

EBPMを支えるデータ・情報基盤の構築： NISTEPの15年間の歩み

SciREXオープンフォーラム

2025年11月21日（金）

富澤宏之

文部科学省 科学技術・学術政策研究所（NISTEP）

第2研究グループ客員総括主任研究官

本プレゼンの概要

1. NISTEPデータ・情報基盤の概要
2. これまでの取り組みを振り返って
 - (1) ビブリオメトリクス指標とデータの“名寄せ”：
SciREX以前の歴史
 - (2) 計量書誌データの“名寄せ”の意義
 - (3) データ・情報基盤のコンセプトの進化
3. 今後の展望

パート1

NISTEPデータ・情報基盤の概要

3

SciREXへのNISTEPの取り組み データ・情報基盤の構築・整備（1）

【事業全体の目的・概要】

- エビデンスに基づく科学技術イノベーション政策の推進、及び、そのための調査分析や政策研究の基礎となるデータ・情報を体系的に活用する基盤を構築する。
- 特に、日本の大学・公的機関の科学知識生産（論文等）についての体系的な分析や、日本の産業における研究開発・イノベーションについての分析を行うための各種データを体系的に活用するための基盤を構築する。

主要な成果物：“辞書”等（NISTEPのwebサイトで公開）

NISTEP大学・公的機関名辞書

（約2万1千の機関を掲載；約1万6千の代表機関と約5千の主な下部組織）

- 大学、大学共同利用機関、短期大学、高等専門学校、公的研究機関（国立研究開発法人等）を中心として収録
- 非営利団体や会社等についても、研究を行っている機関は可能な限り掲載
- 全収録機関の93.2%には英語名も掲載

NISTEP企業名辞書

（約1万1000社を掲載；変遷企業名数約2万5000件を含む）

- 全上場企業、特許出願数や意匠・商標登録数が累積100件以上の企業、特許出願数の増加率が一定以上の企業などを収録
- 特許出願データ約1500万件（1970年～）をベースに国内企業の名称変遷・合併等の沿革や企業情報を毎年処理

データ分析のための接続テーブル

- 大学・公的機関名や資金配分機関等名の英語表記ゆれテーブル
- NISTEP大学・公的機関名辞書と論文データ（Scopus, WoSCC）との接続テーブル
- NISTEP大学・公的機関名辞書と世界の研究機関のレジストリー（ROR: Research Organization Registry）のデータとの対応テーブル
- NISTEP企業名辞書とJIPパテントデータベース（特許統計分析用DB）との接続テーブル

【“辞書”等の直接的効果】

- 論文書誌データや特許データの質の問題（データクリーニングや名寄せの必要性）を大幅に改善
- 科学技術政策やイノベーションの研究者のデータ処理の負担を大幅に軽減
- これまでマクロデータ分析に留まっていた課題のミクロデータ分析が実現（例：国レベルの分析から個別機関レベルの分析に深化）
- インプット系データとアウトプット系データの連結（例：研究開発費データと論文データの連結）

活用事例

- 大学、シンクタンク等による政策研究において大学・公的機関名辞書、企業名辞書等を活用
（例）財務省財務総合政策研究所「イノベーションを通じた生産性向上に関する研究会」報告書、2018年3月。
- NISTEPの調査研究レポート（多数）
（例）日本の部門別・大学グループ別論文数など
- 大学のIR専門家やURAなどによる個別大学の研究力の分析
- 国際的な認知（国際学会、OECDフォーラムでの紹介、OECD専門委員会、RORの会合での発表等）
（例）PICMET（技術経営国際会議）主催の国際会議（2015年7月）での専門セッション、OECDブルースカイⅢ科学・イノベーション指標フォーラム（2016年9月）における欧州研究者による言及

【“辞書”等の波及効果】

- NISTEPの調査研究能力の向上
例：日本の科学研究力についての様々なデータを政策立案のために提供
- データの質的向上により、多様な学問領域からの研究者が参入
例：経済学者が論文書誌データの分析を開始
- 各大学のIR専門家やURAによるデータ活用が拡大

4

主要な成果物：検索ツール等

(NISTEPのwebサイトで公開；下記以外にも各種のデータや検索ツール等を公開)

科学技術・イノベーション白書検索

- 「科学技術・イノベーション白書」（1958年版以降、2025年版までの全ての）テキスト情報を検索し、各年代の科学技術に関する政策・施策の動向を調べることが可能な検索ツール
- 注目する語句の出現回数を出力する機能や、キーワードだけでなく、類義語も併せて検索する「あいまい検索」などの機能も保有

基本政策文書データベース

- 科学技術・イノベーション基本計画（第1期～第6期）とその前身である科学技術会議答申、科学技術イノベーション総合戦略、統合イノベーション戦略、の全テキストデータを公開
- 「科学技術・イノベーション白書検索」と同様の検索機能を保有

デルファイ調査検索

- NISTEPの「デルファイ調査」(※)の1971～2024年の全12回の調査結果をキーワード等によって検索可能
- ※ 今後30年間で実現が期待される科学技術の実現時期や重要性等を専門家が予測するアンケート調査

NISTEP定点調査検索

- 「NISTEP定点調査」(※)の2011～2015年の結果について、機関属性別や個人属性別、時系列の集計結果の表示、自由記述回答の検索などが可能

※ 参考：第1期には政策課題対応型調査研究も実施

- 政府の研究開発投資の経済的・社会的波及効果等の総合的な調査研究（2016年度より、NISTEPからSciREXセンターに発展的に継承、一部をNISTEPの定常的な調査研究へ移行）

政策課題に対応した調査の実施

博士人材の追跡調査

(第2期以降はNISTEPの定常的調査研究として実施)

- 目的・概要：
 - ・ 持続的な科学技術イノベーションの主たる担い手である博士人材の社会全体における活躍状況を把握・提示
 - ・ 博士人材のキャリアパスの把握・可視化と、それを通じたエビデンスベースの科学技術政策・人材政策の立案への貢献
- 調査の展開：
 - ・ 博士課程を有する大学が参加する情報基盤プラットフォームとして博士人材データベース(JGRAD)を整備（2025年9月4日現在、158大学が参加）
 - ・ 特定の修了年の博士課程修了者コホートの置かれている状況の追跡調査(JD-Pro)も実施

NISTEP定点調査

- 概要：
 - ・ 統計からは把握しにくい科学技術やイノベーション創出の状況について、第一線で研究開発に取り組む研究者や有識者の意識をアンケート調査
 - ・ 繰り返し調査により、変化に対する現場の意識を追跡
- 最近の調査結果の例：
 - ・ 2024年調査では、望ましい能力を持つ博士後期課程進学者の数、基盤の経費の確保、研究時間などの状況が厳しく、物価高騰が研究環境に深刻な影響を及ぼしているとの認識

5

我が国におけるデータ・情報の活用を促進するための各種活動

- 科学技術イノベーション政策・施策・事業のエビデンスベースの効果把握やPDCAサイクルの確立や向上に向けて、データ・情報の整備・活用の促進の“旗振り役”をNISTEPが担う

関係機関ネットワークの構築と会合開催
(2013年度～2020年度)

- 目的：
 - ・ 研究開発ファンディング等のデータの整備・活用の促進
 - ・ 課題や問題意識の共有
- 主要な検討ポイント：
 - ・ 各機関の内部データ基盤の整備と活用状況・課題
 - ・ データの標準化、機関を越えたデータ接続
 - ・ ファンディング情報の戦略的活用
 - ・ データ・情報の整備・活用に関する政策提言
- 参加機関：
 - ・ 科学技術振興機構(JST)
 - ・ 日本学術振興会(JSPS)
 - ・ 日本医療研究開発機構(AMED)
 - ・ 新エネルギー・産業技術総合開発機構(NEDO)
 - ・ 農業・食品産業技術総合研究機構(NARO)
 - ・ 情報通信研究機構(NICT)
 - ・ 経済産業研究所(RIETI)
 - ・ 国立情報学研究所(NII)
 - ・ 大学改革支援・学位授与機構(NIAD-QE)
- 活動の反響等
 - ・ 当ネットワークの検討内容を取り入れた実践的取り組みを内閣府が開始（2017年12月22日に内閣府が開催した研究開発法人情報連絡会において、当ネットワークの活動内容を報告）
 - ・ 日刊工業新聞(2018年1月26日／30面)に当ネットワーク会合の活動に言及した記事が掲載

「体系的番号」の提案と普及促進活動

- 目的・概要：
 - ・ 論文の謝辞に記載される研究資源情報(受給した研究費、使用した研究施設、情報・観測データ等の情報)を、研究開発支援施策・事業の成果を把握するための情報源として活用できるようにする。
 - ・ 論文の謝辞に記載する研究資源情報のなかに、受給した研究費・資金配分制度ごとに付与した体系的課題番号を明記することをルール化する。
- 成果：
 - ・ 政府全体として「体系的番号」の付与が制度化
「論文謝辞等における研究費に係る体系的番号の記載について」(令和2年1月14日、競争的研究費に関する関係府省連絡会申し合わせ)
 - ・ NISTEPのwebサイトにおける「体系的番号」の公開・更新

NISTEP機関同定プログラムの公開

- 目的：
 - ・ NISTEP大学・公的機関名辞書の構築・整備のために開発した機関同定プログラムを利用希望者に公開
- 実施期間：2021年12月より実施
- 公開内容：
 - ・ 機関同定プログラム
 - ・ 辞書に独自の機関を追加できる機能(部局等についての独自分析をしたいとの要望を受けた機能)等
- 約40名(大学のURAやSciREX関係者等)が利用中

6

NISTEPデータ・情報基盤の構想（特に“辞書”や“名寄せ”）はどのように受け止められたか

- それ自体の評価は高いが・・・
 - 「継続することが重要」との評価
 - ただし、“地味”な印象がある？
 - “辞書”や“名寄せ”はデータの精度を上げるためのものと認識されている？
- 利用イメージが違う？（“同床異夢”）
 - データを整備すれば、すぐに有用な“エビデンス”が得られるという期待（？）
 - 「データ→エビデンス」という期待（“分析”の必要性・重要性・困難性が分かり難い）
 - 主な利用者（政策研究者なのか政策立案者なのか）についてのイメージの違い
 - SciREX事業全体のなかでの役割・機能についてのイメージの違い
 - 「SciREXで必要なデータをNISTEPが整備する」という期待？

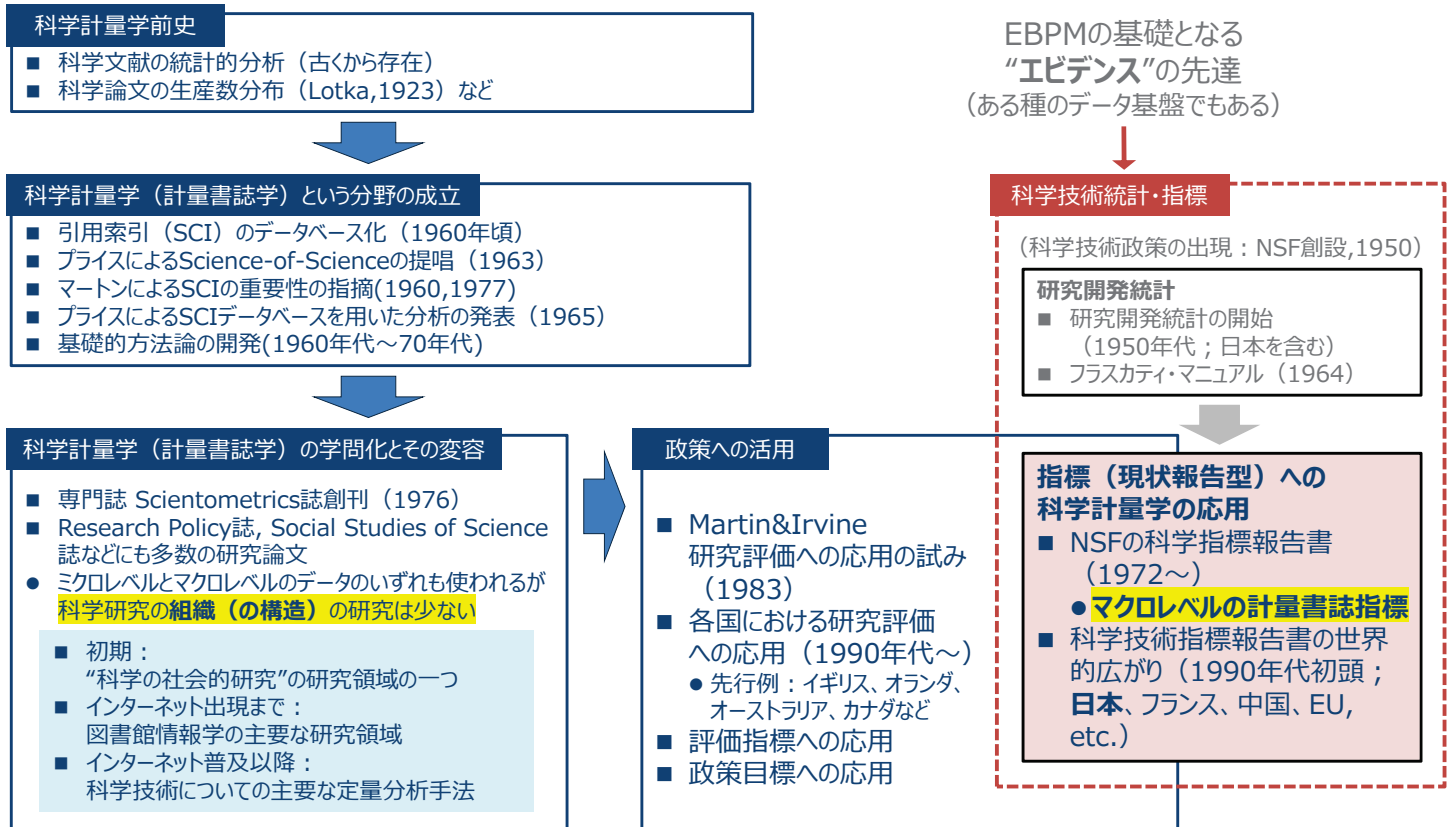
7

パート2(1)

ビブリオメトリクス指標とデータの“名寄せ”： SciREX以前の歴史

8

科学計量学と科学技術指標の歴史より



参考: 藤垣裕子・平川秀幸・富澤宏之・調麻佐志・林隆之・牧野淳一郎『研究評価・科学論のための科学計量学入門』(丸善), 2004年

9

ビブリオメトリクス指標やデータの“名寄せ”との個人的な関わり

個人の経験

- 1989年: NISTEP加入
 - 科学技術指標開発の担当グループに配属
- 1990年: 科学技術指標のための計量書誌学への取り組みを開始
 - 米国NSFの科学工学指標と英国SPRUの科学技術政策研究がモデル
 - マクロレベルの計量書誌指標 (国別の論文数・論文被引用度など) は作成できたが、他の指標のように産・学・官の内訳を示すことが困難であることが不満
 - 論文著者の所属機関名の“名寄せ”の必要性 (& その困難さ) を知る
- 1991年3月: サセックス大学SPRU (科学政策研究発祥の地) 訪問
 - 計量書誌学の研究評価への応用の試みで世界的に有名であったBen Martin氏に面会
 - “名寄せ”の本格的な取り組みを目の当たりにする
- 1991年9月: NISTEPが日本で最初の科学技術指標レポートを公表
 - 計量書誌指標 (日本などの主要国の論文数・論文被引用度等) はかなり注目された
 - 科学技術指標レポートの担当は2014年度まで (本格的な関与は2004年版が最後)
- 1990年代～2000年代: マクロレベルの計量書誌指標を中心に、科学技術に関する統計・指標 (今の言葉で言えば“エビデンス”) の研究者として活動
- 2003～2004年: 「基本計画の達成効果の評価のための調査」(NISTEPの大規模調査)
 - 林隆之氏 (現GRIPS)、山下泰弘氏 (現JST/CRDS) の協力を得て、論文データの網羅的な名寄せを実施
- 各種の活動
 - 2004年: 『研究評価・科学論のための科学計量学入門』(丸善)の執筆に参加
 - 2006年: ブルースカイ国際会議で、計量書誌データの名寄せの必要性を主張 (後述)

関連動向

- ◆ 日本国内では図書館情報学の研究者による計量書誌学の研究が盛ん
 - 学問としての研究がほとんどで、政策への応用は語られるのみ
 - 大規模な研究プロジェクトも見当たらなかった
 - 個別の研究においては、データの精度に敏感であり、データの“名寄せ”に類似したことは普通に行われていた
 - 研究機関レジストリや情報標準化の必要性も既に議論されていた
- ◆ 1990年代後半: 国立情報学研究所が日本の主要な論文生産機関の個別データを作成・公表 (大規模な“名寄せ”を実施)
 - ➔ 根岸正光・山崎茂明『研究評価』(丸善) 2001年
- ◆ 2000年: 大学別の計量書誌指標を匿名のグループが発表した“事件” (今の言葉で言えば“炎上”)

科学技術指標の歴史的発展 (2000年まで)

指標、理論、専門家	1950~60年代	70年代	80年代	90年代
用いられた主な指標	・ R&D	・ R&D ・ 特許 ・ 技術貿易	・ R&D ・ 特許 ・ 技術貿易 ・ ハイテク製品 ・ 計量書誌学 ・ 人材 ・ イノベーション調査	・ R&D ・ 特許 ・ 技術貿易 ・ ハイテク製品 ・ 計量書誌学 ・ 人材 ・ イノベーション調査 ・ 技術文庫におけるイノベーションへの言及 ・ 製造技術の調査 ・ 産業技術への公的支援 ・ 無形資産 ・ 情報通信技術についての諸指標 ・ 投入・産出行列 ・ 生産性 ・ リスクキャピタル ・ 吸収合併と買収 ・ 経済分析から借用された指標
イノベーション・モデルの概念化	リニアモデル	→	チェーン・リンクド・モデル	システミック・モデル
科学技術指標の専門家の役割	方法論とデータの提供者	→		データ、方法論、分析の提供者。様々なタイプの指標(科学技術指標と経済指標の両方)の統合者

出典: OECD, "Report on the Activities of NESTI, DSTI/STPI(2000)26, September 2000. (日本語訳は発表者による)

科学技術指標の展望:2000年代と今後

	1950~60年代	70年代	80年代	90年代	2000年代	2010年代
用いられた主な指標	・ R&D	・ R&D ・ 特許 ・ 技術貿易	・ R&D ・ 特許 ・ 技術貿易 ・ ハイテク製品 ・ 計量書誌学 ・ 人材 ・ イノベーション調査	・ R&D ・ 特許 ・ 技術貿易 ・ ハイテク製品 ・ 計量書誌学 ・ 人材 ・ イノベーション調査 ・ 技術文庫におけるイノベーションへの言及 ・ 製造技術の調査 ・ 産業技術への公的支援 ・ 無形資産 ・ 情報通信技術の諸指標 ・ 投入・産出行列 ・ 生産性 ・ リスクキャピタル ・ 吸収合併と買収	90年代までの指標 ・ 産学連携 ・ 知識のスピロオーバー ・ 非技術的イノベーション ・ リンケージやネットワークの指標	???
科学技術、イノベーションの概念	リニアモデル		チェーン・リンクド・モデル	システミック・モデル	モード論、知識社会論 ナショナル・イノベーション・システム 地域・クラスター オープン・イノベーション	サービス産業、ソフト オープン・イノベーション
主要な政策上の論点			産業競争力	経済学的視点の導入、研究評価/政策評価、イノベーション促進、グローバルイノベーション、新興国の台頭、エビデンスベースの政策形成		

(前ページの表に基づき、発表者が作成)

次世代の指標をめぐる議論の例: ブルースカイ会議1996

- 会議の概要
 - OECD主催の会議
 - 10年先を見据えて、今後必要となる科学技術指標は何かを議論
- 主な論点
 - アウトプットの指標
 - 特許統計の重要性がクローズアップ
 - OECDと3極特許庁との連携などの契機に
 - 計量書誌学指標への批判
 - 英語圏(特に米国)への偏り
 - 個別の論文の価値の違いを無視した“荒さ”

次世代の指標をめぐる議論の例: ブルースカイ会議2006

- 会議の概要
 - OECDとカナダ政府が主催し、米国NSFが支援
 - 次の10年間を展望(1996年よりちょうど10年後)
- 主要な論点
 - 新規指標の提案は少なかった
 - “データの新しい組み合わせ”
 - 例: イノベーション調査データと企業財務データ
 - マイクロ・データの活用 (統計の個票などの活用)
 - マクロ・データのための分析には限界 ⇒ 体系的なマイクロ・データの活用へ
 - 異なる種類のデータをマイクロ・レベルでリンク
 - 主としてイノベーションやその経済学的分析のための統計データに重点
 - “オープン・イノベーション”がクローズアップ



計量書誌データにおけるマルチレベルデータの重要性和データ名寄せの必要性についての国際発信 (2006年9月27日 Blue Sky II Forum 発表資料より)

Constructing a Multi-level Scientometric Indicators System

Blue Sky II 2006 Forum

What Indicators for Science, Technology and Innovation Policies in the 21st Century?

September 25 - 27, 2006
Ottawa, Ontario, Canada

Hiroyuki TOMIZAWA¹ & Takayuki HAYASHI²

1. National Institute of Science and Technology Policy [NISTEP], Japan
2. National Institution for Academic Degrees and University Evaluation [NIAD], Japan

Macro vs. Micro Bibliometrics

- Macro-level bibliometrics
 - has developed with S&T indicators
 - based on large quantity of data and rather simple method
- (Meso-level)
- Micro-level bibliometrics
 - Mainstream of bibliometrics as a discipline
 - Suitable for highly original research
 - “Laboratory” for new methods of analysis

Macro-level
・ country

Meso-level
・ sector
(higher education, government, industry)

Micro-level
・ research group
・ research area
・ research institution
・ person (author)
・ paper

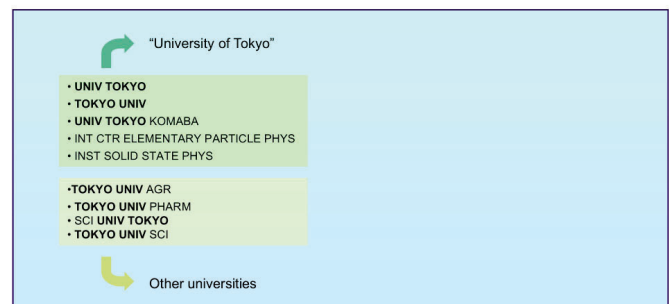
「マルチレベルの科学計量指標システムの構築」というタイトルで発表

- マクロレベルとミクロレベルのビブリオメトリクスの間の乖離を指摘
 - ◆ ミクロ指標を足し合わせても国全体の値にならない、など
 - その解消のためには、データの“名寄せ”が重要であることを主張
- なお、この会議では、実質的な大統領科学顧問であったマーバー氏が、米国SciSIPプロジェクトの基となる構想を講演した

機関名の“名寄せ”の重要性を説明

Methodological Challenge of Multi-level Structural Database

- “Name identification” and “standardization of names”



- The most labor-effective way is, although seemingly paradoxical, to do name identification and categorize all organizations in Japan, rather than accumulate the search result of each organization.

パート2(2)

計量書誌データの“名寄せ”の意義

13

データ名寄せの重要性についての説明の例
(2014年7月18日 第7回「科学技術イノベーション政策の科学」構造化研究会)

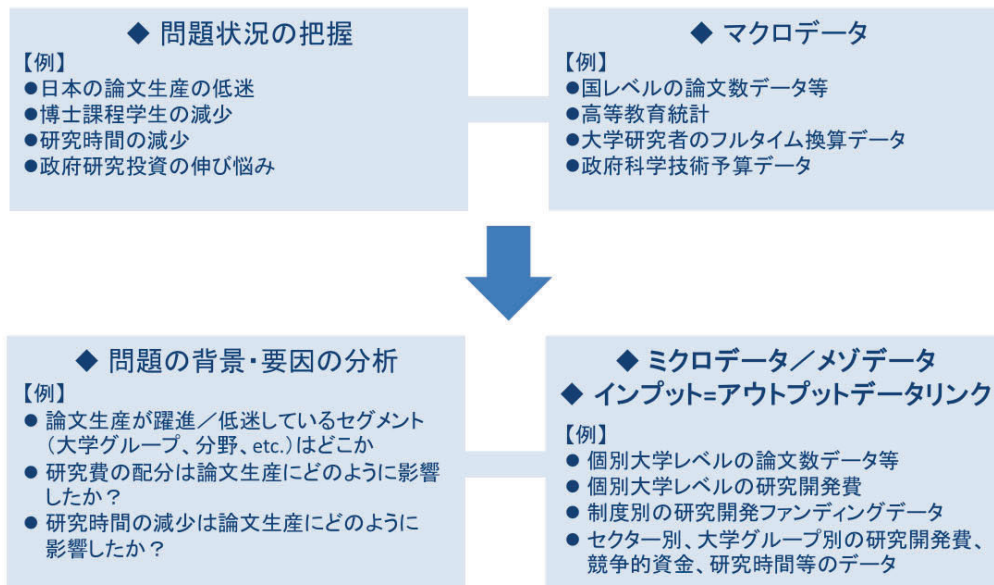
データの“名寄せ”の意義

- 単に「データの精度を高める」こととは全く異なる
- そうではなく、全く利用不可能であったマイクロ・データや中間レベルのデータを新たに創り出す取り組み
- 膨大な作業の重複・繰り返しを避ける
 - 研究者がオリジナルデータを分析しようとする、ほとんどの場合、分析そのものに先立ち、膨大な作業が必要となる
 - それをNISTEPが一手に引き受ける形(いわば“汚れ仕事”)
- 本質的に、直ちに“実装”という性格の仕事ではない
 - 時間のかかる地味な仕事
 - しかし、より広い観点からは、“実装”までに最も時間を要する部分を減らす取り組みとも言える

データの“名寄せ”の意義を包括的に説明しているが、どちらかと言えば、実施者の視点からの説明

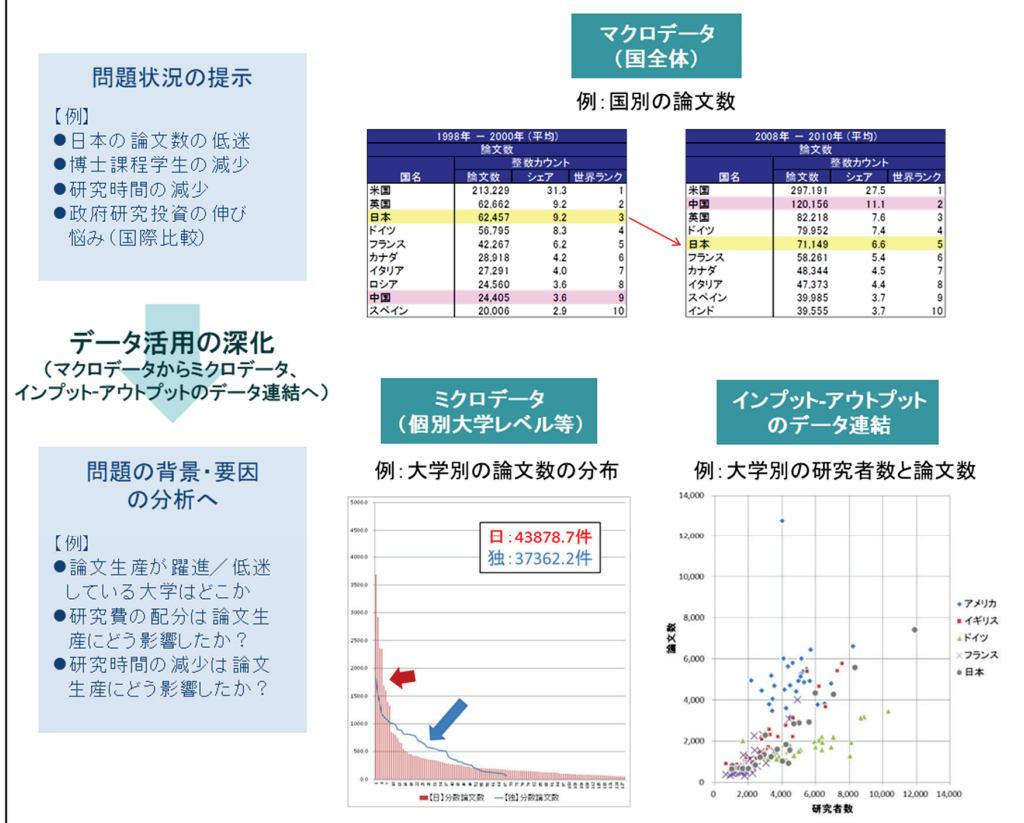
14

問題の要因を分析するためのデータの高度化の必要性



15

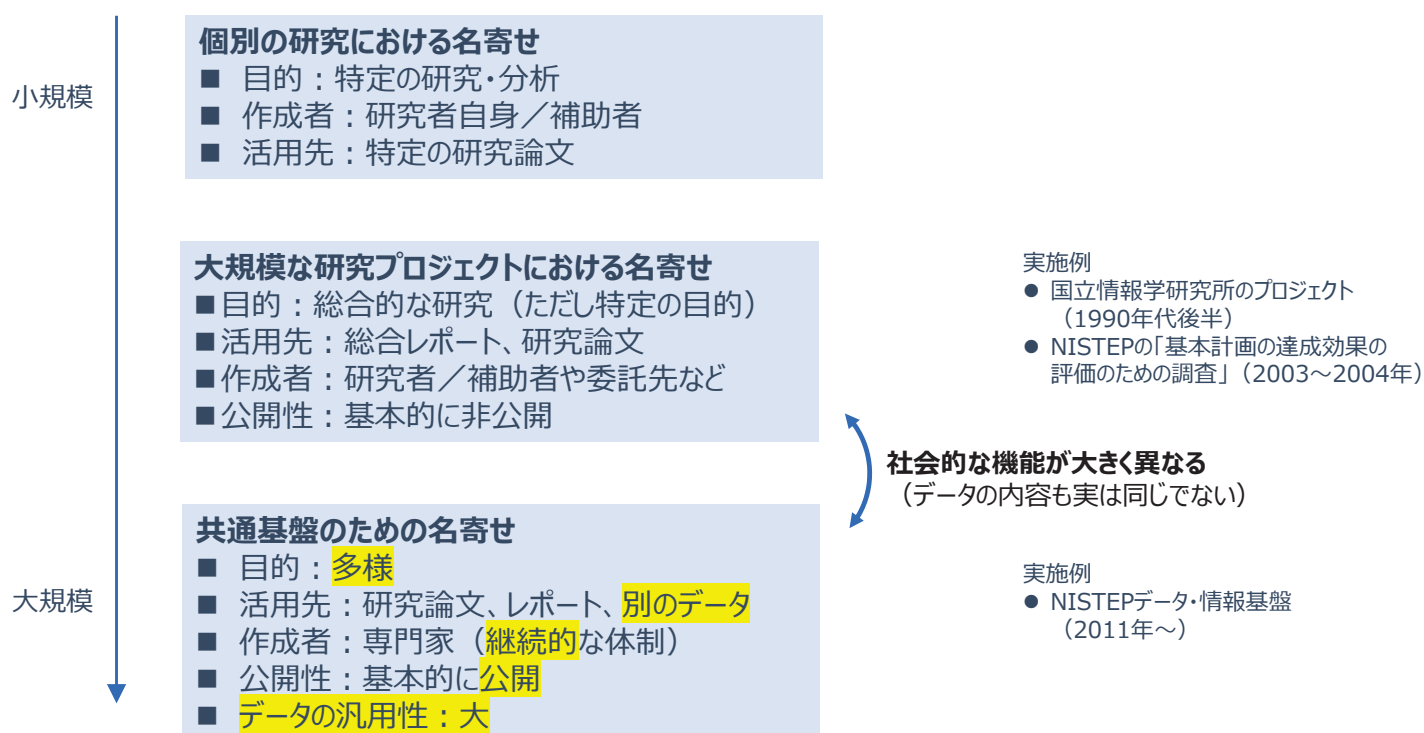
データの活用イメージ: 科学研究についてのマイクロ分析とインプット-アウトプット分析



マクロデータとマイクロデータで分析できることの違いを、例を用いて説明

16

計量書誌データの“名寄せ”（マイクロレベルのデータの整備） の実施レベルによる違い



17

パート2(3)

データ・情報基盤のコンセプトの進化

18

研究基盤としてのデータ・情報基盤の重要性

当初考えていたよりも「研究基盤としての」という性格が強くなったことの背景を説明

開始時点の基本コンセプト

1. 政策形成プロセスをより合理的なものにするための基盤
 - 政策立案のための客観的根拠(エビデンス)としてのデータ
 - 政策議論の質の向上のためのツール
 - 政策の評価や検証の基礎
 - 個別機関での政策立案のためのデータ
2. 科学技術イノベーションに関する政策研究の基盤
 - 政策研究における科学的方法論の強化
 - データの充実により、多様な学問領域からの研究者の参入を促す
3. 国民に対する説明責任
 - 公的科学技术システムの透明化
 - 科学技術政策の効果の提示

研究基盤としての機能を重視する背景

1. データ・情報基盤が直ちに政策形成の直接的エビデンスになる場合は少ない
 - データ・情報基盤は、十分な分析がなされて初めて意味を持つものが多い
 - 行政実務者が政策立案に必要なデータ・情報基盤を事前に明確に想定することは困難
 - データ・情報の活用では、その背景となるモデルや仮説の体系(あるいは“理論”)が重要
2. 政策研究者や専門的な分析家の介在が必要である場合が多い
 - データ・情報の活用では、データ・情報の内容についての詳細な知識が必要な場合が多い
3. 政策研究におけるニーズが高い
 - 科学技術イノベーションに関するデータ・情報は不十分
 - 政策研究者の多くはデータ・情報の整備に多大な労力をかけている

19

データ・情報基盤とその“実装”の間にあるもの

「分析」の重要性を強調

- データ・情報は十分な分析がなされて初めて意味を持つものが多い
- 政策研究者や専門的な分析家の介在が必要である場合が多い
 - データ・情報の活用では、データ・情報の内容についての詳細な知識が必要な場合が多い
 - データ・情報の分析結果の解釈は更に困難
- 政策立案に必要なデータ・情報を事前に想定することは容易ではない
 - 行政ニーズに適合したデータ・情報基盤をつくるのが困難である要因
 - 政策立案者自身も、どのようなデータが必要か明確にできない場合も多い
 - データ分析者の多様な取り組みが必要
 - 共通基盤的なデータ・情報を整備することが有用ではないか？

「共通的・基盤的なデータ・情報基盤」の有用性を提起しているが、少し後で微修正している(後述)

20

これまでの科学技術政策研究の経験より

- 政策に有用なエビデンスを作成するために、ポリシーメーカーと政策研究者のインターアクションは重要であるが、簡単なプロセスではない。
 - 「ポリシーメーカーのニーズ ⇒ 政策研究者」という図式は、成り立たない場合がほとんど
 - ポリシーメーカー自身が、どのようなエビデンスが必要か、明確でない場合が多い
 - ポリシーメーカーのニーズが明確な場合でも、実際にエビデンスを作ることがあまりに困難な場合も多い
 - ただし、ポリシーメーカーからのニーズの発出が続くと、政策研究者も何が重要か分かってくる
 - 政策研究者が自分でこれが有用と考えるエビデンスを自ら作ることは、時間を要するが、有用であることが多い
 - ただし、当然ながら、政策研究者が真剣に“顧客”(ポリシーメーカー)に、自分の成果を売り込もうとするインセンティブが無いと成り立たない

21

「データ・情報基盤の構築」への取り組みからの考察

- 「共通基盤的データをNISTEPが作成」というモデルはこれまである程度機能したが、限界に突き当たっている
 - データ整備と“分析”の間に距離がある
 - 共通基盤的データは、それほど多くない
 - 政策エビデンスや政策研究に必要なデータは多様であるため(特に細部)
- 中央集約型モデルではなく分散型／ボトムアップ型へ
 - (例)NISTEPがデータを整備するだけでなく、研究者等によるデータ構築結果をNISTEPに集約する仕組み
 - (例)大学では、URAやIR担当者による個別大学レベルのデータの分析が進んでいるので、これらの取り組みをネットワーク化する
- データ構築とデータ分析を分離した事業モデルから両者を一致化したモデルへ
 - 総花的でなく“テーマ”を絞ってでも、データ構築とデータ分析を一体化した方が良いのではないか

「共通基盤的データ」という考え方(前述)を少し“軌道修正”することを考えはじめた

「分散型／ボトムアップ型」という考え方は重要

22

データ整備の障害: 複雑な要素

「よく設計された集約型のデータ・情報基盤」という“幻想”を否定しようとしている

- データの整備には、複雑な要素が絡み合う
 - 2012年にNISTEPの国際WSで、先行したデータ整備の取り組みをしている各国の専門家は、皆、「データ整備は複雑なので、体系的な計画を立てて、それに沿って進める方法はダメ。いろいろと試して、上手くいったものを強化していく方法が良い」と言っていた。
- 博士人材DBの例:
 - 当初、「政策のための科学」に必要なデータを、各大学から提供してもらう、というモデルであった
 - 大学からみると、単なるデータ提供の依頼に過ぎない
 - 各大学の運営に重要なデータを各大学が集め、その一部をNISTEPに集約、というモデルに修正
 - 各大学がデータが必要と理解すれば、積極的に関与するようになる
- ファンディング・データの整備の例:
 - 先行して整備されていた科研費DBを用いた分析結果を出したことがJSTのFMDBにつながった

23

データ・情報のオープン化の重要性

「オープン化」の考え方は、「分散型・ボトムアップ型」(前述)という考え方とも関連する

- “(広義の)政策研究の重視” は、別の問題を生む可能性
 - 政策形成プロセスにおいて、行政部局と政策研究機関・専門家という2つのアクターによる閉じた世界が形成される恐れがある
 - 行政部局は、自らの政策の裏付けとなるようなデータ・情報を求める傾向
 - 政策研究機関・専門家は、行政部局のニーズに応えるデータ・情報や分析結果を提供しようとする傾向
 - このような関係性が確立すると、データ・情報基盤は、行政部局の政策を一方向的に正当化するためのものに陥りかねない
- このような問題の回避のためにはデータ・情報のオープン化が重要
 - データ・情報が多様なユーザーによって活用されることにより、様々な視点からの政策エビデンスが提供と利用が進展する
 - 政策議論が多様な視点からなされるようになる
- ◆ 本プロジェクトで整備したデータ・情報は可能な限り公開する方針

科学論の「科学知識はなぜ信頼性が高いのか」という議論における“反証可能性”の概念に近い考え方

24

今後の展望

25

NISTEPデータ・情報基盤の今後に関連するいくつかの論点

- データ・情報基盤は分散型の整備体制に移行していくことが望ましい？
 - NISTEPだけでなく、関係機関、データのユーザー、等による整備体制
 - また、既に一部がそうであるように、NISTEPの調査研究活動のなかに吸収されていく？
 - 欧州のRISIS（研究・イノベーション動向・政策等関連のデータの分散型インフラ）はモデルになる？
 - あるいはWikipediaのコンテンツにするといった方法もありえる？
- “辞書”は世界的な情報の標準化やオープンサイエンスにおけるデータ共有の一部となっていく？
 - 標準化の例：ROR (Research Organization Registry)
 - 機関名の標準化だけでなく、研究者名（個人名）の標準化も行われている
- AIの活用により、NISTEPの“辞書”は不要になる？
 - AIの活用はデータ整備の大きな助けになることは確か
 - しかし、原理的に、得られるデータの正確さには限界があるのではないかと
 - 逆に新たな情報のノイズが生じる可能性など
 - 正確性の確認の“最後の拠り所”としての“辞書”の必要性は残るのではないかと

26