

# 「地方創生」における政策キーワードの変遷

## — テキストマイニングによる「総合戦略」の分析 —

仲北浦 淳基

大正大学 地域創生学部地域創生学科 専任講師

(要旨) 本稿の目的は、テキストマイニングの手法を用いることで、2014年から2020年までの「総合戦略」におけるキーワードの変遷を明らかにすることである。そのために、「総合戦略」の第1期初版と第2期初版だけでなく、それらの毎年の改訂版をも含めた計7年分を分析対象とする。分析の結果、第1期から第2期への変化として指摘されてきたことが定量的にも確認されたが、他方で、期内の改訂においても重要な変化があることが分かった。「地方版総合戦略」の策定のために国の「総合戦略」を勘案するときは、期をまたいだ改訂だけでなく、毎年の改訂にも注意を払わなければ、当該のキーワードが強調されるに至った背景や経緯を捉え損ねる可能性がある。

キーワード：地方創生、総合戦略、テキストマイニング、計量テキスト分析、データビジュアライズ

### 1. はじめに

#### (1) 問題の所在

2014年12月27日に閣議決定された最初の「まち・ひと・しごと創生総合戦略」（以下、「総合戦略」と省略する）は、いわゆる「地方創生」の大きな指針の1つとして打ち出された。その後、賛否両論を巻き起こしながらも、「総合戦略」は、改訂が重ねられ、2019年12月20日には“第2期”が閣議決定された。

「総合戦略」を策定する法的な根拠は、2014年11月28日に公布された「まち・ひと・しごと創生法」であり、その第8条第6項には次のように明記されている。

政府は、情勢の推移により必要が生じた場合には、まち・ひと・しごと創生総合戦略を変更しなければならない。（傍点引用者。以下同様）

つまり、政府は、国の「総合戦略」を1度だけ策定して終わりにするのではなく、「情勢の推移」を見ながら、都度変更を加えていかなければならな

いのである。

確かに、「総合戦略」における第1期（2014年）と第2期（2019年）という大きな区分については広く認知されており、その変化に言及した研究もある（例えば森川2020や鈴木2020）。しかし、本稿でも詳述するように、国の「総合戦略」は、第1期から第2期への改訂があっただけでなく、部分的な改訂が毎年行なわれてきた。それにもかかわらず、国の「総合戦略」における毎年の変化に関する研究や調査は見当たらない。

#### (2) 本稿の目的と方法

そこで本稿では、先行研究を踏まえながら、国の「総合戦略」における各年の変化を追う。政府が「地方創生」において強調しようとした政策キーワードの変遷を明らかにしたい。

そのために本稿では、文書という質的なデータを定量的に解析するテキストマイニングという手法を用いる。この手法は、文学を嚆矢として社会学、教育学、経営学、経済学など様々な分野において活用されてきたが、特に日本の政策に関する論議や関心の分析に関しては、新聞記事を用いて

都市計画の潮流を明らかにしようとした小森他(2006)が最初であるように見える。その後も現在までに数多くの研究が蓄積されている(例えば小田切2014; 森2019; 増田2021)<sup>1</sup>。

テキストマイニングという手法を用いる利点は、言葉という質的なデータを定量的に解析できることである。また、膨大な質的なデータを読み込む際の分析者の着想や“読み”を、より客観的な数量的証拠をもって検証できることも利点である。さらには、膨大な質的なデータを定量的に解析することで、これまで一般的には気づかれなかった新たな知見を見出すことも期待できる<sup>2</sup>。

本稿では、各年版の「総合戦略」をテキストマイニングによって分析することで、国の「総合戦略」のキーワードがどのように変遷してきたのかを明らかにする。

## 2. 分析対象の概要と分析の方針

### (1) 分析対象の概要

本稿は、内閣府が提供する「地方創生サイト」で公開されている「総合戦略」を分析対象とする。ここで、2022年2月現在の策定状況を確認しておこう(表-1)。「地方創生」を掲げる「まち・ひと・しごと創生法」が2014年11月28日に公布されてから、最初の「長期ビジョン」と「総合戦略」が公表されたのは同年12月27日である(表-1、第1行)。

その後、「長期ビジョン」は5年後の2019年12月20日に「令和元年改訂版」として公表されている(表-1、第2列)。2015年6月30日には、「国の「総合戦略」に掲げられた基本目標の達成に向けて作成された政策パッケージ・個別施策について、今後の対応の方向を取りまとめたもの」(「基本方針2015」、3)として「基本方針」が初めて示され、その後も毎年6・7月に公表されている(表-1、第1列)。

そして、本稿で分析対象とするのは「総合戦略」であり、2014年から2018年が「第1期」、2019年から2023年が「第2期」と定められてはいるが、毎年12月には改訂版が公表されている(表-1、第3列)。ただし、2022年2月現在において、2021年版の「総合戦略」は公表されていない。そのため、本稿が分析対象とする「総合戦略」は、2014年版から2020年版までの計7年分である。

表-1 「基本方針」、「長期ビジョン」、「総合戦略」のそれぞれの策定日(「地方創生サイト」より筆者作成)

	「基本方針」	「長期ビジョン」	「総合戦略」	
平成26年(2014年)		12月27日	12月27日	第1期
平成27年(2015年)	6月30日		12月24日	
平成28年(2016年)	6月2日		12月22日	
平成29年(2017年)	6月9日		12月22日	
平成30年(2018年)	6月15日		12月21日	
令和元年(2019年)	6月21日	12月20日	12月20日	第2期
令和2年(2020年)	7月17日		12月21日	
令和3年(2021年)	6月18日			

### (2) 分析対象に関する先行研究

#### a) 国と自治体のトップダウン関係

「まち・ひと・しごと創生法」(第8条、第9条、第10条)によれば、「総合戦略」は政府(国)・都道府県・市区町村の3段階でそれぞれ定められることになっている。ただし、各自治体は「〔国の〕まち・ひと・しごと創生総合戦略を勘案して」、その「地方版総合戦略」を定めるように努めなければならない、と規定されている。確かに、各自治体(都道府県および市区町村)は、当該自治体の実情に応じて「地方版総合戦略」を定めると規定されているが、実際上、自治体は国の「総合戦略」の方針に沿うことが求められている。

「地方版総合戦略」の策定において、国と自治体のトップダウン関係があるという点については、先行研究においても繰り返し指摘されてきた。例

<sup>1</sup> なお、筆者は経済学説史・経済思想史の分野において、英国の経済学者の言説や政策委員会での議論をテキストマイニングによって分析してきた(仲北浦2021; 仲北浦・小峯2019; 小峯・仲北浦2021)。

<sup>2</sup> さらに、テキストデータの範囲内でその傾向などを捉えるだけでなく、その範囲外から新たな発見を抽出する「探索的なデータ解析 exploratory data analysis」をしてこそ「真のテキストマイニング real text data mining」だとする立場もある。(Cf. Hearst1999, 3)

えば、中罵（2018、10）は「総合戦略をどのように策定するかは自治体により様々であるが、国の総合戦略、都道府県の方針に合致したものとするためには、…人口減少を予測している自治体であっても、予測される人口減少を受け入れた上での計画の策定は許されない状況である」として、各自治体の「地方版総合戦略」が国（上位）の「総合戦略」の方針内に納まらざるを得ないことを指摘している。

また、政府は2014年（第1期初版）の「総合戦略」において、

適切な短期・中期の政策目標を伴う政策パッケージを示し、それぞれの進捗についてアウトカム指標を原則とした重要業績評価指標（KPI）で検証し、改善する仕組み（PDCA サイクル）を確立する（7ページ）

と宣言している。「政策パッケージ」の提示と「KPI」の設定という方針は現在も継続されているが、この方針がさらに国と自治体のトップダウン関係を引き起こしているとの指摘がある。

政策パッケージに関しては、福岡県と県下60市町村の「人口ビジョン」および「地方版総合戦略」を調査した小野他（2019、76）が、「[国の定める]政策パッケージは各地方自治体の独自性の高い政策展開を結果的に阻んでしまっている実態にあった」と結論づけている。

また、KPI の設定に関して松井（2017、45）は、KPI の設定に懸念を示す説を紹介した上でその懸念理由を次のように推測している。すなわち、KPI という用語が政策の中で明示的に用いられた『日本再興戦略』（2013年）の中で、KPI の達成可否については「トップダウンで検証を行う」と断定されていることから、「総合戦略で求められた KPI もまたトップダウン的な検証が想定され、事後的な検証のために自治体はその通りに実施するように

「上から」の管理として理解」されたのだろう、という推測である。

このような様々な事情<sup>3</sup>から、特に第1期においては、当事者である自治体の担当者も国からのトップダウンの印象を受ける結果となったようだ。

「地方版総合戦略」（第1期）の策定に関して自治体に対する悉皆調査を行なった坂本（2018、95）は「国主導の強いコントロールの下で進められており、地域の実態に見合った運用になっていない」と感じている自治体が少なからず存在することを指摘している。

以上のように、「総合戦略」の策定において国と自治体がトップダウン関係にあることが分かる。ただし、本稿の目的は、国と自治体のトップダウン関係や国の「総合戦略」の是非を問うことではない。むしろ、本稿は、これらの先行研究が指摘しているような国と自治体のトップダウン関係があるのだとすれば、国の「総合戦略」が「地方版総合戦略」に強い影響を及ぼしているはずであり、であるならば、なおさらこれまでの国の「総合戦略」の特徴や変遷を明らかにしておく必要があるだろう、と主張するのである。

#### b) 「総合戦略」第1期から第2期への変化

しかし、「地方版総合戦略」に関する研究に比べて、国の「総合戦略」の研究は比較的手薄な状態である。ただし、国の「総合戦略」における第1期から第2期への変化については、特に《基本目標の微修正》および《「横断的な目標」と「関係人口」の追加》が先行研究によって指摘されている。

例えば鈴木（2020、4-5）は、第2期の基本目標の変化を「微修正」とみなした上で、特に「稼ぐ地域」や「地方とのつながり」という語が基本目標に追加されたことに注目している。さらに、第2期においては新たな目標を「既存の目標に加えるのではなく、「横断的な目標」という形で追加された点が注目される」として、「横断的な目標」という第2期の新たな方針に注意を促している（同

<sup>3</sup> 他にも、地方創生関連の補助金や交付金に採択される必要があるという仕組みからトップダウン関係が生じるという指摘もある。例えば、中澤（2016、289）は、「政府がパッケージごと具体的なかつ短期的な数値目標を設定している以上、これらと整合する内容と水準の目標を設定していない限り、[採択

のための]コンテストに勝ち抜き、補助金や交付金の交付を受けることはできないであろう。…[国から]求められると想定される内容と水準から「地域版総合戦略」が逆算される傾向を生む」と述べている。

上、3-4)。なお、「横断的な目標」は2つ設定されており、一方が「多様な人材の活躍を推進する」、他方が「新しい時代の流れを力にする」である。後者はさらに「地域におけるSociety 5.0の推進」と「地方創生 SDGs の実現などの持続可能なまちづくり」という2つの方針が掲げられている。

また、森川(2020、302)は、第1期から第2期への変化について、「…[第2期では]「関係人口」の概念が導入されたくらいで、第1期の戦略目標と比べて特に目新しいものはない」と述べ、第2期における「関係人口」の追加を強調している。なお、「関係人口」は、上述の第2期の基本目標で追加された「地方とのつながり」と関係の深い語である。

本稿では、これらの先行研究でも指摘された第1期から第2期にかけての大きな変化、すなわち、《基本目標の微修正》および《「横断的な目標」と「関係人口」の追加》を定量的に確認することを第一の目標とし、さらに、「総合戦略」の毎年の改訂版を比較分析することによって、国の「総合戦略」における政策キーワードの変遷を辿りたい。このような知見は、国の「総合戦略」を勘案する必要に迫られる自治体にとって、今後も「地方版総合戦略」の改訂を重ねていく上でのメルクマールとして意義を持つであろう。

### (3) 具体的な分析の手順と留意事項

#### a) 具体的な分析の手順

本稿ではKH Coder 3<sup>4</sup>というフリーソフトを用いて計量テキスト分析を行なう。KH Coder 3では、多種多様な分析方法が実装されているが、本稿では主に対応分析とKWIC コンコーダンス(コロケーション統計)を用いる。

本論で詳説するが、対応分析は、複数の文献(例えば第1期「総合戦略」と第2期「総合戦略」)における特徴語を見出すのに長けている。本稿では対応分析を応用することで、「総合戦略」が年々どのように変化してきたのかを明らかにする。また、テキストマイニングでは、単語を最小単位として

分析するが、その語の前後関係や文脈を明らかにしたいときはKWIC コンコーダンス(コロケーション統計)を活用する<sup>5</sup>。

本稿では、2014年と2019年に策定された第1期と第2期の「総合戦略」だけでなく、毎年の改訂版も分析対象とし、以下の3段階で計量テキスト分析を行なう。

- ① 2014年版から2020年版までの量的な推移を明らかにする。
- ② 2014年版(第1期初版)と2019年版(第2期初版)の「総合戦略」を比較する。
- ③ 各期初版とその改訂版をそれぞれ比較する。

これらの分析によって、2014年から2020年にかけて、国の「総合戦略」の強調点がどのように変遷してきたのかを明らかにしたい。

#### b) 分析における留意事項

本稿では内閣府の「地方創生サイト」で公開されている「総合戦略」を用いるが、これらの「総合戦略」のテキストのうち分析対象とするのは本文のみであり、表題、目次、図表および脚注は分析対象から除外している。

なお、計量テキスト分析においては主に出現回数の多い語が分析対象となるが、これは逆を言えば、出現回数の少ない語が分析から除外されるということである。ゆえに、できるだけ分析の漏れが生じないようにするためには、高頻出語と低頻出語に分けてそれぞれ分析するなどの工夫が必要である。しかし、本稿では紙幅の制約上、そこまで詳細な分析はできないので、暫定的に出現回数30回を1つの基準として、それ以上出現した語を分析対象としている。

また、KH Coder 3では、動詞や形容詞など様々な品詞を分析対象に含めることができるが、本稿では名詞のみを分析対象とする(「名詞」、「サ変名詞」、「固有名詞」、「組織名」、「人名」、「地名」、「タグ」、「名詞B」、「名詞C」)。というのは、名詞は各品詞の中でも理念・概念・思想を強く象徴すると

<sup>4</sup> <https://kncoder.net/dl3.html>(2022年2月1日最終閲覧)

<sup>5</sup> テキストマイニングの概要とKH Coder 3の操作方法については樋口(2020)を参照のこと。他にもテキストマイニングに関

連する具体的な手順や留意事項については仲北浦(2021)および、仲北浦・小峯(2019)と小峯・仲北浦(2021)の補論・補説も参照されたい。

考えられているからである（喜田2018、210）。

### 3. 分析と考察

#### (1) 2014年版から2020年版までの量的な推移

まず、基本的な量的推移として、2014年版から2020年版までの7年分の「総合戦略」の総語数と名詞の語数（図-1）および、全体の異なり語数（語の種類数）と名詞の異なり語数の推移を見る（図-2）。

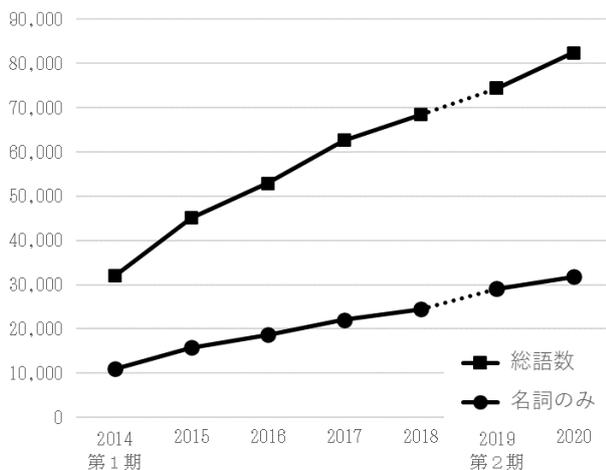


図-1 総語数と名詞の語数の推移  
(KH Coder 3の分析結果より筆者作成)

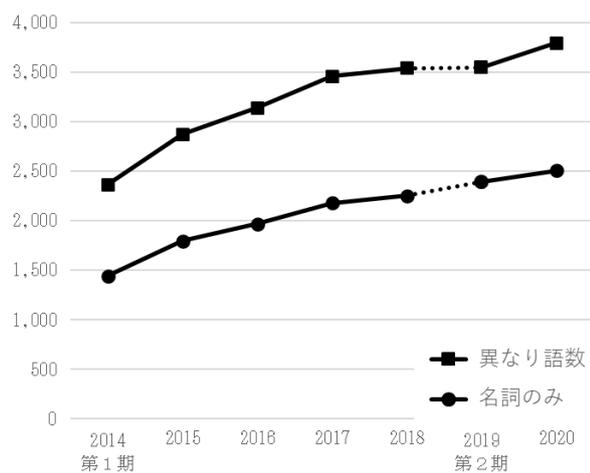


図-2 全体の異なり語数と名詞の異なり語数の推移  
(KH Coder 3の分析結果より筆者作成)

図-1より一見して明らかなように、「総合戦略」は毎年改訂を重ねるたびに分量が増加している。また、図-2からは、全体としての語数が増加しているだけでなく、使用されている語の種類も増えていることが分かる。さらに、第1期から第2期での改訂よりもむしろ、毎年の改訂において分量が増加していることは注目に値する。内容の分析は次節以降で行なうが、少なくとも総語数や異なり語数から見れば、「総合戦略」が毎年大幅に書き加えられていることは明らかであろう。

次に、語の出現回数の分布をみる。次ページの図-3は各年版「総合戦略」の語の出現回数（横軸）と度数（縦軸）を表したグラフをまとめたものである。どの年も出現回数の少ない語が非常に多く、出現回数30回未満の語だけで全体の約90%~95%を占めている<sup>6</sup>。さらに詳細な調査は必要だが、出現回数の少ない語が全体の大半を占めていることは、その分「総合戦略」が幅広いテーマに言及していることの表れと言えるかもしれない。

#### (2) 第1期初版と第2期初版の比較

##### a) 《基本目標の微修正》および「横断的な目標」と「関係人口」の追加

次に、一般的に知られている第1期から第2期への改訂における変化を定量的に示すために、2014年版（第1期初版）と2019年版（第2期初版）の「総合戦略」を対応分析によって比較する。ここでも分析対象とするのは出現回数が30回以上の名詞である。結果は図-4のとおりである。分析対象となる語は282語あるが、この図で散布されているのは「差異が顕著な語」かつ「原点から離れた語」の上位60語のみである（以下同様）。

まず、対応分析の図の見方を説明しておこう。横軸と縦軸のそれぞれに「0」があり、それぞれ垂直、水平に点線がひかれている。それらの点線の交点が原点である。そして、原点付近に散布された語はより普遍的に出現した一般語であり、原点から離れて散布された語はその当該文書で特に

<sup>6</sup> 先述したように本稿では出現回数30回以上の名詞のみを分析対象とする。分析対象となっているのは文書全体の5%~10%ほどである点に留意が必要である。

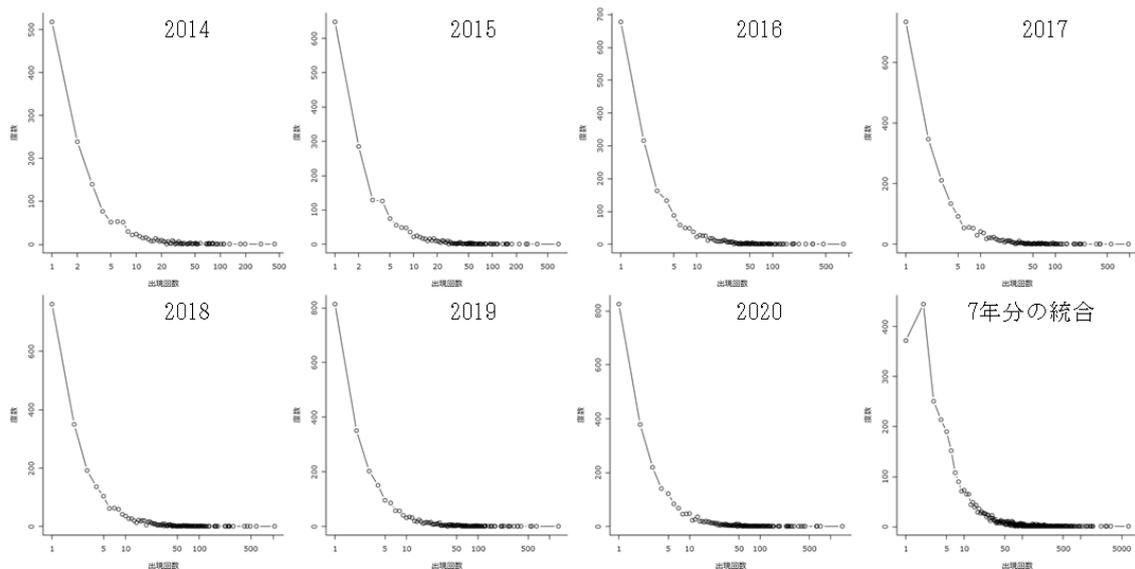


図-3 各年版「総合戦略」における出現回数の分布 (KH Coder 3の分析結果より筆者作成)

多く出現した特徴語である。具体的に図-4で説明すると、原点の右上(第1象限)に「2019年(第2期)」、左下(第3象限)に「2014(第1期)」の語が散布されている。原点付近には「地域」と「地方」という語が散布されており、これは2014年版と2019年版に共通して出現した語(一般語)であることを意味する。他方で、例えば図の左下隅の「府省庁」という語は2014年版で特に多く出現した語、逆を言えば、2019年版ではほとんど出現しなかった語だということである。こういう意味で、「府省庁」は2014年版の特徴語とみなせる。以上のことから、第1象限に散布されている語は2019年版の特徴語であり、第3象限に散布されている語は2014年版の特徴語とみなすことができる。また、原点から離れるほど、比較相手における出現回数がより少なかったという意味で、当該文書の特徴語と言える程度が増す。

では、2014年版(第1期初版)から2019年版(第2期初版)への変化はどのようなものだったのだろうか。まずは、前章(2)のb)で指摘した《基本目標の微修正》および《「横断的な目標」と「関係人口」の追加》に注目してみよう。

次に、2014年版の特徴語(第3象限)として「府省庁」があるのに対して、2019年版は「内閣府」、「内閣官房」、「国土交通省」、「総務省」、「農林水産省」、「厚生労働省」、「参事官」、「政策局」、「政策課」、「企画課」、「振興」、「事務」、「本部」が特

徴語として出ていることから、2014年版では「府省庁」として指摘されるにとどまっていた組織名が、2019年版では具体的に明言されていると推測できる。実際に、第2期「総合戦略」の本文を見ると、その後半部分の「政策パッケージ」で設定されている詳細な目標の1つ1つ全てに、関係する府省庁の具体的な組織名が列挙されている。これだけでは確たることは言えないが、この変化は縦割り行政を脱却して分野横断的に地方創生に取り組もうという姿勢の強さの表れと言えるかもしれない。

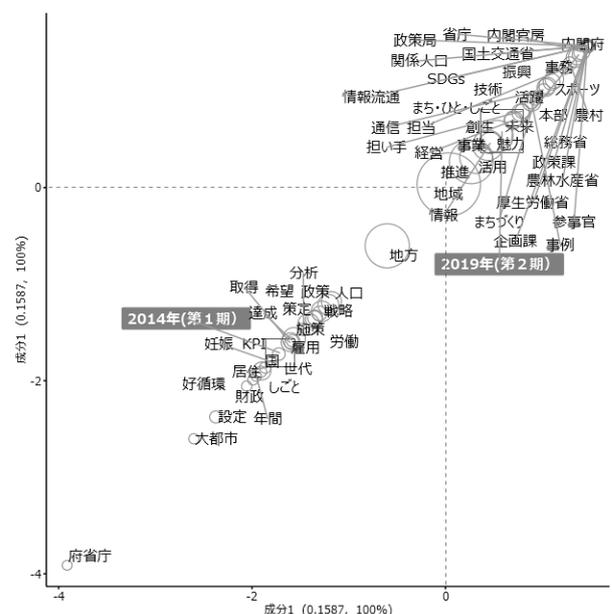
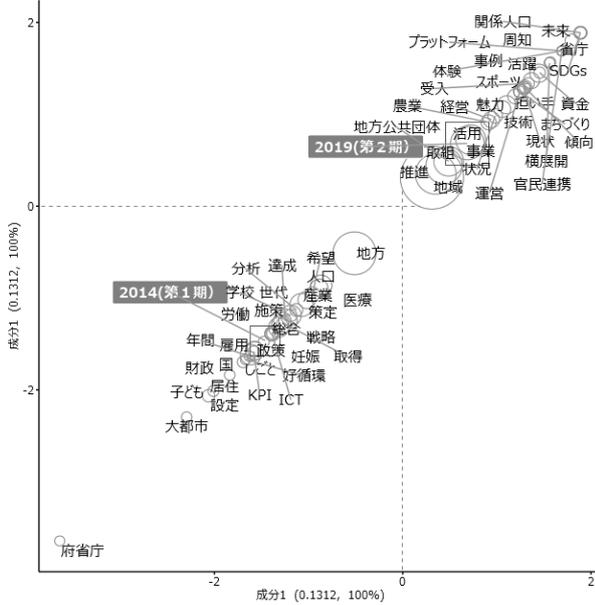


図-4 2014年版と2019年版の比較(対応分析)

## b) その他の主な変化



図ー5 2014年版と組織名を削除した2019年版の比較  
(対応分析)

以上の分析から、「SDGs」と「関係人口」が第2期「総合戦略」における強い程度の特徴語として表れたことが定量的にも示せた。さらに、第2期において関係府省庁の組織名が具体的に指定されるようになったことも指摘した。

ただし、様々な組織名が高頻度で出現することで、第2期「総合戦略」の内容における特徴語が見えにくくなっている。そこで、第2期の「政策パッケージ」で具体的な組織名が記された部分を全て削除して、再度対応分析を行なった（以下の分析でも、具体的な組織名を削除したテキストを使用することとする）。その結果が図ー5である。

前項で指摘した「SDGs」と「関係人口」を除くと、まず注目されるのは右上隅の「未来」という語である（2014年版：1回、2019年版：65回）。KH Coder 3の機能であるKWICコンコルダンス（コロケーション統計）から、「未来技術」という表現が65回中41回使用されていることが分かった（「技術」という語も2019年版の特徴語として第1象限に散布されている）。この語は「新しい時代の流れを力にする」という「横断的な目標」の説明の中で初出であり、「Society5.0の実現に向けた技術」を意味するようだ（「総合戦略」2019年版、24）。「未来」という語は「横断的な目標」という第2期の特徴を

反映していると言えるだろう。なお「Society5.0」は2019年版では13回出現しているが、2014年版では出現していない。

次に「プラットフォーム」という語がある。コロケーション統計によれば、この語は「〇〇プラットフォーム」という形で様々な意味合いで使われている。一例を示すと、「官民連携プラットフォーム」（11回）、「データプラットフォーム」（3回）、「地域プラットフォーム」（3回）、「マイキープラットフォーム」（2回）などである。何らかの活動を促す際に、国（政府）がその活動の基盤や仕組みを構築しようとする姿勢が垣間見える。

最後に、やや意外の感を抱くが、「農業」という語も第2期の特徴語として出ている（2014年版：13回、2019年版：66回）。出現回数から判断して、第1期において「農業」が軽視されていたとは言えないだろうが、第2期において出現回数が飛躍的に増加している。コロケーション統計を見ると、「スマート農業」という表現が10回、他にも「データ」（3回）、「技術」（3回）、「AI」（1回）、「ICT」（1回）という語が「農業」と共起していた。これもまた、第2期の「横断的な目標」の「地域におけるSociety5.0の推進」が反映された結果と推測できる。また、「農業」と「経営」が8回共起していること、および、第2期の特徴語として「官民連携」「資金」「事業」「活用」が出ていることは、「稼ぐ地域」という第2期の《基本目標の微修正》を反映した結果と言えるかもしれない。

### (3) 各期初版とその改訂版の比較

前節では、2014年版（第1期初版）と2019年版（第2期初版）を比較した。その分析によって、第1期から第2期への変化を定量的に確認することができた。しかし、それだけでは不十分である。というのは、2014年版から2019年版への変化がそれまでの改訂ですでに起こっていた可能性があるからである。第1期中の改訂における変化を、第1期から第2期への改訂による変化だと誤認することは避けなければならない。そのためには「総合戦略」の毎年の変化を追っていく必要がある。

そこで本節では、各期初版とその改訂版とをそれぞれ比較していく。そうすることで、各年の「総

合戦略」における比較的細やかな変化を捉えられるだけでなく、上記の問題もある程度は回避できるだろう<sup>7</sup>。なお、紙幅の制約上、各分析において、特に強い程度での特徴語と言える語（図の右上隅に散布された語）を中心に言及する。

#### a) 2014年版(第1期初版)と2015年版の比較

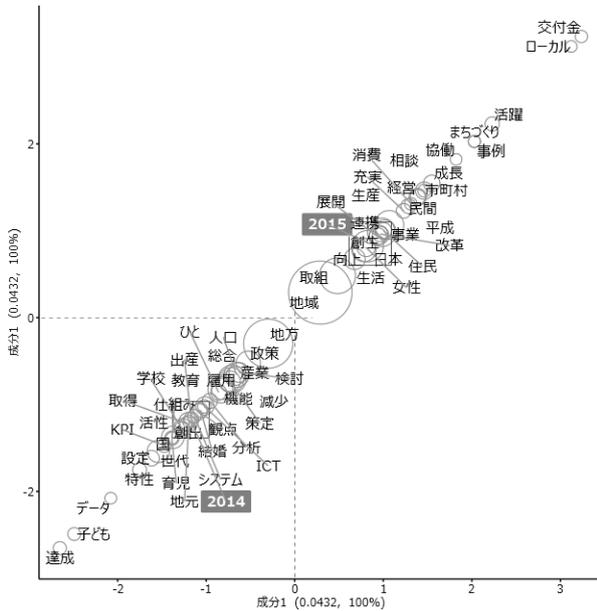


図-6 2014年版と2015年版の比較（対応分析）

2014年版(第1期初版)の最初の改訂である2015年版では、「交付金」と「ローカル」が特徴語として出ている（図-6右上隅）。

まず、「交付金」は2014年版(第1期初版)で3回、2015年版で31回出現している。2014年版では「新型交付金」と表現されており、これは「使途を狭く縛る個別補助金や効果検証の仕組みを伴わない一括交付金とは異なる、第三のアプローチを志向したものである(『総合戦略』2014年版、59)。2014年版では構想段階でとどまっていた「新型交付金」は、2015年版において「地方創生推進交付金」として具体化され、繰り返し言及されるようになった(16回)。

次に、「ローカル」は2014年版で3回、2015年版で27回出現している。この語は単体ではなく「ロ

ーカル○○」という表現で出現している。コロケーション統計によれば、2014年版では「ローカル版クールジャパン」(2回)と「ローカル10,000プロジェクト」(1回)という2つの表現で出現している。一方で、2015年版では「ローカル・アベノミクス」が6回で最多である。これは、マクロ経済政策(アベノミクス)の一部として「地方創生」を位置づける方針を反映したものとみることができる(cf.「地方創生版・三本の矢」)。あくまで国家を全体と捉えた上での地域政策という見方は、国の「総合戦略」の一特徴と言えるかもしれない<sup>8</sup>。なお、他にも「ローカルイノベーション」(5回)、「ローカルベンチマーク」(5回)、「ローカル・クールジャパン」(4回)、「ローカルブランディング」(3回)、「ローカルサービス」(3回)などの表現が出現している。

#### b)2014年版(第1期初版)と2016年版の比較

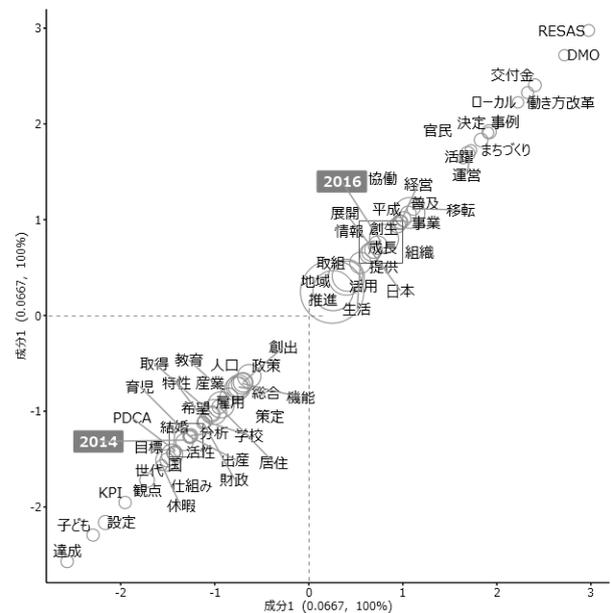


図-7 2014年版と2016年版の比較（対応分析）

次に2度目の改訂である2016年版を比較する（図-7）。まず右上隅には「RESAS」がある(2014年版：0回、2016年版：35回)。RESASとは内閣官房および経済産業省が提供する「地域経済分析シ

<sup>7</sup> 2つの「総合戦略」での出現回数が合計で30回に満たない語が分析対象から除外されていることに留意が必要である。合算して出現回数が30回未満の語は新出語であっても図中には反映されない。

<sup>8</sup> 中澤(2016、5-6)も次のように指摘している。「『地方創生』は、字面が醸し出すような、「新たな『地方』を創り出す」という、「地方」そのものを目的とする理念を含意しているとしても、同時に、そしてそれ以上に、「国家」や「国民経済」という全体の量的拡大の手段としての性格を強く持っている。



地域企業が長期的な事業にチャレンジできるために「リスク性資金の充実」（6回）が必要との立場に立っている。なお、そのようなリスク性のある資金を供給するのは主に「地域金融機関」やそれによって構成される「地域ファンド」とされている。このように、同一の語でも、各年の改訂を重ねるにつれ、文脈や目的に広がりが出ていることが分かる。

最後に、右上の「子供」（2014年版：3回、2016年版：33回）と左下隅の「子ども」（2014年版：25回、2017年版：14回）に注目しておこう。同じ言葉だが、各年版によって漢字と仮名での差異が出ている。KWIC コンコーダンスを調べると、この差異は「放課後子ども総合プラン」と「放課後子供教室」のような制度名などによる使い分けが原因となっていることが分かった。しかしその一方で、「子ども数」と「子供数」のように、とくに使い分けの必要のない、表記の揺れともみえる部分も確認された。やや余談のようにみえるかもしれないが、テキストを定量的に分析する際には、このような表記の揺れにも注意を払う必要があることを指摘しておきたい。「子供」と「子ども」という同じ言葉でも、両者は別々の語としてカウントされるからである。

#### d) 2014年版(第1期初版)と2018年版の比較

次は、第1期最後の改訂である2018年版との比較である(図-9)。まず右上隅に「SDGs」がある(2014年版：0回、2018年版：45回)。ここで注意すべきは、「SDGs」という語は、実は第2期への改訂において出現したのではないという点である。前節(2)b)で指摘したように確かに「SDGs」は第2期の特徴語として表れていたが、実は第1期の2018年版でもすでに特徴語として表れているのである。世界的な潮流に乗り、地域政策においてもその重要性を積極的に説こうとする姿勢がうかがわれる。

さらに、第2期の特徴語として出ていた「未来」も2018年版の特徴語に挙がっている(2014年版：1回、2018年版：29回)。ただし、「SDGs」とは異なり、2018年版では、「未来」という語が第2期とは異なる使われ方をしている。前述のとおり、第2期では「未来技術」という表現が大半を占めて

いたが、2018年版では「地域未来投資促進法」(8回)、「近未来技術」(7回)、「SDGs 未来都市」(3回)、「地域未来牽引企業」(2回)という表現が使用されている。

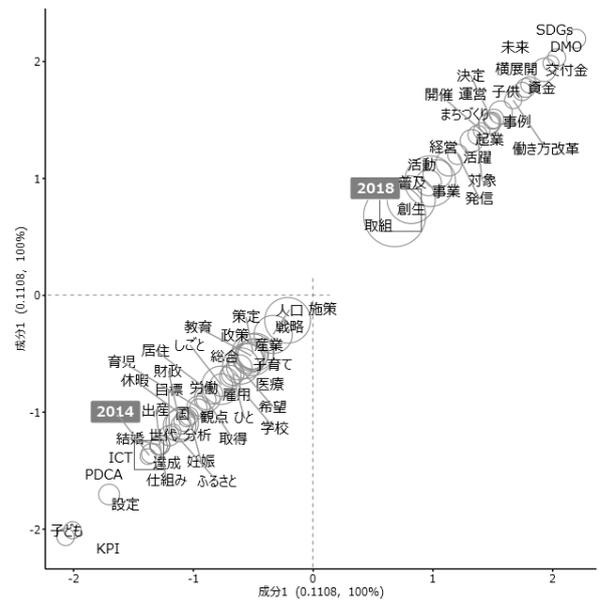


図-9 2014年版と2018年版の比較(対応分析)

「地域未来投資促進法」の正式名称は「地域経済牽引事業の促進による地域の成長発展の基盤強化に関する法律」(2007年施行)であり、2017年6月に改正された。経済産業省によれば、この法律の目的は「地域の特性を生かして、高い付加価値を創出し、地域の事業者に対する相当の経済的効果を及ぼす「地域経済牽引事業」を促進すること」である。この目的は明らかに「地方創生」を念頭に置いたものと考えられ、先の「地域未来牽引企業」もこれに強く関連する語と思われる。法律の改正が「総合戦略」の改訂において即時に反映された例と言えるだろう。また、「近未来技術」や「SDGs 未来都市」という表現は、翌年から開始される第2期の「横断的な目標」(「地域におけるSociety 5.0の推進」および「地方創生SDGsの実現などの持続可能なまちづくり」)につながっていくものと考えられる。この点において、第1期から第2期への変化は、大きく飛躍したものというわけではなく、各年版における漸進的な改訂を通じて、ある程度の連続性が保たれているようにみえる。

最後に、「起業」という語にも注目しよう（2014年版：6回、2018年版：46回）。2014年版との比較において、2015年版以降の特徴語として「経営」という語が毎年出現していたが、2018年版においては「起業」も特徴語として出現している。コーケーション統計によれば、「起業」との共起語として「支援」（18回）、「創業」（7回）、「UIJ ターン」（6回）、「教育」（4回）がある。UIJ ターンの促進も織り込んだ、地域における創業支援や起業家教育が強調されている。これは第2期の「稼ぐ地域」に引き継がれていく特徴と推測できる。また、2018年版における「外国人起業家」（35ページ）や「女性起業家」（87ページ）への支援という方針は、第2期（2019年版）の「多様な人材の活躍を推進する」という「横断的な目標」を想起させる内容である。

以上のように、第1期内の改訂でも重要なキーワードが新たに強調されるようになったり、目的や方針が変更・拡大されたりする場合がある。そのため、期をまたぐ改訂にのみ注目しては、それらの変化を捉えきれない可能性があるだろう。

#### e) 2018年版と2019年版(第2期初版)の比較

さて、ここまでは第1期初版（2014年版）と各年の改訂版とを比較してきたが、ここでは第1期最終版（2018年版）と第2期初版（2019年版）を比較する（図-10）。前項 d) で指摘したように、第2期の特徴と思われていた語の一部は、実は2018年版ですでに特徴語として現れていた。よって、2018年版と2019年版を比較すると、両版の特徴語はすなわち一般語となるので、両版の比較は、第2期より強い程度の特徴語を見出す助けとなるだろう。右上隅には「関係人口」とあり、やはりこれが第1期から第2期への変化の大きな特徴と言えるだろう。だが、他にも注目すべき特徴語が見られる。

まず「エネルギー」という語である（2018年版：8回、2019年版：34回<sup>9</sup>）。「SDGs」が特徴語として出ていた2018年版でも、「エネルギー」という語は

出現していたが、第2期ではさらに出現回数が増加している。「エネルギー」の共起語として特に増加したのは「分散型」である（2018年版：2回、2019年版：10回）。「分散型エネルギー」は、再生可能エネルギーなどを活用した地域でのエネルギー事業を指し、2019年版では「エネルギーの真の地産地消」と位置づけられている（62ページ）。2018年版では「分散型エネルギーインフラプロジェクト」を推進している旨が説明されるにとどまっていたが、2019年版では、新たな事業の創出（15ページ）、地域活性化や災害対応（86ページ）といったより広い文脈において言及されている。

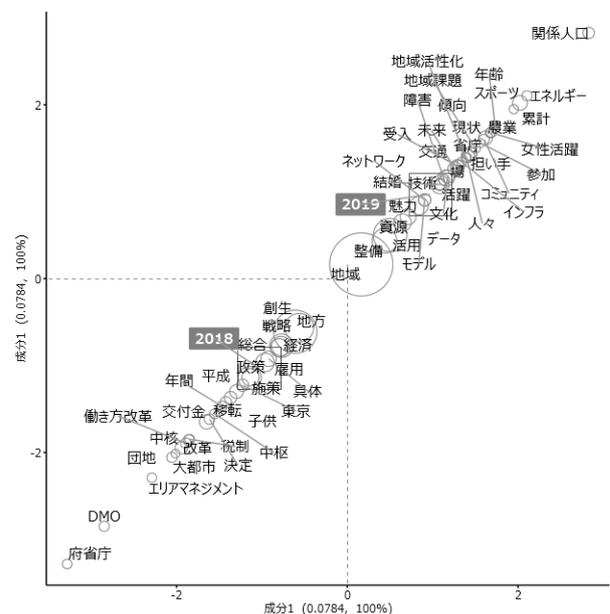


図-10 2018年版と2019年版の比較（対応分析）

次に「スポーツ」がある（2018年版：21回、2019年版：83回）。2018年版（第1期最終版）では特に「基本目標1」の雇用創出や観光強化の文脈でスポーツが言及されている（「スポーツツーリズム」（3回）、「スポーツイベント」（2回））。他方で、2019年版では、「基本目標1」ではなくむしろ「基本目標4」の魅力的な地域をつくるという文脈で、地域の個性や魅力の源泉としてスポーツが位置づけられている（cf. 「スポーツ・レガシー」、60ペー

<sup>9</sup> 「再生可能エネルギー」と「省エネルギー」の出現回数はここには含まれていない。両者を含めると、「エネルギー」の出現回数は、2018年版：15回、2019年版：41回である。

ジ)。具体的には、「プロスポーツチーム」(5回)や「スポーツ施設」(4回)が地域の価値の源泉として挙げられている。これは、「スポーツ」という同一の語でも、強調点や方針が変更された例と言えるだろう。

次に「女性活躍」という語にも注目しよう(2018年版:8回、2019年版:24回)。「女性」単独ならば2018年版:80回、2019年版:91回と大差はないが、「女性活躍」という表現は出現回数が3倍になっている。また、「障害」という語も2019年版の特徴語として現れている(2018年版:15回、2019年版:31回)。第1期でも福祉政策として障害者や障害者世帯への支援が多く言及されていたが、2019年版ではそれに加えて、「(女性)高齢者、障害者、外国人など、誰もが」や「高齢者、障害者、子育て世帯などの多様な世帯」、「年齢や(性別及び)障害の有無等を問わず」といった表現が繰り返し使用されている。これは「多様な人材の活躍を推進する」という第2期の「横断的な目標」と非常に関係が強いと考えられる。

以上のような第2期の特徴語が浮かび上がった一方で、第2期では第1期で強調されてきた「交付金」(68回→27回)、「DMO」(38回→5回)、「エリアマネジメント」(29回→7回)、「働き方改革」(33回→11回)の出現回数が減退していることも明らかになった(図-10左下)。

#### f) 2019年版(第2期初版)と2020年版の比較

最後に第2期初版とその最初の改訂である2020年版を比較しよう(図-11)。2020年版では、コロナ禍の影響が色濃く反映されている。

まず右上隅には「感染症」がある(2019年版:0回、2020年版:58回)。これは紛れもなく「コロナウイルス感染症」のことであり、2020年版での新出語である。新出語でありながら、出現回数も他の特徴語と比べて多く、2020年版における改訂に極めて大きな影響を与えていることが推測される。また、それに付随して「テレワーク」も特徴語として現れている(2019年版:6回、2020年版:41回)。確かに2019年版でも「テレワーク」が言及されているが、コロナ禍に伴う新しい働き方の推進として2020年版ではその強調に拍車がかかって

いる。さらに、「デジタル」という語がある(2019年版:14回、2020年版:41回)。「DX」(デジタル・トランスフォーメーション)というスローガンが新たに掲げられ(2019年版:0回、2020年版:21回)、様々な活動のデジタル化やそれを担う人材育成の推進が強調されている。なお、左下隅を見ると分かるように、「DX」という新出語が強調される一方で「AI」の出現回数はやや減退している(2019年版:21回、2020年版:14回)。以上のことから、期内の改訂であれ、政府が「情勢の推移」を捉えつつ、新たな方針を積極的に提案・強調していることがうかがえるだろう。

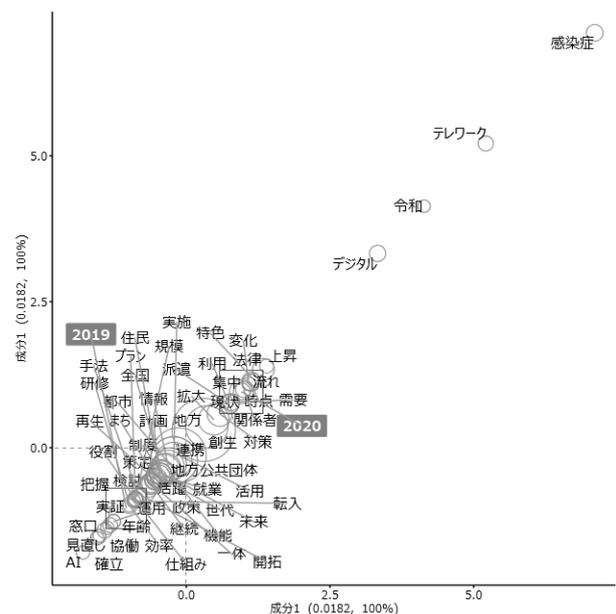


図-11 2019年版と2020年版の比較(対応分析)

## 4. 結語

以上、本稿では2014年から2020年までの国の「総合戦略」をテキストマイニングによって分析してきた。既知の事柄を定量的にも確認することが第一の目的であったが、その過程で留意すべき点もいくつか浮かび上がった。終章では、前章で抽出した語について、各年版での出現回数をヒートマップで表現することで、2014年版から2020年版までのキーワードの変遷を可視化する(図-12)。そしてこのヒートマップを頼りに、本稿の分析で明らかになった点をまとめよう。図-12からは以

	交付金	ローカル	RESAS	DMO	働き方改革	エリア マネジメント	SDGs	未来	関係人口	農業	エネルギー	スポーツ	女性活躍	感染症	テレワーク	デジタル	DX
第1期初版 2014	3	3	0	1	3	0	0	1	0	13	7	8	1	0	9	0	0
2015	31	27	21	23	23	2	0	2	0	18	5	11	7	0	6	0	0
2016	39	29	35	30	34	4	1	23	0	23	5	15	9	0	8	0	0
2017	44	24	23	35	34	35	23	25	2	20	7	22	9	0	7	0	0
2018	68	26	22	38	33	29	45	29	5	23	8	21	8	0	6	0	0
第2期初版 2019	27	14	22	5	11	7	45	65	51	66	34	83	24	0	6	14	0
2020	31	13	29	3	11	7	45	63	54	78	32	87	24	58	41	41	21
合計	243	136	152	135	149	84	159	208	112	241	98	247	82	58	83	55	21

図-12 各年版における出現回数のヒートマップ

(KH Coder 3の分析結果より筆者作成)

下のことが読み取れる。

第一に、各年版において同じ語の出現回数が完全に一致している事例の方が少ないという点である。ほとんどの場合において、改訂とともに語の出現回数変動している。これまで毎年の改訂については比較的軽視される傾向にあったが、毎年の改訂においても、軽微な修正だけでなく内容に大きく変更が加えられていると推測できる。

第二に、第1期から第2期への改訂、すなわち2019年版において急に出現した語がほとんどないという点である。今回の分析で抽出した各年版の特徴語の中で、第1期から第2期への改訂で急に出現した語（新出語）は「デジタル」だけであった。第2期初版（2019年版）で急激に出現回数を増やした語であっても、第1期での各年の改訂においてすでに出現していたことが分かる（「SDGs」、「未来」、「農業」、「スポーツ」、「女性活躍」）。特に第1期から第2期への最も大きな変化とみなされてきた「関係人口」でさえ、実は初出は第1期の2017年版である。このように、第1期における毎年の改訂の延長上に第2期があるのであり、各年の改訂によって強調される内容が漸進的に変化していることが分かる。

第三に、重要と考えられるキーワードの中には、第1期から第2期への改訂ではなく、各年の改訂において新たに出現したものが多くという点である（「RESAS」、「エリアマネジメント」、「SDGs」、「関係人口」、「感染症」、「DX」）。前章の分析でも、初版では構想段階だったものが各年の改訂において

具体名で語られ強調されるようになる、という事例がいくつか確認された。また、国の「総合戦略」が考慮している「情勢の推移」とは、世界的な潮流も含めた国家（マクロ）レベルの情勢という側面が強いように見え、そのような変化に対して組織的な検討や法整備を伴いながら、可能なかぎり迅速に「総合戦略」に取り入れようとした事例も確認された。

そして最後に、「エネルギー」や「テレワーク」のように、かつては強調されていたが一旦はその強調の程度が減退し、その後再び強調される場合があるという点である。確かに、同一の語であっても、強調する文脈や目的が変更されたり広げられたりする事例は前章の分析でもいくつか確認された。

以上のように、国の「総合戦略」では、第1期内の改訂でも重要なキーワードが新たに強調されるようになったり、目的や方針が変更・拡大されたりする場合もある。また、従来、第2期の特徴と考えられてきたキーワードは、第1期内での改訂においてすでにその芽を出していた。そのため、期をまたいだ改訂だけでなく、各年の改訂版にも注意を向ける必要があるだろう。また、自治体の「地方版総合戦略」は5年ごとの改訂が一般的であるように見えるが、その改訂において国の「総合戦略」を勘案するときに、新たな期の改訂版を参照するだけだと、そこで展開されているキーワードが強調されるに至った背景や経緯を捉え損ねる可能性がある。「地方版総合戦略」に取り入れた

いキーワードがあるとするれば、その語が国の「総合戦略」のどの版で使われ始めたのかを確認するのが良いかもしれない。

なお、本稿の目的は「総合戦略」の各年版における変化を捉えることだったため、本稿の分析は、いわば各年版の差異を強調するものであった。逆を言えば、「総合戦略」で中心的一貫して強調されてきた語については分析対象から除外されている。また、出現回数30回以上の高頻出語に絞った上に、各年版の「総合戦略」を2年ずつ分析するという方法をとったため、「総合戦略」の改訂における目立った変化を部分的に明らかにするとどまっている。

「総合戦略」に関して計量テキスト分析をさらに深めていくためには、次のような方向が考えられる。まず、改訂での変化をさらに漏れなく捉えるためには、本稿では扱わなかった出現回数30回未満の低頻出語に焦点を当てて同じく対応分析を行なうことが有益であろう。また、「地方創生」は非常に幅広いテーマ性を持つので、階層的クラスター分析やトピック分析によって「総合戦略」で語られている話題を抽出することも考えられる。さらには、着目したいキーワードを分類した上でコーディング・ルールを作成すれば、「総合戦略」の変遷過程をより詳細に捉えることができるだろう。今後の課題としたい。

## 補節 具体的な手順に関する覚書

### (1) テキストボックスを含んだ文書のテキストファイル化

テキストボックス等の図形を含んだ文書は、PDFからテキストファイル化する際に一工夫が必要である。というのは、PDFファイルからWordファイルに変換する段階では問題はないが、Wordファイルからテキストファイルに変換する際にテキストボックス内の文字が消えてしまうからである。これを克服するための一方法を以下で紹介する。ただし、以下の操作にはPDF編集ソフトウェア「Adobe Acrobat Pro DC」が必要である。

手順	操作
1	整形したPDFファイルを「単一HTMLページ」で書きだす（ブレンテキストで書きだしてもよいが、経験上、行などが入れ替わる等の不具合が生じる場合が多い。その原因は不明である）
2	「単一HTMLページ」を任意のブラウザで開き、テキストを全選択してコピーする
3	コピーしたテキストをWordなどにペーストする
4	目視にてPDFファイルとWordファイルの内容に齟齬がないかを確認する
5	確認後、Wordファイル内の全テキストをコピーし、テキストエディタにペーストする

### (2) ルビを効率的に検索・削除する方法

PDFに埋め込まれたルビは、テキストファイル化した後も文中にそのまま残されてしまう。テキスト分析においてルビは不要なので一括で削除したいが、そうすると「おわりに」などのひらがなだけで独立した部分も削除されてしまう。そこで効率的にルビを検索して削除する方法を以下で紹介する。ただし、この操作には正規表現を使用できるソフトウェアが必要である（本稿では「MIFES10」を使用した）。

入力する正規表現	説明
<code>¥n[あ-ん]{1,4}¥n</code>	改行後の1～4文字までの連続したひらがな

正規表現の先頭の「¥n」は「改行記号」、「[あ-ん]」は「全てのひらがな」（「あ」は小文字）、「{1,4}」は「1文字から4文字まで」（任意で調整する）を意味する。これらを統合した文字列で検索することでテキスト内のルビを効率的に発見することができる。ただし、削除すべきか否かの判断は目視で行なう他ないだろう。

### (3) 不要な改行を効率的に削除する方法

PDFをテキストファイル化したときに、文中で不要な改行が入ってしまう問題を解決する必要がある。1つ1つ削除していく作業には途方もない時間がかかり、逆に置換機能を用いて一括で削除しようとする、必要な改行まで消えてしまう。この問題を改善する一手法を以下で紹介する。ここではMicrosoft Wordなどの置換機能を用いる。

手順	入力欄	入力する記号
1	置換前	^p^p
	置換後	^p
	連続した改行を単一の改行に変換する。検索結果が0になるまで置換を繰り返す	
2	置換前	。^p
	置換後	。☆
	文末（。）のすぐ右にある改行を削除した上でその箇所☆印を残す	
3	置換前	^p
	置換後	★
	文中に残った改行を全て削除した上で、その箇所に★を残す	
4	置換前	★
	置換後	(空白)
	1つつ目視しながら改行の要不要を判断し、不要ならば(空白)に置換する	
5	置換前	★および☆
	置換後	^p
	目視での確認終了後、★と☆を全て改行に戻す	

#### (4) 注釈記号を効率的に検索・削除する方法

最後は脚注記号についてである。脚注記号も文中に紛れてしまうので、テキスト分析での誤差を生じさせる原因の1つとなる。これを1つ1つ削除していくのは途方もない作業になるので、正規表現を用いて効率的に検索する方法を2つ紹介す

#### 参考文献

- 1) Hearst, M. A.: Untangling text data mining, ACL '99, pp. 3-10, 1999.
- 2) 小田切康彦: 政策系大学における研究動向: 論文タイトルを用いたテキストマイニングから、徳島大学社会科学研究所、28巻、pp. 61-82、2014.
- 3) 小野緋莉、松元絵美、徳田光弘: 地方創生政策における構造的矛盾—福岡県及び県下60市町村を対象として、日本建築学会九州支部研究報告集、58巻、pp. 73-76、2019.
- 4) 喜田昌樹: 新テキストマイニング入門: 経営研究での「非構造化データ」の扱い方、白桃書房、2018.
- 5) 小峯敦、仲北浦淳基: ケインズ『一般理論』における訳語の選定～学者訳と日常訳の協働に向けて、龍谷大学経済学部ディスカッションペーパー、2021.
- 6) 小森望、山本彩野、日端康雄: 新聞メディアからみた90年代以降の都市計画・まちづくりの潮流に関する一考察、都市住宅学、55号、pp. 108-113、2006.
- 7) 坂本誠: 地方創生政策が浮き彫りにした国—地方関係の現状と課題—「地方版総合戦略」の策定に関する市町村悉皆アンケート調査の結果をふまえて、自治総研、44巻、474号、pp. 76-100、2018.
- 8) 鈴木雄大郎: 第2期まち・ひと・しごと創生総合戦略を読み解く、大和総研(地方創生・地方経済)、2020.
- 9) 仲北浦淳基: ロバートソンにおける「実物」と「貨幣」—コーディング・ルールの作成と特徴語の抽出、テキストマイ

る。以下では「(1)」のような注釈記号を例にとって説明する。

#### a) 方法1

入力する正規表現	説明
¥(.+?¥)	( )で囲まれた1文字以上の文字列

まず正規表現の先頭の「¥」は右の記号が正規表現ではないことを意味する。「(」は正規表現としての意味を持っているので、左側に「¥」をつけることで「(」が正規表現ではなく文字列であることを示す必要がある。これは末尾の「¥)」も同様である。次の「.+」は「1文字以上」、「?」は「最短の一致」を意味する。「?」がないと非常に離れた文字列まで検索対象になってしまう。これらの文字列で検索することで、( )に囲まれた1文字以上のテキストを検索することができる。

#### b) 方法2

入力する正規表現	説明
¥([0-9]{,2}¥)	( )で囲まれた2桁以下の数字

先述のとおり、「¥(」と「¥)」は文字列としての「(」と「)」を意味する。次の「[0-9]」は「全ての数字」、「{,2}」は「2桁以下」を意味する。これらを統合した文字列で検索することで、( )で囲まれた2桁以下の数字を検索することができる。

ニングから読み解く経済学史、ナカニシヤ出版、pp. 147-165、2021.

10) 仲北浦淳基、小峯敦：マクミラン委員会におけるロバートソンの経済観：テキストマイニングから判明する「実物」と「貨幣」、龍谷大学経済学論集、58巻1号、pp59-100、2019.

11) 中澤高志：「地方創生」の目的論、経済地理学年報、62巻4号、pp. 285-305、2016.

12) 中嶌いづみ：自治体総合計画の意義と課題：基礎自治体（市区）における総合計画の位置づけの分析を中心に、法政大学、博士論文（公共政策学）、2018.

13) 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：KH Coder OFFICIAL BOOK、ナカニシヤ出版、2020

14) 増田正：わが国地方議会における「会派」の政策関心の可視化：自己組織化マップによる計量テキスト分析、地域政策研究、23巻4号、pp. 19-36、2021.

15) 松井望：「基本方針による管理」と計画化：総合戦略と総合計画を事例に、公共政策研究、17巻、pp. 40-51、2017.

16) 森幹彦：地域課題の解決に向けた議論のトピック分析、知識共創、9巻、pp. (IV3)1-(IV3)8、2019.

17) 森川洋：地方創生政策とその問題点、人文地理、72巻3号、pp. 299-315、2020.