**国家自然科学基金委员会--广东省人民政府大数据科学研究中心项目2016年度项目指南**

**一、设立宗旨**

　　根据国家和广东省科技发展战略需要，为充分发挥国家自然科学基金的导向作用以及广东省数据和计算资源的优势，引领全国大数据科学领域的基础研究，促进大数据产业的发展，国家自然科学基金委员会和广东省人民政府于2015年签署协议，设立国家自然科学基金委员会-广东省人民政府大数据科学研究中心项目（以下简称大数据项目）。

　　大数据项目围绕“智慧城市”建设，设置智能交通、智慧医疗与健康、智慧安全、智慧防灾、智慧金融、智慧教育和智慧管理等研究领域，汇聚国内大数据源头创新领域人才和科技资源，共同解决大数据科学领域的重大科学问题和技术问题。

**二、实施原则**

　　大数据项目是国家自然科学基金的组成部分，以国家超级计算广州中心“天河二号”超级计算机为平台，面向全国，公平竞争，分批择优支持。大数据项目强调学科交叉和强强联合，鼓励全国有大数据研究优势的依托单位与广东省高等院校和科研机构联合提出项目申请。

　　大数据项目申请、评审、管理和资金使用按照《国家自然科学基金条例》《国家自然科学基金联合基金项目管理办法》和《国家自然科学基金资助项目资金管理办法》等有关规定执行。

**三、2016年度资助计划、资助领域和研究方向**

　　2016年度大数据项目拟安排直接费用4910万元，拟资助如下领域的“中心项目”或“重点支持项目”。

**资助领域一、视频大数据高效表达、深度分析与综合利用。**

瞄准国家和社会公共安全的重大战略需求，突破图像视频大数据的高效表达、特征编码、对象识别、行为分析等瓶颈问题，在天河二号超级计算机上实现百亿幅以上图像大数据和十万路以上视频大数据的深度解析和关联分析，并在省级公安系统进行大规模应用示范。

　　1.图像视频大数据中车辆的实时搜索与精准识别。

　　深度挖掘海量卡口、电子警察图像和监控视频中蕴含的车辆信息，对千万规模车辆的视觉特征进行动态建模和精准识别。在天河二号超级计算机上构建大型图像搜索引擎，支持百亿幅以上车辆图像的高效索引、车辆的实时检测和跟踪与识别，识别精度和检索效率国际领先。

　　2.视频大数据高效表达与处理引擎。

　　面向省级百万路以上监控视频的高质量存储和高效率传输，研究监控视频的高效表达方法和分析识别算法评测方法，建立视频大数据分析识别算法库，构建高清视频实时处理引擎，支持千路高清晰度视频的实时编码、特征提取和结构化语义描述，在天河二号超级计算机实现十万路视频的实时分析和集成验证。

　　3.时空关联的监控视频对象及行为识别。

　　针对省域视频大数据中蕴含的目标对象（行人、物体、车辆等），研究基于时空关联分析的视频大数据协同表示理论与方法，按照准实时要求，设计高性能计算模型和在线学习系统，支持十万级规模特定视觉对象的增强、识别、重检、追踪和行为分析。

　　4.群智感知与多源异构跨媒体融合的视觉大数据分析与预测。

　　以揭示多源、群体交互环境下跨媒体历史数据和现实数据之间的相互映射机制与信息流动机理为目标，研究以“人-地-事”为核心的跨媒体信息融合关联和群体行为预测预警方法，针对十万特定人的社会活动与行为记录，建立多源异构跨媒体信息融合联动的大数据分析和预测系统。

　　5.视频大数据关键技术验证平台与综合应用示范。

　　汇集视频大数据处理分析方面的研究成果，依托天河二号超级计算机平台和行业云平台，创建支持快速应用功能开发的统一软件平台和编程接口，构建视频大数据关键技术验证平台，支持全省范围图像视频大数据的实时汇聚、同步分析、特定对象识别、事件分析挖掘和预测预警，在5个以上地级市进行应用示范。

　　本资助领域拟资助1个中心项目，资助直接费用2000-3000万元，资助期限为4年，申请书中的研究期限应当填写“2017年1月1日-2020年12月31日”。

　　中心项目的主要研究内容应当涵盖本资助领域下的5个研究方向。围绕研究目标，研究内容应相互关联，互为支撑。人员结构和规模合理，包括若干优秀的学术带头人、高素质研究骨干、高水平技术人员和精干的管理人员，在国内本领域中处于优势地位。**合作研究单位数量根据实际情况确定。**

**资助领域二、面向生物医疗健康大数据分析挖掘的基础理论与关键技术。**

面向广东地区及我国城市化过程中典型的居民医疗与健康的需求，针对生物医疗健康大数据的多源、异构、动态、海量等特点，基于天河二号超级计算机，研究生物医疗健康大数据的收集、分析、挖掘与利用的理论和技术，为揭示复杂、典型疾病及健康问题的发生机理及干预机制提供理论和技术支持；本项目以天河二号超级计算机为平台，研制的算法、软件及支撑工具应在天河二号超级计算机平台上予以实现。主要研究方向包括：

　　1.基于复杂疾病及居民健康大数据的疾病及健康问题预警、预测及干预模型和方法的研究。

　　基于复杂疾病及居民健康大数据，研究重大疾病及健康问题（如恶性肿瘤、重大传染病等）发生机制和干预方法。基于居民健康基本信息及重大疾病大数据，研究居民重大疾病与相关健康问题相互关系、发生机制以及有效干预方法；挖掘重大疾病与相关健康问题的演化规律，构建重大疾病预测、预警模型；基于天河二号超级计算机，创建重大疾病大数据可视化分析平台，为重大疾病及其相关健康问题的防控策略和干预措施的制定提供科学依据。

　　2. 基于多模态大数据的复杂疾病临床诊断标准及应用模型研究。

　　针对典型疾病的生物、生理、行为、心理等多模态数据的复杂特征，研究非侵入式、普适化数据感知及获取技术（信噪比、采样率、传输率、待机时间等需满足日常环境及持续监测的需要）；基于天河二号超级计算机研究多模态数据的高效数据分析、融合及挖掘技术；揭示多模态数据特征与典型疾病之间的关联关系及演化规律，为解释复杂疾病的发病机理及临床诊断指标奠定基础，为实现疾病的早诊早治提供理论基础及技术支撑。

　　3. 面向典型疾病的高精度生物大分子三维结构研究。

　　针对大规模、大尺度生物大分子冷冻电镜数据，基于天河二号超级计算机，研究高精度生物大分子三维结构；研究高精度三维重构模型、算法及可视特征分析的关键技术；解决当前大规模、大尺度生物大分子三维结构并行计算的瓶颈问题，实现基于天河二号超级计算机的生物大分子三维重构软件和可视化分析工具；针对典型疾病（如老年痴呆症），获得生物大分子的原子分辨率三维结构，为药物研制及生物医学的科学发现提供支持。

　　4. 重大疾病多组学与医学大数据挖掘基础理论及关键技术。

　　面向三级甲等类大医院关于重大疾病的多组学和生物医学图像大数据，开展关键致病因子的深度挖掘与分析；研究重大疾病在基因和蛋白质水平上的网络精准构建方法；研究重大疾病不同阶段的生物大分子、影像和表征的症状之间的关联性和差异特征，并开展疾病产生机理、发展演变过程及药物靶点的研究；基于天河二号超级计算机，实现关于重大疾病的多组学与医学大数据的挖掘与分析工具，为针对不同人群重大疾病的个性化医疗和精准用药提供理论和技术。

　　5. 基于天河二号超级计算机的智慧医学与健康大数据平台研究与构建。

　　基于天河二号超级计算机系统、针对典型复杂疾病（如老年痴呆症、抑郁症、恶性肿瘤、重大传染疾病等），研究生物与医学大数据的获取、处理、共享及开放策略的信息系统构建方法；研究典型复杂疾病的生物和医学大数据处理、建模及数据挖掘系统，在此基础上开发原型系统及工具；研究面向重大疾病及健康问题的干预、决策机制及预警、防控策略；构建基于天河二号超级计算机的智慧医学与健康大数据平台。

　　针对本资助领域的5个研究方向，拟分别资助5项重点支持项目。直接费用资助强度400-600万元/项，资助期限为4年，申请书中的研究期限应当填写“2017年1月1日-2020年12月31日”，**合作研究单位的数量不得超过2个。**

**四、申报要求及注意事项**

　　（一）申报要求。

　　1.申请人条件

　　（1）中心项目申请人应当是本领域高水平的学术带头人，具有较强的组织管理和协调能力，有足够的时间和精力从事本项目的工作，并在项目的实施过程中起主导作用；主要研究方向的负责人应为本领域有影响的学术带头人，学术思想活跃，研究成果显著；中心项目申请人应当是其中1个主要研究方向的负责人。

　　（2）重点支持项目申请人应当具有承担基础研究课题或者其他从事基础研究的经历；具有高级专业技术职务（职称）。在站博士后研究人员以及正在攻读研究生学位的人员不得申请。

　　2.依托单位（含合作研究单位）应当拥有可共享的科研大数据集,具有良好的大数据研究基础。

　　3.申请项目应当依托“天河二号”超级计算机开展研究，并将中山大学作为合作研究单位之一（国家超级计算广州中心由中山大学管理）；鼓励申请人与广东省内具有一定研究实力的高等院校及其附属医院或研究机构开展合作研究。项目申请书中应当明确与中山大学等合作各方的合作内容、主要分工和经费预算等。其中，中山大学作为合作研究单位的经费，以及用于“天河二号”超级计算机开展计算的经费预算，合计不低于大数据项目经费的50%。项目实施过程中产生的科学数据应当汇交至“天河二号”超级计算机上实施计算与科学研究；相关科学数据和研究成果应当开放共享。

　　（二）限项规定。

　　大数据项目属于国家自然科学基金联合基金项目，执行国家自然科学基金的限项申请规定，具体内容请参见《2016年度国家自然科学基金项目指南》。

　　（三）申请注意事项。

　　1.申请书报送日期为2016年4月11日-15日16时。

　　2.申请书采用在线方式撰写，对申请人具体要求如下：

　　（1）申请人在填报申请书前，应当认真阅读本项目指南和《2016年度国家自然科学基金项目指南》中申请须知的相关内容，不符合项目指南和相关要求的申请项目不予受理。

　　（2）申请人登录科学基金网络信息系统（以下简称ISIS系统，没有系统账号的申请人请向依托单位基金管理联系人申请开户），按照撰写提纲要求撰写申请书。

　　（3）申请书中的资助类别选择“联合基金项目”，亚类说明选择“中心项目（X）”或“重点支持项目”，附注说明选择“NSFC-广东大数据科学中心项目”，“申请代码1”选择F02，“申请代码2”根据项目研究领域自主选择相应的申请代码。**以上选择不准确或者未选择的项目申请将不予受理。**

（4）申请书的报告正文应当按照“中心项目”（申请书撰写提纲见附件2）或“重点支持项目”申请书撰写提纲的要求撰写。如果申请人已经承担与大数据项目相关的国家其他科技计划项目，应当在报告正文的“研究基础”部分论述申请项目与其他相关项目的区别与联系。

　　（5）申请人完成申请书撰写后，在线提交电子申请书，下载并打印最终PDF版本申请书，向依托单位提交签字后的纸质申请书原件。

　　（6）申请人应当保证纸质申请书与电子版内容一致。

　　（7）大数据资助项目在执行期间形成的有关论文、专著、研究报告、软件、专利及鉴定、获奖、成果报道等，应注明“国家自然科学基金委员会-广东省人民政府大数据科学研究中心项目（项目批准号）”。

　　3.依托单位应当对本单位申请人所提交申请材料的真实性和完整性进行审核，并在规定时间内将申请材料报送国家自然科学基金委员会。具体要求如下：

　　（1）应在规定的项目申请截止日期（2016年4月15日16时）前提交本单位电子申请书及附件材料，并统一报送经单位签字盖章后的纸质申请书原件（一式一份）及要求报送的纸质附件材料。

　　（2）提交电子申请书时，应通过ISIS系统逐项确认。

　　（3）报送纸质申请材料时，还应包括本单位公函和申请项目清单,材料不完整不予接收。

　　（4）可将纸质申请书直接送达或者邮寄至国家自然科学基金委员会项目材料接收工作组。采用邮寄方式的，请在项目申请截止日期前（以发信邮戳日期为准）以快递方式邮寄，并在信封左下角注明“大数据项目申请材料”。请勿使用邮政包裹，以免延误申请。

　　4.联系方式。

　　（1）材料接收组联系方式。

　　通讯地址：北京市海淀区双清路83号国家自然科学基金委员会项目材料接收工作组（行政楼101房间）

　　邮编：100085

　　联系电话：010-62328591

　　（2）联合资助双方联系方式。

|  |  |
| --- | --- |
| 国家自然科学基金委员会  　　信息科学部综合处  　　地址：北京市海淀区双清路83号  　　邮编：100085  　　联系人：何杰  吴国政  　　电话：010-62327146  　　　　　010-62327929  电子邮件：hejie@nsfc.gov.cn | 广东省科学技术厅  基础研究与科研条件处  地址：广州市越秀区连新路171号  邮编：510033  联系人：钟自然  电话：020-83163335  　　　020-83163835  电子邮件：duanyizhu@gdte.cn |