

大学・公的機関における研究開発に関するデータ整備について

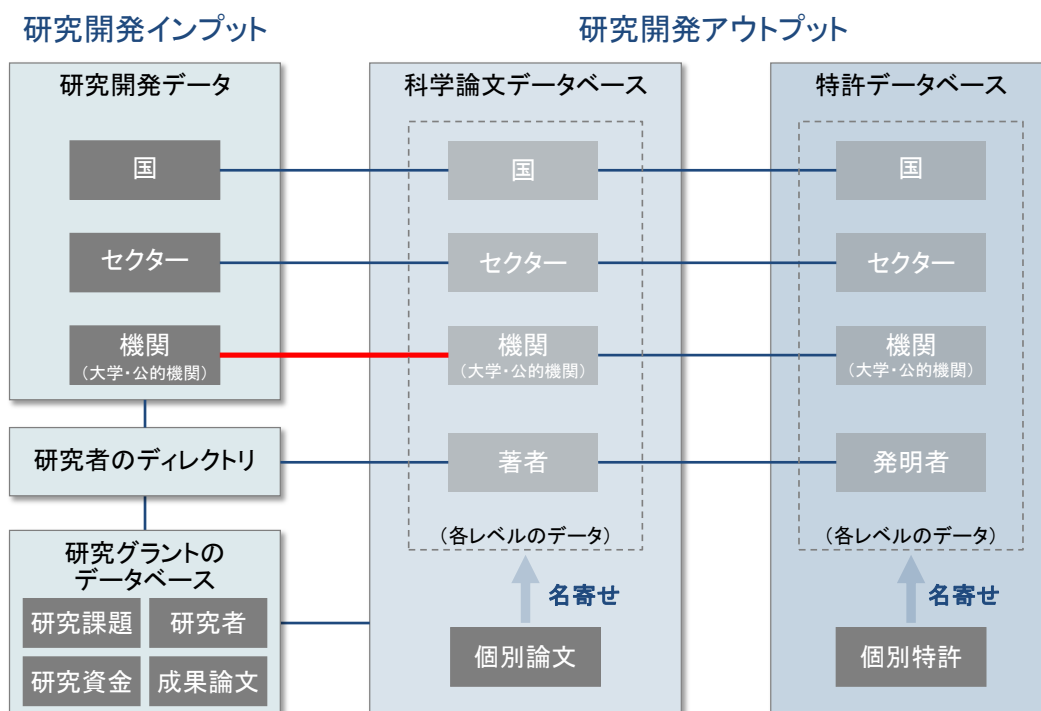
(2017年5月1日)

1. 全体構想

科学技術・学術政策研究所(NISTEP)では、2011年度から、文部科学省の「科学技術イノベーションにおける“政策のための科学”推進事業」の一環として「大学・公的機関における研究開発に関するデータ整備」を進めています。これは、政府予算で実施されている研究開発の実態やパフォーマンスの把握・分析・評価を国、セクター、個別機関・研究プロジェクトなどの各レベルで行うための基礎として、大学や公的機関の研究開発のインプットとアウトプットに関するデータの整備を行うものです。特に、研究開発インプットとアウトプットのデータをマイクロレベルでリンクさせ、政府の研究開発投資の成果や研究開発システムに与えた影響を定量的・構造的に分析できるようにすることを目指しています。

その全体的な構想は、図1に示すようなモデルで示されます。このモデルに示された分析が可能になるよう、科学論文データベース等の名寄せや各データを連結するための共通基盤的なデータを中心に、データ整備を進めています。

図1 大学・公的機関における研究開発に関するデータ整備の概念モデル



2. 具体的なデータ整備内容

この事業を開始した 2011 年度より以下のデータ整備を行っており、データの調整等が済み次第、順次公開を進めています。

(1) NISTEP 大学・公的機関名辞書（以下、「機関名辞書」と略）の整備

研究活動を行っている国内の約 12,300 の機関とそれらの主要な下部組織(約 2,500) について、和英の名称、所属するセクター等の基本情報を収録しています。現存しない機関については継承機関を示しています。32 の大学と一部の公的機関については、下部組織の情報を網羅的に収録しています。大学、国の機関、特殊法人・独立行政法人等を中心に整備を行っていますが、地方公共団体の機関、民間企業、非営利団体もできるだけ含めています。民間企業については、「NISTEP 企業名辞書」（「産業における研究開発・イノベーションに関するデータ」のページをご覧ください）とのリンクを採っています。

機関名辞書は、個別機関レベルの分析のための基礎情報源として使用することができます。

(2) 機関名の表記ゆれデータの収集とデータベース中の著者所属機関の同定

科学論文の計量的分析に用いられる Web of Science Core Collection データベース及び Scopus データベースを対象として、これらに採録された 1996 年以降の論文データに含まれる日本の著者所属機関データの同定を行い、機関名辞書と対応付けました。これらの対応付けデータは、一部を公開するとともに、NISTEP で作成している大学ベンチマーキングやサイエンスマップ等でも活用されています。

また、種々の情報源から得られた機関（下部組織を含む）の別名、略称等を加え、多数の機関表記ゆれデータを蓄積・公表しています。

(3) 各種情報源の機関情報のリンク

機関名辞書は、以下のデータベース等に含まれる大学・公的機関の大部分を収録しており、これらのデータベース等を個別機関レベルで相互に連結する際には、機関名辞書を仲介させることにより効率的にデータを作成することができます(図 2 参照)。

(a) 科学論文書誌データベース

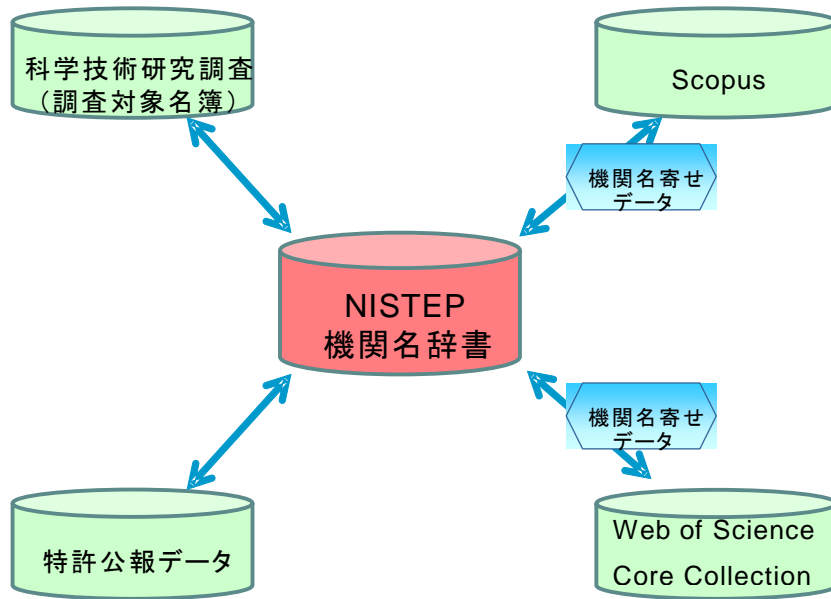
- ・ Scopus (Elsevier 社)
- ・ Web of Science Core Collection (Clarivate Analytics(旧トムソン・ロイター IP&Science))

(b) 機関レベルの研究開発データ（「科学技術研究調査」等）

(c) 特許公報データベース

詳細については、科学技術・学術政策研究所(NISTEP 大学・公的機関名辞書担当 E-mail: data-infra@nistep.go.jp)にお問い合わせください。また、これらのデータベース等は、科学技術・学術政策研究所からは提供していませんので、別途、入手が必要です。

図2 各種データベース等を機関レベルで連結する際の機関名辞書の役割



3. データの公開について

「大学・公的機関における研究開発に関するデータ整備」事業で得られた成果は、できる限り公開し、利用者の研究・調査等の目的に活用していただくとともに、ご意見・ご要望をいただいてデータや事業の改善を図りたいと考えています。

2で述べたデータセットのうち、現在、以下の3種のデータをwebサイトにおいて公開しています。

- (a) NISTEP 大学・公的機関名辞書(ver.2016.1)：2(1)で説明したとおりです。
- (b) 大学・公的機関名英語表記ゆれテーブル：2(2)で述べた整備により得られたもので以下の2つがあります。
 - ・大学・公的機関名英語表記ゆれテーブル(ver.2016.1)：機関名辞書に含まれる機関の正式名や別名と、Web of Science Core Collection から取り出された表記ゆれデータを統合したものです。下部組織のデータ、大学、公的機関以外の機関のデータも含んでいます。
 - ・大学・公的機関名英語表記ゆれテーブル(Scopus版)(ver.2013.1)：Scopus から大学、公的機関の表記ゆれデータを取り出したものです。
- (c) NISTEP 大学・公的機関名辞書と論文データベースとの対応テーブル：2(3)で述べた整備により得られたもので以下の2つがあります。
 - ・Scopus-NISTEP 大学・公的機関名辞書対応テーブル(ver.2013.1)：1996～2010年のScopusに現れる大学・公的機関の機関名データ(延べ220万件)を、機関名辞書の収録機関に対応させたものです。
 - ・WoSCC-NISTEP 大学・公的機関名辞書対応テーブル(ver.2017.1)：1998～2015年のWeb of Science Core Collectionに現れる機関の機関名データ(延べ273万件)を、機関

名辞書の収録機関に対応させたものです。

これらの内容及び使用方法については、それぞれの利用マニュアル等をお読み下さい。