NBDC
National Bioscience Database Center

独立行政法人
科学技術振興機構
Japan Science and Technology Agency

生命科学系データベースアーカイブと省間連携

坂東明日佳[1]、熊谷禎洋[2]、大久保克彦[2]、八塚茂[2]、○畠中秀樹[1]
1独立行政法人科学技術振興機構　バイオサイエンスデータベースセンター、2(株)日立ソリューションズ

生命科学系データベースアーカイブとは

あのデータベースが、丸ごとダウンロード可能に！

国内で生産された生命科学系データベースを誰もが利用できるよう、
ダウンロード可能なかたちで長期にわたり維持していくサービスである。

アーカイブのWebサイト

一括ダウンロード
データセットを丸ごとダウンロードすることが可能。

データの説明（メタデータ）
データベースやデータ内容を十分に説明して、欲しいデータセットが見つかりやすくしている。

簡易検索
データの閲覧や検索,ダウンロードができる「簡易検索ページ」を設置。

利用許諾の明示
データがさまざまな用途に活用されるよう、利用許諾条件を明確化。

アーカイブにおける省間連携

データベース統合推進事業におけるアーカイブ構築

NBDCのポータルサイトに、各省のデータベースをガイドラインに沿ってアーカイブ化して保存し、**各省のデータベース**に対して、**統一形式でのデータダウンロードが可能となる状態** を実現する。

各省協力機関

経済産業省・産業技術総合研究所
MEDALS　http://medals.jp/

農林水産省・農業生物資源研究所
AgriTOGO　http://togo.dna.affrc.go.jp/

厚生労働省・医薬基盤研究所
NIBIO　http://www.nibio.go.jp/

課題

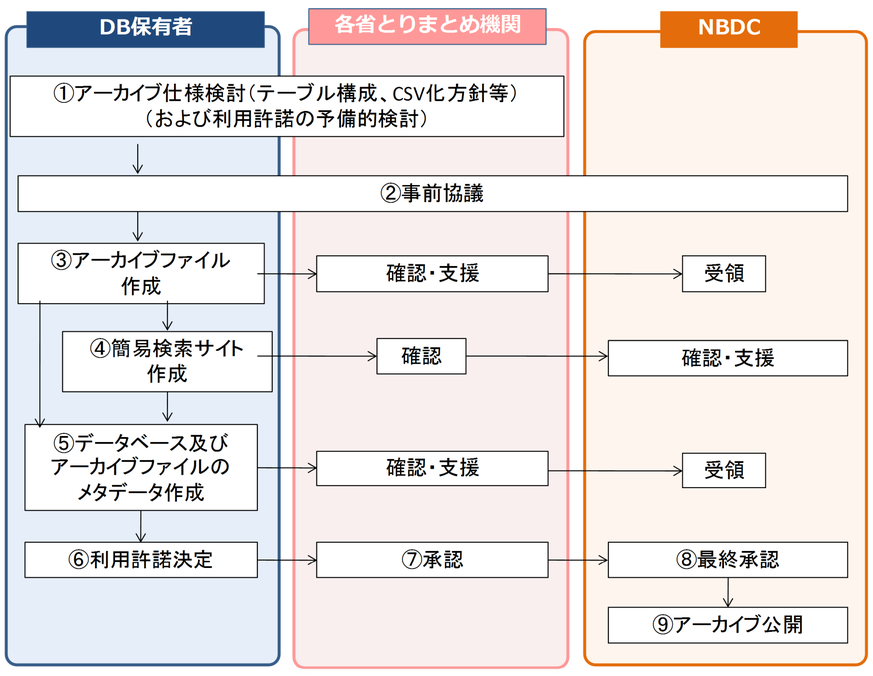
各省が取りまとめるDBをアーカイブで公開するためのデータ準備作業のガイドライン化。

省間連携のためのアーカイブの取り組み

アーカイブ作成ガイドラインの作成、公開

作業フローの定義

各省取りまとめ機関とNBDC間での作業フローを定義し、このフローに基づき作業を進める。

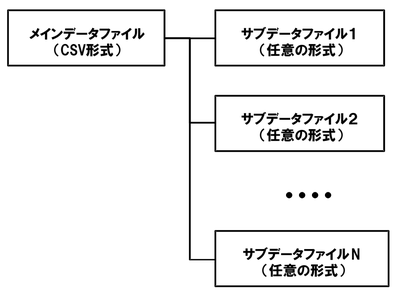


アーカイブ作業の明確化・効率化

データ加工のガイドライン

データのファイル構成や項目立てについての方針を規定。

ファイル構成



データ項目

- 一覧性を重視した項目の決定
- 項目ごとにデータ型を統一

データ構造を均質化し、将来の統合検索への布石とする。

利用許諾決定の支援

- 使用許諾テンプレートの提供

クリエイティブ・コモンズ・ジャパンの提供するライセンスシステムを使用許諾テンプレートとして採用。

法律に関する専門的な知識が不要に。

- 標準利用許諾の推奨

データの営利利用や改変が可能なクリエイティブ・コモンズ・ライセンスを標準利用許諾として推奨。

「表示一継承」(CC BY-SA)

公共財であるデータの利用を促進

生命科学系データベースアーカイブWebサイトの改善

取りまとめ機関が直接メタデータ入力などのアーカイブ化作業ができるよう、Webサイトを改善した。

各省で作業負荷を分散し、アーカイブ化を加速

省間連携の成果(例)

DIAM-微生物情報（経済産業省との連携）
微生物の同定・分類関連情報、安全関連情報、利用関連情報に関するデータベース。

Open TG-GATEs（厚生労働省との連携）
131の化合物をラット個体およびラット・ヒト肝細胞へ曝露した際の遺伝子発現と毒性のデータベース。

今後の展開

省間連携の強化
アーカイブ作成を通じて、ガイドラインの改良や拡充を進める。
公募要領でデータ提供を明文化したプロジェクトのデータへの対応を進める。

アーカイブサイトの拡充
セマンティックWeb技術を適用し、データベース間の統合的な検索を図る。
データベースカタログ・横断検索とのインタフェースの統合を図る。
アーカイブの利用事例をサイトに掲載し、アーカイブの普及・利用促進を図る。

トーゴーの日シンポジウム 2011　ライフサイエンス分野のデータベース統合の“カタチ”を探る

会期：2011.10.5　於：日本科学未来館　みらい CAN ホール

Licensed under a Creative Commons 表示2.1 日本 license
(c) 2011 バイオサイエンスデータベースセンター