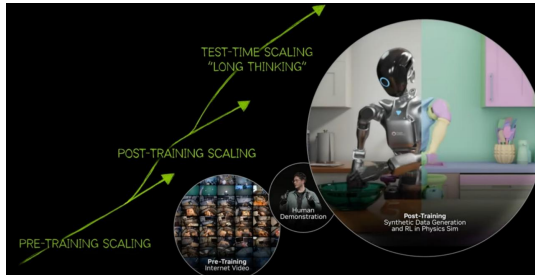


## 【AI】

这周听了黄仁勋在 COMPUTEX 2025 的演讲，整理了一些关键信息：



机器学习到目前为止的迭代周期：第一代：图像识别、文本识别、声音识别；第二代：生成式 AI，生成图像、生成文本、生成视频、生成声音（文本到图片、图片到文本、文本到视频、视频到文本等）。未来无论是 Agent 还是 Robot 主要以生成模型为主。针对 Robot 的学习能力 NVIDIA 推出 DreamGen 项目，目标是机器人首先要学会成为一个机器人，核心思路是：

解决痛点：

- 1、缺乏大规模的真实和合成数据对模型进行训练
- 2、人工示范由于成本问题无法大规模进行

解决思路：

- 1、利用 NVIDIA Cosmos 模型在物理 AI 世界基础模型上实现数据增强
- 2、GROOT Dreams 基于 Cosmos 模型增强后的数据，用于生成大规模合成轨迹数据

训练方式：

Step1: 用特定环境端到端操作人工演示数据 fine-tune Cosmos 模型

Step2: 使用指令 prompt 模型生成梦境或是预测未来状态，利用 Cosmos 是生成式模型的特点，可以生成更多场景

Step3: 对增强后的数据进行评估和筛选

Step4: GROOT Dreams 将 2D 视频转为 3D 视频

Step5: 训练模型

计算中心拆解：

Computer1: AI for learning

Computer2: Simulation engine where AI could learn how to be a robot in a virtual environment

Computer3: Deployment Model

我的思考，未来数据增强和无监督学习或半监督学习会成为生成式模型的核心路径。对于现阶段模型的整体效果判断：

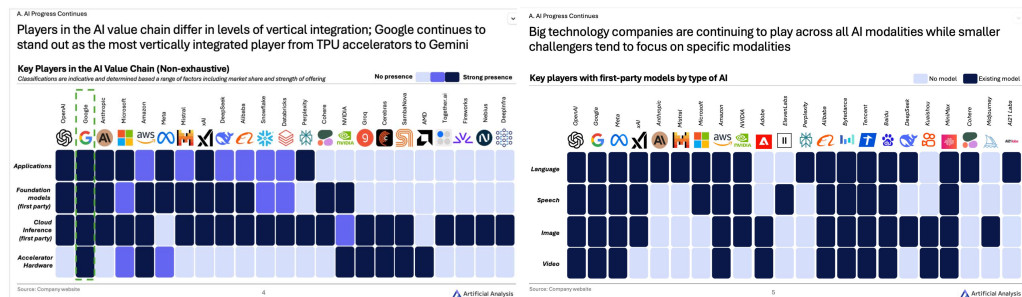
- 1、【训练数据的重要性】模型的训练的关键瓶颈在样本，样本量决定着成本，所以数据增强成为数据增强关键路径。决定一个模型表现的核心要素：数据、模型。极端讲，数据比模型更重要。
- 2、【工具使用技巧】在很多领域，大模型无法给出精准的答案。我们使用大模型便捷的同时，需要有清晰的辨别信息真伪的能力。在使用过程中，如果对要查询的内容完全不懂，那么对模型提供的信息应秉持怀疑态度，对细节事实进行搜索进行进一步验证。利用 AI 工具的原则应该是，多让 AI 提供信息摘要，少让 AI 提供设计方案或者对于 AI 提供的方案要谨慎采纳。
- 3、【大模型目前的瓶颈】目前的模型只能提供助手能力，无法成为教授。

也归因于它吃掉的数据也都是分散的，单独领域的大量数据。所以大模型本身是不具备跨学科综合分析创新的能力，在复杂问题的拆解上面，存在瓶颈。同时由于是商业产品，为迎合用户情绪，模型本身又倾向于顺着用户的思路进行推理，会导致信息进一步失真，从而误导用户。

本周阅读《AI-Q1-2025 行业报告》

<https://artificialanalysis.ai/downloads/state-of-ai/2025/Artificial-Analysis-State-of-AI-Q1-2025-Highlights-Report.pdf>

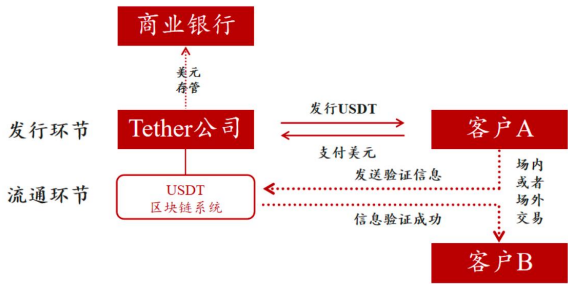
Google 在应用、模型、云计算、硬件等多领域都处于前列位置，发力很猛啊！



【Web3】

最近看了浙商证券提供的《稳定币详解》报告，整理总体思路：

图8: USDT 发行与流通



资料来源:《稳定币的“稳定”与“不稳定”——柯达》，浙商证券研究所

稳定币发行机构铆钉美元发行稳定币，客户将法币换为稳定币，客户用稳定币完成支付需求，发行机构用法币购买美债。

解决美债短期内的需求，维持美债价格。但现阶段并不能够缓解美国政府需要支付的利息。

主权债务的背后是国家信用背书，主权债务的定价本质上是国家信用定价。

对于美国政府来说，一方面要解决美债需求的问题。另一方面是解决政府财政需要支付的美债利息的问题。

引申思考，我一直认为 AI 和 Web3 并不冲突，AI 会成为 Web3 的催化剂。AI 定位在提速工具，Web3 则是虚拟现实的具体产品，AI 可以帮助 Web3 提升迭代和学习的效率。

关于效率，资金使用效率提升和生产工具升级，都会带来社会生产效率的提升，带来社会供给量的增加，带来产品的价格下降，带来社会的福利的增加。

关于信息，在 AI 和 Web3 的驱动下信息的茧房不会消失但会变少。分布式账本的思路提供了去中心化、公开透明、不可篡改、可信任的数字技术，是文类文明从信息不透明逐渐走向信息透明的见证。AI 的领域对于信息的版权是无法维护的，亦是用开源的信息供给给用户开源的信息，所以无论是 AI 还是 Web3 都构建在数据开源的基础上。

关于 Token，在 Web3 中，Token 指代币，数字资产。在 AI 中，Token 指的是一个词、一个像素，数字文本！数字图像。共同之处，都是使用数字来抽象具体的事物。为什么都使用 Token 这个词，我的理解是因为两个领域都有从具体的人类物品到抽象的计算机表达之意。

【谈判&博弈】

最近在阅读《穷查理宝典》学习价值投资，同时我自己也在场官司的漩涡中，在实践过程中，将心得整理出来。

我认为谈判和博弈的核心重点是：逻辑和心力

逻辑:

1、辩论中寻找论据的原则要遵循“第一性原理”，回归到最元的证据。在寻找客观事实的过程中，需要跳出自证陷阱。如：在对方举证的事实中存在我方不利证据时，应该从元数据角度出发收集证据佐证对方提出证据的不合理，而非验证这个证据本身的问题。这里的思维方式我认为类似于芒格先生在投资中的多元思维，从多角度来验证一个公司的潜力，而非仅仅局限在基本面或财报等具体某一个角度。

2、反驳对方或是质疑对方观点时，本着“谁提出谁质证”的原则，提出结论的人进行举证，再对于举证的内容对手方再进行问题拆解、调研。

3、反驳技巧，采用疑问方式，针对存疑的问题让对方再次举证，直至对方无法再此举证，而非直接列举我方的证据，这样容易陷入自证陷阱，更容易被抓到漏洞。

4、收敛目标，确定核心目标，所有的证据应该围绕一个具体的目标列举。这里的思维方式我认为类似于价值投资中的能力圈概念，我们需要集中我们的目标，然后在目标内发起攻击。而不是多个目标分散来打。

5、不凭感觉下结论，而要依据足够多的客观事实下结论。我认为与投资中要足够理性，不能凭感觉操作的思想有异曲同工之妙。

心力:

1、谈判过程中，对方可能采取极限施压的策略：如短时间内提供大量包含数据的证据（尽管数据可能存在很多误差），以达到击溃对手的心理防线。

2、对方可能采取强势态度，以强势的命令和呵责的口吻进行沟通，以达到击溃对手的心理防线。

应对策略:

1、极限思考：极限预测谈判的结果，最好和最差的。最好的结果给予自己坚持的勇气和力量。最坏的结果从量化的数据层面分析自己能接受的止损边界。因为博弈的过程，心理会承受很大压力，所以不宜持久战，应速战速决。我认为，这里与芒格先生所提的：如果我知道我会死在哪里，我一定不会去哪里思想很相似，当我无法规避这个“麻烦”的时候，我可以选择在过程中，明确自己的风险边界和承受范围，如果超出就及时止损。

2、分析清楚具体的目标后，专注于自己一侧的逻辑，不预想对方的行动。在对方出现意图或即将出牌时，再拆解对方的出牌意图。因为在证据不足的情况下，揣测对方意图会增加焦虑，也会增加在无意义或不确定的事情上浪费的时间。

对于知行合一的思考，我认为知行合一不是什么都不做，什么都规避。而是做而后修正，而后再做，而后再修正，在不断的【行】中，修正【知】，最后达到以【知】指导【行】。在面对“麻烦”处理“麻烦”的过程中，会恐惧、失败、不确定，对心理建设的考验，然后经历过之后，便会成为人生中不可多得的财富。