

ライフサイエンス統合データベース開発運用
(戦略立案・実行評価／統合データベース開発／統合データベース支援)
平成 22 年度研究成果報告書

平成 22 年 12 月

大学共同利用機関法人 情報・システム研究機構
高木利久

目 次

1. プロジェクトの目的	3
2. 戦略立案・実行評価（情報・システム研究機構）	4
2. 1 戦略立案・実行評価の実施計画	4
2. 2 戦略立案・実行評価の実施内容	5
2. 2. 1 戦略立案	5
2. 2. 2 実行評価	8
3. 統合データベース開発	13
3. 1 共通基盤技術開発	13
3. 1. 1 共通基盤技術開発の実施計画	13
3. 1. 2 共通基盤技術開発の実施内容	13
3. 2 ヒト統合 DB の開発・運用	17
3. 2. 1 ヒト統合 DB の開発・運用の実施計画	17
3. 2. 2 ヒト統合 DB の開発・運用の実施内容	17
3. 3 モデル生物・産業応用生物統合 DB の開発・運用	22
3. 3. 1 モデル生物・産業応用生物統合 DB の開発・運用の実施計画	22
3. 3. 2 モデル生物・産業応用生物統合 DB の開発・運用の実施内容	23
4. 統合データベース支援	23
4. 1 ポータル整備・運用、広報、普及啓発	23
4. 1. 1 ポータル整備・運用、広報、普及啓発の実施計画	23
(1) ポータル整備・運用	23
(2) 広報・普及啓発	24
4. 1. 2 ポータル整備・運用、広報、普及啓発の実施内容	24
(1) ポータル整備・運用	24
(2) 広報・普及啓発	26
4. 2 データベース受入	28
4. 2. 1 データベース受入の実施計画	28
4. 2. 2 データベース受入の実施内容	29
5. プロジェクトの総合的推進（情報・システム研究機構）	33
5. 1 プロジェクトの総合的推進の実施計画	33
5. 2 プロジェクトの総合的推進の実施内容	33
(1) 研究運営委員会及び研究運営委員会作業部会	33
(2) 関連機関との連絡会	35
5. 3 その他シンポジウム、セミナー等	35
(1) 統合データベース講習会	35

(2) BioHackathon	39
6. 成果の外部への発表	42
7. 特許出願状況	42
8. 整備実績一覧	42
9. 外部発表実績一覧	42
10. 実施体制	42

1. プロジェクトの目的

ライフサイエンスやバイオ産業に従事する研究者や技術者がいわゆるゲノムプロジェクト・ポストゲノムプロジェクトの成果や多様な DB や解析ツールをストレスなく利用してより高度な研究開発が効率よくできる環境（統合 DB）を実現することを目的とする。このため、情報・システム研究機構、科学技術振興機構、産業技術総合研究所生命情報工学研究センター、かずさ DNA 研究所、九州大学、奈良先端科学技術大学、東京大学、お茶の水女子大学、長浜バイオ大学は共同して、以下の3つの業務を行う。

(1) 戦略立案・実行評価

ライフサイエンス（基礎生物学および応用生物学（医・薬・農））、情報処理技術、ライフサイエンス DB の3分野の専門家による研究運営委員会を組織し、各々の分野の動向に即した我が国の DB 整備戦略を立案する。また、3分野の動向に関する情報の網羅的収集・分析を行い、戦略立案や実行評価を支援するチームを組織する。

(2) 統合データベース開発

ライフサイエンス、バイオ産業にかかわる情報へのアクセスと利用に関する格段の利便性向上とそれによる研究開発の飛躍的な効率化と質的向上を目指して、ヒト統合化 DB およびモデル生物・産業応用生物統合 DB の開発・運用と、それに必要な統合 DB 構築技術、DB の標準化、文献等からの知識獲得技術等の共通基盤技術開発を実施する。

(3) 統合データベース支援

個々の公開 DB や利用可能な DB 解析ツールなどを網羅的に調査し、提示するポータルサイトを構築する。運用にあたって専門家及び利用者の意見集約システムを開発し、意見をフィードバックして運用する。また日本語による研究情報の流通促進に注力する。プロジェクトデータベースを受入れ、相互運用可能な DB へ変換した上で公開し、運用する。また、広報・普及啓発の活動を行う。人材育成は実績のある機関と連携し階層別に、アノテータ・キュレータ教育、DB 構築者養成、及び DB 高度利用者養成を行う。

以上は、統合データベースプロジェクトの中核機関を構成する情報・システム研究機構以下の上記9機関が共同して実行すべき実施目的であるが、以下では、情報・システム研究機構の実施内容について記載する。

2. 戦略立案・実行評価

2. 1 戦略立案・実行評価の実施計画

ライフサイエンス（基礎生物学および応用生物学（医・薬・農））、情報処理技術、ライフサイエンス DB の 3 分野の専門家による研究運営委員会を組織し、各々の分野の動向に即した我が国の DB 整備戦略を立案する。上記委員会に対し、3 分野の動向に関する情報の網羅的収集・分析を行い、戦略立案や実行評価を支援するチームを組織する。このような体制のもとで、戦略立案、実行評価の支援業務とそのため情報収集・分析だけでなく、関係府省、利用者、産業界、学会・出版社、国内外の研究機関等関係者との連絡調整等の活動もあわせて行う。

(i)戦略立案

平成 21 年度までに、統合データベースプロジェクトの目標や年次計画の具体化、国内外の動向調査も踏まえたデータベース統合化のあるべき姿の検討を進めてきた。併せてデータ共有化や公開促進の観点から、国家プロジェクト成果に関する権利関係や著作権、個人情報保護に関する調査を行い、プロジェクト成果であるデータの提供を促進するための取り組みや提供いただいたデータの運用方針の立案、個人情報に関わるデータの公開方針の策定などを行ってきた。特に、平成 21 年度は東京大学グループと連携し、ヒトゲノム多型データ共有方針を決定し公開した。また、データ共有、著作権等に関するシンポジウムを開催し、その現状と重要性を訴えると同時に、将来の医療データ統合化のための調査、検討を行った。平成 22 年度は、これまでの成果をベースに、特に本プロジェクト終了後の進め方も念頭においた我が国におけるライフサイエンスデータベース統合の在り方の検討を進めると同時に、引き続き国家プロジェクト等により産生されるデータおよびその他有用な関連情報の共有および公開を実現するための方策を推進する。

(ii)実行評価

データベース基盤の構築・統合・活用にあたっては、産業界を含めたデータベース利用者の視点に立った取り組みが大変重要である。従って、想定されるユーザドメインから、有用な回答が期待できる有識者を選定し、本プロジェクト成果のユーザ評価を実施したり、出展した学会等の展示会でのアンケート結果も踏まえ、アクセスログ等によりユーザの利用状況を把握しながら逐次サービス内容の改善を行ってきた。平成 22 年度も、展示会等におけるアンケートの実施やアクセスログの解析などを通じてユーザの生の声の把握に努めて、引き続きサービス内容の向上に努める。さらに上記有識者を対象にプロジェクト成果全体に対するユーザ評価を実施し、本プロジェクトの総括を行うと同時に、ライフサイエンス分野の統合データベース構築に関する今後の進め方への指針を得る。

2. 2 戦略立案・実行評価の実施内容

2. 2. 1 戦略立案

プロジェクト発足当初、ライフサイエンス、知識情報処理、ライフサイエンスデータベース (DB) の 3 分野の専門家による研究運営委員会を組織し、DB 整備戦略について議論する場の構築と、上記委員会に対し、3 分野の動向に関する俯瞰データを提示し、適宜与えられる調査課題に対し答えることのできる体制(統合 DB 整備戦略作業部会)を構築し、戦略立案の基盤となる情報の収集・分析を行った。その結果、以下のような統合データベース構築の基本的アウトラインをまとめた。

- ・ DB 構築者ではなく利用者の思考や意思決定を支援する DB を構築する。
- ・ 利用者の興味、知識に応じて必要な情報、判断材料をもれなく提示する。
- ・ 複数 DB をつなぎ異種データ・知識の関係が俯瞰 (仮説生成) できるようにする。
- ・ いろいろなツールを簡便に組み合わせて解析 (知識発見) できるようにする。
- ・ できれば、上記のことが日本語で行えるようにする。
- ・ DB 化 (構造化) されないもの (論文 (テキスト、ポンチ絵、画像)、特許、教科書、報告書、解説記事、など) もうまく扱えるようにする。
- ・ DB 構築に最先端の研究開発は必要だが、あくまでも利用者の利便性向上のためのサービス事業であることを認識する。

これに従って、具体的な課題について各種データベースの統合化やサービスの実現方法そしてデータベースの受入れ戦略などについて各年度詳細な議論を行ってきた。また、各年度を通して、本プロジェクト終了後の我が国のライフサイエンス統合 DB の在り方に関する議論も行ってきた。統合化のスキームについては、主に検索のタイプから、DB カタログ、横断検索及び統合検索の三つのケースについて検討し、統合化の年次計画として、DB カタログについて我が国に存在するとされるゲノム、ポストゲノム関係の 250 の DB のほとんどを網羅することとした。横断検索については、20 程度の代表的な DB を掲載し、平成 20 年度中には我が国の関連 DB の大部分と海外の主要 DB を掲載することとした。また、統合検索については、国内 DB の受入れと併せて、順次対象となる DB 数を増やしていくこととした(図 2.1 参照)。さらに、契約・著作権並びに個人情報関係については、当該分野に詳しい弁護士のコンサルテーションを得ながら進めてきた。その中で、ライフサイエンス分野の国家プロジェクトからデータおよびデータベースの提供を受けるための方策について検討し、文部科学省との密接な連携の上で、プロジェクトの公募要領にデータおよびデータベース提供条項を盛り込むことができたのは大きな収穫である (図 2.2)。

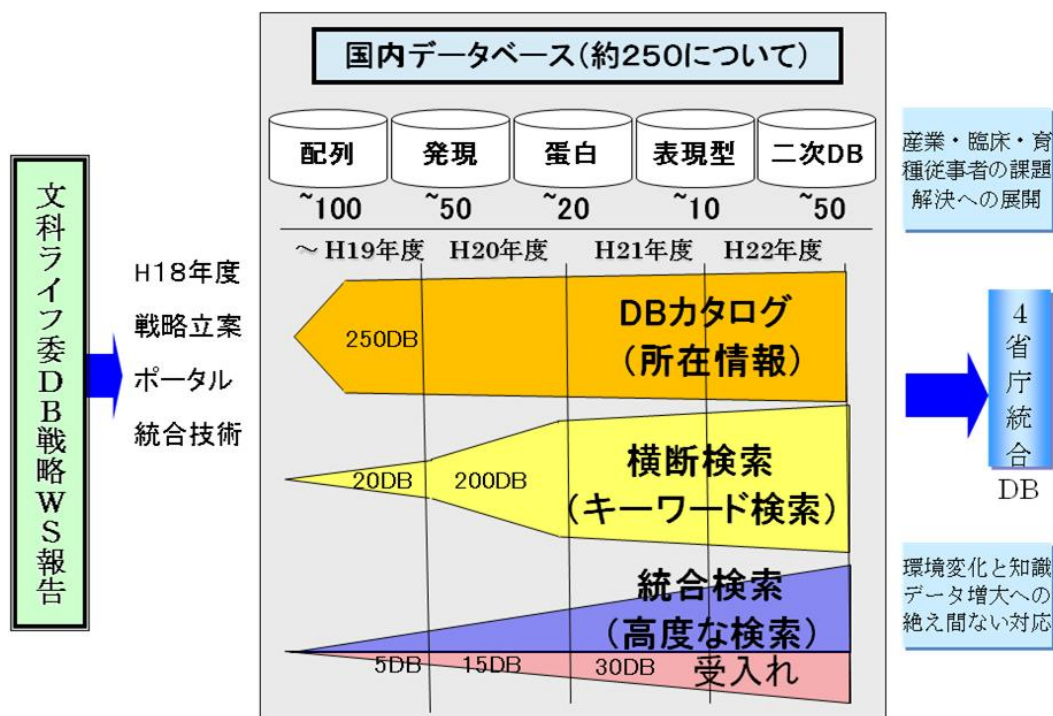


図 2.1 DB統合化の年次計画

(7) 統合データベースプロジェクトへのデータ提供

本プログラムはライフサイエンス分野のデータ産生を伴うプロジェクトであるため、本プログラムの実施者は、ライフサイエンス分野における研究に資するよう、本プログラムの実施の結果産生したデータ又はデータベースを、一般に利用可能な形で公開するべく努めるものとします。そのため、本プログラムの実施者は、毎事業年度にまとめる成果報告書の提出に合わせて、論文発表等で公表された成果に関わる生データの複製物、又は本プログラムで構築した公開用データベースの複製物を、ライフサイエンス統合データベースセンター[※]又は文部科学省の指定する機関に提供するものとします。提供された複製物については、ライフサイエンス分野のデータベース整備（統合データベース構築）に利用するため、これら複製物に関わる知的財産権を非独占的に複製・改変その他必要な形で利用できるものとします。なお、本プログラムの実施者は、複製物の提供を受けた機関の求めに応じ、複製物を利用するに当たって必要となる情報提供に協力するものとします。

※ライフサイエンス統合データベースセンター (<http://dbcls.rois.ac.jp/index.html>)

生命情報のデータ又はデータベースはライフサイエンス研究を支える基盤であり、その整備を推進するため、文部科学省では「統合データベースプロジェクト」を実施しています (<http://lifesciencedb.mext.go.jp/>)。ライフサイエンス統合データベースセンターは大学共同利用機関法人情報・システム研究機構に属し、当該プロジェクトを実施するための中核機関です。本センターでは、我が国のライフサイエンス関係データベースの利便性の向上を図るため、データベース整備戦略の立案・評価、データベース統合化の基盤技術開発、及び統合化データベース開発等を行い、ライフサイエンスデータベースの統合化を推進しています。

なお、本センターは、ライフサイエンスの基盤となるデータの保存と配布の役割を担うデータデポジトリのサイトであり、データ提供者と競合するような研究活動を行う機関ではありません。

(平成20年度以降の公募)

- データ/データベースの公開への努力
- データ/データベースの提供 (DBCLS or 指定機関へ)
- 非独占的な複製・改変権の担保
- 必要な情報の提供義務
- DBCLSと統合データベースプロジェクトの説明
- 提供者と競合しないことの説明

図 2.2 ライフサイエンス課委託業務公募要領のデータ提供条項

平成 22 年度は、平成 21 年 12 月に公開した、「疾患解析から医療応用を実現する DB 開発」ウェブサイトを通じた実際のデータ受け渡しに伴って顕在化した手続き上の不便を解消する手順を再整備し、書面もそれに合わせて改訂した。また、「ゲノム支援」のデータベース構築に関する打ち合わせにオブザーバーとして参加し、ヒトゲノム情報の共有化について関係者と意見を交わした。倫理の問題に関しては、先行している国外のゲノムやコホートのプロジェクト等において関連する部門でリーダーシップを取っている研究者が一堂に会する初めての国際会議が英国で開催されたので、DBCLS からも参加して各国の具体的な状況を知るとともに日本の状況についても議論する機会を得た。一方で、本プロジェクトサイトを利用した録画公開についての外部からの要請が増加していることに伴い、録画公開についてのポリシーを新たに検討して公開した。また、平成 22 年 10 月には 4 省（文科、厚労、農水、経産）間でのデータ共有に関して積極的な関係機関が共同で合同シンポジウム「ライフサイエンスの未来へ～10 年先のデータベースを考える～」を開催した。平成 21 年度、22 年度に開催したシンポジウムの発表資料等（要旨、ポスター PDF、講演録画など）はそれぞれのウェブサイトから検索・閲覧可能である（表 2.1 参照）。

表 2.1 シンポジウムウェブサイト

シンポジウム名	開催日	URL
「データベースが拓く これからのライフサイエンス」	平成 21 年 6 月 12 日	https://symposium.lifesciencedb.jp/2009/
「科学における情報の 上手な権利化と共有化」	平成 21 年 10 月 5 日	http://symposium.lifesciencedb.jp/IPDS/
「ライフサイエンスの未来へ ～10 年先のデータベースを考える～」	平成 22 年 10 月 5 日	https://symposium.lifesciencedb.jp/godo/

また、平成 20 年度以降のライフサイエンス課管轄の委託事業の公募要領には、図 2.3 で示したように、統合データベースへのデータ、データベース提供条項が盛り込まれている。平成 22 年度より、該当プロジェクトへのデータ提供の働きかけを開始した。前例のない試みであることもあり、必ずしもスムーズに進んでいるわけではないが、データ提供をいただく手順の整理も含め、方向性が見えてきている。表 2.2 に、対象となる 5 つのプロジェクトに関する状況をまとめた。

表 2.2 公募要領記載の「統合データベースへのデータ提供」に関するフォロー状況

事業名	対象年度	対象機関	状況(平成 22 年 10 月現在)
脳科学研究戦略推進プログラム	20 年度 21 年度	全機関	対象となるデータがあるかどうか事務局にて調査いただき、1 件データ提供いただける機関を紹介いただいた。現在先方機関との交渉準備中。
ターゲットタンパク研究プログラム	21 年度	阪大、理研	成果報告書の確定後、DBCLS が成果報告書を閲覧する旨、文部科学省から各機関に了解をとっていただいたうえで、対象データの存非を確認することになった。ライフサイエンス課取りまとめで成果報告書を入手予定。
革新的細胞解析研究プログラム	21 年度	全機関	事務局に確認、現在のところ対象データがないとの回答を得た。
ナショナルバイオリソースプロジェクト	21 年度	理研、NINS、横市大、筑波大	現在データ取りまとめ機関と相談後、個別機関へ DBCLS から対応することになった。2 機関からデータを入手済みで、現在公開準備中。残り 2 機関とは現在交渉中。
橋渡し研究支援プログラム	21 年度	全機関	成果報告書の確定後、DBCLS が成果報告書を閲覧する旨、文部科学省から各機関に了解をとっていただいたうえで、対象データの存非を確認することになった。現在、一部成果報告書を入手済み、内容確認中。

2. 2. 2 実行評価

本プロジェクトでは、公開されているサービスや取り組みについての説明や統合 DB の進め方について生命科学関係者から意見をうかがい今後の技術開発やセンター運営の参考にしている。取り組み内容の説明に加えてテスト検索が実行できるページや実施内容を表す模式図を見せるなどの工夫をしてサービス内容の紹介も兼ねたウェブサイト記入方式のユーザ評価のシステムを既に構築している（図 2.3 21 年度評価記入サイトの一部）。

平成 22 年度は、プロジェクトの最終年度であるため、全ての関連機関に対する評価を実行できるよう各機関にウェブサイトでの何らかの情報やサービスの提供を依頼し、機関ごと（DBCLS のみ実施内容が多岐にわたるため、サービス内容による 5 つの分類ごと）に評価ページを用意した。評価対象の内容は、プロジェクトの全期間に実施した内容の全てとし、参考資料としてプロジェクト全期間にわたる各機関の予算配分結果と平成 20 年度以降の各機関の年度目標と進捗のまとめ表を掲載し、前年の評価の際に要望のあった各サービスのアクセス数について（22 年度に新規に公開したサービスを除く）も参考情報として提供した。評価作業は 470 名余に依頼し、11 月半ばに終了し、評価サイトで公開した。今後は、寄せられたコメントへの対応等を進める予定である。

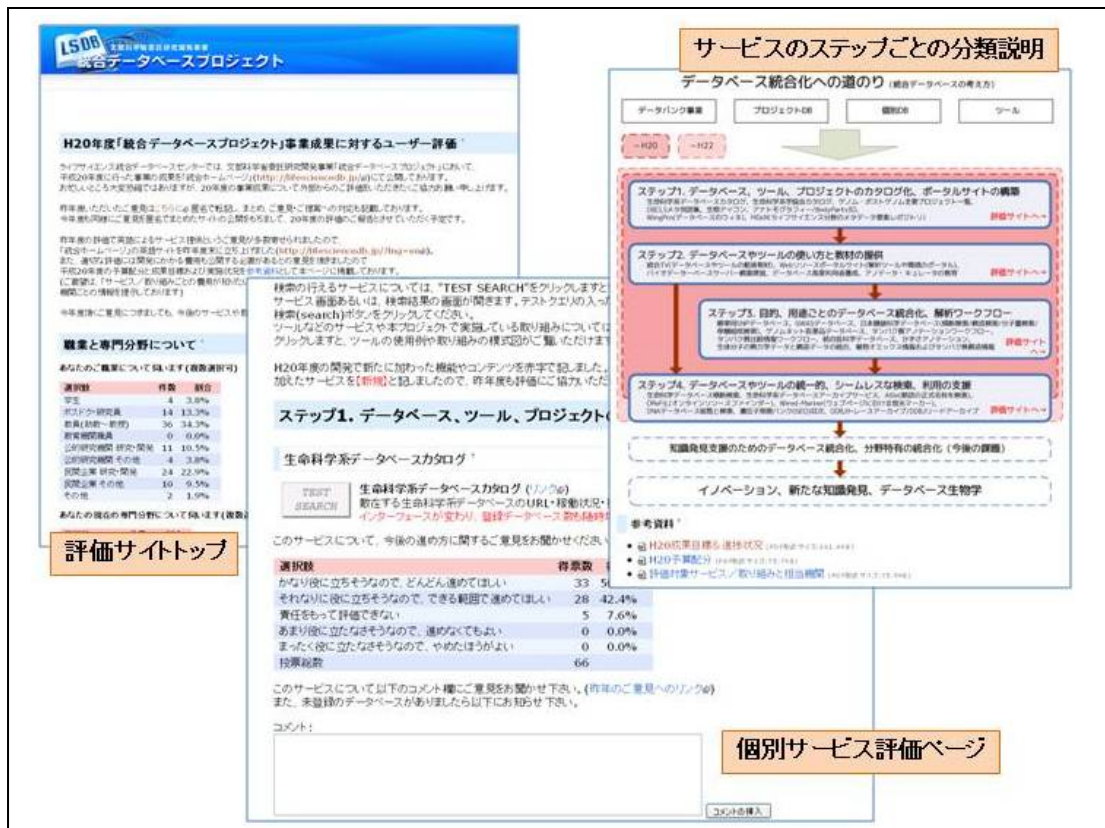


図 2.3 21 年度評価記入サイトの一部

表 2.3 評価結果公開サイト

対象年度	実施年月	結果公表サイト URL
平成 18	H19 年 3 月	http://lifesciencedb.jp/feedback/hyouka18/hyouka_H18.html
平成 19	H20 年 8 月	http://lifesciencedb.jp/feedback/hyouka19/?result
平成 20	H21 年 10 月	http://lifesciencedb.jp/feedback/hyouka20/?result
平成 18-22	H22 年 11 月	(評価実施中)

参考) 外部評価のページ <http://lifesciencedb.jp/l5db.cgi?gg=hyouka>

一方、各サービスのアクセスログの解析については平成 20 年度から開始しており、解析方針は以下の通りで、基本的には現在も変わっていない。

- 解析ソフトは AWStats ver.6.6 (<http://awstats.sourceforge.net/>)を採用し、サービス開発者の IP アドレスを除外する以外は、デフォルト設定のまま利用した。AWStats が出力する主な集計値は以下の 3 つである。
 - 一意な訪問者 IP: アクセス元の IP アドレス数。同一 IP からの複数回のアクセスは 1 と計数。
 - 訪問回数 (セッション数): アクセス元の IP アドレス数。ただし 1 時間以上の間隔をおいての同一 IP からのアクセスは別訪問と計数。

▶ ページ数（いわゆるページビュー）：閲覧ページ数。

- 統合 HP(<http://lifesciencedb.jp>)のポータルにリストされている項目（サービス）毎に月次集計。
- 毎日自動更新し、本センターの関係者が常時 AWStats の結果を閲覧できるようにする。

平成 22 年度は、プロジェクト全期間にわたる成果についてのユーザアンケートとアクセス解析等により提供サービスの評価を行う段階に入った。10 月に実施したユーザアンケート資料にはサービス単位の訪問回数とページ数を表示し、評価指標のひとつとなった。作業部会や運営委員会の報告、研究者による学会発表にもこのアクセスログ解析結果は多数引用されている。利用目的が明確化しつつあり、それに沿った利用しやすい形式で解析結果を提供するのが今後の課題である。

表 2.4 ライフサイエンス統合データベースセンター全体の集計結果

	平成 20 年度	平成 21 年度	平成 22 年度 (10 月末現在)
一意な訪問者 IP	231,305	439,119	310,736
訪問回数	436,203	942,174	721,858
ページ数	1,765,823	6,350,496	4,891,269

本センター全体の集計値は表 2.4 のとおりである。サービス別訪問回数の合計であるため、例えば 1 ユーザが 1 回の訪問で 2 つのサービスを訪問した場合は 2 と計数している。また、一意な訪問者 IP 数も月次の合計のため、同一の IP から異なる月に 2 回訪問した場合は 2 となる。図 2.4 のヒートマップは、本センターの提供するサービス全体を時系列かつサービス別に俯瞰するためのものである。



図 2.4 ライフサイエンス統合データベースセンターが提供するサービス全体のヒートマップ

サービス（アクセス数降順）

横軸は時間軸、縦軸はサービス軸で最大月次アクセス数の降順にソートしている。1行が1サービスに対応しており、赤いセルはそのサービスでのアクセス数の月次最大値、緑は最小値、黄は中間の値を示している。ヒートマップの右上が赤いことから、アクセス数の多いサービスほど順調にアクセス数を伸ばしているといえる。反対に、右下の一部のみが緑であることから、アクセス数が減少しているサービスは少数であり、総アクセス数への影響も少ないことがわかる。しかも、更新が終了した過去のコンテンツとして掲載しているものであるため、更新を継続しているサービスはアクセス数が増加している。なお上のマップでは一意な訪問者 IP 数をアクセス数としたが、訪問回数、ページ数の場合もほぼ同様の傾向であった。

アクセス数上位のサービスは表 2.5 の通りである。サービス開始時期の違いや月毎の変動を考慮して、サービス開始以降の月次最大値を代表値とした。一意な訪問者 IP 数をユーザ数とみなすと、表 2.5 から独自コンテンツを提供しているサービスがユーザを獲得しているといえる。トップの生命科学データベース横断検索は独自コンテンツではないが、内訳は蛋白質核酸酵素バックナンバー日本語 PDF ファイルを出力するリクエストが大部分を占めている。表 2.6 もほぼ同様の傾向を示している。

表 2.5 一意な訪問者 IP 数の月次最大値上位 10 サービス

サービス名	開始時期	月次最大値
生命科学データベース横断検索	2007/10～	17,741
DNA データベース総覧と検索 (DDBJ/EMBL/GenBank)	2007/10～	7,015
蛋白質核酸酵素 全文検索	2008/5～	4,904
文科省「ゲノム」研究報告書 全文検索	2008/5～	4,609
統合 TV(DB やツールの動画教材)	2007/7～	4,326
アノモグラフィー/BodyParts3D	2007/10～	3,879
ライフサイエンス新着論文レビュー	2010/8～	3,645
統合ぐらし	2007/9～	3,128
シンポジウム・講演会	2009/4～	2,033
Gendoo	2007/10～	1,888

表 2.6 訪問回数の月次最大値上位 10 サービス

サービス名	開始時期	月次最大値
生命科学データベース横断検索	2007/10～	32,474
DNA データベース総覧と検索(DDBJ/EMBL/GenBank)	2007/10～	17,890
統合 TV(DB やツールの動画教材)	2007/7～	15,425
蛋白質核酸酵素 全文検索	2008/5～	11,758
ライフサイエンス新着論文レビュー	2010/8～	11,294
文科省「ゲノム」研究報告書 全文検索	2008/5～	6,786
誰でもデータベースが構築できる TogoDB	2008/4～	6,181
Gendoo	2007/10～	6,082
統合ぐらし	2007/9～	5,941
アナトモグラフィー/BodyParts3D	2007/10～	5,449

表 2.7 のページ数は、ユーザ数というよりはむしろリクエストを処理した回数とみなせるため、表 2.5、2.6 とは多少異なる傾向を示している。ページ数の多少はシステムの構築方法にも依存するとはいえ、上位にはコンテンツ提供サービスだけでなく、解析処理を行う TogoWS や MiGAP のようなサービスも現れる。バッチ処理等で 1 度の訪問で多数のリクエストを処理しているものと思われる。

表 2.7 ページ数の月次最大値上位 10 サービス

サービス名	開始時期	月次最大値
国内ウェブサービスの統合 TogoWS	2008/2～	399,873
DNA データベース総覧と検索(DDBJ/EMBL/GenBank)	2007/10～	380,834
アナトモグラフィー/BodyParts3D	2007/10～	248,853
誰でもデータベースが構築できる TogoDB	2008/4～	114,779
統合 TV(DB やツールの動画教材)	2007/7～	77,700
生命科学データベース横断検索	2007/10～	52,946
遺伝子発現バンク(GEO)目次	2007/10～	51,250
MiGAP	2009/4～	46,700
ライフサイエンス新着論文レビュー	2010/8～	44,484
Gendoo	2007/10～	25,003

表 2.8 はアクセス元のドメインの内訳である。プロバイダなど組織を特定できないドメイン(net/ne/com)からのアクセスが大きく増加している。教育機関(ac/edu)、政府系機関(go/gov)については実数は増えているにもかかわらず、割合は横ばいもしくはやや減少して

いる。民間企業(co)についても減少しているが、その他に分類している com ドメインにも民間企業が含まれておりやや低めの数字となっている。

表 2.8 アクセス元ドメインの内訳

ドメイン種別	平成 20 年度	平成 21 年度	平成 22 年度 (10 月末まで)
教育機関(ac/edu)	18%	10%	13%
民間企業(co)	13%	5%	2%
政府系機関(go/gov)	4%	1%	2%
インターネットプロバイダ等(ne/net)	23%	57%	54%
その他(com など)	42%	27%	29%

3. 統合データベース開発

3. 1 共通基盤技術開発

3. 1. 1 共通基盤技術開発の実施計画

平成21年度に開発した生命科学研究者の興味を反映した文献情報の検索、整理が容易にできる環境 TogoDoc について、平成22年度は機能向上のための開発を行う。平成21年度まで開発してきたインターフェースDBCLS Galaxy について、平成22年度はツールの追加および改良を行う。また、平成21年度まで開発してきた TogoDB、TogoWS などのサービスに対し、平成22年度はユーザビリティを高めるための諸開発を行い、その成果を現在稼働中のサービスに反映させる。さらに、統合検索に向けた基盤技術として、セマンティックウェブ技術に関する調査を行い、必要となる技術を開発する。

3. 1. 2 共通基盤技術開発の実施内容

本課題では、研究者が多様な生命科学 DB や解析ツールをストレスなく利用できるための基盤技術を提供することを目標とし、様々なサービスを開発してきた(図 3.1)。それらは、文献執筆支援や文献情報からの情報抽出を行うテキストマイニングサービス群、簡単に誰でもデータベースが構築可能となるサービスである TogoDB (<http://togodb.dbcls.jp/>)、国内外の主要な DB や TogoDB のウェブサービスを統一的インターフェイスで利用できる TogoWS (<http://togows.dbcls.jp/>)、ウェブサービスやツールを組み合わせる解析ワークフローを作成できるサービス DBCLS Galaxy (<http://galaxy.dbcls.jp/>)、シングルサインオンサービス DBCLS OpenID (<http://openid.dbcls.jp/>)から成り、互いに連携して、様々な研究スタイルをもつ生命科学研究者が DB を利用するための基盤サービスとなっている。ま

た、これらのサービスを開発するにあたり、国際的動向の調査、国際的連携と標準化、それに伴う開発が必要であり、それらを行うために、毎年、国際開発会議 DBCLS BioHackathon を開催してきた。

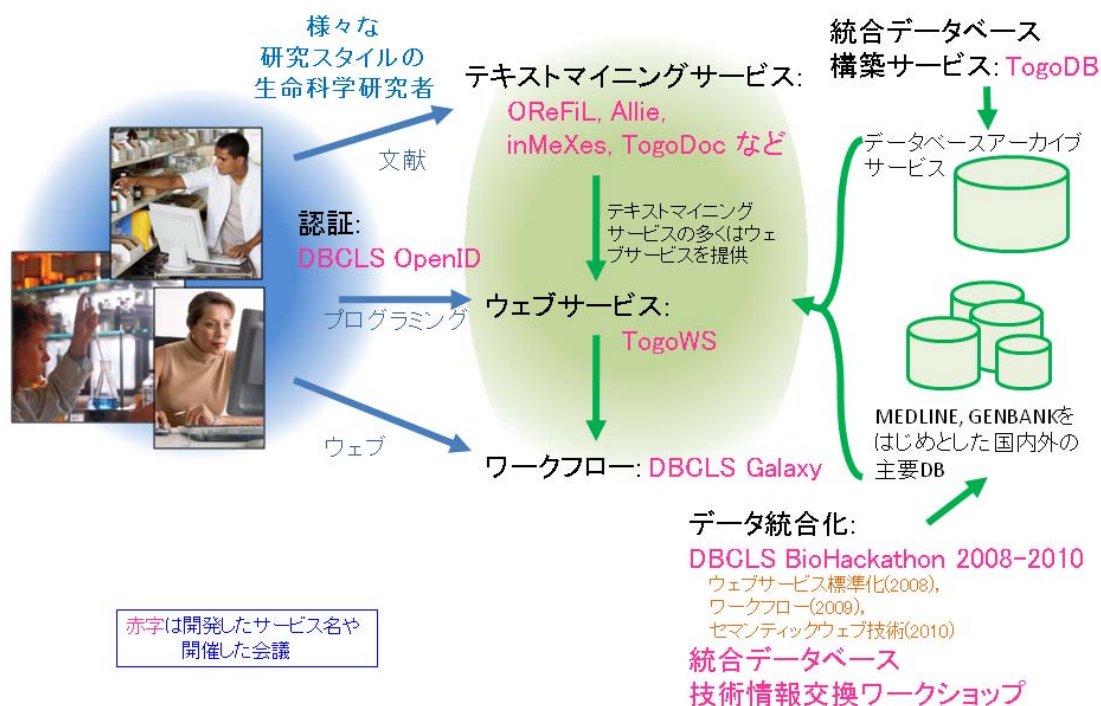


図 3.1 共通基盤技術開発によって提供されるサービスと開催した会議

現在一般公開しているテキストマイニングサービスは、OReFiL (<http://orefil.dbcls.jp/>)、Allie (<http://allie.dbcls.jp/>)、inMeXes (<http://docman.dbcls.jp/im/>)、TogoDoc (https://docman.dbcls.jp/pubmed_recom) の 4 つから成る。Allie、inMeXes、TogoDoc の三つのサービスは文献執筆支援を目的に開発され、具体的には、Allie は生命科学における略語とその展開形の検索システム、inMeXes は生命科学文献における英語表現検索システム、TogoDoc は生命科学分野の書誌情報と論文(PDF ファイル)を管理するためのサービスとツールである。TogoDBはDB受入のためのテンプレートシステムとして開発された。表形式のデータを CSV 形式でインポートすることで、レコード表示、リスト表示、検索機能をもったウェブデータベースの公開・管理を誰でも手軽に行えるシステムである。TogoWS は国内主要拠点のデータベース (DDBJ、PDBj、KEGG) や海外の主要拠点データベース (NCBI、EBI) を仮想的につないで、統一的なインターフェイスを提供するサービスである。

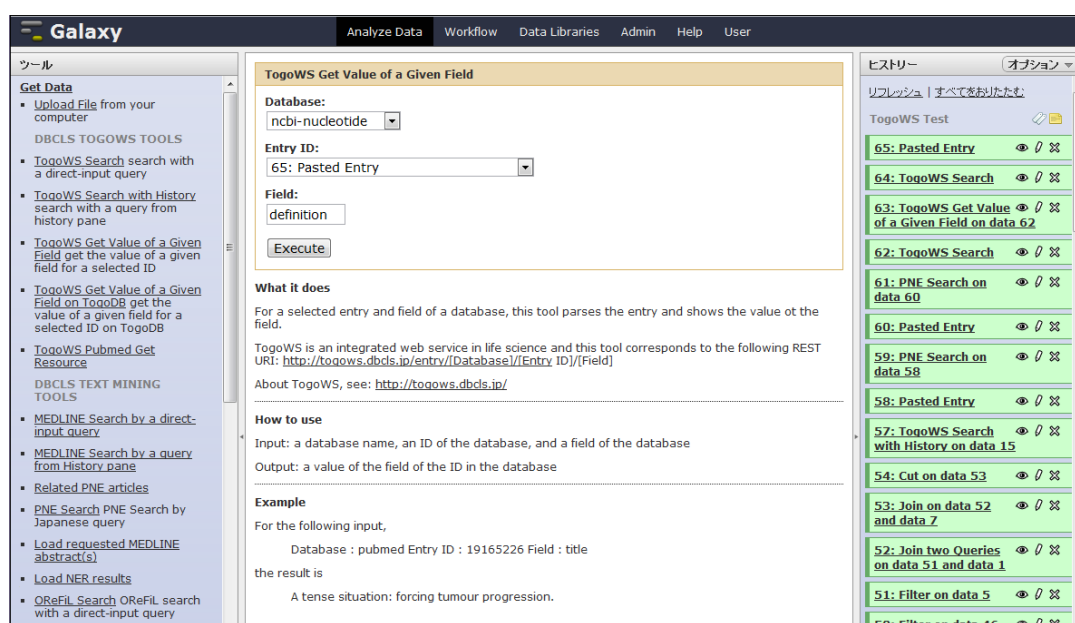


図 3.2 DBCLS Galaxy: 生命科学データ解析のためのワークフロー作成・共有サービス

DBCLS Galaxy はウェブブラウザ上で複数のツールを組み合わせて解析ワークフローが容易に作成可能なサービスである(図 3.2)。これまでに、TogoWS のサービスの一部を Galaxy に組み込んだプロトタイプを開発し、DBCLS Galaxy での解析結果を TogoDB へ出力するツールが組み込まれ、上記の TogoDB の DBCLS Galaxy からのデータの入力機能と繋げて、DBCLS Galaxy で解析した結果をスムーズに TogoDB から公開できるようにした。また、セマンティックウェブ技術を用いた知識ベース freebase (<http://www.freebase.com/>)からの RDF データ取得など、セマンティックウェブ技術関連のツールの組み込み開発を行い、平成 22 年度には、統合 DB で開発した横断検索の組み込みや、前年度に引き続いてセマンティックウェブ技術関連のツール組み込みを拡充し、さらに、日本語化および日本語ドキュメントの作成を進めた。

DBCLS OpenID は一つの ID で複数のサイトの認証が可能なシングルサインオンサービスである(図 3.3)。標準規格 OpenID 1.1 に準拠し、その認証プロトコル仕様に沿って認証サーバを構築した。この開発により、統合データベースプロジェクトに関連する認証をこのサーバで一元的に行うことが可能となり、複数のサイト間を行き来するユーザの負担を減らすことができるようになった。平成 20 年度には OpenID 2.0 にも対応し、より幅広いユーザの利用を可能とした。現在、DBCLS OpenID は、統合データベースプロジェクトのページをはじめ、TogoDoc、TogoDB、DBCLS Galaxy、TogoWiki、かずきアノテーション、かずきナビゲーション、MiGAP など多くのサービスで採用されており、ユーザはこれらのサービスを一つの DBCLS OpenID アカウントで利用可能となっている。



図 3.3 DBCLS OpenID: 複数のサイトにログイン可能な認証サービス

上記の諸サービスの開発と並行して、国際的な動向・現状調査と連携した技術開発を目的として、国際開発会議 DBCLS BioHackathon をこれまでに三回開催し、各トピックのグループ、特に、ウェブサービス設計指針を改良するグループ、Galaxy の国際化や今後の展開について議論・開発するグループ、次世代シーケンサからの大量データをどう扱うかを議論するグループ、大量データの可視化について議論・開発するグループ、マイクロアレイデータ解析について議論・開発するグループ、実際の生物系データを持ち寄ってユースケースについて議論するグループ、セマンティックウェブをどう実装すべきか議論するグループ、文献情報のツール群について議論するグループ毎に熱心な議論が行われた。これらの議論は、共通基盤技術開発の諸サービス、特に TogoWS と DBCLS Galaxy の開発に大きな影響を与えた。更に、平成 21 年度に第三回目として開催された「セマンティックウェブ技術」をテーマにした会議では、国内外のデータプロバイダーの実務者レベルでの利用技術の観点における合意形成がおこなわれた。それに基づいて、今後も国際協力的に、データ統合化やそのサービスの開発を進めて行くこととし、データの応用部分での議論に基づいて開発を行い、ユーザインターフェイス部分やデータ統合化についてのいくつかの新しい技術的開発的取り組みも実施した。ここでの議論は平成 21 年度から 22 年度にかけて行われた TogoDB、TogoWS、DBCLS Galaxy でのセマンティックウェブ技術関連の開発を大きく促進した。平成 22 年度は、これまで得てきた国際技術動向や知見を国内で共有する必要があると考え、国内のデータベース構築者・利用者を対象に、統合デー

データベース技術情報交換ワークショップ (<http://wiki.lifesciencedb.jp/mw/index.php/BH10.10>)を行った。国内から70名弱の参加者があり、活発な議論と情報交換がおこなわれた。三回のDBCLS BioHackathonおよび統合データベース技術情報交換ワークショップでの議論と成果はそれぞれウェブ上でまとめられており閲覧が可能となっている。

3. 2 ヒト統合DBの開発・運用

3. 2. 1 ヒト統合DBの開発・運用の実施計画

平成21年度までに、辞書シソーラス類の整備、キュレーションのための文献解読手法、ツールの開発、バンク目次システムの構築を行ってきた。特に平成21年度は、ボディパーツ3Dの胸部臓器、筋肉の充実を図ると同時に、生物名称、学術用語の大幅な拡充を行い、その成果を公開した。平成22年度も引き続き医学研究や創薬開発研究に携わる利用者がゲノム情報、分子データベース、文献情報等の、研究開発に必要な情報を容易に統合利用できるように、背景知識の整理を継続辞書やオントロジーの形で追加公開する。この辞書オントロジーを利用して事業で構築したサービス（DNAバンク目次と遺伝子発現バンク他）を融合利用可能にする。アナトモグラフィ・ボディパーツ3Dの開発も継続し上記サービスとの連携を図る。

3. 2. 2 ヒト統合DBの開発・運用の実施内容

より多種類の分子データや文献データを統合利用するにはデータの種類の依存しない上位のレベルでの共通の索引付けや分類整理が必要である。日付著者アドレス引用関係等を一階層医学生物学に下ると、生物種、解剖、疾患、分子遺伝子名称、化合物薬品、操作検査手技、研究分野などがこの上位レベルである。これは文献整理のためのキーワード群と同じであるが米国以外ではこれらの概念（用語）の収集が未発達であり、また膨大な文献や実験データのメタデータ部分に自由に記載された内容へのマッピング技術も未整備である。日本語を置き去りにせずこの課題の解決に向けて実証的サービスも構築と基礎資源の作成をおこなった。フェーズの違う産物が散在するが、ゴールは未統制用語による表現を操作可能にできる十分な網羅性の用語集、その意味による整理を行うシソーラス、であり、さらに分類表現を理解や利用に繋げるための情報ハブ表現の開発である。

材料データとして DDBJ リリースファイル (<ftp://ftp.ddbj.nig.ac.jp/>、NCBI GEO メタデータ <ftp://ftp.ncbi.nih.gov/pub/geo/> をおよび GenBank 生物分類ルールを利用し研究分類のためのフィールドの値および埋まり具合に基づく分類ルール、国別分類のための頭文字表記を含む世界の都市名の政治的國家名対応ルール、動物解剖分類のための正規表現による頻出臓器細胞名称と臓器分類の対応ルール、同様な植物解剖分類ルール、およ

び別項で報告する生物名称日本語辞書を開発利用し提供してきた。

その結果 1) 配列データ、遺伝子発現データともに各種の分類ルールによる多元独立自動分類を行い、シソーラス的に上位からたどることで自分が排除している対象を意識しながら関心プロジェクトに到達しデータを取得することを可能となった。

2) キーワード検索や配列検索の後で得られるレコードを同様に検索後分類が可能であり、多数のヒットを上位概念から絞り込むことが可能となった。

3) さらに自動分類に従った各種の統計によるデータ全体の俯瞰的表現を可能にした。俯瞰表現には表以外に **googlemap** および **anatomography**, (内部開発別項)生物アイコン (別項) を用いている。

一方、我が国では昭和初期より継続的に学術用語や生物学名を個別の学会が主体となって収集更新しているがそのほとんどが書籍として販売され統合利用や転用が困難となっている。また JICST (JST) および NII はそれぞれの事業の為に保有していると思われるシソーラスなどはその契約上や事業モデル上外部に利用許諾を出すことが困難で当然そのサイズや機能を客観的に評価した例も見当たらず個別の用語の出典情報も記載がなく不整合時に正当性を検証する手段もない。そこで、本プロジェクトでは上記の例に学び主に書籍を情報源に現行法制度下において誰もが自由に転用利用できることを命題に、「転用可能な巨大言語資源」としてメタ用語集を編纂した。メタ用語集では英語、ラテン語などの科学基本用語 (基本概念) との対で日本語表記が収集されているのでオントロジーや分類学などの構造化は外部資源に依存して自由に行うことができる。材料は文部科学省学術用語 1 1 篇を含む書籍 3 1 巻および **wikipedia** を含む 3 2 の **web** サイトを情報源として用い、電子化したのちに創意著作権のないファクトデータのみを下記の一定形式で抽出する手作業およびパースを行った (<http://lifesciencedb.jp/bdls/90326357>)。

辞書や用語集およびウェブサイトはカラム構造を用いた表形式で語彙およびその意味が提示されており、人間には容易に意味が理解できるが、複数の書籍の間では同じカラム見出しの意味やその種類、行であらわされる集積単位 (概念) がそれぞれに異なっている。本プロジェクトでは、編纂中に意味粒度のまるめや不整合の断りない修正が行われないうに、第一段階として電子化と形式の統一のみを考え、可逆的にオリジナルデータを再現可能な形式に転載することを行った。その形式とはテーブルの一つのセル中の値を文献 X の表 Y の行 N では列 M の値は **Term·T** であるという一文であらわす単文形式であり、容易にオリジナルデータを再現することが可能である。複数のテーブルを一つの単文集形式で集めたのちに「表 Y の列 N と表 Z の列 NN は同じ意味カラムである」などのルール集を別途明示的に作成適用することで転用を一層容易に出来る。

@Aartsenia japonica@	is @学名@	for @Aartsenia japonica@	in @東海大学出版会@	@日本近海産貝類図鑑@	version @第 1 版第 1 刷@
@コシダカドウマルクチキレ@	is @和名@	for @Aartsenia japonica@	in @東海大学出版会@	@日本近海産貝類図鑑@	version @第 1 版第 1 刷@
@Abacetus leucotelus@	is @学名@	for @Abacetus leucotelus@	in @北隆館@	@新訂 原色昆虫大図鑑 第 II 巻 (甲虫 篇)@	version @平成 19 年 5 月 11 日@
@ホソツヤナガコムシ@	is @和名@	for @Abacetus leucotelus@	in @北隆館@	@新訂 原色昆虫大図鑑 第 II 巻 (甲虫 篇)@	version @平成 19 年 5 月 11 日@
@Abacetus tanakai@	is @学名@	for @Abacetus tanakai@	in @北隆館@	@新訂 原色昆虫大図鑑 第 II 巻 (甲虫 篇)@	version @平成 19 年 5 月 11 日@
@ヨツボシツヤナガコムシ@	is @和名@	for @Abacetus tanakai@	in @北隆館@	@新訂 原色昆虫大図鑑 第 II 巻 (甲虫 篇)@	version @平成 19 年 5 月 11 日@

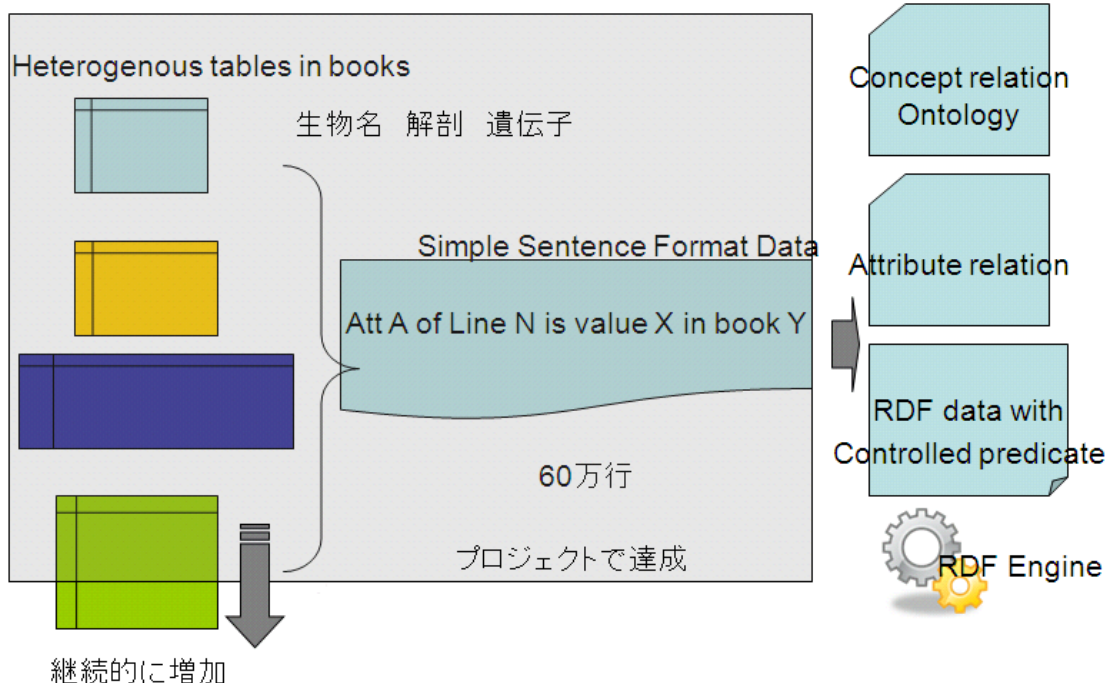


図 3.4 単文表現の例とメタ用語集プロジェクトの概要

これまでに 100 万行を超える単純文データを集め、Species2000, NCBItaxonomy, Nomina anatomica, Foundation Model of Anatomy など各種の世界規模での用語集とマップすることで常に概念単位の数とカバー率を測定している。生物名称に関して日本語書籍で与えられたラテン語と Species2000 や NCBItaxonomy との対応が予想外に悪く、学名の更新によるものやラテン語表記の語尾変化によるものなどの原因が多く、これら世界資源へのマップは未着の技術的課題として残された。マップの結果まとめは以下の URL を参照されたい。 <http://lifesciencedb.jp/bdls/8abf67fb72b66cc153ca7d50679c>

解剖学用語は医学系の文献、分子データ、医療記録等を統合整理するためのもっともカバー率の高い索引用語である。しかしながら、用語の関係に空間的な関係を多く持つためにオントロジー等のシンボル関係への表現が難しく、またシンボルからヒトの理解への整理表現手法としての能力も劣る問題がある。そこで、BodyParts/Anaomography はシンボル表現と利用者のインタフェイスとして計画された。解剖学概念集である Nomina Anatomica (7700 語) に対してすべて個別に人体モデル中でのセグメントを与え、さらに米国で発達中のシンボル表現 Foundation Model of Anatomy を利用してこれらの立体を概念的に離散会合可能にすることを目標にした医学系情報のハブである。関連分野の画像診断領域では CT や MRI などの断面像の濃度差で与えられる臓器輪郭の抽出課題の解決レベ

ルが飽和し、臨床現場でも個人の CT や MRI 画像の輪郭を 3 次元合成した 3D データは容易に得られる。最近はその延長として **Computer Asisted Diagnosis** と呼ばれる課題が注目されており、輪郭抽出 臓器同定から異常画像（器質的疾患）の検出や同定を目指した研究が盛んに行われている。診断以外の目的で解剖学的に精緻な 3D モデルの構築に関しては **VisibleHumanProject(NIH 1990-)** による特定の検体の詳細な固定断面カラー画像に由来するものが有名であるが、主にこれも輪郭抽出の課題としての研究が盛んに行われた。そのほか教育コミュニケーション用メディカルアートとしてのカノニカル画像集や 3D 人体モデル表示サービスもいくつか存在する。これらは教育目的や手術シミュレーションの背景として主に商業ドメインで応用が行われている。

一方、脳に限定すれば関心領域(ROI)の共通表現や操作の為に標準座標系を導入する考えは 1950 年代に始まる定位脳手術に源が求められる。この目的のモデルは標本脳に規準点とグリッドをはじめ持ち込んだタライラッハアトラスに始まる (**J. Talairach and P. Tournoux, "Co-planar Stereotaxic Atlas of the Human Brain1988**)。この 70 才の女性の切片アトラスに基づく座標系は後に fMRI などを使う機能画像診断分野でも広く使われるようになる。タライラッハアトラスは電子化共有され **BrainMap** などのニューロイメージングデータの統合サイトでも座標系として利用されている。このような展開は **BodyParts3D** プロジェクトの意図するところと共通である。最近では詳細な MRI 画像を用いた **Harvard Surgical Planning Laboratory (SPL)** の分節脳 MRI3D データとそのイメージ視覚化スイート"**Slicer**"(2008)などが多目的に利用されている。しかしながら、解剖学全体をカバーする精緻で分節数の大きなデータはパブリックドメインには存在せず、情報の整理統合共有目的には利用できない。

情報ハブとして 3D モデルを用いる例は最近急に海外で散見されるようになった。出版されている例では “**Spatiotemporal integration of molecular and anatomical data in virtual reality using semantic mapping**” Sun Center of Excellence for Visual Genomics, University of Calgary, Calgary, AB, Canada; *Int J Nanomedicine*. 2009; 4: 79-89. のみであるが、IBM のプレスリリースは広く知られている“**3D Avatar to Help Doctors Visualize Patient Records**” <http://www-03.ibm.com/press/us/en/pressrelease/22375.wss> である。図 3.5、3.6 には **BodyParts/Anatomography** の特徴と他競合プロジェクトとの比較および分散統合と集中統合の比較をそれぞれ示している。

BodyParts: 解剖学オントロジー Foundation Model of Anatomy (FMA 75,000 concepts)のうち解剖学会が認める7000あまりの概念(Terminologia Anatomica)を3次元人体モデル中のセグメントとして定義します。シンボル表現とモデル表現間の対応を行うことで以下の利点があります。

- オントロジーが立体データに与える便益
 - 1) 点線面空間などを区別して取り換える
 - 2) セグメントの集合としてカテゴリー概念も表現できる (ex. 副腎=右副腎+左副腎)
 - 3) 未定義の概念も定義済み概念の節分として取り換える
- 立体定義がオントロジーに与える便益
 - 1) ほとんど説明的な定義を与られていないFMAやTAに対して見える定義を与えられる
 - 2) トップレベルが難解で階層が深く探索しづらいオントロジーからの概念探索のインターフェイスを与える

Anatomography: web-レンダリング serverでAvatarをmush-up素材として提供します

2007年の開始時には同じ目的のプロジェクトは不在でしたが、2010年現在はIBM's Zurich Research LabのAnatomic-Symbolic Mapping EngineとCalgary大 Sun center of excellence visual genomics laboratoryのCavemanが類似目的のプロジェクトとして進行中です。比較すると特徴は以下ようになります。

DBCLS:	分散型統合	FMA	カノニカル	CC-BY
Sun micro CEO	集中型統合	SNOMED-CT	Visible Human	独占
IBM's Zurich :	集中型統合	SNOMED-CT	Visible Human ?	独占

図 3.5 BodyParts/Anatomography の特徴と競合プロジェクトとの比較

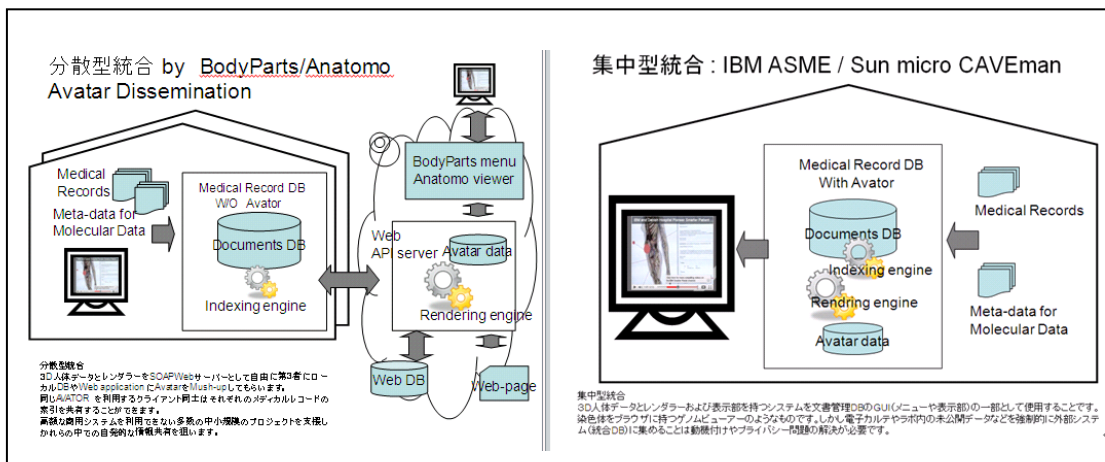


図 3.6 医療医学情報の分散統合と集中統合の比較

Aatomography は BodyParts データを呼び出してユーザの指示によって著作権のないイラスト画像(Avatar 像)を作るためのレンダリングサーバである。分散型 Avatar の特別な特徴としておもに API 経由での利用によるオンザフライでの描画を企図したサーバーであり、メニュー画面を利用して作成した Avatoar はパーツ ID や設定パラメタを記載した RestfulURL として出力され保存または埋め込みされることを前提としている。Web テキスト、個人ブログ、メール添付、データベースとして外部でも利用されることを前提と

し、負荷分散装置を備えたレンダリングサーバーを構築した。

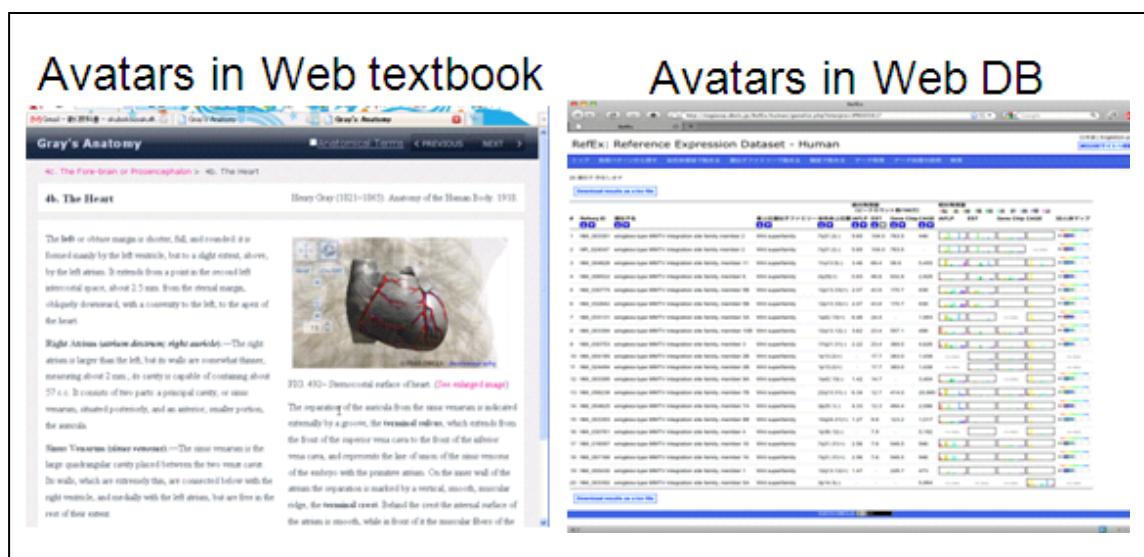


図 3.7 オンザフライで Avatar を表示した例 これらはhtml 文であり画像データは持たない。 http://221.186.138.155/project/anatomo5/demo_graysanatomy2/index.html

3. 3. モデル生物・産業応用生物統合DBの開発・運用

3. 3. 1 モデル生物・産業応用生物統合 DB の開発・運用の実施計画

主として微生物を対象とする基礎生物学研究および生物の産業応用研究に携わる利用者がゲノム情報、分子データベース、文献情報等の、研究開発に必要な情報を容易に統合利用可能とする環境を整備する。平成20年度までに、微生物情報資源のサイトの情報を収集しデータベースカタログに登録して統合データベースのポータルサイトから利用可能とした。また、平成20年度から我が国の微生物産業と新世代シーケンサの普及を想定して、これからの微生物ゲノム解析の標準を提供すべく微生物ゲノム解析パイプライン、Microbial Genome Annotation Pipeline (MiGAP)の開発に着手した。平成21年6月に一般公開したMiGAPのベシック版 (b-MiGAP) と利用者が解析パラメーターを調整可能な版s-MiGAPを限定公開した。さらに、利用者の自由度が高いg-MiGAPを開発した。平成22年度は、これらの新しい版をより高性能なサーバで安定に一般に公開する。さらに、真核微生物のゲノムアノテーション機能を実現して、微生物ゲノムアノテーションパイプラインとして一応の完成とする。MiGAPによって利用者は、微生物ゲノムの遺伝子領域やRNA領域の観点から、参照データベース群に存在する関連情報を統合的に参照可能となる。

3. 3. 2 モデル生物・産業応用生物統合 DB の開発・運用の実施内容

MiGAPの利用実績が高まってきたため、解析サーバを国立遺伝学研究所の大規模PCクラスターの一部に移行して、複数ゲノムの並行解析を可能とした。また、既知の情報が蓄積されているデータベースを分散させることによって、解析対象配列からの高速な参照を可能とした。解析の高速化に加えて、真核微生物ゲノムのアノテーション機能も加えた。これらによって、利用者の利便性が大幅に高まった。

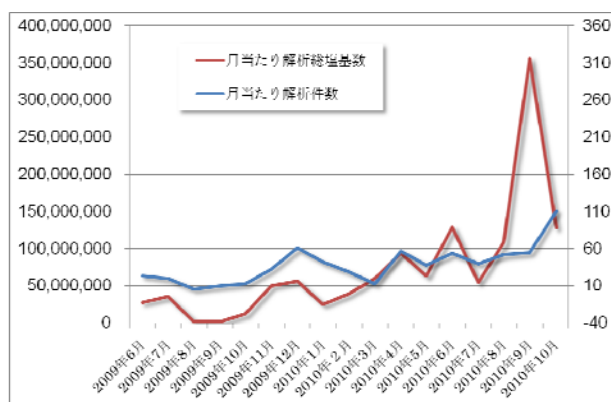


図 3.8 MiGAPの利用実績

公開後 17 カ月間の解析総件数が 658 件、解析総塩基数が 1.2Gbases 余り、予測した CDS 総数が 110 万個余りであった。解析件数と解析塩基数については月ごとの変動を図 3.19 にまとめた。比較的大きなゲノムを集中的に投入されることもあり、月ごとの変動が大きい。解析件数および解析データ量ともに着実に増加してきている。今後、近縁種ゲノムの知識の重みづけや既知遺伝子のアミノ酸配列マッピングの機能拡充といった利用者からの要望に応じていく。

4. 統合データベース支援

4. 1 ポータル整備・運用、広報、普及啓発

4. 1. 1 ポータル整備・運用、広報、普及啓発の実施計画

(1) ポータル整備・運用

「統合ホームページ」(URL:<http://lifesciencedb.jp/>)は利用者からのデータベースへのポータルサイトであるとともに、統合データベースプロジェクトの成果であるサービスやリソースの公開サイトである。平成 21 年度現在、50 件のサービスがトップページより利用可能である。ポータル整備・運用では平成 21 年度に引き続き 22 年度も統合ホームページの更新管理を担当する。また、最終年度にあわせてインターフェースの変更改良を行うとともに、個別に開発されてきたデータベースやサービスの統合を図る。具体的には、横断検索サービスについては現行の機能を維持しつつ、プロジェクトの成果を統合利用し、(1)辞書により強化された日本語による検索、(2)統合化された遺伝子データベースや蛋白質データベースの検索、(3)プロジェクト成果や受入れデータベースの検索、の 3 点を実現する、統合検索サービスのプロトタイプを開発する。併せて、学会要旨等の日本語コンテンツの

充実も図る。

(2) 広報、普及啓発

広報・普及啓発は、平成21年度、大学院生、学生を対象に8回の講習会を開催しプロジェクト成果の普及を図った。平成22年度も各地での講習会の開催を予定している。動画教材である統合TVは平成21年度までに250本の動画を製作し、生命科学系DBやツールの利用法を解説する他、講習会や講演会のプレゼンテーションの公開を行った。また英語版の制作も開始した。平成22年度も引き続きコンテンツの充実を続けるとともに、過去の動画の更新も行い、さらにDBカタログなど他のサービスとの連携も図っていく。

広報活動は平成20年度にはBMB2008等計5回、21年度には8回のブース出展などを行い利用者への普及を図ったが、アンケートの結果などからまだユーザへの浸透度を高める必要があると思われるので、平成22年度については国内の主要学会に加え、サービス内容に対応する特定の分野の学会への出展を計画している。またインターネットやパンフレット、各種広告など様々な媒体形態を利用して、合わせて一部のサービスについては英語化の媒体も作成して、統合DB活動の普及を行う。またそれと同時に、サービスに関するアンケート等を定期的実施し利用者の意見をサービスに反映させていくこととする。プロジェクトの成果発表およびプロジェクトが抱える問題を取り上げたシンポジウムについても、平成21年度に2回実施したところ反響が大きかったので、22年度にもプロジェクトのまとめの意味も含め、シンポジウム開催を1回予定している。

4. 1. 2 ポータル整備・運用、広報、普及啓発の実施内容

(1) ポータル整備・運用

1) ポータルサイトの構築、整備運用

平成22年9月現在で、58件のツールやサービス、成果物、関連サイトへのリンクがあり、そのうち32件が中核機関からのサービスやコンテンツである。各成果物はポータル、検索、データベース、ツール、というように分類表示している。

2) データベースカタログ

ポータルサイトの重要な機能として、どこにどのようなデータベースがあるのか、利用者が目的のデータベースを簡単に探すことができる必要がある。データベースカタログは国内のデータベースに関する情報を、各データベースの内容をもとに、分子種や生物種、運営機関ごとに調べることのできるサービスである。平成22年9月現在で912件のデータが公開されている(図4.1)。平成23年度からは4省庁連携の枠組みで調査データを作成していく予定である。またインタフェースについては検索エンジン対策のために詳細データの表示にMediaWikiを利用し、JSTで構築されてきたWingProとの

統合を行う計画である。データベースカタログと同一のインターフェースを利用して、生命科学系の学協会情報と生物アイコンの公開も同時に行っている。



図 4.1 データベースカタログのインターフェース

3) 生命科学データベース横断検索

横断検索は統合の第一段階として国内 DB を中心に、各 DB はそれぞれのサイトに置いたまま、キーワード検索を可能にした検索サービスである。平成 19 年度 13 件のデータベースを検索対象にスタートしたが、平成 20 年度は 200 件に、平成 22 年 9 月現在で 270 件の DB が、一括して検索可能となった。また、入力した検索キーワードは自動的に日英翻訳されるため、相手のデータベースが日本語か英語かを気にする必要は無い。さらに、平成 22 年度には経済産業省統合データベースプロジェクトにも技術情報を提供し検索連携を実現、利用者は省庁の区別なくデータベースを横断的に検索可能となった (図 4.2)。

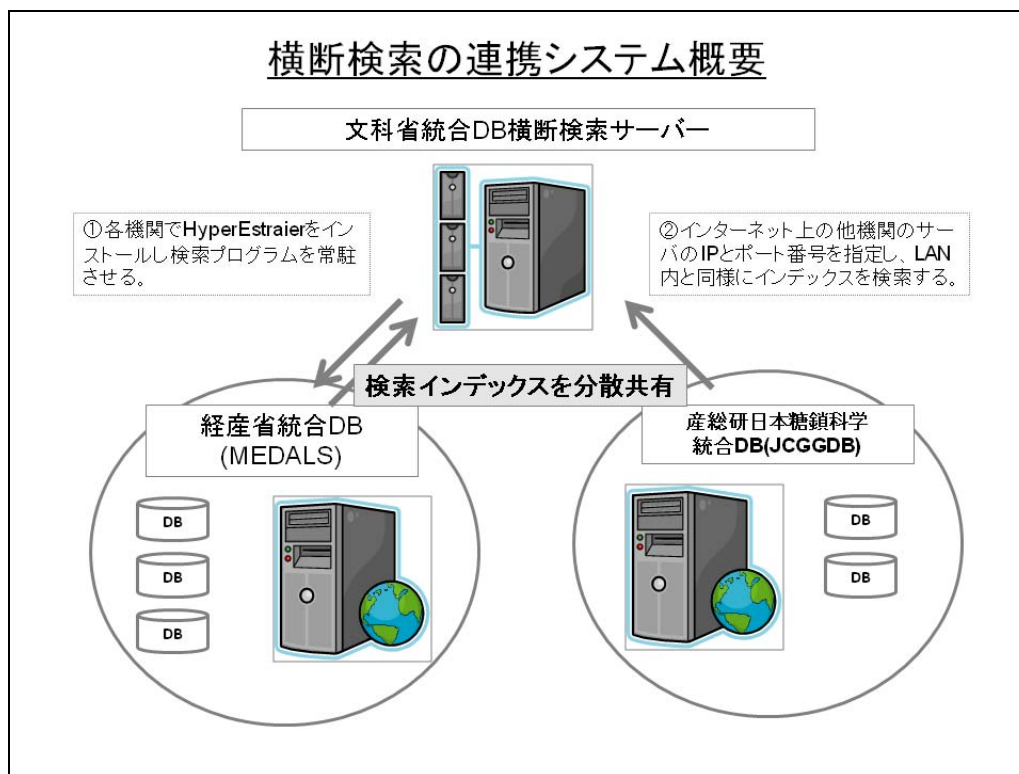


図 4.2 横断検索における連携システムの概要

4) 日本語オリジナルコンテンツの作成並びに英語サイトの作成

ポータルサイト構築にあたって、データベースの検索だけでなく、独自の日本語情報資源やオリジナルコンテンツの制作も行った。H19年度から生命科学分野の日本語総説誌である蛋白質核酸酵素(共立出版)のバックナンバー電子化並びに検索サービスを開始した。また、文部科学省特定領域研究「ゲノム」報告書を公開し、横断検索の対象として分子データベースや特許公報と合わせて検索可能にした。報告書に関しては平成22年度よりウェブ上で公開されていたCRESTやさきがけの報告書についても横断検索対象とした。また、データベースの利用促進を目的として、平成19年7月からデータベースやウェブツールの使い方を動画で解説する統合TVを開始し、平成22年11月現在、388件の動画を公開しており、公報普及の場ともなっている。

(2) 広報、普及啓発

統合データベースプロジェクトの普及のために、実際にインターネットにつながった端末を前にして主として初心者(大学後半から大学院前半が中心、ただし企業等からの参加も可能で、特に制限は設けない)向けに行う統合データベース講習会および単独講師による短いバージョンあるいは実習なしの講演会を平成19年度から開催することを計画し、

19年度末に第1回の講習会を開催、平成20年度からは年間を通じてコンスタントに開催してきた(表4.1)。講習会の内容は、統合データベースプロジェクトそのものあるいは本プロジェクトで開発されたサービスの解説を主体とした初心者向けの内容が中心であるが、広くデータ解析に利用されているウェブサイトやツールの解説なども盛り込み、研究の初心者に役立つ内容を幅広く提供できるように努めた。一方で、平成21年度からは2日間にわたる講習会を数回企画し、全般的な通常の講習会形式(1日目)と質問への回答形式(2日目)によってやや高度な内容にも対応できるような方法も試みたほか、Nucleic Acids Research 誌データベース特集号およびウェブサーバ特集号の解説を中心としたやや専門的な勉強会も実施し、中上級者向けのニーズにマッチした構成の会も設けた。

表 4.1 統合データベース講習会等の実施内容

年 度	統合DB 講習会(回)	講習会のべ参加者数(人) (ustreamによる参加者数)	統合DB 講演会(回) #	統合DB 勉強会(回)
22	7《予定含む》	131(413)《実施分のみ》	2	

#他機関からの要請分含まず

統合データベース講習会等で使用したテキストやその様子を録画した動画は、教育/人材育成の各参画機関において作成された教育プログラムを集約して平成18年度より公開しているウェブサイト「MotDB」から利用可能となるよう維持管理し、常に最新の情報が提供されるようにした。動画の一部は「統合TV」のコンテンツとしてあとからいつでも視聴できるようにしたほか、ustreamを利用して開催と同時の配信(中継)も行った。

プロジェクトの成果を集めたウェブサイトの本格的な公開(2008年6月)に合わせて、平成20年度から、プロジェクトで開発したデータベースやサービスの紹介やその他データベース統合化に必要な活動を紹介するために、学会年会におけるワークショップやフォーラムの企画・運営、ブースの出展やランチョンセミナーの開催といった広報活動を開始した(表4.2)。平成22年度は、効果の点から考えて再度BioJapan2010へ出展し、個別のサービスに関連の深い中規模の学会へのブース出展を新たに試みた(平成22年度はプロジェクト最終年度のため、12月で出展等の活動を終えるように計画した)。これらの展示会等においてはアンケートを実施し、サービス改善のための情報を得るとともに、ユーザ評価者(2.2.2項参照)の募集を行った。なお、22年度のBMB2010では、プロジェクト内外のライフサイエンスデータベース関連機関と相談のうえ、通路を挟んだ6つのブースを並べて出展することを試みた。また、多数のユーザが参加すると思われる分子生物学会年会(あるいはBMB[生化学会との合同年会])ではブース出展のほかに、サービスの紹介などのためにDBCLS研究員や参画機関が多数のポスター発表も行ってきており、平成22年度大会(2010年12月実施)ではテーマごとに分散する形で発表することとし、会期中様々な場面で本プロジェクトの成果がユーザの目に触れるようにした。

表 4.2 広報活動まとめ

年度	内容	回数（対象学会・展示会名称）
22	学会等企画提案	3（生物物理学会、BMB2010 [#] /2件）
	ブース出展	8（脂質生化学会、蛋白質科学会、CBRC2010、BioJapan2010、Biocuration2010、生物工学会、人類遺伝学会、BMB2010 [#] ）
	ランチョンセミナー	2（脂質生化学会、生物物理若手夏の学校）

[#]の BMB2010 は 2010 年 12 月開催のため予定

プロジェクトに関する情報発信（表 4.3）としては、これまで文科省の記者会室を通じたプレスリリースを 3 回（プロジェクト成果公開サイトの本格開始、経産省統合データベースサイトとの相互検索の開始、4 省関連機関での合同開催シンポジウムの案内）実施した。また、プロジェクト成果公開サイト、DBCLS のサイトおよび JST の「文科省統合データベース整備事業サイト」を通じたニュース配信を行うとともに、そのうちの一部は日経の Biotechnology Japan のサイトを利用した PR も実施した。22 年の 4 月には、DBCLS のウェブサイト进行全面改定したが、その際に DBCLS サイト、プロジェクト成果公開サイト、さらに DBCLS 用に取得した Twitter アカウントへ同時にニュースが流れるように工夫し、情報発信の効率化を図った。Twitter アカウントでは、このほか、DBCLS メンバーが発表する外部の講演会等の情報や、DBCLS 主催のイベントの実況中継なども配信している。プロジェクトに関するイベント等の案内や人材育成プログラムへの参加者募集には、展示会等でのアンケートの際承諾をいただいた方のメーリングリスト、製薬企業を中心としたメーリングリスト、「ゲノム」特定領域研究の班員のメーリングリスト等を利用した。

表 4.3 情報発信について

年度	プレス リリース（回）	ニュース 配信（回）	Biotech. Japan（回）	パンフレット 配布数（部）	リーフレットの 種類（英語版）累計
22	2	57 [#]	6 [#]	2400 [#]	43（9） [#]

[#]2010 年 11 月現在の数値

4. 2 データベース受入

4. 2. 1 データベース受入の実施計画

我が国のプロジェクトデータベースについて、データベースのタイプに応じた方法で、データベース受け入れを実施する。平成21年度は、「生命科学系データベースアーカイブ」を立ち上げ、変異体や表現型等のデータベースを中心に様々な対象のデータベース10件以上のアーカイブ公開を行った。

平成 22 年度は、プロジェクト参加機関のデータベース受け入れと連携をさらに進める

と同時に、アーカイブの対象をさらに拡張してアーカイブサイトでの公開を進める。また、アーカイブ対象の多様性に対応して、アーカイブサイトのインターフェースや機能の拡張を行い、主要な生命科学系プロジェクトの調査とプロジェクトカタログの更新を継続する。さらに、これらのデータ統合のためのコンテンツ整備を行う。

4. 2. 2 データベース受入の実施内容

1) ライフサイエンスプロジェクト及びデータベースの調査

データベース受け入れの対象や方法を検討するために、我が国のプロジェクトデータベースの現状を調査した。まずデータベースカタログデータの補充を行う形で、内閣府の連携施策群成果リストに理研、ゲノム特定領域（新旧）、JST SNPポータル、JBIRC、CBRC、かずさDNA研究所、農業生物資源研究所等のリストを追加し、分担機関である東京医科歯科大学が実施したアンケートのデータベース情報からこれまでにデータベースカタログに未収録のデータベース情報を15件登録した。

また、受け入れ対象データベースの把握のために、データベース自体の調査に並行して、ライフサイエンス国家プロジェクトの事業や成果に対する調査を行い、「生命科学系主要プロジェクト一覧」（図4.3、http://togodb.dbcls.jp/lldb_project）として公開した。平成22年度は文科省所管のプロジェクトデータ19件を更新、2件を追加し、これらのデータを前年度公開分22件とともに英語化した。

プロジェクト名	開始年度	終了年度	概要	データベース公開状況 の分類	公開データベース	ダウンロードサイト	公開ホームページ	実施機関
ゲノムネットワークプロジェクト	2004	2008	遺伝子の発現調節機能に関わる網羅的な解析	コンソリウム公開、一部共有	ゲノムネットワークプラットフォーム統合データベース(http://genometwork.nig.ac.jp/)	ゲノムネットワーク(http://genometwork.nig.ac.jp/public/download/dataset.html) リリース情報(http://genometwork.nig.ac.jp/public/contents/release.html)	ゲノムネットワークプロジェクト(http://genometwork.nig.ac.jp/)	京都府立総合研究機構
農林水産生物ゲノム情報統合DB	2006	2010	農林水産生物ゲノム統合DBの整備	共有	AgriTOGOデータベース(http://togo.dna.affrc.go.jp/)	http://togo.dna.affrc.go.jp/download.html	-	農林水産省
マイクロゲノム解析プログラム	1994	-	マイクロゲノム、cDNA配列、遺伝子発現解析の取得と解析	公開	マイクロDNAデータベース(http://kakoodna.dna.affrc.go.jp/) マイクロRNAデータベース(http://kako2ddb.dna.affrc.go.jp/qn-bn/search_2DOB.cgi) マイクロゲノム/アセンブリデータベース(http://kakovgms.dna.affrc.go.jp/)	-	マイクロゲノム解析プログラムホームページ(http://ggs.dna.affrc.go.jp/gg/index.html)	農林水産省
家畜ゲノム解析研究プロジェクト	1998	2001	ブタcDNA配列、繁殖性能、肉質性状の解析と解析	一部公開、一部共有	Animal Genome Database(http://animal.dna.affrc.go.jp/)	-	家畜ゲノム解析研究プログラムWebサイト(http://animal.dna.affrc.go.jp/)	農林水産省

図4.3 生命科学系主要プロジェクト一覧

2) ライフサイエンスデータベース受入方法の検討

上記の調査に基づいて、データベースをそれぞれの形態に従って分類し、それに基づいた

受け入れ方法を検討した。その結果、DNA配列や発現のデータが最もオリジナルデータの割合が多いことがわかった。ここでは、DNA配列の中でも日本が優位性を持つcDNAに着目し受入の試行を行うこととした。次に、データベース作成者との間の様々な状況に応じた受入方法の検討を行った。受入や連係の方法として、一括してデータを集める場合とデータベースが分散している場合に分けて検討した。データを集める場合は、1)データだけをFTPでダウンロード可能にする場合とデータベース自体を移設する場合、2)データベースの型を揃え、管理はそれぞれのデータベース構築者が運用する場合(データベース受入テンプレートの利用) 3)データベースの型を揃え、統合した形で当センターが構築管理を行う場合、いわゆる集中型統合の場合を検討した。また、データベースが分散している場合の連係方法と

Life Science Database Archive
LSDB Archive - アーカイブトップヘルプ
アーカイブ検索 [検索] [English] [Japanese]

文部科学省委託研究開発事業 統合データベースプロジェクト
生命科学系データベースアーカイブ

ホーム アーカイブの説明 寄託応募要領 更新履歴 お問い合わせ

いかなる良質なデータベースでも、説明が十分でない、利用条件が明確でない、ダウンロードできないなどの理由で十分に利用され、引用され、相応しい評価をうける機会を逃していることがあります。

生命科学系データベースアーカイブは、国内のライフサイエンス研究者が生み出したデータセットをわが国の公共財として長期間安全に維持保管し、データ説明(メタデータ)を統一して検索を容易にすると共に、利用許諾条件などの明示を行うことで、多くの人が容易にデータアクセスしダウンロードを行えるようにするサービスです(詳細説明)。

データが長期にわたって保全し、データベース作成者のクレジットを明示する一方、学術的や民間等様々なユーザが利用しやすい形にすることで、それぞれの研究の生命科学へのいっそうの貢献を支援します。データベースの寄託を随時募集しています(寄託応募要領)。

データインテグレーションへの利用
安心して、適切に利用可能
データベース利用者
データのダウンロード
目的検索機能
利用許諾条件、メタデータ
生命科学系データベースアーカイブ

データベースの維持管理
クレジット明示
データベース作成者
データベースの寄託

アーカイブデータベース一覧(ヘルプ)

現在、37件のアーカイブデータベースを公開しています。
[利用許諾]の右の列に「(ア)スタリク」のあるデータベースは、「文部科学省委託研究開発事業 統合データベースプロジェクト」によるデータベースです。

一覧内検索:

データベース	データベース運用場所	代表者	データベースカテゴリ	生物種	主要なキーワードを赤字表示	利用許諾
MAGEST ダウンロード 検索機能 およびAPI提供	東京大学 医科学研究所	高堂 和裕	cDNA	マボヤ	マボヤの遺伝子発現とESTのデータベース	CC表示・継承 詳細
Medaka EST Database ダウンロード 検索機能 およびAPI提供	東京大学大学院理学系研究科 生物科学専攻 動物科学大 遺伝動物発生学研究室	武田 洋幸	cDNA、発現	メダカ	メダカEST、迅速マッピングシステム M-marker 2003、medaka microarray BKのデータ	CC表示・継承 詳細
Heterozygosity of the Microsatellite Markers in the Japanese Population ダウンロード 検索機能 およびAPI提供	九州大学 生体医療医学研究所 原遺伝情報実験センター グローバル遺伝学分室	山本 徹	多型	ヒト	日本人におけるマイクロサテライトマーカーヘテロ接合体のデータベース	CC表示・継承 詳細
TMPDB ダウンロード 検索機能 およびAPI提供	弘前大学大学院 理工学研究科 電子情報工学科	清水 俊夫	蛋白質	生物全般	膜貫通トポロジモデルと各種トポロジ予測法の性能検証結果のデータベース	CC表示・継承 詳細
DIAM - 微生物情報 ダウンロード 検索機能 およびAPI提供	バイオインダストリー協会	-	文献	微生物全般(ウイルス、真正細菌、古細菌、真核微生物を含む)	微生物の同定・分類関連情報、安全関連情報、利用関連情報データベース	CC表示・継承 詳細
RA target genes in the Ciona embryo ダウンロード 検索機能 およびAPI提供	高知大学 理学部 応用理学科	森原 道樹	発現	カタユレイボヤ	カタユレイボヤ胚におけるレチノイン酸誘導遺伝子発達のマイクロアレイによる発現解析とWhole mount in situ hybridizationによる画像のデータベース	CC表示・継承 詳細

図 4.4 生命科学系データベースアーカイブ

しては、1)横断検索、及び2)API等を用いる分散型の統合検索を検討した。その結果、現在主に用いている簡易検索機能を併用した形に比べると、自前サーバもしくはクラウドを用いたミラーサーバの構築は、個別データベース独自の機能を移管できデータベース寄託者側からの支持は得易いが、データベース間で利用法に統一感がないためユーザからの支

持は得難く、またコストが高いことがわかった。しかし今後受け入れるデータベースの内容によっては、クラウドを用いたミラーサーバ構築も検討の余地があると考えられた。

3) 受け入れデータベースの運用システムの構築

以上の調査と検討に基づいて、受け入れデータベースの運用のために、受け入れたデータベースをダウンロード可能な形（アーカイブ）で公開するサービス「生命科学系データベースアーカイブ」（図 4.4、<http://dbarchive.lifesciencedb.jp/>）を構築、公開した。また、運用の過程で各種の機能拡張を進めた。

平成 22 年度は、登録データベース件数の増加に対応するために、データベース一覧においてインクリメンタルサーチによる絞り込みを可能にして、前年度導入したメタデータ横断検索と合わせてユーザが適切なデータベースを選択しやすくなるように配置した。また簡易検索ページでは、検索において AND 検索、OR 検索、フレーズ検索やデータ型に合わせたフォームでの検索をユーザが適切に行えるよう整備した他、一覧における表示件数の変更や小数の表示桁数の管理など、多様なデータに対応するための機能の拡張を続けている。

4) データベースの受け入れ

①各種機関からの受け入れ

平成 22 年度は、10 月末日までに 8 件のデータベースを受け入れるとともに、9 件のアーカイブ公開を行った結果、受け入れデータベースは累積 38 件、公開アーカイブは累積 37 件に達した（図 4.5）。新たに公開したアーカイブには RA target genes in the Ciona embryo（高知大学）、TMPDB（弘前大学）、Heterozygosity of the Microsatellite Markers in the Japanese Population（九州大学）、Medaka EST Database（東京大学）、MAGEST（東京大学で運用）の他、バイオインダストリー協会からの 4 つのアーカイブの 1 つ DIAM-微生物情報のように、ダウンロードできるデータのファイル数が 40 を超える複雑なアーカイブも現れている。前年度まではファイル数 3~4 の場合が多く、その多くは TogoDB のテーブルとして表現しておりテーブル間のリンクによって効率的に情報の閲覧をできるようにしていたが、今後は巨大もしくは複雑なデータベースの扱いが増えることが想定され、対応の検討を進めている。

②蛋白質関連データベース統合検索システムの開発

平成 22 年度は TogoProt の正式公開を始めるとともに、対象データベース更新への対応を定期的に行えるように、スクリプト作成によりシステム化を進めている。また、ユーザの便宜のために、データベース以外の有用なリンク、例えば本プロジェクトで開発された生命情報工学研究センターの各種 Web ワークフローや、経産省統合データベースプロジェクトのリンク自動解析システムなどへのリンクをまとめたフレームを含むインターフェースを作成し、試験的に公開している。

③他省庁との連携

経産省統合データベースプロジェクトとの間では平成 20 年度から、メタデータや横断検索、利用許諾、ウェブサービスなどでの共通化をはかることで合意するなどの連携を図

データベース	データベース運用場所	代表者	データベース	データベース運用場所	代表者
MAGEST	東京大学 医科学研究所	高峯 昭和	AcEST	京都大学 京都大学理学部 文芸学系 生命化学 植物遺伝学研究室	浦川 肇
Medaka EST Database	東京大学大学院理学系研究科 動物学専攻 動物学大講義動物学専攻	武田 洋希	Dicty_eDB - Dictyostelium cDNA Database	京都大学 大学院生命環境科学部 応用生命科学科	遠藤 勇子
Heterozygosity of the Microsatellite Markers in the Japanese Population	九州大学 生命科学研究科 遺伝学専攻 分子遺伝学研究室	山本 肇	Gclust Server	東京大学大学院 総合文化科学研究科 応用生命科学専攻 植物学専攻	佐藤 道博
TMPOB	京都大学 大学院 理工学研究所 電子情報学専攻	清水 俊夫	Budding yeast cDNA sequencing project	東京大学 大学院理学系研究科 動物学専攻	伊藤 俊司
DIAM - 微生物情報	バイオインフォマティクス協会	-	Yeast Interacting Proteins Database	東京大学 大学院理学系研究科 動物学専攻	伊藤 俊司
RA target genes in the Ciona embryo	高知大学 理学部 応用理学科	島原 道徳	rRNADB.CE	奈良バイオ大学 遺伝学分子生物学研究室	河部 忠志
DIAM - 安全性に関する文献	バイオインフォマティクス協会	-	DMPD (Dynamic Macrophage Pathway C-SML Database)	東京大学 医科学研究所	菅野 祐
DIAM - バイオテクノロジーの基礎知識	バイオインフォマティクス協会	-	Fluorome	豊田大学 教育事業推進課	園池 公昭
DIAM - バイオセーフティ関連の法規制	バイオインフォマティクス協会	-	GlycoProtDB	独立行政法人 農業技術総合研究所 植物工学研究センター グライコプロテオーム解析チーム	梶 裕之
The NAISTrap database	奈良先端科学技術大学院大学 バイオサイエンス研究科 動物遺伝学専攻	石田 隆博	Arabidopsis thaliana EST index	かずさDNA研究所 植物ゲノム研究部	中村 信一
DGSP Database - Drosophila Gene Search Project Database	京都大学 大学院理工学研究所 生命科学専攻 植物遺伝学研究室	相模 敏郎	Chlamydomonas reinhardtii EST index	かずさDNA研究所 植物ゲノム研究部	中村 信一
EGTC - Database for the Exchangeable Gene Trap Clones	熊本大学 生命資源研究センター	熊本 正徳	Lotus japonicus EST index	かずさDNA研究所 植物ゲノム研究部	中村 信一
Transposon insertion site database - germline	熊本大学 大学院 医学系研究科 遺伝学専攻	堀江 啓二	Trifolium pratense EST index	かずさDNA研究所 植物ゲノム研究部	中村 信一
HATODAS - 重原子データベース	独立行政法人 理化学研究所 植物学研究所 植物光科学総合研究センター	園田 道徳	BodyParts3D	ライフサイエンス統合データベースセンター	大久保 公葉
KTmAbDB - C.elegans Monoclonal Antibody Collection	理化学研究所 東京 遺伝学専攻 専攻研究センター 第五フロア 分子生物学研究チーム	杉本 優希子	生命科学系データベースカタログ	ライフサイエンス統合データベースセンター	川本 祥子
Atlas (ISH Data Base)	京都大学 大学院生命環境科学部 応用生命科学科	遠藤 勇子	生命科学系協会カタログ	ライフサイエンス統合データベースセンター	川本 祥子
GETDB - Gal4 Enhancer Trap Insertion Database	京都工芸繊維大学	村 祥生	生命科学系主要プロジェクト一覧	ライフサイエンス統合データベースセンター	西川 繁夫
WorTS - Worm TS mutant Database	大阪大学 微生物学研究所 植物細胞学専攻	中村 利明	施設名辞書	ライフサイエンス統合データベースセンター	大久保 公葉
			遺伝子名辞書	ライフサイエンス統合データベースセンター	大久保 公葉

図 4.5 生命科学系データベースアーカイブのデータベース一覧

図り、平成 21 年、平成 22 年度と会合を続けるとともに、上記コンセンサスに沿って協力し複数のデータベースに対するアーカイブ作成を進めている。

5. プロジェクトの総合的推進

5. 1 プロジェクトの総合的推進の実施計画

中核機関である情報・システム研究機構は、分担機関である京都大学、東京医科歯科大学、東京大学大学院医学系研究科および補完課題の実施機関である理化学研究所、産業技術総合研究所糖鎖医工学研究センター、情報・システム研究機構国立遺伝学研究所、九州工業大学に全体戦略を提示し連携して本事業を推進する。平成 22 年度は最終年度になるので、システム開発等はなるべく上期までに終わらせ、公開することを徹底する。これにより、ユーザの利用状況を把握しながら、サービス内容の向上に努める。そして引き続きプロジェクト参加機関との連携を徹底すると共に、研究運営委員会作業部会その他の会議の開催等を通してプロジェクトの推進を強化する。

5. 2 プロジェクトの総合的推進の実施内容

(1) 研究運営委員会及び研究運営委員会作業部会

平成 22 年度は、プロジェクトの最終年度に当たるため、開発関係は 9 月末までには終了させ、年度末までにはその効果を検証することを中核機関はじめ関係機関においては徹底することとした。さらに、平成 23 年度には新センターが JST 内に発足し、本プロジェクトで得られた成果を引継ぐことがほぼ決定的となっているため、今年度は当初立てた計画通りに進んでいるのか、あるいは修正や終了させるべき項目があるのか否か等、そして新センターを取り巻く状況も踏まえ研究運営委員会及び作業部会分科会その他の会議を通してプロジェクト参加機関の連携・調整、及び進捗の確認を行った。

尚、研究運営委員会及び作業部会分科会のメンバーは下記の通りである(敬称略、順不同)。

勝木 元也	自然科学研究機構理事
金岡 昌治	(株)大日本住友製薬研究本部副本部長
金久 實	京都大学化学研究所バイオインフォマティクスセンター長
白木澤佳子	(独)科学技術振興機構イノベーション推進本部バイオサイエンスデータベースセンター準備室長
榊 佳之	豊橋技術科学大学長
豊田 哲郎	(独)理化学研究所横浜研究所生命情報基盤研究部門長

長洲 毅志	エーザイ（株）理事・CSO付き担当部長
中村 桂子	JT生命誌研究館館長
中村 春木	大阪大学蛋白質研究所蛋白質情報科学研究系教授
長村 吉晃	（独）農業生物資源研究所基盤研究領域ゲノムリソースセンター長
松尾 泰樹	（独）理化学研究所横浜研究所研究推進部長
松原 謙一(委員長)	（株）DNAチップ研究所取締役名誉所長
湯元 昇	（独）産業技術総合研究所理事
吉田 輝彦	（独）国立がんセンター研究所遺伝医学研究部分野分野長
吉田 光昭	（財）癌研究会癌化学療法センター所長
大久保公策	情報・システム研究機構国立遺伝学研究所 生命情報・DDBJ研究センター長
五條堀 孝	情報・システム研究機構国立遺伝学研究所副所長
高木 利久	東京大学大学院新領域創成科学研究科教授
堀田 凱樹(副委員長)	情報・システム研究機構長

また、研究運営委員会作業部会のメンバーは下記の通りである（敬称略、順不同）。

浅井 潔	（独）産業技術総合研究所生命情報工学研究センター長
白木澤佳子	（独）科学技術振興機構イノベーション推進本部バイオサイエンスデータベースセンター準備室長
五斗 進	京都大学化学研究所バイオインフォマティクスセンター准教授
田中 博	東京医科歯科大学情報医科学センター長
田畑 哲之	（財）かずさDNA研究所副所長
徳永 勝士	東京大学大学院医学系研究科教授
豊田 哲郎	（独）理化学研究所横浜研究所生命情報基盤研究部門長
成松 久	（独）産業技術総合研究所糖鎖医工学研究センター長
松原 謙一	（株）DNAチップ研究所代表取締役社長
森下 真一	東京大学新領域創成科学研究科教授
五條堀 孝	情報・システム研究機構国立遺伝学研究所副所長
菅原 秀明	情報・システム研究機構国立遺伝学研究所 生命情報・DDBJ研究センター特任教授
高木 利久(主査)	東京大学大学院新領域創成科学研究科教授

上記の体制で、開催した研究運営委員会及び研究運営委員会作業部会は次の通りである。

第十回研究運営委員会 平成22年4月26日

第十一回研究運営委員会 平成22年11月12日

第十二回作業部会分科会	平成 22 年 6 月 1 日
第十三回作業部会分科会	平成 22 年 6 月 2 日
第十四回作業部会分科会	平成 22 年 6 月 3 日
第十五回作業部会分科会	平成 22 年 10 月 21 日
第十六回作業部会分科会	平成 22 年 11 月 2 日
第十七回作業部会分科会	平成 22 年 11 月 2 日

(2) 関連機関との連絡会

平成 22 年度は、21 年度に引き続き各省との連携を進める打ち合わせを設定すると共に、参画機関間の連携の見直しと今後重要になるであろう医療画像データに関する会合を設定した。会合の一覧は以下の通りである。

日付	時間	内容	場所
H22.04.20	16:30-18:00	分担機関（東大 G との打合せ）	東大医学部
H22.04.28	18:00-20:00	ヒトゲノム variation database 打合せ	外部会議室
H22.05.12	17:00-19:00	分担機関（東大 G との打ち合わせ）	東大医学部
H22.05.24	16:00-16:30	トキシコゲノミクスプロジェクトとの打合せ	テレビ会議
H22.05.26	14:00-15:30	理研 OSC とのミーティング	DBCLS
H22.05.28	15:00-18:00	経産省統合 DB とのミーティング	産総研
H22.06.07	13:00-14:30	トキシコゲノミクスプロジェクトとの打合せ	医薬基盤研
H22.06.02	18:00-20:00	ヒトゲノム variation database に関する打合せ	外部会議室
H22.07.08	14:00-16:00	分担機関（医科歯科大、阪大）打合せ	DBCLS
H22.07.27	14:00-15:30	SEVENS アーカイブについての打合せ	産総研
H22.08.06	16:00-18:00	経産省統合 DB とのミーティング	DBCLS
H22.08.18	14:00-18:00	医用画像データ関係打合せ	DBCLS
H22.09.09	13:00-14:30	トキシコゲノミクスプロジェクトとの打合せ	DBCLS
H22.09.21	16:00-17:00	経産省統合 DB とのミーティング	DBCLS
H22.11.10	10:00-12:00	分担機関（東大 G との打ち合わせ）	東大医学部

5. 3 その他シンポジウム、セミナー等

(1) 統合データベース講習会

4. 1. 1. 項で記載した平成 22 年度の統合データベース講習会の詳細は以下のとおりである。

①統合データベース講習会：AJACS 湘南 3

日時：2010 年 6 月 4 日（金）午前 10 時～午後 5 時

場所：日本大学生物資源科学部 2 号館 231 教室

対象：学部 3、4 年生や大学院生、教員を始めとしたライフサイエンス分野のデータベースを効率よく利用することに興味のある方や統合データベースが実際にどのように作成、維持管理されているかに興味のある方

内容：

10:00-10:20 「はじめに：統合データベースプロジェクトとは？」

10:20-10:40 「統合 TV」

10:40-10:50 休憩

10:50-11:40 「生命科学横断検索の利用法」

11:40-12:20 「文献情報を効率的に活用する」

12:20-13:20 昼休憩

13:20-14:20 「遺伝子発現データの使い倒し術」

14:20-15:20 「次世代シーケンサの活用法」

15:20-15:40 休憩

15:40-16:20 「BodyParts3D/Anatomography の利用法」

16:20-17:00 「KazusaMart の使い方」

②AJACS & 第 22 回 DDBJing 講習会

日時：2010 年 6 月 23 日（水）10 時から 17 時まで

2010 年 6 月 24 日（木）10 時から 17 時まで

場所：ライフサイエンス統合データベースセンター（東京大学本郷キャンパス 工学部 12 号館 4 階）

主催：大学共同利用機関法人 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(DBCLS) 国立遺伝学研究所 生命情報・DDBJ 研究センター 日本 DNA データバン(DDBJ)

対象：DBCLS・DDBJ を利用されている方

内容：

第 1 日目 6 月 23 日（水）10:00～17:00

9:30-10:00 受付

10:00-10:15 統合データベースプロジェクトの紹介

高祖 歩美(ライフサイエンス統合データベースセンター 特任研究員)

10:15-11:30 カタログサービス・横断検索の利用法

川本 祥子(ライフサイエンス統合データベースセンター 特任准教授)

11:30-12:20 文献関連サービスの活用法

山本 泰智(ライフサイエンス統合データベースセンター 特任研究員)

12:20-13:30 昼食休憩

13:30-15:00 遺伝子発現情報を使い倒す

- 坊農 秀雅 (ライフサイエンス統合データベースセンター 特任准教授)
- 15:00-15:20 休憩
- 15:20-16:10 BodyParts3D/Anatomography の利用法
三橋 信孝(ライフサイエンス統合データベースセンター 特任研究員)
- 16:10-17:00 DBCLS Galaxy の利用法
山口 敦子(ライフサイエンス統合データベースセンター 特任准教授)
- 第2日目 6月24日(木) 10:00~17:00
- 10:00-11:30 DDBJ の紹介と配列データの検索
中村 保一 (国立遺伝学研究所 教授)
- 11:30-12:30 次世代シーケンサ配列の登録・データ解析
次世代シーケンサのクラウド型解析パイプライン (15分)
神沼 英里 (国立遺伝学研究所 助教)
次世代シーケンサアーカイブ DB (20分)
児玉 悠一 (DDBJ 構築チーム アノテータ)
クラウド型解析パイプライン・実習 assembly/mapping (25分)
望月 孝子 (国立遺伝学研究所 エンジニア)
- 12:30-13:30 昼食休憩
- 13:30-15:10 SAKURA を用いた塩基配列登録の方法・実習
小菅 武英 (DDBJ 構築チーム アノテータ)
- 15:10-15:30 休憩
- 15:30-17:00 相同性検索と系統解析: BLAST と ClustalW
中村 保一 (国立遺伝学研究所 教授)

③統合データベース講習会: AJACS 加賀

日時: 2010年7月6日(火) 午前9時30分~午後4時

場所: 金沢大学 角間キャンパス 自然研本館2階 計算機実習室1

対象: 学部3、4年生や大学院生、教員を始めとしたライフサイエンス分野のデータベースを効率よく利用することに興味のある方や統合データベースが実際にどのように作成、維持管理されているかに興味のある方

内容:

- 9:30-09:45 「はじめに: 統合データベースプロジェクトとは？」
- 9:45-10:45 「生命科学横断検索の利用法」
- 10:45-11:45 「タンパク質データベースの活用法」
- 11:45-13:00 昼休憩
- 13:00-14:50 「遺伝子発現データの活用法/次世代シーケンサの活用法」
- 15:00-16:00 「文献情報を効率的に活用する」

④統合データベース講習会：AJACS みちのく

日時：2010年8月4日（水）12：30～18：00

2010年8月5日（木）10：00～17：00

場所：東北大学 星陵地区 IT 教育施設（医学部4号館6階）

対象：ライフサイエンス分野のデータベースを効率よく利用し、研究等に役立てたい方

内容：

8月4日（水）

12:30～15:00 統合データベースの利用法

統合データベースプロジェクトとは

統合TVを使い倒す

統合データベースプロジェクトのサービスを使い倒す

15:30～18:00 次世代シーケンサの活用法

DDBJのNGSデータアーカイブ: DRA (DDBJ Sequence Read Archive)

おまけ：第三世代シーケンサ

データの解析法

8月5日（木）

10:00～12:30 情報の可視化

ゲノムブラウザ、Cytoscape

BodyParts3D/Anatomography

14:00～17:00 遺伝子発現情報を使い倒す

遺伝子発現データベースの活用法

R/Bioconductor を使った遺伝子発現解析

⑤統合データベース講習会：AJACS 本郷7

日時：2010年10月4日（月）13：00～18：00

場所：ライフサイエンス統合データベースセンター（東京大学本郷キャンパス浅野地区工学部12号館4階）

対象：ライフサイエンス分野のデータベースを効率よく利用し、研究等に役立てたい方

内容：

13:00～13:20 統合データベースプロジェクトについて

河野信（ライフサイエンス統合データベースセンター）

13:20～14:00 TogoProt とデータ統合

畠中秀樹（ライフサイエンス統合データベースセンター）

14:00～14:40 ワークフローによる配列及び構造解析

野口保、田代俊行（産業技術総合研究所 生命情報工学研究センター）

- 情報統合班)
- 15:00～15:40 糖鎖関連データベースの紹介
鹿内俊秀 (産業技術総合研究所 糖鎖医工学研究センター)
- 15:40～16:20 tRNADB-CE : エキスパートがキュレートした tRNA 遺伝子データベースの紹介
阿部貴志 (長浜バイオ大学)
- 16:40～17:20 微生物ゲノム自動アノテーションパイプライン MiGAP の使い方・使われ方
菅原秀明 (国立遺伝学研究所)、兼崎友 (東京農業大学)、大山彰 (インシリコバイオロジー)、黒川顕 (東京工業大学)
- 17:20～18:00 GWAS/CNV-DB と Mutation DB の紹介
小池麻子 (日立製作所 中央研究所)

⑥統合データベース講習会 : AJACS 近江 3

日時 : 2010 年 11 月 5 日 (金) 13 : 30～17 : 00

場所 : 長浜バイオ大学命江館 2 階 情報実習室 1

対象 : 長浜バイオ大学生命情報科学コース 3 回生 + ライフサイエンス分野のデータベースを効率よく利用し、研究等に役立てたい方

内容 :

- 13:30-13:45 はじめに～「統合データベースプロジェクトとは」 担当・川本祥子
(DBCLS) 統合ホームページ(<http://lifesciencedb.jp/>)から データベース
カタログ&横断検索の簡単な実習
- 13:45-14:25 「Kazusa Annotation Suite によるゲノムアノテーション」
- 14:25-15:05 「遺伝子発現データベースの活用法」
- 15:05-15:20 休憩
- 15:20-16:00 「TogoProt から探るタンパク質」
- 16:00-17:00 「Galaxy でできるゲノムスケール共同研究」

(2) BioHackathon

平成 22 年度は、これまで得てきた国際技術動向や知見を国内で共有する必要があると考え、国内のデータベース構築者・利用者を対象に、統合データベース技術情報交換ワークショップ(<http://wiki.lifesciencedb.jp/mw/index.php/BH10.10>)を行った。国内から 70 名弱の参加者があり、活発な議論と情報交換がおこなわれた。プログラム等スケジュールは下記の通りである。

会期 : 2010 年 10 月 18 日 (月) ～22 日 (金)

場所：静岡県修善寺 ラフォーレ修善寺

内容（プログラム）：

2010-10-18（月）

13:30-14:00 オープニング

中尾 光輝（DBCLS）：開催趣旨説明

大久保 公策（NIG DDBJ センター長）：ウェルカムメッセージ

14:00-15:00 シンポジウム・セッション1：統合データベースプロジェクト

川本 祥子（DBCLS）：文部科学省ライフサイエンス統合データベースプロジェクトと DBCLS のサービス概要とポータル・横断検索

村上勝彦、山崎千里（AIST）：経済産業省ライフサイエンス統合データベースプロジェクト

15:00-17:00 ライトニングトーク 参加者全員、一人3分の4コマプレゼンテーション

17:00-18:00 ポスターセッション

4コマプレゼンテーションをA4四枚に印刷したものを掲示。

18:00 懇親会

2010-10-19（火）

9:30-12:00 シンポジウム・セッション2：データベース

中村保一（DDBJ）：DDBJ ってみんなの役にたつんだろうか

飯田啓介（DBCLS）：データベースと編集（仮）

熊谷禎洋（日立ソリューションズ）：データベース受け入れ事業

岡本忍（DBCLS）：KazusaAnnotation によるゲノムアノテーション協働作業手法の開発

五十嵐芳暢（医薬基盤研究所）：トキシコゲノミクスインフォマティクスプロジェクト

河野信（DBCLS）：統合TVとスクリーンキャスト（仮）

荒川和晴（慶應義塾大学先端生命科学研究所）：ウェブサービス統合のためのユーザインタフェース

櫻井望（かずさDNA研究所）：プロジェクト専用データベースの事後公開にあたって

13:00-15:00 シンポジウム・セッション3：サービス・技術

西山智明（金沢大学／ERATO 長谷部分化全能性）：SOLiD シーケンスデータの解析

川路英哉（RIKEN）：公共DBへのデータ登録とフォーマット

粕川雄也（RIKEN CDB）：Brain Star のAPIと可視化（仮）

木下聖子（Soka University）：糖鎖情報の統合を目指して

金 進東（DBCLS）：バイオ文献からの情報抽出ー世界の動向

15:00-19:00 チュートリアル

牧口大旭 (三井情報)、杉崎太一郎 (三井情報) : 横断検索の P2P 型データベース統合

片山俊明 (HGC) : TogoDB/TogoWS

神沼英里 (DDBJ) : DDBJ パイプライン

三橋信孝 (DBCLS) : BodyParts3D/Anatomography

西田孝三 (NAIST) : Cytoscape への KEGG データベース情報の取り込みと実験情報のマッピング

大林武 (東北大学) : 共発現データベース ATTED-II が広く利用されるようになった訳

荒川和晴 (慶應義塾大学先端生命科学研究所) : Dashcode 3 による次世代ウェブアプリケーション開発

2010-10-20 (水)

9:30-11:00 ワークショップ 1

西田孝三 : Cytoscape を用いたデータベースクライアント、可視化

後藤直久 : BioRuby と ソースコードレポジトリ

11:00-12:00 オープンスペース

これまでの話題から、掘り下げて議論や調査、開発するトピックを列挙する。

13:00-18:00 ハッカソン 1

ハッカソンでは、調査や開発、議論を小グループに別れて実施する。

- ・データベースの RDF 化 金城
- ・実用的なバイオオントロジー関連技術の調査 川島
- ・ウェブサービスマニフェストの開発と BioDBCore 対応 山崎
- ・データベースのライセンス問題
- ・糖鎖 API 標準化部分の実装 木下
- ・TogoWS, TogoDB の開発 片山
- ・BioRuby 後藤
- ・iPhone アプリ
- ・ブックマークレット開発
- ・バイオデータ可視化
- ・Cytoscape プラグインを作る
- ・Galaxy ツールを作る 神田
- ・R/Bioconductor パッケージを作る
- ・OSQA を利用したあたらしい QA サイトもしくはサイトホスティング 中尾、神田
- ・オルソログクラスターへの自動命名手法の開発 岡本
- ・ウェブサービス API と Cytoscape プラグイン- 大林, 笠原、西田

18:00 から今日のまとめ。

2010-10-21 (木)

9:30-12:00 ハッカソン 3

11:45 午前のまとめ

13:00-19:00 ハッカソン 4

18:30 今日のまとめ

2010-10-22 (金)

9:30-12:00 ハッカソン 5

13:00-15:00 ハッカソン 6

15:00-16:00 まとめ

6. 成果の外部への発表

論文、学会発表等の発表内容については、別紙 2 の外部発表実績一覧を参照されたい。
平成 22 年度の新聞、雑誌等における本プロジェクトに関する記事は以下の通りである。

タイトル	メディア名	掲載日
情報・システム研究機構(ROIS)、産総研など 生命科学系 DB 相互検索—省の垣根超え効率化—	日刊工業新聞	H22.5.25
情報・システム研究機構(DBCLS)、産総研など DB 相互検索可能に—文科・経産関連を統合—	化学工業日報	H22.5.25

7. 特許出願状況

該当なし。

8. 整備実績一覧

別紙 1 の整備実績一覧を参照されたい。

9. 外部発表実績一覧

別紙 2 の外部発表実績一覧を参照されたい。

10. 実施体制

次頁を参照されたい。

10. 実施体制

役割 ◎：業務主任者、●：サブテーマ代表、○：実施担当

氏名	所属		研究項目	年度毎の登録				
	部門	役職		H18	H19	H20	H21	H22
堀田 凱樹	情報・システム研究機構	機構長	サブテーマ（１）	◎	—	—	—	—
高木 利久	ライフサイエンス統合データベースセンター	特任教授	サブテーマ（１）、（２）、（３）	●	◎	◎	◎	◎
大久保 公策	国立遺伝学研究所生命情報・DDBJ 研究センター	教授・センター長	サブテーマ（２）ヒト統合DBの開発・運用	●	●	●	●	●
小笠原 理	国立遺伝学研究所生命情報・DDBJ 研究センター	助教	サブテーマ（２）ヒト統合DBの開発・運用	—	○	○	○	○
川本 祥子	ライフサイエンス統合データベースセンター	特任准教授	サブテーマ（３）ポータル整備・運用、広報、普及啓発	○	●	●	●	●
菅原 秀明	国立遺伝学研究所生命情報・DDBJ 研究センター	特任教授	サブテーマ（２）モデル生物・産業応用生物統合DBの開発・運用	○	○	○	○	○
高野 明彦	国立情報学研究所連想情報学研究開発センター	教授・センター長	サブテーマ（１）	○	—	—	—	—
武田 英明	国立情報学研究所情報学プリンシプル研究系	教授	サブテーマ（２）共通基盤技術開発	○	—	—	—	—
五條堀 孝	国立遺伝学研究所生命情報・DDBJ 研究センター	教授	サブテーマ（１）	○	—	—	—	—
西川 建	国立遺伝学研究所生命情報・DDBJ 研究センター	教授	サブテーマ（２）共通基盤技術開発	○	—	—	—	—
藤山 秋佐夫	国立情報学研究所情報学プリンシプル研究系	教授・研究主幹	サブテーマ（１）	○	—	—	—	—
市瀬 龍太郎	国立情報学研究所プリンシプル研究系	准教授	サブテーマ（２）共通基盤技術開発	—	—	—	—	○
岡本 忍	ライフサイエンス統合データベースセンター	特任准教授	サブテーマ（２）モデル生物・産業応用生物統合DBの開発・運用	○	—	—	○	○
水田 洋子	ライフサイエンス統合データベースセンター	プロジェクト研究員Ⅱ	サブテーマ（２）共通基盤技術開発	○	—	—	—	—
三橋 信孝	ライフサイエンス統合データベースセンター	特任研究員	サブテーマ（２）ヒト統合DBの開発・運用	○	○	○	○	○
永井 啓一	ライフサイエンス統合データベースセンター	特任教授	サブテーマ（１）、（３）	—	○	○	○	○
箕輪 真理	ライフサイエンス統合データベースセンター	特任准教授	サブテーマ（１）、（３）広報・普及啓発	—	—	○	○	○
西川 哲夫	ライフサイエンス統合データベースセンター	特任教授	サブテーマ（１）、（３）	—	○	○	○	—
吉羽 洋周	ライフサイエンス統合データベースセンター	特任研究員	サブテーマ（１）	—	—	—	○	○
山口 敦子	ライフサイエンス統合データベースセンター	特任准教授	サブテーマ（２）共通基盤技術開発	—	○	○	○	○
坊農 秀雅	ライフサイエンス統合データベースセンター	特任准教授	サブテーマ（２）、（３）	—	○	○	○	○

山本泰智	ライフサイエンス統合データベースセンター	特任研究員	サブテーマ（２）共通基盤技術開発	－	○	○	○	○
河野 信	ライフサイエンス統合データベースセンター	特任研究員	サブテーマ（３）ポータル整備・運用、広報・普及啓発	－	○	○	○	○
仲里猛留	ライフサイエンス統合データベースセンター	特任研究員	サブテーマ（３）ポータル整備・運用	－	○	○	○	○
全 弘宇	ライフサイエンス統合データベースセンター	特任研究員	サブテーマ（２）共通基盤技術開発	－	○	○	○	－
藤枝 香	ライフサイエンス統合データベースセンター	特任研究員	サブテーマ（２）ヒト統合 DB の開発・運用	－	○	○	○	○
坂東 明日香	ライフサイエンス統合データベースセンター	特任技術専門員	サブテーマ（３）データベースの受入と運用	－	－	－	○	○
S.Riedel	ライフサイエンス統合データベースセンター	特任研究員	サブテーマ（２）共通基盤技術開発	－	－	○	－	－
中尾 光輝	ライフサイエンス統合データベースセンター	特任研究員	サブテーマ（２）共通基盤技術開発	－	－	－	○	○
大内田 賢太	ライフサイエンス統合データベースセンター	特任技術専門員	サブテーマ（２）共通基盤技術開発	－	－	○	○	○
高祖 歩美	ライフサイエンス統合データベースセンター	特任研究員	サブテーマ（３）広報・普及啓発	－	－	○	○	○
小野 浩雅	ライフサイエンス統合データベースセンター	RA（→特任技術専門員）	サブテーマ（３）ポータル整備・運用	－	○	○	○	○
中村 春木	ライフサイエンス統合データベースセンター	客員教授	サブテーマ（３）データベースの受入と運用	－	○	○	○	○
金城 玲	ライフサイエンス統合データベースセンター	客員准教授	サブテーマ（３）データベースの受入と運用	－	○	○	○	○
畠中 秀樹	ライフサイエンス統合データベースセンター	特任准教授	サブテーマ（３）データベースの受入と運用	－	－	－	○	○
金 進東	ライフサイエンス統合データベースセンター	特任准教授	サブテーマ（２）共通基盤技術開発	－	－	－	－	○
飯田 啓介	ライフサイエンス統合データベースセンター	特任技術専門員	サブテーマ（３）ポータル整備・運用	－	－	－	－	○
王 悦	ライフサイエンス統合データベースセンター	技術補佐員	サブテーマ（２）共通基盤技術開発	－	－	－	－	○
大村 蓉子	ライフサイエンス統合データベースセンター	RA	サブテーマ（３）ポータル整備・運用	－	－	○	○	－
千葉 温子	ライフサイエンス統合データベースセンター	RA	サブテーマ（３）ポータル整備・運用	－	－	○	○	－
山口 瑤子	ライフサイエンス統合データベースセンター	RA	サブテーマ（３）ポータル整備・運用	－	－	○	○	－
横山 貴央	ライフサイエンス統合データベースセンター	RA	サブテーマ（３）ポータル整備・運用	－	－	○	○	○
大石晶一郎	ライフサイエンス統合データベースセンター	技術補佐員	サブテーマ（３）ポータル整備・運用	－	－	－	○	○
大田 達郎	ライフサイエンス統合データベースセンター	RA	サブテーマ（３）ポータル整備・運用	－	－	－	○	○
奥田 裕樹	ライフサイエンス統合データベースセンター	RA	サブテーマ（３）ポータル整備・運用	－	－	－	○	○
小野 智央	ライフサイエンス統合データベースセンター	RA	サブテーマ（３）ポータル整備・運用	－	－	－	○	○

熊谷 雄介	ライフサイエンス統合データベースセンター	RA	サブテーマ(3)ポータル整備・運用	-	-	-	○	○
高橋 公海	ライフサイエンス統合データベースセンター	RA	サブテーマ(3)ポータル整備・運用	-	-	-	○	-
寺田 愛花	ライフサイエンス統合データベースセンター	RA	サブテーマ(3)ポータル整備・運用	-	-	-	○	○
並木 洋平	ライフサイエンス統合データベースセンター	RA	サブテーマ(3)ポータル整備・運用	-	-	-	○	○
根来 哲也	ライフサイエンス統合データベースセンター	技術補佐員	サブテーマ(3)ポータル整備・運用	-	-	-	○	○
橋本 拓也	ライフサイエンス統合データベースセンター	RA	サブテーマ(3)ポータル整備・運用	-	-	-	○	-
林 和弘	ライフサイエンス統合データベースセンター	RA	サブテーマ(3)ポータル整備・運用	-	-	-	○	○
山田 恵	ライフサイエンス統合データベースセンター	RA	サブテーマ(3)ポータル整備・運用	-	-	-	○	○
吉田 康久	ライフサイエンス統合データベースセンター	RA	サブテーマ(3)ポータル整備・運用	-	-	-	○	-
飯田 宗玄	ライフサイエンス統合データベースセンター	技術補佐員	サブテーマ(3)ポータル整備・運用	-	-	-	-	○
伊藤 真和吏	ライフサイエンス統合データベースセンター	RA	サブテーマ(3)ポータル整備・運用	-	-	-	-	○
久保 奈月	ライフサイエンス統合データベースセンター	技術補佐員	サブテーマ(3)ポータル整備・運用	-	-	-	-	○
坂田 隼人	ライフサイエンス統合データベースセンター	技術補佐員	サブテーマ(3)ポータル整備・運用	-	-	-	-	○
水津 珠世	ライフサイエンス統合データベースセンター	技術補佐員	サブテーマ(3)ポータル整備・運用	-	-	-	-	○
高橋 文緒	ライフサイエンス統合データベースセンター	技術補佐員	サブテーマ(3)ポータル整備・運用	-	-	-	-	○
武田 英之	ライフサイエンス統合データベースセンター	RA	サブテーマ(3)ポータル整備・運用	-	-	-	-	○
竹馬 光太郎	ライフサイエンス統合データベースセンター	技術補佐員	サブテーマ(3)ポータル整備・運用	-	-	-	-	○
趙 漢哲	ライフサイエンス統合データベースセンター	RA	サブテーマ(2)共通基盤技術開発	-	-	-	-	○
比留川 治子	ライフサイエンス統合データベースセンター	RA	サブテーマ(3)ポータル整備・運用	-	-	-	-	○
加藤 菊也	大阪府立病院機構大阪府立成人病センター研究所	所長	サブテーマ(2)ヒト統合DBの開発・運用	○	○	○	○	○
黒川 顕	東京工業大学大学院生命理工学研究科生命情報専攻	教授	サブテーマ(2)モデル生物・産業応用生物統合DBの開発・運用	-	○	○	○	○
内山 郁夫	自然科学研究機構基礎生物学研究所	助教	サブテーマ(2)モデル生物・産業応用生物統合DBの開発・運用	-	○	○	○	○
伊藤 剛	(独)農業生物資源研究所基礎研究領域ゲノム情報研究ユニット	ユニット長	サブテーマ(2)モデル生物・産業応用生物統合DBの開発・運用	-	○	○	○	○
片山 俊明	東京大学医科学研究所ヒトゲノム解析センター	助教	サブテーマ(2)共通基盤技術開発	○	○	○	○	○
川島 秀一	東京大学医科学研究所ヒトゲノム解析センター	助教	サブテーマ(2)共通基盤技術開発	○	○	○	○	○

PAPAZIAN Frederic	ライフサイエンス統合データベースセンター	外来研究員	サブテーマ（２）共通基盤技術開発	－	－	－	－	○
岩崎 渉	東京大学大学院新領域創成科学研究科	助教	サブテーマ（２）共通基盤技術開発	－	－	－	－	○
蓑島 伸生	浜松医科大学光量子医学研究センター	教授	サブテーマ（２）ヒト統合 DB の 開発・運用	－	－	－	－	○
後藤 直久	大阪大学微生物病研究所	助教	サブテーマ（２）共通基盤技術開発	－	－	－	－	○
内藤 雄樹	東京大学大学院理学系研究科	助教	サブテーマ（３）ポータル整備・運用	－	－	－	○	○

なお、サブテーマ（１）は「戦略立案・実行評価」、サブテーマ（２）は「統合データベース開発」、サブテーマ（３）は「統合データベース支援」である。

整備実績一覧

(別紙 1)

(1) データ (又はDB) の連結、統合化整備

通番	データ (又はDB) の名称	公開／未公開	概要 (データの種別 (生物種)・数量 (kB 等)、本プロジェクトで実施した特徴点、進捗状況、今後の計画・課題などを簡潔にわかりやすく記述)
1	生命科学データベースカタログ http://lifesciencedb.jp/lldb.cgi?pg=1 (英語版) http://lifesciencedb.jp/lldb.cgi?gg=dbcatalog&lng=en	公開	生命科学分野のデータベースカタログ (注釈及びメタデータ) で、データベースサイトのトップページをサムネイルで表示。データベースの稼働状況をモニターして表示。利用者からのコメントを受け付ける。掲載データベース数を増やすことが課題で、海外 DB 数の増加と記載内容の充実化を図ることが必要。その他の課題はデータベースの ID の統一化など、他のデータベースに関する情報との対応を持たせること。平成 22 年度は DB 情報の追加、更新、厚労省、農水省関係の DB の調査、補強を行った。 登録 DB 数 : 912 件 (平成 22 年 10 月現在) 開始年月 : 平成 20 年 2 月。平成 22 年 10 月の訪問者数 : 295/訪問回数 : 381/ページビュー : 500。
2	生命科学データベース横断検索 http://lifesciencedb.jp/dbsearch/ (英語版) http://lifesciencedb.jp/dbsearch/en/	公開	生命科学分野のデータベースや文献の横断検索サービスで、生命科学分野の国内外のデータベースと文献を網羅的に検索。平成 20 年 3 月にサービスを開始、平成 20 年度は対象データベースの拡大を図り、平成 21 年度はインデックスの更新管理並びに、関連キーワード表示機能や Blast 検索機能の開発を行った。また経産省統合 DB プロジェクトとの検索連携を実現した。H22 年度は対象データベースの充実を図るとともに、検索キーワードの翻訳結果や関連誤表示機能の実装やデータ間のリンク情報の格納を進めている。平成 22 年 9 月に英語版を限定公開。平成 22 年度 3 月末に改良インターフェースの公開予定。検索対象 DB 数 : 270 データベース (平成 22 年 10 月現在) 開始年月 : 平成 20 年 3 月。平成 22 年 10 月の訪問者数 : 11888/訪問回数 : 22460/ページビュー : 34462。
3	DNA データベース総覧と検索 http://lifesciencedb.jp/ddbj/ (英語版) http://lifesciencedb.jp/ddbj-e/	公開	巨大な国際 DNA バンクを高速に検索し生物種、分子種、プロジェクトに分類して表示するシステム。従来巨大な計算機を必要としていたが、バンク内部を詳細に検討し 1 億 300 万レコードを 58 万プロジェクトに縮退させ実質上、計算機一台での検索を可能にした。さらに、ある研究機関から発表されたある生物のゲノムデータを一括してダウンロードするというような研究単位でのデータ取得を実現した。BLAST による配列検索、核酸関連の特許公報へのダイレクトリンク、レコードの時系列展開も可能である。英語版サービスサイトの整備も完了。今後は、データベース受入れ、横断検索など他のサービスとの統合、検索の一層の高速化、EST の臓器分類や発現データとの統合を図る。検索対象 : 1 億 300 万レコード。開始年月 : 平成 19 年 3 月。平成 22 年 10 月の訪問者数 : 3914/訪問回数 : 16422/ページビュー : 95984/ダウンロード : 8.44Gbyte。
4	統合 TV	公開	データベースやウェブツールの使い方などを動画で発信するウェブサイト。動画コンテンツ

	http://togotv.dbcls.jp/ (英語版) http://togotv.dbcls.jp/?category=English		数は 366 に達した (平成 22 年 10 月現在)。開始年月 : 平成 19 年 7 月。平成 22 年 10 月の訪問者数 : 4326/訪問回数 : 15425/ページビュー : 74288/ダウンロード : 150.47Gbyte。
5	生命科学学協会カタログ http://lifesciencedb.jp/lbdb.cgi?gg=tool_academy	公開	生命科学系の学協会のデータベースと検索サービス。国内の生命科学系学協会を網羅したデータベース。各学会の詳細情報やウェブサイト、学会誌などの公開情報を掲載。平成 19 年度は日本学術会議の資料等をもとに生命科学系の学協会をリストアップしそれぞれの情報を調査、DB カタログと同じユーザインターフェースを利用して公開している。平成 21 年以降は各学会が出版している学会誌や大会要旨の情報調査を行った。掲載学協会数 : 600 学会 開始年月 : 平成 20 年 4 月。平成 22 年 3 月の訪問者数 : 47/訪問回数 : 63/ページビュー : 171。
6	学会要旨統合検索 http://lifesciencedb.jp/lbdb.cgi?gg=tool_tproc	公開	学会要旨集の検索サービス。著者名や施設名の表記揺れを吸収する辞書を搭載し、研究者の研究履歴をたどることが可能で、キーワードのトレンドを表示させることもできる。要旨集の提供を受けた遺伝学会と進化学会については全文検索全文表示を行っている。非公開の分子生物学会については、索引に対する検索のみ行っており要旨本文は表示しない。H21 年度には生物物理学会から要旨データの提供を受けた。学協会カタログや横断検索との重複を避けるため現在は更新を中断している。検索対象 : 4 学会 15729 要旨。開始年月 : 平成 20 年 4 月。平成 22 年 10 月の訪問者数 : 30/訪問回数 : 38/ページビュー : 43。
7	蛋白質核酸酵素全文検索 http://lifesciencedb.jp/pne/	公開 (更新分一部未公開)	蛋白質核酸酵素バックナンバーの全文検索サービス。国内で出版社により発行されている総説誌は、研究者にとって重要な研究発表の場であるとともに、無くてはならない良質な情報源でもある。共立出版の協力により蛋白質核酸酵素のバックナンバー電子化と公開が可能となり、すでに 1985~2007 年分を公開中である。残念ながら本雑誌は平成 22 年 1 月号をもって休刊となったが、休刊号までの総説をアーカイブし今後も提供を続ける予定である。収録論文 : 1985 年~2007 年 (全文表示分)~2009 年 (要旨表示分)。開始年月 : 平成 20 年 6 月。平成 22 年 10 月の訪問者数 : 4643/訪問回数 : 11758/ページビュー : 22325。
8	MotDB http://MotDB.dbcls.jp/	公開	DBCLS を担っていく人材であるアノテータ、キュレータ、システムデータベース管理者向けの教材を提供するサイト (MotDB は Master of the database の略)。中核機関で独自に作成した教材 PDF ファイルもダウンロード可能。講習会開催のたびに、情報が追加されている。開始年月 : 平成 19 年 3 月。平成 22 年 10 月の訪問者数 : 967/訪問回数 : 1595/ページビュー : 8109/ダウンロード : 1.29Gbyte。
9	遺伝子発現バンク (GEO) 目次 http://okubolab.genes.nig.ac.jp/bodymap_i/ (英語版) http://lifesciencedb.jp/geo-	公開	ヒト遺伝子の解剖学的な発現パターンデータの統合サイトを構築した。発現パターンは、測定法毎に異なる場合があることが知られており、ここではできるだけ客観的な発現パターンの解釈を可能にするために、5 種類の発現データ、即ち iAFLP、GeneChip、EST、NCBI の SAGEmap によるタグマップ、SAGE データの独自タグマップに基づく発現データを表示可能にした。平成 22 年 10 月の訪問者数 : 1318/訪問回数 : 2852/ページビュー : 10598/ダウンロード :

	e/ 生命科学系データベースアーカイブ http://dbarchive.lifescience.db.jp/	公開	885.65Mbyte。 データベースのアーカイブを公開するサービス。各登録データベースに対して、データベースとデータのメタデータ、利用許諾が示され、アーカイブデータのダウンロードが可能である。データの一括ダウンロードと共に、検索して絞り込んだデータのダウンロード機能も備えている。また、データベースの寄託を公募するためのページを作成し、その中で「寄託データベース利用許諾作成マニュアル」、「データベース利用許諾用語定義解説」などの作成を行い公開している。現在国内の生命科学系データベースが37件登録されている。開始年月：平成21年3月。平成22年10月の利用状況（訪問者数：1063、訪問回数：1470、ページビュー：4456、ダウンロード：2.93GB）
11	生命科学系主要プロジェクト一覧 http://togodb.dbcls.jp/lldb_project (英語版) http://togodb.dbcls.jp/lldb_project_en	公開	日本国内の生命科学系関連の主要なプロジェクトに関して詳細な情報を収集しまとめたもの。1990年代以降の4省（文科省、経産省、厚労省、農林水産省）の生命科学に関連する80の主要プロジェクトについて、プロジェクトの期間、概要、予算、代表研究者、公開の状況、成果、文献等に関する情報が閲覧できる。開始年月：平成20年3月。平成20年度まで「ゲノム・ポストゲノム主要プロジェクト一覧」として公開していたものの後継版。平成22年10月の訪問者数：12/訪問回数：12/ページビュー：12
12	TogoProt http://lifesciencedb.jp/togoprot/	公開	蛋白質関連データベース統合検索システム。ユーザは日本語または英語のキーワードで蛋白質を検索し、得た蛋白質リストから蛋白質を選ぶと、その蛋白質に関連づけられる各種のデータをデータベース横断的に集めて表示することができる。蛋白質の代わりに蛋白質ファミリーを選べば、ファミリーのいずれかのメンバーに関連づけられる各種データを表示できる。41のデータベースを収録。開始年月日：平成22年6月。平成22年10月の訪問者数：75/訪問回数：123/ページビュー：482
13	日本の生命科学データベース政策 http://lifesciencedb.jp/sciencepolicy/	公開	これまでにライフサイエンスのデータベースに関わる検討がいろいろなところで行われており、一元的に俯瞰できる状態ではなかったため、本プロジェクトに関与しそうな報告書、会議資料（ウェブページ情報等含む）等を1箇所に集約し、時系列に整理した。開始年月：平成21年3月。4月から外部公開した。平成22年10月の訪問者数：86/訪問回数：103/ページビュー：113
14	文科省「ゲノム」研究報告書全文検索 http://togodb.dbcls.jp/genome_sci_ryouiki1/	公開	ゲノム特定領域（2005年-2009年）の年次報告書データ一覧サイト。全文検索可能で、全235の題目ごとのメタデータつきで、報告書のPDFファイルのダウンロードが可能。開始年月日：平成20年6月。平成22年10月の訪問者数：4106/訪問回数：6135/ページビュー：7273
15	Togoexp (RefEx) http://togoexp.dbcls.jp/	公開	RefEx (Reference Expression dataset) は、提供する4種類の異なる手法 (EST, GeneChip, iAFLP, CAGE) によるヒトおよびマウスの遺伝子発現データのリファレンスデータセット。個々の遺伝子については染色体領域や遺伝子ファミリー (InterPro)、Gene Ontology に基づ

			く生物学的機能別に検索ができる他、発現パターンから探すことが可能。開始年月日：平成21年12月。平成22年10月の訪問者数：111/訪問回数：160/ページビュー：914
16	ライフサイエンス 新着論文レビュー http://first.lifesciencedb.jp/	公開	<i>Nature, Science, Cell</i> などに代表されるトップジャーナルに掲載された日本人を著者とする生命科学分野の論文について、論文の著者自身の執筆による、専門分野の異なる生命科学研究者にむけた日本語によるレビューを、だれでも自由に閲覧・利用できるよう Web 上にていち早く無料で公開。開始年月日：平成22年9月。平成22年10月の訪問者数：2756/訪問回数：11294/ページビュー：44484

(2) DB基盤システム、ツール等開発成果物の整備

通番	DB基盤システム、ツール等の名称	公開／未公開	概要（主な機能・特徴点、進捗状況、今後の計画などを簡潔にわかりやすく記述）
1	遺伝子名称シソーラス http://lifesciencedb.mext.go.jp/result/tech.html#id2-1-1	公開	分子情報から医学文献までの遺伝子や蛋白質名称の正確な同定のための遺伝子固有名称（「遺伝子名」と一般名称（「ファミリー名」）の辞書データの構築を目的として、様々なデータベースで利用されている名称の収集と専門的キュレータによる編集を行い、遺伝子が持つ多様な名称の関係を明示した遺伝子名称シソーラスを開発した。本シソーラスは、ヒトをはじめ9種類の生物をカバーしている。
2	生物学名羅日英対応辞書 http://lifesciencedb.mext.go.jp/result/tech.html#id2-1-1	公開	研究分野でよく使われる生物種の基準として、学名に日本語一般名を対応させた生物学名日本語一般名対応辞書を開発した。対応付けは、塩基配列データベース (DDBJ) の登録エントリ数が多い生物種から順番に行った。また、標準和名が存在しない場合、その生物を説明する一般的な名称を用いた。今年度は追加作業を継続。登録件数：学名（ラテン語）と他言語での種名の対応 - 77,625 件、学名と和名の対応 - 66,346 件、そのほか、学術用語の日英対応 - 42,823 件。
3	施設名称辞書 http://lifesciencedb.mext.go.jp/result/tech.html#id2-1-1 (http://lifesciencedb.jp:32888/agency/agency.cgi)	公開	日本のライフサイエンス研究を俯瞰するための重要な情報源として、各種関連学会の抄録などの報告文書があるが、これらをデータベース化する際に問題となるのが、例えば、大阪大、阪大、大阪大学大学院などといった、施設名称研究室名称の表記ゆれである。これに対応するために、施設名称辞書を開発した。すでに公開しているものは1625の名称についてまとめたものであるが、昨年度からその範囲を15,000件以上の研究機関・研究グループに広げ、ホームページ掲載情報や学会要旨データベースより収集した機関名別名情報を格納し、機関名の表記ゆれに対応した検索を可能とした。今年度も辞書の拡張を継続し、引き続き内部公開で試験運用中である。
4	動物解剖学自動分類タガー http://lifesciencedb.mext.go.jp/result/tech.html#id2-1-1	公開	動物の臓器、組織を、大きく10のグループに分類(大分類)し、さらにそれぞれのグループを細かく分類、合計40の小分類グループに分類する。基本的には、与えられた解剖用語に対して、解剖用語辞書、病理関連語彙の分類辞書、形容詞の解剖用語辞書、一般的な臓器名称の分類辞書、の4種類の辞書を順番に検索し、上記のカテゴリに分類する。

5	植物解剖学自動分類タガー http://lifesciencedb.mext.go.jp/result/tech.html#id2-1-1	公開	植物(維管束植物)の部位、組織を大きく6のグループに分類し、さらにそれぞれのグループを細かく分類、合計11の小分類グループに分類する。分類における検索対象としては、生物種に合わせて、種子を持たない維管束植物の解剖用語辞書、イネ科の解剖用語辞書、その他被子植物の解剖用語辞書、裸子植物の解剖用語辞書、トウモロコシ属の解剖用語辞書、フウチョウソウ目の解剖用語辞書を用いる。
6	都市名国名自動検出タガー http://lifesciencedb.mext.go.jp/result/tech.html#id2-1-1	公開	論文やデータベースレコードに見られる国の名称の未記載や国の名称にみられる表記ゆれを吸収することを目的に、国別に分類するための辞書を開発した。
7	学術用語集	公開	文部科学省発行の学術用語集のうち、生命科学分野に関連のあるもの医学編、農学編、植物学編、化学編、歯学編、薬学編、動物学編について、学術用語と対訳情報を取り出し対訳辞書としてまとめ公開した。開始年月：平成21年8月、更新年月：平成22年10月。登録用語数-150,592件
8	辞書の構築と公開 http://lifesciencedb.jp/lldb.cgi?gg=dic	公開	これまでに構築、集積してきた、辞書、シソーラス、オントロジーの公開サイト。学術用語、学名、学名と他言語での種名の対応、学名と和名の対応、学術用語の日本語と英語の対応、解剖学用語の和英ラテン対応辞書、遺伝子名称シソーラス、施設名辞書、解剖学辞書がダウンロード可能。開始年月：平成21年9月。平成22年10月の訪問者数：690/訪問回数：827/ページビュー：1850。
9	生物アイコン http://lifesciencedb.jp/lldb.cgi?gg=resource_icon (英語版) http://lifesciencedb.jp/lldb.cgi?gg=resource_icon&lng=en	公開	ユーザが自身のHPなどで自由に利用できるように提供されている生物のアイコン集。目的とするイラストをキーワード検索でき、生物に関する外部情報のリンクも付いている。ユーザはイラストに関するコメントを書き込むことができ、さらには同じ条件で配布可能なイラスト等を投稿することもできる。アイコン数：220。開始年月：平成20年6月。定期的にアイコン数を増やし、インターフェースを更新。ヘルプページを作製した。平成22年10月の訪問者数：249/訪問回数：307/ページビュー：452/ダウンロード：86.4Mbyte。
10	BodyParts3D http://lifesciencedb.jp/bp3d	公開	解剖学用語、すなわち臓器・器官・部位の名称やそれらをくくる概念をモデル人間中の3次元座標で定義した辞書。BodyParts3D内の空間関係で用語関係を表現すると、ツリー型表現(いわゆる解剖オントロジー)と違い、改変にロバスタな表現が可能。今後も用語の追加を続ける。現在、全パーツ数が1314種である。生命科学系データベースアーカイブから全ポリゴンメッシュデータがダウンロード可能。今年度は脳モデルの改修を行った。アクセス情報はアナトモグラフィーと一緒にカウントしているため下記参照。
11	アナトモグラフィー http://lifesciencedb.jp/bp3d http://lifesciencedb.jp/bp3d?lng=en (英語)	公開	BodyParts3Dから部品を選択して自由に人体の部分のモデル図が書ける。視点固定の図譜や3Dアトラスと異なり透過度や着色、視点は自由に選択可能。今後は、解剖学教育や医学用語辞書の助けになる人体版グーグルマップを目指す。すでに公開したバージョンに加え、拡大・縮小、平行移動など自由度が高くかつ高速な図形処理が可能なバージョンを公開した。平成22年10月の訪問者数：3879/訪問回数：5432/ページビュー：143427/ダウンロード：

			27.93Gbyte。
12	TogoDB http://togodb.dbcls.jp/	公開	データベース受け入れシステム。さまざまな形式のさまざまな種類のデータを簡単に読み込み、自動的に共通のウェブ検索インターフェースを生成する。また、自動的にウェブサービス(SOAP/WSDL API)を生成する。これらの機能により、容易に維持困難なDBを移設できる。また、本ツールのダウンロードにより、誰でもDBを構築可能である。平成22年10月の訪問者数：1348/訪問回数：6181/ページビュー：114779/ダウンロード：5.27Gbyte。
13	TogoWS http://togows.dbcls.jp/	公開	国内外のウェブサービスを透過的に統合するシステム。国内各サービスの稼働状況を継続的に監視し、サービスの連携に必要なデータ形式変換機能等を提供する。これにより、ユーザは統一的命名規則と統一的データ構造で国内外のウェブサービスにアクセスすることができるため、相互運用が容易となり、ワークフローの作成・蓄積を促進する。開始年月：平成20年4月。平成22年10月の訪問者数：433/訪問回数：973/ページビュー：36536/ダウンロード：2.25Gbyte。
14	OpenID 認証システム http://openid.dbcls.jp/ (英語版) https://openid.dbcls.jp/?lang=en	公開	一つのIDで複数のサイトを認証できるアカウントシステム。このシステムを利用することで、ユーザはひとつのIDとひとつのパスワードを保持すればよい。これにより、各サイトで認証サービスを用意する必要がなく、サイト間のユーザ情報の集約が容易になる。開始年月：平成20年4月。平成22年10月の訪問者数：300/訪問回数：663/ページビュー：7138。
15	OReFiL http://orefil.dbcls.jp/ (英語版) http://orefil.dbcls.jp/en	公開	オンライン上に存在する多数の生命科学系の資源(データベースやソフトウェアなど)を効率的に見つけるための検索システム。信頼性や概要を容易に取得するために、資源の所在だけでなく、その資源について書かれている論文情報や、他の利用者の評価を閲覧できる。アカウントを取得することで、自ら評価を行うことも可能である。また、検索に関連する機能を全てウェブサービスとして提供しているため、利用者のプログラムからOReFiLを利用できる。英語版ヘルプなどを整え、環境を整備した。開始年月：平成19年8月。平成22年10月の訪問者数：365/訪問回数：556/ページビュー：2227。
16	Allie http://allie.dbcls.jp/ (英語版) http://allie.dbcls.jp/	公開	医学生物系論文誌情報データベースMEDLINEを対象とし、出現する略語とその転回形のペアを検索するシステム。生命科学系の論文では非常に多くの略語が使われており、同じ表記でも全く違う意味を示していることが少なくない。そこで、利用者の興味のある略語を検索語として入力することで、その使われ方を一覧表示すると共に、論文の発表年を提示する。また、検索された各ペアについて、それが出現する論文中で共起する他の略語も提示する。検索に関連する機能を全てウェブサービスとして提供しているため、利用者のプログラムからAllieを利用できる。英語版ヘルプなどを整え、環境を整備した。開始年月：平成20年4月。平成22年10月の訪問者数：936/訪問回数：1625/ページビュー：4192。
17	DBCLS Galaxy http://galaxy.dbcls.jp/	公開	生命科学データに特化した、ウェブベースの対話的ツール組み合わせインターフェース。マウス操作と簡単な入力で、プログラムを書くことなく、途中結果を確認しながら、用意され

			たツールを対話的に組み合わせて使うことができる。DBCLS Galaxy では、TogoWS やテキストマイニングツールなど、DBCLS で開発されたツール群を組み込み、利用できるようにした。ログインは DBCLS OpenID で可能である。また、メニューの国際化対応へ協力し、メニューの日本語化を行い、DBCLS Galaxy のみならず、元の Galaxy にも反映させた。平成 22 年 10 月の訪問者数：190/訪問回数：260/ページビュー：7013。
18	ScrapParty	公開	ブラウザのプラグインによるキュレータ支援用アプリケーション。オンラインで文献やウェブサイトを閲覧しながら、重要な記述や画像、図表を集めコメントとともに簡単に DB 化するツール。平成 18 年度アノテーションの調査、試験に利用した。下記 Wired-Marker が後継アプリケーションであり、現在 ScrapParty 名での開発は終了。
19	Wired-Marker http://www.wired-marker.org	公開	情報共有ツール、アノテーション支援ツール。Firefox ブラウザのプラグインとして働く。データベース構築、アノテーションを支援するブラウザ用のプラグイン。オンラインで閲覧したページ、論文のテキストや図をマークしブックマークすることが可能。再訪問の際にも同じ箇所がハイライトされアノテーションのエビデンスとして詳細な情報が残せる。またフォルダを自由に階層化して情報を整理することが可能。新たにキャッシュ機能が加わり、ウェブページの書き換えが起きても過去のデータを参照できるようになった。 公開日：H19.12.19 総ダウンロード数：366,163 件。
20	MiGAP http://migap.lifesciencedb.jp/ (英語版) http://migap.lifesciencedb.jp/en/	公開	技術革新により大量に生産されるようになった微生物ゲノムの利用価値を高めるため、アノテーションを自動的に行うパイプラインを開発し、試験的にプロジェクト内部に公開、評価を受けた。評価の意見を反映した修正作業を進め、平成 21 年 6 月に一般公開。利用のためには、OpenID の取得が必要。開始年月：平成 21 年 3 月。平成 22 年 10 月の訪問者数：170/訪問回数：383/ページビュー：186383。
21	inMeXes http://docman.dbcls.jp/im/ (英語版) http://docman.dbcls.jp/im/	公開	医学生物系論文誌情報データベース MEDLINE を対象とし、利用者が入力した文字列とマッチする表現を逐次的(1文字入力する度)に検索するシステム。生命科学系分野の学術文献を英語で執筆する際、既に発表されている論文中で実際に使われている表現を容易に検索出来るれば、特に非ネイティブスピーカーが多数の日本人研究者にとって有用であると考え、構築した。マッチした表現の全 MEDLINE 中(題目・要旨)での出現頻度も表示されるほか、当該表現が使われている文や関連情報を表示させることも可能。開始年月：平成 21 年 7 月。平成 22 年 10 月の訪問者数：508/訪問回数：1516/ページビュー：134342。
22	TogoDoc http://docman.dbcls.jp/pubmed_recom (英語版)	公開	文献情報及び論文 PDF を管理し、また、特定の文献情報群に関連する論文情報を提示するシステム。日々多くの論文が発表されている昨今においては自身が興味を持つ文献情報を効率良く取得することが容易では無くなってきている。また、多くの学術雑誌がオンラインジャーナルを発行していることから論文全文を PDF ファイルで取得出来る機会が増大し、研究者の PC に多くの論文 PDF ファイルが保存されている。以上の背景から文献情報の効率的な

	http://docman.dbcls.jp/pubmed_recom		取得、整理環境を実現すべく、そのために必要な情報は既に研究者自身の手元にある文献情報リストから得られるようなシステム TogoDoc を構築した。利用者 PC に保存されている PDF ファイルを解析し必要な書誌情報を自動的に付加するなどの機能を持つ文献情報管理ソフトウェア TogoDocClient と書誌情報のやり取りや PDF ファイルの授受を行うなどの連携機能を持つ。公開年月：平成 21 年 12 月。
23	Gendoo http://gendoo.dbcls.jp/	公開	Gendoo は、遺伝子や疾患について、関連する疾患や薬剤、生命現象、臓器などの情報をキーワードで表示するシステム。入力として、Gene ID、遺伝子名、OMIM ID、疾患名 (OMIM のタイトル)、キーワード (MeSH terms) : 疾患、薬剤、生命現象、臓器を受け付け、関連する語を表示する。開始年月：平成 19 年 12 月試験公開。平成 22 年 10 月の訪問者数：1639/訪問回数：2018/ページビュー：3445。

(3) その他の成果物 ((1)、(2) に該当しないもの)

通番	名称	公開／未公開	概要
1	統合 PJ 成果公開サイト http://lifesciencedb.jp	公開	統合 DB プロジェクトに関する中核機関の成果公開サイト。生命科学データベースカタログ、DB 横断検索、INSD 全文検索、統合 TV、OReFiL、生命科学学協会検索、学会要旨統合俯瞰システム、MotDB、LSDB ブログ等を掲載している。また、プロジェクト参加機関の成果公開サイトへのリンクも張ってある。全掲載サービスは 60 件で、DBCLS 分 40 件、その他参加機関分 20 件である。開始年月：平成 20 年 6 月。DBCLS 提供サービスへのアクセス状況としては、平成 22 年 10 月の訪問者数：49639/訪問回数：118782/ページビュー：689712。
2	統合 PJ 成果公開サイト(英語版) http://lifesciencedb.jp/?lng=en	公開	国外で本プロジェクトの紹介をする機会も増えてきたことから、英語での利用が可能なサービスを集めて、英語版のポータルサイトを公開した。現在紹介されているデータ・サービスは、ゲノム・ポストゲノム主要プロジェクト一覧、生命科学データベースカタログ、生物アイコン、DNA データベース総覧と検索、遺伝子発現バンク (GEO) 目次、アナトモグラフィ/BodyParts3D、Wired-Marker、TogoWS、togoTV。そのほか、プロジェクト紹介ページ、参画機関の成果ページ(英語版があるもの)へのリンクを張った。開始年月：平成 21 年 3 月。
3	統合データベース間の連携調査 http://charles.kazusa.or.jp/lifesciencedb/	公開	代表的モデル研究植物である「イネ」ならびに「シロイヌナズナ」のゲノムアノテーション型公開データベース各 46 サイト、25 サイトの基本項目の調査と実験生物研究者 188 名のアンケートにより調査し、それぞれのデータベース間の連携のための現状の課題と将来の DB 統合にむけた課題の整理を行った。
4	検索アルゴリズムを含めた知識情報技術の動向調査 http://lifesciencedb.mext.go.jp/result/strategy.html#id1	公開	次世代の生物情報データベース統合に必要な知識情報技術として、検索システム、データマイニング、Web 2.0 およびグリッドコンピューティングに焦点を絞り、聞き取り調査や文献調査によって動向を調べた。

	-4		
5	遺伝統計学分野における解析技術の基礎調査 http://lifesciencedb.mext.go.jp/result/strategy.html#id1-4	公開	遺伝統計学分野で用いられる連鎖解析、連鎖不平衡解析)、QTL解析等の解析手法と、それぞれの手法における代表的なアルゴリズム計8種類の調査を行ない、その特徴および長所・短所を評価した。併せて、各手法の代表的プログラム計15種類に関して、実装されているアルゴリズム、動作環境などを調査し、その評価を行った。
6	臨床情報・疾患健康情報の調査 http://lifesciencedb.mext.go.jp/result/strategy.html#id1-4	公開	わが国の代表的コホート研究に関してその研究の背景と目的、対象地域、ターゲット疾患、特徴的な検査項目、対象人数、代表研究者、研究開始時期、資金源等について調査した。
7	データベース受入に関するアンケート調査 http://lifesciencedb.jp/lldb.cgi?pg=0&gg=ukeire_survey	試験的公開	受入による統合化の方針策定と受入対象の候補抽出を目的として、国内データベースを対象として、データベース内容と検索方法、経歴、維持・管理状況、提供可能性及び提供時の問題点などをアンケートにより調査した。送付した104件中、65件の回答が得られ、半数以上の機関が、統合DB事業に対して全データもしくは一部データを開示できると回答した。結果の概要および詳細報告(PDF)をウェブ上で公開した。
8	BioHackathon2008 ホームページ http://hackathon.dbcls.jp/	公開	分散環境のまま、各種DBをリアルタイムかつ統合的に利用できる環境を構築すること目的に開催した、国内外のライフサイエンスDBのウェブサービスにおけるデータ構造や命名規則を標準化するためのBioHackathon2008の活動内容と経過を紹介するサイト。2008年2月11日から15日まで都内で開催。海外参加者34名、国内参加者37名。平成22年10月の訪問者数：221/訪問回数：682/ページビュー：2278。
9	BioHackathon2009 ホームページ http://hackathon2.dbcls.jp/	公開	標準化技術および標準化を利用する技術をより安定的に使用可能にすることを目的として開催したBioHackathon2009の活動内容を紹介するサイト。巨大なデータについてどうすべきか、セマンティックウェブ技術をどう利用すべきかなどの議論の報告が掲載されている。2009年3月16日から20日まで、都内および沖縄で開催、海外参加者30名、国内参加者28名。平成22年10月の訪問者数：320/訪問回数：422/ページビュー：3348。
10	BioHackathon2010 ホームページ http://hackathon3.dbcls.jp/	公開	統合検索に向けたセマンティックウェブ技術の調査と開発を目的とした、BioHackathon2010の活動内容を紹介するサイト。2010/2/8 - 12の期間、海外32名、国内36名が参加し、LinkedDataを実現するための合意形成、セマンティックウェブ技術のフィージビリティ調査のための開発などが行われた。平成22年10月の訪問者数：309/訪問回数：1143/ページビュー：4203。
11	シンポジウム・講演会 http://symposium.lifesciencedb.jp/	公開	DBCLSが主催したシンポジウムサイトをリストしたサイト。現在は2009年6月12日および10月5日、2010年10月5日開催の3つのシンポジウムについてリンクがあり、2009年6月12日分のサイトでは、講演動画9件・スライド7件・発表ポスターPDF68件、2009年10月5日分のサイトでは講演動画8件を掲載している。2010年のサイトでは発表ポスターPDF28

			件を公開、講演動画 9 件・スライド 10 件を公開準備中。平成 22 年 10 月の訪問者数：672/訪問回数：1292/ページビュー：3838。
12	外部評価 http://lifesciencedb.jp/lldb.cgi?gg=hyouka	公開	登録ユーザによるサービスへの評価内容を公開するサイト。平成 19 年から基本的に前年度の公開サービスへの評価を実施。平成 20 年度からはウェブサイト記入方式（5 段階評価とフリーコメント）に変更し、サービスについての理解を深めてもらう工夫も行った。なお、いただいたご意見への対応・回答も併せて掲載している。平成 21 年度には、前年度の要望に基づき参考資料として 20 年度の各機関の予算と目標/進捗情報を掲載した。22 年度はプロジェクト最終年度のためプロジェクト全期間について全ての参画機関についての評価を行うこととし、全期間の予算情報と 20 年度以降の目標/進捗情報、各サービスのアクセス情報（22 年度新規公開分は除く）も掲載した。評価者の人数：約 60 名（平成 19 年度）、58 名（平成 20 年度）、117 名（平成 21 年度）。平成 22 年 10 月の訪問者数：206/訪問回数：283/ページビュー：8675。
13	運営委員会等の資料 http://togodb.dbcls.jp/koukai	公開	平成 18 年度から行われてきた運営委員会、作業部会、作業部会分科会の会議資料と議事概要および委員名簿を掲載。開催日や資料名称の簡易検索機能付き。平成 22 年 10 月の訪問者数：244/訪問回数：288/ページビュー：3838。
14	統合データベース技術情報交換ワークショップ http://wiki.lifesciencedb.jp/mw/index.php/BH10.10	公開	次世代のライフサイエンスデータベースをより利便性の高いものとすることを目指して、これまでの BioHackathon で蓄積された知識を国内で共有する目的で行ったワークショップの活動を紹介するサイト。2010/10/18 - 22 の期間、修善寺で開催し、約 70 名の参加者があった。

外部発表実績一覧

(別紙2)

学会等発表実績

委託業務題目: 「ライフサイエンス統合データベース開発運用」(戦略立案・実行評価/統合データベース開発/統合データベース支援)

機 関 名: 大学共同利用機関法人 情報・システム研究機構

1. 学会等における口頭・ポスター発表

発表した成果(発表題目、口頭・ポスター発表の別)	発表者氏名	発表した場所(学会等名)	発表した時期	国内・外の別
平成 22 年度				
新たな日本語 Web コンテンツ「ライフサイエンス新着論文レビュー」[口頭]	飯田 啓介	第 6 回 SPARC Japan セミナー2010 「日本発オープンアクセス」	2010 年 10 月 20 日	国内(国立情報学研究所)
“Microbial Genome Annotation Pipeline (MiGAP)”	Sugawara H, Ohyama A, Mori H and Kurokawa K.	Biocuration 2010	2010 年 10 月 11-14 日	国内(台場)
Development of Cooperative Work Tools and Workflow for Multimodal Genome Annotation [ポスター]	Okamoto S, Nakao M	Biocuration 2010	2010 年 10 月 11-14 日	国内(台場)
TogoDB + TogoWS: a data integration platform for the Semantic Web [口頭]	Toshiaki Katayama and Mitsuteru Nakao	The Conference of International Society for Biocuration (Biocuration2010)	2010 年 10 月 11-14 日	国内(東京台場)
データ・情報の円滑な利用に向けた権利関係の整理の試み [ポスター]	箕輪真理、高祖歩美、中尾光輝、河野信、八塚茂、畠中秀樹、川本祥子、永井啓一、大久保公策、高木利久	シンポジウム 「ライフサイエンスの未来へ～10 年先のデータベースを考える～」	2010 年 10 月 5 日	国内(東京大学本郷キャンパス浅野地区武田ホール)

文部科学省「統合データベースプロジェクト」の 広報・普及活動 [ポスター]	高祖歩美、箕輪真理	シンポジウム 「ライフサイエンスの未 来へ～10年先のデータベースを考える ～」	2010年10月5日	国内(東京大学 本郷キャンパ ス浅野地区 武田ホール)
統合データベースの教育・人材育成活動～統合 TVとMotDB～ [ポスター]	河野信、小野浩雅、坊農秀 雅	シンポジウム 「ライフサイエンスの未 来へ～10年先のデータベースを考える ～」	2010年10月5日	国内(東京大学 本郷キャンパ ス浅野地区 武田ホール)
多様なゲノム注釈のためのツールと協働作業手 法の開発 [ポスター]	岡本忍、中尾光輝、藤澤貴 智、佐藤修正、金子貴一、 吉村英尚、中平有香、阿久 津智子、鐘ヶ江弘美、笠井 誠、山本純子、作田千代子、 眞板寛子、鹿島裕子、中村 保一	シンポジウム 「ライフサイエンスの未 来へ～10年先のデータベースを考える ～」	2010年10月5日	国内(東京大学 本郷キャンパ ス浅野地区 武田ホール)
ウェブサービス相互運用、ワークフロー、デー タ共有のための基盤技術開発(2) [ポスター]	中尾光輝、山口敦子、山本 泰智、片山俊明、川島秀一	シンポジウム 「ライフサイエンスの未 来へ～10年先のデータベースを考える ～」	2010年10月5日	国内(東京大学 本郷キャンパ ス浅野地区 武田ホール)
ウェブサービス相互運用、ワークフロー、デー タ共有のための基盤技術開発(1) [ポスター]	中尾光輝、山口敦子、山本 泰智、片山俊明、川島秀一	シンポジウム 「ライフサイエンスの未 来へ～10年先のデータベースを考える ～」	2010年10月5日	国内(東京大学 本郷キャンパ ス浅野地区 武田ホール)
論文執筆の際に役立つサービスの紹介 [ポスタ ー]	山本泰智、山口敦子、坊農 秀雅、高木利久	シンポジウム 「ライフサイエンスの未 来へ～10年先のデータベースを考える ～」	2010年10月5日	国内(東京大学 本郷キャンパ ス浅野地区 武田ホール)
生物学名和名対訳辞書の構築と公開 [ポスター]	森山丈士、栗原英輔、小笠 原理、川本祥子、大久保公 策	シンポジウム 「ライフサイエンスの未 来へ～10年先のデータベースを考える ～」	2010年10月5日	国内(東京大学 本郷キャンパ ス浅野地区 武田ホール)

統合データベースプロジェクトの新しい日本語コンテンツ「新着論文レビュー」[ポスター]	飯田啓介、中尾光輝、川本祥子、高木利久	シンポジウム「ライフサイエンスの未来へ～10年先のデータベースを考える～」	2010年10月5日	国内(東京大学本郷キャンパス浅野地区武田ホール)
「生命科学系データベースアーカイブ」の成果と今後 [ポスター]	坂東明日佳、熊谷禎洋、八塚 茂、西川哲夫、畠中秀樹、大久保公策、高木利久	シンポジウム「ライフサイエンスの未来へ～10年先のデータベースを考える～」	2010年10月5日	国内(東京大学本郷キャンパス浅野地区武田ホール)
生命科学系データベース情報の収集とポータルサイト構築 [ポスター]	川本祥子	シンポジウム「ライフサイエンスの未来へ～10年先のデータベースを考える～」	2010年10月5日	国内(東京大学本郷キャンパス浅野地区武田ホール)
ライフサイエンス統合データベースの横断検索サービスの現状 [ポスター]	杉崎太一朗、牧口大旭、高橋順子、平井信一、奥村利幸、谷口仁、畠中秀樹、川本祥子、高木利久	シンポジウム「ライフサイエンスの未来へ～10年先のデータベースを考える～」	2010年10月5日	国内(東京大学本郷キャンパス浅野地区武田ホール)
TogoProt による蛋白質関連データベースの統合検索 [ポスター]	畠中秀樹、平井信一、多門啓子、大野忠、牧口大旭、奥村利幸、川本祥子、中村春木、高木利久、	シンポジウム「ライフサイエンスの未来へ～10年先のデータベースを考える～」	2010年10月5日	国内(東京大学本郷キャンパス浅野地区武田ホール)
遺伝子発現リファレンスデータセット『RefEx』 [ポスター]	小野浩雅、大久保公策、高木利久、坊農秀雅	シンポジウム「ライフサイエンスの未来へ～10年先のデータベースを考える～」	2010年10月5日	国内(東京大学本郷キャンパス浅野地区武田ホール)
BodyParts3D と Anatomography: 3次元人体構造データベースと人体解剖図作成サービス [ポスター]	三橋信孝、藤枝香、今井紫緒、田村卓郎、川本祥子、高木利久、大久保公策	シンポジウム「ライフサイエンスの未来へ～10年先のデータベースを考える～」	2010年10月5日	国内(東京大学本郷キャンパス浅野地区武田ホール)
Reference Expression dataset (RefEx) for practical gene expression analysis of cancer-related molecular targets [口頭]	坊農秀雅	第69回日本癌学会学術総会	2010年9月22-24日	国内(大阪国際会議場)

データおよび知識の共有と統合化に向けて [口頭]	高木 利久	第 48 回日本生物物理学会年会 フォーラム「デジタル革命の本当のインパクト」	2010 年 9 月 20 日	国内(東北大川内キャンパス)
デジタル革命の本当のインパクト[口頭]	大久保 公策	第 48 回日本生物物理学会年会 フォーラム「デジタル革命の本当のインパクト」	2010 年 9 月 20 日	国内(東北大川内キャンパス)
RefEx: Reference expression dataset for practical use of gene expression data [ポスター]	Hidemasa Bono, Hiromasa Ono, Kousaku Okubo, Toshihisa Takagi	The 10th Cold Spring Harbor Laboratory/Wellcome Trust conference on Genome Informatics	2010 年 9 月 15-19 日	国外 (Hinxton (UK))
Functional indexing and curation of next-generation sequencing data [ポスター]	Takeru Nakazato, Hidemasa Bono, Toshihisa Takagi	The 10th Cold Spring Harbor Laboratory/Wellcome Trust conference on Genome Informatics	2010 年 9 月 15-19 日	国外 (Hinxton (UK))
“WFCC World Data Centre for Microorganisms (WDCM) hosted by the National Institute of Genetics: Lessons Learned” [招待講演]	Hideaki Sugawara	Research Network for Applied Microbes & International symposium on microbial resources management, utilization and fungal DNA barcode (Chinese Academy of Science Institute of Microbiology (Beijing))	2010 年 8 月 24 日	国外 (北京)
統合データベースプロジェクト: ライフ・イノベーションによる知のめぐりのよいオープンな生命科学を目指して [口頭]	坊農秀雅	平成 22 年度脳科学研究戦略推進プログラム 課題 A, B 分科会プログラム	2010 年 7 月 27-28 日	国内(ホテルさっぽろ芸文館)
ゲノムアノテーション協同作業を支援する手法と技術[口頭]	岡本 忍	ラン藻ゲノム研究会	2010 年 7 月 24 日	
TogoProt: 蛋白 ID を用いたデータベース横断検索システム [ポスター]	畠中秀樹、平井信一、大野忠、多門啓子、高橋頼子、押田健寛、牧口大旭、奥村利幸、川本祥子、西川哲夫、中村春木、高木利久	第 10 回日本蛋白質科学会年会	2010 年 6 月 16-18 日	国内(札幌コンベンションセンター)
ゲノム注釈の民主化 [口頭]	岡本 忍	農業環境技術研究所公開セミナー	2010 年 6 月 14 日	国内 (つくば)

DBCLS Galaxy: A Galaxy Service to integrate databases in Japan [口頭]	Atsuko Yamaguchi	Galaxy Developer Conference 2010	2010年5月15-17日	国外(NewYork)
Galaxy Internationalization (i18n) and Localization (L10n) [口頭]	Mitsuteru Nakao	Galaxy Developer Conference 2010	2010年5月15-17日	国外(NewYork)
「科学技術コモンズと情報知識学の挑戦」ライフサイエンスにおける統合データベースの構築と課題 [招待講演]	高木 利久	情報知識学会 第18回(2010年度)年次大会記念シンポジウム	2010年5月15日	国内(東大本郷工学部2号館)
平成21年度				
生命科学系データベースアーカイブ [口頭]	畠中 秀樹	日本農芸化学会2010年度大会 ランチオンセミナー「情報整理術2010～文献とデータを賢く管理する～」	2010年3月27-30日	国内(東京)
情報メディアとしてのゲノム [口頭]	岡本 忍	GCOE 若手研究異分野セミナー	2010年3月28-29日	国内(北海道)
内閣府統合データベース事業と理研への期待[口頭]	高木 利久、大石道夫	第18回理事長ファンドワークショップ「理研総合データベースを核にした研究基盤連携網の構築」	2010年3月12-13日	国内(葉山)
内閣府統合データベース事業と理研への期待、文科省ライフサイエンス統合データベースプロジェクトの目的、現状、課題[口頭]	高木 利久	第18回理事長ファンドワークショップ「理研総合データベースを核にした研究基盤連携網の構築」	2010年3月12-13日	国内(葉山)
微生物ゲノム基盤情報の高度化と統合のためのシステム	岡本 忍	第4回日本ゲノム微生物学会年会	2010年3月7-9日	国内(九州)
国際開発会議 DBCLS バイオハッカソン2010 レポート[口頭]	中尾 光輝	人工知能学会第二種研究会 分子生物情報研究会 (SIG-MBI)	2010年3月5-6日	国内(能美)
ゲノムアノテーション情報の高度化統合化への取り組み [口頭]	岡本 忍	かずさ研究交流セミナー	2010年1月27日	国内(千葉)

BIOLOGICAL DOCUMENT CLUSTERING USING A GENE-GENE NETWORK[ポスター]	Hong-Woo Chun, Shinobu Okamoto, Yasunori Yamamoto, Atsuko Yamaguchi, Toshihisa Takagi	PSB(Pacific Symposium on Biocomputing) 2010	2010年1月4-8日	国外 (Hawaii)
DDBJ Read Archive and DDBJ Read Annotation Pipeline:An Archive Database and an Analytical Tool for Next-Generation Sequence Data T003/P120[ポスター/口頭]	Eli Kaminuma, Yuichi Kodama, Satoshi Saruhashi, Takeshi Konno, Takako Mochizuki, Hidemasa Bono, Hideaki Sugawara, Kousaku Okubo, Toshihisa Takagi, Yasukazu Nakamura	The 20th International Conference on Genome Informatics (GIW2009)	2009年12月14-16日	国内 (横浜)
なぜ今共有なのか？個人プレーから社会プレーへ[口頭]	大久保 公策	第32回日本分子生物学会年会、フォーラム「デジタル時代のジレンマ」”共有と秘匿のバランス”	2009年12月10日	国内 (横浜)
データベース統合センターの役割[口頭]	高木 利久	第32回日本分子生物学会年会、フォーラム「デジタル時代のジレンマ」”共有と秘匿のバランス”	2009年12月10日	国内 (横浜)
実験生物学者にとって有益な情報環境を提案する機会：「統合データベースプロジェクト」のユーザ評価[口頭]	高祖 歩美、箕輪 真理	第32回日本分子生物学会年会、ワークショップ	2009年12月10日	国内 (横浜)
ブタ成熟脂肪細胞および顆粒層細胞における脱分化ならびに多能性獲得機構の統合トランスクリプトミクス[ポスター]、[口頭]	小野 浩雅、坊農 秀雅、加野 浩一郎	第32回日本分子生物学会年会、ワークショップ	2009年12月10日	国内 (横浜)
かずさアノテーションを用いた分散型ゲノムアノテーションの実証実験[口頭]	岡本 忍、中尾 光輝、藤沢 貴智、佐藤 修正、金子 貴一、中村 保一	第32回日本分子生物学会年会、ワークショップ	2009年12月10日	国内 (横浜)
統合データベースプロジェクト：オープンで知のめぐりの良い分子生物学を目指して[口頭]	坊農 秀雅	第32回日本分子生物学会年会、ワークショップ	2009年12月10日	国内 (横浜)

DBCLS で提供中の文献関係サービス[ポスター]	山本 泰智、山口敦子、坊農秀雅、高木利久	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内（横浜）
KazusaMart: 植物および関連微生物のゲノムデータ統合[ポスター]	藤澤 貴智、岡本 忍、中尾 光輝、中村 保一	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内（横浜）
「DDBJ Read Archive」：次世代シーケンサからの出力データのためのアーカイブ[ポスター]	児玉 悠一、猿橋 智、五條 堀 孝、舘野 義男、神沼 英里、中村 保一、菅原	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内（横浜）
次世代シーケンサーのハイスループット・データ解析システム「DDBJ Short Read Annotation Pipeline」の構築[ポスター]	神沼 英里、児玉 悠一、大久保 公策、高木 利久、中村 保一	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内（横浜）
微生物ゲノムアノテーションパイプライン MiGAP[ポスター]	大山 彰、森 宙史、黒川 颯、菅原 秀明	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内（横浜）
DBCLS Galaxy: 生物データ処理の対話的ツール組み合わせインターフェース[ポスター]	山口 敦子、全 弘宇、中尾 光輝、大内田 賢太、山本 泰智、高木 利久	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内（横浜）
生命科学研究者のための統合文献情報管理システム[ポスター]	岩崎 渉、山本 泰智、高木 利久	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内（横浜）
Representation of the multiple roles of proteins from anatomical perspective by data analysis[ポスター]	坂東 明日佳、中村 桂子、大久保 公策	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内（横浜）
キーワードと蛋白 ID を用いた横断検索システム[ポスター]	畠中 秀樹、平井 信一、大野 忠、高橋 順子、牧口 大旭、奥村 利幸、熊谷 禎洋、八塚 茂、川本 祥子、西川 哲夫、中村 春木、高木 利久	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内（横浜）
生命情報データベースの横断検索システム: Web-API サービスの開発[ポスター]	牧口 大旭、高橋 順子、川崎 恭実子、川本 祥子、高木 利久	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内（横浜）

データバンク (INSD, GEO) 目次システムの開発 [ポスター]	小笠原 理、渡邊 康司、栗原 英輔、森山 丈士、大久保 公策	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内 (横浜)
BodyParts3D と Anatomography: 3 次元人体構造データベースと人体解剖図作成サービス [ポスター]	三橋 信孝、藤枝 香、田村 卓郎、川本 祥子、高木 利久、大久保 公策	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内 (横浜)
生命科学系データベースアーカイブサービス [ポスター]	西川 哲夫、八塚 茂、熊谷 禎洋、畠中 秀樹、片山 俊明、川本 祥子、大久保 公策、高木 利久	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内 (横浜)
ライフサイエンス統合データベース: アップデート 2009 [ポスター]	河野 信、Lsdb 開発スタッフ、大久保 公策、高木 利久	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内 (横浜)
かずさアノテーションを用いた分散型ゲノムアノテーションの実証実験 [口頭]	岡本 忍	第 32 回日本分子生物学会年会	2009 年 12 月 9-12 日	国内 (横浜)
Functional analysis pipeline of transcript sequences as gene expression data [口頭]	Hidemasa Bono	第 7 回がんとハイポキシア研究会	2009 年 12 月 5-6 日	国内 (京都)
シアノバクテリアのゲノム情報基盤 [口頭]	岡本 忍	ラン藻ワークショップ	2009 年 12 月 4-5 日	国内 (千葉)
GuideLink: A Corpus Annotation System that Integrates the Management of Annotation Guidelines [ポスター]	Kenta Oouchida, Jin-Dong Kim, Toshihisa Takagi and Jun'ichi Tsujii	The 23rd Pacific Asia Conference on Language, Information and Computation	2009 年 12 月 3-5 日	国内 (香港)
デジタル科学における独占と共有の賢いバランス [口頭]	大久保 公策	産総研 生命情報工学研究センター2009 年度研究報告シンポジウム	2009 年 12 月 3-4 日	国内 (お台場)
統合データベースプロジェクトのサービス紹介 - 複数のデータベースを繋ぐツールたち - [口頭]	山口敦子、中尾 光輝	産総研 生命情報工学研究センター2009 年度研究報告シンポジウム	2009 年 12 月 3-4 日	国内 (お台場)
ライフサイエンス統合データベースの課題: 権利と法律、技術 [口頭]	中尾 光輝	産総研 生命情報工学研究センター2009 年度研究報告シンポジウム	2009 年 12 月 3-4 日	国内 (お台場)

「生命科学情報の共有を目指す”ライフサイエンス統合データベース”」[ポスター]	箕輪 真理、高祖 歩美、坂東 明日佳、川本 祥子	情報・システム研究機構シンポジウム 情報とシステム 2009	2009年11月30日	国内（品川）
Integrated Database Project for Life Science in Japan[口頭]	Toshihisa Takagi	BIOINF02009, CBI-KSBSB Joint International Conference	2009年11月4-6日	国外（釜山）
統合DBプロジェクトのサービス紹介～データベースを役立てよう～[口頭]	川本 祥子	第47回日本生物物理学会年会 ランチ ョンセミナー「データベース統合の多面的展開」	2009年10月30日-11月1日	国内（徳島）
多様なデータが広く活用されるために[口頭]	畠中 秀樹	第47回日本生物物理学会年会 ランチ ョンセミナー「データベース統合の多面的展開」	2009年10月30日-11月1日	国内（徳島）
Systematic organization of gene expression data in Japan[口頭]	Hidemasa Bono, Eli Kaminuma, Yuichi Kodama, Yasukazu Nakamura, Kousaku Okubo, and Toshihisa Takagi	12th International MGED Meeting, focusing on translational genomics and high throughput sequencing	2009年10月5-8日	国外(Phoenix, Arizona)
「デジタル科学への最後のステップ」[口頭]	大久保公策	シンポジウム「科学技術における情報の 上手な権利化と共有化」	2009年10月5日	国内（東大弥生）
ライフサイエンス統合データベースプロジェクト[口頭]	高木 利久	日本人類遺伝学会 第54回大会 ワーク ショップ「ゲノムインフマティクス」	2009年9月26日	国内（品川）
ライフサイエンス分野の統合データベースの学部や高校教育における活用[口頭]	坊農 秀雅	日本遺伝学会第81回大会 ミニ・シンポ ジウム「自分が興味を持つ遺伝子がどの 範囲の環境微生物に存在するのかを探索 する方法：環境微生物ゲノム配列から のお宝遺伝子発掘の学部や高校教育に おける活用」	2009年9月16日	国内（長野）
RizoBaseの新機能 [口頭]	岡本 忍	植物微生物研究会	2009年9月8-10日	国内（長野）

統合データベースプロジェクト: オープンで知のめぐりの良いライフサイエンス研究を目指して [口頭]	坊農秀雅	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009 「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内 (東大本郷)
BodyParts3D/Anatomography: 3次元人体構造 DBと人体モデル図作成ツール [ポスター]	三橋 信孝	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009 「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内 (東大本郷)
Gene Indexing: 植物と関連微生物におけるオープンゲノム注釈実証実験 [ポスター]	岡本 忍	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009 「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内 (東大本郷)
生物種ごとの日本の分子生物学研究一覧の構築に向けて ~ 「生物アイコン」の現状と今後 [ポスター]	坂東 明日佳	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009 「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内 (東大本郷)
生命科学系データベースアーカイブサービス [ポスター]	西川 哲夫	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009 「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内 (東大本郷)
DBCLS Galaxy: 生物データ処理の対話的ツール組み合わせインターフェース [ポスター]	山口 敦子	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009 「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内 (東大本郷)
発現統合: 配列データとしての発現データを扱う解析パイプラインの開発 [ポスター]	坊農 秀雅	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009 「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内 (東大本郷)

BioHackathon 2009 レポート[ポスター]	中尾 光輝	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内（東大本郷）
MeSH キーワード集による疾患の特徴プロファイリング[ポスター]	仲里 猛留	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内（東大本郷）
テキスト処理して情報整理[ポスター]	山本 泰智	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内（東大本郷）
統合データベース講習会 AJACS[ポスター]	河野 信	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内（東大本郷）
統合TV-動画によるバイオデータベースとウェブツールの活用法解説サイト[ポスター]	小野 浩雅	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内（東大本郷）
MeSH キーワード集による疾患の特徴プロファイリング[ポスター]	仲里猛留、坊農秀雅、松田秀雄、高木利久	文部科学省委託研究開発事業 統合データベースプロジェクト シンポジウム 2009「データベースが拓くこれからのライフサイエンス」	2009年6月12日	国内（東大本郷）

統合データベースポータル「統合ホームページ」 lifesciencedb.jp[ポスター]	仲里 猛留、坂東 明日佳、 坊農 秀雅、全 弘宇、藤枝 香、畠中 秀樹、平井 信一、 川本 祥子、河野 信、高祖 歩美、熊谷 禎洋、箕輪 真 理、三橋 信孝、永井 啓一、 中尾 光輝、西川 哲夫、岡 本 忍、小野 浩雅、大内田 賢太、谷口 仁、山口 敦子、 山本 泰智、八塚 茂、大久 保 公策、高木 利久	第9回日本分子生物学会春季シンポジ ウム	2009年5月12-13日	国内(宮崎)
平成20年度				
Gene Indexing: Manual micro-annotation of plant-related organisms using KazusaAnnotation [ポスター]	岡本忍、中尾光輝	第3回国際バイオキュレーション会議 Biocuration 2009	2ヶ00江9年弘4美月, 1 笠6-井19	国外(Berlin, Germany)
Semantic clustering of biomedical abbreviation definitions using bipartite graph [ポスター]	Atsuko Yamaguchi, Yasunori Yamamoto, Hong-Woo Chun and Toshihisa Takagi	The 14th Pacific Symposium on Biocomputing (PSB)	2009年1月5-9日	国外(Hawaii)
Functional profiling of OMIM data using MeSH vocabulary [口頭及びポスター]	Takeru Nakazato, Hidemasa Bono, Hideo Matsuda, Toshihisa Takagi	2008年日本バイオインフォマティクス 学会年会(JSBi2008)	2008年12月15-16日	国内(大阪)
統合データベースプロジェクトのサービスにつ いて [口頭]	川本祥子	第31回日本分子生物学会年会 第81 回日本生化学会大会 合同大会 (BMB2008) シンポジウム「ライフサイエンス統合デ ータベースプロジェクト」	2008年12月9-12日	国内(神戸)

日本のライフサイエンス統合データベース構築・維持のための人材育成活動 [口頭]	坊農秀雅	第31回日本分子生物学会年会 第81回日本生化学会大会 合同大会 (BMB2008) シンポジウム「ライフサイエンス統合データベースプロジェクト」	2008年12月9-12日	国内(神戸)
科学データは誰のものか? [口頭]	大久保公策	第31回日本分子生物学会年会 第81回日本生化学会大会 合同大会 (BMB2008) シンポジウム「ライフサイエンス統合データベースプロジェクト」	2008年12月9-12日	国内(神戸)
KazusaWiki: ゲノムデータベースにおけるのWikiの活用 [ポスター]	藤澤 貴智、岡本 忍、中尾光輝、中村 保一	第31回日本分子生物学会年会 第81回日本生化学会大会 合同大会 (BMB2008)	2008年12月9-12日	国内(神戸)
Gene Indexing:KazusaAnnotationによる遺伝子注釈実証実験 [ポスター]	岡本 忍、中尾 光輝、藤澤 貴智、中村 保一	第31回日本分子生物学会年会 第81回日本生化学会大会 合同大会 (BMB2008)	2008年12月9-12日	国内(神戸)
かずさアノテーション: 持続可能な遺伝子アノテーション整備のための新型システム [ポスター]	中尾 光輝、岡本 忍、藤澤 貴智、中村 保一	第31回日本分子生物学会年会 第81回日本生化学会大会 合同大会 (BMB2008)	2008年12月9-12日	国内(神戸)
ライフサイエンス統合データベースポータル lifesciencedb.jp [ポスター]	河野信、坊農秀雅、全弘宇、藤枝香、川本祥子、高祖歩美、箕輪真理、三橋信孝、永井啓一、仲里猛留、西川哲夫、大野忠、小野浩雅、山口敦子、山本泰智、八塚	第31回日本分子生物学会年会 第81回日本生化学会大会 合同大会 (BMB2008)	2008年12月9-12日	国内(神戸)

	茂、大久保公策、高木利久			
MeSH を用いた生物学的解釈によるマイクロアレイデータの解析の実際 [ポスター]	仲里猛留、坊農秀雅、松田秀雄、高木利久	第 31 回日本分子生物学会年会 第 81 回日本生化学会大会 合同大会 (BMB2008)	2008 年 12 月 9-12 日	国内(神戸)
統合 TV - 動画によるバイオデータベースとウェブツールの活用法解説サイト [ポスター]	小野 浩雅、河野 信、大村 蓉子、山口 瑤子、竹若 浩一、川本 祥子、高木 利久、坊農 秀雅	第 31 回日本分子生物学会年会 第 81 回日本生化学会大会 合同大会 (BMB2008)	2008 年 12 月 9-12 日	国内(神戸)
BodyParts3D: 解剖学概念を表現した 3 次元人体構造データベース [ポスター]	三橋 信孝、藤枝 香、田村 卓郎、川本 祥子、高木 利久、大久保 公策	第 31 回日本分子生物学会年会 第 81 回日本生化学会大会 合同大会 (BMB2008)	2008 年 12 月 9-12 日	国内(神戸)
DDBJ, PDBj, KEGG, CBRC のウェブサービスによる連携 [口頭]	片山俊明	第 46 回日本生物物理学会年会 シンポジウム「動き出したライフサイエンス統合データベース」	2008 年 12 月 3-5 日	国内(福岡)
統合データベースプロジェクトは何の役に立つのか? [口頭]	高木利久	第 46 回日本生物物理学会年会 シンポジウム「動き出したライフサイエンス統合データベース」	2008 年 12 月 3-5 日	国内(福岡)
統合 DB プロジェクトのサービスについて [口頭]	川本祥子	第 46 回日本生物物理学会年会 シンポジウム「動き出したライフサイエンス統合データベース」	2008 年 12 月 3-5 日	国内(福岡)
A Literature Management System with A Recommendation Function at Database Center for Life Science. [ポスター]	Yaunori Yamamoto and Toshihisa Takagi	The 19th International Conference on Genome Informatics	2008 年 12 月 1-3 日	国外 (Gold Coast)
Life Science Database Portal Site: lifesciencedb.jp. [ポスター]	Shin Kawano, Hidemasa Bono, Shoko Kawamoto, Kousaku Okubo and Toshihisa Takagi	The 19th International Conference on Genome Informatics	2008 年 12 月 1-3 日	国外 (Gold Coast)
Functional organization of transcript sequences as gene expression data. [ポスター]	Hidemasa Bono, Shoko Kawamoto, Kousaku Okubo and Toshihisa Takagi	The 19th International Conference on Genome Informatics	2008 年 12 月 1-3 日	国外 (Gold Coast)

低酸素刺激における比較トランスクリプトーム解析 [口頭及びポスター発表]	坊農秀雅	第6回がんとハイポキシア研究会	2008年11月29-30日	国内(広島)
生命科学データの共有に向けた「統合データベースプロジェクト」の試み [ポスター]	箕輪真理、坊農秀雅、全弘宇、藤枝香、川本祥子、河野信、高祖歩美、熊谷禎洋、三橋信孝、永井啓一、仲里猛留、西川哲夫、大村蓉子、大野忠、小野浩雅、山口敦子、山口瑤子、山本泰智、八塚茂、横山貴央、大久保公策、高木利久	情報とシステム2008 「社会のイノベーションを誘発する情報システム08	2008年11月5日	国内(東京)
Text-related Services for Researchers in Life Sciences [ポスター]	Yunori Yamamoto and Toshihisa Takagi	BIBM2008	2008年11月3-5日	国外 (Phila. USA)
データの保護と利用のための情報セキュリティ [口頭]	箕輪真理	科学研究費補助金特定領域研究「ゲノム」4領域シンポジウム「ゲノム医学のこれから」	2008年10月14日	国内(東京)
Practical Organization of Sequence Data as Gene Expression Data [ポスター]	Hidemasa Bono, Shoko Kawamoto, Kousaku Okubo and Toshihisa Takagi	The 8th Cold Spring Harbor Laboratory/Wellcome Trust conference on Genome Informatics	2008年9月10-14日	国外 (Hinxton UK)
Japan started data-sharing center for publicly-funded biomedical science [ポスター]	Kousaku Okubo, Shoko Kawamoto, Hidemasa Bono and Toshihisa Takagi	The 8th Cold Spring Harbor Laboratory/Wellcome Trust conference on Genome Informatics	2008年9月10-14日	国外 (Hinxton UK)
ライフサイエンス統合データベースセンターにおける言語処理関係の活動 [口頭]	山本 泰智、大久保 公策、川本 祥子、全 弘宇、高木 利久	第38回 人工知能学会分子生物情報研究会	2008年8月28日	国内(東京)
Kazusa Annotation Suite: 植物関連生物ゲノム情報統合の試み [口頭]	岡本 忍、中尾 光輝、藤澤 貴智、中村 保一	日本進化学会第10回大会	2008年8月22-24日	国内(東京)

統合TV：動画によるデータベースの使い方講座 [口頭]	河野 信	日本進化学会第10回大会	2008年8月22-24日	国内(東京)
統合データベースプロジェクトとライフサイエンス統合データベースセンター [口頭]	坊農 秀雅	日本進化学会第10回大会	2008年8月22-24日	国内(東京)
ライフサイエンス統合データベースプロジェクト [口頭]	坊農秀雅、全弘宇、藤枝香、川本祥子、河野信、高祖歩美、峰崎雄一、箕輪真理、三橋信孝、永井啓一、仲里猛留、西川哲夫、大野忠、小野浩雅、山口敦子、山本泰智、八塚茂、大久保公策、高木利久	第28回日本糖質学会年会	2008年8月18-20日	国内(つくば)
統合データベースセンターの紹介 [ポスター]	河野 信、坊農 秀雅、全 弘宇、藤枝 香、川本 祥子、高祖 歩美、峰崎 雄一、箕輪 真理、三橋 信孝、永井 啓一、仲里 猛留、西川 哲夫、大野 忠、小野 浩雅、山口 敦子、山本 泰智、八塚 茂、大久保 公策、高木 利久	第8回日本分子生物学会春季シンポジウム	2008年5月26-27日	国内(札幌)
平成19年度				
統合データベースセンターの紹介	河野 信	第37回人工知能学会分子生物情報研究会(SIG-MBI)・第9回オープンバイオ研究会	2008年3月7-8日	国内(能美)
ライフサイエンス統合データベースプロジェクト	高木利久	「生命をはかる」研究会	2008年2月18日	国内

TogoTV - a broadcast station of tutorial movies about bioinformatics resources	S. Kawano, H. Ono, H. Bono, S. Kawamoto, T. Takagi	日本バイオインフォマティクス学会年会 (JSBi2007)	2007年12月17-19日	国内 (東京)
MeSH terms を用いた生物学的機能付与による遺伝子群の解析手法の開発	仲里猛留、滝中徹、坊農秀雅、麻生川稔、松田秀雄	第30回日本分子生物学会年会第80回日本生化学会大会 合同大会 (BMB2007)	2007年12月11-15日	国内 (横浜)
ライフサイエンス統合データベースプロジェクトの課題と成果	川本祥子、坊農秀雅、池田佳代子、河野信、三橋信孝、峰崎雄一、永井啓一、西川哲夫、小野 浩雅、山口敦子、山本泰智、大久保公策、高木利久	第30回日本分子生物学会年会第80回日本生化学会大会 合同大会 (BMB2007)	2007年12月11-15日	国内 (横浜)
ライフサイエンス統合データベースセンターにおけるデータベースポータル構築	河野信、小野浩雅、仲里猛留、坊農秀雅、岡本忍、中村保一、川本祥子、大久保公策、高木利久	第30回日本分子生物学会年会第80回日本生化学会大会 合同大会 (BMB2007)	2007年12月11-15日	国内 (横浜)
次元人体モデルのセグメンテーションによる解剖学用語の形式表現	三橋信孝、藤枝香、川本祥子、高木利久、大久保公策	第30回日本分子生物学会年会第80回日本生化学会大会 合同大会 (BMB2007)	2007年12月11-15日	国内 (横浜)
遺伝子発現統合データベースの開発	有川浩司、飯塚高康、渡邊康司、小笠原理、大久保公策	第30回日本分子生物学会年会第80回日本生化学会大会 合同大会 (BMB2007)	2007年12月11-15日	国内 (横浜)
分子生物学の研究動向の俯瞰を目的としたデータベース (INSDC, GEO) 目次の開発	小笠原理、飯塚高康、渡邊康司、有川浩司、大久保公策	第30回日本分子生物学会年会第80回日本生化学会大会 合同大会 (BMB2007)	2007年12月11-15日	国内 (横浜)
ゲノムセントラルな遺伝子発現情報融通システム YUZ (柚子)	坊農秀雅	第30回日本分子生物学会年会第80回日本生化学会大会 合同大会 (BMB2007)	2007年12月11-15日	国内 (横浜)
ライフサイエンス統合データベース基盤整備のためのデータベースポータル構築: WIINGpro	黒田 雅子, 小池 俊行, 安田 絵美, 佐藤 早苗, 藤田 晶子, 河村 明子, 川越 千晴, 坊農 秀雅, 川島 秀	第30回日本分子生物学会年会第80回日本生化学会大会 合同大会 (BMB2007)	2007年12月11-15日	国内 (横浜)

	一, 谷口 文晃、川本 祥子、高木利久			
統合 TV によるがんとハイポキシア研究の推進	坊農秀雅	第 5 回 がんとハイポキシア研究会	2007 年 12 月 1 日	国内 (千葉)
YUZ: an environment for integration AND ANALYSIS of gene regulatory networks	H. Bono, S. Kawano, S. Kawamoto, and T. Takagi	GENOME INFORMATICS	2007 年 11 月 1-5	国外 (New York (USA))
An approach to decipher gene regulatory networks from the federation of databases in life science	H. Bono, S. Kawano, S. Kawamoto and T. Takagi	FUNCTIONAL GENOMICS & SYSTEMS BIOLOGY	2007 年 10 月 10-13 日	国外 (Hinxton (UK))
ゲノム研究のためのオントロジーとテキストマイニング [口頭、招待講演]	高木利久	第 16 回セマンティックウェブとオントロジー研究会	2007 年 7 月 23 日	国内
平成 18 年度				
使い倒し系バイオインフォマティクスによる知のめぐりのよい生物学研究のすすめ	坊農秀雅	お茶の水女子大学「魅力ある大学院教育」第 6 回バイオインフォマティクスへの招待	2007 年 3 月 16 日	国内
統合 DB の構築に必要な情報技術	高木 利久	情報とシステム 2007	2007 年 3 月 1 日	国内
ライフサイエンスのデータベースの現状と課題	大久保公策	情報とシステム 2007	2007 年 3 月 1 日	国内
データベースから見たライフサイエンスプロジェクト	高木 利久	日本分子生物学会 2006 フォーラムシンポジウム[プロジェクト型研究時代の生命科学の課題]	2006 年 12 月 8 日	国内
知的生産性向上のための情報処理	大久保 公策	日本分子生物学会 2006 フォーラムシンポジウム[プロジェクト型研究時代の生命科学の課題]	2006 年 12 月 8 日	国内
ライフサイエンス DB その歴史とわが国の現状と課題	高木 利久	日本分子生物学会 2006 フォーラム バイオテクノロジーセミナー	2006 年 12 月 8 日	国内
オントロジーや辞書は役に立つのか	大久保 公策	日本分子生物学会 2006 フォーラムバイオテクノロジーセミナー	2006 年 12 月 8 日	国内

ライフサイエンスにおけるゲノム情報の高度利用に向けた生命知識の構造化	大久保 公策	知の構造化ワークショップー知の構造化ツールは、新しいサイエンスを開くのかー	2006年12月4日	国内
------------------------------------	--------	---------------------------------------	------------	----

2. 学会誌・雑誌等における論文掲載

掲載した論文（発表題目）	発表者氏名	発表した場所（学会誌・雑誌等名）	発表した時期	国内・外の別
平成 22 年度				
改変 OK? 許可は必要? 何を明示すればいい? いまさら聞けないクリエイティブコモンズ	永井 幸輔、中尾 光輝、平田 雄一、渡辺 智暁	技術評論社 59 巻、52-60	2010 年 10 月 23 日	国内
BioRuby: bioinformatics software for the Ruby programming language.	Naohisa Goto, Pjotr Prins, Mitsuteru Nakao, Raoul Bonnal, Jan Aerts and Toshiaki Katayama	BIOINFORMATICS APPLICATIONS NOTE Vol. 26 no. 20 2010, pages 2617-2619	published: 2010 年 8 月 25 日	国外
The DBCLS BioHackathon: standardization and interoperability for bioinformatics web services and workflows.	Toshiaki Katayama, Kazuharu Arakawa, Mitsuteru Nakao et al.	Journal of Biomedical Semantics 2010, Vol.1, 1-8	published: 2010 年 8 月 21 日	国外
文部科学省「統合データベースプロジェクト」と PubMed を中心とした関連データベース	仲里 猛留、坊農 秀雅	情報の科学と技術, 60 巻 7 号, 265-271	2010 年 7 月	国内
「セマンティック Web プログラミング」Toby Segaran, Colin Evans, Jamie Taylor 著	大向 一輝、加藤 文彦、中尾 光輝、山本 泰智 監訳、玉川 竜司 訳	オライリージャパン	2010 年 6 月	国内

TogoWS: integrated SOAP and REST APIs for interoperable bioinformatics Web services	Toshiaki Katayama, Mitsuteru Nakao, and Toshihisa Takagi	Nucl. Acids Res. (2010) 38 (suppl 2): W706-W711.	published online: 2010年5月14日	国外
平成 21 年度				
GuideLink: A Corpus Annotation System that Integrates the Management of Annotation Guidelines	Kenta Oouchida, Jin-Dong Kim, Toshihisa Takagi and Jun'ichi Tsujii	Proceedings of the 23rd Pacific Asia Conference on Language, Information and Computation	2009年12月3-5日	国外
CyanoBase: the cyanobacteria genome database update 2010	Mitsuteru Nakao, Shinobu Okamoto, Mitsuyo Kohara, Tsunakazu Fujishiro, Takatomo Fujisawa, Shusei Sato, Satoshi Tabata, Takakazu Kaneko and Yasukazu Nakamura	Nucleic Acids Research, 2010, Vol. 38, Database issue D379-D381	published online on 2009年10月30日	国外 (Oxford University Press)
生命科学分野におけるテキストマイニング	山本 泰智	情報処理学会誌 2009, Vol. 50, No. 9, 853-858	2009年9月	国内
生命情報学が直面する大規模ゲノムデータ時代の課題 : 1. 分散データの統合とセマンティック Web	中尾 光輝、片山 俊明	情報処理学会誌 2009, Vol. 50, No. 9, 836-844	2009年9月	国内
生命情報学が直面する大規模ゲノムデータ時代の課題 : 0. 編集にあたって	片山 俊明	情報処理学会誌 2009, Vol. 50, No. 9, 833-835	2009年9月	国内
Gendoo: Functional profiling of gene and disease features using MeSH vocabulary	Takeru Nakazato, Hidemasa Bono, Hideo Matsuda and Toshihisa Takagi	Nucleic Acids Research, 2009, Vol. 37, Web Server issue W166-W169	2009年6月	国外 (Oxford University Press)
ライフサイエンス統合データベースセンターと統合データベースプロジェクト	坊農 秀雅	情報の科学と技術, 59巻4号, 165-170	2009年4月	国内
平成 20 年度				

BodyParts3D: 3D structure database for anatomical concepts	Nobutaka Mitsuhashi, Kaori Fujieda, Takuro Tamura, Shoko Kawamoto, Toshihisa Takagi and Kousaku Okubo	Nucleic Acids Research, 2009, Vol. 37, Database issue D782-D785	2008年10月3日	国外
蛋白質核酸酵素・連載 ライフサイエンス分野の統合データベース 「『蛋白質 核酸 酵素』バックナンバーの全文検索サービス公開にあたって」	川本祥子	蛋白質核酸酵素 Vol. 53, No. 9, pp. 1200-1205	2008年7月	国内
ライフサイエンス分野のデータベースの現状と統合データベースプロジェクト	坊農秀雅、高木利久	日本化学会情報化部会誌 特集 Vol. 26, No. 1, pp. 5-7	2008年4月	国内
平成 19 年度				
利用の立場からのコメント	高木利久	蛋白質核酸酵素(共立出版) 53巻、5号、686-691	2008年	国内
統合データベースプロジェクトのサービスとその利用法	川本祥子、坊農秀雅	蛋白質核酸酵素(共立出版) 53巻、3号 281-287、	2008年	国内
統合データベースセンターがめざすもの	高木利久	蛋白質核酸酵素(共立出版) 52巻、11号、1388-1393	2007年	国内
生命科学系データベース統合化の背景	大久保公策	蛋白質核酸酵素(共立出版) 52巻、9号、1027-1031	2007年	国内
OReFiL: an online resource finder for life sciences	Y. Yamamoto and T. Takagi	BMC Bioinformatics, 2007, 8, 287	2007年	国外
BioCompass: A novel functional inference tool that utilizes MeSH hierarchy to analyze groups of genes	T. Nakazato, T. Takinaka, H. Mizuguchi, H. Matsuda, H. Bono and M. Asogawa	<i>In Silico</i> Biology, 2007, 8, 6	2007年	国外
平成 18 年度				

生命科学データベースの現状と課題	大久保 公策	科学 77 巻、4 号、364-369	2007 年	国内
------------------	--------	---------------------	--------	----