

宏观质量管理学术丛书

ChanPin ZhiLiang AnQuan WangLuo
XinXi JianCe Yu YuJing YanJiu

产品质量安全网络信息
监测与预警研究：
基于互联网信息的实证分析

程虹 许伟 ● 著

中国社会科学出版社

宏观质量管理学术丛书

ChanPin ZhiLiang AnQuan WangLuo
XinXi JianCe Yu YuJing YanJiu

产品质量安全网络信息
监测与预警研究：
基于互联网信息的实证分析

程虹 许伟·著

中国社会科学出版社

图书在版编目(CIP)数据

产品质量安全网络信息监测与预警研究: 基于互联网信息的实证分析 / 程虹, 许伟著. —北京: 中国社会科学出版社, 2014. 12

ISBN 978 - 7 - 5161 - 5406 - 9

I. ①产… II. ①程…②许… III. ①互联网络—应用—
产品质量—质量管理—研究 IV. ①F273.2 - 39

中国版本图书馆 CIP 数据核字(2014)第 308390 号

出 版 人 赵剑英
责任编辑 田 文
特约编辑 吴连生
责任校对 张爱华
责任印制 王 超

出 版 中国社会科学出版社
社 址 北京鼓楼西大街甲 158 号
邮 编 100720
网 址 <http://www.csspw.cn>
发 行 部 010 - 84083685
门 市 部 010 - 84029450
经 销 新华书店及其他书店

印刷装订 三河市君旺印务有限公司
版 次 2014 年 12 月第 1 版
印 次 2014 年 12 月第 1 次印刷

开 本 710 × 1000 1/16
印 张 13.75
插 页 2
字 数 233 千字
定 价 43.00 元

凡购买中国社会科学出版社图书, 如有质量问题请与本社联系调换

电话: 010 - 84083683

版权所有 侵权必究

《宏观质量管理学术丛书》

编委会

主 编 程 虹

编 委 (以姓氏拼音排序)

陈 川	陈昕洲	邓 悦	范寒冰	范锐敏
贾国栋	李丹丹	李 酣	李艳红	刘三江
刘 芸	罗连发	罗 英	宋 琼	宋时磊
陶 军	王 虹	汪晓清	许 伟	余 凡
余红伟	袁晓峰	张继宏		

跨学科的前沿探索

——《宏观质量管理学术丛书》总序

无论宏观质量管理的定义还需要怎样的精确界定，也无论宏观质量管理的学科体系还有待怎样的科学凝练，宏观质量作为一种客观现象的存在是毋庸置疑的，而对宏观质量的管理更是一个显而易见的问题。科学最重要的是发现“问题”，这些问题既是新的，背后的规律又是有待人们去探索的。宏观质量以及对宏观质量的管理，就是一个新的问题，也是亟须从科学上探索其内在规律的问题。《宏观质量管理学术丛书》的定位，就是立足于探索宏观质量现象背后的科学规律，并在此基础上提出科学的质量管理和质量治理的制度、方法与工具。

现代科学的发展趋势就是要研究“问题”，同时对问题的研究则是采用跨学科的研究方法。宏观质量管理本身就是一个复杂的问题，需要从管理科学、经济科学、信息科学、工程科学和人文科学的跨学科视野展开综合研究，只有这样的跨学科研究才能识别和认识宏观质量管理背后的一般规律。《宏观质量管理学术丛书》将从跨学科的角度，推出系列的学术专著，包括微观产品质量与宏观经济增长质量研究、中国特色质量管理体系机制变迁研究、中国质量史研究、区域质量评价与经济性影响因素研究、质量主体博弈研究、质量责任体系研究、比较实验对中国质量治理体系优化研究、中国质量服务机构发展研究、标准作为一种市场秩序的理论研究、大数据管理下的质量创新研究、电子商务中质量信息分析研究、网络质量语义研究、基于网络的质量安全监测与预警方法研究等。这些系列著作都将从跨学科的角度

度，研究和探索宏观质量管理这一问题。

《宏观质量管理学术丛书》的学术追求是前沿与创新，将整合应用跨学科的前沿理论，以实证为主要研究方法，立足于大数据的应用，着力于对问题背后一般规律的探索，力求提出有创新性的理论观点和科学方法。

2

《宏观质量管理学术丛书》由武汉大学质量发展战略研究院、宏观质量管理湖北省协同创新中心组织编辑出版，是武汉大学质量发展战略研究院学术同仁科研成果的汇聚，将集中反映武汉大学质量发展战略研究院这一学术共同体的学科范式 and 科学追求。

程 虹

《宏观质量管理学术丛书》主编

2014 年 12 月于武汉大学樱顶

本书是国家社科基金重大项目(11&ZD158)的研究成果, 并受科技部公益性行业科研专项(201210117, 201310202), “十二五”国家科技支撑计划项目(2011BAK06B06), 中央高校基本科研业务费专项资金(2014115010202), “宏观质量管理湖北省协同创新中心”的资助

目 录

第一章 产品质量网络监测概述	(1)
第一节 研究背景	(1)
一 质量网络监测概念	(1)
二 质量进入大数据时代	(4)
三 总体质量信息不对称	(8)
四 质量信息的主要来源——消费者	(12)
第二节 相关文献回顾	(17)
一 质量的含义	(17)
二 关于质量信息源的研究	(19)
三 关于质量评价理论与方法的研究	(25)
四 关于质量安全预警的研究	(32)
五 关于质量信息搜集的研究	(37)
第三节 基于消费者的质量网络监测基本理论与方法	(39)
一 质量安全网络监测的指标	(39)
二 研究方法和模型建构	(42)
三 模型应用与解释	(45)
四 方法的进一步讨论	(51)
第二章 乳制品行业质量网络监测报告	(54)
第一节 乳制品行业基本信息	(54)
一 乳制品年度质量风险状况基本分析	(54)
二 六个维度质量风险分析	(63)
第二节 重点乳制品企业质量网络监测报告	(70)
一 Milk 1 年度质量分析报告	(70)

第三节	乳制品热点事件	(76)
一	热点事件 1: 圣元乳业“事件”圆满解决	(76)
二	热点事件 6: “三聚氰胺”奶粉事件	(81)
第三章	汽车行业质量网络监测报告	(89)
第一节	汽车行业基本信息	(89)
一	汽车总体信息计量	(89)
二	汽车品牌六大维度分析	(91)
三	合资品牌与自主品牌的维度分析	(98)
第二节	重点汽车企业质量网络监测报告	(100)
一	汽车企业 CAR 1 年度质量分析报告	(100)
二	汽车企业 CAR 2 质量分析报告	(107)
第三节	汽车行业热点事件	(117)
一	热点事件 2: 大众 DSG 变速箱之困事件	(117)
第四章	质量监测的语料库建设	(124)
第一节	质量语料库建设总体思路	(124)
一	质量语料库建设目标	(124)
二	质量语料库管理规范	(125)
三	语料库管理系统	(129)
第二节	质量语料库分类	(132)
一	质量语料库分类设计	(132)
二	企业通用风险分类标准及语料库	(133)
三	行业风险分类标准及语料库	(135)
四	区域政府通用风险分类标准及语料库	(140)
第五章	政府质量安全评价与网络预警的实现	
——	网络平台的构建	(146)
第一节	平台总体建设思路	(146)
一	平台建设目标	(146)
二	平台服务模式及其价值	(153)

三 核心技术	(157)
第二节 平台系统分析与设计	(158)
一 系统总体结构	(158)
二 平台子系统组成	(160)
三 系统数据结构设计	(163)
四 平台测试与实施	(167)
第三节 质量安全分类算法与效能测试	(170)
一 分类算法描述	(170)
二 算法验证与测试	(177)
 第六章 一般性结论和启示	(182)
第一节 乳制品行业质量监管需要合作治理	(182)
一 特征事实	(182)
二 理论分析	(185)
三 政策建议	(186)
第二节 汽车市场质量监测分析	(190)
一 特征事实	(190)
二 理论分析	(192)
三 政策建议	(194)
第三节 质量舆情的有效管理	(195)
一 特征事实:网络质量舆情影响越来越大	(195)
二 理论分析	(196)
三 政策建议	(197)
 参考文献	(202)

第一章 产品质量网络监测概述

第一节 研究背景

一 质量网络监测概念

质量网络监测，是武汉大学质量发展战略研究院基于“质量安全信息网络监测与预警服务平台”，通过综合利用质量风险理论、价值工程理论，结合搜索技术、云计算、数据挖掘、语义分析、质量安全评价体系等新一代网络信息技术的集成创新，实时收集来自公共互联网的海量消费者感知的质量安全信息，经过深度的挖掘、提炼、分析、处理，及时向广大企业、消费者和政府提供富有价值的质量安全信息服务。在当前这样一个“大数据”时代，质量数据的爆炸性增长极大地增加了质量安全的不确定性，并对质量安全治理的方法与技术提出了更高的要求，基于消费者的海量互联网信息的质量网络监测，为产品质量安全有效网络治理提供了一种行之有效的方法和技术。

本研究依托于“质量安全信息网络监测与预警服务平台”监测到的质量数据进行研究分析，选取了国内乳制品和汽车两个行业，共计 34 家企业、58 个品牌和 284 个产品，其中包括 55 个乳制品产品和 229 个汽车产品的网络质量，共获得有效数据 390223 条。研究主要涉及 8 家乳制品企业，分别是蒙牛、伊利、光明、完达山、三元、南山、雀巢、圣元等企业；涉及 26 家汽车企业，分别是通用、福特、大众、宝马、奔驰、PSA、本田、丰田、日产、马自达、现代、比亚迪、奇瑞、吉利、上汽、长城、东风、铃木、一汽、力帆、东南、长安、江淮、海马、众泰、中华等企业。通过对汽车和乳制品行业进行深度分析，得出了一些有价值的观点。

希望我们此次研究能得到政府和相关企业的重视,为我国的汽车、乳制品行业的健康发展作出贡献,为中国企业的质量发展提出建设性建议。

质量安全问题存在的根本原因,在于质量信息的不对称性。在互联网兴起之前,能够产生和传播企业质量安全信息的主体,除了企业自身外,就是政府、媒体和专业从事质量服务的第三方机构。作为消费者要想向社会传播自身使用产品和对企业质量评价的感受,一般只能向以上的主体提供。进入互联网时代后,作为网民的消费者都可以通过互联网,直接向社会传播对产品和服务质量的评价信息^①,特别是随着“微博”等“自媒体”传播方式的出现,每一个消费者几乎都可以成为一个独立的媒体,导致企业质量安全信息的传播,表现出极强的即时性和互动性^②。根据中国互联网络信息中心(CNNIC)公布的第31次《中国互联网络发展状况统计报告》的数据,截至2012年12月底,我国互联网网民已达5.64亿人,全年共计新增网民5090万人。互联网普及率为42.1%,较2011年年底提升3.8个百分点,普及率的增长幅度相比上年继续缩小。手机网民规模在2012年增长迅速,并于年中超越使用台式电脑接入互联网的网民。手机上网的发展为网络接入、终端获取受到限制的人群和地区提供了使用互联网的可能性,包括偏远农村地区居民、农村进城务工人员、低学历低收入群体^③。网民中40岁以上各年龄段人群占比率均有不同程度的提升,在这些用户中仅微博用户人数就达到3.01亿(CNNIC, 2012),成为事实上的信息发布者。在移动互联网和移动智能终端快速发展的背景下,以上的用户,更是成为企业质量安全信息全天候的潜在生产者 and 传播者。

在非互联网的条件下,要获得企业质量安全状态的信息,虽然也可以通过消费者投诉、调查等方式,但主要还是依赖于对企业所生产产品的检验或认证,以及通过对企业经营的各类质量信用信息的获取。产品检验的局限性在于,基于成本和产品上市时间的要求,固定抽样的比例只能是产品总体数量的极少部分,不可能反映每一件产品的质量状态,更不可能全面反映生产各种不同类型产品企业的质量状态。对企业质量信用信息的分

① 程虹:《2012年中国质量状况》,《宏观质量研究》2013年第1期。

② 程虹、范寒冰、肖宇:《企业质量安全风险有效治理的理论框架——基于互联网信息的企业质量安全分类模型及实现方法》,《管理世界》2012年第12期。

③ 苏璞睿、冯登国:《面向国家战略需求构建信息安全体系》,《高科技与产业化》2013年第2期。

析,因为企业数量的庞大、机会主义倾向,以及信息的不对称,也很难低成本地获取来自众多企业的质量信息。面向企业的产品质量评价,最终的只能是来自于消费者使用后的感知。原因在于,消费者会因为使用产品所感知到的确定性后果和状态,准确地表达对企业质量安全的评价。消费者对风险更为敏感,作出购买决策时倾向于减少其感知风险而不是最大化其感知价值(Mitchell, 1999),改变、推迟或取消购买决策在很大程度上是受到感知风险的影响(Kotler, 1991)。因而,相较于来自产品检验和企业质量信用记录的信息而言,来自于消费者的使用信息,对于企业的质量安全评价,无论是在信息的数量上,还是在信息所反映的质量安全的最终状态上,都更为广泛,也更为准确^①。

面向消费者获取企业的质量安全信息^②,有多种渠道和载体,包括企业自身的用户信息和投诉服务系统、社会组织的消费者投诉系统、不同媒体的消费者投诉信息以及政府部门接受的消费者投诉信息等。但是,这些传统的获取消费者质量信息的方式面临的问题在于,既不能获取最大数量的消费者信息,又不能实时地得到这些消费者的信息,从而导致不能全面地评价企业的质量安全状态。而在互联网的条件下,既可以获取海量的消费者质量安全信息,又可以实时地得到这些信息,从而为企业质量安全的评价,提供了来自于消费者的最为广泛和实时的信息来源。

传统的质量安全监测和监管模式是通过企业内部的售后服务、投诉渠道、热线电话、市场调查、媒体以及政府监管部门的检验检测、抽查、执法等途径,获得产品的质量安全信息,然后责令企业回应、妥善处理或停业检查整顿。虽然消费者与企业 and 政府监管部门的质量安全信息互动早就存在,但仍存在着极大的信息不对称、信息数量有限、信息传递滞后、信息收集困难,难以智能化处理和挖掘利用,获得质量安全的规律性、趋势性判断。

本研究通过深度网质量安全网络信息监测与预警平台,来获取需要分析的8个乳制品企业以及26个汽车企业的质量信息。这些信息主要来源于微博、论坛、投诉平台和博客等全国知名的自媒体网络,覆盖了中国网络媒体Alex排名前2000位的90%以上。数据的获取方式是采用网络爬虫

^① 程虹、范寒冰、肖宇:《企业质量安全风险有效治理的理论框架——基于互联网信息的企业质量安全分类模型及实现方法》,《管理世界》2012年第12期。

^② 程虹:《2012年中国质量状况》,《宏观质量研究》2013年第1期。

技术和元搜索技术：一种是通过指定范围的网站对其进行抓取采集，支持新闻、博客、论坛和微博采集，比如人民网、央视网、中国网、新浪、腾讯网、搜狐、网易等网站；另一种是通过百度、谷歌、奇虎 360、搜狗等搜索引擎根据关键词进行全网的数据采集。通过对上述的网络媒体信息源进行不间断的网络获取，依据企业名称和产品名称对已经采集获取的数据进行过滤，获得与企业相关的所有数据，并通过人工智能自动判别的方式，将所有涉及企业质量风险的数据进行标注，构建相应的分类信息源。

二 质量进入大数据时代

现在全世界的学术界、新闻界、商界，包括政府都在讨论大数据时代，而牛津大学教授的维克托·迈尔-舍恩伯格和肯尼思所写的《大数据时代》与《删除：大数据取舍之道》是其中的代表性著作。所谓“大数据”，是指其规模超出了现有采集、储存和分析的数据总量，是一个定性的描述方法，意指人类现在所面临的数据量、数据增长和数据类型，超出了以往的规模，以其“大”的定义，告诉我们人类已经进入了一个数据无处不在、数据爆发性增长的新时代^①。21 世纪是数据信息大发展的时代，移动互联、社交网络、电子商务等极大拓展了互联网的边界和应用范围，各种数据正在迅速膨胀并变大。在 2006 年，个人用户才刚刚迈进 TB 时代，全球一共新产生了约 180EB 的数据；在 2011 年，这个数字达到了 1.8ZB。而有市场研究机构预测：到 2020 年，整个世界的的数据总量将会增长 44 倍，达到 35.2ZB（1ZB = 10 亿 TB）。大数据有四个特征，即大量化（Volume）、多样化（Variety）、快速化（Velocity）、价值密度低（Value）。在大数据时代，人类仅凭所拥有的数据总量，就可以更加清楚地判断出事物之间的关系。比如，美国国家邮政局在邮政车上，都安装了传感器的装置，它们能在将邮件送到每个社区时，实时地采集该社区的噪声、空气质量指数，如此一来，科学家能更加准确地了解到各个不同区域的空气、噪声等环境质量的精确指标。

在大数据时代，人类突破了以前数据采集的困难，不再仅依靠抽样的小样本数据，而是拥有了一个事物的全数据。美国利用穿梭在各地的车辆，实时地收集各类天气的数据，使得当地天气预报的准确性大大提高。

^① [英] 维克托·迈尔-舍恩伯格、肯尼思·库克耶：《大数据时代》，盛杨燕、周涛译，浙江人民出版社 2013 年版。

在大数据时代,人类可以搜集各类不同的数据信息,不再追求单一数据的精确性,而是依靠更大的数据总量,来支撑对事物的判断。在美国,人们在手机上输入某个产品的名称,就可以得到关于该产品被召回、被消费者投诉和媒体评价的信息,从而更加理性地消费。我们说,大数据时代所带来的绝不仅仅是数据量的大小,而是将带来人类思维方式、工作方式和决策方式的变革^①。国际著名的咨询机构麦肯锡公司被称为“大数据:下一个创新、竞争和生产力的前沿”,美国政府在2012年发布了《美国联邦政府大数据研发计划》,明确表示要抢占下一个比石油、煤炭更为重要的资源。

关于数据对质量的价值,美国著名的管理学家、统计学家戴明曾经说过,“除了上帝,任何人都必须用数据来说话”,戴明作为享誉世界的集统计学家、管理学家和质量管理学家这三个专家的称谓于一身,本身就表明了质量是以数据的统计为基础的结论^②。无论是质量固有特性的测量,还是顾客满意的评价,实际上都来自于数据的表达,可以说没有数据,就没有质量,任何质量都是基于数据的评价和统计。

在人类进入大数据时代的背景下,质量是否也进入到了大数据时代呢?我们非常有必要对这个问题,作一番认真的分析。

第一个分析当然是大数据时代最重要的条件,那就是互联网的普及。实际上,这几年来无论是中国,还是世界其他国家,质量问题之所以如此被人们广泛关注,又如此让大家感到不满,其中一个重要的原因就是互联网的普及。在互联网上,哪怕是一个非常小的国家的某个普通消费者,只要他在网络上发布了一条不好的质量安全信息,瞬间就可能会被转发到全世界的每一个角落,并让消费者人人自危。互联网所带来的质量数据,是我们以往任何一种方法都不可企及的,因为每天每一个消费者都可以随意地、实时地在互联网上发布他感到不满的质量问题的信息,这些数据累积起来就是大数据。所以,互联网的普及,毫无疑问地把我们拉入到了质量的大数据时代。

第二个分析就是数据获取技术的进步和普及。大数据时代的很多数据不是来自于人本身,而是来自于物体本身,如很多发达国家通过各种传感器,使得车辆、机器和物品都能传递物体的信息,这一点在我国也越来越

① 涂子沛:《大数据:正在到来的数据革命》,广西师范大学出版社2012年版。

② 程虹:《中国质量怎么了》,湖北科学技术出版社2013年版。

普及，比如在很多人的车辆上所装载的 GPS 系统，都在时刻地传递有关这辆车的各类信息。再比如，在重庆、北京，通过在电梯上安装传感器，在监控中心能够实时地获取电梯运行的各类数据。如今，这些数据获取技术的成本已经越来越低，比如，射频技术（RFID）已经便宜到可以在一个鸡蛋上应用，使得各大型商场对于每一件商品的来龙去脉都清清楚楚。传感器和射频技术的普及，使得我们对物体的运行性能、质量安全状态有了更为方便的数据获得方法，这在环境质量、工程质量中也将会得到越来越广泛的应用，目前，很多桥梁上就安装了这样的传感器，监控中心能实时地获取这些桥梁承载的安全指标。

第三个分析就是信息工具的个人化。以前只有专业机构用得起的信息采集仪器，现在已经被个人所广泛使用，比如便携式的血压、心脏和其他人体器官的监测仪器，使人们能够实时地获得自身健康的相关指标。更不要说现在已经普及几百块钱一个的智能手机，使人们获取视频影像信息和声音数据信息，变得非常便捷和快速。网上的“表哥”“微笑哥”等官员，就是因为网民的一张手表图片、一张微笑的面部表情图片，而招来了“灭顶之灾”。实际上，现在国内外有很多消费者，每天通过个人简单的手持信息设备，在网上发布大量的有关产品、服务、工程 and 环境的图片、声音和视频的质量数据。有些企业也开始通过这种低成本的信息工具，在网上实时地公开自己生产现场的视频，让每个消费者能在网上看到产品生产的全过程。实际上，政府、企业每天也在应用这些新的信息工具，产生大量的质量监管和质量管理的数量。以上所分析的这些大量的数据，都是在传统的工业时代所不可能大规模拥有的，更为重要的是，这些数据几乎以秒为单位在快速增加。从这方面看，我们说质量确实已经进入了大数据时代。

接下来，我们需要研究的是，质量进入大数据时代以后，质量的大数据有哪些主要的类型呢？第一类，是来自于消费者在使用中的质量数据，这是质量大数据的主要类型；第二类，是来自于物体的质量数据，包括产品的质量数据、检验机构的质量数据；第三类，就是来自于政府在质量监管业务中所产生的质量数据。当然还有一些其他类型的数据，但主要是这三类。这三类数据分别对应了质量的三个不同主体，在使用中，质量数据对应的是消费者主体，物体的质量数据主要对应的是企业主体，监管业务的质量数据对应的是政府主体。分析和清晰地划分大数据的类型有着特别

重要的意义，大数据就意味着各类数据之间要相互兼容，要有更多不同类型数据的呈现，才能找到质量之间的相关性。比如，政府质量监管业务中所产生的业务数据具有抽样样本的代表性，但有时由于样本量太小，并不容易得出对这个行业产品质量风险的判断，如果加上消费者所提供的大量质量数据，就较容易得出某一个产品领域的风险状况，从而作出更为科学的监管决策。目前的问题是，我们的质量大数据已经出现，但是相互之间却是隔绝的，不能真正地发挥质量大数据的应有价值，这就需要政府监管部门创新思维，开门搞监管，开门获取其他类型的质量大数据，而不是对那些小作坊式的数据都要亲力亲为，一谈到需要其他类型的质量大数据，就开始自己搞投入，建机构。实际上，大数据时代的质量监管，就意味着监管部门必须与社会其他主体，共同分享质量的数据。

质量大数据的价值，并不在于这些数据本身，而在于通过对这些数据的分析而得到应用，这才是数据能成为资源的原因。武汉大学质量院作为中国知名高校中专业的质量研究和教学机构，近几年来一直在做质量数据分析的研究，为此专门建了两个平台：一个是中国质量观测基地；另一个是中国质量网络信息监测与预警平台。前者主要是通过线下的消费者实地调查，获得有关产品、服务、环境和工程的质量满意度、质量安全、政府质量公共服务和消费者质量素质等方面的数据，这些数据的获得是基于连续性的观测；后者主要通过互联网的语义分析技术、数据挖掘技术和智能决策支持技术，实时地监测和获取消费者有关质量的各类信息。该平台目前获取的有效质量数据总量已达到 500 多万条，而且还在以分钟为单位不断增加。我们之所以花这么多投入，包括资金、人力和时间，主要不是为了获得数据本身，而是要通过这些数据来找出不同地区、不同领域、不同行业的质量状态的普遍特征，尤其是要找出质量安全风险演变的一些规律性的东西，看某一类型的产品是哪些企业、在什么时段更容易出现质量安全问题，找出问题背后的规律。比如，私人轿车因为质量而出现安全问题，到底在一天内的哪个时段更容易出现，问题的出现是与天气的关系大，还是与道路的状况关系更大？这些规律性的问题，只有在大数据的条件下才能得出结论。在这里，我还可以给大家推荐另外一本研究大数据的著作，那就是全球复杂网络研究权威、美国东北大学教授艾伯特·拉斯洛·巴拉巴西所著的《爆发》。这本书指出，基于大数据，人类的行为都是有规律可循的，因为更多的数据会更真实地呈现出某一种行为的规律