

統計分析によるシナリオの妥当性評価の試み

山田 菊子¹

¹正会員 東京工業大学 研究員 大学院理工学研究科 土木工学専攻 (〒152-8552 東京都目黒区大岡山 2-12-1-M1-20)

E-mail: kiko.yamada@plan.cv.titech.ac.jp

プロジェクトや政策立案の場面において、有識者や住民等の関係主体の意見を把握することの重要性が認識され、ワークショップやPI等が実施されている。これらの場面では、発言や議論の内容を計画策定者が把握し策定に役立てられなければ、実施の意義は半減する。著者らは発言を計画策定に役立つ形にすることを目的として、発言内容を情報システム開発で用いられるシナリオに記述する手法を提案した。シナリオは、計画に登場するステークホルダーの物語である。有識者の発言記録の分析に適用し、計画策定者の理解を助けることを確認した。またシナリオの質の評価、分析者に必要なスキルや計画プロセスでの位置づけなどの課題を把握した。そこで本研究では、これらの課題のうちシナリオの質の評価について取り組む。発言記録と作成したシナリオを語の出現頻度等の統計指標より分析し、シナリオの妥当性の評価に寄与する指標を検討するものである。

Key Words : *human-centered design, scenario, interview, long-term plan, evaluation*

1. はじめに

プロジェクトや政策立案の場面で、有識者やステークホルダ（利害関係者）の関与を得る場面は多くある。まちづくり計画立案のワークショップなどでは、参加者の多様な意見をお互いに反映しながら一定の合意や、意識の変化を得ようとするものである。ワークショップ実施中に作成するメモや図、付箋紙と模造紙に示された意見の集約結果、さらには発言記録などが得られる。一方、インタビュー、会議においては、参加者による出力（図、インタビュー記事、発言者によるレジメ等）に加え、発言記録も得られる。

特に、録音を元に起こした、より生データに近い発言記録は豊富なコンテンツを含むと考えられるが、文字数が膨大であることもあって、そのまま計画立案に用いることは難しい。現場では、調査者や計画策定者が要点を箇条書きにしたり、大幅に集約するなどして、計画に用いられている。

このような発言録を分析する研究には、発言録そのものの分析を行なうものがある。鄭ら¹⁾は、公的討議における第三者委員会の発言録の各文を、ファセット (facet) と呼ぶ4つの項目の評価値から統計的機械学習モデルにより分類する方法を提案し、河川会議の議事録に適用して、参加者間の認識の不一致や意見の対立を把握できることを示した。

丸石ら²⁾は、ワークショップにおける参加者の意識の変化に注目して、数値化するとともに、発言の類似度を、語を要素とする文のベクトルの内積と定義した。

また、榊原ら³⁾は、参加者全体がワークショップにおける発言を共有していると想定し、議論の進行に伴う話題の変遷を、語の出現頻度の因子分析により把握する方法を提案した。

さらに難波ら⁴⁾は、討議の記録に出現する語の頻度から特徴語を抽出し、意見や発言者の推移を可視化した。

一方、著者らは、発言記録の内容の把握の目的が、計画策定者の策定する計画に役立てることにあるとの前提に立ち、発言記録を「計画のステークホルダー」を主語とするシナリオに記述する方法を提案した。提案した手法はこれまでに、豪雨災害における水道事業の記録⁵⁾や、土木学会が設置した有識者会議の記録⁶⁾⁷⁾の分析に適用し、後者では計画立案への有用性を検討した。

そこで本研究では、これらの先行研究を参照しつつ、著者らが提案するシナリオと、原記録を統計的指標により比較し、シナリオの妥当性を把握することを試みる。発言者に参照することなく、分析者が記述したシナリオの妥当性を一定程度確認できる方法を提供することは、シナリオを利用した設計や分析プロセスの効率化に寄与できると考えるものである。

2. シナリオを用いる分析手法

シナリオを用いる分析法は、発言者による発言やレジメ、インタビュー記録などのテキストデータを分析し、分析対象の事象のステークホルダーの、時期ごとの行動のシナリオとして記述するものである。

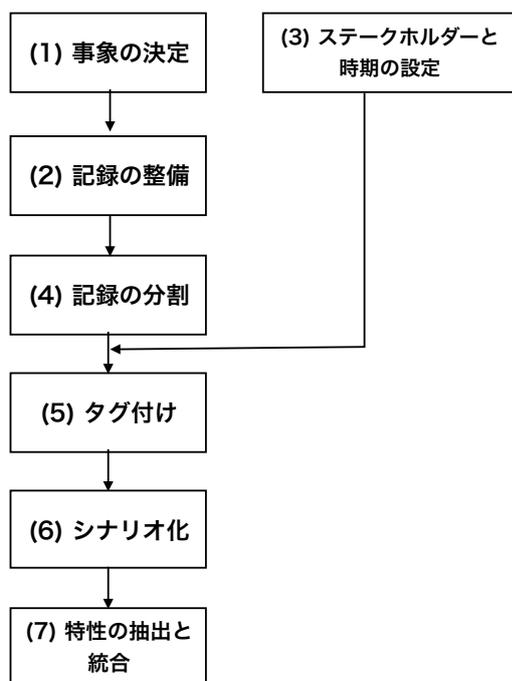


図-1 分析手順

以下、著者らの既往研究⁶⁾をもとに手法の概要を示す(図-1)。

(1) 事象の決定

分析の対象とする事象を決定する。本研究では、「豪雨災害時の水道事業者の対応」や「土木学会の中長期計画」である。既に発生した事象や、プロジェクトや計画の対象が考えられる。

(2) 記録の整備

発言者の発言や意思、意向が表出した資料を収集し、テキスト化する。具体的には、発言を録音したデータ、インタビュー記事、発言者が執筆した文章などがこれにあたり、本研究では「発言録」と呼ぶ。合わせて、発言者と対象とする事象との関係を示す背景情報も収集する。

(3) ステークホルダーと時期の設定

対象とする事象に登場するステークホルダー(利害関係者)と、時期を設定する。すでに発生した事象であれば、事象の展開の段階である。また、計画であれば、計画対象時期と、必要に応じてその前後である。

(4) 記録の分割

テキスト化した発言録を、意味を失わない程度に細分化する。主語と述語を持つ文より小さな単位には分割しないが、複数の文章からなる程度に集約すること

もある。

(5) タグ付け

分割した記録にステークホルダーと時期のタグを振る。一つの記録に複数のステークホルダー、時期のタグが付与されることもある。

(6) シナリオ化

ステークホルダーと時期の組み合わせごとに記録を集約し、ステークホルダーを主語とする物語であるシナリオに記述する。郷ら⁸⁾が示すシナリオの4要素のうち、特に、「アクター」、すなわちステークホルダーと、「アクションとイベントの列」を明記する。

(7) 特性の抽出と構造化

すべてのシナリオからキーワードを抽出し、特性の構造図を作成する。キーワードを付箋紙に書き出し、これをワークショップなどにより分類する。

3. シナリオ評価の試み

(1) 目的

シナリオは分析者が記述する。これまでに実施した事例研究の一部では、発言者や、インタビュー実施者にシナリオや特性の構造図などの生成物を確認させることでシナリオの妥当性の確保をめざした。しかし、組織の戦略分析を目的として、雑誌に発表されたインタビュー記事を発言録とするような場合、発言者、インタビュー実施者のいずれに確認させることは困難である。また、発言者の関与が、第三者としての客観的な分析に影響を与える場合も考えられる。

このような場合に、発言者や、記録者の関与を得ることなく、シナリオの妥当性を一定程度、把握する指標が必要となる。

そこで本節では、2つの事例研究を対象として、簡便な統計分析を用いたシナリオの妥当性の検討を試みる。ここでは、原データである発言録と、分析者が記述したシナリオに頻繁に出現する語の傾向に類似性があることを、妥当性の指標とする。

(2) 妥当性の仮定

シナリオには、原データである発言録の内容を過不足なく反映することが求められる。しかしながら、発言録が音声データをテキスト化したものである場合には、同じ内容の繰り返しや、接続詞、接尾語や、指示代名詞等が多く含まれる。また、発言録と発言者による執筆文書を同時に分析する場合にも同様に、同じ内容の繰り返しが出現する。後者の場合、分析者がすべ

ての資料に目を通した上で、シナリオに記述することで、冗長性を削減することができるのが、本手法のメリットである。

そこで、シナリオが発言録を一定程度妥当に反映していることの指標を、名詞の語彙の出現頻度に求めることとする。

(3) 分析手順

発言録とシナリオ、それぞれから名詞のみを抽出し、出現頻度を比較する。手順の概略を示す。

a) 名詞の抽出

それぞれのデータを語単位に分割するために、形態素解析ツール“MeCab”を使用する。“MeCab”において名詞を指定して抽出し、出現頻度を計算する。

b) データの調整

その後、手作業により、次に該当するものを排除する。

記号、数字、代名詞、数詞、頭文字、接頭語、接尾語、助詞、副詞、助動詞、接続詞、単位、単漢字

また、表記の揺れを統一する。本事例での例には次のものがある。

- 「シヴィル・エンジニアリング」「Civil Engineernt」
→ 「シビル・エンジニアリング」
- 「わたくし」→ 「私」
- 「メーカー」→ 「メーカー」

c) 累積出現割合の算定

すべての発言記録と、すべてのシナリオについて、発言記録に出現する名詞の上位から累積出現割合計算し比較する。

(4) データの概要

本研究では、土木学会の有識者会議の発言録、水道事業における災害対応の記録の二つについて、発言録と記述したシナリオの比較を行なった(表-1)。

a) 有識者会議の発言記録

土木学会が2011年12月に開催した「有識者会議」の記録を用いる。有識者会議は「厳しい社会環境の中で土木技術者の果たすべき役割」をテーマとし、10名の委員が事前に作成した文書と、議長を含む11名の議論の音声データを書き起こした発言記録を分析の対象とした。議論の結果を次の土木学会の中長期計画に役立てることを目的として、シナリオによる分析を試みたものである⁶⁾。

発言者11名ごとに、4名のステークホルダーの、4つの時期の97のシナリオが記述された。

テキスト化された録音データと、発言者が事前に執筆した文書からなる発言録が28,862字、分析者のシナ

リオが16,964字である。

b) 水道事業における災害対応記録

水道事業における災害対応の記録と、その記録をもとに記述したステークホルダーのシナリオを分析対象とする。

災害対応の記録は、北海道石狩市において2010年8月に発生した集中豪雨の際の、市の水道事業担当職員であった清野⁹⁾による執筆論文である。このうち、序論、考察等を除く災害対応に関する記述のみを発言録として、関係者をステークホルダーとするシナリオに記述した(山田⁵⁾)。災害時に断水回避にいたった理由を、ステークホルダーの行動とその動機(論文中では「ビジネスドライバ」と呼んだ)にあるとし、シナリオに記述し分析したものである。

市民を含む11名のステークホルダーの背景情報を把握し、災害対応の2日間の行動を10のシナリオに記述した。なお、シナリオには、ステークホルダーの背景情報を含む。背景情報は、記録の執筆者と一部のステークホルダーに対して補足的に追加したインタビューによる。

清野の論文から抜粋した災害対応に関する記述は3,500字、シナリオが3,061字である(表-1)。

(5) 分析結果

a) 有識者会議発言記録

分析の対象となる名詞の出現回数等を表-2に示す。出現する名詞は発言記録で1,162語、シナリオでは660語と、シナリオに記述することで半減した。この割合は、文字数の割合にほぼ一致する。すべての語の出現回数は、発言記録で4,414回、シナリオで3,212回で発言記録の72.7%に相当する。もっとも多く出現する語は、発言録、シナリオとも「土木」であるが、発言録では182回、シナリオでは342回と、シナリオが2倍近い。これは、発言では明示として省略される語も、シナリオにする際には、主語や述語の省略を極力避けるためであると考えられる。

発言記録の66位まで、シナリオの99位までの語で出現回数の50%を越える。

表-2 出現する語の回数(有識者会議記録)

データ	文字数 (字)	出現語数 (語)	出現回数 (回)	最多出現語 の出現回数 (回)	50%を越える語 の順位
発言録	28,862	1,162	4,414	182	66位
シナリオ	16,964	660	3,211	342	99位
シナリオ /発言録	58.8%	56.8%	72.7%	187.9%	-

注：最多出現回数はそれぞれでの記録での回数である。また、50%を越える語を計算する際の順序は発言録に出現する語数の降順である。

表-1 使用データ

	発言録			シナリオ				
	記録の対象	記録の種類	文字数	分析の目的	ステークホルダー数	時期	シナリオの数	文字数
有識者会議 発言録	土木学会の有識者会議における発言	(1) 発言記録 (2) 発言者による執筆文書	28,862	将来計画への反映	4	4	97	16,964
災害対応記録	水道事業の災害対応	対応記録をまとめた論文	3,500	災害対応の実現の原因分析	11	1	10	3,061

表-3 出現する語の回数（災害対応記録）

データ	文字数 (字)	出現語数 (語)	出現回数 (回)	最多出現語 の出現回数 (回)	50%を越える語 の順位
発言録	3,500	326	835	25	53位
シナリオ	3,061	209	911	21	47位
シナリオ /発言録	87.5%	64.1%	109.1%	84.0%	-

注：最多出現回数はそれぞれの記録での回数である。また、50%を越える語を計算する際の順序は発言録に出現する語数の降順である。

b) 災害対応記録

分析の対象となる名詞の出現回数等を表-3に示す。出現する名詞は、発言記録で326語、シナリオでは209語、発言記録の64.1%と、集約された。すべての語の出現回数は、発言記録で835回、シナリオで911回で発言記録の109.1%に相当する。多く出現する語は、発言録では「職員」「管」「時」「導水」「作業」などである。一方、シナリオでは「日」「目」「場」「浄水」「管」「導水」「水道」「職員」などである。シナリオで上位であった「日」「目」「場」などは、時間を表す「1日目」、場所を示す「浄水場」等をシナリオで明記したことにより増加したものと考えられる。

発言記録の53位まで、シナリオの47位までの語で出現回数の50%を越える。

4. 考察

本研究ではシナリオの妥当性の評価を目的として、形態素解析を用いた語の出現頻度を分析した。

有識者会議の発言記録と、水道事業の災害対応記録を比較した結果、いずれの場合にも出現語数が集約されるにも関わらず、語の出現回数は増加することを把握した。これには二つの理由が考えられる。

- 分析者がシナリオに記述する際に、テーマに関わりのない事項を排除すること
- 特に話し言葉を発言録とする場合、主語や目的語を積極的に補うこと

本研究で取り上げた事例研究は、有識者会議の発言録は音声データをもとに書き起こした記録と、発言者

の執筆した文書であり、水道事業の災害対応の記録は、発言者が執筆した文書である。このように、入力となる発言録の種類が一致していないため、一概に比較はできないものの、いずれの場合にも、出現語数が集約され、さらに出現回数が増加することが把握された。

5. 結論

本研究では、まず、発言録から作成するシナリオの妥当性の評価の必要性について概観し、その後、二つの事例研究について、出現語の語数を指標とする分析を行なった。その結果、発言録の種類が異なっても、シナリオは、出現語数を集約し、さらに、出現回数を増幅することを把握した。

製品や情報システム設計における「シナリオベースデザイン (scenario-based design)」においては、「シナリオ」は、設計の各段階の中間成果物として定義される。Rosson, Carrollが提案したこのフレームワークでは、(1) 分析段階の「問題シナリオ」、(2) 設計段階の「活動シナリオ」「情報シナリオ」「インタラクションシナリオ」があり、(3) プロトタイプと評価段階では、作成したプロトタイプが(2)で定義したシナリオを実現したかについて評価する¹⁰⁾。

ワークショップや公的討議の要約を、シナリオとして表現する場合、本研究で取り上げたケースのうちの有識者会議は、将来のあるべき姿を取り上げた「活動シナリオ」である一方、災害時の水道のケースは、現在の活動の問題点を把握するための「問題シナリオ」となる。異なる段階で用いられるシナリオの評価指標は異なる可能性がある。今後の課題の一つとしたい。

謝辞：事例研究のうち有識者会議の発言録の分析は土木学会企画委員会の2012年度の活動として実施した。また本稿の執筆に際しては東京工業大学 朝倉康夫氏に指導をいただいた。御礼を申し上げる。

参考文献

- 1) 鄭蝦榮, 小林潔司, 羽鳥剛史: ファセット分解と公的討議のプロトコル分析, 土木計画学研究・講演集, Vol. 40, No. 276, 土木学会, 2009.
- 2) 丸石浩一, 佐々木邦明: ワークショップにおける討議内容の数値化と視覚化の試み, 土木計画学研究・講演集, Vol. 38, No. 119, 土木学会, 2008.
- 3) 榊原弘之, 長曾我部まどか: テキスト分析を通じたワークショップ討議の評価手法に関する研究, 土木計画学研究・講演集, Vol. 42, No. 15, 土木学会, 2010.
- 4) 難波雄二, 塚井誠人, 桑野将司: 文脈マイニングモデルを用いた討議過程の可視化手法に関する研究, 土木学会論文集 D3, Vol. 67, No. 5, pp. I209-219, 2011.
- 5) 山田菊子: 人間中心アプローチによる事業の分析と戦略立案の可能性～水道事業における災害対応を事例として～, 商学討究, Vol. 62, No. 2/3, pp. 149-166, 2011.
- 6) 山田菊子, 榎木洋子: シナリオを用いた発言要約に関する研究, 土木計画学研究発表会・講演集, Vol. 48, 土木学会, 2013.
- 7) 山田菊子: シナリオを用いた発言要約の課題に関する考察, 平成 25 年度土木学会北海道支部論文報告集, Vol. 70, No. D-27, 土木学会, 2014.
- 8) 郷健太郎, Carroll, John M., 今宮淳美: ユーザーの視点を取り入れる技術: システム開発におけるシナリオの役割, 情報処理, Vol. 41, No. 1, pp. 82-87, 2000.
- 9) 清野馨: 石狩市における断水回避顛末記, 水道公論, Vol. 47, No. 2, pp. 32-42, 2011.
- 10) 郷健太郎: シナリオベースデザイン, ユーザビリティハンドブック, 共立出版, p. 465, 2007.

(2014.04.25 受付)

STATISTICAL ANALYSIS OF SCENARIOS IN LONGTERM PLANNING

Kiko YAMADA-KAWAI

Opinions expressed by both professionals and public are thought to be an essential aspect to draw more efficient and agreed plans. The author has proposed a method to summarize opinions into stakeholders' scenarios that well suit to the framework of target plans. In this study, frequency data of words appears in the original records and scenarios to derive information in evaluating validity of the generated scenarios. Two case studies were conducted to find that variety of words decreases, but frequency increases in the scenarios. The case studies included audiotyped discussion record by the professionals for longer-term plan for an academic society, and a report on a water supplier at a disaster.