

# Jason Wei在各个场合的观点

📅 2025年10月19日 ⌚ 1 分钟阅读

#AI

#Jason wei

Jason wei在各个场合的观点

## Jason Wei在斯坦福 AI 俱乐部发表的演讲 (2025-10)

**会议:** Jason Wei在斯坦福 AI 俱乐部发表的演讲 **日期:** 2025年10月  
**主题:** 3 Key Ideas in AI in 2025

介绍了三个核心且基本 (fundamental) 的理念, 旨在帮助理解和预测 2025 年及未来人工智能的发展和影响。

这三个关键理念总结如下:

### 1. 智能即商品化 (Intelligence as a Commodity)

这一理念的核心观点是, **获取知识或进行推理的成本和可及性将趋近于零。**

**成本趋势:** 随着人工智能能力的实现, 获得特定智能水平的成本将持续下降, 并趋向于零。这种趋势得益于\*\*自适应计算 (adaptive compute) \*\*的应用, 允许根据任务难度调整计算量, 从而无需持续扩大模型规模。

**即时知识:** 任何**公开可用的信息**都将能够**即时获取**。例如, 在“代理时代” (agents era), 即使是难度较高的知识查询 (如查询某个城市特定年份的结婚人数), 也能在几分钟内解决。

**影响:** 知识的商品化将促进以前受知识壁垒限制的领域**民主化**, 例如**编程**和**个人健康**。同时, 由于公共信息的相对成本降低, **私人内部信息的相对价值将相应提高。**

## 目录

## 文章信息

字数

阅读时间

发布时间

更新时间

## 标签

#AI

#Jason wei

## 2. 验证者法则 (Verifiers Law)

验证者法则指出，**训练人工智能以完成特定任务的能力，与其验证该任务的难易程度成正比。**

**核心主张：**任何可解决且**易于验证**的任务，最终都将被 AI 征服。最先被自动化的任务，将是那些**验证起来极其简单** (trivial to verify) 的任务。

**验证标准：**任务的可验证性取决于五个因素：是否存在客观真相、验证速度、是否可扩展到同时验证数百万个响应、噪音是否低，以及是否提供了连续奖励（即响应质量的整个光谱，而不仅仅是“通过/未通过”）。

**不对称性：**这一法则基于**验证的不对称性 (asymmetry of verification)**，即对于某些任务来说，验证解决方案要比找到解决方案容易得多（例如，数独难题，或编写运行如 Twitter 等大型网站的代码）。

**影响：**测量能力（即验证能力）是推动 AI 进步的关键驱动因素。因此，**创建可供 AI 优化的测量方法或指标**将是未来的一个重要增长领域。

## 3. 智能的锯齿状边缘 (The Jagged Edge of Intelligence)

这一理念认为，**人工智能在特定任务上的能力和改进速度将根据任务的某些属性而有所不同。**

**反驳快速爆发论：**智能的边缘是“锯齿状”的，它否定了**快速超级智能爆发 (fast takeoff)** 的假设，因为 AI 的进步是一个渐进的、在不同任务间不平衡的过程，而不是一个突然的、全面的飞跃。

**能力差异：**AI 的能力在某些任务上可能达到“波峰”（如复杂的数学问题或竞赛编程），而在另一些任务上则处于“波谷”（如克林基特语 (Tlingit) 等数据稀少的语言）。

**改进速度的启发式因素：**AI 对任务的影响和改进速度最大化发生在符合以下特点的任务上：

**数字化任务：**迭代速度更快，计算可更容易扩展。

**对人类而言更容易的任务：**对人类来说越容易的任务，AI 也越容易完成。

**数据丰富的任务：**数据越丰富，AI 表现越好（但明确的客观指标可通过强化学习生成合成数据来弥补数据不足）。

**影响：**某些领域如**软件开发**将得到极大的加速，而其他缺乏数字化和数据丰富的领域（如**理发**或**传统地毯制作**）将可能在短期内保持不受影响。

分享这篇文章



## 相关文章推荐

### Agent训练 新范式: ...

传统AI训练像是  
把人类所有的...

### AI Context Engineeri...

这里将收集  
Context...

### 小模型逆 袭: 秘密...

该论文介绍了一个  
名 ...