

# 第1章

## 绪论

随着信息技术的发展,互联网的应用得到广泛普及,信息环境发生了相当大的变化,信息社会给人们带来了海量的信息,供人们参考、借鉴和学习。如何快速地从海量的信息中获取最有价值的东西,成了人们最棘手的问题。因此,只有掌握好信息检索的理论和技能,提高信息检索能力,才能快速、合理地利用信息资源。

信息化社会的发展对高等教育提出了更高的要求。为了全面提高大学生的素质,以适应信息时代的要求,许多国家将信息素养教育作为培养新世纪人才的重要内容,而文献信息检索已成为实施信息素养教育的核心要素,其目的是培养学生的信息意识、信息检索能力、信息吸收能力和信息整合能力,最终提高学生的信息利用能力和知识创新能力。

### 1.1 信息检索的现状

#### 1.1.1 信息与信息载体

信息这一概念最初是由 C.E.Shannon 和 W.Wiener 提出来的,他们试图给信息一个正式的、定量的定义,从通信工程、计算机和电信的角度来看,一个消息中携带信息量的大小用比特来衡量。信息是信息论中的一个术语,我们常常把消息中有意义的内容称为信息。

不同研究领域的学者对信息的理解各有不同。

哲学家认为:信息是事物存在的方式和运动状态的表现形式;

数学家认为:信息是一种概率;

物理学家认为:信息是“熵”;

通信学家认为:信息是“不定度”的描述;

图书信息领域的专家认为:信息是通过各种形式进行传播、记录、出版及发行的概念、事实及论著。

可见,信息的概念十分复杂,从不同领域、不同角度理解会产生不同的结果。

我们可以这样认识信息：自然界与人类活动的事实及人类对它们的认识和创造是信息的组成内容，而载体记录和媒体传播则是信息存在的物理形式。在人类进入信息社会的时代，信息已成为发展科技、经济、文化、教育的重要支柱之一。

信息载体是在信息传播中携带信息的媒介，是信息赖以附载的物质基础，即用于记录、传输、积累和保存信息的实体，信息载体的物质形式是多种多样的。如果是被人感知了，信息就会通过传导神经网络导入人的大脑。如果信息只是停留在人的大脑中，就像一个既不会说话又不能写字、没有任何表达能力的人一样，毫无价值。所以，反映到大脑中的信息只有表达出来才能发挥它的价值。如果说出来，信息就依附于声音中；写出来，信息就依附于文字中；画出来，信息将会依附于图像中。这样，文字、图片、图形、广播、电视、电话、语音、音乐、影视、数据库等就承载了信息，成为信息的载体。信息的载体分为两种：一种是以能源和介质为特征，运用声波、光波、电波传递信息的无形载体；一种是以实物形态记录为特征，运用纸张、胶卷、胶片、磁带、磁盘传递和储存信息的有形载体。正是由于主体需要表达从事物中感知的信息，才要借助一定的载体。同一个信息可以依附于不同的载体表达出来。

### 1.1.2 信息的特征

所谓信息的特征，就是指信息区别于其他事物的本质属性。通过对信息概念的分析，可以总结出信息具有如下特征。

#### 1. 信息的客观性

信息是事物运动变化和状态的客观反映，其实质内容具有客观性，信息客观性的特征是由信息源的客观性决定的。由于运动是普遍存在的，也就决定了信息的普遍性。信息普遍存在于自然界和人类社会，同时，信息本身也具有客观实用性。

#### 2. 信息的依附性

信息必须依附于一定的载体而存在，必须以符号、文字、图形、音频、视频等形式依附于书籍、磁带、磁盘、光盘等载体上。信息与载体不可分割。

#### 3. 信息的扩散性

信息的传递性决定了信息的可扩散性，也就是说，信息通过各种渠道进行传播，信息网络的发展更促进了信息的扩散。

#### 4. 信息的扩充性

人们对信息的感知和获取是不断增长的，因此信息资源的扩充与积累也是无限的。人们对信息处理能力越强，信息扩充得就越快。

#### 5. 信息的替代性

信息的物质形态是可以互相转移和变换的。

#### 6. 信息的共享性

信息可共享，在信息扩散和用户分享信息的过程中，信息载体本身的信息量并不会因此而减少，各用户分享的信息份额不因分享人的多少而受影响。

社会的进步赋予了信息更丰厚的内涵，信息的膨胀与人们对其需求的激增，使信息成为当今社会生活的一大支柱，成为一种与能源、材料并存的重要战略资源。

### 1.1.3 信息的功能

信息作为维系社会发展的三大要素之一，无论是在自然界，还是在人类认识世界和改造世界的活动中都具有多方面的功能，发挥着重要作用。

#### 1. 资源功能

信息已经成为 21 世纪人类社会的重要资源，科技的进步、社会的发展都与信息密切相关，人类通过对客观世界各种信息的认识、处理、吸收、利用和物化，促进社会的持续发展。

#### 2. 组织管理功能

从管理的角度来讲，管理系统是一个信息输入、处理、输出与反馈的系统。在这一系统的运作过程中，每个环节都必须以信息为依据，也必须以信息作为相互联系的条件。

#### 3. 中介功能

信息的中介功能表现在人与客观事物之间和人与人之间。人与客观事物的认识和联系是以信息的存在为条件的，无论是物质系统还是精神系统，系统内部和外部的联系都必须通过相应信息的联系过程来实现。因此，信息是沟通的桥梁和纽带。

#### 4. 消除不确定性（解惑）功能

这一功能是相对于信息接收者的状态改变而言的。在接收者收到关于某一事件的信息之前，可能对事件存在着多种估计，当接收到关于这一事件的信息之后，接收者会改变原有的估计状态，甚至会使原有的多种估计中的某一种唯一地确定下来，从而消除不确定性。

#### 5. 传播功能

信息的组织管理功能、中介功能和消除不确定性（解惑）功能都是通过其传播活动来实现的。信息的传播功能是以信息内容的可传输性为基础的。

### 1.1.4 知识、文献的概念及二者与信息的关系

信息前文已述及，现主要介绍知识和文献的概念，以及信息、知识、文献的相互关系。

#### 1. 知识

国外对知识的理解有以下几种。

《韦伯斯特词典》对知识的解释是：“知识是通过实践、研究调查获得的关于事物的事实和状态的认识，是人们获得的关于真理和原理的认识的总和。”

1973 年，美国学者贝尔在其著作中指出：“知识是对事实或思想的一套有系统地阐述，提出合理的判断或经验性的结果。它通过某种交流手段，以某种系统的方式传播给其他人。”

在以上观点中，“知识”不再是一个简单的、多元素的无序集合，而是被纳入一个动态的、与人或组织交互的系统中。更确切地说，只有在“使用”过程中，知识才体现出其价值，才成为有实践意义的、真正的知识。

国内对知识的理解有以下几种。

“知识”一词在《辞源》中有两种解释：一是“相识相知的人”，二是“指人对事物的认识”，这与现代汉语中的含义相近。后一种含义最早出现在清朝洪亮吉的《洪北江集》中：“孩提之时，知饮食而不知礼让，然不可谓非孩提时之真性也。至有知识，而后知家人有严君之义焉。”

1980 年出版的《辞海》中将“知识”定义为“人们在社会实践中积累起来的经验”，并指出“从本质上说，知识属于认识的范畴”，国外有些学者认为知识是一种能够改变某些人或某

些事物的信息,既包括使信息成为行动的基础的方式,也包括通过对信息的运用使某个个体(或机构)有能力进行改变或进行更为有效的行为的方式。

《现代汉语词典》对“知识”的定义是“人们在改造世界的实践中获得的认识和经验的总和”。有的学者综合以上说法,认为“知识是人们通过学习、发现及感悟所得到的对世界认识的总和,是人类经验的结晶”。

总结起来,我们给知识的定义为:知识是指人类对信息和客观事物规律的认识,它是人们在社会实践中积累起来的经验。人们对事物由表及里、由现象到本质、由感性到理性的认识深化,便形成了知识。知识是信息内容的组成部分。

## 2. 文献

1999年版《辞海》中把文献定义为“记录有知识的一切载体的统称”。1983年公布的《中华人民共和国国家标准·文献著录总则》(GB 3792.1—1983)把文献定义为“记录有知识的一切载体”。2012年的第6版《现代汉语词典》把文献定义为“有历史价值或参考价值的图书资料”。可见,凡是记录有知识的一切载体都可以称为“文献”。

文献由以下4个基本要素组成。

(1) 所记录的知识 and 信息,即文献的内容。

(2) 记录知识和信息的符号,文献中的知识和信息是借助于声音、文字、图表、图像等形式记录下来并为人们所感知的。

(3) 用于记录知识和信息的物质载体,如竹简、纸张、胶卷、胶片等,它们是文献的外在形式。

(4) 记录的方式或手段,如铸刻、书写、印刷、复制、录音、录像等,它们是知识、信息与载体的联系方式(摘自《中国大百科全书·图书馆学情报学档案学》)。文献具有认识、存储和交流知识的作用。

## 3. 信息、知识、文献的关系

信息包含了知识,知识是信息被认识的部分。知识可以分为主观知识和客观知识。信息经过人脑接收、选择、处理而形成并存在于人脑中的知识称为主观知识。主观知识借助语言符号,通过各种物质载体记录下来,就会变成可以传递的客观知识,即文献。

信息、知识之间的逻辑关系是包含与被包含的关系。知识是信息的一部分,文献是信息、知识的具体体现,它不仅是信息、知识的主要物质形式,也是读者吸收利用信息、知识的主要途径。

进入21世纪以来,科学技术发展迅速,人类社会的信息化、网络化进程也大大加快,各类信息数量剧增,随之而来的是新学科的不断出现和学科之间相互交叉与渗透的加快,使各专业信息发布分散而无规律。主要有以下几个方面的发展遇到了瓶颈。

(1) 非文本信息发展滞后。非文本信息(图像、音频、视频等多媒体信息)的检索技术、数字化技术、高密度存储技术为非文本信息提供了广阔的发展空间,多媒体信息已逐渐成为网络的主流。信息检索技术正在从传统的纯文本检索向超文本支持的非线性多媒体检索技术发展,然而图像、音频、视频的检索技术却处于萌芽阶段,需要高新技术支持并不断创新。

(2) 搜索引擎缺陷。分类目录搜索引擎采用人工干预技术,信息分类不规范,没有一个统一的控制词表和参照标准,分类目录差别较大;搜索范围较小,数据库更新慢,查询交叉类目时容易遗漏;如果用户检索请求没有对应的分类目录,则无法进行查找;信息遗漏不可避免,查全率低。

(3) 检索知识和技能匮乏。知识检索是一种全新的信息检索方式,是把用户请求与索引库匹配,寻找与请求关联的网页并返回排序的命中信息的过程。运用截词、词位限定、布尔逻辑运算等技术可以控制用户请求与数据库匹配的精度,但是信息检索难以避免丢失相关信息或产生大量冗余信息,即出现信息漏检与误检。信息检索效率是衡量信息检索效果的重要指标,是检验信息检索技术成熟与否的标准。知识是信息加工与序化的产物,是高浓度的有序化的信息;知识检索必然是高层次的信息检索。

## 1.2 信息检索的发展趋势

近年来信息检索的面貌大大得到改观,具体表现在:检索行为大众化,检索行为日常化,检索趋于经济化(时间和费用)。总的来说,信息检索工具已开始发展检索以外的多样化的信息服务,以便能为用户提供快捷、准确、全面的服务。下面对其未来可能的发展方向稍做介绍。

### 1.2.1 信息检索专业化

面对网上五花八门、包罗万象的信息,很难有一个完备的检索或分类体系将其高度序化,所以为了能为用户提供针对性更强、专指度更高、更经济、更快速地基于专业内容的信息服务,专业性的信息检索实属人心所向。由于每学科专业都有自己独特的词汇和用语,特定的信息检索工具应该使用与之相应的标引和检索语言,而这一点正是包罗万象的综合性信息检索工具难以做到的。因此,信息检索工具不能真正地追求大收录及大标引量,应更加注重突出专业特色,提供对一些专业性、学术性成果较深入的核心数据库的访问。如面向某个学科领域的专业性搜索引擎或称垂直性门户网站的网络信息检索工具正在研制。另外,检索网站通过提供更多经加工、编辑、评价、筛选的深层次信息内容来吸引用户。

### 1.2.2 受控语言的引入与使用

一般地说,自然语言检索由于其灵活性,有助于检全率的提高。与之相应,受控语言检索由于其规范性和准确性,有利于检准率的提高。目前,除少数情形(如,Online Computer Library Center的NetFirst)之外,其他主要的网络信息检索工具都采用自然语言标引和检索,其结果必然是同义词和近义词得不到控制,词之间关系得不到揭示,最终导致检索结果过多过杂的现象。在利用因特网检索软件搜索信息时,用户往往以一个或几个检索词作为检索入口获得一定量的主页地址,再以这些主页作为入口开始在网上漫游。至于自己所用的检索词是否贴切用户却毫无把握。具体做法为建立同义词典。用户在提出一个检索词后可获得一批候选词,然后判断,用户可以选择其中一部分或全部作为检索用词,大大提高信息检索的检全率。

当然,绝大多数网络信息检索工具不采用受控语言有其原因,这主要包括网络信息数量大、变化快和涉及面广,现有的受控语言难以适应网络信息标引和检索的需要。另外,面对如此庞大的网络信息,检索工具有无必要投入大量的人力使用受控语言标引它们呢?尽管人们没有对网络环境下检索语言类型问题展开全面系统的讨论,可是,网络上已经出现了使用受控语言检索的工具。例如,隶属于OCLC的NetFirst采用了《美国国会图书馆主题词表》

和《杜威十进分类法》类分网络信息。由此可见，受控语言已开始涉足网络信息检索以提高网络信息的检索效果。

### 1.2.3 信息检索的智能化

以往，人们总是把改善检索效果的希望寄托于提高信息标引质量和改进检索机制，而收效并不显著。为此，一些研究者已陆续提出了改善网络信息检索效果的其他方法，其中包括智能检索软件（Intelligent Agent）和自动化数字“图书馆员”（Automatic Digital Librarians）。这些途径的共同点是把改善检索效果的着重点从网络信息检索工具转向某种“中间件”或“智能代理”。虽然这种“智能代理”带有理想色彩，但也并非空中楼阁，事实上，构成这种“智能代理”的部分要素已在一些现有的网络信息检索工具中实施。如 Ask Jeeves 和 Inquizit 都能把用户的自然语言提问自动转换为检索提问，用户可用“Why is the sky blue?”作为检索提问，而不必再考虑检索词的选择问题。同时，智能化的自动索引软件可以对搜集网页的内容相关性及质量加以判断。智能化的检索软件可以自动分析用户提问并形成恰当的检索策略等。总之，随着网络用户对检索的精度、效率要求的不断提高，网络信息检索必须重视提高其在检索功能和服务上的智能化程度。

### 1.2.4 记录检索路径和内容提示信息

当检索者以一个节点作为入口检索信息时，应采用特殊符号标记记录，保存用户浏览路径及其邻近节点，并根据用户要求保存节点信息概要和片段，帮助用户及时调整检索方向，避免偏离目标时重新游历一遍。具体做法，设置一个动态存储器，在检索过程中，其存储内容是动态变化的。开始检索时，动态存储器是空的，随着检索的进行，系统不断把检索问题的出世状态及新的知识，推理路径及相关描述存入动态存储器。查询检索路径时可根据用户需要限制其存储器的各项指标，例如，节点上下各几步，或以何种方式（检索地图、文本等）输出。目前，Notecards 系统提供了一个记录表来显示用户浏览路径及相关节点。

### 1.2.5 使用语义-文献双层数据结构

目前有一些基于超文本环境的联机检索系统，既实现了信息检索的灵活性和方便性，又保证了较高的检索质量。因特网检索工具可以借鉴这类机检系统的成功经验，模仿其数据结构，建立一种类似的符合自身特点的语义-文献双层数据结构。

语义信息和语义链集合构成了双层模型的第一层（语义描述网络）；文献信息和结构链构成了另外一层（实体网络）；连接链则贯穿于两层之间，在双层结构的层面切换时发挥作用。语义层由描述语义概念的词汇构成。当在语义层中加入新的描述词汇时，检索软件会自动建立新词汇和文献层中各单元的对对应关系。同样，当文献层中加入新文献时，系统也会建立起信息单元和语义结构的联系。语义层的词汇集合由抽出词（以自动抽词算法从主页文本中抽出），索引词（检索系统采用的系列索引词汇）和用户词（检索过程中所用的检索词）组成。

检索时，用户通过输入的检索词利用语义链方便地看到相关用词并沿着连接链进入文献层查找目标文献。当检索词不在索引词集合中时，可以先进入符合要求的文献层，然后顺着连接链到达索引词集合，继而利用语义链找到相关用词后再通过连接链进入文献层查找目标文献。

总之，能很好地实现语义描述网络与实体网络之间的互动。

## 1.2.6 面向用户的人性化服务

新世纪科技创新强调“以人为本”，网络信息检索也不例外，先进的、人性化的可视化检索界面无疑会给用户带来检索的效率和心理的愉悦。例如，有的检索系统以三维图来显示检索结果。友好先进的结果提供方式直接影响用户对信息的吸收和利用，NorthernLight 和 Excite 可以将检索结果分组汇集，并在网页右上边显示不同组名。同时，网络检索工具对检索结果的筛选（过滤，屏蔽等）也是搜索引擎技术发展的重要方面。

## 1.3 信息检索的意义和作用

### 1.3.1 信息检索的意义

#### 1. 有利于减少课题的重复研究，提高科研成功率

任何科学研究都是在继承前人知识的基础上有所发明、有所创新的。也就是说，每个人都把前人认识事物的终点作为继承探索的起点。任何人从事某一特定领域的学术活动，或开始做一项新的科研工作，都要花费大量的时间，对有关文献进行全面的调查研究，摸清国内外是否有人做过或者正在做同样的工作，取得了一些什么成果，尚存在什么问题，以便借鉴、改进和部署自己的工作。只有这样方能做到胸中有数，才能有所发现、有所创新、有所前进，否则容易造成重复劳动，导致人力、物力、财力的浪费。

我们知道，科学技术的发展具有连续性和继承性，闭门造车只会重复别人的劳动或走弯路。在研究工作中，任何一个课题从选题、试验直到出成果，每一个环节都离不开信息。研究人员在选题开始就必须进行信息检索，了解别人在该项目上已经做了哪些工作，哪些工作目前正在做，谁在做，进展情况如何等。这样，用户就可以在他人研究的基础上进行再创造，从而避免重复研究，少走或不走弯路。

#### 2. 有助于节约时间，提高科研效率

随着科学技术的发展，文献数量在剧增并且学科间相互渗透。科研人员在进行一项科研活动中，查找资料占了大量时间。据美国和日本 20 世纪 80 年代的一项统计。科学工作者在从事科研活动中所花的时间为，试验研究占 32.1%，计划、思考占 7.7%，数据处理占 9.3%，查找情报资料占 50.9%，如果熟悉文献检索方法，就能大大节省查找资料的时间，从而加快科研速度，早出科研成果。

在当今世界，提高科研效率，加快科研速度的意义还在于使相同科研课题在国内外竞争中处于有利位置。专利法规定，对相同的发明成果，按先申请原则授予专利权。即只授予第一个申请人专利权，其后申请的发明作为已知技术处理。显然，如果忽视科研速度，即使科研获得了成功，但由于发明失去了时间的新颖性，也会变成无效劳动，给国家带来损失。

#### 3. 有助于协助管理者正确决策

准确、可靠和及时的信息，是正确决策的基础。在竞争激烈的今天，如果不能通过在阅文献，获得国内外有关本行业的发展动态，做到知己知彼，那么无论经营何种产业都如同“盲人

骑瞎马，夜半临深池”一样危险，很难有成功的希望。改革开放之初，我国由于与国外经济交往少，信息不灵，盲目重复引进，一些引进技术和设备不适用，造成惊人的损失和浪费，这方面的教训是很惨痛的。

#### 4. 有助于增强知识积累，改善知识结构，提高自身素质

英国情报学家布鲁克斯曾提出关于情报与知识的基本方程： $K(S)+aI-K(S+AS)$ 。公式中， $K(S)$  为原有的知识结构， $aI$  为信息增量， $K(S+AS)$  为新的知识结构。该公式表明，新的知识结构是随着吸收信息量而增加的，而吸收信息量又取决于原有的知识结构。据美国工程教育协会统计，美国大学毕业的科技人员所具有的知识，只有 12.5% 是在大学阶段获取的，而 87.5% 则来自工作实践。大学毕业后，5 年内不补充新知识，原有 50% 的知识会失效；10 年不补充，原有 100% 的知识会失效。

掌握了信息检索的方法和技能，就会找到一条吸收和利用大量新知识的捷径，以最少的时间和精力，继承前人的知识，最大限度地增加知识积累，改善知识结构，不断提高自身素质，在激烈的竞争环境中立于不败之地。

#### 5. 有助培养信息意识，提高信息素质

信息素质 (Information Quality) 或信息素养 (Information Literacy) 实质上是一个人和信息搜集、整理、筛选、判断、评价和利用方面的能力，是个人综合素质的一个方面，信息检索课是对大学生进行信息素质教育的主要形式之一。信息素质教育内容主要由信息意识教育、信息道德与信息法规教育、信息能力教育等组成。其中，信息意识教育主要培养大学生对信息的敏感度，或检索、分析、判断和吸收信息的自觉程度，信息道德和信息法规教育。防止信息垃圾和信息污染，不制作、不传播、不使用不良信息，不借助网络进行人身攻击，不侵犯他人的知识产权、隐私权，不利用信息技术进行违法犯罪活动内容。

### 1.3.2 信息检索的作用

#### 1. 信息检索有助于知识更新

随着科学技术的飞速发展，知识老化现象也不断加重。据我国有关部门的典型调查，20 世纪 70 年代的大学毕业生，5 年后有 45% 的知识老化，10 年后有 75% 的知识老化。由于科技发展越来越快，所以当今毕业生的知识老化速度也在加快。只有不断自学、进修，才能适应迅速发展变化的信息时代的要求。

知识更新，重要的是拓宽知识面。近几年美国学术界认为，有必要建立一种通才学，就是要求学生有较宽的知识面，不仅要掌握理工科渗透的知识，还要了解文科渗透的知识。这是因为，科技发展使得人类社会生产的产业结构正处在急剧变化之中，大批知识密集型工业相继涌现，交叉科学大量出现，现在的高校毕业生知识面过于狭窄，已不适应飞速发展的新形势。

尽早掌握信息检索的本领，就会在未来的竞争中取得更大的主动权。

#### 2. 信息检索有助于发展教育

1) 教育的重要性越来越突出。在过去的农业社会和工业社会中，教育的作用虽然显著，但没有文化或文化不高也能够从事劳动生产。

随着信息社会的到来，教育的重要性也越来越突出。信息社会是知识密集的社会，各种现代化技术将会渗透到所有行业和部门，不经过系统教育就不能胜任现代化设备仪器的操作。从这个意义上说，新技术革命也可称作“知识革命”，由此带来的信息时代也应称为“知识时代”，



这就对当代教育提出了更高的要求。

2) 必须把教育放在优先位置。国家要强大, 必须科技领先, 而要重视科技事业, 必须优先发展教育。当今世界, 谁要充当“领头羊”, 就必须把加强教育放在最优先的位置。

美国前总统克林顿多次强调, 每个美国人必须终身学习。在美国政府 1996 年提出的“教育技术规划”(Educational Technology Initiative) 纲领中指出: 在 2000 年, 全美国的所有教室和图书馆都将联上信息高速公路, 让每个孩子都能在 21 世纪的技术文化中受到最好的教育。

3) 信息检索有助于终身教育。我国教育部提出, 面临 21 世纪知识经济的挑战, 必须加快我国教育信息化的步伐, 根据各地区经济发展不平衡的现实, 分 3 个层次推进信息化教育。

(1) 以计算机多媒体为核心的教育技术在学校的普及与运用;

(2) 组织学校上网, 利用网上资源;

(3) 开办远程教育, 提供广泛的学习资源, 不断满足社会终身教育的需求。

目前, 我国正投入巨资建设远程教育网络和拓宽教育主干网。如果掌握了从网上获取信息的本领, 就会受益无穷。

### 3. 信息检索有助于科学研究

科技文献中记载着前人的劳动成果, 可以向后人提供借鉴; 此外, 科技文献中又记载着同代人的生产和科研的成果, 可以提供参考。通过利用文献, 可以避免别人的重复劳动, 提高科研的速度和效益。古今中外一切有成就的科学家, 都是在广泛吸收前人和同代人知识的基础上, 受到启发而取得成功的。正如牛顿所说: “如果我比笛卡尔看得远些, 那是因为我站在巨人的肩上的缘故。”

### 4. 信息检索有助于跟踪学术最新动态

有一位大学本科生在自己的毕业论文及答辩中提出了很多新的观点、新的设想, 令答辩组的专家、教授不知所措, 因为这些新观点是他们从未听说过的。原来, 这位大学本科生在校的几年中, 一直利用图书馆和互联网关注着这项研究。追踪着这项研究的世界最新动向, 这位大学本科生的知识更新已远超周围的老师和同学。

### 5. 信息检索有助于节省科研时间, 提高工作效率

据统计, 科研人员查找信息资料的时间, 一般要占全部科研活动时间的 35%~40%。掌握科学的信息检索方法可以节省科研人员查阅信息的时间, 为科研工作节省大量的宝贵时间。科学研究是一种创造性劳动, 兼有连续性和继承性特点。对于任何一个科技工作者来说, 系统地掌握国内外科技信息, 了解科技发展水平与动向, 利用已有的研究成果, 避免重复他人的劳动, 少走弯路, 具有重要的现实意义。

### 6. 信息检索有助于提高信息素质

1974 年, 美国信息工程协会主席 Paul Zurkowski 首次定义信息素质, 即它是利用大量的信息工具及主要信息源使问题得到解答的技术和技能, 具备信息素质的人, 能够识别何时需要信息, 知道如何查找、评估和有效利用需要的信息来解决实际问题或者做出决策, 无论其选择的信息是来自于计算机、图书馆、政府机构、电影, 还是其他任何可能的来源。通过信息检索技术的学习, 应用科学的方法, 培养科学品德和精神, 都能够解决信息检索过程中遇到的问题。

### 7. 信息检索有助于高校图书馆信息化建设

1) 高校图书馆功能简介。高校图书馆是高等院校的文献信息中心, 是为教学、科研服务的学术性机构, 是高校的三大“支柱”之一。其主要职能是文献收集、资料加工、借阅服务、读者教育、环境熏陶。图书馆在为广大师生提供各项服务的同时, 也为广大师生提供了安静、幽雅的学习环境。因此, 高校图书馆的设置对大学生的功课学习、课外自学、陶冶情操和健康

成长都具有至关重要的作用。

2) 如何打开“知识宝库”。图书馆历来被称为人类的“知识宝库”和“科学殿堂”，高校图书馆的藏书往往数以万计、十万计，甚至百万计，各种报刊资料更是品种繁多。近年来，电子出版物也纷纷登场，特别是“盘”家族（硬盘、软盘、光盘），已经显示出强大的生命力，正在迅速占领传统出版物市场，在图书馆馆藏中所占比例也在迅速提高。

为了使各种文献整齐、有序，便于读者借阅、查找，图书馆人员把它们进行了科学的加工、分类和编排，并编制了多种目录、索引和文摘等检索工具供师生使用。广大师生把这些检索工具形象地称为打开“知识宝库”的“钥匙”。可以查找图书、期刊；查找论文、报告；查找字词、数据等。

3) 促进高校图书馆的数字化建设。1993年以来，发达国家投入了大量资金，开发、研制数字图书馆。近几年，国内开始关注、跟踪国际数字图书馆的发展。

数字化图书馆就是图书馆藏信息实现数字化管理，并且上网服务，供读者随时随地查阅。与传统图书馆藏书不同的是，数字图书馆中的众多图书，不再是孤立地散布于世界各地的图书馆中，而是永久性地存储在硬盘、软盘、光盘之中，或流动在全球信息网络上，成为人类共享的知识财富。

我国有着悠久的历史，图书馆文献资源十分丰富，要实现数字化，工作量巨大，超出一般人们的想象，这是最大的困难。为了加快图书馆的数字化进程，1997年9月，由IBM倡议，北京大学、清华大学、北京大学图书馆、香港、台湾，以及韩国、日本的17所高校图书馆为发起成员，在北京成立了“亚太数字图书馆论坛”。其宗旨是推进数字图书馆技术和标准在各大学、博物馆和文化收藏机构的应用。数字图书馆建设是将包括多媒体在内的各种信息的数字化、存储、管理、查询和发布集成在一起，使这些信息在网络上传播，以得到最大限度的应用。其中，北京大学投资100万美元，用于数字图书馆建设；上海交通大学等院校的图书馆数字化工作也在紧锣密鼓地进行。

目前，国内生产文献信息数据库的最大部门是中国科技信息所，中国科技信息所的“万方”数据库系列光盘在国内首屈一指。

4) 高校图书馆的新职能。由于图书馆是重要的文献信息源，所以各国都把图书馆的数字化、网络化作为建设重点。近几年，国内高校图书馆加快了现代化建设的步伐，计算机、复印机、语音设备、网络器件等不断增加，受到了许多部门的关注。因此，充分利用高校图书馆的现代化设备和电子文献，对大学生进行能力培养和全面教育是时代对高校图书馆提出的新要求。

目前，北京大学图书馆的网页上设有“网上教室”，清华大学图书馆设有“Internet教室”等网站，以充分发挥高校图书馆的教育职能，开展网上教育。

5) 高校图书馆的网络化。信息网络是由各种专业应用领域的信息系统和公用高速通信网络平台组成的。美国的公用高速通信网是一个以超大容量的光纤传输网络为骨干，以高性能计算机为枢纽，能交互传输和交换语言、图像和数据，拥有多媒体终端，其信道速率可达Gb/s级的宽带、高速和综合智能通信网络。公用高速通信网在信息网中承担信息传输和交换的任务，是信息网的中枢神经系统。

计算机网络实际上是以共享硬件、软件、数据等资源为目的而连接起来的各自具有独立功能的计算机系统的集合。它是在网络协议控制下，利用各种通信手段，把地理上分散的计算机有机地连接在一起，达到相互通信而且共享软件、硬件、数据等资源的计算机复合系统。

计算机联网的主要目的在于共享资源。计算机联网后发展了分布式数据处理和分布式数据库。在获得数据和需要进行数据处理的地点设置计算机，把数据处理的功能分散到各台计算机上，可利用计算机网络实现分布处理和建立性能优良、可靠性高的分布式数据库系统。

我国目前建有四大网络：中国科学院系统的中国科技网（CSTNET）、中国教育系统的中国教育研究网（CERNET）、1995年由邮电部门主建及经营管理的中国公众计算机网（ChinaNET）和中国金桥网（ChinaGBN，也称为中国国家公用经济信息通信网）。此外，还有中国联合网络通信有限公司的骨干网，和中国移动互联网络的骨干网。

对于高校图书馆的上网统计，目前CERNET已有300余所高校图书馆的信息。CERNET上的中国博士论文全文数据库、最新工程类期刊、会议录文献报道服务数据库等。CERNET的最大特点是比较重视网络信息的组织和利用，文献信息服务是CERNET的重要服务项目。

我国的高校有1900多所，高校图书馆的科技教育类信息源与其他类型图书馆相比占有优势。例如，江苏省高校文献信息网络建设工作的研究报告表明，高校馆藏的科技文献数量占全省的50.4%。显然，就CERNET的信息资源来看，高校图书馆网络无疑是校园网和地区教育科研网上最大、最重要的资源子网。近几年，高校图书馆的现代化速度明显加快。

### 思考题

1. 什么是信息、知识和文献？
2. 按文献载体和记录形式划分，信息资源分为哪几类，并简述它们之间的差异。
3. 简述信息检索的意义和作用。