的表情、动作呈现在虚拟数字人形象上，从而与用户进行交互。



智能驱动型数字人系统运作流程



真人驱动型数字人系统运作流程

## 虚拟数字人制作过程

虚拟数字人制作过程涉及的**建模**、驱动、渲染三大关键技术。

**建模**：相机阵列扫描重建替代结构光扫描重建成为人物建模主流方式。

结构光扫描重建可以实现 0.1 毫米级的扫描重建精度，但其扫描时间长，对于人体这类运动目标，在友好度和适应性方面都差强人意。



相机阵列扫描

## 虚拟数字人制作过程

虚拟数字人制作过程涉及的建模、**驱动**、渲染三大关键技术。

### 驱动：

智能合成、动作捕捉迁移是数字人目前主流的动作生产方式。

### 智能合成：

2D、3D 数字人均已实现嘴型动作的智能合成，其他身体部位的动作目前还只支持录播。

### 动作捕捉迁移：

通过将捕捉采集的动作迁移至数字人是目前3D数字人动作生成的主要方式，核心技术是动作捕捉。

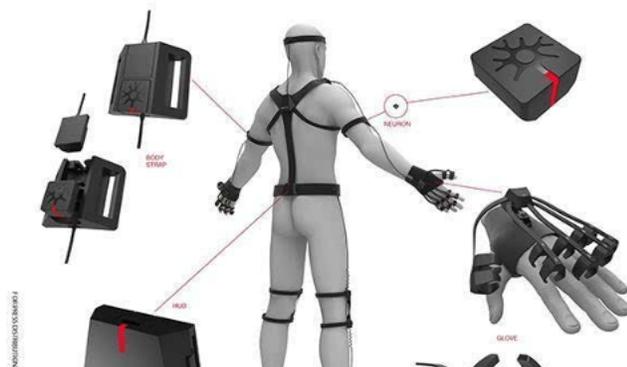
主流动作捕捉方案性能对比

方案	精度	抗遮挡	对环境要求	硬件成本	算法开发难度	应用公司
光学捕捉	高	低	高	高	中	Vicon、OptiTrack
惯性捕捉	低	高	低	中	低	Xsens、Noitom
视觉捕捉	低	低	低	低	高	Leap Motion、微软 Kinect

光学捕捉



惯性捕捉



视觉捕捉



## 虚拟数字人制作过程

虚拟数字人制作过程涉及的建模、驱动、**渲染**三大关键技术。

**渲染**：随着硬件能力的提升和算法的突破，真实性和实时性均大幅提升。

PBR ( Physically Based Rendering , 基于物理的渲染技术 ) 渲染技术的进步以及重光照等新型渲染技术的出现使数字人皮肤纹理变得真实。

目前常见的几款3D引擎，如 UnrealEngine 4, CryEngine 3, Unity 3D 5 , 均有了各自的PBR 实现。

实时渲染技术的突破助力写实类数字人实现实时交互，应用范围快速扩大。实时渲染指图形数据的实时计算与输出，其每一帧都是针对当时实际的环境光源、相机位置和材质参数计算出来的图像。

# 虚拟数字人应用场景

通过虚拟数字人产品与生产生活相融合，其规模化、可定制化、可复制化的能力能够推动改善传统环节流程、提升效能、降低成本等，大幅提升业务体验，给传统领域带来变革。

领域	场景	角色	
泛娱乐	影视	数字替身特效可以帮助导演实现现实拍摄中无法表现的内容和效果，已成为特效商业大片拍摄中的重要技术手段和卖点。	数字替身
	传媒	定制化虚拟主持人/主播/偶像，支持从音频/文本内容一键生成视频，实现节目内容快速、自动化生产，打造品牌特有 IP 形象，实现观众互动，优化观看体验。	虚拟主持人 虚拟主播 虚拟偶像
	游戏	越来越真实的数字人游戏角色使游戏者有了更强的代入感，可玩性变得更强。	数字角色
金融	通过智能理财顾问、智能客服等角色，实现以客户为中心的、智能高效的人性化服务。	智能客服 智能理财顾问	
文旅	博物馆、科技馆、主题乐园、名人故居等虚拟小剧场、虚拟导游、虚拟讲解员。	虚拟导游 虚拟讲解员	
教育	基于 VR/AR 的场景式教育，虚拟导师帮助构建自适应/个性化学习环境。	虚拟导师	
医疗	以数字人实现家庭陪护/家庭医生/心理咨询，实时关注家庭成员身心健康，并及时提供应对建议。	心理医生 家庭医生	
零售	从大屏到机器人到全息空间，从数据分析、个性营销、智能货架、无人商店四大应用场景切入构建线下零售服务新流程。电商直播在虚拟数字人与真人主播和观众互动，介绍商品。	顾客服务数字人 商家管理数字人 虚拟主播	

## 虚拟数字人+金融

- 虚拟数字人能够有效助力金融机构实现“降本增效”，各大金融巨头争相布局“数字员工”。运用金融科技实现“降本增效”应对市场竞争成为金融行业发展共识。
- 虚拟数字人拥有拟人的表情动作，可进行智能对话，能够与服务类场景较多的金融行业天然结合。多家金融机构正利用虚拟数字人技术打造“数字员工”，成为了科技创新、降本增效的重要方向。

## 虚拟数字人+金融：典型金融机构介绍

**中国工商银行**，推出的数字人银行员工，在移动终端和大屏等媒介上展现数字拟态形象，能够在产品营销讲解、金融业务办理、资讯播报、咨询问答等多个业务场景，实现与用户的可视化交互，为用户带来个性化服务，有效解除用户顾虑，提升用户体验和驻留时间，真正让数字化服务“听得见”的同时也“看得见”。

**解决方案厂商：**  
硅基智能



## 虚拟数字人+金融：典型金融机构介绍

**浦发银行**“3D金融数字人”创新计划自 2018 年启动，2019 年建设并投入金融应用。

业内首个综合运用 3D 人像实时驱动、多模态交互、自然语言处理、情绪识别、智能推荐、多因素生物认证等前沿技术，连接多个金融交易系统及物联网系统。

作为银行员工的数字分身，已正式在线上 APP 渠道以及线下营业网点为 C 端用户提供普惠服务。这也是由该行总行行长直接领导的创新应用项目之一，已成为金融数字化转型成效的一张名片。

### 解决方案厂商：

浦发银行和**百度**共同发布数字员工“小浦”。



浦发银行 SPD BANK

Hi, 我是数字人小浦  
欢迎来到未来银行

ENGAGING WITH YOU  
BANKING FOR YOU

浦发银行正在打造数字员工系列  
A new digital workforce is on the way

- 财富规划师
- 智能柜员
- 大堂经理
- 数字面试官
- 智能客服
- 数字主播
- 数字文档审阅员

敬请期待更多数字员工角色

SPD Bank dedicates to innovate financial services with fintech solutions

- API
- DaoS
- Cloud Service
- Block Chain
- IoT

浦发银行数字员工 2020 年 8 月版

## 虚拟数字人+金融：典型金融机构介绍

**光大银行**，2019年率先落地应用虚拟银行雇员，即采用魔法科技的虚拟数字人技术。2020年，小米开发者大会采用魔法科技虚拟数字人直播方案，发布全新小爱5.0。目前，魔法科技AI智能交互虚拟数字人已广泛应用于教育、金融、医疗、市政等领域。

### 解决方案厂商：

**魔法科技**运用自主研发的全栈式AI表演动画解决方案、全智能虚拟数字人技术及文本实时合成语音动画(TTSA)技术，打造全智能虚拟数字人能力平台，赋能数字人在各类终端与用户进行三维、实时多模态地智能互动，可提供7\*24小时的智能化服务。



# 其它虚拟数字人解决方案厂商

## 1. 腾讯

腾讯提供 IP 角色授权与多模态交互方案，致力于打造“能听会说”的 AI 智能形象。2020 年腾讯数字生态大会，采用了 IP 虚拟人技术，基于镜双城动画番剧女主角那笙的 IP 角色，腾讯构建了“那笙虚拟助手”，使其来到大会上与演讲嘉宾进行互动。

## 2. 网易伏羲实验室-虚拟人科技创新中心

网易伏羲致力于全面推进虚拟人前沿技术研究、工程优化和解决方案，汇集美术制作、形象智能创建、智能表情合成、智能肢体动作合成、智能语音合成、智能对话生成、智能用户感知、运动采集等关键技术。已赋能于多行业多场景，包括游戏角色、虚拟教师、虚拟讲解员、虚拟诗人、虚拟品牌形象代言人、虚拟偶像；创新成果丰富，(截至 2020 年 11 月)共计发表国际顶级会议及期刊论文 58 篇，授权发明专利 34 项。

## 3. 搜狗-虚拟数字人计算机视觉、自然语言处理技术厂商

搜狗推出了多模态人工智能技术，基于文本、语音、图像、图形等多种模态信息建模生成高逼真度的虚拟数字人。2018 年 11 月 7 日第五届世界互联网大会上，搜狗与新华社联合发布 AI 主持人。通过语音合成、唇形合成、表情合成以及深度学习等技术，克隆出具备和真人主播一样播报能力的“AI 合成主播”。2020 年 5 月 21 日，搜狗联合新华社又推出了 3D 版 AI 合成主播“新小微”，“新小微”AI 合成主播高度还原真人发肤，在立体感、灵活度、可塑性、交互能力等多个方面，都有了大幅跃升，而且还能根据播报的内容，做出各种更接近于真人的姿势和表情。

*Thank You*

谢谢