

大模型时代： 生成式AI发展与科技创新范式



刘志毅

人工智能科学家

- 人工智能科学家，现为上海人工智能实验室应用研究部高级研究员，上海交大安泰AI与营销研究中心特聘研究员，国家人工智能标准总体组专家，AIIA联盟可信AI专家委员会委员，上海开源信息技术协会AI伦理与治理专委会主任，上海人工智能技术协会专家委员。
- 主要研究领域包括人工智能、智能经济以及数字经济学等。

个人学术专著



《黑镜与秩序》

2022年7月清华大学出版社正式出版，十余位公共管理、法学、计算机、经济金融的跨领域领军学者推荐



《智能经济：用数字经济学思维理解世界》

国内首部数字经济领域的技术思想专著



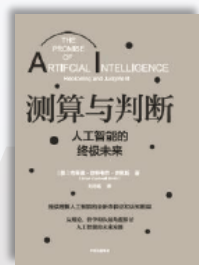
《数字经济学》

重大学术影响专著，数字经济学学科的开拓之作



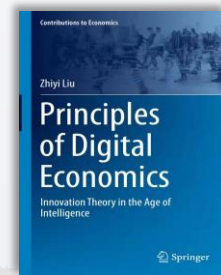
《AI与区块链智能》

2019年12月正式发布，国内首部讨论AI与区块链产业融合的教育部高等教育教材



《测算与判断：人工智能的终极未来》

作为中文版本唯一译者进行了全文翻译工作，于2022年8月中信出版社正式出版



《Principles of Digital Economics》

第一本全英文数字经济学的学术理论专著，由Springer Nature出版并单独成系列

ChatGPT的成功源于NLP领域70年发展的长期积累

NLP领域的发展为大语言模型（Foundation Models）提供了坚实的基础。

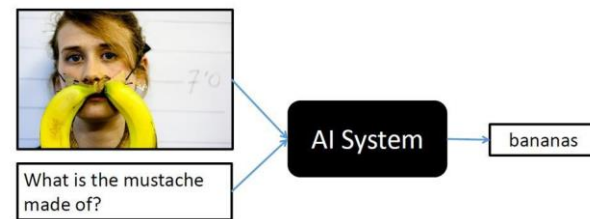


Gary Marcus
现代机器人之父
纽约大学教授

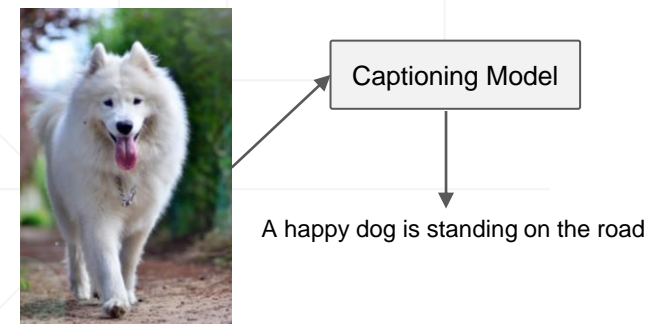
Lecun
Meta AI的负责人
图灵奖得主

LeCun指出：ChatGPT本质上是一种大语言模型，并指出：

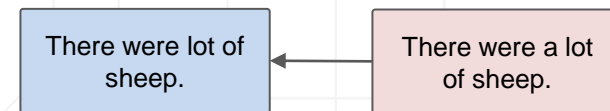
“大型语言模型其实是一条‘邪路’”。就基础技术而言，ChatGPT并不是特别有创新性，这不是什么革命性的东西。Yann LeCun和Gary Marcus对此达成了空前一致。



视觉问答 Visual Question Answering



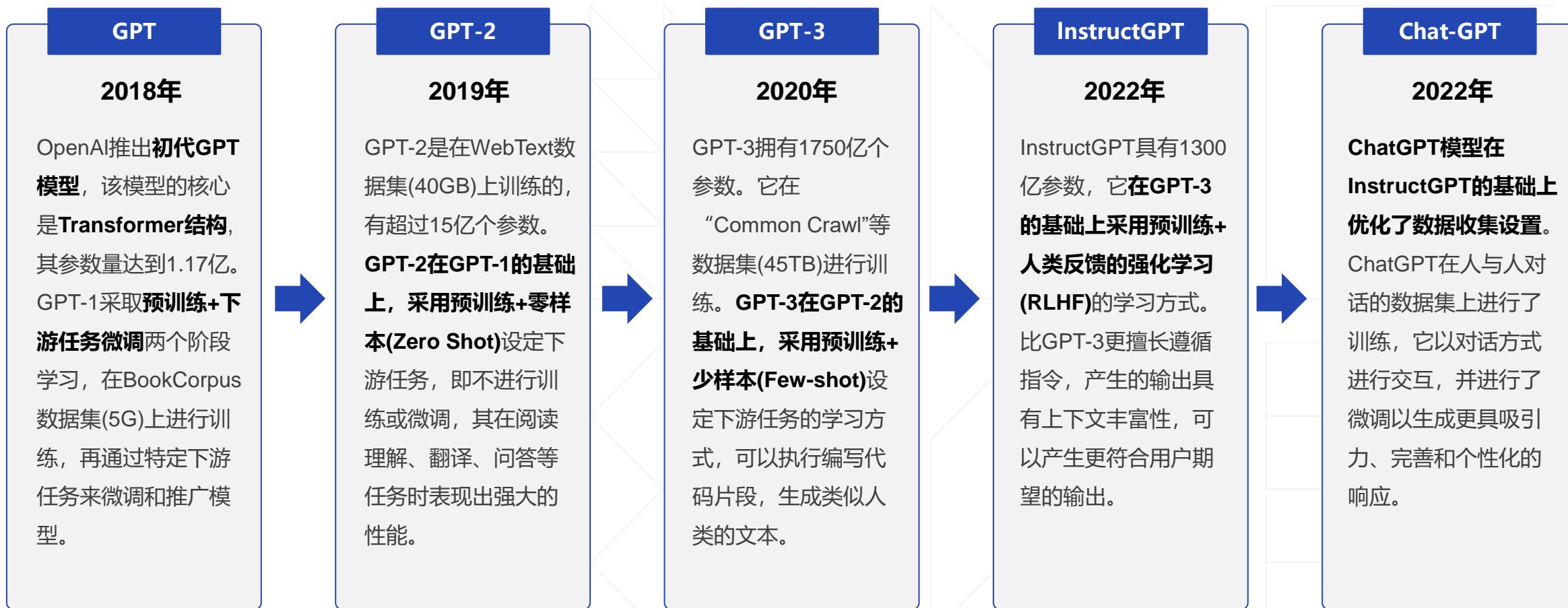
图像标注 Image captioning



语法纠正 Grammatical Error Correction

ChatGPT的发展历程

- ChatGPT在**海量的文本数据**上进行了预训练，可以对自然语言输入**产生类似人类的回答**。
- ChatGPT其核心是建立在**GPT模型** (Generative Pre-trained Transformer) 的发展基础之上。



ChatGPT的技术突破

ChatGPT之所以能够突破，是因为采用了一系列深度学习的新技术，包括无监督学习、有监督学习、多任务学习、小样本学习、自监督学习以及基于人类反馈的强化学习。



除采用一系列深度学习新技术外，ChatGPT还有以下技术特点。

- **基于Transformer模型的架构：**更好的语言建模能力和更高的并行性。
- **预训练与微调：**先在大规模语料库上进行无监督的预训练，再在特定任务上进行微调，可以使得模型具有更好的泛化能力和更高的准确率。
- **多层架构：**模型由多个Transformer层组成，每个层都可以学习不同的特征和语义，从而提高模型的表现。
- **动态扩展：**在预训练过程中逐渐增加模型的层数和参数量，以逐步提高模型的性能。
- **大规模语料库：**使用了大规模语料库进行训练，如维基百科等，使得模型具有更好的语言理解能力和更高的泛化能力。

大模型的定义与价值

市场对于超大模型的看法

厂商宣传中常提到千亿级、万亿级参数为大模型、超大模型。实际在产业落地阶段，超级预训练大模型主要具备以下特征：

- ① 预训练模型学习过海量数据
- ② 用户端模型二次开发仅需少量数据
- ③ 相比上一代的算法模型有更高的预测准确度

大模型的定义

是对原有算法模型的技术升级，基于海量数据开发预训练模型，到最终用户环境中使用少量数据即可获得比之前的算法模型更好的结果

要点：1) 预训练型学习了多少数据，2) 迁移学习的效果如何，3) 部署的整体成本如何
并不是所有大模型参数都参与训练，单体模型的参数会全部参与训练，稀疏模型中只有部分参数参与运算

大模型的价值

开发门槛的降低：用户可以在AI原子能力的基础上实现自定义模型，无需关注模型细节

计算资源的降低：预训练由厂商提供，用户侧只需进行小数据量的再训练

预测效果的提升：研究与实践均表明，使用大模型的效果比之前的模型效果相比，均有不同程度的提升

行业应用

设备知识库
智能问数

智能客服
知识搜索

线路巡检
安全生产
.....

原子能力

智能对话
智能创作

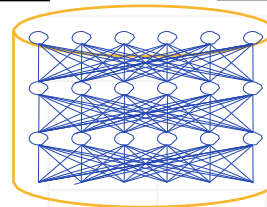
智能问答
文档理解

目标监测
图像分割
.....

领域相关的少量数据

微调

应用



领域无关的海量数据

自监督学习

文本 知识 图像 视频

市场上可见的大模型日益丰富

厂商	大模型	厂商	大模型	厂商	大模型
 Meta	LlaMA	 字节跳动	—	 微软、微软亚洲研究院	Megatron Turing (&英伟达)、NUWA(&北大)
 google	PaLM	 OpenGVLab (商汤平台)	基于书生的超高效预训练模型	 北京智源人工智能研究院	悟道2.0
 商汤科技 (&上交等)	书生 (INTERN)	 清华大学	GLM-1308、CokeBert (&腾讯)	 京东	—
 北京大学	鹏程·盘古α (&鹏程&华为) (资料少, 不确定含北大)、NUWA (微软亚洲)	 DeepMind	Gopher	 OPPO	OBERT
 NVIDIA 英伟达	NeMo LLM	 粤港澳大湾区数字经济研究院 (IDEA研究院)	二郎神	 中国科学院	紫东太初
 浪潮	源1.0智能大模型	 美团	—	 华为云	鹏程·盘古α (&鹏程&华为)
 奇点智源	SkyCode多语言开源编程大模型	 澜舟科技	孟子	 腾讯	混元AI大模型
 追一科技	RoFormerV2	 快手	—	 百度	文心一言
 鹏程实验室	鹏程·盘古α (&北大&华为)	 OpenAI	GPT-4	 云从科技	SemBERT (&上海交通大学)
 阿里巴巴达摩院智能计算实验室	M6、“通义”大模型	 联汇科技	OmModel		

核心技术层协同发展，共同赋能 AI 大模型生态

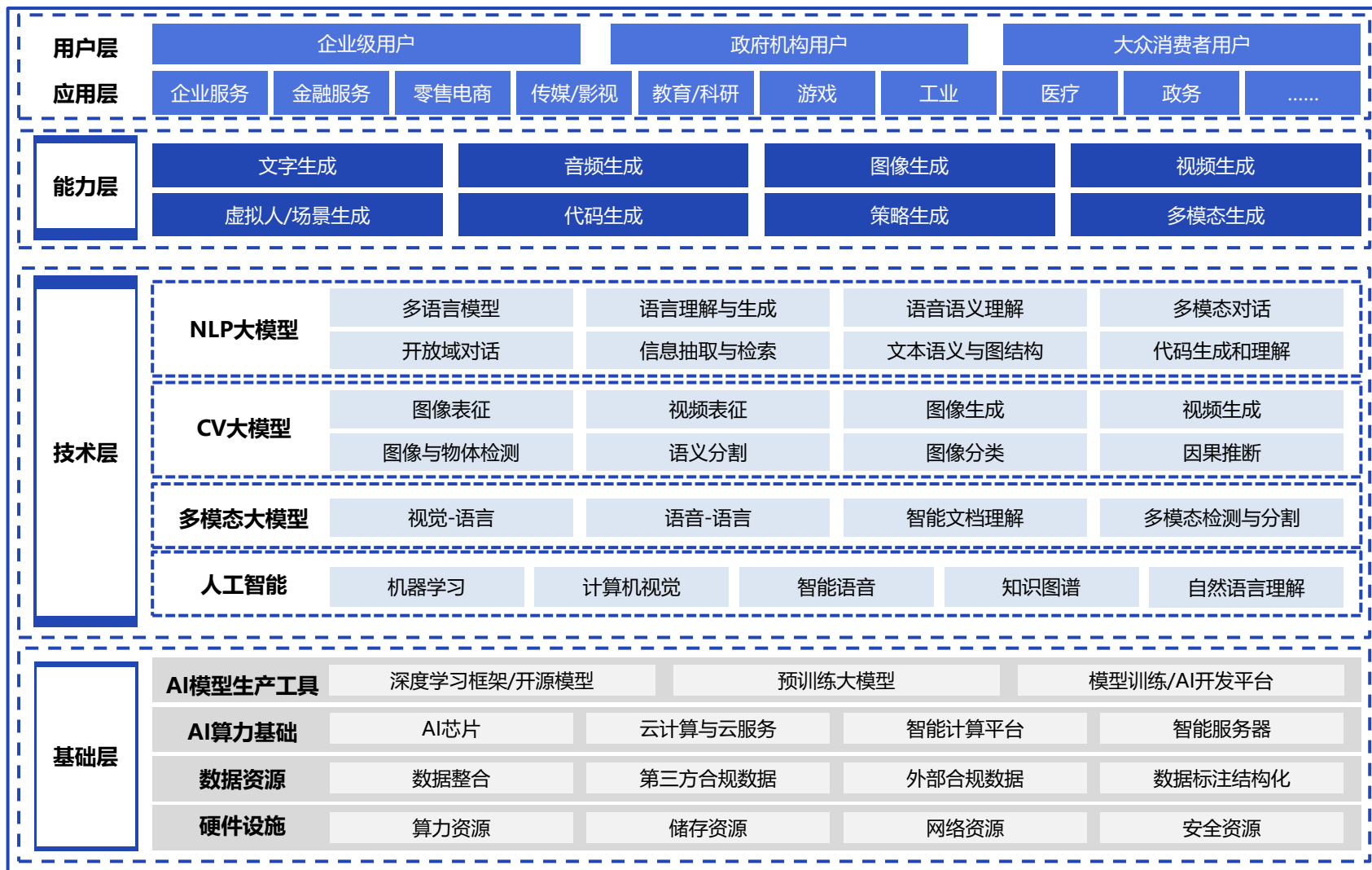
AI 大模型的技术架构通常涉及多个层次，可以分为基础层、技术层、能力层、应用层、终端层五大板块，其中核心技术层涵盖 AI 技术群和大模型的融合创新，为各行业深度赋能。

□ **基础层：** AI 大模型的基础层涉及硬件基础设施和数据、算力、算法模型三大核心要素。随着 AI 大模型规模的不断扩大，对计算资源的需求也在增加。因此，高性能的硬件设备、海量场景数据、强大的算力基础和升级迭代的算法模型成为了支持 AI 大模型发展的关键。深度学习模型的不断升级和迭代，增强了 AI 算法的学习能力；同时，开源模式将使 AI 大模型成为海量应用、网络和服务的基础。

□ **技术层：** AI 大模型的技术层主要涉及模型构建。目前，Transformer 架构在 AI 大模型领域占据主导地位，如 BERT、GPT 系列等。AI 大模型包括 NLP 大模型、CV 大模型、多模态大模型等。这些模型采用预训练和微调的策略，先在大量无标注数据上学习语言或图像的基本表示，然后针对特定任务进行微调。

□ **能力层、应用层及用户层：** 在基础层和技术层的支持下，AI 大模型拥有了文字、音频、图像、视频、代码、策略、多模态生成能力等，具体应用于金融、电商、传媒、教育、游戏、医疗、工业、政务等多个领域，为企业级用户、政府机构用户、大众消费者用户提供产品和服务。

AI 大模型的技术架构



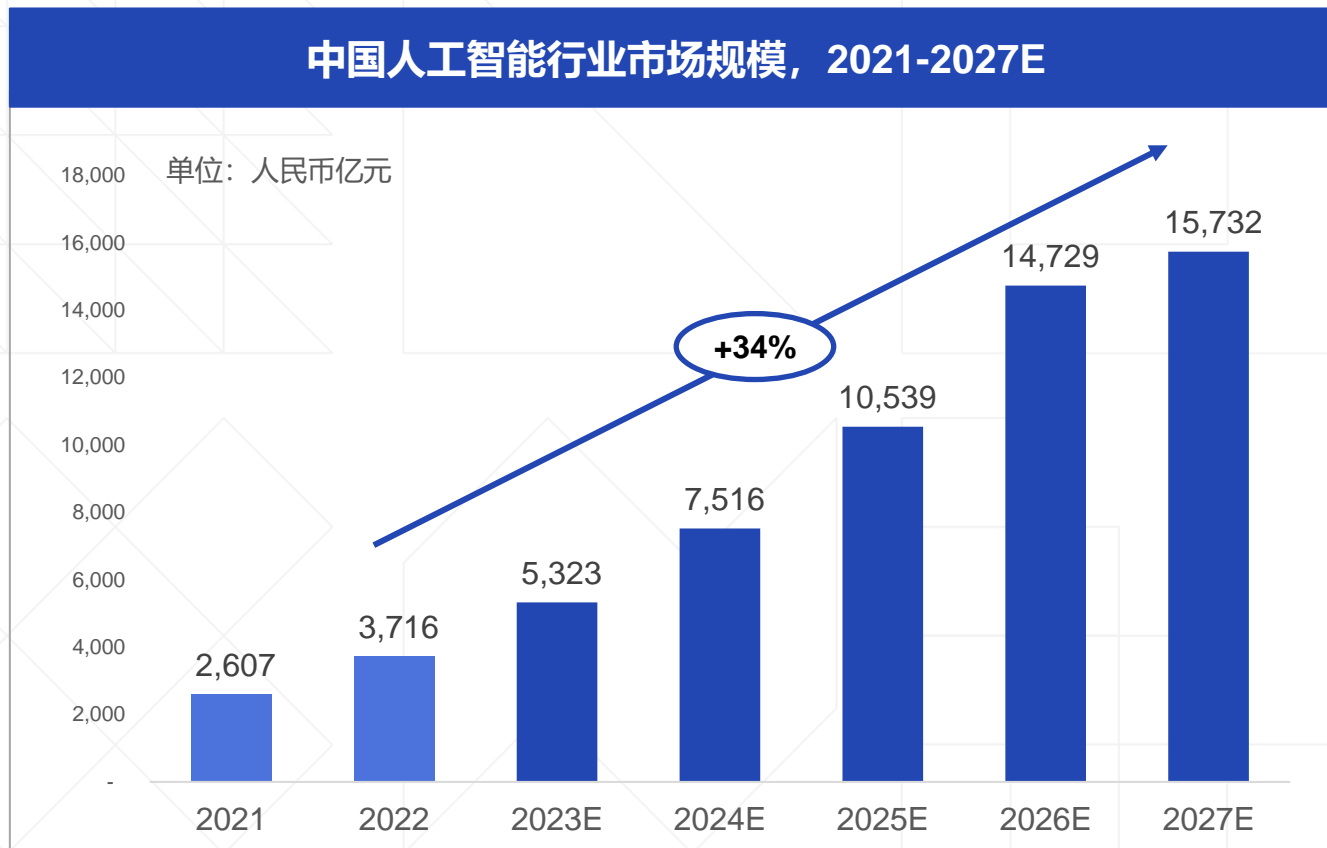
AI 市场高景气，大模型下游行业需求旺盛

- 人工智能能够赋能经济社会发展各领域，下游各个领域的产业升级对大模型的需求持续走高。例如，在泛交通领域，人工智能能够在自动驾驶、智能座舱、运行管理优化等多个方面为汽车行业创造价值；在医疗方面，人工智能可以提高疾病检测的效率以及医学影像领域的智能化分析。据测算，2022年中国人工智能行业市场规模为3,716亿人民币，预计2027年将达到15,372亿人民币，有望在下游制造、交通、金融、医疗等多领域不断渗透，实现大规模落地应用。
- 下游行业对人工智能需求呈现出碎片化、多样化的特点，从开发、精调、优化、迭代到应用都需要投入巨大的人力和财力，成本极高。而大模型能够向外赋能，包括通过开放API的形式，降低AI应用开发门槛，提高落地部署效率和精度等，进而降低AI规模化部署的成本，满足各行业场景的应用需求，进一步推动人工智能进入工业化发展阶段。

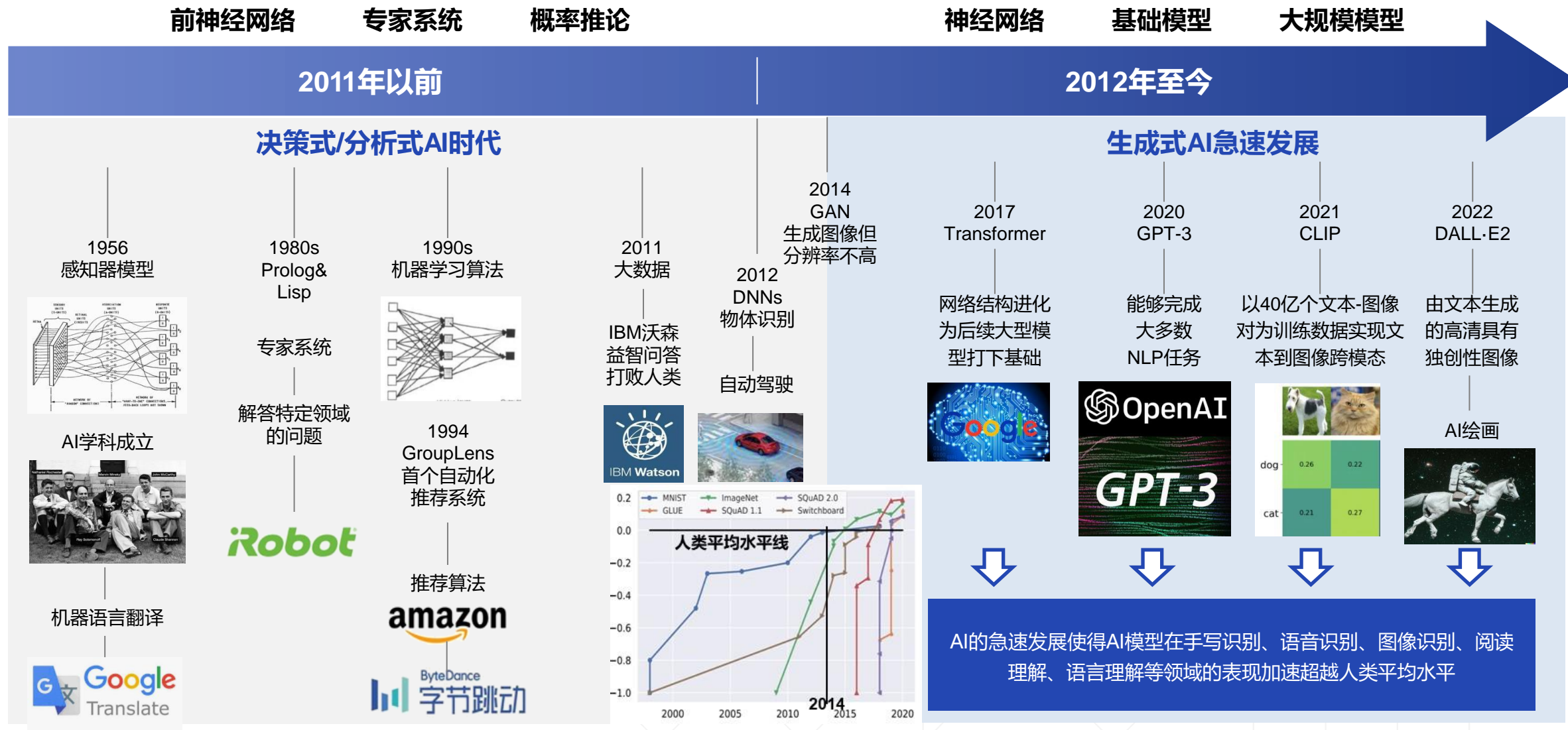
AI 大模型应用场景丰富



中国人工智能行业市场规模，2021-2027E



从决策到生成，AI技术与应用迎来跨越发展



AIGC当前典型应用：数字内容生成

文生文、文生图、文生视频

01

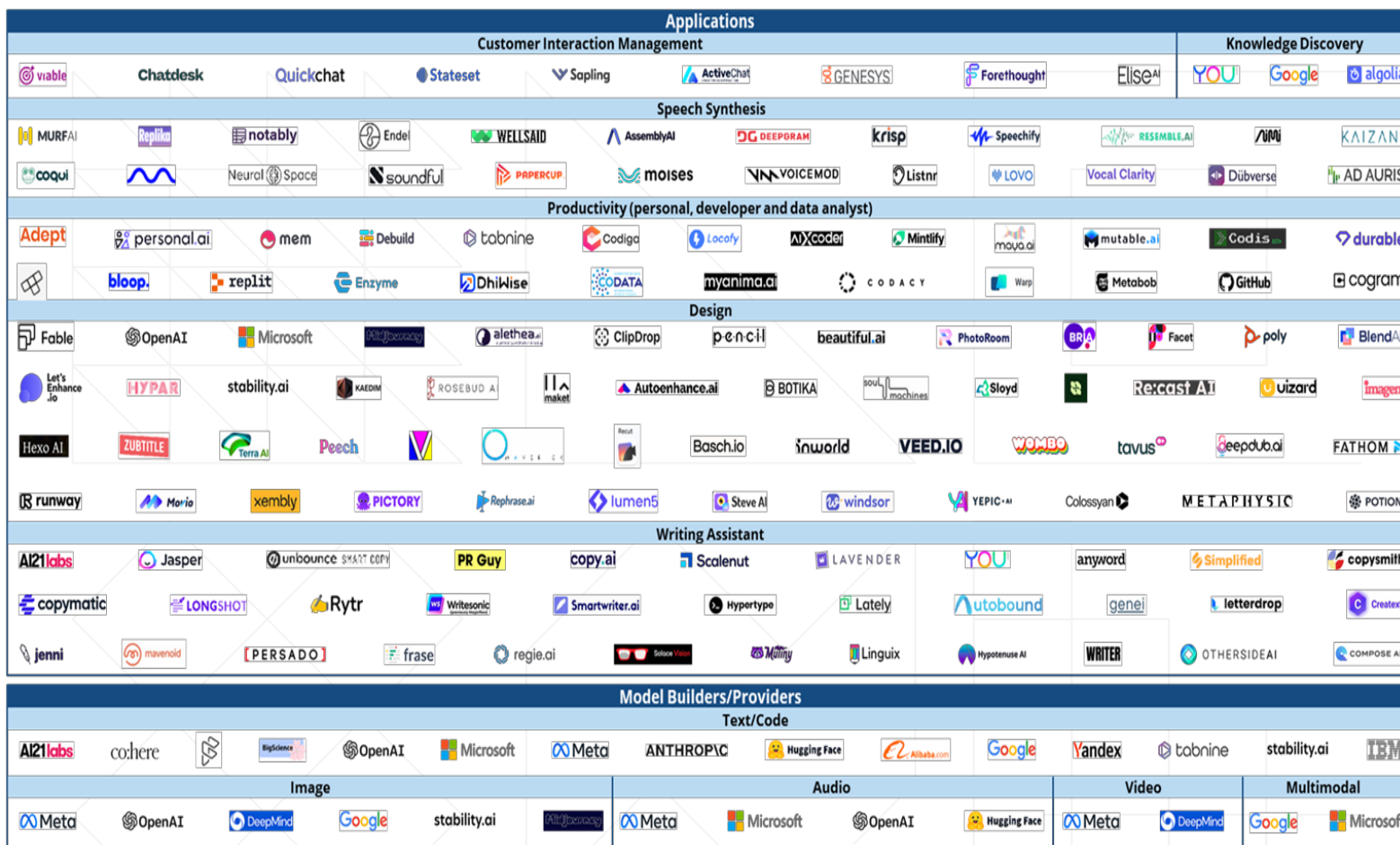
典型的AIGC基础模型

- OpenAI的GPT-3 (大型语言模型[LLM])、DALL-E2 (图像模型), 以及Whisper (语音识别模型)
- Stability.ai的stable Diffusion (图像模型)
- Midjourney (图像模型)
- 谷歌的PaLM和LaMDA (对话应用的语言模型) (大型语言模型)
- Meta的NLLB (大型语言模型)、Make-a-video (视频模型) 和data2vec (多模态)
- 微软的DeepNet (大型语言模型) 和 MetaLM (多模态)

落地场景

- 数字人/虚拟助理
- 游戏互娱
- 美术设计
- 产品海报设计

02




生成式AI企业应用案例：知识管理&代码生成

知识管理类应用

 **美国跨国投资管理和服务公司&OpenAI的GPT-3**
对财富管理内容的培训进行微调，以便金融顾问能够搜索公司中的现有知识，并为客户轻松创建定制内容。

 **Lucy将生成式AI用于知识管理：Lucy Synopsis**

 **百事公司、Haleon (原GSK) 和Saatchi & Saatchi等组织**
采用其生成式AI来阅读组织中的信息，提供独特的两到三句话摘要回应，直接准确地回答用户的问题。

 **谷歌发布Bard：** 由对话应用语言模型驱动的实验性服务。


 **微软宣布新的Bing和Edge——有限预览阶段。**


 **Yext使用生成式人工智能，为其知识图谱引入自动内容生成。**

C3.ai宣布推出C3 Generative AI for Enterprise Search：
使用自然语言界面快速定位、检索和呈现企业信息系统整个语料库中的所有相关数据。

Viable聚合来自服务台票据、调查、CRM和评论的定性数据，以帮助团队了解客户对其产品的看法和感受客户反馈的准确性从66%提高到90%。

代码生成

 **GPT-3的Codex程序：**
一种机器学习模型，专门为代码生成而训练，可以产生各种不同语言的代码。

 **微软Copilot：**
从注释和代码中提取上下文，能力包括：
1) 代码翻译，2) 代码自动完成，3) 从自然语言到代码。

 **Amazon CodeWhisperer：**
人工智能编程助手，理解用户的初始提示来自动完成软件代码。

用户案例：
在一家行业领先的审计、咨询、税务和顾问服务机构的六周试点中，有55名开发人员参加，生成式AI生成的代码准确性为65%，代码开发速度提高20%。

生成式AI企业应用案例：销售营销&其他

销售和营销



Microsoft Viva Sales：帮助卖家更有效地与潜在客户和顾客沟通

将为各种场景生成建议的电子邮件内容，如回复询问或创建提案，并附上与收件人具体相关数据，如定价、促销和截止日期。通过可定制内容的自动建议，卖家可以花更少的时间编写电子邮件和从同事或数据库中搜索销售数据。



Jasper：

可以制作博客、社交媒体帖子、网络副本、销售电子邮件、广告和其他类型的面向客户的内容。



Rephrase.ai Synthesia：

可以制作博客、社交媒体帖子、网络副本、销售电子邮件、广告和其他类型的面向客户的内容。



Adobe Express：

通过提示生成模板，或根据描述创建独特的文本效果。目前主要集中在图像方面，未来将包括视频、3D设计、纹理创作、标志设计等等。



Salesforce Einstein GPT：

在销售、服务、营销、商业和IT互动中，超大规模提供人工智能创造的内容。



Slack新ChatGPT：

提供即时对话摘要、研究工具和写作协助。

特定行业用例

医疗领域

- **改善医生的工作流程：**生成式AI可以审查医疗记录，并在护理点提供与病人和治疗医生相关的简明总结。
- **Profluent：**使用生成式人工智能来设计蛋白质。
- **Absci——生成式人工智能药物创造公司：**试图在计算机上“从头”设计新的抗体

金融业

- **投资组合优化**
- **知识管理**





制造业

- **通用汽车：**与欧特克工程师合作，为座椅支架创造了150个新的设计理念，并选择了比原始组件轻40%和强度20%的最终设计。
- **Armour：**利用生成式设计算法为所有类型的运动训练打造出具有灵活性和稳定性最佳组合的鞋子。
- **菲利普·斯塔克和卡特尔使用生成算法将美学和功能结合起来：**我们如何才能用最少的材料？设计师和人工智能合作产生一种非常规但高效的椅子设计。

AI 大模型技术创新，助推生成式 AI 应用场景加速落地

- **伴随 AI 技术升级和大模型成熟，AI 绘画与 ChatGPT 的成功破圈，生成式 AI 技术迎来发展拐点，行业关注度大幅提升。** 生成式 AI 是指基于大模型、生成对抗网络 GAN 等人工智能技术，通过已有数据寻找规律，并通过适当的泛化能力生成相关内容的技术，可生成如图像、文本、音频、视频等原创内容变体。例如，以 ChatGPT、Midjourney、文心一格、商汤商量、Codex 为代表的生成式 AI 应用拥有文本语言理解能力、涌现能力以及思维链推理能力，能够完成文学创作、新闻写作、数理逻辑推算、代码生成、图片生成等多项任务。目前，国内电商、游戏、文娱、设计等行业正在积极使用相关的生成式 AI 应用来提高自身工作效率，尤其以文生图应用为主。
- **生成式 AI 不仅能够增强并加速下游多领域的设计，而且有潜力“发明”人类可能错过的 redesign、新对象。** 生成式 AI 有生成大规模、高质量、低成本内容优势，在算力和算法支持下生成大量内容，生成的内容质量将持续超越 UGC 与 PGC。未来有望为各行业提供内容支持并促进其内容繁荣，最大化释放内容生产力。
- **文字生成属于发展成熟、易于跨界转化的赛道，而跨模态生成赛道的发展潜力最高。** 生成式 AI 应用根据模态划分为文字生成、音频生成、图像生成、视频生成、跨模态生成。语音合成、文本生成、图像属性编辑等技术应用目前较为成熟，跨模态生成、策略生成是高增长潜力的应用场景，在自动驾驶、机器人控制等领域有极高应用价值，随着未来技术不断发展成熟，预计 3-5 年可实现稳定落地。

大模型发展及相关应用实现落地时间表

	2020 年前	2020	2022	2025?	2030?	2050?
 文字	垃圾邮件识别 翻译 基础问答	基础文案写作 起草初稿	更长的文字 完成第二稿	垂直微调 科学论文	高于人类平均水 平的终稿写作	产出比职业作家 写得更好的终稿
 代码 生成	单行代码 自动完成	多行代码生成	更长的代码 更高的准确率	多程序语言 更多垂直领域	文本到产品 (草稿)	产出比全职开发人 员做得更好的文本到产品 (终 稿)
 图像			艺术作品 Logo 设计 摄影	产品设计模型 建筑概念模型	产品设计模型 建筑模型	产出比专业艺术家、设计师、 摄影师做得更好的终稿
 视频/ 3D/游戏				基础/初稿视 频及 3D 文件	第二稿	AI Roblox 产出基于个性化梦 想的电子游戏和电影

可用的大模型：
初步探索 基本形成 准备阶段

大模型迎来发展新机遇，未来前景可期

关键发现

- 人工智能的政策引导逐渐覆盖到大模型生态，并开始出台相应的生成式 AI 监管建议，进一步支撑大模型生态的有序发展；
- 大模型的神经网络架构和训练大模型的 AI 基础设施，均逐渐发展成熟，推动大模型的生产更加系统化和工程化；
- 下游企业用户的 AI 部署需求进一步规模化发展，急需在上游大模型支撑下，获得 AI 应用开发门槛降低，部署精度提高等基础价值，进而降低 AI 规模化部署的成本；
- 大模型的发展趋于通用化与专用化并进，平台化与简易化并进；
- 依托 Model as a Service，大模型建立起面向政企、消费者群体等差异化的商业模式，并逐渐形成基模型、领域、行业大模型一体的商业化架构。

在“基础设施支撑 + 顶层设计优化 + 下游需求旺盛”三轮驱动下，AI 大模型迎来了良好的发展契机

通过数据交互和任务反馈，优秀的大模型能够赋能各行各业开放任务，满足对未来AI 应用的期待。展望未来，大模型“训练基础设施 - 底层技术 - 基础应用 - 垂直应用”发展路线逐渐清晰，随着底层技术逐步革新，基模型和领域大模型持续完善，大模型应用边界不断拓宽，将加速赋能交通、医疗、金融等各个行业和领域，引发一场以强人工智能和通用人工智能为代表的新一轮智能革命浪潮，大幅提高生产和生活效率，带来深刻的经济、社会和产业变革。



大模型的多种价值，将加速人工智能的技术进步和规模化应用

AI 大模型具有降低开发门槛、提高模型精度和泛化能力、提高内容生成质量和效率等多种价值，实现了对传统 AI 技术的突破。

一方面，大模型可以帮助降低机器学习和自然语言处理应用的开发门槛，能够对复杂的模式和规律进行更准确的建模，通过不断地学习和更新自己的参数来提高其性能和准确度，提高模型的精度，更好地泛化到新的数据集和任务中。另一方面，大模型通常能够更好地泛化到新的数据集和任务中，可以提高内容生成质量和效率，例如生成对话、摘要、翻译等。除此之外，大模型的开源性和可复制性可以促进学术研究的发展和技术的普及，增强生态繁荣度，从而加速人工智能技术的进步和应用。

大模型的五大基本价值

增强模型泛化能力

大模型通过学习大量数据和任务获得广泛知识，捕捉更多细节，更好地泛化到新的数据集和任务中。



降低开发门槛

大模型通过自动学习更多的特征和规律，减少手动特征工程的需求，使开发者能更轻松地构建高质量的模型。



提高内容生成质量和效率

大模型可以更好地捕捉和分析信息，用于生成高质量的内容，提升生成效率，例如生成对话、摘要、翻译等。



大模型的基本价值

提高模型精度

大模型拥有更多的参数和更深的层次结构，能对复杂的模式和规律进行准确建模，并通过不断学习和更新自己的参数提高性能和准确度。



增强生态繁荣度

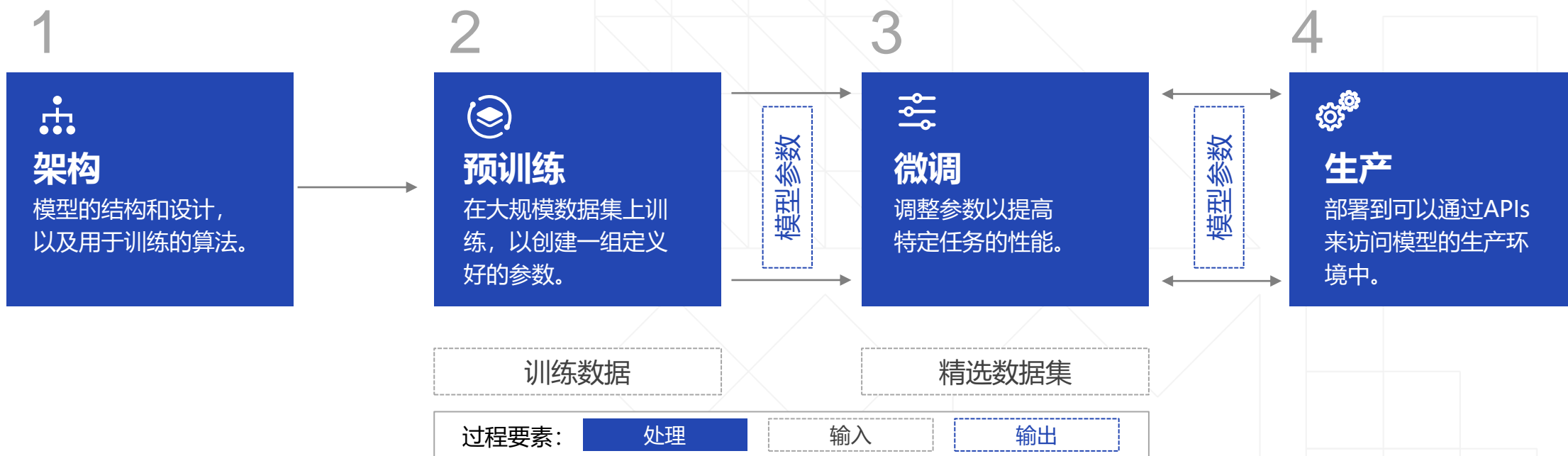
大模型的开源性和可复制性能够聚合开发者、高校、实验室等多方资源，在促进学术研究的发展和普及的同时，加速人工智能技术的进步和应用。



大模型的能力来自于技术工艺的变化

基础模型 (Foundation Model) 是区别生成式人工智能技术栈和之前人工智能的关键。基础模型是斯坦福大学基础模型研究中心创造的术语，它是一个在广泛数据集上预先训练的机器学习模型，可以用于解决一系列问题。基础模型通常经过以下四个阶段的开发，如下所示。

基础模型的开发



✓ 的确，生成式人工智能可以彻底改变企业与客户之间的运作和互动方式，甚至可能重新定义我们对“员工”的认知。在某些消费者和企业领域，这种转型已经在进行中。

生成式人工智能的核心商业逻辑：效率，指导，创作，娱乐

在2022年，OpenAI的DALL·E2因其文本生成图像的能力而引起了全世界的关注。这个模型可以根据简单的文字提示创建图像，有直接的提示，比如“丛林中的一只狮子”，还有更加滑稽有趣的提示，比如“以毕加索风格描绘两只狮子打篮球。”

我们将消费者用例--即个人在其生活中的用例--根据所提供的实用性分为四个大类：

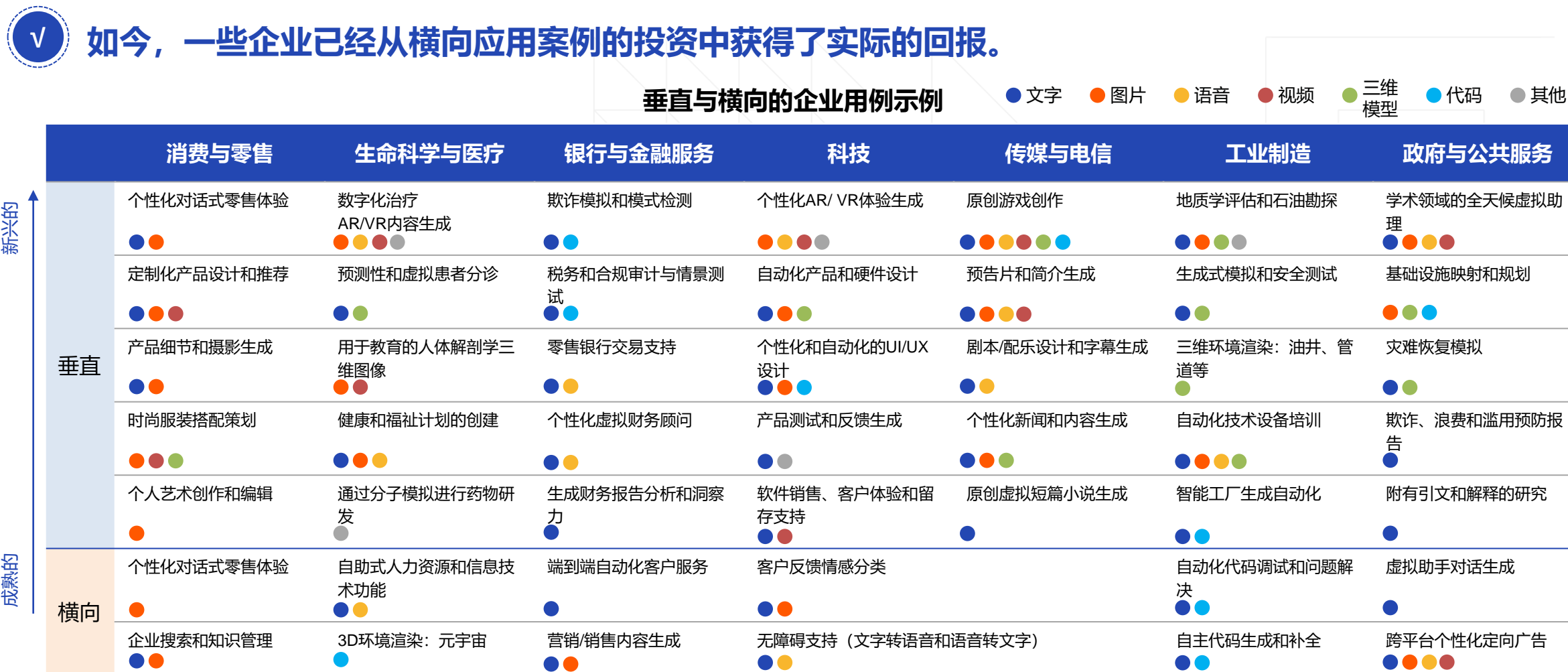
	效率 优化计划、研究和产品开发等任务	指导 提供个性化指导或学习内容	创作 生成或增强内容，复制创意过程	娱乐 建造游戏、虚拟人物和其他娱乐项目
高复杂度	创建健康状况和健康计划	与虚拟伙伴进行交流	生成和编辑视频文件	创建原创游戏
	发现新产品	创建个性化的财务计划	创建室内设计模拟图	与流行文化人物进行聊天
	进行带有引证的研究	教授新语言	策划服装和时尚理念	渲染3D环境
低复杂度	整理内容	综合研究论文	修改和编辑设计文件	重新混音或采样音乐
	回答一般问题	指导和为个人写作提供信息	创作艺术和编辑图像	生成原创的虚构短篇小说
	供应商示例			
	Synthesis.ai Consensus	Grammarly Lingostar.ai	Luminar AI Lensa	Jasper Scenario

这只是市场的初步观察，随着工作的深入，类别之间可能会有重叠。而且，随着未来几代人工智能的发展（例如，那些能够实现多模态交互或完全在单个设备上运行的人工智能），预计会出现新的类别的用例。

生成式人工智能在消费者和企业领域的应用案例

在生成式人工智能的形成时期，最受欢迎的企业应用案例，将是用于推动内部或B2B产出的，通用的或适用于跨行业跨职能（“横向”）的用例。然而，与之前出现的技术一样，通常在特定行业的“垂直”应用案例存在着更可持续的价值创造机会。

如今，一些企业已经从横向应用案例的投资中获得了实际的回报。



中国 AI 大模型主要厂商竞争力评价——厂商总览

国外玩家



**Open AI:
GPT-4**

GPT-4 于 2023 年 3 月 14 日发布, 是千亿级参数的多模态预训练模型, 能够支持图像和文本的输入



**Anthropic:
Claude**

Claude 于 2023 年 3 月 14 日发布, 是 Anthropic 开发的类似 ChatGPT 的聊天机器人, 使用的是 constitutional AI 的偏好模型



**Google:
PaLM-E**

aLM-E 于 2023 年 3 月 10 日发布, 是一种多模态视觉语言模型 (VLM), 具有 5620 亿个参数, 集成了可控制机器人的视觉和语言能力



**Naver:
HyperCLOVA**

与三星电子合作, 于 2021 年 5 月发布, 超过 2000 亿个参数的超大规模模型; 97% 使用的是韩文语料, 计划 23H1 基于大模型推出 Search



**Kakao:
KoGPT**

于 2021 年发布, 专注于开发基于 AI 的图像创建技术和医疗保健技术



**LG:
Exaone**

于 2022 年 12 月发布, 拥有 3000 亿参数, 使用图像和文本数据的多模态模型, 是目前韩国参数规模最大的模型, 用于生物医药和智能制造行业



**Meta: SAM、
LLaMA**

视觉模型 SAM: 于 2023 年 4 月发布, 通过精细标注, 识别分割画面上的指定物体。大型语言模型 LLaMA: 于 2023 年 2 月发布, 包括 4 种参数规模

本土玩家

互联网公司



阿里巴巴: 通义大模型

最新进展: 于 2023 年 4 月正式推出通义千问大模型



华为: 盘古大模型

于 2021 年 4 月正式对外发布盘古预训练大模型



腾讯: 混元大模型

于 2022 年 4 月首次对外披露混元 AI 大模型研发进展



百度: 文心大模型

最新进展: 于 2023 年 3 月正式发布文心一言大模型



快手: K7 大模型

百亿级多模态内容理解大模型 Pluto, 高效视频相似度计算

科技公司



商汤: 日日新 SenseNova 大模型体系

提供自然语言、内容生成、自动化数据标注、自定义模型训练等大模型及能力, 及一系列生成式 AI 应用



昆仑万维: 昆仑天工大模型

于 2023 年 4 月 17 日正式发布千亿级大语言模型“天工”



科大讯飞: 1+N 认知智能大模型

将于 2023 年 5 月 6 日正式发布

硬件公司



浪潮信息: 源 1.0 大模型

于 2021 年 9 月 28 日发布, 具有 2457 亿参数, 是中文语料 AI 模型



澜起科技: 孟子大模型

于 2023 年 3 月 14 日正式发布了“类 ChatGPT”的语言生成模型——孟子 Mchat 可控大模型

创业公司



王慧文创立光年之外

前商汤员工创立 MiniMax



杨植麟创立循环智能



周伯文创立衍远科技



唐杰创立智谱 AI, 现为 CTO



蓝振忠创立西湖心辰



创业黑马子公司数智云科与达摩院签订战略合作协议, 致力于打造对标 ChatGPT 的中国本土化大模型

实验室



高校



生成式人工智能发展带来的风险

✓ **生成式人工智能确实存在一定风险，如果在扩展规模时没有考虑和缓解这些风险，可能会减缓转型速度。**

首先，模型应该持续进行训练以提高性能，这会引起有关敏感数据、隐私和安全的担忧。接下来，模型输出的结果与训练数据的质量息息相关。因此，任何数据偏差（例如在表示或采样方面）通常会在输出中出现。其他挑战包括确定结果的知识产权所有权、高计算成本和强化学习中昂贵的人工干预成本。**参与生成式人工智能的开发、消费、讨论和监管的所有人都应努力管理以下风险：**

信任侵蚀

恶意行为，如幻觉、深度伪造、网络钓鱼和提示注入，以及不确定因素，如不标注数据来源，都可能暴露攻击面，破坏客户的信任。

安全和风险

公司在使用具有固有“多租户”模型的嵌入式技术和调优技术时，应该保持数据、嵌入和调优的机密性，并跟进快速变化的监管环境。

长期的员工失业风险

当前，使用生成式人工智能的最高投资回报率用例是辅助工作流程和提高生产力；然而，随着模型的不断进步，如果没有适当的技能提升和劳动力规划，可能存在失业风险。

偏见和歧视

如果没有防护措施和持续监控，生成式人工智能很容易模仿偏见并传播歧视行为。

成本

使用生成式人工智能的查询/提示的成本可能高达索引查询的十倍。虽然这些成本很可能会随时间降低，但经济因素应该考虑进内部商业案例和客户定价，以推动采用。

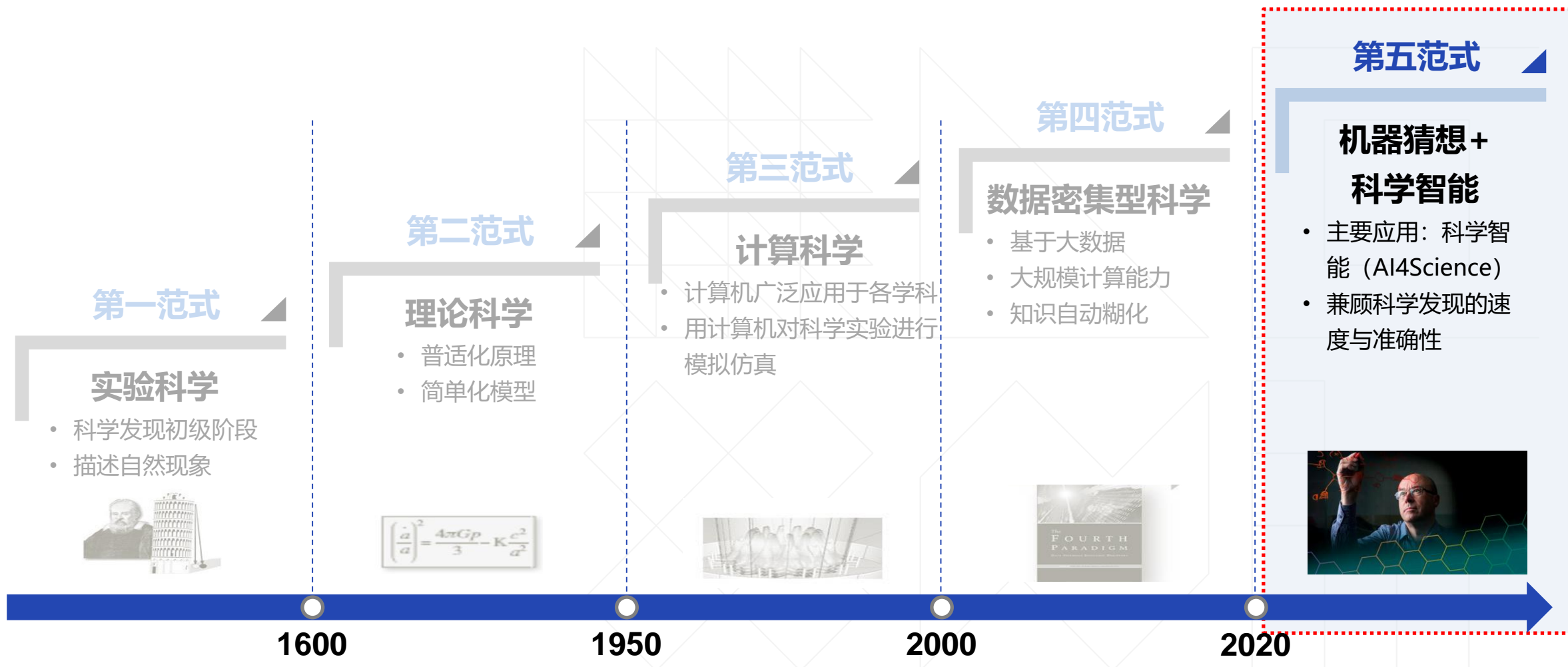
数据隐私和知识产权的保护

模型将在一系列专有的、通常是私有的数据上进行训练，需要符合监管要求，进行节点隔离和源头可追溯性的管理。



第五范式的出现

图灵奖得主吉姆·格雷 (Jim Gray) 2007年提出科学范式



科学智能+机器猜想：第五范式的出现

过去几年时间，我们看到了随着机器学习的发展，深度学习应用于科学基础理论的发现中，并且兼顾了效率与准确性。这种使用机器学习的新方式与第四范式数据建模截然不同，因为用于训练神经网络的数据来自科学基本方程的数值解，而非经验观察。



“AI4Science 是一次深植于微软使命的尝试，这将充分利用我们的人工智能能力来开发新的科学发现工具，从而让我们和科学界的其他同仁能够应对人类面临的最重要的一些挑战。”

——微软执行副总裁兼首席技术官Kevin Scott



一个典型的案例是，剑桥的团队与诺华合作的生成化学（Generative Chemistry）项目，这个项目负责的参与者表示人工智能不仅可以从我们过去的实验中学学习，而且随着实验室中设计和测试的每一次新迭代，机器学习算法可以识别新的模式，并为早期药物发现和开发过程提供指导，希望通过这样的方式，我们可以增强人类科学家的专业知识，从而更快地设计出更好的分子。

我们认为“第五范式”的特点就在于通过巨型复杂系统的“量子对撞”的方式进行“机器猜想”，通过不同的“算法思维”和“应用场景”的对撞，得到不同领域的“专业知识”，将未知的结论推导出来，从而反向推动该领域的发展，得到在经验领域上尚未得到的前瞻性的结果。

The background is a solid blue color with several faint, light-blue geometric patterns. In the top-left, there is a grid of squares with diagonal lines. In the top-right, there is a large right-angled triangle. In the bottom-left, there is a grid of diamonds. In the bottom-center, there is a large, complex polygon with multiple vertices. In the bottom-right, there is a grid of squares of varying sizes.

THANKS!