

基于基金项目的学科交叉热点主题识别与分类研究^{*}

——以美国农业与食品领域为例

齐世杰^{1,2} 赵静娟¹ 郑建华² 串丽敏¹

(1. 北京市农林科学院数据科学与农业经济研究所, 北京 100097; 2. 中国农业科学院农业信息研究所国家新闻出版署农业融合出版知识挖掘与知识服务重点实验室, 北京 100081)

摘要: 学科交叉是原始创新成果之源, 科学基金是引导原始创新的重要途径。准确识别由基金资助的学科交叉热点方向, 把握交叉主题的研究趋势, 为我国重大科技创新、前瞻部署提供参考。基于基金项目数据, 利用Salton指数和共现分析, 识别基金项目中的学科交叉热点主题, 通过切分时间窗口构建“支持趋势—交叉趋势”战略坐标图, 将学科/主题定位为前沿型、潜力型、衰退型和热点型, 以美国农业与食品领域为例验证方法的有效性。结果显示: 美国农业基金项目的学科交叉趋势日益显著, 学科和主题均以前沿型和潜力型居多, 涉及基因组学和遗传学、土壤改良与治理、综合虫害管理、动植物育种等方面, 以食品加工技术、食物保护、动物营养利用等为代表的研究主题的交叉融合表现活跃。未来, 应提高交叉突破创新意识, 借助战略坐标图定位, 分类布局面向国家需求的重点领域关键技术创新方向。

关键词: 农业与食品; 美国; 学科交叉; 主题识别; 趋势战略坐标

中图分类号: G353.1 **DOI:** 10.3772/j.issn.1673-2286.2023.10.004

引文格式: 齐世杰, 赵静娟, 郑建华, 等. 基于基金项目的学科交叉热点主题识别与分类研究: 以美国农业与食品领域为例[J]. 数字图书馆论坛, 2023(10): 29-39.

近年来, 国内外科学界一直倡导学科交叉研究 (Interdisciplinary Research, IDR), 促进新兴交叉学科发展。学科交叉研究, 也称跨学科研究, 是一种新的科学研究范式, 是交叉关联与渗透融合多学科领域的研究^[1]。学科交叉研究是知识创新的重要方法, 不仅能够催生重大科学成果产生, 也是解决当今社会发展中复杂问题的重要途径^[2-3]。世界各国已将资助学科交叉研究提升到国家科技战略层面, 如美国国家科学基金会的“发展融合研究计划”、欧盟的“未来新兴技术计划”, 以及英国设立的面向农业/粮食安全等全球挑战的重大基金

项目, 均将跨学科研究视为重要支持对象^[4]。据统计, 2006—2013年美国国家科学基金会共投入了6亿美元来资助合成生物学学科交叉研究。2020年, 我国国家自然科学基金委设立交叉科学部, 以促进对复杂科学技术问题的协同攻关, 寻求科技突破口。2021年习总书记提出推动“新农科”建设, 这是面向科技创新、产业升级需求的学科融合发展理念。在新时代发展要求下, 交叉融合研究势必成为科技创新的重要驱动力量。

学科交叉主题识别是学科交叉研究的重要内容, 学者大多基于论文或专利数据识别学科交叉主题, 而

收稿日期: 2023-07-12

^{*}本研究得到北京市农林科学院创新专项“智库型农业情报研究与服务能力提升”(编号: KJCX20230208)、“基于多源数据融合的农业热点前沿主题识别与实证研究”(编号: KJCX20200403)、“面向科研管理的情报研究与服务能力提升”(编号: KJCX20230210)和国家新闻出版署农业融合出版知识挖掘与知识服务项目“基于学科交叉的新兴主题识别方法与实证研究”(编号: 2022KMKS02)资助。

针对基金项目的研究较少。政府基金是基础科学研究的重要支撑,世界领先创新型国家重要科研基金资助项目具有前瞻性、先导性和战略性,相较于论文和专利,基金项目能够更早地反映特定领域亟待解决的科学问题和新技术^[5-7]。基于基金项目的科学大数据,准确识别研究领域的学科交叉热点主题,有助于掌握研究领域的交叉前沿与态势,对我国科技战略预先部署和提速发展具有客观参考意义^[8]。因此,基于基金项目的学科交叉主题识别无论是对于科技领域还是情报领域,均具有重要的探索价值。

农业是立国之本,强国之基。党的十八大以来,从深入实施藏粮于地、藏粮于技战略,到大力推进种源等农业关键核心技术攻关,农业科技创新带动农业高质量发展^[9]。党的二十大提出建设农业强国,纵观全球,农业强国必是农业科技创新强国。因此,提升农业科技创新能力,加强对前沿热点、关键环节和卡点瓶颈的科技攻关势在必行。本文以美国农业与食品领域的基金项目为例,对基金项目学科和重点研究主题进行计量分析,基于学科交叉测度和共现分析,识别学科交叉热点主题,通过构建“支持趋势-交叉趋势”战略坐标图,判别基金项目中学科和研究主题的分类和研究趋势,以期对相关领域的研究者提供参考与借鉴。

1 相关研究

通过学科交叉主题识别可发现学科交叉向多样化、深层次发展过程中形成的交叉与汇聚点,能够分析研究领域在哪些研究方向上产生了交叉现象。关于学科交叉主题识别已有大量研究,交叉主题识别过程大致包括3个部分:对文献或主题进行学科交叉特征测度,根据数据的共词特征、引文特征或网络结构特征提取表征交叉主题的词汇,通过设置阈值识别交叉主题^[10]。其中,提取主题词是主题识别的关键步骤,主题词的准确性和可靠性能够直接影响主题识别的效果。

根据提取方法和原理,学科交叉主题识别方法大致分为以下几种。①基于共现分析的方法,其原理是基于共同出现的词语进行主题识别,是反映文本学科主题的最直接、便捷的方法。例如,隗玲等^[11]基于弱共现关系理论和突发词监测,将论文中具有弱共现关系的高频词作为主题词,识别情报领域的高交叉性主题。Dong等^[12]以图情学论文为对象,综合利用共现网络、突发词监测等方法,将高频关键词聚类,对学科交叉

主题进行分析与识别。②基于社会网络的识别方法,通常以关键词为分析单元,结合聚类分析、多维尺度分析等展开研究。例如,李长玲等^[13-14]基于论文关键词,通过重叠社区发现方法构建知识聚类和重叠社群网络,识别情报学和计算机学科的交叉主题。③基于主题模型的识别方法,该方法综合考虑了文本语法、语义信息,能够在大规模语料中抽取主题词并获取主题词出现的概率,但算法需要人工设置阈值,可解释性欠佳,有学者对此进行改进。例如,陈琼等^[15]以医学信息领域论文的题目、摘要和关键词为语料,结合LDA(Latent Dirichlet Allocation)主题模型与Div指标对学科交叉主题进行识别和划分。阮光册等^[16]通过构建关键词共现网络,结合聚类分析和LDA主题模型,识别情报学和教育学两个学科的交叉主题。张振青等^[17]针对农业领域论文,利用改进的Phrase LDA主题模型从抽取的频繁短语中挖掘重要短语作为主题词,提高了语义的丰富性和主题的一致性,且模型不需要设置绝对阈值。④基于引文分析的识别方法,利用文献的引用关系或引用内容来识别交叉主题,往往具有一定时滞性,应用较少。例如,Chi等^[18]对文献共被引网络进行聚类,抽取学科交叉集群中高被引文献的主题词作为学科交叉主题。可以看出,目前的学科交叉主题识别方法多以论文或专利文本中的题目、摘要或关键词表征主题词,主题词来源较为单一。

综上,已有研究大多以论文或专利数据为研究对象,论文或专利数据与基金数据相比具有一定的时滞性,前瞻性较弱;从方法上看,主题词的提取方法较为丰富,学者通过算法从原始文档中提取能够表征交叉主题的词汇,提取方法对算法依赖性强,且需要人工设置阈值,标准和规则具有主观性,所得词汇准确性和合理性不易评判,解释性不强。此外,大多数研究止于对学科交叉主题的提取,缺乏后续进一步的研究,如交叉主题的发展趋势、领域中交叉主题的分类定位研究等。

本文以美国食品与农业研究院(National Institute of Food and Agriculture, NIFA)基金项目为对象,先利用共现分析和Salton指数定量测度学科交叉组合,再基于基金项目自带主题类别字段识别学科交叉领域的热点主题,进一步构建“支持趋势-交叉趋势”战略坐标图,将学科主题归为前沿型、潜力型、衰退型、热点型4类,从而定量刻画NIFA项目中学科主题的交叉趋势,多维度展示交叉主题在领域中的“生态位”,为基于基金项目的学科交叉主题识别和多维度分析提供新思路。该方

法可拓展至其他学科领域, 所得结果也可为我国农业项目规划、科研管理、科技创新提供一定的参考。

2 研究方法

2.1 学科共现分析

共现分析的基本思想是共同出现的特征项间存在着某种关联, 关联程度可用共现频次测度^[19]。通常情况下基金项目标注2个及以上的学科类别, 反映了这些学科之间存在一定的关联性。刘扬等^[20]针对国家自然科学基金申请代码, 采用语义相似度和共现分析, 计算出一级学科之间的交叉程度并可视化展示。

通过对基金项目进行学科共现分析, 能够量化不同学科间的关联强度, 分析学科之间的交叉融合状态, 通过可视化则可以直观、清晰展示数据之间的关系和状态。采取以下步骤: 首先提取基金项目的学科类别, 统计共现关系的强度, 即共现频次, 通过设定阈值(设为2), 生成共现矩阵, 其次根据共现矩阵绘制共现网络, 最后调用Python程序包进行学科共现可视化, 更直观地展示学科交叉融合关系, 便于读者理解与深入分析。

2.2 学科交叉度测算

测算学科交叉度是量化学科交叉程度的重要手段, 是识别学科交叉主题的基础。先通过学科交叉度测算筛选高交叉性学科组, 再根据项目信息的映射发现交叉热点主题。对于学科交叉度已有较多成熟的测算指标, 测算维度主要包括学科多样性、学科平衡性和学科差异性。国内外学者相继提出了跨领域引用指数(COC)^[21]、布里渊指数^[22]、Salton指数^[23]、信息熵指数^[24]、学科集成化指数^[25]、Rao-Stirling指标^[26]、Div指标和True Diversity指标^[27]等测算指标, 也有学者围绕学科多样性、学科平衡性、学科差异性等维度构建了学科交叉测度框架^[28], 不同指标各有优劣。根据基金数据特性, 研究领域各个资助项目的学科多样性和学科平衡性区分度较小, 而Salton指数在表征多样性和差异性方面比较灵敏^[29], 更适用于NIFA基金数据测度。

Hamers等^[30]在研究中验证了Salton指数在测度学科差异性方面的有效性, 因此借鉴Salton指数测算基金项目的学科交叉程度, 识别与量化高交叉性学科组,

并作为下一步识别学科交叉热点主题的数据基础。Salton指数的取值范围是0~1, 越接近1表示交叉度越高, 其表达式如式(1)所示。

$$S_{AB} = \frac{N_{AB}}{\sqrt{N_A} \times \sqrt{N_B}} \quad (1)$$

式中: S_{AB} 表示Salton指数; N_A 表示学科A的项目数; N_B 表示学科B的项目数; N_{AB} 表示学科A、B的共有项目数。

2.3 “支持趋势-交叉趋势”战略坐标图

1988年Law等^[31]提出战略坐标图, 通过坐标中不同象限和位置展示研究主题在学科领域中的地位及成熟度, 为科研人员发现学科领域的热点研究主题及研究前沿提供了有效途径。闵超等^[32]采用战略坐标分析法, 探究了图书情报与新闻传播两个学科的学科交叉热点主题划分和演化过程。袁毅等^[33]利用战略坐标对人工智能领域的学科群外部结构进行分析。范丽鹏等^[8]针对基金项目, 提出了立项强度和研发投入力度两个指标, 利用战略坐标图对技术主题前沿程度进行判定。

将学科交叉度引入战略坐标图, 提出支持度和交叉度指标, 利用指标在2016—2018年、2019—2021年两个时间窗口中的差值进行趋势分析, 构建“支持趋势-交叉趋势”战略坐标图, 展示研究领域中学科、研究主题的交叉趋势和发展态势, 从而实现对学科和研究主题的类型进行细粒度研判和准确定位, 为发现领域交叉热点主题提供有效途径, 有助于多维度刻画和比较领域内学科和研究主题的发展趋势。

(1) 支持趋势。首先将包含该学科/主题的项目与总项目的比例定义为支持度, 以支持度在两个时间窗口内的差值代表支持趋势: 值越大代表该学科/主题的关注度和研究热度越高, 该学科/主题往往蕴含着该时期亟待解决的重大问题。支持趋势公式如式(2)所示。

$$F_i = \left(\frac{N_i}{n} \right)_t - \left(\frac{N_i}{n} \right)_{t'} \quad (2)$$

式中: N_i 代表包含第*i*个学科/主题的项目数; n 代表总项目数; t 代表不同时刻。

(2) 交叉趋势。将项目中与某一学科/主题共同出现的学科/主题的数量定义为交叉度*P*, 以交叉度在两个时间窗口内的差值表征交叉趋势, 值大于0表示呈上升趋势, 值越大代表研究领域中与该学科/主题关联的

学科交叉研究越多。交叉趋势公式如式(3)所示。

$$Y_i = P_{i,t} - P_{i,t'} \quad (3)$$

式中: P_i 代表项目中与第*i*个学科/主题共同出现的学科/主题的数量。

(3) 战略坐标图。以支持趋势为横轴、交叉趋势为纵轴,建立战略坐标图,如图1所示。根据发展趋势刻画了4种类型的学科/主题:第一象限的节点具有“双高”特点,为前沿型学科/主题;第二象限中交叉趋势增强,支持趋势减弱,学科/主题未来有一定交叉融合的发展潜力,判归为潜力型学科/主题;第三象限的支持趋势和交叉趋势均呈减弱状态,判归为衰退型学科/主题;第四象限中支持趋势增强,交叉趋势减弱,为热点型学科/主题。

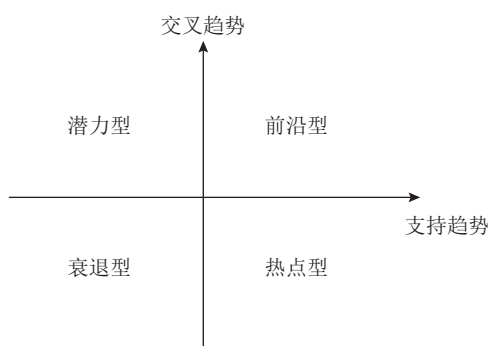


图1 “支持趋势-交叉趋势”战略坐标图

3 实验结果与分析

3.1 数据来源

NIFA是重要的科研机构,致力于农业、环境、人类健康、福利等方面研究,在美国农业领域具有一定的影响力。本文基金数据来自NIFA官网,包括2016—2022年NIFA资助的基金项目,每个项目的属性信息包括项目名称、项目编号、申报单位、申请人、非技术摘要、关键词、学科分类、知识分类、立项年份、执行时间、资助金额等20多个字段。检索时间为2022年12月1日,对数据进行去重、合并和清洗。

3.2 学科主题计量分析

3.2.1 重点学科布局

统计整理NIFA项目数据发现,2016—2022年

NIFA资助项目数量为10 310项。NIFA资助项目共涉及48个学科,可分为生物科学类、物理科学类、社会与行为科学类;以生物科学类居多,社会与行为科学类次之,物理科学类最少。2016—2022年NIFA项目Top 15学科资助情况(见表1)显示,资助频次位于Top 3的学科是教育学、工程学和生物学。教育学涉及农业人才培养计划、青年创业计划以及农场主、农业企业家等农业生产经营主体扶持等。工程学涉及堆垛机器人、无人机、收获臂、智能农机等农用智能装置、软硬件设备的研发,以及光谱分析、超声波、纳米技术、机器学习等先进技术在农业中的应用等。生物学涉及农业生物技术、微生物组和代谢组的相互作用关系、种植业和养殖业的高效生产与环境保护等。资助金额位于Top 3的学科分别是工程学、经济学和生物学,每个学科的平均资助金额达4 000万美元,是NIFA重点支持的学科方向。大多数学科的资助频次与金额成正比。

表1 2016—2022年NIFA资助的重点学科

序号	学科	资助频次/次	资助金额/万美元	学科类别
1	教育学	963	3 464	社会与行为科学类
2	工程学	724	4 071	物理科学类
3	生物学	670	3 905	生物科学类
4	生态学	650	3 504	生物科学类
5	营养与代谢	629	3 102	生物科学类
6	经济学	557	4 065	社会与行为科学类
7	遗传学	494	2 546	生物科学类
8	分子生物学	477	2 320	生物科学类
9	生理学	382	2 171	生物科学类
10	昆虫学与粉螨学	373	1 900	生物科学类
11	细菌学	363	2 154	生物科学类
12	化学	312	1 348	物理科学类
13	管理学	274	1 586	社会与行为科学类
14	信息和通信	249	2 238	社会与行为科学类
15	病理学	238	1 844	生物科学类

3.2.2 重点研究主题

根据NIFA基金项目数据中的知识分类字段,统计排名Top 20的重点研究主题,见表2。“土壤、植物、水、养分关系”涉及利用微生物、植物对土壤进行改良并研究其作用机制,提高粮食、经济作物产量和生产效率,提高农业生产水资源利用率等方面。“交流、教育和信息传递”涉及农业人才培育,农业新技术、新方

法, 生产经营能力提升, 培训计划等方面。“植物基因组、遗传学和遗传机制”涉及粮食、水果、蔬菜等植物新品种的开发与选育, 用转录组学等技术方法鉴定水果、蔬菜等植物发育过程的关键调控因子, 基因编辑、基因标记等技术应用及遗传机制研究等方面。

NIFA资助1 000万美元以上的5个重点方向是: ①工厂管理系统; ②土壤、植物、水、养分关系; ③动物疾病; ④交流、教育和信息传递; ⑤虫害综合管理系统。其中, “土壤、植物、水、养分关系”和“交流、教育和信息传递”是资助频次和资助金额“双高”的重点方向。

表2 2016—2022年NIFA资助的重点主题

序 号	研究主题	资助频 次/次	资助金额/ 万美元	序 号	研究主题	资助频 次/次	资助金额/ 万美元
1	土壤、植物、水、养分关系	583	1 547	11	工程系统和设备	250	196
2	交流、教育和信息传递	574	1 166	12	基础植物生物学	215	415
3	植物基因组、遗传学和遗传机制	406	276	13	营养教育与行为	207	138
4	动物疾病	380	1 317	14	青年发展	200	566
5	保护食品免受病原微生物、寄生虫的毒素污染	318	835	15	动物管理系统	194	617
6	昆虫、螨和其他影响植物的节肢动物	314	359	16	人口中的营养和饥饿	173	130
7	影响植物的病原体和线虫	312	338	17	节约和有效利用水	170	49
8	农业生产和农场管理经济学	310	332	18	动物的繁殖性能	165	143
9	工厂管理系统	293	1 826	19	新的和改进的食品加工技术	155	152
10	虫害综合管理系统	286	1 078	20	森林资源的管理和可持续性	152	301

3.3 学科交叉热点主题识别

学科交叉点往往就是新的学科增长点, 也是最可能产生重大科学突破和变革性创新的研究前沿^[9]。利用Python语言程序对项目进行学科共现分析与可视化, 绘制学科共现图谱, 见图2。节点代表学科, 节点越大中心度越高, 学科交叉研究越丰富。节点的连线代表共现关系。

2016—2022年NIFA资助的学科交叉研究比较丰富, 各学科大类呈现不同交叉特性: ①生物科学类多呈现学科内部的交叉融合, 其中生物学、生理学和生态学的交叉性最强; ②物理科学类和社会与行为科学类的学科数量相当, 多呈现学科外部交叉融合的特点, 如信息和通信、工程学和生物学之间的交叉融合; ③基础研究学科中生物化学和生物物理学、信息和通信、水文学交叉性较强, 应用性学科中流行病学、昆虫学与粉螨学交叉性较强。

为了进一步量化学科间的交叉程度, 借助Salton指数表征共现学科差异性, 以此直接测算学科之间

的交叉度。按照Salton指数从高到低, 统计排名Top 10的交叉度较高的学科组, 结合项目的知识分类字段提取学科组对应的研究主题作为热点主题, 见表3。

从表3来看, NIFA基金数据中, 交叉度最高的两个学科是昆虫学与粉螨学、杂草科学, Salton指数是0.242, 交叉的热点主题聚焦在虫害防治、作物保护方面。杂草科学与病理学的Salton指数次之, 交叉产生的杂草精准防治、植物病害等研究是NIFA支持的热点方向。整体来看, 学科交叉热点主题主要涉及病虫害防治、作物遗传改良、动植物育种、基因组学、基因调控等方面。近几年, 分子生物学与细胞生物学交叉产生的蛋白质组学研究、干细胞研究等也是重点方向; 数学和计算机科学与统计学、计量经济学和生物特征学交叉产生的精准农业、全基因组预测等是“农业4.0”的重要特征。

基于学科交叉分析可知, NIFA项目学科和研究主题的交叉性较为明显, 表现出以生物科学类学科间交叉为主, 物理科学类和社会与行为科学类学科逐渐与

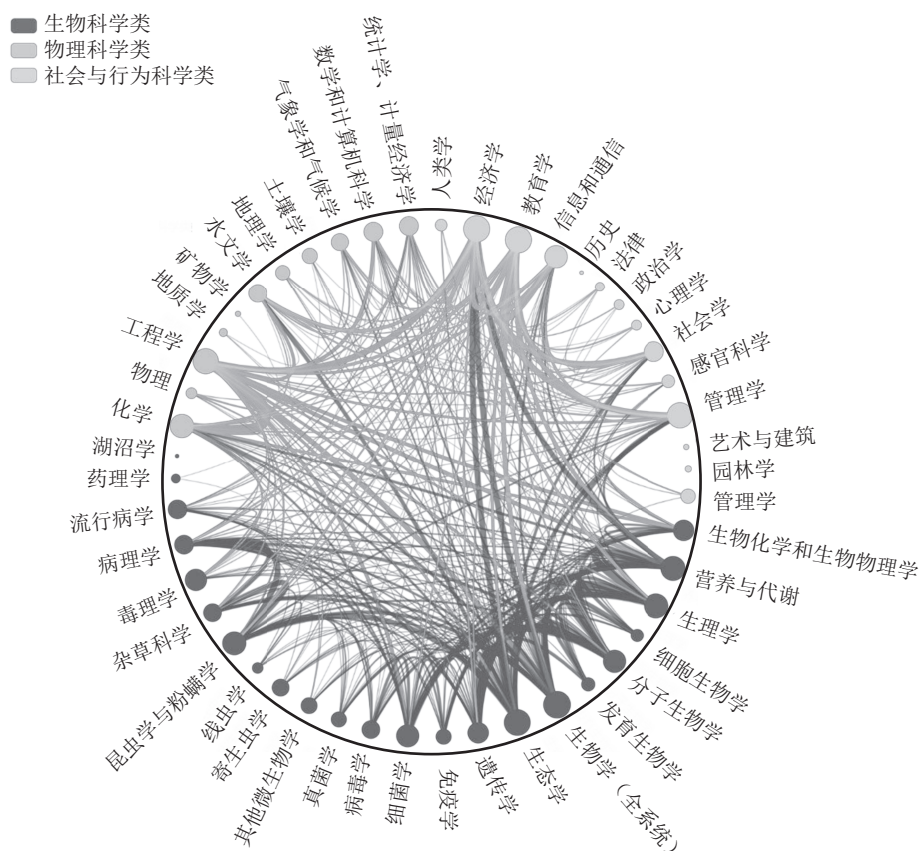


图2 2016—2022年NIFA资助项目学科共现图谱

表3 2016—2022年NIFA资助的学科交叉热点主题

学科1	学科2	Salton指数	热点主题
昆虫学与粉螨学	杂草科学	0.242	区域作物研究计划、有害生物综合治理、作物保护、虫害防治管理
病理学	杂草科学	0.237	有害生物综合治理、杂草精准防治、利用遗传和农艺改进作物生产、植物病害、植物病原
分子生物学	细胞生物学	0.206	蛋白质组学、细菌病原体抗生素耐药性、畜禽繁殖效率提升、转运蛋白、干细胞研究
昆虫学与粉螨学	病理学	0.198	作物病虫害防治、植物-病毒-食草动物相互作用机制研究、综合虫害管理、病原体传播控制
工程学	化学	0.193	生物炼制、电化学、纤维素生物燃料、纳米复合材料、巴氏杀菌、光化学、催化作用
发育生物学	细胞生物学	0.189	胚胎发育与营养、畜禽育种、干细胞研究、杂交繁育、生长激素、细胞因子增强
杂草科学	线虫学	0.177	植物寄生线虫、生物防治、虫害综合治理
数学和计算机科学	统计学、计量经济学和生物特征学	0.176	精准农业、大数据、生物多样性、遗传多样性、机器学习、可持续农业、全基因组预测
生理学	细胞生物学	0.167	内分泌调节、肌肉生长、畜禽繁殖、脂质代谢、干细胞、动物育种
生理学	发育生物学	0.161	畜禽育种、产前规划、基因调控、表型遗传、干细胞、生育能力、生殖细胞

农业领域交融的发展态势。教育学、信息和通信、工程学、生物学、生态学是资助的热点学科，且学科间的交叉融合程度较深。此外，以虫害防治管理、蛋白质组学、畜禽育种、精准农业、基因调控等为代表的研究主题

的交叉融合表现活跃。

3.4 基于战略坐标的主题分类与趋势分析

3.4.1 学科级战略坐标

如图3所示，为了更好地对比不同学科随着时间推



图3 NIFA资助学科战略坐标分布

移的发展趋势,根据“支持趋势-交叉趋势”战略坐标图,将项目数据划分到2016—2018年、2019—2021年两个时间窗口,提取NIFA项目数据中的学科字段,计算学科的支持趋势和交叉趋势指标。

整体来看,NIFA资助的大多数项目的学科交叉性呈上升趋势。

(1) 前沿型学科(第一象限)数量最多,占学科总量的34%,包括生理学、生物化学和生物物理学、微生物学、分子生物学、遗传学、数学和计算机科学等,具体方向如下。①智能育种,如作物表型智能分析系统、基因组-环境组智能育种方法探究、数据驱动的CRISPR基因组编辑设计方法。②智慧农业在生态环境领域的应用,如无人机遥感技术应用于土壤中水和养分监测与利用效率评估、作物智能灌溉程序与设备、精准施肥模式等。③动植物生长发育,如动物生长代谢引起的免疫和疾病研究、单细胞型蛋白质组学研究、应激性花粉不育、开花前耐旱和易感高粱品种的根系生长发育。④遗传性状与代谢,如商品肉牛群体性状的遗传机制、植物疾病基因鉴定及高效靶向基因工程、交叉代谢组学和转录组学在植物通路发现中的应用。

(2) 潜力型学科(第二象限)数量占比为31%,代表性学科为经济学、信息和通信、园林学、生物学、生态学、免疫学等,在农业领域这些学科的交叉研究日益丰富,主要围绕农业经济、生态环境、计算机技术等,具体方向如下。①计算机技术应用:利用机器学习获得高分辨率的土壤湿度和蒸散数据、人工智能助力农业

转型和决策支持、基于射频识别(RFID)相位和非线性滤波的牲畜精确定位跟踪等。②能源和生态及可持续发展:森林生物能源和生物燃料一体化研究、森林群落恢复力评估与提升研究、混合种植模式提高土地抗灾能力、废水在森林中可循环利用、温室气体减排和生态系统服务。③微生物组中的多物种相互作用:寄生虫个体、种群和群落的动态响应等。

(3) 衰退型学科(第三象限)数量占比为31%,代表性学科为社会学、气象学和气候学、水文学、真菌学、地理学等,在农业领域上述学科呈逐渐边缘化的衰退趋势,涉及的研究内容如下。①农业生态环境建设,如提高农业、自然/社会景观恢复力的研究,水质监测和预测模型,以及城市农业景观中的生物多样性、可持续性建设。②病菌的鉴定与检测,如黄瓜假霜霉病菌种群分析及抗霜霉病鉴定、群体基因组学分析病原菌的迁移和进化、农业环境中抗真菌药耐药性的检测及耐药因素的鉴定。

(4) 热点型学科(第四象限)相关项目数量有所增加,代表性学科为细菌学,具体方向如下。①新材料在动植物细菌检测工具中的应用,如用于快速检测病原体的纳米生物分子探针复合生物传感器系统,用于细菌病原体诊断的下一代测序方法。②食品领域,如噬菌体在生乳和生乳干酪中控制病原体的商业化应用、抑制大肠杆菌O157:H7在鲜切加工和冷藏过程中存活的生菜品种的鉴定。③病原体基因表达及作用机理,如病菌存活的转录组学分析、微生物群落的定植抗性。

3.4.2 主题级战略坐标

为了更好地对比不同研究主题随时间的发展趋势,根据“支持趋势-交叉趋势”战略坐标图,将项目数据划分至2016—2018年、2019—2021年两个时间窗口,提取NIFA项目数据中知识分类字段的研究主题,计算主题的支持趋势和交叉趋势指标,结果如下。

(1) 前沿型主题(第一象限)数量占比较多,是战略部署的重点方向,以新问题、新方法、新技术或新产品为主,包括植物基因组及遗传学,食品加工技术,新冠病毒冷链传播防控,土壤、植物和水的关系,动物营养利用等。①农业大数据、机器学习等精准识别和精准育种技术,如利用农业大数据识别与预测生鲜产品的微生物风险,通过机器学习进行精确牛体内噬菌体识别,表观遗传修饰物和动物器官发育的精准设计;②新型生物传感器、纳米传感器技术应用于检测食品生物分子或细菌,如食品过冷保鲜技术、智能化高温短时杀菌牛奶生产线、一次性激光诱导石墨烯生物传感器快速现场检测细菌技术;③新冠病毒传播防控措施及效果评估,如新冠病毒在供应链中传播机制、政府干预对阻断新冠病毒传播效果评估;④利用同位素量化土壤养分和肥力相关研究;⑤提高动物产量及效果评价研究。

(2) 潜力型主题(第二象限)的交叉性增加,多涉及融入农业领域的工具、方法等,应用领域包括虫害检测及综合管理系统、动植物生理过程和调节机制、农业资源可持续管理,比如:①应用机器学习预测分子生物活性、开发纳米材料小分子生物标志物传感器;②通过机器人采样和现场宏基因组测序检测早期植物病害、病害防治及智能决策系统研发;③分子生物技术在动植物生理过程和调节机制中的应用,微生物、矿物质代谢过程及其对农业资源可持续发展的影响,基因组学和表型等在农场管理中的应用。

(3) 衰退型主题(第三象限)的研究项目多具有时效性、地域性和应急性,研究主题包括预防和减轻污染、有毒化学品和有毒植物、流域保护和管理等,近年来NIFA对于上述主题的关注度有所下降。①开发新技术和新材料用于保护生态与减轻污染,如开发智能生物炭控释氮肥以提高玉米生产的可持续性、土壤生物降解塑料地膜的新技术、多功能交联碳纳米管/水凝胶纳米复合材料提高肥料利用效率和缓解干旱、天然防腐剂的绿色木材复合材料。②有毒化学品控制与管理,

如石墨烯的生物传感器的应用、利用高级氧化技术保障饮用水安全、利用纳米复合磷管理杂草。③地域流域性问题,如用于中西部常见除草剂多重传感的便携式现场生物传感器、农业密集区附近河流中新烟碱类化合物的光降解。

(4) 热点型主题(第四象限)以农业领域的传统性、基础性研究方向为主,NIFA对其的支持力度逐年增加,包括动物遗传改良、土壤资源评价、空气资源保护和管理、植物虫害监测识别、设施农业和农业设备方向。①动物遗传改良:基因编辑、分子育种技术和高通量测序技术辅助动物繁殖和选育,动物遗传性状的表达、基因组分析及转录组的挖掘,动物疾病抗性基因组测序和全基因组关联研究,功能基因组学在育种中的应用,影响动物生长的表观遗传调控机制。②土壤评价:基于深度学习的土壤参数监测模型、多功能多深度土壤传感器实时监测模型与管理系统、农业生态系统的新兴土壤健康框架。③设施农业:用于减少牲畜设施中甲烷的生物过滤器、农业场所温室气体的测量等。

基于坐标战略图分析可知,NIFA资助项目中,学科主题以前沿型和潜力型为主,衰退型次之,热点型主题数量最少。随着交叉学科的发展,科学研究项目呈现出以解决重大社会问题、科技问题为导向的特点,如:前沿型主题研究利用微生物学、分子生物学和遗传学、计算机科学的技术方法,解决植物基因组及遗传学、食品加工、食品安全等中的问题;潜力型主题多分布于生物学、信息和通信、工程学,应用于虫害综合管理系统、动物生理过程等研究;衰退型主题如有毒化学品和有毒植物、流域保护和管理等面向美国农业中的突发性、地域性局部问题;热点型主题则倾向于传统农业领域,如动物遗传改良、土地资源评价等。

3.4.3 结果验证分析

对识别的主题战略坐标分类结果进行验证。美国农业在世界处于领先地位,且越尖端、越前沿的科学研究越需要打破单一学科的限制,呈现出较高的学科交叉性^[15]。因此,选用由中国农业科学院农业信息研究所研制的《2022全球农业研究热点前沿分析解读》^[34]、中国工程院研制的《全球工程前沿2022》^[35]来验证所提方法的有效性。

《2022全球农业研究热点前沿分析解读》《全球工程前沿2022》分别以论文(2016—2022年)和专利

(2016—2021年)为数据源,结合多指标定量分析和专家定性判断,具有较高的权威性,与本研究数据时间范围比较符合,同时可从不同数据源角度验证主题分类的有效性和一致性。

将得出的主题战略坐标分类与《2022全球农业研究热点前沿分析解读》《全球工程前沿2022》进行比对与验证,见表4。前沿型主题、潜力型主题和热点型主题与《2022全球农业研究热点前沿分析解读》内容的重合率为92%,与《全球工程前沿2022》内容的重合率为57%,衰退型主题则均未出现在上述报告之中。综上,所提方法具有有效性。

表4 主题战略坐标分类结果对比与验证

战略坐标分类	主题内容	《2022全球农业研究热点前沿分析解读》	《全球工程前沿2022》
前沿型主题	植物基因组及遗传学	√	√
	食品加工技术	√	√
	新冠病毒冷链传播防控	×	×
	土壤、植物和水的关系	√	×
	动物营养利用	√	×
潜力型主题	虫害检测及综合管理系统	√	√
	动植物生理过程和调节机制	√	√
	农业资源可持续管理	√	×
	人口营养方向	√	×
热点型主题	动物遗传改良	√	√
	土壤资源评价	√	√
	空气资源保护和管理	√	×
	植物虫害监测识别	√	√
	设施农业和农业设备	√	√
衰退型主题	预防和减轻污染	×	×
	有毒化学品和有毒植物	×	×
	流域保护和管理	×	×

注：“√”代表有对应内容，“×”代表无对应内容。

4 结语

学科交叉研究是科学创新的重要途径,基于基金项目的学科交叉研究能够指引创新方向,表征领域研究热点与动态。本文主要贡献如下:①针对学科交叉主题识别中主题词提取方法单一、结果不易评判等问题,利用学科交叉测度指数与基金自带主题信息,识别出学科

交叉热点主题,提高了识别的准确性和可解释性;②通过构造支持度和交叉度指标,建立“支持趋势-交叉趋势”战略坐标图,将基金资助学科主题发展类型划分为前沿型、潜力型、热点型和衰退型,进一步明晰了学科交叉主题分类和发展趋势,以NIFA基金为例验证了方法的有效性。所提方法可以迁移至其他学科领域,所得结论也可为农业食品领域未来战略布局和科技创新提供一定的参考。

中国的跨学科研究起步晚,未来仍有许多可探索的空间^[36],比如两个以上学科交叉程度的测度与分析问题、特定领域学科交叉网络构建与测度指标等。同时本文构建“支持趋势-交叉趋势”二维战略坐标图,后续可考虑结合复杂网络指标或跨学科方法构建多结构分析框架,并探索学科交叉的新应用场景。

参考文献

[1] FONTANA M, IORI M, MONTIOBBIO F, et al. New and atypical combinations: an assessment of novelty and interdisciplinarity[J]. Research Policy, 2020, 49 (7): 104063.

[2] 肖敏, 江宁, 叶小梅, 等. 农业科研院所视角下新兴交叉学科发展的思考[J]. 农业科技管理, 2020, 39 (2): 17-20.

[3] 湛群芳, 陈积明. 交叉学科发展是“双一流”建设的新兴增长点[J]. 中国高等教育, 2018 (10): 42-43.

[4] 戴亚飞, 杜全生, 潘庆, 等. 探索中前行的交叉科学发展之路[J]. 大学与学科, 2021, 2 (4): 1-13.

[5] 刘小平, 冷伏海, 李泽霞. 国际科技前沿分析的方法和途径[J]. 图书情报工作, 2012, 56 (12): 60-65.

[6] ZHOU P, CAI X J, LYU X Z. An in-depth analysis of government funding and international collaboration in scientific research[J]. Scientometrics, 2020, 125 (2): 1331-1347.

[7] 张雪, 张志强. 美国科学基金会资助项目的学科交叉度演化规律及影响研究[J]. 情报理论与实践, 2021, 44 (12): 122-132.

[8] 范丽鹏, 王曰芬, 岑咏华, 等. 基金项目计划学部交叉及对前沿分布的影响研究: 以美国NSF数据中AI领域为例[J]. 情报学报, 2022, 41 (9): 956-966.

[9] 何娟. 建设农业强国, 掌握粮食安全主动权: 强国之路正扬帆⑨[N]. 人民日报, 2022-09-29 (7).

[10] 李佳蕾, 安培浚, 肖仙桃. 学科交叉主题识别方法研究综述[J]. 数据分析与知识发现, 2023, 7 (4): 1-15.

[11] 隗玲, 许海云, 郭婷, 等. 基于弱共现和突发监测的情报学学科研

- 究主题及交叉性分析[J]. 图书情报工作, 2015, 59 (21): 105-114.
- [12] DONG K, XU H Y, LUO R, et al. An integrated method for interdisciplinary topic identification and prediction: a case study on information science and library science[J]. Scientometrics, 2018, 115 (2): 849-868.
- [13] 李长玲, 崔斌, 刘小慧, 等. 作者潜在跨学科合作度模型构建与合作组合识别: 以图书情报学与计算机科学为例[J]. 情报理论与实践, 2017, 40 (7): 61-66.
- [14] 李长玲, 刘非凡, 郭凤娇. 运用重叠社群可视化软件CFinder分析学科交叉研究主题: 以情报学和计算机科学为例[J]. 图书情报工作, 2013, 57 (7): 75-80.
- [15] 陈琼, 朱庆华, 闵华, 等. 基于领域主题的学科交叉特征识别方法研究: 以医学信息学为例[J]. 现代情报, 2022, 42 (4): 11-24.
- [16] 阮光册, 夏磊. 学科间交叉研究主题识别: 以图书情报学与教育学为例[J]. 情报科学, 2020, 38 (12): 152-157.
- [17] 张振青, 孙巍. 基于特征测度和PhraseLDA模型的领域学科交叉主题识别研究: 以纳米技术的农业环境应用领域为例[J]. 数据分析与知识发现, 2023, 7 (7): 32-45.
- [18] CHI R B, YOUNG J. The interdisciplinary structure of research on intercultural relations: a co-citation network analysis study[J]. Scientometrics, 2013, 96 (1): 147-171.
- [19] HINZE S. Bibliographical cartography of an emerging interdisciplinary discipline: the case of bioelectronics[J]. Scientometrics, 1994, 29 (3): 353-376.
- [20] 刘扬, 张钰琛, 郑知敏, 等. 从国家自然科学基金项目申请代码分析的视角看中国基础研究的学科交叉性[J]. 中国科学基金, 2022, 36 (6): 986-993.
- [21] PORTER A L, CHUBIN D E. An indicator of cross-disciplinary research[J]. Scientometrics, 1985, 8 (3): 161-176.
- [22] BRILLOUIN L, HELLWARTH R W. Science and information theory[J]. Physics Today, 1956, 9 (12): 39-40.
- [23] SALTON G, MCGILL M J. Introduction to Modern Information Retrieval[M]. New York: McGraw-Hill, 1983.
- [24] SHANNON C E, WEAVER W. The mathematical theory of communication[J]. Bell Labs Technical Journal, 1950, 3 (9): 31-32.
- [25] PORTER A L, COHEN A S, ROESSNER J D, et al. Measuring researcher interdisciplinarity[J]. Scientometrics, 2007 (1): 72-76.
- [26] STIRLING A. A general framework for analysing diversity in science, technology and society[J]. Journal of the Royal Society, Interface, 2007, 4 (15): 707-719.
- [27] LEYDESDORFF L, WAGNER C S, BORNEMANN L. Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative variety, and the Gini coefficient[J]. Journal of Informetrics, 2019, 13 (1): 255-269.
- [28] RAFOLS I, MEYER M. Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience[J]. Scientometrics, 2010, 82 (2): 263-287.
- [29] 郭树行, 谈斯奇. 关键词共现研究趋势分析[J]. 科技资讯, 2011, 9 (32): 204-205.
- [30] HAMERS L, HEMERYCK Y, HERWEYERS G, et al. Similarity measures in scientometric research: the Jaccard index versus Salton's cosine formula[J]. Information Processing & Management, 1989, 25 (3): 315-318.
- [31] LAW J, BAUIN S, COURTIAL J P, et al. Policy and the mapping of scientific change: a co-word analysis of research into environmental acidification[J]. Scientometrics, 1988, 14 (3): 251-264.
- [32] 闵超, 孙建军. 学科交叉研究热点聚类分析: 以国内图书情报学和新闻传播学为例[J]. 图书情报工作, 2014, 58 (1): 109-116.
- [33] 袁毅, 季泽豪. 人工智能跨学科特性研究: 文献计量分析的视角[J]. 图书馆杂志, 2022, 41 (6): 46-52.
- [34] 孙巍, 李周晶, 马晓敏, 等. 2022全球农业研究热点前沿分析解读[J]. 农学学报, 2023, 13 (3): 6-9.
- [35] 中国工程院全球工程前沿项目组. 全球工程前沿2022[M]. 北京: 高等教育出版社, 2022.
- [36] 刘德娟, 沈力. 关于跨学科科学研究的国际经验及启示[J]. 世界科技研究与发展, 2023, 45 (5): 596-605.

作者简介

齐世杰, 女, 硕士, 助理研究员, 研究方向: 科学计量学。

赵静娟, 女, 硕士, 副研究员, 研究方向: 科技情报。

郑建华, 女, 博士, 高级工程师, 研究方向: 农业信息管理。

串丽敏, 女, 博士, 副研究员, 通信作者, 研究方向: 智能知识服务, E-mail: xiaochuan200506@126.com。

Identification and Classification of Interdisciplinary Hot Topics Based on Fund: Taking the U. S. Agriculture and Food Area as an Example

QI ShiJie^{1,2} ZHAO JingJuan¹ ZHENG JianHua² CHUAN LiMin¹

(1. Institute of Data Science and Agricultural Economics, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, P. R. China; 2. Key Laboratory of Knowledge Mining and Knowledge Services in Agricultural Converging Publishing, National Press and Publication Administration, Agricultural Information Institute of Chinese Academy of Agricultural Sciences, Beijing 100081, P. R. China)

Abstract: The interdisciplinary research has gradually become the source of original innovation achievements. The science fund is an important way to guide the original innovation. This article aims to accurately identify the interdisciplinary hot topics supported by science fund, grasp the research trends of interdisciplinary topics, and provide references for major scientific and technological innovation and prospective deployment. Based on data of fund projects, using the Salton index and co-occurrence analysis, we explore the interdisciplinary hot topics of the fund projects, construct the strategic coordinate map of “support trend-cross trend” according to the time windows, and position the disciplines/topics as cutting-edge, potential, declining, and hot ones. The effectiveness of the method is validated using the U.S. agriculture and food area as an example. The analysis shows that the interdisciplinary trend of the U. S. agricultural fund projects is becoming significant. Disciplines and topics are predominantly cutting-edge oriented and potential oriented, involving genomics and genetics, soil improvement and management, integrated pest management, and animal and plant breeding. The cross-integration of research topics represented by food processing technology, food protection, and animal nutrition utilization is active. We should improve the awareness of cross-breakthrough innovation, and use the strategic coordinate map to classify and lay out key technological innovation directions in key areas facing national needs.

Keywords: Agriculture and Food; United States; Interdisciplinarity; Topic Identification; Trend Strategic Coordinate

(责任编辑: 王玮)