

理化学研究所業績登録システム (RARS) の開発

舛本 現¹⁾, 関 進¹⁾, 検崎 博生¹⁾

1) 理化学研究所 情報統合本部

masumoto@riken.jp, susumu.seki@riken.jp, hiroo.kenzaki@riken.jp

Development of RIKEN Research Achievement Registration System(RARS)

Gen Masumoto¹⁾, Susumu Seki¹⁾, Hiroo Kenzaki¹⁾

1) RIKEN Information R&D and Strategy Headquarters

概要

理化学研究所(以下、理研という)では、研究業績の登録・管理のために業績登録システム(Research Achievement Registration System; RARS)を開発し、2021年7月より運用を開始した。従来の理研の研究業績の登録・管理方法は、登録作業の負担が大きく、また登録されたデータも再利用するのが難しいという問題点があった。そこで、外部の研究業績データベース(以下、DBという)による入力支援や理研の人事DBと連携した著者の所属同定といった機能をもった新しい研究業績の登録・管理システムRARSを開発・導入した。本稿では、RARSの導入の経緯、システムの概要・特徴・実装、利用状況、理研が推進するオープンサイエンスとの関係について紹介し、現状の課題と今後の展開について報告する。

1 はじめに

理研では新しく研究業績登録システム(Research Achievement Registration System, RARS)を開発し、2021年7月から運用を開始した。本稿では、RARSについて、導入の経緯、システムの概要と特徴、実装の詳細、利用状況、理研が推進しているオープンサイエンスとの関係、さらに現状の課題と今後の展開について報告する。

理研における研究業績登録は、2013年度以降は所定のExcelファイルに必要事項を記入して図書館にメール添付で提出するという方式でおこなわれてきた。業績登録のための入力項目は、著者名・タイトル・発表雑誌名といった書誌情報に加えて、理研の内部規程で定められている研究不正防止のチェックシートなどの合計23~37項目である(業績タイプによって項目数は異なる)。提出された記録は業績登録の担当者によって内容をチェックされ、所内ウェブサイト上で公開されてきた。

これまでの登録方式では、ほぼすべての項目が自由記述欄への手動による入力なので研究者の負担が大きく、また業績登録担当者にとっても記入セルのずれやタイプミスなどの細かいチェックに多くの時間をとられて負担が大きかった。さらに手動入力なので所属

組織や発表媒体などの名称についても表記ゆれが大きく、研究業績の集計などでデータを再利用する際の障害となっていた。また、Web of Scienceなどの外部の研究業績DBで理研の業績を検索した結果と比較すると、届出が提出されている件数が少ないことも問題となっていた。このような状況を改善して、研究者と事務担当者の作業負担を軽減し、かつ理研に所属する研究者の正確な業績を把握することが求められていた。

また、理研は2019年度からオープンサイエンスの推進のための基盤技術の研究開発を目的としたプロジェクトを開始しており、研究成果の収集・蓄積・公開・共有のためのさまざまな基盤の整備を進めている。このプロジェクトに対して本システムは業績の管理基盤という立場から貢献するものであり、同じように研究情報の管理基盤である研究データ管理システムや研究文献公開システムなどと共に理研オープンサイエンス体系を構成している。

2 本システムの概要

本システムにおける業績登録の基本的な流れを図1に示す。利用者はウェブインターフェースで研究業績の登録をおこなう。業績登録時には外部の研究業績DBの情報を用いて業績の書誌情報について入力支援を利用することができる。また、理研所属著者につい

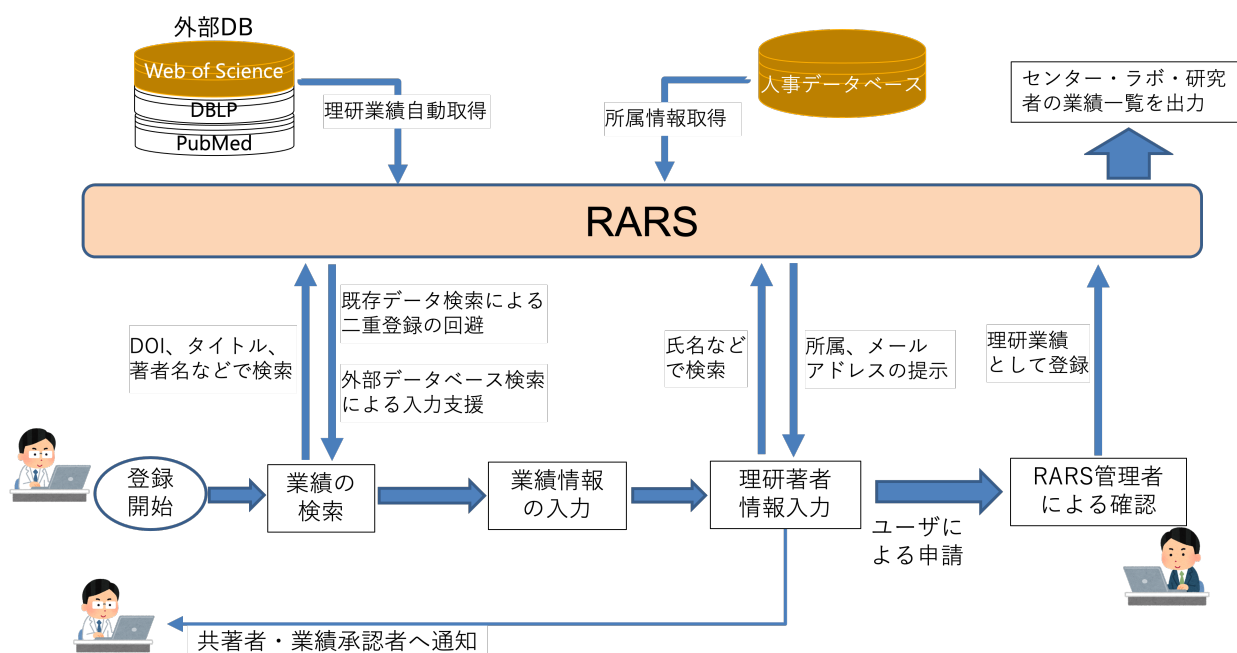


図1 業績登録の流れ

ては理研の人事 DB の所属情報を用いて所属の同定をおこなう。

本システムは、理研業績管理モジュール、外部情報モジュール、研究者情報モジュールの3つのモジュールからなる(図2)。理研業績管理モジュールは理研の業績を保存する理研業績 DB に対して登録・修正・出力などの操作をおこなうモジュールであり、本システムの基幹をなす。外部情報モジュールは外部の研究業績 DB から理研の業績を検索するためのモジュールであり、このモジュールから提供される書誌情報を用いて業績入力時に入力支援をおこなう。研究者情報モジュールは理研の人事 DB の所属情報と連携して研究者の年度ごとの理研内での所属情報を管理するためのモジュールであり、前述した理研著者の所属同定に利用する。

3 本システムの特徴

本システムの主な特徴は以下の通りである。

1. 外部の研究業績 DB の情報を用いた入力支援

本システム導入の目的のひとつは利用者の入力負担軽減である。利用者は、まず登録しようとしている業績について本システム上で検索をおこなう。この時に理研業績 DB だけでなく外部の研究業績 DB についても検索がおこなわれて結果が表示される(図3)。検索した結果、理研業績 DB に登録されていない場合は、新規入力画面に移行して必要な情報を入力する。入力

画面では、外部の研究業績 DB に登録されている業績については、その情報を用いて多くの入力項目が補完されるので入力者の負担は軽減されている。

外部の研究業績 DB の情報については、定期的に外部 DB で理研業績を検索してローカルの DB (外部情報 DB) に保存しておく方法と、利用者が本システムで検索するたびに API で外部 DB に直接検索して取得する方法を想定しており、外部の業績 DB ごとにどちらかの方法で情報を提供する。現状では利用可能な外部の研究業績 DB は Web of Science のみであり、この DB については定期的な検索結果をローカルの外部情報 DB に保存する方法で運用している。

2. 人事 DB と連携した理研著者の所属同定

理研のセンター・研究室などのさまざまな組織レベルでの業績を把握するためには著者の所属を正確に記入・登録する必要がある。これまでの業績登録でも著者の所属の記入欄は存在したが、略称を用いるかどうか、また日英のどちらで入力するかなどが入力者によって異なるために表記ゆれが大きく、業績を集計する際の問題となっていた。

そこで、本システムでは理研の人事 DB と連携して RARS 内の研究者 DB に理研の研究者の所属情報を保存し、理研所属の著者の所属(センター・研究室など)の候補を表示して入力者が選択・確定する機能を実装した。所属情報については、その研究者の過去数年間の所属も表示されるので、論文投稿時と業績登

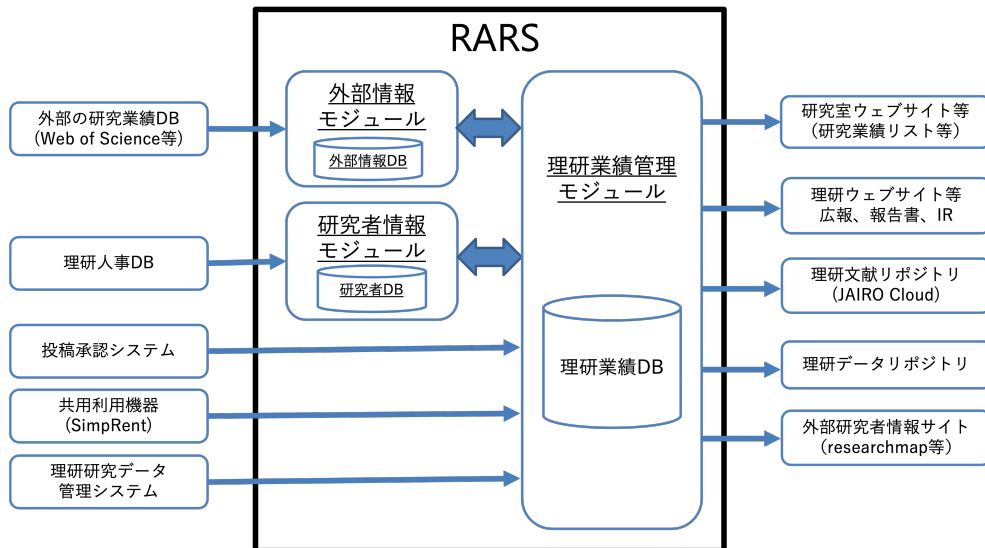


図2 RARS 概要図

録時で所属の名称が異なる場合にも対応している。このことによって理研業績 DB には著者の所属が統一された表記で登録されることになり、今後は業績の集計や業績リストの作成などの業務の負担が軽減されることが見込まれる。

3. 理研内の各種情報システムとの連携

理研には、理研内の研究者が利用することができる共同利用機器や研究技術支援があり、それらの支援業務を担当するチームも存在している。こういった研究支援がそれぞれの業績にどのように貢献しているかを把握することは、支援チームにとってだけでなく研究所においても重要である。これまでも業績登録時に研究支援の利用に関して記載する欄はあったが、自由記述式であったことあり、集計には多くの労力を必要としていた。本システムではウェブ上の記入欄の書式を統一することで集計を容易にした。

また、本システムは研究データ管理システムや文献管理システムと共に理研オープンサイエンス体系の一部として研究情報の管理基盤を担っている。そこで、理研業績 DB に研究データリポジトリや文献リポジトリに関係する項目を用意し、それらのシステムと連携することができるようになっている。

4. 入力・編集権限の適切な管理と通知機能

理研では、研究室のアシスタント（秘書）が研究室主宰者（PI）や研究員の業績をまとめて入力することがしばしばある。そこで本システムでは代理人を設定する機能を実装し、研究者が自身の代理人に研究室のアシスタントや事務担当者を設定することで業績登録業務を彼らに依頼することができるようにした。代理

人情報も研究者 DB によって管理されている。

また、業績の登録に際しては、たとえば理研内の共同利用機器の利用情報のように一部の共著者のみが把握している情報の入力が必要となることもあるので、編集中に一時保存して共著者に他の項目を入力してもらうという機能も必要である。そこで本システムでは前述した理研著者の所属同定機能で同定された著者については、当該業績に対する編集権限を付与し、追加で入力・編集・保存・登録申請をおこなえるようにした。

本システムでは、業績データの一時保存・登録申請・登録の各時点で共著者と研究室の PI にメールで通知を送る機能がある。この機能によって共著者は登録内容をチェックすることができ、また二重登録の防止にも役立っている。

4 実装に関する詳細

4.1 データベース

メインの DB である理研業績 DB は ElasticSearch を利用している。また、外部の研究業績 DB から理研関係の業績を取り込んで保存する外部情報 DB は MongoDB、人事 DB と連携して必要な情報を保存する研究者 DB については MariaDB を利用している。

4.2 認証とアクセス制御

利用者の認証については、理研の全職員を対象とした Shibboleth 認証を用いた認証システムによっておこなわれている。また、本システムは理研外部にも公開されており、従来在所内の手続きサイトのように VPN を用いなくてもアクセスすることができる。よって客



図 3 RARS 画面 (左: 検索結果画面、右: 入力画面)

員研究員などの理研と直接雇用関係にない研究者も自分自身で業績を登録することが可能である。

4.3 ハードウェア環境

本システムは、理研の和光キャンパスに設置されている共同利用計算機 HOKUSAI の一部であるデータ科学基盤 HOKUSAI Sailing Ship(HSS) 上に仮想マシン (VM) を用いて構築されている。

5 導入後の利用状況

2021 年 7 月 29 日に運用を開始して以降の約 2 ヶ月間の利用実績を表 1 に示す。カッコ内の数字は入力された業績の中で Web of Science による入力支援を利用して入力された件数である。昨年度までの 1 年間あたりの業績の登録数は、論文発表が約 2000 件、口頭発表が約 5000 件なので、この数字は 2 ヶ月間の登録件数としては少ない。これは理研内で本サービスの認知がまだ広がっていないことが原因であると考えられるので、今後も引き続き広く周知していくことが必要である。

表 1 登録実績 (2021 年 9 月 27 日現在)

業績タイプ	登録数 (WoS 入力支援利用数)
論文発表	247 (98)
口頭発表	258 (0)
書籍	9 (0)

また、「はじめに」でも述べたように昨年度までの業績登録の件数は実際の発表数に対して過小であり、本システムでは入力支援機能によって負担が減ることを考えると登録される業績数の増加が予想される。昨年度までも業績登録が年度末に集中し、図書館の成果発表担当者の負担が大きかったが、今年度は登録業績数が増えるとさらに年度末に登録が集中することになってしまうので、利用者に対して早めの業績登録を促して改善していきたい。

Web of Science による入力支援については、論文発表でも入力支援を利用していない業績が半数以上を占めている。これらの業績は和文雑誌や Web of Science には未収録の欧文雑誌に掲載された業績であると考えられるが、中には論文が発行されてから Web of Science に登録される前に本システム上で登録をしたために入力支援を利用していない業績も一定数含まれている。

6 理研オープンサイエンス体系

現在、理研ではオープンサイエンスを推進しており、そのためにさまざまなプロジェクトや基盤開発が進められている。その中で本システムは開発が進められている理研オープンサイエンス体系の一部を担っている。図 4 に理研オープンサイエンス体系の概要を示す。

本システムは理研オープンサイエンス体系の中で業

績登録・管理を担っており、研究データ管理基盤である理研研究情報管理システム (RIKEN Research Data & copyrighted-work Management System, R2DMS) や文献管理システム (JAIRO Cloud の利用を予定) との連携を強く意識している。そのため本システムでは業績登録の際に研究データに付与される識別子などをする項目や、文献リポジトリへの登録に必要な追加の書誌情報を入力することが可能になっている。

7 現状の課題と今後の展開

現状では、本システムはこれまでの Excel ファイルによる業績登録をウェブインターフェースによるに置き換えただけという点が多く、研究者と事務担当者が必要とするニーズに対してはまだ大幅に機能が不足している。そのことを踏まえて今後も開発を継続していく予定である。今年度開発を予定している主要な点は以下である。

1. 過去登録業績のインポート機能

研究者や所内の各種組織 (センター・研究室など) が業績リストを作成する際には過去の業績も必要となるので、これまでに登録された業績についても本システムに取り込んで統一して管理したいという要望が多く寄せられている。ただし、過去の業績は Excel への自由記述方式なので各項目のフォーマットも統一されず、本システムとは登録項目が異なる部分もあるのでスクリプトなどでデータを変換して自動でインポートすることはできない。

そこで、メインの理研業績 DB とは別に過去業績専

用の DB を用意し、過去に登録された業績についてはそちらの DB に保存することにした。過去業績を理研業績 DB に登録する際には、外部情報モジュールと同じような仕組みで過去業績 DB のデータを用いてを支援し、利用者自身に登録してもらう機能を実装する予定である。

2. 主要業績設定機能

研究業績の一覧を研究者個人のウェブサイトや研究室のウェブサイトに掲載する場合にすべての業績を載せるのではなく、主要な業績を研究者自身によって選択して掲載することもよくある。そこで本システムでも主要業績を設定する機能を追加する予定である。

この機能を実装する際に重要な点は、その論文が「誰にとっての」主要業績であるかを柔軟に指定できることである。研究者個人レベルだけでなく、研究室レベル、さらには部門やセンターといったより大きな組織のレベルに対しても主要業績を設定できるようにする必要があり、そのような設定に対応した主要業績設定機能を追加する予定である。

3. 外部 DB の追加

現状では外部の研究業績 DB として利用できるのは Web of Science のみだが、今後は利用可能な外部 DB を増やしていく予定である。特に、人工知能やスーパーコンピュータなどの情報系の研究分野では雑誌論文だけではなく国際会議での発表が業績として重視されているので、これらの分野の関係者からは会議録論文の DB を追加することを強く要望されている。そこで情報系の国際会議の情報を多く収録している DBLP

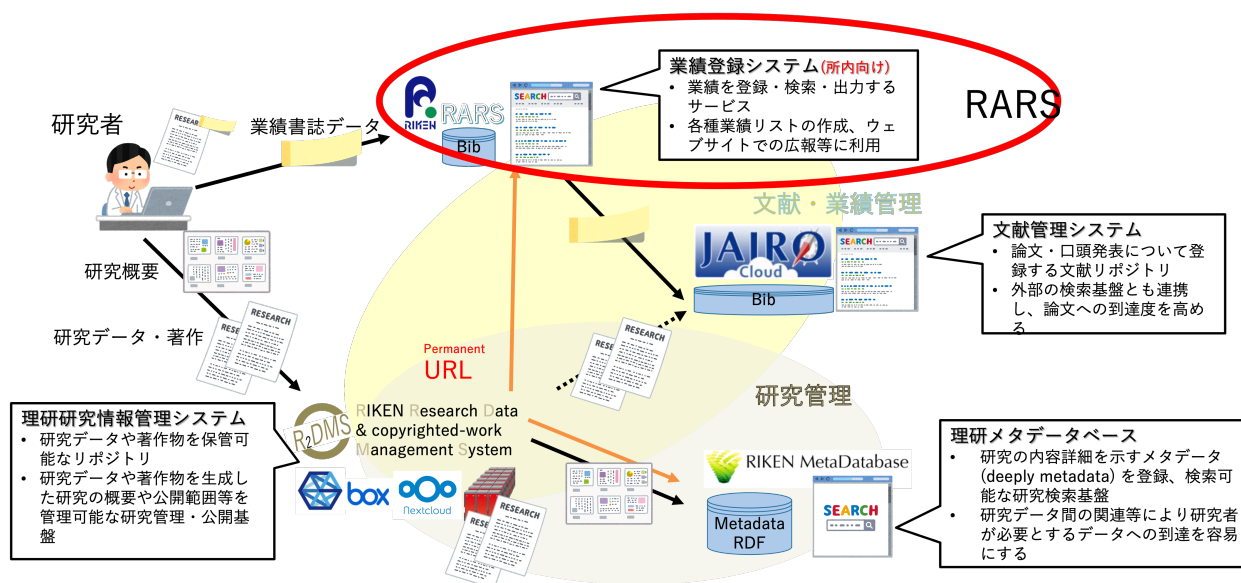


図 4 理研オープンサイエンス体系

を外部情報 DB として利用できるようにすることを予定している。また、医学系や生命科学系からの要望が多い PubMed も利用できるようにする予定である。

4. 理研内の情報システムとの API による連携

「本システムの特徴」で述べたように、本システムでは共同利用機器の利用情報や研究データリポジトリ、文献リポジトリなどの理研内の情報システムに関連した情報を入力する項目がある。現状ではそれらのほとんどは手動での入力が必要であるが、将来的にはこれらの情報を情報システム間の API 連携でやりとりすることで入力の手間を軽減することを計画している。今年度は共同利用機器の予約システムの SimpRent との API 連携を予定している。

5. researchmap との連携

科学技術振興機構 (JST) が運用する業績 DB の researchmap が科研費の申請時に参照されるようになったこともあり、研究者にとって researchmap 上の業績リストを正確かつ最新に保つことは必須となっている。多くの大学などでも業績管理システムと researchmap をなんらかのかたちで連携させることで研究者の業績管理の負担の軽減をはかっている。本システムでも今年度の開発で本システムに登録した業績データを API 連携などによって researchmap に登録できる機能を追加する予定である。

8 まとめ

本稿では、理研における新しい研究業績登録システムの導入経緯、システムの概要と特徴、システム構成、利用状況、理研のオープンサイエンスとの関係・現状の課題と今後の展望について述べた。今後も理研の研究者、事務担当者、および関連各部署の意見を聞きながら、引き続き本システムの改善と安定した運用をおこなっていく予定である。