

第14回 理化学研究所 バイオリソース研究センター 情報検討委員会

(2022年12月27日開催)

評価・助言

統合情報開発室 (室長:柁屋 啓志)

1. 技術開発あるいは研究開発の実績

(1) 前回 (2018年度) のレビュー委員会からの主な指摘事項への対応は充分か

- 前回の検討委員会で作された主な指摘事項への対応は充分である。
- 地道な成果を情報発信して各リソースの利用につなげる努力など今後の改善につながる点が多くあり、見守ってきたい。
- 委員の意見を適切にカテゴライズした上で、丁寧に対応している。研究インフラ整備や、セキュリティ、BCP対策などに、一部仕掛りの事項等もあるように見受けられるが、継続的な対応に期待する。
- 新たな HP コンテンツの開始、twitter の検討など理研 BRC 全体の広報につながる事にも取り組んでおり、実際に SNS など理研 BRC の情報が流れてくるようになったという実感がある。AI プラットホームの開発に実際に経費をあてた点は、指摘事項への取り組みとしてだけでなく、将来的な研究開発への取り組みとして評価できる。

(2) 社会や国内外の研究者コミュニティへの貢献度の観点から、これまでの技術開発あるいは研究開発の実績は世界の主要なバイオリソースセンターの水準に達しているか。

- Web サイト利用数、検案件数ともに世界の主要センターの水準に達している。特にリソース RDF データの整備については他に類をみないもので、今後もこの分野を先導する立場である。RDF による情報統合の利点をさらに生かすためにも、他機関との連携によるユーザの利便性向上も目指してほしい。
- バイオリソース事業はオープンサイエンスに大きく貢献しており、日本の冠たるセンターとして世界に引けを取らない。今回はアクセス元分析が省略されていた模様だが、変わらず世界各国に寄与しているものと理解した。
- マウス、植物、細胞、遺伝子、微生物のリソースを 1 つのセンターで扱う世界をリードするリソースセンターであり、統合情報サービスの充実により多く活用され、世界の主要なバイオリソースセンターの水準に達している。

(3) これまでの活動は、理研 BRC の第4期中長期計画 (2018年度~2024年度) に沿って適切か

- 理研 BRC の第4期中長期計画 (2018年度~2024年度) に沿って適切である。理研全体のデータの統合をリードしながら、バイオリソース関連のメタデータ統合、これに伴うホームページ公開コンテンツの充実と発信に取り組み、さらに、新型コロナウイルス等の感染症研究を加速するため、関連するバイオリソースの情報整備・発信に注力しており、高く評価できる。
- 計画に沿って着実に進めている一方で、疾患モデルマウスや難病関係データとの連携など、より従来のリソースに関連する充実化も含めた進捗があると感じた。
- 疾患から様々な生物種までメタデータ統合の指導的な立場を務め、各分野での成果につながっている。セキュリティを担保しつつ、理研全体の方針ともすり合わせながらのサーバ整備は、簡単ではないが着実に進められている。サーバや DB 維持にかかる様々なコストについては、情報系以外の人にも良く理解してもらいたい。

(4) バイオリソース整備事業における国際交流や国際化に積極的に取り組んでいるか、国際的なハブとして機能しているか

- アジアのリソースセンター連携コンソーシアム ANRRC の情報基盤の運営を支援しながら、マウスリソースにかかる国際連携として IMPC、AAMMRA、国際データポータルとの連携として IMSR、Cellosaurus、WDCM と連携するなど、国際交流や国際化に積極的に取り組んでおり、国際的なハブとして機能し、リソースセンターにおけるデータ統合について大きな存在感を示している。
- SBI やサンガー研究所/EMBL、COVID-19 オントロジーワーキングとの共同研究を主体的積極的に進め、国際的なハブとして貢献している。今後さらに国際化面を発展させるためにはさらなる人材の投入も必要かもしれない。
- 国際化について、次の国際標準となる仕様の提案に至る時期ではないかと期待する。
- 理研 BRC 全体としての国際交流・国際化のなかで情報開発室が重要な役割を果たしており、国際的なハブとしても十分に機能していることは明確である。また、オントロジーなどリソースに関する情報整備でも国際連携の中で進めている。

(5) 技術開発や研究開発のための人材登用・人材育成及び情報発信・広報は適切か

- メタデータ統合・国際標準化・横断検索、HP コンテンツ充実、大規模データ解析と適切に人材配置、育成しており、今後研究員 1 名の登用を検討するなど適切に取り組んでいる。バイオリソース整備事業を活性化するための情報発信・広報活動も、理研 BRC の情報発信・広報活動の一環として取り組んでおり、十分である。
- 情報整備の重要性が増す一方で、データサイエンス人材は不足しており、情報人材をどのように配置するかは大きな課題である。情報を担当する開発室だけでなく、各リソースの開発室にも 1 名ないし 2 名のリソースに精通した情報人材を配置することで、メタデータ標準化やオミックスデータによる付加価値向上が推進できるだろう。そのためのリソース情報人材育成を担う取り組みやキャリアパスも必要になる。
- 理研 BRC 内だけでの人材育成には限界があるように思う。例えば、理系の大学学部や大学院研究科などとの連携により、IT 教育などの出張講義を計画するなどの施策も考慮すべきである。
- Twitter での情報発信など新しい試みを検討・実施していることは評価できる。
- 塩基配列データバンク同様、あくまで研究基盤であり注目を集めにくい事業であるため、引き続き科学者コミュニティや一般への「魅力的な」情報発信を実施する努力を続ける必要がある。

2. 技術開発あるいは研究開発の計画

(1) 第 4 期中長期目標の達成に向けた残り期間（2025 年 3 月まで）の計画はセンターやバイオリソース整備事業の発展に貢献するか

- 第 4 期中長期目標の達成に向けた残り期間の計画が適切に設定されている。
- AI を活用した大規模データ解析技術・データ可視化技術の研究開発については開発室としての次の主要ツールとしても期待している。
- 大規模データ解析技術・データ可視化技術の開発は重要なテーマであるが、人的リソースの問題もあると思われる。課題を抱えるチームも含めて外部とうまく連携しつつ進めていると見受けられるが、さらに解析の専門家に入ってもらえるといいのではないかと感じた。
- AI 活用によって大きな成果が創出されることが期待される。BRISO の SIP2 終了後の展開にも期待している。

- センター全体の方針と協調したものとして適切に計画されている。特に AI 活用は各リソースからのニーズも高くなると予想されることから、協業の場の創出に期待したい。次々にリリースされる AI 関連サービスをどのように生命科学、リソース分野で活用するのかについての議論や実践の場としても、リソース関連論文の自動抽出は NBRP 全体として実務的に要望の高い課題でもあり実用化に期待したい。

(2) 次期中長期目標期間（第 5 期：2025 年度～2031 年度）に向けて計画しているバイオリソース整備事業は適切か

- 次期中長期目標期間に向けての計画は適切である。
- 情報統合、標準化は地味だが重要な事業なので粛々と進めてほしい。ビッグデータ解析や未来予測へのチャレンジも重要なテーマである。その中で必要な情報整備（オントロジーを用いるところも含めて）は自動化も含め重点的に進めてほしい。
- 第 5 期にむけた計画はまだ少しフォーカスを定めるための議論が必要なようにも感じる。いずれも魅力的な内容であるが、全体としてどこを目指そうとしているのか、第 5 期終了時の姿を明確に共有すべき。また、遺伝資源のデジタル配列情報(DSI)や、研究データ管理に関するデジタル庁への対応などがキーワードとして挙がっていないように思われたが、必要に応じて検討していただきたい。
- 適切かつ堅実に計画されている。特に RRID 国際協力の推進、提供/寄託にかかる次世代オンラインシステムについては、理研外のリソース並びに情報センターと連携協力して進めていただきたい。RDF やオントロジーの真価が一般には理解しづらい面、本当に実用に足るのかという疑問があるので、未来予測へのチャレンジは、成功するにせよ、そうでない場合も問題点が明らかになり、意義があるものと考えられる。

(3) 第 5 期を見据えた長期的視点から、新規に取り組むべき事業について委員からの提案

- サーバの分散化とセキュリティの強化及び 24 時間体制での運用に向けて、理研全体としてもさらに工夫できると良い。外部クラウドサービスの利用と理研内でのストレージサーバの運用（理研独自データの保全）のバランスが重要である。
- バイオリソース循環を有効に駆動させるには情報が中心となって、国内外のリソースの情報統合とそこから次に必要となるリソースの提案などを進める必要がある。新たなリソース整備事業の拡充が国のプロジェクトとしてされるためにも、理研 BRC の統合情報開発室が各種リソース提供機関と協調することが重要である。
- 開発室はオープンサイエンスに大きく貢献していると理解しているが、理研 BRC 他室との連携、さらには他機関とのデータ連携などが進められるとなおよい。
- 次期中長期目標期間に向けて、理研 BRC の統合情報プラットフォームのあり方を検討することを提案したい。理研 BRC ミッションの高度実現に向け、統合情報プラットフォームをデザインすることが重要である。バイオリソース横断検索に加えて、各開発室のデータベースの検索結果を、AI やオントロジーを駆使して、フロントエンドで統合して利用者に提示してはどうか。これにより、5 つのリソースを 1 つのセンターで扱う理研 BRC の統合情報プラットフォームはさらに価値が上がる。第 4 期では、既存システムの統廃合が進み、よりよい情報プラットフォームの構築が進んだことを高く評価する。第 5 期は統合情報プラットフォームのグランドデザインを起し、全体の計画を立案するところから取り組むことに期待する。
- 中長期的には現在「最先端」とされる情報技術が陳腐化することが考え得るため、変化に応じた技術の取捨選択を実施するための情報収集と取り組みの機動性が得られるように、人材の確保と技術力の維持・増強を図られたい。