

ライフサイエンスデータベース統合推進事業 平成 23 年度～平成 28 年度 事業評価報告書

1. 評価の概要

1. はじめに

本評価結果は、JSTが、平成23年度から実施しているライフサイエンスデータベース統合推進事業の平成28年度までの実施状況等について、外部有識者からなる「ライフサイエンスデータベース統合推進事業評価委員会」による事業評価を行い、その結果をまとめたものである。

2. 評価対象事業

ライフサイエンスデータベース統合推進事業

3. 事業評価の目的

NBDC は「内閣府総合科学技術会議 統合データベースタスクフォース報告書（平成 21 年 5 月）」を受け、文部科学省統合データベースプロジェクトと JST バイオインフォマティクス推進センターの活動を継続・統合する形で平成 23 年 4 月に JST に設置された。

NBDC では、公的研究開発投資の成果であるデータを、効率よく再利用できるように統合データベースを整備し、データ価値の最大化を図ることを目的として、上記報告書で統合データベースセンターに求められている機能を網羅する 4 本の事業の柱（戦略の立案、ポータルサイトの構築・運用、データベース統合化基盤技術の開発、バイオ関連データベース統合化の推進）を構成し、その柱に沿って、事業を企画・運営している。

センター運営 6 年目を迎え、今後の NBDC の事業推進のために、これまで進めてきた事業推進状況の検証を行うことを目的として、事業の評価を行った。

なお、今後の NBDC の具体的な事業運営の方向性については、事業評価の結果も踏まえ、NBDC 運営委員会にて検討することとしている。

4. 事業概要

(1) 戦略の立案

データベース整備・統合化の戦略企画、ガイドラインの策定、国内外との連携構築等を実施。

(2) ポータルサイトの構築・運用

データベースのカタログ、横断検索、アーカイブ等のサービスの提供を実施。

(3) データベース統合化基盤技術の開発

データベース統合化に向けて基盤となる技術開発とその実装を実施。

(4) バイオ関連データベース統合化の推進

分野毎のデータベース統合化により、国内バイオ関連データベースの統合を実施。

5. ライフサイエンスデータベース統合推進事業評価委員会委員

	氏名	所属
主査	清水 謙多郎	東京大学大学院農学生命科学研究科 教授
委員	大西 康夫	東京大学大学院農学生命科学研究科 教授
委員	小幡 裕一	理化学研究所 バイオリソースセンター センター長
委員	鈴木 蘭美	エーザイ株式会社 コーポレート BD 部 部長 執行役
委員	辻井 潤一	産業技術総合研究所人工知能研究センター センター長
委員	山崎 真巳	千葉大学大学院薬学研究院 准教授
委員	吉田 輝彦	国立がん研究センター中央病院 遺伝子診断部門長

(主査以外は 50 音順)

6. 評価方法

6. 1 評価手順

1) 書面評価

1 1月下旬に事業報告書による書面評価を行った。

2) 事業評価委員会

1 2月19日(月) J S T 東京本部にて事業評価委員会を開催した。

- ・センター長からの実施報告と質疑応答
- ・評価委員による議論

3) 評価報告書の作成

書面評価ならびに事業評価委員会での議論に基づき評価結果を取りまとめた。

6. 2 評価項目

1) 事業計画や運営体制について

2) 事業の進捗状況について

3) 事業成果は十分か

なお、上記事業実績の評価に加え、「今後の事業継続にあたり、強化・改善すべき点等について」、評価委員からの意見を頂いた。

II. 評価結果

全体的評価

NBDC は、平成 23 年度の設立までの準備段階、平成 23 年度からの 3 年を期間とした第一段階および平成 26 年度以降の第二段階を経て現在に至るが、段階ごとに推進目標等を定め、統合データベースセンターの持つべき機能として定められた機能を着実に実施してきている。

4 省連携（文部科学省、厚生労働省、農林水産省、経済産業省）によるポータルサイトの構築・運用、データ共有のためのガイドラインの策定とヒトデータベースの運用、データベース統合に係る基盤技術の構築や国内外のデータベース構築に係る実務担当者とのネットワーク形成、さらに、国内の重要なデータベースの開発・維持、生物種や各オミクス階層を超えたデータベース統合の基盤整備といった成果が得られている。加えて、当初定められた機能以外にも、人材育成の検討や関連機関との連携等の取組みを積極的に進めている。

以上のことから、事業計画や運営体制、事業の推進状況や成果等、全体として優れた成果を出していると結論づけられ、近年のオープンサイエンスや情報技術等に関わる動向にも配慮しつつ、今後も継続して事業を推進することが望まれる。

1. 事業計画や運営体制について

本事業は、総合科学技術会議ライフサイエンス PT 統合データベースタスクフォースにより取りまとめられた「統合データベースタスクフォース報告書」（平成 21 年 5 月 27 日）の内容を踏まえて、平成 23 年度から開始している。

我が国において優れた統合データベースを効率的に構築していくためには、基本方針や推進方策の策定、関係省庁間の調整など司令塔機能を担う機関として、総合科学技術会議（当時）に「統合データベース推進本部（仮称）」を設置することが求められていたが、それが実現しなかった。このような状況の中、上記報告書をもとに、NBDC においてデータベース統合に向けての課題の洗い出し、制度設計、運営体制の構築がなされ、適切に事業が推進されている点は評価できる。

さらに、事業開始当時は予見できなかった国立研究開発法人日本医療研究開発機構（AMED）の創設等への対応も進めており、プロジェクトとの連携を推進できている点も評価できる。

2. 事業の進捗状況について

(1) 戦略の立案

ヒト由来データの取扱いに関する統一的な指針が整備されていない状況で、いち早く、データ共有のためのガイドラインを策定し、ヒトデータベースの運用を開始した。さらに、Global Alliance For Genomic Health(GA4GH)への加入による、ヒトデータの国際的な共有の動向把握やプロジェクトへの参画、加えて、Resource Description Framework (RDF)によるデータ統合に向けた国際連携やBioHackathon等の開催など、適切に取り組んでいると評価できる。

また、AMED やその他大型プロジェクトと連携し、公開前データの共有支援も視野に入れて戦略を検討していることも評価できる。

(2) ポータルサイトの構築・運用

4省連携によるポータルサイトの構築・運用として、カタログ、アーカイブ、横断検索の構築・運用など、順調に進捗しており、多様なデータベースへの入り口が統合され、容易にアクセスできるようになったことは評価できる。さらに、RDF形式データを掲載したポータルサイトの構築は、統合化基盤の普及という点で意義のある取り組みであり、サイトの構築が進捗していることは評価できる。

(3) データベース統合化基盤技術の開発

統合化のための基盤的な技術開発と並行して、ライフサイエンス分野のデータベース統合のために、RDF化に関する国際的な標準化を積極的に先導してきたことは評価できる。また、データベース開発担当者を集めて毎月開催してきたSPARQLthonだけでなく、国際連携にも資するBioHackathonやライフサイエンス分野における主要なデータベース構築を担う機関の実務担当者会議等の開催を通じ、国内外に人的ネットワークを構築してきたが、この人的ネットワークはデータベース統合化基盤技術の開発に大いに貢献している点で、これらの活動も評価できる。

(4) バイオ関連データベース統合化の推進

統合化推進プログラムで支援されたデータベースのうち、現時点でのアクセス数や活用例等からは、利用者による活用がまだ十分とはいえないものも一部見受けられるが、多くのデータベースでは有効性・重要性が十分示されている。全体としては、国内の重要なバイオ関連データベースが多数開発・維持されてきており、さらに、生物種や各オミクス階層を超えたデータベース統合の基盤が整備されていることは評価できる。

ライフサイエンス分野全体を網羅するデータベースが整備されれば、データ駆動型の新しい科学の進展が期待できるため、今後も本プログラムの推進が望まれる。

(5) その他

人材育成施策に関する予算が認可されなかった点は残念であるが、そのような状況の中でも、現場のニーズに応えた人材育成のために、講習会の企画・開催などの、積極的な活動を行っていることは妥当であるといえる。

3. 事業成果は十分か

(1) 研究開発による成果について、ライフサイエンス分野のデータベースの統合に資する成果が得られているか。

NBDC がライフサイエンス分野におけるオープンサイエンス推進の中心となっており、ポータルサイトの構築、技術開発、重要なデータベース開発の支援等、どれも成果を挙げている。さらに、データベースの統合を目指した RDF の普及等が行われており評価できる。統合化推進プログラムによって支援されたデータベースの開発および統合化については、一部のデータベースには、これからの成果を期待すべきものがあるものの、全体として成果が得られており評価できる。

(2) ライフサイエンスデータベース関連府省との連携、データ拡充および利用状況などについて、ライフサイエンス研究開発全体の活性化に資する十分な成果が得られているか。

関連府省との連携については、他省庁の取組みが縮小してきていることが気になる状況でありながらも、NBDC が中心に進めてきた結果として、ポータルサイトの構築・運用等においては、カタログ収録数、横断検索対象データベース数、アーカイブ数とも着実に拡充され、計画通りの成果をあげている。また、統合化推進プログラムによって支援されたデータベースの開発および統合化については、非常によく利用されているものも多く、ライフサイエンス研究開発全体の活性化に資する十分な成果が得られていると評価できる。

III. 今後の事業継続にあたって、強化・改善すべき点についての意見

- (1) 医科学、医療、創薬の分野に加え、食、環境、バイオマスエネルギー（グリーンバイオ）などの分野におけるデータベース整備も重要である。
- (2) 整備されたデータベースをより活用するために、高度な知識処理、データマイニング、人工知能研究、ソフトウェアの開発、大量のデータを扱うための技術、HPC (high-performance computing) 関連技術の開発、などに取り組む必要があると考えられ、ライフサイエンス分野以外の情報科学研究者および組織との連携をより一層強化してほしい。
- (3) ライフサイエンス分野の研究開発には、構造化データだけではなく、画像、テキストなどの非構造化データ、さらに、RDF 化が困難な膨大なデータ等も重要であり、引き続きそれらデータの取扱いについての検討を進めてほしい。
- (4) ポータルサイトや、統合化推進プログラムにより支援されたデータベースについて、より多くのユーザーから利用されるための改善・見直しを続けるとともに、さらに、データベースの利用状況をよりよく把握する取組みも考えるべきである。
- (5) ライフサイエンス研究におけるインフォマティックス人材の育成は重要な課題であるが、NBDC 単独で行えることには限界もあるため、事業全体における位置づけや役割、取組みの推進方法を検討して進める必要がある。
- (6) NBDC は今まで、「統合データベースタスクフォース報告書」(平成 21 年 5 月 27 日)で取りまとめられた内容を踏まえて事業を推進してきたが、今後の進め方については、研究の動向等の変化も踏まえて、事業が推進すべき取組み、さらに、人的資源や予算配分・優先度等の再検討が必要ではないか。

以上