



Image
Tokens



NVIDIA COSMOS

世界模型深度解读



ZOMI

COSMOS

视频目录大纲

1. 什么是世界模型 What is WFM (World Foundation Model)
2. COSMOS NV官网解读
3. COSMOS 技术论文分析
4. COSMOS 对产业影响与思考



COSMOS 为什么是世界模型？

- COSMOS 不就是一个 SORA 视频生成模型吗？有什么好值得在 CES 发布会上讲的？



01

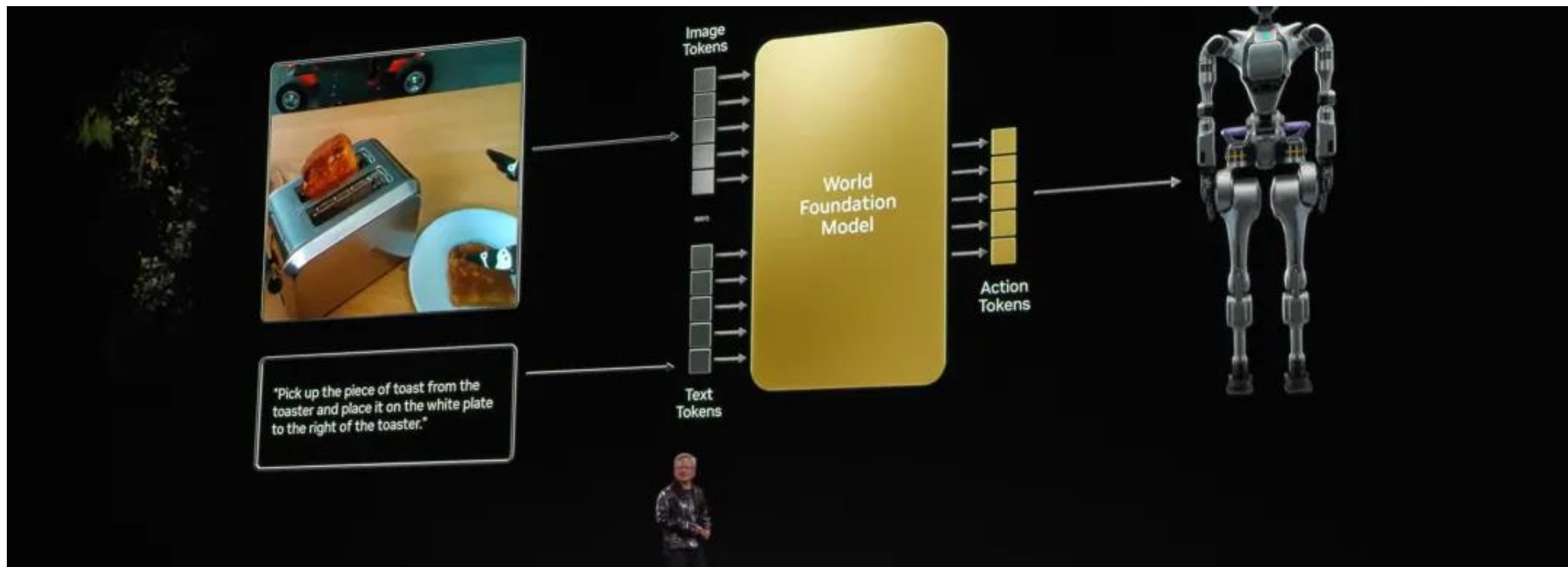
什么是世界模型？

WFM



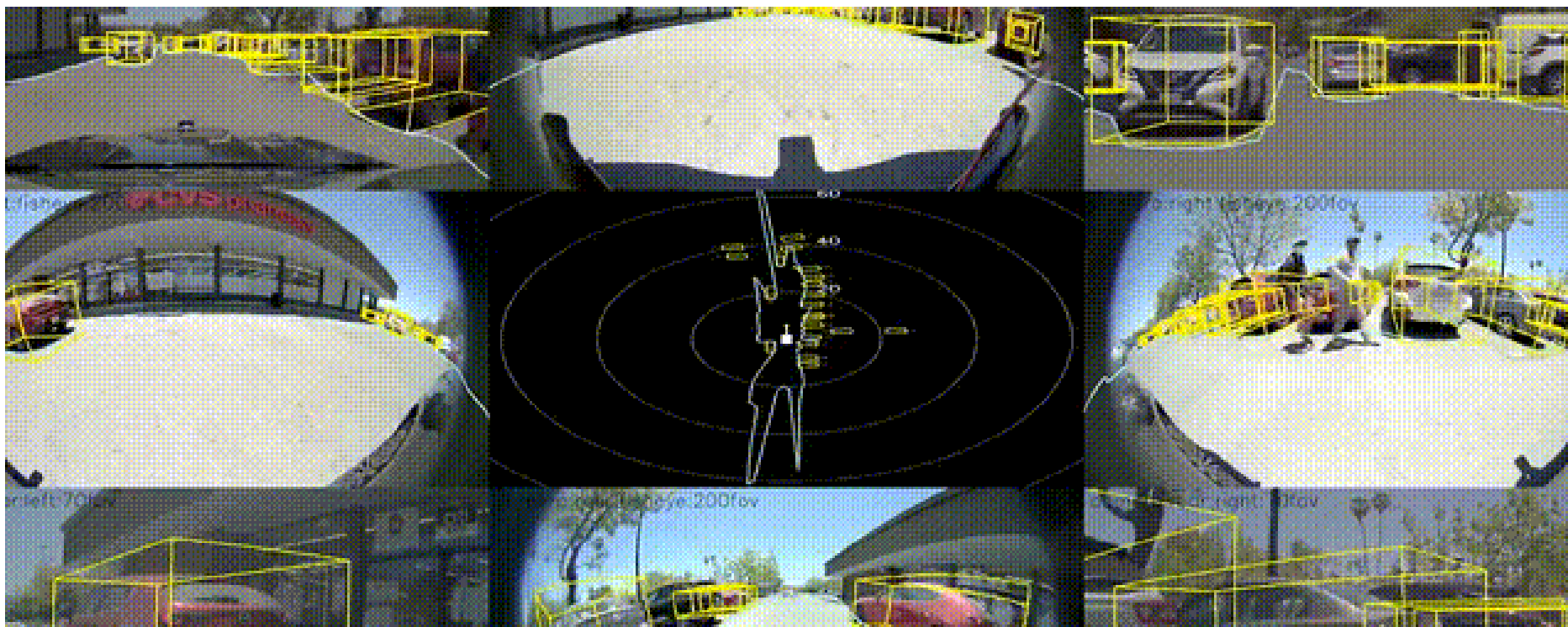
WFM

- LLM 提示一次生成一个 token，仅限内容 token 输出。若要实现从「内容token」转向「动作token」的输出，LLM 无法满足。我们需要的是，一个能够理解物理世界的模型，简之WFM。



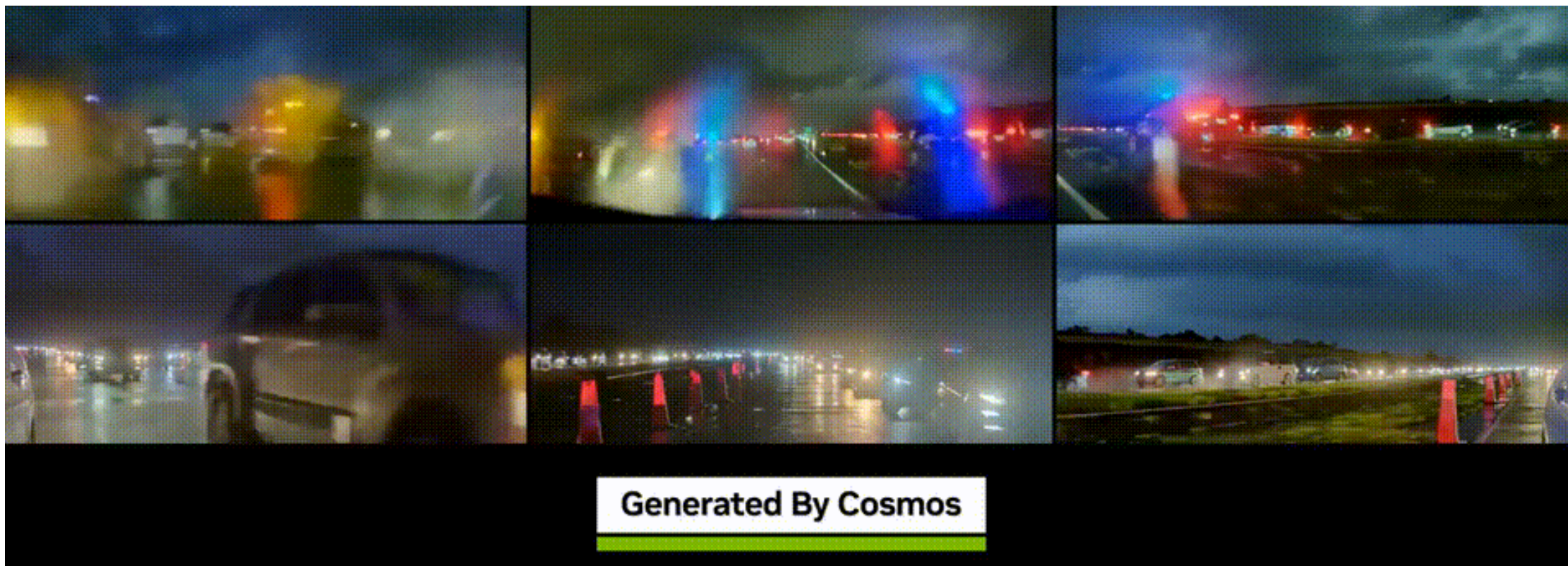
WFM

- WFM 可以从文本/图像输入数据，生成详细的视频，并通过将其的当前状态（图像/视频）与动作（提示/控制信号）相结合来**预测场景的演变**。WFM 能够想象许多不同的环境，并可以模拟未来，帮助物理AI开发者做出更好的决策。



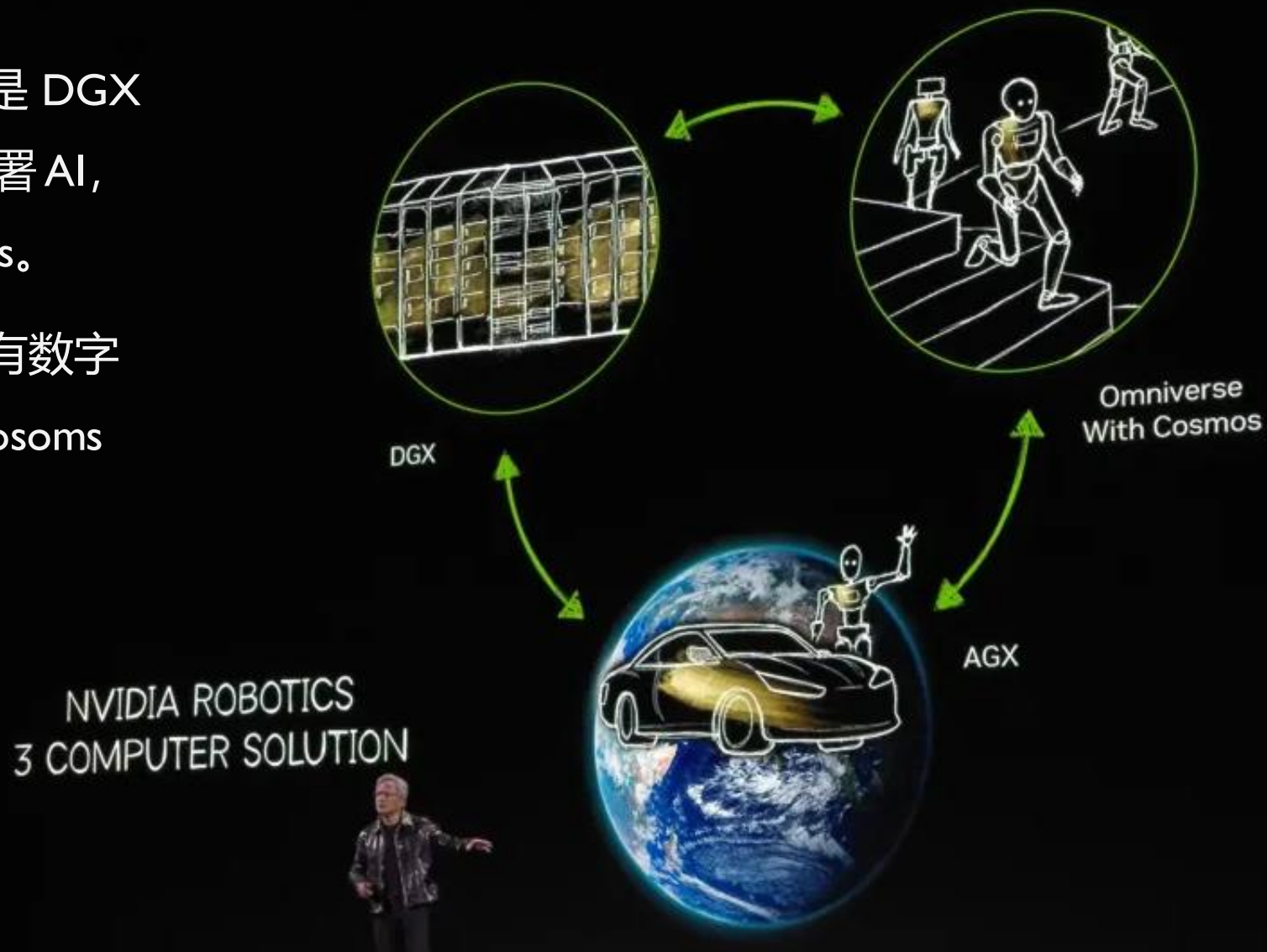
WFM

- 有了 WFM 模拟的 3D 环境，研究者就可以在受控环境中训练和测试物理 AI 系统，物理模拟世界的一切。



WFM

- 未来世界「三台计算机」：一台是 DGX 用来训练AI，另一台 AGX用来部署 AI，最后一台便是 Omniverse+Cosmos。
- 老黄认为：未来，每一个工厂都有数字孪生，你可以将 Omniverse 和 Cosoms 结合，生成一大批未来场景。



02

COSMOS NV官网 解读



COSMOS 形式

- **两种形式**：扩散模型（生成连续的 token）；自回归模型（生成离散的 token）
- **两种生成模式**：文本->视频；文本+视频->视频
- **三种规格模型**：Nano、Super、Ultra

Cosmos Nano

Super low-latency, real-time models optimized for deploying at the edge

Cosmos Super

Highly performant baseline models for out-of-the-box fine-tuning and deployment

Cosmos Ultra

Maximum-accuracy and quality, provides best-fidelity knowledge transfer for distilling custom models



COSMOS 形式

- Diffusion models: 将生成问题分解成一系列去噪任务。
- Autoregressive models: 将问题作为一系列下一个 Token 预测任务来解决。



03

COSMOS 技术论文 分析



04

COSMOS 对产业影响与思考

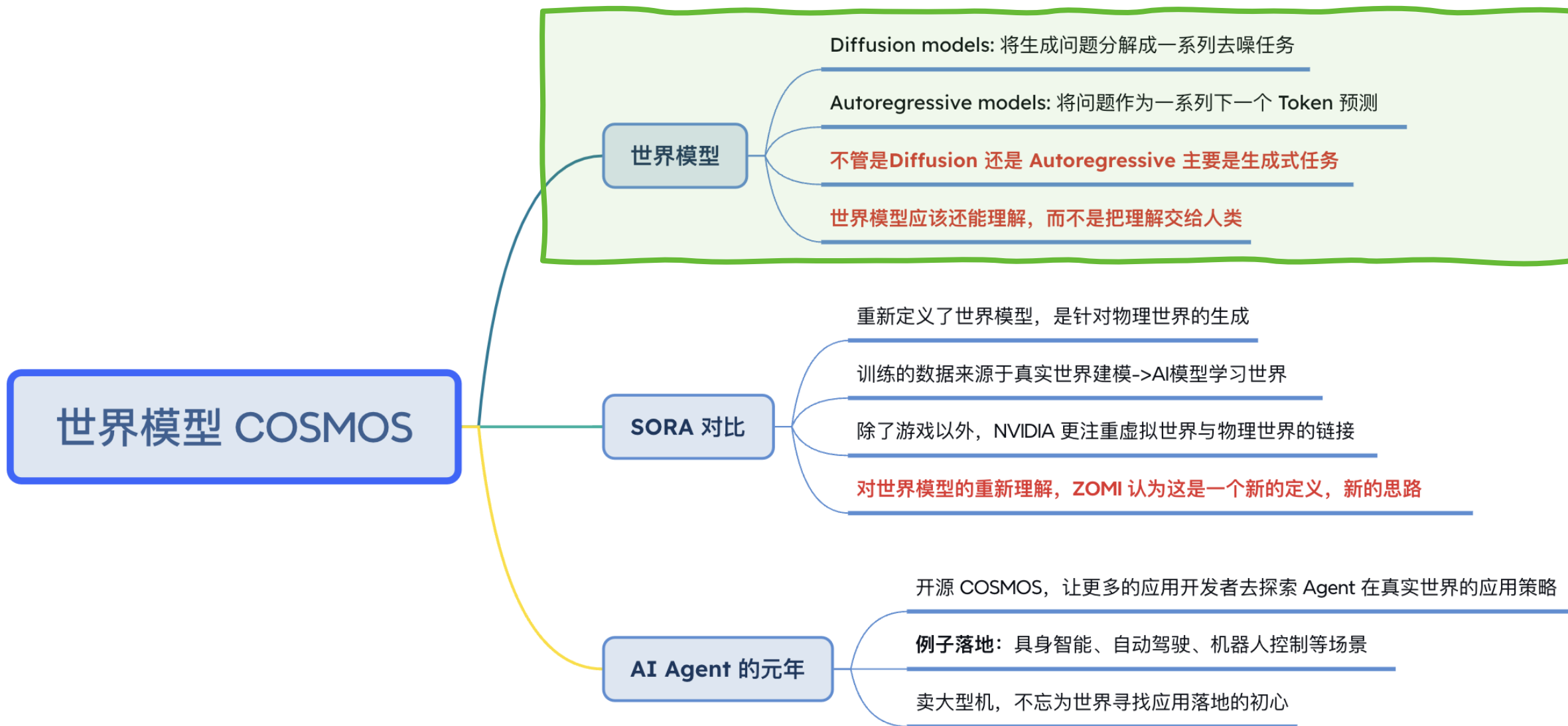


COSMOS 为什么是世界模型？

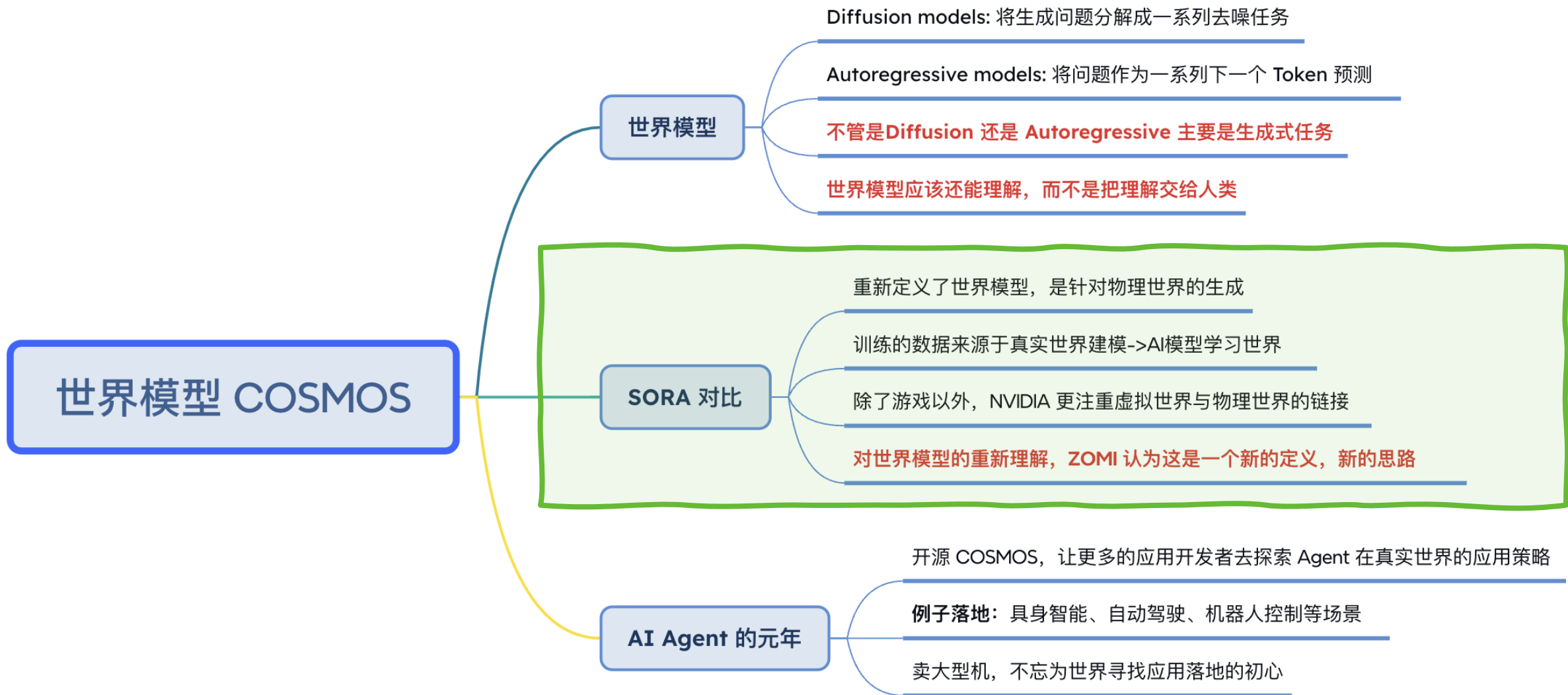
- COSMOS 不就是一个 SORA 视频生成模型吗？有什么好值得在 CES 发布会上讲的？
- **COSMOS 不是 SORA，COSMOS 是世界模型**



总结与思考



总结与思考



总结与思考





Thank you

把AI系统带入每个开发者、每个家庭、
每个组织，构建万物互联的智能世界

Bring AI System to every person, home and
organization for a fully connected,
intelligent world.

Copyright © 2024 XXX Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. XXX may change the information at any time without notice.



GitHub <https://github.com/chenzomi12/AIFoundation>

Github 链接

- <https://github.com/NVIDIA/Cosmos>
- <https://github.com/orgs/NVIDIA/repositories?q=cosmos>

