

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年 3月31日現在

機関番号：37404

研究種目：基盤研究（C）

研究期間：2009～2011

課題番号：21500243

研究課題名（和文） 大学研究影響度指標の研究

研究課題名（英文） A study on university impact indicators

研究代表者

角田 裕之（TSUNODA HIROYUKI）

尚綱大学・文化言語学部・教授

研究者番号：30454961

研究成果の概要（和文）：大学の研究影響度を測定する指標を提案することを目的とした。まず、著者所属機関から大学を識別する方法を考察した。次に、引用を経由する文献の重みとナレッジプールを組み込んだ指標の特性を分析した。最後に、大学ランキングの物理学分野と天文学領域、及び、工学分野と材料学領域における相関から、提案した指標が大学ランキングに近いことを示した。

研究成果の概要（英文）：This study aims to propose indicators to measure the impact of university research. First, we discussed how to distinguish the university from the institution of the author. Second, we analyzed the characteristics of the knowledge pool and the weight of papers through a citation. Finally, the proposed indicators was shown to have the similar characteristics of university rankings between the subject of physics and the field of astronomy, and between the subject of engineering and the field of materials science.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,700,000	510,000	2,210,000
2010年度	1,400,000	420,000	1,820,000
2011年度	400,000	120,000	520,000
総計	3,500,000	1,050,000	4,550,000

研究分野：総合領域

科研費の分科・細目：情報学・図書館情報学・人文社会情報学

キーワード：図書館情報学、計量情報学、科学計量学、計量書誌学、研究影響度、引用分析、大学評価

## 1. 研究開始当初の背景

学術研究（研究と呼ぶ）は国公立大学などの高等教育研究機関および附置研究所（大学と呼ぶ）、国立研究所や研究機構、民間企業および付設研究所などで実施されている。政府はこれらの研究を推進するために、平成19年度には、3兆5千億円の予算を計上した。特に、科学研究費補助金は、研究のための競

争資金であり、昭和62年に451億円であったが、平成19年度には1913億円と4倍以上に増加した。これらの予算の源泉は税金であり、近年多額の税金の使途には、アカウンタビリティが求められ、かつ、効果的な執行が望まれる。これらの社会的な要請に定めるため、文部科学省では大学における研究方法や内容の改善を図る取組を積極的に推進する

ように指導しており、各大学とも自己点検・評価、外部評価等を実施している。

このような社会的な背景を踏まえ、本課題の研究開始までに、文献の重みが引用を經由して文献に伝わる2つの指標を提案した。一つ目は、研究者が研究者へ与える研究の影響の程度を測定する指標(Researcher Impact: *RI*)である。*RI*を用いて、計量書誌学を専門とするトップ15人の研究者の研究影響度を測定し、順位表を作成した。この順位と、研究成果の実績を評価するときに用いられていた被引用量による順位を比較した結果、大きく異なることを示した。二つ目は、研究への国の貢献度を測定する指標(Nation Contribution: *NC*)である。*NC*を用いて、図書館情報学における主要20カ国の研究貢献度を測定し、順位表を作成した。この順位と、国の研究の輸出/輸入比に相当する引用係数(被引用量/引用量)による順位を比較した結果、順位変動が発生し、日本が中国よりも下回り20カ国中で最下位になることを示した。これらの事前研究から、引用を經由して伝わる文献の重みの仕組みと、引用を經由して伝わる研究の影響度の仕組みが類似していることが分かった。

大学の研究影響度を測定する指標を提案するには、次の問題があった。まず、大量の文献に記述された著者所属機関から大学を識別することが極めて困難であった。この研究は機関名の名寄せと呼ばれ、小林ら(科学技術振興機構と富士通研究所の合同チーム)が、日本語表記で92%を達成したが、英語表記では78%であった。しかも、日本国内機関に限定した識別であった。また、研究者や大学の研究実績を評価する研究は多くあるが、大学の研究影響が伝わる仕組みについての研究や研究影響度を測定する研究は不足していた。

## 2. 研究の目的

本研究は、文献に記述された著者の所属大学を識別する方法と、研究が研究に与える影響を測定する方法を考察することにより、大学の研究影響度を測定する指標を提案することを主たる目的とする。具体的には、天文学領域と材料学領域の研究者が執筆した文献から、

- (1) 大量の文献から著者所属機関を識別する方法(マクロプログラムにより半自動化する方法)
- (2) 識別子と工程の組み合わせによる識別精度の向上(引用索引データベースから抽出された識別子に限定する制約条件)
- (3) 引用を經由する文献の重みとナレッジプールを研究影響度指標への組み込み方法(大学ランキングとの順位相関)を考察する。

## 3. 研究の方法

本課題を研究するには、二段階の研究が必要である。第一段階は大学の識別に関する研究、第二段階は大学ランキングとの相関が高い指標の研究である。まず、第一段階では、(1)研究に必要な文献を収集し、(2)著者の所属機関を調査し、(3)辞書を編集し、(4)大学を識別する工程を考察する。第二段階では、(5)引用を經由して伝わる文献の重みと、(6)ナレッジプールを指標に組み込む方法を考察し、(7)大学ランキングとの順位相関を分析する。

### (1) 対象とする文献

本課題の測定範囲は、物理学分野から天文学領域と工学分野から材料学領域の2つの下位領域とした。文献を収集する情報源は、物理学分野と工学分野の両分野を含む自然科学系分野における最も包括的な引用索引データベースのひとつである Thomson Reuters の Web of Science に収録されている SCI (Science Citation Index Expanded) とした。ソース文献(被引用文献)は、SCIで天文学領域に分類された雑誌から *Astronomical Journal*, *Astronomy & Astrophysics*, *Astrophysical Journal*, *Monthly Notices of the Royal Astronomical Society* の4誌、および、材料学領域に分類された雑誌から *Advanced Materials*, *Chemistry of Materials*, *Journal of Materials Chemistry*, *Nanotechnology* の4誌の計8誌に2005年に発表された文献(Article)とした。引用文献は、ソース文献を引用している文献(Article, News Item, Reprint, Correction, Editorial, Item about an individual, Letter, Meeting Abstract, Review)のうち、2005年から2009年の間にSCIが採録した全ての雑誌に掲載された文献とした。この条件を満たす情報源の文献数を表1に示す。

表1 文献数

領域	被引用文献数	引用文献数
天文学	6,030	54,458
材料学	2,739	45,705
	8,769	100,163

### (2) 著者所属機関

被引用文献と引用文献の合計108,932件に記述された著者所属機関は、延べ268,258件であった。これらを研究代表者が一件ずつ目視で確認し、大学を確認したのは、延べ174,335件、被引用文献もしくは引用文献のいずれかに現れる大学は2,522校であった。本課題における大学とは、ユネスコ国際大学協会高等教育情報センター(IAU/UNESCO Information Center on Higher Education)

が編集する World Higher Education Data Base 2010(WHE)に収録された機関を指す。

### (3) 辞書編集

WHE と SCI では編集方針が異なる。WHE の大学名称は正規かつフルスペルで記述されているのに対し、SCI の大学名称は、頻出語や専門用語に略語で表記され、ストップワードと記号を削除され、創立者や所属自治体などの冠名が無くても識別できる単語を削除されている。無編集の WHE を用いて SCI から収集した著者所属機関の大学名称と照合したところ、一致したのは 268,258 件中 22 件であった。SCI との照合には WHE を編集する必要がある。そこで、以下の方針に基づき SCI の大学名称、市、郵便番号、国の記述と一致するように、WHE の大学名称、頭字語、住所、郵便番号、国を編集した。この編集は、特定の研究領域下に出現する著者所属機関に基づく最適化でもある。

- ① ダイアクリティカルマーク付き文字をアルファベットに統一
- ② 短縮語を正規単語に統一
- ③ ストップワードと記号の削除
- ④ 英語以外の言語によるスペルと語順を英語のスペルと語順に統一
- ⑤ 不要単語の削除と不足単語の追加
- ⑥ 頭字語がある大学名称を頭文語に統一
- ⑦ 附置施設の表記を統一
- ⑧ 国、市、郵便番号の表記を統一

### (4) 識別工程

辞書を用いて、以下の工程の順に照合を進め、著者所属機関から大学を識別した。

- ① 大学名称の照合
- ② 通称・変遷名の照合
- ③ 頭字語の照合
- ④ 辞書から生成した頭字語の照合
- ⑤ 語順の照合
- ⑥ 市を付加した大学名称の照合
- ⑦ 著者所属機関から生成した頭字語の照合
- ⑧ 附置施設の照合
- ⑨-③ 市の制約条件を緩和し、頭字語の照合
- ⑨-④ 市の制約条件を緩和し、辞書から生成した頭字語の照合
- ⑨-⑥ 市の制約条件を緩和し、市を付加した大学名称の照合
- ⑩-⑦ 郵便番号の制約条件を緩和し、著者所属機関から生成した頭字語の照合
- ⑪ 市と郵便番号の制約条件を緩和し、辞書と著者所属機関からの生成頭字語同士の照合

### (5) 引用を經由して伝わる文献の重み

本課題の研究では、大学の研究影響度は国

の研究貢献度( $NC$ )より、研究者の研究影響度( $R$ )に近いことから、引用元の文献の重みが引用を經由して引用先の文献に伝わる方法を採用した。その方法は、

- ① □引用元の文献の重みは引用を經由して伝わる引用先の文献に均等配分する。
  - ② □引用先の文献の重みは引用を經由して伝わる引用元の文献重みを合計する。
- による。そのイメージを図 1 に示す。

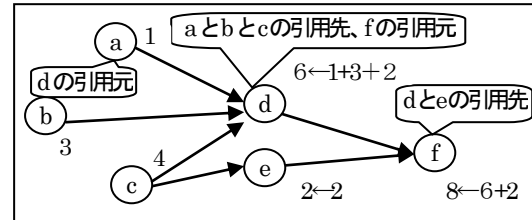


図 1 引用を經由して伝わる文献の重み付け

図 1 は、文献 a と文献 b が文献 d を各 1 回、文献 c が文献 d と文献 e を各 1 回、文献 d と文献 e が文献 f を各 1 回引用し、仮に文献 a の重みが 1、文献 b の重みが 3、文献 c の重みが 2 としたとき、文献 d は引用元の文献 a と文献 b の重みの和 4 と文献 c の 1/2 の重み 2 と和で合計 6 になる。他の文献の重みも同様に決める。文献で考察する限り文献 a、文献 b、文献 c の重みを決めることができない。なぜなら、この方法では、出発の文献に重みをあたることができず、全ての文献の重みが零になる。そこで、文献の集合体として大学を仮定し、全ての大学に重みを付ける方法を示す。そのイメージを図 2 に示す。

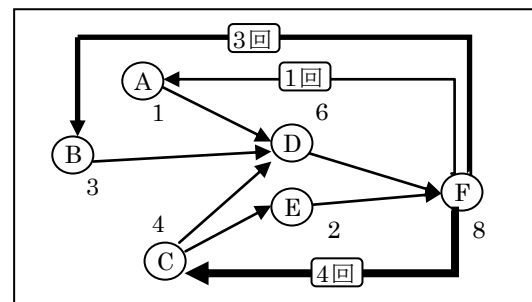


図 2 引用を經由して伝わる大学の重み付け

文献 a~文献 f を大学 A~大学 F に所属する著者が執筆したと仮定する。大学 F に所属する著者の文献 f' が大学 A に所属する著者の文献 a' を 1 回、大学 B に所属する著者の文献 b' を 3 回、大学 C に所属する著者の文献 c' を 4 回引用したとする。図 2 はこのときの引用を經由して伝わる大学の重み付けを示したものである。引用関係や引用回数に変化しても、文献と引用が再帰(任意の文献から出発して引用を經由し有限時間内にもとの文献

に帰着く確率が1である)・強連結(任意の文献から出発して引用を経由し任意の文献へ到達できる道がある)であれば文献の重みは決定する。

#### (6) ナレッジプール

「文献はいずれ周知の学術知識となり、文献は周知の学術知識を非明示的に引用する」と仮定し、周知の学術知識の集積をナレッジプールと呼ぶ。ナレッジプールと大学の引用のイメージを図3に示す。

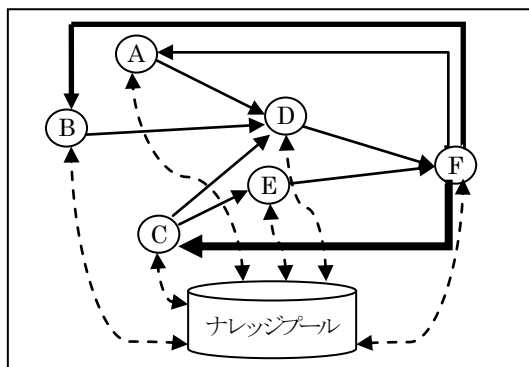


図3 ナレッジプールと大学の引用

このため、ナレッジプールを組み込んだ指標は、引用元の重みを明示的な引用を経由し引用先に伝える機能と、文献の重みを非明示的な引用を経由しナレッジプールを仲介させ文献に伝える機能を併せ持つ。ナレッジプールを組み込むことにより、ナレッジプールと全ての大学が相互の道で繋がるため、再帰・強連結になり、全ての大学の重み付けが可能になる。

#### (7) 大学ランキングを構成する指標

主要な大学ランキングとそれを構成する指標、及びその割合を示す。

##### ① The Times Higher Education World University Rankings (THE)

2011-2012年版は次の5つの指標からなる。

・教育環境評価	30%
・研究環境評価	30%
・被引用量	30%
・競争的資金獲得額	2.5%
・教員と学生の国際性	7.5%

##### ② QS World University Rankings (QS)

2011-2012年版は次の6つの指標からなる。

・学術による評価	40%
・企業による評価	10%
・教員あたり被引用量	20%
・学生あたり教員数	20%
・外国人留学生比率	5%
・外国人教員比率	5%

##### ③ Academic Ranking of World Universities (AR)

2010年版は次の6つの指標のうち、5つの指標からなる。

・ノーベル賞などの賞の数	10%
・分野別受賞者数	20%
・分野別高被引用研究者数	20%
・自然科学の文献数	0%
・SCIとSSCIの文献数	20%
・常勤研究員数	10%

##### ④ Ranking Web of World Universities Second (RW)

2011年版は次の4つの指標からなる。

・ウェブページ枚数	20%
・学術文献のファイル量	15%
・グーグルスカラーによる論文数と被引用量	15%
・サイトが受けた外部リンク総数	50%

##### ⑤ SIR (Scimago Institutions Rankings) World Report (SIR)

2011年版は次の6つの指標のうち、1つの指標からなる。

・文献数	100%
・海外共同研究比率	0%
・機関被引用率	0%
・高レベル誌の文献数	0%
・ジニー係数	0%
・高被引用誌の文献比率	0%

##### ⑥ The Complete List: The Top 100 Global Universities (NW)

2006年版のみであり、次の3つの指標からなる。

・被引用量	50%
・外国人学生比率	40%
・図書館蔵書数	10%

#### 4. 研究成果

本研究の成果は2つある。著者所属機関から大学を識別する精度の向上と、大学ランキングと相関が強い指標の提案である。

##### (1) 大学識別率

本課題の研究では、SCIの天文学と材料学に現れる著者所属機関に最適化させて編集した辞書を用いた。この識別率が高かったのは、工程①の大学名称の照合(76.5%)、②通称変遷名の照合(8.4%)、⑥市を付加した大学名称の照合(5.8%)、③頭字語の照合(4.7%)の4工程であった。図4に各工程までの累積識別率を示す。

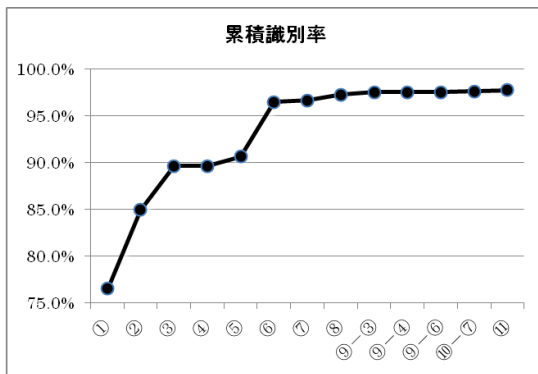


図4 累積識別率推移

本課題の成果は、英語の大学名称の識別において、小林らが達成した78%を大きく更新し、97.7%の識別が可能であった。これは、特定の分野と研究領域下であったこと、出現する大学が限定的であったこと、及び、事前に辞書を編集（最適化）したことが寄与したと考える。

## (2) 大学ランキング関連

ナレッジプールを組み込んだ指標(*KP*)と被引用量の指標(*Cited*)について、大学ランキングとの相関を比較する。主要な6つの大学ランキングで全分野における第1位から第100位までの大学の順位について、相関を比較したグラフを図5に示す。

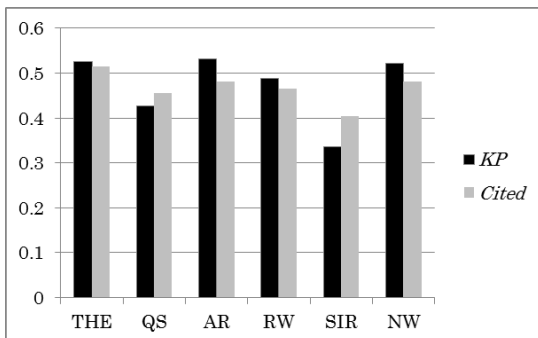


図5 全分野大学ランキングとの順位相関係数

*KP* と相関が強いのは、THE、AR、RW、NW で、逆に *Cited* と相関が強いのは、QS、SIR であった。前者の4つのランキングと *KP* の相関が高い理由は、ランキングを構成する指標に被引用量及び、文献数以外の要素が占める割合が、70%、80%、85%、50%と高いことが関係していると推測する。この割合が0%のSIRは *KP* と *Cited* の開きが最も大きくなった。ところが、この割合が80%と比較的高いQSと *Cited* において相関が強くなった原因を分析することが出来なかった。

そこで、比較する対象の対応関係を見直し、特定分野と特定研究領域の相関に限定し測

定した。THE、QS、ARには、物理学分野(:P)と工学分野(:E)の分野別ランキングがある。物理学分野の大学ランキングと天文学領域の *KP* と *Cited*、工学分野の大学ランキングと材料学領域の *KP* と *Cited* について順位の相関を比較したグラフを図6に示す。

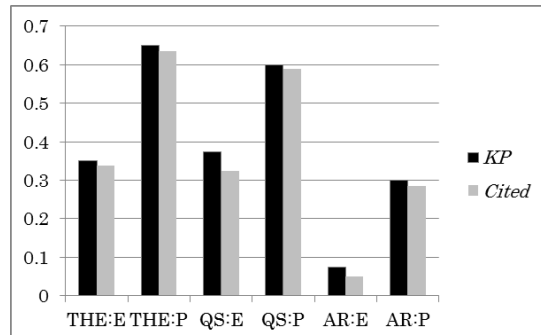


図6 分野別大学ランキングとの順位相関係数

全分野を比較した結果とは異なり、全ての比較で *KP* は *Cited* より相関が強かった。この要因として、ピアレビューが及ぼす影響を考察する。3つの大学ランキングにおけるピアレビューの指標が占める割合は、THE(教育環境評価と研究環境評価)が60%、QS(学術による評価と企業による評価)が50%、賞の選考もピアレビューによるものとすれば、AR(ノーベル賞などの賞の数と分野別の受賞者数)が30%となる。このことから、大学ランキングと *KP* の相関が強くなったのは、大学ランキングを構成するピアレビューが関係するとも考える。しかし、各大学ランキングを構成する指標は、それぞれ異なるピアレビューなので単純にその割合との相関を測るだけでは、関係を明らかにできない。これには、異なるピアレビューの質的影響度を比較する研究が必要となる。

本課題の研究成果は、大学の研究影響度を測定する用途に一步近づいた指標を提案したことである。しかしながら、本課題の研究成果は、物理学分野の天文学領域と工学分野の材料学領域に限定的である。他の分野と他の研究領域の関係においても、同様の傾向があるかについては未知である。今後、対象分野と研究領域に広げ、本指標を用いた調査と分析をすることにより、指標の有用性を示していきたい。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計2件)

① 魚田裕之、学術知識集積指標と大学ランキングの類似に関する一考察. 尚絅学園研究紀

要 A. 人文・社会科学編 第 6 号 2012 p. 79-85.  
査読無

②角田裕之(2011). 文献データベースにおける著者所属機関名を正規機関名に変換する方法に関する一考察. 尚絅学園研究紀要 A. 人文・社会科学編 第 5 号 2011 p. 35-57. 査読無

〔学会発表〕(計 2 件)

①角田裕之. 引用関係に基づく学術知識集積(ナレッジプール)を組み込んだ大学研究影響度指標の研究. 情報メディア学会第 13 回研究会発表資料. 東京, 2011-12-3, p. 5-8.

②角田裕之. 引用索引データベース(SCI)における所属機関(高等教育研究機関)の識別の研究. 第 59 回日本図書館情報学会研究大会発表要綱. 東京, 2011-11-12/13, p. 41-44.

## 6. 研究組織

### (1) 研究代表者

角田 裕之 (TSUNODA HIROYUKI)  
尚絅大学・文化言語学部・教授  
研究者番号：30454961

### (2) 研究分担者

### (3) 連携研究者

緑川 伸之 (MIDORIKAWA NOBUYUKI)  
筑波大学・大学院図書館情報メディア研究科・教授  
研究者番号：70166073  
芳鐘 冬樹 (YOSHIKANE FUYUKI)  
筑波大学・大学院図書館情報メディア研究科・准教授  
研究者番号：30353428