

ライフサイエンスDB統合推進事業とヒトDB

東大・理・生物/JST-NBDC/遺伝研DDBJ 高木利久

- データの共有、公共財化を促進し、その価値を最大化
- 内閣府CSTP主導の統合データベースプロジェクト(H18～)
 - 文科省、経産省、農水省、厚労省で実施
 - H23年12月に四省連携のポータルサイト
- 文科省の統合データベースプロジェクト(H18～)
 - 中核センターの設立
 - H19～情報・システム研究機構ライフサイエンス統合DBセンターDBCLS
 - H23～科学技術振興機構バイオサイエンスDBセンターNBDC
 - クリエイティブコモンズ(CC)ライセンスによるデータの共有
 - フォーマット、辞書、統合技術、動画教材などの開発
 - カタログ、横断検索、アーカイブの構築など種々のサービス提供
 - 研究分野ごとのデータベース統合化進行中(ファンディングによる)
 - ヒト由来データの共有・セキュリティガイドラインの作成
 - ヒトDB(オープン、制限アクセス)の構築、受入れ(DDBJと連携して)
 - これまでの申請34件
 - 公開(待機中も含め)23件(オープン6,制限公開17)、その他は審査中など
 - 次世代がん、ゲノム支援、東北メディカルメガバンク、バイオバンクジャパンなどより

NBDC

共有と統合を通じて
生命科学データの
価値の最大化を図る

統合データベースの沿革

2000年11月

科学技術会議 ライフサイエンス部会 ゲノム科学委員会
「ゲノム情報科学におけるわが国の戦略について」(2000年11月17日)
※人材養成、データベース構築、情報解析技術開発の3つの観点から推進戦略を提案

2001年4月

JSTにバイオインフォマティクス推進センター(BIRD)を設立

2004年8月

科学技術・学術審議会 研究計画・評価分科会
ライフサイエンス委員会 データベース整備戦略作業部会
「我が国におけるライフサイエンス分野のデータベース整備戦略のあり方について」(2005年5月17日)
※戦略委員会の設置、ポータルサイトの構築、統合データベースのための技術開発、人材養成を緊急に取り組むべき課題として提言

2006年4月

情報・システム研究機構を中核機関とした文部科学省「統合データベースプロジェクト」が開始

2006年4月

農林水産省、経済産業省でも統合データベースのプロジェクト開始

2008年12月


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
「統合データベースタスクフォース報告書」(2009年5月27日)
※ライフサイエンス統合データベースセンター(DBCLS)とBIRDとの一体的な運用の提言

2011年4月

JSTにバイオサイエンスデータベースセンター(NBDC)を設置
ライフサイエンスデータベース統合推進事業が開始
基盤技術開発プログラムに1件、統合化推進プログラムに10件の課題を採択

2011年12月

4省の生命科学系データベース
合同ポータルサイトを開設
<http://integbio.jp/>

 integbio.jp

2012年4月

統合化推進プログラムに
1件の課題を採択

2013年9月

統合化推進プログラム
(統合データ解析トライアル)に
8件の課題を採択

2014年4月

統合化推進プログラムに
9件の課題を採択

各省連携 統合へのステップ

カタログ 文部科学省、厚生労働省、農林水産省、経済産業省のDB情報を集約し、「Integbio データベースカタログ」にて公開

横断検索 索引データを相互に共有し、検索仕様を共通化して、文部科学省、厚生労働省、農林水産省、経済産業省のDB内を一括して検索

アーカイブ 共通のガイドラインに基づき、文部科学省、厚生労働省、農林水産省、経済産業省でアーカイブ作成を分担してデータを公開

文部科学省：バイオサイエンス
データベースセンター
<http://biosciencedbc.jp/>



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



農林水産省：農業生物資源研究所
<http://agrid.dna.affrc.go.jp/>



経済産業省：産業技術総合研究所
創薬分子プロファイリング研究センター
<http://medals.jp/>

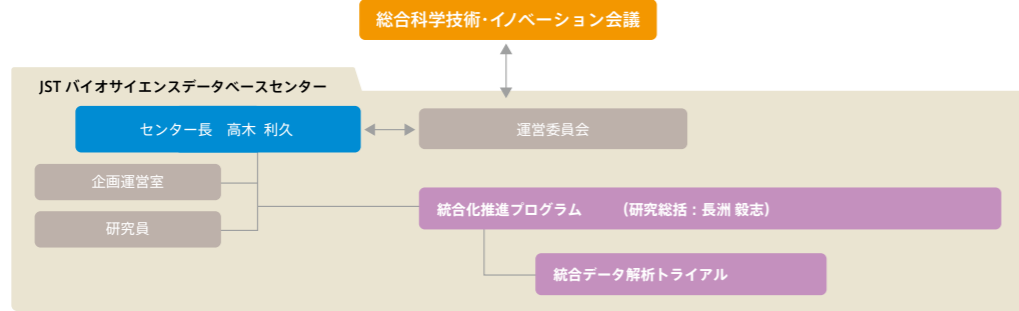
DB再構築 セマンティック・ウェブ技術を利用した統合DBの構築

バイオサイエンスデータベースセンター(NBDC)は日本の生命科学データベースの統合を実現します。

センター長 高木利久 Toshihisa Takagi
(東京大学大学院理学系研究科生物科学専攻)



体制図

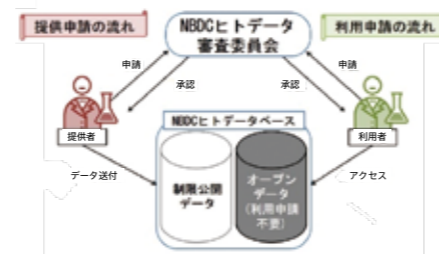


ライフサイエンスデータベース統合推進事業

- 1 戦略の立案**
DBの整備・統合化の戦略企画や、データや技術のコーディネート、DB統合化ガイドラインの策定、国内外との連携構築など
- 2 ポータルサイトの構築・運用**
DBに関連するサービス、研究開発プログラムの情報やセンターの活動を紹介するポータルサイトを運用
- 3 データベース統合化基盤技術の研究開発**
ライフサイエンス統合データベースセンター(DBCLS)/ROISとの共同研究により、DB統合化の実現に向けて基盤となる技術の開発を実施
- 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のファンディングにより、分野ごとのDB統合化等を通じて、生命科学分野DBの統合を実施

NBDC ヒトデータベース <http://humandbs.biosciencedbc.jp/>

個人情報の保護に配慮しつつヒトに関するデータの共有や利用を推進するための共有プラットフォーム『NBDC ヒトデータベース』を設立し、国立遺伝学研究所 DNA Data Bank of Japan (DDBJ) センターと協力して運営しています。
本DBは、次世代シーケンサーをはじめとする解析技術の発達に伴って膨大に産生されつつあるヒトに関するさまざまなデータを整理・格納して、生命科学の進展のために有効に活用するためのルール(ガイドライン)や仕組みを提供するものです。
本サイトから、ヒトに関するデータの利用及び提供の申請を行なうことができます。



広報・普及活動

シンポジウム・ワークショップ <http://events.biosciencedbc.jp/sympo/>

● トーゴの日シンポジウム

毎年10月5日を「トーゴの日」とし、生命科学分野のDB統合にまつわる問題をともに考え、議論を深めるシンポジウムを開催しています。

● 国際開発者会議 BioHackathon (バイオハッカソン) <http://www.biohackathon.org/>

国内外から生命科学の主要なDBの提供者、DB統合技術の研究者やソフトウェア開発者を一同に集め、DBが現在直面している問題に対して、技術開発を交えながら議論する会議です。



講習会 <http://events.biosciencedbc.jp/training/>

生命科学系のDBやツールの使い方、DBを統合する活動を紹介する初心者向けの講習会を開催しています。参加者は、それぞれPCを触りながらハンズオンで、DBやツールの使い方について学びます。2007年度から全国各地の大学や研究機関などで開催されています。

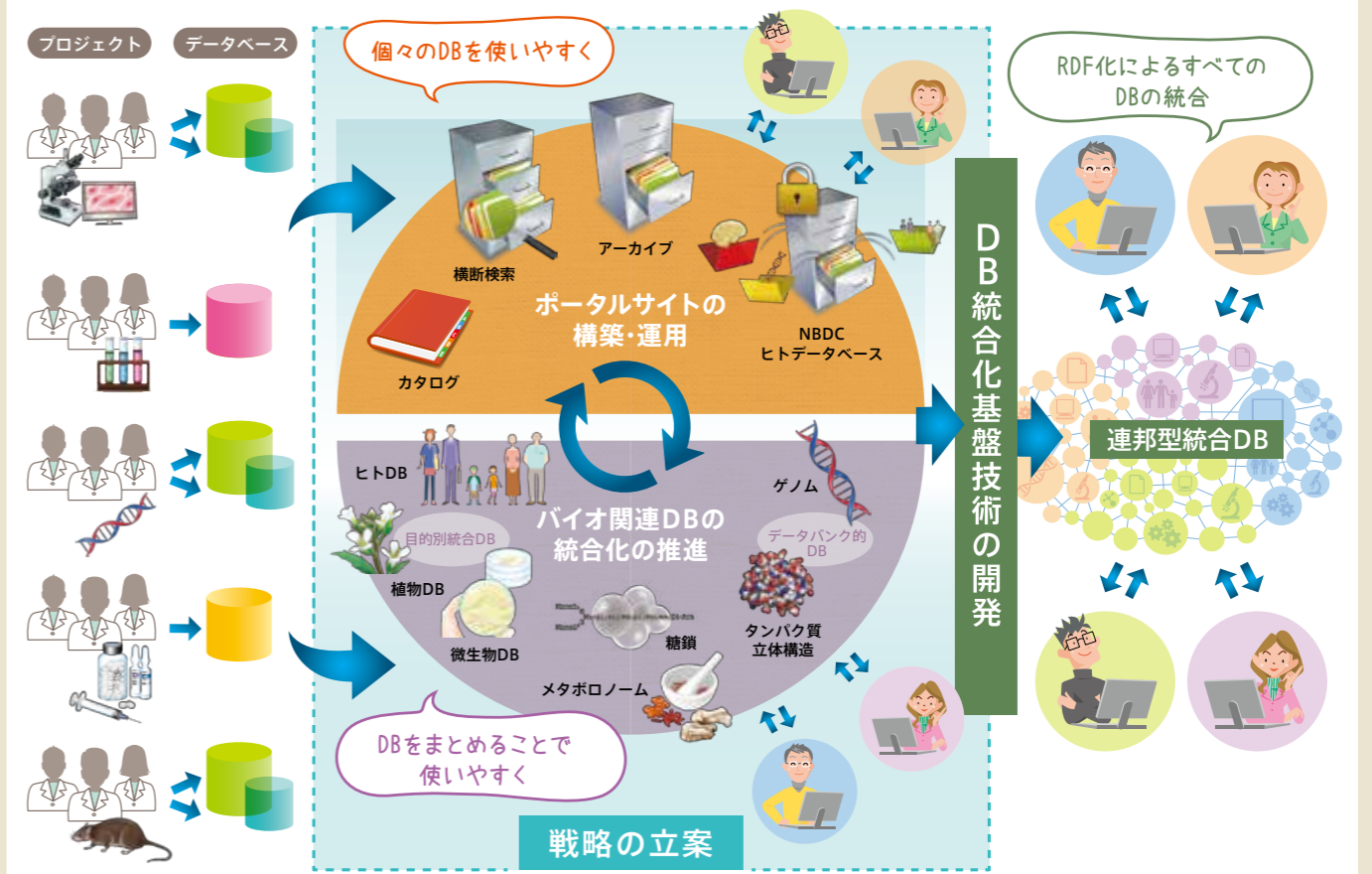


展示会 <http://events.biosciencedbc.jp/exhibition/>

BioJapanなどの民間の展示会や学会の附設展示会でDBやサービスの使い方をポスターや配布資料、デモPCを使用して紹介しています。日本分子生物学会年会では特別コーナー「使ってみようバイオデータベース—つながるデータ、広がる世界」を開催しています。



NBDCの事業推進の構成



ライフサイエンスデータベース統合推進事業のロードマップ



NBDCの取り組み

バイオサイエンスデータベースセンター (NBDC) では、日本の生命科学データベースの統合を実現するための研究開発とサービス提供を行っています。それにより生命科学研究成果が、広く研究者コミュニティに共有かつ活用されることで、基礎研究や産業応用研究が活性化されることが期待されます。

さまざまな成果

ユーザーがNBDCの提供するサービスを生命科学や医療、教育などの分野で活用することで新しい発見や新薬の開発などの成果が生まれることが期待されます。



ユーザー



サービス提供

NBDC
ポータルサイト <http://biosciencedbc.jp/>

分野横断的に収集された情報をポータルサイトから一括利用できます。



カタログ

目的のDBを一覧から探せます。
<http://integbio.jp/dbcatalog/>



横断検索

さまざまなDBが一括検索できます。
<http://biosciencedbc.jp/dbsearch/>

アーカイブ

DBを丸ごとダウンロードできます。
<http://dbarchive.biosciencedbc.jp/>



NBDCヒトデータベース

ヒトに関するさまざまなデータを受け入れ、オープンあるいは制限付きで研究者間で共有します。
<http://humandbs.biosciencedbc.jp>

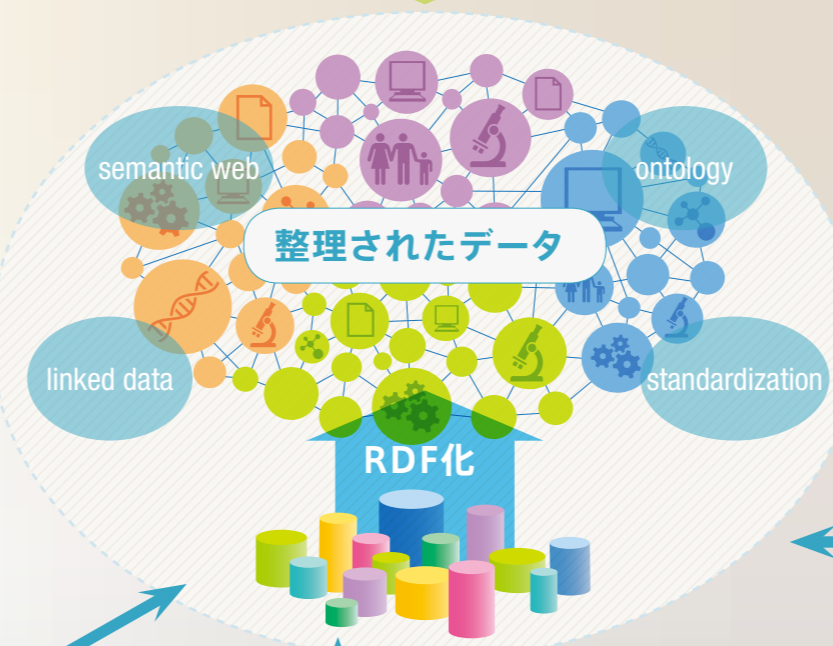
より高度な検索・DBツール

情報技術

より高度なデータベース検索プラットフォームを構築し新しいツールを提供します。さらに生命科学系コンテンツのRDF化を支援する情報技術を開発します。



RDF化支援等



RDF: Resource Description Framework

研究開発

生命科学系コンテンツ

分野ごとに産出される情報を網羅的に収集して整理することによって、生命科学のさまざまな領域の知識を集積します。



データの寄託

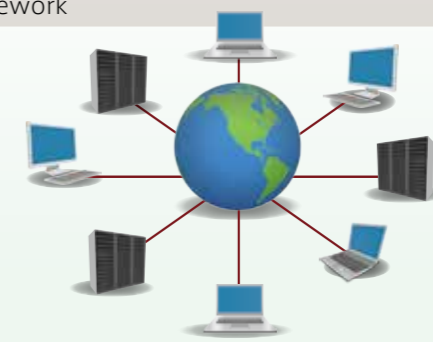
国内のデータ、データベース、バンク情報、ツール、辞書、オントロジー

国内



国際連携

海外のデータベース
海外のデータ共有プロジェクト



研究開発プログラム実施課題

研究代表者、課題名、実施年度を紹介しています。

各課題の成果の詳細は [【研究開発成果データベース】 http://biosciencedbc.jp/db-link](http://biosciencedbc.jp/db-link) をご覧ください。

生命科学系コンテンツ 統合化推進プログラム 2014～
 統合化推進プログラム 研究総括
長洲 毅志 Takeshi Nagasu
 エーザイ株式会社 アドバイザー



有田 正規 Masanori Arita
 理化学研究所環境資源科学研究センター チームリーダー

課題名
 2014～ 生物種メタボロームモデル・データベースの構築

金谷 重彦 Shigehiko Kanaya
 奈良先端科学技術大学院大学情報科学研究科 教授

課題名
 2011～13 メタボローム・データベースの開発

研究開発成果
 バイオマスバンク <http://bio.massbank.jp/>

黒川 顕 Ken Kurokawa
 東京工業大学地球生命研究所 教授

課題名
 2014～ ゲノム・メタゲノム情報統合による微生物DBの超高度化推進
 2011～13 ゲノム・メタゲノム情報を基盤とした微生物DBの統合

研究開発成果
 微生物統合データベース <http://microbedb.jp/>



大浪 修一 Shuichi Onami
 理化学研究所生命システム研究センター発動態研究チーム チームリーダー

課題名
 2012～ 生命動態システム科学のデータベースの統合化

研究開発成果
 生命動態システム科学の統合データベースSSBD [SSBD Database](http://ssbd.qbic.riken.jp)
<http://ssbd.qbic.riken.jp>

岩坪 威 Takeshi Iwatsubo
 東京大学大学院医学系研究科 教授

課題名
 2011～13 ヒト脳疾患画像データベース統合化研究

研究開発成果
 ヒト疾患脳画像データベース <http://bidb.biosciencedbc.jp/>

金久 實 Minoru Kanehisa
 京都大学化学研究所 特任教授

課題名
 2014～ ゲノムとフェノタイプ・疾患・医薬品の統合データベース
 2011～13 ゲノム情報に基づく疾患・医薬品・環境物質データの統合

研究開発成果
 KEGG MEDICUS <http://www.kegg.jp/kegg/medicus/>



田畑 哲之 Satoshi Tabata
 かずさDNA研究所 所長

課題名
 2014～ 植物ゲノム情報活用のための統合研究基盤の構築
 2011～13 ゲノム情報に基づく植物データベースの統合

研究開発成果
 Plant Genome DataBase Japan <http://pgdbj.jp/>



菅野 純夫 Sumio Sugano
 東京大学大学院新領域創成科学研究科 教授

課題名
 2014～ 疾患ヒトゲノム変異の生物学的機能注釈を
 目指した多階層オミクスデータの統合

松田 文彦 Fumihiko Matsuda
 京都大学大学院医学研究科 附属ゲノム医学センター センター長・教授

課題名
 2011～13 大規模ゲノム疫学研究の統合情報基盤の構築


研究開発成果
 Human Genetic Variation Browser <http://www.genome.med.kyoto-u.ac.jp/SnpDB/index.html>



徳永 勝士 Katsushi Tokunaga
 東京大学大学院医学系研究科 教授

課題名
 2014～ 個別化医療に向けたヒトゲノムバリエーションデータベース
 2011～13 ヒトゲノムバリエーションデータベースの開発

研究開発成果
 ヒトゲノムバリエーションデータベース <https://gwas.biosciencedbc.jp/index.Japanese.html>



成松 久 Hisashi Narimatsu
 産業技術総合研究所 糖鎖創薬技術研究センター 招聘研究員

課題名
 2014～ 糖鎖統合データベースおよび国際糖鎖構造リポジトリの開発
 2011～13 糖鎖統合データベースと研究支援ツールの開発


研究開発成果
 JCGGDB <http://jcgddb.jp>



中村 春木 Haruki Nakamura
 大阪大学蛋白質研究所 所長・教授

課題名
 2014～ 蛋白質構造データベースの高度化と統合的運用
 2011～13 蛋白質構造データベースの国際的な構築と統合化

研究開発成果
 PDBj (Protein Data Bank Japan) <http://pdbj.org/>



榎屋 啓志 Hiroshi Masuya
 理化学研究所バイオリソースセンター ユニットリーダー

課題名
 2014～ 生命と環境のフェノーム統合データベース

豊田 哲郎 Tetsuro Toyoda
 理化学研究所情報基盤センター 統合データベース特別ユニットリーダー

課題名
 2011～13 生命と環境のフェノーム統合データベース

研究開発成果
 BioLOD (Biophenome Linked Open Database) <http://biolod.org/>

統合化推進プログラム(統合データ解析トライアル) 統合化推進プログラムのサブプログラムとして、統合化推進プログラムで統合されたデータベースのデータ解析ツールなどを開発

大規模なタンパク質データ解析のための高速な局所配列特徴抽出法の開発	理化学研究所脳科学総合研究センター 研究員 蝦名 鉄平
マルチオミクスデータを用いたゲノム規模代謝モデリングのためのネットワーク解析システムの開発	理化学研究所生命システム研究センター テクニカルスタッフ 西田 孝三
KNApSACKを用いた植物の機能メカニズム解明のための基盤構築	京都大学化学研究所 大学院生 西村 陽介
機械学習を用いたタンパク質-リガンド結合部位予測ツールの自動生成パイプラインの開発	東京大学大学院農学生命科学研究科応用生命工学専攻生物情報工学研究室 大学院生 番野 雅城
植物代謝物プロファイリングデータベース AtMetExpressの開発とオミクスデータ統合の推進	理化学研究所環境資源科学研究センターメタボローム情報研究チーム 研究員 福島 敦史
共起関係解析によるタンパク質の機能モジュール探索法の開発	九州工業大学大学院情報工学研究科生命情報工学研究系 助教 藤井 聡
タンパク質-糖鎖間の糖鎖結合部位の解明のためのツール改良及び解析	創価大学大学院工学研究科生命情報工学専攻 大学院生 細田 正恵
MicrobeDB.jp データを用いたメタゲノム解析 Web アプリケーションの開発	東京工業大学大学院生命理工学研究科 助教 森 宙史

情報技術 基盤技術開発プログラム DB統合化の実現に向けて基盤となる技術の開発を実施

小原 雄治 Yuji Kohara
 情報・システム研究機構 ライフサイエンス 統合データベースセンター センター長

課題名
 2011～13 データベース統合に関わる基盤技術開発

従来型の大規模集中型のデータベース統合ではなく、全く新しい「フェデレーション(連邦)型」のデータベース統合を行うために、RDFを中心とする技術を用いて、DDBJ や PDBj など国内の拠点データベース、統合化推進プログラムの分野別データベースなどを分散的に結ぶ新たな統合のシステムを検討し、統合化に必要な要素技術の開発を行っています。また、すでに蓄積の始まっている次世代シーケンサーデータをはじめとした大規模データの利用に関する技術開発やサービスの提供を実施、データの利用率を高めるためのさまざまな活動を行っています。

※2014年度からはNBDCとの共同研究として、「基盤技術開発」を継続します。

研究開発成果 統合TV <http://togotv.dbcls.jp/ja/> ライフサイエンス新着論文レビュー <http://first.lifesciencedb.jp/> GGRNA <http://ggrna.dbcls.jp/> Refex <http://refex.dbcls.jp/> Allie <http://allie.dbcls.jp/> TogoGenome <http://togogenome.org/> 他



統合のためにDBCLSが提供するサービス

統合(トータル)

生命知識の統合と発見