

©刘 彤 蒋继娅 吴素研 编著

科技情报

KEJI QINGBAO
YU XINXI JISHU

与

信息技术

北京科学技术出版社

科技情报与信息技术

刘 彤 蒋继娅 吴素研 编著

 北京科学技术出版社

图书在版编目 (CIP) 数据

科技情报与信息技术/刘彤编著. —北京: 北京科学技术出版社, 2010. 2

ISBN 978 - 7 - 5304 - 3854 - 1

I. 科… II. 刘… III. 科技情报 - 情报工作 IV. G35

中国版本图书馆 CIP 数据核字 (2010) 第 018630 号

科技情报与信息技术

编 著: 刘 彤 蒋继娅 吴素研

责任编辑: 马光宇

执行编辑: 王宁波

封面设计: 沈学成

责任印制: 韩美子

出 版 人: 张敬德

出版发行: 北京科学技术出版社

社 址: 北京西直门南大街 16 号

邮政编码: 100035

电话传真: 0086 - 10 - 66161951 (总编室)

0086 - 10 - 66113227 (发行部) 0086 - 10 - 66161952 (发行部传真)

电子信箱: bjkjpress@163.com

网 址: www.bkjpress.com

经 销: 新华书店

印 刷: 三河国新印装有限公司

开 本: 787mm × 1092mm 1/16

印 张: 10.5

版 次: 2010 年 2 月第 1 版

印 次: 2010 年 2 月第 1 次印刷

ISBN 978 - 7 - 5304 - 3854 - 1/G · 969

定 价: 25.00 元



京科版图书, 版权所有, 侵权必究。

京科版图书, 印装差错, 负责退换。

序 言

随着信息时代的到来和社会对信息需求的极大增长，情报特别是科技情报对社会发展的作用日益突出，互联网的出现又带来了情报来源、工作主体、工作重心和服务模式的巨大变化。特别是如何从互联网的大量信息中甄别、获取有用情报成为科技情报工作的重中之重，也成为当今信息社会的一大难题。这些要求和变化，促使传统的科技情报技术发生革命性的变革。刘彤等人的著作《科技情报与信息技术》正是基于这一背景下，结合自身的实践而创作的。

该书涵盖了情报收集、情报分析方法、情报发布与交流服务以及信息安全保障等科技情报工作的主要流程和支撑技术，结合作者的实际工作经验，具体化了现代科技情报技术的内涵和实现过程，从而提高了著作的利用价值。特别是关于特色数据库的建设思想、建设原则和开发实例，为各级科技情报机构以特色求生存和发展，提供了富有战略性的发展方向。

作者以信息科学研究人员的视角分析科技情报工作，寻找情报学与信息科学的结合点。全书以现代信息技术对科技情报工作的支撑作为主线，具有较强的学术意义和实用价值。

全书特点可以归纳如下。

(1) 内容涵盖了科技情报的全部工作流程；探索了情报工作的内涵，对信息、知识、情报的基本关系提出了独到的辨析见解，体现了本书的基础性和科学性。

(2) 突出了现代信息技术，特别是一些前沿技术在情报信息各个环节上的作用，体现了本书的前瞻性和先进性。

(3) 结合作者的实践经验，将一些独到的见解和总结融入全书各章，体现了本书的创新性。

(4) 丰富的实例分析、设计实例和一些计算机辅助工具的介绍，为

读者提供了实用化的范例指导，体现了本书的实用性。

总之，该书选题学术水平高，应用价值强，富有时代特色，内容丰富，结构清晰，文笔流畅，图文并茂。对于科技情报专业人员和信息系统开发人员均有很好的参考价值。

我相信，《科技情报与信息技术》一书的出版，必将受到我国科技情报界及信息管理相关领域科研人员的欢迎，并且对提高科技情报工作者的信息化意识，引进智能化情报分析方法，推动科技情报工作方式创新具有积极的意义。

北京交通大学教授 裴政

2009年11月

前 言

在信息化、网络化的大趋势下，“信息化带动情报工作，情报工作促进信息化”已成为科技情报界普遍接受的观点。如何提高科技情报工作者的信息化意识，使其在日常的科技情报工作中，自主利用信息技术，建设更完善便捷的多元化科技信息情报网络，更好地为政府部门提供决策支持，为科技创新主体（企事业单位、高等院校、科研院所及科技人员）提供全方位的信息服务，已成为当前科技情报工作的重中之重。

作者在工作中发现：由于很多科技情报工作人员缺少信息科学背景知识，因此他们很难在情报的收集、分析过程中选择并应用各种信息技术，达到事半功倍的效果；而专业的信息技术人员由于缺乏对科技情报工作的具体认识，难以挖掘用户需求，找不到信息技术与科技情报工作的契合点。因此，本书旨在阐述信息技术对科技情报工作的影响和改变，以及如何在情报工作的各个环节充分利用信息技术，提高情报分析成果的科学性、合理性，提高科技情报机构在情报收集、分析加工、交流反馈以及资源建设方面的能力。情报研究人员只需要“按图索骥”，根据自己的工作需求就能找到相应的智能化解决方案，而不需要深入探究其中的算法原理和技术细节。

本书从网络环境和科技情报工作的实际需求出发，系统地论述了如何在科技情报工作的信息收集整理、情报分析加工、情报发布交流、信息资源建设、科技查新等方面使用信息技术，并结合实际工作给出了示例，力求突出理论的严谨性、技术的先进性、方法的系统性、语言的通俗性。

全书共分八章。第一章对知识经济时代下科技情报工作的内涵、服务范围、服务形式的变化，以及对社会发展和科技创新的促进作用进行了系统论述。第二章阐述了基于互联网的情报收集技术，重点介绍了垂直搜索引擎、RSS 订阅、舆情监控等最新的信息采集方法在情报收集工

作中的应用。第三章举例介绍了在情报分析过程中可利用的知识发现、知识挖掘与可视化表达技术,并对 Web2.0 环境下的情报研究组织形式的发展变化进行了论述。第四章介绍了在情报发布和交流过程中如何选择和利用动态网站、智能信息推拉、短信平台等信息交互手段。第五章以北京科技情报服务平台为例,介绍了自动化、智能化的信息采集、加工、发布体系。第六章论述了如何建立特色数据库系统,实现科技情报服务机构的信息资源建设目标。第七章介绍了科技查新业务管理系统的设计和实现思路。第八章介绍了与科技情报工作相关的信息安全技术。

全书由刘彤策划组织,并执笔编写第一、二、五、六、七、八章,第四章由蒋继娅编写,第三章为刘彤和吴素研共同编写。全书由刘彤进行统稿、修改。

本书在编写过程中得到了北京市科学技术研究院李永进副院长、北京大学谢新洲教授、北京交通大学袁正定教授的指导,在此表示衷心的感谢!

刘 彤

2009 年 11 月

目 录

1 绪 论	(1)
1.1 科技情报工作的流程	(1)
1.2 科技情报对社会发展的作用	(2)
1.2.1 科学决策的支持作用	(2)
1.2.2 降低科研创新风险	(3)
1.2.3 科技查新避免技术重复	(3)
1.2.4 加速科研进程	(3)
1.3 信息、知识与情报的关系辨析	(3)
1.3.1 信息、知识和情报的定义	(4)
1.3.2 知识与情报的关系	(4)
1.3.3 信息与情报的关系	(4)
1.3.4 信息、知识与情报的关系	(5)
1.4 科技情报对知识创新的促进作用	(5)
1.5 知识经济时代对科技情报的需求特点	(6)
1.6 网络信息时代科技情报工作的变化	(6)
1.6.1 科技情报的来源发生改变	(7)
1.6.2 情报工作主体发生变化	(7)
1.6.3 工作重心发生转移	(7)
1.6.4 工作方式发生变化	(8)
1.7 科技情报新的服务形式——公益情报服务	(8)
1.8 小结	(9)
2 基于互联网的情报收集技术	(10)
2.1 我国的互联网发展现状	(10)
2.2 搜索引擎在情报收集工作中的应用	(11)
2.2.1 百度网页搜索技巧	(12)
2.2.2 网络信息检索技巧总结	(14)
2.2.3 搜索引擎的使用步骤	(15)

2.3	垂直搜索引擎	(16)
2.3.1	垂直搜索引擎的特点	(16)
2.3.2	垂直搜索引擎实例	(17)
2.4	网络数据库	(21)
2.4.1	综合性数据库	(22)
2.4.2	专利数据库	(23)
2.4.3	专业数据库	(24)
2.5	数字图书馆	(26)
2.5.1	数字图书馆简介	(26)
2.5.2	数字图书馆的特色数据资源	(27)
2.5.3	数字图书馆的延伸——手机图书馆	(27)
2.6	专业科技情报网站	(28)
2.7	RSS 阅读器	(30)
2.8	舆情监控	(31)
2.8.1	方案概述	(32)
2.8.2	应用案例	(33)
2.9	小结	(33)
3	基于信息技术的情报分析方法	(35)
3.1	情报分析的目的、对象及作用	(35)
3.2	情报分析方法的国内外现状与发展趋势	(36)
3.3	网络调查法	(39)
3.3.1	网上问卷调查法	(39)
3.3.2	网上讨论法	(40)
3.3.3	专业在线问卷调查网站	(41)
3.4	基于数据挖掘技术的情报分析方法研究	(42)
3.4.1	数据挖掘的定义	(42)
3.4.2	数据挖掘在情报学领域的应用	(42)
3.5	常用情报分析方法的计算机实现	(46)
3.5.1	常用分析软件 Excel、SPSS、Matlab 的简介	(46)
3.5.2	专家调查法介绍及计算机辅助实现	(48)
3.5.3	层次分析法介绍及计算机辅助实现	(50)
3.5.4	模糊评判法介绍及计算机辅助实现	(54)
3.6	专利情报分析软件应用及实证研究	(56)

3.6.1	专利情报分析领域的相关信息技术及发展趋势	(56)
3.6.2	国内外常用专利分析软件汇总	(58)
3.6.3	实证分析	(60)
3.6.4	专利分析技术的发展趋势	(65)
3.7	情报预测	(65)
3.8	WEB2.0 下的情报研究工作	(67)
3.8.1	Web 2.0 技术简介	(67)
3.8.2	Web2.0 对科技情报工作的影响	(69)
3.9	情报研究报告的智能化管理	(71)
3.9.1	属性的描述	(71)
3.9.2	内容的描述	(72)
3.10	小结	(73)
4	科技情报的发布与交流	(75)
4.1	动态网站管理与发布反馈系统	(75)
4.1.1	网站建设技术综述	(75)
4.1.2	J2EE 应用与环境	(76)
4.1.3	基于 Model-View-Controller 模式的 struts 框架	(76)
4.1.4	Hibernate 基本用法	(77)
4.1.5	Spring 基本用法	(78)
4.1.6	Freemarker 模板引擎技术介绍	(79)
4.1.7	基于 SSH 和 Freemarker 技术的动态网站管理与发布系统	(80)
4.2	基于搜索引擎优化技术的网站优化策略	(83)
4.2.1	搜索引擎技术优化技术概述	(84)
4.2.2	模板引擎技术在 SEO 中的应用	(87)
4.2.3	网站搜索引擎优化实例	(88)
4.3	信息拉取与信息推送	(89)
4.3.1	信息拉取技术	(90)
4.3.2	信息推送技术	(91)
4.3.3	智能信息推拉技术	(92)
4.4	RSS 推送	(92)
4.5	博客	(93)
4.6	BBS	(94)
4.7	短信平台	(95)

4.8 小结	(95)
5 北京科技情报服务平台实例研究	(97)
5.1 系统简介	(97)
5.2 科技信息采集与存储体系	(98)
5.2.1 抓取器的结构	(99)
5.2.2 信息抓取管理平台	(100)
5.2.3 工作状况的实时监控与回溯查询	(101)
5.2.4 噪声处理	(101)
5.3 智能情报加工系统	(103)
5.3.1 分词功能介绍	(103)
5.3.2 去重功能介绍	(104)
5.3.3 分类功能介绍	(105)
5.3.4 自动摘要功能介绍	(106)
5.4 科技信息发布与共享平台	(108)
5.4.1 北京科技信息网简介	(108)
5.4.2 网站后台管理系统	(110)
5.5 小结	(112)
6 特色数据库建设	(113)
6.1 特色数据库建设的必要性	(113)
6.2 国内特色数据库建设现状	(114)
6.3 特色数据库的建设原则	(115)
6.3.1 突出特色,先进实用的原则	(115)
6.3.2 规范建设、确保质量的原则	(115)
6.3.3 重点建设、循序渐进的原则	(115)
6.3.4 公开、共享,保证信息安全的原则	(116)
6.3.5 可发展性原则	(116)
6.4 特色数据库系统的建设方法	(116)
6.4.1 特色数据库系统的构成	(116)
6.4.2 特色数据库的信息来源	(117)
6.4.3 特色数据库的数据加工	(118)
6.4.4 特色数据库系统的搭建方案	(118)
6.5 特色数据库建设实例	(120)
6.5.1 数据库简介	(120)

6.5.2 数据来源	(123)
6.5.3 系统功能	(123)
6.6 小结	(126)
7 科技查新业务管理系统设计与实现	(127)
7.1 系统简介	(127)
7.2 系统基本结构	(128)
7.3 系统基本功能	(128)
7.4 系统功能介绍	(129)
7.4.1 常用科技查新数据库汇总	(129)
7.4.2 用户委托书的网上提交	(129)
7.4.3 查新报告的填写	(130)
7.4.4 查新报告的审核与打印	(130)
7.4.5 检索统计功能	(131)
7.4.6 系统管理功能	(131)
7.5 查新业务管理系统的开发和应用价值	(131)
7.6 小结	(132)
8 科技情报业务系统相关信息安全技术	(133)
8.1 信息安全简介	(133)
8.2 科技情报业务系统相关的安全技术	(134)
8.3 身份认证	(134)
8.3.1 用户名/密码方式	(134)
8.3.2 智能卡认证	(135)
8.3.3 身份特征认证	(135)
8.3.4 源地址认证	(135)
8.4 访问控制	(135)
8.5 单点登录	(136)
8.6 数字签名	(137)
8.7 网络安全审计	(137)
8.8 安全传输通道	(138)
8.8.1 SSL 协议	(138)
8.8.2 HTTPS 安全超文本传输协议	(139)
8.9 数据安全备份	(139)
8.10 数据加密	(139)

8.11	日常性的安全管理措施	(140)
8.11.1	防操作系统漏洞	(140)
8.11.2	防数据库漏洞	(140)
8.12	常见防护手段	(140)
8.12.1	防恶意登录攻击	(140)
8.12.2	防 SQL 注入攻击	(141)
8.13	小结	(144)
附录 A	RSS feed 文件应用示例	(145)
附录 B	定时自动备份重要文件方法示例	(147)
附录 C	数据库定时自动备份方法示例	(149)
参考文献		(151)

1 绪 论

情报，是为了一定目的，具有一定实效性，满足特定需要的知识或消息。情报是传播中的知识，是意志、决策、部署、规划、行动所需要的知识和智慧。情报具有知识性、传递性和效用性等特点。科技情报专指通过公开信息渠道获取的有关科学发展、技术创新、最新动态的有用知识。科技情报研究工作是指运用先进的信息手段和定性定量相结合的研究方法，如信息整序、文献计量、分析综合、专家调查等情报研究和预测方法，针对特定的科学技术和技术经济问题，以信息收集——归纳整理——分析研究——服务利用为基本过程，对所掌握的信息进行科学地分析，并进一步提出具有总结性或认识性的情报，以向决策者、科研工作者提供增值的信息成果为特征的一项研究活动。其研究成果是产生某种认识、建议和方案。科技情报研究内容围绕科技发展的方针和政策、发展规划、科技经济和社会发展战略的制定，为企业的技术创新、政府的决策支持而服务，如重大技术的引进，科研项目的立项等。

我国科技情报工作始于 1956 年，在“科教兴国”战略的实施中发挥了重要的作用。进入 21 世纪以来，科技进步日趋加快，科技与经济日渐融合。我国科技情报（信息）机构作为科技发展的重要基础和国家科技创新体系的重要组成部分，已经成为促进信息传播与应用的骨干力量，科技情报信息服务已成为促进科技发展与经济繁荣的强大活力之一，并为国家的科技进步、经济发展和社会信息化做出了贡献。当前，我国在创新型国家的建设过程中，更离不开准确、及时、快速反应的科技情报系统的支撑。如何顺应知识经济时代的发展，充分利用信息技术，促进知识型科技信息服务业的发展，使科技情报服务于知识经济社会的发展，并具有前瞻性、时效性和专业化的特点，是每一个科技情报服务机构、每一位科技信息服务从业人员所必需面对的问题。

1.1 科技情报工作的流程

科技情报工作由多个环节前后衔接而成，包括情报收集、情报处理、情报分析、情报发布与反馈等环节。如图 1-1 所示。

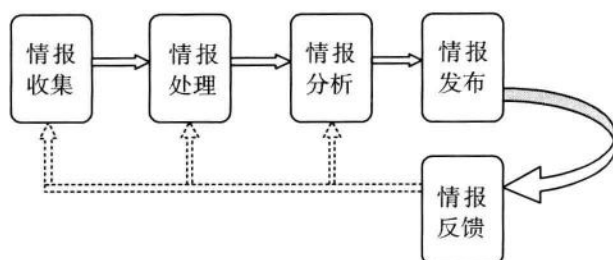


图 1-1 科技情报工作的流程

情报收集就是情报机关和情报人员根据收集规划和情报需求，通过多种渠道，有计划、有目的地广泛收集和系统积累情报资料的活动。情报收集的途径主要包括：报刊和专业杂志、行业协会出版物、产业研究报告、互联网信息、专业数据库、业务办公文档等信息。情报收集环节是开展情报工作的前提和基础。

情报处理主要是对收集到的原始信息进行初步处理，以便进一步分析或将处理结果直接提供给情报使用者，可人工进行处理或使用计算机技术进行分类标引、摘要生成等自动处理。

情报分析是对原始信息进行综合、评价、分析，使信息转化为深层次情报的过程，情报分析产品可表现为情报简报、情报专题等。这一阶段的工作必须由专门的行业资深人员经过对原始信息的深度分析才能完成，如利用 SWOT 分析、关键成功要素分析、价值链分析等情报研究方法，进行深度分析，为情报用户得出分析结果。

情报发布是指将收集、处理、分析之后的情报信息或情报产品，以适宜的形式（如通过网站、制作成印刷品）传递给最终情报用户。同时，还需要提供畅通的情报评价和反馈手段，以收集用户反馈需求，对情报收集内容、情报处理深度和广度等进行必要的调整和修正。

1.2 科技情报对社会发展的作用

天下大事，情报为先。随着社会信息化进程的加快，情报信息作为经济资源、战略资源、管理资源的重要性日益明显，科技情报工作为在经济和社会发展中的服务作用日益突出，主要表现在以下几个方面。

1.2.1 科学决策的支持作用

科学决策是贯彻“科学发展观”的重要体现。科学的决策离不开信息分析和情

报研究,将最新的科技信息提供给决策者,不仅可使决策者掌握最新动态,也可启迪决策思维。情报越全面、准确、及时,领导的决策就越有把握。因此,科技情报工作在传统“耳目、尖兵、参谋”作用的基础上,又被政府高层认为是“智囊团”和“思想库”。

1.2.2 降低科研创新风险

科研创新是一种新探索,这就决定了它的不可预测性和风险性。降低创新风险的一个重要因素是及时获得全面、准确的相关信息、情报。一个新的科研项目在创新开发之前,首先要收集和掌握国内外有关的科技情报,了解该项技术的历史、现状及发展趋势,前人研究的经验或失败的教训,以及当前该技术在国内外的发展水平等,并将其作为论证的重要依据,这样才能保证技术创新开发具有生命力、创造性和竞争性。

1.2.3 科技查新避免技术重复

目前,我国有50%~70%的科研人员是在对世界科研动态、科研成果、科技文献缺乏足够了解的情况下开始科研工作的,所以课题立项重复,创新程度不高,造成无效劳动。据统计,我国的科研项目重复率高达40%。通过科技查新能及时了解、掌握国内外的科研成果和科技动态,帮助科研人员“站在巨人的肩膀上”,可以避免技术重复,使科研工作不走或少走弯路,保证开发项目的先进性。

1.2.4 加速科研进程

目前,世界上每年出版图书、期刊、发表的学术论文、公开的专利说明书、技术报告等文献浩如烟海。通过情报人员进行收集、加工、整理等创造性的知识劳动,可以把情报文献加工成简报、汇编、综述、行业发展报告、研究报告、专题数据库等多种形式的二次文献或三次文献,便于用户吸收、消化和使用,提高科研人员对文献的利用水平,同时还可以为科技创新活动节约大量的时间和经费。

1.3 信息、知识与情报的关系辨析

信息、情报和知识是情报学中的基本概念,信息与情报、知识与情报、以及三

者之间的关系辨析一直是学术界颇有争议的话题。1992 年以来,很多科技情报研究所更名为科技信息研究所,这一做法更是引起了社会广泛的争论。本节对信息、情报、知识的定义及其相互关系进行初步的探讨。

1.3.1 信息、知识和情报的定义

信息是事物运动时发出的信号所带来的消息,是事物存在方式和运动规律的一种表现形式。不同的事物具有不同的存在方式和运动规律,从而构成了各种事物的不同特征。信息普遍存在于自然界、社会界以及人的思维之中,是客观事物本质特征千差万别的反应。

知识是人们对客观事物运动规律的认识,是经过人脑加工处理过的系统化了的的信息。知识是人类经验和智慧的总结,是人们科学地认识世界、改造世界的力量。知识具有高度的抽象性、概括性、系统性、可靠性、科学性,是人们对客观事物本质和规律的认识。

情报是指传递着的有特定效用的知识,是人们为了解决某个具体问题所需的知识信息。知识性、传递性和效用性是情报的基本属性。

1.3.2 知识与情报的关系

著名学者钱学森指出:“情报就是为了解决一个特定问题所需要的知识”、“情报是激活了的知识”。情报是一种特殊的知识,只有针对特定问题,经过激活、活化后的知识才是情报,所以知识并不都是情报。

1.3.3 信息与情报的关系

信息不完全是情报,情报不等于信息。情报包含于信息之中,通过对信息进行筛选、分析、综合而得到的对实现主体特定活动有用的部分,便是情报。情报具有知识性、效用性、针对性、时效性。可见,信息是比情报更广的概念,是情报的源泉和基础。情报是被研究、加工了的有用信息。因此,情报的产生离不开情报研究者的智力因素。即使在信息技术高度发达的今天,情报的收集、甄别、加工、发布和交流等工作也不能完全依赖计算机自动完成。没有人的活动,信息永远无法上升到情报的高度,所以在科技情报工作中,人的作用是不能忽视,更不能替代的。