

## 第13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

(2021年12月28日開催)

評価・提言

### 統合情報開発室（室長：柵屋 啓志）

#### 1. 昨年度（2020年度）及び2021年度上半期（9月）までの実績（事業実績と研究開発の成果）について

##### （1）前回委員会（2020年7月21日開催）の指摘事項への対応は充分か

- ・極めて真摯かつ十分な対応をされている。
- ・システム運用の見直しにより事業実施の効率化を実現しており、対応は十分であるとする。その中で大規模データ解析の重点明確化もされているが、BRC全体としての情報解析基盤の重要性という意味では、統合情報開発室のみに依存するのではなく、バイオリソース関連研究開発プログラムの一つとしてプロジェクトを立ち上げるぐらいの重点化があってもいいかもしれない。
- ・情報の標準化と外部機関との連携についても、バイオリソースについては国内関連機関・プロジェクトとうまく連携して進めることができている。また、病名についても横断検索機能において国際疾病分類 ICD-11 に対応しており、充分である。引き続き NBRP、NITE、農研機構などの省間連携も含めて進めてほしい。
- ・BCPの見直しについては統合情報開発室でのシステム運用の見直しは評価できるが、BCP自体は理研全体または BRC で検討すべき内容だと思われる。
- ・前々回の委員会でコメントしたデータの品質評価についても活動を開始されていることを嬉しく思う。
- ・多岐に渡る指摘事項について、それぞれ適切な対処、もしくは説明がなされている。計画の統廃合や省力化を実施したことにより、他の課題の成果につながったと評価できる。指摘事項に対する機関の対応も評価できる。

##### （2）実績は世界の主要なバイオリソースセンターの関連事業の水準に達しているか

- ・世界の主要なバイオリソースセンター ATCC や DSMZ と並ぶ、世界的な情報提供水準に達し、また、維持している。
- ・それぞれのバイオリソースセンターの取り扱いリソースや運営形態が異なるため、単純な数字での比較には問題がある場合があるかもしれないが、本センターが世界の主要センターのなかでもきわめて高い水準で実施していることは間違いない。
- ・3つの柱をメタデータ統合化・国際標準化・横断検索を中心として国内外の関連機関と協力しながら事業を進めており、データベース登録件数は関連事業の水準に十分達している。
- ・ウェブサイトのアクセス数も順調に伸びているように見える。データ科学者がデータ利活用するためのメタデータの整備も重要であるが、ウェブサイトのアクセス数増加がリソース利用の増加に結びついているかどうかは判断が難しい。評価指標を設定するのが難しい点はあるが、リソース提供における情報の役割をどのように評価するかを検討するといいかもしい。
- ・メタデータの整備と DB アクセス増加に関して、理研内での BRC 以外の RDF データ整備プロジェクトとの関係が見えるとよいと感じた。
- ・海外に比してひけをとらないと思う。また、海外勢との国際連携も進められている点は高く評価できる。
- ・特に、メタデータ統合や情報標準化、RRID の導入において高い水準に達している。

### **(3) 実績は理研 BRC の第 4 期中長期計画（2018 年度～2024 年度の 7 年間）に沿い適切か、また、センターの発展に貢献したか**

- ・BRC の情報センターとして十分に機能しており、センターにとって不可欠な存在となっている。
- ・適切であり、十分な貢献を果たしている。
- ・様々な情報が付加された世界最高水準のバイオリソースを提供するというバイオリソース整備事業の計画に沿って適切に進めており、横断検索の拡張、ホームページリニューアル、大規模データ解析支援などにおいてセンターの発展にも貢献している。
- ・各種メトリクスに現れている通り、中長期計画に沿って適切、かつ、発展的に活動し、実績を挙げている。
- ・リソースのヒト医学研究への活用のための情報整備など、研究動向と利用者のニーズに応えたデータベース整備が実施されている。
- ・COVID-19 の研究の支援等の社会ニーズ・研究ニーズを捉えながら、世界最高水準のバイオリソースを収集・保存・提供し、基盤技術開発、利活用に資する研究開発を推進している。

### **(4) 社会や研究コミュニティへの広報・情報発信は充分だったか**

- ・広報、情報発信についても十分と思われるが、トラブル対応（ハード及びソフトの両面）とハードウェアの変化に応じたパッチやソフトウェアの更新計画を十分に議論し、情報発信や運用の停止に繋がらないように、更なる調整が行われることを期待する。また、アクセスログ解析に基づく、サービスの向上を目指せると良いと思う。
- ・多大な努力をしており充分成果は挙げているが、研究基盤として華やかではなく理解されにくい領域ではあるため、今後もより科学者また一般への情報発信を広げる努力を続けていきたい。
- ・上記ホームページリニューアルに加え、BRC20 周年記念行事などへの貢献により十分な成果をあげている。
- ・理研 BRC 設立 20 周年記念事業をはじめ、HP の充実等、広報、情報発信も問題ない。
- ・統合情報開発室として、バイオリソース推進室と連携し事業を進めている点や、各リソースのコミュニティ向けメールニュースのシステム運用も行っている点、HP の開発運用の面で、BRC の広報・情報発信を支えている。
- ・BRC ウェブサイトのアクセス数解析においても月別ユーザー数、訪問回数、閲覧ページ数が増加傾向にあり充分である。

## **2. 第 4 期中長期計画内（～2024 年度まで）の事業計画について**

### **(1) 計画は世界の主要なバイオリソースセンターの水準に達するために適切か、また、理研 BRC の発展に貢献するか**

- ・BRC の発展には必要不可欠な計画となっており、貢献への期待度は大きいと思う。
- ・適切な計画であり発展に貢献するものと考える。
- ・理研 BRC の発展に貢献する適切なものであるが、大規模データ解析技術・データ可視化技術については統合情報開発室の人員を増やす必要があるのではないか。または、1(1)で書いたようにバイオリソース関連研究開発プログラムの一つとしてプロジェクトを立ち上げるぐらいの役割分担があってもいいかもしれない。
- ・計画通りに実施されれば、海外に比しても十分な水準であり、さらなる発展が期待できる。ただ、大規模データ解析技術・データ可視化技術に関しては、まだもう少し実態と方向性が理解できなかった。
- ・計画は適切である。国内のリソース情報の国際標準化の先導役として期待したい。
- ・メタデータ統合・国際標準化・横断検索、HP 公開コンテンツ充実、大規模データ解析技術・データ可視化技術の 3 本の柱で取り組んでおり、発展に貢献している。

## **(2)過去の委員会での委員からの提言は計画に活かされているか**

- ・コロナ禍でもあり、危機管理についてのステップアップを期待する。
- ・十分に活かされている。
- ・サーバー運用体制の整備や BCP 対策、国際標準化など、時間のかかる課題にも着実に取り組んでいることが理解できた。
- ・関連分野、連携組織の委員からの提言助言が取り入れられて計画に活かされており、BRC 以外への波及効果もあると認められる。

## **(3)計画に無い新規に整備すべきリソース、研究開発、技術開発はあるか**

- ・理化学研究所全体のクラウド化に即して、公開系などのクラウド化計画を議論する必要があると思う。
- ・現時点で提言は無い。ただし中長期的には現在主流、あるいは期待されている情報技術が陳腐化または置き換わるといった状況は過去の経緯からも充分考え得るため、機動的に時宜に応じた情報技術の取捨選択を実施できるよう考えていただき運用に反映されることを期待する。
- ・大規模データ解析技術・データ可視化技術に関して、リソースの新しい利活用法の検討ができるといい。ウェブサイトのアクセス増がデータベースアクセスによるものということなので、それらがどのような形で利用されていくのか興味がある。
- ・計画は野心的であり、現計画の着実な実施が望まれる。ナレッジグラフを大規模に活用している事例としては、日本はもちろん、欧米等と比較してもトップクラスだと思われる。今回、JSAI で優秀賞を取られたことは素晴らしいと思うが、是非、実際の活用まで含めて海外のトップカンファレンスでのアピールを計画していただきたい。
- ・特になし。むしろ利用の低いサービスを廃止することを検討するとよい。

## **(4)次年度（2022年度）の計画は適切に設定されているか**

- ・適切に設定されている。
- ・次年度の計画に問題ない。
- ・実施可能な内容で適切に設定されている。RRID のカタログ表示について国内の先行事例となり、メリットなどを示してもらえることを期待している。
- ・ゲノム研究の進展でモデル生物においても多型情報を活用したいというニーズが高まっている。MoG+の RDF 化と TogoVar との連携は、今後、他の生物にも応用できるかどうか、成果に期待したい。

## **3. その他**

### **（理研 BRC の第 5 期中長期計画（2025～）策定や 10 年先を見据えた長期的展望からの提言等）**

- ・BRC の 3 つの柱として、バイオリソース整備事業、基盤技術開発事業、バイオリソース関連研究開発プログラムが挙げられているが、現在は、バイオリソース整備事業の中に位置付けている「統合情報開発室」のみが、活動しており、基盤技術の開発から次世代のデジタル技術動向に至る項目も担当していると思われる。基盤技術あるいはバイオリソース関連研究開発プログラムの中に、セキュリティやインターネット関連技術、バイオリソース管理用データベースシステム開発、デジタル技術とバイオデータサイエンスを中心に活動する部門の新規設置を検討いただけることを期待する。
- ・今回も、過去から引き継がれた情報基盤の維持・更新をしつつ、新たな機能を追加していく様子が垣間見えた。情報の統合、廃止などは成果として目に見えにくいことであるためご苦勞の割にはやりにくい仕事かと思うが、我々委員はそのあたりもきちんと見聞きし、評価させていただく必要があると思う。委員長からはクラウド利用のご

提言があったが、個人的には、現状、まだオンプレで機材をもつほうが運営コスト的にはかなり有利と思っている。しかし将来、ご紹介いただいた理研の、あるいは民間クラウドへのシステム基盤の移行の可能性も充分あるかと思うので、各種 DB や基幹システムのコンテナ化など、転換時に移行のしやすいシステム構築をご念頭に置いて長期的な整備をすすめると良いのではないかと考える。

・「情報なくしてリソースの価値なし」は重要であるので、情報の重要性をより強調したものにできるとよいと思う。情報によるリソースの高付加価値化と新たな活用法の提案に加え、それによってどのように人類の諸課題解決につながるのか、また、そのためにどのようなリソースが新たに必要となるかなど、BRC 自体の戦略立案に役立つ情報基盤も提供できるというのではないだろうか。

・既に、研究データは、関連する分野毎に国を挙げて活用を考えるべき時期だと思う。貴センターのミッション外になるかもしれないが、中長期計画には「外部機関等との相補的な連携の促進を図る」ともあり、オープンとクローズを合わせて、どのように他機関、企業と連携していけるか、いくべきか、一部サイエンスの枠を超えて、一緒に検討していければ有り難い。

・理研 ICT 基盤更新にともない発生する作業や、センター内の調整については、機関からの資金的な支援、人的な支援をお願いしたい。

・DX など情報統合の重要性が様々な分野で言われ始めているが、実際には手間暇のかかる作業であり、データを整えている間に技術が変化しそれへの対応も必要となる。限られた予算の中で、情報整備・統合への大胆な配分があるか無いかで、結果は大きく異なってくると思う。また、それを支える人材育成も、BRC、理研だけでなく全体の課題だと思う。

・現在、オープンな情報リソースの整備に取り組まれているが、今後は、BRC として他の開発室と連携して、バイオリソースの利活用促進という観点で、BRC の活動を先導することが望まれる。