

平成 25 年度産業技術研究開発

「ライフサイエンスデータベースプロジェクト」
に関する成果報告書

平成 26 年 3 月

(委託先)独立行政法人 産業技術総合研究所

目次

1.	総括	3
2.	研究開発成果	4
2. 1	ポータル構築連携	4
2. 1. 1	経済産業省ライフサイエンスポータルサイトの拡張作業および運営	4
2. 1. 2	統合データベースセンターとの連携	20
2. 2	横断検索連携	22
2. 3	アーカイブ構築連携	25
2. 4	国内外の最新動向調査及び普及に向けた取り組み	27
2. 4. 1	国内外の最新動向調査	27
2. 4. 2	広報・普及活動	30
2. 4. 3	ユーザーニーズ調査	33
2. 5	運営委員会	37
2. 6	成果発表	37
3.	謝辞	41
4.	参加者名簿	42

1. 総括

平成 25 年度経済産業省ライフサイエンスデータベースプロジェクトは、経済産業省関連機関により実施されたライフサイエンス分野の研究開発プロジェクトの成果等に関するポータル構築連携、横断検索連携、アーカイブ構築連携を行うことを目的とし、日本全体のライフ分野の統合データベースセンターとしての役割が期待される科学技術振興機構(JST)バイオサイエンスデータベースセンター(NBDC)との緊密な連携のもとで、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で実施された。

今年度の主な成果は、まず、前年度までの経産省ライフサイエンスデータベースプロジェクトによって構築されたポータルサイト MEDALS(<http://medals.jp/>)を継続的に改良しつつ運営し、データベースやソフトウェア等の成果物に関する情報発信を行った。ポータルサイトの内容は隨時更新し、新しいデータベース等の紹介も行った。特に、経産省関連の各種研究開発プロジェクトで作成されたデータベースやソフトウェアについては、詳細な調査を実施した結果、合計件の成果物を「便覧」に掲載した。また、多数のデータベースを一括検索できる「MEDALS 横断検索」については対象データベースをさらに充実させたほか、データ ID による検索との連携を実現した。データベースのアーカイブ作成も、NBDC との連携のもとで実施した。広報活動では、データベース講習会の開催やアンケートによるニーズ調査のほか、幅広い利用者への情報提供を実現するために新しい情報発信の手段として Facebook や Twitter の利用を継続し、また USTstream による講習会ビデオ配信を行った。

以上のように、本プロジェクトはライフ分野のデータベース統合化において多くの成果を挙げ、有用な情報を社会に発信することができた。今後も利用者となる産業界等の研究者の意見を反映させつつ、本事業の成果を広く産業界や社会へ提供することをめざして、本プロジェクトを継続的に実施していくことがきわめて重要であると考える。

平成 25 年度経済産業省ライフサイエンスデータベースプロジェクト・リーダー 今西 規

2. 研究開発成果

● 背景

ライフサイエンス分野では、自身の研究成果を既に蓄積されている研究成果や研究データと対比することにより、自身の研究成果の仮説を考案する手がかりが得られたり、効果的な治療薬など新しい実用化の発想が得られたりする可能性がある。このため、国家プロジェクト等により产生された研究データを一括して活用できるデータベースが、産業界や社会から要望されている。

このため、総合科学技術会議では、上記の問題を解決するため、ライフサイエンス分野の研究成果をオールジャパン体制で整備するための報告書及びロードマップを策定している。(総合科学技術会議専門調査会分野別推進戦略総合PTライフサイエンス PT 第 14 回配布資料 : <http://www8.cao.go.jp/cstp/project/bunyabetu2006/life/14kai/haihu14.html>) これを受け、経済産業省では平成 23 年度に、ライフサイエンスデータベースプロジェクトを実施し、経済産業省関連機関により実施された研究開発プロジェクトの成果等を整備したポータルサイトの構築等を実施してきたところである。

● 目的

本事業では、独立行政法人科学技術振興機構(JST)に設置された統合データベースセンターを中心とした政府全体の統合データベースを構築するため、本事業では関係府省の分担業務であるポータル構築連携、横断検索連携、アーカイブ構築連携、を行うことを目的とする。

● 概要

本事業では、平成 24 年度において実施した「産業技術研究開発(ライフサイエンスデータベースプロジェクト)」の成果を活用し、統合データベースセンターと連携して政府全体の統合データベースを構築するため、「1. ポータル(サイト)構築連携」、「2. 横断検索連携」、「3. アーカイブ構築連携」、「4. 国内外の最新動向調査及び普及に向けた取り組み」、及び「5. 運営委員会」について実施した。以上を実施するにあたり、文部科学省、農林水産省等、他省庁のデータベース関連プロジェクトと連携し、さらに産業界のニーズを取り込むことに積極的に取り組んだ。

2. 1 ポータル構築連携

2. 1. 1 経済産業省ライフサイエンスポータルサイトの拡張作業および運営

2.1.1-(1) 公開可能な成果物の調査と便覧への掲載

① 成果物の調査

平成 20 年度から平成 22 年度に実施された「産業技術研究開発(統合データベースプロジェクト)」において経済産業省関連のライフサイエンス分野のプロジェクトによるデータベースが調査された。その結果、データベース等の成果物候補が存在するプロジェクト数は 59 件(当時進行中のプロジェクトを含む)であり、これによる成果物候補は 217 件あることが明らかになっていた。また、

これらの成果物のうち、公開されたものについては、2013年3月31日時点で、155件がポータルサイト MEDALS (<http://medals.jp/>) の成果物便覧（カタログ）に格納されていた（http://medals.jp/list/list_alldb、http://medals.jp/list/list_alltool）。残りの62件は便覧未掲載で、うち53件が非公開データ、3件が公開可否未確認、6件が公開可能（公開予定を含む）という状態であった。

これらの成果物の調査からポータルサイト MEDALS の成果物便覧に載るまでの流れは図2.1-1の形で行われていた。まず成果物便覧に掲載してあるもの以外で他の成果物がないか、各プロジェクトの成果報告書をもとに、それらの存在・公開状態の調査を行う。主な候補は、新エネルギー・産業技術総合開発機構(NEDO)の未格納分、産業技術総合研究所(AIST)、および製品評価技術基盤機構(NITE)によるライフサイエンス系成果物である。経済産業省の関連機関やWeb上の情報を元に、ライフサイエンスのデータベース（ツールも含む）として適切と判断したものを「初期候補」とする。候補からコンタクト先を調査し、各プロジェクトのリーダーや担当の方にコンタクトをとて、どのような具体的な成果物（データベース）があったかを問い合わせる。回答によって、「保留」、「非公開」、または「公開可能（公開している、公開予定である）」となる。「保留」の例としては、個別の担当者にコンタクトがとれないものや、担当者が不明ですぐには出てこないものなどがある。「非公開」というのは、無償公開ではないことで、製品やサービスとして商業利用している場合はこれに含まれる。また、「公開可能」には公開予定のもの、すなわち調査中の時点ではまだ公開されていないものが含まれる。

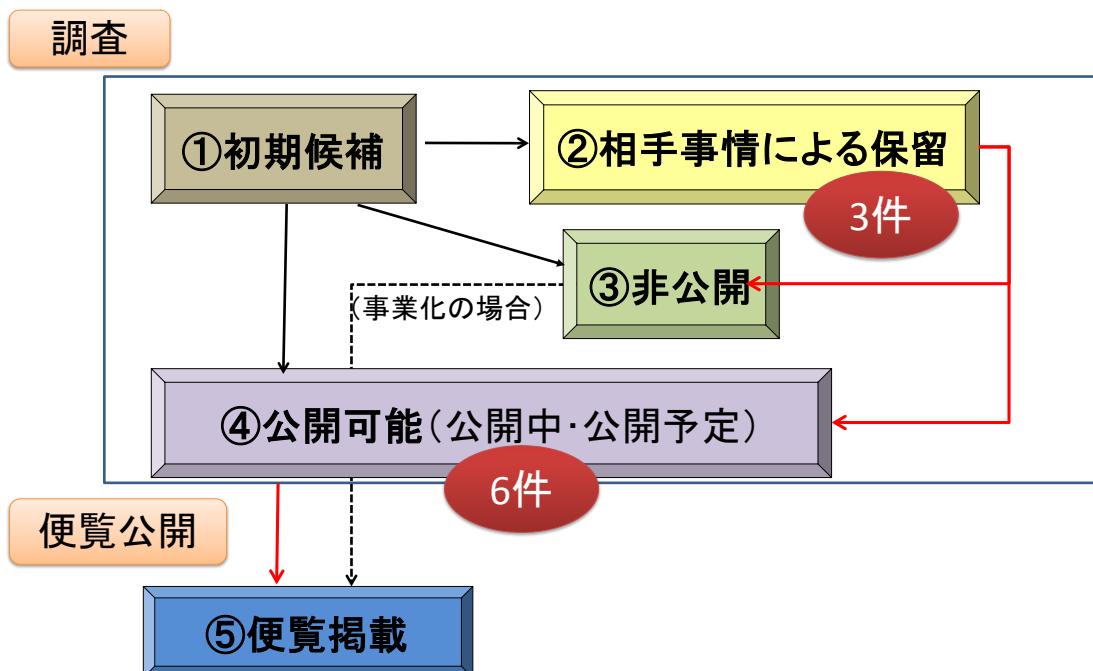


図 2. 1-1 成果物候補の選定から便覧掲載までの状態遷移

平成 25 年度の事業としては、「保留」3 件の公開可否未確認の成果物および「公開可能」6 件の成果物について、再調査を行った。方法としては「①初期候補」に対する調査方法と同様で、これらについて再度コンタクトを取り、公開か非公開かを調査した。すぐには回答が得られないことが多く、その場合でも最低 3 回はメールないし電話でコンタクトをとった。また、6 件の「公開可能」について進捗や詳細を調査し、公開されたものを便覧に載せ、各データベースの本家サイトへ便覧からハイパーリンクをする作業を行った。この作業が済んだ成果物は「便覧掲載」の状態になる。この次の段階としては、2. 2 節に記述の「横断検索」や 2. 3 節に記述の「アーカイブ」という形での活用を検討することになる。それらについては、MEDALS のプロジェクト便覧(各成果物を出したプロジェクトについて記述している)にその情報を掲載すること、およびデータをダウンロードする形態での公開(2. 3 節のアーカイブ構築連携に詳述)を検討した。

「保留」3 件の公開可否未確認の対象成果物を表 2. 1-1 に示す。この 3 件は前年度の調査によれば、3 つとも 1998 年度から 2002 年度まで行われたプロジェクト「ゲノムインフォマティクスプロジェクト」で、「成果物」は 7 件あるが、4 つは便覧掲載済み(ただし 2011 年度に PJ 便覧へ掲載)で、残りの 3 つに該当する。この 3 件について 3 回メールと電話で問合せをし、ほぼ最終的回答が得られた。検討の結果、いずれの成果物もプロジェクト内で特殊用途に向けたツールであるため公開しないということである。7 件のいずれもプロジェクト便覧に掲載している
(http://medals.jp/project_list/detail/69)。

表 2. 1-1 「保留」の公開可否未確認の対象成果物(3 件)の結果

	名称	プロジェクト便覧掲載	ステータス
1	発現頻度情報収集用ソフトウェア	済み	非公開(プロジェクト内ツールのため)
2	遺伝子発現プロファイル計測実験支援システム	済み	同上
3	転写因子データベース (T F D B)	済み	同上

次に、「公開可能(便覧未掲載)」6 件以外に新たに開発された成果物が見つかった。表 2. 1-2 には、当初予定していた 6 件(下部、継続性=「継続」と、今年度新たに見つかった成果物 14 件(継続性=「新規」)の合計 20 件をまとめた。このうち、19 件のデータが公開済みとなっていた。4 場、10 番は MEDALS の対象である医学生物学に関わる分子データとはそぐわないもので対象としないこととした。また、11 番と 12 番(いずれも糖鎖関連データ)は、詳細な調査の結果、産総研の資金によるものでないことが分かったので対象としないこととした。未公開の成果物は 6 番(被験者の脳 MRI 画像等)の 1 件だけとなった。

表 2. 1-2 「公開可能(便覧未掲載)」についての調査結果

番号名	継続案件便覧	概要	区分	機関	公開年	
1 fastapl	継続	公開済	ツール	産総研	2012	
2 seg-suite	継続	公開済	ツール	産総研	2012	
3 Dnemulator	継続	公開済	ツール	産総研	2012	
4 AIST-SHANEL	継続	対象としない	日本における河川流域の化学物質の暴露評価と対策評価のためのモデル解析ソフト	ツール	産総研	2003
5 GSE45650	継続	公開済	ラット毒性評価データ(高機能簡易型有害性評価プロジェクト)	DB	化学物質評価研究機構	2013
6 未定	継続	未公開	被験者の脳MRI画像等データ	DB	東京大学医学部	-
7 ARTRA	新規	公開済	ナズナのマイクロアレイデータベース	DB	かずさDNA研究所	2009
8 MiFuP	新規	公開済	ゲノム配列情報から微生物の機能を推定するデータベース	DB	化学物質評価研究機構	2014
9 ToxBay	新規	公開済	ベイジアン推定毒性推定システム	DB	慶應大学	2012
10 Risk Learning	新規	対象としない	化学物質の健康リスク評価データ	DB	産総研	2003
11 GDGDB	新規	対象としない	糖鎖関連疾患と原因遺伝子のデータベース	DB	産総研	2011
12 PACDB	新規	対象としない	糖鎖・病原体データベース	DB	産総研	2011
13 SAHG	新規	公開済	ヒトゲノムからの予測蛋白の構造・機能データ	DB	産総研	2012
14 H-EPD	新規	公開済	ヒトタンパク質データベース	DB	産総研	2012
15 TuMaRdb	新規	公開済	疾患糖鎖マーカーデータベース	DB	産総研	2011
16 GlycoPOD	新規	公開済	実験プロトコルデータベース	DB	産総研	2011
17 GlycoExplorer	新規	公開済	分子構造による検索システム	DB	産総研	2011
18 J-CHECK	新規	公開済	化学物質のリスク・暴露情報(化審法に準拠)	DB	製品評価技術基盤機構	2008
19 DAFS	新規	公開済	双対分解法によるRNA配列アラインメント	ツール	産総研	2012
20 NGSFeatGen	新規	公開済	次世代シーケンサー配列のクラスター作成	ツール	産総研	2012

以上の 2 つの作業の結果、今後も継続調査が必要な成果物の数は 0 件になった。当初保留事業だった 3 件については非公開で決着し、公開可能で便覧未掲載の 6 件については(10 番)「被験者の脳 MRI 画像等」だけが公開されずに残った。今後はこの 1 件と、今後新たに発表される成果物を調査することが期待される。

2.1.1-(2) データベース間の ID 統合:「リンク自動管理システム」の開発

産総研・創薬分子プロファイリング研究センターでは、持続可能なデータベース統合の方式を模索すべく、データベース間のリンクを自動で更新するための「リンク自動管理システム」の開発を進めている(<http://biodb.jp/>)。これは、多数の公開データベースからデータ管理用 ID を集めてその対応関係を自動で解析し、データベース間のリンクを常に最新かつ網羅的に提供するシステムである。また、データ ID の一括変換サービスや、データ ID 対応表のダウンロードなどのサービスも提供している。

① リンク自動管理システムへの対象データベースの追加

今年度追加したデータベースは以下の通りである(図 2. 1-2)。

- ・2013 年 5 月 21 日 Human Protein ATLAS(正常およびがん組織におけるタンパク質の発現データ)を追加
- ・2013 年 7 月 9 日 Peptide ATLAS(質量分析法によるタンパク質断片の測定データ)を追加
- ・2013 年 8 月 27 日 Ensembl Protein(Ensembl 版のヒトタンパク質)を追加
- ・2013 年 10 月 18 日 neXtProt(ヒトタンパク質の実験的エビデンス)を追加

H-InvDB	NCBI	Ensembl	KEGG
<ul style="list-style-type: none"> • H-Inv cluster ID (HIX) • H-Inv transcript ID(HIT) • H-Inv protein ID (HIP) • Accession Number 	<ul style="list-style-type: none"> • GenelD • Accession Number • RefSeq ID • RefSeq Protein ID • OMIM ID • PubMed ID • HUGO gene symbol 	<ul style="list-style-type: none"> • HUGO gene symbol • Ensembl Gene ID (ENSG) • Ensembl Transcript ID (ENST) • Ensembl Protein ID (ENSP) 	<ul style="list-style-type: none"> • KEGG Gene ID • KEGG Pathway ID • GenelD
HUGO	UniProt	GeMDBJ	MutationView
<ul style="list-style-type: none"> • HUGO gene symbol 	<ul style="list-style-type: none"> • UniProt Accession Number • Accession Number 	<ul style="list-style-type: none"> • dbSNP rs# • HUGO gene symbol 	<ul style="list-style-type: none"> • OMIM ID • HUGO gene symbol
PDBj	fRNAdb	HPRD	H-ANGEL
<ul style="list-style-type: none"> • PDB ID • UniProt Accession Number 	<ul style="list-style-type: none"> • fRNAdb ID • Accession Number 	<ul style="list-style-type: none"> • HPRD ID • GenelD 	<ul style="list-style-type: none"> • H-Inv cluster ID (HIX) • Accession Number
NBRC	FLJ Human cDNA	HGPD	H-GOLD
<ul style="list-style-type: none"> • FLJ ID • Clone ID • Accession Number 	<ul style="list-style-type: none"> • FLJ ID • Clone ID • Accession Number 	<ul style="list-style-type: none"> • FLJ ID • Clone ID • Accession Number 	<ul style="list-style-type: none"> • H-GOLD Marker ID • Accession Number
H-DBAS	G-Compass	CIPRO	JSNP DATABASE
<ul style="list-style-type: none"> • H-Inv cluster ID (HIX) • Accession Number 	<ul style="list-style-type: none"> • H-Inv transcript ID(HIT) • Accession Number 	<ul style="list-style-type: none"> • CIPRO ID • Accession Number 	<ul style="list-style-type: none"> • JSNP ID • Accession Number
VarySysDB	DGV	PDBeChem	DrugBank
<ul style="list-style-type: none"> • H-Inv transcript ID(HIT) • Accession Number 	<ul style="list-style-type: none"> • DGV ID • HUGO gene symbol 	<ul style="list-style-type: none"> • PDBeChem ID • PDB ID 	<ul style="list-style-type: none"> • DrugBank ID • UniProt Accession Number
Evolva	LEGENDA	Wikipedia Human proteins	GGDB
<ul style="list-style-type: none"> • H-Inv transcript ID(HIT) 	<ul style="list-style-type: none"> • GenelD 	<ul style="list-style-type: none"> • HUGO gene symbol 	<ul style="list-style-type: none"> • HUGO gene symbol
COXPRESdb	Human Protein ATLAS	Peptide Atlas	New neXt Prot
<ul style="list-style-type: none"> • GenelD 	<ul style="list-style-type: none"> • Ensembl Gene ID (ENSG) • HUGO gene symbol 	<ul style="list-style-type: none"> • Peptide Accession • UniProt Accession Number 	<ul style="list-style-type: none"> • UniProt Accession Number

図 2. 1-2 リンク自動管理システムに登録されているデータベース(Human の部分)

以上の開発により、リンク自動管理システムが対応するデータベースの種類はヒトの分子情報は 42 種類、マウスは 19 種類、ラットは 12 種類、ホヤは 12 種類、化合物は 15 種類となった。そして、取り扱うデータ ID の件数は約 1 億 6000 万件、扱うリンク数は約 8 億件となり、世界最大のデータ ID の統合データベースである。

② MEDALS 橫断検索横断検索に対する分子 ID 情報の提供

MEDALS 橫断検索のインデックスを作成する際、分子情報に関する ID を HTML データより抽出し、横断検索の結果表示の際に ID を表示する。このときにリンク自動管理システムの ID 情報を利用して Web ページにかかれたテキストが分子情報を示す ID であるかどうかを判定する。さらに、抽出した ID をメタデータ(マイクロデータ)としてインデックスに埋め込み、MEDALS 橫断検索の結果画面に表示する(図 2. 1-3)。ヒットした画面(データベース内の Web ページ)に記述された ID の一覧が表示される。ID にはオリジナルのデータベースへのリンクが張られており、ユーザはクリッ

クで該当ページに遷移することが出来る。インデックス作成時の ID 抽出には大量のデータを処理する必要から、リンク自動管理システムから公開している ID resolver Web サービスのプログラムを改良して使用する予定であった。しかし、実際に検討した結果、ID resolver プログラムへの改良を行わなくとも十分な処理能力および機能をもつと判断し、現在公開しているものをそのまま利用することとした。

The screenshot shows a search results page titled "Information Overview [hgpd]". The URL is http://hgpd.lifesciencedb.jp/cgi/pg_locus_list.cgi?txt_srch_cdna_name=FLJ20442. The results list various gene IDs and their associated systems:

- Entry ID: FLJ20442
- HUGO: DUSP23
- Hyperlink Management System: DUSP23
- NCBI GenBank: AK000449
- Hyperlink Management System: AK000449
- NBRC: FLJ20442
- Hyperlink Management System: FLJ20442
- HUGO: SLAMF8
- Hyperlink Management System: SLAMF8
- NCBI GenBank: AK074669
- Hyperlink Management System: AK074669
- NBRC: FLJ90188
- Hyperlink Management System: FLJ90188
- Ensembl Transcript: ENST00000289707
- Hyperlink Management System: ENST00000289707
- Ensembl Transcript: ENST00000289707
- Hyperlink Management System: ENST00000289707
- HUGO: C1ORF204
- Hyperlink Management System: C1ORF204
- NCBI GenBank: AK096506
- Hyperlink Management System: AK096506
- NBRC: FLJ39187
- Hyperlink Management System: FLJ39187
- HUGO: CCDC19
- Hyperlink Management System: CCDC19
- HUGO: TAGLN2
- Hyperlink Management System: TAGLN2
- HUGO: VSIG8
- Hyperlink Management System: VSIG8
- NBRC: FLJ76411
- Hyperlink Management System: FLJ76411

At the bottom right of the results table is a "Show / Hide" button.

図 2.1-3. MEDALS 横断検索に表示された分子 ID

MEDALS 横断検索において、“chr1”のキーワードで検索をし、さらに“Show/Hide”的ボタンをクリックすることにより、ヒットした画面(データベース内の Web ページ)に記述された ID の一覧をリンクつきで表示した画面。

2.1.1-(3) ポータルサイトの運営

① 便覧の更新

(a) 新規便覧の作成と公開

第 2.1.1-(1) 節の表 2.1-2 にあるように、今年度は新規に開発・発表された 15 件のデータベースおよびツールについて便覧を作成し、公開を行った(2014 年 3 月 10 日時点)。このうち 4 件は、年度初めに予定していたものをカタログ化したものである。その他の 11 件は、年度に入った後の調査によって、各研究機関から発表されていることがわかった成果物である。

(b) 便覧の定期更新と RSS による更新情報の配信

既に本便覧に掲載されているデータベース詳細情報の内容の更新、改良を定期的に行った。

更新にあたっては、我々が独自開発した外部のデータベースサイトに対してそのコンテンツ変更を検知する「更新検知ツール」を用いて毎日更新をチェックした。このツールをもちいると、MEDALS 便覧に載っているデータベースについて、サイトのページ変更された部分が検知される。それをうけて、我々の担当者が外部サイトを目視で確認し、データ(ソフトウェア)自体が変わったか、あるいは機能が大幅に追加されたかの2つをチェックした。そのどちらかが目視で認められた場合に、便覧を更新してニュース事項としてトップページに載せ、また、更新情報の広報のために、RSS (RDF Site Summary) を使って広報した。RSS とは Web サイトの要約などのメタデータを構造化して記述する XML ベースのフォーマットである。我々は、RSS データを検知ツールから自動で生成している。これを MEDALS サイトにおき、ユーザー(RSS リーダー)の要求に応答する形で配信した。ユーザーのブラウザーのような RSS リーダーでこの情報をみると初めは図 2. 1-4 のようなサマリ画面が見られる。いずれかの行をクリックすると図 2. 1-5 のように何が更新されたのかの詳細を見ることができる。

便覧の定期更新は、2013 年 4 月から 2014 年 3 月までの 12 ヶ月にわたり、週一回を原則として実施し、合計 150 件の RSS 配信をした(3 月 6 日までの実績)。この内訳としては、便覧に書かれた外部データベースサイトの更新:119 件、MEDALS サイト(便覧等)の更新:16 件、イベント等ニュース:15 件となっている(表 2. 1-3)。

MEDALS »			
☆ [外部サイト更新][成果物]FORTE - 2013 February 20 the NCBI non-redundant database: 12/02/13 (98% & filtered) 22444		2013/02/21	□
☆ [外部サイト更新][成果物]SPALN - Present Version: 2.0.6d, Last update: 2013-02-18		2013/02/21	□
☆ [外部サイト更新][成果物]FORTE - 2013 February 13 the NCBI non-redundant database: 12/02/13 (98% & filtered) 22431		2013/02/14	□
☆ [外部サイト更新][成果物]LAST - last-274.zip 12-Feb-2013 11:54 424K		2013/02/14	□
☆ [外部サイト更新][成果物]FORTE - 2013 February 06 the NCBI non-redundant database: 12/02/13 (98% & filtered) 22411		2013/02/08	□
☆ [外部サイト更新][成果物]LAST - last-270.zip 04-Feb-2013 17:07 422K		2013/02/08	□
☆ [外部サイト更新][成果物]MAFFT - The latest version is 7.023 (2013/Feb/03).		2013/02/08	□
☆ [外部サイト更新][成果物]MassBank - 2013年02月01日 常に最新の公開データをダウンロードできるようになりました。		2013/02/08	□
☆ [外部サイト更新][成果物]FORTE - 2013 January 30 the NCBI non-redundant database: 12/02/13 (98% & filtered) 22398		2013/01/31	□
☆ [外部サイト更新][成果物]KOMICS - Data update (2013/01/29) Data for biological samples were added. There are data for 75		2013/01/31	□
☆ [外部サイト更新][成果物]PubMedScan - PubMedScan v3.4 was released. [13.01.29] Fixed a bug the search function for		2013/01/31	□
☆ [外部サイト更新][成果物]リンク自動管理システム - January 24, 2013 COXPRESdb were included.		2013/01/25	□
☆ [外部サイト更新][成果物]PowerFT - PowerGet 3.3.6ベータ版を公開 LC-MSの解析ソフト、PowerGetの3.3.6ベータ版を公開しまし		2013/01/25	□
☆ [外部サイト更新][成果物]MassBank - 2013年01月23日 ワシントン州立大学のデータ(2,626 スペクトル)を公開しました。		2013/01/25	□
☆ [外部サイト更新][成果物]SPALN - Present Version: 2.0.6c, Last update: 2013-01-23		2013/01/25	□
☆ [外部サイト更新][成果物]FORTE - 2013 January 23 the NCBI non-redundant database: 12/02/13 (98% & filtered) 22380		2013/01/25	□
☆ [外部サイト更新][成果物]DoBISCUIT - 2013-01-16 Release [8] Benastatin Chartreusin Chromomycin Enterocin		2013/01/25	□
☆ [外部サイト更新][成果物]Workflow - 2013/01/17 · RNA Structure Prediction Active ワークフローにRNA相互作用予測を追加しまし		2013/01/25	□
☆ [外部サイト更新][成果物]FORTE - 2013 January 16 the NCBI non-redundant database: 12/02/13 (98% & filtered) 22371		2013/01/18	□

図 2. 1-4 RSS Reader での RSS データの表示



図 2. 1-5 RSS 情報をブラウザで表示した図

表 2. 1-3 RSS 配信による広報の内訳

内容	回数
外部サイト更新	119
成果物追加	16
お知らせ/News	15
合計	150

② セキュリティチェック

ポータルサイト MEDALS は、セキュリティに関する対策をおこないながら運用した。セキュリティに関するチェックについては、広く利用されている安定した検査ソフトウェアを利用してセキュリティ上の脆弱性を点検した。キーワード検索などユーザーからの入力を受け付ける機能には特に脆弱性が発生し易いため、1) ポータルサイト、2) ポータルサイト内検索、3) PubMedScan、4) リンク自動管理システム、5) MEDALS 横断検索、6) MEDALS RDF の6つのサービスについて点検を実施した。

Web アプリケーションの脆弱性診断ツール「OWASP ZAP」を利用して、MEDALS サイトの網羅的脆弱性チェックを行った。このツールは図 2. 1-6 にリストされる脆弱性をチェックすることができる。

■Passive

- Incomplete or no cache-control and pragma HTTPHeader set
- Content-Type header missing
- Cookie no http-only flag
- Cookie without secure flag
- Cross-domain JavaScript source file inclusion
- IE8's XSS protection filter not disabled
- Secure pages including mixed content
- Password Autocomplete in browser
- Private IP disclosure
- X-Content-Type-Options header missing
- X-Frame-Options header not set

■インジェクション

- Server side include
- Cross Site Scripting (Reflected)
- SQL Injection
- CRLF injection
- Parameter tampering

■インフォメーション

- Directory browsing
- Session ID in URL rewrite

■クライアント・ブラウザ

- Secure page browser cache

■サーバ・セキュリティ

- Path Traversal
- Remote File Inclusion

■一般

- URL Redirector Abuse

図 2. 1-6 OWASP ZAP が対象とする脆弱性項目

OWASP ZAP を用いて、各サイトを点検したところ、High レベル脆弱性が 3 件、Low レベル脆弱性が 13 件検出された。High レベル脆弱性のうち 1 件は、指摘された攻撃を試みたが再現せず、論理的にも存在しないと考えられたため、誤検出と判定した。残り 2 件についてはエスケープ処理により対応した。Low レベル脆弱性の 5 件は異なるサーバーについての同じ内容であるが、手間がかかる割に重大な問題でもないため対策しないこととした。残りの 8 件は設定をすることで対応した。以上の対策により、上記の各サイトは一定の基準をクリアしたものとなった。

また、ポータルサイトの運営については、情報セキュリティ上問題がないように運営し、アクセス情報の管理、利用状況や利用方法の分析を行った。ポータルサイトへのアクセス数の集計については、MEDALS サーバーに Google Analytics の仕組みを設定し、集計した。アクセス(ページビュー)数は、2013 年 4 月から 2014 年 2 月まで(11 ヶ月)の集計で 月平均で 1,616 件(前年度は 1,561 件)であった。合計では 17,777 件であった。

③ 利便性向上に向けたポータルサイトの改良

(a) ユーザーの意見を取り込んだ修正

便覧データをダウンロードして利用したい、という要望があったため、XML の形で取り込めるように対応した(図 2. 1-7)。

The screenshot shows a search results page for MEDALS. At the top right, there is a box containing the text "この便覧データが XML で取得できるリンク" with a red arrow pointing to it. Below this, there is a "XML output" button. To the left, there is a thumbnail image of a DNA helix and the text "詳細情報". On the right, there is a table with the following data:

成果物名	DAFS
成果物の別名	なし
成果物に関する説明	期待精度最大化 列アラインメント
成果物のタイプ	Tool

図 2. 1-7 便覧データを XML 形式でダウンロードできるリンクをつけた便覧ページ部分

(b) マイクロデータの便覧への埋め込み

マイクロデータとは、Web ページに特殊な形式で埋め込んだデータで、ブラウザには表示されない情報(テキスト)を提供するものである。それによって、意味(属性)を限定した情報を第 3 者が取得でき、高度に利活用することを想定している。そこで MEDALS では、便覧ページに生物種などのマイクロデータを埋め込んだ。その結果、連携先である医薬基盤研が運用する横断検索 Sagaceにおいて検索結果に、埋め込んだ情報が表示できるようになった(図 2. 1-8)。この図の例では、「Evola」というデータベース名で検索すると MEDALS の Evola を記述した便覧ページがヒットし、その中にかかれた関連文献の PubMed ID が独立した場所に表示されている。



図 2. 1-8 MEDALS 便覧のマイクロデータ(PubMed ID)が表示された Sagace での検索結果

(c) 収録データベースの分類

データベースを新たな分類法を開発してユーザーに提示する。現在の MEDALS にも「DNA、RNA、タンパク質、その他」という分類が用意されている。図 2. 1-9 にトップページのデータベース便覧のカテゴリを表示しているが、示されているように「その他」カテゴリに属する DB が 51 件と多数になっていて全体の把握が困難である。そこで、データの共通性・類似性に基づいてデータベース全体の地図を別途作成して、データベースの全体を俯瞰したり、データベースそれぞれの位置づけを直観的に把握しやすい表による分類図を提供することとした(<http://medals.jp/classification/>)。「クラス」と「更新日」による成果物の分類の例を図 2. 1-10 に示す。生物種やタイプなどから2つの軸をユーザが選ぶと、各データベースが表にマップされて一覧がでてくる。これによって、どういうタイプのデータベースがどの程度あるかを把握することができる。



図 2.1-9 TOP ページのデータベース便覧のカテゴリ

MEDALS Japan
MEI database portal for life science

Database classification by matrix

This page designed for [Firefox](#).

Database collection

MEDALS (bio) ▾

First condition (vertical): Second condition (horizontal)

Study date ▾ Class ▾

Show the matrix

Transpose the matrix

clear

Show a bird-eye view on top-right

[Return to Database map](#)

Please select conditions and show the matrix!

DB		Tool
2011	INDO Pathway Database TMFuctions	DasfroldAffid DasfroldFold DasfroldHomid Monitor
	ATCC BSI T database Brain Atlas Database of Japanese Monkey for WWW DCCIPPCA Dioxin Risk Information Platform DoKE DS-DBRE EAM DOBAN Dose EPIC EPIC/KB EU Human cDNA Database H-Aff93 H-OBAS H-OBES HDFC HDFC-DATABASE HMG4-Ac5 HMG4-Ac6 Family KOMIS LEADERIA MEDALL Metabon Metaboss PDB-PDBDB PPI view SEVEN The database of full-length cDNA and EST sequences of Lofthus VicUGeC	ACHER ALN CACUVE DasfroldAlm FORT GPI2G Interaction Partner H2AF Hyperlink Management System ID Converter System IPknot Mapgenetics ZFIN ZFIN-AffyView XEN99 LAST MAPFT PAPTA PAPTA PicTP Piccount SLIDESORT SPAWN statRNN myProteo FastText
2010	GenBank GpD GeneBase MS-HIT Fragment Viewer NCBI protein database search PDB UCSC GenomeBrowser for Functional RNA	A-METHOD for Extracting Optimal Sequence Related to Biological Activity KATAHANA MK3-DNA PMID-Extractor PubMedScan Raccast SARCAST
2009	B-complex GOKE GWAS GymproDB H-Egg LIOB TMbeta-OX50 The Functional RNA Sequence Database	ASTAN Dr.DNA8+ MeV Munsi Phenomen RIDE SAMURAI SGARNA Local Multiple Stem Kernels TADT
2008	DNA ProbeLocator Escherichia coli Mammalian Database Human-Gene diversity of Life-style related Diseases (H-GOLD) KATAHANA Microcosm lichen database system Sparcle Prodig	ALNQG Dr.DNA8 FOODLE SGOL SGOIS/DAN
2007	Interlinked Cancer Genome Database Reference database for gene Expression Analysis TMbeta-OBENOME	DPQox HegRisk THE MICROARRAY ANALYSIS SYSTEM - TRIBAX - WULF PORT ZENDE
2006	ASTRA DAEWE GENOIS II	GHIFIRM GRIbs SIP-system Sciens TMbeta-NET
2004		DCCIP GLIPY Prediction program
2002	Flat Brain Sections	
2001	ARCH-NAIO Database for genetic engineering of microbe Database of transcription and translational regulations for filamentous fungal TrAP	

図 2.1-10 分類表示された成果物。「クラス」と「更新日」による成果物の分類の例。

④ サブシステム PubMedScan の運用と改良

PubMedScan は PubMed 文献データベースに登録された文献データのうち、ユーザー毎の興味に応じた新規文献をアラートするシステムで、MEDALS のサブシステムである。本年度は、MEDALS のサブシステムである文献推薦システム PubMedScan の改良を行った。具体的には、

1. アラート間隔をユーザーごとに変更できるようにする。
2. 過去の一定期間に公開された文献を再検索できる機能を新規開発する。
3. お知らせ内容に見つかった文献の Abstract を追加する
4. 登録・推薦文献リストの転送機能とリンク(他ツールとの連携含)

第1点目のアラート間隔は現在1日1回に固定されているが、ユーザーが興味を持っている文献の傾向や生活パターンに合わせ、1週間に1回等の柔軟な設定を可能にした(図 2. 1-11, 12)。第2点目の再検索機能であるが、PubMedScan を使い始めたユーザーは興味のある文献を登録した後、関連する文献が新規公開されるまで PubMedScan の機能を体験することがこれまで出来なかつた。そこで、過去 1 ヶ月程度の一定期間に遡って関連文献を検索できるようにすることで、PubMedScan の利便性を早期に実感できるようにした(図 2. 1-13)。第3点目の Abstract(論文概要)の配信であるが、現在メールで文献を知らせる際、タイトル、著者、雑誌名などは配信しているが、タイトル以上の詳細内容を知るには NCBI のサイトまでリンクをたどっていく必要があつた。ここにあらかじめ(希望するユーザには)文献の Abstract を加えておける機能を設け、手間が省けるようにした(図 2. 1-14)。最後に第4点目の転送機能は、他の文献管理システム(Mendeley や EndNote)との連携を図るものである。個人の文献リストを管理するソフトとしては、上述の2つはメジャーなソフトである。Mendeley は 260 万人の研究者ユーザーが利用しているとされている。これらのソフトと PubMedScan の間で、文献リストのやりとりが出来れば相互的に利便性が高まるということで、この機能を設けた(図 2. 1-15, 16)。なお、最後の2つ(第3、4点目)はユーザへの Web アンケート(後述)によって要望が高いことが明らかとなつたため、追加として開発を行つたものである。

補足であるが、PubMedScan は NITE の二次代謝産物合成遺伝子データベース「DoBISCUIT」において毎日の情報更新に利用されており、成果物の利用促進に貢献している(http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3531092/#_sec2title)。このように PubMedScan を広めることで、ひいては他のデータベース更新や新しいDB構築への助けになるようなことも進めていきたい。

 PubMedScan [Move to BioDBScan](#)

kmamiya My account Logout

[Top](#) [Preferred articles](#) [Recommended articles](#) [Add to preferred articles](#) [Check related articles](#) [Import References](#) [Topic management](#) [Release Information](#)

Your account information

Account edit		Auto search setting
User: kmamiya	Email: testuser@sist.go.jp	Topic: All topic
Password		Interval: Daily
Display abstracts on e-mail reports <input type="checkbox"/>		<input type="button" value="Update"/>
Password confirmation		
<input type="button" value="Update user information"/> <input type="button" value="Delete my account"/>		

PubMedScan (Japanese, English) contact: pubmedscan-ml@aist.go.jp
BioDBScan (Japanese, English) contact: biodbscan-ml@aist.go.jp
MEDALS (Japanese, English)
2008-2014 Copyright© AIST. All rights reserved.

図 2. 1-11 PubMedScan でアラート間隔を変更する画面

Auto search setting

Topic: All topic

Interval: Monthly

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

図 2. 1-12 PubMedScan でアラート間隔で月の配信回数を選択する画面

 PubMedScan [Move to BioDBScan](#)

k.mamiya [My account](#) [Logout](#)

[Top](#) [Preferred articles](#) [Recommended articles](#) [Add to preferred articles](#) [Check related articles](#) [Import References](#) [Topic management](#) [Release Information](#)

Check new recommended article now

If you are new user of the PubMedScan, please try the following items.

1. Add your preferred articles (Default topic) to PubMedScan.
2. Set search conditions;
 Start date: 2014/02/07 End date: 2014/03/06
3. Push button.
4. A short time later, the new recommended articles report is sent to your email address.

Topics: Start date: 2014/03/06 End date: 2014/03/06 Register results to recommended list:

PubMedScan ([Japanese](#), [English](#)) contact: pubmedscan-ml@aist.go.jp
 MEDALS ([Japanese](#), [English](#))
 2008-2014 Copyright© AIST. All rights reserved.

図 2. 1-13 PubMedScan で過去一定期間に遡って関連文献を検索する画面

Account edit

User: k.mamiya

Email: testuser@aist.go.jp abstract 付与が選択可能

Password:

Password confirmation:

Display abstracts on e-mail reports

Password confirmation:

図 2. 1-14 PubMedScan で文献の Abstract を加えておける機能

 PubMedScan [Move to BioDBScan](#)

kmamiya [My account](#) [Logout](#)

[Top](#) [Preferred articles](#) [Recommended articles](#) [Add to preferred articles](#) [Check related articles](#) [Import References](#) [Topic management](#) [Release Information](#)

Import References Notice : To collect the PMID's of your interest easily from your PDF or text files, we also provide software "PMID-Extractor"

How To Use

1. Click the Browse button below and select a reference file(*bib, *ris, *xml) that is exported from a reference management software such as EndNote, Mendeley.
2. Click "Import".

ファイルが選択されていません。

Processing result

total reference number:
number of PMID acquired:
acquisition rate:

Reference titles that successed to acquire PMID

Reference titles that failed to acquire PMID

Topic name: Default ▾

PMID:

PubMedScan (Japanese, English) contact: pubmedscan-ml@aist.go.jp
MEDALS (Japanese, English)
2008-2014 Copyright© AIST. All rights reserved.

図 2. 1-15 PubMedScan で文献リストをインポートする画面

Processing result

total reference number:5
number of PMID acquired:5
acquisition rate:100.00%

Reference titles that successed to acquire PMID

1. No.01 21592391 Causal graph-based analysis of genome-wide association data in rheumatoid arthritis.
2. No.02 16606683 Improved scoring of functional groups from gene expression data by decorrelating GO graph structure.
3. No.03 20628352 Annotating non-coding regions of the genome.
4. No.04 21358748 Genome structural variation discovery and genotyping.
5. No.05 19678920 Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin.

Reference titles that failed to acquire PMID

Topic name: Default ▾

PMID:

```
21592391
16606683
20628352
21358748
19678920
```

図 2. 1-16 PubMedScan でインポートする文献リストの認識が成功した画面

2. 1. 2 統合データベースセンターとの連携

NBDC とはデータベースのカタログ作成活動の他、横断検索の構築運用やアーカイブ活動も連携している。これら連携活動をいっそう規模を広げて進めるため、および省庁連携の活動を広報するために、合同のポータルサイト(integbio)を立ち上げて 2011 年に公開している。図 2. 1-17 に integbio のトップページ(<http://integbio.jp/>)を示す。このページでは、連携の活動や成果を説明するテキスト情報以外に、カタログ、横断検索、アーカイブの 3 つ(図 2. 1-17、右上)を中心に紹介している。

MEDALS のデータベース便覧と integbio カタログとの連携については、互いに便覧データを更新するたびにその都度メールで情報提供を行った。この仕組みによって、MEDALS と integbio のどちらの便覧(カタログ)データに追加や変更があった場合でも、その変更内容を速やかに反映できた。図 2. 1-18 に、MEDALS 便覧と連携しているカタログサイトのページを示す。

The screenshot shows the homepage of the 'integbio.jp' portal. At the top, there is a banner indicating a maintenance period from March 16 to 19, 2012, during which three services (catalog, search, archive) will be suspended. Below the banner, the main content area is titled '4省の生命科学系データベースの統合を目指して'. This section contains a list of four milestones for the integration of life science databases across four ministries:

- (1) カタログ連携：データベースカタログへの登録、リンクの実現
各省のデータベースに関する情報を統一的に、各省のデータベースに対して、リンクがかかる状態を実現する。
- (2) 横断検索連携：横断検索サーバーの設置、横断検索機能の実現
各省における検索サービスに、検索検索サーバーを設置することで検索インターフェースを一元化し、各省のデータベースに対して、横断検索機能が可能となる状態を実現する。
- (3) アーカイブ機能連携：統一データベースのデータロードの実現
各省のデータベースをデータロードによってアーカイブ化して実現し、各省のデータベースに対して、統一形式でのデータダウンロードが可能となる状態を実現する。
- (4) データベース再構築連携：データベース再構築による統合の実現
各省から譲り受けたデータベースを再構築して統合し、各省のデータベースが、統一したフォーマットで再構築された状態を実現する。

On the right side of the page, there are three search boxes for 'Catalogue search (Catalog)', 'Data search (Search)', and 'Data use (Archive)'. Below these boxes, there is a section titled 'NBDC' featuring links to various NBDC-related websites.

図 2. 1-17 省庁連携の活動を紹介する integbio のトップページ

右上の「データベースを探す(カタログ)」は、カタログ統合ページへのリンクである。左下に MEDALS ポータルの紹介がある。

- 生命科学系データベースを一覧から探す -

Integbio データベースカタログ

[全条件をリセット](#)

一覧内を検索する

一覧を絞り込む

生物種

- + 動物 (506)
- + 植物 (225)
- + 原生生物 (36)
- + 菌類 (69)
- + 真正細菌 (124)
- 古細菌 (41)
- ウイルス (47)

カテゴリ

<対象>

- ゲノム (172)
- 遺伝子 (292)
- cDNA (187)

+ 続きを見る

<データの種類>

- 配列 (509)
- 構造 (197)
- 遺伝子発現 (146)

+ 続きを見る

稼動状況

- 稼動中
- 休止
- 運用終了

地域

- 日本
- 日本以外の国・地域

その他の条件

- LSDBアーカイブ収載のDB

[②検索機能について](#)

データベースのレコード一覧

1361 件

最初へ 前へ 1 2 3 4 5 6 7 8 9 10 次へ 最後へ

並べ替え: [レコード公開順](#)

薬用植物総合情報データベース

運用機関: 独立行政法人農業基盤研究所

生物種: Euphorbiaceae | Saxifragaceae | Asteraceae | その他多数

説明: 「薬用植物総合情報データベース」は、漢方薬・生薬に用いられる薬用植物に関するさまざまな情報が、簡単な操作で検索・閲覧することができる、薬用植物の「総合図鑑」です。薬用植物の情報... [詳細へ](#)



YOKBAI (NC: green rice leafhopper) EST

運用機関: 独立行政法人農業基盤研究所

生物種: Nephrotettix cincticeps

説明: シマグロヨコバイ (Nephrotettix cincticeps) のESTデータベースです。データは、様々な発生段階や組織から構築されたcDNAライブラリーに由来しています。クローンごとに、配列情報に加えてGene Ontologyによる... [詳細へ](#)



PVcDNA: YUSURIKA cDNA

運用機関: 独立行政法人農業基盤研究所

生物種: Polypedium vanderplanki

説明: ネムリユスリカ (Polypedium vanderplanki) のESTデータベースです。データは幼虫全体から得られたcDNAライブラリーに由来しています。クローンごとに、配列情報に加えてGene Ontologyによる... [詳細へ](#)



PcsEST: KAMEMUSHI (PS: oriental stink bug) EST

運用機関: 独立行政法人農業基盤研究所

生物種: Plautia stali

説明: 黒樹の苦虫として知られるチャバネアオカズミシ (Plautia crossota stali) のESTデータベースです。データはメス成虫の中腸から構築された2つのcDNAライブラリーに由来しています。クローンご... [詳細へ](#)



J-CHECK: 化審法データベース

運用機関: 独立行政法人製品評価技術基盤機構

生物種:

説明: J-CHECKは、「化学物質の審査及び製造等の規制に関する法律」にかかる厚生労働省、経済産業省及び環境省が、化学物質の安全性情報を発信基盤の充実・強化を目指して化学物質の安全性情報を... [詳細へ](#)



ARTRA: Arabidopsis DNA Microarray probe information

運用機関: かずさディー・エヌ・エー研究所

生物種: Arabidopsis thaliana

説明: カラババイオ株式会社との共同で作製した、シロイナズナのマイクロアレイ情報を探しているデータベースです。DNAマイクロアレイやRNA設計に利用できる転写産物や遺伝子特異的な配列を格... [詳細へ](#)



Strawberry GARDEN: Strawberry Genome And Resource ENtry

運用機関: かずさディー・エヌ・エー研究所

生物種: Fragaria x ananassa | Fragaria iinumae | Fragaria nipponica | Fragaria nubicola | Fragaria orientalis | ...

説明: 食用イチゴのゲノムに関するデータベースです。栽培品種であるFragaria x ananassaのゲノムを近縁野生種4種と比較し、公開されている1品種のゲノム上にマップして得られた遺伝子の、シーケ... [詳細へ](#)



Coelacanth Genome Project

運用機関: 信頼・システム研究機構 国立遺伝学研究所

生物種: Latimeria chalumnae | Latimeria menadoensis

説明: シーラカンスのゲノム情報をまとめたデータベースです。Gbrowseによる閲覧、Blast検索、キーワード検索が可能です。また、頭部のCT画像や3次元像なども収録されています。 [詳細へ](#)



図 2. 1-18 MEDALS 便覧と連携している integbio サイトのデータベースカタログ

21

2. 2 横断検索連携

経済産業省ライフサイエンスポータルサイトに掲載されているデータベースについて、他省のポータル間での横断的な検索を可能とするため、検索インデックスの作成、ノイズのクリーニング等の作業を行った。インデックスの作成においては、統合データベースセンター等と協議の上で統一化している。図 2. 2-1 に MEDALS の横断検索ページへのリンク(赤い四角部分)を示す。ボタンをクリックすると、図 2. 2-2 に遷移するので、キーワードを入れ、サーチアイコン(虫眼鏡のアイコン)をクリックして検索する。



図 2. 2-1 MEDALS トップページと横断検索ワード入力箇所

2 つの赤い枠から横断検索のページに遷移または検索ができる。

インデックスの作成は、まだ横断検索の対象となっていないデータベースについて行うが、すでに検索対象となっているデータベースについてもオリジナルのデータベースの更新をすみやかに検索インデックスへ反映した。このために、上記の前節「ポータルサイトの運営」で述べた「更新検知ツール」を使うことで、データベース自体が更新したことを効率的に把握し、インデックスを最新の状態に維持した。

図 2. 2-2 横断検索のページ

検索キーワード「プロテアーゼ」を入力すると、右下のような検索結果が表示される。

平成25年度のインデックス作成対象については、検索対象として価値の高いデータベースを選定する。1) 比較的新規プロジェクトの成果で注目度が高いもの、2) 独自性が高いデータで広く利用されることが望まれるもの、3) 化合物に関するもの、という3種類の観点から調査・検討した。その結果を鑑み、1)に該当する「RAvariome」「PCDq」、2)に該当する「KNAPSAcK family」に属する新規データベース3件、3)に該当する「MassBank」、の計6件のデータベースをインデックス作成データベースを候補とした。次にクローリングとIDリストの調査を行った結果、検索インデックス作成が可能なものとして、以下の3つについてインデックス作成をした(3件)。

1. MassBank(マススペクトルデータ共有のための分散型データベース)
2. RAvariome(リウマチ関連多型データベース)
3. PCDq(ヒトタンパク質複合体)

本プロジェクトでは横断検索に追加する対象データベースの件数として、3件以上が目標であ

り、3月までに横断検索対象として追加した。

運用でも NBDC と連携を行った。迅速かつ正確な連絡をとることを目的に、横断検索連携専用のメーリングリストにおいて、停止予定、新規 DB 追加、およびインデックスデータ更新等の連絡を年度内(2014 年 2 月まで)で丁度 100 回のメール連絡を行った。

今年度は単にインデックス作成と更新を行うだけでなく、マイクロデータに対応した(外部データベースサイトについての)インデックスを作成すべく、省庁連携にて共通仕様を検討し、その仕様に対応したインデックスを作成した。既出の節(2.1.1-(3)③利便性向上に向けたポータルサイトの改良)では MEDALS 便覧にマイクロデータを埋め込んだ報告を記述したが、この節では、外部データベースに埋め込まれたマイクロデータを、こちらで作成する横断検索インデックスに埋め込み、MEDALS 横断検索システムで表示するようにしたことを報告する。

例として DoBISCUIT に埋め込まれた PubMed ID と最終更新日が表示される図を示す(図 2.2-3)。この図は、キーワード「二次代謝産物合成遺伝子」(黄色のマーキング箇所)で検索したときのものである。



図 2.2-3 MEDALS 横断検索に表示されたサムネイル画像、PubMed ID および最終更新日

2.3 アーカイブ構築連携

各省統一形式でのデータダウンロードを可能とするため、NBDC で提供しているアーカイブ作成ガイドラインに沿ってデータベースのアーカイブ化を実施した（ガイドライン URL は http://dbarchive.biosciencedbc.jp/files/nbdc_dbarchive_guidelines.pdf）。全体の作業は NBDC と作業内容や進捗について協議を行い、NBDC の担当者と作業を分担しつつ進めた。

図 2.3-1 は、MEDALS のトップページにおけるダウンロードページへのリンク位置を示す。MEDALS からダウンロードできる成果物は、データ、ツール（ソフトウェア）、成果報告書となっている。データ統合先の NBDC ではアーカイブサービスがすでに始まっているため、データベースについては NBDC から提供するようにし、その他のツールや報告書についてのみ MEDALS から提供するようにした。



図 2.3-1 MEDALS トップページでのアーカイブへの入り口

ガイドラインにあるアーカイブの作業内容を簡略化して示すと図 2.3-2 のようになる。開発者に対して、アーカイブの意義、作業手順、公開・利用条件などの説明を行い、その後にデータを統一的に扱えるように整形し（紫部分）、非専門分野の研究者でもなるべく内容がわかるような説明を検討して説明文を記述する（緑部分）。最後に公開形態やデータ自体を開発者に確認いただき、許諾書面をいただいてから公開した。

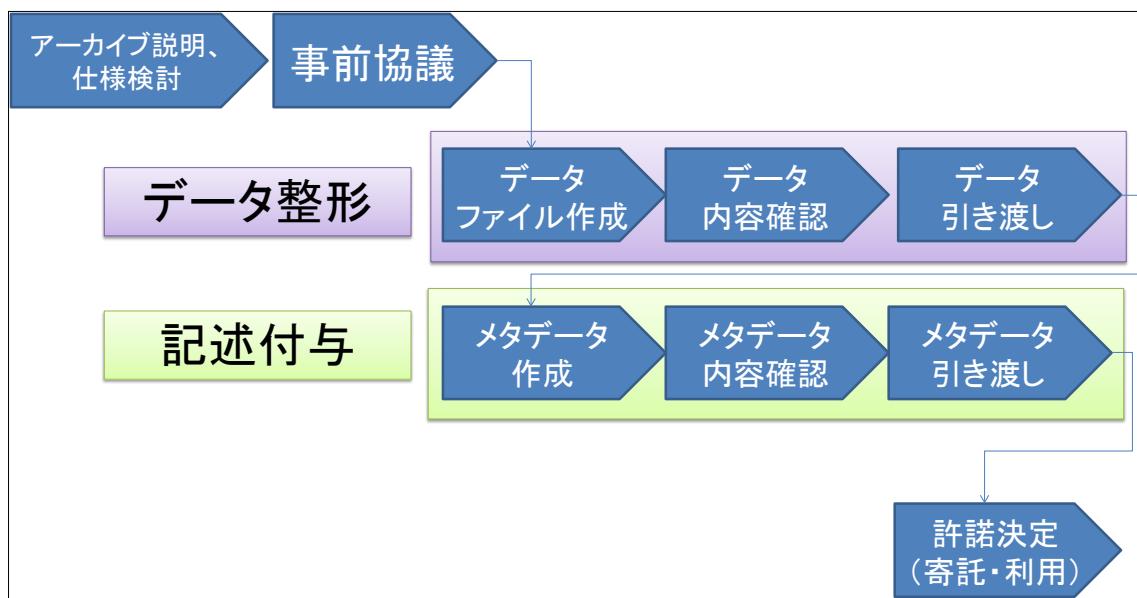


図 2.3-2 アーカイブガイドラインによるデータアーカイブ化の流れ

アーカイブする対象のデータベースの選定については、MEDALS 便覧に記載があるものから、いくつかの選定基準をもって選定した。ユーザーニーズが高いと考えられるデータベースの有力候補は、1) 独自性が高いもの(KNApSAcK family のうち新規に開発されたデータベース3件)、2) データ公開が最近であり鮮度の高いもの(最近公開したもの)として、自己免疫疾患、関節リウマチ(RA)に関わる遺伝子多型情報データベース RAvariome、プロジェクトが終了したものとして、NEDO プロジェクト機能性 RNA の成果である fRNAdb とその関連データベース、NEDO プロジェクト「糖鎖機能活用技術開発」の成果である GGDB(糖鎖関連遺伝子のデータベース)、3) 担当者の異動に伴い今後の維持が困難になると予想されるもの(TMBETA-NET)などであり、初期段階では、これらの候補から検討した。調査の結果、KNApSAcK family の一部は経済産業省の対象データではないことが判明した。また、費用対効果についても検討した結果、RAvariome、PCDq(ヒトタンパク質複合体)、MassBank、MassBase(高分解能マススペクトルデータベース)、酵母画像データDBの5件についてアーカイブ化を行うこととした。作業が完了したデータはNBDCに納めた。3月12日時点ではこのうち一部が未公開であるが、NBDCの確認作業が完了した後、NBDCから公開されることになっている。

2.4 国内外の最新動向調査及び普及に向けた取り組み

本年度の国内外の最新動向調査及び普及に向けた取り組みとして、国内外の最新動向調査、広報・普及活動、ユーザーニーズ調査の3項目を実施した内容を以下に報告する。

2.4.1 国内外の最新動向調査

ライフサイエンス分野のデータベース統合に関する知見を、本事業に活かすため、関連学会等に参加や研究会等に参加し、国内外の動向について調査した。また、関連する技術や研究内容の把握のための資料収集を行った。具体的には、国際学会に3人回、国内学会等に9人回參加した。表2.4-1に、参加した国際会議と国内会議をそれぞれ記した。

表2.4-1 参加した国際および国内学会等

日付	国際会議名	開催場所
2013/07/17-18	Critical Assessment of Genome Interpretation (CAGI 2013)	独・ベルリン
2013/07/19-20	The annual international conference on Intelligent Systems for Molecular Biology (ISMB 2013)	独・ベルリン
2013/09/14-18	HUPO 12th Annual World Congress [Human Proteome Organization (HUPO 2013)]	横浜
日付	国内会議名	開催場所
2013/08/09-11	集団遺伝学研究会	山梨
2013/10/04-05	トーゴーの日シンポジウム 2013～ライフサイエンス分野のデータベース統合の“カタチ”を探る～	東京
2013/10/28-31	生命医薬情報学連合大会（2013年日本バイオインフォマティクス学会年会 情報計算化学生物学会(CBI学会)年次大会 オミックス医療研究会年会）	東京
2013/12/11-14	第36回日本分子生物学会年会(MBSJ2013)	神戸
2014/01/26-31	データ統合技術開発会議(バイオハッカソン)	沖縄

(1)国際会議の調査事例①:Critical Assessment of Genome Interpretation (CAGI 2013)

2013年7月17日～18日にドイツベルリンの Max Planck Institute for Molecular Genetics で開催された Critical Assessment of Genome Interpretation (CAGI 2013)に参加しポスター発表を行つ

た。CAGI は、ゲノム情報から体質や疾患といった表現型を予測する手法を客観的に評価することを目的としたコミュニティで、数年に 1 度、ゲノム-表現型の予測を課題としたコンテスト形式のカンファレンスを開催している。報告者は、ここでポスター発表を行った。

本年度は、ゲノムデータに基づく患者/健常者の判定、ゲノムデータと表現型の組み合わせ問題、疾患原因遺伝子の同定、スプライシングや変異によるタンパク質への影響予測などの 11 問の課題(チャレンジ)が提出された。CAGI の予測コンテストに参加する方法は簡単で、CAGI に登録後、ウェブ経由で各チャレンジのデータセット(エクソーム解析データなど)を取得し、締切までに予測手法と結果を投稿する方式である。データ解析に与えられた期間は 2012 年 10 月から 2013 年 4 月までの 6 か月で、1つのチャレンジに対して複数の予測結果を投稿することができる。CAGI では、予測の公正な競争を行うためにチャレンジのデータセットは未発表のデータに限定している。

(2)国際会議の調査事例②: The annual international conference on Intelligent Systems for Molecular Biology (ISMB 2013)

ISMB は世界最大のバイオインフォマティックスに関する国際学会であり、世界のメジャーなものばかりではなくマイナーなデータベースについても多くの発表がある。ここから最新情報等を収集した。

たとえば、Tomás Di Domenico は、MobiDB(天然変性領域のデータベースを開発した。17,285 の蛋白質について、変性領域かどうかのアノテーションがついている。DisProt, PDB (X-Ray and NMR) など 9 つの予測を統合している。また、Cellular Phenotype Database (<http://www.ebi.ac.uk/fg/sym>) は、HT スクリーニングから得られる表現型を格納している。現時点ではヒトとハエのみに対応している。Warren Kibbe (Northwestern University, United States)らは、ヒト疾患のオントロジーデータベースである Disease Ontology: coupling human disorder phenotypes to gene annotations in diverse organisms について進歩を講演した。その他、Human Genome Variation Society のガイドラインに従う「変異の専門用語」をチェックするためのソフトウェアとして、Mutalyzer (<https://mutalyzer.nl/>) が開発された。この gene variant database では未公開データの直接サブミッションが可能であることが特徴である。Aniadou は、自然言語処理の発表で、NACTEM 研究所のさまざまなサービスを紹介した(<http://www.nactem.ac.uk/services.php>)。この他にも多数の国内外のデータベースについて最新情報を収集する事ができた。

(3)国際会議の調査事例③: HUPO 12th Annual World Congress [Human Proteome Organization (HUPO 2013)]

HUPO2013 は Human Proteome Organization (HUPO)が主催するヒトプロテオームに関する国際会議で、今年は日本で開催され、日本プロテオーム学会が主催した。この会議で個別データベースの進歩状況を調査した。Matius らは Antibodypedia を紹介した。このデータとしては同じ蛋白質について複数の抗体がある。抗体の数が 9 個より少ない蛋白質は、全体の 20%くらいであり、100

個以上くらい抗体があるタンパク質が多い。human protein atlas (<http://www.proteinatlas.org/>) や Protein Atlas (<http://www.proteinatlas.org/>) の講演では、多くの転写物が格納されていた(27 normal tissues and 44 cell lines)。イメージファイルも 13million 個あり、2005 年時点ですでに 18,000 遺伝子格納されていた。将来は、tissue, cancer, subcell, cell line を加えていく予定という。AntibodyWiki の紹介があった。また、MS によるヒトペプチドの検出や PeptideAtlas の紹介があった。染色体ベースでプロテオームを解読していくプロジェクトとして、C-HPP があり、Wiki が公開し始めている段階である。

(4) 国内会議の調査事例②:トーゴーの日シンポジウム 2013

2013 年 10 月 4-5 日に東京で開催されたトーゴーの日シンポジウム 2013 に参加し、ポスター発表 2 件および受付にてパンフレット 3 種(100 部)の配布を行った。本シンポジウムは、NBDC の主催で 4 省のプロジェクトが連携し毎年開催している。統合推進プロジェクトの成果発表の場でもあり、過去に採択されたデータ公開活動が紹介された。

京都大学の松沢教授は、ヒトの病気はマウスで調べても使えないとして、大規模コホート研究の情報を蓄積する重要性を説いた。J-ANDI プロジェクトでは、アルツハイマーの早期発見にむけ、解析を行っている。タウ編成について紹介していた。この関連のオープンにするデータ(メタデータ)は平成 25 年に公開する予定という。制限されたアクセスでの公開(生データ)は J-ADNI 内で公開するという予定が発表された。

多くのプロジェクトで、データ公開の枠組みとして RDF というフレームワークによる公開形式をとっており、国内バイオプロジェクトにおいては、セマンティック Web や RDF がひろがっているという印象を強く受けた。

2.4.2 広報・普及活動

本事業で構築した成果物の利用普及のため、2.4.1で調査を実施した学会やイベント等において、ポスター発表、口頭発表等の機会を利用し、広報・普及活動を行った。

(1)パンフレット配布

図 2.4-1 のようなパンフレットを増刷して、学会や発表会において配布し、広報に役立てた。



図 2.4-1 今年度広報に用いたパンフレット(A4)

(2)学会発表およびその機会を利用した広報活動

学会等において MEDALS およびヒト遺伝子統合データベース H-InvDB に関するポスター発表・口頭発表を行い、その機会を利用してデモンストレーションやパンフレット配布等の広報活動を実施した。全体では、国際学会発表 3 件、国内発表 8 件、講習会 4 回を行った。

例として、2013 年 12 月 11 日から 14 日に神戸で開催された第 36 回日本分子生物学会年会 (MBSJ2013)でのブース出店とデモの様子を示す(図 2.4-2)。

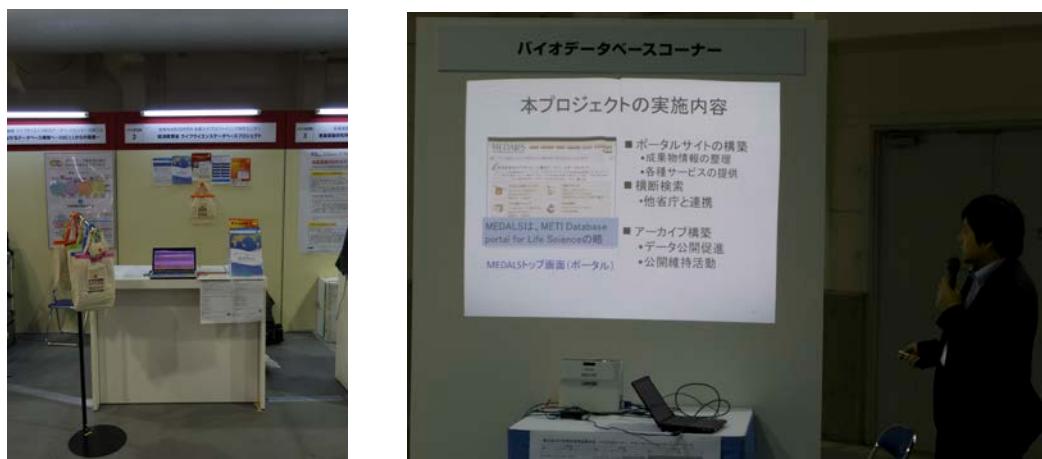


図 2.4-2 学会発表風景(日本分子生物学会年会において講演、ブースでのデモを実施)

(3) 講習会開催

MEDALS および H-InvDB に関する講習会を全国各地で計 4 回開催し、内容の講義・デモンストレーションおよび実習を実施した(表 2. 4-2)。当研究チームで、創薬研究に役立つヒト遺伝子データベース H-InvDB を運用しているため、これとタイアップすることにより内容的に幅の広い魅力的な講習会プログラムを企画し、実習を含めた対面での講習会により、MEDALS ツールの操作などの応用的な MEDALS の活用法を発信した。

また、第 2 回講習会を除くすべての回で Ustream による中継および映像による配信を行った(例えば、<http://www.ustream.tv/recorder/43027512>)。第 2 回講習会では映像専用に利用するための回線の利用手続きがとれなかっただため、送信することができなかった。

(4) SNS(ソーシャル・ネットワーキング・サービス)を用いた情報発信

SNS(ソーシャル・ネットワーキング・サービス)はここ数年で急速に普及した新たな情報発信・情報収集の手法で、twitter や Facebook は国内外の学術出版社や主要データベース等からの情報発信に利用されるようになった。昨年度のプロジェクトから MEDALS の SNS を使った広報を開始しているが、今年度も引き続き実施した。

具体的には、twitter および Facebook にアカウントを作成しており、内容は同じものが出るように連結されている。MEDALS のアップデートやイベントの種別ごとに区別できるハッシュタグを付けニュース配信を合計 280 件行った(表 2. 4-2)。図 2. 4-3 に配信の様子を示す。

表 2. 4-2 今年度開催した講習会

日付	講習会名	開催場所	参加人数
2013/12/12	第 1 回講習会	東京(産総研・臨海副都心センター)	13 名
2013/12/25	第 2 回講習会	神奈川(東海大学)	13 名
2014/1/24	第 3 回講習会	大阪(産総研・関西センター)	5 名
2014/2/13	第 4 回講習会	茨城(産総研・つくば)	7 名

The image shows two side-by-side screenshots of social media profiles for MEDALS. On the left is the Twitter profile (@medals2), which has 62 followers and 109 tweets. It features a logo with overlapping colored bars and a bio stating: 'MEDALS(METI Database portal for Life Sciences)は、経済産業省サイエンス政策データベースサイトです。' Below the bio are three recent tweets. The first tweet links to a news article about the MEDALS database reaching 100 million entries. The second tweet links to another news article about the database reaching 200 million entries. The third tweet links to a news article about the database reaching 300 million entries. On the right is the Facebook page (Medals Meti), which has 1,174 likes and 1,174 posts. It also features the same logo and bio. Below the bio are three recent posts, each linking to a news article about the database reaching 100 million, 200 million, and 300 million entries respectively.

図 2. 4-3 SNS(ソーシャル・ネットワーキング・サービス)を用いた情報発信例

また、Facebook の「いいね！」ボタン設置は当初全ページ共通であったが、MEDALS 便覧の各ページに Facebook によるシェアボタンの設置を行い、人気のあるデータベースでは数字が増えて際立たせるようにした(図 2. 4-4)。

The screenshot shows the MEDALS database portal interface. At the top, there is a navigation bar with links for Home, Database, Tools, Download, About, English, Q and A, and Practice. Below the navigation bar, there are social sharing buttons for RSS, Twitter, and Facebook. The Facebook button has a red arrow pointing to it, indicating its placement. A red box highlights a message box that says "このデータベースのみの 票数を示す" (Shows the number of votes only for this database). To the left of the message box, there is a thumbnail image of a database entry titled "forte". On the right side, there is a table with three rows:

成果物名	FORTE
成果物の別名	なし
成果物に関する説明	FORTEは、タンパク質構造類似性質のアミノ酸配列を用いて、既知で表示されます。

図 2. 4-4 MEDALS 個別データベース便覧での SNS(ソーシャル・ネットワーキング・サービス)への誘導

2.4.3 ユーザーニーズ調査

MEDALS の潜在ニーズを調べることを主な目的として、Web 上のアンケート調査および対面でのヒアリング調査を実施した。ニーズ調査結果で判明した改善すべき点は、その後の活動に反映させた。なお、過去のアンケート設問を MEDALS サイトで公開している。

(1) アンケート調査

MEDALS の潜在ニーズを調べることを主な目的として、専門会社に委託したウェブアンケート調査を 2013 年の 9 月から 10 月にかけて実施し、123 名の回答結果を集計した。質問項目は自由記述回答も含む全 30 問で、閲覧頻度と感想に関しては回答しやすいようテーブル形式の選択肢を採用した。実施内容の概要を表 2.4-3 に示す。

表 2.4-3 2013 年度アンケート調査実施概要

調査期間	2013/09/18～2013/10/20（約 1 か月）
調査方法	Web アンケート（30 問）
目的	・ユーザーニーズに即したより良いポータルサイトの構築 ・研究成果の連携強化 ・MEDALS の実装および経産省関連プロジェクト成果物の広報
調査項目	・MEDALS ポータルサイト ・横断検索 ・経産省関連プロジェクト新規成果物
有効回答	123 名 (H-InvDB メールマガジン購読者、SNS 閲覧者、各種学会会員等)

① 調査対象者について

アンケート案内の配信先を検討した結果、5 割以上の回答者の所属が公的研究機関と民間企業であった（図 2.4-5）。

【アンケート案内の配信先】

- H-InvDB メールマガジン購読者
- 日本バイオインフォマティクス学会会員
- 日本分子生物学会会員
- BIOWEB 登録者など

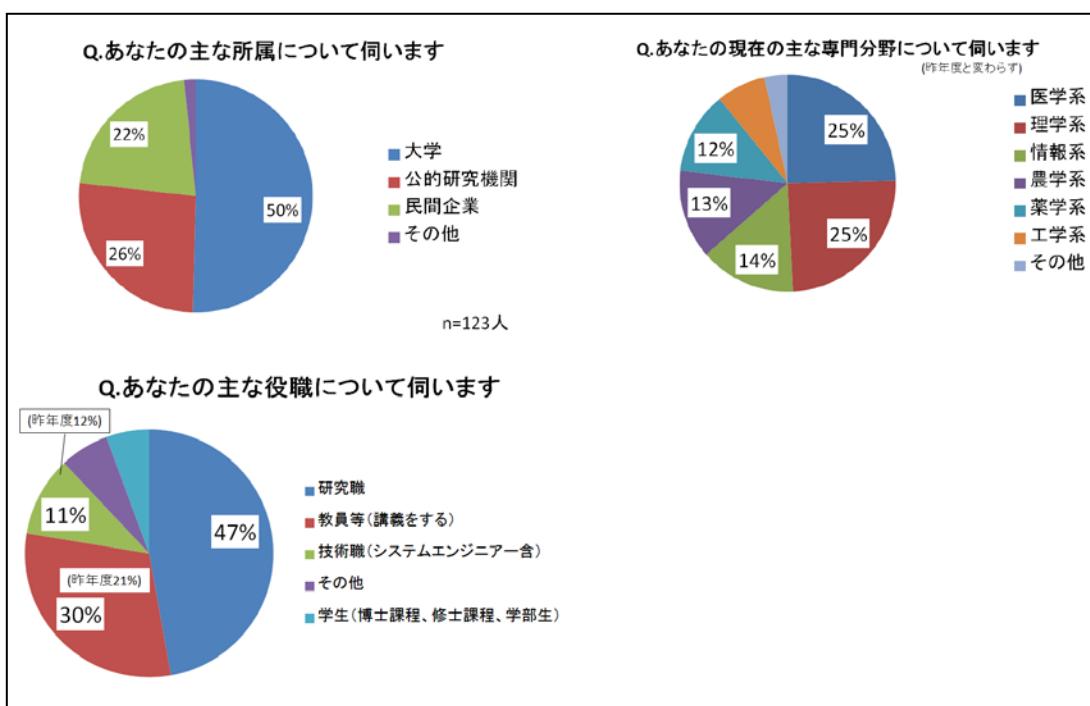


図 2.4-5 2013 年度アンケート調査 実施対象

② MEDALS ポータルサイトの利用について

MEDALS ポータルサイトの各コンテンツの利用、および国内主要バイオポータルの利用について選択式および自由記述形式で調査した(図 2.4-6)。昨年度よりも利用したことがある割合が増えていた。とくに PubMedScan については、月に1回以上と頻繁に利用しているユーザーが多い。

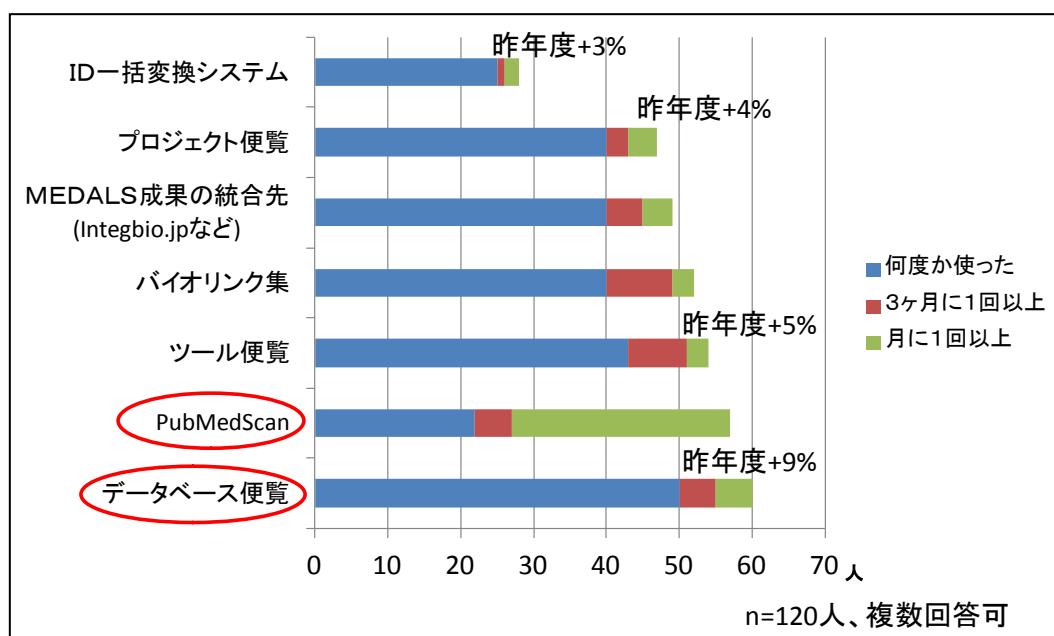


図 2.4-6 MEDALS コンテンツの利用状況

以下に、ポータルサイトへの希望・要望についての自由記述回答の一部を紹介する。

・便覧の内容について拡充すべきものは？

- ✓ 化合物に関する ID 変換サービス(化合物に関する複数のデータベース間の ID 対応表)
- ✓ 各種プロジェクトの報告書

・PubMedScan について

- ✓ 他アプリとの連携を強化してほしい(EndNote、Mendeley 等)
- ✓ キーワードによる検索機能をつけてほしい

これらについての対策としては、ID 一括変換システムでより多くの化合物データを含める、プロジェクト便覧以外のものを再調査する、PubMedScan の拡張開発(セクション「2.1.1-(3)ポータルサイトの運営、④サブシステム PubMedScan の運用と改良」を参照)を行った。

③ データベース活用への課題について

本プロジェクトでは、データベース普及のためのあらゆる手段を検討・計画しているが、より広い視野に基づく意見を収集するための設問をいくつかもうけている。その中で「今後更なるデータベースの活用を目指すためには何が必要でしょうか」という質問についての回答を図 2.4-7 に示す。この回答には、各データベースのインターフェースの改良や、データベース講習会を設けることの意見が多くかった。これをさらにまとめると、各データベースの使い方が分かりづらいという意見のようである。対策としては、本プロジェクトのような年度ごとのプロジェクトでは対応が難しいが、長期的な取り組みとしてであれば、仕様のモデルやガイドラインを策定することなどが考えられる。今後の課題としたい。

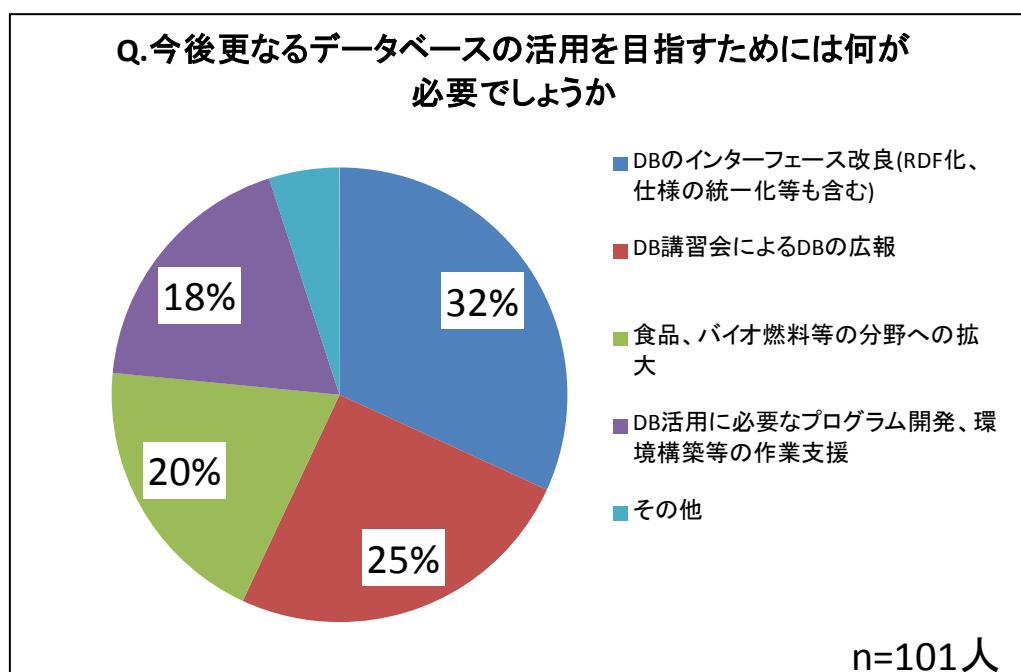


図 2.4-7 更なるデータベースの活用のためにすべきこと

(2)ヒアリング調査

現場の意見を集めることを目的として、分子生物学会でのブースによる広報時、および各データベース講習会時に参加者と講師を対象にポータルサイトやデータベース利用の問題点などについてヒアリング調査を実施した。個別のデータベースについての意見が多かったが、共通する意見としては、参加者から広報活動に力を入れてほしいという意見が多くあった。講習会では実習の時間が足りないという意見があり、そのため後半では実習時間を確保するように努めた。ヒアリング調査によって、ウェブアンケート調査だけでは調べきれない内容も聞くことができ、大変有意義であった。

2.5 運営委員会

経済産業省ライフサイエンスデータベースプロジェクトの進捗状況を明らかにし、議論を通じて外部有識者(アドバイザー)らの意見を成果物に反映させることを目的として、2013年11月28日に運営委員会を産総研臨海副都心センターにて開催した。アドバイザー、研究協力者、ならびにバイオサイエンスデータベースセンターなどからのオブザーバーを含め、22人が参加した。アドバイザーならびに研究協力者は4-(2)、4-(3)参照のこと。運営委員会では、プロジェクトの実施テーマごとに進捗報告とアドバイザー等からのコメントがあり、続いて総合討論を行った。ポータルサイト MEDALS の運営・開発、ならびに成果物の調査が順調に進んでいること、ならびに本年度産総研で RIODB が停止した件について報告がなされた。また、横断検索、アーカイブ作成、学会での広報といった NBDC との連携も滞りなく緊密に行われていることが確認された。とくにアドバイザーからは、「経産省ライフサイエンスデータベースプロジェクトを長期的視野で運営すべきである」、「次年度からの継続的事業は行われないが、なんらかの継続的な事業の形を模索すべきである」という意見がでた。

2.6 成果発表

平成25年度の本プロジェクトの成果に関する発表を以下に示す。

2.6-(1) 論文

1. Yasuhiro Takenaka, Akiko Noda-Ogura, Tadashi Imanishi, Atsushi Yamaguchi, Takashi Gojobori, Yasushi Shigeri (2013) Computational analysis and functional expression of ancestral copepod luciferase. *Gene* 528 (2) : 201–205. 2013.10
2. Yoko Nagai, Tadashi Imanishi (2013) RAvariome: a genetic risk variants database for rheumatoid arthritis based on assessment of reproducibility between or within human populations. *Database* 2013 : bat073. 2013.10
3. Kanako O Koyanagi, Tadashi Imanishi, Takashi Gojobori (2013) Bidirectional Gene Pairs in the Human Genome. *Encyclopedia of Life Sciences*, John Wiley & Sons, Ltd: Chichester: 1–9. 2013.12

2.6-(2) 発表

1. 村上 勝彦, 山崎 千里, 今西 規 (2013) 「ヒト遺伝子データベース H-InvDB の RDF 化と Endpoint の公開」2013年度人工知能学会全国大会 2013.06
2. Katsuhiko Murakami, Tadashi Imanishi (2013) 「Annotation network from different data sources of genes and proteins by correlation analysis」ISMB/ECCB 2013 2013.07
3. Katsuhiko Murakami, Tadashi Imanishi (2013) 「Automatic Extraction of Correlation of Annotation Terms in Databases to Find Similar Concepts, Synonyms, and Multifunction」 HUPO 12th Annual World Congress 2013.09

4. Katsuhiko Murakami, Yoko Nagai, Tadashi Imanishi (2013) 「PCDq and H-EPD: Human Protein Complex Database and Human Protein Sequence Database Developed to Research Uncharacterized Human Proteins」 HUPO 12th Annual World Congress 2013.09
5. Tadashi Imanishi (2013) 「Integration and data-mining of human transcriptome and proteome databases in H-InvDB」 HUPO 12th Annual World Congress 2013.09
6. 世良 実穂, 長井 陽子, 村上 勝彦, 羽原 拓哉, 今西 規 (2013) 「ヒト転写産物統合データベース H-InvDB のプロテオミクスへの利用」トーゴーの日シンポジウム 2013 2013.10
7. 村上 勝彦, 間宮 健太郎, 今西 規 (2013) 「経済産業省ライフサイエンスデータベースプロジェクト・ポータルサイト MEDALS」トーゴーの日シンポジウム 2013 2013.10
8. Katsuhiko Murakami, Tadashi Imanishi (2013) 「Correlation of annotation terms in heterogeneous databases and its application to Gene Set Enrichment Analysis (GSEA) and ontology construction」生命医薬情報学連合大会「オミックス・計算 そして創薬」 2013.10
9. Yoko Nagai, Katsuhiko Murakami, Tadashi Imanishi (2013) 「Gene Set Enrichment & Clustering method to identify common characteristics of GWAS-detected genes」生命医薬情報学連合大会「オミックス・計算 そして創薬」 2013.10
10. 斎藤 穎一, 村上 勝彦, 今西 規 (2013) 「Batch-Learning SOM(BLSOM)法を用いた海洋微生物の比較メタゲノム解析」第 29 回日本微生物生態学会大会 2013.11
11. 今西 規 (2013) 「バイオデータベースの統合化とデータマイニングを推進するためのリンク自動管理システム」第 36 回日本分子生物学会年会 2013.12
12. 村上 勝彦, 今西 規 (2013) 「データベース記述用語の相関解析とその応用:高度 GSEA とオントロジー構築へ向けて」第 36 回日本分子生物学会年会 2013.12
13. 世良 実穂, 長井 陽子, 村上 勝彦, 羽原 拓哉, 今西 規 (2013) 「ヒト転写産物統合データベース H-InvDB のプロテオミクスへの展開」第 36 回日本分子生物学会年会 2013.12
14. 間宮 健太郎, 村上 勝彦, 世良 実穂, 今西 規 (2013) 「MEDALS(経済産業省ライフサイエンスデータベースプロジェクト・ポータルサイト)の進展」第 36 回日本分子生物学会年会 2013.12
15. 今西 規 (2013) 「ヒト遺伝子統合データベース H-InvDB と関連ツール」平成 25 年度第 1 回データベース講習会「創薬研究におけるバイオデータベース講習会」2013.12
16. 今西 規 (2013) 「ヒト選択的スプライシングデータベース H-DBAS」平成 25 年度第 1 回データベース講習会「創薬研究におけるバイオデータベース講習会」2013.12
17. 村上 勝彦 (2013) 「統合データベースポータルサイト MEDALS の活用」平成 25 年度第 1 回データベース講習会「創薬研究におけるバイオデータベース講習会」2013.12
18. 村上 勝彦 (2013) 「統合データベースポータルサイト MEDALS の活用」平成 25 年度第 2 回データベース講習会「医学研究のためのバイオデータベース講習会」2013.12
19. 今西 規 (2013) 「ヒト遺伝子の統合データベース H-InvDB とリンク自動管理システム」平成 25 年度第 2 回データベース講習会「医学研究のためのバイオデータベース講習会」2013.12

20. 村上 勝彦 (2014) 「統合データベースポータルサイト MEDALS の活用」平成 25 年度第 3 回
データベース講習会「創薬研究におけるバイオデータベース講習会」 2014.01
21. 今西 規 (2014) 「ヒト遺伝子統合データベース H-InvDB と関連ツール」 平成 25 年度第 3 回
データベース講習会「創薬研究におけるバイオデータベース講習会」 2014.01
22. 村上 勝彦 (2014) 「統合データベースポータルサイト MEDALS の活用」平成 25 年度第 4 回
データベース講習会「創薬研究におけるバイオデータベース講習会」 2014.02
23. 世良 実穂 (2014) 「ヒト遺伝子統合データベース H-InvDB」平成 25 年度第 4 回データベース
講習会「創薬研究におけるバイオデータベース講習会」 2014.02
24. 今西 規 (2014) 「H-InvDB 関連 DB とツールの利用」平成 25 年度第 4 回データベース講習
会「創薬研究におけるバイオデータベース講習会」 2014.02
25. 斎藤 祐一, 村上 勝彦, 今西 規 (2014) 「BLSOM を利用した腸内細菌叢の比較メタゲノム
解析」2014 年度農芸化学会大会 2014.03

2.6-(3) データベース公開・更新

2013 年

- | | |
|-----------|---|
| 6 月 17 日 | リンク自動管理システムに Human Protein ATLAS(正常およびがん組織におけるタンパク質の発現データ)を正式に追加 |
| 6 月 17 日 | 経済産業省ライフサイエンス関連プロジェクトのポータルサイト MEDALS 更新
対応データベースに H-EPD を追加 |
| 7 月 9 日 | リンク自動管理システムに Peptide ATLAS(質量分析法によるタンパク質断片
の測定データ)を追加 |
| 8 月 27 日 | リンク自動管理システムに Ensembl Protein(Ensembl 版のヒトタンパク質)を追
加 |
| 9 月 11 日 | 経済産業省ライフサイエンス関連プロジェクトのポータルサイト MEDALS 更新
対応データベースに H-EPD を追加 |
| 9 月 19 日 | 経済産業省ライフサイエンス関連プロジェクトのポータルサイト MEDALS 更新
対応データベースに Recount DB を追加 |
| 10 月 18 日 | リンク自動管理システムに neXtProt(ヒトタンパク質の実験的エビデンス)を追
加 |

2014 年

- | | |
|----------|--|
| 1 月 30 日 | 経済産業省ライフサイエンス関連プロジェクトのポータルサイト MEDALS 更新
対応データベースに化審法データベースを追加 |
| 1 月 30 日 | 経済産業省ライフサイエンス関連プロジェクトのポータルサイト MEDALS 更新
対応データベースに SAHG を追加 |

2月28日	経済産業省ライフサイエンス関連プロジェクトのポータルサイト MEDALS 更新 対応データベースに TuMaRdb を追加
2月28日	経済産業省ライフサイエンス関連プロジェクトのポータルサイト MEDALS 更新 対応データベースに GlycoPOD を追加
2月28日	経済産業省ライフサイエンス関連プロジェクトのポータルサイト MEDALS 更新 対応データベースに GlycoExplorer を追加
3月5日	MEDALS 横断検索更新

3. 謝辞

本プロジェクトの実施にあたり、多数の方のご協力、ご貢献を頂きました。ここに深く感謝申し上げます。

- 便覧・横断検索・アーカイブの作成にご協力いただいた各データベース・解析ツール開発者の方々
- アンケート調査にご協力いただいた方々
- ヒアリングにご協力いただいた企業、研究所、大学の方々

4. 参加者名簿

4-(1) プロジェクト実施者

役職	氏名	所属機関
プロジェクトリーダー	今西 規	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
プロジェクト顧問	五條堀 孝	KAUST 国立遺伝学研究所
研究員	村上 勝彦	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
	斎藤 穎一	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
	間宮 健太郎	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
	世良 実穂	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
	羽原 拓哉	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
	遠藤 智宏	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

4-(2) 運営委員(アドバイザー)

油谷 浩幸	東京大学先端科学技術研究センター 教授
長洲 毅志	エーザイ株式会社 理事・CSO 付 担当部長
田中 博	東京医科歯科大学 難治疾患研究所(生命情報学) 教授

4-(3) 研究協力者

高木 利久	独立行政法人 科学技術振興機構(JST) バイオサイエンスデータベースセンター(NBDC)
畠中 秀樹	独立行政法人 科学技術振興機構(JST) バイオサイエンスデータベースセンター(NBDC)
川本 祥子	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)
光山 統泰	独立行政法人 産業技術総合研究所 生命情報工学研究センター(CBRC)

藤田 信之

独立行政法人 製品評価技術基盤機構(NITE)

バイオテクノロジーセンター 上席参事官

矢野 貴久

独立行政法人 新エネルギー・産業技術総合開発機構

(NEDO)バイオテクノロジー・医療技術部 主査

以上