

王波 编著

计算机辅助

JOURNALISM

新闻学概论

An Outline of
Computer-Assisted
Journalism

AN OUTLINE OF COMPUTER-ASSISTED JOURNALISM

新华出版社



计算机辅助新闻学概论

王 波 编著

新 华 出 版 社

图书在版编目 (CIP) 数据

计算机辅助新闻学概论/王波编著. - 北京: 新华出版社,
2000.9

ISBN 7-5011-5002-8

I. 计… II. 王… III. 计算机应用-新闻工作 IV. G210.7

中国版本图书馆 CIP 数据核字 (2000) 第 51788 号

计算机辅助新闻学概论

王波 编著

*

新华出版社出版发行

(北京宣武门西大街 57 号 邮编: 100803)

新华书店经销

新华出版社激光照排中心照排

新华出版社印刷厂印刷

*

850 × 1168 毫米 32 开本 11 印张 269 千字

2000 年 9 月第一版 2000 年 9 月北京第一次印刷

ISBN 7-5011-5002-8/G·1823 定价: 22.00 元

目 录

前言	(1)
一、计算机辅助报道：从数据的海洋中发现新闻	(9)
第一章 计算机辅助报道的产生和发展	(12)
与调查性报道一脉相承	(12)
与精确新闻学异曲同工	(14)
计算机辅助报道的早期尝试	(17)
微机和相关软件的出现扫除了技术障碍	(21)
政府文件电子化、信息服务网络化	(22)
立法措施	(24)
计算机辅助报道的作用与影响	(27)
第二章 计算机辅助报道的基本分析方法	(32)
处理计算机辅助报道中所用数据的一些指导原则	(35)
计算机辅助报道分析方法的一般策略	(36)
处理数据的基本步骤	(38)
对计算机辅助报道的一次系统调查	(46)
第三章 电子表格与数据库的使用	(57)
Excel 的基本使用方法	(59)
Access 的基本使用方法	(69)
二、计算机辅助调研：21 世纪的新闻调研方式	(89)
计算机辅助调研的基本步骤	(91)

第四章 数据库	(98)
商业性数据库	(99)
因特网上的数据库资源	(105)
由我国有关机构创建的中文网上数据库	(107)
第五章 因特网上的信息资源	(116)
因特网的发展及其特点	(117)
网上信息的特点	(120)
使用因特网信息资源需要了解的一些基本概念	(125)
网上新闻网站	(131)
政府和其他机构网站	(144)
行业网站	(149)
为新闻工作者服务的网站	(152)
第六章 电子邮件：新型的采访与获取信息的工具	(162)
电子邮件的特点	(162)
如何识别电子邮件的地址	(171)
使用电子邮件应注意的若干问题	(173)
电子邮件专用缩略语	(174)
使用电子邮件的一些小技巧	(176)
第七章 搜索引擎与信息查询技巧	(179)
搜索引擎的工作原理	(180)
常用中文主题目录与搜索引擎介绍	(183)
英文搜索引擎	(190)
搜索小窍门	(195)
三、计算机辅助引证：从纸介质的一统天下到	
光、电、磁的加盟	(199)
第八章 电子出版物	(202)
电子出版物的特点及其在新闻调研中的应用	(203)
我国新闻工作者可用来查证新闻事实的几种	

电子出版物	(206)
使用电子出版物的利与弊	(211)
第九章 因特网上的参考信息资源	(213)
网上参考信息数据库	(214)
网上图书馆	(214)
网上工具书	(218)
网上专家资源	(221)
四、计算机辅助聚会：获取信息的虚拟社区	(225)
第十章 邮件列表与新闻组	(230)
邮件列表的主要用途	(231)
邮件列表的使用方法	(235)
国内外部分免费邮件列表服务商	(237)
新闻组的工作原理与结构	(238)
如何阅读新闻组的文章	(242)
新闻组在新闻采访调研中的应用	(246)
第十一章 电子布告栏系统与电子论坛	(250)
电子布告栏系统的特点	(250)
电子布告栏站点在我国的发展与现状	(254)
如何访问电子布告栏系统	(255)
电子论坛	(259)
第十二章 计算机辅助新闻学给我们的启示	(262)
计算机辅助新闻学在美国产生和发展的必然性	(264)
计算机辅助新闻学的积极意义和局限性	(266)
计算机辅助新闻学可能引发的一些道德和 法律问题	(271)
在我国新闻媒介推行计算机辅助新闻学的必要性 与可能性	(280)
附录 1 国家或地区域名代码	(286)

附录 2	美联社有关使用电子信息服务的政策规定·····	(290)
附录 3	计算机辅助报道实例 ·····	(294)
例 1	·····	(294)
例 2	·····	(312)
例 3	·····	(316)
例 4	·····	(323)
例 5	·····	(332)
参考书目	·····	(343)

前 言

电子计算机的出现使现代新闻传媒的面貌为之一新，从印刷媒介告别铅与火走向光与电，到今天将世界“一网打尽”的因特网。尽管信息新技术为新闻媒介带来的变化数不胜数，但透过这些变化多样的表现形式，我们都能看到计算机在其中所起的神奇作用。

是计算机编辑照排系统使印刷媒介从铅与火向光与电的转变成为可能，而因特网就是由无数台相互连接起来的计算机组成。然而虽然自 70 年代以来以计算机为代表的信息存储、处理、传播技术使新闻媒介的生产方式发生了巨大的变化，这种变革更多地涉及新闻媒介中的专业技术人员，而对于广大新闻工作者，特别是那些印刷媒介的从业人员，计算机给他们带来的变化只是取代了传统的打字机（在我国则是经历了一个更为复杂的“换笔”过程）。新闻工作者调研、采访、报道仍然基本沿用传统的方法和工具，直到 80 年代初逐渐盛行于美国等一些西方发达国家的“计算机辅助新闻学”（computer - assisted journalism）才使这一状况有了重大的改变。

这种以利用计算机网络上各种丰富的信息资源和强大功能去搜索、核实新闻信息，利用计算机对大量数据进行分析，并以分析结果作为报道的线索和素材为主要特征的新思路、新方法对传统的新闻传播在调研、采访、报道等方面都产生了重大影响，其

作用早已远远超越了计算机进入新闻编辑部的初期阶段——计算机主要被用于文字处理和排版的范围，影响了越来越多的新闻工作者的日常工作方式。

美国米德尔伯格公司与哥伦比亚大学新闻研究生院合作，在过去的几年中，一直对新闻工作者使用因特网的情况进行每年一次的调查。据 2000 年 3 月公布的对 400 多位报社杂志社主编的最新调查结果显示，美国的新闻工作者在新闻采访和报道中使用因特网的人数有了大幅度的增加。在接受调查的人中，每天至少上网一次查询有关信息的占 75%，而在 1999 年的调查结果中这一比例仅为 48%。75% 的新闻工作者每周在家上网的平均时间达 4.7 小时，在工作单位上网的平均时间为 8.7 小时。对网上的有关文章进行调研，发现和构思报道的主题，收集有关参考资料已经取代了收发电子邮件成为新闻工作者上网的主要目的。同时更多的新闻工作者，通过电子邮件与采访对象以及受众建立并保持经常性的联系。^①

按照美国的一些新闻学者的观点，计算机辅助新闻学主要包括以下四个方面的内容：

1. 计算机辅助报道（computer - assisted reporting）主要是指新闻工作者在计算机上利用电子表格进行复杂的计算，利用统计软件对大量数据进行分析，利用数据库软件建立原创信息集合，然后将上述计算机应用所产生的计算和分析结果作为新闻报道的线索或素材，用来揭示事物的发展趋势，或对信息加以核实验证。

2. 计算机辅助调研（computer - assisted research）指为报道目的而进行的专门调查或信息搜索。调研与报道之间的区别在于，报道一般依赖于第一手的、独立的原始信息来源，例如新闻工作者的采访、观察或自己利用计算机进行的数据分析；而调研则主要依靠第二手的、间接的信息来源，例如现有的数据库或文献资

料。计算机辅助调研的主要信息来源是以电子读物形式出现的各种报告、文献，商业数据库或信息服务商所提供的在线信息、数据库以及因特网上的各种信息资源。

3. 计算机辅助引证 (computer - assisted reference) 是指使用在线的或光盘形式的各种辞典、百科全书、地名索引、年鉴和汇编等对新闻报道所涉及的事实、定义、概念，字词的确切写法和含义等加以核实确定。

4. 计算机辅助聚会 (computer - assisted rendezvous) 指利用因特网上的各种聚会场所寻求相关信息，了解他人对某一问题的意见主张，就某一问题向专家请教，登陆到其他电脑网络以利用其信息资源，被认为是计算机辅助新闻学中最新颖、最令人心动的一个领域。目前一些西方发达国家新闻工作者常用的“虚拟社区” (virtual community) 包括邮件列表 (mailing list) 新闻组 (news group)、电子布告栏系统 (BBS)、聊天室 (chat room) 等。^②

计算机辅助新闻学的前身是现在构成其主体的计算机辅助报道。计算机辅助报道产生于 60 年代末的美国，当时美国社会矛盾的激化使公众对新闻媒介消极地等待政府机构和企业组织的官员和政客单方面提供信息，然后就事论事地加以报道感到强烈不满，这一局面迫使一些新闻工作者去探索将社会科学研究分析的方法和计算机数据分析引入新闻报道，以便对新闻事件发生的原因及可能产生的影响进行更深入、全面的报道。

由于这种新的报道方法具有从看似乱无头绪的大量数据中发现新闻线索，揭示事物的某些发展趋势的潜在能力，因此自其诞生之日起，计算机辅助报道在很长一段时间内主要用于揭露渎职舞弊等违法行为的调查性报道。在计算机辅助报道发展的早期阶段，数据的计算和分析必须在费用高昂、技术复杂的大型计算机主机上进行，因而在很大程度上限制了这一报道方式的推广和普

及。整个 70 年代在新闻报道中利用计算机进行数据分析，仅限于少数先驱者的尝试，所发表的几篇报道也未能引起比较大的反响。

从 80 年代初起，简便实用的微型计算机和大量用户友好的实用软件的问世，使计算机走进了新闻编辑部，为计算机辅助报道的普及提供了技术和设备上的条件，而美国各级政府机构、企业组织相继实行的文件纪录计算机化，也在客观上加快了新闻机构利用计算机读取、分析这些资料的步伐。在这些因素的共同作用下，计算机数据分析在新闻报道中有了更加广泛的应用，出现了“计算机辅助报道”这一正式称谓。随着计算机辅助报道的逐步普及，目前它不仅被用于大规模的调查性报道，也频繁出现于日常报道中，并产生了一批有影响的作品，从 1989 年《亚特兰大宪法日报》的记者比尔·戴德曼在揭露当地银行业批准住房抵押方面的种族歧视的报道中运用计算机数据分析，并获得当年的普利策调查性报道奖开始，运用计算机辅助报道方法写出的新闻作品已有多篇获得普利策奖及其他新闻大奖，这标志着计算机辅助报道这种新的报道方式已经得到广大新闻工作者和读者的认可。

数字技术的不断发展使人类存储、处理信息的能力产生了质的飞跃，各种形式的电子出版物、数据库一问世就以其信息容量大、检索方便的特点显示出在新闻研究中的巨大潜力。而以数字技术、多媒体技术和网络技术为核心的因特网将全世界数亿台计算机连在一起，初步实现了全球范围内的信息共享，更是新闻工作者重要的信息资源以及研究、采访和报道的重要工具。一种区别于新闻工作者以采访本、笔、录音机、剪刀、胶水、卡片、剪报为主要工具的传统新闻采访、调研方法的新方式逐渐形成。

这一新方式以利用数字化技术和计算机网络技术检索、获取和交流信息为特征，它极大地扩展了新闻工作者信息搜索和采访

的范围，提高了他们研究和采访的效率，于是一些西方新闻学者将通过电子出版物、在线数据库和因特网而实现的大范围、快速搜索信息、引证核实新闻事实、与各种采访对象建立联系的种种功能概括为计算机辅助调研、计算机辅助引证和计算机辅助聚会，并将这三个方面与计算机辅助报道一起统称计算机辅助新闻学。

从以上的叙述中我们可以看出，计算机辅助新闻学是一个包含多种意义，本身并不十分严谨的术语，对于这一术语及其定义，西方新闻界尚未形成比较一致的看法，对于其内涵、外延也存在一些不同意见。部分新闻界及学术界人士甚至声称这一概念本身就难以成立，他们认为，如果因为在新闻报道的过程中使用了计算机，就立下一个计算机辅助报道的名目，那么以此类推是否也应当有电话辅助报道，甚至自来水笔辅助报道？还有的人认为计算机辅助报道或计算机辅助新闻学等类似的术语给人造成一种它们是与传统的新闻报道和新闻学完全不同的两回事的错觉。

本书的作者认为，诸如此类的质疑虽然不无道理，但一个不可否认的事实是，以计算机网络为代表的信息新技术在新闻报道中的应用，已经使信息采集的方式发生了实质性的变化。电话、电报甚至当计算机仅仅用来进行文字处理时，他们所起的作用仅限于提高新闻采写的效率、提高新闻报道的时效性，扩展新闻报道的领域与范围，而计算机辅助报道则可以说在某种程度上改变了新闻报道的定义，新闻工作者不再仅仅等到某一新闻事件发生后，再去记录事件发生的过程，而是利用计算机及相关统计软件，对数据加以分析，从过去仅靠新闻敏感和传统的调研采访工具手段无法看出新闻的地方发现新闻线索，它给新闻报道带来的影响可以说是革命性的，这种影响超出了以往其他信息传播技术和工具。

至于那种认为提出计算机辅助新闻学的概念会造成误导的担

心，笔者认为，计算机辅助新闻学不过是传统新闻实践在数字网络时代的延伸和补充，它不可能也无法完全取代传统的新闻采编基本技能和工作方式，但毕竟二者之间有着明显的区别。虽然西方新闻界尚未就计算机辅助新闻学这一术语达成共识，一些新闻工作者和学者也在使用“数据库新闻学”、“在线新闻学”、“网络空间新闻学”等各种术语去描述这一新的调研、采访和报道方式，但这种现象其实并不奇怪。

当一种新的理论与实践问世时，他们的倡导者总要急切地打出标新立异的旗号，以示自己的理论与实践是如何地不同于其他流派，因此当计算机深入社会生活的各个领域，并带来革命性变化的时候，我们就有了计算机辅助设计、计算机辅助教学、计算机辅助诊断等一大批冠以计算机辅助的新术语去描述这些新事物、新现象，究其实这主要是由于计算机在这些行业和领域的创造性应用还属于新生事物，还需要大张旗鼓地去倡导。试想，如果在这些领域中计算机各种先进、强大的功能已得到普遍应用，并成为其不可缺少的一个组成部分时，还会有谁去画蛇添足地加上“计算机辅助”的描述成分呢？

鉴于在西方新闻工作者和媒介研究人员中使用比较普遍的是计算机辅助报道、计算机辅助新闻学等术语，本书的作者在书中也采用了这些基本从英文直译过来的术语。作者认为，当务之急是把计算机辅助新闻学这一新发展介绍给国内的新闻界同行。

其实我国已有一些新闻工作者在不同程度上将这种被西方学者称为计算机辅助新闻学的新方法应用于采访调研，并取得了可喜的成果。中国社会科学院新闻与传播研究所 1998 年在《人民日报》、新华社和《中国日报》的 1100 名编辑记者中就使用计算机和因特网的情况而作的一次调查表明，89.5% 的编辑记者使用计算机，27.9% 的人经常上网。新闻工作者使用最多的数字化信息资源是光盘数据库，常用这一信息资源的人达 27.8%，51.3%

的人使用过本单位的在线数据库，超过 40% 的人使用过电子邮件，经常使用者达 18.8%，超过 30% 的人使用过浏览器。

那些在将因特网作为发掘新闻素材、采访的工具方面先行一步的编辑记者已经从中大大受益，例如一位女记者利用电子邮件在一周之内采访了美国、英国、日本、瑞士、加拿大等国的 10 位科学家，由此而写成的报道后来被刊登在美国的《科学》杂志上。一位年轻记者从《北京晚报》上看到一条消息，说德国人巴蒂斯（Bates）所收集的有关南京大屠杀的一些资料存放在耶鲁大学神学院，于是他通过雅虎网站找到了这些资料，并且发现了一个专门传送巴蒂斯手记原文的站点，这位记者通过电子邮件与该网站联系，第二天便获得了所需的信息。

这些编辑记者认为在信息时代计算机和网络对于新闻报道的作用至关重要，他们将这种作用形象地概括为“使记者看得更远、听得更清、想得更深、写得更快”。^③

然而与新闻实践相比，我们对这一领域的研究显然滞后了，目前国内除个别文章、论文对这一题目略有涉及之外，对计算机辅助新闻学尚缺乏比较系统的介绍与研究，而这种状态是与我们的新闻媒介要在 21 世纪用先进的技术武装自己，在信息阵地的争夺战中抢占制高点的要求和任务不相适应的。正是基于这样的想法，作者利用在美国马里兰大学新闻学院以休伯特·汉弗莱学者的身份进修学习的机会，选修了计算机辅助报道、在线新闻学等相关课程，回国后又参阅了大量的有关资料，对国内的有关情况作了调查研究，在此基础上编写了这本《计算机辅助新闻学概论》。

作为国内第一本比较系统地介绍计算机辅助新闻学的实用性专著，作者力求以通俗的语言介绍计算机辅助新闻学各个不同组成部分的发展与现状，并通过大量研究和报道的实例说明计算机辅助报道、计算机辅助调研、计算机辅助引证、计算机辅助聚会

的基本方法及其具体应用，书中包括大量国外利用计算机辅助报道方法写成的报道实例，以及对国内外网上信息资源的介绍。

由于本书是一本入门性质的读物，因此叙述的重点不在于诸如具体操作这样的技术细节，对于计算机辅助报道所使用的有关统计软件的介绍也仅限于一般个人微机中都附带的 Excel，Access 等常用软件，作者的目的是让中国读者看过本书后能够对计算机辅助新闻学有一个基本的了解，并希望能够引起国内新闻界对这一领域的关注。作者相信，只要我们善于用批判的眼光对计算机辅助新闻学进行认真的分析研究，就能够从中找出我们可以借鉴、利用的东西，为我国的社会主义新闻事业服务。

本书包含大量新的报道实例和对网络信息资源的介绍，作者希望不论是对于广大的新闻工作者、新闻院校的师生，还是网络爱好者和普通读者，本书都能成为使他们感兴趣的一本知识性读物。

注释

①米德尔伯格/罗斯第6次年度网络空间调查报告新闻稿《尽管新闻工作者认为因特网缺乏可信性，但经过不同程度的确认后仍然使用网上信息》<http://www.middleberg.com>

②Nora Paul, "Computer Assisted Research - A Guide to Tapping Online Information," third edition. The Poynter Institute for Media Studies p. 1. 1.

③卜卫、刘晓红：《新闻记者的网络使用：〈人民日报〉、新华社、〈中国日报〉记者使用计算机和互联网的报告》，《新闻传播研究》1998年第3期。

一、计算机辅助报道： 从数据的海洋中发现新闻

1985年的一段时间内，位于美国东北部的罗得岛州校车接连发生恶性交通事故，造成多名儿童伤亡，这引起当地居民的严重关注，也成为当地新闻媒介报道的热点。

在美国，中小学学生乘坐的校车倍受优待，当校车停下来让孩子上下车时，来往车辆都要停车等候。担任校车司机的人也必须经过有关当局审查批准，通过专门的考试，领取专门的驾驶证方可上岗，正因为这一点，在不论是面积还是人口都属于小州的罗得岛发生的这一系列校车交通事故更显得异乎寻常。

舆论对于造成事故的原因众说纷纭，莫衷一是。当地的《普罗维登斯日报》的记者在报道此事时，决定对此展开深入的调查，这一任务落在了计算机辅助报道的先驱者之一，当时在该报任记者的埃里奥特·加斯宾和他的一名同事的身上。

两人首先查阅了肇事司机的有关文件记录，发现这些人在过去的3年内都有比较严重的交通违章记录，而这样的人按照有关规定是不能担任校车司机的。

两位记者并没有停留在就事论事地报道这些事件，而是决心对全州所有的校车司机的情况作一次调查，以便确定类似这几位肇事司机的情况在全州的校车司机中是否带有普遍性。他们在报社的一位新闻资料员的帮助下，得到了一份输入计算机的所有注册校车司机的名单，他们将这份名单与过去3年中罗得岛州交通

违章的名单相对照，结果发现罗得岛州每5名校车司机中就有1人在过去的3年中有过一次严重交通违章，一些人违章纪录达一二十次。

随后他们又将校车司机名单与刑事法庭的记录相对照，发现一些校车司机曾因持有毒品，进行毒品交易，赌博和敲诈勒索等罪行被判刑。

根据通过计算机数据分析得到的这些线索，加斯宾和他的同事对有关的机构和个人进行了广泛的采访，他们发现罗得岛州对校车司机的审批制度上存在着严重缺陷，例如没有对申请当校车司机的人加以考察的明文规定，不要求申请人提供自己过去的驾驶记录和犯罪记录，不对申请人的全部驾驶记录和犯罪记录进行审核，而是依赖当地警方官员的话，但不同地方的警方官员往往有不同的标准。而且一个人一旦被批准作校车司机后，即使违反了法律也不披露此人的身份，驾驶执照管理制度也存在着严重缺陷，而州有关当局至少在四年前就对此有所耳闻，却没有采取任何行动。

两位记者后来发表的有关报道，揭示了罗得岛州校车频繁发生事故的深层原因，报道发表后引起社会各界的广泛重视，有关部门对校车司机驾驶执照的审核颁发制度进行了彻底的改革，65名不合格校车司机的驾驶执照被吊销。（报道的全文见书后的附录3《州当局核查制度不严，交通肇事与犯有重罪者竟成校车司机》）

至此，为了查清有违章驾驶记录和犯罪记录的校车司机，两位记者已经核对了100多万条记录。显然，如果没有计算机的帮助，类似这种规模的搜索、分类、检索是难以想象的。罗得岛州虽然不大，但也有近2000名校车司机，要查明这些人的交通违章记录、刑事犯罪记录就必须到所有有关的部门查询几种不同的记录，然后把这些记录一一加以对照，才能查清每一位司机的情