

発表 4

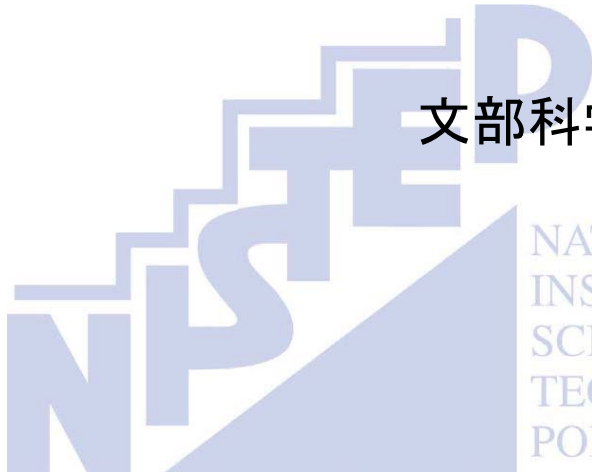
「データ・情報基盤の構築」

科学技術・学術基盤調査研究室長

富澤 宏之

発表4

データ・情報基盤の構築



文部科学省科学技術・学術政策研究所
科学技術・学術基盤調査研究室長
富澤宏之

NATIONAL
INSTITUTE OF
SCIENCE AND
TECHNOLOGY
POLICY

2014年12月16日

パート1

全体的状況

科学技術イノベーションにおける「政策のための科学」 「データ・情報基盤の構築」の全体的構想

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

事業の概要

1. 政策に資する基礎的データの整備
 - 既存の各種のデータ(論文・特許データや統計データ等)の活用を高度化するため、ミクロ・レベルでの連結情報等の整備
 - 必要に応じ、新たなデータの収集を実施
2. 政策プロセスへの実装のための情報の高度化
 - 基礎的データの整理・分析による政策情報の整備
 - 政策情報の可視化・集約化
 - 政策オプション作成への貢献
3. 幅広い利用者への各種データ・情報の提供
 - 行政資料、データの所在情報、各種調査データなどをwebで公開
 - 分かりやすく使いやすいデータツールの提供
4. 内外の研究との連携

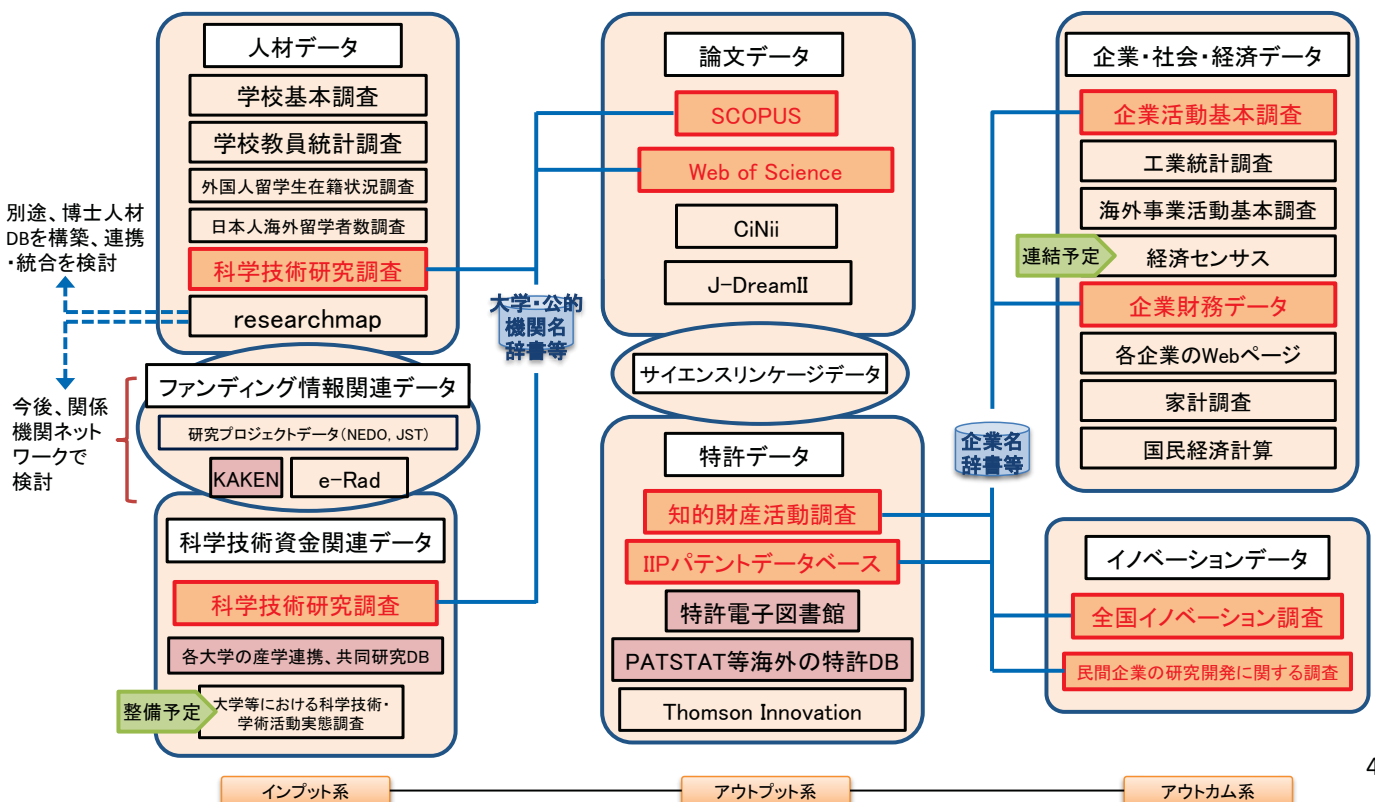
3

データ・情報基盤の進捗

- ニーズの高いデータのうち、**相互に連結**する意義・価値の高いデータや**政策の評価**につながりうるデータ等を重視して整備・連結対象とした。
- データを相互に連結するためのハブとなる大学・公的機関名辞書と企業名辞書の作成に重点を置いた。

凡例

- データ名 ⇒ 独自の構築データ
- データ名 ⇒ 主要な整備・連結対象
- データ名 ⇒ 部分的な整備・連結対象



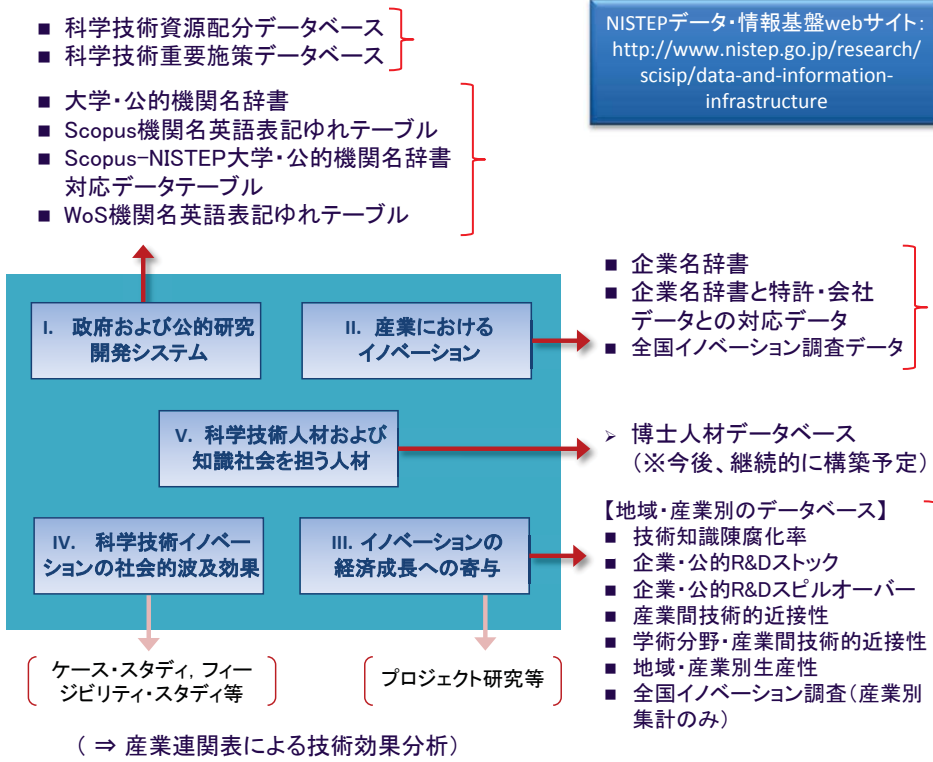
発表4-2

4

データ・情報基盤の全体的な構築状況

科学技術イノベーションに関する研究の基盤

政策立案のためのエビデンス提供ツール



- NISTEP定点調査検索
- NISTEP定点調査自由記述簡易検索用データベース
- NISTEP定点調査自由記述テキストマイニング用辞書
- 科学技術指標HTML版
- 科学論文の国際共著データの地図表示システム
- デルファイ調査検索システム

一般的なデータ・情報基盤

- NISTEPの全レポートの検索・提供システム(リポジトリ)
- 国内外のデータ・情報基盤へのリンク集

5

パート2

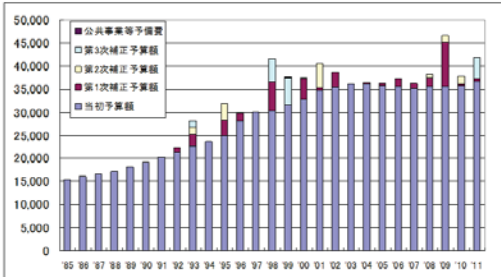
個別のデータ・情報基盤の状況 (主要コンテンツ)

(1) 科学技術資源配分・重要施策データベースの構築

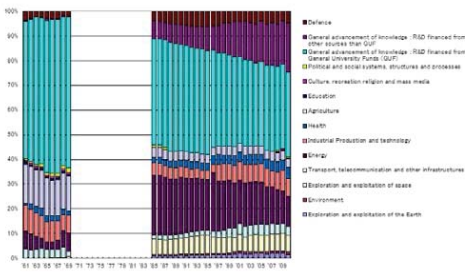
■ 資源配分データベースの構築

- 1970年代からの科学技術関係経費の総額及び配分を可能な限り分類整理(省庁、会計、用途、分野等)。

科学技術関係経費総額(本予算・補正予算別)



OECD社会・経済目的別政府研究開発予算の推移

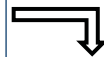


■ 重要施策データベースの構築

- 1950年代からの科学技術白書の記述を基に、関係府省の施策を33の施策群に分類し、それぞれの施策群毎に政策の系譜、主要施策を整理。

通史・概説(データベース作成者による)

地域における科学技術振興への取組の本格的な契機となったのは、科学技術会議11号答申(昭和59年11月)、第4次全国総合開発計画(昭和62年6月閣議決定)において地域における科学技術振興が位置づけられたことである。平成7年度には、「地域における科学技術活動の活性化に関する基本指針」(平成7年12月内閣総理大臣決定)が策定され、さらに科学技術基本法(平成7年11月)、科学技術基本計画(第1期)に地域における科学技術振興が明確に位置づけられたことから、より多くの関連施策が実施されるようになった。科学技術庁所管事業としては、地域研究開発促進拠点支援事業(通称RSP事業)、地域結集型共同研究事業などが開始された。併せて、自治体における科学技術振興への取組が活発化した。



主な関連施策リスト

No.	題名	策定機関	区分	出典	実施期間	実施主体	概要	説明
1	1984.479	3科学技術基本法	地域における科学技術の振興	事業・制度	平成9年版白書 3-2-6-2	地	科学技術庁	科学技術庁は、平成9年度から、地域における研究開発促進拠点を中核として、国立及び公設試験研究機関、大学、民間の研究機関等の研究コーディネーター職の充実を図るとともに、地域におけるニーズ、シーズの調査・育成、実証試験等をもとに、これらを通じて地域における科学技術の振興と新産業創出の促進を支援する同支援事業を実施。

(2) 公的研究機関に関するデータ整備

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

問題状況の提示

- 日本の論文数の低減
- 博士課程学生の減少
- 研究時間の減少
- 政府研究投資の伸び悩み(国際比較)

データ活用の深化

(マクロデータからマイクロデータ、インプット-アウトプットのデータ連結へ)

問題の背景・要因の分析へ

- 論文生産が躍進/低迷している大学はどこか
- 研究費の配分は論文生産にどう影響したか?
- 研究時間の減少は論文生産にどう影響したか?

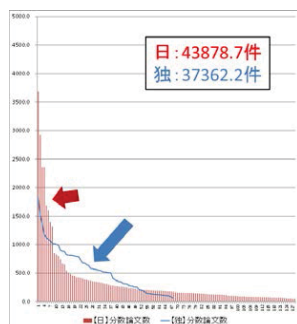
マクロデータ(国全体)

例: 国別の論文数

国名	1998年 - 2000年(平均)			2008年 - 2010年(平均)		
	論文数	シェア	世界ランク	論文数	シェア	世界ランク
米国	213,229	31.3	1	297,101	27.5	1
英国	82,662	9.2	2	120,156	11.1	2
日本	62,457	9.2	3	82,218	7.6	3
ドイツ	56,795	8.3	4	79,952	7.4	4
フランス	42,267	6.2	5	71,149	6.6	5
カナダ	28,918	4.2	6	58,261	5.4	6
イタリア	27,291	4.0	7	48,344	4.5	7
ロシア	24,560	3.6	8	47,373	4.4	8
中国	24,405	3.6	9	39,985	3.7	9
スペイン	20,006	2.9	10	39,555	3.7	10

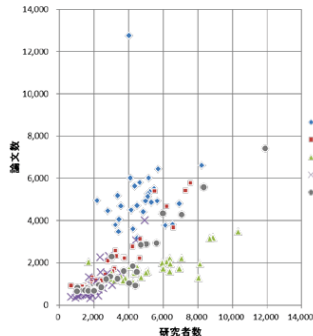
マイクロデータ(個別大学レベル等)

例: 大学別の論文数の分布



インプット-アウトプットのデータ連結

例: 大学別の研究者数と論文数



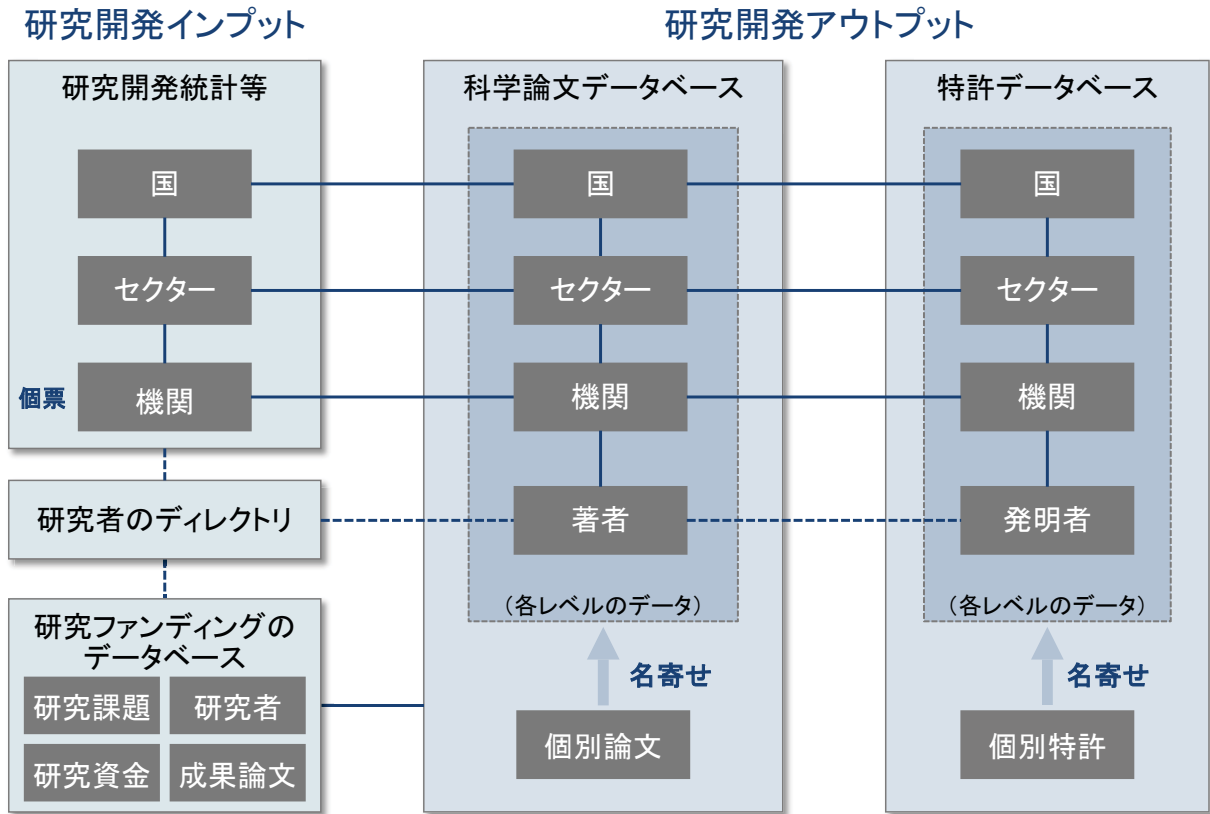
データ基盤の内容

- 機関レベルのマイクロデータの分析の基礎となる大学・公的研究機関名辞書
- 個別大学・公的研究機関の論文データの整備(機関名寄せ)
- 様々な統計データ等を個票レベルで相互連結するための情報
- 研究者個人レベルの論文データの作成手法の開発
- 日本と比較するための各国データ(研究者数、研究費)の収集・分析

想定されるデータ使用者

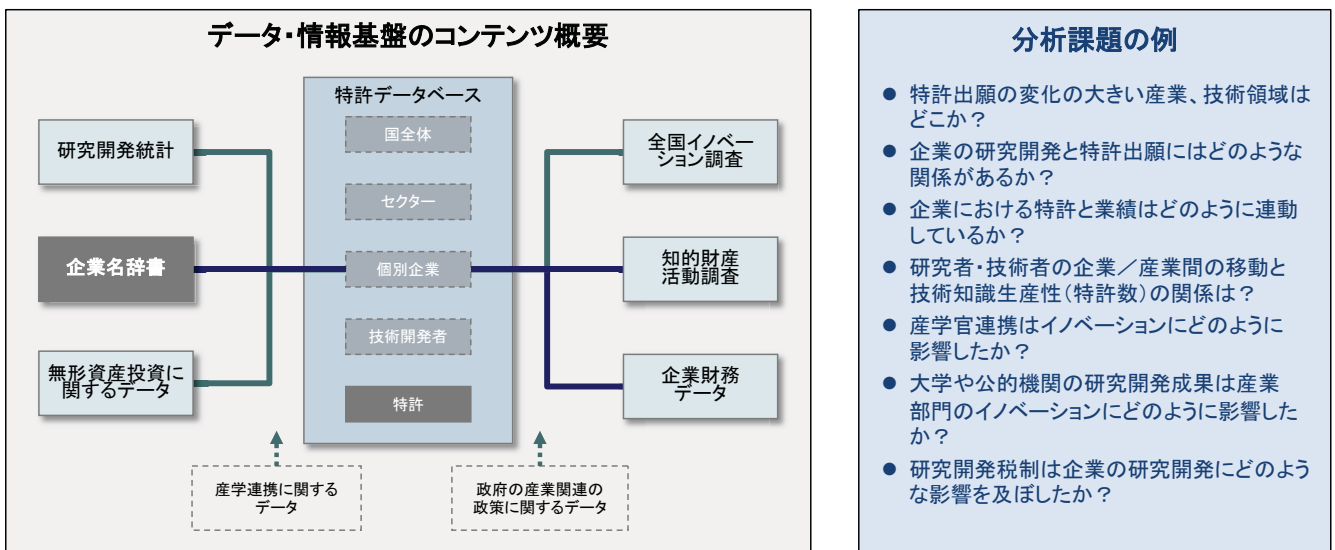
- 行政対応型の政策研究機関(NISTEPを含む)の専門家
- 科学論文の生産性分析に取り組む政策研究者・経済学者
- 科学技術知識の動態分析等に取り組むデータサイエンティスト
- 各大学における研究戦略策定の担当者(IR担当者やURA等)

日本の研究開発データの連結の概念モデル： 垂直方向(階層構造)と水平方向(異種データ間)の連結



9

(3) 産業の研究開発・イノベーションに関するデータ整備



データ整備：2014年度実績

- 企業名辞書の接続対象の Patent データベースの拡大
- 企業名辞書の拡張
- 発明者名寄せの高精度化

※ 企業名辞書更新版を2014年11月17日にweb公開

成果の活用を通じた検討：2013～2014年度

有識者ワーキンググループを設置し、整備成果データを活用した分析を実施し、発表会を開催した。また、整備データの過不足、利便性等についての評価を実施。

元橋一之 鈴木潤 大西宏一郎 西村陽一郎 中村健太 山内勇 長岡貞男

東京大学大学院工学研究科 教授(2013年度のみ)
政策研究大学院大学(GRIPS) 教授
大阪工業大学知的財産学部 講師
神奈川大学経済学部・大学院経済学研究科 准教授
神戸大学大学院経済学研究科・経済学部 准教授
経済産業研究所(RIETI) 研究員
一橋大学イノベーション研究センター 教授

10

NISTEP企業名辞書と各種の企業データとの関係・相違点

(●は連結対象としたもの)

データ	特徴
NISTEP企業名辞書	<ul style="list-style-type: none"> 各種データ(統計の個票データや特許データ等)を企業レベルで連結するためのハブ機能(コンコーダンステーブル) 特許データにおける企業名の表記揺れを解消(名寄せ) 企業の吸収・合併・分社などの変遷情報を定量分析に活用できる形で収録
企業情報データ (●東洋経済データ、帝国データバンク、東京商工リサーチ、日経データ等)	<ul style="list-style-type: none"> 企業概要、業績、財務など企業の基本的なデータを収録 研究開発やイノベーションに関する情報はほとんど無い コスト効果性の高い東洋経済データについてコンコーダンステーブルを作成
経済センサス [総務省統計局]	<ul style="list-style-type: none"> 事業所及び企業の産業、従業者規模等の基本的構造を全国的及び地域別に明らかにすることを目的 原則として全産業をカバーする一次統計
●企業活動基本調査 [経済産業省]	<ul style="list-style-type: none"> 企業を単位とした事業活動の多角化の実態等についてのデータ 研究開発・能力開発、特許権等の所有・使用状況・技術取引等についての情報を含む
●全国イノベーション調査 [NISTEP]	<ul style="list-style-type: none"> 企業のイノベーション活動についての詳細な調査 国際的な基準に準拠(主要項目についての国際比較が可能)
●民間企業の研究活動に関する調査[NISTEP]	<ul style="list-style-type: none"> 企業の研究活動についての詳細な調査 資本金1億円以上の研究開発活動実施企業(約3400社;回収率)を対象

11

パート3

データ・情報基盤 関係機関ネットワークの概要

データ・情報基盤 関係機関ネットワーク

【目的】: 研究開発ファンディング実施機関および関連データ保有機関のネットワーク構築

【2013年度実績】: 3回開催(各機関のデータ保有・整備状況等について情報を共有)

【2014年度実績】: 3回開催(第5期科学技術基本計画策定プロセスへの提案につき討議)

○参加機関:

- 国立情報学研究所(NII)
- (独)科学技術振興機構(JST)
- (独)大学評価・学位授与機構(NIAD)
- (独)日本学術振興会(JSPS)
- (独)経済産業研究所(RIETI)
- (独)新エネルギー・産業技術総合開発機構(NEDO)
- (独)情報通信研究機構(NICT)
- (独)農業・食品産業技術総合研究機構(NARO)
- (独)医薬基盤研究所(NIBIO) [※H25年度末まで]

○主要な検討ポイント

- 研究開発ファンディング情報の整備・標準化の可能性(統一課題番号の導入等)
- 科学技術基本計画レビューへの対応の可能性

関係機関ネットワーク参加機関による競争的資金配分額は、我が国の競争的資金の約9割を占める。

平成25年度では競争的資金予算額(約4,090億円)の約89%(約3,630億円)。

出典: 平成26年度版科学技術白書の第2-5-2表/競争的資金総括表より科学技術・学術政策研究所が集計。
科研費、国家課題対応型研究開発推進事業等、本省との共同所管の事業も含む。

13

次期科学技術基本計画への提言(案)

[データ・情報基盤関係機関ネットワーク会合における共通認識ペーパー(H26.09)]

- ◆ 科学技術予算の資源配分の効果や研究開発ファンディングの状況・成果を分析して、科学技術動向を俯瞰するとともに、研究構想の検討や政策策定に生かすためには、研究開発予算の配分・運用及び成果についてのデータ・情報基盤について、事業の特性に配慮しつつ、必要に応じ継続的・体系的に整備することが重要。
- ◆ 更に、整備されるデータのうち、研究者、研究機関、研究プロジェクトのデータ・情報については、世界の動向を見つつ、標準化、ID化及びインプット・アウトプットの関連付け等を体系的に行っていく必要あり。
- ◆ こうしたデータ・情報基盤の構築を効率的に進めるため、自動的(自律的)に情報が蓄積・共有される仕組みについて、研究者コミュニティの意見等も踏まえ検討していくことが重要。

14

Web of Scienceの「謝辞情報」における 主要な資金配分機関等の表記状況

- 23万篇の日本論文(2008~2012年)に謝辞情報が有り、資金配分機関等名が54万回出現
- そのうち、日本の資金配分機関等は32.6万回出現
- 「文部科学省」の出現回数は約11万回、ただし約1.2万の表記バリエーションが存在

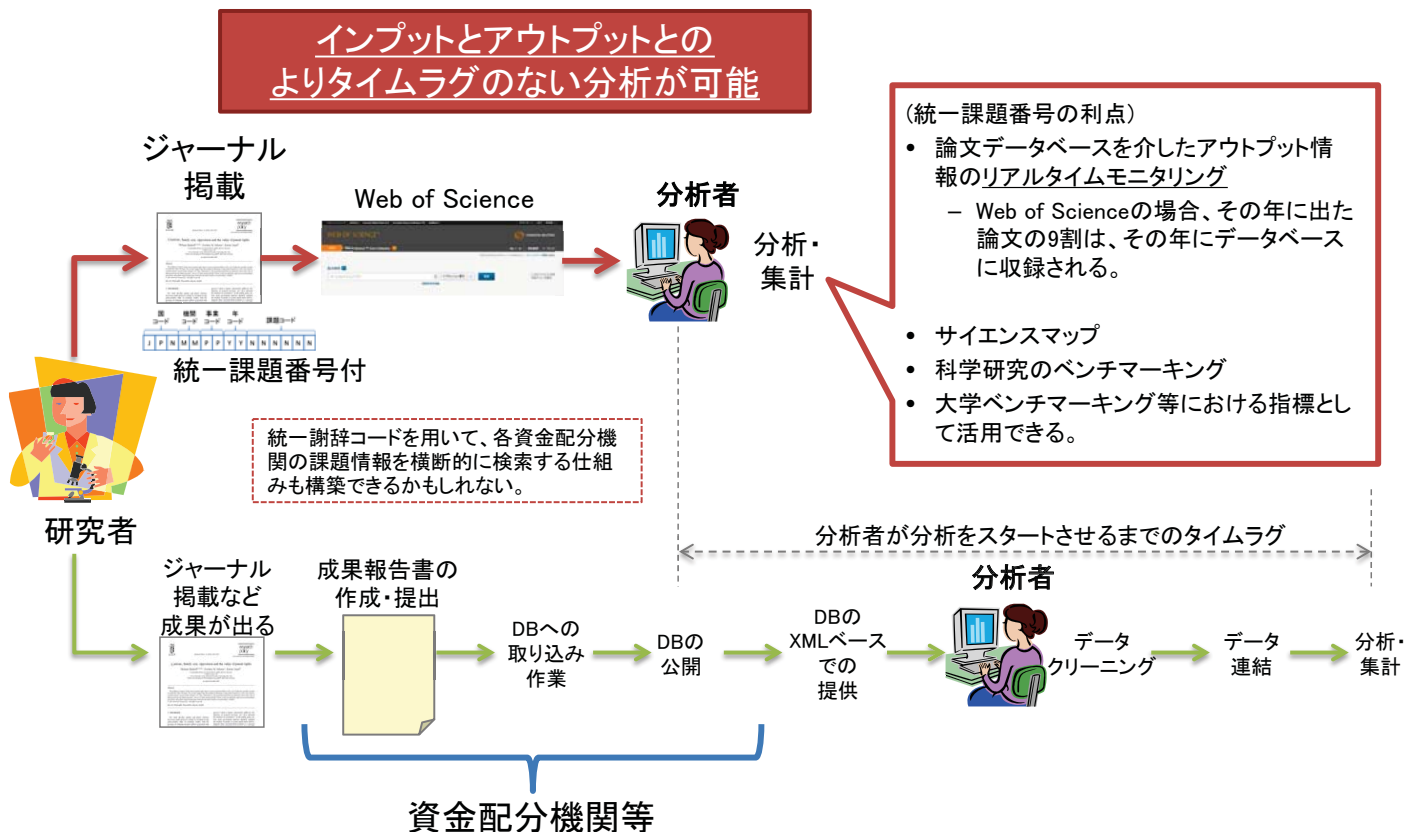
機関名	表記バリエーション	出現頻度	出現頻度 (%)	累積 (%)
文部科学省	11,519	113,886	34.9	34.9
独立行政法人日本学術振興会	5,095	73,386	22.5	57.4
厚生労働省	2,257	18,179	5.6	62.9
独立行政法人科学技術振興機構	3,318	17,372	5.3	68.3
独立行政法人新エネルギー・産業技術総合開発機構	872	8,180	2.5	70.8
公益財団法人武田科学振興財団	97	3,458	1.1	71.8
農林水産省	330	2,619	0.8	72.6
環境省	355	2,523	0.8	73.4
公益財団法人上原記念生命科学財団	60	1,956	0.6	74.0
独立行政法人農業・食品産業技術総合研究機構	397	1,943	0.6	74.6

上記10機関で、名寄せを実施した謝辞情報(国内機関)の約75%をカバー

(出典) 論文の謝辞情報を用いたファンディング情報把握に向けて一謝辞情報の実態把握とそれを踏まえた将来的な方向性の提案一, 科学技術・学術政策研究所, NISTEP NOTE No. 13 (2014年12月)

15

統一課題番号を用いた分析の流れ



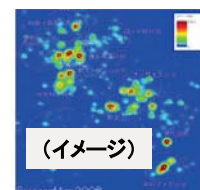
今後の展望

“エビデンスに立脚した政策形成”のための NISTEPの調査研究の展望

- ① 未来を拓く戦略R&D領域、スター研究者のプロジェクト・マッピングに向けて(マイクロレベル)

◆ 新興・注目領域・イノベーションフロントの可視化・同定

▶ サイエンスマップ／リンクマイニング



- ② 人・研究資金・創出成果の紐付けによる、時・空間軸での施策・プログラムのアウトカム分析・可視化に向けて(マイクロ／メソレベル)

◆ 研究開発リソースと成果の時空間分析システムの構築

▶ [博士人材DB<ヒト>](#)／[関係機関ネットワーク<fund/成果>](#)の統合的活用

→(例)個別アクター(大学・機関／地域／分野...)ごとの「研究力」・生産性／
経済・社会インパクトの分析、国際的な人材流動性