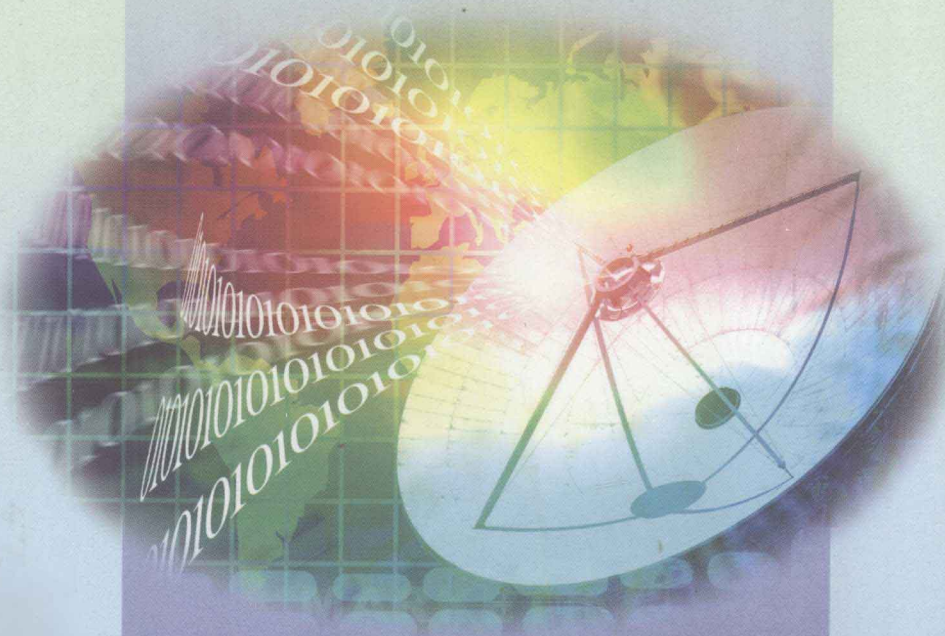




全国高等农业院校教材
全国高等农业院校教学指导委员会审定

信息检索与利用

谢德体 宛章齐 主编



中国农业出版社

全国高等农业院校教材
全国高等农业院校教学指导委员会审定

信息检索与利用

谢德体 宛章齐 主编

中国农业出版社

图书在版编目 (CIP) 数据

信息检索与利用/谢德体, 宛章齐主编. —北京: 中国农业出版社, 2002.12
全国高等农业院校教材
ISBN 7-109-07947-3

I. 信... II. ①谢...②宛... III. 情报检索—高等学校—教材 IV. G252.7

中国版本图书馆 CIP 数据核字 (2002) 第 079319 号

中国农业出版社出版
(北京市朝阳区农展馆北路 2 号)
(邮政编码 100026)

出版人: 傅玉祥
责任编辑 刘振生

农业出版社印刷厂印刷 新华书店北京发行所发行
2002 年 12 月第 1 版 2004 年 2 月北京第 2 次印刷

开本: 787mm×960mm 1/16 印张: 15.75

字数: 278 千字

定价: 21.40 元

(凡本版图书出现印刷、装订错误, 请向出版社发行部调换)

前 言

《信息检索与利用》课程包括信息的“检索”和“利用”两部分内容。就信息的“检索”而言,当今社会的文献信息资源呈现爆炸性增长趋势,网络信息更新速度加快,计算机检索日益普遍,迅速取代手工检索成为最主要的检索手段,使得传统的文献信息检索教材的内容明显滞后。就信息的“利用”而言,对检索的信息还需要去粗取精,去伪存真,提高利用率,而这方面正是目前的文献信息检索教材的薄弱环节。

本书作为全国高等农业院校“十五”规划教材,首次实行公开招标、联合编写,具有内容新颖、实用、全面的特点。对于“检索”部分的内容,主要介绍计算机检索,选材立足于网络信息资源,压缩了手工检索理论和检索工具的内容。我国加入WTO后加强知识产权保护,使得专利商标信息的检索具有非常重要的意义,本书扩展和加深了专利信息检索的内容,增加了“商标信息检索”。对于信息的“利用”部分,增加了文献统计分析和竞争情报的基础知识,以适应素质教育,提高学生通过文献发现问题、分析问题的能力,适应市场竞争。在学科覆盖面上,内容兼顾了非农业专业,具有通用性,可作为各类高等院校的“信息检索与利用”教材。

全书由西南农业大学图书馆谢德体教授担任第一主编,华中农业大学图书馆宛章齐副研究馆员担任第二主编。各章节编写人员如下:西南农业大学图书馆谢德体编写第九章,陈蔚杰编写第三章,闫勇编写第七章,李玉兰编写第六章第三节二小节;华中农业大学图书馆宛章齐编写第二章,叶勤编写第五章,余丽清编写第八章第一、二、三节,车永坚编写第八章第四、五、六节;甘肃农业大学图书馆刘喜编写第六章第一、二节,第三节一小节;湖南农业大学图书馆李晋衡编写第四章;天津农业大学图书馆吴宝华编写第一章。

本书参考了同行的教材、专著的研究成果,在此一并表示衷心感谢。由于编者水平有限,时间仓促,难免有疏漏、错误之处,敬请同行和读者批评指正。

编者

2002年9月

主 编	谢德体	宛章齐		
副 主 编	陈蔚杰	刘 喜	李晋衡	吴宝华
编写人员	谢德体	宛章齐	陈蔚杰	刘 喜
	李晋衡	吴宝华	闫 勇	李玉兰
	叶 勤	余丽清	车永坚	

目 录

前言

第一章 绪论	1
第一节 基本概念	1
一、信息、知识、文献和情报的基本概念	1
二、信息、知识、文献和情报的相互关系	3
三、文献信息的特点	4
四、文献信息的发展趋势	5
第二节 信息的类型与级次	8
一、信息的类型	8
二、信息的级次	11
第三节 信息检索与利用的意义	12
第四节 信息意识	13
一、信息意识及其表现形式	13
二、信息意识的形成条件	14
三、加强文献信息教育, 培养大学生信息意识与创新能力	15
第二章 信息检索基础	17
第一节 信息需求类型	17
一、撰写论文	17
二、科学研究和学科建设	18
三、生产开发	18
四、回答或解决单一问题	18
五、对已知文献的查找	19
第二节 检索工具	19

一、检索工具的类型	19
二、手工检索工具的主要排检方法及原理	19
三、计算机检索工具及其原理	26
第四节 检索方法及步骤	31
一、检索方法	31
二、检索途径	32
第五节 检索效果评价及修正	36
一、评价指标	37
二、检索策略	38
三、检索结果修正	38
第三章 Internet	41
第一节 Internet 概述	41
一、Internet 的起源	41
二、Internet 的工作原理	41
三、Internet 的服务方式	42
第二节 搜索引擎	48
一、概述	48
二、常见搜索引擎的主要原理	49
三、搜索引擎的类型	51
四、搜索引擎的使用技巧	69
第三节 新闻组	74
一、概述	74
二、新闻文章的格式	75
第四节 广域信息服务 Wais	76
一、概述	76
二、Wais 的用法	77
第四章 事实数据检索	81
第一节 概述	81
一、事实数据检索的定义	81
二、事实数据检索工具的特点	81
第二节 检索工具及使用方法	82
一、印刷型检索工具（参考工具书）及使用方法	82

二、数据库型检索工具及使用方法	91
第五章 图书信息检索	95
第一节 图书信息检索的一般问题	95
一、定义	95
二、图书信息检索特点	95
第二节 图书信息检索工具及其使用方法	96
一、图书出版信息检索工具及使用方法	96
二、图书馆目录信息检索系统及使用方法	97
三、网上书店及其检索	101
四、数字图书馆信息检索系统	106
第六章 期刊信息检索	109
第一节 概述	109
一、含义	109
二、期刊信息检索的特点	109
第二节 中文期刊信息检索工具和使用方法	110
一、印刷型检索工具	110
二、数据型检索工具及使用方法	112
第三节 外文期刊信息检索工具及使用方法	119
一、印刷型检索工具及使用方法	119
二、数据库型检索工具及使用方法	137
第七章 专利及商标信息检索	150
第一节 专利信息检索	150
一、专利概述	150
二、专利的类型	150
三、专利的审批程序	151
四、专利信息	151
五、专利信息检索工具	152
第二节 商标信息检索	175
一、商标概述	175
二、检索工具及使用方法	177

第八章 其他信息检索	181
第一节 标准信息检索	181
一、概述	181
二、检索工具和使用方法	183
第二节 会议信息检索	188
一、概述	188
二、检索工具和使用方法	189
第三节 科技报告检索	196
一、概述	196
二、检索工具和使用方法	197
第四节 学位论文检索	198
一、概述	198
二、检索工具和使用方法	199
第五节 政府及国际组织出版物检索	201
一、概述	201
二、检索工具和使用方法	202
第六节 报纸信息检索	204
一、概述	204
二、检索工具和使用方法	204
第九章 竞争情报	210
第一节 竞争情报概述	210
一、竞争情报的定义和内容	210
二、竞争情报的作用	211
三、竞争情报研究的现状	211
第二节 竞争情报源及收集方法	214
一、竞争情报源	214
二、竞争情报搜集方法	216
第三节 竞争情报分析方法	219
一、概述	219
二、专利信息分析法	221
三、定标赶超	226
四、竞争环境监测	233

第四节 竞争情报案例	236
第五节 竞争情报系统	238
参考文献	240

第一章 绪 论

第一节 基本概念

一、信息、知识、文献和情报的基本概念

(一) 信息

信息 (information) 一词来自拉丁语, 在日常生活中, 常被人们理解为以口头、书面或者以其他方式传递的消息。就信息的基本内涵来讲, 信息是自然界、人类社会以及思维活动中普遍存在的现象, 是一切事物自身存在方式以及它们之间相互关系、相互作用等运动状态的表达。信息不是物质本身, 而是物质的一种基本属性。信息与客观事物一样, 无处不在, 无时不在。人类认识世界的过程, 实际上就是不断从外界获取信息、加工信息和传递信息的过程。在当今信息社会中, 信息与物质、能源一起被视为社会经济发展的三大支柱, 成为促进社会经济发展的重要战略资源。

信息的基本属性是客观性、传递性、共享性和中介性。信息的客观性是指信息是现实世界中各种事物运动方式与状态的反映, 其存在是客观的, 不以人的主观意志为条件。它可以被人所感知、处理、存储、传递和使用。信息的传递性是指任何信息只有经过其载体的传递, 才能被接受、处理和使用。信息的传递性不仅表现在跨越时空的传递, 而且还表现在通过不同的载体形式和不同手段的传递。信息的传递特性是实现信息资源共享的基础。信息的共享性是指, 同一信息传递给多个用户使用, 信息的传递者并不因为提供了信息而失去原有的信息或者减少原有的信息量。因此, 信息的共享性是推动社会交流的主要原动力。信息的中介性是指信息是人类认识世界和改造世界的媒介, 在实现人类与自然界的沟通中起着桥梁的作用。人类通过自己的感觉器官从物质世界中感知和选择信息, 然后通过大脑的加工, 形成再生信息作用于物质世界,

达到改造世界的目的,在这一过程中,信息始终起着认识事物的中介作用。

(二) 知识

知识(knowledge)是人类社会实践的总结,是人类通过信息对自然界、人类社会以及思维方式与运动规律的认识和掌握,是人脑通过思维重新组合的系统化的信息集合。可见,信息是构成知识的原料,这些原料通过人脑接收、选择、整理、提炼等加工过程,形成各种各样的知识。知识的产生离不开信息和信息传递,它是信息的一个重要组成部分。随着人类认识的不断深入发展,新的知识不断产生、不断积累,从而使人类的知识体系不断丰富、发展和完善。

在信息时代,知识已经成为发展生产、提高竞争力的关键因素。现在世界竞争的核心归根结底是智力和知识的竞争,谁拥有了知识和信息,谁就能够在竞争中抢占高新技术的制高点和最前沿。因此,知识在经济发展中的互动作用愈来愈明显。

知识从不同的角度一般可以划分为下列四类:

知事(know-what),即关于事实的知识,这类知识是人们在日常生活和工作中常常用到的。

知因(know-why),即关于科学原理和自然科学方面的知识。这类知识是国家技术进步的基础。

知能(know-how),即做事的技巧和能力。知能往往是居于单个的企业或个人所拥有的知识,并且仅限于其自身范围内使用而不向外传播的知识,秘诀和窍门就属于这一类知识。

知人(know-who),即关于谁知道什么和谁知道如何做什么的信息。了解这类知识可以与有关专家建立联系,从而有效地利用他们的知识。尤其在知识经济条件下,运用知人这一类知识对于科学管理和组织来说,是非常重要和有效的手段。

(三) 文献

文献(document)是一个发展的概念。“文献”一词最早出现于孔子《论语·八佾》中:“夏礼吾能言之,杞不足徵也;殷礼吾能言之,宋不足徵也;文献不足故也。足,则吾能徵之矣。”在这里,孔子只是说明文献的足与不足作为能否讲解杞、宋两国礼制的一个依据,而并未说明文献究竟是指什么。

南宋朱熹在他所著的《四书章句集注》、《论语集注》中,对“文献”进行注释:“文,典籍也;献,贤也。”按照朱熹的意思,典籍和贤者的言论或指有关典章制度的文字资料和精通典章制度的学者即为文献。

以后,随着历史的发展,文献概念逐渐演化为专指有价值的各种图书资料。到现代,随着现代科学的飞速发展,记录知识的手段和载体越来越多样

化,文献的外延不断扩大。国际标准化组织在其制定的《文献情报术语国际标准》(ISO/DIS5217)中对文献的解释是:“在存储、检索、利用或传递记录信息的过程中,可作为一个单元处理的,在载体内、载体上或依附载体而存储有信息或数据的载体。”

1999年新版《辞海》对文献的定义解释为“记录有知识的一切载体的统称”。我国在1983年公布的《中华人民共和国国家标准·文献著录总则》(GB3792·1-83)把文献定义为:“记录有知识的一切载体。”可见,凡是记录有知识的一切载体都可以称之为文献。古代甲骨文记录以龟甲、兽骨作为材料,是文献;有文字和图形的碑刻是文献;竹简和帛书是文献;有铭文的青铜器是文献;现今的机读资料、电子出版物、缩微制品等等都是文献。我们可以认为,尽管文献的载体材料、记录手段在不断演进,但是构成文献的四大要素依然故我,即被记录的知识内容、承载知识内容的载体、记录知识的符号和记录知识内容的手段。文献具有记录、贮存和传递知识的作用,是人们获取知识、信息和情报的重要来源,也是交流传播知识、信息和情报的最基本的手段。

(四) 情报

情报(information)也是一个发展的概念,即在不同历史时期,其概念有所不同。在古代,情报首先产生于军事领域,所以就有“战时关于敌情之报告,曰情报”之说,以后,随着科学技术的迅速发展,新知识的大量涌现,知识和信息处理、传递技术的不断革新,使知识和信息的传播工作进入一个新的发展阶段,加深了人们对情报概念的认识,认为情报是“作为交流对象的有用知识”,是“在特定时间、特定状态下对特定的人提供的有用知识”,是“激活了、活化了的知识”。情报的基本属性是知识性、传递性和效用性。“情报”被广泛用于政治、经济和文化领域。人们在社会实践中源源不断地创造、交流与利用着各种各样的情报,情报是经济建设、科研、生产、经营管理等不可缺少的宝贵财富,是进行决策、规划、管理的主要依据。

情报的本质是知识,但知识并不等于情报,知识必须传递给用户、能满足用户的特定需要才能成为情报。可见,情报是知识的一部分,是特定的知识,是经过激活过程的活化了的知识。这里的激活过程,就是指为了便于用户查找和使用,对文献(即知识)进行加工整理,使之有序化、系统化的过程。

二、信息、知识、文献和情报的相互关系

综上所述,信息是事物存在方式及运动状态的表达,知识是人脑通过思维重新组合的系统化的信息集合,情报是对人提供的有用知识。世界是物质的,

物质的运动便产生了信息；各种信息经过人们进行系统化的加工处理，便转化为知识；知识经过系统化的加工处理转化为情报；情报应用于实践，解决实践中存在的问题，创造出物质财富或精神财富，这时的情报，便转化为生产力，产生新的信息，形成一个无限循环的转化过程。因此，信息、知识、文献、情报之间的逻辑关系是：信息包含知识，知识包含情报。文献是记录有知识的载体，当文献中记录的知识传递给用户，并为用户所利用时，文献中的知识就转化为情报；情报虽大多来自文献，但也可来自口头和实物，所以情报与文献是交叉关系。它们的关系不仅如此，而且可以相互转化。例如，知识在需要被用来解决特定问题时，便转化成情报；情报在不需要利用时，便还原为客观的知识。知识和情报，对于既不能认识又不能理解的人来说，它们都不过是一种信息。其间的关系如图 1-1 所示。

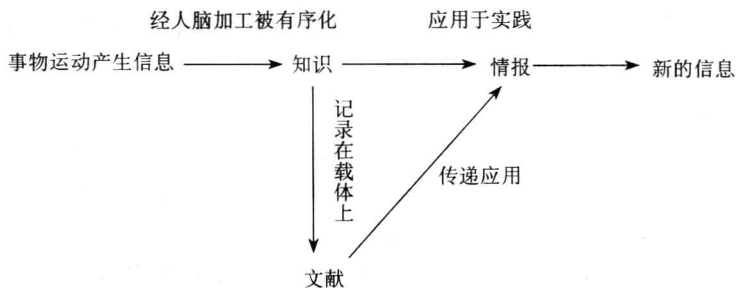


图 1-1 信息、知识、文献和情报之间的关系

三、文献信息的特点

文献信息是人们检索和利用的对象，要更好地检索和利用文献信息，必须对文献信息有一个充分的认识，特别是要正确掌握文献信息的特点。一般来说，文献信息的特点主要有：

（一）知识性

知识是文献信息的实质内容，没有记录下任何知识内容的载体不能称为文献信息。文献信息之所以产生，就是为了记录认识世界、改造世界的过程中所获得的知识 and 科学成果。人们之所以需要检索和利用文献信息，就是需要文献信息中所载有的知识内容。人们吸收了文献信息的知识内容，就可以增长自己的才干或者解决特定的问题，促进自己的工作。

（二）载体性

人们头脑中的知识不论如何丰富,都不能称为文献信息,只有将知识用一定的方法、手段记录在载体上,方能形成文献信息。文献信息之所以能成为人类知识存在的主要形式,就是依赖于载体。载体是文献信息的物质实体,它是客观存在的,可以独立于人类的存在和意识之外,使人类在认识世界和改造世界中取得的经验和认识得以保存和流传,从而推动人类社会不断进步。离开了载体,知识的有效记录和继承就无法实现。

知识性和载体性是文献信息的两个最根本的属性,由它们导出了其他一系列的属性。

(三) 累积性

科学文献信息记录了人类几千年的科学发现、发明创造和一切成果,汇集着人类的文明。文献信息的累积性,不仅表现于载体的累积性,更重要的表现于其知识内容的累积性和继承性。按英国哲学家卡尔·波普尔的理论,人类历史上形成的文献构成了客观知识世界。有史以来,全世界任何一个国家在任何一个时代创造的科学发现、发明得以保存、流传到今天,为全世界人民所使用,正是文献信息累积性的结果。如果没有文献信息的累积性,如果每一个国家和每一个时代的科学家都必须独立地去积累所需要的知识,重新去发现规律,那么,科学就不可能有目前的高速发展。正是由于文献信息的累积性,使得当代科学家可以把前人的成就作为开展自己工作的基础。

(四) 可传递性

载体是文献信息的物质实体,是知识的一种物质外壳。载体既然是一种物质,就必然具有物质的属性——运动、空间、时间等。由此,文献信息可以从空间的一点移到另一点,从时间的一刻移到另一刻。正是通过其载体的传递达到文献的知识内容的传播和利用。

上述特点是科学文献信息所具有的共同点,不仅如此,社会科学文献信息还具有不同于自然科学文献信息的许多特点,如思想性、时代性、综合性、研究成果的不确定性以及长久的效用性等。

四、文献信息的发展趋势

文献信息不仅具有上述特点,而且随着现代科学技术的发展,在内容、形式、品种、数量等方面还会发生不断的变化,其发展趋势主要表现在如下几个方面。

(一) 文献信息的出版周期缩短,数量急剧增多

在信息时代,科学的发展、知识门类的增加以及人们对信息和知识的需求

量的增长,导致了各种记录人类科学知识的文献信息载体出版周期的缩短和出版数量的激增。据统计,在17世纪以前,人类知识的积累大约1 000年翻一番。17世纪至19世纪,大约每100年翻一番;19世纪以来大约每50年增加1倍;20世纪中叶每10年增加1倍,70年代每5年增加1倍,80年代后每3年增加1倍。一些尖端领域和新兴学科每2~3年文献量就要翻一番。从图书期刊出版来说,20世纪70年代以来,全世界每年出版图书50万种,大约每分钟都有一本新书问世;到20世纪90年代,全世界每年出版图书多达80多万种;目前全世界出版的科技期刊约有6万余种,每年发表科技论文约500万篇,平均每天发表论文1.3万~1.4万篇;全世界每年出版的专利说明书100多万件,累计已达3000万件;全世界的技术标准总数为20万件;每年出版的会议录1万多种,科技会议论文10万篇以上;每年公开出版的科技报告20万篇,学位论文20多万篇。社会科学文献信息出版物同样如此。据联合国教科文组织统计,现在世界各国每年出版的社会科学著作、文集和文章已达100万件。为适应信息化社会的需要,这种趋势还将继续下去。这一趋势将使建立统一的文献信息资源体系成为必要,使文献信息资源的协调、搜集、利用等工作的现代化变得刻不容缓。

(二) 文献信息的语种范围不断扩大,译文数量增多

过去,科技文献发表的语种大部分只是用英、法、德等少数几种语种,随着科学技术的发展,这种局面已经被打破,现在科学文献信息所使用的语种已越来越多,其中英文、德文、俄文、法文、西班牙文、中文、日文所占比重较大。例如,国际食品情报所编辑、英国出版的《食品科学与工艺文摘》(FS-TA),收录的文献语种达48种,美国《化学文摘》(CA)所收录的文献语种除了英语外,还有其他55种语言文字。在这些语言文字中,英文文献信息量约占全世界信息量的2/3,但这个比例已在逐渐下降,其他语种比例不断上升。据有关资料统计,当今发表的文献信息中有1/2是用半数左右的科学家所不懂的语言出版的。因此,为解决读者利用文献的困难,大量翻译外国文献信息已成为各国社会科学研究中的必然趋势。

(三) 文献信息类型复杂,载体形式多样化

人类社会在发展科学技术的同时,也在不断发展记载科学技术的知识手段。现代的科学文献信息,除了拥有图书、报纸、期刊、会议论文、档案等类型繁多的印刷型文献信息外,还产生了大量非印刷型、现代化的文献信息载体,如缩微胶片、视听资料、光盘、软盘、计算机网络等。而且,随着社会科学文献信息品种和数量的激增,文献信息贮存、检索的方式方法必将日趋科学合理。

(四) 文献信息的离散度越来越高, 淘汰速度加快

随着现代科学技术的不断发展, 各门学科、领域之间不断分化和综合的现象日趋严重, 造成了文献信息内容交叉重复, 文献信息的分布异常分散。据统计, 某一个专业的文献, 在本专业期刊上发表的只占 50%, 而另外 50% 则发表在相关的专业期刊上; 对一个专题范围内的文献来说, 约有 1/3 登载在刊名与该专题相同的期刊上, 约有 1/3 登载在刊名与该专题有关的期刊上, 其余 1/3 则登载在刊名与该专题无关的期刊上。

另外, 根据引文、索引的统计, 地质学文献有 15.6% 引自非地质期刊; 数学文献有 21% 引自非数学期刊; 物理文献有 25% 引自非物理期刊; 化学文献约有 27% 引自非化学期刊。

此外, 科技期刊的专业面越来越宽, 内容越来越杂, 越来越成为名副其实的“杂志”。有人对 1129 种期刊的论文内容进行分析统计, 只涉及 1 个学科的只占 11%, 包含 2~4 个学科的占 53%, 包含 5~7 个学科的占 36%。文献发表的这种分散性, 增加了查找和利用文献的难度。

英国的化学家和文献计量学家布拉德福 (S.G. Bradford) 经过长期的统计分析后提出了文献分布的规律, 认为在某一特定学科领域内, 大部分文献集中在较少量的期刊中, 而其余少部分则分散在大量的边缘学科或其他学科的文献中, 这一规律被称为“布拉德福定律”。根据这一定律, 提出了“核心期刊”的概念。

文献信息不仅离散度高, 而且淘汰速度越来越快。现代科学技术发展的一个明显的特点是速度快, 成果多, 知识量大。知识的新陈代谢频繁, 陈旧率加快, 不成熟的观点和理论很快就被成熟的观点和理论所代替, 科研成果转化为生产力的周期也在大大缩短。

与此相关, 现代文献的老化加快, 时效性越来越强, 超过一定时限就会降低, 甚至失去使用价值。一般来说, 图书的有效期为 10~20 年, 期刊论文 3~5 年, 科技报告 10 年, 学位论文 5~7 年, 标准文献 5 年, 专利文献短于法律规定的 15~20 年。某些关于新技术、新工艺的更替速度更快。

有人对科技文献的老化速度提出了“半衰期”这一概念, 即在某一专业所有的文献中约有一半文献已停止使用所经历的时间长短。据统计, 当前各学科科技文献的半衰期分别为: 生物医学 3 年, 物理学 4.6 年, 化工 4.8 年, 生理学 7.2 年, 化学 8.1 年, 数学 10 年, 地理学 16 年。

科学技术发展越快, 文献的寿命越短, 这对文献检索工作提出了更高的要求。