

2026 年度拟提名陕西省自然科学奖项目公示内容

一、项目名称

转录组异构生物数据的智能融合与解析理论研究

二、提名者及提名意见

提名者：陕西省融办

提名意见：该项目面向人民生命健康和人工智能国家重大战略需求，聚焦转录组组学数据的智能融合与解析，顺应多源多尺度生物大数据精确解析的发展趋势，着力攻克了异构生物数据跨模态融合建模、一致化精确表征和生物分子活动智能解析等核心挑战，在异构生物数据智能融合与解析的理论创新方面取得了系统性、原创性成果。项目提出了面向大规模转录组生物数据的新型融合表示理论，构建了基于复杂网络与路径推理的关联解析新体系，揭示了“网络邻居”效应，为发现未知的生物分子功能联系提供了全新的计算视角；创立了面向多类任务的深度学习框架，显著提升了对 RNA 调控机制、药物作用机制和疾病发生机制三个层次的复杂生物现象的泛化解析能力。研究成果对推动新药研发、发现疾病生物标志物、阐明复杂疾病致病机理和揭示生物学新机制具有重要的学术价值和理论意义。相关成果得到了国内外学者的广泛引用和积极评价，推动了计算生物学与人工智能的交叉融合发展。该项目理论创新显著，方法体系先进，处于国际前沿水平，为生物医学大数据分析提供了重要的理论工具，对学科建设和经济社会发展具有重要的指导作用。

成果材料齐全、规范，无知识产权纠纷，人员排序无争议，符合陕西省自然科学奖提名条件。特提名为陕西省自然科学奖一等奖。

三、项目简介

该项目属计算机科学与生物医学交叉研究领域。

随着高通量测序技术的飞速发展，转录组学研究催生了海量多源异构 RNA 数据，涵盖以 RNA 分子为核心的复杂生物关系网络等。这些数据在类型、维度和尺度上呈现显著差异，构成了典型的“异构大数据”体系。如何对转录组异构数据进行智能融合与深度解析，以揭示复杂生命活动规律、加速疾病机理研究和药物发现，是当前生物信息学与人工智能领域共同面临的重大科学挑战。

传统计算方法多局限于单一类型数据处理，难以有效整合多源信息，且普遍依赖人工特征工程，模型泛化能力受限。因此，亟需发展能够自动融合异构数据、挖掘深层关联关系的新型理论模型与算法。这不仅具有重要的理论价值，也对推动精准医学和生物医药产业进步具有迫切的现实意义。

该项目在国家自然科学基金、科技部重点研发计划等国家级和省部级项目的

资助下,历时 10 年探究求索,围绕异构生物数据智能融合与解析核心科学问题,开展了系统性的理论研究,取得具有重要科学意义和国际影响的创新性成果。重要科学发现包括以下三个方面:

面向药物作用发现场景的智能计算研究:针对异质性生物关系网络的结构特性,项目提出新型深度学习计算范式,赋能**单药耐药机制**阐释与**多药联用安全性**智能计算。代表性论文[1]提出了多通道特征融合模型来处理多类型药物-药物相互作用预测,建立了基于生物医学知识图谱结合胶囊网络与注意力接收域的计算模型。该模型能并行提取药物的分子结构、靶点信息等多种特征并进行深度融合,一举解决多种类型药物互作的预测问题。代表性论文[2]针对非编码 RNA 分子的链长度差异大、已知关系稀缺的问题,提出了一个基于异构关系网络的通用卷积计算框架,用于分子关联网络中多样化的生物分子关系预测,体现了项目理论方法的通用性和扩展性。

面向疾病发生机制发现场景的智能计算研究:针对**环状**和**超短链**的两种 RNA 分子结构,实现面向 800 多种人类复杂疾病亚型的高可靠性生物标志物识别。在代表性论文[3]中,利用深度堆栈自编码器整合了 circRNA、疾病的多源相似性信息,实现了降维与特征融合,显著提升了 circRNA-疾病关联预测精度。在代表性论文[4]中,创新性地将自然语言处理技术引入 microRNA 序列特征提取,融合功能相似性、语义相似性与关联网络等多源异构生物信息,构建逻辑模型树分类器,最终实现面向数百种疾病的 RNA 标志物的精准预测。

面向 RNA 调控机制发现场景的智能计算研究:针对异构 RNA 分子的极大物理化学结构异质性,项目提出了基于深度学习的统一表征框架,实现对特定 RNA 分子的调控关系预测。项目将全局生物网络拓扑信息引入异构数据智能解析。代表性论文[5]开创性地提出了路径计算模型,它不依赖于已知关联,而是通过搜索分子互作网络中所有路径来衡量分子对之间的关联强度,利用双向的网络信息传播计算策略揭示了异构 RNA 分子之间的潜在生物功能联系,为发掘 RNA 分子的新型生物功能角色提供了强有力的计算工具。

四、客观评价

该项目 5 篇代表性论文发表在信息科学领域顶级期刊 **IEEE Transactions on Knowledge and Data Engineering**、**IEEE Transactions on Cybernetics** 及生物学领域顶级期刊 **Bioinformatics** 等期刊上,科学发现原创性及学术价值得到国内外专家广泛引用和积极评价,项目培养国家级领军人才 1 人,国家级青年人才 2 人次。研究成果被 **Nature Communications**(IF:15.7)、**Molecular Cancer**(IF:33.9)、**Information Fusion** (IF:15.5) 等权威期刊正面评价和引用 465 次,包括美国、英国、德国、加拿大等 20 余个国家和地区 80 余位国内外院士、数授等著名学者:

➤ 代表性论文[1]受到发表在 **Nature Communications**、**Information Fusion**、**IEEE Transactions on Knowledge and Data Engineering**、**IEEE Transactions on Industrial**

Informatics 等顶级期刊的广泛引用。

中国工程院院士、国家自然科学基金创新研究群体学术带头人、中南大学学术委员会主任、中国自动化学会副理事长桂卫华教授在中科院 1 区期刊 IEEE Transactions on Industrial Informatics 上发表的文章 A rolling bearing fault diagnosis method based on multimodal knowledge graph 中将我们的模型作为性能标杆。

浙江大学求是特聘教授、爱思唯尔中国高被引学者侯廷军教授在 Nature Communications 上发表的文章 ClickGen: Directed exploration of synthesizable chemical space via modular reactions and reinforcement learning 中将我们的方法作为计算化学中负采样策略的一个成熟范例。

ACM Fellow, IEEE Fellow, 国际神经网络学会会士, AAIA Fellow, 香港中文大学教育副校长 Irwin King 教授在中科院 1 区 TOP 期刊 Information Fusion 上发表的文章 A survey on graph embedding techniques for biomedical data: Methods and applications 中称赞我们基于概率的负样本策略成功地生成了高质量负样本、避免了引入假阴性样本,并开发了基于胶囊图神经网络的模型来处理生物学知识图谱中的多关系数据。

复旦大学特聘教授、CCF 人工智能与模式识别专业委员会的常务委员窦景德教授在中科院 1 区期刊 IEEE Transactions on Knowledge and Data Engineering 上发表的文章 Giant: Protein-ligand binding affinity prediction via geometry-aware interactive graph neural network 中高度评价我们的工作,称赞我们在药物发现应用中取得了显著的成功,突出了数据驱动的计算方法在解决基于药物的数据挖掘和知识发现的复杂问题方面的有效性。

国家杰青, Distinguished Fellow of IETI, 国家优秀青年科学基金获得者, 东南大学耿新教授在 IEEE Transactions on Neural Networks and Learning Systems 发表的文章 Scalable label distribution learning for multi-label classification 评价我们的工作有效挖掘不同标签的底层信息的积极探索。

➤ 代表性论文[2]受到发表在 Briefings in Bioinformatics 和 Bioinformatics 等顶尖期刊的广泛引用。

中南大学计算机学院原院长、国务院特殊津贴专家王建新教授,在其发表在 Briefings in Bioinformatics 期刊上的论文 NASMDR: a framework for miRNA-drug resistance prediction using efficient neural architecture search and graph isomorphism networks 中评价代表性论文[2]所提出的 GCMDR 方法是 miRNA-药物抗药性关联预测这一少有研究的领域中的首个方法。

南洋理工大学 Chee Keong Kwoh 教授,在其发表在 Bioinformatics 期刊上 Predicting human microbe-drug associations via graph convolutional network with conditional random field 的论文中评价代表性论文[2]所利用的图神经网络模型是有前景的机器学习方法,具有优异的处理图类型输入数据的能力,能成功解决 miRNA 相关的抗药性预测问题。

国家级领军人才、军事科学院军事医学研究院伯晓晨研究员,在其发表在

Wiley Interdisciplinary Reviews: RNA 期刊上的论文 A comprehensive checklist of computational resources and methodologies for noncoding RNAs in systems medicine 中评价代表性论文[2]提出的 GCMDR 方法能正确地利用多源信息数据，为药物设计和治疗评估提供有用的见解。

➤ 代表性论文[3]受到发表在 Molecular Cancer、IEEE Transactions on Cybernetics、Information Sciences 等顶尖期刊的广泛引用。

英国皇家工程院院士，University of Southampton 荣誉教授，IET 高级成就奖章及法拉第奖章获得者 Chris J. Harris 教授在中科院 1 区 TOP 期刊 IEEE Transactions on Cybernetics 上发表的文章 Adaptive Multioutput Gradient RBF Tracker for Nonlinear and Nonstationary Regression 中将我们的模型作为最先进的方法之一进行对比。

美国医学与生物工程院院士、俄罗斯工程院外籍院士、乌克兰工程院外籍院士、欧洲科学与艺术院院士、英国皇家公共卫生院院士，国家特聘专家，国家级高层次人才，全球前 2% 顶尖科学家，世界顶尖 1000 名计算机科学家潘毅教授在 Information Sciences 发表文章 Identify potential circRNA-disease associations through a multi-objective evolutionary algorithm，高度评价我们的方法在训练其旋转森林后，该方法的预测准确度是能充分满足计算需求的。

美国哈佛大学麻省总医院访问学者，中南大学二级教授李小荣教授在 Molecular Cancer 发表文章 Comprehensive landscape and future perspectives of circular RNAs in colorectal cancer，高度评价我们的工作——开发了一个名为 IMS-CDA 的在线数据库，用于环状 RNA 的鉴定、表征和功能分析。

国家杰出青年基金获得者、ESI 高被引科学家、上海交通大学沈红斌教授在计算生物学学科 TOP 期刊 Briefings in Bioinformatics 上发表文章 MDGF-MCEC: a Multi-view Dual Attention Embedding Model with Cooperative Ensemble Learning for CircRNA-disease Association Prediction，对我们的工作进行详细评价——我们的方法提高了发现环状 RNA 与疾病之间关联的效率，是生物信息学领域的重要研究方法，可为进一步的生物湿实验提供理论指导。

以上国内外同行的公开学术评价彰显了本项目成果的学术原创性和重要科学价值。

五、代表性论文专著目录（不超过 8 条，其中代表性论文不超过 5 篇，代表性专著不超过 3 部）

（按照表格所示栏目填写支撑本项目重要科学发现的代表性论文专著详细情况，不超过 8 篇，按重要程度排序。所列论文专著应公开发表 2 年以上即 2024 年 8 月 1 日以前公开发表。所列代表作及论文应以省内单位或个人为主要完成单位，署名第一单位（标号为 1 的单位）应为国内单位。

“作者”、“通讯作者（含共同通讯作者）”、“第一作者（含共同第一作者）”和“国内作者”，均应基于论文的全部作者进行填写，不得只填写本项目完成人或少填漏填。

其中，“作者”、“通讯作者（含共同通讯作者）”和“第一作者（含共同第一作者）”的姓名表述应与论文原文的署名保持一致，“国内作者”填写作者的中文姓名。

该表所列论文专著的知识产权归国内所有且无争议，未曾在往年国家科学技术奖励项目、往年省部级（政府）科学技术奖励项目和本年度其他陕西省科学技术奖提名项目中作为支撑材料出现。用于提名陕西省科学技术奖的情况，已征得未列入项目主要完成人和主要完成单位的作者的同意，其中，未列入项目主要完成人的第一作者、通讯作者（含共同第一作者、共同通讯作者）已出具知情同意书面签字意见，与其他作者的有关知情证明材料均存档备查。）

序号	论文专著 名称	刊名	作者	年卷页码 （xx 年 xx 卷 xx 页）	发表 时间	通讯作 者	第一作 者	国内作 者	他 引 总 次 数	检索数 据库	知识 产权 是否 归 国 内 所 有
----	------------	----	----	--------------------------------	----------	----------	----------	----------	-----------------------	-----------	---

1	Biomedical Knowledge Graph Embedding With Capsule Network for Multi-Label Drug-Drug Interaction Prediction	IEEE Transaction s on Knowledge and Data Engineering	Xiaorui Su, Zhuhong You, Deshuang Huang, Lei Wang, Long Wong, Boya Ji, Bowei Zhao	2023 年 35 卷 6 期 5640-5651 页	2022 年 2 月 28 日	Zhuhong You, Deshuan g Huang, Lei Wang	XiaoRui Su	苏 小 芮, 尤 著宏, 黄德 双, 王 磊, 黄 利广, 姬 博亚, 赵	92	SCIE, Web of Sciences	是
2	Graph convolution for predicting associations between miRNA and drug resistance	Bioinformat ics	Yu-An Huang, Pengwei Hu, Keith C. C. Chan, Zhu-Hong You	2020 年 36 卷 851-858 页	2019 年 8 月 9 日	Yu-an Huang, Zhu-Hong You	Yu-an Huang, Pengwei Hu	黄裕 安, 胡 鹏伟, 陈振 冲, 尤 著宏	105	SCIE, Web of Sciences	是
3	IMS-CDA: Prediction of CircRNA-disease Associations from the Integration of Multi-source Similarity Information with Deep Stacked Auto-encoder Model	IEEE Transaction s on Cybernetics	Lei Wang, Zhu-Hong You, Jian-Qiang Li, Yu-An Huang	2021 51 卷 11 期 5522-5531 页	2020 年 10 月 7 日	Zhu-Hong You, Lei Wang	Lei Wang	王磊, 尤著 宏, 李 坚强, 黄裕安	63	SCIE, Web of Sciences	是

4	LMTRDA: Using logistic model tree to predict MiRNA-disease associations by fusing multisource information of sequences and similarities	PLOS Computational Biology	Lei Wang, Zhu-Hong You, Xing Chen, Yang-Ming Li, Yan-Nan Dong, Li-Ping Li, Kai Zheng	2019 年 15 卷 3 期 1006865 页	2019 年 3 月 27 日	Zhu-Hong You, Xing Chen	Lei Wang, Zhu-Hong You	王磊, 尤著宏, 陈兴, 李阳明, 李丽萍, 董雅楠, 郑恺	106	SCIE, Web of Sciences	是
5	Constructing prediction models from expression profiles for large scale lncRNA-miRNA interaction profiling	Bioinformatics	Yu-An Huang, Keith C. C. Chan, Zhu-Hong You	2018 年 34 卷 812-819 页	2017 年 10 月 23 日	Yu-An Huang, Zhu-Hong You	Yu-An Huang	黄裕安, 陈振冲, 尤著宏	99	SCIE, Web of Sciences	是
6											
7											
8											
合 计									465		

六、主要完成人情况（不超过 6 人）

（所列完成人应为在陕个人，或与在陕个人合作的我国其他地域的个人（**第一完成人必须为全职在陕的个人**），且是“代表性论文专著”主要学术思想的提出者，并在“代表性论文专著”中有署名。应按表格要求逐项填写。附件所列验收、鉴定的专家组成员不能作为完成人。同一人同一年度只能作为一个提名项目的完成人参加陕西省科技奖的评审。附件所列验收、鉴定的专家组成员不能作为完成人。

工作单位：根据人事关系填写完成人现工作的单位，已退休的填写退休前的工作单位。

完成单位：填写完成人参与本项目主要研究工作时所在单位，应为国内法人单位。如涉及多个单位，应根据贡献大小填写一个单位。完成单位与奖励证书关联，请根据实际情况审慎填写。

对本项目贡献：不超过 300 字。应具体写明完成人对本项目做出的实质性贡献，并注明代表性论文专著编号。填报时括号部分内容删除。）

姓名	排名	行政职务	技术职称	工作单位	完成单位	对本项目贡献
尤著宏	1	无	教授	西北工业大学	西北工业大学	作为项目的总负责人，提出了“转录组异构生物数据智能融合与解析”的总体研究思路和理论框架。主导了核心算法的设计，是代表性论文[1]、[2]、[3]、[4]、[5]的通讯作者，对论文的学术思想、关键技术和结论起了决定性作用。全面负责项目的规划、组织和协调。
黄裕安	2	无	教授	西北工业大学	西北工业大学	作为项目核心算法的共同设计者和主要发展者，主要负责基于网络拓扑和路径推理的模型构建，是代表性论文[2]、[5]的第一作者及代表性论文[3]的主要作者。为项目在多模态特征融

						合方向做出了关键性贡献。参与了项目的总体设计讨论。
王磊	3	无	教授	中国矿业大学	中国矿业大学	作为项目中深度学习模型构建与实验验证的主要完成人，主要负责深度表征学习模型的实现、优化和大量计算实验。是代表性论文[3]、[4]的第一作者，及代表性论文[1]的主要作者。同时参与了部分算法实现和实验分析工作，为项目提供了坚实的数据处理和计算支撑。
胡鹏伟	4	无	研究员	中国科学院新疆理化技术研究所	中国科学院新疆理化技术研究所	作为项目中多模态学习理论和应用拓展的主要完成人，主要负责多模态特征融合的统一深度学习架构的构建、优化和验证。是代表性论文[2]的主要作者，在生物分子关系预测的通用学习框架构建中发挥了关键作用。
	5					
	6					

七、主要完成单位情况（不超过 3 个）

完成单位	排名	对本项目主要贡献（限 600 字）
西北工业大学	1	西北工业大学作为本项目的牵头组织和主要实施单位，全面负责项目的规划、管理和资源协调。学校将人工智能与生物信息交叉研究列为重点发展方向，为本项目的顺利实施提供了坚实的政策支持和资源保障。
中国矿业大学	2	中国矿业大学作为本项目的重要参与单位和成果应用验证基地，在生物医学数据的精准标注、算法模型的生物学验证以及推动成果向临床转化方面发挥了不可或缺的关键作用。
中国科学院新疆理化技术研究所	3	中国科学院新疆理化技术研究所是本项目重要协同攻关和技术支撑单位，在关键理论验证、核心算法研发方面做出重要贡献，对项目提出的算法预测结果进行了多轮实验验证，将计算预测与生物学功能相结合，显著提升了项目成果的可靠性和科学价值。

八、完成人合作关系说明

尤著宏、黄裕安等共同发表了代表性论文 2、3、5 在内的 30 余篇研究论文；

尤著宏、王磊等共同发表了代表性论文 1、3、4 在内的 20 余篇研究论文；

尤著宏、胡鹏伟等共同发表了代表性论文 2 在内的 10 余篇研究论文。