

AI 数字人直播: 开源与私有化部署解决方案研究报告

1. 引言

1.1. 定义与背景

AI 数字人直播, 或称 AI 虚拟形象直播, 是指利用人工智能技术创建和驱动虚拟形象(数字人), 并将其应用于实时直播场景的过程。这涉及到计算机图形学、语音合成(TTS)、自然语言处理(NLP)、语音识别(ASR)以及实时音视频流处理等多项技术的融合¹。近年来, 随着 AI 技术的飞速发展和算力的提升, 数字人的真实感、交互能力和生成效率显著提高, 使得 AI 数字人直播从概念走向实际应用。

1.2. 市场驱动力与应用场景

AI 数字人直播的兴起受到多重因素驱动。首先, 它能显著降低直播成本, 无需聘请真人主播、租赁场地或购买昂贵的设备, 尤其对于需要 24/7 不间断直播的场景, 成本优势明显²。其次, AI 数字人可以克服语言障碍, 通过多语言 TTS 技术触达全球观众², 并保持品牌形象的一致性。此外, AI 数字人还能提供高度个性化的互动体验²。

这些优势使得 AI 数字人直播在多个领域展现出巨大潜力, 包括:

- 电商直播带货: 7x24 小时在线, 进行产品展示和销售¹。
- 在线教育与培训: 提供标准化的、可重复的教学内容和虚拟辅导⁷。
- 虚拟客服与品牌大使: 提供信息咨询、客户支持和品牌推广⁸。
- 新闻播报与内容创作: 自动化生成新闻播报或信息发布视频¹⁰。
- 虚拟活动与娱乐: 在虚拟会议、展览或娱乐活动中担任主持人或表演者⁸。

1.3. 研究目的与范围

尽管云 SaaS (Software as a Service) 平台 (如 HeyGen¹⁴, Synthesia⁷, D-ID²⁴ 等) 提供了便捷的 AI 数字人创建和直播服务, 但对于追求更高数据控制权、安全性、定制化能力或有特定集成需求的企业而言, 开源解决方案和商业私有化部署方案成为重要的考量选项。

本报告旨在深入研究 AI 数字人直播领域的开源项目和提供私有化部署 (本地部署 / On-Premise) 的商业解决方案。具体研究范围包括:

- 识别和评估相关的开源项目, 分析其功能、成熟度、技术栈和部署要求。
- 调研提供私有化部署选项的商业供应商, 分析其方案特点、技术架构、安全性和定价模型。
- 比较开源、私有化部署和云 SaaS 三种模式在 AI 数字人直播应用中的优劣势。
- 总结研究发现, 为不同需求场景提供决策参考。

2. AI 数字人直播关键技术解析

构建一个功能完善的 AI 数字人直播系统, 需要整合多种关键技术模块。理解这些技术有助于评估不同解决方案的技术深度和能力。

2.1. 虚拟形象创建与渲染

这是数字人的视觉基础。创建方式多样, 可以通过 3D 建模软件从零开始设计, 利用游戏引擎(如 Unreal Engine²⁰)或专门平台(如 Ready Player Me³⁰)生成, 甚至基于真人照片或短视频生成数字分身(Custom Avatar/Digital Twin)⁷。渲染技术决定了数字人的视觉真实感, 从传统的基于多边形的渲染到新兴的神经辐射场(NeRF)技术³⁶, 都在不断提升渲染质量和效率。实时渲染是直播应用的关键, 要求在保证视觉效果的同时, 满足低延迟的要求⁸。一些开源项目如 ai-iris-avatar⁴⁰ 利用 Unity 引擎, 而 NVIDIA 的方案则利用 Omniverse Renderer²⁰。

2.2. 实时动画与唇语同步(Lip-Sync)

为了让数字人看起来“活”起来, 需要实时生成自然的动作和表情。

- 身体与面部动画: 可以基于预设动画库(如 Mixamo³⁰)随机组合, 或通过 AI 模型根据语音语调、情感或特定指令(如 emoji³¹)实时生成。NVIDIA 的 Audio2Face 技术²⁰ 可以根据音频驱动面部动画。一些项目还加入了物理模拟(如头发⁴⁰)和细微动作(如眨眼、眼球微动³⁰)来增强真实感。
- 唇语同步: 这是实现自然对话感的关键。需要将输入的文本或实时生成的语音, 精确地映射为虚拟形象的口型变化。常用的技术包括基于规则的映射、基于音素(Phoneme)的分析(如 Rhubarb Lip-sync), 以及基于深度学习的模型(如 Wav2Lip³⁷)。这些模型需要将音频特征实时转换为驱动口型变化的参数(如 BlendShapes 或 Visemes³¹)。延迟和准确性是衡量唇语同步技术优劣的核心指标。开源项目 LiveTalking⁴³ 和 Ditto⁴⁶ 都重点关注了这一领域。

2.3. 语音合成(Text-to-Speech, TTS)

TTS 技术将文本转换为自然流畅的人类语音。现代 TTS 系统(如 Google Cloud TTS³¹, ElevenLabs²⁹, Azure TTS³¹, Coqui TTS, fish-speech⁴⁴, 腾讯语音合成⁴⁴)能够生成多种语言、口音和情感风格的语音⁵。对于直播场景, TTS 需要具备低延迟特性, 最好能支持流式合成, 即边生成边播放, 以减少用户等待时间。声音克隆(Voice Cloning)技术允许创建特定人物(如名人、CEO 或用户自己)的 AI 语音, 增强个性化和品牌辨识度²¹。

2.4. 语音识别(Automatic Speech Recognition, ASR)

在交互式直播场景中, ASR 技术用于将用户的语音输入转换为文本, 以便后续的自然语言理解和处理⁴¹。常用的 ASR 引擎包括 OpenAI Whisper³⁰、Google Cloud Speech-to-Text⁵² 等。实时性和准确性是 ASR 在直播应用中的关键要求。一些开源项目集成了不同的 ASR 模型, 如 LiveTalking⁴⁴ 和 handcrafted-persona-engine⁴⁹。

2.5. 自然语言理解(NLU)与对话管理

为了让数字人能够理解用户意图并进行有意义的对话, 需要 NLU 和对话管理技术。

- **NLU**: 利用 NLP 技术(通常基于大型语言模型 LLM, 如 GPT 系列²⁰、Claude⁵⁴、Gemini²⁴ 等)来理解用户输入的文本(来自 ASR 或直接输入)的含义、提取关键信息和识别用户意图。
- **对话管理**: 负责维护对话状态、决定下一步行动(如回答问题、反问、执行任务)并生成回复内容。这通常也由 LLM 完成, 或者通过更传统的基于规则或状态机的系统实现。
- **知识库与 RAG**: 为了让数字人能够回答特定领域的问题或提供基于特定知识的回复, 通常会结合知识库。检索增强生成(Retrieval-Augmented Generation, RAG)技术允许 LLM 在生成回复前, 先从指定的知识库(如公司文档、产品手册²⁰)中检索相关信息, 提高回复的相关性和准确性¹⁶。

2.6. 直播流集成

最终, 生成的数字人音视频内容需要通过直播流协议(如 RTMP⁴³ 或 WebRTC⁴³)推送到直播平台(如 TikTok, YouTube, Shopee, Lazada 等⁵)。这需要确保音视频同步、低延迟传输, 并能与直播平台的互动功能(如评论读取、实时回复⁶)相结合。一些解决方案提供了与 OBS (Open Broadcaster Software) 等直播软件的集成能力⁵。

3. 开源 AI 数字人直播解决方案

对于希望获得最大控制权和定制能力, 并且具备相应技术实力的组织而言, 开源解决方案提供了一条可行的路径。然而, 该领域目前呈现出一定的碎片化特征, 完整的端到端开源直播系统相对较少, 更多的是专注于特定环节(如 TTS、唇语同步、渲染)的库或框架。用户通常需要自行整合多个组件来构建完整的解决方案。

3.1. 开源项目概览与评估

以下是一些在 GitHub 等平台上受到关注的相关开源项目:

- **项目名称: LiveTalking⁴³**
 - **核心功能**: 这是一个相对完整的、专注于实时交互式数字人直播的开源系统框架。它整合了多种数字人模型(ernerf, musetalk, wav2lip, Ultralight-Digital-Human)、支持实时音视频同步对话、声音克隆、打断、全身体拼接、RTMP/WebRTC 推流以及多并发处理。
 - **成熟度**: 项目较为活跃, 近期有代码和模型更新(如支持新模型、优化并发、添加 TTS 服务)。拥有 5.4k+ 星标和 800+ 复刻, 表明社区关注度较高。存在 300+ 个 Issue, 说明用户在积极使用并反馈问题。提供了 Docker 镜像, 简化了部署。⁴⁴
 - **社区支持**: 提供了文档、知识星球和微信公众号等支持渠道⁴⁴。有 14 位贡献者⁴⁴。
 - **技术栈**: 主要使用 Python (70.3%), 也包含 Cuda (12.0%) 和 JavaScript (9.4%)

⁴⁴。依赖 PyTorch、CUDA。集成了 ER-NeRF、MuseTalk、Wav2Lip、Ultralight-Digital-Human 等模型, 以及 fish-speech、腾讯 TTS 等服务 ⁴³。包含多种 ASR 模型支持 ⁴⁴。

- 软硬件要求: 推荐 Ubuntu 20.04、Python 3.10、PyTorch 1.12、CUDA 11.3。需要 GPU(尤其是 ernerf 模型)。支持 macOS GPU 推理。⁴⁴
- 许可证: Apache-2.0 ⁴³。
- 定位分析: LiveTalking 是目前较为接近端到端开源直播解决方案的项目, 其框架设计考虑了实时交互的多个方面。Apache-2.0 许可证对商业应用友好。但仍需要一定的技术能力进行配置、部署和可能的二次开发。
- 项目名称: **handcrafted-persona-engine** ⁴⁹
 - 核心功能: 一个 AI 驱动的交互式头像引擎, 特别适用于 VTubing 和虚拟助手场景。集成了 Live2D 动画、LLM 对话、ASR(Whisper)、TTS(自定义模型)、可选 RVC(实时声音克隆)、字幕显示和 Spout 输出(用于 OBS 等直播软件)。
 - 成熟度: 有 V2.0.0 版本发布, 表明项目在迭代。部分功能标记为实验性。
 - 社区支持: 提供 Discord 服务器和 GitHub Issues 支持 ⁴⁹。
 - 技术栈: 主要使用 C# (97.1%) ⁴⁹。集成了 Live2D、Whisper.NET、Silero VAD、ONNX (TTS/RVC)、espeak-ng、NAudio/PortAudio、Spout ⁴⁹。可连接 OpenAI 兼容的 LLM API ⁴⁹。
 - 软硬件要求: 强制要求支持 CUDA 的 NVIDIA GPU ⁴⁹。需要 .NET Runtime、espeak-ng ⁴⁹。
 - 许可证: 搜索结果未明确提及 ⁴⁹, 需查阅代码库。
 - 定位分析: 该项目高度针对 VTubing 场景, Live2D 和 Spout 集成是其核心优势。使用 C# 技术栈使其在众多 Python 项目中独树一帜。强制要求 NVIDIA GPU 是一个关键限制。
- 项目名称: **ai-iris-avatar** ⁴⁰
 - 核心功能: 旨在实现 100% 本地运行的 AI 驱动 3D 头像对话系统。特点包括高度可配置(可替换 LLM、TTS、3D 模型)、支持自定义知识库(RAG)、聊天上下文、多语音、快速响应(<4s)、集成 Unity 引擎(利用其动画和物理系统)、Oculus LipSync、3D 骨骼动画(Mixamo)、精细的面部控制(眼动、眨眼、头发物理)和远程事件触发。
 - 成熟度: 功能描述详细, 看似可用, 但版本发布和社区规模信息不足。
 - 社区支持: 搜索结果中信息较少。
 - 技术栈: Python (服务器端)、Unity (客户端, C#)。默认使用 Ollama 运行 Gemma 2B (LLM), XTTS v2.0 (TTS), Oculus Lipsync。利用 DeepSpeed 加速 LLM 推理。使用 NativeWebSocket 通信, Mixamo 动画 ⁴⁰。
 - 软硬件要求: 设计为本地运行。推荐使用 GPU(开发者使用 RTX 3060 12GB)。需要 Python、Unity 2022.3.21f1。需要较大硬盘空间存储模型 ⁴⁰。
 - 许可证: 核心代码为 GPL-3.0。包含第三方组件(如 Oculus Lipsync, XTTS v2.0 非商业许可证), 使用前需接受其各自许可证 ⁴⁰。

- 定位分析：独特的本地优先和 Unity 引擎集成方案，提供了高视觉保真度和物理交互的可能性。但 GPL-3.0 许可证和依赖的非商业 TTS 许可证严重限制了其商业应用前景。
- 项目名称：**TalkingHead (JavaScript)** ³¹
 - 核心功能：一个用于实时唇语同步的 JavaScript 类库。支持 Ready Player Me 全身 3D 头像 (GLB) 和 Mixamo 动画 (FBX)。内置了对 Google Cloud TTS 的支持，并可通过 API 集成 ElevenLabs、Azure Speech SDK 或任何提供词级时间戳的 TTS 服务。支持将 Emoji 转换为面部表情。使用 ThreeJS/WebGL 进行渲染。
 - 成熟度：有示例项目应用，GitHub 星标数(538)表明有一定关注度。
 - 社区支持：GitHub 上有少量 Issues 和 PR ³¹。
 - 技术栈：JavaScript, ThreeJS/WebGL。依赖外部 TTS API (Google, ElevenLabs, Azure 等) ³¹。
 - 软硬件要求：支持 JavaScript 和 WebGL 的现代浏览器 ³¹。需要有效的 TTS 服务 API 密钥(通常需要付费) ³¹。
 - 许可证：MIT License ³¹。
 - 定位分析：提供了一个纯 Web 前端的技术路径，降低了 Web 开发者的入门门槛。TTS 集成的灵活性是优点，但也意味着核心功能依赖外部(通常是付费)服务。MIT 许可证使其对商业应用友好。

3.2. 其他相关开源组件与资源

除了上述相对完整的项目或框架外，还有许多专注于特定技术环节的开源项目和资源：

- **NVIDIA AI Blueprint: Digital Human** ²⁰：这是一个由 NVIDIA 提供的参考蓝图，展示了如何使用其生态系统(NVIDIA NIMs 如 Audio2Face, Riva; Omniverse Renderer 或 Unreal Engine)构建企业级数字人客服应用，特别是连接 LLM/RAG 应用。它本身不是一个可以直接运行的完整项目，而是一个架构指南，强依赖 NVIDIA 的商业或专有技术栈。这对于已经投入 NVIDIA 生态的组织具有参考价值。
- **Ditto (AntGroup)** ⁴⁶：蚂蚁集团开源的实时可控说话头合成技术，基于运动空间扩散模型。提供了推理代码和模型(TensorRT 格式)。这是一个专注于底层合成算法的研究项目，可以作为高质量说话头视频生成的一个组件，但需要集成到更大的直播或交互框架中。TensorRT 依赖意味着需要 NVIDIA 硬件。许可证为 Apache-2.0。
- **AvaChat** ⁵⁴：一个演示项目，展示了如何将 LLM (GPT, Claude) 与 D-ID 的流式 API 集成，实现实时说话头。它本身不是一个独立的开源解决方案，而是依赖于 D-ID 的商业服务。可作为集成 D-ID API 的参考示例。
- **persona (sausheong)** ⁴⁵：一个利用 Tortoise-TTS、SDXL-Turbo(图像生成)、One-Shot Free-View Neural Talking Head Synthesis(动画)、Wav2Lip(唇语同步)和 Real ESRGAN(超分)等多个特定 AI 模型来生成说话头视频的工具。其重点在于离线视频生成，而非实时流交互。
- **Awesome**

Lists:(<https://github.com/Kedreamix/Awesome-Talking-Head-Synthesis>)³⁶ 和 [awesome-talking-head-generation](#)³⁷ 是非常宝贵的资源聚合, 收录了大量关于说话头生成的研究论文和代码库(如 Wav2Lip, face-vid2vid, HeadNeRF 等)。这些列表对于寻找特定算法或技术组件非常有帮助, 但将这些组件整合成一个完整的实时直播系统需要大量的开发工作。

- 其他提及的项目: Metahuman Stream²⁸ (基于 UE5)、build-an-avatar-with-ASR-TTS-Transformer-Omniverse-Audio2Face⁴¹ (集成 NVIDIA Audio2Face) 等进一步证明了 Unreal Engine 和 NVIDIA 技术栈在该领域的应用。

3.3. 开源方案总结

开源领域为 AI 数字人直播提供了丰富的技术组件和部分集成框架。LiveTalking 是一个值得关注的、相对完整的直播框架选项, 许可证友好。handcrafted-persona-engine 则专注于 VTubing 场景。ai-iris-avatar 提供了独特的本地 Unity 解决方案, 但受限于许可证。TalkingHead (JS) 降低了 Web 开发门槛, 但依赖外部服务。选择开源方案意味着需要投入显著的技术资源进行整合、配置、部署和维护, 并且需要仔细评估各项目的许可证限制、社区活跃度和硬件要求。

表 1: 主要开源 AI 数字人直播相关项目对比

项目名称	主要焦点	关键特性	成熟度 (Stars/Forks/Activity)	主要语言/技术栈	硬件需求 (GPU?)	许可证
LiveTalking	实时直播框架	多模型支持, 音视频同步, 声音克隆, 打断, RTMP/WebRTC, 多并发	5.4k/803/活跃	Python, Cuda, JS; PyTorch, Wav2Lip, ER-NeRF, MuseTalk, Ultralight, fish-speech 等	推荐 NVIDIA GPU	Apache-2.0
handcrafted-persona-engine	VTubing/虚拟助手引擎	Live2D 集成, LLM 对话, ASR (Whisper), TTS, RVC (可选),	有 V2.0 版本/社区活跃	C#; Live2D, Whisper.NET, ONNX, Spout	强制 NVIDIA GPU	未明确 ⁴⁹

		Spout 输出				
ai-iris-avatar	本地运行 3D 头像	Unity 引擎, 本地 LLM/TTS, RAG, Oculus LipSync, 物理模拟, 高度可配置	信息不足	Python (Server), C# (Unity Client); Ollama, XTTSv2, Oculus LipSync, DeepSpeed	推荐 GPU	GPL-3.0 (核心)
TalkingHead (JS)	Web 前端 实时唇语 同步库	Ready Player Me/Mixamo 支持, Emoji 表情, 多 TTS API 集成 (Google, ElevenLabs, Azure), 可扩展语言	538/159/中等	JavaScript, ThreeJS/ WebGL; 依赖外部 TTS API	不需要特定 GPU	MIT
NVIDIA Blueprint	企业客服 参考蓝图 (非独立项目)	集成 NVIDIA NIMs (Audio2Face, Riva), Omniverse /UE 渲染, 连接 LLM/RAG	N/A	NVIDIA 生态系统	需要 NVIDIA GPU	N/A

4. 商业私有化部署解决方案

对于优先考虑数据安全、合规性、服务保障和集成便利性的企业，商业供应商提供的私有化部署 (On-Premise 或 Private Cloud) 解决方案是重要的选择。与完全依赖公有云的 SaaS 模式相比，私有化部署将部分或全部系统部署在客户自有的基础设施内，提供了更高的控制权。

4.1. 私有化部署的驱动因素

选择私有化部署通常基于以下考虑：

- 数据主权与合规性：尤其在金融¹¹、医疗⁸等受严格监管的行业，将敏感数据保留在组织内部的基础设施中，有助于满足数据本地化和隐私保护法规（如 GDPR, CCPA）的要求⁸。
- 安全性：将系统部署在内部网络，可以更好地控制访问权限，减少暴露在公共互联网上的攻击面，提供更强的安全保障⁸。
- 性能与延迟：本地部署可以减少网络传输延迟，对于需要实时交互的应用（如数字人直播），可能提供更低的响应时间⁵⁷。
- 定制化与集成：私有化部署允许更深层次的基础设施定制和与现有内部系统的集成⁸。
- 成本可控性（长期）：对于大规模、高并发的稳定应用场景，长期来看，私有化部署的总拥有成本（TCO）可能低于按使用量付费的 SaaS 模式⁵⁷。

4.2. 供应商概览与方案分析

目前，明确提供 AI 数字人直播相关私有化部署选项的商业供应商相对较少，主要有 Ravatar 和 Uneeq。其他主流 SaaS 供应商（如 DeepBrain AI, Soul Machines, Synthesia 等）的私有化部署选项则不明确或似乎并非其标准产品。

- 供应商：**Ravatar**⁸
 - 私有化部署方案：Ravatar 明确提供“On-Premise deployment (Private Cloud)”选项⁸，将其定位为重视数据主权、安全性和控制权的企业的理想选择。
 - 核心功能与技术：提供基于 RAVAPLA 平台的实时交互式 3D AI 头像。平台整合了 AI 语音、对话 AI（可连接本地/云端/混合 GPT 类模型）、自然语言生成等服务⁵⁷。支持实时对话、高度个性化定制（视觉、声音克隆、知识库、性格）、多语言能力和多种交互界面（Web、移动应用/消息、全息显示（RAVABOX/RAVALL）、AI Kiosk）⁸。采用低代码/无代码理念简化集成⁸。
 - 部署架构与要求：将头像及相关服务托管在客户自有的受控环境中⁵⁷。允许客户根据需求定制基础设施以优化性能⁵⁷。具体硬件、软件栈和网络要求未在公开资料中详述，可能在售前或 PoC 阶段确定⁵⁷。SourceForge 页面提及支持 KVM 和 Microsoft Hyper-V⁶⁰。
 - 安全性：强调通过本地化部署保障数据隐私和安全，将敏感数据保留在客户边界内⁸。
 - 支持模式：提供分级支持：基础（邮件）、标准（邮件/聊天，商业计划及以上）、高级（专属客户经理、主动监控）⁸。
 - 许可与定价：定价基于项目范围、定制程度、质量要求和功能。PoC/MVP 起价为 5,000 欧元⁸。虽然提供了云订阅计划（基础版 199 欧元/月起）作为参考，但私有化部署很可能是定制报价⁸。
 - 典型用例：AI 医疗助手、AI 酒店礼宾、AI 主题演讲者、全息远程呈现、AI 科学大使、AI 品牌大使、虚拟警官、数字复活等⁸。

- 评估: Ravatar 将私有化部署作为其核心产品之一进行推广, 特别强调安全和控制优势。其平台功能完善, 并已在多个行业落地, 全息显示集成是其独特亮点。
- 供应商: **UneeQ**²⁴
 - 私有化部署方案: UneeQ 提供灵活的部署选项, 包括本地部署(On-Premise)⁵⁶。具体方案包括:
 - **MiniPrem**: 在客户本地符合要求的硬件上运行数字人渲染器(Renny)和 Audio2Face, 实现本地渲染, 但仍需连接云端平台进行 AI 处理和管理⁴²。
 - **On-Premise Kiosk**: 将数字人部署在本地信息亭硬件上, 适用于物理世界的交互场景, 带宽需求较低⁶²。
 - **通用 On-Premise/Private Cloud**: UneeQ 2.0 平台宣称支持云、混合或本地部署⁵⁶, 暗示可能提供更全面的私有化方案, 但具体细节需与销售确认⁵⁵。
 - 核心功能与技术: 提供企业级的交互式“数字人”。平台核心包括 Synapse™ 编排引擎(连接各种 AI 服务和数据源)和 Synanim™ 实时动画技术²⁹。平台开放性高, 允许客户集成自选的 LLM(如 ChatGPT, Claude, Gemini 等)、TTS(如 ElevenLabs, Azure, Amazon Polly)和内部知识库(通过 RAG)²⁴。支持高度定制化的头像创建(UneeQ 原创、MetaHuman 导入、数字孪生、品牌吉祥物)²⁹。UneeQ 2.0 宣称响应时间低于 1 秒⁵⁶。
 - 部署架构与要求:
 - **MiniPrem 架构**: 本地 GPU (NVIDIA RTX 4080/6000 Ada+ 16GB+) 负责渲染(通过 Docker 容器运行 Renny 和 Audio2Face), 连接云端 P2 平台进行管理和 AI 处理⁴²。
 - **MiniPrem 要求**: Ubuntu 24.04, 高端 CPU, 32GB+ RAM, 256GB+ NVMe SSD, 指定 NVIDIA GPU 及驱动/CUDA 版本, Chrome 124+, 需要 P2 Portal 账户、API 密钥、TTS 密钥等⁴²。
 - **Kiosk 架构**: 本地部署在信息亭硬件上⁶²。
 - **通用 On-Premise 要求**: 未详述, 可能根据企业需求定制。
 - 安全性: 私有化部署本身提供数据本地化优势。MiniPrem 需要 API 密钥认证⁴²。UneeQ 平台设计考虑企业级安全需求⁵⁵。
 - 支持模式: 提供企业级支持, 包括培训和持续改进⁵⁵。MiniPrem 文档提及联系交付专家和使用远程访问工具⁴²。
 - 许可与定价: 主要面向企业客户, 定价基于项目范围(设计复杂性、使用量、支持需求)²⁴。私有化部署(包括 MiniPrem)很可能是定制报价。定制化头像的 IP 归客户所有²⁹。
 - 典型用例: 金融服务、零售(导购、产品专家)、技术支持、医疗保健、电信、政府、人力资源、娱乐、交通等²⁹。
 - 评估: UneeQ 平台高度灵活、可定制, 并强调开放性(AI 后端选择、IP 归属), 非常适合需要深度集成和定制化的大型企业。MiniPrem 提供了一种具体的本地渲染方案以解决性能/延迟问题, 而更全面的私有化部署可能性也存在。
- 供应商: **DeepBrain AI**¹¹

- 私有化部署方案：在其金融解决方案中提到了 AI Kiosk 的“对话式 AI 数字人选项可作为本地或基于云的解决方案”¹¹。但其主要网站和产品页面缺乏关于通用 AI Human(对话式)或 AI Studios(视频生成)私有化部署的明确信息¹¹。其 CFO 提及 Nvidia Project Digits 可用于本地原型设计，之后再部署到云或数据中心，这更侧重开发流程而非最终部署模式⁶⁴。
- 核心功能与技术(云/通用)：AI Studios (SaaS 视频生成平台) 和 AI Human (实时对话数字人平台) 是其核心产品⁵³。提供逼真头像(150+ 库、自定义)、多语言 TTS (80+)、声音克隆、多种内容到视频的转换工具、屏幕录制、与 ChatGPT 等 LLM 集成能力¹¹。
- 部署架构与要求(私有化)：除 Kiosk 外，通用私有化部署的架构和要求信息缺失。
- 许可与定价(云/通用)：AI Studios 提供分层订阅计划(免费、个人 \$24/月、团队 \$55/月/座席、企业定制)³²。AI Human 可能主要采用定制的企业定价²⁴。
- 评估：DeepBrain AI 在 AI 头像技术(尤其视频生成)方面实力雄厚。但基于现有公开信息，其通用型私有化部署选项似乎并非标准产品或推广重点。需要私有化部署的客户应直接联系销售进行确认。
- 其他供应商分析 (Soul Machines, Synthesia):
 - Soul Machines:** 早期在研究所内部署，但上市时迁移到云架构 (AWS)⁶⁵。公开信息未提及面向客户的私有化部署选项⁵²。以其“数字人”和“生物 AI”技术著称，提供逼真的情感表达²⁴。主要面向企业，可能采用定制定价²⁴。结论：大概率不提供标准私有化部署。
 - Synthesia:** 绝大多数信息指向其为云 SaaS 平台⁷。强调云平台的安全性 (SOC 2, GDPR) 和企业功能 (SSO)⁷。一个播客片段提及为银行部署“本地 LLM”⁷²，但这可能是指特定组件的混合部署或特殊案例，而非整个平台的私有化，且与其他信息相悖。结论：基本可以确定为云 SaaS 模式，私有化部署可能性极低，除非有特殊企业协议。

4.3. 私有化部署市场洞察

提供 AI 数字人直播私有化部署选项的供应商数量有限(主要是 Ravatar 和 Uneeq)，这表明与蓬勃发展的云 SaaS 市场相比，私有化部署仍然是一个相对小众的市场。这可能反映了在本地部署和维护复杂 AI 模型(渲染、TTS、LLM 等)的技术挑战和高昂成本，以及市场对 SaaS 模式的易用性、快速迭代和可扩展性的普遍偏好。然而，这些私有化选项的存在也证明了在特定企业领域(尤其是受监管行业)，对于数据控制、安全性和性能的强烈需求足以支撑这种部署模式。

表 2: 主要商业供应商私有化部署方案对比

供应商	明确提供私有化/本地部署	部署模式细节	关键特性	目标行业示例	定价模式	已知要求(若有)

Ravatar	是	完全本地/ 私有云部署	实时 3D 交互, 高度定制 (视觉/声音/知识), 多语言, 全息集成	医疗, 酒店, 公共服务	定制报价 (PoC 起价 €5k)	KVM/Hyper-V (来自 SourceForge)
Uneeq	是	MiniPrem (本地渲染混合云), 本地 Kiosk, 可能有完全本地	平台开放 (集成自选 AI/数据), 高度定制 (头像类型), 低延迟, 企业级	金融, 零售, 科技, 医疗	定制报价	MiniPrem: Ubuntu, 高端硬件 (RTX 4080+), P2 账户
DeepBrain AI	仅 Kiosk 明确提及	Kiosk 可本地部署, 通用方案不明确	(云) AI Studios (视频生成), AI Human (对话), 80+ 语言	金融, 媒体, 教育	(云) 分层订阅/定制报价	不明确
Soul Machines	否 (基于现有信息)	N/A (云为主)	(云) "生物 AI" 驱动的情感表达	N/A	(云) 可能定制报价	N/A
Synthesia	否 (基于现有信息)	N/A (云为主, 可能有特定组件混合部署)	(云) 140+ 语言, 海量模板, 易用性, 视频生成	培训, 营销, 企业沟通	(云) 分层订阅 (免费版-\$67+/月)	N/A

5. 部署模型对比分析: 开源 vs. 私有化商业 vs. 云 SaaS

选择合适的 AI 数字人直播解决方案部署模型, 需要权衡多个关键因素。以下将从成本、控制权、安全性、可扩展性、维护复杂度和功能更新等维度, 对开源、私有化商业和云 SaaS 三种模式进行比较。

5.1. 总拥有成本(TCO)与定价结构

- 开源方案: 软件本身通常无许可费用³¹, 但 TCO 并非为零。主要成本在于:
 - 硬件投入: 尤其是高性能 GPU(用于 AI 模型推理和渲染)⁴⁰。

- 基础设施：服务器、网络、存储的搭建和维护。
- 人力成本：需要具备专业技能的开发、集成和运维团队²⁸。
- 潜在成本：数据采集、模型训练或微调、第三方 API 调用(如某些 TTS/LLM 服务)⁴⁷。成本结构以初始资本支出(CapEx)和持续的内部运营支出(OpEx)为主。
- 私有化商业方案：成本构成复杂，通常包括：
 - 软件许可/设置费：可能涉及较高的初始费用⁸。
 - 硬件与基础设施：与开源类似，需要自行承担或按供应商规格配置⁴²。
 - 维护与支持合同：通常需要支付持续的费用以获得更新和技术支持⁸。成本结构是 CapEx 和 OpEx 的混合。对于稳定且大规模的应用，长期 TCO 可能低于 SaaS⁵⁷。
- 云 SaaS 方案：主要成本是基于订阅的运营支出(OpEx)：
 - 订阅费：按月或按年支付，通常根据功能层级、使用量(如视频分钟数、API 调用次数、并发数)或用户席位定价⁷。
 - 附加费用：可能需要为定制头像、额外功能或超出配额的使用量支付额外费用²⁴。初始成本低，费用与使用量挂钩，易于预算，但大规模使用时成本可能迅速增长。BocaLive 声称其 AI 头像直播成本可低至每小时 1 美元⁵。

5.2. 数据主权、控制权与安全影响

- 开源方案：提供最高级别的数据主权和控制权。所有数据和处理流程都在组织内部，安全责任完全由组织承担。适用于对数据安全和隐私有最高要求的场景。
- 私有化商业方案：提供高级别的数据控制权。数据存储和处理在组织内部基础设施中进行⁸。供应商可能需要有限的访问权限(例如用于支持)。安全性是其核心卖点，有助于满足合规要求⁵⁷。适用于受监管行业或对数据安全有严格要求的企业。
- 云 SaaS 方案：数据存储在供应商的云服务器上。组织对数据的直接控制较少，需依赖供应商的安全措施和合规认证(如 SOC 2, GDPR⁷)。可能存在数据主权方面的顾虑，尤其对于跨国运营或特定行业。多租户环境可能带来潜在风险(尽管供应商通常会采取隔离措施)。

5.3. 可扩展性、性能与可靠性

- 开源方案：可扩展性完全取决于组织自身构建的架构和投入的基础设施资源。性能(尤其是延迟)依赖于硬件配置和优化水平。可靠性取决于实施质量和持续维护。
- 私有化商业方案：可扩展性由组织通过增加内部资源来管理⁵⁷。性能通常针对本地部署进行优化，可能提供比公有云更低的延迟⁵⁷。UneeQ MiniPrem 方案即是为了提升本地渲染性能⁴²。可靠性依赖于内部基础设施的稳定性和供应商的支持响应。
- 云 SaaS 方案：通常具有良好的可扩展性，用户可根据需求调整订阅计划。供应商负责管理基础设施、性能和可靠性，通常提供服务水平协议(SLA)。延迟可能受用户地理位置影响。能够较好地应对流量波动。

5.4. 维护开销、更新与供应商锁定

- 开源方案：维护开销最高。需要内部团队负责基础设施管理、依赖库更新、安全补丁、

故障排查等。存在项目停止维护或引入不兼容更新的风险。核心软件无供应商锁定，但可能对特定技术栈或内部专业知识产生依赖。

- 私化商业方案：内部基础设施仍需维护。软件更新由供应商提供，部署更新可能需要协调。存在对特定供应商平台及其生态系统的锁定风险。支持合同至关重要。
- 云 SaaS 方案：维护开销最低，由供应商负责平台、基础设施和安全更新。用户可以快速开始使用。供应商锁定风险最高，迁移数据和工作流到其他平台通常很困难。用户需跟随供应商的技术路线图和功能发布节奏。

5.5. 功能可用性、创新速度与定制灵活性

- 开源方案：功能集取决于所选项目及其社区活跃度³¹。提供最大的定制灵活性（可访问和修改源代码）。创新速度不一，活跃社区可能迭代迅速，但也可能停滞。在功能完善度或企业级特性方面可能落后于商业产品。
- 私有化商业方案：通常提供针对企业需求的全面功能集⁸。定制化在平台允许范围内或通过专业服务实现⁸。创新速度由供应商决定，更新频率可能低于云 SaaS。可提供深度集成能力⁵⁵。
- 云 SaaS 方案：通常拥有最新的功能和快速的创新周期，得益于持续部署模式⁷。定制化通常限于配置选项、模板和品牌设置⁷。架构层面的深度定制灵活性较低。不同层级的订阅计划提供不同的功能集⁷。

5.6. AI 数字人直播的普遍优势与挑战

独立于部署模式，AI 数字人直播技术本身也带来一些普遍的优势和挑战：

- 优势：成本效益⁴、全天候运行²、多语言能力²、品牌一致性、可扩展性¹⁵、个性化潜力²和可访问性¹²。
- 挑战：技术门槛²、伦理问题（偏见、透明度、隐私、深度伪造风险）¹²、可能缺乏人类的情感细微度和互动深度⁷⁷、实现高度真实感的成本和复杂性、交互延迟问题³⁰、可能引发“恐怖谷”效应²⁹、对高质量训练数据的依赖⁷⁷、以及系统集成的复杂性⁵⁸。此外，向用户披露数字人是 AI 而非真人，可能会影响用户的互动方式和投入程度⁸¹。底层 AI 模型可能存在的偏见（如英语中心偏见⁸²）也需要关注。

表 3：部署模型优劣势矩阵

因素	开源方案	私有化商业方案	云 SaaS 方案
TCO/定价	优势：无许可费。劣势：高硬件/人力/维护成本，隐性成本高。	优势：长期大规模使用可能成本更优。劣势：高初始许可/设置费，硬件成本，持续支持费。	优势：低初始成本，按需付费，易预算。劣势：大规模使用成本高，持续订阅费。

数据控制/安全	优势: 最高控制权, 数据完全本地。劣势: 安全责任自负。	优势: 高控制权, 数据本地, 满足合规。劣势: 供应商可能需有限访问。	优势: 供应商负责安全/合规 (通常)。劣势: 控制权低, 数据在第三方, 主权担忧。
可扩展性/性能	优势: 潜力无限 (取决于投入)。劣势: 完全依赖自身架构和资源。	优势: 性能优化 (低延迟), 可控扩展。劣势: 扩展受内部资源限制。	优势: 易扩展, 供应商管理性能/可靠性。劣势: 延迟可能受网络影响。
维护/更新/锁定	优势: 无供应商锁定。劣势: 维护开销极高, 依赖内部专业知识, 项目风险。	优势: 供应商提供更新/支持。劣势: 仍需维护基础设施, 存在供应商锁定。	优势: 维护开销最低, 自动更新。劣势: 供应商锁定风险高, 迁移困难。
功能/定制/创新	优势: 最大定制灵活性。劣势: 功能可能不完善, 创新速度不一。	优势: 功能全面 (企业级), 可深度集成。劣势: 定制受限, 创新速度由供应商决定。	优势: 功能最新, 创新速度快。劣势: 定制灵活性低。

6. 结论与建议

6.1. 关键研究发现总结

本研究对 AI 数字人直播的开源及私有化部署解决方案进行了深入探讨, 主要发现如下:

1. 开源方案可行但挑战并存: 存在如 LiveTalking、handcrafted-persona-engine、ai-iris-avatar、TalkingHead (JS) 等多个可行的开源项目, 提供了从完整框架到特定组件库的不同选择。这些方案赋予用户最大的控制权和定制潜力, 但通常需要较高的技术门槛、显著的集成工作量和持续的维护投入。许可证 (Apache-2.0, MIT, GPL-3.0 等) 和硬件要求 (尤其是 GPU) 是选择开源方案时的关键考量因素。
2. 私有化商业方案满足特定需求: 少数商业供应商 (主要为 Ravatar 和 Uneeq) 提供私有化部署选项, 以满足企业对数据主权、安全合规和深度集成的严格要求。这些方案通常功能强大、可高度定制, 并提供专业支持, 但初始投资和长期成本可能较高。DeepBrain AI 等其他主流供应商在此领域的公开信息较少, 其私有化部署可能并非标准产品。
3. 云 **SaaS** 方案主导市场: 以 HeyGen、Synthesia、D-ID 等为代表的云 SaaS 平台凭借其易用性、快速迭代、丰富功能和较低的初始成本, 成为目前市场的主流选择。然而, 用户需要在便利性与数据控制权、长期成本和供应商锁定之间做出权衡。
4. 部署模式的选择取决于核心诉求: 三种部署模式各有优劣 (见表 3)。开源方案适合技术实力强、追求极致定制和控制的组织; 私有化商业方案适合注重安全合规、需要深度集成且预算充足的大型企业; 云 SaaS 方案则适合大多数寻求快速部署、易用性和成

本效益的用户，特别是初创企业和中小企业。

6.2. 场景化推荐

基于上述分析，针对不同场景提出以下建议：

- 初创公司/中小企业：预算和技术资源有限，建议优先考虑云 **SaaS** 方案。可以从免费或低成本的入门级计划开始⁷，快速验证 AI 数字人直播的效果。随着业务增长，再评估升级到更高级别计划或探索其他模式。
- 技术驱动型组织(具备研发能力)：若对特定功能有深度定制需求，或希望完全掌控技术栈和数据，可以评估开源方案。建议选择许可证友好(如 Apache-2.0 或 MIT)且社区活跃的项目(如 LiveTalking⁴⁴ 或 TalkingHead (JS)³¹)，并准备好投入相应的研发和运维资源。
- 大型企业/受监管行业(金融、医疗等)：数据安全、合规性和品牌控制是首要考虑因素，私有化商业方案(如 Ravatar⁸ 或 Uneeq⁵⁶)通常是更合适的选择。需要进行详细的供应商评估、技术对接和合同谈判。
- **VTubing**/个人创作者：可考虑专注于此场景的开源项目(如 handcrafted-persona-engine⁴⁹，需注意硬件和潜在许可问题)或集成云 **SaaS** 平台(如 HeyGen, D-ID)的 API 到直播工作流中。

6.3. 实施挑战与战略考量

无论选择哪种部署模式，成功实施 AI 数字人直播都需要关注以下挑战：

- 技术复杂性：集成渲染、动画、TTS、ASR、NLU 和流媒体等多个模块本身就具有挑战性，需要跨学科的技术能力²。
- 数据策略：创建高质量的定制头像或训练 AI 理解特定领域知识，需要有效的数据采集、处理和管理策略¹⁶。
- 伦理与合规：必须认真对待 AI 可能带来的偏见、隐私侵犯、透明度缺失和潜在的滥用风险，建立负责任的 AI 治理框架¹²。明确告知用户其互动对象是 AI 是建立信任的基础⁸¹。
- 用例选择与期望管理：应从明确、价值驱动的用例开始，逐步迭代。避免对当前 AI 能力有过高期望，尤其是在完全自主、富有深度情感的交互方面¹⁶。
- 系统集成：需要确保数字人系统能与现有的 CRM、电商平台、知识库、直播工具等顺畅对接⁸。

6.4. 未来展望

AI 数字人直播技术正处在快速发展阶段，未来趋势可能包括：

- 更高层次的真实感与情感表达：AI 模型将能生成更细微、更自然的表情和动作，更好地模拟人类的情感交流。
- 更强的实时交互能力：LLM 和相关技术的进步将进一步降低交互延迟，提升对话的流

畅度和上下文理解能力²⁰。

- 更广泛的部署场景：随着技术的成熟和成本的下降，AI 数字人将出现在更多交互界面，包括 XR(扩展现实) /元宇宙环境⁵⁵。
- 混合部署模式的兴起：结合云端 AI 能力和本地渲染/交互处理的混合模式(类似 UneeQ MiniPrem)可能会更普遍，以平衡性能、成本和控制。
- 伦理规范与标准的发展：行业将更加关注并制定相应的伦理规范和技术标准，以引导技术的健康发展⁵⁸。
- 与自主网络和数字孪生技术的融合：这些底层技术的进步将为 AI 数字人的能力演进提供更强大的支撑³。

总之，AI 数字人直播技术为各行各业带来了创新的机遇。组织在选择解决方案时，应充分评估自身需求、资源、风险承受能力，并结合对不同部署模式特点的理解，做出明智的战略决策。

Works cited

1. The 54th Statistical Report on China's Internet Development - CNNIC.com., accessed May 1, 2025, <https://www.cnnic.com.cn/IDR/ReportDownloads/202411/P020241101318428715781.pdf>
2. ANALYSIS OF DIGITAL MARKETING STRATEGIES FOR CROSS BORDER E-COMMERCE OF HAIR DRYER PRODUCTS - Upubscience Publisher, accessed May 1, 2025, <http://www.upubscience.com/upload/20250415164952.pdf>
3. ZTE Autonomous Networks White Paper (2025), accessed May 1, 2025, [https://www.zte.com.cn/content/dam/zte-site/res-www-zte-com-cn/mediaries/zte/service/ZTE%20Autonomous%20Networks%20White%20Paper%20%20\(2025\).pdf](https://www.zte.com.cn/content/dam/zte-site/res-www-zte-com-cn/mediaries/zte/service/ZTE%20Autonomous%20Networks%20White%20Paper%20%20(2025).pdf)
4. Betting on AI Live Streaming Sales! Has Viya 'Returned'? - Albase, accessed May 1, 2025, <https://www.aibase.com/news/504>
5. AI Live Stream: Revolutionizing Interactive Live Streaming with AI, accessed May 1, 2025, <https://www.topview.ai/make/ai-live-stream>
6. BocaLive - AI Digital Human Live Streaming Software Platform, accessed May 1, 2025, <https://www.bocalive.ai/>
7. Free AI Video Generator - Create AI Videos in 140 Languages, accessed May 1, 2025, <https://www.synthesia.io/>
8. RAVATAR: Real-Time Interactive 3D AI Avatars & Digital Humans, accessed May 1, 2025, <https://ravatar.com/>
9. BocaLive Launches AI Avatar for Live Streaming on Multi-platforms!, accessed May 1, 2025, <https://www.bocalive.ai/blog/events/bocalive-launches-ai-avatar-for-streaming-go-live-on-shopee-tiktok-lazada-without-humans-on-camera/>
10. How to Generate AI News Anchors and Videos in One Minute, accessed May 1, 2025, <https://www.vidnoz.com/ai-solutions/ai-news-generator.html>

11. DeepBrain AI - AI Avatar Banking & Financial Service - AI Studios, accessed May 1, 2025, <https://www.aistudios.com/solutions/finance>
12. (PDF) Human-Computer Interaction and Artificial Intelligence: Advancing Care Through Extended Mind Theory - ResearchGate, accessed May 1, 2025, https://www.researchgate.net/publication/386357754_Human-Computer_Interaction_and_Artificial_Intelligence_Advancing_Care_Through_Extended_Mind_Theory
13. Earnings call: JD.com posts solid Q1 growth, focuses on user experience - Investing.com, accessed May 1, 2025, <https://www.investing.com/news/stock-market-news/earnings-call-jdcom-posts-solid-q1-growth-focuses-on-user-experience-93CH-3446066>
14. HeyGen: AI-Powered Video Creation Platform Made Easy - Deepgram, accessed May 1, 2025, <https://deepgram.com/ai-apps/heygen>
15. AI Spokesperson Video Creator | 230+ AI Video Presenters, accessed May 1, 2025, <https://www.synthesia.io/tools/video-spokesperson>
16. The Rise of (Humanized) AI Avatars And Their Practical Use Cases, accessed May 1, 2025, <https://blog.tobiaszwingmann.com/p/the-rise-of-humanized-ai-avatars-and-their-practical-use-cases>
17. DeepBrain AI: Create Realistic AI Avatars for Videos in Minutes | Deepgram, accessed May 1, 2025, <https://deepgram.com/ai-apps/deepbrain-ai>
18. DeepBrain AI - About us, accessed May 1, 2025, <https://www.aistudios.com/company/about>
19. DeepBrain AI DOCS, accessed May 1, 2025, <https://docs.deepbrain.io/>
20. NVIDIA AI Blueprint for digital human for customer service. - GitHub, accessed May 1, 2025, <https://github.com/NVIDIA-AI-Blueprints/digital-human>
21. HeyGen: AI Video Generator | #1 on G2's 2025 Top 100 List, accessed May 1, 2025, <https://www.heygen.com/>
22. HeyGen Pricing | Free & Paid Plans from \$24/mo | HeyGen, accessed May 1, 2025, <https://www.heygen.com/pricing>
23. Synthesia Reviews & Product Details 2025 - Tekpon, accessed May 1, 2025, <https://tekpon.com/software/synthesia/reviews/>
24. Are you looking for an AI-powered avatar video live chat platform for ..., accessed May 1, 2025, <https://www.stromfee.cloud/post/are-you-looking-for-an-ai-powered-avatar-video-live-chat-platform-for-customer-support-personal-use>
25. D-ID Pricing Plans | Generative AI Video Platform, accessed May 1, 2025, <https://www.d-id.com/pricing/studio/>
26. API Pricing | D-ID AI Video Generator, accessed May 1, 2025, <https://www.d-id.com/pricing/api/>
27. D-ID - Features, Pricing, Pros & Cons (April 2025) - Siteefy, accessed May 1, 2025, <https://siteefy.com/ai-tools/d-id/>
28. Top Trending GitHub Projects This Week: AI Chatbots, Real-Time Digital Humans, and More! - YouTube, accessed May 1, 2025, <https://www.youtube.com/watch?v=r24IBZ0K-sU>
29. UneeQ Platform | Digital Human Creation For Enterprise Brands, accessed May 1,

- 2025, <https://www.digitalhumans.com/features/digital-human-creation>
30. Revolutionize Animation: Build a Digital Human with Large Language Models - HedgeDoc, accessed May 1, 2025, <https://monadical.com/posts/build-a-digital-human-with-large-language-models.html>
 31. met4citizen/TalkingHead: Talking Head (3D): A JavaScript ... - GitHub, accessed May 1, 2025, <https://github.com/met4citizen/TalkingHead>
 32. Best AI Video Generator | AI STUDIO, accessed May 1, 2025, <https://www.aistudios.com/>
 33. Make Your Own AI Virtual Presenter in Just 4 Steps - Mango Animate, accessed May 1, 2025, <https://school.mangoanimate.com/make-your-own-ai-virtual-presenter/>
 34. 10 Best AI Avatar Generators (April 2025) - Unite.AI, accessed May 1, 2025, <https://www.unite.ai/best-ai-avatar-generators/>
 35. Convert PowerPoint To Video With AI, accessed May 1, 2025, <https://www.automate.video/>
 36. Kedreamix/Awesome-Talking-Head-Synthesis - GitHub, accessed May 1, 2025, <https://github.com/Kedreamix/Awesome-Talking-Head-Synthesis>
 37. harlanhong/awesome-talking-head-generation - GitHub, accessed May 1, 2025, <https://github.com/harlanhong/awesome-talking-head-generation>
 38. Real-Time AI Avatar Generation | Restackio, accessed May 1, 2025, <https://www.restack.io/p/ai-avatars-answer-real-time-avatar-generation-cat-ai>
 39. Advancing Digital Interactions with Yepic AI's Real-Time Avatars, accessed May 1, 2025, <https://www.yepic.ai/blog/advancing-digital-interactions-with-yepic-ai-s-real-time-avatars>
 40. Scthe/ai-iris-avatar: Talk with AI-powered detailed 3D ... - GitHub, accessed May 1, 2025, <https://github.com/Scthe/ai-iris-avatar>
 41. metaiintw/build-an-avatar-with-ASR-TTS-Transformer-Omniverse-Audio2Face - GitHub, accessed May 1, 2025, <https://github.com/metaiintw/build-an-avatar-with-ASR-TTS-Transformer-Omniverse-Audio2Face>
 42. MiniPrem - Uneeq Docs, accessed May 1, 2025, <https://docs.uneeq.io/miniprem>
 43. LiveTalking: open source real-time interactive digital human live system, to achieve synchronous audio and video dialogues - 首席AI分享圈, accessed May 1, 2025, <https://www.aisharenet.com/en/livetalking/>
 44. GitHub - lipku/LiveTalking: Real time interactive streaming digital ..., accessed May 1, 2025, <https://github.com/lipku/LiveTalking>
 45. sausheong/persona: Talking head video AI generator - GitHub, accessed May 1, 2025, <https://github.com/sausheong/persona>
 46. antgroup/ditto-talkinghead: Ditto: Motion-Space Diffusion for Controllable Realtime Talking Head Synthesis - GitHub, accessed May 1, 2025, <https://github.com/antgroup/ditto-talkinghead>
 47. asanchezyali/talking-avatar-with-ai: This project is a digital human that can talk and listen to you. It uses OpenAI's GPT to generate responses, OpenAI's Whisper

- to transcript the audio, Eleven Labs to generate voice and Rhubarb Lip Sync to generate the lip sync. - GitHub, accessed May 1, 2025, <https://github.com/asanchezyali/talking-avatar-with-ai>
48. Best AI Text To Video Generators In 2025 - RebelLink, accessed May 1, 2025, <https://www.rebellink.com/ai-text-to-video-generators/>
 49. fagenorn/handcrafted-persona-engine: An AI-powered ... - GitHub, accessed May 1, 2025, <https://github.com/fagenorn/handcrafted-persona-engine>
 50. HeyGen Pricing & Reviews 2025 | Techjockey.com, accessed May 1, 2025, <https://www.techjockey.com/detail/heygen>
 51. How to Create AI Virtual Assistant: Product Owner's Guide 2025 - MobiDev, accessed May 1, 2025, <https://mobidev.biz/blog/ai-virtual-assistant-technology-guide>
 52. NVIDIA Omniverse ACE vs. Soul Machines Comparison - SourceForge, accessed May 1, 2025, <https://sourceforge.net/software/compare/NVIDIA-Omniverse-ACE-vs-Soul-Machines/>
 53. DeepBrain AI INC. - Microsoft Azure Marketplace, accessed May 1, 2025, https://azuremarketplace.microsoft.com/en-us/marketplace/apps/deepbrainaiinc1665180927999.deepbrain_ai?tab=Overview
 54. AvaChat - is a realtime AI chat demo with animated talking heads - it uses Large Language Models (GPT, API2D GPT4, Cluade) as text inputs to D-ID's image-to-video talking head model (via D-ID stream api) - GitHub, accessed May 1, 2025, <https://github.com/Jaykef/AvaChat>
 55. UneeQ Platform Features | Integrations and deployment, accessed May 1, 2025, <https://www.digitalhumans.com/features/integrations-and-deployment>
 56. Launch of UneeQ 2.0 sets a new standard for AI digital workforces, driving innovation in the autonomous digital human avatar industry, accessed May 1, 2025, <https://www.digitalhumans.com/blog/launch-of-uneeq-2-sets-a-new-standard-for-ai-digital-workforces-driving-innovation-in-the-autonomous-digital-human-avatar-industry>
 57. Advanced AI and Real-Time 3D Avatars in On-Premise Solutions ..., accessed May 1, 2025, <https://ravatar.com/ravatar-3d-real-time-ai-avatars-in-on-premise-solutions/>
 58. The Application of Artificial Intelligence in E-Commerce Media Operations - Atlantis Press, accessed May 1, 2025, <https://www.atlantis-press.com/article/126009893.pdf>
 59. How AI Interactive Avatars Can Improve Real-Time Customer Experience & Service, accessed May 1, 2025, <https://aiinfluencercompany.com/how-ai-interactive-avatars-can-improve-real-time-customer-experience-service/>
 60. Plexer Replicator vs. RAVATAR Comparison - SourceForge, accessed May 1, 2025, <https://sourceforge.net/software/compare/Plexer-Replicator-vs-RAVATAR/>
 61. Avatarly vs. RAVATAR Comparison - SourceForge, accessed May 1, 2025, <https://sourceforge.net/software/compare/Avatarly-vs-RAVATAR/>

62. UneeQ Kiosks | On-premise kiosk solutions for digital humans, accessed May 1, 2025, <https://www.digitalhumans.com/use-cases/kiosks>
63. 30 Best AI Tools to Watch in 2025 (Tried and Tested) - Digital Success Blog, accessed May 1, 2025, <https://www.digitalsuccess.us/blog/30-best-ai-tools-to-watch-in-2025-tried-and-tested.html>
64. Beyond the Cloud: Nvidia's Mini Supercomputer Brings AI on-Premise | PYMNTS.com, accessed May 1, 2025, <https://www.pymnts.com/news/artificial-intelligence/2025/nvidia-mini-supercomputer-project-digits-brings-ai-business-colleges-homes/>
65. Soul Machines - Okta, accessed May 1, 2025, <https://www.okta.com/en-ca/msm-new-countries-testing/customers/soul-machines/>
66. Soul Machines | Okta UK, accessed May 1, 2025, <https://www.okta.com/uk/customers/soul-machines/>
67. Soul Machines | AI Assistants, accessed May 1, 2025, <https://www.soulmachines.com/>
68. Simplify Model Deployment with Hugging Face Inference Endpoints, accessed May 1, 2025, <https://www.toolify.ai/ai-news/simplify-model-deployment-with-hugging-face-inference-endpoints-968680>
69. Compare DaveAI vs. Soul Machines in 2025 - Slashdot, accessed May 1, 2025, <https://slashdot.org/software/comparison/DaveAI-vs-Soul-Machines/>
70. Synthesia - Make, accessed May 1, 2025, <https://www.make.com/en/help/ai-in-make/synthesia>
71. Compare RepliQ vs. Synthesia in 2025 - Slashdot, accessed May 1, 2025, <https://slashdot.org/software/comparison/RepliQ-vs-Synthesia/>
72. Vertical AI Innovations: Abacus LLM in Finance & Synthesia's Video Avatars | E2112, accessed May 1, 2025, <https://podcasts.apple.com/us/podcast/vertical-ai-innovations-abacus-llm-in-finance-synthesias/id315114957?i=1000703660766&l=zh-Hant-TW>
73. Subscriptions Explained: What You Need to Know - HeyGen Help Center, accessed May 1, 2025, <https://help.heygen.com/en/articles/9204682-subscriptions-explained-what-you-need-to-know>
74. HeyGen: Personalized Videos HubSpot Integration | Connect Them Today, accessed May 1, 2025, <https://ecosystem.hubspot.com/marketplace/apps/heygen-ai-avatar-3035101>
75. HeyGen Pricing & Latest Features – Fall 2025 Updates, accessed May 1, 2025, <https://www.heygen.com/latest-drops/fall-2024>
76. Empower your sales, marketing, and customer engagement with D-ID's interactive digital human avatars - Azure Marketplace, accessed May 1, 2025, https://azuremarketplace.microsoft.com/en/marketplace/apps/deidentification1706608474824.d-id_agents
77. Teacher Professional Development - EdTech Books, accessed May 1, 2025,

- https://edtechbooks.org/ai_in_education/teacher_professional_development
78. Towards AI Literacy: 101+ - Teach Anywhere, accessed May 1, 2025,
https://teachanywhere.opened.ca/wp-content/uploads/sites/1462/2024/08/AILiteracy101_June2024F-1.pdf
79. Teacher Professional Development Teacher Professional Development for AI in Education focused on the k-12 sector, accessed May 1, 2025,
<https://edtechbooks.s3.us-west-2.amazonaws.com/pdfs/713/13343.pdf>
80. Pros and Cons of Remote Work: What You Need to Know | AI Art Generator, accessed May 1, 2025,
<https://easy-peasy.ai/ai-image-generator/images/remote-work-pros-cons-advantages-disadvantages>
81. Understanding users' motivations to engage in virtual worlds: A multipurpose model and empirical testing | Request PDF - ResearchGate, accessed May 1, 2025,
https://www.researchgate.net/publication/220496224_Understanding_users'_motivations_to_engage_in_virtual_worlds_A_multipurpose_model_and_empirical_testing
82. ACTFL 2024 Conference Report - The FLT MAG, accessed May 1, 2025,
<https://fltmag.com/actfl-2024-conference-report/>