

基于引用网络的医学主题领域跨学科知识交流研究^{*}

林紫洛 杨雪梅 关陟昊 于诗睿 唐小利

(中国医学科学院/北京协和医学院医学信息研究所/图书馆 北京 100005)

〔摘要〕 基于学科引用网络揭示医学主题研究领域的跨学科知识交流现状。具体流程包括根据文献间的引用关系构建领域学科网络,利用社会网络分析的中心度指标识别领域内的关键学科,并以核酸疫苗主题领域为实证,结合关键词聚类分析该领域核心学科跨学科引用的知识内容。

〔关键词〕 跨学科;知识交流;引用网络;社会网络分析;核酸疫苗

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2022.11.005

Study on Interdisciplinary Knowledge Exchange in Medical Subject Field Based on Citation Network LIN Ziluo, YANG Xuemei, GUAN Zhihao, YU Shirui, TANG Xiaoli, Institute of Medical Information & Library, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100005, China

〔Abstract〕 Based on disciplinary citation network, the paper reveals the current situation of interdisciplinary knowledge exchange in medical subject research field. Specifically, the discipline network is constructed according to the citation relationship between papers, and the key disciplines in the field are identified by using the centrality index of Social Network Analysis (SNA). The nucleic acid vaccine subject is taken as an example, and the knowledge content of interdisciplinary citation of the core disciplines in the field is analyzed by keyword clustering.

〔Keywords〕 interdisciplinary; knowledge exchange; citation network; Social Network Analysis (SNA); nucleic acid vaccine

〔修回日期〕 2022-04-20

〔修回简介〕 林紫洛,硕士研究生;通信作者:唐小利,研究馆员,硕士生导师。

〔基金项目〕 中国医学科学院医学与健康科技创新工程“生物医学文献信息保障与集成服务平台”(项目编号:2021-I2M-1-033);国家社会科学基金项目“基于科学与技术交叉模型的创新前沿识别方法与应用研究”(项目编号:18BTQ064)。

1 引言

科学研究并不局限于单一的学科,当代重大发现越来越依赖于多学科交叉融合。例如 DNA 双螺旋结构的发现是遗传学、生物化学、物理学等不同学科知识融合的结果,还开辟了分子生物学这一新的学科领域。我国大力支持跨学科相关研究,国家自然科学基金委员会在重大项目、重大研究计划项目和基础科学中心项目中均提出要促进学科的交叉与融合^[1]。随着科研进程的加快,学者们越来越倾向于综合运用多领域、多学科的知识来解决医学问

题,促使学科之间的联系越来越紧密。同时,随着科学文献数量呈井喷式增长,快速把握学科以外的科学研究进展与发展方向更加困难,这在一定程度上阻碍了科学的创新与技术的发展。因此,如何打破不同学科或领域之间的壁垒,找到学术交叉点,促进学科交叉与融合,是目前亟待解决的问题。本文旨在提出一种基于学科引用网络的方法,揭示医学主题研究领域的跨学科知识交流特征及主题,为研究人员快速定位领域中的学科交叉点、了解相关研究在不同学科的发展动向提供借鉴。

2 相关研究

“跨学科”也称“交叉学科”,英文表达为 Interdisciplinary,最早在 1926 年被 Woodworth R S 公开使用,刘仲林将其解释为超越一个已知学科的传统界域而进行的科学或教育活动^[2]。针对跨学科的研究包括跨学科性测度研究和主题识别,前者关注学科之间的交叉程度,后者则进一步明确学科交叉的内容特征。学者们往往从不同研究视角、使用不同的学科分类体系、依据不同测度指标或采取不同分析方法开展跨学科的研究。研究视角包括作者视角和文献视角,其中基于文献的研究更为普遍,且相较于分析合作作者的隶属关系,更能准确地表征跨学科过程中新知识的产生^[3]。一般以科研论文及其参考文献作为分析对象,用文献所属期刊的分类近似表示其学科类别^[4],例如 Porter A L 等^[5]对 6 个研究领域的期刊文章及其引文进行分析,测度领域间的跨学科程度与变化情况。当前的跨学科研究更多关注已有分类体系下具体的学科类别,而不是主题研究领域。研究者往往选择两个或多个学科^[6-8](如“图书情报学”和“教育学”),或选定某个明显的交叉学科^[9-11](如“纳米科学与纳米技术”)进行研究。前者锁定了少数学科,针对性较强,却易忽略其他可能发生交叉的学科;后者侧重于分析相对成熟的交叉学科,但有些研究领域尚未形成特定的学科,其中却集成多个不同学科的知识,存在潜在的学科间知识融合。此外,从具体学科类别出发进行分析,难以聚焦到某些研究对象或科学问题上,对主

题研究领域的跨学科知识交流发展帮助甚微。因此,本文沿用前人界定的“主题研究领域”概念,即科学研究中具体的研究对象^[12],探索医学主题研究领域的跨学科知识交流现状。

3 数据来源与方法

在活跃的学科关键知识节点引入跨学科相关知识,将产生创新知识的生长点^[13],其中包括识别学科关键节点和发现跨学科相关知识这两个重点。本文的研究思路与之契合,拟解决两个主要问题:一是了解医学主题研究领域内的学科分布,识别其中的关键学科,二是进一步挖掘跨学科交流的知识内容。方法路径,见图 1。

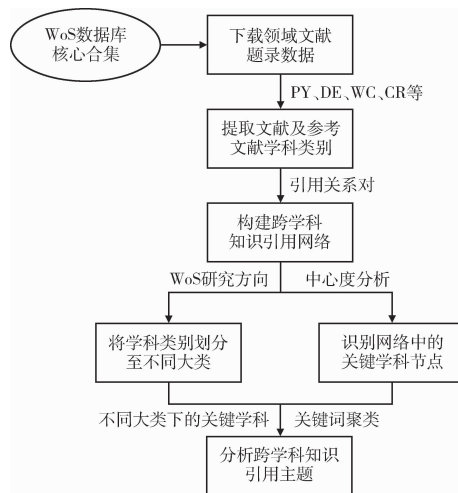


图 1 研究框架

3.1 数据获取及预处理

科学研究成果多以论文为载体发表,本文以科研论文及其参考文献为研究对象。Web of Science (WoS) 数据库核心合集将期刊分为 254 个学科类别,每本期刊均对应 1 到多个学科类别,且从 WoS 数据库中下载的题录提供参考文献信息,有较完整的数据可供分析。故本文以 WoS 数据库为来源,检索领域相关文献,以期刊的学科类别表示文献或参考文献所属学科。1 篇文献引用多篇参考文献,相当于文献所属学科引入多个参考文献所属学科的知识内容,这一过程中形成的多个引用关系对,保存

以后用于后续研究。

3.2 领域文献及其参考文献的跨学科网络构建

社会网络分析方法是一种研究社会结构、分析个体之间关系的方法，可用于研究引用关系或合作关系^[14]。学科之间通过引用关系建立连接，产生知识流动，知识由被引学科流向施引学科，多个学科发生联系时便形成网络。因此，本文基于核酸疫苗领域文献与其参考文献之间的引用关系，构建跨学科知识引用网络，网络中的节点代表学科，箭头由施引学科指向被引学科，表示前者引用后者的知识内容，产生跨学科知识交流现象。此外，由于 WoS 的学科分类粒度较细，不同学科可能同属于一个大类、拥有相近的研究方向，为找到知识区分度较大的两门或多门学科，进而分析相互之间的跨学科交流，参考 WoS 提供的各学科类别对应的研究方向^[15]，将学科归到不同大类下。

3.3 基于网络中心度的关键学科节点识别

在社会网络分析中，中心度分析常被用于测度某一节点在网络中的位置，用以判断该节点的重要性与影响力，其中点度中心度和中介中心度是最常用的指标^[16]。点度中心度指与某个节点直接相连的节点数，在有向网络中又可分为入度中心度与出度中心度。在本文构建的学科网络中，入度中心度越高说明该学科的知识更多地流向多个学科；相反，出度中心度越高说明跨学科知识的流入量越多。中介中心度指该节点有多大程度处于其他两个节点最短路径的中间，起到“中介”的作用，Leydesdorff L 和 Rafols I^[17]认为中介中心度可以被视为潜在的跨学科衡量指标。某个学科节点的中介中心度越大，说明其在网络中的位置越重要、控制能力越强。本文使用点度中心度和中介中心度来衡量学科在网络中的重要程度，并借鉴黄水清、张俊和阎素兰^[18]提出的层级划分方法，将点度中心度和中介中心度排名前 14.6% 的学科视为关键学科。

3.4 基于关键词聚类的跨学科引用主题分析

学科的所属大类代表该学科的大致研究方向，而

学科节点的中心度反映该学科的重要程度，将两者综合考虑，对不同研究方向下的两门或多门学科进行跨学科引用主题分析，见图 2，箭头指向表示引用方向，A 类关键学科 a1 引用 B 类关键学科 b1、b2、b3，即学科 b1、b2、b3 的知识流向学科 a1，识别 b1、b2、b3 相关文献的主题，可以了解学科 a1 主要引用 B 类学科的哪些知识内容。关键词是论文作者对文章主要内容的提炼，能直观反映文章主题，因此本文使用关键词聚类方法进行文献主题挖掘。具体做法如下：获取某个学科引用其他大类学科的参考文献，提取关键词后采用 Gephi 提供的模块化功能进行关键词聚类，每个簇视为 1 个研究主题。

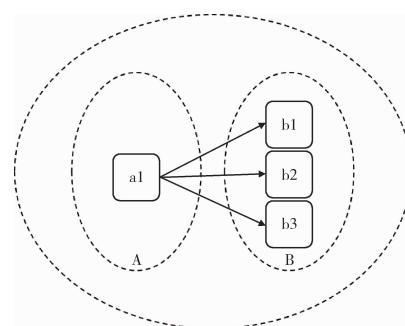


图 2 跨学科引用主题分析示例

4 实证分析

4.1 核酸疫苗领域学科总体情况

近年来 mRNA 疫苗受到了广泛关注，相关研究成果也越来越多，被《麻省理工科技评论》评为 2021 年“全球 10 大突破性技术”^[19]，因此本文选定核酸疫苗领域进行实证研究。在数据获取阶段，以 WoS 数据库核心合集作为数据来源，检索式为：TS = (“nucleic acid * vaccin * ” OR “DNA * vaccin * ” OR “DNA - based vaccin * ” OR “genetic vaccin * ” OR “nucleic acid * immunization” OR “polynucleotide vaccin * ” OR “RNA * vaccin * ” OR “RNA - based vaccin * ” OR “mRNA vaccin * ” OR “mRNA - based vaccine * ”)，并将文献类型限制在 Artical、Review、Letter、Proceedings Paper 范围内，由于本文暂不从时间维度考察跨学科引用，故不限定时间跨度。截至 2021 年 8 月 16 日，共得到

12 589 篇文献,并提取出 22 416 条“文献-学科类别”关系(1 篇文献可能对应多个学科类别)。对结果进行统计,把握核酸疫苗领域内论文的整体情况。核酸疫苗的相关论文最早发表于 1987 年,1987—2021 年间核酸疫苗相关论文的发文量与涉及的学科类别数量,见图 3。可以看出核酸疫苗自 1995 年开始受到大量关注,2006 年以后热度稍退,近几年又重新引起研究者重视,同时领域涵盖的学科类别数量总体上呈增长趋势,说明核酸疫苗领域的研究问题越来越广,融合了越来越多学科的知识内容。此外,获得 677 619 条参考文献记录,删除信息不全文献后,将这些参考文献映射到学科类别上,得到 941 682 条参考文献学科类别记录。根据“文献学科类别-参考文献学科类别”对应方式,构建学科间引用关系对,如“Virology-Immunology”表示在病毒学的研究中引入免疫学相关知识。

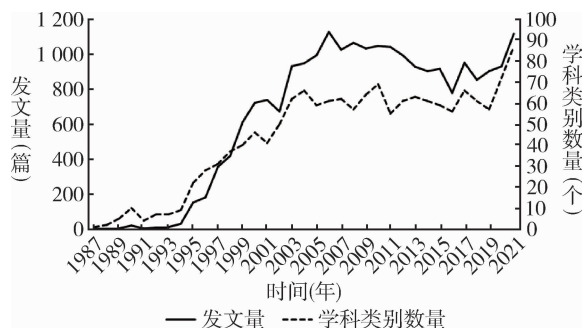


图3 核酸疫苗相关发文年度分布及学科类别数量

4.2 跨学科知识引用网络构建

为了揭示核酸疫苗领域跨学科知识交流的特点,基于学科之间引用关系构建跨学科知识引用网络,并进一步分析网络结构特征。所构建的学科网络包括 230 个节点,占 WoS 学科类别总数的 90.55%,网络的图密度高达 0.148,该学科网络不仅规模较大,而且各学科之间的知识交流频繁。网络中心的学科节点度较大且较为密集,越往外越稀疏,说明在核酸疫苗领域有一部分学科的跨学科知识交流较为活跃,这部分学科可能是领域内的关键学科。在网络中按学科数量从高到低排列分别是生命科学与生物医学、应用科学、社会科学、自然科学、艺术与人文。位于中心位置的学科主要是生命

科学与生物医学大类下的,其次是应用科学和自然科学,说明这些大类在领域内较为关键,对应的学科值得重点关注。此外,社会科学的学科节点大多散布在网络的周围,说明不是领域内参与知识交流的重点学科,只是相关知识内容得到了引用。

4.3 核酸疫苗领域关键学科分析

对学科网络中的节点进行中心度分析,并按照中心度数值高低进行排序,中心度越高则表示该学科越重要、在网络中的位置越核心。本文构建的学科网络包括 230 个学科节点,取各类中心度靠前的 34 个(前 14.6% 约为 34)学科类别作为核酸疫苗领域的关键学科。入度中心度、出度中心度和中介中心度排名前 3 的学科类别依次都是免疫学(immunology)、医学-研究与实验(medicine, research & experimental)、生物技术和应用生物学(biotechnology & applied microbiology),在网络中处于核心地位。此外,具有高入度中心度的学科还包括多学科科学(multidisciplinary sciences)、生物化学与分子生物学(biochemistry & molecular biology)等,这些学科的知识流向了较多的学科,对其他学科有较高的影响力;具有高出度中心度的学科还有药理学与药剂学(pharmacology & pharmacy)、生物化学与分子生物学等,这些学科在跨学科知识交流上较为活跃,在研究问题时引入更多其他学科的知识;而中介中心度较高的学科还有生物化学与分子生物学、药理学与药剂学等,这些学科可能在其他学科之间的知识交流中起到桥梁作用,具有较强控制能力。在学科网络中找到代表这些关键学科节点与它们之间的连线,见图 4,构成核酸疫苗领域学科引用网络的主干部分。主干网中包括生命科学与生物医学、自然科学、应用科学 3 个大类的学科,其中,生命科学与生物医学的关键学科中心度大多高于其他两类的学科,在核酸疫苗领域发挥主要作用;自然科学的关键学科均与物理、化学相关;应用科学的关键学科则涉及材料、工程方面的知识。

4.4 以免疫学学科为例的跨学科引用主题分析

生命科学与生物医学、自然科学、应用科学 3

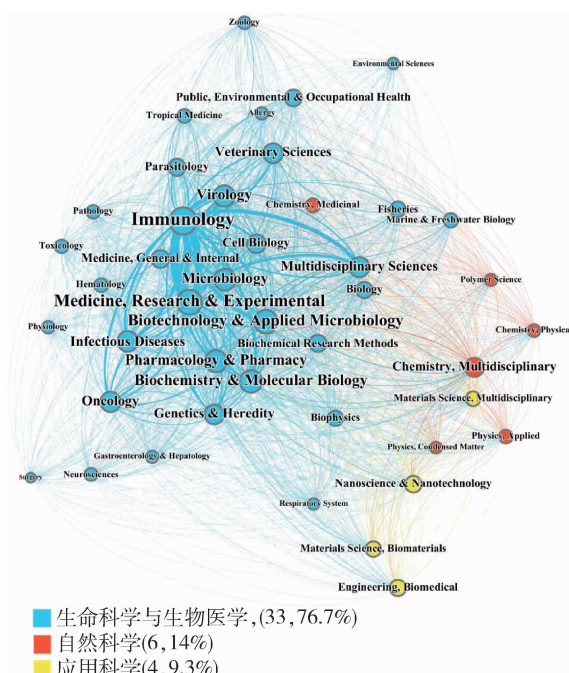


图4 核酸疫苗领域学科引用网络主干部分

个研究方向的区分度较大,围绕核酸疫苗问题展开讨论,为探究单个学科在引用其他研究方向的学科时具体参考了哪些知识内容,采用关键词聚类进行主题分析。相比起其他学科,免疫学各项中心度均最大,说明其在核酸疫苗领域是最为核心的学科,因此本文以免疫学为例,从前期获取的数据中抽取与该学科相关的文献,根据文献内容分析其跨学科引用主题。免疫学属于生命科学与生物医学类,除自身研究方向外,主要引用自然科学类的化学-多学科(chemistry, multidisciplinary)、医药化学(chemistry, medicinal)、应用物理(physics, applied)等学科,共817篇论文,以及应用科学类的生物医学工程(engineering, biomedical)、纳米科学与纳米技术(nanoscience & nanotechnology)、生物材料科学(materials science, biomaterials)等学科,共370篇论文。分别提取两个大类下相关学科文献的关键词,将关键词统一为小写形式并还原词性,为确保聚类效果良好,对高频词 vaccine 予以剔除,进而分析关键词的共现情况,最终选取主要的4个主题簇进行可视化展示。免疫学引用的自然

科学文献关键词共现结果,见图5。主题1中的高频词有纳米颗粒(nanoparticle)、脂质体(liposome)、药物递送(drug delivery)、壳聚糖(chitosan)等,主要涉及纳米颗粒在药物递送系统中的应用,相似技术迁移到核酸疫苗领域中,反映出纳米颗粒载体良好的疫苗递送能力;主题2中的高频词有基因递送(gene delivery)、基因治疗(gene therapy)、转染(transfection)、DNA等,主要是与基因治疗相关的知识内容,这是由于核酸疫苗技术是从基因治疗领域发展起来的;主题3中的高频词有辅助(adjuvant)、树突状细胞(dendritic cell)、DNA疫苗(DNA vaccine)、免疫反应(immune response)等,主要内容与树突状细胞有关,树突状细胞是一种强大的抗原呈递细胞,可以增强核酸疫苗和疫苗辅助剂的免疫反应;主题4中的高频词有微针(microneedle)、皮肤(skin)、疫苗接种(vaccination)等,主要内容是疫苗的给药方式,随着微针、纳米微针的日益成熟,透皮给药技术快速发展,在核酸疫苗给药上也能够发挥重要作用。免疫学引用的应用科学文献关键词共现结果,见图6。主题1中的高频词有纳米颗粒(nanoparticle)、壳聚糖(chitosan)、药物递送(drug delivery)、免疫疗法(immunotherapy)等,与自然科学共现结果的主题1类似,主要内容与纳米颗粒递送载体有关;主题2中的高频词有DNA疫苗(DNA vaccine)、辅助(adjuvant)、免疫反应(immune response)、碳纳米管(carbon nanotube)等,主要内容与核酸疫苗引起的机体免疫反应有关,碳纳米管作为近些年的热门纳米材料之一,能够激发人体的免疫反应,提高特异性抗体水平;主题3中的高频词有covid-19、sars-cov-2、mma vaccine(mRNA疫苗)等,主要内容是mRNA疫苗技术的应用;主题4中的高频词有树突状细胞(dendritic cell)、抗体(antibody)、金纳米颗粒(gold nanoparticle)、共递送(co-delivery)等,主要内容是使用纳米颗粒作为载体,将抗原等多种组分同步递送至树突状细胞。

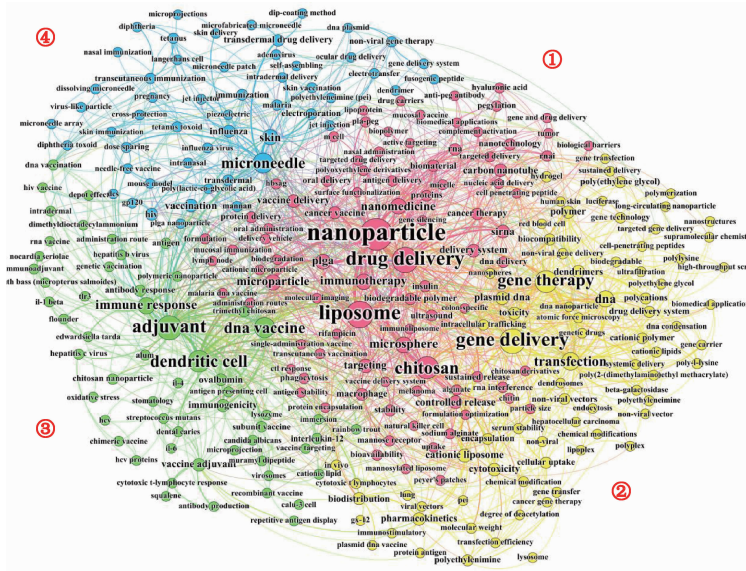


图 5 免疫学引用的自然科学文献关键词聚类结果

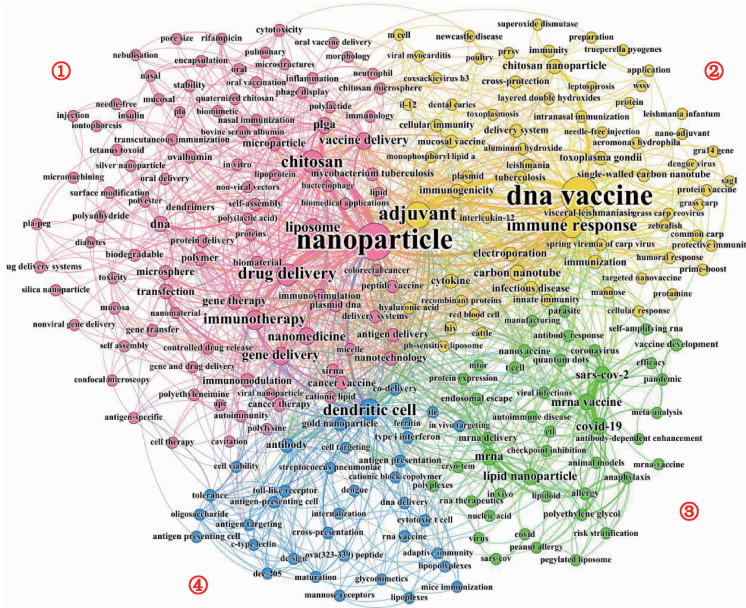


图 6 免疫学引用的应用科学文献关键词聚类结果

5 结语

本文基于文献间的引用关系构建医学主题研究领域跨学科知识引用网络，考察学科分布情况，并结合关键词聚类分析跨学科引用主题。在实证部分，以核酸疫苗为例展开研究，识别出核酸疫苗领域的关键词免疫学、医学—研究和实验、生物技术和应用生物

学等，领域内核心学科免疫学引用的自然科学类知识内容与纳米颗粒递送载体、基因治疗、树突状细胞增强免疫反应、微针透皮给药等相关，引用的应用科学类知识内容与纳米颗粒递送载体、机体免疫反应、mRNA 疫苗技术、纳米颗粒共递送多种组分至树突状细胞等相关，有助于揭示核酸疫苗领域的跨学科知识交流态势，帮助研究人员快速定位学科交叉点，明确学科间知识交流的内容。

参考文献

- 1 国家自然科学基金委员会. 资助格局 [EB/OL]. [2021-08-14]. <http://www.nsf.gov.cn/publish/portal0/jg-sz/08/default.htm#02>.
- 2 刘仲林. 交叉科学时代的交叉研究 [J]. 科学学研究, 1993 (2): 11-18.
- 3 Rafols I, Meyer M. How Cross-disciplinary is Bionanotechnology? Explorations in the Specialty of Molecular Motors [J]. Scientometrics, 2007, 70 (3): 633-650.
- 4 Porter A L, Roessner J D, Cohen A S, et al. Interdisciplinary Research: Meaning, Metrics and Nurture [J]. Research Evaluation, 2006, 15 (3): 187-195.
- 5 Porter A L, Rafols I. Is Science Becoming more Interdisciplinary? Measuring and Mapping Six Research Fields over Time [J]. Scientometrics, 2009, 81 (3): 719-745.
- 6 商宪丽. 基于 LDA 的交叉学科潜在主题识别研究——以数字图书馆为例 [J]. 情报科学, 2018, 36 (6): 57-62.
- 7 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.
- 8 徐庶睿, 卢超, 章成志. 术语引用视角下的学科交叉测度——以 PLOS ONE 上六个学科为例 [J]. 情报学报, 2017, 36 (8): 809-820.
- 9 Xu H, Guo T, Yue Z, et al. Interdisciplinary Topics of Information Science: a Study Based on the Terms Interdisciplinarity Index Series [J]. Scientometrics, 2016, 106 (2): 583-601.
- 10 Dong K, Xu H, 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.
- 11 韩正琪, 刘小平, 寇晶晶. 基于 Rao-Stirling 指数和 LDA 模型的领域学科交叉主题识别——以纳米科技为例 [J]. 情报科学, 2020, 38 (2): 116-124.
- 12 叶春蕾. 基于 Web of Science 学科分类的主题研究领域跨学科态势分析方法研究 [J]. 图书情报工作, 2018, 62 (2): 127-134.
- 13 李长玲, 高峰, 牌艳欣. 试论跨学科潜在知识生长点及其识别方法 [J]. 科学研究, 2021, 39 (6): 1007-1014.
- 14 Otte E, Rousseau R. Social Network Analysis: a Powerful Strategy, also for the Information Sciences [J]. Journal of Information Science, 2002, 28 (6): 441-453.
- 15 Web of Science. Research Areas [EB/OL]. [2021-11-04]. http://images.webofknowledge.com/WOKRS534DR3/help/zh_CN/WOS/hp_research_areas_easca.html.
- 16 王旻霞, 赵丙军. 跨学科知识交流网络结构特征研究 [J]. 情报科学, 2016, 34 (5): 46-50, 104.
- 17 Leydesdorff L, Rafols I. Indicators of the Interdisciplinarity of Journals: Siversity, Centrality, and Citations [J]. Journal of Informetrics, 2011, 5 (1): 87-100.
- 18 黄水清, 张俊, 阎素兰. 黄金分割法在学科及机构评价中的应用 [J]. 图书情报工作, 2012, 56 (22): 33-36.
- 19 Review M T. MIT Technology Review Presents 10 Breakthrough Technologies of 2021 [EB/OL]. [2021-11-15]. <https://www.technologyreview.com/press-releases/mit-technology-review-presents-10-breakthrough-technologies-of-2021/>.
- 38 Singh A, Hosanagar K, Gandhi A. Machine Learning Instrument Variables for Causal Inference [EB/OL]. [2021-12-20]. https://www.researchgate.net/publication/342899660_Machine_Learning_Instrument_Variables_for_Causal_Inference.
- 39 Belloni A, Chen D, Chernozhukov V, et al. Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain [J]. Econometrica, 2010, 80 (6): 2369-2429.
- 40 Tamma P D, Turnbull A E, Harris A D, et al. Less Is More: Combination Antibiotic Therapy for the Treatment of Gram-negative Bacteremia in Pediatric Patients [J]. JAMA Pediatrics, 2013, 167 (10): 903-910.
- 41 McCaffrey D F, Ridgeway G, Morral A R. Propensity Score Estimation with Boosted Regression for Evaluating Causal Effects in Observational Studies [J]. Psychological Methods, 2004, 9 (4): 403-425.
- 42 Qiu Y, Chen X, Shi W. Impacts of Social and Economic Factors on the Transmission of Coronavirus Disease 2019 (COVID-19) in China [J]. Journal of Population Economics, 2020, 33 (4): 1127-1172.

(上接第 26 页)