

卒業論文のテーマに何が選ばれてきたか

—現代社会学部の卒業論文題目の語彙分析から—

丸 野 由 希

要 旨

京都女子大学現代社会学部では、変化していく現代の問題について学びを深めるため、学生の提出する卒業論文の題目に大きな多様性が表れていると考えられる。本稿では、現代社会学部の設立以来、これまでに提出された計12年分の卒業論文題目について語彙分析を行い、現代社会学部の学生たちがどのような卒業論文のテーマに取り組んできたかについて考察する。

キーワード：現代社会学部、卒業論文、語彙分析、ワードクラウド、ネットワーク分析

はじめに

京都女子大学現代社会学部（以下、現代社会学部）では、多様化、多層化、複雑化した「現代社会」に関する幅広い知識を身につけ、社会問題を多角的に認識する方法を学ぶ。ダブル・クラスター制のもと、学びの領域として「人間・文化」「家族・地域社会」「政治・公共政策」「経済・経営」「情報・環境」の5つのクラスター（表1）から興味や関心のある2つのクラスターを選び専門分野に学び進め、その集大成として卒業論文に取り組む。

表1 現代社会学部の5つのクラスター（2015年度カリキュラム）

クラスター	説 明
人間・文化	人間の精神・こころ・身体・生命のあり方やメディア・宗教・スポーツなどの文化的営みを、多彩な切り口で探求する。
家族・地域社会	家族・ジェンダー・地域社会を社会学的なアプローチで探求し、次代の新しい家族像・地域社会のあり方を探る。
政治・公共政策	政治学・行政学・法学分野の科目をバランスよく学び、良き市民として生きる力を養う。公務員志望者にも対応。
経済・経営	経済学・経営学・会計学を包括的に学ぶことによって、ビジネス社会で必要とされる知識や考え方、提案力を培う。
情報・環境	高度情報社会を支える情報科学・情報工学の知識とスキルを修得し、あわせて学際的知識を活かして環境問題を考える力を養う。

本研究では、現代社会学部創立から平成26年度までの計12年分の卒業論文題目データを集積し、分析を行った。探索的分析や推測分析など様々な分析方法があるなかで、頻度分析とネッ

トワーク分析を、年度ごと、指導教員の専門分野ごとに行った。指導教員の専門分野はクラスターと同様に「人間・文化」「家族・地域社会」「政治・公共政策」「経済・経営」「情報・環境」の5つに分類した。専門分野ごとの分析には、ゼミの指導教員の情報が入手できた平成21年度～平成26年度までのデータのみを使用した。

I 分析データ・手法

1. 使用したデータ

本研究では、現代社会学部の学生の卒業論文の題目、副題、指導教員の3項目をデータとして使用した。分析対象は、現代社会学部が設置されてからはじめての卒業生となる平成15年度から平成26年度までの計12年分である。指導教員については、はじめの数年間のデータが残っていなかったことから、平成21年度から平成26年度までの計6年分のデータのみを用いた。表2は年度ごとの卒業論文提出人数である。

表2 年度ごとの卒業論文提出人数

年	人数	年	人数
平成15年度	274人	平成21年度	269人
平成16年度	253人	平成22年度	293人
平成17年度	232人	平成23年度	271人
平成18年度	253人	平成24年度	256人
平成19年度	257人	平成25年度	254人
平成20年度	256人	平成26年度	274人

2. 形態素解析による品詞情報の抽出

本研究では、卒業論文の題目、副題に用いられた名詞に着目した。そこで、題目、副題を形態素解析にかけ、名詞のみを抽出した。形態素解析とは、文を単語単位に切り分け、品詞の情報などを加えることを指す¹⁾。例えば、「すももももももものうち」という文を形態素解析にかけると以下のようなになる(図1)。

```

すもも 名詞、一般、*、*、*、*、すもも、スモモ、スモモ
も      助詞、係助詞、*、*、*、*、も、モ、モ
もも    名詞、一般、*、*、*、*、もも、モモ、モモ
も      助詞、係助詞、*、*、*、*、も、モ、モ
もも    名詞、一般、*、*、*、*、もも、モモ、モモ
の      助詞、連体化、*、*、*、*、の、ノ、ノ
うち    名詞、非自立、副詞可能、*、*、*、うち、ウチ、ウチ

```

図1 MeCabの解析

1) 金明哲「テキストデータの統計科学入門」岩波書店、2009

形態素解析エンジンはいくつか公開されているが、本研究ではMeCabを用いた。MeCab（和布蕉）は、京都大学情報科学研究科とNTTコミュニケーション科学基礎研究所の共同研究ユニットプロジェクトを通じて開発されたオープンソース形態素解析エンジンである²⁾。

3. 頻度分析を可視化するためのワードクラウド

頻度分析は、ある単語がテキストの中でどのくらいの頻度で出現するのかを分析することである。テキスト中の単語の出現頻度は数字を並べた表で表現されることが多いが、全体の傾向を直感的に捉えにくい。そこで、文章中で出現頻度が高い単語を複数選び出し、その頻度に応じた大きさで図示する手法であるワードクラウドを用いた³⁾。本研究では、形態素エンジンMeCabをRから使うためのRのパッケージであるRMeCabと、ワードクラウドの手法を使って単語の頻度分析を視覚的に表現するためのRのパッケージwordcloudを用いて、年度ごと、専門分野ごとに頻度分析を行った。

4. ネットワーク分析

ネットワーク分析は、社会学や通信ネットワークなどの分野で多く用いられており、グラフ理論を基礎に置いている。グラフ理論では、ものどもの関係を頂点同士結ぶ線分によって表した図形をグラフと呼び、頂点の間に向きがあるグラフを有向グラフ、向きがないものを無向グラフと言う¹⁾。テキスト解析において、特定の複数の単語の組が同じテキストの中に現れることを「共起」と呼び、共起関係について調べることは重要な分析方法である。ネットワーク分析では、二つの語の共起関係をネットワーク図と呼ばれるグラフを使って視覚的に表現することで、分析対象のテキストについて有用な知見を得ることができる。本研究では、Rのパッケージであるigraphを用いて、4年ごとに区切った場合、指導教員の専門分野ごとの共起関係をネットワークマップで表し、ネットワーク分析を行った。MeCabの辞書を用いて形態素解析を行うと、例えば「経済学」などの言葉は「経済」と「学」に分かれるが、このままでネットワーク分析を行うと「学」の文字が「経済」、「経営」、「政治」などを結びつけるハブとなってしまい、正しい関係性が見失われてしまう。そこで「経済学」などを一つの名詞として辞書に登録することで問題を回避した。

2) MeCab: Yet Another Part-of-speech and Morphological Analyzer (mecab.googlecode.com/svn/trunk/mecab/doc/index.html) 2016/01/06閲覧

3) 「ワードクラウド（ワードクラウド）とは—コトバンク」、
(<https://kotobank.jp/word/ワードクラウド-674221>) 2016/01/12閲覧

II 頻度分析

1. 年度ごとにみた頻度の変化

年度に題目、副題をRMeCabで形態素解析し、名詞の出現回数を数えた。昇順に並べてグラフ化したものを図2.1 - 2.12に示す。このとき、1文字の単語は意味をなさないものが多かったため除外した。また、単語の語数上、図に表示されているのは15回以上使われている単語のみである。

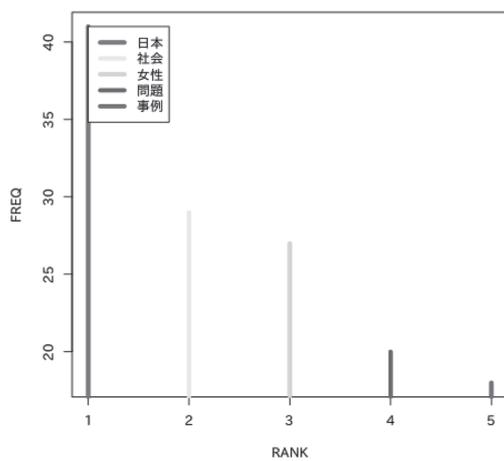


図2.1 平成26年度の場合

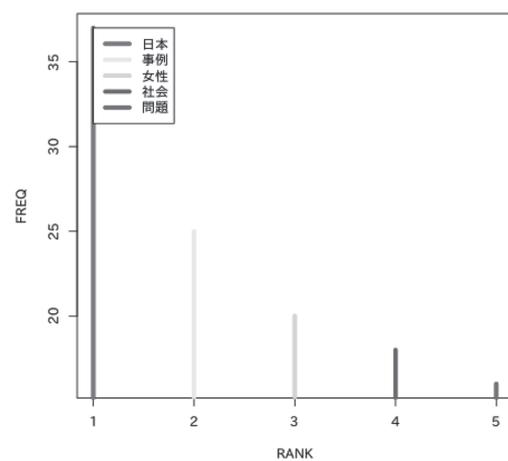


図2.2 平成25年度の場合

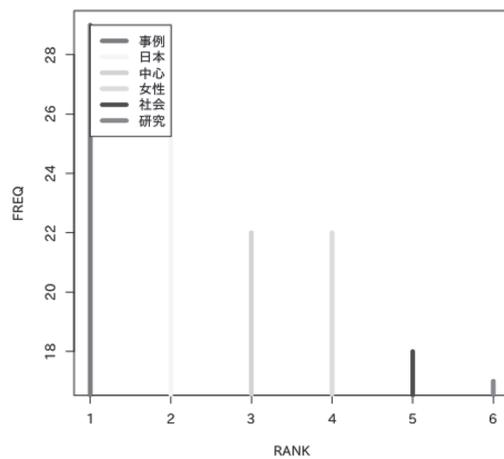


図2.3 平成24年度の場合

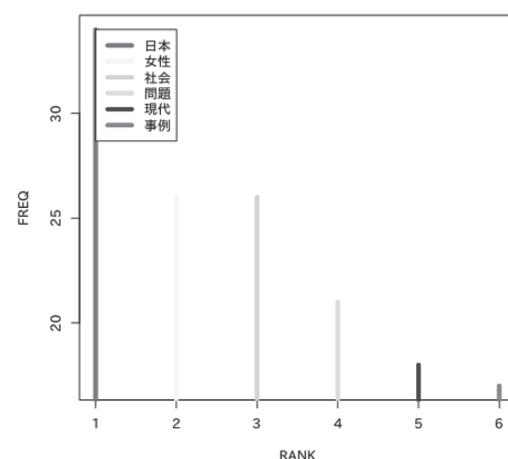


図2.4 平成23年度の場合

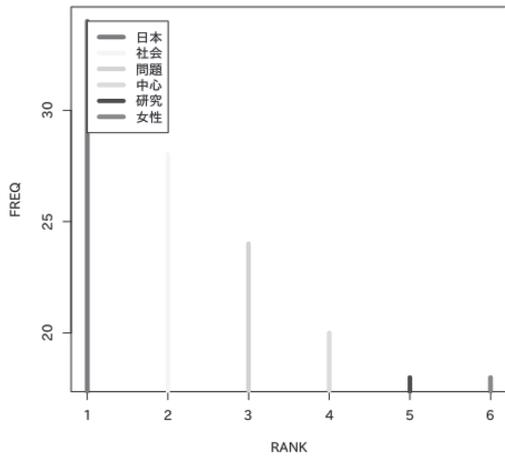


図 2.5 平成22年度の場合

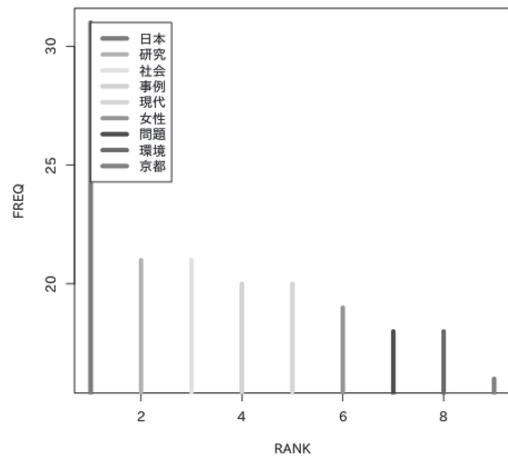


図 2.6 平成21年度の場合

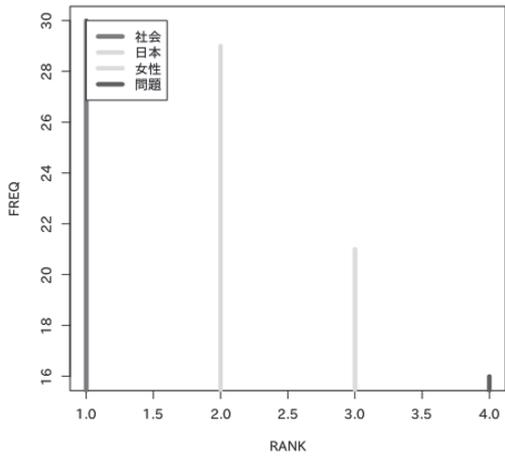


図 2.7 平成20年度の場合

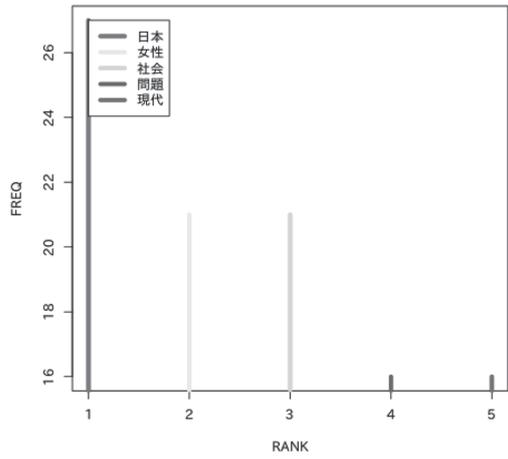


図 2.8 平成19年度の場合

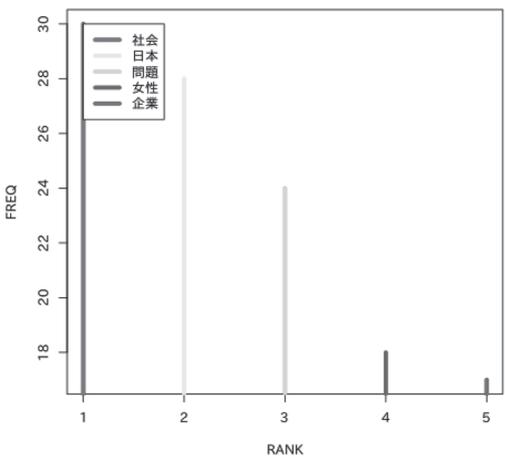


図 2.9 平成18年度の場合

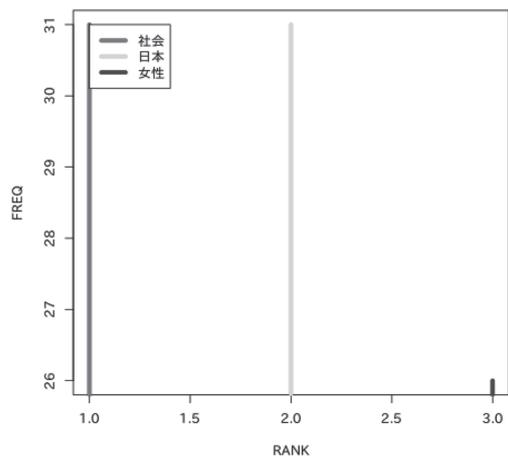


図 2.10 平成17年度の場合

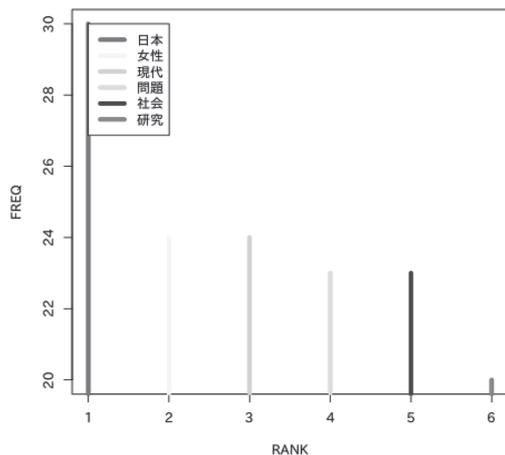


図 2.11 平成16年度の場合

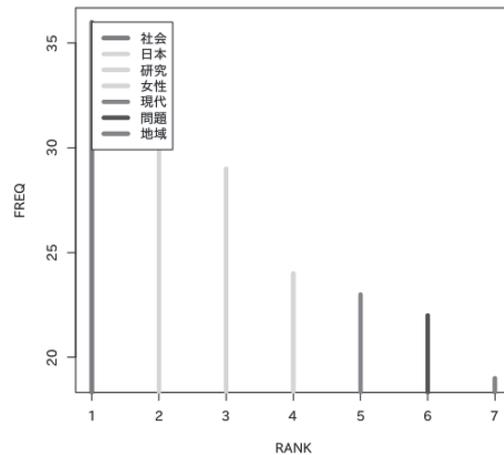


図 2.12 平成15年度の場合

年度ごとに見た場合、出現回数が上位の語彙にはあまり変化がみられず、ほとんどの年度で「女性」「日本」「社会」「問題」などの言葉が上位を占めた。

2. ワードクラウドで表現した頻度の変化

出現頻度の少ない名詞にも着目するために、ワードクラウドを用いて分析を行った結果を図 2.13- 2.14に示す。文字の大きいものほど出現頻度が高くなっている。



図 2.13 平成25年度の場合

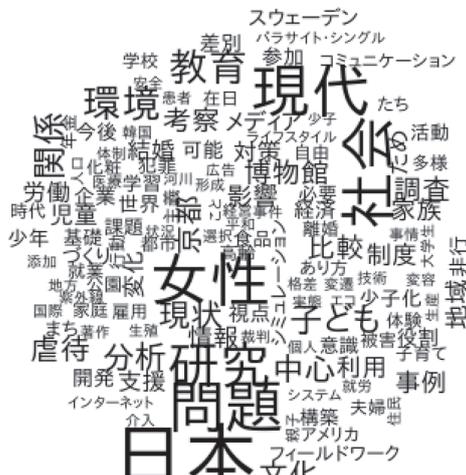


図 2.14 平成16年度の場合

平成25年度（図 2.13）には、「原発」「被害」「福島」「災害」などの文字がみられる。これらの単語が使われている題目、副題をもとのデータから検索すると、以下のような題目、副題がみられた。

- 原発災害後の長期汚染地域における家屋環境での空間線量率の低減方策に関する研究
- デマ災害における「情報」の重要性～福島原発災害におけるデマ・風評被害をなくすために～
- 福島原発事故と原子力損害賠償訴訟法を基礎とする関西での被害予測
- 原発災害後の復興における福島県産品販売の現状に関する研究
- 原発ゼロは可能かつ正しい選択か
- 世代を超えて伝わる戦争～ベトナム枯葉剤被害者を事例に～
- いじめの発生誘因と被害後の長期的影響について

このことから、福島県の原発問題を題材に卒業論文を執筆した人が多いことが分かる。ゼミを選択するのは論文執筆のおよそ2年前からのため、平成23年度近辺の出来事に着目すると、平成23年3月11日に東日本大震災が起きている。この影響で、平成25年度に「原発」「被害」「福島」「災害」などの言葉が例年より多くみられるのではないかと考える。

次に平成16年度（図2.14）をみると、「食品」という文字がみられる。もとのデータから「食品」という単語を検索してみると以下のような題目、副題がみられた。

- サプリメント（栄養補助食品）、必要性の是非
- 食品添加物の必要性～絶対安全を求める消費者～
- 食品添加物とその表示について～食品添加物から見る食の安全性～
- 食品公害事件とこれからの対策

卒業論文執筆2年前の平成14年度近辺の出来事に着目すると、平成13年に食品偽装事件のきっかけにもなった雪印牛肉偽装事件が起きている。この影響で、食品公害事件や食品添加物に関する題目で執筆した人が平成16年度に多く、「食品」という言葉が例年より多くみられるのではないかと考える。

3. 指導教員の専門分野ごとにみた頻度の変化

次に指導教員の専門分野ごとに頻度解析を行った。年度ごとと同様に、題目、副題をRMeCabで形態素解析し、名詞の出現回数の集計を行った。昇順に並べてグラフ化したものを図2.15～2.19に示す。このとき、単語として分類されたもののうち一文字のものは、名詞として意味をなさないものが多かったため除外している。

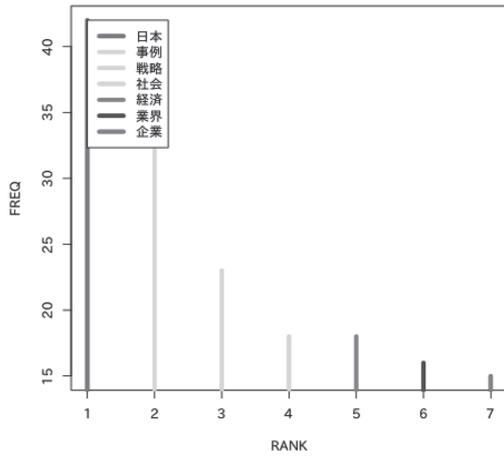


図 2.15 経済・経営

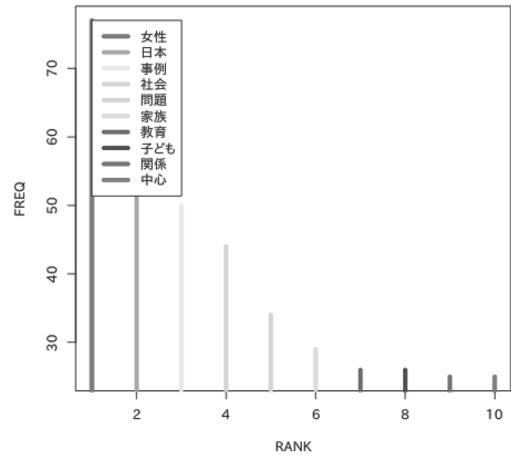


図 2.16 家族・地域社会

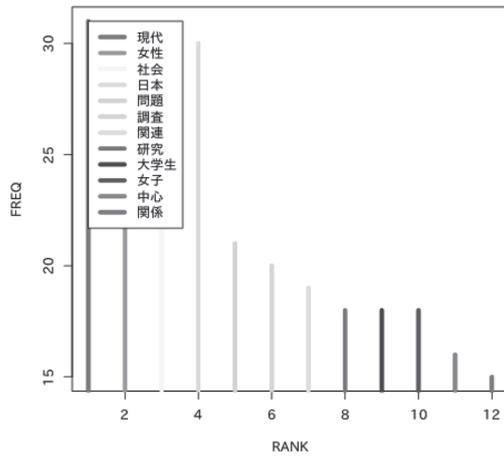


図 2.17 人間・文化

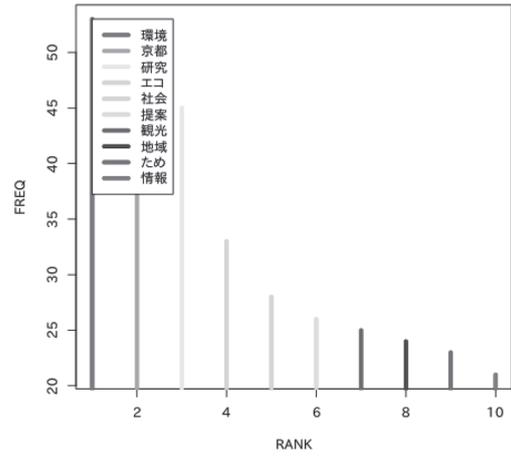


図 2.18 情報・環境

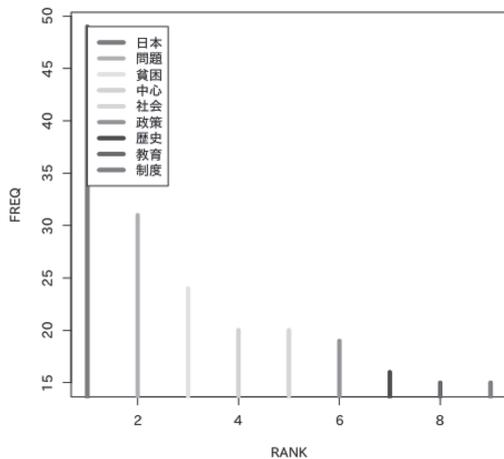


図 2.19 政治・公共政策



図 2.24 政治・公共政策

頻度が少なく小さな文字で書かれた言葉に着目と、経営・経済では「マーケティング」「リーダーシップ」「モチベーション」「マネジメント」など経営に関する言葉が多くあげられている。家族・地域社会では、地域を表す「都道府県」「商店街」や家族に関係する「主婦」「母子」「保育」などの言葉が、人間・文化では「ポジティブ」「いじめ」など人に関する言葉があげられている。情報・環境では、「システム」「アプリケーション」「シミュレーション」、政治・公共政策では「格差」や様々な国の名前が表れ、強く専門分野を表している。このようにワードクラウドで表した図からも、専門分野ごとに多様性ととんでいることがわかる。

Ⅲ ネットワーク分析

次に名詞同士の関係性に着目してネットワーク分析を行った。

1. 年度ごとにみた場合

分析結果をより分かりやすくするため12年分のデータを、平成15年度～平成18年度、平成19年度～平成22年度、平成23年度～平成26年度の4年ごとに区切り、ネットワークマップを作成した。その結果を図3.1～3.3に示す。語数の関係上、他の言葉と3回以上の共起関係が現れた言葉のみを出力している。

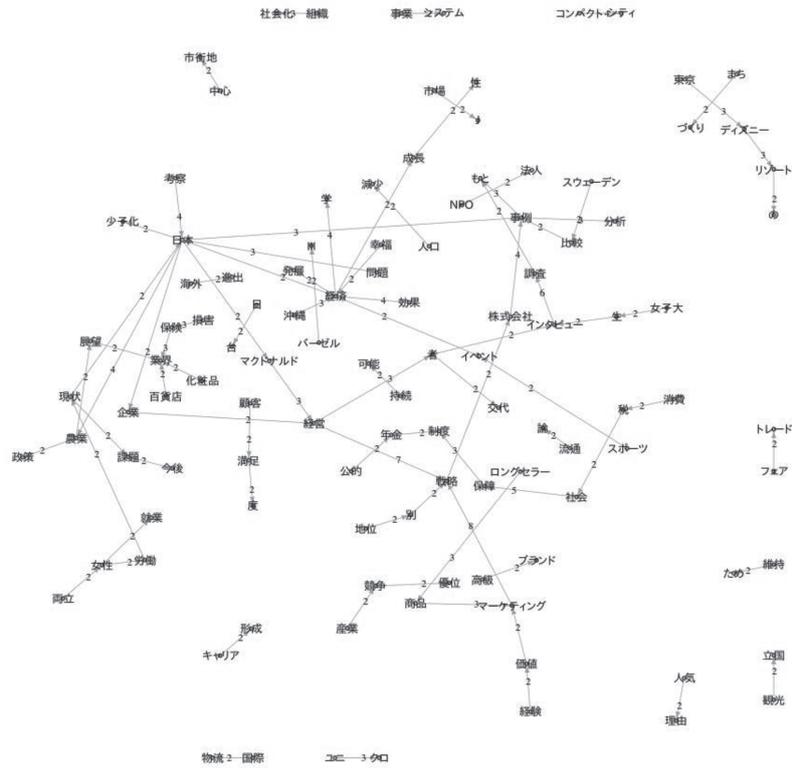


図3.4 経営・経済

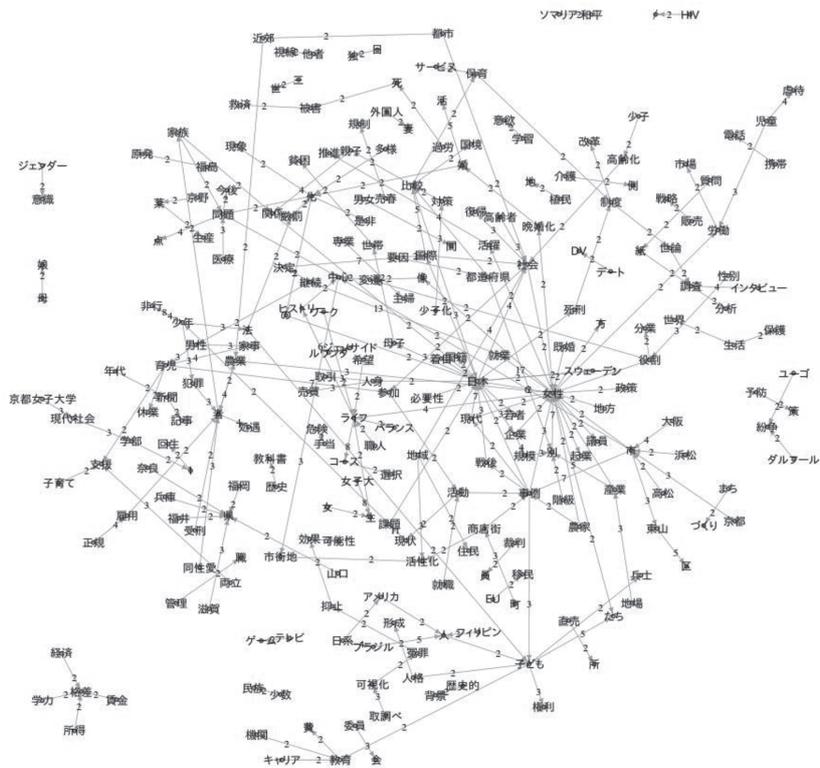


図3.5 家族・地域社会

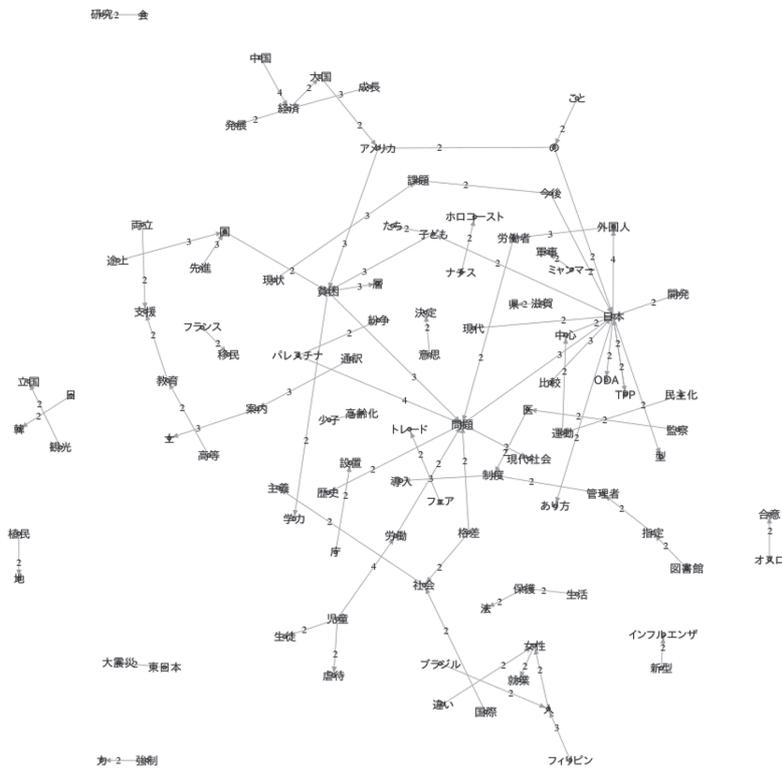


図 3. 8 政治・公共政策

ワードクラウドの時と同様に、ネットワークマップにおいても、それぞれの専門分野に大きく違いが表れている。経営・経済では「日本」「経済」「事例」などの言葉が中心になっており、「業界」や「戦略」など様々な言葉が飛び交い、経営・経済の特長が顕著に表れている。家族・地域社会では「日本」「女性」に多く矢印が向いているが、その次に「ライフ」「子ども」「社会」などに多く矢印が向いており、他の語彙との関係性が高いことを示している。家族の中における子どもや、家庭を持った時の自分のライフコースなどに関心を持つ人が多いことがわかる。また、他の専門分野と比べて「県」や「市」のまわりに様々な地域名があることも、この分野の特長の1つといえる。人間・文化では「日本」「社会」「調査」に加え、「医療」の周りにも多く語彙がとりまいている。また全体をみてみると、「妊娠」「死」「移植」など生殖や生命に関する言葉が多くみられる。このことから人間・文化では、生殖や生命系の分野に関心を持つ人も多いことがわかる。情報・環境では、その名の通り「環境」や情報系の卒業論文を書く上で多い「研究」「提案」の他にも「京都」「観光」などの周りに多く言葉がとりまいており、こちらも顕著に専門分野の特性が表れている。政治・公共政策では、政治に関する言葉が多くみられる。外国の地名も多いが、「日本」「問題」の周りに多く言葉を取りまいており、日本のことはもちろん、世界のあらゆる問題に関心を持っているのだろう。このことから、同じ現代社会学部であっても、専門分野によって大きな違いが生じていることがわかる。

おわりに

現代社会学部の卒業論文の題目の頻度分析、ネットワーク分析を行った結果、現代社会学部の学生は、同じ学部にも所属しているにも関わらず、専門分野ごとに大きな違いがみられ、多様性にとんだ題目で卒業論文を執筆していることがわかった。また、現代社会学部を標榜していることにも見られるとおり、常に新しい問題やその年に大きな問題となったものに関心を持ち、その時代に応じた内容で卒業論文を執筆していることがわかった。今後は、名詞だけではなく他の品詞にも着目し、さらに分析を進めたい。

謝辞

現代社会学部卒業生の奥平美貴さんには、平成21年度から平成26年度までの6年分の紙媒体で管理されていた卒業論文題目の電子化を行っていただきました。このデータによって、卒業論文題目の解析および論文作成に着手することが可能となりました。この誌上をもちまして多大なるご貢献に感謝の意を申し上げます。