

虚拟数字人

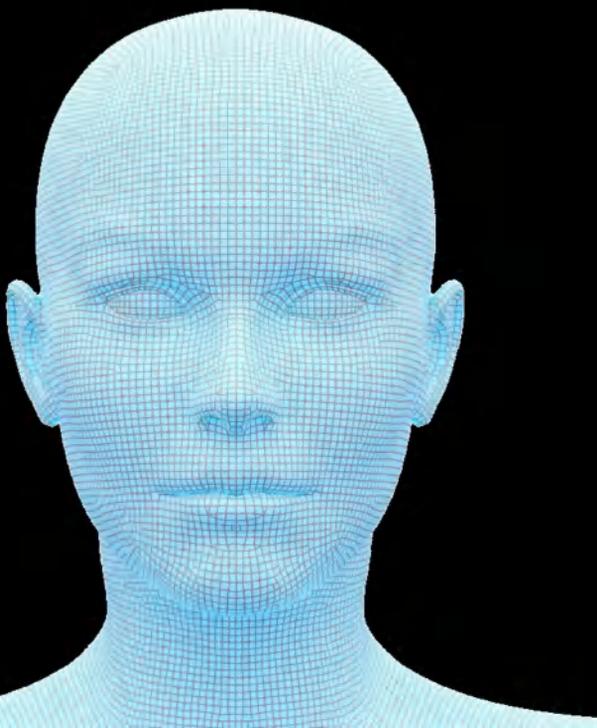
深度产业报告

技术背景篇

定义、关键词、分类、技术细节

产业现状篇

市场规模、发展阶段、细分场景及判断、产业链、竞争格局、中外对比、产业瓶颈



技术背景

1.1 定义

虚拟数字人指存在于非物理世界中，由计算机图形学、图形渲染、动作捕捉、深度学习、语音合成等计算机手段创造及使用，并具有多重人类特征（外貌特征、人类表演能力、人类交互能力等）的综合产物。市面上也多将其称为虚拟形象、虚拟人、数字人等，代表性的细分应用包括虚拟助手、虚拟客服、虚拟偶像/主播等。



Source: UneeQ官网, 相芯科技产品手册, 魔法科技产品手册, 网络公开报道

其具体特征可拆分为“虚拟”、“数字”及“人”三个部分。在试用了相关产品后，我们认为，“人”是其中的核心的因素。高度拟人化为用户带来的亲切感、关怀感与沉浸感是多数消费者的核心使用动力。能否提供足够自然逼真的相处体验，将成为虚拟数字人在各个场景中取代真人，完成语音交互方式升级的重要标准。



虚拟—存在于非物理世界中，不同场景实现难度不同

目前，虚拟数字人主要以图片、视频、实时直播、实时动画等方式存在于电子屏中，如APP、小程序、软硬一体显示设备。在未来，VR设备与全息投影也将成为其重要存在方式。

需注意的是，尽管其存在方式均在非物理世界中，由于各场景所需的时延（如直播等实时场景要求低时延，但内容生成场景无该要求）、驱动方式（计算驱动对模型的深度学习能力有极高要求）等不同，对技术、运营等要求差异较大。



数字—依托多项技术存在，相关技术成熟成为其发展重要推动力

虚拟数字人是典型的多技术综合产物。除CG建模+真人驱动的类型外，多模态技术与深度学习成为其核心点。该部分将在技术中详细展开。多位从业人员认为，虚拟数字人近年的发展来源于CG（Computer Graphics，利用计算机进行视觉设计和生产）、语音识别、图像识别、动作捕捉等相关技术的共同成熟。



人—在外表、行为、交互行为等方面高度拟人化，外在表现和交互效果已成为核心发展路线

外表：

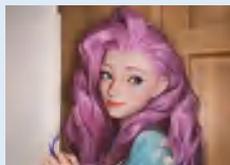
虚拟数字人的面部长相和整体形象。会受到虚拟数字人类别（如直接借用真人形象、高保真建模、风格化）、制作细节（对汗毛、皮肤、头发等细节的建模）、渲染水平、设计审美等影响



三星 Neon
基于真人采集



Unreal Engine Metahuman Creator
高保真建模



英雄联盟 Seraphine
风格化



Genies 基于真人明星的虚拟数字人
风格化

行为：

虚拟数字人的面部表情、形体表达、语音表述等。会受到驱动方式（真人驱动、计算驱动、预制调节等）、驱动模型类别（精细面部肌肉驱动、语音合成模型中对语气词、韵律的处理等）、训练数据、驱动模型精度等影响

交互：

虚拟数字人与现实世界的交互水平，包括回答内容、肢体反应等，会受到语音识别能力、自然语言理解及处理水平、知识图谱、预先设置知识库等影响。

1.2 技术细节

从技术层面来看，虚拟数字人可以分为真人驱动型和计算驱动型。后者是近年来多模态技术和深度学习发展的技术集大成者。

在真人驱动中，在完成原画建模和关键点绑定后，动捕设备或摄像头将基于真人的动作/表情等驱动虚拟数字人。

由于背后有真人操作，真人驱动型在动作灵活度、互动效果等方面有明显优势，一方面能够在影视内容的创作中减低生产成本，为影视行业降低门槛，推动影视级内容向消费级转化。另一方面则多用于虚拟偶像、重要直播中，帮助虚拟数字人完成大型直播、现场路演等互动性、碎片化活动。

事实上，这种技术思路可以看作是传统影视制作中，CG技术的进一步延续。近年来主要的技术突破在于动作捕捉环节。随着图像识别技术，姿势、表情等识别算法的进步，昂贵的惯性或光学动捕设备不再是驱动的必备工具。普通摄像头结合理想的识别算法通用能实现较为精准的驱动（如iPhone 12摄像头已可支持简单的动作捕捉），显著降低了精细虚拟内容生成的门槛。

真人驱动型技术流程

1. 形象设计及建模

基于IP设计或真人偶像绘制原画，进行面部及身体3D建模，选择关键点

2. 建模绑定

将识别关键点映射至模型上，进行绑定。关键点绑定的数量及位置影响最终效果。

3. 表演捕捉

利用动作捕捉设备或特定摄像头+图像识别，捕捉在形体、表情、眼神、手势等方面的关键点变化。

4. 驱动及渲染

真人演员（在虚拟偶像中称为中之人）根据制作需要进行相应表演，实时驱动虚拟数字人表演。在较为精细的制作中，会根据真人演员和建模的区别进行重定向，并对动作、眼神、手指等采用不同的驱动方式。

需要时需进行语音合成，形成特定设置语音

5. 生成内容，进行互动

进行直播，或录制其动作生成内容



美国虚拟主播CodeMiko
采用真人驱动方法调整虚拟主播的动作及表情

在计算驱动型中，虚拟数字人的语音表达、面部表情、具体动作将主要通过深度学习模型的运算结果实时或离线驱动，在渲染后实现最终效果。

计算驱动的虚拟数字人最终效果受到语音合成（语音表述在韵律、情感、流畅度等方面是否符合真人发声习惯）、NLP技术（与使用者的语言交互是否顺畅、是否能够理解使用者需求）、语音识别（能否准确识别使用者需求）等技术的共同影响。尽管在特定方向上，各感知类技术已有的商业化能力已足以支撑，然而，但要达成理想的综合效果，需要该公司在三个方面同时具有较强的综合能力。这也间接导致国内开展计算驱动型虚拟数字人业务的公司，大多是在感知技术方面有较强的综合实力，相对成熟的AI技术公司。

计算驱动型技术流程

1. 设计形象。扫描真人形态及表演、采集驱动数据

利用多方位摄像头，对通用/特定模特进行打点扫描（视最终需求可进行全身或局部扫描），采集其说话时的唇动、表情、面部肌肉变化细节、姿态等数据

计算驱动型技术流程

2. 形象建模，进行绑定

设计所需的模型，或基于特定真人进行高还原度建模。进行关键点绑定。关键点绑定的数量及位置影响最终效果。

当需要基于真人照片生成虚拟内容时，一类做法是将通用的人脸模型迁移至该真人照片上，形成虚拟形象，实质为表情迁移。另一类则是生成动漫类效果，基于预先设置的形象分类算法，将真人照片中的眼型、发型等元素进行分类，并与预先设置的动漫元素进行匹配，最终生成动漫式的虚拟形象。

3. 训练各类驱动模型：决定最终效果的核心步骤

利用深度学习，学习模特语音、唇形、表情参数间的潜在映射关系，形成各自的驱动模型与驱动方式。

充足的驱动关键点配合以精度较高的驱动模型，能够高还原度的复原人脸骨骼和肌肉的细微变化，得到逼真的表情驱动模型。魔法科技等业界领先的模型可组合出超千种表情效果，并包含眼神驱动。科大讯飞、竹间智能等公司会对语音/文本中的因素进行提取，增加情感驱动模型等。

目前为止，大多数厂商的驱动模型大多是语音-唇形，语音-驱动。动作、手势等驱动大多依靠人为现场指令或预设置驱动。

对于需对特定真人定制化的数字化虚拟数字人，部分公司会基于在通用驱动模型的基础上，结合少量真人驱动数据训练定制化驱动模型。这种情形可视为预训练模型+小样本学习。

4. 内容制作：基于输入的语音（或由输入文本转化的语音），预测唇动、表情等参数

核心的技术流程是基于输入的语音，或首先基于TTS技术（Text-to-speech，语音合成技术），将输入的本文转化为语音。基于语音，结合第3步得到的驱动模型，并利用生成对抗模型GAN选出最符合现实的图片，推理得到每帧数字人的图片。通过时间戳，将语音和每帧的数字人图片进行结合。

5. 进行渲染，生成最终内容。直播时进行实时渲染

为保证在特定场景下能够实现实时低延迟渲染，计算框架的大小、算力供给等技术问题同样会影响到虚拟数字人的最终生成效果

附加步骤：

针对需要进行交互的虚拟数字人，生产商会预先设置有问答库、知识图谱等，并承接入虚拟数字人的对话系统。

当判定技术中的核心要素时，我们需要回归到虚拟数字人的核心价值，也即前面所提及的“拟人化”，这需要虚拟数字人在形象或交互上具有技术优势。在技术层面主要体现为以下三点。



CG建模/图像迁移技术 影响外观呈现

体现为虚拟数字人外观的拟人程度。国内外在该项上的技术差异，部分导致了目前国内外玩家主要专注于不同的细分场景和发展路径



NLP交互技术 影响交互体验

以对话能力为核心。继文本对话助手、语音AI助手后，该技术继续在虚拟数字人中发挥核心作用，可以视为虚拟数字人的大脑。在AI交互助手方面已有理想成效，如小冰等，公司能够为其添加较好的通用式互动能力。追一科技等公司则通过知识图谱、业务问答库、对话型引擎等增强虚拟数字人的业务互动能力。



CV等深度学习模型 影响驱动效果

受数据量、计算框架、关键特征点等因素深刻影响。能否呈现自然的面部表情变动、肢体变动等，在极大程度上取决于语音驱动的深度模型效果。除此之外，能否对情感等因素进行特别设计，同样会产生重要影响。

值得一提的是，2021年，Fable Studio的新虚拟数字人Charlie and Deck背后将有GPT-3支撑，生成类语言模型的加入有望显著提升虚拟个人助手的日常交互能力，这也将是GPT-3实现商业化的全新途径。

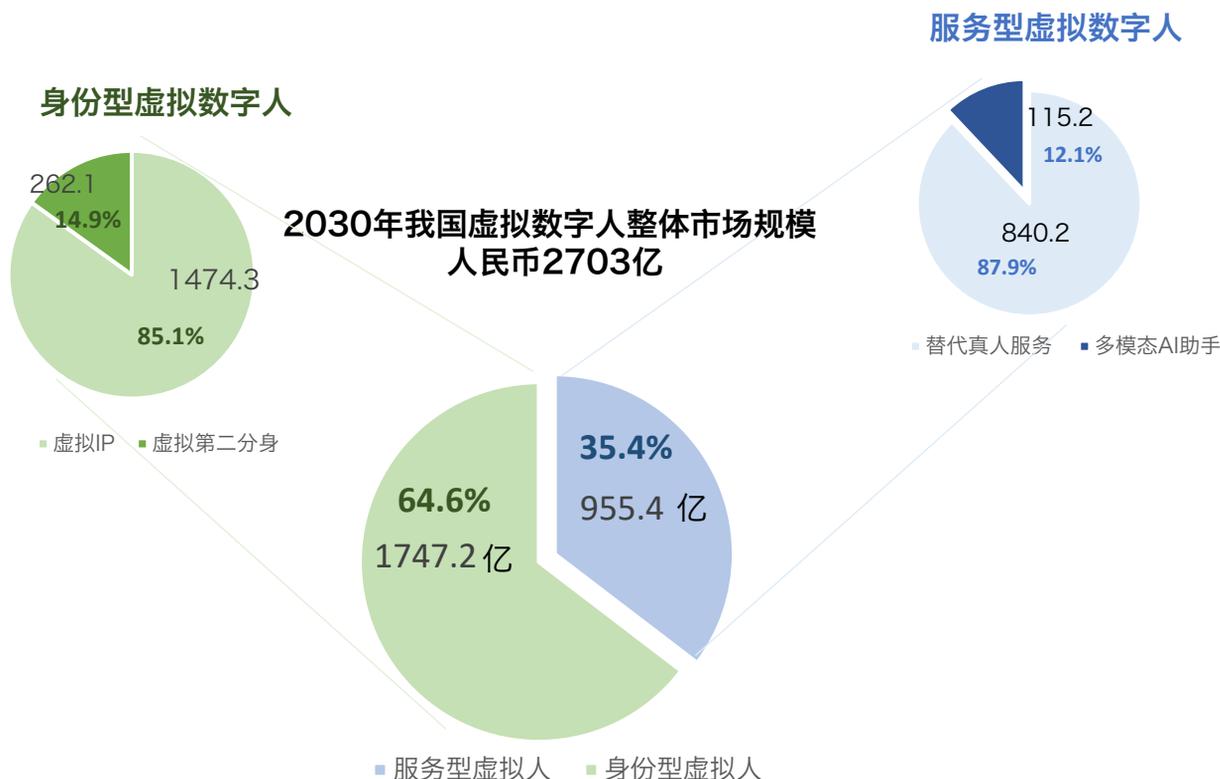
行业现状

2.1 市场规模

基于国内外多家虚拟数字人厂商的业务，本报告将虚拟数字人按照其产业应用划分为两类：服务型虚拟数字人和身份型虚拟数字人。

	服务型虚拟数字人	身份型虚拟数字人
核心功能	功能性，提供服务 <ul style="list-style-type: none"> 替代真人服务，完成内容生产、简单功能 多模态AI助手，提供日常陪伴、关怀等服务 	身份性，用于娱乐/社交 <ul style="list-style-type: none"> 虚拟IP/偶像，推动虚拟内容生产 虚拟世界第二分身，用于社交娱乐及Metaverse
产出定位	<ul style="list-style-type: none"> 现实世界中主持人、导游、购物主播等服务型角色的虚拟化 具有关怀感和真实感的AI助手 	<ul style="list-style-type: none"> 虚拟偶像或为真人偶像的辅助分身 个体在虚拟世界（或Metaverse）中的第二分身
代表应用	<ul style="list-style-type: none"> 虚拟主播、虚拟老师、标准化内容制作 虚拟关怀师、虚拟陪伴助手 	<ul style="list-style-type: none"> 虚拟偶像 VR Chat、虚拟形象构建应用
产业价值	<ol style="list-style-type: none"> 降低已有服务型产业的成本，为存量市场降本增效 提升AI助手的交互效果，扩展其接受度与适用场景 	<ol style="list-style-type: none"> 为未来的虚拟化世界提供人的核心交互中介，在增量市场创造新价值增长点 降低虚拟内容的制作门槛

经过测算，我们认为，在2030年，我国虚拟数字人整体市场规模将达到2700亿。其中，得益于虚拟IP的巨大潜力，以及虚拟第二分身的起步，身份型虚拟数字人将占据主导地位，约1750亿，并逐步成为Metaverse中的重要一环。服务型虚拟数字人则相对稳定发展，多模态AI助手仍有待进一步发展，多种对话式服务升级至虚拟数字人形态，总规模超过950亿。



理想的市场规模体现在虚拟数字人这一市场行业天花板高，实现路径明确，且已出现了有相当竞争力的技术公司。在各行各业近两年开始意识到其技术可行性和商业价值后，行业将快速进入高速发展期。

首先，作为多模态升级的代表技术，虚拟数字人的应用场景众多，可广泛与各行业领域相结合，变现路径和市场潜力明确。其中虚拟直播、虚拟偶像已得到明确的商业价值验证。而虚拟分身生成等场景也已得到了资本和相关产业方的认可。

其次，大多数国内玩家为综合实力较强，发展成熟的AI科技公司，在客户积累、技术实力、销售渠道、资金供给等方面较为可靠，行业预计可直接进入平稳发展期。

最后，在多家业内代表公司看来，虚拟数字人作为Metaverse主要的交互载体，具有明确的巨大增长潜力，并基于NFT、VR等有理想的延展空间。行业天花板高，能够维持长期和衍生发展。

2.2 行业发展阶段

火山引擎、追一科技等公司基于自身观察，认为虚拟数字人行业目前仍处于孵化培育的阶段。相芯科技基于XR行业技术成熟度曲线，认为虚拟化身技术在未来2-5年内将进入稳步爬坡期。

而在行业的长期发展方面，可以依据面向市场、解决方案及产品成熟度、基础设施及衍生产业进行划分。

相芯科技基于自身认知，将虚拟数字人的未来发展划分为三个阶段，依次为1. B端行业探索与教育的成长期。该阶段主要用于完善行业解决方案，同时通过2B2C对C端客户进行市场教育；2. B端场景成熟向C端渗透的发展期。该阶段B端应用场景将与行业深度结合，逐渐成熟。同时厂商开始关注UGC，消费级工具、C端消费订阅制、虚拟内容生产迎来井喷。3.B、C端应用成熟，并形成数字资产及经济体系，技术迭代的成熟期。

基于我们的观察，目前三大阶段的对应现象均已有所出现。但目前，面向B端探索成熟的落地场景仍然为各厂商的主要关注点。

据分析，虚拟数字人行业未来的主要驱动力包括：用户代际变化，新一代消费者对内容消费和虚拟世界更为渴求；虚拟数字人相关技术（如计算机图形学、深度学习等）门槛相对降低，成本有所回落；资本热度上升，受Metaverse概念和国外成功厂商引导，虚拟化的趋势逐渐成为共识；VR眼镜等相关配套设备逐渐回暖，有望实现大规模商用。

2.3 细分场景及相关判断

在本部分中，会对服务型虚拟数字人和身份型虚拟数字人进行具体展开，说明其核心价值、主要赛道、推动力、相关玩家及其特征，并基于分析师的观察给出一定判断。



2.3.1 服务型虚拟数字人

目前，国内的服务型虚拟数字人主要用于替代真人进行播报等内容生成，并进行简单问答交互等。而国外则由于在CG（Computer Graphic，计算机图形学）方面具有技术优势，能够打造具有高度关怀感的虚拟数字人，率先在医疗等场景落地了虚拟陪伴助手、心理咨询顾问等。

整体而言，由于虚拟数字人提供动作、表情等更多维度的信息，之前仅基于文本、图像或语音的单模态服务及内容都能以虚拟数字人的形态进行更丰富的呈现。内容方面，有声读物、文本内容等将有望快速扩展至视频领域。据相

芯科技，除去常见的课件等内容生产外，部分面临着人才短缺，但希望能够融媒体转型的县市级媒体，也在试图通过有服务型虚拟人，低成本地生成所需媒体内容。服务方面，通用的语音助手、个性化顾问、车载系统，针对特定场景的指引、导览等均能够以更为人性化和易于使用的方式呈现。不过，尽管虚拟数字人似乎通用性很强，但仍需要注意，目前的制作成本仍高达百万级，需要首先关注对多模态信息和拟人化具有强烈需求的场景。

短期主要价值增长点

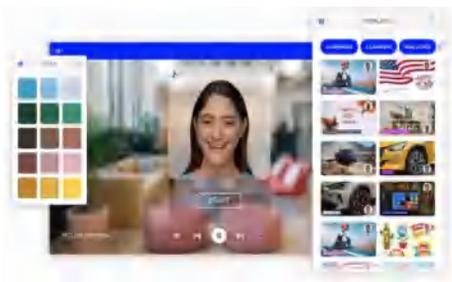
a. 替代真人服务

- **核心价值及主要功能：**代替真人进行内容生产和简单交互 — 降低对真人表演的需求，减少标准化内容的制作成本；特定场景下（政务、金融等）客服、虚拟主播、新闻播报
- **主要赛道：**企业服务、内容行业
- **未来推动力：**业务快速更迭导致的服务类人力培养难度与成本增加；内容行业向视频时代升级的需求
- **主要判断：**需关注对拟人化有强烈需求的特定服务场景；通过知识图谱、客户积累等专注垂直行业服务的短期内更易变现；内容行业更易实现规模化复制，需强调SaaS及预设模板等标准化工具

在**内容生产**方面，虚拟数字人内容生成平台已成为多家厂商共同的发力点。国内厂商包括火山引擎、科大讯飞、相芯科技等，平台以新闻播报为主要场景。国外厂商以英国的Synthesia和Hour One为代表。中外主要差异点主要在可供选择的虚拟数字人数量，以及预设模板等线下细节。

在这类平台上，使用者只需将需要播报的内容输入平台，选择主持人形象、音色、背景后，即可快速生成相关播报视频。使用者可以利用时间轴，对虚拟数字人进行动作等调节。部分产品还支持插入演示面板、根据时间轴调整位置等，最终生成图文并茂，带有解释说明的视频。

尽管国内各平台的产品设计中，更多的倾向于新闻播报。但我们认为，对于产品说明书、教学内容、产品说明视频等内容相对固定，观看者对于视频灵活度无过高要求的说明性内容，都可以采用这种方式进行制作，以降低相关投入。



HourOne的虚拟内容生成平台
配有36个虚拟形象，多类标准化模板及视频管理系统



Synthesia的虚拟内容生成平台
配备40个虚拟形象，虚拟数字人视频内容生成领域融资最高

简单交互方面，虚拟数字人不适用于通用性、专业性、交互性过强的领域，会暴露其现有技术短板。国内目前多专注于特定细分市场，主要用于提供简单业务交互。

虚拟主播是目前国内商业化竞争最为激烈的场景，以中科深智为代表，部分厂商开始提供一站式服务，包括客户端、布置方案、硬件系统、乃至流量运用等配套服务。个别厂家会将直播场景中的运营细节融入产品设计中，用于更好的迎合消费者心理和平台的流量分发机制。目前，虚拟直播的价格门槛已从数万级下降至数千级，使用门槛也大幅降低。

其次为各家均在摸索部署的虚拟客服，据追一科技，在现有的技术条件下，业务需求和规则流程相对明确的客服场景成为虚拟数字人落地的理想方式。除去替代真人外，由于能够基于知识图谱、数据迭代等进行更快的业务知识升级，虚拟人在客服等背景知识会快速更迭的领域中有独特优势。虚拟客服得以在服务现场通过软硬一体大屏、电子播报屏

等，在线上通过App等形式更好的服务客户。在这一场景中，金融领域由于对新技术接受程度较高，且能够承担目前高达数百万元的客单价，正成为主要的应用行业。

其他场景还包含已在多家规划中的虚拟教师、导航导览（地铁站、医院等）、展览介绍（导览“小春妮”）等。

在使用中，通过预置知识库、知识图谱、用户自主配置触发条件等等，虚拟数字人可以对用户的基本诉求进行行为和语音识别，并以固定话术进行回应。例如，虚拟主播可以以固定的几种话术欢迎进入直播间的消费者，并在消费者提及价格、存货、产品详情等意图时，根据事实情况进行回应。虚拟客服则可以借助NLP技术理解用户需求并进行解答。目前，追一科技、科大讯飞等均有配有行业知识库。

在教育、直播等此类场景中，尽管交互只占据虚拟数字人整体工作中的一部分，但在魔法科技看来，问答的高转化率才是当前场景中的价值关键。追一科技同样认为互动沟通是虚拟数字人能力中最重要的部分。因此，对关键问题的识别交互能力对此类场景至关重要。为了解决这一问题，专注于此类场景的公司需要对相应的业务场景具有充足的行业知识积累，以提升其语言交互系统，并更好的对接原场景中的业务系统。

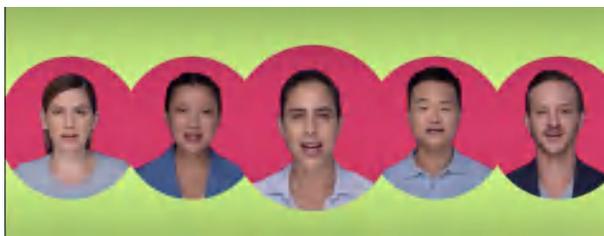
长期想象空间大，已出现领先布局者

b. 多模态AI助手

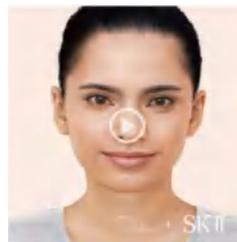
- **核心价值及主要功能：**基于真人形象，提供关怀感和真实感 — 在通用场景或特定关怀类场景中提供顾问、关怀、陪伴、事务处理等服务
- **主要赛道：**个性化经济、单身经济、电商购物、医疗康复
- **未来推动力：**新时代对个性化顾问和服务的需求；单身/老龄化/心理脆弱等人群对关怀陪伴的需求
- **主要判断：**关注陪伴感、关怀感等心理因素的场景更易变现，如健康顾问、虚拟男友等；对虚拟数字人在形象和交互方面技术要求极高；细分场景率先发展，在与手机厂商、系统提供商达成合作后，通用式助手将大规模普及；短期内会和语音交互助手有替代关系，长期语音助手可能成为其附属交互手段

多模态AI助手可以理解为助手型虚拟数字人，不仅能够在特定场景下基于用户系统高度个性化，还可以实现通用化交互。

特定场景下的多模态助手是国外虚拟数字人公司的业务发展重点，用于提供情感联结。医疗顾问、日常陪伴、购物客服等领域目前均已成为国外厂商的关注重点。如三星Neon的CEO, Pranav Mistry所言，“It's not because you want to know something, but you need a person to talk to”。与之相伴，是国外公司对虚拟数字人外表逼真度的追求。包括UneeQ、SoulMachines、HourOne、Neon会选择在预制虚拟数字人的设计及制作上投入大量精力。相对于国内，其预制虚拟形象在外观及行动方面，对真人的还原度有明显优势。



HourOne的预制虚拟数字人形象，用于网页向导等



SoulMachines为SK-II提供的个性化护肤顾问形象

报告搜一搜

更多金融干货下载

800000+份行业研究报告

长按识别关注公众号



ID: reportsys

不同于简单的虚拟客服，特定场景中的助手是高度个性化的。除去多模态技术外，与业务相关的个性化分析及推荐系统在此场景中同样重要。目前国外厂商已有的落地场景包括个性化理财顾问、心理咨询顾问、购物助理等。

除去赋能外部机构外，TwentyBN选择自主构建虚拟数字人陪伴应用。其Fitness Ally为个性化健身教练，能够基于使用者的实时运动情况等完成个性化交互。

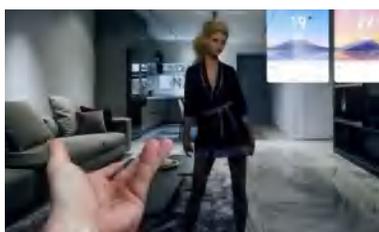
通用化的虚拟数字人助理目前仍主要处于前期阶段。基于智能音箱等非可视化设备的语音AI助手仍为主流。尽管拥有具体形象的虚拟数字人是智能虚拟助手的多模态高级形态，仍需要一定时间转化。

一方面，当前AI助手多以语音为载体，如手机、音箱等，消费者已习惯在使用时不动用视觉，且设备（以智能音箱为主）并未配备有显示装置。其多模态的特性不具备相对优势，更改用户习惯需要一定时间。

然而同时，在XR（Extended Reality，包括VR、AR、MR）或全息投影的方式下，具有具体形象的AI助手将具有较大意义。且此类助手将不仅存在于现实投射中，由于VR的大规模推广将重新构建底层的操作系统和消费者的感官需求，在VR世界中，如同过往游戏中的NPC，形象化的虚拟助手将承担特定场景/应用的引导员等。

部分公司目前已在尝试这一方向。与AR相结合的公司Hybri目前已推出首个AI虚拟助手应用。使用者可以基于照片生成虚拟形象，该形象将以AR形式持续存在，并进行简单交互。该应用目前已获得近4万使用者。而知名AI虚拟助手，包括小冰、小爱同学等，都已在为其虚拟形象定制专属虚拟数字人形象。但其目前仍处于对外展示阶段，由于对实时渲染等方面的需求，尚未广泛搭载。据三星Neon团队，他们期待Neon未来能够拥有自己的情感和记忆。

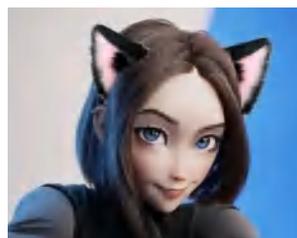
但由于AI助手大多由手机厂商或底层系统上进行开发，技术成熟后的增长可能是爆发式的。据ResearchAndMarkets Global Intelligent Virtual Assistant Market（2020 to 2027）预测，到2027年，全球智能虚拟助手市场规模预计将达到451亿美元，复合年增长率为34.0%，可从中对虚拟数字人助手的市场规模有所想象。



Hybri的虚拟AI助手



小爱同学的虚拟形象（小米公司）



三星助手虚拟形象Sam

在多模态AI助手领域，其规模增长关键在于相关设备的成熟，例如中老年市场的手机终端，物联网终端，VR眼镜的推广，甚至可能车载HUD的车载助手中。

2.3.2 身份型虚拟数字人

与缺乏人格象征的服务型虚拟数字人相比，身份型虚拟数字人更强调其本身的身份。

一方面，在现实世界中，具有独立人设的虚拟IP能够通过静态（照片）、动态（视频）、实时（直播）等方式引发关注及变现。另一方面，为每个人未来的虚拟世界乃至Metaverse中创造专属身份，将成为必然趋势。尽管对个性化身份的诉求已经长期存在，但在黑科技看来，身份仍大多以profile的形式存在，仅仅是一个过渡形态，人始终需要全感官的参与。虚拟数字人本身及其衍生市场都将迎来长足的发展。

a. 虚拟IP/偶像

- **核心价值及主要功能：**满足IP运营的视频化、碎片化、实时化需求；创造更为稳定、易得、低成本的IP — 创造全新虚拟数字人IP；或将漫画、小说等IP立体化运营
- **主要赛道：**文娱领域
- **未来推动力：**新世代对虚拟化内容的向往；使用真人IP的风险增加；品牌方对定制化代言人的需求
- **主要判断：**我国的虚拟IP在网红/直播领域更易爆发；面向广大用户的低门槛工具将成为重要增长点；由于能够帮助各类IP实现直播、现场路演等形式，文娱体系整体连接将加强，各类IP易实现全渠道运营；对IP的运营能力成为公司实力的重要象征

虚拟IP指其在现实世界中并不存在对应的真人，其外貌特征、基本人设、各类偏好、背景信息等均由人为设定。虚拟IP存在两种方式，全新创造的IP，以及基于文本、漫画、游戏中的原有IP，立体化后的可交互、可多渠道运营的虚拟IP。尽管在虚拟IP领域，大多会采用真人驱动的方式进行运营，但对应的真人演员（在二次元界也成为“中之人”）与虚拟IP本身并无强绑定关系。

Gensis等为明星创造虚拟形象的公司，国内黄子韬、迪丽热巴的虚拟化身本质上是真人偶像身份的延续，用于代替真人进行相关表演，属于替代真人服务中的一种，主要用于帮助真人明星增加其曝光率，并非此处所指的虚拟IP。

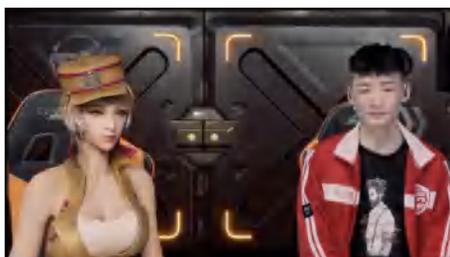
目前，虚拟IP的商业价值已经得到了充分验证。最为知名的虚拟IP Lil Miquela 2020年收入超千万美元，而在国内，虚拟IP翎已登上Vogue杂志，并获得特斯拉代言。除去传统的偶像渠道外，直播及网红也成为虚拟IP的重点发展市场。代表IP包括美国主播CodeMiko，中国抖音网红阿喜，中国B站网红鹿鸣、日本Imma。

整体而言，虚拟IP相对于真人IP，解决了MCN对特定IP长期稳定持有的问题，以偶像/网红为核心场景，在直播、代言等领域均有所发展。相对真人IP可能出现各类突发事件，收费昂贵、品牌方难以积累自有流量的问题，虚拟偶像由于可以依托机构进行运作，人设稳定，可高频次出席相关品牌活动。此外，由于我国的短视频和直播业态正在迅速发展，面对高频、碎片且实时的IP运营需求，虚拟IP更能够适应这一趋势。根据魔法科技的客户经验，相对于传统动画公司，除传统的长内容领域外，垂直虚拟数字人公司在技术及内容制作方面更能够满足目前的短视频及直播需求。

虚拟偶像方面，在初音未来、洛天依等成功案例下，目前国内已出现了专门的虚拟偶像经纪公司，SM等国内知名娱乐公司在为真人偶像创造虚拟形象的同时，开始打造虚拟偶像。在虚拟偶像方面，大多采用文娱公司+技术公司联合打造的方式。由文娱公司进行前期的形象设计、人设打造，以及后期的流量、活动运营等，由技术公司提供虚拟数字人相关技术，并为具体互动提供技术支持。据次世文化，单个虚拟偶像的孵化周期大约在45天左右，以一年半为理想的使用周期。但多家虚拟形象技术公司向我们表示，将在未来关注自身的IP运营能力，并针对特定内容呈现场景发掘潜力赛道，培养对应的虚拟IP，成熟后进行经验复制。在这一领域，对内容行业的深入了解和技术一同成为了参与公司的硬性门槛。



韩国SM公司推出的新女团Aespa
一半成员包含虚拟偶像



穿越火线游戏IP 灵狐
在斗鱼进行现场直播



美妆品牌花西子同名虚拟代言人

而直播领域则为我国虚拟文娱的发展特色，是虚拟数字人UGC（User-Generation Content）的典型代表，在创造商业价值的同时，有助于激发大众对于虚拟世界的适应力和想象力。包括爱奇艺、B站、淘宝直播、虎牙直播、酷狗音乐、腾讯、抖音平台都已经开始使用虚拟主播。由于虚拟数字人技术使视频类和实时演绎类运营的门槛降低，小说IP（阅文集团等）、漫画IP（齐麟等）、游戏IP（灵狐等）等也正在积极通过虚拟数字人技术，使原有单一模态的IP能够实现直播、线下路演等，获得了IP运营新的价值增长点。C端直播的软硬全栈解决方案正在成为多家厂商的共同发力点，由于潜在市场规模巨大（直播市场目前市场规模约1.5万亿），这一细分市场也是目前最易出现跨界竞争者的领域。

代言领域，欧莱雅（M姐）、花西子（花西子）已开始有所尝试，通过虚拟IP打造完全符合品牌调性的虚拟代言人。在完成虚拟代言人打造后，可将该特定IP延展至客服等衍生的虚拟数字人领域，持续加深品牌印象，出于品牌对宣传效果的考虑，对虚拟数字人的形象质量和交互效果有极高的要求。因此，此类代言人的相关活动大多交由魔法等CG类厂商，通过真人驱动达成。

在面向平台或特定品牌提供了虚拟偶像、虚拟代言人等B端产品后，身份型虚拟数字人公司将转向面对C端的轻量化产品，通过降低软硬件要求、减少成本等方面赋能虚拟数字人UGC，实现人人可虚拟。据公开渠道，目前单一虚拟数字人的制作成本已从六位数下降了五到十倍。

b. 虚拟世界第二分身

需求得到过往验证，需要相关基础设施支撑

- **核心价值及主要功能：**满足个人对虚拟身份的需求 — 为社交、娱乐、Metaverse设计脱离于现实世界的第二分身
- **主要赛道：**互联网社交、游戏、虚拟内容
- **未来推动力：**新时代对虚拟化内容的向往；Metaverse带动的虚拟数字人基础设施普及率（以年为单位进行迭代）
- **主要判断：**在用户需求方面已得到过往捏脸应用的充分认证，市场潜力明确；社交为其天然的适用场景，高开放性的社交类游戏可能成为其首先爆发点；Metaverse基础设施构建、XR设备规模化出货都会成为重要推动因素

分身型虚拟数字人主要面向的是未来的虚拟世界，把为每个人创造自己的虚拟化分身作为最终目的，满足个人在虚拟世界中的身份需求。为自己创造独特的形象在社交、游戏等领域被反复验证过，QQ秀、Nepeto、Anmoji、逆水寒等的成功都与这种底层需求紧密相连，但过往多为低还原度的平面形象，虚拟第二分身将有望通过其特有的真实感和沉浸感进一步满足这种分身需求。且用户需要在虚拟形象的创造上拥有更多话语权。基于虚拟身份，还可能出现以NFT服饰等为代表的虚拟消费市场。

在与VR从业者交流后，我们发现，尽管高还原度的个性化外表有助于提升第二身份的沉浸感。但核心问题在于交互手段。第二身份的核心在于精细的描述使用者当前的反应姿态，包括位置、外貌、注意力、情绪等一系列要素，才能为用户提供基于第二身份的存在感。因此，第二分身虚拟数字人的普及需要一系列基础设施完善。在虚拟产业的内容、硬件、软件等方面基本成熟后，虚拟数字人将有望迎来快速爆发。

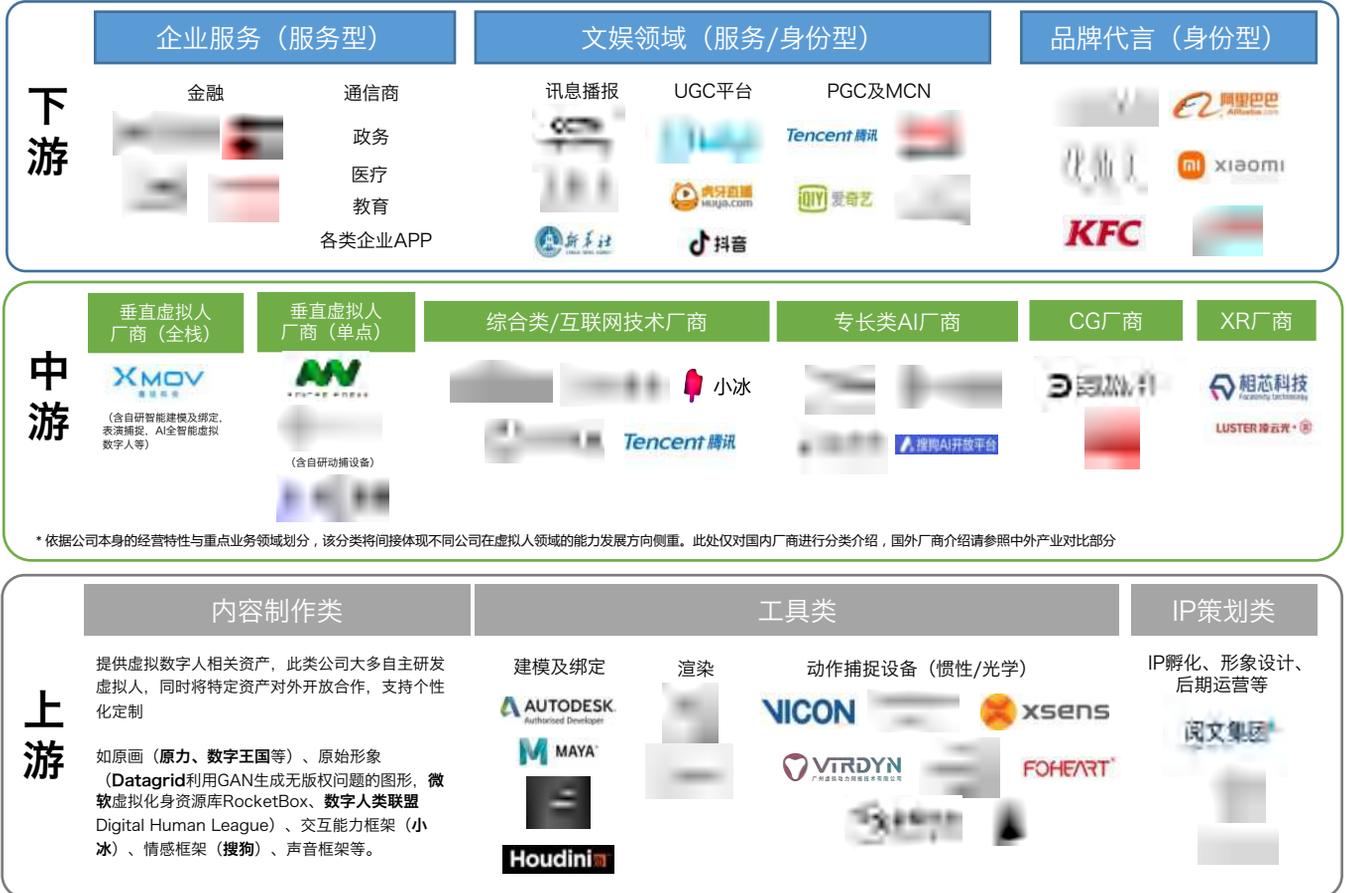
社交领域为目前主要的应用场景，国内包括号称“社交元宇宙”的Soul，带有面部捕捉功能的Vyou，各类VR chat应用的等。英伟达甚至将其应用于视频会议中，通过其Vid2Vid Cameo帮助使用者以固定的真人形象参加视频会议。

与Metaverse 概念的结合，为虚拟数字人的第二身份定位提供了充足的想象空间。相芯科技将虚拟形象视为用户与虚拟世界链接交互的第一入口。在基于虚拟形象的交互中，参与者得以进行虚拟生产和劳动，并最终形成虚拟生产和社会关系。在

业内人士的想象中，在未来的Metaverse中，虚拟数字人将作为品牌为用户提供服务的核心中介，在作用上等同于目前的APP。

Metaverse方面，部分厂商已有所行动。Wolf3D（Ready Player Me头像制作器）、Unreal Engine（MetaHuman Creator）、Oculus（Avatar SDK2.0）、Loom.ai（对外开放其API，已被Roblox收购）均已为用户提供生成个人虚拟形象的平台，用户可在生成后将带有绑定的角色直接导入VR游戏、VR Chat、内容工具等进行使用。Oculus表示，用户创建的虚拟化身最终会以某种形式出现在Facebook、Messenger、Instagram等应用中。

2.4 虚拟数字人产业链



尽管目前大多国内厂商选择对形象、语音、语言等多方面进行自研，但多家厂商已明确表述，未来会将非自身核心的技术部分进行外包。目前，我们已经从个别头部公司处了解到，多家不同类别的虚拟人厂商正在以各自的核心技术能力达成合作，在教育等领域快速试水。这一趋势将与国外产品目前的商业模式相似。

2.5 竞争格局及中外对比

从竞争格局来看，无论在国内还是国外，虚拟数字人的竞争格局都尚未成熟，行业集中度相对较高。尽管玩家数量在快速增长，但由于行业对综合的技术实力和合作渠道能力要求较高，具有相当实力的头部公司仍相对有限。在中国，在虚拟直播、虚拟偶像等潜力确定的细分市场，出现部分针对中小客户、或承接经纪等延展业务的专业虚拟数字人公司。

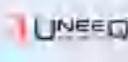
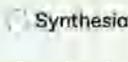
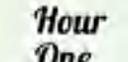
国内外在细分市场上的竞争差异较大。国内外目前的共同市场是自动生成虚拟内容，但由于在外表细节、预设模板、配套系统等方面的差异，国外在该细分市场具有较为明显的产品和客户优势。此外，虚拟偶像受到较多的共同关注，亚洲地区文娱公司等产业方的参与尤为突出。而在其他市场上，国外更关注情感关怀的顾问/助手类虚拟数字人和用于打造虚拟数字人的应用，而这两个细分市场在国内并无太多发展。国内的玩家更关注虚拟客服类场景。对虚拟直播的高度关注成为我国玩家的特有方向。

整体而言，国内外在细分市场上的竞争各家仍处于试验产品可行性、打造标杆客户的阶段，距离商业化规模推广仍具有一段距离。

从整体对比来看，我们发现，与差别主要存在于具体技术和产品的其他技术领域不同，中外的虚拟数字人厂家在具体玩家、技术水平、关键场景、产品形态、运营方式等方面都具有明显的差别。综合来看，这种差异实质上是行业性的差异。通过综合考察中外玩家之后，我们认为这种体系化的差异起源于技术，但同时也是我国直播业态兴起、互联网元素复杂、人口基数大等多方面因素共同助推而成的。

但基于对虚拟数字人本质的理解，我们认为，在未来，中外依旧会在具体的使用场景上有所合并。这种体系化差异不应当被理解为不同的发展速度，未来即便是在同一细分市场，也可能会出现完全不同的业态。

2.5.1 国外主要玩家

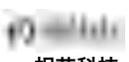
公司名称	公司定位	融资状况	典型市场	公司特性
 UneeQ	人性化的AI互动服务，强调拟人化的情绪带来的商业价值提升	2018年A轮 1000万美元	医疗康复/心理关怀顾问、虚拟客服、品牌大使	项目制和订阅制并行，单个定制化项目周期为6个月（2个月深度涉及，4个月数据收集及算法训练，2个月进一步调整） 产生的数字人可在PC、手机等多种终端运行 在NLP方面高度兼容，兼容各类数据库、知识库、NLP及聊天机器人系统
 Soull-Machines	人性化的人机交互系统，强调Engage with customers与虚拟形象AGI	B轮5750万美元 总额6500万美元	产品顾问、网站客服、虚拟购物助理、在线银行客服、医疗保健助理等	通过其构建的完整流程、预设库（Human OS Platform）和创建平台（The digital DNA Studio），使用者可以构建具有不同性格、外表、价值观的虚拟数字人
 Samsung Neon	提供视频聊天机器人	-	财务顾问；瑜伽教练、电视主播等	Neon 被三星视为视频聊天机器人，（核心功能是“a person you need to talk to”）。由于具有自己的情感和记忆，可以学习人们的喜好并拟人化的回答询问 采用订阅制，且在回应上具有更高的准确率
 TwentyBN	本身为强调视频数据、深度神经网络训练的计算机视觉解决方案的AI公司	2018年A轮 1000万美元	定制型AI教练Fitness Ally	强调多方位感知、个性化交互等，采用订阅制。具有本地延迟设计
 DataGrid	AI图像/视频生成公司 提供虚拟数字人、图像及视频生产、图像加工方案。 以生成对抗网络作为技术核心	2019年种子轮 36万美元	静态模特、虚拟偶像、线上教学等，有交互界面的规划	强调通过GAN避免版权问题，已和图片网站Imagenavi合作，对外出售虚拟形象图片 目前已成立合资的虚拟偶像公司
 Synthesia	根据文本，自动生成虚拟数字人的视频	2021年A轮 1250万美元 (视频领域最高)	虚拟内容自动生成，网页向导	提供40个现有虚拟形象，个性化虚拟形象定价1000美元，只需提供模特10分钟的绿幕视频 提供订阅制，个人用户30美元/月 2020年营收超700万人民币，但整体亏损
 HourOne	根据文本生成虚拟数字人内容	2020年种子轮 500万美元	产品演示视频、新闻播报、教学课件等	36个预置形象 提供有丰富的视频模板，可进行自行设计 提供有视频管理系统，用于管理发布、检测流量等 在内容输入中直接对接文本、数据、CRM等 支持用户以自身形象做发言人

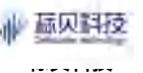
公司名称	公司定位	融资状况	典型市场	公司特性
 Unreal Engine	影视及动画引擎	-	MetaHuman Creator, 可通过自定义创造高还原度的虚拟形象	服务免费, 要求使用EPIC引擎渲染 极大降低了CG的技术门槛和时间消耗 已自动完成绑定, 可直接加工为游戏NPC等
 Genies	基于个人形象, 生成虚拟分身	2021年B轮 6500万美元, 总额1.17亿美元	3D卡通化身生成	早期主要为明星打造代营业的虚拟化身, 现已支持社区交流和NFT物品交易
 Loom.ai	基于个人形象, 生成虚拟分身	已被Roblox收购 收购前融资近600万美元 三星参投	3D卡通化身生成	
 Oben	基于个人形象, 生成虚拟分身	B轮1000万 总融资约2300万美元 腾讯参投	3D卡通生成	附有其他图像增强等功能。目前已和韩国文娱公司SM成立合资公司, 用于打造虚拟偶像
 Spatial	VR+AR协同办公	A轮1400万美元	生成会议用的虚拟形象	受扫描技术和算力限制, VR会议软件大多采用相同的技术和运营思路, 以虚拟形象代替真人, 提升会议体验感

2.5.2 国内主要玩家

由于国内在虚拟数字人领域的头部玩家大多并不专注于在这一领域, 主要是基于已有的相关技术和业务基础进入该行业, 我们在此不对各公司的成立年限和融资状况进行介绍。

值得注意的是, 除去AI厂商和CG厂商外, 我们观察到部分动作捕捉设备厂商进行软件自研, 进入了虚拟直播领域。我国虚拟数字人领域未来的玩家生态可能会更加丰富, 具体所在的业务领域与切入方式将和公司基因高度相关, 头部客户和中部客户差异明显。

公司名称	公司定位	虚拟形象产品	代表行业/客户	备注
 魔法科技	虚拟数字人垂直厂商 致力于用顶尖的计算机图形学和AI核心技术打造三维虚拟内容制作及虚拟数字人基础设施, 赋能虚拟+X全新生态。	虚拟内容协同制作智能云平台、专业级及消费级虚拟直播技术和产品、服务型虚拟数字人能力平台、身份型虚拟数字人全栈式技术和产品服务	身份型: 魔法原创(翎Ling、Sam&Liz)、腾讯(灵狐)、阿里巴巴(天猫喵酱)、阅文(叶修)、城市IP(宁波市等)、小米, 美的、伊利等; 服务型: 魔法原创(Ada)、金融(光大银行、太平金科等); 教育(杭州萧山教育局、字节跳动等); 电商(怡宝等); 医疗(上海精神卫生中心虚拟心理医生); 运营商(中国移动、中国联通)	魔法是为三维虚拟内容制作提供智能化、工业化的基础设施; 为三维虚拟数字人提供造人、养人、用人的全栈式技术和产品服务。
 相芯科技	领先的XR技术创新者。专注于图形学技术和人工智能技术的深度融合创新, 为客户提供XR内容生产与互动的解决方案。	AI虚拟助手、AI虚拟主播	虚拟助手(金融、党建、政务、电商、博物馆); 虚拟主播(济南电视台、广西日报、人民日报、浙江电视台钱江频道、京东); 美颜滤镜AR特效(虎牙直播、珍爱网、新东方教育)	专注于提供VR/AR/移动应用的支撑技术
 追一科技	专长类技术厂商 领先的人工智能公司和AI数字员工提供商, 主攻深度学习和自然语言处理, 提供智能语义、语音和视觉的AI全栈服务	多模态数字人	银行、政务、保险、券商、运营商、教育、零售/电商、公共交通、文旅、媒体/娱乐、居家/健康	

公司名称	公司定位	虚拟形象产品	代表行业/客户	备注
 黑镜科技	虚拟数字人内容平台厂商 自研真人 AI 形象引擎、算法、资产工作流等虚拟人全栈技术，致力于为全球创作者提供基于 AI 的虚拟内容生成服务与平台	虚拟内容生成平台，真人 AI 形象引擎 /SDK，虚拟人资产算法，虚拟人动画算法	互联网内容平台，电商，传媒，广告，专业/半专业创作者，虚拟人引擎开发者	
 中科深智	虚拟数字人垂直厂商 致力于持续研发数字虚拟人驱动引擎和技术美术平台，为客户提供极高性价比的数字人全栈式解决方案。	创梦达及创梦加（专业级动画），创梦易（快速动画生产及虚拟直播），创梦易自动播及云小七（7*24小时AI虚拟直播）	为一禅小和尚、默默酱、萌芽熊、僵小鱼等头部IP，央视、中兴、中石化、中国银联等龙头企业，以及美的、蒙牛、怡思丁、安莉芳等 2000+ 电商品牌提供了数字人解决方案。	持续扩大资产库规模，已有近500个原创虚拟IP授权给客户使用
 科大讯飞	专长类技术厂商 亚太地区知名的智能语音和人工智能上市企业	AI虚拟主播、AI虚拟数字人	虚拟新闻主播（新华社、交汇点新闻、温州都市报、金陵晚报、广西卫视）；虚拟数字人（客服、交通、政务咨询、生活服务）	
 火山引擎	综合类技术厂商 字节跳动面向企业的智能科技品牌	捏脸及表情驱动 虚拟主播 虚拟形象生产平台	互动娱乐、广电传媒播报、金融客服、车载娱乐等行业及场景	个性化3D虚拟形象支持实时面部表情、肢体动作与语音驱动
 百度	综合类技术厂商 全栈的AI能力，提供端到端软硬一体的应用	虚拟主播直播、虚拟形象智能交互解决方案	虚拟主播（央视网、好看视频、中国天气） 虚拟形象智能交互（问询导览、广告营销、迎宾） 虚拟数字人导购员，虚拟主播，虚拟培训师等	
 搜狗	专长类技术厂商 搜狗推出的以语音交互技术为核心的人工智能平台	搜狗分身	传媒、教育、金融、娱乐 虚拟主播（搜狗虚拟气象主播、新华社AI合成主播、新华社3D AI合成主播）、金融虚拟客服（平安惠普、平安保险）、车载语音助手（蔚来汽车）、明星二次元形象（黄子韬）、声音内容制作	2021年推出搜狗手语虚拟主播
 小冰 小冰公司	面向新交互形式的完整人工智能技术框架 “在亿万万人之中，我只属于你”，其个性交互框架已对公众开放	小冰	虚拟亲友、AI托管小编、虚拟歌手、AI主播、测颜值替身 Beta版本现已开放多种虚拟数字人类生产线	
 亚贝科技	专长类AI厂商 专注智能语音交互和AI数据服务的人工智能公司	智能客服、虚拟员工、虚拟主播、恐龙贝克	智能客服、虚拟主播、虚拟员工	
 拟仁智能	CG类厂商 是一家将人工智能（AI）与计算机图形学（CG）技术相结合的创新型高科技公司	定制化虚拟数字人	个人助手、智能客服、电商零售、旅游文创、教育医疗、广告营销、直播主持、游戏娱乐	
 网易伏羲	综合类厂商 国内专业从事游戏与泛娱乐AI研究和应用的顶尖机构	网易伏羲AI虚拟数字人	虚拟主播、虚拟客服、虚拟助手、虚拟直播、游戏动画、空间虚拟智能（虚拟导览、虚拟歌手、虚拟诗人、交互式虚拟数字人）（北京市文物局、杭州市中国丝绸城项目、为麦当劳提供虚拟数字人应用）	

2.5.3 中外综合对比

在综合对比了虚拟数字人行业的中外公司后，我们发现双方存在体系性的行业差异。技术重点、细分市场、商业模式、运营方式等均存在较大差异。我们认为，这种整体性差异的直接原因在于技术重点和商业侧重的不同。但同样受到我国乃至亚洲文娱消费、思想观念等方面的影响。短期内，预计这种差异将持续，中外将依据不同的发展路径催熟各细分市场。但基于我们对虚拟数字人本质的认知，长期双方将出现加速融合趋势。

共同点



创始团队

- 一派为AI技术公司，核心是“技术”。凭借其综合的感知类技术扩展业务，进入虚拟数字人生成领域。该部分公司大多专注于自动生成虚拟数字人内容的细分市场
- 另一派为在计算机图形学或传统CG技术上有深厚技术积淀的公司，核心为“美术”。在中国，此类公司在虚拟偶像等对细节要求更高的领域有更明显的优势。而在国外，此类公司则更关注关怀式的虚拟数字人生产



细分市场

- 虚拟偶像和虚拟内容生成成为共同的重点市场



竞争状况

- 均处于竞争格局不成熟阶段，头部公司具有明显的先发优势，通过标杆客户验证其需求及技术可行性，并产生行业示范效应



技术层面

- 国外更多的选择了高保真+计算驱动的技术思路。在外表方面，国外公司呈现出了更为理想的效果，虚拟数字人形象还原度极高。驱动方面，肌肉与骨骼的细节驱动方面效果明显，高度还原人体正常运行状况。据中科深智介绍，这与我国技术美术人才（同时兼顾美术修养及软件技能）的缺失相关
- 这种差异可能来源于国内审美对于风格化的倾向、国外CG和光场等技术的发展，国外更多选择真人进行1:1建模，数据采集量和模型设计不同等方面



商业模式

- 国内厂商大多选择自主承包语音、文字和图像能力，国外则选择将核心的NLP能力对包，在产品中接入微软、AWS、Google等对话系统，并支持客户自主选择，在业务上更加集中。这种差异可能来源于国外以CG厂商为主的公司格局，多位国内厂商表示，也将在未来将部分能力开放给外部厂商。这一差异在CG类厂商中未来可能消失
- 国外更为标准化，会提供大量预设的虚拟数字人形象，并提供更为丰富的音色、模板等。国内提供相对有限，部分厂商会基于用户需求进行特定设计，定制化项目占比较高
- 国外的交付以特定虚拟数字人为主，会在客户方完成训练与迭代，国内以内容交付为主



细分市场

- 国内更注重风格化和特定专业领域，据某业内人士，这种选择有助于回避在特定上的技术差异
- 国外以提供人性化助手为核心，强调对人的情感关怀。国内则更强调现有生产方式的简化和内容的丰富



发展路线

- 国内大多是传统的AI公司，在这个赛道中基于虚拟直播或虚拟主持人播报起家，并开始关注这两项技术的周边产业，如完整的直播运用方案、IP运营等并逐渐向其他领域延展。但由于国内市场和内容消费的市场太大，缺乏业务延展的动力。部分公司在向消费级的完整解决方案拓展
- 除去游戏引擎公司Epic是服务于引擎的，其他公司大多是纯粹专一于这一市场的。就目前动向而言，可能会沿垂直行业发展
- 国外企业更多聚焦于2C，延续了从日本兴起的虚拟陪伴文化（这一方向在国内主要体现在AcFun和哔哩哔哩中）。我国企业则多聚焦于实用性更强的2B2C领域，通过客服等服务型场景降本增效，更快的兑现技术价值

差异点

2.6 进入门槛及行业瓶颈

在追一科技等厂商看来，技术是进入行业的核心门槛，需要进入玩家在机器视觉、语音交互和自然语言处理方面都具有深厚的技术积累，并将三者进行有机结合。科大讯飞则在技术方面强调了虚拟数字人向自我管理的演化，认知智能、灵活性、个性化、情感化都是虚拟数字人需要进一步提升的技术方面。虚拟数字人动作的细腻度和丰富度被魔法科技等厂商强调。



上限极高的技术 门槛

- **外观呈现效果**：涉及CG建模及驱动模型的精细程度。外表超写实方面依旧需要技术公司主导，而驱动方面细节不足，跳帧、形变等问题同样亟待解决
- **交互自然度与认知智能**：NLP对话模型中语义理解、多轮对话、上下文理解、情感识别、现实逻辑等。
- **个性化程度**：数据迭代、个性化设计等，影响虚拟数字人能否承担私人顾问、私人助手等职责
- **算力消耗优化**：直接影响能否进行实时渲染，限制了虚拟数字人的使用场景
- **硬件性能及渲染能力**：服务端渲染、云端渲染、消费级手机终端渲染，要去产品在当前场景下具有快速细腻的渲染能力，会涉及硬件设计及供给、网络模型等
- **技术平民化**：通过相关技术降低产品的使用和成本门槛
- 审美、核心架构等相关技术，CV、TTS、NLP三方面的综合技术能力

而在追求提供完整交付的厂商，如火山引擎看来，好的厂商需要站在客户的业务场景之下考虑虚拟形象给客户带来的价值和效率提升，在很多场景中需要参与到业务的体验以及设计中。由于虚拟数字人行业未来需要和各垂直行业进行深度融合，对业务的深度了解同样成为重要门槛。相芯科技同样认为前期的行业探索有助于自身进一步了解客户的诉求



客户积累带来 的业务认知

- **客户熟悉度**：在熟悉客户中易拓展创新性业务，寻求当前技术水准可满足的需求场景
- **业务系统对接**：对客户业务系统相对熟悉，并连接背后业务系统，提升产品附加值
- **行业knowhow积累**：有助于针对性改进产品，在提升产品价值的同时，快速拓展垂直行业赛道。这一点在直播领域体现为对流量分发逻辑、客户购买动机等要素的洞察会影响虚拟数字人的动作、话术、设计审美等问题



在各细分市场亟待 实现的标准化复制

国外采用预制大量虚拟形象，提供充分模板的方式进行解决。对于国内企业，由于主要**应对的场景高度定制化**。就目前而言，**定制化项目的整体工期和客单价（百万级别）都较高**，在大规模复制上存在一定问题，导致虚拟数字人的应用无法快速推广到各行业。

一方面是寻找合适的场景合作方，以**SDK/API实现标准化**，目前我们尚未发现最为合适的应用场景。另一方面，则是以**SaaS的形式，直接提供最终成果**，目前直接生成虚拟数字人播报视频的业务正在以这种方式提供，但其所占市场份额相当有限。