

# 媒体议程对个人议程的影响研究

## ——基于Twitter平台的大数据文本挖掘与主题建模的进路

易红发

(上海大学 新闻传播学院, 上海 200072)

**【摘要】** 已历经半世纪的议程设置研究面临新的挑战。随着互联网和社交媒体的发展,受众个人之间的差异性和互动性增强,受媒体影响的公众议程再难成为社会共识,个人层面的议程设置成为重要研究取向。本研究明确界定了“个人议程”概念,并基于2015年Twitter全平台抽样的7177.2万条推文和主题建模技术,进行了媒体议程与个人议程关系的实证研究。研究发现:(1) Twitter媒体呈现最多的是“体育赛事”议题,大多数Twitter个人用户都有特别关注的议题;(2) 大多数个人议程与公众议程没有一致性,人内议题显要性内容和人际议题显要性内容高度相关,“个人议程”概念具有实证的合理性;(3) Twitter媒体议程与30.3%的Twitter个人议程显著正相关。研究同时讨论了个人议程设置效果的强弱、个人层面的偶发条件研究以及个人议程设置与公众议程设置的关系。

**【关键词】** 个人议程; 公众议程; 议程设置; 个人差异; 主题建模

**【中图分类号】** 210.9

**【文献标识码】** A

### 一、引言

社会公共生活中存在许多有待解决的问题。其中,哪些问题最重要,应该被优先解决?关于这一问题,每个人都有一个无形的“议程”,该“议程”记录了个人对当前社会公共议题显要性的判断以及解决的优先顺序的认识。那么,个人的这种判断和认识究竟来自何处?麦库姆斯和肖推动的议程设置研究提供了一种解释:受众对社会公共议题显要性的判断和认识,通常来自新闻媒体(McCombs & Shaw, 1972)。议程设置研究将个人对议题显要性的判断和认识进行加总,形成公众议程,分析公众议程与媒体议程的一致性程度,为解释媒介对受众认知的影响做出了重大贡献。

**【作者简介】** 易红发,上海大学新闻传播学院讲师,上海大学上海市舆情监测与分析中心研究员。

**【基金项目】** 国家社科重点研究项目“传播秩序视野下的网络强国战略研究”(15AZD076)。

然而，随着互联网和社交媒体的发展，“公众议程”概念的适用性问题逐渐显现出来。一方面，社交媒体的发展使得受众个人之间的互动性大大增强，议题的显要性可能不仅仅体现于个人判断，还可能体现在个人与他人的讨论和互动之中；而“公众议程”概念无法涵盖互动内容（如转发和评论），只关注个人自己的判断和认识。另一方面，公众议程作为对个人判断和认识的加总，忽视了个人差异；而互联网和社交媒体平台上的个人差异逐渐凸显出来，“受众分化”“个性化”“长尾效应”越来越明显。关键在于：个人差异过大可能使得公众议程只能代表小部分受众对议题显要性的认识，对议题显要性的共识将不复存在。

互联网和社交媒体平台上的议程设置研究需要转变思路，应该进行“真正个人层面的议程设置研究”（Guo, 2017: 10），直接考察媒体议程对“个人议程”的影响。这类研究着重分析的是媒体通过对相关议题的报道可以显著影响多少数量和多少比例的受众个人，并且探究在什么条件下媒体议程能够显著影响“个人议程”。媒体议程对“个人议程”影响的研究能够指导新闻媒体影响尽可能多的受众，推动共识的形成。基于此，本研究将在清晰界定“个人议程”的基础上，以Twitter这一国际主流社交媒体为研究平台，以Twitter2015年全年抽样的7177.2万条推文为研究数据，利用文本挖掘与主题建模技术，分析Twitter上媒体议程和“个人议程”的关系，以期拓展议程设置理论和助于媒体“共识形成”能力的发挥。

## 二、“个人议程”概念与研究问题

### （一）公众议程与个人差异

半个世纪以来，议程设置研究已形成极其复杂的体系，比如三种研究传统：公众议程设置（public agenda setting）、政策议程设置（policy agenda setting）和媒体议程设置（media agenda setting）（迪林、罗杰斯，2009：22）；三个研究层面：议题议程设置（issue agenda setting）、属性议程设置（attribute agenda setting）和网络议程设置（network agenda setting）（McCombs, 2014: 55）。在所有取向中，最基础、最核心的部分仍然是媒体议程对公众议程影响的研究，即公众议程设置研究。

公众议程设置研究起源于麦库姆斯和肖1968年的教堂山研究（McCombs & Shaw, 1972），该研究第一次证实了媒体议程和公众议程的显著相关。随后，麦库姆斯等人继续完成了夏洛特研究（Shaw & McCombs, 1977）和1976年的总统选举研究（McCombs et al., 1981），证实了媒体议程确实显著影响了公众议程，即媒体为公众设置议程。自此之后，议程设置理论逐渐成型，麦库姆斯等人的研究方法（媒体内容分析结合受众调查）也成为议程设置研究的基本方法。

公众议程设置研究一直伴随着各种批评，一个核心问题是公众议程的概念与测量。Cobb和Elder以及迪林和罗杰斯将议题界定为“具有冲突性且被披露的社会公共问题”，将议程界定为“某个时点上议题显要性的等级排列”。因此，媒体议程是“媒体报道中议题显要性的等级排列”；公众议程是“公众中议题显要性的等级排列”（迪林、罗杰斯，2009：2-7）。Kosicki批评公众议程是一种“在个人层面测量，在总体层面分析”的概念（Kosicki，1993）。公众作为一个总体性存在，其议题显要性无法直接测量，麦库姆斯等人的做法是：先测量每个个人的议题显要性，然后将这些个人议题显要性相加，得到公众议程。比如在教堂山研究中，麦库姆斯和肖首先询问受访者他们认为的关键议题有哪些，然后将这些议题按数量加总并排序（McCombs & Shaw，1972：181）。

问题在于，由个人议题显要性加总得到的公众议程，其实质是一种平均人议程。假定我们进行了一次受众调查，随机抽取5位受访者，询问他们对5个议题的显要性认知（取值为1~5，数字越大议题越重要），得到如表1的调查结果。公众议程设置研究一般会将5位受访者的个人议题显要性加总为公众议程：议题A至议题E的显要性为[17, 16, 15, 14, 13]；然后再计算媒体议程和公众议程的相关系数。早期议程设置研究一般采用斯皮尔曼等级相关系数（*Spearman's  $\rho$* ），近年来的研究（如Vu et al.，2014）也采用皮尔逊相关系数（*Pearson's  $r$* ）。其中，样本量为常数，我们将公众议程中的议题显要性与样本量相除，就会得到平均人议程，各议题显要性为[3.4, 3.2, 3.0, 2.8, 2.6]，与公众议程中的议题显要性一一对应；公众议程成了n倍的平均人议程。研究结果上，平均人议程与媒体议程的相关系数等于公众议程与媒体议程的相关系数。

表1 公众议程与平均人议程

	议题 A	议题 B	议题 C	议题 D	议题 E
受访者 1	5	4	3	2	1
受访者 2	4	3	2	1	5
受访者 3	3	5	2	4	1
受访者 4	1	2	3	4	5
受访者 5	4	2	5	3	1
公众议程	17	16	15	14	13
平均人议程	3.4	3.2	3.0	2.8	2.6

公众议程的概念名称使人误以为公众议程是所有受众的总体议程。而实际上，这种个人议题显要性加总的公众议程只是一个议程（平均人议程）的重复。平均人议程的关键问题在于忽视了个人差异，不足以代表所有公众。其实媒体议程也有加总过程，一般

是由多个媒体的议题显要性加总而成，麦库姆斯曾解释过多个媒体的议题显要性能够加总为媒体议程的关键在于：不同媒体的报道具有相似性（McCombs, 1981）。然而，有研究者质疑个人议题显要性能否加总为公众议程（Becker, 1991）。如果个人之间存在较大差异，那么实质为平均人议程的公众议程只能代表极少部分的受众个人，无法代表整体公众。

假定个人议题显要性服从正态分布，实质为平均人议程的公众议程处于均值点上。我们设定平均人议程可以代表邻近1个单位的个人议题显要性，如图1，在标准正态分布下， $y \sim N(\mu=0, \sigma^2=1)$ ，此时平均人议程可以代表大部分的受众个人（68.27%）。然而，当个人差异增大时，如 $\sigma \rightarrow 2$ 时， $y \sim N(\mu=0, \sigma^2=4)$ ，此时平均人议程只能代表38.29%的受众个人。当个人差异继续增大，如 $\sigma \rightarrow 3$ ，平均人议程仅能代表26.11%的受众个人。当平均人议程无法代表大部分受众个人时，即使媒体议程能够显著影响平均人议程（公众议程），也还有大部分的受众个人不受影响。

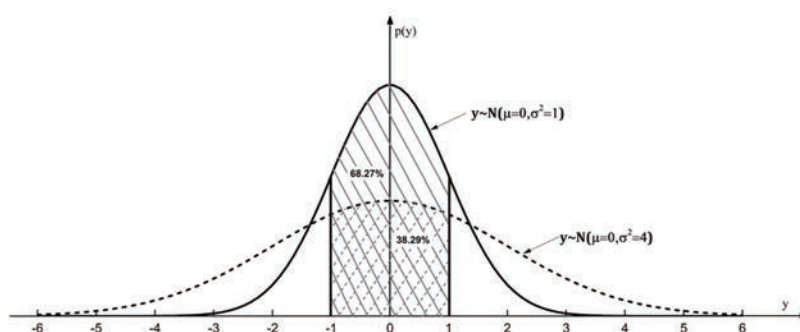


图1 平均人议程（公众议程）与个人差异

我们现在很难考察20世纪七八十年代受众个人之间的差异程度，或许传统媒体时代的个人差异并不大，由此当时的议程设置研究忽视个人差异，将个人议题显要性加总为公众议程有一定的合理性。然而，随着互联网的兴起，特别是社交媒体的飞速发展，个人差异已到了研究者难以忽视的程度。已有诸多概念从不同角度表述过个人差异，如长尾效应（安德森，2012）、个体异质性（Hausman & Newey, 2016）、分众化（靖鸣、臧诚，2011）、碎片化（廖圣清等，2015）。社交媒体时代，个人议题显要性认知差异的一个重要表现是：部分社交媒体用户对某些议题有特殊兴趣（McCombs et al., 2014: 793）。当有特殊兴趣的用户比例越来越高，甚至形成兴趣群体时，媒体的影响就变得越来越小。公众议程设置研究由于忽视个人差异，难以反映这种影响的连续性变化。

## （二）社交媒体时代“个人议程”概念的界定

早在1972年，麦库姆斯和肖就已经意识到公众议程设置研究忽视个人差异的缺陷，

虽然初始研究是对加总数据的分析，但他们认为“后续研究应该从宏大社会层面转移到社会心理层面”（McCombs & Shaw, 1972: 184）。为缓和这一问题，议程设置的偶发条件研究中考察了个人差异，如导向需求（Camaj & Weaver, 2013）、个人经验（Zucker, 1978）的影响。然而，这类研究其实是将受众划分成几个子群（如“高导向需求者”“低导向需求者”），实际上还是加总数据的分析（Guo, 2017: 10）。第三层面议程设置的主要倡导者郭蕾认为：需要进行“真正个人层面的议程设置研究”（Guo, 2017: 10），直接考察媒体议程对“个人议程”（individual agenda）的影响。

以阿卡普尔科模型（Acapulco typology）而言，如表2，媒体议程对“个人议程”的影响研究属于角度二（“自动”）。角度二的研究中，各议题竞争的并非总体层面公众的显要性，而是个人层面的显要性。与角度一研究（公众议程设置）检验媒体议程与公众议程的一致性不同的是，角度二研究的主要目的是：媒体议程显著影响了多少个人议程。欲进行媒体议程对个人议程的影响研究，我们首先需要明确界定“个人议程”概念。

表2 阿卡普尔科模型（McCombs, 2014）

议题显要性的测量		
	总体层面	个人层面
关注点	角度一 竞争	角度二 自动
	角度三 自然历史	角度四 认知画像

“个人议程”一词已在个别研究中有所提及（如Kosicki, 1991; Roessler, 1999; Arguelhes & Hartmann, 2017），但尚并未被明确界定。从议题和议程的角度来说，“个人议程”应该是“个人议题显要性的等级排列”。相对于公众议程对个人议题显要性的加总，“个人议程”没有加总的过程，并不会形成平均人议程；“个人议程”恰当地反映了不同受众个人对议题显要性的认知差异。

关于个人议题显要性，其实公众议程中的个人议题显要性是一种人内议题显要性。人内议题显要性来自人内认知，即不管别人怎么说，受众个人认为哪些议题最重要。教堂山研究中，麦库姆斯和肖通过以下问题测量议题显要性：“不管候选人怎么说，你自己认为关键议题有哪些？”（McCombs & Shaw, 1972: 178）后来，大量议程设置研究采用盖洛普民意调查测量公众议程，由此形成人内议题显要性测量的一般性提问方法（most important problem, MIP）：“你个人认为今天国家面临的最重要的问题是什么？”（Yeager et al., 2011）



然而，社交媒体时代“个人议程”中的议题显要性并不完全是人内议题显要性。社交媒体平台上个人用户发表的原创内容，如微博平台中的原创博文、Twitter中的原创推文，确实属于人内议题显要性。但社交媒体上个人用户之间的互动内容，如对原创内容的转发和评论，并不是个人的人内认知，不在人内议题显要性的范畴之内。并且，互动是社交媒体平台的核心特征之一，互动内容占据了社交媒体内容的相当比重。社交媒体上的议程设置研究中，研究者无法忽视互动内容的存在，不能只利用原创内容测量人内议题显要性；社交媒体时代的“个人议程”概念中的个人议题显要性不能仅仅包含人内议题显要性。

事实上，1974年麦克劳德等人就已经发现受众个人的议题显要性包含人内议题显要性和人际议题显要性两个维度（McLeod et al., 1974）。他们认为议程设置存在两种源头：一种来源于李普曼对“我们头脑中图画”形成的描述（Lippmann, 1922）；另一种源自帕克对新闻和媒体功能的分析（Park, 1925、1940）。李普曼认为媒体作为外部世界的反映，影响的是人们对外部世界的认知；帕克认为媒体是共同体的组成部分，影响的是人们对“新闻所反映的议题的讨论”，“新闻是某种使得人们交谈的事物”（Park, 1940: 677）。换言之，李普曼认为的重要议题来自个人对外部世界的认识，帕克认为重要的议题来自共同体内个人与他人的讨论。据此，麦克劳德等人总结：人内议题显要性由个人独立确定，议题显要性来自个人需求、欲望或想法；而人际议题显要性是共同体成员通过对议题的讨论而确定；并且他们将人内议题显要性和人际议题显要性相加的总显要性称为“感知议题显要性”（McLeod et al., 1974: 9）。后来，Gadziala和Becker给出了测量人际议题显要性的一般提问方法：“在国家面临的所有问题和议题中，上周你和你的朋友讨论最多的是什么？”（Gadziala & Becker, 1983）。

麦库姆斯也曾明确区分了人内议程（intrapersonal agenda）和人际议程（interpersonal agenda）；他认为人内议程的主要内容是个人认为重要的议题，人际议程的主要内容是个人与他人讨论最多的议题（McCombs, 1977: 91）。麦库姆斯承认个人议题显要性包含人内议题显要性和人际议题显要性两个维度；但他认为二者不能相加为总的议题显要性，原因在于二者存在差异：“许多个人认为的重要议题可能并不出现在他与朋友或家人的讨论中”（McCombs, 1977: 91）。

随着媒介环境的变化，麦库姆斯在20世纪70年代描述的人内议题显要性和人际议题显要性的差异可能在缩小。明显地，社交媒体上个人用户之间的互动内容（转发、评论等）属于人际议题显要性的范畴。已有研究表明，Twitter上的转发、评论和原创内容在主题（议题）上并没有显著差异（Qu et al., 2015）；麦库姆斯本身也承认人内议程和人际议程存在“可观的重叠”（McCombs, 1977: 91）。并且，从社交媒体的特性上说，社交媒体上的互动是在原创内容基础上的互动；如果没有用户的原创内容，转发和评论无法独立存在。社交媒体上的人际议题显要性并不独立于人内议题显要性，二者

并不存在本质上的差异。另外，社交媒体上的人内议题显要性内容（原创内容）和人际议题显要性内容（转发、评论等）具有相似的形式（如微博博文、Twitter推文），二者可以采用相同的测量方法。

总之，社交媒体时代的“个人议程”应该包括人内议题显要性和人际议题显要性两个维度，个人用户发表的原创内容从属于人内议题显要性；个人用户之间的互动内容从属于人际议题显要性。

相较于总体层面的公众议程，如图2，“个人议程”没有加总过程，是一种个人层面的概念，有助于分析不同受众个人之间的差异性；相较于公众议程概念只包含人内议题显要性，“个人议程”概念同时包含人内议题显要性和人际议题显要性两个维度。感知议题显要性由麦克劳德等人提出，包含人内和人际两个维度，与个人议程有相通之处（McLeod et al., 1974）；但感知议题显要性的测量仍然有加总的过程，是一个总体层面的概念，不能等同于“个人议程”。

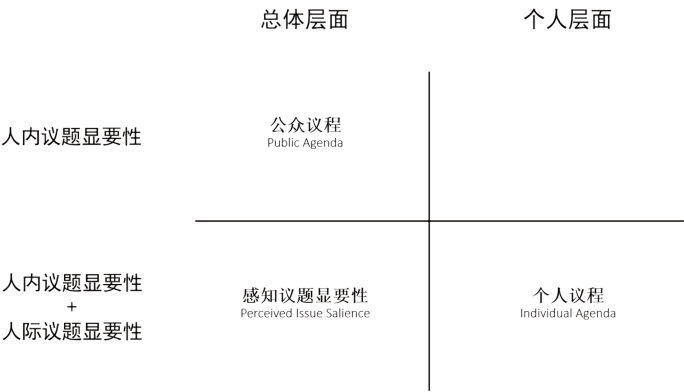


图2 “个人议程”的相关概念

（三）研究问题与研究假设

我们已经明确界定了“个人议程”：个人议题显要性的等级排列。社交媒体时代的个人议题显要性包含人内议题显要性和人际议题显要性两部分。在明确界定个人议程概念后，接下来我们可以进行“真正个人层面的议程设置研究”（Guo, 2017: 10），即媒体议程对个人议程的影响研究，或称“个人议程设置研究”。

下文我们将以Twitter这一国际主流社交媒体平台为研究对象，对个人议程设置进行实证研究。社交媒体时代个人议程设置研究有两个立论基础：个人差异和双维度的个人议题显要性。上文在界定“个人议程”概念的过程中，我们已经从理论概念和现实问题两个角度进行了阐述，下文我们需要通过实证研究进行验证。公众议程是一种总体概念，实质是一种平均人议程，忽视了个人之间的差异；社交媒体时代个人差异的扩大（McCombs et

al., 2014: 793) 导致公众议程难以代表大多数的个人议程; 换言之, 社交媒体时代的大多数个人议程与公众议程不再具有一致性。由此, 我们提出第一个假设:

H1: Twitter上的多数个人议程与公众议程不存在显著正相关。

关于双维度的个人议题显要性。从理论上说, 麦克劳德等人(1974)和麦库姆斯(1977)均已描绘了人内议题显要性和人际议题显要性两个维度; 从现实来看, 社交媒体, 如Twitter中的原创内容属于人内议题显要性, 转发、评论等互动内容属于人际议题显要性。问题的关键在于传统媒体时代的人内议题显要性和人际议题显要性存在明显差异(McCombs, 1977: 91), 因此二者不能相加; 而社交媒体上的原创内容与互动内容在议题上可能没有显著差异, 二者可以相加, 因此我们需要验证:

H2: Twitter上的人内议题显要性和人际议题显要性高度相关。

与公众议程设置研究主要目的在于检验媒体议程与公众议程的一致性不同, 个人议程设置研究主要关注的问题是: 媒体议程能够显著影响多少个人议程? 在Twitter平台的实证研究中, 我们首先需要考察Twitter媒体议程和Twitter个人议程的分布状况, 比如Twitter上媒体呈现的主要议题, Twitter个人用户关注的主要议题。媒体议程和个人议程关系研究包括关系方向和关系强度两个方面。关系方向方面, 现有的以Twitter平台的议程设置研究一般视媒体为影响者, 受众为被影响者(Nalwoga, 2017); 从现实来看, Twitter中的普通个人用户可能为媒体账号提供个别的消息来源, 但很难影响整个媒体议程。本研究将主要关注Twitter媒体议程和Twitter个人议程的关系强度; 二者关系强度越高, 媒体议程更可能影响个人议程。基于上述内容, 我们提出个人议程设置研究的核心问题:

Rq1: Twitter媒体议程与多少Twitter个人议程存在显著正相关?

### 三、研究方法

#### (一) 数据来源

本研究数据来自复旦大学自然语言处理实验室, 原始数据通过Twitter API获取, 数据获取时间为2015年10月至2016年3月。数据获取中, 研究者首先随机选择10个种子用户, 通过关注账号的不断迭代, 共获取406.7万位用户; 随后按6%的比例随机抽取24.4万名用户, 获取这些用户的全部推文, 共3.45亿条。所得数据的时间范围为2006年3月至2016年3月, 为避免议题过于分散, 本研究选取2015年的全年推文作为研究数据, 共计7177.2万条。每条推文的字段包含两部分: 文本字段和用户字段。主要文本字段如推文内容、转发数、点赞数; 主要用户字段如用户名、用户描述、粉丝数、朋友数。



## （二）概念测量

本研究测量的主要概念包括：Twitter媒体议程、Twitter个人议程和Twitter公众议程。

Twitter媒体议程测量的是Twitter媒体账号的内容。本研究所涉及的媒体账号是指专业媒体组织的Twitter账号，不包括新闻从业者个人账号。本研究在识别媒体账号时主要基于用户的自我描述和Twitter官方认证。对Twitter媒体议程的测量主要基于Twitter媒体账号对某一议题的报道数量。研究中我们先计算每个媒体对相关议题报道的频数，然而对所有被纳入分析的媒体进行累加，对累加数据进行等级排列，得到Twitter媒体议程。

Twitter个人议程和Twitter公众议程的测量互有关联。Twitter公众议程是所有个人用户的人内议题显要性的加总；Twitter个人议程是每个个人用户的人内议题显要性和人际议题显要性。人内议题显要性来自个人的人内认知；人际议题显要性主要来自人际讨论。Twitter平台上的人内议题显要性主要体现为Twitter个人用户发表的原创推文；人际议题显要性主要体现为Twitter个人用户与其他用户的互动内容，如转发内容和评论内容。我们在测量Twitter公众议程时，先计算每个个人用户发表的原创推文中涉及各议题的频数，然后将每个个人用户在每个议题上的频数进行累加，对累加数据进行等级排列，得到Twitter公众议程。我们在测量Twitter个人议程时，先分别计算每个个人用户原创推文涉及各议题的频数和该用户转发、评论内容中涉及各议题的频数，将两种频数相加，得到每个个人用户的Twitter个人议程。

## （三）文本挖掘与主题建模

社交媒体上的个人议程设置研究面临方法上的问题，传统的“媒体内容分析+受众抽样调查”不再适用。内容分析方法主观性较强，且难以对大规模文本进行有效的处理和分析。令人可喜的是，大数据文本挖掘技术的发展使得对大规模文本的自动化分析成为可能。文本挖掘是从大量文本的集合或语料库中抽取事先未知的、可理解的、有潜在使用价值的模式和知识（刘春艳，2011）。

其中，主题建模（topic modeling）是一种适用于议程设置研究的文本挖掘技术。主题模型技术的基本思想是“文本是由多个主题混合而成的，而主题是特征词上的一种概率分布，即每篇文本是主题的混合分布，而每一个主题是一组特征词的混合分布”

（王燕鹏，2017）。研究者已将主题建模技术用于识别议题，比如Bae等人利用主题建模技术识别Twitter议题，并对每日议题进行可视化，实验数据为2013年韩国Twitter中的近1.5亿条推文（Bae et al., 2014）。Grimmer采集了美国众议院议员发布的超过64000篇的新闻稿，应用主题建模技术对新闻稿的议题进行识别；通过与专业的人类编码员的对比，研究发现主题建模所得的结果要优于人类编码结果（Grimmer, 2009）。据Hong等

人的研究, Twitter上主题建模的准确率最高可达95.83% (Hong et al., 2010)。

本研究将应用LDA (latent dirichlet allocation) 主题建模技术对Twitter文本所含的议题进行识别。LDA主题建模和议题识别的过程中, 我们利用MongoDB数据库存储和查询twitter数据, 利用GraphLab2.1 (由python语言调用) 对7117.2万条推文进行LDA主题建模。LDA主题模型的基本参数为:  $\alpha=0.1$ ;  $\beta=0.1$ ; 主题数=100; 求解方法为吉布斯采样 (Gibbs sampling); 迭代次数=1000。议题是一种社会公共问题, 主题建模所得的100个主题并不都属于议题。后续分析中我们将排除如下内容: (1) 不属于议题的私人问题 (如“家庭关系”); (2) 无实际意义的主题; (3) 非媒体账号和个人账号 (如商业机构、公益机构) 的内容。最后我们识别了36个议题, 涉及153.1万条媒体推文和2092万条个人用户推文。

36个议题的命名同时基于LDA主题建模结果和已有文献。其中, “种族问题” “民族问题” “治安问题” “交通问题” “难民问题” “税收问题” “收支问题” “司法问题” “商业技术” “医学研究” “太空探索” 和 “总统选举” 的命名主要基于盖洛普的MIP调查 (Gallup, 2017); “恐怖主义” “政府活动” “政治争论” “宗教问题” “健康问题” “全球贸易” “就业问题” “能源问题” “商界领袖” “社交媒体” “自然灾害” 和 “城市问题” 的命名主要基于Godeaux (2014) 和Guo等人 (2015) 的大数据议程设置研究; “性别问题” “教育问题” “犯罪问题” “动物问题” “体育赛事” “气象状况” 和 “气候变迁” 的命名主要基于Rogstad (2016) 和Wenner (2014) 的Twitter议程设置研究; “移动设备” “流行音乐” “影视明星” “演艺活动” 和 “游戏问题” 的命名主要基于Bantimaroudis等人 (2014) 的文化议程设置研究。

## 四、研究结果

### (一) Twitter上的个人议程、公众议程与媒体议程

Twitter个人议程是每个Twitter个人用户的议程, 本研究共含有72344项个人议程, 涉及2092万条推文。如表3, 我们以个人议程A (来自ID为“100949009”的用户) 和个人议程B (来自ID为“10904”的用户) 为例, 个人议程A中排名第一的议题是“税收问题”, 占比将近一半 (49.1%), 排名第二的议题是“流行音乐”, 占比10.1%, 远低于排名第一的议题比例。个人议程B中有议题占比超过10%: “犯罪问题” 和 “演艺活动”, 分别占比22.2%和17.2%。可以发现, 个人议程A和个人议程B都有占比特别突出的议题。实际上, 72344项个人议程中, 52.2%的个人议程中排名前二的议题占比之和超过30%; 77.0%的个人议程中排名前三的议题占比之和超过30%; 这意味着多数Twitter个人用户都有其特别关注的议题, 特别关注议题分散在36个议题中, 体现出个人差异。

表3 Twitter上的个人议程、公众议程与媒体议程（频数/百分比）

议题	个人议程 A	个人议程 B	公众议程	媒体议程
体育赛事	29(0.7)	8(0.6)	74.7 万(7.0)	13.3 万(8.7)
社交媒体	413(10.1)	3(0.2)	49.6 万(4.6)	10.4 万(6.8)
交通问题	243(5.9)	9(0.7)	24.7 万(2.3)	8.9 万(5.8)
总统选举	12(0.3)	11(0.9)	27.8 万(2.6)	7.4 万(4.9)
流行音乐	45(1.1)	30(2.4)	70.7 万(6.6)	7.3 万(4.7)
气候变化	69(1.7)	18(1.5)	53.2 万(5.0)	6.7 万(4.4)
太空探索	3(0.1)	25(2.0)	21.2 万(2.0)	6.3 万(4.1)
游戏问题	2015(49.1)	14(1.1)	22.5 万(2.1)	6.2 万(4.1)
气象状况	146(3.6)	5(0.4)	56.8 万(5.3)	5.1 万(3.3)
种族问题	14(0.3)	40(3.2)	19.1 万(1.8)	5.1 万(3.3)
演艺活动	26(0.6)	16(1.3)	32.8 万(3.1)	4.9 万(3.2)
健康问题	14(0.3)	40(3.2)	22.7 万(2.1)	4.9 万(3.2)
动物问题	51(1.2)	7(0.6)	21.4 万(2.0)	4.8 万(3.1)
难民问题	11(0.3)	9(0.7)	17.0 万(1.6)	4.8 万(3.1)
治安问题	21(0.5)	17(1.4)	37.0 万(3.5)	4.6 万(3.0)
宗教问题	102(2.5)	8(0.6)	21.5 万(2.0)	4.4 万(2.9)
城市问题	17(0.4)	11(0.9)	48.5 万(4.5)	4.2 万(2.7)
教育问题	53(1.3)	28(2.3)	37.3 万(3.5)	4.0 万(2.6)
民族问题	142(3.5)	46(3.7)	23.1 万(2.2)	3.4 万(2.2)
政治争论	39(1.0)	31(2.5)	20.3 万(1.9)	3.2 万(2.1)
税收问题	96(2.3)	40(3.2)	19.3 万(1.8)	3.1 万(2.1)
商界领袖	14(0.3)	90(7.3)	31.2 万(2.9)	2.6 万(1.7)
收支问题	18(0.4)	20(1.6)	30.4 万(2.8)	2.6 万(1.7)
恐怖主义	53(1.3)	212(17.2)	31.6 万(3.0)	2.5 万(1.7)
就业问题	92(2.2)	9(0.7)	19.9 万(1.9)	2.3 万(1.5)
医学研究	43(1.1)	17(1.4)	20.3 万(1.9)	2.2 万(1.5)
自然灾害	9(0.2)	42(3.4)	15.0 万(1.4)	2.2 万(1.4)
能源问题	35(0.9)	28(2.3)	19.8 万(1.9)	2.1 万(1.4)
性别问题	44(1.1)	21(1.7)	26.6 万(2.5)	2.0 万(1.3)
影视明星	13(0.3)	14(1.1)	27.9 万(2.6)	1.9 万(1.3)
政府活动	13(0.3)	23(1.9)	26.0 万(2.4)	1.9 万(1.2)
犯罪问题	77(1.9)	15(1.2)	14.3 万(1.3)	1.7 万(1.1)
全球贸易	35(0.9)	273(22.2)	17.4 万(1.6)	1.7 万(1.1)
商业技术	26(0.6)	30(2.4)	20.6 万(1.9)	1.6 万(1.1)
司法问题	41(1.0)	8(0.6)	15.8 万(1.5)	1.4 万(0.9)
移动设备	28(0.7)	13(1.1)	36.0 万(3.4)	1.3 万(0.9)
总计	4102(100.0)	1231(100.0)	1074.0 万(100.0)	153.1 万(100.0)

Twitter公众议程涉及1074万条推文，排名第一的议题是“体育赛事”，占比7.0%，远低于个人议程A和个人议程B中排名第一的议题比例；公众议程中排名最末的议题是“移动设备”，占比1.3%，明显高于个人议程A和个人议程B中排名最末的议题比例。Twitter媒体议程涉及153.1万条推文，排名第一的议题与公众议程一致，也是“体育赛事”，占比8.7%。“体育赛事”“社交媒体”和“流行音乐”同时出现在公众议程和媒体议程前五的议题中；“移动设备”“商业技术”和“犯罪问题”同时出现在公众议程和媒体议程后五的议题中。由此看来，Twitter上的公众议程与媒体议程具有较强的一致性，实际上，二者的皮尔逊相关系数达0.62 ( $Sig. < 0.001$ )；以公众议程设置研究的视角来看，Twitter媒体很有可能为Twitter个人用户设置议程。

## （二）Twitter上多数个人议程与公众议程没有一致性；人内议题显要性与人际议题显要性高度相关

社交媒体平台上个人议程设置研究立论的基础之一是个人的差异性，本研究主要通过Twitter上公众议程与个人议程的差异性进行验证，二者皮尔逊相关系数分布如图3。Twitter上公众议程与72344项个人议程的相关系数的最大值为0.87 ( $Sig. < 0.001$ )，说明部分个人议程确实与公众议程存在高度相关性；相关系数最小值为-0.37 ( $Sig. < 0.05$ )，即有的公众议程中的显要议题在某些个人议程中的显要性极低。72344项相关系数的均值为0.27，中位数为0.28，在 $\alpha = 0.05$ 下均未达显著（显著正相关的临界值为0.33）。实际上，72344项个人议程中，仅有29120项与Twitter公众议程的相关系数大于0.33，即59.7%的个人议程与公众议程不存在显著的正向相关。这表明，大多数情况下，Twitter个人议程与公众议程没有一致性，验证了假设H1，也表明Twitter这一社交媒体平台上确实存在明显的个人差异，进行个人层面的议程设置研究具有合理性。

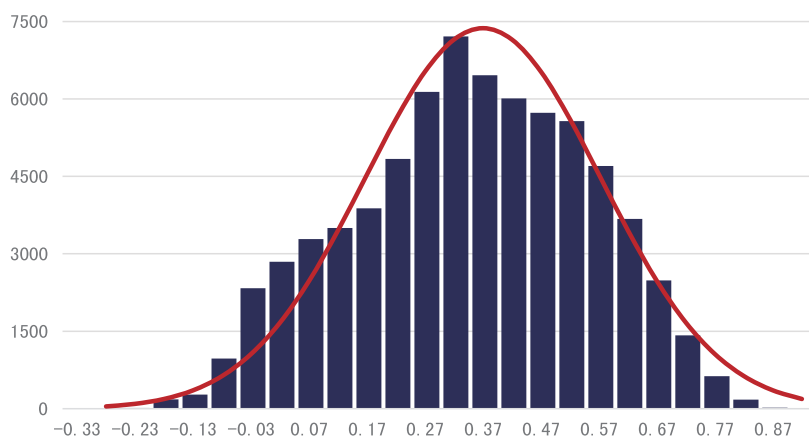


图3 Twitter上公众议程与个人议程相关系数的分布

个人议程设置研究的另一个立论基础是人内议题显著性和人际议题显著性的一致性。本研究中，Twitter人内议题显著性共涉及1074万条原创推文，Twitter人际议题显著性共涉及1018万条转发推文和评论推文，二者比例接近1:1，人际议题显著性的存在无法忽视。更为重要的是，人内议题显著性和人际议题显著性的皮尔逊相关系数高达0.81 ( $Sig. < 0.001$ )，验证了假设H2，二者确实存在高度相关。由于存在高度的一致性，人内议题显著性和人际议题显著性可以相加；加上已证实的个人差异，至此，“个人议程”概念不仅从理论上得以推导，而且从现实中也得到验证。

### (三) Twitter上媒体议程与30.3%的个人议程显著正相关

“个人议程”概念的理论推导和现实验证使得个人议程设置研究成为可能，个人议程设置研究关注一个媒体议程对多项个人议程的影响。本研究分析的是Twitter媒体议程与72344项个人议程之间的一致性程度，二者皮尔逊相关系数分布如图4。Twitter媒体议程与72344项Twitter个人议程相关系数的最大值为0.84 ( $Sig. < 0.001$ )，在不考虑其他因素的情况下，这说明Twitter媒体账号仍然具备为部分Twitter普通个人用户设置议程的能力。Twitter媒体议程与个人议程相关系数的最小值为-0.50 ( $Sig. < 0.01$ )，此时，Twitter媒体不仅不能为该用户设置议程，甚至该用户将媒体很少报道的议题作为其个人的重要议题。事实上，72344位Twitter个人用户中，有13223位用户(18.3%)的个人议程与Twitter媒体议程呈现负相关；这些用户的存在不仅仅表明这些个人议程与Twitter媒体议程的不一致，更可能表明这些个人用户的信息源可能并不主要是Twitter媒体，甚至不主要是一般意义上的专业媒体；这些用户也就成为Twitter媒体难以影响的用户。

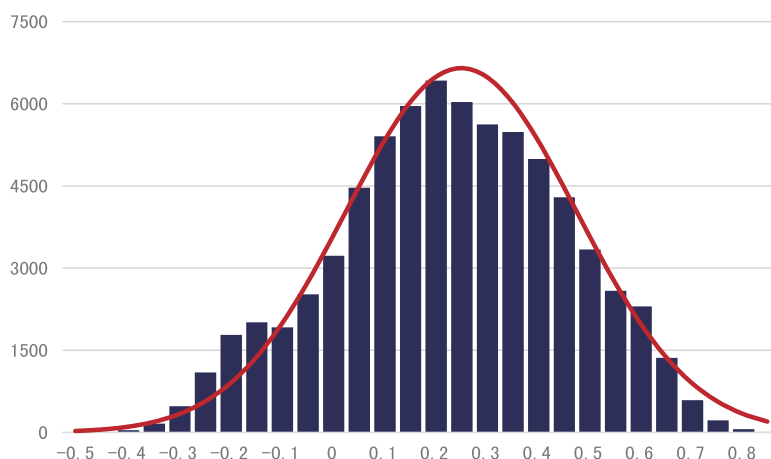


图4 Twitter上媒体议程与个人议程相关系数的分布



Twitter上媒体议程与72344项个人议程相关系数的均值为0.20,中位数为0.21,都未达显著( $Sig. > 0.05$ ),这表明,一般来说, Twitter媒体议程并不显著影响Twitter个人议程。仍然以0.33为显著正相关的临界值,有21905位Twitter用户的个人议程与媒体议程的相关系数大于0.33,由此我们回答Rq1: Twitter媒体议程与30.3%的Twitter个人议程显著正相关。换言之,大部分的Twitter个人议程与Twitter媒体议程没有一致性,也就意味着Twitter媒体不能为大部分Twitter个人用户设置议程。

## 五、结论与讨论

本研究从个人差异、阿卡普尔科模型和新媒介的发展论证社交媒体时代应该进行“真正个人层面的议程设置研究”(Guo, 2017: 10),从帕克传统、社交媒体平台的互动性和受众的主动性论证个人议题显要性的两个维度(人内议题显要性和人际议题显要性),以此为基础界定了“个人议程”概念。并且,本研究基于Twitter平台2015年全年的抽样数据(含7177.2万条推文)和LDA主题建模技术进行了实证研究。研究发现,

(1)“体育赛事”是Twitter媒体报道最多的议题,多数Twitter个人用户都有特别关注的议题(占比远高于平均比例);(2)Twitter上约有60%的个人议程与公众议程不存在显著正相关,人内议题显要性与人际议题显要性高度相关( $r=0.81$ ),证实了“个人议程”概念的合理性;(3)Twitter媒体议程与30.3%的Twitter个人议程存在显著正相关。

Twitter上媒体议程与大多数个人议程没有一致性,是否意味着一种“弱效果”或“有限效果”?从30.3%的绝对数值上看,这确实不是一种“强效果”。不过,我们需要考虑到以下情境:其一,本研究的数据是Twitter全平台2015年全年的抽样数据,涵盖了多达36种议题;在某个特定时间段的某个特定领域或具体案例中,媒体议程与个人议程一致性的比例应有明显提高。其二,30.3%作为一种“弱效果”是在没有参照系的情况下的一种常识化理解;相对于拉扎斯菲尔德等人在“人民的选择”研究中发现的8%的态度转变(拉扎斯菲尔德等,2012: 57),30.3%明显是一个大得多的比例;我们难以明确判断媒体议程与30.3%的个人议程存在一致性到底是一种“强效果”还是“弱效果”。其三, Twitter媒体议程与30.3%的twitter个人议程的一致性同时形成了Twitter媒体议程与Twitter公众议程的显著相关( $r=0.62$ ,  $Sig. < 0.001$ );二者是一体两面的关系,如果承认 $r=0.62$ 是一种“强效果”,那么也就无法否认30.3%的一致性也是一种“强效果”。

“强效果”与“弱效果”之争不应该成为个人议程设置研究的重点。个人差异的存在必然使得媒体议程与一部分个人议程呈现较高的一致性,与另一个部分个人议程不相关,甚至负相关。个人议程设置研究更应该关注的是:什么情况下媒体议程与个人议程显著正相关?或者说什么因素影响了媒体对受众个人的议程设置效果?我们需要从“个人议程”的视角出发,进行议程设置的偶发条件研究(麦库姆斯,2018: 110)。相对

于考察媒介体制、议题属性的影响，个人议程设置视角出发的偶发条件研究更为关注受众个人的社会心理因素对议程设置效果的影响。公众议程设置在个人层面的偶发条件研究是将受众划分成几个子群（如“高导向需求者”“低导向需求者”），所谓的个人特征其实是子群特征，并不能在个人层面考察受众的社会心理特征对议程设置效果的影响。个人议程设置研究本就是以个人差异为基础立论的，能够直接分析个人社会心理因素的影响，这也是后续个人议程设置研究的重点所在。

个人议程设置研究并没有否认公众议程设置研究的价值。由于个人差异的持续增大，公众议程只能代表越来越少的个人议程，越来越无法代表受众在议题显要性上的共识。媒体可能仍然与公众议程显著正相关，却只能影响少量的受众个人。然而，公众议程及公众议程设置在公共事务决策中仍然具有重要作用。虽然公众对于何为重要议题可能没有共识，但政府及相关利益团体在推动解决社会公共问题时仍然需要一个待解决问题的排序，以便于决策。待解决问题的排序即是公众议程的体现，并受到媒体议程的影响。并且，个人议程设置效果与公众议程设置效果存在关联。本研究中，Twitter媒体议程可能能够影响30%左右的Twitter个人议程，并与Twitter公众议程的皮尔逊相关系数为0.62。实际上，0.33为显著正相关的临界值，据估算，此时只需要Twitter媒体能够影响10%左右的Twitter个人用户。这意味着，媒体议程只需要影响极少部分的个人议程，就能够达到对公众议程的显著影响，而政府及相关利益团体据此决策。并且，随着个人之间差异的增大，为达到媒体议程与公众议程的显著正相关，媒体需要影响的受众个人比例将越来越低；这意味着媒体对公众议程的影响将变得越来越容易。此时，媒体推动形成的“共识”越来越少，但仍然能够通过影响公众议程进而影响决策。

个人议程设置效果的偶发条件研究之外，未来个人议程设置研究更应该在多种社交媒体平台和互联网平台上进行，本研究局限于Twitter平台，研究结论的推广性有所欠缺。从时间序列数据中考察媒体议程和个人议程的关系也是未来研究的重点方向。另外，多种媒体议程能够相加为总的媒体议程是基于“不同媒体报道具有相似性”（McCombs，1981）的假定，后续研究需要验证这一假定是否成立；如果不成立，则需要从媒介使用的角度考察特定媒体对特定个人的影响。

---

#### 参考文献：

- [1] McCombs, M. E., and Shaw, D. L. "The Agenda-setting Function of Mass Media." *Public Opinion Quarterly* 36.2 (1972):176-187.
- [2] Guo, L. "Agenda - Setting: Individual - Level Effects Versus Aggregate - Level Effect." *The International Encyclopedia of Media Effects*. Ed. Patrick Rössler, Wiley-Blackwell, US:John Wiley & Sons, 2017:1-13.
- [3] [美]迪林, 罗杰斯著, 倪建平译. 传播概念 • Agenda-Setting[M]. 上海: 复旦大学出版社, 2009.
- [4] McCombs, M. *Setting the Agenda: Mass Media and Public Opinion*. New Jersey: John Wiley & Sons, 2014.

- [5] Shaw, D. L., and McCombs, M. E. *The Emergence of American Political Issues : the Agenda-setting Function of the Press*. West Pub.co, 1977.
- [6] McCombs, M., Graber, D. A., and Weaver, D. H. *Media Agenda-setting in the Presidential Election*. NY: Praeger Scientific, 1981.
- [7] Kosicki, G. M. "Problems and Opportunities in Agenda-Setting Research." *Journal of Communication* 43.2(1993):100-127.
- [8] Vu, H. T., Guo, L., and McCombs, M. E. "Exploring 'the World Outside and the Pictures in our Heads' A Network Agenda-setting Study." *Journalism & Mass Communication Quarterly* 91.4 (2014):669-686.
- [9] McCombs, M. E. "Tracing the Time Lag in Agenda-setting." *Journalism Quarterly* 58.1 (1981):51-55.
- [10] Becker, L. B. "Reflecting on Metaphors." *Annals of the International Communication Association* 14.1 (1991):341-346.
- [11] [美] 克里斯·安德森著, 乔江涛译. 长尾理论[M]. 北京: 中信出版社, 2012.
- [12] Hausman, J. A., and Newey, W. K. "Individual Heterogeneity and Average Welfare." *Econometrica* 84.3 (2016):1225-1248.
- [13] 靖鸣, 臧诚. 媒介融合时代信息流动模式、分众化传播及媒体对社会凝聚力的影响[J]. 新闻与传播研究, 2011 (5), 第34-42页, 第110页.
- [14] 廖圣清, 黄文森, 易红发, 申琦. 媒介的碎片化使用: 媒介使用概念与测量的再思考[J]. 新闻大学, 2015 (6), 第61-73页.
- [15] McCombs, M. E., Shaw, D. L., and Weaver, D. H. "New Directions in Agenda-setting Theory and Research." *Mass Communication and Society* 17.6 (2014):781-802.
- [16] Camaj, L., and Weaver, D. H. "Need for Orientation and Attribute Agenda-setting during a US Election Campaign." *International Journal of Communication* 7 (2013):22.
- [17] Zucker, H. "The Variable Nature of News Media Influence." *Annals of the International Communication Association* 2.1 (1978):225-240.
- [18] Roessler, P. "The Individual Agenda-designing Process: How Interpersonal Communication, Egocentric Networks, and Mass Media Shape the Perception of Political Issues by Individuals." *Communication Research* 26.6 (1999):666-700.
- [19] Arguelhes, D. W., and Hartmann, I. A. "Timing Control without Docket Control: How Individual Justices Shape the Brazilian Supreme Court's Agenda." *Journal of Law and Courts* 5.1 (2017):105-140.
- [20] Yeager, D. S., Larson, S. B., Krosnick, J. A., et al. "Measuring Americans' Issue Priorities: A New Version of the Most Important Problem Question Reveals more Concern about Global Warming and the Environment." *Public Opinion Quarterly* 75.1 (2011):125-138.
- [21] McLeod, J. M., Becker, L. B., and Byrnes, J. E. "Another Look at the Agenda-setting Function of the Press." *Communication Research* 1.2 (1974) :131-166.
- [22] Lippmann, W. *Public Opinion*. NY: Hartcourt Brace.1922.
- [23] Park, R. E. *The City*. Chicago: Chicago University Press,1925.
- [24] Park, R. E. "News as a Form of Knowledge: A Chapter in the Sociology of Knowledge." *American journal of Sociology* 45.5 (1940):669-686.
- [25] Gadziala, S. M., and Becker, L. B. "A New Look at Agenda-Setting in the 1976 Election Debates." *Journalism & Mass Communication Quarterly* 60.1 (1983):122-126.
- [26] McCombs, M. "Agenda Setting Function of Mass Media." *Public Relations Review* 3.4 (1977):89-95.
- [27] Qu, Q., Chen, C., Jensen, C. S., et al. "Space-Time Aware Behavioral Topic Modeling for Microblog Posts." *IEEE Data Eng. Bull* 38.2 (2015):58-67.
- [28] Nalwoga, L. *Examining Agenda Setting Effects of Twitter Users during the 2016 Uganda Presidential Election*. Uppsala University, 2017.
- [29] 刘春艳. 基于信息可视化的文本挖掘研究领域前沿与演化分析[J]. 图书情报工作, 2011 (S2), 第270-272页, 第189页.
- [30] 王燕鹏. 基于文献计量的主题模型研究进展分析[J]. 科学观察, 2017 (5), 第9-20页.
- [31] Bae, J. H., Han, N. G., and Song, M. "Twitter Issue Tracking System by Topic Modeling Techniques." *Journal of Intelligence and Information Systems* 20.2 (2014):109-122.
- [32] Grimmer, J. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 18.1 (2009):1-35.
- [33] Hong, L., and Davison, B. D. Empirical Study of Topic Modeling in Twitter in: *Proceedings of the First Workshop on Social Media Analytics*. ACM, 2010: 80-88.
- [34] Gallup. *Most Important Problem*. Retrieved from news.gallup.com/poll/1675/most-important-problem.aspx. /2017-10-21.
- [35] Godeaux, L. "The Dynamics of Public Attention: Agenda-Setting Theory Meets Big Data." *Journal of Communication* 64.2 (2014):193-214.
- [36] Guo, L., and Vargo, C. "The Power of Message Networks: A Big-Data Analysis of the Network Agenda Setting Model and Issue Ownership." *Mass Communication & Society* 18.5 (2015):557-576.
- [37] Rogstad, I. "Is Twitter just Rehashing? Intermedia Agenda Setting between Twitter and Mainstream Media." *Journal of Information Technology & Politics* 13.2 (2016):1-17.
- [38] Wenner, L. A. "Much ado (or not) about Twitter? Assessing an Emergent Communication and Sport Research Agenda." *Communication & Sport* 2.2 (2014):103-106.
- [39] Bantimaroudis, P., and Zyglidopoulos, S. C. "Cultural Agenda Setting: Salient Attributes in the Cultural Domain." *Corporate Reputation Review* 17.3 (2014):183-194.
- [40] [美] 拉扎斯菲尔德, 贝雷尔森, 高德特著, 唐茜译. 人民的选择[M]. 北京: 中国人民大学出版社, 2012.
- [41] [美] 马克斯韦尔·麦库姆斯著, 郭镇之, 徐培喜译. 议程设置: 大众媒介与舆论[M]. 北京: 北京大学出版社, 2018.

are as follows: (1) social media use has a significant positive predictive effect on social responsibility identification after controlling gender, age, household registration, source (urban and rural residents), income, and time spent in Shanghai; (2) social integration can play an intermediary role in the relationship between social media use and social responsibility identification; (3) both the direct predictive effect of social media use on social responsibility identification and the mediating effect of social integration in their relationship are regulated by social trust.

**[Keywords]** social media use; social responsibility identification; social integration; social trust

### 63 Impacts of Media Agenda on Individual Agenda: Mining Text-data on Twitter

---

· *YI Hong-fa*

**[Abstract]** A half-century Agenda-setting research faces new challenges. With the development of the Internet and social media, the diversity and interaction between audiences has increased. The public agenda affected by the media can no longer become a social consensus, and the agenda-setting research at the individual level has become an important orientation. This study clearly defines the concept of "individual agenda" and conducts an empirical study of the relationship between media agenda and individual agenda based on 71.772 million tweets in 2015 on Twitter platform and topic modeling technique. The study found that: (1) Twitter media reported the most about "sports events", and most Twitter users had special concerns. (2) Most individual agendas were inconsistent with the public agenda; the intrapersonal salience and the interpersonal salience were highly related. (3) Twitter media agenda was significantly related to 30.3% of Twitter individual agendas. The study also discusses the strength of the individual agenda setting effect, the occasional conditions at the individual level, and the relationship between individual agenda setting and public agenda setting.

**[Keywords]** individual agenda; public agenda; agenda-setting; personal differences; topic modeling