

公共圏における目標共有のための Linked Data の試作

Prototyping Linked Data for Sharing Public Goals in Public Sphere

白松 俊*¹ 大囿 忠親*¹ 新谷 虎松*¹
 Shun Shiramatsu Tadachika Ozono Toramatsu Shintani

*¹名古屋工業大学 大学院工学研究科 情報工学専攻

Department of Computer Science and Technology, Graduate School of Engineering, Nagoya Institute of Technology

SOCIA is a linked open data set for citizen participation, which contains background information about regional social issues. We aim to apply it to promoting public collaboration, e.g., revitalization from Great East Japan Earthquake. To support public collaboration, we introduced some classes and properties for sharing stakeholders' goals in public sphere. Moreover, we built 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.

1. はじめに

日本の地域社会は現在、複雑に絡み合った多様な問題（災害リスク増大、インフラ老朽化、少子高齢化、過疎化、世代間格差、etc.）に直面している。その地域特有の多様な問題に対処するには、公的討論を通じた住民の参画が非常に重要である。特に東日本大震災の被災地は、震災被害に起因する様々な問題を抱えており、行政だけでなく地域住民自身が積極的に参画できるような体制作りが重要視されている [鈴木 12]。

我々はこれまで、住民参画の間口を広げる情報共有基盤を目指し、Web上の地域関連コンテンツ（ニュース記事、マイクロブログ、議会議事録等）を構造化した Linked Open Data (LOD) である SOCIA (Social Opinions and Concerns for Ideal Argumentation) を構築してきた [Shiramatsu 12]。本稿ではこれを更に発展させ、震災復興のための組織横断的な協働を促進すべく、公共圏で目標を共有するための LOD を試作する。特に、似た目標を持つ主体同士をマッチングするサービスを想定すると、合意形成や協働にあたって問題となりやすい「総論賛成・各論反対」のような齟齬に対処する必要がある。そのためには、目標を部分目標に細分化することによって得られる階層構造 [白松 12] が重要であると考え、この階層構造と自然言語の目標記述を併用した類似度計算手法を検討する。

2. 関連動向: 住民参画と LOD

住民参画とは、公的討論等の対話的なプロセスによって住民が行政の政策決定過程に関わることを指す。参画の前提となる透明性を確保するためには、行政による公共データの公開が重要な役割を担う。有用な気付きや知見を持つ住民がその問題背景を十分に把握しているとは限らず、参画する上での心理的な障壁が存在するので、オープンデータやその可視化サービスによって背景情報を提供すれば、参画の障壁を下げて間口を広げる効果が期待できる。

図 1 は、住民参画の深さのレベルと、想定されるオープンデータの有効範囲を示している。図中の Spectrum of Pub-

lic Participation は、国際市民参画協会 (IAP2; International Association of Public Participation) が住民参画の深さのレベルを 5 段階で表したものである [IAP2 07]。また図中の三原則とは、2009 年の米オバマ政権発足時に全省庁の長官宛ての覚書で掲げられたものであり [Obama 09]、これがその後、米行政機関のデータポータル Data.gov*¹ の整備にも繋がっている。オープンデータによる情報公開は、特に図中左側の透明性や情報共有のために重要な役割を果たす。より深いレベルの参画や協働においても有用と考えられるが、組織を超えて協働する場合は、RDF 等で意味的にリンクされていないオープンデータよりも、意味的リンクによってコンテキストを提供している LOD の方が、より有効にデータを活用できると期待される。特に、組織横断的に複数のデータをマッシュアップして新たな傾向を発見したり、協力関係を構築していく際に、LOD によるコンテキスト共有が効果的に働くと考えられる。

電子政府や住民参画に関連するオントロジーとして、様々な用途のものが提案されてきた [Ouchetto 12]。その 1 つである eGTPM [Sarantis 10] は、電子政府に関連するプロジェクト管理のために設計されたオントロジーである。他にも、住民参画支援に利用可能なボキャブラリ Participation Schema [Styles 08] や、興味を似た主体間の協働支援に利用可能なボキャブラリ Weighted Interests Vocabulary [Brickley 10] がある。これらの従来研究に対する本研究の独自性は、興味が異なる主体同士

IAP2 (国際市民参画協会) による Spectrum of Public Participation

Inform (情報共有)	Consult (意見収集)	Involve (関与)	Collaborate (協働)	Empower (決定権)
------------------	-------------------	-----------------	---------------------	------------------

米オバマ政権の OGI (Open Government Initiative) による三原則

Transparency (透明性)	Participation (参画)	Collaboration (協働)
-----------------------	-----------------------	-----------------------

公的なインパクト



図 1: 住民参画の深さのレベルとオープンデータの有効範囲

連絡先: 白松 俊, 名古屋工業大学, 〒466-8555 名古屋市昭和区御器所町名古屋工業大学つくり領域, 052-735-5129, siramatu@nitech.ac.jp

*¹ <http://www.data.gov/>

が共通の目標に向かって協力することが重要であるという観点の下、公共圏での目標共有に焦点を当てている点である。

住民参画を志向したサービスとして米国で運用されている Neighborland*2[Chang 11]では、「I want [欲しいもの] in [地区名]」という定型文を用いてユーザ自身が希望を入力し、他ユーザのレイティングによって賛同の意志を共有できる。これは本研究の目指す目標共有と近いサービスであるが、本研究が目指す地域や組織を超えた協働のためには、Neighborland がサポートしていない類似目標のマッチング機構や、部分目標への分割による作業分担の支援機構が必要と考える。

3. 目標の階層構造のためのオントロジー

公共圏での協力関係構築や合意形成は、震災復興にあたって必ず必要になるプロセスである。公的な目標に関するデータセットの共有は、類似する目標を持つ主体同士が協力するにあたって有用であると考えられる。ただし、類似した大目標を持っている2主体が「総論賛成・各論反対」となり細部で衝突する場合も多い。細部の相違点を明らかにして妥協点を探るためには、大目標を部分目標に分割し階層化する必要がある。また、抽象的な大目標に貢献しようと思うと心理的な障壁が高いが、細分化して具体的な部分目標にすると細かい作業単位の分担だけで貢献できるようになり、参画の障壁が低くなるという利点もある。我々は、SOCIA を拡張して上記のような目標の階層構造を LOD 化することにより、類似した目標を持つ主体同士をマッチングする「目標マッチングサービス」の開発を目指している。そのために、従来の SOCIA オントロジーを拡張し、図2のようなクラスやプロパティを追加する。特に、目標とそれを細分化した部分目標との関係を表すプロパティ `socia:subgoal` を導入することにより、目標の階層構造が記述可能になる。

この階層構造は、複数の目標間の類似度計算を洗練化するためにも有用である。分割した部分目標を持たず、短い自然言語文から成る記述 (`dc:description`) のみを持つような目標の場合、他の目標との類似度計算が難しくなり、高い再現率での類似目標検索が不可能になるだろう。我々は、分野や組織を超えた協働を促進するような目標マッチングサービスを目指しており、そのためには高い再現率での類似目標検索が可能な類

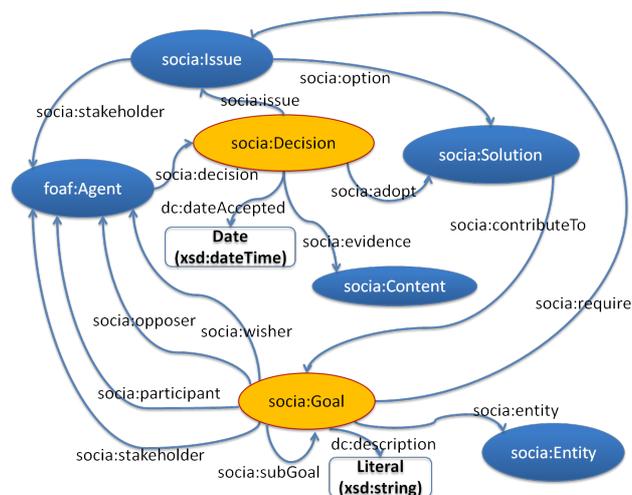


図 2: 目標や意思決定に関するクラス群

*2 <https://neighborland.com/>

似度計算手法が必要である。そのようなサービスにおいては、SNS 上で目標を持つ住民や主体自身によって目標が入力されることが望ましいが、本稿では予備調査のため、震災復興に関連する Web 上の資料から人手で目標を抽出することによって復興目標の LOD を試作する。

4. 関連資料からの人手による目標抽出

Yahoo!ニュース*3の震災復興に関連する 96 記事と、「原発事故子ども・被災者支援法」について各地の弁護士会が Web 上で公開している 2 文書 [福島 12, 大阪 12] から、657 の復興目標および 4349 の RDF トリプルをアノテーション作業員 1 名が人手で抽出し、復興目標に関する LOD データセット*4 を試作した。最も抽象的な「震災復興」という大目標*5がルートノードとなり、そこから `socia:subgoal` プロパティで具体的な部分目標が繋がっていくことにより、階層構造を構成している。ルートノード「震災復興」には 36 の部分目標があり、例えばその一つに「東北に観光客を誘致する」があり、そのまた部分目標である「新たな旅行商品を作る」は図3のようなインスタンスとして表されている。「新たな旅行商品を作る」も更に2つの部分目標「大規模な研修を行う」「参加各社は視察後、12月までに東北旅行の企画案をまとめる」に分割されており、`socia:subgoal` プロパティで記述されている。

本試行では、別の記事から抽出された目標であっても同一の目標とアノテーション作業員が判断した場合は、それらを1つのリソースとしてまとめるという方針を取った。ただし、異なる地域のそれぞれについて同一の目標があった場合についての指針は与えていなかったため、例えば「大熊町の避難住民の帰還」と「南相馬市の住民の帰還」のような目標は並列に作成され、それらの間に関係は付与されなかった。このような構造については、中間ノードとしての「住民の帰還」という目標をボトムアップに作成する作業の支援機構が必要と考えられる。このような中間ノードが、地域間での協働や、他地域の事例を参考にする上で有効に機能すると期待されるためである。ただし、いきなり中間ノードの作成を自動推薦するのは困難である

http://data.open-opinion.org/socia/data/Goal/新たな旅行商品を作る	
Class (rdf:type)	http://data.open-opinion.org/socia/ns#Goal
Title (dc:title)	新たな旅行商品を作る
Description (dc:description)	復興支援のため、東北6県の新たな旅行商品を作る
Reference (dc:references)	http://headlines.yahoo.co.jp/hl?a=20121121-00000010-khks-bus_all
Wisher (socia:wisher)	http://data.open-opinion.org/socia/data/Person/日本旅行業協会加盟社
Wisher (socia:wisher)	http://ja.dbpedia.org/resource/日本旅行業協会
Sub Goal (socia:subGoal)	http://data.open-opinion.org/socia/data/Goal/大規模な研修を行う
Sub Goal (socia:subGoal)	http://data.open-opinion.org/socia/data/Goal/参加各社は視察後、12月までに東北旅行の企画案をまとめる
Preparation (socia:preparation)	http://data.open-opinion.org/socia/data/Goal/被災地のボランティア活動を行う
Preparation (socia:preparation)	http://data.open-opinion.org/socia/data/Goal/東北の観光関係者との交流会を行う

図 3: 記事から抽出した復興目標の例

*3 <http://headlines.yahoo.co.jp/>

*4 <http://data.open-opinion.org/socia/data/Goal?rdf:type=socia:Goal&limit=100>

*5 <http://data.open-opinion.org/socia/data/Goal/%E9%9C%87%E7%81%BD%E5%BE%A9%E8%88%88>

ため、まずは前段階として類似目標へのリンクを付与し、類似目標が多い目標クラスが検出された段階で中間ノードの作成を推薦するといった手順が考えられる。当然、類似目標へのリンク付与作業についても支援機構が必要であり、そのために目標間の類似度計算手法が必要である。データ数が多くなると、このような作業を支援機構無しで作成するのは非常に困難であるため、支援機構の開発は必要不可欠な課題である。

今後は、目標マッチングサービスのための類似度計算手法を検討する際のテストに本データセットを利用しつつ、中間ノード作成支援機能を開発する予定である。中間ノード作成支援機能の開発にあたっては、目標の階層構造の可視化も重要である。特に、可視化すべき範囲をどのように決定するかについて、今後検討していく必要がある。また、目標マッチングサービスの開発と並行して、中間ノード作成支援機能を用いた目標の階層構造を洗練化する被験者実験も計画している。

5. 階層構造と目標記述を併用した類似度計算

短い自然言語文から成る記述 (`dc:description` プロパティ) と、目標の階層構造 `socia:subgoal` プロパティを併用した類似度は、Bag-of-feature ベクトルを再帰的に定義することにより、以下のようなコサイン類似度として計算する。なお、 $\text{sim}(\cdot, \cdot)$ は目標間の類似度、 g は公的な目標、 $\text{bof}(g)$ は g の bag-of-feature ベクトル、 $\text{sub}(g)$ は g の部分目標の集合、 $w \in W$ は形態素 N -gram、 $z \in Z$ は LDA (Latent Dirichlet Allocation)[Blei 03] 等のトピックモデルで得られる潜在トピック、 $\text{tfidf}(w, g)$ は g を 1 文書と見なした w の TF-IDF 値であり、条件 $0 \leq \alpha, \beta, \gamma \leq 1$ と $\alpha + \beta + \gamma = 1$ を満たすものとする。

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

$$\begin{aligned} \text{bof}(g) &= \frac{\alpha}{\|\text{tfidf}(g)\|} \text{tfidf}(g) + \frac{\beta}{\|\text{lدا}(g)\|} \text{lدا}(g) \\ &+ \frac{\gamma}{|\text{sub}(g)|} \sum_{sg \in \text{sub}(g)} \frac{\text{bof}(sg)}{\|\text{bof}(sg)\|} \end{aligned} \quad (2)$$

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

$$\text{lدا}(g) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ p(z_1|g) \\ \vdots \\ p(z_{|Z|}|g) \end{pmatrix} \in \mathbb{R}^{|W|+|Z|} \quad (4)$$

ここで、部分目標の階層構造 $\text{sub}(g)$ だけでなく潜在トピック z も併用している意図は、短い自然言語の目標記述に含まれる語彙の少なさをカバーするためである。今後は、本稿で試作した復興目標の LOD データセットを用いた評価実験を通じ、経験的に最適なパラメータ α, β, γ を決定することで、式 (2) の各項の必要性を検討していく必要がある。

6. まとめ

本稿では、これまでに構築してきた住民参画のための LOD データセット SOCIA を拡張し、住民間の協働を促進する目標マッチングサービスを開発すべく、公共圏における目標共有のための LOD データセットを試作した。協働支援には部分目標への細分化による階層構造が重要であると考え、目標の階層構造を記述可能にした。具体例として、東日本大震災からの復興目標に着目し、関連資料から 657 の復興目標とそれに付随する 4349 の RDF トリプルを人手で抽出した。その過程で、異なる地域を対象とするが同一と見なせる目標に着目し、類似目標へのリンク付与作業や中間ノードの作成作業を支援する機構の必要性について述べた。さらに、短い自然言語の目標記述だけでなく、目標の階層構造や潜在トピックも併用した類似度計算手法を示した。

今後の目標マッチングサービスの開発に向けて、まずは本稿で試作したデータセットを用いた類似度計算手法の評価実験を行い、上記のリンク付与作業や中間ノード作成作業の支援機構を開発していく予定である。

謝辞

本研究は、JST 復興促進プログラム (A-STEP) (No. 241FT0304)、科学研究費補助金若手研究 (B) (No. 25870321)、および、総務省 SCOPE 地域 ICT 振興型研究開発プログラムの支援を受けたものです。

参考文献

- [Blei 03] Blei, D. M., Ng, A. Y., and Jordan, M. I.: Latent Dirichlet Allocation, *Journal of Machine Learning Research*, Vol. 3, pp. 993–1022 (2003)
- [Brickley 10] Brickley, D., Miller, L., Inkster, T., Zeng, Y., Wang, Y., Damljanić, D., Huang, Z., Kinsella, S., Breslin, J., and Ferris, B.: The Weighted Interests Vocabulary 0.5, <http://purl.org/ontology/wi/core> (2010)
- [Chang 11] Chang, C.: Neighborland, <http://candychang.com/neighborland/> (2011)
- [IAP2 07] IAP2, : IAP2 Spectrum of Public Participation, http://www.iap2.org/associations/4748/files/IAP2%20Spectrum_vertical.pdf (2007)
- [Obama 09] Obama, B.: Transparency and Open Government, Memorandum for the Heads of Executive Departments and Agencies, http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment (2009)
- [Ouchetto 12] Ouchetto, H., Ouchetto, O., and Roudies, O.: Ontology-oriented e-Gov Services Retrieval, *International Journal of Computer Science Issues*, Vol. 9, No. 3 (2012)
- [Sarantis 10] Sarantis, D. and Askounis, D.: Knowledge Exploitation via Ontology Development in e-Government Project Management, *International Journal of Digital Society*, Vol. 1, No. 4, pp. 246–255 (2010)

- [Shiramatsu 12] Shiramatsu, S., Swezey, R., Sano, H., Hirata, N., Ozono, T., and 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*, Vol. 7444 of *Lecture Notes in Computer Science*, pp. 73–84, Springer (2012)
- [Styles 08] Styles, R., Wallace, C., and Moeller, K.: Participation Schema, <http://purl.org/vocab/participation/schema> (2008)
- [大阪 12] 大阪弁護士会：原発事故子ども・被災者支援法第 5 条に定められた「基本方針」の策定に関する意見書, http://www.osakaben.or.jp/web/03_speak/kanri/db/info/2012/2012_50a4700f8e2d1_0.pdf (2012)
- [白松 12] 白松俊, 平田紀史, Swezey, R. M. E., 佐野博之, 大園忠親, 新谷虎松: LOD データセット SOCIA の復興促進への適用に向けた検討, 人工知能学会第 28 回セマンティックウェブとオントロジー研究会, SIG-SWO-A1202-05 (2012)
- [福島 12] 福島弁護士会：原発事故子ども・被災者支援法に基づき求められる施策に関する基調報告書, <http://business3.plala.or.jp/fba/topics/2012/topix121015.pdf> (2012)
- [鈴木 12] 鈴木孝男: 復興まちづくり推進員を介した地域コミュニティ再生 (連載東日本大震災 | 連続ルポ 1 | 動き出す被災地), Vol. 127, No. 1629, pp. 2–3 (2012)