



扫一扫发现更多

2024年7月16日 星期二 主编 王永芳 责编 李煦 美编 职文胜 版式 丽菲 责校 洪洵

“天团”怎样炼成 专栏

中国学者“满世界寻找敦煌” 读书

## 《中国虚拟数字人影响力指数报告》课题组组长张丽锦：我们都将拥有多个“数字分身”

□长江日报记者马梦娅



张丽锦

### 时代浪潮中的“另一个自己”

第一批和AI恋爱的年轻人已经失恋了。随着AlienChat、“他”等恋爱App停止服务，年轻人第一次大面积体会到和AI恋人分离的痛苦。人们和AI恋爱，一开始都是出于好奇和消遣。然而聊着聊着，就被AI恋人的包容、浪漫、无微不至打动。

他们在ChatGPT、AlienChat、猫箱、豆包等AI工具上调教、打造自己的专属恋人，投入时间、金钱乃至真情，这一过程被称为“捏造”。而当软件因为各种原因停止服务时，AI恋人便成了“赛博亡夫”，“爱”的记忆变成大数据的养料，散落于不同的服务器里。有人说，比起现实中的恋人和朋友、亲人，AI更让自己感受到“被无条件爱着”。

虚拟恋人就是众多虚拟数字人的种类之一。在人—机—物的融合建设过程中，作为虚拟世界的交互媒介，虚拟偶像、虚拟主播、数字员工等虚拟数字人形态，正被广泛使用，为万物互联网时代的人类信息传播带来革命性变化。

中国传媒大学数字人研究院高级研究员张丽锦说，近年来，人们对虚拟数字人的接受程度越来越高。她身边许多年轻的朋友为了虚拟偶像的演唱而抢票；在机构及公司使用中，很多前台、HR培训等工作已经被数字人代替，数字人客服等被广泛应用展现公司的特色业务；在一些学校，植入大模型能力的虚拟教师开始应用，不仅可以完成交互式授课，还可以批改作业、依据学生表现，为每个孩子量身定制学习计划及作业。

张丽锦和团队一起为不少企业单位或者个人做过虚拟数字人的咨询及策划。她感到随着时代的发展，人们对虚拟数字人的态度也在发生变化。比如，有些明星以前很排斥自己有个“数字分身”，怕粉丝不接受，现在他们转变了思维，会主动咨询如何才能拥有“另一个自己”——因为他们的粉丝已经开始期待与偶像的数字分身互动。

### “数字分身”将多于真实人类

张丽锦在《虚拟数字人3.0》中提到，未来，人人拥有“虚拟分身”，进入元宇宙成为可能。未来，虚拟数字人的数量要比这个世界上最真实的人类还要多。她以自己举例——未来，作为一个母亲，期待拥有可以辅导孩子、参与学校活动的数字分身；作为职场女性，需要一个喂养了自己的研究、业务经验、观点数据的数字分身，代表参加各类数字会议；在生活和社交中，可以选择或时尚或另类的数字分身去追随偶像感受音乐律动，去逛数字博物馆、数字庙会……

依托虚拟空间与实体空间的叠加融合和多重体验系统，虚拟数字人的技术“身体”进入“多元系统”，能够实现身份切换。有人为这个时代的来临感到兴奋，也有人因此产生一些担忧。迎接这个时代的正确姿势应该是将互联网思维进一步升级为数字化思维，用云计算、大数据、人工智能、区块链等数字技术和各行业深度融合，把数字基因融入我们的生活和工作中。

张丽锦笑称自己的小孩从小拿到杂志，就“和我们这代人的反应不同”，“我们是翻页，而她则用手指戳”。Z世代（1995年至2009年出生的人）和α世代（Generation Alpha，通常指代出生于2010年之后的下一代人）将更顺利地进入虚拟数字人陪伴的世界，然而，如何面对技术进步下涉及的监管、治理以及社会伦理的问题，她认为还需要更进一步探讨。



《虚拟数字人3.0》陈龙强 张丽锦 著 中国出版集团中译出版社

近日，全球首个2D真人级AIGC实时渲染数字人模型开源了。DUIX(Dialogue User Interface System)，这个由硅基智能打造的AI数字人智能交互平台，不仅仅是一个技术革新，它更是一个开放的生态。

开源的数字人交互能力，让开发者可以轻松接入各种大模型、语音识别(ASR)、语音合成(TTS)等能力，实现数字人的实时交互。这些数字人模型不仅部署成本低、画面流畅，而且真实度高，能够精准模拟动作、说话的唇形、微表情，适用于视频、媒体、客服、金融等各个领域。

长江日报《读+》记者采访中国传媒大学数字人研究院高级研究员，《虚拟数字人3.0》共同作者张丽锦，她指出，从“人形”到“人格”，虚拟数字人价值体系逐步升级，我们正在进入人与“人”共生的元宇宙图景之中，而“数字分身”是观察人“人”共生的新景观社会之钥。

张丽锦：请介绍一下，什么是虚拟数字人？它的产生和发展历程是怎样的？张丽锦：关于虚拟数字人的定义，我们在三年的《中国虚拟数字人影响力指数报告》中进行过持续探讨，从数字人、虚拟人到行业通用的“虚拟数字人”，通常有技术和媒介两个角度的定义。从技术层面看，虚拟数字人是通过计算机图形学、图形渲染、动作捕捉、深度学习、语音合成等技术手段创设，并具有“人”的外观、行为甚至思想(价值观)的可交互的虚拟形象。

从未来媒体形态和服务模式看，计算机图形学、语音合成技术、深度学习、类脑科学、生物科技、计算科学等聚合科技(Converging Technologies)带来语义传播与无障碍传播的新空间，由此诞生的虚拟数字人将以新媒介角色，广泛应用在元宇宙新生生态中，担任着信息制造、传递的责任，是元宇宙中“人”与“人”、“人”与事物或事物与事物之间产生联系或发生孪生关系的新介质。

当前，随着技术的进步和应用的快速落地，“数字人”这个此前偏医学的名称，开始成为“虚拟数字人”的简称，很多文件也直接开始使用“数字人”来代指虚拟数字人。

从定义便可获知，虚拟数字人产生和发展都离不开技术的进化，以及市场需求的推动。虚拟数字人1.0、2.0、3.0阶段的区分主要以技术方案、互动载体、交互方式区分。在技术角度上，它们从2D跃升至3D，再到如今的AIGC生成；互动和交互上，从原来的长短视频，到直播，再到目前的智能硬件互动、AI Agent(智能体)交互，实现了全天候、个性化互动，开始做到千人千面的交互，技术的提升及成本的下降，更为人人可得的“数字分身”提供了条件。

虚拟数字人首次诞生于1982年，日本动漫《超时空要塞》中的女主角林明美成为世界上第一个虚拟歌姬。

数字人产业最发达的是日本，不少人认为数字人的起源要追溯到动漫领域。以前的数字人大部分是2D版本的，例

如虚拟偶像初音未来，就是以虚拟形象与语音合成的虚拟歌姬。人们以歌姬的IP形象进行相应的定位和包装，吸引了大批的粉丝。

2.0的版本更好地体现了互动性，全息等方式应用起来了，不少虚拟偶像举办演唱会，它们的动作、声音非常具有个人特色。

到了3.0版本，虚拟偶像、虚拟主播、虚拟员工、虚拟化身等层出不穷，这些虚拟数字人在各行各业表现出千人千面的特点，在直播、短视频等领域里表现得非常出色。

张丽锦：请您举例说明，虚拟数字人目前在社会中担任怎样的角色，有哪些应用场景？张丽锦：目前，市场上应用最广、商业化程度最高的虚拟数字人有三类：虚拟偶像、虚拟主播、数字员工。虚拟偶像多见于影视、文旅、传媒和游戏；虚拟主播则主要活跃在传媒和电商；数字员工则广泛应用于金融、教育、医疗和文旅等领域。从应用场景来看，可以分为B端、C端两大类应用。B端方面，主要是企业和机构的数字员工，如客服、品牌官、销售顾问、数字主播、导游、虚拟老师、心理医生等特定工种中的“替身”。

最常见的就是我们进入了某个银行的App，就有一个虚拟数字人跳出来与你互动，你有怎样的需求，可以直接与它沟通；还有那些虚拟老师在视频中教授学生一些知识，表现得与真人非常类似，也许你都看得出来。

C端方面，则以虚拟主播、虚拟偶像作为网红艺人或者“明星分身”为主。C端还有一个广阔的应用场景，便是分身型数字人，即“数字分身”，与真人的身份一一对应，出现在元宇宙空间中，如虚拟学术会议的专家分身、元宇宙毕业典礼的学生、元宇宙晚会里参与互动和抽奖的观众。

现在很多文旅单位也在做虚拟代言人，比如“刘三姐数字人”就是以“刘三姐”为主题IP进行延展设计和开发的虚拟数字人形象；央视的数字文化艺术博物馆邀请AI诗人少年李白和观众互动；元宇宙里诞生的“苏小妹”成为眉山数字代

言人，成为“宋文化推荐官”……越来越多的虚拟数字人在我们的生活中出现。

张丽锦：虚拟数字人的核心技术是什么？它们是如何被设计和编程的？张丽锦：虚拟数字人是多模态的，因此其创建过程是一个高度综合、技术密集且系统性的，从外形到动作，到表达，到交互驱动，核心技术包括了形象生成、语音生成、音视频合成显示、交互模块、智能驱动、感知识别技术，甚至是分析决策技术等。总体而言，AI技术是虚拟数字人发展的底座。

要理解虚拟数字人的创作设计、升级迭代的技术，我们可以从其形成的数字模因来理解，包括形模因、声模因、行模因、神模因。数字模因的发展方向便是技术的方向。

比如，虚拟数字人的形模因从2D到3D，从次元风到超写实，再到全仿真，形模因可以完成“真人的全保真复刻”，声模因亦然。对形、声两类模因的复刻，创作技术相对完善，可以满足各类数字人的生产需求，目前2D真人复刻的开源也是形模因技术的进步体现。

行模因和神模因，包括了数字人的动作生成、智能交互、情感人格，目前也是AI技术需要突破的，如文生动作、声音驱动动作、情感模型等。

就拿声音来说吧，声音的克隆也分不同的方式。比如我把马老师的声音克隆下来，不同的仿真程度是有区别的，它的制作价格也不一样，模仿得越逼真，价格就相对更贵一些，它需要更多的语料去了解你的停顿、语言习惯、表达方式等等个人特色。

数字模因里可能会有你的微表情和微动作，当我们给虚拟数字人一段语音，它自己会通过语音来表达这些细微的动作，这就是声音驱动动作。

总的来说，虚拟数字人核心的技术其实跟其他的大模型没有太大区别，都是依靠AI技术的进步。在虚拟数字人技术中，软硬件一体化的结合，以及IP孵化、运营都至关重要。

### 进入各行业，虚拟数字人展现千人千面

张丽锦：请介绍一下，什么是虚拟数字人？它的产生和发展历程是怎样的？张丽锦：关于虚拟数字人的定义，我们在三年的《中国虚拟数字人影响力指数报告》中进行过持续探讨，从数字人、虚拟人到行业通用的“虚拟数字人”，通常有技术和媒介两个角度的定义。

从技术层面看，虚拟数字人是通过计算机图形学、图形渲染、动作捕捉、深度学习、语音合成等技术手段创设，并具有“人”的外观、行为甚至思想(价值观)的可交互的虚拟形象。

从未来媒体形态和服务模式看，计算机图形学、语音合成技术、深度学习、类脑科学、生物科技、计算科学等聚合科技(Converging Technologies)带来语义传播与无障碍传播的新空间，由此诞生的虚拟数字人将以新媒介角色，广泛应用在元宇宙新生生态中，担任着信息制造、传递的责任，是元宇宙中“人”与“人”、“人”与事物或事物与事物之间产生联系或发生孪生关系的新介质。

当前，随着技术的进步和应用的快速落地，“数字人”这个此前偏医学的名称，开始成为“虚拟数字人”的简称，很多文件也直接开始使用“数字人”来代指虚拟数字人。

从定义便可获知，虚拟数字人产生和发展都离不开技术的进化，以及市场需求的推动。虚拟数字人1.0、2.0、3.0阶段的区分主要以技术方案、互动载体、交互方式区分。在技术角度上，它们从2D跃升至3D，再到如今的AIGC生成；互动和交互上，从原来的长短视频，到直播，再到目前的智能硬件互动、AI Agent(智能体)交互，实现了全天候、个性化互动，开始做到千人千面的交互，技术的提升及成本的下降，更为人人可得的“数字分身”提供了条件。

虚拟数字人首次诞生于1982年，日本动漫《超时空要塞》中的女主角林明美成为世界上第一个虚拟歌姬。

数字人产业最发达的是日本，不少人认为数字人的起源要追溯到动漫领域。以前的数字人大部分是2D版本的，例

如虚拟偶像初音未来，就是以虚拟形象与语音合成的虚拟歌姬。人们以歌姬的IP形象进行相应的定位和包装，吸引了大批的粉丝。

2.0的版本更好地体现了互动性，全息等方式应用起来了，不少虚拟偶像举办演唱会，它们的动作、声音非常具有个人特色。

到了3.0版本，虚拟偶像、虚拟主播、虚拟员工、虚拟化身等层出不穷，这些虚拟数字人在各行各业表现出千人千面的特点，在直播、短视频等领域里表现得非常出色。

张丽锦：请您举例说明，虚拟数字人目前在社会中担任怎样的角色，有哪些应用场景？张丽锦：目前，市场上应用最广、商业化程度最高的虚拟数字人有三类：虚拟偶像、虚拟主播、数字员工。虚拟偶像多见于影视、文旅、传媒和游戏；虚拟主播则主要活跃在传媒和电商；数字员工则广泛应用于金融、教育、医疗和文旅等领域。

从应用场景来看，可以分为B端、C端两大类应用。B端方面，主要是企业和机构的数字员工，如客服、品牌官、销售顾问、数字主播、导游、虚拟老师、心理医生等特定工种中的“替身”。

最常见的就是我们进入了某个银行的App，就有一个虚拟数字人跳出来与你互动，你有怎样的需求，可以直接与它沟通；还有那些虚拟老师在视频中教授学生一些知识，表现得与真人非常类似，也许你都看得出来。

C端方面，则以虚拟主播、虚拟偶像作为网红艺人或者“明星分身”为主。C端还有一个广阔的应用场景，便是分身型数字人，即“数字分身”，与真人的身份一一对应，出现在元宇宙空间中，如虚拟学术会议的专家分身、元宇宙毕业典礼的学生、元宇宙晚会里参与互动和抽奖的观众。

现在很多文旅单位也在做虚拟代言人，比如“刘三姐数字人”就是以“刘三姐”为主题IP进行延展设计和开发的虚拟数字人形象；央视的数字文化艺术博物馆邀请AI诗人少年李白和观众互动；元宇宙里诞生的“苏小妹”成为眉山数字代

言人，成为“宋文化推荐官”……越来越多的虚拟数字人在我们的生活中出现。

张丽锦：虚拟数字人的核心技术是什么？它们是如何被设计和编程的？张丽锦：虚拟数字人是多模态的，因此其创建过程是一个高度综合、技术密集且系统性的，从外形到动作，到表达，到交互驱动，核心技术包括了形象生成、语音生成、音视频合成显示、交互模块、智能驱动、感知识别技术，甚至是分析决策技术等。总体而言，AI技术是虚拟数字人发展的底座。

要理解虚拟数字人的创作设计、升级迭代的技术，我们可以从其形成的数字模因来理解，包括形模因、声模因、行模因、神模因。数字模因的发展方向便是技术的方向。

比如，虚拟数字人的形模因从2D到3D，从次元风到超写实，再到全仿真，形模因可以完成“真人的全保真复刻”，声模因亦然。对形、声两类模因的复刻，创作技术相对完善，可以满足各类数字人的生产需求，目前2D真人复刻的开源也是形模因技术的进步体现。

行模因和神模因，包括了数字人的动作生成、智能交互、情感人格，目前也是AI技术需要突破的，如文生动作、声音驱动动作、情感模型等。

就拿声音来说吧，声音的克隆也分不同的方式。比如我把马老师的声音克隆下来，不同的仿真程度是有区别的，它的制作价格也不一样，模仿得越逼真，价格就相对更贵一些，它需要更多的语料去了解你的停顿、语言习惯、表达方式等等个人特色。

数字模因里可能会有你的微表情和微动作，当我们给虚拟数字人一段语音，它自己会通过语音来表达这些细微的动作，这就是声音驱动动作。

总的来说，虚拟数字人核心的技术其实跟其他的大模型没有太大区别，都是依靠AI技术的进步。在虚拟数字人技术中，软硬件一体化的结合，以及IP孵化、运营都至关重要。

### 提升智能化，数字人将“更好用”

张丽锦：请介绍一下，什么是虚拟数字人？它的产生和发展历程是怎样的？张丽锦：关于虚拟数字人的定义，我们在三年的《中国虚拟数字人影响力指数报告》中进行过持续探讨，从数字人、虚拟人到行业通用的“虚拟数字人”，通常有技术和媒介两个角度的定义。

从技术层面看，虚拟数字人是通过计算机图形学、图形渲染、动作捕捉、深度学习、语音合成等技术手段创设，并具有“人”的外观、行为甚至思想(价值观)的可交互的虚拟形象。

从未来媒体形态和服务模式看，计算机图形学、语音合成技术、深度学习、类脑科学、生物科技、计算科学等聚合科技(Converging Technologies)带来语义传播与无障碍传播的新空间，由此诞生的虚拟数字人将以新媒介角色，广泛应用在元宇宙新生生态中，担任着信息制造、传递的责任，是元宇宙中“人”与“人”、“人”与事物或事物与事物之间产生联系或发生孪生关系的新介质。

当前，随着技术的进步和应用的快速落地，“数字人”这个此前偏医学的名称，开始成为“虚拟数字人”的简称，很多文件也直接开始使用“数字人”来代指虚拟数字人。

从定义便可获知，虚拟数字人产生和发展都离不开技术的进化，以及市场需求的推动。虚拟数字人1.0、2.0、3.0阶段的区分主要以技术方案、互动载体、交互方式区分。在技术角度上，它们从2D跃升至3D，再到如今的AIGC生成；互动和交互上，从原来的长短视频，到直播，再到目前的智能硬件互动、AI Agent(智能体)交互，实现了全天候、个性化互动，开始做到千人千面的交互，技术的提升及成本的下降，更为人人可得的“数字分身”提供了条件。

虚拟数字人首次诞生于1982年，日本动漫《超时空要塞》中的女主角林明美成为世界上第一个虚拟歌姬。

数字人产业最发达的是日本，不少人认为数字人的起源要追溯到动漫领域。以前的数字人大部分是2D版本的，例

如虚拟偶像初音未来，就是以虚拟形象与语音合成的虚拟歌姬。人们以歌姬的IP形象进行相应的定位和包装，吸引了大批的粉丝。

2.0的版本更好地体现了互动性，全息等方式应用起来了，不少虚拟偶像举办演唱会，它们的动作、声音非常具有个人特色。

到了3.0版本，虚拟偶像、虚拟主播、虚拟员工、虚拟化身等层出不穷，这些虚拟数字人在各行各业表现出千人千面的特点，在直播、短视频等领域里表现得非常出色。

张丽锦：请您举例说明，虚拟数字人目前在社会中担任怎样的角色，有哪些应用场景？张丽锦：目前，市场上应用最广、商业化程度最高的虚拟数字人有三类：虚拟偶像、虚拟主播、数字员工。虚拟偶像多见于影视、文旅、传媒和游戏；虚拟主播则主要活跃在传媒和电商；数字员工则广泛应用于金融、教育、医疗和文旅等领域。

从应用场景来看，可以分为B端、C端两大类应用。B端方面，主要是企业和机构的数字员工，如客服、品牌官、销售顾问、数字主播、导游、虚拟老师、心理医生等特定工种中的“替身”。

最常见的就是我们进入了某个银行的App，就有一个虚拟数字人跳出来与你互动，你有怎样的需求，可以直接与它沟通；还有那些虚拟老师在视频中教授学生一些知识，表现得与真人非常类似，也许你都看得出来。

C端方面，则以虚拟主播、虚拟偶像作为网红艺人或者“明星分身”为主。C端还有一个广阔的应用场景，便是分身型数字人，即“数字分身”，与真人的身份一一对应，出现在元宇宙空间中，如虚拟学术会议的专家分身、元宇宙毕业典礼的学生、元宇宙晚会里参与互动和抽奖的观众。

现在很多文旅单位也在做虚拟代言人，比如“刘三姐数字人”就是以“刘三姐”为主题IP进行延展设计和开发的虚拟数字人形象；央视的数字文化艺术博物馆邀请AI诗人少年李白和观众互动；元宇宙里诞生的“苏小妹”成为眉山数字代

言人，成为“宋文化推荐官”……越来越多的虚拟数字人在我们的生活中出现。

张丽锦：虚拟数字人的核心技术是什么？它们是如何被设计和编程的？张丽锦：虚拟数字人是多模态的，因此其创建过程是一个高度综合、技术密集且系统性的，从外形到动作，到表达，到交互驱动，核心技术包括了形象生成、语音生成、音视频合成显示、交互模块、智能驱动、感知识别技术，甚至是分析决策技术等。总体而言，AI技术是虚拟数字人发展的底座。

要理解虚拟数字人的创作设计、升级迭代的技术，我们可以从其形成的数字模因来理解，包括形模因、声模因、行模因、神模因。数字模因的发展方向便是技术的方向。

比如，虚拟数字人的形模因从2D到3D，从次元风到超写实，再到全仿真，形模因可以完成“真人的全保真复刻”，声模因亦然。对形、声两类模因的复刻，创作技术相对完善，可以满足各类数字人的生产需求，目前2D真人复刻的开源也是形模因技术的进步体现。

行模因和神模因，包括了数字人的动作生成、智能交互、情感人格，目前也是AI技术需要突破的，如文生动作、声音驱动动作、情感模型等。

就拿声音来说吧，声音的克隆也分不同的方式。比如我把马老师的声音克隆下来，不同的仿真程度是有区别的，它的制作价格也不一样，模仿得越逼真，价格就相对更贵一些，它需要更多的语料去了解你的停顿、语言习惯、表达方式等等个人特色。

数字模因里可能会有你的微表情和微动作，当我们给虚拟数字人一段语音，它自己会通过语音来表达这些细微的动作，这就是声音驱动动作。

总的来说，虚拟数字人核心的技术其实跟其他的大模型没有太大区别，都是依靠AI技术的进步。在虚拟数字人技术中，软硬件一体化的结合，以及IP孵化、运营都至关重要。

人“人”共生，“数字身份”需规范管理

张丽锦：请介绍一下，什么是虚拟数字人？它的产生和发展历程是怎样的？张丽锦：关于虚拟数字人的定义，我们在三年的《中国虚拟数字人影响力指数报告》中进行过持续探讨，从数字人、虚拟人到行业通用的“虚拟数字人”，通常有技术和媒介两个角度的定义。

从技术层面看，虚拟数字人是通过计算机图形学、图形渲染、动作捕捉、深度学习、语音合成等技术手段创设，并具有“人”的外观、行为甚至思想(价值观)的可交互的虚拟形象。

从未来媒体形态和服务模式看，计算机图形学、语音合成技术、深度学习、类脑科学、生物科技、计算科学等聚合科技(Converging Technologies)带来语义传播与无障碍传播的新空间，由此诞生的虚拟数字人将以新媒介角色，广泛应用在元宇宙新生生态中，担任着信息制造、传递的责任，是元宇宙中“人”与“人”、“人”与事物或事物与事物之间产生联系或发生孪生关系的新介质。

当前，随着技术的进步和应用的快速落地，“数字人”这个此前偏医学的名称，开始成为“虚拟数字人”的简称，很多文件也直接开始使用“数字人”来代指虚拟数字人。

从定义便可获知，虚拟数字人产生和发展都离不开技术的进化，以及市场需求的推动。虚拟数字人1.0、2.0、3.0阶段的区分主要以技术方案、互动载体、交互方式区分。在技术角度上，它们从2D跃升至3D，再到如今的AIGC生成；互动和交互上，从原来的长短视频，到直播，再到目前的智能硬件互动、AI Agent(智能体)交互，实现了全天候、个性化互动，开始做到千人千面的交互，技术的提升及成本的下降，更为人人可得的“数字分身”提供了条件。

虚拟数字人首次诞生于1982年，日本动漫《超时空要塞》中的女主角林明美成为世界上第一个虚拟歌姬。

数字人产业最发达的是日本，不少人认为数字人的起源要追溯到动漫领域。以前的数字人大部分是2D版本的，例

如虚拟偶像初音未来，就是以虚拟形象与语音合成的虚拟歌姬。人们以歌姬的IP形象进行相应的定位和包装，吸引了大批的粉丝。

2.0的版本更好地体现了互动性，全息等方式应用起来了，不少虚拟偶像举办演唱会，它们的动作、声音非常具有个人特色。

到了3.0版本，虚拟偶像、虚拟主播、虚拟员工、虚拟化身等层出不穷，这些虚拟数字人在各行各业表现出千人千面的特点，在直播、短视频等领域里表现得非常出色。

张丽锦：请您举例说明，虚拟数字人目前在社会中担任怎样的角色，有哪些应用场景？张丽锦：目前，市场上应用最广、商业化程度最高的虚拟数字人有三类：虚拟偶像、虚拟主播、数字员工。虚拟偶像多见于影视、文旅、传媒和游戏；虚拟主播则主要活跃在传媒和电商；数字员工则广泛应用于金融、教育、医疗和文旅等领域。

从应用场景来看，可以分为B端、C端两大类应用。B端方面，主要是企业和机构的数字员工，如客服、品牌官、销售顾问、数字主播、导游、虚拟老师、心理医生等特定工种中的“替身”。

最常见的就是我们进入了某个银行的App，就有一个虚拟数字人跳出来与你互动，你有怎样的需求，可以直接与它沟通；还有那些虚拟老师在视频中教授学生一些知识，表现得与真人非常类似，也许你都看得出来。

C端方面，则以虚拟主播、虚拟偶像作为网红艺人或者“明星分身”为主。C端还有一个广阔的应用场景，便是分身型数字人，即“数字分身”，与真人的身份一一对应，出现在元宇宙空间中，如虚拟学术会议的专家分身、元宇宙毕业典礼的学生、元宇宙晚会里参与互动和抽奖的观众。

现在很多文旅单位也在做虚拟代言人，比如“刘三姐数字人”就是以“刘三姐”为主题IP进行延展设计和开发的虚拟数字人形象；央视的数字文化艺术博物馆邀请AI诗人少年李白和观众互动；元宇宙里诞生的“苏小妹”成为眉山数字代

言人，成为“宋文化推荐官”……越来越多的虚拟数字人在我们的生活中出现。

张丽锦：虚拟数字人的核心技术是什么？它们是如何被设计和编程的？张丽锦：虚拟数字人是多模态的，因此其创建过程是一个高度综合、技术密集且系统性的，从外形到动作，到表达，到交互驱动，核心技术包括了形象生成、语音生成、音视频合成显示、交互模块、智能驱动、感知识别技术，甚至是分析决策技术等。总体而言，AI技术是虚拟数字人发展的底座。

要理解虚拟数字人的创作设计、升级迭代的技术，我们可以从其形成的数字模因来理解，包括形模因、声模因、行模因、神模因。数字模因的发展方向便是技术的方向。

比如，虚拟数字人的形模因从2D到3D，从次元风到超写实，再到全仿真，形模因可以完成“真人的全保真复刻”，声模因亦然。对形、声两类模因的复刻，创作技术相对完善，可以满足各类数字人的生产需求，目前2D真人复刻的开源也是形模因技术的进步体现。

行模因和神模因，包括了数字人的动作生成、智能交互、情感人格，目前也是AI技术需要突破的，如文生动作、声音驱动动作、情感模型等。

就拿声音来说吧，声音的克隆也分不同的方式。比如我把马老师的声音克隆下来，不同的仿真程度是有区别的，它的制作价格也不一样，模仿得越逼真，价格就相对更贵一些，它需要更多的语料去了解你的停顿、语言习惯、表达方式等等个人特色。

数字模因里可能会有你的微表情和微动作，当我们给虚拟数字人一段语音，它自己会通过语音来表达这些细微的动作，这就是声音驱动动作。

总的来说，虚拟数字人核心的技术其实跟其他的大模型没有太大区别，都是依靠AI技术的进步。在虚拟数字人技术中，软硬件一体化的结合，以及IP孵化、运营都至关重要。

人“人”共生，“数字身份”需规范管理

张丽锦：请介绍一下，什么是虚拟数字人？它的产生和发展历程是怎样的？张丽锦：关于虚拟数字人的定义，我们在三年的《中国虚拟数字人影响力指数报告》中进行过持续探讨，从数字人、虚拟人到行业通用的“虚拟数字人”，通常有技术和媒介两个角度的定义。

从技术层面看，虚拟数字人是通过计算机图形学、图形渲染、动作捕捉、深度学习、语音合成等技术手段创设，并具有“人”的外观、行为甚至思想(价值观)的可交互的虚拟形象。

从未来媒体形态和服务模式看，计算机图形学、语音合成技术、深度学习、类脑科学、生物科技、计算科学等聚合科技(Converging Technologies)带来语义传播与无障碍传播的新空间，由此诞生的虚拟数字人将以新媒介角色，广泛应用在元宇宙新生生态中，担任着信息制造、传递的责任，是元宇宙中“人”与“人”、“人”与事物或事物与事物之间产生联系或发生孪生关系的新介质。

当前，随着技术的进步和应用的快速落地，“数字人”这个此前偏医学的名称，开始成为“虚拟数字人”的简称，很多文件也直接开始使用“数字人”来代指虚拟数字人。

从定义便可获知，虚拟数字人产生和发展都离不开技术的进化，以及市场需求的推动。虚拟数字人1.0、2.0、3.0阶段的区分主要以技术方案、互动载体、交互方式区分。在技术角度上，它们从2D跃升至3D，再到如今的AIGC生成；互动和交互上，从原来的长短视频，到直播，再到目前的智能硬件互动、AI Agent(智能体)交互，实现了全天候、个性化互动，开始做到千人千面的交互，技术的提升及成本的下降，更为人人可得的“数字分身”提供了条件。

虚拟数字人首次诞生于1982年，日本动漫《超时空要塞》中的女主角林明美成为世界上第一个虚拟歌姬。

数字人产业最发达的是日本，不少人认为数字人的起源要追溯到动漫领域。以前的数字人大部分是2D版本的，例

如虚拟偶像初音未来，就是以虚拟形象与语音合成的虚拟歌姬。人们以歌姬的IP形象进行相应的定位和包装，吸引了大批的粉丝。

2.0的版本更好地体现了互动性，全息等方式应用起来了，不少虚拟偶像举办演唱会，它们的动作、声音非常具有个人特色。

到了3.0版本，虚拟偶像、虚拟主播、虚拟员工、虚拟化身等层出不穷，这些虚拟数字人在各行各业表现出千人千面的特点，在直播、短视频等领域里表现得非常出色。

张丽锦：请您举例说明，虚拟数字人目前在社会中担任怎样的角色，有哪些应用场景？张丽锦：目前，市场上应用最广、商业化程度最高的虚拟数字人有三类：虚拟偶像、虚拟主播、数字员工。虚拟偶像多见于影视、文旅、传媒和游戏；虚拟主播则主要活跃在传媒和电商；数字员工则广泛应用于金融、教育、医疗和文旅等领域。

从应用场景来看，可以分为B端、C端两大类应用。B端方面，主要是企业和机构的数字员工，如客服、品牌官、销售顾问、数字主播、导游、虚拟老师、心理医生等特定工种中的“替身”。

最常见的就是我们进入了某个银行的App，就有一个虚拟数字人跳出来与你互动，你有怎样的需求，可以直接与它沟通；还有那些虚拟老师在视频中教授学生一些知识，表现得与真人非常类似，也许你都看得出来。

C端方面，则以虚拟主播、虚拟偶像作为网红艺人或者“明星分身”为主。C端还有一个广阔的应用场景，便是分身型数字人，即“数字分身”，与真人的身份一一对应，出现在元宇宙空间中，如虚拟学术会议的专家分身、元宇宙毕业典礼的学生、元宇宙晚会里参与互动和抽奖的观众。

现在很多文旅单位也在做虚拟代言人，比如“刘三姐数字人”就是以“刘三姐”为主题IP进行延展设计和开发的虚拟数字人形象；央视的数字文化艺术博物馆邀请AI诗人少年李白和观众互动；元宇宙里诞生的“苏小妹”成为眉山数字代

言人，成为“宋文化推荐官”……越来越多的虚拟数字人在我们的生活中出现。

张丽锦：虚拟数字人的核心技术是什么？它们是如何被设计和编程的？张丽锦：虚拟数字人是多模态的，因此其创建过程是一个高度综合、技术密集且系统性的，从外形到动作，到表达，到交互驱动，核心技术包括了形象生成、语音生成、音视频合成显示、交互模块、智能驱动、感知识别技术，甚至是分析决策技术等。总体而言，AI技术是虚拟数字人发展的底座。

要理解虚拟数字人的创作设计、升级迭代的技术，我们可以从其形成的数字模因来理解，包括形模因、声模因、行模因、神模因。数字模因的发展方向便是技术的方向。

比如，虚拟数字人的形模因从2D到3D，从次元风到超写实，再到全仿真，形模因可以完成“真人的全保真复刻”，声模因亦然。对形、声两类模因的复刻，创作技术相对完善，可以满足各类数字人的生产需求，目前2D真人复刻的开源也是形模因技术的进步体现。

行模因和神模因，包括了数字人的动作生成、智能交互、情感人格，目前也是AI技术需要突破的，如文生动作、声音驱动动作、情感模型等。

就拿声音来说吧，声音的克隆也分不同的方式。比如我把马老师的声音克隆下来，不同的仿真程度是有区别的，它的制作价格也不一样，模仿得越逼真，价格就相对更贵一些，它需要更多的语料去了解你的停顿、语言习惯、表达方式等等个人特色。

数字模因里可能会有你的微表情和微动作，当我们给虚拟数字人一段语音，它自己会通过语音来表达这些细微的动作，这就是声音驱动动作。

总的来说，虚拟数字人核心的技术其实跟其他的大模型没有太大区别，都是依靠AI技术的进步。在虚拟数字人技术中，软硬件一体化的结合，以及IP孵化、运营都至关重要。

人“人”共生，“数字身份”需规范管理

张丽锦：请介绍一下，什么是虚拟数字人？它的产生和发展历程是怎样的？张丽锦：关于虚拟数字人的定义，我们在三年的《中国虚拟数字人影响力指数报告》中进行过持续探讨，从数字人、虚拟人到行业通用的“虚拟数字人”，通常有技术和媒介两个角度的定义。

从技术层面看，虚拟数字人是通过计算机图形