

「統合データベースプロジェクト」
第11回研究運営委員会 議事要旨

【日時】 平成22年11月12日（金）14：00～17：40

【場所】 情報・システム研究機構事務局会議室

【出席者】 勝木委員、白木澤委員、豊田委員、長洲委員、中村春木委員、松原委員長、湯元委員、吉田輝彦委員、吉田光昭委員、大久保委員、五條堀委員、高木委員、堀田副委員長

【陪席】

厚生労働省 : 高畑専門官

農林水産省 : 大塚専門官

経済産業省 : 鈴木係長

文部科学省 : 釜井補佐、田中調査員

（独）科学技術振興機構 : 内田副調査役、堀尾主査、酒井主任調査員

ライフサイエンス統合データベースセンター : 永井特任教授、川本特任准教授、坊農特任准教授、箕輪特任准教授、畠中特任准教授、吉羽特任研究員

【事務局】 呉事務局長、加藤企画課長

【挨拶】 松原委員長より簡単な挨拶があり、開会が宣言された。また、議事に入る前に運営委員の変更についてのお知らせがあった。内容はJST大倉委員から白木澤委員へ、また、高木委員の所属変更である。

【議事】

1. 研究運営委員会（第10回）議事要旨について

各委員へ事前送付されている4月開催の第10回研究運営委員会議事要旨（資料1）の確認の依頼があった。修正が必要な場合は近日中に連絡いただきたい。

2. プロジェクトの平成22年9月末時点の達成状況

松原委員長からプロジェクトの達成状況として10月21日と11月2日に開催された作業部会分科会で各参加機関が集まって議論した内容の説明が高木委員に求められた。

本プロジェクトは来年3月で終わるため、今年の9月までに開発やデータベース（以下、DB）の公開作業を終え、残り半年で評価をいただいてフィードバックするという段取りになっており、9月末の時点で成果を取りまとめた。各機関からの状況報告が資料2-2で、それをまとめたものが資料2-1である。各機関の成果のポイントについて資料2-2を用いて説明する。（以下、担当機関名を太字下線で示す）

中核機関であるライフサイエンス統合データベースセンター（以下、DBCLS）では多岐にわたるサービスとその開発を実施しているが、数値的な達成度としてはカタログづくり912件（目標850件）、横断検索で検索可能なDBが270件（目標250件）、DBの受入については現在38件（目標45件、年度末までには達成予定）となっている。内容についてはのちほどデモを行う。サービスへのアクセス数については、統合プロジェクトのポータルサイトには5万人弱/月、100万ページ程度/月で、個別のサービスについては資料をご覧ください。JSTではWINGproというWiki型のDBのカタログやメタデータのエレメントのアーカイブを実施している。かずさDNA研究所では植物関連のDB、特にデータのアノテーションに注力して開発し、研究者の情報交換ネットワークをもちいて新規のゲノムのアノテーションに成果を上げている。産総研の生命情報工学研究センター（CBRC）ではデータを解析する際のプログラムを連携させる仕組み（ワークフロー）の技術開発を行った。奈良先端科学技術大学院大学ではプロジェクトで扱う日本語のドキュメント解析に必要な辞書を構築するためのソフトウェアを開発している。エンドユーザー向けではないため使用実績はないが、検索を高度化するための技術開発である。九州大学ではゲノムワイド関連解析（GWAS）のデータの品質保証のためのパイプラインを構築し、これまで4件の公開データについて解析を実施した。東京大学で実施された人材育成は自前でデータベースをつくるための教育で、次世代型のシーケンサーのデータ解析といった難しい問題も扱っている。実際の講習を受けた人数は多くはないが、公開されている教材の閲

覧数は多い。長浜バイオ大学ではDBのコンテンツをつくる人材の育成として、長浜バイオ大学の学生への教育に加えて出張講習会を実施したり教材を公開したりして学外の教育活動も実施している。また、退職された先生方と学生が組んでDBをつくるなどユニークな試みを行っている。お茶の水女子大ではDBをどう活用するかという教育を目的として、結婚などで一度退職した方など社会人をターゲットにして基本的なコンピューターや生物に関する知識から始めて、そのデータの解析までを教える教育を実施している。東大や長浜と同様に教材をネット上で公開しているので受講者以外の方の教育にも役立っている。京都大学ではKEGGに追加する形で化合物、特に医薬品のDBを充実させ構造検索を実現することがミッションで、このDBはKEGGのいろいろな機能と連携して使えるようになっている。東京医科歯科大学では医学、臨床DBの一つのモデルケースとしてがんに焦点を当て、臨床データから分子データまでの統合をすすめた。さらに来年3月までには倫理規定のガイドライン(案)を作成することになっている。また、本プロジェクトで医科歯科と連携している大阪大学では周辺の複数の病院と連携して多施設のデータをうまく統合するDBのモデルとしてパーキンソン病DBを構築している。さらに、大阪大学と医科歯科の2つの疾患データ(パーキンソン病とがん)を横断的に検索できるDBの構築のトライアルも実施している。東京大学医学部を中心とした取り組みではGWS、コピー数変異とリシークエンスデータのDBを構築しているが、一般公開可能なデータや数段階のアクセスコントロール付きのデータを内包したDBになっている。理研では①植物関係のシロイヌナズナを中心としたDB、②タンパク質関係の様々な情報の公開を担当し、理研ScieSから公開されている。産総研の糖鎖医学工学研究センターは日本の糖鎖関係のデータのハブになっており、非常に多くのDBが構築された。プロジェクト後はアジアのハブになるようなDBに育てていくという方向性である。遺伝研では新型シーケンサーからの出力データの受け皿となるシステム(シーケンス・リード・アーカイブ)を国際連携のもとにつくることが目的で順調にデータが蓄積されつつある。また、波形データのレポジトリ(トレース・アーカイブ)についても完全長cDNAのデータが全て格納されるなど順調である。九州工業大学では以前から構築されてきた熱力学のデータを論文から収集した世界的にも例の無いDBをプロジェクトでも継続して構築してきたが、人手の作業を一部自動化するために、読むべき論文の絞り込みにDBCLSのテキスト処理のソフトウェアを使うなどの省力化作業が進められている。それぞれのサービス等へのアクセスの状況は資料2-2を各機関の資料をご参照いただきたい。

このうち、DBCLSが提供するサービスの一部についてデモをご覧に入れる。内容としては4年間の本事業の目的のうちの4項目、ポータルサイトの整備、コンテンツ作成、日本語での情報提供、そしてDBの受け入れ(アーカイブ)、ならびに広報・事業関係のコンテンツについて、順に説明する。(以下、サービス名称を太字で示す)

統合DBプロジェクトのトップページからは中核機関だけでなく、すべての参画機関の成果を公開している。現在は成果物を、ポータル、アーカイブといったカテゴリーに分けて配置している。サービス以外にも、本運営委員会など委員会の記録やDBに関する科学政策の資料集などもある。データベースカタログでは国内を中心に公開されているDBの情報を集め、現在約900件の情報を公開、そのうち国内のDBは約500件である。カテゴリー分類されており対象とする分子や生物種、DBを構築している組織などで選択することが可能である。各DBの稼働状況もわかるようになっている。これらのDBの中から有用かつ検索に適したものを一括して検索できるようにしたものが横断検索である。検索結果はカタログ同様分類から絞り込める。また、自動で日英翻訳を行い、日本語で入力しても同時に英語のDBも検索できる。検索は経産省の統合DBプロジェクトと連携しており、省庁が分かれていても一つの窓からすべて検索できるようになっている。現在、農林水産省の統合プロジェクトとも連携のための調整作業が進行中である。横断検索ではそれぞれのエントリー間にリンクはないが、エントリー間のIDに相互リンクをつけたものが、TogoProtである。これは蛋白質の情報をキーワード検索した際に、ファミリー単位でさまざまなDBの結果を見られるもので、多種多様な情報を一括して閲覧できる。横断検索に比べより統合化された検索に向けたサービスであり、今後さらに開発を進めていく予定である。次に、オリジナルコンテンツを2つご紹介する。1つめは統合TVで、研究者、特にバイオインフォマティクスが専門ではない研究者が多種多様なDBやツールを使いこなすためにつくられた動画コンテンツである。また、講習会やシンポジウムなどを遠隔地の方々にもごらんいただけるようにライブ配信とアーカイブに努めており、アクセス数上も人気が高い。プロジェクトの成果のヘルプとしても活用しており、YouTubeやiTunesなどからも配信している。もう一つのオリジナルコンテンツとしてライフサイエンス新着論文レビューというサービスを今年度開始した。これはトップジャーナルに掲載された論文の著者自身がレビューを書くというもので、オリジナル論文が発表されて2週

間以内に読める。クリエイティブ・コモンズという著作権に関する世界共通のライセンスにより文章や図の引用改題を自由に行えるという特徴もある。つぎにアーカイブではDBの保全や公開、共有のために国内の研究者から寄託されたDBを一括して公開しており、生命科学分野では初めてのDB専門のアーカイブである。DBの説明やデータ項目の説明といったメタデータを共通のフォーマットでつけて現在38件が一括公開されている。利用者側にはいろいろなDBを同じフォーマットでダウンロードして利用でき、寄託者側には維持管理を任せられる、利用がより多くなるといったメリットが挙げられる。クリエイティブ・コモンズ・ライセンスを利用した利用許諾により利用される側も利用する側も安心して使えるという仕組みも設けている。このアーカイブと前半に紹介したデータベースカタログ、横断検索は、今後JSTの新センターでサービスが継続される。

これらのサービスに対して役に立つか/立たないかの評価と自由なコメントをいただくユーザ評価を実施しているが詳細は議事6で説明する。最後に、本日の運営委員会を含め当プロジェクトの会議等の記録は、全文の検索も可能な形で公開している。ご利用いただきたい。

●ポータル英語でのカバー率について説明していただきたい。

→日本語でつくったサービスも英語化を進めており、トップページの右上から英語サイトにリンクしている。日本語サイトで40～50サービスがあるうちの10ぐらい(DBのカタログやDDBJの目次、その他)が英語化されているが、現在ほとんど6割以上が国内からのアクセスで、確実に国外からとわかるものが5%、また、ドメイン別ではアカデミックのアドレスからは10%、また、企業からは4%であるが、多数を占める未分類のものについても今後詳しい分析をすすめる。上位のドメインにはライフサイエンス関連の企業からのアクセスが見られる。

●最後の評価アンケートには英語もあるのか。

→現在、日本の方だけを対象にしているため日本語のみである。

●日本にいる外国人とか、例えばアジアなど外国からの意見というのも結構大事ではないかなと思うので、英語版の開発もお願いしたい。

3. JST新センターの体制・方針

松原委員長から白木澤委員にJSTの新しいセンターの体制等についての説明が求められた。

資料3「バイオサイエンスデータベースセンター設置に向けた検討状況」を用いてご説明する。次年度からの統合データベースセンター設置の背景と目的としては、多くの府省で分散して提供されているDBを有機的に統合し、ユーザの立場に立った使い勝手のよいシステムを構築し、DBに蓄積された膨大な情報を利用者にも効率的に提供することによってライフサイエンス分野の知識発見を支援するということである。その検討の経緯については平成12年に「ゲノム情報科学におけるわが国の戦略について」という報告書が科学技術会議から出され、それを受けて平成13年度にJSTにバイオインフォマティクス推進センター(BIRD)が設立、その後平成17年8月にライフサイエンス委員会で「我が国におけるライフサイエンス分野のDB整備戦略のあり方について」という報告書が出され、これを受けて平成18年度にこの統合DBプロジェクトが5年間の時限付プロジェクトとして開始されている。その後、平成20年10月にライフサイエンス委員会で「ライフサイエンスデータベースの統合・維持・運用の在り方」という報告書が出され、DBCLSとBIRDの一体的な運用をJSTに設置する新たな組織で行うことが提言されている。さらに平成20年12月には総合科学技術会議(CSTP)の統合DBタスクフォースにおいて、統合DB構築のための体制整備、ロードマップ等について検討がなされ、これを受けて平成21年9月から、JSTにおいて統合DBセンターの設置準備委員会で種々の検討を重ねてきた。現時点では、準備段階として、文部科学省の統合DBプロジェクトとJSTのBIRDを一体化して、平成23年度から統合DBセンターをJSTに設置、23年度から25年度までの3年間を第1段階とし、第2段階である26年度からの新たな統合DBセンターの設置については第1段階での実施の状況を見て検討するという流れで現在のロードマップが決められている。現在はこの第1段階に向けての準備を進めているところである。JSTに設置した準備委員会(委員長:かずさDNA研究所大石先生)では平成21年の9月から4回の委員会とその間の分科会などを開催、さらには関係府省との意見交換なども行って、今年6月に検討の取りまとめを行った。その後、詳細な機能について検討するための制度検討ワーキンググループを7月から8月にかけて3回開催した。新センター(名称:バイオサイエンスデータベースセンター、以下、NBDC)の役割としては、CSTPの助言のもとD

B統合に必要な研究開発を実施するとともに、システムの構築・維持・管理等を行うということになっている。主な機能として、戦略の立案、研究開発の推進、DB統合化の推進の3つの役割を果たすことにより、さまざまな国内外のDBの情報を登録または提供していただくことにより、ユーザから見ると一元化されたポータルへのアクセスが実現されて、効率的に情報を収集できることを目指している。DB統合化の推進について少し詳しく説明すると、統合化については、(1)ポータルの構築(各省のDBに対して、所在情報がわかるようにリンクが行われる状態を実現)、(2)横断検索(相互に一括して横断検索が可能となる状態を実現)、(3)アーカイブの構築(各省のDBに対して統一の形式でのデータのダウンロードが可能となる状態を実現)、(4)DB構築連携(最終的に統一したフォーマットで再構築された状態を実現)という4つの段階に分けて進めることを考えている。そして、この中で(1)から(3)までを第1期(平成23~25年度)に実施し、その後は第2期以降というようなスケジュールを考えている。この役割を果たすために、NBDCに持たせる機能として4つを考えている。①戦略の立案、②実際にサービスを提供するためのポータルサイトの構築、運用、③DBを統合化するための基盤技術の研究開発、④関連DB統合化の実際の推進であるが、③、④については、それぞれ基盤技術開発プログラム、統合化推進プログラムという公募プログラムによって実施することを考えている。これら4つの機能が、研究者が国内のさまざまなバイオ関係のDBを統合的に活用することができるようになるための大きな柱である。来年4月以降に正式に発足することを予定しているNBDCの体制(案)として、CSTPの統合DB推進タスクフォースとの連携をとりながら、JSTにおいてはセンター長、副センター長の体制の元、戦略の研究部門、DBを実際に運営する部門、研究開発を推進する部門の3つの部門を設け、いろいろなご意見を伺う場として統合DB運営委員会を設置する。来年の4月のセンター設置に向けてのスケジュールとして、この10月にNBDCの準備室を設置した。ここではJST内の体制の整備など、センターの発足に向けての準備を進める。また、平成23年度に向けて2つのプログラムの公募を実施する。1つ目が基盤技術開発プログラム(DB統合化の実現に向けて基盤となる技術開発を行い、実装までを行う)の募集(11月1日~12月24日)・選考で、その内容は年間約3億円の研究開発費で第1期の3年間で予定採択件数1件、研究総括は本研究運営委員会の委員でもある長洲さんにお引き受けいただいた。2つ目が統合化推進プログラム(分野ごとのDBの統合化などを通じて国内バイオ関連DBの統合を推進する)で、12月の公募開始を予定して現在検討中である。このような募集・選考を行って、23年4月にNBDCを正式に発足するというスケジュールで進めている。

参考資料として準備委員会の委員の名簿、平成23年度概算要求における内閣府の優先度判定の結果を添付した。

●統一フォーマット、統一形式でのダウンロードを可能にして、その後、統一形式で再構築ということだが、個々のDBの工夫、デザインなど全部骨抜きにした生のデータだけが残るという形を考えているのか。DBにも専門家が工夫したさまざまな解析の仕方とかツールがある。DBに関することはすべて新センターがやるなら個別のDB化には各省からはお金を出さないということになると、センターが収集する統一形式のデータは残るが、専門家が工夫したDBは公開・維持されなくなるということが心配である。

→DBアーカイブはオリジナルDBがなくなっているわけではなく、権利関係の整理や他のDBとのマージのためにフォーマットを整えるなどユーザの利便性を高めるために別個につくり直している。なので、心配されているようなことは基本的には起こらないと思う。

●目的によってはフォーマットが統一されたいろいろなDBを使いたい人もいるし、もとのDBのいろいろな検索機能を使いたい人もいるので、(3)と(4)は、ある意味並行して行われると考えて良いか。→そうである。

●ファンディングを大型プロジェクトという形で実施されるが、例えばフェローシップ、奨学金のように若手に資金を出し、もらった人は国内のいろいろな拠点で研究できるという体制は検討しないのか。→予算の規模と第1期が3年という期間であることから、検討の結果、フェローシップなどは盛り込まないことになった。今後の検討課題であると考えている。

●広くDBを受け入れるという姿勢は鮮明に出しているのか。

→そのように出していきたいが、もしそのように見えないことがあれば、いつでも教えていただきたい。

●DBCLSの発表でアーカイブは今後JSTで運用されると聞いたが、それはJSTの新センターの中に設備をつくって移すのか、それとも現状の設備で運用だけをJSTがやるのか。

→形式上は移すことになる。ただ、実際の運用の場所や移行期間などについては、まだこれから具体的

に検討していくという段階である。原則としては、移設することを考えている。

4. DBCLSの体制と状況について

松原委員長からDBCLSの体制と状況について、情報・システム研究機構長でもある堀田副委員長に説明が求められた。

皆様のご協力のおかげで、本プロジェクトは順調に進み、最後の取りまとめの段階まで来られた。

プロジェクト開始時にこれを引き受けるために情報・システム研究機構（以下、機構）としてはDBCLSをつくったわけだが、プロジェクト終了後センターをどうするかというところはいろいろ議論してきた。機構としてはDBというものは非常に重要で、その構築、統合化はやはり機構の大きな使命であると考えているので、結論としては、このプロジェクトの終了にかかわらず、DBCLSは引き続き機構内で運営するという事になった。ただし、新たにさらに幅の広いDB統合の技術開発を目指すと言う意味で、体制についてはいろいろな変更を行うこととなる。第一に高木委員にはJSTの新センター準備委員会でぜひ協力をしてほしいという要請が大石委員長からあったので、センター長を離れていただいてJSTへの協力をお願いした。そのために高木委員は古巣の東大の新領域創成科学研究科の教授に戻り、兼務でJSTの仕事もされる。NBDCの準備室では統括大石先生、補佐高木先生という体制になる。新しいDBCLSセンター長にはすでに東大の情報理工の米澤明憲教授が着任された。ただし、高木委員は本プロジェクトの責任者なので、当面は機構の特任教授としてプロジェクトの取りまとめをお願いする。

DBCLSの経費として、運営費交付金（当初3,750万円、現在5,454万円）が充てられており、この規模は今後の予算でも確保していきたい。ただし、これまでの事業を推進してきた費用の大半はプロジェクトの委託事業費であり、当初は文科省から、最終22年度はJSTのBIRDとの一体的な運営という形で予算を出していただいた。今後はJSTの公募に応募して、ぜひDBCLSの経験を生かしてDB事業に貢献したいというのが、機構としての方針である。

資料には最後に現在の人員について参考までに示した。

DBの重要性を考えて、全体的なサイエンスのDBを広く考えることが機構の使命ではないかという議論も当初はしたことがある。米澤先生が新しくセンター長になられることも契機に、ライフサイエンスだけでなくDB全体を今後どうしていくか、そのための技術開発、方針の立案などを機構として発信するためにDBCLSが役に立てば良いと考えている。皆様からもご支援をお願いしたい。

●ライフサイエンス以外も含めてということだが、今、第4次科学技術政策の立案に際し、成果を出した後でどうするかということがあまり書かれていない。社会とのコミュニケーションというのはあったが、やはり広い範囲での日本の科学技術の成果を、きちんとDBとして納税者に公開していくその姿勢が書かれてほしいと思ってパブコメで意見は言った。第4次の政策に関わっている先生がおられたら、ぜひこういったことを言ってくれと、今のDBCLSの次のものも非常に意味が出てくると思う。

●成果公開についてはその程度しか書かれていないのか。

●私が見た範囲では、非常に弱い。日本の科学技術の成果がどうなるのかというコメントはほとんど無いように思える。

→プロジェクトの応募要領に得られたデータをどうするかということを書き込んでほしいとお願いしているが、ライフサイエンス課以外のほかの分野では全く行われてないと思う。大きなお金を投じてかなりの研究がされても、終わると何かわからなくなってしまうということは大変問題である。

5. プロジェクトの事後評価等について

松原委員長から文科省ライフ課へ、プロジェクト事後評価についての説明が求められた。

まず資料5-1「事後評価の実施について」からご説明する。

文部科学省ライフサイエンス委員会の親委員会である科学技術学術審議会研究計画評価分科会の評価指針（研究開発施策の終了時には目標の達成状況を正確に把握して、その後の施策展開への活用等を行うために事後評価を実施すべき）に基づき行うものであるが、統合DBプロジェクトについては来年度早々にJSTのNBDCに成果を引き継ぐ必要があることから、施策の終了前に事後評価を実施して、その評価結果をJSTの運営する施策の企画と立案等にうまく生かしていきたい。すでに10月27日

付で、統合DBプロジェクトの実施機関の代表研究者あてに事後評価の実施を連絡し、書面評価で必要な成果報告表の作成を11月1日に依頼、来週の金曜日（11月19日）締め切りで作成をお願いしている。事後評価の評価については、成果報告表の書面評価だけではなく、評価期間（12月中を予定）中に成果物に実際にさわっていただき、評価委員の要望によってはさらにヒアリング（1月中旬実施予定）を実施する。

資料5-2「統合DBプロジェクト平成22年度上期進捗の評価依頼について」、5-3「中間評価シート」については、統合DBプロジェクトの作業部会委員の方々に、9月末上期までの進捗状況を相互評価するための依頼文書等で、より良い成果のためのアドバイスなどを11月10日を〆切としてお願いしている。集計結果については今後の参考とするほか、事後評価委員会にも参考資料として使用する予定である。資料5-4「統合DBプロジェクト研究運営委員会作業部会 委員名簿」は参考までに添付した。

6. 合同シンポジウムとユーザ評価について

松原委員長より合同シンポジウムとユーザ評価についての報告が高木委員に求められ、高木委員の要請によりDBCLSから説明がされた。

資料6「合同シンポジウムとユーザ評価について」にしたがって2点ご報告する。

10月5日に「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～」というタイトルで、東大の武田ホールにおいてシンポジウムを開催した。本シンポジウムは、4省でデータ統合の努力をしている関連機関が集まって実施することになり、主催はDBCLSだが共催として医薬基盤研、農業生物資源研、産総研のBIRCおよびJSTが参加し、各省と内閣府の後援をいただいて開催した。ライフサイエンスに関わる情報全体の統合に関する問題点等については各ご専門の先生方の講演をお願いし、また、主催・共催機関関連の取り組みの成果の報告を44件のポスターで発表した。会場参加者は230名、同時配信視聴者70カ所以上であった。シンポジウムのWEBサイトからの講演の資料や録画、ポスターのPDFの公開については現在作業中、一部すでに公開済である。

次に今年のユーザ評価について紹介する。今年はプロジェクト最終年度であるので、平成18～22年度に開発されたサービス等について各機関が準備した説明資料をもとに機関ごと（DBCLSのみサービスを5つに分類）に評価用のページを作成し、10月25日から本日まで評価を実施した。内容は機関ごとのサービスや取り組みの5段階評価とコメント欄への自由記述である。また、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、年度目標と進捗のまとめ表を添付し、各サービスについての一定期間の総アクセス数を掲載した。依頼者数は470名余りで、昨日までにのべ133名が参加している。

- シンポジウム参加者の年齢層はどのようなものか。来場者の興味はどういったところにあるか。
→若い学生の方から委員をなさるような先生方までいろいろな方が来場した。DBの現状や使われ方などに興味があるのだと思われる。
- 今後の事業の予定などについて話題に上ったか。
→最初に郷先生より内閣府のタスクフォースでの議論を説明いただき、最後に高木先生がJSTの準備室等について話されたので、来年以降本事業がどうなっていくかについて理解されたと思う。
- 来年度以降の話はあまり外に出ていないので、機会あるごとに公にしていけないと思う。
→DBCLSでは本年度中には他のイベントの予定はないが、12月のBMB2010（分子生物学会、生化学会合同年会）で高木委員と長洲委員がオーガナイザーのワークショップがあるので、その中で説明していただく予定である。

7. 年間スケジュールと今後の予定について

松原委員長からの求めに応じて、DBCLSから今後のスケジュールについての説明が引き続き行われた。

資料7「統合データベースプロジェクトH22年間スケジュール」は前回の研究運営委員会に提出したスケジュール表をリバイスしたもので、基本的には研究運営委員会、作業分科会、行事予定等も計画どおりあるいはそれ以上の内容で実施された。あとは2月に第12回研究運営委員会と第6回作業部会分科会を合同で実施する予定で、その際にはその時点での成果の概要を各参画機関の先生方からご説明いただき、関係者が一堂に集う会にする計画である。

8. その他

松原委員長より、プロジェクト全般やそれに関する自由な意見を求める発言があり、各委員から発言が相次いだ。

● J S Tからの発表の参考資料2にあるC S T Pの評価について、S A B Cの評価にもましてコメントが重要であり、次期のN B D Cに生かさないといけないと思う。特に「厚労省を巻き込んでほしい」という意見があったが、厚労省との協力関係はどのように進められようとしているのかについてご説明いただけないか。また、(本日オブザーバー参加の)厚労省から何かコメントがいただけないか。

→ (厚労省) 統合DBプロジェクトについて、厚労省は省全体の統合DBを持っていないため経産省などと比べて参加が遅れている状況だが、医薬基盤研のDBを本プロジェクトのアーカイブに移行する具体的な話など、徐々に進めているところである。

→ (文科省) 皆様方のご協力により文部科学省の委託事業は最終年度を迎えることができ、誠に感謝している。来年度からC S T Pの指摘を受けてオールジャパンのバイオサイエンス分野の統合DB事業をJ S Tで実施することになり、現在関係の予算の確保に向けて鋭意作業中である。今後ともご支援いただきたい。ご質問の各省との連携に関して、ライフサイエンスの分野については各省それぞれが予算の特別枠(コンテスト形式で募集を行うもの)に応募している状況だが、DB事業については文部科学省が主導して各省連携の体制をつくり、政府全体としてライフィノベーションの推進基盤となるDBを構築できるよう努力したい。

●日本の予算制度上はライフサイエンスに関わるところが4つほどあるが、それぞれのDBを構築しながらもどこが統合するかという点で争わないようにするにはどうしたらいいのか。

●科学予算の配布機関がオープンアクセスとデータシェアリングに関して何もポリシーを持ってない中国、インド、日本ではデータの流通量が明らかに低い。ポリシーを明確にすることが重要である。

●現政権では、科学技術関係の予算が厳しくなっているのが現実で、省庁連携など今までとは違うアプローチでやっついていかないと予算がとれないという方向になりつつある。C S T Pがどのようにイニシアティブをとるかにもよるが、むしろこの状況は「統合」のような方策にとっては一つの追い風だと思う。

●それに関係して最近になって内閣官房からメディカルイノベーションセンターの構想が出されているが、医療の情報を中心的にまとめる日本版N I Hみたいなものになっていくという方向なのか。

●メディカルイノベーション構想については、内閣官房および経産省、文科省、厚労省の政務三役が主導して行っている。その一環として内閣官房にヘッドクォーターを設置するという案も含まれていたが、日本版N I Hのようなものが創設されるかどうかについては確たる情報がない。ただ、その検討の中でも連携を進めるという話になっているので、省庁連携の追い風になるのではないかなと思う。

●一気に4省庁全部統合するのは、今の日本の状況だとどの政党が政権を取っていてもそれほど楽な話ではないと思うが、文科省としては実績として少しずつ連携の仕組みをつくる努力をするということか。→省益ではなく、日本としてライフサイエンスの水準を高めて、オールジャパンで国際競争に勝っていくようにすることが国益上必要なので、そういったところは実施していく。

●プロジェクトが節目を迎えて制度が変わるということを業界にもっとよく浸透するような活動をすべきではないか。来年の4月からの体制などについて公表できるようになったら、新体制で何がしたいのか、体制の変化に伴うプラス・マイナスについて説明して、パブコメを取るのはいかがか。米国では大統領が変わったからといって突然N I Hをやめるということは影響が大きいとあり得ない。規模は小さくとも本プロジェクトの成果はそのようなものになって来ているので、プロジェクト終了時の体制変更については明言しなくてはならない。J S Tの準備室にもご検討いただきたい。

→ (J S T準備室) 検討したい。

●先ほど日本とインドと中国について言及されていたが、中国では本プロジェクトのような分野で若手が育っていると聞かす。このままでは、インドと日本はさらに置いていかれるのか。

●中国では20代のバイオインフォを学んだ人が何百人単位で育っていることは確かである。韓国でも大きな動きが出てきていると思うが、残念ながら、日本ではなかなかそういった動きが見られない。

●今回の統合DB事業では、データ収集だけではなく、それらが流通している状況がもっと見えるといいと思う。データを共有できているという点を評価に入れていただきたい。事後評価ではそういった観点の評価をされるのか。

→ (文科省) 現在は評価指標には入っていない。

●「厚労省を巻き込んでほしい」というコメントに関して、医療関係のデータにはプライバシーなどの関係で非連結にしたとしてもなかなか公開できないものがあるとよく議論になるが、その点を何とか突破することがこの統合DBの成否を決めると思う。他の生物を対象とした研究成果としてそのDBをつくるのは中間段階であり、ヒトのデータと合わせて公開することが最終的な目的だと思う。一方で、ヒトの病気やプライバシーにかかわる情報の公開には制限があるのは理解できるが、1種の動物として科学的にほかの動物などと比べたりすることは必要である。なので、どうしたら公開できるようになるかということを経済DBで考えていただきたい。そのためにも「厚労省を巻き込んでほしい」と言われたのだと思う。

●確かにヒトのデータの共有は非常に大事なポイントだと思う。現在の統合プロジェクトの中でも、ヒトのタイピングデータやリシークエンスデータなどの流通に関するガイドラインを設け、完全にオープンなもの、ユーザ登録を要するもの、倫理審査委員会等の手続きを要するものといったレベルを分けた公開を実施している。ただ、これも技術の進化や欧米等の動きとともに変わるものなので、引き続き検討が必要である。もっと突き詰めるとおもとのインフォームド・コンセント（IC）の取り方にも関わってくるので、統合DBよりも一段上の内閣府の統合DB推進タスクフォースで、ヒト由来データの扱いに関する分科会を設けてルール等を議論することになっているが、そのタスクフォースがまだ1回しか開催されていないので、かなり時間がかかるかも知れない。

●総合科学技術会議においても、ゲノムコホートを推進すべきという意見があり、ICの件も含めて非常に課題があるがいずれクリアしていかなければならない。DBとは違った側面からライフサイエンス研究を政策として担う文部科学省としては検討していきたい。

●いろいろな制約があることは理解しているが、こういった研究のためのデータには根本的にICが必要なのか疑問に思う。既存のルールでは制約が強くなりがちなので、相当根本的に考えないと反科学的なことになるのではないかと心配している。

●前競争的な部分は公的資金で実施し、競争的な部分は利益を求める民間がやればよいと分けると、税金を投入した前競争的な部分を共有することが重要になってくる。ヒトに関するデータにも前競争的と競争的の境界があると思う。ヒトデータ全体で考えずに前競争的な部分が何で、そこに税金を投入して結果が共有されるところという効果があるということを考えるべきである。

●日本では生命倫理に関して長らくいろいろな議論があるがまだ集約されていないと思う。本気で議論する機会ができることを望んでいる。

●総合科学技術会議で連携施策の科学調査研究が行われたが、最大の論点はいかに医学的な情報と基礎的な情報をつなげるかであった。しかし、メディカルイノベーションではむしろメディカルと基礎を分けようという動きがある。「バイオ」でひとくくりしてきたためにかえって患者や国民のニーズに応えられないと考えられており、メディカルはメディカルだけでやっていきたいと公言される先生もいる。厚労省が統合DBに入ろうとするのはいいが、一方でメディカルイノベーションのほうでは別のDBが必要だという話では統合DBの意義がわからなくなってくる。新センターの名称のバイオサイエンスの定義は人によりあいまいなので、メディカルとはちょっと違うとも言われかねない。その意味では、名称も非常に重要である。また、先ほどの機構の話で、DBのターゲットをライフサイエンス以外に広げると言われたが、例えば国会図書館の電子化などいろいろな動きがあり全然まとまっていない。ライフサイエンス以外へDBを拡張する際には慎重に進める必要がある。

●生のデータは全部公開されてほかの人がアクセスできるようにならなければ、統合DBとしての意味がないと思う。共有すべき情報が個人情報という形でまるで所有権が、それも本人のではなく担当の医師の所有権があるかのように語られているというのはおかしいと思う。生命倫理は基本的には事例ごとに対応すべきで、全ての事案を解決するようなルールでは科学を停めてしまうと思う。本気で議論して、だめならだめで結論を出さないと先へ進めないような気がする。そのためにはこういう問題があるということをもっと言うべきである。

●過去には倫理の人が出てくると研究そのものができなくなると随分思ったものだが、抜きにしてはものごとが進まないのも確かである。メディカルの問題にしても個人の問題以外、例えば薬をつくるか診断をするといったことは倫理とは全く関係がない。なので、ヒトに関する科学の進展とそのデータの用途研究をきちんと区別する必要があると思う。また、ゲノム情報はオーダーメイド医療や診断に役に立つといったことを短絡的に言ってきて、非常にいい面だけを力説し過ぎたと思う。だから、基礎的な情報の統合DBを持つことのメリットは従来の個人情報にかかわる倫理の制限は及ばないものであるということを経済DBに訴えていくことがとても大事だ。ヒトに応用するための科学をやることとその科学を個

人に応用するという科学と両方あると思う。倫理の観点では個人への応用を出発点にして議論しているので、個人に応用できるはずのない科学は倫理が及ぶ範囲には無いといった意見をまとめておくべきだと思う。

●ただし、それはとても難しいと思うので、ヒトデータの共有についてはやはり所有権の問題として、つまり、非連結の情報は所有権が離れていると考えるべきだと思う。後で調べたらすぐわかるという話と所有権の問題とは全然違う話だ。非連結であればヒトデータも自由に使っていいものだと思うし、科学コミュニティできちんと管理できるはずだ。

●そういう議論をオープンに堂々とやるためにも、こういうレベルの研究においては倫理を気にしないでいいということを打ち出して議論をするのがいいと思う。

●今までの議論から、このライフサイエンスのDBが日本の社会に及ぼす影響の中には健康問題に関わるDBの発展が重要な位置を占めることがわかった。ただし、その議論は日本の生命倫理の問題にとどまらず、日本の医師のカルテ管理の問題と密接に関係しているので、最初は混乱することを覚悟で問題提起を行う必要があると思う。

●ヒトのゲノムを知ることがヒトに関するすべてを知りたいわけで、病気だけに限ったことではない。あくまでその一部に疾患の話があるのである。統合DBの中でゲノムデータの公開ができればそれを個人データとして扱う以外の見方の研究者、例えばインフォマティクスなどの観点からも利用できる。多くの人に利用されれば有用性や価値も高くなる。

●生命倫理やカルテなど医療情報の問題に話が集中したが、このライフサイエンスの統合DBの次のねらいについて少しは議論できたと思う。次回の研究運営委員会は2月だがその時に再度この議論をしても間に合わないので、新センターのねらいについて簡単にまとめていただきたいが、いかがか。

→(JST準備室) 大まかなポイントでよければ今月中ぐらいには一回まとめるようにしたい。

本日の議題についてはすべて議論が終了したので、松原委員長の閉会の挨拶により本会は終了した。

以上