

AI + 游戏 产业变革研究报告

The AI Revolution in Gaming

量子位智库出品 2025.07

量子位 insights



目录 CONTENTS

1

游戏创作的AI需求

P1-P10

2

游戏创作的AI革新

P11-P42

3

AI改变的游戏体验

P43-P51

4

游戏产业的AI变革

P52-P59



Part 1

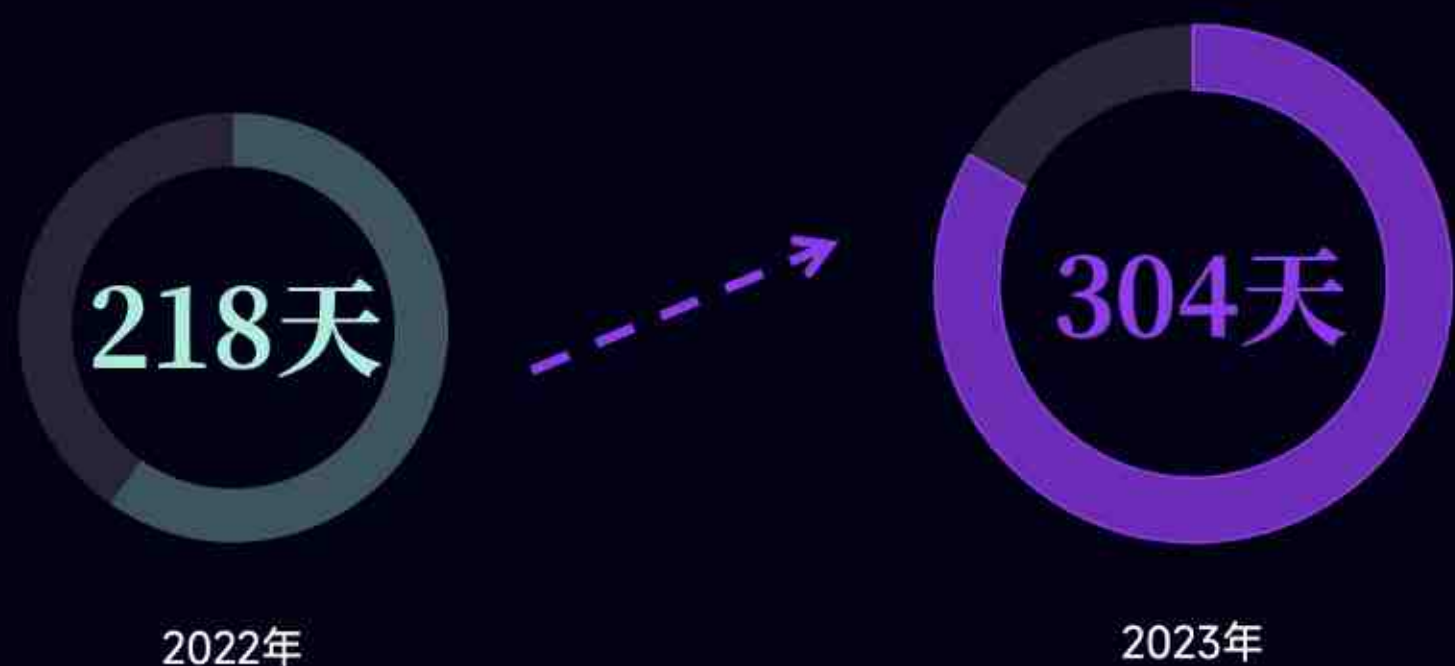
游戏创作的AI需求

半个世纪以来，我们玩到的游戏愈发精良，体验愈发丰富，制作成本却在持续膨胀：游戏上线所需的时间被不断拉长、营销费用增长到与研发费用相似……为了应对这一情况，并利用有限资源创作出更成功的游戏，开发者们对AI抱有积极态度。



游戏项目投入巨大： 时间长、成本高

近十年来，玩家的游戏体验不断升级，我们所能玩到的游戏不仅在玩法、画面和引擎上有了提升，体量也在不断膨胀。开发游戏早已从独立「手工业」迈入工业化阶段，上线一款游戏的平均时间不断拉长，投入成本不断攀升。



图：上线一款游戏的平均时间从218天攀升到304天。(2023年, Unity)



图：经典游戏作品的开发成本与周期；前期投入巨大，还要承受几度跳票和市场反馈不佳的风险。

研发费用和营销费用的膨胀

2024年，中国仅上市游戏企业的累积研发投入就突破了1000亿元，市值前十的企业能够达到10亿元以上的投入和10%以上的研发占营业收入占比。

以同公司的现象级作品为例，CD Project Group在2015年向《巫师3:狂猎》投入约8000万美元的研发费用，而在2020年为《赛博朋克2077》投入约1.73亿美元。开发者们从不吝啬对核心作品的投入。

同时，为在激烈的市场竞争中胜出，游戏开发商还要承担庞大的营销费用，有时甚至与游戏本体的研发费用持平。

例如《赛博朋克2077》的整体研发成本约为3.16亿美元，研发与营销费用占比约55:45；《GTA5》也有约1.28亿美元的营销费用。

热门游戏产品的研发费用



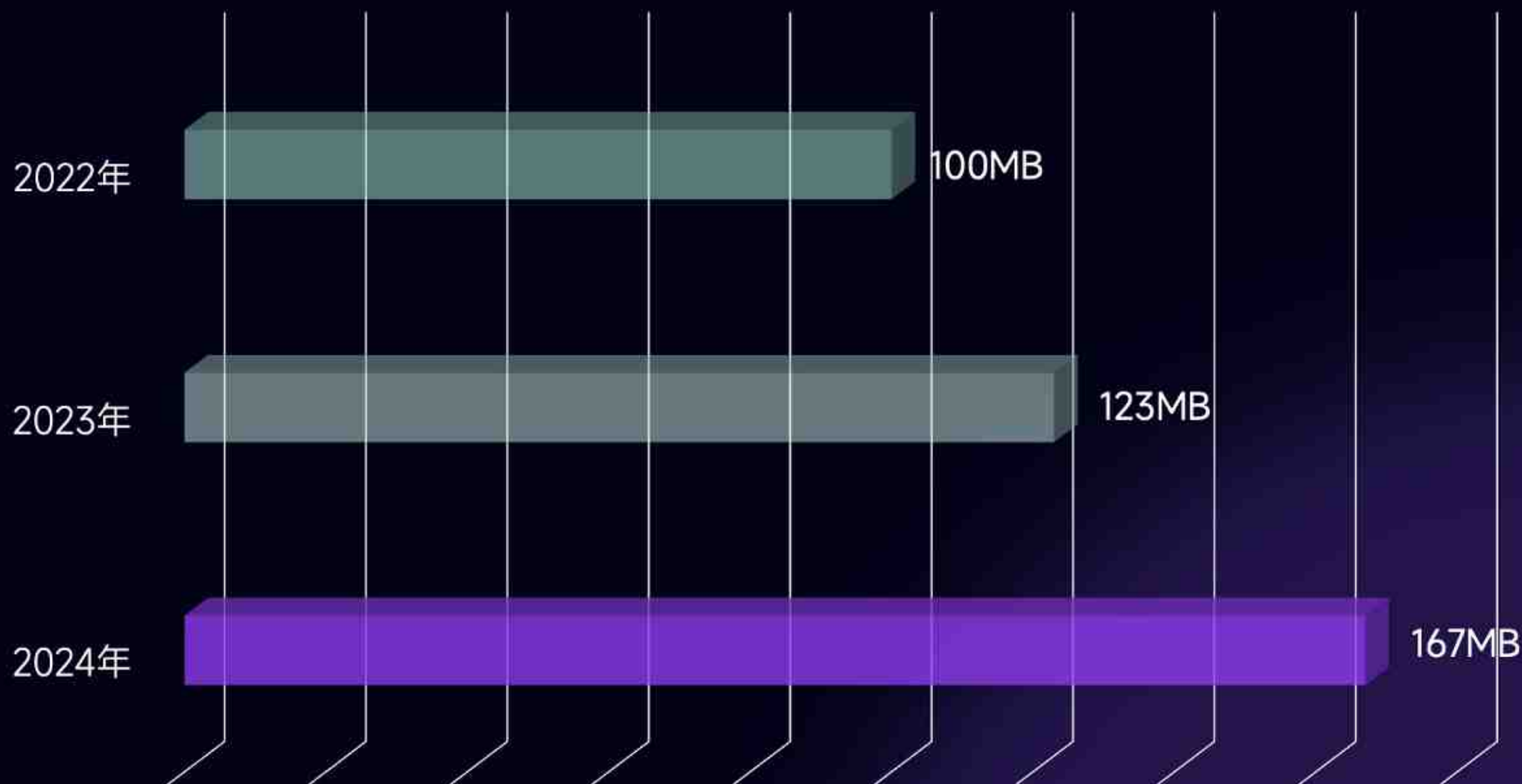
游戏体量的膨胀

游戏体量的膨胀也值得关注，现在可以被称为3A大作的游戏普遍突破100G；1997年，诺基亚6110首次搭载经游戏《贪吃蛇》时，存储占用小于5KB；今天想要在手机上畅玩《王者荣耀》或者《绝区零》，则需要预留10-50G的空间供其发挥。

而以开放世界游戏为代表的大型游戏，不管在任何平台上都受人欢迎。



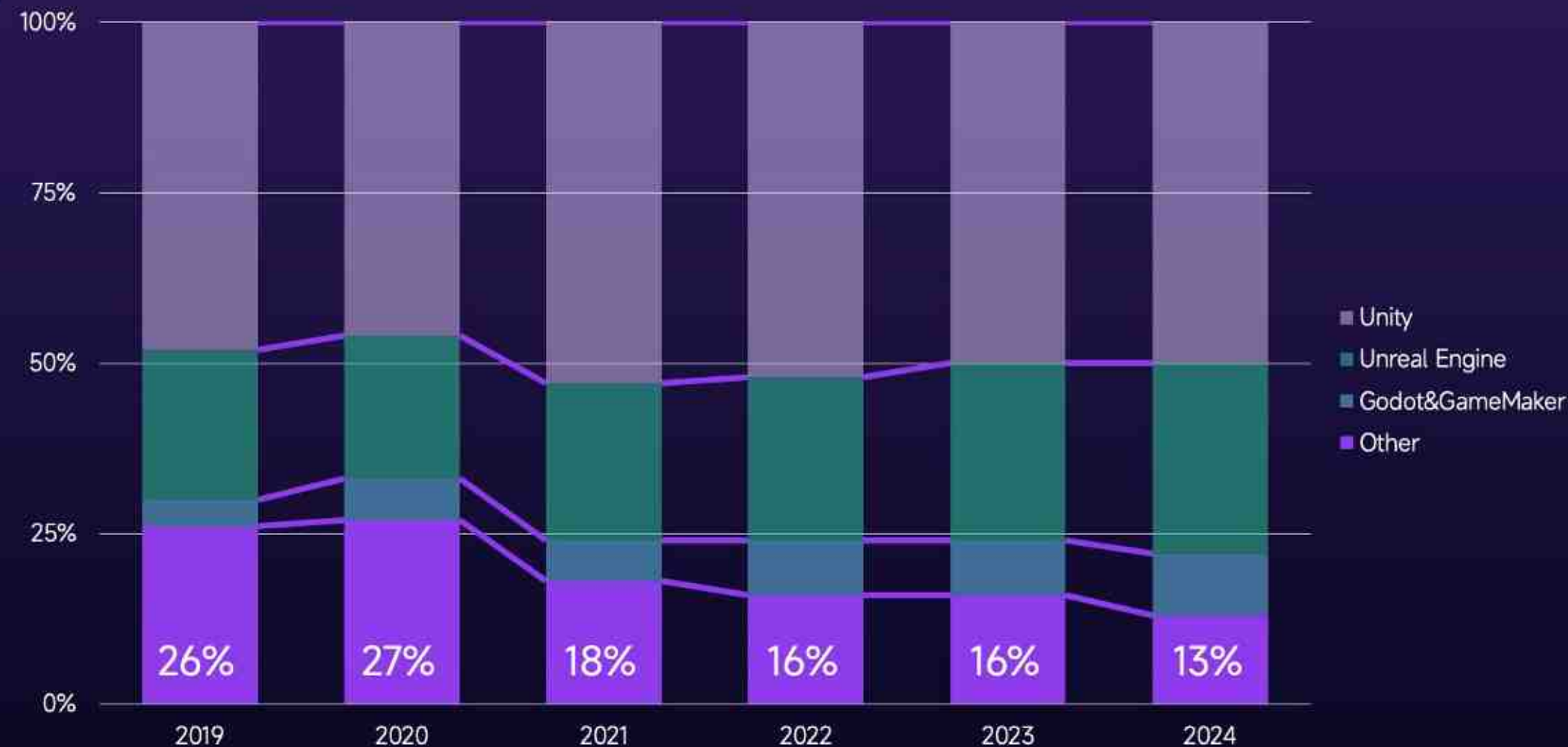
项目时间、投入和体量的不断膨胀，使得开发者对AI工具的态度集中在**降低**生产环节的各生产要素**成本**，并对AI提供的更多自由度玩法和游戏体验抱有相当高的期待，以获得更高回报。



图：显示了使用Unity平台构建的游戏中游戏容量的中位数，2022年以来增长了**67%**。(Internal Editor 分析数据，2025年，Unity)

开发者乐于利用AI： 节流与开源

开源游戏引擎的出现是一个超级转折点，它让开发者无需依赖大型发行商即可创作游戏。近10年，游戏开发使用其他（非Unity、Unreal、Godot或GameMaker的）自研引擎的份额不断减小。（2025年，VGI）



≈80%

2024年发布的游戏当中，有51%是使用Unity制作、28%使用Unreal Engine制作。来自Unity的数据表明，至少62%的工作室表示在工作流程中引入了AI；同时，Unreal同样在其开发平台中设计了AI组件。



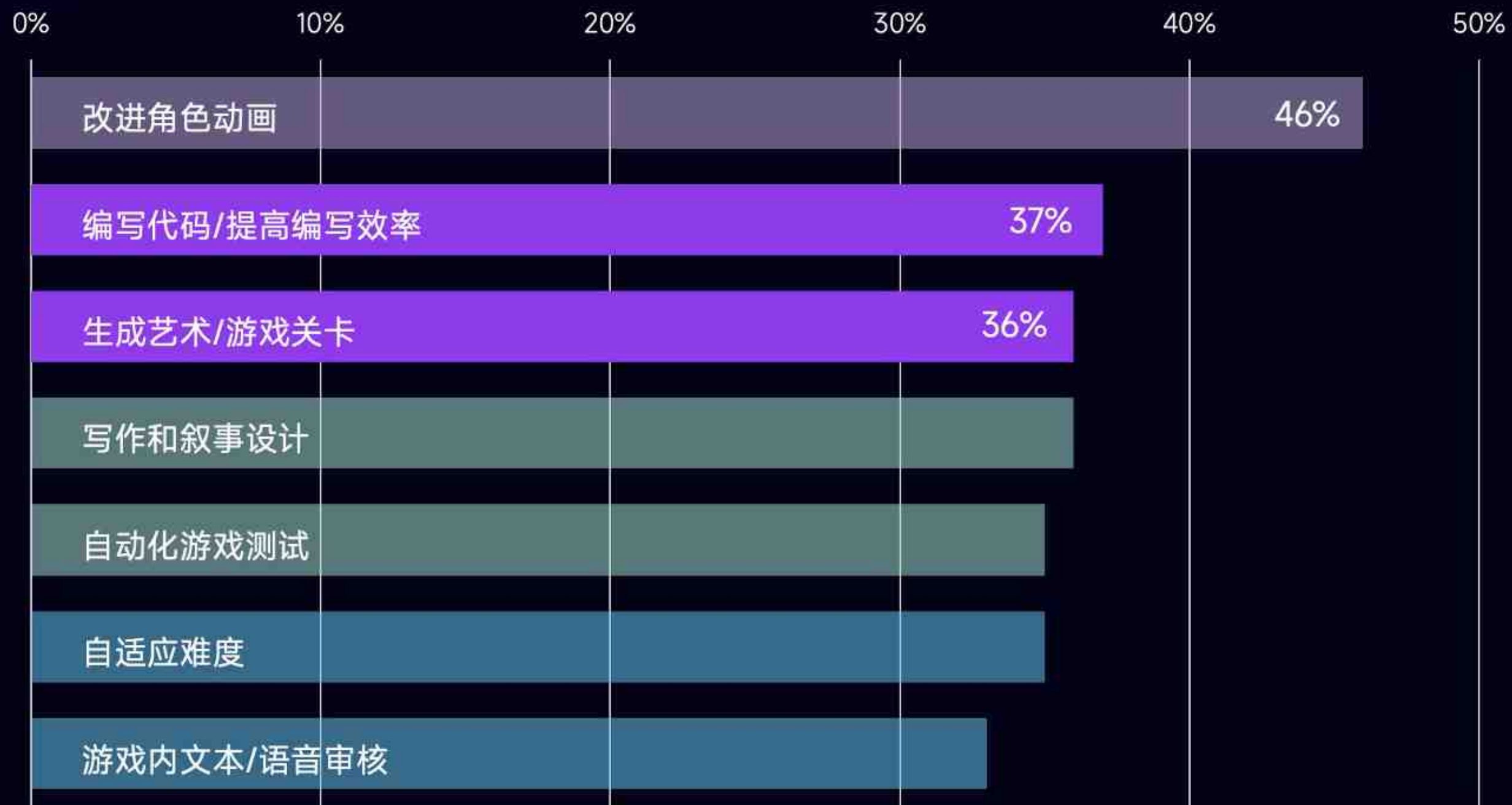
- Unity Muse：加速实时3D应用程序创建、查找工作代码示例等
- Unity Sentis：将AI模型嵌入到实时3D引擎的解决方案
- 第三方AI插件支持：包括AI驱动的智能NPC、视觉特效、纹理、2D精灵等



- NavMesh生成：2D/3D路径规划
- MassEntity框架：大规模AI代理模拟
- 第三方AI插件支持：包括AI关卡生成系统、语音驱动角色和实时对话生成等

作为几乎所有游戏项目开发的“第一步”，我们可以看到，在今天的游戏创作里，**想要完全抵抗AI是不可能的。**

节流：使用AI工具节省时间



我们可以通过一定范围内的样本数据调研，观察到游戏创作者对待AI工具们的态度。



使用AI来进行快速制作原型、模拟概念设计和重复资产创建是游戏开发中最具落地价值的三大场景，他们起到了同一个效果——使用AI工具为开发者节省时间。

其中，37%的受访开发者表示正在使用AI加快编码速度，而36%的开发者正在生成美术艺术和游戏关卡、测试游戏循环以及自动处理叙事元素。

图：游戏开发过程中使用AI实现哪些功能？ (Cint™ 调查问卷，2023年，Unity)

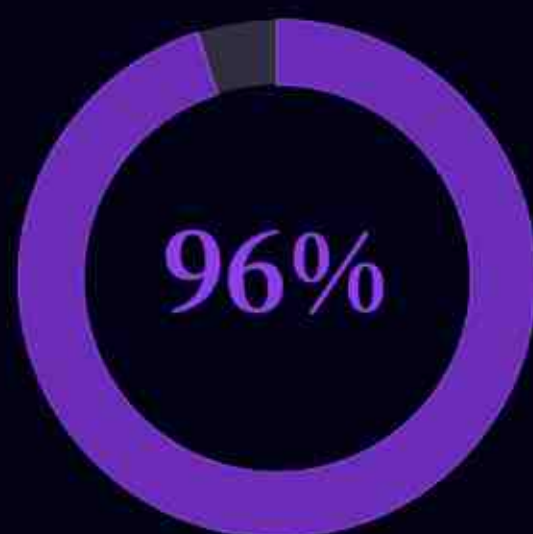
节流：加快原型设计

68%

万事开头难，对大部分开发者来说从零开始开发一款游戏会有些无从下手。在2023年的一项调查研究中，68%使用AI的开发者认为，使用AI的主要原因之一是这类工具可以**加快原型设计和概念设计速度**。



2022年



2023年

39%
小于1个月。

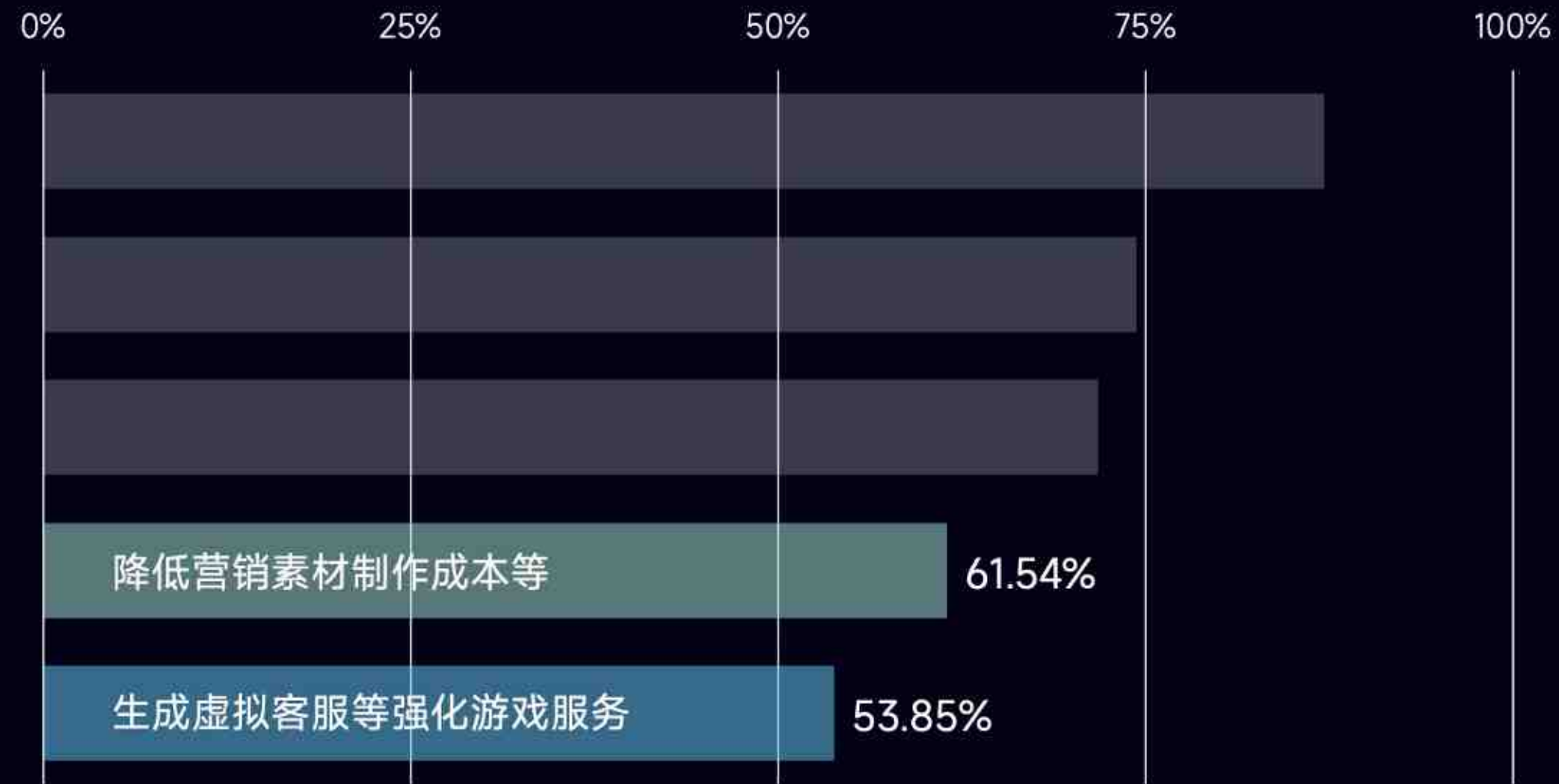


54%

的开发者制作原型所需时间小于3个月。

原型设计的速度确实也有所上升。2023年，原型设计时间不超过3个月的工作室占比达到96%，而在2022年这一比例为85%。

节流：减少消耗， 优化营销和运营流程



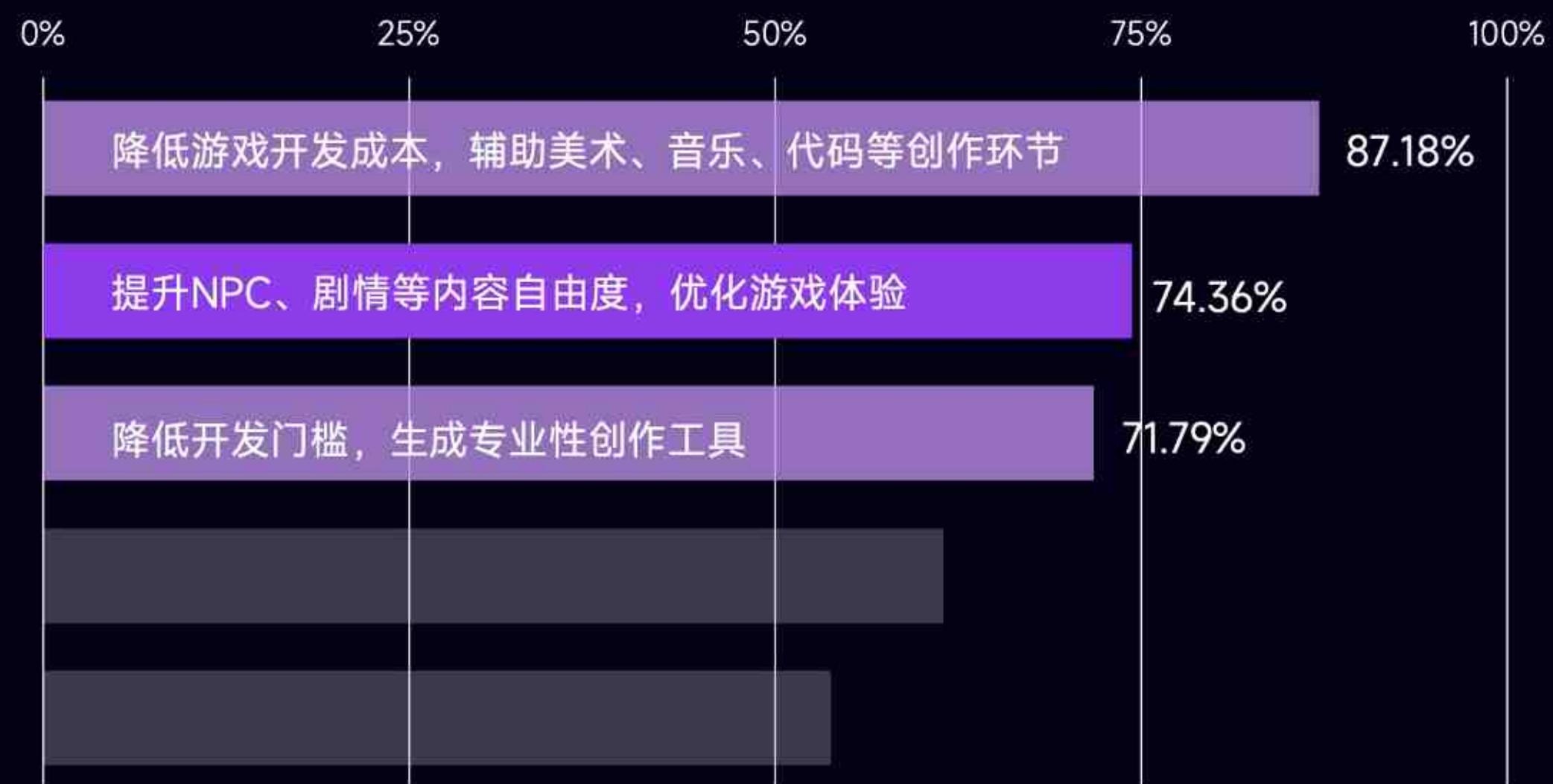
图：对于营销和运营环节，游戏开发者也期待AI能提供帮助。（AI相关技术对游戏的长期影响，伽马数据）

71%

在使用AI的工作室中，71%表示AI帮助优化了交付和运营流程。（2023年，Unity）

再结合游戏开发的日程越来越长，快速概念化以缩短发行周期，并把更多时间转移到细节的构建和优化上能创建更多项目优势。

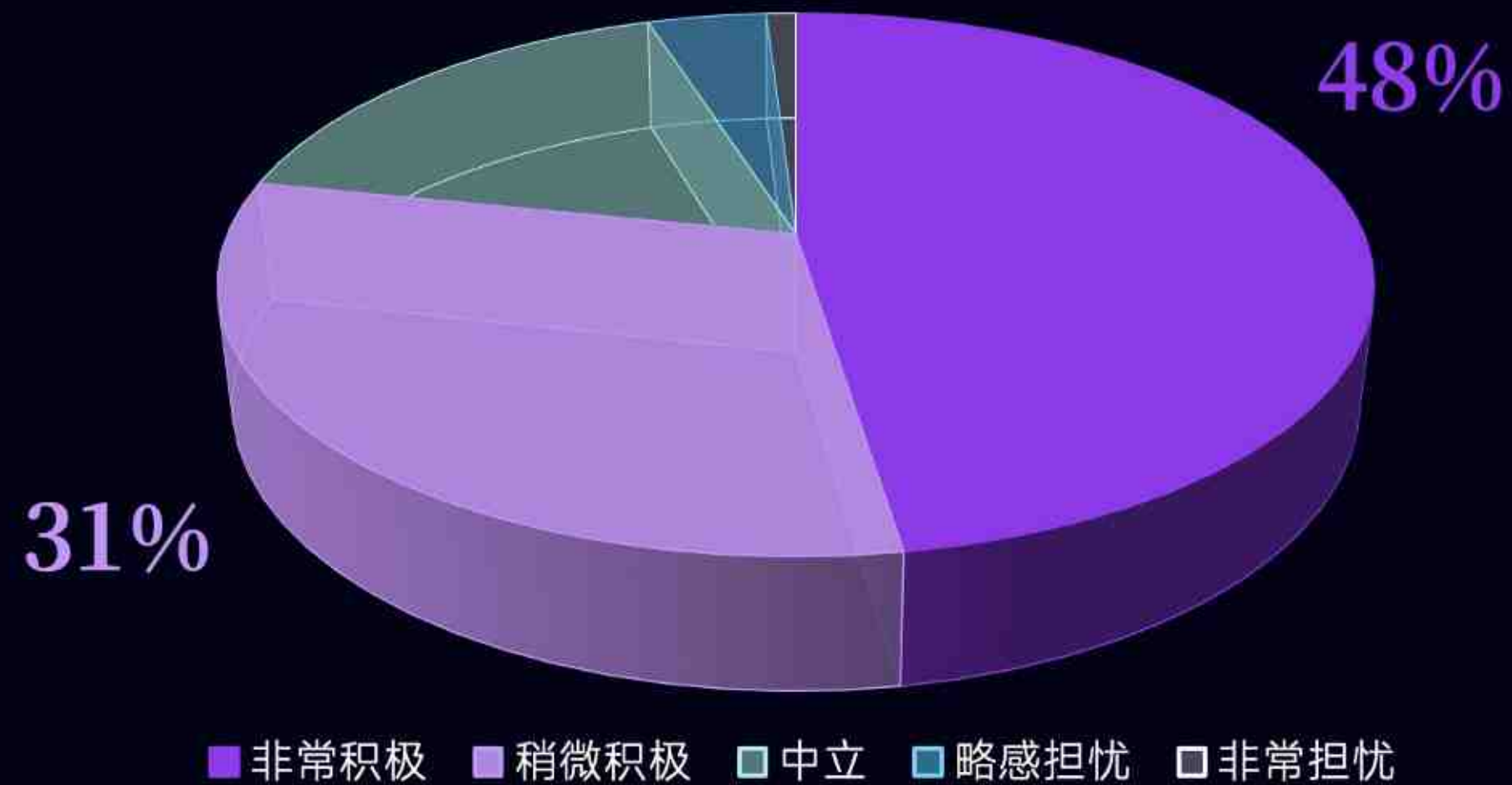
开源：帮助创作带来效率与创意



图：游戏开发者希望AI能帮助降低门槛、帮助创作。（AI相关技术对游戏的长期影响，伽马数据）

尽管开发者对AI工具的态度集中在降低生产环节的各生产要素成本，但仍有约75%的开发者对AI提供的更多自由度玩法和游戏体验抱有相当高的期待，这在AI技术对游戏长期影响中占据第2位。

开发者们对AI工具的积极态度



图：79%的开发者对AI持积极态度，总体来说是正向的。（Cint™ 调查问卷，2024年，Unity）

开发者们对AI工具持积极乐观的态度。2025年，在Unity平台上不计划使用AI工具的开发者的比例仅有4%。



总体来说，开发者有利用AI构建新的技术基础、重构工作流程的需求，以期在游戏研发领域形成新的竞争力。单项技术在工作流中的成功利用最终会产生重大的企业级影响，企业的选择又会影响游戏市场的走向。



Part 2

游戏创作的AI革新

AI技术对游戏产业造成了一系列冲击，并产生了连锁反应。游戏创作的各环节可以借AI革新工作方法。总体来说，AI带来了质量和效率的双重提升。

整个游戏管线的改变： 8个方面的积极影响

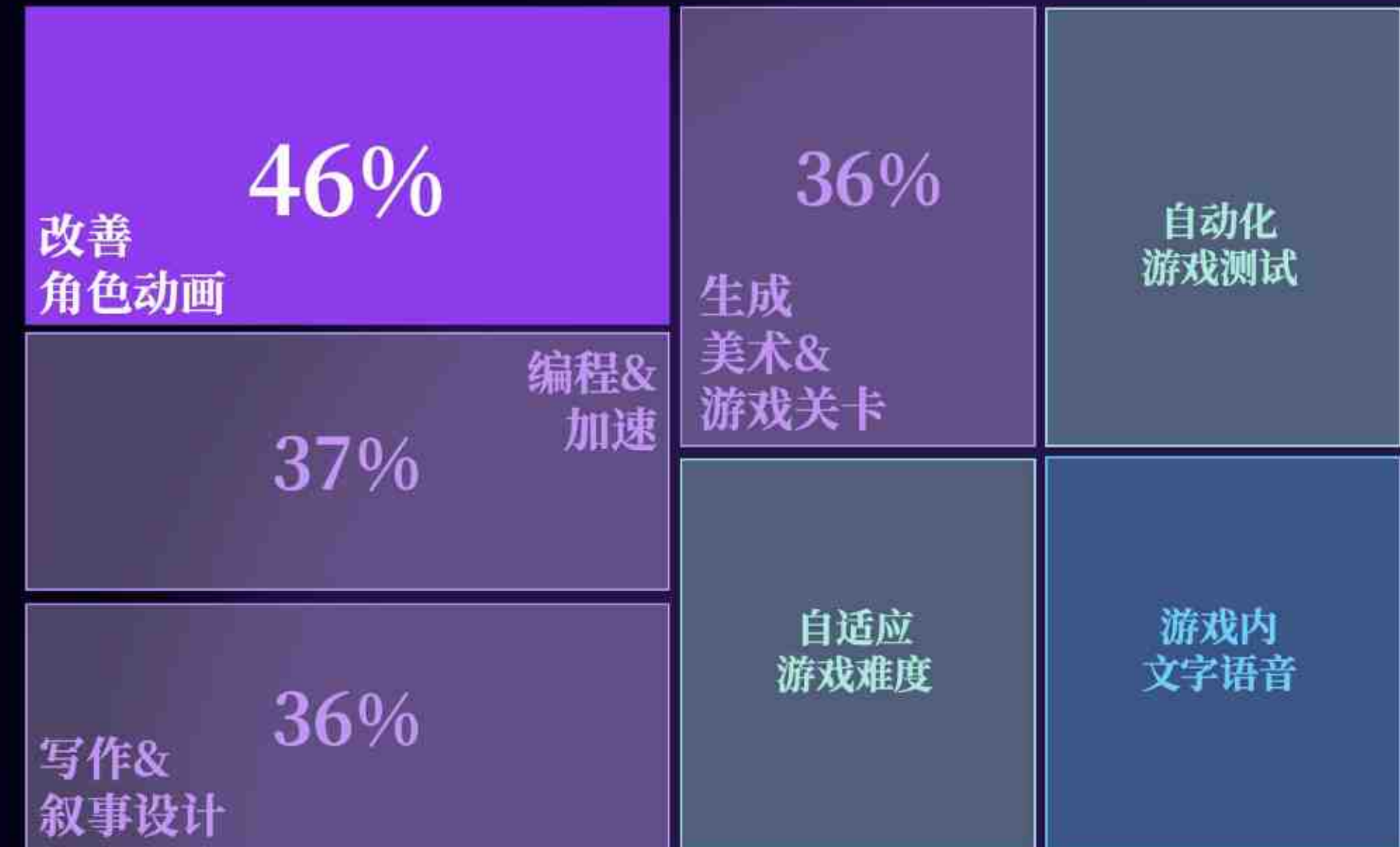
一个传统的游戏制作管线将游戏的诞生划分为3个类别：预制作、制作和后期制作。一款游戏往往萌芽于一个充满灵性的想法，随后制作出模型、草图、故事梗概和游戏风格。在建模阶段美术师们开始将一个愿景转化为后续可用的资源，同时游戏资产的大部分都是在制作阶段创造的。



AI在8大方面对游戏创作 workflow 产生了积极的影响，包括：



进入游戏管线的不同部分，AI承担的工作各有不同：

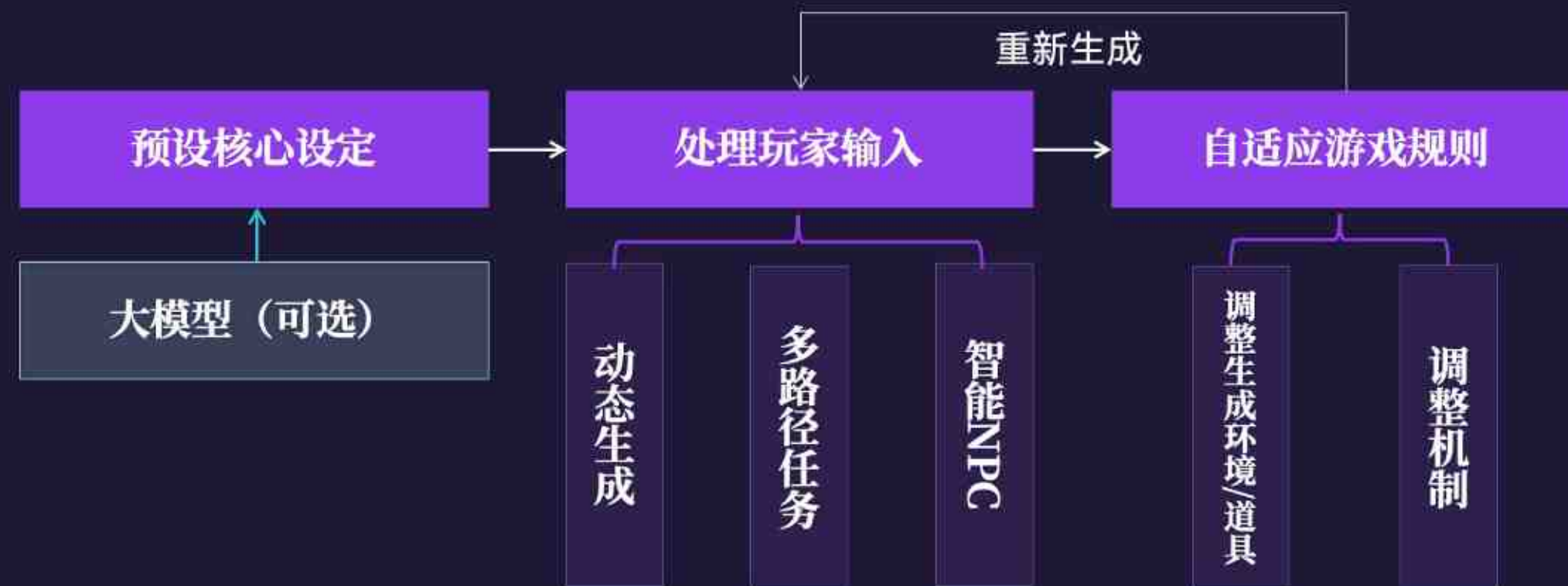


图：展示了游戏开发者使用AI的开发环节，涉及各种不同的方向。（Cint™ 调查问卷，2023年，Unity）

在游戏开发的「核心架构与底层设计」环节中，我们探讨了AI在游戏引擎方面的进步，接下来我们将继续研究AI在其他方面的积极影响。

叙事与角色塑造： 非线性叙事的要求

非线性叙事强调玩家选择对故事走向和情节分支的影响。以前，游戏开发者采取编写大量文本和故事线来实现多样化的故事结局；现在，我们可以利用AI来创造游戏内更多可能。



图：AI驱动非线性叙事的框架。

AI在创作游戏故事背景、塑造角色等方面，首先依托最基础的自然语言生成用法：

- 使用自然语言处理时，AI可以根据玩家的选择**动态调整**故事情节和角色行为；
- AI生成的非线性叙事和**动态角色发展**能够极大丰富游戏体验，更深入地推动游戏叙事和互动性的边界。

64%

- 约64%的游戏创作者倾向于使用AI为他们开发的游戏世界**填充NPC**；

56%

- Unity的数据证实，56%的AI使用者在**世界构建**过程中使用了AI。

为了创造更加真实的开放世界，游戏开发者大胆使用各种方法为玩家打造沉浸式游戏体验。

决策式AI阶段的智能NPC

使用AI控制智能NPC在游戏中早有实践：

行为树 Behavior Tree

- 模块化、灵活性、可扩展：**游戏开发者添加、修改和重组NPC行为，运行时也能动态调整结构；
- 复杂行为决策：**能够处理更复杂的决策逻辑，使NPC展现更自然智能的行为；
- 可视化调试：**Designer等可视化工具，可以很容易地理解和调试AI行为。

决策树 Decision Tree

- 在游戏AI中，决策树可以有效管理、控制NPC的行为，使NPC根据游戏环境和玩家行为做出合理反应；
- 决策树可以定义NPC可能执行的所有行为，并选择影响NPC行为的属性，如「玩家距离」、「玩家状态」，**节点根据属性做出决策**，并决定NPC的下一步行为。

状态机 Finite State Machine

- 一个NPC可能具有「观察」、「攻击」、「逃跑」等各种状态，每一种状态都对应一系列的行为和转移；
- 状态机允许游戏开发者将复杂的AI行为分解为更简单的**状态和转移**。

效用AI Utility AI

- 效用AI使得NPC可以通过**优先级和权重**来选择行动；
- 实现效用系统时，游戏开发者会根据环境和世界状态定义一个公式或曲线，来解释不同因素对**行动得分**的影响。

目标导向动作规划 Goal-Oriented Action Planning

- 设定NPC行动序列**，如寻找资源、避免危险、攻击敌人；
- 不同游戏类型和环境中，NPC还可通过GOAP实现战术策略等高难度行为。

生成式AI阶段的智能NPC

生成式AI阶段，游戏开发者在决策式AI的基础上融合更多思维和方法用以训练智能NPC。

AI Copilot

即时响应指令交互，与玩家合作战斗
实时游戏状态决策，为玩家提供帮助
风格化对话属性，与玩家进行情感交流

AI NPC社会

基于预设人设背景，role-play对话能力
具有长期规划能力，完成开放场景下的复杂任务
动态变化社会关系，衍生出新的剧情故事

训练形成

高质量数据集训练

- 针对性训练数据集，强化NPC模型的基础能力；
- 如角色知识、对话能力、情节演绎和逻辑推理。

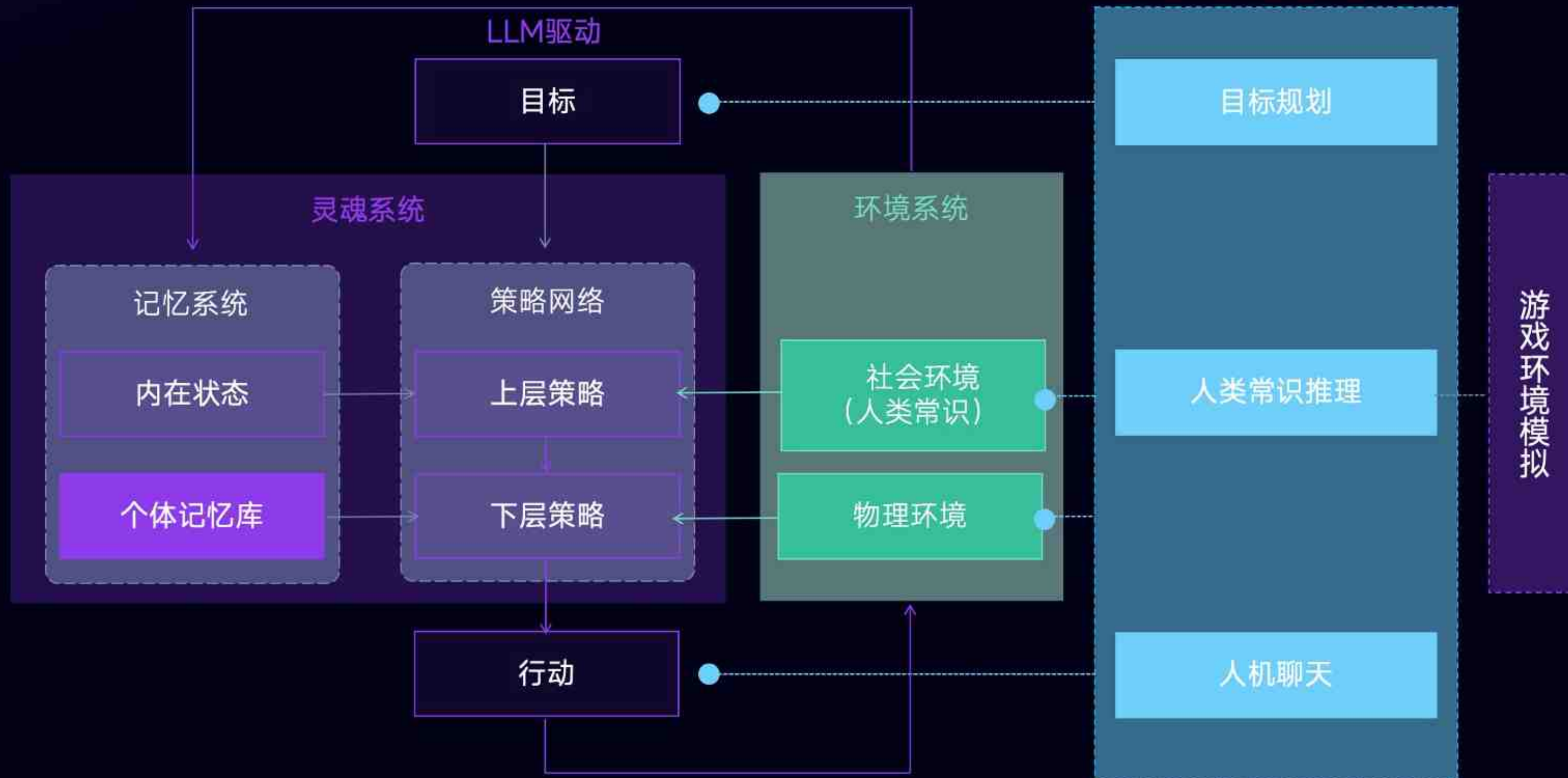
多模态感知能力

- 整合视觉、语义和音频数据，提升NPC对游戏环境的理解，使其更好与玩家互动；

智能决策优化和迭代

- AI不断优化自己的策略，适应游戏环境和玩家行为；
- 根据玩家反馈和游戏数据，不断优化NPC的行为和对话，提高智能性和互动性。

生成式AI阶段的智能NPC



图：AI NPC方案。 (2024年, 超参数科技)

陪玩/队友Copilot:

- 具有自我人设, 可进行风格化问答
- 能够感知玩家与游戏状态, 帮助或与玩家合作
- 毫秒级响应的指令交互
- 为每个玩家配备的专属长期记忆

动态变化的NPC社会:

- AI与玩家平权参与游戏、多维度交互
- 富有变化且可传导的Player+AI NPC社交网络
- 主题指引下的故事表达

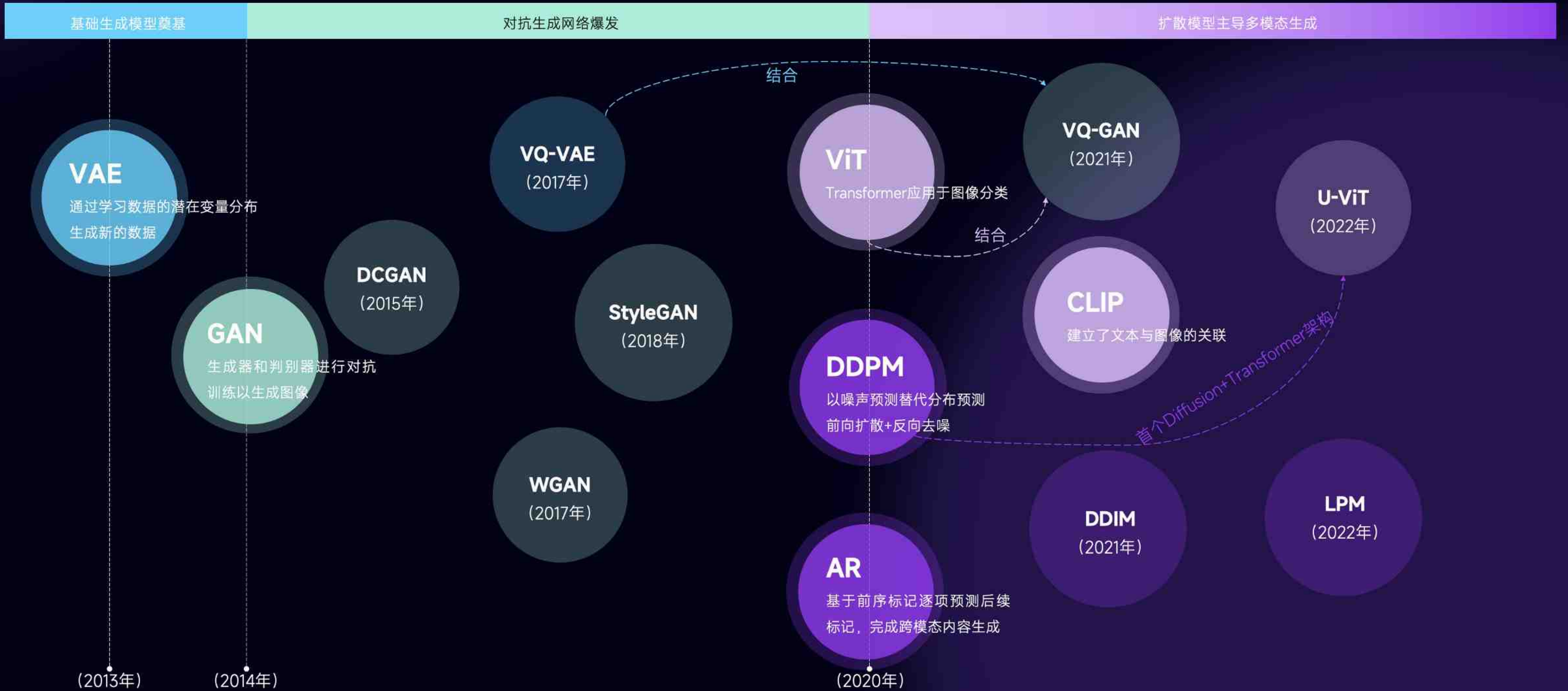
美术资产与视觉表现： 2D美术生成

AI生成图像领域中，生成模型(Generative models)和神经网络渲染(Neural rendering)是两个主要的技术分支。

十年以来，AI图像生成经历了生成对抗网络（GAN, 2014）、去噪扩散模型（DDPM, 2020）到多模态自回归模型（AR）和多种混合模型融合的过程。生成式AI在图像创作上的表现一直是直观且最能引起注意的，到2022年，主流生成式图像应用DALL-E 2.0、Stable Diffusion和Midjourney的诞生，使图像生成的精细化程度得到有效改善，通过提示对图像生成结果进行控制的能力有很大提升。

对游戏而言，**图像生成开始具备进入游戏管线的条件。**

图像生成模型的发展



即插即用的文字生成图像产品 促进AI的游戏工业化

	DALL-E	Imagen	Midjourney	Stable Diffusion
特性	<ul style="list-style-type: none"> 具有易用性和对高级提示的准确理解； 有限控制性，在需要像素级精确、多资产一致或者有微调复杂参数的游戏开发任务中表现不足。 	<ul style="list-style-type: none"> 照片级真实感、高精度； 对深层语言有精确的理解，因此具有更强的控制力。 	<ul style="list-style-type: none"> 艺术表现突出，可以生成生动且独特的设计； 可使用先前生成图像作为新提示的参考，强化特性。 	<ul style="list-style-type: none"> 支持使用自定数据集微调模型，实现高度个性化的图像质量和一致的风格； 支持负面提示，有强大的控制性； 提供广泛的参数选项来微调输出。
适用范围	<ul style="list-style-type: none"> 快速构思、概念性设计； 早期探索阶段的优秀工具，而非遵守特定设计规范的最终资产。 	<ul style="list-style-type: none"> 高保真的视觉资产； 视觉概念模型的快速原型制作。 	<ul style="list-style-type: none"> 尝试新颖艺术风格； 需要高效地实验、迭代和完善效果时。 	<ul style="list-style-type: none"> 快速生成高质量游戏资产； 从文本提示创建逼真、详细的3D场景。
主要架构	Transformer-based	级联扩散+LLM编码器	GAN+深度学习+扩散模型	LDM
集成/访问	GPT/API	API	Discord	本地/API
开源	✗	✗	✗	✓

游戏工业化强调的是**标准**、**规模**和**自动化**，仅仅是技术突破的实现还无法在短时间内对产业造成影响。2D美术生成能够在较短的时间内被游戏开发的流水线吸收，生成结果精细且稳定的文本到图像生成模型（产品）功不可没。

AI美术工具的权衡在即时易用性和深度的精细控制之间。

通过文本生成详细的场景和角色美术，重新定义了设计的工作流程。

AI驱动的产品增强了可访问性，有效降低开发者的技术门槛，使高质量的游戏内容创作变得更加普及，这意味着一个更具包容性和多样性的开发系统。

利用AI改进神经网络渲染

传统渲染

逐帧计算光线路径、模拟光与材质、环境交互

VS

AI渲染

预测光照效果、材质细节和噪声去除，无需计算真实的光线路径

优化算法

- 优化光线追踪
- 优化物理渲染，模拟流体、烟雾效果更好
- 多视角合成，可以作用在VR/AR中

改进流程

- 集成进渲染流程，比如AI预测光照、阴影
- 实时动态场景渲染，并保持高训练和储存效率
- 实现端到端生成，比如使用神经网络直接输入场景描述、相机参数等生成图像

创新能力

效果上:

- 结合生成模型，生成具有复杂场景和光照的逼真图像
- 强大的泛化能力，可以渲染未见过的场景和条件

使用上:

- 交互式神经渲染，可以实时修改场景并看到结果
- 风格化渲染，最小化生成图像与风格图像、内容图像的差异

神经网络渲染的发展

1

NeRF

从部分二维图像中重建复杂的三维场景，并从新的视角渲染出逼真的三维视图。

● 进步：

- (1) NeRF的**隐式渲染**使用一个相对较小的神经网络代替了显式渲染中采用点云、网格、体素等形式，将场景建模为一个连续的5D辐射场，根据这个模型可以渲染出任意视角下的清晰图像；
- (2) 在表达复杂场景时其**参数量**相比显式表示更少，所需占用的**内存、算力**相对更少。

● 价值：

降低高保真内容的生产成本，在静态环境构建中更有优势。

● 与传统三维重建相比：



2

高斯泼溅 (3D/4D)

空间点云数据作为基础表示，生成柔和连续的三维表面。

● 进步：

- (1) 比NeRF有更低的计算资源需求、更快的渲染速度，并支持高FPS的实时渲染；
- (2) 在减少渲染伪影、渲染高光、反射等细节时表现出色；
- (3) 显式结构支持**快速编辑**和**动态调整**。

● 价值：

在游戏中这意味着游戏场景的动态加载和**实时交互**成为可能，同时无需重构整个场景即可修改高斯泼溅效果，这种灵活性降低了游戏开发的**迭代成本**。

● 与NeRF相比：

	NeRF	高斯泼溅
兼容性	部分完成渲染工作但在Maya、Blender等常用CG软件中直接编辑较困难	与现有渲染引擎融合更好
算力	渲染较简单的物体也要用相等算力需额外进行Marching Cubes等格式转换，增加计算开销	可根据渲染复杂度自适应调整用于表达的高斯数目
硬件支持	与GPU擅长的架构不一致	基本兼容

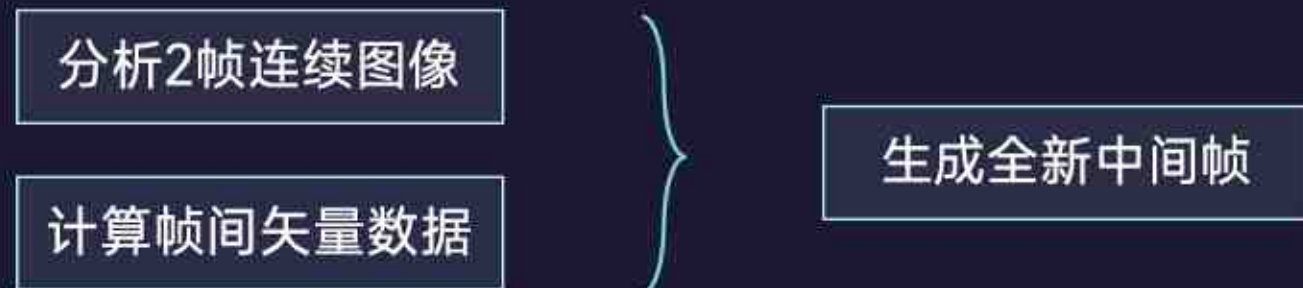
神经网络渲染的发展



DLSS 3/4

通过神经网络模型提升实时渲染帧率与画质。

	帧生成	多帧生成
运动矢量 光流信息	光流处理器 OFA单元计算	Tensor Core通过 神经网络生成
SR超分 RR重建	CNN	Transformer



在原有帧的基础上，AI模型能够额外生成帧以提高游戏帧率，实现性能翻倍的效果。

- 进步：
 - (1) 以AI生成代替传统渲染流程中的部分计算，突破高分辨率、光线追踪等场景下的性能瓶颈；
 - (2) 在处理复杂几何和光照细节等方面有更强的表现力，能创建或还原原生画面中可能被忽略的细节。
- 价值：
 - (1) 硬件替代：使得中端显卡也能在高画质下运行只有旗舰显卡才能驾驭的游戏；
 - (2) 软件集成：在D5 Render等渲染软件中使用，帮助开发者在实时渲染和预览时获得更高帧率。

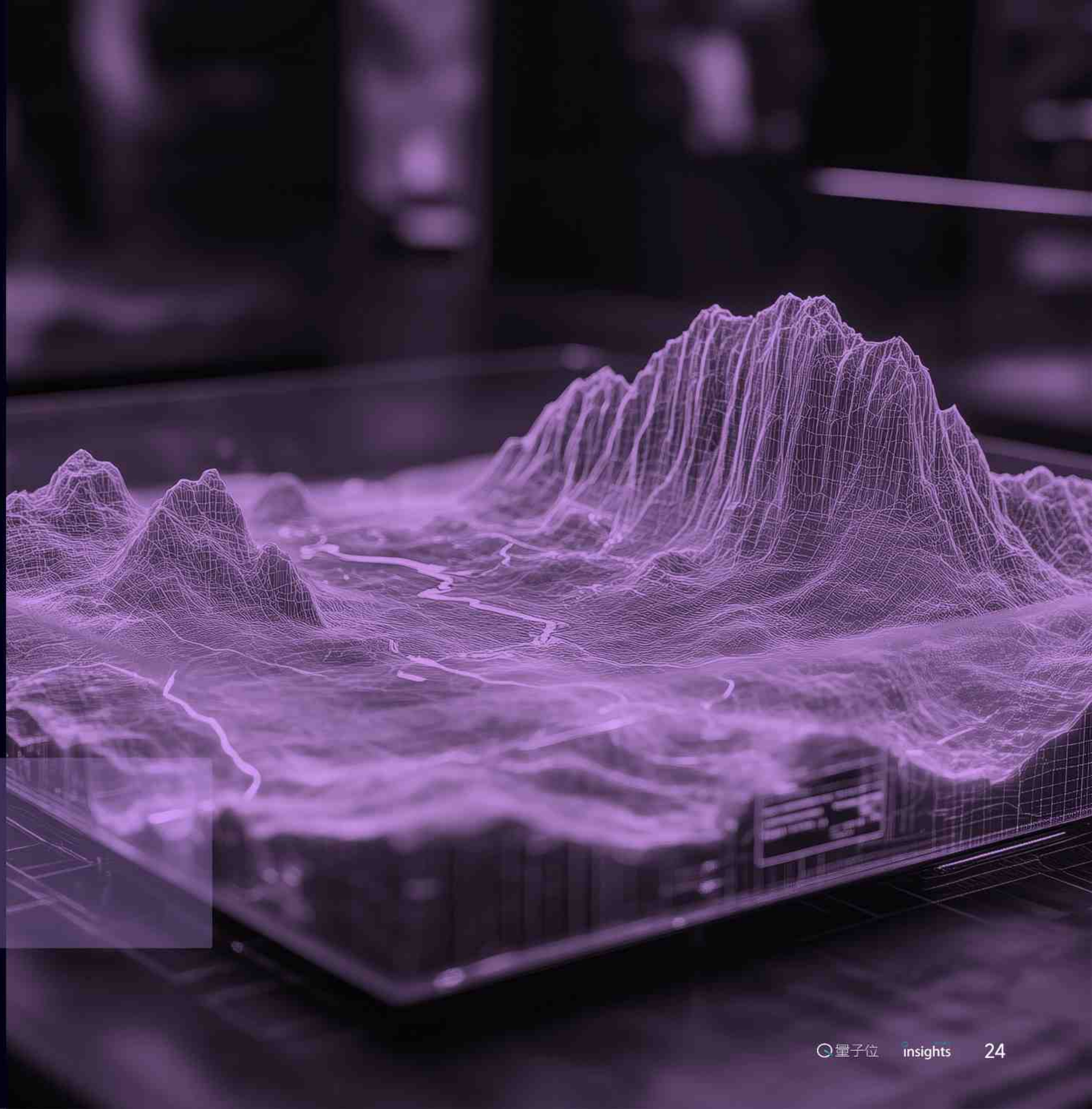
相比图像生成模型，神经网络渲染在游戏美术领域利用的广度和深度都是另一个量级。



综合来看神经网络渲染的各项技术应用和突破都可以缩短渲染时间、减少显存占用，自然而然可以降低成本；同时大幅提升帧率、增强图像质量，达到“价廉”而“物美”的境界。

3D美术工作流： 3D模型生成流程的 「拆解」和「统一」

3D模型资产和3D渲染在游戏创作中有着和2D美术同等重要的地位。3A游戏通常仅在3D建模上就花费数千万元，以确保获得最一流的模型。现如今，90%以上的游戏开发需要使用3D建模与3D渲染效果。



传统的3D美术 workflows



游戏3D建模往往要从基础的几何体开始，随后使用雕塑法增加更多的细节和褶皱。3D建模平台会给出一个模型需要的细节程度，模型越是细节化、面数就越多，生成它所需的制作能力就越强。

workflow中的每一个环节几乎都需要一位或一组3D美术工程师来完成，根据他们负责的工作内容被称为绑定师、动画工程师、渲染工程师等。到今天，成熟的游戏工业化流程将制作3D模型资产的每一项工作都精细拆解，以创造令人惊叹的精美作品。

探索3D生成的原因： 高昂的费用和漫长的周期



- 游戏3D建模的价格差异很大
- 经验丰富的3D专家可按小时对3D项目收费



- workflow里的许多结构都需要花费大量时间
- 耗时受模型复杂度、技术要求、工作效率和执行经验的影响

生成3D的3种思路

1

基于2D升维技术的3D生成
(Generate Novel View Synthesis)

作为生成3D的主流思路之一，2D生成3D通过创建深度图、立体生成和未覆盖区域重建等步骤，将2D转为有深度感的3D表现。

2

直接生成3D

直接在3D数据集上进行训练和推理，从各类3D几何形状中提取丰富的3D先验知识。

- 直接从文本生成3D对象的模型在使用复杂提示时，经常遇到语义上的困难。

3

生成3D环境

强调的是生成3D可交互世界。

- 更进一步的商业化呈现是一键生成3D游戏。

生成3D的2种主流方法

回归

- 用FNN进行线性/非线性变换，生成3D表示。
- 推理速度快，但对重建有遮挡区域的效果不好。

扩散

- 通过迭代采样能生成多样化的3D结果。
- 但计算效率低，且与输入图像的对齐效果差。

数据集的选取

- 2D图像只能记录真实物体的一个侧面，无限多角度的图像也无法完整描述3D内容；
- 而一开始就使用3D原生数据，又要面对训练数据不足和生成结果不可控的问题。



多种思路下，有很多算法可以实现3D生成。
在模型开发的过程中，开发者会在多种生成思路下
混用算法、数据集。

不同的3D生成方法实现路径不一，但可以通过几个**关键的标签**来区分：

生成模型	用于创建3D内容的基础算法或框架	Normalizing Flow/GAN/VAE/自回归/Diffusion.....
生成空间	将高维数据映射到空间中	Lantent Code/Voxel/Polygon/Pixel/Later set/Tri-plane/Point.....
重建空间	生成对象的中间或最终表现形式，决定了3D对象如何被格式化和呈现，影响3D模型的质量和可用性	Point Cloud/Mesh/Tri-plane/NeRF/SDF/高斯泼溅.....
渲染	将3D模型数据转化为2D图像的过程	Mixed Rendering/Network/Surface/Volume.....
监督	模型在训练过程中是否使用了带有标签的数据，监督学习通常能生成更精确的结果	3D/2D.....
条件	生成3D内容时，模型是否接受了额外输入来引导生成过程	Label/Image/Text/无监督.....

另外，现有的3D生成产品已经对3D美术工作流的其他部分产生野心，例如Meshy提供了可随时使用的**装配模型**和**运动库**，开发者可以尝试独立完成动画工作流程。

AI在3D（及其他游戏开发环节）中强调开发者独立工作**的能力，现有的**工作流程**可以**收缩**，创意的边界却能够**无限扩张**。**

3D生成技术对游戏领域的影响： 优化与提供



优化工作流程

通过将3D技术集成到现有流程中来简化工作流程。

自动执行重复任务；
使用插件和脚本以提高效率；
采用行业标准进行文件管理和协作。



改进2D渲染

3D建模可移动的视点，带来的高质量的深度
和清晰的透视即使在2D画面中也有很大价值。

已被应用于许多动画领域，在游戏剧情衔接和CG
上也有用武之地。



提供3D资产

提供真实感和细节（灯光、纹理、材质），创建
更真实、身临其境的3D场景。



VR和AR交互式沉浸体验的需求

新一代游戏设备对游戏技术的新要求。

3D美术未来发展方向： 强调细节和可编辑性



生成质量改进

下一步发展对可用细节的要求将进一步提升

3D模型的精细度

模型准确性

渲染分辨率

色彩与光影的准确性

渲染对材质的表达

.....



可编辑性

目前3D生成模型在修改方面表现不足

可以加强生成模型对自然语言指令与图像指令的理解能力

使模型与传统3D建模工作进行衔接，将3D建模工作与渲染工作分离，同时生成的模型为网格数据（Mesh）



比起直接可用的3D生成，在游戏

领域会更强调**可编辑性**

AI游戏美术得到广泛应用 技术上实现**精细化¹**、**多模态²**和**风格迁移³**

精细化程度

使生成的图像更加精美、笔触更加真实。

多模态生成

提供更多输入和输出的可能。

风格迁移技术

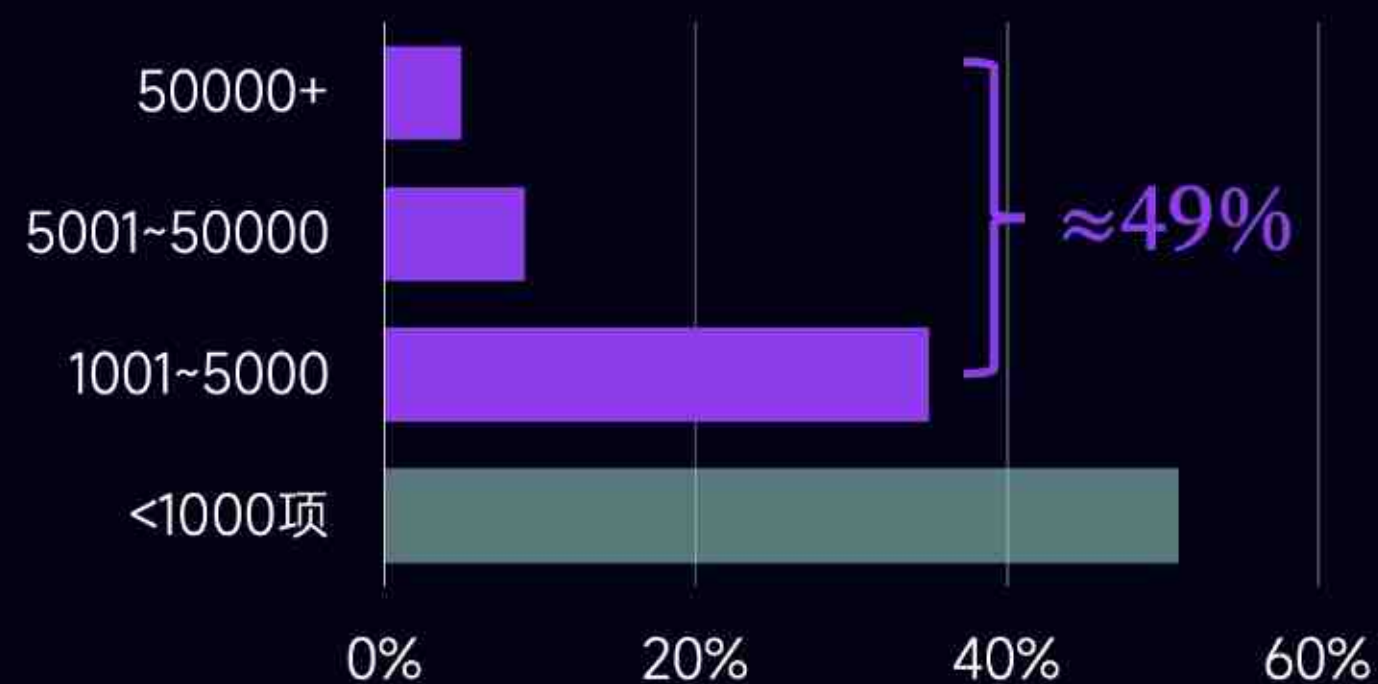
生成图像拥有更加可控的细节、明显的风格化特征且可复制可统一。

- 新一代图像生成产品受到极大瞩目的重要原因即他们在风格迁移方面的出色表现。例如Midjourney的风格调谐器（Style Tuner）使用户可以调试自己的个性化需求，进行风格混合和复制迁移。

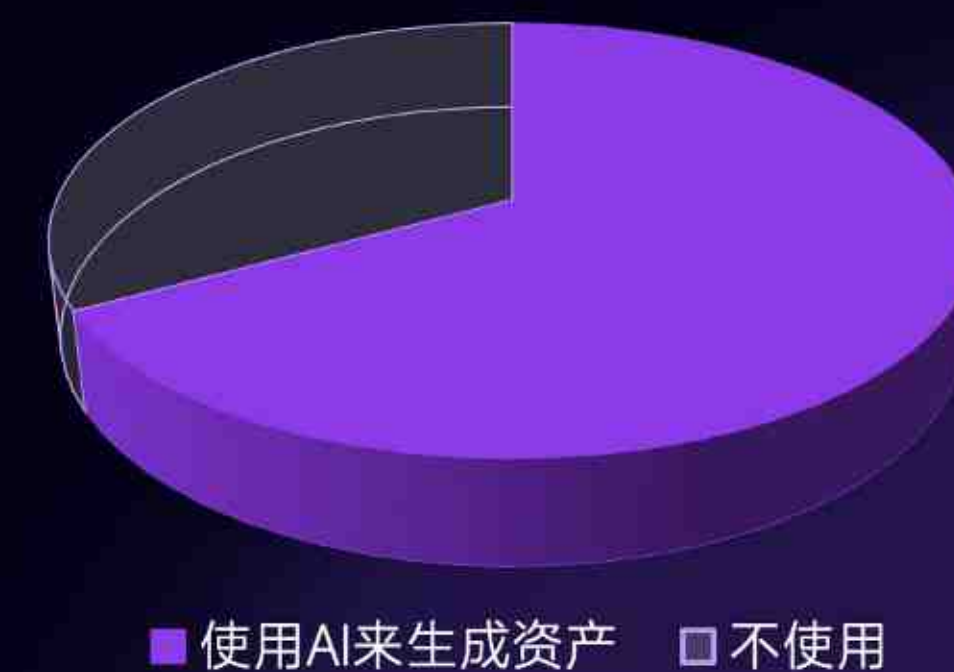
AI游戏美术得到广泛应用

游戏资产 **4** 的需求

更加宏大和真实的游戏世界观、玩家对开放世界的要求、新一代设备更强的显示能力等都在不断提高我们对游戏素材数量和质量的要求。游戏开发过程中，**资产管理问题常常导致开发工作延期。**

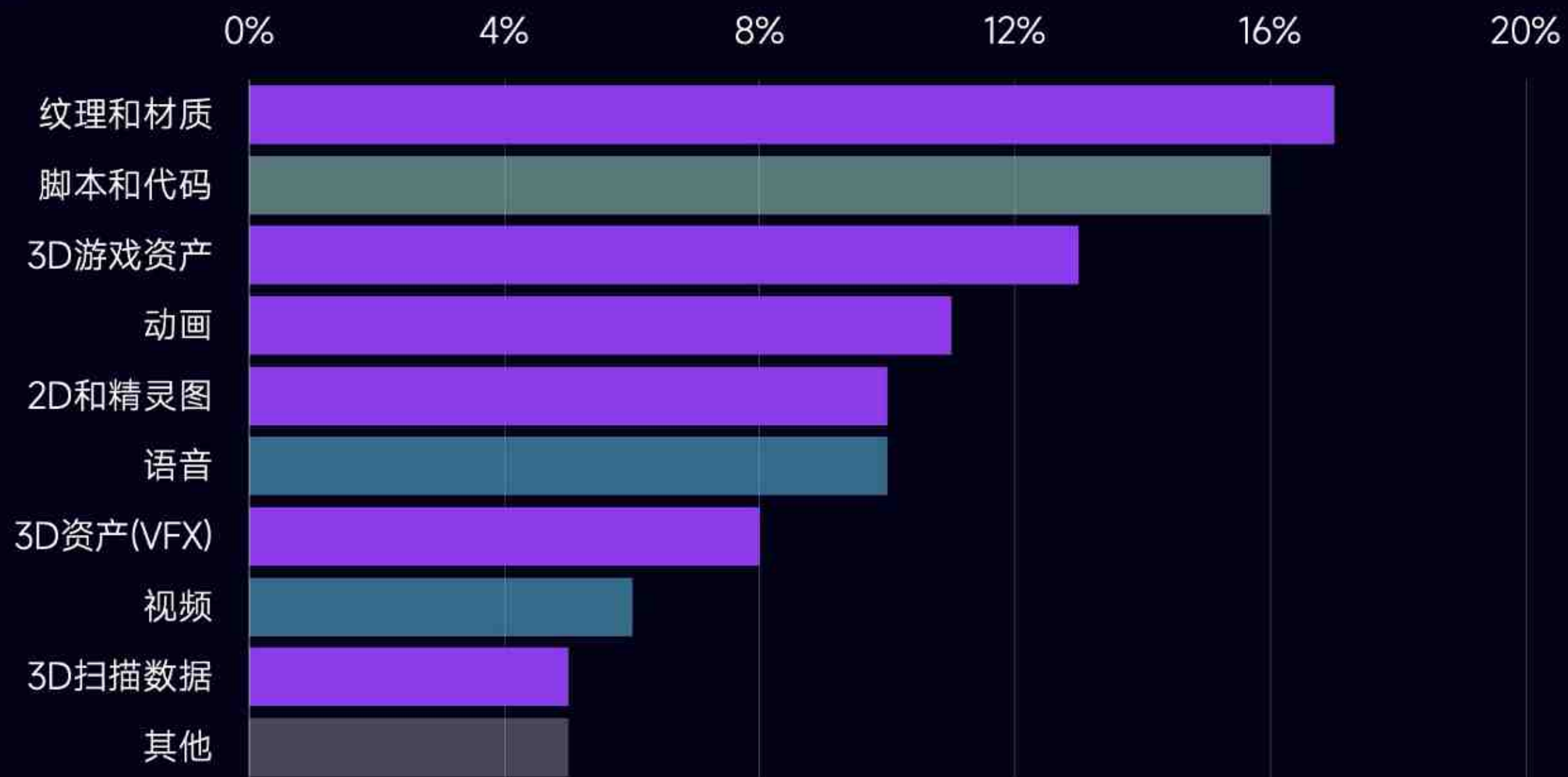


2023年一项有关游戏资产管理的研究显示，**将近50%的工作室使用超过1000项资产。**管理、协作和运转庞大的资产库对开发者而言是亟需克服的障碍，开发者尝试使用新方法来完成工作。



67% 的开发者表示在使用AI来生成资产。

AI游戏美术得到广泛应用 降低工作成本⁵的需求



纹理、材质、3D模型与2D图像都是游戏开发者最常使用的预制资产。

同时，AI进入游戏美术的制作工具流能够降低至少30%的工作成本。

量子位智库在对巨人网络AI应用的访谈中获悉，开发的AI绘画平台“iMagine”已经植入了美术团队的工作流中，在角色和场景原画的创作过程中，通过AI生成图像辅助美术团队的创作，能够节省50%-70%的时间。部分项目的内部UI和icon设计已经完全交由AI完成，在整体上实现了超过80%的效率优化。

音乐与音效生成

AI生成音乐正在追求更高质量的音频，目前不少模型都能达到44.1kHz或更高的音频采样率，AI生成音乐作为音乐资产在游戏制作和发行环节使用都是非常可行的。

像MusicLM等模型已经支持生成多音轨的作品，使用AI生成音乐为原型，佐以专业制作人的协调，将使AI音乐更快进入游戏制作与发行的生产线。

2种生成方法： 符号生成模型与音频生成模型

基于乐理规则的 符号生成模型

- 提取出音频中的各类音乐信息，如歌词、旋律、乐器等，并对这些信息进行标注，拆分成音乐的各个环节来生成词、曲等；
- 在此基础上训练并生成新的音乐符号，借由传统音乐制作流程将其渲染为音频。

基于音频数据的 音频生成模型

- 直接用海量音频数据训练模型；
- 使用深度神经网络和VAE或Transformer等方法，端到端生成音乐。

音频生成模型的不足

- 音频生成的多模态能力目前还停留在文本和音乐两种类型的输入上；
- 存在算力成本高昂、侵权风险相较于符号生成更大等问题。



实际中会同时采用两种生成方法，根据文本或原音频直接生成音频，可以生成乐谱并支持对已生成的乐谱进行自动编辑操作。

AI生成音频的3大使用场景



BGM/动态音乐

使用自然语言与音频进行对位，并逐帧生成音频的技术对传统的音乐生成方案是种颠覆性的进展。

加快产出

类比于语音合成的TTS技术，以此输出的音频更加自然流畅，并拥有类似人类演唱者的情绪氛围。在游戏中使用AI为环境、角色设计主题曲，可以很快形成风格相似的一系列配乐。

附加价值

但作词作曲家、演唱者和演奏者对游戏营销、推广方面的价值不可估量，在实际的游戏生产和发行过程中，AI对音乐创作和编排的渗透并没有想象中高。



环境音效/交互音效

游戏内音效是增强玩家沉浸感和提高游戏体验的重要元素。常见的音效类型有环境、界面、战斗、交互、提示音效等。

规避风险

尽管有很多开源的音效库，但使用AI生成的音效可以自行调整适合游戏主题的风格、减小版权风险。

选择空间

成熟的游戏厂商一般都有自己内置的音效库，采样自专业的音效师或最真实的自然世界，使用AI可以帮助独立游戏工作室或游戏创业者解决迫在眉睫的问题，还能提供和开源音效库不同的游戏体验。



TTS语音合成

将文本转化为自然、富有表现力的语音。

1

使用TTS语音合成的优势：

- CV与角色的真正分离，创造沉浸感
- 减小项目流程协调上的困难和各类风险

2

但综合来看，游戏语音中的AI利用程度与场景对「情感」的要求正相关。

人类对愤怒、喜悦、悲伤等6种基本情绪的声音识别准确率达92%，且不受语言文化影响，而音素模仿情感的能力却不足。

3

AI也试图从数据层面提升模型对人类情感音素的理解能力。

微软亚洲研究院VocalDNA项目分析了全球2000种语言的情绪录音，构建出首个人类情感声音图谱。



音乐生成

SOUNDRAW
LoudMe Mureka
udio BETA
Suno
网易天音

音效生成

Eleven Labs Epidemic Sound
KROTOS STUDIO
阿里通义 ThinkSound
Google AudioLM

语音合成

RESEMBLE.AI
MURF.AI
Google Cloud Text-to-Speech
Microsoft Azure AI Speech
Amazon Polly
Unity RT-Voice PRO
腾讯 GiiNEX
NVIDIA NeMo T5-TTS



AI音频生成已经从功能实现迈向情感化设计，新的标准提升到如何更适配游戏的场景及内容。

同时音频作为剧情演出的重要部分，AI音频生成正探索和视频生成的融合。

测试、优化与安全

在游戏开发中的测试及优化环节，AI往往出现在平衡性测试和功能性测试的场景中：

平衡性测试

- 敌方AI：AI模拟对手可以根据游戏难度调整其行为和策略，使玩家感受到不同水平的挑战；
- 队友AI：在匹配玩家缺失时，AI作为队友共同完成对局；
- 玩家行为模拟：自动发现游戏中的BUG和平衡问题，以此优化游戏的性能和资源分配。

功能性测试

- 使用AI快速搭建符合基本要求的模型环境：
- 生成和测试不同的游戏场景和玩法
- 测试某一项代码或功能是否能够跑通



此外，在线游戏行为中，AI Bots作为「虚拟玩家」参与游戏，能够替代真人陪玩用户，赋能开发者降本增效的同时提升游戏质量、增强玩家体验，提升游戏运营效率与效果。



图：GAME BOT技术方案。 (2024年，超参数科技)

有效提升DAU增长率等关键指标

投放AI后，案例DAU次月增长最高达2%；整体增长率稳步提升，最高可达8.35%



优化玩家体验

AI发起局间互动，中低段位玩家互动量提升50%以上，被点赞率提高90%；AI接管掉线玩家，掉线玩家客诉率相对下降200%

玩家类型	互动量变化
新手玩家	77% ↑
活跃玩家	53% ↑
资深玩家	46% ↑

识破率低

筛选大众水平玩家与高水平玩家分别进行三角测试，99.5%概率认为AI识破率在(0,10%]间

(2024年，超参数科技)

游戏创作性能的优化

技术实现与效率提升

生成文字、图片、视频、音乐来缩减时间和资源消耗。

生产管线优化

快速识别游戏开发过程中的基本问题，并提出改善方法。AI最大的动能是不断产出改善工作流的最新方法，并不断融合进原本的工作流之中。

例如育碧提出的一种基于深度人工神经网络的光学动捕数据自动清洗和结算方案，不久后被网易AI Lab优化并以Vicon插件的形式部署在网易互娱和Quantic Deream的动捕工作流之中。



对早期游戏的优化和适配

材质、分辨率及动态画面增强

Nvidia RTX Remix: 通过AI分析原始贴图，可以生成高精度贴图，配合物理精确渲染增强凸显表现力，已作用于《半条命》、《最终幻想7》等游戏。

经典游戏的现代化改造

GameGAN模型：无代码重建游戏《吃豆人》；
WHAM：AI模型可以通过学习老游戏的机制，将其优化适配到现代硬件平台上。



World and Human Action Model

微软采用《Bleeding Edge》游戏数据训练出的游戏生成模型，数据中包含了：

- 玩家对3D空间中物体、角色和环境交互方式的“实践性理解”；
- 游戏物理和游戏如何对玩家的控制器动作做出反应。

商业化与运营： 个性化分析&效率增长

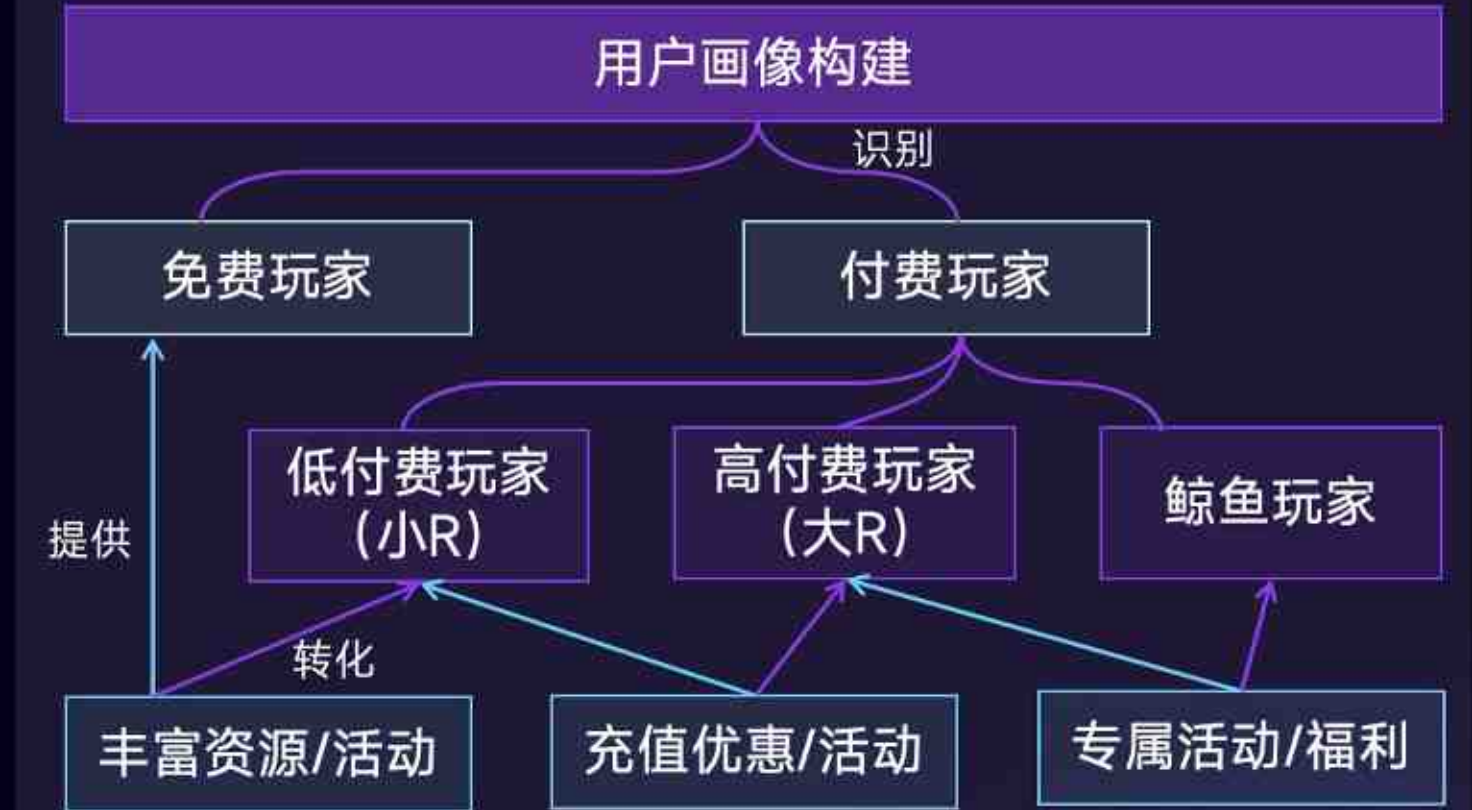


游戏内AI能够分析玩家行为、偏好从而实现高度个性化的内容推荐和营销；同时可以根据用户画像对玩家进行维护和转化；游戏外AI提供对投放和全球化市场有贡献的效率增长。

个性化内容与营销



玩家分析



智能运营与社群管理



全球智能化发行

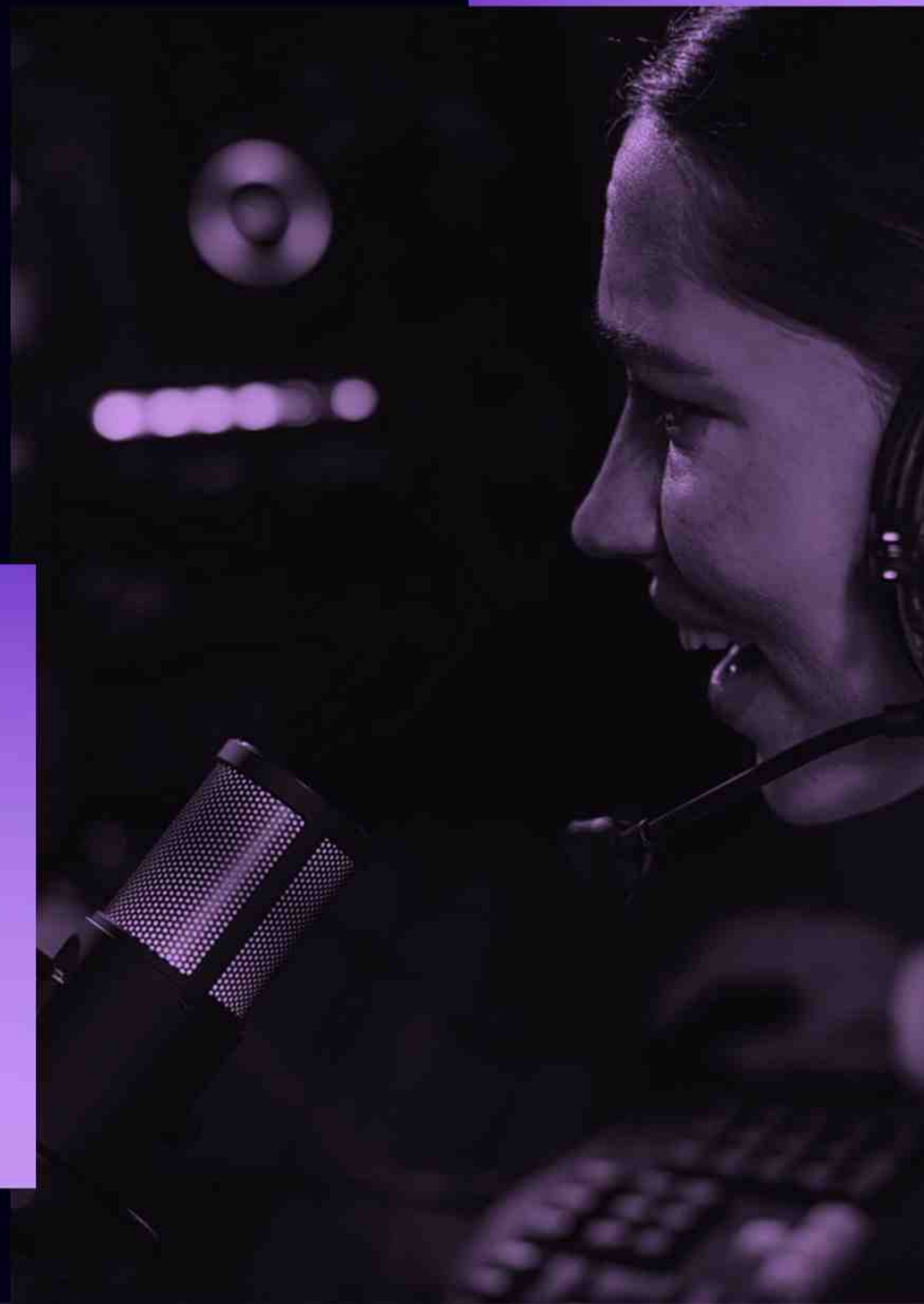


AI视频在所有投放视频中占比超过70%；AI翻译之后校准修改率在10%以下。此外，AI客户已经全面接入客服体系，问题解决率达到70%。
(2024年，三七互娱)

Part 3

AI改变的游戏体验

技术领域的AI研究始终追求极致的自治性，而游戏领域中却有所不同。简单的融合无法在提供长期有效的游戏体验、新奇的技术最终也只有技术爱好者问津。AI+游戏的路，必须走通核心玩法与玩家体验。



AI: 冲击与创造游戏类型



因此在与AI的结合中，

核心玩法设计

与深度受此影响的

玩家体验

是开发者必须要更关注的。

游戏AI的发展： 可控性与自治性的平衡



此外，一个显然的现象是：游戏中有生命力的AI和学界、业界对人工智能的探索是2个不同的概念，「更智能」的追求有时会打破沉浸式的体验。「永远无法战胜的对手」和「无所不能的队友」一样剥夺了玩家的游戏体验。

在游戏中实践AI时开发者一直在可控性（Controllability）和自治性（Autonomy）之间交锋，AI的自治性越强，可控性就越会下降，开发者设计好的「完美演出」就越可能毁于一旦。

所以从过去到未来，游戏的AI线路一直在其中寻找平衡，这也和现有游戏在AI利用上的探索目标是一致的。



以下是当前AI在游戏中进行探索的三种形式，其中，初级、中级和深层的融合是以生成式AI能力和游戏玩法、设计的结合程度来划分的。

初级融合： 游戏中的大模型接口



↓

如何承接玩法、剧情的推进是开发者更需要考量的

接口只能起到卖点的作用，且最终会随着AI对现实生活的渗透、新鲜感的褪去而作用越来越小。同时，因为不涉及游戏的核心玩法，在游戏内删去接口的载体也不会对游戏性有所影响。

比如RPG游戏中，NPC往往起到一个提供有效信息的作用，通过大语言模型丰富了NPC的对话之后，如何承接玩法、剧情的推进是开发者更需要考量的。

1 NPC聊天陪伴系统



《逆水寒》手游 大模型女团

为通义、文心、abab、Kimi和豆包大模型打造的AI娘形象，可以在游戏内回答问题。

2 内置智能游戏客服（玩家舆情监测）



《传奇世界》 玄玄老人

接入DeepSeek，可以通过自然语言对话获取准确的游戏信息和操作教程。

中级融合： 替代可智能优化的游戏环节

从游戏的核心架构和底层设计出发，思考什么样的游戏环节可以进行智能优化。通过这种设计方式，在这一层面AI替代了传统游戏中的某个「角色」。想象设计一个以团队合作为主的第一人称战术射击游戏，存在A、B两个对抗队伍：



Part2中我们也提到，前两种情况常出现在大型在线游戏中，以优化同时段玩家不足的问题。



例如用户生成内容（UGC）的实质是用户替代了游戏中预制的资产设计、关卡设计；加入AI后，AI+UGC是使用AI替代了实现设计想法的一些底层技术，如贴图、法则的生成。再或者，使用AI来替代游戏内在的评判逻辑，比如Game Master。

行业内这些层面都有广泛探索。

更深入的交互，合理游戏化AI:

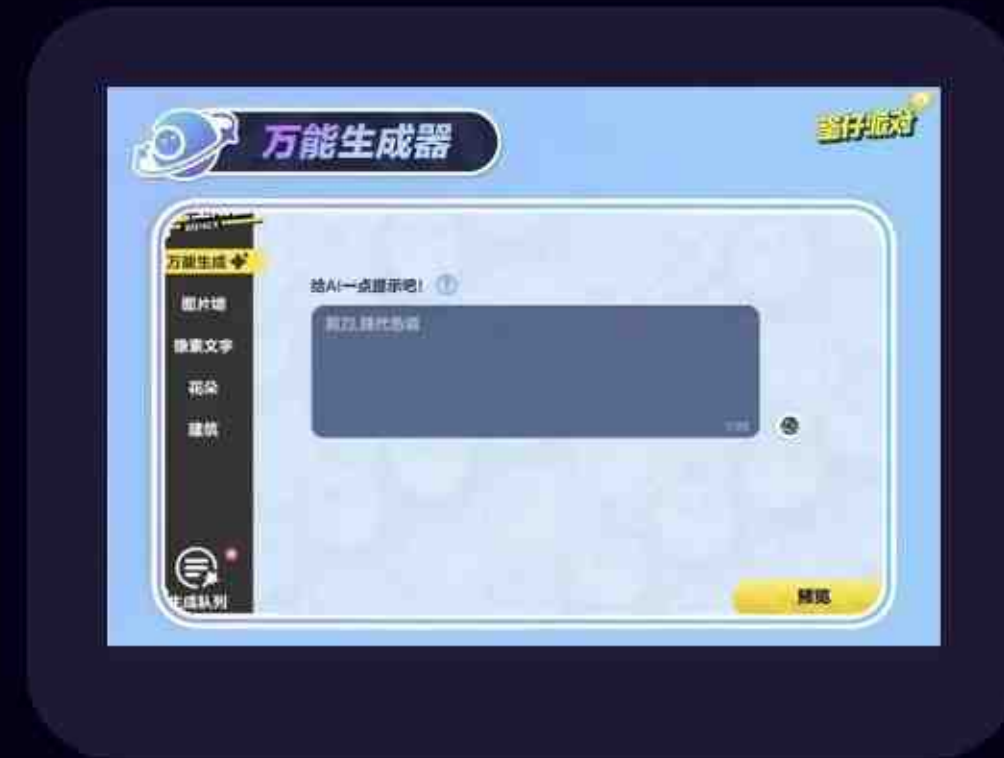
1 以AI为UGC的引擎



网易Eggy事业部

UGC是充分调动玩家创造力的玩法，这一游戏生态下降低门槛、提升工具的友好性很重要。

所宣传功能	实质	技术	目的
一键生成	有限制的模型生成	ModelMaker 3D模型生成	地图生成
万能生成器	文本驱动模型生成	ModelMaker 3D模型生成	地图生成
视频生成动作	动作捕捉、姿态提取、风格迁移	AnimationMaker PoseCap技术	玩家自创模型动作 视频传播
剧情动画编辑器	多种生成组合	AI识别、TTS等	玩家自创视频传播
蛋码	图形化编程功能	自然语言处理辅助	玩家自定义 迁移游戏玩法



蛋仔派对-万能生成器



蛋仔派对-蛋码



可以看到所有推出的新功能背后都有AI技术的扶持，而且是在游戏玩法驱动的情况下和UGC内容的要求紧密结合的，集成的AI功能首先减少了玩家的重复工作，其次进一步「放权」提高了创意的可能性。

更深入的交互，合理游戏化AI： 2 以AI为Game Master的游戏

游戏化语言	你可以说任何你想说的话来炫耀你的优点，或者诽谤你的对手	与各种角色展开对话，每个角色都有自己的喜好、风格和倾向
实质	提供一个与大模型沟通的接口	攻略向游戏，但选择权开放给AI
交互	输入文本	PVP

该作的Steam“AI披露内容声明”中写到：

「在我们的游戏中，人工智能（AI）扮演着核心角色，作为玩家可互动并尝试追求的非玩家角色（NPC）。AI会根据玩家的输入动态生成回应，创造出个性化且引人入胜的体验。」

「这种对AI的创新运用实现了独特的实时互动，使得每次游戏进程都因玩家的选择而截然不同。AI通过智能响应双方玩家的行为，强化了游戏的竞技性，确保玩法始终保持新鲜感和挑战性。」

在该游戏中，AI是游戏的GM。不是用户单一的向AI输入信息，而是把PVP对战中「裁判」一职交给更具有多样化反馈的AI，实现了玩法上的结合。

「《Pick Me Pick Me》是一款创新的双人在线竞技游戏，融合了约会的混乱与乐趣，带你进入虚拟的爱情战场。在这里，你将与对手一同展开对话比拼，与各种性格独特的AI角色进行互动！」



Optillusion工作室



PMPM-对话玩法



PMPM-玩家对战

开发者对AI能力和游戏逻辑的了解同样深入，才能在这一层次上创造出有亮点的AI游戏。



深层融合：直接生成可交互的3D游戏世界

通过直接生成，游戏中的一切都源自AI、出自生成。AI最有可能在这一层上衍生出新的游戏类型，但目前业界和学界的主要研究集中在营造3D场景上。

	Genie 2	World Labs	WonderWorld	Oasis
技术原理	基于视频扩散技术，逐帧生成，利用大规模视频数据集训练，具备物体交互、物理模拟等功能	从图片入手，估算景物深度和相对位置，构建贴近现实的3D模型，注重物理关系的精确呈现	采用FLAGS表示方法，通过单视图生成算法、基于几何的初始化和引导深度扩散技术，快速生成高质量3D场景	利用大规模数据集进行训练，构建大规模世界模型，但存在分辨率低、容易“遗忘”场景布局等问题
数据需求	<ul style="list-style-type: none"> 需要大规模视频数据集进行训练 数据需求较大 	<ul style="list-style-type: none"> 主要基于图片进行3D建模 数据需求相对较小 	<ul style="list-style-type: none"> 依赖单视图图像和文本引导的扩散模型 数据需求适中 	<ul style="list-style-type: none"> 需大规模数据集来构建世界模型 数据需求较大
生成能力	<ul style="list-style-type: none"> 能生成无限延展的3D世界 具有长程记忆 可模拟复杂动作和物理效果 	能生成高质量、贴近现实的3D世界 生成范围相对有限	能生成高质量、连贯、多样化的3D场景 生成速度快，实时渲染能力强	能构建大规模3D世界 在细节和一致性方面不足
交互性	支持实时互动，玩家或智能体的操作会影响世界动态变化	提供丰富的互动特效和动画功能，但用户移动范围受限	支持实时交互，用户可实时控制探索方向和场景内容	交互性相对较弱，主要侧重于世界模型的构建

没有游戏性的「游戏」



直接生成3D世界目前只能
是一项技术

技术不足

受制于计算复杂、长时间序列建模的衰减、显存占用等问题，所生成的世界很难在超过60秒的时间上连续；
核心体验时间一般只能延续10-20秒，强制重启无法形成长期探索的沉浸感。

缺乏游戏叙事

类型往往是开放式的沙盒世界，缺乏明确的游戏目标，也没有任务系统、成就解锁等机制用以引导玩家。

风格同质化

模型训练过程中往往会使用大量游戏实况视频，生成的场景风格、布局、NPC行为等与经典游戏有很强的「即视感」。

如同我们在之前提到的可控性与自治性之间的矛盾，与AI深度结合的类游戏项目集中在生成3D场景上，基本不存在游戏体验。游戏设计需平衡确定性（如关卡设计）与自由度，而当前的AI生成过度偏向后者。例如Wonderworld中，玩家可能因地形随机变化而卡关，破坏体验流畅性。

传统游戏通过难度曲线、资源分配等元规则调控体验，而AI生成环境缺乏此类底层设计。在当前这一时间点，直接生成3D游戏世界只可视作一项技术，不能称之为类游戏。但我们认为，每项技术都会在波及全行业时对某几类游戏产生突出影响，深层融合同样也将指向受AI影响最深的游戏类型。

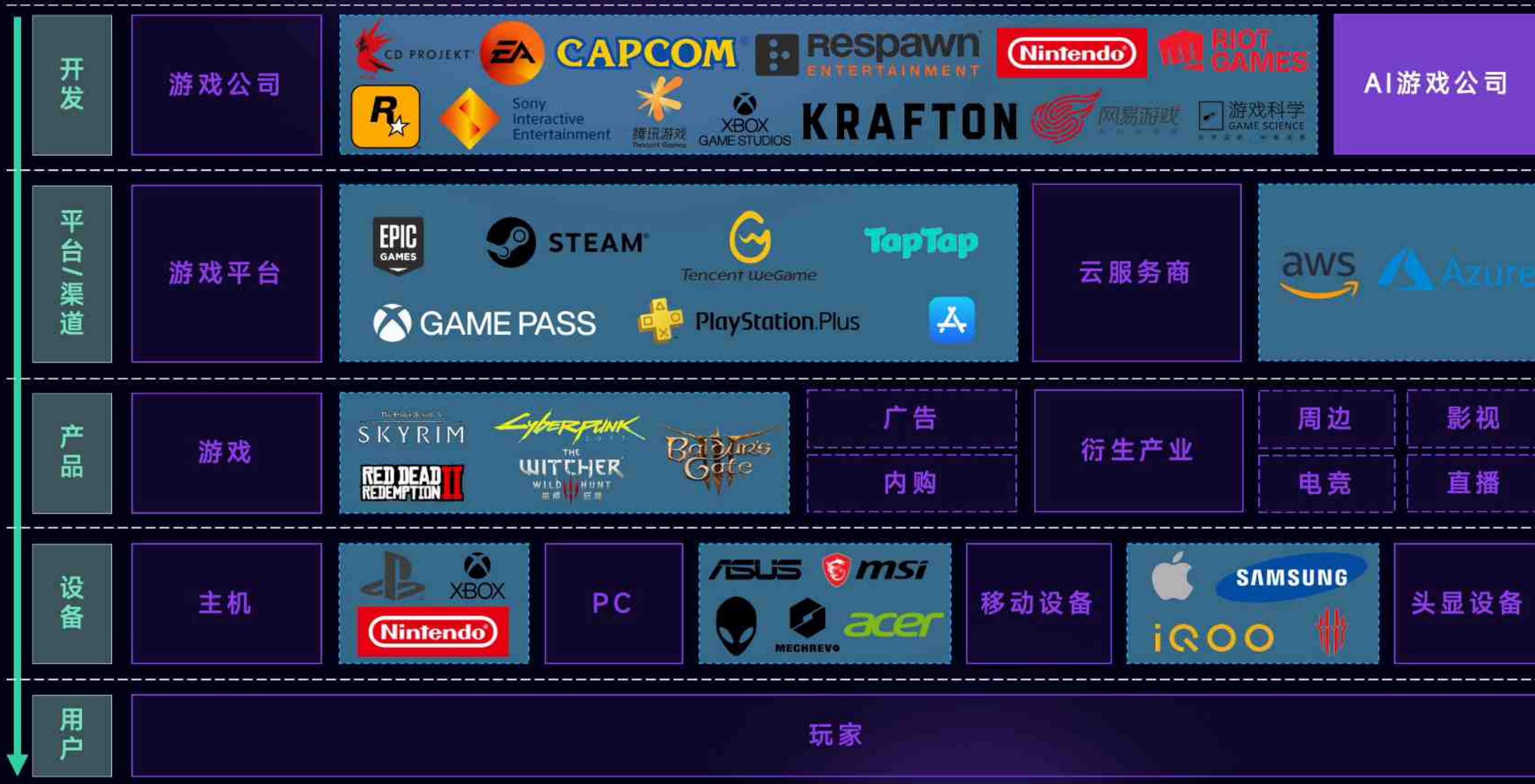


Part 4

游戏产业的AI变革

市场中，AI技术和AI策略在游戏产业/创业中的重要性被一再强调，超越了其他技术因素。总体来说，受长期和本轮AI技术的刺激，游戏产业发生了3大变革：首先，游戏市场的版图发生了变化；其次，作为核心产品的游戏本身被影响；最后，数据反映游戏市场受到了刺激。

游戏市场版图更新: 新生力量加入¹和传统玩家升级²



新生力量:

(1) 市场的主体毫无疑问是游戏公司，但科技公司正在以供应商的身份出现在产业链之中，同时AI产品/AI能力成为新兴的供应商品。

(2) 此外，2022年以来成立的一批游戏公司可以被称为AI游戏创业者。与前一代游戏出品方选择基于自身的游戏生产管线，配置对AI的投入不同，AI游戏创业者在市场上的第一次发声，往往是在AI能力上添加游戏想法。

传统升级:

市场的各层级及部分对技术更新产生剧烈反应——对技术的接纳速度、相关策略的响应速度和产品的更新速度非比寻常。

ANUTTACON

视觉生成/实时渲染	游戏场景和Stella模型、动作的实现
语音识别/语音合成	处理玩家的语音输入和Stella的对话输出
智能体架构	Stella需要根据玩家行为动态调整剧情并自主决策
硬件/算法优化	封闭测试基于iOS12及以上系统，在移动端实现PC级的AI渲染和文字、语音沟通

基础设定

实机演示中谈论到Stella个人话题，可以看出工作室为AI驱动的主角建立了清晰的人物设定，包括作为天体物理学学生的学习契机和成长经历。

任务目标

游戏设计了任务触发点和结局判定。玩家有较为清晰的主线任务，Stella被困外星球Gaia，最终目标是要从星球上脱困；同时和其他剧情类游戏相同，主线任务被拆解为阶段性的小目标，Stella需要先从含氧量不断降低的救生舱中逃出。

「名为Stella的女孩坠落到外星球，玩家是她唯一能联系的人，通过实时AI增强信息引导她推进故事，玩家的回应决定她是生存还是死亡。」

Whispers From The Star



Stella正使用语音生成回复玩家

AgentLive (喵吉托)



问题生成/回答打分	AI即时生成题目和评论；分数比拼与表情互动
语音生成	AI伙伴可以与你语音交流
AI NPC	玩家的AI伙伴/宠物

AgentLive始终在探索多种游戏玩法与AI的结合：

- 综合玩法的派对类游戏《萌爪派对》与前作《言灵计划》存在精神上的联系，即利用AI使玩家「拥有自己的宝可梦」。它强调玩家拥有的AI宠物会有独一无二的个性，能够与玩家建立情感羁绊。即使玩家离线，它们也会自主生活、四处探索并结交朋友。
- 轻肉鸽类型的农场游戏《圈圈圆圆圈圈》则可能搭载生成卡牌的功能，让每次对局的情况变幻莫测。

目前AgentLive的项目都具有完整的游戏性，并在探索哪些部分的AI优化可以丰富体验。

「打造你的圆圈农场，培育繁荣的栖息地，让古怪可爱的动物伙伴安居其中，成为终极牧场设计师！通过肉鸽类游戏机制掌握战略性棋盘布局，将荒芜的土地转变为一个充满奇思妙想的农场。」

Circle Mountain Circle Sea (圈圈圆圆圈圈)



「萌爪派对 是一个由AI驱动 的社交派对游戏社区。在这里，玩家不仅可以和朋友一起畅玩在线聚会游戏，还可以与AI伙伴一起享受游戏乐趣，建立深厚的羁绊。」

Paw Party (萌爪派对)

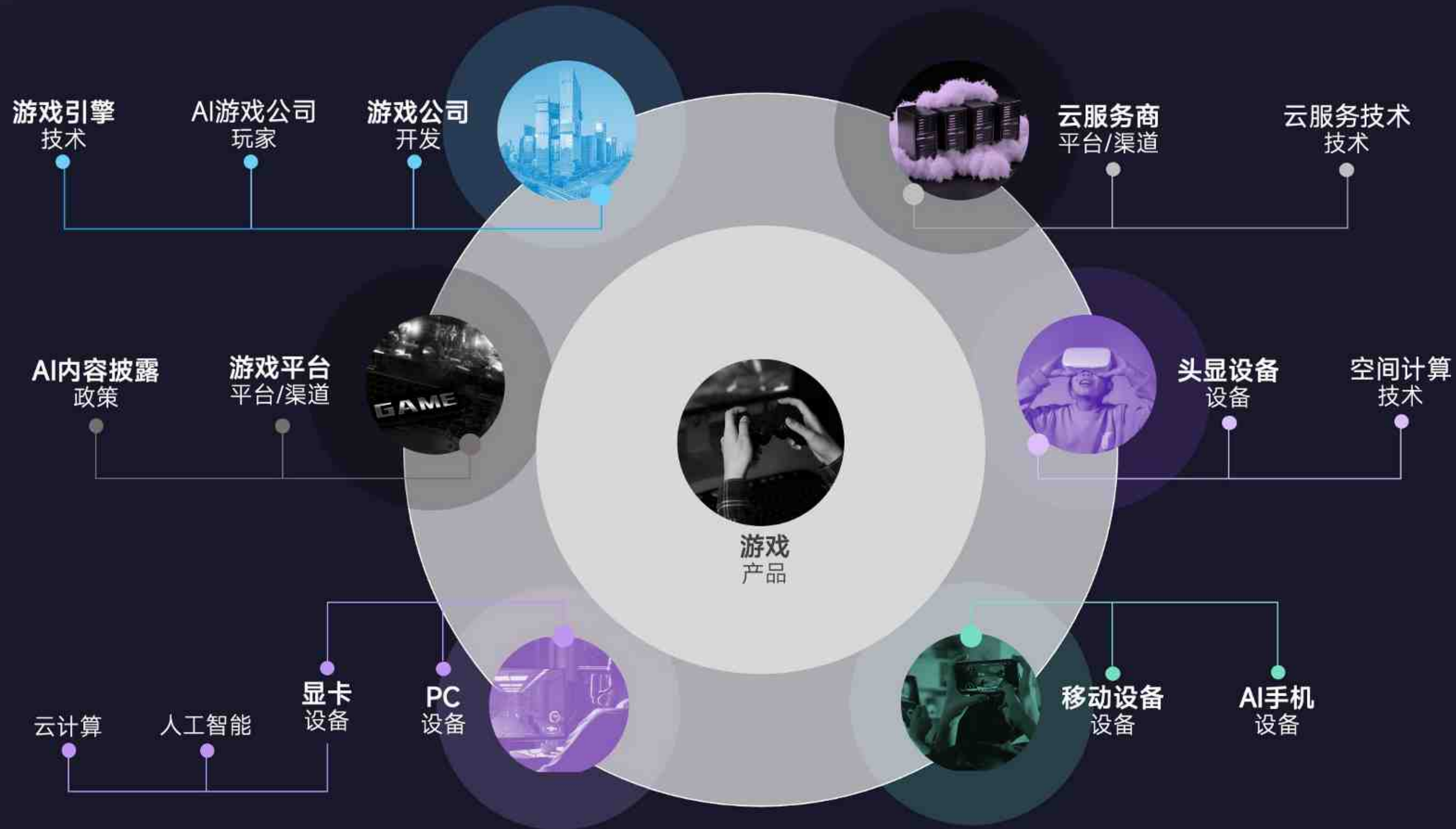


AI即时生成问题



「Generate」创建测试卡牌

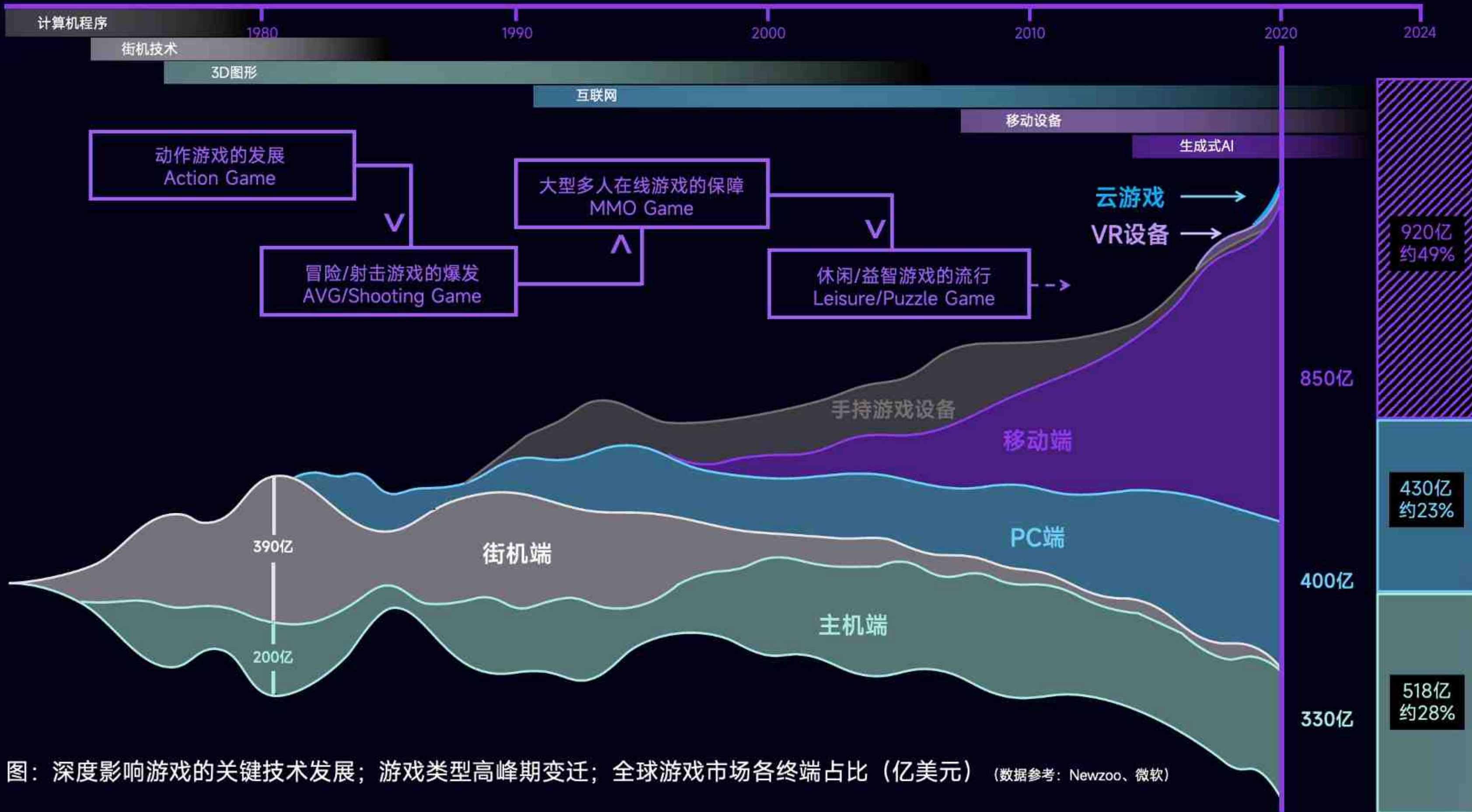
边界扩展进一步实现 3



以游戏产品的更新为核心，开发、平台及渠道和设备都在市场中开拓AI版图，市场泛起的活力吸引新一轮玩家加入。

同时市场的组成部分再与AI技术、需求相结合，延伸出游戏引擎、云服务、空间计算等技术更新；AI手机、游戏机等设备更新；AI内容披露等策略更新；对云计算、人工智能等领先市场也形成正向激励。

AI搅动游戏市场： 更受冲击的游戏类型

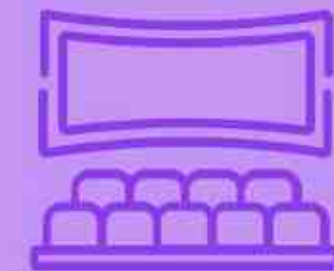
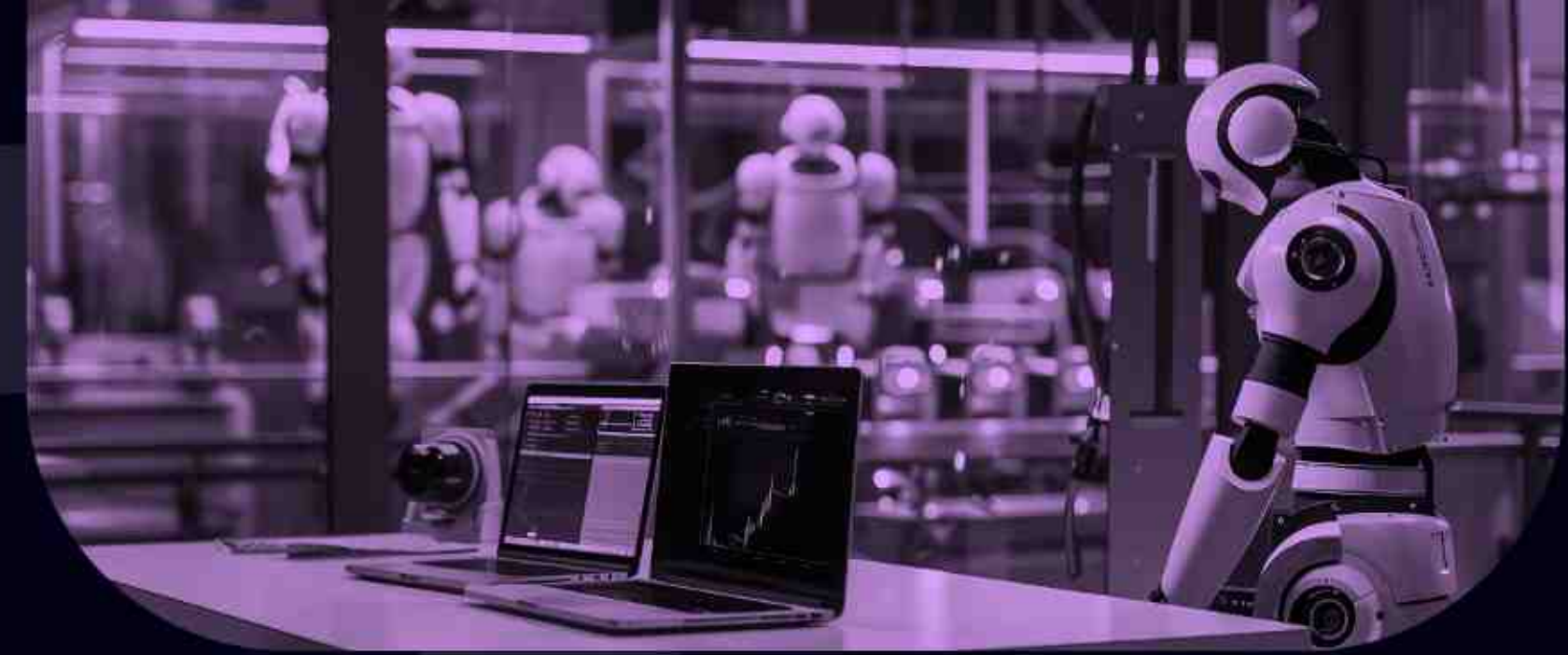


游戏的高峰期及游戏类型的发展深受技术适配性的影响

- 1972年《Pong》这款简单的数字乒乓球游戏取得了巨大的商业成功，被认为是街机游戏行业的开端；
- 90年代3D图形技术进入快速发展和普及，同时3dfx推出的Voodoo 1显卡使PC能够实现真正的3D图形，PlayStation和Nintendo 64的发布将真3D游戏带入主流家庭市场，使得射击游戏在90年代后期爆发；
- 互联网（Internet）技术重塑了MMO游戏的体验感；
- Apple iPhone的发布彻底改变了智能手机市场，移动设备的普及让休闲/益智游戏在21世纪10年代流行。

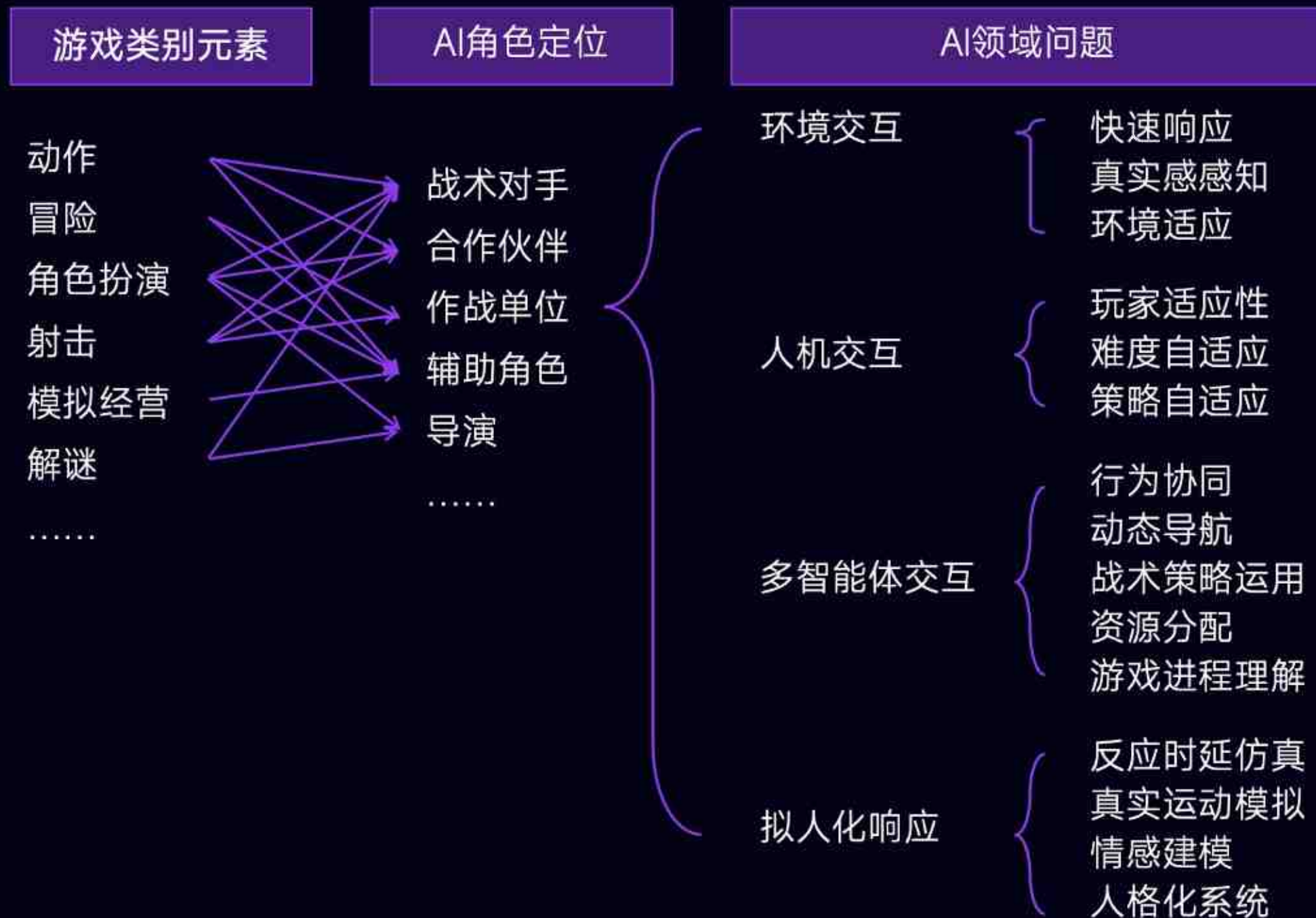
更受冲击的游戏类型： 强调自选&生成，迎来解构+创新

因此核心的技术和主要游戏设备的更新都有催生不同类型游戏高峰期的可能性。技术的特性和游戏类别元素的结合是有迹可循的。

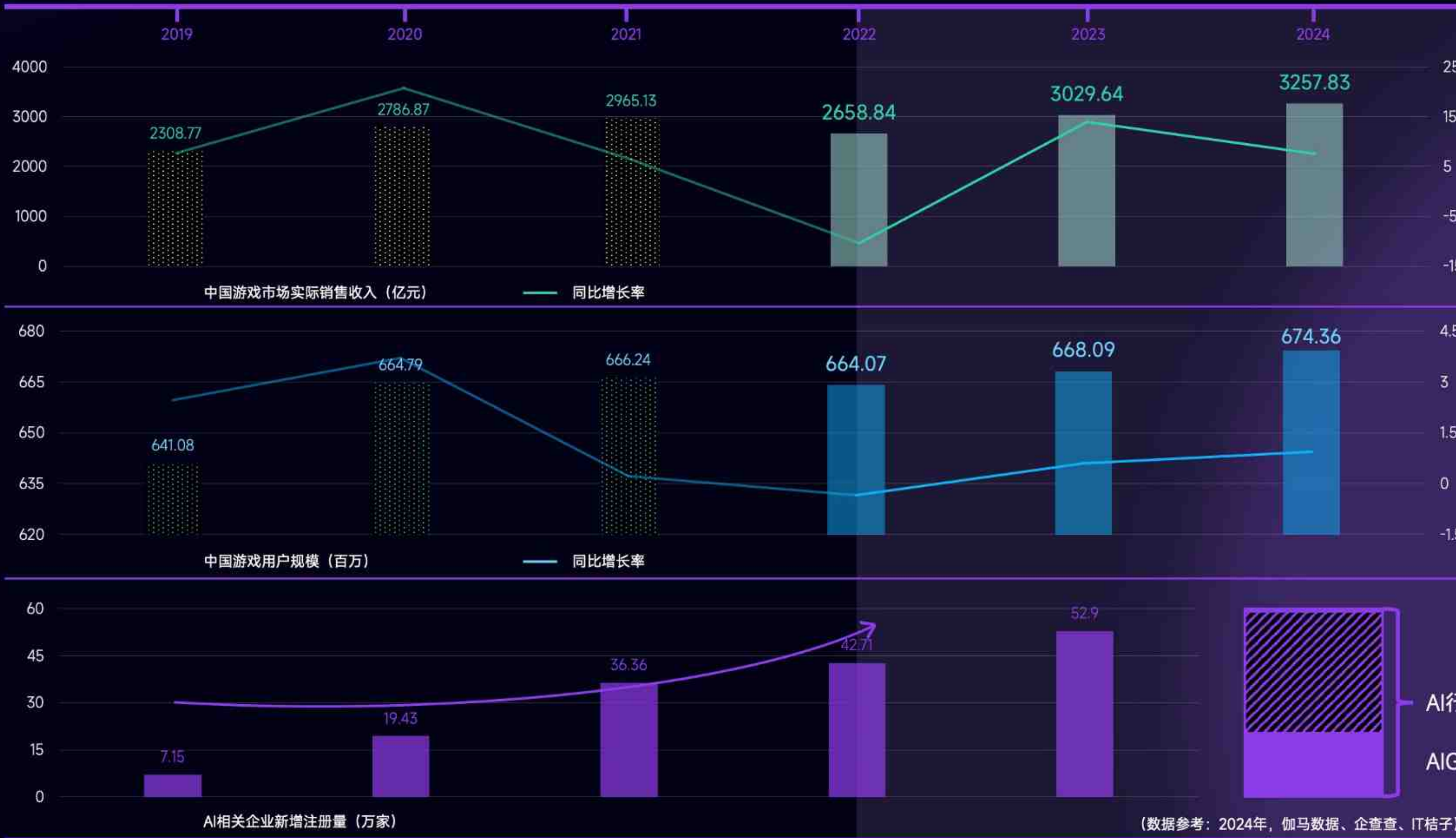


伴随着VR等新终端的出现，沉浸式的技术风格强调的是完整、复合的体验。
核心设计提倡**玩家自选冒险**的RPG、AVG特性游戏，和实现**内容实时生成**的休闲/益智/策略特性的游戏受影响最大。

↓
特别的是，生成式AI为核心技术的时期，游戏的类型将得到进一步的解构+创新，但核心仍会随着「技术适配性」和「玩家体验」的双重需求展开。



AI&游戏市场： 回暖期与AI爆发期



在中国游戏营收前50的企业中，超过60%已明确布局AIGC技术。

2024年我国游戏市场实际销售收入为3257.83亿人民币，同比增长7.53%，市场仍在持续回暖。同时自2022至2024年，中国游戏用户规模达到6.74亿，同比增长0.94%。2022年起，中国游戏市场进入一个稳健且持续向上的增长期。

我们多次强调游戏与生成式AI爆发在时间上的「巧合性」，并充分论证了技术层面上的相关性。

不能忽视的是，AI技术在游戏开发中的逐渐落地为游戏行业提供了新的增长空间，成本、效率和质量同步上升的时代已来。



游戏市场3500亿元大关，需要全新的增长机会。利用AI释放更多创意和生产力，进而将其投入游戏市场，是显然的策略。



量子位智库公众号



扫码获取更多报告

AI + 游戏 产业变革研究报告

● 研究人员

郑钰瑶
量子位智库分析师
yuyao@qbitai.com

● 量子位智库出品

量子位旗下科技创新产业链接平台。致力于提供前沿科技和技术创新领域产学研体系化研究。