

석사학위논문
Master's Thesis

PolicyScape: 정부 정책에 대한 이해 관계자의 관점
수집 시스템

PolicyScape: Crowdsourcing Perspectives on Public Policies from
Stakeholders

2019

김현우 (金玄宇 Kim, Hyunwoo)

한국과학기술원

Korea Advanced Institute of Science and Technology

석사학위논문

PolicyScape: 정부 정책에 대한 이해 관계자의 관점
수집 시스템

2019

김현우

한국과학기술원

전산학부

PolicyScape: 정부 정책에 대한 이해 관계자의 관점 수집 시스템

김 현 우

위 논문은 한국과학기술원 석사학위논문으로
학위논문 심사위원회의 심사를 통과하였음

2019년 6월 19일

심사위원장 김 주 호 (인)

심 사 위 원 오 혜 연 (인)

심 사 위 원 차 미 영 (인)

PolicyScape: Crowdsourcing Perspectives on Public Policies from Stakeholders

Hyunwoo Kim

Advisor: Juho Kim

A dissertation submitted to the faculty of
Korea Advanced Institute of Science and Technology in
partial fulfillment of the requirements for the degree of
Master of Science in Computer Science

Daejeon, Korea
June 19, 2019

Approved by

Juho Kim
Professor of School of Computing

The study was conducted in accordance with Code of Research Ethics¹.

¹ Declaration of Ethical Conduct in Research: I, as a graduate student of Korea Advanced Institute of Science and Technology, hereby declare that I have not committed any act that may damage the credibility of my research. This includes, but is not limited to, falsification, thesis written by someone else, distortion of research findings, and plagiarism. I confirm that my thesis contains honest conclusions based on my own careful research under the guidance of my advisor.

MCS
20174255

김현우. PolicyScape: 정부 정책에 대한 이해 관계자의 관점 수집 시스템.
전산학부 . 2019년. 33+iv 쪽. 지도교수: 김주호. (영문 논문)

Hyunwoo Kim. PolicyScape: Crowdsourcing Perspectives on Public Policies
from Stakeholders. School of Computing . 2019. 33+iv pages. Advisor:
Juho Kim. (Text in English)

초 록

공공 정책은 다양한 사회 집단에게 각각 다른 방식으로 영향을 준다. 따라서, 정책에 대한 의견 형성을 위해서 정책에 대한 각 집단의 관점을 이해하는 것이 중요하다. 이 연구에서는 정책의 이해 관계자라는 개념을 이용해 정책에 대한 관점을 공유하고, 다른 관점을 탐색하는 아이디어를 제시한다. 예비 실험을 통해, 정책의 이해 관계자라는 정체성을 질문하면 시민들이 정책에 대해 더욱 설득력 있는 의견을 제시할 수 있다는 것을 확인하였다. 정책의 이해 관계자라는 정체성에 기반하여 정책에 대한 관점을 공유하고, 다른 시민의 관점을 탐색할 수 있는 온라인 플랫폼 PolicyScape를 개발하였다. 사용자는 정책에 대한 자신의 관점과, 정책의 이해 관계자라는 정체성을 태그 형태로 공유한다. 또한, 이해 관계자 태그를 이용하여 정책에 대한 다른 사람들의 관점을 탐색할 수 있다. 151명이 참여한 온라인 실험을 통해, 참가자들이 정책에 대한 다양하고 유효한 관점을 공유할 수 있다는 것을 보였다. 또한, 실험 참가자들의 의견을 통해 플랫폼이 수집한 이해 관계자의 관점이 정책에 대한 의견 형성에 도움이 될 수 있음을 보였다.

핵심 낱말 시민 참여, 정책, 이해 관계자, 온라인 토론, 클라우드소싱, 시민 기술

Abstract

Public policy can affect a diverse range of social groups in different ways. Therefore, understanding their perspectives is important for the citizens to form opinions on the policy. This work presents the idea of sharing and exploring citizens' perspectives on public policy using the concept of policy stakeholders. By doing so, I aim to support citizens to consider how the policy affects themselves and the other social groups. Findings from a pilot study showed that asking citizens about their identities as stakeholders led them to share opinions more persuasive to others. I implemented the idea as PolicyScape, an online platform for sharing and exploring citizens' perspectives on public policies with citizens' identities as policy stakeholders. The user of PolicyScape can share their perspective on the policy along with their stakeholder identities as tags. The user can also browse the perspectives of others with tags describing stakeholder identities. An online experiment with 151 participants showed that the participants can share diverse and valid perspectives on public policies with PolicyScape. Comments from the participants showed that the collected stakeholder perspectives could help participants forming their opinions on the policies.

Keywords Civic engagement, public policy, stakeholder, online discussion, crowdsourcing, civic tech

Contents

Contents	i
List of Tables	iii
List of Figures	iv
Chapter 1. Introduction	1
Chapter 2. Related Work	3
2.1 Leveraging Citizens' Knowledge on Public Policies	3
2.2 Enhancing Listening in Online Discussion	3
2.3 Relating Policy and Social Groups	4
2.3.1 Effect of Social Group on Political Thinking	4
2.3.2 Policy Stakeholders	4
Chapter 3. Formative Study	5
3.1 Participants	5
3.2 Conditions	5
3.3 Measures	5
3.4 Results	6
3.4.1 Conclusion	8
Chapter 4. System	9
4.1 Design Considerations	9
4.1.1 Identifying Stakeholder Groups	9
4.1.2 Questions to Ask to Stakeholders	9
4.1.3 Presentation of Stakeholder Groups	10
4.1.4 Exposing Diverse Perspectives	10
4.1.5 Interacting with Others' Perspectives	10
4.2 User flow of PolicyScape	10
4.2.1 Step 1: Describing Stakeholder Identity and Effects on Themselves	10
4.2.2 Step 2: Taking Others' Perspectives	11
4.2.3 Step 3: Exploring Effects	11
Chapter 5. Evaluation	12
5.1 Participants	12

5.2	Conditions	12
5.2.1	Choice of baseline condition	12
5.2.2	Populating initial data	13
5.3	Procedure and Tasks	13
5.4	Measures for Analyzing Shared Perspectives	14
5.4.1	Diversity	14
5.4.2	Validity	14
Chapter 6.	Results	16
6.1	Can the users of PolicyScape share diverse and valid perspectives of stakeholders on a public policy?	16
6.1.1	Diversity of Stakeholder Descriptions and Effects	16
6.1.2	Validity of Stakeholder Descriptions and Tags	17
6.2	How would users think of the crowdsourced perspectives?	18
6.2.1	Level of Understanding	18
6.2.2	Stance and Confidence	19
6.2.3	Value of Crowdsourced Perspectives and the News Article	20
Chapter 7.	Discussion	23
7.1	Effect of PolicyScape on Users' Opinion	23
7.2	Improving Opinion Sharing in PolicyScape	23
7.3	Possible Improvements for User Experience	24
7.3.1	Design Iteration	24
7.4	Generalizing to Other Policies	26
7.5	Limitation	26
Chapter 8.	Conclusion	28
	Bibliography	29
	Acknowledgments	32
	Curriculum Vitae in Korean	33

List of Tables

3.1	The number of comments containing features related to the value of the comment.	6
6.1	The most and least referenced tags for each policies.	16
6.2	The number of new stakeholder descriptions and the effects of the policy, compared to news articles	17
6.3	Number of participants who changed their stance on each policy in each condition.	19

List of Figures

3.1	The four experimental conditions for the formative study.	6
4.1	The main interactions of PolicyScape.	9
5.1	Summary of the three experimental conditions and the task flow for each condition. . . .	13
6.1	Self-reported level of understanding of each policy.	18
6.2	Self-reported level of understanding on how each policy affects the participants themselves.	18
6.3	Self-reported level of understanding on how each policy affects other stakeholder groups. .	19
6.4	Self-reported level of confidence of their stance for each policy.	20
6.5	Pre and post confidence level for high and low confidence groups.	20
7.1	New version of user interface for understanding the policy.	25
7.2	Revised user interface for browsing the perspectives of the stakeholders after iteration. . .	26

Chapter 1. Introduction

Public policies can have significant influence on citizens' everyday lives, ranging from an increase of income tax rate to health insurance reforms to stronger gas emission regulations to revised schedules for garbage collection. Therefore, citizens are motivated to understand the effects of different policies. However, they often evaluate policies from their own interests and perspectives only, while a policy involves a large number of stakeholders. For example, a health insurance reform affects not only patients and hospitals, but also private insurance companies, employers, and taxpayers [1]. This suggests that exploring and considering wider perspectives from diverse stakeholders enables individual citizens to understand the effect of the policy and have informed opinions on it.

A lot of existing platforms support users to exchange their opinions on the public policy. For example, ConsiderIt [2] supports users to exchange their opinions on a policy issue by organizing the argument with pros/cons list. Opinion Space [3] supports the users to discover diverse opinions by organizing them in terms of political stances of the author. IBIS-based argument mapping approaches such as Cohere [4] and Deliberatorium [5] organizes participants' arguments by logical relationship between them. However, while they support citizens to share and discover opinions, those systems rarely focus on who are related to the public policy. By doing so, users of those systems miss the opportunity of explicitly considering the identities of the people affected by the policy.

As an alternative, I propose enabling people to share their opinions on the policy and explore others' opinions with their identities as policy stakeholders. In this work, I use the term *policy stakeholder* as any kind of group or individual that can affect or can be affected by the policy, rather than focusing on organized groups with shared interest [6]. By sharing opinions with identities as stakeholders, I aim to nudge people to consider their personal relevance to the policy, which could lead people to think deeply on the issue [7, 8]. I also aim to capture the personal experience or knowledge on the policy issue by involving stakeholder identity. By presenting the stakeholders of the policy, I aim to support people to understand who are affected by the policy and how they are affected by the policy, which serves as a guideline for political thinking [9].

To observe whether explicitly considering stakeholder groups could be beneficial, I conducted a formative study on how asking the identity as policy stakeholder affects people's opinion on policy issues. I recruited 48 participants to share their opinions on two public policies with and without mentioning their identities as stakeholders. Analysis of collected response showed that asking stakeholder identities could lead to opinions mentioning more social groups. Comments of three external evaluators showed that opinions mentioning social groups were more valuable, by being more persuasive and nudging the readers to consider wider range of social groups relevant to the policy.

With evidence showing that asking stakeholder identity led to more valuable opinions, I built PolicyScape, an online platform for citizens to share their perspectives on public policies with stakeholder identities [10]. In the system, the user shares their perspective on a policy issue by describing how they are affected by the policy, along with tags that describe the identity of the user as a policy stakeholder. With the shared perspectives, the user can consider perspectives of a specific stakeholder group deeper by explicitly guessing their perspectives and verifying the guess in the system. The user can also explore the perspectives of diverse stakeholders collected in the system, using the stakeholder tags.

With the first version of the system, I conducted an online experiment a) to evaluate whether the

design of PolicyScape can support users to share diverse and valid perspectives of policy stakeholders and b) to observe how the users of PolicyScape think of the stakeholder perspectives in the system. The experiment involved 151 participants, and they were randomly assigned to three experimental conditions: Baseline (reading only policy-related news articles), Readonly (Baseline + exploring stakeholder perspectives with PolicyScape), and PolicyScape (Readonly + sharing their own perspective and guessing others' perspectives). Qualitative analysis and expert evaluation on the collected perspectives show that PolicyScape can collect novel and valid perspectives that could complement the existing information channel. While the understanding of policy effects were significantly increased in each condition, there was no significant difference across the conditions. The comments of the participants show that the participants regarded the collected perspectives as diverse, novel, and realistic, making them valuable source for understanding diverse stakeholder perspectives around public policies.

The contributions of the paper are as follows:

- Design considerations for making stakeholder-based opinion sharing system on public policy
- PolicyScape, an online platform for sharing and exploring stakeholder opinions on policy issues
- Experimental findings showing that the users of PolicyScape can share diverse and valid stakeholder perspectives, and collected data can help users to understand diverse stakeholder perspectives

Chapter 2. Related Work

In this chapter, I discuss previous work on leveraging citizens' knowledge on public policies, facilitating online discussion, and introducing the concept of social group to public policy issues.

2.1 Leveraging Citizens' Knowledge on Public Policies

Crowdsourced policymaking has been investigated as a way for engaging citizens in policymaking process and collecting relevant knowledge for improving the policy [11]. Crowdsourcing could be applied in every stage of policymaking process, from agenda setting to evaluation of the policy [12]. In Finland, Aitamurto and Landemore [13] conducted an experiment on crowdsourcing law making on off-road traffic. In a open forum, citizens were invited to freely suggest ideas and exchange arguments on the reform of off-road traffic law. Analysis of the process showed that the crowdsourcing experiment collected experience-based knowledge on the issue that could help policymakers understand the issue better despite lack of demographic representativeness of the participants [11, 13]. Also, they observed that citizens were deliberating on the issue by exchanging arguments on the issue and the citizens were able to learn to understand others' opinion from the crowdsourcing process [14]. However, organizing and analyzing the large scale data from crowdsourcing remained as a challenge.

Regulation Room platform, developed by Cornell eRulemaking Initiative, was another example of using citizens' knowledge on policymaking process [15]. Epstein et al. [16] showed that the participants of the actual rulemaking in Regulation Room contributed narratives that contain experience-based knowledge on the policy issues, such as unintended consequences of the policy and disagreement within an interest group.

Outside the context of policymaking process, there were several interactive systems that helped citizens collect their knowledge on policy issues and build a common artifact. BudgetMap [17] shows an example of creating a collective understanding of government budgets by classifying the budget items with participation of citizens. CommunityCrit [18] demonstrates that the local community can improve urban planning of their own community by small contributions from each member.

Extending this line of research, I aim to build an interactive system that could support its users to contribute their knowledge in the form of stakeholder perspectives, and to explore others knowledge with stakeholder identities.

2.2 Enhancing Listening in Online Discussion

Listening to and respecting others' opinions is a key virtue of deliberation, which requires the participants to seriously consider each argument [19]. In an online discussion, however, self-selection and selective perception of information make such consideration difficult [20].

ConsiderIt [2] is an online platform to facilitate personal deliberation on policy issues. The system supports users to construct their own pros/cons list on issues based on the arguments from other users, and asks the users to craft their own positions on the issues. One of the main design point of the system was presenting only the pros/cons arguments, without any extra information related to the author. By

such design the system aimed to make the users judge the argument purely based on the argument itself, not based on the political stance of the user.

Opinion Space [3] helps users discover and respect diverse opinions on controversial issues by visualizing each opinion based on the similarity of political stances between the author of the opinion and the user. Their evaluation showed that such visualization helped its users to feel the opinion diversity and agree with comments from others.

Visualizing the relationships between user posts are used commonly to support online discussion. For example, Deliberatorium [5] supports large-scale deliberations by supporting discussion in hierarchical tree between arguments and refutations. Systems based on Issue Based Information System, such as gIBIS [21] and Cohere [4] supports collective sensemaking by presenting the arguments, supporting information, and their relationships in a map.

Extending this line of research, I explore whether stakeholder identities can be used to facilitate citizens to understand diverse perspectives regarding policy issues.

2.3 Relating Policy and Social Groups

In this section, I review previous work from social psychology and policy science on how the concept of social groups can be valuable for forming opinions on public policies.

2.3.1 Effect of Social Group on Political Thinking

Information on social groups related to the policy can be an effective cue for the citizens to form opinions on the issue. The effect of group cue is different if a person identifies themselves as a member of the group or not. If the person identifies themselves as a member of the group affected by the policy, they would feel naturally sympathetic to that group and have pro-group preference on the issue [9]. On the other hand, if the person recognizes that social groups that they do not belong to are related to the issue, the person's prior knowledge on the status of the social group and emotional affect to that group is used to judge whether that group is treated fairly [9]. This suggest that social groups relevant to a social issue could be an effective cue for the citizens to build their opinions on the issue.

2.3.2 Policy Stakeholders

Understanding who the stakeholders are and how they are affected by public policy is important in assessing the effect of the policy [22]. Stakeholder analysis, a systematic method for investigating stakeholders, can identify a wide range of stakeholders, including marginalized groups as well as the key players [23, 22]. Incorporating the perspectives of each stakeholder group, stakeholder analysis provides a comprehensive understanding of the relationship between stakeholder groups and their conflicts of interest [22]. Furthermore, some deliberative democracy researchers claim that involving the stakeholder representatives can be an efficient and publicly legible way to encompass social and political diversity on the issue [6].

In this work, we question if the concept of policy stakeholder can be applied for collecting citizens' opinion and serve as an effective resource for the public to get a comprehensive understanding of the effect of the policy.

Chapter 3. Formative Study

I investigated how people would perceive other citizens' opinions on the public policy, written with or without stakeholder identity.

3.1 Participants

I recruited 48 participants from an external commercial survey agency. The mean age of the participants was 38.75, and 50% of the participants were female. The agency controlled for the age group and gender of the participants to be equally distributed. Among the participants, 8 participants said that they were conservative, 19 participants said that they were progressive, and 9 participants said that they were moderate.

3.2 Conditions

I chose two government policies of South Korea: 1) a blind hiring policy that banned all public organizations from using job candidates' personal information such as their family background, demographic details, physical attributes, and academic background (e.g., school and grade information), and 2) a health insurance reform that gradually extends insurance coverage for treatments and medications. I chose those two policies as they were well-known to the public as election pledges of the current president of South Korea and these policies could affect a large number of people with different interests.

I also prepared two sets of questions: Opinion set and Effect set. Opinion set simply asked the participants to comment on the policy with questions like "What do you think about the policy?". After asking that question, I asked how much they are affected by the policy and their identities as stakeholders. On the other hand, Effect set targeted the participants to think about how the policy affects them. So, Effect set first asked the participants about how much they are affected by the policy. Then, participants were asked to describe their identities as stakeholders of the policy. After those questions, I asked them about the effect of the policy to themselves.

I asked the participants to comment on one policy using Opinion set and on the other one using Effect set. The ordering of the question set and the pairing between the question set and the policy were counterbalanced, leading to 4 different experiment condition overall. Figure 3.1 shows the experimental task for each condition in detail.

3.3 Measures

I recruited three raters to read the comments of the people on the public policies. The raters were chosen not to have any deep knowledge on the policy, so they do not have any bias on the policy. The raters read two sets of comments on the same policy, without knowing that the opinions were collected with Opinion question set or Effect question set. The raters were instructed to choose the valuable comments from each set, and to think aloud the reason for choosing a certain comment as valuable. We also asked them to compare the two sets of comments.

From the interviews with three raters, we extracted four features of the comments:

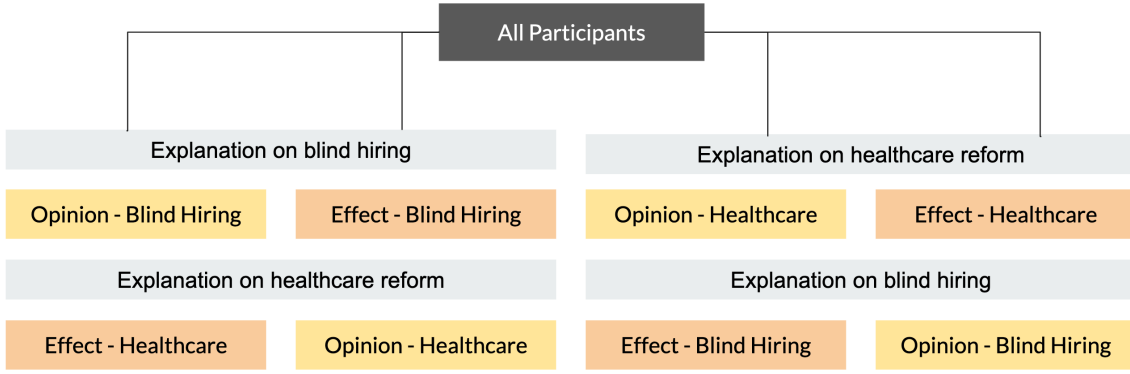


Figure 3.1: The four experimental conditions for the formative study. For each experimental condition, 12 participants were assigned.

- Comments mentioning the effect of the policy to the social group of the author of the opinion
- Comments mentioning the effect of the policy to any social groups
- Comments mentioning the effect of the policy to general values, such as fairness and equality
- Comments containing specific evidence, such as anecdote or statistics, to support the opinion

Finally, one another external rater and I counted the number of features for each set of comments to compare the two sets of the comment for each policy.

3.4 Results

After collecting the data, we removed the opinions of 13 participants who did not write their identities correctly, such as being an employee and a job seeker at the same time. After removal, 18 people answered Opinion set questions for healthcare reform and Effect set questions for blind hiring. 17 people answered Opinion set questions for blind hiring and Effect set for healthcare reform.

Table 3.1: The number of comments containing features related to the value of the comment.

Question set	Policy	Number of Participants	Own social group	Any social group	General value	Evidence
Opinion	Blind hiring	17	1	2	8	1
	Healthcare reform	18	3	6	3	3
Effect	Blind hiring	18	4	5	9	2
	Healthcare reform	17	6	8	1	0

Table 3.1 shows the difference of opinion features between Opinion and Effect set.

Comments generated from the Opinion set tend to refer to more general values, while comments from the Effect set tend to argue with the specific social group related to the policy more. For example, one comment from the Opinion set simply said, “*I’m supporting blind hiring policy because it could open fair opportunity for many people*”. However, a response from the Effect posed the same argument in a

very different way, saying that *“I’m a housewife, but with this policy I can trust my child would look for their job with their merits, not affected by any corruption.”*

Although the specific preference was different from raters to raters and policy to policy, the raters preferred the comments set with specific reasoning and information. The first rater preferred Opinion set for both policies, mentioning that the data from Opinion set tried to support their argument with reasoning but the data from Effect set were mostly about effect of the policy that anyone could come up with. She mentioned that she would prefer Effect set if it contained more specific effects related to specific stakeholder groups. On the other hand, the second rater preferred Effect set for both policies because she thought the comments in the Opinion set were mentioning generic effects, while Effect set were focusing more on the specific effects of the policy. The third rater preferred Effect set for blind hiring but Opinion set for healthcare reform, because he could discover more comments supported with the specific effects for social groups in each case.

The recurring theme from the raters was the value of mentioning some kind of specific social group related to the policy.

Making the comment persuasive

The raters mentioned that the link between the comment and the mentioned social groups would make the comment more persuasive. One rater picked an comment on blind hiring policy from Effect set as an example. The comment mentioned that the policy would give a wider range of options for the job-seekers. As it was written by a woman graduating from a local university, the rater thought that her comment would be more credible. However, the raters mentioned that they were hard to link the comments and the collected identities as stakeholders, as the identities were too broad in most of the cases, such as “A person holding national healthcare”.

Nudging for considering other social groups

The raters also agreed that social groups mentioned in the comments made them to question whether there could be any other social group affected by the policy. For example, when the first rater saw a comment on healthcare reform from Opinion set, saying that it would help low-income groups, she became curious about the other social groups, like the group who would pay more for the insurance.

Learning about reader’s social group

One rater specifically pointed out the benefit of comments mentioning her own social group, saying *“Although I am supportive for the policy, I could learn a possible negative effect of the policy with the comment opposing the policy because it was specifically mentioning people in their 20s or 30s, where I belong. Although there is another comment mentioning the same effect, I could agree with the comment more if it mentions my social group.”*

Understanding specific context of the policy

Raters liked the specific context around the policy. One of the raters’ favorite comments was about blind hiring from Effect set. That comment was special because it described current biases in recruiting in media industry. The second rater thought that comment was agreeable because it explained specific situation where the blind hiring could be effective. The third rater chose this comment as interesting,

saying that he never thought about the media industry so he could think about the issue from a new perspective.

3.4.1 Conclusion

Although it is hard to make conclusive arguments because of the small sample size, the findings of this formative study show that asking stakeholder identities leads to more comments mentioning social groups. The raters' comments show that comments mentioning social groups and specific evidence were valuable by making the comment more persuasive and nudging the readers to consider a wider range of stakeholder groups.

Chapter 4. System

In this section, I first list a set of design considerations for PolicyScape to capture and present the perspectives of the stakeholders, along with their identities. Then, I describe how the user interacts with the system to share their own perspectives and explore others' perspective in detail.

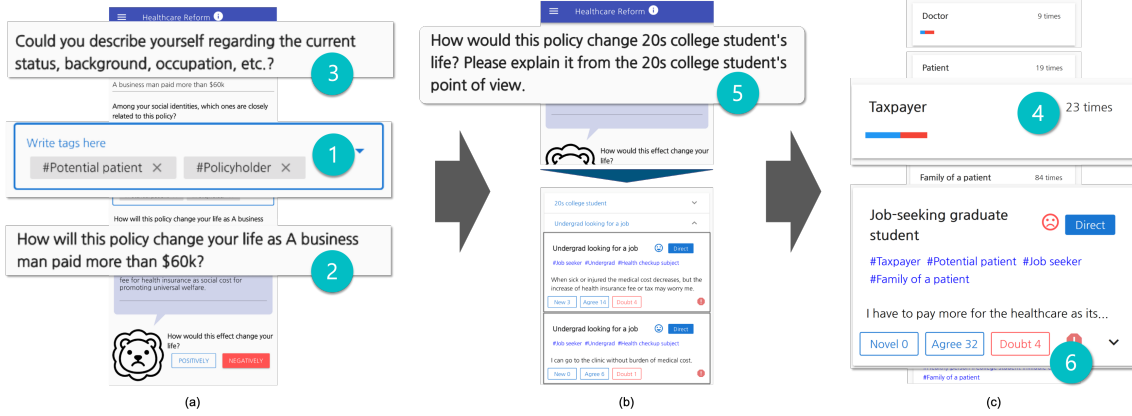


Figure 4.1: The main interactions of PolicyScape. (a) The user describes their own identity as a stakeholder with both free-form text and tag. Then, the user describes the effect of the policy to them. (b) The user is instructed to guess the perspective of the other stakeholders, randomly chosen by the system. After guessing, the user can verify their guess by seeing the stakeholder perspectives. (c) After describing their own identity as a stakeholder and guessing others' perspectives, the user can freely explore the crowdsourced perspectives in terms of stakeholders.

4.1 Design Considerations

4.1.1 Identifying Stakeholder Groups

Previous work on stakeholder analysis does not suggest a concrete way of identifying stakeholders. Instead, stakeholders are often identified iteratively, e.g., brainstorming by experts or snowball sampling [23]. Previous work warns about two pitfalls: omitting relevant stakeholders and identifying too many groups as stakeholders [23].

PolicyScape uses user-generated tags to identify stakeholder groups (Figure 4.1-1), so anyone may add relevant stakeholder groups. By using user-generated tags, I tried to accommodate diverse and unexpected stakeholder groups into the system. Still, the system should be able to prevent overflowing of only slightly relevant groups and keep the tags consistent throughout the data. As an approach, the system supports autocomplete for the tag input.

4.1.2 Questions to Ask to Stakeholders

While most of the previous work focused on collecting people's opinions on issues, PolicyScape asks users to report a policy's effect on them instead of their stance (Figure 4.1-2). By doing so, I aim to make users focus more on how the policy is related to them, rather than comparing the policy to general values.

I also tried to nudge people to share their experience-based knowledge from the stakeholders [11], such as unintended consequences of the policy or detailed context that might not be available for outsiders [24].

PolicyScape also asks about the user’s identity as a stakeholder of the policy, both in free-form text and tags (Figure 4.1-3). For instance, if a user is affected by the healthcare reform as her parents are suffering from a cardiovascular disease, she may describe her social status in an open-ended format, for example, “a taxpayer with parents suffering from a cardiovascular disease”, and identify her social group by tags, for example, “family of a patient”. By adding free-form description of stakeholder identity, I wanted to capture detailed context of each stakeholder that might affect their view on the policy, and also make people feel the collected perspectives more credible.

4.1.3 Presentation of Stakeholder Groups

The system supports the users to explore the users’ opinion with the identities of the stakeholder groups. (Figure 4.1-4) To prevent system overflowing with slightly related stakeholder groups mentioned by only few participants, the system only show tags mentioned by more than a certain number of people in the list of stakeholders, which is three in the initial version. Each tag item in the stakeholder list shows a bar graph with the ratio of positive to negative effects, which presents an opinion overview and a possible disagreement within each stakeholder group. By selecting each group, users can find individual stakeholders’ opinions with detailed descriptions of their social status.

4.1.4 Exposing Diverse Perspectives

Although the system tries to collect and present diverse stakeholder perspectives, it would be meaningless if the users are not exploring diverse perspectives. Inspired by perspective-taking [25] for moderating extreme opinions, PolicyScape provides an explicit interaction for users to think from perspectives of some other stakeholder groups, randomly chosen by the system (Figure 4.1-5). By perspective-taking, PolicyScape aims to help the users understand and respect other stakeholders. PolicyScape also lets users check the actual inputs from the stakeholders after guessing, so they could reflect on their initial guess.

4.1.5 Interacting with Others’ Perspectives

The system allowed the users to vote whether a certain opinion is novel, agreeable, or doubtful (Figure 4.1-6). The system aimed to use these votes as a measure for judging the value of the perspective. I excluded commenting feature from the system to prevent flaming between users and nudge the users to share their own perspectives to the system instead.

4.2 User flow of PolicyScape

Upon entering the system, the user chooses a policy of their interest. Then, the user goes through following steps:

4.2.1 Step 1: Describing Stakeholder Identity and Effects on Themselves

First, the user starts by stating how much a policy may affect their life. They do so by selecting one of the five possible values: “Don’t know”, “Not affected at all”, “Affected a little”, “Somewhat affected”,

and "Strongly affected". If they declare that they are not affected or do not know, they go straight to Step 2. Otherwise, they are prompted to explain how they are affected by the policy. (Figure 4.1(a)) The user is asked to enter the following information: (1) a description of the identity as a stakeholder as well as with tags, (2) a description of the effect, and (3) whether the effect is positive or negative.

4.2.2 Step 2: Taking Others' Perspectives

In this step (Figure 4.1(b)), the user is invited to think about the policy from the perspective of some other stakeholders and compare their guesses with the effects posted by actual stakeholders in the system. This step consists of two substeps: guessing and verification.

Guessing

The user speculates how a stakeholder will be affected by the policy. The stakeholder is randomly chosen by the system among those previously contributed by other users in Step 1 and were voted to the positive more than once in Step 3. The user is asked to answer with the format similar to Step 1, except that the stakeholder is selected by the system. The user takes as many guesses as they want and then proceeds to the verification sub-step.

Verification

The user is asked to compare their guessed effects with the effects collected from real stakeholders, which are displayed in the card layout. This design gives the user an opportunity to check their understanding, identify any gaps in understanding between their and others' thoughts, and modify their perspective if needed.

4.2.3 Step 3: Exploring Effects

Users are now able to explore all the effects the system contains (Figure 4.1(c)). Tags that have been used the least number of times are displayed at the top, so that users are likely to explore underrepresented or uncommon stakeholders rather than the obvious ones first. To ensure validity of the displayed tags, the system only displays tags that have been used three or more times by users.

Each tag display includes the tag label (i.e., social group), the number of times it has been used, and a bar graph showing the positive vs negative effect distribution. Users can also filter effects by one (e.g., "undergrad") or multiple tags (e.g., "undergrad" and "job seeker") , which allows seeing effects for a particular set of stakeholder groups.

When the user selects a tag from the list, a list of individual effects for the given tag is displayed as cards. The user can see details about an effect, such as the full description, the list of associated tags, whether the poster marked it as positive or negative, and whether the poster was a direct stakeholder or guessed the effect. The direct vs guessing label helps users make more informed assessment of the effect they are seeing. Furthermore, the user can respond to an effect by clicking on the "New", "Agree", and "Doubt" buttons. Effects are sorted in descending order by the total number of New and Agree reactions. The user can also report suspicious content. This reporting feature helps filter bad effect descriptions with incorrect information, harassment, or spam.

Chapter 5. Evaluation

To see whether PolicyScape can collect diverse perspectives from citizens and how sharing own perspective and exploring others' perspectives affect users' understanding of policy issues, I conducted an online experiment. By the experiment, I targeted to answer the following questions:

1. Can the current design of PolicyScape support citizens to share diverse and valid stakeholder perspectives on a public policy?
2. What would be the value of sharing and exploring stakeholder perspectives to the citizens?

5.1 Participants

I recruited participants by posting calls for participation in various online communities in South Korea, including university communities from some of the top-tier universities. The system was open for participation for 13 days. During that period, a total of 1,036 users visited the main page of the system. A total of 240 participants (75, 89, and 76 for baseline, Readonly, and PolicyScape conditions respectively) signed up and finished the preliminary survey. In the end, 151 of them (60, 46, and 45 respectively for each condition) finished the session. Dropout ratio was 20%, 48.3%, and 40.7%, respectively. Out of 151 participants, 82 were female, and the age ranged from 18 to 66 with an average of 26.78 ($SD = 7.86$). Participants received a voucher worth 7,500 KRW for their participation. A total of 29 participants completed the follow-up survey (16 for Readonly, 13 for PolicyScape conditions).

5.2 Conditions

The study used a between-subjects design, where each participant was randomly assigned to one of the three conditions: baseline, Readonly, and PolicyScape. In the baseline condition, participants read two articles on a given policy that are selected by researchers. In the Readonly condition, participants read the articles and additionally were able to browse the collected effects from PolicyScape freely. In the PolicyScape condition, the participants read the articles and were asked to describe their own stakeholder perspectives. Then, the participants guessed others' perspectives and explored others' perspectives.

In the experiment, I used the same set of national policies of South Korea: 1) a blind hiring policy that banned all public organizations from using job candidates' personal information such as their family background, demographic details, physical attributes, and academic background (e.g., school and grade information), and 2) a health insurance reform program that gradually extends insurance coverage for treatments and medications.

5.2.1 Choice of baseline condition

With the baseline condition, I targeted to simulate a common way citizens learn about policies. News articles are the primary source for citizens to learn about policies in their everyday life [26]. So, I decided that the baseline condition would give two news articles about a policy. Furthermore, this

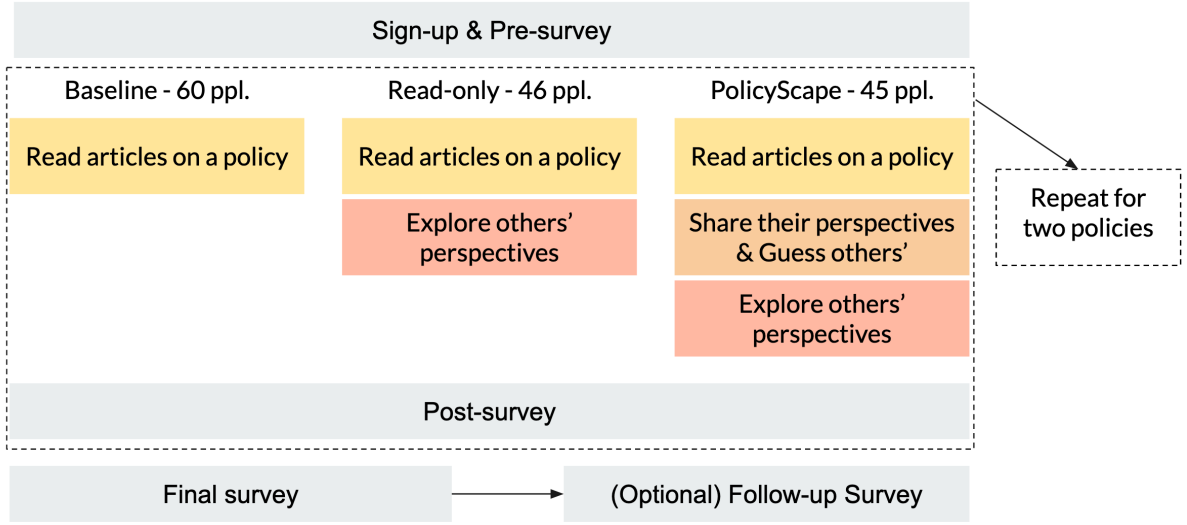


Figure 5.1: Summary of the three experimental conditions and the task flow for each condition.

design ensures that participants would get some knowledge about the policy before participating in the experimental tasks.

For each policy, I chose two articles written by the 30 most viewed online news media in South Korea [27]. I selected one article supporting the policy and the other with an opposing view, with both discussing perspectives of more than two stakeholder groups. This ensured that participants would get a balanced and broad understanding of the policy with articles.

5.2.2 Populating initial data

As PolicyScape relies solely on the participants' contribution, users participating in the initial stage might not get the benefit of collected perspectives. To ensure that all participants assigned to Readonly and PolicyScape can explore a number of perspectives and experience different from the baseline condition, I decided to pre-populate the system. Before starting the experiment, I recruited 8 paid participants (2 female, mean age 24.25) from an online community in a South Korean university to pre-populate the system. They used PolicyScape to submit at least three new perspectives for each policy. They generated 12 and 15 unique stakeholders and 25 and 30 effects for the blind hiring and the insurance reform policy, respectively. They received a voucher worth approximately 7,500 KRW for their participation.

5.3 Procedure and Tasks

Upon agreeing to participate in the online experiment, participants answered the preliminary survey, asking them to self-report their level of understanding of policy effects for the two policies used in the experiment. In the baseline condition, for each policy, participants read two news articles for at least 2 minutes (1 minute each). In the Readonly condition, participants read two news articles first. Then, they moved to the exploring step and looked through perspectives available in the system. They were asked to select and look at at least nine tags for each policy. In the PolicyScape condition, they were asked to generate at least three perspectives (including stating their own effect, if they were a stakeholder) after reading the articles. Then they explored the perspectives as in Readonly.

Right after the task, for each policy, participants answered a short survey about their self-reported understanding of the effects of the policy. Each participant conducted the same task for the two policies (in randomized order). After the main session, they answered the final survey on their experience on seeing the perspective of others via news article or PolicyScape, and sharing their own perspectives on the policies.

To gain a richer qualitative insight into the study experience, I conducted a follow-up survey. Participants in the Readonly and PolicyScape conditions who opted in for participating in an additional study were invited to optionally answer another survey. I asked about their opinions on the policies used in the system and some of the notable perspectives they saw from the experiment. For participants in the PolicyScape condition, I additionally asked about the perspective-taking experience.

5.4 Measures for Analyzing Shared Perspectives

I measured the quality of the shared perspectives with two criteria: 1) *Diversity*: how the perspectives capture stakeholders and effects not covered in the mainstream media and 2) *Validity*: how valid the perspectives are.

5.4.1 Diversity

To see how PolicyScape users can come up with perspectives that are not covered in the media, I asked two external raters to evaluate each of the stakeholder descriptions and effects collected from the study. For each policy, both external raters were instructed to read the two news articles covered in the experiment and find four additional online news articles that deliver diverse perspectives of the policy. After reading the articles (two given and four selected by themselves), they labeled each stakeholder description and effect into one of three categories: 1) covered, 2) partially covered, and 3) not covered in the news articles. For the effect items, they could also choose a fourth option for meaningless responses (e.g. “Will get a positive effect.”). Each item was labeled twice: first, it was labeled against two articles given in the experimental condition, and then it was labeled against six articles including the articles found by the raters. By first labeling, I aim to measure how many perspectives were from outside of the baseline condition. From the second labeling, I targeted to find the perspectives that were not covered by current media.

The raters labeled all of the stakeholder descriptions (59 for the blind hiring and 63 for insurance reform policy). For effect items, I asked them to label 100 effects for each policy which were randomly sampled from 270 and 242 total effects for the blind hiring and insurance reform policy, respectively. Out of 100 effects, 57 and 53 were written by stakeholders for the blind hiring and insurance reform policy, respectively. Two raters labeled the first 10 items of each list together to create a consistent standard for labeling. Then, they labeled the remaining items independently and discussed in person to settle the disagreements. Their inter-rater reliability (Cohen’s κ) was 0.82, 0.46 for the blind hiring policy (stakeholder descriptions and effects, respectively) and 0.45, 0.53 for the insurance reform policy. I attribute the relatively low IRR to the high subjectivity of the task.

5.4.2 Validity

Effects that contain incorrect or misleading statements can be critical to PolicyScape, as users can be affected by them. To determine the validity of the effects submitted by the participants, they were

evaluated by two experts (one for each policy). The effects for the blind hiring policy were evaluated by a government official at the Ministry of Employment and Labor and is in charge of this policy, while the effects for the insurance reform policy were evaluated by another government officer at the Ministry of Health and Welfare.

Due to limited expert resources, I only evaluated the effects marked as not covered or partially covered by the six news articles in the previous stage. They rated each item as ‘valid’ or ‘invalid’. Effects that contain inaccurate, incorrect, or misleading descriptions were labeled as ‘invalid’.

Chapter 6. Results

In this section, I report the main results from the experiment. The median time spent on the platform was 12 minutes for Baseline, 23.5 for Readonly, and 34.5 minutes for PolicyScape conditions.

Participants reported moderate background knowledge on each policy in the preliminary survey. The average self-reported knowledge for the blind hiring policy was 4.89 on a 7-point Likert scale, with an average score of 4.78 (SD=1.37) for the Baseline condition, 4.80 (SD=1.43) for Readonly, and 5.08 (SD=1.26) for PolicyScape. Participants reported less knowledge of the healthcare reform, with an average score of 4.04/7, including 3.81 (SD=1.85) for baseline condition, 4.35 (SD=1.86) for Readonly, and 3.96 (SD=1.52) for PolicyScape. A Wilcoxon rank sum test showed a significant difference of knowledge between both policies ($W=8619.5$, $p < 0.001$).

Although I tried to reach out citizens with diverse background, the age and occupation of participants who finished the final survey were quite biased. 55.0% (83/151) of participants said that they are student. 72.2% (109/151) were in their twenties and only 6.6% (10/151) said that they are over forty.

I report users' activity data from the system log. On the blind hiring policy, participants submitted 270 effects and 40 tags. They agreed ('agree' button in the system) with an average of 3.03 effects, identified 0.29 novel effects ('new' button) and expressed doubt on 0.47 effects. On the healthcare reform policy, participants posted 242 effects and 48 tags; they upvoted an average of 3.10 effects, found 0.37 new effects, and expressed doubt on 0.45 effects. Table 6.1 show tags that were most and least cited in effects made by stakeholders.

Table 6.1: The most and least referenced tags for each policies. The number in the parentheses indicate the number of stakeholders who used each tag.

	Blind hiring	Healthcare reform
Top 2	- Undergrad (69) - Well-educated (59)	- Potential patient (38) - Family of patient (27)
Bottom 2	- Artist (1) - Career break (1)	- Senior citizen (1) - Diabetic (1)

6.1 Can the users of PolicyScape share diverse and valid perspectives of stakeholders on a public policy?

6.1.1 Diversity of Stakeholder Descriptions and Effects

Two Article Condition On average, 26.2% of collected stakeholder descriptions were marked as not covered, and 27.9% as partially covered. Table 6.2 shows the repartition of descriptions and effect across policies. For effects, an average of 42% and 6% were marked as not covered and partially covered respectively.

Six Article Condition When compared to the contents of six news articles, the average proportion of not covered and partially covered stakeholder descriptions decreased to 14.7% and 19.7% respectively. For the effects, 29.5% and 1% were not covered and partially covered, respectively. Table 6.2 presents detailed numbers for 2 articles and 6 articles conditions.

Table 6.2: Number of not covered, partially covered, covered, and meaningless stakeholder descriptions and effects compared to the contents of 2 and 6 news articles.

	# Articles	Not covered	Partially covered	Covered	Meaning less	Total
Blind hiring						
Stakeholder descriptions	2	26 (44.1)	23 (39.0)	10 (16.9)	-	59
	6	13 (22.0)	14 (23.7)	32 (54.2)	-	59
Effects	2	34	5	51	10	100
	6	30	1	59	10	100
Healthcare reform						
Stakeholder descriptions	2	6 (9.5)	11 (17.5)	46 (73.0)	-	63
	6	5 (7.9)	10 (15.9)	48 (76.2)	-	63
Effects	2	47	7	36	10	100
	6	29	1	60	10	100

Below I present some examples of stakeholder descriptions and effects marked as not covered compared to the contents of six news articles.

- Stakeholder descriptions

- (Blind hiring) Mother of job seeker
- (Blind hiring) High school teacher
- (Healthcare reform) Medical device manufacturer
- (Healthcare reform) Nurse

- Effects

- (Blind hiring) Engineering grad student: “ [...] the job performance will be judged based on my publication and, therefore, the blind hiring is almost impossible.”
- (Healthcare reform) High-income earner: “ As a subscriber of private insurance, most medical expenses are already covered. For me, this policy is about being burdened with social cost to expand universal welfare.”

6.1.2 Validity of Stakeholder Descriptions and Tags

Experts evaluated a total of 61 effects and marked 58 (95.1%) as valid. For the blind hiring policy, 2 out of 31 were invalid, including one non-stakeholder’s guess. For the healthcare reform policy, 2 out of 30 were invalid and both were non-stakeholders’ guesses.

Effects labeled as invalid contained speculations based on a misunderstanding or limited knowledge of the policy. For example, one stakeholder said that the benefit of attending grad school disappears with the blind hiring policy. The expert marked this as invalid, saying that this might not be true as candidates' publications would still be considered in the recruiting process.

6.2 How would users think of the crowdsourced perspectives?

To understand the value of crowdsourced perspectives to the users, I analyzed the difference in participants' self-reported level of understanding and participants' stance on the policy. Also, I asked the participants on their experience on seeing the effects generated by stakeholders and the reasoning behind opinion change.

6.2.1 Level of Understanding

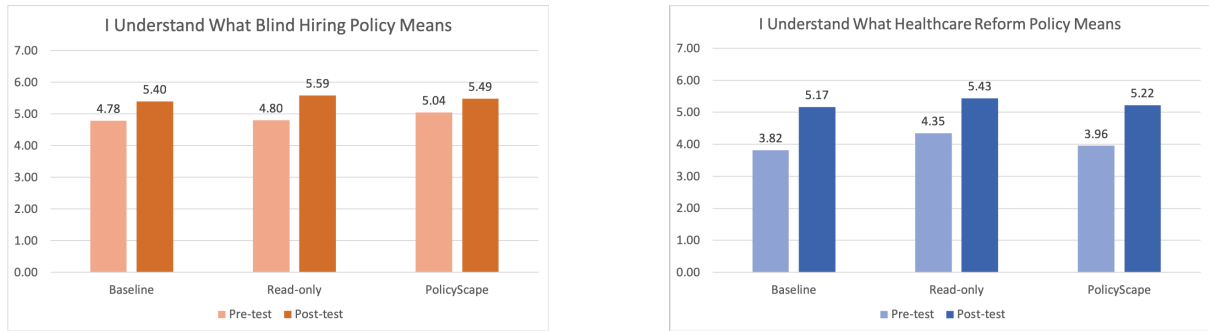


Figure 6.1: Self-reported level of understanding of each policy. While the score was improved significantly, there were no significant differences between the conditions.

Understanding scores significantly increased from 4.87 (SD=1.35) to 5.48 (SD=1.11) for the blind hiring policy ($W=14339$, $p < .01$) and from 4.01 (SD=1.76) to 5.26 (SD=1.06) for the healthcare reform policy ($W=16072$, $p < .01$). However, the difference between the conditions were not significant (Kruskal-Wallis, $p > .05$) (Figure 6.1).

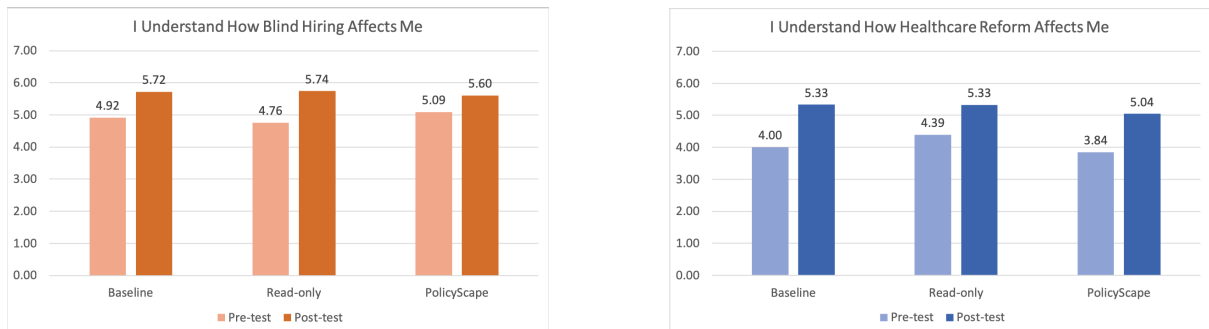


Figure 6.2: Self-reported level of understanding on how each policy affects the participants themselves. While the score was improved significantly, there were no significant differences between the conditions.

Score for understanding how the policy affects themselves significantly increased from 4.92 (SD=1.45) to 5.69 (SD=1.02) for the blind hiring policy ($W=3294$, $p < .0001$) and from 4.07 (SD=1.83) to 5.24

(SD=1.08) for the healthcare reform policy ($W=4397$, $p < .0001$). However, the difference between the conditions were not significant. (Figure 6.2).

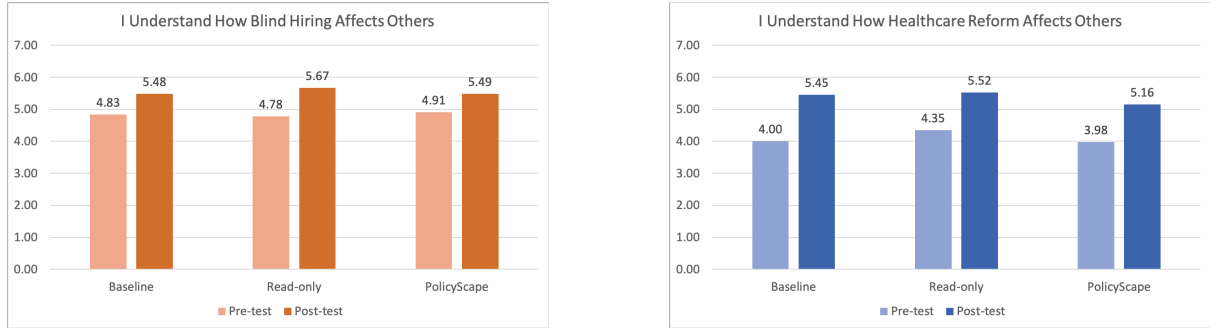


Figure 6.3: Self-reported level of understanding on how each policy affects the participants themselves. While the score was improved significantly, there were no significant differences between the conditions.

Score for understanding how the policy affects the other stakeholder groups also significantly increased from 4.84 (SD=1.37) to 5.54 (SD=1.13) for the blind hiring policy ($W=3016$, $p < .0001$) and from 4.10 (SD=1.78) to 4.87 (SD=1.08) for the healthcare reform policy ($W=5009$, $p < .0001$). In this case, the difference between the conditions were not significant either (Figure 6.3).

6.2.2 Stance and Confidence

38 out of 151 participants (12 in Baseline, 12 in Readonly, and 14 in PolicyScape conditions) changed their stance on at least one of the policies. While more participants with PolicyScape changed their mind, the difference was not significant (Chi-square test, $p=0.42$). Details on stance change are shown in Table 6.3.

Table 6.3: Number of participants who changed their stance on each policy in each condition.

	Baseline	Readonly	PolicyScape	Total
Participants	60	46	45	151
Blind hiring	6 (10.0%)	8 (17.4%)	5 (11.1%)	19 (12.6%)
Healthcare reform	7 (11.7%)	4 (8.7%)	10 (22.2%)	21 (13.9%)

The levels of confidence increased significantly from 4.77 (SD=1.42) to 5.15 (SD=1.44; $W=9505$, $p < .05$) for the blind hiring policy and significantly from 4.29 (SD=1.81) to 4.97 (SD=1.39; $W=9001$, $p < .01$) for the healthcare reform. However, there was no significant difference between the conditions ($p > .05$) (Figure 6.5).

While the confidence score tended to increased for participants who originally gave low or moderate confidence score ($<4/7$), the confidence score slightly decreased for those who gave high ($>5/7$) confidence score in each policy (see Figure 6.5). I could get some explanations for this in the follow-up survey. One participant said, “*I have been supporting the healthcare reform policy. However, by using PolicyScape I realized that I will become a taxpayer in the future and have to take the burden of the reform by myself. Now I’m neutral on the policy.*”

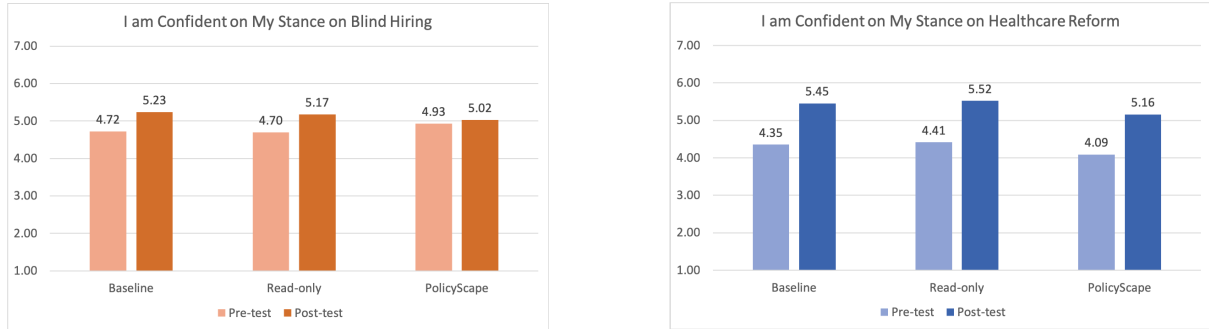


Figure 6.4: Self-reported level of confidence of their stance for each policy. While the score was improved significantly, there were no significant differences between the conditions.

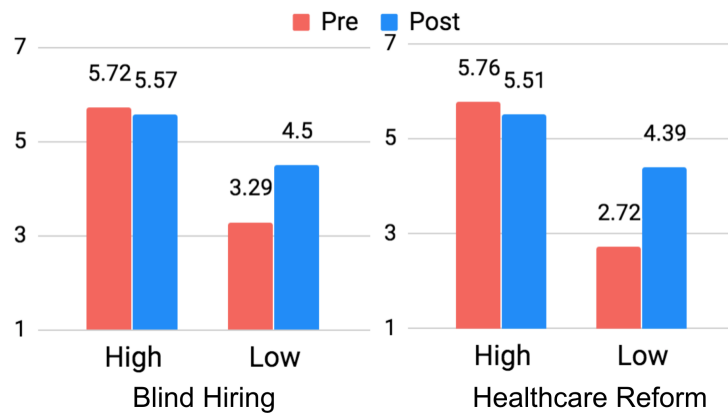


Figure 6.5: Pre and post confidence level for high and low confidence groups.

6.2.3 Value of Crowdsourced Perspectives and the News Article

To understand how participants perceived the stakeholders' perspective, I asked them to comment on the strengths and weaknesses of crowdsourced perspectives in the final survey. From 91 participants in Readonly and PolicyScape condition, I gathered 51 pro- and 25 cons- comments. For each group of comments, my collaborator and I separated comments consisting of more than one point into multiple comments so that each comment contains only one point. I also asked the critical factor of changing the opinion after the experiment in the follow-up survey.

Out of 51 comments mentioning the strength of the crowdsourced perspectives, 14 comments mentioned its diversity and 12 comments mentioned presenting effects they didn't know before. One participant specifically said, *"News articles tend to analyze the effect [of the policy] by categorizing [stakeholders], but with stakeholder identities I could see small opinions that would not be captured by such patterns."* 8 comments liked the crowdsourced perspectives being detailed and realistic. One participant from the follow-up survey specifically mentioned that she became less opposing the healthcare reform after listening to the voices of the families of the patients who would get benefit from the policy. On the other hand, participants mentioned that some of the crowdsourced perspectives were too narrow-minded and biased. Also, the participants had doubts on whether the crowdsourced perspectives were really written by the stakeholders.

Value of Stakeholder Perspectives for Opinion Formation

Comments from the follow-up survey shows how the collected stakeholder perspectives helped users to make their opinion in several ways. Out of 29 participants for the follow-up survey, 7 participants mentioned the effect of the articles for blind hiring and 7 participants mentioned it for healthcare reform.

Participants used the perspectives of other stakeholder groups to understand how the policy affects others, especially for the groups heavily affected by the policy. Interestingly, they were able to expect the opinions of the stakeholder groups, and they wanted to verify their guess with the stakeholder perspectives. One participant said, *“I was interested in the opinions of the companies [on blind hiring policy]. I think the policy would help recruiting better human resources, and I wanted to know how the companies think.”* Another participant mentioned, *“I was interested in the opinions of the doctors [on healthcare reform]. I think the doctors would get economic loss from the policy, and I wanted to know what the doctors would think about it.”*

If the participants could find about the heavily affected stakeholder groups from the collected perspectives, they could use it for judging the effectiveness of the policy. For example, one participant said, *“Initially, I thought blind hiring could be fair for everyone, but after the experiment I learned that the policy might not be fair to everyone so I slightly changed my mind.”* Another participant who was opposed to the policy due to more tax burden said, *“I am less opposed to healthcare reform after using the system. I read a perspective of a family of a patient getting benefit from the policy, and then I thought I could pay more tax if it is used for good.”*

Participants were also interested in the perspectives of their own groups. By so, they wanted to know what their peers are thinking. One participant said, *“I was interested in the perspectives of job seekers. I’m also a job seeker, so I wanted to know more about the specific cases.”*. Another participant said, *“I remember one comment from a healthcare subscriber supporting for the healthcare reform. As the author of the comment and I shared similar social status, I wanted to verify whether that opinion was majority or not.”* Not many of them explicitly answered how such experience affected their opinion. One participant said, *“As a college student, I only considered the positive impact of health insurance reform policy. However, by reading collected perspectives, I realized that I am the future taxpayer who is negatively affected by this reform. Now, I feel neutral on the policy.”* This could be an example of invoking in-group bias for political thinking [9].

Value of News Articles for Opinion Formation

Interestingly, some of the participants mentioned that the news articles were critical to make their opinion. Out of 29 participants for the follow-up survey, 7 participants mentioned the effect of the articles for blind hiring and 7 participants mentioned it for healthcare reform.

Some participants valued the specific context from the stakeholders for developing their opinions on the policy. Such cases were related to heavily affected but under-represented stakeholder groups such as recruiters for blind hiring policy and doctors for healthcare reform policy. For example, one participant said, *“I already knew gynecologists are under hard time. I read similar opinions from the doctors in the article, so I could agree with the doctors, and became less confident on supporting healthcare reform.”* Similarly, another participant said, *“In the article, recruiters said that the criteria for hiring would be ambiguous with blind hiring. Although I am still supportive on blind hiring, I think the policy could be improved more.”*

Another valuable aspect of articles were delivering complicated part of the issue from the experts’

view. One comment said, *“I changed my mind to support the policy (healthcare reform). I read the article by a doctor, criticizing that the patient have no choice but taking the doctors’ service in the current medical system. Then, I thought that the low-quality medical service in Korea was caused because the government were not able to control the medical expenses transparently.”* Another comment said, *“I was supportive for Healthcare Reform because it sounded nice, but I changed my mind because from the articles I learned about the problems related to the reimbursement rate and the side effects of the policy.”*

Chapter 7. Discussion

In this section, I further explain some of the main findings from the study and discuss limitations.

7.1 Effect of PolicyScape on Users’ Opinion

The experimental result did not show clear difference in understanding of policy or policy effects. One possible reason is that the used measures were not accurate. As the main measure was the self-reported score on understanding of policy effects, self-reporting bias would have affected the results. Also, it might be possible that the experimental condition of reading two news articles with different perspectives were too ideal as a baseline condition. This was implied from the comments of the users, mentioning that just reading news articles helped them widening their viewpoints on the policy issues. While providing such articles can give better user experience, the effects of the articles would be so large that it made the effect of the experimental condition marginal. To address the two factors, I plan to introduce the objective questions on the public policies to measure the knowledge difference accurately for the next iteration. Also, more realistic baseline condition, such as using only one neutral news article on the policy, will be used.

7.2 Improving Opinion Sharing in PolicyScape

The evaluation of PolicyScape showed that the system was possible to encompass identities of policy stakeholders and their valid perspectives on public policies, and the participants felt that the collected perspectives were diverse. Although the demographics of the participants were skewed towards young college students, this could be possible because only a couple of participants within the same stakeholder group is enough to contribute to the diversity of policy stakeholders. Still, wider range of participation could contribute to the diversity of opinion, especially for incorporating more opinions of heavily affected but under-represented groups, such as recruiters in case of blind hiring and doctors in case of healthcare reform. Such diversity would help the users understand their perspectives in a less biased way.

However, there are still rooms for improving opinion sharing interaction of PolicyScape.

First, the current design is not capturing the relationships between tags well, making it hard for the users to browse the collected data with stakeholder tags. One of the problems is not capturing semantically duplicate tags, such as “public official” and “government employee” appearing at the same time. Another problem is lack of hierarchical structure between tags. For instance, in ideal case every effect tagged with “doctors” should be also tagged with “medical staff”, but in the actual data that was not the case. Although such problems are quite common with user-generated tags [28] and approaches for detecting similar tags [29] exist, the size of collected data was too small to directly apply such techniques. Interaction design to nudge the users to consider hierarchical tag structures and similar tags, such as tag suggestions or more noticeable autocomplete, could be an alternative approach for the problem.

Second, the current design of asking the effect of the policy requires the participants to understand the policy and its relationship to themselves. So, it is possible that PolicyScape exclude the participation of stakeholders who are unaware of the policy or unaware of how the policy affects themselves. Such cases would be more likely if understanding the policy issue requires high expertise, such as nuclear

energy policy. As Regulation Room [15] suggests, providing objective and accessible information on the policy in the system could be a solution. Incorporating news articles on the policies as a part of the user experience can be a starting point for identifying more specific information needs on the policy.

7.3 Possible Improvements for User Experience

From the comments of the users in the post-survey and the follow-up survey, I learned that there is a large room for improving the user experience.

First, news articles could improve the user experience, by providing information on the policy issues and presenting alternative perspectives from the experts. This aspect was more clearly observed with healthcare reform policy, which requires understanding of the whole medical system to understand the doctors' perspectives. Including a news article within the system could be a better approach to supplement the collected perspectives and to provide basic information to understand the policy.

Second, the presentation of the stakeholder groups and their perspectives can be more interactive to encompass diverse needs of the users. The comments of the users showed that the users were interested in very different groups of people, and the perspective of each group can be valuable in different ways, from learning new information on the policy impact and stakeholder groups to confirming their prior opinions with supports from the stakeholder voices. By designing more interactive components, the system can guide the users gaining diverse insights from the collected data.

The system also had a lot of usability issues. In specific, the experimental platform was targeted for mobile phones to encompass a wider range of participation. However, the users suffered from reading long news articles, answering questions on stakeholder groups, and browsing a large number of collected perspectives from mobile phones. Furthermore, the plain list of stakeholder tags and the perspectives of stakeholders could not support users to browse a large set of collected perspectives easily. By supporting PC platform and introducing more lightweight interactions, I aim to reduce the burden of the users of the system.

7.3.1 Design Iteration

Based on these lessons, I present the design of PolicyScape after iteration. In this design, the system aim to explicitly guide the users to explore the crowdsourced data, so they could discover new and diverse opinions without burden of browsing tens of stakeholder groups and perspectives. After the user choose the policy of their interest, they go through the following steps in the system to browse the crowdsourced perspectives:

Step 1: Understanding Policy

First, the user sees a brief explanation of the policy and a link to a news article containing the voices of the stakeholders of the policy (Figure 7.1). With these materials, the system supports the user understanding the policy and the perspectives of stakeholders around the policy. The explanation was written by the authors based on the policy description from the government agency, and the article was also curated by the authors. After reading these materials, the user answers how the policy affects them as in the initial version of the system.

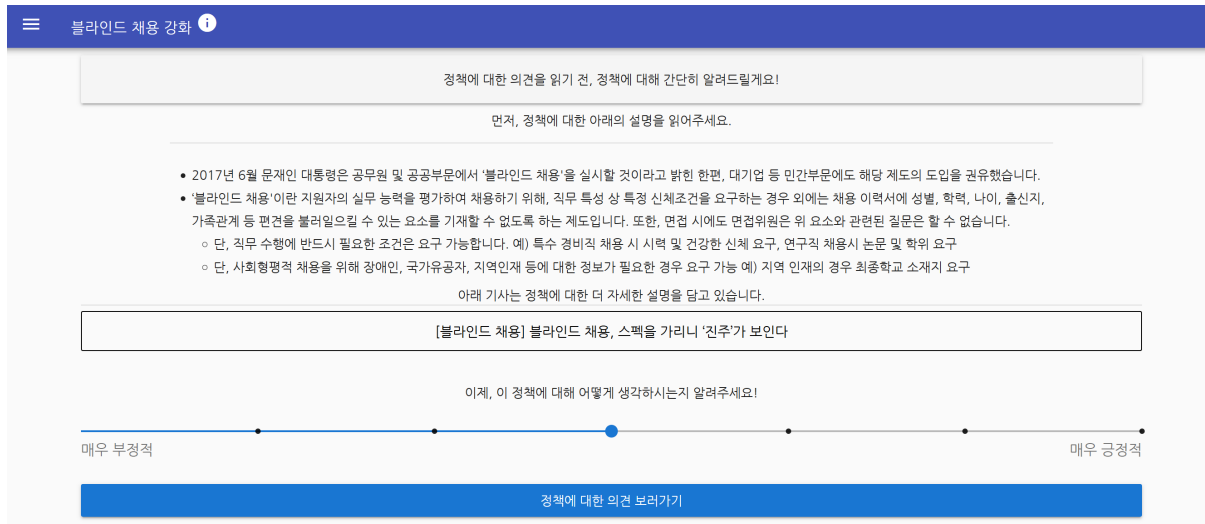


Figure 7.1: New version of user interface for understanding the policy. The system provides a brief explanation of the policy and a link to a news article explaining the policy.

Step 2: Exploring Stakeholder Groups

After reading the materials, the user is invited to explore the stakeholder perspectives present in the system. The system guides the user to discover the stakeholder group of interest.

First, the user asked to guess the possible stakeholder groups of the policy(Figure 7.2a). After entering more than one stakeholder group and click "Verify" button, the user can see a word cloud of stakeholder groups, providing an overview of stakeholders in the system at a glance(Figure 7.2b).

Then, the user can pick stakeholder groups of interest. The system provides some classes of stakeholder groups, such as "Groups related to me" and "Groups that I can predict their opinions" as hints for selecting stakeholder groups.

Step 3: Exploring Perspectives of Each Stakeholder Group

When the user selects an stakeholder group, the system guides the user to explore opinions of the stakeholder groups. Two types of interaction is presented to the user, determined by the number of the stakeholder perspectives in the system.

Summarizing Interaction If the selected tag contains a small number of stakeholder perspectives, the system targets to make the user read all of the perspectives. As a way of doing so, the user is requested to read the stakeholder perspectives thoroughly and summarize them. After summarizing, the user can choose other stakeholder group of interest(Figure 7.2c).

Exploring Interaction If the selected tag is related to a large number of stakeholder perspectives, the system aims to help the user understand the general opinions of the stakeholders first. As an approach, the system presents a word cloud of the stakeholder perspectives (Figure 7.2d). By hovering on each word, the user can understand in what context the word was used in the stakeholder perspectives. With the word cloud, the system asks the users to guess the main argument of the selected stakeholder. After guessing, the system invites the user to see the whole list of stakeholder perspectives.

the data. While the collected perspectives were diverse enough, skewed demographics might lead to over-representation or under-representation of some of the stakeholder groups, such as university students. Facilitating wider range of participation from more diverse population with more advertisement, or co-working with governments or non-governmental organizations for facilitating political participation could be a solution.

Second, the system did not have any systematic way for filtering out false identity or false information. The current version of system relies on naive approach of using users' votes, but robust way of quality control still remains as a challenge. One possible approach to filter out false identity could be presenting a history of identities for each user, so the other users could use any conflicts of identities to judge the credibility of the user.

Chapter 8. Conclusion

Public policies can affect a wide range of social groups, but understanding their perspectives on the public policies is still challenging. This work introduces the idea of using stakeholder identities as a way of sharing and browsing citizens' perspectives on public policies. I implemented the idea as PolicyScape, an online platform for sharing and exploring the perspectives of stakeholders on public policies. Findings from the evaluation showed that the users of the system were able to contribute diverse and valid perspectives on public policies. Although the effect of exploring collected perspectives were not observed in a statistically significant way, the comments from the users suggested that the collected perspectives could help users develop their own opinion on the policy. Ensuring data quality remains as a future work.

Bibliography

- [1] Barack Obama. United states health care reform: Progress to date and next steps. *JAMA*, 316(5):525–532, 2016.
- [2] Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. Supporting reflective public thought with considerit. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW '12)*, page 265, New York, New York, USA, 2012. ACM Press.
- [3] Siamak Faridani, Ephrat Bitton, Kimiko Ryokai, and Ken Goldberg. Opinion space: A scalable tool for browsing online comments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1175–1184, New York, NY, USA, 2010. ACM.
- [4] Anna De Liddo and Simon Buckingham Shum. Cohere: A prototype for contested collective intelligence. *ACM Computer Supported Cooperative Work (CSCW 2010)*, 2010.
- [5] M. Klein. The mit deliberatorium: Enabling large-scale deliberation about complex systemic problems. In *2011 International Conference on Collaboration Technologies and Systems (CTS)*, pages 161–161, May 2011.
- [6] David Kahane, Kristjana Loptson, Jade Herriman, and Max Hardy. Stakeholder and Citizen Roles in Public Deliberation. *Journal of Public Deliberation*, 9(2):n/a, 2013.
- [7] Peter R. Darke and Shelly Chaiken. The pursuit of self-interest: Self-interest bias in attitude judgment and persuasion. *Journal of Personality and Social Psychology*, 89(6):864–883, 2005.
- [8] Dennis Chong, Jack Citrin, and Patricia Conley. When Self-Interest Matters. *Political Psychology*, 22(3):541–570, sep 2001.
- [9] Pamela Johnston Conover. The role of social groups in political thinking. *British Journal of Political Science*, 18(1):51–76, 1988.
- [10] Hyunwoo Kim, Eun-Young Ko, Donghoon Han, Sung-Chul Lee, Simon T. Perrault, Jihee Kim, and Juho Kim. Crowdsourcing Perspectives on Public Policy from Stakeholders. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, pages 1–6, New York, New York, USA, 2019. ACM Press.
- [11] Tanja Aitamurto and Kaiping Chen. The value of crowdsourcing in public policymaking: Epistemic, democratic and economic value. *Theory and Practice of Legislation*, 5(1):55–72, 2017.
- [12] John Prpić, Araz Taeihagh, and James Melton. The Fundamentals of Policy Crowdsourcing. *Policy & Internet*, 7(3):340–361, 2015.
- [13] Tanja Aitamurto, Hélène Landemore, David Lee, and Ashish Goel. *Crowdsourced off-road traffic law reform experiment in Finland. Report about idea crowdsourcing and evaluation*. 2014.
- [14] Tanja Aitamurto and Hélène Landemore. Crowdsourced Deliberation: The Case of the Law on Off-Road Traffic in Finland. *Policy and Internet*, 8(2):174–196, 2016.

- [15] Cynthia R. Farina, Dmitry Epstein, Josiah Heidt, and Mary J. Newhart. Designing an Online Civic Engagement Platform: Balancing "More" vs. "Better" Participation in Complex Public Policymaking. *International Journal of E-Politics*, 5(1):16–40, 2014.
- [16] Dmitry Epstein, Cynthia Farina, and Josiah Heidt. The value of words: Narrative as evidence in policy making. *Evidence and Policy*, 10(2):243–258, 2014.
- [17] Nam Wook Kim, Jonghyuk Jung, Eun-Young Ko, Songyi Han, Chang Won Lee, Juho Kim, and Jihee Kim. BudgetMap: Engaging Taxpayers in the Issue-Driven Classification of a Government Budget. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*, pages 1026–1037, 2016.
- [18] Narges Mahyar, Michael R. James, Michelle M. Ng, Reginald A. Wu, and Steven P. Dow. CommunityCrit : Inviting the Public to Improve and Evaluate Urban Design Ideas through Micro-Activities. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pages 1–14, 2018.
- [19] Hans Stiegler and Menno D. T. de Jong. Facilitating personal deliberation online: Immediate effects of two considerit variations. *Computers in human behavior*, 51:461–469, 2015.
- [20] Diana C. Mutz. *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press, 2006.
- [21] Jeff Conklin and Michael L Begeman. gibis: A hypertext tool for exploratory policy discussion. *ACM Transactions on Information Systems (TOIS)*, 6(4):303–331, 1988.
- [22] Frank Vanclay, Ana Maria Esteves, Ilse Aucamp, and Daniel M. Franks. Social Impact Assessment: Guidance for assessing and managing the social impacts of projects. *International Association for Impact Assessment*, 1(April):98, 2015.
- [23] Mark S. Reed, Anil Graves, Norman Dandy, Helena Posthumus, Klaus Hubacek, Joe Morris, Christina Prell, Claire H. Quinn, and Lindsay C. Stringer. Who's in and why? A typology of stakeholder analysis methods for natural resource management. *Journal of Environmental Management*, 90(5):1933–1949, 2009.
- [24] Cynthia R. Farina, Mary J. Newhart, and Josiah Heidt. Rulemaking vs. democracy: Judging and nudging public participation that counts. *Cornell e-Rulemaking Initiative Publications*, (17), 2014.
- [25] Hannah M. Tuller, Christopher J. Bryan, Gail D. Heyman, and Nicholas J. S. Christenfeld. Seeing the other side: Perspective taking and the moderation of extremity. *Journal of Experimental Social Psychology*, 59:18–23, 2015.
- [26] Richard Van Der Wurff, Knut De Swert, and Sophie Lecheler. News Quality and Public Opinion: The Impact of Deliberative Quality of News Media on Citizens' Argument Repertoire. *International Journal of Public Opinion Research*, 30(2):233–256, 2018.
- [27] Weekeun Park Jungae Yang Minjae Chi, Sunho Kim and Sewook Oh. *2016 Whitepaper on Korean Online News Media*. Korea Press Foundation, Seoul, Republic of Korea, 2016.
- [28] Scott A. Golder and Bernardo A. Huberman. Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, 2006.

- [29] Hui Guo, Anca Ivan, Rama Akkiraju, and Richard Goodwin. Learning Ontologies to Improve the Quality of Automatic Web Service Matching. In *IEEE International Conference on Web Services (ICWS 2007)*, pages 118–125. IEEE, jul 2007.
- [30] Eric Corbett and Christopher A. Le Dantec. The problem of community engagement: Disentangling the practices of municipal government. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 574:1–574:13, New York, NY, USA, 2018. ACM.

Acknowledgment

I would first like to thank my advisor Juho Kim for his considerate, thorough, and constructive advice on my life as a researcher as well as my research. I would also like to thank my collaborators for their wonderful ideas, critical feedback, and enormous efforts on every part of this research. I would also like to thank all members of KIXLAB for their valuable comments for the research. Finally, I would like to appreciate endless support from my family.

Curriculum Vitae in Korean

이 름: 김현우

전 자 주 소: pkh1694@gmail.com

학 력

2010. 2. – 2013. 2. KAIST 부설 한국과학영재학교 (3년 졸업)
2013. 3. – 2017. 8. 한국과학기술원 화학과 (전산학 복수전공) (학사)
2017. 8. – 2019. 8. 한국과학기술원 전산학부 (석사)

경 력

2017. 8. – 2017. 12. 한국과학기술원 전산학부 조교

연구 업 적

1. H. Kim, E.-Y. Ko, D. Han, S. Lee, S. Perrault, J. Kim, and J. Kim, “Crowdsourcing Perspectives from Policy Stakeholders,” *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*.