

# 让AlphaFold的力量为全世界所用

德米斯·哈萨比斯

2020年12月，我们向世界介绍 AlphaFold 2 及其在预测蛋白质结构方面的超凡能力。当时，评估者认为这是解决困扰科学界 50 年的重大问题——解析蛋白质三维结构——的最佳工具。2021 年 7 月，我们在《自然》(Nature) 刊载了两篇论文，发布了 AlphaFold 的开源代码，解释这个高度创新的体系是怎么来的；同时，我们还分享包括人体中存在的每一种蛋白质以及常用于科学研究的 20 种生物体的蛋白质的高质量预测。

科学家试图寻找各类疾病的治疗方法，也希望解决人类面临的其他重大问题，包括抗生素耐药性、微塑料污染和气候变化。如果能有对蛋白质结构的更深入理解，研究工作无疑会因此获益。蛋白质就像小巧精致的生物机器，而机器的结构决定了它的功能，因此揭示蛋白质的结构可以帮助我们理解蛋白质的功能。

AI 是支持研究人员工作的强大工具，而我们相信 AlphaFold 是迄今为止人工智能领域提供的最强科学辅助。通过实验确定蛋白质结构是一项耗时且艰苦的工作，但 AlphaFold 却证明人工智能可以在几分钟内以原子级精度预测蛋白质的形状。借助 AlphaFold 收获对蛋白质的更深见解，为我们在生物学和医学领域取得更多新突破奠定了基础。

此次突破得益于 AlphaFold 团队五年来不知疲倦的探索和惊人的创造力——他们创造性地解决了大量极具难度的问题，得到了许多新的复杂算法，这些都是攻克终极难关所必需的；还得益于过去几个月与我们在一起的欧洲生物信息研究所 (EMBL-EBI) 的密切合作，我们现在才能与世界共享这宝藏。

不久前我们还与 EMBL-EBI 合作推出了 AlphaFold 蛋白质结构数据库，提供迄今为止最完整且准确的人类蛋白质组的结构图——数量是此前人类已积累的两倍还多。除了人类蛋白质组（人类基因组表达的蛋白质共约 2 万种）之外，我们还提供其他 20 种生物的蛋白质组，总计超过 35 万种蛋白质结构的数据访问权限。从大肠杆菌到酵母菌，从果蝇到小鼠，这 20 种生物极具科研意义，是无数研究论文的主题和众多科学进展的主角。我们计划之后将数据库的覆盖范围扩大至几乎所有已知蛋白质——超过 1 亿个结构，涵盖 UniProt 蛋白质数据库的大部分内容。这是一本名副其实的世界蛋白质年鉴。未来我们将继续投入对 AlphaFold 的改进，系统和数据库也会定期更新。

最令人激动的是，这本全新的蛋白质年鉴能在世界各地的科学家手中发挥神奇力量，推动相关研究，提升人类对蛋白质的理解。通过早期合作，我们已经看到了研究人员在自己的工作中利用好 AlphaFold 的可能。而这些都是结构生物信息学革命的开始。随着 AlphaFold 的发展，现有大量数据有望转变为未来的科学进步。

AlphaFold 的成功建立在几代科学家的努力之上，从蛋白质成像和晶体学领域的早期先驱，到蛋白质结构预测专家和结构生物学家，众多研究者很早就开始投入到相关探索中，并留下大量实验结果，为此项创新奠定了基础。

我们认为人工智能可以大幅加速许多科学领域的突破，进而推动人类进步。我们相信 AI 有望彻底改变 21 世纪的科研方式，也热切期待着 AlphaFold 帮助科学界解锁下一个神奇。



德米斯·哈萨比斯  
(Demis Hassabis)  
Deep Mind首席执行官

——刘迪一译自 [deepmind.com](https://deepmind.com)