BudgetMap: Engaging Taxpayers in the Issue-Driven Classification of a Government Budget

Nam Wook Kim[1] **Jonghyuk Jung**[2] **Eun-Young Ko**[2] **Songyi Han**[2] **Chang Won Lee**[2]
Juho Kim[2 3] **Jihee Kim**[2]
[1] Harvard SEAS Cambridge, MA USA
[2]KAIST Daejeon, Republic of Korea
[3]Stanford University Stanford, CA USA
namwkim@seas.harvard.edu {heretor92, key44, blairhan0126, chiyah, juhokim, jiheekim }@kaist.ac.kr

ABSTRACT
Despite recent efforts in opening up government data, developing tools for taxpayers to make sense of extensive and multi-faceted budget data remains an open challenge. In this paper, we present BudgetMap, an issue-driven classification and navigation interface for the budgets of government programs. Our novel issue-driven approach can complement the traditional budget classification system used by government organizations by reflecting time-evolving public interests. BudgetMap elicits the public to tag government programs with social issues by providing two modes of tagging. User-initiated tagging allows people to voluntarily search for programs of interest and classify each program with related social issues, while system-initiated tagging guides people through possible matches of issues and programs via microtasks. BudgetMap then facilitates visual exploration of the tagged budget data. Our evaluation shows that participants' awareness and understanding of budgetary issues increased after using BudgetMap, while they collaboratively identified issue-budget links with quality comparable to expert-generated links.

Author Keywords
Budget classification; budget navigation; civic engagement; social issue; tagging; crowdsourcing; visual interface.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
A government budget is taxpayers' payment for services yet to be implemented. It is also considered as the single most important policy document of a government [19]. Accordingly, the ability for taxpayers to evaluate how a government spends their money is fundamental to a democracy [2].

To ensure budget transparency and public trust, many government administrations nowadays provide the public with data tables or interactive interfaces to understand how their fiscal resources are allocated. The impact of budget transparency is further amplified through public participation and collaboration [18]. Traditional participation channels include community scorecards, public expenditure tracking, and participatory budgeting [13]. More recently, online channels such as crowdsourced budget prioritization and digital budgeting have gained popularity. The advantage of using Internet-based technologies is well recognized as they enhance the delivery of quality public services and achieve broad public participation [20].

Despite numerous efforts for opening up government data, engaging taxpayers to make sense of the extensive and multi-faceted budget data remains an open challenge. The complexity arises as the budget is allocated to an immense number of public services and programs, as a result of reflecting various interests and tradeoffs in making budgetary decisions. Even though the budget proposals and plans are available online in many countries, these existing resources suffer from two main drawbacks: 1) they fail to reduce the complexity of the budget in their way of delivery to the general public, and 2) their static data format cannot accurately reflect public interests that constantly evolve over time. In addition, while taxpayers are capable of understanding complex issues and making informed decisions, government organizations lack suitable tools to leverage the wisdom of the crowd [26, 22].

To address these challenges, we present BudgetMap, an issue-driven navigation interface for the budgets of government programs. It allows navigating a government budget through a lens of social issues, which dynamically reflect public interests. To collect the necessary link information between social issues and budget programs, we explore human computation methods that elicit contributions from taxpayers. While domain experts might be able to find the links, this approach will not scale to constantly emerging social issues and millions of budget programs in a government. Automatic algorithms in a simple form would not work well as they may not be able to capture the nuances of complex social issues. Moreover, public participation can complement domain experts by introducing taxpayers' perspectives and local knowledge. It can also serve as an auditing channel to improve budget transparency. In this work, we turn to taxpayers for a scalable and participatory solution.

BudgetMap embeds tagging activities for voluntary users to participate in. To accommodate users with differing levels of motivation for participation, BudgetMap provides two modes of tagging: user-initiated and system-initiated. User-initiated

CSCW '16, February 27-March 02, 2016, San Francisco, CA, USA
 ISBN 978-1-4503-3592-8/16/02$15.00 .
DOI: http://dx.doi.org/10.1145/2818048.2820004

tagging (UT) allows people to voluntarily search programs of interest and classify each program according to the related social issues by tagging an issue, while system-initiated tagging (ST) guides people through possible matches of issues and programs via microtasks. They complement each other in that ST facilitates lightweight exploration of arbitrary areas suggested by the system, while UT allows users to actively explore areas of their interest. We explore these tagging methods in the iterative design and evaluation of BudgetMap.

The first version supported both UT and ST from a single interface [10]. We conducted a controlled lab study to see the feasibility of issue-driven classification and navigation of a government budget. We found that participants formulated issues that span multiple budget categories, identified issue-program links with high accuracy, and reported improved awareness and interest on government budget policies. We also had a live deployment of the first version. While overall responses from the public were positive, we also observed that the level of participation was lower than expected, likely due to the complexity of the interface and usability issues. Based on these lessons, we redesigned BudgetMap, where UT and ST modes are separated for improved discoverability. We particularly focused on simplifying ST to facilitate lightweight contributions. We then conducted an online study to evaluate the new ST interface, where participants were asked to identify at least 150 issue-program links. We found that with ST, participants created issue-budget links with quality comparable to the links by budget experts. Moreover, our results suggest that tagging activities influence their informedness and perspectives on budget operations.

BudgetMap presents a crowdsourcing model in which users produce useful information for the system and future users, while the tasks help increase their awareness and interest on task-related issues. While some existing systems [29, 8] have shown to provide intrinsic benefits for participation while crowdsourcing information at scale, we contribute a novel application in the largely unexplored domain of government budget navigation. BudgetMap has implications for designing systems for civic engagement and other CSCW contexts that attempt to incentivize the crowd with benefits beyond monetary rewards.

The contributions of the paper are as follows:

- An introduction of issue-driven classification and navigation of budget data, which reflects trending social issues.
- BudgetMap, a system that solicits taxpayers to identify issue-budget links and facilitates visual exploration of the tagged budget data.
- Results from evaluation studies showing the feasibility of the issue-driven approach, including high quality issue-budget tags as well as improved budget awareness.

RELATED WORK

We review previous work on improving government transparency, supporting public engagement, and enhancing information navigation, with a specific focus on the government budget domain.

Improving Budget Transparency

With the growing open government movement, it is now common for government organizations to enable public access to internal data on the web. For example, the city of Seoul discloses detailed budgetary information including real-time budget spendings and allocated budgets to every program and service[1]. However, simply opening up more budget information to the public does not suffice, especially when taxpayers face difficulty understanding and using it (e.g., legal language, administrative jargon, and hard-to-parse raw data files). For this reason, the public data is often processed and presented in the form of infographics or data visualizations. Often, the efforts to make sense of government data for the public are made outside of public sectors to improve government transparency. For example, OpenSpending[2] is a community-driven project that tracks and visualizes public financial information across the world. Many Bills [1] combines visualization and machine learning techniques to improve the readability and understandability of legislative documents. Our approach complements existing efforts by adding a familiar social issue dimension to budget data, thus lowering the barrier to understanding such data for taxpayers.

Supporting Public Participation

Better presentation and participation mechanisms for budget data can empower active input from the public. Recent advances in civic engagement and crowdsourcing have lowered the cost of participation and enabled public collaboration at large scale [21, 16], thereby presenting a new channel through which taxpayers engage with government activities. For example, BudgetChallenge[3] is a collective prioritization tool that asks taxpayers to make budgetary decisions given a fixed budget, while Buy a Feature[4] turns the budget prioritization into a serious game where people purchase features with limited amounts of money. Factful [9] is a news reading application that supports the political discussion of a government budget through crowdsourced fact-checking and contextual budget information.

Leveraging public participation at scale to address real-world problems can be found in other public domains as well. PeerToPatent attempts to improve the patent examination process by enabling the public to assess claims of pending patent applications[5]. ConsiderIt [12] supports public deliberation by allowing users to create, share, and adopt pro/con points of ballot measures, while integrating a fact-checking service through a public dialogue [11]. Similarly, OpinionSpace [5] provides a platform for collecting, visualizing, and analyzing public opinions on issues and policies. Moreover, many citizen science projects such as FoldIt and Test My Brain tackle scientific research questions in collaboration with the large number of citizen scientists on a web-platform.[6] BudgetMap contributes a novel effort to promote public participation in

[1] data.seoul.go.kr
[2] www.openspending.org
[3] www.budgetchallenge.org
[4] www.innovationgames.com
[5] www.peertopatent.org
[6] See Wiggins and Crowston [30] for a detailed overview and various types of citizen science projects.

addressing real-world problems. In particular, BudgetMap engages taxpayers to classify budget programs by social issues of public interest, which makes the budget data more transparent and comprehensible for the public, while increasing public awareness and interests in budgetary issues.

Enhancing Information Navigation through Metadata

Organizing information with metadata is a popular method to help users browse and search information. Social tagging refers to collectively categorizing resources in a shared online environment [27]. Tags help the discovery of relevant resources, and the social relationships among taggers are a means of information discovery [17]. Since social tagging is a decentralized task, the vocabulary problem naturally arises [6], which indicates the variability of word usage among individuals. However, researchers have shown that shared vocabularies emerge from large-scale distributed tagging systems [7].

There are numerous existing systems that use tagging to improve information navigation. CommentSpace adds tags and links to assist social data analysis by allowing users to organize new evidence, identify others' findings, and synthesize both [31]. Crowdy collects summary labels from learners who are watching existing educational videos to improve the learning experience both for themselves and future learners [29]. Wordle is a text visualization technique often used to visualize tags to give an overview of information space [28]. Commercial examples include social bookmarking sites such as Delicious[7] and StumbleUpon[8], or social network sites such as Instagram [9] and Pinterest[10]. While social tagging can alleviate the problem of the overwhelming amount of government data, such advantage of tagging is less explored in the public domain [23]. A notable exception is the use of tags to improve the usability of library catalogues [25]. BudgetMap takes a similar approach to traditional tagging systems by supporting community-owned tags, but uniquely focuses on the budget domain and supports user- and system-initiated tagging methods. This process generates domain-specific tags and social issues, which improve the navigation of government budget information. Tagged and untagged budget areas are then visualized to facilitate further tagging and improve navigation.

ISSUE-DRIVEN BUDGET NAVIGATION

Our overarching goal in this research is to enhance public awareness of government activities. Specifically, we focus on helping taxpayers explore the budgets of government programs by leveraging familiar social issues. We introduce the idea of issue-driven navigation, which uses social issues as filters to navigate complex and multidimensional budget data. In order to generate tags that link issues and budget programs, we engage taxpayers in UT and ST tasks. In these tasks, taxpayers generate useful navigation cues for future taxpayers, while engaging in a meaningful experience themselves.

[7] www.delicious.com
[8] www.stumbleupon.com
[9] www.instagram.com
[10] www.pinterest.com

The motivation behind the issue-driven approach lies in problems with the existing budget classification system [4]. First, the current classification system often fails to reflect the rapidly changing economic environment. For instance, in a crisis environment or transitional economic environment, it is important to stabilize the economy through timely and dynamic adjustment in expenditure plans. Second,the rigidity of the current classification scheme forces budget items into a single category. Such inflexibility is a consequence of managerial convenience for the government to easily manage public expenditure. This often results in the difficulty of accurately evaluating the government's service performance as mapping budgets to results becomes harder. Finally, the current classification system is not designed with openness and accessibility in mind: it is currently difficult for citizens to understand budget data due to its complex language and obscure jargon.

To further illuminate the benefits for using "social issues" to navigate government budget programs, we borrow a case of a recent tragic accident in Korea and its impact on the budget. On April 16th, 2014, the Sewol ferry sank, and 295 people died and 9 people went missing. After the disaster, the public has raised concerns about the government's safety management and the budgets allocated to it. Because government programs related to public safety were spread out across various budget categories, such as transportation, health, accessibility, and defense, taxpayers had trouble understanding how their money was spent on public safety. In response, the government introduced a new budget accounting layer for public safety. However, it is unsustainable for a government to create a new accounting scheme whenever there is a new issue to be addressed. Thus, the need arises for a mechanism in which the public can actively participate in helping the budget classification system to reflect time-sensitive social issues.

We hypothesize that dynamic issue-driven classification by the public will have the following advantages over the existing classification system. First, our classification uses the language of the general public, therefore making the budget more accessible for navigation. Second, our classification can meet the timely needs and interests of the public because social issues by nature reflect the current status of a society. Third, budget classification using issues built by taxpayers can serve as constructive feedback for government officials in their budget planning and system improvement efforts. Fourth, taxpayers can be better informed of budgetary issues by engaging in activities related to government data [26]. However, we do not expect our issue-driven classification to replace the existing system, but instead supplement it while improving public understanding and awareness.

A set of design challenges exist in building issue-driven budget navigation. Since we turn to taxpayers for help in identifying issue-program links, a crucial design consideration is motivating them to participate in the tagging tasks. That is, the benefits conferred by engaging in the tagging tasks may not be seen higher than the advantages of the issue-driven navigation itself. In order to draw broad participation, system design should make the tasks manageable by lay users,

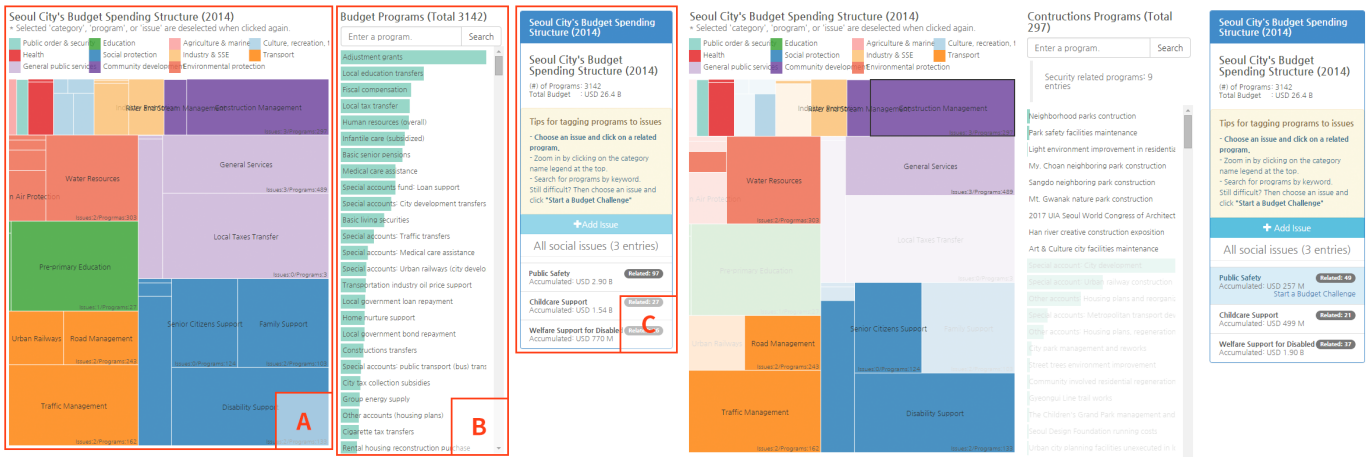

Figure 1: BudgetMap: (a) Overview of the BudgetMap interface. A: Budget category information of Seoul city. B: A list of programs sorted by budget size. C: A list of social issues. (b) Issue-driven navigation: When the user selects an issue, relevant budgets and programs are highlighted. On the middle panel, the matching programs are shown at the top.

provide multiple ways of participation for users with different expertise, and, most importantly, incentivize participation through benefits such as increased awareness and enjoyable experience.

BUDGET DATA MODEL

The budget data model used across this paper relies on two data sources (budgets and programs) managed by the City of Seoul[11]. A program is a unit of operating budget and refers to a set of activities or services that meet specific policy objectives of the government. We synthesize them into a single coherent data model where each entity has a program name, its allocated budget, and hierarchical upper categories to which the program belongs. For example, a disability allowance program belongs to a budget category named 'disability support', and this category in turn belongs to an upper category named 'social welfare'.

A social issue is a form of tag that may connect to multiple programs. It refers to a social problem or conflict raised by a society ranging from local to national issues. A budget program can be tagged by more than one issue and vice versa, resulting in a many-to-many relationship between issues and programs. We define three tag types, namely 'related', 'unclear', and 'unrelated', and keeps track of the number of tags for each type. When calculating the total budget of an issue, we aggregate the budgets of all related programs tagged by the issue. While the same data model is used in two iterations of a system design, they use data from different years—2014 and 2015, respectively.

BUDGETMAP: FIRST DESIGN ITERATION

Design and Implementation

BudgetMap is a web-based system designed to support issue-driven navigation of a government budget. BudgetMap solicits issue-program tags from taxpayers via UT and ST methods and facilitates visual exploration of the tagged budget data using social issues as navigation cues. We expect that the two tagging methods complement each other in serving the goal of helping taxpayers understand how budgets are allocated for specific social issues they care about. An earlier version of this iteration has been previously introduced [10], and this section summarizes main system components and describes major design decisions, rationale, and lessons.

The interface consists of three panels (Figure 1): a budget category visualization (A: treemap), a list of programs (B: bar chart), and a list of social issues (C: list group). The left panel displays an overview of the entire budget space, and the selected budget category is used for filtering programs in the bar chart. If no category is selected, all programs are displayed. On the middle panel, programs are sorted according to their budget size and can be searched by keywords. On the right panel, a summary of the selected program and a list of related issues are shown. If no program is selected, all issues registered in the system are displayed.

In UT, the user deliberately tags a budget with an issue. The system provides two ways to create a tag: the user selects a program of interest and then adds an issue to the program (Figure 1(a): the light-blue 'Add Issue' button on the right panel), or the user selects an existing issue first and adds a program to the issue by clicking on the program. In the former, if the user adds an issue without selecting a program, then the issue is added to the global issue list without any program attached. To assist with browsing and tagging, programs can be searched with keywords or filtered by a budget category.

ST solicits lightweight and structured contributions from users, especially those who may not deliberately search for programs and add tags (Figure 2). This task is activated when the user selects an issue and clicks the 'Start a Budget Challenge' button. A random program is displayed and the user is

[11] opengov.seoul.go.kr and cleanplus.seoul.go.kr

asked to decide whether the program is related to the selected issue by choosing one of three options: 'related,' 'unrelated,' and 'unclear.' Upon answering a question, a new question is displayed. All users' collective contributions (i.e., the total budget of related programs identified by the crowd) as well as individual contributions are displayed. With ST, the user can quickly add tags for the current issue without manually navigating the budget structure.

Once budget programs are tagged with issues, the user can navigate the budget space using a specific issue. All the budget categories and programs related to the selected issue are highlighted, while other elements are grayed out (Figure 1(b)).

Evaluation

For evaluation, we ran a lab study with UT and ST tasks in BudgetMap. Our goal was to see if our issue-driven and crowdsourced classification has advantages over the existing classification system. Specifically, we examined whether participants create issues that span multiple categories, accurately identify issue-program links, and develop interest and awareness on government budget programs.

Interfaces and Procedures

Nineteen participants were recruited from a behavioral study participant pool at a university (5 female, 14 male, age mean=21.79, stdev=2.74). They were assigned to use one of three budget navigation interfaces (a between-subject design): a simple treemap visualization interface of the existing budget classification that the city government uses (Treemap); BudgetMap with only the UT enabled (UT); and BudgetMap with only the ST enabled (ST).

To give participants a concrete objective to explore the government budget, we asked them to estimate the total budget related to an assigned issue, the calculation of which would be based on the information they would collect via the given interface[12]. Each interface was paired with one of the three social issues: public safety, childcare support, and welfare support for the disabled. Our research team selected these issues based on public interest in Korea as of September 2014. While new issues can be submitted by users, we fixed issues for the purpose of control experiments and focused on tagging budget programs with the given issues. For each issue-interface pair, participants were given seven minutes.

In a pre-task survey, each participant was first asked to come up with any social issues that they would like to know how much budget is allocated to. We intended to qualitatively understand what type of social issues participants are interested to know in the budget context. In the main task, participants saw the three issues (i.e., public safety, childcare support, and welfare support for the disabled) paired with the interface stimuli. The issue-interface pairs and their orders were counterbalanced across participants. A budget estimation task for each issue-interface pair was followed by a post-task survey.

[12]Because Treemap did not include detailed budget information in program units, we provided web links to Seoul's open data web portal (opengov.seoul.go.kr) so that participants could search through the open data without our interface support.

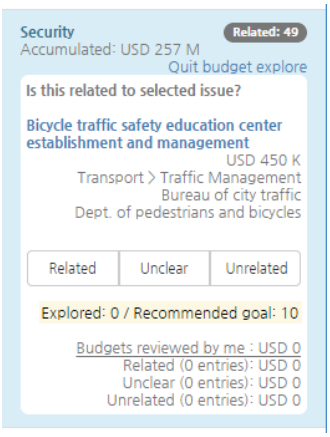

Figure 2: ST: the user is presented with a randomly chosen program and asked to determine the relationship with the selected issue.

In a post-task survey, we asked the following self-reported questions in 7-point Likert scale: Q1) whether a participant became more interested in the budget programs related to the given issue when using the given interface, and Q2) whether the given interface helped a participant better understand the budget programs related to the given issue.

Results

In the pre-task survey, 19 participants submitted 82 issues. A researcher in charge of Seoul's budget data in our research team qualitatively analyzed these issues: 30 of those spanned multiple government-defined categories and 10 of those were identified as trending social issues. For instance, the budgets related to 'public safety,' 'support for low-income families,' and 'support for minorities' spanned over multiple categories, and 'Sewol ferry accident' and '(recently reformed) welfare support for the elderly' are some of the examples of the highly debated current issues. This indicates that people think about issues in ways that the current budget classification system does not support.

To evaluate the quality of issue-program links that participants had identified, we created an expert reference solution of issue-program links. A researcher in our team and an external expert who has a work experience in the budget accounting team at a government organization rated every possible issue-program link for the three issues (total 9426 links) as 'unrelated,' 'weakly related (a program is not originally intended for the issue, but it may have some indirect effect),' or 'strongly related (a program directly aims to solve the problems related to the issue).' Note that 'weakly related' is a superset of 'strongly related'. Cohen's κ values between the two raters for each issue were 0.63, 0.54, and 0.79, respectively. Most differences between the raters were caused by how inclusive the definition of each issue was. The two raters then constructed the final reference solution by resolving their differences. Tags generated by study participants were evaluated against the reference solution by calculating the proportion of tags matched and the results are shown in 1; 'unclear' tags are not used in the evaluation.

Accuracy	Related				Unrelated	
	UT		ST		ST	
	weak	strong	weak	strong	weak	strong
Public Safety	0.79	0.59	0.50	0.40	0.93	0.95
Childcare Support	0.73	0.61	0.77	0.68	0.99	0.996
Welfare Support for Disabled	0.91	0.89	0.77	0.77	1.00	1.00

Table 1: Tag evaluation results for three issues ('weak' tags are a superset of 'strong' tags). Accuracy= (# of tags that are consistent with the reference solution)/(# of total related and unrelated tags).

As people's perception of a social issue can vary widely, we expect the 'related' links identified in BudgetMap to reflect various interpretations of an issue. We first observed that participants are more likely to find a correct 'unrelated' issue-program link, as indicated by higher accuracy for 'unrelated' than 'related' links. Next, the accuracy for 'weakly related' links is over 80% except for 'public safety.' In case of the 'public safety,' we posit that its definition significantly varies across people compared to the other two issues. This highlights one of the challenges in handling social issues that are perceived at a more abstract level. Finally, there was no significant difference in accuracy between strongly related and weakly related links. This suggests that participants were able to identify non-obvious and indirect links with similar accuracy to what they did for more direct links.

In the post-task survey, participants' answers were higher in both Q1 (Treemap: 3.16, UT: 4.89, ST: 4.42) and Q2 (TreeMap: 2.84, UT: 5.37, ST: 4.32) when using UT and ST interfaces than when using Treemap ($p < 0.05$ for Tukey HSD). UT scored higher than ST, although the result of Tukey HSD was not significant.

Preliminary Live Deployment: Although our lab study provides some evidence for our hypotheses on the issue-driven classification and navigation, the crowdsourcing nature of collecting issue-program tags calls for a large-scale deployment. We publicly launched the BudgetMap website in September 2014 and broadly advertised using social media and mailing lists. In the first five days of the deployment, 3,441 users visited the website. When we reviewed all the comments on Facebook that linked or shared our website, many appreciated the ability to view and navigate the city's budget programs, which suggests that presenting budget data in a publicly accessible way can provide value to the public. However, we also discovered challenges in guiding the public to actively engage in tagging tasks. While 11,459 actions were logged (clicks, search, tagging, and voting), only 697 issue-program tags (identified as either related or unrelated) were created using the two tagging methods. The number of tags was even less than that from the controlled experiment with only 19 participants. When we measured the accuracy of the collected tags compared to the reference solution used in the lab experiment, it ranged between 37% and 65% depending on the issue.

Design Lessons

L1: Difficulty of motivating taxpayers' participation
A major challenge in our live deployment was to guide the public to participate actively in the tagging tasks. While taxpayers' reactions to tagging were overall positive, casual users on the web have not participated in the tagging as actively as we initially expected. While it may be a typical issue faced by online communities (e.g., tragedy of commons [14]), it may also be due to the inherent difficulty of budgetary information or the complexity of the interface. We revisit this problem by redesigning the ST interface with a more simplified workflow in the next version.

L2: Need for quality control in the wild
There were instances of taxpayers' subjective and diverse interpretations of issues such as whether public safety includes public health issues or childcare support includes the construction of a children's playground. Quality control in our context differs from conventional crowdsourcing tasks in that our tagging tasks are inherently subjective and affected by diverse interpretations and sociopolitical views. At the same time, we also observed clearly incorrect tags as well. It is clear that there needs to be improved mechanisms to help users avoid unintentional mistakes and manage diverse responses. Another lesson we learned was that there can be a trade-off between lowering the bar for participation and ensuring quality responses.

L3: Need to improve the usability of the interfaces
We also observed usability issues based on feedback from live users. First of all, packing too many features in a single interface confused users. In particular, the ST interface was hard to discover. Also, a majority of users accessed the site through mobile devices, but the site was not mobile-friendly. These usability issues might have also contributed to the low level of participation. A live interface designed to encourage public participation should focus on providing a simple and usable user experience and making the tagging interface more prominent [3].

BUDGETMAP+: SECOND DESIGN ITERATION

Design and Implementation

We designed BudgetMap+ (Figure 3), based on the lessons learned from the first iteration. We focused on improving the ST interface as it is a key enabler for lowering the barrier to participation for lay users and for generating issue-program tags at large scale.

We separated UT and ST into two individual interfaces for simplicity and better discoverability (L1 and L3 from the design lessons). To better control the quality of issue-program tags, we added a simple quality control mechanism similar to majority voting (L2). Instead of simply classifying each tag as true when the number of votes for 'related' is greater than that of 'not related' we maintain a differential threshold (x) between the former and the latter. For instance, in order for a tag to be considered as 'related', the number of users identified an issue-program link as 'related' needs to be higher than the number of users identified the same link as 'unrelated' by x, and vice versa. We determine the optimal

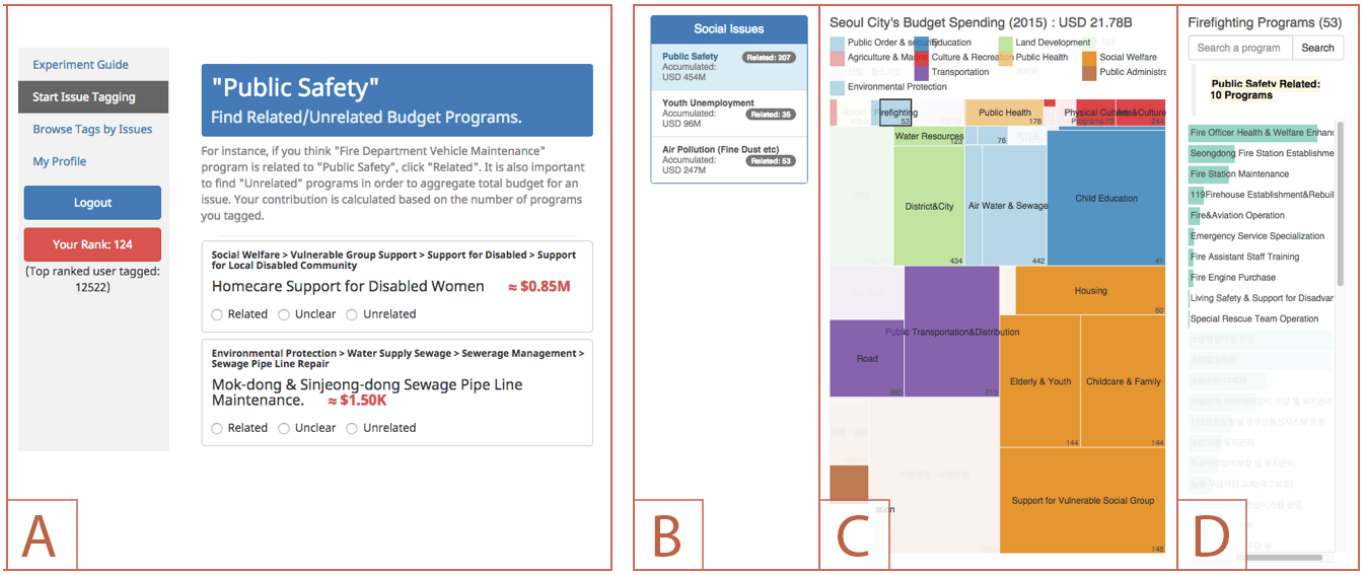

Figure 3: BudgetMap+: ST and UT are separated into two individual interfaces. (A): The new ST interface in which taxpayers can explore 10 randomly selected budget programs at a time along with their categories for a selected issue shown at the top. (B): A list of social issues. (C): budget category information of Seoul city. (D): A list of programs sorted by budget size.

threshold (x) value by considering the overall quality of the tags. While this mechanism might address spamming behaviors and incorrect tags, we recognize that further exploration is needed to design a more advanced mechanism for incorporating diverse perspectives and gaming behaviors to reflect a particular political view.

We also addressed many usability issues to improve the overall user experience (L3). First of all, BudgetMap+ is mobile-friendly and uses a responsive layout. To emphasize social issues as navigation cues, they are relocated to the left panel (Figure 3-B), while a list or programs are relocated to the right panel (Figure 3-D). UT is further simplified to help users spot missing programs for a selected issue, when the visualizations highlight uncovered areas from ST.

ST is significantly redesigned as a separate interface to better serve the idea of minimizing participation efforts. For each issue selected, the user can explore 10 randomly selected budget programs at a time (Figure 3-A) instead of one by one as in the previous version. The user can load the next 10 programs to explore more. In addition to the amount of collective and individual contributions, we also show the user's contribution ranking versus others and the number of programs tagged by the top user. In the tagging interface, the user can select an issue to redirect to the navigation interface to browse related tags.

To explore the feasibility of automatically identifying issue-program links, we implemented an algorithmic solution based on simple keyword matching. We first collected online news articles that are relevant to a social issue, by using the issue as a search term in Google News. We then extracted budget-related keywords from the articles by matching each word in an article with a list of keywords in the budget program names. Next, we ranked budget programs against the article text using tf-idf (term frequency–inverse document frequency). We considered the top programs as valid issue-program links. Unfortunately, the quality of the automatic solution was very low (less than 10% overall). The surface-level keyword matching has many limitations. Most importantly, it misses many possible issue-program links that require a nuanced and contextual understanding of both the issue and the program. We do not suggest that automated solutions are not feasible; that is, it could be improved if more structured features of the budget data were further utilized, such as relationships between programs indicated by categories, agencies, and departments. However, this simple test suggests that at least for some part of the link discovery process, human judgement would be more effective and appropriate than algorithmic discovery. For this reason, we chose to randomly select and present programs to the user in BudgetMap+.

Evaluation

To evaluate whether BudgetMap+ can help taxpayers generate issue-program tags, we conducted an online study. Since UT mostly remained the same, we only evaluated the ST interface in BudgetMap+. We publicly released BudgetMap+ to allow any interested taxpayers to sign up to participate in the study.

Participants

Total 104 participants (72 female, 32 male, age mean=25.32, std=7.36, min=19, max=63) who either currently live or have a primary address in the Seoul metropolitan area completed our study. We posted the link to the study website on our research team's Facebook page, and also ran a Facebook ads campaign for recruitment. We later learned that our study link was shared on one university's (in Seoul) community website

and a major information sharing community with 1.5M registered users. 36 out of 104 participants reported that they were currently paying taxes to the city, 17 participants have paid taxes to the city before, and the rest haven't paid any taxes to the city. The low proportion of tax-paying participants is possibly due to the high proportion of college students in our participant pool. We acknowledge that our study participants may not be a representative sample of the city's actual taxpayers, but everyone in the study was a registered voter who volunteered to participate in the study.

ST Tasks and Procedures
Each participant first completed a pre-task survey to register on our study website. The pre-task survey was designed to measure a participant's knowledge and opinions on the city's budget. Participants were then asked to create at least 150 issue-program tags, with a minimum requirement of 50 tags for each of the three social issues that we pre-selected. Upon creating the minimum number of tags, the post-task survey link was activated. The post-task survey asked about participants' experience on BudgetMap+ and the issue-driven approach. It also included the same set of questions as the pre-task survey and ended with general demographic questions. Participants could freely create more tags and navigate tags by issues using the visual interface after the study. The mean session length including the surveys and the tasks was 87m 8s (median=40m 7s). Participants could complete the study in multiple sittings by resuming their progress, as the system tracked activities for each user account.

Each participant who completed the entire study received a 10,000 Korean won (~$9) gift card for their participation. We also rewarded the top 10 participants who tagged the most budget programs with additional 20,000 Korean wons (~$18).

We selected three social issues based on public interest in Korea as of May 2015: public safety, air pollution, and youth unemployment. Public safety has been one of the most debated social issues since the aforementioned Sewol ferry accident. Air pollution is a recurrent issue in the spring as the density of micro air pollutants tends to increase during the season. Youth unemployment in South Korea has just marked its record high since 1999, hitting 11% in February 2015, which is more than double the general unemployment rate of the country[13].

Budget experts' qualitative evaluation of BudgetMap+
To gain insight into what crowdsourced issue-driven classification would mean for the government and domain experts, we conducted semi-structured interviews with three budget experts after the experiment. The experts were a government officer who works in the budget planning team at a local government (with the same state level as Seoul metropolitan government) (E1), a former national congressman (E2), and a government officer who is in charge of open government data in the Seoul metropolitan government (E3). Each interview session took about 20~30 minutes and the experts had a chance to interact with BudgetMap+, with the tags collected during the experiment loaded. We report the experts' feedback on BudgetMap+ in the discussion section.

Evaluation of crowd-generated tags
We compared the issue-program links identified by crowd-generated tags against a reference solution. First, we generate a crowd solution from crowd-generated tags. An issue-program link is identified as 'related' if the number of 'related' tags is more than that of 'unrelated' tags by x. To construct the reference expert solution, the two raters in the first experiment and an additional researcher in our team rated all possible links between the three issues and budget programs in the city's 2015 budget (total 12,039 links = 4,013 programs $\times$ 3 issues) as 'unrelated', 'weakly related', or 'strongly related'. Fleiss' κ values between the three raters for each issue were 0.64, 0.74, and 0.62, respectively. The three raters then constructed the final reference solution by resolving their differences. We acknowledge that our reference solution does not serve as ground truth, due to the subjective nature of issue-budget links. However, the reference solution can be a baseline to evaluate the crowd solution as it has the merit of raters' manual issue-budget link identification with concrete criteria, exhaustive coverage of the budget program space, and multi-step deliberation among the raters to reach consensus.

For the public safety issue, we were able to consider a government solution of issue-program links as the city government has recently started to provide a list of budget programs that are relevant to the issue to meet the high public interest. Cohen's κ between the issue-program links provided by the government and the 'strongly related' programs in the reference solution was 0.63. Cohen's κ increases to 0.69 when we include 'weakly related' links in the reference solution. While both solutions are fairly consistent with each other, they slightly differ in terms of the scope of the definition of public safety. For example, the government solution includes 'aid for the Korean war commemoration ceremony' and 'civil engineer award', whereas they are 'unrelated' in the reference solution. In contrast, the reference solution includes programs related to drugs and infectious diseases, while the government solution does not. Considering their fair similarity, however, we only use the reference solution in the evaluation.

We observed a mean of 730 tags (std=1,941, median=173, min=160, max:12,522) per participant. Unlike the first experiment, we combined both 'weakly related' and 'strongly related' links in the reference solution when evaluating 'related' links in the crowd solution. We present precision, recall, and accuracy results as well as the coverage of the tags for different thresholds in Figure 4. Precision indicates how accurate related or unrelated links in the crowd solution are compared to the reference solution, while recall shows how comprehensively participants identified related or unrelated links in the reference solution. Accuracy combines both related and unrelated links and indicates how accurate total links in the crowd solution are compared to the reference solution.

The results show that precision, recall, and accuracy initially increase but tend to drop after a certain threshold value. As shown in Figure 4 (4), total links without a threshold cov-

[13]The Korean national statistical office: www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=1063

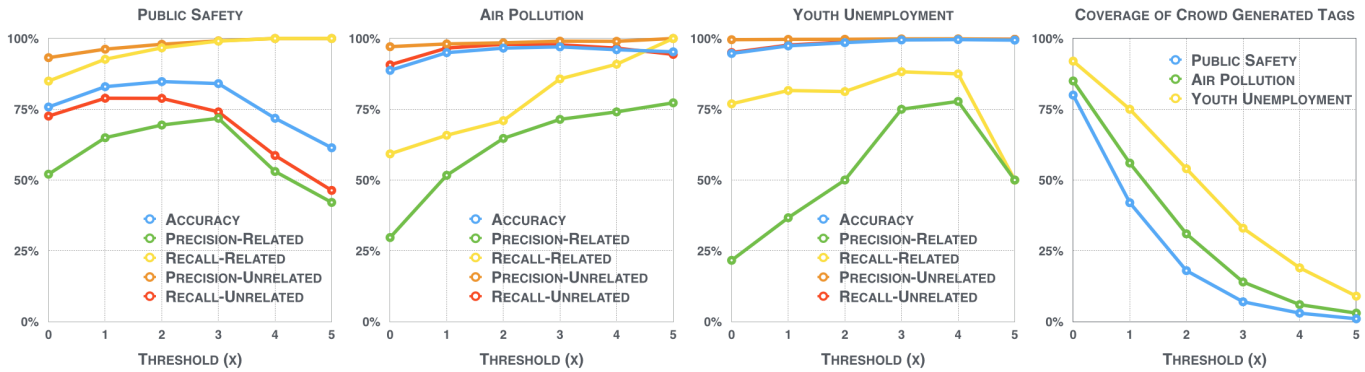

Figure 4: Crowd solution quality for different thresholds: **(1)** Accuracy= (# of programs links identified in the crowd solution that are consistent with the reference solution)/(# of programs links identified in the crowd solution), **(2)** Precision=(# of 'related (unrelated)' program links identified in the crowd solution that are consistent with the reference solution)/(# of 'related (unrelated)' program links identified in the crowd solution),**(3)** Recall=(# of 'related (unrelated)' program links identified in the crowd solution)/(# of 'related (unrelated)' program links in the reference solution), and **(4)** Coverage=(# of programs links identified in the crowd solution)/(# of programs).

ers more than 90% of entire programs, but are filtered to less than 1% with the threshold value of 5. Overall, the quality of results was the best at the threshold value of 3 and is summarized in Table 2. It should be noted that the optimal value may change particularly depending on total number of tags collected; that is, one can expect that it will increase with a large number of tags. For the issue of public safety, 270 budget programs were linked with the threshold value of 3, out of total 3,905 programs reviewed by study participants. The coverage of the tags is 7% (270 out of total 4,013 programs).

For the programs linked as 'related', 72% of them were matched with those of the reference solution (precision). All the 'related' programs in the reference solution were covered by the crowd solution (recall).[14] And 99% of the 'unrelated' programs in the crowd solution were matching with the reference solution (precision). The recall was 74%. Overall, 84% of the budget programs in the crowd solution were correctly identified as either 'related' or 'unrelated' (accuracy). Note that the accuracy rates were even higher for the other two issues. The overall precision of tags went up for all issues with the quality control. We used Cohen's κ to measure agreement between the crowd solution (with $x = 3$) and the reference solution. The values were 0.69, 0.76, and 0.81 for public safety, air pollution, and youth unemployment, respectively, which suggests a substantial level of agreement.

Pre-task vs. Post-task Survey Results

We analyzed self-reported measures of informedness and opinions in our pre-task and post-task surveys. Because participants answered the same set of questions in both surveys, we looked at changes in participants' answers to see if their level of informedness and opinions on the city budget changed, using a two-sided Wilcoxon signed-rank test. We only present a subset of noteworthy questions here.

[14]Note that we limit our budget space only to the programs that pass the quality control test when we calculate accuracy, precision, and recall, which implies that our budget space changes when we change the threshold value x. This is because 'true/false', 'positive/negative' classification becomes ambiguous for the programs that cannot pass the quality control test due to the small number of tags obtained.

	Public Safety		**Air Pollution**		**Youth Unempl.**	
Total Reviewed	3905		4010		4013	
Identified Links	270		574		1344	
Coverage (%)	7		14		33	
Accuracy (%)	84		97		99	
	R	UR	R	UR	R	UR
Precision (%)	72	99	71	99	75	98
Recall (%)	99	74	85	98	88	99

Table 2: Crowd solution evaluation results when threshold $x = 3$ (R: 'related', UR: 'unrelated').

To measure the change in participants' informedness against Seoul city's budget, we asked the following questions in 7-point Likert-scale (1: strongly disagree, 7: strongly agree), and observed their level of informedness significantly increased after tagging tasks.

- "I know in which budget programs Seoul's tax revenues are spent." 2.43→3.02, ($p < 0.01$)
- "I know some budget programs related to public safety." 2.56→3.38, ($p < 0.001$)
- "I know some budget programs related to air pollution." 2.24→3.29, ($p < 0.001$)
- "I know some budget programs related to youth unemployment." 2.89→3.40, ($p < 0.01$)

Comments from participants in the post-task survey also suggest increased informedness. For instance, 16 participants mentioned that they had learned that program budget sizes were surprisingly large, 17 participants mentioned that they had learned that the unit of budget programs were very de-

tailed, and 13 participants mentioned that they had learned that budget programs are quite diverse.

We also examined if participants' viewpoint on government spending changed after tagging tasks by asking the following questions in 7-point Likert-scale (1: strongly disagree, 7: strongly agree).

- "Considering overall operations of the city government, I'm willing to pay more tax" 2.85→3.05, ($p = 0.06$).
- "The city government is executing its budget in the right direction regarding public safety." 3.02→3.34, ($p < 0.05$).
- "The city government is executing its budget in the right direction regarding air pollution." 2.83→3.10, ($p < 0.05$).
- "The city government is executing its budget in the right direction regarding youth unemployment." 2.75→2.73, ($p = 0.89$).

While their perspective improved positively for public safety and air pollution and marginally for overall execution, it did not change for youth unemployment. All scores were on the negative side, however.

In summary, participants generated expert-quality tags through ST, showing the capability of non-expert taxpayers in building the issue-driven classification of budget programs. While the absolute scores were still low, their informedness of the government budget as well as their opinions on government execution regarding public safety, air pollution were notably improved and changed.

DISCUSSION

We now discuss lessons from our design and study of BudgetMap+, along with feedback from external experts including government officials. These points might have practical implications for future researchers designing related systems.

Benefits of Issue-driven Approach

In the post-task survey, 42 of 104 participants appreciated the ability to navigate budgets through the lens of social issues in their comments, describing BudgetMap+ as more interesting and intuitive than the traditional budget classification system. Participants commented that "Understanding government budget spending through issues would be more effective from a social problem resolution perspective."—p026, and "We encounter social issues in our daily lives. They made it easier for me to evaluate the suitability of budget programs"—p045. Similarly, experts acknowledged that the traditional budget classification system can be misleading (E1, E2), because it is mainly designed for administrative efficiency. They also commented that the issue-driven approach would be more intuitive to the general public (E1, E3). Twenty four participants said BudgetMap+ was engaging and helped increase their interests on the government budget. A participant commented that "I used to be indifferent about government spending but became more interested after the study. I would like to further know how budgets are allocated in detail."—p085, and an expert (E2) emphasized the potential usefulness of BudgetMap as an educational resource for the public.

Encouraging Participation & Collecting Opinions

In the initial design phase, we envisioned that issue-driven budget classification by taxpayers may serve as constructive feedback for government officials, and expected that taxpayers become more interested in and better informed of budgetary issues. In the post-task survey, 14 participants expressed that they identified budget programs that they found to be useless or wasteful, and 15 said they felt more accountable for the government's actions after using BudgetMap. One participant commented that "This kind of system can be useful for gathering public opinions and improving government transparency and accountability."—p002. An expert (E1) also said that it would be useful to collect public opinions about individual programs. Another expert (E3) noted that technological solutions like BudgetMap+ may play an educational role for taxpayers who would want to engage in participatory budgeting.

Incorporating Diverse Viewpoints of Taxpayers

While we evaluated the quality of crowd-generated tags by comparing them against the expert-generated reference solution, it is important to note that there is no ground truth answer for issue-program links. That is, low accuracy does not necessarily mean the lack of expertise of taxpayers, but may reflect different and yet legitimate viewpoints. A disagreement may come from the ambiguity in defining the scope of a social issue, which eight participants mentioned in the post-task survey. Also, we observed that there are at least 10 programs that participants tagged as related to public safety, that are not in the reference solution (e.g., 'remodeling of the Jang-chung sports stadium'). This suggests that taxpayers may be able to contribute diverse, contextual, and localized viewpoints. An expert (E1) noted that government officials can have different rationale for developing programs, which may not be obvious or revealed to taxpayers. He also mentioned that taxpayers are mostly concerned about their neighborhood. For future systems, it would be beneficial to have a conflict resolution mechanism or additional communication channels. Crowdsourced moderation mechanisms might be another solution that can help reach consensus by encouraging citizen-to-citizen discourse. We also plan to incorporate design insights from Lampe et al., who discuss the design and moderation results of political discussions on an online community site [15].

Limitations

Our study mainly evaluated the feasibility of issue-driven navigation by recruiting online participants to generate issue-program links. The experiment design had several limitations. First, participants were not allowed to create new issues. The issues we provided were rather broad, national issues. Future study is needed to see if narrow, localized issues produce different results. Also, while we attempted to simulate a live setting in the experiments, many study participants were students or non-taxpayers. This may have biased the tagging results. To more accurately reflect varying political viewpoints and socio-economic status of taxpayers, an actual live deployment with active taxpayers will be necessary. Likewise, study participants were largely drawn from

the young adult population who are likely to be more technologically inclined than the general public. It is a fundamental limitation many Internet-based civic systems face, as the elderly population is underrepresented in online spaces. Further design considerations are necessary to support technologically challenged groups.

The financial incentives given to our study participants also create some concerns. First of all, the financial incentives will become impractical when you want to encourage public participation at larger scale. In addition, more altruistic motivation can be leveraged in the systems like BudgetMap that aim to create public value. Therefore, we plan to incorporate non-financial motivations in the future versions of BudgetMap. First, personalized feedback with comparisons to others' performance can be provided, which has been shown to be an effective motivator [24]. We also plan to consider combining individual motivations and social motivations [14], by acknowledging individual progress as well as the social contribution. Additionally, we will explore gamification designs. In our study, we observed that two top tagging participants were having a close game, who each contributed more than 12,000 tags. While further investigation is necessary to more deeply understand taxpayers' motivation in using our system, we suspect that even the simple leaderboard showing the top contributors' performance was effective.

CONCLUSION AND FUTURE WORK

The open government movement introduces new challenges for the government to help the public make the best use of open data and to present open data in a readily accessible way. This paper explores the idea of classifying and navigating government budget programs using social issues, with the goal of leveraging taxpayers in generating issue-program links and increasing their awareness on budgetary issues. We implement this idea with BudgetMap, a web-based tagging and visual interface powered by the budgetary facts of the Seoul metropolitan government.

For future work, we plan to explore ways to incorporate diverse viewpoints of the public and to address the ambiguity in defining the scope of an issue. We also plan to introduce non-financial incentives to encourage public participation. In making use of the open data, we plan to expand current budget data to span multiple years. This will allow the navigation of budgets by issues over time, enabling analysis of temporal changes in budgets allocated to specific issues. For generalization, we will explore other government organizations and other data domains to which the idea of issue-driven navigation is applicable. For example, legislative bills, or even the specific contents within, can be categorized using social issues to reduce the barriers to public understanding. BudgetMap provides a simple yet powerful model for making open government data more accessible.

ACKNOWLEDGMENTS

This research is supported by the Undergraduate Research Program and High-Risk High-Return Program at KAIST. Nam Wook Kim is partly supported by the Kwanjeong Educational Foundation.

REFERENCES

1. Yannick Assogba, Irene Ros, Joan DiMicco, and Matt McKeon. 2011. Many Bills: Engaging Citizens Through Visualizations of Congressional Legislation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 433–442. DOI: http://dx.doi.org/10.1145/1978942.1979004

2. Phineas Baxandall and Benet Magnuson. 2008. Transparency.gov. 2.0: Using the internet for budget transparency to increase accountability, efficiency and taxpayer confidence. *Boston: MASSPIRG Educational Fund* (2008).

3. Susan L. Bryant, Andrea Forte, and Amy Bruckman. 2005. Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia. In *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work (GROUP '05)*. ACM, New York, NY, USA, 1–10. DOI: http://dx.doi.org/10.1145/1099203.1099205

4. Jack Diamond. 2006. *Budget system reform in emerging economies: the challenges and the reform agenda*. Technical Report. IMF.

5. Siamak Faridani, Ephrat Bitton, Kimiko Ryokai, and Ken Goldberg. 2010. Opinion Space: A Scalable Tool for Browsing Online Comments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 1175–1184. DOI: http://dx.doi.org/10.1145/1753326.1753502

6. G. W. Furnas, T. K. Landauer, L. M. Gomez, and S. T. Dumais. 1987. The Vocabulary Problem in Human-system Communication. *Commun. ACM* 30, 11 (Nov. 1987), 964–971. DOI: http://dx.doi.org/10.1145/32206.32212

7. Harry Halpin, Valentin Robu, and Hana Shepherd. 2007. The Complex Dynamics of Collaborative Tagging. In *Proceedings of the 16th International Conference on World Wide Web (WWW '07)*. ACM, New York, NY, USA, 211–220. DOI: http://dx.doi.org/10.1145/1242572.1242602

8. Kurtis Heimerl, Brian Gawalt, Kuang Chen, Tapan Parikh, and Björn Hartmann. 2012. CommunitySourcing: Engaging Local Crowds to Perform Expert Work via Physical Kiosks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 1539–1548. DOI: http://dx.doi.org/10.1145/2207676.2208619

9. Juho Kim, Eun-Young Ko, Jonghyuk Jung, Chang Won Lee, Nam Wook Kim, and Jihee Kim. 2015a. Factful: Engaging Taxpayers in the Public Discussion of a Government Budget. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA,

2843–2852. DOI: http://dx.doi.org/10.1145/2702123.2702352

10. Nam Wook Kim, Chang Won Lee, Jonghyuk Jung, Eun-Young Ko, Juho Kim, and Jihee Kim. 2015b. BudgetMap: Issue-Driven Navigation for a Government Budget. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 1097–1102. DOI: http://dx.doi.org/10.1145/2702613.2732932
11. Travis Kriplean, Caitlin Bonnar, Alan Borning, Bo Kinney, and Brian Gill. 2014. Integrating On-demand Fact-checking with Public Dialogue. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work (CSCW '14)*. ACM, New York, NY, USA, 1188–1199. DOI: http://dx.doi.org/10.1145/2531602.2531677
12. Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. 2012. Supporting Reflective Public Thought with Considerit. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 265–274. DOI: http://dx.doi.org/10.1145/2145204.2145249
13. Renee Kuriyan, Savita Bailur, Bjorn-Soren Gigler, Kyung Ryul Park, and Ashnah Kalemera. 2011. *Technologies for Transparency and Accountability: Implications for ICT Policy and Recommendations*. The World Bank.
14. Cliff Lampe, Rick Wash, Alcides Velasquez, and Elif Ozkaya. 2010. Motivations to Participate in Online Communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 1927–1936. DOI: http://dx.doi.org/10.1145/1753326.1753616
15. Cliff Lampe, Paul Zube, Jusil Lee, Chul Hyun Park, and Erik Johnston. 2014. Crowdsourcing civility: A natural experiment examining the effects of distributed moderation in online forums. *Government Information Quarterly* 31, 2 (2014), 317–326.
16. Daniel Lathrop and Laurel Ruma. 2010. *Open government: Collaboration, transparency, and participation in practice*. O'Reilly Media.
17. Cameron Marlow, Mor Naaman, Danah Boyd, and Marc Davis. 2006. Position paper, tagging, taxonomy, flickr, article, toread. In *Collaborative Web Tagging Workshop at WWW'06*.
18. Rosemary McGee and John Gaventa. 2010. Review of impact and effectiveness of transparency and accountability initiatives. *Institute of Development Studies* (2010).
19. OECD. 2002. OECD Best Practices for Budget Transparency. *OECD Journal on Budgeting* (2002).
20. United Nations. Department of Economic. 2014. United Nations e-Government Survey. (2014).
21. Tiago Peixoto. 2009. Beyond Theory: e-Participatory Budgeting and its Promises for eParticipation. *European Journal of ePractice* 7 (2009), 55–63.
22. Suzanne J Piotrowski and Gregg G Van Ryzin. 2007. Citizen attitudes toward transparency in local government. *The American Review of Public Administration* 37, 3 (2007), 306–323.
23. Christopher G Reddick. 2010. *Citizens and E-Government: Evaluating Policy and Management*. IGI Global.
24. Katharina Reinecke and Krzysztof Z. Gajos. 2015. LabintheWild: Conducting Large-Scale Online Experiments With Uncompensated Samples. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work (CSCW '15)*. ACM, New York, NY, USA, 1364–1378. DOI: http://dx.doi.org/10.1145/2675133.2675246
25. Louise F Spiteri. 2013. The structure and form of folksonomy tags: The road to the public library catalog. *Information technology and libraries* 26, 3 (2013), 13–25.
26. Susan Tanaka. 2007. Engaging the Public in National Budgeting: A Non-Governmental Perspective. *OECD Journal on Budgeting* 7, 2 (2007), 139.
27. Jennifer Trant. 2009. Studying social tagging and folksonomy: A review and framework. *Journal of Digital Information* 10, 1 (2009).
28. Fernanda B. Viegas, Martin Wattenberg, and Jonathan Feinberg. 2009. Participatory Visualization with Wordle. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (Nov. 2009), 1137–1144. DOI: http://dx.doi.org/10.1109/TVCG.2009.171
29. Sarah Weir, Juho Kim, Krzysztof Z. Gajos, and Robert C. Miller. 2015. Learnersourcing Subgoal Labels for How-to Videos. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work (CSCW '15)*. ACM, New York, NY, USA, 405–416. DOI: http://dx.doi.org/10.1145/2675133.2675219
30. Andrea Wiggins and Kevin Crowston. 2011. From Conservation to Crowdsourcing: A Typology of Citizen Science. In *Proceedings of the 2011 44th Hawaii International Conference on System Sciences (HICSS '11)*. IEEE Computer Society, Washington, DC, USA, 1–10. DOI: http://dx.doi.org/10.1109/HICSS.2011.207
31. Wesley Willett, Jeffrey Heer, Joseph Hellerstein, and Maneesh Agrawala. 2011. CommentSpace: Structured Support for Collaborative Visual Analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 3131–3140. DOI: http://dx.doi.org/10.1145/1978942.1979407