Implementing a System Enabling Open Innovation by Sharing Public Goals Based on Linked Open Data

Teemu Tossavainen[1,2], Shun Shiramatsu[1], Tadachika Ozono[1], and Toramatsu Shintani[1]

[1]Graduate School of Engineering, Nagoya Institute of Technology, Japan
[2]School of Science, Aalto University, Finland
teemu.tossavainen@aalto.fi,{siramatu,ozono,tora}@nitech.ac.jp

Abstract. Social network channels are used to organise public goals, but unstructured conversation presents a challenge for identifying groups focusing on similar issues. To address the problem, we implemented a web system for creating goals from public issues and for discovery of similar goals. We designed SOCIA ontology to structure information as hierarchical goal structure. Potential collaborators can use the hierarchical information in consensus building. We proposed a method to calculate similarity between public goals on the basis of the hierarchical structure. To apply our proposed method to real-life situation, we designed and implemented an easy-to-use user interface to structure public goals by citizens. We are arranging workshops to use the system in real-life setting for gathering local issues from citizens of Ogaki city to formulate public goals to solve them. We will use the experiences gained from the workshops to improve the system for deployment for open use by communities to utilise open innovation in decision making and for facilitating collaboration between governmental agents and citizens.

Keywords: Linked open data, public collaboration, goal matching service

1 Introduction

Societies around the world are looking ways to improve efficiency of the public sector by increasing public participation [1]. Efficiency of public decision making is important, for example, for a society's ability to recover from large-scale natural disasters, like the great earthquake in Japan 2011. In addition, many industrial countries are facing large problems with ageing population and growing dependency ratio [2]. These issues increase the pressure to improve operation of public sector.

In our research, we focus on public goals. Regional societies make various public decisions in different levels of community's internal hierarchy to address public issues. Public goals are formed in the decision making process, which leads to a large amount of different goals, some of which are conflicting or overlapping.

An unsuccessfully set goal can lead to wasting of resources, for example a less important task may be given the attention, while addressing a more important issues is neglected, also a goal definition that is too general or vague hinders effective participation. More importantly, lack of coordination can lead to a situation where opportunities for cooperation are missed and possible synergy benefits and resource savings are not realised. Addressing aforementioned issues could potentially be beneficial and bring resource savings.

We aim to develop technologies for facilitating open innovation through collaboration in public spheres. Open innovation and eParticipation can be seen as possible solutions for improving public sector's efficiency. In eParticipation, "the purpose is to increase citizens' abilities to participate in the digital governance" [3], participation of citizens is needed in open innovation. One of the characteristic of open innovation is increased transparency. In one hand, information is necessity for enabling the public to participate in the decision making and in other hand the public's views and opinions are seen beneficial in open decision making. Implementing these methods could be a way to improve a community's ability to generate more informed public goals and to increase public participation. Information and communication techniques (ICT), like social networks, blogs, and eParticipation systems are channels for public conversation in eParticipation. These channels are used also for discussion for organising goals and events, but the discussion is unstructured and thus it is challenging to use the accumulated information for identifying groups that are focusing on similar issues. While it is possible to programmatically access the content of the discussion for automatic information extraction, the current solutions are not able to realise high accuracy for extracting hierarchical structure of goals or issues. Hence, we need to develop a system for manual structuring of goals and for collecting training data for the automatic extraction. We implemented a public goal sharing system, for enabling open innovation. The system has functions for creating and exploring goals and issues and discovering similar goals.

The structure of this paper is as follows. In section 2 we provide an overview of related works. In section 3 we present the system architecture, the implemented method for analysing the goal similarity, and introduce the system functionality. In conclusion section we present the summary and discuss about possible future work.

2 Related works

2.1 Goal Management Tools

In the research field of project management, goals are commonly structured as hierarchies by subdividing goals into subgoals. Instances of such structures are seen in the Thinking Process of the Theory of Constraint (TOC) [4] and the Work Breakdown Structure (WBS) in the Project Management Body of Knowledge (PMBOK) [5]. Although these models are generally used for project management within an organization, some researchers apply them to public sector problems. Especially, TOC is tried to be applied to the recovery from disasters[6, 7]. In our previous research [8], we built a linked open dataset of public goals for promoting recovery and revitalization from the Great East Japan Earthquake[1]. The dataset is based on the hierarchical structure that can be regarded as a simplification of the above models. We assumed that simplifying the models is needed for applying them to public use and for ensuring the interoperability of data.

There are many Web services for sharing tasks and managing projects, e.g., Trello[2], Cyboze Live[3], and Backlog[4]. However, they do not support functions for sharing goals and finding potential collaborators in public spheres.

2.2 Issue Sharing and Collaborator Matching

FixMyStreet[5] is a collaborative web service to enable citizens to report, view or discuss local problems such as graffiti, fly tipping, broken paving slabs or street lighting by locating them on a map [9]. Neighborland[6] is also a collaborative web service to share citizens' wishes by filling blanks in a form: "I want ____ in ____". Although the key concepts of these services are similar to that of this research, we focus on retrieval of potential collaborators and a user interface for subdividing public goals that are not dealt with the above web services.

There have been many researches about 2-sided matching problem based on the game theory [10]. We, however, do not limit the number of collaborators because dealing with public issues requires a lot of collaborators.

3 System architecture and implementation

In our research, we developed a system for sharing public goals. Agents use the application to input new goals and issues. Goals are structured in hierarchical order, where a goal has a set of goals as subgoals that are partial solutions for the parent goal. The goal structure is important because it allows dividing an abstract goal to more tangible and concrete subgoals, which are easier to participate in. Additionally, the agent can explore goals by filtering them by various criteria, e.g., with a keyword search, which facilitates the agent to locate relevant goals for more detailed inspection. Moreover, the agent can discover goals with a similarity search, where the system suggests similar goals. The goal discovery enables the agent to discover opportunities for collaboration over common agenda. Additionally, the agent can use the system to discover other agents, who have similar goals. The discovery function facilitates the agent to identify those parties that has same kind of aspirations, which are potential partners for cooperation. Additionally, the agent will be able to visually compare possible partners' goal trees, which facilitates the agent to have better understating about the level of

[1] http://data.open-opinion.org/socia/data/Goal?rdf:type=socia:Goal
[2] http://trello.com/
[3] http://cybozulive.com/
[4] http://www.backlog.jp/
[5] http://www.fixmystreet.com/
[6] https://neighborland.com/

similarity, it also aids in a negotiation process between the agents by identifying possible conflicts of interest.

3.1 SOCIA Ontology

We use SOCIA (Social Opinions and Concerns for Ideal Argumentation) resource description framework (RDF) ontology to build linked data. Firstly, SOCIA links the background context to the goal data, secondly it describes the relations between entities and defines their structure. Figure 1 shows the relevant part of the SOCIA ontology structure. It describes goal related information and it is marked with socia: prefix. dc: prefix is used with Dublin metadata initiative's metadata terms[7], :foaf prefix in World Wide Web consortium's friend of a friend (FOAF) ontology[8], and geo: prefix is used with the GeoNames ontology[9]. The socia:issue class represents a public issue, which has dc:title, dc:description, socia:references, dc:spatial, dc:createdDate, and dc:creator properties. The socia:references property contains links to external data sources, for example news articles or blog posts. dc:spatial reference links an issue to a geo:SpatialThing class, which indicates the related spatial information, it also contains a socia:solution relation to a goal. The socia:Goal class represents a public goal. It contains dc:title, dc:description, dc:status, dc:spatial, dc:creator, dc:dateSubmitted, socia:desiredTargetDate, socia:requiredTargetDate, socia:status, and socia:subGoalOf properties. The socia:status property indicates the current status of the goal, which can have one of following values: "not started",

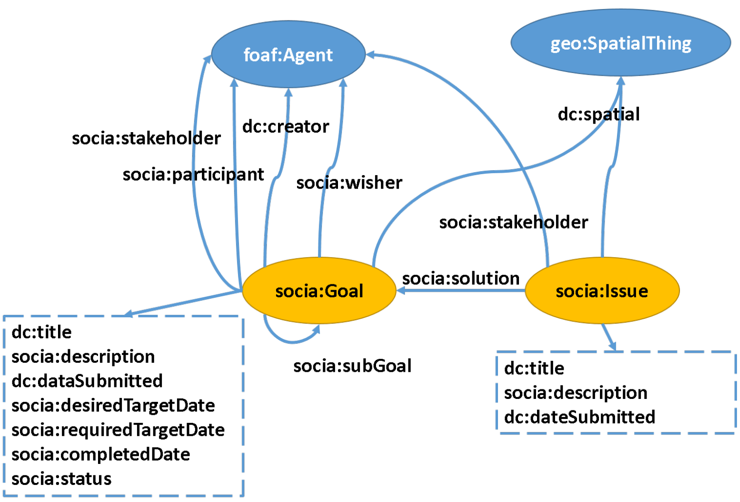

Fig. 1. SOCIA ontology to represent public goals

[7] http://dublincore.org/documents/2012/06/14/dcmi-terms/?v=terms

[8] http://xmlns.com/foaf/spec/

[9] http://www.geonames.org/ontology/documentation.html

"started", "abandoned", and "completed". The socia:subGoal property has a set of subgoals, which forms the hierarchical goal structure. The socia:participant property contains a set of agents that are participating in the goal. Figure 2 shows the classes and relations that are used to store the similarity value between a pair of goals. Socia:AnnotationInfo has a source and a target relation with the two socia:goal instances and it contains the socia:weight property that indicates the level of similarity. The socia:weight property can have values from zero to one, where greater value indicates more similar goals.

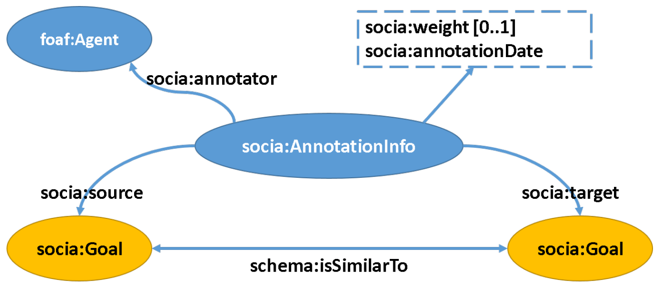

Fig. 2. SOCIA AnnotationInfo to save goal similarity

3.2 Goal Similarity

The similarity between public goals g_i and g_j can be calculated using a cosine measure between $\mathrm{bof}(g_i)$ and $\mathrm{bof}(g_j)$, a bag-of-features vectors of the goals:

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

$$(2)$$

In our past research [8], we defined $\mathrm{bof}(g)$ as a weighted summation of surficial (TF-IDF) features, latent (LDA) features, and recursive (subgoal) features:

$$\mathrm{bof}(g) = \frac{\alpha}{\|\mathrm{tfidf}(g)\|}\mathrm{tfidf}(g) + \frac{\beta}{\|\mathrm{lda}(g)\|}\mathrm{lda}(g) + \frac{\gamma}{|\mathrm{sub}(g)|} \sum_{sg \in \mathrm{sub}(g)} \frac{\mathrm{bof}(sg)}{\|\mathrm{bof}(sg)\|} \quad (3)$$

$$\mathrm{tfidf}(g) = \begin{pmatrix} \mathrm{tfidf}(w_1, g) \\ \vdots \\ \mathrm{tfidf}(w_{|W|}, g) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{|W|+|Z|}, \quad \mathrm{lda}(g) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathrm{p}(z_1|g) \\ \vdots \\ \mathrm{p}(z_{|Z|}|g) \end{pmatrix} \in \mathbb{R}^{|W|+|Z|}, \quad (4)$$

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. Moreover, contextual information in the recursive subgoal feature is also beneficial to deal with the short goal descriptions. The parameters α, β, and γ are empirically determined on the basis of actual data.

In this paper, we redefine the bag-of-feature vector $\mathrm{bof}(g)$, because we empirically found that contextual information is contained in not only subgoals, but also supergoals (parent goals). To incorporate features of supergoals into $\mathrm{bof}(g)$, we newly define $\mathrm{bof}_{\mathrm{self}}(g)$, a bag-of-features vector extracted only from the target goal g, and $\mathrm{bof}_{\mathrm{cntxt}}(g)$, a contextual bag-of-features vector extacted from subgoals and supergoals.

$$\mathrm{bof}(g) = \frac{1-\gamma(g)}{\|\mathrm{bof}_{\mathrm{self}}(g)\|}\mathrm{bof}_{\mathrm{self}}(g) + \frac{\gamma(g)}{\|\mathrm{bof}_{\mathrm{cntxt}}(g)\|}\mathrm{bof}_{\mathrm{cntxt}}(g), \tag{5}$$

$$\mathrm{bof}_{\mathrm{self}}(g) = \frac{\alpha}{\|\mathrm{tfidf}(g)\|}\mathrm{tfidf}(g) + \frac{\beta}{\|\mathrm{lda}(g)\|}\mathrm{lda}(g), \tag{6}$$

$$\mathrm{bof}_{\mathrm{cntxt}}(g) = \sum_{subg \in \mathrm{sub}(g)} \mathrm{bof}_{\mathrm{sub}}(subg) + \sum_{supg \in \mathrm{sup}(g)} \mathrm{bof}_{\mathrm{sup}}(supg), \tag{7}$$

$$\mathrm{bof}_{\mathrm{sub}}(g) = d_{\mathrm{sub}}\left(\mathrm{bof}_{\mathrm{self}}(g) + \sum_{subg \in \mathrm{sub}(g)} \mathrm{bof}_{\mathrm{sub}}(subg)\right), \tag{8}$$

$$\mathrm{bof}_{\mathrm{sup}}(g) = d_{\mathrm{sup}}\left(\mathrm{bof}_{\mathrm{self}}(g) + \sum_{supg \in \mathrm{sup}(g)} \mathrm{bof}_{\mathrm{sup}}(supg)\right), \tag{9}$$

$$\gamma(g) = \mathit{upper}_{\mathrm{cntxt}} \cdot \tanh(k \cdot \|\mathrm{bof}_{\mathrm{cntxt}}(g)\|), \tag{10}$$

where $\mathrm{sup}(g)$ denotes a set of supergoals of g, d_{sub} and d_{sup} respectively denote decay ratios when recursively tracking subgoals and supergoals, $\mathit{upper}_{\mathrm{cntxt}}$ denotes an upper limit of the weight of $\mathrm{bof}_{\mathrm{cntxt}}(g)$, $\alpha + \beta = 1$, and $0 \leq \alpha$, β, $\mathit{upper}_{\mathrm{cntxt}}$, d_{sub}, $d_{\mathrm{sup}} \leq 1$. The definitions of $\mathrm{tfidf}(g)$ and $\mathrm{lda}(g)$ are not modified from Equation 4. The hyperbolic tangent is used for adjusting $\gamma(g)$, the weight of the contextual bag-of-features $\mathrm{bof}_{\mathrm{cntxt}}(g)$, according to the amount of subgoals and supergoals. Thus, $\gamma(g)$ is 0 when $\|\mathrm{bof}_{\mathrm{cntxt}}(g)\| = 0$ and asymptotically gets close to $\mathit{upper}_{\mathrm{cntxt}}$ along to the increase of $\|\mathrm{bof}_{\mathrm{cntxt}}(g)\|$. The parameters α, β, $\mathit{upper}_{\mathrm{cntxt}}$, d_{sub}, d_{sup}, and k are empirically determined.

In order to recommend similar goals, a pair of goals g_i and g_j satisfying $\mathrm{sim}(g_i, g_j) > \theta_g$ is linked by the property schema:isSimilarTo that is defined by schema.org[10].

In this paper, we consider to recommend not only similar goals but also potential collaborators who aim at similar goals. To recommend potential col-

[10] `http://schema.org/Product`

laborators for a user u, here we formulate the similarity between users u and u_k as follows:

$$\mathrm{sim}(u, u_k) = \frac{1}{|\mathrm{goals}(u)|} \sum_{g \in \mathrm{goals}(u)} \max_{g' \in \mathrm{goals}(u_k)} \mathrm{sim}(g, g') \tag{11}$$

where goals(u) denotes a set of goals that the user u aims at. The user u is linked to users u_k who satisfy $\mathrm{sim}(u, u_k) > \theta_u$ with the property schema:isSimilarTo in order to recommend potential collaborators u_k to the user u.

3.3 System Implementation

We implemented a web platform for linked open data for creating, exploring and discovering public goals. The system is used by agents and decision makers in the society's government and members of the society. Figure 3 shows the structure of the application. It is divided into three parts, a website user interface, a server side API, and a linked open data storage. The website user interface is

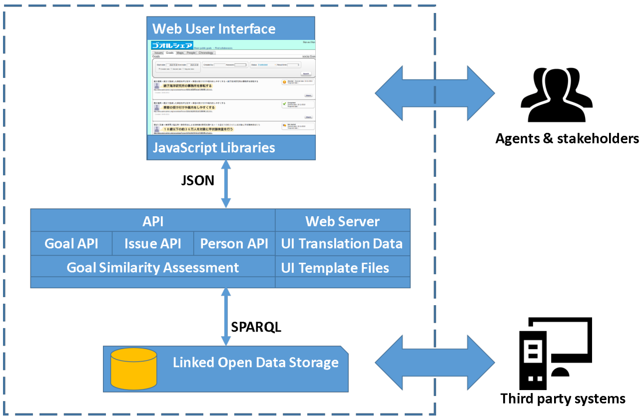

Fig. 3. Structure of the system

implemented with HTML and JavaScript. We utilise an open source JavaScript template library that reuses HTML document object model (DOM) elements to produce dynamic HTML content. In addition, we use a JavaScript library for localising the user interface, the client side localisation helps in producing multilingual dynamic content. Moreover, we utilise Google maps API for displaying spatial information. A requirement for the user interface is to display goal information in an easy-to-explore manner. System:'s usability and clear visualisation is an important design factor, because the system is intended to be used by broad audience with varying IT skills. The user interface design contains visual elements, e.g. agents are presented visually with an avatar image and a goal status is indicated with a presenting icon. The user is required to log in with Facebook authentication, the Facebook account is used to identify the user. In future work, additional login methods will be added. The user information used to identify involved parties in the linked data. The client side application uses asynchronous HTTP requests to access the server side API, responses are in JSON (JavaScript Object Notation) format.

The server side API contains the implementation for searching and creating goals and related entities. The API provides interfaces for fetching the data by making HTTP requests, parameters are transferred with HTTP GET and POST methods. The API verifies the parameter data validity and ensures data consistency in the data storage, for example by verifying that required entities exists before inserting triples to form a relations between them. We utilise SPARQL resource description framework query language for accessing the linked data storage. The API also provides support functions for the user interface, for example it provides data for implementing client side auto completion feature. By handling aforementioned issues in the API, it simplifies the client side solution implementation.

The data is stored in a RDF data storage. We utilised the OpenLink Virtuoso data server[11]. The data is stored in a triple form, i.e. as subject-predicate-object sets. Predicate defines the type of relation between the linked subject and object entities, the entity keys are universal resource locators. The data storage provides a SPARQL endpoint for accessing the data with the RDF query language. The data storage and the SOCIA ontology follows open data principles, both the data and the vocabulary is available for third party applications.

The following RDF/N3 shows an example of RDF triples.

```
<http://collab.open-opinion.org/resource/goal000098>
    rdf:type <http://data.open-opinion.org/socia-ns#Goal>;
    dc:title "Charting the condition of walking ways near Mt. Ikeda"@en;
    socia:subGoal <http://collab.open-opinion.org/resource/goal000102>.
```

3.4 Issue Creation

Here we describe an example of a process of forming issues from public concerns and setting goals to address them. An agent explores internet resources in a search for public concerns, possible channels could be news networks, microblogs, social media, eParticipation and eGovernment systems. When the agent finds a resource that could constitute as an issue, he accesses the goal sharing system with a browser and logs in with Facebook account. Afterwards, the agents navigates to the issues section and opens the issue creation dialog, seen in Figure 5 by clicking a button. The agent inputs the issue details title and description and possibly a spatial location, also he adds the references to the relevant material found in the earlier exploration phase. We utilise the GeoNames ontology for inputting the spatial location to provide a location search with human readable

[11] http://virtuoso.openlinksw.com

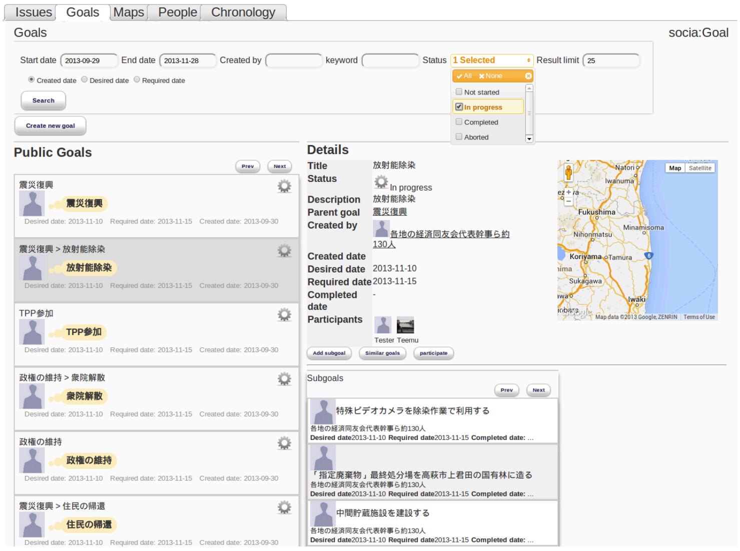

Fig. 4. User interface to interact with a goal

names. The issue generation phase facilitates open innovation, by using the public concerns as source of the issue. In the second phase, a decision maker logs into the system, and begins exploring issues. The user can search issues by filtering the results with creation date, creator, and keyword, the resulting set is shown as a list of issues. The user opens a detail view by clicking an issue in the list. The detail view displays additional information about the issue, e.g. the description and links to the referred material. At this point, the decision makers can debate over the issue, the information provided by the issue details facilitates the decision making by improving the understanding of the issue in question. By clicking the control in the details view, user can create a goal that would be solution for the issue. Subgoals can then be added to the goal to construct a goal hierarchy to achieve more concrete goals.

3.5 Goal Discovery

Here we present an example how the system can facilitate collaboration by the process to discover similar goals. A user begins by accessing the website with a browser, after which he must login into the application. When the user explores goals, seen in Figure 5, he can first input filter options to the filter control. The user has option to filter results by creation date, desired completion date, required completion date, keyword, goal status, and creator. The search result is shown as a list of goals, which displays the basic information about the goal. The basic information contains a title, a description, a status, and a creators' avatar image. The user can open a detail view by selecting a goal by clicking the desired goal in the list. The detail view displays additional information about the goal,

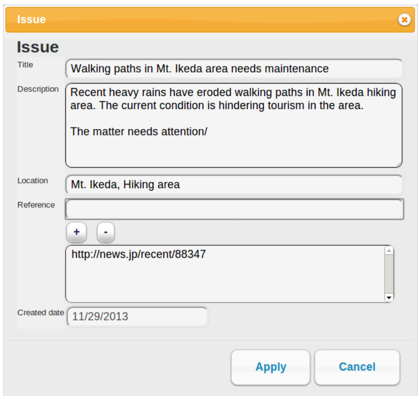

Fig. 5. User interface for adding new issues

like lists of subgoals and participants. Clicking the detail view's "find similar goals" control initiates a new search for goals based on the similarity value. The user can then explore the goal set in a list, which is ordered by descending similarity.

4 Conclusion

In this paper, we present the implemented a web based goal sharing system. The aim of the system is to increase transparency, advance the opportunities to utilise open innovation and eParticipation principles, and to facilitate governmental agents to collaborate. The system offers functions for creating public goals and issues, exploring and discovering similar goals, and visualising the level of similarity. We implemented a method to calculate similarity between goals that utilises the structured goal information. We present the SOCIA vocabulary that we used to structure the goal related information as linked open data. We discuss about two examples of using the system.

Currently, we are arranging workshops in Ogaki city to use the system in real-life situation, for gathering real-life issues and goals. We use the gathered data from the workshops to quantitatively measure, test, and improve the implemented system. One important concern is to determine suitable values for parameters α, β, and γ in the similarity analysis method.

We plan to deploy the web platform to complement existing solutions, like CitiSpe@k[12] and SOCIA linked data storage for facilitating communities to use open innovation in their decision making process and to enable governmental agents to collaborate with other governmental parties and citizens. Our system has a potential to get together concerned citizens and parties that can solve the concrete problem, e.g., CODE for JAPAN[13]. CODE for JAPAN advances the

[12] http://www.open-opinion.org/citispeak/

[13] http://code4japan.org/

cause of open innovation and eParticipation by getting people to provide solutions and tools for local communities and providing information to governmental officials.

A possible future research topic is automatic issue and goal suggestion. After citizens input their issues and goals, we will be able to construct a training corpus for automatic extraction of issues or goals from textual content. Such training data enables us to deal with the novel research topic.

Acknowledgment. This work was supported by the Grant-in-Aid for Young Scientists (B) (No. 25870321) from Japan Society for the Promotion of Science.

References

1. IAP2: IAP2 Spectrum of public participation. http://www.iap2.org/associations/4748/files/Kettering_FINALPreliminaryFindingsReport.pdf (2013)
2. The World Bank: Age dependency ratio, old (% of working-age population). http://data.worldbank.org/indicator/SP.POP.DPND.OL/countries (2013)
3. S 譁, ., Rose, J., Flak, S.: The shape of eparticipation: Characterizing an emerging research area. Government Information Quarterly **25**(3) (2008) 400–428
4. AGI Goldratt Institute: The theory of constraints and its thinking processes - a brief introduction to toc. http://www.goldratt.com/pdfs/toctpwp.pdf (2009)
5. Project Management Institute: A Guide to the Project Management Body of Knowledge. 4th edn. Project Management Institute (2008)
6. Cheng, J., Shigekawa, K., Meguro, K., Yamazaki, F., Nakagawa, I., Hayashi, H., Tamura, K.: Applying the toc logistic process to clarify the problem schemes of near-field earthquake in tokyo metropolitan area. Journal of social safety science (11) (2009) 225–233 (in Japanese).
7. Ohara, M., Kondo, S., Kou, T., Numada, M., Meguro, K.: Overview of social issues after the Great East-Japan Earthquake disaster - part 3 of activity reports of 3.11net Tokyo. SEISAN KENKYU **63**(6) (2011) 749–754 (in Japanese).
8. Shiramatsu, S., Ozono, T., Shintani, T.: Approaches to assessing public concerns: Building linked data for public goals and criteria extracted from textual content. In: Electronic Participation. Proceedings of the 5th IFIP WG 8.5 International Conference, ePart 2013. Volume 8075 of Lecture Notes in Computer Science., Springer (2013) 109–121
9. King, S.F., Brown, P.: Fix my street or else: using the internet to voice local public service concerns. In: Proceedings of the 1st international conference on theory and practice of electronic governance. (2007) 72–80
10. Roth, A.E., Sotomayor, M.A.O.: Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis. Number 18 in Econometric Society Monographs. Cambridge University Press (1990)
11. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. Journal of Machine Learning Research **3** (2003) 993–1022