2025年红杉 AI 闭门峰会深度解读报告

I. 执行摘要：解码2025年红杉 AI Ascent 峰会

2025年5月2日于旧金山举行的红杉AI Ascent峰会（Sequoia AI Ascent 2025）标志着一个关键的转折点，预示着人工智能（AI）正从一项前景光明的技术演变为一股基础性的经济力量 。此次峰会强调了AI领域前所未有的市场机遇，以及在发展过程中保持"最大速度"（maximum velocity）的紧迫性 。同时，AI智能体（AI agents）作为具备变革能力的"行动引擎"（action engines）的崛起也成为焦点 。本报告旨在综合梳理红杉资本合伙人及全球顶尖AI领袖在峰会上发表的关键论述、展示的技术突破以及阐述的战略要务，并从资深技术专家的视角，对他们共同描绘的AI未来蓝图及其深远影响进行深度解读。

核心主题洞察：

- 指数级市场潜力：AI的市场潜力远超以往的技术浪潮，其增长呈指数级态势 。
- AI智能体与智能体经济的崛起：AI智能体正从概念框架迅速走向实际应用，能够自动化复杂工作流程，预示着"智能体经济"（Agent Economy）时代的到来 。
- 垂直领域AI与应用层的战略聚焦：在应用层，特别是针对特定行业的垂直AI解决方案，将成为创造持久价值的关键 。
- AI基础设施的巨大投入与挑战：AI基础设施建设，尤其是数据中心和能源供应，既带来了巨大机遇，也伴随着严峻挑战 。
- "随机性思维"的必要性：在AI时代，适应和驾驭其固有的概率性特征，需要向"随机性思维"（Stochastic Mindset）转变 。
- AI商业模式的演进：传统商业模式面临变革，基于成果的定价模式（outcomes-based pricing）正逐步取代订阅制 。
- **安全、伦理与治理的集成**：将安全性、伦理考量和治理框架融入AI开发的全生命周期，已成为不容忽视的关键环节 。

专家视角解读：

此次峰会不仅是技术进展的展示，更是一次关于商业运营、创新模式和竞争格局范式转换的号召。红杉资本合伙人Pat Grady所言的"市场中巨大的吸吮声"（tremendous sucking sound），不仅指代了市场对AI的强烈需求，更形象地描绘了AI对顶尖人才、巨额资本和战略焦点的强大引力。

一个值得深思的现象是，峰会对"最大速度"的强调与对治理和伦理的讨论（后者在更广泛的AI讨论中频繁出现，但在已披露的AI Ascent具体内容中，除了安全议题，其深度似乎不及对速度的强调）之间形成了一种张力。这揭示了一个潜在的"创新-治理鸿沟"：行业在全力构建强大引擎的同时，其全面的安全与伦理操作手册仍在起草之中。Pat Grady的"最大速度"和"巨大的吸吮声"突显了快速创新和抢占市场份额的巨大压力。与此同时，外部关于AI伦理、安全和治理的讨论日益增多 。虽然AI Ascent峰会触及了安全问题（如Nikesh Arora的演讲 和Robust Intelligence的介绍），但根据现有信息，峰会内部对伦理框架的深入讨论似乎不如对发展速度的强调那样突出。这种不平衡可能导致"创新-治理差距"，即快速、不受约束的发展可能引发不可预见的负面后果、社会反弹，并最终减缓或误导AI的潜力。峰会对速度的侧重暗示治理可能滞后，这是技术革命中的常见模式，但在AI领域，其风险更高。

此外，“智能体经济”的构想若成为现实，可能演化为一个复杂的生态系统，需要全新的微观经济理论和治理模型。Konstantine Buhler提出的“智能体经济”设想AI智能体能够进行交易并形成经济关系。这意味着一个由自主或半自主实体进行经济决策的网络。传统经济理论基于人类的理性行为，而AI智能体经济可能展现出不同的动态，受其目标函数、学习算法和交互协议的驱动。如果多个AI智能体为相似目标（例如，在特定市场中为其各自所有者实现利润最大化）进行优化，并使用相似的数据和学习技术，它们可能在没有明确协调的情况下独立地趋向于类似串通的行为（例如价格固定），即“算法串通”。这不仅需要AI伦理，还需要一个全新的“算法经济学”领域和专为智能体驱动市场设计的治理框架，以确保公平性、防止市场操纵，并应对新型系统性风险。若缺乏前瞻性的管理，可能催生“算法合谋”或其他未曾预料的、破坏性的市场行为。

II. 红杉愿景：驾驭AI革命浪潮

红杉资本的合伙人们在AI Ascent 2025峰会上，从不同维度阐述了他们对当前AI发展阶段的判断以及未来趋势的展望，共同勾勒出一幅AI驱动的变革图景。

A. Pat Grady：“市场中巨大的吸吮声”——市场规模、发展速度与创业要务

Pat Grady指出，AI代表的利润池至少比云计算等先前技术转型大一个数量级，它同时冲击着软件和服务两大市场。AI产品正从工具（tools）进化为辅助驾驶（copilots），最终将发展为自动驾驶（autopilots），相应的支出也从软件预算转向劳动力预算 。

AI的普及速度前所未有，这得益于几个关键因素：首先是ChatGPT发布后引发的全球性认知普及；其次是业已成熟的传播渠道，如社交媒体平台；最后是无处不在的互联网连接 。

对于初创企业，Grady建议聚焦于特定垂直领域或特定功能的应用，解决复杂问题（可能需要“人在回路”的解决方案），并坚持“从客户出发”（from the customer back）的原则 。他发出的行动号召是：“任何时候都要以最大速度前进” 。Grady的观点将AI定位为一次根本性的经济结构重塑，而非渐进式改进。他对“最大速度”的强调，凸显了在这个快速演变的领域中，竞争的激烈程度以及先发优势的稍纵即逝。

“从客户出发”的方法论在AI时代具有新的内涵。Grady建议初创企业“从客户出发”进行构建 。基础模型正变得日益强大且易于获取，这可能导致纯模型层的商品化。因此，初创企业可持续的差异化优势，不太可能仅仅依赖于拥有一个略胜一筹的通用模型。相反，理解并解决特定垂直领域中具体、复杂的客户问题，使初创企业能够：深度融入客户的工作流程，创造高昂的转换成本；获取并利用非公开的、独特的领域特定数据；通过解决独特的客户问题，构建专有的数据飞轮，即AI解决方案随着更多客户特定数据的积累而改进，从而进一步提升其价值和防御壁垒。这使得“从客户出发”的策略成为在AI优先的世界中构建护城河的关键。

更深一层看，AI同时冲击“软件和服务市场”的论断，预示着两者界限的模糊化。传统意义上的服务，如法律文件审阅、财务分析等，通常涉及大量的人力劳动和专业知识。“自动驾驶”式的AI系统则意味着对目前由人类服务提供者执行的复杂任务进行高度自动化。如果AI能够自主或近乎自主地执行这些服务任务，那么这些服务就能以软件的可扩展性和成本结构来提供。这意味着传统服务行业不仅面临着为其人力员工提供更高效工具的挑战，更面临着来自能够直接提供终端服务的AI系统的颠覆。这可能导致对何为“服务”与何为“软件产品”的重新定义。

B. Sonya Huang：从用户粘性到产品市场契合——应用型AI的崛起

Sonya Huang强调，AI工具的用户参与度发生了巨大转变，例如ChatGPT的日活跃用户与月活跃用户比例已接近Reddit的水平，这标志着强大的用户粘性 。这种用户留存率的提高是衡量AI产品价值的关键指标 。

在突破性应用方面，Huang特别指出编程领域已达到“尖叫级的产品市场契合”（screaming product-market fit）。其他取得显著进展的领域包括广告（生成精准的广告文案）、教育（辅助概念可视化）和医疗保健（如Open Evidence辅助诊断）。她再次肯定了应用层将创造最大价值的观点，并列举了如Harvey、Glean、Sierra、Cursor等一批AI杀手级应用的出现 。

对于初创企业，Huang提醒需关注收入质量（是持久的行为改变还是短暂试用）、实现健康利润率的路径（单位token成本持续下降，基于价值的定价能力提升），以及与业务指标紧密挂钩的数据飞轮的重要性 。Huang的洞察将AI的热潮落脚于具体的用户行为和成功的应用案例。她将编程领域视为一个“转折点市场类别”，这一点意义重大，因为它直接影响未来软件和AI系统的创造方式。

编程领域产品市场契合的快速实现，以Anthropic的Claude 3.5 Sonnet和Cursor等工具为代表，预示着软件开发生命周期本身的加速。AI编程工具 使得软件创建更加便捷、快速且经济高效 。这种开发效率和可及性的提升，意味着更多的创意能够更快地转化为原型并部署。作为许多AI应用交付媒介的软件，其开发周期的缩短将导致AI原生应用的更快普及和多样化。这形成了一个正反馈循环：AI帮助更快地构建AI及AI驱动的软件。

进一步分析，对“与业务指标挂钩的数据飞轮”的强调，意味着未来的竞争壁垒将较少依赖静态数据集，而更多地依赖于通过与客户持续的、能产生价值的互动所驱动的动态自学习系统。许多公司声称拥有数据飞轮，但这些飞轮往往未能直接驱动可衡量的业务改进。在AI时代，数据至关重要，但仅有原始数据是不够的。其价值来源于持续生成、提炼并用于改进AI产品的数据，而这些AI产品反过来又为客户带来更好的业务成果。这创造了一个良性循环：更好的成果带来更多的使用，从而产生更具体、更有价值的数据，这些数据进一步改进AI，从而带来更优异的成果。因此，“飞轮”不仅仅是数据的积累，而是互动、价值交付、数据捕获和模型改进的整个系统。这个深度嵌入客户价值的动态过程，将成为强大的竞争护城河。换言之，与成果相关的、持续的数据采集和优化过程本身，构成了核心的知识产权。

C. Konstantine Buhler：智能体经济的黎明与随机性思维的构建

Konstantine Buhler预见，AI的下一个主要进化阶段是“智能体经济”（Agent Economy）的出现。在这个经济体系中，AI智能体不仅传递信息，还将转移资源、进行交易，并与人类及其他智能体发展出经济关系 。这标志着AI从“答案引擎”向“行动引擎”的转变 。

实现智能体经济面临三大技术挑战：持久身份（智能体需保持一致的个性化特征和记忆）；无缝通信协议（需要类似TCP/IP的智能体间交互标准）；以及强大的安全机制以建立信任 。

与此相伴的是“随机性思维”（Stochastic Mindset）的必要性。这要求从确定性计算转向拥抱概率性结果，意味着在“不确定性显著增加的情况下，撬动远超以往的杠杆”，这将从根本上改变工作方式、企业运营乃至整个经济体的运作模式 。Buhler还提出了“全天候经济”（Always-On Economy）的概念，即AI通过混合式智能体/人类系统，消除时间摩擦，赋能金融、医疗、制造、餐饮、教育、招聘、物流等众多行业实现24/7不间断运营 。Buhler关于智能体经济和随机性思维的构想，描绘了一个AI不仅是工具，更是经济和运营流程中活跃且时而不可预测的参与者的未来。这对系统设计、风险管理和人类适应能力都具有深远影响。

“随机性思维”不仅是工程师面临的技术挑战，更是企业面临的深刻文化和组织挑战。Buhler将“随机性思维”定义为从确定性向概率性结果的转变 。传统的商业实践往往追求可预测性、清晰的因果关系以及不确定性的最小化。而AI系统，尤其是复杂的AI系统，本质上是概率性的，其输出并非总是完全可预测或能以确定性方式解释。因此，深度采用AI的组织必须改变其运营和战略方法。这包括：领导者需要理解并适应基于概率而非确定性的决策；开发新的、能够反映概率性成功率的KPI和绩效指标；实施专为随机系统设计的风险管理框架；以及培育一种接受实验和从“概率性失败”中学习的文化。这与传统的管理范式大相径庭。

更进一步，“智能体经济”可能需要超越传统网络安全范畴的全新“数字信任基础设施”。Buhler的智能体经济涉及智能体转移资源和进行交易 。其关键挑战包括持久身份、无缝通信和安全 。为了使智能体之间，特别是跨不同平台或所有者的智能体之间能够自主进行经济交易，高度的信任是必需的。传统安全侧重于保护系统和数据，而新的需求则关乎信任自主智能体的*行为*和*承诺*。这可能需要：为AI智能体提供可验证凭证（证明其能力、权限和隶属关系）；基于过往表现和可靠性的智能体声誉系统；智能体可以订立并可验证执行的“智能合约”或自动化协议框架；以及去中心化身份解决方案，以防止智能体的欺骗或失实陈述。这种“数字信任基础设施”将是功能性智能体经济的基础层，其范围超出了当前的网络安全范式，或许可以借鉴去中心化身份和智能合约等概念。

III. AI巨擘论道：行业领袖的前沿洞察

AI Ascent 2025汇聚了AI领域的顶尖思想家和实践者。他们的演讲不仅揭示了各自企业在AI浪潮中的战略布局，也为整个行业的发展方向提供了极具价值的参考。

A. Sam Altman (OpenAI)：个性化AI的未来与“核心AI订阅”

Sam Altman回顾了OpenAI从一个研究实验室演变为AI领域主导平台的历程，其中ChatGPT的诞生源于对用户在Playground中与GPT-3互动行为的观察 。他展望ChatGPT将转变为一个深度个性化的AI服务，能够记忆用户的全部生活背景信息（如对话、邮件），并在所有服务中无缝运行 。

OpenAI的目标是打造“核心AI订阅服务”，提供强大的基础模型和核心服务，同时通过平台和SDK赋能其他开发者在其基础上创造价值 。Altman观察到不同代际用户与AI的互动模式存在差异：大学生将AI视为操作系统，二三十岁的专业人士将其用作生活顾问，而年长用户则主要用于搜索和基础任务——这些模式预示了AI集成的演进方向和潜在的产品机会 。

他认为从AI助手到智能体再到完整应用的演进将是连续的，并预想未来会出现一个新的互联网协议层，用于连接联邦化的组件和智能体工具，内置身份验证、支付和数据传输功能 。为保持产品迭代速度，OpenAI推崇小型化、高授权的团队 。关于未来1-3年的价值创造，Altman预测：2025年，AI智能体将执行实际工作（尤其在编程领域）；2026年，AI将辅助重大科学发现；到2027年，机器人将成为重要的经济贡献者 。Altman的愿景指向一个AI与个体生活深度融合的未来，以及OpenAI以平台为中心的战略，在掌控核心智能的同时培育生态系统。他对AI影响力的时间规划颇为积极且多层面。

“核心AI订阅”模式若能成功，可能将OpenAI（或类似的基础模型提供商）定位为一种新型的智能“公用事业”或“操作系统”。Altman设想的“核心AI订阅”是深度个性化且跨所有服务无缝工作的 。这意味着用户将依赖一个中央AI来处理广泛的任务和信息。如果这个AI成为许多数字（甚至物理）互动的主要界面或引擎，它就如同一个操作系统（正如他指出Z世代已如此使用 ）或基础公用事业（如电力或互联网接入）。“核心AI订阅”的提供者将对信息流、用户体验以及第三方开发者在其平台上构建应用的能力拥有巨大影响力。这种“智能交付”的集中化可能导致显著的平台锁定效应，并引发关于竞争、数据治理以及平台所有者与其用户/开发者之间权力动态的疑问。

值得注意的是，Altman所设想的用于智能体交互的“互联网新协议层” ，直接回应了Buhler为智能体经济提出的关键挑战之一（即无缝通信协议 ）。Buhler将“无缝通信协议：智能体交互的TCP/IP等价物”确定为智能体经济的关键技术挑战之一 。Altman的愿景直接应对了这一挑战。这样一个协议的开发和采用，将是实现真正可互操作和可扩展的智能体经济的基础性步骤。如果OpenAI或类似实体成功建立此协议，它可能成为事实上的标准，塑造未来智能体驱动互联网的整体架构，并赋予其创建者巨大的影响力。这也与关于智能体通信协议（如ACP 和Gibber link ）的研究相呼应，表明业界已广泛认识到这一需求。问题在于，它将是一个开放标准还是专有标准。

B. Jensen Huang (Nvidia) 与 Jim Fan (Nvidia)：物理AI时代、机器人技术与“物理图灵测试”

Nvidia的创始人兼CEO黄仁勋（Jensen Huang）认为，物理AI是继生成式AI之后的下一个前沿领域，AI系统将能够在物理世界中感知、推理、规划和行动 。Nvidia将此视为一个价值50万亿美元的机遇（AI 50榜单中提及的Figure AI、Skild AI即为例证 ）。为此，Nvidia推出了Cosmos，一个面向机器人和工业AI的世界级基础模型平台，并已在GitHub上开源 。

Nvidia AI总监Jim Fan进一步提出了“物理图灵测试”的概念，即当机器人完成一项真实世界任务（如打扫房间、准备晚餐）的效果与人类无法区分时，才算通过测试 。他坦言，目前的机器人技术距此目标尚远。机器人技术面临的关键数据挑战在于，机器人所需的连续关节控制信号无法从互联网获取，必须通过人类演示 painstakingly 收集（Fan称之为“燃烧人类燃料”）。

为解决这一瓶颈，Fan倡导将模拟（simulation）作为机器人训练的“核能”。Nvidia的方案包括：数字孪生模拟（比实时快10000倍，结合领域随机化）、零样本迁移（训练于模拟环境，直接部署于现实世界，无需微调），以及令人惊讶的低参数需求（人形机器人控制的“潜意识处理”仅需150万参数）。模拟技术本身也在进化，已发展到Simulation 2.0阶段，即完全由生成式视频扩散模型构建，可用于模拟反事实场景。Fan还展望了“物理API”的未来，机器人将能像软件操控比特一样操控原子，从而催生一个物理技能的应用商店 。硬件方面，GeForce RTX 50系列GPU（如RTX 5090，拥有920亿晶体管，算力达3352 TOPS）为游戏、自动驾驶、机器人和智能体AI提供强大动力 。Nvidia的愿景将AI的触角延伸至物理领域，并将模拟视为克服数据瓶颈的关键。 “物理图灵测试”为真正的具身智能提供了一个引人注目但仍遥远的目标。

对物理AI训练大规模依赖模拟（正如Fan所提议 ）引入了一个新的抽象层次，并可能导致与真实世界物理规律的偏离。Fan主张将模拟作为机器人技术的“核能”，以克服数据限制 。这涉及在模拟环境中广泛训练AI模型，然后将其部署到真实世界的机器人上（零样本或最少微调）。然而，模拟本质上是对现实的近似。尽管领域随机化等技术旨在弥合模拟与现实之间的差距，但差异总是存在的。对于控制物理机器人的AI系统，尤其是在不可预测的人类环境或安全关键应用中，即使是微小的未建模物理现象或传感器噪声特性，也可能导致意外甚至危险的行为。这意味着此类AI系统的验证和确认（V&V）过程必须异常严格，超越当前的软件V&V标准，包括在多样化和对抗性条件下进行广泛的物理测试。这种V&V的成本和复杂性可能成为一个重要瓶颈。

此外，一个用于机器人的“物理API” 可能会普及物理任务自动化，但也引发了关于在共享物理空间中的“运营权”、自主行为的责任归属，以及滥用于制造自主物理威胁的潜在风险等深层问题。Fan设想的“物理API”将允许广泛编程和部署机器人技能 。这意味着未来许多个人和组织可以在家庭、公共场所和工业环境中部署执行各种任务的自主机器人。与在数字领域运行的软件API不同，物理API直接影响物理世界。这就引出了关键问题：如果自主部署的机器人技能造成损害或伤害，责任谁负？API提供商、技能开发者、机器人所有者，还是机器人本身（如果它具有法律人格）？自主机器人在与人类和其他机器人共享空间时的“交战规则”是什么？冲突如何解决？我们如何防止滥用物理API制造自主武器或物理伤害工具？通过物理API开发或部署技能需要哪些认证或许可？这将需要目前尚不存在的社会和法律框架。

C. Jeff Dean (Google)：迈向高级AI工程与类脑系统

Google首席科学家Jeff Dean预测，AI的技能水平可能在2026年（即从2025年5月算起的“一年左右”）达到初级程序员的水平 。这与AI在编程领域展现出的“尖叫级产品市场契合度”相呼应 。

然而，Dean指出，一个“虚拟工程师”需要的不仅仅是编写代码，还包括运行测试、调试、理解文档以及向经验丰富的同事学习等能力 。AI可以通过虚拟环境和学习文档来掌握这些技能。在更长远的展望中，Dean认为AI技术将从他曾引领研发的TPU等专用硬件，向更具生物特征、受大脑启发的有机系统演进 。这与神经拟态计算和生物启发架构的研究方向一致 。Dean对AI在软件工程领域发展的预测具有短期且高影响力的特点。他对类脑AI架构的长期看法，则暗示了一条通往更通用、适应性更强的智能之路，有望超越当前依赖暴力扩展的模式。

如果AI能在2026年达到初级工程师的能力 ， 这将深刻重塑软件开发团队的结构、对入门级人类工程师的需求以及软件生产的经济学。Dean预测AI将在2026年达到初级程序员的技能水平 。初级工程师通常负责复杂度较低的编码任务、错误修复、测试和文档编写。如果AI能够自动化这些任务的很大一部分，那么对当前角色的人类初级工程师的需求可能会减少。软件开发团队可能会演变为由更少、更资深的工程师指导和管理AI编码助手/智能体。这可能会提高整体生产力，但也可能在入门级机会减少的情况下，对培养下一代高级工程师构成挑战，从而产生技能缺口。软件开发的经济学可能会发生变化，常规编码的成本可能降低，但对复杂问题解决、系统设计和AI管理技能的价值会更高。这可能将高级工程师解放出来，从事更具创造性的架构设计和创新工作，但也可能在缺乏主动管理的情况下造成技能断层。

对“受大脑启发的有机系统” 的追求，尤其来自Google AI这样的主要实验室，表明未来AI发展可能出现与纯粹基于Transformer的LLM规模化路径的分野。Dean预见到向“受大脑启发的有机系统”的转变 。当前主流模型（如Transformer）非常有效，但也非常耗费数据和计算资源。受大脑启发的架构（例如神经拟态计算、脉冲神经网络、借鉴神经科学原理的系统 ）通常有望实现更高的能源效率、更自然的学习机制，并可能更好地处理时间序列或

模糊信息。Sakana AI的CTM 就是此类自然启发方法的一个例子。如果这一方向的研究取得突破，可能会催生新一代AI系统，它们不仅是当前LLM的放大版，而且拥有根本不同的能力和效率。这可能解决当前方法的一些可持续性问题（如能源消耗 ）和扩展限制（如数据可用性）。这类系统的成功可能带来更节能、推理更精妙、持续学习能力更强的AI，从而克服当前架构的一些局限性。这与Sakana AI的CTM 和神经拟态硬件 等概念相呼应。

D. Bret Taylor (Sierra)：以成果为导向的AI重塑软件商业模式

Sierra联合创始人Bret Taylor认为，AI正推动软件行业从传统的订阅制向基于成果的定价模式（outcomes-based pricing）发生根本性转变。例如，Sierra的定价模式是当其AI智能体自主解决问题时向客户收费，若需人工介入则免费 。

这种模式要求供应商对其软件交付的实际结果承担更大责任。Taylor强调，驱动可衡量的顶线业务成果通常比单纯节省成本更有价值 。他坚信，最大的机遇在于构建针对特定行业的AI智能体（如电信、银行、保险、医疗保健），并对在企业市场缺乏强大垂直领域专注度的通用型AI平台持怀疑态度 。对客户需求的深刻理解和同理心，而非仅仅展示技术能力，是企业AI成功的关键 。Taylor的洞察揭示了AI市场的成熟趋势，即价值越来越取决于可量化的实际成果，而非仅仅技术本身。这对AI公司的市场定位和产品结构具有重要影响。

向基于成果的定价模式 的转变，将促使AI供应商与其客户之间建立更深层次的合作伙伴关系。Taylor主张基于成果的定价，即根据AI交付的特定结果付费 。为了准确衡量成果，AI供应商需要访问相关的客户数据，并就成功的成果标准达成明确、一致的定义。这需要超越传统软件供应商-客户关系的数据共享和协作水平。AI供应商需要深入了解客户的业务，才能将其AI解决方案与有价值的成果联系起来。这将培育一种更具共生性的关系：供应商的成功与客户通过AI获得的成功直接相关。这可能导致更长期、更具战略性的合作伙伴关系，而非交易性销售。因为供应商需要获取运营数据并清晰定义成功指标，这可能导致更紧密集成和相互依赖的合作关系。

从更宏观的视角看，AI领域向基于成果定价的广泛转变，可能会从根本上改变风险投资和整体投资界对AI初创企业的评估逻辑。Taylor的成果导向定价模型 若成为AI软件的主流变现方式，那么收入将直接取决于已证实的价值交付。风险资本家在投资AI初创企业时，不仅需要评估技术本身，还需要评估初创企业以下能力：清晰定义和量化其AI所能带来的成果；开发可靠的方法来衡量这些成果；构建足够可靠的AI系统以持续实现这些成果；以及基于这些成果谈判和管理合同。这将投资焦点从纯粹的技术指标（例如模型在基准测试中的准确率）或用户获取数量，转向为客户带来的可证明的投资回报率。拥有深厚垂直领域专业知识和强大客户验证机制，能够清晰展示成果的初创企业，在此类环境中将更具吸引力。投资者可能会更关注企业定义、衡量和交付可量化成果的能力，而不仅仅是用户增长或技术新颖性。这可能更有利于拥有深厚领域专业知识和强大客户验证机制的初创企业。

E. Mike Krieger (Anthropic)：构建机构级AI并驾驭合作伙伴-产品生态系统

Anthropic的Mike Krieger探讨了在构建如Claude Code 这样的产品与维护合作伙伴及客户关系之间的微妙平衡 。Anthropic的Claude 3.5 Sonnet已在编程领域引发快速转变 ，其重点在于能够产生实际商业成果的企业工作流和工具 。

Krieger（或在Krieger出席的Human[X]峰会上的一位发言者）曾指出机器人领域缺乏开放性是一个隐忧，并期望出现更多开源组件 。这与AI Ascent峰会上强调开源重要性的总体基调 形成对比。作为主要LLM开发商Anthropic的代表，Krieger的观点揭示了在AI生态系统中竞争与合作所涉及的战略决策，尤其是在基础模型自身日益成为强大应用平台的背景下。

在开发专有的、高价值AI产品（如Claude Code ）与通过合作伙伴关系和API赋能更广泛生态系统 之间存在的张力，是基础模型提供商面临的核心战略挑战。Anthropic与OpenAI一样，开发强大的基础模型。它们也基于这些模型发布特定的产品/功能（例如Claude Code ） 。这可能会产生“渠道冲突”或与那些也希望通过API使用Anthropic基础模型构建类似应用的合作伙伴或客户形成竞争。Krieger在发布Claude Code前致电主要编程客户 的举动，正突显了这种敏感性。战略挑战在于界定哪些能力保持核心/专有，哪些通过API开放以促进生态系统发展。过多的专有开发可能会抑制生态系统的积极性；而过少则可能无法从独特创新中获取价值。它们的成功可能取决于能否找到合适的平衡点，既避免疏远开发者，又能获取重要价值。

对机器人领域加强“开放性”的呼吁，可能与一些先进AI模型的封闭特性形成对比，这暗示着未来硬件（机器人）可能变得更加标准化和开放，而核心智能（控制它们的AI模型）则依然是关键的差异化因素和专有价值的来源。在Krieger出席的Human[X]峰会上，一位发言者指出许多机器人是闭源的，并希望有更多的开放性 。这与驱动这些机器人的最先进AI模型通常具有专有性形成对比。如果机器人硬件平台变得更加开放或标准化（可能通过开源设计或通用接口），物理AI领域的创新可能会加速。然而，为这些机器人提供“大脑”的AI模型可能仍然高度差异化且专有，成为价值捕获的主要场所。这可能导致一个生态系统：多个硬件供应商生产与少数占主导地位的AI“大脑”或“机器人操作系统”兼容的机器人，类似于PC行业拥有标准化的硬件架构和少数占主导地位的操作系统。这种模式可能类似于PC产业的演进。

F. Harrison Chase (LangChain)：编排智能体生态系统——认知架构与用户体验

LangChain的创始人Harrison Chase认为，智能体正从辅助驾驶（copilots）向更自主的系统转变。然而，当前高度通用的自主智能体（如AutoGPT）已让位于更实用、受约束且针对特定领域的“认知架构”（cognitive architectures）。

他将定制化认知架构定义为LLM应用的系统架构（包括数据流、LLM调用流程）。通过特定检查步骤指导智能体的定制化图谱（bespoke graphs），对于实现真实世界的商业价值、效率和可靠性至关重要 。LangChain通过提供“编排层”（如用于构建可定制循环流程智能体的LangGraph），旨在简化这种受控、灵活智能体的构建过程，并解决状态管理和持久化等挑战 。

Chase强调用户体验（UX）的极端重要性。精心设计的UX可以弥补LLM目前的不完美。聊天界面适用于初始交互，但自主智能体需要更复杂的UX模式，如透明的行动日志、回溯/编辑智能体决策的能力、协作界面以及从反馈中学习的机制 。Chase还讨论了在后台持续运行的“环境智能体”（ambient agents）。此外，可观测性（通过LangSmith等工具实现）对于调试多步骤智能体工作流中非确定性的LLM行为至关重要 。Chase对智能体发展持务实看法，强调当前企业应用需要的是结构化、可控的自主性，而非不受约束的通用智能。LangChain的工具在实现这一目标中发挥着关键作用。

对“定制化认知架构”的关注表明，AI智能体开发的价值将显著体现在这些定制化工作流和推理路径的设计与工程上，而不仅仅是底层的LLM。Chase认为通用自主智能体不如拥有定制化认知架构的智能体实用 。这些架构定义了智能体如何处理信息、做出决策以及与工具/LLM交互以完成特定任务。设计和实现这些架构需要对领域、任务、LLM的能力/局限性以及AI工程最佳实践有深刻的理解。这并非易事，远不止于简单地提示LLM。因此，一个围绕为各种业务流程设计、构建和优化这些认知架构的新专业领域将会出现。这为专业化工具（如LangChain）和专家服务创造了市场。

“环境智能体”在后台持续运行的概念，结合对强大编排和可观测性的需求，指向了“智能流程自动化2.0”的未来。Chase介绍了在后台持续运行、响应事件的“环境智能体” 。他还强调了复杂智能体工作流需要编排（LangGraph）和可观测性（LangSmith）。当前的机器人流程自动化（RPA）主要自动化重复性的、基于规则的任务。而环境AI智能体凭借其处理更复杂推理和适应变化条件的能力（在其认知架构内），代表了对传统RPA的重大升级。如果这些智能体能够持续监控系统和事件，并且能够被可靠地编排和观察，它们就能够以更高程度的自主性主动管理和优化复杂的业务流程。人类的参与将从执行任务转向设计智能体架构、监督其性能、处理智能体无法管理的异常情况，以及基于智能体衍生的洞察做出战略决策。这是一种更高级的智能自动化形式，AI智能体将深度嵌入企业系统，主动管理和优化工作流程，而人类的监督主要集中在异常处理和战略层面。

G. Chase Lochmiller (Crusoe)：数据中心即新AI工厂——基础设施与能源要务

Crusoe的CEO Chase Lochmiller强调“数据中心是新的计算单元”这一概念。现代AI工厂实际上是数据中心规模的计算机，其功率密度极高（例如，Nvidia最新技术使机架功率达到600千瓦，远超传统的2-4千瓦），因此需要全新的冷却方案（如大型水管系统）。麦肯锡预测，到2030年，AI数据中心的资本支出将达到5.2万亿美元，传统IT数据中心则为1.5万亿美元 。

AI对能源的需求带来了“阶跃式的变化”，改变了这个一度“沉睡”的公用事业行业。数据中心的电力需求预计到2026年将比2022年翻一番以上 。这给公用事业公司和净零排放目标带来了挑战 。Crusoe正在进行垂直整合，甚至自建发电厂并探索小型模块化反应堆（SMR）为数据中心供电 。

为应对供应链瓶颈和公用事业限制，Crusoe采取了加速基础设施建设和垂直整合的策略。通过自建制造设施，Crusoe将关键电气组件的交付时间从100周缩短至22周，并能在200-300天内完成大型数据中心项目，而传统上这需要3-5年 。Crusoe公司本身也从最初利用油田的火炬气进行计算，发展到建设大规模AI数据中心（目前在建约2吉瓦，储备项目约20吉瓦）。作为对比，传统的算力中心弗吉尼亚北部，经过数十年发展，总容量也仅约4.5吉瓦 。Lochmiller的演讲揭示了AI革命的巨大规模和物理世界的制约因素。AI的未来与能源生产和基础设施部署速度的突破密不可分。

“数据中心是新的计算单元”这一论断，标志着整个数据中心，凭借其计算、网络、冷却和供电等复杂组件的协同运作，被设计和优化为一个单一的、集成的AI工作负载系统，而非离散服务器的集合。Lochmiller指出“数据中心是新的计算单元” 。传统数据中心通常在通用服务器上承载多样化的工作负载。而AI工作负载，特别是大型模型训练和推理，需要高度专业化、密集封装且功耗巨大的硬件（如GPU、TPU），以及高带宽互连。这些AI工作负载的性能不仅取决于单个服务器，还取决于数千个加速器的协同操作、海量数据吞吐量以及整个设施复杂的冷却和供电系统。因此，数据中心本身必须被构建为一个内聚的“AI工厂”或“数据中心规模的计算机”，其中所有组件都为AI进行了优化。这涉及到在史无前例的规模上对计算、网络、电力和冷却进行协同设计。这需要一门全新的“AI数据中心架构”学科。

AI数据中心巨大的能源需求以及对SMR等专用能源解决方案的探索 ，可能导致AI公司成为能源领域的重要参与者。AI数据中心对能源的需求呈“阶跃式变化” 。公用事业公司难以满足这种需求，促使像Crusoe这样的AI基础设施建设者考虑自建发电厂或投资SMR 。如果AI公司开始直接投资或开发专用发电设施（例如SMR、大规模可再生能源项目），以确保其数据中心获得稳定、充足且可能具有成本效益的电力供应，它们实际上就进入了能源市场。这可能涉及与能源公司成立合资企业、收购能源资产或成为独立的电力生产商。这种对能源的垂直整合将使它们能够更好地控制AI的关键投入，但也会使其面临能源行业相关的新的监管、运营和市场风险。这将是这些科技公司一次重大的战略多元化。它们不仅仅是能源消费者，还可能成为能源生产者或新型能源基础设施的直接投资者，从而模糊科技与能源行业的界限。

H. Carl Eschenbach (Workday)：企业AI、领域特定智能体与“智能体记录系统”

Workday的CEO Carl Eschenbach认为，AI正推动一场转变，从人类与技术协同工作，到技术为人类服务，同时强调保持企业中的人类能动性和连接的重要性 。他指出，在企业应用中，拥有精心策划数据的领域特定智能体比通用模型更有价值，因为它们理解特定的业务流程和工作流 。

Workday对AI的商业化采取三管齐下的策略：基于坐席的定价（对全体员工有价值的智能体或AI技术带来的提升）；基于角色的智能体（例如薪资智能体）；以及基于消耗的API访问（AI智能体访问Workday数据的频率）。

为应对“智能体蔓延”（agent sprawl）并确保AI智能体以安全、负责、合乎道德和法规的方式通过“AI网关”引入企业，Workday正在构建“智能体记录系统”（Agent System of Record）。该系统旨在统一管理企业内AI智能体和人类员工，并将其置于一个可信平台上 。Eschenbach还强调，成功销售和实施企业级AI解决方案需要高层管理人员的参与和传统的企业销售方法，而不仅仅是产品驱动的增长 。Eschenbach从企业视角清晰地阐述了AI的采纳、管理和商业化路径，其中“智能体记录系统”是应对新兴治理需求的创新概念。

“智能体记录系统”的概念意味着AI智能体将在组织内获得类似于“数字员工”的地位。Eschenbach描述的“智能体记录系统”用于在统一平台上管理AI智能体和人类员工 。该系统将处理智能体的入职、身份、访问控制等事务 。这些都是传统上与管理人类员工相关的职能。因此，AI智能体正被概念化为企业内一种新型的员工或实体。这将需要人力资源部门、IT部门以及法律合规团队为这些“数字员工”制定新的政策和程序，涵盖从创建/采购到退役的整个生命周期，包括正式的入职流程、角色定义、访问控制、绩效监控乃至“离职”管理——这为人力资源和IT治理开辟了一个全新领域。

Workday的三管齐下AI商业化策略 （基于坐席、基于角色、基于消耗）表明，企业AI的价值将通过一种混合方式来获取。Workday计划通过基于坐席的定价、基于角色的智能体和基于消耗的API访问来实现AI的商业化 。基于坐席的定价反映了可供许多用户使用的通用生产力提升。基于角色的智能体定价则反映了自动化特定高价值工作职能或任务的价值。基于消耗的API访问反映了通过AI访问和利用Workday精心策划的数据所产生的价值。这种多方面的方法承认AI在企业内以不同方式交付价值：作为通用工具、作为专业工作者以及作为智能数据处理器。其他企业AI供应商可能会采用类似的混合定价模型来捕获这些不同的价值流，这使得AI采购和成本管理更加复杂，但也更符合多样化的用例，既体现了广泛的效用提升，也体现了专门的、由成果驱动的自动化。这种复杂的定价模式可能成为企业AI解决方案的标准。

I. Nikesh Arora (Palo Alto Networks)：捍卫AI前沿——平衡创新与控制

Palo Alto Networks的CEO Nikesh Arora强调，随着AI获得“手臂和腿”（即对系统的实际控制权），安全问题变得至关重要。真正的挑战在于跟上AI驱动攻击的加速步伐，这要求从以预防为主转向实时检测和响应系统 。

Palo Alto Networks推出的“AI防火墙”解决方案，通过检查进出AI模型的数据，防止劫持、未经授权的访问和数据泄露 。这一概念与Robust Intelligence的AI防火墙理念一致 。Arora主张在积极进行AI实验的同时，必须实施强有力的护栏、监控和人工监督，尤其是在AI直接控制关键系统之前 。

对于负责任的AI部署，Arora建议重点关注高风险环境，这些环境需要高精度的模型、广泛的测试、持续的监控和故障切换机制。企业数据基础设施的就绪状态是关键 。Arora从关键的网络安全视角出发，强调AI的强大能力必须辅以同样复杂的安全措施。“AI防火墙”是确保企业安全采用AI的一项关键创新。

Arora所倡导的对AI输入输出进行“不懈检查” 的需求，表明AI安全将成为一个持续的、动态的过程，类似于DevOps（可称为DevSecOps），而非一次性设置。Arora强调需要一个“AI防火墙”并对流入和流出AI模型的数据进行“不懈检查” 。AI模型，特别是那些持续学习或与动态数据交互的模型，可能存在不断演变的漏洞或产生意外输出。因此，保护AI安全并非一劳永逸的静态任务，它需要持续的监控、测试和安全控制的调整。这类似于DevSecOps将安全性贯穿于整个软件开发生命周期。对于AI系统，可能需要一种类似的方法，或许可以称之为“AISecOps”。这将催生对具备AI模型行为、数据完整性、AI对抗性攻击以及AI系统实时异常检测等专业知识的安全专业人才的新需求和技能。

如果AI本身被用于攻击和防御（AI驱动的攻击 vs. AI防火墙 ），我们可能会进入网络安全的“算法战”时代。Arora指出，不良行为者正在利用AI加速攻击 。Palo Alto Networks和Robust Intelligence正在开发AI驱动的防御系统（AI防火墙）。这造成了一场AI被攻击者和防御者同时使用的军备竞赛。AI驱动攻击的速度可能变得过快，以至于人类防御者在没有自身AI工具的情况下无法有效应对。同样，AI防御系统也需要迅速适应新型AI生成的攻击向量。这种动态可能导致网络空间中高度自动化且迅速升级的“算法战”，其主要参与者是相互对抗的AI系统。这样一个环境的稳定性和可控性成为关键问题。在这个时代，攻击和防御的速度与复杂性迅速升级，可能超出人类在没有AI辅助的情况下进行管理的能力，若不加以审慎治理，可能导致高度不稳定的网络环境。

表1：AI Ascent 2025峰会主要演讲嘉宾核心观点摘要

演讲嘉宾	所属机构	核心信息/主要预测 (AI Ascent 2025)	主要支撑论点/案例/战略启示 (引文ID)
Pat Grady	红杉资本	AI市场机遇远超以往，初创企业需“全速前进”，聚焦垂直应用与客户价值。	AI利润池比云大一个数量级，同时冲击软件和服务市场；AI产品从工具演进至“自动驾驶”模式，预算从软件转向人力 ； AI普及速度空前，得益于现有认知、渠道和连接 ； 初创企业应专注垂直/功能应用，解决复杂问题，从客户出发 。
			ChatGPT用户活跃度接近Reddit ； 编程是首个

Sonya Huang	红杉资本	AI应用用户粘性显著增强，编程等领域已实现“尖叫级产品市场契合”，应用层价值巨大。	转折点市场类别，广告、教育、医疗等领域涌现突破性应用（如Open Evidence）；Harvey, Glean, Sierra, Cursor等杀手级应用出现；关注收入质量、利润路径和有效数据飞轮 。
Konstantine Buhler	红杉资本	“智能体经济”是AI下一重要进化，需“随机性思维”应对，催生“全天候经济”。	AI智能体将转移资源、进行交易，形成经济关系，从“答案引擎”变“行动引擎”；技术挑战：持久身份、通信协议、安全；随机性思维：拥抱概率结果，以更少确定性撬动更大杠杆；AI赋能24/7运营 。
Sam Altman	OpenAI	ChatGPT将成深度个性化“核心AI订阅服务”，OpenAI构建平台，AI智能体将逐步实现工作自动化、科学发现和机器人经济贡献。	ChatGPT记忆用户全部生活背景，跨服务无缝工作；提供基础模型和核心服务，赋能第三方创新；预测2025年AI智能体执行编程等工作，2026年助推科学发现，2027年机器人贡献经济；提出新的智能体互联网协议层 。
Jensen Huang & Jim Fan	Nvidia	物理AI是下一前沿，机器人技术潜力巨大（50万亿美元市场），通过模拟技术（Nvidia Cosmos）和“物理图灵测试”推动发展。	物理AI能感知、推理、规划、行动；Cosmos为机器人和工业AI基础模型平台；物理图灵测试：机器人任务效果与人无异；模拟克服机器人数据收集瓶颈，实现零样本迁移；RTX 50系列GPU为核心硬件 。
Jeff Dean	Google	AI有望在2026年达到初级程序员水平，未来AI系统将向更“有机”、类脑方向发展。	AI不仅能写代码，还能学习测试、调试等工程技能；技术从专用硬件（如TPU）向类脑系统演进 。
Bret Taylor	Sierra	AI正推动软件商业模式从订阅制转向基于成果的定价，垂直专业化是AI智能体的最大机遇。	Sierra按AI自主解决问题收费；供应商需为实际结果负责，驱动顶线成果比降本更重要；企业AI成功需深耕行业，理解客户痛点 。
Harrison Chase	LangChain	AI智能体正从辅助工具向更自主系统演进，实用化需定制化“认知架构”和强大编排、可观测性及优秀用户体验。	通用自主智能体让位给针对特定领域、有约束的认知架构；LangChain提供编排层（如LangGraph）简化智能体构建；用户体验和可观测性（LangSmith）对智能体采纳和调试至关重要；提出“环境智能体”概念 。
Chase Lochmiller	Crusoe	数据中心是新的计算单元和“AI工厂”，其建设面临巨大的能源需求和基础设施挑战，需加速和垂直整合。	AI工厂即数据中心规模计算机，功率密度极高（600kW/机架）；AI能源需求呈阶跃式增长，Crusoe自建电厂、探索SMR；通过垂直整合加速基础设施建设（如组件制造）。
		企业AI应关注领域特定智能体，Workday推行	领域特定智能体结合精心策划数据比通用模型

Carl Eschenbach	Workday	“智能体记录系统”统一管理AI与人类员工，并采用混合AI商业化模式。	更有价值； “智能体记录系统”应对“智能体蔓延”，确保合规引入； 商业模式：基于坐席、角色和消耗的混合定价 。
Nikesh Arora	Palo Alto Networks	AI安全至关重要，需从预防转向实时检测响应，通过“AI防火墙”等技术平衡创新与控制，负责任地部署AI。	AI获得系统控制权后安全是首要问题，需应对AI驱动的攻击加速； “AI防火墙”检查模型输入输出； 在受控环境中积极实验，但关键系统上线前需强力护栏和人工监督 。

IV. AI Ascent 2025 的主导议题与技术前沿

A. 智能体的飞跃：从AI助手到自主行动引擎

人工智能正经历从信息检索和辅助工具到能够独立完成复杂工作流程并解决实际问题的“行动引擎”的深刻转变 。这一转变的核心驱动力是AI智能体（Agentic AI）的成熟。它们不再仅仅是响应指令的聊天机器人，而是能够主动执行任务、做出决策，并在特定领域展现出超越传统自动化工具的能力 。红杉资本与福布斯联合发布的2025年度AI 50榜单也印证了这一趋势，上榜的初创企业大多致力于将智能体应用于真实的企业工作流中，例如法律AI领域的Harvey、客户服务领域的Sierra以及编程领域的Cursor等 。

Konstantine Buhler提出的“智能体经济”构想进一步描绘了这一蓝图：AI智能体不仅能够相互沟通，还能转移资源、进行交易，并与人类及其他智能体形成复杂的经济关系网络 。为实现这一愿景，必须克服若干关键技术障碍，包括为智能体建立持久的数字身份（使其拥有连贯的“个性”和记忆）、开发类似TCP/IP的智能体间无缝通信协议（ACP 、Gibber link 等研究以及Sam Altman的设想 均指向此方向），以及构建强大的安全与信任机制 。Buhler同时提出的“全天候经济”概念，则指出AI智能体与人类协同的混合系统将消除传统的时间限制，赋能各行各业实现24/7不间断运营 。

然而，智能体经济的实现并非坦途。当前智能体在处理冗长、非结构化任务时仍面临可扩展性挑战，且容易出现故障 。Nvidia估计下一代AI智能体可能需要百倍于当前的计算资源，这构成了显著的算力鸿沟 。治理与监管同样面临难题，智能体的自主性和进化行为使其往往游离于现有法规的灰色地带，欧盟AI法案等现有框架可能需要针对智能体AI进行调整（如风险分类、人类监督、文档要求等）。市场集中度也是一个潜在风险，传统的网络效应可能导致智能体市场出现“赢家通吃”的局面，即少数主导性智能体凭借网络规模而非技术优势形成垄断 。此外，智能体AI带来的劳动力替代 、潜在的安全风险（如被恶意用于网络攻击或散布虚假信息），以及其“黑箱”特性引发的信任与可解释性危机 ，都是亟待解决的挑战。

为AI智能体建立“持久身份” 不仅是一项技术挑战，更是一个哲学命题。Buhler将“持久身份”视为智能体经济的关键，意味着智能体拥有连贯的个性和记忆。如果一个拥有持久身份的智能体能够自主行动、进行交易并签订协议，它就开始类似于一个法律或经济行为者。这就引出了诸多问题：如果一个拥有持久身份的智能体造成损害或违约，责任谁负？开发者、所有者，还是智能体本身？拥有持久身份的智能体能否拥有资产或知识产权？拥有持久身份的智能体应享有哪些权利或受到哪些保护？这些问题超越了技术实现的范畴，深入到关于社会中自主AI实体地位的法律和哲学领域。目前的法律框架很大程度上尚未为这类实体做好准备。它触及了数字人格、责任归属以及智能本质等深层问题，需要全新的法律和伦理框架来界定。

Buhler提出的“全天候经济”概念，描述了AI驱动的各行业持续运营。这有望提高效率和资产利用率。然而，获得这些全天候AI驱动服务（例如医疗、教育、金融）的机会可能并非均等。那些拥有更好条件（数字素养、财务资源、地理位置）的人可能会不成比例地受益。对劳动者而言，全天候经济可能意味着新型轮班工作、对人工监督岗位持续连接的要求，或传统工作时间的模糊化，从而影响工作与生活的平衡。通常基于人类工作模式和时区的劳动法，可能需要重大修订，以适应跨全球边界持续运作的混合人机劳动力。如果对这些持续服务的获取以及参与这一经济体系的益处分配不均，可能会加剧社会不平等。同时，这也将重新定义全球范围内的劳工法律和工作与生活的平衡。

B. AI的垂直化深耕：领域特定解决方案的崛起

红杉资本的合伙人们在峰会上一致强调，AI创造巨大价值的主战场在于应用层，特别是针对特定行业或特定功能的垂直解决方案 。Pat Grady建议初创企业聚焦此类应用，并坚持“从客户出发”的原则 ，Bret Taylor也力挺垂直领域的AI智能体 。

在垂直AI领域，深厚的领域专业知识以及专有的、经过精心整理的领域特定数据集，是构筑竞争壁垒的关键 。目前，这类高质量数据集的短缺已成为行业发展的一大制约因素 。

尽管如此，多个垂直领域已涌现出成功的AI应用案例：

- 编程：被认为是“尖叫级的产品市场契合” ，是AI应用的关键垂直领域。
- 医疗与生命科学：因其数据丰富性、监管紧迫性和高价值决策需求，被视为AI市场增长最快的领域之一 。例如，Open Evidence用于辅助诊断 。
- 法律：Harvey AI等工具正自动化法律工作流程 。
- 金融与物流：AI正在重新定义这些行业的运营效率和战略规划 。
- 企业AI：Workday等公司证明，拥有高质量数据的领域特定智能体更具价值 。
- 新兴创业：Apriora（招聘）、Sweetspot（政府采购合同）、ShowAndTell（牙科患者护理）、Avitor.ai（私人飞机预订）等初创公司凭借垂直AI智能体迅速崛起 。

市场研究机构MarketsandMarkets的报告预测，2025年全球AI市场规模将达到3717.1亿美元，其增长的一个关键驱动力正是企业对垂直场景化AI解决方案的青睐，市场正向领域特定模型转变 。峰会强烈传递出一个信号：尽管基础模型提供了强大的引擎，但AI领域真正的经济牵引力和可防御的商业模式，将源于那些凭借深厚领域知识解决特定行业痛点的定制化方案。

垂直AI的兴起意味着AI应用市场可能出现细分化趋势。峰会强调垂直特定AI的价值 。不同行业（医疗、法律、金融、制造）拥有独特的工作流程、数据类型和监管要求。为这些垂直领域开发有效的AI需要深厚的领域专业知识和量身定制的解决方案。这表明单一的通用AI应用平台不太可能有效满足所有垂直领域的多样化需求。相反，我们可能会看到大量专注于特定细分市场的专业AI公司涌现。这将催生一个充满活力和多样性的初创企业生态系统。然而，这也意味着企业可能需要集成来自多个AI供应商的解决方案，从而在互操作性、数据共享和整体系统管理方面面临挑战。众多专业参与者能够在各自的细分市场蓬勃发展，而非少数几家通用平台巨头垄断一切。这将创造一个丰富的初创企业生态系统，但也给企业带来了集成和互操作性的挑战。

垂直AI对高质量、领域特定数据的严重依赖 ，以及这类数据目前存在的短缺问题 ，可能成为行业发展的主要瓶颈。这种情况可能催生对“数据创建”或“数据丰富化”服务的巨大需求，并可能围绕特定行业数据的合乎道德且合规的采集与标注，形成新的商业模式。这种短缺会阻碍垂直AI解决方案的开发和性能。这将为那些能够做到以下几点的公司创造商机：开发针对特定垂直领域的高效准确数据标注工具和平台；提供合乎道德地获取、整理和丰富领域特定数据集的服务；创建针对特定行业需求的合成数据生成技术，同时确保真实性和实用性；在行业内以合规方式促成数据共享协议。获取或创建这些有价值数据集的能力可能成为重要的竞争优势，并且可能围绕特定垂直领域的

“AI数据就绪性”出现新的服务类别。

C. AI基础设施的淘金热：为智能革命提供动力

Crusoe公司CEO Chase Lochmiller提出的“数据中心即新的计算单元”深刻揭示了AI基础设施的变革。AI工厂本质上是数据中心规模的巨型计算机。麦肯锡预测，到2030年，AI数据中心的资本支出将高达5.2万亿美元，另有1.5万亿美元用于传统IT设施 。

这些AI数据中心具有极高的功率密度，例如Nvidia的最新技术将单个机架功耗推至600千瓦（传统仅为2-4千瓦），因此需要先进的液冷解决方案（被形容为“水利工程”）。AI引发了能源需求的“阶跃式变化”，预计到2026年数据中心的电力需求将比2022年翻一番以上 ，这对传统电力行业和全球净零排放目标构成了严峻挑战 。Crusoe等公司甚至开始自建发电厂并探索小型模块化反应堆（SMR）作为解决方案 。

为应对紧迫需求和供应链瓶颈，基础设施建设正在加速并出现垂直整合趋势。Crusoe通过自建制造能力，将关键电气组件的交付周期从100周缩短至22周，并能在200-300天内建成大型数据中心，远快于传统的3-5年 。Crusoe自身也从最初利用油田废弃的天然气进行计算，发展到建设2GW在建、20GW储备的庞大AI数据中心集群 。相比之下，传统的算力枢纽如北弗吉尼亚州，经过数十年发展，总容量也仅约4.5GW 。五大超大规模云服务商预计在2024至2027年间的资本支出总额将超过1万亿美元 。

硬件创新是这一切的基础：

- Nvidia：发布了GeForce RTX 50系列GPU（如RTX 5090，集成920亿晶体管，算力达3352 TOPS）、Blackwell架构、Cosmos世界基础模型以及用于自动驾驶的DRIVE AGX平台 。黄仁勋称“物理AI赋能工业和机器人技术是一个价值50万亿美元的机遇” 。
- Google：Jeff Dean引领的TPU项目持续创新 ，Gemini 2.5 Pro在编程和多模态能力上表现卓越 。
- 其他AI加速器：Cerebras等公司也在积极发展 。

然而，AI基础设施的建设也面临劳动力短缺、电网接入、芯片供应、关税壁垒和出口管制等供应链及地缘政治风险 。台湾和韩国等新兴市场在全球AI硬件供应链中扮演着至关重要的角色 。AI基础设施的建设是一项堪比历史上工业革命的宏伟工程，其特点是巨额资本投入、极端工程挑战以及对能源和全球供应链的严重依赖。

“数据中心即新的计算单元”的论断，意味着整个数据中心，包括其计算、网络、冷却和供电等复杂子系统，被视为一个为AI工作负载而设计和优化的单一集成系统，而非仅仅是离散服务器的集合。Lochmiller指出“数据中心是新的计算单元” 。传统数据中心通常在通用服务器上承载多样化的工作负载。而AI工作负载，特别是大型模型训练和推理，需要高度专业化、密集封装且功耗巨大的硬件（如GPU、TPU），以及高带宽互连。这些AI工作负载的性能不仅取决于单个服务器，还取决于数千个加速器的协同操作、海量数据吞吐量以及整个设施复杂的冷却和供电系统。因此，数据中心本身必须被构建为一个内聚的“AI工厂”或“数据中心规模的计算机”，其中所有组件都为AI进行了优化。这涉及到在史无前例的规模上对计算、网络、电力和冷却进行协同设计。这催生了一门名为“AI数据中心架构”的新兴学科。

AI数据中心巨大的能源需求以及对SMR等专用电力解决方案的探索 ，可能促使AI公司成为能源领域的重要参与者。AI数据中心对能源的需求呈“阶跃式变化” 。公用事业公司难以满足这种需求，促使像Crusoe这样的AI基础设施建设者考虑自建发电厂或投资SMR 。如果AI公司开始直接投资或开发专用发电设施（例如SMR、大规模可再生能源项目），以确保其数据中心获得稳定、充足且可能具有成本效益的电力供应，它们实际上就进入了能源市场。这可能涉及与能源公司成立合资企业、收购能源资产或成为独立的电力生产商。这种对能源的垂直整合将使它们能够更好地控制AI的关键投入，但也会使其面临能源行业相关的新的监管、运营和市场风险。这将是这些科技公司一次重大的战略多元化。它们不仅是能源的消费者，还可能成为能源的生产者或新型能源基础设施的直接投资者，从而模糊科技与能源行业的界限。

D. AI商业与思维的新范式

除了技术和基础设施的革新，AI Ascent 2025也揭示了商业模式和思维方式的深刻转变。

首先是**向基于成果的定价模式的转变**。Sierra的Bret Taylor是这一模式的积极倡导者，他认为AI正推动行业从传统的订阅制转向为AI成功完成工作而付费的模式 。这种模式将供应商的成功与客户获得的价值直接挂钩，要求供应商承担更大的责任 。尽管基于价值、基于绩效、基于使用量的定价模型也在AI服务中被探讨 ，但成果导向因其直接与商业影响关联而备受关注。然而，其实施也面临挑战，如成果的清晰归因和企业对可变账单的接受度等 。

其次是**拥抱"随机性思维"**。Konstantine Buhler强调，从确定性计算转向接受和管理概率性结果至关重要 。这意味着要学会在"不确定性显著增加的情况下，撬动远超以往的杠杆" 。这种思维方式对于基于AI进行决策至关重要 。AI的思考方式本身就涉及概率推理、模式识别、客观性和持续学习 ，它要求理解结果的分布，而非假定单一的确定性答案。然而，随机性计算在实践中也面临挑战，如其精度依赖于比特流的长度和质量，随机数生成成本高昂，且通常假设比特流独立（相关性可能导致失败），某些函数的随机表达困难，以及精度与内存之间的权衡等 。如果随机性导致AI行为与用户价值观不一致，还可能破坏信任 。

这些范式的转变与技术进步同等重要。AI的随机性要求我们以新的方式思考可靠性和控制问题，而基于成果的定价则重新定义了软件的价值主张。

AI成果导向定价模式 的采纳，将推动复杂且双方互信的"AI审计"和"价值归因"框架的发展。基于成果的定价要求在预定义结果实现后付款 。许多业务成果受多种因素影响，并非仅由单个AI系统决定。因此，为了公平透明地实施基于成果的定价，需要一种可靠的方法来：分离并衡量AI系统对成果的具体贡献；审计AI达成成果的性能和决策过程；以及订立明确无误的合同，定义成果和衡量标准。这很可能催生新的"AI价值归因"工具和方法论，以及第三方AI审计服务，以确保供应商和客户都能信任这一过程。准确衡量一个AI系统对业务成果的直接影响可能非常复杂，需要新的工具和方法论。

在组织内部成功灌输"随机性思维" 可能成为一项重要的竞争优势。"随机性思维"涉及在与AI合作时接受不确定性和概率性结果 。许多AI应用，特别是涉及复杂推理或生成性任务的应用，不会是100%可预测或无错误的。那些严格要求AI系统达到确定性完美的组织可能会行动迟缓、过度谨慎，或在初步受挫后迅速放弃有前景的AI项目。相反，培养了随机性思维的组织更有可能：设计具有适当人工监督和后备机制的AI系统；实施迭代开发周期，从成功和失败中学习；鼓励对AI能力的实验和探索，即使某些实验不能立即产生完美结果；以及更深入地理解如何有效地管理和利用概率系统。这种在不确定环境中学习和适应的能力，可以带来更快的创新、更好的AI集成，并最终形成竞争优势。它使企业能够通过拥抱实验、从概率性失败中学习，从而更快、更有效地利用AI进行创新，而那些固守确定性思维的竞争对手则难以适应。

E. 前沿AI模型与算法探讨

尽管AI Ascent峰会为闭门会议，完整的技术论文通常不会在峰会期间公开发布，但讨论往往围绕着影响该领域的前沿研究和模型展开。现有信息反映了可能在峰会上讨论的更广泛的AI进展。

- **Sakana AI的持续思考机器（Continuous Thought Machines, CTMs）**：这是一种受自然启发的AI模型，利用神经元活动的同步性和时间动态进行逐步"思考"，使推理过程更具可解释性和类人性。其核心在于一个"内部思考维度"（internal ticks），允许模型进行迭代式信息处理。CTMs在迷宫求解和图像识别等任务中展示了其能力，并呈现出类似人类的涌现行为 。相关论文信息提及 ，arXiv编号2505.05522曾被提及 ，但一度无法访问 。
- **清华大学/北京智源人工智能研究院的"绝对零度推理器"（Absolute Zero Reasoner, AZR）**：这是一种采用可验证奖励的强化学习（RLVR）范式，其中单个模型通过自我提问和解决这些问题来学习和改进推理能力，整个过程*无需任何外部人工标注数据*。它使用代码执行器进行任务验证和奖励生成。AZR在编程和数学推理任务上取得了当前最佳（SOTA）性能，超越了依赖人工标注样本训练的模型。该系统利用演绎、溯因和归纳三种推理模式来自我生成不断加深的挑战 。相关论文信息及arXiv编号2505.03335可见于 。

- **Google的Gemini 2.5 Pro**：作为业界领先的大语言模型（LLM），Gemini 2.5 Pro具备“思考”能力（在回应前进行问题分解和规划），支持多模态输入（文本、图像、音频、视频），拥有百万到两百万级别的token上下文窗口。在数学、科学和编程（在SWE-Bench Verified上得分63.8%）等领域表现出色 。其更新版本在WebDevArena排行榜上领先，并在视频理解方面达到SOTA水平（VideoMME得分84.8%）。此外，Google还推出了隐式缓存（implicit caching）技术以降低API使用成本 。
- **Anthropic的Claude系列模型（如Claude 3.5 Sonnet, Claude Code）**：这些模型对AI编程领域产生了重大影响 。同时，Anthropic的研究也揭示了思维链（Chain-of-Thought, CoT）忠实性问题：模型在推理时往往不会完全表述其使用的“提示”或中间步骤，CoT揭示的已用提示不足20%，这引发了对AI安全性和可信度的担忧 。
- **DeepSeek系列模型（v3, v3 0324, r1）**：这是一系列开源模型，具有开放权重推理的特性，允许开发者访问和调整模型参数。其中，r1模型在推理、网页搜索和上下文感知方面表现优异，超越了OpenAI的o1和Meta的Llama 3.3等模型。拥有6850亿参数的v3 0324模型则在非推理模型中处于领先地位 。
- **其他值得关注的技术提及**：
    - xAI的Grok 3：有望具备高级推理能力 。
    - Nvidia的“物理图灵测试”与机器人模拟技术 。
    - EAR（无量化视觉自回归）：无需量化的连续视觉内容生成方法 。
    - 阿里巴巴的ZeroSearch：AI无需真实搜索引擎即可进行信息搜索，大幅降低训练成本 。
    - Meta的Locate 3D：用于3D环境中物体定位的模型 。

峰会上讨论的这些模型和算法方法，清晰地指向了当前AI研究的几个核心方向：提升推理能力、实现更高程度的自主性（尤其是自学习能力）、增强多模态处理、提高效率（降低成本和计算需求），以及发展更接近生物大脑或受自然启发的计算架构。其中，AZR和CTM等概念尤为新颖，代表了对AI学习和“思考”方式的根本性探索。

V. 投资视野：红杉资本的AI创业指南针

红杉资本在AI Ascent 2025峰会上不仅展现了对AI技术前沿的深刻洞察，也清晰地传递了其在AI领域的投资哲学和关注方向。

红杉的投资哲学与策略：

红杉资本合伙人强调，AI领域是一个“拼命快跑的行业”（run like heck business），需要保持“最大速度” 。其投资焦点明确指向应用层，特别是针对特定垂直行业或特定功能的应用，并坚持“从客户需求出发”（customer-back）解决复杂问题 。对于初创企业而言，收入质量、健康的利润模型以及与业务指标紧密相连的有效数据飞轮，是衡量其价值的关键 。

重点关注的AI子领域：

- AI智能体与智能体经济：这无疑是红杉高度关注的核心领域，预示着AI从辅助工具向自主行动者的转变 。
- 垂直AI应用：横跨编程、医疗、法律、金融等多个行业，解决具体场景的痛点问题 。
- AI基础设施：虽然红杉的投资传统上更侧重于应用和软件层面，但数据中心、能源供应和专用硬件等基础设施的极端重要性得到了充分肯定 。
- 机器人技术（物理AI）：AI 50榜单中的Figure AI和Skild AI即为例证，黄仁勋也表达了对这一领域50万亿美元市场潜力的乐观看法 。
- **AI驱动的软件开发工具**：编程领域被视为已达到“尖叫级产品市场契合” 。

更广泛的AI领域风险投资趋势（2025年第一季度）：

根据相关行业报告，2025年第一季度，一笔高达400亿美元的AI交易使该季度风险投资总额达到801亿美元，成为自2022年第一季度以来表现最强劲的季度 。若剔除此项交易，则投资额环比下降36%。信息技术（IT）领域持续主导VC投资，占总投资额的74%，AI在其中影响显著 。随着基础设施建设的推进，投资者开始更多地关注应用层作为投资机会 。尽管巨额融资案例（不含上述400亿美元交易）相较2024年第四季度有所减少，但市场仍有大量“干火药”（dry powder，指已募集但尚未投资的资金）。然而，新基金的募集规模降至2018年第三季度以来的最低点，且在缺乏明确退出路径的情况下，投资者对后续轮融资持谨慎态度 。

摩根大通将AI公司分为五类：AI硬件（如Nvidia、ASML、台积电）、AI超大规模计算平台（如AWS、Google Cloud）、AI开发者（如Adobe、Microsoft及小型应用开发者）、AI集成商（大型企业自建方案及IT服务商）和AI要素（能源、原材料、数据等）。目前投资者热情主要集中在硬件和超大规模计算平台。但长期来看，“AI要素”和“AI开发者”领域存在大量机会，后者可能诞生最大的赢家 。

综合来看，红杉资本正积极推动在AI能够快速交付实际价值的领域进行大胆而专注的投资，尤其是在应用层和垂直市场。更广泛的风险投资环境也显示出对AI的巨额投入，但同时也伴随着谨慎情绪，并将重心逐步转向应用层面。红杉对应用层和垂直AI初创企业的重视 ，与更广泛的风险投资市场向应用层转移的趋势 以及领域特定AI的增长 高度吻合。这表明投资逻辑日趋成熟：基础模型的开发日益集中，而将这些模型应用于具体商业问题的广阔领域，为风险投资和价值创造提供了更广阔的舞台。基础设施的“镐和铲子”游戏规模巨大但资本密集 ； 而真正的“黄金”在于构建于其上的应用。

VI. 规划负责任的未来：AI时代的伦理、安全与治理

尽管AI Ascent 2025峰会的公开信息中未详细阐述全面的AI伦理框架，但鉴于Sam Altman、Mike Krieger等行业领袖的出席，以及对医疗、金融等高风险应用的讨论，伦理、安全与治理的考量无疑是题中应有之义。红杉资本自身的内容生态也多有触及，例如Nikesh Arora关于AI安全护栏的论述 ，对Robust Intelligence AI防火墙的介绍 ，以及Konstantine Buhler对智能体经济中安全需求的强调 。

关键关切领域与焦点：

- AI安全：随着AI自主性的增强，安全问题变得至关重要。这包括部署AI防火墙（如Palo Alto Networks、Robust Intelligence的产品）、实现实时威胁检测与响应、严格审查AI系统的输入输出数据、防止未经授权的访问以及确保有效的人工监督 。同时，也需要应对由AI驱动的新型攻击手段 。
- 偏见与公平性：确保AI系统公平对待所有用户，避免歧视。这需要使用多样化的训练数据，进行常规审计，并采用算法层面的公平性技术 。
- 透明度与可解释性（XAI）：努力使AI决策过程不再是“黑箱”。这对于建立信任至关重要，尤其是在医疗、金融、法律等高风险AI应用中 。Anthropic关于思维链（CoT）局限性的研究也凸显了此方面的挑战 。
- 问责制：人类必须对AI的决策最终负责。需要建立清晰的问责框架，特别是针对智能体AI可能产生的错误或损害 。
- 数据治理与隐私保护：在AI处理日益增多的敏感数据时，保护数据、防止滥用是核心要求 。
- 人工监督：在高风险AI应用中，人工监督是强制性的。对于自主性日益增强的智能体系统，需要重新定义监督的内涵与方式 。
- 监管环境：针对AI的专门法规（如欧盟AI法案）和全球标准化努力正在兴起 。然而，监管如何跟上技术创新的步伐仍是一个持续的挑战。
- 劳动力市场冲击与再培训：AI自动化对劳动力市场的影响深远，需要相应的再培训和适应策略 。
- **随机性与信任**：AI系统的随机性行为如果管理不当，可能破坏其与用户价值观的一致性，进而损害信任 。

从红杉资本合伙人的发言和投资组合中可以推断，其立场强调将安全内置于AI系统（Arora ），投资于AI安全技术（Robust Intelligence ），并认识到智能体经济中信任的必要性（Buhler ）。“全天候经济”的构想也强调了AI系统与人工监督的审慎结合 。

负责任的AI正从原则探讨走向实践落地。AI Ascent峰会聚焦于强大的“行动引擎”和自主智能体，这无疑提升了对稳健治理、安全保障和伦理考量的要求。行业对“最大速度”的追求，必须与这些关键的保障措施审慎平衡。AI能力的快速发展，特别是在智能体系统和如AZR这样的自学习模型领域 ，其速度正在超越全面治理、伦理和安全框架的建设速度。这种“步调问题”——即技术发展快于社会理解和管理其影响的能力——在AI领域尤为突出，因为AI具有潜在的涌现行为和广泛影响。尽管AI Ascent峰会强调速度 ，但更广泛的AI社区（包括像Anthropic这样的一些参与者 ）正在努力解决如何确保这种速度不会导致不负责任的部署或不可预见的负面社会后果。这需要主动的、适应性的治理机制，而非被动的监管。

VII. 战略综合与专家展望：未来之路

关键主题的融合： AI Ascent 2025描绘了一幅由日益自主的智能体驱动的AI革命图景。这些智能体主要通过垂直应用进行部署，构建于全新规模的基础设施之上，并要求新的经济模型和人类思维方式的根本转变。

速度与真实性的两难： 在竞争激烈的环境中，对AI发展“最大速度” 的呼吁可以理解，但这必须与对真实性、可靠性和安全性的需求相平衡，尤其是当AI系统成为具有现实世界影响力的“行动引擎”时 。“随机性思维” 承认了固有的不确定性，但这更需要强大的验证、安全（如Arora所强调 ）和伦理监督框架，而非盲目冒进。

“应用层”作为战场与价值核心： 尽管基础模型是关键的赋能者，但峰会再次强调，可持续的差异化和价值捕获将主要发生在应用层 。在这里，深厚的领域专业知识、独特的数据飞轮以及针对特定客户问题的解决方案（垂直AI ）将构筑持久的护城河。“智能体记录系统” 正是解决治理需求的应用层创新的一个例子。

基础设施：关键赋能者与潜在瓶颈： “数据中心即新计算机” 的论断凸显了支撑AI的物理现实。能源可获得性、专用硬件的供应链弹性以及所需的巨额资本投资都是严峻的挑战 。硬件（Nvidia 、Google TPU ）和基础设施部署（Crusoe ）的创新至关重要，但短期到中期内，需求规模可能超过供应能力。

演进中的人机共生： 从AI媲美初级程序员 到混合人机团队的“全天候经济” ，工作的本质将发生深刻变革。这需要主动的劳动力适应、再培训策略以及对组织结构的重新思考。“物理图灵测试” 代表了这种共生关系在物理世界中的终极（尽管遥远）愿景。

展望——驾驭“随机指数”时代：

AI的发展轨迹不仅在能力上呈指数级增长，其发展路径和社会影响也日益呈现随机性。“随机性思维” 因此不仅是一种操作策略，更是一种战略要务。

- **展望一：“认知架构师”的崛起**：随着AI智能体的普及，一个新的关键角色——“认知架构师”——将会出现。这些专业人士将负责设计、实施和优化那些控制智能体行为、决策逻辑以及与企业系统和数据集成的定制化认知架构（如Harrison Chase所描述 ）。这超越了简单的提示工程，涉及到完整的智能体系统设计，需要融合AI/ML专业知识、领域知识和系统思维。
- **展望二：智能体经济的“信任堆栈”**：智能体经济 的实现将需要一个多层次的“信任堆栈”（Trust Stack）。这将包括：
    - 技术层：安全的通信协议（Altman的设想 、ACP ）、强大的智能体身份验证机制以及AI防火墙（Arora 、Robust Intelligence）。
    - 运营层：用于成果导向定价的AI审计框架 、持续监控和可观测性平台（LangSmith ）以及“智能体记录系统”（Eschenbach ）。

    - **伦理与治理层**：针对特定行业的AI智能体伦理规范、适应性监管框架以及算法问责机制。
- **展望三：“垂直AI鸿沟”**：尽管垂直AI 前景广阔，但许多行业将面临一个“垂直AI鸿沟”——即通用AI模型的潜力与在特定领域实际部署合规、可靠的解决方案之间的巨大差距。跨越这一鸿沟需要大量投入来整理领域特定数据 、开发专业化的智能体架构、应对行业特有的法规，并与最终用户建立信任。那些成功在高价值细分市场弥合这一鸿沟的初创企业，将成为主要的收购目标或市场领导者。

结语： 市场中那“巨大的吸吮声”正以惊人的速度将未来拉至眼前。AI Ascent 2025清晰地表明，驾驭这个由AI驱动的未来，关键不仅在于构建更快、更智能的AI，更在于以智慧、安全的方式构建，并深刻理解其变革力量与内在的不确定性。最终的胜利者将是那些既能掌握创新速度，又能保证应用真实性的企业与个人。