

## 基于政府网站数据的网络问政分析研究

小组成员: 许家辉 杜思竹 冯国强 吴思余



- 1 选题解读
- 2 参考文献
- 3 思路分析



第一部分 选题解读

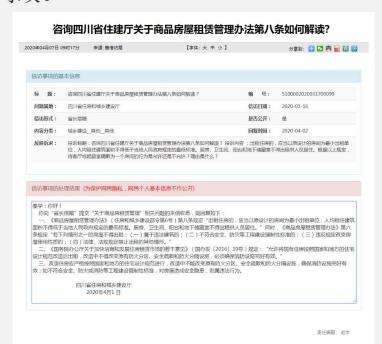






网络问政——随着网络的日益普及,互联网在中国民众的政治、经济和社会生活中扮演着日益重要的角色,中国公民以网民的身份通过互联网行使知情权、参与权、表达权和监督权,这就是网络问政。



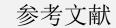




第二部分 参考文献

PART 02







论文题目	数据来源及数据获取	数据结构	
《公众环境诉求影响因素及其作用机制》	数据来源:5个网络问政平台时间段:2010年2017年获取方法:利用软件编程爬取了研究时间段内北京市公众的所有留言数据,共获得26541条数据预处理:机器学习文本分类技术	所有留言数据的文本(从中提取时间、地区名、问 题关键词等)	
《公众网络反腐败参与研究》	数据来源:全国性网络问政平台——人民网的"地方领导留言板"时间段: 2008 至 2014 年上半年期间获取方法:未提	文本	
《条块关系视野下的网络诉求与政府回应模式研究》	以中部某地市设立的网络问政平台为分析案例,并针对该平台进行全数据抓取本研究获取了2010年该平台运行至2017年11月底该平台的所有留言及政府回应信息。	文本	
《网络空间公民利益诉求的对象选择与行为特征》	数据来源:全国网络问政平台———人民网" 地方领导留言板" 时间段: 2008年1月到2015年9月 获取方法:网络爬虫技术	全部发帖内容、发帖时间、发帖诉求对象等信息	







论文题目	研究方法	理论模型	实证模型
《公众环境诉求影响因素及其作用机制》	先从 26541 条文本中随机抽取1300 条样本,逐条 人工分析内容,并按照分类标准要求归纳汇总。参 考"环境污染投诉网"的分类标准,本文将环境污染 类型分为 6 类,分别为大气污染、电磁辐射污、固 体废物污染、生态及资源环境破坏、水污染和噪音 污染。然后将样本中的 700 条作为训练集,剩余		公众环境诉求量,构建如下面板模型: $\ln EA_n = \alpha_0 + \alpha_1 \times \ln GDPC_n + \alpha_2 \times \ln EDU_n + \alpha_3 \times \ln POP_n + \alpha_4 \times AGE_n + \alpha_5 \times \qquad (1)$ $SEX_n + \sum_{n=1}^{N} \alpha_{6n} \times \ln X_{n-1} + \varepsilon_n$ 其中, $i$ 代表北京市業区, $i$ 代表时间。 $EA$ 代表网络问政平台上的公众环境诉求量, $GDPC$ 代表人均实际 $GDP$ 水平, $EDU$ 代表地区教育水平,
《公众网络反腐败参与研究》	首先采取自动文本分析方法、采用有监督学习法。 首先从本文采集的 21 万网民发帖中筛选与反腐败相 关的发帖共约 5.4 万; 其次从5.4 万原始网民发帖中 随机抽取 1000 份留言由研究助理进行人工分类; 再次利用朴素贝叶斯算法学习人工分类规模,并对 其余所有网民留言进行计算机自动分类; 最后研究 者再核查机器分析结果的效度,并展开后续分析。		主要采用统计分析中的回归
《条块关系视野下的网络诉求与政府回应模式研究》	文本分析、机器学习、 <mark>情感分析</mark>	Logistic 回归	模型(1)仅仅放入公民诉求的时间变量;模型(2)是全模型,放入诉求时间、诉求对象、诉求议题分布等;模型(3)分析公民诉求对象为政府职能部门时对于公民诉求的回应机制;模型(4)则分析诉求对为各县市区时对于公民诉求的互动机制;模型(5)分析公共服务行业部门在接受公民诉求后的回应机制。
《网络空间公民利益诉求的对象选择与行为特征》	自动文本分析	基于 Logit 回归	

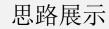


第三部分

思路分析

PART 03







- 1.研究中国网络问政平台的建设情况(评分机制的建立)
- 2.网络问政的冲突——民众关注与政府关注
- 3.网络问政对公共服务的推进作用
- 4.网络问政的创新成效





网络问政的冲突——民众关注与政府关注

数据来源:来自政府网络问政平台

数据内容:主要是政民问答数据,包括公众提问类型(咨询问答、投诉举报)、标题、时间、内容,以及政府回复的具体部门、回信时间、回信内容等相关文本信息

获取方式: python网络爬虫技术





## 主要模型如下

- 1. 领域分类模型 (利用文本分析,采用特征词比较的方法)
- 2.政府关注识别模型(用于识别政府的关注高低)
- 3.民众关注识别模型(用于识别民众关注的高低)
- 4.政府关注-民众关注冲突模型(用于分类不同领域冲突类型)

最终将多个模型结合,综合得到不同领域不同时间的冲突模型,从中研究出一些有趣的结论

结论一般分三个角度: 民众关注角度, 政府关注角度, 以及民众与政府冲突角度



## 谢谢您的观看!