

# 研究業績とは何（であるべき）か？

How is (or Should) the Performance of Academic Researchers (be) Evaluated?

佐倉 統\*

Osamu Sakura

## 1. 序論

研究業績とは何だろうか。この問いは、一見、自明に思えるかもしれない。研究者は、学術的かつ専門的な知識を生産する職業である。大学院生はその知識生産の方法や技能を身につける訓練を受け、学位論文（修士論文、博士論文）として提出することが課程修了の要件として課せられている。このように、論文を作成する作業はプロセスと手法がすでに確立しているから、何が業績かについても確立した基準が存在するように思える。

たしかに、研究分野を特定して考えれば、その分野での業績評価の基準はおおむね一定している。だからこそそれぞれの学会では優秀な研究に賞を出し、称賛することが何の違和感もなく行なわれている。当該学会員の間で評価基準が共有されているからだ。だが、その学会での業績評価基準が別の学会でも通用するかどうかは、別の話だ。実際、学問分野が異なれば評価基準は異なるし、分野間の隔たりが大きくなればばるほど、業績評価基準の違いも大きくなっているのが現状である。

ひとつの学会における活動だけに注目するの

であれば他の学会との交流についてはあまり考える必要がないから、複数の学会の間で評価基準が異なってもとくに大きな問題はない。しかし、昨今学際的な研究教育活動が奨励され、他分野の研究者と共通のテーマで活動する機会は格段に増えている。同一のプロジェクトに異分野の研究者が参加することも珍しくない。そのような場合、プロジェクト全体の研究業績評価と異なる基準を採用している分野の研究者は、活動が公平に評価されないこともあるかもしれない。

また、インターネットやSNSの発達により、以前ならひとつの学会や研究コミュニティの内部で閉じていた情報の流通が、その外に一気にあふれて流れ出す状況もしばしば見られる。そのような、あふれ出した情報は文脈や場面から切り離されて無定見に流通してしまうため、本来であれば直接比較はできない異なる分野間での研究業績をうわべだけ単純に比較して、背景知識に詳しくない第三者が優劣を軽率に判断してしまう場合も散見される。

そして、情報学環・学際情報学府のような学

\* 東京大学大学院情報学環、理化学研究所革新知能統合研究センター

キーワード：業績評価、定量評価、定性評価、ライデン声明、学際組織

際的な研究教育組織では、学問分野の距離が遠い分野の教員や学生が同じ組織で教育研究をおこなっているため、しばしばこの問題に直面せざるを得ないことになる。そのひとつの例が毎年度選定している最優秀修士論文（あるいは博士論文）だ。その年度の学位論文の中からもっとも優れたものをひとつ選び、全学の総長賞候補として推薦しなければならないのだが、この候補者しぼり込みの際におよそ専門分野の異なる論文を横並びで比較検討して「最」優秀を決めるという無理難題が生じる。

それでも教員や研究員のようにある程度確立した研究者は自分たちの研究分野を持っているから、まだよい。しかし学際情報学府の学生たちは、とくに修士課程の間は所属学会を模索しつつ修士の研究を進めることも珍しくない。あっちの学会こっちの研究会と、異なるところで発表しているうちに異分野での文化や価値観の違いを実際に体験してとまどい、果たして自分はどのような評価尺度に沿って修士論文を書いていけばよいのか、迷う学生がいることもしばしば経験するところだ。

学際情報学府を卒業した学生のすべてが研究者を目指すわけではもちろんないが、学位論文（修士論文、博士論文）は教育活動の集大成であると同時に、立派な研究成果である。「国立大学法人東京大学における研究活動上の不正行為の防止に関する規則」では第3条第2項で「研究者」とは、東京大学において研究活動に従事している者をいう」と定義され、学位論文の作成や研究室のプロジェクトに参加している学生も含まれる<sup>(1)</sup>。ゼミや研究会での発表も、閉じた空間ではあるが、研究成果の発表だ。学生

たちの日々の活動の成果は「研究業績」として評価されるのであり、その評価尺度の分野間での多様性はしたがって学生の日々の評価にも直接影響することになる。

そして、大学院卒業後に研究職ではない人生を歩む人にとっても、仕事の成果がどのような基準や尺度で評価されるのかという点では、学術研究の場合と共通する特徴が多々あるはずである。個人の活動を、その個人が所属する社会や共同体がどのように評価するのか、その際の評価尺度として何が適切なのかという構造的問題は、学術研究以外の場でも共通だ。「重要な発見をしようとするよりも、自分の発見を重要なものにするに、努力すべきである」というのは、発生物学者・白上謙一（1972: pp. 83-84）が説いた研究の心構えのひとつだが、学術研究に限らず企画立案や営業の新規開拓などにも当てはまると思う。

この論文で私が主張したいことは、業績（成果）の評価は多元的であるべきだし、多元的でなければならない、ということである。これは、なんでもありを意味するものではない。どのような多元性が健全な学術研究の発展のために適切なかを熟考する必要がある。その際に依拠するのはライデン声明（Leiden Manifesto）である。詳細は第2節で述べるが、2015年に世界で最も権威のある自然科学学術誌 *Nature* に発表された宣言書（Hicks et al., 2015）で、学術研究の健全かつ社会的に有用な発展のためには研究業績評価の多元的な基準が必要であることを論証し、それを実現する方針となる原則を10項目にまとめたものである（小野寺・伊神 [2016] も参照）。このライデン声明の意義は、

そもそも学術研究が社会のためのものであるという原点に立ち返って、そのために必要な研究業績とはどのようなものであるべきかを考察し、それにもとづいて定量的な評価尺度の位置づけを再定位したことにあつた<sup>(2)</sup>。

これらの動向を踏まえての本稿の第1の目的

## 2. 業績評価の実態と問題点

### 2.1 業績評価の問題点

学術研究者は知識の生産者であるところでは定義する。生産される知識の性質が研究分野によって異なるのは知識が多面的な存在である以上当たり前のことであるが、昨今、えてしてこのことが十分認識されないまま、単一の尺度でのみ研究業績が評価されることが多い。学術界の内部においても、一般社会においても同様である。

学術界内部の状況は、研究者個人の評価と研究組織の評価の両方において、定量的評価の過剰な使用が指摘されている。個人については、研究職への就職や昇進の際のよりどころとして論文掲載誌の重要度を表わすとされているインパクトファクター（IF）や、研究者個人の学術的影響力を表わすとされているh指数といった定量的指標の提出が要求されたり（Marson, 2020；それぞれの指標については次節で述べる）、北欧や中国、パキスタン、韓国の大学ではこれらの定量的評価にもとづいて研究報奨金を出すところもあるという（Fuyuno & Cyranoski, 2006; Hicks et al., 2015）。組織については、世界中の大学を共通の尺度で順位付けする大学ランキングが重視される傾向が強まっ

は、研究業績とはなにかを考えることで学際的な研究活動・教育活動のありかたについての見通しを得ることであり、第2の目的は、そもそも研究とか学術とかは何のためにあるのかを考えるきっかけを、読者、とくに学生や若手研究者に提供することである。

とくに影響力が大きいのは、中国の上海交通大学が始めて現在は上海ランキング・コンサルタントが行なっている世界大学学術ランキング（Academic Ranking of World University: ARWU）、イギリスの大学評価専門民間会社 Quacquarelli Symonds による QS 世界大学ランキング、そこから分離独立した Times Higher Education による THE 世界大学ランキングで、これらのランキングの尺度が主観的であるなど不完全さが繰り返し指摘されていながらも、大学管理者自身が数字に右往左往させられているのが現状である（麻生, 2021）。

学術界の外、一般社会においても、研究者の優劣を一元的な尺度で測定する風潮が強まっている。たとえば公的機関への国や自治体からの予算削減や、メディアやネットでの批判の論拠として、当該分野への適用が不適切である尺度が使われることがある。2020年10月に学術会議が推薦した会員候補者6名を、それまでの慣行を破って菅義偉首相が任命拒否した問題は耳目を騒がせたが、その際、当該6名のh指数が低いことを根拠に、これらの研究者が学術会議会員に値しないという論調がネット上で盛り上

がりを見せた。これらの研究者は人文系・社会系の人たちなのでh指数を適用すること自体が業績評価として不適切で、いわばホームランを打っていないからといってサッカー選手を「だめなスポーツ選手」と評価するようなことなのに、そのような配慮がなく、数字だけがひとり歩きしてしまったのだ。さすがにこの状況には反論が起り、修正がなされたが（牧野, 2020）、そもそもこのような形で誤用されてしまうこと自体、定量的評価が過度に行き渡っている状況を示唆する。

この件で私がとくに危惧したのは、学術界「内部」の人たちから、同様の批判がなされたことである。「名月論」という匿名のプロガーは若手脳神経外科医を自称しているが、自身のブログで以下のように述べている：

h-index という研究者の評価指標でも [ 任命拒否された6名のうち ] 5名が0点、1名がわずか2点ということで研究者業績も低いとネットで暴露されています（中略）。

ちなみに、私自身の h-index を scopus のサイトで調べてみたところ、私ですら10年程度の研究歴ですが、7点ありました。（名月論, 2020 ; [ ] は引用者）

これを「同意見です」として引用している別のブログ（奥永, 2020）の執筆者も、自分のh指数は26と書いてあるので専門分野はわからないが研究者であると思われる。このような、活動の実態も評価軸も異なる分野の研究業績について、異分野の研究者が軽々に、さもわかっているかのように批判する風潮は、学術界全体

の活力を削ぎ、社会からの信頼を失わせるのではないかと危惧する。

なお、誤解されないように付け加えておきますが、私の趣旨はこれらのブログの執筆者を批判することではない。彼ら彼女らがこのような発言をする背景には日本の若手中堅研究者をめぐる過酷な研究・就職環境があり、年長の、学術会議に推薦されるような研究者たち、ひいては学術会議そのものへの不満が潜在的に横たわっていると思われる。先ほどの「名月論」は上記引用箇所続けて以下のように述べている：

当然、私は臨床医でアカデミアの端くれですが、学術会議の会員でもなんでもありません。

0点や2点の人が学術会議の会員で公金の恩恵を受けるの???

その人達本当に研究しているの???

と正直に思います。（名月論, 2020）

ここにはそのような苛立ちが鮮明に表れている。彼ら彼女らが手近な指標を使ってその苛立ちをぶつけるのは当然のことだろう。問題なのは、h指数やインパクトファクターなどの定量的評価指数が、一見わかりやすく客観的であるかのように見えて、実はさまざまな問題を含んでいる指標であり、「危険！ 取り扱い注意」だと十分に周知されていないことだ。銃は護身用にも使えるかもしれないが、取扱い方を間違えれば人を殺傷することのできる破壊力を持っている。使用法についての周知徹底や規制の厳格化なしに銃を使える状態にしておいて銃乱射事件で多くの死傷者が出たとしたら、それはそ

の社会の規制の仕方が間違っているのである。乱射事件を起こした張本人のことをいくら責めても、問題の解決にはならない。研究業績の定量的評価指数も同じことだ。

だがそもそも、いったいなぜ、このような定

量的指標がひとり歩きしてしまう状況が出現したのか。そもそも、これらの指数は何を表わしており、どのような経緯で開発されたのか。次節で背景を振り返ってみよう。

## 2.2 科学計量学の出現と席捲

学術研究成果の定量的評価が本格的に始まったのは20世紀後半である。1950年代、論文の引用・被引用関係を分析することで科学研究の知識ネットワーク構造を可視化する作業をおこなっていたアメリカの言語学者 Eugene Garfield (1955) は、自身の開発した方法の有効性を *Science* 誌に発表し、翌1956年に科学情

報研究所 (Institute for Scientific Information: ISI) を設立した。これをもって現代の科学計量学がスタートしたと言ってよいだろう [図1]。なお、本稿では科学計量学そのものについてはほとんど紹介しないので、より詳しく知りたい方は藤垣ら (2004) を参照していただきたい。



図1 科学計量学の始祖、Eugene Garfield (Wikimedia)  
(Fig 1 Eugene Garfield, the founder of scientometrics)

自然科学分野を対象として、重要な論文はたくさん引用されているという前提のもと、論文を被引用数によって順位付けした科学引用指数 (Science Citation Index: SCI) が開発され、世界中の多くの分野で「便利な指標」として広く使われようになった。後に社会科学引用指数 (Social Science Citation Index: SSCI)、人文学

引用指数 (Arts and Humanities Citation Index: A&HCI) も開発されているが、SCIほどには使われていないようである。

1970年代になると、Garfield (1972) と ISI は個別の論文だけでなく学術専門誌がどれだけ引用されたかを定量的に測定する指標インパクトファクター (Impact Factor: IF) を開発した。



厳密には彼らのオリジナルではなく、図書館が限られた予算内で購入する雑誌の優先順位を決めるのに被引用数を尺度に使っていたところからヒントを得たらしい（麻生，2021）。IF は基本的には、ある学術専門誌に掲載されたすべての論文の科学引用指数の平均値である（過去何年間の期間で見るとは任意だが、通常は過去2年間のデータを使う）。この値が高い雑誌ほど、「重要」で「影響力が大きい」とみなされる。先にも書いたように、Garfield 自身の関心は、さまざまな論文がお互いにどのように影響しあって科学知識の生産過程が動いているのか、そのダイナミクスを可視化することだった（実際、そのようなネットワーク図も描いている [Garfield, 1970]）。論文や学術雑誌の評価尺度として SCI や IF を開発したわけではない。だが、彼の思いや意図とは別に、IF はすなわち学術雑誌の「格」であると世の中一般から解釈された。学術雑誌の編集者は IF を高めるための努力を出版社から要求され、研究者は少しでも IF の高い雑誌に論文を掲載しようとやっきになった。その結果、先にも触れたように、研究者の就職や昇進の際に提出する研究業績リストには論文掲載誌の IF も並記するよう要求される場合も散見されるようになった。

Garfield と ISI による科学計量学が発展したもうひとつの要因は、論文のデータベース化である。現在は Web of Science として知られている学術データベースは、当初は Web of

Knowledge として、やはり Garfield=ISI が発明したものだった（現在は ISI の後継組織である Clarivate Analytics [旧 Thomson Reuters] が管理運営）。ISI 発足当初は電子化されたデータをパッケージメディアで配布していたが、インターネットの普及によってウェブ上でこれらのデータベースが使用出来る環境が出現すると、SCI や IF などの定量的指標は一段と手軽に、いつでもどこでも使えるようになり、その普及に一段と拍車がかかった。

2005 年には研究者個人の研究生産量を表わす  $h$  指数 ( $h$  index) が提言され (Hirsch, 2005)、これが「当該研究者の研究コミュニティへの貢献度」を示す指標として広く使われるようになっていく。ある研究者の  $h$  は、「その研究者が発表した論文のうち  $h$  回以上引用された論文が  $h$  本あるという条件を満たす最大の値」である。現在、 $h$  指数は Web of Science などのデータベースや文献検索ツールから簡単に調べることができるようになっており、そのお手軽さもあって、研究職の就職や昇進の際にしばしば参照されるのは、すでに述べたとおりである。しかし研究能力の一面しか表わしていないこれらの指標を判断基準にすることは不適切であり、訴訟沙汰になるのではないかという予想もなされている (Marson, 2020)。このほかにさまざまな場面で業績の定量的評価には負の側面が目立つ状況になっており、その行き過ぎを是正する動きが活発になってきている。

### 2.3 梅棹忠夫による定量的研究活動評価

本論の本筋とは関係ないが、研究業績の定量的評価の歴史について、日本の状況に関して補

足をしておく。1989 年に、当時国立民族学博物館（以下、民博）の館長だった梅棹忠夫が、

研究業績の定量的評価の試みをおこなっているのである。梅棹は、すでに1977年に民博の活動をさまざまな角度から評価することの必要性を説いており、それから12年後に、博物館の活動評価の一環としてスタッフ個人の研究業績を定量的に評価する指標の作成を実際に試みた。ここで梅棹（1989）がおこなったのはGarfield=ISI流の引用・被引用ネットワークを使うものではなく、知識の生産量の単純な定量化であった。研究者の平均給与を各研究者が1年間に書いた論文の総ページ数で割り、研究者ごとのページ単価、つまり「原稿料」を算出したのである。しかし彼はSCIやIFには触れておらず、この時点で視野に入っていなかったものと思われる。

ここで注目したいのは、論文の「質」の評価を評価者（=梅棹）は放棄していることだ。査読付きの学術専門誌に掲載されたものは、学会という専門家集団が質を保証したものだとしてそれ以上の評価には踏み込まず、一律に総ページ数という単一の尺度で測定している。つまり、学会誌・専門誌に掲載された論文は、掲載された雑誌の違いにかかわらず、どれも同じ重みで評価されているのだ。GarfieldのSCIはまさにこの点で突破口を開いたものであり、学術論文の「質」という、それまで専門家以外には

評価できなかったものを、「数字」という誰が見ても一目瞭然の指標に変換したからこそ、世界中でこれだけ広く受け入れられ、使われるようになったのである。梅棹はその点では奥ゆかしいというか踏み込みが足りないというか、今から見ればいかにも中途半端な定量化にとどまっている。

「質」に踏み込まないというのは、この件に限らず梅棹の思考全般に見られる特徴のひとつである（佐倉，2011）。そのことと合わせて考えると、梅棹は、学術論文ひとつひとつの「質」を定量評価の俎上に載せることはパンドラの箱を開けるようなことだと、直観していたのかもしれない。生態学的思考を得意としていた彼が、論文の引用・被引用関係による知識の生態系構造を頭の中に描いていなかったとは思えない。どの程度明確に自覚していたかはさておき、質の定量化の一手手前で踏みとどまった方が安全だと感じたのではないかと推測する。そして、Garfield流の過激な定量化の弊害があまりにも目立ってきている現状を踏まえると、梅棹がそこまで踏み込まなかったことの意味を、世界の潮流を知らなかったとか中途半端だったと切り捨てるのではなく、もう一度積極的に評価し直すことが、学術業績評価のこれからを考える際に役に立つのかもしれない。

### 3. 多元的評価の重要性

#### 3.1 定量的評価の限界

定量的評価は知識生産の動態を可視化するのに威力を発揮するが、研究者や研究組織の活動業績を評価する指標として使うことには弊害が

大きい（Seglen, 1994; Lawrence, 2003, 2007; Greenwood, 2007; Caves, 2014; Marson, 2020; 麻生, 2021 など）。全体の状況を把握するには有

効だが、個別の事例を評価するにはばらつきが大きすぎて、精度の高い測定ができないのである。たとえば SCI/IF は学術専門誌に掲載された論文の引用だけを対象とするが教科書に引用された例数は数えない (Marson, 2020)。しかし標準的な大学の教科書に古典的な研究例として引用されることは、大きな影響力をもつ。IF はその雑誌に掲載される個々の論文の SCI の分散が大きく、IF が非常に高いか非常に低い数誌を除けば個々の雑誌の「実力」を適切に表わしているとはいえない (Seglen, 1994; Greenwood, 2007)。引用の慣行も分野ごとに異なるから分野を超えた比較には使えない。IF は共著者の多い論文ほど高くなることが知られており (Amin & Mabe, 2000)、研究の性質によって単独指向の強い分野 (数学や人文系 [Larivière et al., 2006]) は不利になるし、誰を共著者とするかの慣行も分野ごとに異なる。また非英語圏ではその国や言語圏独自の課題を解決するための学術的活動も重要であり、英語論文だけを対象とする SCI や IF はそれらの重要性を軽視することになる (Hicks et al., 2015)。その国の言語で書かれた総説論文は、教育的効果も大きい (佐倉, 2019)。

これら諸点への留保なしに研究業績の定量的

### 3.2 ライデン声明

科学計量学者たちはこれらの問題点を以前から認識して研究者コミュニティに警告を發しつつ、定量的指標の適切な使用法について検討を重ねてきた (麻生 [2021] がそのうちのいくつかの動きをまとめている)。ある意味で、それらの集大成として發表されたのが「ライデン声

指標を個人あるいは個々の研究組織に無定見に適用することは、弊害が大きい。研究者は研究成果の評価をどの雑誌に掲載されたかで測ようになり、IF の高い雑誌に掲載された論文は良い論文、そのような論文を量産している研究者は良い研究者、という評価尺度が定着してしまっている (Lawrence, 2003, 2007)。Caves (2014) は、この傾向を「高インパクトファクター症候群 (High Impact Factor Syndrome: HIFS)」と名付けて批判している。HIFS が昂じると、なんとしても高 IF の雑誌に論文を載せなければというプレッシャーに負けて、研究不正をおかす温床になっているという指摘もある (麻生, 2021)

これらの弊害に対抗するべく、すでにヨーロッパの非英語圏諸国では、自国での書籍を研究業績としてどのように評価するか、さまざまな試行がなされている (Giménez-Toledo et al., 2016)。標葉 (2017) はこれらの動向を踏まえ、日本の人文・社会系分野でも自身の研究分野と社会との関係を積極的に構築しなおし、社会における生産的相互作用への貢献なども自分たちの業績として積極的に発信していくべきだと主張している。

明」である (Hicks et al., 2015)。これは、2014年にライデン大学で開催された第19回科学技術指標国際会議でのジョージア工科大学の Diana Hicks の講演をベースに、ライデン大学科学技術センターの Paul Wouster らが加わって論点を10の原則に整理し、社会的影響力の



大きい（IF が極めて高い！）*Nature* に発表したものだ。その経緯の詳細については小野寺・伊神（2016）による解説やライデン声明のホームページ<sup>(2)</sup>を参照いただきたい。

このライデン声明 10 原則は定量的指標の誤用を厳しく戒め、それに代わってどのような評価の方法が良いのかをまとめているが、ここではそのうち、学際的な研究教育においてとくに重要と私が思うもの 4 つについて、紹介しつつ検討する。選定の基準は、主として学生および若手の研究者が、自身の研究業績としてどのようなものを生産していったら良いのかを考える際の目安になりうるものである。組織やプロジェクト全体の評価を念頭においたものではない。以下のライデン原則の日本語訳は小野寺・伊神（2016）による訳文（ライデン声明の HP にある日本語版と同一）に拠っているが、一部、Hicks et al. (2015) を参照して佐倉が修正したところがある。解説部分は佐倉のオリジナルである。参考までに、10 原則のすべてを付録として文末に掲げた。

### 原則 1 定量的評価は専門家による定性的評定の補助に用いるべきである

業績や研究者個人の能力の評価は、あくまでも定性的（質的）であるべきだというのがライデン声明の根底にある価値観だ。アメリカの科学史家 Theodore Porter (1995) は、専門家同士の評価は元来は定性的なものであり、背景知識や経緯を共有していない専門外の人たち——スポンサーや権力者やメディアなど——に業績を「わかりやすく、かつ客観的に」示す必要が生じたときに定量的な評価尺度が要求されてき

たことを明らかにしている。それが専門家集団「内」にも跳ね返ってきたのが昨今の状況なのである。定量的評価は質的評価を洗練させたものではなく、むしろ逆なのだ。本稿第 1 節で研究の心構えについての卓見を引用した白上謙一は、こうも言っている。「質的な結果をめざしてのみ、量的な研究を遂行せねばならぬ」(白上, 1972: p. 86)。これは研究の評価ではなくその目的を説いた発言だが、ライデン声明とほぼ同様の主張であり、白上の慧眼には改めて驚かされる。

### 原則 2 機関、グループ、研究者の目標に照らして業績を測定すること

これは少しわかりにくいかもしれないが、狭義の学術的成果（原著論文）を算出するだけが科学者の役割ではなく、社会的な課題の解決を目標として研究成果を社会実装していくことも科学者にとって重要な社会的使命であり、SCI/IF などの定量的指標だけではそれらの活動を正當に評価できないというのが趣旨である。

### 原則 3 優れた地域的研究を守ること

学術分野によっては英語以外の言語で書いた成果は無視されることも珍しくない。しかし、とくに人文社会系の分野では国や地域の問題に取り組むことも重要な使命であり、成果をその地域の言語で書いた方がより効果的なことも珍しくない。自然科学や医学・工学分野でも、疫学や公衆衛生、防災、自然保護など、同様の特性をもつ分野は少なくない。Hicks et al. (2015) は、スペインの法律学者たちが IF の高い英語雑誌への論文掲載を目指して抽象的なモデルの

研究やアメリカのデータを使った研究を優先しており、地域の労働法や移民問題といった重要な社会課題が軽視されている状況を指摘している。日本の現状は幸いにしてまだそこまでではないと思われるが、逆に言えば「まだ傷の浅い今のうち」に手を打っておくことが必要だろう。

#### 原則 6 分野により発表と引用の実状は異なることに留意せよ

人文社会系、とくに人文系では、研究成果を単行書として刊行することが長らく重要視されてきた (Larivière et al., 2006)。これらの分野では単行書の評価は学術専門誌に掲載された論文より劣るものではない。研究成果をどのような形で発表するのが適切かは、学問分野の特性によるところが大きい。雑誌論文のような「細切れ」の内容では、思想や歴史についての深い考察には不足しがちであるというのが、単行書のボリュームを必要とする理由である<sup>(3)</sup>。

一方、コンピューター科学や情報科学では国際会議の発表論集 (プロシーディングス) が、場合によっては学術専門誌 (ジャーナル) より高く評価されることもある。権威のある国際会議では発表の採択率が平均的な学術誌より低いこともしばしばあり、技術の変化が早いこともあって、「雑誌になるのを待ってられない」といった事情が作用しているようだ。

このようにライデン声明は定量的評価の功罪

を踏まえた上で、実用的な多面的評価の基準を明確に整理してくれた。これらから明らかなように、学術研究が社会に資するためのものであり続けるならば、その多面的評価は必要不可欠だ。

先に述べたように、IF などの現在主流となっている定量的尺度が広く使われるようになったのは 20 世紀後半、とくに最後の四半世紀以後のことである。まだ数十年しか経っていないのであり、尺度自体もその使い方も、成熟したものにはなっていない。金科玉条のように崇め奉るのではなく、臨機応変に対処しなければならない状況にある。

研究業績の評価指標だけでなく、知識の生産とその品質管理システム自体も歴史的な産物であり、現在のような、学術専門誌を中心とした形式になったのはさほど昔のことではない (詳細は佐倉 [2020] を参照)。19 世紀半ばから後半に活躍した Charles Darwin は原著論文も書いているが、おもな業績は著書 (モノグラフ) で発表している。もっとも影響力が大きく今に至るも頻繁に参照引用されている『種の起源』(Darwin, 1859) はその典型である。

ダーウィンの時代に戻れということではない。戻る必要もないし戻るべきでもない。だが、今の時代で標準的とされている「研究業績」だけが、いつでもどこでも研究業績として認められてきたのではないことは、認識しておくべきだと思う。現代の基準の強みと弱みを、相対化して認識しておこう。

#### 4. 補足的論点——共著者数の問題

学際的な研究教育組織では、前節（および付録）で紹介したようなライデン声明が問題視していることがらに、日々直面していると言ってもよい。それだけに、ライデン声明の重要性がよく実感できるところでもある。

一点、原則6で少し触れられている、共著者数の問題について補足しておきたい。分野による慣行の違いが大きい現象だからだ。すでに述べたように、IFは共著者数が増えると高くなる傾向にある（Amin & Mabe, 2000）。共著者数は人文系では少なく、自然科学系では多い。2000年のデータでは自然科学系学術論文の90%が複数著者で書かれているのに対し、社会学系は約47%、人文系は10%のみが複数著者の論文である（Larivière et al., 2006）<sup>(4)</sup>。

だが興味深いことに、自然科学系でも単行書になると共著者数は少なくなる。ここで言う「単行書」とは、書籍の一部分の分担執筆ではなく、

一冊すべてを同じ著者が執筆する場合のことを意味している。Google Scholarで同じ分野の論文と書籍とで共著者数を比較してみると（2021年2月4日検索）、たとえば"neuroscience"をキーワードに被引用数の多い順に10ページ目までに出てきた文献を対象にした場合、論文（N=76）では共著者数の平均値が2.3人（SD=1.44）、中央値が2人なのに対し、教科書を除いた単行書（N=19）では平均値が1.6人（SD=0.90）、中央値1人であった。平均値の差は統計的に有意である（ $t=2.53$ ,  $p=0.007$ ）[表1]。より専門的な単語をキーワードに使うと、論文の著者数はさらに大きくなる。たとえば"hypothalamus"（＝視床下部）で検索してリストアップされた論文（N=50）の共著者数は平均値が4.5人（SD=4.99）、中央値3人である[表1]。この条件では単行書は検索結果に上がってこなかった。

表1 論文と単行書の著者数の違い  
(Tab 1 Comparison of the number of co-authors between articles and books)

Keyword	"neuroscience"		"hypothalamus"	
	Article	Book <sup>1)</sup>	Article	Book
N	76	19	50	0
<i>No. of co-authors</i>				
Mean	2.3*	1.6*	4.5	--
SD	1.44	0.90	4.99	--
Median	2	1	3	--

Searched by Google Scholar, on 4 February, 2021.

1) Textbook not included, \* Significantly different ( $p<0.01$ )

単行書でも教科書の類は共著者数が多く、論文でも総説は共著者数が少ない傾向にある。こ

の違いは、研究成果における「作品性」の重要度によると考えられる。作品性というのは曖昧

な言葉だが、とりあえず漠然と、研究成果が著者個人の理念や考えかたなどに依存する度合いと考えておこう。人文系の哲学や文学研究ではこの度合いは高いが、自然科学系の実験報告論文では低い。「著者性」と言ってもいいかもしれない。これを Roland Barthes や Michel Foucault の作者論と結びつけて論じる能力は私にはないが、興味深い論点ではないかと思う。ちなみに自然科学系でも数学は単著論文が多い。

作品性が人文系の書籍より重視されるのは芸術の分野である。小説や戯曲といった文学作品はほとんどが単独著者によるものだし、絵画や彫刻、音楽作品なども作者はひとりのことが多い。興味深いことに芸術やデザインの領域では、複数の人間の共同作業でもしばしばリーダー単独の「作品」として認知される傾向にある。建築作品が典型的だが、「製作者」として名前がクレジットされるのは、通常は設計者だけである。ファッション・デザインでもパターナーやアシスタントの名前は知られず、Pierre

## 5. おわりに

研究成果の評価というのは、研究者と社会との対話なのだと思う。どちらか片方だけの主張や言い分が通るものではない。研究者が、誰が何と言おうとこれが私の成果だ！といくら力んでも、世間から相手にされないのであればどうしようもない。一方で、評価する側の都合だけが優先されて、研究者の自発性や自主性がまったく顧みられなくなるとは、研究者のモラルが下がってしまって良い成果はかえって出てこ

Cardin や Giorgio Armani といったトップデザイナーの名前だけがブランドとして掲げられる。

作品性が重視される分野ではその制作に多くの人間が関わっていても、最終産物は強烈な個性を持った単独の「作者」に帰属する。一方で同じ共同作業の成果であっても著者性よりも客観的な結果が重視される自然科学分野では連名で論文が発表される。これは、「作品」に相当するのは自然科学分野でも著書であると考えれば辻褄が合う。考えてみれば高名な科学者の表わした著書も、多くの共同研究者たちとの研究の成果である。それらの成果はすでに公表された論文として明示的に引用され、学術界のルールでは知識の第一生産者と所有権は明確にクレジットされる。この点が芸術界とは異なるとはいえ、著書自体はあくまでも単独著者のものとして認知され、流通していく。逆に、芸術やデザインの世界では、科学論文の文献引用に相当する現象は何なのかといったことを考えるのも、興味深いかもしれない。

なくなるだろう。

芸術家の中には、社会からの評価に敢然と立ち向かい、みずからの信念にひたすら従うことで革命的な成果を残した人もいる。オーストリア生まれの作曲家、Arnold Schönberg もそのひとりである [図2]。彼は12音音楽を創始して西洋古典音楽の基盤であった調性音楽を解体した、音楽史上の革命家だ。



図2 Arnold Schönbergの自画像(1908年)(Wikimedia)  
Fig 2 Self portrait by Arnold Schönberg (1908)

しかし、彼の作曲した音楽は当時の批評家や聴衆からはなかなか受け入れられず、批判——というより誹謗中傷があとを絶たなかった。ときには野次や嘲笑が激しく、演奏会自体が大混乱したこともあったという(Griffiths, 1978, pp.35-36)。そこで彼は、一般公開せずに批評家を閉め出して会員限定の演奏会をおこなうことにする。1918年から21年まで続いた私的演奏協会(Verein für musikalische Privataufführung)である(Rosen, 1975; 石田, 2012)。「この協会が設立されたのは、同時代の音楽を本来の持主である音楽家に取り戻し、大衆的な音楽マーケットの墮落した影響から切り離すためだった」(Rosen, 1975, p. 66, 日本語版を参照しつつ佐倉が原文から訳した)。この方法によって練習時間を十分確保し、理解のある聴衆だけのために、自分たちの作品を含めた同時代の音楽を披露することができたのである。

もちろん、批評家らを締め出した行為についての評価は賛否両面ある。だが、このような形

態の私的演奏協会を行なったからこそ、Schönbergたちは同時代の他の作曲家たちの音楽への理解を深め、彼ら自身の音楽様式を掘り下げ、確固たるものにすることができたのだった。また、同時代の音楽家同士のネットワーク形成の場としても機能した(石田, 2012)。当時の多くの批評家や聴衆からは拒絶されていたSchönbergたちの音楽が、その後の20世紀後半の音楽シーンに大きな影響を及ぼしたことはまぎれもない歴史上の事実である(Salzman, 1974; Griffiths, 1978)。他者からの評価尺度ではなく自らの理想を確固として追求し続けたからこそ、Schönbergとその弟子や友人たちは、革命的な業績を上げることができたのだと言ってよい<sup>(5)</sup>。

もちろんこれは極端な例で、誰もが真似できるものではない。だが、自分の成果を評価してくれる聴衆(読者)は、人数の多寡はともあれ、必ずどこかにいるものだ。今の時代の主流となっている評価尺度に合わせることは必要では



あるが、その尺度に合った生産物だけが業績となるべきものではないことは、決して忘れないでほしい。

この間（はざま）にあって、しかし研究業績——世間からそのように認められるもの——を出さなければ、とくに若い世代の研究者は生き残れないし、大学院生は修論や博論が書けなければ卒業できない。企業に就職したらそこで認められる業績を出さないことには評価されない。ときに自分の意に沿わない形の業績を強いられることもあるかもしれないが、そもそも業績とは大なり小なりそのようなものであると考えることも必要だろう。一方で、これ以上は絶対妥協できないという一線を越えるべきでもない。あくまでも自分と社会との対話なのである。

自分の業績の優れた点は、少なくとも自分には理解できるはずである。まずは自分の中の評価軸を確固たるものにすること。これは大学院で身につけるべき事柄だ。そして、その自分の

評価軸が社会（自分の所属している組織やコミュニティ）の評価軸とずれているときに、どのように調整するか。どこまでなら妥協できて、どこからは許されないか、その境目を見極める力は、生涯かけて少しずつ学んでいくしかないように思う。分野によってその境目の位置は異なるし、経験値を高めていくしかないのではないか。

最後にもう一度、白上謙一（1972）の別の言葉を引用する。「たえず自分の見付けた問題を持ち、どこまでも自分流におしすすめて行かねばならぬ」（p.84）。これも勇ましく、美しい檄文である。だが、どこまでも自分流におしすすめて行ける人はたしかに素晴らしいが、一方で、ときに退くことも必要な局面もあるはずだと思う。勇気ある撤退は、長い目で見れば必ずや大きな実りをもたらすはずだ。いつ、どこで、どのように撤退すべきなのか。その判断もまた、自分と社会との対話によって形づくられるはずである。

## 謝辞

隠岐さや香教授（名古屋大学）から草稿への有益なコメントと参考文献の御教示をいただいた。本稿の議論の一部は JSPS 科研費 19H01228, 19K21604 の助成の成果である。

## 註

- (1) 国立大学法人東京大学における研究活動上の不正行為の防止に関する規則 [https://www.u-tokyo.ac.jp/gen01/reiki\\_int/reiki\\_honbun/au07410491.html](https://www.u-tokyo.ac.jp/gen01/reiki_int/reiki_honbun/au07410491.html) [2021年2月4日確認]
- (2) ライデン声明のホームページには Hicks et al. (2015) による宣言の日本語訳を含む 23カ国語版と動画解説などがある (<http://www.leidenmanifesto.org> [2021年2月4日確認])。
- (3) インターネット上で学術情報の多くが流通している現在、紙の成書という媒体（およびそれを標準とする知識の生態系）は「賞味期限切れ」という見解は、別途成り立ちうる。私もその見解に賛成するところもあるが、しかしそのことと、「現状の評価基準がどのようなものであるか」は別の問題である。新しい情報メディアを使って活発に活動できる人はどんどんすばい。一方で、現行の評価基準に沿った形で研究業績を積み重ねなければならない立場の人も大勢いる。
- (4) ちなみに、2021年1月現在、科学論文の共著者数の最高記録は5,154人で、2015年に発表された高エネルギー物理学分野のもののだそう（<https://current.ndl.go.jp/node/28498> [2021年2月4日確認]）。

- <sup>5)</sup> この私的演奏協会を Schönberg 自身は、自分たちの作品発表の場としてではなく、教育活動の一環として位置づけていた (Rosen, 1975, p. 65)。彼は、教師はみずからのコピーとして弟子を育ててはいけないという独自の教育哲学をもった、教育熱心な音楽教師でもあり (Schoenberg, 1950=2019/1973)、創作活動と教育活動は不可分のものと認識していた (上野, 2005)。革命的な創作活動と教育とのこのような関係は、学術研究と大学院教育の関係にとっても示唆があると思われる。

## 参考文献

- Amin, M. and Mabe, M. (2000) Impact factors: use and abuse. *Perspectives in Publishing*, No.1, 1-6.
- 麻生一枝 (2021) 『科学者をまどわす魔法の数字, インパクト・ファクターの正体——誤用の悪影響と賢い使い方を考える』日本評論社
- Caves, C.M. (2014) High-impact-factor syndrome. *APS News*, **23**(10), November, The Back Page. <https://www.aps.org/publications/apsnews/201411/backpage.cfm> [2021年2月4日確認]
- Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.
- 藤垣祐子・平川秀幸・富澤宏之・調麻佐志・林隆之・牧野淳一郎 (2004) 『研究評価・科学論のための科学計量学入門』丸善
- Fuyuno, I. and Cryanoski, D. (2006) Cash for papers: putting a premium on publication. *Nature*, **441**, 792. DOI: 10.1038/441792b
- Garfield, E. (1955) Citation indexes for science. *Science*, **122**(3159), 108-111. DOI: 10.1126/science.122.3159.108.
- Garfield, E. (1970) Citation indexing for studying science. *Nature*, **227**, 669-671. DOI: <https://doi.org/10.1038/227669a0>
- Garfield, E. (1972) Analysis as a tool in journal evaluation. *Science*, **178**(4060), 471-479. DOI: <https://www.jstor.org/stable/1735096>
- Garfield, E. (2006) The history and meaning of the journal impact factor. *JAMA: Journal of the American Medical Association*, **295**(1), 90-93. DOI: 10.1001/jama.295.1.90
- Giménez-Toledo, E., Mañana-Rodríguez, J., Engels, T. C. E., Ingwersen, P., Pölonen, J., Sivertsen, G., Verleysen, F. T. and Zuccala, A. A. (2016) Taking scholarly books into account: current developments in five European countries. *Scientometrics*, **107**, 685-699. DOI: 10.1007/s11192-016-1886-5
- Greenwood, D. C. (2007) Reliability of journal impact factor rankings. *BMC Medical Research Methodology*, **7**, 48-53. DOI: 10.1186/1471-2288-7-48
- Griffiths, P. (1978) *A Concise History of Modern Music: From Debussy to Boulez*. Thames & Hudson [石田一志訳 (1984) 『現代音楽小史——ドビュッシーからブーレーズまで——』音楽之友社]
- Hicks, D. and Wouters, P., Waltman, de Rijcke, S. and Rafols, I. (2015) The Leiden Manifesto for research metrics. *Nature*, **250**, 429-431. DOI: 10.1038/520429a
- Hirsch, J. E. (2005) An index to quantify an individual's scientific research output. *Proceedings of National Academy of Sciences of the United States of America*, **102**(46), 16569-16572. DOI: <https://doi.org/10.1073/pnas.0507655102>
- 石田一志 (2012) 『シェーンベルクの旅路』春秋社
- Larivière, V., Gingras, Y. and Archambault, É. (2006) Canadian collaboration networks: a comparative analysis of the natural sciences, social sciences and the humanities. *Scientometrics* **68**, 519-533. DOI: <https://doi.org/10.1007/s11192-006-0127-8>
- Lawrence, P.A. (2003) The politics of publication. *Nature*, **422**, 259-261. DOI: <https://doi.org/10.1038/422259a>
- Lawrence, P.A. (2007) The mismeasurement of science. *Current Biology*, **17**(15), R583-R585. DOI: 10.1016/j.cub.2007.06.014
- Leiden Manifesto for Research Metrics. <http://www.leidenmanifesto.org> [2021年2月4日確認]
- 牧野宏美 (2020) 「ファクトチェック：ツイッターで拡散「任命拒否6人、ツールで低評価だから学者と言えない」は誤り」『毎日新聞 (デジタル版)』2020年10月10日、<https://mainichi.jp/articles/20201010/k00/00m/040/006000c> [2021年2月4日確認]
- Marson, S. M. (2020) Editorial: Is the Impact Factor (IF) ethical to use for promotion and tenure decisions? *Journal of Social Work Values and Ethics*, **17**(1), 2-5.
- 名月論 (2020) 「学術会議任命拒否問題について物申す」『Blogs』2020年10月9日、<https://blogs.com/article/490064/?p=2> [2021年2月4日確認]
- 奥永さつき (2020) 「学術会議任命拒否問題について」[https://blog.goo.ne.jp/prajnaparamita\\_2/e/c4b35298936a024b935349494e483aa4](https://blog.goo.ne.jp/prajnaparamita_2/e/c4b35298936a024b935349494e483aa4) [2021年2月4日確認]
- 小野寺夏生・伊神正貴 (2016) 「研究計量に関するライデン声明について」『STI Horizon』**2**(4), 35-39. DOI: <http://doi.org/10.15108/>

stih.00050

- Porter, T. M. (1995) *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press [藤垣祐子訳 (2013) 『数値と客観性——科学と社会における信頼の獲得——』みすず書房]
- Rosen, C. (1975) *Arnold Schoenberg*. Modern Masters (Series Editor: F. Kermode), New York: Viking Press [武田明倫訳 (1984) 『シェーンベルク』岩波現代選書、岩波書店]
- 佐倉統 (2011) 「梅棹忠夫と 3.11 —— 私たちは科学技術とどう向き合っていくのか——」『中央公論』126(8): 24-41.
- 佐倉統 (2019) 「日本語総説誌の教育的役割は今でも大きいはずだ」『生物科学』70(2), 65.
- 佐倉統 (2020) 『科学とはなにか』ブルーバックス, 講談社
- Salzman, E. (1974) *Twentieth-Century Music: An Introduction*. Englewood Cliffs, NJ: Prentice-Hall [松前紀男・秋岡陽訳 (1993) 『20世紀の音楽』音楽史シリーズ6, 東海大学出版会]
- Schoenberg, A. (1950) The blessing of the dressing. In: *Style and Idea: Selected Writings of Arnold Schoenberg*. New York: Philosophical Library [上田昭訳 (2019) 『芸術の創造と大衆性』所収: 『シェーンベルク音楽論選 様式と思想』ちくま学芸文庫、筑摩書房; 初出: 上田昭訳 (1973) 『音楽の様式と思想』三一書房]
- Seglen, P. O. (1994) Causal relationship between article citedness and journal impact. *Journal of the American Society for Information Science*, 45(1), 1-11. DOI: 10.1002/ (SICI) 1097-4571 (199401) 45:1<1::AID-ASIJ>3.0.CO;2-Y
- 標葉隆馬 (2017) 「人文・社会科学を巡る研究評価の現在と課題」『年報 科学・技術・社会』26, 1-39.
- 白上謙一 (1972) 『生物学と方法』河出書房
- 上野大輔 (2005) 「A. シェーンベルクの教育活動と音楽思想」『東京情報大学研究論集』8(2), 25-35.
- 梅棹忠夫 (1977) 「国立民族学博物館における研究のあり方について」『国立民族学博物館研究報告』1(4), 896-938 [再録: 梅棹忠夫著・小山修三編 (1993) 『研究と経営』〈梅棹忠夫著作集第22巻〉中央公論社, ]
- 梅棹忠夫 (1989) 「研究業績の評価について」『民博通信』46, 10-20 [再録: 梅棹忠夫 (1990) 『情報管理論』岩波書店, ; 梅棹忠夫著・小山修三編 (1993) 『研究と経営』〈梅棹忠夫著作集第22巻〉中央公論社, ]

付録 ライデン声明 10 原則 (Hicks et al., 2015; 小野寺・伊神 [2016] の日本語訳を一部改訳)

- 原則 1 定量的評価は専門家による定性的評定の補助に用いるべきである
- 原則 2 機関、グループ、研究者の研究目的に照らして業績を測定すること
- 原則 3 優れた地域的研究を守ること
- 原則 4 データの収集と分析のプロセスをオープン、透明、単純に保つこと
- 原則 5 被評価者がデータと分析過程を確認できるようにすべきである
- 原則 6 分野により発表と引用の実状は異なることに留意せよ
- 原則 7 個々の研究者の評定は、そのポートフォリオの定性的判定を踏まえること
- 原則 8 不適切な特定や誤った精度を避けよ
- 原則 9 評定と指標がシステム全体に及ぼす効果を認識すること
- 原則 10 指標を定期的に吟味し改善せよ



佐倉 統 (さくら・おさむ)

[専門] 科学技術社会論  
[主たる著書・論文]  
『科学とはなにか』(講談社ブルーバックス)  
『「便利」は人を不幸にする』(新潮選書)  
『人と「機械」をつなぐデザイン』(編著、東京大学出版会)  
[所属] 東京大学大学院情報学環  
[所属学会] 科学技術社会論学会、日本行動進化学会、神経科学学会、人工知能学会

# How is (or Should) the Performance of Academic Researchers (be) Evaluated?

Osamu Sakura\*

The philosophy and methods of evaluating outcomes of research activities differ widely across the academic fields. Formerly, these evaluation frameworks were based on each research field scenario, but recently, quantitative measures have become popular and are being applied across various disciplines. The frequently applied measures are, for example, Science Citation Index for articles, Impact Factor for academic journals, and h-index for the individual research. Although these quantitative indices are helpful to visualize the dynamics of the production process of scientific knowledge, they are inappropriate measures for evaluating the “excellence” of individual researchers and academic journals because the variances are too large. In addition, the citation data mainly cover articles published in academic journals; thus, evaluating the achievements in the form of books, such as textbooks and monographs, is difficult. Therefore, applying these indices is inappropriate for researchers who put priorities in publishing studies on specific fields such as humanities. Nevertheless, conspicuous misuse of these quantitative measures has been increasing, and they have been applied as criteria in personnel evaluations, such as hiring and promotion, especially in recent years.

The Leiden Manifesto, published in *Nature* in 2015, summarizes the problems with quantitative measures and proposes 10 principles on how research achievements should be evaluated. It states that the qualitative approach should take the main role in the evaluation, whereas the quantitative measures should be used as supplementary. Moreover, the evaluation should consider the characteristics of each field. This present study introduces four principles among the 10, which are particularly important for interdisciplinary research and education organizations. Furthermore, it reconfirms the necessity of multidimensional measures in the evaluation of interdisciplinary achievements.

The evaluation of research performance is a dialog between academic and public society. It is not only a reflection of the academic researchers’ desire or the requests from the public. A research evaluation is a communication between different values of both parties, and therefore, difficulties to

---

\* Interfaculty Initiative in Information Studies, the University of Tokyo; RIKEN Center for Advanced Intelligence Research Project (AIP)  
Key Words : Evaluation of Research Products, Quantitative Evaluation, Qualitative Evaluation, Leiden Manifesto, Interdisciplinarity.

a certain degree are inevitable. Those who are in the midst of this evaluation process, especially the younger academicians, should not accept only one of the assessment measures, but rather seek a compromise between the two to overcome the underlying issue.