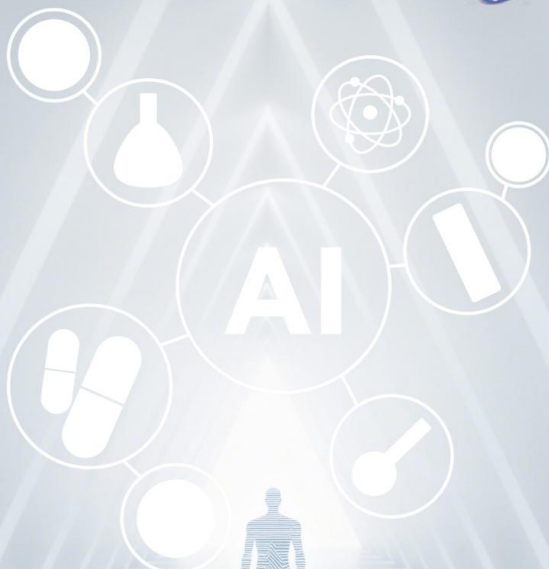


2025 中国AI4LS 行业发展 | 蓝皮书



■ 摘要

AI for Life Sciences (AI4LS) 正加速构建以数据为核心生产要素的新型科研基础设施。该体系涵盖高通量数据采集、智能算法建模、自动化实验验证与知识图谱构建等关键技术模块，形成创新的研发生态。当前，AI4LS已在药物研发、研发实验室、基因组学、合成生物学等多个前沿领域展现出赋能效应。可以预见，AI4LS将持续引领生命科学向更加高效、精准、可预测的方向演进，成为推动生物经济高质量发展的关键引擎。

沙利文谨此发布《2025中国AI4LS行业发展蓝皮书》（以下简称“蓝皮书”），全面解析了AI4LS从科学研究范式演变、多维驱动体系到场景应用落地的全景生态，并系统梳理了技术矩阵、商业模式及未来趋势。通过对高潜力赛道与核心增长动力的深度挖掘，本蓝皮书旨在为相关企业与投资者提供具备前瞻性的市场洞察与战略决策参考。

■ 从工具到范式，AI重塑科研

AI4LS正推动生命科学从“工具辅助”迈向“范式重构”，深度融合机器智能与科学原理，突破传统研发瓶颈。它通过数据与机理的双轮驱动，在蛋白质预测、分子设计等领域实现突破，标志着从“模仿”向“创造”的演进。

■ 产业需求、政策、技术与资本协同驱动

AI4LS发展受产业需求、政策引导、技术矩阵与资本赋能四重驱动：需求端破局生命科学产业所面临的成本、效率与数据瓶颈，国家政策从量到质演进，“数据-算法-算力-知识-团队”五大技术支柱构建闭环科研生态，资本赋能技术壁垒高、商业化路径清晰企业，推动行业健康成熟。

■ 场景应用与挑战：多场景落地，挑战仍存

AI已广泛渗透至多样化的应用场景并发挥效应，加速药物研发全程、驱动研发实验室实现智能化与自动化升级、赋能基因组学迈向精准医疗、优化合成生物制造全链条。然而，数据质量、算法透明性、伦理法规等系统性挑战亟待解决，需构建针对性策略以实现高质量发展。

■ 中国AI4LS发展趋势：逐步从技术追随者向具备全球竞争力的体系构建者转变

从产业合作来看，中国AI4LS企业正通过“验证+协作”、“标准+接口”、“整合+落地”的深度产业合作路径，加速技术落地转化与商业模式构建。与此同时，AI4S技术的通用性正推动跨领域融合、多场景复制，众多平台型企业加速崛起，构建起贯通医药、材料、能源等产业的多元化创新生态。未来，中国AI4LS企业在科研投入、产业基础和创新能力的三重驱动下，正逐步建立起支撑管线开发、自主创新和全球竞争的系统性优势，孕育出面向国际舞台的行业龙头企业。

■ 目录

第一章 AI4LS的概览

1.1 科学研究范式演变	-----	06
1.2 AI4S定义解析	-----	07
1.3 AI4S应用领域	-----	08
1.4 AI4LS在生命科学领域的发展历程	-----	10

第二章 AI4LS多维驱动体系

2.1 产业转型升级需求	-----	12
2.1.1 湿实验成本及重复性人工操作依赖等问题阻碍科学突破	-----	12
2.1.2 线性流程与复杂的临床招募及设计制约行业效率	-----	13
2.1.3 数据规模、高质量数据资源匮乏及数据孤岛形成数据瓶颈	-----	14
2.1.4 伦理压力影响实验进展	-----	17
2.2 顶层设计与政策布局	-----	18
2.3 技术矩阵的搭建	-----	22
2.3.1 数据资源——生命科学领域的“基础设施”	-----	22
2.3.2 算法平台——从海量数据中萃取科学发现	-----	24
2.3.3 算力平台——规模化、高速的计算能力驱动算法落地	-----	25
2.3.4 领域知识嵌入——确保AI模型的输出符合科学规律	-----	26
2.3.5 复合型团队——支撑系统运行与交叉创新的关键	-----	27
2.4 AI4LS赛道的资本共识	-----	28
2.4.1 中国AI4LS投融资情况	-----	28
2.4.2 AI4LS重点应用领域企业投融资概览	-----	29

第三章 场景应用解析和挑战应对策略

3.1 AI应用场景：药物研发	-----	31
-----------------	-------	----

■ 目录

3.1.1 靶点的发现与验证	-----	32
3.1.2 苗头化合物发现	-----	34
3.1.3 先导化合物优化	-----	35
3.1.4 成药性优化	-----	36
3.1.5 临床前验证	-----	37
3.1.6 AI技术企业与药物管线	-----	38
3.1.7 药物研发场景的代表性AI企业	-----	41
3.2 AI应用场景：基因组学与个性化医疗	-----	43
3.3 AI应用场景：合成生物学与生物制造	-----	46
3.4 “集大成者”的智慧研发实验室	-----	48
3.5 AI4LS面临的核心挑战与策略应对	-----	52
 第四章 中国AI4LS行业发展趋势一览		
4.1 产业合作推动AI4LS商业化落地	-----	55
4.2 跨领域协同带动平台型企业崛起	-----	57
4.3 科研、产业、创新三重驱动下的孕育环境	-----	58
 第五章 部分中国AI4LS领域的企业介绍		
5.1 镁伽科技	-----	60
5.2 望石智慧	-----	62
5.3 予路乾行	-----	64
5.4 剂泰科技	-----	66

第一章节

AI4LS的概览

从“工具辅助”至“范式重构”，
AI4LS通过机器智能与科学原理的深
度融合，重塑生命科学研究规则

1.1 科学研究范式演变

图灵奖得主Jim Gray提出的四范式理论为理解科研模式演变提供了经典框架，而当前科学家们正在提出的第五范式——以AI为核心的智能化科研AI for Science (AI4S)

科研四范式与AI工具的碰撞

图灵奖得主Jim Gray提出，将科学研究分为四类范式，揭示了人类认知世界的四次方法论革命。依次为几千年前的经验科学，主要通过感官或简单工具记录自然现象，依赖实验观察与经验总结；几百年前的理论科学，使用数学模型或归纳法构建理论，强调逻辑演绎与普适性规律；几十年前的计算科学，主要利用计算机模拟复杂的现象，解决解析方法难处理的非线性问题；今天是数据探索，统一于理论、实验和模拟，数据依靠信息设备收集或模拟产生，依靠软件处理，用计算机进行存储，使用专用的数据管理和统计软件进行分析。从经验观察到数据探索，每一范式均建立在前一范式基础上，同时突破其局限：经验科学受限于主观观察，理论科学突破经验局限但面临验证难题，计算科学弥补理论验证的不足，数据探索则通过海量数据揭示传统方法难以捕捉的复杂规律。

随着人工智能在研究和生产上的渗透，科研流程也在发生变化。AI不断刷新数据的获取和处理速度，通过数据、算力、算法的深度耦合，更快、更准确地理解复杂问题环境，其与人类类似的学习能力大幅提高了科学研究的效率，给现有的科学研究方法带来巨大变革。在AI的加持下，当前的科学研究正向第五范式——以AI为核心的智能化科研方向进化，体现了人类现代科学认证体系的深度、广度、方式和效率的演进。

图：科研范式的演变流程

科研范式	概念	范例	优势	局限性
 经验科学	指侧重于经验事实的描述和明确具体的实用性的科学，一般较少抽象的理论概括性。	伽利略的物理学、动力	直接来源于自然观察，能真实反映客观现象，具有高度的实证基础。	依赖大量观察和实验，可能带有盲目性；缺乏系统性理论框架，难以解释复杂现象。
 理论科学	指偏重理论总结和理性概括，强调较高普遍的理论认识而非直接实用意义的科学。	爱因斯坦的相对论、经济学的博弈论	通过数学建模揭示本质机制，具有高度的抽象性与普适性。	过度依赖理论假设，可能脱离实际验证；难以处理非线性和高度复杂的系统问题。
 计算科学	指与数据模型构建、定量分析方法以及利用计算机来分析和解决科学问题相关的研究领域。	石油勘探、天气预报、亚原子粒子行为预测	可模拟复杂系统中不可实验或难以解析的问题，提升研究的精度与广度。	依赖高性能计算资源；模型简化可能导致结果偏差。
 数据密集型科学	以大数据为核心，通过数据挖掘和统计分析探索未知规律，从假设驱动转向数据驱动，整合理论、实验与模拟。	制造设备预测性维护、人类全基因组测序	利用大规模数据发现潜在规律，突破传统理论框架限制，驱动新发现。	数据质量参差不齐；隐私与伦理问题突出。
 AI驱动的智能化学	AI通过自主学习与自动化流程推动科学发现，实现从假设生成到实验验证的全流程智能化。	AlphaFold预测蛋白质结构，智慧实验室	融合人类知识与机器智能，具备从问题定义到解法生成的闭环能力，加速科学创新效率和深度。	结果质量受AI工具的影响；资源门槛较高。

来源：文献检索，沙利文分析

1.2 AI4S定义解析

AI技术具有工具属性，服务于人类已有知识框架下的效率提升。而AI4S是从工具到范式的质变，利用机器智能加速科学进程、提高研究效率并推动成果转化，从而拓展人类认知边界

AI4S范式的核心定义

AI for Science所代表的下一代科学范式是指融合机器学习、深度学习等人工智能技术，分析处理多维度、多模态、多场景下的模拟和真实数据。不同于传统计算机，其作用仅能代替人类完成复杂繁琐的计算。AI更加追求“智能”，希望计算机能够模拟人类思考，从而进行学习、推理与决策。AI4S将四个发展阶段的科学研究方法有机结合，使用已知的科学规律进行建模，同时又挖掘海量数据的规律，在计算机的强大算力的加持下，进行科学问题的研究。

值得强调的是，AI4S范式被进一步沉淀为可落地的“智能化科研工作流”，可以打破既往“提出问题-验证假设-发布成果”的线性链条，而是一个非线性、动态演化的系统。环节之间相互作用和反馈：选题方向的变化可能引发对数据结构或算力平台的全新需求；基础设施的升级又可能激发更复杂问题的提出；模型推理与实验验证过程中的反馈信息，能够直接回溯优化前端的数据处理、建模策略甚至问题定义本身。此外，该系统具备高度适应性，能够根据科研进程、技术进步或外部条件的变化灵活调整路径，实现真正意义上的自我迭代和智能协同。

图：AI4S的发展目标定位



AI4S范式的技术路径

自文艺复兴以来，科学研究的演进始终沿着数据驱动与模型驱动两条并行的范式展开。在统计学方法与机器学习技术的推动下，数据驱动的范式演变为解决复杂问题的利器，尤其在缺乏明确理论框架的场景中。然而，这种经验主义方法论固有的“黑箱”特性，导致其难以对结论形成机理层面的解释。与之形成对照的是，模型驱动研究始终坚守第一性原理的探索传统，通过揭示自然界的底层规律推动理论科学的突破性发展。苦于计算量巨大，在需要使用这些原理时，人们往往陷入空有理论而无法使用的境地。

人工智能技术破解传统范式所面临的发展瓶颈，更开创性地将两者深度融合。即借助算力将模型驱动和数据驱动深度融合，形成AI4S系统工程。从“数据”中提炼经验性“原理”，也使用“原理”来仿真模拟出“数据”。相较于其他领域如自然语言处理中大模型对符号关联的浅层建模，AI4S所实现的机理层融合使得复杂的系统研究成为可能。

图：实现AI4S的技术路径

技术路径	核心优势	典型局限	未来方向
模型驱动	可解释性强，机制透明	规则维护成本高	知识嵌入与混合架构
数据驱动	适应性强，发现隐含模式	数据依赖性强	数据增强与弱监督学习
模型与数据融合	兼顾可解释性与适应性	模型兼容性与推理效率	模块化架构与自动化融合

来源：文献检索，沙利文分析

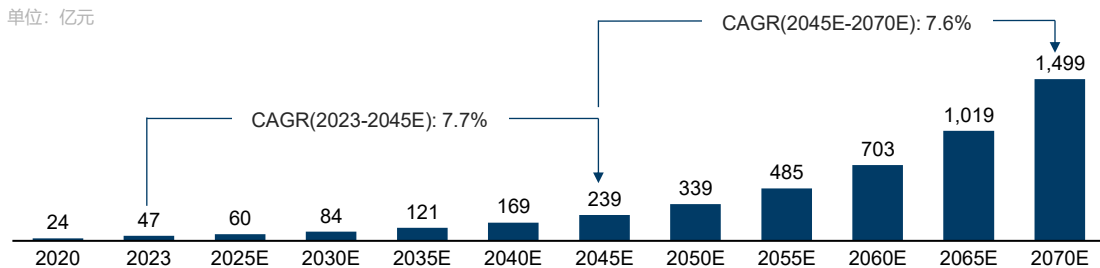
1.3 AI4S应用领域

AI4S通过数据、算法与算力的深度融合，逐步渗透并重构了多个科学领域的研究模式，从生命科学、材料科学到地球科学、能源领域，AI正成为推动创新发展的重要引擎

中国AI4S的市场规模及未来预测

目前中国市场仍处于初期发展阶段，但表现出显著的增长潜力。截至2023年，中国AI4S市场规模已达到47亿元人民币，涵盖药物研发、合成生物学、基因测序、材料开发及电池与储能等核心领域，预计2025年将达到60亿元。从中长期发展来看，随着AI4S迈入3.0阶段并扩展至更多学科，市场规模有望突破千亿元体量，并在2070年达到1,499亿元人民币，发展成为战略新兴产业。

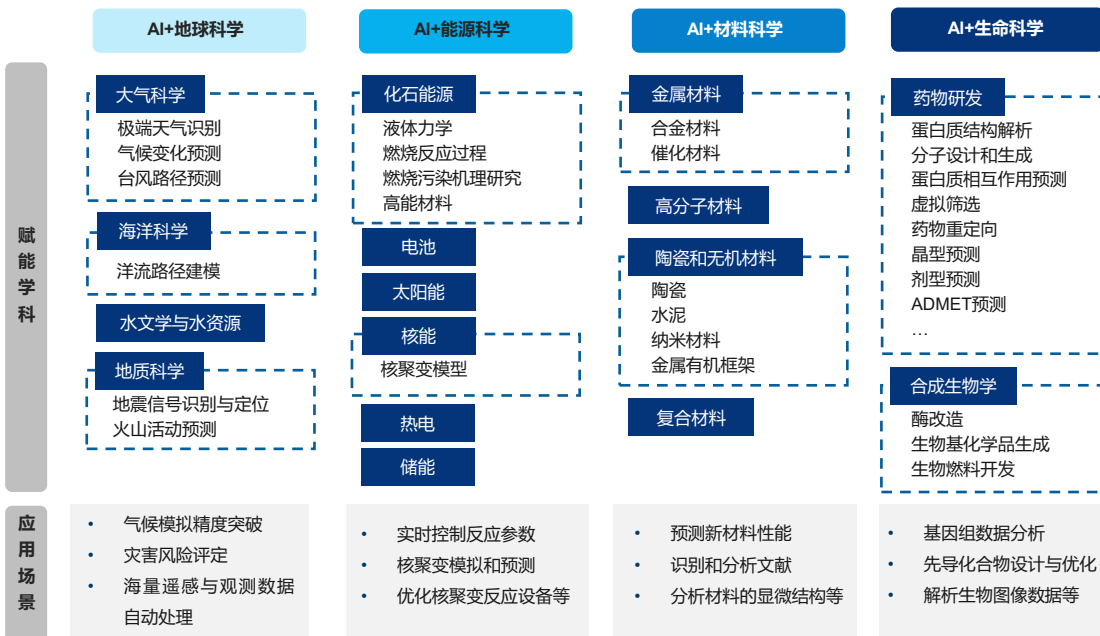
图：中国AI for Science市场规模及预测，2020-2070E



AI4S在多个学科领域得到了广泛应用

从生命科学中的蛋白质结构预测，到材料科学中的新型配方设计，再到地球科学中的气候建模与能源系统优化，AI正深入嵌入科学发现与工程实践的各个环节。跨领域的深度应用不仅体现了AI4S技术的普适性和潜力，更为未来更广泛、更深入的科学创新奠定了坚实的基础。

图：AI4S在能源、材料、地球科学及生命科学等领域的具体应用



来源：文献检索，沙利文分析

1.3 AI4S应用领域

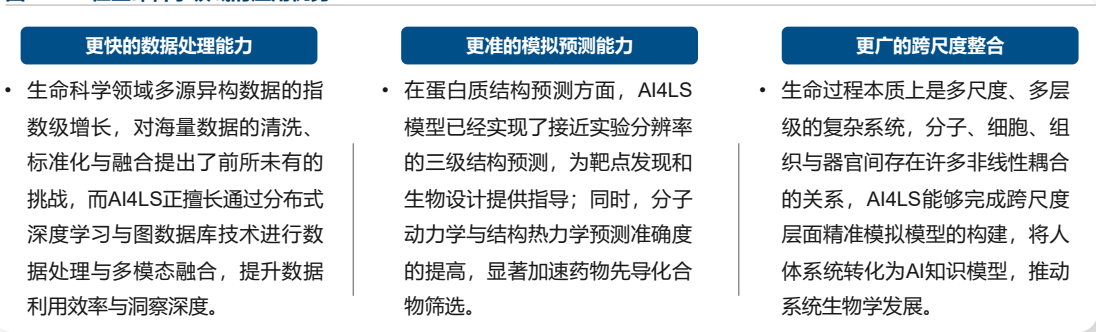
AI在多个领域提升效率与创新，尤其是在生命科学中，AI4LS与生命科学核心需求的契合促成了从基础研究到临床转化的全面加速

生命科学是AI4S展现价值的理想领域之一

AI4S的应用边界正迅速从单一学科拓展至多学科协同，涵盖从物理、化学到地球科学、工程技术等关键领域。在这一广泛赋能的背景下，某些领域因其独特的科研特征与产业需求，成为AI4S能力释放最为彻底、价值体现最为显著的代表。其中，生命科学因其数据基础深厚、问题复杂度高与应用前景广，正逐步确立其作为AI4S最理想应用场景之一的地位。

生命科学领域的研究对象具有极高的复杂性与多样性，从微观的分子结构到宏观的生理过程，往往面临数据过于庞大、变量关系难以解析以及研发周期漫长等诸多难题。而AI4LS恰好具备处理多维复杂数据的天然优势，快速发现隐藏在海量数据背后的潜在规律，并将这些规律转化为精准的预测与优化策略，大幅提升研究与产业开发的效率和成功率，驱动生命科学领域的创新变革。

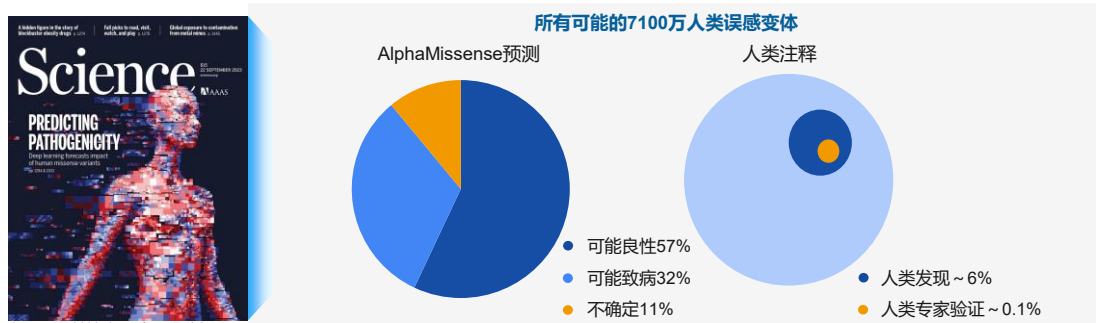
图：AI4S在生命科学领域的应用优势



在众多AI4LS成果中，由DeepMind推出的AlphaFold 3堪称生命科学方向的代表性突破。AlphaFold 3在前代基础上大幅升级，不再局限于蛋白质结构预测，而是首次实现了对DNA、RNA、小分子、离子等多种生物分子的精准建模，具备预测复合体结构和相互作用的能力。其采用Pairformer模块和扩散模块，结合多序列比对和先进的激活函数SwiGLU，并采用DiffusionModule作为生成模块，在提升精度的同时，显著优化了模型的和生成能力。

同年内，又有Chai-2和Boltz-2面向公众开源。Chai-2由Chai Discovery推出，是首个多模态原子级生成模型，能够在无需先验binders的情况下实现抗体、纳米抗体及小蛋白结合分子的零样本设计。Boltz-2由MIT与Recursion联合开发，着重于将蛋白质结构预测与结合亲和力计算统一建模，在保持接近自由能扰动精度的同时提高计算速度。两者的出现分别代表了AI在分子生成设计与物理能量学建模两条路径上的突破，也进一步拓展了AlphaFold所奠定的结构预测范式，使AI在蛋白质科学中的应用从静态结构走向功能导向和药物发现全流程。

图：AI4S在生命科学领域的成功应用案例



1.4 AI4LS在生命科学领域的发展历程

当以较长的时间周期和更宏观的视角去总结和预测AI4LS行业的发展，按照行业总体上需要解决的问题，可以把AI4LS的历史和未来可预见范围内的发展大致分为三个阶段

1.0 概念导入期—“模仿”阶段(2016-2021)

在1.0阶段，AI4LS首次被引入科学研究，主要完成从“实验—模拟”的基础跃迁。AI模型能在已有实验数据基础上进行简单外推，通过计算模拟算法对实验结果进行外推和扩大，如虚拟药物筛选使用打分函数对候选分子打分排序。这极大提升了计算效率，并降低了部分实验负担。然而，该阶段AI仍为实验辅助工具，无法脱离实验体系独立运行，仍有大量时间和成本需要投入到实验中。其可解释性差、工程化弱，缺乏生态闭环，导致整体进展呈“点状突破”。

- AI4LS1.0的角色是“智能外延”。算法基本依赖人工设置的规则进行学习，属于数据的智能延伸，通过深度学习海量实验观测数据实现规律总结。

典型应用

虚拟药物高通量筛选：AI在已知化合物数据库中高效识别潜在药物分子，显著加速研发流程。

关键转折事件

2016年AlphaFold 1.0发布：首次用深度学习预测蛋白质结构，在实验环节的可计算模块引入AI。该事件使“AI可以解决科学问题”的信念首次在科研群体中传播，AI4S作为概念真正成型。

2.0 大规模基础设施建设期—“预测”阶段(2021-2026)

在2.0阶段，AI开始构建具有明确置信区间的可预测模型，实现从“模仿现象”向“指导实验”转变。算法虽然无法十分精准地复现形势世界，但人们对算法能力有清晰认知，有可预测的误差范围和置信区间。科学家可在AI预测的结果基础上进行定向实验验证，大幅提升研发效率。AlphaFold 2等工具具备通用性和工程可落地性，形成初步生态。然而，AI4LS的应用仍受限于系统工程挑战，算法、算力、平台协同的不足使其无法产出预期的精准结果，加之跨学科融合门槛高，导致行业落地缓慢。

- AI4LS2.0的关键词为“预测”，强调有边界、有误差评估、可验证性。AI不再是“盲目学习者”，而是带着约束与目标在特定科学框架内运行。

典型应用

分子动力学模拟：通过融合量子力学与统计力学原理，AI能复现实验现象，更能预测性能表现。

关键转折事件

2020年AlphaFold 2.0在CASP14获胜并开源：以原子级精度预测蛋白质结构，且开源行为建立统一生态接口。在AlphaFold2.0之后，AI从辅助实验升级为替代传统模拟方法。

3.0 成熟应用期—“创造”阶段(2026年及以后)

在3.0阶段，AI将具备从目标函数出发，自主探索科学解答路径的能力。AI技术与科学研究的深度融合将形成全新的研究范式。在这个阶段，AI不仅仅是一种辅助工具，它将成为一种能够提出假设、设计实验并自主学习的新科学方法论的重要组成部分。高级AI系统将在多学科交叉领域中提供前所未有的洞察力，推动科学进步。但困难依然存在，如需要对AI模型进行科学评估，验证其能否完全替代复杂实验；跨行业协同如何达成标准；如何对AI生成的结果进行监管等问题。只有完全解决了这些问题才能发挥AI技术的全部价值，为科研注入能量。

- AI4LS3.0从“预测助手”跃迁为“设计者与探索者”，具备从功能需求到结构设计的全闭环自主能力，真正实现在复杂系统中的自主创新。

典型应用

- **材料组分智能生成系统：**AI在千万种元素组合中锁定最优配方。
- **AI驱动的智能实验室系统**

关键转折事件

2023年RFDiffusion实现从零生成功能性分子：标志AI不再仅预测已有结构，而是创造全新结构。

2024年AlphaFold 3.0获得诺奖认可：AI可以高度准确地预测蛋白质与各种分子相互作用，得到了国际科学界的最高认可。

来源：文献检索，沙利文分析

第二章节

AI4LS多维驱动体系

产业需求牵引、政策引导护航、技术
矩阵支撑与资本赋能

2.1 产业转型升级需求

在基础科学研究的指数级进步与产业转化需求在日益增长的背景下，生命科学行业长期面临着数据规模、实验成本与研发周期的“不可能三角”

01 实验成本阻碍科学突破

实验成本的持续攀升正成为生命科学产业转型升级的硬性约束。从早期高通量筛选到临床前体外实验，大量昂贵且重复的湿实验不断推高研发费用，使新项目的资本回报周期被拉长，投资人和企业的风险承受阈值迅速逼近极限。专业技术人员短缺更进一步放大了时间与资金消耗，造成创新管线在前进途中频频受阻。如何拆解实验环节的高成本，以及对重复性人工操作的过度依赖，已成为企业拥抱AI4LS的直接动力。

I. 湿实验成本高昂：

在创新药研发领域，传统实验范式正面临日益严峻的经济性挑战。实验操作尤其是高通量筛选、基因编辑技术、动物模型等“湿实验”环节在当前生命科学流程中仍占据极高的成本支出，单次百万量级化合物库的筛选成本通常超过百万美元，而后续的反复迭代验证与失败回退更将研发支出推向新高。通过分子生成和筛选，减少需要合成和进入湿实验的分子数量是降低湿实验成本的重要方式之一。没有好的分子，自动化实验只能节省合成过程中的人力成本，无法真正减少湿实验的数量从而控制成本。

例如，Exscientia开发的AI技术平台与行业基准相比，已经达到能将早期设计工作量缩减70%，同时将前期成本也削减80%。望石智慧的Lingo3DMol模型在分子生成上同样取得巨大突破，其与广州实验室合作开发的抗新冠小分子二代3CL小分子口服抑制剂，为国内首个进入临床阶段的二代3CL小分子口服抑制剂候选药物。该项目运用AI技术为药物开发深度赋能，自研发立项至IND申请仅用时12个月。予路乾行开发的Divamics平台，仅1年时间就助推一款非小细胞肺癌靶向药物进入临床试验；另一自主研发的Biotrajectory综合智能体数据库，能够以超过常规数十倍的速度整合和分析文献、专利、临床数据，帮助早研团队快速做出高成功率决策。剂泰科技开发的NanoForge平台下人工智能引导的干湿实验室系统已实现从计算机设计到体内验证的端到端闭环优化，在MTS-004项目中，以NanoForge为技术引擎开发的AiTEM平台承担了剂型开发的建模与预测分析，将临床前制剂优化周期从行业平均的1-2年缩短至3个月内完成。

图：GenAI推动药物研发加速的核心能力



分子生成与筛选引擎

VAE、GAN、Transformer类模型一次即可在云端生成并评估数百万候选分子，几小时内完成过去需数月的虚筛与初步性质预测，从而把命中物发现周期从4-7年缩短到数月或数周。



靶-化合物相互作用预测引擎

深度神经网络在合成前就能计算结合亲和力、ADMET与脱靶效应，使成药性不佳或毒性高的分子在进入实验阶段前就被淘汰。早期标记可将候选药物库质量提高约30%，降低后期失败导致的高昂成本。



自动化合成与工艺规划引擎

AI驱动的合成工具依据数百万条历史反应路径自动给出最优合成路线，最大限度地减少步骤、提高产量并选择可扩展的试剂。通过将逆合成预测与自动化实验室工作流程相结合，将实验室规模的合成速度提升50%。

II. 重复性人工操作依赖：

实验环节中的重复性手动操作亦是实验成本的重要构成之一。根据美国国家科学基金会2023年发布的报告《NSF23-322》，2020年美国医药制造业的人力成本占其研发支出的56%。PhRMA成员企业在2001至2019年间的R&D年支出翻倍，由400亿美元达到了800亿美元。现实情况是大量人力资源消耗于样本处理、移液、记录等低附加值工作。不仅导致科研人员资源被低效占用，还提高了人为误差和实验不可重复率。在此背景下，AI与自动化实验室技术的融合成为主流解决方案。

来源：文献检索，公开信息，沙利文分析

2.1 产业转型升级需求

在基础科学研究的指数级进步与产业转化需求在日益增长的背景下，生命科学行业长期面临着数据规模、实验成本与研发周期的“不可能三角”

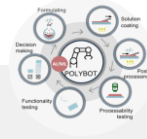
图：自动化实验的应用范例

液体处理智能体Auflo-P系列液体处理智能体



镁伽科技Auflo-P系列液体处理智能体搭载4/8个独立的1000 μ L移液通道，兼容各类样本、试剂耗材和功能模块。搭配96通道移液臂和抓手，可整机自由配置，满足各种复杂的实验室自动化场景需求。广泛应用于生物制药、生命科学等领域，持续优化实验可靠性和自动化运行时间，加速研究与发现进程。

自驱实验室Polybot智能实验平台



AI驱动的“自驱实验室”能够进行自动化实验设计、执行和分析，并全天候运行，显著提升实验通量和数据质量。美国国家实验室Argonne开发的Polybot智能实验平台，可在数周内完成9万个材料组合的筛选，人力操作需数月时间，表明“自驱实验室”在成本与效率上的颠覆性潜力。

02 研发周期放大失败风险

在生命科学产业中，研发周期过长与项目成功率偏低是长期制约行业效率与创新活力的核心难题。在资本趋于审慎、市场竞争加剧的背景下，传统研发流程显得愈发笨重低效。因此，如何在保障科学严谨性的前提下压缩时间成本、提高决策效率与试验命中率，正成为驱动产业向AI4LS加速融合的关键动因。

I. 线性流程与高失败率：

生命科学领域的传统研发流程长期呈现出明显的线性特征：从靶点发现到先导化合物筛选，再到临床前研究、临床试验直至最终上市，每一步环节依赖前一阶段的完成，整体流程耗时且高度依赖顺序推进。这一“串联式”逻辑不仅导致整体开发周期动辄延长至10至15年，还因各阶段无法交叉验证而在发现问题时付出更高的时间和成本代价，尤其是在临床后期失败时，往往意味着数年投入的付诸东流。

人工智能的引入正打破这一研发惯性。通过高度并行化的计算能力，AI能在项目早期同时处理靶点验证、毒性预测与分子优化等多个决策环节，使原本必须串行进行的工作实现并发运行，从而显著压缩整体时间轴。AI还能在数据驱动下精准识别失败风险较高的候选，提前淘汰无效管线，降低后期试错成本。PhRMA的报告显示，AI应用可将早期研发周期缩短60%至70%，这一转变已成为全球制药企业提升效率、加快创新的重要战略路径。

II. 临床招募与设计滞后：

临床试验是药物研发流程中最复杂且最耗时的阶段。因招募进度缓慢而被迫延期的临床试验，通常会进一步引发方案调整、预算溢出和项目进度失控等连锁反应。而在多中心、多人群、多变量参与的现实条件下，试验管理的复杂性常常导致数据偏倚与资源浪费。通过对电子健康记录（EHR）和真实世界数据的深度挖掘，AI能精准识别符合纳入标准的患者群体，提高招募效率；同时借助数字孪生建模与动态预测分析，AI可实时优化试验方案，提升执行灵活性与资源调配效率。

在一项143例血液瘤患者的AI药物匹配实验中，Exscientia借助AI技术实现了前端精准匹配与后端方案优化的闭环式赋能，有效延长了54%患者的疾病控制期。这种方式解决了传统临床流程中“找人难、疗效不确定、设计刚性”的痛点，是AI4LS在临床研究阶段落地的直接佐证。

来源：NSF，文献检索，产品官网，公开信息，沙利文分析

2.1 产业转型升级需求

在基础科学研究的指数级进步与产业转化需求在日益增长的背景下，生命科学行业长期面临着数据规模、实验成本与研发周期的“不可能三角”

03数据规模限制研究效率

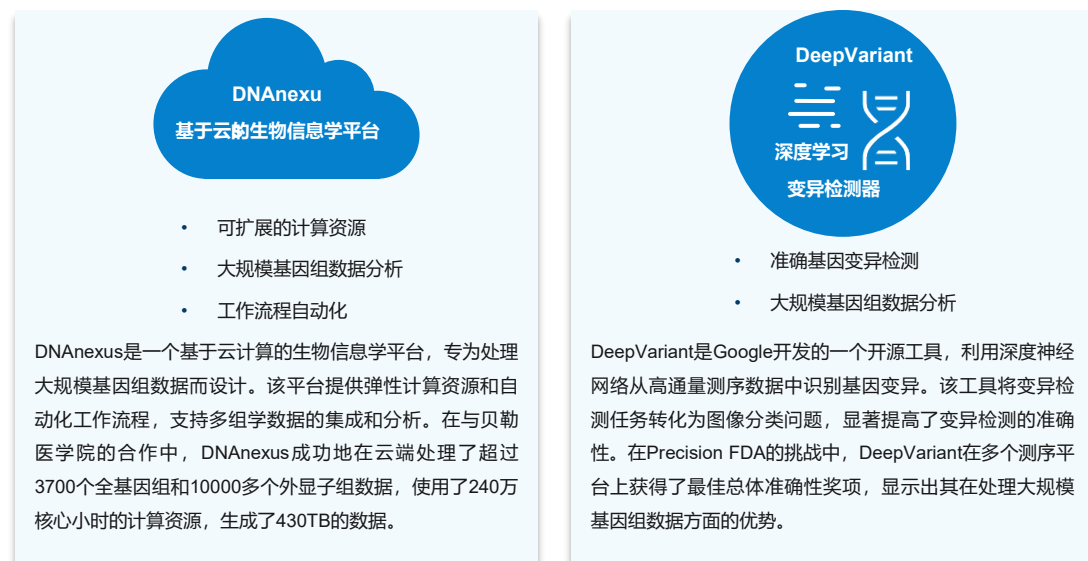
在生命科学迈向数据驱动时代的进程中，日益庞大的数据规模已成为推动产业转型的关键力量。这一趋势不仅体现在数据生成速度的激增，还包括跨机构数据整合的障碍和高质量标注资源的匮乏，全面制约了研究效率和知识转化能力。随着传统工具难以支撑对多源异构数据的高效处理，人工智能技术，特别是具备多模态学习与大规模训练能力的模型，正在成为破解数据瓶颈、驱动生命科学创新的核心手段。与此同时，数据规模带来的挑战，也正逐步转化为AI4LS发展的现实需求与战略机遇。

I. 数据爆发与处理压力：

随着基因组学、单细胞组学、高通量影像技术的广泛应用，生命科学领域所产生的数据量呈爆炸式增长，年度数据规模已达到EB级。以基因测序为例，英国的UKBiobank项目已对50万名参与者进行了全基因组测序，其数据库总数据量超过30PB；美国国家人类基因组研究组织预计未来十年内，全球基因组研究所产生的数据最高将达到40EB。面对如此海量的数据，传统生物信息学分析技术面临双重瓶颈：在计算效率层面，传统的特征提取与统计建模方法难以应对海量异构数据的实时处理需求；在算法性能层面，线性模型与浅层机器学习架构无法有效捕获多模态数据间的非线性关联，导致关键生物标志物发现与疾病机制解析的周期显著延长。

在此背景下，为了应对算力与算法的双重滞后性，深度神经网络及大语言模型展现出独特的技术优势。其通过自适应多模态表征学习机制，能够跨越基因组变异、蛋白质相互作用、影像表型等多维空间，实现端到端的生物特征抽象与知识蒸馏。同时，模型训练所需的海量计算资源正驱动硬件基础设施的革新迭代——基于GPU/TPU异构加速的超级计算集群已成为科研加速器。由此催生出两类新型技术生态：一方面，云计算赋能的生物信息平台（如DNAnexus、SevenBridges）通过弹性算力池化解大规模并行计算需求；另一方面，集成数据治理、模型训练与结果可视化的全栈式AI解决方案（如NVIDIA Clara、Google DeepVariant）正在重塑生物医学研究的工作流范式。

图：AI数据处理实践案例



来源：文献检索，公开信息，沙利文分析

2.1 产业转型升级需求

在基础科学研究的指数级进步与产业转化需求在日益增长的背景下，生命科学行业长期面临着数据规模、实验成本与研发周期的“不可能三角”

II. 数据孤岛与共享障碍：

随着高通量测序、质谱分析、影像诊断等技术的广泛应用，研究所涉及的数据类型从基因组、转录组、蛋白质组、表观遗传组等分子层级信息，延伸到医学图像、结构化检验指标、非结构化病历文本、传感器数据和临床行为记录等多层级、多模态数据体系。

这些数据在采集方式、格式结构、语义标准、时间频率和完整性上存在显著差异，导致信息分布碎片化，难以在统一的分析框架下进行整合利用。此外，在真实世界的医疗和科研环境中，数据模态往往由不同系统独立生成，缺乏跨平台兼容机制，形成孤立的数据岛，严重阻碍了对复杂生物机制的系统性理解与精准建模。

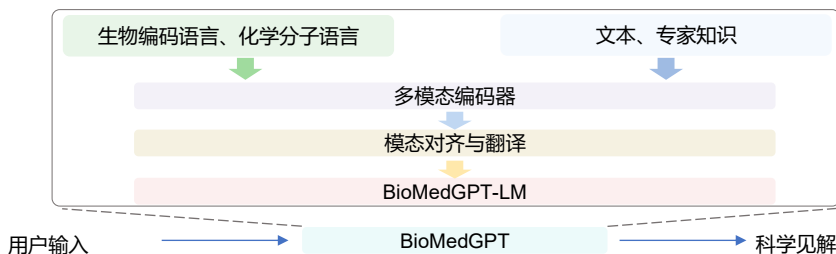
图：AI处理多模态数据的原理与机制

- | | |
|---|--|
| <p>1 模态嵌入与统一表征学习：数据（图像、文本、序列、数值）可通过特定神经网络模块编码为向量形式，例如卷积神经网络（CNN）处理影像数据；循环神经网络（RNN）或Transformer处理临床时间序列；自注意力机制用于跨模态关联建模。所有模态最后嵌入到同一语义空间中进行联合学习，捕捉其间的协同特征。</p> | <p>2 多模态对齐与注意力机制：AI利用对齐机制识别不同模态中“语义等价”的成分（如影像中肿块与病理报告中的“低分化腺癌”），再通过注意力模块赋予其不同权重，以实现精细判别与决策。</p> |
| <p>3 端到端预测与优化反馈：相较于传统分步分析，AI模型可端到端训练，直接输出诊断、分型、药物反应预测等结果，并通过误差反向传播机制优化整个网络结构，实现自适应学习。</p> | <p>4 处理缺失与非结构化数据的能力强：深度模型可通过自编码器、图网络补全缺失值，或利用自然语言处理模型将文本转化为结构化特征，打通传统“只处理数值型数据”的瓶颈。</p> |

面对传统的分析方法难以处理异构数据之间复杂的非线性交互关系、特征提取、语义对齐和跨模态映射方面能力有限、无法有效挖掘多模态数据间的潜在关联等挑战，AI具备天然的“特征提取、非线性表达、统一嵌入”的能力，为多模态数据的融合与智能利用提供了全新路径。在数据预处理最小化的前提下，AI自动完成信息的学习与融合。通过自适应的神经网络结构，将基因变异、影像表现和临床表型等不同来源的信息集成至统一的表征空间，实现复杂生物信号的跨层级推理与动态预测。

以由清华大学智能产业研究院与水木分子联合开发的多模态生物医药大模型BioMedGPT为例，其能够将分子图谱、蛋白质序列等生物模态与自然语言嵌入到统一的特征空间，实现跨模态问答、知识推理和生成任务。BioMedGPT-10B在多个生物医药问答基准数据集上实现了最先进水平，其专业领域问答能力比肩人类专家，成功通过了美国医师资格考试。此外，模型还支持分子问答和蛋白质问答任务，显著提升了新药发现和靶点识别的效率。

图：BioMedGPT的介绍



来源：文献检索，公开信息，沙利文分析

2.1 产业转型升级需求

在基础科学研究的指数级进步与产业转化需求在日益增长的背景下，生命科学行业长期面临着数据规模、实验成本与研发周期的“不可能三角”

III. 数据质量与通量瓶颈：

在生命科学领域传统研发中，一个核心矛盾是：想要获得高质量、标准化标注的数据，就必须牺牲实验通量和速度，并高度依赖操作人员的技术水平，导致数据难以复用和规模化积累。在实验设计阶段，样本量不足或取样存在偏差是常见问题。受限于成本、时间或动物伦理审批，许多研究使用的样本数量远未达到统计学要求，导致数据无法反映真实的生物学特征。另外，现代生命科学问题往往需要从药物剂量、药效、细胞类型等多维度探索，而传统实验平台在物理通量和成本上无法有效承载对应的海量样本，研究者往往被迫采用一次只改变一个变量的简化策略，这不仅效率低下，更无法捕捉变量间的交互作用，从而限制了数据的深度与科学性。

实验执行阶段是数据质量问题的“高发区”。人工操作具有不可避免的低效率和易疲劳特点，实验结果高度依赖研究人员的个人技能与经验，批次间差异和主观误差亦难以避免，后期操作的稳定性难以保证与初期一致。同时，仪器设备的误差也制约了数据产生的规模与速度。这种低效率限制了大规模、系统性的数据积累，造成了高质量数据较为稀缺的问题。

全流程智能自动化实验系统提出一个全新的概念，即通过机器人技术、物联网和AI调度算法，使得将原本依赖人工操作的湿实验流程整合为一个7x24小时不间断运行的闭环系统成为可能。全流程智能自动化实验系统通过机器人执行标准化操作，消除了人工操作引入的批次间差异和主观误差，从数据生成的源头确保了其一致性与可靠性。另外，系统能够自动记录每次操作，例如为每一份检测样本自动关联了其精确的培养历史、处理时间和试剂批次。这使得海量实验数据不再是孤立的数值，而是带有丰富、准确的元数据标注。这种高质量与一致性的标准化数据流正是训练和优化AI预测模型最理想的原材料，使得后续的数据分析和模型迭代能够建立在坚实可靠的基础之上。

目前这一理念已在多个前沿生命科学领域得到实践，并催生了一系列成功的研究案例。斯坦福大学团队开发的CRISPR-GPT系统成为了“人类提出假设-AI设计实验-机器人执行”新科研模式的典范。该系统能够与Agilent Bravo自动化液体处理工作站等实验硬件整合，AI生成的标准化、机器可读的实验指令可直接驱动Bravo工作站执行精准的液体处理操作，从而实现了从实验设计到湿实验执行的端到端自动化。研究表明，研究人员利用该系统，首次尝试便成功在肺癌细胞中高效敲除特定基因，并在黑色素瘤细胞中精准激活目标基因。全流程智能自动化实验系统已不再是理论构想，它正通过将实验生物学转化为可计算、可编程的标准化数据流程，驱动生命科学走向更高通量、更可重现的新范式。

图：传统实验室与智能实验室工作重心对比

工作类别	传统实验室（以人工为主）	智能实验室（人机协同）
实验执行	科学家亲手进行移液、称量、培养等大量重复劳动	由机器人、机械臂和自动化工作站7x24小时完成，标准、精准
实验设计	基于有限数据和经验进行设计，试错成本高	提出关键科学问题，制定研究目标和实验策略
数据分析	耗费大量时间进行基础的数据整理和图表绘制	人机协作：AI模型快速处理海量数据，科学家从中挖掘洞见、形成新假设
系统维护优化	维护基础仪器设备	维护、校准复杂的自动化系统，并编程和优化机器人实验流程
创新与决策	埋没于繁琐操作，创新思考时间被压缩	科学家解读AI发现的意外结果，提出颠覆性新思路，做出重大决策

来源：文献检索，公开信息，沙利文分析

2.1 产业转型升级需求

在基础科学研究的指数级进步与产业转化需求在日益增长的背景下，生命科学行业长期面临着数据规模、实验成本与研发周期的“不可能三角”

04 伦理压力影响实验进展

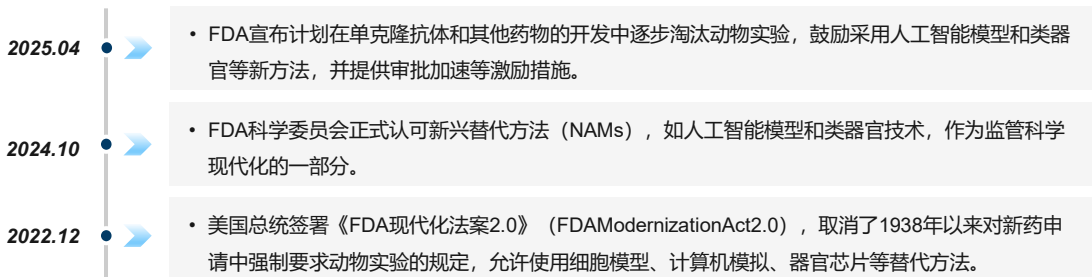
在生物医学研发领域，传统实验体系的可持续性正遭受高度挑战。当前，以高通量筛选为代表的实验技术虽能加速数据采集，但其依赖的专用耗材与大规模动物模型导致单次实验成本高达数十万美元。而涉及灵长类等高等动物的复杂药效评估，则因伦理审查压力不断升级面临实施壁垒。更严峻的是，实验方案的不可复现性进一步加剧资源浪费，导致大量初期实验结果无法向临床研究阶段有效转化。资源密集型的研发模式推高研发成本，更在伦理合规与生态可持续性层面引发系统性风险。

在此背景下，计算生物学与智能实验室技术正构建新一代研发基础设施。基于多尺度建模的数字孪生体（DigitalTwin）技术，通过整合器官芯片（Organ-on-a-Chip）、类器官培养与AI驱动的生理系统仿真，实现对药物代谢、毒性效应的体外高精度预测。深度学习赋能的虚拟患者队列则能模拟人群异质性，替代部分临床试验的受试者招募。



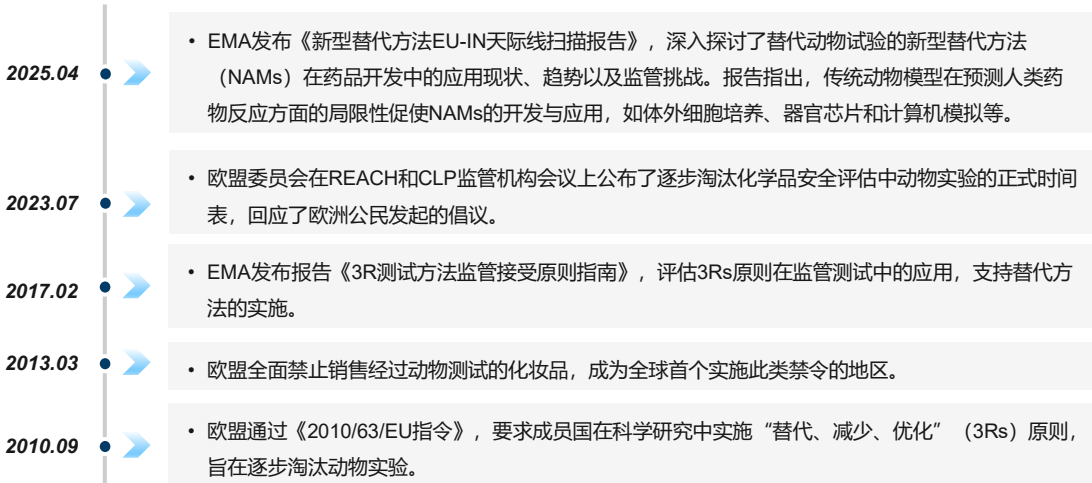
FDA正通过立法与监管更新，加速将AI、器官芯片等新兴技术纳入药物评估流程，逐步取消对动物实验的依赖

图：FDA逐步取消动物实验的关键政策与事件



EMA以政策指令和科学评估为抓手，系统推动动物实验的替代方法在化学品与药品开发中的广泛应用

图：EMA推动动物实验替代的政策与进展



来源：FDA, EMA, 文献检索, 公开信息, 沙利文分析

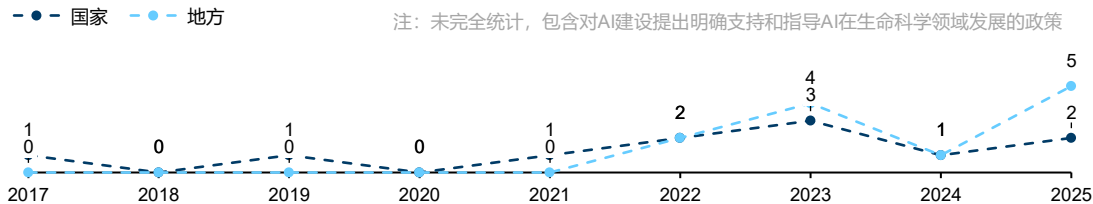
2.2 顶层设计与政策布局

中国政府高度重视以AI为核心的技术革新在生命科学领域的战略价值，构建了从国家顶层设计到地方实践支持的多层次政策体系，为AI4LS的技术演进与产业化落地提供了坚实支撑

在新一轮科技革命与产业变革背景下，与人工智能技术的融合正成为推动生物医药产业升级的核心动力。中国政府高度重视以AI为核心的技术革新在生命科学领域的战略价值，围绕AI4LS的技术基础设施、数据平台、算法突破、算力供给以及人才培养等关键要素，构建从国家顶层设计到地方实践支持的多层次政策体系。

面向AI4LS的政策，呈现数量持续攀升、内容不断细化、关注重心由宏观顶层设计向示范应用和治理标准迈进的发展趋势。早期政策以国家级规划和专项经费为主，着力构建大数据平台、高通量自动化实验中心与公共算力基础；进入“十四五”阶段后，陆续下发一系列聚焦场景示范的应用通知，并配套生成式AI管理办法及标准建设征求意见稿，表明监管与标准化正成为行业健康发展的核心。国务院也于2025年8月印发《关于深入实施“人工智能+”行动的意见》（国发〔2025〕11号），将医疗健康作为应用方向之一，推动人工智能与生物医药的深度融合。与此同时，各地政府也不断补强区域创新生态，聚焦示范区打造、人才引进、专用平台孵化及产业全链条扶持。总体来看，AI4LS政策正从“量”到“质”双维度演进，既加大经费与项目支持力度，也日益强调应用场景落地、标准规则和可持续治理，为AI驱动的生命科学研究和产业化转型奠定了更为全面的制度基础。

图：支持政策数量变化解析（2017-2025年）



国家级政策支持：顶层设计持续加码，科技投入步步推进

2017年，《新一代人工智能发展规划》明确将人工智能上升为国家战略、为生命科学领域的大数据与算法创新奠定基础。此后，我国持续通过一系列顶层设计和专项部署不断加大对AI4LS的支持力度。2019年，科技部在全国范围内启动人工智能创新发展试验区建设，为AI技术在药物筛选、基因编辑等生命科学场景中进行示范应用提供了“沙箱”环境；紧接着2021年发布的重大项目指南，则对人工智能技术发展、落地应用场景等核心方向给予了专项经费保障。进入“十四五”期间，中央通过《质量强国建设纲要》和《数字中国建设整体布局规划》两大纲领性文件，强化了智能制造与数字底座的协同发展。

与此同时，政策层层推进。从2022年起，国家自然科学基金和科技部陆续设立“人工智能驱动的科学基础研究”专项，并下发多轮支持示范应用场景的通知，到2023年国家网信办等部门又联合发布生成式AI服务管理办法和AI标准体系建设征求意见稿，全面推动AI模型、算力和数据标准化。

图：2017-2025年部分国家出台支持AI相关产业发展政策

时间	发布单位	文件名称	相关内容
2017年7月	国务院	《新一代人工智能发展规划》	<ul style="list-style-type: none"> 规划提出了三个重大战略目标，包括2020年与世界先进水平同步，2025年取得重大突破，2030年达到世界领先水平。
2019年9月	科学技术部	《国家新一代人工智能创新发展试验区建设工作指引》	<ul style="list-style-type: none"> 指引全面实施试验区建设。强调适应人工智能发展的特点和趋势，深化创新链与产业链融合。 预计到2023年，将建立20个试验区。

来源：政府官网，文献检索，公开信息，沙利文分析

2.2 顶层设计与政策布局

中国政府高度重视以AI为核心的技术革新在生命科学领域的战略价值，构建了从国家顶层设计到地方实践支持的多层次政策体系，为AI4LS的技术演进与产业化落地提供了坚实支撑

2021年8月	科学技术部	《新一代人工智能重大项目2021年度项目申报指南》	<ul style="list-style-type: none"> 指南指出将在四个核心技术方向启动共计21项研究任务。 预计将分配约5.34亿元国家资金。
2022年1月	国务院	《“十四五”数字经济发展规划(2021-2025)》	<ul style="list-style-type: none"> 规划明确了包括人工智能领域创新能力在内等关键任务。 在生物经济发展方面，通过强化科技攻关、优化市场准入与医保政策，打造国家生物战略科技力量并推动创新成果转化。
2022年8月	科学技术部	《关于支持建设新一代人工智能示范应用场景的通知》	<ul style="list-style-type: none"> 为加快人工智能应用并培育国民经济新增长点，通知明确：首批应用场景涵盖农业、医疗诊疗等行业的智能化应用。
2023年3月	国家自然科学基金委员会	《人工智能驱动的科学“专项部署工作”》	<ul style="list-style-type: none"> 该部署紧密围绕基础学科的关键问题。 重点关注药物研发、基因研究、生物育种以及新材料研发等关键领域的科学研究需求。
2023年7月	国家网信办、国家发改委等八个部门	《生成式人工智能服务管理暂行办法》	<ul style="list-style-type: none"> 提出有效措施鼓励生成式人工智能创新，对生成式人工智能服务实施包容、审慎、分级监管，促进产业健康发展。
2023年8月	科学技术部	《关于支持新一代人工智能示范应用场景建设的通知》	<ul style="list-style-type: none"> 呼吁加强新一代人工智能示范应用场景建设。鼓励地方政府和企业探索在智慧城市、智能制造和智慧医疗等领域的创新。
2024年3月	中共中央办公厅、国务院办公厅	《数字中国建设整体布局规划》	<ul style="list-style-type: none"> 提出加快推进数字基础设施建设，通过数字化转型赋能传统产业，培育新数字产业。促进人工智能与产业的融合发展。
2025年8月	国务院	《关于深入实施“人工智能+”行动的意见》	<ul style="list-style-type: none"> 加速科学发现：探索AI驱动的新科研范式，建设科学大模型。 -驱动技术研发：加强AI与生物制造等领域的技术协同创新。
2025年9月	国家自然科学基金委员会	《国家自然科学基金委员会“未来生物技术”项目指南》	<ul style="list-style-type: none"> 明确资助：人工智能驱动的生命功能分子预测与设计合成；人工智能模拟复杂生物系统的数字孪生技术；具身智能驱动的智慧医学技术。

来源：政府官网，文献检索，公开信息，沙利文分析

2.2 顶层设计与政策布局

中国政府高度重视以AI为核心的技术革新在生命科学领域的战略价值，构建了从国家顶层设计到地方实践支持的多层次政策体系，为AI4LS的技术演进与产业化落地提供了坚实支撑

地方政策推动：差异化试点布局，打造区域创新高地

自2022年以来，各地政府纷纷出台系统性举措，以法律法规、行动方案和专项扶持的方式，构建差异化的AI4LS发展生态，为AI4LS的算法突破、数据共享和产业化落地提供区域支撑。

深圳率先在2022年通过《深圳经济特区人工智能产业促进条例》，为全市AI产业高质量发展提供法治保障，并于2023年推出高水平应用行动方案，明确打造国家级试验区和创新应用先行区。上海同月颁布《上海市促进人工智能产业发展条例》，建立市级战略咨询专家委员会，并在2024年“模塑申城”实施方案中推动智能算力集群、行业大模型与数据共享平台建设，力争到2025年形成50余个行业示范。北京则自2023年起接连发布《加快建设具有全球影响力的人工智能创新策源地实施方案》《促进通用人工智能创新发展的若干措施》《北京人工智能训练计算中心建设指南》等政策，聚焦人才培养、算力中心、通用AI和企业培育，力求将核心产业规模推升至万亿级。苏州通过千万元级别的专项补贴和开放创新平台，助力国家新一代AI创新试验区落地。成都在2024年行动计划中锁定千亿级产业规模目标，聚焦基础创新和生态示范。2025年初，深圳发改委再度推出覆盖医药器械全链条的32条扶持措施，支持AI辅助研发公共服务平台及重大产业项目。

图：2022-2025年部分各地出台支持AI相关产业发展政策

时间	发布单位	文件名称	相关内容
2022年9月	深圳人民代表大会	《深圳经济特区人工智能产业促进条例》	<ul style="list-style-type: none"> 中国首部人工智能产业专项立法。 旨在推动深圳经济特区人工智能产业的高质量发展。
2022年9月	上海市人民政府	《上海市促进人工智能产业发展条例》	<ul style="list-style-type: none"> 成立市人工智能战略咨询专家委员会，明确市级经信、发改、科技、网信等多部门职责分工。 推动“东数西算”在长三角枢纽节点的落地，并完善生命健康等领域公共数据动态开放清单。
2023年5月	北京市人民政府	《北京市加快建设具有全球影响力的人工智能创新策源地实施方案（2023-2025年）》	<ul style="list-style-type: none"> 为AI在包括医药健康在内的各行业应用提供了宏观指导和发展环境；到2030年，人工智能核心产业规模将达到3000亿元。
2023年5月	深圳人民代表大会	深圳市加快推动人工智能发展高水平应用行动方案（2023-2024年）》	<ul style="list-style-type: none"> 该方案的目标是凝聚全市合力，建设人工智能国家级试验区和人工智能创新应用先行区。 在医疗卫生机构探索开展临床辅助决策、医学影像辅助诊断、医用机器人等人工智能应用试点。

来源：政府官网，文献检索，公开信息，沙利文分析

2.2 顶层设计与政策布局

中国政府高度重视以AI为核心的技术革新在生命科学领域的战略价值，构建了从国家顶层设计到地方实践支持的多层次政策体系，为AI4LS的技术演进与产业化落地提供了坚实支撑

2023年5月	苏州市人民政府	《关于进一步支持国家新一代人工智能创新发展试验区、国家生物药技术创新中心、国家第三代半导体技术创新中心（苏州）的若干政策》	<ul style="list-style-type: none"> 为国家AI创新试验区提供最高1000万元绩效补贴，推动AI技术攻关、创新平台建设，企业孵化与应用。 通过十大举措，系统性推动人工智能与生物医药等战略性新兴产业的深度融合与协同创新。
2023年11月	北京市人民政府	加快北京成为全球人工智能先锋城市建设行动计划	<ul style="list-style-type: none"> 到2025年，旨在培育5-10家全球有影响力的人工智能领军企业及200家高增长人工智能企业。人工智能产业规模将超过3000亿元，相关产业增加值将超过1万亿元。
2024年11月	上海市人民政府	《上海市发展医学人工智能工作方案（2025—2027年）》	<ul style="list-style-type: none"> 系统规划了AI在临床诊疗、药物研发、医疗器械、公共卫生等领域的深度融合与应用。
2025年4月	深圳市发改委	《全链条支持医药和医疗器械发展若干措施》	<ul style="list-style-type: none"> 围绕研发创新、AI应用、临床试验等九大方面提出32条措施，支持建设AI辅助研发公共服务平台和重大产业项目。
2025年7月	北京市人民政府	《北京市加快推动“人工智能+医药健康”创新发展行动计划（2025-2027年）》	<ul style="list-style-type: none"> 通过系统布局前沿技术研发、场景应用与生态建设，以人工智能驱动医药健康产业的全链条创新与高质量发展。
2025年9月	江苏省人民政府	《中国（江苏）自由贸易试验区生物医药全产业链开放创新发展专项实施方案》	<ul style="list-style-type: none"> 推动大数据、人工智能在药物靶标筛选、药物分子设计、医疗器械设计制造等方面典型场景和解决方案中率先推广应用。
2025年9月	山东省发展和改革委员会等14部门	《关于支持细胞与基因治疗产业高质量发展的若干措施》	<ul style="list-style-type: none"> 实施“人工智能+生物医药”专项行动；打造AI应用中试基地，加速药品筛选与临床试验。
2025年10月	上海市政府办公厅	《关于加快推动前沿技术创新与未来产业培育的若干措施》	<ul style="list-style-type: none"> 发展壮大细胞与基因治疗、脑机接口等领域；加快培育类脑智能等领域。

来源：政府官网，文献检索，公开信息，沙利文分析

2.3 技术矩阵的搭建

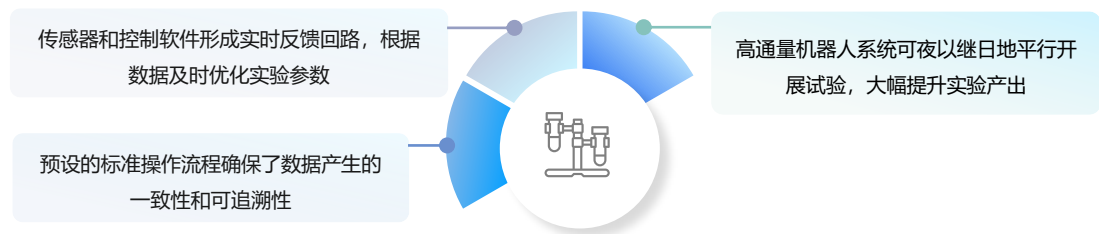
AI4LS在生命科学领域的落地依托于五大核心技术支柱的快速发展，从而形成“数据-算法-算力-知识-团队-新数据-反哺模型优化”的闭环科研生态

1 数据资源——生命科学领域的“基础设施”

I. 自动化数据生产：

大规模、高质量的数据资源是AI4LS模型训练和科学发现的基石，而数据获取方式正由传统的人工实验逐步转向自动化的生产模式。过去人工依赖的实验流程存在效率低、周期长且易受人为误差影响等弊端，数据质量和一致性难以保障，很难满足当前AI模型对海量且精准数据的需求。

自动化数据生产被视为生命科学领域的“基础设施”，其意义在于数据产出量的指数级扩张，更体现于通过标准化、闭环化和高可信度的数据供给，破解长期困扰科研的“实验效率—数据质量—模型性能”三元约束，为AI模型提供源源不断的高质量训练数据。引入自动化实验平台后，实验的设计执行、数据采集与清洗过程实现了机器化和程序化：



实验室自动化平台融合机器人操作与AI决策，被定位为生命科学的“数据工厂”。实验执行的灵活性和一致性均得到增强，数据质量可控可溯，显著提升了实验效率和数据产出规模。

以镁伽科技的自动化平台为例，其基于感知、构思及执行三大技术构建，融合多个由LibraX与Labillion驱动的自主智能体，可将AI预测转化为实验规程，由中央数智化系统统一调度控制，提升效率与准确性，模块之间无缝衔接形成闭环实验流程。Megalaxy Lab3.0可提升实验通量和可重复性，将开发周期从6-10个月压缩至6-8周。其中，MegaMolecule全自动分子实验集群每天能完成接近400项复杂生物分子构建任务，与传统手工实验相比效率大幅提升；MegaCell全自动细胞实验集群每天可完成超过10,000个样本点的药物筛选实验，在高通量药物筛选和稳定细胞系构建方面效果显著，展现出广阔的商业化前景。

Laboratory automation system

自动化实验室

通过机器人执行与数智化调度，高通量、标准化地产出可信数据，成为连接实验与算法的实体数据工厂



定位

将传统“人工操作+离散仪器”升级为端到端机器人执行+数智化调度的实验室整体解决方案

核心技术

- 高通量：7×24连续运行，实验吞吐量提升一个数量级
- 高一致性：标准化SOP与全程追溯，批间差异显著降低
- 高闭环：实时数据回传至AI决策层，实验参数动态优化

技术构成

- 智能机械臂与自动化工站（移液、克隆、细胞培养）
- 物联传感网络与调度中枢
- 数据湖+流式分析引擎
- AI决策模块（实验规划、异常检测、参数自适应）

来源：文献检索，公开信息，沙利文分析

2.3 技术矩阵的搭建

AI4LS在生命科学领域的落地依托于五大核心技术支柱的快速发展，从而形成“数据-算法-算力-知识-团队-新数据-反哺模型优化”的闭环科研生态

人工智能虚拟细胞模型 (AIVirtualCell, AIVC) 能够在不同环境和状态下动态模拟细胞的生物学行为。这一数字孪生的虚拟细胞可以高精度地控制实验条件并重复试验，避免了传统湿实验中由环境变化引起的误差，并大幅降低了时间和资源成本。借助虚拟细胞模拟，科学家得以预测细胞对基因编辑、药物作用等干预的响应，从中筛选出最有前景的方案再投入实际实验，从而优化研究路线。

AIVirtualCell

虚拟细胞模型

依托多模态真实数据与神经网络，在计算空间重建细胞动态，为机理探索和药物干预提供数字孪生平台



定位

以神经网络为引擎、以真实组学与显微影像为驱动，构建可计算、可扰动、可进化的in-silico细胞

AI引擎

- 多尺度神经网络融合 (GNN+ 时序Transformer)
- 物理约束损失函数确保生物学合理性
- 强化学习驱动的扰动策略搜索

数据支柱

- 先验知识层：文献、数据库中的生物网络与通路
- 静态结构层：单细胞组学与高分辨显微影像
- 动态状态层：扰动实验产生的时间序列多组学数据

II. 多模态生命科学数据：

当前生命科学研究呈现出多模态数据并存的特征，传递出的信息角度多样：基因组学和单细胞测序数据提供遗传与表达层面的精细视角，全球性的Human Cell Atlas计划构建了单细胞级别的人类细胞图谱；蛋白质组学结合AlphaFold预测的蛋白质三维结构数据库为理解分子功能提供了前所未有的详实依据；大规模生物医学影像与临床数据的关联分析正在支撑精准医疗的发展，将医学影像与患者基因组数据相关联，有助于更准确地分类疾病亚型和制定个性化治疗方案。

将这些多源异构的数据信息融合在一起，从整体上勾勒生命过程的全貌，远超单一数据维度所能提供的洞见。相较于以往的单因素研究，多模态生命科学数据被视为驱动新一代科研范式升级的重要资源，其综合运用为发现复杂生物过程中的关联提供了基础。

对多模态数据价值的挖掘，离不开擅长识别与融合分析大规模数据的人工智能技术。通过引入深度学习等AI方法，可以自动提取基因序列、蛋白结构、影像病灶、临床表型等高维特征，并在统一的模型中学习它们之间的关联，在更高层次上揭示生命科学的规律。

图：多模态生命科学数据分析依托技术

Transformer架构

- Transformer架构已经被用于基因组序列建模和蛋白质相互作用预测，提升了对蛋白结构和功能关系的辨识能力。

图神经网络 (GNN)

- 图神经网络 (GNN) 将分子结构表示为图谱，从而加速药物分子筛选和性质预测，在药物研发中取得了显著成效。

来源：文献检索，公开信息，沙利文分析

2.3 技术矩阵的搭建

AI4LS在生命科学领域的落地依托于五大核心技术支柱的快速发展，从而形成“数据-算法-算力-知识-团队-新数据-反哺模型优化”的闭环科研生态

2 算法平台——从海量数据中萃取科学发现

I. 深度学习架构:

深度学习的突破为AI在生命科学中实现复杂模式识别、结构建模和系统预测奠定了方法论基础。以多层神经网络为代表的深度学习架构，具备强大的非线性建模能力，能够从大规模生物数据中自动学习高阶特征，解决传统统计模型难以应对的维度灾难与数据异构问题。

随着Transformer、GNN等架构的引入，AI模型在处理序列信息和多模态输入方面能力显著增强，使得从DNA序列到蛋白质结构、从化合物结构到表型行为的“端到端建模”成为可能。

图：深度学习架构应用实例



来源：文献检索，公开信息，沙利文分析

2.3 技术矩阵的搭建

AI4LS在生命科学领域的落地依托于五大核心技术支柱的快速发展，从而形成“数据-算法-算力-知识-团队-新数据-反哺模型优化”的闭环科研生态

II. 强化学习优化:

区别于监督学习，强化学习通过构建与实验系统或生理模型的交互环境，不断试错学习、获取反馈并优化策略，从而实现实验设计、参数调控、治疗路径选择等关键任务的动态优化。此外，其模拟复杂过程中的因果推理能力，能够适配现实场景，弥补传统建模方式的局限，主要应用在分子合成路径规划、生物实验路径搜索和个性化治疗等领域。

图：强化学习应用实例

脓毒症治疗策略

- 在重症医疗领域，脓毒症的治疗策略一直是一个复杂的问题。研究人员利用深度强化学习方法，基于大型重症监护室数据库MIMIC-III和eICU，开发出一个AI临床医师模型。该模型通过分析患者的历史数据，学习并推荐最优的治疗策略。研究发现，当实际治疗决策与AI模型推荐一致时，患者的死亡率显著降低。

3 算力平台——规模化、高速的计算能力驱动算法落地

算力平台的差异化能力正构成AI4LS的基础设施构成：从宏观药物筛选，到结构预测，再到化学本质级判断形成完整算力支持链条。正是横跨“速度、精度、落地能力”的全景算力组合，让AI4LS不再只是概念突破，而是真正走入医学和药物开发的生产前沿。

图：算力赋能AI4LS全链条案例

类别	典型平台	AI4LS场景	成果
超级计算机	Summit	COVID-19超大规模虚拟筛选	48小时内完成8000多种化合物的模拟，锁定了77个潜在抑制剂
云HPC	Orion&AWS	基于云端的CADD分子设计平台	Beacon Discovery使用Orion验证蛋白质对接方法、形状和化学特征，从而识别出两种新型化学实体，并针对已知的 G 蛋白偶联受体 (GPCR) 靶点，找到 30 + 个有效匹配项
专用芯片	Wafer-Scale Engine	Cerebras CS-1 AI 系统	阿贡国家实验室自2020年以来一直使用 CS-1进行COVID-19研究和基于全球最大癌症治疗数据库的癌症肿瘤研究。在CS-1上运行的一系列用于预测癌症药物对肿瘤反应的模型，与GPU基线相比，实现了数百倍的速度提升
量子计算	Tencent Quantum Lab	量子算法驱动共价键的吉布斯自由能测定，以及模拟共价键相互作用	<ul style="list-style-type: none"> 前体药物活化：碳-碳键断裂前药策略探讨了一种应用于β-拉帕酮的创新前药活化方法，以实现癌症特异性靶向 共价抑制：KRAS是多种癌症中普遍存在的蛋白质靶点，采取共价抑制的方式使其蛋白无法发挥正常功能，从而达到治疗目的

I. 高性能计算集群:

集中式超级计算机为蛋白质折叠、分子对接与群体基因组学提供计算吞吐，推动生命科学任务从实验台走向规模化验证。大带宽互联与并行文件系统使得原始数据可在短时间完成读写，保证深度网络训练与大规模模拟的稳定供给。借助此类集群，结构生物学与虚拟筛选的工作周期被压缩至传统流程的十分之一以内。

来源：文献检索，公开信息，沙利文分析

2.3 技术矩阵的搭建

AI4LS在生命科学领域的落地依托于五大核心技术支柱的快速发展，从而形成“数据-算法-算力-知识-团队-新数据-反哺模型优化”的闭环科研生态

II. 分布式计算：

云端弹性HPC把“按需付费”的经济模式与超算级硬件结合，令中小型实验室和初创企业能够在高峰阶段瞬时启动万核集群、闲时自动释放资源。可编排的容器栈和自动合规审计让生物医药团队专注算法与实验，而将基础设施的运行维护外包给云平台。通过弹性调度，基因组组装、单细胞转录组分析与亿级分子对接等任务实现小时级周转，为算法快速迭代和多轮假设检验提供了灵活算力。

III. 专用芯片突破：

AI4LS模型主要依赖张量计算，这对算力平台提出了高效能的要求。相比传统通用CPU，专用芯片在能效、计算成本和运行延迟方面具有明显优势。配备的高带宽内存与低精度混合计算方式，使得单颗芯片的计算能力可达到百TFLOPS级别，大幅提升了在蛋白质复合物结构预测、代谢网络模拟等复杂任务中的处理效率。同时，专为AI优化的指令集和算法加速库，降低了模型部署的技术门槛，使得大规模模型可以在合理的功耗下扩展至上千颗芯片协同运行，从而实现AI4LS大模型的可持续落地和应用。

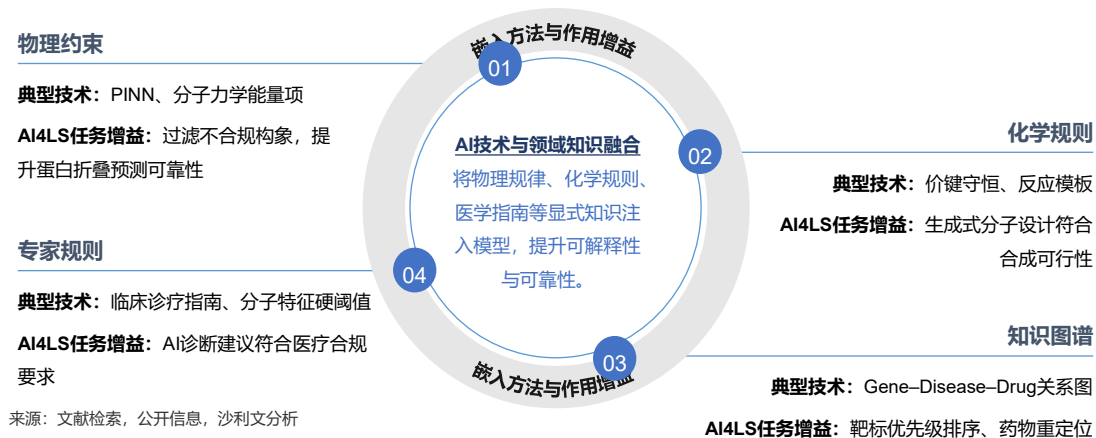
IV. 前沿量子计算：

量子加速器与类脑芯片为组合优化和电子结构计算提供新路径。尽管量子计算技术尚处早期，混合异构架构已经开始将量子计算嵌入虚拟筛选、反应路径搜索等尖峰任务，预示未来可能实现在能量评估和分子电子态解析方面的大幅提速。与此同时，类脑架构以事件驱动方式处理稀疏数据，为医疗影像与实时信号分析带来低功耗推理的新选择，进一步补强AI4LS端-云-超算一体化的算力版图。

4 领域知识嵌入——确保AI模型的输出符合科学规律

在AI4LS闭环中，数据、算法与算力共同塑造模型的外在性能，而领域知识提供的是科学可信度的内在保障。生命科学问题上受制于严谨的物理化学规律和复杂的生物网络，若模型仅凭统计相关性进行推断，容易出现“黑箱”结论或偏离机理的误判。将物理定律、化学规则与基因-疾病-通路等生物先验知识注入模型，可在训练初始为算法设置科学护栏，缩小搜索空间、提高收敛效率，并通过可解释路径增强对研究者和监管机构的透明度。因此，领域知识嵌入与数据资源、智能算法、算力平台、复合团队一起构成AI4LS技术矩阵的核心支柱，是连接已知科学与未知发现的关键桥梁。

图：领域知识嵌入内容与意义






2.3 技术矩阵的搭建

AI4LS在生命科学领域的落地依托于五大核心技术支柱的快速发展，从而形成“数据-算法-算力-知识-团队-新数据-反哺模型优化”的闭环科研生态

“知识嵌入”的方式，显著缩小模型的搜索空间，提高训练的收敛效率；同时也增强了模型的可解释性，使输出结果更具溯源能力，便于研究者理解与监管机构审查。尤其在小样本稀缺的真实场景中，领域知识作为“补充信息源”有效缓解了模型对数据量的过度依赖，维持预测的稳定性与泛化能力。此外，知识嵌入还促使模型生成的假设天然符合科学逻辑，具有较强的实验可验证性，可大幅降低实际验证的成本与试错风险。目前，已有多家AI4LS头部企业在实践中成功引入领域知识以优化模型性能。

图：应用领域知识提升模型准确性实例

 <h3>物理约束模型实践</h3> <p>望石智慧：提出并实践了基于物理约束的三级递进式评估体系。第一阶段通过QED、SAS等指标结合分子几何参数，筛除结构不合理结构；第二阶段通过结构基础药物设计和配体基础药物设计模式，真实还原新口袋结构下有效分子生成效果，解决“评分高但无价值”问题；第三阶段进一步考察模型在生成新骨架及新结合模式上的表现，衡量探索未知化学空间的能力。</p>	 <h3>生物先验知识实践</h3> <p>英矽智能：PandaOmics平台构建Gene-Disease-Pathway知识图谱，将文献证据、通路活性与大规模转录组数据整合为图神经网络输入。系统对纤维化相关基因进行优先级排序后，用三个月时间即可锁定并验证新的潜在靶点，其中靶点置信度排名前1%的候选占比达15%，显著提高了早期研发决策效率并减少盲目实验。</p>
 <h3>闭环迭代实践</h3> <p>予路乾行：AI-MD-实验数据闭环反馈机制通过计算预测与实验验证的迭代，持续优化AI模型和模拟精度。将AI加速MD模拟（构象集合、瞬态口袋）和AI驱动预测（变构位点）的结果作为假设提交至实验验证环节，随后将实验验证获得的真实结构数据和功能信息将反馈给AI模型，用于模型的再训练、参数优化和特征精炼，从而形成一个持续改进的“干湿”闭环研发流程。</p>	 <h3>生物机制约束模型</h3> <p>剂泰科技：AI-LNP平台模型受pKa、脂质尾部长度等关键参数约束，在体内药代动力学阶段，整合LNP组分与靶向性、内体逃逸等生物学行为的关联知识，预测其在多个器官内的递送效率，并进一步评估可放大生产性与安全性。该体系助力平台实现8个器官组织的靶向递送，其最快管线已进入pre-NDA阶段，验证了方法的有效性。</p>

5 复合型团队——支撑系统运行与交叉创新的关键

在AI4LS技术矩阵中，复合型人才是支撑系统运行与交叉创新的关键。生命科学与人工智能分属高度异构的学科体系：前者强调生物机制、实验逻辑与循证路径，后者注重算法建模、数据驱动与数学优化。两者的有效协同，既要求科研人员深入理解生物复杂性，也需算法工程师将模型精准嵌入科学问题语境中。因此，AI4LS发展亟需具备“生物+计算”双重素养的跨界人才。

目前，领先企业已在团队构建中普遍布局“双背景人才+协同合作”。一方面，积极招聘兼具生物医药背景与机器学习技能的专业人才组建跨学科团队；另一方面，推动以需求为导向的协作机制，使生物医药领域专家与数据科学家、AI工程师围绕共同目标密切配合。实践表明，药物研发团队引入AI专家后，能提升候选分子筛选与实验设计效率；而AI团队若有生物专家参与，则可校准模型假设，避免脱离实际。此类协作模式显著促进了AI技术向产业端的有效转化。复合型团队不仅是知识融合的组织形式，更是AI4LS技术协同落地的重要保障。未来，随着模型复杂度与场景精细度不断提高，生物、计算与工程三类人才的融合程度，将直接决定AI能否从辅助工具演进为生命科学领域的核心生产力。构建以目标为导向、跨学科能力为基础的高效团队，将成为推动AI4LS从实验室走向产业化应用的关键所在。

来源：文献检索，公开信息，沙利文分析

2.4 AI4LS赛道的资本共识

中国AI4LS投融资热度在经历阶段性调整后于2024年强势回升

AI在生命科学领域的投融资情况

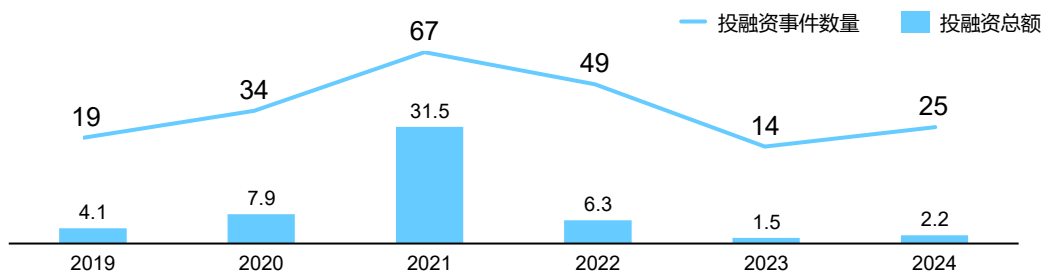
融资规模与节奏：高峰后回调，进入理性发展期

AI赋能生命科学，尤其新药研发的投资在近年经历热度的掀起。2019至2021年，该领域融资快速攀升，并在2021年创下新高：全国AI+药物研发相关创业公司全年共有67起融资事件，融资总额约31.5亿美元。然而到了2022年，热度开始降温：全年融资事件降至49起，金额约6.3亿美元，大幅低于2021年的水平。尽管融资总额下降，资本的偏好从广撒网转到重点培养，是行业走向健康和成熟的必经之路。

随着AI技术在生命科学领域取得更具突破性的成果，资本的目光再次被吸引。2024年投融资事件回升至25件，投融资总额较2023年提升44.8%。投融资轮次主要集中在早期阶段，A轮及之前的轮次数量为21起，占比52%，种子轮和天使轮共7起，占比28%。

图：中国AI制药领域一级市场投融资情况（2019-2024年）

单位：亿美元



AI4LS领域代表性企业分析

在资本逐渐回归理性、更加注重技术落地与产业协同的大背景下，AI4LS领域的投融资逻辑也正在发生深刻转变。相比早期对算法突破或概念创新的单点关注，当前资本更加青睐具备结构化能力的企业——它们往往拥有可验证的技术路线、可扩展的平台架构以及可对接产业需求的实际应用路径。资本押注的不再是“是否会做AI”，而是“能否将AI嵌入生命科学核心流程中并持续释放价值”。

在这一趋势下，资本尤其青睐具备以下特征的AI4LS企业：其一，技术壁垒明确，掌握难以复制的算法模型或独有数据资源，如世界领先的分子生成模型、蛋白质结构预测算法，或大规模生物组学数据库。技术壁垒不仅体现为研发实力，更代表潜在长期竞争优势；其二，应用场景清晰，能够将AI能力有效嵌入靶点发现、药物筛选、临床试验等生命科学关键流程，并已有实际验证案例。这意味着技术与生命科学痛点之间的匹配度，是AI价值释放的第一入口；其三，商业化路径可行、市场空间可观，能够通过SaaS平台、联合开发、“AI+研发管线”双轮驱动等模式形成持续收益；其四，具备良好的产业协同能力，能够融入药企、CRO、云计算平台等产业链节点中，既可获得资源支持，也向投资人释放“被主流生态认可”的信号。例如与药企联合共研管线、与科技巨头共建平台、或进入国家级科研平台，都被视为积极信号，充分体现AI+生物的综合势能。

以下受资本青睐的企业，往往在算法、实验系统、数据资源或临床转化等关键环节中建立了系统性壁垒，并形成跨科研、产业、资本的闭环联动，具备更强的放大潜力与稳定成长性，从而获得资本支持，并成为AI4LS赛道中最具代表性的标杆样本。

来源：公司官网，公开信息，沙利文分析

2.4 AI4LS赛道的资本共识

资本重点聚焦具备技术落地能力与平台扩展价值的头部项目

图：AI4LS重点应用域投融资概览（按最新一轮公开融资）

应用场景	公司（中国）	2020	2021	2022	2023	2024	2025	业务模式
AI药物研发	百图生科		A轮			战略融资		生物计算驱动的生命科学平台
	望石智慧 StoneWise	A+轮	B2轮	B+轮				小分子创新药平台型服务+自研创新管线
		A轮	B1轮					
	予路乾行 Divamics			天使轮	Pre-A轮		A轮 战略融资	AI+分子动力学CRO服务+自研创新管线
	英矽智能 Insilico		C轮	D+轮 D轮 战略融资			E轮	生成式AI服务+自研转化
	剂泰科技 METiS TechBio	天使轮	A轮	B轮		C轮	D轮	
Pre-A轮								
智慧实验室 &生命科学 基础设施	镁伽科技 MegaRobo		B轮-B3轮	C轮			D轮	自主智能体及多智能体解决方案 智慧实验室
	君济生物 JunjiBio				天使轮	Pre-A轮		生物合成与化学合成技术
	品峰医疗 Pinnacles		A轮			A+轮	B轮	
			天使轮					
基因组学& 精准医疗	吉因加 GenePlus	B+轮	C轮					大规模测序数据的智能化分析与解读
	安诺优达 Annoroad			C+轮				AI+IVD
合成生物学 &生物制造	分子之心 MoleculeMind			天使轮	Pre-A轮	A轮		AI蛋白质设计平台
	百奥几何 BioGeometry			天使轮		Pre-A轮		AI驱动的大分子CRO服务
	途深智合 ToursunSynbio				种子轮	天使轮 天使+		基于自然语言交互的蛋白质大模型服务

说明：表格中**加粗斜体**字体表示该轮融资金额超过人民币1亿元

来源：公司官网，公开信息，沙利文分析

第三章节

场景应用解析和挑战应对策略

AI4LS广泛渗透至多样化的应用场景并发挥效应，在技术、科学融合、伦理与法律以及生态系统方面的挑战亟待应对

3.1 AI应用场景：药物研发

AI正重构药物研发各环节的技术路径与效率边界，不仅提升设计质量与开发成功率，也催生了以平台化、管线化与协同化为特征的全新产业生态

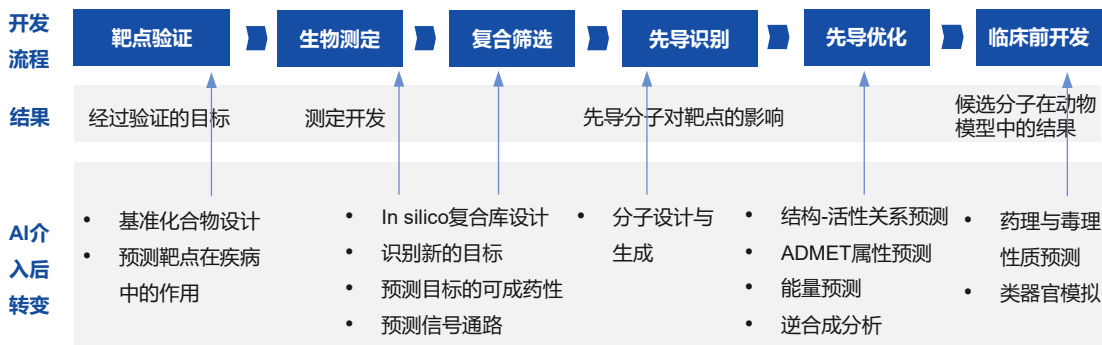
从底层建模、机制理解到结构优化与风险预测，人工智能正日益成为新药研发体系中的核心基础能力。其根本优势不仅在于提高效率、压缩成本，更在于拓展了化学空间、丰富了设计思路，并通过数据驱动的方式提升药物设计的科学决策密度。AI能够在低样本、小数据、复杂非线性问题中构建出稳定可迁移的预测与生成模型，从而实现在多个阶段同时提升设计质量与成功概率，改变了药物研发长期依赖经验、线性推进的被动模式。

AI重塑药物研发核心环节

从制药的流程看，AI技术已经在多个环节找到了适合自己的应用场景，并发挥出巨大潜力。从靶点发现阶段到临床前验证阶段，AI技术渗透到早期药物研发的诸多环节。例如基准化合物设计、预测疾病靶点、预测信号通路、预测成药靶点、确认新靶点、硅化合物库设计、预测药物结构与活性的关系、预测ADMET性质等。现阶段AI在药物研发需要大数据分析和高通量测试的阶段优势最为明显。例如在根据海量文献筛选靶点、专利追踪、批量性质预测等方面，药企便可以通过使用人工智能相关技术降低大量制药成本。

面对创新药研发领域日趋激烈的同质化竞争，不断取得的突破性进展也给AI制药的未来带来了空前的想象力。人工智能技术与生物制药领域的结合将药物发现、临床前研究的时间缩短接近30-50%，百亿美金级别的化合物筛选成本也可以通过AI制药技术节省出来，带来的诸多底层创新无疑也展现出了巨大的价值。

图：AI在药物研发中的应用与作用



商业生态：平台化、管线化与协同化三种路径并行演进

当前AI药物研发领域已逐步形成三种典型的商业生态路径，代表不同企业战略定位与技术能力边界。第一类是平台型企业，构建通用性AI模型，通过SaaS或定制化服务赋能传统药企与科研机构，强调技术的可扩展性与服务广度；第二类是管线型企业，其AI平台不仅参与候选分子的生成与筛选，更直接驱动项目向临床转化，是目前唯一可验证“AI药物闭环可行性”的路径；第三类是以协同型路径，通过与平台企业合作，引入AI辅助工具加速内部研发体系数字化转型。

三类路径构成互为补充的生态网络，平台型企业提供底层算力与建模框架，管线型企业实现场景验证与技术落地，协同型企业则以规模化管线驱动产业联通，形成“技术—应用—场景”三层联动结构，共同推动AI在药物研发全流程中的融合升级。

来源：CBInsights，文献检索，沙利文分析

3.1.1 靶点的发现与验证——靶蛋白建模与解析

AI通过深度学习模型赋能蛋白质结构建模与解析任务，不仅提升结构预测效率与准确性，也为新药靶点发现与分子设计提供了更高维度的结构基础支持

靶蛋白建模与解析的科学内涵与挑战

在药物研发早期阶段中，靶蛋白建模是实现结构基础药物设计（SBDD）的前提，其核心目标在于预测蛋白质的三维结构，进而揭示其空间构象、功能区域及其与小分子或其他生物大分子之间的相互作用机制。然而，在传统研究中，蛋白质结构的解析主要依赖于X射线晶体学、核磁共振（NMR）与冷冻电镜（Cryo-EM）等实验手段，普遍面临周期长、成本高、实验条件苛刻、蛋白结晶困难等技术瓶颈，难以满足高通量、快速推进的现代药物发现需求，尤其是对于缺乏同源结构信息的新型靶点或膜蛋白等难解析蛋白。

AI助力构建高效、准确的靶蛋白结构建模技术

AI不再仅仅作为传统建模工具的辅助补充，而是逐步成为主导性建模方。通过引入深度神经网络、图神经网络（GNN）、变换器结构（Transformer）及蛋白质语言模型等先进技术，AI驱动下的蛋白质结构建模能够对大规模蛋白序列与结构数据的深度学习，建立从序列到空间构象的预测模型，从而精准捕捉蛋白质内部残基之间的复杂空间关系，实现对三级结构的快速、高通量预测。目前，这一技术真正逐渐从技术试验走向系统应用，拓展了人类对蛋白质结构空间的认知边界。

AI在靶蛋白建模与解析中的核心能力

人工智能在靶蛋白建模与解析中的核心能力，体现为其对结构预测、构象注释和功能解析等多个关键环节的深度赋能。得益于深度学习模型的高效表达能力，AI不仅显著提升了三维结构建模的效率与精度，也突破了传统方法在结构不完整、数据稀缺或构象复杂等情境下的技术瓶颈。尤其在结构注释、结合位点识别、复合物建模等方面，AI能够基于序列、结构或组学信息实现快速准确的预测，为蛋白质功能理解和靶点机制研究提供全流程支撑。

图：AI在靶蛋白建模与解析中的任务覆盖范围

任务	AI赋能过程	所需输入	输出结果
蛋白质结构预测	根据序列预测蛋白三级结构	氨基酸序列	3D结构坐标（PDB格式）
二级结构与功能位点识别	识别α螺旋、β折叠与活性/结合位点	氨基酸序列、MSA数据	结构注释信息
同源建模与结构补全	推测未知结构的区域或残基	已知模板结构与序列比对	完整结构模型
可结合性分析	推测配体结合口袋、柔性构象	蛋白质3D结构	结合位点热图、可视化
蛋白-蛋白相互作用预测	预测复合物构象与界面	多个蛋白结构或序列	蛋白复合物结构预测

来源：文献检索，沙利文分析

3.1.1 靶点的发现与验证——结合位点确证

AI通过结合结构识别、空间建模与动态验证等手段，构建出从识别潜在结合口袋到预测配体适配性的全流程路径，显著提升了结合位点确证的效率与可靠性

结合位点确证的科学内涵与挑战

结合位点是蛋白质表面或内部能与配体发生高亲和力、特异性相互作用的三维功能区域，其结合能力源于多尺度特征的协同作用。从空间结构看，这些位置通常呈现出凹陷或口袋状结构，其容积和几何封闭性为配体提供物理容纳空间，并通过局部环绕的构象辅助分子定位。在化学特性层面，结合位点通过电荷互补、极性/疏水性匹配以及残基特异性排列形成功能微环境。

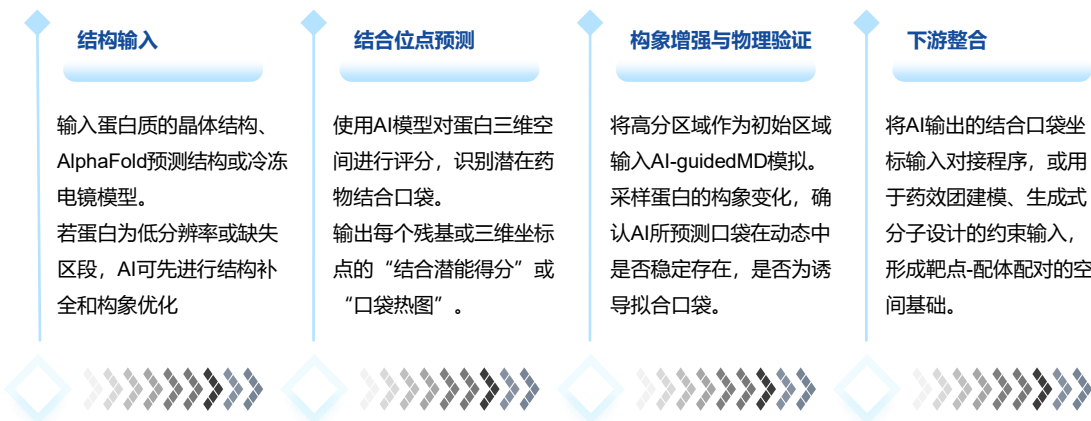
值得注意的是，部分结合位点具有动态特征——通过构象柔性或局部结构涨落形成瞬时暴露的“隐秘口袋”（cryptic pocket），或在配体结合时通过诱导契合机制调整构象。这些理化要素并非孤立存在，而是通过三维空间排布与电子云分布的动态耦合，共同塑造结合位点的特异性和功能可塑性，这一特性在药物靶点识别与分子设计中尤为重要。

在AI技术兴起前，结合位点的识别主要依赖于三类传统方法。基于几何算法的工具，通过计算蛋白质表面网格的曲率、凹陷深度及空腔体积，筛选出符合拓扑特征的口袋结构，但其仅关注空间形状而忽视电荷分布、静电势场和构象柔性等化学特性，易遗漏浅层结合位点或依赖动态暴露的隐秘口袋。实验结构解析法则基于配体-蛋白共晶结构直接定位结合位点，并辅以定点突变实验验证关键残基的功能贡献，但该方法依赖高分辨率结构数据且实验周期长、成本高昂，难以系统性拓展至未知靶标。分子动力学模拟虽然可以通过长时间尺度采样蛋白质构象变化，捕捉瞬时开放的口袋结构，但计算资源消耗巨大且缺乏目标导向性，无法高效完成大规模筛选。

AI的引入为结合位点的预测与确证提供了全新的技术路径

人工识别往往依赖静态晶体结构，通过肉眼观察凹陷区域或借助启发式算法判断潜在结合口袋。而AI技术在结合位点确证中的应用，构建的是从结构输入到机制验证、再到药物设计对接的闭环路径。该过程通常以蛋白质的三维结构为起点，对潜在结合口袋进行AI驱动的评分与定位，输出结合潜能热图或可信残基坐标。在此基础上，模型可引导分子动力学模拟，对预测位点的动态稳定性和物理可行性进行验证。最终，AI识别的结合坐标被用于药效团构建、配体对接或分子生成，成为结构驱动药物设计的重要空间依据。该路径实现了从识别到验证、从预测到设计的融合，显著提升了结合位点研究的准确性与实用性。

图：AI嵌入结合位点确证流程



来源：文献检索，沙利文分析

3.1.2 苗头化合物发现

AI通过虚拟筛选、分子生成、多任务预测等多种技术手段，重构了苗头化合物发现的筛选与评估路径，显著提升了早期药物候选结构的命中率与创新性

苗头化合物发现的科学内涵与挑战

苗头化合物指在早期药物研发阶段表现出初步生物活性、可进一步优化的分子实体。在苗头化合物阶段，从数百万个分子中初步筛出的“种子”，核心目标是确认化合物与靶标的初步相互作用及活性，且苗头化合物需满足以下关键标准：靶点结合强度、生物活性验证和理化性质基础。传统研发路径存在“低命中率、高成本、慢迭代”的瓶颈，苗头化合物的发现高度依赖高通量筛选、随机组合库设计与反复的实验验证，效率低下、资源消耗大，且命中率不足1%。

AI介入有助于发掘具有成药潜力的新颖结构

针对候选化合物获取难题，AI借助深度学习、生成建模与多模态数据集成等手段，使药物设计更具系统性、可预测性与自动化特征。尤其在苗头化合物的发现阶段，AI不仅能够介入虚拟筛选、活性预测及复杂表型筛查等关键环节，更可通过生成式建模实现分子的从头生成，直接产出具有新颖骨架与多样性特征的候选分子。由此，AI工具不仅具备理解结构、融合特征与生成优化的能力，还能在大规模分子空间中完成全新结构的创造、组合构建与筛选优选，为新结构、新机制的first-in-class候选药物提供强有力的技术支持。这些底层原理最终体现在AI系统对关键任务的具体完成能力上，贯穿于苗头化合物发现的各个环节。

图：AI在苗头化合物发现中的主要应用场景

1 虚拟筛选与打分预测 (Virtual Screening)

- 方法：AI模型学习已知活性分子与靶点之间的关系，通过预测结合能、构象适配性等指标，对虚拟化合物库进行高通量筛选。
- 技术：图神经网络 (GNN)、3D卷积网络、深度打分函数 (DeepDocking)

2 分子生成与反向设计 (De Novo Drug Design)

- 方法：基于靶点结构或结合口袋特征，使用生成模型设计可结合的小分子，进行药效团驱动的定向合成。
- 技术：生成对抗网络 (GAN)、Transformer架构、变分自编码器 (VAE)、扩散模型、强化学习

3 药效预测与性质评估 (Activity&Property Prediction)

- 方法：AI对候选分子的活性、选择性、毒性、溶解性等进行多任务预测，筛除早期失败分子。
- 技术：多任务学习、QSAR深度模型、Transformer架构

4 多靶点与表型筛选 (Polypharmacology & Phenotypic Screening)

- 方法：AI用于分析分子对多个靶点或整体现象的响应，通过表型图像、转录组数据等，挖掘非典型作用机制。
- 技术：图像识别网络 (CNN)、多模态融合模型 (如Graph+Omics)

来源：文献检索，沙利文分析

3.1.3 先导化合物优化

AI通过结构预测、性质建模与多目标优化等手段，实现对候选分子的活性增强与理性迭代，显著提升了药物优化阶段的效率、准确性与创新能力

先导化合物优化的科学内涵与挑战

先导化合物优化是指在药物研发早期阶段，通过对初始命中化合物进行结构修饰与理性设计，提升其与靶点之间的结合亲和力和选择性，从而获得具备临床开发潜力的先导化合物。这一过程，实质上是一种多目标协同的分子工程过程。不仅关注分子与靶点的稳定相互作用构建，还涉及对构象控制、官能团调节以及结构-活性关系（SAR）的系统挖掘。同时，活性优化需兼顾成药性参数，如溶解性、代谢稳定性与毒性风险。

传统的化合物优化过程主要依赖化学家经验和逐步试错的结构修饰策略，通常需要大量分子合成与生物测试才能逐步提升活性。这种方法效率低、周期长，难以系统地处理结构-活性关系中的非线性与多目标权衡，尤其在面对复杂靶点、诱导契合构象或脱靶风险控制等问题时显得力不从心。此外，传统手段难以整合大规模数据资源用于优化决策，导致候选分子的优化路径缺乏系统性和预测力。

AI提升药物活性优化流程的效率与科学性

AI能够在大规模化合物空间中快速识别高潜力结构，并基于靶点构象进行精准适配性设计，实现对分子亲和力与选择性的同步提升。相比传统依赖经验和试错的方式，AI所具备出色的非线性建模与模式识别能力，能够挖掘隐藏在结构背后的复杂构效规律。同时，通过集成多任务学习和生成建模，支持对活性、溶解性、代谢稳定性等多个关键指标的协同优化，使得分子结构能够在多维参数中达到更优平衡。更重要的是，AI推动了从静态筛选向动态学习的转变，使分子设计过程具备自我调整与智能进化的能力，从而显著加快了从命中到候选的优化迭代速度。

AI在药物活性优化中的主要应用场景

AI在先导化合物优化阶段的主要作用贯穿于候选分子从初步命中（hit）向高活性先导（lead）进化的整个过程，核心目标是提升分子的靶点结合力、选择性与机制稳定性，同时兼顾其他关键属性如毒性、溶解度等。在这些应用场景之下，AI技术的引入在效率和规模上实现了量级提升，更通过生成模型、深度学习打分网络、多任务预测等手段，有效突破了传统手段在构象探索、精度控制和属性平衡方面的核心瓶颈。

图：AI在先导化合物优化中的主要介入路径

分子结构-活性关系建模（QSAR）



AI可通过监督学习或深度学习方式，构建分子结构与生物活性之间的非线性映射模型。相比传统QSAR模型，AI具有更强的表达能力和更高的泛化性能。

结合构象预测与结合自由能估算



AI可以辅助进行蛋白与小分子的对接分析，不仅预测结合位点与结合姿势，还可进一步评估热力学稳定性。

结构生成功能驱动优化



AI生成模型可根据已知活性片段，在保持活性基团的前提下，通过侧链优化、骨架跳跃、或新颖环结构生成，产生活性更高的新分子。

多靶点选择性与脱靶分析



AI可同时预测候选分子对多个靶点的亲和力，优化其靶点选择性。例如在抗肿瘤、精神疾病等多靶点适应症中，AI可避免脱靶效应导致毒性反应，提升药效安全性。

来源：文献检索，沙利文分析

3.1.4 成药性优化

AI通过多任务预测、结构评估与反馈优化机制，有效赋能成药性多维参数的系统建模与优化选择，显著提升了分子向“可用于人体”转化过程的成功率与确定性

成药性优化的科学内涵与挑战

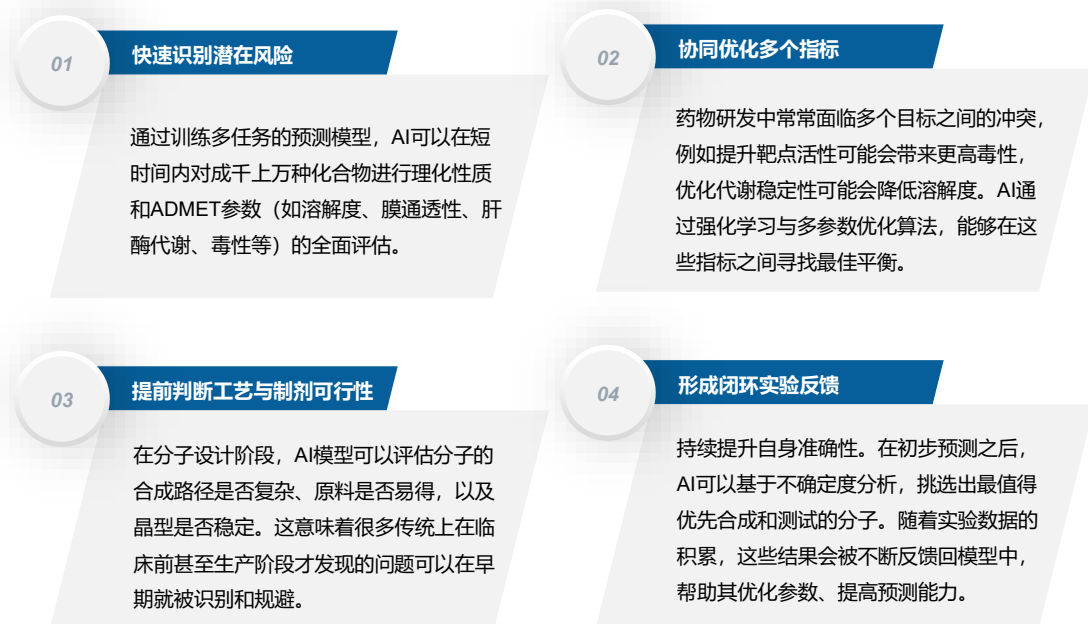
成药性优化作为新药研发从“具有活性”迈向“可用于人体”的关键环节，其科学内涵远不止于ADMET（Absorption、Distribution、Metabolism、Excretion）筛选，而是一项融合理化性质、生物代谢、安全边界与制剂可行性的多学科挑战。从分子层面看，成药性优化首先要求候选化合物在极端生理环境下具备良好的稳定性与可吸收性，还要求分子具备合理的代谢路径和低毒性。此外，成药性还应包括合成路线的可行性与制剂的工业适配性，如晶型稳定性、手性控制和合成步骤控制在合理范围内，以支持后续的大规模生产和CMC注册策略。

然而，这一过程面临诸多科学与工程挑战。首先，活性、溶解度、膜透过性等属性间往往存在相互冲突。其次，许多早期评估指标在体外和体内之间并不总是线性可转化，跨尺度预测难度高。此外，高质量的ADMET标签本身数据稀疏且存在样本偏倚，导致传统建模方法在外推阶段表现不稳定，难以在临床前早期有效筛除高风险候选。更难预警的是诸如代谢产物毒性、免疫性肝毒等特定类型毒性，其机制复杂、实验捕获难度大，对模型提出了可解释性和机制建模的更高要求。最后，制剂工艺相关问题如晶型多态性、手性选择性和低收率合成路径，常在后期CMC阶段才暴露，导致前期优化成果被推翻。

AI在成药性优化中的核心任务

在成药性优化环节，AI主要承担四项互相衔接的核心任务：快速识别潜在风险、协同优化多个指标、提前判断工艺与制剂可行性、形成闭环实验反馈。这四项任务的协同将原本串行、耗时的成药性评估压缩为高通量、闭环迭代的智能流程。相较传统方法的单点判定或线性筛选，AI能够同时评估多个成药性维度并挖掘复杂的非线性关联，从而在多参数空间中实现分子结构的全局最优修饰。

图：AI在成药性优化中的技术路径与机理



来源：文献检索，沙利文分析

3.1.5 临床前验证

AI通过系统性建模手段替代部分传统实验流程，实现对毒性、PK/PD参数与跨物种转化等关键指标的高效预测与风险评估，显著提升了临床前验证阶段的效率与可靠性

临床前验证的科学内涵与挑战

临床前验证是药物研发流程中承上启下的关键阶段，其核心任务是系统评估候选化合物的有效性、安全性、代谢特性以及作用机制，判断其是否具备进入临床试验的科学依据。这一阶段通常包括体外活性测试、动物药效模型验证、毒理学评估、药代动力学（PK）与药效动力学（PD）研究，以及潜在风险因素的识别和机制层面的确认。与早期发现阶段相比，临床前验证更强调数据的可重复性、外推性和监管合规性，是确保后续临床开发成功率的“决策闸口”。由于涉及实验系统复杂、数据维度高、时间成本大，该阶段也成为AI技术介入的高价值场景之一。

从传统验证模式向AI驱动验证转型的推动力，源于临床前阶段长期存在的效率瓶颈与科学限制。传统方法依赖大量体内外实验，操作复杂、成本高昂，且动物实验结果在人类中的可外推性有限，常导致候选分子在临床试验阶段因毒性或疗效问题而失败。此外，面对新型药物结构、生物靶点复杂性与多维数据爆炸式增长，传统线性分析手段难以处理异质性强、非线性关联显著的验证任务。AI的引入恰恰契合了当前对高通量预测、机制建模与跨模态整合的迫切需求。其在毒性预测、PK模拟、动物-人类转化建模等方面的能力，提升了验证环节的科学与前瞻性，正在推动临床前研究迈向更精准、更智能的技术范式。

AI的介入推动临床前研究迈向更精准、更智能的技术范式

AI技术正通过深度学习、图神经网络、多任务建模等手段，构建可替代或补强实验的预测模型体系。原本依赖体外实验的hERG通道抑制检测，可被AI构建的毒性预测模型所替代；动物PK实验则正逐步由AI增强的PBPK模型承接；晶型筛选与稳定性实验也可通过AI与物理机制耦合的晶体结构预测模型完成。此外，针对动物毒性与跨物种转化等任务，AI正通过多任务学习和多模态因果图谱等方法，提高风险评估的准确性与可解释性。这种由实验流程向模型体系的系统性转化，不仅提高了研发效率，也为更早期的机制洞察和风险预测提供了支持。

图：AI将实验流程转化为建模系统

传统实验任务	AI建模系统形式
hERG通道抑制实验	深度毒性预测模型（GNN/MLP）
动物PK实验	AI增强PBPK模型
晶型筛选与剂型稳定性实验	AI+物理建模（AI-CrystalNet）
动物毒性实验	多任务毒性预测模型
跨物种转化验证	多模态迁移模型+因果图谱

AI在临床前验证中的主要应用场景

AI在临床前验证中的主要应用场景，集中体现在对药物候选分子的安全性评估、药代动力学建模、毒性预测以及跨物种转化能力判断等关键环节。传统上，这些任务依赖大量体外试验和动物实验，不仅耗时高、成本大，还存在伦理争议和跨物种差异带来的预测误差。AI的引入使得研究人员能够基于分子结构、生物通路、组学数据等多源信息，构建高精度的预测模型，实现对hERG阻断、肝毒性、晶型稳定性、代谢路径等关键指标的模拟判断。不同任务所采用的模型路径具有显著的异构性，既包括图神经网络与Transformer等结构识别模型，也涵盖PBPK增强建模、多模态迁移学习与知识图谱推理等复杂系统建构技术。既是传统实验的数字替代，更通过机制层级的建模方式，将风险评估从结果导向转向过程可解释，显著提高了验证的前瞻性、稳定性与临床转化率。

来源：文献检索，沙利文分析

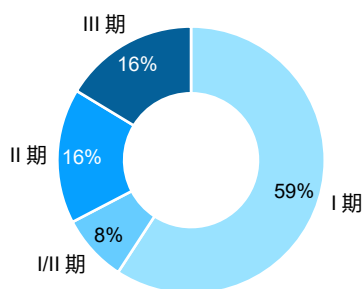
3.1.6 AI技术企业与药物管线

AI药物管线正由早期探索集中迈向验证与确证阶段，关键代表项目已在多个靶点上体现出从机制创新到临床推进的可行性与速度优势

AI开发药物管线在不同临床阶段的结构特征

目前，国内应用AI技术开发的药物管线按临床阶段划分：Phase I（含I/II）占67%，Phase II与Phase III各占16%。AI对药物研发价值的体现呈现明显的阶段梯度：Phase I验证算法潜力，Phase II证明临床价值，Phase III兑现商业价值。目前国内管线已经迈入首批AI优化重磅分子待上市的临界期。部分项目正进入注册决策的窗口期，也意味着AI从技术驱动逐步走向产业兑现的转折点。在这一阶段，差距将不再体现于算法形式本身，而更多体现为模型能否对准真实需求、高效嵌入决策流程。与此同时，全球监管机构对算法透明度与可追溯性的要求日趋严格，这促使AI企业不仅要跑得快，更要跑得稳。因此，企业的胜负关键在于能否在已验证靶点上用AI做出速度和质量的双重领先，并在早期阶段就构建起可审计的算法证据链，以迎接日趋严格的全球监管。

图：中国应用AI技术的药物管线（按临床阶段划分）



注：统计截至2025.05.28

AI开发药物管线在不同疾病领域的结构特征

国内AI赋能管线目前呈现出肿瘤项目登记最多，却有四分之三仍停留在I期，反映出AI主要用于难成药靶点探索但仍需跨越临床验证门槛；代谢-内分泌领域虽登记略少，却已有一半进至II-III期，可见生成式蛋白语言模型与肽环稳定化算法，已成功将多靶点长效激动剂从设计端快速推向确证性研究；免疫系统疾病同样保持50%的中后期占比；皮肤类疾病的项目数虽不多，却拥有最高的后期转化率，凸显已知靶点与AI优化结合的模式商业确定性；罕见病则因样本稀缺与机制复杂，多停留在I期机制验证阶段。整体来看，登记数量自肿瘤向慢病递减，而II/III期占比却递增，提示企业应先在疗效评价标准已有充分可借鉴范式、投资进入风险小的代谢或免疫赛道实现商业兑现，再将技术和资金反哺至高风险的肿瘤与罕见病领域，以实现AI药物价值的全周期放大。

从国内AI药物管线的适应症与疾病领域分布可以看出，高度公开且多模态数据资源丰富的疾病（如肿瘤、代谢和免疫类）天然吸引更多企业聚集，呈现“数据可得性越高、管线越拥挤”的正相关。在这些热点赛道中，GLP-1/GIP、TYK2/JAK1等经验证靶点被不断平移，从肥胖到非酒精性脂肪性肝炎、从银屑病到溃疡性结肠炎等相邻适应症，显示出靶点组件跨疾病快速复制的趋势。相比之下，慢病领域更强调通过AI协调多靶点或双功能分子的药效、安全性和给药频率，以“一分子解决多合并症”来凸显差异化，而罕见病虽项目稀少，却集中体现全新靶点与结构的高新颖度，成为技术探索的试验田。值得注意的是，即使实体瘤项目数量庞大，它们也并未集中于某一器官，而是在肺、胰、乳腺等多种组织中分散，说明AI选靶更多基于分子-途径可成药性而非传统器官分类。总体来看，管线价值分布呈现一端为患者基数庞大的高发慢病，另一端为价格天花板高企的罕见病的双峰格局，中间地带的常规适应症反而相对稀疏，反映了企业在市场规模与高溢价孤儿药这两种商业模型之间进行资源倾斜与策略布局的权衡。

来源：Insights数据库，公开信息，沙利文分析

3.1.6 AI技术企业与药物管线

国内AI赋能药物管线呈现肥胖与代谢病领跑、炎症免疫与实体瘤并进的集聚格局，多数项目仍处早期但已显跨靶点、多路径的结构创新潜力

具体适应症——体现人工智能价值的集结点

第一类集中在实体瘤及特定亚型，结构预测加分子动力学方法让一批过去被视为成药性低的口袋得到快速命中。Relay Therapeutics的RLY-2608由其自研AI药物发现平台Dynamo促成，这是首个已知的变构、泛突变型和异构体选择性PI3K α 抑制剂。2024年公布最新中期临床数据显示，RLY-2608联合氟维司群在HR+、HER2-局部晚期或转移性乳腺癌患者中达到39%的确认客观缓解率（ORR）和11.4个月的中位无进展生存期（PFS）。这一试验结果充分验证了AI在传统难成药靶点转化上的实际临床价值。

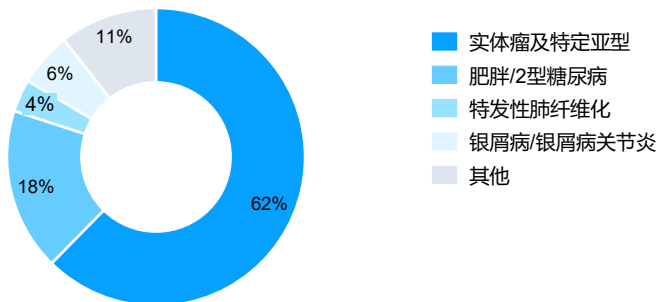
第二类是肥胖和二型糖尿病，这一方向因可量化的体重或血糖终点而成为生成式蛋白模型和肽环稳定化算法的首选试验场，双靶点乃至三靶点的长效多肽已经以极快速度进入确证性临床。BrightGene与予路乾行合作开发的BGM0504通过生成式蛋白语言模型与分子动力学联用，仅保留210条候选肽序列就锁定了双GLP-1R/GIPR激动剂靶点；其中国IIb期数据已证实体重下降幅度优于司美格鲁肽。

第三类是斑块型银屑病及其关节病变，深度学习帮助识别TYK2与JAK结构中的水分子网络，从而显著提高选择性并减少毒副反应，配合影像自动评分技术，相关药物的入组效率和疗效评估速度远超传统方案。Zasocitinib开发的TAK-279分子利用深度学习分析TYK2激酶JH2结构中的“水网”后，设计出全靠疏水插栓而非极性氢键的别构抑制剂，离体选择性比JAK1/3高50到100倍；IIb期的12周数据显示，PASI75达67%到68%（30mg、15mg组），明显优于同类TYK2抑制剂。

第四类是属于纤维化类慢病的特发性肺纤维化，算法能够把病理影像和转录组图谱结合起来，提供连续量化的数字病理读数，使得在中期试验就能捕捉到微观结构的缓解信号。英矽智能的INS018-055是一款通过跨组学知识图谱发现的TNIK抑制剂，项目自靶点提出到临床用时18个月，远低于领域平均耗时；正在中国和美国同步推进的IIa期试验已看到用力肺活量下降速率减缓信号，为国内罕见病赛道提供了首个真正“AI-first”的小分子样板。

图：中国应用AI技术的药物管线（按适应症划分）

注：统计截至2025.05.28；统计口径：纳入的在研管线包括处于临床前阶段的产品



来源：Insights数据库，沙利文分析

3.1.6 AI技术企业与药物管线

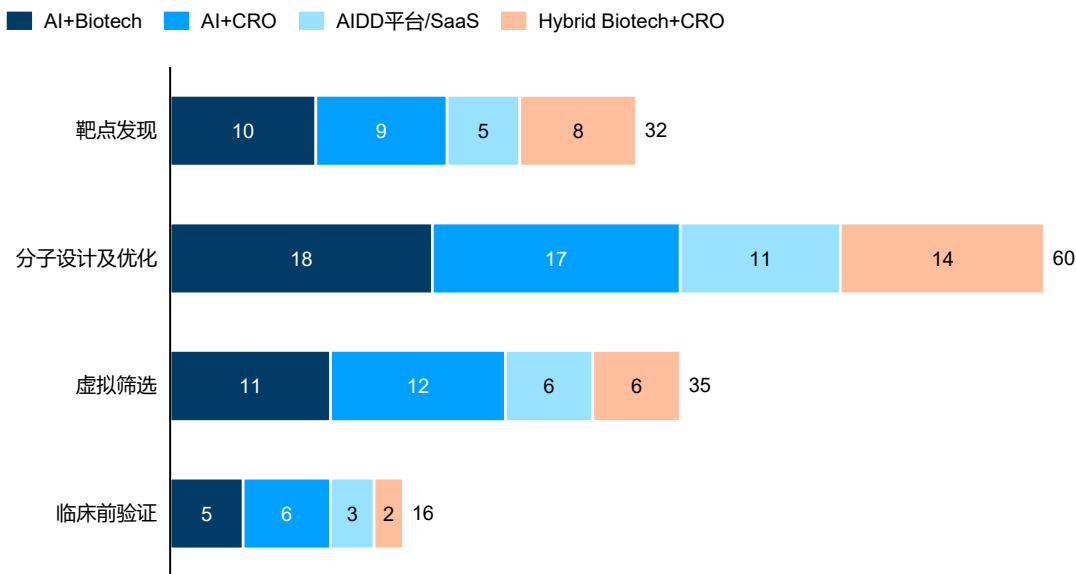
中国AI药物研发企业在分工结构、平台建设与场景选择上逐步形成多元路径，呈现出从功能垂直到复合协同、从建模工具到验证闭环的生态演进趋势

中国AI药物研发企业在不同研发环节的结构特征

中国AI药物研发企业在各研发环节呈现出鲜明的层级结构。越往下游走，计算与实验深度耦合的要求越高，导致创业门槛和痛点迅速抬升，也让跨环节、复杂商业模式更受资本与客户青睐。相反，在上游环节，平台与数据网络效应仍是核心竞争焦点。整体来看，企业数量在靶点发现和分子设计阶段最多，向临床前验证端骤减，形成典型的倒V型分布，说明大量团队停留在纯计算层面而难以承担实验设施和GMP投入。

与此同时，同一家企业常以不同业态在多环节频繁露面，Biotech的技术重心覆盖药物发现全流程平台以把控知识产权，而AI-CRO则在下游布局ADMET预测以延伸客单价价值链。此外，SaaS平台企业因数据维度与算法迭代需求，将产品化重点放在靶点发现等早期环节，进入分子筛选及更后期阶段，平台化企业比例随之明显下滑。

图：中国应用AI技术的药物研发企业（不完全统计）



来源：Insights数据库，沙利文分析

3.1.7 药物研发场景的代表性AI企业

不同类型企业围绕平台构建、模型驱动与落地能力展开多元探索，代表性公司通过差异化技术路径和应用模式，正在推动AI从工具向赋能主体的跃迁

AI+Biotech：AI驱动的药物研发企业，聚焦药物发现与开发全流程，利用AI技术突破传统研发瓶颈。

英矽智能成立于2014年，定位为“AI驱动的端到端药物研发公司”。公司搭建Pharma.AI平台，平台形成从靶点识别到候选化合物生成、临床可行性评估的闭环系统，具备跨模态数据集成、端到端分子优化及临床路径预测能力，覆盖靶点发现（PandaOmics）、分子设计（Chemistry42）和临床前评估（inClinico）等研发流程。

- ISM001-055（TNIK抑制剂）由平台端到端设计，仅用18个月完成药物分子发现全流程。目前针对特发性肺纤维化（IPF）适应症的临床IIa期已取得积极试验数据。



AI+SaaS/技术平台：AI驱动的垂直领域工具平台，提供标准化AI工具或数据服务，降低行业技术使用门槛。

百图生科成立于2020年，通过打造跨模态大模型为药物研发提供标准化AI基础设施，建立了在序列理解、结构建模与自动生成上的多模型协同机制。公司推出的xTrimoV3模型拥有2,100亿参数，是目前生命科学领域参数量最大的基础模型之一，覆盖蛋白质、RNA、小分子、医学影像与文献等模态。模型具备强泛化能力与跨任务适配能力，已被封装为多个API模块，广泛应用于靶点发现、分子设计、抗体建模与药效预测等任务。

- 2023年10月，百图生科与赛诺菲达成战略合作，基于百图生科的生命科学大模型（xTrimo），联合开发用于生物治疗药物发现的AI模型，重点优化大分子药物的设计与多参数优化。赛诺菲支付1000万美元预付款，并根据模型开发、临床前及商业化里程碑支付后续款项，总交易金额超10亿美元。



AI+CRO：AI赋能的合同研究组织，将AI技术嵌入传统CRO服务，提升研发效率与数据价值。

水木未来成立于2019年，是一家聚焦结构生物学与AI药物设计融合的科技型CRO企业。公司依托冷冻电镜（Cryo-EM）平台、图神经网络与生成模型算法，构建结构解析与AI设计一体化能力，为全球药企提供从蛋白结构解析到先导分子筛选的全流程服务。公司搭建的智能结构解析平台整合了300kV冷冻电镜与AI算法框架，支持对膜蛋白、蛋白复合体等高难度靶点进行高分辨率成像。

- 在北京、杭州建立全球冷冻电镜与AI药物创新中心，逐步打造高通量结构-AI设计协同平台，具备强区域辐射能力。



来源：公司官网，公开信息，沙利文分析

3.1.7 药物研发场景的代表性AI企业

不同类型企业围绕平台构建、模型驱动与落地能力展开多元探索，代表性公司通过差异化技术路径和应用模式，正在推动AI从工具向赋能主体的跃迁

AI+Hybrid：研发与服务双轮驱动的复合型企业，同时推进自研管线与对外技术服务。

望石智慧成立于2018年，是一家使用人工智能技术驱动新药研发的科技公司。凭借人工智能药物研发底层理论的突破、药物研发数据的深度治理、药物研发行业的沉淀认知和强大的软件与工程能力，望石搭建了能够精准地生成与靶点口袋结构契合的分子或分子骨架的多模态AI 3D分子生成大模型基座。基于多模态AI分子生成大模型，公司既推动自主管线研发，形成从早期分子生成到临床验证的闭环；也在此基础上，通过平台化服务和实体分子库建设为全球药企与科研机构提供高效的分子设计与优化工具。



- HPK1抑制剂SWA1211候选药物在临床前研究中展现出同类最佳的潜力，在与PD-1抗体类药物联用展现出更强的协同抗肿瘤疗效。2025年3月已获得中美双IND批件；2025年6月完成针对晚期实体瘤I期临床研究的首例患者给药。
- 公司的通过AI分子生成平台构建的化合物实体库服务已通过数十家客户已完成商业转化，例如贝勒医学院（Baylor College of Medicine），英国癌症研究中心、清华大学、中科院物理研究所、广州国家实验室、复旦大学医学院、齐鲁锐格等。

予路乾行是一家面向全球市场的以多尺度分子动力学与AI算法驱动的新药研发科技公司。公司首度提出“分子电影”的新概念和“Dynamics-based Drug Discovery”药物研发新范式，并依托AI技术整合海量文献、专利和临床数据库，打造了综合智能体Biotrajectory数据库，形成从靶点识别到临床前化合物的“高速、高效且高成功率”的全研发链条。自2021年成立以来，公司已参与全球80多条药物管线研发，在各疾病领域和多模式药物发现中积累了深厚的经验，临床前化合物的研发成功率超过70%。除为药企提供难成药靶点与高柔性体系下的全流程研发服务外，公司也充分利用自研平台全速推进代谢、自身免疫疾病等领域的自主研发管线。



- 作为全球分子设计与创新工场，公司与北京鞍石生物合作的小分子管线以及与苏州博瑞医药合作的多肽管线于2024年进入临床试验三期；目前，公司已在苏州、北京、新加坡和日本建立研发与BD网络，逐步打造国际化的AI+分子动力学药物研发生态。
- 公司自主研发的IL23/TL1A抑制剂在临床前研究中展现出同类最佳的潜力，已初步达成转让意向。

剂泰科技是一家人工智能（AI）驱动纳米材料创新的生物科技公司。公司自主开发全球首个人工智能驱动的纳米递送解决方案提供平台NanoForge，拥有目前全球最大规模千万级LNP脂质库，并基于NanoForge打造了三大核心解决方案：AiLNP（AI核酸递送系统设计平台）、AiRNA（AI mRNA序列设计平台）、AiTEM（AI小分子制剂设计平台）。



- 截止目前，剂泰科技拥有超过1,000万种以上的脂质结构，10万个可做模型训练的数据点，突破性成功开发人体内肝、肺、免疫器官、心脏、肌肉、肿瘤、中枢神经系统、胃肠道8个靶向器官或组织的LNP，获得授权和已备案专利申请共超过100项。
- 公司已成功开发超过10个管线项目，产出7个临床前候选药物，并行推进4个临床项目，最快的管线已到达pre-NDA阶段。

来源：公司官网，公开信息，沙利文分析

3.2.1 AI应用场景：基因组学与个性化医疗

AI通过深度学习与图模型手段重构了基因组学的理解与决策方式，有效弥补传统方法在结构识别、信号提取与精准匹配上的不足，推动个性化医疗迈向更加智能化与系统化的落地路径

随着测序技术的快速发展和成本的急剧下降，人类进入了“全基因组可及”的时代。每个人的基因信息逐步被数字化、结构化，为医学从群体平均疗法向个体精准治疗转型奠定了基础。与此同时，表观遗传修饰、单细胞分辨率、多模态生物组学等新兴技术相继成熟，为揭示个体差异、疾病机制及干预靶点提供了前所未有的深度维度。

基因组学与个性化医疗因此成为21世纪医学进步的核心方向之一，不仅能更早发现疾病风险，还能为每一位患者量身定制最合适的治疗方案。然而，这一宏伟愿景的实现并非易事，其依赖于对海量、复杂、生物意义不确定的数据进行高效、精准、系统的解读，这正是人工智能技术展现颠覆潜力的核心场域。

传统方法面临的困难

在AI介入之前，基因组数据的分析严重依赖统计模型、生物信息学脚本和大量人工注释。首先，面对全基因组尺度的数据，传统全基因组关联研究（GWAS）方法主要基于线性回归与显著性筛选，难以捕捉复杂的多基因交互关系，导致许多与疾病密切相关的隐性遗传因素被遗漏。其次，在单细胞测序中，由于高噪声、高稀疏性和批次效应的干扰，研究者往往需要耗费大量时间手动筛选细胞亚群，依赖既有标记基因的经验分类，限制新型细胞状态的发现。

而在个性化治疗中，整合不同组学、病理、药敏、影像等异构数据更是一大挑战，传统工具难以建立跨维度的精准模型，治疗推荐多依赖临床经验和指南，难以实现真正的因人而异。此外，CRISPR等基因编辑工具在靶点选择和脱靶预测上也面临精度不足与效率低下的问题。整体而言，传统方法存在三个主要障碍：维度过高而难以建模、信噪比过低导致误判、数据异构难以统一表达。

图：传统方法制约基因组学研究的因素



AI带来的变革与优势

人工智能，正在根本性地重塑基因组学的认知框架与实践路径。首先，AI通过非线性建模能力打破了传统统计方法的边界，能够从庞杂的单核苷酸多态性（SNP）与表型数据中挖掘复杂遗传交互关系，实现对疾病风险的高精度预测。其次，在单细胞组学中，AI可自动学习高维稀疏数据中的结构特征，实现无监督细胞聚类与轨迹识别，推动细胞异质性研究从“注解式”向“发现式”转型。

在精准医疗层面，AI可整合基因组、蛋白质组、临床记录等多源异构信息，构建个体化的药物反应模型与治疗推荐系统，打破传统模式中的逐步试验。在基因编辑方面，AI优化sgRNA设计，提高编辑效率与安全性，推动基因治疗临床转化。更重要的是，AI的自我学习与迭代能力，使其在不断积累数据中越用越准，形成以知识进化为基础的动态系统。因此，AI为实现真正意义上的个性化医疗提供了技术基础与路径。

来源：文献检索，沙利文分析

3.2.1 AI应用场景：基因组学与个性化医疗

AI赋能基因组学与个性化医疗的多个关键任务，涵盖从数据解析、风险预测到治疗决策与基因编辑的全流程，为精准医疗提供智能化与可扩展的解决方案

AI在基因组学与个性化医疗中的应用

基因组学与个性化医疗迈向“智能转化阶段”。AI不仅实现了从数据解析到临床决策的贯通式协同，还使得疾病机制解析更为精准、风险预测更为前瞻、治疗方案更具个体化特征。AI所展现的具体作用路径与实际成果，深入揭示其在生命科学体系中所引发的结构性变革与产业价值潜力。

图：AI的具体任务与功能



来源：文献检索，沙利文分析

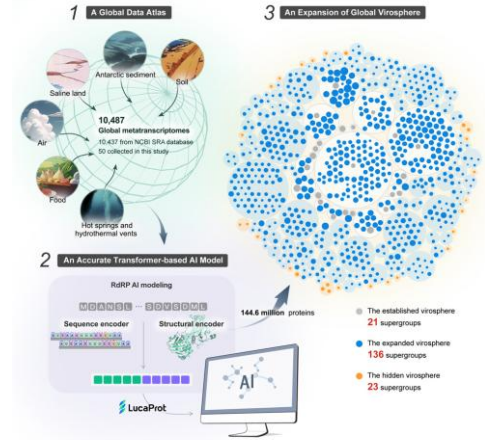
3.2.2 基因组学场景的代表性AI成果

AI技术正在深度赋能基因组学领域，显著提升大规模数据挖掘、复杂关系建模与生物学知识发现的效率，推动基础研究、算法创新与疾病诊疗深度融合

大模型+智能体：RNA病毒AI挖掘系统LucaProt

中山大学海洋科学学院、生命科学学院与阿里云等伙伴在2024年10月于《Cell》发表论文，提出深度学习算法LucaProt，通过结合RNA依赖性RNA聚合酶序列与预测结构特征，在10487份来自全球多生态系统的宏转录组数据中一次性识别出180个RNA病毒超群与超过16万种全新RNA病毒，使RNA病毒超群数量扩容约9倍，其增长种类是已知病毒的近30倍，并显著降低传统同源性比对遗漏“病毒暗物质”的风险。团队报告该模型假阳性率仅0.014%，假阴性率1.72%，并可追溯病毒进化脉络，为早期疫情监测和新型疫苗设计提供了高维度信息图谱。

图：LucaPort图解

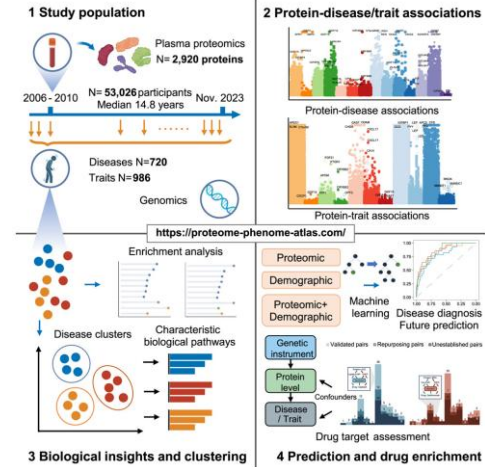


图源：Hou X, He Y, Fang P, Mei SQ, Xu Z, Wu WC, et al. Using artificial intelligence to document the hidden RNA virosphere. Cell [Internet]. 2024 Oct 1 [cited 2024 Oct 14];187:6929–42. Available from: <https://pubmed.ncbi.nlm.nih.gov/39389057/>

多组学图谱工程：复旦大学人类健康与疾病蛋白质组图谱项目

复旦大学附属华山医院与类脑智能科学与技术研究院团队历时五年，结合53026份中国成人血浆样本的深度质谱数据与人工智能建模方法，绘制了全球首个人类健康与疾病血浆蛋白质组图谱，并于2024年11月以封面论文发表于《Cell》杂志。图谱涵盖2920种蛋白、406种既往疾病、660种新发疾病和986个健康表型，累计揭示168100条蛋白质-疾病关联与554488条蛋白-表型关联，使研究者能够在一次抽血中预判数百类疾病风险。通过深度学习模型，研究团队为183种疾病建立了高精度早筛框架，并锁定26个具有转化潜力的新药靶点。

图：蛋白质组图谱图解



图源：Deng YT, You J, He Y, Zhang Y, Li HY, Wu XR, et al. Atlas of the plasma proteome in health and disease in 53,026 adults. Cell. 2024 Nov 1;188:253–71.

“数据-AI-临床”闭环：康圣环球与BiostateAI合资公司

2025年6月，中国专科特检领军企业康圣环球与专注RNA测序的BiostateAI在武汉签署协议成立RNA-AI精准医疗平台合资公司，整合300万例临床RNA-seq样本与Total-RNA测序BIRT工艺，将测序成本压缩近五倍，并一次性捕获mRNA、lncRNA、miRNA、circRNA等30万条转录本。BiostateAI的OmicsWebCopilot平台借助大语言模型自动生成统计脚本、可视化报告与论文草稿，帮助医院和CRO公司在缺乏生信人才的情况下迅速完成大规模转录组解读与差异分析。合资公司首轮产品将聚焦于自身免疫病、口腔癌、糖尿病、淋巴瘤与器官移植随访五大临床场景。

来源：公司官网，公开信息，沙利文分析

3.3.1 AI应用场景：合成生物学与生物制造

AI正以系统级建模、强化学习与生成算法重塑合成生物学核心环节，实现从代谢通路设计到酶工程改造与发酵工艺优化的智能化加速与精准控制

合成生物学作为一门融合生物学、工程学与计算科学的交叉前沿学科，旨在重构生物系统，使其具备特定的合成功能。通过对基因元件、代谢路径、调控网络等的工程化设计，科学家可以构建用于药物中间体、功能蛋白、生物燃料、材料单体等高附加值产物的大规模生物制造体系。

这一过程的核心在于对细胞“程序”的精准编码，从遗传信息层面驱动物质合成流程，实现绿色、可持续的替代生产方案。因此，合成生物学被视为下一代工业革命的关键技术底座，其产业化路径正快速延展至医药、农业、化工、食品等多个战略性领域。

传统方法面临的困难

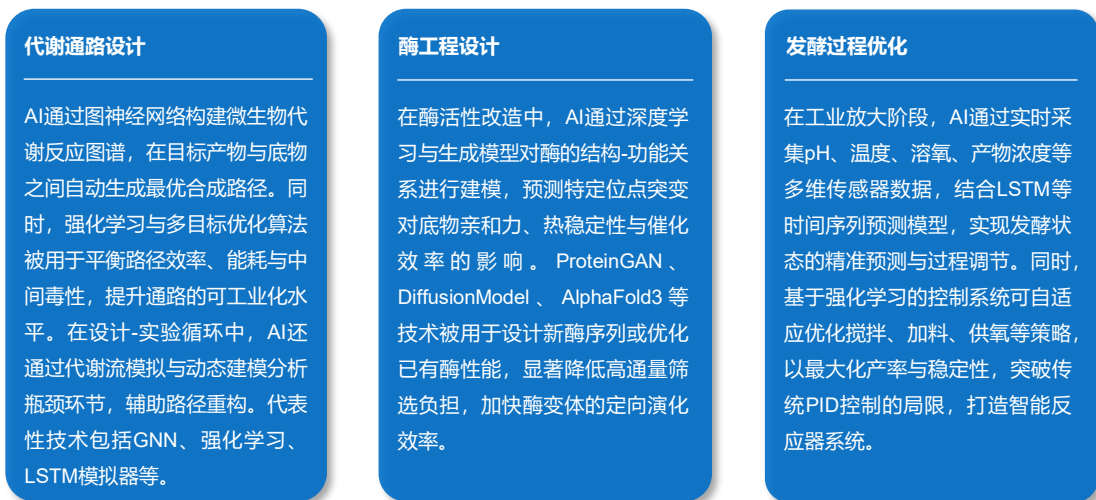
传统合成生物学与生物制造在实践中长期面临系统设计复杂、工程效率低与可预测性差等技术瓶颈。首先，生物系统本身具有极高的非线性和反馈耦合特性，导致通路设计常常呈现“黑箱式”不确定性，实验验证需通过大量试错完成，耗时耗力。其次，核心任务如代谢通路构建、酶定向进化、发酵工艺调控等，均高度依赖专家经验与重复性实验，缺乏自动化建模和优化能力。即便构建成功，工程菌的稳定性、产率优化与放大效应仍需多轮迭代。与此同时，大量数据往往未被充分利用，缺乏形成有效知识反馈的机制，严重制约了工艺可复制性与产业化落地速度。

AI带来的变革与优势

AI介入之下，合成生物学不再依赖经验积累与实验直觉。一方面，AI模型能够模拟复杂生物反应系统，精准预测设计路径的可行性、效率与稳态行为，显著提高设计起点质量。另一方面，通过深度学习、强化学习与生成模型等技术，AI可自动生成合成路径、筛选高性能酶突变体、实时优化发酵参数，并在不断反馈中实现自我学习与迭代。此外，AI还可打通DBTL之间的信息闭环，构建高度自动化的工程平台，将工程中的实时反馈快速转化为制造执行，加快从实验室到工厂的转化路径。

当前在合成生物学与生物制造工程中，AI已在代谢通路设计、酶工程设计与发酵过程优化等典型任务中发挥出核心作用，以提升设计效率、工艺稳定性与产率水平。

图：AI在核心任务中均承担了关键角色



来源：文献检索，沙利文分析

3.3.2 合成生物学场景的代表性AI企业

随着AI技术持续深入合成生物学全链条研发，国内企业正通过AI驱动的设计平台与智能化制造体系加速技术落地，推动合成生物学从实验室研究走向规模化产业应用

蓝晶微生物：AI赋能可持续生物制造

蓝晶微生物是一家中国领先的合成生物学企业，致力于通过AI技术赋能生物制造全流程。公司建立了一支由机器人、软件开发、机械电气、大数据和合成生物学等不同领域资深科学家和工程师组成的跨学科团队，开发合成生物学研发平台SynBio_OS，将设计-构建-测试-学习循环从实验室拓展到工业级别，显著提升了技术开发和产品交付的效率。目前已部署柔性自动化实验平台（BlueArk）、超高通量发酵平台（AutoFarm）和智慧云端数据系统（CyberFarm）三部分。预计未来3年内，SynBioOS能够将蓝晶微生物的单个产品完整研发周期在现有基础上再行缩短。通过AI技术的深度融合，蓝晶微生物实现了从菌株开发、小试、中试到工业量产各个环节的数据整合与优化，推动了合成生物学技术的产业化应用。

华大智造：构建AI驱动的合成生物学智能平台

华大智造作为国内领先的生命科技核心工具提供商，近年来积极布局合成生物学与生物制造领域，致力于通过“AI+自动化+测序”一体化平台，打造新一代智能化生物制造体系。目前，华大智造的合成生物学平台已在酶工程、代谢通路优化、发酵过程控制等多个领域实现应用，显著提升了研发效率和产业化水平，推动了合成生物学从实验室走向工业化生产的进程。

在功能蛋白设计方面，华大智造自主研发了EvoPlay蛋白质设计平台，基于强化学习算法，结合蛋白质语言模型（如ESM）和AlphaFold结构预测工具，实现了对蛋白质序列的高效优化与功能预测。在实验室自动化方面，华大智造推出了合成生物学智“惠”自动化产品组合，整合了高通量质粒构建、菌株筛选、酶活检测等模块，构建了从设计到测试的全流程自动化系统。在发酵过程优化方面，华大智造通过AI算法实现对发酵状态的动态监控与实时调节，突破传统工艺瓶颈，保障规模化生产的工艺稳定性。此外，华大智造还推出了针对合成生物学的测序工具解决方案，自动化设备可搭建全流程，成为合成生物学的有力支撑。

智衍造物：AI与合成生物学双核驱动平台

智衍造物成立于2024年，是国内首批聚焦“AI+合成生物学”的创新型平台企业。其核心愿景在于通过AI驱动与合成生物技术的深度融合，构建具有可预测性、可定制性的工业级菌株工程和酶催化体系，实现从菌株设计到量产的全链条数字化控制

公司创新性地构建SWIFT菌株工程平台，采用人工染色体技术作为菌种载体，在此基础上嵌入AI驱动的代谢网络模型，从系统层面进行代谢路径规划与调控。与传统手工设计相比，通路构建效率提升约8倍。智衍造物运用AI-酶催化设计系统，对候选酶的催化效率、底物亲和性、稳定性进行高精度预测，并结合自动化高通量筛选平台进行验证。这一流程显著加速了从序列到功能型酶的转化速度，极大降低筛选成本。其首期产品——基于解脂耶氏酵母合成的虾青素，已完成从研发验证到中试规模试生产，生产成本已接近传统化学合成路径，且生物活性和生物利用度在功能性市场具备竞争优势。此外，公司正推进EPA、叶黄素和β-胡萝卜素等多条类脂产物管线。

来源：公司官网，公开信息，沙利文分析

3.4 “集大成者”的智慧研发实验室

智慧实验室通过系统化改造与智能化集成，打破传统实验体系中的效率瓶颈与再现性难题，构建以数据驱动为核心的新一代生命科学研发基础设施

传统生命科学实验的痛点分析

在传统生命科学研究体系中，实验流程长期依赖人工操作、单机设备和线下管理，积累了诸多根深蒂固的结构性难题。首先，实验通量与运行速度受到严重限制。科研人员在完成如细胞培养、液体分装等基础任务时需频繁进行手工操作，尤其在化合物筛选、蛋白表达等需高通量支持的任务中，效率瓶颈尤为显著。其次，由于缺乏自动化手段，实验的每个关键步骤往往依赖操作者个人的操作习惯与经验判断，加之缺乏实时过程监控与标准化参数记录，导致实验复现时常常出现偏差甚至失败。生物医学研究成果难以在不同实验室中复现，“再现性危机”制约了基础发现的工程化落地与新药研发的迭代速度。此外，数据管理问题日益突出。在传统实验室中，不同仪器产生的数据往往分散于各自的软件平台，缺乏统一的结构、标签和接口标准，数据流动效率低下且容易丢失，使得数据难以被后续复用，更难作为AI模型训练所需的高质量样本库。

图：效率、数据、再现性成为制约传统实验室发展的关键因素



智慧实验室的发展目标

智慧实验室是生命科学领域长期痛点积压、传统实验体系难以为继背景下的结构性解决方案。其发展目标是系统性重构的方式，全面提升生命科学研究的效率、标准化水平、数据价值和智能决策能力。旨在构建一个具备极限通量的高效率实验平台，实现实验流程的全自动、连续化与高并行执行，从而显著压缩实验周期，加快科研节奏。同时，智慧实验室以“零人为偏差”为方向，建立高度标准化、模块化的实验流程体系，确保不同批次与不同场景下结果的高度可重复性。在数据层面，智慧实验室强调“数据即产物”，致力于打造从采集到存储全链路结构化、可追溯、可被AI直接调用的原生数据体系，为模型训练与科学推理提供高质量输入。更进一步，通过集成主动学习与强化学习等AI算法，智慧实验室还具备实时调整策略、优化实验路径的能力，从而实现“设计-执行-反馈”的动态闭环，使科研流程从被动执行走向智能驱动。

图：智慧实验室所配备的技术单元



来源：文献检索，沙利文分析

3.4 “集大成者”的智慧研发实验室

智慧实验室以“顶层设计+技术分层”的路径推动从数据采集、自动执行到智能决策的全流程升级，构建面向未来生命科学研发的智能化实验基础架构

图：从多个设计维度对比传统实验室与AI驱动的智慧实验室

对比维度	传统实验室	AI驱动的智慧实验室
数据管理与分析	人工记录/基础软件数据分析滞后	实时数据采集，AI算法自动分析，提供预测性见解
实验设计与优化	试错法为主，依赖经验	AI模拟实验条件，优化参数组合，加速研发周期
自动化程度	部分设备自动化，人工干预多	全流程自动化（样本处理、数据分析等）
资源管理	手动记录资源使用，易浪费	智能库存系统，动态分配世纪/设备，减少闲置
安全性与合规性	人工监控，依赖操作规范	AI实时监测异常（如温度、气体泄漏），自动记录数据
协作与远程访问	需物理在场，协作效率低	云端数据共享，支持远程实验监控与操作
可扩展性	设备与技术更新周期长	模块化设计，快速集成新技术（如新传感器）
成本结构	初期投入低，长期人力/资源成本高	初期投入高，长期通过效率提升降低成本
环境影响	能源/试剂浪费较常见	优化实验流程，减少废弃物，支持绿色科研
决策支持	依赖人工经验判断	AI提供数据驱动决策，预测实验结果可靠性

顶层设计框架指导智慧实验室的构建

智慧实验室的发展目标与传统实验的结构性难题是一体两面的映射关系。正是因为存在效率瓶颈、数据碎片化、不可复现等问题，智慧实验室的发展路径才被清晰地设定为必须达成的技术目标。为了逐步实现这些目标，智慧实验室并非一蹴而就地构建完成，而是通过一个由基础到高级、由局部自动化到整体智能化的演进过程推进的。智慧实验室正是在五个层次的逐步搭建与集成中，由自动化走向智能化，由局部流程提升走向全系统优化。每一层都对应着一个传统难题的解决路径，也共同构成智慧实验室实现其发展目标的技术逻辑闭环。

图：智慧实验室的五层递进范式



来源：文献检索，沙利文分析

3.4 “集大成者”的智慧研发实验室

中国企业正通过模块化、自动化与智能化的技术路径，推动智慧实验室从单点设备升级为系统集成，构建具备自主执行与动态调度能力的新型科研基础设施

在生命科学自动化版图里，中国企业正沿着“单模块—工作站—系统—智慧”四个迭代台阶迅速铺开：先把一个高耗时步骤彻底解放，做成硬件黑盒，再把若干动作封装成流程工作站。随后用导轨和调度软件把所有设备与数据流串成无人值守的流水线，最终让AI与机器人形成干—湿闭环、自主迭代的科研工厂。

从单模块到智慧自动化，实验室自动化的发展路径体现了生命科学研发体系由局部替代向整体重构、由工具导向向智能驱动的深层跃迁。每一阶段不仅在技术复杂度上逐层递进，更在科研范式上带来质变。前期以硬件解放人力、规范操作流程为核心，中期通过系统集成实现流程连续化与数据一体化，而最终则依托AI与算法构建具备自主学习与自优化能力的科研闭环。随着更多企业在各阶段完成突破，智慧实验室正逐步从概念验证走向规模应用，成为推动生命科学进入智能化时代的关键引擎。

单模块自动化：专攻瓶颈动作

单模块自动化的关键特征在于“聚焦一个关键步骤，用一台设备把它做到极致”。这些设备往往围绕移液、离心、称量、样本存取等核心动作展开，目标是实现操作高度标准化、误差最小化和连续运行能力。其价值不仅体现在效率提升，还在于显著提高数据一致性，降低人为干预引发的误差，是实验室从人工走向自动的第一步。




图：单模块自动化完成企业代表产品

<p>中析生物 SG93系列高精度移液平台</p>  <p>覆盖PCR搭建、梯度稀释等高频实验动作，具备多通道协同与重复精度保障能力。能够有效消除手工操作中的液量偏差问题，大幅提高样本处理速度与实验一致性。</p>	<p>艾迪迈 MSP2000</p>  <p>全自动磁性固相萃取-色谱/质谱样本前处理设备，具备条码追溯系统和封闭舱体设计。显著提升前处理环节的效率与可控性，减少操作时间的同时提升样本处理质量与人员安全性。</p>	<p>镁伽科技 CellVue</p>  <p>采用高分辨率CMOS相机，搭配数字共聚焦，能快速获取大量细胞图像，提供更综合、更准确的数据。跟随硬件升级匹配采集软件的更新完善，图像处理算法强大，大大提升相关分析效率和精度。</p>
---	---	---

工作站自动化：把一段常用流程封装成“黑盒”

工作站自动化强调流程封装，即将某一常用实验流程中的多个子模块（如加样、混合、提取、检测）整合在一个平台上，构成“即插即用”的实验黑盒。相比单模块，其优势在于实现了多个步骤的一致性和过程的串联连续化。科研人员只需配置试剂和参数，即可一键运行，大幅降低操作复杂度，同时显著提升实验的可重复性与通量，尤其适合前处理、高通量检测等高频任务。

图：工作站自动化完成企业代表产品

<p>华大智造 MGISP-960</p>  <p>具备液体处理、磁珠纯化和建库一体化能力，兼容多类测序方案和流程需求。可大幅减少人工介入的操作环节，使文库构建流程具备更高的一致性和通量。</p>	<p>海尔生物医疗 H-ACCS-AC</p>  <p>覆盖细胞接种、换液、传代、培养与收集全过程，搭载图像识别与物联网控制模块。能在细胞实验过程中有效避免批次差异问题，显著提升细胞培养流程的质量控制水平。</p>	<p>镁伽科技 Seloris</p>  <p>具备全流程可追溯记录系统，采用权限分离的数字化平台。可实现固体/液体原料的投料及搅拌溶解，精确pH值调节，超声波脱气等操作，全程高度自动化，降低重复性工作，减少人为操作带来的潜在误差。</p>
--	---	--

来源：公司官网，公开信息，沙利文分析

3.4 “集大成者”的智慧研发实验室

中国企业正通过模块化、自动化与智能化的技术路径，推动智慧实验室从单点设备升级为系统集成，构建具备自主执行与动态调度能力的新型科研基础设施

系统自动化：导轨+调度系统打造无人流水线

系统自动化的关键在于“全流程联通”。它通过轨道系统、物流机器人、调度操作系统等，将多个模块和工作站串联成可连续运行的无人流水线，实现实验样本从接收到数据输出的全链路无人化与统一质控。这一层级解决了实验室“设备孤岛”问题，使操作流、数据流、样本流三者融合，提升处理规模与效率，为大规模标准化生产、质控审计和数据资产积累提供了系统性支持。

图：系统自动化完成企业代表产品

<p>迈瑞医疗</p> <p>MT8000全实验室智能化流水线</p> <p>三轨磁悬浮轨道系统串联多学科自动化分析模块，配套有智能质控和样本调度平台。能够显著提升实验室各类样本的整体处理效率，在占地空间与资源配置上实现优化，并实现统一化的质量监控与流程管理。</p>	<p>晶泰科技</p> <p>XtalDynamics自动化系统</p> <p>集成多个工作站，通过调度系统和物流机器人连接合成、筛选、分析等模块，支持云端操控与自动反馈。实现实验室全天候运行，并在自动执行的同时不断积累结构化数据，为新材料、新药研发等任务提供高效率与高稳定性的实验支撑平台。</p>
---	---

智慧自动化：AI与机器人闭环的自迭代科研工厂

智慧自动化是实验室自动化的最高阶段，其核心特征是闭环与自我优化。在这一阶段，AI不仅参与分析结果，更能主动生成实验假设、规划实验变量并控制机器人执行实验操作。实验数据实时回流AI模型进行再训练，形成“设计—执行—反馈”的自学习系统，从而彻底解放了科研人员的低阶操作，使实验系统自身具备决策能力，实现了科研从静态执行向动态学习的范式转变，是支撑AI4LS的关键基础设施形态。

图：智慧自动化完成企业代表产品

<p>镁伽科技</p> <p>Megalaxy Lab 3.0</p> <p>整合MegaCell全自动细胞实验集群与MegaMolecule全自动分子实验集群等多个智能实验单元，并嵌入主动学习算法来实时优化机器人执行路径，形成完整的“实验-分析-决策”闭环。实现了实验流程中感知、构思与执行的深度耦合。系统不仅可以实施复杂实验任务，还可以根据实时数据自动生成下一个实验方案，从而将开发周期从6至10个月压缩至6至8周，极大提升实验通量、可重复性与科学发现的准确性。流程中的自动反馈机制不仅减少了人工参与，也保证了研究过程的持续优化与一致性。</p>	
--	--

来源：公司官网，公开信息，沙利文分析

3.5 AI4LS面临的核心挑战与策略应对

AI4LS面临的数据瓶颈与科学治理难题正逐步凸显，亟需通过标准化体系、跨学科融合与基础设施升级来夯实底层能力，提升生命科学智能化发展的稳定性与持续性

当前，行业已初步形成跨模态数据融合、智能算法创新、算力基础设施优化的多层次协同体系，但在数据质量、技术透明性、跨学科融合、伦理法规与生态建设等关键节点，依旧存在显著掣肘，需构建针对性的应对策略以推动其迈向更高质量的发展阶段。

■ 数据与技术挑战

首先，数据质量问题构成AI4LS应用扩展的根本制约。生命科学领域的数据具备显著的多模态特征，涵盖基因组、蛋白质组、表型组、生物影像、临床病历等异构信息，其在来源、结构、采集手段与语义标注等方面缺乏统一标准，导致数据可融合性差，严重限制了模型训练的广度与深度。加之，传统数据采集流程依赖人工操作，存在较高的噪声水平与系统性偏倚，进一步削弱了模型的泛化能力。针对这一问题，需从“源头标准化”与“体系联通化”两方面入手。一方面，通过部署自动化、标准化的实验平台与智慧实验室，实现数据采集过程的结构统一、标签一致与质控闭环；另一方面，推动构建符合国家信息安全与数据保护规范的数据共享平台，结合联邦学习、多方安全计算等新型隐私保护技术，打破数据孤岛，在保护数据主权前提下释放跨机构数据协同潜力。

其次，AI模型的可解释性与工程可控性不足。生命科学研究高度依赖因果推理与生物机制解释，而当前主流的深度学习模型多为“黑箱”结构，缺乏可溯源的决策逻辑，难以满足科研工作者对于结果机制可验证性与可重复性的基本要求。此外，AI模型的训练与推理高度依赖高性能算力资源，当前算法迭代速度与算力基础设施之间仍存在阶段性脱节，特别是在中小型科研机构中算力获取成本高、资源配置不均问题尤为突出。对此，采取“知识驱动”与“数据驱动”的融合建模策略，将结构生物学、调控网络等领域显性知识嵌入模型架构，可有效提升模型的可解释能力；同时加强算力服务平台的构建，依托国家级超级计算中心、AI专用加速芯片以及量子计算等新兴技术手段，为科研用户提供低门槛、弹性化的计算资源支撑体系。

■ 科学融合挑战

AI4LS的持续演进也面临跨学科协同机制滞后的深层结构性问题。生命科学的核心知识体系以连续场、生物过程与机制建模为主，而AI算法体系则多建立在离散优化、统计推理等计算模式基础之上，二者在模型表达、时间尺度、数据结构等多个维度存在根本性差异。此外，不同学科之间在术语、研究范式、目标定义等方面缺乏统一语义体系，使得跨域团队协作效率低、模型融合路径模糊。

应对策略应聚焦于统一跨模态知识表示框架的构建，例如发展支持连续-离散混合建模的跨尺度融合技术，引入知识图谱、多语义对齐机制等打通学科边界；同时，应加强跨学科人才培养体系建设，推动在高校、科研机构中设置AI与学科融合的复合型人才培养路径，并建立以科研问题为牵引的联合攻关平台，提升模型设计的科研适配性与工程转化效率。

来源：文献检索，沙利文分析

3.5 AI4LS面临的核心挑战与策略应对

AI4LS的伦理法规与生态体系仍处于早期建设阶段，亟需在数据合规、安全标准与产业协同机制上加快完善，以支撑人工智能技术在生命科学领域的可信、可控与可持续发展

■ 伦理与法律挑战

伦理治理与法规合规体系的不完善亦构成AI4LS大规模应用的制度性障碍。生命科学数据具有高度敏感性，涉及个体基因组、疾病史等难以匿名化的个人信息，一旦泄露或滥用，可能对患者造成长期性、不可逆的权益损害。目前，全球范围内关于医学AI的数据使用规范存在显著差异，国内相关法规仍处于迭代发展阶段，缺乏面向AI场景的专门规范机制。

此外，模型决策过程的不透明与可能的偏倚性引发了对“算法歧视”的广泛担忧。为此，应推动构建生命科学领域的AI伦理评价与算法透明性框架，探索“沙盒监管”机制，在确保科研合规与伦理底线的同时，允许在受控环境中进行前沿技术试验；同时，推动立法进程，引入包括同态加密、差分隐私、零知识验证在内的数据保护技术标准，以实现数据安全性、合规性与可用性的有机统一。

■ 生态与系统挑战

从生态体系建设的角度来看，当前AI4LS产业链仍处于早期集成阶段，缺乏完整统一的标准体系与开放兼容的技术接口，不同平台之间数据、模型与工具难以互通，严重影响了技术的可扩展性与资源的复用效率。同时，人才结构失衡，具备交叉背景的工程科研复合型人才数量有限，限制了AI技术在生命科学领域的深度渗透与转化落地。

对此，应加快行业标准制定进程，围绕数据标注格式、模型评测指标、接口通信协议等关键环节构建统一规范，推动构建开源共享生态，提升行业整体的技术成熟度与协作效率；同时，通过政府引导与市场机制双轮驱动，支持AI与生命科学融合的专业设置、课题资助与产教结合示范项目建设，构建多层次、多路径的人才培育与引进机制，夯实产业可持续发展的基础。

来源：文献检索，沙利文分析

第四章

中国AI4LS行业发展趋势一览

中国AI4LS行业在产业协作、跨界融合与自主创新的协同驱动下，正加速形成全球竞争力，迈向从技术追随到范式引领的转变

4.1 产业合作推动AI4LS商业化落地

中国AI4LS企业正通过“验证+协作”、“标准+接口”、“整合+落地”的深度产业合作路径，打通从算法模型到临床应用的关键环节，加速技术落地转化与商业模式构建

中国AI4LS企业正积极嵌入医药产业链条，通过与制药、生物技术、CRO、CDMO以及科研设备供应商的深度合作，逐步建立起以“技术验证—标准制定—生态协同”为核心路径的产业合作机制。这种合作不仅加速了AI技术的场景化落地，也为AI4LS技术成果的产业化应用奠定了关键基础。通过模型算法在真实场景下的验证、实验流程的标准化、接口协议的共建，以及核心应用数据的共享流通，中国AI4LS企业正逐步完成从技术供应者向价值创造者的转型。

■ “验证+合作”：打通从模型到临床的转化路径

AI在生命科学中的实际价值，最终要通过研发效率、临床验证、药物上市等实质性成果体现。因此，AI4LS企业在与制药、CRO、CDMO企业的合作中，首要任务是将AI模型从实验室推向现实验证环境，通过“共研+共测+共投”的方式建立协同研发体系。这类合作机制正在从“平台试用”走向“联合共建”，越来越多的AI4LS企业不再仅提供工具，而是以联合开发、共同承担临床试验和里程碑费用的形式深度参与药物全生命周期，实现AI在药效、安全性等关键临床指标上的真实验证。

图：联手验证临床可行性、拉通上游发现与下游制造的企业案例



- 深势科技**以算法平台切入，与CDMO企业诺泰生物达成战略合作，共建AI驱动的新型研发生产一体化平台。通过为诺泰定制AI研发平台、打造全流程数字化体系，深势的技术有效缩短研发周期、降低试错成本。短期优化运营智能增效，中期以环肽偶联配体为突破构建AI制药体系，延伸至合成生物学领域，搭建DBTL闭环及生物制造全链路平台；长期植入AI基因，开发跨领域大模型，实现多场景融汇与战略跃迁。

■ “标准+接口”：推动实验流程和数据体系标准化

在AI4LS加速落地的过程中，标准制定成为贯穿产业链整合的技术纽带。AI模型的可用性不仅取决于算法本身，更高度依赖实验流程是否标准、接口是否通用、数据是否结构化。中国AI4LS企业正通过与设备商、生物平台方的合作，共同推动行业标准的建立。通过标准与接口协作机制，AI4LS企业不仅提升自身效率，更逐步建立起可嵌入产业链的技术语言体系，实现与产业链上中下游的无缝对接。

图：行业建设一体化标准的企业案例

标准	牵头/参与方	涵盖要点	进展与意义
T/SBIAORG0002-2025《基于大语言模型(LLM)的药物警戒个例安全性报告AI辅助信息提取规范》	上海市生物医药行业协会；多家AI4LS企业	定义LLM在药物安全报告中的数字字段、验证指标及算法可解释性等要求	2025年5月正式发布，成为全球首个药物警戒-LLM团体标准，已被20+ CRO/CDMO采纳，用作合规接口模板。
TCIS35002-XXXX(征求意见稿)《智能制造生命科学自动化制造与分析系统设计指南》	中国仪器仪表学会；镁伽科技等	规定生命科学自动化实验室的系统组成及功能、错误、安全要求和电磁兼容性的设计内容	2025年4月发布征求意见稿并在苏州镁伽示范线先行验证运行良好。

来源：公开信息，沙利文分析

4.1 产业合作推动AI4LS商业化落地

中国AI4LS企业正通过“验证+协作”、“标准+接口”、“整合+落地”的深度产业合作路径，打通从算法模型到临床应用的关键环节，加速技术落地转化与商业模式构建

■ “整合+落地”：推动产业协同与商业模式创新

从早期技术验证和标准制定阶段迈入规模化应用阶段，中国AI4LS企业不再局限于“工具提供方”角色，而是通过与药企、CRO、CDMO、设备商等上下游主体的深度协作，逐步形成联合开发、平台定制、实验室整体解决方案等多元商业路径。AI平台与产业链环节实现互嵌共建，推动模型、数据、实验、转化等要素融合重组，形成贯通药物研发、评估与生产的系统性能力。目前，AI4LS行业正采取三种产业协同的模式促进技术导向走向价值协同：“联合开发”、“定制平台+服务复购”以及“实验室整体解决方案”。

图：产业协同三种模式



来源：公司官网，公开信息，沙利文分析

4.2 跨领域协同带动平台型企业崛起

AI4S技术的通用性正推动中国企业跨领域融合、多场景复制，众多平台型企业加速崛起，构建起贯通医药、材料、能源等产业的多元化创新生态

跨界协同的趋势源于科学问题的共性使然，更得益于数字技术的高可移植性。一方面，算法模型具有通用性：例如蛋白质结构预测的深度学习模型也可用于材料晶体结构性质预测，药物分子生成算法同样适用于设计新型催化剂分子。另一方面，算力和数据平台可共享：大型算力集群和自动化实验装置一旦建成，可以在不同学科的研发中全天候运转，提高总体投入产出比。

中国AI4S企业通过“技术复用-场景拓展-生态共建”的路径，将在一个领域成熟的AI模型、计算平台推广到新的应用场景，并联合不同领域的合作伙伴共建产业生态。通过技术要素在不同产业间的流动复用，跨领域协同正在催生生态级的平台型企业。此外，当企业在多个行业积累数据和场景后，便具备构建开放生态的条件：既能向不同行业输出标准化的AI基础设施，又能联合行业伙伴定制开发专用模型和应用，实现平台与生态的共生发展。这类平台企业不再受制于单一市场波动，并可通过跨界创新持续催生新增长点，将大幅提升中国AI4LS的整体竞争力。

可以预见未来能够跨越医药、材料、能源等领域提供AI赋能解决方案的中国平台型企业，将有望引领全球多领域融合的科技创新范式。

图：典型的跨领域协同实践案例



商汤科技长期投入于原创技术研究，涵盖感知智能、自然语言处理、决策智能、智能内容生成等关键技术领域，同时包含AI芯片、AI传感器及AI算力基础设施在内的关键能力。以原创技术体系为根基，SenseCore商汤AI大装置为核心基座，并在此基础上建立日日新SenseNova大模型及研发体系，布局多领域、多方向前沿研究，快速打通AI在各个垂直场景中的应用，向行业赋能。

- 在工业领域，商汤科技依托商汤工业引擎为一汽解放青岛基地开发了行业首条“5G+AI”的冲压件表面质量在线检测线。基于AI技术实现升级后，相关产线不仅精简了冲压车间的质检环节用工，更提高了质检结果的一致性和稳定性。
- 在医疗领域，商汤科技开发的SenseCare智慧医院综合解决方案利用了覆盖多模态的医疗基础模型群，赋能医院的就医、诊疗、随访等环节。目前该解决方案已落地全国多家头部三甲医院，用更多元化的AI服务助力医院智慧化转型升级。



镁伽扎根于生命科学实验室自动化，核心技术平台在跨领域应用中也展现出强大适应性。依托感知、构思、执行三大技术架构，镁伽通过自主智能体及智能操作系统，开发出一系列多智能解决方案，已覆盖生命科学及化工、食品饮料及农业、消费级集成电路及新能源等众多领域。通过在多个行业的成功试点，镁伽证明了一套智能实验室技术多场景复制的可能，这使镁伽从垂直领域的自动化供应商成长为综合性平台企业。

- 在制药领域，镁伽的全自动高通量筛选系统、大规模细胞培养工作站等装备已投入新药研发流程，提高实验效率和产能。
- 在新材料和石化行业，镁伽与安捷伦共建的联合实验室将智能机器人与分析仪器相结合，打造出了差异化的数字化实验室解决方案，实现对石油化工催化剂、高分子材料研发过程的自动化升级。
- 在农业领域，镁伽以AI+精准农业”为核心理念，依托自研VLA智能架构，通过多模态感知全面采集作物和环境信息，结合语言理解实现跨系统数据融合，并将AI决策快速转化为精准动作，实现农业全流程可视化、标准化和智能化。该方案已在葡萄种植中完成试行应用，取得阶段性成果。

来源：公司官网，公开信息，沙利文分析

4.3 科研、产业、创新三重驱动下的孕育环境

中国AI4LS企业在科研投入、产业基础和创新能力的三重驱动下，正逐步建立起支撑管线开发、自主创新和全球竞争的系统性优势，孕育出面向国际舞台的行业龙头企业

中国AI4LS行业正形成孕育世界级企业的肥沃土壤。这一环境由产业基础、科研实力和企业创新三股动力共同驱动，为本土企业走向全球领先提供了坚实支撑。

■ 产业基础

中国庞大的医药健康需求与完善的供应链体系，为AI4LS技术的落地提供了充足的应用场景与资源保障。一方面，创新药研发活跃，临床试验需求旺盛，开放的研究生态让AI驱动的药物发现能够更快获得验证反馈。医院、科研院所与企业之间的数据互联互通，叠加持续扩容的云算力，为AI模型的训练与部署提供了稳定支撑。

政策层面对技术创新持包容审慎态度，鼓励真实世界证据的使用，持续优化注册审批流程，为新技术落地留出灵活空间。资本市场与产业园区则为创业团队和成熟企业提供了实验平台、人才引进、技术转移等多维度支持，形成了贯通研发、生产、应用的良性循环生态。在这一综合环境的推动下，本土AI4LS企业得以在效率、成本与成果转化速度上不断提升，并逐步培养出面向国际市场竞争的能力。

■ 科研实力

中国持续加大对AI与生命科学交叉领域的投入，科研产出迅速提升，构筑起创新发展的知识高地。一方面，高校与科研院所长期坚持生命科学、计算机科学、材料科学等多学科交叉培养，造就了具备复合背景的研究群体，并通过贯通实验室、医院与产业界的协同机制，不断推动算法建模、分子机制阐释与前沿仪器开发等工作的深入开展。开放共享的研究范式已成共识，各类数据、代码与成果在统一平台上流通，加速创新思想的验证与迭代；临床医学与基础科研的紧密互动，使算法模型得以迅速融入真实病例场景并进行精准调优。另一方面，国家级与地方级重点实验室、工程技术中心相继布局，为研究者提供了先进的云算力、高通量实验和多模态成像设施，形成灵活高效的创新试验场。多元化科研资助渠道与包容性评价体系鼓励团队聚焦前沿课题，丰富的学术交流和竞赛进一步促进了知识的快速扩散与迭代。同时，伴随实验平台和监管机制的不断完善，科研成果正加速实现从“可用”到“可转化”的跃迁。三者相互支撑，共同构成了中国AI4LS产业创新从源头到落地的科研能力闭环，也为龙头企业的成长提供了高效而稳定的知识供给系统。

■ 创新能力

中国在人工智能与生命科学融合领域的创新活力，主要体现在跨学科团队的快速形成、算法与实验平台的紧密耦合，以及成果向产业端的高效转化。科研人员与工程师在同一研发闭环中协作，将深度学习、大模型和自动化实验等技术整合为可持续演进的工具链。

在此格局下，具有代表性的企业陆续涌现。百图生科通过构建超大规模蛋白语言模型，探索序列、结构与功能之间的内在规律，实现靶点先导发现与蛋白设计的一体化迭代。模型输出可直接进入其自动化实验流水线，再经高通量验证反馈回算法端，形成自我强化的循环。华为云依托盘古大模型，将分子生成、性质预测与合成可行性评估融为一体，并通过云端算力与无人实验室联动，为科研机构与药企提供“算法-实验-生产”一站式研发环境。望石智慧则聚焦化学空间探索与合成路径规划，其智能决策系统可根据实时实验结果动态调整分子设计方案，与合作方的微流控平台协同完成闭环优化，有效提升早期筛选效率。不同方向的实践展现了本土创新体系的多元性。随着产业链各环节对开放协作的认可度不断提高，此类创新范式将进一步加速知识扩散与技术迭代，为AI4LS领域输送更大的突破动能。

来源：公开信息，沙利文分析

第五章节

部分中国AI4LS领域 的企业介绍

5.1 镁伽科技



■ 公司介绍

镁伽科技是中国机器人技术应用领域领先的自主智能体 (Autonomous agents) 提供商，致力于为智慧实验室与智能制造场景提供全面的自主智能体及多智能体解决方案。通过高度智能化的软硬件系统，推动行业加速迈向高效、精准、可持续的智能化未来。凭借在人工智能与机器人自动化领域创新性构建的“感知 (Perception)、构思 (Conception) 及执行 (Execution)” 闭环技术架构，镁伽的多功能自主智能体能够独立执行复杂任务，同时模块化产品组合设计可轻松配置并规模化部署于不同场景，将人力从重复性工作中解放出来，使其专注于实验室与制造场景中的高价值任务，为下一代科研与制造业等领域发展做好前瞻布局。

目前，镁伽正将技术优势与产品创新力应用于生命科学及化工、食品饮料及农业、消费级集成电路及新能源等行业，并持续拓展至更为多元化的战略新兴领域，助力客户建设智能组织、突破技术边界、重塑运营范式，成为引领产业升级转型的关键驱动者与最佳合作伙伴。



900+
团队人数



50%+
研发人员比例



450+
授权专利数量



550+
申请专利数量

■ 发展历程

2016

- 镁伽成立
- 进军生命科学智慧实验室领域

2021

- 完成B-B3轮融资
- 发布Megalaxy Lab 1.0
- 业务扩展至新能源领域

2024

- Megalaxy Lab升级为3.0
- 业务扩展至精细化工行业
- 获工信部评为国家级专精特新“小巨人”企业
- 杭州交付中心成立

2017

- 完成天使轮融资

2018

- 镁伽苏州交付中心成立

2019

- 完成A及A+轮融资
- 业务扩展至消费级集成电路产业
- 深圳交付中心成立

2022

- 完成C轮融资

2025

- 业务扩展至农业领域
- 完成D轮融资

■ 公司成就

商业化成就

第一国产自主智能体供应商
*以2024年中国智慧实验室收入规模计

2024年度收入人民币9.3亿元

最广泛的机器人应用场景覆盖

截止2024年12月31日累计服务880+客户

技术领先性

Labillion and LibraX: 中国**首套**智慧实验室自主智能体操作系统

Megalaxy Lab: 中国**首个**基于多智能体解决方案的智能生命科学实验室

Aufluo: 中国**首款**AI驱动液体处理工作站

CellVue: 国产**首批**商业化通用型高内涵成像分析系统

Dicing Blade: 国内**首款**7μm超薄切割刀片

5.1 镁伽科技



■ 核心技术架构

公司开发的自主智能体依托感知、构思及执行三大核心技术架构驱动，通过先进闭环架构形成自我强化的生态系统。感知捕获实时环境与运行数据，注入构思的AI驱动优化机制，进而指导执行进行精准物理运作。

图：核心技术架构——感知-构思-执行

- **感知** 采用先进传感，实时实现精准环境感知
- **构思** 基于“感知”输入及特定任务需求持续优化执行策略智能决策，实现战略级自适应响应
- **执行** 通过机器人技术和自动化实现精准灵活的操作；确保任务可靠落地，将概念化的解决方案无缝转化为有形的实际成果



■ 产品组合

智慧实验室		智能制造	
智能体	操作系统	智能体	操作系统
<ul style="list-style-type: none"> Aufluo 液体处理智能体 CellVue 高内涵成像分析智能体 Seloris 制药智能体 AegisRex 样品前处理智能体 Dismix 饮品分液智能体 Misa 巡检与移栽智能体 	<ul style="list-style-type: none"> Labillion: 人工智能赋能的下一代智能操作系统 LibraX: 高效智能质控系统 	<ul style="list-style-type: none"> Manavis-AOI系列 晶圆缺陷检测 Manavis-Overlay系列 套刻量测 Manavis-Particle系列 粒子缺陷检测 Manavis-Battery 电池检测 Fabsil-Laser系列 激光切割 Fabsil-Dicing系列 晶圆切割 	<ul style="list-style-type: none"> FusionTwinX 智能制造操作系统

行业探索



■ 镁伽鲲鹏实验室3.0

2024年9月，鲲鹏实验室3.0作为面向智慧生物实验室的多智能体解决方案标杆实践诞生。依托Labillion和LibraX操作系统，自主智能体构建了数据管理、自动化及人工智能闭环系统，具备高通量处理能力，可多维度处理海量数据并保持卓越质量。系统在采集实时数据时，确保流程稳定一致，并能驾驭复杂场景，超越人工处理能力，将开发周期从传统的6-10个月压缩至6-8周，推动科研数字化、智能化升级。

图：镁伽科技的三大自主智能体集群

<p>MegaMolecule全自动分子实验集群</p> <p>每天可完成约400项复杂生物分子构建任务；特别适用于菌株筛选、抗体筛选和合成生物学等场景</p>	<p>MegaCell全自动细胞实验集群</p> <p>每天可完成超过10,000个样本点的药物筛选实验；在高通量药物筛选和稳定细胞系构建方面效果显著</p>	<p>MegaOmics高通量多组学集群</p> <p>全面数字化基础设施，为药物研究提供涵盖多模态数据采集与生物信息分析；确保全流程数据可追溯性和可靠性</p>
--	---	--

来源：公开信息，沙利文分析



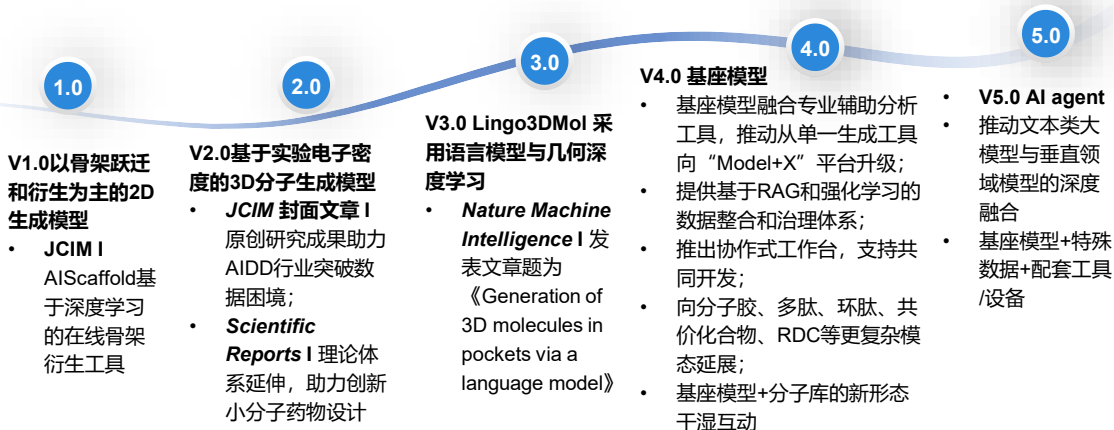
5.2 望石智慧

■ 公司介绍

望石智慧 (StoneWise) 成立于2018年，是一家使用人工智能技术驱动新药研发的科技公司。凭借人工智能药物研发底层理论的突破、药物研发数据的深度治理、药物研发行业的沉淀认知和强大的软件与工程能力，望石搭建了能够精准地生成与靶点口袋结构契合的分子或分子骨架的多模态AI 3D分子生成大模型基座。

公司的分子生成模型以GPT/Transformer框架为基础，融合了几何深度学习等算法。该模型还可作为基座模型，合作伙伴能够在此基础上，充分整合自身拥有的数据、认知以及模型等要素，进行定制化迭代。目前，已有近百家药企、科研机构客户日常使用。依托该模型，公司内部自研管线已收获显著成果，其中最快的管线已进入临床一期；同时基座模型衍生出的实体库业务，也在海内外积极扩展中。

图：V 1.0-5.0 分子生成模型的迭代路径：始终以技术为导向，持续推进技术驱动的药物研发新范式



■ 核心技术优势

- 基于CADD方法，建设广度与多样性兼备的高质量数据库，包括3D复合物、片段势能扫描数据与NCI能量数据库
- 基于**实验电子云密度的深度挖掘**，构建NCI库、晶格点库等，指导模型理解关键结合位点

01 千万级蛋白小分子复合物与亿级类药小分子数据

- 基于Transformer架构，融合**语言模型与几何深度学习**等算法
- 覆盖空口袋生成、多片段固定生成、骨架跃迁等不同使用场景，满足**FIC、BIC**等不同药物研发需求
- 根据药物机制特点进行数据建设与训练，可满足**分子胶、共价化合物、多肽**等不同机制药物设计需求

02 多场景、多模态的算法模型架构

- 合理性评估**：通过2D成药性指标与3D构象参数筛除结构不合理或难以合成的分子，确保生成结果具备基本可信度
- 活性复现评估**：在信息不泄露条件下通过SBDD与LBDD模式检验模型对真实活性分子的复现能力，真实反映分子功能有效性
- 创新性评估**：考察模型生成新骨架新结合模式、新骨架旧结合模式的能力，衡量其探索未知化学空间和支持新药研发的潜力

03 工业级价值的三阶段递进式评估框架

■ 融资进展

公司自成立以来获得众多业内知名的资本支持，7年内共完成6轮融资，总金额超1亿美元。



5.2 望石智慧

■ 平台服务：模块化工具支持研发流程

小分子药物发现平台

基于药物设计需求的MolVado™药物发现平台以分子生成模型为核心，配套分子动力学模拟、相对结合自由能计算、扭转角势能扫描、化合物子结构检索等系列分子评估工具。望石进一步开发了可定制的药物发现平台MolVortex®，可通过治理湿实验数据并进行模型训练，提升模型生成能力。

该平台将全新骨架Hit的发现周期从过去的6个月-2年压缩到了3个月以内。目前多家国内头部药企在MolVado™平台效果已使用验证的基础上，启动MolVortex®定制平台开发，用于湿实验数据处理。

AI模型基座构建化合物实体库

公司借助AI分子生成模型的基座，为小分子创新药的早期研发，提供更高效的基于AIDD、CADD和化学实验室的“干湿”整合服务，构建更好代表化学空间的实体分子库，目前已实现万级别量级的实体交付，同时在特定的结构下具备差异化优势。

目前该业务正在海内外积极探索中，有数十家客户已完成商业转化，例如贝勒医学院、英国癌症研究中心、清华大学、中科院物理研究所、广州国家实验室、复旦大学医学院、齐鲁锐格等。

■ 生态合作：科研驱动与技术创新并行发力

望石智慧自成立以来始终以科研驱动和技术突破为核心引擎，公司多项基于模型的研究成果已在Nature、Nature子刊、CELL、JCI封面文章等专业顶级期刊发表。与AWS、华为、阿里、英伟达等上游算力与技术厂商深度协作，形成“模型—算力—应用”的生态闭环，确保平台长期算力供给与技术前沿性。

在“十四五”国家重点研发计划“生物与信息融合”专项中，与北京大学、清华大学、中科院等机构合作构建基于AI大数据驱动和可信安全计算的创新药物筛选系统。公司还深度参与了北京市“AI+健康协同创新”项目，着力打造融合电子密度的多模态AI药物设计平台，探索从“依赖分子库筛选”向“空口袋生成并迭代优化”的全新研发范式。

■ 自主管线研发：AI驱动的临床级分子产出

二代3CL小分子口服抑制剂

- 公司与广州实验室合作开发的抗新冠小分子口服药物为国内**首个**进入临床阶段的二代3CL小分子口服抑制剂候选药物，自研发立项至IND申请仅用时**12个月**。
- 具备高吸收、代谢稳定等特征，相比目前已上市的疗法适用人群更**广泛**。对比拟肽类药物，单疗程成产**工艺及成本**都有巨大**优势**。

HPK1抑制剂SWA1211

- 在AI驱动的电子密度分析和分子生成算法支持下，该分子具备全新骨架和优异的成药性，展现出**同类最佳**的潜力，并且与PD-1抗体类药物联用展现出更强的协同抗肿瘤疗效。
- 2025年3月已获得**中美双IND批件**；同年6月完成针对晚期实体瘤**I期临床研究**的首例患者给药。

■ 战略发展



模式延展

以分子生成功能为基石，正在从小分子模态会延展到分子胶、共价化合物、多肽、环肽等，逐步向分子量更大的“大”分子做模式延展。



平台共创

从基座AI分子生成大模型出发，到工具链整合平台；同时可以融合客户数据，在不触碰客户数据基础上展开平台共同开发项目。

来源：公开信息，沙利文分析

5.3 予路乾行

■ 公司介绍

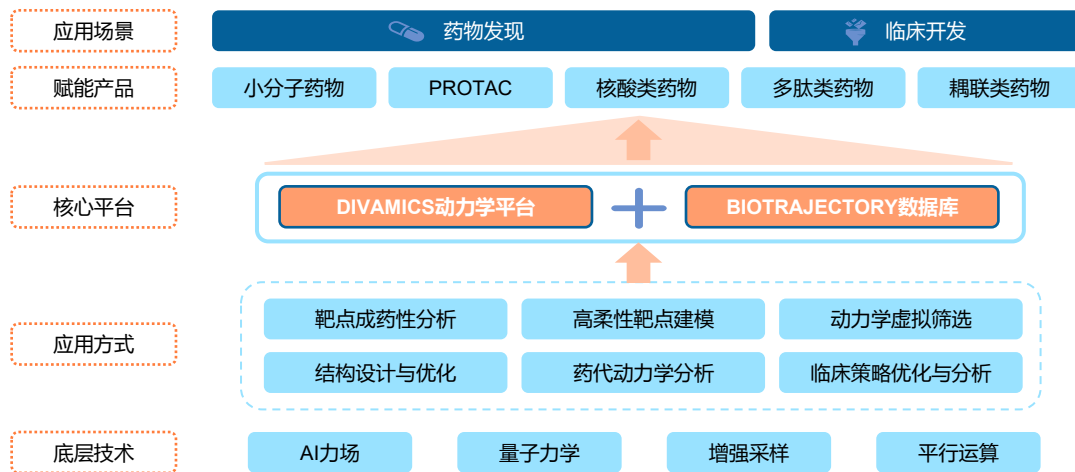
苏州予路乾行生物科技有限公司 (Divamics Inc.) 是一家以人工智能、量子力学及分子模拟算法为核心技术的新药早研企业。公司自创立之初即定位服务全球市场，依托自研的 Divamics 多尺度分子动力学平台与 Biotrajectory 动力学轨迹数据库，以独创的“分子电影”药物设计理念，为生物医药企业提供 AI 驱动的 Hit-to-PCC 一体化药物研发服务。目前已服务国内外50多家医药行业客户，合作推进80余条新药研发管线。公司总部位于苏州，在北京设有计算中心，在新加坡、日本设有 BD 中心。

■ 核心平台与商业模式

予路乾行以“平台+管线”的模式推动新药早研，核心依托两大技术平台：1) **Divamics 平台**：基于“分子电影”理念研发的多尺度分子动力学模拟平台，能够精确再现实验无法观测的微观生物学过程，补齐药物研发逻辑链条中长期缺失的关键信息；2) **Biotrajectory 动力学轨迹数据库**：整合全球药物研发大数据，系统揭示药效学与药理学背后的分子机制，高效分析临床与专利信息，助力研发决策。目前已收录20个蛋白家族、500+难成药靶点、上亿级专利化合物，覆盖肿瘤、神经、心血管、免疫等九大疾病领域。作为全球首个集成了大语言模型的生物医药AI智能体数据库，具备20倍速复杂数据处理能力，加速药物研发。

从商业化模式来看，公司面向包括小分子、大分子、核酸、抗体及多肽在内的五大药物类型予路乾行通过 **服务、联合研发与分润合作、自研管线转让交易** 等多元模式，与全球药企实现价值转化。

图：予路乾行技术体系与平台架构



■ 核心团队

予路乾行由一批具有国际化视野的科学家创立，核心成员毕业于北京大学、美国佛罗里达大学等顶尖高校，曾在欧美领先药企及研究机构从事计算化学与药物设计研究。创始人郑铮博士研发的自由能算法MovableType，得到了诺贝尔化学奖得主罗德·霍夫曼的认可和推荐。公司的科研与工程团队均来自世界顶尖机构，硕博比例超过90%，具备理论/计算化学与有机/药物化学的综合背景，能够在计算化学家与药物化学家之间建立有效的信息桥梁，形成独特的跨学科优势。

来源：公开信息，沙利文分析

5.3 予路乾行

■ 科研实力：在AI药物研发领域，具有国际化的学术影响力与产业价值

- **2022年**，在JCIM连续发表两篇封面文章，提出基于动力学模拟的新药研发范式与 Movable Type 自由能模拟平台，并解析 c-Met 激酶构象转变机理
- **2023年**，在JCIM发表c-Met激酶构象转变与活性调控研究成果
- **2024年**，与博瑞医药及武汉理工大学合作完成GLP-1/GIP双重激动剂BGM0504的研究，成果发表在《Nature》子刊Scientific Reports
- **2025年2月**，联合北京大学、武汉理工大学在Chem发表适配体降解剂系统进化平台研究；5月，在JCIM发表DihedralsDiff小分子三维构象生成模型；6月，再次在JCIM发表分子胶评估方法学研究；7月，博瑞医药与予路乾行（Divamics）合作开发的新一代双GLP-1/GIP受体激动剂BGM1812，正式发表在国际药物化学权威期刊《Journal of Medicinal Chemistry》

■ 行业合作：在产业合作及临床转化方面构建了全球化的创新生态

国内市场

公司自2021年成立后，持续建立与国内生物医药客户的合作，订单复购率达到80%，已公开的合作项目：

- 与北京鞍石生物合作，利用高精度分子模拟技术解析药物-靶点相互作用机制，仅用一年时间便推动一款非小细胞肺癌靶向药物进入临床试验；2024年，双方合作的非小细胞肺癌靶向药物管线正式进入临床三期试验
- 与博瑞医药战略合作并连续产出成果——代谢疾病方向的BGM1812与BGM0504
- 与某中美跨国公司和苏州勤浩医药达成数百万元战略合作，聚焦AI+小分子难成药及难成药靶点的研究
- 携手立迪生物，整合功能性检测与多组学平台，探索PDX反向工程靶点发现
- 与华先清研、华先医药共同打造“干-湿实验协作研发平台”
- 与晶云药物开展战略合作，提供从分子发现到临床供应的一体化解决方案
- 与康龙化成达成战略合作，整合AI分子模拟与全球研发服务网络，构建“动态计算-定向合成-智能优化”的新药开发体系

海外市场

公司成立之初就定位于服务全球的生物医药企业，去年开始与北美某头部Pharma企业达成了战略合作，双方将开展16条药物管线的合作，目前（截止2025年5月份）已成功交付2个项目。2024年底又与海外某客户签订首个“干湿实验整合”订单，通过整合AI设计、动态模拟与实验验证的全链条能力，正式迈入一体化解决方案提供商的新阶段。与韩国Therasid等公司的合作也验证了予路乾行的技术与服务能力发展亚洲市场的可行性。目前公司有超过60%的业务贡献来自海外市场。

■ 服务案例

公司凭借自研的多尺度分子动力学与AI融合平台，在多个合作案例中展现了独特优势。其中，在与博瑞医药的合作中，予路乾行利用AI驱动分子动力学优化，发现并改造关键盐桥和侧链，成功推动BGM0504和BGM1812两个分子进入临床及临床前阶段，展现出在代谢疾病治疗中“疗效增强+差异化机制”的双重价值。

图：服务产品——博瑞医药BGM0504和BGM1812

药品名称	靶点	药物类型	适应症	产品全球最高临床阶段
BGM0504	GLP1R x GIPR	多肽	II型糖尿病、肥胖	临床三期
BGM1812	AMYP x CTR	多肽	糖尿病、肥胖	临床前

来源：公开信息，沙利文分析

5.4 剂泰科技

■ 公司介绍

剂泰科技是一家人工智能驱动纳米材料创新的生物科技公司，专注于利用靶向药物递送和发现技术，加速患者在重要疾病领域获得创新疗法。公司由美国工程院院士陈红敏博士以及MIT科学家赖才达博士、王文首博士联合创立，已获得国家专精特新“小巨人”、国家高新技术企业等重要资质认定。

公司自主开发全球首个人工智能驱动的纳米递送解决方案提供平台NanoForge，拥有目前全球最大规模千万级LNP脂质库，并基于NanoForge打造了三大核心解决方案：AiLNP（AI核酸递送系统设计平台）、AiRNA（AI mRNA序列设计平台）、AiTEM（AI小分子制剂设计平台）。公司已具备实现肝脏、肺部、肌肉和免疫细胞等8个关键器官和组织LNP精准靶向递送能力，在多器官、多组织靶向递送难题上取得突破性进展，为肿瘤、代谢系统疾病、自体免疫性疾病、神经系统退行性疾病等提供成药机会，也在器官水平上为生命体抵抗衰老提供可能。

■ 发展历程



■ 核心技术优势

全球最大规模LNP脂质库驱动AI递送创新： METIS可电离脂质库构成NanoForge的数据核心，涵盖超1000万种脂质结构，其中大量脂质具备20余项关键特性的湿实验数据支撑，保障模型训练与算法优化的稳健性。该专有数据集规模助力剂泰科技突破知识产权限制，系统整合正负实验结果，推动分子生成与预测模型微调，持续拓展器官靶向能力。

首创纳米材料创新垂类人工智能基础模型： NanoForge以系列垂类人工智能基础模型为底层架构，覆盖纳米材料开发周期中分子生成、特性预测、实验设计、处方优化和专家验证的各个关键阶段，可跨功能并行运行，并实现小分子纳米材料优化、预测脂质和脂质纳米颗粒20种关键物理化学和功能特性。

自研专有一体化干湿实验室基础设施： 剂泰科技构建了独特的“干湿实验一体化”平台，将AI计算设计、高通量制剂筛选与体内验证形成端到端的闭环优化。该平台通过AI引导的自动化实验系统，可同时测试数百种制剂组合，并根据靶向递送特性快速筛选高性能脂质。

业内首个纳米递送人工智能体： 剂泰科技自研业内首个纳米递送人工智能体ALAN，作为NanoForge平台的智能接口，可执行自然语言或结构化查询，引导基于脂质的递送系统开发与优化。ALAN能够理解并分解任务，动态调用基础模型及平台内数据集，包括精选脂质库、内部数据与外部文献。其推理层可集成多模型输出，并基于领域规则进行结构化推理，逻辑清晰透明。

来源：公开信息，沙利文分析

5.4 剂泰科技

■ 核心解决方案

剂泰科技依托其自主构建的NanoForge平台打造了三大专业解决方案，为不同类型药物的开发提供全方位的AI赋能。

AiLNP平台

AI核酸递送系统设计平台

通过计算建模技术，平台可模拟生物大分子与超分子行为，并预测组装体与制剂特性。核心能力包括：基于专有脂质库、语言模型与生成算法，智能生成筛选候选脂质体；融合AI预测与分子动力学模拟，解析作用机制并预测20余种脂质特性。平台还支持对LNP各关键组分进行设计与比例优化，构建高效定制化制剂方案。

AiRNA平台

AI mRNA序列设计平台

支持AI驱动的全长mRNA序列设计与多目标优化，涵盖表达、折叠、稳定性及生产可行性。平台利用生成式AI构建器官靶向非翻译区（UTRs），并经高通量实验验证；同时模拟预测纳米颗粒与mRNA的体内动态（如分布、摄取、逃逸与响应），结合实验数据协同优化LNP、制剂与序列，提升治疗效能。

AiTEM平台

AI小分子制剂设计平台

利用人工智能、高通量迭代与量子化学等技术，驱动算法模型完成辅料类型与比例等制剂设计与优化。通过分子动力学与量子模拟，可预测药物-辅料相互作用并筛选最优候选。平台还具备优化共溶剂溶剂化等纳米级制剂的能力，从而提升药物溶解度与生物利用度。制剂样品经自动化高通量平台快速测试后，由AI迭代分析溶解度与渗透性等数据，推荐高性能制剂推进后续开发。

■ 产品管线

中枢神经类	MTS004	假性延髓情绪障碍
		吞咽困难
代谢类疾病	MTS201	肥胖
		2型糖尿病
		代谢功能障碍 相关脂肪性肝炎
		短肠综合征
肿瘤类	MTS105	肝细胞癌
	MTS107	宫颈癌/头颈癌
	MTS108	小细胞肺癌
	MTS110	前列腺癌
自免疾病类	MTS109	自身免疫

➤ MTS004

- 国内首款完成III期临床的AI赋能制剂新药，也是中国首款且目前唯一完成III期临床试验的PBA药物，有望填补国内PBA治疗领域尚无获批药物的空白。从立项、新药临床试验申请批准至完成III临床试验仅耗时38个月，大幅提升了药物研发效率。
- MTS-004口崩片剂型创新为PBA患者群体提供了更优的服用便利性，有效解决了此类患者群体普遍存在的吞咽困难难题。

➤ MTS105

- 利用mRNA技术，通过编码双特异性T细胞衔接蛋白（BiTE），这是一种新型的治疗机制，能够引导免疫系统更精准地识别和攻击肿瘤细胞
- 基于剂泰科技肿瘤靶向LNP递送技术，能在肝脏和肿瘤组织中实现高浓度的局部特异性表达，提高治疗效果的同时减少对正常细胞的影响。

■ 商业模式

剂泰科技采用风险可控的自研与合作开发“双轨”模式，充分释放平台解决方案的商业价值。双轨并行的战略布局，可以有效支持公司的可持续发展，平台业务产生的收入将反哺管线资产的研发进展和临床转化。

- 产品合作：通过许可授权、共同开发或商业化合作等方式，推动自有管线的研发进程
- 平台战略合作：为外部合作伙伴提供递送系统、脂质库和各项药物发现解决方案

来源：公开信息，沙利文分析

■ 法律声明

- ◆ 本报告著作权归沙利文所有，未经书面许可，任何机构或个人不得以任何形式翻版、复刻、发表或引用。若征得沙利文同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“沙利文”，且不得对本报告进行任何有悖原意的引用、删节或修改。
- ◆ 本报告分析师具有专业研究能力，保证报告数据均来自合法合规渠道，观点产出及数据分析基于分析师对行业的客观理解，本报告不受任何第三方授意或影响。本报告数据和信息均来源于公开信息渠道，沙利文拥有对报告的最终解释权。
- ◆ 本报告所涉及的观点或信息仅供参考，不构成任何投资建议。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。在法律许可的情况下，沙利文可能会为报告中提及的企业提供或争取提供投融资或咨询等相关服务。本报告所指的公司或投资标的的价值、价格及投资收入可升可跌。
- ◆ 本报告的部分信息来源于公开资料，沙利文对该等信息的准确性、完整性或可靠性拥有最终解释权。本文所载的资料、意见及推测仅反映沙利文于发布本报告当日的判断，过往报告中的描述不应作为日后的表现依据，沙利文不保证本报告所含信息保持在最新状态。在不同时期，沙利文可发出与本文所载资料、意见及推测不一致的报告和文章。同时，沙利文对本报告所含信息可在不发出通知的情形下做出修改，读者应当自行关注相应的更新或修改。任何机构或个人应对其利用本报告的数据、分析、研究、部分或者全部内容所进行的一切活动负责并承担该等活动所导致的任何损失或伤害。

联系我们

毛化 Fred Mao

弗若斯特沙利文大中华区医疗业务合伙人兼董事总经理



联系邮箱:

fred.mao@frostchina.com

知识中心 Knowledge Center

弗若斯特沙利文大中华区生命科学事业部知识中心



联系邮箱:

hcknowledgecenter@frostchina.com