

AI 的发展趋势

如果你只有几分钟的时间，以下是投资者、运营者以及创始人应该了解的最令人兴奋的 AI 趋势。

- 人手一位机器人副驾。AI 已经在简化插图、协作与写代码的工作。很快 AI 就会变成所有知识员工的助手。未来，律师、金融分析师、建筑师等都会有自己版本的 GitHub “Copilot”。
- 跟踪价值转移。由于 AI 初创企业往往要依赖于像 GPT-3、Codex 这样的公开模型，有人对其有没有护城河提出了质疑。核心问题与谁俘获了价值有关。利用 GPT-3 的应用能不能成功捕捉到价值？还是说最后还是落到基础设施层手上？
- 超越文字与图像。因为文字与图像自动化创作的能力，GPT-3 与 DALLE-2 赢得了大量关注。但 AI 最有影响力的应用也许来自于生命科学。AI 可以用来设计出更好的药物，或者进行更高效的临床实验。
- 界面的改善。与 AI 的交互往往是简单的文本框的方式，用户输入“提示”，机器返回结果。虽说这种用法很简单，但需要有更大的控制才能释放这种技术的威力。面临的挑战是在促进这种潜能的同时避免引入不必要的复杂性。应用需要流畅、富有创意的界面才能火起来。
- 解决劳动力短缺。熟练蓝领供应短缺，但社会的需求却在增加。比方说，熟练焊工每年都在以 7% 的速度减少，但对这些焊工的需求却以 4% 的速度在增长。有 AI 支持的机器人也许可以成为解决方案的一部分，焊接、建造等人工也许都可以实现自动化。

“这次不一样。”



约翰·邓普顿爵士（Sir John Templeton），1999年被《Money》杂志评为“世纪选股人”。他说，“这是投资最危险的5个字。”

说得好，很妙的讽刺。市场经常会出现海市蜃楼，看似异常的情况其实稀松平常——那不过是我们熟悉的周而复始其中的一个步骤而已。

但是，有时候情况确实会不一样。有时候星星之火也会燎原。有时候这个世界确实变了。

说到人工智能（AI），风投界的看法是，我们也许正处于这样一个时刻。过去一年的时间里，随着公众兴趣的增加，新的模式与初创公司如雨后春笋般涌现。在整个行业的风险投资随着大盘回落而放缓之际，我们与风投家进行交流，去了解他们最感兴趣的是什么，以及他们对经常被提及的生成式人工智能最兴奋的是什么。

一如既往，这是我们回顾这段时间寻找虚幻曙光的机会——是资本在冷却中寻找热度的结果。但感觉那种情况不太可能。我从事第一份风投工作的时间是在 2016 年，那时候每一个 pitch deck 都是自己具备一点人工智能优势，而聊天机器人则被看作是用户体验的进化。现在玩 DALLE-2、GPT-3 和 Stable Diffusion 的感觉跟那个时代截然不同，相当于从一个会说话的洋娃娃跃进到一位早熟的幼儿。人工智能正在释放真正的创造力与真正的商业价值，可生成新颖的图像、合理的文字以及可用的代码。随着一代代模型长江后浪推前浪，随着初创企业找到利用它们的新方法，大量的创新和实验往往令人感觉应接不暇。今天可能性的地平线看起来还很遥远，但也许几周之后就会突兀地呈现在你面前。

为了更好地了解行业现状，我邀请了十位思想深刻的 AI 投资者分享他们认为值得关注的趋势。我希望这些洞察能够帮助我们更好地确定机会在哪里，以及哪些主题值得进一步研究。

关于此次交流的一点说明。

虽然投资者知道其他撰稿人都写了什么，并鼓励我选择不同的主题，但我发现其实谈一谈重复的东西往往很有趣。同一个主题两位投资者的解读方式可能会截然不同，而这种差异是有价值的。

此外，投资者提到所投资的公司我也有意保留下来。一切都是权衡取舍的问题，我相信收益会大于感知到的成本。这种方法的缺点是，投资者可能会被看成为自己的投资组合“背书”。首先，我们挑选的是我认为有思想且可靠的贡献者。其次，让投资者选择自己最了解、研究最深入的公司会更有趣。我们还要求他们在喜欢的公司当中做出选择。最后，这表明他们利益攸关，充分利用了自己的信念。

好了，接下来我们就一起掀开 AI 的兔子洞，去了解这一新技术是如何影响我们的思想、身体和机器的。

让人类的工作站上新台阶

有没有什么职业是像“艺术家”这样属于典型的右脑型的？或者像“程

程序员”这样典型的左脑型职业？

如果用四个字总结去年的特点的话，那就是快速演进，尤其是大型语言模型的快速演进。这种人工智能的快速演进尤其引人注目的是，现在它们开始为辅助工具提供动力，从而在根本上提高各行各业的生产力、影响力与价值。

对于艺术家，我们有 AI 图像生成工具，比方说 OpenAI 的 DALL-E、Midjourney 等。对于程序员，我们有微软的 GitHub Copilot，它可以帮助软件开发者用许多当前最流行的计算机语言编写、测试和改进代码。

虽然一些 AI 怀疑论者把大型语言模型说成是暴力破解的预测机器，永远也没法给计算机注入人类的智能或意识等任何东西，但在令人兴奋的实践中，我们所看到的是，这些类型的 AI 工具在促进人类繁荣方面，已经开始发挥着影响极为深远深刻的作用。

Copilot 为开发者以及 DALL-E 为各种视觉创意人员所做的，是减少或消除他们工作当中那些机械式、很耗时但仍然至关重要的事情。当然，这种情况并不是软件开发者以及艺术家所独有的。大型语言模型针对大量文本数据进行训练，然后结合它们“学习”的内容，在用户提供的提示下，生成统计上可能的（上下文敏感的）输出。因此，虽然 Github Copilot 是通过摄取大量计算机代码训练而成，但不同版本的 Copilot 几乎可以适用于任何职业。

比方说，面向律师的 Copilot 可以帮助他们基于自然语言查询以前的案例以及最佳实践起草合同、动议、辩护状以及其他法律文件。AI 还可以建议相关判例、法规与传讯，或对现有文件存在的可能错误、不一致或风险进行标记。

面向建筑师的 Copilot 可以帮助他们按照规范、约束和目标对建筑物和结构进行设计、建模和优化。AI 助手还可以生成交互式的可视化效果，并帮助确定项目对环境、社会和经济的影响。

不妨想象这样一个世界。在这个世界里，成千上万个行业里数百万的专

业人士都在使用各种专业版的 Copilot，用更快的速度将生产力、准确性和创造力提高到全新的水平。一个所有行业的专业人士都可以使用通用工具的世界，去利用通过接口编写的每个 app、API 或软件程序的威力，让他们得以用通俗易懂的言语描述他们想要完成的任务。

在反乌托邦的未来愿景里，泛化的技术，或者特指的人工智能，往往会被描述成会导致贫富两极分化愈发严重的力量，大部分人类将被机器剥夺掉权利，被边缘化以及陷入贫困。

不过，在今天我们所看到的世界里，新的人工智能工具正在以前所未有的方式有效地实现了设施与效率的大众化。在此过程中，它们为专业人士赋权，让他们的生产力达到新的水平，并给社会带来好处，其收益甚至可能超越工业革命之所得。不仅如此，人们还会发现自己的工作变得更有吸引力和成就感了，因为他们将会有更多时间专注于自己最具创造力、战略性以及新颖性的事情上。

未来已来。五年之内，所有的主流职业都会有人工智能的增效工具。这些工具可以促进人类在各个职业中的卓越表现——不管是右脑型、左脑型还是任何复合类型。

生成式人工智能与生命科学

AI 又过了一个盛夏。我们目睹了新的研究团体的兴起。这些团队正以前所未有的速度将大型集中式实验室开发的突破性 AI 模型开源出来。虽然这些文本生成图像/视频模型可以提供病毒式流行的消费级产品，抓住我们的想象力，但这些模型最有影响力的应用不太可能是它们的一阶效应。我相信真正有待开发的是人工智能与科学的结合部，尤其是在生命科学领域。

今天的科学方法植根于数据驱动的实验。我们可生成的，用于解释生物系统的数据的分辨率和规模正在不断改进，与此同时，我们还在开发能够对人类语言、自然图像或社交网络图谱建模的 AI 模型架构。这些架构可以直接转化为对蛋白质的语言、细胞图像或化学分子图的建模。这种不可思议的泛化能力现在正在让我们在蛋白质结构预测以及药物分子设

计方面取得突破。AI正在推动新一代技术驱动的生物技术公司向价值数十亿美元的制药行业进军，以更快的速度、更低的成本提供更好的药物。

Ali Madani 是这个领域的创始人之一，他在 Salesforce Research 领导了一项名为 ProGen 的 AI 蛋白质工程登月计划。这项计划的目标是开发专门用于设计全新的人工合成蛋白质的大型语言模型。这些人工合成的蛋白质将可重现甚至优于天然存在的同类蛋白质的特性。他们的团队做出了人工智能生成的蛋白质的第一个 3D 晶体结构。鉴于蛋白质是一切生命的机能驱动，像这样的技术释放出来的可能性是巨大的。

协作式界面

大型语言模型（LLM）是我们见过的最强大的工具之一。我们还在测试这些模型的指令范围在哪里。如何做出巧妙的提示设计已迅速成为极客的热门运动。（很滑稽，“让我们按步骤思考”的提示几乎增强了模型的推理能力。）

但自然语言并非万灵药——在没有手册指导的情况下，我们仍在盲目地发出命令。提示设计没有指南、缺乏一致的抽象，没有明确的地图来引导我们探索模型的“潜在空间”，只有大量的反复试验以及各种妙招。

有些初创企业已经开始给范畴狭窄的用例做用户体验的简化，并且获得了回报。举个例子，Jasper 提供模板化提示来创作营销文案。正如大多数语言模型产品给用户展示“采样步骤”以及“种子”这些不透明概念所说明的那样，我们还处于早期阶段。

每个可以访问互联网的人很快就会在日常任务当中间接地用上大型语言模型。至少，搜索会被彻底颠覆，能按需提供答案和摘要。我们应该还能看到为更好地掌握以及更深入的互动而设计的 LLM 工具。创意人员已经希望生成出来的图像在结构上是可操纵的，而员工则希望能得到可靠的输出。很多人也许希望自己的人工智能助手能接受特定知识的教育。这些是复杂的“头脑的自行车”（编者注：乔布斯语，计算机是头脑的自行车），将释放知识工作者的生产力。

Ilya Sutskever 对“提示”很不屑，觉得这个词只是权宜之计，是因为我们模型存在缺陷。我希望他是对的（因为通常他都是对的），我们的模型会逐渐能理解意图。但是一个根本问题是人类的意图未必总是确定的；人的意思往往是迭代的、探索性的。一旦模型要处理需要这种思维的，更复杂的任务，我的直觉是，对于创造最终用户价值来说，理解 workflow 并实现针对该工作流的更多控制和反馈至关重要。改进用户体验的早期想法包括模板、在不同代际的模型间选择用户界面和添加更多约束的能力、对上下文长度的控制、对链式过程中的居间控制，以及公开模型的“思维过程”。

一些创业者和投资者对围绕着别人的模型来建立商业价值感到绝望，但我们对于如何与 AI 进行交互的了解才刚刚开始。不同领域之间可能会存在差异，研究人员不太可能满足每个用户角色的需求。这些强大模型的唯一界面会一直都是一个简单、静态的文本框吗？我认为不会——这就是产品的机会。

AI 视频创作

生成式人工智能现在风靡一时，而且它的火有充分的理由，因为确实非常令人兴奋。撇开技术实力不谈，常见的商业问题也适用：你能不能开发出这样一种产品，用生成式 AI 解决问题的效果可以比其他方式好 10 倍？随着时间的推移，你能不能建立起牢固的竞争优势？

以我的一家之言，视频创作平台 Synthesia 就是一个很好的例子。通过它可以看出如何在生成 AI 的基础之上做出令人兴奋的业务。通过 Synthesia，用户只需键入几行文本，单击几个按钮，然后，就会生成一个专业的视频，几分钟之内就会有一个人类化身来讲解文本（请注意，它可以支持多达 60 种语言）。

Synthesia 可以用于各种企业用例，尤其是在入职和培训方面的用例。对于许多客户来说，之前要么是提供很少有人能读下去的冗长 PDF，要么就得花费数十万美元通过演员、导演、摄像机以及后期制作来做出专业视频。

通过生成式 AI，Synthesia 可以极大地减少创建商业视频所需的工作量、时间与金钱，这个效率提升也许达到 100 倍之巨，而且让任何人都能做到这一点。此外，它还开发了自己的专有 AI 技术——公司的两位联合创始人都是 AI 方面的教授，并且拥有一直强大的内部研究团队。虽然它也利用了大型语言模型，但与一般初创企业基于 GPT-3 来开发不一样的是，该公司避免了对平台的依赖性，为长期稳固的竞争优势以及赢得品类的领导地位铺平了道路。

自动化代码生成与应用开发

现代机器学习（ML）的进展的步伐似乎一直都很快。深度学习模型第一次赢得最受欢迎的计算机视觉竞赛还只是十年前的事。可是，当 GitHub 在 2021 年底推出其“AI 结对程序员”Copilot 产品时，许多人（甚至像我这样从事 ML 工作的人！）都对现在的深度学习模型居然可以帮高技能的软件开发者自动补全代码感到震惊。在开发者用于来编码的界面里，Copilot 可以就如何完成一行代码给出建议，甚至可以根据该代码应该做什么的自然描述生成多行代码。一些用过第一版 Copilot 的工程师声称，该助手每天可为他们节省数小时的时间，甚至可以替他们编写 40% 的代码。

Copilot 是用 OpenAI Codex 的大型语言模型（LLM）开发出来的，这个模型可以将自然语言翻译成许多流行的编程语言，并利用了数千万个公开的 GitHub 代码存储库的数据进行训练。交代一下，OpenAI 是一家位于旧金山的人工智能研究公司；这家非营利组织成立于 2015 年，并在 2019 年转为营利性组织，然后从微软（微软于 2018 年收购了 GitHub）处筹集了 10 亿美元的资金进行研究。作为回报，微软拿到了对 OpenAI 的部分 LLM 的独家访问权，其中就包括 Codex。

Copilot 最终还是提供了令人信服的证据，证明当前的 ML 功能可以自动化完成越来越多的代码生成与应用开发的工作。新成立的初创企业以及既有公司已经在着手解决产品开发体验的多个方面，包括自动代码审查、代码质量改进、shell 命令自动完成、文档创建，甚至包括前端和网站生成。

Grit 是在这个领域创业的一个早期例子，它承担的是对于任何公司来说都是最可怕的工程任务，也就是所谓的“技术债务”。一旦开发者为了更快推出功能而走上编码的捷径，但却牺牲长掉期可靠性和性能时，这种债务就会累积。Grit 的产品扮演了自动化开发者的角色，可以修复许多常见问题，并根据人类对代码更改的反馈进行改进。通过将静态分析与 LLM 相结合，Grit 的愿景是能够做出可自我维护的软件。

鉴于这项技术有可能彻底改变软件开发，多家投资者已经汇编出相关公司名单，许多其他初创公司也在悄悄地开发相关技术。其中部分初创公司以 Codex API 为基础，希望通过独特的产品体验以及专有的数据飞轮形成差异化。其他的正在从头开始开发自己的模型，或者对开源模型机械能微调。随着这些公司走向成熟，不管最大收益者是人工智能基础设施提供商还是人工智能应用本身，价值主要积聚在什么地方将变得更加清晰。

临床试验的数字孪生

人工智能将改变我们用药物治疗人类疾病的方式。

一想到 AI 和制药时，我们最常想到的应用就是用于药物发现的 AI。（理由很充分：人工智能驱动的药物发现具有巨大潜力。）

但是，还有一个机器学习的用例，尽管它覆盖的范围不广（而且资金投入也较少），但因为有望更快、更有效地将改变数百万患者人生的疗法推向市场，所以也很引人注目。这就是数字孪生在临床试验当中的应用。

当今的临床试验效率低下且成本高昂，这是有据可查的。平均每一种新药需要 10 多年的时间以及 20 亿美元的开支才能推向市场。招募试验参与者是药物通过临床试验的一大障碍。一项试验需要招募成百上千乃至成千上万名志愿者才能满足试验组与对照组的人数需求。这一点已经成为一个重要瓶颈。80% 的临床试验都会因为登记问题而延迟，试验发起人每天因试验延迟的潜在收入损失高达 800 万美元。每年有数百个临床试验因受试患者人数不足而终止；事实上，这是临床试验被终止的首要

原因。

“数字孪生”为这一挑战提供了一种变革性的解决方案。其基本想法很简单：生成式机器学习模型可以替临床试验患者模拟安慰剂的结果。这可以在个体患者层面完成：可以为试验实验组的每个人类试验参与者建立一个数字孪生，然后模拟该个体在对照组中的表现。

至关重要的是，这意味着制药公司需要招募的人类参与者要少得多，因为大部分对照组患者群体都可以被数字孪生取代。这大大加快了临床试验的速度和成本，让改变人生的疗法能够更快地进入市场，并惠及数百万有需要的患者。

总部位于旧金山的人工智能初创公司 Unlearn 就站在这一变革性技术的前沿。Unlearn 目前正在与包括默克集团在内在某些全球最大的制药公司合作，部署该初创公司的数字孪生技术，加速临床试验的进程。今年早些时候，欧洲药品管理局（European Medical Agency）正式批准 Unlearn 的技术可用于临床试验，标志着该技术已经赢得主要监管的认证，可为大规模部署做好准备。

预计几年之后，将数字孪生纳入临床试验，简化治疗药物的上市途径，将会成为制药和生物技术公司的标准做法。

值得注意的是，尽管与流行的文本生成图像模型无关，但用于临床试验的数字孪生是生成式人工智能一个引人注目的代表。生成式机器学习模型如何才能对现实世界产生巨大影响，并创造出数十亿美元的价值呢？为个体患者生成模拟的安慰剂结果就是一个很好的例子。

为 workflow 而来，因个性化而留

随着越来越多的用户与生成式 AI 模型进行互动，我们对 AI 马上可以解决的问题有了更深入的了解：那些我们已经拥有大量训练数据的问题；那些在 99% 的时间内都能获得正确答案会非常有用、但在 1% 的情况下出现错误也不会是灾难性的问题；以及基础模型可以不断吸收人类反馈并随着时间的推移会变得更好的问题。随着 AI 跨越鸿沟，进入主流，直

观的工作流会推动大规模的采用，让那些不太熟悉 AI 的人快速看到其价值所在。

在下一代 AI 初创公司里，专注于工作流设计，并能根据用户反馈对模型做出微调的创始人能够做出最好的产品。

符合这种模式的初创企业有两类：一是 AI 代理，二是 AI 增强型 SaaS。人工智能代理可完成重复性的知识工作——律师、工程师、会计师或者医生所从事工作当中的那些重复性的工作。人工智能增强型 SaaS 将依赖人工智能层从现有工作流中析取出更多的价值——比方说，给采集音频数据的平台添加文字记录和摘要，或增加语言界面来简化 SaaS 应用。在这两种情况下，为了保证输出质量，人工监督依旧存在。用户会给出正面和负面的反馈，然后这些反馈将被用来对模型做出调整。

能在竞争中获胜的创始人会对界面和工作流进行合适的设计，基于当前的提示和自动完成模式进行创新，为用户提供高级控制，并降低认知开销。这些工作流会通过模板或专门的可组合模型来加速常见的用例，同时确保“紧急情况下打破玻璃”（break-glass）选项可用于不常见的边缘情况。用户不必了解模型的工作机制或根据模型改造自己。当用户与产品交互时，通过接受答案而生成的数据会自动反馈回去，驱动个性化与留存的数据飞轮。

这些初创企业会专注于自己的核心竞争力，并将把通用人工智能模型的开发留给研究实验室和开源社区。实际上，后者已经推出了非常强大的模型。我们已经看到了像 Stable Diffusion 这样的文本生成图像模型，像 Whisper 这样的音频转录模型，以及像 GPT-J 和 GPT-Neo 这样的语言模型。初创企业则会利用 AI 研究的最新进展，当新模型可用时更换新模型，并根据专有的用户反馈历史数据做出微调。目前的局限性在于产品设计师的精力集中在让不了解 AI 的消费者能够轻松参与，并迅速从模型获得价值的界面上。这方面的护城河可以是用户与这些模型互动时收集到的综合的工作流与数据，这些可以为将来更强大的模型提供信息。

链式提示以及多步骤自动化

很多人现在都已经体验过提示界面的强大功能，感受过大型语言模型（LLM）可以带来的各种创造性和实用用例，是时候更进一步了。作为用户，提示界面和 LLM 的美妙之处在于它们可以解释意图，并带来行动。到目前为止，我们看到的成果主要是单人体验上的，比如像文本生成图像，文字补全等非常具体或一对一的任务上。接下来，预计将会看到有人会开发将提示链接在一起的基础设施，让我们得以实现多步的操作（通过 LLM 及/或最终通过 API 交互），并释放出巨大的投资回报潜能。

这对消费者用例或者企业用例都会产生意义。有些是成果很容易兑现的简单用例，也许不具备风投的规模（“周二下午 6 点到 8 点左右帮我在布鲁克林南部找一家安静浪漫的亚洲餐厅订个桌，两人”），但有些用例可以为运营、采购、数据分析等各种任务带来实质性的提升。

在消费者方面，我们很可能会把这些用例当作横向拓展的楔子。在企业端，对这些动作加以完善，可以变得更具行业性或上下文相关，这时候难以集成或专有的数据管道就会发挥更大的作用。

把提示和不同类型的模型（未必总是需要）链接在一起，这种想法还可以帮助克服当今主流的 API 驱动模型（如 GPT-3 模型的内存是相当静态的，你可以查询信息，生成摘要，然后根据需要使用结果）存在一些局限性。这方面的用例可以是建立新的搜索引擎，就像最近的 webGPT 所做的事情那样。

总的来说，在 AI 的帮助下，再加上一点人类的独创性，链式提示或串行操作可以带来更先进的工作流、新型的软件产品以及针对常见问题的新接口范例，这件事情的潜力巨大，我们刚刚开始触及它的皮毛。

解决现实世界挑战的工业自动化

我们都看过 AI 模型最新突破，比方说 DALL-E 2、GPT-3 等方面的消息了。由于这些创新，企业正在进行重新想象，部分开发者和设计师担心会丢掉自己的工作。但与此同时，在办公楼外面，在餐馆、建筑工地以及工厂等地方，我们正面临着这辈子见过的最严重的劳动力短缺。

比方说，焊工的平均年龄为 55 岁。每一年，熟练焊工都在以 7% 的速度减少，但对这些焊工的需求却以 4% 的速度在增长。到 2024 年，光是在美国就会出现 400000 个焊工的职位空缺。

这就是像 Path Robotics

这样将熟练的劳动力自动化的公司为何如此重要的原因所在。Path 让公司可以用现成的机器人自动焊接新零件，关键是不需要对这些机器人进行耗时且昂贵的重新编程。如今，从电线杆到液压油箱再到消音器，Path 能够为客户实现各种金属应用的自主焊接。借助该公司的软件，机器人可以通过可视化的方式去学习焊接，而不是学习如何用代码来进行焊接。这样公司就可以通过可视化的 QA 检查逐步提高机器人的焊接性能。未来，Path 的核心技术还可以应用到众多的其他制造任务上。

类似，一旦劳动力短缺与供应链挑战叠加到一起时，情况会变得更加糟糕，这就是目前建筑行业的现状。这种双重痛苦导致房地产或商业地产业主无法按时交付建设项目。Ergeon 正在利用先进的 AI 实现围栏建筑整个施工过程的自动化，包括远程测量、自动化设计、报价等。这项技术让公司的交付速度比一般承包商提高了 10 倍。他们已经建立起全球最大的房屋建筑数据库，而且开放授权，任何人都可以用来建造。

写字楼以外的事情我们讨论得不多，但许多价值 1000 亿美元的企业都会在写字楼以外的地方出现。重新想象办公桌与电脑之外的世界是一个千载难逢的机会，我们对此感到无比兴奋。

本文链接：<https://dqcm.net/zixun/16699522101087.html>