

「ライフサイエンス分野の統合データベース整備事業」

大学共同利用機関法人 情報・システム研究機構

情報・システム研究機構長 堀田凱樹

(様式1)

提案書

(ライフサイエンス分野の統合データベース整備事業)

1. 提案課題

責 任 機 関 研 究 代 表 者	大学共同利用機関法人 情報・システム研究機構 堀田凱樹
課題の概要 (200 字以内)	我が国に多数存在するライフサイエンス関係のデータベース(DB)を多くの研究者等が容易に統合利用可能にするための組織的および技術的基盤づくりとして、(1)DB整備の戦略立案支援、(2)DB統合化基盤技術開発、(3)ポータルサイトの整備、を行う。
ミッションステートメント(200字以内)	多くの研究者等が十分な情報をストレスなく利用して研究できる環境整備のために (1)ライフサイエンス及び知識情報処理の識者にライフサイエンスDBの専門家を加えた統合DB整備戦略作業部会を組織し戦略を議論。また、戦略立案の基盤となる情報の収集・分析を実施(2)統合化に不可欠な知識表現法、情報共有技術および文献からの知識抽出技術等を開発(3)DB及び解析ツールの所在を網羅・分類・解説したポータルサイトを構築
課題名	ライフサイエンス統合データベース基盤整備

2. 代表者等

責任機関	大学共同利用機関法人 情報・システム研究機構			
研究 代表者 (機関の 長とする こと)	ふりがな 氏名	ほった よしき 堀田 凱樹	生年月日	2006年4月1日現在の年齢
	機関・所属部署名	大学共同利用機関法人 情報・システム研究機構		役職名 機構長
	ふりがな 所在地	〒105-0001 <small>とうきょうとみなとくらのもん</small> 東京都港区虎ノ門4-3-13 <small>しゅうわかみやちょう</small> 秀和神谷町ビル2階		
	TEL. (03)6402-6201	FAX. (03)3431-3070		
	E-mail: yhotta@rois.ac.jp			
	所属機関 の産学官	「学」		エフォート(年間全業務時間を100%とした際の の当該研究に割く時間配分率) 10%
事務連絡 担当者 (当該担当者に審査 結果等すべての連絡 をいたします)	ふりがな 担当者名	こうのかつとし 河野克俊	役職名	総務課長
	機関・所属部署名	大学共同利用機関法人 情報・システム研究機構 事務局		
	事務 連絡先	〒105-0001 東京都港区虎ノ門4-3-13 秀和神 谷町ビル2階		
	(当該担当者に審査結果等す べての連絡をいたします)	TEL. 03-6402-6206	FAX. 03-3431-3070	
E-mail: kono@rois.ac.jp				

- 3 業務実施予定期間 予定
平成18年8月～平成19年3月まで
なお、業務終了は年度末とする。

- 4 経費の見込額(概算)
初年度 200百万円(直接経費 180百万円+一般管理費20百万円)
総額 200百万円(直接経費 180百万円+一般管理費20百万円)

課題名「ライフサイエンス統合データベース基盤整備」

研究代表者名 「堀田凱樹」 責任機関名 「大学共同利用機関法人 情報・システム研究機構」

研究の目標・概要
<p>1. ミッションステートメント</p> <p>より多くのライフサイエンス研究者等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 等をストレスなく利用でき、より高度な研究ができる環境の実現のために</p> <p>(1) ライフサイエンスおよび知識情報処理の識者にライフサイエンス DB の専門家を加えた統合 DB 整備戦略作業部会を組織し、十分な情報の収集・分析を行い、それに基づいて統合化戦略を議論する。(2) 統合化に不可欠な知識表現法、情報共有技術および文献からの知識抽出技術等を開発し、戦略立案のための調査(1)やポータルでの利用者誘導(3)に利用する。(3) DB や解析ツールの所在を網羅・分類・解説したポータルサイトを構築する。</p> <p>2. 内容</p> <p>(1) ライフサイエンス(基礎生物学、薬学、医学、農学) 知識情報処理、ライフサイエンス DB の三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 DB 整備戦略を議論する。上記委員会に対し、三分野の動向に関する基礎データを提示し、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)を整備する。また、この体制を利用して戦略立案の基盤となる情報の収集・分析を行う。</p> <p>(2) 知識やデータの表現方法が統一されていない多くの DB を機能的にまた物理的に統合利用可能にするために必要な知識表現法、情報共有技術および文献からの知識抽出技術等を開発する。さらに医学分野に関しては多型情報および癌関連のゲノム情報や発現情報の DB 構築や統合 DB のプロトタイプ構築を通じて分子医学データの統合化に必要なデータ項目や利用モデルを検討する。</p> <p>(3) 内閣府連携施策群「ライフサイエンス分野のデータベースの統合化に関する調査研究」において作成された国内外の DB 網羅リストおよび JST においてこれまで作成してきたポータルサイト WING - DB の詳細な内容説明を基礎に、わかりやすく分類され、また、ライフサイエンスの各分野での利用に応じた説明がついた網羅性の高い DB のカタログを作成し、さらに JST で整理を進めてきた解析ツールなどの WEB リソースポータルを統合することで多様な利用者を目的にかなった DB やリソースに誘導するポータルサイトを構築する。</p> <p>3. 実施体制</p> <p>責任機関である情報・システム研究機構は全体の統括と各サブテーマを遂行し、JST はサブテーマ(3)を、九州大学はサブテーマ(2)の一部を分担する。</p>
諸外国の現状等
<p>1. 外国の現状</p> <p>米国では、医学生物学の総合文献 DB (PubMed) を作成している国立医学図書館 (NIH NLM) 内のバイオ技術情報センター (NCBI) へ核酸配列バンク事業 (GenBank) が移管されたのを契機に統合化の動きがはじまった。その後、ゲノムプロジェクトの進行中に遺伝病電子辞典や遺伝子地図情報などが同センターに集合する形でさらなる統合化が進み、最近では化合物情報も取り込む形で統合化が行われている。これらは集中型の統合化と言えるものである。一方、欧州では、世界で最も古い核酸配列バンクを作ったことで知られる欧州分子生物学研究所 (EMBL) が設立した欧州バイオインフォマティクス研究所 (EBI) において分子情報を中心とした各種データの整理、連携、提供が進められている。これに加え、最近では高度な情報共有のための技術開発も進められるなど、分子情報を基盤とした分散型の統合化が行われている。</p> <p>2. 我が国の水準</p> <p>国際的な研究サービス活動である DNA バンク事業や蛋白質構造バンク事業に欧米外から唯一参加する施設である国立遺伝学研究所の DDBJ や大阪大学の PDBJ の活動は国際的にも意義深いものである。また京都大学の知識データベース KEGG は開始後 10 年を経て自ら切り開いた知識データベース分野で不動の地位を得ている。このようにバンク事業や個別のデータベースについては世界的なものが散見される一方で、DB 利用環境全体を見渡してみると、我が国には長期戦略が存在せず、そのため大型測定研究プロジェクトのデータさえ十分に有効活用できるようにはなっていない。これらはプロジェクト末期に独自の成果 DB として公開されるものの、他の多くの DB との統合利用環境の構築が十分ではなく、その成果は十分に理解・利用されずに陳腐化して行く傾向にある。従って国内の利用者にとって国産データに関する利便性の水準は必ずしも高いものではない。</p>
研究進展・成果がもたらす利点等
<p>分子生物学以降ライフサイエンスの成果の主体は、遺伝子や蛋白質に関するデータや言葉を使った生命現象の記述である。この膨大な成果はいまやデータベースの助けなく整理することも利用することも不可能である。従ってデータベース統合の効果はゲノム研究やポストゲノム研究に限定されずライフサイエンス全体に波及する。本プロジェクトの遂行により、真に利用者の利便性を考えた、また、ライフサイエンス研究の動向や情報技術の動向を見据えた長期的な DB 戦略の立案が可能な体制が整う。また、統合利用へむけた基盤的な技術課題が克服されることで、より多様な背景知識の研究者等がストレスなくプロジェクト研究やバイオインフォマティクスの成果等を有効活用できる情報基盤の本格的実現に向けた環境が整う。統合化の基盤技術を利用したポータルサイトを通じて、より多くの研究者等が DB や解析ツールの所在や内容について知り、これにより DB の利用者層が格段に拡大され、また、利用機会が増大することでライフサイエンス分野全体にプロジェクト研究の成果還元が進む。DB 統合化の活動や DB の維持管理に関して、その活動の広報やマニュアル化が進むことで、将来のより広い統合化へ向けた活動への協力や参加が期待できる。</p>

(様式3)

研究計画構想・詳細

(ライフサイエンス分野の統合データベース整備事業)

課題名 「ライフサイエンス統合データベース基盤整備」
研究代表者名 「堀田 凱樹」
責任機関名 「大学共同利用機関法人 情報・システム研究機構」

1. 研究内容

(1) ライフサイエンス(基礎生物学および応用生物学(医・薬・農学))、知識情報処理、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 DB 整備戦略について議論する。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)を構築する。また、この体制を利用して戦略立案の基盤となる情報の収集・分析を行う。

(2) 知識やデータの表現方法が統一されていない多くの DB を機能的にまた物理的に統合利用可能にするために必要な知識表現法、情報共有技術および文献からの知識抽出技術等を開発する。さらに医学分野に関しては多型情報および癌関連のゲノム情報や発現情報の DB 構築や統合 DB のプロトタイプ構築を通じて分子医学データの統合化に必要なデータ項目や利用モデルを検討する。

(3) 内閣府連携施策群「ライフサイエンス分野のデータベースの統合化に関する調査研究」において作成された国内外の DB 網羅リストおよび JST においてこれまで作成してきたポータルサイト WING-DB の詳細な内容説明を統合し、わかりやすく分類され、また、ライフサイエンスの各分野での利用に応じた説明がついた網羅性の高い DB のカタログを作成し、さらに JST で整理を進めてきた解析ツールなどの WEB リソースポータルを統合することで利用者を目的にかなった DB やリソースに誘導するポータルサイトを構築する。

2. ミッションステートメント

より多くの利用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ができる環境の実現のために

(1.1) ライフサイエンス(基礎生物学、薬学、医学、植物、微生物学) 知識情報処理、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、利用者の望む DB サービスの具現化の戦略を議論する。

(1.2) ライフサイエンス、知識情報処理、バイオ DB の研究動向に関する調査グループ(統合 DB 整備戦略作業部会)を組織し、ライフサイエンス研究の俯瞰、DB 化されている情報の俯瞰、および知識情報処理技術の動向につき上記委員会に報告する。

(2.1) 統合化知識表現法、情報共有技術および文献からの知識抽出技術等を開発する。

(3.1) DB や解析ツールの所在や利用法を網羅・分類・解説したポータルサイトを構築する。

(3.2) DB 構築・維持に不可欠な人員の教育マニュアルを作成し、本格統合化時の人材教育に備える。

3. 研究進展によるライフサイエンス分野の科学技術の発展への貢献度、経済社会への波及効果について

ゲノム研究やポストゲノム研究にとどまらず過去 20 年間の遺伝子中心の生物学の成果を理解し研究開発に利用する事はいまや統合されたデータベースなくしては不可能である。

本提案の実践により我が国の研究者等が既存の情報を十分に利用して研究を進めることのできる基盤作りが開始される。まず、我が国のライフサイエンス研究のための情報基盤整備の長期的な戦略を常に最新の基礎生物学・応用生物学、知識情報処理、ライフサイエンスDBに関する十分な情報に基づいて立案・評価・修正を行う組織の基礎が形成される。これにより本格的なデータベース統合利用に向けた有効で柔軟な戦略が立案可能になる。利用者であるライフサイエンス分野の動向と意見を的確にとらえ、最新の技術的な動向を考慮して立案されるDB統合化戦略は将来のライフサイエンス分野のDB構築やDB統合化などの情報基盤整備への研究投資をより効果的かつ効率的なものにする。さらに提案するポータルサイト構築活動はこれまでの主要なDBや解析ツールへの案内や調査を再利用かつ統合し、さらに利便性を高めたものであり、DB統合利用環境の整備完了をまつことなく、DBや解析ツール等の既存の情報資源へのより多くの研究者等のアクセスを誘導することで、ライフサイエンス分野全体の研究基盤の向上が期待できる。

4．提案にいたる準備・調査について

本提案は内閣府連携施策群「ライフサイエンス分野のデータベースの統合化に関する調査研究」および文部科学省ライフサイエンス委員会「データベース整備戦略作業部会」の主要な参加研究者らが議論を重ねて立案したものであり、双方の報告書の精神に従ったものになっている。それに加え、報告書に表現されていない多くのDB利用者や構築者の意見もこの提案の具体化には盛り込まれている。とくに統合に関するイメージや課題は本提案書に名前を連ねる研究者の主要メンバーが、平成13年度科学技術振興機構(JST)異分野研究者交流フォーラム「ゲノムと言語」、平成14年度ワークショップ「ゲノムと言語」、平成15年度異分野研究者交流促進事業領域探索プログラム「生命科学知識研究」において、所属する分野を超えて議論を重ね共有してきたものである。また本提案書のメンバーの一部は、JSTバイオインフォマティクス推進事業(BIRD)などの事業に直接的間接的に参画し、試験研究や要素技術開発を通じて本提案の実現にいたる基礎データや要素技術を開発してきた。

5．実施体制について

責任機関である情報・システム研究機構は全体の統括とサブテーマを遂行し、分担機関である科学技術振興機構(JST)はサブテーマ(3)を、九州大学はサブテーマ(2)の一部を分担する。責任機関である情報・システム研究機構は研究運営委員会およびサブテーマの円滑な実行に必要な環境と計算機資源を提供する。分担機関であるJSTは所有する一般利用者への情報の提供機構や広報システムをサブテーマ(3)に提供する。

責任機関である情報・システム研究機構は、欧米との協力によって運営される日本DNAデータベース(DDBJ)の維持運営にあたってきた国立遺伝学研究所及び我が国の情報科学の中核である国立情報学研究所を擁している。研究代表者堀田は機構長として、これらを統括するとともに、機構内部にライフサイエンスと情報科学研究の融合研究を推進するために新領域融合研究センターを設立しその運営に当たってきた。本研究課題はオールジャパンの体制による多様な融合研究が必要であるが、その実現のためには大学共同利用機関法人の特質を生かすことが最も有効であり、本研究機構がその担い手として最もふさわしいとの考えにより本提案をおこなうものである。参画機関であるJSTは2001年からのバイオインフォマティクス推進センター(BIRD)の運営および長年の日本語文献抄録検索システム(JOIS, JDream)の構築運営の実績がある。九州大学は遺伝情報実験施設においてヒト多型情報データの生産からQSNPDB構築を通じて、近年ゲノム医科学分野で重要性を増している多型情報の一般への公開を年余にわたり行っている。

6．研究運営委員会の構成について

実施機関以外に研究運営委員会に参加する機関（者）（案）

東京大学大学院新領域創成科学研究科	吉田光昭
京都大学化学研究所	金久實
大阪大学蛋白質研究所	中村春木
理化学研究所	榊 佳之
産業技術総合研究所	秋山 泰
東京理科大学	増保安彦
JT 生命誌研究館	中村桂子
DNA チップ研究所	松原謙一
かずさDNA研究所	田畑哲之
東京大学大学院情報理工学系研究科	辻井潤一
九州大学大学院 農学研究院	久原 哲
自然科学研究機構基礎生物学研究所 / JST	勝木元也
東京大学大学院新領域創成科学研究科 / JST	高木利久

7．国内外の研究状況について

米国では医学生物学の総合文献 DB (PubMed) を作成している国立医学図書館 (NIH NLM) 内のバイオ技術情報センター (NCBI) へ核酸配列バンク事業 (GenBank) が移管されたのを契機に統合化の動きがはじまった。その後、ゲノムプロジェクトの進行中に遺伝病電子辞典や遺伝子地図情報などが同センターに集合する形でさらなる統合化が進み、最近では化合物情報も取り込む形で統合化が行われている。これらは集中型の統合化と言えるものである。一方、欧州では、世界で最も古い核酸配列バンクを作ったことで知られる欧州分子生物学研究所 (EMBL) が設立した欧州バイオインフォマティクス研究所 (EBI) において分子情報を中心とした各種データの整理、連携、提供が進められている。これに加え、最近では高度な情報共有のための技術開発も進められるなど、分子情報を基盤とした分散型の統合化が行われている。

一方、国内では、国際的な研究サービス活動である DNA バンク事業や蛋白質構造バンク事業に欧米外から唯一参加する施設である国立遺伝学研究所の DDBJ や大阪大学の PDB J の活動は国際的にも意義深いものである。また京都大学の知識データベース KEGG は開始後 10 年を経て自ら切り開いた知識データベース分野で不動の地位を得ている。このようにバンク事業や個別のデータベースについては世界的なものも散見される一方で、DB 利用環境全体を見渡してみると、我が国には長期戦略が存在せず、そのため大型測定研究プロジェクトのデータさえ十分に有効活用できるようにはなっていない。これらはプロジェクト末期に独自の成果 DB として公開されるものの、他のデータへのリンク情報の付加などはプロジェクト毎に様々の精度や充実度や頻度で行われており、その成果は十分に理解・利用されずに陳腐化してゆく傾向にある。加えてデータベースは解析研究が高度になり精度を上げるに従って複雑さを増しているにもかかわらず論文中心の研究評価体制の下で構築者がわかりやすさや使いやすさに注力する動機付けはほとんど行われていない。以上の理由で個別データベースと違い研究社会の総合力が試されるデータベースに関する利便性の水準は利用者にとっては必ずしも高いものではない。

8．生命倫理・安全面への配慮について

該当するデータおよび試料を扱わない。

(様式4)

計画内容

(ライフサイエンス分野の統合データベース整備事業)

課題名 「ライフサイエンス統合データベース基盤整備」
研究代表者名 「堀田 凱樹」
責任機関名 「大学共同利用機関法人 情報・システム研究機構」

(1) データベース統合戦略立案および評価

サブテーマ責任者(所属機関・年齢): 堀田凱樹(情報・システム研究機構 67)

(研究内容)

目的

ライフサイエンスおよび知識情報処理の識者にライフサイエンス DB の専門家を加えた統合 DB 整備戦略作業部会を組織し、十分な情報の収集・分析を行い、それに基づいて統合化戦略を議論する。

目標

我が国のライフサイエンス研究のための情報基盤整備の長期的な戦略を常に最新の基礎生物学・応用生物学、知識情報処理、ライフサイエンス DB に関する十分な情報に基づいて立案・評価・修正を行う組織の基礎を形成する。これにより、平成19年度以降のライフサイエンス分野のデータベース統合利用環境の構築に資することを目指す。

研究の現状

DDBJ、PDBJ、KEGG などバンク事業や個別のデータベースについては世界的なものが散見される一方で、DB 利用環境全体を見渡してみると、我が国には長期戦略が存在せず、そのため大型測定研究プロジェクトのデータさえ十分に有効活用できるようにはなっていない。これらはプロジェクト末期に独自の成果 DB として公開されるものの、他の多くの DB との統合利用環境の構築が十分ではなく、その成果は十分に理解・利用されずに陳腐化して行く傾向にある。従って国内の利用者にとって国産データに関する利便性の水準は必ずしも高いものではない。

研究の内容

実施期間: 平成18年8月~平成19年3月

手法:

DB 統合化の戦略は1)利用者であるライフサイエンス分野の状況、2)素材である個別データベースの状況、3)利用可能な情報技術、の3つの動向を常に監視しながら継続的にかつ柔軟に立案されねばならない。しかしながらこれまでは1)から3)に関する十分な情報の入手や分析が十分ではない状況で主に DB 構築者のみによる合議で立案されることが多かった。ここでは3分野からの専門家を集めた研究運営委員会を組織し、同委員会に3分野の動向を報告する調査組織(統合 DB 整備戦略作業部会)を設けることで視野の広い情報に基づいて有識者が情報基盤整備の長期戦略や DB 統合化の具体的な方針について議論できる環境を整える。

担当機関・担当者

サブテーマ責任者(所属機関・年齢): 堀田凱樹(情報・システム研究機構 67)

項目 統合 DB 整備戦略作業部会

項目 統合データベース間の連携と課題の整理

研究参画者: 五條堀孝(国立遺伝学研究所 54)、藤山秋佐夫(国立情報学研究所 55)

研究協力者: 田畑哲之(かずさ DNA 研究所 51)、久原哲(九州大学 56)

項目 国内外の DB の俯瞰と質的量的比較

研究参画者：大久保公策（国立遺伝学研究所 46）
 項目 ライフサイエンス分野の俯瞰調査
 研究協力者：中村桂子（JT 生命誌研究館 70）
 項目 知識情報技術の動向調査
 研究参画者：高木利久（東京大学 / JST 52）、高野明彦（国立情報学研究所 48）、
 菅原秀明（国立遺伝学研究所 61）

期待される具体的な成果等

- ・ライフサイエンス（基礎生物学、薬学、医学、植物、微生物学）知識情報処理、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、利用者の望む DB サービスの具現化の戦略を議論する環境を整えられる。
- ・ライフサイエンス、知識情報処理、バイオ DB の研究動向に関する調査グループ（統合 DB 整備戦略作業部会）を組織し、ライフサイエンス研究の俯瞰、DB 化されている情報の俯瞰、および知識情報処理技術の動向を把握する環境が整えられる。

研究の必要性

分子生物学以降ライフサイエンスの成果の主体は、遺伝子や蛋白質に関するデータや言葉を使った生命現象の記述である。この膨大な成果はいまやデータベースの助けなく整理することも利用することも不可能である。従ってデータベース統合の効果はゲノム研究やポストゲノム研究に限定されずライフサイエンス全体に波及する。

項目 統合 DB 整備戦略作業部会

上記の1) から3) の各分野内でこれまで研究を行ってきた専門家集団からなる研究運営委員会を組織し、統合化 DB の整備戦略を議論する。また各分野から適宜代表的な利用者や潜在的利用者を集め DB 統合化戦略の状況につき説明し、また、意見をまとめ、DB 利用者との情報共有に努める。

研究運営委員会の想定メンバーは下記の通りである。(提案書提出時点では本人の了解は得ていない)

情報・システム研究機構	堀田凱樹 小原雄治 五條堀孝 大久保公策
自然科学研究機構基礎生物学研究所 / JST 東京大学 / JST 科学技術振興機構 (JST) 東京大学 東京大学 京都大学 大阪大学 理化学研究所 産業技術総合研究所 東京理科大学 JT 生命誌研究館 DNA チップ研究所 かずさDNA研究所 九州大学	勝木元也 高木利久 大倉克美 吉田光昭 辻井潤一 金久 實 中村春木 榊 佳之 秋山 泰 増保安彦 中村桂子 松原謙一 田畑哲之 久原 哲

上記の研究運営委員会の実働部隊（情報の収集・分析、動向調査、戦略立案支援、な

ど)として以下の統合 DB 整備戦略作業部会を設ける

「統合 DB 整備戦略作業部会」

五條堀孝 構築者委員、利用者委員(分子進化)、
DDBJ および他の統合 DB プロジェクトとの連携
田畑哲之 構築者委員、利用者委員(植物学)
大久保公策 構築者委員、利用者委員(医学)
菅原秀明 構築者委員、情報技術委員
高木利久 情報技術委員、構築者委員、BIRD プロジェクトとの連携
高野明彦 情報技術委員
黒田雅子 文献情報との連携
藤山秋佐夫 利用者委員(ゲノム)
久原哲 利用者委員(微生物学、農学)
中村桂子 利用者委員(基礎生物学)
増保安彦 利用者委員(創薬)

項目 統合データベース間の連携と課題の整理

研究参画者：五條堀孝(国立遺伝学研究所 54)、藤山秋佐夫(国立情報学研究所 55)

研究協力者：田畑哲之(かずさ DNA 研究所 51)、久原哲(九州大学 56)

現在わが国では DB の統合とプロジェクト成果の還元を目的にしている事業として文部科学省の「ゲノムネットワークプロジェクト」における「ヒトゲノムネットワーク情報システムの構築」および経済産業省が平成 12 年度からの「統合データベースプロジェクト」とその延長である平成 17 年度からの「ゲノム情報統合プロジェクト」さらに平成 18 年度開始の農林水産省「統合データベース」プロジェクトが存在する。これらはおもにゲノムへの生物学医学的注釈付け(ゲノムアノテーションと呼ぶ)を目的としたものであるが、本提案ではこれらのいずれとも重複のないように留意し、ゲノムアノテーションを通じたデータ整理については計画していないがこれらのゲノムアノテーション型の DB 統合はプロジェクトデータの出口として確立された重要なものである。いずれの事業も過去のデータ生産や整理プロジェクトの総和に匹敵する規模であり、本提案のミッションに照らすと早期からこれらの事業のデータ統合方針や予想される最終成果、事業遂行中に克服された課題や未決題について知ることは現実的に極めて重要な課題である。ここではこれらのプロジェクトの指導者も含めた最適な参画者たちによって国内外のゲノムアノテーション型のデータ統合を構造的機能的に比較し、それらをさらに統合し再利用する方策について調査し統合 DB 整備戦略作業部会および運営委員会に報告する。

項目 国内外の DB の俯瞰と質的量的比較

研究参画者：大久保公策(国立遺伝学研究所 46)

国内の維持管理統合の対象となる DB の現状を国外の状況とも対応付けながら知ることとは立案の要件である。ライフサイエンス領域の DB は国内でも 200 を超えると報告されているがどのようなものがどれだけ存在しているかに関する情報は皆無に等しく維持管理や統合のための立案が困難である。ここではデータベースを網羅し多次元分類した内閣府調査の結果とデータベースに関して JST において続けてきた専門家による DB 解説や分類活動を融合し DB の俯瞰をより豊かに与える DB カタログを作成する。さらに大型の DB に関しては内容まで深く踏み込んだ評価を行うための機能比較などの検討も行い、統合 DB 整備戦略作業部会および研究運営委員会に報告する。

項目 ライフサイエンス分野の俯瞰調査

研究協力者：中村桂子（JT 生命誌研究館 70）

利用者の研究動向を知り、現実的な利用モデルを知ることは DB 統合の立案に欠かせない要件である。さらに統合利用者からの意見を求めるためには様々な研究における活発なデータ利用者について知ることが重要である。ここでは国内の分子生物学会・生化学会・癌学会の学会要旨を利用して国内の研究動向を経時的にまとめ、サブテーマ2で開発する予定の索引系列による索引付けで構造化して統合 DB 整備戦略作業部会および研究運営委員会に我が国のライフサイエンス研究の俯瞰を提示する。

項目 知識情報技術の動向調査

研究参画者：高木利久（東京大学 / JST 52）、高野明彦（国立情報学研究所 48）、菅原秀明（国立遺伝学研究所 61）

あらゆる分野での DB 活動や相互運用による DB 統合化の流れはすべて情報処理分野での技術開発に対応する形でおこってきたものである。特に画像や文献などの多様でかつ膨大な情報を扱うライフサイエンス分野では情報処理技術の動向を早期に知りその将来性を見極めることは無駄のない統合戦略の立案に不可欠である。ここでは急速に展開する情報処理技術の動向その中でもとくに知識情報処理の動向を文献資料、ネットワーク、聞き取り調査、実地見聞などを駆使し調査し分析を加え、統合 DB 整備戦略作業部会および研究運営委員会に報告する。

（2）データベース統合化基盤技術開発

サブテーマ責任者（所属機関・年齢）：大久保公策（国立遺伝学研究所 46）

（研究内容）

目的

知識やデータの表現方法が統一されていない多くの DB を機能的にまた物理的に統合利用可能にする。

目標

統合化に不可欠な知識表現法、情報共有技術および文献からの知識抽出技術等を開発し、戦略立案のための調査（1）やポータルでの利用者誘導（3）に利用する。これにより、平成19年度以降のライフサイエンス分野のデータベース統合利用環境の構築に資することを目指す。

研究の現状

ゲノムにさまざまな分子データを貼り付け、そこに生物学医学的な注釈を加えるタイプの統合化（物理的な統合）に関しては研究が進んだが、生命機能を中心とした DB の統合化はまだまだ研究が緒にすぎたばかりである。また、依然として知識やデータの表現方法が統一されていない多くの DB が存在する。機能的にまた物理的に統合利用可能にするためには知識表現法、情報共有技術および文献からの知識抽出技術等の基盤技術を開発する必要がある。

研究の内容

実施期間：平成18年8月～平成19年3月

手法：

知識やデータの表現方法が統一されていない多くの DB を機能的にまた物理的に統合利用可能にするために必要な知識表現法、情報共有技術および文献からの知識抽出技術等を開発する。さらに統合化の一つの試みとして、とくに重要性必要性の高い医学分野をとりあげ、多型情報および癌関連のゲノム情報や発現情報を統合する DB 構築や統合 DB のプロトタイプ構築を通じて分子医学 DB の統合化に必要なデータ項目や利用モデルを検討する。

担当機関・担当者

サブテーマ責任者（所属機関・年齢）：大久保公策（国立遺伝学研究所 46）

項目 DB 俯瞰化のための索引技術開発

研究参画者：大久保公策（国立遺伝学研究所 46）

項目 基盤知識表現技術開発

研究参画者：高木利久（東京大学 52）、大久保公策（国立遺伝学研究所 46）、川本祥子（新領域融合研究センター 41）

研究協力者：藤淵航（産業技術総合研究所 38）、新規雇用研究員 A

項目 癌研究知識表現技術開発

研究協力者：加藤菊也（大阪府立成人病センター 50）

項目 多型知識表現技術開発

研究参画者：林健志（九州大学 62）

項目 キュレーション援助技術開発

研究参画者 武田英明（国立情報学研究所 42）、西川建（国立遺伝学研究所 62）、高野明彦（国立情報学研究所 48）、新規雇用研究員 B

期待される具体的な成果等

分子生物学以降ライフサイエンスの成果の主体は、遺伝子や蛋白質に関するデータや言葉を使った生命現象の記述である。この膨大な成果はいまやデータベースの助けなく整理することも利用することも不可能である。従ってデータベース統合の効果はゲノム研究やポストゲノム研究に限定されずライフサイエンス全体に波及する。

研究の必要性

データベースを統合し判りやすく利用者に提示するためには地道な専門職の努力（サブテーマの（3）項目 で対応）が緊急に必要とされているが、それだけではデータの増大に持続的に対応することは不可能である。従って応用分野に特化した知識の整理法を開発するとともに最先端の情報共有技術や知識情報技術【判り易くする技術】を積極的にライフサイエンス DB の分野に持ち込んで、また、それらをライフサイエンス分野に適用可能にするための技術開発を行い、将来のデータ増大にも対応できる準備を開始する必要がある。

項目 DB 俯瞰化のための索引技術開発

研究参画者：大久保公策（国立遺伝学研究所 46）

様々な DB 内で管理されている塩基配列、蛋白質構造、遺伝子発現などの多様な測定データの内容（実験手法やどういうプロジェクトから得られたかなど）について知ることが利用の第一歩である。これまでのこれらの DB を横断的に検索することはできたが、その場合用語や ID による検索を受け付けるのみであり検索されてこなかったデータにつき全く情報を与えない。一方、論文投稿と連動してデータ投稿を受け付けるバンク型データベースには年々測定型研究のデータが蓄積しているが、その内容は多様化している。ここでは配列（DDBJ）蛋白質構造（PDBJ）多型（dbSNP）発現（GEO CIBEX）の 4 種のデータバンクのレコードを研究プロジェクト単位に再編し、その内容や規模を俯瞰するための用語辞書を開発する。これを索引に利用することで、4 種のバンクにどのような研究由来のデータが存在するのかを示す「データ目次」にあたるものを開発する。そしてこれをサブテーマ（1）戦略立案および（3）ポータルサイト構築に提供する。

項目 基盤知識表現技術開発

研究参画者：高木利久（東京大学 52）、大久保公策（国立遺伝学研究所 46）、川本祥子（新領域融合研究センター 41）

研究協力者：藤淵航（産業技術総合研究所 38）、新規雇用研究員 A

ライフサイエンスではあらゆる記述(データや文献内容)は、どの生物の、どの臓器・細胞の、どの状態の、どの疾患のといった文脈と独立には存在しない。データ検索時にもこれらの文脈概念が利用できなければ意味のある検索は行えない。しかしながら解剖名称や細胞名称などの分野共通の基盤的な概念に関しても概念の整理と索引への利用は進んでいない。ここでは応用分野の区別無く重要な基礎的文脈の表現に用いられる蛋白質名称、動物・植物の解剖用語、細胞名称と細胞機能および実験手法を対象に文脈情報を制度よく表現する索引系を開発し、ポータルサイト構築(サブテーマの(3))や上記項目 に提供する。

項目 癌研究知識表現技術開発

研究協力者：加藤菊也(大阪府立成人病センター 50)

癌は他の疾患とは異なり、体細胞レベルでの遺伝情報(染色体異常、癌関連遺伝子の突然変異)が分子データの中心になっている。また、オーダーメイド化が最も進んだ疾患領域であるため、臨床情報は他の疾患よりも遥かに詳細多量である。体細胞レベルでの遺伝情報のDB化は個々の癌研究者によって様々な規模で行われているが、全体を統合したものはまだない。しかしながら癌組織のヒトゲノム再配列解析プロジェクトが欧米で進行しており、そのサテライトとして統合DB化が検討されている。臨床情報については、未だDB化に伴う倫理問題が議論されている段階である。そこで本プロジェクトでは大規模再配列解析プロジェクトの成果を念頭に置いて、すべての体細胞レベルでの遺伝情報と臨床情報を統合化したDBの作成において必要な表現法やデータ整理法について検討する。また、いくつかの実験的なDB構築を通じたフィジビリティスタディを通じて一般の基礎臨床研究者にとっての癌統合DBのあり方を検討する。

項目 多型知識表現技術開発

研究参画者：林健志(九州大学 62)

研究参画者の林は、これまで多型情報のDB化として、日本人ゲノムのプロモーター領域に存在するSNPを正確な頻度情報とともに記載したデータベース「dbQSNP」および日本人確定ハプロタイプ情報を記載した「D-HaploDB」の構築に取り組んできた。本プロジェクトではこれらをさらに拡充するとともに、これを統合し日本人試料を用いた関連解析による疾患原因遺伝子の探索に不可欠な、高精度の情報基盤を確立することを目指す。また、XML化等の標準化にも取り組む。これにより、ヒトゲノム多型DBの標準化、統合DBへの組み込みのフィジビリティスタディを行い、多型情報統合DBの医学分野におけるあり方を検討する。

項目 キュレーション援助技術開発

研究参画者 武田英明(国立情報学研究所 42)、西川建(国立遺伝学研究所 62)、高野明彦(国立情報学研究所 48)、新規雇用研究員 B

ライフサイエンスの文献は情報検索システム(例えばPubMed)によって容易に検索することはできるが、英語などの自然言語で書かれているため、計算機では多数の文献の内容を理解したり利用したりすることは非常に困難である。そこで文献情報の統合的な利用は、すなわち文献からの情報抽出や要素間の関係把握は、主にDBキュレーターと呼ばれる研究支援者によってマニュアルで行われているのが現状である。その作業は文献に書かれた内容をテーブルなどの定型化された知識表現に落とし込むことである。これにより文献に書かれた内容が再利用可能となる。現在キュレーターの業務は印刷物とラインマーカー、キャビネットを使って行われている。文献の読解以降の整理管理はキュレーターにとって単調な作業であり、また一方で紙の上に保存

されたマーカー情報は利用者が利用する方法はなく有効に活用されていない。人手による関心領域の記録や関心領域の内容の分類やDB化を目的にキュレーターの論文読解から始まる一連の作業を行う環境をクライアントサーバとして構築し、作業環境の標準化を通じて貴重なキュレーターによる作業の高効率化や機械的補助の技術を開発する。

(3) ポータルサイト構築

サブテーマ責任者(所属機関・年齢): 勝木 元也(基礎生物学研究所 / JST 63歳)

(研究内容)

目的

利用者を目的にかなったDBやリソースに誘導するポータルサイトを構築する。

目標

DBや解析ツールの所在や利用法を網羅・分類・解説したポータルサイトを構築する。これによりライフサイエンス分野におけるプロジェクト研究の成果の活用やDB利用を促進する。

研究の現状

現存するDBのカタログは遺伝子や蛋白質等のデータ対象によるおおまかな分類が与えられるのみで利用者が目的にかなったDBを選択することは容易でない。

研究の内容

実施期間: 平成18年8月~平成19年3月

手法:

- ・ 内閣府連携施策群「ライフサイエンス分野のデータベースの統合化に関する調査研究」において作成された国内外のDB網羅リストおよびJSTにおいてこれまで作成してきたWINGDBの詳細な内容説明を統合する。これをさらに整理拡充し、わかりやすく分類され、ライフサイエンスの各分野の利用に応じた説明のついた網羅性の高い国内外のDBのカタログを作成する。さらにJSTで整理を進めてきた解析ツールなどのWEBリソースポータルをも統合し、DBや解析ツールに関する統合ポータルサイトを構築する。
- ・ さらにバンク型のDBのデータ内容を利用者の視点に立って分類した“ミレニアムデータの目次”を作成する。
- ・ DBに関するニュースを提供し、また、さまざまな観点のキーワードでのDBやWEBリソースの検索機能を設けることにより、さらにライフサイエンス関連のサイトと連携することによりライフサイエンス全般の幅広い利用にむけたポータルサイトを構築する。
- ・ さらに、文献情報との連携や日本語での検索機能実現のため調査を行い、結果をとりまとめる。
- ・ ポータルサイトの公開およびこの「ライフサイエンス分野の統合データベース整備事業」の広報、普及のため、広報活動を行う。
- ・ 統合データベース整備の普及啓発の一環として、DBの統合や維持管理において必要な人材の啓発のための教材を作成する。

担当機関・担当者

サブテーマ責任者: 基礎生物学研究所 / JST 勝木元也 63

項目 データベース(DB)ポータル構築

研究参画者: 黒田雅子(科学技術振興機構 52)

研究協力者: 坊農秀雅(埼玉医科大学 33)、川島秀一(東京大学 35)

項目 WEBリソースポータル構築

研究協力者: 伊藤武彦(三菱総合研究所 36)

- 荒木次郎（三菱総合研究所 36）、片山俊明（東京大学 34）
- 項目 ポータルサイト連携
研究参画者：藤山秋佐夫（国立情報学研究所 55）、川本祥子（新領域融合研究センター 41）、大倉克美（科学技術振興機構 53）、黒田雅子（科学技術振興機構 52）
- 項目 文献情報との連携調査
研究参画者：大倉克美（科学技術振興機構 53）、黒田雅子（科学技術振興機構 52）
- 項目 広報・教育普及
研究参画者：大久保公策（国立遺伝学研究所 46）、黒田雅子（科学技術振興機構 52）、新規雇用研究員 C
研究協力者：中村保一（かずさ DNA 研究所 40）、坊農秀雅（埼玉医科大学 33）

期待される具体的成果

より多くの研究者等が DB や解析ツールの所在や内容について知り、これにより DB の利用者層が格段に拡大され、また、利用機会が増大することでライフサイエンス分野全体にプロジェクト研究の成果還元が進む。DB 統合化の活動や DB の維持管理に関して、その活動の広報やマニュアル化が進むことで、将来のより広い統合化へ向けた活動への協力や参加が期待できる。

研究の必要性

我が国および世界で数多くのデータベースが構築されており、公開されている。これらを十分に活用するための、最新の情報を保持したポータルサイトが不可欠である。

我が国で構築提供される個別のデータベースについては世界的なものが見られる一方で、DB 利用環境全体を見渡してみると、我が国には長期戦略が存在せず、そのため大型測定研究プロジェクトのデータさえ十分に有効活用できるようにはなっていない。これらはプロジェクト末期に独自の成果 DB として公開されるものの、他の多くの DB と統合利用できるようにはなっておらず、その成果は十分に理解・利用されずに陳腐化して行く傾向にある。従って国内の利用者にとって国産データに関する利便性の水準は必ずしも高いものではない。

- 項目 データベース（DB）ポータル構築
研究参画者：黒田雅子（科学技術振興機構 52）
研究協力者：坊農秀雅（埼玉医科大学 33）、川島秀一（東京大学 35）

DB 統合の第一段階は網羅性の高いカタログの作成と内容による DB の分類や解説による利用者の DB への案内である。現存する DB のカタログは遺伝子や蛋白質等のデータ対象によるおおまかな分類が与えられるのみで利用者が目的にかなった DB を選択することは容易でない。ここではサブテーマ（1）や（2）において DB の俯瞰や戦略立案の目的で作成する DB ディレクトリおよび高度な DB 分類に JST がこれまで構築維持してきた WINGDB 案内における利用者向けの日本語解説などを融合し我が国の利用者を最適な DB に案内する機構を構築する。

- 項目 WEB リソースポータル構築
研究協力者：伊藤武彦（三菱総合研究所 36）、荒木次郎（三菱総合研究所 36）、片山俊明（東京大学 34）

DB に加えてバイオインフォマティクスでは解析のためのツールや環境を WEB 上に公開し一般に提供している例が多い。DB 中のデータを利用するためにはこれらのリソースもあわせて利用することが多いが、ネット上にどのような目的に利用できるどのようなツールが存在するのかは必ずしも容易に知ることはできない。ここでは JS T でこれまで独自に進めてきた WEB リソースカタログ作りを增強し、サブテーマ

(2)で開発する索引付けに従った索引を行いDBポータルと統合して利用することの可能なWEBリソースポータルを構築する。

項目 ポータルサイト連携

研究参画者：藤山秋佐夫（国立情報学研究所 55）川本祥子（新領域融合研究センター 41）大倉克美（科学技術振興機構 53）黒田雅子（科学技術振興機構 52）

上に述べたDBポータルやWEBリソースポータルに必ずしも該当するものではないが、これに関連するサイトとして研究課題や重要な発見についての文部科学省のライフサイエンスポータルや科学コミュニケーションに主眼をおいた情報・システム機構のJABIONなどが我が国には存在する。これらのライフサイエンスのポータルは相互にリンクするとどまり、サイトの構築が全体的にコーディネートされているわけではない。ここではこれらのそれぞれ内容が異なるポータルサイトの機能を生かし、相互に利用すべき部分では利用し、全体としてより高い機能を果たすべく連携する仕組みを考案する。

項目 文献情報との連携調査

研究参画者：大倉克美（科学技術振興機構 53）黒田雅子（科学技術振興機構 52）

これまでの科学研究におけるDBでは遺伝子機能にかかわる記述や参考文献情報も英語のみで提供されその内容は限定されてきた。一方でJSTが開発運営を続けている文献情報提供事業では国内外の情報が日本語で幅広く提供されている。これを活用し、遺伝子機能に関連する情報について、DBと文献情報との連携および日本語化するための方策について検討する。すでにそのための調査に着手しているので、その結果をとりまとめる。

項目 広報・教育普及

研究参画者：大久保公策（国立遺伝学研究所 46）黒田雅子（科学技術振興機構 52）新規雇用研究員C
研究協力者：中村保一（かずさDNA研究所 40）坊農秀雅（埼玉医科大学 33）

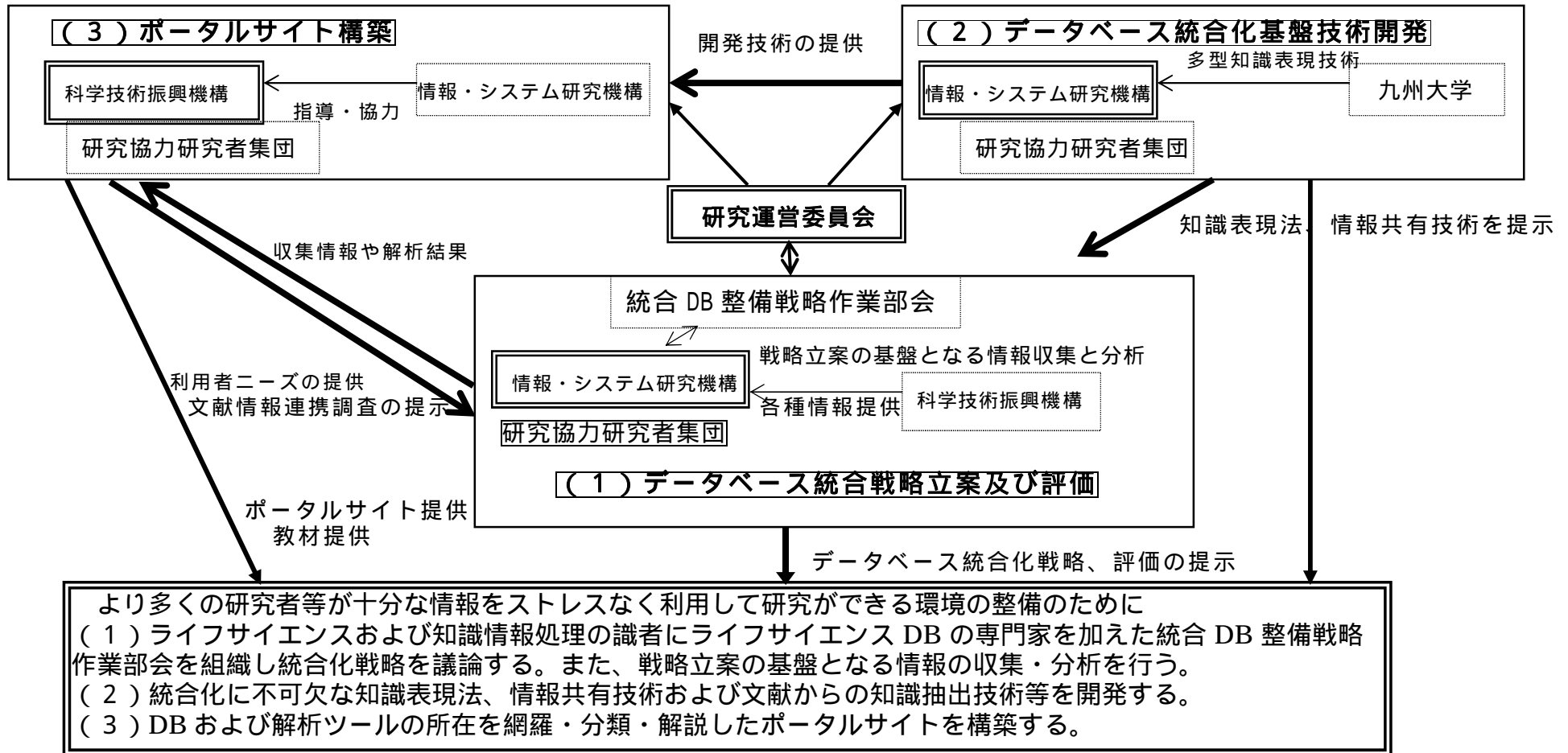
DB統合や維持管理のためには(1)キュレーター(論文の内容を理解して情報抽出整理し定型化した表現に変換してDB構築を支援する学芸員)(2)アノテーター(プログラム処理の結果に総合判断を加え、データに生物学的医学的な解釈を付与する学芸員)(3)DBマネージャー(DBについて理解しておりデータの参照情報などを自立的に更新できる技術者)の3者が必要である。しかしながらこれらの専門業務の存在はDB構築運営を行ってきた組織でしか知られておらず、広くこれらの業務内容について衆知し、自習の材料を与えることは将来のDB統合事業に参加可能な人材の裾野を広げることであり、これらの専門職のキャリアパス作成の第一段階である。ここでは3者の業務内容を区別し、業務に必要な基礎知識や技術について解説した教材を作成しDBの本格的統合化事業に備える。また、統合DBの重要性に関して学会などで広報活動を行うとともに、統合化のあり方に対する意見収集に努める。

(様式5)

研究体制図

(ライフサイエンス分野の統合データベース整備事業)

課題名 「ライフサイエンス統合データベース基盤整備」
研究代表者名 「堀田 凱樹」
責任機関名 「大学共同利用機関法人 情報・システム研究機構」



(様式6)

所要経費の見込額

(ライフサイエンス分野の統合データベース整備事業)

課題名 「ライフサイエンス統合データベース基盤整備」
 研究代表者名 「堀田 凱樹」
 責任機関名 「大学共同利用機関法人 情報・システム研究機構」

(単位：百万円)

経費の内容	年 度	総 額
	平成18年度	
情報・システム研究機構	141	141
備品及び試作品費	0	0
消耗品	0	0
電子計算機諸費	118	118
(内容)	プログラム作成費等(以下内訳)	
	(1)データベース統合戦略立案および評価 37	
	(2)データベース統合化基盤技術開発 59	
	(3)ポータルサイト構築 一部 22	
その他	10	10
(内容)	会議開催費・役務費(事務局)	
旅費	1	1
(内容)	研究参画/協力研究者及び委員の交通費	
人件費	12	12
	新規雇用研究員3名他 アルバイト等	
科学技術振興機構	24	24
消耗品	0	0
電子計算機諸費	24	24
(内容)	プログラム作成費等(以下内訳)	
	ポータルサイト構築	
その他	0	0
旅費	0	0
人件費	0	0
九州大学	15	15
備品及び試作品費	0	0
消耗品	4	4
(内容)	試薬その他	
電子計算機諸費	10	10
(内容)	プログラム作成費等(以下内訳)	
	データベース統合化基盤技術開発	
旅費	1	1
(内容)	国際学会発表・情報収集	
人件費	0	0
	0	0
	180	180

(様式7・別紙2)

参画研究者リスト

(ライフサイエンス分野の統合データベース整備事業)

課題名 「 ライフサイエンス統合データベース基盤整備 」
研究代表者名 「 堀田 凱樹 」
責任機関名 「 大学共同利用機関法人 情報・システム研究機構 」

氏名	所属機関名	職名	提案課題における役割
堀田 凱樹	大学共同利用機関法人 情報・システム研究機構	機構長	研究代表者 運営委員会委員長 サブテーマ1 責任者
小原 雄治	情報・システム研究機構 国立遺伝学研究所	理事・所長	運営委員会委員
五條堀 孝	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	教授・センター長	運営委員会委員 サブテーマ1 参画者
吉田 光昭	東京大学大学院新領域創成科学研究科	客員教授	運営委員会委員
辻井 潤一	東京大学大学院 情報学環	教授	運営委員会委員
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	教授・センター長	運営委員会委員
中村 春木	大阪大学 蛋白質研究所附属プロテオミクス総合研究センター	教授	運営委員会委員
秋山 泰	独立行政法人産業技術総合研究所 生命情報研究センター (CBRC)	センター長	運営委員会委員
久原 哲	九州大学大学院 農学研究院 遺伝子資源工学部門	教授	運営委員会委員 サブテーマ1 協力者
松原 謙一	(株)DNAチップ研究所	代表取締役	運営委員会委員
榊 佳之	理化学研究所 ゲノム科学総合研究センター	センター長	運営委員会委員
田畑 哲之	(財)かずさDNA研究所 植物ゲノム基盤情報部	部長	運営委員会委員 サブテーマ1 協力者
中村 桂子	JT生命誌研究館	館長	運営委員会委員 サブテーマ1 協力者

勝木 元也	自然科学研究機構基礎生物学研究所 (独)科学技術振興機構バイオインフォマティクス推進センター	理事・所長 統括	運営委員会委員、サブテーマ3責任者
大倉 克美	(独)科学技術振興機構 研究基盤情報部	部長	運営委員会委員 サブテーマ3参画者
黒田 雅子	(独)科学技術振興機構 研究基盤情報部 バイオインフォマティクス課	課長	サブテーマ1参画者 サブテーマ3参画者
川本 祥子	情報・システム研究機構 新領域融合研究センター	融合プロジェクト研究員(特任助教授)	サブテーマ2参画者 サブテーマ3参画者
藤山 秋佐夫	情報・システム研究機構 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1参画者 サブテーマ3参画者
片山 俊明	東京大学 医科学研究所	助手	サブテーマ3協力者
坊農 秀雅	埼玉医科大学ゲノム医学研究センター	助教授	サブテーマ3協力者
川島 秀一	東京大学 医科学研究所	助手	サブテーマ3協力者
中村 保一	(財)かずさDNA研究所 植物ゲノム基盤情報部 植物ゲノム情報研究室	室長	サブテーマ3協力者
伊藤 武彦	株式会社 三菱総合研究所 先端科学研究センター 生命科学研究グループ	主任研究員	サブテーマ3協力者
荒木 次郎	株式会社 三菱総合研究所 先端科学研究センター	主任研究員	サブテーマ3協力者
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	教授	サブテーマ1参画者
高木 利久	東京大学大学院新領域創成科学研究科 (独)科学技術振興機構バイオインフォマティクス推進センター	教授 副統括	運営委員会委員 サブテーマ1参画者、サブテーマ2参画者
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	教授	運営委員会委員 サブテーマ1参画者、サブテーマ2責任者、サブテーマ3参画者、
西川 建	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	教授	サブテーマ2参画者

高野 明彦	情報・システム研究機構 国立情報学研究所 コンテンツ科学研究系	教授	サブテーマ1 参画者、サブテーマ2 参画者
武田 英明	情報・システム研究機構 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 参画者
水田 洋子	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	研究員	サブテーマ2 参画者（雇用予定者）
三橋 信孝	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	研究員	サブテーマ2 参画者（雇用予定者）
加藤 菊也	大阪府立成人病センター 研究所	所長	サブテーマ2 協力者
林 健志	九州大学 遺伝情報実験センター	教授	サブテーマ2 参画者
藤渊 航	独立行政法人産業技術総合研究所 生命情報研究センター（CBRC）	研究員	サブテーマ2 協力者
増保 安彦	東京理科大学 薬学部	教授	運営委員会委員 サブテーマ1 協力者

- 1 研究代表者、サブテーマ責任者だけでなく、すべての研究参画者について記載して下さい。また、研究運営委員会の委員予定者についても記載して下さい。
- 2 所属機関が大学の場合は、学部・研究科、研究所又は研究センターまで記載して下さい。その他の研究機関の場合は、研究所、研究部門まで記載して下さい。