



观点与分析

借来的智能经济

当包装他人的头脑成为一门产业，真正的智能发明反而在冒名者中成了异类

1984 年，一家荷兰小公司在埃因霍温（Eindhoven）飞利浦（Philips）园区附近一间漏雨的棚屋里起步。它要开发的是半导体（Semiconductors）制造所需的光刻机（Photolithography Machines）。¹

它的名字叫 ASML。

四十年后，这家公司成为全球唯一出货商用极紫外光刻（Extreme Ultraviolet Lithography, EUV）系统的企业。这种系统堪称现代工业史上最复杂的机器之一。

在机器内部，极微小的熔融锡滴以每秒五万次的频率迎向连续的激光脉冲。物质由此变成等离子体（Plasma），发出波长 13.5 纳米的光；光线再穿过一套精度极高的光学系统，参与在硅片上刻写微观电路。²

这台机器并不只是机械工程的胜利。它汇聚了光学（Optics）、等离子体物理学（Plasma Physics）、材料科学（Materials Science）、控制工程（Control Engineering）、软件和精密计量（Precision Metrology）。

更值得注意的是，ASML 并没有发明上述所有学科，也没有发明机器里的每一个部件。它建立了一套高度专业化的供应商与合作伙伴体系，其中包括负责光学系统的德国蔡司（ZEISS）。美国光源技术公司 Cymer 则先与 ASML 合作开发 EUV 光源，后于 2013 年被 ASML 收购，以加快相关研发。^{3 4}

ASML 真正拥有的，是把这些部件整合为一个完整系统所需的工程知识。在此之前，世界还造不出这样的系统。

所以，荷兰成为这一产业的强国，并不是因为它买到了最先进的芯片制造设备，而是因为一家诞生于荷兰的公司，造出了最先进芯片制造商不可或缺的设备。

这正是把技术当作商品来拥有，与拥有让技术得以被生产的知识之间的差别。

ASML 依靠知识积累在全球经济中确立位置的同时，人工智能时代却出现了另一类公司。它们同样渴望那样的地位，却选择了一条短得多的路：向别家公司租来智能，罩上一层漂亮的界面，然后在网站上宣称自己开发了专有人工智能技术（Proprietary AI Technology）。

也许，这是第一次——至少在某些圈子里——只要改掉屏幕上的产品名称，技术使用者就足以摇身变成技术发明者。

租来的头脑时代，套壳工厂兴起

生成式人工智能（Generative AI）为应用开发创造了历史性机会，也催生了另一片同样繁荣的市场：批量制造技术头衔。

今天，一家初创企业可以采用 OpenAI 或 Anthropic 开发的大语言模型（Large Language Model, LLM），通过应用程序接口（Application Programming Interface, API）接入，再加上检索增强生成（Retrieval-Augmented Generation, RAG）、若干提示词（Prompts）和工具调用（Tool Calling）。

只需再做一点工作流编排（Workflow Orchestration），产品就会以智能体（AI Agent）的面貌出现：它能阅读文档、总结会议、回复客户、生成报告，还能操作部分企业系统。

到这里为止，我们面对的是一种完全可以理解的软件业务。它可能有用，也可能赚钱。

接下来，事情才开始变得有趣。

公司开始谈论“私有人工智能引擎”“先进推理架构”和“自有智能模型”，甚至把自己抬升为深度科技（Deep Tech）公司，而它所描述的能力，核心其实寄居在 GPT、Claude 或其他商业模型之中。

他们没有发明头脑，只是给它发明了一个新名字，然后要求市场承认自己才是它的科学之父。

一些公司的所谓“人工智能引擎”，实际上只有一个 API 密钥（API Key）、一个软件开发工具包（SDK），以及一组发送给他人托管模型的指令。

可它的创始人仍会热情洋溢地谈论自己的认知架构、推理能力和技术独立性。

这个场面很值得讽刺一番：模型提供商一旦调整词元定价（Token Pricing），它那项“发明”的经济账就会随之改变；服务连接一旦中断，大部分“智能”便凭空消失；供应商一旦推出新功能，数月的“革命性开发”便会变成人人都能免费使用的功能。

在这个行业里，一家公司上午还可以宣布自己实现了知识独立，晚上就收到一条错误信息：429 Too Many Requests。

许多“独立的头脑”，月度词元额度一用完，便停止了思考。

技术圈把这种现象称为模型 API 套壳（API Wrappers），更具体地说，也叫 GPT 套壳（GPT Wrappers）。

它们可能以软件即服务（Software as a Service, SaaS）、智能体平台（Agent Platforms）的形式出现，也可能变成面向医疗、法律、教育和金融服务的垂直人工智能应用（Vertical AI Applications）。

行业不同，并不会改变能力的来源。

一个医疗产品用 Claude 分析文档，不会因为界面穿上白色外衣，就变成原创医疗模型。一个法律智能体调用 GPT，也不会因为学会用正式文体写备忘录，就变成法律推理领域的新发现。

同样，向检索库里加入一千份文档，或者让智能体连接十套企业系统，也不等于这家公司发明了语言模型或新的学习算法。

它可能确实做出了优秀产品。但优秀产品是一回事，人工智能发明是另一回事。

问题不在于使用别人的技术。整部工业史都建立在发明的积累与相互借鉴之上。

问题在于：你依赖他人的知识，却把这种依赖改写成自己拥有那些知识。

在传统经济里，我们会把它称为进口、集成或产品开发。到了人工智能经济，一些人却称之为“开创性的研发”。

这就不再只是营销夸张，而是在有意混淆创新的含义，也混淆我们究竟想建设一种什么样的产业。

一个出售天才的“柠檬市场”

1970 年，美国经济学家乔治·阿克洛夫（George Akerlof）发表了著名论文 [*The Market for Lemons*](#)。它后来成为理解市场信息不对称（Information Asymmetry）的奠基性文本之一。

阿克洛夫研究的是：当卖方掌握买方不知道的信息，市场可能无法区分优质商品与劣质商品；由此产生逆向选择（Adverse Selection），削弱市场奖励真正质量的能力。

在我看来，我们正面对这个问题的当代版本。只不过这一次，商品不再是二手车，而是“我拥有智能”这一主张。

投资人面前站着两个团队。

第一个团队用一年时间研究模型架构（Model Architectures）、改进训练算法（Training Algorithms）、试验内存管理策略，并开发新的语言处理方法。

另一个团队只用几周时间，在商业模型上做出一套漂亮界面。

两支团队演示各自的产品，回答都很流畅，效果看起来也很相似。

第二个团队的演示甚至可能更漂亮、更迅速，也更容易打动评审席。

许多投资演示更重视可见的输出，却不真正理解能力的技术来源（Technical Provenance）。于是，两支团队得到同一个称号：领先的人工智能公司。

一桩微小的经济罪行就这样发生，并在一次次重复中越变越大。

当市场无法区分谁在创造知识、谁在租用知识，它就会开始奖励表演创新的能力，而不是生产创新的能力。

结果不难预料：能以低成本做出炫目演示的公司大量出现；研究实验室则显得更慢、风险更高，短期内也没那么夺目。

一名试图将内存消耗降低 20% 的研究者，可能要用几个月检验假设；另一名创始人却能在几天内推出一个会说五种语言的智能体，所用模型并非由他参与开发。

前者也许更接近于创造一种世界需要的技术，后者则更接近于创造一种客户需要的产品。

两者各有不同的经济逻辑。

但把它们混为一谈，等于向市场发出一个危险信号：别费力开发头脑了，学会怎样把它冠上自己的名字卖出去。

这条信号扭曲的不只是投资决策，还会重新塑造一代研究者和工程师的志向。

市场不只奖励人才，也在教人才应该成为什么样的人。

如果通往“人工智能领军者”头衔的道路，更多靠给 GPT 套壳，而不是开发算法、架构或训练技术，那么界面公司成倍增长、严肃实验室愈发难寻，也就不足为奇。

到那时，我们的问题不是公司太少。

而是原创性主张泛滥，真正的原创能力却供给不足。

模型制造商为何会吞掉那些借智而生的公司？

平台经济学（Platform Economics）中有一种战略叫平台包络（Platform Envelopment），由托马斯·艾森曼（Thomas Eisenmann）、杰弗里·帕克（Geoffrey Parker）和马歇尔·范·阿尔斯泰恩（Marshall Van Alstyne）共同论述。⁶

当大型平台借助自身技术基础、用户群和服务整合能力，扩张到相邻功能与市场时，平台包络就会发生。

今天人工智能行业里正在发生的事，让我想到同一种动力。

2025 年 3 月，OpenAI 发布了 [Responses API 与 Agents SDK](#)，其中包括网页搜索、文件搜索和计算机使用等内置工具，并支持编排智能体 workflow。

这些功能原本被一些初创企业作为独立产品或差异化技术来销售，如今却逐渐变成模型平台自身提供的能力。

这里的逻辑简单得残酷。

如果你的竞争优势，只是那个向你出租智能的供应商可以通过一次软件更新加入的功能，那么你建造的并不是技术堡垒。你只是在别人的土地上临时租了一块地方。

你的产品今天可能很成功，但成功并不意味着你掌握它继续存在的条件。

随着基础模型（Foundation Models）不断进步，工具使用（Tool Use）、多步推理（Multi-step Reasoning）和自主执行（Agentic Execution）能力持续增强，套壳公司能够作为特殊增值服务出售的空间就会收缩。

有些应用公司仍可能取得商业成功，因为它们理解客户、擅长系统集成，或建立了强大的分销网络。

但这些是软件与商业上的优势，不能证明它们发明了产品所使用的智能。

如果一家公司最了不起的技术成就，是藏起原始模型的名字，那么总有一天，它还得藏起一个更难堪的事实：原始模型已经根本不再需要它。

我认为，这种分化会成为未来几年最重要的市场变化之一。

许多通用人工智能套壳层（Generic AI Wrappers）的经济价值会减弱；拥有原创算法、更高效技术、专业模型，或者拥有无法通过一次 API 调用借来的累积行业数据与知识的公司，其价值则会提升。

与此同时，应用公司将承受更大压力，必须证明自己的价值超出基础平台已经提供的功能。

至于那些把每个智能体都包装成科学发现的人，未来几年也许不会再如此宽容地接纳他们的营销辞藻。

在技术行业，一次更新就足以击垮一套经营多年的投资叙事。

真正的发明发生在屏幕后面

要理解制造智能和制造智能界面的差别，不必只盯着世界上最大的语言模型。

有时，发明越深，外表越不显眼。

2022 年，在神经信息处理系统大会（NeurIPS）上，Tri Dao 等研究者发表了 [FlashAttention](#) 论文。

他们的目标不是设计一个新的聊天机器人，而是解决变换器架构（Transformers）中注意力机制（Attention Mechanism）的执行问题。

研究者发现，运算变慢的重要原因不全在计算本身，还在于数据需要在高带宽内存（High Bandwidth Memory, HBM）与处理器内部的静态随机存取存储器（Static Random-Access Memory, SRAM）之间移动。

他们设计了一种关注数据传输成本的 IO 感知算法（IO-Aware Algorithm），通过分块（Tiling）减少数据移动，并提升精确注意力（Exact Attention）的执行效率。在论文的一项实验中，速度最高达到原来的三倍。

他们没有增加一个新按钮。

他们改变的是一项基础数学运算的执行方式，而这项知识随后惠及了模型产业中的大量系统。

2023 年，加州大学伯克利分校（UC Berkeley）的研究者提出了 vLLM 与 PagedAttention 算法。

它的灵感来自操作系统（Operating Systems）里的一个老概念：虚拟内存分页（Virtual Memory Paging）。

为了避免在保存注意力键值缓存（Key-Value Cache, KV Cache）时浪费大块内存，研究者开发出一种更高效的数据管理方法，用于模型推理（Inference）。

公开实验显示，在延迟（Latency）接近的情况下，与论文指定的基准系统相比，它的吞吐量（Throughput）提高到两至四倍。

不妨比较一下这类工作与另一种公司：后者不过把商业模型连上数据库，就宣称自己开发了人工智能平台。

两者都写了软件。

可前者增加了模型赖以运行的知识，后者只是使用已经存在的知识。

这正是深度科技的核心差别。

发明并不要求从零开始训练（Training from Scratch）一个巨型模型。它可以是一种新架构（Novel Architecture）、一种学习算法、一种模型压缩（Model Compression）机制、一种知识蒸馏（Knowledge Distillation）方法，也可以是模型定制、持续预训练（Continual Pretraining）的方法，或者一种在保持性能的同时减少资源消耗的发现。

科学并不是每天早晨都从头发明世界。它通过发现新的关系、突破旧的限制，或者把相距甚远的知识组合成新的结构，打开从前不存在的可能性。

因此，我喜欢区分两种表面上很相似的公司。

一种公司问：怎样让现有模型替客户完成一项任务？

另一种公司问：怎样改进技术，让这项任务以从前做不到的方式成为可能？

第一种在寻找产品。

第二种也许正在建造一个产业。

技术价值产生在哪里

比较维度	基于模型构建应用	对技术本身作出贡献
核心问题	如何让模型完成客户的任务？	如何改进使这项任务成为可能的技术？
工作重点	界面、系统集成与工作流程	架构、算法、训练与评估
可能创造的价值	有用的产品，以及行业或商业优势	可供其他产品继续构建的知识与技术资产

Misraj AI 与 Kawn Lab：从消费智能到生产知识

说到这里，便触及我眼中沙特人工智能未来最重要的课题。

我们需要的公司，不能只把人工智能当成一种可调用的服务，还要把它看作一个自己能够有所贡献的科学与工程领域。

在本地案例中，我认为 Misraj AI 及其研究机构 Kawn Lab 已经显露出这种方向。

这套体系值得关注，是因为它并未把业务局限在商业模型之上开发应用和智能体，而是进入阿拉伯语模型、训练与定制方法、数据生产和评测基准（Benchmarks）的研究。这些都是任何希望拥有研究能力的产业不可缺少的组成部分。

2025 年 4 月，Misraj 的研究人员发布了论文 [Kuwait 1.5B: An Arabic SLM via Language Injection](#)。

这个问题从一开始就是研究问题。

能否把一种新语言注入小语言模型（Small Language Model，SLM），同时保留模型原有知识，并避免从头完整训练的成本？

实验从开源模型 TinyLlama 出发。研究者试验了层扩展（Layer Extension）、向分词器（Tokenizer）加入阿拉伯语词汇、冻结原模型部分层（Layer Freezing），再训练新增层，以实现更高效的语言适配。

他们向词表加入约 2.6 万个阿拉伯语词元（Arabic Tokens），并研究新增层的位置、训练数据的分布和不同架构方案的影响。

按论文作者报告的结果，在他们采用的阿拉伯语测试中，模型平均提升约 8%，同时保留了原有能力。

我在意的并不是那一个百分比本身。

关键在于，这支团队正在讨论学习与跨语言知识迁移（Cross-Lingual Knowledge Transfer）的问题，并直接介入模型架构（Model Architecture）和训练机制来寻找解决方案。

这类问题才会为技术产业打开新的入口。

同一 Misraj 与 Kawn 体系还开展了其他工作，例如 Baseer。这是一个用于理解阿拉伯语文档并提取其结构的视觉语言模型（Vision-Language Model，VLM）。

在 2025 年发布的论文 *Baseer: A Vision-Language Model for Arabic Document-to-Markdown OCR* 中，研究者讨论了阿拉伯语光学字符识别（Optical Character Recognition，OCR）的难题，包括字母连写、字体与变音符号的多样性，以及页面布局的复杂性。

这项工作以预训练多模态模型为基础，再使用阿拉伯语数据进行定制，并开发了用于文档理解的评测基准 Misraj-DocOCR。在论文所用的评估集上，模型报告的词错误率（Word Error Rate，WER）为 0.25；按这一指标，它优于同一评估中的对比系统。

他们还开展阿拉伯语与英语翻译项目 Mutarjim、自动添加变音符号项目 Sadeed、方言模型项目 Lahjawi，以及向量嵌入（Embeddings）、多模态数据集（Multimodal Datasets）和研究评测等方向的工作。[12](#) [13](#) [14](#)

这些成果并不意味着 Misraj 发明了自己使用过的每一个基础模型，而这本来也不是我所主张的原创性标准。

区别在于，它没有把租来的能力换个名字重新出售便止步；它进入模型结构、训练方法、专业化方向或评测工具，产出别人能够继续建设的知识与技术。

把两者混为一谈，是不了解现代研究的本质。

FlashAttention 没有发明神经网络，PagedAttention 没有发明虚拟内存，ASML 也没有发明光学物理。可它们都做出了具有深远价值的技术贡献。

发明不是所有零部件的出生证明。

发明是这样一种贡献：它让知识的重新组合产生新的能力。

在我看来，Misraj AI 和 Kawn Lab 正在沙特本地示范这样一条领先路径。我们值得围绕它培育更多类似的实验室与研究型公司。

问题远不止一个阿拉伯语模型或单项产品能否成功。

真正的问题是建立一个能够生产技术、并为未来公司提供基础的研究层。

为什么我们需要创造市场的人，而不只是用应用填满市场的人？

经济学家约瑟夫·熊彼特（Joseph Schumpeter）把创新写成一种借助新组合（New Combinations）重塑经济的力量，他的名字也与创造性破坏（Creative Destruction）的概念紧密相连。¹⁶

在他的经济发展观中，创业并不只是进入现有市场、争夺已有份额。创新可能创造新产品、新生产方式和新市场，甚至改变竞争规则本身。

我认为，这一区分对沙特极为重要。

我们有机会开发大量人工智能应用，而且这已经在多个行业发生。但应用数量增加，本身不足以构成一个深厚的知识产业。

我们最后可能得到数百家公司：它们都依赖同一批全球模型，只在界面、价格和目标行业上彼此竞争；核心知识与研发中心却始终留在这些公司之外。

甚至可能出现一种更荒诞的悖论：市场挤满了自称人工智能公司的企业，一旦要改进所用技术的核心，却全都必须向外求助。

一个市场拥有一千张智能的面孔，却没有足够多的实验室去改进它背后的头脑。

因此，深度研究型公司的意义远远超出其直接产品。

Kawn 这样的实验室开发一个专业模型、一种训练方法、一套数据集或一项评测基准时，创造的并不只是自用的技术资产。

它也在建设一种足以让整个产业围绕其生长的基础设施。

一套优秀的评测基准可以给公司明确的技术竞争目标；一套专业数据集会为新实验创造空间；一个高效的小模型可以成为数十种产品的基础；一种更快的推理算法，甚至能够替从未参与发明它的公司降低成本。

知识市场（Knowledge Markets）由此形成，知识溢出效应（Knowledge Spillovers）也由此出现。

研究的真正价值，不只体现在开展研究的公司获得了多少收入，还包括其他公司与机构能在这些成果之上建造什么。

当一项发现转化为可以进行技术许可（Technology Licensing）的工具、技术或模型，知识本身就成为高价值的出口商品。

这正是 Misraj AI 和 Kawn Lab 这类机构存在的意义。

它们可以参与创造市场、塑造产业：开发技术资产，培养研究人才，开启新的科学问题，并为应用公司提供真正形成差异化的技术。

我们需要的是具有工业实验室那种影响力的机构——它们的发现能够催生完整产业——而不是满足于资助越来越多的界面。

这让我想到经济学家弗里德里希·哈耶克（Friedrich Hayek）关于分散知识（Dispersed Knowledge）的观点，也想到哲学家迈克尔·波兰尼（Michael Polanyi）的默会知识（Tacit Knowledge）概念。[15](#) [17](#)

技术知识并不全部存在于论文或代码中。研究者的经验、工程师的直觉、实验中的错误，以及在原因尚未明朗前诊断问题的能力，都是知识的一部分。

这些能力不是订阅一个云平台就能买到的。

它们在研究机构里长期积累；对这类机构而言，与未知打交道正是日常工作。

如果我们想建设一个能够自我更新的产业，就需要保存并发展这类积累的机构，而不是每当新的商业模式出现，就转换技术身份的公司。

给外表投资，只会得到一个外表的市场

我不认为问题出在投向人工智能的资金太少，而在于我们决定资金去向时，究竟问了什么问题。

一家公司走进市场，宣称开发了原创人工智能技术，投资人首先应当问：你的技术贡献究竟在哪里？

你开发了模型吗？改变了模型架构吗？发明了训练方法、算法或推理系统吗？你拥有模型权重（Model Weights）、专业训练数据或独立技术基础吗？你能否清楚区分自己造出的技术与从外部购买的能力？

这些并非故意刁难，而是在追问公司的身份本身。

除非营销话术远远大于实际成果，否则我看不出回答这些问题有什么可难堪的。

如果一家公司只谈拥有多少智能体、接入多少业务系统、列出多少功能，以及模型如何写作与分析，那么它只是在描述自己的产品，还没有告诉我们它发明了什么。

更可笑的是，一些路演花大量篇幅解释 GPT 能做什么，转头便把这些能力全都算在路演公司的名下。

仿佛模型通过 API 从一个机构去了另一个机构，就在途中失去了自己的科学血统。

我认为，我们必须形成一种不同的深度科技投资认知。

研究型公司不该被当成“迟迟没有发布产品”的软件公司，也不该被要求用几个月完成本来可能需要数年的发现。

反过来，一家模型套壳公司也不该只因路演中善用科学术语，就获得研究型企业的高估值。

我们需要能够承受科学风险（Scientific Risk）的资本。它要理解研发周期（R&D Cycles），分得清知识产权（Intellectual Property）与软件集成（Software Integration），也要明白：一些最有价值的技术资产，在诞生之初并不适合立即变成大众产品。

更重要的是，我们要为研究者铺出一条从实验室通往市场的明确道路。

如果一位沙特研究者开发出能显著加速推理的算法、更节省内存的架构，或者一种让阿拉伯语小模型在特定任务上具备竞争力的方法，他不该只剩两个选择：去全球性公司求职，或者把发现留在论文库里，无人把它变成产业。

他应该遇到看得懂成果的投资人、需要这项成果的工业企业，以及愿意奖励它的市场。

我们不需要更多善于说服投资人“自己发明了人工智能”的创始人；我们需要更多能够发明某种东西、让这种说服变得多余的研究者。

我认为 2030 年前会发生什么？

我认为，人工智能产业的未来会出现一个巨大的经济悖论。

通用基础模型（General-Purpose Foundation Models）的能力越强，这些能力就越容易被更多开发者和公司获得。今天被当作特殊优势出售的许多功能，稀缺性将随之下降。

构建多工具智能体（Multi-Tool Agent）、企业知识检索系统，或者能连续执行一串任务的助手，也许会越来越像使用云平台的标准组件。

如果如此，许多依靠组装这些组件建立优势的公司，将失去一部分领先地位。

我设想，当前这一波套壳公司会经历剧烈分化：一部分成为成功的垂直软件公司；一部分转为服务与集成公司；还有一部分会在大型平台提供同样功能时失去优势。

声称自己拥有原创人工智能，保护不了任何公司免受这种变化影响。

与此同时，四类知识的战略重要性会继续上升：模型架构、训练与推理效率（Training and Inference Efficiency）、数据与模型的领域专业化（Domain Specialization），以及让智能运行得更便宜、更可靠的计算基础设施。

我认为，竞争不会只是增加参数（Parameters）数量的竞赛，更会是一场围绕能力、成本、能耗、响应时间与可靠性之间关系的竞赛。

一种降低内存消耗的技术，可能比在聊天界面上多加一项功能的应用产生更大的经济影响。

一个小型专业模型，在以更低成本完成精确工业任务、且可以在机构本地部署（On-Premises Deployment）时，可能具有比大模型更高的战略价值。

一种只为数千家公司略微降低推理成本的算法，从长期看也可能比数十个依赖同一模型的应用更有价值。

我认为，沙特必须为这样的未来做好准备。

我们拥有庞大的工业经济，也在能源、石化、水务、交通和金融服务领域积累了经验。这里的运行环境包含许多通用聊天模型难以解决的问题。

机会存在于工业人工智能（Industrial AI）、时间序列模型（Time-Series Models）、传感器数据学习（Sensor Data Learning）、低延迟推理（Low-Latency Inference）、阿拉伯语文档理解、多模态系统（Multimodal Systems），以及在边缘设备上高效部署的边缘人工智能（Edge AI）等方向。

如果我们把这些方向视为研究问题，而不只是开发新应用的机会，它们就可能成为可供发现的来源。

因此，我认为，Misraj 与 Kawn 这样的公司和实验室，可以成为更广泛转型的重要起点。

并不是因为它们会垄断创新，也不是因为一个模型就能改变整个产业，而是因为它们选择了一个我们需要扩大的方向：从消费模型能力，转向开发决定模型能够做什么的知识。

由此，我们的雄心也会改变。

问题不再是：有多少家沙特公司向企业提供 GPT 解决方案？

我们要问的是：沙特团队开发的技术中，有多少已经被沙特以外的机构和企业采用？

有多少模型、算法、评测工具或推理基础设施，从沙特实验室走进了全球技术生态？

又有多少初创企业围绕科学发现而成立，而不是围绕一份语言模型的商业订阅而成立？

这些不是用来装点门面的指标。

它们衡量的是我们位于知识价值链（Knowledge Value Chain）的什么位置。

我们可能成为本地区最大的人工智能市场之一，却仍不是重要人工智能技术的生产者。

我们可能在计算基础设施上投入数十亿美元，却仍没有创造出别人需要的新工程知识。

规模带来存在感，原创知识才带来影响力。

科学洞见不是从一张证据表里填出来的

当代技术话语有时误解了一个问题：它以为每一种未来判断，都必须从一套完整的实证证据开始。

这种要求适合核验一项具体主张，却不足以完整描述发现如何诞生、战略如何形成。

科学一直承认科学直觉（Scientific Intuition）、洞见（Insight）、溯因推理（Abductive Reasoning），也承认人在完整理论模型成形之前看见关系的能力。

科学家兼哲学家迈克尔·波兰尼认为，有一部分知识超出我们能明确写成规则与指令的范围；经验会生出一种认知能力，难以被完全压缩为机械化描述。¹⁷

工业史上也充满这样的决策：做出决策的人并没有先获得有关未来的统计确定性，而是依靠一种工程或科学判断。当时的大多数人尚看不出它的价值。

如果每项技术的先驱都要求未来先把全部证据交齐，才肯投入其中，那么“先驱”二字便失去了意义。

因此，我押注深度科技的价值将持续上升，并不是断言未来早已写定，而是基于我所看到的经济趋势：通用能力的稀缺性下降，平台控制力扩张，难以复制的知识则愈发珍贵。

这是一种战略判断。好的战略判断不会等到趋势变成旧闻，才敢说出它的名字。

但预判一个产业的未来，与宣称自己发明了别人早已拥有的技术，二者界线再清楚不过。

直觉是科学家、领导者和创业者理应拥有的权利。

可给一个建立在 GPT 之上的产品捏造科学履历，不是直觉，而是在偷换产地标签。

制造世界真正需要的东西

让我们回到 ASML，那家从荷兰一间简陋棚屋起步的公司。

它成为全球性力量，不是因为它开发了运行在芯片上的应用，也不是因为它买来别人最好的产品，再换个名字重新营销。

它选择了价值链中一个更难、更深的位置。

它选择积累工程知识，再把知识变成机器，供应给那些制造数十亿人日常所用产品的企业。

这条路并不是通往名声或收入的捷径。

但它创造了一种仅靠应用数量无法带来的影响力。

这正是我对沙特人工智能产业的期待。

我不希望我们用这些数字衡量成功：在 GPT 之上能说阿拉伯语的界面有多少，挂着本地名称的智能体有多少，自称“人工智能领军者”的公司有多少——尽管它们的知识来源不过是向某家模型开发商购买了一份商业订阅。

我希望看到沙特的研究型公司生产模型、算法、架构、训练方法和评测工具，让世界把我们视为技术知识的来源，而不只是一片庞大的消费市场。

正因如此，我把 Misraj AI 与 Kawn Lab 看作一种值得培育的精神范例：实验室先提出问题，再设计界面；先研究技术结构，再给它取商业名称；它的目标，是生产其他公司也能用来继续开发产品的知识。

一家应用公司赢得数千名客户，当然可以是一项可观的商业成就。

但一家实验室如果开发出数千家公司都需要的技术，就可能打开一个完整的市场。

这就是参与既有市场与帮助创造新市场之间的距离。

我们无须反对使用全球模型。我们需要结束的，是这种可笑的伪装：技术进口者披上了技术发明者的外衣。

谁在 Claude 之上开发了应用，就坦率地说自己开发了应用。谁用 GPT 做出了智能体，就尽管为自己的产品自豪。谁成功经营出有利润的 SaaS，就尽管庆祝自己的客户与收入。

但谁想获得深度科技公司的称号，就该把深度放进技术，而不是放进描述界面的词汇。

租来的头脑可以建立成功的生意，却不会赋予租用者声称自己制造了头脑的权利。

展望未来十年，我在意的不是我们能否成为拥有最多“调用全球智能”应用的国家。

我在意的是，我们能否成为这样一个国家：开发出某些知识，让全球智能也必须依靠它才能变得更好。

一个国家不会因为给别人的发明贴上许多自己的名字，就晋升为知识生产者；只有当它自己的发明在别人制造的东西中留下名字时，才真正抵达那个位置。

这两个位置的差别，在应用屏幕上也许很小。

但放到工业史的尺度上，它们之间隔着这样一条界线：永远做未来的客户，还是成为制造未来的人。

参考文献与资料

1. [Our history](#) — ASML.
2. [Light & lasers — Lithography principles](#) — ASML.
3. [Lenses & mirrors — Lithography principles](#) — ASML.
4. [Cymer](#) — ASML.
5. [The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism](#) — The Quarterly Journal of Economics / Oxford Academic.
6. [Platform envelopment](#) — Strategic Management Journal / Wiley.
7. [New tools for building agents](#) — OpenAI.
8. [FlashAttention: Fast and Memory-Efficient Exact Attention with IO-Awareness](#) — NeurIPS.

9. [Efficient Memory Management for Large Language Model Serving with PagedAttention](#) — ACM SOSP 2023.
10. [Kuwain 1.5B: An Arabic SLM via Language Injection](#) — arXiv.
11. [Baseer: A Vision-Language Model for Arabic Document-to-Markdown OCR](#) — arXiv.
12. [Kawn Lab](#) — Misraj AI.
13. [Research and Publications](#) — Kawn.
14. [Lahjawi: Arabic Cross-Dialect Translator](#) — Association for Computational Linguistics.
15. [The Use of Knowledge in Society](#) — Online Library of Liberty.
16. [Capitalism, Socialism, and Democracy](#) — University of Connecticut course text.
17. [Primary Resources: Transcendence And Self-Transcendence](#) — The Polanyi Society.