

全球视野下的“数据新闻”：理念与实践

方洁 颜冬

摘要

“大数据”是2012年以来国内新闻传播学界热议的话题。与之不同的是，关于“数据新闻”的研究却非常欠缺。后者是当下全球新闻业应对大数据时代发展变革中产生的新兴领域，且被视为未来新闻业的发展趋势。本文试图通过对国内外关于数据新闻的阐释文献的梳理，结合主流媒体实践数据新闻的经典案例剖析，解读“数据新闻”的概念内涵与基本特征，理清数据新闻与精确新闻、计算机辅助新闻报道、数据可视化之间的关系，展示数据新闻实践的最新成果。

关键词

数据新闻、精确新闻、数据可视化、新闻报道、计算机辅助报道、大数据

作者简介

方洁，中国人民大学新闻与社会发展研究中心研究员，新闻学院讲师。电邮：fangjie2008@vip.sina.com。

颜冬，中国人民大学新闻学院研究生。电邮：yayd1213@163.com。

本文得到教育部人文社会科学重点研究基地自主项目资助。

Data Journalism: Theory and Practice

FANG Jie, YAN Dong

Abstract

A number of studies have been done toward the thesis of “big data” since 2012, but few researchers have paid attention to another concept of “data journalism”. Data journalism, a new concept emerging from the innovation and development of modern journalism practices, is believed to be the future of journalism. This paper reviewed different views on data journalism from home and abroad and sorted out some main stream media practices for case studies. Thus it tried to define the concept, describe its basic features and differentiate data journalism from the concepts of precision journalism, computer-assisted reporting and data visualization for a whole review.

Keywords

data journalism, precision journalism, data visualization, news reporting, computer-assisted reporting, big data

Authors

Fang Jie is assistant professor at the School of Journalism, Renmin University of China. Email: fangjie2008@vip.sina.com.

Yan Dong is postgraduate student at the School of Journalism, Renmin University of China. Email: yayd1213@163.com.

This paper was funded by independent projects of humanities studies under the Ministry of Education.

“数据新闻”(data journalism)，也被称为“数据驱动新闻”(data-driven journalism)，这一概念近年来发端于新闻实践领域。西方主流大报和一些独立新闻机构(如ProPublica¹)设立了专门团队来设计一些新型的新闻应用：即运用各种技术软件来抓取、处理、分析和形象化呈现数据，数据呈现方式包括可视化数据图、互动图表和网络在线演示等，开启了数据新闻实践领域的第一页。如今包括中国在内的世界各国的传统媒体、新兴网站和独立新闻机构正在逐步接受“数据新闻”的理念，并进行相应的实践尝试。

相比如火如荼开展创新的业界，学界对这一新兴领域的关注尚嫌不足，针对“数据新闻”的理论阐发滞后于新闻实践。本文试图通过对国内外关于数据新闻的阐释文献的梳理，结合主流媒体运用数据新闻的经典案例剖析，理清数据新闻理念的内涵与特征，分析数据新闻产生的背景及原因，呈现数据新闻实践的最新成果。

一、对“数据新闻”内涵与特征的梳理

关于“数据新闻”概念的界定，学界尚无定论。我们只能从媒体实践人员对之介绍、总结的博客、视频和其他相关网络文章阐述中，对这一新兴的传媒词汇的内涵与特征进行描摹。操作数据新闻的主要人员大都精于计算机技术，供职于主流的新闻机构或是资深的独立撰稿人，制作过精美的数据新闻作品，他们自称为“数据记者”(data journalist)，一些较为活跃的数据记者组建了类似“数据驱动新闻网站”(<http://datadrivenjournalism.net/>)和“数据新闻博客”(<http://www.datajournalismblog.com/>)等专业网络平台分享和交流实践经验。综观这些人对“数据新闻”的看法，我们发现有如下三种理解此概念的角度：

(一) 从新闻呈现形态的角度理解“数据新闻”

对于大多数从事数据记者而言，数据新闻是有别于传统新闻报道的新报道形态。如荷兰数据记者亨克·范艾斯(Van Ess, 2012²)认为数据新闻使记者能够通过发现、制作、呈现大量数据这一套工作流程展现未曾报道过的新闻故事，找到新的报道角度。一些数据新闻不一定符合传统意义上好新闻的标准，因为它们着重于呈现问题在哪里，而不是解释问题为何出现。一个好的数据新闻成果有多个层面，你既可以深入了解许多个性化的细节，也能看清整体状况。

当下最为流行的关于“数据新闻”的研究是一本名为《数据新闻手册》(Data Journalism Handbook)的小册子，它也是全球第一本专门探讨数据新闻的著作。该书由欧洲新闻学中心(European Journalism Centre)和开放知识基金会(Open Knowledge Foundation)倡导，初步成型于2011年在伦敦Mozilla大会(Mozilla Festival)³上的48小时工作坊，尔后通过几十位来自BBC、德国之声、卫报、纽约时报、美国在线新闻、ProPublica等机构的数据新闻倡导者与操作专家以网络协作方式编写而成，并由O'Reilly出版，可在“知识共享署名—相同方式共享协议”(Creative Commons Attribution-ShareAlike license)下在线免费阅读。该书不仅尝试解答数据新闻是什么，为什么的问题，还展示了国际各大主流媒体运营数据新闻的案例。

在这本手册中,来自伯明翰城市大学的保罗·布拉德肖(Bradshaw)和德国之声电视台的米尔科·劳伦兹(Lorenz)总结了“数据新闻”的几种呈现形式(O'Reilly, 2011):数据新闻与其他新闻的不同之处或许在于当你将传统的“新闻鼻”、讲述扣人心弦的故事的能力与大量的数据信息结合在一起时,新闻报道能呈现出许多新的可能:如运用软件去发现成百上千官方文件中的关联性,英国《每日电讯报》集团对议员花销的调查性报道就是一例;数据新闻可以使记者运用互动式信息图表报道复杂的新闻故事,如汉斯·罗斯林在Gapminder网站上关于“可视化的世界贫穷状况”的演讲已经吸引了来自全球的数百万网友点击观看;数据新闻可以解释新闻与每个人存在何种联系,如BBC和《金融时报》所作的交互式预算报道可以帮助你找出预算如何对你个人产生影响;数据新闻还能向受众公开新闻获取的过程,如英国《卫报》利用“数据博客”与受众之间进行分享的成功经验。

(二) 从新闻生产流程的角度理解“数据新闻”

一些数据记者认为,数据新闻不只是丰富了新闻表现的形式,还改变了新闻生产流程。米尔科·劳伦兹主持了2010年在阿姆斯特丹召开的第一届国际数据新闻圆桌会议(data-driven journalism Amsterdam round-table),该会议由欧洲新闻中心(European Journalism Centre)和荷兰阿姆斯特丹大学(the University of Amsterdam)主办。他在会议演讲(Lorenz, 2010)中指出:数据新闻是一种工作流程,包括以下基本步骤:通过反复抓取、筛选和重组来深度挖掘数据,聚焦专门信息以过滤数据,可视化的呈现数据并合成新闻故事。数据新闻可被视为一个不断提炼信息的过程,在这一过程中,原始数据转换成有意义的信息。当把复杂的事组织成条理清晰、易于理解和记忆的故事时,公众才能获取更多益处。

北卡罗来纳大学教堂山分校的名誉教授菲利普·梅耶(Philip Meyer)是精确新闻学理论的奠基人,他认为“在信息量不足的时代,记者主要的精力在于寻找和获取信息,然而处于信息丰富的今天,信息处理的过程就显得尤其重要。信息处理过程包含两个层次:一个是通过分析不断变动的数据以找到其中的意义和结构,另一个则是通过展示让用户了解哪些信息对他们具有重要性和相关性。数据新闻要像科学一样严谨,它公开其方法,呈现其结果,经得起核实验证。”(Howard, 2012a)

(三) 从新闻行业发展的角度理解“数据新闻”

“数据新闻”正被越来越多的从业者认可,他们认为“数据新闻”将成为新闻业未来最重要的发展方向。他们给予“数据新闻”在新闻行业发展中如此重要的地位,主要源自两个观点:

第一,很多业界人士认为促使数据新闻成为未来新闻业主流的主要原因源自当下的时代背景。在“数据新闻博客”中,网站发起者玛丽安·布查德(Marianne Bouchart)在网站介绍中就表示:“现在的科技和设计软件发达,尤其是在维基解密事件之后,开放知识(Open Knowledge)的理念流行,记者必须要学会新技术才能追上潮流,而数据新闻则是未来新闻的基本工具。”⁴互联网的兴起改变了人们对于知识生产与共享的传统观念,信息公开成为越来越多民主政府的选择,开放源代码软件的流行将使数据可视化变得更为容易,正是技术和理念的双重支持促使新闻业做出相应的变革。

第二,一些从业者进一步从数据新闻之于公众的价值角度进行阐释。如有“互联网之父”之称的蒂姆·伯纳斯·李(Tim Berners-Lee)对数据新闻发展前景的非常乐观(转引自Data

driven journalism网站, 2012年4月23日), 他说:“数据新闻就是未来, 记者需要成为‘数据通’, 过去的记者能通过酒吧聊天获取新闻, 这种方法今天有时还管用, 但是未来记者需要钻研数据, 能够运用工具分析数据, 找到其中有趣的部分。记者还要正确看待数据, 帮助公众获知: 在哪些部分零散的数据可以进行整合, 呈现整体状况, 从而了解这个国家正在发生怎样的变化。”另一位英国的独立多媒体记者亚当·韦斯特布鲁克在接受欧洲新闻中心的采访中明确表态(Westbrook, 2010): 数据新闻是未来新闻业最具发展潜力的领域之一。对此, 他解释道: 数据新闻使新闻回归本质: 挖掘公众无暇处理的信息, 核实信息, 理清信息的内涵后将之发布给公众。如果更多的记者这样工作(而不是依赖于新闻发布会上提供的数据), 公众将更加文明和有见识。可见, 在波谲云诡的大数据环境下, 新闻从业者能通过“数据新闻”对数据的收集、处理和解析帮助公众了解身处环境的风云变幻, 实现媒体服务公众利益的基本职能。

综合上述, 本文认为从狭义上看, “数据新闻”的内涵就是“基于数据的抓取、挖掘、统计、分析和可视化呈现的新型新闻报道方式”。数据新闻的核心是对数据的处理, 正如记者大卫·麦克坎德莱斯(David McCandless, 《信息之美》一书的作者)所言:“数据正变得越来越重要, 这并不是因为数据的量大, 而是我们拥有了工具和能力去分析数据, 找出模式、结构并揭示趋势(转引自The Guardian网站, 2013年4月4日)。”可见媒体要推行数据新闻业务, 首要是解决数据处理技术的问题。

从广义上看, “数据新闻”是在大数据时代新闻学发展形成的新领域, 它代表未来新闻业发展的一大方向, 并包含以下基本特征:

- 1) 以服务公众利益为目的。这是数据新闻的出发点, 所有数据的处理和呈现归根究底是为了让公众理解我们身处的大数据时代中数据变迁的内涵, 了解宏观数据如何影响每个人;
- 2) 以公开的数据为基础。这是数据新闻存在的前提, 如果政府、社会其他组织不公开信息或者没有提供联网数据库, 缺乏数据分析材料, 数据新闻也不可能得以推行;
- 3) 依靠特殊的软件程序对数据进行处理, 开掘隐藏于宏观、抽象数据背后的新闻故事。这是数据新闻的技术保障, 也是数据新闻得以和一般新闻相区别的核心特征;
- 4) 以形象、互动的可视化方式呈现新闻。这是数据新闻的个性展现形式, 得益于科学可视化的发展, 数据新闻将复杂、抽象、难懂的数据转化为简单、具体、生动的新闻报道。

二、与“数据新闻”相关的概念辨析

在查阅西方阐释数据新闻的文献时, 笔者看到阐释者经常将数据新闻与“精确新闻”、“计算机辅助新闻报道”、“数据可视化”等概念相联系, 这一方面呈现出数据新闻这一新的新闻领域与新闻学其他领域之间存在着一定的历史延续性, 它不是无根之木, 而是基于已有的学科基础发展形成的; 另一方面通过比较也可以进一步剖析和认识数据新闻的特征, 帮助我们理清这一领域与相关领域之间的关系。

(一) 数据新闻与“精确新闻”(Precision Journalism)

谈起对数据的统计分析, 人们自然会联想起上世纪60年代兴起的精确新闻。

精确新闻学是由美国学者菲利普·迈耶在20世纪60年代提出来的。在《底特律自由报》有着丰富报道经验的菲利普·梅耶出版了专著《精确新闻学——一种社会科学报道的理论》, 并

于1973年转入美国北卡罗来纳大学任教。该书详尽地叙述了精确新闻报道的调查方法、统计技术和写作要求,已译有中文版,并在国内新闻界产生了较大的影响。

精确新闻主张将社会调查研究方法应用到新闻实践中,科学的收集资料、查证事实,用数据来说话,从而提高新闻报道的准确性和客观性。从上世纪70年代开始,美国报业对“精确新闻”进行广泛的探索,大多数美国新闻院系将“精确新闻学”作为专业课程,足见其对美国新闻业的影响。

谈及精确新闻和数据新闻的差异,概念提出的时代背景不容忽视。首先,“数据”之于上世纪70年代的美国和当下全球的意义有很大差别,无论是数据的量还是人们认识数据、呈现数据的手段都有了飞跃。正如郭晓科博士所言“数据新闻学是精确新闻学的进一步延伸,数据新闻学使新闻生产过程更为精细化,它对新闻工作者的技能要求除传统的文字写作、音视频制作外,还包括社科研究方法,计算机数据抓取、处理、可视化,平面/交互设计,计算机编程等多个领域(郭晓科,2013:30)。”其次,互联网的快速发展使线上、线下社会融合,网络文化的盛行改变了知识生产的传统模式。如果说精确新闻更多的是从媒体的专业性角度出发,探寻的是利用数据增加报道的准确客观,那么数据新闻则更多的是从用户的角度出发,探寻数据对于用户有何意义,而用户也从新闻的被动接受转向主动获取、参与制作,他们在数据新闻产制中占据重要的地位。再次,从影响的范围看,精确新闻的概念由菲利普·迈耶提出,其理念传播和实践范围主要在美国,对世界其他国家也有影响,但影响的范围和广度有限。而数据新闻并不是由某一位学者提出的概念,这个概念兴起于一个遍布全球的数据记者群,其影响的范围更广(关于这一点,本文还将进一步论证)。

(二) 数据新闻与“计算机辅助新闻报道(Computer-Assisted Reporting,简称CAR)”

在阐述数据和新闻的关系时,另一个经常被提及的相似的报道方式叫“计算机辅助新闻报道”。计算机辅助新闻报道是用计算机来辅助收集和处理信息的新闻报道方式。学者卜卫(1998)曾梳理计算机辅助新闻报道的发展历史,指出其发端于上世纪50年代,伴随计算机技术的发展,其经历了使用大型机处理政府数据库、发现和调查新闻事实的早期阶段;再到上世纪70—80年代通过PC机普及、商业和政府数据库的进一步开放形成的“以新闻报道为目的,对任何计算机化的信息来源的处理和使用行为”的第二阶段,发展到上世纪90年代中期以后利用互联网进行新闻采集、分析和制作的第三阶段。

卜卫(1998)认为计算机辅助新闻报道和精确新闻报道有一种不可分割的联系,计算机辅助的是精确新闻报道过程中的数据采集、处理和分析。但是后者不仅能辅助精确新闻报道,还能辅助其他类型报道,它超出数据采集和处理的辅助范围,在各种类型的新闻及新闻制作各环节发挥作用。

米尔科·劳伦兹(Lorenz,2010)则强调计算机辅助新闻报道主要起一种辅助工具的作用,它不是影响整个新闻生产流程的基本方法。运用计算机搜索大型数据库仍然是调查性报道需要的重要技能,但是数据新闻是一个完全不同的概念,它使计算机辅助新闻报道只是未来新闻报道链上的一环而已。

从新闻业历史发展的维度看,计算机辅助新闻报道是较早出现的概念,其次是精确新闻,然后才是数据新闻,三者有一定的延续关系,但并没有相互取代。鉴于计算机在数据收集处理

过程中的普及化运用,当下无论精确新闻,还是数据新闻,都不能脱离计算机辅助新闻报道。但是正如前面提及的学者所言,计算机辅助新闻报道更偏向于一种辅助工具,它不是一种独立存在的新闻报道方式,而强调一种方法和运用,从这个角度看,数据新闻的概念代表着一种新闻发展的形态,其概念的内涵和外延比计算机辅助新闻报道更加广阔。

(三) 数据新闻与“数据可视化”

提到数据新闻学必然会触及“数据可视化”,这个概念的含义更为广阔,其并不局限于新闻领域,而是与计算机技术、设计和编程等相关,正是这种技术的应用使得数据和新闻完美结合,产生信息图表和动态信息图等集科学、实用与美观为一体的数据新闻报道案例。

早在1686年,英国科学家埃德蒙·哈雷制作了世界信风地图,这张地图生动展现了全球信风的分布与状态,是有据可查的最早的科学可视化案例。1987年,由布鲁斯·麦考梅克(Bruce H. McCormick)、托马斯·德房蒂(Thomas A. DeFanti)和玛克辛·布朗(Maxine D. Brown)所编写的美国国家科学基金会报告《科学计算之中的可视化》(Visualization in Scientific Computing)强调了新的基于计算机的可视化技术方法的必要性,促进了这一领域的研究。

对于“数据可视化”,有学者作了如下定义:在科学可视化基础上诞生的“数据可视化”将大量数据组合构成数据图像,同时将数据的各个属性值以多维数据的形式表示,使人们能够以更直观的方式从不同的维度观察数据及其结构关系,发现数据中隐含的信息(韩卫国等,2004)。数据可视化是一个处于不断演变之中的概念,其涵盖的技术方法十分广泛,运用的边界在不断地扩大。尤其是“大数据”时代,“数据可视化”也被视为应用技术,跟数据挖掘、数据管理等技术紧密联系。如今,可视化作为一门涉及计算机图形、图像处理、计算机视觉、人机交互等多个领域的综合学科,不但广泛应用于如医学、生物、地理等领域的科学计算,而且在如金融、通信、网络等行业中信息处理方面的应用亦是如火如荼(李妹芳,2011)。

数据可视化丰富了数据新闻的表现形式。杰夫·麦考吉(Geoff McGhee)曾任职于《纽约时报》与ABC新闻,2009—2010年他在斯坦福大学研究数据可视化,并发布了一组名为“数据时代的新闻”(Journalism in the Age of Data)的系列视频⁵,邀请各大型媒体的主要技术人员和数据记者来阐释他们如何用数据可视化做新闻。McGhee教授认为记者现在面临前所未有的数据冲击,要利用好来自艺术界、计算机领域的数据可视化技术来发掘数据在新闻报道领域的潜力。

另一方面,数据新闻并不等同于数据可视化,英国《卫报》的数据记者和该报网站数据博客的编辑西蒙·罗杰斯(Rogers, 2011a)认为数据新闻并不拘泥于可视化,它也可以以一个新闻故事的形式呈现,甚至仅仅展示数据。数据新闻是运用新的新闻叙事形式,如何运用形式是灵活和富有弹性的。

三、西方媒体在“数据新闻”领域的实践

从宽泛的意义上看,西方媒体在“数据新闻”领域实践的历史可以追溯到19世纪。如1821年5月5日《卫报》创刊第一期上的一篇调查未成年教育系统的报道就被视为该报最早的数据新闻(Rogers, 2011b)。西蒙·罗杰斯认为虽然数据处理的技术已今非昔比,但是如今的数据新闻

和早期数据新闻在报道动机上是相同的 (Rogers, 2011b)。尽管数据新闻在媒体报道中很早即已出现,但是类似的尝试是零碎的,缺乏系统化,也缺乏相应的报道理念支撑。

谈到较为系统地推行数据新闻报道,美国《纽约时报》和英国《卫报》最值得关注。2007年,美国《纽约时报》建立了一个记者加程序员的团队,即现在的“互动新闻技术部”(Interactive News Technologies Department),它交叉于技术和采编部门之间,探索在线新闻的报道形式。这个团队在《纽约时报》网站上为总统选举、奥运会等重大新闻制作推出了大量动态、互动的信息图表,每一张图表都由大量数据作支撑(徐涛,2013)。《纽约时报》网站为此专门在“多媒体报道”板块下开辟了“互动”频道,汇总此类报道。

尽管如此,《纽约时报》并没有对数据新闻做出明确的分类,而是将一些数据新闻报道放在多媒体报道板块中。与之相比,英国《卫报》则清晰采用“数据新闻”概念并对这一理念进行大力推广。2009年,《卫报》网站开设“数据商店”(data store)版块,下分“大数据”、“数据新闻”、“数据博客”等细分频道。“数据商店”向用户开放数据库的链接与搜索,为用户提供大量来自政府、博物馆、大学、研究机构、非政府组织的公开数据。

《卫报》“数据博客”频道主编西蒙·罗杰斯于2013年4月出版了专著《事实是神圣的:数据的力量》(Facts are Sacred; the power of data.)。这本书展示了《卫报》数据记者如何分析数据,阐述了数据对世界带来的影响。他提到数据改变的不仅是新闻业,它使得做新闻和其他行业之间的界限消失,整个世界都在数据的影响下。

谷歌公司的凯瑟琳·赫尔利(Kathryn Hurley)在准备参加O'Reilly举办的年度Strata大会(该会议的宗旨是“让数据运转”、“用数据改变世界”)时对《卫报》“数据商店”下属的“数据博客(Data blog)”频道的工作流程和方法进行了为期一周的观察,她总结卫报数据团队使用工具和资源处理数据有三个步骤:获得数据、讲故事、分享数据(Hurley, 2012)。

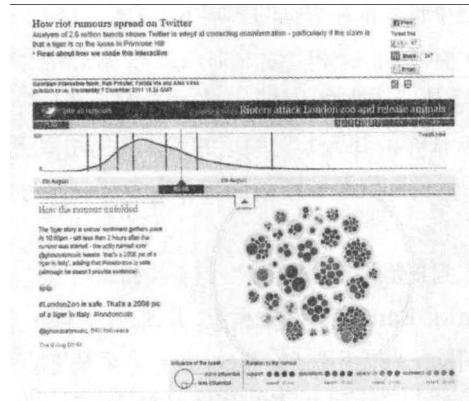
首先,当数据团队察觉到重要的新闻事件时,记者编辑就会到互联网上搜索包括Office for National Statistics和World Bank等公共数据源,并通过Google Refine和DataWrangler等工具对数据做初步处理,清除一些无用信息。其次,数据分析是讲故事的主要部分,分析结果能帮助大众理解故事和数据。在实际运用中,数据团队经常采用Google Fusion Tables、Google Maps API、CartoDB、Google Spreadsheets、DataWrapper和Tableau等数据可视化工具。最后,在故事中添加下载原始数据的链接,方便用户研究数据、发现故事、做出评论,或是找出数据中存在的差错和质疑之处。

和受众分享数据的过程从某种角度上改变了新闻产制的过程。卫报网站为用户提供培训课程,教授用户制作数据新闻的方法。对数据感兴趣的用户可以自己查证新闻中的数据源头,从中发现隐藏的新闻。媒体数据团队和用户之间搭建了合作平台,集体协作生产新闻,使“众包新闻”(crowdsourcing journalism)成为现实。

以2009年针对英国议员花销的调查性报道(MPs' expences)为例,6月,英国国会下议院公布了5500个PDF文档,其中包括议会所有646名成员四年间的花销报账单。《卫报》为此开展了众包新闻的尝试,设立专题,将所些数据公布给用户,邀请用户帮忙查找账单中是否有可疑或可议之处。数万名用户参与了这一数据新闻的制作,帮《卫报》挖出大量新闻,该报道被视为众包

新闻的成功案例。

“数据新闻”业务的成功开展为《卫报》赢得了业界的诸多赞誉。2011年8月，伦敦发生大骚乱，参与骚乱者运用手机和社交媒体进行联系，让警察大伤脑筋。英国政府因而对社交媒体进行强硬批评，并试图推行限制措施。与此同时，《卫报》网站与伦敦政治经济学院合作，历时一年多完成了《解读骚乱》(Reading the Riots) 的全媒体报道。报道中一个重要的组成部分——对260万条Twitter数据的统计分析就是一个数据新闻报道的典范。卫报数据团队运用交互图的方式呈现大骚乱期间在社交网络中流传的七大谣言是如何传播扩散的。这则名为“骚乱谣言如何在Twitter中传播”的数据新闻报道将谣言用支持、反对、质疑、批评四种态度量化，谣言的起始是一个圆点，随着谣言的扩散，圆点会演变成圆圈，并聚合成集群。图中针对每个谣言设计了时间轴，伴随时间的移动，用户可以看到各个谣言生成、扩散、消亡的过程。这一形象的展示方式依赖数据团队和科研人员的大量数据研究所得，科学呈现了社交网络中谣言的生成到消失的过程。这则报道最大的影响是修正了人们运用经验判断的误差，用数据分析证明社交媒体具有一定的谣言的自我澄清机制，人们虽然借助社交平台传播和扩散谣言，但同时用户对这些谣言的反对、质疑和批评也会产生一定的反作用，并最终驱散谣言。这一系列报道的推出在一定程度上影响了英国政府对社交网络管理的公共政策，《卫报》网站也因之获得了2012年度“数据新闻奖”(Data Journalism Awards)。



图：“骚乱谣言如何在Twitter中传播”报道网页截图⁶

“数据新闻奖”由非赢利的民间组织“全球编辑网络 (Global Editors Network)”主办，谷歌给予支持，是第一个专门为数据新闻报道设立的奖项。该奖项从2012年开始在全球颁布，当年参与评奖的作品达到286个之多，涉及全球51个国家和地区（见表），最终评出了包括上述《卫报》骚乱报道的六个年度最佳数据新闻。提交参奖作品的既有机构也有个人：媒体是主体，共有101家媒体参奖，此外，自由撰稿人、广告机构、公司、社会组织、大学和政府机构也参评。参评的数据新闻作品制作团队平均由3.96人构成，人数最多的团队高达30人之多，作品耗时从8小时到7年不等，平均需要5个多月。从全球编辑网络对参评作品的上述统计，我们可以做出如下推断：1) 数据新闻的实践并不局限于上文提及的英美两国，而是在全球多个地区推行，2) 参与数据新闻制作的人员并不局限于媒体，3) 数据新闻制作多以团队为主，4) 优秀的数据新闻作品往往需要较长的制作时间。

分布大洲	欧洲	北美洲	非洲	亚洲	南美洲	大洋洲
参加评奖的主要国家和地区	英国 荷兰 西班牙 罗马尼亚 俄罗斯 乌克兰	美国 加拿大 墨西哥	肯尼亚 尼日利亚 埃及	菲律宾 印度 吉尔吉斯 斯坦 阿联酋 中国香港	巴西 阿根廷	澳大利亚 新西兰
参评作品总量	134	80	22	32	12	6

表：2012年度数据新闻奖参加评奖者的全球分布表⁸

数据新闻实践的开展催动新闻教育和研究的跟进。一方面，一些高校、科研机构与媒体合作，为媒体收集、处理数据信息提供智力支持，共同完成规模较大的数据新闻报道，如前面提及的《卫报》的《解读骚乱》，另一方面，部分高校和科研机构开设相关的研究中心，专门从事数据新闻业务的研究。如2012年4月30日，陶氏基金会 (Tow Foundation) 与奈特基金会 (John S. and James L. Knight Foundation) 宣布将提供200万美元来资助哥伦比亚大学新闻学院进行数字时代新闻报道的研究，其中公共数据和新闻的透明度、数据可视化等数据新闻相关领域是资助研究的重点⁹。

结语

数据新闻是新闻业面临时代变化做出的积极应对，它既丰富了新闻报道的呈现形式，让受众从此类报道中感受到理性之美、深度之美，又促使新闻报道流程随之变革，使“众包新闻”这种合作共享式的生产方式得以在现实中成功运转，进而改变了专业媒体与用户之间的关系模式。与此同时，“身处信息时代的记者比以往任何时候都更加需要对大量数据进行核实、分析和整合 (Howard, 2012b)。”数据新闻的兴起对新闻从业者提出了更高的要求，然而学界对数据新闻的关注和研究才刚刚起步，有诸多问题仍需要解答。新闻实务教学者和从业者都需要更新知识结构，以适应新传播环境的变化。

注释 [Notes]

- 1 ProPublica是美国一个独立的非盈利新闻编辑部，为公众利益进行调查报道。2008年开始运营，有专业的调查记者团队，并与多家传统主流媒体合作，2010年和2011年均获得普利策奖，是首家获此殊荣的网络媒体，其网站为 <http://www.propublica.org/>，介绍详见网站中About Us版块。来源：<http://www.propublica.org/about/>。
- 2 原文为荷兰语，后经笔者和作者电邮联系后确认了其中关于“数据新闻”定义的英文版，此处是根据证实后的英文版翻译而成。
- 3 Mozilla是一个全球化的、非盈利的，致力于创建自由、开源技术的新型软件公司，“火狐浏览器”是其最具影响力的产品。该公司每年一届的大会（Mozilla Festival）汇集了数百名创新技术人员。
- 4 参见Datajournalismblog网站的简介，<http://www.datajournalismblog.com/about/#>。
- 5 该视频网页参见<http://datajournalism.stanford.edu/#>。

- 6 参见《卫报》网站的报道, <http://www.guardian.co.uk/uk/interactive/2011/dec/07/london-riots-twitter>。
- 8 数据来自“全球编辑网络”对“2012数据新闻奖”的总结文章, 网页参见 <http://www.globaleditorsnetwork.org/dja/entries2012/>。
- 9 参见奈特基金会的相关报道, <http://www.knightfoundation.org/press-room/press-release/new-research-effort-columbia-university-seeks-best/>。

引用文献 [Works Cited]

- 卜卫 (1998). 计算机辅助新闻报道: 信息时代记者培训的重要课程. 《新闻与传播研究》, (1), 11-20.
- [Bu, Wei (1998). Computer-assisted reporting: an important skill training for journalists in an informational age. *Journalism & Communication*, (1), 11-20.]
- 郭晓科 (2013). 《大数据》. 北京: 清华大学出版社.
- [Guo, Xiaoke (2013). *Big data*. Beijing: Tsinghua University Press.]
- 韩卫国、王劲峰、王海起等 (2004). 基于数据可视化的交通流量分析. 《武汉理工大学学报》, 28(5), 668-670.
- [Han, Weiguo & Wang, jinfeng & Wang, Haiqi. (2004). Traffic flow analysis of data visualization. *Journal of Wuhan University of Technology*, 28 (5), 668-670.]
- 李妹芳 (2011). 译者序.载Julie Steele, Noah Iliinsky,《数据可视化之美》(祝洪凯, 李妹芳译). 北京: 机械工业出版社.
- [Li, Meifang. (2011). Translator's preface. In Julie Steele & Noah Iliinsky, *Beautiful visualization*. Beijing: China Machine Press.]
- 徐涛 (2013年1月28日). 纽约时报实验. 《第一财经周刊》, (4), 2013年5月24日访问于第一财经周刊网站, <http://www.cbnweek.com/yuedu/ydpage/?rajd=2825>.
- [Xu, Tao. (Jan. 28, 2013). The innovation of the New York Times. *CBN Weekly*, (4), Retrieved on May 24, 2013 from CBNweek.com, <http://www.cbnweek.com/yuedu/ydpage/?rajd=2825>.]
- Data Driven Journalism (Apr. 23, 2012). Data Journalism Handbook: Why is data journalism important? Retrieved on May 11, 2013, from datadrivenjournalism.net, http://datadrivenjournalism.net/news_and_analysis/data_journalism_handbook_why_is_data_journalismImportant.
- Howard, A. (Feb. 14, 2012a). The bond between data and journalism grows stronger. Retrieved on May 11, 2013, from The Strata, <http://strata.oreilly.com/2012/02/data-journalism-computer-assisted-reporting-government.html>.
- Howard, A. (May 15, 2012b). Profile of the data journalist: the long form developer. Retrieved on May 22, 2013, from The Strata, <http://strata.oreilly.com/2012/05/profile-of-the-data-journalist-10.html>.
- Hurley, K. (Nov. 4, 2012). Think like a data journalist. Retrieved on May 22, 2013, from The Guardian, <http://www.guardian.co.uk/news/datablog/2012/nov/14/think-like-data-journalist>.
- Lorenz, M. (2010). Status and outlook for Data Driven Journalism. Paper presented at the Conference of data-driven journalism Amsterdam round-table. The Netherlands. http://mediapusher.eu/datadrivenjournalism/pdf/ddj_paper_final.pdf.
- O'Reilly. (2011). *Data Journalism Handbook*. Retrieved on May 11, 2013, from DataDrivenJournalism.net, <http://datajournalismhandbook.org/>.

- Rogers, S. (July 28, 2011a). Data journalism at the Guardian: what is it and how do we do it? Retrieved on May, 22, 2013, from The Guardian, <http://www.guardian.co.uk/news/datablog/2011/jul/28/data-journalism>.
- Rogers, S. (Sep. 26, 2011b). The first Guardian data journalism: May 5, 1821. Retrieved on May 22, 2013, from The Guardian, <http://www.guardian.co.uk/news/datablog/2011/sep/26/data-journalism-guardian>.
- The Guardian. (Apr. 4, 2013). What is data journalism at the Guardian? Retrieved on May 11, 2013, from The Guardian, <http://www.guardian.co.uk/news/datablog/video/2013/apr/04/what-is-data-journalism-video>.
- Van Ess, H. (2012). Handboek datajournalistiek. Retrieved from <http://www.boomlemma.nl/communicatie-media/catalogus/handboek-datajournalistiek-1>.
- Westbrook, A. (2010). Data-driven journalism: what is there to learn? Paper presented at the Conference of data-driven journalism Amsterdam round-table. The Netherlands. http://mediapusher.eu/datadrivenjournalism/pdf/ddj_paper_final.pdf.

(责任编辑：李新颖)