Approaches to Assessing Public Concerns: Building Linked Data for Public Goals and Criteria Extracted from Textual Content

Shun Shiramatsu, Tadachika Ozono, Toramatsu Shintani

Approaches to Assessing Public Concerns: Building Linked Data for Public Goals and Criteria Extracted from Textual Content

Shun Shiramatsu[1], Tadachika Ozono[1], and Toramatsu Shintani[1]

[1]Graduate School of Engineering, Nagoya Institute of Technology, Japan
{siramatu,ozono,tora}@nitech.ac.jp

Abstract. The importance of public involvement in Japanese regional societies is increasing because they currently face complicated and ongoing social issues due to the post-maturity stage of these societies. Since citizens who have beneficial awareness or knowledge are not always experts on relevant social issues, assessing and sharing public concerns are needed to reduce barriers to public participation. We propose two approaches to assess public concerns. The first is building a linked open data set by extracting public goals for a specific social issue aimed at by citizens or agents from articles or public opinions. This paper deals with hierarchical goals and subgoals for recovery and revitalization from the Great East Japan Earthquake manually extracted from related articles. The data set can be used for developing services to match citizens and agents who aim at similar goals to facilitate collaboration. The second approach is building a linked data set by extracting assessment criteria for a specific social issue from public opinions. This paper deals with candidate terms that potentially represent such criteria for a specific public project automatically extracted from clusters of citizens' opinions. The data set can be used as evidence for policy-making about the target project.

Keywords: Linked Data, Public Involvement, Concern Assessment, Goal Matching Service, Text Mining

1 Introduction

Japanese regional societies currently face complicated and ongoing social issues, e.g., disaster risks, dilapidated infrastructures, radiation pollution, and an aging population. Some Japanese researchers regard such troubling situations, that are partially due to the post-maturity stage of societies, as "a front-runner of emerging issues" [1]. Public involvement is an interactive communication process between stakeholders in deciding public policy [2] and has thus become more important to explore optimal solutions to complicated issues. For example, interactive and bottom-up communication is essential to design an optimal policy toward recovery and revitalization from the Great East Japan Earthquake [3].

Since citizens who have beneficial awareness or knowledge are not always experts on relevant social issues, public concerns need to be assessed and shared to reduce barriers to public participation. It is difficult to participate in issues without contextual or background information. Linked open data (LOD), which are semantically connected data on the basis of universal resource identifiers (URIs) and the resource description framework (RDF), plays an important role in fostering open government [4]. We aim to accrue LOD to share public concerns among citizens, governments, and experts to increase transparency that facilitates eParticipation. The structure of public concerns is an important context when building consensus. In this paper, we call the process of structuring public concerns "*concern assessment*".

Social networking systems (SNSs) such as Facebook are used for developing collaborative relationships not only in private but also in public spheres. To reduce barriers to participation and collaboration in public spheres, we consider that SNSs need to incorporate functions to share public concerns. In this paper, we focus on two types of public concerns, i.e., public goals and assessment criteria. Public goals that are aimed at by citizens are important for facilitating public collaboration. Assessment criteria for a specific social issue are also important for comparing and deliberating on multiple options to solve the issue.

We have developed an LOD set called SOCIA (Social Opinions and Concerns for Ideal Argumentation) that consists of Web content related to Japanese geographic regions, e.g., regional news articles, microblog posts, and minutes of city council meetings. The vocabularies for structuring relationships among such content and opinions are partially defined in the SOCIA ontology that we designed.

The main reason SOCIA deals with such web content is to share background context behind regional social issues. The background context that should be supported, however, includes not only relationships among content but also personal context, e.g., public goals and assessment criteria. The conventional SOCIA dataset and ontology could not yet support these kinds of personal context. In this paper, we expand the SOCIA ontology for structuring public goals and criteria and presents and present a prototyped dataset consistings of goals for revitalization from the Great East Japan Earthquake.

2 Literature Review

The International Association for Public Participation (IAP2) [5] and the Obama administration's Open Government Initiative (OGI) [6] have presented similar stages for public participation, i.e., the Spectrum of Public Participation and the Principles of Open Government shown in Fig. 1. The gradation in the figure represents the public impact of each stage. The figure also indicates the expected coverage of the use of LOD. Open data generally contributes to transparency and informativity, i.e., to the first stage. However, non-linked open data (e.g., CSV table data) generally lack interoperability. LOD is expected to be able to also contribute to the higher/collaborative stages because semantic links compliant

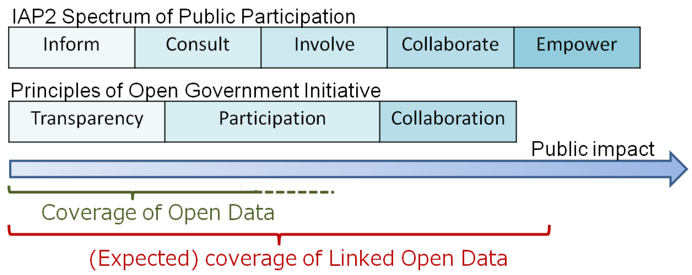

Fig. 1. Expected coverage of Linked Open Data

with RDF increase the interoperability of data and help us to reuse data for inter-organizational collaboration. Contextual information provided by the semantic links provides the potential for developing social web services to facilitate public collaboratoin. For example, an architecture based on linked data paradigm for participatory decision-making proposed by Kalampokis et al [7] can potentially be expanded to an architecture for supporting inter-organizational collaboration.

Over 40 countries currently have open data portals.[1] The number of open data portals has been increasing since 2009. In Japan, the Ministry of Economy, Trade and Industry operates a web site called the "Open Government Laboratory"[2] as an experimental Web site towards achieving eParticipation and eGovernment. The LOD Challenge Japan has been held since 2011, which is modeled on the Open Data Challenge in Europe. SOCIA and our system [8] received the ChallengeDay Award at the LOD Challenge Japan 2011.[3] Utilizing open data is rapidly promoted by "e-Government Open Data Strategy" of the IT Strategy Headquarters of the Japanese Government since 2012.

There are several vocabularies that can be used for public participation or collaboration, e.g., the participation schema [9] and the weighted interests vocabulary [10]. However, these vocabularies have not focused on assessing public concerns to facilitate public collaboration. This study presents how to deal with public goals and assessment criteria on the basis of LOD.

3 Manual Extraction of Public Goals

Public collaboration and consensus building between stakeholders are essential to enable revitalization from disasters, e.g., the Great East Japan Earthquake. Collaboration between multiple agents generally requires the following conditions:

- Similarity of the agents' goals or objectives
- Complementarity of the agents' skills, abilities, or resources

[1] `http://www.data.gov/opendatasites`
[2] `http://openlabs.go.jp/ (in Japanese)`
[3] `http://lod.sfc.keio.ac.jp/challenge2011/result2011.html (in Japanese)`

As the first step, this study focuses on the similarity of the goals. Sharing a data set of public goals can help citizens, who have similar goals, build consensus and collaborate with one another.

We focus on the following three problems related to public collaboration.

1. Citizens cannot easily find somebody whose goals are similar to their ones.
2. Stakeholders who have similar goals occasionally conflict with one another when building consensus because subgoals are sometimes difficult to be agreed on even if the final goal is generally agreed on.
3. A too abstract and general goal is hard to be contributed collaboratively.

We presume that the hierarchies of goals and subgoals play important roles to address these problems. First, the hierarchical structure can make methods of calculating the similarity between public goals more sophisticated. The hierarchy provides rich context to improve retrieval of similar goals. If the data set of public goals had only short textual descriptions without hierarchical structures, calculating the similarity between goals would be difficult and the recall ratio in retrieving similar goals would be lower. Second, visualizing the hierarchies is expected to support people in conflict to attain compromises. Third, dividing goals into fine-grained subgoals reduces barriers to participation and collaboration because small contributions to fine-grained subgoals are more easily provided.

We are planning to develop a Web service to match citizens and agents who are aiming at similar goals to facilitate collaboration. Toward this end, we expanded the SOCIA ontology to describe the public goals in Fig. 2. The property `socia:subgoal` enables us to describe the hierarchical structure of

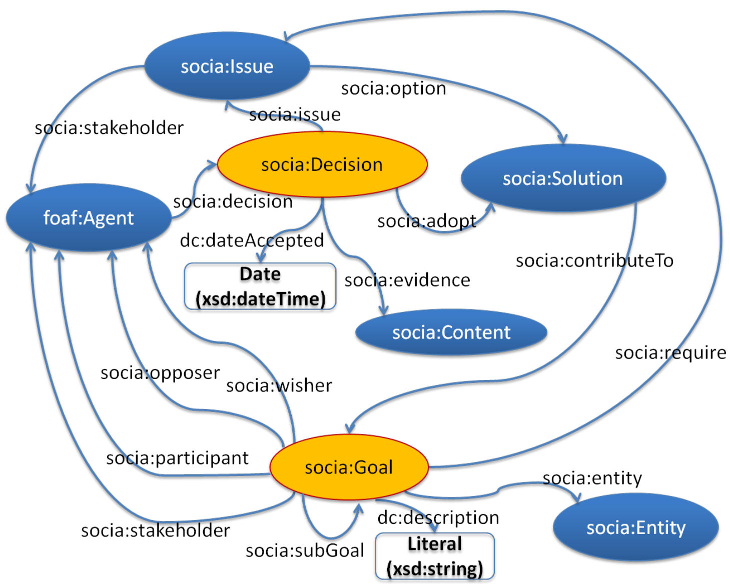

Fig. 2. Expanded classes in SOCIA ontology to represent public goals

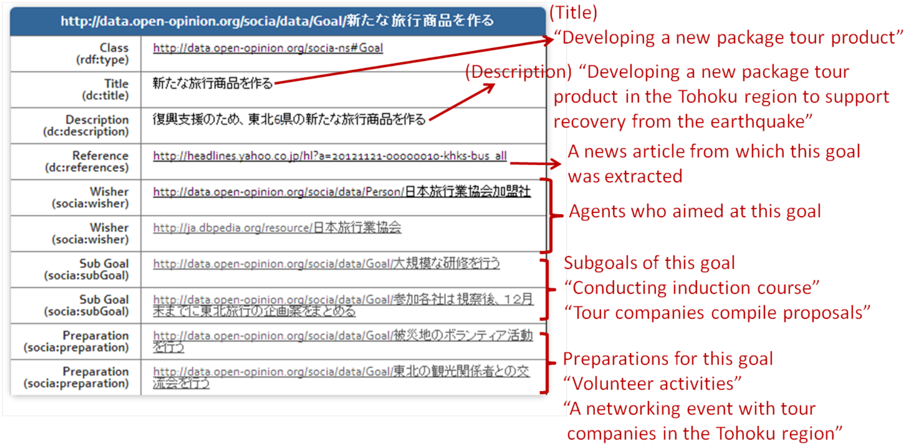

Fig. 3. Instance of public goal: "Developing new package tour product"

goals and subgoals. The public goal matching service that we aim to develop requires high-recall retrieval of similar goals to facilitate inter-domain, inter-area, and inter-organizational collaboration.

To develop a service for matching public goals, data on public goals need to be input by stakeholders who are aiming at the goals in person. Before developing such an SNS like mechanism to input stakeholders' goals and match them, we built an LOD set[4] by manually extracting public goals from news articles and related documents. The 657 public goals and 4349 RDF triples were manually extracted from 96 news articles and two related documents by one human annotator. The most abstract goal that is the root node of the goal-subgoal hierarchy is "revitalization from the earthquake".[5] The subgoals are linked from this goal with the `socia:subgoal` property.

The manually built LOD set can be used for developing a method of calculating the similarities between public goals. It can also be used as example seed data when citizen users input their own goals for revitalization. Fig. 3 shows an instance of a public goal to revitalize the Tohoku region from the Great East Japan Earthquake. This goal of "developing a new package tour product", has a title in Japanese, a description in Japanese, and two subgoal data resources.

The cosine similarity between public goals can be calculated on the basis of a recursive definition of a bag-of-features vector as:

---

[4] `http://data.open-opinion.org/socia/data/Goal?rdf:type=socia:Goal&limit=100 (in Japanese)`

[5] `http://data.open-opinion.org/socia/data/Goal/%E9%9C%87%E7%81%BD%E5%BE%A9%E8%88%88 (in Japanese)`

$$\mathrm{sim}(g_i, g_j) = \frac{\mathrm{bof}(g_i) \cdot \mathrm{bof}(g_j)}{\|\mathrm{bof}(g_i)\| \|\mathrm{bof}(g_j)\|} \quad (1)$$

$$\mathrm{bof}(g) = \frac{\alpha}{\|\mathrm{tfidf}(g)\|}\mathrm{tfidf}(g) + \frac{\beta}{\|\mathrm{lda}(g)\|}\mathrm{lda}(g) + \frac{\gamma}{|\mathrm{sub}(g)|} \sum_{sg \in \mathrm{sub}(g)} \frac{\mathrm{bof}(sg)}{\|\mathrm{bof}(sg)\|} \quad (2)$$

$$\mathrm{tfidf}(g) = \begin{pmatrix} \mathrm{tfidf}(w_1, g) \\ \vdots \\ \mathrm{tfidf}(w_{|W|}, g) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{|W|+|Z|}, \quad \mathrm{lda}(g) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathrm{p}(z_1|g) \\ \vdots \\ \mathrm{p}(z_{|Z|}|g) \end{pmatrix} \in \mathbb{R}^{|W|+|Z|}, \quad (3)$$

where g denotes a public goal, $\mathrm{bof}(g)$ denotes a bag-of-features vector of g, and $\mathrm{sub}(g)$ denotes a set of subgoals of g. Here, $w \in W$ denotes a term, $z \in Z$ denotes a latent topic derived by a latent topic model [11], and $\mathrm{tfidf}(w, g)$ denotes the TF-IDF, i.e., the product of term frequency and inverse document frequency, of w in a title and a description of g. The $\mathrm{p}(z|g)$ denotes the probability of z given g, $0 \leq \alpha, \beta, \gamma \leq 1$, and $\alpha + \beta + \gamma = 1$. The reason this definition incorporates a latent topic model is to enable short descriptions of goals to be dealt with because TF-IDF is insufficient for calculating similarities in short texts. The parameters α, β, and γ are empirically determined on the basis of actual data.

This prototyped method of calculating similarities should be tested, verified, and refined though experiments in future work using the LOD set of public goals that we present.

4 Automatic Extraction of Assessment Criteria

Citispe@k, which is our system for online public debate, supports manual tagging of assessment criteria for public opinions and Web content [12]. Transparent and participatory management of public issues requires assessing public concerns about the issue and criteria for the assessment. We call the criteria for the concern assessment "assessment criteria" in this paper. Although assessment criteria are diversified for each public issue, citispe@k does not yet support suggestion functions for setting new criteria. Here, we investigate the ability to apply text mining to extract assessment criteria from public opinions gathered from public workshops about a specific public project to maintain mountainous areas.

These workshops were held four times in four different areas. Participants at each workshop were divided into three to four debate teams. There were a total of 15 debate teams in these workshops. There were about five to six participants for each debate team, and one of them was a facilitator who did not state opinions. The opinions stated in each debate were manually structured according to the KJ-method [13], which consisted of brainstorming and grouping phases. The opinions were written on colored cards in the brainstorming phase. Red cards

Table 1. Citizens' opinions written on sticky notes with KJ method

	Area A	Area B	Area C	Area D	Total
No. of debate teams	4	4	3	4	15
Positive (red)	50	63	45	54	212
Negative (blue)	40	57	48	59	194
Demand (yellow)	37	48	57	53	205
Total	127	168	150	166	611

Table 2. Utterances in debate transcripts

	Area A	Area B	Area C	Area D	Total
No. of debate teams	4	4	3	4	15
In-range, citizens	678	685	450	681	2494
In-range, facilitators	509	401	279	279	1468
Out-of-range	293	534	288	252	1367
Total	1480	1620	1017	1212	5329

were for positive opinions, blue cards were for negative ones, and yellow cards were for demands or hopes. The opinions on the cards were manually classified into several groups in the grouping phase. The opinion groups had manually assigned labels. Although the group labels potentially represented assessment criteria if their expressions were uniform, their expressions were actually non-uniform and different for each debate team. Hence, we should apply text mining techniques to automatically extract candidate terms for assessment criteria. We employed a method of cluster analysis using text mining techniques, i.e., we clustered the opinions and extracted feature terms for each cluster.

The frequency of terms for short texts was insufficient for calculating similarities. The shorter the text content became, the lower the probability of the same term concurrently occurring in two kinds of content became, even if they were semantically close. Actually, participants at the workshops could not write lengthy opinions on the cards. The average number of morphemes in each opinion on a card was only 13.4. To address this problem, we used the latent Dirichlet allocation (LDA) [11], which is a frequently used model of latent topics in a document set. We used an implementation of the hierarchical Dirichlet process-LDA (HDP-LDA)[6] in the training phase of a topic model [14]. Although conventional LDA needs to be manually given the number of latent topics, HDP-LDA can determine this automatically.

Table 1 summarizes the number of opinions written on cards for each area. A total of 611 opinions were written by 15 debate teams. Since there were insufficient opinions to train the latent topic model, we also used utterances that were related to opinions on the cards in 15 debate transcripts. Table 2 lists the number of utterances in debate transcripts for each area. We regarded adjacent sentences sandwiched by interlocutors' names as one utterance. Although there were a total of 5329 utterances, the transcripts also included irrelevant utter-

[6] `http://www.cs.princeton.edu/~blei/topicmodeling.html`

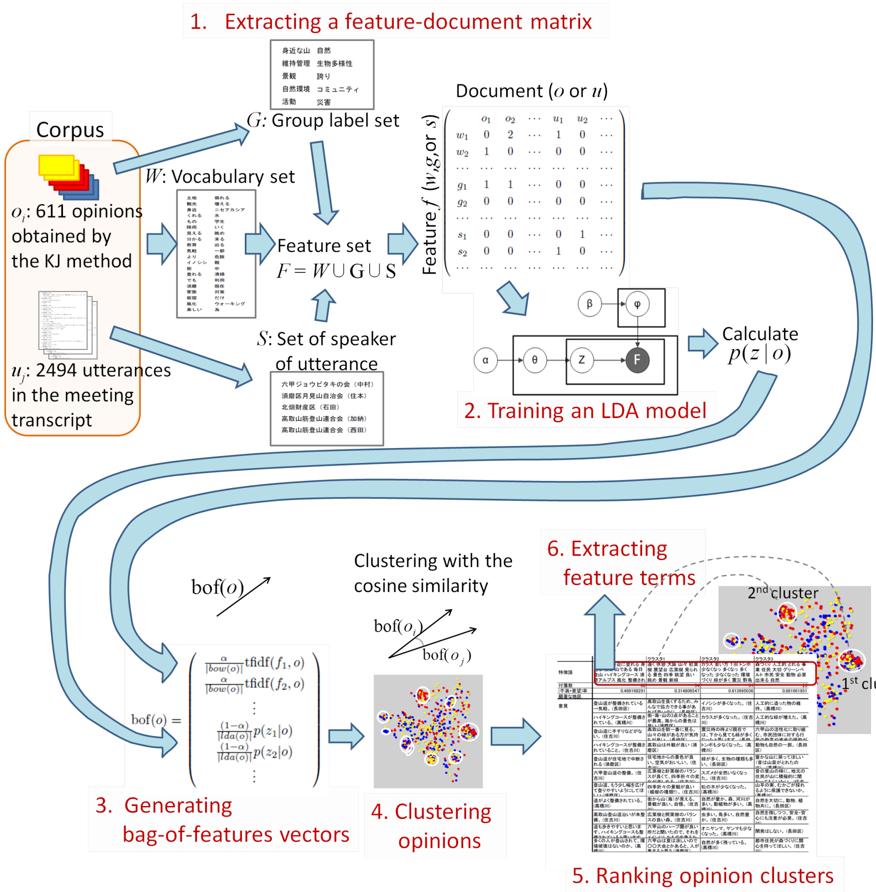

Fig. 4. Clustering opinions and extracting feature terms that potentially represent assessment criteria

ances, e.g., introductory and concluding remarks and facilitators' utterances. To divide transcripts into relevant (in-range) and irrelevant segments (out-of-range), we appended boundary markers to them. There were 2494 utterances other than those by facilitators in the in-range segment. These 611 opinions and 2494 utterances were used as a corpus to train HDP-LDA.

The procedure for clustering opinions and extracting feature terms that potentially represent assessment criteria is detailed in Fig. 4 Hereafter, let $o \in O$ be an opinion written on a card, $u \in U$ be an utterance in transcripts of the debate, and $d \in D = O \cup U$ be a document (i.e., d is any one of o or u). Let $w \in W$ be a morpheme N-gram ($N = 1, 2, 3$), $g \in G$ be a label for an opinion group manually assigned, $s \in S$ be a speaker (interlocutor) of an utterance in the debate transcripts, and $z \in Z$ be a latent topic derived by HDP-LDA.

To prepare for step 1 in the figure, determine feature set $F = W \cup G \cup S$ and document set $D = O \cup U$ appearing in the corpus. $w \in W$ can be extracted from the corpus through morphological analysis by using MeCab[7], which is a morphological analyzer. Morpheme N-grams that appear less than three times in the corpus are excluded because such rare expressions are not suitable for statistical processing. In the step 1, the feature-document matrix consists of frequencies of features in each document (i.e., opinion on cards or utterances in transcripts). In the step 2, an LDA model is trained from the feature-document matrix with the HDP-LDA tool. Probability $\mathrm{p}(z|o)$ is calculated using the parameters obtained with the trained model. In the step 3, the $\mathrm{bof}(o)$, which is a bag-of-features

Table 3. Feature terms of top four clusters that potentially represent assessment criteria

	1st cluster	2nd cluster
Feature N-grams (translated from Japanese)	Desirable, climbable, near, stroll, mountain, everyday climbing, hiking trail, Suma Alps	Far, break, Osaka, mountains, foliage tree, observation deck, broad-leafed tree, landscape, seasons
Interpretation by a user	Maintaining hiking trails	Landscapes of mountains
No. of opinions	82	54
(Negative+demand) ratio	0.488	0.315
Opinions near to cluster centroid (translated from Japanese)	The climbing trails are maintained → ease of use	Are there any collaborative tasks to make Mt. Takatori better
	The hiking trails are maintained.	Great because the sea, mountain, and town can all be seen. Good perspective from the sea.
	There are no handrails on the climbing trails.	I have a view of Mt. Takatori every morning. The mountain's green surroundings are pleasant.

	3rd cluster	4th cluster
Feature N-grams (translated from Japanese)	Crow, young people, dragonfly, decrease, increase, environment - creation, a lot of greenery	Grow forest, artificial, harvested, project, citizen, important, green belt, safety, animal
Interpretation by a user	Ecosystem	Growing forest and nature
No. of opinions	31	25
(Negative+demand) ratio	0.614	0.882
Opinions near to cluster centroid (translated from Japanese)	Number of boars increased.	Maintain artificially constructed things.
	Number of crows increased.	Artificial forests increased.
	Now there is a lot of greenery, more than when the Hanshin earthquake occurred.	Citizens groups for vitalizing mountainous areas need government financial help.

[7] `https://code.google.com/p/mecab/ (in Japanese)`

vector for o, is generated as:

$$\mathrm{bof}(o) = \frac{\alpha}{\|\mathrm{tfidf}(o)\|}\mathrm{tfidf}(o) + \frac{1-\alpha}{\|\mathrm{lda}(o)\|}\mathrm{lda}(o), \tag{4}$$

where vectors tfidf(o) and lda(o) are defined in the same way as that in Eq. (3) in the previous section. Parameter α satisfies $0 \le \alpha \le 1$. In the step 4, opinions o_i and o_j whose cosine similarity is greater than a particular threshold, θ, are grouped as cluster c. Clusters whose cosine similarity between their centroids is greater than θ are also grouped as one cluster. One opinion can belong to multiple clusters, i.e., this method is a kind of soft clustering. In the step 5, opinion clusters $c \in C$ are ranked in descending order of the number of opinions. In the step 6, w as candidate feature terms for each opinion cluster c are ranked with the following score based on pointwise mutual information (PMI):

$$\mathrm{score}(w,c) = \frac{\mathrm{PMI}(w,c) - E_w}{\sigma_w} \quad \text{and} \tag{5}$$

$$\mathrm{PMI}(w,c) = \log \frac{\mathrm{p}(w,c)}{\mathrm{p}(w)\mathrm{p}(c)} \quad , \tag{6}$$

where $E_w = \frac{1}{|C|}\sum_{c\in C}\mathrm{PMI}(w,c)$ and $\sigma_w = \frac{1}{|C|}\sqrt{\sum_{c\in C}(\mathrm{PMI}(w,c) - E_w)^2}$. Canonicalization by using standard variation σ_w in Eq. (5) is necessary because rare terms tend to be over-emphasized by only the PMI value.

We empirically set $\alpha = 0.5$ and $\theta = 0.65$ in this experiment , and the four top-ranked clusters and extracted feature terms with high scores are listed in Table 3. The feature terms for each cluster represent kinds of facets of opinions in the cluster. They potentially represent assessment criteria that are focused on by the opinion cluster. For example, the feature terms in the first cluster can be interpreted as "maintaining hiking trails" and those in the second cluster can be interpreted as "landscapes of mountains". The obtained clusters can be

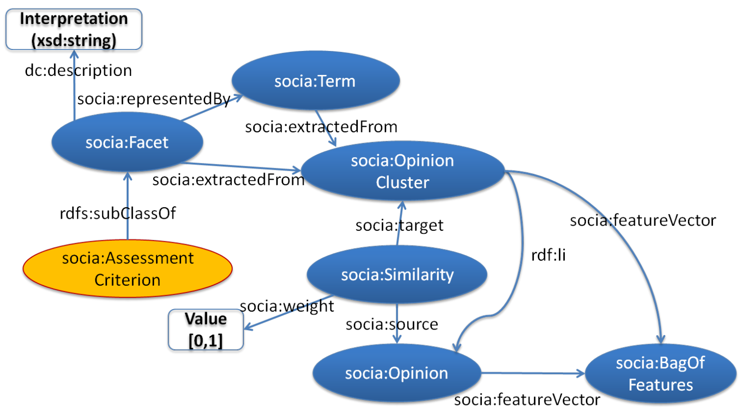

Fig. 5. Expanded classes in SOCIA ontology to represent assessment criteria

interpreted as facets or assessment criteria of opinions. The ratio for negative opinions and demand (blue and yellow cards), which is weighted according to the distance from cluster centroids, represents the degree of needs to be addressed by the target public project. For example, the ratio for the second cluster is low because the participants are satisfying the landscapes of mountains.

Fig. 5 outlines classes that are newly needed in SOCIA to describe the assessment criteria extracted from opinions. All opinion clusters correspond to `socia:Facet`. Clusters interpreted as assessment criteria can be instances of `socia:AssessmentCriterion`, which is the subclass of `socia:Facet`. An LOD set for assessment criteria can be built according to the classes in the figure. The links between assessment criteria and opinion clusters enable government and citizens to check context behind the concern assessment. Such structure can be utilized to develop tools for assessing and sharing public concerns.

Furthermore, we visualized the distribution of opinions to enable users to understand the overview using non-metric multidimensional scaling (NMDS) based on the inverse cosine similarity of $\mathrm{bof}(o)$ shown in Fig. 6. The colors of the points in the figure correspond to the colors of cards. Semantically close opinions are closely located by the NMDS algorithm. We use function `isoMDS` for NMDS, which is included in the library MASS in the statistical software R.[8] On the basis of this visualization, we developed an exploratory browsing interface on the Web browser shown in Fig. 7. Users can interactively browse neighboring opinion clusters of their clicked points in this browsing interface.

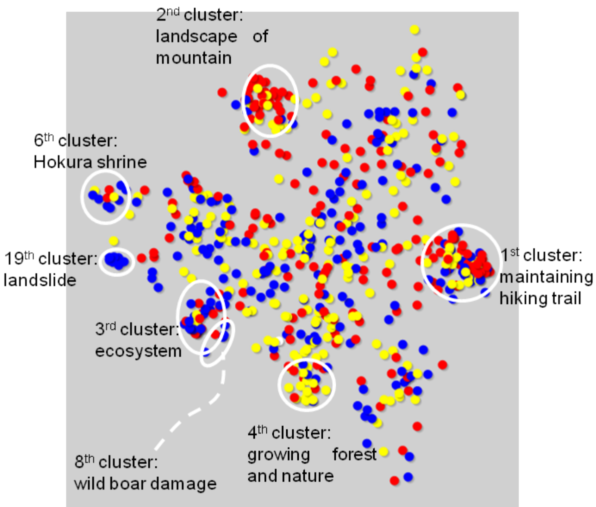

Fig. 6. Visualizing opinion distribution based on NMDS

[8] `http://www.r-project.org/`

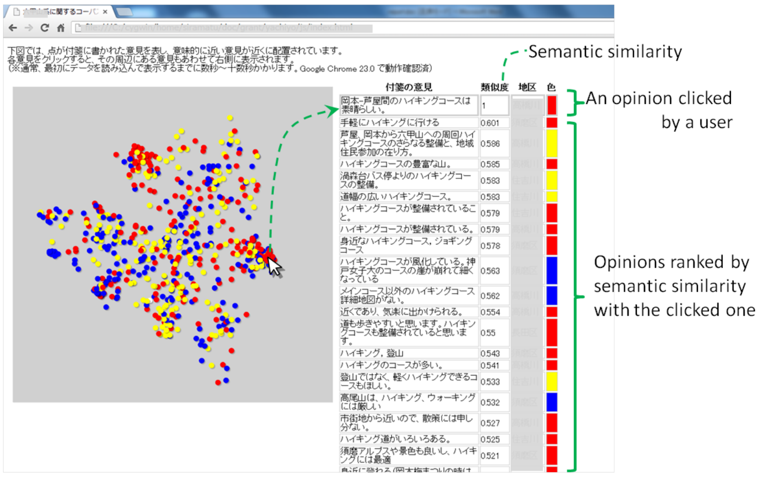

Fig. 7. Web application for exploratory browsing of opinions based on NMDS visualization

5 Conclusion

We focused on two types of public concerns, i.e., public goals and assessment criteria, and presented our approaches to assessing them. First, the LOD of public goals for revitalization from the Great East Japan Earthquake that was aimed at by citizens or agents was manually built. It contained 657 public goals and 4349 RDF triples manually extracted from 96 news articles and two related documents. The data set dealt with the hierarchical structure of goals and subgoals, which played important roles in attaining compromises. The hierarchy of subgoals was recursively used to generate a bag-of-features vector of a public goal in order to avoid decreasing the recall ratio. We are planning to test and verify our method of calculating the similarities between goals and to develop a goal matching service using this data set. The effectiveness of the recursive definition of the bag-of-features vector can be verified through empirically determining the parameters α, β, and γ on the basis of the LOD of public goals. If the optimal value of γ becomes significantly greater than 0, the subgoal structure can be regarded as actually significant. Second, we investigated the ability of applying text mining to extract assessment criteria from public opinions gathered at workshops on a public project to maintain mountainous areas. The feature terms automatically extracted from an opinion cluster helped us to interpret what kinds of assessment criteria were indicated by clusters. We also presented an extension of our ontology to build LOD for assessment criteria. Moreover, we developed an exploratory browsing interface to enable overviews of opinion clusters to be understood.

Acknowledgments. This work was supported by the Revitalization Promotion Program (A-STEP) (No. 241FT0304) of JST, the Grant-in-Aid for Young Scientists (B)

(No. 25870321) from JSPS, the SCOPE from the Ministry of Internal Affairs and Communications, the Rokko Sabo office of the Ministry of Land, Infrastructure, Transport and Tourism, and Yachiyo Engineering Co., Ltd. We would also like to sincerely thank Dr. Hayeong JEONG for her valuable advices.

References

1. Komiyama, H.: Vision 2050 and the role of japan toward the sustainable society. In: Proceedings of the 4th International Symposium on Environmentally Conscious Design and Inverse Manufacturing. (2005) 2–4
2. Jeong, H., Hatori, T., Kobayashi, K.: Discourse analysis of public debates: A corpus-based approach. In: Proceedings of 2007 IEEE International Conference on Systems, Man and Cybernetics. (2007) 1782–1793
3. Nishimori, M., Kanke, M., Tsutsuki, A., Kotake, N., Shirasaka, S., Yasui, T.: Optimal policy design for disaster-hit area of japan: bottom-up systems analysis of special zone for reconstruction by the isdm. In: Proceedings of the CESUN 3rd International Engineering Systems Symposium. (2012)
4. Hochtl, J., Reichstadter, P.: Linked open data - a means for public sector information management. In: Proceedings of the 2nd international conference on Electronic government and the information systems perspective. Volume 6866 of Lecture Notes in Computer Science., Springer (2011) 330–343
5. IAP2: Iap2 spectrum of public participation. http://www.iap2.org/associations/4748/files/IAP2%20Spectrum_vertical.pdf (2007)
6. WhiteHouse: Open government initiative. http://www.whitehouse.gov/open (2009)
7. Kalampokis, E., Hausenblas, M., Tarabanis, K.: Combining social and government open data for participatory decision-making. In: Electronic Participation - Proceedings of the 3rd IFIP WG 8.5 international conference, ePart 2011. Volume 6847 of Lecture Notes in Computer Science., Springer (2011) 36–47
8. Shiramatsu, S., Swezey, R., Sano, H., Hirata, N., , Ozono, T., Shintani, T.: Structuring japanese regional information gathered from the web as linked open data for use in concern assessment. In: Electronic Participation - Proceedings of the 4th IFIP WG 8.5 International Conference, ePart 2012. Volume 7444 of Lecture Notes in Computer Science., Springer (2012) 73–84
9. Styles, R., Wallace, C., Moeller, K.: Participation schema. http://purl.org/vocab/participation/schema (2008)
10. Brickley, D., Miller, L., Inkster, T., Zeng, Y., Wang, Y., Damljanovic, D., Huang, Z., Kinsella, S., Breslin, J., Ferris, B.: The weighted interests vocabulary 0.5. http://purl.org/ontology/wi/core (2010)
11. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet Allocation. Journal of Machine Learning Research **3** (2003) 993–1022
12. Swezey, R., Sano, H., Hirata, N., Shiramatsu, S., Ozono, T., Shintani, T.: An e-participation support system for regional communities based on linked open data, classification and clustering. In: Proceedings of the 11th IEEE International Conference on Cognitive Informatics & Cognitive Computing. (2012) 211–218
13. Kunifuji, S., Kato, N.: Consensus-making support systems dedicated to creative problem solving. International Journal of Information Technology and Decision Making **6**(3) (2007) 459–474
14. Teh, Y., Jordan, M., Beal, M., Blei, D.: Hierarchical dirichlet processes. Journal of the American Statistical Association **101**(476) (2006) 1566—–1581