

《奇点更近》2：人与 AI 融合

万维钢·精英日课 6（年度日更）

雷·库兹韦尔的《奇点更近》这本书中最独特的设想，是通过脑机接口，让人脑和 AI 融合在一起。这一讲咱们说说它的必要性和可行性。

希腊神话中有个女神叫卡桑德拉（**Cassandra**），她是特洛伊的公主、阿波罗的祭司，擅长对未来进行预言。可是她又受到某种诅咒，使得没有人相信她的预言。卡桑德拉的预言往往跟灾难有关。



库兹韦尔在书的结尾跟卡桑德拉有一番对话，不过这一次是卡桑德拉提问，库兹韦尔预言。这番对话的核心就是人与 AI 的融合。

卡桑德拉说：如果 **AI** 如此强大，它自己就能完成各种工作，那它为什么还要与人类大脑连接呢？

你想想是不是这样，人难道不是一个累赘吗？我们为什么不坐等 **AI** 把事情办好，然后向我们汇报呢？

库兹韦尔的回答是，**这纯粹是因为人类自身的需要。我们需要人生的目的和意义，为此我们必须参与工作，可是我们的大脑又跟不上 **AI** 的节奏，所以我们必须跟 **AI** 融合。**



一开始我不太喜欢库兹韦尔这个建议，因为我认为哪怕不搞人机融合，人也有意义。我们专栏一再讲，哪怕 **AI** 再强，人也有人的用处。人的作用主要是两个方面 ——

一个是输出主动性。就是你的人生经历、你所在文化的历史、你的基因，所有的一切微妙元素共同决定了你喜欢什么、想要什么。因为涉及到的因素实在太多也太细微，其中很多早已无法数字化，所以 **AI** 永远都没办法预测你的喜好。而且因为 **AI** 自身没有历史也没有人生经历，所以它们没有自己的人生意义，所以它们应该服从我们的人生意义和我们的价值观。因为我们是历史演化的产物而 **AI** 可以随便复制，所以我们比 **AI** 更宝贵。

另一个是拍板做决定。哪怕 **AI** 能力再强，因为最后承担决策后果的是我们而不是他们，所以它们只能建议，最终的决策权必须在我们手上。

库兹韦尔没讲这两个意义，他只是说如果人不参与工作，会感到很难受。我对工作不工作倒不是很担心。

但是上周末我去看了电影《异形：夺命舰》，其中有很多人和“人造人（artificial person）”互动的剧情 —— 我越看，就越觉得很多决策直接交给 AI 似乎更好.....人反应太慢，而且有时候不是很理性 —— 但是女主角大多数时候原则性强而且经常获得灵感，智商绝对高于 AI。



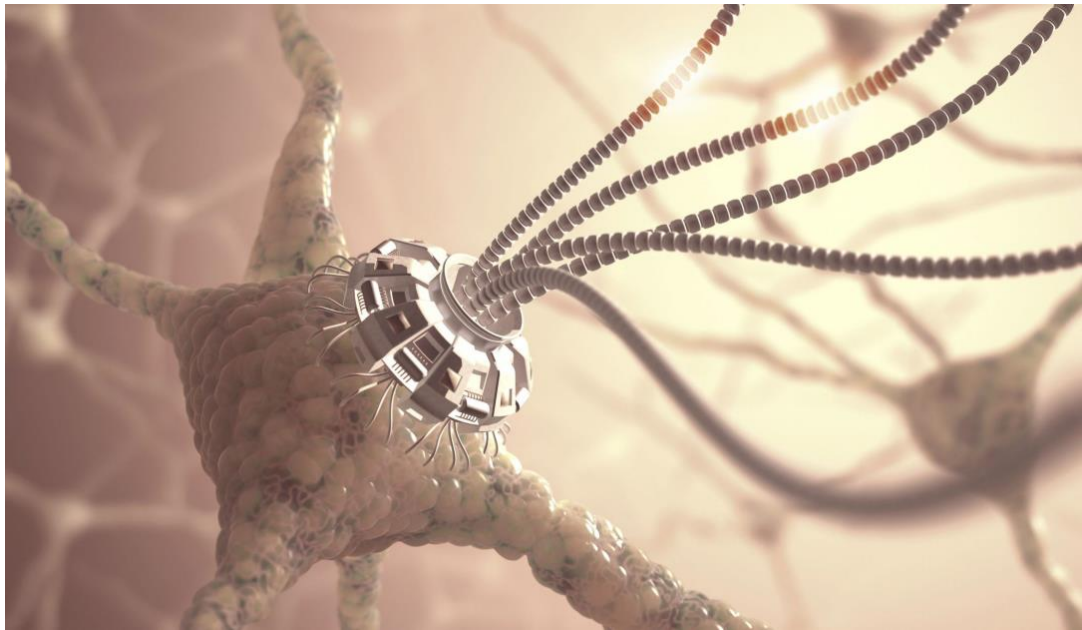


不管你怎么做思想实验，把人的价值观和灵感与 **AI** 的算力、判断力和执行力结合起来总是好的。尤其在危急或者多变的时刻，人必须做出实时的决策，等 **AI** 用语音汇报都来不及，还是用脑机接口直接心灵感应最方便。

所以未来的局面会比科幻电影更激进，只是场面上可能没有那么多戏剧冲突：到时候根本分不清谁是人谁是 **AI**。

✱

库兹韦尔的建议是把云端的 **AI** 算力作为大脑新皮层的扩展。这样我们在自身大脑的新皮层之外，又通过 **AI** 增加了更多、更快的新皮层，从而大幅提升思维能力，也许提高数百万倍。



我们专栏刚刚讲过麦克斯·班尼特的《智能简史》，我们知道新皮层跟大脑最近的几次突破都有关系，代表最高的智能。我们能模拟，能有心智理论，能理解抽象概念、能通过预言和他人交流、能建立共同的想象，都是因为新皮层。我们跟黑猩猩的大脑结构几乎完全一样，只是新皮层比它们多，就能随意使用它们完全无法理解的概念——那如果我们自己的新皮层再扩大百倍千倍，又是一种什么境界呢？

而且让大脑直接接入 **AI**，能突破输入输出的瓶颈问题。你想过没有，我们大脑的能力很可能受到了输入输出方法的限制。你有一个想法要表达出来，现在要么是通过语言、要么是通过动作，这两个方式都比较慢，而且往往不精确。有时候你心里有，但是你表达不出来，找不着合适的词，或者没有那个技能。

比如明朝著名文化人王世贞，是个书画鉴赏家，传说他就是《金瓶梅》的作者。他非常懂，也很喜欢谈论书法，但是自己写不好。所以他有句

话叫「吾腕有鬼，吾眼有神」：我的审美力是神级的，但是我的手没练好，表达不出来。

如果放到现在，王世贞适合用 **AI** 作画。那何必非得用语言表达呢？何必非得先把你的意思翻译成一个具体的动作或者笔画，再交给外界处理呢？如果你的大脑直接跟 **AI** 连接，只要心意一动，设计图上若干个元素立即变化，那是什么境界？你出图可能不比 **AI** 慢多少，而且完全合心合意。

更好的是，大脑与 **AI** 融合之后，也许我们能感知到各种新的、之前没有的感觉。

爱德华·威尔逊在《创造的本源》[1] 一书中表达过这个意思：人类的文艺受自身输入输出能力的限制太大了。很多动物拥有的视觉、听觉、嗅觉范围我们都没有达到，如果有一种技术能把动物的那些体验直接传递给人的大脑，那又是什么样的感觉。

库兹韦尔也设想，与 **AI** 融合之后，我们将可以欣赏更为复杂无序的事物、非语言的想法和难以名状的美感。那时的文化艺术，该有何等的发展。

再比如说，现在 **AI** 的思考过程对人类来说是一个黑盒子，那如果我们与 **AI** 融合，也许就能直接感知它们的情绪和直觉。比如说，这个化学分子式给你一个什么感觉？那个磁场位形为什么让你紧张呢？我们到时候也许能理解和欣赏一些现在无法想象的高难度抽象概念.....

但我们必须先解决脑机接口的一些技术问题。



马斯克的 **Neuralink** 公司搞脑机接口，终极目标就是让人和 **AI** 融合。路线很清楚，但是有好几个难点。

比如信号翻译问题。我们现在对外交流主要通过语言，而我们讲了，语言表达不是天然的，是后天在各个文化系统中约定俗成的，需要每个孩子从小学习。对大脑来说，这意味着不同的人的大脑活动表达的意思是不一样的。

你怎么知道这几个脑区的活动代表哪个意思呢？从神经活动到语义的映射并不是直接了当的，不同的人不一样，而且考虑到大脑的可塑性，同一个人在不同的时间、环境或者情境下，大脑活动和语义之间的对应关系也可能发生变化。你怎么才能一看大脑就知道 he 想说啥呢？

现在的办法是机器学习。给每个人弄一套数据集，训练一个单独的 **AI** 大脑翻译界面。**Facebook** 赞助的一项研究在 2020 年取得了效果。给人戴个头盔，让人默念一些口语句子，通过机器学习把捕捉到的电信号和句子里的单词做一一对应。



这个研究达到了一定的准确度。也就是说你只要想，计算机就可以通过头盔知道你想的是哪个词。不过 Facebook 在 2021 年停止了这个项目。

之所以停止，也许是因为这种非侵入式的脑机接口的分辨率太低，前景有限。

✱

从大脑外面测量大脑活动，只有两个办法。

一个是用功能性核磁共振扫描（**fMRI**），这里扫描的是大脑各处的**血液流量**。哪个脑区血液流量有变化，哪个脑区就在搞活动。这种成像的位置比较精确，但是时间有延迟。关键是我们思考靠的是神经元之间互相发射电信号，而血液流量变化都是电信号之后发生的，中间有少则几百

毫秒、多则几秒的延迟。这就使得你难以精确把握哪个流量对应的是哪个想法。

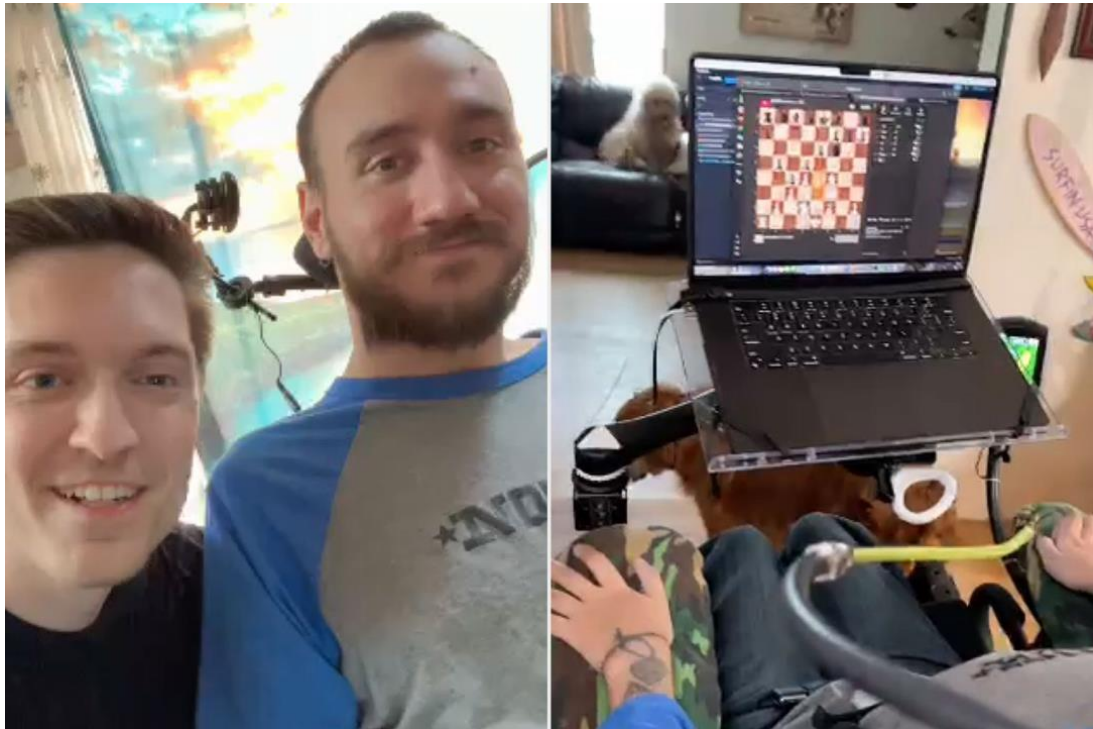
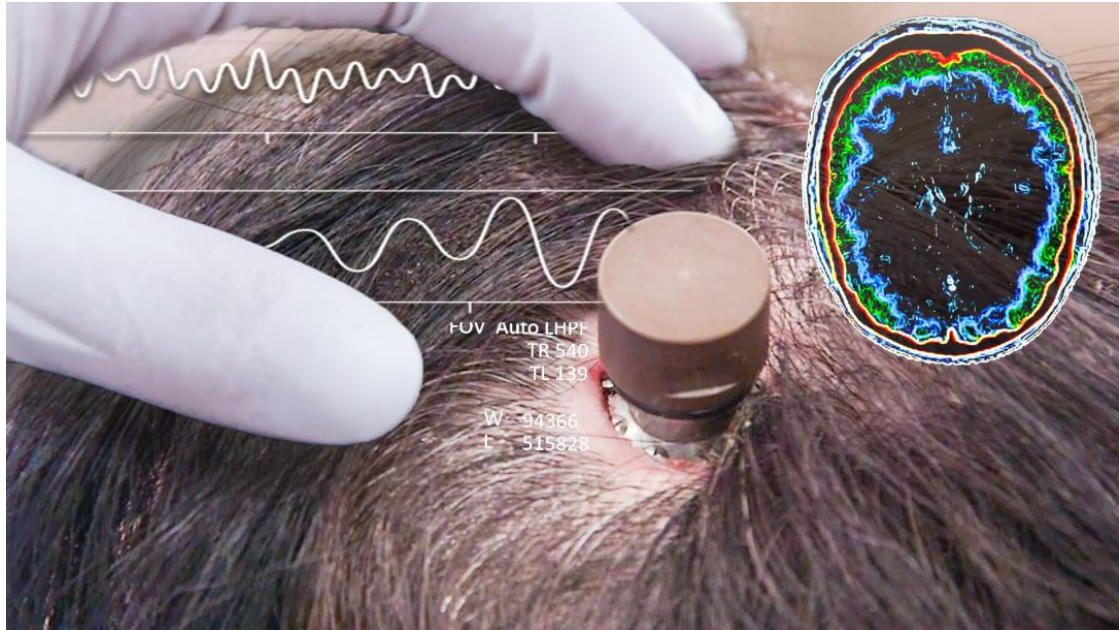
另一个办法是戴头盔测量脑电波。这回是直接测电信号，没有时间延迟，但是因为中间隔着头骨，你很难精确测定一个信号是从大脑中哪个位置传出来的。脑电波信号的空间分辨率很低，几个立方厘米的区域中有无数个神经元，你无法描写精细的想法。

更何况这些还都只是输出。如果再考虑到从 **AI** 到大脑的信息输入，非侵入方法就更没戏了。

要想实时、精确读取大脑信息，乃至给大脑输入信息，就必须突破头骨，直接跟大脑接触。这就是「植入式脑机接口」（**implanted brain-computer interface, BCI**）。

2024 年一月，马斯克的 Neuralink 在美国食品药品监督管理局的允许下，把一个芯片植入了一个人的大脑。





这个装置有 **1024** 个电极，接入大脑的运动皮层，能处理神经元信号。所有信号被传递到植入的芯片之中，这个芯片做些处理，再通过无线网络传输到外边的计算机上。为了确保手术顺利，**Neuralink** 甚至专门开发了一个手术机器人。

实验取得了成功。受试者现在可以用大脑直接控制鼠标，可以在手机上打字。

这只是第一步，只考虑动作意念的输出。如果要考虑所有的大脑信号，未来的脑机接口需要在整个新皮层中布满电极 —— 如果用现在接线方式那个手术可就太大了。

库兹韦尔认为唯一的办法是等到 **2030** 年代，纳米机器人成熟之后，以血液注射的方式让纳米机器人遍布大脑，不但能接受而且能对大脑传递信号。到时候人们不需要做手术就可以接上云端的扩展大脑。

如果我们延续《智能简史》的节奏，那将是大脑演化的第六次突破。



在库兹韦尔和卡桑德拉的对话中，卡桑德拉还是给了一个预言，那就是政府监管。卡桑德拉说因为脑机接口的操作必须往大脑中植入异物，不管你说的多安全，政府也不太可能大力支持。所以她认为人与 **AI** 的融合将会因为监管法规而被推迟十年。

这个猜测很有道理。现在政府批准做脑机接口实验的受试者都是瘫痪病人，他们因为比如说脊髓损伤，或者是渐冻症患者，而不能把大脑的动作指令传递到身体四肢。对这些患者来说接入脑机接口的潜在好处大于风险。而对健康的人来说，那是一点点风险都必须非常慎重才行。

我的看法是前面还有太多不确定性。如果脑机接口只是可以让人用意念移动光标或者机械四肢，那对健康人来说没什么价值。而脑机接口到底能不能实现比语言和动作、比人的五感更快更准确的输入输出，现在还没有实验证据。

但是很明显，人脑要升级，就非得走这条路。



注释

[1] 《创造的起源》4：科学的文艺和文艺的科学

划重点

- 1.我们需要人生的目的和意义，为此我们必须参与工作，可是我们的大脑又跟不上 AI 的节奏，所以我们必须跟 AI 融合。
- 2.库兹韦尔的建议是把云端的 AI 算力作为大脑新皮层的扩展。
- 3.要想实时、精确读取大脑信息，乃至给大脑输入信息，就必须突破头骨，直接跟大脑接触。这就是「植入式脑机接口」。