深势科技情况介绍

# 一、深势科技基本情况介绍

深势科技是全球AI for Science领先者，致力于解决原子、分子、基因和蛋白等微尺度科学研究和工业研发的关键问题，科研团队由中国科学院院士领衔，汇集了近百名数学、物理、化学、生物、材料、计算机等多个领域的优秀青年科学家和工程师，其中博士及博士后的占比超过35%。团队核心成员获得2020年全球高性能计算领域最高奖项“戈登贝尔奖”，相关工作当选2020年中国十大科技进展和全球AI领域十大技术突破。

深势科技是国家高新技术企业、国家级专精特新“小巨人”企业，设有1个企业博士后科研工作站，是先导物成药性研究全国重点实验室、科学智能北京市重点实验室的联合依托单位。公司承担国家及省部级科技项目20余项，已申请国内发明专利122项，PCT专利9件，其中16项发明专利已授权，另有软件著作权93项。

深势科技依托在交叉学科领域的深耕，构建了AI for Science大模型体系“深势·宇知”，并以“微尺度工业设计与仿真”为切入点，打造了玻尔科研空间站、Hermite药物计算设计平台、RiDYMO难成药靶标研发平台及Piloteye电池设计自动化平台等科研和工业研发基础设施，形成了AI for Science的“创新-落地”链路和开放生态，赋能“千行百业”。

基于在“AI for Science”领域取得的各项成果，深势科技得到政府、用户、资本市场等各方的高度认可。在能源材料领域，已经与宁德时代、比亚迪、多氟多等能源材料的头部企业开展合作；在生物医药领域，已服务国内外80家的医药类院校和研究所，以及翰森制药、上海恒瑞医药、云南白药、东阳光药业、泓博医药等大型上市药企、头部创新Biotech及国内外CRO巨头；在教育科研领域，已服务包含北大、清华、上海交大、复旦、武汉大学、普林斯顿、牛津大学等国内外知名高校的一千余个课题组。

# 二、核心技术与成果

## （一）核心技术：新一代分子模拟技术

深势科技自主研发的新一代分子模拟技术，开创性地提出了“多尺度建模+机器学习+高性能计算”的科学研究新范式。基于第一性原理的计算数据，通过深度神经网络对原子相互作用势能进行参数化，构建了能够有效克服“维数灾难”的模型，解决了微观尺度分子模拟中精度与效率难以兼顾的问题，突破性地实现了精度与效率的统一，在合金、新能源等材料领域有极大的应用前景。2020年，在当时世界第一超算Summit上，深势团队在保持量子力学精度的前提下成功模拟了上亿原子的运动轨迹，将超大系统的分子动力学模拟带入了新阶段。在这之前，最大的模拟系统是100万个硅原子，而此研究则成功模拟6.79亿个水分子及1.27亿个铜原子，速度也比之前的模拟体系高出了几个数量级，使原来可能需要60年才能完成的模拟体系缩短到了1天，打破了此前35年来分子动力学局限于小型系统（数千个原子）的困境，并因此斩获了2020年的戈登贝尔奖。

## （二）主要成果：“深势·宇知”AI for Science大模型体系

### 1、原子大模型

**DeePMD：**深度势能分子动力学（Deep Potential Molecular Dynamics, DeePMD）是一种通过深度神经网络拟合原子间相互作用势（potential energy surface, PES）的方法。训练数据集通常是由基于密度泛函理论（density functional theory, DFT）的方法计算获得。通过机器学习拟合基于第一性原理产生的数据，DeePMD 实现了分子动力学精度和效率的极大统一，现已成为理论化学、计算物理、分子生物学、材料科学等领域的热门研究方法，基于 DeePMD 发表的相关工作已超百篇。

**DPA-1：**2022年12月，深势科技和合作者基于 DeePMD 提出了DPA-1，开发者在含有56种元素的较大数据集上进行了预训练，并将此预训练模型在各种下游任务上进行了迁移学习，实验表明，此预训练模型能大幅降低下游任务训练所需数据量及训练成本、提高模型预测精度，从而对分子模拟相关领域产生深远的影响。对比DeePMD，在某些条件下DPA-1的测试精度甚至能实现一两个数量级的提升，这说明模型可以从现有数据中学习到隐含的原子间交互信息，也进一步证明了模型强大的迁移能力。

**DPA-2：**2023年12月发布，采用了多任务训练的策略，可同时学习计算设置不同、标签类型不同的各类数据集。由此产生的模型在下游任务上显示出极强的 few-shot 乃至 zero-shot 迁移的能力，显著超越过去的方案。DPA-2 模型的数据集已覆盖了半导体、钙钛矿、合金、表面催化、正极材料、固态电解质、有机分子等多类体系。

### 2、分子大模型

**Uni-Mol：** Uni-Mol直接将分子三维坐标信息作为模型输入，在几乎所有与药物分子和蛋白口袋相关的下游任务上都超越了 SOTA 水平，并能直接完成分子构象生成、蛋白-配体结合位点预测、可药性等三维构象相关任务，预测结果准确度超越现有解决方案。Uni-Mol 作为基座大模型，具有强大的通用性，在蛋白-药物构象生成、单步逆合成设计、药物毒理性预测、碳捕捉材料性质预测、OLED发光性质预测等不同场景中均能通过迁移和微调达到该场景的领域最优。Uni-Mol已被应用于深势科技的多个产品中，也获得了大量学界和业界研究人员的广泛关注。 **Uni-Mol+ 进一步增强了性能，**Uni-Mol+ 基于低成本的方法如 RDKit/Openbabel 生成初始构象，并通过迭代优化这些构象，使其逼近DFT方法得到的高精度稳态构象，从而可以通过基于模型优化后的构象来获得更精确的量子化学性质预测结果。

**Uni-ELF：**Uni-ELF 通用电解质配方设计框架，通过分子与配方阶段的预训练，在预测分子属性（如熔点、沸点、可合成性）和配方属性（如电导、库仑效率）方面显著优于现有的先进方法。在分子层面上，利用 Uni-Mol 模型重建三维分子结构；在混合物层面上，从分子动力学模拟中预测统计结构性质（例如径向分布函数）。全面的预训练使 Uni-ELF 能够捕捉复杂的分子和混合物级别的信息，显著提升其预测能力。Uni-ELF 框架的应用场景还可扩展应用于其他需要配方级别预测或生成的领域，如化工品（涂层/燃油等）、日化品（洗护/美妆等）与药剂等领域。

### 3、基因大模型Uni-RNA

2023年7月，深势科技推出Uni-RNA基因序列大模型。Uni-RNA 利用约10亿条高质量RNA序列进行了大规模的预训练，几乎涵盖了所有RNA空间，充分挖掘了RNA序列的潜在信息。通过在广泛的下游任务中微调模型，Uni-RNA在 RNA 结构预测、mRNA序列性质预测和 RNA功能预测等三个RNA领域的七个主要任务中全部取得了领先的结果，更是为未来RNA领域研究的深度革新提供了无限可能。

### 4、蛋白大模型Uni-Fold

2021年12月，深势科技推出Uni-Fold并开源训练代码。2022年8月，深势科技开源Uni-Fold v2.0，进一步提高精度和效率。蛋白结构建模，即通过序列、冷冻电镜图像等解析得到蛋白的三维结构，是基于结构的药物研发的先决条件。得到蛋白结构后，对重点区域进行精修和优化，是保证后续研究过程正确性的关键。蛋白结构预测工具Uni-Fold在蛋白质结构研究领域首次完全开源训练代码与推理代码，支持高聚体蛋白体系的结构预测，且在相同训练数据集情况下，预测结果准确度达到业界顶尖。

### 5、文献大模型Uni-SMART

Uni-SMART是一个多模态科学分析大模型。该模型通过整合文本、表格、图表和分子结构等多模态专家标注数据，结合大规模模型训练、微调及多模态RAG技术，显著增强了对科学文献中多模态内容的理解与处理能力。在多个领域的严格定量评估中，Uni-SMART展现了卓越的性能，超越了现有主要基于文本的语言模型。

### 6、表征大模型Uni-AIMS

Uni-AIMS（AI-Powered Microscopy Imaging System），是一款集成了前沿人工智能技术的电镜图像分析软件。Uni-AIMS利用深度学习算法自动识别电镜图像中的物体，并进行细致的表征计算，输出详尽的统计报告，为科研探索和工业生产提供数据支持。Uni-AIMS的AI引擎采用最新技术，实现对电镜图像中目标的高精度自动识别，提升了图像分析的自动化和智能化水平，显著提高分析效率和准确性。

# 三、主要产品

## 1、Bohrium 玻尔科研空间站

Bohrium®玻尔科研空间站是深势科技面向AI for Science时代打造的科研云平台，其由教学、科研和开发三个核心平台组成。其中，**教学平台**通过提供Notebook、课程和比赛等互动式教学工具，有效解决了传统代码开发教学中“眼、手、脑不同步”的问题，使得学习过程更加高效和直观。**科研平台**则集成了DPA、Uni-Mol等先进的科学大模型引擎，对第一性原理计算、分子动力学等微尺度科学计算算法和软件进行了深度优化，为科研人员提供了强大的算力支持和便捷的计算模拟环境。**开发平台**进一步助力科研人员，使他们能够在短时间内（五分钟）开发出AI for Science的应用APP，从而释放科研生产力，显著提高创新效率。

## 2、Hermite 新一代药物计算设计平台

Hermite® 是深势科技基于AI for Science打造的新一代药物计算设计平台，可以为临床前药物研发提供一站式计算解决方案。其上加载行业领先的自由能微扰计算工具Uni-FEP、超高通量虚拟筛选工具Uni-VSW等核心模块，赋能蛋白结构预测、靶标确证、苗头化合物发现、先导物优化等多个环节。

Hermite®提供了基于SaaS的交互式全新分子展示体验，能对项目、成员、数据进行细粒度管理，具有完备资质认证和多层级安全防护，支持在线使用和私有化部署，深受客户信赖。目前已有超六成国内头部行业用户选择使用Hermite®平台，应用于超过50个药物管线项目中，客户累计开展Hermite® Uni-FEP计算任务超过20万次。

## 3、RiDYMO 高质量Hit发现和优化平台

RiDYMO®是深势科技基于AI for Science开发的高质量Hit发现和优化平台，依托自主研发的Hermite®计算药物设计软件，旨在通过创新的人工智能、物理算法和高通量实验，解锁难成药靶点的动力学机制，探索更广泛的小分子、大环分子及环肽分子的化学空间。该平台特别在口服大环分子设计方面拥有独特优势，结合平行化合成可快速交付创新的药物候选化合物。

作为其核心算法之一的强化动力学（Reinforced Dynamics, RiD）在分子模拟采样效率上实现了近百倍的提升，通过充分结合神经网络的高维表征能力，RiD能有效捕捉复杂生物大分子体系中的动态构象变化。RiDYMO®已成功应用于多个项目，包括c-Myc、GPX4的小分子发现，以及环肽和ADC等其他形式的药物，充分展示了其在创新药物发现中的强大潜力。

## 4、“玄铸”人工智能与物理建模驱动的材料设计平台

“玄铸”是深势科技研发的新一代材料设计平台，深度结合人工智能技术与物理模拟和先进表征技术，并提供高质量多源数据库、高通量和高性能计算引擎和高便捷性的操作界面。材料研发人员可以在“玄铸”材料智能设计平台的帮助下深入理解材料构效关系，进而加速材料及工艺，持续提 升新材料性能，推动能源电池、催化剂、聚合物、合金、陶瓷、电子产品等领域发展。

## 5、Piloteye 电池设计自动化平台

深势科技电池设计自动化平台Piloteye®是中国自主知识产权且全球领先的电池设计自动化智能研发平台。平台聚焦Read-文献调研、Design-实验设计、Make-合成制备、Test-表征测试、Analysis-分析优化这五个电池研发的关键环节，通过结合机器学习、跨尺度建模等先进算法，提高电池设计的精确度和可靠性，缩短创新到量产间的周期，推动电池研发领域“设计理性化”“开发平台化”和“制造智能化”的实现，打通电池研发全生命周期，连接产业链上下游，是新一代电池行业的基础设施。