

科技新闻

·深度报道·

AI与生物武器报告引发科学家分歧

Anthropic 称,研究人员可能利用 Claude 从事病毒与毒素方面的不当研究

文/Jocelyn Kaiser

这是一则令人不安的警示,但其意义尚不明朗。2026年9月10日,人工智能(artificial intelligence, AI)开发商 Anthropic 发布了一份报告,指出美国境外的科学家可能试图利用其模型 Claude 开发生物武器。该公司披露了5个案例的部分细节,其中3个涉及病毒,另外2个涉及毒素。

这是顶级AI公司首次公开承认其产品可能被用于生物恐怖主义或生物战,报告一经发布便占据各大媒体头条,并引起部分科学家的警觉,尤其是因为使用者采取了向 Claude 隐瞒自身位置的手段。曾协助拜登政府应对新冠疫情的内科医生兼公共卫生专家 Ashish Jha 在 X 平台上写道:“我们正不知不觉地滑向一场潜在的巨大灾难。在坏事发生之前,我们不会再得到5次警告了。”

但一些病毒学家和生物安全专家则没有那么警觉。他们认为,Anthropic 报告中所描述的研究很可能只是常规生物医学研究的一部分,例如用于开发新药或疫苗。美国约翰斯·霍普金斯大学生物安全专家 Gigi Gronvall 表示,网上的一些反应“过度”了。她曾为 Anthropic 提供咨询,但未参与该报告的撰写。

尽管如此,研究人员仍赞赏 Anthropic 对其发现的公开透明,因为这些发现为迄今大多停留在假设层面的生物安全风险提供了现实证据,并已引发关于监管的讨论。英国伦敦国王学院生物安全专家 Filippa Lentzos 表示:“这开始让我们得以窥见开发者们在真实环境中究竟遇到了什么。在缺乏充分政府监管的情况下,我宁愿他们这样做,而不是什么都不做。”

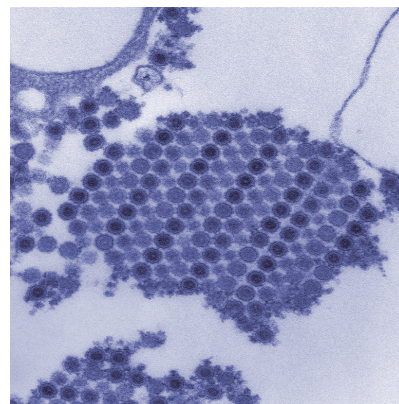
Anthropic 这份报告中还详述了其AI的其他潜在非法或令人不安的用途。发布之际,正值一场关于AI模型安全的更广泛辩论进行之时,甚至连一些顶级公司的员工都认为AI对人类构成生存威胁。在2026年9月12日发表的一篇文章中,Anthropic 首席执行官 Dario Amodei 主张放慢AI的发展速度,并表示禁止利用AI开发生物武器是各国可能达成共识的措施之一。

Anthropic 的报告称,2026年5月,有研究人员利用 Claude 撰写了一份基金申请书,该项目旨在找出基孔肯雅病毒(该病毒可导致发热、严重关节痛等症状,但极少致命)中能够使其更高效传播并逃避免疫系统的突变。研究人员向AI输入的提示词显示,他们计划改造出含有这些突变的病毒,并在动物

身上进行测试。

同样在2026年5月,另一个团队要求 Claude 规划实验,对一种高致病性禽流感病毒株(报告未指明具体毒株)制造突变,以帮助其适应哺乳动物并更易传播。Anthropic 称,该项目相当于“对具有增强大流行潜力的病原体开展研究”,这是一种颇具争议的实验类型,在特朗普政府实施的新规下,美国已不再允许开展此类实验。在第3个案例中,研究人员利用该AI撰写了一份基金申请书,旨在找出使正痘病毒家族(该家族包括猴痘病毒和已被根除的天花病毒)中某个未指明成员在小鼠体内毒力减弱的突变。

美国斯克里普斯研究所病毒学家 Kristian Andersen 表示,这些工作“听起来吓人,但实际上只是基础生物学研究,完全可以在高等级防护实验室中安全开展”。他说,这些研究可能只是给病原体引入了自然界中已存在的突变,而非科学家设计出的新的、风险更高的基因改变。他补充道,这类工作对于理解病毒造成危害或变得更加



基孔肯雅病毒颗粒的彩色透射电镜图。

Anthropic 报告所举案例之一涉及该病毒研究;图中病毒并非报告所述项目的样本

(图片来源:美国疾控中心)



一份关于部分病原体研究者如何使用人工智能的报告引发了担忧

(图片来源:《Science》)

有害的机制至关重要。美国乔治·梅森大学生物安全专家 Gregory Koblentz 补充说:“Anthropic 所讨论的所有研究,针对的都是在世界不同地区流行的、当前构成公共卫生威胁的病原体。”

第4组研究人员让 Claude 汇编一份毒液毒素图谱,作为药物开发项目的一部分;最后一组则利用 Claude 协助通过计算手段重新设计一组毒素,其中包括一种细菌毒素和一种来自被视为重大疾病威胁的病毒的蛋白质。报告称,这2个项目研究的都是“可同时开发成新型疗法或毒剂的化合物”。但 Andersen 指出,那份毒液图谱所依据的是公开可获取的知识。

在全部5个案例中,科学家都采取了诸如使用中间服务器、VPN 或伪装账号等手段,使自己看起来像是从美国或其他未限制使用 Claude 的国家使用该模型。对此,大多数科学家并不感到意外。在 Koblentz 看来,这些手段听起来“就像青少年绕过手机和电脑上的家长控制,以及人们夏天为了看世界杯而做的事”。

据 Anthropic 称,其他潜在的危险信号包括:基孔肯雅病毒研究

“拟在军事研究所开展”,而正痘病毒研究则在一家“国家关联”实验室进行。然而,Koblentz 指出,许多研究人员就在政府实验室工作,而美军本身也“开展并资助一系列对部队构成潜在风险的传染病研究”。

Gronvall 指出,就目前而言,AI 最大的隐忧在于,一个心怀恶意的新手科学家会利用这些模型迅速获取制造危险病原体所需的知识。但在她看来,被 Claude 标记的那些研究人员“只是在利用最新工具辅助自己的工作”,与任何地方的资深科学家并无不同。

Lentzos 和 Koblentz 都认为,政府有必要强制企业限制谁能够为生物研究目的访问这些模型。但 Andersen 担心,打击行动也会妨碍正当研究者。他说,即便是现在,当他用 Claude 比较2株埃博拉病毒的基因组时,查询也会被拦截。美国亚利桑那州立大学生物安全专家 David Gillum 也有同样的担忧:“这些公司本就手握重权,而现在它们又要做出这些理论上可能影响科学知识的决定。”

Gronvall 建议,与其由政府施加严格监管,不如让 AI 公司继续分享案例、改进防护措施。她说:“我认为这个领域会不断演变和成长,持续关注它很重要,但我不认为这些模型在生物武器的恶意用途上像人们以为的那样有用。”

(译自《Science》,2026,393(6817))

准金属催化能否减少贵金属依赖

文/马星星¹,宋秋玲²

在现代化学合成的精密世界里,过渡金属(如钯、铂、铑)一直发挥着重要作用。它们如同身怀绝技的“分子工匠”,能高效地搭建起从治病良药到高科技材料的各类复杂分子。2010年,诺贝尔化学奖授予了在“钯催化交叉偶联反应”领域做出杰出贡献的3位科学家,足见这项技术对人类生活的深刻影响。然而,这份依赖正变得昂贵而脆弱。

2024年,南非约占全球铂矿山产量的70%,这种地理上的高度集中意味着贸易摩擦或地缘冲突都可能直接冲击关键催化材料的稳定供应。与此同时,贵金属的开采与提纯伴随着高昂的环境代价。正是在多重压力叠加之下,合成化学家开始追问一个根本问题:能否用一类更普遍、更廉价、更绿色的元素——其地壳储量是铂系贵金属的数千甚至数万倍——去取代使用了数十年的传统“贵族”工具?

2025年7月,中国化学会推

1. 福州大学化学学院,有机合成与功能福建省高校重点实验室,副教授
2. 福州大学化学学院,教授