

アントレプレナーシップ論文抄録を用いたテキストマイニング分析

Text mining analysis of entrepreneurship thesis abstracts

林 永周 (京都情報大学院大学)

Yeongjoo Lim (The Kyoto College of Graduate Studies for Informatics)

Abstract

Entrepreneurship is an important economic activity in the revitalization of a national economy. Entrepreneurship enables the development of new business and aids in facilitating the generation of employment as a result entrepreneurship is expected to bolster the transformation of industrial structure and assist in economic growth. However, the accumulation of researches on entrepreneurship in Japan are still very few when compared to the United States and Europe. In this study, in order to clarify the current situation of entrepreneurship research in Japan a detailed analysis of literature is carried out by exploiting the strengths of text mining in order to put forward propositions for future studies.

1. 研究の背景と問題意識

起業は、国家経済を活性化するための重要な経済活動であり、起業活動によって新しいビジネスや雇用が生み出され、その結果として経済成長や産業構造の転換が期待される。ベンチャー企業の役割については、「日本産業のイノベーションを促進し、日本経済全体の成長と活性化を図るためには、新しい技術やビジネスモデルを有し、大きなビジネスリスクをとって新規事業に挑戦するベンチャー企業の創出・成長が不可欠である」[1]、「創業（起業）は、経営者を含めた新しい雇用創出の機会につながり、これまでと同じ製品やサービスあるいは生産方法や管理方法では既存企業との競争に打ち勝つことが困難なことから、創業は、新しい製品やサービス、あるいは何らかの『イノベーション』をもたらすと期待できる」[2]とベンチャー企業の必要と経済効果について議論されている。ベンチャー企業の役割と経済効果については、十分な社会的な合意が形成されながらも、日本における起業活動率調査では、日本の起業活動率は最も低い水準であることが指摘されている[3]。日本の起業活動率が低いとは言え、日本では数多くのアントレプレナーシップに関する教育プログラムが実施されている。大和総研の調査

によれば、アントレプレナーシップ関連の教育を提供している大学は2000年の139校から2009年には252校に増加し、科目数も同じ時期に330科目から1,078科目へと3倍以上となっている[4]。講義の質の問題はあるにせよ、このような現状は日本においても、アントレプレナーシップの教育に関する関心が高まり、その必要性について認知されていると捉えることができる。その半面、アントレプレナーシップに関する研究の蓄積は米国、欧州に比べ、日本ではマイナー的な位置づけであり、研究の蓄積も少ない。日本において論文のデータベースを提供しているCiNiiのウェブサイトにて、アントレプレナーシップで検索すると検索される論文は73件に過ぎず¹⁾、意味を広げアントレプレナー、起業家で検索しても検索件数はそれぞれ、191件²⁾、527件に過ぎない³⁾。一方、英文の論文データベースである

- 1) CiNiiにてキーワードをアントレプレナーシップ、CiNiiに本文あり・連携サービスへのリンクありで検索した結果（最終検索：2016年3月25日）
- 2) CiNiiにてキーワードをアントレプレナー、CiNiiに本文あり・連携サービスへのリンクありで検索した結果（最終検索：2016年3月25日）
- 3) CiNiiにてキーワードを起業家、CiNiiに本文あり・連携サービスへのリンクありで検索した結果（最終検索：2016年3月25日）

EBSCOhost のウェブサイトにて、Entrepreneur をキーワードとし検索すれば、ヒットする検索結果は 385,326 件で、条件を学術誌（査読）に限定しても 32,608 件がヒットする⁴⁾。単純に比較することで研究の蓄積を判断できない意見があるにせよ、日本においてアントレプレナー、またはアントレプレナーシップに関する研究の蓄積が乏しい状況であることは明確であると考えられる。そこで、本研究では、日本で公開されている論文の抄録を分析することによって、日本におけるアントレプレナーシップに関する研究の動向、方向性を明確にすると共に、今後議論すべき領域を明らかにする。

2. アントレプレナーシップに関する定義

アントレプレナー、アントレプレナーシップ、起業家、起業家精神など、アントレプレナーに関する同類の言葉が多く使われているが、それに対する定義と違いが明確にあるわけではない。アントレプレナー（Entrepreneur）は、「起業家」「企業家」「起業家」という言葉で訳されている。しかし、そうした翻訳の仕方に関しては、まだ十分に議論されておらず、統一した見解もない [5]。また、アントレプレナーシップ（Entrepreneurship）は、一般的に「企業家精神」という言葉で訳されている。これは、アントレプレナーに「精神面」や「個人的資質」を意味するものと解釈される傾向である [5]。このように、アントレプレナーとアントレプレナーシップの概念において、原（2002）は、アントレプレナーとアントレプレナーシップを経済学、心理学、社会学、経営学の様々な学問から分析し、アントレプレナーを高い創造性と高い経営技術を有するか、あるいはそれらを借用する能力を有し、そして事業機会を認識し、その事業を創造する者であると定義している。また、アントレプレナーシップとは、事業の機会を認知し、その事業を創造する動的プロセスであると定義している。原（2002）の定義からすれば、アントレプレナーは事業を創造する人であり、アントレプレナーシップは事業を創造する動的プロセスとして捉えることができ、アントレプレナーが行動することがアントレプレナーシップであると解釈することができる。

4) 最終検索：2016年3月25日

3. 研究目的と分析方法

本研究では、原（2002）が示唆しているアントレプレナーシップとアントレプレナーの定義にしたがって、アントレプレナーシップとアントレプレナーを定義する。本研究は、動的プロセスであるアントレプレナーシップに焦点を当てて分析する。本研究の目的は、日本におけるアントレプレナーシップに関する研究の現状を把握し、今後研究されるべき領域を明らかにすることである。

日本におけるアントレプレナーシップに関する研究の現状を把握するために、CiNii 学術論文データベースにて、キーワードをアントレプレナー、アントレプレナーシップ、起業家の3つで検索し、それぞれ検索された論文のうち、重複で検索された論文を除いた 272 件の論文のリストを作成し、本文と抄録が入手できる論文を確認し、キーワードで不適切だと判断された論文を除いた 55 件を分析の対象とした。分析方法としては、KH-Coder を用いたテキストマイニングを行い、頻度分析と共起ネットワーク分析を行う。テキストマイニング分析は、テキストデータを計算機で定量的に解析して有用な方法を抽出するための様々な手法 [6] であり、大量の文字情報を一括処理して、テキストの中に隠された法則や知見を発見する分析手法である。

頻度分析は、文献における出現頻度を検討するものであり、基本的考え方は、「出現頻度が高い単語ほど重要度が高い」、または「出現頻度が低い単語ほど重要度が低い」というものであり、対象文献の特徴を知るための最も基本的な分析である。

共起ネットワーク分析は、語と語の結びつきを探る手法であり、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いる。数学のグラフ（Graph）理論に基礎をおいている。共起ネットワークとは、対象文献における単語と単語の間の関連性を検討するものであり、頻度分析の検討では十分に明らかにならない特徴を、単語間の関連性を分析することによって、それぞれの文献の特徴を明らかにすることが可能となる。KH-Coder を用いると、このような単語間の関係性を踏まえて、出現パターンの似通った単語、すなわち共起の程度が強い単語を線で結んだネットワーク図を描くことができる [7]。KH-Coder では、単語と単語のネットワークを描く際には、T. M. J. Fruchterman & E. M. Reingold (1991) [8] の方法を用いている。

4. 分析

4-1 データ・セット

CiNii 学術論文データベースにて、キーワードをアントレプレナー、アントレプレナーシップ、起業家の3つで検索し、それぞれ検索された論文のうち、重複で検索された論文を除いた272件の論文のリストを作成し、本文と抄録が入手できる論文を確認し、キーワードで不適切だと判断された論文を除いた55件をKH-Coderで分析可能な状態でコーディングする。コーディング方法はTxtファイルを、エディターを用いて作成し、<H1> 論文タイトル </H1>、<H2> 論文抄録 </H2> のルールで、文字コードShift-JISで作成した。分析対象となる抄録は、2016年3月に収集されたものである。

4-2 頻度分析

KH-Coderを用いて、頻度分析を行った。抽出語のオプションにて、品詞別で、出現回数を抽出した。

4-2-1 名詞

分析対象である55件の論文抄録の「名詞」の頻出回数から、上位30語をリストアップしたのが表1である。ここでいう名詞には、固有名詞やサ変名詞（～する）と連結した名詞は含まれない。頻出回数から見てみると、アントレプレナーシップに関する研究は「企業」(103)やベンチャー(66)、ビジネス(53)を分析していると考えられる。また、

表1 「名詞」上位30語

名詞	頻度	名詞	頻度
企業	103	課題	19
ベンチャー	66	市場	19
社会	62	概念	18
ビジネス	53	目的	18
大学	41	現状	17
事業	31	精神	16
技術	26	制度	15
ネットワーク	25	地域	15
経済	25	中心	15
イノベーション	24	プロセス	14
事例	24	金融	14
中小	23	個人	14
本稿	23	能力	14
国際	21	モデル	13
人材	20	知識	13

その分析の切り口として、大学(41)、事業(31)、技術(26)、ネットワーク(25)、イノベーション(24)、事例(24)の側面から分析していると推測することができる。

4-2-2 サ変名詞

続いて、分析対象である55件の論文抄録の「サ変名詞」(～するをつけると動詞化する名詞)についての頻出回数から、上位30語をリストアップしたのが表2である(ただし、30位にランクした単語の頻度が10で同じだったため、リストアップしたのは、合計34になる)。頻出回数から見てみると「起業(222)・創造(22)・創業(16)する」ということが、主たる分析の対象であり、起業における「教育(97)・育成(20)する」ことや「経営(42)・管理(11)する」、「支援(32)する」などの仕組みについての研究(51)、分析(31)、考察(26)が主に進められていると考えられる。また、アントレプレナーシップを持つ人に対する分析が意識(16)、認識(11)、経験(10)を通じて研究されていると考えられる。

4-3 共起ネットワーク分析

頻度分析は、文献における出現頻度を検討するものであり、「出現頻度が高い単語ほど重要度が高い」「出現頻度が低い単語ほど重要度が低い」という考

表2 「サ変名詞」上位30語

サ変名詞	頻度	サ変名詞	頻度
起業	222	展開	16
教育	97	関係	15
研究	51	評価	14
経営	42	検討	13
支援	32	報告	13
分析	31	学習	12
調査	29	比較	12
成長	28	管理	11
考察	26	既存	11
活動	24	構築	11
創出	23	注目	11
組織	22	認識	11
創造	22	アプローチ	10
育成	20	競争	10
実践	18	経験	10
意識	16	紹介	10
創業	16	成功	10

え方に依存している。そのために、頻度分析のみでは、不十分な分析結果での解釈となるため、共起ネットワーク分析を行った。共起ネットワークは、出現パターンの似通った語、すなわち共起の程度が強い語を線で繋いだ分析手法である。共起ネットワーク分析での共起関係の強弱について、Jaccard 係数の値が計算される。Jaccard 係数の値は 0 から 1 までの値となり、対象となった単語のすべての組み合わせについて、2つの単語について同じ文章中に同時に出現する（共起する）と関連性が強いと見なし、その値は 1 に近づく。また、2つの単語について同じ文章中に出現しない（共起しない）と関連性が低いと見なし、その値は 0 に近づく。単語 X と単語 Y に関する Jaccard 係数 (j) は下記の数式で求められる。

$$j = a / (a + b + c)$$

a: 単語 X と単語 Y が同時に出現している文章数

b: 単語 X のみ出現している文章数

c: 単語 Y のみ出現している文章数

KH-Coder では、Jaccard 係数の値を自動的に計算し、単語間のネットワーク図を自動的に描き、単語の「中心性」を基準としたネットワークと「サブグラフ検出」によるネットワークのいずれかを選択することができる。また、ネットワーク図においても、単語と単語をつなげる線の太さは、単語間の共起の程度が強いほど太く描画され、単語間の共起の程度が弱いほど細く描画される。本研究での共起ネットワークは、「中心性」と「サブグラフ」の 2 つを用いて分析する。「中心性」を基準としたネットワーク図は、「媒介中心性」に基づいて作成されている。また「サブグラフ検出」によるネットワーク図は、「共起関係の媒介性に基づく方法」を用いて作成されたものである [9]。

4-3-1 共起ネットワーク分析結果（中心性）

共起ネットワーク分析の中心性分析は、それぞれの単語が当該文献の中でどの程度中心的な役割を果たしているかを、ネットワーク図によって示すものである。それぞれ単語は色分けされ、赤色—桃色—白色—水色の順に中心性の高さが示される [7]。

KH-Coder で共起ネットワークを描画するための設定値としては、集計単位を H2、最小出現数を 10、最小文章数を 1 とし、分析に用いられる単語数は 95 個であった。また、オプションにて、共起関係

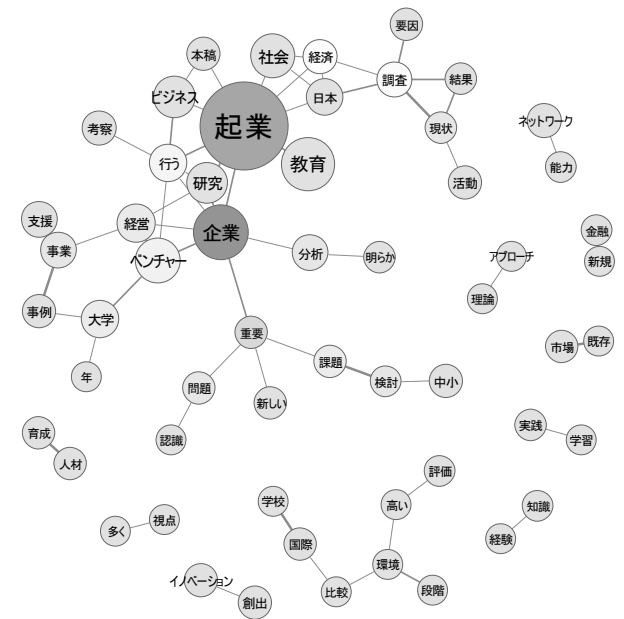


図 1 共起ネットワーク（中心性（媒介））

Node:50, Edge:68, Density: .056, Min Jaccard: .333

の種類は、語—語と設定し、共起関係の絞込は描画数 60（全てデフォルト値）に設定した。描画された共起ネットワーク（中心性）は図 1 である。

図 1 から見てみると、アントレプレナーシップに関する研究では、起業、企業がネットワーク図上で重要な役割を果たしている。起業と企業をメインとし、大学発ベンチャーの事例、事業支援、ビジネスについての研究が進められていることが推測できる。

起業とビジネス、企業、ベンチャーなどが比較的密度の高い（線が太い）ことから見てみると、起業やベンチャー企業のビジネスが研究対象となっており、企業が抱えている重要な問題、課題を検討することも研究対象として扱われている。また、「調査」を重心に日本、経済、現状、結果のネットワークが形成されており、この結果は、分析が和文の論文を対象としており、日本における現状、経済などについて分析が行われているためであると考えられる。

4-3-2 共起ネットワーク分析結果（サブグラフ）

共起ネットワークのサブグラフ検出は、相対的に強くお互いに結びついている単語同士を自動的に検出してグループ分けを行い、ネットワーク図によって示すものである。それぞれの単語グループは色分けされる [7]。ただし、単語グループの色分けは、KH-Coder が自動処理によって行うことであり、色分けにつねに重要な意味があるわけではない。サブグラフ検出は、グラフを解釈する際の補助として利

用することが望ましい[10]。色分けにおいて、囲み枠の中が白色であれば、その単語は他の単語とグループを形成していない単独の単語であることを意味している。また、サブグラフ検出を行った場合、同じサブグラフに含まれる単語は実線でつながるのに対し、お互いに異なるサブグラフに含まれる単語は破線でつながる。

KH-Coderで描画された共起ネットワークをサブグラフ検出(媒介)した結果は図2である。サブグラフ検出は、図1の中心性共起ネットワーク図をグループ分けしたものであり、色分けには特に意味はない。

図2から見てみると、日本におけるアントレプレナーシップの研究は、11グループに分けられる(囲み枠が白色は、グループを形成しない単独の言葉であるため、グループ数から除く)。

第1グループ(濃い緑色)は、起業を主たるキーワードとし、ビジネスや、事業事例、大学ベンチャー起業、起業教育が重要な役割を果たしていると考えられる。第2グループ(赤色)は、日本における現状と経済、社会への影響を調査している傾向であると読み取れる。第3グループ(薄い紫色)は、課題と問題を認識し、検討することで形成されていると考えられる。第4グループ(薄い黄色)は、国際的な環境、学校を比較し、評価している内容で形成されていると考えられる。他のグループはそれぞれ、ネットワーク能力、理論的アプローチ、人材育

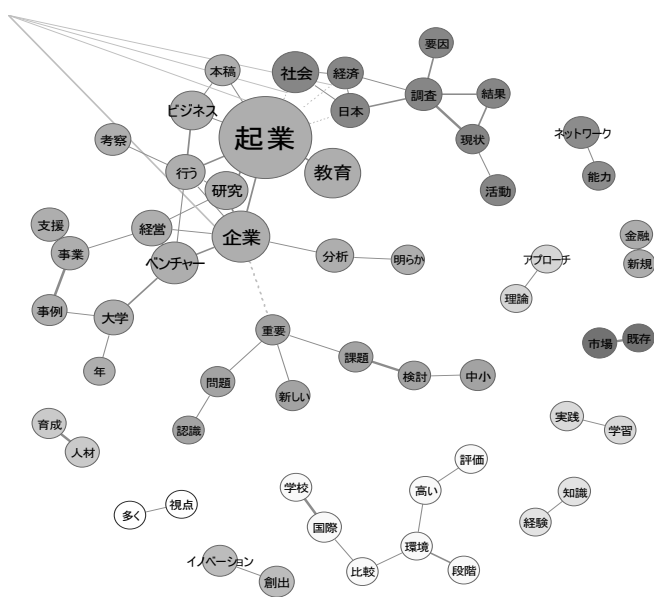


図2 共起ネットワーク(サブグラフ検出(媒介))
Node:50, Edge:68, Density: .056, Min Jaccard: .333

成、知識経験、イノベーション創出、実践学習、既存市場で形成されている。特に、人材育成は、比較的太い線につながっていることから、一つの研究領域として確立されていると考えられる。

5. 考察

本章では、日本におけるアントレプレナーシップ研究の現状と、今後研究されるべきアントレプレナーシップ研究領域について考察したい。

名詞の頻度分析の結果、頻出回数から見てみると、アントレプレナーシップに関する研究は企業(103)やベンチャー(66)、ビジネス(53)が特に多く、続いて大学(41)、事業(31)、技術(26)、ネットワーク(25)、イノベーション(24)、事例(24)が出現していることから、企業のうちベンチャー企業を、大学、事業、技術、ネットワーク、イノベーションについて、事例を用いて分析していると考えられる。

サ変名詞の頻度分析では、起業(222)・創造(22)・創業(16)の3つを同類の概念と捉えれば、事業創出(起業)のための教育(教育(97)・育成(20))、マネジメント(経営(42)・管理(11))、サポート(支援(32))を研究(研究(51)、分析(31)、考察(26))していると考えられる。

共起ネットワーク分析の結果から、起業と企業、ベンチャーが中心的な役割を果たしており、現状調査、問題解決、人材育成、イノベーション創出、ネットワーク能力、国際比較などのアプローチから研究されていると考えられる。

これらの結果を踏まえると、本研究で用いられているアントレプレナーシップの定義が動的プロセスであることから考えると、日本におけるアントレプレナーシップの研究は、どちらかと言えば、動的プロセスよりは、既存のベンチャー企業の実態を分析している傾向が強いと考えられる。アントレプレナーシップは、アントレプレナーが行動するプロセスであるという概念から考えると、教育においては、如何にアントレプレナーを育成するか、アントレプレナーはどのようなプロセスで起業するのかについて十分な研究がされているとは言えない。これらの原因としては、日本にアントレプレナーシップが定着したのが外国に比べて比較的遅かったこと、研究対象としてマイナー的な領域であったこと、ベンチャー企業などの情報入手に限界があること、起業活動率が世界最下位水準であることが挙げられ

る。しかし、製造業を中心とした徹底した品質・コストマネジメントだけでは、日本経済の成長を期待することは難しく、アントレプレナーやアントレプレナーシップに対する社会的な期待度が増えていることは事実である。

こうしたことから、今後研究されるべきアントレプレナーシップ研究領域としては、以下の領域があると考えられる。

- 1) アントレプレナーの育成（学校における教育）
- 2) アントレプレナーが行動するプロセス
- 3) アントレプレナーシップの具体的プロセス
- 4) アントレプレナーとアントレプレナーでない人の差
- 5) アントレプレナーでない人をアントレプレナーにするための要因

6. まとめ

本研究では、テキストマイニング手法を用いて、日本におけるアントレプレナーシップの研究の現状と今後の研究すべき領域を導出するために、和文で作成された論文抄録を分析した。分析方法としては、名詞の頻度分析、サ変名詞の頻度分析、共起ネットワーク（中心性、サブグラフ）の手法で分析した。その結果、日本におけるアントレプレナーシップ研究は、動態的プロセスよりは、その実態を研究対象としている傾向が見られることと、今後の研究されるべき領域を5つ導出することができた。

本研究ではいくつかの課題が残されている。第1に、分析において用いたデータが論文抄録のみであることから、論文筆者によって、主張したい内容が全て単語として含まれているとは限らない。第2に、抄録が入手できる論文55件のみを分析対象としており、全体論文（272件）の約20%程度しか検討できていない。これらの要因は、研究結果に何らかのバイアスになっている可能性を否めない。そのた

め、今後、論文抄録のみではなく、全文を用いた研究の拡大が求められる。また、和文のみを分析対象にしていることから、日本におけるアントレプレナーシップに関する研究の世界的位置づけについても考慮されていないことから、他言語との比較研究も今後の課題としたい。

【参考文献】

- [1] 経済産業省（2008）「ベンチャー企業の創出・成長に関する研究会最終報告書」
- [2] 本庄裕司（2010）『アントレプレナーシップの経済学』同友館
- [3] 磯辺剛彦，矢作恒雄（2011）『起業と経済成長—Global Entrepreneurship Monitor 調査報告—』慶應義塾大学出版会株式会社
- [4] 大和総研（2010），平成22年度「産業技術人材育成支援事業（起業家人材育成事業）」報告書
- [5] 原憲一郎（2002），アントレプレナーシップの概念試論，龍谷大学経営学論集，42(2)，pp. 44-57
- [6] 松村真宏，三浦麻子（2009）『人文・社会科学のためのテキストマイニング』誠信書房
- [7] 下平裕之，福田進治（2014），古典派経済学の普及過程に関するテキストマイニング分析—リカード，ミル，マーティノーを中心に—，人文社会論叢，社会科学篇，Vol31，2014
- [8] Fruchterman T. M. & Reingold E. M., "Graph Drawing by Force-directed Placement", Software: Practice and Experience, 21(11), 1129-1164, 1991.
- [9] 樋口耕一（2012）「KH Coder 2.x リファレンス・マニュアル」(<http://khc.sourceforge.net/>)

◆著者紹介

林 永周 Yeongjoo Lim

京都情報大学院大学非常勤講師。
立命館大学大学院テクノロジー・マネジメント研究科博士課程終了，博士（技術経営）。
研究分野はビジネスモデルとアントレプレナーシップ教育。