

学校的理想装备

电子图书·学校专集

校园网上的最佳资源

中小学信息科学知识

信息学方法论



## 第一章 导论

### 第一节 信息学方法论基础内容

信息学作为一门新兴的交叉边缘学科，经过近 50 年的发展，取得了令人瞩目的成就，其独立的学科地位也得到了广泛的承认。然而，难以否认的是，这门年轻的学科离一门成熟的精密学科还有距离，它不仅表现在缺乏统一规范的理论体系，也表现在其研究方法的缺陷。科学史表明，一门学科如果要成为真正独立的、成功的科学学科，就必须有自己独特的理论体系，还必须有本学科特有的研究方法或表征其存在意义的特征方法存在。因此，信息学方法论的建立和研究已成为信息学进一步完善和发展的重要课题之一，而且在理论上已形成了一套较完整的体系结构，其应用也愈加广泛和深入。

#### 一、基本概念和原理

什么是方法论？《中国大百科全书》规定其含义为关于认识世界和改造世界的方法的理论。方法论有哲学方法、一般科学方法、具体科学方法之分。哲学方法论是关于认识世界、改造世界、探索实现主观世界与客观世界相一致的最一般的方法理论；一般科学方法论是研究各门具体学科，带有一定普遍意义，适用于许多有关领域的方法理论；具体科学方法论是指研究某一具体学科，涉及某一具体领域的方法理论，三者之间的关系是互相依存、互相影响、互相补充的对立统一关系。

信息学方法论则是关于信息学方法的学说与理论。它主要研究信息学认识活动中所运用的所有研究方法的结构、功能与特点，阐述这些方法的应用、发展规律和方向，以及各种方法之间的相互关系，属于具体科学方法论。信息学方法论的核心是信息学方法，信息学方法论应从整体上把握信息学认识活动。

信息学在建立研究信息学方法论的一些基本概念的基础上，通过科学判断这一思维形式表达了信息学方法的一些基本规律和基本关系，形成了一系列关于信息学方法论的基本原理或基本定律。但由于情报学是一门由社会科学和自然科学交叉而成的综合性学科，其方法有的源于一般科学，如系统科学、数学、经验科学、思维科学等，有的是信息学专有的方法，如信息计量学方法，引文分析方法、信息整序方法等；其方法的总体又受哲学的宏观指导，因此，信息学方法论的一些基本原理或基本定律一方面直接来自于源学科，一方面是信息学家将引入方法的原理、定律融合于信息学的认识研究过程中，深化综合而成。例如，系统科学方法中主要有系统论、信息论、控制论、协同论、突变论、耗散结构论等；数学方法中主要有集合论、对数论、群论、模糊理论等；思维科学方法中有抽象与具体、比较与分类、归纳与演绎等；经济学方法中有投入产出分析理论、边际效益理论等；信息学方法中有信息检索原理、布拉德福定律、洛特卡定律、乔普夫定律、文献增长规律等。上述这些方法的基本原理或基本定律构成了信息学方法论的基本原理或基本定律。

#### 二、信息学方法论的形成与发展

信息学方法论的建立和形成有其必然的条件，主要表现在以下几个方面。

1. 以信息学学科本身来看，要正确深入地研究信息学这一科学认识活动

的本质及内在形成发展机制、建立理论体系及动态发展模型、研究开发信息技术及应用，不仅需要建立各种信息学专门方法，引入各种哲学方法及一般科学方法，而且还要深入研究这些方法的产生、演变、性质、特征、适用范围以及这些方法之间的逻辑结构、相互关系、共同性及差异性，以便使现有的各种信息学方法在信息学研究活动中得到合理的运用，发挥最佳功能。

2.从信息学方法的发展史来看，尽管信息学方法是与信息学这一科学认识活动同生共长、相互依存，并随其发展而发展的，但信息学方法论却是由信息学方法的不断创造、积累发展而形成的。本世纪50年代中期至60年代末期，是信息学面向实用研究的阶段，信息学主要研究信息的手工存贮与检索问题，研究方法主要为经验描述、定性分析；70年代是信息学面向信息技术与应用的研究阶段，信息学将数学方法应用于计算机检索和文献计量分析中，形成了以定量分析为主要带头方法的研究阶段；80年代至现在，在深入研究、发展信息技术与应用的同时，又重新重视信息学基础理论的研究。因此，信息学逐渐把系统科学方法引入其研究之中，并将已有的大量的信息学方法进行分析、总结，着重研究信息学方法的结构、性质、特征、适用范围及相互关系等，建立了信息学方法论，以满足信息学既需定性分析又需定量分析的整体综合研究的要求。

3.从科学理论的构成要素来看，信息学方法论形成发展至今，已有了关于其方法论的基础概念、基本原理或基本定律以及逻辑推论，尽管信息学对这三个要素的研究深度还不够，但已形成的三个构成要素足以证明信息学完全有条件建立自己的信息学方法论。

在信息学方法论的建立与研究初期，由于研究队伍与人员知识结构、思维定势等方面的局限，信息学研究中描述性的、经验性的、随意性的研究方法曾经占有相当大的比例并居于主导地位。现在，信息学方法论已发生了重大变革，更严谨、更科学的方法及其应用已经形成了一定的规模，其具体发展体现在：

(1)研究方法上，从单一性逐步转向整体性。以往信息学传统的经验方法，多侧重于孤立研究信息工作流程中各环节的状态，如文献信息的管理、信息工作的服务方法等，而较少把信息、信息工作者、信息用户联系成一个系统加以考虑。后来人们发现整体研究不会人为割断信息系统固有的内在联系，而且可以开拓思路，提高信息服务效果，发挥信息系统的整体功能和价值，所以这种整体性研究方法越来越受到重视。

(2)分析方法上，信息过程从定性描述逐步过渡到定量描述。研究信息系统整体功能和价值，必须运用数量分析方法。以往信息学研究多运用静态分析方法，处于“用图表来加以说明的阶段”，即使有一些统计数字描述，也仅仅是初级的，一般没有超出算术统计的范围。近几年来，利用数量分析方法开展信息过程及信息学的过程的研究逐渐增多，主要反映在以下几个方面：

运用数学原理和统计学方法，对有关数据进行数量分析，如探讨文献统计规律的理论，引文分析量化研究，信息工作的经济效益分析，文献管理的优化问题，以及信息用户调查研究等。

建立数学模型，对过程进行模拟研究，如从抽象化角度对某些信息学概念、信息过程分析运用了数学描述，信息检索理论中运用数理方法建立检索数学模型；信息分析预测中运用回归分析、模型法开展一些定量研究。

利用电子计算机系统收集、整理、存储各类信息学研究数据、资料。运用数理方法论证、修改已有的信息学定律、公式、模型等。

(3) 研究方式上,逐步从个体分散的业余研究转向集体有组织、有计划的研究和个体业余研究相结合。有关信息学的理论刊物在数量上和质量上都有了很大的提高,它们从不同侧面研究实际或理论问题。一些信息机构还根据需求和可能设置了信息理论与方法的专职研究室,在理论研究领域中起骨干和先导作用。

此外,信息学与其他相关学科的联系越来越广泛,从而拓宽了信息学的研究领域。一方面,信息理论研究人员广泛运用其他相关学科的成果和方法开展研究;另一方面,采取多学科合作研究的方式,聘请相关学科的学者、专家参加,出现了集体有组织、有计划的研究方式,使信息学理论研究进入了新的层次。

(4) 研究对象上,逐步从提供信息的系统本身发展到利用信息的用户上。近年来,信息用户研究已开始成为信息学研究的热门课题,内容包括信息用户及需求,信息服务效果评价,用户心理学、信息搜集策略等,提出要改善利用信息的环境,从信息政策的高度加以重视。

### 三、信息学方法论的研究内容

信息学方法论是信息学理论体系的有机组成部分,其根本目的在于提高信息学研究方法使用的合理性、有效性,并为信息学特征方法的建立、完善提供指导。为此,信息学方法的研究应该覆盖以下内容:

1. 信息学方法论的一般问题。其主要目标是从理论的高度提炼出信息学研究的方法要素,在此基础上说明理论与方法论、问题与方法之间的辩证关系。同时,对信息学方法的来源、发展规律与趋势以及信息方法体系的动态结构进行总体把握,从而分析构建信息学特征方法的规律与原则。

2. 信息学研究的总体认识。任何一门学科的方法都受到研究者的价值判断、思维习惯、知识结构等主观因素的影响。因此,即使是在同一领域里,研究方法也表现出较大的差异。但是,研究方法也要受研究对象及研究目标等客观因素的影响。所以,对信息学研究对象的特点、研究的目标及其组织管理从总体上加以把握是信息学方法论研究的重要内容。

3. 信息学方法的个别研究,包括各种方法的性能、特点、优缺点、适用范围、使用程序与技巧、使用科学性的评价标准似及具体研究中方法选择的影响因素等。

4. 各种方法在信息学研究中的具体应用情况,通过实例分析展示被应用方法的特征和应用效果。

5. 信息学方法论的研究原则和信息学研究中各种方法使用的选择原则。

总之,信息学方法论不仅要从微观上对各种方法的诸多方面加以研究,而且还要从宏观上研究这些方法构成的体系及信息学认识活动的特点与规律。

## 第二节 信息学方法论的特点和作用

### 一、信息学方法论的特点和功能

作为信息学认识活动与研究方法的知识体系,信息学方法论具有以下特

点：

1.从其内容上看，信息学方法论具有概括性与抽象性。信息学方法论从如何使信息学研究合理化的角度来促使自发经验研究向自觉系统研究的演进，因而具有比个人经验更大的抽象性与概括性。

2.从其结构来看，信息学方法论具有层次性与系统性。在各种方法之间建立起现实的联系与从属关系，因而信息学方法论具有系统性。在信息学方法体系中，各种方法又具有层次之分。

3.从其存在形式来看，信息学方法论具有广泛性。方法论知识不仅以独立的理论形态存在，体现为各种规则、方法和具体作法，而且直接渗透在与信息学研究的内容相联系的认识活动之中，因而具有广泛性。

4.从其作用方式来看，信息学方法论具有间接性。信息学方法论的原则原理是在方法的运用中得到具体反映的，因此，其作用是间接的。首先，研究者研究什么并不是方法论能够规定的；其次，信息学方法论虽能为研究者建立决策的方案，但方案的选择与实施受研究者主观因素的影响。

5.从其发展来看，信息学方法论具有动态性。信息学通过移植其它学科的方法，创造自己的专门方法来丰富其方法体系，而且，各种方法在方法体系中的地位也将随信息学研究的重点、热点与难点的变迁而变迁。从这个意义上来说，信息学方法论是一个时间上有序的、开放的、动态的知识体系。

信息学方法论的特点与内容决定着它的功能是多方面的。信息学方法论作为信息学方法的理论指导和规范，能深刻地阐明信息学形成发展的内在机制；能清楚地说明信息学研究者是如何思维、如何研究的，并描述其思维过程及研究过程；能科学地将信息学的工作方式、行为方式和思维方式的共性抽象出来。形成科学的规范和准则，以促使信息学研究沿此模式进行；能及时发现问題，提出课题，实现研究目标。

因此，信息学方法论具有如下功能。

(1) 信息学方法论具有对方法选择和使用的规范功能。信息学方法论通过制定具体研究之中方法选择的一般原则以及方法使用的规则，为研究者规范化选择、使用方法提供依据和指导，从而使信息学研究趋于严谨与科学。这对于大量引进其它学科方法的信息学研究尤其具有重大的意义，是解决目前信息学研究方法选择与使用上存在缺陷的根本途径。

(2) 信息学方法论具有对方法创造的启发功能。信息学方法论通过对信息学专门方法形成与发展的机制和规律的研究，确立建立专门方法的一般原则，从而启发和指导信息学专门方法的建立与发展。需要强调的是，信息学方法论本身并不具备创造方法的功能，因为任何一门学科的方法的创造并不是产生于方法论，而是决定着方法论。信息学方法的创造是在信息学研究实践中完成的。

(3) 信息学方法论具有完善信息学理论的功能。信息学方法论本身是信息学理论体系的有机组成部分，它的发展将促进信息学理论体系的完善。另一方面，信息学方法论能够从整体上认识信息学理论体系及其运动规律，从而为信息学研究提供正确的思维方式，促进信息学理论研究的深入与理论体系的完善。

(4) 信息学方法论具有认识表达功能，它解决如何将正确的认识通过正确的方式表达出来的问题。由于信息学主要是由经验科学发展而来的，因而其表达方式从叙述个别的经验事实材料开始，进而叙述信息学领域里的结

论和思想，叙述由它们加以论证的公理、引理和个别结论。

(5) 信息学方法论具有论证、验证功能。因为任何一种理论学说不仅要有科学的思维方法、严密的逻辑方法加以论证，还要有科学的观察实验和社会实践加以检验才能判别其真理性。信息学方法论能够解决信息学的研究结果及信息学理论学说的论证和验证问题。

此外，信息学方法论通过制订方法，使用科学性判定标准，发挥评价功能，通过把握信息学研究的发展规律，预见信息学未来研究方法，它还具有对信息学方法的现实批判反思功能。

## 二、信息学方法论的研究原则

信息学方法论的建立和研究，在与人类认识自然和社会的总体方法相联系的同时，还必须确立自己的特点和范围。因此在讨论信息学方法论体系时，应当遵循六项原则。

### (一) 过程和方法相结合

信息研究，同任何科学研究一样，从开始到结束，都是一个连续进行的动态过程，整个过程由若干环节构成，不同环节采用不同的研究方法。以往较多的注重某一环节的研究方法，而较少考虑整个过程的研究方法是失之偏颇的。实际上，科学研究方法通常是指获得科学知识应遵循的程序，研究问题、解决问题的方法和手段的总和，因此应当将方法和程序有机地结合起来。

信息研究工作一般由两个相互联系又相立区别的部分构成，即常规信息研究工作和专题信息研究工作。常规信息研究产生动态性和基础性信息，是专题信息研究的基础，为专题信息研究提供丰富的材料，而专题信息研究则产生专题性信息，是常规信息研究的发展，并深化了常规信息研究工作。常规信息研究的基本环节是扫描、分析和积累，而专题研究的基本程序是选题、调查、整序、抽象、表达和评价。可以说这两种类型的信息研究及其程序构成了信息研究工作的基本模式和格局。因此，以这些程序及其方法为基线，是从事信息方法论研究应当遵循的重要原则。

### (二) 一般和特殊相结合

任何学科都有适用于它的哲学方法、一般方法和专用方法。大体来说，辩证唯物主义是人们认识世界和改造世界的基本方法，它对科学研究方法论有着巨大的指导作用；观察方法。实验方法、逻辑方法、系统方法则是自然科学研究的一般方法，而通过研究物质的化学变化来研究物质运动规律的一系列方法乃是化学学科所独有的方法。

在信息方法论的研究中，应当把吸收和独创相结合。通过对信息的选择、整序和抽象来研究科学技术问题，是信息研究的特色。通过信息检索获取信息的方法，以分类和主题为基础的整序方法，以文献计量为基础的信息计量方法和引文分析方法都是信息学的特有方法。如何发展信息学的特有方法是信息学方法论研究的重要任务。

吸收和移植是各种学科及其方法论发展的普遍现象，信息学也不例外。吸收和移植的基本思想是将一般方法的普遍性与信息学研究的特殊性相结合，逐步做到融化与同化。在吸收与移植过程中，一旦证明这种方法适用于信息学研究，它也应当成为信息学研究的一种方法，成为信息学方法体系中的有机组成。

信息学方法论的研究不仅仅局限于科技信息研究领域，还应逐步汇集和探索其它领域信息研究的方法，建立起包括经济、军事等在内的整个信息研

究工作的方法论体系。

### （三）理论和应用相结合

在信息学方法论的建设过程中，应当十分注意理论与应用相结合，两者不可偏废。方法论既然是关于方法的学说或体系，那么对方法理论的学习无疑是十分重要的。现代横断学科实质上都具有方法论的意义，正在迅速应用到自然科学和社会科学的各个领域，推动着这些学科的进步和发展。但学习的目的在于应用，从某种意义上讲，方法与实践的关系，有如游泳理论与游泳实践的关系，因为科学研究方法的掌握毕竟具有某种技巧的性质。

回顾我国科技信息工作，在引进国外先进学术思想和科学方法上作出了重要的贡献，但如何在信息领域中加以推广和应用，认识不足，措施不力，效果不佳。因此在信息学方法论的建设过程中，应当花大力气研究各种方法在信息研究活动中的应用特点、应用范围、应用步骤、典型案例、改进办法，解决好方法的适用性和实用性问题是极其重要的。

### （四）历史和现实相结合

从根本上看，应该把方法看作是人类认识世界、改造世界的经验总结和理论概括，是科学技术和生产力发展水平的标志。人们的方法水平总是与人们的知识水平相适应的。科学史表明，科学与方法同生共长，形影相随。信息学方法论的建设既要受信息研究人员的知识结构和实践范围的制约，也会受到现代科学技术和社会进步的推动，总不会停留在一个水平上。它也需要经历一个发展的过程，以便把我国信息工作的历史和现代科学技术发展的现实很好地衔接和填补起来，在总结和继承的基础上改进和发展，既不能熟视无睹，又不能操之过急。

历史上任何方法论的建立都经历了一个从零星的、局部的、不系统的探索到大量的、全面的、系统的发展过程。在信息学研究领域中，关于研究模式的探讨，也经历了相当长的时期，现在基本形成共识，即信息学研究的合理做法应当是定性研究与定量研究相结合，自行研究与组织研究相结合，基础研究与应用研究相结合，科技信息与经济信息相结合，国外信息研究与国内信息研究相结合，这就是一个发展的过程。

而作为信息学特征方法的信息计量学方法与引文分析方法，虽然已经显示了其在信息学研究中的突出特点和独特作用，但因为对它们的研究和应用的历史并不太长，所以，在进一步应用和研究的过程中，就要本着历史与现实相结合的观点，既不能不顾其成熟完善程度而盲目使用，又不能因其不够完善而弃之不理，我们应本着在应用和研究中完善并发展这一基本方法，精心研究，大胆使用。

### （五）定性和定量相结合

在信息学方法论的建设过程中，应当坚持定性分析与定量研究相结合，把两者有机地结合起来。事实上，在一项研究工作中，定性与定量是相辅相成，缺一不可的。定量研究为信息研究提供数量依据，侧重于数学模型的建立和求解；定性分析把握信息研究的重心和方向，侧重于物理模型的建立和数据意义的阐述。但传统信息学研究只作定性的说明，很少作定量的分析。目前的信息学研究中有人过分强调定量分析而忽视甚至否定定性研究，这是不可取的。

其实，定性与定量的结合，不仅仅在于通常意义下的定性与定量相结合的具体方法，而且还在于强调一种系统的研究方式。这种系统研究方式的实

质是以与被研究客体，即信息学本身的复杂性相适应的“认识系统”去对付“对象系统”。定性与定量相结合的实质，是指把理论和经验、逻辑和非逻辑、人的智慧和现代化研究工具结合起来。因此，信息学方法论的建设应当把定性与定量的结合作为一个重要原则去遵循。

#### （六）个体论与整体论相结合

个体论与整体论之争，是由于强调自然科学与社会科学研究对象的不同特征而引发出来的。通常情况下，注重自然科学类的差别而不大注意个体差别，较易从个别推知一般；而社会科学则很重个体的差别，认为每个人的存在都是不可置换的。总之，个体论者更注重个别性的认识，整体论者更注意由整体联系去解释个体的活动。

信息学作为自然科学与社会科学的交叉学科，其研究对象也必然身兼二性，既有自然科学的性质，又有社会科学的性质。因此，在信息学研究中，我们既不能从个体论和整体论中择一而用，也不能回避其矛盾，而只能是用马克思主义的辩证综合，达到和谐发展。既要用个体论的方法研究信息活动中用户、信息工作者的个体行为，又要用整体论的观点，从整体上把握信息活动全过程及其普遍规律。

### 三、信息学方法论建设与研究的意义

正如科学研究方法论是科学技术发展的重要工具和巨大动力，研究方法的使用和变革，总是直接或间接引起科技发展的飞跃一样，信息学方法论的建设对信息学的建设和信息工作的向前发展也具有重要的理论和实际意义。

#### （一）奠定信息学的发展基础

任何学科的建立和发展都是以方法论的建立和发展为重要内容和基础的。例如，没有运筹学方法的发展，就不可能建立现代的系统工程技术，统计方法和类比方法的应用为信息论的创立提供了重要条件。几十年来，我国的信息研究与其它软科学相比，发展不快的重要原因之一就是信息部门不重视方法论和现代化的建设。对于很大一部分信息研究机构和人员来说，几十年一贯制，在研究方法上创新不多，移植不力。如果我国信息学的发展不从理论与实践两个方面建设自己的方法论体系，就难以把信息学作为一门学科来发展，就不可能在科学领域中获得自己应有的地位。

#### （二）有助于信息学学科理论建设

从本质上讲，所谓科学的“危机”，归根到底是科学基础概念和科学方法的危机，一旦新的科学研究方法产生，就会产生一场新的科学革命。有的学者甚至断言：如果方法不能有所创新或者有效地突破，那么指望在信息学理论上有所突破，将是难以企及的。巴甫洛夫指出：“科学是随着研究方法所获得的成就而前进的。研究方法每前进一步，我们就要提高一步，随之在我们面前也就开始了一个充满着种种新鲜事物的更广阔的前景。因此，我们头等重要的任务是制定研究方法”。对于信息学来说也是这样。

信息学研究方法论作为信息学认识的主要手段、认识工具及认识工具之“软件”，理所应当承担起信息学的开路先锋任务，它还应当也能够为信息学的发展起到定向作用，成为获得信息学研究成果的桥梁。

#### （三）有利于学习、识别、评价、借鉴和继承已有的信息研究成果

目前，信息学研究中的许多结论和成果是通过运用各种科学方法而取得的，如果我们不满足于在字面上被动接受这些研究者所得出的结论，不仅知其然而且要知其所以然，就必须能够理解研究者在提出问题、搜集资料、采



集数据和分析研究中所采取的各个步骤，以及由此得出结论的过程。如果不了解、这些方法和步骤，在这些方面不具备相应的知识，那么就很难理解和评价信息学研究中许多新颖的、令人感兴趣的研究成果。有经验的研究人员在评价一篇科学论文时，往往不会只看其结论，重要的是要看论文作者是通过什么方法得出这个结论的。

#### （四）有利于提高科学素养，合理安排研究工作，正确地开展研究工作

学习和研究信息学方法论，对于课题的选择、线索的鉴别、假设的确定、论证与验证，无疑都会有所帮助。运用这些方法，将会使我们准确及时地发现信息活动的各种问题，抓住其中最主要的问题，选择推动问题解决的最适合的方式方法，排出一般问题解决的先后顺序，以点带面，各个击破。

相反，一个不具备信息学方法知识的信息工作者，是不可能去辨别、定义、建议筹划或进行一个有价值的信息学研究项目的。当前，信息工作中的许多问题没有得到很好的、有效的研究，其中一个非常重要的原因就在于广大信息工作者缺乏或根本没有接受过科学方法的训练。不难想象，一个不懂布拉德福定律的人如何去搞核心期刊选定，一个不懂控制论、系统论的人如何去搞信息系统设计和信息控制研究，一个没有学习过引文分析方法的人如何去研究几位专业作者之间的关系。

#### （五）有助于帮助信息工作者掌握正确的思维方法和工作方法

信息学方法论是信息学认识程序的规范化、最优化。所以学习和研究信息学方法论将有助于信息工作者掌握一套科学的认识方法，因而也有助于信息工作方法的规范化和最优化。法国生理学家 C. 贝尔纳（Bernard）曾经说过：“良好的方法能使我们更好地发挥运用天赋的才能，而拙劣的方法则可能阻碍才能的发挥。因此，科学中最难能可贵的创造才华，由于方法拙劣可能被削弱，甚至被扼杀；而良好的方法则会增长、促进这种才华。”

学习和研究信息学方法论，还有助于我们掌握良好的思维方法。正如斯蒂芬·梅森所说的。“比起任何特殊的科学理论来，对人类的价值影响最大的恐怕还是科学的方法论。科学方法依靠理论论证而不诉诸情感，它提出在不同的观点中进行抉择时必须尊重经验的证明，这种做法现在比起一百年前好像更广泛地被运用到处理人与人的关系方面来了。”受方法论的影响，信息工作者在考虑问题和处理问题时更多地采用理性主义和客观主义的观点，这有利用树立唯物论和辩证法的观点，以正确的思维方式开展信息学与信息学研究工作。

#### （六）提高信息研究工作的生存能力

社会和用户的信息需求是信息机构生存和发展的基础。随着现代管理和改革开放的发展，信息研究机构面临着重大的竞争和挑战，主要表现在咨询业的竞争；信息机构信息优势的削弱；信息研究和人员应变能力的不足。同时还应当看到，竞争和挑战不仅来自外部，也来自内部。信息技术的发展，计算机检索系统的建立，也将影响信息部门内部对文献搜集和研究两大工作的相互关系及发展。

应付上述问题的核心在于增强信息研究机构满足社会 and 用户需求的能力，提高信息研究的速度和质量。为此，不解决信息研究手段和方法的现代化是很难实现的。因此，方法论的建设对于信息研究工作乃至信息学的生存和发展来说，具有重大的现实意义。

### 第三节 信息学方法论的体系结构

#### 一、信息学方法论体系的构成

信息学方法论是指信息学研究中，为研究信息及信息过程的规律性所采取的一切方法和技巧的总和。信息学方法既包括从微观上研究具体信息的方法，也包括从宏观上研究信息过程规律性的方法。在信息学方法论中，通常以科学方法的“三分法”——按照方法的普遍性程序将科学方法划分为哲学方法、一般科学方法与专门方法——是依据构造信息学的方法体系，这一分类能够较全面地揭示不同类别的信息学方法的适用范围。这三个层次的方法的有机结合构成了信息学方法论体系。

##### （一）哲学方法

哲学方法是适用于一切科学研究的最普遍的科学方法，也是具有高层次指导意义的方法。一般而言，哲学方法并非用来解决某一学科的具体问题，而只是在哲学高度上来研究问题。哲学方法从信息学开始产生起，就介入其中，并形成了信息学哲学流派，如布鲁克斯的知识吸收学派、乌勒舒拉的哲学学派。这与物理学的情况不一样，物理学在有了统一的规范之后，才引入哲学的方法，而信息学研究中，哲学方法的应用是在统一规范形成之前。这是因为哲学不仅作为一般的世界观与方法论指导信息学研究，而且能够深入到信息学研究的某些细节上，如信息的定义、信息学理论基础与理论体系的研究。

##### （二）一般科学方法

一般科学方法是指各门学科或大部分学科都采用的科学方法。它是将各个学科研究方法的共性抽出来而形成的具体的泛化方法，具有通用性、横断性和综合性特征。我们可以将应用于信息学研究中的一般科学方法分为四大部分：横断科学方法，如数学方法、控制论方法、系统方法、信息方法等；经验科学方法，如调查法、观察法、实验法等；理性思维方法，如比较与分类法、归纳与演绎法等；移植科学方法。一般科学方法在信息学研究中的应用，使得信息学研究行为过程规范化、科学化，信息研究成果也更加科学，更容易为社会所承认和接受。

##### （三）信息学专有研究方法

作为一门年轻的学科，信息学的专有研究文法还不太成熟、不太完善，多半是由其它学科的研究方法移植过来加以改造而成的。但是随着这些方法在信息学研究中的应用，这些方法已经或正在成为揭示信息和信息现象内部规律的有独特作用、独到之处的研究方法。我们可以把信息计量学方法、引文分析方法看成是比较成熟的信息学专有研究方法，并把信息整序方法和内容分析方法等看作正在发展、很有前途的信息学专用研究方法。它们除具有研究功能外，还被视为使信息学成为一门独立学科的标志。

信息学方法论体系的三层次，有区别又有联系。哲学方法对一般科学方法和信息学专有方法有指导意义，是对一般科学方法和信息学专有方法的理论概括；信息学专有方法要接受一般科学方法的指导，同时也是一般科学方法的基础。哲学方法、一般科学方法、信息学专有方法共同构成整个信息学方法论的体系结构（以后各章节将择要介绍）。

#### 二、信息学方法论体系结构

## 第二章信息学研究的哲学方法

### 第一节信息学的哲学基础研究

哲学基础是信息学理论基础的重要组成部分，是信息学理论研究的世界观和方法论。尽管人们对信息学的哲学基础尚无一个统一的认识，但各国信息学家从不同角度、不同深度对其进行了探讨和研究，形成了不同的派别。

信息学哲学基础研究派别主要有：科学哲学流派、马克思哲学流派、贝塔朗菲系统论流派和结构主义流派等。

形成信息学哲学研究流派众多的原因主要有三个方面：一是从哲学的发展历史来看，哲学已从注重本体论的传统哲学研究，进入到以认识论及方法论为主的当代西方哲学研究。正是由于波普尔的批判理性主义、库恩的历史主义、拉卡托斯的“科学研究纲领方法论”、皮亚杰的结构主义、马克思的唯物主义和历史唯物主义等多种哲学流派的产生与发展，使得信息学者采用不同的哲学观点，形成了众多的信息学哲学基础研究流派。二是从信息学的研究者来看，信息学家本身的知识水平、对信息学的认识深度、对哲学的吸收理解程度、运用哲学观指导信息学及解决信息学问题的能力，亦是形成“观点各异”的主要原因之一。三是从信仰、社会、文化等因素来看，不同的哲学信仰、科学技术水平、文化制度、文化背景亦是不可忽视的重要原因。

#### 一、科学哲学流派

科学哲学流派包括以波普尔的“世界3理论”为哲学基础的本体论流派和以库恩的科学“范式”、“科学革命发展的动态模式”及拉卡托斯的“科学研究纲领方法论”为哲学基础的认识论及方法论流派。

##### （一）本体论流派

它是卡尔·波普尔（Karl Popper）的“世界3理论”为哲学基础，波普尔本体论可概要如下：

世界1——物质世界，宇宙，地球对我们虽是维持生命所必需的，但在宇宙中只不过是无限的物质和辐射的整体中微不足道的一个点。

世界2——主观经验世界或“意识状态”。

世界3——客观知识世界，即人类头脑的产物，它记录在语言、艺术、科学技术以及其他一切人工制品中，人类已将它们贮存或扩散在整个地球上。

波普尔认为，虽然这三个世界是独立的，但它们也相互作用。由于人类生活在地球上，因此我们是物质世界的一部分，并且为了不断地生存下去，需依赖于太阳发出的光和热、空气中的氧、食品中的蛋白质、水等。由于我们具有脑力，所以我们也是世界2的一部分。波普尔将他发表的思想记录在他的著作中，而我们则可以一直借助于世界3的资源。书籍也是物质实体，是世界3的一小点，但它独立于人创造它的人类，作为客观事物而存在。布鲁克斯正是以“世界3理论”作为信息学的哲学基础，确定了信息学研究对象在本体论上所处的哲学位置，指明了信息学的理论工作是研究“世界2”和“世界3”的相互作用；实际工作是收集和整理“世界3”的记录；波普尔的“世界3理论”之所以引起信息学者的注意，是因为该观点第一次从纯实用观点以外的角度，为信息学的专业活动提供了理论根据。

布鲁克斯的《信息科学基础》发表使波普尔的“世界3理论”在国际信

息界引起了极大反响。我国信息界也从 1982 年开始引进和介绍了波普尔的“世界 3 理论”，其中有不少学者追随布鲁克斯开始在波普尔的“客观知识世界”的观点和哲学理论上构建信息学理论大厦。同时也有另一部分人提出异议，并以唯物辩证法的观点展开了对波普尔哲学的批判，指出了波普尔“世界 3 理论”的重大疏误，如波普尔把意识的客观实在性说成如同物质一样是客观实在的东西，混淆了意识与物质的根本区别和对立；再如，把社会意识依存于社会存在，并反映社会存在的相对独立性，歪曲成为一种独立于物理世界之外的“自由的”独立性等。因此，我们既要肯定“世界 3 理论”具有一定的合理性，尤其是关于“客观知识”概念的提出对信息学理论建设的巨大启发和启迪作用，同时又应指出“波普尔世界 3 理论并没有科学地说明和解释世界”，在论证和逻辑上出现了一些疏漏，有其不够严谨的一面，所以不应当“把信息学的哲学基础建立在波普尔为我们所构筑的如此松软的沙滩上。”信息学的理论基础只能是辩证唯物主义，而关于信息的普遍性与特殊性、运动性与静止性等问题，也只有从辩证唯物主义观点出发，才能作出正确的解释。

## （二）认识论及方法论流派

### 1. 以库恩的“范式”及“科学发展动态模式”为哲学基础的观点：

库恩的“范式”是美国科学哲学家托马斯·库恩提出来的，它是指“在科学实际活动中某些被公认的范例——包括定律理论、应用以及仪器设备系统在内的范例——为某一科学传统的出现提供了模型”。科学与前科学的区别就在于科学具有范式（范型、模范），而前科学、非科学则没有。范式为科学共同体（科学家按同一的范式组成的集体）所一致拥有，他们按统一的范式从事科学研究活动，这就是科学。前科学的特点是工作者对他们所从事的学科的基本原理，甚至有关的观察现象、观点完全不一致，经常争论，每个人不得从头开始证明自己的特定观点是正确的。用“范式”的理论来看信息学现状，可以认为信息学正处于由前科学向科学发展或转化的过程中。有了这个基本的看法，在信息学理论研究中就既要允许对基本概念、原理的继续争论，又要力求求同存异，争取在较短的时间内早日步入正式科学之列。

库恩的“范式”理论是客观世界某一事物的内部结构，在人们头脑中反映出来的，方法论的知识或理论的框架，有其合理因素，但亦有其弱点：第一是“范式”的含义不够精确，第二是以科学家的心理与信念为基础建立科学理论是唯心的；第三是把“范式”的更替看成是突如其来的绝对的变化，而不是人类知识长期积累的结果。

### 2. 以拉卡托斯的“科学研究纲领方法论”为哲学基础的观点。

科学研究纲领，是拉卡托斯提出的一个描述科学发展的模式和方方法论。拉卡托斯主张不能孤立地对一个理论进行评价，评价任何科学理论都必须连同它的辅助假设、背景知识和初始条件一起进行。而由某种连续性联系起来的各个理论所组成的理论系列就是一个科学研究纲领。科学研究纲领由四个相互联系的部分组成：最基本的理论构成“硬核”；许多辅助性假设构成的“保护带”；保护硬核心反面启示规则——反面启示法；改善、发展理论的正面启示规则——正面启示法。其中“硬核”是整个研究纲领的基础部分，不许改变，不容反驳，否则将导致整个理论系列被彻底否定；各种辅助性假设所构成的“保护带”之职责就是保卫硬核，竭尽所能不让硬核遭到经验事实的反驳；反面启示性是规定人们不应该干什么，避免哪些研究途

径，不要把矛头指向硬核而尽量把经验反驳转向保护带；正面启示性则鼓励人们精简、修正、增加或完善辅助性假设，以稳定、完善和发展整个学科研究纲领。

由于拉卡托斯“科学研究纲领”的内核比库恩的“范式”意义明确的多；拉卡托斯的硬核是理性的产物，而库恩的“范式”是心理主义的解释；拉卡托斯肯定了新旧理论系统之间的连续性和进步性，而库恩的科学发展动态模式是突变的。因此，拉卡托斯的“科学研究纲领方法论”逐渐引起了信息学界的关注。

科学研究纲领的理论，给信息学的研究方法和原则指出一条很有价值的思路。我们也可以考虑选择或约定出信息学理论的“硬核”，并以各种辅助假设把“硬核”保护起来，形成保护带。当信息学理论与从信息工作实践中所得的结论有矛盾时，应首先考虑调整辅助性假设，通过修正和完善，保护信息学理论的硬核，保持信息学理论的稳定性。只有在修改或增加辅助性假设不能导致得到证实的新事实的预见时，才把否定指向硬核。这样才能既保证信息学理论的相对稳定，又保证信息学理论体系的开放性，最终使信息学理论增加生命力和活力，求得更大的发展。因此，对正处于发展阶段的信息学来说，科学研究纲领方法论不失为一种较理想的、构建理论体系的哲学基础。

## 二、马克思主义哲学流派

该派主要以马克思的辩证唯物主义作为信息学理论体系的哲学基础。由于各国哲学信仰、社会环境、文化背景不同，而我国自建国以来主要是用马克思的辩证唯物主义作为世界观和方法论，所以，我国信息界比较重视马克思主义哲学对信息学理论体系建立的指导。

利用辩证唯物主义认识论能够有效地分析信息工作的内在矛盾，即信息与信息需求的矛盾。辩证唯物主义认为该矛盾产生、存在于整个认识过程，是人类认识过程的一个环节；认为信息不仅是人类认识活动的产物、人类认识的对象，还是人类认识的工具、联系主客体的中介；认为辩证唯物主义认识论为信息学提供了适当的哲学位置，是信息学的哲学基础。

但也有部分信息学者发表了不同的看法，从辩证唯物主义认识论角度来看，认为信息与信息需求的矛盾是认识过程的一个环节，并以此寻找信息学的哲学位置是符合我国国情的。然而从科学认识论的角度来看，信息学理论体系应该以科学事实、科学概念、科学定律为基础和构成材料；通过自我组织、自我调节和自我反馈形成科学的动态理论系统；通过科学的实践检验判明其真理性，检验成功了才能被人们所接受。因此，仅从辩证唯物主义能为信息工作的内在机制提供适当的哲学位置及认识论，就认为辩证唯物主义认识论是信息学的哲学基础似乎还不够严谨。

因此，信息学理论应以辩证唯物主义为立足点，吸收西方某些理论的有益成分，才能健康地发展和完善信息学的理论体系。

## 三、贝塔朗菲系统论流派

该派主要以系统论思想为哲学基础

从目前信息学理论研究现状来看，国内外皆很重视系统论思想对信息学理论体系建立的作用。它以贝塔朗菲的一般系统论为哲学基础，探讨了信息现象和信息过程的系统行为规律；并按照钱学森的三台防理论将信息学理论体系分为信息服务工程 信息科学技术 信息学原理 马克思主义哲学；认

为系统论的产生和运用，揭示了客观物质世界新的本质联系和运动规律，为现代科学技术的发展提出了新思想、新方法，因此，以系统论为哲学方法能使信息学早日成为一门体系完善严谨的科学。

不同的观点认为，尽管系统论能从不同侧面揭示信息学研究对象之间的相互联系、相互作用及内在运动规律，然而，系统论只是提供了一种使信息学理论体系建立更加完善的一种哲学方法，并没有指明信息学在哲学意义上的位置及核心任务；只是说明了信息学理论体系作为一种系统具有动态性，并没有指明其科学发展的动态模式如何。

#### 四、结构主义流派

该派主要以结构主义的深层、表层结构为哲学基础。目前，无论是国内还是国外，以结构主义为哲学方法论研究信息学理论体系的还不多。

以结构主义为哲学方法论，以科学技术知识体系为参照系，信息学可划分为深层结构、表层结构和边缘结构。深层结构是指信息学各分支领域和各构成要素所共有的结构形式，是深藏于信息学表象内部的深刻本质，决定了信息学的外在表现形式——信息学表层结构；表层结构是信息学各分支领域和构成要素所特有的结构形式，既统一于深层结构，又具有多样性；边缘结构是指信息学以特有的结构形式为媒介，为相关学科建立的各种联系。

虽然说已有的文献从结构主义角度探讨了信息学各构成要素之间的联系，将信息学的结构划分成深层、表层、边缘结构，为信息学理论体系的建立提供了一种全新的思路，但并没有论证信息学整体结构模型的最本质问题，即深层结构的具体内容是什么，这是与结构主义以深层结构为重点的观点不相符合的。

## 第二节 哲学方法简介

### 一、哲学方法的概念

所谓哲学方法，就是运用哲学的原则解决具体的科学问题。哲学方法作为信息学方法论系统中最高层次的方法，可以理解为如下四个方面的含义：

1. 哲学就是方法论。现代哲学与科学分工明确，全部实证知识应由包括信息学在内的各门科学提供，哲学只提供方法论的原则；另一方面，任何一个完整严谨的哲学体系本身，都应既是世界观又是方法论，是世界观与方法论的统一。从哲学的本性看，它的任务主要是向人们提供如何才能正确把握对象存在和发展的规律，它所要启示人们的，主要是一种方法论，即教人们应如何正确思维，以及提供正确的思维方式和思维方法。

2. 哲学方法具有普遍的适用性，能涵盖一切领域。科学方法具有三个层次，即具体科学方法（如信息学的专有方法）、一般科学方法和哲学方法。哲学方法是从前两个层次中提炼出来的，而前两个层次方法又蕴含着哲学方法的内容，二者有着本质上的共同点和一致性，是一般与个别的关系。哲学方法作为最高层次的方法，它适用于并能指导自然科学、社会科学、思维科学及交叉科学的研究，是人们处理外部世界关系的基本规范和准则。在人们认识和改造世界的活动中，它作为总体和一般性的方法论起作用，因而被包括在一切范围的科学认识结构之中。

3. 哲学方法具有多元性。它不仅是指马克思主义哲学方法，也指以往一切旧哲学方法；不仅是指中国的哲学方法，也指外国的、西方的哲学方法。

在西方资产阶级哲学中，其本体论和世界观或许是唯心的，但在方法上往往有可吸收的合理因素，如黑格尔的唯心辩证法。现代西方哲学中关于方法论的合理思想，也是人类科学思想宝库中的有机组成部分。例如系统学分析方法的产生及其应用、波普尔三个世界划分的理论及其影响，都说明了这个问题。

4. 哲学的应用化也体现了哲学就是方法论的功能。马克思主义哲学的一个显著特点是重视哲学在实践中的应用，马克思主义哲学是哲理与现实相统一的哲学应用观。哲学要改变世界，就离不开哲学的应用，而这又突出了哲学的方法论功能。应用哲学作为基础哲学和具体科学的中间环节，是完成指导具体科学又发展哲学这一双重任务的有效途径。这样，就可以使哲学不再是站在最前面去面对各门科学作概括总结，而是渗透到各门科学中去作哲学思维，突出了哲学的功能，并使上下趋于逐级抽象、具体的双向流动之中，以保证哲学的丰富性、开放性、科学性，又可以为哲学开辟一条把握时代精神的新通道，使哲学不再只是远离现实的玄想，而是活跃于人们改造自然、变动社会、更新观念的现实生活中。

## 二、哲学方法的特点及局限

### （一）具有普遍性

哲学方法作为最高层次的方法，是人们作为主体处理同外部世界关系的基本规范和准则，在人们的实践活动中起着总体性和一般性方法论的作用，具有最大的普遍性和广泛性。它对其它层次的方法居于统驭一切的地位，起着灵魂和支柱作用。哲学方法可以运用于所有科学，包括自然科学、社会科学、思维科学和横断科学、交叉科学的研究之中，而不像其它方法只能运用到本学科或移植到部分科学研究中。

### （二）具有抽象性和思辨性

哲学方法的总体性、普遍性等特点，是与哲学思维把握对象的高度抽象性和高度思辨性的特点分不开的。哲学方法的抽象性和思辨性表现在它在研究具体问题时，是从哲学的高度，从处于支配地位的层次进行指导，而并非像具体科学方法一样直接用于解决本学科中的某一个具体问题。哲学方法并不告诉你关于信息的概念，信息学的研究对象到底是什么，而只是抽象地思辨地指导你，应当如何看待、分析、研究、解决概念的问题和学科研究对象的问题。也正因为如此，哲学方法才成为人们认识和把握事物普遍本质与真理的根本方法。

### （三）具有主体能动性

哲学思维和哲学方法作为主客体相互作用的手段与中介，作为主体能动地把握客体的最高方法论，具有鲜明的主体性和能动性。可以说，任何一种哲学方法的形成都是人类自觉意识的觉醒和主体性增强的标志。主体能动性这个特点是哲学方法所特有的，应用于信息学中的其他科学方法都没有这个特点。

### （四）具有时代性

任何哲学方法都是时代的历史的产物，既受时代的局限，又随时代发展而不断发展。从方法论的角度上说，由于哲学方法与科学方法的相互渗透，决定了二者结盟、汇流的大趋势，并随着今天自然科学和社会科学的“一体化”而愈加明显，突出地表现在科学的哲学化和哲学的科学化。而横断科学、交叉科学，如系统论、控制论、耗散结构论等，由于其普遍性和较普遍的方

方法论意义，很可能成为哲学方法和科学方法汇流的交接口，成为新的生长点。即使像马克思主义哲学的一些基本的方法论原则，也较创立时代大大丰富和发展了。

然而，哲学方法毕竟是一种高度抽象、思辨的方法，它只是给信息学研究提供一些原则，使信息学研究人员站到更高的位置来看问题，或者是给我们提供一种看待、思考、分析信息学与信息工作中各种现象和问题的方法。信息学与信息工作中大量的具体问题的解决，还是应当采用一般的科学方法，如调查法、观察法、归纳与演绎、分析与综合等等。我们既要支持和鼓励广大信息学研究人员掌握和运用哲学方法，对信息学中的重大问题进行有效的探讨，又坚决反对为了应用哲学方法而进行的哲学思辨，故弄玄虚，或是不分问题大小，适合与否，都一味追求哲学方法的应用。

### 三、哲学方法的基本功能

从方法论的角度看，哲学方法的基本功能可概括为实践功能、反思功能、批判功能和预见功能。

#### （一）实践功能

实践功能是哲学尤其是马克思主义哲学的根本功能，也是哲学方法的其他功能的出发点和归宿点。哲学不只是理论地解释世界，单纯地履行描述功能，更重要的是要行使规范化功能，即指导人们对现存世界进行实践的批判，从而实际地反对和改变事物的现状。马克思就把自己的哲学世界观规定为“实践的唯物主义”。然而，强调哲学的实践功能，尤其是马克思主义哲学的实践性，并不意味着它能直接地为变革信息工作的实践服务，而是需要通过它的强大的方法论功能和经过信息学以及信息实践观念这样的中介环节，提高人类的理性思维能力。哲学的至高之功乃是提供最普遍的方法论原则。

#### （二）反思功能

哲学反思，指的是对“认识”的认识，对“思想”的思想，对“思维”的思维，即对科学成果的再思考。具体学科及其方法也有认识和思维功能，但只在于提供具体知识，而不可能像哲学反思那样提供获取具体知识所需要的世界观和方法论的知识，也不能像哲学一样对科学成果进行再思考。通过反思功能，通过对“认识”的认识，哲学还为具体学科提供最一般的概念，如内容和形式、原因和结果、规律、矛盾、本质和现象等。对“认识”的认识和对“思想”的反思，是哲学把握包括信息学研究在内的整个世界的特有方式。

#### （三）批判功能

哲学之所以为哲学，至为关键的是它具有批判、扬弃和超越功能，这也是哲学的活力和动力之所在。哲学与哲学方法的批判功能首先表现在对理论的批判。哲学批判作为一种前提性考察，就是对理论前提（经验前提和逻辑前提）的考察，是对命题本身是否合理的追究，对理论的根本前提进行质疑以致重新建构，以求理论在根本前提上实现重大层次的跃迁，并力图倡导人们的理性自由精神，培养人们的怀疑精神和创新意识。其次，哲学批判还表现为哲学对现实、实践的批判。因为虽然一切信息实践活动都是现实，但未必是合理的，只有通过对其进行的批判性分析和总结，才能提高广大信息工作者对现实信息实践活动与其合理性的认识。哲学批判功能的第三个表现是哲学的自我批判，即哲学对哲学的批判。例如，对于信息界兴起的波普尔“世界3理论”热，可以用马克思主义哲学的辩证唯物论对之进行分析和批判；



而且马克思主义哲学本身也正是通过不断的自我批判而逐渐完善起来的，并且得到空前的发展。

#### （四）预见功能

哲学的预见功能对于认识和指导具体科学的发展具有重要意义。哲学的正确位置不是跟在科学后面做总结，而是要走在科学的前面。科学是以探索未知为己任的，因此，预见也是科学生命力和力量的重要标志。但由于具体科学的视野比较窄，影响了它对事物发展作出整体性、长远性的科学预见，所以，哲学从全局、整体、长远方面的预见，不仅对探索科学研究中的未知世界，而且对探索各门学科本身的发展都十分重要。哲学方法的预见往往蕴含着丰富的科学思想，具有极大的创造性和深刻的启发性。它能够给信息学研究人员以坚定的信念、明确的科学方法和探索目标，并提供全新的方法论武器。这对于确定信息工作的目标、信息学研究的新领域等都有十分重要的指导意义和借鉴作用。

### 第三节 哲学方法的应用

哲学方法作为一切科学的最普遍的方法和最一般的理论工具，它在信息学研究过程中一直发挥着方法论的作用。例如在处理信息领域中的认识与实践的关系、部分与整体的关系、形式与内容的关系等方面，都需要哲学方法的指导或应用

从微观上看，信息学作为一门新兴的综合性学科，还非常年轻，其理论边界和研究范围还未最后确定，理论体系还未最后建立，甚至连最基本的概念和理论基础都还在争论之中，学科地位也不算十分稳固。这些问题的解决，除了对信息学进行深入研究外，哲学理论的指导是决不可少的。因此，信息学对哲学方法的需求就更迫切、更直接。

信息学的发展史也表明，哲学方法是信息学研究广泛使用的方法之一。实际上，在信息学的许多研究领域都已自觉或不自觉地应用哲学方法了，人们不仅从哲学的角度对信息学的一些基本概念、一些基本问题进行思考，运用哲学的一些观点和方法来解释社会的现象，而且还借用哲学的某些理论来建筑信息学的理论大厦。

下面具体叙述哲学方法在信息学研究中的实际应用。

#### 一、哲学方法在信息概念探讨中的应用

“信息”的概念是信息学中最基本的概念，所以信息学研究者应用各种方法从不同角度对信息概念进行了探讨，目前关于“信息”的定义已达几十余种。早在对信息概念进行探讨的初期，许多研究者就已开始尝试使用哲学方法，从概念和概念发展的辩证法着眼，分析信息现象的历史和现状，讨论信息与认识的关系，从而确定信息概念的内涵和外延。

还有人从信息的复杂性和多变性出发，提出了信息的不可定义性。从马克思主义哲学的观点来看，这种“不可知论”的观点是不正确的。马克思主义的认识论是彻底的可知论，它认为通过实践，客观世界及其规律是可以被认识的，世界上只有被认识的和尚未被认识的事物，没有也不可能有不认识的事物。信息现象的复杂性是人们所公认的，信息概念的多种解释是人们尚未完全认识信息本质的一种表现。随着人们对信息认识的进一步深入，信息的本质会愈来愈为人们所掌握。因此，应用正确的哲学观点和方法，我

们就不会由于信息的复杂性而认为信息不可定义。同时，我们也承认对信息概念的正确认识要有一个过程，目前对信息的定义只是相对正确地揭示了其本质，而并非已经一劳永逸地找到了关于信息定义的绝对真理。

## 二、关于图书信息的“中介世界”理论

信息学界对图书信息到底是属于物质世界还是精神世界一直争论不休。有人提出了图书信息是物质世界与精神世界间的一种“中介世界”的观点，认为信息作为一种特殊的存在，其内容是精神的，其载体是物质的，它既像物质世界，又像精神世界，既不像物质世界，又不像精神世界，好像同时属于精神和物质这两个世界。在信息这个特殊存在中，物质与精神的差异配合在一起，二者之间绝对分明的界限模糊了，固定不变的边界变化了。

恩格斯在《自然辩证法》一书中曾写道：“一切差异都在中间阶段融合。一切对立都经中间而过渡，对自然观的这种发展阶段来说，旧的形而上学的思维方法就不再够了。辩证法不知道什么绝对分明的和固定不变的界限，不知道什么无条件的普通有效的‘非此即彼’！它使固定的形而上学的差异互相过渡，除了‘非此即彼’，又在适当的地方承认‘亦此亦彼’，并且使对立互为中介。”可以认为，关于图书信息的“中介世界”的理论是在图书信息领域中对唯物辩证法观点和方法的较好应用，是较有新意的信息学研究观点。

## 三、哲学方法为信息学研究提供了思想方法

马克思主义哲学唯物辩证法认为，世界上的一切现象都处于普遍联系和永恒运动之中，实践是主观和客观对立统一的基础，理论产生于实践，应用于实践，并在实践中得到检验。据此观点和方法，我们在信息学理论研究中，就应做到：

1. 要坚持用全面的、联系的观点看待和分析问题。我们在探讨信息事业产生、发展、变化的历史原因时，要把信息工作与整个社会的经济、科技、文化背景密切联系起来，避免孤立、片面地研究和分析，要把信息工作和信息系统建设放到社会大系统中，来看其地位、发展方向和作用的发挥，避免脱离社会需求搞自成体系的信息系统建设，避免脱离社会系统空谈信息工作的地位和作用。

2. 要以发展的观点分析信息活动的各种现象和问题。例如在研究信息工作的任务和重点时，既要客观地、历史地看待前一阶段信息工作的任务和重点，肯定它在当时历史条件下的必要性、必然性与成绩，又不能为前一阶段的工作重点与任务所限制，还要根据当前的形势确定目前信息工作的任务。另外，也要充分估计下一阶段的形势，在做好现阶段信息工作的同时，注意为下阶段信息工作的开展打好必要的基础。总之，要避免用固定的、静止的、一成不变观点看问题，既不能因形势的变化而否认过去的一些成功的作法和正确的观点，也不能把根据当前情况确定的方针、政策当作永远正确的东西不再改变地一直贯彻下去，不允许后人根据以后的实际情况再做任何改变。

3. 信息学理论一定要密切联系信息工作的实际。信息学研究的课题应多从信息工作面临的问题中去寻找和确定，信息学研究人员要到信息工作实践中作深入广泛的调查研究，信息学研究的成果要能够指导信息工作的实践。

## 四、哲学方法与信息学理论大厦的建立

科学理论的产生，首先要有一个在一定思想原则指导下的经验积累的过程，这是科学体系形成前的最简单也是最基本的准备阶段。一大堆杂乱无章

的实验材料远远不是科学。所以信息论理论的创造，与其说是需要单纯积累越来越多的、越来越新的事实，不如说需要对信息活动的现实领域内的关键性事实进行深刻的概括，或者叫一般哲学方法论进行分析。

信息理论界在运用哲学方法构建信息学理论体系上作了许多富有成效的尝试。早在 1981 年，就有人提出要用科学哲学的理论观点，特别是库恩的“范式”理论来看待并建设信息学理论。还有人曾尝试用先进的哲学思想和辩证逻辑，探索信息学的结构问题，它根据研究知识运动形式、过程、要点等的差别从体系结构上把信息学分为确定研究对象、范围、性质、任务、原理、定律与相关学科关系的基础理论，着重研究社会各个局部中知识运动规律的专门信息学和研究信息学应用技术的信息管理学，并从事物发展的根本原因在于事物的内部矛盾性这一哲学论断出发，分析了信息学发展的动力结构表现为信息供求矛盾等五大矛盾方面。

还有研究者在对现代科学技术体系的纵横两方面进行哲学解释之后，又从信息学的研究对象、信息学研究对象在客观世界中的存在范围和信息学研究对象的深度三个方面对信息学体系结构进行了哲学探索，提出了信息学对象结构图。

可见，运用哲学原理和方法探索建立科学的信息学理论大厦无疑是一项极富挑战性的工作，而且虽然得出的结论未必十分正确，但不失为一种好的尝试。

## 第三章信息论与信息方法

### 第一节信息论与信息科学

信息论是一门新兴学科，它产生于本世纪 40 年代末，最早仅局限于通信领域，经过近 50 年的发展，由于现代自然科学发展的综合整体化趋势，各门学科的相互联系、相互渗透，信息的概念以及信息的一些基本理论已经超越通信领域，逐步推广、运用于其它学科。在此基础上，于 60 年代末至 70 年代初出现了信息科学。信息科学是在信息论基础上发展起来的，它涉及数学、通信理论、控制论、计算机科学、人工智能、电子学以及文献信息处理技术等多个领域，与哲学的关系也很密切。信息科学是一门多边缘的学科，亦称之为横断学科。信息科学的出现，大大拓宽了信息论的研究范畴，同时也推动了信息论及其方法的广泛应用。

#### 一、信息论的产生与形成

信息论作为一门科学理论，其形成和产生的渊源可以追溯到 19 世纪。19 世纪中叶到本世纪 40 年代可以作为信息论产生前的准备阶段。自牛顿力学创立以来，从 17 世纪到 19 世纪在自然科学领域中机械唯物论的世界观占统治地位，机械唯物论否认客观世界存在着偶然因素，把偶然性与必然性绝对地对立起来，企图用拉普拉斯的决定论解释一切。拉普拉斯（P·LaPlace，1749~1827 年）是法国天文学家、数学家和物理学家。他曾经提出，只要知道了宇宙中所有的作用力及所有原子在某一时刻的坐标和速度，就能确定世界的全部过去和未来，这就是人们通常所说的拉普拉斯决定论。这是一种机械决定论的世界观。值得提出的是，正当绝大多数自然科学家在这种形而上学思想的束缚下，遵照机械决定论的观点思考问题的时候，美国物理学家吉布斯（Josiah Willard Gibbs）和奥地利物理学家波尔兹曼（Ludwig Boltzmann）却首先把统计学引进物理领域，使物理学对客观世界中存在的不确定性和偶然性不得不加以考虑。把研究偶然性作为一种科学方法引进物理学，这是吉布斯的一大功绩。对此，美国数学家诺伯特·维纳（Norbert Wiener）曾予以肯定，他指出，吉布斯的统计力学方法中隐含着把一个复杂的偶然事件看作为由许多局部偶然事件无限概率之和。他在《人当作人来使用》一书中曾指出：吉布斯为了修正牛顿的传统学说而将偶然性的观点引入物理学。这种研究方法为信息论的创立提供了方法论的前提。波尔兹曼把熵函数引入统计物理学，对熵首先作了微观解释，指出熵是关于一个物理系统分子运动状态的物理量，表示分子运动的混乱程度，并且把熵和信息联系起来，提出“熵是一个系统失去了的‘信息’的度量。偶然性，熵函数引进物理学为信息论的产生提供了思想前提。他们虽然都涉及到有关信息方面的问题，但由于当时科学技术水平的限制，未能直接提出信息理论的问题。

本世纪 20 年代，根据通信实践的需要，奈奎斯特（H.Nyquist）与哈特莱（Hartley）最早研究了通信系统的传输效率问题，奈奎斯特提出，电信信号的传输速率与信道频带宽度之间存在着比例关系。哈特莱提出，用对数作为信息量的测度，这样，信息就可以用数学方法从数量上加以测度。1928 年，哈特莱发表了《信息传输》（《Transmission of Information》）一文，首次提出了消息是代码、符号，它与信息有区别（消息是信息的载体，消息的形式是多样的、具体的，如各种语言、文字、图像等等，而信息是指包含

在各种具体消息中的抽象量)。他区分了信息和消息在概念上的差异，并提出用消息出现的概率的对数来度量其中所包含的信息，从而为申农(C.E.Shannon)的信息理论的建立奠定了初步基础。

1948年数学家申农发表了著名的论文《通信的数学理论》(《A Mathematical Theory of Communication》)，1949年又发表了另一篇论文《在噪声中的通信》。当时申农在贝尔电话公司工作，他为了解决信息的编码问题，为了提高通信系统的效率和可靠性，在研究过程中，需要对信息进行数学处理，这就要求舍弃通信系统中消息的具体内容，把信源发生的信息仅仅看作一个抽象的量。同时，由于通信的对象——信息具有随机性的特点，因此申农把用于物理学中的数学统计方法移植到通信领域，从而提出了信息熵的数学公式，从量的方面来描述信息的传输和提取的问题，并提出了信息量的概念。除了上述贡献外，申农还提出了通信系统模型以及编码定理等方面有关的信息理论问题。申农的这两篇著作确立了现代信息理论的基础，而申农也因此成为信息论的奠基人。此后，信息论就逐步发展成为一门独立的边缘学科。

信息论的另一位奠基人维纳，差不多与申农同时发表了两篇著作：《控制论》(《Cybernetics》，1948年)与《平稳时间序列的外推、内插和平滑化》。维纳从控制和通信的角度研究了信息问题，主要是从自动控制的观点研究信号被噪声干扰时的信号处理问题，建立了维纳的“滤波理论”(通信与控制中的滤波问题，指的是从获得的信息与干扰中尽可能滤除干扰，分离出所期望的消息)。1951年，维纳又发表了另一著作《人当作人来使用》。在上述著作中，维纳提出了信息量的概念、测量信息量的数学公式，叙述了信息概念形成的思想前提，同时把信息的概念又进一步推广，认为信息不仅是通信领域研究的对象，而且与控制系统有密切联系。于是，信息论便成为控制论的一个基本理论了。

美国的另一名统计学家费希尔(R.Fisher)，因为需要一个标准来估计实验数据内的信息，他从古典统计理论的角度研究了信息的量度问题。

从上可知，申农、维纳、费希尔各自从不同的方面(申农从编码方面，维纳从滤波理论方面，费希尔则从古典统计理论方面)研究了信息的理论问题，虽然角度不同，但却得到共同的认识：通信必须以随机事件为对象，通信和控制系统所接收的信息带有某种随机性质，通信的目的在于消除收信人的不定性，而要对消息作定量的研究，就必须把各种消息(信号、代码等)的这一共同特征抽取出来，略去其相应的具体内容，单纯从量的方面进行度量，为此就要用统计平均信息的概念来量度消息的概率分布，从而提出用统计数学方法处理通信的理论问题，并给出了信息量的数学公式。信息量等于被消除的不定性的数量，这是申农等人信息论的最重要的结果之一。于是任何一个事件，只要知道它的各个可能独立状态的概率分布，就可以求出它的熵值，从而求出它能提供的信息量。从此，通信科学就由定性研究阶段进入定量研究阶段，这就为信息论及信息科学的研究奠定了初步的理论基础。

## 二、信息论的研究内容

申农在《通信的数学理论》一文引言中曾说过：“通信的基本问题是在通信的一端精确地或近似地复现另一端所挑选的消息。”这句话表达了通信系统的基本问题是要解决通信的有效性与可靠性这两方面的问题，即以最大的速率传递信息，而且要保证在干扰存在的条件下，能够最佳地和准确地再

现消息。为此就涉及从信源到信宿的通信全过程。申农信息论的基本问题和研究内容涉及到信源、信宿、信道以及编码问题。

### （一）信源问题

信源就是信息的来源——发出消息的来源。信息源可以是人、机器、自然界的物体等等。信源发出信息时，它一般以某种符号（文字、图像等）或某种信号（语言、电磁波信号等）表现出来，而表示信息的这些符号或信号称之为消息，消息的形式是具体的、多样的。消息一般可理解为信息的载体，它承载着信息。

信源所发出的消息带有随机性，它是不确定的。如果消息是确定的，而且是预先知道的，那么对信宿来说，就无信息可言。

信源问题的核心是信源包含的信息到底有多少？这要作定量的描述，用信息熵进行测度。

### （二）信宿问题

信宿即收信人，信息的接收者，可以是人，也可以是机器，如收音机、电视机等。

信宿问题解答的是收信人能收到或提取多少由信源发来的信息量。它不涉及信息的意义及对收信人是否有用，即所谓的实效性的问题。

### （三）信道问题

所谓信道就是信息传递的通道，是传输信息的媒质。信息是构成信息流通系统的重要组成部分。信道的关键问题是信道的容量，就是说，要求以最大的速率传送最大的信息量。从广义上看，信通不仅承担了信息的传输任务，而且也担负着信息的存贮任务。信息的传输过程，在某种意义上看也是信息的存贮过程。如用书信传递信息，发信人把书信写在纸张上，通过邮寄，收信人收到了信息，这是一个信息的传输过程。信息一经书写在纸上，纸张就成为存贮信息的载体。在电信系统中，传输信息的过程中也存在着存贮过程，如两人打电话，总需用一定的时间，即使时间很短，在这段时间内电路也起着存贮信息的作用。而在现代通信系统中，由于使用了电子计算机，它可以存贮信号，待时传送，存贮作用更为明显。所以，从特定意义上讲，信道的传输过程就是信息的存贮过程，信道的问题就是要以最大的容量存贮最多的信息。信道是多种多样的，有线信道，如同轴线缆、光导纤维；无线信息，如自由空间，电离层等。在复杂的电信系统中，信道不是单一的，而是多路的，信道的数量很大，它联系众多的信源、信宿而组成通信网络。

在实际的通信系统中，信息传输过程中总是存在着干扰，因此，提高系统的可靠性是信息论的基本问题之一。

### （四）编码问题

所谓编码就是把信息转换成信号的措施。信息发出消息，在其被送入信道前往往要经过编码过程，消息之所以要经过编码，是因为：信源发出的原始消息不适于直接在信道上传输，只有经过编码，才能使消息成为适合在信息上传送的具体形式。如，打电报，把文字符号转换成莫尔斯电码。为了对抗信道中的噪声干扰，需要把消息编码成不易被干扰破坏的具体形式，以免在接收端再现消息时引起失真。信源发出消息经过编码后，使信息熵的传输速度接近于信道容量。所以编码问题是从理论上说明，如何编码，使信源发出的消息能充分表达，信息的容量能被充分利用，并能在接收端无失真地再现消息，以提高通信系统的有效性与可靠性。

编码定理是信息论中的重要组成部分。申农从提高通信系统的有效性与可靠性出发，提出了有关的三个编码定理。申农的第一定理是无失真信源编码定理，研究了信息熵与编码长度之间的关系；第二定理为信源编码定理，指出在给定信道的条件下，总可能找到一种编码方法，使译码的错误概率任意小；第三定理为限失真的信源编码定理，指出在限定失真的条件下，失真度与信息传输率之间的关系。这三个定理在当时是对信息论的两个基本问题（有效性与可靠性）的一种理论上的解释。

从 60 年代以后直到目前，申农信息论发展比较活跃的方向是信源编码理论和多路通信理论。这些研究进展是与计算机网络、大系统、系统工程的实际需要相关的。此外，近些年来关于申农理论中信息量的研究也十分活跃，如各种广义信息量理论的出现，以及现有信息量如估值、判决、滤波（非线性）、识别理论的应用等等。

### 三、从信息论到信息科学

申农的信息理论，经过 50~60 年代许多学者的研究，逐渐发展成一门相当完整的科学理论。申农的信息论是以通信系统模型为对象，以概率和数理统计为工具，从量的方面描述信息的传输和提取方面的问题，所以称统计信息。它研究信源能够给出多少信息量，信道能够传送多少信息量，而信宿又能收到或提取多少信息量。由于申农的信息理论，对信息所作的定量描述是以概率论和数理统计为基础，因而信息可以定量地进行计算，所以使信息论成为一门科学，这正是申农的功绩。同时，也正因为如此，它只是对信息作定量的描述而不考虑信息的其它方面，如信源发出的信息语义（意义）如何？怎样对之进行定理描述？信宿收到信息后的效用如何？价值怎样？有无相对性？这些问题涉及到信息的语义与实效，或者说涉及信息的意义与价值等问题，而这些正是申农理论的不足之处。

任何信息都包括统计信息、意义信息和有效信息三个方面。早在 1949 年韦弗尔（Weaver）与申农曾经共同提出过信息问题可分为三级：第一级即申农所研究的问题，称通信的技术问题；第二级是信息的意义问题；第三级是信息的实效性问题。由于历史条件的限制，申农的理论只是从通信的技术角度，解决了统计信息问题，而对第二级、第三级问题，即语义（意义）信息和有效信息，申农假定对任何观察者只是一个常数，即是不变的，因而没有加以考虑。

随着近代科学的发展，科学技术的相互渗透与相互影响，信息论中的一些概念和基本理论已经越过通信领域而渗入其它学科，而信息概念及其理论在其他领域中的推广和运用，以及其他新兴学科如系统科学的出现和发展，原来从统计角度，局限于通信领域的申农信息理论，其不足之处更加明显地暴露出来，并促使信息科学的产生。

由于信息论广泛渗入其他学科，普遍运用于各个领域，信息理论在现代科学技术发展中所占有的重要地位，日益为人们所认识。同时，由于它的普遍运用，信息论所研究的范围不仅涉及客观现实的一切领域，而且也涉及人的思维领域，内容复杂，涉及面广，因此，对信息的概念和实质、信源熵、信息量以及语义信息、有效信息、模糊信息等问题的研究就提上了日程，要求有更确切的理解和更一般的理论，从而提出了信息科学问题。

信息科学是以信息论为基础，并与电子学、计算机和自动化技术、生物学、数学、物理学等科学相联系而发展起来的。它的任务是研究信息的性质，

研究机器、生物和人类对于各种信息的获取、变换、传输、处理、利用和控制的一般规律，它所研究的范围要比申农理论所局限的通信领域广泛得多，要求它更为一般化，并且克服申农理论的缺陷而适用于各个领域。

信息科学是 70 年代提出的，因此信息科学还处于发展的初期阶段，是一门正在形成中的学科。当前，对信息科学所研究的内容和范围问题，存在着不同的看法。现在尚无定论。我国信息学家钟义信认为，信息科学是以信息作为主要研究对象，以信息的运动规律和应用方法为主要研究内容，以扩展人类的信息功能（特别是智力功能）为主要研究目标的一门新兴的、边缘的、横断的综合性科学。其研究范围是：

- (1) 探讨信息的本质；
- (2) 研究信息的质量；
- (3) 阐明信息的运动规律；
- (4) 揭示利用信息进行控制的原理和方法；
- (5) 寻求利用信息实现最佳组织的原理和方法。

第 1 项涉及到信息是属于物质或精神范畴的问题，因此与哲学有关；第 2 项是研究信息的度量方法；第 3 项是研究信息如何产生，如何提取，如何检测、变换、传递、存贮、处理和识别的规律，以上两项属于信息论的基本领域；第 4 项是控制论，包括信息仿生学和人工智能；第 5 项是系统论，包括运筹学、系统工程、耗散结构和协同学等。因此信息科学的基础是哲学、数理化 and 生物科学；它的主体是信息论、控制论和系统论；主要的工具包括电子科学和计算机科学。

## 第二节 信息论中的方法论问题

信息理论经过近 50 年的发展，已由申农的信息论进展到信息科学，随着理论研究的进展，必然涉及到一系列有关方法论的问题。为了对某些方法论问题加以探讨，首先要看到其研究对象的扩展，因为，在一定意义上，科学方法是由一定的科学研究对象所决定的。研究对象的不同，往往要求采用的方法也不同，这里所说的研究对象的扩展，是指随着科学技术与社会物质生产的发展，信息的内容越来越多地被纳入人们的研究范围之内。

当今社会、物质生产、科学技术和社会生活有很大发展，规模日益巨大、日趋复杂，与之相应的通信与各种信息加工系统也必然如此。例如，申农当时曾研究过双向通信信道，这是由两个反问的经典模型构成的，毕竟比较简单。而现在则需要对多径信道、多用户网络、多个通信终端、相互干扰信道等加以研究，这就复杂得多。这类信息加工系统一般都是大系统和复杂系统，而且不仅限于工程技术系统，还涉及生物系统、生态系统、社会系统、智能系统等等。在这些系统中，信息的产生、获取、传递、加工、存贮、使用等等是十分复杂的，因此仅仅研究给定的信源，以及它所发出信息的统计特性是远远不够的。这就不能不涉及到信息的产生、信息的使用，从而涉及到信息的语义和信息的效用（价值）问题。正是由于这种客观需要，使得其研究对象不得不由统计信息向语义信息和有效信息方面伸展。可见，信息科学的产生是与大系统理论、系统工程的发展、系统科学的形成分不开的。正是在这个意义上，信息科学是与系统科学一道发展起来的，两者相辅相成。它们在实质上都以大系统和复杂系统为研究对象，不过研究的角度不同。这意味



着人类认识史和科学史已进入一个新的历史时期，要在新的水平上重新对整体性的对象加以研究，这类对象一般不仅规模巨大，联系复杂，而且具有不确定性的特点，因而研究起来特别困难。与此相联，方法论问题也就特别突出。

现代科学技术与社会生活的发展决定了信息科学研究对象扩展的必要性，而新的科学方法的出现则使这一扩展具有可能性。我们知道，研究对象、科学理论和科学方法之间存在着密切的关系。不同的研究对象会出现不同的理论，而不同的理论则派生出以之为根据的方法，对于开创性的工作来说，针对新的研究对象，有时还必须提出开创性的新的方法来进行探索。

从方法论上进行分析，申农信息论的建立主要有两点：第一是运用了科学抽象与类比的方法，将消息、信息、情报等在各个不同科学领域中的具体概念、加以类比，舍去个性，抽象出共性，从而得到具有一般的、共同的科学概念——信息，使得对之加以形式化与数量化的处理成为可能；第二，是针对信息的随机性特点，采用了统计数学的方法。对概率论与随机过程的研究，在数学上已取得了一定的成果，并被作为一种数学工具与方法引入物理学。在 40 年代，针对控制系统与通信系统所具有的不确定性这一突出特点，无论是维纳还是申农等都不约而同地从物理学和数学中移植了这种统计方法，这在方法论上是一个重大的突破和创新。以至可以这样说，科学抽象与类比的方法为申农信息论的建立提供了必要的条件，而把不确定性简称为随机性，并用统计方法加以处理，则可说是具有充分条件的特点。可见维纳等人重视方法论的研究并不是偶然的。在科学技术上早就使用了科学抽象的方法，而类比与统计方法，尤其是把统计方法用于这一领域，则是一种具有开创性质的工作。如果没有统计方法与类比方法，申农信息论的创立简直是难以想象的。

就信息论的一些基本问题及其之后的进一步研究进展来看，信息理论不断向前发展的结果是由于引入了一些新的方法。

### 一、在原有的数学工具基础上，在已有的有关参量之外引入新的参量

在贝里斯 (M. Belis) 和高艾斯 (S. Guiasu) 提出的“量—质”统一量度信息中，除申农统计模型中的随机事件集及遍于事件集上定义的概率分布以外，还引入了表示实效性的“有效分布”这一新的重要参量。同样，在相对信息中，则把信息与产生信息的信源和使用信息的观察者作为三位一体的一个整体对象，在数学上是作为一个三元组来处理的，由此才导出与观察者相关的四个变量以及相对于另一观察者的相对信息。这些工作表明，由于现有研究对象较原有的复杂，才不得不引入相应的新的参量。一般说来，这种方法往往用于在科学技术的发展中推广某项新成果，但是它也要受到一定条件的限制，如果涉及到的不仅是信息的量，而且还要考虑信息的质和其它一些主观因素时，就必须研究和探索出新的方法，否则就难以获得新的进展。

### 二、采用非概率的方法，突破了在随机性之上建立的统计数学方法

1965 年，查德开创了模糊集概念与模糊数学的方法，这样就可以用非概率的、非统计的方法建立一种在性质上全新的信息模型，着重于从质的方面对信息进行定性分析，这相对于申农的信息论的定量分析来说，是一项重大的突破，而这一突破也是与整个系统科学不仅要从量上，也要从质上进行分析是一致的，反映了人类认识史与科学史进入到一个新的时期。必须指出，

模糊信息与模糊熵的概念不能理解为与统计信息、语义信息和有效信息相并列的另一种信息。应该看到，一切信息一般说来都具有模糊性的特点，因此，语义信息和有效信息也同样具有模糊的性质，只是在特定条件下才把它们加以简化，才作为确定的、统计的对象来处理。在一定意义上，我们可以说，从信息论到信息科学的研究就是从对统计模型的研究进入对模糊模型和统计模型的研究。

统计方法和模糊集方法是处理熵函数与信息量的两种不同方法。各自分别建立在随机性与可能性的基础上。可能性的概念是查德于 1978 年在《模糊集合——可能性理论基础》一文中对模糊性的一种解释。这种可能性测度是对普通测度特别是概率测度的一种推广，为我们对信息的定性分析提供了一种新的数学工具与方法，因而在方法论上有所突破，使得在从信息论到信息科学的进展中，由统计方法进入到统计方法与非统计方法（包括模糊数学方法）的阶段。同时还必须看到，不确定性的研究是十分复杂的，随机性与模糊性都不过是不确定性的某些情况，因此可以预言，为了从广度上与深度上进一步对不确定性加以研究，还必然会出现一系列新的数学工具与方法，这是因为在现代科学技术中，不确定性的研究在方法论上具有特别突出的意义。如果说吉布斯等人首先在物理学中考虑随机性这一类特定的不确定性的话，那么，维纳与申农则把这一方法推广到控制与信息的领域。但是到了 60、70 年代，这种方法的局限性日益显露。即使现代出现的模糊数学方法，也有严格的限制，仅仅能处理某些特定的模糊性，而不是任何模糊性。尤其是随着现代科技的发展，在控制论、信息论、人工智能等新兴学科中陆续出现了许多新的方法，如控制的状态空间等等。这样，在方法论中就给我们以重要的启示，对于信息科学所提出的一系列的课题，有必要寻求新的数学工具与方法。

### 三、类比方法与方法的移植问题

信息概念的提出和申农信息论的建立是与类比方法分不开的。一方面，可以在相似性的基础上进行类比，通过科学抽象与概括得到信息的一般概念和共同规律；另一方面，特别值得注意的是推导出来的信息量公式与热力学第二定律的熵公式只差一个负量。从两个在当时看来是完全不同的领域中，所作出的数学推导在形式上几乎一致，这在当时是很不清楚的。之后信息论与理论物理学领域沟通起来，进一步明确了不定性与不定度的普遍意义。其后再经过普里高津、哈肯、爱根等人的工作，提出了耗散结构、协同学理论等新的成果，不仅对普通系统论作出了一定的量的处理，而且把理论物理学与理论生物学沟通起来，带来了一系列重大的哲学和方法论问题，这在方法论上不仅反映了整体化的趋势，而且深刻地揭示了统一性与相似性的内在联系。类比方法与方法的移植，其客观的根据正是在于这种统一性和相似性。这一科学事实及其进展一方面把信息论、物理学、生物学、普通系统论等学科联系起来，另一方面也有助于理解本来在物理学领域中运用的统计方法能够移植于信息论的研究之中。

可见，类比方法和方法的移植对于信息论与信息科学的研究，乃至对于现代科学技术的发展都起了十分重要的作用。然而在一些情况下，这种方法也会导致不合适的或者错误的推断，为了对之加以判别，除了有赖于科学技术本身的发展和实践的检验以外，还要对类比方法与方法的移植的客观依据进一步加以探讨，识别类比和方法移植的界限与条件，找出判定的法则及规

律，这正是方法论研究需要解决的问题。

### 第三节 信息方法及其应用

科学的综合发展及实践证明，信息论与信息科学已进展为具有一般方法论意义的学科。当今社会，信息概念已经渗入各个科学领域，信息方法已广泛用于各门自然科学和社会科学，具有普遍的意义。

1948年，控制论的创始人维纳在其著作《控制论》一书中，曾用过“信息的方法”这一术语，他说，“任何组织所以能够保持自身的内在稳定性，是由于它具有取得、使用、保持和传递信息的方法。”然而维纳对这一术语的含义并未作明确的解释，他所说的“信息的方法”可能是指具体的取得、使用、保持和传递信息的技术方法或具体的通讯手段与工具，其目的是为了保持系统自身内在的稳定性。

1955年苏联的索波列夫在《控制论的若干基本特征》一文中，明确提出信息的观点具有一般方法论的意义。他指出，在控制论中信息这个概念有着非常广泛的意义，信息是一个新的观点，在自然科学中引入信息的概念和信息量的统一量度的概念，就可以从一个新的统一观点研究自然界中物体相互作用的各种极不相同的过程，并且可将各种不同的对象当作信息的处理与传送的系统，而从一个统一的观点进行研究，因此在方法论上具有巨大的意义。

1979年我国出版的《自然辩证法讲义》明确地介绍了信息方法，并把它作为一种具有普遍方法论意义的科学研究方法加以阐述。

#### 一、信息方法及其特点

信息方法就是运用信息的观点，把系统看作借助于信息的获取、传递、加工、处理而实现其有目的运动的一种研究方法。它把对象看成是一个信息流动系统。通过对信息流程的分析与处理，达到认识对象的发展运动的特点和规律，以及利用这些特点和规律达到人类的目的。

信息方法不同于传统的经验方法，它有其自身的特点，这是由信息科学本身的性质所决定的。信息科学是一门多边缘的横断学科，具有综合性的特点，它研究的是从各种复杂的不同质的系统中抽取出来的统一的信息关系。

信息方法和特点是用信息概念作为分析和处理问题的基础，它完全撇开对象的具体运动形态，把系统的有目的运动抽象为一个信息变换过程。这种信息变换的过程如图3—1所示。

正是由于信息流的正常流动，特别是反馈信息的存在，才能使系统按预定目标实现控制。而且两个系统的相互联系必须通过信息通道进行信息交换才能实现，如图3—2所示。

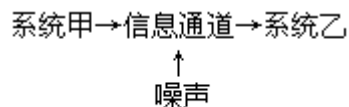


图 3-2 信息交换过程

维纳在控制论研究中处处运用这种方法。他在研究人与外界相互作用关系时指出：人通过感觉器官感知周围世界，在脑和神经系统中调整获得的信息，经过适当的储存、校正和选择等过程进入效应器官。这些效应器官反作用于外部世界，同时，也通过像运动感觉器官末梢这类感受器，再作用于中

枢神经系统。运动感觉器官所收到的信息又同已储存的信息结合在一起，影响将来的动作。他认为“接收信息和使用信息的过程，就是我们适应外部世界环境偶然性变化的过程。”人类要有效地生活就要有足够的信息。人类正是凭信息感觉器感知周围的一切变化，根据这种变化，由大脑作出决定调整自己的行动，改变着自己与自然界斗争的策略。显然，他始终着眼于信息，把人作用于外界的行为归结为信息和信息的反馈过程。正如他所言的“任何组织所以能够保持自身的内稳定性，是由于它具有取得、使用、保持和传递信息的方法”。

信息方法的另一特点是直接从整体出发，用信息联系、信息转化的观点，综合研究系统的信息过程。信息方法与传统的经验方法不同，它不是割断系统的内在联系，用孤立、静止、局步的观点去研究事物，也不是在剖析的基础上进行机械的综合，而是用联系、转化的观点，综合研究系统运动的信息过程。用这种信息方法对复杂事物进行研究时，不需要对事物的整体结构进行解剖分析，而是从其信息的流程加以综合考察，就可获得有关系统的整体性的性能和知识，而用简单的剖析和综合方法则不可能取得。因此，信息方法是现代科技领域中研究事物的复杂性、系统性、整体性的一种重要方法。

## 二、信息方法的作用

### （一）信息方法揭示了机器、生物有机体和社会不同物质运动形态之间的信息联系

在现实存在的许多复杂系统中，如技术系统中的通讯、控制、火箭、导弹制导、电子计算机等；人或生物系统中的生命现象、各个感官与外界的接触、神经中枢与各个感官之间的联系、大脑的记忆与指挥；人类社会系统中的生产过程和经济管理、交通管理等等，看起来它们之间的物质构成的运动形态都极不相同，用传统方法很难发现它们之间的内在联系，而利用信息方法进行考察，则可以把它们看成是信息系统，这些系统都存在着信息的接收、存储、加工处理和传递的过程。正是由于这一信息流动过程，才使系统能维持正常的有目的运动，从而揭示出它们之间的信息联系。

因此，运用信息方法可以揭示人脑与机器这两种截然不同的物质运动形态之间的对应关系和共同本质。人脑是由100~1000多亿神经细胞组成的，神经细胞可以处理兴奋和抑制两种状态，而电子计算机则是由许多人造神经元组成，相应地有接通或断开两种状态；大脑工作特征是利用神经脉冲，而机器可以利用电脉冲；机器与人脑都具有从外界获得信息、加工处理、传递信息的能力，它们存在着共同的信息联系。所以，可以把它们看作是一个信息变换的系统，这就为利用机器模拟并代替部分人脑的功能提供了科学根据。

### （二）信息方法为开展人体科学研究提供了基础

人体科学是把人这个生命有机体作为研究的对象，研究人体的结构和功能以及如何保护、发挥、发展人体功能的一门学问。过去西方医学界采用了一些物理、化学、生物学的观点和方法，探索机体的结构和功能及其规律性，但往往缺乏系统的观点，只重视分析和实验的方法，而没有找到精确的定量的描述人体结构和功能规律的科学方法。我国中医从总体上看具有系统辩证的观点，但往往停留在古老的、定性的、经验的阶段。而运用信息方法可以对性质截然不同，复杂程度极不一致的对象，给出统一的定量的综合指标，

利用这种定量指标可以综合地反映出以往认为是零乱的难以概括的医学实验资料，并阐明早先未曾揭示的新的规律。

人们已在实践中运用信息方法对生命活动过程以及实现这一过程的机体作出定量描述，并据此判断生命有机系统状态的稳定程度，以及这些系统的变化趋势是巩固、激活，还是瓦解、衰落。为此有人提出用信息熵函数来表示信息源的不确定程度，它在数量上等于信息量，在内容上却与它相反，用以确定信息源的不确定性和多样性。在生命机体中，这个熵函数可以反映其组织结构的复杂程度和有序状态，因为生命机体是一个开放系统，它必须与外界进行物质、能量、信息的交换才能维持机体的生存。当生命机体内部新陈代谢作用，使机体内部有序结构遭受破坏，这时信息熵就增加，同时它又不断从外界接收信息，在机体内部合成高度的有序结构，使信息熵降低。因此，人们可以通过测定这个熵函数确定生命机体的组织结构的复杂程度和有序状态，可以判断机体是否处于正常稳定的状态。可见信息熵在医学中具有重要地位。

用信息方法分析人体生命系统各部分之间，人体与外部环境之间相互影响、相互作用的信息联系，它们之间存在着信息接收、传递、存取、处理、输出，正是这股信息流才使得机体得到调节、控制，维持正常的生命活动。我们可以用一种简单的网络关系表示两个系统之间的联系，如图 3—3 所示。

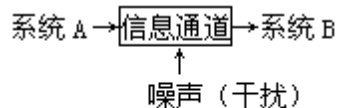


图 3—3 系统之间的联系

在正常情况下，系统之间信息传输系统具有足够的可靠性，使得发送端系统 A 发出的信息能准确无误地——被接收端系统 B 所接收。当通道中噪声干扰增大，也就是疾病、衰老以及反常的外界因素增大，使机体正常状态或细节机制受到破坏，造成系统 A 即信息源发出的信息，或在通道内有噪声干扰情况下传递的信息发生畸变或丢失，使接受端收不到——对应的准确信息。甚至接受端系统本身受到外界干扰的破坏，从而会增加疾病。运用信息方法，不仅能定性判断疾病是由于正常信息的缺乏（某种内分泌疾病），异常信息的干扰（传染病），信息代码的错乱（分子病、遗传病），或是信息通道的阻塞，使机体失去控制所致。还能进一步给出在这些情况下，信息关系的定量描述，确定系统的抗干扰能力，从而对疾病的轻重程度，机体的愈合能力作出数量上的分析。信息方法还对分子生物学的创立起了重大的作用。40 年代美国德尔布鲁克等人组织了噬菌体研究小组，创立了“信息学派”。他们再一次证明 DNA 是遗传信息的载体，它携带了生命自我复制的全部信息。1953 年这个学派的成员沃森提出了 DNA 的双螺旋结构的分子模型，标志着分子生物学的正式诞生。可见，信息方法是研究生命现象不可缺少的重要方法。

（三）信息方法揭示了某些事物运动的新规律，对过去难以理解的现象作出科学的说明

随着信息理论研究的进展及其在各个领域的推广，人们可以用信息的观点对某些现象作出科学的说明。例如，运用信息方法可以对某些生物群体活动的现象作出科学的解释，掌握活动规律为人类服务。俗话说：“禽有禽言，兽有兽语”，动物之间特别是群体活动之间存在着某种联系方式，以便使一

个生物个体影响另一些生物体的活动，这就是生物通讯。人们发现动物之间具有完整的发送和接收信号的通讯系统，它们可以利用气味、声音、不同运动姿态、色彩以至超声波、电磁场等多种信号传递信息。

许多生物群体具有严密的组织系统和完善的通讯系统。如蜜蜂，当侦察的蜜蜂发现蜜源后，回来用不同的舞蹈和发出长短、高低不同的声音报告蜜源的方向、距离以及花蜜的质量，然后工蜂都来嗅闻侦察蜂身上的花香气味，根据它所提供的信息去采蜜。这就是一个信息的接受、传递过程，离开了信息方法，则很难发现这一现象。

人们运用信息方法对动物通讯进行研究，记录了许多飞禽、走兽和昆虫的声音信号，并试图译释各种信息的意思，可以预料，不久的将来，人们将能用电子仪器指挥一些动物的活动，从而保护有益于人类的动物，诱歼有害动物或驱逐出一定的区域。

#### **（四）信息方法为实现科学技术、生产、经营管理、社会管理的现代化提供了思想武器**

人的认识世界和改造世界的实践活动存在着三个共同的流动过程，即劳动力组成的人流，生产资料、劳动资料等组成的物流，以及组织、计划、指导、协调、控制管理以达到预定目标的信息流，其中任一流通过程发生堵塞、中断，都将造成实践活动的破坏和停顿。而信息流的畅通是保证人流、物流畅通的前提条件，信息流调节着人流和物流的数量、方向、速度、目标，它驾驭人和物进行有目的的、有规则的活动。许多事实说明，信息方法对于实现科学技术、生产管理、社会管理的现代化具有重大意义。

例如，人们在进行科学技术研究的实践时，必须首先获得有关该项目的科学研究资料，然后对这些资料进行分析处理，从中得出必要的认识，产生相应的判断，作出计划，着手进行深入的研究，以期有新的发现与认识。用信息方法分析，即存在着一个获取信息 存储 加工处理 输出信息的信息流，并且可以运用现代信息技术使科研活动的信息过程更加畅通，实现科学研究的实践活动信息化，加速科研工作的进程，从而取得更大的效果。

目前这种信息方法的运用，已随着计算机的发展渗透到人们实践活动的许多领域，包括社会生活、经营管理等。因为任何企业或事业单位，只要存在着生产和科技活动，就存在着管理活动，只要有生产、科研和管理，就客观上存在着信息的流动。因此，应用信息方法进行科研、生产、经营管理，已成为各项事业现代化进程中极为重要的手段。

#### **（五）信息方法的使用提供了科学决策的基础**

决策过程实际上是一个信息的收集、传递、整理加工、变换的过程。决策正确与否，往往与能否及时、准确地获取足够的信息有着直接的关系。如果没有及时、准确、足够的信息，不可能发现问题，确定目标，提出正确的方案，进行科学的决策。因此，必须运用信息方法才能使决策达到科学化水平。

在科学决策的实践中，信息方法贯穿着决策过程的始终，不仅决策的全过程中需要运用它，就是决策的每一阶段、每一环节都必须运用它。如确立目标、制定方案、分析评价，哪一项都不能离开信息的收集、加工、存贮和传递。原始信息的收集，是事关决策能否正确的基础工作，它要求人们根据某种决策的需要，收集国内外的情报信息，然后对信息进行加工处理，去伪存真，去粗取精，使其成为符合决策需要的信息。并及时、准确地提供给决

策者，使信息得到有效的使用，从而作出科学的决策。可见信息方法是科学决策中不可缺少的方法，信息方法的使用为科学决策奠定了基础。

### 三、信息方法在信息学中的应用

信息论与信息方法已渗透到各门学科之中，信息学的研究和信息工作的开展也离不开信息论与信息方法在其实践中的应用。

#### （一）可以把信息系统看作是一个典型的信息流通系统

信息系统是一个开放系统，一方面，它将大量的社会信息资源搜集、聚集到系统中，另一方面，它又将经过加工、处理后的有序信息传递、提供给广大的社会用户。社会用户将使用信息的效果即反馈信息，传给信息系统，信息系统经过分析，再将系统中的有关信息加工成更加适合用户使用的信息，提供给用户。这样，从信息方法来看，一个复杂的信息系统就变成了一个十分简单的信息流通系统。

#### （二）在用户研究中引入信息流动的概念

信息用户从信息系统得到所需信息后，用户头脑中原有的知识结构在输入信息的激发下，形成了新的知识结构或组合关系，促进用户提出新的学术观点，写出有创见性的学术报告或专著，或者作出了某项决策，总之是产生了新的信息。这些新产生的信息可能又被作为原始信息输入到信息系统中去，进行加工、整理，然后再由信息系统把这些新信息传递给新用户，促进新用户再产生更新的信息。

#### （三）从信息论的观点看，信息过程就是一个信息加工、转换和传递过程

当信息人员拿到一篇含有  $I_1$  信息的文献  $D_1$  后，他将文献中的主题词与叙词表进行比较，其目的是进行加工与转换，然后制成带有代替信息  $I_2$  的替代文献  $D_2$ 。从  $D_1$  到  $D_2$  就是一种信息加工、转换过程。在检索过程中，信息检索人员也可通过  $D_2$  找到  $D_1$ ，用户则从文献  $D_1$  中获得所需信息  $I_1$ 。

#### （四）通过信息方法研究信息人员

信息人员作为中介，是连接用户和信息系统的信息资源的桥梁。信息人员在得到用户提问后，先将用户的问题信息加工、转换成信息系统可以识别的形式，输入信息系统，得到有关信息，再把这些信息直接或是再加工后提供给用户，并准备接受用户反馈信息。这样，就可运用信息方法研究信息人员，信息人员则被看作是一个信息流通过程。

## 第四章信息计量学方法

### 第一节信息计量学概述

#### 一、信息计量学的产生和发展

信息计量学是应用数学方法研究、描述信息过程、现象与规律的一门学科。“信息计量学”的概念是由德国昂·纳克(O.Nacke)于1929年在德国的《文献工作通报》上首次提出的。1980年FID成立了FID/IM(信息计量学委员会),并于1980年3月在法兰克福市召开了第一次国际信息计量学讨论会。信息计量学是数学、统计学与信息学广泛结合而形成的一个新的分支学科。它的主要内容是应用数学方法分析和处理信息传递的种种矛盾,从定量的角度分析和研究信息的动态特性,并找出其中的内在规律。

FID/IM为了实现创立之初拟定计划,具体开展了几项重要活动:1982年在前捷克斯洛伐克召开了“信息计量学”教学计划会议;1985年进一步评估、论证了“信息计量学”计划;1985年7月开始在印度出版《信息计量学通讯》(Informetrics Newsletter)。这些活动对于推动信息计量学的发展,使信息计量学逐渐得到国际信息学界的承认,产生了积极的影响。

1984年,英国著名信息学家布鲁克斯发表了两篇论述信息计量学的论文,讨论了信息计量学的一些基本理论问题。1987年,布鲁克斯在比利时召开的“文献计量学和信息检索的理论问题国际会议”上建议,将术语Informetrics包括在拟于1989年在加拿大召开的会议名称之内。这个建议得到了当时参加会议的信息学者的普遍支持和赞同。接着,1991年在印度邦加罗尔举行的第三次会议的名称就是“信息计量学国际会议”。由于这三次国际会议之后出版的论文集上都有In-formetrics标题,因此,国外一些著名的信息学专家把1987年看成是Informetrics被国际信息学界正式承认的一年。

在1987~1992年期间,许多信息学专家又发表了许多关于Informetrics的文章。Informetrics在北美和西欧非常流行,以致1987年以后的一些西方私人信息服务公司多改名为时髦的“信息计量学公司”(Informetrica Limited)。

在中国,对“信息计量学”术语大多持比较保守的态度,以“信息计量学”作为论文或辞典条目只有5~6次,而作为教材或专著的书名至今尚无。我国信息学界目前仍停留在文献计量学的研究时期,与国外信息学界正在实现的由信息学中传统的文献计量学研究向信息计量学研究的根本性转变,形成强烈的反差。

#### 二、信息计量学与文献计量学、科学计量学的联系

信息计量学的产生和发展过程表明,它是在传统的文献计量学及科学计量学的基础上演变过来的。因此,信息计量学与文献计量学、科学计量学是3门密切相关的学科,它们之间有重合、搭界的相同之处。

首先,在研究对象方面,信息计量学、文献计量学、科学计量学这3门计量学科都要以文献(主要是科学文献,下同)作为计量对象,即文献是这3门学科的重合、搭界之处。正如英国信息学家C.汉森所指出的那样,“科学、信息对许多人来说是文献的同义语。在信息计量学产生之前,信息学家一直以文献作为自己的计量对象;而科学计量学从它产生的那一刻起,就从



统计书目学及文献计量学那里吸收了不少以文献为计量对象的方法、定律、模型等成果。正因为如此，本世纪 20~40 年代对文献进行计量研究的科学家洛特卡、布拉德福和齐普夫，成了这 3 门学科的先驱人物。

其次，由于这 3 门计量学科都要以文献作为自己的研究对象，从而决定了它们在研究方法、研究内容方面也有重合、搭界之处。在研究方法上这 3 门计量学科非常类似，有时甚至完全相同，主要表现在以下几个方面：研究方法的数学基础都有统计学（包括分统计、抽样统计、多元统计）、集合论和数理统计、运筹学，包括线性规划、排队论等，代数、微积分学、向量分析、矩阵论、图论、聚类分析、模拟方法、模糊数学等；具体的研究方法都有引文分析（包括引文标引与引文者动机、引文网络图与引文矩阵、引文耦合和同被引分析、老化分析）、文献统计分析（包括集中与分散分析、分布规律解释）和数学模型法；计量模型都有词（字）频率模型（齐夫定律）、期刊等级模型（布拉德福定律）、倒平方模型（洛特卡定律）、老化模型、半衰期模型、拟合逼近方法。

在研究内容上，这 3 门计量学科的共同之处表现在以下几个方面：语言、词和词组的频率统计，包括自然语言主题与索引、印刷文献与电子文献；作者特征——根据论文数量或其它方法（如合著者次数）决定的生产率测度；出版源特征，如某一学科的论文在全部学术杂志上的最显著的分布；引文分析，包括作者、论文、团体、杂志、国家的分布，效用评价，学科间的同被引测度；文献失效或老化（由引文或使用进行测度）。

### 三、信息计量学与文献计量学、科学计量学的区别

虽然信息计量学与文献计量学、科学计量学有上述重合、搭界的相同之处。但是，我们仍然可以根据这 3 门计量学学科在研究对象、研究目的等方面的特定属性或侧重点，将其一一区别开来。

首先，信息计量学的研究对象大于文献计量学、科学计量学的研究对象。从根本上说，信息计量学的对象包括各种类型的信息（如数据、事件、事实、文本、文献等），其中既有正式交流的数据、文本、文献等，也有非正式交流的消息、事实、事件等。而文献计量学的对象则基本上是正式交流的文献，其研究对象范围远远小于信息计量学的研究对象范围。至于科学计量学的对象则包括正式交流文献中的科学文献（从这个角度看它比文献计量学的对象小）和非正式交流的科学事实、科学事件（从这个角度看它比文献计量学的对象大）。这 3 门计量学科在研究对象上的联系与区别，我们可用图 4-1 或图 4-2 表示。

其次，由于这 3 门计量学科的研究对象并不完全相同，决定了它们在研究方法、研究内容方面也有不尽相同之处。信息计量学的方法既包括确定性数学（如统计学），也包括非确定性数学（如概率论），而文献计量学和科学计量学的研究方法主要是确定性数学；信息计量学的定量处理方法既包括定量变量，也包括定性变量（我国学者郭俊义将其称为“广义量化”），而文献计量学和科学计量学的定量处理方法“特色”则是定量变量（“狭义量化”）。信息计量学的研究内容包括信息的定义与测度（它原来部分地被包括在“信息论”中，J.T——苏特克莱弗认为它应是信息计量学的重要研究内容）及信息检索性能测度量的类型与特征等，前者是文献计量学和科学计量学不能包括的，后者是科学计量学不能包括的。文献计量学的研究内容包括图书流通及内部使用、藏书布局与藏书建设等，这是信息计量学和科学计量

学所不能包括的；科学计量学的研究内容包括科学活动的评价指标、科技管理决策与科技政策，这些是信息计量学和文献计量学所不能包括的。

最后，更重要的是这3门计量学科的研究目的不同。信息计量学的研究目的是揭示信息本身在数量方面的规律性，它从属信息学的研究目的需要；科学计量学的研究目的是揭示科学活动在数量方面的规律性，它从属于科学学的研究目的需要。关于这3门计量学学科的研究目的从属于3门科学的研究目的需要，最先是前苏联学者提出的，后来布鲁克斯又作了进一步地论述。

此外应当看到，这3门计量学科之间的研究目的的不同，在一定程度上也决定了它们在研究方法、研究内容上有所不同；而这3门计量学科之间研究对象的不同，在一定程度上决定了它们在研究方法、研究内容上有所不同，这是来自两个方向的同一作用。

#### 四、信息计量学与文献计量学、科学计量学的学科关系

清楚了信息计量学和文献计量学、科学计量学在上述几个方面——尤其是在研究对象和研究目的方面的联系与区别，我们可以进一步地对这3门计量学科之间的关系进行划分。正如匈牙利著名科学家 T. 布劳温曾指出的那样，虽然“文献计量学和科学计量学的研究方法非常类似，有时甚至完全相同，但我们还是可以根据其研究对象和研究目的来区分它们。”

我们知道，学科之间是否存在相互关系，首先是根据它们的研究对象来确定。根据这3门计量学在研究对象上的联系与区别（见图4—1、图4—2），我们可以直接得到它们之间的联系与区别，如图4—3。

也就是说，从这3门计量学科的研究对象来看，它们之间的关系是信息计量学包括文献计量学和科学计量学，文献计量学和科学计量学是交叉关系。图4—3的阴影部分就是它们在研究对象、研究方法、研究内容方面的重合、搭界之处。

学科间的关系划分，除了考虑它们各自的研究对象这个根本性因素外，还必须进一步考虑它们各自的研究目的。科学中常常出现这样的情况，同一研究对象，由于研究目的不同，就会形成完全有别的学科。例如，物理化学和化学物理，它们有共同的研究对象，但由于它们的研究目的不同，前者属于化学分支，后者则属于物理学分支。因此，如果同一研究对象的学科，有完全不同的研究目的，我们就应该进一步对它们的学科关系进行划分。考虑这3门计量学科的研究目的，它们之间的关系就会发生位移，从而导致信息计量学不能完全包括文献计量学和科学计量学的新型学科关系。也就是说，由于信息计量学从属于信息学的研究目的需要，即作为信息学的一个分支，信息计量学只能部分地包括文献计量学和科学计量学；由于文献计量学、科学计量学分别从属于图书馆学、科学学的研究目的需要，即作为图书馆学、科学学的一个分支，文献计量学和科学计量学也只能部分地从属于信息计量学。

### 第二节 信息计量学方法

#### 一、信息计量学方法的特点

信息计量学作为信息学的一个分支，具有其内部必要的扩张力——科学逻辑结构的扩张力。同时，由于来自社会、科学技术、生产的需要，其外部压力与动力是巨大的。信息计量学本身也力求寻找并实现其具体应用，信息研究定量化的进展，又势必要借助信息计量学的方法，结果两者相得益彰。

信息计量学用作研究方法时，有种种特点及优势。

(一) 突出了信息研究的信息特色

信息研究今天已成为我国信息机构的一种重要服务方式。在世界上，只有我国和意大利是由信息研究所搞信息研究的。如果说 30 年前信息研究是信息部门开创的，并独家实践了 20 年的话，那么今天许多相关的软科学研究部门也竞相开展了类似的工作。同时，信息研究所采用的方法，也具有强烈的广普性。几乎所有的社会科学与大部分自然科学的研究方法，均可为信息研究所用，这早已在实践中得到证实。随之而来的便是信息研究的“信息学”色彩的逐渐淡化，这已为许多信息研究工作者所觉察。例如，趋势外推法是我国信息界和其它科学领域普遍采用的，如果我们不讲究信息研究的信息学个性，久之信息研究恐就难以维继了。如果信息学界以往没有刻意追求信息研究的信息学特征的话，那么，今天就十分有必要予以重视了。致力于信息学研究的特殊方法的开发，是突出其信息学本色的必由之路。如是，可能要首推信息计量学方法。

人人皆知，文献是信息计量学的根本，以信息计量学为研究方法，务必以文献或其某些规律为依据。这些既是信息学机构收藏与管理的对象，也是信息学研究人员熟悉并能灵活运用知识。信息学计量学方法的采用，真正体现着信息研究人员在施其所长。其它领域里的研究人员（如社会学）也可能采取这种方法，但那是他们的移植运用，其成果也应属于信息研究范畴。信息计量学充实了信息研究的信息学特征，突出了信息研究的信息学本色。

例如，1980 年初，日本化学工业社组织社会工业与科技界人士，展望 80 年代化学工业的前景。日本科学技术信息中心信息部的小森隆，就与众不同地采用信息计量学方法，完成了这一调研任务。

他以 JOIS—S 理工数据库中的 622405 篇文献为来源资料，计数这些文献中出现的塑料、橡胶、纤维等关键词的频次。其中主要合成纤维关键词的频次如表 4—1 所示。

表 4—1 有关纤维的词频

关键词	频次	关键词	频次
丙烯纤维	315	人造纤维	333
聚酰胺纤维	705	醋酸纤维	65
聚脂纤维	184	三醋酸纤维	25
聚烯烃纤维	189		

据此，小森隆断言 80 年代聚酯纤维仍将有压倒一切的优势，而醋酸纤维则越来越受冷落。这一信息计量学方法的研究结论，是否与实际情况吻合呢？表 4—2 是有关世界合成纤维生产量的实际统计表。通过比较可以证实，出现于文献中的各种纤维的词频，是反映实际情况的，聚酯纤维产量比重之大与其词频之多是对应的。由此可见，从信息计量学方法取得的研究结果，完全可以正确地反映客观趋势，堪与其它研究方法相媲美，可誉为信息学研究人员所特有的一种研究方法。

表 4—2 世界合成纤维生产编译数据

年度	聚酰胺		聚脂		丙烯酸纤维		其它	
	产量	%	产量	%	产量	%	产量	%
1971	3450	37	3450	37	1860	20	560	6
1972	3790	37	3996	39	1946	19	512	5
1973	4054	35	4749	41	2317	20	463	4
1974	3845	34	4863	43	2149	19	452	4
1975	3522	33	4803	45	2028	19	320	3
1976	3904	32	5490	45	2440	20	366	3
1977	3971	31	5893	46	2434	19	512	4
1978	4213	31	6251	46	2713	20	408	3

## (二) 能其它方法之所不能

信息研究属于科学实验范畴，但它与一般科学实验是有所不同的。科学实验是探索未知的活动，如果用 K.波普的哲学术语来说，那是开发“世界 1”的。信息研究是总结已知的知识，是分析与综合“世界 3”的。科学实验有可能碍于客观条件，如实验装置与设备，而难以进行，那么信息研究则限制甚少。凡是有知识之处，均可见信息研究之功。如果不能进行信息研究时，可以断言是由于研究方法的困囿，缺乏有效的方法对已知的知识加以总结与剖析，信息计量学所展现的方法论性能，就令人大有另辟蹊径之感，它能迂回地解决问题。

在信息计量学中，有一个越来越为人们所瞩目的分析技术——引文分析。引文分析可以由表及里地探索事物（只要有文献表现的话）的结构和规律。正像人们可以利用分泌物来探测有机体内脏器官的健康情况一样。引文分析就是利用著文所引用的文献，对著文所体现的内容进行分析与综合。有时著文所反映的内容，对研究任务来说不是荫蔽的，就是难得要领的（仅指其对特定信息研究而言），当其它研究方法不能奏效时，引文分析有可能建立奇功。从引文分析来看，引文是著文某些特征的线索。一篇著文的引文可能无关宏旨，但一个学科一定样本容量的引文，却可以挖掘出许多深层信息，因此，引文分析作为信息研究的一种方法，有其它方法的绝对不可替代性，有其它方法所不能之功效。

例如，考察并比较世界各国与地区的基础研究水平，如有可能的话最好还要排出顺序，显然这是一个比较繁琐的研究课题。若选择其它方法难以深入，要进行定量研究就更困难了。匈牙利信息科学和科学计量学研究所（Information Science and Scientometric Research Unit）的 T.布劳温（Braun）等人，就用信息计量学中文献计量学之引文分析法，使这个似乎无从入手的课题，迎刃而解了。

他们选择了《SCI》为资料来源，统计了 1978~1980 年三年间 107 个国家和地区的有关几个指标。它们是：所有科技领域里的出版物总量；出版物年度增长百分比；每件出版物实际平均被引率（内含 1979~1980 年引用 1978 年，1980~1981 年引用 1979 年，1980~1981 年引用 1980 年）；每一件出版物的期望被引率，这一指标是根据受引期刊平均被引率而计算的；相对被引率，布劳温等人规定，由直接被引文量计算出来的结果为实际被引率，而相对被引率则是

实际被引率与期望被引率之比。他以这 5 个指标，考察了全世界 107 个国家和地区的基础研究水平——如表 4—3 所示——我们仅摘录了 18 个国家的有关数据。

应当说这 5 个指标基本上勾勒了各国学术水平状况，其中既有“量”的指标（如出版物总量），也有“质”的指标（如各种被引率）。当然这 5 个指标能否完整准确地反映各国和地区的学术水平，那是可以讨论的。但除引文分析外，其它方法均不见得能取得如此完善的结果。

表 4—3 一些国家的基础研究水平及其对比

续表

国家或地区	出版物总量		出版物年增长率		实际平均被引率		期望被引率		相对被引率	
	篇	序位	%	序位	次篇	序位	次篇	序位	次篇	序位
西班牙	6510	21	17.54	16	1.53	26	2.18	28	0.70	41
中国	2457	32	62.13	1	0.43	95	0.91	103	0.47	83
罗马尼亚	2015	36	-5.74	89	0.84	63	1.21	89	0.69	44
墨西哥	1736	37	9.27	37	1.93	19	2.56	12	0.76	36
伊朗	720	43	-30.04	104	0.76	75	1.71	54	0.45	85
伊拉克	517	48	3.77	61	0.67	79	1.56	63	0.43	88
尼泊尔	26	90	49.90	5	0.81	68	1.02	101	0.77	21

布罗温等人以《SCI》摘储为标准，可比性是很充分的，但是《SCI》本身既有学科的偏向（生命科学比重大），也有语种偏向（英文文献比重大）。当然这是布罗温等人无法克服的。这里我们意在突出引文分析的方法论优势。

### （三）计量元素好辨别、易收集

实施信息计量学方法时，所选用的计量元素大都显示了文献的某些特征，一望而知，一索即得，勿需更多的变换与分解作业。信息研究所采用的任何一种方法，开头的步骤都是：“拟定目标，收集资料”。这对信息计量学方法而言，有许多供应数据的资料库，所需数据明晰可鉴，极易获得，例如《SCI》、《SSCI》等。即便手边没有这类检索工具，而是亲自动手统计的，所用的计量元素也是好查找的。一般说来，文献的某种特征以及与文献相关的某些量值，均可作为“计量元素”，归纳起来如表 4—4 所示。

到目前为止，信息计量学方法约有 20 余个计量元素，可测项多达 50~60 个。致使采用信息计量学为研究方法时，有许多可借鉴的计量入口。表 4—4 中所列的各项都出自文献或与文献的相关量，都有显著的标志，也是固定的量值，好查、易收集，这是一个不容忽视的便利条件。

表 4—4 文献计量学方法的计量元素

续表

元素类别	元素名称	计量的主要内容
检索工具	13. 类目	数量
	14. 文献条目	数量、类型等
	15. 索引词	索引数量、类型等
其它	16. 分类号	数量、被标引的频次、标记制度等
	17. 读者	数量、构成、阅读方式、借阅周期与习惯等
	18. 复印件	数量、语种、年度、学科等
	19. 文献载体	类型、流通范围
	20. 机构	性质、数量、服务方式及对象等

例如，激光技术是一个发展迅速的技术领域，激光应用也日趋普遍。那么 1988 年前后激光应用的重点是什么呢？完成这类课题的普通调研方法不可谓没有。但是，科技文献是科技活动的真实记录，信息计量学方法是可以充当此任的。文献以文摘杂志为背景，以文献条目为计量元素，考虑 1977 ~ 1983 年间文摘摘储不同类目文摘条目数量，以此来判断激光应用的重点，如表 4—5 所示。

从表 4—5 可知：1977 ~ 1980 年间激光应用的重点是光通讯。但是激光干涉计量、激光测量、激光聚变、化学应用与激光热处理等应用，也日益为人们所重视。自 1981 年起便取代了光通讯应用的首席地位。

这个实例有力地说明，采用信息计量学作为信息研究方法真可谓简单易行。本例所需要的不同类目下的文摘条目是很容易获得的，这些条目形态分明，排列醒目，便于开展统计作业。

#### （四）研究者毋须有深广的专业背景知识

科技信息研究是专业性很强的分析与综合工作，它要求从事信息研究工作的人员有深广的专业背景知识。经验证明，进行某一专题技术信息研究时，如果研究人员缺乏该专业知识，对该专题技术若明若暗，那是很难有效进行信息研究的。有许多研究方法恰恰是因为研究人员专业知识不足而不能采用。另外，专职信息人员欲研究的课题是多方面的，所熟悉的专业背景知识却是有限的。信息研究人员专业背景知识的局限性，可因采用信息计量学方法而得以缓和补救。这是因为应用信息计量学方法不要求具备深广的专业背景知识。

信息计量学方法不涉及过多过细的文献内容，表 4—4 所列的计量元素，即使没有专深的背景知识也是可以认识和掌握的。在一般情况下，只要具备标引专业文献（指主题标引）的知识，就能较好地运用信息计量学方法。在进行某些新技术、高技术的信息研究时，由于缺乏足够的专业知识，难以采用其它研究方法，信息计量学是一种很好的替代措施。正因为如此，信息计量学方法常常成为信息研究人员乐于采用的方法。

表 4—5 1977 ~ 1983 年文摘条目分类（数字系所占%）

年度 顺序	1977	1978	1979	1980	1981	1982	1983
2	通讯：光 纤与波导 8.40	通讯：光 纤与波导 10.16	通讯：光 纤与波导 9.49	通讯：光 纤与波导 9.69	测量技术 6.90	制造业： 焊接热处 理	干涉计量 5.99
3	测量技术 6.41	通讯：集 成光学 4.70	测量技术 4.80	通讯：集 成光学 5.21	通讯：光 纤与波导 6.11	激光聚变 5.64	
4	分离同位 素 5.25	激光测量 技术 4.60	制造业焊 接热处理	制造业焊 接热处理	通讯：元 器件设计 5.38	通讯：元 器件设计 5.89	化学应用 5.05
5	通讯：集 成光学 4.15	激光聚变 4.49	化学应用 4.39	测量技术 4.53	激光聚变 5.05	化学应用 4.67	通讯：光 纤与波导 4.94
6	全息：记 录系统 3.92	大气传输 3.70	多普勒测 速仪 3.84	激光聚变 4.43	干涉计量 4.83	干涉计量 4.67	通讯：元 器件设计 4.76

例如，欲考察英国海流学（Ocean Current）科学研究采用的观察技术及其变化情况，由于没有更多的海流学科学知识，很难采用其他方法。若采用信息计量学方法就可少受海流学知识缺乏的牵制，能迅速地开展工作。

表 4—6 40 种期刊不同观测技术的文章数量

年度 观测技术	1973/1974	1975/1976	1977/1978	1979/1980	1981/1982	1973 ~ 1982 合计
船测	51	64	93	100	111	419
卫星	6	5	14	27	16	68
（漂移浮标）	1	1	6	10	4	22

年度	1973/197	1975/1976	1977/1978	1979/1980	1981/1982	1973 ~ 1982
观测技术	4					合计
( IR 图像 )	4	4	7	15	10	40
( 测高法 )	-	-	-	2	2	4
( 摄影法 )	1	-	1	-	-	2
船测	7	4	6	4	5	26
陆地、栈桥、冰	-	4	4	1	1	10
观察技术合计	54	71	102	109	114	450
实验性论文	1	6	7	9	3	261
理论文章	54	81	88	136	156	515
总计	109	158	197	254	273	991

首先，请英、法、美、德国等国家一流海流学学者，确定 40 种有关该学科的重点科技期刊。进而从中选出 991 篇海流学方面的文章，区分它们所采用的观测技术，按年度汇总如表 4—6 所示。同时还从 1976 年开始，对当年和回溯 3 年发表于上述 40 种期刊上的文章，进行引文分析，即利用《SCI》统计不同观测技术文章被引频次及平均每篇文章被引次数(即表 4—7 中括号内的数字)。

表 4 - 7 不同观测技术文章被引用数量

年度 观测技术	1976	1978	1980	1982	合计
船测	196 (1.70)	348 (2.22)	291 (1.51)	229 (1.09)	1084 (1.57)
卫星	13 ( 1.18 )	59 ( 3.11 )	67 ( 1.63 )	55 ( 1.28 )	194 ( 1.70 )
船测	14 ( 1.27 )	23 ( 2.30 )	15 ( 1.50 )	12 ( 1.33 )	64 ( 1.60 )
陆地、栈桥 冰	- (-)	13 (1.63)	15 ( 3.00 )	3 ( 1.50 )	31 ( 1.63 )
观察合计	199 (1.59)	381 (2.20)	323 (1.53)	255 (1.14)	1158 (1.58)
实验性文章	5 (0.71)	18 (1.38)	21 (1.31)	20 (1.67)	64 (1.36)
理论文章	202 (1.50)	331 (1.96)	280 (1.25)	315 (1.08)	1128 (1.38)
总计	406 (1.52)	730 (2.06)	624 (1.38)	590 (1.12)	2350 (1.47)

从表 4—6 可以发现，卫星观测文章日益增多，从 1973 年 ~ 1974 年的 6 篇，增加到 1979 ~ 1980 年的 10 篇。然而在 1981 ~ 1982 年间发表的论文中，



114 篇观测性文章中有 111 篇使用了船测技术。尽管卫星遥感技术在海流观测中有增加的趋势，但也仅有 17% 左右（由文章体现出来的），船测目前仍是主要的一种常用观测技术。

表 4—7 就更强化了以上的判断。1982 年被引的 255 篇观测文章，有 229 篇船测技术。平均的被引率又进一步证明，1978 年卫星遥感技术有相当大的影响效果，每篇文章平均被引 3 次以上，但 1982 年又有所下降。可见目前遥感技术虽然有所应用，但仍需大量船测技术进行补充和辅证。

这样一个常人甚感生疏的课题，由于采用了信息计量学方法，就可以发现英国海流学观测技术的变化与现状。

## 二、信息计量学方法的基础

我们在实施信息计量过程中，先后确立了几个信息定量规律。这些规律构成了信息计量学的基础，信息计量学方法主要是依照这些定律实现量化的。

### （一）信息增长定律

信息增长定律是描述信息数量随时间而有规律地增长。令  $F$  表示信息的数量， $t$  表示时间，则信息增长定律的数学表达形式为：

$$F = f(t)$$

式中  $f(t)$  的总趋势满足  $t$  增大时， $F$  也相应增大。描述信息增长规律的主要函数是：线性函数、指数函数、逻辑曲线函数等。其中以 D.J. 普赖斯建立的指数增长定律最为著名：

$$F(t) = ae^{bt} \quad (1)$$

式中， $F(t)$  为某年 ( $t$ ) 的信息累积数量； $t$  为时间（以年为单位）； $b$  为信息持续增长率，即每一年的信息增长率。

关于指数增长定律，另一个值得注意的量就是翻倍时间，即在一个固定的时间内，信息量增加了一倍，即

$$d = \ln 2 / b \quad (2)$$

例如： $b=0.1$  时，信息量翻倍时间为  $d=0.693/0.1=$

指数增长规律只有在没有限制或干扰的情况下才会出现，如果受到智力的、物质的和经济的限制，普赖斯指出信息增长更趋于逻辑曲线。俄罗斯学者弗勒杜茨和纳利莫夫在大量研究的基础上，提出了著名的逻辑曲线方程式：

$$F(t) = \frac{K}{1 + a e^{-bt}}$$

式中， $F(t)$  表示  $t$  年的信息累积量； $K$  为  $F(t)$  增长的最大值； $a$  与  $b$  为参数。

信息是科学技术成果主要表现形式之一，所以信息的增长可以直接或间接地反映出科学技术以及相关事物的现状与前景。

### （二）信息老化规律

对于科技信息来说，不计其物理形态上的破损、载体的变质，就它们所记载和表现的科学内容而言，也会令人有逐渐过时之感。但要说清楚信息老化的确切含义还是不容易的。我们大体上可以说，随着时间的流逝，文献所载的信息逐渐过时，以致于文献本身不再被使用。对于信息老化的量度主要有 3 个指标。

1. 半衰期：所谓半衰期是一个时间概念，意指在该时间内发表了某一学

科或领域正在被利用的全部信息的一半，或者目前所利用的信息的一半是在该时间内发表的。例如，以“信息技术”信息而言，它们的引文一半是1988年以后发表的，那么就可以说它的半衰期为4年（以1992年计）。C.A.西格摩（Segmour）又称半衰期，为信息的“中值引文年龄”。

2. 普赖斯指数：这是普赖斯提出的一个衡量信息老化程度的指标。规定年限不超过5年的信息引文数量同引文总量之比作为指数。依据这一指数，可将所有被利用的信息分成两大类：档案性信息、现时应用的信息。年龄超过5年仍被引证的信息则属于现时应用的信息。

3. 期刊有益性系数：这是B. C.布鲁克斯1970年提出的一个衡量期刊信息老化的指标，意指可望在某种程度上被科学家引用的文章数。

有关信息老化的数学模型主要有：

(1) 负指数模型：1970年布鲁克斯提出：在一个确切时间内，被引用的论文与它的发表年度呈负指数函数关系。即

$$C(t) = Ke^{-at} \quad (4)$$

式中， $C(t)$  是引用的信息之中  $t$  年以前所发表的论文数量； $K$  为常数，随不同学科而异； $a$  表示老化率 ( $a < 1$ )。

(2) 巴尔顿—凯普勒方程：美国某冶金图书馆工作人员 R.E. 巴尔顿 (Bwtor) 和物理学家 R.W. 凯普勒 (Kebler) 联合提出了描述信息老化的方程式：

$$Y = 1 - \left( \frac{a}{e^x} + \frac{b}{e^{2x}} \right) \quad (5)$$

$$a + b = 1$$

$Y$  为一定时期内，总的引文比率 ( $Y < 1$ )； $x$  为时间（以10年计）。显然，令式(5)中  $Y = 1/2$ ，这时的  $x$  即为半衰期  $X_H$ ，

$$\frac{1}{2} = 1 - \left( \frac{a}{e^x} + \frac{b}{e^{2x}} \right)$$

$$(e^x)^2 - 2a(e^x) - 2b = 0$$

$$e^{X_{H1}} = a + \sqrt{a^2 + 2b}$$

$$e^{X_{H2}} = a - \sqrt{a^2 + 2b}$$

以及

$$X_{H1} = \ln(a + \sqrt{a^2 + 2b})$$

$$X_{H2} = \ln(a - \sqrt{a^2 + 2b})$$

我们取出  $X_{H1} = \ln(a + \sqrt{a^2 + 2b})$ ，则同不同的  $a$ 、 $b$  值，可以得到不同的半衰期  $X_H$  值，由于  $b = 1 - a$ ，所以式(6)可写成

$$X_{H1} = \ln(a + \sqrt{a^2 + 2 - 2a})$$

(3) 莫蒂列夫修正式：俄罗斯信息学家莫蒂列夫用大量数据进行  $X^2$  检验，并对式(5)进行了如下修正：

$$Y = 1 - \left( \frac{a}{e^{x-0.1}} + \frac{b}{e^{2x-0.2}} \right)$$

莫蒂列夫的修正式提高了巴尔顿—凯普勒方程式的精度。

(4) 传播老化模型：1979年 A.阿威里麦斯仿物理学中的传播理论，推测了科学成就的自然衰减，他提出如下信息老化率：

$$V_t = [2(\pi at)^{\frac{1}{2}}]^{-1}$$

这些数学模型都能在一定程度上计算出信息老化趋势，但是这些模型需要有大量的前提条件，所以它们的应用并非是任意的。

### (三) 论文作者分布规律

关于科技信息作者分布规律的描述方式是很多的。它们的主要目的是考察科技人员的著述能力与信息量之间的关系。其中最著名的是洛特卡定律。

1926年末，美国一家人寿保险公司的统计员 A.J.洛特卡，以《化学文摘》和《物理学史一览表》为来源资料，考察了1907~1916年10年间的累积索引。他得出撰写 X 篇论文的作者相对频率 y(x) 的公式：

$$y(x) = \frac{C}{x^n}$$

当年洛特卡说：“这两个例子表明的指数近似等于 2.0”。于是，式(10)在 12 年之后，被 G.K.齐普夫称为“倒平方定律”。但是后人的继续研究表明，指数 2 仅是一个特例。1974年，J.维拉奇(Vilady)发现，对不同的学科而言，n 可以从 1.2 浮动到 3.5 以上。

式(10)中的  $C=1/(\sum_{x=1}^{\infty} 1/x^n)$ ，而  $\sum_{x=1}^{\infty} \frac{1}{x^n} = S(n)$

亦即  $C=1/S(n)$ 。S(n) 被认为是黎曼 ζ 函数，当  $n > 1$  时收敛，当  $n \leq 1$  时发散。若 n 为偶数时， $C=6/\pi^2=0.6079$ 。然而当 n 为任意数时，C 值不能按上式计算。有人提出近似计算公式为：

$$C=1/\sum_{x=1}^{\infty} 1/X^n$$

其中

$$\sum_{x=1}^{\infty} 1/X^n = [\sum_{x=1}^{p-1} 1/X^n + 1(n-1) \cdot P^{(n-1)} + 1/2P^n + n/24 \cdot (P-1)^{n+1}]$$

普赖斯对文献作者分布规律有过许多论断（其实只能算假设），其中最著名的是“平方根定律”——“撰写全部论文半数的高产作家的数量，等于全体论文作者的平方根”。

$$\sum_{m+1}^l n(x) = \sqrt{N}$$

式中，n(x) 为写 X 篇著作的科技论文作者人数，l=，即在一次取样中科技人员撰写最多的论文数量，N 为撰写论文作者的总人数。m 可按下列条件确定：

$$\sum_{n=2}^m X(x) = \sum_{n=m+1}^l X(x)$$

### (四) 科技信息离散定律

科学论文在科技期刊中的分布是不均匀的，少数期刊中“拥挤”着大量的论文，大量的期刊中“稀释”着少量的论文。1934年英国信息学家 S.C.布拉德福(Bradford)明确地指出这种倾向。他写道：“对某一主题而言，将科学期刊按刊登相关论文减少的顺序排列时，可以划分对该主题最有贡献

的核心区，以及含有与该区域论文数量相同的几个区域。每个区域里的期刊数量成  $1:n:n^2 \dots$ ”，这就是为后人所称道的布拉德福定律。

半个多世纪以来，布拉德福定律沿着所谓区域法和图像法发展。在区域法中有代表性的是 F.F. 莱姆库勒 (Leimkuhler) 所建立的布拉德福分布

$$F(x) = \frac{\ln(1 + \beta x)}{\ln(1 + \beta)}$$

是分布函数的参数， $x$  表示在一个特定学科中载文量最高期刊的数量与相关于该学科的总刊量之比。 $0 < x < 1$ ； $F(x)$  为这一部分期刊的载文量与该期刊组合总产量之比（即该部分期刊中全部相关论文与总期刊集合所拥有的论文之比）。

$$P(x) = F(x) = \frac{\beta}{(1 + \beta x) \ln(1 + \beta)}$$

此式为布拉德福分布的概率密度函数。

图像法的代表中，质量最好的是 1977 年前苏联信息学家斯马里柯夫建立的统一方程式：

$$R(n) = K \ln(q_1^n + q_2 e^{-n})$$

式中， $R(n)$  为累积载文量， $K$ 、 $q_1$ 、 $q_2$  均为常数。

信息离散定律的模型和经验公式很多，有的能实用，有的仅有描述与解释作用，能应用的质量与置信水平也不完全一样。

#### （五）词频分布规律

科技信息是由符号有规律地排列而形成的。符号主要是文字，现代书面语言——文字主要是表意语言。由于表意语言的运用，人们就可以用自然语言把思想内容固定在可以携带的载体上，而不再单纯地依赖口叙和记忆力。语言本身是有其规律的，就其被人类运用而言也是这样。齐普夫在 1935 年经过大量数据的验证得出：如果在一个足够长的文集中， $f$  表示一个词汇出现的频率， $r$  是按  $f$  大小而排列的顺序。那么，它们之积是一个常数：

$$f \cdot r = C \quad (17)$$

严格地说，式 (17) 仅仅对中频、中序的词才是有效的。50 年代初期，法国数学家 B. 芒代尔布罗 (Mandelbort) 开始用信息论的方法研究词序分布规律，经过严格地数字推导，从理论上提出了如下公式：

$$(r + m) \cdot f = C \quad (18)$$

式中， $r$  是词的等级； $f$  是频次； $m$ 、 $g$  和  $C$  是取决于文

集的亥数。式 (18) 对于表现低序高频词更为适宜。但是式 (18) 同样也没有解决高序低频词的问题。1967 年 A.D. 布思 (Booth) 提出了新的数学解析式，指出出现不同频次的词数，与文集所用的词长度和常数  $C$  无关，仅与频次不同的词汇  $\ln$  的频次  $n$  有关：

$$\ln/I = \frac{2}{n(n+1)}$$

显然，式 (19) 对于处理低频高序词是适宜的。但是  $\ln$  的  $n$  临界值为多大呢？式 (19) 的适用范围应怎样规定呢？1973 年 J.C. 多诺霍 (Donohue) 提出一个判断式：

$$\ln = \frac{-1 \pm \sqrt{1 + 8I_1}}{2}$$

式中， $l_1$  为仅出现一次的词数，小于词频  $l_n$  的词均属于低频词。

### 三、信息计量学方法的类型

信息计量学方法是由多种方法构成的综合体。按研究手段可分为信息统计分析法、数学模型分析法、系统分析法、矩阵分析法、网络分析法等，这里介绍其中最基本的也是最重要的一种方法：信息统计法。

信息统计是指以某一特定单位对信息或其相关媒介进行统一的计量。信息统计分析方法是利用统计学方法对信息进行统计分析，以数据来描述和揭示信息的数量特征和变化规律，从而达到一定研究目的的一种分析研究方法。

1. 信息统计的媒介：各种类型的书目、索引、文摘；期刊杂志；期刊指南和联合目录；评述性杂志及其它出版物；与信息利用情况有关的读者借阅单、阅览登记资料等等。

2. 信息统计分析的主要内容：对各种各样的信息进行统计与分析；著者数的统计与分析；科学术语的统计与分析；引证关系的统计与分析；读者的统计与分析；信息利用情况的统计与分析等等。

3. 信息统计分析的步骤：统计调查：搜集研究对象的原始数据，数据必须准确可靠；统计整理：对统计数据进行分门别类的加工整理，包括对统计数据进行必要的计算，根据一定的规则排序，并用统计表或统计图的形式表示出来；统计分析：包括对统计数据的结论分析和误差分析，前者是以统计数据为依据分析判断其相应的规律性的结论，后者是对结果的准确性和精确度进行分析。

信息统计是信息研究的基础工作，信息统计数据是开展信息定量研究的必要条件，运用信息统计分析方法可以反映信息的增长变化、分布特征、流通状况、利用程度，揭示信息的数量变化规律，为信息单位的信息管理提供依据。

## 第三节 信息计量学方法的应用

### 一、信息计量学方法的应用

#### （一）测定核心期刊

期刊是最基本、最常用的一种信息存贮类型。在期刊论文的实际分布中，存在着一种普遍现象，即对于某一特定学科或专业来说，少数几种期刊所含的相关信息量很大，多数期刊的信息量却很小。我们将含有某专业、学科相关信息量很大的少数几种期刊，称为该专业、学科的核心期刊。确定核心期刊，不仅可以满足信息单位科学管理的需要，缓解信息经费、馆藏等方面的问题，而且可以提高信息服务的效率和准确性、针对性。因此，测定核心期刊是信息工作的一项重要任务。信息计量学方法为测定某一专业、学科的核心期刊，提供了一套科学的方法，如布拉德福定律法及百分比补偿、累积百分比法、文摘法、流通率法与综合评价法。加拿大学者 A. 波普曾用布拉德福图像分析法测定了信息学的核心期刊。近年来，信息计量学方法已被广泛应用于化学、医学、农业、生物、海洋等多种学科的核心期刊的确定。

#### （二）指导信息收藏管理

应用信息计量学方法指导信息收藏与管理，表现在以下几个方面：制定信息采集策略，比如优先订购核心期刊，确定某一主题、专业的信息采集

的适当比例等；选择各类信息的最佳搜集方式，如根据用户要求的特点和经济、合理等因素，确定对某一类信息是采取订购、交换、复制或其他什么形式；确定期刊收藏的最佳方案，即根据布氏定律与信息老化指数规律，从有关某一学科 100% 的期刊中选出 P 泥作为一个适当的比例加以收藏；指导信息文献剔旧工作，即依据文献的载文量、引文量、流通数据、用户反馈数据、馆藏大小等因素，确定信息文献剔旧的合理方案，维护动态的最佳馆藏；评价信息收藏工作，即根据信息用户占有率、信息利用率数据评价一个信息机构的信息收藏是否充分、合理、经济，是否能基本满足用户对信息的需求。

### （三）测定检索工具的完整性

检索工具的质量直接关系到检索效果，利用布氏定律，可以测定文摘、索引等书目类检索工具的完整性。其具体作法是：统计要校验的文摘或索引的实际款目数量和摘引的期刊数量；根据从实际统计中得到的某学科期刊  $n$  和  $R(n)$  的一组数据，利用  $R(n) = K \log n$ ，求出期刊总数 ( $N = K$ )；再根据公式  $R(N) = K \log N$  求出该学科的论文总数；通过实际值与理论计算值的比较，便能测定其完整性。例如，按布氏定律公式计算，1970 年的《热带文摘》该年只摘及了 374 种期刊和 2284 篇论文。因此，该文摘漏摘期刊 269 种，漏摘期刊率高达 41.8%；漏摘论文 500 篇，其漏摘率为 18%。也就是说，该文摘具有 58.2% 的期刊完整性和 82% 的论文完整性。

### （四）指导用户利用期刊

利用布氏定律原理，可以指导用户选择阅读重点文献，从而节省用户时间，提高获取信息的效率。具体地说，就是用布氏定律确定某学科的“核心期刊”，用这种量的概念指导用户阅读。

### （五）研究信息利用规律

信息利用规律的研究有利于开发信息资源，提高信息服务的针对性。利用信息计量学方法，可以从定量角度研究信息利用规律。中国科技信息学会 1981 年进行的“科技人员利用信息综合调查”就是一个很好的例子。

### （六）研究科学发展的特点

现代科学发展的许多突出特点，都可以从科学信息的数量及其变化上看出。从信息计量的角度，可以看出并证明科学发展的速度在加快，科学发展具有继承性，具有阶段性、学科交叉渗透性，科学发展的重点不断转移以及科学劳动的集体性等现代科学发展的特点。

### （七）预测学科发展

一般说来，一门学科的成长一般要经历萌芽、发展、成熟、分化的过程。在这个过程中，表述该学科研究成果的信息在数量和内容构成上也要相应地发生变化。当一门学科萌芽时，只有少数几篇文献，其内容也大多是一些实验性事实和学科概念的讨论；当学科发展时，信息数量显著增长，内容日渐成熟，理论性信息明显增多；当学科发展到成熟阶段时，信息增长速度变慢并逐渐达到饱和状态，应用信息的比例增大，这标志着该学科已经成熟，很少有新的发展；如果分化出新的知识领域，该学科的信息量又会速增。这说明一门学科本身的成长过程与其信息在数量和内容构成上的变化有着密切的联系。而这种联系正是我们利用信息计量预测学科发展动向及前景的重要依据。

## 二、信息计量学方法应用综合举例

例：南通医学院董建成、周沂在大量的信息统计的基础上，运用等级排列和图表分析技术，研究了世界和中国医学信息学发展的状况。

### (一) 材料和方法

统计数据来自美国 Silver Platter 公司生产的 MEDLINE 光盘，该光盘共存取 72 个国家和地区 43 种文字的主要生物医学文献。

统计方法，利用 MEDLINE 光盘检索系统中医学主题词表 (MeSH) 的标准主题词，检索医学信息学的全部文献，按年代分为 10 组。在各组文献中，应用逻辑运算符进行主题、副主题检索，获取医学信息学各研究领域的不同文献；再采用位置运算符获取医学信息学文献的不同类型、语种和出版国家。所有检索结果均输入 FOX BASE 数据库，分别从各国文献的不同方面进行计量和主题分析。

### (二) 结果和讨论

1. 结果：1987~1996 年被 MEDLINE 系统收录的生物医学文献共 3529936 篇，其中有关医学信息学的文献 1106 篇，逐年的变化情况见图 4—4。这在一定程度上反映了全世界医学信息学的文献量。全部的医学信息学文献分别来自 41 个国家和地区，23 种文字的不同研究领域 (见表 4—8、表 4—9)。

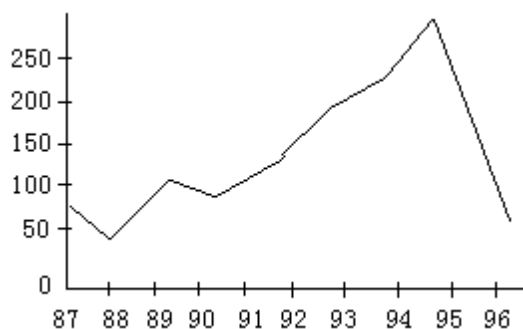


图 4-4 1987-1996 年医学信息学文献数量的变化

根据不同刊物发表医学信息学文献量的不同，作者在系统对 19 种重要的与医学信息学相关的期刊进行了排序 (见表 4-10)。表 4—8 医学信息学文

献语种分布

年份	文献量	英文	法文	俄文	西班牙	德文	捷克文	意大利	其他
1996	61	57	2	2	0	0	0	0	0
1995	250	235	3	0	1	3	1	2	5
1994	183	167	1	1	3	2	1	1	7
1993	141	123	6	3	1	2	0	2	4
1992	84	72	1	3	2	1	0	1	4
1991	82	57	4	5	5	1	0	3	7
1990	74	49	5	1	3	9	1	2	4
1989	100	74	11	3	3	2	4	1	2
1988	55	29	4	5	5	0	6	1	5
1987	76	41	10	2	1	1	6	0	15
合计	1106	904	47	25	24	21	19	13	53
%	100.00	81.73	4.25	2.26	2.17	1.90	1.72	1.18	4.79

表 4—9 医学信息学各研究领域文献量

表 4—10 医学信息学核心期刊表

序号	国别	文种	刊名	国际标准刊号 ISSN	摘引量
1	德国	英 English	MethodInfMed	0026-1270	96
2	美国	英	ProcAnnuSympCom- putApplMedCare	0195-4210	58
3	德国	英	MD Comput	0724-6811	51
4	英国	英	BMJ	0959-8138	42
5	美国	英	Jam Med Inform Assoc	1067-5027	35
6	美国	英	JAMA	0098-7484	33
7	加拿大	英	Medinfo		33
8	英国	英	IntJBiomedComput	0020-7107	23
9	美国	英	JMEDSyst	0148-5598	23
10	英国	英	ComputMethodsPro- gramsBiomed	0169-2607	20
11	英国	英	MedInfLond	0307-7640	18
12	美国	英	BullMedLibrAssoc	0025-7338	17
13	美国	英	AcadMed	1040-2446-16	
14	美国	英	ComputNurs	0736-8593	16
15	美国	英	JDentEduc	0022-0337	15
16	美国	英	NursManage	0744-6314	15
17	捷克	捷克	CeskPediatr	0069-2328	14
18	美国	英	ComputBiomedRes	0010-4809	13
19	英国	英	Lancet	0140-6736	12

2. 讨论：要对 10 年来国际医学信息学发展趋势作出总的结论是困难的，因为大量数据资料存在着多种变化因素。然而，从科学论文中常可得知某一学科学术的动态，从医学信息的过去、现在和将来中受到启迪。

(1) 从医学信息学文献的文种上看，英文文献占 80% 以上，法文文献次之，其他文种的文献占文献总量的 14.02%。从医学信息学研究的内容看，涉及其应用、计算机和教育方面的文献已超过一半，学科发展趋势研究的文献占文献总量的 10.22%，前者已成为医学信息学研究的主要内容。其特点有四： 计算机科学和技术在医学信息学研究中的应用最为普遍，尤其是计算机化的医学记录、远程医疗和电子化的生物医学出版业取得了实质性进展； 医学信息学研究冲破了传统的单纯性信息交流和信息存取的模式，形成了所谓的横向结合和解决日常工作中各种各样难题的研究体系；医学信息学教育及利用信息技术进行的医学教育得到了长足的发展； Internet( 国际互联网 ) 的发展为医学信息学的研究提供了更为广阔的天地，从而使得人们能够容易地利用医学信息作出决策； 在技术和行政上解决电子医学信息的保密性与有效入口途径之间的矛盾正有待解决。再从出版医学信息学文献的国别看，美国出版的医学信息学文献有 520 篇，占第一位，英国和德国分别为 129 篇和 118 篇，加拿大 46 篇，法国 37 篇。中国仅有 3 篇论文被收入



MEDLINE 检索系统，名列第 23 位。这种现象充分表明了英美等发达国家在医学信息学研究上的优势，我国医学信息学的研究不论在深度或广度上与发达国家相比尚存在差距。

(2) 1987 ~ 1996 年，MEDLINE 收录医学信息学研究的文献量仅占医学总文献量的 0.031%，这与其他学科相比，不仅医学信息学研究的成果少，而且反映了国内外医学信息学研究的人员和机构亦远远不足。从附图可见，虽然近年来的医学信息学研究论文数量有所增加，但涉及医学信息学领域的研究和开发，通常需要从从事过医学信息学正规训练的内科医生或其他卫生专业人员。这应当引起我们的重视，而对其研究的投入要注意参考国际医学信息学研究的走向。

(3) MEDLINE 1987 ~ 1996 年共收录医学信息学文献 1106 篇，涉及期刊 236 种，其中 550 篇文献分布在 19 种重要期刊中，占被检索医学信息学文献总量的 49.73%。19 种重要期刊中，英文期刊 18 种，被检索文献 536 篇，占 19 种重要期刊被检索文献总量的 97.45%。其中美国出版的期刊 10 种、英国 5 种、德国 2 种、加拿大 1 种、捷克斯洛伐克 1 种。以上表明，英文是医学信息学文献的最主要文种，美国、英国和德国是当前世界上发表有关医学信息学研究文献最多的国家。

(4) 与国外相比，我国医学信息学研究存在的主要问题是：我国医学信息学研究的整体水平与先进国家相比还有很大差距，意识淡薄，经费不足，研究手段落后；指令性任务少，协作攻关课题少，难以形成较高水平的选题；在医学信息设备研制和改进方面，缺少必要的技术储备；对外交流少，人才成长缓慢。

当前，随着我国计算机科学和医学科学的发展，大力开展我国医学信息学研究的条件已经基本具备。我们要充分利用当前的有利时机，以国外先进的技术和研究动向为借鉴，正确处理近期发展和长远发展、基础研究与应用研究的关系。根据我国医学科学发展的特点及信息基础设施的实际情况，积极开展适合我国国情的数据基础、信息基础和知识基础的不同层次的研究，树立系统观念，加强横向联系和内外交流，协作攻关，力争在不长的时间内，解决几个重大课题；同时加强与医学信息学发展有关的基础学科的研究，增强发展后劲，推动应用研究工作的发展。在今后的医学信息学研究中，要不断获取最新信息，掌握先进国家医学信息学发展的动态，注意在医学信息学研究中应用新理论、新技术和新方法，同时努力改变设备手段落后的局面，提高我国医学信息学研究的总体水平。

## 第五章引文分析方法

引文分析方法被人们视为信息计量学的重要手段之一。从世界范围来看,自 1927 年第一篇使用引文分析方法的论文问世以来,引文分析方法已被作为研究文献利用规律、科学评价、科学预测和研究科学与社会关系的重要手段。从我国的情况看,引文分析方法自 80 年代引入以来,已给我国信息学和科学学研究注入了一股强大的活力。人们希望借助它实现信息学研究和科学研究的定量化。

## 第一节 引文分析方法概述

引文分析方法以其独特的优势成为一种重要的信息计量学方法，并被广泛使用，下面我们就引文分析方法的主要内容加以概述。

### 一、引文分析方法的理论基础及性质

#### (一) 引文分析方法的文献学基础

引文分析方法，就是利用各种数学及统计学的方法进行比较、归纳、抽象、概括等的逻辑方法，对科学期刊、论文、著者等分析对象的引用和被引用现象进行分析，以揭示其数量特征和内在规律的一种信息计量研究方法。

出于某种原因，科技文献的作者总是在其文献末尾列出在撰写该文时参考或引用过的文献“清单”。人们将这种参考或引用其他文献的过程叫做“引证”。被引文献与引用文献或引用者必定存在着某种关系。M.Weinstak 将引用机理归纳为 15 种类型：对开拓者表示尊重；对有关著作给予荣誉；核对其所用的方法及仪器；提供背景阅读材料；对自己的著作予以更正；对别人的著作予以更正；评价以前的著作；为自己的主张寻求充分的论证；提供研究者现有的著作；对未被传播、很少被标引或未被引证的文献提供向导；(11) 鉴定数据及物理常数等；(12) 核对原始资料中某个观点或概念是否被讨论过；(13) 核对原始资料或其他著作中的起因人物的某个概念或名词；(14) 承认他人的著作或概念；(15) 对他人的优先权要求提出争议。

从所划分出的引用机理类型可以看出，文献的相互引用是由科学发展的规律和科学研究活动的规律所决定的。科学学的研究反复表明，科学知识具有明显的累积性、继承性；任何新的学科或新的技术，都是在原有学科或技术的基础上分化、衍生出来的，都是对原有学科或技术的发展。科技文献是科技成果的反映和记录。新文献反映和记录的新的科研成果是建立在过去所发表文献记录和反映旧有的科研成果的基础之上的，是对旧有文献和成果的继承和发展。人类知识之所以能代代相传，并得以累积，形成今天这样庞大的知识体系，所依仗的就是文献；即旧文献对于新文献是要发生知识转移的。对科学研究过程中的文献利用规律和科技工作者信息交流渠道的研究表明，科技工作者得以创造新的科研成果的知识基础主要来源于其他文献。将那些对自己的文献发生了知识转移的文献列为参考文献，是现代各国关于著作权的有关法律、法规所明确规定的，也是科学活动中必须遵循的行为准则。文献之间的这种引用关系是文献与文献之间知识转移的遗迹，这就是引文分析方法得以建立的文献学基础。

那么，事实上科技文献的作者是否将那些对自己的文献发生了知识转移的文献都列为了参考文献，以及所列出的参考文献是否都是对自己的文献发生了知识转移的文献呢？人们的考察结果表明，实际的引文关系远比上面的分析复杂。但这并不危及引文分析方法赖以建立的文献学基础。

#### (二) 引文分析方法的数学原理

引文分析方法的数学基础是概率论与数理统计。它使用了概率分布、抽样统计、样本与总体的关系等基本的统计数学原理。它有时直接对总体进行分析，也常常采用抽样的方法，用样本的特征代表总体的特征，通过样本内或样本之间的特征比较，得出总体的结论（引文评价）；或通过样本内部文献引用关系的分析，推断总体内部的文献引用关系，进而推断知识转移关系

和学科、主题之间的关系（引文网状分析）。

既然分析的直接对象是样本，而分析的最终结论是关于总体的，那么，样本对于总体就应具有充分的代表性。样本是否对总体具有充分的代表性，是由所考察指标的分布形式和抽样的方法决定的。为使样本能充分地代表总体，就要求所抽取的样本充分地大，以使样本能将我们所要分析的总体的特征忠实地反映出来。另一方面，如果所抽取的样本对于总体具有充分的代表性，也不必一味强调样本的规模，因为它毕竟是样本。我们分析样本，而不直接分析总体，是为了减少统计分析的工作量。只要能充分地代表总体，样本应尽可能地小。

引文分析方法的数学原理决定了引文测度具有宏观性和相对性，这是在进行引文评价和利用引文测度结果时必须给予充分注意的。

引文测度是一种统计意义的宏观测度，是对被分析对象整体的考察。既然是宏观测度，就不应过分强调微观的差异性。我们在对风速这一自然现象进行测度时，考察的是气流整体的运动，而不考虑有的气体分子在作反方向的热运动，以及局部气流遇到障碍物而作反方向的运动，就是这个道理。又如评选先进集体，考察的是集体这一整体是否先进，而不强调集体中人人都达到先进个人的标准，也是这个道理。有些批评引文分析方法的文章认为，被引频次高的期刊，其被引频次主要集中在少数文章上，许多文章几乎不被引用，因而引文评价不科学。显然，这种认识是不能成立的。核心期刊之所以成为核心期刊，就是因为它刊载了较多的被引频次高的文献。

在我们实际应用中存在着“用而不引”，即并未将有知识转移的文献尽数列出的现象。这是不足为怪的。因为文献与文献之间的知识转移不是简单的机械转移。被转移的知识要在人脑中暂驻相当时间，并经人脑加工、升华。科技人员进行一项科研并将其科研成果撰写成论文，需经历比较长时间的知识准备和积累。一个人不可能同时阅读两篇以上的文献，他所阅读的文献总是有先有后的，在阅读后一篇文献时先前阅读过的文献中的有关知识程度不同地暂住在他的大脑中。当他在著录参考文献时难免发生“用而不引”的情况。不过这并不影响引文分析方法得以建立的文献学基础，因为引文分析方法建立在统计数学的统计抽样的原理基础之上。

有时候有人提出所列出的参考文献并不是都对该篇文献有知识转移。作者在著录引用文献时或存在随意性，或存在引用名家的不相干的文献以提高自身价值的情况。即便存在这些情况，对引文评价也是没有多大影响的，因为这些情况对于所有分析对象是均匀分布的。如果这种情况占的比重较大，也可能受到影响的是引文网状分析的结论。不过从过去的引文分析实践所取得的结果来看，这种情况似乎并未对被引频次起到决定性作用。通常代表我国科研水平的全国性学报被引频次最高，而一般在这些刊物上发表文章的是各个学科领域的较有影响的名家，这似乎为以上认识提供了依据，但是，这些引用并非都是慕名，权威性期刊和权威人物的学术带头作用也是不可否认的。由学术水平较高的刊物向学术水平较低的刊物发生知识转移，是科学研究和科技文献发展的一般规律。此外，引文分析的结果还表明，科技期刊（即使是层次较低的刊物）的自引频次往往大于对其他某种刊物（即使是全国性学报）的他引频次，这说明每一种刊物及其为这种刊物撰文的作者都有其特定的研究领域和研究路向，即使在对引用文献进行著录时，也体现了这一点。

对评价性引文分析（如引文评刊、引文评选核心作者）来说，不是简单

地通过测度考察被引频次的绝对值而得出结论，而是将所有考察对象的被引频次的大小进行比较、排序，而后得出结论，即评价性引文分析是一种相对测度，而不是一种绝对测度。因而，对所有分析对象（特定的期刊或特定著者的文献的被引频次）都有影响的因素（如：著者引用文献具有一定程度的随意性；一篇文献的著者引用多少文献具有随机性；被引文献对于引用文献的知识转移量不同等因素）是不会对分析结论产生影响的。因为这些因素对所有分析对象的被引频次的绝对数量和被引用的质量都有影响，即，它们的作用是均匀分布的，在进行分析比较时，已将其作用排除在测度结果之外了。

例如，引文指标往往受多因素影响，指标值的大小往往由多因素决定，是一个多元函数系统。而一些具体的引文分析方法则将这一指标用于测度某一因素，假定其他因素不变，即将多元统计分析问题简化为一元统计分析问题来处理。当然，尽管统计分析的原理是无可置疑的，但各种假定是否成立却直接影响到分析结果的有效性和可靠性。

## 二、引文分析方法的基本内容及类型

### （一）引文分析方法的类型

从不同的角度和标准来划分，引文分析方法有着不同的类型。如果从获取引文数据的方式来看，有直接法和间接法之分。前者是直接来源于期刊中统计原始论文所附的被引文献，从而取得数据并进行引文分析的方法；后者则是通过“科学引文索引”（SCI）、“期刊引用报告”（JCR）等引文分析工具，查得引文数据再进行分析的一种方法。若从文献引证的相关程度来看，则有自引分析、双引分析、三引分析等类型。如果从分析的出发点和内容来看，引文分析大致有三种基本类型：

1. 引文数量分析：主要用于评价期刊和论文；研究文献情报流的规律等。

2. 引文网状分析：主要用于揭示科学结构、学科相关程度和进行文献检索等。

3. 引文链状分析：科技论文间存在着一种“引文链”，如文献A被文献B引，B被文献C引，C又被文献D引，等等。对这种引文的链状结构进行研究可以揭示科学的发展过程并展望未来的前景。

### （二）引文分析方法的基本内容

由于科学技术体系结构的影响，各个年代各种类型的数量繁多的科学引文形成了相应的分布结构；同时，科学引文都具有许多基本要素。如果我们从不同的角度或从各种基本要素出发，对科学引文的分布结构进行描述和分析，便形成引文分析的基本内容，一般包括：

1. 引文年代分析：一般来说，随着年度的由远而近，引文量呈增长趋势，即时间愈近，被引用的文献愈多；文献被利用的峰值是该文章发生以后的第二年。如果以引文年代为横轴，各年引文量为纵轴，在坐标图上描绘各年数据点，然后用一条线连接起来，便可得到一条引文年代分布曲线。通过对该曲线的分析，不仅可以了解被引文献的出版、传播和利用情况，而且可以研究科学发展的进程和规律，特别是在文献老化和科技史的研究中，引文年代分析更是一种广泛应用的有效方法。

2. 引文量分析：引文量是某一主体对象含有的参考文献数量。它是引文链的基本特征之一。通过引文数量的分析，不仅可以揭示文献引证与被引证双方的相互联系，而且还可以从定量的角度反映出主体之间的联系强度。如

果两篇论文或两种期刊之间的引文数量大，就可以认为它们之间的引证强度大，说明其联系较紧密。引文量的分布规律可从下列几个方面分析：

(1) 引文量的理论分布：我们将一定量的论文的引文量数据进行分析比较，发现其变化规律表现为以平均数为中点，接近中点的频数最多，离平均数远的频数趋于减少，形成中间高两极低的正态理论分布。

(2) 引文篇数分布：即每篇研究论文平均占有的引文篇数的分布。它不仅反映了论文作者引用文献的广度和深度，而且还能说明引文与被引文的学科内容之间的联系强度。

3. 引文的集中与离散规律分析：引文分布的集中性与离散性是相对于一定的测度指标而言的。引文按来源期刊的分布，引文篇数的频数以平均数为中心的分布，引文按年度、语种、文献类型等的分布，都表现出这种集中与离散的趋势。

4. 引文类型分析：科学研究中引用的文献很广，有期刊论文、图书和特种文献。对被引文献的类型进行分析，将有利于确定文献情报搜集的重点。

5. 引文语种分析：引用文献是由不同语种的文献构成的。某一语种的文献被引用量愈大，则说明该语种比较常用和重要。考察和分析引文语种的分布，对于人们有计划地引进外文文献、译文选题、外语教育等，颇有参考价值。

6. 引文国别分析：对引文的国别分析，特别是各国文献互引情况的统计分析，可以探明各国互引文献的状况，弄清国际文献交流的数量和流向。

### (三) 引文分析方法的特点及局限

1. 广泛适用性：引文分析的素材是引文与被引文，而引文现象又是普遍存在的。以期刊论文为例，全世界范围约有 90% 以上的科学论文附设了引用文献，平均每篇论文有引用文献 15 篇。我国目前 88% 左右的重要科学论文带有引用文献，平均每篇中文科学论文有引用文献 8.9 篇，可以说，凡是有引用文献的地方，引文分析方法就有用武之地，所以，引文分析方法具有广泛适用性。

2. 简便易用性：由于引文分析不要求其它先决条件和辅助条件，不需要使用者具有十分专深的知识，研究的深度、广度可以由自己控制，所以一般的信息人员都可以借助于这种方法，完成一些有价值的研究课题，解决一些工作中的实际问题。总之，这种方法的使用限制极少，简便易用，很值得在广大的信息人员中普及推广。

3. 功能特异性：由于引文分析方法具有广泛适用性和简便易用性的特点，通过一些不太复杂的统计和分析，就可以确定核心期刊、研究文献老化规律、研究信息用户的需求特点，甚至可以研究学科结构、评价人才等，我们不能不为其功能而感叹。

虽然引文分析方法具有广阔的应用前景，但是也存在一定的局限性。著者引用文献是一个人为控制的思维和判断过程，而作为其表现形式的引用文献，仅仅是宏观的、表面的测度，受到许多限制因素的影响。

(1) 引文关系上假联系的影响：引用文献的原因多种多样，两篇论文可能出于完全不同的原因或从不同的角度引用同一篇早期文献，一篇可能是引用其方法，另一篇可能是引用其结果，那么这两篇文献在内容上的联系就有可能是虚假的。引文有些是发生在前言和篇名中；有些是发生在正文中；有些发生在结论或讨论中。在这些情况下，作者对原著的引用内容和程度是

不相同的。引文对原著的关系和重要性也各不相同，但在目前的引文分析中，对它们都是同等看待，不加区分的。这样也容易造成成假关系。新刊的论文得不到大量引用；小型期刊被引率往往低于大型期刊；引而未用或用而未引的情况也时有发生。文献引用中的这些现象都会影响引文分析方法的应用和效果。

(2) 文献被引用并不完全等于重要：例如，有些具有错误观点或结论的论文，后人出于批评商榷，被引次数可能很多。另一方面，被引次数较少的文献也不能一概认为不重要。它受到许多因素的限制，如发表的时间、语种、学科专业等等。被引次数上的微小差别也不能完全说明质量上的优劣，它有很大的随机性，只有当这一差别很大时，才能说明问题。

(3) 著者选用引文受到可获得性的影响：索普(M.E.Soper)研究指出，著者引用的文献，大部分是个人收藏的文献；少部分是本部门 and 就近图书馆的资料，而其它城市或其它国家的文献所占比例甚小。这说明著者选用参考文献以方便为准则，以占有为前提；同时还要受到著者语言能力、文献本身年龄和流通周期，以及二次出版物报道的影响。

(4) 马太效应的影响：有的研究者认为，在文献引用方面也存在着马太效应的影响。人们往往以“名著”、“权威”作为选择引文的标准，有的确是出于需要，也有的则是为了装饰门面，抬高自己论文的身价。一种期刊因为发表名人的文章而为众人所引用，以至引起连锁反应，结果其引文率很高。这种马太效应的心理作用，掩盖和影响文献引用的真实性。

### 三、引文分析方法的使用范围

#### (一) 引文分析方法的测度指标

一般来说，对科学期刊进行分析时常用的测度指标有五种：即自引率、被自引率、影响因子、引证率与当年指标。在对专业和学科结构进行研究时，除用引证率外，还可用引文耦合和同被引等测度指标。

1. 自引率：在引用文献的过程中，限于主体本身范围内的引用称之为“自引”。包括同一类学科文献的自引、同一期刊文献的自引、同一著者文献的自引、同一机构文献的自引、同一种文献的自引、同一时期文献的自引、同一地区文献的自引。自引率就是对主体本身范围内文献引用的次数与主体引用的文献总数的比值。

$$\text{自引率} = \frac{\text{主体的自引次数}}{\text{主体引用的文献总数}}$$

2. 被自引率：这是被引用现象的一个测度，被自引率就是主体文献被自引的次数与主体被引用的总次数的比值。它反映出被引用中有多少是被自己引用的。

$$\text{被自引率} = \frac{\text{主体文献被自引的次数}}{\text{主体被引用的总数}}$$

3. 影响因子：主要在研究科技期刊时使用，等于期刊在规定时间内（一般是两年）论文被引量与可引论文总数之比。

$$\text{某年度某刊的影响因子} = \frac{\text{某年引用某刊前两年论文的总数}}{\text{前两年该刊所发表的论文总数}}$$

4. 引证率：期刊引证率等于该刊中参考文献量除以期刊载文量。这是衡量吸收文献能力的一个相对指标。

$$\text{引证率} = \frac{\text{参考文献量}}{\text{载文量}}$$

5. 即时指标：这是测度期刊被引用速度的指标，它是期刊某年发表的论文当年被引用的次数，除以该刊这一年所发表文章的总数，是衡量期刊重要性的一种依据。

$$\text{即时指标} = \frac{\text{某刊某年发表的论文当年被引用的次数}}{\text{该刊当年发表的文章总数}}$$

6. 引文耦合：当两篇文章同时引用一篇或多篇相同的文献时，这种现象称引文耦合，这两篇文献就具有耦合关系。引文耦合的文献之间总存在着这样或那样的联系，其联系的程度称为耦合强度。

7. 同被引：当两篇（多篇）论文同时被别的论文引用时，则称这两篇论文具有“同被引”关系，引用它们的论文的多少，即同被引程度，称为同被引强度。

## （二）引文分析方法的应用领域

目前，引文分析技术日趋完善，应用不断扩大，已发展成为文献计量学的重要方法之一。引文分析方法的应用主要有以下几个方面：

1. 测定学科的影响和重要性：通过文献引用频率的分析研究可以测定某一学科的影响和某一国家某些学科的重要性。

2. 研究学科结构：通过引文聚类分析，特别是从引文间的网状关系进行研究，能够探明有关学科之间的亲缘关系和结构，划定某学科的作者集体；分析推测学科间的交叉、渗透和衍生趋势；还能对某一学科的产生背景、发展概貌、突破性成就、相互渗透和今后发展方向进行分析，从而揭示科学的动态结构和某些发展规律。1974年，加菲尔德运用计算机系统，通过引文分析，描绘出1972年和1973年生物医学领域各主要课题、各课题相互间的关系以及新课题产生的示意图，从而反映出生物医学研究的内部结构。

3. 研究学科信息源分布：通过文献间的相互引证关系，分析某学科（或专业）文献的参考文献的来源和学科特性，不仅可以了解该学科与哪些学科有联系，而且还能探明其信息的来源及分布特征，从而为制定本学科的信息管理方案和发展规划提供依据。

4. 确定核心期刊：引文分析方法是确定核心期刊的常用方法之一。这种方法的主要特点，是从文献被利用的角度来评价和选择期刊的，比较客观。加菲尔德通过引文分析，研究了文献的聚类规律。他将期刊按照期刊引用率的次序排列，发现每门学科的文獻都包含有其它学科的核心文献。这样，所有学科的文獻加在一起就可构成一个整体的、多学科的核心文献，而刊载这些核心文献的期刊不过1000种左右。利用期刊引文的这种集中性规律可以确定学科的核心期刊。

5. 研究文献老化规律：目前，有关文献老化的研究一般是从文献被利用角度出发的。D. 普赖斯曾利用引文分析探讨文献的老化规律。通过对“当年指标”和“期刊平均引用率”的分析，他认为期刊论文是由半衰期绝然不同的两大类文献构成的，即档案性文献和有现时作用的文献。科学文献之间引文关系的一种基本形式是引文的时间序列。对引文的年代分布曲线进行分析，可以测定各学科期刊的“半衰期”和“最大引文年限”，从而为制定文献的最佳收藏年限、对文献利用进行定量分析提供依据。同时，一个学科的引文年代分布曲线与其老化曲线极为相似。这有力地说明文献引文分布反映



了文献老化的规律性。因此，从文献引用的角度研究文献老化规律是一种有效的途径和方法。

6. 研究信息用户的需求特点：利用引文分析方法进行信息用户研究是一种重要途径。根据科学文献的引文可以研究用户的信息需求特点。一般来说，附在论文末尾的被引用文献是用户（作者）所需要和利用的最有代表性的文献。因此，引文的特点可基本反映出用户利用正式渠道获得信息的主要特点，尤其是某信息中心对其所服务的用户所发表的论文的引文分析，更具有直接的指导意义。通过对同一专业的用户所发表的论文的大量引文统计，可以获得与信息需求有关的许多指标，如引文数量、引文的文献类型、引文的语种分布、引文的时间分布、引文出处等。

7. 评价人才：在人才评价方面，常采用引文分析方法。这是因为某著者的论文被别人引用的程度可以是衡量该论文学术价值和影响的一种测度，同时，也从科研成果被利用的角度反映了该著者在本学科领域内的影响和地位。因此，引文数据为人才评价提供了定量依据。从对历年诺贝尔奖金获得者的论文被引用情况的统计来看，物理学、化学、医学领域中诺贝尔奖金获得者中，其论文被引次数最高者（L.D.Landan）为 1888 次；最低的也有 79 次（J·H·D·Jensen）；只有六名低于 200 次。可见，这些科学界的精英的论文的引用次数是很高的。

## 第二节 引文分析方法的主要工具

引文分析方法作为信息计量学研究的重要手段之一，越来越受到大家的重视。通过引文分析方法，我们可以评价一个单位、一种期刊、一个人的学术水平，研究一个学科的结构和文献发展规律。而这方面应用最多的是美国的《科学引文索引》，其它还有《社会科学引文索引》《期刊引证报告》等等。下面我们只介绍一下《科学引文索引》。

### 一、概述

《科学引文索引》由美国费城科技情报所出版，1961 年创刊，现为月刊，收录范围很广。目前，《科学引文索引》收录了世界上 40 多个科技发达国家用 37 种文字出版的 3300 多种期刊、1800 多种会议录和 1250 多种重要专著、丛书、论文集、专利文献、评论、通讯等，编排了他们之间文献互相引证关系，提供查找途径。所收录文献的内容涉及自然科学的各个领域，包括生命科学、行为科学、生物、临床医学、兽医学、化学、物理、农业、工程技术、药学等应用科学。学科覆盖面广，可从各种不同角度获得相关与交叉学科期刊上的文献。

《科学引文索引》由 5 种索引构成，它们是作者引文索引、专利引文索引、轮排主题索引、团体引文索引和来源索引，只要有一点线索（被引证文献著者、单位、地名、课题内容等），即可由远及近，或由近及远查找所需文献。

### 二、查找技巧

#### （一）著者引文索引（Citation Index： Author）

《著者引文索引》是由被引证文献著者姓名查找引证该文献的新著者及其文献题目的索引。此索引按被引证文献的第一著者姓名英文字顺排列，同一著者有多篇公开发表的文献，则按其发表先后的时间次序排列。引证文献

的新著者也按其姓名英文字顺排列。某权威学者即使很早以前在期刊上公开发表的论文，只要被人在当今期刊上发表的文章中引证了，该权威学者姓名必将收录在著者引文索引中，在其大名和刊登论文的刊物名称下列出所有引证过他论文的新著者、出版物、卷次、起始页码等。

查《著者引文索引》的前提条件是，事先必须确切知道某学科专业研究领域知名专家学者姓名，以此为线索，发掘、获得该专业研究领域更多、更新、更有学术技术水平的文献。

《著者引文索引》的用途，是提供读者（科技人员）从已知被引证著者姓名，查找和扩大引用（证）的新著者及其文献题录与出处的途径。

#### （二）专利引文索引（Patent Citation Index）

《专利引文索引》收录所有被引证过的专利，并揭示各篇专利被他人引证的次数。《专利引文索引》中的专利均按被引证专利号由小到大顺序排列。

在《科学引文索引》的《专利引文索引》中，被引证专利的重要内容依次排序为：被引证的专利号（黑体字）、被引证专利的年份、专利发明者、新申请专利（APPI）或再版专利（Reiss）、国别；而引证专利的重要内容依次排序为：引证著者、引证出版物、文献代号、卷、页次、年份（见SCI—1下）。接着，按引证著者姓名的英文字顺查找同一期的来源索引，即可得知引证著者引证专利后写成的论文题目及其刊载的期刊名称、卷、期、起止页码及年份。

#### （三）轮排主题索引（Permuterm Subject Index）

《轮排主题索引》为一种题内关键词索引。先选取对引证文献篇名中有实质意义的关键词进行轮排。系指从来源论文中选出能反映其主要内容的关键词按字顺轮流进行相互组配。所有被选出的关键词都可以作为主要词（Primary Term）或配合词（Co - Term）轮流组排，以提高检出论文的专指性。轮排主题索引的优点是可从若干个关键词查找到同一篇论文，并由配合词后所列出的论文著者姓名转查《来源索引》，获得该论文的题目、出处及参考文献顺序号（见SCI—2）。在所查找课题事先并不掌握文献资料或相关著者情况下，宜采用此索引查找。但用的是非规范化主题词，致使文献分散，查全率低，应使用参见（Sa）弥补，Sa为Seeabo的缩写“也见”之意。

#### （四）团体索引（Corporate Index）

此处“团体”专指来源著者的工作单位（公司、研究所、院校等）。《团体索引》由地理与机构两部分构成，后者是按来源著者工作单位名称英文字顺排列，其后列出该工作单位所在国家（美国为州名）和城市。以此查找来源著者姓名和来源出版物的索引。

《团体索引》的优点是，通过它可以了解和掌握国内外某个医疗或院校、科研部门在任意一个时间内（一年或若干年）总共发表了多少篇论文，以及都是谁在什么刊物上发表了具体的论文或专利名称。据此，一方面可熟知各国有关单位的研究水平与动态，另一方面可对同类研究单位的研究水平、学术水平进行评价，同样，也可以从具体专家学者一特定时间内发表论文多少，对其学术技术水平与知名度作出评价。

根据查找前提供的条件不同，《团体索引》又分为地名部分（Geographic Section）与机构部分（Organization Section）。其中：

机构部分系按各机构名称英文字顺排序，据此可查找到该具体机构所在

的国家(美国为州名)和城市名称,以此作为进行地名部分查找的辅助索引,但并不提供论文题目及其出处。

地名部分是将来源论文按著者工作单位所属国家(美国为州名)、城市、工作单位名称三者先后顺次排列(见 SCI-3)。其中,工作单位名称视其隶属关系不同而分级不一,有的可分为4级,如大学、学校(隶属大学)、系、教研室。

#### (五) 来源索引(Source Index)

《来源索引》是将《著者引文索引》、《专利引文索引》和《团体索引》中可查到引文的著者及其所发表论文或专利的出处与发表年份等,水管其资料所属类型,一律按著者姓名英文字顺排列(SCI—4)。匿名文献,如编辑部文章、评论、按语、校正、通讯、会议资料等,按其出版物名称的英文字顺排列在来源索引的最前面。

《来源索引》中只选用论文或专利的第一著者姓名作为查找对象,并顺序排列。次要著者一律以 See 引见第一著者。《来源索引》中的论文或专利不提供题目,只给出其刊载的刊名(或图书号)、卷、页次、发表年份、参考文献顺序号。

《来源索引》实际上是科技文献的著者索引,或叫《科学引文索引》的著者索引。

由《科学引文索引》能查找到某学者某年某月在某种出版物上发表过某篇论文;某学者的某篇论文、某篇专利已被他人所引证,以及引证的论文又发表在何种出版上;世界各国有哪些单位在研究和自己相同或相关的科研项目等等。

综上所述,《SCI》的手检经过可简单归纳如下:

由被引证著者查找:被引证著者 著者引证索引 引证著者 来源索引  
题录出处 原始文献。

由被引证专利查找:被引证专利号 专利引证索引 文献出处 题录出处  
原始文献。

由主题查找:主题词 轮排主题索引 文献著者 来源索引 题录出处  
原始文献。

由团体著者查找:机构名称 机构引证索引机构部分 该机构所在地(州)、国(名,城市名) 地理部分 机构所在地 机构名称 著者姓名与著作出处 来源索引 题录出处 原始文献。

由个人著者查找:著者姓名 来源索引 题录出处 原始文献。

### 第三节引文分析方法应用举例

我们依照引文分析方法的原理和步骤,列举了一个实例,以此具体说明引文分析方法在信息研究中的应用。从而使读者尽快有效地使用引文分析方法于工作实践中。

#### 一、资料来源与研究方法

本文选择《环境科学学报》等10种国内环境科学核心期刊为统计分析和研究对象。采用文献计量学的方法,通过对10种期刊最近3年(1989~1991年)所载文献的引文分析和

作者分析,得出我国环境科学文献引文的一般规律及我国环境科学研究

作者队伍构成，为进一步深化我国环境科学研究提供必要的依据。

## 二、引文的统计分析

通过对科学论文的引文统计分析，可以了解论文作者吸收情报的能力。

### (一) 引文的数量

引文数量的多少，在不同程度上反映出文献所蕴藏的信息含量。

表 5—1 10 种期刊引文量统计

期刊名称	论文总篇数	有引文的百分比	篇均引文量(条)
球境科学学报	213	98.12	9.54
中国环境科学	269	88.10	6.12
环境科学	369	94.03	6.52
环境化学	233	93.99	8.42
海洋环境科学	187	90.90	8.82
中国环境监测	378	43.65	3.45
环境科学丛刊	89	73.031	2.33
环境科学研究	114	71.92	5.55
上海环境科学	570	50.00	2.32
环境科学与技术	162	61.11	4.25
合计, 平均	2584	76.48	6.73

1. 有引文的论文比例。在我国 10 种环境科学期刊中,《环境科学学报》有引文的论文比例最高(占 98.12%),《中国环境监测》有引文的论文比例最低(占 43.65%)。10 种期刊有引文的论文其平均比例为 76.48%(见表 5—1),比国外科技期刊的平均值 90%和我国科技期刊的平均值 88%低。

2. 篇均引文量。在 10 种期刊中,《环境科学丛刊》的篇均引文量最高,为 12.33 条,《上海环境科学》的篇均引文量最低,为 2.32 条。10 种期刊的篇均引文量(6.73 条),不仅低于国外科技期刊篇均引文量 15 条的指标,而且与其它自然科学中文期刊的数据相比(见表 5—2)也显得太低。表明环境科学期刊所载论文,在利用吸收现有科研成果方面需要加强。

表 5—2 不同学科中文期刊平均引文量数据比较

学科/天文学/生物学/海洋科学/化学/物理/环境科学/地震学  
平均引文量(条)/14.87/12.10/8.85/8.84/8.40/6.73/5.56

表 5—2 不同学科中文期刊平均引文量数据比较

学科	天文学	生物学	海洋科学	化学	物理	环境科学	地震学
平均引文量 (条)	14.87	12.10	8.85	8.84	8.40	6.73	5.56

### (二) 引文的学科

学科之间的交叉渗透在一定程度上反映了学科之间的亲缘关系,而引文分析正是透视和分析这种关系的行之有效的科学方法。10 种环境科学期刊引文共涉及 40 多个学科,其中引自环境科学文献占 56.20%,其它自然科学文献占 41.27%,社会科学文献占 1.40%,综合学科文献占 1.23%。由此得出:

在全部引文中,环境科学文献被引率最高,说明我国环境科学发展较快,

已初步形成了自己的学科体系。环境科学与其它自然科学相关学科关系较为密切，特别是化学和医药科学，分别占全部引文的 11% 和 7%。

### (三) 引文的文献类型

研究引文的文献类型分布，可以了解本学科论著的文献来源和构成比例，从而确定各类文献载体的信息价值。在 10 种期刊的 14753 篇引文中，引自期刊的引文为 9849 篇，占全部引文的 66.75%；引自图书的引文 3660 篇，占 24.80%；引自其它文献（包括论文集、标准规范、学位论文、报纸等）共 1246 篇，占 8.45%（见表 5—3）。由表 5—3 的数据得出这样一个结论：我国环境科学期刊文献的主要引文来自期刊，平均占 66.75%，与世界科学研究活动信息量有 65%~70% 来自科技期刊的结果相一致。比我国自然科学期刊 56.50% 的文献来自期刊高出 10.23%。说明期刊已成为我国环境科学研究的最主要信息来源。

表 5—3 10 种期刊引文类型分布

期刊 (篇)(%)	图书 (篇)(%)	论文集 (篇)	标准 (篇)	其它资和 (篇)	报纸 (篇)	合计 (篇)
9849 66.75	6660 24.80	387	147	674	38	14755

### (四) 引文的语种

引文语种分布是测度期刊作者利用国外文献和该刊信息吸收能力的指标。

在 14755 篇引文中，引用英文文献为 7403 篇，占全部引文的 50.17%；引自中文文献为 7186 篇，占 48.70%；引自其它语种文献（包括日文、俄文和德文）166 篇，占 1.13%（见表 5—4）。在 10 种期刊中，《环境化学》英文引文比例（65.85%）最高，《环境科学与技术》中文引文比例（68.40%）最高。

表 5—4 10 种期刊引文语种分布

英文 (篇)(%)	中文 (篇)(%)	日文 (篇)	俄文 (篇)	德文 (篇)	合计 (篇)
7403 50.17	7186 48.70	106	48	12	14755

从表 5—4 看出：在我国环境科学研究中，英文是最常用的一种外语，其次为中文。日文、俄文和德文的文献引用较少，仅占 1.13%，与世界上这些语种文献占 23.00% 之间存在较大距离。说明我国环境科学研究者中掌握日文、俄文和德文的人太少，或者与有关部门这些文种的文献数量不足有关。

### (五) 引文的年代

从时间角度对引文的分布规律进行研究是引文分析的重要内容之一。它可以反映出被引用文献的出版、传播和利用情况，特别是在文献老化和科技史的研究中，引文年代分布的分析更是一种广泛而有效的科学方法。从表 5—5 可以看出：

1. 最大引文年限：最大引文年限反映了文献最活跃，最有生命力的时期。10 种期刊的引文频次，中文在有关文献发表后第 5 年达到最大值，外文在有关文献发表后第 9 年其利用率达到最大值。中外文引文年限分布呈现集

中与离散两个特征。即当年值低，中文4~5年后，外文7~9年后达到高值，然后逐渐减少。表5—5 10种期刊中外文引文年限等情况统计

中文引 文年限	引文 频次	频率 (%)	累积频 率(%)	外文引 文年限	引文 频次	频率 (%)	累积频 率(%)
当年	26	0.4	0.4	当年	5	0.1	0.1
第2年	244	3.5	3.9	第2年	53	0.7	0.8
第3年	455	6.6	10.5	第3年	132	1.8	2.6
第4年	809	11.8	22.3	第4年	263	3.5	6.1
第5年	933	13.5	35.8	第5年	415	5.6	11.7
第6年	763	11.1	46.9	第6年	490	6.6	18.3
第7年	778	11.3	58.2	第7年	529	7.1	25.4
第8年	591	8.6	66.8	第8年	547	7.4	32.8
第9年	627	9.1	75.9	第9年	603	8.1	40.9
第10年	455	6.6	82.5	第10年	519	7.0	47.9
第11年	317	4.6	87.1	第11年	470	6.3	54.2
第12年	237	3.4	90.5	第12年	451	6.1	60.3
:	:	:	:	:	:	:	:
第22年	3	0.1	98.3	第22年	94	1.3	90.7
第23~ 32年	89	1.3	99.6	第23~ 32年	478	6.4	97.1
第33~ 第42年	28	0.4	100	第33~ 82年	217	2.9	100
合计	6.887	100	合计	7423	100		

从表5—5还可以看出，中文近6年内发表的文献被引频率达到46.90%，前12年内发表的文献被引频率达到90.50%。从而推测出中文环境科学期刊的使用期限在12年左右。外文近10年内发表的文献被引频率达47.9%，前22年内发表的文献被引频率达到90.7%。从而推测出外文环境科学期刊的使用期限在22年左右。

2，半衰期：半衰期是指现在正被利用的某学科文献中，50%的近期文献的出版年限。它反映某一学科专业发展的兴衰以及学科专业期刊论文的使用寿命。从表5—5看出，10种环境科学中文期刊的半衰期为6.3年，外文为10.4年，比表5—6有关学科中文期刊半衰期要长。

表5—6 有关学科中文期刊半衰期

学科名称	半衰期(年)	学科名称	半衰期(年)
数学	4.6	农业科学	4.9
力学	4.9	医药学	4.4
天文学	4.8	地球科学	5.4
物理学	3.9	材料与工程	4.2
化学	4.4	信息科学	3.5
基础生物学	5.3	综合类	4.4

### (六) 引文期刊分布

通过对引文期刊分布特征的分析,能够获得环境科学的核心期刊。9849篇期刊引文共来自274种中文期刊(不包括中文期刊),其中引用中文环境科学期刊70种、被引用次数大于50次的环境科学期刊共10种,约占全部被引中文期刊内3.7%,而这10种期刊总共被引次数则达到1683次,占全部被引期刊被引次数的17%(见表5—7)。

可见这10种期刊对我国环境科学研究者的研究活动起了十分重要的作用,它们是我国环境科学的核心期刊。

### (七) 期刊自引率

期刊论文引用本刊已发生的文献与该刊全部引文之比,称为期刊自引率。10种期刊自引率见表5—8。从表5—8看出,《海洋环境科学》自引率(8.06%)最高,其次为《中国环境科学》(5.4%)。其它8种期刊自引率相对偏少。10种期刊自引率平均为3.85%,这与自然科学中大部分学科的期刊自引率5%~10%相比要低。

## 三、作者的统计分析

表5—7 中国环境科学核心期刊

表5—8 10种期刊的自引率统计

刊名	引文总篇数	自引篇数	%
环境科学学报	2033	54	2.65
中国环境科学	1648	89	5.40
环境科学	2409	87	3.61
环境化学	1962	75	3.82
海洋环境科学	1650	133	8.06
中国环境监测	1306	51	3.90
环境科学丛刊	1098	14	1.97
环境科学研究	633	10	1.57
上海环境科学	1326	40	3.01
环境科学与技术	690	15	2.17
合计	14755	568	3.85

### (一) 作者人数分布

10种期刊共发表2584篇文章,其作者人数分布见表5—9。

由表5—9看出,约有89.59%的论文作者人数分布在1~4人之间,最

多为 15 人；环境科学研究课题近 54.3%，作者为 2~4 人。

表 5—9 作者人数分布

统计还表明，10 种期刊论文合著率最高的是《环境科学学报》，为 82.63%，高于《动物学报》的 79.7%，《数学学报》的 47.5%；低于《化学学报》的 98.1%，《微生物学报》的 95.4%，《物理学报》的 87.6%，《植物学报》的 84.8%。10 种期刊的论文合著率平均为 64.71%，低于上述有关自然科学学报的该项百分值。说明环境科学研究还应加强集体协作，这对提高研究质量是有积极意义的。

### （二）平均作者数分布

10 种期刊作者共 6264 人次，篇平均作者 2.42 人。低于《中国科学》(B) 的 3.27 人，《物理学报》的 3.11 人，《化学学报》的 3.45 人，高于《地质学报》的 1.83 人，《计算机学报》的 1.75 人和《数学学报》的 1.24 人。

### （三）作者的系统分布

按作者单位所属系统划分，10 种期刊作者的系统分布见表 5—10。

由表 5—10 看出，高校作者发表论文最多，占 32.12%，说明高校作者是我国环境科学研究队伍中最主要的力量。省市环境所（站）作者发文量居第二位。

表 5—10 作者所属系统分布

续表

作者系统	发文篇数	%
国家海洋局及所属所	122	4.72
其它(政府各部门、部委单位)	108	4.18
企业公司	66	2.55
县级环保监测站	32	1.24
单位不明	14	0.54
合计	2584	100

### （四）作者的地区分布

研究论文作者的地区分布，可以确定各地区的环境科学科研队伍的状况，进而了解其科研能力的大小，为国家统一协调环境科学研究力量提供必要的参考数据。10 种期刊作者发文地区分布见表 5—11。

由表 5—11 看出，论文作者分布遍及全国（除西藏外）各省市自治区，这足以说明环境科学研究队伍在全国各地已基本形成。作者群主要分布在北京、上海，湖北、江苏、辽宁、浙江和广东等七省市。七省市作者论文数量占总数的 112.09%。表明七省市是我国环境科学文献产生的核心地区。

### （五）作者的合著类型分布

考察不同作者合著的类型对于揭示科学工作者论文的生产能力，以及掌握科技发展水平具有一定的意义。10 种环境科学期刊论文合著类型见表 5—12。

由表 5—12 可以看出，同一单位同事间的合著类型的比例最高，占 57.41%。跨地区跨单位间的合著类型排第二，占 37.46%，高于《金属学报》的



18.94%。说明我国环境科学工作者国内的合作交流通畅，横向联合与协作较强，较好地发挥了各地区各单位的科研优势。国际间的合著类型占 5.13%，与我国自然科学期刊的 5.56% 相接近。

表 5—11 作者发文的地区分布

地区	论文数	百分比	地区	论文数	百分比	地区	论文数	百分比
北京	584	22.60	吉林	55	2.13	新疆	11	0.43
上海	436	16.87	河北	53	2.05	贵州	7	0.27
湖北	229	3.36	安徽	34	1.32	宁夏	6	0.23
江苏	223	8.63	甘肃	30	1.16	海南	4	0.15
辽宁	178	6.89	黑龙江	29	1.12	青海	3	0.12
浙江	111	4.30	山西	29	1.12	香港	1	0.04
广东	102	3.95	陕西	29	1.12	日本	1	0.04
山东	78	3.02	江西	25	0.97	德国	1	0.04
四川	69	2.67	云南	21	0.81	加拿大	1	0.04
天津	58	2.24	内蒙古	17	0.66	地区不明	18	0.70
福建	58	2.24	河南	16	0.62	合计	2584	100
湖南	55	2.13	广西	12	0.46			

表 5—12 作者合著类型分布

## 第六章信息实践中的经验科学方法

在一般科学方法中，有一类方法是以通过与客观事物的直接接触而获得研究所需数据、素材和知识、技能的，如专家调查法、信息咨询法、观察方法和预测方法等，我们把这一类科学方法统称为经验科学方法。经验是一切认识的起点，经验知识是理论构成的出发点，也是理性思维的必要素材。当然，经验只有上升到理性认识，才能把握事物的本质，更正确地认识世界和改造世界。在信息学研究中，我们需要要经验知识，而专家调查法等经验方法使这些知识和素材的获得成为可能。

### 第一节专家调查法

专家调查法或称专家评估法，是以专家作为索取信息的对象，依靠专家的知识 and 经验，由专家通过调查研究对问题作出判断、评估和预测的一种方法。

专家调查法应用广泛，多年来信息研究机构采用专家个人调查法和会议调查完成了许多信息研究报告，为政府部门和企业经营单位决策提供了重要依据。

60年代中期，国外许多政府机构和公司企业热衷于建立电子计算机数据处理系统，但是，实践表明，利用专家头脑的直观判断仍具有强大的生命力，专家的作用和经验是电子计算机无法完全取代的。在许多情况下，只有依靠专家才能作出判断和评估。

在下列3种典型情况下，利用专家的知识 and 经验是有效的，也是唯一可选用的调查方法。

**数据缺乏：**数据是各种定量研究的基础。然而，有时数据不足，或数据不能反映真实情况，或采集数据的时间过长，或者付出的代价过高，因而无法采用定量方法。

**新技术评估：**对于一些崭新的科学技术，在没有或缺乏数据的条件下，专家的判断往往是唯一的评价根据。

**非技术因素起主要作用：**当决策的问题超出了技术和经济范围而涉及到生态环境、公众舆论以致政治因素时，这些非技术因素的重要性往往超过技术本身的发展因素，因而过去的数据和技术因素就处于次要地位，在这种情况下，只有依靠专家才能作出判断。

此外，由于原始信息量极大，决策涉及的相关因素（技术、政治、经济、环境、心理、文化传统等等）过多，计算机处理这样大的信息量，费用很高。这时，从费用效果考虑，也应采用专家调查法。

#### 一、专家个人调查法

顾名思义，专家个人调查法是由专家个人进行调查、分析和判断的方法。这是一种由来已久的调查研究方法，科技专家为了科技探索选取课题、开发应用等自身研究的需要，通常对本技术领域及相关领域进行调查，以了解现状和发展趋势，这种调查是自发进行的。

信息研究人员大多是应某一部门、团体或企业的要求和委托而开展调查研究的。自建立信息机构以来，各级信息研究所为领导部门决策服务进行的调查研究多数是由信息学专家个人完成的，因而信息所采用专家个人调查法

最多且是经验丰富的部门。信息学专家个人调查的成果主要是信息研究报告，如航天信息所编印的《航天情报研究报告系列文集》、兵器信息所编印的《兵工情报研究报告》等，其中绝大部分是信息研究人员个人调查研究的成果。

专家个人调查，作为一种供选择的方法，它的优点是：

1. 可以充分利用专家个人的知识和经验，最大限度地发挥专家个人的创造能力；
2. 不受外界影响，没有心理压力，这种影响在专家会议上通常是较明显的；
3. 组织工作简单，调研费用很少，因而简便易行。

由于这种方法对某些问题可以作出有效的分析和判断，又简便易行，因而在没有条件使用或不必要使用其他方法的情况下，专家个人调查法往往被认为是能较快获得研究结果的方法。1975年国外有人对世界许多国家使用的预测方法进行过调查，专家个人调查在各种预测方法中使用的比例，分别为8.7%（机构使用的方法中）和8.2%（个人使用的方法中）。

部门、团体和企业需要了解某项新技术、新产品、新工艺的现状和发展趋势时，最适宜采用专家个人调查法，由调查者写出综述或评述，提供所需要的信息；领导机关为了制订规划、方案论证、技术引进等多方面的需要，可把由专家个人撰写的信息研究报告，作为决策的参考和依据；专家个人调查还可为专家会议提供背景材料。当专家们对某项技术、某个方案存在较大分歧时，科研领导机关有时在召开专家会议进行调查之前，选择意见不同的、有代表性的若干专家，分别提出几个调查报告供会议讨论，这种作法经常能收到较好的效果。

专家个人进行调查，由于个人的知识面、知识深度、工作经验，以及占有的信息，很难满足调查所要求的全面性；此外，还可能受专家个人兴趣的影响，难免产生片面性。有时由专家个人作出的直观性判断，其他人难以审查其正确与否。这些缺点是专家个人调查法难以避免的。

## 二、专家小组讨论法

参加专家小组讨论会的专家一般来自某一专门领域，例如科学技术领域、经济领域、军事领域等。这种不跨领域的特点正是专家小组讨论法与智囊团式讨论法的一个重要区别。

科研部门在制订规划、技术论证、确定方案、引进项目时，通常召开专家会议进行调查，征询专家们的意见，各种不同观点和意见要在会议上展开充分讨论，以求达成一致意见，作为决策的依据。

与专家个人调查法相比，专家小组讨论法具有以下3方面的优势：

1. 不仅专家们收集的信息量可弥补专家个人之不足，更重要的是在该领域内有各种学术观点的代表，可防止专家个人的片面性。
2. 对讨论对象可能受到各种因素的影响，考虑得更为全面。个别专家忽视的目前影响较小而在未来发展过程中将起重要作用的因素，可由其他专家考虑和补充。
3. 专家会议集体提出的判断多于专家个人的判断，通过比较和选择，可以提供最有价值的判断。

由于专家小组讨论法具有以上优点，这种方法在现代预测活动中应用较为广泛。国外有人统计，在各种预测方法中使用的比例分别为10%（机构）

和 8% (个人)。

当然，由于专家小组讨论法是即席发言，因而会有许多不足，主要是由心理因素造成的，主要包括以下几方面：

(1) 多数压服少数：在会议过程中，多数人的意见对少数人或个别人总是形成一种心理压力，即便多数人的意见是错误的，少数人也可能屈服于多数人的意见，而使错误主张获得通过。

(2) 权威影响集体：通常权威有较大的影响力。如果权威人物为了推行其个人意志而施加各种压力，则必然破坏专家小组周密而慎重的研究，不能发挥集体智慧，最后形成的结论很可能只代表某个权威的意见，这样就失去了该方法的——不曰。

(3) 论数据量比质量占有优势：在争论中，提供论据数量较多的一方往往占有优势，如果这些论据不能进行定量化分析则更为明显。因此，只要少数专家接二连三地抛出各种论据，即使他们的论据华而不实，也可能压倒其他人的正确意见。

(4) 利害关系的干扰：有时专家小组的成员出于利益方面的考虑，参加辩论不是为了如何获得正确的意见，而是力图达到自己的目的，千方百计想把别人拉到自己一边来，甚至大喊大叫，使会议不能正常进行。

(5) 低质量的折衷：在意见分歧的情况下，如果某权威感到自己的意见论据不够充分，但又碍于面子不愿修正自己的意见，或者个别专家由于利害关系坚持己见，在这种情况下，会议为了能达成一致意见而相互妥协。这种折衷满足了少数人的一部分利益，也不伤害任何一位专家的自尊心，结果是专家集体不能作出高水平评价和判断。

此外，参加会议的专家们带来了许多信息，其中可能有不准确的甚至错误的信息，这些信息在会议讨论的有限时间内很难及时得到改正，以致对讨论产生不利影响。

针对专家小组会议法的缺点，创造了头脑风暴法和德尔菲法。经验表明，头脑风暴法和德尔菲法能有效地避免专家小组会议法由于心理因素造成的不利影响。

### 三、头脑风暴法

头脑风暴法 (Brainstorming) 在创造工程学中系智力激励法。所谓头脑风暴，简单地说，就是激发创造性思维。由 A.F. 奥斯 (Osborn) 首创的这种方法，其目的是获得有价值的设想，方法是制定一套讨论规则，在短时间内造成思想非常活跃的气氛，诱发出大量的创造性设想。

头脑风暴法的基本思想是，若要得到有价值的设想，首先要能提出较多的设想。设想的数量越多，则获得有价值的创造性设想的概率就越大。

头脑风暴法，由于把提出创造性设想作为一个非常重要的阶段来进行，而把评价设想的可行性以形成最终可行方案作为另一个阶段，这样，既充分发挥了专家的创造性，又全面评价了设想的可行性，把创造性和现实性很好地结合起来，从而可获得极有创见的可行设想或方案，避免专家小组讨论会通常存在的缺点，排除低质量折衷方案。

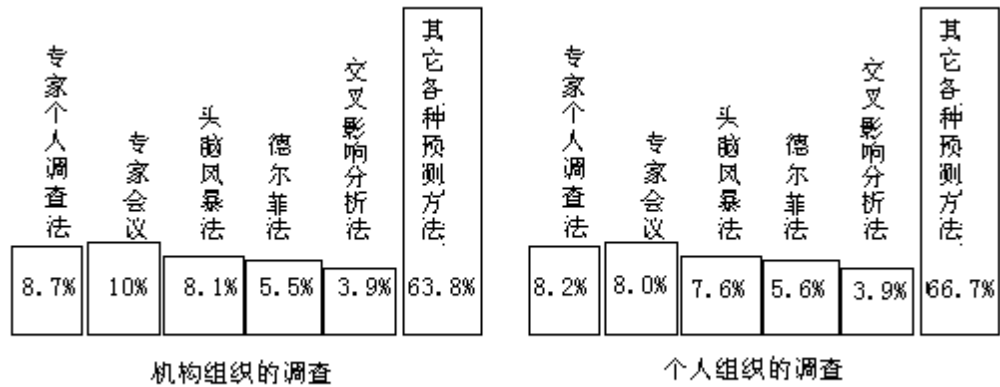


图6-1 专家调查法与其它各种预测方法的使用比例

采用头脑风暴法，提出一组切实可行的方案，供决策者研究，在国外已广泛应用于军事和民用预测工作中。例如，美国国防部在制订长远科技规划时，邀请50名专家采用头脑风暴法开了两周会议，与会者的任务是对事先提出的工作文件提出建议，并通过讨论把文件变为协调一致的报告。通过质疑，原工作文件中只有25%~30%的意见得到保留，结果形成了一个新的、更可行的规划。由此可见，头脑风暴法是一种很有价值的方法。它在各种预测方法中使用的比例分别为8.1%（机构）和7.6%（个人）。

## 第二节 信息咨询方法

信息咨询是文献信息机构的一种重要的服务形式，也是信息研究的一种成果。鉴于社会咨询业的普遍兴起，信息咨询与信息咨询方法的作用将日益显著。

### 一、信息咨询的基本概念

#### （一）什么是咨询

咨询现象古来有之。从哲学角度分析，任何个人或机构对客观事物的认识必然存在局限性，这是产生咨询需求的客观基础。

“咨询”一词在古汉语中并不是从一开始就构成一个词。《尚书·舜典》中有“咨十有二枚”和“询于四岳”。《春秋左传·襄公》中有“咨事为谏”、“咨亲为询”、“咨礼为度”和“咨难为谋”。咨询一词连用，最早见于东汉人所著《楚辞章句》一书，该书“九思·疾世”章中有“纷载驱兮高驰，将咨询兮皇羲。”这距今已有1800多年。

咨询一词在英语中为CONSULT，在法语中为CON - SULER，在拉丁语中为CONSULTO，其基本意思都是与别人商量。在国外同样出现过类似“巫”的职业。

可见，咨询活动历史悠久，但长期以来没有社会化。只是到了19世纪以后，随着政治、经济、军事和科技的不断发展，各个领域、各个层次的业务、技术和管理问题不仅数量增加，而且越来越复杂，单纯依靠个人的经验和智慧已难以正确解决这些问题，再加上社会分工的不断细化，信息流量的不断膨胀，咨询活动成为一门独立的行业已成为可能。例如，经济的发展促使建设项目大量的涌现。这些项目往往面临错综复杂，并涉及多学科、多行业的问题。而项目承担者又不可能拥有如此多的专业人员，或者说没有经济

力量承担这么多仅使用一时的专业人员来解决这些问题。这就要求社会分工出一种行业，从事这种行业的机构集中了一批能解决用户面临问题的专业人员，能够满足用户“召之即来，来之即战”的要求。于是，零散的咨询活动逐渐形成了咨询业。

那么，到底什么是咨询呢？目前，尚无统一的严格定义。现将几种主要提法罗列如下：

1. 咨询就是询问、征求意见、商量办法。或者说是一种向客户提供专门知识、技术和经验，帮助客户解决各种疑难问题的活动。

2. 咨询是“咨询方（咨询人员或咨询机构）根据委托方即客户（国际组织、政府机关、社会团体、企事业单位乃至个人）提出的要求，以其专门的知识、信息、技能和经验，运用科学的方法和先进手段，进行调查、研究、分析、预测，客观地提供最佳的或几种可供选择的方案（或建议、报告等），帮助委托方解决复杂问题的服务。”

3. 美国 DAI 咨询公司（Development Alternatives Inc.）认为“咨询就是向单位、机构或企业有偿地提供符合要求的专门技术、管理经验和专业知识。”

4. 国际劳工局编写的《管理咨询—专业指南》中则把咨询翻译为“咨询在本质上是一项参谋性的工作”，“咨询是一项提供与实际管理问题有关的专业知识和技能的服务工作。”

5. 在信息社会里，政府机关、商业机构及其他客户都离不开信息，而咨询就是掌握和提供信息。

不难看出，这些不同的提法是从不同的角度给咨询下定义。但它们所反映的咨询工作的实质却是一致的，简而言之，即：咨询就是委托方提出问题，并征求解决问题的方案；受托方出主意、想办法，提出解决问题的建议。

## （二）信息咨询的含义

信息咨询是信息研究机构依据用户的需求提供的关于信息的咨询。课题信息研究的成果，也可视为对用户的一种系统的咨询。

信息咨询的定义与上述咨询类似，它是人们在办事前追求了解情况的一种求知本能。信息咨询研究是咨询人员或信息研究人员对咨询课题的一种科研工作。

信息咨询的日常工作可以分设在文献信息部门或信息研究机构。社会上的咨询业也设日常信息咨询部门，它对人们进行信息或相关知识的咨询。咨询机构主要承担对用户需求的研究，承担力所能及的咨询课题或一些能够即席回答的咨询问题。

信息咨询不同于利用一般的文献资料。咨询人员对所答询的信息质量有责任，甚至是法律责任。因此，答询要慎重。

## （三）信息咨询与一般咨询业的关系

1. 产生于不同层次的需要：随着科学技术的飞速发展，特别是战后出现了把军事科技广泛转为民用的现象，各发达国家的经济迅速发展，形成了大工业、大工程、大科学的趋势，人类面临的重大决策也变得十分复杂起来，这就需要有专门的研究机构提供准确的信息，需要有专门的机构研究决策。

在这种日益增长的不同层次的需求下，一部分科研人员，利用自己的专业知识，转向文献信息部门，从事科技信息工作和信息研究工作（信息分析研究）；另一部分更高层次的科研人员（含社会、人文科学）利用自己的专

业知识和权威，转向独立的咨询业（又名思想库，头脑工厂）。前者成为科技信息业的开创者，后者成为社会咨询业的先驱。

成千上万的用户中，有的只需要信息咨询，自己进行决策研究；也有的要求全面的决策咨询，就是把信息研究与决策研究进行综合，拿出解决课题的决策方案。只要这两种需求存在，这两种研究也就会同时存在，或协作并存。

2. 具有不可分割的联系：根据用户的不同层次需要，咨询可分为两个系统，但咨询研究离不开信息研究，二者又连成一个整体。咨询过程实际上就是获取、传递和反馈信息的过程，是信息人员和咨询人员的一种信息交流，二者都是对知识的再创造、再组织，都是共同开发利用知识资源。

在管理还未形成科学之前，决策是由领导凭经验作出的。实现科学管理以来，财产权与经营权分离，决策也逐渐科学化，即由专门的研究人员进行信息研究和决策研究，然后再由决策人抉择。但是，不是所有用户都要求进行决策咨询。这样，在西方除产生了咨询业之外，还产生了“信息分析中心”这样的机构。美国在第一次世界大战前，就在工商界建有 13 个这种“中心”。他们只提供信息知识，即信息咨询。

在中国也是先产生信息研究机构，然后才产生咨询机构。它们形式上是分立的，但在工作上相互联系，互补优缺长短的。

#### （四）咨询业与文献信息工作的关系

从工作上讲，咨询服务主要是一种信息服务。因而，它应当是整个文献信息工作的一种服务形式。咨询服务由于服务内容广泛，它必须组织各种专业人员为咨询进行研究，其中主要是信息研究人员，也有文献研究人员和文献工作人员等等。他们为某个目标进行着各有侧重的智能化劳动，共同开发利用信息资源（知识资源），用以解决复杂的决策问题。因此他们同属于利用知识的科学研究工作。

### 二、信息咨询的任务

#### （一）宏观任务

信息咨询工作是一项社会性很强的工作，为了把工作搞活、搞好，我们必须充分发挥地方综合性信息研究所特有的职能管理作用，采取有效的组织形式，开发社会上的人才资源，将分散在大专院校、科研单位和大型企业的具有较高水平的技术人员组织起来，特别要充分发挥各专业信息网的作用。现阶段，信息咨询服务工作的重点在于协助企业开发具有最佳经济效益的新产品、新技术、新工艺。要想卓有成效地搞好这项工作，必须将科技信息工作与经济信息、市场信息、管理信息等项工作结合起来，把科技信息与技术、经济、社会兼容起来，这样科技信息咨询服务才能做到“更有效”。

在科学技术迅猛发展的今天，情报信息的准确、及时与否，往往决定着经济发展的成效。因此，科技信息机构要加强与政府决策部门、企业生产部门、市场销售部门、对外经贸交流部门的有机联系，建立起技术、经济、贸易相结合的综合信息交流网络，使科技信息咨询服务做到信息准确、有的放矢。

#### （二）微观任务

开展咨询服务的机构必须重视加强基础建设，即信息机构的物质建设和组织建设。当前，要继续大力加强文献基础工作，根据本地区、本部门的特点，广辟渠道，有选择、有重点地搜集、存储科技资料；要逐步建立和

健全以本所为中心的本地区本部门科技信息计算机检索系统，做好微型电子计算机的推广工作；重视声相、缩微、复制和印刷出版等新技术的应用，逐步实现科技信息工作现代化。要重视智力开发，建立一支责任心强、熟悉业务、热心信息咨询服务工作的，质、量并举的信息咨询专职队伍，这支队伍是搞好信息咨询工作的重要保证。

在相当长的时间内，我国科技信息工作的主要矛盾是信息资料愈来愈多，信息用户的需要愈来愈迫切。因此，解决“找”的问题是科技信息工作的重要任务，也是信息咨询的重要任务。为此，要十分重视文献的搜集与管理工作，特别要收藏比较齐全的检索工具书；要开展用户培训，提高广大信息用户检索和使用文献的能力。要抓好咨询人员与科技人员，信息咨询部门与信息用户的两个结合，以提高信息咨询的针对性和查准率，提高信息咨询服务的质量和水平。同时，信息咨询人员要做到熟悉检索工具，熟悉信息用户，熟悉馆藏，这样才能做好咨询工作。

### 三、信息咨询的方法

#### （一）被动性咨询

属于这类咨询方法的主要有：检索性咨询：这是咨询人员针对用户已经确定的科研课题、科技攻关题目而提出的检索课题，利用中外检索工具查找所需要的文献资料、文献题录等；专题事项和数据性咨询：系指咨询人员不要查找大量文献资料，只要答复用户提出的某一具体事实、数据、某工艺的适用范围、某设备的性能参数、标准、公司名称及组织机构、外文资料中的缩写字母、疑难符号等；研究性咨询：系指为信息用户选定的研究课题，进行设计方案的论证或制定各行业的发展规划提供水平、动向、趋势和概况方面的信息；指导性咨询：系指对信息用户在阅览和查找资料时，或在实际工作中所遇到的问题给予指导。

#### （二）主动性咨询

主动性咨询是指信息咨询人员为了搞好、搞活咨询工作，主动出谋献策，力所能及地为信息用户排忧解难，帮助解决各种科技咨询问题，以达到振兴本地区、本部门经济为目的而采取的咨询方式。属于这类咨询的方法可以灵活机动，形式多样，是今后信息咨询改革的方向，各信息机构都在不断探索和创新。现已经采用和正在试行的主动性咨询方法主要有：推荐性咨询：系指信息咨询人员根据自己所掌握的信息用户需求情况，凭借自己平时对资料的积累以及对新到资料的“扫描”，选择一些适合本地区本部门情况、花钱不多，有推广价值的文献资料（适用生产技术信息），用电话、送资料上门或寄送“信息资料对口推荐函”等方式，主动为用户服务；专题性咨询：系指信息部门针对本地区、本部门某一时期科技工作中心或某一专业及普遍存在的问题，举办小型专题资料展览或巡回专题资料展览，进行专题信息交流和集中力量开展某一专题信息咨询服务。开发性咨询：系指为对口的企业研制新产品、新工艺、新材料或者为产品的更新换代、技术改造所进行的咨询；预测性咨询：系指根据特定的需要，利用各种科技信息手段，进行综合的、定性和定量相结合的科学预测。预测咨询服务的范围很广，有产品生产预测，原材料预测、销售预测、市场价格预测等。

此外，还有引进咨询、考察咨询、审核咨询、定向性（跟踪）咨询、新颖性咨询、市场信息咨询等多种信息咨询方法。总之，信息咨询是既能面向基层生产技术，又能面向上级决策的咨询服务工作，也是集文献、调研工作



于一体的工作方式，不仅可以无偿服务，而且可以实行有偿服务，从而使服务的针对性得到加强，提高工作成效。

### 第三节 观察方法

#### 一、观察方法的定义

观察方法是通过人们的感官或仪器，有目的有计划地对自然状态下的客观事物进行系统考察和描述的一种研究方法。它是人类最早使用的和最基本的研究方法之一。在科学方法论中，观察法属于获取感性知识和材料的方法。观察法是一种基础性的方法。

观察法分为直接观察法和间接观察法。直接观察指用感官直接感知对象；间接观察则指感官通过仪器设备观测对象。在人类认识自然的历史上，伴随着人类生产活动的进行，产生了直接观察法，而间接观察法则是在人类可以制造仪器之后产生的。随着资本主义工业的产生和发展，也产生和发展了越来越先进的仪器设备，从而发展了以仪器设备延伸或代替感官对自然现象进行观察测量的间接观察方法，并极大地提高了观察的深度、广度和精度，形成了科学的观察方法。在现代，任何科学观察都是定性观察与定量观察的结合，并以定量观察为主。观察方法在实验方法尚未形成之前，是科学研究的主要方法。

#### 二、观察方法的实施过程

观察方法的实施过程，一般可以分为以下几个阶段。

1. 观察准备。根据观察的目的，确定观察对象，选择观察方式，准备观察用品，提出问题，拟定观察计划。

2. 观察阶段。采用直接观察法或间接观察法进行实际观察，应该注意观察的严肃性，严格按观察程序进行。

3. 记录观察结果。在观察过程中，或在观察之后，做观察记录，如速记、卡片、记录、日志、声象记录等，并对观察记录进行检查。

4. 写观察报告。最后，要对观察活动进行分析总结，写出报告。报告的内容大体包括：观察的时间、地点、对象，观察方法，观察过程以及对观察者的特性分析，对被观察事实的描述和观察者本人的意见与判断等。

#### 三、观察方法在实际中的应用

观察方法的应用是：

1. 科学始于观察，它可以帮助人们获得比较可靠的第一手材料，为进一步研究提供事实和数据。

2. 观察可能直接导致科学上的新发现，这在天文学、气象学、心理学领域是常有的事。在信息学领域，也会通过观察发现一些过去未曾想到的问题，扩大和开辟新的研究领域。

3. 对科学假说和科学理论进行检验。在信息工作领域，有许多观点仅仅是设想（假说），有不少问题存在着分歧意见。那么，如果选择一定环境下的研究对象进行一定的观察，就可以初步验证这些假说和理论是否正确，而无需再作全程重复的研究。

#### 四、观察方法在信息研究中的功用

观察方法在信息工作中的具体应用并不少，只是我们没有专门留意罢了。信息学研究人员应当努力总结观察方法的应用情况，并以科学的形式把

观察结果表达出来。

观察方法在信息学研究中的应用主要表现在以下三个方面：

1. 广泛用于信息用户研究和对信息人员的研究。信息用户和信息人员作为活动着的研究对象，特别适合应用观察方法。通过观察、了解用户的个性特征，了解用户利用信息的习惯，了解用户对信息中心各种信息服务的反应，以便提供针对性更强的服务，改进信息服务工作，取得一些有价值的原始材料。观察法在信息人员研究中的应用比在用户研究中的应用更容易更方便。通过观察，可以了解信息人员的服务态度、工作习惯、情绪波动、心理变化等各方面的情况，对加强信息工作的科学管理大有好处。

2. 可以用于验证某些经验说法和理论观点。研究人员对信息工作中存在的问题有不同的解释和假设，而这些经验之说或理论观点正确与否，要从理论上证明是非曲直，往往既费事又难以见效，而有时候，不太费功夫的观察却可以得出结论，辨出是非。以如信息用户是选择手工检索的多还是计算机检索的多这一问题，你只需在一段时间内分别到本系统信息中心机检室和手检室看一下，比较一下用户人数，这一问题就非常清楚了。

3. 用观察来研究、改进信息服务。信息人员并不需要特别留意，就可以观察到信息用户对信息中心的服务方式、服务质量、规章制度、阅览室的布置，甚至咨询、检索、复制、外借等部门之间位置联系的抱怨与建议。这些抱怨与建议可以成为信息中心有的放矢地改进信息服务的很好依据。

#### 五、观察方法的不足

观察方法的不足有：

1. 研究人员难以控制事情的进展；
2. 观察的样本容量比较小；
3. 观察的结论难以用准确的数量表示；
4. 需要与其它方法一起来完成一个项目的研究。

### 第四节 预测方法

预测是科学决策和现代管理的基础，也是信息研究的重要内容。在很多情况下，科学预测与信息预测是密不可分的，因此，很大一部分所谓预测方法，也是信息研究的常用方法。

预测方法虽然很多，但截至目前为止，还没有一个统一的、普遍适用的分类体系。大体说来，可以将预测方法分为技术方法和定量方法两大类。预测的技术方法就是我们经常说的专家评估法或者直观判断法，这种方法的预测精度在很大程度上取决于预测者（专家）的技术和技巧。这种方法通常称为定性方法。然而现今以专家评估为基础的技术方法，是在定性分析基础上，以打分等方式作出定量评估，预测结果是具有统计特性的。预测的技术方法主要适用于宏观战略预测，当然也可以从事微观战术预测。技术方法主要包括：专家会议、交叉影响、目标预测、抽样调查和历史类推等。

预测的定量方法主要是利用原始数据，借助数学手段，建立数学模型从事预测。其预测精度和可靠性，在很大程度上取决于数据的准确性和方法的科学性。预测的定性方法包括：移动平均、指数平滑、回归分析、趋势外推、投入产出模型等。

对于信息研究中的预测方法，这里只介绍德尔菲法。

德尔菲法（又可音译为特尔斐法）又可称为专家集体预测法，是美国兰德公司于 1964 年创造的一种科学预测方法。德尔菲（Delphi）是古希腊传说中的神谕之地，城中有座阿波罗神殿，可以预卜未来。这里借用此名，表示一种预测方法。这个方法是在预测领导小组的主持下，就某个科学技术课题向有关专家发出征询意见的调查表，通过匿名函询的办法请专家们提出看法或进行论证，然后由领导小组汇总整理，把整理结果作为参考意见再发给这些专家，供进一步分析判断，提出新的论证。如此多次反复，按意见收集情况作出预测。30 多年来，德尔菲法已成为一种广为使用的预测方法。它是根据专家的经验 and 知识用直观方法作出的一种预测，这对以非连续性变化事件为对象的预测（如过去没有足够信息或相关因素比较复杂的技术领域，事件的发展主要取决于科技政策和主观能动性的领域等）较为有效。所以，在制订长远规划这类工作中，它是决策者的重要工具。

### （一）德尔菲法的基本程序

1. 确定预测目标。常常是影响大而且意见分歧的课题，可用于技术预测和技术市场预测。

2. 选择被调查的对象。由于是通过专家对未来事件作出概率估计，所以专家的选择直接影响预测的成败。对专家的主要要求有：专家的代表面应广泛，通常包括技术专家、管理专家、信息专家和高层决策人员；专家群众的权威程度要高，即使没有高级职称，但是长期接触、熟悉所调查的问题，亦可称为专家。在选择专家时，要严格专家的推荐和审定程序，审定的主要内容是了解专家对预测目标的熟悉程度和是否有时间参加预测的全过程。专家小组一般以 20~50 人为宜，大型预测可达 100 人。

3. 设计评估意见征询表。表格要简明扼要，设计得好的表格，通常能使专家的主要精力放在思考决断上，而应答填表的时间并不长。填表的方式要简单，对事件进行评估时，尽可能让专家用数字和英文字母表达自己的意见，而不必填写很多文字，填表时间一般应控制在 2~4 小时左右。

4. 专家征询和轮间信息反馈。经典的德尔菲法一般分 3~4 轮征询，第一轮是事件征询，发给专家的征询表格只提出预测目标，由专家提出应预测的事件。专家回函后，组织者经筛选、分类，归纳和整理，用精确的技术语言制订事件一览表，作为第二轮征询表发给专家。第二轮是事件评估，专家对第二轮表格中的每个事件作出评价。评价的主要内容包括时间预测、方案择优、决策评价和资金分配比例等。第二轮征询表收回后立即进行统计，求出专家总体意见的概率分布，连同第三轮征询表一起反馈给专家。第三轮是轮间信息反馈与再征询，专家填表时，可根据总体意见的倾向（以均值表示）和分散程度（以方差表示）重新评估，可修正自己前一轮的评估意见。一般经过三轮的征询和信息反馈，专家的意见逐次收敛。第四轮一般将落选的意见（即少数人的意见）整理成征询表反馈给专家，再给予一次评估的机会，有时会出现落选的意见重新获得多数赞同的现象。四轮征询结束后，将收敛程度高的意见用数理统计的方法整理出来，写出预测报告。

### （二）派生德尔菲法的简化过程

在实际进行预测时，常采用派生德尔菲法。

1. 第一轮直接进行事件评估，由组织者根据已经掌握的资料，直接拟定事件一览表，加速征询进程。

2. 同时提供背景材料和数据，以减轻专家负担。

3. 若专家意见收敛较好，进行一轮或二轮征询即可停止。

在某种情况下采用部分取消匿名和部分取消反馈，有时有利于加快意见收敛的进程。

### （三）评估结果的数据处理

德尔菲法的一项主要工作是在每轮征询之后的数据分析和处理，将定性评估结果进行量化。最常用的量化方法是将各种评估意见用打分法转为分值，然后再求出各种评估意见的概率分布。概率最大的事件表示是最有可能发生的事件。在概率分布中，用均值表示事件的概率，用方差表示不同意见的分散程度。根据均值和方差就可掌握总体意见的趋向和分散程度。预测结果处理方法很多，仅介绍几种。

1. 对事件完成时间预测结果的处理：一般用四分位图表示处理结果，将专家评估值在时间轴上顺序排列，找出中分位点，它对应的年份为中位数，表示专家中有一半人估计的时间早于它，另一半人估计的时间晚于它，或者理解为有 50% 的专家认为该年份能实现此事件。先于中分位点的四分点为下四分点，后于中分位点的四分点为上四分点。

预测学家根据大量统计，得出一个根据中位数推算上下四分点的经验公式。即如果中位数年份距组织预测的年份为  $X$  年，则下四分点距组织预测年份为  $\frac{2}{3}X$  年，上四分点为  $\frac{5}{3}X$  年。例如，1964 年进行了一项预测，测得中位数为 2000 年，则  $X=2000-1964=36$  年，那么，下四分点为  $1964+\frac{2}{3}X=1988$  年，上四分点为  $1964+\frac{5}{3}X=2024$  年。预测结果常常标明中位数和上、下四分位点，如上例为 2000 年（1988 ~ 2024 年）。上、下四分点之间的区间表示专家意见的分散程序，又称 50% 置信区间。

2. 专家意见的集中程度：它可以由评分平均值、评分比重、最高评分频度和平均名次指标等几个侧面加以描述。

（1）评分平均值：

$$M_j = \frac{1}{m_j} \sum_{i=1}^{m_j} C_{ij}$$

$M_j$  为  $j$  方案的评分平均值；

$m_j$  为参加  $j$  方案评价的专家数；

$C_{ij}$  为  $i$  专家对  $j$  方案的评价值；

（2）评分比重：某方案所得到全体专家的评分在全体专家给全部方案的评分总和中所占的比重，为该方案的评分比重。

$$K_{ji} = \frac{\sum_{j=1}^{m_j} C_{ij}}{\sum_{j=1}^n \sum_{i=1}^{m_j} C_{ij}}$$

式中：

$K_{ij}$  为方案  $j$  的评分比重；

$n$  为方案数。

（3）最高评分频度：给某方案最高分数的专家，在给该方案评分的全

体专家中所占的比重，为该方案的最高评分频度。

$$K_{\max j} = \frac{m^{\max j}}{m_j}$$

式中：

$K_{\max j}$  为  $j$  方案的最高评分频度；

$m^{\max j}$  为对  $j$  方案给最高分数的专家数；

$m_j$  为对  $j$  方案给分的全体专家数。

(4) 平均名次指标：计算这种指标时，首先由应邀专家分别给各方案打分，再将全部方案分别就各个方案的评分排队，列出名次，这样每个方案就得到与专家数相同的名次数，然后用专家总数除名次的数值总和，其商便是平均名次。数值越低，表明名次靠前，因此评估质量越好，最低者最佳。

$$R_j = \frac{\sum_{i=1}^{m_j} R_{ij}}{m_j}$$

式中：

$R_j$  为某方案  $j$  的平均名次；

$R_{ij}$  为专家  $i$  给某方案  $j$  评分排出的名次；

$m_j$  为给  $j$  方案评分的专家总数。

3. 专家意见的协调程度：这是一项重要指标，它反映专家意见的收敛情况。主要用变异系数表示，变异系数是代表评价相对波动大小的重要指标，具体求法按如下顺序进行：

(1) 计算  $j$  方案评估的均方差：均方差代表评价的离散程序，其公式为：

$$D_j = \frac{1}{m_j} \sum_{i=1}^{m_j} (C_{ij} - M_j)^2$$

式中：

$D_j$  为方案  $j$  的均方差；

$C_{ij}$  为  $i$  专家对  $j$  方案的评价值；

$M_j$  为  $j$  方案的评分平均值；

$m_j$  为全体专家数。

(2) 计算  $j$  方案的标准差：标准差代表评价的变异程度，其公式为：

$$\sigma_j = \sqrt{\frac{1}{m_j} \sum_{i=1}^{m_j} (C_{ij} - M_j)^2}$$

式中：

$\sigma_j$  为  $j$  方案的准标差；

其余量同前。

(3) 计算  $j$  方案评价的变异系数：将计算出的  $j$  方案的评分平均值和标准差代入下面公式，即可求出变异系数。

式中： $V_j$  为  $j$  方案专家们意见的协调程度， $V_j$  值越小，专家们的协调程度越高，专家们的意见越收敛，表征德尔菲法征询反馈过程已近完成。可以将预测结果整理出来，提供给信息用户。

除上述指标外，还可以计算协调系数、专家权威程度，寻求高度协调专家群和寻找持异端意见的专家等。在预测中究竟计算哪些指标，根据用户需求和预测需要来决定。

#### （四）德尔菲法的优势与缺陷

##### 1. 德尔菲法的优势：

（1）专家对问题的回答有一定的时间准备，能使回答比较成熟，并可以集各种专家之专长；

（2）在征询意见的几轮反复中，专家能了解不同的意见，而经过不同的分析后提出的看法较为完善；

（3）征询过程中用匿名方式进行，有利于各位专家敞开思想，独立思考，不为少数权威意见所左右；

（4）对专家意见的汇总整理，采用数理统计方法，使定性的调查有了定量的说明，所得结论更为科学。

##### 2. 德尔菲法的缺陷：

（1）预测结果取决于专家对预测对象的主观看法，受专家的学识、评价尺度、生理状态及兴趣程度等主观因素的制约；

（2）专家在日常工作中一般专业方向比较明确，容易在有限范围内进行习惯思维，往往不具备了解预测问题全局所必需的思想方法。

（3）专家对问题的评价通常建立在直观的基础上，缺乏严格的考证，因此专家的预测结论往往是不稳定的；

（4）专家对发展的趋势的预测用直观外推方法，对大大超于现实的思想是难以估计的。

德尔菲法作为一种重要的预测方法，在实际应用中还将不断完善和发展。

## 第七章 信息研究中的科学思维方法

思维是人脑对客观事物的一种间接的和概括的反映过程。科学研究是一种探索性和创造性的脑力劳动，须臾也离不开科学的思维。科学思维的能力和水平，是一个国家科技事业发达的重要标志，也是科技人员素质的重要量度。正如恩格斯所说：“一个民族想要站在科学的最高峰，就一刻也不能没有理论思维。”

信息研究的过程，本质上就是一个思维过程，是人脑对信息的获取和加工，是根据社会需要，在一定的心理结构和操作下，对贮存的和外来的信息进行鉴别和筛选，重新联结和组合的过程。信息研究中常用的思维方法有比较与分类法、归纳与演绎法、分析与综合法、相关分析方法等。

### 第一节 比较与分类法

一般来说，人们认识事物总是从区分事物开始的，要区分就得比较和分类。要对各种信息进行区分，进而从事各种逻辑加工，也应从比较和分类做起。因此，比较和分类是信息研究中常用的思维方法。两者关系密切，比较是分类的基础；分类是比较的结果。

#### 一、比较方法

##### （一）比较方法的概念及其类型

比较法是确定对象之间的差异和共同点的逻辑方法。在科学研究中，运用比较方法的目的在于依据一定的标准，找出表面上差异很大的事物之间所存在的本质上的共同点，或者在表面上极为相似的事物之间所存在的本质上的差异点，以加深对事物的认识。运用比较法，可以对所搜集的信息资料按照研究的需要进行定性的鉴别；也可以对观察或实验结果进行定性或定量的分析；还可以对理论研究的结果与观察、实验的事实之间是否一致作出明确的判断；也能追溯事物发展的历史渊源和明确事物发展的历史顺序。因此，比较法是科学研究中不可缺少的一种逻辑方法，而且是分类、类比分析与综合等逻辑方法的基础。比较法得以运用的客观依据在于事物发展的普遍联系性，在于事物发展变化的共同性与差异性的同时存在。

比较法的类型，依据对象具有同一性和差异性，可分为同类比较法和异类比较法；依据对象的历史发展和相互联系，可分为异期纵向比较法和同期横向比较法；依据对象的整体性与局部性，可分为客观比较法和微观比较法等。在实际研究工作中，上述各类比较法常常交叉使用或综合使用。

运用比较方法必须遵循下列原则：可比性；在同一关系下比较；有明确的比较标准。

##### （二）比较方法的应用

1. 可对客观事物作定量的和定性的分析。定量分析，一是从剖析事物的个别属性开始，从中确定其主要属性和次要属性及其参数。属性分析得越透彻，通过对比得出的结果就越精确。二是从分析各项技术经济指标开始，对某一指标进行时空上程度上或构成上的比较。定量比较是一种数据性比较，分差额比较和相对比较两种。它们均以绝对数为基础。差额比较用减法进行时空上绝对数的比较，相减后还是绝对数，看不出程度和效益。相对比较是用除法进行两个绝对数的比较，可说明事物间的关系、发展程度和效益。

2.可揭示不易发现的事物本质上的异同。为揭示事物本质上的异同，首先要透过现象分析事物的本质，然后再比较本质上的差异。例如为掌握世界上各主要国家自然条件上的异同，应先分析构成自然条件的主要因素，如各地区的海拔高度、历年来的温度和降水量等，然后对这些因素进行时间上的数据运算，求得某一阶段内的平均值，或画出平均典型，然后再通过各个国家的上述因素的平均值进行空间上的比较，就可从复杂、分散的数据中，找出各个国家自然条件的异同点及其变化规律。

3.可进行时空水平比较。比较不同国家、地区、部门的经济力量、生产能力、科技水平以及自然资源的特点，总结经验教训。比较各国学科、专业、行业、企业的发展历史、现状，了解其发展速度和顺序变化，掌握其薄弱环节。比较各企业的工艺流程、设备性能、生产效果，了解其水平差距和对策。比较以产品为中心的八大技术经济指标和市场需求，确立企业发展方向。比较科研、生产和工程项目的多种方案，寻求最优方案。

### （三）应用比较方法注意事项

1.注意对象的可比性。对象的可比性通常包括时间、空间和范畴的可比性。时间的可比性是指所比较的数字和情况必须是同一时期的，有时甚至要按周或者按日计算；空间的可比性是指比较时要充分考虑到国家和地域上的差别；范畴的可比性是指相比较的事物必须属于同一层次，具有相同的内涵和外延，例如对于全员劳动生产率，国内与国外的理解就不完全相同。

2.决定事物的性质和事物发展方向的因素是多方面的，进行比较必须考虑局部和整体目前和长远的关系，抓住主要矛盾进行比较。

3.事物是由多种因素决定的，比较时除了抓住主要矛盾以外，还必须注意多项指标的比较，才能避免认识的局限性。例如要从国外引进一台隧道掘进机，不仅要各国各种型号机器的功率和价格进行比较，还必须对它们在不同地质的掘进速度、移动灵活性、零部件更换频率、燃料消耗等项指标逐一比较。

4.比较是分析、判断和推理的基础。比较时通常是用数据进行比较，但是，同一个数据可能表征着两种截然相反的情况，不同的人对同一数据也可能作出完全不同的判断。因此，在信息研究中要注意分析数据，透过现象看本质，防止认识表面化。例如，甲乙两厂机床的数量一样多，在一段时间内，甲厂比乙厂生产的产品多就不一定是好事。也许甲厂忽视设备的维修和保养，片面追求机器运转率，长此下去，甲厂的机床会很快磨损，最终将导致机床寿命的缩短。

## 二、分类法

### （一）分类法的概念及其类型

世界上一切事物都可以按其属性区分开来，并归入一定的门类。这种按属性异同将事物区别为不同种类的思维方法叫做分类法。类是具有共同特征的集合。

分类法是以比较法为基础的。人们通过比较，揭示事物之间的共同点和差异点，然后根据共同点将事物归并为较大的类，再根据差异点将较大的类划分为较小的类。正如陈世骧同志所言：“万物虽众，但均有共有别。推而共之，可以类上归类，推而别之，可以类下分类。共性是归合事物的根据，特性是区别事物的根据，共性与特性是对立统一，是一切分类的根据。”这样分类的结果就将事物区别为具有一定从属关系的不同层次的大小类别，形



成各种概念系统，反映客观世界中事物的区别和联系。

分类法作为一种认识事物、区别事物的方法，在世界上已有几千年的历史，并已发展成为一门专门的学科。分类法和主题法是在长期信息研究工作中形成与发展起来的文献整理和检索的工具与基础，也是信息研究的重要方法。根据分类法的目的和标准，通常分类法的类型有现象分类法，即根据事物的外部标志或外在联系进行分类，如按作者和厂商分类；本质分类法，即根据事物的本质特征或内在联系进行的分类，如按学科和内容分类；实用分类法，即根据应用需要进行的分类，如按字顺和笔划分类等。

## （二）分类法的应用

分类法是科学认识和科学研究的起点和基础，同样也是信息研究的重要方法。分类法在信息研究工作中的作用有：

1. 信息整序的主要手段。思维的过程是一个信息加工过程。而对获得的大量信息的加工处理，则从整序着手。恩格斯说：“无数杂乱的认识资料得到清理，它们有了头绪，有了分类，彼此间有了因果联系；知识变成了科学……”因此分类法是整理和储存情况、事实、数据，使之条理化系统化，并为分门别类深入研究创造条件的主要手段，否则就会千头万绪，如堕烟海。

2. 认识事物的基本依据。俗话说“物以类聚，人以群分”。分类是认识事物和研究事物的依据和起点。千变万化的事物和成千上万的数据，不形成种或类的概念，不进行系统分类是不可想象的。正如恩格斯所说：“没有种的概念，整个科学就没有了。科学的一切部门都必需以种的概念为基础……”。大量纷繁庞杂的事实材料一经分类，往往会显露出某些规律性的线索和端倪，有助于概括和发现。

科学史上，元素周期律的发现就得益于比较和分类的方法。门捷列夫以元素的原子量为分类标准，按其递增顺序排列，发现元素的化学性质随着原子量递增存在着周期循环的现象，从而发现了周期律。化学元素周期表就是分类思想的杰作。

3. 信息研究的重要基础。分类法是各门学科研究的一种基本方法，因为科学要从堆积事实的经验阶段走向整理事实的理论阶段，就不能不借助分类法来架设中间的桥梁。在信息研究过程中，课题内容的划分、问题的思考、考察结果的表述以及信息研究的许多方法，如归纳与演绎、分析和综合等的运用都是以分类和划分作为重要基础的。例如层次分析法的核心就是把问题层次化，根据问题的性质和要达到的总目标，把同一层次的问题分解为不同的要素，形成一个多层次的分析结构模型。在这里不论层次的划分还是问题的分解，都是与分类法密不可分的。

## （三）分类法的原则

不论是现象分类还是本质分类，都须遵循以下规则：

1. 每一次分类必须按照同一个标准进行，否则就会出现分类重迭或分类混乱的逻辑错误；

2. 分类必须相对应，即划分所得的各事项应与被划母项相应，否则会出现分类过窄或过宽的逻辑错误；

3. 分类必须按照一定的层次逐级进行，否则会出现越级划分的逻辑错误。

## （四）分类法的缺点

分类法又叫体系法，由于分类体系是按单线展开的，每次分类只能有一

个标准，所以会使具有多种属性的事物无法得到充分的反映。也就是说，许多事物（如图书），可以按某一属性而集中，也可以根据其它属性而分散，这就产生了集中与分散的矛盾，对于认识事物是不利的。

由于事物有多种属性，要想全面地认识和揭示事物。就必须建立许多标准，从不同侧面划分出许多类型。比如，关于信息用户的类型，我们可以从用户所属学科或领域划分，可以按用户有没有利用某一信息系统提供的信息服务划分，可以根据用户对信息的使用情况划分，可以按用户的技术职务划分，也可以按对用户信息保证的级别划分，等等，还可以有许多不同的标准，而这些标准，从不同角度、侧面反映了信息用户的各种特征，不可能只有一个统一的标准。同时，由于信息用户特征的多样性，人们对信息用户特征的认识的局限性，所设立的标准不可能周全，也不可能过多。因此，分类法的局限性也就明显地表现出来了。

## 第二节 归纳与演绎法

马克思主义认识论认为，一切科学研究都必须遵循两条途径：由认识个别到认识一般，再由认识一般到认识个别。这就是归纳和演绎的过程。它们是统一认识过程中两个既互相对立，又互相依存的思维方法。两者的关系是，归纳是演绎的基础，演绎是归纳的前导。就像恩格斯所说：“归纳和演绎，正如分析和综合一样，是必然相互联系着的。不应当牺牲一个而把另一个捧到天上去，应当把每一个都用到该用的地方，而要做到这一点，就只有注意它们的相互联系、它们的相互补充。”

### 一、归纳法

#### （一）归纳法的概念及其分类

归纳法是从个别中发现一般的思维方法和推理形式，即从个别事实中概括出一般原理。如我国古代农学家贾思勰广泛搜集和总结了古代农作物的耕作经验，总结出“早熟者，苗短而收多，晚熟者，苗长而收少”的良种培育规律，即现在的矮秆高产，采用的就是归纳法。

在科学研究中，归纳法是一种得到普遍应用的方法。归纳法的实质就是从个别到一般、从特殊到普遍、从具体到抽象的分析推理过程。它不是对复杂、众多素材的牵强附会的编排，也不是一大堆事实的机械的分类罗列，而是要从现象的综合归纳中找出事物之间的内在联系，发现事物的共性和本质规律。这是符合辩证唯物论的认识论的。因为人们的认识总是从接触和认识个别、特殊的事例开始，然后在大量同类事物的认识中，把对个别事物的认识上升到对事物共同本质的认识。因此，可以认为归纳法是分析和认识事物的一种基础方法。

通常认为，归纳法可以分为完全归纳法和不完全归纳法两种。

1. 完全归纳法：完全归纳法是根据某类事物中每一事物都具有某种属性，推出该类全部事物都具有该属性的归纳推理。

因为完全归纳法是考察了某类事物的全部对象，发现它们具有某种属性之后才作出的概括，所以得出的一般结论确实可靠，是一种必然性推理。但它要求完全枚举出某类事物中的所有个体，因此只能用于数目有限，不能用于数目无限的类别上。数学上的穷举法、数学归纳法、四色问题的证明可以看作是完全归纳法的应用。

2. 不完全归纳法：不完全归纳法是根据某类事物的部分对象具有某种属性，而作出该类事物都具有某种属性的一般性结论的归纳推理。不完全归纳法突破了完全归纳法的局限性，对人类认识范围的扩大具有重要意义。

由于完全归纳法的缺陷，在实际应用中大都采用不完全归纳法。

不完全归纳法在实际应用中又可分为简单枚举法和科学归纳法两种。

### （二）归纳法的应用

不论在人们的日常生活中，还是在科学研究活动中，归纳法都是一种重要方法，在信息学研究中也常常使用归纳法。其作用有以下几点：

1. 帮助人们从大量经验事实中找出普遍的特征。在探求自然的普遍真理的科学史上，许多经验定律和公式都是应用归纳方法取得的。由于信息学是一个新兴的发展中的学科，处于积累经验阶段，所以归纳法就更显得十分重要了。信息学研究者需要从大量调查、观察、实验所得的材料中总结出一般概念、理论，建立自己的学科体系。信息学中，许多定义的确立都是通过归纳概括，人为确定的，如什么是信息，什么是信息工作，什么是信息检索等，都是在考察了一定的客体之后定性地规定出它们的定义。因为各人观察的角度不同，表述方式不同，所以某一事物的定义往往有几种、数十种，如关于信息的定义就多达几十种。

2. 使人们在考察个别事物时，受到启发，接近真理。科学研究中的许多猜想是通过归纳法提出的，虽然猜想还不是真理，但毕竟向真理迈出了一步。例如在克兰菲尔德实验中，当所查的文献要求的概念全部满足时，查全率为 65%；当所要求的概念减少一个即可认可时，查全率为 85%；当只有一个单元，即可认可时，查全率为 97%。他们从这些事实归纳中得出一定的规律，然后进一步实验，最后得出一个定律：“当查全指数提高时，相关率就必须下降，相反，当查全指数下降时，而相关率就会改善。”这就是后来人们所说的“查全查准原则。”

### （三）归纳法的局限性

归纳法是信息研究中经常应用的一种方法。但它也有局限性，主要表现在归纳都是对已经发现或出现的事实而言的，它不能预见事物发展过程中的复杂情况。恩格斯指出“归纳法确认鱼类是一种终身专门用鳃呼吸的脊椎动物。可是出现了一些动物，这些动物的鱼的特征差不多是大家公认的，但是它们除去鳃还有很发达的肺；我们看得出来：每一条鱼在鳔中都有潜在的肺。”因此，可以说应用归纳法所得出的结论被新的事实所推翻的情况，在科学史上是屡见不鲜的。所以，我们在运用归纳法分析信息资料时，应该充分注意到这一点，尽量能同其他分析方法结合运用，以避免结论的局限性。

## 二、演绎法

### （一）演绎法的概念及其分类

演绎是在一般中发现个别的思维方法和推理形式，即用已知的一般原理考察某一特殊对象，推演出有关这个对象的结论。如从牛顿 3 定律出发，推演出当火炮的仰角为  $45^\circ$  时射程最远，并得出了射程与炮管倾角之间变化规律的一般公式，采用的就是演绎法。演绎逻辑奠基于亚里士多德，他提出了三段论法的一般规律。欧几里德几何学则从 5 个公理和 5 个公设出发，构成了一个逻辑上完美严密的体系，使亚里士多德关于科学是演绎系统的理想得到了实现。后来法国哲学家和物理学家 R. 笛卡儿创立了以数学为基础、以演绎法为核心的方法论，把演绎方法看作是科学发现的基本逻辑，并与培根的

归纳主义相对立，成为近代演绎主义的代表。接着德国数学家 G.W. 莱布尼茨则把数学与形式逻辑结合起来首创了数理逻辑，用形式化的数学方法研究问题，把演绎推理发展到一个新的阶段。

同归纳法一样，客观世界存在的一般与个别、普通与特殊的关系也是演绎法的客观依据。但与归纳法不同，演绎法是前提与结论之间有着蕴涵关系的推理。正如恩格斯所说：“如果有正确的前提，并且把思维规律正确地运用于这些前提，那么结果必定与现实相符”。因此，演绎法具有不可抗拒的逻辑力量，是人类认识客观事物的一种重要方法。

从逻辑学上讲，演绎推理可分为直言三段论、假言三段论、选言三段论、联言推理和关系推理。但从科学研究角度讲，人们还可以将演绎推理分为公理演绎法、假说演绎法、定律演绎法和理论演绎法等。

### （二）演绎法的应用

1. 它是逻辑证明的工具，也是一个学科理论体系形成的重要方法。信息学理论体系的形成离不开演绎推理这一重要方法，同时，演绎法也为信息学研究中某些具体理论观点的形成提供了有力的逻辑推理工具。

近几年有关信息理论的讨论中，关于“信息定义”的制定是极其混乱的，有人声称有 60 余种提法。但是在中国和日本都出现一种反传统的观点：认为“信息定义”是什么的问题提法欠妥。持这种观点的人认为，提出和企图回答本质主义问题往往是学科不成熟时期的特征，目前信息学还是一个不成熟的学科。“信息定义是什么”是一个本质主义问题，出现这一问题的讨论是可以理解的，但是要求得出本质主义的答案，并作为第一性的问题来讨论，目前没有必要，也不可能得出准确的结论。在论述这一问题时，运用的就是演绎法的逻辑力量。

2. 它是科学预见的一种手段。科学预见就是将一般原理用于具体情况正确推论，信息学要从经验科学进入理性科学阶段，一个很重要的方面就是要能够运用演绎推理的方式，而不仅仅是用经验的方式得出一些新的理论观点，也就是要做出一些科学预见。

美国图书馆学家谢拉曾根据物质产品的生产、分配和利用，推论智力产品的生产、分配和利用，并取得一定的成果，称之为“社会认识论。”后来，英国信息学家布鲁克斯发展了谢拉演绎的推论的结果，绘制出知识结构“地图”，从知识“地图”上可以看到智力产品形成的知识系统及其动态结构，为智力产品的运动找到了一种表达方式。

3. 它是检验和发展假说与理论的一个必要环节。对科学假说和理论的检验，首先需要从假说和理论所论述的一般原理，推演出一个可以与实验相对比的具体结构，指导实验的设计和进行。这里所运用的就是演绎方法，正是这种方法为假说和理论的发展提供了必要的条件。

### （三）演绎法的缺陷

演绎是以公理、定理、假说等为前提的，而这些理论是通过别的方法获得的。如果前提中有一个是不正确的，就不可能得出正确的结论。所以单靠演绎方法是不行的。同时，演绎法是从一般到个别、从共性到个性的思维方法，但共性不能完全包含个性，一般也不能包括全部个别，所以演绎方法不可能揭示个别的多方面的属性。

## 第三节分析与综合法

分析和综合是揭示个别和一般、现象和本质的内在联系的思维方法，是科学抽象的主要手段，是各种认识活动的基础。人们认识客观事物，必须首先解剖事物的各个方面、各个部分，弄清各个部分的结构、功能、性质及其在整体中的作用，然后再从整体考虑，找出各个部分之间的联系和统一，从而作到从总体上把握事物的本质和规律。

分析和综合并不是彼此割裂的，整个认识过程是分析和综合的统一。恩格斯说：“思维既把相互联系的要素联合为一个统一体，同样也把意识的对象分解为它们的要素。没有分析就没有综合。”

## 一、分析法

### （一）分析法的概念及其分类

分析是把整体分解为部分，把复杂的事物分解为简单要素分别加以研究的一种思维方法。在客观事物中，组成整体的各个部分本来是相互联结的，为了分析这些部分或方面，就必须把它们暂时割裂开来，把被考察的因素从整体中抽取出来，暂时孤立起来，以便让它单独地起作用。在研究工作中，所采用的单因素分析法或单因子实验法都是为了深入事物的内部，研究它们的细节，为从总体上把握事物积累材料。常用的分析法包括定向分析、因果分析和层次分析等。

1.定向分析：根据时间和空间的不同又可分为历史分析和地区分析两种。

（1）历史分析，根据某一事物的发展历史和发展过程，分析各历史阶段内出现的因果关系或矛盾关系以及这些因素和关系的性质，在整个课题发展过程中所占的地位、所起的作用，排除其中一般的和次要的因果关系或矛盾关系，抓住其中特殊的和主要的因果关系和矛盾关系，就能总结出成功的经验或失败的教训。

（2）地区分析。是对不同国家或地区反映的同一课题进行对比分析，排除其中无差距或差距小的问题，抓住差距大的问题，即抓住了主要矛盾。

2.因果分析：又分为内因分析、条件分析和典型分析三种。

（1）内因分析。总结科技发展成败的经验教训时，对于发展历史不长，或无其他类似地区以资比较的新兴技术，可以从分析内因（包括科学理论、人员、技术、产品、原材料、设备等）着手。

（2）条件分析。对影响科技生产发展的诸种条件进行分析，掌握各种条件的影响和相互关系，并对比国内情况，为国内是否发展或引进该项技术提供决策依据。

（3）典型分析。是从反映同一事件，但时间、空间和性质上有差异的众多素材中，选择一种或几种对本国具有参考价值的事例，深入分析其中的核心问题，对事件发生的地点、时间、过程和后果影响进行分析，从中获得经验或教训。典型分析并不局限于事件。它可是一家工厂的创建和发展，一个产品的研制政策的制定和修正等。

3.层次分析 根据辩证唯物主义观点，任何事物都是可分的，可分就有层次，每个层次都有自己的矛盾，深入剖析事物的层次，抓住各层次的主要矛盾和矛盾着的双方的相互转化，以及主要矛盾和次要矛盾的相互转化，就能促使事物不断发展。层次剖析越透彻，就越接近于抓住事物的本质。

### （二）分析法的应用

1. 总结成败的经验教训。在科学技术发展过程中，经常会碰到这种情况，当矛盾得到解决，就会促进科学技术的发展，获得成功、兴旺、增产以及质量和效率的提高；当矛盾日趋激化，就会制约科学技术的发展，导致失败、衰退、减产以及质量和效率的下降等。分析法就是依据事物间的这种促进作用或制约作用，研究因果关系或矛盾关系，分析影响某项科学技术发展的主要因素，总结其成败的经验教训，进而结合国内或本地区的实际情况，提供借鉴。

2. 研究科学、技术和产品行业的兴衰背景、发展趋势、途径和方法。科学技术和产品行业的发展，总是受控于内外因素的影响，因此，分析这些内外因素及其相互关系就能把握住科学发展的总规律和总趋势。

3. 分析引进技术和适用性。引进技术是否适用于国内实际情况，如技术力量、资源条件、经济水平、社会需要，取决于对国内外条件因素的对比分析，并做出判断，在这里，分析法的应用起着主要作用。

4. 研究科研生产中的政策管理问题。可以分析影响科技效率的六大有关因素，包括人员、经费、工艺技术、政策管理、实验手段和信息资料等；分析影响生产效率的八大技术经济指标；分析制定政策的背景和依据；分析影响产品质量的五大因素，包括人、设备、工艺材料和测试方法等；分析影响产品的销售，造成产销脱节，供过于求或供不应求等情况的社会需要、经济水平、市场价格、产品质量、品种构成等主要因素。

### （三）分析法实施的基本步骤

1. 分解事物具有多样性的各个部分；
2. 考察各个部分的特殊结构和本质；
3. 研究多样性的各个方面的地位作用和相互之间的联系。

## 二、综合法

### （一）综合法的概念及其分类

综合法是把对象的各个部分、属性、要素联系起来，从总体上进行考察和研究的一种思维方法。综合法是在分析的基础上进行的科学概括，并把对各个部分、各种属性、各种要素及其相互关系的认识总括统一为对事物整体的认识，从而达到对事物的全貌及其本质和规律的把握。

综合法不是事物的各个部分、各种属性、各个要素的简单地机械相加，而是在思维中把事物的各个方面的内在联系统一成一个有机整体。认识事物必须从分析到综合。只有分析而没有综合，就不可能达到对事物的本质和规律的全面认识。

综合法可分为以下三类：

1. 简单综合：简单综合是对某些对象进行研究，发现其中相同之处或不同之处，然后进行综合归纳的一种逻辑思维方法。有两种类型：一种是从部分到整体的综合，目的在于使分散的知识集中化、系统化，开辟知识的新的应用途径；一种是从整体到部分再到整体，目的是继承已有知识，创造新的知识体系、技术体系和产品。

2. 提炼综合：是把个别的、分散的、局部的情况结合在一起，从中提炼出能产生同一现象的共性原因或特殊原因，从而提出新的认识、新的概念或新的结论的一种逻辑思维方法。

3. 系统综合：是从系统的观点出发，对专题进行大范围的综合研究。先提出要点，然后把内容融合进去，使课题从纵的方面融合，包括有关历史、

现状和未来趋势，做到知识和技术乃至产品的继承；从横的方面不仅要融合国内外情况，而且要融合有关科研生产情况及其与相邻学科和有关经济部分的关系，而且还要考虑方针政策、技术经济、自然、资源、环境和社会等外部因素。最终得以了解知识技术、产品行业和市场需要间的关系，掌握科技生产发展的规律。

### （二）综合法的应用

1. 在比较和分析的基础上使分散、片面、内容各异的有关观点、知识、技术、产品集中化、系列化；使不同时间出现的、逐步深化的知识连续化、深刻化；使来自各个方面的知识（科学技术、生产实践以及外部条件因素）综合化、精炼化。它是通过对信息进行更高层次的概括、加工，产生认识上的质的飞跃。

2. 掌握多门类科学技术，创造新的科学技术体系。通过综合研究能掌握多种科学技术，把它们应用于实践必将获得新的成就。例如电子技术、激光技术、计算机技术等都是已有的技术，如能通过信息研究把它们与农业相结合，就能通过科学研究创造出新的农业科技体系。又如电子显微镜、超声波显示、X线造影、同位素扫描、放射自显影等结合应用必然促进分子生物学发展。信息研究本身是自然科学与社会科学的综合，通过这种综合必将形成新的科学体系。

3. 掌握科技发展的规律，为决策者提供依据。由于综合研究涉及面广，能充分掌握某项科技的发展规律，因此能为政策制定者提供科学生产发展现状、历史和教训方面的信息，为计划制定者和产品项目决策者提供依据。

4. 促进交叉学科、边缘学科的发展和大学科、大系统的形成。

### （三）综合法实施的基本步骤

1. 把握事物被分析出来的各个方面；
2. 确定各个方面的有机联系和结构形式；
3. 揭示事物总体的本质和规律，从多样性的统一上再现事物的整体。

### 三、分析法与综合法的辩证关系

分析法与综合法是认识过程中相辅相成、辩证统一的两种方法。

分析法和综合法的辩证统一首先表现在分析法和综合法的相互依存、相互渗透中。综合法必须以分析法为基础，没有分析，认识不能深入，对总体的认识就只能是抽象的、空洞的。只有分析而没有综合，认识就可能囿于枝节之见，不能统观全局。事实上任何分析总要从某种整体性出发，不能离开以整体性认识为指导，否则分析就会有很大的盲目性。

分析法和综合法的统一还表现在它们的相互转比上。人的认识是一个由现象到本质、由一级本质到二级本质不断深化的过程。在这个过程中，从现象到本质、从具体到抽象的飞跃是以分析法为主的；一旦达到了对事物的本质的认识，就要用这个本质说明原有的现象，这就是提出假说、建立理论（或模型）的过程，这个过程是以综合法为主。随着认识的推移，当新的事实与原有的理论发生矛盾时，认识又可能在新的层次上转入分析法。人们的认识就是在这种分析—综合—再分析—再综合的过程中不断前进的。

## 第四节 相关分析方法

所谓事物之间的相关是指事物的发展变化会引起另一事物的发展变

化，是事物之间相互关系的一种变化过程。因此，相关分析法就是在掌握一定的数据和事物之间相关性的基础上，通过特定的一些相关关系的定性或定量分析进行逻辑推理的一种信息研究方法。

按事物之间的联系方式，相关分析方法可分为因果相关、伴随相关、并列相关、包容相关等。因果相关是利用已知事物和未知事物具有因果关系研究事物的方法，如根据复合装甲出现的事实，推断穿甲弹将比破甲弹会有更大的发展；根据核威胁的理论，预计常规武器的地位和作用将显著增加等。伴随相关是利用已知事物和未知事物相伴出现的特点研究事物的方法，如根据乡镇企业的发展，注意对环境保护的研究；根据美苏缓和的局势，警惕周边战争的爆发等。由于间接、迂回是相关分析法的一个重要特点。因而，该法特别适用于军事技术、专利技术及其他难得技术的战术研究与战略分析。具体相关分析方法的特点和作用有以下几点。

### 一、相关分析法的特点

1. 研究方法间接、迂回。在利用相关分析法研究问题时，人们的研究目标、所要解释的问题是甲事物，但直接面对的却是乙事物，然后通过甲乙事物之间的联系，从乙事物出发，采用迂回的间接的方式去接近甲事物，最后认识甲事物。

2. 强烈依靠研究者的经验。彼此之间相互联系是事物的普遍特点之一，如何将恰好是相关关系的因素联想到一起，这在很大程度上要依靠研究者的经验。对于各种相关事物的细心观察、大量积累，并从中总结出规律性的东西，是成功地应用相关分析法的重要条件。

### 二、相关分析法的作用

由于事物之间的联系是错综复杂的，因果相关与伴随相关交互混复，有时甚至与非相关事物偶合，造成某种假象，不易分辨。因此，使用相关分析法的关键在于调研人员具备足够的有关科技知识和进行深入的调查研究，才能进行由此及彼、由表及里和去伪存真的分析。某果汁厂进行果汁澄清问题的调研，开始由于信息人员缺乏技术知识，加上果汁加工工序较多，影响质量的因素也很多，一时难以抓住主要矛盾。后来信息人员与技术人员一起进行分析研究，根据混浊原因是出现酒石酸沉淀，最后才判断是巴氏杀菌工序温度控制不好所致。

利用因果联系研究事物，既可以根据已知之因推测未知之果，也可以根据已知之果反推未知之因。例如中国山西一个偏远的山村中，10多年来出生的婴儿都是女性，而且成年妇女个个患有头痛、骨痛的怪病。这些奇特的现象引起了医学界的注意，经过专家们的调查，原因在于这两个村庄的居民饮用了含镉量较高的污水，水源被遗弃的锌矿污染了。科学家们认为，当人体内含镉量高时，精子的成熟和活动能力受到损害；比较起来，含有X染色体的精子抵抗力强，生存率比含有Y染色体的精子要高。这样，含X染色体的精子与卵子结合的机会就多。所以容易生成女胎。在上述实例中，科学家探索“女儿村”奥秘所采用的思路就是相关分析。

由于事物的联系是普遍存在着的，在信息研究中，常常利用相关分析，由已知信息来推知未知信息。60年代初，有关大庆油田的情况是保密的，日本根据油罐车上泥土的颜色和厚度、汽车司机的着装，根据报纸上有关王进喜事迹的报道以及其它资料，准确地判断出大庆油田在齐齐哈尔与哈尔滨之间的松辽平原上。



