

原著論文

テキストマイニングから見た政令指定都市の自治体総合計画 ～浜松市・岡山市・新潟市のスマートシティ構想パターン化～

○平松 隆志, 遊橋 裕泰
静岡大学創造科学技術大学院

A Study on Text Mining of Local Government Comprehensive Plans: Patterning of Smart City Concepts

○Takashi HIRAMATSU, Hiroyasu YUHASHI
Graduate School of Science and Technology, Shizuoka University

概要: 日本の「高齢化」「東京一極集中」などの社会課題に対してスマートシティ構想に活路を見出そうとする自治体が増えている。国はデジタル田園都市国家構想総合戦略の実証等で先行事例の創出を行うも実装にはまだ課題が多いとされている。スマートシティに関する事例や調査資料など関連文書は増えてきているが、其々の地域課題に対しての解決には、それら関連文書を読み解き実装計画を策定する必要がある。本研究では、スマートシティ普及段階における行政文書のスマートシティ構想(実装計画)をテキストマイニングで解析を行い、浜松市、岡山市、新潟市の比較検討からスマートシティ構想のパターンモデルを導出した。

キーワード: 自治体総合計画, スマートシティ構想, テキストマイニング

Keywords: Local Government Comprehensive Plans, Smart City Concept, Text Mining

1. はじめに

1.1 研究背景と目的

近年、新型コロナ危機を契機にして市民が都市や地域での生活に対する考え方が変わりつつあるなか、高齢化や東京一極集中による地方の衰退といった課題が顕在化してきている。これらの都市の課題に対処しながら、経済を更に発展させるために内閣府は、Society 5.0 を目指すべき未来社会の指針として提唱し、ICT 等の積極的な活用が不可欠であるとしている[1]。

スマートシティは、Society 5.0 実装の場と定義されており、地域における ICT 等の新技術を駆使した高度なマネジメント(計画、整備、管理・運営等)により、都市や地域が抱える様々な課題を解決し、新たな価値を創出し続け、持続可能な都市や地域へと進化させることを目指している[2]。地方自治体はSociety5.0 の実現に向け、行政経営にスマートシティ構想など分野別計画を策定しているが、実際の実装においては十分ではないことが指摘されている[3]。

本研究では地域特有の課題や、取り組みの優先順位がある中、国のガイドラインやナレッジを参考に独自のスマートシティ

構想を策定している。一方で、課題や計画は十分に定量化されておらず、計画から実装におけるPDCA やOODA プロセスの継続が困難なため、実装が進んでいない状況だと仮定した。そこで各自治体のスマートシティ構想計画のテキストマイニングによる定量化のために、まずはテキストマイニング方法の定量化が可能か考察を行う。

各自治体はスマートシティをどのように進めようとしているのか、さらに課題となっている実装をどのように理解し、施策を活用しようとしているのかを検証することが可能か、まずは行政運営の基本である自治体総合計画やスマートシティ構想等に対しテキストマイニングを用いて定量化を行う。

2. 先行研究

本研究ではスマートシティ構想の構造的理解をテーマとして類推を行う。スマートシティとテキストマイニングの研究について CiNii Research (<https://cir.nii.ac.jp/>)2023年12月31日時点のフリーワードで「スマートシティ」のみで検索した場合1,500件以上、「テキストマイニング」のみでは5,000件以上の論文が検索されるが、「スマートシティ テキストマイニング」を検索すると近い先行研究はほぼ見当たらなかった。そのため他分野の研究のうち「観光 テキストマイニング」は約100件が検索され、これを先行研究の対象とした。

2.1 観光関連の記述に関するテキストマイニング

観光関連の記述に関するテキストマイニングで泉澤(2019)

2024年3月1日受理。(2024年3月16日「スマートライフ学会2024年大会」にて発表)

著者照会先: 〒432-8011 浜松市中央区城北3-5-1
静岡大学 創造科学技術大学院 平松 隆志

[4]は、地方の観光振興を観光客の前観光経験と有意味な経験を分析し、その構造的変化から再訪問を促進する方法に関する知見を得ることを目的とした研究を行っている。観光客の理解構造(ある特定の観光地における経験を経た観光客の中にあるその観光地に対する新たな理解の構造)の解明を行っている。

地域によって観光経済は行政運営の大きな要素でありインバウンド施策は国策でもあることから、「出発地—目的地モデル」「構造モデル」「発達モデル」など観光客の行動モデルが研究され、自治体の観光施策等で使用されている。

この研究では、ガイドブックと旅行記をテキストマイニングの手法で分析し、観光客の前観光経験と有意味な経験の差異を捉えている。スマートシティにおいては、対象が観光客と市民、ガイドブックと行政文書のスマートシティ構想と異なるが、市民のスマートシティ構想の経験の差の把握にこのテキストマイニング手法が有効であるか検討する。

2.2 観光政策とロコミのテキストマイニング

同じく観光政策とテキストマイニングの研究で、富川(2022)[5]は、地域の観光振興策の中で従来の政策内容の精査や政策による効果検証にテキストマイニングを用いている。観光政策の内容や旅行ロコミサイトにある膨大なデータを基に分析を行い、広島県竹原市の大久野島観光の政策的課題を考察している。

大久野島観光では、観光ビジョンで特に大久野島のウサギを活かした観光客誘致を図り、ロコミから政策内容と観光客ニーズが合致していることが明らかになった。一方で観光客にとって大久野島観光のニーズは他の観光資源にもあり、ウサギ観光によるネガティブイメージもあることや GIS を用いて観光効果と大久野島人口増減の空間的関係を分析し、観光による正と負の効果両方から実態を調査し考察している。

この研究では、テキストマイニングの分析結果から、実際に島の観光施策に何が合ったのか資料調査を行い、時系列変化を明らかにしている。本研究でスマートシティにおいてもテキストマイニングから導出された分析に対し、実際の施策における影響の調査を行うものとする。

2.3 行政文書のテキストマイニング

地域によって観光経済は行政運営の大きな要素でありインバウンド施策は国策でもあることから、「出発地—目的地モデル」「構造モデル」「発達モデル」など観光客の行動モデルが研究され、活用されている。

一方、スマートシティ構想では、国内で 2010 年頃に実証が始まり、事例をもとに成功例、失敗例から令和 5 年度 第 1 回スマートシティモデル事業等推進有識者委員会(国交省) [3]にて全国の取り組み状況や課題、対策についてまとめられ報告されるようになってきたが、評価の標準化や、取組みのモデル化は今後進んでいくものと思われる。

観光政策とスマートシティ構想はどちらも、計画やビジョンとして行政文書で策定されるケースがある。計画の評価として観光の場合、訪問者数や経済効果、観光動態データ等の数値データや、ロコミ、アンケートのテキストマイニングによる分析評価の定量化が研究されている。スマートシティ構想の評価においても、QOL での評価や、自治体共通の市民アンケートと KPI で定量化した地域幸福度(Well-Being)指標 [6]の活用が進んでおり、これらガイドラインや公開データと掛け合わせることでスマートシティ構想においても社会実装を目的として、行政文書のテキストマイニングの定量化に向けての事例分析は今後の研究課題としている。

3. 研究方法

本研究で呼称するスマートシティ構想は、地方自治法で策定が義務付けされていないため、体裁や呼称、文書の位置づけも各自自治体で異なる。自治体には、都市計画、福祉、教育などの具体的な政策を策定した分野別計画と同列のものや、政策は言及せずビジョンを示すものまで幅広い。そのためテキストマイニングするスマートシティ構想の文書について選定を先に行う。

3.1 自治体文書の選定

自治体総合計画は地方自治体における行政運営の最上位計画であり、住民全体で共有する自治体の将来目標や施策を示し、住民や事業者、行政が行動するための基本的な指針となるものである。2011 年 8 月に基本構想の策定を義務付けていた規定が廃止され自治体は引き続き独自の計画を策定し運用している[7]ため文書は自治体ごとに様々な形態をとっている。本研究では、その自治体の包括的な計画である性質と、次にテキストマイニングのための文章量で意図的に選定した。

3.1.1 分析対象の総合計画の選定

分析の対象は、各自自治体ホームページ(自治体 HP)に掲載されている総合計画、基本計画、戦略ビジョンなどから選定する。

3.1.2 スマートシティ構想の選定

今回主に取り扱う浜松市は、約 1,558 ㎢の広大な市域の中に、海、山、川、湖といった豊かな自然環境と都市環境が共存した「国土縮図型」の政令指定都市といわれている。全国トップクラスの広大な市域や道路総延長距離、橋梁数があり、みなし過疎地域が市域面積の約 50%となる過疎地域の内包課題を有している。また全国の地域課題を凝縮した政令指定都市としてスマートシティに取り組んでおり、特徴が出やすい可能性があり選定している。比較対象は政令指定都市 20 市とした。

各自自治体のスマートシティ構想は、各自自治体 HP から「スマートシティ」「DX」等で検索し内容を確認したうえで近いものを選定した。ただし DX においては、自治体業務の効率化の

みにフォーカスされている場合があり、内容を確認し、スマートシティに関する文書であれば選択した。ファイル形式においても、MS-Word や PDF 形式(文字コード有無)、また HTML の場合は文字をコピーしてテキストファイルにしている。

3.2 分析ツール

表 1 使用ソフトウェア等

ソフトウェア等	用途
KH Corder	テキストマイニングツール
MeCab	形態素解析エンジン
pdf2text	PDF をテキストコンバート
ImageScanOCR	OCR 機能
mecab-ipadic-NEologd	システム辞書(2020-09-10(木)定期データ更新版)
JUMAN 品詞体系	益岡・窪田文法の品詞体系

テキストマイニングツールについては、社会調査によく用いられ、分析手順についての記述の多い KH Corder[8]を使用した。形態素解析には言語、辞書、コーパスに依存しない MeCab を使い、新語のシステム辞書として「mecab-ipadic-NEologd」を使用し、品詞の整理には益岡・窪田文法を拡張した JUMAN 品詞体系を使用した。文字コードありの PDF からテキストを抽出する際、コンバータやコピーアンドペーストでは文字コードにより文字化けが発生することがあり、文字化けが発生した場合は、KH Corder で解析ができる文字コードに変換する作業を加え、改めて MeCab に取り込んだ。

文字がアウトライン化されているなど文字コードが利用できない PDF については、ImageScanOCR を使用してテキスト化した。ImageScanOCR は WindowsOS に標準搭載されている OCR 機能を使用しており、昨今の AI-OCR ほど精度はないため、読み取り後に可能な範囲で手作業でのデータクレンジングを行った。使用したシステム辞書の mecab-ipadic-NEologd は 2020 年 9 月以降更新がないため、以降の新語を形態素解析に使用するため統一の強制抽出ユーザ辞書を作成し使用した。一方で使用しない語の指定は行っていない。

3.3 分析アプローチ

選定した行政文書がテキストマイニングに適しているか検証し、次いで各分析から特徴的な抽出語が現れるか検証するため浜松市をはじめ政令 20 都市から複数の手法で分析を行う。

3.3.1 行政文書の形態素解析

政令 20 都市の総合計画とスマートシティ構想を比較するにあたり、自治体ごとの文書の形態に相関があるのか確認するため KH Corder で形態要素解析を行い、各自治体の形態要素構造を比較し相関があるか分析した。

総合計画、スマートシティ構想ともに自治体 HP から取得した異なるファイル形式を取り込んだが、抽出語数と抽出された

語の種類(異なり語)の率(異なり語数÷総使用語数)で負の相関がみられ、ファイル形式による明らかな差分は検出されず、行政文書の体裁は各自治体で異なるが、使用される語や、品詞の割合等で大きく変わらないことが確認できた。

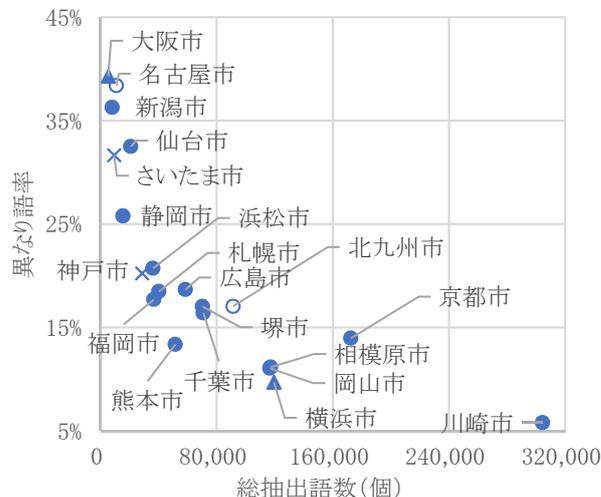


図 1.1 政令 20 都市(総合計画)
総抽出語数×異なり語率

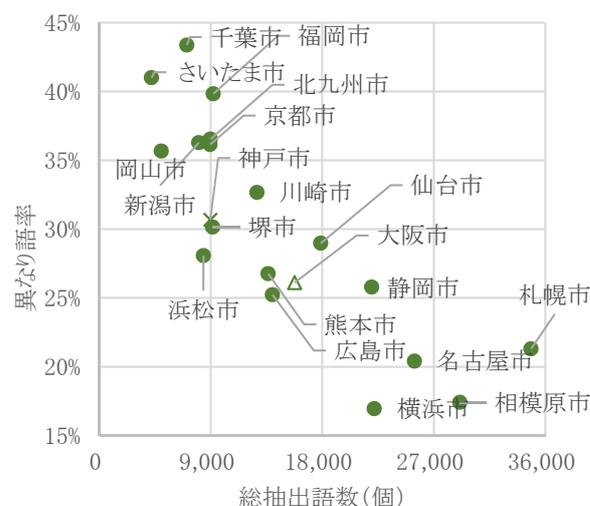


図 1.2 政令 20 都市(スマートシティ構想)
総抽出語数×異なり語率

表 2 図 1.1, 図 1.2 のラベル

ラベル	内容
●	文字コード込みの PDF ファイルをテキストにコンバート
○	文字コードなし(イメージファイル)の PDF を OCR でテキスト化しデータクレンジングを実施
×	PDF をテキストにコンバート後、改行修正等のデータクレンジングを実施
▲	.docx(MS-Word)ファイル
△	.pptx(MS-PowerPoint)ファイル

するため「全体」と名付けた。縦軸の下は、具体的な手段が出てくるため「手段」、上は相対的に「計画」を名付けた。「計画」寄りの共通クラスターと「手段」寄りの特徴語「スマートシティ」から、スマートシティを計画の手段としている構造が理解できる。

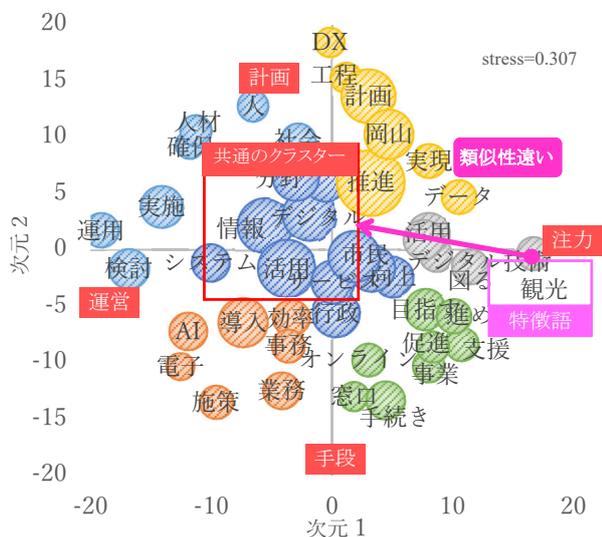


図 3.2 抽出語の 2 次元尺度構成 (岡山市 DX 推進計画)

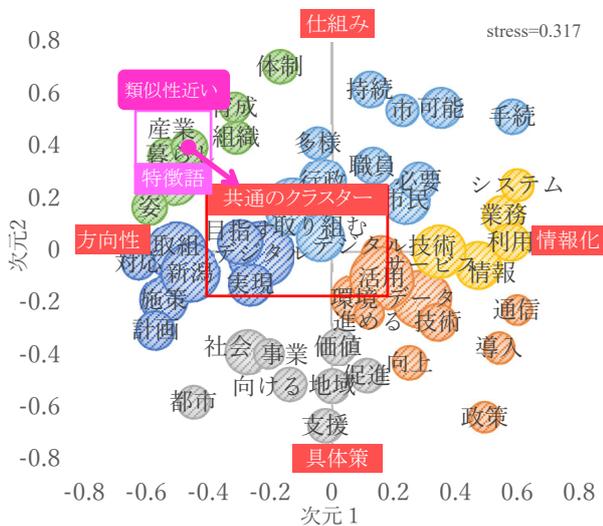


図 3.3 抽出語の 2 次元尺度構成 (新潟市デジタル化基本方針)

表 5 図 3.1, 3.2, 3.3 多次元尺度構成設定内容

項目	設定内容
次元	2
集計単位	文
布置される語数	50 程度
クラスター数	6
次元の名づけ	推察

岡山市DX推進計画の 2 次元尺度構成では、縦軸は浜松市と同じ「計画」「手段」となった。横軸は右に特徴の「観光」など「注力」の語がでており、左には DX 推進の「運営」に関する語が多く出ている。岡山市を浜松市と近い人口で選定しているため同じスマートシティ構想をテキストマイニングすることで縦軸が同じになることは、スマートシティ構想の構造が似ていることが推察される。一方、新潟市デジタル化基本方針の 2 次元尺度構成では、浜松市、岡山市とは異なる軸の名づけになり自治体毎の構造が解釈できることも推察される。

3.3.4 共起ネットワークによる構造的理解

3 市の 2 次元尺度構成から各市の特徴を視覚化し、軸の名づけからスマートシティ構想の構造を推察した。さらに共通のサブグラフや他のサブグラフとの共起関係を分析することで、それぞれの重要度や関係性の分析を行う。

共起ネットワークの Jaccard 係数や最小出現数は図的解釈を優先し設定したものをサブグラフ (modularity) 検出した。

図的解釈では 2 次元尺度構成で 3 市共通のクラスターを含むサブグラフで似た傾向を示しており、同じサブグラフ内であればその市の特徴の可能性はある。(図 4.1, 4.2, 4.3)。

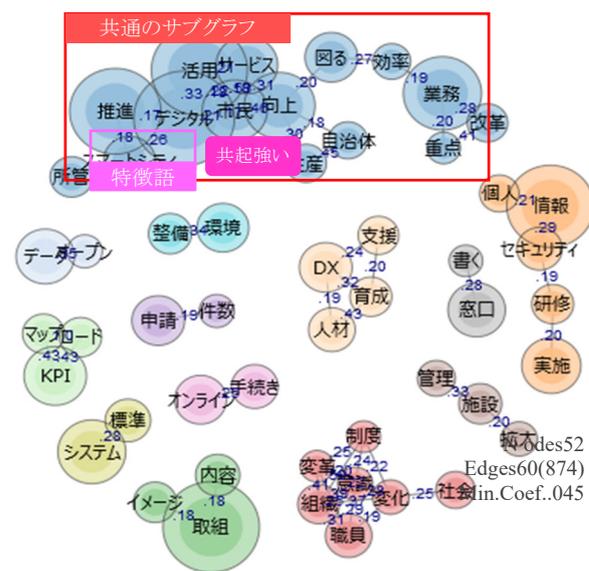


図 4.1 抽出語の共起ネットワーク (浜松市 DX 推進計画【解説版】)

浜松市 DX 推進計画【解説版】の 2 次元尺度構成で特徴として抽出した「スマートシティ」は、共起ネットワーク(図 4.1)では「デジタル」「活用」といった 3 市共通のクラスターと共起関係が強く、「スマートシティ」が浜松市の特徴として表れている。

岡山市の共起ネットワークでは、特徴の「観光」が共通のクラスターと離れて共起し、2 次元尺度構成においても他との類似性が低く、スマートシティ構想の中では観光も一つの要素であることが伺える。実際に本文にて確認すると「産業」「農業」

「観光」「オープンデータ」等の分野(カテゴリ)のひとつではあるが、実際に「観光」が相対的に多く出現していることがわかる。

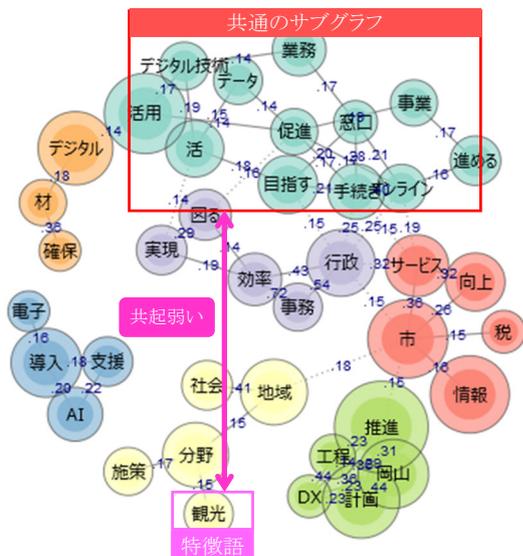


図 4.2 抽出語の共起ネットワーク
(岡山市 DX 推進計画)

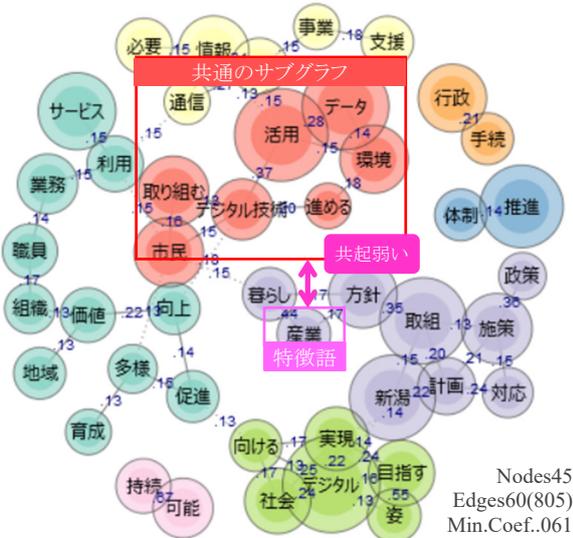


図 4.3 抽出語の共起ネットワーク
(新潟市デジタル化基本方針)

新潟市の共起ネットワークでは、3 市同様のサブグラフになったがクラスター間の関係は他と比較し緩くつながり、特徴が少なく全体的に方針を示していることが伺える。

4. 考察

本研究で、自治体の政策や計画をテキストマイニングから定量的に評価することを可能にした。また各分析アプローチからそれぞれの相関や傾向が確かめられた。その中から浜松市は「スマートシティ」に先行的に取り組み、自治体政策に反

映していることを定量的に証明し、抽出語からも自治体の傾向や特徴を抽出することが可能となった。

政令指定都市の総合計画、スマートシティ構想計画をテキストマイニングした結果、ファイル形式や使用文字数他、様々な体裁がありながら形態素解析で抽出された語の割合において相関を見ることができた。以降の分析で利用したサンプルも基本的な形態素の構成は相関ある文書をテキストマイニングできることが判明した。

浜松市戦略計画において、2015 年から毎年 2024 年分を形態素解析したところ、スマートシティに関する辞書に登録のない語の増え方が確認できた。未知語が多かった 2022 年付近では、国の「デジタル田園都市国家構想総合戦略」が策定や、浜松市が国のスーパーシティ型国家戦略特別区域に応募、指定されるなど実際の環境の変化が調査から判明した。さらに、国のスーパーシティ型国家戦略特別区域の指定を受けた浜松市と仙台市の文書に形態素解析によるテキストマイニングより、抽出される未知語どうしを比較しそれぞれの計画や構想の特徴を把握することができる。

浜松市、岡山市、新潟市のスマートシティ構想を比較した 2 次元尺度構成からは、3 市で「デジタル」「活用」など含む共通のクラスターと、3 市独自の特徴語を視覚化することができた。また軸の名づけでは計画の構造も視覚化できる。さらに共起ネットワークで 2 次元尺度構成から得たクラスターと特徴語のつながりの強さを分析した。

共通のクラスターと特徴語の関係について、2 次元尺度構成では語の類似性の遠近、共起ネットワークでは共起関係(文章内で語と一緒につかわれる頻度や関連性)の係数(Jaccard 係数)をそれぞれ抽出しスマートシティ構想評価の標準化としてモデルを考察した(表 5)。また新潟市の共起においては、共通のサブグラフと特徴語の Jaccard 係数による弱い/強い判断については比較対象が少ないため、推察した結果から弱いとした。

表 5 3 市のスマートシティ構想のパターンモデル

2 次元尺度構成	類似性 遠い	類似性 近い
共起ネットワーク		
共起関係 弱い	整理されている施策から構想(岡山市)	国の方針を踏襲した構想(新潟市)
共起関係 強い	地域の特徴を生かした構想(浜松市)	本研究で取扱いなし

特徴語:浜松市(スマートシティ)、岡山市(観光)、新潟市(産業)

スマートシティ構想のテキストデータから定量化したこのパターンモデルは、他の自治体基礎データなどの数値データとの比較検討も可能となる。またサンプル以外の自治体も同様にパターン化することで、パターンの似ている自治体を見つけ

参考にすることや、目指すスマートシティ構想のパターンを指標とすることが可能である。

5. まとめ

自治体の行政文書をテキストマイニングして規定構造の把握を行い、スマートシティ構想の定量化から考察を得ることができた。分析手法として、各自治体文書において形態素的に相関があり、テキストマイニングに適していること判明した。その上で観光政策の先行研究が、旅行経験前をスマートシティ構想の計画とし参照や引用が有効であることが判明できた。しかし、スマートシティ構想の社会実装が進まない分析には、今回の調査ではサンプルに限られているため、より多くのケースを分析し精度を上げる必要がある。また、実装側の定量化においては地域幸福度(Well-Being)指標を活用した分析を進め、本研究の結果と照合する研究を進める。その他にも、国の政策や新しい技術により常に新しく変わっていくものであり、PDCA や OODA プロセスでも検証できることが望ましいと考察する。

スマートシティ構想自体においても、ひとつの自治体で複数の分野別計画や所掌、単一の自治体ではなく広域で取り組もうとする地域もあり、煩雑であることも分かった。これに国の政策や実装技術、地域課題の要素も加わり煩雑になるため、行政文書のテキストマイニングによる定量化は、計画の理解構造をわかりやすくし行政運営の一助になる可能性がある。

自治体の行政運営は、地方分権改革の取り組みの一環として国から地方への「義務付け・枠付けの見直し」が進み、基本構想策定の義務付けが廃止され、各地方公共団体が独自の行政運営が可能になった。一方、全国にはスマートシティに取り組む余裕のない小規模な自治体も存在しており、持続可能な行政運営が喫緊の課題となっている。スマートシティは本来そのような自治体にこそ生かしていくことが期待されているが、時間とともに情報は分散され、收拾は困難になる傾向がある。引き続き、先進的にスマートシティ構想を進め社会実装においても成功している自治体や、国土縮図型の都市として浜松市を含む他の市町村の研究を進め、困難に直面している地域の課題を解決する手助けとなることを期待して研究を進めていく予定である。

分析対象資料

- ・浜松市総合計画(2014年12月策定):市の最上位計画である総合計画として、基本構想及び基本計画を2015～2016年度を計画期間として策定。
- ・浜松市戦略計画:浜松市総合計画の実現を目指し、市の重点施策やその目標を掲げ、政策や事業とともに、行財政改革や資源配分などの考え方を含めた市政全般にわたる方向性を示し、毎年度の環境の変化を踏まえて策定。

- ・浜松市スーパーシティ/スマートシティシンポジウム資料(2022年2月開催):浜松市デジタル・スマートシティ官民連携プラットフォーム運営委員会(過去の開催概要)。
- ・仙台市×東北大学スーパーシティ構想～キャンパスから未来都市が広がる～(2021年4月公表):先端テクノロジーに関して高い技術を有する東北大学と連携した、「仙台市×東北大学スーパーシティ構想」を取りまとめ国が提案募集を行っているスーパーシティ型国家戦略特区に応募、仙台市が公表した概要資料。
- ・浜松市DX推進計画(2023年1月策定):国の「官民データ活用推進基本法」「デジタル社会形成基本法」等や総務省「自治体DX推進計画」、市の「デジタルを活用したまちづくり推進条例」の制定等を踏まえ策定。
- ・岡山市DX推進計画(2022年3月策定):「岡山市第六次総合計画」を上位計画とする個別計画であり、DX推進によって総合計画の3つの将来都市像の実現につなげるもの。
- ・新潟市デジタル化基本方針(2023年3月策定):新潟市の最上位計画である「新潟市総合計画」が掲げる目指す都市像の実現をデジタルの観点で推進するための分野横断的な指針。計画期間は、2023年度から2030年度まで(期間中においても状況の変化に応じて随時見直しを行う)。

参考文献

- [1] 内閣府他:第5期科学技術基本計画(平成28～平成32年度),科学技術基本計画,閣議決定(2016.1.22).<https://www8.cao.go.jp/cstp/kihonkeikaku/index5.html>
- [2] 内閣府他:スマートシティガイドブック(2023.08 ver.2.00).
https://www8.cao.go.jp/cstp/society5_0/smartcity/guidebook.html
- [3] 国交省:資料 3-1:スマートシティ実装に向けた調査・検討,令和5年度第1回(2023.8.23)
<https://www.mlit.go.jp/toshi/daisei/content/001628426.pdf>
- [4] 泉澤圭亮,中鉢令兒:テキストマイニングの手法の活用による観光者の経験に着目した観光地の理解構造の変化に関する研究([観光研究]2019.9,Vol.31,No.1,日本観光研究学会機関誌)
- [5] 富川久美子:大久野島における観光対象の変遷と観光行動,島嶼研究 第23巻1号(2022.3)
- [6] 一般社団法人スマートシティ・インスティテュート(2024)「地域幸福度(Well-Being)指標」デジタル庁,<https://www.sci-japan.or.jp/LWCI/>
- [7] 大塚敬:地方自治法改正による基本構想策定義務付け廃止の経緯とその影響(2017.5.12),三菱UFJリサーチ&コンサルティング
https://www.murc.jp/library/column/sn_170512/
- [8] 樋口耕一:社会調査のための計量テキスト分析—内容分析の継承と発展を目指して【第2版】、ナカニシヤ出版(2021.2.1)
- [9] 内閣官房デジタル田園都市国家構想実現会議事務

局:地方版総合戦略の策定状況等に関する調査結果(2023.10.20)

著者紹介



平松 隆志(学生会員)

2003年島根大学大学院 総合理工学研究科博士前期課程修了、修士(工学)。同年NTTドコモ東海支社入社。2023年静岡大学創造科学技術大学院博士後期課程に社会人入学、現在に至る。主に、経営戦略・マーケティングの理論を用いた、自治体スマートシティ構想の社会実装の研究に従事。



遊橋 裕泰(正会員)

静岡大学創造科学技術大学院情報科学専攻・教授。東京工業大学修了、博士(学術)。所属学会:スマートライフ学会、経営情報学会、日本情報経営学会、日本マーケティング学会。