

●邱均平 邹菲

关于内容分析法的研究^{*}

摘要 内容分析法产生于20世纪初。其类型有解读式内容分析法、实验式内容分析法和计算机辅助内容分析法。国外从20世纪70年代以来对内容分析法的研究日益重视。我国近些年对其研究也引起重视。其原理是对文献内容所含信息量及其变化进行分析,从而达到对文献内容进行可再现的、有效的推断。它具有基本的方法、步骤及工具。内容分析法具有广泛的应用前景。表6。参考文献12。

关键词 内容分析法 发展历程 方法原理 应用
分类号 G350

ABSTRACT The content analysis methods originated in the early 20th century. Their patterns include hermeneutic content analysis, empirical content analysis and computer-aided content analysis. More and more attention has been paid to the study of content analysis methods in foreign countries since the 1970s and in China recently. In this paper, the authors analyze their principles, methods, steps, tools and application prospects. 6 tabs. 12 refs.

KEY WORDS Content analysis method. Development process. Method and principle. Application.
CLASS NUMBER G350

顾名思义,内容分析法(Content Analysis)是一种对研究对象的内容进行深入分析,透过现象看本质的科学方法。美国传播学家伯纳德·贝雷尔森(Bernard Berelson)首先把它定义为一种客观地、系统地、定量地描述交流的明显内容的研究方法^[1]。在近一百年的发展历程中,内容分析法已经被广泛运用到新闻传播、图书情报、政治军事、社会学、心理学等社会科学各领域中,取得了显著的成效。随着信息时代的到来,内容分析法在计算机、网络技术研究中也成为了一个新热点。本文将对其中的一些问题进行探讨。

1 内容分析法的发展历程

1.1 内容分析法的产生与发展

20世纪初,有人采用一些半定量的方法对文献的内容进行分析和解释。这些研究主要是通过统计报纸上某方面内容的新闻报道篇数,来考察报道的重点以及社会舆论状况,并对艺术、音乐、文学和哲学等方面文献的主题内容进行分析,以期发现社会和文化变化的历史趋势。这种研究方式的出现和应用形成了内容分析法的雏形,因其可操作性强、适用范围广的特点而得以不断发展。第二次世界大战期

间,美国著名传播学家哈罗德·拉斯韦尔等人在美国国会图书馆组织了一项名为“战时通讯研究”的工作,这项工作系统地发展和完善了内容分析法。他们以德国公开出版的报纸为研究对象,通过内容分析法获取重要的军政机密情报,取得了出乎意料的成功。这项工作所采用的方法和所取得的效果,为战后内容分析法的发展和应用奠定了基础。奈斯比特所著《大趋势——改变我们生活的十个新方向》一书的出版,则是内容分析法走向成熟的里程碑。他的咨询公司运用内容分析法对200份美国报纸进行分析、综合,经过几年的积累,归纳出美国从工业社会过渡到信息社会的十大趋势^[2]。在这部书取得成功的同时,众多的研究者开始关注内容分析法在社会研究中的巨大作用和潜力。

1.2 内容分析法的类型演变

内容分析法作为一种研究社会现实的科学方法,经过了不断的理论探讨和实际应用才逐步趋于成熟与完善。在此过程中,通过该方法究竟是定性研究还是定量研究的激烈争论,研究者先后提出了以下几种方法类型,从中可以看到内容分析法的演变轨迹以及技术进步对其发展的推动作用。

^{*} 本文系教育部重大项目(02JAZD870003)的研究成果之一。

(1) 解读式内容分析法 (Hermeneutic content analysis)

解读式内容分析法是一种通过精读、理解并阐释文本内容来传达作者的意图的方法。“解读”的含义不只停留在对事实进行简单解说的层面上,而是从整体和更高的层次上把握文本内容的复杂背景和思想结构,从而发掘文本内容的真正意义。这种高层次的理解不是线性的,而具有循环结构:单项内容只有在整体的背景环境下才能被理解,而对整体内容的理解反过来则是对各个单项内容理解的综合结果。

这种方法强调真实、客观、全面地反映文本内容的本来意义,具有一定的深度,适用于以描述事实为目的的个案研究。但因其解读过程中不可避免的主观性和研究对象的单一性,其分析结果往往被认为是随机的、难以证实的,因而缺乏普遍性。

(2) 实验式内容分析法 (Empirical content analysis)

实验式内容分析主要指定量内容分析和定性内容分析相结合的方法。20 世纪 20 年代末,新闻界首次运用了定量内容分析法,将文本内容划分为特定类目,计算每类内容元素出现频率,描述明显的内容特征。该方法具有三个基本要素,即客观、系统、定量。用来作为计数单元的文本内容可以是单词、符号、主题、句子、段落或其他语法单元,也可以是一个笼统的“项目”或“时空”的概念。这些计数单元在文本中客观存在,其出现频率也是明显可查的,但这并不能保证分析结果的有效性和可靠性。一方面是因为,统计变量的制定和对内容的评价分类仍由分析人员主观判定,难以制定标准,操作难度较大;另一方面计数对象也仅限于文本中明显的内容特征,而不能对潜在含义、写作动机、背景环境、对读者的影响等方面展开来进行推导,这无疑限制了该方法的应用价值。

定性内容分析法主要是对文本中各概念要素之间的联系及组织结构进行描述和推理性分析。举例来说,有一种常用于课本分析的“完型填空式”方法,即将同样的文本提供给不同的读者,或不同的文本提供给同一个人,文本中被删掉了某些词,由受测者进行完型填空。通过这种方法来衡量文本的可读性和读者的理解情况,由于考虑到了各种可能性,其分析结果可以提供一些关于读者理解层次和能力的有用信息。与定量方法直观的数据化不同的是,定性方法强调通过全面深刻的理解和严密的逻辑推理,来传达文本内容。

一般认为,任何一种科研方法都包含一定的定性

步骤。比如,研究开始阶段要确定主题和调查对象,明确相关概念,制定研究计划;最后阶段还要针对研究的问题,解释实验结果。但是单纯的定性方法缺乏必要的客观依据,存在一定主观性和不确定性,说服力有限。因此,很多学者倡导将定性方法和定量方法结合起来,取长补短,相得益彰。定性定量相结合的内容分析法应具备以下几个要点:①对问题有必要的认识基础和理论推导;②客观地选择样本并进行复核;③在整理资料过程中发展一个可靠而有效的分类体系;④定量的分析实验数据,并做出正确的理解^[3]。

(3) 计算机辅助内容分析法 (Computer-aided content analysis)

计算机技术的应用极大地推进了内容分析法的发展。无论是在定性内容分析法中出现的半自动内容分析 (Computer-aided content analysis),还是在定量内容分析法中出现的计算机辅助内容分析 (Computer-assisted content analysis),都只存在术语名称上的差别,而实质上,正是计算机技术将各种定性定量研究方法有效地结合起来,博采众长,使内容分析法取得了迅速推广和飞跃发展。互联网上也已出现了众多内容分析法的专门研究网站,还提供了不少可免费下载的内容分析软件,相关论坛在这方面的讨论也是热火朝天。

2 国外内容分析法研究进展

在诸多发达国家中,内容分析法已成为大众传播研究的主要手段,在其他领域的应用也相当普及,并已取得了较为丰硕的研究成果。本文采用文献统计的方法,对国外相关期刊论文进行了较全面的统计分析,以期掌握国外内容分析法的主要研究进展。

2.1 选取数据资料来源

在选择数据源时,笔者曾对所有可获取的资料来源,包括文摘杂志、网络资源、数据库等进行了一番比较分析,发现网络数据库的查全率和查准率是最令人满意的,因此选取了学术期刊全文库 (Academic Search Premier) 和剑桥科学文摘数据库 (Cambridge Science Abstracts) 作为数据样本来源。这两个网络数据库覆盖面广,收录期刊全,检索途径多,能从中获得较为充分的研究资料,后期复核也证明所选取的资料基本涵盖了关于内容分析法的主要期刊论文。经过严格筛选,笔者搜集到 1970 年至 2001 年间的 100 篇英文期刊论文的全文或文摘信息,并以此为基础展开了进一步的研究。

2.2 国外内容分析方法的研究概况

(1) 论文数量的年代分布情况

关于内容分析法的期刊论文按时间段分布情况如表1所示。

表1 国外内容分析法研究论文的时间分布

年份	1970 - 1980	1981 - 1985	1986 - 1990	1991 - 1995	1996 - 2001
论文数(篇)	2	5	10	33	50

表1清楚地显示出,自20世纪70年代以来,国外研究内容分析方法的论文数量呈明显的上升趋势,各时间段之间相比,近乎是成倍增长,由此可见内容分析法在国外得到了迅速的认同和广泛的应用。

(2) 论文学科分布情况

内容分析法研究论文分布的期刊数量很多,根据学科范围划分的论文分布情况如表2所示。

表2 国外内容分析法研究论文的学科分布

期刊类型	图情学期刊	社会学期刊	新闻传播学期刊	计算机学期刊	医学期刊
载文篇数(篇)	46	27	19	6	2

从表2可以看出,研究内容分析法的学者主要来自图书情报学、社会学、新闻传播学、计算机科学和医学5类学科,论文分布的学科面较为广泛,文、理、医俱全,但各类学者研究的角度和侧重点有所不同。

(3) 论文研究角度分析

综合考察国外内容分析法相关论文,可以根据研究角度的不同,将其分为理论研究和应用研究两大类。其分布情况如表3所示。

表3 国外内容分析法研究论文的主题分布

期刊类型	图情学期刊	社会学期刊	新闻传播学期刊	计算机学期刊	医学期刊	占总数比例(%)
理论研究	9	12	6	2	1	30.0
应用研究	37	15	13	4	1	70.0

表3的数据反映出两个方面的情况:首先,应用类论文占大多数,可见内容分析法的确是一种实践性比较强的科研方法,主要用于分析解决实际问题。同时,方法本身也仍在理论探讨中不断地完善和发展。其次,各学科的研究角度也各有侧重。图情人员借助专业优势,在获取充分资料的基础上能做出比较客观全面的分析,与之类似的是新闻传播学,也能掌握大量有价值的的第一手资料来进行应用研究;社会学研究范围较广,在应用内容分析法的同时也很注重发展理论;而计算机科学、医学虽然发文不多,但在两个方面也都有研究。

析,可以反映国内研究的进展情况,通过与国外研究的综合比较,进而可以发现我国内容分析法研究中存在的问题。

3.1 选取数据来源

数据来源选取的是重庆维普资讯有限公司研制的《中文科技期刊数据库》(1989-2002)和清华大学开发的《中国期刊网专题全文数据库》(1994-2002)。此外,辅以《全国报刊索引》(社科版)进行手工检索,共收集到1992-2002年间内容分析法研究论文80篇。

3.2 我国内容分析方法的研究现状

(1) 论文数量分布情况

截至2002年6月,关于内容分析法的期刊论文按年代分布情况如表4所示。

表4 我国内容分析法研究论文的时间分布

年份	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
论文数(篇)	1	3	2	3	4	5	9	10	18	19	6
百分比(%)	1.3	3.8	2.6	3.8	5.0	6.3	11.3	12.5	22.5	23.8	7.5

从表4可以看出,自20世纪90年代以来,我国学术界开始引入内容分析方法,相关论文数量呈逐年上升趋势,尤其是近3年来,发文数量有了显著提高,这说明内容分析方法日益得到广泛的认同和重视。

(2) 论文期刊分布情况

内容分析法研究论文分布的期刊数量很多,大约有60种,根据学科范围划分的期刊分布情况如表5所示。

表 5 我国内容分析法研究论文的期刊分布

期刊类型	图情学期刊	新闻传播学期刊	社会学期刊	医学期刊	心理学期刊	综合学报	其它类
期刊种类	12	4	3	10	4	16	12
载文篇数(篇)	17	14	5	10	5	17	12
占论文总数比例(%)	21.3	17.5	6.3	12.5	6.3	21.3	15.0

从表 5 可以看出,内容分析法研究论文分布的学科面非常广泛,这也是内容分析法作为一种社科研究方法的普适性所决定的。相对而言,图书情报学、新闻传播学和医学是研究内容分析法较多的学科。

(3) 论文主题分析

从微观角度出发,就内容分析论文的研究主题进行分析,有利于了解我国内容分析法研究的重点和方向,认清优势与不足,指导今后的研究发展。

表 6 我国内容分析法研究论文的主题分布

期刊类型	图情学期刊	新闻传播学期刊	社会学期刊	医学期刊	心理学期刊	综合学报	其它类	占总数比例(%)
系统分析	4	5	5	3	4	8	3	40.0
理论研究	10	4	0	1	0	4	0	23.8
定性分析	2	4	0	1	0	3	8	22.5
定量分析	1	1	0	5	1	2	1	13.8

表 6 的数据反映了两个方面的情况:其一,系统应用内容分析法进行社会科学研究论文居多,证明研究人员具备了一定的科学精神和研究能力,因为只有将数据统计与推导分析相结合,才能更有力地说明实际问题。理论研究类文章所占比例较少,反映出内容分析方法本身仍在不断完善和推广的过程中。其二,各学科对内容分析法的研究也各有侧重。图情专业人员发挥信息优势,追踪研究动态,在引入、介绍内容分析方法方面的论述居多;新闻传播学既关注社会现实,又进行理论研究,因此各类主题均有涉及;而医学、社会学、心理学虽然发文不多,但注重实用性,大多运用内容分析法作为解决实际问题的工具。

3.3 我国内容分析法研究中存在的问题

(1) 理论研究缺乏广度和深度

总体看来,对内容分析法进行理论研究的论文所占比例并不高,而且大多数集中在图书情报领域。这一方面是由于内容分析方法本身属于实验性较强的实用型方法,大多数学科注重其作为一种分析工具的应用研究;另一方面也是由于内容分析法在我国的应用还比较少,很多需要理论指导的细节性的问题还未表现出来。而在图情领域的研究论文中,概述文章多,深入研究少。在 10 篇理论文章中,有 5 篇是从整体上介绍内容分析方法,其余几篇分别介绍了内容分析法在军事、传播、图书馆等领域的应用。无论

从数量上还是内容上,都反映出我国图书情报研究人员对内容分析法的认识还停留在比较浅的层面上,未能深入到内容分析法的方法体系之中。

然而实际上,经过半个多世纪的实践与发展,内容分析法的方法体系得到了不断的补充和完善,实用性与严谨性显著增强。从报刊文献、影音资料扩展到电子文献、网络应用领域,应用领域越来越宽广;从人工统计到计算机辅助内容分析,技术手段也越来越先进。这是一个理论性与实践性都很强的领域,有很多问题有待深入研究。

(2) 应用研究缺乏体系和标准

虽然与理论研究比较起来,目前对内容分析方法的实际应用开展得比较多,但是仔细调查就会发现,真正领会了内容分析法的精髓,全面掌握其方法步骤的并不多。内容分析方法是一种基于定性研究的量化分析方法。它用语言表示而非数量表示的文献转换为用数量表示的资料,并将分析的结果用统计数字描述。通过对文献内容“量”的分析,找出能反映文献内容的一定本质方面又易于计数的特征,从而能克服定性研究的主观性和不确切性的缺陷,达到对文献“质”的更深刻、更精确的认识^[4]。然而就目前我国的应用研究情况来看,很少有能真正符合内容分析方法客观、系统、量化三个要求的论文。因此,不仅理论研究要深入开展下去,应用研究也亟需建立起科学而完整的实施标准和方法体系,

理论与实际要结合起来,相辅相成,互相促进。

(3) 技术研究进展缓慢

作为一种基于定量分析的研究方法,内容分析法对文献资料的依赖性很大,不仅数量多,而且要求来源准确可靠,这就给研究人员带来了极大的工作负荷。从笔者搜集的文章来看,绝大多数统计工作都是由研究人员手工操作完成,没有提及运用了何种辅助工具或技术。仅有2篇理论研究类文章介绍了计算机辅助技术的进展及在网络媒介上的运用。

实际上,资料的收集、整理、统计是较为简单的重复性工作,只是因为待分析文献量浩大,才使研究变得耗时费力。随着信息技术的发展,文献统计部分的工作可以在计算机软件工具的辅助下变得轻松起来;但同时,伴随网络时代而来的“信息爆炸”又进一步加大了内容分析的难度。因此,只有不断将新技术应用于研究之中,才能提高工作效率和分析结论可信度。

4 内容分析法的原理、步骤及工具

4.1 内容分析法的原理

内容分析法的实质是对文献内容所含信息量及其变化的分析^[5],其研究目的是根据数据对内容进行可再现的、有效的推断^[6]。从哲学上来讲,该方法的可行性是以客观世界的可知论为前提的,亦即人们可以通过对客观信息的分析研究,正确认识客观世界的规律。在这一认识过程中,内容分析法强调的是正确有效的分析推理能力,其方法原理也就在于运用多种统计、推理、比较的分析方法来透过现象看本质。

4.2 内容分析法的基本步骤

来自不同学科的研究者带着各自的知识背景和实用目的开展了多种多样的内容分析研究。随着计算机技术的应用,各种研究方法开始逐渐融合、相互补充,在遵循内容分析法基本原理的基础上,研究程序基本一致。

就具体研究过程而言,内容分析法包含以下6个基本步骤。

(1)提出研究问题。由于具体问题要具体分析,因此构建一个研究大纲对于指导方法的实施是十分重要的。在研究大纲中需要确定研究目的、划定研究范围并提出假设。

(2)抽取文献样本。在不可能研究整个文献信息的总体时,就需要采用抽样方法。样本选择的标准要符合研究目的、信息含量大、具有连续性、内容体例基本一致,简言之,应能从样本的性质中推断与

总体性质有关的结论。

(3)确定分析单元,即发掘研究所需考察的各项因素,这些因素应都与分析目的有一种必然的联系,且便于抽取操作。分析单元可以是单词、符号、主题、人物,以及意义独立的词组、句子或段落乃至整篇文献都可以作为分析单位。

(4)制定类目系统,即确定分析单元的归类标准。有效的类目系统首先应具有完备性,保证所有分析单元都有所归属;同时类目之间应该是互斥和独立的,一个分析单元只能放在一个类目中;类目系统还应具有可信度,应能得到不同的编码员的一致认同^[7]。

(5)内容编码与统计。编码是将分析单元分配到类目系统中去的过程,可以借助计算机技术完成这项重复性工作,不仅速度快,而且保证了编码标准的一致性。对数据的统计工作也可以交由相应的统计软件完成,百分比、平均值、相关分析、回归分析等各种统计分析均可实现,而且精度更高。

(6)解释与检验。研究人员要对量化数据做出合理的解释和分析,并与文献的定性描述判断结合起来,提出自己的观点和结论。分析结果还要经过信度和效度的检验,才具有最终说服力。

4.3 内容分析法的软件工具

国外内容分析法相关英文软件多达数十种,而且新品种、新版本不断被开发出来。相对而言,国内在这方面的研究尚属空白,当然,这种情况与中文语法与语义结构的特殊性不无关系。这里,首先就国外英文软件研究情况谈谈内容分析法软件工具的功能与特点。

从内容分析法的实施步骤中可以看出,软件工具主要用于编码与统计过程。但是首先要解决的一个问题是文件格式的转换,各种数据资料如文本、声音、图片、图像、视频等都需要转换成计算机可读文件保存。其次,建立一套完整的分类体系也是实现计算机内容分析的中心任务之一^[8]。只有在比较完善的类目系统基础上,计算机才能对文本内容进行自动、高效的编码工作,这也正是国外学者目前积极探讨的问题。

在已开发出来的数十种内容分析软件中,并没有统一的格式与功能标准,即使在软件名称上也各有不同的描述,如文本分析(Text Analysis)、文本挖掘(Text Mining)、内容分析(Content Analysis)、文本管理(Text Management)、数据分析(Data Analysis)等等。从中可看出不同软件在功能特征上也有不同的

侧重点。举例来说,伽利略公司(Galileo Company)的文本分析软件 CATPAC 以多维排列和概念地图著称,其功能包括自动阅读文本并找出隐含概念,直接绘制概念地图并概括大意,发现语词之间的关联,其优势在于可以处理任何能编成 ASCII 或 RTF 码的语言,拥有自动编码系统而无需预编码^[9]。而 Provalis Research 为其统计软件 SimStat 开发的内容分析和文本挖掘模块 WordStat 则是专为研究文本信息而设计的,可处理的信息类型包括标题、文章、谈话、演讲、电子文献等。该软件采用词典及相关文本挖掘方法对文本自动划分类目,也可结合手工编码,系统应用编码规则,揭示个人用词差异,建立 KWIC(Keyword-In-Context,文中关键词)表以辅助修订现有的编码规则,并通过统计内部分级的一致性来评价编码的可靠性。它还包括一些探索性数据分析和图形工具,对于文献内容与类目之间的关系及相似度等,可采用等级聚类或多维排列的方法进行判定,并以图表形式直观展示^[10]。

5 内容分析法的应用前景

由于内容分析法能被应用于研究任何文献或有记录的交流传播事件,因此它的应用领域十分广泛,包括从市场和媒体研究,到人种和文化研究、社会学和政治学、心理认知科学以及很多其他领域,在不同领域中的研究能发挥不同的功效。

在社科情报工作中运用内容分析方法,充分发挥情报部门文献积累的优势以及研究人员高度的情报意识和信息处理能力,一方面可以紧跟社会热点,更好地为研究课题服务,成为管理者的“思想库”,另一方面可以结合文献计量学等研究方法,为科学管理与预测的研究提供必要依据。

在新闻传播研究中运用内容分析法,可以反映社会热点问题,探索时代发展趋势,并通过比较研究揭示各国传播、交流内容的差异,从而探讨民族、文化及人类学等方面的问题。

在政府与社会机构管理中借鉴内容分析成果,可以监测政府与组织的宣传机构运行状况,识别个人、团体或机构的意识倾向、关注焦点和交流趋势。而且,通过观察人们在态度和行为上对交流信息的反应,以及调查个人或团体的心理和情绪状态,可以有效指导行政与管理工作。

另外,内容分析成果反映出的社会语言和心理语言的密切关系、语义与语法结构的联系等规则,在

计算机人工智能开发中也发挥着重要作用^[11]。

除了学科应用领域,在空间应用领域上,席卷全球的网络媒体无疑又为内容分析开拓了一片全新的发展空间。网络时代给内容分析法提出的最重要的问题之一,是对于网上的交流信息的本质及其影响力的认识与思考^[12]。对网络进行内容分析将是一项极具挑战性的工作,这不仅是因为网络信息类型混杂,图文声并茂,动画、视频以及正待开发的触觉、嗅觉等多媒体信息,其处理难度前所未有;而且网络信息的非线性结构、交互功能、超媒体链接的广泛存在,使其长期处于一个庞大的无序状态,若要对其进行有序化,任务将极其艰巨。然而这项工作的意义却也十分重大,其研究成果将能解决目前网络信息开发、组织和利用中存在的种种问题,其研究前景将是出现一个数字化、多元化、个性化、实时化、互动化而又有序化的信息世界。正是基于对网络内容分析法的重要意义和美好前景的充分认识,各国学者就此积极开展多方面的研究,使其成为目前内容分析法的主要研究发展方向之一。

参考文献

- 1,6 Bryce Allen. 图书情报学研究中的内容分析法. 国外情报科学, 1993(1)
- 2 范并思. 论社科情报研究的方法体系突破口. 情报资料工作, 1995(2)
- 3,8 Fies Wilfried, Tamai Christian. Content analysis in empirical social research. International Journal of Educational Research, 1999(1)
- 4 马文峰. 试析内容分析法在社科情报学中的应用. 情报科学, 2000(4)
- 5 吴世忠. 内容分析方法论纲. 情报资料工作, 1991(2)
- 7 李卜乾. 内容分析法简介. 当代传播, 2000(1)
- 9 <http://www.galileoco.com/products/catpac.php>
- 10 <http://www.simstat.com/wordstat.htm>
- 11 Overview: Content Analysis.
<http://writing.colostate.edu/references/research/content>
- 12 Veare Christopher, Lin Wan-Ying. Content Analysis of the World Wide Web: Opportunities and Challenges. Social Science Computer Review, 2000(8)

邱均平 武汉大学信息管理学院教授、博士生导师, 图书情报研究所所长、中国科学评价研究中心主任。通讯地址: 武汉。邮编 430072。

邹 菲 武汉大学信息管理学院 2001 级硕士生。通讯地址同上。

(来稿时间: 2003-11-03)