

●袁毅

网络信息资源内容评价关键指标研究*

摘要 互联网上存在大量低质量的、重复的甚至错误的信息。人们为了准确地利用信息,必须学会判断信息质量的基本方法。在专家建议的大量评价指标基础上,筛选出网络信息内容评价的6项关键指标:权威性、准确性、客观性、时效性、覆盖度及利用程度。并从定性和定量分析两种途径详细给出判断方法。参考文献8。

关键词 网络信息资源 内容评价 评价指标 定量分析 定性分析

分类号 G253

ABSTRACT There is a large quantity of low-quality, repeated and erroneous information. To better utilize information, we should learn methods for the judgment of information quality. In this paper, the author proposes some key indicators for the evaluation of the contents of network information resources, such as authority, exactness, objectivity, timeliness, coverage and usage. 8 refs.

KEY WORDS Network information resource. Content evaluation. Evaluation indicator. Quantitative study. Qualitative study.

CLASS NUMBER G253

由于缺乏传统媒体的同行评价过程、发表文章的门槛低、发布的信息受各种潜在的广告商、赞助商的利益影响等各种原因,互联网上充斥着大量低质量的、重复的甚至错误的信息。Beredjikian 等人对互联网上整形外科信息质量进行了调查与评估,被评估网站内容质量平均得分只有 28.4(总分 100 分),其中 23% 的网站提供的信息不正确甚至是误导的^[1];Piero Impicciatore 调查了互联网上家庭护理发热儿童的信息,在被调查的 41 个网页中,只有 4 个网页里的内容与权威机构关于护理发热儿童的建议基本吻合,而其他 38 个网页,均或多或少地存在问题,其中两个网页上还出现非常严重的错误^[2]。Harkness Connell 等检查了网络参考咨询答案的准确性,在有答案的网页中,只有 75% 的回答是正确的^[3]。

解决互联网信息质量低的方法之一是对网络信息质量进行评价。大量文献对网络信息资源的评价指标进行了研究,有的评价机构提出的指标多达 100 余个。然而,过多的评价指标不具备现实可操作性,而且也没有进一步给出具体实施方法。

文献[4],[5],[6],[7]较全面地提出或汇总了网络信息资源评价指标,其中高频出现的指标是:权威性、准确性、客观性、时效性、覆盖度及利用程度。本文认为,这六项指标基本上满足对网络信息资源的

内容评价的需要。以下对这六项指标的使用方法作深入探讨。

1 权威性

网站的权威性是指网站在用户中建立起来的可信程度。它通过两个方面体现:一是网站及其网页的权威性,二是网站作者或编辑的权威性。网站的权威性不仅产生于网站本身质量和网页质量,还受到建站机构在社会上声誉的影响;网站上高质量的内容依赖于网站内拥有较高学术水平的作者群和著名的编辑、编委进行质量控制。

1.1 网站或网页的权威性

网站或网页的权威性判断主要有 10 个方面。

(1) 网站的主页是否清楚地表明了建站的机构或个人。

(2) 是否清楚地标示了网站的版权所有者,如在主页上标注 copyright。

(3) 是否写明了网站的责任者或管理者,如“the trustees of Stanford University”表示责任者为斯坦福大学。

(4) 是否提供了更具权威的上层管理机构的链接。如一个大学下设的研究所的网站设置了它的上层管理机构网站的链接,如有的网站在主页的首要位

* 本文系江苏省哲学社会科学研究“十五”规划基金项目“核心网站评价及发现的理论与方法”系列成果之一(项目编号:04XWB013)。

置设置了上层管理机构的链接,有的网站在每一页均设置了与上层管理机构网站的链接。

(5)网站是否清楚地显示了联系电话、Email、传真及网站的真实地址。<http://www.checkdomain.com> 提供了网站背景检索数据库,用户只要输入某网站的网址,就能立即检索到某网站详细的背景信息。

(6)网站是否清楚地反映了它在现实社会中的身份。现实社会中的权威性是可以移植到网络上的。有的网站吸引了大量的链接,正是因为它在现实社会中的声誉所起的作用。即使网站的所有者缺乏现实生活中的声誉,但真实地表示出现实社会中的身份,也会增强网站的信度。

(7)网站是否有保护知识产权方面的说明,是否有相关的资格。如有的网站发布的作品与作者签有协议,或在网站发布联系作者以支付转载费的信息,有的公布了信息内容的使用权资格的相关文件。一个重视知识产权保护的网站发布的信息会更加规范和可靠。

(8)网站上是否有外界对网站的明显或潜在的正面评价。有些评价机构定期或不定期地对网站进行评比、排名,有的媒体对网站的情况进行评价与介绍,网站将这些对自己网站的评价公布在网站上,这样会增强网站的权威性。

(9)网站上是否有隐私保护声明。一些规范的网站明确地向用户提示隐私保护的相关政策及说明,一些网站在用户使用站点之前,弹出提示框,显示网站对用户注册信息的保密承诺,或提示网站使用了Cookie等采集用户检索信息的软件,如数据库 Ei Village 和 ProQuest Research Library 的网站均有此类说明和提示。

(10)网站或网页被外部网站链接的次数可以反映网站或网页的权威性或影响力。例如,清华大学网站的外部链接数远大于一般大学网站的外部链接数。链接数可以通过搜索引擎 AltaVista 和 Alltheweb 检索,这两个搜索引擎是目前最可靠的计量各种类型链接数的工具。

1.2 网站作者的权威性

作者权威性是网站内容质量评价的重要指标。网站作者权威性主要通过以下方面考察:网站的权威作者总数,作者发文量,作者著述被引量,作者经历和身份,作者的资历,作者的专业背景,作者机构的权威性和声誉,作者发布信息是否有商业利益,作者言论是否客观,作者所在机构的宗旨及主要业务,作者是否

能够联系。基于上述判断标准,本文提出定量测评网站作者权威性的3种方法。

(1)领域专家在网站中出现的次数。ISI最新产品 ISI Highly Cited 提供了各学科高被引作者的名录,查到某领域专家的名录后,利用 Google 提供的站内检索功能,就可以检索到某专家在网站出现的次数。

(2)网站中所有作者的影响因子之和。某作者的影响因子是指该作者发表的全部文章所在期刊的影响因子之和。对网站的所有作者的影响因子求和,该数值的高低反映了网站是否拥有一批高质量的作者群。

(3)统计每个网站中出现教授、专家、博士的网页,根据这类网页数量的多少,判断网站的作者权威性。Nicholson 提出了利用文献挖掘(bibliomining)的方法,抽取网站网页中出现“Dr., PhD, Professor”的次数,以此作为网站的作者权威性指标^[8]。但他的方法依赖于文献挖掘工具,对于一般用户,直接利用 Google 提供的站内检索功能,如“professor site: www.cumt.edu.cn”,即可以检索到网站中出现教授的网页,不同的是,后者统计的是出现教授、专家、博士的网页数,前者是出现教授、专家、博士的次数。

2 准确性

尽管网络信息资源缺乏质量控制过程,但仍有一些基本的方法判断网络信息资源的准确性。

(1)网站上是否有质量控制过程或手段。尽管网站缺乏同行评价的过程,但是,一个好的学术网站特别是一些电子期刊的网站仍会适度地延续一些传统载体的审查制度,或采取一些特殊的质量控制方式,例如,有的文学网站聘请自愿者使用专门的校对软件为他们校对网上的信息内容;有的网站显示了质量控制的活动和奖惩制度;有的网站社区通过各种奖励政策来促进网民向社区提供高质量信息,如医学网站丁香园采取的正是这一策略,如果用户要看高质量的文章,必须有一定的积分,而这积分来源于用户曾经对网站的贡献。特别要强调的是,如果一个网站公布了其编辑部、编辑部的编委,说明网站内容具有质量控制过程,这类网站内容的准确性高于其他网站,如果编委是某学科的权威学者,无形中还会提高网站的权威性。

(2)该网站的网页上是否有错别字、病句等低级错误。如果低级错误都无法控制,说明该网站缺乏基本的质量控制。

(3) 抽样检查网站中的概念解释、原理、实验方法、结论是否与百科全书、教科书相符。

(4) 对网站上提供的电话或电子邮箱等联系信息进行检验。一个网站如果连联系的方式都出错或不更新,网站内容的准确性也很难保证。

(5) 图、表、数据表示是否清楚,是否清楚地标注了数据来源,是否与文字相对应。互联网上存在大量仅有文字缺少图、表、数据的文章,一个高质量的网站是不应该有这类问题的。

(6) 日期是否清楚。

(7) 搜索引擎 Google 中的 PageRank 有时可以作为评价网页准确性的工具。用户只要安装了 Google tools,IE 浏览器中就会出现 Google 工具栏,工具栏中显示 PageRank 图标,只要浏览器打开某一网页,PageRank 会自动显示该网页的网页级别(从 0 到 10 的所有自然数)。文献[3]以网络参考咨询网页作为样本,发现 PageRank 大于 5 时,网页信息准确的可能性是网页信息不准确可能性的 1.89 倍(通过似然比判断)。PageRank 高低还与网页的准确性存在正相关。PageRank 与一般网页的准确性是否存在相关性,尚需进一步研究。

3 客观性

网站内容的客观性判断依赖于用户对某领域情况的熟悉程度。这里包括用户对本学科知识的了解程度,用户对科学共同体中的著名人物、事件、组织的熟悉程度,对本领域中存在的与经济利益或政治目的相关的一些活动的了解程度。当然,有些网站容易判断。比如有些网站发表的言论与社会公认的正义相悖,有明显的歪曲事实的现象,如纳粹网络、邪教组织网站。

更多的网站信息内容与广告、赞助者的信息相混淆,这时,网站的客观性难以判断。传统媒体的广告与媒体的内容是分开的,用户能够非常清楚地区分哪些是媒体发布的内容,但是网络完全不同于传统媒体。首先,网络发布信息、收集信息、传播信息的低成本及便捷性,使广告商、有各种政治目的的人、生产商蜂拥而至,发布各种有一定目的信息。第二,生产商的网站为了吸引潜在用户,除发布自己产品和服务的相关信息,还在网站上提供大量与商业利益无关的信息,这样,网站上客观的信息往往与不客观的信息混淆在一起。第三,提供信息的过程可能嵌入潜在的广告。如一个牙科医生给网站提供牙科知识的同时,演

示自己的手术过程。

显然,一些文献提到以网站是否有广告判断网站信息质量的客观性是不正确的。有广告或赞助的网站,提供的信息未必都是有偏见不可信的,没有广告或赞助的网站提供的信息未必都是客观的,关键的问题是要清楚地了解广告、赞助者与信息的关系,判断信息产生的动机,遵循一些基本的原则对网站内容进行客观性判断。网站的客观性判断至少要考虑以下两个方面的问题:

(1) 区分不同类型广告和赞助的目的和动机。广告一般包括商业广告、公益广告、机构广告。商业广告以追求商业利润为目的。网站上的商业广告主要是以旗帜广告形式出现,或直接由生产商创建网站,在网站上发布产品信息、服务信息、厂商信息和客户信息等。公益广告是为了促进公共政策、解决争端、解决社会问题为目的的广告,如环境保护、SARS 防治、普及教育等,发布者通常是政府或非营利性组织。机构广告不是为促进产品和服务而是为促进本机构或组织发展获取社会支持的广告。

赞助有合作赞助和非营利组织的赞助两种。前者是为了获得公众形象而对某项活动提供经济或物资方面的赞助。如体育比赛等一些活动中的赞助。后者通常是为了促进科学、社会、文学艺术的发展而设立各种基金,对学术研究或活动进行赞助。

商业性广告如果以旗帜广告或按钮等形式出现,用户很容易区分广告和信息内容。对于生产商或服务商自己建立的网站,则要非常慎重地与同类产品的相关信息对比。公益广告通常不存在偏见和其他目的。对于机构广告则要了解机构的背景和广告的目的,当不带商业目的及政治偏见时,是值得信赖的。合作赞助虽然目的是提高公众形象,但往往在网站上设链,通过链接,用户可以访问赞助者的网站,因此它常常与商业性的广告混淆。非营利组织赞助的网站或网页的内容,大多数是学术研究的成果,并在网页上注有本项目受某某基金资助的字样,这些成果往往代表学术界较高的学术水平,提供的信息被认为是真实客观的并且质量较高。

(2) 对于专业性较强的信息,应采用多种途径进行客观判断。例如咨询专业人员、对比教科书、与权威网站提供的信息对比。

4 时效性

时效性是指信息发布的一定时间内所能起的作

用。

传统正式出版物一般标注较规范的日期信息。但是互联网没有日期标注的规范;日期标注极其混乱,且意义模糊。

互联网上的时间信息主要有:信息创作的完成时间,信息发表在印刷型载体上的时间,信息上载或在网站发布时间,信息更新时间,信息的版权期,引文时间,图表及数据的相关时间。有的网页标注较全面的时间信息,有的网页甚至没有任何时间信息,大多数网页标注1~2项时间信息:更新时间和信息出版(或发布)时间。

如果网页上没有显示更新时间,对于静态网页,只要在地址栏输入“javascript: alert (document. lastModified)”,IE浏览器就会显示出所打开网页的最近一次更新日期。但对于非动态的网页,就无法了解网页的更新日期了。尽管如此,用户常常将更新日期误以为信息发表(发布)的日期。考察网站时效性,主要根据以下几个方面:

(1)显示最新更新的时间。

(2)显示信息在服务器上发布的时间。

(3)信息在传统媒体上发布的时间。

(4)文中引用图表、数据或其他资源的时间,有些数据需要显示更详细的时间信息,如引用数据的时间,被引数据发表的时间,被引数据采集的时间。

(5)在网站的主页注明版权日期。

(6)所有的日期(如是英文)均要写成国际标准格式:dd-mm-yy。

当一个学术网站的网页达到上述要求,并且更新时间较短时,这个网站是一个经常更新的、较规范的网站。

5 主题覆盖度

主题覆盖度是网站内容涉及某学科或主题内容的广度和深度。广度与深度是对于特定的目标用户而言。判断网站的主题覆盖度主要方法有:

(1)从网站的导航图、目录、频道及网站自身的介绍来判断。

(2)从各种媒体对网站的评价中了解。

(3)请专业人员主观判断网站是否具有广度和深度。

(4)统计某领域主题特征词在网站中出现的网页数,如果出现网页数多,则被认为主题覆盖度广。主题特征词是能体现某学科、专业或主题内容特征并

能区别于其他学科、专业或主题内容的关键词。主题特征词的提取方法有两种:一种是通过训练样本网页获得。例如,当人们需要提取网络计量学的主题特征词时,可以从网络计量学最相关的样本网页上进行分词,找出频率较高的词,然后再由学科专家从中评选出具有区分能力的词,形成主题特征词库。另一种方法是利用权威数据库(如SCI),检索主题相关文献,将所有相关文献的题名或摘要合成文本(或者直接提取文献中主题词或叙词),统计出现频率较高的词,最后再由学科专家从中选出具有区分能力的词作为主题特征词。得到主题特征词库后,利用搜索引擎Google的站内搜索功能,检索网站中所有包含主题特征词的网页之和,当一个网站包含的主题特征词的网页之和较大时,说明该网站较其他网站具有更强的主题覆盖度。但这种方法主要是判断主题覆盖的广度。

6 利用程度

评价指标中多强调了目的用户、实用性、适用性、易用性,但这些概念非常模糊,评价时不具有可操作性。事实上,这些特征可以通过用户的利用情况来定量判断。网站的利用程度主要通过采集以下数据进行考察:

(1)网站的一般访问量,即网站被用户访问的次数。

(2)网站的深度访问量,包括网站的页面浏览次数、页面浏览时间、用户在网站停留的时间。一些学术网站(如<http://www.elsevier.com>)还有更为细致的用户访问日志:用户访问的主题类目、高频关键词、高频访问的文章、不同时间段访问人数、来自不同地域访问者的数量、文献下载量、全文浏览量等。

(3)网站的用户数。网站多要求使用者注册,个人网站也会设立一个邮件列表来联系访问者。注册用户或者参加邮件列表的用户必定是这个网站的忠实用户,一个网站拥有一批忠诚度较高的用户,本身就体现了网站信息的可用性、实用性。

(4)网站被文献引用量。即根据文献引用某网站的次数来确定网站的被利用程度。一个网站如果被大量的文献引用,说明网站的学术性较强,网站的信息具有一定的学术价值和实用价值。该方法主要适用于学术网站利用程度评价,可以根据评价对象,选择不同的检索工具检索这一指标。ProQuest Research Library、中国学术期刊数据库、(下转第85页)

里,与10多个国家接壤,少数民族主要居住在边疆地区,有20多个民族跨境而居,这为民族古籍的国际化交流提供了得天独厚的条件。在国外,日本、印度、前苏联及欧美各国都建立了许多研究所、学会和国际常设机构,研究蒙古学、藏学、满学、突厥学、敦煌学等,并出版了大量的中国少数民族文献。我们应充分利用独特的地理位置环境和特殊政策,广辟国际信息渠道,广泛开展国际交往。其内容大致包括:①各国互换民族古籍,做到文献资源共享;②各国共同编辑民族古籍资料索引及民族古籍研究刊物;③合作进行专题民族古籍调研;④合作编辑出版民族古籍;⑤及时交流民族古籍工作经验;⑥合作进行民族古籍工作人员培训;⑦合作购置和共同开发运用大型的现代化设施,如建立公用的民族古籍数据库;⑧制定共同的民族古籍工作规范,便于国际通用;⑨互相聘请民族古籍专家或互派访问学者。

参考文献

- 1 国家民族事务委员会《中国民族工作五十年》编委会. 中国民族工作五十年. 北京:民族出版社,1999
- 2 华林. 西部大开发与少数民族历史档案保护政策研究. 档案学研究,2002(2)
- 3 华林. 论藏文历史档案的发掘利用. 中国藏学,2003(4)
- 4 张公瑾. 民族古文献概况. 北京:民族出版社,1997
- 5 包和平. 少数民族古籍的科学管理和开发利用. 中国图书馆学报,2001(1)

包和平 大连民族学院图书馆馆长,研究馆员. 通信地址:大连市. 邮编116600。

何 丽 中国民族图书馆馆长,副研究馆员. 通信地址:北京市. 邮编100031。

(来稿时间:2005-01-11)

(上接第57页)中国社会科学引文数据库、中国科学引文数据库、中文生物医学数据库均有检索网站被文献引用量的功能。

网站内容质量评价是一项非常重要的工作,开发有效的评价指标和评价方法,不仅有利于评价机构和信息服务机构更好地开展信息资源导航推介服务,更为重要的是,可以指导众多的互联网用户快速有效地判断网络信息的质量,减少错误信息的传播,也有助于信息构建者遵循基本的质量控制方法,创建规范的高质量的信息资源。

网站内容质量评价又是一项艰巨的任务。必须将定性评价与定量评价方法相结合,才能找到有效的的评价方法。定性评价有助于解决大量无法量化的指标评价问题,有助于减少定量分析绝对化造成的误差、误判;而定量分析则可以克服定性评价的主观性、缺乏可比性等弱点,能够快速地对客观数据进行对比分析。本文对网站的权威性、主题覆盖度、利用程度分别提出了定量判断的方法,对网站内容质量定性评价向定量评价发展作了一次尝试。

参考文献

- 1 Beredjiklian PK, Bozentika DJ, Steinberg DR, et al. Evaluating the source and content of orthopaedic information on the Internet: The Case of Carpal Tunnel

Syndrome. J Bone Joint Surg Am,2000,82(11)

- 2 崔雷,张晗. 医学网址评价标准与方法. 情报理论与实践,2000(4)
- 3 Martin Frické, Don Fallis. Indicators of accuracy for answers to ready reference questions on the internet. Journal of the American Society for Information Science and Technology, 2004,55(3)
- 4 陈雅,郑建明. 网站评价指标体系研究. 中国图书馆学报,2002(5)
- 5 张咏. 环境科学网络信息资源评价研究:[博士学位论文]. 南京:南京大学,2002
- 6 王晓丽. 基于专业类网站评价体系研究. 情报杂志, 2002(10)
- 7 Janet E. Alexander. etc. . Web wisdom: how to evaluate and create information quality on the Web. London: Lawrence Erlbaum Associates, 1999
- 8 Nicholson, S. bibliomining for automated collection development in a digital library setting: using data mining to discover web. Journal of the American Society for Information Science and Technology, 2003, 54(12)

袁 毅 中国矿业大学文法学院副教授,硕士生导师,南京大学信息管理系博士生. 通信地址:江苏徐州. 邮编221008。

(来稿时间:2005-02-25)