



中文社交媒体上 对地方和中央政府的 态度

食品安全案例研究

Douglas Yeung 和 Astrid Stuth Cevallos 合著

This is a Chinese translation (simplified characters) of *Attitudes Toward Local and National Government Expressed over Chinese Social Media: A Case Study of Food Safety*.

有关本出版物的更多信息，请查询 www.rand.org/t/rr1308。

兰德公司出版，加州圣莫尼卡
版权所有 © 2016 兰德公司
RAND® 是兰德公司的注册商标。

有限的平面和电子媒体发行权

本文件和文中所含商标受法律保护。本作品的知识产权归兰德公司所有，不得用于商业用途。未经授权，严禁在网络上发布本作品。本文件仅允许个人复制使用，但不得擅自修改和删节。未经许可，不得复制或以其他方式将兰德公司的任何研究文献用于商业用途。有关翻印和链接授权的信息，请查询 www.rand.org/pubs/permissions。

兰德公司是一家解决公共政策挑战的研究机构，旨在协助推进全球社区的安全、卫生与繁荣事业。兰德公司致力于公共利益，属于非营利性、无党派组织。

兰德公司的出版物未必代表其研究客户和赞助商的观点。

赞助兰德公司

欢迎通过下列网址提供可免税的慈善捐赠
www.rand.org/giving/contribute

www.rand.org

序言

本报告介绍了对社交媒体的心理语言学分析，旨在探索社交媒体在中国公众对国内政治问题（如环境、食品安全、地方与中央政府）的反应上如何发挥前沿视角的作用。该研究项目的首要目的是了解社交媒体上关于各种话题的言论内容和基调。通过对中文社交媒体的初步探讨，还介绍了分析中文社交媒体的方法、可能性与挑战、所需的工具和步骤、以及与英语社交媒体的比较。本项研究面向那些关注有关各级政府的公共舆论的政策制定者和有意探索中文社交媒体的研究人员。

本研究项目承蒙唐仲英基金会的慷慨资助。凭此资助，兰德亚太政策中心于 2007 年成立了唐氏美中关系研究所，而本研究项目即由唐氏美中关系研究所主办。

兰德亚太政策中心

兰德亚太政策中心 (CAPP) 是兰德公司国际项目的一部分，致力于分析亚太及周边地区的政治、社会、经济和科技发展动态。通过研究与分析，亚太政策中心协助公共和私营部门的决

策者解决问题，应对挑战，并探索推进社会安定、智能与繁荣的途径。

有关兰德亚太政策中心的更多信息，可访问 www.rand.org/international_programs/capp 或通过网页上的联系方式与中心主任联系。

目录

序言.....	iii
综述.....	xi
鸣谢.....	xvii
缩略语.....	xix
第一章	
社交媒体在中国作为政治表达渠道.....	1
社交媒体在当代中国应用情况.....	2
中国的食品安全监管与福喜事件.....	9
第二章	
社交媒体在中国应用情况的自动分析.....	15
社交媒体分析采用的心理语言学方法.....	16
收集和准备中文社交媒体数据的样本.....	18
在中文社交媒体背景下处理和解读数据.....	22
第三章	
研究结果.....	25
关于食品安全和福喜事件的推文：趋势变化.....	25
第四章	
结论.....	39
中文社交媒体用户对中央政府表达的不满普遍多过地方政府， 但 Twitter 与微博上的看法有所不同.....	39
分析中文社交媒体所面临的特有挑战.....	41

未来的研究可以扩大对中国国内政治和社交媒体平台的调查， 并解决自动分析面临的挑战.....	44
思想总结.....	48
附录	
搜索分类法.....	49
参考文献.....	51

图目录

图 3.1	议论食品安全推文中的情绪词使用量	26
图 3.2	议论福喜事件推文中的第一人称代词使用量	28
图 3.3	议论福喜事件推文中的其他代词使用量	29
图 3.4	议论福喜事件推文中的情绪词使用量	31
图 3.5	在提到地方与中央政府和美国企业的食品安全推文中 代词使用量的比较	32
图 3.6	在提到地方与中央政府和美国企业的食品安全推文中 情绪词使用量的比较	32
图 3.7	在提到地方与中央政府和美国企业的食品安全推文中 法律和道德词语使用量的比较	33
图 3.8	在提到地方与中央政府和美国企业的食品安全微博中 代词使用量的比较, 2014 年 8 月 26 日 - 9 月 8 日	34
图 3.9	在提到地方与中央政府和美国企业的食品安全微博中 情绪词使用量的比较, 2014 年 8 月 26 日 - 9 月 8 日	36
图 3.10	在提到地方与中央政府和美国企业的食品 安全微博中法律和道德词语使用量的比较, 2014 年 8 月 26 日 - 9 月 8 日	37

表目录

表 1.1	上海福喜事件和其他食品安全问题的时间脉络， 2014 年 7-8 月	13
表 2.1	按类别分析食品安全主题的微博帖子数量	20
表 A.1	用于收集中文社交媒体帖子的词语	49

综述

社交媒体日益成为常见的政治表达途径。分析社交媒体，有助于解释对政治事件和运动的公众舆论和看法。心理语言学致力于研究语言可以揭示的心理状态，譬如，态度和情绪。正是通过这种研究，深入剖析了伊朗和香港政治事件中的公众舆论和情绪。本项研究工作利用计算机技术分析了中国社交媒体文本，特别是在 2014 年麦当劳和肯德基餐厅曝出使用“福喜”过期肉丑闻期间，社交媒体对地方和中央政府以及西方公司的态度。

在中国，社交媒体让民众对治理问题的议论扩大化。有些人认为社交媒体推动了有意义的政策变化，另一些人声称社交媒体不过是给公民提供了一条宣泄途径，在不威胁当局的前提下，公民可以由此抒发内心的不满。中国的互联网用户数量庞大，接近 6.5 亿人，占总人口的 48%；互联网用户数量高于印度，普及率也超过印度、泰国和越南。微博在这些用户中非常流行，而新浪微博是最活跃的微博平台。

中国微博用户数量在 2012 年达到 3.09 亿的峰值，此后，由于政府加强对内容的监控而有所下降。Facebook 和 Twitter 等境外流行的社交媒体平台，由于遭到中国政府屏蔽，在国内没有被广泛使用。尽管如此，许多中国网民却通过代理服务器

或其他“翻墙”手段访问 Twitter。估计使用 Twitter 的中国公民人数介于 10,000 至 3,500 万。

通过分析食品安全问题，可以深入了解社交媒体在测定中国公众舆论方面的作用。食品安全在 2008 年引起人们的广泛关注，当时一家国有企业被发现在奶粉里掺入三聚氰胺。随后几年，越来越多的中国人觉察食品安全是重大问题，尤其是最有可能活跃于新浪微博的高收入人群和年轻人。分析社交媒体上关于这些问题的帖子，特别是 2014 年 7 月和 8 月曝出的福喜丑闻，有助于考察这方面的中国舆论动向。

研究方法

我们选择了讨论一般食品安全和福喜事件的中文微博帖子样本，收集了 2014 年 6 月 1 日至 8 月 31 日期间的 2,184 则中文推文，以及 2014 年 8 月 26 日至 9 月 8 日期间的 537 则博文。遗憾的是，我们只能收集在数据收集工作开始之后发布的博文，而无法像收集推文那样收集早前发布的博文。尽管如此，这些样本可以揭示不同受众是如何看待这些问题及其相应责任的，而有关受众是能够绕开屏蔽使用 Twitter 的精英群体以及可能反映较广泛中国舆论的大众群体。

除了分析提到地方和中央政府以及西方公司的帖子，我们还使用自动化软件识别代词、表达情感和责任的话语。使用第一人称单数代词（例如，我、我的）表明以自我为中心，这与威胁、沮丧或不安等消极状态相关；第一人称复数代词（例如，我们、我们的）表明作者有群体或社群的意识；第三人称代词（例如，他、她、他们）表明作者不认同其讨论的群体。

表达情感的词语可以说明作者对某个话题是正面还是负面的态度。在讨论特定实体时使用表达责任的词语，可能表明作者将食品安全的责任与该实体联系起来。

研究结果

Twitter 分析

福喜事件后有关食品安全的舆论走向表明，中国舆论与地方和中央政府的反应合拍。在福喜事件爆发的一周时间内，讨论食品安全的推文透露负面情绪骤升，而正面情绪下降。直到中央政府要求外国快餐连锁企业在其中文网站上公布供应商信息时，有关食品安全的负面情绪才回到事发前的水平。然而，有关福喜事件的负面情绪持续高涨了几周时间。

关于此事的帖子使用第一人称单数代词的比率也很高，进一步表明在丑闻曝出之初公众一般持消极态度。当上海警方发布消息称已逮捕在过期肉事件中负主要责任的高管时，推文中其他人称代词的使用量达到峰值，这说明公众在疏远这些高管。

Twitter 用户对此事的失望情绪可能较少针对地方政府，更多是针对中央政府和美国企业。特别是，Twitter 用户在议论中央政府和美国企业时，更倾向于使用第三人称代词。网民对中央政府和美国企业使用负面情绪词汇的数量，也是对地方政府的两倍。对表达责任的词语的分析表明，Twitter 用户认为地方政府有责任查处违法行为，或者在这方面工作大概是卓有成效的。相比之下，在关于中央政府的推文中，腐败一词出现的频率几乎是关于地方政府的两倍。这可能反映了人们对中

央或地方各级政府的腐败问题的普遍态度，抑或是人们经常议论中国的腐败现象，或者习惯上将整个国家一体议论。

微博分析

微博用户似乎有更强的社群意识，对中央政府的沮丧或不安感较少。在谈论食品安全时，他们相比 Twitter 用户来说，对中央政府的看法更正面，而对地方政府的看法更负面。

议论地方政府的微博文章中第一人称单数代词的使用次数，接近议论中央政府的两倍。相比之下，议论中央政府的微博文章中第一人称复数代词的使用次数，大约是议论地方政府的五倍。这说明用户在讨论食品安全时，对国家的亲切感超过省、市层面。

同时，微博用户对中央政府表达的愤怒要多过地方政府。在议论食品安全的微博文章中，提到国家或中央政府的，其表达愤怒的词语占比是提到省、市当局的四倍以上，其表达悲伤的词语占比也是提到地方政府的三倍。

尽管极少的微博样本提到美国或美国企业（可能是因为福喜事件后的数据收集存在延迟），但与提到中国实体的微博相比，更多是议论责任或谴责之辞。要厘清上述乃至其他结果的具体解释，可能需要对内容作进一步的人工分析。

结论

Twitter 与微博用户之间的差异可能反映不同社群的态度。由于访问 Twitter 需要花费更多功夫，发布中文推文的用户可能

在政治话题上更加活跃，更倾向批评中央政府。微博用户或更能代表整体的政治观点。

另外，微博对中央政府的态度更为正面，或反映中央政府对国内媒体的管控。微博表达的民意与中央政府对此案的关注呈对应关系，或反映中央政府为将治理丑闻转移给地方官员和西方公司而采取的公关行动。同时，议论中央政府的帖子所使用的情感词语多于议论其他实体的，说明中央政府未能完全转移这次丑闻，抑或是微博用户认为食品安全是中央政府负责监管的大事。

未来的研究可以扩大对中国国内平台和社交媒体平台的调查，并解决自动分析面临的挑战。扩展关键词（例如，将肉质或快餐纳入食品质量分析）或可抓取更大范围的议论。未来的研究也可以探索通过更严谨的方法来选择推文分类所用的词语，以及比较 Facebook 和 Twitter 用户与更常用的中国平台的用户。

污染和环境等其他国内问题也值得开展类似的调查。就中央和地方对这些问题的责任，网民可能持有不同看法。例如，他们可能为中央政府发布减少污染的命令叫好，但指责地方政府无视新法规——这是普遍做法，因为地方政府缺乏监管国有企业的动力。

对中国的政治问题开展社交媒体研究，或可启发意料之外的使用方式，包括促进社会控制。研究人员应考虑政策制定者对研究结果的利用方式的影响。

鸣谢

Mike Lostumbo 和 Seth Jones 为开展本项研究提供了支持和宝贵的指导。Zev Winkelman、Christopher Skeels 和 David Manheim 曾协助数据的收集与管理。Academia Sinica Lexicon/Corpus Group 和 Chinese LIWC Dictionary Group 对中文版语文探索与字词计算给予帮助。最后，Rafiq Dossani、Rogier Creemers 和一位匿名审稿人提出了精辟的意见和建议，帮助大大完善了本报告。在此，一并表达谢意。

缩略语

中共	指	中国共产党
C-LIWC	指	中文版语文探索与字词计算
LIWC	指	语文探索与字词计算
PX	指	对二甲苯 (PX)
SC-LIWC	指	简体中文版语文探索与字词计算

在世界各地，社交媒体被用于抗议活动或动员等政治事件的情况越来越普遍。因此，分析社交媒体在这方面的应用，有助于解释可能引致或议论这些运动的公众舆论和看法。以往的心理语言学研究致力于分析语言可以揭示的心理状态，譬如，态度和情绪。正是通过这种研究，剖析了如何通过社交媒体了解正在伊朗（Elson 等人，2012 年）和香港（Yeung 和 Cevallos，2014 年）发生的政治事件中的公众舆论和情绪。本项研究工作利用计算机技术分析了社交媒体文本，我们以此为基础，探讨了中文社交媒体上表达的态度和议论。我们调查了 Twitter 和新浪微博等微博平台上中文用户对地方政府、中央政府和西方公司的态度。我们选择食品安全作为具体话题展开调查，特别是采用 2014 年 7 月麦当劳和肯德基餐厅使用“福喜”过期肉丑闻（详见下文）作为案例研究。

食品安全关系到中国国内政治的几个重要问题，包括媒体（和公众）作为监察人的角色，政府监管在一个执行责任分散的体系下面临的挑战，法治薄弱，以及地方官员没有动力查处有利可图的本地企业（Olesen，2011 年）。过去十年，涉及婴儿配方奶粉和猪肉的食品安全问题，让中国消费者人心

惶惶，对国内食品供应愈加忧虑和不信任（Jiang，2008年；Barclay，2011年）。一些新闻报道指出，在这些丑闻曝光后，中国官方放宽了媒体报道和公众议论食品安全事件的范围，以求更好地监管国内蓬勃发展的食品行业（Olesen，2011年）。因此，食品安全可能也是研究中文社交媒体的极好切入点。举例来说，近年广泛报道的食品安全丑闻，譬如曝光快餐企业使用过期肉的福喜事件，发酵成具体的可测事件，在社交媒体引发热议。

本研究报告试图探讨中文社交媒体用户是否和如何议论福喜事件和食品安全，以及他们认为谁（例如，中国地方政府、中央政府、食品安全丑闻涉事西方公司等）应该对福喜事件乃至广义上的食品安全承担责任。为了解决这些问题，我们对中文社交媒体开展了探索性研究，从新浪微博和 Twitter 收集了网帖。碍于资源有限，我们进行了定性分析，而非更严谨的统计或内容分析。本研究项目的首要目的，是证明我们用于社交媒体分析的计算机技术（例如，Elson 等人，2012年）能够揭示中国国内政治生活中人们对地方和中央政府的态度。本报告还解释了研究中文社交媒体文本的一些挑战和局限。研究结果和评估中文社交媒体的方法，可能会引起中文分析人员或寻求深入了解中国国内政治舆论的政策制定者的关注。

社交媒体在当代中国应用情况

在中国，社交媒体让民众对治理问题的议论扩大化。当然，人们对社交媒体在中国政治领域的性质和作用存在分歧；有些人认为社交媒体推动了有意义的政策变化，另一些人声称社交媒

体不过是给公民提供了一条宣泄途径，在不威胁中共政权的前提下，公民可以由此抒发内心的不满（Xiao，2011年）。近年来，中国对社交媒体的政策转变，也扭转了民间行动和结社的可能性（Creemers，2015年）。然而，社交媒体似乎已成为表达意见和动员民众的重要工具，特别是涉及到曝光和谴责地方政府的越权行为时。

在中国什么人使用社交媒体？

截止至2014年12月，中国的互联网用户数量接近6.5亿，占总人口的近48%（中国互联网络信息中心，2015年，第25页）。相比之下，人口规模与中国大致相等的印度，互联网用户为2.43亿，仅占总人口的19%；人均国内生产总值(per capita GDP)与中国相当的泰国，互联网用户为1,900万，占总人口的29%。中国的互联网普及率更接近同为集权（共产主义）政体的越南，后者拥有4,000万互联网用户，占总人口的43%（世界银行，日期不详；互联网实时统计网站 [Internet Live Stats]，2014年）。大部分中国网民（“网民”是许多中国互联网用户的自我称谓）是年轻人，不满30岁的网民占56%；男性占56%；52%的网民有高中以上学历；逾70%是城市居民；他们的收入高于平均水平，53%的网民年收入超过人民币24,000元（约合3,900美元）——中国家庭年收入平均只有人民币13,000元（约合2,100美元）。逾85%的中国网民使用手机上网（中国互联网络信息中心，2015年，第28-35页）。¹

¹ 中国互联网络信息中心 (CNNIC) 是中国的一家域名注册商和研究中心，目前由中国国家互联网信息办公室主管。

具体而言，微博在中国相当普及，本土的平台多如牛毛，例如，新浪微博、腾讯微博、搜狐微博、网易微博、饭否、以及嘀咕。其中，2009年8月推出的新浪微博（又简称“微博”）是最活跃的平台。这些微博平台与 Twitter 类似，用户可以向公众发布不超过 140 个字符的短消息。但是，因为大部分中文词语仅只含一两个字符，中文微博帖子可以传达的信息量远超过同类型的英文推文。

微博用户于 2012 年攀升至 3.09 亿的峰值，这意味着 55% 的中国互联网用户都有微博帐号（中国互联网络信息中心，2014 年，第 48 页）。此后（原因在下文探讨），微博用户数量下降近 20%，到 2014 年底为 2.49 亿，仅占中国互联网用户的 38%（中国互联网络信息中心，2015 年，第 56 页）。这些用户大都是年轻人，其中 53% 不满 24 岁，90% 不满 34 岁；受教育程度颇高，其中 71% 的人持有大专以上学历（Bai，2014 年）。虽然新浪微博的官方统计显示男女用户数量相当，但其他研究机构称男性用户占比略高，介于微博用户总数的 52% 至 57%（Ng，2013 年）。大部分用户也比较富裕：48% 的用户年收入超过人民币 36,000 元（约合 5,900 美元）（Millward，2012 年）。不出所料，微博用户大多数是城市居民：北京市（总人口估计为 2,200 万）和江苏省（总人口估计为 8,000 万，其中包括大约 2,400 万上海居民）的用户均超过 1,500 万；广东省（总人口估计为 1.05 亿）的用户超过 3,000 万（新浪微博数据中心，2012 年）。事实上，广东一个省就发布了逾两成的微博帖子（Ng，2013 年）。

在中国，Facebook 和 Twitter 等在美国乃至世界各地流行的社交媒体平台，并不如人们所预料的那样被广泛使用。中国政府担心社交媒体平台可能会迅速传播未经审查的讯息，从而引发有组织的抗议活动，遂于 2009 年 6 月开始屏蔽 Twitter、

Facebook 和其他社交媒体网站，当时正值天安门广场抗议事件 20 周年纪念日前后。中国网民仍然可以访问 Twitter，有些人确实在访问。但要访问 Twitter，必须借助代理服务器或虚拟专用网络 (VPN) 来翻越长城防火墙，所谓的“翻墙”（参见 Tkacheva 等人，2013 年，第 93-117 页；Robinson、Yu 和 An，2013 年）。因此，无法确切地测定有多少中国公民使用 Twitter。² 估计数字相差甚远：一项极具争议的研究显示，约有 3,500 万中国公民经常使用 Twitter，这里可能是 Twitter 全球最大的单一市场（Ong，2012 年）；但随后最少有两项分析，连同著名的中文 Twitter 用户、互联网自由活动家赵静 (Michael Anti) 的概算表明，中国用户数量实际介于 10,000 至 100,000（Russell，2012 年；Zhai，2013 年；Ng，2014 年）。由于难以获得中国 Twitter 用户的可靠统计数字，中国 Twitter 用户数量不得而知。由于中国互联网用户需要额外花一些功夫和步骤来访问 Twitter，尽管中国的 Twitter 用户在其互联网用户中占比不大，但他们更有可能通过 Twitter 发表看法或参加评论，因为在中文社交媒体平台上这些看法和评论会被删除。事实上，许多中国异见人士，包括颇有争议的艺术家和煽动者艾未未及互联网自由活动家兼记者赵静，都活跃于 Twitter（Goodman，2012 年）。2011 年一份关于国际博客使用互联网过滤翻墙工具（如 VPN 或代理服务器）的报告同样表明，在互联网用户中，使用这类工具的用户占比极小，但这类用户更可能是人脉较广、影响力较大的个人，他们有能力向更广泛的受众传播信息和意见（Roberts 等人，2011 年）。虽然上述门槛可能导致

² 审查和互联网流量监控未必是中国网民很少使用 Twitter 或 Facebook 的唯一原因；部分研究人员认为，从满足中国互联网消费者的需求角度，腾讯微博、新浪微博和人人网等中国同行做得确实更好。详情见 Sullivan，2012 年。

选择偏差和用户统计信息扭曲，但新浪微博的用户趋势（如城市居民、受教育程度）可能也适用于中国的 Twitter 用户。

中国的审查、匿名问题和社交媒体格局变化

随着微博平台日益流行，它们在中国的政治表达和动员中也发挥了愈加显著的作用。鉴于新浪微博等广泛使用的通信工具，有可能帮助动员民众参与对抗政府的集体行动，中国政府已设法监控社交媒体，以便降低新浪微博活动演化成有组织的、危及政权的政治运动的可能性。中国政治专家谢淑丽 (Susan Shirk) 认为，中国领导人意识到，喜欢上网表达失望情绪的人“更有可能冒险参加或组织群众抗议” (Shirk, 2011 年)。结果，在政府的授意下，公司雇用的内容监察人员会迅速删除敏感内容。一般情况下，相比那些表达不满的帖子，关于“可能引发集体行动的现实事件”的帖子更有可能遭到删除，这说明当局力图阻止准抗议者通过社交媒体组织活动，但容许网民抒发不满情绪 (King、Pan 和 Roberts, 2014 年, 第 1 页; 另见 King、Pan 和 Roberts, 2013 年)。2012 年，为了控制新浪微博上可能危及政权的讨论，官方制定了两条新法规：一条是要求网民在开办帐号时实名登记；另一条要求新浪微博监控粉丝数量达 10 万以上的知名博主的帖子，并在问题帖子发布的五分钟内删除之 (Tkacheva 等人, 2013 年)。正如一些报刊文章指出，自从 2013 年 3 月习近平正式掌权以来，政府打压有影响力的博主的相关措施来势更猛。(Chin 和 Mozur, 2013 年)。

虽然新浪微博仍然被广泛使用，但此前其作为对政治问题发表意见和动员群众的首选社交媒体平台的公认地位，已有所削弱。2013 年，新浪微博失去了近 10% 的用户。即使是留下

来的用户，其活跃度也在下降——23%的用户减少了微博的使用，只有13%的用户增加了微博的使用（中国互联网络信息中心，2014年，第66页）。同样，《电讯报》近期的一项研究显示，就连最活跃的微博用户在2013年12月发布的帖子数量，也比2012年3月少了74%（Moore，2014年）。用户少发微博乃至彻底弃用也许对应了两种趋势：一是政府在2013年8月打压知名博主；二是替代性的社交媒体平台崛起——其中一些可能更难审查，例如，移动即时通信应用微信。不过，新浪微博仍然被广泛用作公众参与治理的重要渠道。微博于2014年6月报告称，其月活跃用户总计1.57亿，日活跃用户总计7,000万（新浪微博，2014年）。³也许是为了利用受众的规模和广度，中国一些地方政府创建了自己的微博帐号，用于促进与本地群众沟通的效率（Schlaeger和Jiang，2014年）。

社交媒体用作政治表达和动员的工具

微博的另一个用途是，为网民提供了一个平台，可以用来传播关于发酵中的危机的信息（特别是那些涉及官员企图掩盖腐败或错误的危机），动员群众支持政府应对此类事件，批评政府的政策。在少数情况下，尽管中国的公民社会力量微弱（参见Jiang，2010年，第117页）⁴，网民参与线上微博社区不仅形成了线下行动，也带来了政策结果。尤其是，微博用户让本地问题受到全国的关注，迫使当局做出回应，下面举两个著名的案例：

³ 请注意，没有迹象表明这些用户是否位于中国内地。

⁴ Jiang认为，“公民社会是以完全独立于国家的社团为基础，在这个意义上，中国尚未形成真正的公民社会”（2010年，第117页）。

- 2011年7月，一位新浪微博用户爆料温州动车追尾脱轨，造成40多人死亡的消息。随后，微博上掀起一阵热议，更曝出当局试图掩埋动车残骸，妨碍对事故原因进行彻底调查。网民揭露有关部门存在系统性腐败，并批评政府把快速现代化的任务置于安全之上，同时从头梳理了这场灾难的来龙去脉。最后，政府迫于压力将铁道部部长以受贿罪革职查办（Tkacheva 等人，2013年，第104-107页）。
- 2011年夏天，大连居民上新浪微博组织抗议，要求关停一家邻近市中心的对二甲苯(PX)工厂。由于担心过量接触对二甲苯（一种用于生产塑料的化学物质）会危害健康，超过12,000名抗议者在热带风暴破坏了化工厂周围的防护栏后涌上街头。虽然监察人员试图制止微博上的信息传播和动员号召，但组织者迅速、有效地利用了该媒体。在抗议发生的当天，官方关停了涉事工厂（Tkacheva 等人，2013年，第107-110页）。

上述事例展示了机智的网民——通常来自年轻的、受过良好教育的城市中产阶层，如何利用新浪微博将地方滥用职权的行为和其他政治怨愤昭示天下，使当局难以阻止民众不满情绪的蔓延，因而更可能迫于公众压力做出回应。鉴于中国没有多党选举和其他向政府问责的民主机制，有迹象显示中央政府认为社交媒体可以在监督腐败问题上发挥重要作用。2014年底，中央纪律检查委员会创建了由其控制的线上贪腐举报平台，试图把社交媒体当作一种情报收集机制，同时最大限度地降低平台引发集体行动的风险（Creemers, 2015年）。这种社交媒体上的互动可充当中央政府观察和控制地方政府的工具，

也突显了研究网民如何通过新浪微博抱怨制度监管滞后造成食品安全危机的重大意义。

中国的食品安全监管与福喜事件

政府监管和地方与中央政府的责任

在当局的授意下，微博平台负责审查每一则可能引发有组织抗议的帖子。然而，表达不满的微博帖子一般会得到容忍。虽然有人在这种二元现象中找到支持“社交媒体作为宣泄途径”理论的证据（Xiao，2011年），但还有另一种可能的解释：在中国这样高度分权的政治制度下，公众通过社交媒体表达不满可以帮助中央政府监察地方官员。⁵

诚如中国传统谚语“山高皇帝远”所影射的，地方官员历来漠视远在北京的中央政府。实际上，许多中国人认为，当腐败无能的地方官员无力履行中央政府用意良好的指令时，社会就会陷入动荡。中国领导人利用这种观念将民怨引向地方官员，这种策略被布鲁金斯学会的李成归纳为“为国家着想，怪责地方”（李，2006年；另见Nathan，2003年）。⁶ 其实，有些分析师认为，中国这样的集权政体容忍甚至鼓励公众抗议（包括线上线下的），是为了监测社会不满情绪，约束地方官员，

⁵ 关于中央政府难以驾驭地方官员和中央与地方关系紧张对政权造成的问题，详见 Fewsmith，2013年。

⁶ 中国公众相信或假定中央政府行仁善的传统自古有之，可追溯到帝制时期。而中国领导人继续培养这种观念，从理论上灌输人民拥护政权的思想，这有助于确保政治诉求局限于地方，不会发展成全面呼吁体制改革。关于这种观点的来历，详见 O'Brien 和 Li，2006年，特别是第42-47页。

以及在必要时推行政策改革（Johnson，2004年和2011年；Lorentzen，2013年）。如此一来，公众表达不满甚至是抗议有助于加强而非削弱中央政府的领导。

中国的立法机关全国人民代表大会负责制定《食品安全法》，据此规范国家食品安全标准和法规（Balzano，2014年和2015年）。截至本报告撰写之时，这部法律正在修订过程中，其中纳入了公众咨询意见。但是，负责检查和落实地方企业达标的却是省、市级食品药品监督管理局的官员。当食品安全丑闻爆发时，中国网民会追究谁的监管缺失责任？这个问题的答案，有助我们深入研讨在中国这类集权政体的国家中央与地方政府之间的关系。

食品安全成为公众关注的问题

食品安全在2008年北京奥运会之后成为中国公众广泛关注的问题，当时，全国最大的国有奶粉企业三鹿集团被曝在奶粉配方中掺入三聚氰胺（一种有毒的、非法的化学添加剂），人为地提高蛋白质含量的检测结果。掺假的婴幼儿配方奶粉致使294,000名婴儿患病，造成6名婴儿死亡。涉案的两名当事人被判死刑，其中一人生产和销售三聚氰胺，另一人则将三聚氰胺掺入奶粉里面。此外，三鹿董事长兼首席执行官被判无期徒刑，并处罚金人民币2,000万元（约合300万美元）（Barboza，2009年）。

从那时起，中国公众对食品安全作为公共卫生和安全问题的关注度显著提升。皮尤研究中心（Pew Research Center）于2013年开展的民意调查显示，38%的中国人认为食品安全“问题非常严重”，高于2008年的12%。认为食品安全“问题非常严重”或“问题比较严重”的中国人总体占比，从2008年的49%上升至2013年的76%，升幅达26个百分点（皮尤

研究中心，2013年）。其中，城市居民、高收入者和30岁以下的年轻人最有可能认为食品安全问题严峻。值得注意的是，他们也是新浪微博上最活跃的人群。

虽然在2008年三鹿三聚氰胺危机爆发时微博尚未问世，但在2013年上海发生的两宗食品安全丑闻中，它充当了传播信息和表达愤怒的平台。当年八月，上海水域发现漂浮着近6,000头病死猪，后来查明是被农民丢弃到黄浦江上游嘉兴段的，网民纷纷到微博上通过黑色幽默和讽刺手法表达不满。例如，一名网友写道：“我们是喝着三聚氰胺奶、吃着激素鸡和地沟油长大的一代。我们的免疫系统早已百毒不侵。上海人民不会轻易被一些死猪击倒”（Lu，2013年）。事发后两个月，官方逮捕了63人，这些人常年在上海和江苏省销售经过化学处理味道和外观像羊肉的狐狸、水貂和老鼠肉，涉案金额达160万美元（Kaiman，2013年）。网民同样是义愤填膺，经常以呕吐的表情回应这则新闻。一名网友写道：“想像一下——你以为你在吃烤羊肉串，但其实你吃的是毛茸茸的老鼠。我以后还怎么敢吃羊肉啊？”（Funakoshi，2013年）。正是在上述背景下，网上开始议论福喜事件。

福喜事件的时间脉络

2014年7月20日，东方卫视通过暗访调查报道，揭露了上海福喜食品有限公司（“上海福喜”）的中国工人使用过期肉加工麦当劳和肯德基的产品，并将掉落在地的肉块放回加工设备。涉事公司是美国肉类生产和供应商福喜集团的子公司。上海监管部门介入后发现上海福喜篡改了4,000多批次牛肉饼的生产日期；这些肉类的保质期为9个月，标签上的生产日期是2014年1月，但实际是在2013年5月生产的（路透社，2014年）。7月21日，上海福喜丑闻成为“微博上浏览次数最多

的话题，2,700 多万人谴责洋快餐，中国食品安全监管，以及地方媒体倾向于追捧知名消费品企业制造新闻”（Kuo，2014 年）。在新浪微博上，这宗丑闻被称为“福喜事件”、“过期门”、甚至“臭肉门”。⁷

随着民众在社交媒体上表达失望和厌恶情绪，地方与中央政府监管部门仓促应对，而受牵连的美国企业（包括麦当劳和肯德基）则急忙同上海福喜撇清关系，并暂停在门店销售某些肉类产品。表 1.1 描述了政府应对危机的时间脉络。

我们的案例研究有别于先前讨论的关于中国网民曝光政府不当行为、腐败和文过饰非的事例。在这里，网民没有揭发对上海福喜的监管失灵，也没有组织线下抗议来推动改革。因此，在此案中社交媒体的作用看似不大，因为是传统媒体首先报道了丑闻。不过，网民在社交媒体上对丑闻的反应才是我们的研究对象，我们无意探讨社交媒体作为一股动员力量的作用，而是旨在研究社交媒体在集权国家作为测定民意的信息来源（参见 Olesen，2011 年）。

⁷ 网民经常使用富有创意的委婉语来避开审查，在此案中，词尾加上“门”字，让人回联想起水门丑闻。

表 1.1

上海福喜事件和其他食品安全问题的时间脉络，2014 年 7-8 月

日期	事件
7 月 20 日	东方卫视播出了调查上海福喜集团违反食品安全规定的暗访录像，包括重新加工和包装过期肉类产品并供应给境内的麦当劳和肯德基餐厅
7 月 22 日	上海监管部门查封了上海福喜的 100 吨过期肉
8 月 3 日	上海福喜六名高管被上海警方拘留
8 月 6 日	中国国家食品药品监督管理总局（“食品药品监督管理局”）颁布新的食品安全召回制度（草案），明确了食品生产商和企业（而非政府）负有保证食品安全的主要责任
8 月 9 日	沃尔玛中国分公司被指违反食品安全规定
8 月 11 日	食品药品监督管理局要求五家外资餐饮企业，包括麦当劳和肯德基，向社会公开供应商信息
8 月 12 日	上海立法机关起草食品安全溯源系统提案
8 月 18 日	亨氏在中国召回铅含量超标的婴儿米粉
8 月 29 日	上海司法机关正式批准逮捕上海福喜的六名高管

第二章

社交媒体在中国应用情况的自动分析

为了解中文用户在微博平台上对地方政府、中央政府和西方公司表达的态度，我们探讨了网上关于食品安全和福喜事件的议论。我们从 Twitter 和新浪微博收集数据，获取了社交媒体消息的心理语言学指标，进而对由此得出的定量结果作出解读。

分析这些社交媒体数据，是为了解答两个主要问题：

- 中文社交媒体用户对于近期的食品安全事件在议论什么话题？
- 中文社交媒体用户在议论中国地方政府、中央政府或西方公司（例如，卷入食品安全丑闻的公司）的时候情绪有何差异？他们认为相关丑闻应归咎于谁或者应向谁问责？

例如，我们的分析包括对照中国食品安全事件的时间脉络描绘量化指标的趋势，据此确定两者是否同步。为分析一段时间内的数据，我们按周汇总中文微博帖子，这似乎最适合分析几个月（从 2014 年 6 月初至 2014 年 8 月底）的大趋势。换言之，我们合并了每一周的社交媒体文本，生成该周的单一量化指标集合。然后，我们跟踪这些指标在议论纷纷的食品安全丑

闻之后一段时间内的变化，并检验这些指标的变化规律与丑闻相关的具体事件是否同步，由此了解人们在事前、事中、以及事后的态度。

我们还通过比较社交媒体用户讨论特定话题的量化指标来分析数据。这些话题以搜索词分类法为基础，搜索词是我们设定的包含食品安全和福喜事件相关问题的一组关键词。其他分析涉及比较议论西方（如美国、麦当劳）、中国地方政府或中央政府的推文子集。（关于将推文归类所用的分类法，参见附录。）

社交媒体分析采用的心理语言学方法

社交媒体文本的心理语言学分析是基于一种观点：语言的使用可以揭示内在状态，即思想和感情，例如，对某个话题的正面或负面情绪，或者与某个群体或公众人物的联系。我们利用自动分析方法制定了这些情绪的量化指标，重点考察人们在社交媒体上使用某些类别词语评论中国食品安全的比率。为解读这些词语类别所表达的态度或情感，我们借鉴了分析方法所用软件工具的相关文献的既有先例：2007年版语文探索与字词计算（LIWC，读作“Luke”）（语文探索与字词计算：LIWC2007应用，2007年；Pennebaker等人，2007年）。该软件可以对书面文本——譬如使用社交媒体发布的消息，进行内容分析。LIWC会依据预置词典比对给定文本中出现的词语，预置词典包含分成多个语言学（如代词、冠词）和心理学（如情感、认知）类别的词语。比对结果用于计算每个类别的词语总数在文

本中的占比。这些比率代表每个词语类别在文本中的相对使用频率，有助了解作者的态度、情感、意向等。

大量研究表明，某些类别的词语具有特别意义。例如，尽管代词在整体的词语使用中占比不大，但它们实际上大概揭示了一个人的心理状态（Pennebaker 和 Chung，2007 年）。通过对 LIWC 生成的定量结果进行定性分析，也就是研究人们对这些词语类别的使用的变化趋势并在现实背景下解读这些变化，我们可以得出有关公众舆论和情绪的结论。

某些类别词语的使用情况可能对了解社交媒体的态度和意见特别有用：

- **代词。**这些语言功能词的使用表明了作者的关注点。第一人称单数代词（例如，我、我的）表明以自我为中心，这与威胁、沮丧或不安等负面情绪相关；第一人称复数代词（例如，我们、我们的）表明作者自认为是更大群体的一分子，意味着群体认同感或社群意识；第三人称代词（例如，他、她、他们）表明作者在描述一个非本人隶属的群体，意味着其身份有别于该群体（即“他者化”）。
- **表达正面和负面情绪的词语。**较多使用这些词语可能说明作者对所议论的话题是趋于正面或负面的态度。
- **表达道德 / 法律责任的词语。**在议论特定实体（例如，中国地方政府、中央政府、西方食品企业）时使用这些词语，可能表明作者将食品安全的责任与此等实体联系起来。

关于我们分析中使用的这些词语类别，以及它们如何进一步用于分析社交媒体，详见 Elson 等人（2012 年）。

使用 C-LIWC 分析中文文本

中文版语文探索与字词计算 (C-LIWC) 以 2007 年版 LIWC 英文词典的繁体中文译本为基础 (Huang 等人, 2012 年)。开发者是台湾研究人员, 他们利用台湾中央研究院的词典开发自己的 LIWC 版本。由于 LIWC 最初是针对英文设计的, 对中文文本开展语言学研究面临诸多潜在的挑战。首先, 有一些中文短语、成语和俚语是英语人士不会使用的表达, 因此很难翻译成英文。这些词语大概不会收录到英文版 LIWC 词典, 故需要另行添加到 C-LIWC 词典。实际上, C-LIWC 包含几个独特的类别, 如第二人称复数代词和几个时态标记。其次, LIWC 所依据的心理学研究大都在西方环境下开展, 未必适用于中国人群。Li 和同事 (2012 年) 比较了英文版和中文版 LIWC 的输出, 描述了每个版本所抓取和分析的词语类别的异同。他们将多个 LIWC 类别合并成较大的组件, 发现每种语言都具有几个共同的心理语言学特征, 如几个代词类型 (包括与本研究相关的, 例如, 第一人称单数、第一人称复数、第三人称复数)、否定性、叙述、以及社会进程。作者认为, 其他心理语言学类别 (如性别、社会关系和成就) 受文化差异的影响更大。

收集和准备中文社交媒体数据的样本

我们首先选择中文用户泛议食品安全话题或专门讨论福喜事件的微博帖子样本。最初, 我们计划单独分析新浪微博的数据, 毕竟它是中国使用最广泛的微博平台。然而, 我们的社交媒体数据供应商 DataSift 出人意料地对访问历史数据作出限制, 导致我们只能访问从数据收集工作开始 (2014 年 9 月) 之后的

微博数据，无法抓取紧随 2014 年 7 月福喜事件爆生后的社交媒体议论。另外，微博上未被自动过滤（基于特定触发词的存在与否）的最敏感帖子，会被中国监察人员在发帖后 24 小时内手动删除。在收集数据时，DataSift 无法确保微博消息在被删除之前已经抓取。然而，鉴于福喜事件发生后在微博平台上围绕食品安全的议论不大可能包含很多会引发集体行动的帖子，更多的可能是评论事件的个人回帖，这种数据收集方法应该已抓取到大部分关于食品安全主题的微博帖子（参见 King、Pan 和 Roberts，2013 年，以及 King、Pan 和 Roberts，2014 年）。

为了补充有限的微博数据，我们纳入了原计划时间段内（2014 年 6-8 月）的 Twitter 数据样本。¹ 由于中国内地的一些异见人士翻越中国的“长城防火墙”，积极使用 Twitter 与中国内地和西方的追随者交流，我们试图将社交媒体用户（其中可能包括这些持不同政见的精英（假定是））对福喜丑闻的言论与微博上更大范围的舆论进行比较。² 但有一点要说明，为了统一起见，我们对 Twitter 和微博的样本都应用了类似的数据收集、汇总和处理步骤。

确定 Twitter 和微博的样本

我们通过查询上述时间段内使用“食品安全”或“福喜事件”等表达的推文或微博，从 Twitter 和微博上收集样本。首先，

¹ 由于数据收集方面的局限，我们无法收集 2014 年 9 月以后的 Twitter 数据，故未能与微博数据进行比较，以求证 Twitter 数据能否代表缺失的微博数据。

² 因为有地理标签的推文数量极少，我们不能肯定这些推文来自中国或反映中国网民的态度。

表 2.1
按类别分析食品安全主题的微博帖子数量

类别	2014年6月1日至 8月31日的推文总数	2014年8月26日至 9月8日的微博总数
地方政府	404	248
中央政府	1,211	270
美国企业	569	19
合计	2,184	537

我们创建了一种与食品安全或福喜事件相关的搜索词分类法，按照不同话题和主体（例如，是否提到中国地方或中央政府）或其他可能相关的主体（例如，洋快餐企业）将这些搜索词分类。表 2.1 显示了包含每个搜索词的推文或微博数量。

我们以是否提到地方政府、中央政府和美国企业为依据，将食品安全推文分类（表 2.1）。值得注意的是，有大量推文使用了“美国”一词。我们以是否提到“美国企业”为依据，将推文分类，由此测定网民对美国或美国公司的态度。

表 2.1 显示，在 2014 年 8 月和 9 月的两个星期内收集的讨论食品安全的微博。同样，这些微博也按照是否提到地方政府、中央政府和美国企业进行分类。但总体而言，特定类别的微博数量相对较少，这可能会局限相应解释的归纳范围。

梳理和汇总数据

在分析之前，必须对原始数据进行处理，确保其格式适当。由于文本限度的约束和其他特点，社交媒体数据如经过某些数据梳理步骤，更易于分析（举例见 Elson 等人，2012 年）。对于本项探索性研究，我们对中文文本进行了最低限度的数据梳理。其中一个重要步骤涉及我们使用的特定搜索词。我们发

现，简体中文版语文探索与字词计算 (SC-LIWC) 词典将“安全”一词归类为表达正面情绪。这会对我们的分析带来问题，因为当文本中出现“食品安全”（这是我们的主要查询对象之一，确实频繁出现）这一短语时，它会被拆解成“食品”和“安全”，而“安全”被归类为正面情绪词。如此一来，似乎会扭曲样本中正面与负面情绪词的比例（意味着推文比实际情况更积极），所以，我们创建了一套新词典，把“安全”一词从“正面情绪”类别中剔除。举例而言，比较 7 月 20-26 日一周之内福喜事件推文样本的 LIWC 分析，SC-LIWC 词典将 1.14% 的词语归类为表达正面情绪，而“安全”一词从正面情绪类别中剔除的经修订词典，则将 0.49% 的词语归类为表达正面情绪。

接下来，按照适当的分析层次汇总数据。考虑到所收集的社交媒体消息的时间脉络（几个月）和数量，我们选择按周汇总数据。

文本分割

中文文本的词语之间没有空格。在进行语言处理之前，我们必须在词语之间添加空格，使 C-LIWC 能够识别词语并分类。有大量文献论述将中文文本分割成词语所面临的挑战，包括现代汉语中究竟什么形式构成一个词的问题（Wong 等人，2010 年）。

我们根据 SC-LIWC 创建者的建议（Huang 等人，2012 年），使用斯坦福分词器（2014 年）分割社交媒体文本，即按周汇总的推文 / 微博。斯坦福分词器设有两种预置的分割标准：宾州中文树库（Xue 等人，2013 年）和北京大学语料库（Duan 等人，2003 年）。我们使用了宾州中文树库语料，其中包含 150 多万个词语，基础素材来自“中文的新闻通讯、政府

文件、杂志文章、各种广播新闻和广播谈话节目、网络新闻组和博客”（Xue 等人，2013 年）。

在中文社交媒体的背景下处理和解读数据

我们使用 C-LIWC 处理经过分割的文本，并对 C-LIWC 的量化输出进行定性解读。因为在网上用中文留言的人似乎大都使用简体中文，我们所用的 C-LIWC 版本以包含简体中文词典的版本 (SC-LIWC) 为基础。创建 SC-LIWC 的台湾中央研究院研究人员发现，在网上收集的 30 个关于分手和兼职的文本文件样本中，标记率超过 80%（Huang 等人，2012 年）。

然而，中国科学院大学的研究人员在分析微博帖子时发现，SC-LIWC 单独归类了大约 40% 至 50% 的词语（Gao 等人，2013 年）。为了提升分类率，这些研究人员创建了一套独立的词典，基础素材是新浪微博上最常用但尚未被 SC-LIWC 词典收录的 5,000 个词语（Gao 等人，2013 年）。添加这套补充词典后，SC-LIWC 能够将更高比例的微博帖子归类：大约 50% 至 60% 的词语。我们的样本也表现出类似的结果——例如，在 7 月 20-26 日一周之内的福喜事件样本中，SC-LIWC 词典归类了 40.9% 的词语，而增补微博专门词典的 SC-LIWC 词典分类了 54.3% 的词语。

SC-LIWC 的分类词语在样本词语总数中的占比似乎不大。相比之下，LIWC2007（即英文版语文探索与字词计算）平均可捕捉到 86% 的书面或口头词语 (LIWC.com)。由于人们会根据通信媒介使用不同的书面和口头用语，微博文本中切合 LIWC 和 SC-LIWC 词典分类的词语可能更少。原因是这些词

典以较正式或官方的文本为基础，例如，教科书、新闻文章。但是，这不应当视为缺失数据。相反，可能是在社交媒体上留言（尤其是以中文书写）的网民，他们使用的词语只有较少部分符合上述预设类别且在心理学上有意义。

通过 C-LIWC 分析 Twitter 和微博的社交媒体消息之后，我们对照相关时间段内主要事件的时间脉络描绘了趋势。我们还以社交媒体消息有无议论地方政府、中央政府或西方公司为依据，比较了 C-LIWC 指标。

在 Twitter 和微博上都存在许多重复的信息，似乎是上海市政府或本地新闻媒体关于福喜事件的报道，这可能导致我们对中国公众舆论的解读复杂化。某些情况下，重复的帖子是转发造成的，对此我们未加过滤，因为可以假定有关内容反映了转发者的意见。换言之，这些重复信息的意义在于：它们暗示那些转发或发布新闻文章的人或实体认为相关内容很重要，而且可能至少在某种程度上代表其观点。

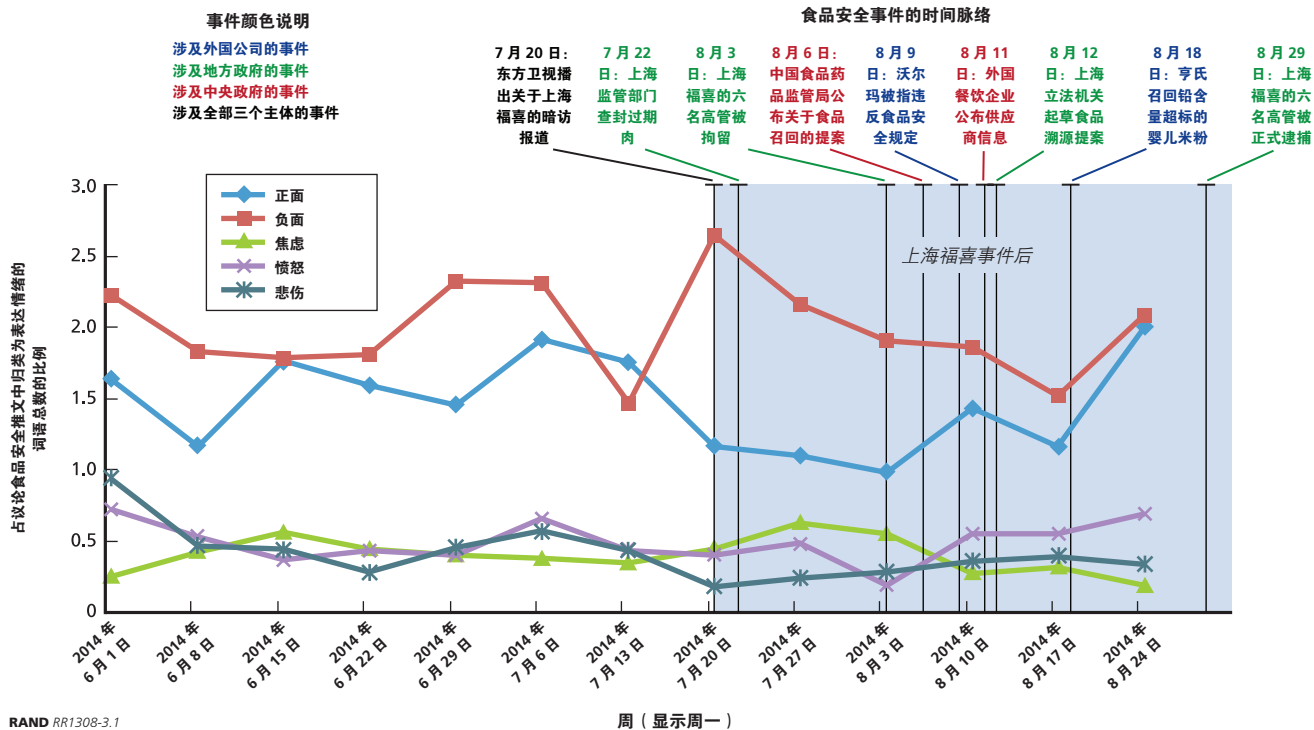
囿于手头只有收集工作开始之后的微博数据，我们没有发现中国国内微博上有很多关于福喜事件的议论。这是意料之中的，因为揭发食品加工丑闻的最初新闻报道与数据收集相隔了几个月，相关议论已逐渐平息。为了补充这个小型数据集的调查结果，我们对福喜事件发生后的 Twitter 数据执行相同的查询，并比较了两个数据集。为了解网民的态度，包括情绪、以及归咎或问责，我们运用相同的 C-LIWC 指标分析 Twitter 和微博上的议论：代词、情绪词、以及提到道德或法律责任的特定词语的使用频率。

关于食品安全和福喜事件的推文：趋势变化

首先，我们探讨了 Twitter 上关于一般食品安全讨论的态度。福喜事件后有关食品安全的舆论走向表明，中国舆论与地方和中央政府对食品安全危机的反应密切协调。图 3.1 所示的趋势表明，在福喜事件爆发的一周时间内，网民负面情绪骤升，而正面情绪下降。（负面情绪词语类别是焦虑、愤怒和悲伤等词语类别的总和。）直到 8 月 3 日那一周，负面情绪的表达持续

图 3.1

议论食品安全推文中的情绪词使用量



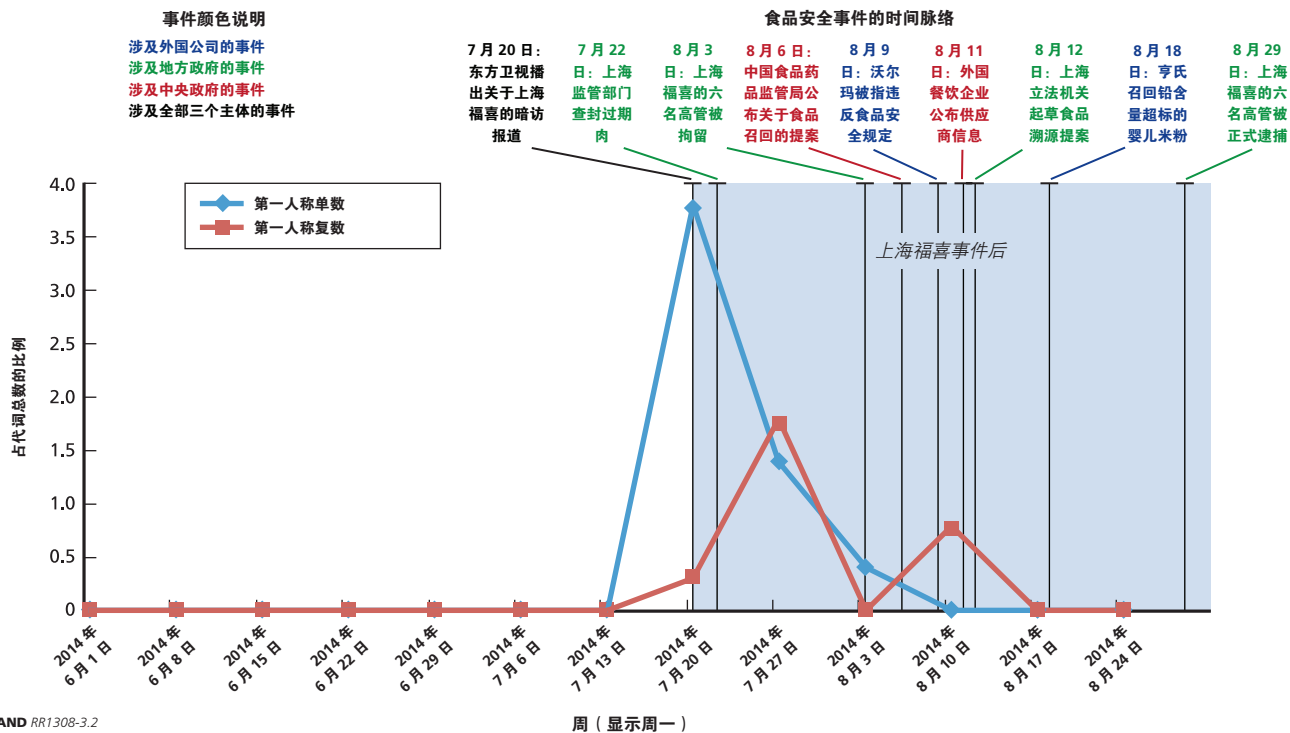
RAND RR1308-3.1

高涨——即人们的措辞愈加负面，随即上海警方宣布已拘留上海福喜的六名高管。8月11日，中央政府要求麦当劳和肯德基等外国快餐连锁企业在其中文网站上公布供应商信息时，议论食品安全的负面情绪才大致回到事发前的水平。

福喜食品安全丑闻的直接后果是，中国网民在专门讨论福喜事件的 Twitter 帖子中，第一人称单数代词的使用率似乎极高（图 3.2）。这说明人们对丑闻消息的最初反应是负面或沮丧的，该规律也反映于泛议食品安全话题（图 3.1）和专门讨论福喜事件（图 3.3）时的态度趋势。但是，到丑闻曝光后的一周内，网民对第一人称复数代词的使用率上升，这说明网民在分享对此事的痛苦和沮丧之情时萌生了一些社群意识或集体凝聚力。8月3日，上海警方逮捕上海福喜的六名高管，追究他们未能在公司肉类生产活动中落实安全常规的责任。此时，网民对第三人称复数代词的使用次数达到峰值，也许这意味着网民在“他者化”或疏远这些高管。

这种第一人称代词使用的二元模式，即第一人称单数代词和第一人称复数代词双双增加，与某些其他重大惨剧所显示的模式相似。例如，紧随 2009 年伊朗总统大选后发生的大规模抗议的相关推文（Elson 等人，2012 年），以及德州农工大学篝火事故的新闻报道（Gortner 和 Pennebaker，2003 年），其第一人称单数代词的使用量在事发后立即飙升，反映一种或暗示强烈负面或沮丧情绪的自我聚焦状态，这也在预料之中。在这两个案例中，第一人称复数代词的使用量——表示渴望与他人互动，也在事发后增加，但时间点有所不同。集体聚焦的范围和变化可能反映不同的处理方式，以及讨论参与者的社会凝聚力。

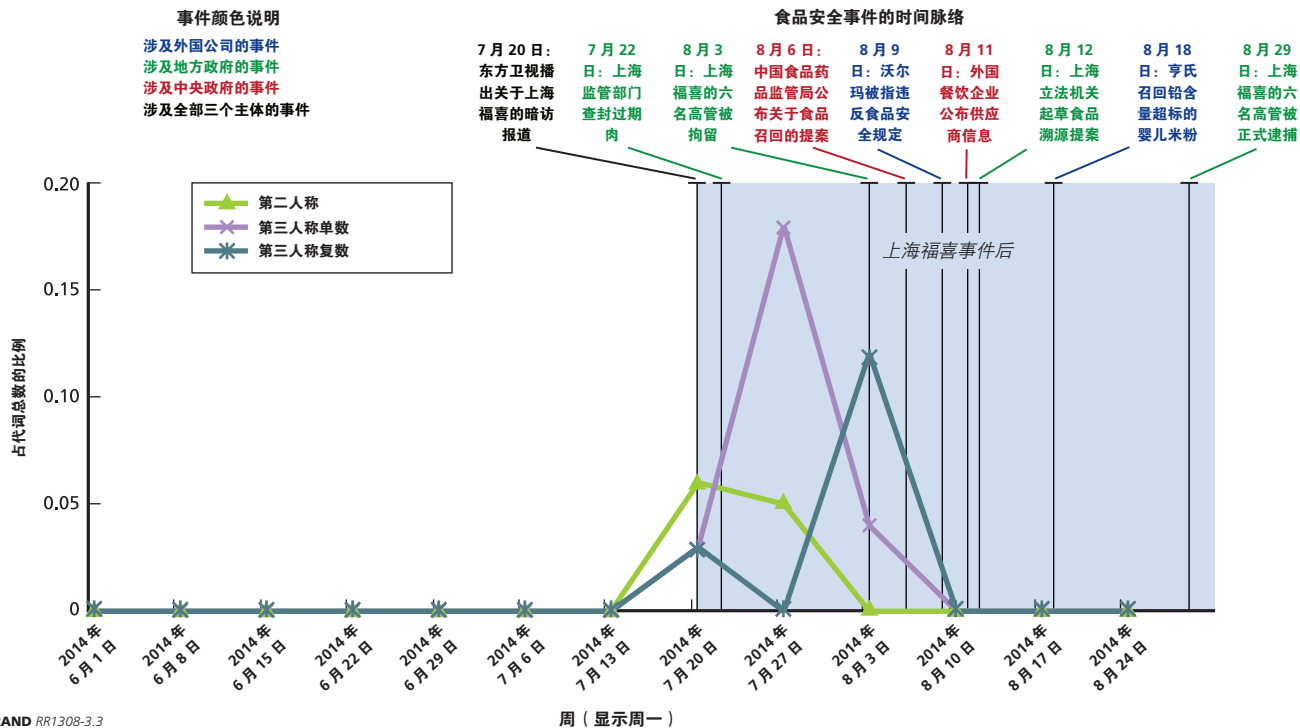
图 3.2
议论福喜事件推文中的第一人称代词使用量



RAND RR1308-3.2

图 3.3

议论福喜事件推文中的其他代词使用量



RAND RR1308-3.3

如图 3.4 所示，在 7 月 20 日丑闻曝光后，专门议论福喜事件的推文所表达的情绪，比那些议论食品安全的帖子明显更加负面。在危机之后的四周时间内，尽管中央和地方政府采取了若干行动，负面情绪仍没有显著变化。对于这些表达负面情绪的帖子所针对的目标，应当谨慎推断——可能需要仔细阅读帖子；持续高涨的负面情绪可能针对麦当劳和肯德基等美国企业，也可能针对中国政府。在此案中，八月最后一周的负面情绪骤升，可能反映网民对上海福喜六名高管的负面情绪。当局先行拘留了这些涉事高管，随后于 8 月 29 日宣布正式逮捕。

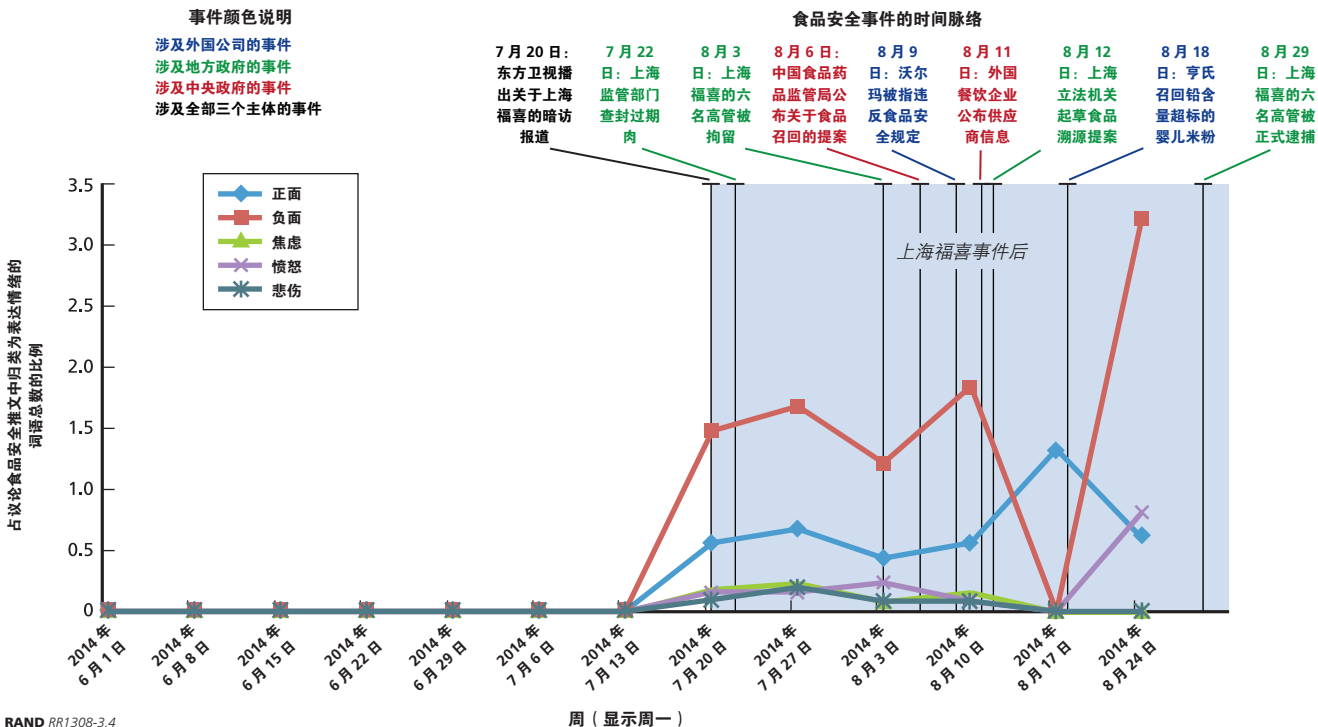
食品安全推文中关于中国政府和美国企业的议论

网民对此事的失望情绪可能较少针对地方政府，更多是针对中央政府和卷入丑闻的美国企业。值得注意的是，网民频繁使用第三人称复数代词，暗示了“他者化”或疏远中央政府和美国企业（图 3.5）。第三人称复数代词的使用通常反映了对对立团体或政府的提述（Pennebaker 和 Chung，2008 年），而在此案中，可能意味着网民将中央政府和美国企业视为食品安全的对立面或问题根源。

网民议论中央政府和美国企业的负面情绪词数量，大约是议论地方政府的两倍，或印证了上述将中央政府和美国企业视为对立面的看法（图 3.6）。特别是，网民似乎将愤怒之情指向了中央政府，而非地方政府；在提到国家和中央政府的推文中表达愤怒的词语数量，是提到省、市当局的推文的四倍以上。同样，在议论中央政府的推文中表达悲伤的词语数量，是

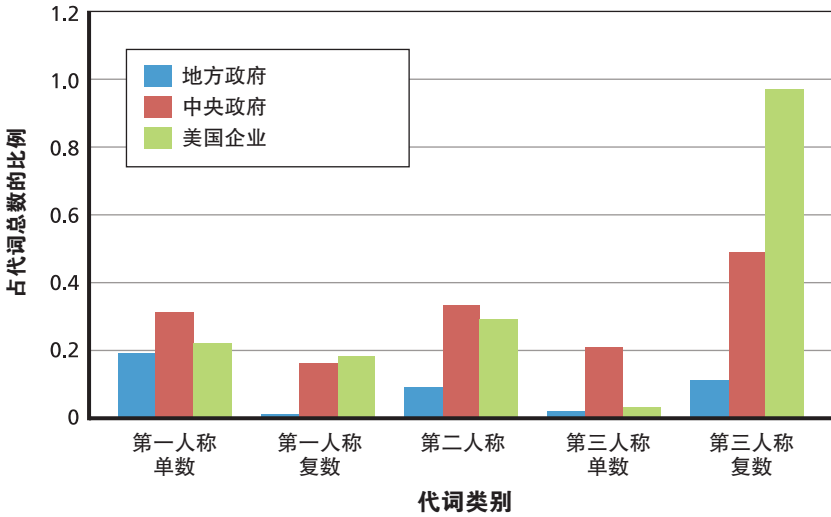
图 3.4

议论福喜事件推文中的情绪词使用量



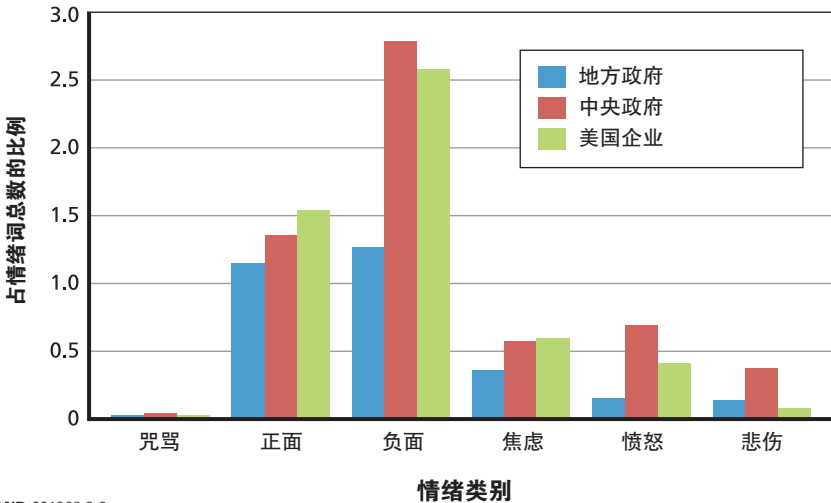
RAND RR1308-3.4

图 3.5.
在提到地方与中央政府和美国企业的食品安全推文中代词使用量的比较



RAND RR1308-3.5

图 3.6
在提到地方与中央政府和美国企业的食品安全推文中情绪词使用量的比较



RAND RR1308-3.6

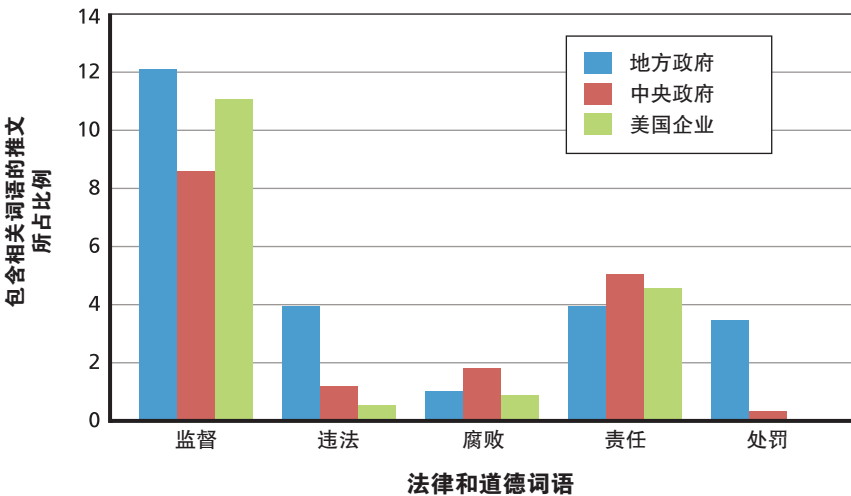
议论地方政府的推文的两倍以上。

作为探索网民归咎食品安全问题的一种方式，我们研究了网民对多个法律和道德词语的使用情况（图 3.7）。尽管整体情形不甚明朗，但网民似乎认为地方政府有责任查处违法行为，这方面工作大概是卓有成效的。另外，在关于中央政府的推文中，腐败一词出现的频率几乎是地方政府的两倍。这可能反映了人们对中央或地方层面的腐败问题的普遍态度，抑或是人们经常议论中国的腐败现象，或者习惯上将整个国家一体议论。

要厘清上述结果的具体解释，可能需要对内容作进一步的人工分析。例如，为何针对中央政府的批评看上去那么多，阅读少数推文后便可得出解释——相当一部分的推文似乎是新闻媒体发布的。这或许可以解释相关的调查结果，也可能让问

图 3.7

在提到地方与中央政府和美国企业的食品安全推文中法律和道德词语使用量的比较



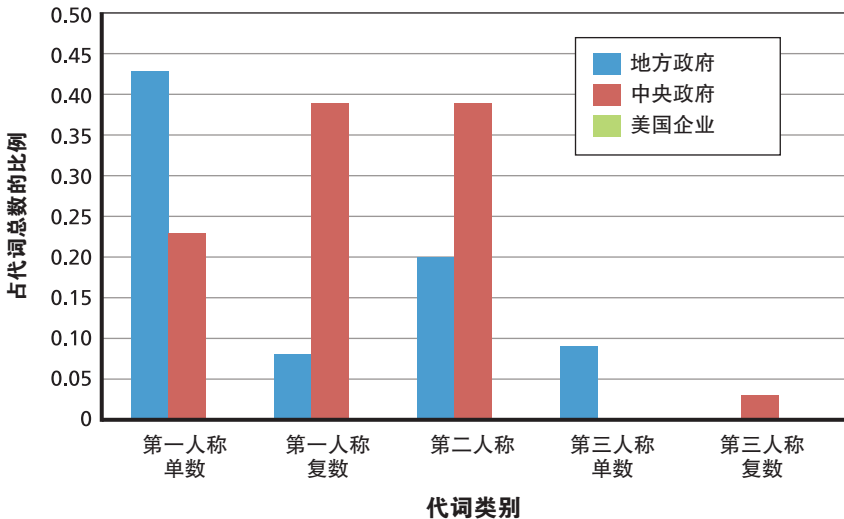
题复杂化，因为西方新闻媒体如《纽约时报》、《华尔街日报》和《卫报》的中文推文和文章，对中国中央政府的批判力度可能更甚于中国网民的帖子。

食品安全微博中关于中国政府和美国企业的议论

鉴于 Twitter 样本相关的不确定因素，譬如可能只有非常积极和或许有偏见的 Twitter 用户会“翻墙”，我们也利用微博数据研究对地方政府、中央政府和美国企业的态度。

对微博上代词使用情况的分析表明，相比地方政府，人们对中央政府有更强的社群意识，沮丧或不安感较少（图 3.8）。有趣的是，微博上对地方和中央政府的态度，似乎与 Twitter 分析的结果相反。在谈论食品安全时，微博网友似乎对中央政

图 3.8
在提到地方与中央政府和美国企业的食品安全微博中代词使用量的比较，2014 年 8 月 26 日 - 9 月 8 日



府的看法更正面，对地方政府的看法更负面，而 Twitter 网友的看法则相反。但必须注意，Twitter 数据和微博数据采自两个不同的时间段。Twitter 数据是在食品安全丑闻事前、事中和事后（从 2014 年 6 月初至 8 月底）收集的，而微博数据是在 2014 年 8 月底至 9 月初的两周时间内收集的。因此，不能直接比较同一时间段内从这两个来源收集的数据。

在议论地方政府的微博中第一人称单数代词的占比，几乎是议论中央政府的微博的两倍，说明网民在微博上议论食品安全与地方政府时，比议论食品安全与中央政府时感到更加不安。议论中央政府的微博帖子有机会、但不大可能被删除。首先，在此案中，关于食品安全的帖子不大可能涉及呼吁集体行动或公开抗议，因为这些内容绝大多数是要被删除的（参见 King、Pan 和 Roberts，2013 年和 2014 年）。其次，关于食品安全的帖子不大可能包含会被立即删除（基于关键词过滤）的词语。我们坚信收集数据时抓取的帖子是未被立即删除的，虽然 Datasift 无法给予证实。相比之下，议论国家或中央政府的微博中第一人称复数代词的占比，几乎是议论地方政府的微博的五倍，这说明网民在讨论中国整体的食品安全时，对国家的亲切感超过省、市层面。

这可能反映微博的审查制度（即批评中央政府保证食品安全不力和未作出切实回应的微博可能已被删除，故不在我们的分析范围内），或者纯粹是微博与 Twitter 的用户群差异，即 Twitter 上的异见人士或海外华人要多于微博。要注意到我们的样本中极少帖子提到美国或美国企业，这可能是由于收集帖

子的时间范围（在食品丑闻曝光后一个月之中的两周）所致。

微博网友议论中央政府（或中国整体）的内容远比议论地方政府的内容更加正面，这与 Twitter 网友形成鲜明反差（图 3.9）。不过，网民似乎仍将愤怒之情指向中央政府，而非地方政府；在提到国家和中央政府的微博中表达愤怒的词语数量，是提到省、市当局的微博的四倍以上。同样，议论中央政府的微博中表达悲伤的词语数量，是议论地方政府的微博的三倍以上。

在笼统地提到美国或具体提到麦当劳或肯德基等美国企业的微博中，更多是议论责任或谴责的帖子（图 3.10）。遗憾的

图 3.9

在提到地方与中央政府和美国企业的食品安全微博中情绪词使用量的比较，2014 年 8 月 26 日 - 9 月 8 日

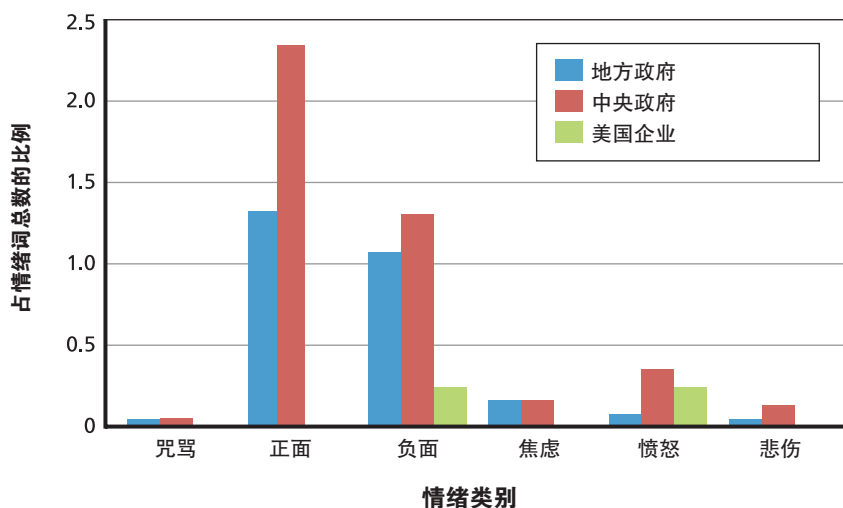
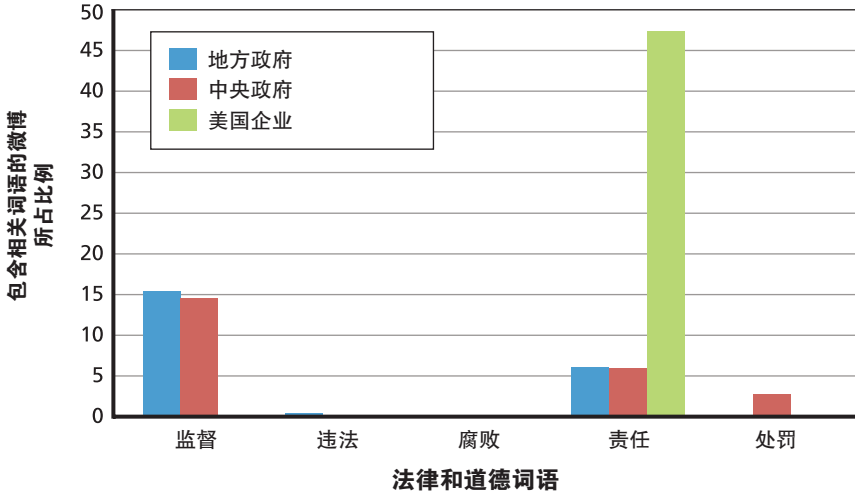


图 3.10

在提到地方与中央政府和美国企业的食品安全微博中法律和道德词语使用量的比较，2014年8月26日-9月8日



RAND RR1308-3.10

是，若不深入阅读更多的帖子，很难了解这是否表示网民更认可美国及其处理食品安全的方式，或者网民是否在指责美国或将这宗食品安全丑闻归咎于美国。

对议论福喜事件的微博进行人工内容分析后，我们发现了不少新闻媒体和官方机构发布的帖子以及个人发布的帖子，举例如下：

监督食品安全不仅是政府的责任。这需要相关行业组织召开会议，积极加强企业自律，并主动及时地逐一排查问题。同时，每个人都有责任监督社会。如果人人都能更加

关注“与己无关的事情”，并及时举报，那么食品就有可能变得安全一些。

鉴于中国食品市场的现状和国情，中国的食品安全问题需要严厉处罚和巨额罚款，否则难以杜绝各类“食品安全”事故。

第四章

结论

本研究报告探讨了网上的中文议论是否和如何揭示关于中国公众舆论的有益见解。我们以食品安全作为案例研究，通过分析中文社交媒体发现，了解公众对中国政府的态度在政策和方法论上的意义，同时也为今后分析中文社交媒体提供了借鉴。尤其是，我们运用计算机技术测定了社交媒体上的民意，结果表明网上关于国内政治的中文议论在地方与中央政府层面呈现不同的性质。

我们对在中国使用的两个社交媒体平台（即 Twitter 和新浪微博）上的议论内容开展了调查研究。与普及度较高的微博相比，我们的 Twitter 数据集更为完整，对于福喜事件的分析相关度更高，但也可能仅代表一组特定的社交媒体用户。

中文社交媒体用户对中央政府表达的不满普遍多过地方政府，但 Twitter 与微博上的看法有所不同

微博平台上中文议论所表达的情绪和其他态度指标，似乎对关于食品安全的国内政治事件非常敏感。特别是，Twitter 用户在议论中央政府时对愤怒、悲伤和腐败等相关词语的使用量比

议论地方政府时更多。虽然微博用户在议论中央政府时也会表达更多的愤怒和悲伤，但他们同时使用了更为正面的情绪词。另外，微博用户在议论中央政府时使用更多的第二人称代词，说明对广义的国家有亲切感——也许反映了爱国主义或民族主义。对中央政府的情绪表达整体上有增长，可能反映了公众中的个人意见多样性，抑或是表明公众内部存在情感冲突，或者国家利益的话题一般会引发兴趣和讨论。

表达情绪的差异也可能反映不同用户群体持有不同的观点和意见。例如，由于访问 Twitter 需要花费更多功夫，发布中文推文的用户可能在政治话题上更加活跃，更倾向批评中央政府。相比之下，目前普及度更高的微博的用户或更能代表整体舆论，至少在政治观点上如此。但是，很难分辨中央政府在微博上比较正面的印象，是名副其实的、还是通过控制中国媒体粉饰的。微博上表达的民意与中央政府对此案的关注呈对应关系，或反映中央政府设计的公关活动取得成功。将治理丑闻转移给地方官员和西方国家或公司以避开批评，符合中央政府的利益，而此案恰恰为中央政府提供了一次实践的机会。用户议论中央政府时仍然使用较多的情绪词，关于这种情况有很多可能性。首先可能是中央政府未能将这宗丑闻转移给地方政府或涉事的西方公司。另一种解释是，网民将食品安全视为头等大事，对此中央政府负有更大的立法、监管和保证责任。抑或是，网民认为在外资企业的监管方面，中央政府比地方政负有更大的责任（因为福喜丑闻的涉事公司是洋快餐连锁企业的供应商）。还有一种解释是，网民认为上海政府（甚至是洋快餐

企业)在危机发生后采取了有效的解决措施,而中央政府则未能及早发现问题。最后一个可能是,上述规律只是反映公众如何响应危机:Twitter上关于福喜事件的整体情绪大体与其他重大创伤或灾难的规律相似(例如,Elson等人,2012;Gortner和Pennebaker,2003年)。这些推文对第一人称单数和第一人称复数代词的使用率较高,分别说明网民对事件的负面反应以及在分享沮丧之情时萌生的社群意识或集体凝聚力。但是,我们只能从Twitter获得网民对福喜事件的态度(无法取得同一时间段的微博数据);因此,我们无法比较Twitter或微博上对福喜事件表达的态度。

分析中文社交媒体所面临的特有挑战

在本项对中文社交媒体的探索性分析中,我们遇到对研究结果的普遍性形成制约的多重挑战。尤其是,我们的数据对中文社交媒体用户的代表性受到限制,因为我们研究中国国内舆论时主要是分析Twitter数据,而非普及度更高的新浪微博平台。通过数据供应商DataSift访问的微博数据,仅限于数据收集工作开始之后发送的消息。这意味着我们基本无法跟踪社交媒体关于福喜事件的议论从开始至今的轨迹。相比之下,Twitter有完整的历史数据可用。但是,以Twitter作为数据源会面临推文的受众不确定的问题。因为难以确定用户位置,加上中国禁止访问Twitter,这些推文的受众可能并非中国境内居民。如果Twitter用户有意与非汉语人士或西方人士交

流，他们可能会使用英文而非中文。上述数据可用性问题的局限了我们比较 Twitter 与微博调查结果的范围。例如，我们无法直接比较福喜事件发生时的公众舆论。无法访问微博上大量的旧帖，也限制了相关社交媒体帖子的样本总数。鉴于 Twitter 和微博数据总体上不足，有必要汇总帖子（即按周汇总），但这也可能导致分析的细致程度和准确性降低，尤其是在整个数据集开始和结束的两周，期间的数据少于一整个星期的数据。最后，在少数情况下，某些社交媒体帖子似乎在数据集中重复了两三次。虽然这种重复会扭曲指标——依靠频数计算相对的词语使用量，但经过充分的隔离处理，应该不会对结果产生实质性影响。

然而，即使能够顺利访问相关数据，使用中文社交媒体（特别是微博）的数据解读个别民意，也面临其他效用受限的挑战和不确定性。例如，中国的 Twitter 用户数量和人口统计数据是靠推测得来的。中国的 Twitter 用户数量估计介于几十万到几千万（参见第一章）。这些 Twitter 用户的人口统计数据可能因为需要翻越长城防火墙而被扭曲，说明有足够动力访问 Twitter 的中国网民也可能正是对当局最不满意的群体。位置等具体的人口统计信息尤其难以确定。鲜有 Twitter 用户会在推文中显示地理标签——我们的样本中有 12 则推文附上了地理标签（总共是 3,488 则推文）。此外，即使有更多推文显示地理标签，在中国访问 Twitter 可能让用户的推文看似来自翻墙使用的代理服务器或 VPN 的位置，而非中国境内。例如，我们研究的部分推文内容清楚地表明用户位于台湾（如使

用繁体中文或内容提到台湾的地点或事件)。因此,无法肯定这些 Twitter 帖子是中国网民发布的,只能说它们是基于中文的议论。

关于中文社交媒体平台用户特有的这种模糊性,进一步延伸到使用这些平台的目的。例如,某些帖子的内容未必反映真实的民意,只代表个别社交媒体用户的意见。举一个简单的例子,新闻机构经常也发布推文或微博。因此,归纳这类帖子的信息可能无法反映确切的个人意见,不过,这类帖子在被个人转发或引用时仍可以说明民意。另一种对分析方法影响更大的可能性是,为中国政府工作的人给政府对事件的反应捧场发文(Sonnad, 2014年)。这种“僵尸帐号”是为操纵公众舆论和民意创建的,结果是很难确定实际的发帖人,以及社交媒体内容在多大程度上反映真正的民意。在我们的样本中,某些消息的确像是企图左右舆论的政府帖子。例如,下面的帖子(转发新闻报道时没有点名但要求严惩企业)重复了许多次:

韩正:在上海,不管什么企业,只要违法,都必须依法严惩。上海市委书记韩正今天下午主持召开专题会议,听取关于“福喜事件”查处情况的汇报。韩正强调,食品安全无小事,事关广大人民群众身体健康和安全,政府各监管部门必须坚持“五个最严”……

未来的研究可以扩大对中国国内政治和社交媒体平台的调查，并解决自动分析面临的挑战

未来的中国舆论研究应继续探讨在各种国内政治问题上对各级政府的政治表达，并寻求解决我们对中文社交媒体进行自动分析所发现的挑战。关于食品安全的问题，扩展用于过滤议论的关键词，以便纳入更常用的短语（如肉质或快餐），这样可以抓取到与了解网民对食品安全及其监管的态度相关的更广泛讨论。此外，未来的研究也可以考虑探索更严谨的方法，来选择以是否提到地方和中央政府为依据分类推文所用的词语；不同的词语可能得出更精确的分类，从而更准确地反映对各级政府的公众舆论。

与中国舆论相关的许多其他国内问题值得开展进一步的调查。例如，污染（如土壤、水、空气）和其他环境问题在新闻媒体和网上被广泛报道和议论。对于中央和地方在环境问题上的责任，网民可能存在意见分歧。尽管中央政府可能负有指示地方官员将经济增长置于环境保护之上的责任，网民对此或有不同的看法。例如，网民可能会为中央政府发布减少水和空气污染的命令叫好，但指责地方政府无视新法规——这是普遍做法，因为地方官员经常缺乏经济动机去监管那些向水源排放工业化学品和向大气排放有毒气体的国有企业（Schmitz, 2014年；Minter, 2014年）。即使公众指责地方政府造成环境恶化，地方性的环保抗议愈加频发——越来越多地发生在城市地区而非农村——可能会挑战中央政府的权威（Xu, 2014年）。因此，社交媒体关于污染（如雾霾）的议论，连带提到

北京、上海、成都、广州、天津等特定城市，或揭示了地方与中央在空气污染问题上的紧张关系。这类消息与空气质量、排放、工厂和工业产出等数据结合起来，有可能描绘出舆论上的地区和人口差异，由此可以找准骚乱最可能发生（和扩散）的地方。这种方法不用研究具体的危机或事件，就足以了解关于空气污染或类似国内话题的总体舆论。

另一个广泛议论的问题涉及香港泛民主运动“占领中环”期间表达的忧虑。这场运动的目的是抗议中央政府宣布香港2017年行政长官候选人要经过亲中央的选举委员会审批。香港抗议者由此产生的失望和愤怒，可以解释为希望由地方政府控制，而非由中央政府控制。通过分析 Twitter 上对地方和中央政府领导人（分别为梁振英和习近平）所表达的情绪可见，梁振英似乎是主要的不满对象（Yeung 和 Cevallos, 2014 年）。此外，对于在地理和文化上较疏远的习近平，抗议者有较大的距离感。进一步分析与香港相关的社交媒体议论，还可以探索经济贸易和移民问题。自 1997 年香港回归以来，香港与内地的贸易联系日益紧密，但香港作为资金来源和贸易中心对中国内地的重要性却随着内地经济增长而减弱。尽管如此，关于香港和内地谁更需要对方的问题一直纠缠不清，令香港居民感到不安，这也是可以理解的（《经济学人》，2014 年；Guilford, 2014 年）。举例来说，香港居民可能怪罪本港政府未能与内地保持足够的经济独立。对内地移民入港的态度也值得探究，因为这可能关涉经济顶层和底层的移民。例如，顶层的移民可能试图将资金存放到境外，而底层的移民可能只是谋求生计。任何一种情况都可能招致香港居民的反感，并引出这

样一个问题：他们认为本港政府有责任减少移民人数，还是认为中央政府有责任从另一头制止移民涌入？

对中国政治问题开展社交媒体研究，或可启发意料之外的使用方式。例如，关于中国公众如何看待政府的政策行动的信息，可用于指导进一步的行动，包括社会控制（Creemers, 2015 年）。中文社交媒体的研究人员应考虑政策制定者对研究结果的利用方式的影响。

最后，利用自动分析方法对中文社交媒体开展的进一步研究，还应寻求解决本报告中所述的部分挑战。例如，在根据社交媒体内容对特定人群下结论时，选择偏差是经常出现的重大问题。在中国，许多活跃的 Twitter 用户属于城市精英，也更可能是异见人士。另外，中国的 Twitter 用户发布的内容至少某种程度上是针对境外受众，而非大都针对国内受众。所以，这些内容或多或少存在修饰成分，而不能揭示人群真正的态度和情绪。要了解 Twitter 这类以精英为主的平台在多大程度上代表广大公众，可能需要与微博等使用更广泛的部分平台进行严谨的比较。¹

沿这条路往下的步骤可能是从 Twitter 或 Facebook 等平台获取的中国境内用户（不论地理标签在境内或自称位于境内）的历史数据，然后与类似的微博或腾讯数据集进行对比，以期发现任何跨平台差异。例如，这种对比可能有助于更细致地探索福喜事件期间的舆论态度。通过收集 9 月份（截至 16 日）的 Twitter 数据，可以比较 Twitter 和微博上的言论数量，从而

¹ 我们的研究只收集了 2014 年 8 月底短短几天同一时间内的 Twitter 和微博数据，由于数据量过少，难以确定 Twitter 数据能够代表 2014 年 7 月和 8 月缺失的微博数据。

大致推断出在危机高峰期微博上原本议论食品安全和福喜事件的帖子数量。径直对比新兴的、竞争性社交媒体平台上的政治舆论，可能也有帮助。虽然新浪微博在中国一直是占主导地位的平台，但近期也有用户弃用。对 4,500 名有影响力的微博用户（定义标准是粉丝数超过五万）的抽样调查显示，2013 年 1 月至 8 月期间，每月发帖总数下跌 20%（Chin 和 Mozur，2013 年）。微博的使用量下降，可能有几个原因。首先，中国共产党（“中共”）采取措施阻挠人们使用微博。2012 年 3 月，中共要求用户实名登记微博帐号。随后在 2012 年 6 月，中共开始逮捕数百名涉嫌在微博上散布“谣言”的用户。《电讯报》的研究显示，“继中共采取行动恫吓有影响力的用户之后，微博帖子数量锐减七成”（Moore，2014 年）。结果，新浪微博移动应用的活跃用户总数下降近 11%，到 2014 年 8 月为 6,750 万（China Internet Watch，2014 年）。微博用户也可能弃用平台，转投微信等竞争对手，后者有较私密的接口供用户与一小群朋友分享信息，被政府审查的可能性较低。鉴于微博的使用率呈下滑趋势，其他社交媒体平台的使用率或有上升趋势，这说明社交媒体在人们日常交流和生活中的重要性和作用日益增加。探索这些平台可能会得出进一步的见解和看法，尽管微信的私人平台性质或妨碍汇总大量的数据供研究之用。未来的研究可以探讨如何控制被我们识别为来自市政府、新闻媒体或其他组织的社交媒体内容，因为不能假设其反映真实的民意。

在本项探索性研究之外，更大范围的中文社交媒体分析可借助较成熟的技术检验上述研究结果的可靠性和有效性。例如，建立一个实验对照组——如通过收集福喜事件前一段时

间内的社交媒体数据，有助于直接比较 Twitter 与微博或其他社交媒体平台的公众舆论。这些额外的数据也可以作为有用的基准，据此比对词语使用量的后续变化。统计检验法，如柯尔莫诺夫 - 斯米尔诺夫检验、格兰杰因果关系检验，可帮助确定所观察到的词语使用差异（例如，事前与事后对比）是否由于偶然因素所致。另外，通过备用的数据查询构建法可能收集到不同的数据集，这有助于更可靠地评估中国舆论。举例来说，其中可能包括收集所有地理标签来自中国的推文，以及用户自称位于中国境内的推文，或者径直使用较笼统和通俗的搜索词——例如，福喜案例中的快餐或肉，而非较正式的食品安全。最后，探索议论食品安全的用户数量和特征，也有助于深入了解议论这些话题的网络人群的大概规模。社交网络分析等技术可以识别出影响力超凡的用户或用户之间的联系，并显示具有不同意见或特征的细分人群特征。

思想总结

对社交媒体言论的心理语言学分析，可以为了解全球政治背景下的公众舆论提供独特视角，这是通过其他途径不易获得的。中文微博平台上的议论似乎提供了丰富的信息，可用于探讨有关国内政治舆论的其他问题。但随着中国用户对政府打压使用社交媒体动员集体行动的措施作出回应，这些可公开访问平台的价值可能在持续下降。本项研究初步探讨了中文社交媒体上关于食品安全问题的议论，勾勒出开展此类分析的初步路线图，包括分析中文社交媒体的方法、可能性与挑战、所需的工具和步骤、以及与英语社交媒体的比较。

附录

搜索分类法

下表列出了用于收集中文社交媒体上议论食品安全、福喜事件和中国地方或中央政府等几个话题的帖子的关键词，同时还提供了包含每个词的推文数量，以粗略说明提到每个词的议论的相对数量。

表 A.1
用于收集中文社交媒体帖子的词语

类别（英文翻译）	中文词语	6月	7月	8月	推文总数
推文总数（包含“福喜事件”和/或“食品安全”）	“福喜事件”和/或“食品安全”	708	1,648	1,132	3,488
有关西方的词语					
USA	美国	48	133	73	254
McDonald's	麦当劳	0	339	99	438
Western fast food	洋快餐	0	118	8	126
KFC	肯德基	0	203	25	228

表 A.1—续

类别 (英文翻译)	中文词语	6月	7月	8月	推文总数
有关中央政府的词语					
Government (NOT province NOT city)	政府	15	109	58	182
Country	国家	108	92	98	298
China	中国	251	533	385	1,169
Leader(s)	领导	2	3	3	8
CCP	中共	3	38	2	43
Whole country	全国	74	22	8	104
有关地方政府的词语					
Province	省	15	66	69	150
City	市	72	174	172	418
Han Zheng (CCP Committee Secretary of Shanghai)	韩正	0	32	0	32
	局长	3	3	73	79
道德 / 法律词语					
Monitor and regulate/supervise	监管	17	135	81	233
Illegal	违法	16	44	20	80
Corruption	腐败	2	23	15	40
Responsibility/blame	责	20	122	50	192
Penalize	处罚	1	4	2	7
Severely punish	严惩	1	26	0	27

参考文献

Bai, Jianrui, “2013 Weibo Users Development Report” [“2013 nian weibo yonghu fazhan baogao”], *Weibo Data Center*, January 1, 2014. As of November 26, 2014: <http://data.weibo.com/report/reportDetail?id=76>

Balzano, John, “Three Things to Watch for in Chinese Food Safety Regulation in 2014,” *Forbes*, February 5, 2014. As of May 19, 2015: <http://www.forbes.com/sites/johnbalzano/2014/02/05/three-things-to-watch-for-in-chinese-food-safety-regulation-in-2014/#2715e4857a0b31102cef51df>

———, “Issues on the Horizon for Chinese Food Safety Law in 2015,” *Forbes*, March 8, 2015. As of May 19, 2015: <http://www.forbes.com/sites/johnbalzano/2015/03/08/issues-on-the-horizon-for-chinese-food-safety-law-in-2015/#2715e4857a0b3c59f5df2541>

Barboza, David, “Death Sentences in Chinese Milk Case,” *New York Times*, January 22, 2009. As of November 26, 2014: <http://www.nytimes.com/2009/01/23/world/asia/23milk.html>

Barclay, Eliza, “Tainted Pork Is Latest Food Safety Scandal in China,” NPR, April 29, 2011.

Chang, Pi-Chuan, Michel Galley, and Chris Manning, “Optimizing Chinese Word Segmentation for Machine Translation Performance,” Association for Computational Linguistics, June 2008. As of May 19, 2015: <http://nlp.stanford.edu/pubs/acl-wmt08-cws.pdf>

Chin, Josh, and Paul Mozur, “China Intensifies Social-Media Crackdown: Campaign Takes Toll on Public Debate, Popular Platform,” *Wall Street Journal*, September 19, 2013. As of November 26, 2014: <http://online.wsj.com/articles/SB10001424127887324807704579082940411106988>

China Internet Network Information Center, “The 33rd Statistical Report on Internet Development in China,” January 2014. As of November 26, 2014: <http://www1.cnnic.cn/IDR/ReportDownloads/201404/U020140417607531610855.pdf>

———, “The 35th Statistical Report on Internet Development in China,” January 2015. As of October 9, 2015:

<http://www1.cnnic.cn/IDR/ReportDownloads/201507/P020150720486421654597.pdf>

Creemers, Rogier, “Cyber China: Upgrading Propaganda, Public Opinion Work, and Social Management for the 21st Century,” *Journal of Contemporary China*, forthcoming, available via Social Science Research Network: December 2, 2015.

Duan, Huiming, Xiaojing Bai, Baobao Chang, and Shiwen Yu, “Chinese Word Segmentation at Peking University,” 2003. As of May 19, 2015:

<http://www.aclweb.org/anthology/W03-1722.pdf>

Economist, “Why Hong Kong Remains Vital to China’s Economy,” September 30, 2014. As of January 16, 2015:

<http://www.economist.com/blogs/economist-explains/2014/09/economist-explains-22>

Elson, S. B., D. C. Yeung, P. Roshan, S. Bohandy, and A. Nader, *Using Social Media to Gauge Iranian Public Opinion and Mood After the 2009 Election*, Santa Monica, Calif.: RAND Corporation, TR-1161-RC, 2012. As of May 19, 2015: http://www.rand.org/pubs/technical_reports/TR1161.html

Fewsmith, Joseph, *The Logic and Limits of Political Reform in China*, Cambridge, UK: Cambridge University Press, 2013.

Fu, King-wa, Chung-hong Chan, and Michael Chau, “Assessing Censorship on Microblogs in China: Discriminatory Keyword Analysis and Impact Evaluation of the ‘Real-Name Registration Policy,’” *IEEE Internet Computing*, Vol. 17, No. 3, May/June 2013, pp. 42–50.

Funakoshi, Minami, “Yet Another Food Safety Scandal in China—Now Rat Meat Masquerades as Lamb,” *Tea Leaf Nation*, May 5, 2013. As of November 26, 2014: <http://funakoshi1.rssing.com/browser.php?indx=8232254&last=1&item=9>

Gao, Rui, Bibo Hao, He Li, Yusong Gao, and Tingshao Zhu, “Developing Simplified Chinese Psychological Linguistic Analysis Dictionary for Microblog,” *Brain and Health Informatics*, Vol. 8211 of *Lecture Notes in Computer Science*, 2013, pp. 359–368.

Goodman, J. David, “As China Reins in Microblogs, Dissidents Find Haven on Twitter,” *New York Times*, January 23, 2012. As of January 14, 2015:

<http://thelede.blogs.nytimes.com/2012/01/23/as-china-reins-in-microblogs-dissidents-find-haven-on-twitter>

Gortner, E. M., and J. W. Pennebaker, “The Archival Anatomy of a Disaster: Media Coverage and Community-wide Health Effects of the Texas A&M Bonfire Tragedy,” *Journal of Social and Clinical Psychology*, Vol. 22, 2003, pp. 580–603.

Guilford, Gwynn, “Why China Doesn’t Feel the Need to Back Down in Hong Kong,” *Quartz*, September 29, 2014, As of January 16, 2015:
<http://qz.com/272952/why-china-doesnt-feel-the-need-to-back-down-in-hong-kong>

Huang, C. L., C. K. Chung, N. Hui, Y.C. Lin, Y. T. Seih, B. Lam, and J. W. Pennebaker, “**The Development of the Chinese Linguistic Inquiry and Word Count Dictionary**,” *Chinese Journal of Psychology*, Vol. 54, No. 2, 2012, pp. 185–201.

Internet Live Stats, “Internet Users by Country (2014),” website, 2014. As of May 19, 2015:
<http://www.internetlivestats.com/internet-users-by-country>

Jiang, Jessie, “China’s Rage Over Toxic Baby Milk,” *TIME*, September 19, 2008. As of April 17, 2015:
<http://content.time.com/time/world/article/0,8599,1842727,00.html>

Jiang, Zhan, “Environmental Journalism in China,” in Susan L. Shirk, ed., *Changing Media, Changing China*, Oxford, UK: Oxford University Press, 2010.

Johnson, Ian, “Do China’s Village Protests Help the Regime?” *New York Review of Books Blog*, December 22, 2011. As of May 19, 2015:
<http://www.nybooks.com/blogs/nyrblog/2011/dec/22/do-chinas-village-protests-help-regime>

———, *Wild Grass: Three Portraits of Change in Modern China*, New York: Pantheon Books, 2004.

Kaiman, Jonathan, “China Arrests 900 in Fake Meat Scandal,” *Guardian*, May 3, 2013. As of April 17, 2015:
<http://www.theguardian.com/world/2013/may/03/china-arrests-fake-meat-scandal>

King, Gary, Jennifer Pan, and Margaret E. Roberts, “How Censorship in China Allows Government Criticism but Silences Collective Expression,” *American Political Science Review*, Vol. 107, No. 2, May 2013, pp. 1–18.

———, “Reverse-Engineering Censorship in China: Randomized Experimentation and Participant Observation,” *Science*, Vol. 345, No. 6199, August 22, 2014, pp. 1–10.

Kuo, Lily, “KFC, Pizza Hut, and McDonald’s Are Hit with a New China Food Scandal: Expired Meat Products,” *Quartz*, July 20, 2014. As of November 26, 2014:
<http://qz.com/237643/kfc-pizza-hut-and-mcdonalds-are-hit-with-a-new-china-food-scandal-involving-expired-meat-products>

Li, Cheng, “Think National, Blame Local: Central-Provincial Dynamics in the Hu Era,” *China Leadership Monitor*, Hoover Institution, Issue 17, January 30, 2006, pp. 1–24.

Li, Haiying, Zhiqiang Cai, Arthur C. Graesser, and Ying Duan, "A Comparative Study on English and Chinese Word Uses with LIWC," Association for the Advancement of Artificial Intelligence, 2012.

Linguistic Inquiry and Word Count: The LIWC2007 Application, website. As of September 10, 2015:

<http://liwc.net/howliwcworks.php#index7>

LIWC—See Linguistic Inquiry and Word Count.

Lorentzen, Peter L., "Regularizing Rioting: Permitting Public Protest in an Authoritarian Regime," *Quarterly Journal of Political Science*, Vol. 8, No. 2, April 2013, pp. 127–158.

Lu, Rachel, "Chinese Web Users Resort to Dark Humor to Mask Fears About Pig Carcasses in Shanghai River," *Tea Leaf Nation*, March 12, 2013. As of November 26, 2014:

<http://www.tealeafnation.com/2013/03/chinese-users-resort-to-dark-humor-to-mask-fears-about-pig-carcasses-in-shanghai-river>

Millward, Steven, "8 Facts About Sina Weibo Users That All Marketers Should Know," *Tech in Asia*, November 16, 2012. As of November 26, 2014:

<http://www.techinasia.com/sina-weibo-users-facts-marketers>

Minter, Adam, "If You Thought China's Air Was Bad, Try the Water," *Bloomberg View*, April 14, 2014. As of January 14, 2015:

<http://www.bloombergtview.com/articles/2014-04-14/if-you-thought-china-s-air-was-bad-try-the-water>

Moore, Malcolm, "China Kills off Discussion on Weibo After Internet Crackdown," *Telegraph*, January 30, 2014. As of November 26, 2014:

<http://www.telegraph.co.uk/news/worldnews/asia/china/10608245/China-kills-off-discussion-on-Weibo-after-internet-crackdown.html>

Nathan, Andrew J., "Authoritarian Resilience," *Journal of Democracy*, Vol. 14, No. 1, January 2003, pp. 6–17.

Ng, Jason Q., "Where Do Weibo Users Live? City and Provincial Breakdown of Various Chinese Internet Statistics," March 1, 2013. As of November 26, 2014:

<http://blockedonweibo.tumblr.com/post/44289028375/where-do-weibo-users-live-city-and-provincial>

———, "There Are Not Millions of Twitter Users in China: Supporting @ooof's Result and Refuting GWI's Conclusion," January 6, 2014. As of November 26, 2014:

<http://blockedonweibo.tumblr.com/post/39828699303/there-are-not-millions-of-twitter-users-in-china>

O'Brien, Kevin, and Lianjiang Li, *Rightful Resistance in Rural China*, Cambridge, UK: Cambridge University Press, 2006.

Olesen, Alexa, "China Eases Censorship On Food Safety Muckraking," Associated Press, May 16, 2011. As of May 19, 2015:

http://www.huffingtonpost.com/2011/05/16/china-press-freedom-food-safety_n_862448.html

Ong, Josh, "Report: Twitter's Most Active Country is China (Where It Is Blocked)" *The Next Web*, September 26, 2012. As of November 26, 2014:

<http://thenextweb.com/asia/2012/09/26/>

surprise-twitthers-active-country-china-where-blocked

Pennebaker, James W., and Cindy K. Chung, "Computerized Text Analysis of al-Qaeda Statements," in Klaus Krippendorff and Mary Angela Bock, eds., *A Content Analysis Reader*, Thousand Oaks, Calif.: Sage, 2008, pp. 453–466.

Pennebaker, J. W., C. K. Chung, M. Ireland, A. Gonzales, and R. J. Booth, "The Development and Psychometric Properties of LIWC2007," Austin, Texas: LIWC, Inc., 2007.

Pew Research Center, "Environmental Concerns on the Rise in China: Many Also Worried About Inflation, Inequality, Corruption," September 19, 2013. As of April 17, 2015:

<http://www.pewglobal.org/2013/09/19/>

environmental-concerns-on-the-rise-in-china

Reuters, "China Regulator Says Scandal-Hit Food Supplier Forged Production Dates," July 26, 2014. As of November 26, 2014:

<http://www.reuters.com/article/2014/07/26/>

china-food-investigation-idUSL4N0Q103T20140726

Roberts, Hal, Ethan Zuckerman, Jillian York, Robert Faris, and John Palfrey, "International Bloggers and Internet Control," Berkman Center for Internet and Society, August 2011.

Robinson, David, Harlan Yu, and Anne An, "Collateral Freedom: A Snapshot of Chinese Internet Users Circumventing Censorship," New York: Open Internet Tools Project, May 21, 2013.

Russell, Jon, "No, Facebook Does Not Have 63.5 Million Active Users in China," *The Next Web*, September 28, 2012. As of May 19, 2015:

<http://thenextweb.com/asia/2012/09/28/no-way-jose>

Schlaeger, Jesper, and Min Jiang, "Official Microblogging and Social Management by Local Governments in China," *China Information*, Vol. 28, No. 2, 2014, pp. 189–213.

Schmitz, Rob, "China's Fight for Cleaner Air," *Marketplace*, July 15, 2014. As of January 14, 2015:

<http://www.marketplace.org/topics/sustainability/we-used-be-china/chinas-fight-cleaner-air>

Shirk, Susan, "Changing Media, Changing China: Introduction," in Susan L. Shirk, ed. *Changing Media, Changing China*, Oxford, UK: Oxford University Press, 2011, pp. 1–37.

Sina Weibo, "Weibo Reports Second Quarter 2014 Financial Results," press release, August 14, 2014. As of November 26, 2014:
<http://ir.weibo.com/phoenix.zhtml?c=253076&xp=irol-newsArticle&ID=1958713&highlight>

Sina Weibo Data Center, "2012 Sina Weibo Users Development Survey Report" ["2012 nian xinlang weibo yonghu fazhan diaocha baogao"], October 2012. As of November 26, 2014:
<http://vdisk.weibo.com/s/AEdhAAT9K9k7>

Sonnad, Nikhil, "Hacked Emails Reveal China's Elaborate and Absurd Internet Propaganda Machine," *Quartz*, December 18, 2014. As of January 14, 2015:
<http://qz.com/311832/hacked-emails-reveal-chinas-elaborate-and-absurd-internet-propaganda-machine>

"Stanford Word Segmenter," Version 3.4.1, August 27, 2014. As of May 19, 2015:
<http://nlp.stanford.edu/software/segmenter.shtml>

Sullivan, Jonathan, "A Tale of Two Microblogs in China," *Media, Culture and Society*, Vol. 34, No. 6, September 2012, pp. 773–783.

Tkacheva, Olesya, Lowell H. Schwartz, Martin C. Libicki, Julie E. Taylor, Jeffrey Martini, and Caroline Baxter, *Internet Freedom and Political Space*, Santa Monica, Calif.: RAND Corporation, RR-295-DOS, 2013. As of January 14, 2016:
http://www.rand.org/pubs/research_reports/RR295.html

Wong, Kam-Fai, Wenjie Li, Ruifeng Xu, and Zheng-sheng Zhang, *Introduction to Chinese Natural Language Processing*, San Rafael, Calif.: Morgan and Claypool Publishers, 2010.

World Bank, "GDP per capita, PPP (Current International \$)," undated database. As of May 19, 2015:
<http://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD>

Xiao, Qiang, "The Rise of Online Public Opinion and Its Political Impact," in Susan L. Shirk, ed., *Changing Media, Changing China*, Oxford, UK: Oxford University Press, 2011, pp. 202–224.

Xu, Beina, "China's Environmental Crisis," Council on Foreign Relations Backgrounder, April 25, 2014. As of January 14, 2015:
<http://www.cfr.org/china/chinas-environmental-crisis/p12608>

Xue, Nianwen, Xiuhong Zhang, Zixin Jiang, Martha Palmer, Fei Xia, Fu-Dong Chiou, and Meiyu Chang, Chinese Treebank 8.0 LDC2013T21, web download, Philadelphia: Linguistic Data Consortium, 2013. As of May 19, 2015:
<https://catalog.ldc.upenn.edu/LDC2013T21>

Yeung, Douglas, and Astrid Stuth Cevallos, "The Mountains Are High and the Emperor Is Far Away," *Foreign Policy*, November 11, 2014. As of May 19, 2015: <http://foreignpolicy.com/2014/11/11/the-mountains-are-high-and-the-emperor-is-far-away>

Zhai, Ivan, "China May Only Have 18,000 Active Twitter Users: Infographic," *South China Morning Post*, January 3, 2013. As of November 26, 2014: <http://www.scmp.com/comment/blogs/article/1119055/china-may-only-have-18000-active-twitter-users-infographic>