

## 附件 4

# **“生物与信息融合（BT 与 IT 融合）”重点专项 2023 年度项目申报指南**

（仅国家科技管理信息系统注册用户登录可见）

为落实“十四五”期间国家科技创新有关部署安排，国家重点研发计划启动实施“生物与信息融合（BT 与 IT 融合）”重点专项。根据本重点专项实施方案的部署，现发布 2023 年度项目申报指南。

本重点专项总体目标是：聚焦未来生命科学、医药健康产业和经济社会发展等重大需求，通过加强生物技术与信息技术跨界融合研究，兼顾科学创新和技术图谱，引领新经济模式发展。突破信息大数据、生物大数据的获取、管理、分析、挖掘、调控和知识发现等底层支撑技术，提升数据整合与转化利用能力；构建 DNA 存储、类脑智能与人机交互、生物知识图谱、可编程细胞智能、智慧医药等交叉融合技术，推进大数据驱动的生命科学知识发现及转化应用；催生一批面向生命健康的颠覆性新技术，形成一批新工具、新技术、新标准与新产品，解决医药大数据、医疗人工智能原创性理论基础薄弱、重大产品和系统缺失等难点问题。

2023 年度指南部署聚焦生物技术与信息技术跨界融合底层支撑技术的“卡脖子”问题，坚持以“瞄准‘十四五’目标，继续

聚焦国家重大需求，对标世界科技前沿，加强 BT-IT 深度融合的新技术新产品研发”为宗旨，进行原理方法发现、关键使能技术提升、系统集成与平台应用示范布局。围绕生物信息编解码与存储、生命—非生命融合智能生物系统构建、BT-IT 融合技术健康医学系统集成应用示范 3 大任务，按照基础研究、共性关键技术、应用示范，拟启动 24 个研究方向，拟安排国拨经费概算约 4.06 亿。其中，围绕 DNA 复合分子体系的数据编码与存储等 9 个技术方向，拟部署 5 个青年科学家项目，每个项目 300 万元，拟安排国拨经费概算 0.15 亿元。

项目统一按指南二级标题（如 1.1）的研究方向申报，实施周期不超过 3 年。除特殊说明外，每个指南方向拟支持 1 项。申报项目的研究内容必须涵盖二级标题下指南所列的全部研究内容和考核指标。基础研究类项目下设课题数不超过 4 个，项目参与单位总数不超过 6 家，共性关键技术类和应用示范类项目下设课题数不超过 5 个，项目参与单位总数不超过 10 家。项目设 1 名负责人，每个课题设 1 名负责人。

青年科学家项目（项目名称后有标注）不再下设课题，项目参与单位总数不超过 3 家。项目设 1 名项目负责人，青年科学家项目负责人年龄要求，男性应为 1985 年 1 月 1 日以后出生，女性应为 1983 年 1 月 1 日以后出生。原则上团队其他参与人员年龄要求同上。常规项目下设青年科学家课题的，青年科学家课题负责人及参与人员年龄要求，与青年科学家项目一致。

每个指南任务原则上支持 1 项，特殊情况下，在同一研究方向下，当出现申报项目评审结果前两位评价相近、技术路线明显不同的情况时，可考虑支持 2 个项目。2 个项目将采取赛马制方式分两个阶段支持。第一阶段完成后将对 2 个项目执行情况进行评估，根据评估结果确定后续支持方式。

本专项研究涉及人类遗传资源采集、保藏、利用、对外提供等的项目，应遵照《中华人民共和国生物安全法》《中华人民共和国人类遗传资源管理条例》等相关规定执行。

本专项涉及人的生命科学和医学研究的项目，应遵守国家卫生健康委、教育部、科技部、国家中医药管理局印发的《涉及人的生命科学和医学研究伦理审查办法》，申请单位需在申请书中提交该项目不违背科技伦理要求的初步审核意见。在项目正式实施前，应按照规定通过伦理审查并签署知情同意书。

本专项研究涉及实验动物的项目，应通过实验动物福利和伦理审查，遵守国家实验动物管理法律法规、技术标准及有关规定，使用合格实验动物，在合格设施内进行动物实验，保证实验过程合法，实验结果真实、有效。

本专项研究涉及人工智能的项目，应遵守中共中央办公厅、国务院办公厅印发的《关于加强科技伦理治理的意见》等国家相关规定的伦理要求和国家新一代人工智能治理专业委员会发布的《新一代人工智能伦理规范》等伦理规范。

本专项研究涉及科技伦理的项目立项后，项目承担单位和主

要参与者应加强科技伦理知识的学习和培训，严格执行国家有关法律法规和科技伦理要求，尊重国际公认的伦理准则。

## **1. 基于 DNA 原理的信息存储系统开发**

### **1.1 基于 DNA 复合分子体系的信息编码与存储(基础研究类, 平行设置青年科学家项目)**

研究内容：研究 DNA 与其他分子（如多肽、高分子等）的复合数据存储方式，开发新型存储基元，数量级提高编码容量与存储密度，实现编码原理创新；完成超大规模（TB 级别以上）信息编码的模拟验证；建立复合分子的高效写入与稳定存储技术；研究复合存储介质的主要降解机制；开发微阵列大规模数据直读式快速寻址与随机读取方法；建立复合分子数据存储数据库并形成适配性纠错算法。

考核指标：开发 1 套基于 DNA 复合分子的新型数据存储体系，包括编码原理、读写技术、适配性试剂等，实现数据信息的高密度（单位编码效率 $>4$  比特/单位物理或存储单元）存储。提升单位信息编码效率至 10 比特/单位物理或存储单元以上，建立 1 套可大规模寻址的快速、无损、直读式索引方法，实现 MB 级数据的随机读取；开发 1 套适配性纠错算法与数据库检索方法。

有关说明：项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：开发 1 套基于 DNA 复合分子的新型数据存储体系，实现数据信息的高密度（单位编码效率 $>4$  比特/单位物理或存储单元）存储。

关键词：DNA 复合分子体系、信息编码、适配性纠错

## 1.2 基于非传统分子信息写入技术的 DNA 存储研究（基础研究类）

研究内容：围绕 DNA 存储分子信息写入关键问题，颠覆传统的“从头合成”DNA 存储信息写入技术路线，构建基于新型分子比特信息写入的数据操控构架，实现基于非传统分子信息写入技术的 DNA 存储原型系统；新型分子信息比特需使用非传统新型分子及结构（不包括已有 DNA 粘合等核酸片段）；针对相关技术路线，完成理论模型构建；突破传统分子信息写入技术路线的“从头合成”串行操控限制，建立与非传统分子信息写入路线相匹配的新型分子信息比特串编/解码技术体系，在写、读、编码密度或存储功能（如加密、寻址、搜索、计算等）上实现跨越式突破；研究在非传统分子信息写入 DNA 存储的写、读、存过程中，出现的主要错误类型和错误率，形成相应的纠错编码算法。针对新型信息比特的介质/原理，对新型技术方法完成 TB 级存储模拟和仿真评估，评估并展示其适用的存储信息类型和专用分子存储系统应用。

考核指标：构建具有规模化潜力的基于非传统分子信息写入技术的新型 DNA 数据存储体系，形成 1 套具有自主知识产权的原创性原理方法和相关读写、编码等技术体系等。颠覆性 DNA 存储技术需要通过非从头合成方法（不包括已有 DNA 粘合技术等传统方法）实现并行方式的信息写入，并完成 KB 或 KB 以上

量级数据写入的实验示范。利用新型分子比特(不包括基于 DNA 粘合技术的核酸片段)信息写入的数据操控构架,通过非传统分子信息写入方式实现单次不低于 30 比特位的并行分子信息串操控。新型分子信息比特不基于普通核酸序列,建立利用 2 级及以上 DNA 分子电路,实现分子数据库写入端的单次动态操控不低于 30 比特信息。单个新型信息分子比特,其纳米结构尺寸小于 1 纳米,可实现同时对不少于 25 种 30 比特长度的信息串的并行读取。对新型 DNA 存储技术的规模化潜力进行相关理论模型论证。

关键词: 新型信息分子比特、串行读写操控、专用存储体系

### **1.3 新一代超高密度 DNA 读取技术研发(共性关键技术类)**

研究内容: 针对高通量测序目前视场与分辨率互相制约的瓶颈,研究基于超分辨显微的新测序技术和方法,样品密度突破衍射极限;研究适用于高通量测序的超分辨显微方法;研究高密度可重用的测序芯片工艺与配套的小尺寸测序样本装载方法;研发具有高稳定性,强发光效率,满足超分辨测序的短波长水溶性荧光染料;研究基于异构计算实现的高分辨率图像快速准确重建;研发具有自主知识产权,鲁棒性好,易于使用和维护的测序仪原理样机,利于商业化产品开发和实现商业价值。

考核指标: 研制出一款测序位点间距 $\leq 300$  纳米,点阵密度 $\geq 1000$  万点每平方毫米的高密度芯片,以及基于超分辨显微技术的测序样机;研制配套的国产化染料试剂和数据处理软件。形成

可市售产品，面向 PE100 读长测序，过滤后测序质量数  $Q30 \geq 85\%$ ，过滤后单张芯片单次下机通量  $\geq 1.2\text{Tb}$ ，过滤后单张芯片日均通量  $\geq 400\text{Gb/天}$ ，在测序质量方面具有与市售主流高通量测序仪器相近的水平；申请 3~5 项发明专利。

有关说明：鼓励产学研联合申报。

关键词：超分辨显微技术、超分辨图像重建、超高密度 DNA 测序仪

#### **1.4 蛋白质单分子测序系统研发（共性关键技术类，平行设置青年科学家项目）**

研究内容：蛋白质是生命活动的具体执行者，直接影响生物体的生长发育与疾病发生，蛋白质的翻译后修饰在生命活动中起着至关重要的作用，其诸多信息无法通过基因测序直接获得；同时蛋白质也是信息存储的潜力载体，具有较核酸更高的理论存储密度（4.32 比特/氨基酸），针对当前蛋白质序列信息获取速度慢、检测限高、动态范围小、准确率低等世界性难题，亟需研发具有自主知识产权的蛋白质单分子测序系统。探索基于生物学、化学、光学、电学或多模态的蛋白质测序创新路线，建立相关理论模型并完成概念验证；利用上述技术，研制具有单分子检测灵敏度的检测体系；建立配套的软硬件支持，包括但不限于检测芯片研制、信号探测技术、样品处理技术、核心生化试剂、数据分析流程、高精度信号识别算法等。

考核指标：建立一套蛋白质单分子测序方法，完成概念验证

并搭建原型机，实现单分子检测灵敏度及单氨基酸识别精度。天然氨基酸区分种类不少于 10 种，修饰氨基酸检测种类不少于 3 种；单次测序累计读长不低于 20 个氨基酸；测序速率不低于 1~5 个氨基酸/小时；申请核心发明专利 3~5 项。

有关说明：鼓励产学研联合申报。项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：建立一套蛋白质单分子测序方法与原型机搭建，实现单分子检测灵敏度及单氨基酸识别精度。

关键词：蛋白质单分子测序、单分子灵敏度、天然与修饰氨基酸检测

### **1.5 高通量蛋白组学数据采集关键技术和设备（共性关键技术类）**

研究内容：针对传统蛋白质免疫检测蛋白质的数量、质谱技术鉴定蛋白质的数量和通量有限，以及不能同时进行蛋白质及翻译后修饰检测等瓶颈问题，自主开发蛋白质组数据采集的关键设备，通过综合多种技术方法，设计多种技术路线，发展自主可控的高通量蛋白质组分析关键技术。技术路线包括但不限于：质谱技术，超高重抗体联合检测技术，超高重核酸适配体联合检测技术，高通量多肽测序技术和纳米孔技术等。结合全自动化高通量样本处理及检测平台，开发大人群血浆/血清蛋白组学检测技术，应用于疾病高危人群高通量蛋白标志物筛查。

考核指标：建立高通量蛋白组检测技术，开发具有自主知识



产权的相关设备，实现高通量血浆/血清蛋白组学检测，单次检测靶标蛋白数量 $\geq 1000$ 种；实现至少2种对翻译后修饰的检测；申请核心发明专利3~5项。

关键词：蛋白质组检测、高通量靶标蛋白检测、翻译后修饰检测、数据采集技术与设备

### **1.6 高精度大规模长读长测序系统研发（共性关键技术类）**

研究内容：突破长读长测序技术原始准确率低的桎梏，完成大规模长读长测序系统的国产化研发及应用拓展。建立测序复合体系的模拟方法，对测序过程的微观机制展开研究，指导生化体系的设计与优化。研究分布式多芯片并行测序系统架构设计，重点突破软件与计算架构。对核心测序蛋白、测序芯片、计算卡、FPGA等关键部件开展自主化研发及国产替代。开展甲基化测序等应用类技术开发，实现在多类型分子信息存储、肿瘤检测、转录组测序等场景的应用。

考核指标：研制出国产高精度大规模长读长测序系统，单次测序原始准确率 $\geq 99\%$ ，读长 N50  $\geq 50\text{Kb}$ ；单台系统单次运行总通量 $\geq 5\text{Tb}$ ，实现测序蛋白与测序芯片的自主化或国产替代；甲基化测序准确率 $\geq 95\%$ ；PCT 专利申请获得受理 $\geq 20$ 项。

有关说明：项目执行期3年，在项目启动后1年实行里程碑考核。考核指标为：单次测序原始准确率 $\geq 99\%$ ，读长 N50  $\geq 50\text{Kb}$ 。

关键词：国产化测序复合体系、高精度高通量长读长、甲基化测序

### 1.7 面向海量数据存储应用的基地式 DNA 存储一体化软硬件系统（共性关键技术类，平行设置青年科学家项目）

研究内容：面向政府机构、数据中心或企业的海量数据的长期存储需求，研制 DNA 存储一体化软硬件系统，实现 DNA 存储操作软硬件系统的集成开发与生态系统建立，重新定义信息存储。研究 DNA 存储适配的合成器件、测序器件与信息处理电路的异构集成理论与方法，填补领域空白，建立规模化、标准化 I/O 的 DNA 存储读写的全自主硬件系统及生化耗材，研制用于 DNA 存储应用的适配合成器件与设备以及读取器件与设备，突破 DNA 存储读写速度。突破传统信息系统的技术框架，研制支持多种 DNA 存储的编解码辅助设计软件、DNA 存储自动写入和读取软件，建立支持多种数据格式的高效 DNA 存储编解码子系统，搭建存储流程的评估体系，形成标准化 DNA 存储软件流程。

考核指标：开发集成不少于 6 种 DNA 存储编解码方法、不少于 3 种纠错编码的编解码辅助设计软件平台和评价平台，可用于多应用场景下的 DNA 存储模拟。完成 1 套 TB 级以上的 DNA 存储一体化设备软硬件集成，数据写入错误率  $\leq 1\%$ ，数据读出恢复率  $\geq 99\%$ ，实现比现有写入读取效率（5 字节/21 小时）3~5 个数量级的突破，单通道达到不低于 250 字节/小时，并行读写通量达到 Gb 级别，建立 DNA 存储产业的文件格式标准。申请核心发明专利 2~3 项，申请软件著作权 3~5 项。

有关说明：鼓励产学研联合申报，项目执行期 3 年，在项目

启动后 1 年实行里程碑考核。考核指标为：开发集成不少于 6 种 DNA 存储编解码方法、不少于 3 种纠错编码的编解码辅助设计软件平台和评价平台。

关键词: DNA 存储一体化基础系统、编解码系统、评估体系、标准化软件流程、海量数据应用

## **2. 面向生命—非生命融合的智能生物系统构建与开发**

### **2.1 视听生物感知非侵入高精度解码与干预增强技术（共性关键技术类，平行设置青年科学家项目）**

研究内容：针对视听感知解码精度不高、增强调控手段缺失问题，面向临床诊疗及康复客观评估需求开展非侵入式视听感知高精度神经解码与干预增强研究。通过生物—信息融合交互揭示视听觉刺激与生物响应内在关联，发展基于 EEG 的非侵入式二维视觉空间信息编解码关键技术、多模空间听觉感知及听觉注意解码技术，拓展视觉信息传输维度、听觉信息辨识精度。面向典型视听障碍如青光眼、人工耳蜗患者，提供基于非侵入式 EEG 解码的康复效果定量评估方法与指标，挖掘康复效果与时间、外部刺激策略等因素的关联特性。面向嘈杂声学场景，开展语音信号听觉选择与注意力辨识研究，通过 EEG 信号实现听觉注意解码和注意力切换监测，解决听障人士嘈杂场景中关注对象的注意状态解读。研究非侵入式视听感知能力干预增强技术，有效提升相关场景下特定人群视听感知辨识能力。

考核指标：基于非侵入式二维视觉空间信息编解码技术，提

出  $\geq 6$  种二维图形的视觉编解码策略，图形识别正确率  $\geq 70\%$ ，重建图形与原图相关性  $\geq 0.7$ ，基于 EEG 特征的像素拓扑编解码结构  $\geq 3 \times 3$ ；基于非侵入式听觉诱发生物响应解码技术，实现中文刺激下听觉注意力解码时间窗  $\leq 1s$ ，解码准确率  $\geq 75\%$ ，空间听觉最小可辨识角度  $\leq 5^\circ$ ，空间定位准确率  $\geq 80\%$ 。面向视听障碍诊疗、康复评估等  $\geq 3$  个场景开展技术验证，青光眼、人工耳蜗等典型病例数量  $\geq 1000$ ，为临床患者提供视听诊断与康复效果评价服务  $\geq 2000$  人次；提出视听障碍诊疗与能力增强相关的标准或指南  $\geq 1$  项，申请发明专利  $\geq 5$  项，取得医疗器械注册证  $\geq 1$  项。

有关说明：项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：解决核心技术实现非侵入式二维视觉空间信息识别，图形类型  $\geq 6$  种，识别正确率  $\geq 70\%$ ，重建相关性  $\geq 0.7$ ，拓扑编解码结构  $\geq 3 \times 3$ ，中文刺激下听觉注意力解码时间窗  $\leq 1s$ ，解码准确率  $\geq 75\%$ 。

关键词：生物信息解码、干预增强、非侵入式、二维视觉、空间听觉

## 2.2 基于生物启发的感算一体神经类器官芯片及应用（基础研究类）

研究内容：研究高效灵活的生物启发计算模型和整体架构，设计集成多脑区的类器官计算模型；研究基于忆阻器等新型神经拟态元器件的高效感算一体神经类器官芯片架构设计，支持脑区计算模型的部署与运行，构建神经类器官芯片功能评价体系；研

究神经类器官计算电路与多通道生物信号双向交互电路的单芯片集成设计；研究基于三维成像技术的高通量神经电极精准植入定位技术；研究面向特定环路的类器官芯片—生物神经网络融合技术；研制生物—电子混合生物拟态原型系统，并开展若干在生物体上的植入式原型应用。

考核指标：研制基于生物启发的感算一体神经类器官芯片，单芯片具备生物信号采集通道 $\geq 256$ 个，集成神经拟态器件 $\geq 100$ 万个，刺激通道 $\geq 32$ 个；研制生物—电子混合神经拟态原型装置，支持神经拟态网络模型 $\geq 5$ 种，提供集成开发环境及工具链；实现面向生物功能障碍的生物体植入式应用演示 $\geq 2$ 种；获得第三方测试报告，申请发明专利 $\geq 5$ 项。

有关说明：项目执行期3年，在项目启动后1年实行里程碑考核。考核指标为：实现单芯片具备生物信号采集通道 $\geq 256$ 个、集成神经拟态器件 $\geq 100$ 万个。

关键词：生物启发、生物信号、神经拟态、神经类器官芯片

### **2.3 超微生机交互系统高通量读译控关键技术及应用（共性关键技术类，平行设置青年科学家项目）**

研究内容：针对传统生机交互系统性能不足、集成度低导致应用受限的现实问题，发展电生理信息高通量读译关键技术，开发高性能超微型非侵入生机智能交互系统。开展无创生物电极与采集电路、调控电路的全集成化研究，实现超微一体化无创生机接口器件；开展基于超微器件的感觉诱发编码范式、极微弱 EEG

信号解码算法以及无创高分辨率 EEG 节律调制技术研究,实现目标 EEG 节律的高信噪比诱发、亚微伏级 EEG 特征的高精度解码以及基于 EEG 节律的高空频分辨率自适应靶向调控;开展高维度指令、高时速响应、高自由度控制、高应答率等高通量读译控关键技术及超微型高性能非侵入生机交互系统集成研究,实现面向医疗临床、航空航天等重大场景的示范应用。

考核指标: 单电极生机接口器件体积  $\leq 1\text{cm}^3$ , 连续工作时长  $\geq 2$  小时; 研究弱隐刺激诱发的无创生机接口 EEG 编解码范式, 视觉刺激单元素视偏角  $\geq 2^\circ$ , 听觉诱发刺激强度  $\leq 45\text{dB}$ ; 发展极微弱 EEG 信号高精度解码算法, 非侵入信号特征解码精度  $\leq 1\mu\text{V}$ , 识别准确率  $\geq 0.9$ , 信息传输率  $\geq 200\text{bits/min}$ ; 设计多维度相干电闭环无创生机调控技术, 精准刺激模型点空间分辨率优于  $0.25\text{mm}$ 、频率分辨率优于  $0.01\text{Hz}$ 、生物调控维度  $\geq 3$ 、靶区准确度  $\geq 90\%$ , 目标节律增强  $\geq 30\%$ ; 实现 4 自由度 12 维度以上连续实时人机操纵控制, 完成航空场景的测试应用; 实现基于 EEG 信号的 200 以上大指令集中枢生机控制, 完成航天场景的测试应用; 实现筛查准确率 0.9, 调控应答率 0.8 以上的精神状态检测与干预, 完成医疗与航天场景的测试应用, 医学受试人数  $\geq 500$ , 航空受试人数  $\geq 10$ , 航天受试人数 (天基+地基)  $\geq 10$ 。相关性能指标通过国家权威检测机构检验报告, 制定生机智能交互技术相关的标准、指南或专家共识  $\geq 1$  项。申请发明专利  $\geq 10$  项、软件著作权  $\geq 5$  项, 获批国家医疗器械注册证  $\geq 2$  项。

有关说明：项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：开发生机接口编解码范式，视觉刺激单元素视偏角  $\geq 2^\circ$ ，听觉诱发刺激强度  $\leq 45\text{dB}$ ，无创 EEG 信号特征解码精度  $\leq 1\mu\text{V}$ 、EEG 编解码控制指令集  $\geq 200$ 。

关键词：超微系统、生机交互、高通量信息编解码、读译控闭环调控

## 2.4 基于生理传感驱动的急救数字孪生技术及模拟人装备(应用示范类)

研究内容：发展辅助医疗系统的技术实力，研究基于生理传感驱动的急救数字孪生技术以及对应高仿真模拟病人。建立人体多系统仿真数学模型，模拟通换气机制、呼吸系统与血液循环的物质交互模型，模拟药物在体内的代谢过程和起效过程，完成数字孪生生理人模型的关键仿真技术；研究临床真实病例到仿真病例的个性化迁移方法；完成可精确还原患者生命体征及病理指标的数字孪生物理人，具备仿生肺与仿生心脏，可与呼吸机、ECMO、麻醉机等真实医疗设备对接，进行急救干预；研究急救患者与疾病的数字复现技术，突破动态仿真疾病演化和救治发展过程的技术难题；研究适用于救治干预指导的 AR 头显和实时救治引导技术；将相关成果应用于医学培训领域及重点疾病的研究工作中，实现急救精准治疗与训练，开拓医疗大数据应用的新局面。

考核指标：从  $\geq 18$  个维度构建数字孪生患者，孪生模型多维度特征融合程度  $\geq 95\%$ ；模拟人核心力学参数与人体真实值平均

误差  $\leq 10\%$ ，肺阻力和顺应性可调节，顺应性范围为  $15\sim 300\text{mL/cmH}_2\text{O}$ ，调控精度  $\leq 0.5\text{mL/cmH}_2\text{O}$ ；模拟人可进行气体交换，呼出真实的二氧化碳和麻醉气体，呼出气体浓度根据病人病情而变化；完成对医疗救治操作无干涉的轻便型指导 AR 头显，自动对准精度达到位移偏移量  $\leq 20\mu\text{m}$ ，旋转偏移量  $\leq 1^\circ$ ，支持六自由度（6-DoF）头部追踪、手部追踪，定位精度  $\leq 1\text{mm}$ ，单眼分辨率  $1920\times 1080$ ，刷新率  $\geq 30\text{Hz}$ ，视场角  $\geq 40^\circ$ ，支持实时手势交互，虚实融合配准精度  $\leq 1\text{mm}$ ；医疗器械注册证 1 项，相关标准或指南 1 项；进行规模化示范和推广，项目成果运用于  $\geq 800$  家机构，训练人数  $\geq 20$  万人，实现销售收入  $\geq 10$  亿元。

有关说明：由企业牵头，其他经费（包括地方财政经费、单位出资及社会渠道资金等）与中央财政经费比例不低于 3:1。项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：完成 18 维数字孪生患者构建，多维特征融合度  $\geq 95\%$ ；实现 AR 头显自动对准精度达到位移偏移量  $\leq 20\mu\text{m}$ ，旋转偏移量  $\leq 1^\circ$ ，单眼分辨率  $1920\times 1080$ ，刷新率  $\geq 30\text{Hz}$ ，视场角  $\geq 40^\circ$ 。

关键词：生理驱动、数字孪生、模拟人、精准治疗

## **2.5 高精度生物感知觉反馈操纵技术与系统（共性关键技术类，平行设置青年科学家项目）**

研究内容：基于感知觉信息对于生智能操纵的重要作用，为进一步提升生物感知能力、增强机器感知反馈水平，研究生物感知反馈的神经机制，实现生物感知觉信息的有效辨识；发展分



布式信息融合技术、生物感知觉反馈技术，实现人机感知觉信息整合及反馈；开发高精度生物感知觉反馈操纵系统，在生物体及机器系统操纵两个维度上得到应用。

考核指标：研究生物感知反馈的神经机制，发展基于脑电生物信息的高低阶强度感知觉辨识技术，感知觉及其反馈通路覆盖视觉、听觉、触觉，辨识准确率 $\geq 90\%$ ；整合手势、眼动、语音、电生理活动等方式，构建沉浸式生物反馈操纵平台，芯片、操作系统、三维渲染引擎等软硬件国产化率 100%；面向生物体康复的医疗应用典型场景，设计下肢平衡感知反馈控制训练策略和系统，最终实现平衡障碍人群站立态足底压力中心波动区域半径 $\leq 3\text{cm}$ ，量产样机 100 台，完成 50 家相关机构的康复示范应用，病例 $\geq 1000$  例，康复训练有效率 $\geq 70\%$ ；面向机器操纵的特种应用典型场景，设计人机匹配性良好的航天遥操作任务感知觉支持系统，完成舱内外典型航天遥操作任务示范应用 $\geq 30$  人次，完成遥操作任务的情景意识水平提升 $\geq 10\%$ ；获得医疗器械注册证 1 项，申请发明专利 $\geq 10$  项，完成知识产权转移与转化 1 项；至少 2 项先进前沿技术实现首创或达到同类技术的国际领先水平。

有关说明：由企业牵头，其他经费（包括地方财政经费、单位出资及社会渠道资金等）与中央财政经费比例不低于 3:1。项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：研制完成自然操控软件，实现手势、眼动、语音、脑机交互功能；基于下肢平衡反馈控制实现平衡障碍人群站立态足底压力中心波

动区域半径 $\leq 5\text{cm}$ 。

关键词：生物反馈、多模态感知、沉浸式交互、感知觉增强操纵

## 2.6 基于微弱物理场激励的细胞快速响应通讯元件精准构建技术（基础研究类，平行设置青年科学家项目）

研究内容：针对生物体内天然靶点对微弱物理激励响应效能不足的问题，面向微弱物理场激励下介导细胞快速响应通讯的关键生物元件结构信息匮乏、调控机制独特、实验筛选效率低下等现状，研究融合信息技术及生物工程手段，应用人工智能、数据挖掘等技术，解析细胞快速响应元件的功能—结构耦合特征的方法；通过小样本学习方法，基于有限度功能性实验数据，进行元件关键功能域的解析和预测，结合定向进化和虚拟筛选，构建具有超响应特性的生物元件，显著提升重要生命活动调节与控制的精度；发展关键生物元件靶向设计、构建和智能组装，智能化精准构建一批细胞快速响应通信、调控元件及界面，实现基于微弱物理场激励的复杂生命系统高时空精度非侵入控制。

考核指标：针对视听觉感知离子通道等高复杂度膜蛋白生物元件，发展 5 种以上新的分子描述策略，3 种以上智能化计算分析技术，实现在原子尺度对分子量 500kD 以上的膜蛋白复合物功能—结构—序列动态耦合关系进行解析，实现蛋白、膜质、配体等多元素互作用信息的解析，提升 3 类以上响应元件的特

征功能域识别效率；针对离子通道等膜蛋白动力学特征等功能性数据难以借助高通量实验获取的问题，开发基于小样本功能数据和有限度特征信息的功能蛋白人工智能设计及构建技术，在基于大数据学习的基础上，针对细胞通讯元件特定功能进行小样本迭代优化，最终实现在样本量 $\leq 200$ 的功能性数据集条件下，对 $\geq 10^7$ 的序列数据进行发掘和分析，通过定向进化、虚拟筛选和理性设计，进行人工智能指导的响应原件设计和构建，实现关键响应指标至少一个数量级的提升，实现元件设计和构建精度达到 1nm，以及亚 $\mu\text{s}$ 级快速动力特征控制；精准构建 $\geq 50$ 个快速响应通讯元件/界面和调控元件，对至少 2 类阈外刺激或多场耦合刺激具有精确响应特征，至少通过 2 种模式实现对阈外刺激的快速响应；开发安全高效的靶向递送和组装系统，实现细胞及生命体高级功能的 $\mu\text{m}$ 级和 ms 级非侵入调控，形成调控原理样机 1 套，在细胞、组织和整体水平实现应用示范 $\geq 3$ 类。申请发明专利 $\geq 5$ 项。

有关说明：项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：实现对 2 类响应元件特征功能域的有效识别，完成对 200kD 以上的膜蛋白功能—结构—序列动态耦合关系解析，构建快速响应通讯元件/界面和调控元件 $\geq 10$ 个，基于微弱物理场激励的离体调控精度达到 $\mu\text{m}$ 和 ms 精度。

关键词：细胞通讯、生物元件、机器学习、高精度物理调控、小样本学习

## 2.7 基于高性能图计算的生物医学信息增强智能技术及可信应用系统开发（应用示范类）

研究内容：针对目前生物医学领域对海量数据存储及处理的安全性要求高、信息有效关联和融合技术难、数据知识化程度低、无法有效支撑健康医疗行业认知智能化应用等问题，构建面向生物医学信息的“存储算用”一体国产自主架构的安全可信软硬件平台；研究面向生物医学信息的高性能图计算技术，形成定制化高效安全图计算框架，实现数据“存—传—算”三维一体化的可信图计算；研究海量数据关联、融合和共享的机制、基础流程，研究基于规则、统计、深度学习的多源知识抽取和融合方法，研究图数据联邦更新机制，开发群智协同下的知识价值评估方法，形成生物医学领域的认知引擎；面向生物医学领域开展增强语言模型研究，构建大型认知智能模型，基于分布式数值表示，实现知识有效获取和学习，提高模型推理能力；实现在健康医疗行业的示范应用。

考核指标：基于国产 ARM 架构 CPU 处理器（主频 $\geq 2.1\text{GHz}$ ，双路）、麒麟系统（银河麒麟 V10 以上版本）构建面向生物医学信息的“存储算用”一体安全可信平台，核心关键器件国产化率 100%；平台支持数据存储 $\geq 100$  亿条，来源 $\geq 3$  个省、覆盖 $\geq 3000$  万人的健康医疗数据；形成面向生物医学领域的高性能图计算框架，支持 100 计算节点、100T 以上的数据任务计算，支持 50 亿条边的图数据存储和查询，平均查询响应时间为亚秒级；构建 1

套图数据活动机制，形成 3 项图数据处理基础流程；认知引擎在知识抽取、融合、更新、价值评估上形成不少于 10 种方法，覆盖千万级人群，节点达到亿级，边达到十亿级；认知智能模型参数达百亿级，涵盖知识点不少于 1 亿个，基于真实临床案例的知识推理准确度  $\geq 95\%$ ；在健康医疗行业进行推广，并针对临床诊疗分析、临床科研、精准医疗、应急医学和流行病学调查等实现场景示范应用  $\geq 5$  个。

有关说明：由企业牵头，其他经费（包括地方财政经费、单位出资及社会渠道资金等）与中央财政经费比例不低于 3:1。项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：建成存储算用一体化安全可信平台，支持国产 ARM 架构 CPU，核心关键器件国产化率 100%，平台数据存储达 20 亿条，初步形成高性能图计算框架，构建 1 套图数据活动机制，初步形成覆盖百万人群的认知引擎。

关键词：高性能图计算、安全可信架构、认知引擎、增强智能

## **2.8 生物可解释多功能类器官智能系统研究（基础研究类，平行设置青年科学家项目）**

研究内容：针对目前智能系统缺乏可解释性及功能单一等缺陷，借鉴生物系统功能多样性机理，挖掘离子、细胞形态、细胞电生理、网络连接等特性的有机整合对生物系统功能的影响作用，构建“强”仿生多功能类器官智能系统。针对高时一空分辨率、高维度、多模态电生理信号，研究信号在生物系统中实现高效稳

定传递的原理，揭示生物系统提升信息特征分离、抗噪和信息存储容量等性能的生物策略；探索解决前馈生物网络学习过程中面临“信用分配”问题的学习算法；通过逆向工程，还原生物系统中离子、细胞和网络动力学等特性相互作用支撑同一生物系统实现不同功能的原理，构建生物可解释的多功能智能软件系统，包含不同学习算法。

考核指标：生物网络模型包含细胞 $\geq 8$ 类，输出细胞（浦肯野细胞等）含形态结构和 $\geq 8$ 种离子通道；信息传递过程中单细胞传递频率 $\geq 200\text{Hz}$ ，信息特征分离过程信息相似度减少 $\geq 90\%$ ，信息解码精确率 $\geq 90\%$ ；针对前馈网络学习过程中面临的时一空尺度“信用分配”问题，结合编码机制、细胞形态和细胞间连接特性提出具有生物现实性解决方案，并使学习容量达到最优（单细胞存储容量逼近 $5\text{KB}$ ）；基于离子通道 $\rightarrow$ 细胞 $\rightarrow$ 网络连接等数据解析同一网络模型实现精细运动学习、基于价值的决策和异常检测等不同功能的理论基础，并建立生物可解释“强”仿生多功能智能学习软件系统1套，可同时实现上述功能 $\geq 2$ 种，可实现监督式学习及强化学习等学习算法 $\geq 2$ 种；获得发明专利3项以上。

有关说明：项目执行期3年，在项目启动后1年实行里程碑考核。考核指标为：构建完成含离子通道及细胞形态特征的精细生物系统模型，提出信息高效传递和学习算法，实现单细胞模型传递信号频率 $\geq 200\text{Hz}$ 、可学习容量逼近 $5\text{KB}$ ，实现1种高效学习功能。

关键词：生物网络、神经计算、“强”仿生智能、多功能智能

## 2.9 人体组织结构功能特征无创检测和信息提取的关键技术 (共性关键技术类)

研究内容：针对人体组织形态结构和功能的无创在体解析，开展高性能信息检测和应用技术的研发。改进生物信息无创在体检测技术的性能指标，基于自主知识产权的高梯度性能检测系统，实现高时空分辨结构和功能信息提取；研究重大多发性疾病发生发展及康复的结构功能机制及其高灵敏度、高特异性的影像学生物指针；开发可临床落地的医疗人工智能产品。

考核指标：基于高梯度性能人体磁共振成像系统（梯度强度 $\geq 120\text{mT/m}$ ，爬升率 $\geq 200\text{T/m/s}$ ，主磁场强度 $\geq 3\text{T}$ ，孔径 $\geq 60\text{cm}$ ，核心关键部件国产化率 100%），研发高时空分辨的结构功能信息检测技术，实现亚毫米级弥散成像技术和亚秒级功能成像技术，并在不少于 5 家单位完成多中心可行性验证；研发智能化信息提取和分析技术，面向脑卒中、脑小血管病等不少于 3 种重大多发疾病，研发疾病影像学特征快速提取的智能算法，相关智能化分析产品或系统，获得医疗器械注册证 $\geq 1$  项，并在不少于 3 家三甲医院及不少于 10 家基层医院落地使用，总人数 $\geq 500$  人次；开发面向神经康复的影像学生物指针（包括微结构、微状态、连接组等），为卒中后运动障碍患者闭环神经康复提供不少于 5 个可预测训练效果的神经影像标记物，并建立训练效果智能预测模型，在影像学结构功能信息辅助下提升运动意图识别效率，实现主动

意图解码准确率  $\geq 80\%$ 。

有关说明：由企业牵头，其他经费（包括地方财政经费、单位出资及社会渠道资金等）与中央财政经费比例不低于 3:1。项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：基于自主知识产权成像系统，实现亚秒级功能成像技术， $TR \leq 500ms$ ，加速倍数  $\geq 16$  倍，并在不少于 2 家单位完成多中心可行性验证；研发智能化影像学信息提取和分析技术，实现不少于 100 个脑区的秒级快速分割和脑区容积信息获取。

关键词：高时空分辨结构功能信息提取、高性能梯度系统、跨模态多尺度、结构功能耦合、医疗人工智能产品、疾病主动康复

## **2.10 基于结构动力学的合成分子机器调控和再造技术（基础研究类）**

研究内容：探索基于新的非平衡统计物理原理的合成分子机器动力学重建和合成技术，整合四维冷冻电镜和液态电镜、核磁共振弛豫、质谱交联等分析表征技术和结构动力学大数据，发展高性能深度学习、增强采样、虚拟进化的融合智能信息技术，实现利用结构动力学信息进行生物大分子机器虚拟定向进化和人工智能辅助功能优化、再造和设计；面向由蛋白质复合超分子机器，通过基因序列在精准再造设计其功能动力学，并通过动力学重建实验在原子水平加以定量验证；以泛素-蛋白酶体系统，即目前已知的最复杂、最普适的细胞过程调控系统之一为概念验证模式研究对象，结合高精度实验和计算，系统测试和验证上述新技术、



新方法和新概念，构建其全功能网络时空动力学调控合成图谱，并利用相关合成元件再造设计，实现复杂生命系统动力学的精准调控的关键应用示范，形成原子水平合成生命系统动力学的原创核心技术范式。

考核指标：突破目前常规动力学实验及计算分析技术在时空尺度上的局限性，实现在原子尺度对总分子量不小于 0.3 兆道尔顿的生物大分子复合物功能动力学—结构—序列耦合关系进行解析表征，实现通过基因序列再造全功能周期的合成分子机器动力学的高精度算法；围绕原子水平时间分辨多模态动力学重建技术革新，建立不少于 3 项分子合成元件的高精度动力学重建和再造的深度学习、增强采样和虚拟进化新方法，集成不少于 3 种互补实验模态的动力学、动态结构变化的实验数据，设计并合成 5 种以上泛素信号编辑或蛋白质降解调控分子机器，并在 2 种以上其他合成生物分子机器体系上加以推广应用，形成关键应用示范；获得发明专利 5 项以上。

有关说明：项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：初步验证通过基因序列再造全功能执行周期时间尺度内复合超分子元件原子动力学的核心技术方案，初步实现利用原子动力学信息对分子元件定向进化方向的预测和人工智能辅助设计。

关键词：合成分子机器、冷冻电镜、原子动力学高维重建、蛋白质设计

## 2.11 基于活体细胞的复杂生物计算系统设计和构建（共性关键技术类，平行设置青年科学家项目）

研究内容：针对目前硅基系统无法高效开展超大规模并行式逻辑计算以及无法有效应用于活体诊疗应用场景的局限，开展基于细胞的复杂活体生物计算系统的研究，构建分布式并行式新型生物计算体系。以人工设计基因调控网络为基础，开展大规模复杂逻辑运算线路的建模和动态仿真研究，建立多样化的且正交性、模块性高的标准化生物信号处理元器件库。通过构建多细胞群体合作的分布并行式生物计算系统和自动化设计软件，突破单个细胞容纳外源基因元件数量和线路规模有限瓶颈。通过重构数学规则将复杂逻辑运算问题映射为基于基因调控网络的可控生化反应，发展新的生物计算范式。

考核指标：搭建基于活体细胞的复杂生物计算系统，利用人工设计基因调控网络将数学规则映射为可控生化反应，实现复杂逻辑运算。构建 1 个大型多样化的正交基因逻辑门器件库，每个常用正交基因逻辑门类别数量不少于 10 对，逻辑门需产生数字化的响应和高于 10 倍的激活或抑制倍数，且交叉活性低于 10%。构建 4~8 种原核细胞的细胞通讯信号体系和 2~4 种真核细胞的细胞通讯信号体系，实现 20 个逻辑门的多细胞耦合逻辑运算，获得 1~2 种具有逻辑运算和调控的类器官。开发和验证面向客户高级输入需求和使用友好的生物线路自动化设计和诊断软件平台 1 个。实现细胞计算在活体诊疗场景的应用展示案例 1 个。获得不

少于 3 项技术发明专利。

有关说明：项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：构建 1 个大型多样化的正交基因逻辑门器件库，实现每个常用正交基因逻辑门类别数量不少于 10 对，逻辑门需产生数字化的响应和高于 10 倍的激活或抑制倍数，且交叉活性低于 10%。实现 20 个逻辑门的多细胞耦合逻辑计算。

关键词：生物计算、人工细胞网络、细胞群体计算、正交逻辑门

### **3. BT-IT 融合技术的健康医药场景应用**

#### **3.1 数字化人体全景表型建模与数字孪生技术（基础研究类）**

研究内容：研究分子、细胞、功能、影像等多尺度、跨时空人体表型信息采集、分析与可视化技术；通过信息采集、存储、处理、共享、分析等多个环节，制定功能互联、格式规范、操作便捷的健康人体全身表型信息数据库；探索人体微观层面的分子信息、代谢、微生物、生理功能与全身各组织器官的结构影像、分子影像之间的关联机制，实现分子、细胞、代谢、功能、影像等表型变量（包括 MRI 心脏和大血管功能精准检测、肌力检测及超声心动图心脏及大血管功能检测、超声颈动脉血流动力学检测等）间关系的时空动态可视化建模；开发人体影像神经辐射场等深度学习技术并对人体影像组学精细刻画，建立数字化人体全景表型呈现系统并形成高维表型数据动态演变插值的渲染引擎；在健康医疗领域开展相关系统的验证，辅助临床教学与诊疗等实际

应用。

考核指标：在同一研究中心完成至少 1000 例健康人体及肿瘤、神经系统疾病人体表型全景数据的基线采集与跟踪随访两次，每人每次采集表型包括分子、细胞、功能和全身各器官影像等 4 类表型数据不低于 50000 个特征，人体表型全景数据采集随访期超过 3 年，并提出不少于 5 种多尺度表型参数之间的关联分析方法；构建针对不少于 10 种针对肿瘤及神经系统疾病的分子及 PSMA、Tau 及 A $\beta$  蛋白等影像标志物；提出人体表型全景数据库建设规范与评价规范，形成 6 项以上标准规范，申请或获得不少于 5 项核心发明专利；构建 1 套可面向临床的人体表型分析与健康人体数字孪生系统，可体现分子、细胞、功能、影像等表型之间的动态关系；统一规范与标准，在 3 家以上医疗机构进行该人体表型数字化呈现系统不少于 1 万人的实际应用与临床验证。

有关说明：项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：建立不少于 1000 人的多尺度表型数据不低于 50000 个特征)，表型数据可视化渲染引擎的交互响应时间低于 2s，支撑亿级体量信息化数据。

关键词：表型组学、分子影像学、计算生物学、数字化人体

### **3.2 细胞命运决定的多层次动态调控理论和数字仿真（基础研究类）**

研究内容：利用单细胞多组学和影像大数据与体内/外细胞可控分化系统，发展相应数据深度挖掘和仿真技术，解析细胞命运

决定的多层次（包括但不限于转录因子，染色质高级结构，液—液相分离）调控关系，构建单细胞和细胞群体跨层级调控网络和知识图谱，揭示决定血液系统发育，血液肿瘤、卵巢癌变或其他生物转变命运关键环节的系统动力学机制和群体动力学涌现机制，搭建形成细胞命运决定的大数据资源库和仿真应用平台，针对性开发癌前诊治新技术。

考核指标：开发 3 种以上具有自主知识产权的鉴定细胞命运关键状态及重要调控基因的统计理论或机器学习算法，系统构建 3~5 个重要细胞谱系命运选择的系统动力学和群体动力学仿真模型和知识图谱，针对血液系统发育，血液肿瘤、卵巢癌变或其他生物过程建立大数据资源库 3 个或以上，实现 1~2 套交互式可视化分析、仿真平台，系统预测潜在的细胞转变调控新途径，实验验证和系统评估数字模型的合理性；利用单细胞多组学和多模态健康大数据 BT-IT 融合的前沿交叉技术，在血液系统发育、血液肿瘤或者卵巢癌变等过程的科研和临床领域实现原创理论的突破和具有自主知识产权的创新技术转化应用 2~3 项。

关键词：单细胞、多组学、生物信息、数据挖掘

### **3.3 数据赋能的人脑多尺度数字孪生系统（共性关键技术类）**

研究内容：基于跨尺度多维度大数据，建立“分子→细胞→环路”的人脑多尺度数字孪生系统。微观水平上，基于单细胞与空间转录组技术，研发空间多组学数据整合的人工智能算法，构建亚细胞级人脑细胞数字通讯网络。基于超高场靶向磁

共振代谢成像新技术，研发活体鉴定人类大脑中代谢物质类型与活性的成像技术，实现大脑中代谢物图谱的系统绘制；介观水平上，基于高梯度场磁共振弥散技术探测大脑水分子运动，构建亚毫米级无变形全脑结构连接图谱；宏观水平上，构建超高场全脑动静脉静态图谱以及呼吸和心跳关联全脑血流动态图谱；在核心技术方面，研发新型人工智能算法，整合基因、影像、代谢、细胞、连接组等跨尺度多维度数据，实现不同尺度数字解码人脑系统。

考核指标：在微观水平实现基于基因时空表达的亚细胞级细胞连接图谱，实现的脑区数量 $\geq 5$ ，实现活体无创磁共振靶向代谢物测量技术，靶向代谢分子种类 $\geq 10$ 种，各代谢分子信号选择特异性达到100%，检测敏感度优于1mmol/L；在介观水平实现B值 $\geq 10000\text{s/mm}^2$ 的亚毫米级分辨率弥散成像技术，整体形变小于5mm，实现亚毫米级介观水平无变形全脑结构连接图谱；在宏观水平建立椎动脉到豆纹动脉三级分支等动静脉的全脑磁共振动静脉静态图谱；实现呼吸和心跳关联全脑血流动力学分析技术；开发一套整合跨尺度、多模态数据的大脑数字孪生系统。

有关说明：项目执行期3年，在项目启动后1年实行里程碑考核。考核指标为：发展磁共振靶向代谢物测量技术，检测敏感度小于1mmol；实现亚毫米级介观水平无变形全脑结构连接图谱，整体形变小于5mm。

关键词：生物信息学、计算生物学、数字脑、人工智能

### 3.4 支撑创新药靶发现与分子生成的新一代 AI 技术研发及应用（应用示范类）

研究内容：研发数据知识双驱动的新一代人工智能制药技术，超越纯数据驱动的深度学习，把大数据驱动的机器学习系统与医学、药学、生物和化学知识推理结合，针对高医疗与社会成本的重大慢性疾病，推动创新药物的药靶发现与分子生成。研发多模态人工智能技术，整合文献与专利与临床资料中的生物通道、蛋白结构、西药化学知识及中药活性成分知识；解码基因、RNA、蛋白、代谢等的信号调节和功能通路及调控网络，构建疾病—药靶—干预分子因果图谱；研发知识启发的图神经网络因果推理技术，理解致病通路并挖掘新药靶；研发物理引导的人工智能加速分子动力学计算技术，预测分子—蛋白亲和力；开发知识启发的人工智能分子大模型，对药效/药代/毒性进行预测；研发专业知识约束的生成式 AI 模型，结合表型、蛋白结构和 RNA 结构建设 AI 分子虚筛和优化技术体系；集成以上技术，针对 2~3 个重大疾病新颖分子靶标，开展大规模药物筛选与设计，结合药效/药代/毒性预测模型等 AI 成药性预测技术，发现并确证具有潜在临床应用价值的原创候选化合物。

考核指标：建成数据知识双驱动的新一代人工智能制药新技术体系；开发 1 套多模态人工智能专利和文献理解技术；整合生物通道、蛋白结构、西药化学知识及中药活性成分知识，构建 1 个因果图谱，开发 1 套数据与知识结合的深度学习与因果推理技

术，构建 5 种以上疾病致病机制与通路，挖掘 3 个以上新药靶；开发 1 套包含知识启发的人工智能分子大模型、物理引导的 AI 加速分子动力学计算模型、知识约束的生成式 AI 模型、分子—蛋白亲和力预测、药效/药代/毒性预测模型等技术的重大慢性疾病分子虚筛和优化技术体系。针对重大慢性疾病药物研发，构建 1 套人工智能驱动的药靶发现与分子生成系统，在 5 家科研机构部署验证。发现 10 个具有良好临床潜力的慢性病候选化合物。

有关说明：鼓励产学研联合申报。项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：开发 1 套多模态人工智能专利和文献理解技术；构建 1 个因果图谱，开发 1 套知识启发的图神经网络因果推理技术，构建 1 种以上疾病致病机制与通路，挖掘 1 个以上新药靶；（项目启动后 2 年）：开发 1 套包含知识启发的人工智能分子大模型、物理引导的 AI 加速分子动力学计算模型、知识约束的生成式 AI 模型、分子—蛋白亲和力预测、药效/药代/毒性预测模型等技术的重大慢性疾病分子虚筛和优化技术体系。

关键词：药靶 AI 发现、可解释性与因果分析、生成式模型、AI 药效/药代预测

### **3.5 基于靶标组学的复方用药 AI 药效评估与智能推荐系统（基础研究类）**

研究内容：联合用药可达到增效减毒效果，也是扩大药物适用疾病谱、延缓抗药性的重要手段。基于前期可药靶性识别和组



学数据预处理与分析的 AI 算法，结合药物代谢、药物靶标和药物—药物相互作用等数据，发现不良反应发生的分子靶标机制和具有协同疗效的药靶组合；基于前期化药—化药、化药—天然产物、化药—抗体的癌症协同用药普适性智能推荐算法，结合个性化临床组学、类器官、PDX 和真实世界大数据，为恶性肿瘤复方用药开发协同药效剂量配比智能测算模型；基于以上，结合疾病分子分型、药物—靶点互作、个体组学特征、不良反应和药物经济性等多维变量，研发联合用药药效 AI 评估系统，支持创新药、新靶点药物二次开发，发现若干有潜力的协同用药组合；建立协同药效临床智能推荐系统，应用于肝癌、肠癌、胃癌等重大疾病开展临床研究。

考核指标：开发 1 套联合用药协同药效计算模型，1 套联合用药不良反应 AI 预测模型；开发 1 套协同药效剂量配比测算模型，并在动物和临床水平上验证其有效性；基于对抗癌药物组合剂量药效的综合评价，开发 1 套癌症联合用药临床智能推荐系统，并在动物水平上验证其有效性，开展包括癌症组学标志物等在内的临床测试不少于 30 例；应用于管线创新药，研发具备临床应用潜力的协同用药组合 3~5 项，完成相应的临床前研究。

有关说明：项目执行期 3 年，在项目启动后 1 年实行里程碑考核。考核指标为：完成抗癌复方协同药效模型建立，启动临床案例研究；完成对协同药效剂量配比测算模型的开发，取得动物水平验证；为管线创新药物计算预测出协同药物组合不低于 5 组。

关键词：生物信息、联合用药、协同药效、药物靶标

### 3.6 疾病靶标可干预位点的智能识别与新药研发技术（共性关键技术类）

研究内容：整合重大疾病生物大分子位点修饰化、致病突变、内外源分子干预等数据，研究建立生物大分子位点干预与功能的知识关联图谱，生成标准化功能性变构位点数据集；利用大数据和人工智能，构建变构特异性分子力场和增强采样等计算方法识别靶标中可满足疾病干预需求的新位点，发展靶向变构位点动态构象的虚拟筛选和评价函数技术，构建成为基于人工智能和大数据的变构药物设计体系；以肿瘤、代谢性疾病等难成药靶标为例，识别新位点并发现活性分子以实现临床干预所需的激动调控、特异性提高或耐药性降低等功能，阐明位点的新功能机制并确证具有首创特征的新位点先导化合物。

考核指标：建立百万量级生物大分子位点与功能的知识图谱和领域标准数据集，开发功能依赖的生物大分子新位点识别技术，建成新位点识别与活性分子发现一体化的智能药物设计平台，在30家以上国内外药物研发中心推广应用：针对亟需激活抑癌基因、阻断蛋白互作、克服耐药等类型的重大疾病难成药靶标，获得10个以上实现临床需求的新位点，基于新位点发现3~5类首创药物先导分子，1~2个原创新药候选进入临床前研究。

关键词：生物信息、计算机辅助药物设计、新靶标、变构位点