

基于文献计量的国外信息历史研究现状分析

Bibliometrics - Based Analysis of Status Quo of Information History Research

刘志辉

(中国科学院国家科学图书馆 北京 100190;中国科学院研究生院 北京 100049)

摘要 通过文献计量的方法分析了信息历史的研究现状。从相关文献的数量、关键词的词频与共现网络揭示了目前信息历史的研究内容,通过著者统计找出了核心作者,对目前的研究现状进行了分析。

关键词 信息历史 文献计量 研究现状

中图分类号 G201 G359.29

信息历史(information history)是一个新的研究领域,国内对此研究尚未有系统的分析和介绍。作为一个重要的新兴领域,2006年Black A.在美国的信息科学与技术年度述评中对信息历史(Information History)进行了介绍。本文希望通过对外国相关文献的定量分析,以揭示国外信息历史的研究现状。

1 信息历史的定义

信息历史是一个新兴的研究领域,目前尚处于初始阶段。美国的图书馆学家Alistair Black(2006)认为信息历史可以看作是书籍史、图书馆史、信息管理史等不同研究领域的上位词(umbrella term)^[1],而Toni Weller(2007)则认为信息史是一个对信息进行历史研究的独特形式,关注信息在过去社会中的作用,以历史证据为基础在历史研究的框架下进行分析^[2]。从两个定义提出的时间来看,Toni Weller的定义是在Alistair Black定义的基础上提出的,而且也是一种概括性定义。这两种定义存在着明显的不同,但是作为对一个新兴研究领域的认识却是可接受的。正如人们对信息无法给出广泛接受的定义一样,人们对信息历史的理解也是多样的。

2 数据的来源与研究方法

以“information history”为关键词,利用web of knowledge进行题名检索,经过人工清洗的结果显示最近两年只有三篇文献以此为题对此领域进行探究和梳理^[3~5]。这是因为早期的信息历史研究中采用了不同的词语,如“historical information science”、“historical infor-

matics”等,加之信息历史又是一个新的研究领域,所以采用传统关键词检索的方式查找相关文献有一定的困难。因此本文以这三篇文献为起点,利用论文引用的特点,通过滚雪球抽样的方法,利用参考文献查找相关文献,通过去重处理,最后建立了一个包括580条记录的数据集。

本文将通过对相关文献的年代分布、作者和关键词统计,分析目前信息历史研究的发展现状,通过同被引文献的分析,进一步揭示目前的研究内容。希望通过系统分析充分展示信息历史的研究现状。

3 信息历史研究现状分析

3.1 文献数量的时间分布 因为本文分析的数据集的来源文献是2006~2007年发表的(见图1),所以2005年以后的数据不具有统计意义。通过对文献数量的时间分布来看,相关文献的数量是趋于增加的,但是在2000~2003年却有下降的趋势。这种下降的原因可能有三:第一,从统计学上来看,可能是因为统计

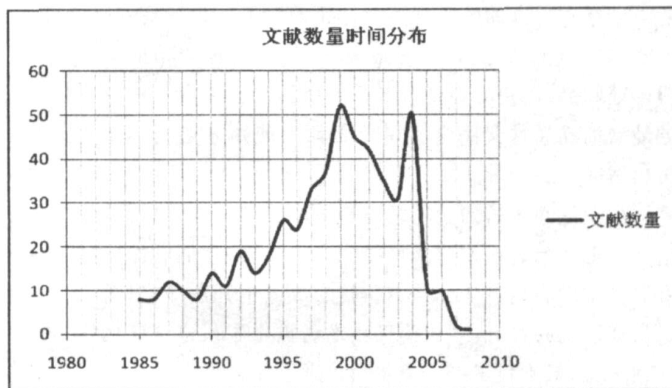


图1 文献数据时间分布

作者简介:刘志辉,男,1979年生,博士研究生,研究方向为学科情报与战略情报。

数据不足造成的,即本文在进行抽样统计时的样本量不够;第二,因为数据集来源于文后的参考文献,所以这种下降也可能是因为引用的规律所造成的;另外一种可能的原因是学界对此主题关注度的下降。2004 年文献数量的增加在一定程度上降低了最后一种原因的可能性。从整体数据来看,学术界对信息历史这一领域的关注还是呈增长趋势的。

3.2 著者分析 发文量在 5 篇以上的作者如表 1 所示。Black A. 以 23 篇的数量名列榜首,他是美国著名的图书馆学家,这说明了图书馆学界对信息历史研究的关注。Hahn T B 的文献总量排名靠前,但是其以第一作者身份发表的却较低。通过对其发表文献的进一步分析,发现他除了第一作者身份表文献外,其它均为会议论文或期刊的专题,其中会议名称是“Conference on the History and Heritage of Information Science”。这说明除了有个别研究者关注外,也有相关的会议对信息历史这一领域进行探讨。在这些发文量较高的作者中另一值得关注的作者是 Weller T.。尽管他的发文量不高,仅有 5 篇,但是从其文献的时间分布上来看,是最近三年(2005~2008)发表的,而且他的博士论文的主题也是关于信息历史的。

表 1 发文量较高的作者

作者	文献数量	第一作者	作者	文献数量	第一作者
Black A	23	20	Davis D G Jr	6	6
Rayward W B	13	10	Keeling D F	6	6
Buckland M K	12	11	Cronin B	5	4
Williams R V	11	9	Hjorland B	5	4
Hahn T B	8	3	Saracevic T	5	4
Bowden M E	6	3	Weller T	5	5
Campbell - Kelly M	6	6			

科研人员的合作关系在文献中表现为合著现象,作者的合著现象可以通过作者共现网络来表现。本文首先统计出数据集中的作者列表,构建出了作者共现矩阵,并利用 ucinet 进行可视化分析。图 2 是发文量较高的作者的合著网络图。从图 2 来看,目前关于信息历史的研究尚未有广泛的合作关系,主要是以小团队的形式进行,这是与信息历史研究仍处于起步阶段的特点是一致的。最大的合作网络(如图 3 所示)是 Williams R V 和 Bowden M E 所在的网络,两位作者不仅有高的度中心性(degree),分别为 4 和 3,而且也具有较高的中介中心性(betweenness),均为 10。Williams R V 发表了一系列的信息科学史书目,Bowden M E 同 Hahn T B 一样也是会议论文的编辑者。

3.3 关键词统计 因为本文是以参考文献为数据来源的,所以对关键词进行抽取,具体的方法是:首先统计文献标题的高频词,经过人工筛选确定关键词列表,并统计出词频较高的关键词。部分高频关键词如表 2 所示。

词如表 2 所示。

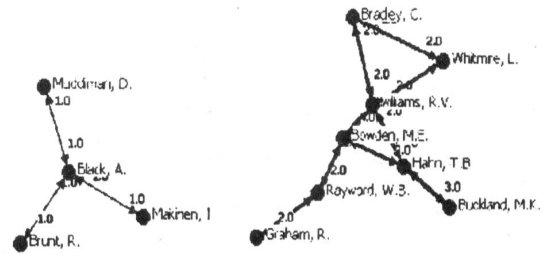


图 2 部分作者的合著网络

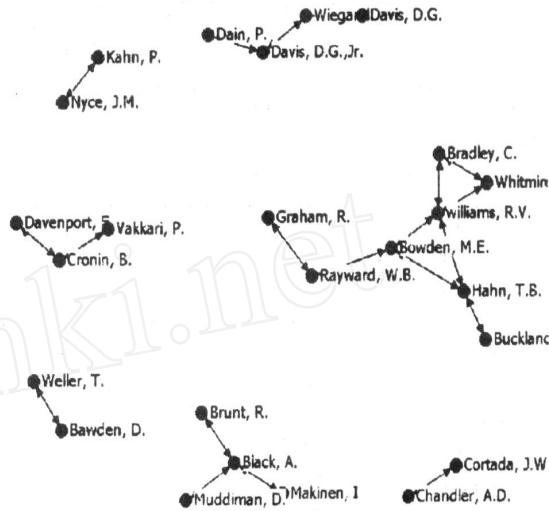


图 3 合著网络网

表 2 高频关键词列表

关键词	词频	关键词	词频
library history	37	information society	16
USA	24	america	14
history of information science	24	internet	11
britain	22	bibliography	10
information science	17	information history	10

根据上表的内容,可以看出:图书馆史、信息学史、信息科学、信息社会是信息历史关注的重要领域,而美国和英国则是重要的研究区域。从关键词的分布中也可以看出信息历史的学科渊源,即图书馆历史、信息学史,以及信息历史的发展趋势,即关注信息社会这个更广泛的主题。下文将从部分关键词的时间分布以及对文献标题的词频分析,进一步说明这一结论。

图 4 为部分高频关键词的时间分布。虽然图 3 未能呈现出连续曲线,但是还是可以从相关关键词的发展中看出信息历史这一领域的发展背景。图书馆史(LB Y HIS)一直受学界的关注。到上世纪 80、90 年代信息学(INFO SCI)、信息学历史(HIS of INFO SCI)开始更多地出现在学界的相关研究成果中。到上纪 90 年代中后期信息社会(INFO SOCIETY)的研究也开始引起了人们的关注。在这种背景下信息历史(INFO HIS)的研究开始出现。利用信息历史研究,人们想从长期的历史趋势中发现信息本身的发展趋势以及信息与社会相互关系。

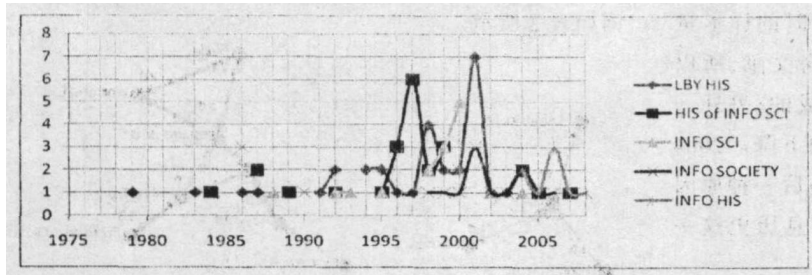


图 4 部分关键词的时间分布

3.4 文献标题分析 本文对文献标题的词频进行了统计,部分高频词如表 3 所示。从表 3 中除了可以发现高频关键词中的单词外,还可以发现一些其它内容。

表 3 文献标题的词频统计(部分)

单词	词频	单词	词频
information	240	historical	21
history	169	documentation	20
science	102	communication	19
library	77	scientific	19
society	38	british	18
libraries	28	world	18
century	27	technology	17
american	26	business	16
knowledge	23	internet	16
modern	23	public	16
states	16	origins	11
systems	16	power	11
development	14	special	11
research	14	story	11
america	13	system	11
britain	13	technological	11
bibliography	12	victorian	11
computer	12	classification	10
revolution	12	culture	10
future	11	management	10

除了图书馆学、信息学外,文献学(documentation)也是在信息历史研究相关文献中出现比较多的单词。从“century”(27)一词中可以看出信息历史研究中多是以某一段时间为研究对象,而“Victorian”(11)一词恰恰能说明对英国维多利亚时期的研究是典型的研究对象,如对这个时期公共图书馆和公共信息管理的研究^[6-7]。通过对文献标题的浏览可以发现标题中多会出现某一时间段之类的时间限定,如“1794~2000”、“Postwar”、“modern”等。在信息历史的研究中另一关注的内容是交流(communication,19),特别是网络(internet,16)和计算机(computer,12)对此的影响。为进一步分析、说明这些研究现状,本文对 2006-2007 年发表的关于信息历史的三篇重要文献同被引文献的关键词进行了分析。

3.5 同被引文献关键词分析 数据集中的文献多是专著,约占数据集中文献总数的 60%,期刊文章占

34%。在 2006~2007 年文献的同被引文献中专著和期刊所占的比例也与这个比例基本相同,分别为 60%和 37%。因此本文对关键词进行处理采用如下方法:对于期刊论文,采用论文中所提供的关键词;如果为专著,则要检索出与之相关的书评,利用书评中来抽取关键词。如果最终未能检索到相关书评,则利用进行关键词统计时所抽取出的关键词。最后利用得到的关键词列表生成关键词共现矩阵并利用 Ucinet 进行分析。根据关键词的共现情况生成的图如图 4 所示,其中点的大小为度中心性(degree),而不同灰度代表不同的中介中心性(betweenness)。为了表述方便,本文将由图 5 中不同的簇称为领域。

从图 5 中可以发现,目前信息历史的研究可以分为几个主要的领域,以文献学(Documentation)为中心的领域、以政府和信息社会为中心的领域、以科技信息为中心的领域和以信息历史为中心的领域。

以文献学为中心的领域最大,包括图书馆史、信息学史、出版史等传统的研究领域以及通信/交流(communication)等新的研究领域。这个领域甚至包括了医学信息学(medical informatics)和化学信息学(chemical information science)。研究的对象不仅是各种理论(social epistemology, theoretical topic)、技术方法(technique and tool, historical study),而且也包括人物(pioneers, Paul Otlet, Margaret Egan, Jesse Shera)和机构(history of American Society of information science)。

以政府和信息社会为中心的领域中主要关注的是信息对政府的影响(epistemic communities, expert movement, history of computer)、政府的信息政策(information strategy, surveillance)以及信息社会(information society, information state, information revolution)。

以科技信息为中心的领域主要是从知识的社会史(social history of knowledge)角度研究科学信息交流问题,如科学信息交流系统(scientific information system)以及其中的信息检索问题(information retrieval)。虽然利用共现关系聚类时这个领域单独出现,但是通过科学信息(scientific information)与科学(science)以及通信/交流(communication)之间的语义关系不难看出。这个领域实际上是可以包括在以文献学为中心的研究领域之中。

以信息历史为中心的领域比较小,但是因为是最新出现的研究领域,所以是最值得关注的领域。通过这个领域内的关键词可以看出这个领域强调的是以历史学的方法去研究不同的对象(history of information

society、history of information disciplines、history of print and written culture),而且也注重不同对象的社会史(social history)以及社会基础设施(social infrastructure)和政治基础设施(political infrastructure)。

4 结 论

信息历史是一个新的研究领域,目前尚未有明确的定义。从对著者的分析中可以看出信息历史受到图书情报领域研究人员的关注,特别是图书馆学家,如 Black, A.。信息历史的研究多是限定于对某一时期的研究,特别是对现代社会的研究。在信息历史的研究中多是文献学、图书馆学、信息学领域的研究,而且越

来越关注信息的社会史方面的研究,如对信息国家(information state)的研究^[8]。此外,随着信息历史研究不断的发展,信息历史的研究开始注重历史学的方法。但是通过前文分析也可以看出,从理论基础来看,信息历史缺乏统一的理论的框架,从具体的研究来看,研究的时期主要集中在现代。

虽然目前信息历史研究存在着不足,但是它为图书情报领域的研究提供了新的视角,而且对信息在不同历史时期、不同社会方面作用的研究,也为其它研究领域提供了新的方法。因此,信息历史作为一个新的研究领域很值得学界进一步关注。

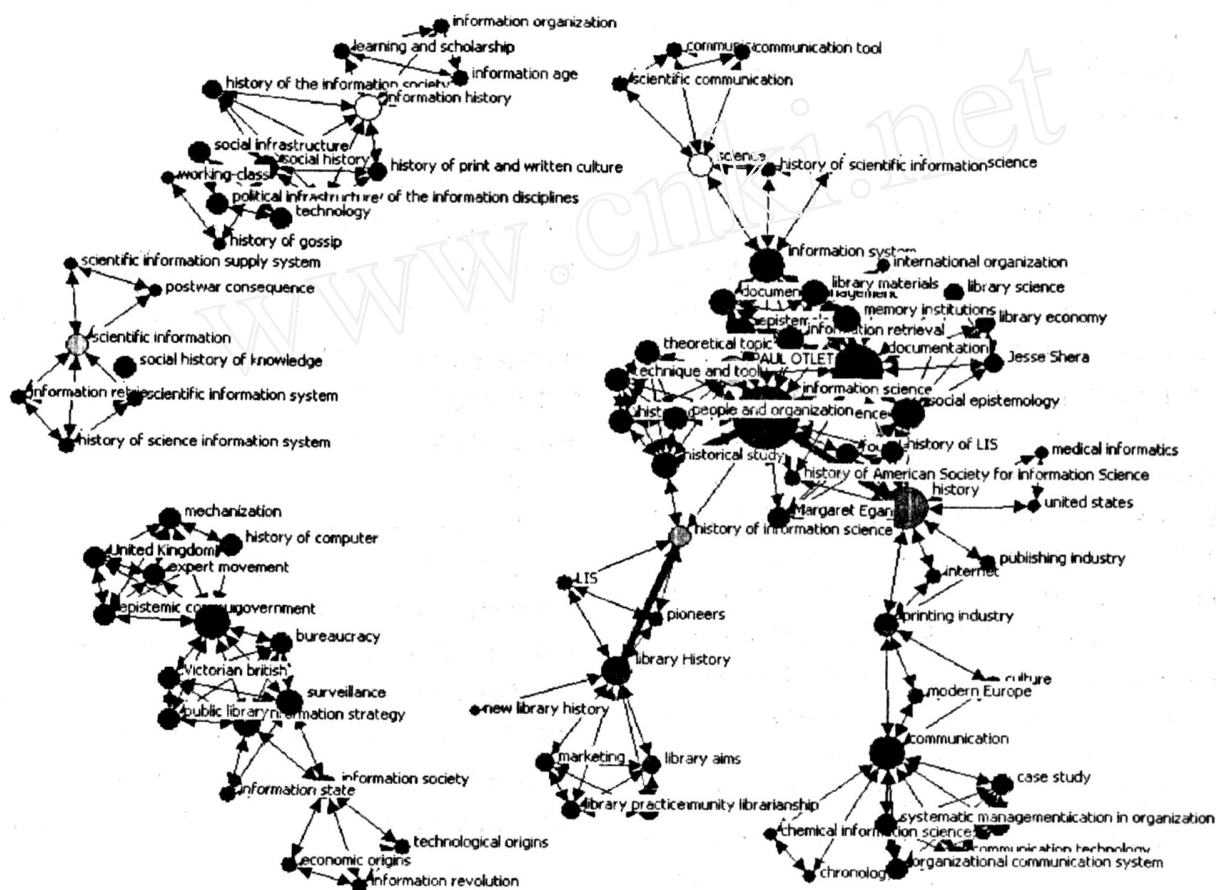


图 5 同被引文献的关键词共现关系图

参 考 文 献

1 Alistair Black. Information History[J]. Annual Review of Information Science and Technology,2006,40:441 - 473

2 Toni Weller. Information History: Its Importance, Relevance and Future[J]. Aslib Proceedings: New Information Perspectives,2007,59(4/5):437 - 448

3 Colin Burke. History of Information Science[J]. Annual Review of Information Science and Technology,2007,41:3 - 53

4 Alistair Black. Lost Worlds of Culture: Victorian Libraries, Library

History and Prospects for a History of Information[J]. Journal of Victorian Culture,1997,2(1):124 - 141

5 Alistair Black. The Victorian Information Society: Surveillance, Bureaucracy, and Public Librarianship in 19th - century Britain[J]. The Information Society,2001,17(1):63 - 80

6 Higgs E. The Information State in England: the Central Collection of Information on Citizens Since 1500[M]. Basingstoke: Palgrave Macmillan,2004

(责编:王平军)