

## 2 F 2 1

# 共引用文献クラスター分析に基づく学際・融合研究の多面的評価

○仙石慎太郎, 小玉裕之 (京都大学)

学際連携・異分野融合研究(学際・融合研究)のパフォーマンス評価に関しては多くの取り組みがなされている。筆者らは、学術文献の共引用構造に着目し、クラスタリング分析とキーワード検索により学術分野のまとまり(コンピテンシー)をボトムアップ的に抽出する手法を適用した。評価指標として、従前の文献数、後方引用数に加え、学際連携の度合い、組織連携の度合いの導入を試みた。加えて、各コンピテンシーの成長性と文献シェアによる、研究拠点の戦略評価を行った。一連の取り組みは、こと基礎研究分野における体系的な研究評価、既存のマネジメント・フレームワークの適用可能性を強く示唆するものである。

### 1. はじめに

#### 1.1. 学術研究のガバナンス

学術研究のガバナンス・モデルにおいて、政府は科学技術政策の企画立案と執行、研究開発資金の配分と運用の責任を担う。受け手となる大学・公的研究機関や企業は、部局や研究拠点のミッションや各学問分野のディシプリンに従い組織管理を行いつつ、政府の事業の目的に沿うように、研究資金を効果的に運用しなければならない。各研究グループは、研究及び教育活動の一環として、担当する研究プロジェクトの実務を円滑に遂行することが求められる。

このようなモデルを想定した場合、主要経営指標(Key Performance Indicator, KPI)の開発と導入が必須となる。KPIを活用することによって、学術研究活動のパフォーマンスを包括的かつ客観的に評価することが可能となるからである。また、評価のみならず、評価結果に基づく活動のモニタリング、政府へのフィードバックに活用することで、いわゆる Plan-Do-See あるいは Plan-Do-Check-Action (PDCA) サイクルを充実させることができる。

現在筆者らの研究グループでは、学術研究のための KPI の開発と実践を進めている。本報告では、その取り組みの一環として、学術論文データベースに基づく新たな評価手法の構築状況、研究活動のマネジメントや政策評価への活用について述べる。

#### 1.2. 大学・研究機関の評価

学術研究成果の客観・定量評価に関しては、これまでに多くの先行的な取り組みがなされているが、依然として、ボトムアップ型の学術研究評価を行うことは困難である。その理由として、我々はここに4つの課題を指摘したい。

第一に、学際・融合研究や組織連携の広がりや、網羅的に探索する技術に乏しい。これまでは、調査対

象とする研究分野に関連すると思われる研究機関を、評価者が予め設定していた(例えば Corley [1])。このような方法では、意外な学際・組織連携の意外な繋がり、広がりを見逃してしまう可能性が高い。

第二に、個人レベルではなく、組織レベルにおいて、学術研究活動の生産性、学際・融合研究の充実度、組織連携の充実度を体系的に評価するための手法論に乏しい。

論文数や後方引用数或いは合成指標(Impact Factor 等)による生産性の評価手法は、これまでも存在し多く利用されている。しかしながら、こと学際・融合研究では、個々の論文が学術的貢献の範囲が複数の研究分野や所属組織に跨っているが少なくない。その場合、貢献度を単一の研究組織・所属組織に帰属させてしまうと、過大評価となってしまう。そのため、本研究では、複数の研究分野・所属組織間における、貢献度の配賦が必要である。

第三に、評価結果から意味合いを抽出し、打ち手に繋げるための、シームレスなマネジメント・フレームワーク、ツールキットが備わっていない。現状では、評価の議論と打ち手の議論とが分断しており、評価結果を効果的に打ち手につながるインフラストラクチャーが整っているとは必ずしも言えない。戦略評価と打ち手の考案のためのフレームワークは、コンサルティング企業等により提唱され、企業経営で活用されている。しかしながら、これらのプラクティスが学術研究のマネジメントに導入された形跡に乏しい。各種の検討結果は、組織マネジメントに簡潔かつ的確にコミュニケーションされ、意思決定と実践に反映させていく必要がある。しかしながら、個別各論に終始しており、マネジメントの関心時に応える統合的な指標、コミュニケーション・ツールは開発途上である。

第四に、これらの分析・調査を、日々の運用論として実現するための、ICT プラットフォームが存在しない。これまえ為されてきた多くの研究は、技術経営や研究政策の専門家によるフルスクラッチの分析に基づいていた。フレームワークやアプローチは数多く提唱され

てきたものの、オペレーションへの対応が等閑であり、これを日々実践することは困難である。また、文献データは刻一刻と変化するので、定期的なサイクルでのアップデートが不可欠であるが、これも従来のアプローチでは困難を極めることは自明である。

このような状況を打破するためには、専門家による文献データベースの提供と、然るべきICTインフラストラクチャーを備えた運用論とが確立されなければならない。

以上の考察より、本研究では、ファクトベースに根差し、組織的活動を評価するための、ボトムアップ型の評価系の構築を試みる。また、学術論文データベースのデータと直結した、実用的なマネジメント・フレームワークの確立を目指す。

### 1.3. 研究の目的

本研究の目的は、論理的手続きに基づく、大学・公的研究機関の学術研究マネジメントの充実である。とりわけ、学術研究活動の効果的な可視化及び評価手法の開発を目指す。そのなかで、学術論文データベースというファクトベースに基づくKPIを開発し導入する。

## 2. 研究手法

### 2.1. データベースと分析ツール

我々は本研究を進めるうえで基盤となる学術論文データベースとして、Elsevier B. V.社のSciVerse Scopus<sup>1</sup>を用いた。また、コンピテンシーの抽出・分析ツールとして、同社のSciVal Spotlight<sup>2</sup>を活用した。Spotlightは、コンピテンシーと呼ばれる、学術研究の集積分野を同定・可視化するためのデータベースである。手法としては、論文間の共引用関係に基づくクラスター分析で抽出されたクラスター候補に対して、論文シェア(Article share)、被引用文献数、新規性の3点からスクリーニングをかけて、最終的にコンピテンシーを抽出する[2-4]。

いずれのデータベースも外販されており、数多くの大学・公的研究機関において使用の実績がある。また、分析・評価の一部は自動化されており、フルスクラッチの作業に比べ、遥かに簡便に運用することができる。よって、これらのツールを活用することで、インプリメンテーションが促進される。

### 2.2. KPIの定義

本研究では、学術研究の生産性を評価するKPIとして、以下の5つを設定した。

#### 2.2.1. 研究成果の量

研究成果の量的指標は、当該コンピテンシーにおける、Fractionalized Articles (FA)の数として定義した。FAの原義上、対象となる文献が複数のコンピテンシーに貢献している場合、文献数の単純合計を対象文献が関与するコンピテンシー数で配賦して、当該コンピテンシーに限定した貢献度のみを評価するよう工夫した。配布方法は、単純割り算とした。

#### 2.2.2. 研究成果の質

研究成果の質的指標は、当該コンピテンシーのFAに対するCitation counts (CC)の数として定義した。対象となる文献が複数のコンピテンシーに貢献している場合、文献数の単純合計を当該FAのコンピテンシー数で配賦して、当該コンピテンシーに限定した貢献度のみを評価するよう工夫した。配布方法は、単純割り算とした。

#### 2.2.3. 研究の生産性<sup>3</sup>

研究の生産性指標は、当該コンピテンシーにおけるFAあたりCCと定義し、これをPublication performance (PP)と命名した。

#### 2.2.4. 研究成果の学際性

研究成果の学際性指標は、先行研究における事例を引用し、Herfindahl-Hirschman Indexの寡占度指数(HHI index)を流用し、Interdisciplinarity Breadth (IB)として定義した[5,6]。本公式で得られる研究分野の広がり、SciVal Scopus及びSciVerse Spotlightの定める、研究分野(Disciplines)のカバレッジとして計算した。具体的には、IBは以下の式で定義される。ここでiは、各disciplineを指す。

$$IB = 1 - HHI = 1 - \sum_i \left[ \frac{FA_i}{FA_{total}} \right]^2$$

#### 2.2.5. 組織連携の充実度

組織連携の充実度は、FA Concentration Ratio (FACR)を新指標として開発・導入した。FACRは、組織内部の連携の度合いを、単一指標により見積もるこ

<sup>1</sup><http://japan.elsevier.com/products/scopus/>

<sup>2</sup><http://japan.elsevier.com/products/scivalspotlight/index.html>

©Copyright 2009-11 The authors and the Institute for Integrated Cell-Material Sciences (iCeMS), Kyoto University. All rights reserved.

<sup>3</sup>生産性の原義は投入資源に対する成果物の量或いは質であるが、本分析では便宜的に、成果物の単位量(論文1報)生成の投入資源一定を仮定した。

とができる。具体的には、FACR は以下の数式で表される。

$$FACR = \frac{\sum_r FA_r}{FA_{total}}$$

ここで r は、当該コンピテンシーにおいて高 FA を誇る研究者 (Top Author) を指す。

FACR が大きい場合 (組織内連携が密な場合) とは、当該コンピテンシーにおいて Top Author がドミナントであり、かつその間で重厚な共著者関係が観察されている状態である。一方、FACR が小さい場合 (組織内連携が疎な場合) とは、Top Author がドミナントでなく、貢献が多くの研究者に分散した構造となっている、及び/或いは、Top Author がドミナントだが、その間の共著者関係が疎である状態を指す。より精緻な評価はネットワーク分析に委ねる必要があり、本指標はあくまで簡便化された指標ではあるが、後述の事例研究を通じて、運用上の有用性は担保されているものと判断する。

### 2.3. 事例研究

#### 2.3.1. 対象とする学術研究分野

事例研究として、幹細胞 (stem cells) 研究とその周辺領域を設定した。具体的には、コンピテンシーの選別キーワードに "stem cell" を指定することで、関連するコンピテンシーのみを抽出した。

#### 2.3.2. 対象とする地域及び大学・研究機関

地域レベルの分析では、日本、米国、及び英国の3地域を選出した。大学・研究機関レベルの分析では、本分野に関係性の深い、ハーバード大学 (HU)、ウィスコンシン大学マディソン校 (UWM)、エジンバラ大学 (UE)、及び京都大学 (KU) の4つの大学を選出した。

## 3. 結果

### 3.1. 事例研究におけるコンピテンシーの記述統計

大学レベルの評価結果を、表 1 に示す。ここで基本統計量として、コンピテンシーの数に加え、第3章で設定した KPI の各々について、最小値、最大値、平均値及び標準偏差を示している。

大学名	KPI	事例数	最小値	最大値	平均値	標準偏差
HU	FA	64	5	1,043	198	347
	CC	64	18	21,303	3,619	6,374
	PP	64	3.9	40.4	16.7	8.6
	IB	64	0	0.95	0.53	0.25
	FACR	64	0.17	2.6	0.88	0.48
KU	FA	9	5	336	96	117
	CC	9	36	6,448	1,220	2,092
	PP	9	3.3	19.2	9.3	4.9
	IB	9	0.4	0.91	0.61	0.21
	FACR	9	0.58	2.66	1.47	0.62
UE	FA	10	4	64	19	18
	CC	10	8	670	239	230
	PP	10	1.8	28.2	13.2	9.5
	IB	10	0.28	0.84	0.56	0.15
	FACR	10	0.57	2.42	1.44	0.68
UWM	FA	15	5	118	32	32
	CC	15	29	2,151	427	541
	PP	15	3.5	31.4	12.4	8.1
	IB	15	0	0.82	0.46	0.25
	FACR	15	0.81	3.94	1.95	0.82

表 1. 各大学事例における設定 KPI の記述統計値

### 3.2. KPI マトリクスによる評価

本節では、導入した KPI をもとに組成したマトリクスでの、研究活動の評価結果を示す。いずれの分析においても、研究成果の生産性 (PP) を非従属変数に、研究成果の量 (FA)、学際性 (IB) 及び組織連携の充実度 (FACR) を独立変数に設定した分析結果である。

図 1a に、生産性 (PP) - 論文数 (FA) のマトリクス評価を示す。各バブルの大きさは、各コンピテンシーにおける FA 数を示す。結果、FA が小さいセグメントにおいては、PP に大きなばらつきが観察された一方 (FA < 200 において、PP = 14.5 ± 9.18)、高い FA が大きいセグメントにおいては、PP が一定値に収束する傾向が観察された (FA > 200 において、PP = 17.4 ± 3.46)。この結果は、コンピテンシーの規模要件が整うことで安定的な生産性を達成することを示唆している。一方、低コンピテンシー・セグメントのばらつき、及び高 PP の根拠については、何ら示唆を与えていない。

そこで我々は、生産性 (PP) - 学際性 (IB) 相関の分析を試みた (図 1b)。結果、本分析においては、高 PP セグメントが、IB = 0.3 - 0.7 の範囲において分布していることが観察された。また、このセグメントの平均 PP 値は、単一ディシプリンに根差すコンピテンシー (IB = 0) と比較しても大きかった。本結果は、高生産性と学際性との間に何らかの関連性があること、及び/或いは、一定範囲において学際研究を推進することの意義を示唆するものである。

更に我々は、生産性 (PP) - 組織連携の充実度 (FACR) 相関の分析を試みた (図 1c)。結果、本分析においては、高 PP セグメントが、FACR = 0.5 - 2.0 の範囲において分布していることが観察された。本結果は、高生産性と組織連携の充実度との間に何らかの関連性があること、及び/或いは、一定レベルの組織内連携が生産性を亢進する可能性を示唆するものである。

この一連の分析を通じて、高生産性 (PP) の達成要

困を、文献データに基づき、多面的期に評価することが可能となった。とりわけ、PP-FA のようなコンベンショナル分析では明らかにされなかった示唆、例えば一定範囲の IB や FACR が PP を高める可能性が、本分析から得ることができた。

### 結びに変えて

本研究では、文献の共引用ネットワーク構造のクラスタリング分析を通じて、学術研究の強み分野を「コンピテンシー」として表出する分析アプローチを採用した。また、これまでの研究の量と質の評価に加え、学際性の度合い及び組織連携の度合いを評価する主要経営指標 (KPI) を考案・導入した。各コンピテンシーにおける研究の生産性との関係性を測定したところ、ある一定の度合いの学際性や組織連携が、研究の高生産性を支えている可能性が示唆された。

これらの結果は、コンピテンシーに基づくボトムアップ型の研究評価アプローチの、組織的な学術研究マネジメントに対する有用性を支持している。但し、本研究で得られた KPI やマネジメント・フレームワークに普遍性を持たせるためには、より多様な研究分野や大学・研究機関への適用と検証が不可欠である。今後は本手法を発展させ、産業クラスター等の地域レベル、また国家レベルへの適用、ナショナル・イノベーション・システムの在り方の検討にも役立てていきたい。

### 謝辞

本研究は文部科学省世界トップレベル研究拠点 (WPI) プログラム、及び内閣府最先端・次世代研究開発支援 (NEXT) プログラムのもと実施している。京都大学 iCeMS、エジンバラ大学 MRC 再生医療センター、京都大学アカデミック・イノベーション・マネジメント研究会の諸氏には貴重な助力を頂いた。ここに感謝の意を表します。

### 参考文献

- [1] Corley E., 2006. Boardman P., Bozeman B., *Res. Policy* 35: 975-93.
- [2] Boyack, K. W., 2009. *Scientometrics* 79 (1), 27-44.
- [3] Klavans, R., Boyack, K. W., 2006. *Scientometrics* 68 (3), 475-99.
- [4] Small H., 1973. *J. Am. Soc. Info.* 24 (4), 265-9.
- [5] Porter, A., Cohen, A., Roessner D., Perreault, M., 2007. *Scientometrics* 72(1), 117-47.
- [6] 安西智宏, 草間亮一, 仙石慎太郎, 2010. 「アカデミック・イノベーション・マネジメント③: 異分野融合型研究拠点のマネジメントと評価」 研究技術・計画学会第 25 回年次学術大会.

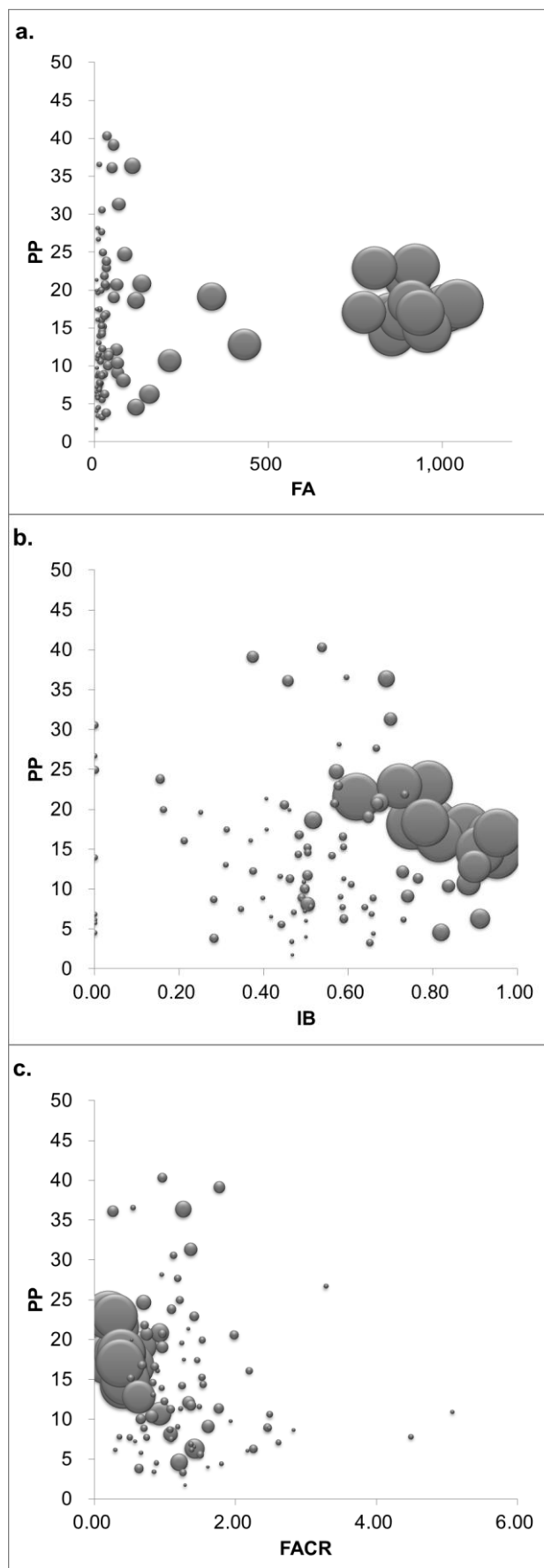


図 1. 設定 KPI 間の相関図。定義及び軸の単位は、文中を参照。丸の大きさは FA (本数) を表す。