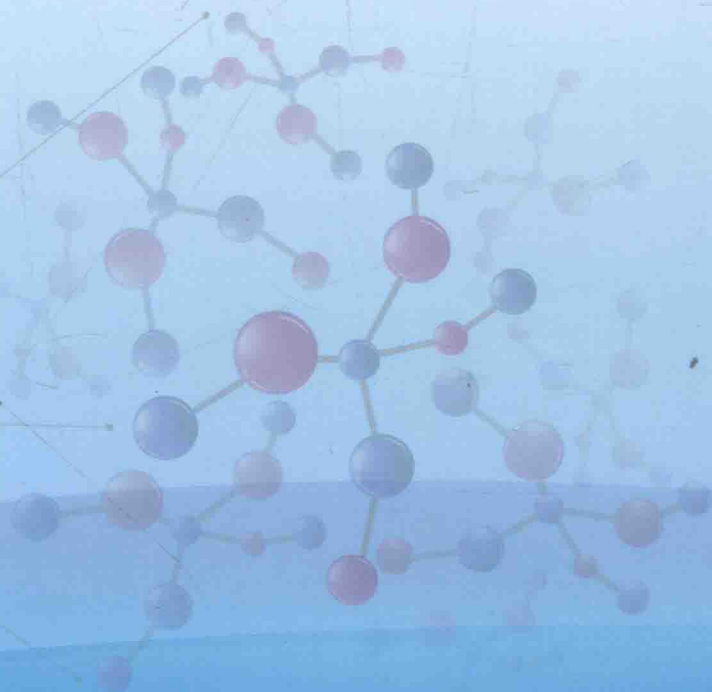


“十二五”国家重点图书出版规划项目

● — 中医药信息学丛书 — ●

中医药情报学

赵英凯 主编



科学出版社

“十二五”国家重点图书出版规划项目

中医药信息学丛书

中医药情报学

赵英凯 主编

科学出版社

北 京

内 容 简 介

本书属于“中医药信息学丛书”之一,是科技情报学的一个组成部分。主要进行中医药相关的情报研究和情报服务。本书由七个章节组成,其中“中医药情报学概述”、“中医药情报研究方法”、“中医药科技查新”三章分别介绍了中医药情报研究的概况、常用方法和手段,以及面向中医药行业的检索、查新。“面向临床的情报服务和循证医学最佳证据”一章探讨了基于循证医学的中医药情报服务和最佳证据研究。“国外针灸发展情报学研究”介绍了近年来国外针灸临床实践和临床试验研究现状等方面的情报研究成果。“中药专利研究”一章则重点讨论了中药专利的检索和分析方法。“外文信息的获取与翻译技巧实战篇”则以实例为主介绍了中医药相关外文信息的获取和翻译技巧。

本书对中医药情报学的研究者和使用者具有指导性,也可作为中医药信息化领域科研人员的参考书。

图书在版编目(CIP)数据

中医药情报学 / 赵英凯主编. —北京: 科学出版社, 2016

(中医药信息学丛书)

“十二五”国家重点图书出版规划项目

ISBN 978-7-03-048859-6

I. ①中… II. ①赵… III. ①中国医药学—情报学 IV. ①R2-03

中国版本图书馆 CIP 数据核字 (2016) 第 134137 号

责任编辑: 刘 亚 曹丽英 / 责任校对: 郑金红

责任印制: 徐晓晨 / 封面设计: 陈 敬

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

北京京华虎彩印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2016 年 7 月第 一 版 开本: 787×1092 1/16

2016 年 7 月第一次印刷 印张: 8 1/8

字数: 190 000

定价: 39.80 元

(如有印装质量问题, 我社负责调换)

丛书编委会

主 编 崔 蒙 吴朝晖 乔延江

编 委 王映辉 李海燕 张华敏 赵英凯

李园白 王 耘 姜晓红

《中医药情报学》编委会

主 编 赵英凯

编 委 王俊文 何 巍 童元元 杨坤杰

李 霞 刘 扬 徐 俊

丛 书 序

21 世纪是世界科学技术迅猛发展的时期,学科之间的交叉融合成为科技发展的重要趋势之一。其中,信息科学技术产生了广泛而深远的影响,对于医学领域也不例外。医学信息学是医学、计算机科学、人工智能、决策学、统计学和信息管理学的新兴交叉学科,在电子病历、医院信息系统、临床决策支持系统、远程医疗及数据交换标准等方面取得了丰硕的成果,已经在医院管理、教学和科研,疾病的预防、诊断和治疗等方面发挥了不可替代的作用。不言而喻,中医药信息学的发展历程更为年轻,富有潜力。中医中药流传数千年,至今仍然保持旺盛的生命力,在维护生命健康中发挥着独特而重要的作用。纵观中医药发展历程,总是与时代紧密相连,唯其如此,方能历久弥新。当今,现代科技背景之下,中医药学术繁荣复兴,与现代医学乃至其他学科的汇聚、交流、融合、互补,逐渐成为中医药时代发展的显著态势。

中医药文献典籍浩如烟海,学术经验传承异彩纷呈,蕴藏着极为宝贵的学术资源,有待深入发掘。信息科学技术方法为此提供了崭新的机遇,对中医药学术的当代传承与发展发挥了重要的作用,中医药信息学这门新兴的学科也由此应运而生。同时,也当应看到,缘于学科性质、理论钩沉、社会文化背景、语言表述、思维模式、时代变迁等差异,中医药学术内容本身与信息科学技术的融合过程中必然存在重大挑战,中医药信息的获取、转化与共享等面临许多困难。这一点是医学信息学、地理信息学等其他与信息学交叉的学科发展过程中较少遇到的。所以尽管呈现出蓬勃的生机与巨大的潜力,但至今尚少有学者,也无专著对其内涵、外延进行详细论述。虽然已经成为国家中医药管理局重点建设学科,但其具体的学科建设仍是筚路蓝缕,充满艰辛,亟需奠基性著作充实其理论内核,支撑后备学术人才的教育培养。幸而,以崔蒙研究员等为首的学术团队,多年来致力于中医药信息学原理与方法学的研究、中医药信息数据库及中医药信息国际标准的研制,其进行了大量基础性的研究工作,积累了较丰富的经验和学识,很多工作与研究都充实了学科领域,为中医药信息学学科的设置、建设与发展提供了极其坚实的基础和有益的借鉴。

对于一门学科而言,理论与实践工作同等重要。相比中医药信息研究工作的大量开展,学科理论建设工作有所滞后,长期势必会影响与制约学科发展。由此,《中医药信息学》编撰工作的意义与价值显得极为关键。该书从全方位的角度介绍了这门学科的去、现况和未来,对中医药信息的内涵、外延、研究方法、内容及意义等着墨甚多,阐发明晰而深刻,对中医药信息学下中医药信息标准、中医药科学数据、中医药知识服务、中药信息学、中医临床信息学、中医药图书馆学和中医药情报学七个分支学科均有系统论述。概言之,其研究内容几乎涵盖了一切与中医药活动有关的信息,如临床、科研、教育、管理、文化、生产经营等领域所产生的信息,提高了对中医药信息获取、转化、传播与利用的能力。

尤其值得一提的是,书中认为中医药信息是认识论层次的信息,具有现代整体性、动态时空性、现象理论等特征,其“主客融合的体验”及“包含本质的现象”等导致了辨证诊断和疗效的模糊,以及相对重视客体的整体变化状态,这些特点与大数据的“整体性”、“混溶性”、“相关性”三大特点不谋而合。如果能够借助大数据研究所获得的成果,从理论、方法

学上解决体验信息获取、存储及传播的问题，必将对中医药学发展起到至关重要的推动作用。

目前，欧美发达国家对医学信息学的教育与训练非常重视，认为掌握必要的现代信息技术是医务工作者必须具备的一项基础知识和基本技能。这一点在中医药领域同样适用，但纵观国内临床医疗系统尤其是中医药领域，对此认识还尚待深化，这对拓展中医药工作者的视野、提升其临床水平及科研能力显然不利。我希望《中医药信息学》的问世能够在较大程度上引发学界对此问题的关注与重视，推动中医药信息学术的普及与发展，获得更大范围的学界共识。

相比传承千年、博大精深的中医药学，中医药信息学刚刚起步，尚有很多的工作需要一一完成，还有很多的困难需要一一克服，可谓前路漫长且艰、任重而道远。可喜的是，《中医药信息学》的编撰为万里征程开了一个好头，为这门学科的发展奠定了基础，指明了方向，确立了模式。“前人栽树，后人乘凉”，希望广大中医药信息工作者以此为起点，在全面而深刻把握中医药学术特质与发展规律的基础上，有效借鉴、运用信息科学原理、方法、技术，不断丰富中医药信息学的内涵，探寻其内在规律，为中医药学术的传承、发展乃至创新提供更多的助益，充分发挥其独特作用。

传统与现代的交融总是令人充满无限的遐想与期待，处于高概念和大数据时代的中医药信息学更加深化其学科特质，望能引领中医药学科、事业与产业的发展。对于崔蒙、吴朝晖、乔延江主编及编写团队，我比较熟悉他们的工作，感佩学者们孜孜不倦、辛勤耕耘、认真治学的精神，创建一个崭新的二级学科实在不易，此书乃中医药信息学的奠基之作。书濒脱稿邀我作序，是对我的信任和鼓励，谨志数语乐观厥成。

王永炎

甲午季秋

目 录

第一章 中医药情报学概述	1
第一节 中医药情报采集	2
第二节 中医情报的分析和评价	7
第三节 中医药情报服务	11
第四节 情报研究方法及其在中医药领域的应用	14
第五节 中医药发展战略情报研究	19
第六节 国际中医药及传统或补充替代医学情报研究	20
第七节 中医药竞争情报研究	20
第二章 中医药情报研究方法	24
第一节 中医药情报研究的概念与特点	24
第二节 中医药情报研究的一般流程	24
第三节 中医药情报研究方法的分类	25
第四节 中医药情报搜集方法	25
第五节 逻辑方法	26
第六节 中医药专家咨询方法	28
第七节 中医药文献计量分析方法	29
第八节 中医药知识发现方法	30
第九节 多元统计方法	31
第十节 中医药情报研究重点领域研究方法	32
第三章 中医药科技查新	37
第一节 科技查新总论	37
第二节 科技查新流程	44
第三节 查新报告的内容	47
第四章 面向临床的情报服务和循证医学最佳证据	55
第一节 相关概念	55
第二节 中医药临床情报服务的情报源	56
第三节 常见临床问题及循证医学临床证据的种类	58
第四节 循证医学系统评价与中医药文献研究方法	61
第五节 面向临床的其他中医药情报服务	65
第五章 国外针灸发展情报学研究	68
第一节 国外针灸的发展现状	68
第二节 国外针灸临床试验的情报学研究示例	68

第六章 中药专利研究	82
第一节 中药专利概述	82
第二节 专利数据库	84
第三节 检索策略	85
第四节 研究方法	96
第七章 外文信息的获取与翻译技巧实战篇	106

第一章 中医药情报学概述

中医药情报学是科技情报研究的一个组成部分,始于1959年中医研究院在学术秘书处成立的情报资料室,当时主要负责中医药情报的搜集、整理和交换等工作。

早期的中医药情报研究主要是基于对国内外相关信息的收集和整理,编写内部资料和专题报告,为科研人员和行业管理部门提供情报服务。例如,1960年公开出版了《中医文摘》(后改名《中国医学文摘——中医》);1973年创办了《中医药研究参考》,主要用于内部交流;1976年开始编印《中医药研究资料》,用于介绍中医药行业科学研究的新技术、新方法、新进展;同年还不定期地编印《中医药动态》,主要为卫生部、院所领导及有关科室提供参考;1978年创办了《国外医学参考资料》(后改为《国外医学·中医中药分册》),这是国内唯一报道国外中医药信息的正式刊物。

目前的中医药情报研究主要包括战略研究、政策研究、专题情报研究和企业情报研究等。

战略研究主要针对中医药在国内外科、教育、医疗、市场等方面的发展现状;国际社会对中医药的态度和需求情况,结合社会政治、经济、教育和管理,从总体出发,在宏观与微观、理论与应用方面进行深入系统的研究,调研相关情报分析和预测中医药面临的各种机遇和挑战,并提出相应对策。它主要是为各级政府卫生部门进行中医药发展战略研究及专题调研等信息分析活动、充分发挥综合情报的社会功能。如人口与医药卫生发展战略与技术经济政策研究、国家中长期科技发展战略研究、中医现代化发展战略研究、中药现代化产业推进战略研究、中药现代化发展战略研究、促进中医药出口创汇的战略与政策研究等。

政策研究主要围绕国家卫生部、中医药管理局及相关政府部门、中医药临床和科研部门的中心工作开展相关的情报分析研究,从错综复杂、内涵丰富的海量数据中提取出有用的信息,研究和运用中医药情报分析方法,处理和解析这些信息,为政府部门现阶段政策的实施情况和下一阶段政策的研究与制定提供决策依据。如中医中药在西部大开发中的地位与作用研究、十省市中医医疗需求与服务调查、中医中药在中国医疗卫生保健中的作用与应用情况、全国民族医药基本情况调研、国外中医药立法状况研究、全国农村中医药基本情况数据统计及研究等。

专题情报研究主要为中医科研单位、高校及临床机构等提供科研课题相关的课题查新、循证服务、技术咨询、决策咨询等服务,同时,随着各国对中医药的科研、临床运用日益增多,掌握国外有关的科研方向,科研动态及成果,教育、医疗、市场的最新信息,进行综合分析,对我国中医药科研发展具有一定促进作用。如关于甲型H1N1流感等突发传染性疾病的中医药防治情报研究、中医临床特色优势研究、传统医学纳入各国国家医疗保健体系情况分析、美国政府关注的与中医学相关的传统医学问题、俄罗斯传统医药信息调研、国内外戒毒研究进展等。

企业情报研究主要为医药企业提供市场相关情报的调研、预测和分析工作。如银杏提取物信息调研、冬虫夏草信息调研、治疗糖尿病中药单味药及中药化学成分信息调研等。中医药企业情报研究属于竞争情报研究的范畴。

中医药行业具有传承与创新并存、行业专属性强等特点,对情报学方法的要求更为复杂和

迫切。到目前为止,情报学方法在中医药领域的应用已有一定发展,相关文献量逐年递增,采用的方法也从单一的数理统计方法和文献计量学方法逐步扩展到数据挖掘、知识发现及竞争情报方法。对于情报方法学的研究,是一个理论指导实践,实践验证、完善理论的循环过程。对于中医药情报研究合适方法的研究,引进借鉴别的学科行之有效的方法和技术是必要的,应全面了解和掌握所选用研究方法的优缺点及适用范围,熟知研究程序和所需要的约束条件,注重研究中医药在现代发展中遇到的各种实际问题,结合情报学特点、中医药学学科特点,把握未来情报学学科的发展趋势。

在互联网没有出现之前,情报科学的研究对象主要是文献,情报学家对科学的关注几乎全部都是通过文献的研究来实现的。而在互联网普及之后,情报研究不仅要面对“信息爆炸”所带来的信息积累与查找利用之间的矛盾,而且情报的来源和研究对象也从单一的文献扩展到数据、图像、语音等多种形式和内容。一般认为,情报学的这种从文献层次向知识层次的深化、演进与发展,是情报学研究的知识化趋势。因此,现代情报研究工作已不再仅仅是为决策者、科研工作者提供准确的信息服务,而且要善于从专业的角度,以敏锐的判断力及时捕捉、展现和评价相关的科研活动和成果,以探索或发现知识的发展规律,中医药情报研究亦是如此。

第一节 中医药情报采集

中医药情报采集是中医药情报工作的第一个环节。情报采集就是根据用户的需求、机构的性质和任务,用科学的方法收集、检索和获取特定情报资源的过程。情报资源是开展情报研究和情报服务的物质基础,其质量的优劣将直接影响情报服务的整体效果。在情报来源极为丰富、情报流通渠道相对畅通的今天,把好情报资源采集的质量关是提高情报研究和服务水平的关键。

不论何种类型的情报研究和情报服务活动,都必须有充分的情报资源。为了提高情报搜集的效果,情报人员应遵循全面性、系统性、针对性、新颖性、可靠性、科学性、计划性等原则。情报搜集一般包括文献调查和社会调查两种途径。

(一) 主要的中医药情报源

中医药情报采集的工作对象是情报源,它是中医药情报工作的物质基础。所谓情报源指的是一切产生和持有情报的个人和机构,或者负载情报的物体。按照情报存在的物质形态,情报可划分为文献形式、语言形式和实物形式。这三种形式主要分为文献型和非文献型两大类,语言形式和实物形式则属于非文献情报源。根据中医药情报的特点分为古代中医药文献、现代中医药文献及网络中医药情报资源等几个部分,下面我们逐一介绍。

1. 古代中医药文献

古代中医药文献的主要载体是图书,大体可分为四类:一是经书,专著和对经书、专著进行注释析义的书籍;二是据经书、专著生成的临床各科文献;三是史志中所载的医药记事;四是散见于历代诸子著作中的单篇论文。古代中医药文献应具备三个要素:一是有具体的图书史料;二是具有使用价值、实践价值;三是具有历史意义、历史地位,并可为当今所利用。

根据中国中医科学院图书馆主编,包括北京图书馆、中国医学科学院图书馆、各中医高校

图书馆等国内 113 个图书馆参与的《全国中医图书联合目录》的图书分类方法,现存中医图书采用三级式分类。其中,在第一级方面共分为 12 类,分别为医经、基础理论、伤寒金匱、诊法、针灸按摩、本草、方书、临床各科、养生、医案(医话、医论)、医史及综合性著作。

2. 现代中医药文献

(1) 按文献的性质、特点和编辑出版形式分类:文献型情报源是科技情报的主要存在方式。文献型情报源主要包括印刷型文献、缩微型文献、机读型文献和声像性文献四类。根据文献的性质、特点和编辑出版形式的不同,中医药现代文献可分为:图书、期刊、报纸、会议文献、专利文献、学位论文、标准文献、政府出版物、产品样本、病案资料等。

1) 图书:目前尚无严格和统一的定义,但一般认为,图书是一种用印刷或手抄方式把原文或图表加以再现,并装订成册,而篇幅大于 48 页的不定期出版的文献。

2) 期刊:是一种定期或不定期的连续出版物,出版周期最长不超过 1 年,且长期使用固定和统一的刊名,版面、开本、篇幅、栏目、价格均基本稳定,每期至少刊载两篇以上稿件,且内容不能重复。根据内容、性质的不同,期刊又可分为:学术性刊物、技术性刊物、快报性刊物、资料性刊物、译文性刊物、检索性刊物等。

3) 报纸:是一种特殊的期刊。在中医药信息的传播上,具有内容性、报道性、发行量大、流通面广的特点。它是了解中医药研究动态、及时获取科研信息的重要途径。目前,公开出版的医药卫生类报纸约有 20 余种。

4) 会议文献:是指在学术会议上宣读或书面交流的论文或报告,根据会议文献的出版时间,可划分为会前出版物及会后出版物。根据会议的地域范围不同,可分为世界性、国际性、全国性和地方发行几种。

5) 专利文献:是指专利审批过程中产生的官方文件及其相关出版物总称,包括专利说明书、专利公报、专利分类、分类表索引等。

6) 学位论文:是指高等院校、科研机构的研究生、毕业生为取得某种学位资格而撰写的学术论文。

7) 标准文献:又称技术文献标准,是由权威机构批准的有关标准化的技术规定,以文件形式出现,是技术标准、技术规格和技术规则等文献的总称。标准文献有一定的法律约束力,按其使用范围可分为国际标准、国家标准、区域标准、专业标准。

8) 政府出版物:是指各国政府及其所属机构发表的行政文件或科技文献,其内容十分广泛,按性质可分为行政性文件和科技文献两大类。

9) 产品样本:是指对定型产品的性能、构造、用途、规格和使用方法的说明性文字和宣传资料,包括产品说明书、产品目录和贸易性刊物。

10) 病案资料:为临床工作的原始记录,是进行临床总结的最可靠依据,也是医学教育和科研工作的重要资料。病案资料一般不出版,有一定保密性。

(2) 按文献的加工程度分类:如果按照文献的加工程度,还可将文献划分一次文献、二次文献、三次文献等。

1) 一次文献:是科研人员依据自己的科研实践撰写的,直接阐述科学发现、发明、成果或学术观点的文献。根据文献的内容性质,只要是作者根据自己的科研成果而发表的原始创作,无论是手稿、译文、论文、图书;也无论是铅印,还是复印,都是一次文献。

2) 二次文献:是对一次文献进行加工、提炼、压缩,按一定的系统结构和组织方式编辑成的工具性文献。二次文献的主要类型有目录、题录、文摘、索引等,形式上主要分为卡片

式、书本式、期刊式等。

3) 三次文献:是指利用二次文献,对一定范围或专题的一次文献的学术内容进行收集整理和综合分析基础上编写的文献。它主要包括两类:一类是指对某些具体研究课题的当前状况和发展趋向进行分析和评述,如论文综述、专题评论等;一类是将发展较为成熟的知识系统化,以便人们学习、查找和利用,如教科书、词典、百科全书、年鉴、手册和指南等。

此外,还有四次文献,又称机读文献,是指利用计算机磁带、磁盘编制的书目索引文献;零次文献,一般是指未公开于社会的非出版型文献,如书信、笔记、手稿、发言稿、会议记录、实验记录,以及通过口头交流或实际传授等形式进行传统的,直接作用于人的感官的非文献型情报信息。这些文献构成了特殊形式的情报源,具有重要的情报价值。

3. 网络中医药情报资源

近年来,传统医学受到越来越多的关注,人们对包括中医药信息在内的传统医学信息的需求也日益增多。为顺应这一潮流,互联网上出现了大量传统医学情报资源,包括中医药相关管理机构、中医药院校、研究机构、中医药企业和公司等建立的站点。国外的一些著名的生物医学检索系统也都收录了我国中医药期刊等相关文献。网络中医药情报资源主要包括与中医中药、传统医学及补充替代医学相关的信息网站或网页,如国家中医药管理局,印度卫生部阿育吠陀及自然疗法、尤纳尼、锡达、顺势疗法部 (AYUSH),美国国立补充替代医学中心 (NCCAM),以及世界卫生组织 (WHO) 网站中涉及“传统与补充医学”的内容等。国内外主要的、包含中医中药及传统或补充替代医学相关内容的文献数据库,如中国知网、万方数据、PubMed、SCI 等也已成为网络中医药情报的重要资源。具体可分为以下几类:

(1) 搜索引擎或网络资源导航:搜索引擎是一种查询网络资源的工具,网络上有许多优秀的搜索引擎及资源导航,目的是帮助用户快速、准确查找网络上的情报资源。如 Google、百度、Yahoo、Medscape Medical、Matrix Doctor's Guide、中国医学生物信息网、国家科技图书文献中心医学图书馆站点导航等。

(2) 中医药相关的医学信息网站

1) 国内外医药管理部门网站:主要包括各国卫生部、药物管理部门、补充替代医学或中医药的专门管理机构等,如卫生部、国家药品食品监督管理局、国家中医药管理局等。

2) 国内外医药卫生权威组织、医院、教育机构及学术团体等网站:如美国国立医学图书馆、北京图书馆、中国医科院图书馆等 30 余家医学图书馆的网站,WHO,中华医学会,中华中医药学会等。

3) 国内外医学核心期刊网站:如《中华内科杂志》、《中华外科杂志》、《中医杂志》、《中药材》、《中草药》,以及 *Medicine*、*Nature*、*Lancet*、《新英格兰医学杂志》等。

(3) 生物医学数据库、中医药学数据库及医学相关检索系统:生物医学数据库按照文献类型划分,主要有文献型数据库、数值或事实型数据库、多媒体数据库等,这些数据库以商业和非商业服务方式提供。我国的中医药期刊文献已被国内外多个生物医学数据库所收录。

1) 文献型数据库:包括题录文摘数据库和全文数据库。题录文摘数据库最著名的是美国国立医学图书馆的 Pubmed 数据库,此外还有 TOXNET、CANLIT、美国专利数据库、中国期刊题录数据库、商用数据库 DIALOG 系统等。全文数据库目前主要有 ProQuest medical Library、OVID、Elsevier Science 公司的 SDOS 全文检索系统、清华同方、万方、维普等全文期刊数据库等。

2) 数值或事实型数据库:主要包括基因库、核酸序列、蛋白质结构库等分子生物学数据

库,以及毒理学、药物方面的事实型数据库。

3) 多媒体数据库:包括物质或药物三维立体结构数据库、各种医学图谱库、医学影像库、病理切片库等。

特别提出的是,中国中医药信息数据库系统,是由中国中医科学院中医药信息所研制的系列中医药数据库系统,其中包括综合查询各类数据库的多库融合平台、具有统计功能的结构型数据库、文献题录型数据库及事实型数据库。

目前,随着中医药在世界各国的影响力不断加大,很多补充替代医学及我国的中医药学期刊杂志已被多个全球权威性生物医学信息检索系统所收录。①美国《医学索引》: *Index Medicus*, 简称 *IM*, 是由美国国家医学图书馆 (National Library of Medicine, NLM) 编辑出版的题录式医学文献检索性期刊。有关针刺、经络、脉诊、方剂、药物的中医药文献报道,在 *IM* 中与日俱增。②美国《化学文摘》: *Chemical Abstracts*, 简称 *CA*, 由美国化学学会化学文摘服务社 (Chemical Abstracts Service of American Chemical Society, CAS) 编辑出版。*CA* 收录的文献虽然以化学化工为主,但也广泛收录了与化学有关的生物医学文献,约占其报道量的 35.5%。其摘录的 1000 种自然科学核心期刊中,包含 350 余种最重要的医学核心期刊。随着中医中药在世界范围的发展, *CA* 有关中医药的报道近年来明显上升,其中尤以针灸和药用植物方面的文献为著。③美国《生物学文摘》: *Biological Abstracts*, 简称 *BA*, 收录了世界上 100 多个国家, 23 种文字出版的 8500 余种生物、医学期刊所载论文及 1 万多种其他类型的文献。④荷兰《医学文摘》: *Excerpta Medica*, 全称 *Excerpta Medica Abstract Journals*, 简称 *EM*, 是世界上唯一的英文医学文摘。其中,生物医学期刊约 3000 种,其他医学相关期刊约 2400 余种,我国有 15 种医学期刊被收入。

(4) 生物医学及中医药学电子出版物:主要包括生物医学电子期刊、报纸、图书、手册、法规、指南、图谱、百科全书等。许多电子图书可以通过网络检索,一些搜索引擎和网站可以提供百科全书、医学词典及医学术语集。

(5) 特种文献:主要包括专利信息、标准、学位论文、科技报告、会议论文等信息资源,目前均可通过网络检索与获取,如中国专利信息网、国家科技成果网、中国医学会议网、中国标准服务网、美国专利数据库、日本专利数据库、美国政府报告数据库、Medicine Conferences (英国医学会议网)、国际标准化组织 (ISO) 等。

(二) 中医药情报采集的原则

不论是利用图书资料,还是通过二次文献;也不管是利用手工查找,还是通过计算机检索,都应本着先近后远、先内后外、先易后难、先简后繁、先本专业后其他专业的原则去寻找中医情报信息。

(1) 先近后远:即查找情报时要从最近期的资料开始追溯既往的资料,以便了解当代的水平与最新的理论观点及方法手段。另外,有些近期信息资料附有既往信息目录,可供选择和利用,这是扩充情报资料的一条重要途径。

(2) 先内后外:即先查找国内的有关信息,然后再查找国外的资料。就目前而言,祖国医学古典信息资料的绝大多数可以在国内找到,现代的中医药研究资料亦数国内居多,因此应当先搞清国内的情况。另外,国内资料易于理解,查阅速度快,如果能充分利用翻译资料,还要弥补掌握语种不多的缺陷。国内有些信息本身也引证了大量其他国家的有关资料及线索,为进一步查找国外信息扩大了来源。

(3) 先易后难:即查找情报时先从本单位、本地区的资料开始,这样比较容易。舍近求

远,即费时又费力。

(4) 先简后繁:即查找情报时先用较简单的方法,后用较繁复的方法。如查找某一专题信息应先查找有关的综述文献,然后再找单篇论文,以便迅速了解有关问题的历史和现状,以及存在的问题、争议与展望,有助于对所研究的问题产生较深刻而全面的认识。综述的正文之后往往列有许多信息目录,是扩大信息资料来源的极好途径。

(5) 先本专业后其他专业:有的信息资料不一定完全登载在本专业刊物之内,往往发表于其他专业或一些综合性期刊之中。因此,查找信息资料,除通过专业期刊外,还要关注其他相关学科和边缘学科的期刊。人们一般对专业资料都相当熟悉,查阅快捷,掌握准确,能较迅速地找到所需资料。在找到的资料中很可能就引证了相关学科期刊上的信息,为进一步广泛查找资料提供了线索。将找到的资料进行全面系统的分析,就更明确了还需要哪些学科领域的信息资料,以便有条不紊地去查找。

(三) 中医药情报采集的方法

(1) 针对研究主题的情报采集方法:针对某一研究主题查找有关的情报资料,通常采用系统检索法、追溯法及循环法。

1) 系统检索法:是利用检索工具全面、系统地查找某一专题信息资料的方法,即利用二次文献寻找一次文献的方法。由于二次文献多数是图书信息中心为了使读者能更充分地利用各种信息资料而编辑的,它是浓缩了的一次文献,因此其所登录的内容基本上是未经选择的,不受作者兴趣和需要的限制。一般而言,利用二次文献所找到的信息资料比较全面而少遗漏,缺点是信息报道比较慢。

2) 追溯法:是根据某一线索逐步深入,尽可能多地将某一问题的全部资料寻找出来。通常是从教科书、专著、文献综述或最新期刊论文等资料的参考文献表开始,将有问题的全部书目或题目逐一采集。然后根据这些题录提供的线索找到其他有关资料,通过阅读进行进一步的采集。长此以往,通过不断积累,内容也将逐渐充实起来。按照这种方法找到的信息,一般都是经过中间作者选择的,受其兴趣和需要的限制,因此不容易全面而无遗漏。

3) 循环法:系统检索法与追溯法交替进行检索即称循环法。一般根据检索题目的特点和信息检索工具的拥有情况,某一段时间利用检索工具进行检索,而另一段时间则用追溯法进行查询,这样交替循环,直至满足检索需求为止。

(2) 针对信息的类型和性质的情报采集方法:按照信息的时间特性可以将信息分为动态信息和静态信息;按照信息的保密特性可以将信息分为公开信息和非公开信息。依据信息的类型和性质不同,可采用不同的信息采集方法。

1) 对于动态的公开信息,采集方法包括问卷调查法、参考观察法、专家咨询法等;静态的公开信息,采集方法有预订采购法、信息检索法、日常积累法等。

2) 对于非公开的动态信息,采集方法有访问访谈法、技术截获法等;对于非公开的静态信息,采集方法有交换索要法、委托收买法等。

经过搜集而获取的原始情报通常是繁杂无序、真假混杂的,因此需要进行处理。情报处理包括整理和评价两个相辅相成、交替进行的环节。其中,情报整理包括形式整理和内容整理,目的是使情报从无序变为有序,成为便于利用的形式;情报评价则强调对整理出来的原始情报进行鉴别,一般依据可靠性、先进性、适用性等指标进行,目的是筛选出有用的情报,淘汰掉无用或不良情报。这两个环节共同作用的结果是使所搜集到的情报不仅是有序的,而且是有用的。

第二节 中医情报的分析和评价

计算机和网络技术的不断发展,使网络信息和各种数字化资源的流量呈几何级数增长,为人类提供了更快捷的信息获取渠道和更及时、全面的信息源。在海量信息当中如何评价情报的价值,如何运用信息处理技术及信息研究方法处理和解析这些信息,发现和认识隐含在数据海洋中的客观规律;如何利用这些数据,并将其转化为知识,指导现代中医药研究,是情报分析工作所要解决的焦点问题。

(一) 国内外情报分析和评价的现状

1. 国外现状

有学者对 1994 年以来国外 5 种情报学核心期刊所载的 2175 篇学术论文的主题进行了统计分析研究发现,情报加工处理这一部分论文的主题主要集中在信息描述、标引、知识发现与数据挖掘、分类、压缩、编目、自然语言处理、摘要、可视化、提炼、排序、信息组织、信息处理、索引、词汇控制和抽取等方面,占该类目论文总数的 85%。并且发现自 1998 年后,能够帮助医学研究者更好地发现不同专业领域文献中隐含关联的非相关文献知识发现法得到了重视,从各年论文分布的趋势上看有较大的持续增长,所研究的内容包括知识发现、信息发现、概念探索、文本挖掘、数据挖掘和网络挖掘等众多方面。

当今国外的情报机构非常重视各种情报的分析,即为各级用户提供相关情报分析结果、未来发展和帮助等咨询功能。一些国外知名的情报机构不断开发分析软件或系统,以提高情报分析功能和水平。

兰德公司是当今世界最负盛名的决策咨询机构,涉猎政治、军事、经济、科技、社会等各方面情报战略研究。兰德通过先进的计算技术、建模技术、分析工具和软件、分析方法来对热点和前沿问题进行分析,扫描、监视和跟踪全球发展动态,为美国州、联邦政府及资助人提供客观的高质量研究报告。目前,兰德已有一套完整的分析方法体系,主要包括:德尔非法、“跨学科”研究方法、统计分析法、大规模情境法、回溯法、规划法、成本效用分析法、处理不确定问题方法、度量研究效益方法、未来方法、评估法、逻辑模型、离散选择模型等。

美国能源部能源信息管理机构(the Energy Information Administration, EIA)用线性规划理论和情景模拟法建立了一个情报分析平台,用以揭示美国能源市场供给和需求之间重要的交互作用、预测 2030 年以前美国中期能源市场。

加拿大科技信息研究所(the Canada Institute for Scientific and Technical Information, CISTI)的研究流程:CISTI 在信息资源、数据库建设等工作的基础上,建立了一个专利分析平台,利用各种功能挖掘有价值的信息,包括文本挖掘分析、高级可视化、灵活的文本管理等,并建立了一个基于数字信息流的信息研究流程。利用信息分析模型,阐释当今科技医学信息环境下信息流高复杂、完全交互的特性。

2. 国内现状

20 世纪 50 年代,我国的信息情报分析活动才开始起步,当时主要是采用文献工作和研究工作相结合的方式,信息不够及时、全面,信息量也较少。80 年代后,我国大量吸收国外的新观念、新技术、新方法,情报分析工作开始迅速兴起,并从国家科研情报单位渗透到社会各

行各业。经过半个世纪的发展,从最初的专题研究阶段发展到综合研究阶段,形成了初具规模的信息分析产业。情报分析在科技信息界又称为情报研究、情报调研,多以科研课题形式呈现;在经济界又称为经济分析、经济预测、市场分析等,多表现为以竞争情报分析为主体的调研任务。

当前,我国信息分析活动的主体多元化,依据服务方式分为两大类:一是非营利性信息分析机构,主要是政府部门的信息研究机构、公益性事业单位,如情报所、图书馆、企业内部信息分析研究机构等;二是营利性信息分析机构,包括金融证券投资信息分析机构、市场调查与分析公司、信息咨询公司等。非营利性信息分析机构的各种信息研究所、信息分析中心是信息分析主体的中坚力量。当前,信息服务的对象以政府领导部门和企业为主,由于市场竞争化日益激烈,企业的竞争情报需求也随之增大。

尽管大量的网络信息处理系统及软件使得信息获取能力不断增强,但“提供分析与建议”却是除情报工作者之外任何技术都无法实现的。这是国内外情报工作不争的一个事实。但由于情报工作在人员素质、分析系统、社会环境、政策环境等各方面的发展都不够成熟,情报人员提供的分析和建议往往不能被企业高层决策者重视和采纳。因此,情报的分析和价值探讨仍将是学术研究中重点关注的焦点问题。

(二) 中医药情报分析

1. 中医药情报分析的定义

情报分析是对收集、整理后的情报进行精加工。通过中医情报专业人员的思维活动,采用科学的方法、手段和工具,对采集到的大量纷繁无序、杂乱无章的信息进行有针对性的选择、分析、综合、预测,以揭示、总结信息中有价值的部分及其本质规律,预测其未来发展状态,最终获得系统、准确、及时、大流量的知识、信息和预防性对策。

例如,在运用分析与综合的方法进行情报研究时,一方面借助专业人员的专业基础进行思维分析,把研究对象的整体分解成各个能反映整体特征的部分,抽取出必然的、本质的部分,并对其分别进行深入细致的考察;另一方面运用综合方法,根据研究目的,将分解出来的无序零散的各部分认识进行重新组合,研究并揭示蕴涵于其中的隐含信息和关联关系,达到重现整体、推断未知或预测未来的目的。

情报分析既与研究对象有关,也与研究目的或任务相关。研究对象和研究目的的广泛性和多样性,决定着情报分析工作的复杂性。

2. 中医药情报分析的目的和作用

(1) 发现知识孤岛,挖掘新知识:如空白点分析法,通过了解研究目标的现有知识存量,将现有知识存量与达到目标所需要的知识总量进行比较,从而找出带有研究性质的未知的特定参数,发现知识孤岛。非相关文献的知识发现方法较好地地把立足于文献外部特征的文献方法和立足于文献内容实质的知识方法有机地结合起来,为跨学科研究中的知识发现提供了有力的工具。

(2) 加强隐性知识的转化与交流:如德尔非法,通过函件进行专家意见的征询和反馈;头脑风暴法,通过面对面的交流进行情报传递。

(3) 实现数字挖掘和知识发现:如内容分析法,可以发现中医药学不同学科科研变化的动态和趋势。

(4) 揭示知识谱系:如知识基因法,通过对知识演变过程中具有遗传性质的概念、特征、

人物、主题等因素的总结与分析,较好地追踪概念的发展图式,探寻其演变历程,揭示知识谱系中各遗传因素之间的逻辑关系。

3. 中医药情报分析的技术和方法

下面我们主要介绍一些目前国内外在医学情报分析工作中常用的技术和方法。

(1) 网络监测技术:通过传统人工方式搜集网络信息,工作效率较低而且具有很大的随意性。利用联机服务系统跟踪和监测特定的目标网站,通过对网站信息的挖掘可获得有用的情报。

由于网络信息大部分以超文本的形式存储,通过超链接提供服务。因此,除了跟踪监测网页内容的变化外,还可以使用 Robot 程序,沿着网页中的链接自动漫游和下载,从而完成自动获取网络信息。获得相关信息以后,可以运用 WEB 文本分类技术对采集结果进行处理。即对采集到的非文本信息,对照“停用词表”去除虚词、介词等无具体含义的词后,对 WEB 文档进行词法分析和词条分割,如果是英文文档还需要进行词干抽取。通常采用向量空间模型(vector space model, VSM)对 WEB 文档进行特征提取。特征项由字、词或短语构成,通过相似度计算找到所有特征项,由它们组成特征项集。在此基础之上,通过文本分类的算法,如 K 最近邻分类算法和贝叶斯分类算法,将上面的处理结果归到一个或多个主题类别,加工整理后归档,将符合研究目的的内容用于撰写分析报告,其余内容可以保存作为长期跟踪材料。

网络监测的优点是对海量网络信息进行实时监测和挖掘,使得现代信息资源当中最丰富、最重要的网络信息能够更加便利和高效地得到利用。

(2) 信息抽取技术:情报分析需要智能化信息处理技术,来解决信息过载的问题。信息抽取作为智能化信息处理的前沿技术,是运用概念描述、关联分析、分类和聚类等功能,实现信息的智能化分析。具体来说,在对采集到的各种信息进行预处理后,通过对文本信息进行语义分析,可以获得预定主题的相关信息,然后从中抽取出具体的特征项,并将处理后的结构化文本信息存入数据库中,再进行下一步的各种分析,最终得到分析结果。这种抽取既可以用于期刊、会议记录等纸质版的信息抽取,也可以用于网络、数据库等。目前用于网页抽取的有网页结构分析、基于隐式马尔科夫模型和基于模式匹配等。

(3) 语义网络技术:目前采用计算机技术实现数量化方法进行情报的分析和预测已经比较普及。最常见的是语义网络技术,这是基于网络结构的一种知识的图解表示,以网络形式实现知识语义结构,使之能够通过多种机制来表达概念、规则及其之间的关联知识。语义网络作为一个带有标识的有向图,其节点表示各种事物、概念、属性和知识实体;链表示所连接节点之间的各种语义关联。国内有学者设计出定量化情报分析的数学模型——三因素(三方)二分(两种情况)网络,并且推广到 m 因素、 n 分网络及 k 值模型。编制软件实现了三因素和四因素二分模型进行信息分析,通过应用“某空军夜行团训练方式”的实例验证了语义网络模型进行情报分析的可行性。实例说明,依靠计算机软件来辅助定量化情报分析,比传统方法的效率和可靠性要高出很多。

(4) 数据挖掘技术:是基于数据仓库发展起来的一种知识发现技术,它是从大量的、不完全的、模糊的、随机的数据中,提取隐含的、有价值信息和知识的过程。数据挖掘技术能够分析特定的情报,将各种单一的情报进行综合,然后采取统计分析、技术群组、文本挖掘、组合理论、专利地图等技术,对其进行分析,然后以统计图谱、关联图谱和报告等形式展现出来。现在较常用的自然语言处理、语义关联分析、词频分布统计和语料学研究等数据挖掘方法,可以用于满足多种情报分析的需要。