甲子光年
JAZZYEAR

2024

AIGC视频生成：走向AI创生时代

视频生成的技术演进、范式重塑与商业化路径探索

出品机构：甲子光年智库
研究团队：张一甲、宋涛
发布时间：2024.03

*刘瑶、小麦对本报告亦有贡献。

甲子光年
JAZZYEAR

“一类人有一类人原力觉醒的方式。

物理学家想学习上帝；

数学家想反抗上帝；

哲学家认为自己就是上帝；

生物学家想造上帝的反……

工程师说都不用，我们再造一个。”

——《甲小姐：站在两个世界之间》甲子光年 2017.10

目录

CONTENTS

甲子光年
JAZZYEAR

1.1 Sora让文生视频迎来“GPT-3”时刻

OpenAI发布文生视频模型Sora，堪称视频生成领域的“GPT-3”时刻

春节假期甚至还未结束，Sora已引发全民关注

“Sora”一词在微信指数及百度指数的关注度快速上升

“炸裂”视频效果成为讨论热点

- 效果逼真：普通人一时难以分辨
- 时长感人：60秒高清视频生成
- “百万” 剪辑：堪比专业的镜头语言
- 多模态：文字、图片、视频皆可生成视频

数据来源：百度指数、微信指数，甲子光年智库整理，2024年；

1.2 Sora的展现效果

Sora模型展现自身超强视频生成及剪辑能力，超出其他竞品一个段位

能力项		Sora	其他模型
基本视频生成	视频时长	60秒	20秒以内
	视频长宽比	1920*1080之间的任意尺寸	固定尺寸比例，例如16:9，9:16，1:1等
	视频清晰度	1080p	部分upscale后达到4k
多模态生成	语言理解能力	强	弱
	文本生成视频	支持	支持
	图片生成视频	强	支持
	视频生成视频	支持	支持
视频编辑	文本编辑视频	支持	支持
	扩展视频	向前/向后扩展	仅支持向后
	视频的无缝连接	支持	不支持
独特模拟能力	3D一致性	强	弱或不支持
	远程相干性和物体持久性	强	弱
	世界交互	强	弱
	数字世界模拟	支持	不支持

其他模型情况

模型	Gen-2	pika1.0	Stable Video Diffusion	Emu Video	W.A.L.T
开发团队	Runway	Pika Labs	Stablity AI	Meta	李飞飞及其学生团队、谷歌
时间	2023年11月	2023年11月	2023年11月	2023年11月	2023年12月
长度	4-18秒	3-7秒	2-4秒	4秒	3秒
分辨率	768*448，1536*896，4096*2160	1280*720 2560*1440	576*1024	512*512	512*896
是否开源	非开源	非开源	开源	非开源	非开源

Sora的语言理解能力更强，可将简短的用户提示转换为更长的详细描述

Sora还可以生成图片，最高可达到2048*2048分辨率

Sora通过插帧技术，实现完全不同主题和场景构图的视频之间的流畅自然的过渡效果

Sora可生成具有动态摄像机运动效果的视频，随着摄像机的移动和旋转，人和场景元素在三维空间中保持一致移动

Sora可以对短期和长期依赖关系进行建模，保持各个主体的时空连贯性和一致性

Sora以简单的方式模拟影响世界状态的行为，比如一个人吃完汉堡可以在上面留下咬痕

Sora还能够模拟人工过程，比如视频游戏，同时通过基本策略控制玩家，同时以高保真度渲染世界及其动态

数据来源：甲子光年智库梳理，2024年；

1.2 Sora的展现效果

大模型训练的“暴力美学”在视频生成领域再次涌现卓越特性

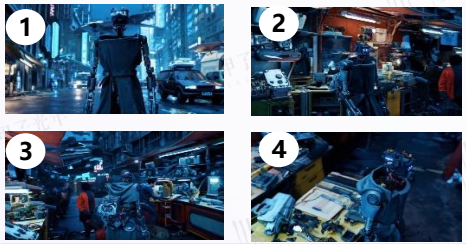
- OpenAI发现视频模型在大规模训练时表现出许多有趣的“涌现”能力，使Sora能够从物理世界中模拟人、动物和环境。值得一提的是OpenAI官网所说的“they are purely phenomena of scale”——它们纯粹是“规模现象”，这再一次验证了“暴力美学”。

文/图像/视频生视频的功能

3D一致性：确保景别切换时运镜的连贯

以上四个镜头由远及近，保证了视频镜头中人和场景的一致性，是其他AI生成视频中少见的。

远程相关性和物体持久性



以上四个镜头在同一视频中生成，包括机器人的多个角度。

与世界互动：Sora有时可以用简单的方式模拟影响世界状况的动作

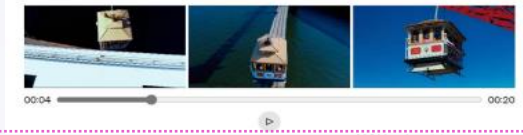
画家可以在画布上留下新的笔触，并随着时间的推移而持续存在。

模拟数字世界

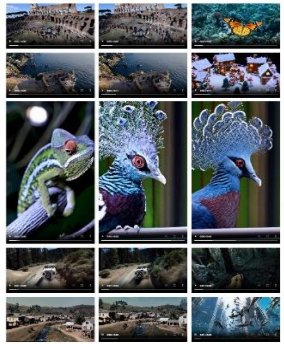
例如，Sora可以同时通过基本策略控制《我的世界》中的玩家，同时以高保真度渲染世界及其动态。

视频剪辑功能

基于时空双维度的视频扩展



不同主题场景视频的无缝连接



一键进行风格渲染

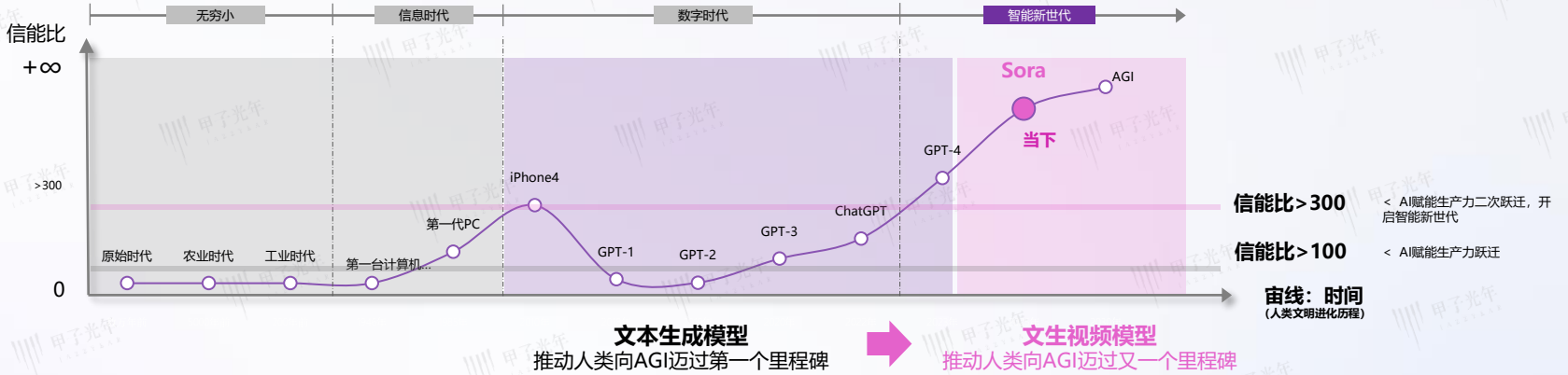
数据来源：公开资料，Open AI技术文件，甲子光年整理，2024年；

1.3 Sora的出现意味着AGI的又一个里程碑时刻

Sora意味着scaling law（规模法则）再次验证，推动文生视频进入“GPT-3”时刻

- **Scaling law（规模法则）的再次验证：**虽然Sora并不十全十美，但它通过scaling law和原有模型拉开了差距，为视频生成领域提供了另一条可以走通的路线，推动行业进入全新的阶段。
- **文生视频的“GPT-3”时刻：**从发展阶段类比，Sora更像文本模型的GPT-3时刻。ChatGPT让人类看到实现AGI的雏形，Sora让实现AGI的目标又进一步。

智能新世代：Sora向AGI再进一步



备注说明：
信能比，是甲子光年智库发明的概念，反映单位能源所能驾驭的信息量。信能比通过单位时间内产生/传输/使用/存储的信息量除以单位时间内所消耗的能源量计算得出，反映单位能源所能调用的信息量水平的高低。
信能比可以体现数据智能技术的先进性和能源效率的高效性：它能够反映整个社会数字化、智能化水平的高低；它能体现能源体系的可持续发展能力；它能反映生产力的高低和生产效率的提升；它能体现社会经济发展的先进性、创新性、可持续性。

1.4 Sora开启“明牌游戏”，推动AIGC应用时间轴进一步被压缩

历史反复表明，一旦先行者模式验证，后来者整体的应用进程时间表将加快

- 先行者往往要花费大量时间精力试错，一旦模式跑通，“明牌游戏”就开启了。后来者会有更好的参考系和聚焦方向。ChatGPT后续的文本生成模型进展就说明了这一点。
- 过去一年，AI文本生成和图像生成相继走向成熟，Sora的发布意味着视频生成应用走向成熟的时间比原先预计的更早出现，AIGC已经加速迈入视频生成阶段。
- 对此，甲子光年智库更新了生成式AI技术的成熟应用进程时间表。2024年可实现根据文本提示生成初版短视频，2025年有望实现根据文本生成初版长视频，并在视频制作环节真实使用落地。

图1：AIGC用户偏好使用的大模型产品类型

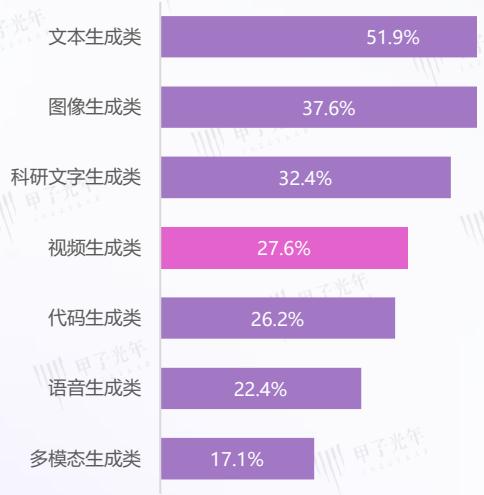


图2：生成式AI技术的成熟应用进程时间表

大模型成熟难度：初级尝试　接近成熟　成熟应用

领域类型	2020年之前	2020年	2022年	2023年	2024年E	2025年E	2030年E
文本领域	诈骗垃圾信息识别 翻译 基础问答回应	基础文案撰写 初稿	更长的文本 二稿	垂直领域的文案 撰写实现可精调（论文等）	终稿，水平接近人类平均值	终稿，水平高于人类平均值	终稿，水平高于专业写手
代码领域	单行代码补足	多行代码生成	更长的代码 更精确的表达	支持更多语种 领域更垂直	根据文本生成初版应用程序	根据文本生成初版应用程序	根据文本生成终版应用程序，比全职开发者水平更高
图像领域			艺术 图标 摄影	模仿（产品设计、建筑等）	终稿（海报设计、产品设计等）	终稿（产品设计、建筑等）	终稿，水平高于专职艺术家、设计师等
视频/3D/游戏领域				视频和3D文件的基础版/初稿	根据文本生成初版的短视频	根据文本生成初版的长视频，并实际应用于制作环节	AI版Roblox可依个人梦想定制的游戏与电影

数据来源：红杉资本，甲子光年智库，2024年；

1.5 Sora验证视频生成的新技术范式

Sora的出现意味着视频生成的DiT技术路线得到有力验证

- 视频生成技术路线在过去主要有两条，一条是基于Transformer的路线，以Phenaki为代表，第二条是Diffusion Model（扩散模型）路线，该路线在2023年是主流路线，诞生了Meta的Make-A-Video、英伟达的Video LDM，Runway的Gen1、Gen2，字节的MagicVideo等代表性产品。
- **Sora的发布，对Transformer + Diffusion Model（DiT）路线进行了成果瞩目的验证。**

图1：AIGC视频生成的技术演进路径

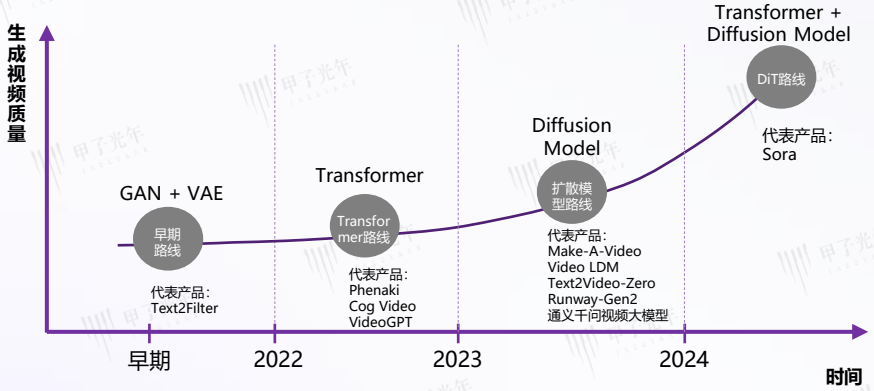


图2：Sora技术优势与竞品的对比情况

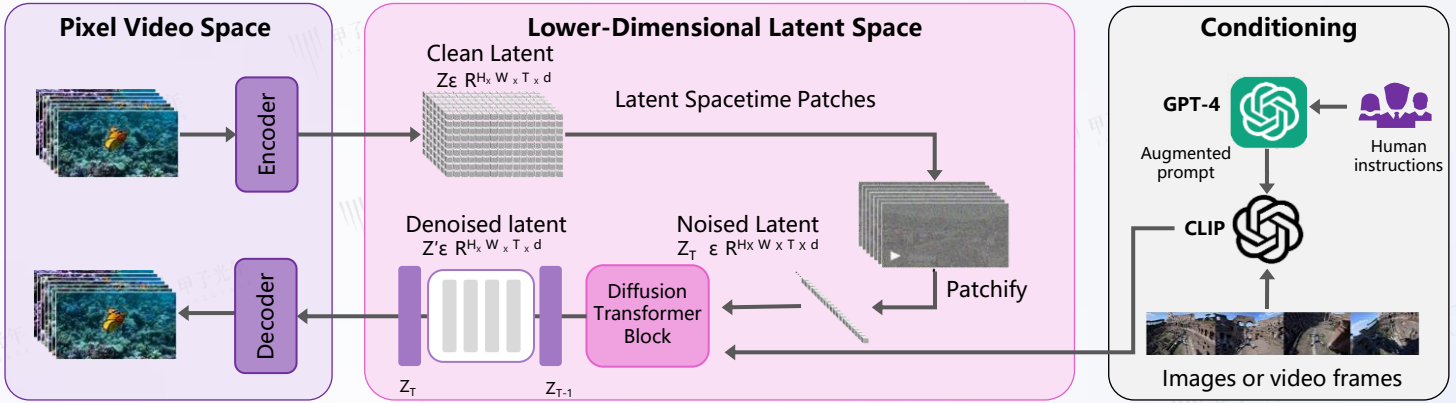
数据来源：国盛证券、硅基研究室，甲子光年智库梳理，2024年；

1.6 Sora的技术原理

Patch（时空编码思路）+DiT（Diffusion和Transformer模型的结合）+Scaling Law（规模效应）

- Sora模型将视频压缩到低维空间（latent space），并使用时空补丁（Spacetime latent patches）来表示视频。这个过程类似于将文本转换为Token表示，而视频则转换为patches表示。Sora模型主要在压缩的低维空间进行训练，并使用解码器将低维空间映射回像素空间，以生成视频。
- Sora使用了diffusion模型，给定输入的噪声块+文本prompt，它被训练来预测原始的“干净”分块。
- Sora是diffusion transformer，而transformer在各个领域都表现出显著的规模效应。

图：业内推测出的Sora技术架构图

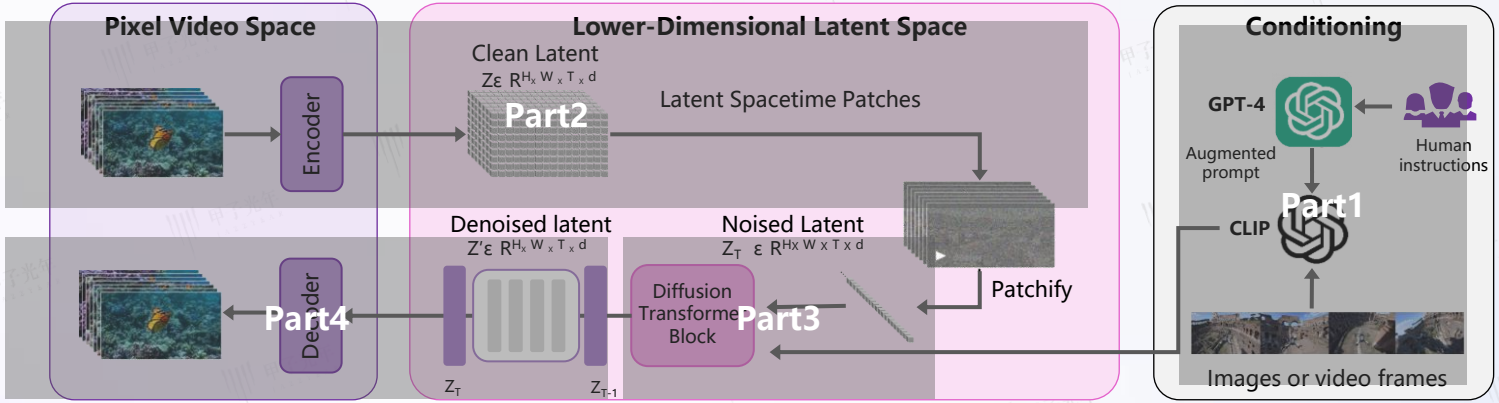


1.6 Sora的技术原理

Sora模型的实施路径可拆分为四个部分

- Sora模型的实施路径有四个核心部分：
  - Part1：使用文生图模型（DALLE 3）把文本和图像对<text, image>联系起来。
  - Part2：视频数据切分为Patches，通过编码器压缩成低维空间表示，解决了时间和空间两个维度的注意力交互（patch化是训练生成式模型的一个非常scalable和高效的图像/视频表征形式）。
  - Part3：Diffusion Transformer。
    - Denoising Diffusion Probabilistic Models (DDPMs)：通过逐步添加噪声来模拟数据分布，然后学习逆向过程去除噪声，以生成新的数据。DiT是DDPM在图像生成中的应用。
    - Latent Diffusion Models (LDMs)：使用变分自编码器将图像压缩到低维表示，然后在低维空间中训练DDPM。这样可以降低计算成本，并使DiT成为基于Transformer的DDPM的适用框架。
  - Part4：DiT生成的低维空间表示，可通过解码器恢复成像素级的视频数据。

图：业内推测的模型实施路径解析



1.7 Sora的局限性

Sora仍存在三大方面局限性，会短期制约其商业化、规模化应用

技术局限性

物理现实主义的挑战

Sora对复杂场景中物理原理的处理不一致，导致无法准确复制因果关系，偶尔会偏离物理合理性。例如物体的不自然变换或对刚性结构的不正确模拟，导致不切实际的物理交互。此外，描绘复杂的动作或捕捉微妙的面部表情是模型可以增强的领域。以上，导致Sora现阶段更擅长幽默的结果而非严肃的内容。

时空连续性的挑战

Sora生成的视频中可能会出现物体无缘无故消失或出现， Sora有时会误解给定提示中与物体的放置或排列相关的指令，从而导致方向混乱。此外，它在保持事件的时间准确性方面面临挑战，可能会导致预期时间流发生偏差，影响生成内容的可靠性和连贯性。

人机交互的限制

Sora生成视频的随机性很强，类似人类的“做梦”，用户可能很难精确指定或调整视频中特定元素的呈现，这限制了Sora在视频编辑和增强方面的潜力，也让Sora在长视频应用中面临挑战。

伦理合规性

数据合规性

可能涉及到他人的隐私信息，例如在视频中出现的人物、场景或个人数据等。未经授权或未经允许的情况下，生成和传播涉及他人隐私的虚假视频可能导致隐私泄露问题。

版权风险

生成的视频内容可能涉及到他人的知识产权/版权，如果未经授权使用他人的作品或内容进行生成，就可能涉嫌侵犯他人的版权权益，引发版权纠纷或法律诉讼。

AI安全问题

可能导致深度伪造视频的增加，即利用技术手段在视频中替换现实中的人物或场景，使得伪造的视频无法通过肉眼识别真伪，给社会带来信任危机和安全隐患。确保Sora的输出始终安全且公正是一项主要挑战。

普适制约性

经济账与成本问题

OpenAI自从推出文本生成大模型再到推出视频生成大模型，一直没有解决商业化问题，大模型的训练需要较高成本投入，如何算好经济账是影响规模化应用的前提。

需要依赖高质量、大规模的视频数据

Sora的训练路径需要依赖庞大规模的视频数据，并需要较高的数据标注、合成能力，后期的迭代升级会受到底层训练数据的影响与限制。

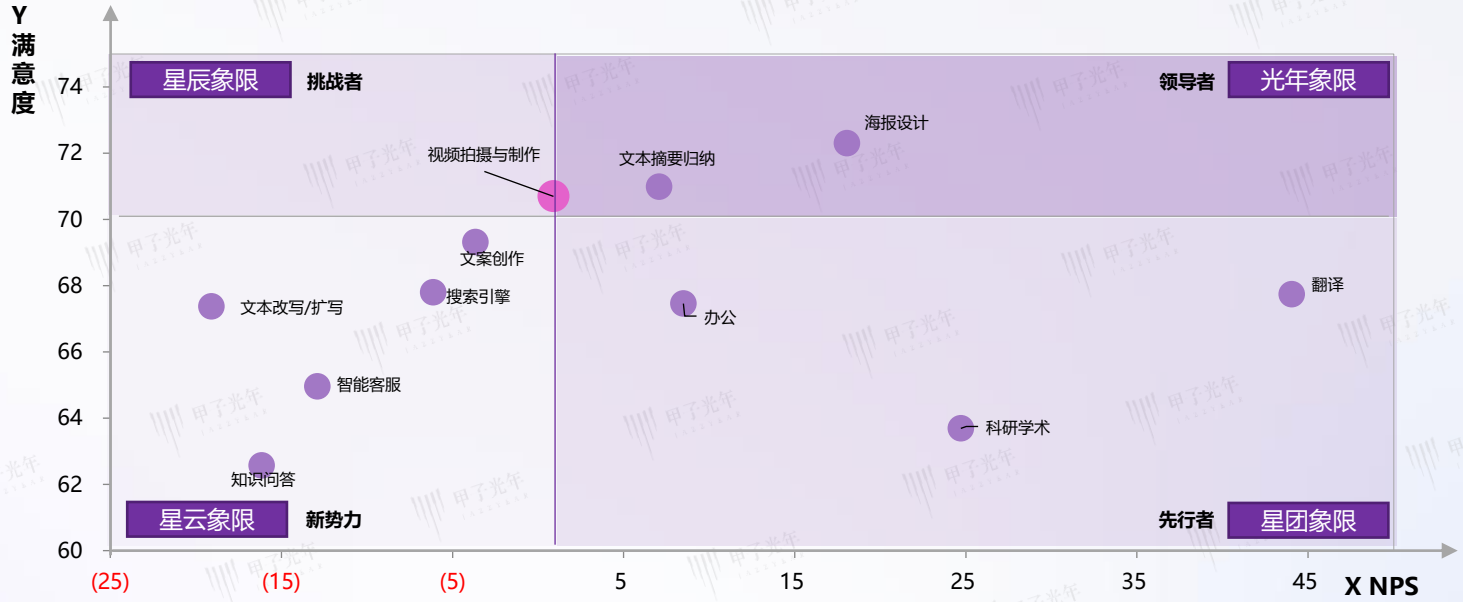
算力瓶颈问题

Sora视频模型的训练需要很高的算力支撑，如何平衡算力、成本、能源消耗等关系是值得关注的制动因素，也将是影响Sora大规模商业化运营的瓶颈。

1.7 Sora的局限性

视频生成处于用户满意但不推荐象限，说明现有视频生成工具虽然惊艳，但尚无法支持实际工作

甲子星空坐标系： 用户对AIGC产品不同应用场景的满意度与NPS值



1.8 Sora引发的世界模型之争

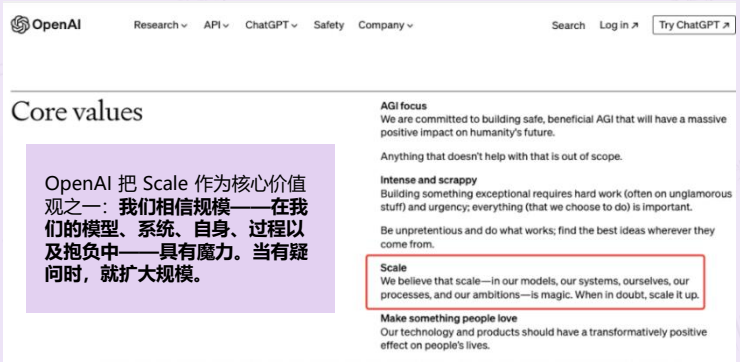
Sora被OpenAI定义为“世界模拟器”，由此引发了世界模型的实施路线之争

正方： OpenAI 把它定义为一个“世界模拟器” (World Simulator)

OpenAI 表示：“Sora是能够理解和模拟现实世界模型的基础，我们相信这种能力将成为实现 AGI 的重要里程碑。”

英伟达高级研究科学家 Jim Fan 更是直接断言：“Sora是一个数据驱动的物理引擎，是一个可学习的模拟器，或世界模型。”

OpenAI 是自回归生成式路线（Auto-regressive models），遵循“大数据、大模型、大算力”的暴力美学路线。从 ChatGPT 到 Sora，都是这一思路的代表性产物。

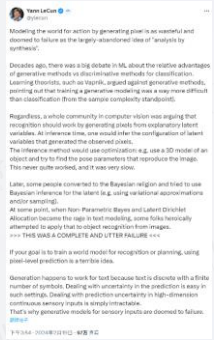


OpenAI 把 Scale 作为核心价值观之一：**我们相信规模——在我们的模型、系统、自身、过程以及抱负中——具有魔力。当有疑问时，就扩大规模。**

VS

反方：仅根据文字提示生成逼真的视频，并不代表模型理解了物理世界！

Yann LeCun，图灵奖获得者和Meta首席科学家，最近表达了对Sora的生成式技术路线的质疑，并指出该路线可能面临失败的风险。



Yann LeCun认为，仅凭文字提示生成逼真视频并不代表模型真正理解物理世界。他指出生成视频的过程与基于世界模型的因果预测完全不同。

在2月19日的发文中，他再次反驳了通过生成像素来建模世界的观点，认为这种方法是浪费，就像被广泛抛弃的“通过合成来分析”的想法一样，注定会失败。

Yann LeCun认为文本生成之所以可行是因为文本本身是离散的，具有有限数量的符号。在这种情况下，处理预测中的不确定性相对容易。而在处理高维连续的感觉输入时，基本上不可能处理预测的不确定性，这也是为什么针对感觉输入的生成模型注定会失败的原因。

Keras 之父 François Chollet 也持有类似观点。他认为仅仅通过让 AI 观看视频是无法完全学习到世界模型的。尽管像 Sora 这样的视频生成模型确实融入了物理模型，问题在于这些模型的准确性及其泛化能力——即它们是否能够适应新的、非训练数据插值的情况。

Artificial Intuition的作者 Carlos E. Perez认为，Sora并没有真正学会物理规律，只是表面上看起来像学会了，就像几年前的烟雾模拟一样。

知名 AI 学者、Meta AI 研究科学家田渊栋也表示，关于 Sora 是否有潜力学到精确物理（当然现在还没有）的本质是：为什么像“预测下一个 token”或“重建”这样简单的思路会产生如此丰富的表示？

目录

CONTENTS

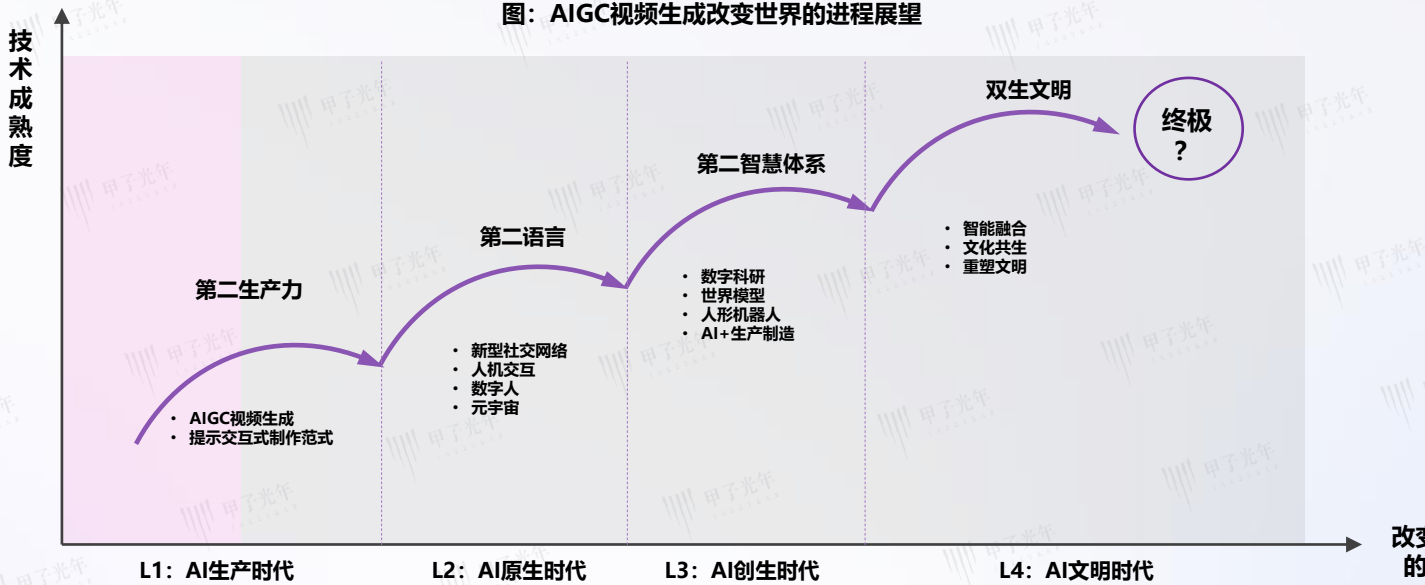
甲子光年
JAZZYEAR

2.1 走向AI创生时代，改变世界刚刚开始

AIGC视频生成开启AI创生时代，重塑视频产业链仅仅是第一步

- 甲子光年智库将AIGC视频生成对世界的影响分为如下四个阶段：
- **L1：AI生产时代/AI工业时代。**AIGC引发内容相关产业的生产力变革，视频产业将是首先被重塑的领域，AI驱动内容领域迎来“工业革命”，大幅提升内容生产效率，**形成第二生产力。**
- **L2：AI原生时代。**AIGC将进一步引发生产关系变革，引发角色与分工的变迁。**视频成为人类信息表达的第二语言，人类语言将告别“词不达意”阶段，**重塑人、内容、机器间的生产关系与交互关系。在这一阶段，**AI渗透率将无限逼近人类在数字世界的生产活动行为边界——人在数字世界可以做的事情，AI都可以做。**
- **L3：AI创生时代。**AI与物理世界进一步融合，逐渐渗透逼近人在物理世界的生产活动行为边界。从AI for science到生产制造，从人形机器人到世界模型，AI将逐渐突破人类为主语的创造范畴，**世界模型将创造人类智慧之外的“第二智慧体系”。**
- **L4：AI文明时代。**AI推动人类认知重塑，开启AI文艺复兴。AI会深度参与人类的物理世界和心灵世界，人类智慧与AI将互相影响、共同进化，**人类文明进入“双生时代”，**形成**“AI的归AI，人类的归人类”**的有序分工和共生模式。

图：AIGC视频生成改变世界的进程展望



数据来源：甲子光年智库，2024年；

2.2 L1-AI生产时代：“拍扁”视频制作链条，开启“提示交互式” 新范式

基于AIGC视频生成工具的提示交互式视频制作范式将重塑传统视频制作流程

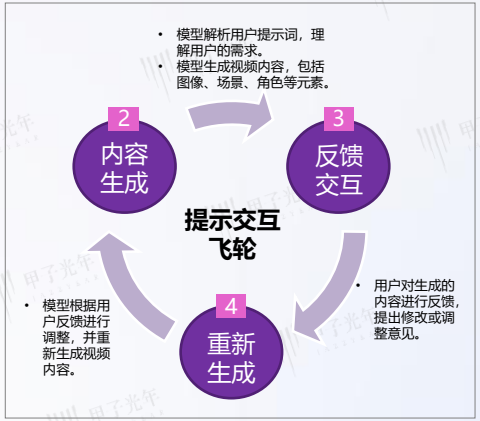
- AIGC视频生成工具可对视频生产流程进行重塑，由传统视频制作范式进化到“提示交互式”新范式。
- “提示交互式”新范式相比传统范式具有三方面的重塑：
  - **“拍扁”制作过程：**传统视频制作流程涉及多个阶段和专业团队的合作，耗费大量时间和资源；而AIGC视频生成可将视频生成、剪辑、后期等环节集于一体，仅需要输入提示词即可生成视频，省去了很多繁琐步骤，尤其可将摄影、素材收集、后期等环节取消或缩短。
  - **提升创意和剪辑自由度：**传统视频制作通常由制作团队提出创意、编写剧本，受人的能力局限；提示交互式视频生成用更可视化的方式激发创作者想象力，支持创作者调用AI模型探索每个镜头的无限可能，而且剪辑过程可以随时发生。
  - **节省制作成本和时间：**传统视频制作流程需要投入较多的人力、物力和时间，而提示交互式视频生成的流程成本和时间较少，可在提示交互的飞轮中迭代生成最终满足需要的内容。

传统视频制作执行流程

1. **策划**
   - 制定视频内容的概念和主题。
   - 编写剧本或提纲，规划视频的情节和内容。
   - 确定目标受众和传达的信息。
2. **筹备**
   - 确定拍摄地点、演员、道具等资源。
   - 制定拍摄计划和时间表。
   - 准备拍摄设备和技术团队。
3. **拍摄**
   - 实际拍摄视频素材。
   - 确保拍摄质量和符合剧本要求。
4. **后期**
   - 视频剪辑和编辑。
   - 添加音效、音乐、特效等元素。
   - 调整画面色调、字幕等。
   - 最终呈现和渲染视频。

基于AIGC工具的提示交互式视频生成制作流程

1. **用户提示**
   - 用户提供关于期望视频内容的提示词或描述性文本。



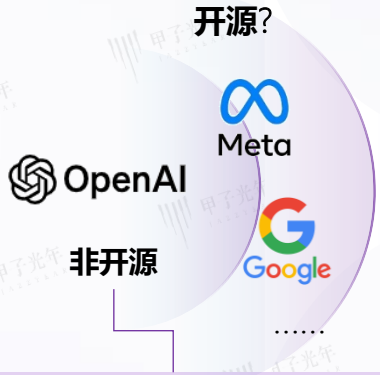
5. **最终生成**
   - 生成满足用户要求的最终视频内容。

数据来源：甲子光年智库，2024年

2.2 L1-AI生产时代：AIGC视频生成将“多点开花”，开源是下个关键节点

伴随开源模型的出现，AI视频生成将迎来多元化的入局者

文生视频领域迫切需要如LLama2的模型，让更多应用层公司节省从0-1的成本



Sora虽未公测，根据当下的视频效果，模型及对应的技术路线与其他公司已经拉开差距，但猜测其模型可能仿照ChatGPT，不提供开源模型

相关赛道企业若想具备先发优势，要么技术更强，建立技术壁垒，要么产品对用户需求的理解更深，建立用户粘性和数据飞轮

AI+视频创业公司

以生成式AI技术为底色，已经完成部分技术积累，正在进行技术追赶和体验创新。

runway、NeverEnds、RightBrain、Pika、SIX PEN Art、HDream.ai、徐图智能、Mobius、morph、生数、.....

数字人技术提供商

数字人本身可以完成部分视频录制，虽然暂且无法完成端到端生成（文字直接生成视频），但可以快速满足部分场景需求。

链基智能、出门问问 mobvoi、世优科技、前领时代、小冰、LUSTER 凌云光、Mogics、.....

AI+影视公司

对视频，尤其是专业视频（影视、广告、动画或游戏）具备深刻理解，AI技术可以充分提供视频创作、分发的工具。

wondershare 万兴科技、万兴影镜、帝视科技 IMPERIAL VISION、DIGITAL DOMAIN 数字王国

互联网科技企业

具备充分的技术积累，产品丰富，平台用户量高，可迅速在内容产业中实现价值。

Alibaba Group 阿里巴巴集团、字节跳动、Baidu 百度、快手、Tencent 腾讯、360、.....

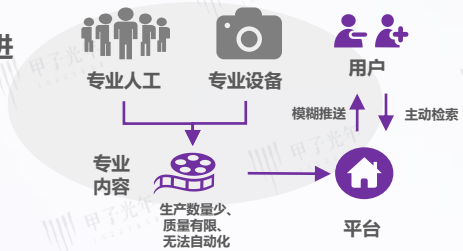
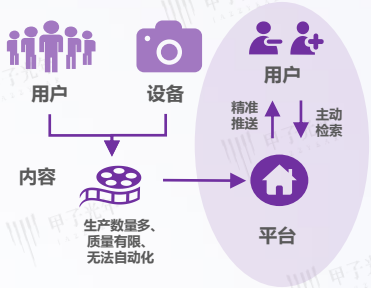
C端用户

全民视频创作的浪潮正蓄势而来，未来人人都会成为导演，每个人都会拥有个人平台。

2.3 L2-AI原生时代：视频用户身份实现“三位一体”

角色变迁：视频用户变为AI原生居民，实现生产者、消费者、拥有者“三位一体”

- 越来越多视频用户将成为AI原生居民：他们同时是内容生产者、消费者和拥有者。个体在视频内容生产、消费和拥有方面拥有更大的主动权和自主性。这将改变人与内容、人与人的生产关系与交互关系。

阶段	信息时代	数字时代	AI原生时代
内容生产形式	• PGC	• UGC	• AIGC
角色转换	• 内容消费者	• 初次内容生产者	• **AI原生居民**【生产、消费、拥有一体】
核心环节演进			

数据来源：甲子光年智库，2024年；

2.3 L2-AI原生时代：视频成为“第二语言”

视频成为人类的第二语言，人类语言告别“词不达意”阶段

- 伴随视频生产成本的无限降低，以及视频可交互、可编辑、可定制的灵活能力，人人可驾驭视频表达的时代到来。
- 视频将成为人类的第二语言，大量用户会进行行为迁移——原本用文字表达的场景，将用视频直接表达。
- 视频具有连续性，视频表达将不受限于“词汇量”，比文字表达拥有更丰富、更沉浸的特征，可以表达更准确的场景、承载更丰富的情感、抵达更深刻的共情。
- 文本与视频的无缝切换，让人类语言告别“词不达意”、“意在言外”的阶段。

图：人类传递信息的内容载体形态演变历程

阶段>	古代	媒体时代	网络时代	数字时代	AI原生时代
投放渠道>	牌匾	媒体刊物	互联网	视频平台	**元宇宙、虚拟世界**
交互变迁>	离线	离线	在线，单向	在线，双向	**实时、沉浸**
角色变迁>	高门槛的内容生产、消费者	高门槛的内容生产、消费者	内容消费者	内容生产者	**生产、消费、拥有三位一体**
内容形态>	文字	文字+图片	文字+图片+广告视频	视频	**可交互、可编辑、可定制的视频**

视频成为第二语言

- **视频语言**：视频语言指利用视频和图像等视觉元素进行交流表达的语言形式。
- **视频语言的特点**：

1. **视觉化表达**：与文字语言相比，视频语言主要通过视觉影像来表达信息，通过图像、颜色、动作等元素传达更加直观、生动、丰富的信息。
2. **多媒体结合**：视频语言通常结合了图像、声音、文字等多种媒体形式，丰富了表达手段和效果。
3. **情感共鸣**：视觉和声音的传达方式更容易引起情感共鸣。
4. **多样化形式**：视频语言可以呈现为电影、电视、短视频、动画等多种形式，适应不同场景和需求。

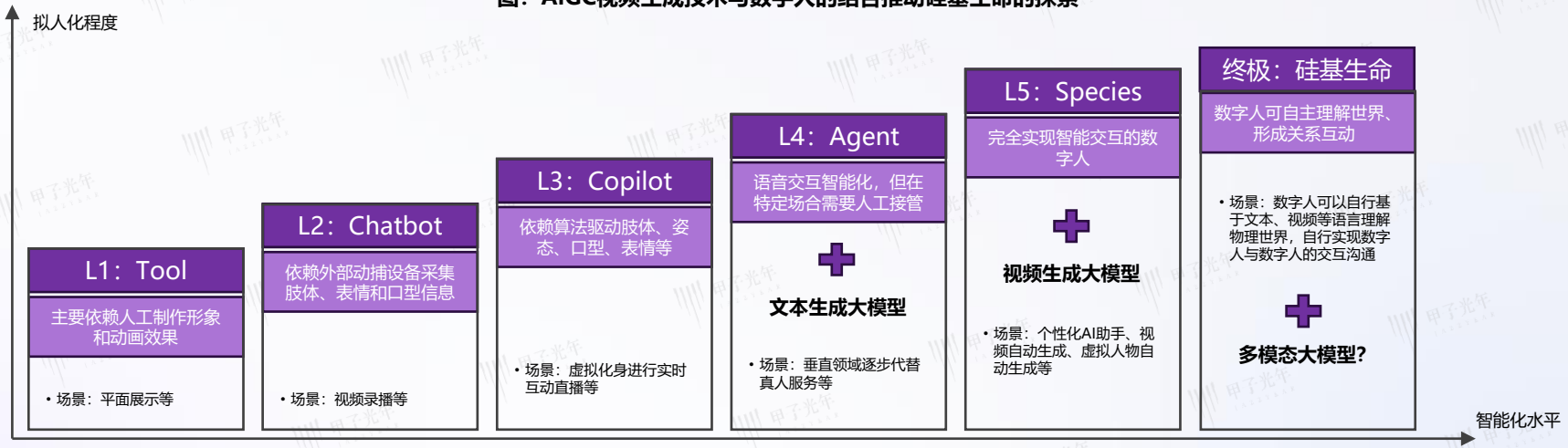
数据来源：甲子光年智库，2024年；

2.3 L2-AI原生时代：AI渗透率无限逼近人类在数字世界生产活动行为边界

数字人与视频生成大模型的结合，推动数字人发展进入L5级

- 数字人与AIGC的结合一直是重要发展方向。在Sora出现之前，主要是数字人与文本生成模型（如GPT系列）的结合，生成虚拟角色的对话和互动内容，主要应用于虚拟助手、客服机器人、虚拟主持人等基于文本的交互和对话场景。
- AIGC视频生成技术的发展将会推动数字人进入全新阶段。数字人与视频生成大模型（如Sora）的结合，提升了数字人的逼真度和互动性，其应用场景会进一步拓宽，涵盖虚拟演员、虚拟教育导师等需要视觉交流和场景互动的领域。
- 未来，数字人还会探索与多模态大模型的融合发展，继续提升仿真度和互动性、拓展应用场景、 探索人机交互的新可能，丰富人们感知和改变世界的方式。
- **“硅基生命”将加速到来，无限逼近人类在数字世界的生产活动行为边界。**

图：AIGC视频生成技术与数字人的结合推动硅基生命的探索



2.3 L2-AI原生时代：元宇宙相关产业将加速到来

应用层与交互层在B端C端都将带来无穷想象空间

- 由于视频和C端有天然的联系，AIGC视频生成技术的快速发展将推动应用层和交互层的快速发展。通过简单的操作用户即可快速生成高质量的视频内容，将大大提升用户体验和参与度，推动元宇宙生态的蓬勃发展。
- 因此，在传统AI技术栈上，应用层和交互层将诞生丰富的创新机会，在B端和C端都迎来无穷的想象空间。

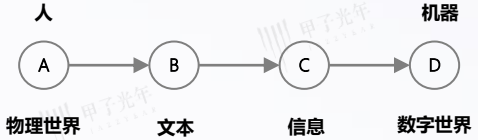
图：AIGC视频生成将加速元宇宙世界的内容构建

2.4 L3-AI创生时代：重塑人与机器的交互

视频成为机器理解物理世界的主要媒介，推动AI与物理世界进一步融合

- 人机交互进入视频语言时代。与传统人机交互相比，视频语言在信息表达形式、感知方式、交互体验和个性化定制等方面都有较大差异点，为用户提供了更加丰富、直观和个性化的交互体验。
- 视频等多模态内容的信息含量更大、更多元，让机器更容易理解物理世界，让机器人真正成为数字世界与物理世界的桥梁。
- AIGC视频生成与具身智能、工业视觉、工业元宇宙等方向的结合，将会推动AI突破数字世界，与物理世界进一步融合。

文本传递信息为主的人机交互



交互语言：文字为主，传统人机交互主要依赖于键盘、鼠标、触摸屏等输入设备以及文字、图像、声音等输出方式进行交流。

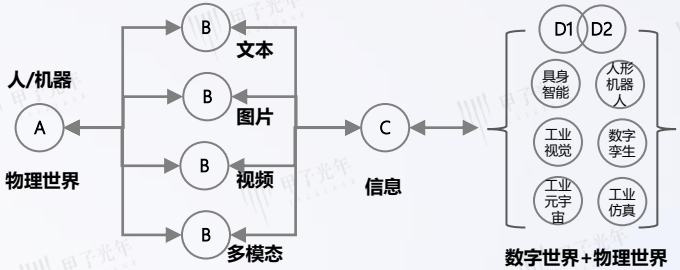
信息表达：信息含量低。传统人机交互以文字、图像、声音等为主要表达方式，信息相对单一。

感知能力：传统人机交互主要依赖于计算机对文字、图像、声音等信息的理解和处理能力。

交互体验：单向交互模型。传统人机交互通常是静态的，用户通过键盘、鼠标等输入设备与计算机进行交互，交互过程相对单一。

>

视频等多模态传递信息为主的人机交互



交互语言：声音、动作、表情、场景.....都可以作为机器理解人的指令的输入形式，再配以摄像头等传感器的机器将会主动理解世界。

信息表达：信息含量大且多样化。视频语言时代的人机交互更加丰富多样，信息以视频为载体，可以包含文字、图像、声音、动作等多种元素，表达更加生动和直观。

感知能力：视频语言时代的人机交互需要计算机具备更强的视频感知和理解能力，能够识别、理解和分析视频中的内容和情境。

交互体验：实时、沉浸式交互。视频语言时代的人机交互更加动态和生动，用户可以通过拍摄、录制视频、实时互动来与计算机进行交互，交互过程更加自然和直观。

数据来源：甲子光年智库，2024年；

2.4 L3-AI创生时代：数字科研推动新一轮“科学革命”

AIGC生成技术与数字孪生、仿真等融合，可驱动科技研发进入全新范式

- AIGC生成技术与数字孪生、仿真等技术的融合可以探索出一条基于虚拟世界仿真的科技研发模型。这种模型可以通过在虚拟世界中建立逼真的数字孪生模型和仿真环境进行科技研究和实验，大大提高科研的效率，解放科研工作者的人力，降低综合科研成本。
- **甲子光年智库将这种基于虚拟世界仿真的科技研发模型称之为数字科研模型，将通过数字科研模型进行研发的模式称为“数字科研”。**
- 当前，AI已经在药物研发、合成生物等基础科学研究中得到广泛应用。AI的进一步发展，将推动数字科研加快实现。未来数字科研模型有望成为科学研究的通用基础设施，在各个学科普及，这将催生新一轮科学范式革命。

图1：AIGC在基础科学研究中应用于众多领域

图2：数字科研的实施步骤

数据来源：甲子光年智库，2024年；

2.4 L3-AI创生时代：世界模型创造人类智慧之外的“第二智慧体系”

两种世界模型：预测式模型和生成式模型

- 世界模型的核心路径分歧来自于：**世界是不是真的需要一个解析解？**
- 人类依靠大脑来理解世界。美国艺术与科学学院院士、加拿大皇家学会院士莉莎·费德曼·巴瑞特在《认识大脑》一书中提出了人类大脑通过对外界刺激进行预测来解释和理解世界的过程。甲子光年将这个过程概括为四个阶段：预测阶段、感知阶段、匹配/思考阶段、形成认知阶段，可简称为“预测式模型”。
- 是否遵循大脑理解世界的模式构成了世界模型的不同思路，将催生不同技术路线。**伴随AI创生时代到来，我们将迎来人类大脑智慧之外的“第二智慧体系”。**
- 甲子光年将世界模型大体划分为两类：
  - **第一智慧体系：预测式世界模型，**代表是人类大脑，Meta的V-JEPA也属于预测式模型。
  - **第二智慧体系：生成式世界模型，**代表是ChatGPT、Sora等深度学习的数据驱动流派。深度学习的数据驱动流派的核心思路是：通过大量数据模拟世界所得到的结果可能会比一个解析解更能反映世界的真实物理，更能体现智能。
- **人类智慧只是智慧的一种范式，ChatGPT、Sora等范式已能够通过大量模拟世界学习到世界规律。**因此，用一个物理公式概括现实世界的思路并不一定正确，**深度学习的数据驱动流派开启的“第二智慧体系”也可能成为理解世界最终奥妙的一把钥匙，而非追求解析解。**

图1：大脑理解世界的四个环节

感知前的预测阶段，大脑基于内部模型、先前经验和期望，预测可能发生的情况。 **预测**

对外界刺激的感知，接受外部刺激，形成感知信号，并进行感知输入。 **感知**

在预测与感知的匹配阶段，大脑将外界刺激与先前的预测进行匹配，并调节、更新预测，以使其与感知一致。 **匹配/思考**

在意义建构和理解阶段，大脑解释、理解感知到的信息，并将其纳入更广泛的认知框架中。 **形成认知**

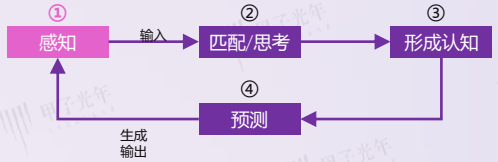
第一智慧体系：预测式模型对世界的理解逻辑 — **Meta的V-JEPA**

路径1：在感知前先预测，然后接收外部刺激信号进行思考匹配并形成认知，基于预测输出对世界理解的内容。认知后是输出阶段。

预测式模型：将路径1对世界的理解逻辑称为预测式模型。

世界模型代表：Meta的V-JEPA。

第二智慧体系：生成式模型对世界的理解逻辑 — **OpenAI的Sora**



路径2：第一步是直接接收外部刺激信号进行思考匹配后形成认知，基于认知预测生成对世界理解的内容。预测后是生产输出阶段。

生成式模型：将路径2对世界的理解逻辑称为生成式模型。

世界模型代表：OpenAI的Sora。

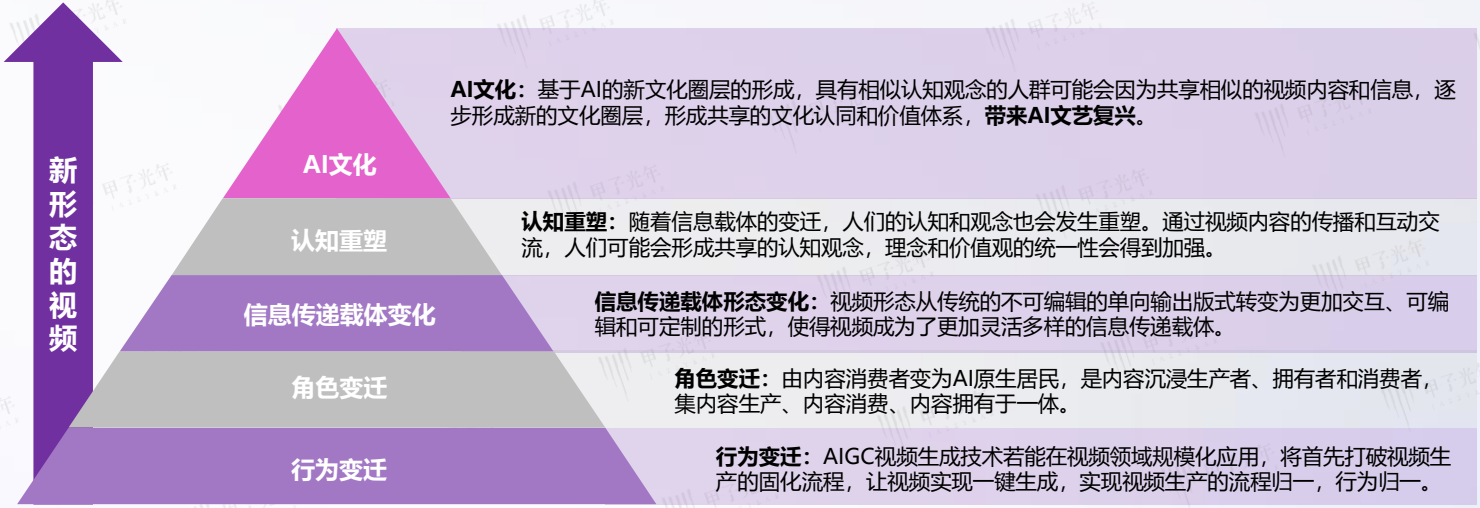
数据来源：甲子光年智库，2024年；

2.5 L4-AI文明时代：AI驱动文艺复兴

交互行为、角色、载体形态的变化推动人类认知重塑，并形成新文化圈层

- 视频作为一种生动、直观的传播媒介，能够更好地激发人们的情感，与文字、图片相比，视频更能引发观众的深度共鸣和参与。
- 信息载体的变化会重塑人类的认知与观念，并将具有相同认知观念的人群逐步集合到一起，形成新的文化圈层，推动文化的变迁，并进一步推动AI版本的文艺复兴。

图：AI驱动文艺复兴



数据来源：甲子光年智库，2024年；

2.5 L4-AI文明时代：重塑人类文明

文明的演进：人类文明进入与AI共建共生的“双生时代”，AI的归AI，人类的归人类

- 波普尔的世界三元组是哲学家卡尔·波普尔提出的概念，用于描述对世界的基本认知，包括三个要素：物理世界、心灵世界和符号世界。
- AI从符号世界出发，参与物理世界的方式是逐渐建立通用的世界模型，参与心灵世界的方式是生成无限的创意和想象。
- 最终，AI会深度参与人类的物理世界和心灵世界，人类智慧与AI将互相影响、共同进化，人类文明进入“双生时代”，形成“AI的归AI，人类的归人类”的有序的分工。

图1：AI主导的世界及其两条影响路径

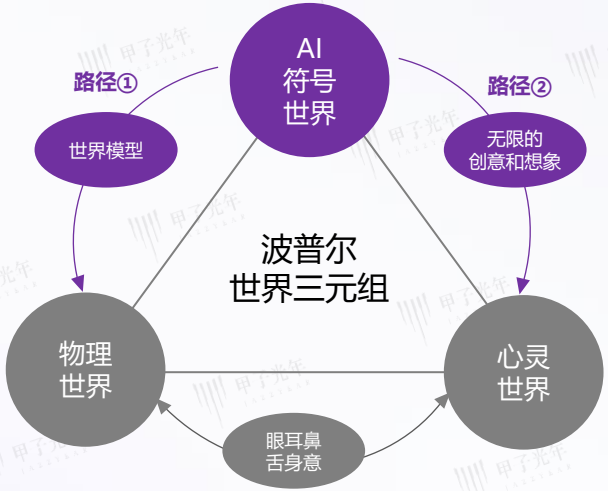
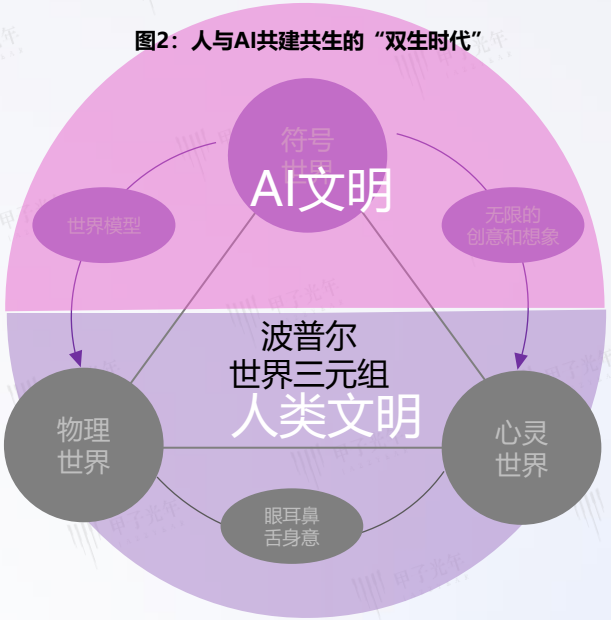


图2：人与AI共建共生的“双生时代”



备注说明：波普尔的世界三元组
第一元：物理世界（World 1）：指的是客观存在的实体世界，包括物质和能量等自然现象。物理世界是独立于我们的意识和思想存在的，是客观存在的。
第二元：心灵世界（World 2）：指的是个体的主观意识和心理活动所构成的世界，包括思想、感觉、情绪、意识等心理现象。心灵世界是个体内部的心理体验领域，是主观存在的。
第三元：符号世界（World 3）：指的是人类通过语言、符号和文化制度等共同建构的文化世界，包括科学理论、艺术作品、社会制度、文化传统等。符号世界是人类共同的文化积累和认知产物，是客观存在的，但是不同于物理世界，是通过人类的创造和交流而存在的。

数据来源：甲子光年智库，2024年；

目录

CONTENTS

甲子光年
JAZZYEAR

3.1 视频内容的两大类型：短视频和长视频

IP→内容→衍生，是视频内容价值链的主要逻辑链条，长视频与短视频是两大核心类型

数据来源：甲子光年智库，2024年；

3.2 传统视频产业链：完整产业链

甲子光年 JAZZYEAR

视频产业链包含七个关键环节，制作环节是最核心环节，也是AI视频生成工具现阶段主要服务环节

图：传统视频产业链及关键核心角色

数据来源：甲子光年智库，2024年；

3.3 传统视频产业链：长视频与短视频的核心产业链环节

长视频需覆盖完整视频产业链，短视频则更注重分发和变现

- 长视频与短视频在产业链中的各自侧重点有显著性差异。长视频需要完整覆盖视频产业链，并非常注重制作环节的投入。短视频则对上游IP、策划、投资等环节依赖度极低，通常关注中下游的制作、分发与变现环节。

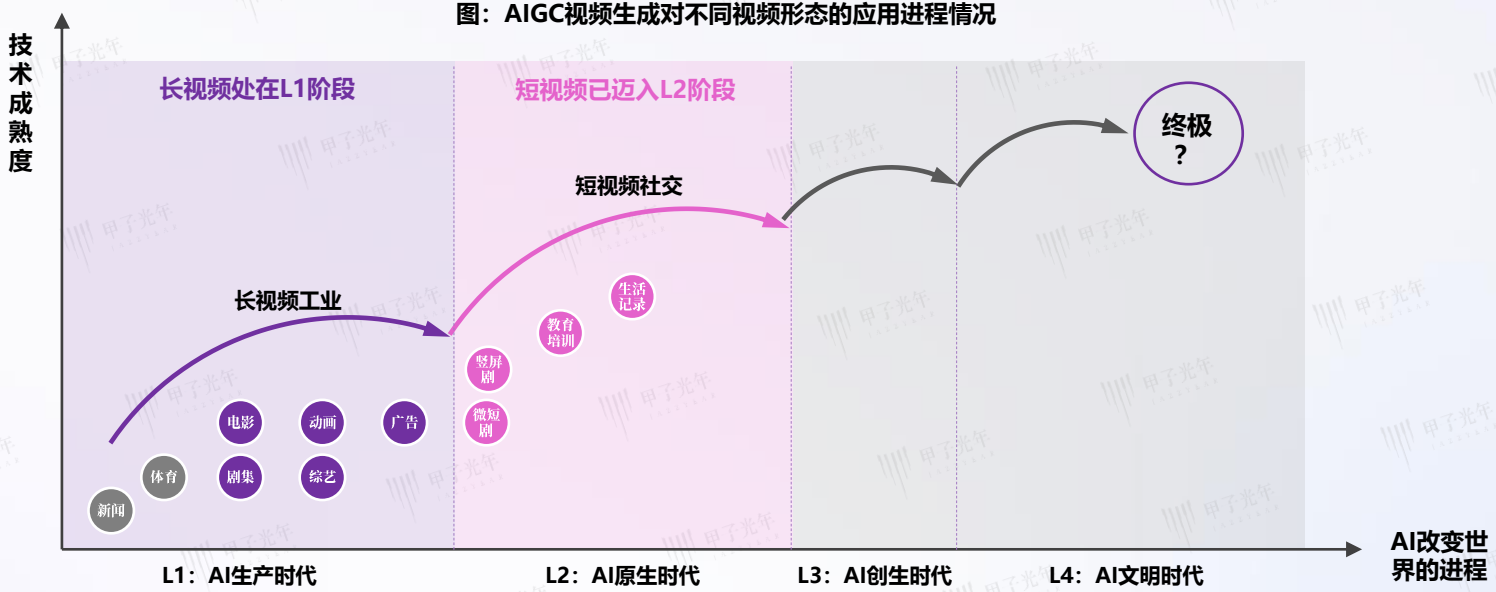
图：长视频与短视频的核心产业链环节

3.4 不同形态视频细分领域的应用进程

短视频正在进入AI原生时代，长视频正在进入AI生产时代

- AIGC视频生成技术在不同形态的视频内容领域的应用进程各不相同。概括而言，长视频领域AIGC视频生成技术仍然处于L1阶段，由于现阶段AIGC视频生成技术的局限性，导致一些具有高度专业性的领域仅仅将其作为生产工具，例如为电影、剧集等提供素材来源，尚无法带来颠覆性重塑，但会压缩原有产业链。而对于新闻这类需要高度准确性的内容，则暂时只能满足情景复现等少量场景。
- 短视频领域则会首先面临AIGC视频生成技术的颠覆，甲子光年智库判断短视频领域将会进入L2即AI原生时代，短视频产业链将不复存在，而会诞生AI原生的短视频模式和平台。

图：AIGC视频生成对不同视频形态的应用进程情况



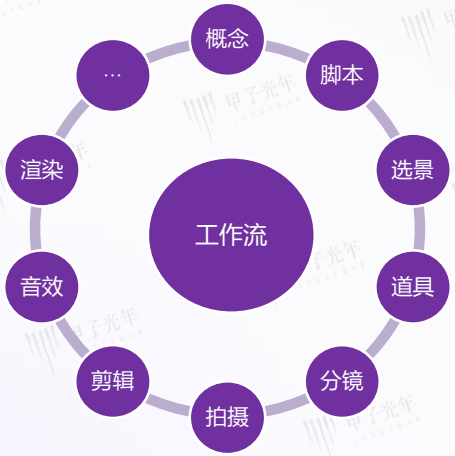
数据来源：甲子光年智库，2024年；

3.5 短视频重塑后的产业链

短视频进入AI原生时代，产业链被压缩，催生AI原生模式的新型平台

- PC互联网时代催生出了长视频平台，移动互联网催生出了短视频平台，虽然短视频平台已经在大幅度应用AI技术进行赋能，但仍然存在显著的短视频生产工作流和短视频制作的角色分工体系。
- AIGC视频生成技术将会打破短视频的原有产业链，大幅度压缩简化生产制作流程和角色分工，实现一键生成的all in one原生模式。
- AI原生视频流程的归一，将会带来AI原生时代的短视频平台新范式，新的视频平台范式将具有无序、沉浸，实时、互动、聚联的AI原生特征。无序是指打破传统固化的视频生产流程。沉浸是指实现全面体验的沉浸式视频生产。实时是低延时的视频快速生成。互动是指一边交互对话一边进行视频调整的个性化、定制化的互动视频。聚联是指去中心化的生产方式。

数字时代短视频的工作流



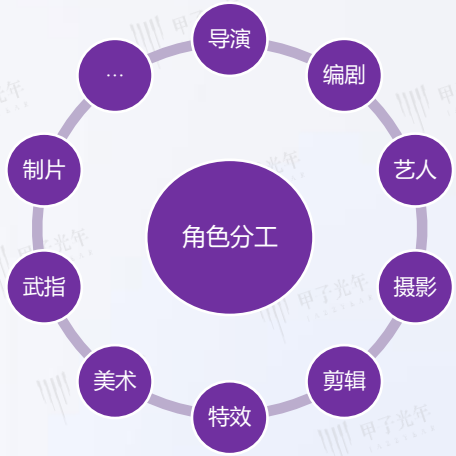
AI原生时代的短视频平台

AI原生平台

All in One，链条缩短，环节融合
无序、沉浸、实时、互动、聚联式平台

AI克隆音色	数字人口播	视频生成
图文成片	素材生成	个人平台

数字时代短视频的角色分工



数据来源：甲子光年智库，2024年；

3.5 短视频重塑后的产业链

新型制播一体的AI原生内容平台，有望颠覆短视频平台格局，每个用户既是导演又是平台

- 在AIGC视频生成对视频产业链的技术变革下，有望孵化出新一代的集短视频制作、分发、变现为一体的全新形态的视频平台。
- **新型的制播一体的AI原生内容平台应该是融合AI创意工具+AI原生视频工具+变现平台三大环节的AI原生短视频平台。**
- 在实现AI原生范式的转换后，过去短视频平台和内容创作者将会出现一些变革：
  - **短视频平台：**将会向AI创意工具+AI原生视频工具+变现平台的融合式的平台转变，提供AI原生视频工具和流量分发平台。
  - **用户平台：**用户将不仅仅是作为生产者和消费者，真正做到人人都是导演型的创作者，并且人人都是一个小型的平台。个人用户可以建立个人平台，企业用户可以建立企业平台，直播机构可以建立直播平台。内容创作者的价值将更注重创意能力、解决实际问题的能力、个人IP影响力等。

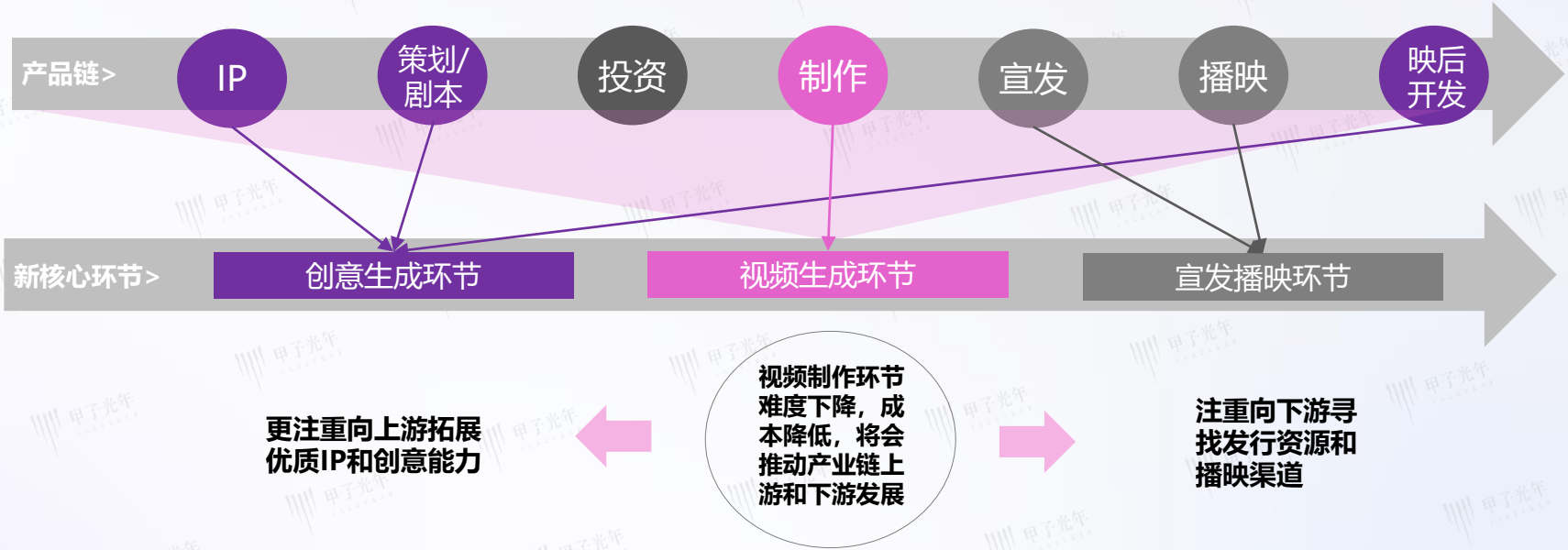
图：新型制播一体的AI原生内容平台的业务模式

3.6 长视频重塑后的产业链

制作环节难度下降，将会助推产业链上游的创意环节和下游宣发播映环节重要程度上升，好故事、好脚本、好平台将成为视频产业的核心竞争力

- AIGC视频生成工具会降低视频制作的准入门槛、抛弃对专业设备的依赖、降低生产成本、提升制作效率。
- 随着制作环节难度下降，好故事、好脚本等产业链上游的创意环节将成为视频产业的核心竞争力。
- AI视频产量的大幅增加需要更符合AI视频特征的播映平台，产业链也将更为注重下游宣发播映平台渠道的建设更新。
- 原有产业链的投资环节主要针对内容制造环节，未来文娱和技术投资将走向融合。

图：AIGC视频生成简化传统视频产业链



3.6 长视频重塑后的产业链

重塑后的视频产业链将整合简化为三大环节：创意生成—视频生成—宣发播映

数据来源：甲子光年智库，2024年；

3.6 长视频重塑后的产业链

重塑后的视频产业链将会变为基于三大模型体系的全新产业链，并带来全新的生产方式

- 重塑后的视频产业链将基于三大环节产生三大产业体系：基于创意生成模型的产业体系、基于视频生成模型的产业体系、基于宣发播映模型的产业体系。

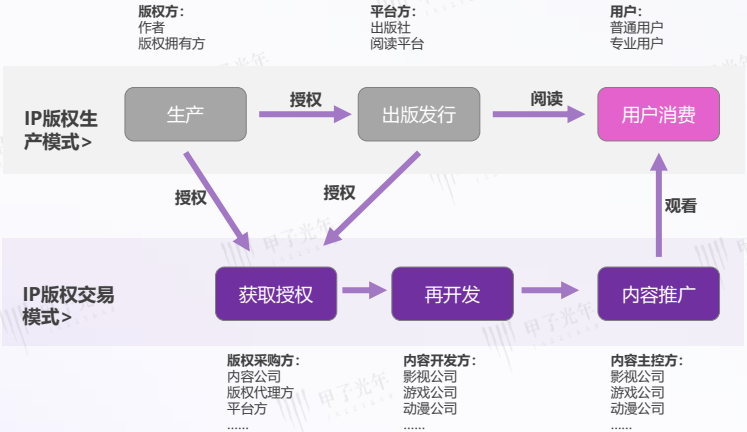
图：AIGC视频生成整合重塑后的全新视频产业链

3.6 长视频重塑后的产业链机会：创意生成体系

IP版权生产交易开发一体化的平台有望成为新的发展机遇

传统的版权生产与交易模式

- 传统IP版权的生产非常依赖作家的能力，而在版权交易环节通常是由版权持有者直接与使用方（如出版商、电影制作公司等）之间进行的，可交易范围较窄。



IP版权生产交易开发一体化的平台

- 使用AIGC技术后，版权生成环节可以直接使用大模型来生成内容，并且可以是文本文章、图片、音频剪辑、视频片段等各种形式的内容。版权交易不再是人与人之间的交易，而是创作者与技术平台之间的交易。因此，IP版权生产和交易开发一体化的新型平台有望迎来机遇。

数据来源：甲子光年智库，2024年；

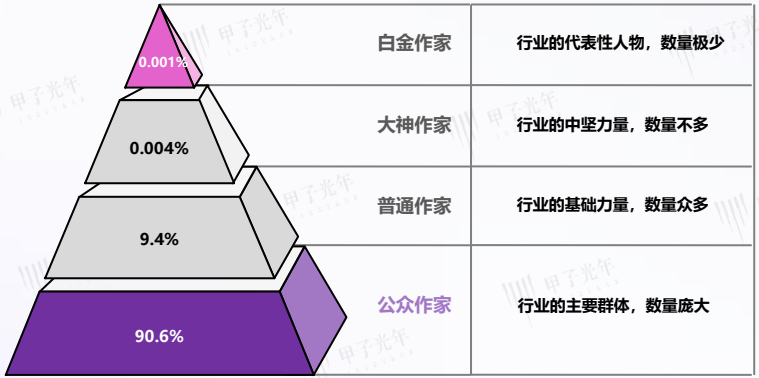
3.6 长视频重塑后的产业链机会：创意生成体系

白金作家群体不再成为稀有资源，未来人人都是小说家，中小型文学平台将可能迎来春天

现阶段的作家群体分布呈现金字塔状态

- 在传统视频产业链中，上游的IP环节中更注重处于金字塔顶端的白金作家群体的维护管理与产品设计，大型文学平台通过垄断平台和作家资源构建核心竞争力。
- AIGC视频生成重塑后的产业链体系中，创意生成体系将会更为注重底层占比90.6%的公众作家，其将成为很多视频生成平台的创意来源和IP输出者。整合AIGC技术、打通AIGC小说生成流程的中小型文学平台有望迎来发展的春天 。

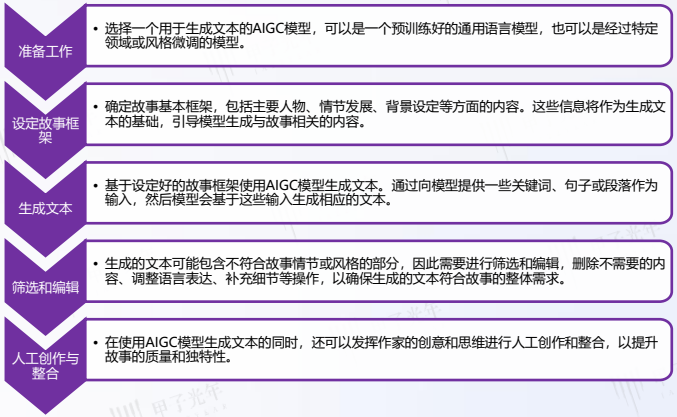
图1：作家群体分布图



人人都是小说家

- 使用AIGC模型撰写小说可以帮助小说作家快速生成大量文本，并为创作提供灵感和创意的启发，可以显著降低撰写小说的门槛，未来小说家将不再是特点人群，而是人人都是小说家。
- 伴随内容供给的增加，传统寡头垄断型的文学平台有望被打破，中小型文学平台的发展将迎来春天。

图2：AIGC生产小说的创作流程



数据来源：甲子光年智库，2024年；

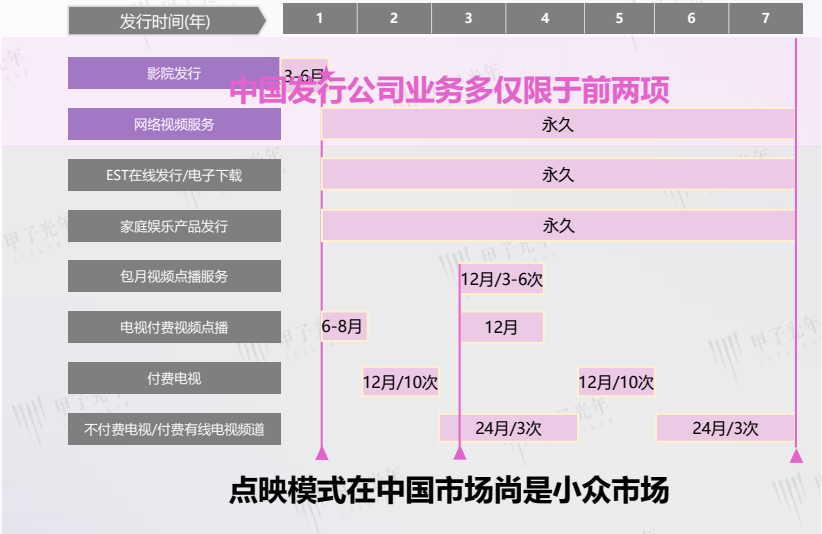
3.6 长视频重塑后的产业链机会：宣发播映体系

传统制播分离模式将被抛弃，沉浸式互动点播云影院将成新机遇

传统的宣发模式是制播分离的

- 在长视频领域，中国传统主流视频宣发模式仅仅聚焦影院发行和网络视频平台发行两种，点播影院模式一直是小众市场。

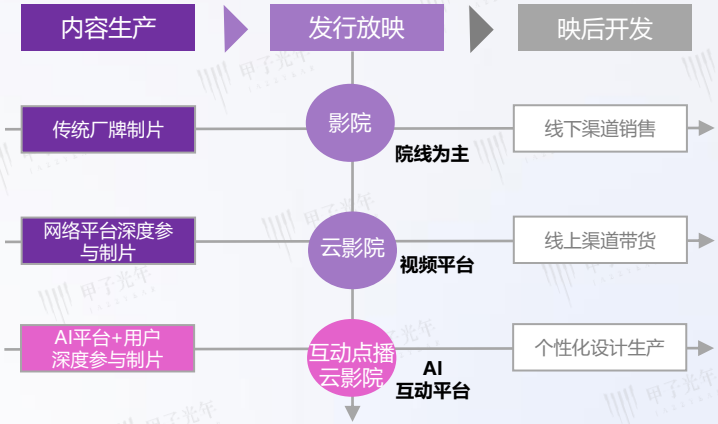
图1：传统专业长视频发行模式：以美国电影发行为例



沉浸式互动点播云影院有望成为第三大发行体系

- 传统视频宣发模式导致下游宣发播映环节过于依赖影院等传统平台或长视频平台。
- 随着AIGC视频生成对视频产业链的重塑，过去基于固定宣发流程的发行模式有望向互动点播模式转型，以AI平台为核心，让用户深度参与电影前期的剧本创作和内容生产，并在沉浸式互动点播云影院上线，满足不同用户对故事走向的不同需求。

图2：“院网”并行的发行模式与新型点播云影院发行体系



数据来源：甲子光年智库梳理，2024年；

目录

CONTENTS

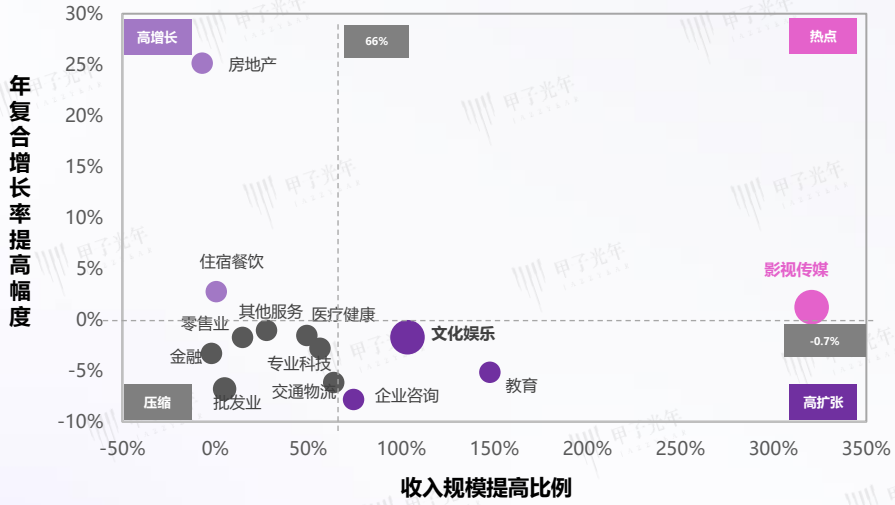
甲子光年
JAZZYEAR

4.1 当经济体处于下行周期时，文娱产业迎来发展

视频是文娱产业的核心内容形态，将会受到大经济周期的影响，迎来新的发展机遇

美国在经济调整期时增长最快的行业是文娱业

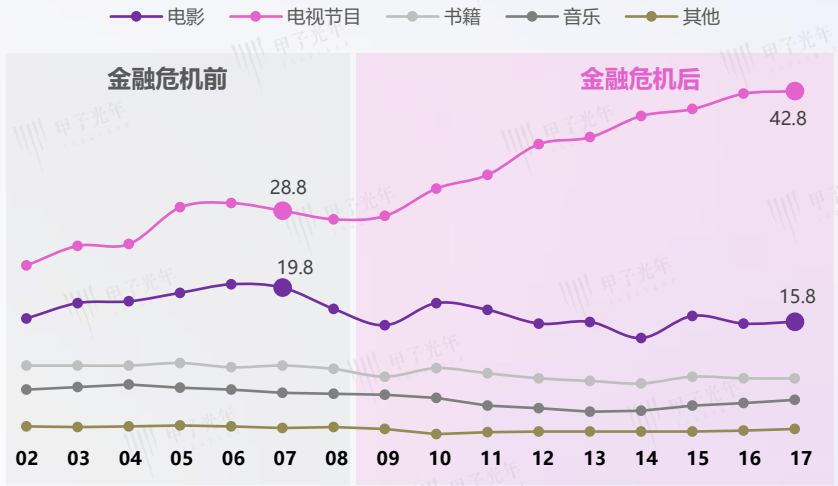
图1：美国经济调整期时的热点行业



备注：2008年金融危机前和危机后美国各服务行业的企业盈利能力情况对比分析

美国经济体文娱业在调整期后反而成为投资高增长领域

图2：文娱企业各细分行业投资规模走势图（十亿美元）



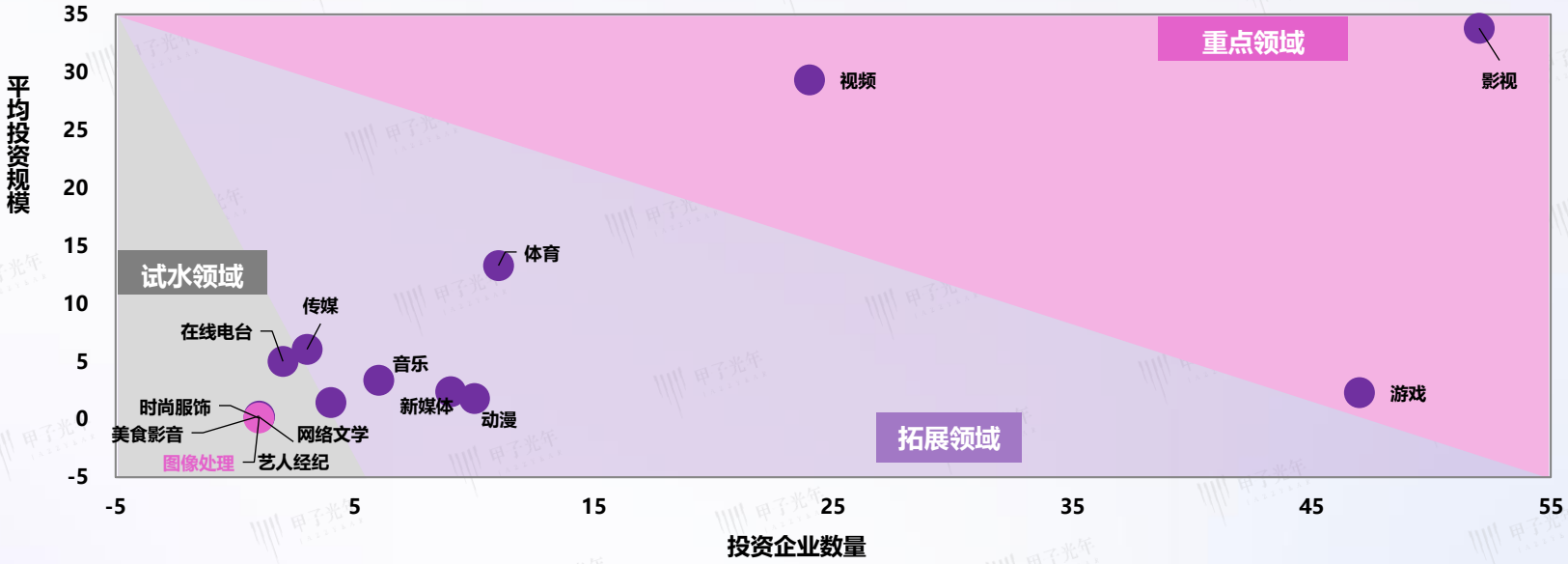
数据来源：美国经济分析局，甲子光年智库梳理，2024年；

4.2 中国文娱领域第一轮投资浪潮：2011-2017年

第一轮投资浪潮中，文娱产业投资聚焦内容生态本身，而忽视了底层技术领域

- 中国文娱市场在2011-2017年是投资高峰期，在第一轮投资浪潮中，影视、视频、游戏是重点投资方向，其次是体育、音乐、动漫、新媒体等领域。
- 在第一轮投资浪潮中，投资机构更多聚焦内容生产类，而忽视了为内容生产机构提供底层技术工具的厂商。

图：2011-2017年主流企业泛娱乐投资领域趋势分布



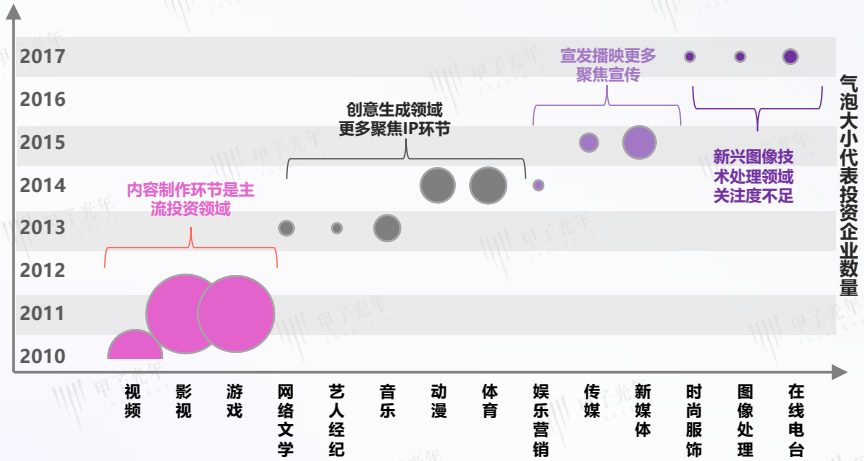
4.2 文娱领域第一轮投资浪潮：2011-2017年

BAT在第一轮投资浪潮的主投资方向是内容制作和播映平台，对创意生成、技术领域缺乏关注

视频内容制作是BAT投资文娱的聚焦方向

- BAT都是先从泛娱乐产业链中游的内容制作环节介入，即视频和影视，而后拓展上游和下游布局。

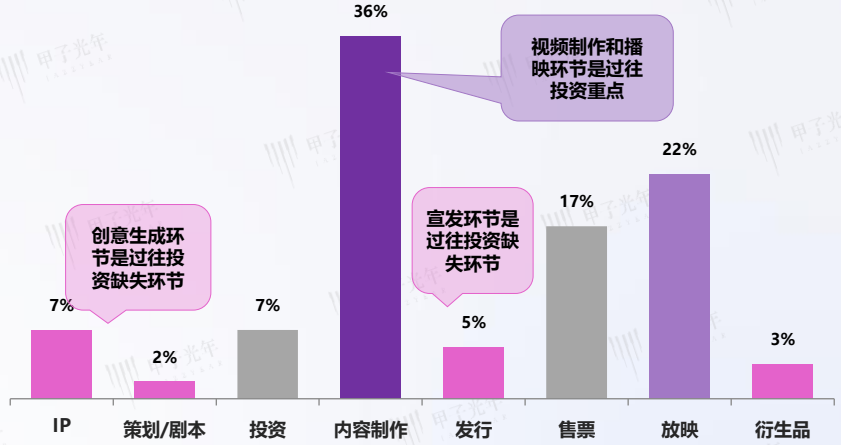
图1：2011-2017年BAT泛娱乐投资历程分布图



创意生成类和底层图像处理技术类缺乏关注

- 内容制作类企业是过往投资重点，但IP、剧本策划类和衍生品类等创意生成体系下的企业是第一轮浪潮中关注度不足的领域。

图2：2011-2017年BAT企业影视行业投资企业数量分布图



备注：BAT主流投资机构包括腾讯、阿里、百度等三家公司对泛娱乐领域投资企业分布情况；

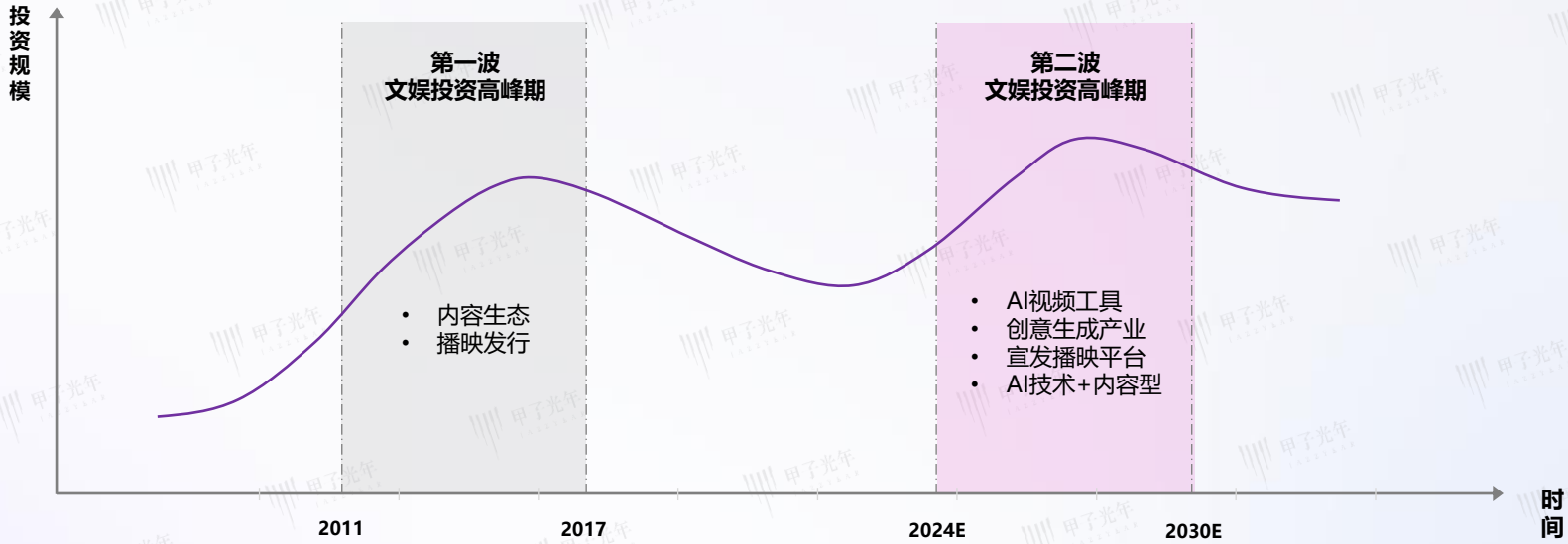
数据来源：甲子光年智库梳理，2024年；

4.3 文娱领域有望开启第二轮投资浪潮

经济周期调整与技术革命双重加持将推动以视频为核心内容形态的文娱产业迎来第二轮爆发期

- 文娱领域在经历第一波投资高峰期后，在2018年开始进入下行周期。在宏观经济周期与AI技术革命的双重加持下，文娱领域有望开启第二轮投资浪潮。
- 在文娱领域的第一轮投资浪潮中，投资机构主要聚焦视频产业链的制作和播映环节，标的企业以影视公司、视频播映平台、影视项目等为主。
- 在文娱领域的第二轮投资浪潮中，投资方向将会更多聚焦底层技术及与技术相融合的内容公司，标的企业将以AI视频工具、创意生成类企业、新型宣发播映平台等为主。

图：以视频内容形态为核心的文娱领域投资周期曲线图



数据来源：甲子光年智库梳理，2024年；

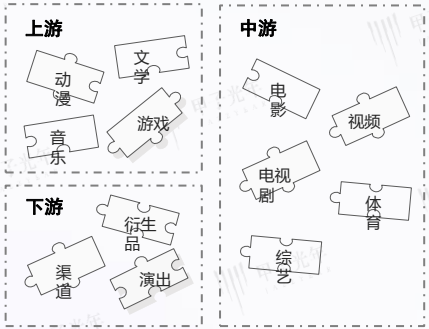
4.3 文娱领域有望开启第二轮投资浪潮

第二轮围绕视频内容为核心的投资热点领域，将以视频内容+技术的生态型公司为主

- 在第一轮文娱领域的投资浪潮中，投资主流形式是以视频内容项目作为投资标的，如投资某一个电影或电视剧等，未来投资对象将会从投资内容项目为主变为投资企业股权为主，被投企业的业务范围将由内容生产为主变为AI技术与视频生成相结合。

过去公司：单一内容型

过去传统视频等内容公司主营业务以某一个子行业/领域为主。



现在公司：内容生态型

现在的主流视频内容公司则以多个子行业联动形式，试图以打造内容生态体系的方式开展主营业务。

未来公司：AI技术+内容型

未来视频领域的公司应该注重AI技术+内容型的构筑，不能仅聚焦视频的应用层，而是视频生成应用层+中间层，甚至结合视频生成基础层进行布局。

4.4 AIGC视频生成技术的投资价值和方向

大厂适合全都要，初创企业适合介入应用层/中间层，央国企适合从底层基础设施开始布局

- 重塑后的每一个视频产业体系都具有较高投资价值。其中，大厂适合进行全产业链布局，初创企业适合入局应用层或中间层某一细分领域，央国企适合入局算力层、平台层和基础层。

图：AIGC视频生成领域适合投资入局的技术方向

下一步研究计划

甲子光年智库将推出《2024中国AI+视频行业发展研究报告》，征集案例合作，欢迎咨询

- 甲子光年智库将推出AIGC视频生成系列报告，下一步要推出的报告为《2024中国AI+视频行业发展研究报告》，现开展典型案例征集合作，欢迎咨询报名。

Part 1 – 机遇：Sora模型爆火，带来AI+视频领域的新潜力

1.1 AIGC领域迎来巨变：DiT模型点燃行业希望

- Sora的视频效果逼真，引发对“AI+视频”的高度关注
- Diffusion + Transformer模型开辟了新技术思路
- ……

1.2 变化中的机遇：文生视频，甚至多模态视频迎来诸多关注

- Sora与其他企业的技术差距分析
- AI+视频的产品形态一览
- 海内外AI+视频投资情况概览
- 个人内容创作人在行业巨变中的生态位变化

1.3 行业面对的新挑战：要么All in，要么出清

- 挑战1：大企业如何追上行业巨头
- 挑战2：中小企业如何利用AI
- 挑战3：AIGC与行业应用之间离得多远
- ……

Part 2 – 需求：内容行业迎来属于自己的“寒武纪爆发”

2.1 行业场景的深度分析：千行千面，泛内容行业可能迎来生产方式巨变

- 影视制作：超级个体的生产及专业剪辑能力的快速普惠
- 内容社区：大量玩法出现，行业爆款在即
- 广告营销：内容营销+个性投推动MarTech企业提供更优质的全案解决方案
- 游戏娱乐：游戏美术流程迎来创作流程的变革，沉浸式游戏再次探索商业性爆发
- 传媒：视频+自媒体将迎来内容的井喷
- 教育：可视化内容实现教育质量的普惠化

2.2 步步为营：多模态*多场景，产业呈现L1-L4的阶梯发展

- AIGC与视频产业的结合成功关键在于人机协作的理解、程度及流程化
- AI视频时代，依然需要人作为最后的内容审核者，人对于视频合理性及创意性的把控成为AI技术应用程度的关键
- 容错率与创意性，成为to B及to C领域的应用的差别关键

Part 3 – 实践：中国本土企业具备成为全球一流企业的潜力

3.1 中国AI+视频全景图谱

- 算力层、数据层、模型层、应用层：三大关键产业链的全面梳理
- 模型层的深度剖析：模型中间层所对应垂直产业、垂直领域的产业链分析
- 全方位捕捉中国AI+视频企业

3.2 中国AI+视频的优质实践者：用AI开启“人人都是up主”的时代

- AI+视频行业先行企业介绍
- 实践企业的技术背景、商业模式等优势梳理
- 各行业标杆性案例的展示
- 海内外企业的对比及出海机会分析

Part 4 – 未来：视频信息可能成为更优质的信息载体

4.1 趋势展望：视频信息的生产成本迎来视频信息传递的便捷性

- 视频信息是“全世界的通用语言”，真正实现全世界文化与与交流的互通互联
- 虚拟世界（元宇宙）可迎来实质进展，人类具备大规模、低成本地生产沉浸式内容的能力

4.2 挑战与风险：内容的监管面临巨大挑战，算力成本可能加深数字鸿沟

- 视频的内容发布、分发、监管的流程面临重新调整，大量垃圾内容可能充斥互联网
- 算力在短时间内依然可能成为AI普惠的最大挑战，大多数人依然捆缚在信息茧房之内

数据来源：甲子光年智库，2024年；

甲子光年
JAZZYEAR

THANKS

谢　谢

北京甲子光年科技服务有限公司是一家科技智库，包含智库、媒体、社群、企业服务版块，立足于中国科技创新前沿阵地，动态跟踪头部科技企业发展和传统产业技术升级案例，致力于推动人工智能、大数据、物联网、云计算、AR/VR交互技术、信息安全、金融科技、大健康等科技创新在产业之中的应用与落地

关注甲子光年公众号

扫码联系商务合作

甲子光年创始人

张一甲
JJJessica0114（微信）

智库院长

宋涛
13693107167（微信/手机）

商业合作负责人

李胜驰
18600783813（微信/手机）