

## 科学研究費助成事業 研究成果報告書

令和 2 年 6 月 29 日現在

機関番号：34504

研究種目：基盤研究(B) (一般)

研究期間：2015～2019

課題番号：15H03389

研究課題名(和文)ビッグデータを用いた女性の就業継続と管理職昇進の阻害・促進要因分析

研究課題名(英文)Analysis using Big Data of Factors that Contribute to Women's Work Continuation and Promotion

研究代表者

大内 章子(OUCHI, Akiko)

関西学院大学・経営戦略研究科・教授

研究者番号：20335110

交付決定額(研究期間全体)：(直接経費) 12,100,000円

研究成果の概要(和文)：本研究では、女性労働に関する政策が必ずしも一般の人々の意見を反映していないのではないかと問題意識の下、人々が日常考え、感じていることを細かく捉えることを目的に、Twitter、本研究独自開発のSNS、ヤフーニュースという3つのソーシャルネットワークの大量のデータを用いて解析した。その結果、育児休業など女性労働についての人々の意見/評価を人による解釈の助けを得ながら自動要約ができるようになり、また、人のつながりをベースにしたネットワーク解析にて重要なユーザ・意見のある程度把握することができた。並行して、従来型のアンケート調査やインタビュー調査で、就業継続や管理職昇進の阻害要因を明らかにした。

研究成果の学術的意義や社会的意義

ソーシャルネットワークが普及し、人々が意見/評価を気軽に発信することができるようになった現代においては、意見のビッグデータを網羅的・大局的に解析することで、公共政策や労働問題などの社会問題の改善に反映させることができると考えられる。本研究では、人々の意見/評価をある程度自動要約でき、また重要なユーザ・意見のある程度把握する方法を見出した。今後、自動要約の更なる改善により、女性労働について一般の人々(男女)の間で何が問題となっているかを把握することができ、従来型のアンケート調査やインタビュー調査も並行して利用していくことで、女性労働問題を解決し政策に反映させることができると期待できる。

研究成果の概要(英文)：The purpose of this research is to grasp details of what people think and feel on a daily basis, under the awareness that policies on women's labor do not necessarily reflect the opinions of the general public. We analyzed using a large amount of data from three social networks, Twitter, SNS which was originally developed by this research, Yahoo News. As a result, we became able to automatically summarize people's opinions/evaluations about female labor such as maternity leave with the help of human interpretation, and to understand important users/opinions to a certain extent by network analysis based on human connections. At the same time, conventional questionnaire surveys and interview surveys revealed the impediment to continued employment and promotion of managers.

研究分野：経営学

キーワード：女性労働 ビッグデータ SNS 女性活躍推進 ネットワーク解析

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

1986年の男女雇用機会均等法施行以降、各種法律が整備され、女性の雇用環境は改善してきた。しかし、国際的にみると、女性管理職比率は米国その他の先進諸国で30~40%を超えているのに対して、日本は約11%で低い。女性が管理職に昇進するためには就業継続が最低限必要となるが、第1子出産を機に有職女性の3分の2が退職しており、出産・育児を経て就業継続することはいまだに困難である。

これまでの研究代表者の研究などから、女性の就業継続や技能形成・昇進は、就業意識や昇進についての「個人の意識」、結婚・出産経験の有無、配偶者や第三者の協力などの「家族状況」、企業の育成・活用方法やワーク・ライフ・バランス支援策の整備といった「職場・仕事状況」の3要因が複雑に絡んでいることが明らかになっている。一方、研究代表者らは、実際に女性の育児休業復帰、再就職・起業の支援や女性リーダー養成を行う中で、女性労働に関する政策が必ずしも一般の人々の意見を反映したものではないという問題意識を持っていた。

ソーシャルネットワークが普及し、人々が意見/評価を気軽に発信することができるようになった現代において、意見のビッグデータを解析することで、公共政策や労働問題などの社会問題の改善に反映させることができると考えられる。そこで、本研究では、一般の人々が手軽に自分の意見を発信できるTwitterを代表とするソーシャルネットワークの大量のデータを扱い、ユーザどうしの繋がりを捉えることにし、特に、女性労働の分野において、現場で起きている事象を細かく捉えることで、政策と現場の意思決定との乖離を少しでも狭めることができると考えた。

### 2. 研究の目的

本研究では、Twitterなどのソーシャルネットワークのログデータを利用して、以下に示す3点を明らかにすることを目的とする。

- 1) 女性労働について、一般の人々(男女)の間で何が問題となっているかを把握する。
  - 2) 国の政策、専門家、メディアの論評に対する人々の意見/評価を自動要約する。
  - 3) 他者との繋がりを表したネットワーク構造の変化から重要なユーザ・意見を把握する。
- 一方、ソーシャルネットワークのもつバイアスを補うために、従来型のアンケート調査やインタビュー調査も並行して利用し、女性労働の就業継続や管理職昇進などの問題を明らかにしていく。

最終的に、上記の研究結果を踏まえて、女性の労働問題を解決するための方策につなげる。

### 3. 研究の方法

研究で主に用いるソーシャルネットワークは、Twitter、本研究で構築・開発・運営したSNSサイト、ヤフーニュースの3つである。

人々の「個人の意識」「家族状況」「職場・仕事状況」は複雑に絡んでいるため、これらの解析は個人属性に応じて実施する必要がある。しかし、Twitterデータでは個人属性データが明示的にデータベース化されていないため、意見/評価の自動要約として文書要約、センチメント解析といったテキストマイニング手法、またメタデータ解析としてユーザネットワークの構造変化の解析を行う。

従来型のアンケート調査やインタビュー調査も実施するほか、これまでに蓄積してきた調査を利用して分析する手法もとる。

### 4. 研究成果

研究は、(1)Twitterデータを用いた分析、(2)アンケート調査やインタビュー調査を用いた分析、(3)独自に構築・開発・運営したSNSのデータを基にした分析、(4)ヤフーニュース記事に対するコメントデータを基にした分析を行った。以下に説明する。

#### (1)Twitterデータを用いた分析

まず、データベースの構築として、クローリングプログラムの作成、Twitterクローリングとデータベース化、イベントデータベースの構築を行った。そして実際に、育児休業(以下、「育休」)3年を求めた安倍首相の発言に関する人々の考えについて、Twitterデータ(28万件)を取得し、テキストマイニングを行い、意見の抽出方法を提案した。具体的には、類似した文章をクラスタリングすることでテーマごとに意見抽出や重要文章の抽出方法を確立することを狙いに、文章分類と意見抽出を行った。2部グラフから研磨後の極大2部クリークをトピックとして扱い、共通するトピックを持つツイートをクラスタリングした。クラスタ内のツイートは分類できたが、文章の類似性の判断が難しいため、被験者の協力を得て文章の評価実験を実施した。その結果、これまで分類モデルの説明変数としてクラスタの有用性を示せなかったが、クラスタリングの性能を評価できた。

次に、収集したツイートデータでテキストの格フレーム解析を行い、データベース化した上でツイートを要約する技術の開発を試みた。具体的には、ツイートの解析から、格フレームを利用した2部グラフに、データ研磨を適用することで構造を明確化し、極大2部クリークの列挙数が大幅に減少することが示された。また、極大2部クリークをトピックとして利用することで、類似するツイートのクラスタリングと有用な情報を抽出できることを示した。育休に関するツイートの要約からは、安倍首相の育休3年政策と仕事復帰に関して否定的な意見が多く、育休

3年という政策は論点がずれていることなど、国民の率直な意見を捉えることができた。

また、Twitter を利用した既婚女性に対する非労働力人口の予測モデル構築を試みた。具体的には、ツイートから、評価表現を機械学習の手法により抽出し、評価表現の格フレームネットワークを構築した後に、評価表現トピックの出現数に基づいて、非労働力人口主婦の増減を予測するモデルを構築した。その結果、「安倍首相」「時短勤務」「女性の働き方」などのキーワードが非労働力人口の減少に寄与し、政策に対して人々が反応していることが確認できた。またこのモデルで Twitter から肯定・否定意見を判定するため、短文のツイートは同一文章では同一の評価表現からなるという仮定に従い、評価表現辞書を構築し、予測モデルに利用した。

#### (2) アンケート調査とインタビュー調査を用いた分析

短文とサンプルの偏りを持つ Twitter データの欠点を補うために実施したのがアンケート調査である。まず、育休を取得した人と取得せずに退職した人へのアンケート調査では、育休取得せずに退職する理由を明らかにし、仕事・育児両立のための施策を評価し提案することを試みた。その結果、家族状況要因の改善が育休取得に強い影響を与えること、ワーク・ライフ・バランスを改善することで育休取得に間接的な効果があることが示された。

次に、男女雇用均等施策やワーク・ライフ・バランス施策への取り組みが企業業績に影響を与えるのかという問題意識で行ったアンケート調査より、均等施策により女性のモチベーションが上がったことが企業業績を有意に高めることが示された。

さらに、企業の人材育成が女性のスキル形成やモチベーションにどのように影響を与えるかについて、アンケート調査とインタビュー調査を用いて分析した。アンケート調査では、基幹労働者に限定しても、男女で配置転換の経験の幅が異なり、それが男女の昇進格差が生じる原因の大きなものの一つになっていることが明らかになった。均等法世代、第二世代、第三世代のいわゆる女性総合職 120 名に対して過去 5 回にわたって行ったインタビュー調査の結果を総合的に分析したものでは、業務配分、配置転換、評価、昇進の一連の人材育成が女性の昇進意欲を高めることを示した。

#### (3) 独自開発した SNS を用いた分析

上記(1)(2)の研究を進める中で、単発のアンケート調査では Twitter の短文とサンプルの偏りを十分に補えないことが明らかになった。そこで継続的に人々の率直な意見を収集できる SNS サイトを構築することにした。オープンソースの SNS パッケージを基盤に必要最小限の機能を実装し、仕事と家庭の両立や管理職昇進に関心の高い女性たちの協力を得て、実際に SNS サイトを利用してもらった。利用規約周知など倫理的配慮をし、ユーザが様々な意見交換をしやすくするための使い勝手を改良しながら、女性の雇用や育児、教育などの情報を提供するポータルサイトとしての機能を充実させた。この SNS サイトで蓄積したユーザのログデータを利用することにより、人のつながりをベースにしたネットワーク解析が可能になった。

ネットワーク解析は、影響力の強いユーザの特定と、クラスタリングを利用したユーザの類型化によるコミュニティ抽出、言葉ネットワークによる意味抽出を目的に実施した。まずネットワークの中心性を表す指標を用いて、ハブとなるユーザとコミュニティ間をつなぐブリッジとなるユーザを発見した。そして、Newman アルゴリズムを適用することでコミュニティ構造を捉え、言葉ネットワークで文章から内容を表す格フレームを抽出し、それらの関係性から文章の意味をある程度把握することができた。

本 SNS サイトを用いて、さらに、関西学院大学専門職大学院の授業でアクティブラーニングの一つとして利用し、その効果を分析した。多くの学生が情報取得などの目的で利用する Facebook とは異なり、本 SNS では幅広い意見や考えの交換ができ、発展的な学習をする動機付けとなり、より深く学ぶことができたと評価していた。自分たちの考えや意見を積極的に表現することを奨励することでポジティブな学習効果があると言える。

本サイトは匿名ではなく本名でユーザ登録をしている。そこで積極的に表現されるユーザの考えや意見は Twitter など一般の匿名の SNS とは異なる。それは次のヤフーニュース記事コメントのクラスタリング活用ワークの研究から示される。

#### (4) ヤフーニュース記事コメントのクラスタリング活用ワークの研究

(1) で用いた Twitter は短文であるのに対して、ヤフーニュース記事のコメントは比較的長文である。そこで、育児休業に関する一般の人々の意見を知るために、ヤフーニュース記事に投稿された多数のコメントをクラスタリングした。ただし、この状態でクラスタリングしたものは、上述の Twitter の分析で文章の類似性の判断が難しいことが示されたように、必ずしも文脈上、整合性の取れるものではないという限界がある。そこで、関西学院大学社会人対象講座の受講者の協力を得て、ワークショップ形式でクラスタ結果の再整理を施して意味解釈をした。その結果、例えば男性の育児休業について、肯定的な意見は、男性が子育て経験を持つべき、労働者の権利、個人の自由など、比較的落ち着いたもので、理由を持っているものが多い一方、否定的な意見は周りや社会から受けるプレッシャーを代弁した攻撃的なものが多いことがわかった。否定的な意見に攻撃的なものが多いのは、記名の本 SNS にはない匿名のヤフーコメントの特徴で、ある意味で人々の本音が出ているとも考えられる。意味解釈にあたってどのような方法が適しているのかは引き続きの課題となった。

以上、本研究では、Twitter、本研究独自開発の SNS、ヤフーニュースという 3 つのソーシャルネットワークの大量のデータを用いて、育児休業など女性労働についての人々の意見/評価を人による解釈の助けを得ながら自動要約することができるようになり、また、人のつながりをベースにしたネットワーク解析にて重要なユーザ・意見のある程度把握することができた。並行して、従来型のアンケート調査やインタビュー調査で、就業継続や管理職昇進の阻害要因を明らかにした。課題として、こうしたデータで性別、年代、家族構成などのユーザプロフィールを推定することはいまだ難しい。今後、ビッグデータを用いた人々の意見/評価の自動要約の更なる改善により、女性労働について一般の人々(男女)の間で何が問題となっているかを把握することができ、従来型のアンケート調査やインタビュー調査も並行して利用していくことで、女性労働問題を解決し政策に反映させることができると期待できる。

5. 主な発表論文等

〔雑誌論文〕 計28件（うち査読付論文 19件 / うち国際共著 2件 / うちオープンアクセス 13件）

1. 著者名 大内章子, 奥井めぐみ, 脇坂明	4. 巻 20
2. 論文標題 男女の配置転換経験の違いは昇進格差を生むのか - 企業調査と管理職・一般従業員調査の実証分析より -	5. 発行年 2017年
3. 雑誌名 ビジネス&アカウンティングレビュー	6. 最初と最後の頁 71, 88
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 奥井めぐみ, 大内章子	4. 巻 -
2. 論文標題 男女の昇進スピード格差と配置転換・職能経験	5. 発行年 2017年
3. 雑誌名 日本労務学会第47回全国大会報告論集	6. 最初と最後の頁 102, 109
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計50件（うち招待講演 9件 / うち国際学会 1件）

1. 発表者名 奥井めぐみ, 大内章子
2. 発表標題 男女の昇進スピード格差と配置転換・職能経験
3. 学会等名 日本労務学会第47回全国大会
4. 発表年 2017年

1. 発表者名 中原孝信, 羽室行信, 大内章子, 宇野毅明
2. 発表標題 Twitterを利用した既婚女性に対する非労働力人口の予測モデル構築
3. 学会等名 新しい経済産業指標開発コンテスト
4. 発表年 2017年

1. 発表者名 中原孝信, 大内章子, 宇野毅明, 羽室行信
2. 発表標題 データ研磨の2部グラフへの適用とTwitterからの意見抽出
3. 学会等名 2016年度人工知能学会(第30回)
4. 発表年 2016年

1. 発表者名 中原孝信, 羽室行信, 大内章子, 宇野毅明
2. 発表標題 Twitterを利用した育休に関する意見抽出 - 2部グラフを対象にしたデータ研磨の適用 -
3. 学会等名 宇野CREST x SOLEIL DATA DOJO第一回シンポジウム
4. 発表年 2015年

1. 発表者名 中原孝信, 大内章子, 羽室行信, 宇野毅明
2. 発表標題 出産時の就業継続と育児休業取得に関する意思決定要因の抽出
3. 学会等名 日本労務学会関西支部会
4. 発表年 2016年

〔図書〕 計2件

1. 著者名 公益財団法人21世紀職業財団(大内章子:分担執筆)大内章子、奥井めぐみ、山名真名	4. 発行年 2015年
2. 出版社 公益財団法人21世紀職業財団	5. 総ページ数 276(15-84)
3. 書名 『若手女性社員の育成とマネジメントに関する調査研究 - 均等法第三世代の男女社員と管理職へのインタビュー・アンケート調査より』(第1-4章;分担執筆)	

〔産業財産権〕

〔その他〕

関西学院大学 女性活躍推進研究センター  
<http://www.kwansei-ac.jp/iba/wpa/>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	宇野 毅明  (UNO Takeaki)  (00302977)	国立情報学研究所・情報学プリンシプル研究系・教授   (62615)	
研究分担者	中原 孝信  (NAKAHARA Takanobu)  (60553089)	専修大学・商学部・准教授   (32634)	
研究分担者	羽室 行信  (HAMURO Yukinobu)  (90268235)	関西学院大学・経営戦略研究科・准教授   (34504)	