

大規模書誌情報のネットワーク構造に基づく研究多様性指標

濱田 ひろか データ科学研究系 特任研究員

1. 研究力を評価する指標

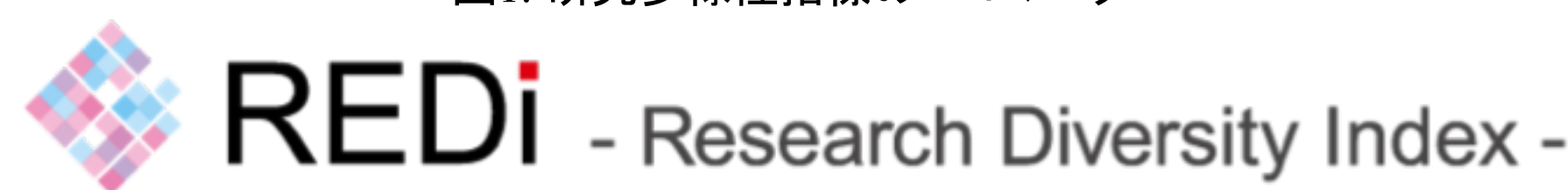
総合学術雑誌 Nature 2017年3月23日号の特別企画冊子「Nature Index 2017 Japan」[1]では、日本の科学成果発表の水準がこの10年間で低下していることを示すレポートが掲載されました。このレポートに関するプレスリリース[2]では、日本の論文出版数の低下と同時に、日本政府が開始した取り組みの1つである長期的な職位を増やすための支援に対する期待感なども書かれました。平成25年度以降、文部科学省では研究大学強化促進事業推進委員会の設置や研究大学強化促進事業を開始するなど様々な形で対策を講じており、いま日本の大学・研究機関では、これまで以上に研究に対する自己評価能力の重要性が高まっています。

近年、大学における「Institutional Research」(IR)[3]が注目を集めています。そのIR活動の一つとして、大学や研究機関あるいは研究者個人の研究成果に関する活動があります。これは研究活動の進展や効果を客観的に評価することを目的とします。研究成果の一つである論文の評価に広く利用される代表的な指標(表1)は、Cited Count, Journal Impact Factor (IF), Top10%論文などいくつかありますが、これらは全て被引用数に基づく指標です。被引用数が大きい論文は確かに影響力を持つと言えます。しかしながら、研究成果の価値が一概にそれのみでないことは、ご存知の通りです。

表1. 論文評価のための主な指標

名称	対象	概要
Cited Count (被引用数)	論文	他の論文からの被引用回数。
Journal Impact Factor	学術雑誌	過去2年間に学術雑誌にて掲載された論文が当年に引用された回数をその掲載論文数で除した値。
Top10%, Top1%	論文	被引用数が各分野の上位10%(あるいは1%)に入る論文の数。

図1. 研究多様性指標のロゴマーク



2. 論文の多様性を評価する指標を開発

そこで私たちは、大学・研究機関における研究成果の評価には、より客観的な外部情報を根拠とした**多様な価値観や評価軸に基づく研究成果の分析手法が必要である**と考え、研究の成果物である論文を評価するための新しい研究力評価指標の開発を行いました。

その一つが、**研究多様性指標 (Research Diversity Index, REDi)**です。イノベーションを引き起こす重要な要素と考えられている「多様性」という観点を加えた論文の評価が可能で、私たちはこの多様性という概念を論文の評価軸として採用するにあたり、研究における多様性とは「**より多くの、より遠くの異分野に広く影響を与える研究**」であり、研究多様性指標はそれを評価する指標であること、と定義しました。

さらにこの新指標にはいくつかの目標を定めました。一つ目は「ボトムアップ型研究助成機関における意思決定に資すること」であり、二つ目は既存指標において長年指摘されている「異なる分野間では比較することができないという問題を解消すること」などです。

3. REDi開発手順

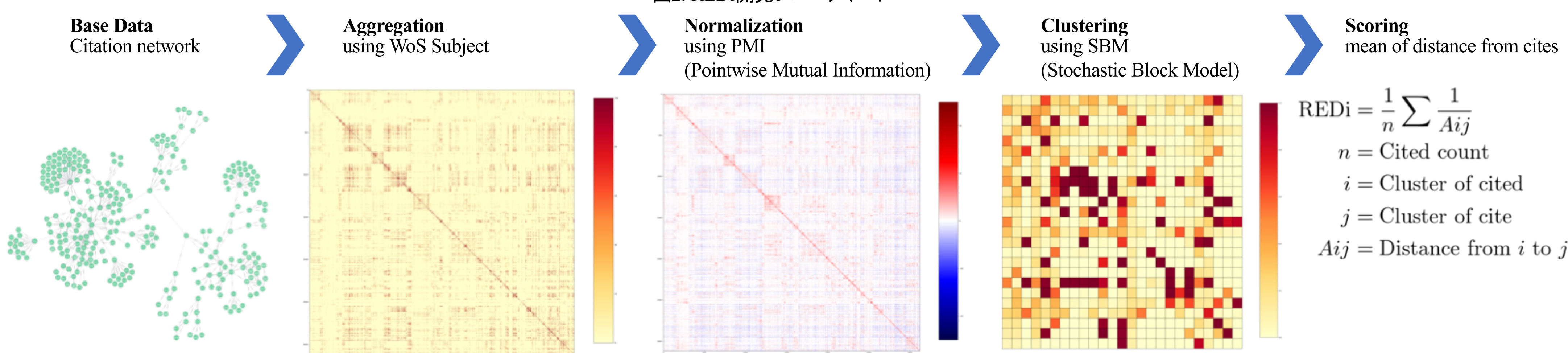
研究多様性指標の目標として掲げた内容を解決する手段として、自然言語処理の分野でよく使われる単語の共引用回数を補正するテクニックであるPointwise Mutual Information[4]と、ネットワークデータのクラスタリングに広く用いられているStochastic Block Model[5]を採用しました。この二つのテクニックを組み合わせることにより、論文の引用-被引用ネットワークから関係性を考慮した**潜在的な分野と、その分野間における繋がり**の強弱を結果として得ることができます。これらの結果から、評価対象とする論文およびそれを引用した論文の潜在的な分野を特定し、より多くの、より遠くの異分野に広く影響を与えた研究を可視化するREDiを定義します(図2)。

REDiは、同程度の被引用数を持つ論文の多様性を比較するために役立ちます(表2)。

表2. 研究多様性指標による論文の評価

WOS ID	Journal Name	Cited Count	REDi
WOS:000306296200061	NANO LETTERS	99	24.44
WOS:000303760600002	SMART MATERIALS AND STRUCTURES	95	422.62
WOS:000298943100058	NANO LETTERS	95	1.43
WOS:000311496000009	ZEITSCHRIFT FUR ANGEWANDTE MATHEMATIK UND PHYSIK	96	585.59
WOS:000306824100008	CARBON	95	84.78
WOS:000304905300047	SCIENCE	96	12.59

図2. REDi開発フローチャート



<謝辞>

本研究は、統計数理研究所共同プログラム重点テーマ2「IRのための学術文献データ分析と統計的モデル研究の深化」の助成を受けたものです。学術文献データについては、クラリベイト アナリティクス社からの支援を受けています。

<参考文献>

- [1] Nature Index 2017 Japan, Nature
- [2] Mark Staniland (2017) Japanese science stalls over past decade, threatening position among world's elite, http://www.nature.com/press_releases/nature-index-2017-japan.html
- [3] 小林雅之, 山田礼子 (2016) 「大学のIR:意思決定支援のための情報収集と分析」, 慶應義塾大学出版会
- [4] Gerlof Bouma (2009) Normalized (Pointwise) Mutual Information in Collocation Extraction, Proceedings of GSCL
- [5] Krzysztof Nowicki & Tom A. B. Snijders (2001) Estimation and Prediction for Stochastic Blockstructures, Journal of the American Statistical Association