

청와대 국민청원에서는 무엇이 일어나는가? :

자연어 처리를 활용한 청와대 국민청원 분석*

송준모** · 박영득***

논문 요약

청와대 국민청원은 단순한 분노의 발산 통로인가, 직접 민주주의의 장인가. 본 연구는 청와대 국민청원에 대한 상반된 가치판단을 넘어서, 감정이 정치적 참여와 판단에 미치는 영향을 경험적으로 분석하고자 한다. 이를 위해 청와대 국민청원 페이지 개설 시점으로부터 연구 시작 시점까지 1년 이상 축적된 게시물을 전수 수집하였다(중복 제거 후 총 88,113 건). 그리고 자연어 처리(구조적 토픽 모형과 워드 투 벡터)를 통해 국민청원 페이지에 나타난 의제와 감정들을 파악한 후, 음이향 회귀분석을 통해 글의 구성요소와 동의수의 관계를 파악하였다. 분석결과는 다음과 같다. 첫째, 다수의 글이 작성된 의제가 곧 다수의 동의를 받지는 않는다. 둘째, 슬픔보다는 분노가 동의에 더 큰 효과를 가진다. 셋째, 분노는 심각한 갈등적 의제보다 일상적인 의제와 결합할 때 더 큰 호응을 얻는다. 본 연구는 이러한 분석 결과를 통해 청와대 국민청원이 기존의 양극화된 접근과 달리 다양한 가능성을 배제한 공간이라고 주장한다. 그리고 경험적 접근을 통해 감정과 이성 사이의 허구적 대립을 해소하고, 다양한 정치적 참여 통로가 민주주의에 기여할 수 있는 조건이 탐색되길 기대한다.

주제어: 청와대 국민청원, 전자청원, 시민참여, 의제설정, 감정, 자연어처리

* 이 연구의 초기 버전은 2018년 한국정치학회 연례학술회의에서 발표된 바 있습니다. 심사과정에서 발전적인 논평을 보내주신 익명의 심사자 선생님들께 깊이 감사드리며, 이 논문의 최종본에 반영하지 못한 제안들은 향후 더 발전된 연구를 통해 반영될 수 있도록 하였습니다. 이 논문은 2016년도 정부재원(교육부 인문사회연구역량강화사업비)으로 한국연구재단의 지원을 받아 연구되었음(NRF-2016S1A3A2925033).

** 연세대학교 사회학과 박사과정

*** 포항공과대학교 인문사회학부 대우조교수

<https://doi.org/10.18854/kpsr.2019.53.5.003>

I. 서론

본 연구는 청와대 국민청원에 게시된 청원문서를 웹크롤링(web crawling)을 통해 수집한 후 자연어 처리(NLP: Natural Language Processing)를 통해 분석함으로써 청와대 국민청원이 어떤 주제들로 구성되어 있는지를 기술(describe)하는 것을 일차적 목표로 한다. 그리고 이러한 기술적 분석에서 더 나아가 청와대 국민청원에 대한 상반된 입장, 즉 의미 있는 사회적 담론이 형성되는데 긍정적 영향을 미치는 공간인지(한겨레경제사회연구원 2018; 한국여성정책연구원 2019), 아니면 제도를 우회하고 시민들의 분노만을 조직화하여 공론장을 피폐하게 만드는지(박상훈 2018)에 대한 경험적 판단근거를 제시하는 것을 최종적 목표로 한다.

민주정(democracy)을 다른 정체(polity)와 구분짓는 핵심적 속성은 아마도 “정부의 시민의 선호에 대한 지속적인 반응성(Dahl 1971, 1)”일 것이다. 일반적으로 민주정은 자유롭고 공정한 경쟁이 보장되는 주기적인 선거를 통해 시민의 대표를 선출하고 그들로 하여금 시민의 요구와 공동체의 이익에 복무하도록 설계된 정체이다. 정기적으로 치러지는 선거는 대표자들에게 보상과 처벌을 가함으로써, 대표자들에게 시민들의 이해관계를 대변할 유인을 제공한다. 주권자인 시민들이 구성한 의회와 정부가 시민들의 권리를 실질적으로 보장하고 공동체를 더 나은 곳으로 만들기 위해서는 시민에 대한 정부의 수직적 책임성(vertical accountability)과 국가 기구 사이의 수평적 책임성(horizontal accountability)이 모두 확보되어야 한다. 하지만 주기적인 선거만으로는 수직적 책임성조차 효과적으로 확보하기가 어렵다(O'Donnell 1998, 112-113).

선거기간이 아니더라도 시민들의 요구는 꾸준히 의회와 정부에 투입될 수 있어야 하며 이러한 요구에 대하여 의회와 정부는 시민들에게 책임을 가지고 응답해야 한다. 일반적으로 민주주의 국가에서는 시민들의 의사를 의회와 정부에 투입하는 경로로서, 언론의 자유나 집회결사의 자유와 같은 자유주의적 기본권이 보장하는 조직화된 정치참여가 중요한 역할을 하고 있다. 그러나 사회의 규모와 복잡성이 급속도로 증대하고, 사회적 변화의 속도가 빠른 오늘날에는 의회와 정부가 이질적인 시민들의 다양한 요구를 빠르게 집약하여 파악할 수 있는 장치 또한 필요하다. 이에 영국, 미국 등 서구권의 주요 민주주의 국가들은 시민들이 개인 수준에서 빠르고 편리하게 국가기구에 의사를 전달하기 위한 수단으로서 전자 청원을 도입하였다.

한국 역시 2017년 8월, 청와대의 주도로 ‘국민이 물으면 정부가 답한다’는 구호를 내걸고 청와대 공식 홈페이지에 국민청원 페이지¹를 개설했다. 청와대 국민청원 페이지는 버락 오바마 대통령 시기에 개설된 미국 백악관의 청원 페이지인 “위더피플(We The People)”²을 벤치마킹하여 일반 국민들이 직접 대통령에 직접 청원을 제기하는 온라인 공간으로 기획되었다. 국민청원을 개설한 문재인 대통령은 새 정부 출범 성과 대국민보고에서 “국민이 선거 때 한 표 행사하는 간접민주주의로는 만족하지 못하고

¹ <https://www1.president.go.kr/petitions>

² <https://petitions.whitehouse.gov/>

있다. 우리 정치가 낙후됐다고 생각한다. 국민은 정당과 정책에 참여하는 직접민주주의를 요구하고 있다”고 밝힌 바 있는데, 이러한 맥락에서 청와대 국민청원은 정당과 선출된 국민의 대표로 구성된 의회를 중심으로 운영되는 대의민주주의의 한계를 국민이 직접 참여하는 창구를 통하여 보완하려는 목적으로 개설되었다고 할 수 있다.

청와대 국민청원 페이지는 청와대가 직접 시민들과 소통하고, 시민들이 국가권력기관에 요구하는 바를 정부에 표현하고 청구할 권리가 구현되는 공간으로서 긍정적인 의미를 가지고 있다. 그러나 다른 한편으로는 대통령이 의회와 정당과 같은 대의기구를 우회하여 국민을 직접 호명하고 동원한다는 점에서 책임정부로부터 이탈했다는 비판을 받기도 한다(박상훈 2018, 144-147). 절차적 문제와 더불어, 청원 페이지에 게시되는 글들의 내용에 대한 비판도 제기된다. 예컨대 재판결과에 항의하며 특정 재판의 주심을 맡았던 법관의 해임을 요구하는 등 정부와 청와대의 권한을 벗어나는 내용에 대한 청원이 제기되거나, 강력범죄자들의 엄벌을 촉구하는 등 부정적인 감정을 촉발하는 게시물이 청원게시판에서 지나치게 많은 비중을 차지하기도 한다. 이뿐만 아니라 축구 경기에서 실수를 한 선수를 처벌하라는 식의, 정부의 정책과 어떠한 관계도 없는 부적절한 청원이 무분별하게 게시되며 많은 호응을 얻기도 한다. 이를 근거로 청와대 국민청원이 공론장을 지나치게 공격적이고 피폐하게 만든다는 비판도 제기된다(박상훈 2018, 145-147). 또한 임혁백(2011)에 따르면 인터넷을 활용하는 신유목적 정치참여는 책임이 불특정 다수에게 분산되기 때문에 정치적 책임성을 담보하기 어려워 전자 포퓰리즘을 야기할 수 있고, 정보기술을 활용할 수 있는 여건이나 능력에 따라 발생하는 정보격차(digital divide)로 인한 정치적 불평등을 심화시킬 수 있다.

청원게시판에 지나치게 무분별한 청원이 다수 제기되는 문제에 대응하여 운영주체인 청와대 측에서도 국민청원 게시판 개편안을 발표했다. 개편안은 크게 두 가지 내용으로 구분될 수 있다. 첫째, 중복청원이나 욕설, 비방 등 부적절한 내용이 담긴 청원이나 지나치게 많은 게시물이 청원게시판에 노출되는 것을 방지하기 위해 100명의 사전동의를 받은 청원문서만 청원게시판에 공개된다. 둘째, 응답기준을 충족한 청원이라고 하더라도 삼권분립에 따라 정부가 답변하지 못할 수 있다는 한계규정이 도입되었다. 또한 허위사실에 기반한 청원이나 특정 집단에 대한 차별 또는 비하의 내용을 담은 청원도 답변을 받지 못할 수 있다.

청와대 국민청원에 대한 문제의식은 크게 청원 내용에 대한 회의론과 대의기구의 우회라는 절차 차원에서의 회의론이 주를 이루고 있다. 본 논문은 청와대 국민청원의 문제를 다른 측면에서 조망하고자 한다. 본 논문은 청와대 국민청원에 부적절한 내용이나 감정적인 내용의 글이 다수 올라온다고 하더라도, 적절히 운영된다면 시민들의 요구를 국가기관이 실시간으로 청취하고 시민들의 요구에 답변함으로써 정부기구와 시민 간의 신뢰를 구축하는데 도움이 될 수 있다고 본다.³ 특히 본 연구에서는 청와대

³ 물론 이러한 관점은 청와대 국민청원 자체가 정책을 결정하지는 않는다는 것을 전제한다. 만일 청와대 국민청원에서 응답기준을 충족한 내용이 의회의 심의, 의결과 같은 적법한 정책결정과정을 거치지 않고도 시행될 수 있다면 청원 내용의 질적 수준과 관계없이 대의기구를 무력화하는 것이라고 볼 수 있겠으나 그러한 정도가 아니라면 청와대 국민청원을 단순히 시민들이 느끼는

청원게시판에 작성되는 글의 감정적 수준과 민주주의의 윤희유로서의 역할이 서로 상반되지 않는다는 입장을 취한다. 즉, 청와대 게시판에 작성되는 글이 합리적이거나 비이성적이라는 이분법을 넘어서 실제 글에서 나타는 감정과 내용들이 다른 동료 시민들의 호응에 어떠한 영향을 미치는지를 경험적으로 살펴보고자 한다.

청와대 국민청원 페이지가 개설된 지 오랜 시간이 지나지는 않았기 때문에 아직 관련 연구가 충분히 축적되지는 않은 실정이다. 국민청원제도의 놀이적 속성을 강조하며 정부와 국민의 소통을 강화하는 역할을 할 수 있다고 주장하는 연구(정성전 2018)와 국민청원 게시판의 게시물로부터 텍스트데이터를 구축하고 주요 키워드를 추출하여 키워드들 간의 의미연결망분석(Semantic Network Analysis)을 시행함으로써 국민청원에 어떠한 개혁이슈들이 제기되는지 탐색한 연구(김찬우 2019), 청와대 국민청원의 참여요인을 포폴리즘적 성향, 정치적 불신, 정치효능감으로 제시하고 각각의 국민청원 참여에 대해 미치는 영향력을 검증한 연구(이승원 · 임한샘 · 이현우 2018) 등이 존재한다. 선행연구들은 청와대 국민청원이 가지는 의미와, 시민들이 청와대 국민청원에 왜 참여하는지, 어떠한 내용들이 어떤 맥락에서 논의되는지를 살펴보고 있지만 실제로 국민청원 게시판에서 시민들이 어떤 내용의 청원에 호응을 하며, 청와대의 답변을 받은 청원과 그렇지 않은 글 사이에 차이가 존재하는지와 같은 질문은 답변 되지 않은 채로 남아있다. 또한 확률모형보다는 단순 빈도에 기반한 분석 방식을 통해서는, 청원 게시판에서 논의되는 내용을 충분히 파악하는 데에 한계가 있다.

이에 본 논문은 대규모 텍스트의 차원을 축소시켜 통계적 분석을 가능하게 만들어주고, 텍스트에 포함된 여러 정보들을 계량화 해주는 자연어처리를 활용하여 국민청원 게시판에 대한 분석을 시도한다. 이를 통하여 국민청원 게시판에서 실제 논의가 제기되는 방식을 파악하고, 어떤 속성을 가진 글이 호응을 얻고 청와대의 답변을 받을 가능성이 높은지를 파악하고자 한다. 구체적으로, 청와대 국민청원 페이지에 게시된 청원문서들을 웹크롤링을 통해 수집하여 자연어처리 기법(토픽모형, 단어 임베딩) 통해 어떠한 내용들이 어떤 방식으로 청와대 국민청원에서 논의되고 있는지 살펴볼 것이다. 그리고 어떤 글이 다수의 호응을 얻는지 통계적 분석을 시도할 것이다. 궁극적으로는 청와대 청원게시판에서 벌어지는 정치참여의 양상을 파악하여, 정치적 공론의 형성과 의제설정에 대해 이해할 수 있는 경험적 단초를 마련하고자 한다.

II. 이론적 논의

1. 시민의 정치참여와 감정

냉전이 시장경제와 민주주의 진영의 승리로 종식되자 프랜시스 후쿠야마(Francis Fukuyama)는 자유민주주의가 승리했으며 이제 더 이상 이 체제가 근본적인 위협을 맞이하는 일이 발생하지 않을 것이라는 ‘역사의 종언(the end of history)’을 선언했다(Fukuyama 1992). 그러나 최근에는 민주주의의 승리로 역사는 끝났다는 선언에 의구심이 제기될만한 징후들이 포착되고 있다. 거의 모든 사람들에게 공고화된 민주주의 국가(established democracies)로 여겨지는 선진 민주주의 국가에서도 민주주의에 대한 대중적 지지가 하락하고 있을 뿐만 아니라(Norris 2011; Dalton 2004; Klingemann 1999), 민주주의보다 권위주의적 지배(authoritarian rule)를 민주주의의 대안으로서 선호하는 시민의 수도 최근 들어 증가하고 있다(Foa and Mounk 2016; 2017).

이러한 징후에 발맞추어 주요 선진국의 제도화된 참여(투표, 정당가입, 정치인에 대한 접촉 등) 역시 점차 쇠퇴한다는 우려가 제기되었다(Dalton 2000; Lijphart 1997). 제도적 참여의 창구가 되는 정당, 의회, 정치인에 대한 신뢰가 하락하면서 대표와 시민의 긴밀한 연결을 전제하고 있는 대의민주주의는 엘리트의 지배로 인식된다(김범수·장우영 2018, 6). 민주주의의 위기가 본격적으로 나타나기 이전부터 참여의 적정 규모와 형태에 대한 비판적 접근이 이루어졌지만(Huntington and Nelson 1976), 참여 자체는 의제설정권과 관련이 있기 때문에 중요한 문제이다.

정치에서 ‘누가’, ‘어떠한’ 결정을 하느냐도 중요하지만, 만일 누군가가 결정의 대상이 되는 의제가 무엇인지를 선별할 수 있는 권력, 즉 수많은 문제들 가운데 정부가 어떠한 문제를 다룰 것인지를 결정할 수 있는 권력을 가지고 있다면 그는 정치적 논쟁 없이도 자신이 원하는 바를 달성할 수 있기 때문이다. 현대 사회에는 수많은 갈등이 상존하고 있지만 민주주의 정치체제하 하더라도 모든 갈등 사안을 다루고 있는 것은 아니다. 정치는 갈등을 편향적으로 동원하기 때문에(mobilization of bias), 특정한 사회적 갈등이 의제로 선택된다(Schattschneider 1960). 즉 정치는 어떠한 의제에 대한 갈등이 시작됨으로부터 시작되는 것이 아니라, 무엇에 대하여 갈등할지를 선택하는 시점부터 이미 시작되어 있는 것이다.

이러한 의미에서 바흐라흐와 바라츠는 그들의 권력에 대한 고전적 연구에서 권력을 가진 정치엘리트들은 정책결정과정에서도 큰 권력을 갖고 있지만, 어떠한 문제가 정치적 의제로 다루어져야하는지 선별하는 데에서도 매우 큰 권력을 갖고 있으며 정치엘리트는 그들이 회피하고자 하는 이슈는 논의될 정치적 의제에서 배제하는 방식으로 권력을 사용한다고 주장한다(Bachrach and Baratz 1962). 또한 시민들이 선거를 통해 정치인들을 선택하고 정책결정과정에서 여러 수단과 방법을 통해 참여할 수 있기는 하나, 일반 시민들은 의제설정권력(agenda-setting power)의 측면에서 언론에 비해 미약한 영향을 미친다(Lee 2014).

하지만 앞서 언급한 시민 참여에 대한 회의론이나 민주주의의 위기에 대한 진단들은 대중의 분노와 혐오를 자극하는 포퓰리즘과 같은 감정적 동원이 합리적인 의사결정을 저해하여 민주주의의 정상적 작동을 어렵게 한다고 지적한다. 이러한 판단은, 특히 분노나 혐오와 같은 부정적인 감정을 포함하는 감정적인 요소를 합리적인 판단과 대비시키며 부정적인 기능을 강조하는 암묵적인 가치판단을 내포하고 있다(Demertzis 2006). 의제설정 과정에 감정이 투입될 경우, 바람직하지 않은 결과가 나온다는 암묵적 전제가 숨어있는 것이다. 그렇지만 정치적 참여에서 감정이 수행하는 역할에 대한 경험연구들은 다소 다른 결과들을 보고한다. 대의 민주주의 체제에서 가장 중요한 제도적 참여인 선거에서부터, 높은 기회비용을 요구하는 집합행동에의 참여까지 감정은 참여를 추동하는 역할을 한다(Brader 2005). 흥미로운 점은 열정과 같은 긍정적인 감정뿐만 아니라 분노, 슬픔, 공포와 같은 부정적인 감정들 역시 참여에 긍정적인 영향을 미치며 오히려 전자에 비해 효과가 더 크다는 점이다(Valentino et al 2011; Weber 2013).

감정은 참여에 지불해야 하는 비용보다 참여로 인한 편익이 더 적기 때문에 발생하는 집합행동의 딜레마를 해결하여 정치적 참여를 가능하게 만들어주는 기능을 한다(Groenendyk 2001). 참여에 미치는 효과가 크게 나타나는 분노의 경우에는 민주주의 이행기에 자신과 타인에 대한 부정의에 민감하게 반응하여 민주주의를 심화하는 데에 도움이 되거나(Mihai 2016), 더 나아가 사회정의를 실현하기를 주저하는 사법부와 같은 제도 내부 기구들을 민주적 방식으로 작동하도록 압박하는 동력이 되기도 한다(Bakiner 2016). 다시 말해, 감정이 강하게 작동하는 정치적 참여의 결과와는 별개로, 감정이 없으면 민주적 정치과정 자체가 성립하기 어렵다고 할 수 있다. 감정과 정치참여의 관계에 대한 기존 연구들을 종합하면, 감정은 수용자의 의식에 영향을 미쳐 정치적 참여를 촉진하거나 혹은 수용자 스스로가 특정 감정을 가짐으로써 정치적 참여를 결심하게 만든다. 그리고 이러한 효과는 부정적 감정, 그 중에서도 분노와 관련하여 가장 강하게 나타난다.

하지만 참여가 이루어진 다음 단계에서 감정이 수행하는 역할에 대해서는 상대적으로 조명이 덜 되었다. 페이스북에서 자폐증 범주 권익옹호 시민단체와 장기기증 홍보 시민단체를 상대로 한 연구는 감정 내부의 경향성보다는, 감정적 수사와 인지적 수사를 구분하여 두 수사 사이에 지배적 경향을 역전시키는 포스트에 구독자들이 더 적극적으로 반응한다는 결과를 제시한다(Bail et al. 2017). 해당 연구는 온라인 공간에서도 시민들이 단순히 정서적으로만 행동하거나 계산만 하는 존재가 아니라, 더 복잡한 동학에 의해 판단을 내리는 존재라는 점을 보여주고 있으나 감정의 영역을 세분화하지 않았기 때문에 감정 자체가 수행하는 역할을 보여주고 있지는 못하다. 따라서 본 연구에서는 시민들의 정치 참여가 이루어진 이후의 시점에서, 선행연구들에서 공통적으로 정치적 참여에 가장 큰 영향을 미친 분노에 대해 시민들이 반응하는 방식을 살펴보도록 할 것이다.

청와대 국민청원은 참여가 이루어진 이후 시민들의 반응을 살펴보기에 용이한 공간이다. 공개적으로 접근 가능한 온라인 게시판이기 때문에, 다수의 시민이 참여할 수 있으며 작성된 글에서 많은 정보를 추출할 수 있다. 또한 청와대 국민청원에서 제기된 많은 청원들은 언론을 통해 기사화되고 관련된 내용

이 국회에서 언급되기도 하는 등 시민들이 정치적 의제를 설정하는 공간으로 기능하고 있기 때문에, 정치과정이 이루어지는 공간으로서 연구할 가치를 지닌다. 이러한 현실을 고려해보면 청와대 국민청원에 대한 분석은 의제설정 과정에서 감정에 시민들이 반응하는 방식을 규명함으로써, 정치 참여와 감정 사이의 상호작용을 규명하는 데에 필요한 경험적 근거를 제시할 수 있을 것이다.

2. 청와대 국민청원에 대한 논쟁

온라인 공간을 통한 국민청원 제도는 시민들의 의제설정권력을 강화하는 기제로 작동할 수 있다. 하지만 정치참여에서 감정의 역할에 대한 논쟁과 마찬가지로, 실제로 운영되고 있는 청와대 국민청원에 대한 평가는 극명하게 엇갈린다.

우선 청와대 국민청원에 대한 부정론의 주장은 다음과 같다. 첫째, 정보격차 문제가 제기된다. 정보통신기술을 활용하는데 있어서 나타나는 정보격차가 심할수록⁴ 정보활용능력과 물적, 인적자원을 가진 사람이 정부정책에 영향력을 미칠 수 있는 가능성은 높아지지만 그러한 능력과 자원을 충분히 갖지 못한 사람들의 요구는 정부정책에 투입되기 어려워진다. 둘째, 익명의 불특정 다수가 참여하여 이루어지는 정치적 결정에는 책임성이라는 민주주의에서 매우 중요한 가치가 담보되기 어렵기 때문에 전자 포퓰리즘을 야기할 수 있다(임혁백 2011: 6). 뿐만 아니라 청와대 국민청원에서는 특정 집단에 대한 혐오감정에 기댄 청원들이 제기될 뿐만 아니라, 그러한 청원이 때로는 상당히 높은 수의 동의를 얻는 등 등 시민들이 공론장에서 지나치게 사나워지고 그럼으로써 공론장이 피폐해지는 결과를 낳기도 한다(박상훈 2018).

이러한 부정론에 반박하는 긍정론은 청와대 국민청원이 사회권 보장이나 젠더 불평등 문제와 같이, 기존의 정치제도에서 제대로 대표되지 못하였던 의미 있는 이슈들이 공론화되는 공간이라는 분석 결과를 제시한다. 한겨레경제사회연구원의 분석에 따르면 청와대 국민청원에서는 “폐법 아닌 성평등·인권 등 ‘사회권 보장’ 요구가 높았다(한겨레경제사회연구원 2018).” 여성이나 아동과 같은 사회적 약자에 대한 언급이 높은 비중을 차지하고 있으며, 이러한 단어들은 성폭력, 차별, 인권, 노동, 임금, 고용 등 사회적 권리와 밀접한 단어들과 긴밀하게 연결되어 나타났다. 또, 한국여성정책연구원의 최근 조사 결과에서도 청와대 국민청원에서 젠더이슈가 상당한 비중을 차지하고 있다는 것이 나타났다(김원정, 정윤미 2019).

두 상반된 주장들은 각자의 경험적 근거를 제시하고 있기 때문에, 무엇이 더 타당한 주장인지를 가려

⁴ 정보격차는 경제적 여건으로 인하여 정보통신기술을 활용할 수 없는 물리적 접근성의 문제로도 다루어지지만(van Dijk&Hacker 2003; Norris 2001; Livingstone and Helsper 2007), 정보통신기술을 자신의 목적을 달성하기 위한 수단으로서 적절히 활용할 수 있는 역량의 격차로도 다루어진다(van Dijk 2003; 2005). 한국의 경우 정보통신기기의 보급은 보편적으로 이루어져 있어 물리적 접근성의 문제는 상대적으로 적지만 정보를 선별적으로 취득하고 가공하는 능력에 있어서는 사람들 사이에 차이가 존재할 수밖에 없다. 즉 여기서 말하는 정보격차란 물리적 접근성으로서의 정보격차라기보다는 인터넷과 정보를 얼마나 잘 활용하는지를 의미하는 역량격차로서의 정보격차를 의미한다.

내기는 쉽지 않다. 청와대 국민청원에서 매우 많은 동의를 얻는 청원들의 상당수가 대중의 분노가 쉽게 집약될 수 있는, 특정 집단이나 개인과 관련된 사건·사고에 집중되어 있다는 점에 주목하면 박상훈의 주장처럼 청와대 국민청원은 정제되지 않은 분노의 분출구에 지나지 않는 것으로 보인다. 그러나 몇몇 조사들이 보여주듯이 인권, 성평등, 노동, 복지, 환경 등 사회적 기본권에 관련된 의제들이 제기되는 공간인 것도 분명한 사실이다. 그런데 방법론적 문제를 다루기에 앞서, 우선 앞 절에서 지적하였다시피 분노의 분출구라는 관점과 사회적 기본권의 대표라는 관점은 양립 불가능하지 않다. 사회적 기본권의 침해에 대한 공분으로 격앙된 형식의 문제제기를 하는 상황도 얼마든지 가능하다. 그리고 분노라는 감정에 대한 사전적 가치판단을 배제한다면, 이러한 형식의 참여가 민주주의에 부정적 영향을 미칠 것이라고 쉽게 단정지을 수 없다. 따라서 본 연구에서는 청와대 청원 게시판에서 제기되는 의제들의 성격과, 각 의제들이 감정적으로 제시되는 방식을 모두 분석하여 통합적 이해를 시도한다. 그렇다면 두 주장은 모두 충분한 경험적 근거를 제시하고 있는가?

한겨레경제사회연구원과 한국여성정책연구원의 분석은 각각 2만 명, 1만 명 이상의 동의를 얻은 소수의 청원문서만으로 분석대상을 제한했다는 한계를 지닌다. 또한 두 분석 모두 개별 단어들 사이의 연결망을 기준으로 분석을 수행하였기 때문에, 단어가 사용된 맥락을 제대로 반영하지 못하고 연구자의 사전적 가치판단이 과도하게 개입되고 있다는 방법론적 한계를 지닌다. 다른 한편으로 부정론은 청원 게시물을 수집하여 분석하지는 못했기 때문에, 청와대의 응답기준을 상회할 정도로 매우 높은 동의를 얻은 극소수의 게시물에 한해서만 타당하다는 문제가 있다. 다시 말해 두 주장 나름의 경험적 근거를 제시하고 있으나, 분석 대상이 되는 경험적 자료의 범위가 제한적이며 분석 방식 역시 상당히 주관적이다.

따라서 본 연구는 청와대 국민청원 게시판에 게시된 청원 전수⁵를 분석함으로써 통합적 이해에 필요한 경험적 근거를 제시하고자 한다. 특히, 청와대 국민청원에 사회적 기본권에 대한 주장이 다수 게시된다는 긍정론은 1만 또는 2만과 같은 동의수 하한선을 둔 분석에 근거하고 있고, 청원게시판에서 높은 동의를 얻는 청원글은 시민들의 정념을 동원하기 용이한 이슈에 관련된 것들이라는 부정론은 (명시적이지는 않지만) 청와대의 응답기준을 상회하거나 그에 준하는 수준의 게시물에 초점을 맞추고 있다는 점에 주목한다. 이는 분석 대상을 더 넓은 범위로 확대하였을 경우, 의제의 성격과 감정의 상호작용에 따라 다양한 현상이 나타날 가능성이 있음을 시사하기 때문이다. 이러한 분석을 통해 서로 대립하는 두 주장에 포괄적인 경험적 근거를 제시하고, 온라인 공론장이 더 생산적으로 운영되기 위한 조건을 탐색하는 데에 도움이 될 수 있을 것이다.

⁵ 물론 중복된 게시물, 장난 청원 등 정치적 의미를 찾기 어려운 청원 문서까지 대량 포함될 경우 머신러닝을 통한 학습 과정에서 상당한 잡음(noise)이 발생할 수 있기 때문에 이러한 문서들은 제외할 것이다.

III. 분석방법⁶

본 연구는 청와대 국민청원 게시판에 업로드 된 문서 전수를 웹크롤링(web crawling)을 통해 수집하여 분석한다. 분석대상 게시물은 청와대 국민청원 게시판이 개설된 2017년 8월 19일부터 2018년 9월 30일까지 작성된 게시물이다.⁷ 본 연구에서는 총 30여만 건(중복글 및 무의미한 게시물 제외⁸ 88,113건)의 게시물을 분석하기 위해 컴퓨터를 이용한 자연어처리(NLP: Natural Language Processing)를 적극적으로 활용한다. 자연어처리는 텍스트나 음성과 같이 인간이 생산하는 다양한 종류의 자연어를 컴퓨터를 통해 분석이 가능한 형태로 가공하는 과정을 지칭한다. 따라서 NLP는 컴퓨터 과학, 정보과학, 통계학, 언어학과 같은 다양한 분야들의 협업을 통해 인간의 자연어 생성 과정을 분석하고, 일반화된 규칙을 통해 언어가 가지는 의미와 맥락을 계량화된 형태로 재현하는 일을 목표로 한다(Chowdhury 2003). 이러한 NLP 기법은 인간이 직접 검토하기 어려운 방대한 양의 자연어 자료 전수에서 의미를 추출하여 인간이 검토할 수 있는 수준의 차원으로 축소해준다는 장점을 지니기 때문에 본 연구에 유용한 분석 도구로 기능할 수 있다. NLP에는 다양한 기법들이 존재하지만, 본 연구에서는 토픽모형의 일종인 구조적 토픽 모형(Structural Topic Model)을 활용하여 개별 문서들을 특정 주제들로 분류하고, 워드 투 벡터(Word2Vec)를 통해 각 문서들에 감정 점수를 매긴다. 이후 음이항 회귀 분석(Negative Binomial Regression)을 통해 게시물의 주제와 감정이 동의수에 미치는 통계적 효과를 추정하고, 청와대의 응답을 받은 게시물이 가지는 기술통계적 특성을 살펴본다.

1) 구조적 토픽 모형(Structural Topic Model)

구조적 토픽 모형은 토픽 모형(topic model)의 일종이다. 토픽 모형은 분석 대상이 되는 전체 문서 집합을 연구자가 사전에 설정한 특정한 개수의 토픽으로 분류해주는 기법이다. 해당 기법은 생성적(generative) 모형으로서, 인간의 문서 생성 과정을 모사한 통계모형을 설정한 후, 주어진 문서의 단어 분포를 기반으로 모형의 파라미터(parameter)를 역으로 추정하여 토픽을 산출하기 때문에 생성적 모형이라고 지칭한다.

인간은 글을 작성할 때 주제를 사전에 설정한 후, 주제와 관련된 단어들을 배치하는 과정을 거쳐 글을 만들어낸다. 토픽 모형은 이를 모사하여, 문서에 반영된 특정한 토픽들이 사전에 존재하며 토픽에

⁶ 해당 연구에서 자료 수집 및 텍스트의 전처리 과정에는 Python의 Konlpy라이브러리의 Kommoran 모듈을 활용하였으며, 분석과정에는 R의 stm(구조적 토픽모형), wordvectors(워드 투 벡터), MASS(음이항회귀) 패키지를 활용하였다.

⁷ 이는 청와대 국민청원 게시판이 개설된 후 본 연구가 개시된 시점까지를 포함하는 시기이다.

⁸ 청와대 국민청원 게시판에는 각종 사안에 대한 시민의 청원사항이 게시되기도 하지만, 청원문서의 내용이 거의 없는 게시물이나 ‘유럽 챔피언스 리그 결승전 재경기 청원’ 등 장난성 게시물도 상당수 존재한다. 이러한 문서들까지 포함하여 토픽모형을 구축하게 되면 토픽의 추출에 적지 않은 잡음이 포함되게 되고 이는 토픽의 해석에도 부정적 영향을 미칠 가능성이 크다. 이러한 문제를 해결하기 위해 제목과 내용이 완전히 동일한 게시물(즉, 복제된 글)의 평균적 동의수(10회) 미만의 게시물을 일괄적으로 제외하였다.

따라 선호하는 단어들이 다르며 각각의 문서들은 서로 다른 토픽들의 조합으로 이루어져 있기 때문에 단어들 역시 토픽의 비중을 따라 발현된다는 가정을 취한다(Blei et al, 2003).

따라서 토픽 모형에서 토픽이란 문서를 구성하는 전체 단어 집단에 대한 확률분포를 의미한다. 하지만 분석 대상이 되는 문서는 문서 작성자가 염두에 둔 토픽 자체에 대한 정보를 가지고 있지 않기 때문에, 문서에 발현된 단어의 분포를 근거로 문서 생성의 근원이 된 토픽의 구성을 추정하게 된다. 그리고 구조적 토픽 모형은 토픽 추정 과정에서 개별 문서의 메타데이터(작성일시, 작성자 등)를 투입하기 때문에 실제 인간의 문서 작성 과정과 더욱 유사하게 추정을 수행한다(Roberts et al 2013; 2014).

토픽 모형은 정치학 분야에서 대규모의 텍스트를 분석할 때 활발하게 활용되고 있다. 미국 상원의원들의 성향을 분류하고, 다른 분류 방식과 비교한 연구들은(Grimmer 2010; Quinn et al 2010) 토픽 모형이 정치적 텍스트를 분류할 때 신뢰할만한 성능을 가지고 있음을 보여준다. 구조적 토픽 모형은 메타데이터를 추정 과정에 반영할 수 있다는 특성으로 인해, 이슬람 급진주의 성직자의 설교문에 대한 분석에서부터 중국 관영언론의 보호무역 효과에 대한 분석, 유튜브에서 소수자에 대한 태도 분석에 이르기까지 최근 사회과학 전반에서 활발하게 활용되고 있다(Lucas et al 2015; Egami et al 2018; Kim 2018; Schwemmer and Jungkunz 2019).

다만 본 연구에서는 다중 멤버십을 특징으로 하는 일반적인 토픽 모형과 달리, 정치적 문서의 특성을 근거로 개별 문서를 하나의 토픽으로 배정한 접근(Quinn et al 2010 213)을 수용한다. 따라서 토픽의 발현비중을 그대로 연속형 변수로 활용하지 않고, 토픽의 발현비중을 기준으로 각 문서를 가장 비중이 높은 토픽으로 분류하도록 할 것이다. 따라서 구조적 토픽 모형을 통해 추정된 결과는 개별 문서들이 각 문서 내에서 가장 비중이 높은 토픽에 할당된 범주형 변수의 형태를 지닌다.

2) 감성분석(Sentiment Analysis)

본 연구의 목적은 감정이 의제와 상호작용하며 사람들의 반응에 미치는 영향을 분석하는 것이기 때문에, 각 청원에서 나타난 감정을 조작화하는 작업이 필수적이다. 감성분석을 수행하는 방식에는 여러 가지가 있으나 가장 손쉽게 사용되는 방식은 감성을 표현하는 단어 사전과 문서를 대조하여 각 문서에 감성 점수를 매기는 방식이다. 토픽 모형을 활용한 변수와 감성을 동시에 투입한 선행연구들(Bail et al 2017; Kim 2018) 역시 이러한 방식으로 감성분석을 수행한다. 하지만 본 연구에서는 두 가지 이유로 인해 대안적 방식을 활용한다.

첫째, 청와대 청원 게시판에 작성된 글들은 인터넷 게시판과 정치 텍스트라는 특성이 결합하여 언어적 표현과 은어들이 다수 등장하기 때문에 기존의 한국어 감성 사전(박상민 외 2018)이 질적으로 신뢰할만한 성능을 보여주지 못한다. 둘째, 대부분의 감성사전은 긍정과 부정 두 형태로만 감정을 분류하기 때문에 본 연구의 목적에 부합하지 않음과 동시에 청와대 청원 게시판에도 잘 적용되지 않는다. 우선 본 연구는 기존 연구에서 참여를 추동하였던 분노와 같은 구체적 감정의 역할을 살펴보는 것을 목적으

로 하고 있으며, 청와대 청원 게시판은 시민들이 주로 호소를 하는 공간이기 때문에 긍정적인 주제는 찾아보기 어렵다.

따라서 해당 연구에서는 워드 투 벡터(Word2Vec)를 활용하여 각 문서들에 분노점수를 매기며, 보조적으로 슬픔점수를 매긴다. 워드 투 벡터는 신경망 학습 방식을 활용하여 전체 문서 내부 단어들의 등장 순서를 예측하는 가중치를 학습시킨 후, 해당 가중치를 각 단어의 벡터공간에서의 좌표값으로 활용하는 단어 임베딩(word embedding) 방식이다(Mikolov et al 2013; Goldberg and Levy 2014). 이를 통해 단어 사이의 의미상의 유사도를 계량화할 수 있기 때문에, ‘분노’와 ‘슬픔’이라는 단어와 각 문서 내 단어들의 평균 유사도를 정규화하여 각 감정에 대한 점수로 활용하도록 한다.

3) 음이항 회귀분석(Negative Binomial Regression)

본 연구의 종속변수는 청원게시판에서 시민들의 호응도를 보여주는 각 게시물에 대한 동의수이다. 하지만 동의수의 분포는 정규분포를 따르지 않으며, 분산(62,264,249) 역시 평균(471.1244)에 비해 매우 크기 때문에 푸아송 회귀분석의 일반화된 형태인 음이항 회귀분석을 활용하였다. 분석식은 앞서 토픽모형과 감성분석을 통하여 구축한, 개별문서 d 가 할당된 k 번째 토픽 Topic 과 개별 문서 d 의 분노점수 Anger 및 슬픔점수 Sadness, 통제변수인 문서 길이와 외부링크 포함 여부(이진변수)를 나타내는 Z , 분노가 각 의제에서 발휘하는 효과를 검증하기 위한 Anger 와 Topic 의 상호작용항으로 구성되었다. 따라서 음이항 회귀분석을 통해 추정하려는 최종 모형은 아래와 같다.

$$\log(\text{Vote}_d) = \beta_0 + \beta_1 \text{Anger}_d + \beta_2 \text{Sadness}_d + \sum_{k \in \text{number of all Topics}} \beta_{3k} T_{dk} + \sum_{k \in \text{number of all Topics}} \beta_{4k} \text{Anger}_d \cdot T_{dk} + \delta Z_d + \epsilon_d$$

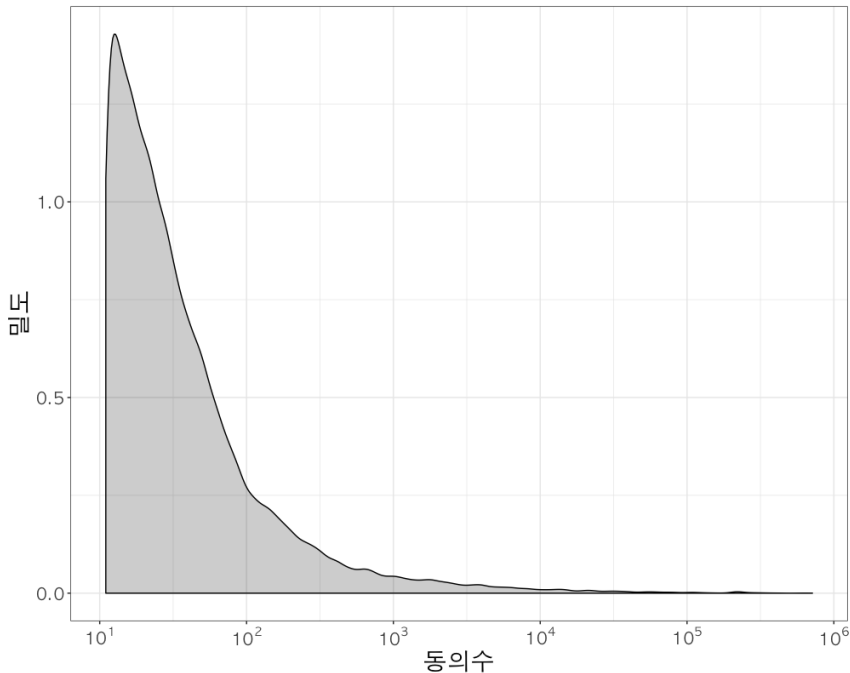
$$T_{dk} = \begin{cases} 1 & \text{if } k = \text{argmax}(\text{Topic}_{dk}) \\ 0 & \text{if } k \neq \text{argmax}(\text{Topic}_{dk}) \end{cases}$$

위 모형에서 문서 d 에서 가장 높은 비중을 가진 k 번째 토픽은 이진변수의 형태로 투입되어, 해당 문서의 동의수에 미치는 문서 주제의 영향력으로 해석된다.

IV. 청와대 국민청원 분석결과

2017년 8월 19일부터 2018년 9월 30일까지 청와대 국민청원 사이트에 게시된 모든 게시물을 수집하여 우선 동의수 분포를 살펴보았다. <그림 1>은 수집된 전체 청원문서의 동의수 분포를 보여준다. 추천을 거의 받지 못하는 게시물이 대다수이며 동의수 100회를 상회하는 게시물도 전체 게시물의 수에 비하여 극소수에 불과함을 알 수 있다. 즉 청와대 국민청원에서 시민들의 호응을 조금이라도 얻는 일은 매우 어려운 일이라는 것을 알 수 있다. 최근 청와대가 국민청원 사이트 운영 규정을 변경하면서 100명 이상이 사전동의한 게시물만이 국민청원 게시판에 노출되도록 하였는데, 아래의 분포를 통해 미루어보면 향후에는 청와대 국민청원 게시판에 노출되는 게시물 자체가 매우 적어질 것이라고 예상할 수 있다.

| 그림 1 | 전체 청원수 분포



수집된 게시물들의 내용을 토픽모델을 통해 분석한 결과로 총 28개의 토픽이 추출되었다.⁹ 그런데 토픽모형은 문서에 존재하는 토픽을 분류하고 각 토픽별 최다빈출 단어와 독점성 단어¹⁰를 추출해줄

⁹ 반비례 관계인 구조적 토픽 가능성(held-out likelihood)과 의미론적 일관성(semantic coherence)을 기준으로 25~32개의 토픽이 최적 범위로 나왔으며, 여러 모형을 구축한 후 질적으로 가장 명확하게 구분이 되는 토픽의 개수를 선택하였다.

¹⁰ 최다빈출 단어는 다른 토픽들과 비교했을 때 해당 토픽에서 단순빈도 기준으로 가장 빈번하게 등장하는 단어들을 뜻하고, 독점성 단어는 다른 토픽에서는 잘 출현하지 않는 반면 해당 토픽에서 등장할 확률이 높은 단어를 의미한다. 따라서 독점성 단어는 해당 토픽의 특징을 대표할 수 있는 단어라고 할 수 있다.

뿐, 토픽의 명칭을 정해주는 것은 아니다. 따라서 토픽의 명칭은 저자들이 최다빈출단어와 독점성 단어를 살펴본 뒤 해당 단어들로부터 유추할 수 있는 토픽명을 임의적으로 명명하였다. 토픽명칭과 각 토픽별 최다빈출단어와 독점성 단어, 그리고 전체 문서에서 차지하는 비중과 개별 문서 내부 비중을 기준으로 분류한 빈도는 아래 <표 1>에 제시되어있다. <표 1>의 토픽은 빈도를 기반으로 내림차순으로 나열하였다. 그리고 대명사나 일반명사 등을 포함하는 ‘더미’ 토픽은 특정한 의제를 대변하지 않기 때문에 개별 문서의 의제 분류시에 포함시키지 않았다.

[표 1 | 청와대 국민청원 게시판에서 추출된 토픽 목록

연번	토픽	최다빈출단어	독점성 단어	비중	빈도
1	범죄	처벌, 피해자, 범죄, 사건, 가해자, 피해, 폭행	가해자, 무고, 조두순, 몰카, 폭행, 무고죄, 미투운동	0.095	6148
2	부동산	주택, 부동산, 집값, 아파트, 서민, 정책, 투기	집값, 부동산, 폭등, 무주택, 보유세, 주택임대	0.064	5525
3	조세/준조세	국민, 세금, 공무원, 연금, 국민연금, 폐지, 서민	연금, 국민연금, 누진세, 에어컨, 누진, 기세, 고갈	0.053	5283
4	국회비판	국회의원, 국민, 국회, 의원, 선거, 정치, 자유한국	국회의원, 국회, 선거, 세비, 특활비, 해산, 자유한국	0.052	5237
5	대통령	대통령, 국민, 정부, 문재인, 정권, 청와대, 당신	문재인, 대통령, 촛불, 박근혜, 당신, 정권, 지지율	0.047	5022
6	난민	난민, 외국인, 한국, 국민, 자국민, 우리나라, 이슬람	난민, 자국민, 이슬람, 제주도, 예멘, 무슬림, 체류	0.045	4806
7	검찰수사	사건, 조사, 수사, 검찰, 경찰, 비리, 검사	이재명, 특검, 드루킹, 진실, 조폭, 검찰, 김경수	0.045	4177
8	외교/안보	북한, 우리, 일본, 평화, 역사, 미국, 전쟁	독도, 남한, 비핵화, 북한, 천안함, 김영철, 폭침	0.044	3958
9	인터넷/언론	청원, 기사, 언론, 내용, 방송, 뉴스, 사이트	일베, 네이버, 사이트, 게시, 링크, 방송, 언론사	0.044	3784
10	노동	근무, 최저임금, 회사, 근로자, 임금, 직원, 시간	근로자, 비정규직, 임금, 근로시간, 최저임금, 최저시급, 연차	0.044	3703
11	스포츠	선수, 올림픽, 감독, 경기, 축구, 스포츠, 국가대표	선수, 축구, 국가대표, 빙상연맹, 노선영, 축구협회, 월드컵	0.035	3630
12	금융/가상화폐	공매도, 주식, 투자, 기업, 거래, 가상화폐, 시장	공매도, 주식, 가상화폐, 증권, 거래소, 코인, 블록체인	0.034	3256
13	교육/입시	학생, 교육, 학교, 시험, 교사, 대학, 공부	영어, 수능, 수시, 사교육, 과목, 교육부, 학생	0.033	2932
14	성별갈등	여성, 남성, 남자, 여자, 군대, 사회, 평등	여성가족부, 페미니즘, 페미니스트, 여성, 국방의무, 성차별, 남성	0.03	2931
15	보육	아이, 교사, 어린이집, 부모, 학교, 시간, 선생님	어린이집, 보육교사, 보육, 교복, 유치원, 아동학대, 유아	0.029	2845
16	사법불신	판결, 판사, 재판, 헌법, 법률, 사법부, 법원	판사, 사법부, 판결, 대법원, 정형식, 재판, 사법	0.027	2777

연번	토픽	최다빈출단어	독점성 단어	비중	빈도
17	갑질	사용, 판매, 불법, 대한항공, 나이, 청소년, 게임	대한항공, 생리대, 고객, 항공사, 한진, 조현민, 신분증	0.027	2657
18	음주운전	사고, 차량, 경찰, 안전, 소방관, 발생, 음주운전	운전자, 음주운전, 차량, 주차, 운전, 차주, 소방	0.026	2576
19	이웃분쟁/동물	동물, 흡연, 담배, 사람, 강아지, 반려동물, 학대	동물, 강아지, 반려견, 식용, 동물학대, 안락사, 층간소음	0.026	2485
20	종교/이념	대한민국, 나라, 국민, 국가, 우리, 사회, 우리나라	나라, 대한민국, 자유, 다운, 교회, 종교, 동성애	0.025	2406
21	환경/에너지	미세먼지, 중국, 원전, 사용, 환경, 정부, 문제	미세먼지, 방사, 발전소, 발암물질, 원전, 태양광, 댐시	0.025	2261
22	의료	병원, 치료, 환자, 의료, 간호사, 수술	병원, 환자, 간호사, 의료사, 치과, 수술, 대학병원	0.024	2172
23	저출산	지원, 소득, 가정, 혜택, 결혼, 자녀, 아이	부부, 양육비, 난임, 한부모, 자녀, 이혼, 건강보험료	0.022	2077
24	아파트	사업, 관리, 계약, 공사, 아파트, 진행, 업체	조합, 택배, 계약, 조합원공사, 건축, 입찰	0.022	1724
25	생활민원	지역, 주민, 서울, 시민, 지방, 계획, 개발	개인회생, 변제, 지역, 마을, 주민, 노선, 회생	0.019	1695
26	정책의견	정책, 문제, 제도, 정부, 국가, 현재, 경제	도입, 감소, 분야, 변화, 증가, 방식, 방향	0.019	1530
27	장애인복지	활동, 장애인, 단체, 사회, 시설, 센터, 복지	장애인, 특수, 활동, 집회, 애인, 복지사, 장애	0.018	516
28	데미	사람, 생각, 보고, 마음, 자기, 정도, 하나	애기, 가요, 소리, 사람, 생각, 그때, 화가	0.012	0

추출된 토픽들은 정치/경제/사회 문제를 주로 담고 있다. 몇 가지 주요한 토픽들을 살펴보면 청와대 국민청원 게시판에서 주로 어떠한 토픽이 어떠한 내용을 중심으로 다루어지는지 알 수 있다.

추출된 토픽 중 상당수의 토픽은 정부정책에 관련된 내용들을 담고 있다. ‘외교/안보’ 토픽의 독점성 단어들을 살펴보면 외교/안보 관련된 담론은 주로 남한, 비핵화, 북한, 천안함, 김영철(북한 조선노동당 통일전선부장)과 같이 북한 관련된 문제를 중심으로 제기되고 있는 것으로 보인다. ‘조세/준조세’ 토픽은 세금이나 국민연금보험료와 같은 조세와 준조세, 또는 공공서비스 비용과 관련된 내용들을 포괄하는 토픽이다. 이 토픽의 독점성 단어에 에어컨, 누진세 등이 포함된 것으로 보아 여름철 전기사용료 관련된 내용들과 국민연금, 고갈 등이 포함된 것은 국민연금 개편 등의 의제에 관련된 내용들이 주로 언급된 것으로 보인다. 이외에도 ‘교육’ 토픽의 경우 ‘수능’, ‘수시’와 같은 대입에 관련된 단어들이 독점성 단어로 존재하고 있으며 해당 단어들을 고려해보았을 때 대학입시제도에서 정시/수시 논쟁을 의미하는 것으로 보인다. 이외에도 부동산 정책이나 아파트 건축 관련된 토픽들이 추출되었으며 층간소음, 애완 동물 등 이웃과의 분쟁요소에 대한 내용도 별도의 토픽으로 존재하고 있다.

정책에 관련된 토픽 외에도 국민들의 관심을 끌었던 사건사고에 대한 청원들에 관련된 토픽들도 추출되었다. 대표적으로 ‘범죄’ 토픽은 강력범죄 가해자에 대한 내용들을 중심으로 하고 있으며 개별 계

시물 내 비중을 기준으로 게시물을 분류하였을 때, 가장 많은 게시물의 주제를 대표한다. 흥미로운 것은 ‘음주운전’의 경우 개념적으로는 범죄에 포함되지만 청와대 국민청원 게시판에 존재하는 토픽으로서는 ‘범죄’ 토픽과는 별도의 토픽으로 분류되었음을 알 수 있다. 이는 토픽 내에 등장하는 소방관, 경찰과 같은 단어를 통해 알 수 있다시피, 가해자에 대한 분노보다 피해자 및 단속 주체들에 대한 공감 및 연민을 보이고 있기 때문이다. 그리고 ‘스포츠’ 토픽은 국가대표 팀에 관련된 내용들이 주로 제기되고 있으며, 빙상연맹, 노선영, 축구협회와 같은 독점성 단어들로 이루어보아 지난 동계올림픽에서 여자 팀 추월 경기나 월드컵에 관련된 이슈들이 해당 토픽에서 특징적인 단어로 나타나고 있다고 볼 수 있다.

위에 소개한 토픽들 외에 해석에 주의가 필요한 토픽들에 대해 설명하면 다음과 같다. ‘정책의견’ 토픽의 경우 대부분 정책에 관련된 내용을 나타내는 토픽이다. 그러나 어떤 정책인지는 명확하지 않다. ‘정책의견’ 토픽은 특정한 정책부문과 연관된 토픽이라기보다는 청원자가 특정한 제도의 도입을 주장하거나, 특정한 제도의 효과를 이야기하거나 정책이 나아가야 할 방향을 언급하는 것과 관련된 토픽에 가깝다. ‘인터넷/언론’ 토픽의 토픽비중이 높은 게시물들의 내용을 별도로 살펴보면 청원문서 작성자가 자신이 청원하는 내용이 인터넷에서 이슈가 되고 있다는 점을 강조하거나 특정 언론 보도를 비판하며 여러 사이트나 관련된 하이퍼링크를 언급하는 표현들이 많다.

아래 <표 2> 는 III.(3).항의 모형에 따라 음이항 회귀분석을 실시한 결과이다. 종속변수가 로그형태이기 때문에, 계수는 한계효과(marginal effect)를 나타낸다.

| 표 2 | 음이항 회귀 분석: 감정과 의제가 동의수에 미치는 영향

	종속변수: 동의수		
	(1)	(2)	(3)
분노지수		-0.026** (0.009)	0.112*** (0.029)
슬픔지수		0.028** (0.009)	0.053*** (0.009)
토픽: 대통령 (기준: 외교/안보)	0.096* (0.044)	0.096* (0.045)	0.370*** (0.051)
토픽: 보육	-0.365*** (0.047)	-0.380*** (0.048)	-0.615*** (0.080)
토픽: 조세/준조세	-0.339*** (0.042)	-0.317*** (0.043)	-0.490*** (0.046)
토픽: 생활민원	-0.395*** (0.046)	-0.382*** (0.046)	-0.298*** (0.049)
토픽: 성별갈등	-0.693*** (0.043)	-0.694*** (0.043)	-0.723*** (0.044)
토픽: 아파트	-0.248*** (0.045)	-0.249*** (0.045)	-0.623*** (0.065)
토픽: 종교/이념	-0.659*** (0.042)	-0.640*** (0.044)	-1.290*** (0.053)
토픽: 검찰수사	0.420*** (0.041)	0.402*** (0.043)	-0.007 (0.049)
토픽: 갑질	0.067 (0.048)	0.032 (0.049)	0.024 (0.061)

	종속변수: 동의수		
	(1)	(2)	(3)
토픽: 환경/에너지	-2.019*** (0.042)	-2.014*** (0.042)	-2.117*** (0.043)
토픽: 이웃분쟁/동물	-0.202*** (0.048)	-0.170*** (0.049)	-0.291*** (0.059)
토픽: 부동산	-0.091 (0.055)	-0.106 (0.055)	-0.315*** (0.078)
토픽: 정책의견	-0.403*** (0.047)	-0.414*** (0.047)	-0.453*** (0.053)
토픽: 사법불신	0.117** (0.045)	0.142** (0.046)	-0.398*** (0.052)
토픽: 노동	-0.023 (0.055)	-0.017 (0.055)	0.653*** (0.080)
토픽: 장애인복지	-0.754*** (0.044)	-0.785*** (0.045)	-0.859*** (0.045)
토픽: 국회비판	-0.028 (0.049)	-0.067 (0.049)	-0.150** (0.055)
토픽: 스포츠	0.365*** (0.051)	0.326*** (0.052)	0.204** (0.065)
토픽: 범죄	0.067 (0.049)	0.023 (0.051)	0.011 (0.055)
토픽: 음주운전	0.633*** (0.045)	0.628*** (0.045)	0.847*** (0.050)
토픽: 난민	-0.237** (0.085)	-0.245** (0.085)	-0.177 (0.103)
토픽: 저출산	-0.528*** (0.052)	-0.552*** (0.052)	-0.454*** (0.080)
토픽: 교육/입시	-0.502*** (0.057)	-0.508*** (0.057)	-1.106*** (0.076)
토픽: 인터넷/언론	-1.671*** (0.042)	-1.669*** (0.042)	-1.786*** (0.043)
토픽: 의료	-0.076 (0.050)	-0.112* (0.050)	-0.274*** (0.055)
토픽: 금융/가상화폐	0.130* (0.050)	0.121* (0.051)	0.219*** (0.065)
본문길이	0.001*** (0.00001)	0.001*** (0.00001)	0.001*** (0.00001)
외부링크	0.944*** (0.038)	0.942*** (0.038)	0.982*** (0.038)
분노지수*대통령			-0.815*** (0.046)
분노지수*보육			-0.315*** (0.065)
분노지수*조세/준조세			0.046 (0.041)
분노지수*생활민원			0.240*** (0.052)
분노지수*성별갈등			-0.733*** (0.047)

	종속변수: 동의수		
	(1)	(2)	(3)
분노지수*아파트			-0.493*** (0.056)
분노지수*종교/이념			0.384*** (0.042)
분노지수*검찰수사			0.225*** (0.042)
분노지수*갑질			-0.067 (0.053)
분노지수*환경/에너지			-0.345*** (0.042)
분노지수*이웃분쟁/동물			-0.058 (0.045)
분노지수*부동산			-0.347*** (0.076)
분노지수*정책의견			-0.114 (0.060)
분노지수*사법불신			0.407*** (0.044)
분노지수*노동			0.777*** (0.075)
분노지수*장애인복지			-0.212*** (0.048)
분노지수*국회비판			-0.169** (0.055)
분노지수*스포츠			-0.220*** (0.065)
분노지수*범죄			0.068 (0.057)
분노지수*음주운전			-0.935*** (0.050)
분노지수*난민			0.113 (0.107)
분노지수*저출산			0.043 (0.078)
분노지수*교육/입시			-0.807*** (0.074)
분노지수*인터넷/언론			-0.521*** (0.044)
분노지수*의료			0.002 (0.051)
분노지수*금융/가상화폐			0.060 (0.063)
Constant	5.568*** (0.034)	5.578*** (0.035)	5.624*** (0.036)
N	88,113	88,113	88,113
Log Likelihood	-529,318.400	-529,313.100	-528,271.300
theta	0.322*** (0.001)	0.322*** (0.001)	0.327*** (0.001)
AIC	1,058,695.000	1,058,688.000	1,056,657.000

*p < .05; **p < .01; ***p < .001

우선 각 게시물이 대표하는 의제(토픽)를 기준으로 볼 경우, 갑질, 범죄와 같은 의제는 감정관련 변수가 포함되지 않은 모형에서도 의제 자체로 동의수에 유의미한 영향을 미치지 못한다. 또한 사회적으로 매우 갈등적인 이슈로 인식되는 성별갈등, 종교/이념, 난민, 인터넷/언론과 같은 의제 역시 의제 자체로 동의수에 유의미한 양의 방향의 효과를 보여주지 못한다. 이에 비해 모든 변수와 상호작용항이 포함된 모형에서 의제 자체로 동의수에 양의 방향의 효과를 내는 의제는 대통령, 노동, 스포츠, 음주운전, 금융/가상화폐 관련 의제들이다. 해당 의제들은 단일한 경향으로 정리하기 어려운 다양한 성격을 지니고 있기 때문에, 의제 자체가 청원 게시판에서 수행하는 역할에 대해서는 뚜렷한 결론을 내리기 어렵다.

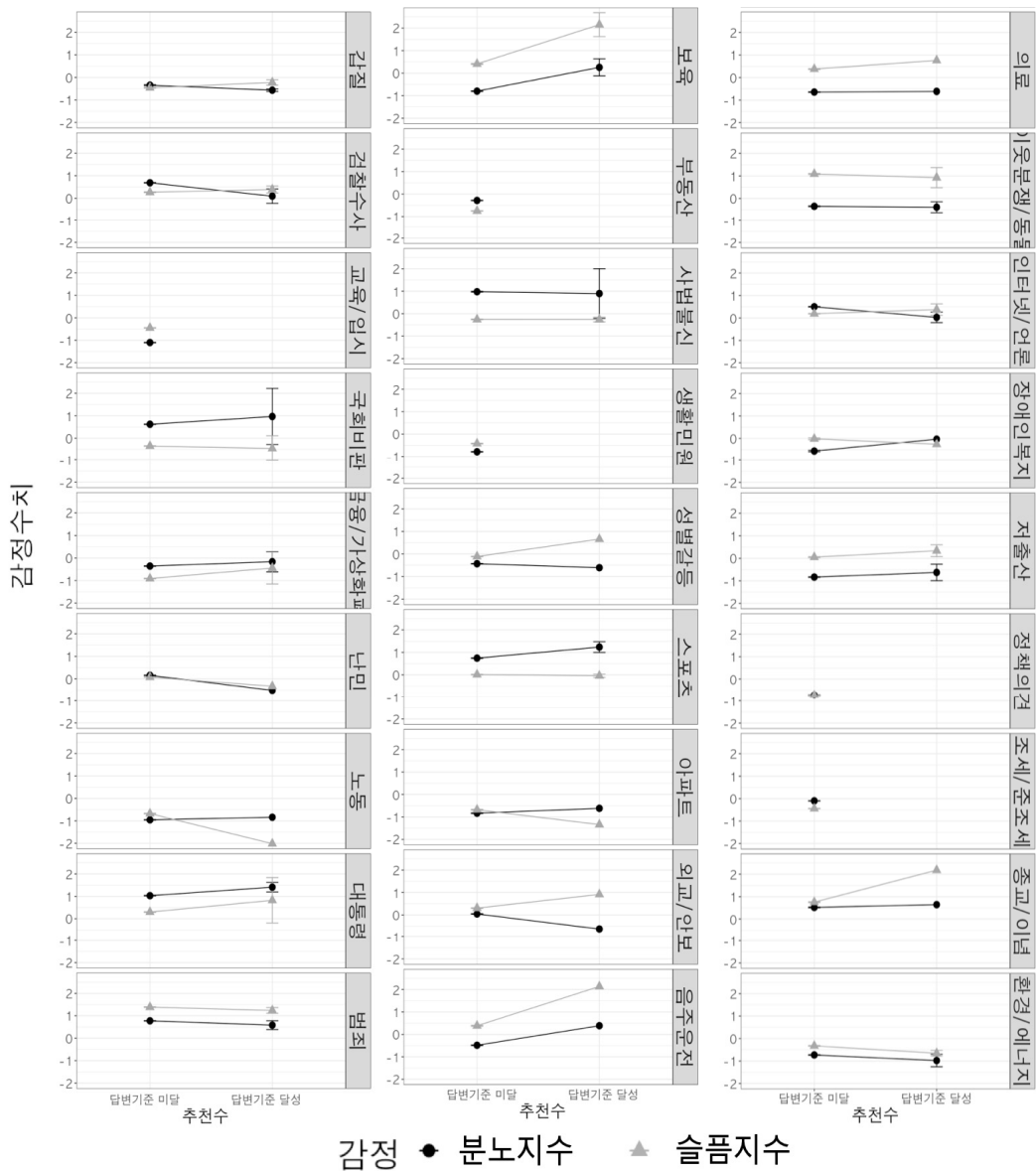
감정의 경우, 슬픔은 모형 (2)와 모형 (3) 모두에서 동의수에 유의미한 양의 방향의 영향력을 미치고 있으나 효과의 크기가 약 3~5% 정도로 비교적 작은 편이다. 또한 슬픔지수를 의제들과 상호작용항으로 투입할 경우 모형이 수렴하지 않았다. 즉, 슬픔은 개별 문서에서 대표하는 의제와는 상관없이 그 자체로 청원게시판에 접속하는 시민들의 공감을 유도하지만 그 효과는 약하며, 의제와의 상호작용도 뚜렷하게 나타나지 않고 매우 큰 분산을 지닌다고 볼 수 있다.

본 연구에서 가장 주목하는 감정인 분노의 경우는 상호작용항이 투입되었을 때 효과의 방향과 크기가 모두 대폭 변화한다. 이는 분노지수가 낮을 경우, 전반적으로 더 높은 조희수를 지니고 있으나 상호작용항에 의해 분노지수가 의제별 분포로 분해되면서 나타난 효과로 파악된다. 분노지수와 상호작용항이 양의 영향력을 지니는 의제는 생활민원, 종교/이념, 검찰수사, 사법불신, 노동이 해당되며, 성별갈등이나 국회비판, 인터넷/언론과 같은 의제는 예상과 달리 상호작용항이 음의 방향을 향한다. 또한 갑질, 범죄와 난민 의제는 분노와 상호작용이 유의미하게 관측되지 않는다. 이를 통해 분노가 청와대 청원 게시판을 완전히 지배하는 영향력을 행사하고 있지 않으며, 거시적인 사회적 갈등보다는 내밀하거나 생활과 밀착한 이슈와 결합하였을 때 시민들의 호응을 이끌어낸다는 점을 추정할 수 있다.

아래의 <그림 2>는 청와대의 공식적 답변이 제시된, 동의수 200,000 개 이상의 게시물과 그렇지 않은 게시물을 의제별로 분류하여 감정 분포를 나타낸 결과이다.¹¹ 전체 자료에서 동의수가 200,000 개 이상인 게시물은 총 53개에 불과하기 때문에 통계적 추정을 하기보다는 기술통계를 보여주고 해석하는 방식을 선택하였다. 보육과 의료, 저출산과 같이 돌봄과 관련된 의제에서 전답변기준과 관계없이 슬픔이 분노보다 더 높은 수치를 보이고 있다는 점은 예상 가능하나, 범죄, 성별갈등, 외교/안보, 음주운전, 종교/이념과 같이 시민들의 분노가 쉽게 발현되거나 공감하기 쉽다고 여겨지는 의제에서도 슬픔이 더 높은 비중을 차지하고 있다. 이를 통해 해당 의제들에서 가해자에 대한 분노보다는 피해자에 대한 연민이 더 많은 호응을 얻고 있다고 추정해볼 수 있다.

¹¹ 연구에 활용한 자료가 수집된 시점에서는 200,000명 이상의 동의를 받은 게시물에 대해 정부가 모두 브리핑을 실시하였기 때문에, 투입된 의제에 대한 정부의 반응 기준을 파악할 수는 없었다. 하지만 자료가 축적되고 200,000명 이상의 동의를 받은 게시물이 대폭 늘어난 현재 시점에서는 200,000명 달성뿐만 아니라 정부의 반응까지도 변수로 넣어서 분석을 시도해볼 수 있을 것이다.

| 그림 2 | 의제별 정부 응답에 따른 감정부포



이에 비해 노동, 아파트, 스포츠 의제에서는 답변기준을 달성한 글들 사이에서 분노지수가 높게 나타났다. 다만 이 경우에도 분노지수가 특별히 높은 글들이 답변기준을 충족 시켰다기보다는, 슬픔지수의 하락 효과로 인한 상대적 증가 효과에 가까우며 기저값 자체도 전반적으로 낮은 양상을 보인다. 예외적으로 스포츠 의제 정도가 상대적으로 높은 기저값과 상승폭을 보이고 있으나, 특기할 정도로 크지는 않다. 슬픔지수에 비해 분노지수는 답변기준 여부에 따른 증가폭이 그리 크지 않다. 이러한 결과는 슬

품의 효과가 분노에 비해 매우 작았던 앞서 <표 2>의 분석에서와 다소 차이를 보이는데, <그림 4>는 글 자체의 속성이 통제되지 않았으며, 무엇보다 동의수 200,000을 기준으로 분류를 하였기 때문에 답변기준 달성의 표본수가 매우 작아 적극적인 해석을 하기 위해서는 추가자료를 분석해볼 필요가 있다.

분석결과를 종합해보면 현재 청와대 청원게시판에서 게시물이 작성되는 양상과, 게시물에 반응하는 방식 사이에서 차이점을 확인할 수 있다. 우선 감정의 차원에서는 작성되는 글의 분노지수나 슬픔지수가 특별히 왼쪽으로 꼬리가 길지는 않으며 비교적 고르게 분포되어 있으나 게시물을 읽는 시민들은 감정에 반응을 한다. 이에 비해 의제 차원에서는 범죄, 부동산, 국회비판과 같은 직접적인 이권이나 사회적 갈등, 혹은 즉자적인 분노를 자극할 것으로 예상되는 글들이 다수 작성되지만 시민들의 반응은 이러한 의제를 분노와 쉽게 결합하지 않는 양상을 보인다. 또한 200,000 명이라는 대규모 참여가 이루어진 경우에도 동일 의제 내에서 급격한 감정의 변동은 관측하기 어렵다.

즉, 청와대 청원 게시판의 청원 전체에서 특별히 과잉된 분노가 집약되는 현상은 관찰하기 어려우며, 이용자들의 호응 역시 포퓰리즘에 동원되는 양태로 나타난다고 보기는 어렵다. 또한 분노는 청와대 게시판에서 시민들의 호응에 일정한 역할을 담당하고 있으나, 대규모의 파괴적 사회갈등을 유도하는 방향보다는 오히려 시민들의 일상적 문제에 대한 공분을 통해 공감을 유도하는 방향으로 해석할 수 있는 소지가 더 높다. 청와대 청원게시판에는 다수의 글들이 올라오며 이 중에서 일정 이상의 호응을 받기 위해서는 공감을 받을만한 내용을 담고 있어야 한다. 현재의 청와대 국민청원 운영방식 하에서는 실제 생활에서 공감이 되는 문제적 경험이나 사회에 대한 공통적 문제의식도 일정한 공감을 얻을 수는 있으나, 이는 의제 자체만의 힘으로는 어려우며 감정과 상호작용을 할 때 더 많은 참여를 이끌어 낼 수 있을 것으로 예상된다.

앞서 청와대 국민청원에 대한 부정론과 긍정론의 갈등을 종합하며, 긍정적 의제가 부정적 감정을 통해 촉발될 가능성을 제시하였으나 사회적 소수자와 관련된 의제들이 분노지수와 유의미한 상관관계를 보이지 않거나 음의 상관관계를 보였다. 다만 한편으로는 소수자에게 적대적이거나 포퓰리즘으로 분류되기 쉬운 의제들 역시 분노와 뚜렷한 선택적 친화성을 가지지는 않는 것으로 확인되었다. 즉, 현재로서는 분노는 특정 이슈들에 대한 시민들의 반응을 활성화하는 역할을 하고 있지만, 이러한 역할이 민주주의 전반이나 사회 구조와 연관되는 지점을 파악하기 위해서는 더 많은 자료 축적과 정밀한 연구가 필요하다고 판단된다.

정리하자면 청와대 청원게시판은 막연한 공격적 정서나 소수자에 대한 혐오가 지배하고 이러한 담론들만이 호응을 얻는 호전적이고 폐쇄한 공론장이라고 일방적으로 규정할 수는 없지만, 다른 한편으로는 시민적 숙의와 합리성이 발휘되고 소수자를 비롯한 신사회운동의 이슈들이 효과적으로 대표되는 공간이라고 볼 수도 없다. 청와대 청원게시판은 부정적 감정과 긍정적 의제들이 함께 상호작용하는 공간이다. 다시 말해 현실의 민주주의 제도가 작동하는 방식과 마찬가지로, 청와대 청원게시판은 서로 다른 정서를 담은 다양한 담론들이 경쟁하는 공간이며, 경쟁의 성패는 단순 직선보다는 2차 함수 곡선에 가까운 모습이다.

V. 결론

오늘날 제도적 차원의 민주주의가 참여의 위기를 겪는 와중에 인터넷 공간을 통한 정치 참여는 새로운 가능성을 열었다. 과거 산업사회의 토대 위에 세워진 대의민주주의가 정보통신사회라는 새로운 조건에 놓이게 되면서 나타난 시민들은 한편으로는 감정에 휩쓸리는 포퓰리즘의 자양분으로 여겨지기도 하나, 한편으로는 민주주의에 필수적인 정치 참여의 원동력을 갖춘 집단이기도 하다. 감정은 인터넷의 도래 이전에도 정치 참여를 추동하는 중요한 요인이었기 때문에, 무작정 위협으로 여기기보다는 어떤 조건에서 어떤 방식으로 정치에 영향을 미치는지를 살펴보아야 할 것이다. 청와대 국민청원도 정부의 정책결정과정을 전문가들에게만 맡기기보다는, 시민들의 생생한 목소리를 담기 위한 시도로서 시작되었다. 현재 국회가 제공하고 있는 청원 제도가 과도한 절차적 요건들로 인해 유명무실하게 운영되는 반면 청와대의 국민청원은 수많은 시민들의 관심과 참여를 이끌어내면서 나름대로 성공하는 것처럼 보인다. 그러나 다른 한편으로는 청와대 국민청원을 통해 오히려 공론장이 피폐해졌고 차질 대의기구를 우회하는 구조적 포퓰리즘의 요소를 가질 수도 있다는 지적도 제기되고 있다. 또한 정보격차의 문제와 책임성의 문제도 청와대 국민청원에 잠재된 문제로서 지목되고 있다.

본 연구는 청와대 국민청원에 대한 상반된 주장을 통합적으로 이해하고 경험적 근거를 제시하기 위해 진행되었다. 이를 경험적으로 살펴보기 위해 청와대 국민청원 페이지에 업로드된 게시물을 수집하여 자연어처리를 통해 청원문서에 나타난 의제와 감정을 분석하였다. 그 결과 청와대 국민청원에는 매우 다양한 정책적 의제들이 제시되고 있으며, 다른 한편으로는 분노와 슬픔과 같은 시민들의 부정적 감정이 표출되기도 한다는 점을 확인할 수 있었다. 그리고 이러한 의제와 감정이 청원 게시판에 접속하는 다른 시민들에게 얼마나 호응을 얻는지 탐색하였다. 분석결과 슬픔이 분노에 비해 더 일관된 효과를 보이고 있으나 효과의 크기는 작았으며, 분노는 감정 자체의 효과의 크기보다 다른 의제와 상호작용을 할 경우 효과의 크기가 더 컸다. 또한 분노와 상호작용 효과를 보이는 의제들은 심각한 사회적 갈등을 야기하는 의제들보다는 시민들이 일상에서 쉽게 겪거나 공감할 수 있는 비교적 생활밀착적인, 구조와는 거리가 있는 의제들이었다. 또한 정부의 브리핑을 통해 응답을 받은 게시물 역시 기술통계 수준에서 부정적 감정의 급격한 변화가 관측되지는 않았다.

따라서 청와대 국민청원에 대한 부정론과 긍정론은 모두 청와대 청원게시판의 일면을 반영하고 있다고 할 수 있다. 의제 공급의 측면에서는 주로 범죄나 정치, 부동산과 관련된 글이 다수 게시되지만, 부정적 감정의 수준은 그리 높지 않으며 게시판 이용자들의 호응 역시 공급되는 의제의 비중과는 다른 양상을 보인다. 그리고 청와대 청원게시판의 이러한 복합적 양상은, 부정론과 긍정론이 암묵적으로 전제하는 것처럼 청원게시판 자체가 고정적 효과를 가지지는 않음을 함축한다. 청원게시판에서 상이한 흐름이 동시에 관측된다는 점은, 바꿔 말하면 게시판 설계에 따라 특정한 효과가 강화되거나 약화될 수 있다는 것이다.

청와대 청원게시판이라는 경쟁적 공간에서 공론장을 성공적으로 구축하여 본래 개설의도인 정책과

정에 대한 시민의 참여를 촉진하기 위해서는, 구조를 개선할 필요가 있다. 즉 청와대 국민청원이 본래의 취지와 같이 정부가 시민들의 다양한 정책제안을 청취하고, 시민은 그들의 요구를 정부에 전달하며 정부로부터 응답을 받을 수 있는(그 응답이 긍정적이든 부정적이든) 공간으로서 활용되기 위해서는 응답기준의 대폭 하향을 포함한 전향적 조치가 요청된다. 사실 청와대 국민청원에 대한 국민들의 관심과 참여가 높은 것은 매우 고무적인 것이라 할 수 있다. 청와대는 과도한 동원과 의제 공급 차원에서 질적 저하를 막기 위해 최근 청와대 국민청원의 운영방식을 개편했다. 청와대는 국민청원의 부정적인 요소를 줄이기 위해 부적절한 청원의 노출을 줄이고, 부적절한 청원에 대한 응답을 하지 않는 방식으로 대처하고 있다.

이것은 어느 정도는 필요한 대책이기는 하지만, 청와대 국민청원의 문제는 공급 단계에서 부적절한 청원이 많다는 것만이 아니다. 오히려 앞선 분석결과에서 살펴보았다시피 공급이 많은 의제가 곧바로 호응으로 이어지지는 않으며, 시민들이 공분을 하는 주제는 사전적으로 예측하기 어렵다. 만약 현재 청와대 국민청원 게시판에서 시민들이 감정과 의제 사이의 상호작용과 선택을 적절한 방향으로 추구하고 있다면, 진정한 문제는 정부의 정책에 관련된 수많은 시민들의 값진 의견과 제안 중 상당수가 정부 기관으로부터 공식답변을 받는데 실패하고 있다는 것이다. 요컨대 현재의 청와대 국민청원은 운영차원에서는 오히려 참여라는 투입 측면의 문제보다, 정치적 효능감이라는 보상 차원의 문제에 대해 고민할 필요가 있다고 제안해볼 수 있다.

이러한 실용적 제안과는 별개로, 해당 연구에서 밝히고자 하였던 감정과 의제 사이의 상호작용 및 정치적 참여 사이의 관계는 후속 연구를 통해 더 명료한 형태로 규명될 필요가 있다. 무엇보다 해당 연구에서는 의제투입 과정에서 비교대상이 존재하지 않은데다가, 정부가 게시판에서 기준을 넘은 모든 의제에 응답을 하였기 때문에 정부의 선별이라는 요소에 대해 분석할 수가 없었다. ‘국회 토크’ 이나 과거 ‘다음 아고라’ 와 같은 의제 투입 통로에서도 감정과 호응도 사이의 관계가 동일하게 관측되는지, 국회에서의 의안 발제와 투입 단계에서의 차이점과 그 결과는 어떻게 다른지와 같은 다양한 질문들에 대한 검토가 요청된다. 궁극적으로는 감정과 정치적 참여, 의제설정 사이의 관계에 대한 심도있는 이론적 검토가 선행되어야 할 것이다.

종합하자면 청와대 청원게시판을 둘러싼 상반된 두 입장은 청와대 청원게시판을 고정된 실체로 파악하며, 감정과 민주적 참여를 대립항으로 설정하였기 때문에 통합될 수 없었다. 하지만 본 연구는 감정과 의제를 분리하고 상호작용을 검토함으로써 두 입장을 통합할 수 있는 길을 제시하였다. 그리고 이러한 논의가 더 다양한 관점과 정교한 분석 방식을 통해 진행될수록, 청와대 청원게시판 뿐만 아니라 디지털 기술이 민주주의와 가지는 관계에 대한 논의도 더욱 풍성해질 것이다.

참 고 문 헌

- 김범수 · 장우영. 2018. 온라인 기반 시민참여정치의 탐색적 연구: 14개 국내의 사례 비교 분석. 『동서연구』 30권 1호, 5-30.
- 김원정 · 정윤미. 2019. “지난 2년, 20만 명 이상 동의 얻은 국민청원 40%가 젠더 이슈, 그중 여성폭력 이슈 63%로 가장 많아.” 『KWDI Brief』 52호 한국여성정책연구원
- 김찬우. 2019. 국민 청원 데이터를 통해 본 주요 개혁 이슈. 『예술인문사회융합멀티미디어논문지』 9권 2호, 823~832
- 박상민, 나철원, 최민성, 이다희, & 온병원. 2018. Bi-LSTM 기반의 한국어 감성사전 구축 방안. 『지능정보연구』 24권 4호, 219-240.
- 박상훈. 2018. 『청와대 정부』 서울: 후마니타스.
- 이승원 · 임한샘 · 이현우. 2018. 청와대 국민청원 참여요인 분석: 포폴리즘인가, 정치효능감인가?. 『OUGHTOPIA』 33권 3호, 111-141.
- 임혁백. 2011. 신유목사회의 출현과 정치참여의 변화: 일반적 참여양태와 국제적 비교. 『평화연구』 19권 2호, 7-37.
- 정성전. 2018. 국민청원제도의 놀이적 속성과 정치 참여에 관한 고찰. 『문화콘텐츠연구』 12권, 131-152.
- 한겨레경제사회연구원. 2018. “국민청원, 폐법 아닌 성평등 · 인권 등 ‘사회권 보장’ 요구 높았다.” <http://heri.kr/964226>(검색일: 2019.09.05.)
- Bachrach, Peter. & Morton S. Baratz. 1962. “Two Faces of Power.” *American Political Science Review* 56: 947-952
- Bail, C. A., Brown, T. W., & Mann, M. 2017. Channeling hearts and minds: Advocacy organizations, cognitive-emotional currents, and public conversation. *American Sociological Review*, 82(6), 1188-1213.
- Bakiner, O. 2016. Negative Emotions and Transitional Justice. By Mihaela Mihai. New York: Columbia University Press, 2016. 240p. \$65.00 hardcover. *Perspectives on Politics*, 14(4), 1205-1207.
- Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. “Latent Dirichlet Allocation.” *Journal of Machine Learning Research* 3(Jan): 993-1022.
- Brader, T. 2005. Striking a responsive chord: How political ads motivate and persuade voters by appealing to emotions. *American Journal of Political Science*, 49(2), 388-405.
- Chowdhury, Gobinda. G. 2003. “Natural Language Processing.” *Annual Review of Information Science and Technology* 37(1): 51-89.
- Dahl, Robert. 1971. *Polyarchy: Participation and Opposition*. New Haven: Yale University Press.
- Dalton, Russell. 2000. “The Decline of Party Identifications.” In *Parties without Partisans: Political Change in Advanced Industrial Democracies*, edited by Russell Dalton and Martin Wattenberg. Oxford: Oxford University Press.
- _____. 2004. *Democratic Challenges, Democratic Choices: The Erosion of Political*

- Support in the Advanced Industrial Democracies*. New York: Oxford University Press.
- Egami, N., Fong, C. J., Grimmer, J., Roberts, M. E., & Stewart, B. M. 2018. How to make causal inferences using texts. *arXiv preprint arXiv:1802.02163*.
- Foa, Roberto Stefan and Yascha Mounk. 2016. "The Democratic Disconnect." *Journal of Democracy* 27(3): 5-17.
- _____. 2017. "The Signs of Deconsolidation." *Journal of Democracy* 28(1): 5-16.
- Fukuyama, Francis. 1992. *The End of History and the Last Man*. New York: Free Press.
- Goldberg, Y., & Levy, O. 2014. word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method. *arXiv preprint arXiv: 1402.3722*.
- Grimmer, J. 2010. A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in Senate press releases. *Political Analysis*, 18(1), 1-35.
- Groenendyk, E. 2011. Current emotion research in political science: How emotions help democracy overcome its collective action problem. *Emotion Review*, 3(4), 455-463.
- Huntington, S. P., & Nelson, J. M. 1976. *No easy choice: Political participation in developing countries*. Cambridge, Mass.: Harvard University Press.
- Kim, S. E. 2018. Media bias against foreign firms as a veiled trade barrier: Evidence from Chinese newspapers. *American Political Science Review*, 112(4), 954-970.
- Klingemann, Hans-Dieter. 1999. "Mapping Political Support in the 1990s: A Global Analysis. In *Critical Citizens*, edited by Pippa Norris. New York: Oxford University Press.
- Lee, Han Soo. 2014. "Analyzing the Multidirectional Relationships between the President, News Media, and the Public: Who Affects Whom?" *Political Communication* 31(2): 259-281.
- Lijphart, Arend. 1997. "Unequal Participation: Democracy's Unresolved Dilemma." *American Political Science Review* 91(1): 1-14.
- Livingstone, Sonia and Ellen Helsper. 2007. "Gradations in Digital Inclusion: Children, Young People and the Digital Divide." *New Media&Society* 9(4): 671-696.
- Lucas, C., Nielsen, R. A., Roberts, M. E., Stewart, B. M., Storer, A., & Tingley, D. 2015. Computer-assisted text analysis for comparative politics. *Political Analysis*, 23(2), 254-277.
- Mihai, M. 2016. *Negative emotions and transitional justice*. Columbia University Press.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Norris, Pippa. 2001. *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge: Cambridge University Press.
- _____. 2011. *Democratic Deficit: Critical Citizens Revisited*. New York: Cambridge University Press.
- O'Donnell, Guillermo. A. 1998. "Horizontal Accountability in New Democracies." *Journal of Democracy* 9(3): 112-126.
- Quinn, K. M., Monroe, B. L., Colaresi, M., Crespín, M. H., & Radev, D. R. 2010. How to analyze political attention with minimal assumptions and costs. *American Journal of Political*

Science, 54(1), 209-228.

- Roberts, Molly, Brandon Stewart, Dustin Tingley, and Edoardo Airoldi. 2013. "The structural topic model and applied social science." Paper presented at the Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation. December.
- Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2014. "stm: R Package for Structural Topic Models." *Journal of Statistical Software* 10(2): 1-40.
- Rong, Xin. 2014. "word2vec Parameter Learning Explained. *arXiv preprint arXiv:1411.2738*.
- Schattschneider, Eric. Elmer. 1960. *The Semisovereign People: A Realist's View of Democracy in America*. New York: Holt, Rinehart and Winston.
- Schwemmer, C., & Jungkunz, S. 2019. Whose ideas are worth spreading? The representation of women and ethnic groups in TED talks. *Political Research Exchange*, 1(1), 1646102.
- Valentino, N. A., Brader, T., Groenendyk, E. W., Gregorowicz, K., & Hutchings, V. L. 2011. Election night's alright for fighting: The role of emotions in political participation. *The Journal of Politics*, 73(1), 156-170.
- van Dijk, Jan. 2003. "A Framework for Digital Divide Research." *Electronic Journal of Communication* 12(1): 1-7.
- _____. 2005. *The Deepening Divide: Inequality in the Informational Society*. Thousand Oaks: Sage
- van Dijk, J. and Kenneth Hacker. 2003) "The Digital Divide as a Complex and Dynamic Phenomenon." *The Information Society* 19(4): 315-326.
- Weber, C. 2013. Emotions, campaigns, and political participation. *Political Research Quarterly*, 66(2), 414-428.

What happens in the Blue House Online Petition? : An Analysis of the Blue House Online Petition Based on Natural Language Processing*

Junmo Song (Yonsei University), Youngdeuk Park (POSTECH)

Abstract

Is Blue House Online Petition a mere channel of raw anger or a field of direct democracy? This study aims to understand empirically the impact of emotions on political participation and opinion. We collected all the posts accumulated for more than a year from the opening of the Blue House Online Petition page (total 88,113 cases after de-duplication). Then, through natural language processing (structural topical models and word2vec), the agenda and emotions expressed in the petition page were identified. And, through the negative-binomial regression analysis, we find out the relation between components of document and number of approve on that document. The results are as follows. First, a large number of written agendas do not get the majority's consent automatically. Second, anger has a greater effect on consent than sadness. Third, anger gains greater favour when combined with agenda related to everyday life than with serious conflict agenda. This study shows that Blue House Online Petition is a space that incubate various possibilities. And through an empirical approach, we hope to resolve the fictional confrontation between emotion and reason, and to explore the conditions under which various channels of political participation can strengthen democracy.

Key words: Blue House Online Petition, e-petition, civic participation, agenda-setting, emotion, natural language processing

* This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2016S1A3A2925033).