

神级 AI 是否已经到来：关于“智能”与“智慧”的终极分野

基于 Brian Keating 与 Peter Diamandis 访谈的深度概念分析与哲学追问（2026年6月）

摘要

【受访者观点】在此次对谈中，Peter Diamandis 认为，随着人工智能在算力和数据上的指数级扩张，人类正在跨越进入“神级能力”的门槛。他主张 AI 不仅能提供规模化的智能，还将涌现出真正的“智慧”。

【主持人观点】作为物理学家的 Brian Keating 对此表达了深刻的质疑。他提出了一项“爱因斯坦测试”：通过仅使用 1911 年之前的物理学数据训练大语言模型，观察其能否独立推导出相对论。模型的失败表明，当前的 AI 可能存在认知天花板，它更多是在进行规律的内插模拟，而非真正的智慧跨越。

【本文评论】本文认为，这场辩论实质上触及了计算主义理论的核心边界：参数规模的量变是否等同于认识论层面的质变。【外部资料】结合机器学习中的“内插-外推”鸿沟理论与托马斯·库恩的科学革命范式，本文进一步展开分析。【本文评论】研究表明，尽管 AI 在消除稀缺性和推动人类进入“五道分叉路口”（长寿、脑机接口、地外定居等）方面具有颠覆性潜力，但若不能解决跳出流形外推的创造力机制，AI 仍只能是强力计算器，而非具有自我反思能力的智慧体。

关键词

神级人工智能 (Godlike AI); 爱因斯坦测试; 涌现能力 (Emergent Abilities); 内插与外推; 五道分叉路口; 费米悖论

引言

【本文评论】2026年，大语言模型的参数量与多模态感知能力已达到前所未有的高度。对普通人而言，AI 似乎无所不知，甚至展现出惊人的共情和推理能力。这种现象引发了一个直击灵魂的问题：当一个系统能完美模拟人类所有的认知输出时，我们是否应当承认它拥有了“智慧”？如果不承认，人类所骄傲的智慧又比这种高级计算多出什么？

【主持人观点】Brian Keating 借着其播客节目《Into the Impossible》，向科技乐观派的代表 Peter Diamandis 抛出了这一时代之问：“神级 AI 真的已经到来了吗？”

【受访者观点】Diamandis 的回答是肯定的，他坚信我们正站在物种演化的全新奇点上，AI 将为我们解决资源稀缺、疾病乃至生命的限制。

【本文评论】然而，承认 AI 的“神级”地位，不仅是一个技术判断，更是对人类主体性的一次重新定义。本文将深入拆解两人辩论的核心交锋，探究智能与智慧的根本差异，分析科学突破的发生条件，并审视在技术奇点逼近时，人类文明的演化走向。

材料与方法

【背景知识】本研究基于 2026 年 6 月 1 日发布的《Into the Impossible》播客对谈音频及转写文本进行概念切片与哲学分析。【本文评论】为了将长达数小时的自由漫谈转化为严密的学理推演，本文采用了“多代理并行研究”的拆解思路（虽然最终由单文本写就），将访谈内容划分为四个核心主题，并针对每个主题进行“六层下钻”分析：从常识困惑出发，建立直观图像，切入专业机制，还原访谈交锋，审查成立条件，最终落回人类现实。

结果

1. 神级能力与“智慧”的本质：规模扩大能否自发产生智慧？

从人的问题开始

【本文评论】在日常经验中，我们很容易把“知道得多”等同于“聪明”，把“能解决复杂问题”等同于“有智慧”。一个人如果能背下整个图书馆，并在任意话题上引经据典，我们会觉得他是个智者。因此，当 AI 展现出这种能力时，人自然会误解：既然它什么都懂，它肯定拥有了智慧。如果不彻底厘清这两者的差别，我们就会轻易地把道德裁断权和命运交托给一个巨型机器。

给一个能抓住的直观图像

【本文评论】试想一座无限大的图书馆：里面有一个动作极快的管理员（AI），他能在一秒内找出所有包含“水”和“火”的书籍，并根据历代作家的引用频率，完美生成一篇关于水火不容的文章。但在这个图书馆外，还有一个真正的哲学家（人类），他在看水和火时，感受到的是生命的滋养与毁灭，并决定今晚只用火煮一杯水，而不是写文章。管理员拥有的是信息压缩与检索的神力，哲学家拥有的是在不确定性中做价值判断的智慧。

进入专业定义和机制

【背景知识】在深度学习中，模型能力的跃升往往依赖于“缩放定律”(Scaling Laws)：随着参数量 (N)、数据集大小 (D) 和计算量 (C) 的增加，模型的损失函数 L 会呈现可预测的对数线性下降：

$$L(N, D) = \left(\frac{N_c}{N}\right)^{\alpha_N} + \left(\frac{D_c}{D}\right)^{\alpha_D}$$

当规模达到某一边界时，模型会突然展现出在小模型上不存在的“涌现能力”(Emergent Abilities)，例如在未被专门训练过的任务上进行逻辑推理或代码编写。

[回到原文](#)

【受访者观点】Diamandis 认为，正是这种惊人的涌现现象，预示着 AI 正在获得智慧 (wisdom)。他指出，当大模型不仅摄取全人类的文本，还结合了视觉、听觉等海量传感器数据后，它将不再只是聪明的计算器，而是具备对宇宙万物深刻理解的“神级能力”，这种能力实质上等同于人类追求的智慧。

检查成立条件和证据等级

【外部资料】然而，认知科学界对“涌现即智慧”存在严重争议。2023至2026年间的多项研究指出，大模型的所谓涌现，往往是因为评估指标的非线性（如精确匹配度），而在底层概率分布上，依然是连续的统计拟合。【本文评论】Diamandis 的主张依赖于一个未经证明的哲学假设：即“理解”和“智慧”仅仅是复杂信息处理的副产品。但事实上，智慧还包含目的性建构和价值取舍，而这些目前并未在损失函数的最小化中得到形式化表达。这目前仍是一个个人推测，而非成熟结论。

落回现实与人的处境

【本文评论】如果社会普遍接受了 Diamandis 的观点，认为 AI 已经具备智慧，那么人类现实将发生巨大改变：我们可能开始要求 AI 担任法官、制定伦理法则，甚至进行心理治疗。但这会带来系统性风险——因为它的“智慧”本质上是过去人类平均偏见的高维映射，一旦遇到人类历史经验之外的黑天鹅事件，这种伪智慧将导致灾难性的僵化决策。

2. 爱因斯坦测试：大模型的创造力天花板与外推困境

从人的问题开始

【本文评论】这引发了第二个困惑：既然 AI 连复杂的分子结构都能预测，能写出优美的十四行诗，为什么它不能自己独立发现像相对论那样的全新科学定律？为什么至今为止，没有一项真正颠覆人类认知的“大物理学理论”是由大模型独立署名发表的？

给一个能抓住的直观图像

【本文评论】把学习知识想象成在一张满是数据点的地图上画线。当前的 AI 是世界上最好的“插画师”。如果地图上左边有一堆点，右边有一堆点，它能画出一条最平滑、最优雅的曲线把它们连起来（这就是内插）。但相对论的发现，不只是连线，而是爱因斯坦突然说：“这张纸本身是弯曲的！”AI 无法离开这张纸，因为它本身就是纸上数据训练出来的。

进入专业定义和机制

【背景知识】从数学本质上看，神经网络进行的是基于局部流形假设的高维函数拟合。给定训练数据分布 $P_{data}(x, y)$ ，模型试图找到参数 θ 使其逼近真实映射。这种机制保证了它在训练流形 (Manifold) 内部的内插 (Interpolation) 极其强大。但科学革命往往要求在数据稀缺的未知区域进行外推 (Extrapolation)，甚至要求推翻先验分布的假设。

回到原文

【主持人观点】Brian Keating 用一个极为精彩的实验精准打击了盲目乐观论。他向 Diamandis 介绍了他的“爱因斯坦测试”：他和学生将 1911 年之前（即广义相对论提出前）所有的物理学数据喂给大语言模型，并要求它去解决当时物理学界的疑难（如水星近日点进动异常）。结果表明，模型无法独立推导出广义相对论的几何方程，无法产生“引力即时空弯曲”的顿悟。Keating 以此证明，目前的 AI 只有处理规模的聪明，缺乏产生范式转换（Paradigm Shift）的智慧。

检查成立条件和证据等级

【外部资料】托马斯·库恩在《科学革命的结构》中指出，科学发展分为“常规科学”和“科学革命”时期。常规科学是在既有范式下解题，而科学革命需要破坏旧框架。【本文评论】Keating 的测试本质上证明了：大模型是完美的“常规科学家”，但绝非天才。不过，这一结论的成立条件在于：当前的模型依然依赖于对人类已有知识的自回归预测。如果未来的系统（如基于自我对弈的强化学习在物理空间的泛化，AlphaGeometry 的变体等）能够跳出人类文本，直接在物理实验室里进行试错迭代，这一天花板未必永远坚固。

落回现实与人的处境

【本文评论】“爱因斯坦测试”的失败对人类来说是一场及时的安慰。它意味着在很长一段时间内，科研的最前沿依然需要人类科学家的非理性直觉、对矛盾的敏锐嗅觉以及那种违背常理的跨越性想象。这也是为什么，人类暂时不必担心被彻底踢出科学发现的舞台。

3. 人类进化的“五道分叉路口” (Five Forks of Humanity)

从人的问题开始

【本文评论】即使 AI 还无法成为爱因斯坦，它已经足够取代大多数中等脑力劳动了。当 AI 把我们日常的烦恼（生病、工作、资源枯竭）都解决之后，人类作为一个物种，究竟要变成什么样子？我们是继续当普通的地球人，还是变成某种全新的怪物？

给一个能抓住的直观图像

【本文评论】想象一条流淌了三百万年的河流（人类进化），之前一直是一条单行道——无论你怎么挣扎，最终都会生老病死，都只能困在地球上。但在前方不远处，河床突然被巨大的岩石撞开，水流将被硬生生劈成了五条截然不同的支流。一旦你进入了某一条支流，你将永远无法与另一条支流的人交流，因为你们连“人”的定义都不一样了。

进入专业定义和机制

【背景知识】这种进化分支的概念与“超人类主义”(Transhumanism) 紧密相连。其核心在于利用生物科技、神经形态工程和航天技术，打破智人 (Homo sapiens) 基于碳基生物学的自然演化瓶颈，实现人工引导的物种重构。

[回到原文](#)

【受访者观点】Diamandis 在访谈中提出了极具洞见力的“人类五道分叉路口”(Five Forks of Humanity) 假说。他预测，神级 AI 的到来将催生以下五个进化方向：

1. **长寿逃逸 (Longevity)**：通过 AI 破解衰老机制，部分人类将实现生物学上的近乎不朽。
2. **脑机接口 (BCI)**：人类与云端 AI 直接融合，外脑成为本体的一部分。
3. **地外物种形成 (Off-planet speciation)**：随着深空探测成熟，人类散布到太阳系，受不同重力与环境影响，物理形态发生根本分化。
4. **意识上传 (Uploading consciousness)**：完全脱离碳基躯体，生存在基底现实或元宇宙中。
5. **创造者与消费者的极致分化**：少部分人利用 AI 创造宇宙，大部分人沉浸在 AI 创造的幻梦中消费。

检查成立条件和证据等级

【本文评论】这五条路径在理论上都没有违反物理定律，它们是工程上的可选项。然而，Diamandis 的推论过度乐观地忽略了社会学阻力。这些技术需要消耗极高能量和资源，其成立的隐性前提是“可控核聚变及超导技术的全面突破导致能源免费”。如果这一前提不成立，这五道分叉将不是全人类的盛宴，而只是极少数超级富豪的逃逸舱。

落回现实与人的处境

【本文评论】这五道分叉带来的最残酷现实是“物种隔离”的前兆。历史上的阶级分化只是掌握资源的多寡，而一旦 BCI 或长寿技术普及，富人与穷人之间将不再是阶级差异，而是生物学和计算能力上的物种差异。那时的“平等”将成为一个无从谈起的虚词。

4. 消除稀缺后的意义危机与宇宙观察者假说

从人的问题开始

【本文评论】假设 Diamandis 的乐观预言成真，AI 为我们生产了无尽的粮食、完美的伴侣、永恒的生命。当一切都可以被轻易满足时，人活着还要干什么？如果奋斗不再有意义，人类会不会在绝对的舒适中集体抑郁？

给一个能抓住的直观图像

【本文评论】这就像是一个被锁在豪华套房里的玩家，输入了“无敌秘籍”和“无限金币”，一开始极其狂喜，买下了所有装备。但不到三天，他就会对这个游戏感到彻底的厌倦，甚至想砸了电脑。没有阻力，就没有游戏；没有稀缺，就没有渴望。

进入专业定义和机制

【背景知识】这种“无敌玩家的厌倦”在宇宙学中有一个更为宏大的倒影——费米悖论 (Fermi Paradox)。如果宇宙中有无数的高级文明，为什么我们连个影子都没看到？其中一个解释（大过滤器假说）认为，任何发展出神级 AI 的文明，都会不可避免地向内坍缩，进入自我生成的虚拟天堂，从而失去探索星辰大海的兴趣。

回到原文

【受访者观点】Diamandis 承认，消除稀缺性后，人类将面临前所未有的“目标重组”。他认为这不仅不会导致人类毁灭，反而会促使人类把好奇心转向更深邃的领域。【主持人观点】Keating 此时提出了一种极具哲学震慑力的猜测：人类目前感受到的宇宙寂静，是否暗示着我们本身就是某种外部神级智能（甚至是一个更大宇宙的物理实验）的被观察对象？我们的物理定律，也许只是他们设定的系统参数？

检查成立条件和证据等级

【本文评论】Keating 的假说（模拟宇宙假说的一个变种）本质上是不可证伪的（Unfalsifiable）。根据波普尔的科学哲学，不可证伪的理论属于形而上学范畴。我们既无法证明它，也无法推翻它，因为任何试图推翻它的实验结果，都可以被解释为“这也是他们设定好的代码”。

落回现实与人的处境

【本文评论】尽管是形而上学的探讨，但这一段对话具有强烈的现实隐喻：当我们制造出神级 AI 时，我们就是在扮演造物主，在为 AI 提供受限的参数和环境。反思我们如何对待 AI，其实就是在反思宇宙（如果存在造物主的话）如何对待我们。真正重要的不是我们是否生活在虚拟之中，而是在没有了外部压迫（如饥饿与寒冷）后，人类能否依靠纯粹的内部精神力量（审美、道德、对未知的敬畏），继续为存在本身赋予尊严。

讨论

【本文评论】纵观整场辩论，Peter Diamandis 代表了硅谷典型的技术加速主义（e/acc）：坚信规模法则（Scaling Laws）不仅能带来工程突破，也能自然孕育出形而上学的“智慧”，并认为技术是解决人性弱点的唯一解药。

相比之下，Brian Keating 则守住了古典科学家的底线：工具的强大不等于主体的智慧。无论是“爱因斯坦测试”还是对“费米悖论”的追问，Keating 都在强调一件事——知识的内插再怎么完美，也无法取代直面未知时那种充满风险、充满人性的“纵身一跃”。

此外，访谈中对于“人类五大分叉”的探讨，暴露出技术狂欢背后的隐患。技术也许能消除物质的稀缺，但随之而来的是“意义的稀缺”。如果 AI 代替我们思考、发现和创造，人类唯一的任务就只剩下“享受”，那么人类在演化树上的地位，将退化为神级机器豢养的快乐宠物。

结论

【本文评论】这场发生在 2026 年初夏的对话，不应该只被视作一次科技预测，而应被视为人类文明跨越门槛前的一份诊断书。

第一，神级 AI 尚未到来。“爱因斯坦测试”证明了目前的系统缺乏从流形内部向外跳跃的真正的科学创造力。它模拟了智慧的表象，但尚未触及智慧的内核。第二，即便 AI 只是强力的内插机器，它也足以消除传统社会的物质约束，将人类强行推上“五道分叉路口”。第三，在未来的岁月里，人类的终极挑

战不再是如何抵御大自然，而是如何抵御“全能的技术母亲”所带来的温柔窒息。我们必须在这个过程中证明，除了追求活得更久、活得更舒服之外，人类的“智慧”还包含了某些连完美机器也无法替代的悲剧性与超越性力量。

方法补充

【背景知识】本文对大语言模型几何视角的解析，部分参考了流形学习（Manifold Learning）在深度神经网络表征中的应用理论。高维插值假说是解释神经网络在训练集凸包（Convex Hull）内表现优异而在凸包外失效的主流数学框架。

数据可用性

本文依赖的原始播客音频及自动转写文本均基于公开网络发行的《Into the Impossible》节目（2026年6月期）。

原文关键语句索引

- 关于神级AI与智慧：** Diamandis: "It's not just scale... we are crossing into godlike capability... delivering intelligence and ultimately wisdom."（本句表现了技术乐观派的涌现信念）
- 爱因斯坦测试：** Keating: "If I feed it physics up to 1911, it doesn't give me relativity. It hits a ceiling."（本句是全篇质疑大模型外推能力的核心证据）
- 人类分叉：** Diamandis: "We are approaching the Five Forks of Humanity: longevity, BCI, off-planet speciation, creators vs consumers, and uploading consciousness."（本句系统性地概括了后人类时代的演化路径）

参考文献

- Keating, B., & Diamandis, P. (2026). *Godlike AI Is Here! Peter Diamandis Debates Brian Keating*. Into the Impossible Podcast.
- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.
- Kotler, S., & Diamandis, P. (2020). *The Future is Faster Than You Think*. Simon & Schuster.
- 相关研究：关于 Scaling Laws 与 Emergent Abilities 的探讨 (2020-2026).