

·实践平台·

Web of Science 和 Google Scholar 引文检索功能比较

耿海英 肖仙桃 (中国科学院国家科学图书馆兰州分馆 甘肃兰州 730000)

摘要: 文章从检索范围、检索方法、检索结果以及检索效果等方面对 Web of Science 和 Google Scholar 的引文检索功能进行了比较分析, 指出它们各自的优势和不足。

关键词: Web of Science Google Scholar 引文检索

中图分类号: G354.4

文献标识码: A

文章编号: 1003-6938(2007)03-0100-03

A Comparative Study on the Citation Search Function between Web of Science and Google Scholar

Geng Haiying Xiao Xiantao (Lanzhou Branch of National Science Library, CAS, Lanzhou, Gansu, 730000)

Abstract: This paper compares Web of Science database with Google scholar retrieval system on the aspects of citation search function, including searching scope, searching methods, searching results and searching effect, and points out the advantages and disadvantages of each database.

Key words: Web of Science; Google scholar; citation search

CLC number: G354.4

Document code: A

Article ID: 1003-6938(2007)03-0100-03

随着 1964 年 Science Citation Index (SCI) 的正式发行和普及, 引文分析在科学研究和期刊评价方面的作用早已为人知晓, 基于引文的检索已经成为一种惯用的检索方法, 而 SCI 也因此在全球学术界牢固地确立了自己的地位。现在 SCI 已成为 Web of Science (WOS) 中的一部分。

传统的文献公开方式主要以期刊发表为主, 因此期刊也就成为人们获取学术文献的主要途径, 通过 WOS 即可实现对多数重要期刊上文献的检索, 同时通过测量期刊文章被引次数即可分析文章的影响力。但是随着网络的发展, 学术交流模式出现重大变化, 研究人员有更多新的方式来公开、散布其作品, 诸如通过预印本库、开放获取性期刊或个人主页。这些新的文献公开方式不仅是对 WOS 用于文献检索的挑战, 更是对传统引文分析的挑战。2004 年底 Google 公司推出的

Google Scholar (GS) 为这种新形式公开的学术文献检索和新型的引文分析提供了探索性的解决方案。现将 WOS 和 GS 的引文检索功能进行比较分析。

1 WOS 与 GS 简介

1.1 WOS

美国科技信息研究所 (Institute for Scientific Information, ISI) 于 1964 年正式发行 Science Citation Index (SCI)。2000 年, ISI 推出 ISI Web of Knowledge 学术信息资源整合体系, 其中以 Web of Science (WOS) 为核心。通过 WOS 可以直接访问 ISI 的三大引文数据库 Science Citation Index Expanded (SCIE)、Social Science Citation Index (SSCI)、Arts & Humanities Citation Index (A&HCI) 三个数据库。WOS 涵盖了自然科学、社

会科学、人文艺术领域中的所有学科,收录了全球 9300 份权威的、高影响力的学术期刊,数据可以一直回溯到 1900 年。

1.2 GS

GS 是世界上最强大的搜索引擎公司 Google 推出的学术搜索工具。GS 由计算机专家 Anurag Acharya 开发。目前,GS 已经与多家学术、科技和技术出版商进行了广泛的合作,如 ACM、Nature、IEEE、OCLC 等。^[1]这种合作使用户能够检索特定的学术文献,通过 GS 从学术出版者、专业团体、预印本库、大学范围以及从网络上获得多种类型的学术文献,其中包括学术期刊、文摘、同行评议论文、学位论文、图书、预印本、PPT 介绍和技术报告等。^[2]GS 已成为获取网上学术信息的一个重要通道。GS 重点提供医学、物理、经济、计算机等学科文献的检索,还可通过知识链接功能提供文章的引用次数及链接,查找文献的被引用情况。

2 WOS 和 GS 的引文检索功能比较

2.1 检索范围

WOS 检索根据其所收录的、经过筛选的期刊,选择各学科领域权威的、影响力高的期刊,由于受语言的限制,WOS 收录的英文期刊较多,如 SCIE 收录的中国大陆期刊只有 75 种(2005 年)。因此,WOS 检索范围也以学科核心期刊为主,以英语语种文献为主。

GS 由于其搜索的文献不仅仅局限于传统形式的期刊文献,还包括了网上多种非传统形式公开的文献,因此,GS 不仅可以免费搜索跟踪同行评议过的文献的引证文献(这种与 WOS 所作的工作一样),还能搜索跟踪会议录、学位论文、预印本和在版文献,以及其他非传统媒体形式文献的引证文献。^[3]且检索语种范围不局限于英语,还包括各种语种。

2.2 检索方法

利用 WOS 查找文献引用和被引用信息的方法主要有两种:第一种,首先利用普通检索或高级检索来获取一批相关文献。然后通过每篇相关文献全记录页面中提供的“Cited References”和“Times Cited”功能来查找引用和被引用信息。“Cited References”提供该篇文献所引用的参考文献。“Times Cited”提供该篇文献在发表后被他人引用的情况。全记录页面的右方有“Create Citation Alert”,允许用户通过设置引文跟踪服务来跟踪某一文献的最新被引用情况。第二种,利用引文检索(Cited Reference Search)方式查找某一特定文献被引用的情况。可分别从被引作者(Cited Author)、被引文献

(Cited Work)、引文年限(Cited Year)等字段入手查找引证文献信息。被引作者除第一作者外,还可以是次要作者(第二以上作者),也可以是团体作者。被引文献除期刊名称外,也可以是图书、专利和政府报告的题名。^[4]

GS 提供普通检索和高级检索两种检索功能。普通检索中可直接输入关键词、人名、期刊名,也可输入诸如“人名 期刊名称”等组配形式缩小搜索范围,获取相关文献及文献的被引频次。^[5]高级搜索可以进行字段限定,如可以按作者、发表日期和出版物等字段进一步缩小搜索范围,更有针对性地检索想要查找的学术资料的被引信息。通过两种检索获取一批源文献,在每篇文献详细记录的下面提供该篇文献的被引情况(Cited by),显示了该篇文献的被引次数,点击链接即可查看所有引用过该篇文献的文献列表。

2.3 检索结果

WOS 的检索结果,可根据引证文献的最近日期、被引次数、相关度、第一作者、标题和出版年进行排序。在文献检索页面和文献引用关系检索页面有“Analyze Results(分析结果)”按钮,可对引证文献从作者、国家、文献类型、机构、语种、出版年代、来源文献、学科主题等角度进行排序分析,从中可发现诸如主要的引用作者和机构、引证文献的年代分布等重要信息。分析结果包括各项指标的名称、相关记录数、占总记录数的百分比以及百分比图示,所有分析结果还可以导入 Excel 表格。WOS 也可将引文检索结果按照多种灵活的格式进行打印、保存、发送电子邮件或者输出到文献管理软件中,便于更深层次的数据分析。

GS 检索结果由文章和随后发表的引用此文章的文章列表两部分组成,并能自动抽取分析引用文献把他们作为单独的分析结果呈现出来。GS 检索结果排序主要以文献的被引频次为依据,同时兼顾作者、出版物知名度等因素。^[5]被引频次越高,说明该文献的学术价值越高,其中被引频次包括文献在书籍和各类非联机出版物中的引用;文献作者越出名,其学术专业的价值地位就越高,排序时往往越靠前;能在国际知名刊物上刊出论文,其内容价值相对较高。但 GS 没有提供检索结果深层分析功能,而且无法将检索结果保存或导出。

从以上内容可以看出,WOS 作为一个专门的引文数据库,拥有强大的引文检索功能。可迅速查找出含有特定内容(被引作者、被引文献、被引年代)的引证文献,并可进一步得到这些文献的引用(或被引用)信息,不仅可以回溯某一研究文献的起源与历史(Cited References)还可追踪其最新的进展

(Cited References),既可以越查越旧,也可以越查越新,越查越深入。但 WOS 也受其收录期刊和语种的限制。GS 的最大亮点就是能搜索网上各种形式的文献,而不仅仅局限于传统学术数据库,为探索网上各种形式公开的学术文献间的引用关系提供了参考。但 GS 提供的检索方式太少,不支持主题词或受控检索;检索结果排序方式少,不能对检索结果进行分析。

2.4 案例检索效果

为了直观地比较这两个数据库的检索效果,本文选取 White HD 于 2003 年发表的《Pathfinder networks and author cocitation analysis: A remapping of paradigmatic information scientists》(以下简称 A 文)进行试检索。测试时间为 2007 年 1 月 9 日,考察 WOS 和 GS 对同一篇文章的引用统计。WOS 检索显示该文被引 19 次,GS 中则显示被引 36 次。追踪引用文献, WOS 查找到 19 条记录,其中包括一篇社论,故 WOS 中最终搜到的被引次数为 18 次; GS 的引证文献中有两篇文献分别重复了两次和一次,另有一篇文献为社论,故 GS 中最终搜到的被引次数为 32 次。进一步逐篇核对两者的引证文献,得到 A 文的被引次数比较如下:

A 文在 WOS 和 GS 的被引次数

GS 中的被引次数	WOS 中的被引次数	只出现在 GS 中的引证文献数	只出现在 WOS 中的引证文献数	两者共有的引证文献数
32	18	18	4	14

WOS 中独有的 4 篇引证文献中 3 篇都能在 GS 中检到,只有 1 篇未能找到,进一步核对这 4 篇文献的参考文献,发现参考文献里确实包括 A 文,此情况说明 GS 并没完全揭示出文献间的引用关系。而 GS 中独有的 19 篇引用文献主要是会议论文、开放获取性文献和非英语语种文献,其中包括两篇中文文献:《基于可视化的作者同被引技术的发展》、《JASIST 与情报科学历程》。这也正体现了 GS 获取学术文献途径的多样化。

3 结论

通过以上多方面的比较以及笔者的检索实践,得出以下结论:

WOS 被全球学术界公认为权威的引文文献检索工具,其收录数据回溯时间长,适合于进行引文回溯检索,可以有效探索一个创新思想、概念或一个方法从其最初提出到当前的历史、发展与应用。其更新快速,研究人员可随时了解所在研究

领域里最新出版的文献情况,追踪当前的研究热点。其独特的引文检索方式,可以方便、快捷地统计出某个作者、某种期刊的被引用数据,由此可定量评价其学术影响力。但也正是由于它的学术性和权威性使得其不能对现今网上多种非传统形式公开的学术文献进行索引,因此利用它进行引文检索的全面性也就受到限制。且由于受购买年限的限制,查找购买年限以前的文献被引用情况时,只能查到引用次数,无法了解引证文献的详细情况。

GS 作为一种新型的引文检索工具,具有许多优势:不存在期刊选择过程中的主观偏见,检索范围不局限于已发表论文,还包括预印本、技术报告、学位论文和会议论文等;检索不受语种的限制,不局限于欧洲各国语种文献,还包括中文、日文等语种文献;能够识别不同形式的引用形式,如图书、Word 文档。最重要的是,GS 为网上文献的引文检索开辟了新的途径,研究人员通过 GS 可以对网络中隐藏的主题关系实行全方位、多学科的搜索挖掘,为基于网络影响因子的学术评价提供了方便的工具。再加上它的强大搜索功能和超大规模数据库的支持以及原始文献的在线提供等优势,基于 GS 的文献计量学研究也将成为新的热点,同时也为网络计量学研究提供了新工具。^[6]但利用 GS 进行引文检索时也存在众多缺陷,如对于年代较早的文献收录不全、检索方式少,检索结果有重复等,因此还需对其引文检索功能进一步完备。

总之,WOS 和 GS 的引文检索功能各有优势,GS 将成为 WOS 引文检索的一种有效补充。

参考文献:

- [1] 张学福.学术文献新的检索工具——Google Scholar[J].图书情报工作动态,2005,(2):7-9.
- [2] <http://scholar.google.com/intl/zh-CN/scholar/about.html>. [EB/OL].
- [3] Noruzi, A. Google Scholar: The New Generation of Citation Indexes[J]. LIBRI, 2005, 55(4): 170-180.
- [4] 樊怡菁. SCIE 和 Scopus 引文功能的比较分析[J]. 现代情报, 2006,(3): 80-82.
- [5][6] 夏旭. 基于 Google 学术搜索的引文检索研究[J]. 情报理论与实践, 2006, 29(6): 697-701.

作者简介:耿海英(1980-),女,中国科学院国家科学图书馆兰州分馆在读硕士研究生,研究方向为文献计量学理论及应用;肖仙桃(1965-),女,研究馆员,硕士生导师,研究方向为文献计量、科研评价等。