



2020年8月29日

日本学術会議 学術フォーラム

「学術振興に寄与する研究評価を目指して」

研究評価をめぐる今後と課題 ー日本と海外の状況からの示唆

標葉 隆馬 (Ryuma Shineha Ph.D)

大阪大学 社会技術共創研究センター准教授



大阪大学 社会技術共創研究センター
Research Center on Ethical, Legal and Social Issues

Osaka University
Research Center on
Ethical, Legal and
Social Issues

本日の内容

- 日本の研究評価の変化と課題
- 海外における議論事例から
 - 英国におけるREFとインパクト評価
 - 国際的な議論の例として
 - ーライデン・マニフェスト
 - ー人文・社会科学評価と今後に関わる一般的課題
 - ー「生産的相互作用」という視点
- まとめ／今後の課題

ファンディング構造と評価システムの変化



一般大学資金：経常的・機関単位
General University Funds (GUF)

質の管理
(Quality Control)

科学技術基本計画：
競争型、PJ型・PG型、重点領域の設定

説明責任の増大

中間組織／境界組織の役割増大

質のモニタリング
(Quality Monitoring)

直接政府資金：PJ型・領域・政策

Direct Government Funds (DGF)

マルチファンディングシステムの進行

評価の対象・タイミングに応じた評価が必要となる



	事前評価	中間評価	事後評価	追跡評価
アウトプット	成果物（論文、著作、発表など）			
アウトカム	アウトプットの結果としてもたらされる短中期的な効果 （特許の数、臨床試験の実施、新薬創出、etc）			
インパクト	長期的な時間軸で生じる幅広い効果 （アウトカムを超える社会・経済的影響、政策の追加効果）			
（スピルオーバー）	研究開発がもたらす他の領域・業種等への波及効果			



日本の研究評価の変化と課題

1990年代後半～：現在の日本の 科学技術政策・制度のベースとなるもの



科学技術基本法

(科学技術・イノベーション基本法へ)

科学技術基本計画

→研究開発推進の方針、研究
環境の整備

科学技術・イノベーション
創出の活性化に関する法律

科 学 技 術 基 本 計 画

平成23年8月19日

閣 議 決 定

日本の科学技術政策の変化

科学技術基本計画

1996年～5年毎

2001年第2期：重点・推進領域の展開

2006年第3期：科学技術イノベーション

2007：イノベーション25...

第4期：2011～2015

第5期：2016～2021

2018年12月

「科学技術・イノベーション創出の活性化に関する法律」（旧：研究開発力強化法）

2020年～科学技術基本改正（科学技術・イノベーション基本法）

・・・人社除外事項の撤廃

1980年代：基礎研究重視（日米科学技術協定が背景）
2000年～：PJ・PG型のファンディング（選択と集中）、「イノベーション」へのシフト。ただし、あくまで科学技術・研究開発に関するイノベーション（範囲が狭い）

2013～科学技術イノベーション総合戦略

2018～統合イノベーション戦略

- ・・・財政諮問会議の影響
- ・・・未来投資戦略などの影響、etc

産業競争力会議の成長戦略の下位戦略の位置付け
（狭い）イノベーション色強化

⇒技術革新とそれによる経済的価値や産業競争力の創出という限定的な意味合いが強い

科学技術基本法の改正によって起きること

・ 科学技術・イノベーション基本法と科学技術・イノベーション基本計画 ・ ・ ・

➤ 「人文科学のみ」の除外規定の撤廃

- ✓ 人文・社会科学系もまた「科学技術・イノベーション創出の活性化に関する法律」（旧：研究開発力強化法）の範疇になる
- ✓ イノベーション志向、PJ型・PG型の強化、経済・社会インパクトへの視点は必然的に強化されることが予想される

➤ 第6期科学技術基本計画の議論（人文・社会科学系諸分野は意識できているか？）

➤ 第7期科学技術基本計画（科学技術・イノベーション基本計画）に向けた議論は？

・ 改正案においても「イノベーション」の意味射程は狭い

- ✓ 科学的な発見または発明、新商品又は新役務の開発その他の創造的活動を通じて新たな価値を生み出し、これを普及することにより経済社会の大きな変化を創出すること
- ✓ 研究開発の成果の実用化によるイノベーションの創出、国民への還元、学際融合、社会的課題対応（少子高齢化、人口減少。国境を越えた社会経済活動、食料問題、エネルギー利用の制約、地球温暖化など）

・ 本法律改正に伴う付帯決議文書

- ✓ イノベーション創出への偏重しないことへの留意
- ✓ 対象範囲が増えることを踏まえた予算の拡充努力
- ✓ 研究者の雇用の安定化、ポスドク・若手研究者の自立と活躍機会・環境の整備努力 ・ ・ ・ など

⇒国立大学法人化の時の付帯決議がほぼ履行されなかったが ・ ・ ・ ？

国の研究開発評価に関する大綱的指針

- ・ 2016年までに計6回改訂
- ・ 研究開発評価の主だった基本的意義
 - 「研究費の効果的・効率的配分」
 - 「研究開発活動の振興・促進」
 - 「説明責任」
- ・ 評価制度とファンディングシステムの変化
 - ⇒研究成果の社会・経済的側面の貢献とそのための評価システムの希求が顕在化し、いくつかの意義・目的が追加
 - 「より効率的で迅速な研究開発成果の国民・社会への還元」
 - 「有望な研究開発課題の発掘と持続」
 - 「政策・施策形成への貢献」

国の研究開発評価に関する大綱的指針

平成24年12月6日

内閣総理大臣決定

2016年12月21日

「国の研究開発評価に関する大綱的指針」



大阪大学 社会技術共創研究センター
Research Center on Ethical, Legal and Social Issues

＜強調されるポイント＞

- ・ 評価疲れ問題の解消

（メタ評価の重要性）

- ・ 経済的・社会的インパクトの評価の
記述充実

（世界的に見てインパクトの議論自体
は進みつつあるが・・・）

- ・ 実効的なプログラム評価推進

⇒ファンディング・エージェンシーの
評価・責任の明確化につなげていかな
いのなら意味は減退する

いずれにせよ研究者の側からも意識し
たモニタリングが必要である

国の研究開発評価に関する大綱的指針

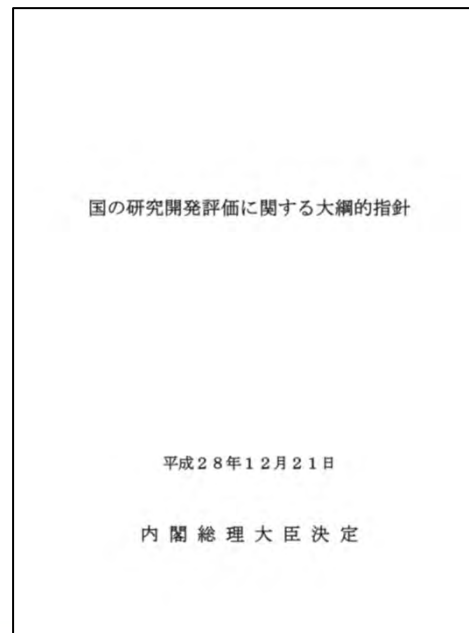
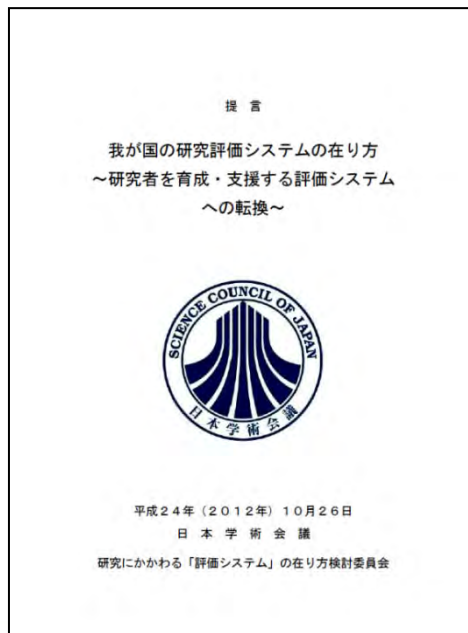
平成28年12月21日

内閣総理大臣決定

日本における研究評価制度の課題

- DGFの割合増加に平行して・・・
- ⇒ 中間組織・境界組織の役割拡大
- ⇒ プロジェクト／プログラム型の増加
- ⇒ 政策レベル
- ⇒ 「評価の評価（メタ評価）」

このレベルでの
評価の重要性



今後の動向への注目・・・

- インパクトやイノベーションの意味射程の狭さ
- 「評価疲れ」問題
- 個人の研究者に皺寄せが行く構造の是正
- 政策・ファンディング構造の変化



研究評価の中で今後重要になる事柄： 「インパクト評価」について

（2016年12月21日「国の研究開発評価に関する大綱的指針」でも言及げ増えたポイントについて、最近の国内外の議論から）



英国における展開－REF・インパクト評価に注目して

英国にみる研究評価システムの変化



rae2008
Research Assessment Exercise

- Peer-Reviewによる高等教育／研究機関評価
- RAEの評価結果に基づき、大学への資金配分変更

英国財務省（2006）

“Science and innovation investment framework 2004-2014: next steps”

- RAEの失敗を指摘（大学の多様な研究の把握が出来ていない）
- 「評価疲れ」問題の進行・顕在化

ビブリオメトリクス指標の重点利用なども議論されるものの・・・

- Research Excellence Framework（2013～）
- RAE同様にPeer-Reviewの維持
- 研究の質や研究環境評価に加えて、新たな「インパクト」を評価基準に設定
- 大学は、分野ごとにインパクトの説明を提出する

REF2014
Research Excellence Framework

REF2014

2011年～12年：準備

2012～13年：投稿（締切2013年11月29日）

2014年：アセスメント→評価発行

評価基準

- ・ **アウトプットの質（65%）**：
“originality, significance and rigor”
- ・ **研究環境（15%）**：
“vitality and sustainability”
- ・ **社会的・経済的・文化的インパクト（20%）**：“reach and significance”

1993年～2013年に行われた研究で、2008年1月1日～2013年7月31日にインパクトが生じたものを説明する

Outputs (65 per cent)	
Definition for the REF	'Outputs' are the product of any form of research, published between January 2008 and December 2013. They include publications such as journal articles, monographs and chapters in books, as well as outputs disseminated in other ways such as designs, performances and exhibitions.
Information provided in submissions	Institutions submitted up to four outputs for each member of staff they selected for inclusion in their submissions. Submitters could request that an output of extended scale and scope be 'double-weighted' by the sub-panel (that is, counted as two outputs in the assessment).
Assessment criteria	The panels assessed the quality of outputs against the criteria of 'originality, significance and rigor'. The assessment was based on peer review of the outputs. Some sub-panels considered the number of times the output had been cited, as contextual information to support peer review.

Impact (20 per cent)	
Definition for the REF	'Impact' is any effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia.
Information provided in submissions	Each submission included: Impact case studies. These four-page documents described impacts that had occurred between January 2008 and July 2013. The submitting institution must have produced high quality research since 1993 that contributed to the impacts. Each submission included one case study, plus an additional case study for every 10 FTE staff. An impact template. This document explained how the submitted unit had enabled impact from its research during the period from 2008 to 2013, and its future strategy for impact.
Assessment criteria	Impact case studies were assessed in terms of the 'reach and significance' of the impacts. Impact templates were assessed in terms of how far the approach and strategy are conducive to achieving impacts.

Environment (15 per cent)	
Definition for the REF	'Environment' refers to the strategy, resources and infrastructure that support research activity in the submitted unit and contribute more widely to the discipline.
Information provided in submissions	Each submission included: An environment template. This document describes the submitted unit's research strategy; its support for research staff and students; its research income, infrastructure and facilities; and its research collaborations and wider contributions to the discipline. Statistical data. Institutions provided data on the amount of research income they received each academic year (from 2008-09 to 2012-13) from different types of sources, and on the number of research doctoral degrees awarded in each of these years. These were based on data that institutions report annually to the Higher Education Statistics Agency.
Assessment criteria	The research environment was assessed in terms of its 'vitality and sustainability'.



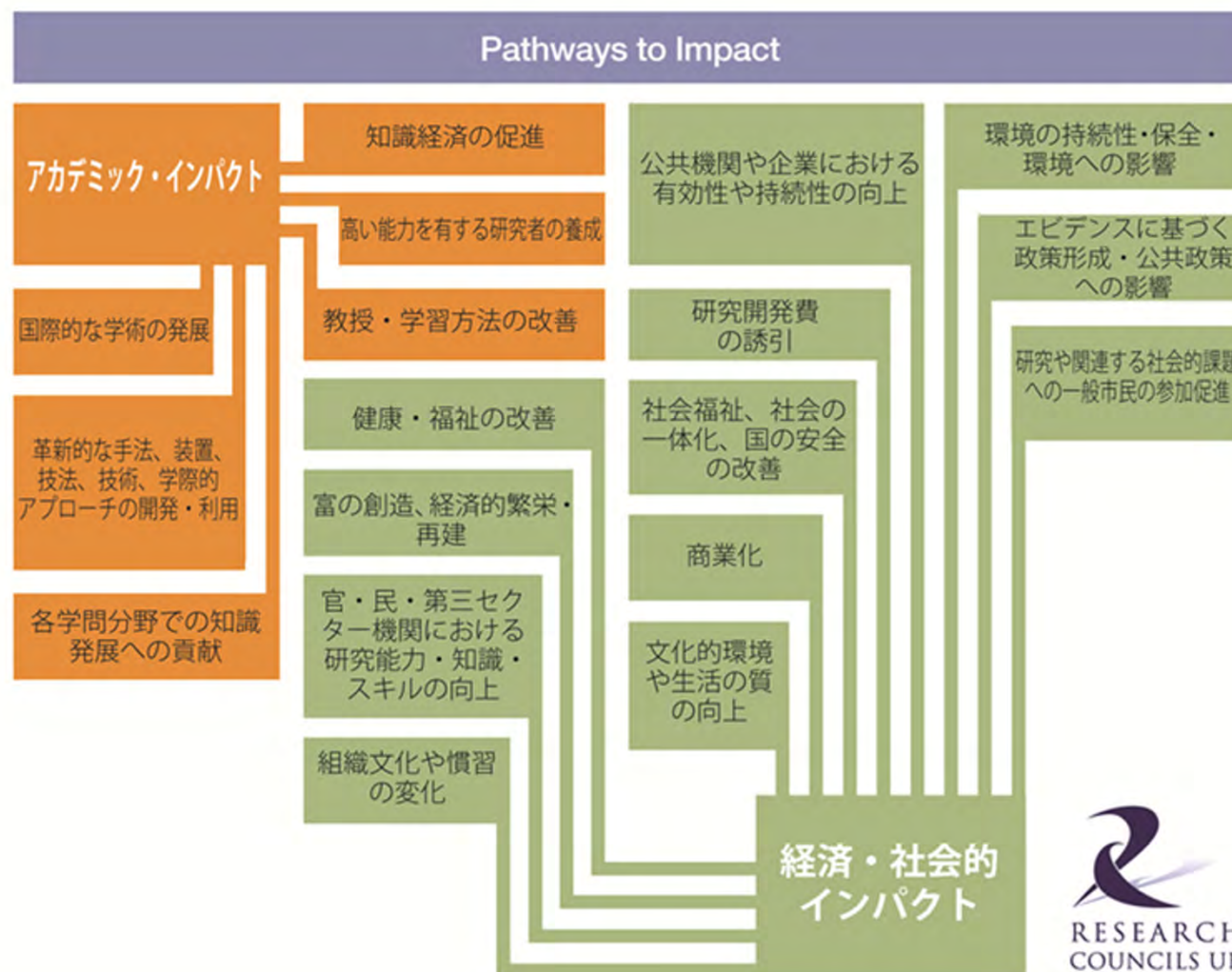
REFにおけるインパクト

学術を超えて、経済、社会、文化、公共政策・サービス、健康、生活の環境・質に関する変化あるいはベネフィットをもたらす効果

an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia

※なおPathway to Impactが事前評価であるのに対して、REFは事後評価であることに注意が必要

英国Research Councilにおける潜在的な経路 (pathways to impact)



REF2021



2019年：準備

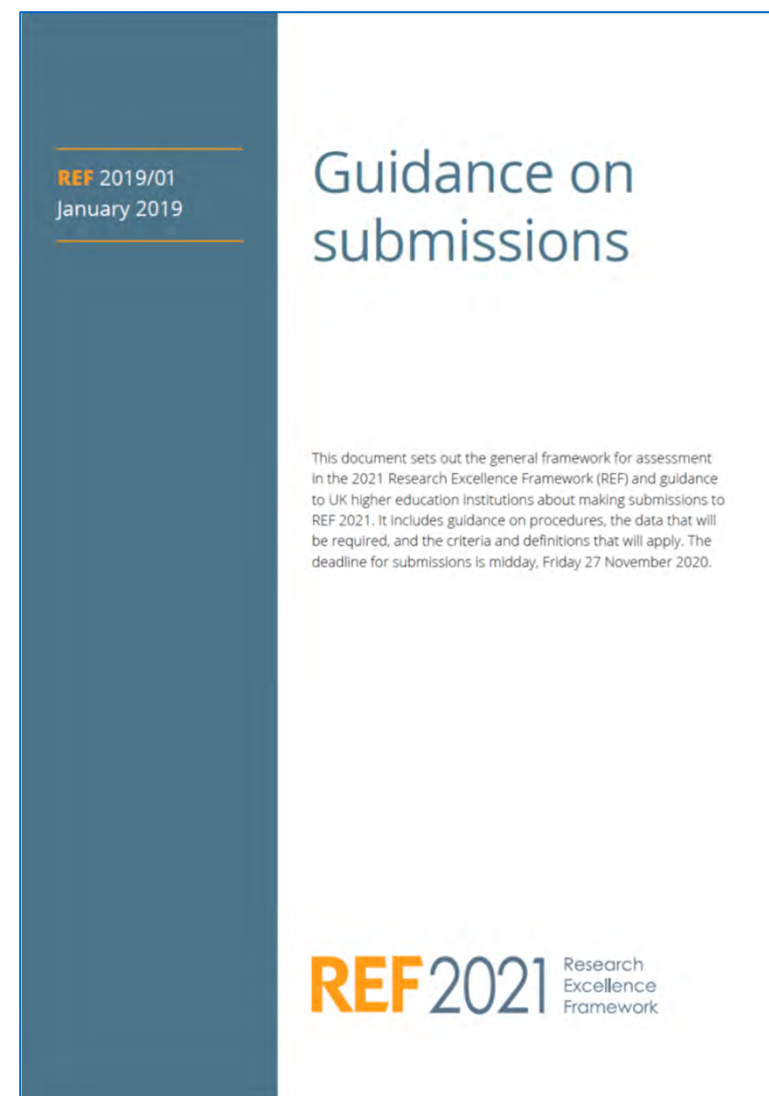
2020年：投稿（締切2020年11月27日）

2021年：アセスメント

評価基準

- ・ **アウトプットの質（60%）** :
“originality, significance and rigor”
- ・ **研究環境（15%）** :
“vitality and sustainability”
- ・ **社会的・経済的・文化的インパクト（25%）** : “reach and significance”

→インパクトの比率が5%up



英国REFにおける「インパクト」議論例

＜REF 2014の評価研究での提案ポイント＞

- 「機関」レベルでのインパクト：学術的研究や産学協同などの研究協力によるインパクト・ケーススタディを例示におけるより柔軟なアプローチの採用
- ケーススタディは、広い意味での研究成果に加えて研究活動や一連の作品に関連付けられても良い（幅広い成果の認知）
- REFの評価ガイドライン：インパクト・ケーススタディの具体的な内容の幅について狭い範囲に限定して解釈すべきでない

⇒評価対象となる「インパクト」には、社会経済的な領域に対して学術研究が及ぼすインパクトだけでなく、政府の政策、公共的な関与と理解、文化生活へのインパクト、専門分野を超えた学術的名インパクト、そしてまた教育に対するインパクトなどをも含むべきである

＜REF 2021でのポイント＞

- インパクトの評点割合が20%から25%に上昇

インパクトの範囲を狭くとらえているのはむしろ研究者の側でもある

Samuel and Derrick. "Societal Impact evaluation: Exploring evaluator perceptions of the characterization of impact under the REF2014" *Research Evaluation*, Vol.24(3), 2015, pp.229-241.

REF2014ヘルス関連パネルA評価者62名の半構造化インタビュー調査

- ・多くの場合、評価者は、例えば「ワクチンの精製・製造」などのような健康や経済に関わる直接的なアウトカムをインパクト指標として捉えてしまっている。
- ・「インパクトを単なる最終アウトカムではない」、「政策の変化」、「薬剤開発の停止」などもインパクトの一種であるという回答、「インパクトは偶発的かつ社会的なプロセス」と考えるようになったというケースも幾分か見られる・・・

研究者側がインパクトの射程を限定的に捉えているということの含意



国際的な議論の例として

ライデン声明の事例

1. 定量的評価は、専門家による定性的評価の支援に用いるべきである
2. 機関、グループまたは研究者の研究目的に照らして業績を測定せよ
3. 優れた地域的研究を保護せよ
4. データ収集と分析のプロセスをオープン、透明、かつ単純に保て
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
6. 分野により発表と引用の慣行は異なることに留意せよ
7. 個々の研究者の評価は、そのポートフォリオの定性的判断に基づくべきである
8. 不適切な具体性や誤った精緻性を避けよ
9. 評価と指標のシステム全体への効果を認識せよ
10. 指標を定期的に吟味し、改善せよ

※ライデンマニフェストの10原則の日本語訳は小野寺・伊神(2016)から引用している (Hicks et al. 2015; 小野寺・伊神 2016)

Leiden Impact Matrix : 大学は様々なインパクトを組織全体としてをカバーしていることに意味があるという捉え方



大阪大学 社会技術共創研究センター
Research Center on Ethical, Legal and Social Issues

相互作用 成果物	学術分野 (科学的相互作用)	専門的領域 (専門家の相互作用)	政官民セクター (企業・公的機関との相互作用)	広い社会 (公共的相互作用)
知識の生産と交換（結果）	同分野の研究者へのアウトリーチ ・共著（論文、本、コメント、査読有⇔査読無、オープンアクセス） ・特別な研究プロジェクトのアウトカム（含、博士論文の指導） ・学士・修士への教育 ・科学的・学術的団体への積極的参加	専門家へのアウトリーチ ・専門誌・マニュアル・書籍における出版・インタビュー ・専門家への講義 ・専門家との／専門家向けのプロジェクト・イベント	特定の企業や公的組織に向けたアウトリーチ活動 ・ビジネス書、行政発行のメディア、マニュアル、書籍などにおける出版・インタビュー ・講演活動やラウンドテーブル ・企業や公的組織との共同生産物・イベント	<u>一般の人々への学術的知見の普及</u> <u>・一般向け出版物（本、論文、コメント、各種インタビュー掲載記事）</u> <u>・公開講演</u> <u>・オープンアクセス</u> <u>・MOOCsなどオンライン講義</u>
知識の使用（効果）	同分野による研究アウトカムの利用 ・情報、装置、インフラ／研究施設、データセット、試験方法、研究室、モデル、プロセス、ソフトウェア、デザインの使用 ・引用	専門家による研究アウトカムの利用 ・情報、装置、モデル、プロセス、ソフトウェア、デザインの使用 ・専門家への助言 ・アントレプレナーシップあるいはインキュベーターセンターを通じた若手専門家の支援	<u>企業や公的組織による研究アウトカムの利用</u> <u>・情報、装置、モデル、プロセス、ソフトウェア、デザインの使用</u> <u>・各種クラスターと標準委員会への貢献（e.g. バイオサイエンスパーク、博物館、ISO、CENなど）</u> <u>・コンサルタント、政策助言</u> <u>・社会的に重要な課題への専門性の貢献（委員会委員、会議、ほか）</u>	<u>一般の人々による研究アウトカムの利用</u> <u>・フォーラム、テレビ、ソーシャルメディア（Youtubeやブログ）などでの議論への貢献</u> <u>・イベント・展示への貢献</u> <u>・卒業生への社会的価値付与</u>
知識の売り込み（収入）	同分野の研究者からの認識 ・研究グラント、科学賞 ・委員会のメンバーシップ、編集委員 ・客員教員としての契約 ランキングにおけるポジション獲得、EUネットワーキング活動と契約 外部資金 ・国内ファンド、EUファンド、その他 ・インパクトのための予算	リターンとゲイン： ・委託研究 ・専門家のためのトレーニング・講義（ポスト学術教育） ・有識者会議への参加（モニタリング委員会、評価委員会など） ・専門家による研究施設利用 ・実践／アントレプレナーシップ（スピンアウト・スピンオフ） ・副業	リターンとゲイン： ・委託研究（含、コンサルタンツ） ・専門的訓練・講義（ポスト学術教育） ・有識者会議への参加（モニタリング委員会、評価委員会など） ・企業や他の組織による研究施設利用 ・特許、ライセンス、公開、収入	リターンとゲイン： ・公的な賞 ・（有料の）オンライン講義 ・社会的団体からの職位・身分 ・雇用、仕事

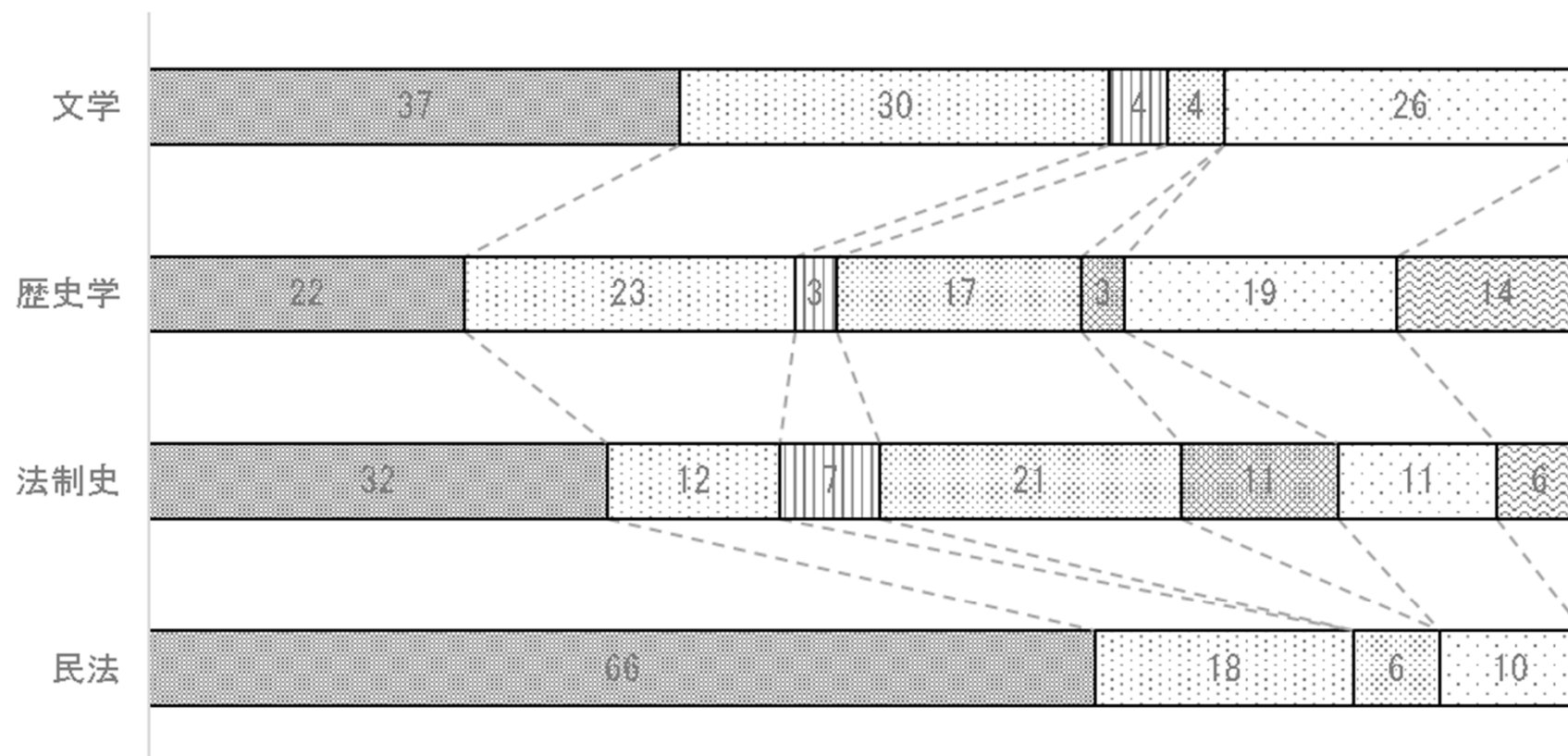
人文・社会科学分野の研究評価における基本的な問い



- ・ 人文・社会科学の分野によって研究目的や生産物が異なる
- ・ **出版経路の多様性**（本や論文の評価、媒体における知識伝達の役割が人文学内部でも異なる）
- ・ 対象**読者が多様**である
- ・ 出版**言語**の問題
- ・ 人文学における**論文や著作における時間的サイクル**の違い
- ・ 科学研究の実践に見られるような分野間の本質的な違いが人文・社会科学諸分野にも存在している（**業績体系や仕事の進め方の異なり、質の指標選択にも関わる**）
- ・ **組織化の度合いの欠如**

- ✓ 我々は研究の何を評価すべきか？
- ✓ どのような目的の下で、我々は研究を評価すべきか？
- ✓ どのように研究評価を行うべきか？
- ✓ 研究をめぐる質と評価の間のつながりはどのようなものか？
- ✓ 大学の内外の様々なステークホルダーに対して、研究の質とインパクトはどのようにして可視化されうるか？

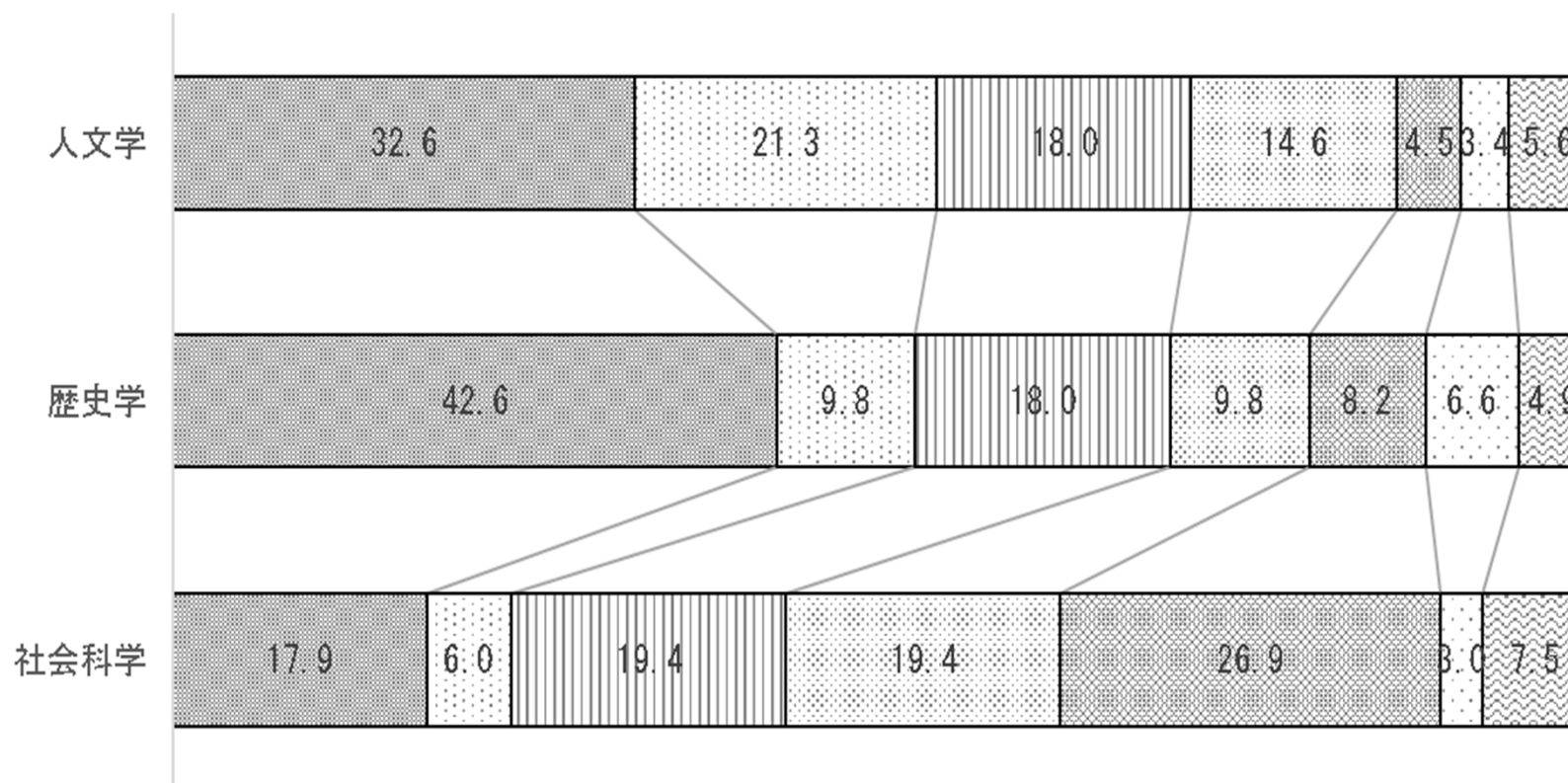
人文系分野の出版動態



■学術誌掲載論文 □研究図書掲載論文 ■辞書 ■紀要 ■雑誌・論評 □図書 ■その他の出版物

四分野におけるアウトプット媒体の比率
 (Williams and Galleron 2015, p.195を元に作成)

人文系分野の出版動態



■アプローチ ■データ ■理論 ■研究テーマ ■方法論 ■アウトカム ■未開拓の領域であること

人文学・歴史学・社会科学分野における評価されるオリジナリティ項目
 (Guetzkow et al. 2004, p.201を元に作成)

ノルウェーモデル

ノルウェーでの出版ポイントの重み付け
(Sivertsen 2010, 2015から訳出・作成)

	レベル 1 出版物	レベル 2 出版物 (ハインパクト出版物)
ISSN番号のある学術雑誌掲載論文	1	3
ISBN番号のある学術編著掲載論文	0.7	1
ISBN番号のある学術書	5	8

- ① (人文・社会科学分野を含めた) 全分野のピアレビュー学術文献に関する適切かつ構造化されたデータベースの構築
- ② 機関レベルで比較可能で分野毎の出版文化に応じた重みづけがなされた『出版ポイント』の規定
- ③ 出版ポイントによるパフォーマンスベースのファンディングモデル

「生産的相互作用」

Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society (SIAMPI)

1. 直接的相互作用（Direct interactions：個人的なつながり）、
2. 間接的相互作用（Indirect interactions：文書やマテリアル、モデルやフィルムなどのやり取り）、
3. 経済的相互作用（Financial interactions：研究契約、経済的貢献、研究プログラムへの寄付などの経済的関与・参加）

- 新しい「知識交換」のネットワークの拡大そのものが中期的なインパクト（あるいは知識生産にいたる中間生成物）として位置づけて評価する
- 評価の対象が研究そのものから相互作用のプロセスにシフト
- 関与するステークホルダー数の増大
- 「知識交換」のプロセスが研究者のモチベーションを効果的に高める効果



「ネットワークの失敗」（Fred and Matthew 2009）の回避
・・・必要となる時間やシステムの基盤の問題

(Spaapen & van Drooge 2011; Spaapen 2012, PenfieldTeresa et al. 2014; ERiC 2010; de Jong 2011)

「生産的相互作用」：より積極的に評価しうる事例 → 3. 11を巡る事例をみると？

＜歴史学・アーカイブ＞

- ・被災歴史資料保存・修復
- ・「歴史資料」の保存を通じた大学と地域の新しいネットワーク形成
- ・地域社会・行政・文化とのつながり
- ・「歴史」を残すことの再考

＜民俗学・人類学＞

被災無形文化財（お祭りなど）を巡る活動

- ・資料・知見提供、復旧支援
- 御神輿の復活、祭儀方法の提示
→ 伝統行事の復活は避難先との繋がりの中
→ 避難先に併せた伝統の変化＝根付き

「論文」などの評価スキームではされにくい、
「社会」とのつながりの中で非常に重要な学術的活動がある

東北大学災害科学国際研究所の例



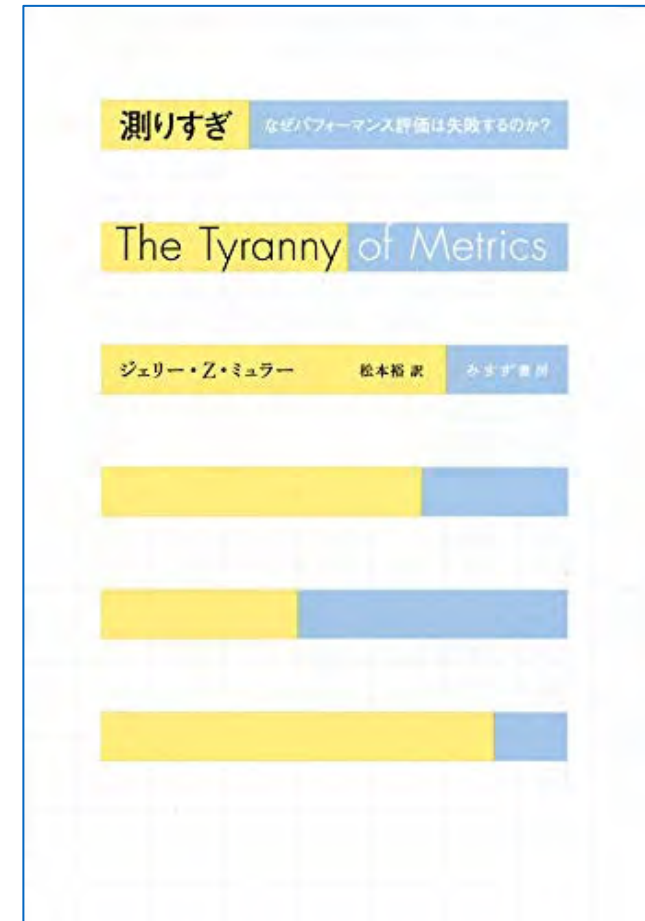


まとめに変えて—研究評価をめぐる課題

そもそも測りすぎることの弊害



- ・ パフォーマンス評価ベースのFundingへの批判
 - 予算配分を気にした短期的成果の希求
 - 論文の粗製乱造の問題
 - 「赤の女王」効果（red queen effect）
 （Geuna 2001; Geuna & Martin 2003; Butler 2003; 小林 2012）。
- ・ 少なくとも以下の事柄を検討する必要がある
 - 研究分野毎の特性、研究活動の行われ方
 - 研究に必要な必要人数や工数・資金
 - 研究や出版にかかる時間スケールとペース
 - 研究分野ごとの引用数の標準化・引用経路の違い（データベースの限界にも関わる）
- ・ 大学という場をイノベーション・エコシステムの中核とするならば
 - 日本の研究状況・環境、過去の知見の網羅的な分析
 - 研究評価制度が持つ課題と解決策を広く関係者間で検討
 - 分野によって異なる知識生産のための環境構築

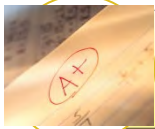


どのような制度的改革であれ、過去の議論と教訓を踏まえながら行わなければ期待されるような効果は得られない。

研究評価がもたらし得る悪影響



現在のエリート科学者から認知されようとする競争が生じ、エリート科学者との共同プロジェクトが実施されやすくなる。



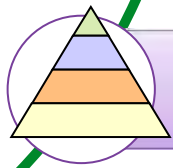
相対比較による評価を行うことで、学問分野ごとに標準的な評価基準が定まる。



研究の目標やアプローチの多様性が減少する。自身の研究が、その学問分野で現在優勢な研究目標にいかに関与するかを示す必要性が増す



現在の知的・組織的境界を超えるような新たな研究領域や研究目標の発展を妨げる。



ランキングによって研究者・研究チーム・研究機関の階層構造が強化される。

(From Whitley 2007: 標葉・林 2013)

⇒このような影響が一様に起きるわけではないが・・・

⇒研究評価を他の要因を含めた広い文脈の中でとらえ、知識の生産性や多様性に及ぼす影響を分析する必要／個人や集合レベルでの知識生産の理解

今回のまとめとして

- ・ 「先端的な研究／知識」への投資／それを支援する評価システム
 - ✓ 知識生産をめぐるエコシステム形成の視点
 - ✓ *Scientia*（知識）のための政策基盤になるか否か
- ・ 科学技術基本法の改訂に伴う変化
 - ✓ 負の影響を防ぐために出来る議論・施策は何か？
 - ✓ 測りすぎていいことはあまりない
 - ✓ 評価基準が形成されてきた歴史的経緯や規範の意味の再考
- ・ 新しい評価取り組みや視点の検討
 - ✓ 生産的相互作用などの新しい視点も踏まえた、より多様な評価軸の検討
 - ✓ 学術が持つ様々な知的社会貢献を評価すべき
 - ✓ 評価されにくかった活動（縁の下の力持ち）の評価