R 创新谈

AI时代，软件产业靠什么赢得竞争

成 蹊



“人工智能+软件”并没有改变软件开发的底层逻辑。人工智能可以为软件开发赋能，却不能替代科学家和工程师的思维

今年9月，工业和信息化部印发《“人工智能+软件”专项行动实施方案》（以下简称《实施方案》），提出优化智能编程工具链条，加快软件产品智能化升级等举措，加快推进人工智能技术在软件和信息技术服务业融合应用。

AI（人工智能）技术快速发展，正深刻改变软件的生产方式、产品形态和服务模式，也给软件产业出了一道新考题：当AI越来越会写代码，编程门槛不断降低，软件产业靠什么赢得竞争？

过去，一个想法变成软件产品，要经历需求分析、程序编写、测试修改等多道工序，耗费大量人力和时间。如今，把需求告诉AI，它就能生成代码、发现错误，甚至协调完成部分开发任务。有媒体报道，重庆一家软件开发企业让3个AI智能体分别承担产品经理、项目经理和技术经理的工作，工程师成了负责调度和把关的“监工”。在代码编写环节，AI工具覆盖率大幅提升，部分企业甚至已实现全员覆盖。

产品形态上，过去的软件大多被动执行指令，现在一些智能体软件却能理解意图、自主完成任务；服务模式则从交付一套固定功能的程序，转变为持续提供智能应用。数据显示，中国企业级AI智能体市场2025年已达212亿元，预计2026年将增至449亿元，2029年有望突破3320亿元，年复合增长率高达107%。《实施方案》从推进软件生产变革、加快软件产品智能化升级、培育智能体软件新业态等方面作出部署，正是着眼于这一系列新变化、新挑战。

人工智能技术在降低软件生产门槛的同时，也让行业知识、专业经验的价值愈发凸显。人工智能可以帮助开发者更快写出程序，读懂行业和需求的功夫却非一朝一夕所能练成。

以制造业领域的排产软件为例，不同工序如何衔接、设备在哪种情况下需要停机、怎样兼顾交付进度与能耗，都需要对生产场景有深入了解。今年前8个月，我国软件业务收入达10.28万亿元；我国制造业增加值连续16年位居世界第一。广阔的软件应用市场为新软件提供了丰富的需求，完整的制造业体系则积累了大量生产数据、工程方法和行业知识。《实施方案》提出，支持有条件的工业企业联合软件企业，将工业经验、工艺参数等知识封装为可复用的模型库和工业组件，研发行业智能化工业软件。借助人工智能技术将专业经验融入算法和软件产品，让软件真正懂得生产流程、适应复杂工况，我国实体产业的优势就有望进一步转化为软件产业的优势。

抓住新机遇，同样需要新技术与真功夫相互成就。

软件开发提速了，质量和可靠性的要求却不会降低。一段人工智能生成的代码放到复杂工程中，能否长时间稳定运行、能否应对异常情况、能否保障数据安全，都需要经受反复检验。

应当清醒地看到，“人工智能+软件”并没有改变软件开发的底层逻辑，物理原理和数学方法仍然是开发者需要持续深耕的问题，具体工程问题的求解内核需要长期研究积累，极高的精度要求需要非常规的代码设计以及海量工业场景验证。人工智能可以为软件开发赋能，却不能替代科学家和工程师的思维。

既要用好人工智能，加快软件智能化升级脚步，也要持续加强核心算法、基础架构等技术攻关。推动软件企业与行业用户深度合作，让产品到真实场景中反复打磨；加强开源社区建设，促进技术共享和协同创新。尤其要重视质量与安全，把代码审查、风险检测贯穿开发全过程。

梳理软件产业的历史脉络，从个人电脑到移动互联网，每次技术变革都催生了新产品、新企业，也改写了产业竞争格局。如今，人工智能又打开了一扇新的大门。将庞大的产业规模转化为过硬的技术实力，将丰富的应用场景转化为软件产品，中国软件产业有望在新一轮科技革命和产业变革中赢得更大发展空间。

R 深度观察·挑大梁、当主角，我们有信心

习近平总书记强调，青年人才是基础研究的生力军，很多重大成果都是年轻人做出来的。

“十五五”开局之年，我国基础研究成果涌现，这背后离不开广大青年人才创新进取、勇攀高峰，为推进中国式现代化建设贡献青春力量。

青年人才如何在基础研究一线挑大梁、当主角，如何创造条件让他们充分施展才华？从今天开始，科技版推出“挑大梁、当主角，我们有信心”系列报道，从不同角度进行深入探讨。

——编 者

青年联合团队打造百亿参数人类基因组基础模型

架AI之桥 解生命密码

本报记者 李君强

从寻找罕见病致病原因，到筛选肿瘤潜在药物靶点，AI正为生命科学研究打开新的窗口。

2025年10月，之江实验室与华大生命科学研究院联合研发的全球首个百亿参数人类基因组基础模型Genos正式发布。一年来，团队已与近40家医院、高校等机构开展合作，尝试借助AI解读遗传变异，为疾病诊断研究、药物靶点发现等寻找新线索。

这项前沿探索的背后是一支年轻的科研队伍。联合研发团队约130人，平均年龄只有30岁。来自计算机、生命科学等不同专业背景的他们，尝试让AI从海量基因组数据中学习，寻找人类尚未认识的生命密码。

探索数十亿个碱基对里的未知

人类基因组包含数十亿个碱基对，记录着人的遗传信息。2003年，人类基因组计划宣布完成，科学家首次拥有了人类基因组高质量参考序列。此后，随着技术不断进步，海量基因组数据加速积累。然而，如何理解基因组序列中的遗传信息，仍是生命科学研究面临的重大挑战。

华大生命科学研究院领域首席科学家刘石平表示，到目前为止，科学家对人类基因组的理解还有许多空白。比如，基因组的某个位置发生变化，会不会影响身体健康？不同基因之间如何相互影响？这些奥秘都还有待发现。“目前已知的蛋白质编码基因仅有2万多个，即使对于这些已知基因，其具体作用以及与疾病的关系也远未研究透彻。”刘石平说。

AI快速发展带来新的思路：既然AI模型能从海量文字中学习语言规律，是否能从大量基因组序列中学习，让原始科学数据直接进入基础模型，帮助科学家找到过去难以发现的联系？

彼时，之江实验室正探索让科学数据进入基础模型的途径，生命科学是其中一个重点领域；华大生命科学研究院则在基因组研究中积累了数据与经验，双方一拍即合。刘石平记得，双方首次碰面进行了约1个小时的交流后，就决定组建联合攻关团队，训练啃下人类基因组这块“硬骨头”的基础模型，寻找与疾病、健康有关的科学线索。

1995年出生的之江实验室生命科学计算研究中心青年科研骨干刘峻琛，是最早加入的团队成员之一。2023年，他从北京邮电大学电子科学与技术专业博士毕业后，进入之江实验室工作。原本并不熟悉生物学的刘峻琛，在参与Genos项目时，有着最朴素的志向，“在了解了许多基因相关的罕见病案例后，我想，如果能运用自己在计算机方面的知识去帮助解决生物学上的难题，就有可能让一些家庭少受痛苦。”刘峻琛说。

与刘峻琛并肩作战的华大生命科学研究院青年科研骨干邓凯文，同样出生于1995年。他在校期间经历了大语言模型快速发展的几年，看到AI处理人类语言的能力不断进步，他琢磨：既然人类基因组序列和自然语言同样是一维信息，能不能让AI也学会“解读”人类基因组序列？“当时在实验室，我非常关注这个事情。”邓凯文表示，随着国外团队陆续开展相关探索，他越发想自己尝试，但受限于计算资源，难以组织大规模的基础模型训练。

2025年7月毕业后，邓凯文进入华大生命科学研究院工作。听说Genos项目正在之江实验室开展，入职才不到一周的他，赶紧来到浙江杭州参与初版模型研发。“我一直想尝试自己训练一个基因组基础模型，看看能有什么效果。”邓凯文说。

刘峻琛带着AI技术走入生命科学，邓凯文带着生物学问题寻找新的方法。就这样，一群来自计算机、生命科学等不同领域的青年科研人员来到同一个课题前，尝试让AI读懂人类的海量遗传序列信息。

让两种学科“语言”说到一起

联合攻关的首道难题，是如何用不同学科语言进行交流。

刘峻琛还记得一次关于“序列长度”的讨论。面对同一个词，计算机专业背景的同事说的是模型处理的“词元”长度，生命科学专业背景的同事想的却是DNA（脱氧核糖核酸）的碱基长度。讨论后大家才发现，原来说的根本不是一回事。

类似的“小插曲”在这支跨学科团队里并不少见，计算机研究人员熟悉模型和算法，但需要学习基因组中的生命规律；生命科学研究人员了解基因的功能和作用，却需要学习如何训练大模型。联合攻关的头一个月，来自不同专业的年轻人坐在一起补基础知识，之江实验室生命科学计算研究中心副主任马静负责带领团队评测不同方案的效果。

跨过了沟通这道关，更大的科研难题接踵而来。此前，一些基因组模型主要利用参考基因组训练，有限的参考数据难以充分呈现不同人群之间的遗传差异。想让AI真正理解人类基因组，就要让它接触更加多样化的遗传信息。

正在团队琢磨数据来源时，人类泛基因组参考联盟开放了一批高质量基因组数据，涵盖不同地区和人群。“这是很关键的节点，这批开放的科学数据让团队坚定了研究方向，把人类遗传多样性作为重要考虑，让模型学习不同个体的基因组信息，为研究遗传变异与疾病的关联提供更可靠的基础。”马静回忆。

有了数据，怎么把模型训练得更好仍是挑战。人类基因组序列看起来只有四种字母代表的碱基，但内在关系却极其复杂。一个碱基变化可能对生命活动产生重要影响，相隔遥远的两段序列也可能共同影响某个基因的表达。如何让模型既能分析较长的序列，又不遗漏细微变化，成了研发的一道难题。

当时，邓凯文参与了模型测评工作，从生物学的视角比较不同版本的模型在具体科研任务中的表现。团队发现，训练数据覆盖更多个体基因组，也就更有助于提高模型在人类相关变异识别上的性能，为团队后续扩大数据规模提供了依据。

在之江实验室，像这样的科研讨论随时都可能发生。上午9点，10名团队成员围在白板前，讨论基因组基础模型新版本的优化路径，经过几轮交流，前一晚冒出的程序问题有了应对办法。



“有问题能随时当面讨论，非常方便。”刘峻琛说，他与邓凯文在同一间办公室工作，不同专业背景的年轻人逐渐学会从对方角度思考问题。

这种紧密的合作，也得益于科研组织方式的创新。之江实验室生命科学计算研究中心副主任陈高祥介绍，实验室的“种子班”机制让外部机构和实验室相关人员可以集中办公，在物理空间上实现“零距离”交流，遇到问题随时讨论、共同攻关。

经过数月集中攻关，团队成功攻克上述难题，Genos模型首次汇聚了全球范围内636个“端粒到端粒”级别的人类基因组作为训练数据，更全面地代表人类遗传多样性；同时，还通过混合专家架构，实现了百万碱基超长上下文序列与单碱基的精准建模，如同阅读一部鸿篇巨著，既能一字不差地记住细节，又能洞悉全局的宏大叙事。

从模型出发，沿着新的科学问题继续追问

Genos初版发布后，团队仍把大部分精力放在模型升级上。刘石平坦言，目前的模型对人类基因组的理解还不够深入，距离发现更多可信的生命规律，还有很长的路要走。邓凯文和刘峻琛如今各自负责研究新模型中的一个模块，两个人的工作恰好对应着基因组中两类难题。

邓凯文琢磨的是“空间关系”。比如，电脑屏幕上的DNA是一串依次排列的碱基，但在人体细胞中，DNA却以复杂的三维结构存在，序列上相隔很远的两个区域，在空间中可能彼此靠近，共同影响某个基因的活动。怎样让模型从一维序列中识别这些复杂的联系？他尝试调整模型的关注方式，让模型能够学习远距离区域之间的联系。

刘峻琛则把研究重点放在模型的“记忆”上。现有模型还无法一次处理完一个人的完整基因组，他希望先挑选出最有代表性的信息，让模型在分析局部DNA序列时，也能结合个体的整体遗传特征作出判断。

两人的工作时常交汇，一个新的研究思路能否在计算上实现、是否具有生物学意义，需要他们共同寻找答案。

AI进入基础研究改变了传统科研范式，如何从众多设想中找到真正值得研究的科学问题成了新的考验。“现在AI让我们尝试新想法的门槛降低了，但哪个想法能真正促进科学发现，需要培养这种判断的‘审美’。”邓凯文说。

刘峻琛对此深有感触。AI让科研节奏越来越快，新论文新成果不断涌现。他有时担心，自己的新想法如果没有尽快做出来，会不会很快被其他团队抢先实现。但他也时刻提醒自己，不能丢失了科研本该有的严谨。“既要用AI提高效率，又不能因为图方便，在研究内容上出现漏洞。”刘峻琛说。

马静介绍，目前，团队进一步研发了OneGenome模型，将基因组序列与生物医学知识融合，让更多人便捷地使用AI开展生命科学研究。

从提出设想、训练模型，到追问新的科学问题，团队年轻人将AI与基础研究深度融合，不断拓展自己的科研视野。人类基因组数十亿个碱基对中，还有太多未知。他们仍在坚持探索，利用AI读出更多原始信息，让科学研究离生命奥秘更近一步。



Genos模型成为全球首个百亿参数人类基因组基础模型

首次汇聚了全球范围内636个“端粒到端粒”级别的人类基因组作为训练数据

Genos模型已开源1.2B、10B两种参数规模的3个版本

本版责编：刘诗瑶 版式设计：张芳曼 沈亦伶

图①：青年联合团队正在展示模型的相关应用场景。

图②：青年联合团队正在讨论模型的下一步优化方向。

以上图片均为**周志豪**摄

R “咖”说科技

以科技创新筑牢“大国棉仓”

高 雷

金秋时节，我国棉花迎来收获季。棉花既是纺织工业的核心原料，也是油料和蛋白饲料的重要来源。筑牢“大国棉仓”，关系棉农生活和就业，还与产业链供应链安全稳定息息相关。

1983年至今，我国稳居全球第一产棉大国。国际棉花咨询委员会2026年9月统计数据显示，2025/2026年度，中国棉花单产高达2487.6千克/公顷，全球平均单产895.8千克/公顷，中国棉花单产是全球平均单产的2.8倍。我们以不足全球1/10的棉田，产出了全球1/4以上的棉花。

产量上去了，质量还要跟上来。“十四五”时期，我国棉花主产区高品质原棉占比从2016年的18.1%提升至2024年的58.2%。当前，高端纺纱、精密织造对纤维长度、强度、一致性等方面的要求日益严苛。我国是全球最大的纺织品服装出口国，与纺织工业向高端化、品牌化升级的需求相比，我国高品质棉花供给仍存在缺口。实现高品质棉花有效供给，是满足纺织工业向上发展的迫切需要，更是我国从棉花大国迈向棉花强国的必由之路。

科技创新，是回答这一时代课题的重要支撑。近年来，全国农业科研力量协同攻关，攥紧“种子芯片”，用好“栽培良法”，装上“数智大脑”，构建起从种质创新到栽培变革、从智能生产到分子设计的全链条科技支撑体系。

在品种创新上，科研人员利用基因编辑技术，实现了纤维长度、强度等单一性状的精准改良，育种方式正从“杂交选育”向“基因组精准设计”转型。“十四五”时期，全国初次审定棉花品种576个，较“十三五”时期增长35.5%，其中高品质品种144个。我国棉花品种遗传品质已迈入“双三零”时代，即纤维长度达到30—34毫米，断裂比强度达到30—32厘牛/特克斯，一批高品质品种为优质棉生产提供了坚实种源支撑。

在栽培变革上，“密矮早膜”等栽培模式与绿洲生态条件深度融合，宽膜覆盖、窄行密植、全程化控、适期早播为早熟高产高质提供技术保障。目前，全国棉田机械化采收率超过90%，一个家庭农场可经营万亩棉田，这在过去难以想象。

在数智大脑方面，今年8月，我国首款棉花知识大语言模型“棉芯1.0”正式发布，打通科研与生产两端，成为科学植棉的智能助手。棉田天地空一体化监测体系，人工智能决策系统，使水肥用工大幅削减，支撑“一人管护千亩棉田”的规模化生产模式。这些成果，是中国农业科学院棉花研究所、华中农业大学等全国数十家科研院校集智攻关的结晶。

国际科技合作是大趋势，乘着高质量共建“一带一路”东风，我国棉花科技进一步迈向广袤的中亚大地，优质棉种、节水栽培、智能采收等成套技术已在中亚多国落地推广。今年9月，中国—乌兹别克斯坦棉花科技创新示范基地正式揭牌，搭建起常态化跨国科研攻关、技术示范、人才培育平台。



“十五五”规划纲要提出，提升农业综合生产能力和质量效益。站在新起点上，唯有持续加强科技创新，才能不断夯实棉花产业的安全根基，助力农业强国建设。

（**作者为中国农业科学院棉花研究所所长**，本报记者**喻思南**采访整理）