

新闻与可视化

FACTS

ARE SACRED

THE POWER OF DATA

数据新闻 大趋势

释放可视化报道的力量


[英] 西蒙·罗杰斯 (Simon Rogers) 著
岳跃译

数据新闻和数据可视化的领军人物、
Google 趋势数据主编分析师西蒙·罗杰斯的扛鼎之作

媒体人必读的数据新闻领域的巅峰之作

喻国明、金兼斌、黄志敏、陈为、沈浩
周葆华、马金馨、吴晓琪、张静波

众多传媒精英联袂推荐

 中国人民大学出版社

FACTS
ARE SACRED
THE POWER OF DATA

数据新闻大趋势

释放可视化报道的力量

【英】西蒙·罗杰斯 (Simon Rogers) 著
岳跃译

中国人民大学出版社
· 北京 ·

图书在版编目 (CIP) 数据

数据新闻大趋势:释放可视化报道的力量 / (英)罗杰斯著;岳跃译. — 北京:中国人民大学出版社, 2015. 4
ISBN 978-7-300-21068-1

I. ①数… II. ①罗… ②岳… III. ①新闻学—研究 IV. ①G210

中国版本图书馆 CIP 数据核字 (2015) 第 070431 号

数据新闻大趋势:释放可视化报道的力量

[英]西蒙·罗杰斯 著

岳跃 译

Shuju Xinwen Daqushi: Shifang Keshihua Baodao De Liliang

出版发行 中国人民大学出版社

社 址 北京中关村大街 31 号

邮政编码 100080

电 话 010-62511242 (总编室)

010-62511770 (质管部)

010-82501766 (邮购部)

010-62514148 (门市部)

010-62515195 (发行公司)

010-62515275 (盗版举报)

网 址 <http://www.crup.com.cn>

<http://www.ttrnet.com> (人大教研网)

经 销 新华书店

印 刷 北京联兴盛业印刷股份有限公司

规 格 215mm×275mm 16 开本

版 次 2015 年 5 月第 1 版

印 张 20.75 插页 2

印 次 2015 年 5 月第 1 次印刷

字 数 310 000

定 价 128.00 元

版权所有

侵权必究

印装差错

负责调换



西蒙·罗杰斯 (Simon Rogers)

现任谷歌趋势数据主编分析师 (Trend Data Editor)。曾任 Twitter 公司首任数据编辑，屡获殊荣的英国《卫报》网站“数据博客”和“数据仓库” (guardian.co.uk/data) 前任主编、《卫报》新闻编辑。通过和图表设计团队通力合作，他赋予了数据生命、使之跃然纸上。他曾深度参与和推动了《卫报》诸多开拓性创举，发动读者用“众包”的方式分析了 45 万条英国国会议员开支记录，以及阿富汗和伊拉克维基解密战争日志。他是 2012 年度英国皇家统计协会“杰出数据新闻奖”得主，被牛津大学互联网研究所提名为“2011 年度英国最佳互联网记者”。

推荐序一

大数据新闻：未来新闻生产的逻辑与方法

2014年3月,《纽约时报》发布了长达96页的《纽约时报创新报告2014》。面对赫芬顿邮报、Buzzfeed这样的互联网媒体公司的激烈竞争,《纽约时报》在报告中忧心忡忡地提出:“我们真的落后了”。美国皮尤研究中心在2014年的媒体研究报告中,直言不讳的指出“我们正在寻找那些能够适应数字媒体采编流程的新闻采编多面手,比如数据库记者、可视化记者和计算机记者。在大数据时代,数字内容生产和数据挖掘分析成为常态,在新闻领域也不能例外,新闻报道的功能和价值发生了新的变化。”

英国《卫报》是将大数据的逻辑与方法引入新闻传播实践领域的先驱者,也是成果卓著的探索者。本书的作者西蒙·罗杰斯更是世界公认的数据新闻领域的领军人。今天欣然看到他的专著能以如此专业和完善的形式呈现,真的是新闻传播的实践界乃至研究者们的一大幸事。他的一系列充满智慧的创意是我们在未来大数据与新闻传播“嫁接”中具有无限启发力的宝贵财富。

从他的书中我们可以看到,在功能层面,数据新闻超越了传统媒体的环境监测、社会协调、文化传承和娱乐等基本功能,基于数据挖掘和分析,让尘封的历史数据、混杂无序的数据重新焕发出新的活力。显然,数据新闻的核心价值就是从数据库到“数据闭环”。它要求我们:

一、 搭建高品质数据库

数据新闻是数据驱动型新闻,因此,必须要有专业的结构化数据库做支撑,才能保证新闻报道的客观性和真实性。数据库的来源可以是多种渠道:一是公共数据,主要是来自于政府网站或工作报告、企业网站、科研机构和专业调研机构的公开数据等权威渠道;二是媒体资料数据库,主要是媒体自身在长期的新闻报道中积累起来的新闻报道素材和数据

等信息；三是受众的个性化数据，主要来自于社交媒体和移动媒体，包括用户发布的内容及其所属类别、个性标签、社交关系、社交行为、加入的朋友圈和地理信息等等，这些都是个性化的多维数据，能够深入到以前新闻报道所无法抵达的行为分析、情感分析、心理分析和社会分析的深层。在未来，随着可穿戴设备和物联网的发展，随时捕捉传感数据将成为可能，这将对数据分析和应用提出更高挑战。基于这些渠道来源，我们可以建立高品质、结构化的专业数据库，为大数据新闻报道打下良好的基础。

以上提到的数据库分类中，知名的有官方数据门户 data.gov(美国)和 data.gov.uk(英国)等，受美国及英国政府启发，越来越多的国家都开设了数据门户网站，鼓励公众对政府数据的再利用。在这些官方的数据网站上，公众可以找到权威的、实时更新的最新索引信息。

除了官方数据库，还有一些社区或专门机构开设的开源数据网站或数据软件网站。比如 The Data Hub 就是一个由开放知识基金会运作的社区推动型 (community-driven) 数据资源网站；再比如在线工具 ScraperWiki，其目的是“简化有用数据的提取方式，使这些数据便于应用到其他应用程序，或者提供给记者和研究人员”。这些数据提取网站及其数据库大多数情况下都是公开的，可以重复使用。再比如世界银行和联合国数据门户网站，可以为所有国家提供多年积累的高水平的指标参数。^①

此外，互联网上还有很多公开的数据源或者是数据博客，可以供记者或者是爱好者们查询和下载数据。比如《纽约时报》数据图表编辑凯文·奎利 (Kevin Quealy) 创办的记录型博客“Charts ‘n’ Things”，收录了《纽约时报》大部分数据新闻的制作过程。一方面可以将之作为一个不完整的《纽约时报》新闻数据库，另一方面也可以将纽约时报的数据新闻编辑们面对庞大的数据是如何构思，如何筛选和如何进行数据可视化处理的作为参考。包括一些有价值的概念也可以在这里找到。“Charts ‘n’ Things”是最有价值的数据新闻博客之一。

再比如荷兰阿姆斯特丹大学教授、数据可视化工程师、多媒体记者巴斯·布鲁克赫伊曾 (Bas Broekhuizen) 建立的 Interactive-Inforgraphics 博客，通过分享互动可视化图表的理念、制作、理论在业界建立了一定的权威。Interactive-Inforgraphics 最大的特点是不仅有工具、技术的分享，同时也有可视化的理论和新鲜理念的交流。巴斯·布鲁克赫伊曾认为数据可视化是一个

^①《数据新闻手册（中文版）》，见 http://datajournalismhandbook.org/chinese/getting_data_0.html#sthash.nPcw0DLL.dpuf。

技术加艺术的实践，但最终是要服务于功能性的工具。清晰的展示数据背后的故事是每个可视化最基本，也是最高的要求。因此无论是学者、工具开发者还是数据新闻编辑，都可以在这里找到有用的素材。^①

二、提高记者的数据素养

作为数据新闻记者，一定要有过人的数据素养，具备敏锐的数据感觉和分析能力，在海量的数据库信息中，能够分辨出有价值的信息，挖掘背后的故事，拟成有价值的新闻选题，在此基础上，如果能够基于数据库和数据分析建立一个数据模型，就能够对同类信息进行监测，从而写出预测性的新闻报道，进一步提升新闻的价值。比如在2014年的英国大选中，本书作者西蒙·罗杰斯作为《卫报》“数据博客”的记者通过对Twitter上用户发表的内容和评论进行实时数据监测，把保守党、工党、绿党等五大政党的在不同地区的竞选活动热度变化制成了动态图，用来观测哪些是投票活动和关注的热门区域。

三、营造“数据闭环”驱动数据创新

在这个数据就是生产资料的时代，如何对数据进行循环利用也是数据挖掘与分析的一个重要议题。“数据闭环”指的是从数据的生成、采集、聚合、集成到数据挖掘和分析的过程中，由各种数据来源、数据采集方式、数据库、数据挖掘技术、数据模型和数据产品组成的一个环状数据运营体系。除了建设专业数据库和提高记者的数据素养，媒体本身乃至整个业界都应该具有利用数据、培养数据的意识，形成一个可持续运营的循环体系，这样才能促进数据创新，催生更多佳作。数据新闻的核心价值在于数据，而数据价值的挖掘不仅仅局限于一次性的价值挖掘，也可以是基于对数据的循环利用，数据库本身需要不断更新和完善，数据挖掘也是，因此，如果能够建立一个可再生的数据循环体系，采集数据、运营数据、创造数据产品和模型，然后再反过来，基于数据模型培养新的数据，同时完善旧的数据，以数据“养数据”，打造数据新闻循环生产体系，同时也会形成数据自然生长和循环利用的密闭型生态系统，使数据闭环“转起来”，这种基于生物逻辑的数据循环和生产方式，会赋予数据新闻持续的创新动力，从数据决策到数据挖掘皆是如此。（见图0-1）

^① 数据新闻博客汇总，见<http://djchina.org/2013/08/30/resource-blogs-2/>，2013-8-30。

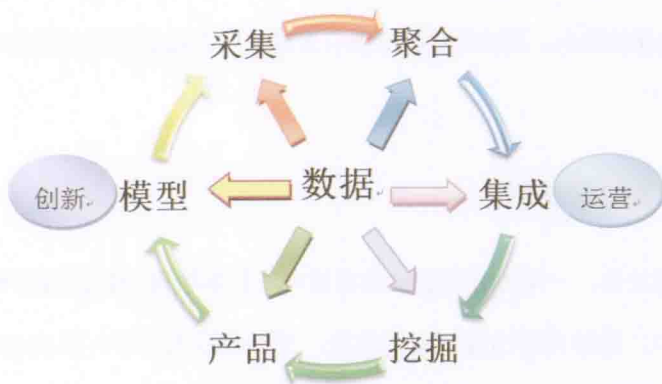


图 0—1 “数据新闻数据闭环” 流程图

比如 AVOS 公司基于用户行为数据分析推出的个性化新闻类阅读产品“美味爱读（Read Wise）”，就是基于对用户阅读数据和反馈信息的循环利用。AVOS 的目标是解决互联网信息碎片化、发现有效信息成本增加的问题。Delicious 是 AVOS 最核心的产品之一，它能帮助用户通过书签服务来梳理、检索自己在互联网中的所得。AVOS 中国团队的工程师在分析书签数据时发现，用户在阅读器的书签内留存并查阅知识、新闻和数据是最典型的使用场景，其中，新闻占了很大一部分比例。消费者感兴趣的是一些特定的话题，但为了能看到感兴趣的内容，要自己从各个频道进行筛选。所以美味爱读会跟踪文章在互联网上的传播情况，把引用数作为判断文章价值的重要标准之一。同时发布时间和文章受评论的状况也会影响它在阅读器里的排序。在此基础上，美味爱读又增加了“焚毁”和“评论”两个功能。“焚毁”就是让用户去除自己不喜欢的内容，系统通过收集这些用户反馈，可以做出更加个性化的推荐^①。长久以往，就形成了一个自动化运营的数据新闻“闭环”系统，完善个性化推荐的内容。

所有这些都在这本书中做出了具体生动的表述并赋以富有创意的实践模式。有鉴于此，特向本领域的实践者、研究者及关注者推荐。

喻国明

中国人民大学新闻学院教授

^① 陈燦然：《文本挖掘算法，热度识别体系：美味爱读是如何搭建个性化阅读架构的》，见<http://itjuzi.com/overview/news/405>。

推荐序二

数据新闻与数据素养

随着宽带无线、移动社交、智能终端时代的来临，数据已经成为当今社会运营和个人日常生活经营的一种核心资源。一方面，我们的衣食住行，每时每刻都通过依附于身上的移动或可穿戴设备，以及触手可及的公共设施，生产和提供着关于我们生活起居所思想方方面面的数据；另一方面，我们日常生活中大大小小的行为决策，小到饭后散步，大到孩子择校，都离不开有意无意的数据获取和消费。不论你是否愿意，我们都已经生活在数据时代；不论你是否相信，我们的生活很大程度上已经是数据驱动（data-driven）模式。一言以蔽之，数据的力量，数据对我们生活的影响力，也许已然超出我们自己的感知和想象了。1989年，时任美国劳工统计局（Bureau of Labor Statistics）统计专员和美国统计学会会长的著名统计学家珍妮特·诺伍德（Janet Norwood）在美国统计学会的演讲中曾说：“这是一个运行于数字之上的国家。”^①今天，我们之中的许多人也可以这样说：我是一个运行于数字之上的人。

因此，当我第一次看到西蒙·罗杰斯的这本《数据新闻大趋势》的原文和译稿样章时，竟然一点也不觉得奇怪：关于这样的主题和内容的一本书，如果不是西蒙·罗杰斯来写，也一定会有另外一个人来写——在这样一个数据革命的历史性时刻！当然，罗杰斯是这个主题——即数据新闻——非常理想的作者。

罗杰斯作为英国《卫报》著名的“数据博客”的编辑，可以说是主流新闻媒体中进行数据新闻尝试和实践的先行者。《卫报》自2009年推出“数据博客”，在新闻叙事和内容生产方面，探索和开创了一系列不同于传统新闻叙事（storytelling）的内容生产方式，在诸多方面颠覆了人们对新闻、新闻生产和新闻业的概念。罗杰斯正是《卫报》这一影响深远的新闻实践的核心推动者和一线实践者。书中大量例子皆来自卫报网站过去几年所做的经典案例。

如果我们回顾新闻学中的各种思潮，数据使用在新闻中的重要性很早就被精确新闻（Precision Journalism）、计算机辅助新闻（Computer-assisted Journalism）

^① 埃维森，格根. 统计学：基本概念和方法[M]. 吴喜之等，译. 北京：高等教育出版社，2000:3.

所强调，对数字的分析处理能力在美国也往往被认为是记者必备的基本功。正因如此，罗杰斯在本书第1章中开篇即强调，数据新闻学的提法虽然时髦，但其内涵和理念却早已存在。数据在新闻事实的构建中从来都有其独特的作用，书中第3章特别介绍了南丁格尔开创性的运用生动的可视化手段，来提高其论述和结论的说服力、视觉冲击力。尽管从复杂性和表现力而言，南丁格尔当年所采用的柱状图、饼图、扇形图无法与现在的信息图相比，但其中所体现的理念则是相通的，即内容生产者应该是冷冰冰的数据和活生生的内容消费者之间的沟通和连接之“桥梁（bridge）”。

信息传播方式和环境的演变与发展，已使我们置身于一个大数据时代。以往稀缺的数据资源现在已经变得过剩，因此，如何把真正有用的数据资源从海量的信息中提取出来并有效地呈现给受众将成为媒体从业者重要的核心能力，这也对新闻从业者在数据的使用上提出了更高的要求。《纽约时报》前些年发表了一系列文章，强调大数据时代“呼唤数据工程师”。在这样一种时代背景下，基于“数据驱动新闻”（Data Driven Journalism）逻辑的“数据新闻”概念的出现，也变得顺理成章，水到渠成。数据新闻强调了新闻生产中两个维度上的能力：一方面是对传统数据使用上的进一步深化，即能对公开的数据及数据库中的内容进行有效应用；另一方面则是对新闻文本中的数据进行合理的可视化呈现。“数据驱动新闻”也充分体现了全媒体时代新闻制作的特点，在该理念看来，新闻的呈现形式不应只有简单的文字和配图，而应是在传统表达方式基础上融入多媒体的可视化工具，并运用一定过滤和分析技术将大量数据融合到一个整体中^①，一如以《卫报》为代表的不少新闻媒介近年来所探索的那样。而近年来随着社会化媒体特别是各种社交媒体的兴起，以及各种各样便捷的信息图制作工具的出现，数据新闻实践大量体现在日常微博、微信等平台的内容中，生产者既有专业记者，也包括公民记者和普通民众，不一而足。

由于数据的统计分析和可视化处理在数据新闻生产中的高显示度和重要性——据罗杰斯估计，就所花的时间而言，统计分析和信息图制作可能会占整个数据新闻作品生产时间的80%，这就意味着，有时人们会过分重视数据本身的分析和呈现，而忽略作品整体叙事逻辑。对此，罗杰斯警告说，成功的数据新闻作品，关键是要以一个记者的视角而不是一个数据分析师的角度来看待各种数据。至少在数据新闻的范畴内，

^① 祝建华：《什么是“数据驱动新闻”》，见人民网<http://media.people.com.cn/n/2012/0718/c120837-18543914.html>，2012-07-18。

传播效果仍是统摄一切取舍的终极依据。数据入新闻，仍需不忘新闻选题与立意之初心。毕竟，我们不是为分析数据而分析数据，也不是单纯为了呈现数据而制作各种图表。数据新闻的最终也是最根本的诉求，仍然是讲好故事，让真相更明了。因为唯有“事实和真相才是神圣的”（Facts are sacred），而数据只有服务于事实的传播，才能真正显出其力量。

由此，我们可以看到，作为新闻内容的生产者，新闻记者在传统的记者素养之外，还需要与时俱进的提升自身的“数据素养”^①。何谓数据素养？就新闻记者而言，至少应包括对数据的收集、分析、呈现和有效评估的能力。数据的收集能力是记者的看家本领；数据的分析和评估，则强调了记者除了人文科学的训练和素养，还需要具备一个社会科学学者的基本能力和素养；而数据的呈现，则强调记者除了人文素养和社科素养，还必须有一定的艺术素养和审美设计能力。数据素养由此成为数据新闻时代记者素养的新维度。

在一篇讨论机器新闻写作的文章^②中，笔者就新闻业面临的挑战有以下这样的观察：

自动新闻生产的不断改进和复杂化，将不断挤压新闻业中那些劳动密集型的工作或新闻生产环节，让这些工作或生产环节，就工作量和所花费的时间而言，主要通过机器即可完成，人工的干预和关键决策，恰如画龙点睛一般，本质上仍会起主导作用，但从工作量上而言，不再需要在创造性要求高度不均的各个环节都同等付出，而只需在一些关键环节上介入即可。记者和编辑，由此需要重新定位自己工作的核心价值到底在哪里，时时考虑自己所从事的工作，是否具有智能的软件或机器人也能胜任。新闻从业者的不可替代性，从长远看，必然来自其个性化特色的内容创作和创新编排、设计、内容传播和推送策略的制定等，而不是其他。

数据新闻也好，机器新闻也好，新闻业的确正在经历前所未有的变革。在这样的变革大潮中，包括新闻记者在内的专业内容生产者如何安身立命，的确是一个需要不断自我反思的问题和命题。这样的思考，也许可以从阅读这本《数据新闻大趋势：释放可视化报道的力量》开始。

金兼斌

清华大学新闻与传播学院教授

① 金兼斌（2013）. 财经记者与数据素养. 新闻与写作, 2013（10）: 5-9.

② 金兼斌（2014）. 机器新闻写作：一场正在发生的革命. 新闻与写作, 2014年第9期, 30-35.

推荐序三

序

如果你爱数据新闻，请读本书，学习先行者《卫报》的操作实例。

如果你关注可视化，请读本书，了解可视化在新闻领域的运用。

如果你在媒体工作，请读本书，知晓全新的新闻生产方式和表现手法。

如果以上都与你无关，你只是好奇心重，也请读本书，认识新闻、技术与艺术如何跨界结合。

可视化无所不在。地图是可视化，到医院照CT是可视化，掏出手机看Wifi信号有几格还是可视化。

可视化分为科学可视化、信息可视化和可视化分析，可视化与媒体结合的产物是数据新闻。

印在纸面的数据新闻叫信息图，用网页呈现的数据新闻叫互动图表。信息图历史悠久，超过百年。互动图表受到互联网技术的限制，诞生至今大概五年。

数据新闻一直在发展，没有准确的定义。按我的标准，用数据可视化的手段呈现的新闻才叫数据新闻。按本书描述的《卫报》标准，从数据挖掘出内容，再用纯文字表现也是数据新闻。怎么定义不重要，怎么操作才重要。

信息图由编辑和设计师合作，互动图表由编辑、设计师和程序员合作。二者的区别，在于后者必须有程序员介入。

一个完整的数据新闻团队必须有编辑、设计师和程序员三种人。因为制作数据新闻，财新网(caixin.com)的程序员拿了新闻奖，这在中国新闻史上是第一次。

上面这几句，是我学习和工作的总结。你很容易记住，却不见得能领会，因为你可能缺乏实践。反复地学习、分析、实践才能进步。

从一无所知，到成为所谓的“专家”、“大咖”，我大概用了一年半的时间，这期间财新数据可视化实验室从零起步，拿了好几个奖。我对成长过程帮助指导过我的人深怀感激，也乐

于分享知识提供帮助。

vislab.caixin.com 是财新数据可视化实验室的官网，上面除了实验室的介绍和作品，还有一个资源库，收集诸多学习资料和可视化工具，欢迎大家利用。

好奇是学习的动力，学习是成长的基础。与各位共勉。

黄志敏

财新传媒 CTO， 财新数据可视化实验室负责人

中文版序

中国，在西方的数据新闻记者眼里，往往可以作为一个不错的选题，但不会是数据来源地。以中国为主题的数据新闻报道并不少见，比如 Propublica 上就有关于国际新闻媒体在中国的传播效果分析 (<https://projects.propublica.org/firewall>)，比如路透社曾经获得过数据新闻奖的作品中国政局指南 (<http://china.fathom.info>) 等等。

写这篇中文版序言，我非常高兴，不仅因为这是本书在中国的首次亮相，更是因为有越来越多的中国记者也开始尝试数据新闻报道。迸发着勃勃生机的中国，让人由衷地兴奋不已。她会是数据新闻发展的沃土吗？中国的数据新闻有可能成为世界上最有影响力的数据新闻吗？

一段时间以来，中国的媒体机构和自媒体也开始在数据新闻领域小试牛刀。其中不得不提到一个人物——马金馨，她的自媒体博客 (<http://yolandama.wordpress.com/>) 专注于东亚地区数据新闻的现状研究。数据新闻理念在中国的传播和发展，她功不可没。

当欧洲和美国的新闻从业者专注于本土的新型数据新闻网站建设时，在中国，不少有意思、受欢迎的数据新闻站点也如雨后春笋般涌现。它们大多是定期发布的数据新闻，在报道方式上，有些已经能够和西方老牌媒体的作品相媲美。比如网易数读、财新数字说、搜狐数字之道，等等。这仅仅是管中窥豹，整个中国的数据新闻正变得和美国、英国的同行们一样，有百家争鸣之势。

以前，《卫报》的“数据博客”(Datablog) 不论是发布日常数据分析还是做数据报道，都是孤军奋战。而现在，数据报道已经是主流报道的一部分，和全世界都在用的互联网、电子邮箱一样寻常。

在当下这个信息图泛滥、数据报道成堆的传媒时代，如何才能脱颖而出？于我看来，

应该注意以下两个方面，它们是相辅相成的。

首先，贴近性。无论是何种类型的新闻，只有叙事清晰、简单易懂，才能达到最大最优的传播效果。与其做最聪明的人，还不如做一个人人都能懂你的人。数据新闻记者很容易陷入方法论和工具的泥潭而本末倒置，往往会忘记我们真正要做的是新闻。数据是报道的关键要素，但不是决定性的因素。为什么不让统计学家统统转行去做新闻记者？这是因为，新闻报道是你的工作，是你记者的职责，你要成为更优秀的记者，而不是更优秀的统计学家。

其次，开放性也至关重要。政府部门的数据发布活动往往是不公开、不透明的，所以我们在使用数据时也会遇到类似的“惰性状态”。当数据新闻记者和团队在做报道时、在赋予数据生命时，这种不开放的状态需要改变。

与此同时，作为一名数据新闻记者，寻找数据、清洗数据、让数据能够讲故事，这些当然是首要任务。不过，还有一项工作常常会被忽略，那就是发布数据、公开数据。数据“来之于民”，也理应“用之于民”，提供数据是为了让他人可重复利用。

在我看来，数据新闻的未来，可能是这样一幅图景：封闭的传统新闻业与开放的新兴新闻业相互抗衡。在这个领域，我们会积极与受众共享数据、信息，寄希望于受众能帮助我们共同做出更好的新闻。

中国的数据新闻记者现在可以选择继续致力于新兴的数据报道方式，这不仅有可能改变自己的国家，还有可能改变世界。正如我在本书中写道，只要会使用合适的工具、只要发挥潜能，人人都能做数据新闻。照这样的逻辑，谁都有可能成为改变世界的一份子，而且是永久的改变。我也迫不及待地期待那一天早日到来。

西蒙·罗杰斯

序 言

数据，不仅是信息时代的新产物，也是数字工业、金融业和商业变革的核心。就其精髓而言，它更像是真相与事实的集大成者。

有着190年历史的《卫报》，一直以来都崇尚数据。无论是《曼彻斯特卫报》^①创刊号头版上的表格，还是大名鼎鼎的总编C. P. 斯科特^②的那句传世名言——“事实是神圣的”（Facts are sacred，原版书的书名正源于此），它们都是最好的注脚。

如今，《卫报》的“数据博客”^③正在以现代的姿态继续拥抱数据，本书就收录了其中的一些扛鼎之作。从图解经济现象和政府支出，到阐释新闻背景资料，再到深度分析像“2011年英国骚乱”这样的重大事件，数据已然成为《卫报》新闻报道中必不可少的一部分。

和做数据新闻同样重要的是：共享我们收集而来的数据，告诉大家我们的做法，同时还要让读者参与到分析数据和改进报道的过程中来。这些日趋重要的工作，在2009年“数据博客”上线后就成了我们工作的核心要义。本书记录了这几年我们为此付出的努力和汗水，希望能对你有所帮助。

艾伦·拉斯布里杰（Alan Rusbridger）^④

《卫报》总编辑

①《卫报》（*Guardian*）创刊于1821年，因总部设于英国曼彻斯特而称为《曼彻斯特卫报》（*The Manchester Guardian*），1959年更名为《卫报》，总部于1964年迁至伦敦。——译者注

②查尔斯·普雷斯特维奇·斯科特（Charles Prestwich Scott，1846～1932），《卫报》历史上最著名的总编辑，从1872年开始，他主持了57年的编务工作。——译者注

③“数据博客”（Datablog，www.guardian.co.uk/news/datablog）是《卫报》的数据新闻频道，于2009年1月14日上线，其作品在网站和报纸都有刊发。作为较早实践数据新闻的先驱，“数据博客”开创了自己的风格，《卫报》也被视为该领域的标杆式媒体。本书中所有的报道案例和数据可视化作品均来自“数据博客”。——译者注

④艾伦·拉斯布里杰于1995年出任《卫报》总编。2014年12月，他宣布将于2015年夏天卸任该职位。——编者注

目 录

附录	307	引言	1
译者后记	304	数据无处不在	21
怎样做数据新闻	277	追根溯源	51
新朋克	263	案例分析： 维基解密	61
案例分析： 2011年伦敦骚乱	237	解读财政开支	85
众包	221	欢迎提出修改意见	124
数据的世界	201	点“数”成金	135
案例分析： 2012年伦敦奥运会	169	衡量幸福	149

引言	数据新闻记者的十条真理
数据无处不在	互联网时代如何发现新闻选题
追根溯源	数据可视化的起源
案例分析：维基解密	把电子表格变成独家新闻
解读财政开支	政府是怎样花掉纳税人的每一分钱的
欢迎提出修改意见	请读者帮忙提升新闻报道效果
点“数”成金	为什么大数据意味着大商机
衡量幸福	除GDP外，还有哪些更好的衡量指标
案例分析：2012年伦敦奥运会	数说奥运会与残奥会
数据的世界	开放数据运动的国家样本
众包	用群体的力量挖掘数据背后的新闻真相
案例分析：2011年伦敦骚乱	“解读骚乱”系列报道
新朋克	谁都能当数据新闻记者
怎样做数据新闻……	你也能做到