

平成29年度 成績概要書

課題コード（研究区分）：1106-118571（戦略研究）

1. 研究課題名と成果の要点

- 1) 研究成果名：テキストマイニングによる農業・農村づくりに係る先進地視察対象の選定
（研究課題名：農村集落における生活環境の創出と産業振興に向けた対策手法の構築）
- 2) キーワード：テキストマイニング、事例集、文章データ、キーワード、出現頻度
- 3) 成果の要約：農業・農村づくりの先進地視察に係る対象選定に向けて、事例集等の文章データからキーワードを抽出し、その出現頻度と取組内容等の事例の属性との対応関係と紹介文の因果関係を整理するテキストマイニングの実施手順を確立した。これにより、視察の目的に合致した対象の選定が可能になる。

2. 研究機関名

- 1) 担当機関・部・グループ・担当者名：中央農試生産研究部生産システムG 研究主任 日向貴久
十勝農試研究部生産システムG、根釧農試酪農研究部地域技術G

- 2) 共同研究機関（協力機関）：

3. 研究期間：平成27年度-平成31年度（2015-2019年度）

4. 研究概要

1) 研究の背景

自治体が地域づくり等の計画を策定する際に参考となる先進事例は全国に多数存在し、地域内で様々な活動が展開されることから、視察を企画する農業関係者の多くが、適切な視察対象の選定に苦労している。

2) 研究の目的

集落営農等の事例集に掲載された文章データからキーワードを抽出し、取組内容や受賞時期等、事例の属性とキーワードの出現頻度との対応関係や文章の因果関係を整理することが可能なテキストマイニングの実施手順を確立することで、先進地視察の対象選定の実務に資する。

5. 研究内容

1) テキストマイニングの実施手順の整理（H27～29年）

- ・ねらい：集落営農の事例集に掲載された文章データを対象にテキストマイニングを実施し、多数ある事例の中から調査候補を選定する。これにより農業関係者が実践可能なテキストマイニングの手順を整理する。
- ・試験項目等：分析対象 農林水産省「集落営農の組織化・経営発展の取組事例」に記載された125事例 磯島の成果*を参考にしたテキストマイニング

2) テキストマイニングの実践（H27～29年）

- ・ねらい：農林水産祭むらづくり部門の受賞理由に関する文章データを対象にテキストマイニングを実施するとともに、実際に現地を訪問しテキストマイニングで整理された情報が視察にもたらす効果を把握する。
- ・試験項目等：平成9～26年度に受賞した51事例（事例当たり1,032字、接続詞・助詞を除く175語）出現頻度と出現時期に関するコレスポネンス分析、ISM法による文章情報の視覚的な整理 検証調査の実施によるテキストマイニングの効果の確認（H24年度農林漁業振興会長賞受賞事例）

6. 成果概要

- 1) 集落営農の先進事例集を題材に、先進地視察の対象選定に役立つテキストマイニングを確立した。本テキストマイニングはキーワード抽出、コレスポネンス分析、ISM法により構成される（図1）。分析では、事例集等、複数の事例に関する文章データを用い、Excel、フリーソフト、汎用的な統計解析ソフトを使用する。
- 2) 具体的な解析手順は以下の通りである（図1）。事例集等の記載文章をテキストデータに加工し、形態素解析ツール「茶筌**」により形態素に分解した上で、Excelに転記する。複数の分析者により同じ意味をなす形態素をコンセプトとしてまとめ、文章中のキーワード候補を抽出する。キーワード候補の出現頻度を集計するため、出現の有無を1・0で示した2値データ化した後、全体に占める出現率等を基準にキーワードを確定させる。最終的には、事例の属性とキーワードの出現頻度との対応関係をコレスポネンス分析により可視化するとともに、文章データの因果関係をISM法により視覚的に整理する。
- 3) 事例集に掲載された125事例の活動概要に関する文章データについてキーワードを抽出した上で、事例の属性として事例集内で分類された取組内容（図2凡例の①～⑨）とキーワードの出現頻度との対応関係を整理した（図2）。先進事例間で共通なキーワードは、図の中央に位置する「水稲」、「麦類」、「大豆」、「作業の省力・効率化」であり、土地利用型作物における作業の効率化に関するワードである。取組内容のうち、①共同化・共同利用と②新技術・経営管理の得点は近似し、両者で出現するキーワードは類似しているといえる。ただし、これらの取組みでは、規模拡大に関するキーワードと投資の抑制に関するキーワードの双方が出現しており、共同化・共同利用に取組む先進事例の視察に際しては、規模拡大と投資の抑制に関連するキーワードで区別した上で、特徴的なキーワードが出現する事例を抽出できる。
- 4) 規模拡大に関連するキーワードが出現し、共同化・共同利用に取組むA営農組合を例に、ISM法により文章の因果関係を視覚的に整理した（図3）。A営農組合は、ブロックローテーションによる土地利用調整を図るとともに、社会保険の充実等を目的にオペレータ組織を分社化することで、農作業受託により規模拡大を図り、大型高性能機械の共同利用を通して低コスト化を実現した事例である。このように、A営農組合は、共同化・共同利用に特徴的なキーワードである低コスト化を実現した事例であることが把握できた。
- 5) 農林水産祭受賞事例の検証調査を通して、テキストマイニングは調査対象の選定に役立つとともに、キーワードの出現時期や頻度と関連づけた考察を可能とし、事例の活動を評価する上で有効であることを確認した。

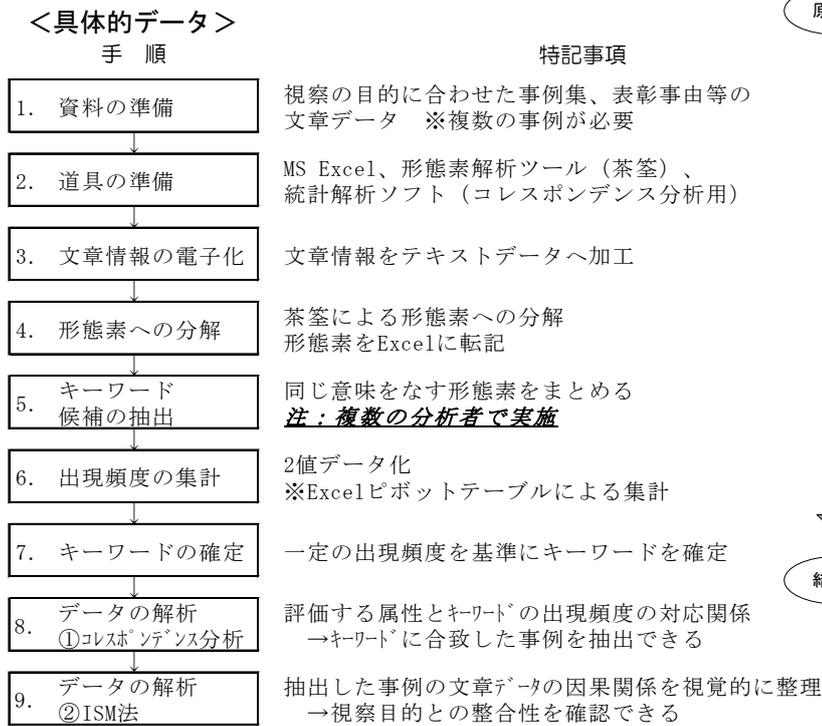


図1 テキストマイニングの実施手順

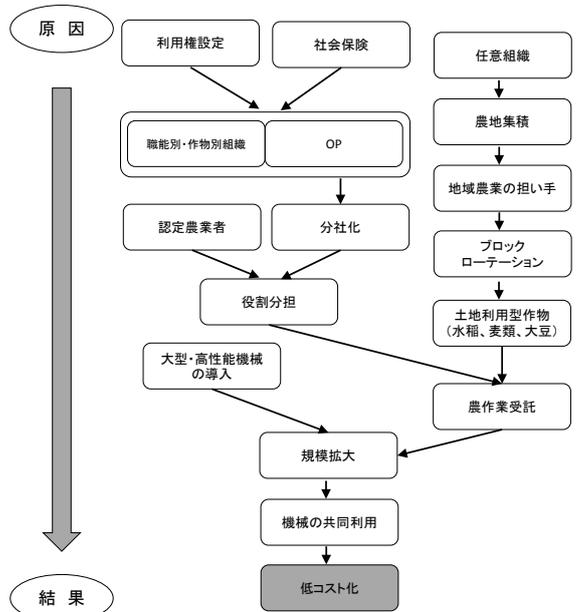


図3 文章の因果関係の視覚的な整理 (ISM法)

用語：ISM法 (Interpretive Structural Modeling)

ISM法は、問題の全体像を把握できるようにするための手法であり、要素間の因果関係を視覚的に整理できる手法である。

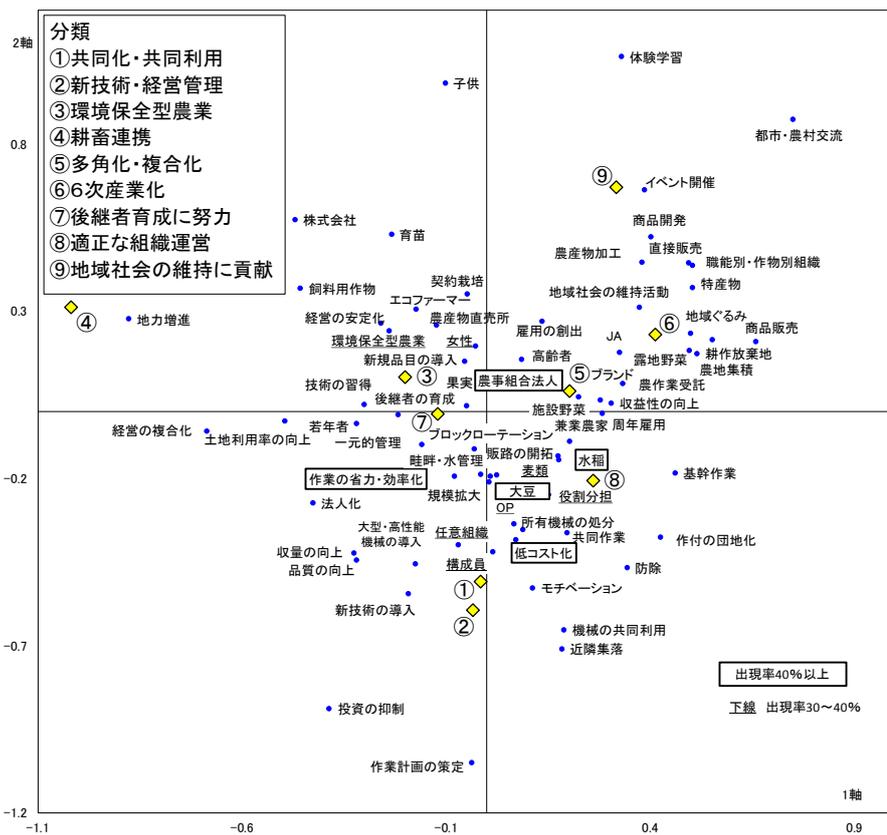


図2 事例の取組内容とキーワードとの対応関係 (コレスポデンス分析)

注1) 品詞に分解した後 (事例当たり 236 字、接続詞・助詞を除く 48 語)、類似した品詞をまとめた上で、その出現頻度を計測した。なお、品詞分解には、「茶筌」を使用した。

注2) 10 事例以上で出現したキーワード 74 語を解析の対象とした。

注3) 図2は、コレスポデンス分析により、属性として設定した取組内容とキーワードの出現頻度の対応関係を可視化している。

注4) 図中の「◇」印は事例集の分類による取組内容であり、「●」印はキーワードである。図中の「◇」印の取組内容に近い「●」印のキーワードは、その取組内容で特徴的なキーワードであることを意味する。

用語：テキストマイニング

アンケート調査の自由回答欄などのテキストデータを分析し、有益な知識や情報を取り出す技術の総称である。市販の分析用ソフトウェアは高額なものが多い。

用語：コレスポデンス分析

アンケート等のクロス集計結果を散布図で視覚的に表現できる手法。散布図では、回答傾向の類似しているものが近くにプロットされる。分析には、汎用的な統計解析ソフトを用いる。

*磯島昭代「テキストマイニングを用いた米に関する消費者のアンケート解析」『農業情報研究』15(1)、2006

**奈良先端科学技術大学院大学情報科学研究科自然言語処理学講座『茶筌 (ChaSen)』<http://chasen-legacy.osdn.jp/>

7. 成果の活用策

1) 成果の活用面と留意点

- ・本成果は、自治体実施する地域づくりに係る先進事例調査において、対象選定の実務に資する。
- ・フリーソフトのPCへの導入に際しては、各自治体の取り決めに準拠する。

2) 残された問題とその対応

8. 研究成果の発表等

白井康裕他「テキストマイニングによる自治体取り組み先進地視察の支援」北海道農業経済学会、2017