

· 管理纵横 ·

国家自然科学基金区域创新发展联合基金 管理数据可视化设计研究

王涛^{1,2*} 金钊^{1,2} 黄志鹏^{1,2} 李彦萍^{1,2} 李小娟^{1,2}

1. 首都师范大学 地球空间信息科学与技术国际示范学院/资源环境与旅游学院,北京 100048
2. 首都师范大学 三维信息获取与应用教育部重点实验室,北京 100048

[摘要] 数据可视化将各类数据转换为图形图像的形式进行视觉展示,可视化分析能够提高数据表达和信息传递的效率,辅助用户开展定量化、科学化的管理和计划工作。国家自然科学基金区域创新发展联合基金作为新时期联合基金重要组成部分,联合资助方规模快速增长,在加大基础研究多元投入,吸引全国优势科研力量解决地方和行业发展中的关键科学问题等方面发挥着重要作用。随着区域创新发展联合基金规模快速增长,管理与分析工作需要有效的手段对数据进行统计分析,以图形的形式展示数据中多层次、多维度、多地域的复杂关系。本文以区域创新发展联合基金管理业务作为研究对象,分析了业务数据可视化管理模式需求,引入并设计了面向典型业务流程的可视化图形和联合基金可视化仪表盘,采用部分业务模拟数据进行了可视化试验,初步成果满足了区域创新发展联合基金业务管理的基础可视化需求。

[关键词] 区域创新发展联合基金;数据可视化;项目管理;仪表盘

数据可视化(Data Visualization)是利用计算机图形学和图像处理技术,将各类数值型、非数值型数据以点、线、面、文本等元素组成的图形图像进行视觉展示的过程,并利用颜色、位置、动画等变量,其目的是以清晰且高效的方式将信息传递给用户^[1]。通过有针对性的可视化设计,有机的组合统计图表、地图、文字等形式,可以既保证数据表达的科学性、艺术性,又提升信息传递的载负量,实现“一图胜千言”的效果。随着政府管理、行业应用的数字化和信息化的发展,各类业务管理产生了海量、实时、多维的有价值数据,这些数据可以通过有效的分析展示来进一步支撑管理工作。常规的表格和统计图形已经难以有效的展示数据中蕴含的信息,业务信息化管理模式的提升面临一定挑战,需要从应用需求特点出发,有针对性的设计实现数据可视化技术,以辅助海量数据应用的适应性展示。科研项目管理是一类典型业务管理工作,包括资助计划、项目设置、项目申请与评审、项目立项与管理评估等多个环节,由于



王涛 博士,首都师范大学教授、地信国际学院副院长,主要从事地图学与地理信息系统的教学与应用研究工作,任国际地图学协会教育与培训委员会主席等。先后承担国家自然科学基金项目、国家重点研发计划子课题等项目,曾获得中国地理信息科技进步奖一等奖。

涉及的学科领域、依托单位、科研人员、管理人员的数量众多、分布广泛,每个阶段都会产生复杂多维的数据,可以服务于科研项目的管理、规划、评估等工作。

联合基金项目是国家自然科学基金资助体系的重要组成部分,经过二十年的发展,该类型项目在引导地方政府、企业、行业部门等主体加大基础研究多元投入,吸引全国优势科研力量解决地方、企业和行业发展中的关键科学问题,推动我国相关领域、行业、区域自主创新能力的提升等方面发挥了不可替代的重要作用。为进一步贯彻习近平总书记关于科技创新的重要论述精神,推动科学基金更好适应基

收稿日期:2021-08-12;修回日期:2021-09-09

* 通信作者,Email:wangt@cnu.edu.cn

本项目受到国家自然科学基金项目(J2024004,41671403)的资助。

基础研究发展需要,2018年,国家自然科学基金委员会(以下简称“自然科学基金委”)启动新时期自然科学基金改革,提出强化协同创新、拓展基础研究多元投入渠道的改革目标,通过加强顶层设计和统筹管理,坚持问题和需求导向,优化整合现有联合资助布局,推进形成了包括“区域创新发展联合基金”、“企业创新发展联合基金”等内容的新时期联合基金资助体系。其中,区域创新发展联合基金在覆盖的省区、研究领域以及项目规模等方面快速增长,2020年,已有16个省、市、自治区的联合资助方发布了资助领域。

区域创新发展联合基金具有投入多元化、资助研究领域广泛、自然科学基金委与联合资助方共同管理、评审过程多阶段多维度等特点,其管理过程包括协议管理、资助领域和方向的确定、项目申请受理与初审、通讯评审与会议评审、联席工作会审议、管理委员会审议、联合基金阶段性总结等环节,每个阶段涉及的主体不同、任务不同,需要进行大量的数据统计和分析,为基金管理各个环节中领导进行决策提供基础。

1 研究基础与现状

可视化技术与图形化展示是掌握各类管理业务信息整体状况的重要工具,以可视化分析为核心的“业务仪表盘”、“领导驾驶舱”已经成为智慧城市、行业资源管理等数据密集型应用的关键组成。图形化展示在科研项目管理与分析中得到了大量应用。李艳波、刘权等(2009)提出了利用地理信息系统(Geographic Information System, GIS)辅助基金项目管理,并将空间统计及可视化应用于科学基金数据的最优布局决策,实现更有效、更便捷的科学基金展示^[2]。刘权等(2009)引入双重粒度的概念开展基于面向对象数据逻辑模型设计,在满足约束性和完整性的条件下,提出建立国家自然科学基金委员会管理数据仓库,并在面向科学基金管理数据的展示系统中验证数据仓库概念模型、逻辑模型和物理模型的有效性^[3]。文宗川等以2008—2017年美国国家科学基金在创新领域发布的研究计划数据为例,通过时间区间图的方式重绘奖励创新项目的时间演进过程,从而得到各个阶段内研究计划的研究脉络和集群重点^[4]。2019年,欧盟研究与创新计划框架“地平线2020”发布了在线交互式可视化信息管理平台^[5],通过地图、直方图、散点图、矩形树图等可视化方法从多个角度对研究指南、项目承担与合作单

位、项目产出等数据进行可视化,利用主题过滤器选取、区域过滤等交互式手段,灵活展示特定区域、时间、科研单位等条件下的项目详细信息,该平台集中展示了项目管理层面多维度的信息,有效地促进了科研数据的使用和共享。

目前,区域创新发展联合基金信息的统计分析主要依赖于办公软件,现有管理信息基本上由统计数据表格和简单统计图形组成,在一定程度上满足了展示基础统计信息的需求。区域创新发展联合基金在研究领域指南拟定、项目申请与评审等方面具有其独特性,对业务数据整理统计、信息可视化展示提出了更高的要求。常规办公软件无法引入更为有效的图表和信息化工具,难以灵活、高效的表达联合基金多层面、多角度的管理信息,现有业务管理工作流程信息化工作面临一系列挑战。为了更为有效的开展区域创新发展联合基金管理,提高管理决策效率,满足规模日益扩大的联合基金项目管理工作,亟需梳理联合基金项目管理信息化与可视化需求,引入可视化分析手段,探讨如何更为高效地展示、统计管理业务数据的途径,并有针对性的设计实现可视化软件工具,展示联合基金项目多层次、多维度、多地域的复杂关系。

本研究针对区域创新发展联合基金管理过程业务数据高效可视化的任务,分析了新时期联合基金管理业务数据的生命周期,结合其管理特点与需求,在对不同维度信息结构化处理的基础上,探讨了业务管理阶段不同类型数据可视化表达方式,采用部分业务模拟数据设计联合基金信息化仪表盘,以满足新时期联合基金展示需求导向、多元投入、开放合作的工作态势。

2 区域创新发展联合基金管理过程数据可视化设计

管理业务数据可视化以及基于可视化的分析工具需要融入联合基金管理的业务过程,为了实现该目标,首先需要对区域创新发展联合基金管理不同管理阶段任务进行解构,分析业务管理数据的供给角色和应用角色的功能。其次,区域创新发展联合基金项目业务管理过程中,自然科学基金委、联合资助方、项目承担单位、申请人等活动主体在协议管理、资助领域和方向拟定、项目申请受理与初审、通讯评审与会议评审、联席工作会等阶段的需求是不同的。再次,根据数据类型确定图表选取和可视化设计,该过程是一个多次反复、相互嵌套的过程^[6],

联合基金管理信息可视化设计可以划分为以下三个步骤：

(1) 数据产生与供给关系分析。按照业务逻辑划分管理阶段,从联合基金的协议方、资助领域与研究方向、申请项目、资助项目、承担单位与合作单位等对象的数据描述进行角色分析和结构化处理。

(2) 面向业务逻辑的数据流重构。联合基金管理全过程数据流与管理活动参与主体形成业务信息化逻辑矩阵,支撑联合基金信息化管理模式的数据库逻辑模型和接口设计。

(3) 可视化图形与数据仪表盘设计。分析设计各种信息相应的最优展示表达方式,围绕典型业务活动,设计数据集成整合、分类查询、统计和可视化的平台。

2.1 业务数据产生与需求供给关系分析

联合基金管理过程包括协议管理、资助领域和方向的确定、申请项目受理与初审、通讯评审与会议评审、联席工作会审议、管理委员会审议、联合基金阶段性总结等环节,每个环节管理任务的开展都需要前期的数据支撑,在指南拟定、预算分配、项目申请及多个阶段的评审等管理活动中形成了管理过程数据流,通过梳理分析联合基金管理各个阶段的数据流,既可以辅助年度内联合基金管理的逐阶段决策,也对未来指南规划、经费预算、成果评估具有较为重要的意义。可以将联合基金管理过程分为项目指南、项目申请及初审、项目评审与资助和工作总结等阶段,并对应的建立每个阶段活动主体和典型业务,如表 1 所示。

数据的生命周期通常可以总结为数据产生、数据采集、数据处理、数据发布、数据利用等五个核心阶段^[7]。对于区域创新联合基金管理流程而言,数据产生始于协议签订、年度项目指南拟定(图 1),项目指南阶段主要的管理任务是确定资助领域和各领域方向,在此管理活动中产生联合资助方、领域、方向等类型的信息,继而服务于项目申请阶段,项目申

表 1 年度联合基金管理过程分析

管理阶段	管理活动主体	管理任务
指南发布	联合资助方 相关科学部 管理办公室	协议管理 资助领域和方向确定
项目申请	项目申请单位 相关科学部 管委会	项目申请 项目初审 项目复审
评审结果	联合资助方 相关科学部 管理办公室	通讯评审 联席工作会审议 会议评审 管委会审批
工作总结	管理办公室	工作总结与汇报

请人依据指南规定申请项目,产生领域、学部、申请人、申请单位、申请项目等信息,管理活动主体对申请项目进行多轮评审,产生领域、方向、获批项目、承担单位、合作单位等信息。完成以上各个阶段业务流程后,进入工作总结阶段,汇总后数据包括联合资助方项目地域、领域、方向分布以及历年申请和资助项目等信息,数据具体生命周期如图 1 所示。

2.2 面向业务逻辑的数据可视化图形设计

区域创新发展联合基金管理者包括了来自自然科学基金委、联合资助方以及联合基金管理委员会等机构的业务管理人员、评审专家和决策领导,在上述业务过程中,管理者需要根据每个阶段生成的数据从地域、领域、方向、科学部、申请单位等角度进行数据统计并可视化,支撑后续决策过程。业务决策数据需求和供给的图表类型是可视化的关键,需要参照用户习惯和具体场景进行有针对性的设计。在信息整合的基础上,针对不同数据生命周期的特点,可以从指南发布、项目申请及初审、项目评审与资助和阶段性工作总结四类管理信息(如表 1)展开分析,以下从四个方面典型应用进行可视化设计与表达。

协议执行管理是区域创新发展联合基金项目的基础。截至 2020 年,共有 16 个区域创新发展联合

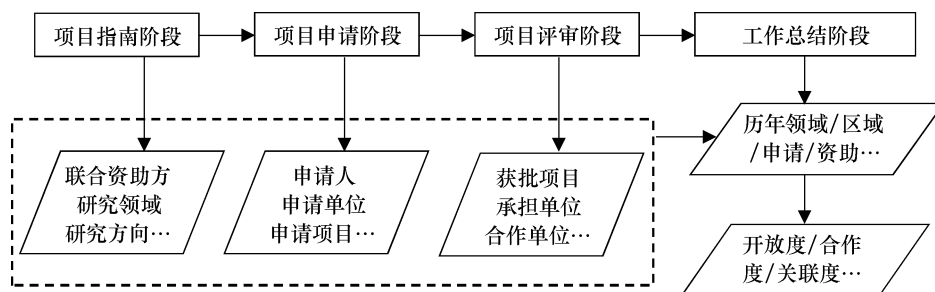


图 1 联合基金管理过程数据生命周期分析

基金联合资助方,联合资助协议的起始执行年份、匹配经费、研究领域等信息各不相同。表达该信息可以使用表格或者直方图,这种方式在基于给定条件排序后,有助于快速获取部分排序在前、在后的信息,然而由于条目过多,不利于用户获取整体信息^[8]。地图作为地理信息传输的重要载体,能够同时表达多个维度的数据,用户查看信息时能够较好的结合已有地理知识,提升获取信息的效率,可以将协议数据定位到地图上表达。可发现联合资助方呈现快速递增态势,且有一定的局部空间集聚效应。

区域创新发展联合基金在拟定资助领域和方向时,自然科学基金委相关处室需要在总结、分析领域和方向总体分布的基础上,与各联合资助方反复论证,以确定各个区域在各领域的资助方向。2020年度,共发布了生物与农业、环境与生态、能源与化工、新材料与先进制造、现代交通与航空航天、电子信息和人口与健康等七个领域的544个研究方向。在该过程的不同阶段,管理活动主体需要快速获取联合资助方在各领域发布的方向、项目申请、项目评审等信息,并分析不同领域在发布方向总数和申请项目数量中的占比情况。可视化展示时,既需要表达资助领域、研究方向和申请项目存在的层次关系,也需

要表达不同领域、方向的比例关系,将常规的环形图重新设计为多个层次的嵌套环形图,如图2所示,外环按照获批项目数、内环按照资助方向数设定环形大小的比例,利用相同色系的颜色建立外环、内环的关联,可以较好的满足实际需求。从图中可以看出,新材料与先进制造、人口与健康等领域在发布方向数量(内环)、申请项目数量(外环)方面占比较大。

区域创新发展联合基金管理工作在申请、评审的各个阶段,需要从项目申请单位所在区域、方向提出资助方以及研究方向所属学部等多个角度进行统计分析,常规的表格和统计图形难以较为高效的展示联合基金多元投入、学科之间交叉关系。为此,可以按照省区、领域、学部对项目申请情况进行统计分析,引入桑基图来表达对象之间关系和数量,对象间的连线越粗,则二者对应数量越大,图3分别展示了2020年区域创新发展联合基金的申请阶段各联合资助方与所提出研究方向,以及资助领域申报项目与评审学部之间的数量关系,从该图中可以看出,各联合资助方在不同的研究领域内均有广泛投入,领域之间跨学科势态明显,人口与健康领域、新材料与先进制造领域的研究方向集中于医学科学学部、工程与材料科学学部。

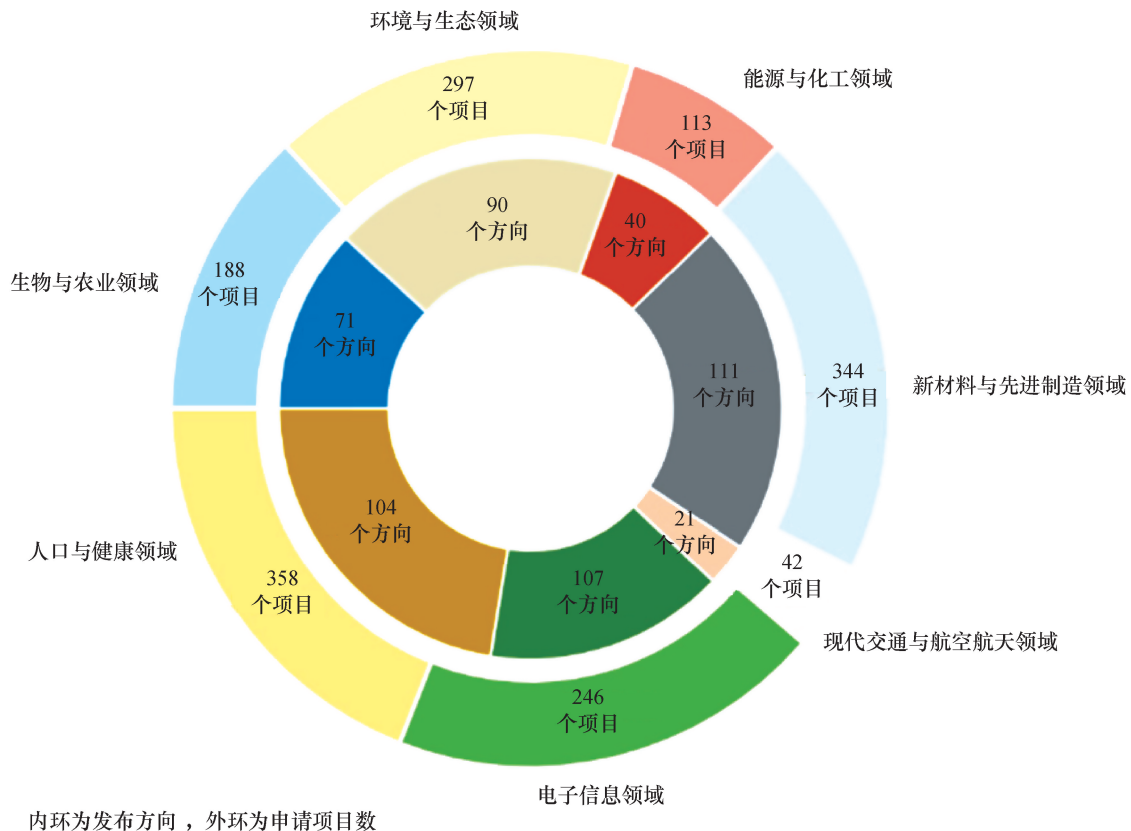


图2 发布方向(内环)及申请项目数量(外环)分布图(2020)

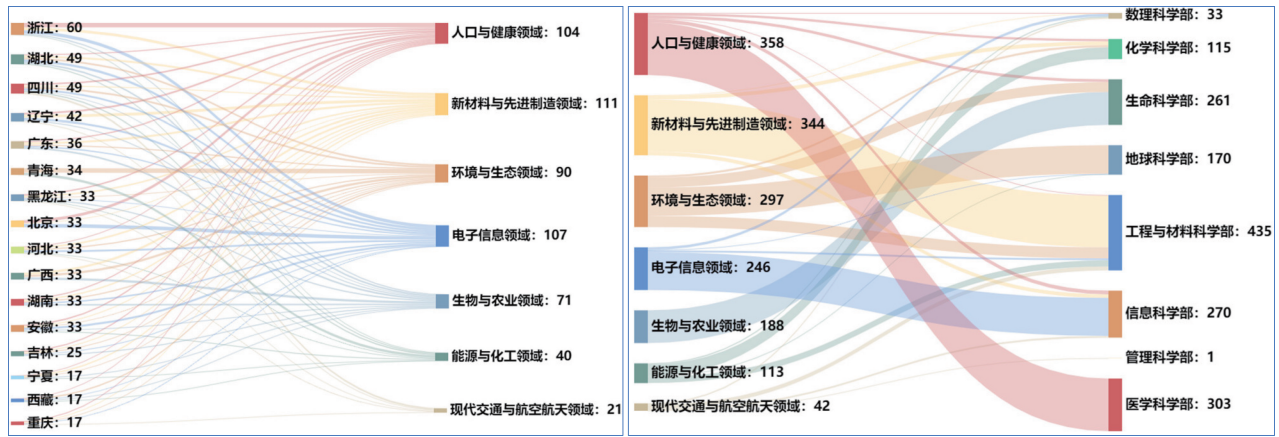


图 3 2020 年联合资助方的领域方向、资助领域申报项目所在学部分布图

联合基金发展历程的统计可以从资助项目数量和资助金额的角度展示。常规的左右或者上下双轴坐标图能够用来表示二元变量，为了更为灵活的展示该信息，可以在风玫瑰图基础上改进形成嵌套的双层玫瑰图(图 4)，分别表示联合基金资助项目数量和资助金额的变化。通过相邻柱状图形的对比，可以看出新时期联合基金改革之后，资助经费在较大体量的基础上增长规模明显，显示了区域创新发展联合基金得到了越来越多的联合资助方的重视。

为了分析联合基金在不同研究领域历年来的资助情况，常规而言需要多张统计表格以展示 20 年间分布在多个学部的数据，管理者难以短时间内获取总体情况，而柱状图等难以在空间有限的情况下有效的展示统计信息。矩形树图采用矩形表示层次结构里的节点，层次关系用矩形之间的嵌套来表示，矩

形大小对应统计数据体量，而矩形颜色可用来增强不同类别之间的区分度。图 5 是在矩形树图的基础上改进的历年不同学部代码资助项目经费的图形展示，通过观察矩形的大小不仅可以横向对比不同学部代码历年获批金额的变化，还可纵向对比历年金额变化信息，在数据可视化系统内，用户可以使用鼠标移动、点击、放大查看因图幅有限或者对应年份数量较小而省略的信息。

3 区域创新发展联合基金管理数据可视化仪表盘设计

静态的图形设计制作可以一定程度的满足管理者的信息获取需求，而管理者更多的是需要能够设定时间、研究领域、联合资助方、项目状态等条件，提取、统计部分相关数据，通过多种统计图形从不同角度展示数据蕴含的信息。因此，需要针对管理主体业务需求，将各管理阶段产生的数据以适合的图形化表达方式展现，并有机的集成为多模块组成的数据可视化仪表盘，提供动态的、可交互的可视化。可视化仪表盘的概念已经从单一视图的报告屏幕发展到具有多个视图和用途的交互界面，针对其设计与实现，Edward Tufte 提出了“数据墨水比例(Data-Ink Ratio)”概念^[9]，指出展示数据的屏幕容量是有限的，应当根据信息的重要程度合理安排屏幕位置，避免无效且无意义的的数据表达。因此，联合基金管理数据可视化仪表盘的设计实现需要遵循信息一览性、图形表达适宜性、业务逻辑对应等特点，以有效地展示数据特征、支撑典型业务开展。

考虑系统稳定性、可拓展性和易用性，采用了 B/S(浏览器/服务器)工作模式，服务端基于 Python 的轻量级后台框架进行开发，能够快速部署至当前常见配置的个人计算机平台，用于联合基金项目管

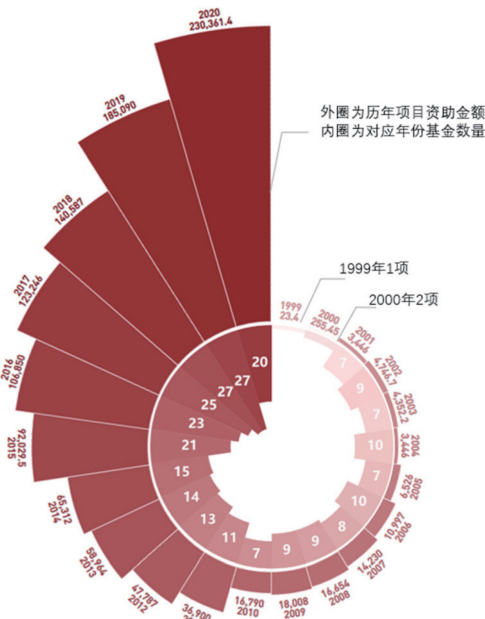


图 4 展示联合基金数量和资助金额变化的嵌套双层玫瑰图

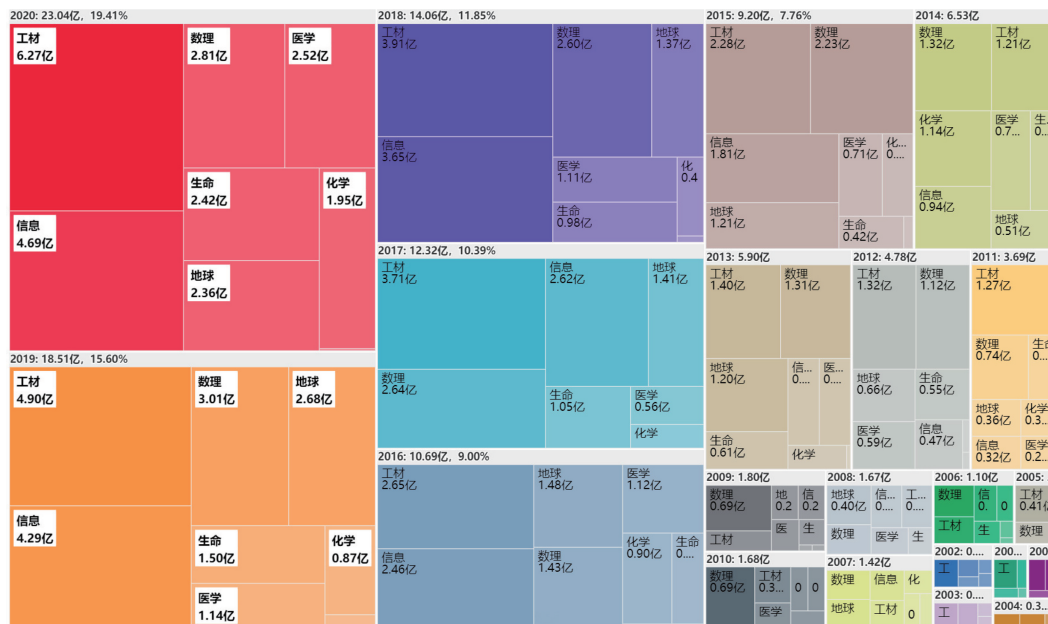


图5 按照资助规模展示历年各科学部联合基金情况

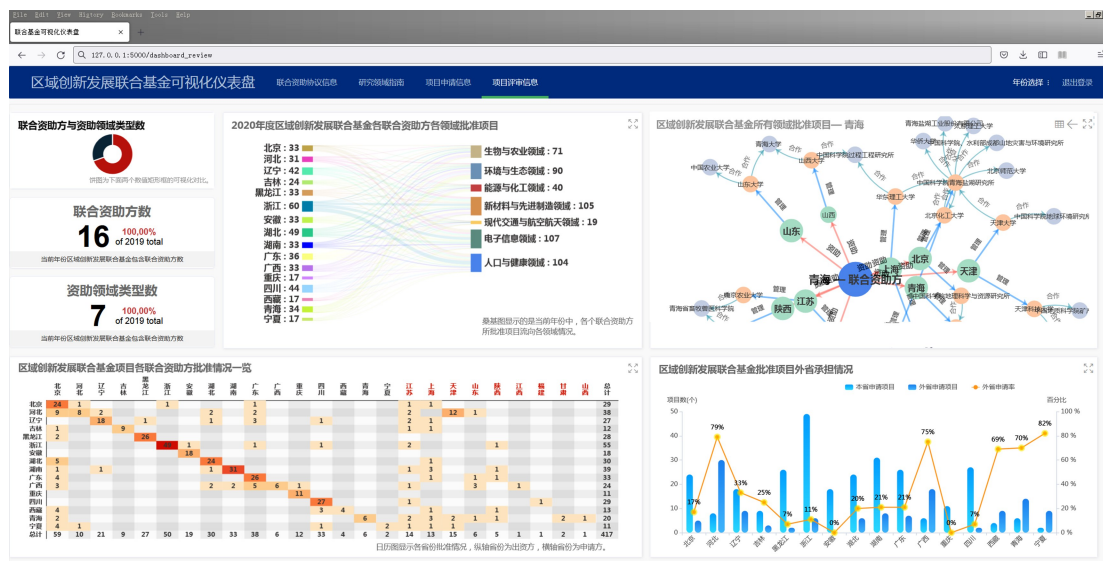


图6 区域创新发展联合基金管理数据可视化仪表盘主视图

理各阶段的统计表格数据规范化处理和集成。浏览器端实现了流式布局数据可视化仪表盘，满足桌面版和移动段使用场景。图6为区域创新发展联合基金可视化仪表盘主界面示例，依据数据的生命周期，选择项目指南、项目申请与资助、历年管理等条件进行仪表盘配置，界面上方为可以配置的系列重要的总体指标，包括所查询年份对应的联合资助方、项目申请与资助、项目资助经费等数量信息；左侧地图部分可以展示各类与区域相关的信息，可以展示不同日期条件下联合资助方、申请单位、承担单位等方面信息空间分布；右侧放置四种适用于不同类型数据

的图形化展示方式，饼图展现不同研究领域发布方向数占总体的比例，桑基图展现内容为申请情况中学部和领域之间的交叉流动势态，堆积图表达不同领域的资助项目数，下方堆叠图展示历年各科学部代码项目的资助数量。根据信息的不同，可以定制系统预设的可视化图形类型。同时，以联合资助方为辐射点，地图视图可以与其他带有省份信息数据的图表协同关联，通过用户以任务为驱动的地图视图交互，其他视图图表数据会随之改变，并通过设定时间、区域等条件，直接聚焦感兴趣的内容，增加数据可视化维度。

4 结论及展望

数据可视化技术有助于提升科研项目业务管理信息化效率。本研究围绕区域创新发展联合基金管理过程数据可用性和决策效率的提升,开展了数据可视化图形设计,分析并设计实现了各工作阶段典型业务数据的图形化表达方式,集成设计了联合基金管理数据可视化仪表盘,对不同业务阶段中管理活动主体产生和需求的数据进行可视化,通过提升数据易用性和信息传递效率,更好的把握区域创新发展联合基金局部与全局态势,从而支撑定量化决策,部分工作已在 2020 年联合基金业务管理工作中得到了初步应用。以此为基础,下一步工作将围绕具体阶段管理可视化需求和集成开展细化设计,并针对可视化仪表盘易用性进行用户调查和分析,根据不同管理主体角色及使用习惯进一步优化区域创新发展联合基金可视化仪表盘。

参 考 文 献

- [1] 陈为,沈则潜,陶煜波. 数据可视化(2版). 北京: 电子工业出版社, 2019.
- [2] 李艳波,刘权,鲁明羽. 基于 GIS 技术的科学基金管理数据可视化方法. 计算机工程与应用, 2009, 45(35): 73—75.
- [3] 刘权,肖智博,鲁明羽. 面向科学基金管理数据数据仓库概念模型设计. 计算机工程与应用, 2009, 45(36): 131—133.
- [4] 文宗川,王立军,薛彪. 美国 NSF 资助创新项目的知识图谱分析及对我国的启示——基于 2008—2017 NSF 相关数据. 科技管理研究, 2019, 39(12): 257—266.
- [5] EU Commission. Horizon Dashboard. (2021-03-06)/[2021-09-09]. <https://webgate.ec.europa.eu/dashboard>.
- [6] Munzner T. A nested model for visualization design and validation. IEEE Transactions on Visualization and Computer Graphics, 2009, 15(6): 921—928.
- [7] 丁宁,马浩琴. 国外高校科学数据生命周期管理模型比较研究及借鉴. 图书情报工作. 2013, 6(6): 18—22.
- [8] Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 1956, 63(2): 81—97.
- [9] Tufte E. The Visual Display of Quantitative Data (2nd edition), Graphics Press, 2001.

Data Visualization Design for NSFC Regional Innovation Development Fund Management

Wang Tao^{1,2*} Jin Zhao^{1,2} Huang Zhipeng^{1,2} Li Yanping^{1,2} Li Xiaojuan^{1,2}

1. College of Geospatial Information Science and Technology/College of Resource, Environment and Tourism, Capital Normal University, Beijing 100048

2. MOE Lab of 3D Data Acquisition and Application, Capital Normal University, Beijing 100048

Abstract Data visualization converts data in various forms into graphics to improve data representation and visual transmission of information. Visualization analysis, or visual analytics, helps users understand data by uncovering information and knowledge in multiple perspectives, which further facilitates management and planning in a quantitative and efficient way. NSFC Regional Innovation Development Fund, which is an essential part of the NSFC grant system, has been growing drastically in recent years. It attracts research investments from a full spectrum of resources and helps researchers national wide to investigate unique scientific questions in specific regions and fields. During management of projects and grants supported by the Fund, decision makers need effective means to conduct data analysis and statistics. Structural information on multiple granularities, dimensions and regions can be extracted to support decision making. Based on demand analysis of information management of the Fund, this paper reports new designs and implementations of data visualization for typical managing transactions. Real world data together with simulated data were used in the system development. A data dashboard is implemented to meet visualization requirements in general context.

Keywords regional innovation development fund; data visualization; grant management; data dashboards

(责任编辑 吴征天)

* Corresponding Author, Email: wangt@cnu.edu.cn