

ライフサイエンス分野の統合データベース整備事業

統合医科学データベース構築方式の開発
21年度 委託業務研究成果報告書

平成22年3月

国立大学法人東京医科歯科大学

田中 博

国立大学法人大阪大学

佐古田 三郎

本報告書は、文部科学省の科学技術試験研究委託事業による委託業務として、東京医科歯科大学、及び大阪大学が実施した、平成21年度の「統合医科学データベース構築方式の開発」の成果を取りまとめたものです。

目次

1 委託業務の目的	3
2 平成21年度（報告年度）の実施内容	3
2.1 実施計画	3
2.2 実施内容（結果）	4
2.3 成果の外部への発表	16
2.4 活動（運営委員会等の活動等）	26
2.5 実施体制	28
2.6 整備実績一覧【代表機関名：東京医科歯科大学】	30

1 委託業務の目的

ライフサイエンス関係データベースに散在する臨床・疾患データベース（DB）を分子から個体にいたる多階層の疾患横断的な統合医科学DBへ構築するための高度化・統合化方式を開発することを目的とする。がん疾患、神経疾患を対象に実証する。

このため、国立大学法人東京医科歯科大学及び国立大学法人大阪大学は共同で業務を行う。

国立大学法人東京医科歯科大学では、がん疾患に関わる研究開発を中心に、国立大学法人大阪大学では神経疾患に関わる研究開発を中心に実施する。

2 平成21年度（報告年度）の実施内容

2.1 実施計画

2.1.1 東京医科歯科大学

① 疾患・臨床医科学データベースに特化した統合技術開発に関する研究

平成20年度に実施した多階層の統合医科学データベースを統合構築するために必要な統合技術の要件定義結果を踏まえて、統合化要素技術として疾患オントロジーの整備及びプロトタイプとして開発したセマンティック検索エンジンの高度化（分子情報に関するデータを取り扱うためのインタフェースの構築）を行う。国立大学法人東京医科歯科大学は、肝臓がん、大腸がんを中心として研究を実施する。

② 疾患データベースの高度化と実証的統合に関する研究

平成20年度に実施した要件定義結果を踏まえて、症例に基づいた統合検索システムを構築し、東京医科歯科大学の統合検索システムの登録研究者に向けて公開する。

国立大学法人東京医科歯科大学は、肝臓がん、大腸がんを中心として研究を実施する。

③ プロジェクトの総合的推進

分担機関である東京医科歯科大学は、中核機関である情報・システム研究機構の全体戦略に従い連携して本事業を推進する。

プロジェクト全体の連携を密としつつ円滑に運営していくため、技術検討会等を開催し、分担機関グループ内の連携・調整及び進捗状況の確認を行う。

がん疾患を中心として、国内の主要な疾患データベースを参照できるような仕組みを検討する。

プロジェクトで得られた成果については、倫理委員会が承諾する範囲内で、東京医科歯科大学の公開する統合医科学データベースのホームページや学会・講演会等を通して情報発信に努める。

2.1.2 大阪大学

① 疾患・臨床医科学データベースに特化した統合技術開発に関する研究

平成20年度に実施した多階層の統合医科学データベースを統合構築するために必要な統合技術の要件定義結果を踏まえて、統合化要素技術として疾患オントロジーの整備を行うと共にプロトタイプとして開発したセマンティック検索エンジンの統合検索システムとの接続プログラムを開発する。

国立大学法人大阪大学は神経疾患（パーキンソン病）について研究を実施する。

② 疾患データベースの高度化と実証的統合に関する研究

大阪大学の神経疾患データベースについて、症例に基づいた検索等の具体的要求事項とその結果について検討する。さらに東京医科歯科大学が構築する統合検索システムとの連携を行うとともに、倫理的な問題を解決（患者同意を含め）し、公開する。

国立大学法人大阪大学は神経疾患（パーキンソン病）について公開患者数を増加させる。（約データ数500）

2.2 実施内容（結果）

2.2.1 東京医科歯科大学

① 疾患・臨床医科学データベースに特化した統合技術開発に関する研究

昨年度までに、東京医科歯科大学の統合的医療データベース（iCOD）、大阪大学のパーキンソン病DB、および国立癌センター研究所のGeMDBJを中心として要件定義書を作成し、統合検索のプロトタイプシステムβ版(図1)を構築した。平成21年度は、このプロトタイプシステム（β版）の「セマンティック検索エンジン」および「統合検索インタフェースプログラム」をエンハンスし、プロトタイプシステム（バージョン1）(図2)の開発を実施した。

統合検索システムは、大きく3つのコンポーネント（User Interface Manager、Navigating Engine、DB Manager）から構成される(図3)。DB Managerのクライアント機能を各要素DBに設置することで、統合検索環境を整備することが可能となる。検索のコアとなるNavigating Engineをセマンティック検索エンジン、ユーザ

及び要素DBとのインタフェースとなるUser Interface Manager、DB Managerを統合検索インタフェースプログラムと称す。

検索処理としては、ユーザが検索キーワードを入力するとUser Interface Manager は、Terminology/ Thesaurus DB等を利用して、検索キーワードの標準化処理を行う。標準化された検索キーワードは、Navigating Engineが、Nosological Ontologyによって意味的解釈を行い、セマンティック検索条件へと変換される。DB Manager は生成されたセマンティック検索条件を各要素データベースの構造をData Format Mapper DBを利用して変換し、要素DB毎に検索を実行し、その検索結果を受け取る。各要素データベースの検索結果は、DB Managerが標準的形式に変換し、User Interface Managerがユーザに検索結果を表示する。

セマンティック検索エンジンの高度化として、以下を実現した。

- 要件定義書に基づく、検索キーワード生成処理の最適化
 - 3つのモジュール (User Interface Manager、Navigating Engine、DB Manager) を実装した。(図3)

- Terminology/Thesaurusの追加・変更・削除等
 - DBに格納することにより、システム管理者が任意にTerminology/Thesaurusの追加、変更、削除を行えるようにした。(図3)

- Nosological Ontologyの追加・変更・削除等
 - DBに格納することにより、システム管理者が任意にTerminology/Thesaurusの追加、変更、削除を行えるようにした。(図3)

- 分子情報の取り扱いに対応可能なように、検索エンジンを改良
 - iCODを要素データベースとして接続することにより、分子情報の取り扱いに対応した。(図2)

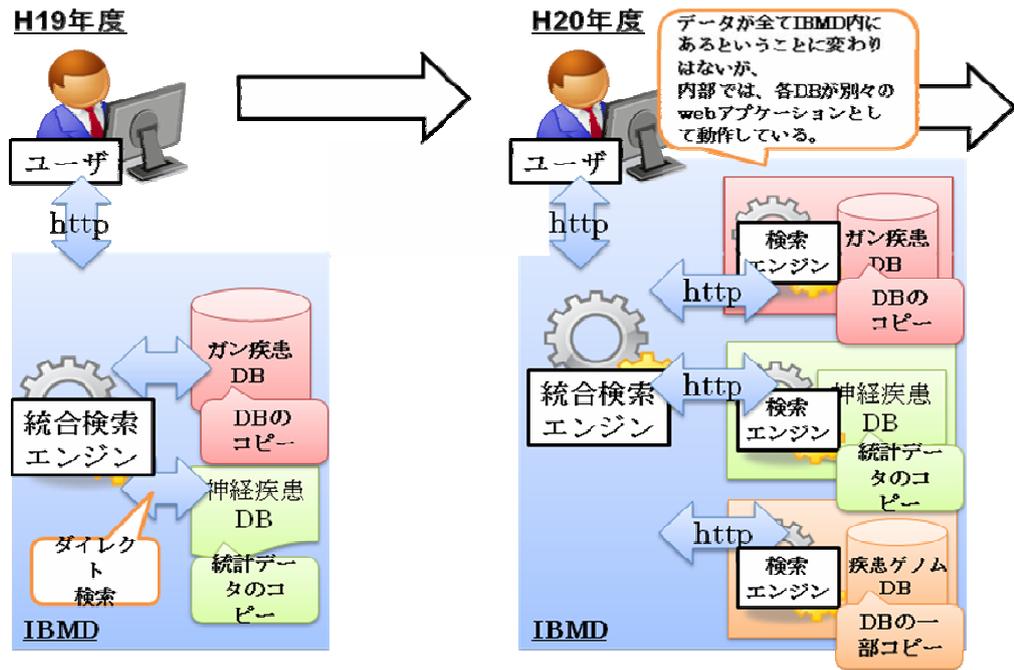


図 1 統合検索システムの構成 (H19, 20 年度)

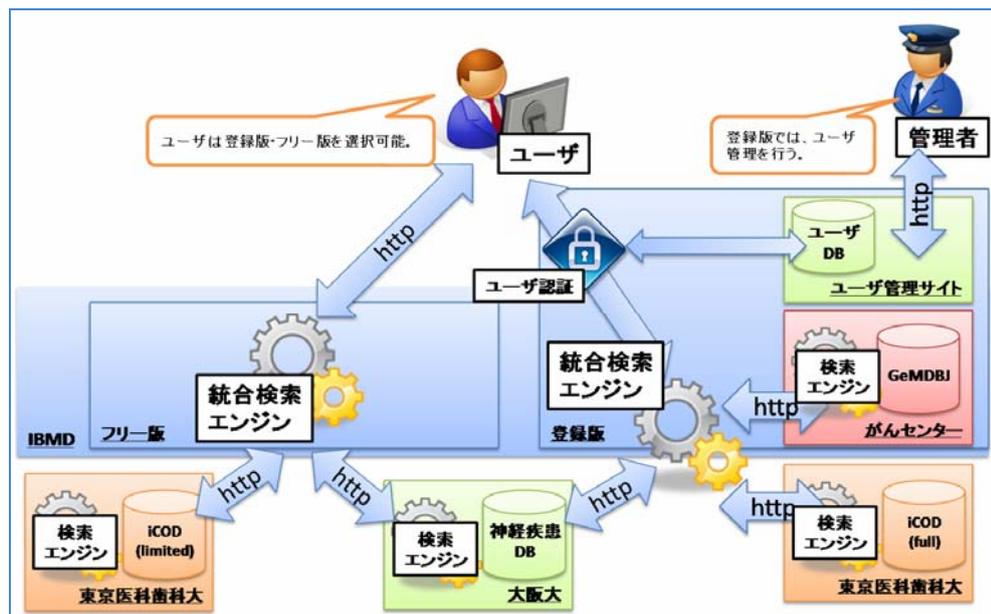


図 2 統合検索システムの構成 (H21 年度)

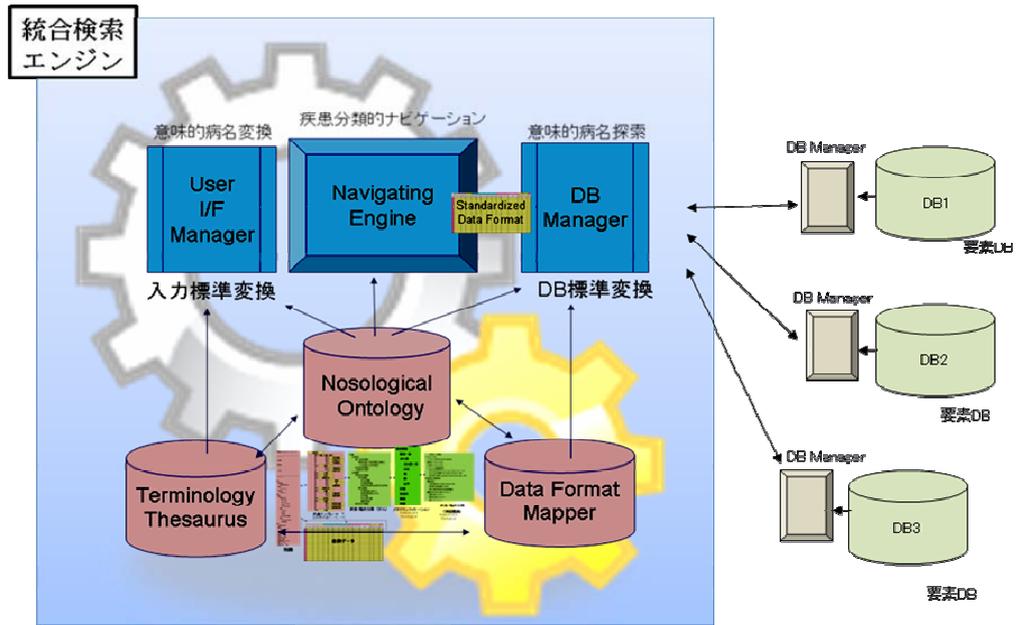


図 3 統合検索エンジン

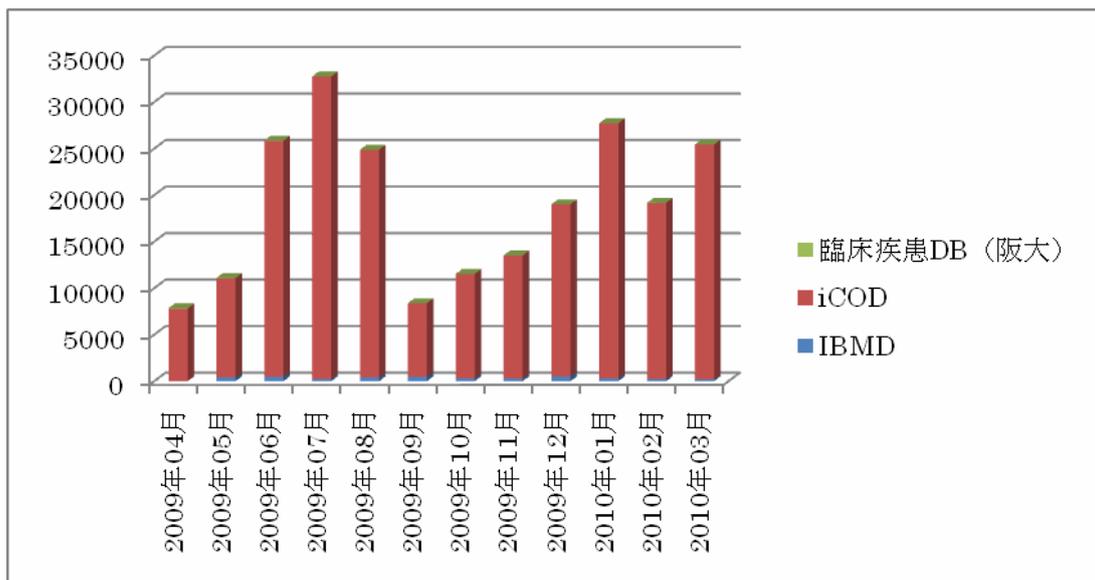


図 4 H21 年度アクセス数

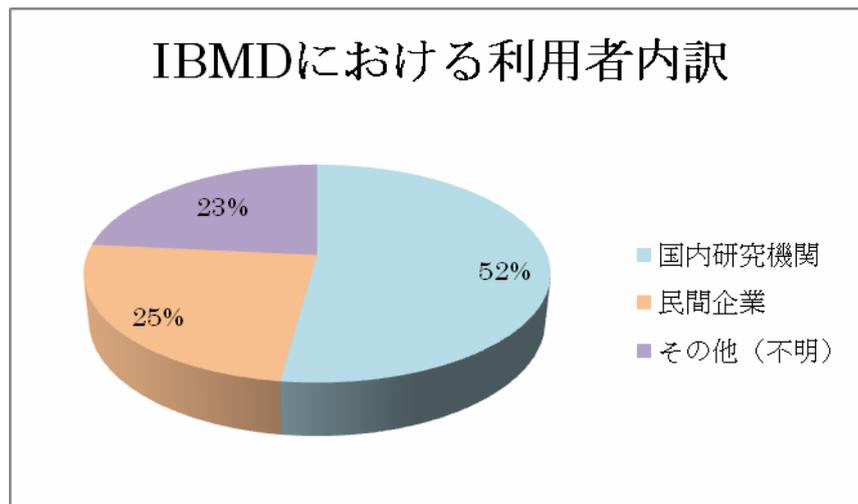


図 5 H21 年度利用者内訳

② 疾患データベースの高度化と実証的統合に関する研究

Data Format Mapperを改良し、昨年度のプロトタイプシステム（β版）では統合データベースに配置しているData Format Mapperを要素データベースにも配置可能とすることで、統合データベースの処理を分散可能とする仕組みを構築した。また、要素データベースへの検索クエリ発行処理の最適化を行うことで、疾患・臨床医科学データベースであるiCOD、パーキンソンDBおよびGeMDBJの統合的な検索を可能とした。更に倫理面、セキュリティ面を考慮し、ログイン認証システムの作成、ユーザ登録およびユーザログイン時の認証の仕組みを実装した(図 6)。

次に、本統合検索システムをより簡単に利用できるよう、データベースのカテゴリ（基本情報、病歴、生活歴など）毎にアクセスレベルを検討し、フリーアクセスレベルと登録研究者レベルを設定した。その後、倫理規定案に準拠の上、それぞれのレベルに対応するサブセットを作成し、その公開を開始した。

また、平成22年度の高度化に向けた1つの指標とすべく、各々のレベルでのユーザ評価を得るためのアンケート集計システムを構築した。

1. 正規のユーザーであるか確認する機能 (認証機能)
2. ユーザー毎にアクセスを制御する機能 (認可機能)
3. ユーザー登録情報を管理する機能
(ユーザー登録申請、ユーザー登録、ユーザー登録メール通知、
ユーザー情報修正、ユーザー登録削除、パスワード再発行、等)

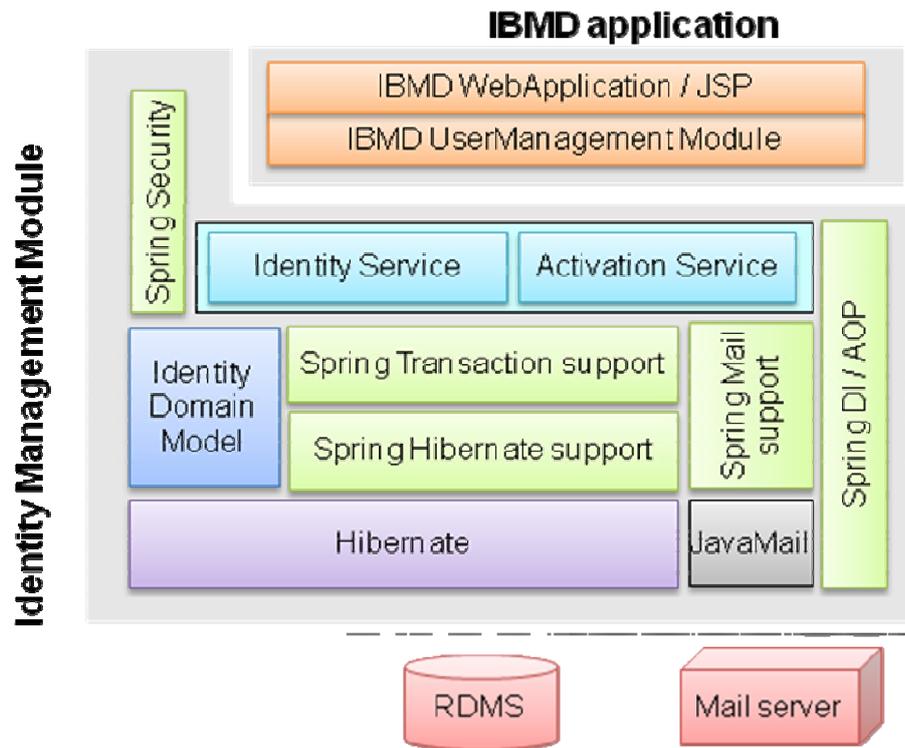


図 6 ユーザー認証

③ プロジェクトの総合的推進

大阪大学グループと連携し、オントロジー・シソーラス (Normalized Clinical Knowledge) を利用した検索機能の高度化、および関連情報への連携方法について、分担機関内統合化要素技術検討会「2. 4 活動 (運営委員会等の活動) 参照 (20 ページ)」を開催し、検索機能やインタフェース設計に関する課題の抽出、仕様の決定を行った。関連情報への連携については、まず国内の主要な疾患データベースを精査し、次に各データベースの構造などを把握後、本プロジェクトとの関連性を検討し、ホームページ上での掲載を開始した。

また、倫理基準検討会「2. 4 活動 (運営委員会等の活動) 参照 (20 ページ)」を設け、平成 22 年度に完了する倫理規定の基礎となるべく、アクセスレベルの設定、疾患解析 DB グループが作成した倫理規定との連携等を中心に検討を行った。

次に、国際的に通用しうる統合データベース基盤技術確立のために必須のハー

ドルとなる統合データベース要素技術の国際標準化のため、WHO-ICD11およびISO TC215での活動を行った。WHO-ICD11検討会議では、統合医科学DBで用いられている情報モデル、NCK知識コンテンツをプリセットとして用いることができるかどうか検討した。また、ISO TC215では、統合医科学DBの要素技術である臨床情報モデルと深い関連性があるGSVMLが臨床ゲノム分野で世界初の国際標準（IS）となった。新たに、WHO、ISOの場において、東アジア伝統医学の情報モデルの検討を開始した。現在、東アジア伝統医学は、中国、韓国、日本が、それぞれ、国の総力と威信をかけて情報モデルとオントロジーの国際標準化とデータベースの統合に力を入れている分野である。

以上のように、計画の策定、進捗管理および必要に応じた計画の見直しを随時行うとともに、本プロジェクトの認知活動を推進する等、今年度の業務を効率的に遂行することができた。

2.2.2 大阪大学

① 疾患・臨床医科学データベースに特化した統合技術開発に関する研究

図5に、今年度までに開発を進めてきた神経疾患（パーキンソン病）データベースの全体像を示す。H19年度には大阪府下10数機関からのパーキンソン病患者に関するデータを収集・管理しているデータベース（Secure Study システム）から、統合データベース事業で公開するデータを匿名化等の処理を施した上で抽出し、公開用DBに格納する仕組みを開発した。さらにH20年度には、大阪大学に設置している公開用データベースを、インターネットを介して東京医科歯科大学が開発した階層の統合医科学データベース（IBMD）の統合検索エンジンと統合するために必要な統合技術の要件定義を行った。今年度は、それらの成果を踏まえて、これらまでデータ移動方式によりプロトタイプとして試作していた、IBMDのセマンティック検索エンジンの統合検索システムと、パーキンソン病DBを接続するプログラム（IBMD要素DBインタフェース）を開発した。

図6にIBMD要素インタフェースの概要を示す。IBMD要素インタフェースは、統合検索エンジンからの要求に応答するための接続プログラムで、これまで開発を進めてきたパーキンソン病DBの追加モジュールとして位置づけられる。統合検索エンジンとの要求・応用に用いる通信はHTTPによるPOSTを利用したXML形式でやり取りが行われる。その基本仕様については、H20年度に作成した統合のための要件定義に基づき、実装レベルでの問題点がないかといった仕様の詳細化は、東京医科歯科大-大阪大で互いに協力して進めた。

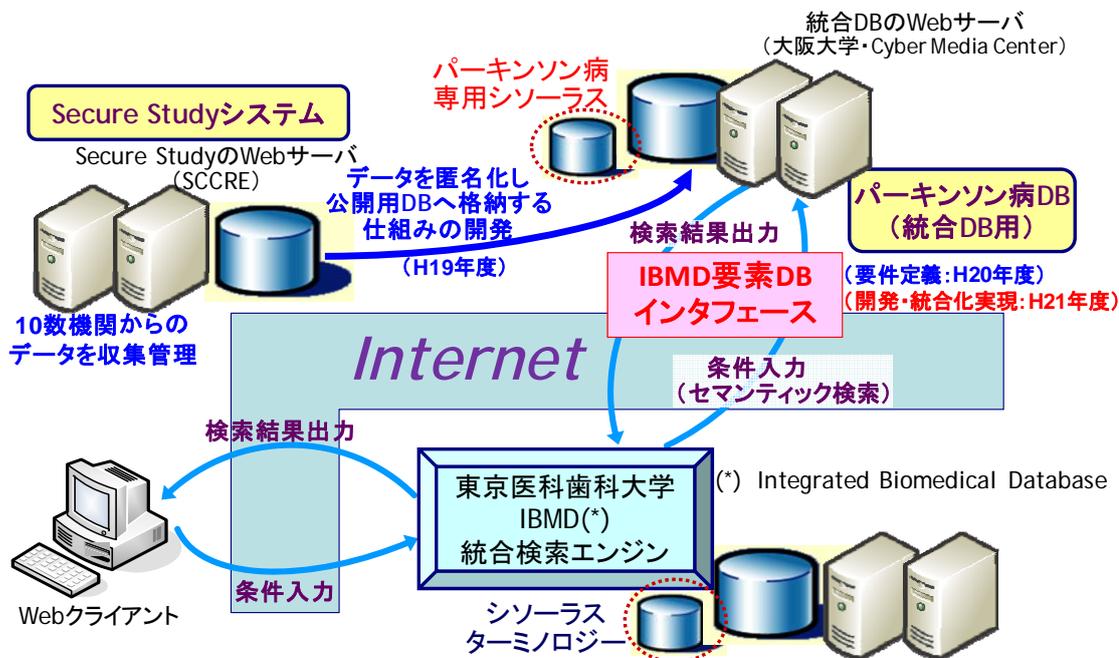


図5 パーキンソン病データベースの全体像

以下に、統合検索エンジンと IBMD 要素インタフェースの間で行われる要求-応答の概要を示す。IBMD 要素インタフェースは、東京医科歯科大学の統合検索エンジン（WEB サービス）からの要求により、大阪大に設置されたパーキンソン病 DB（統合 DB）のデータを返す。具体的な機能としては、以下の2つの機能を提供する。

1) 件数問い合わせ機能

統合検索エンジンから渡された検索条件を基に、該当するデータの件数を返す。検索条件の値は、完全一致とする。入力情報に誤りがあった場合は、「invalid request」のメッセージを返す。要求・応答は、それぞれ、req-count.xml および res-count.xml という XML 形式でやり取りされ、これらの情報のチェックは、XML スキーマを用いて行われる。

2) 一覧取得機能

統合検索エンジンから渡された検索条件を基に、該当するデータの患者情報を返す。検索条件の値は完全一致とし、返却する臨床データの単位は、患者 ID 単位とする。入力情報に誤りがあった場合は、「invalid request」のメッセージを返す。要求・応答は、それぞれ、req-search.xml および res-search.xml という XML 形式でやり取りされ、これらの情報のチェックは、XML スキーマを用いて行われる。

これらの2つの機能は、図6に示すように、統合検索エンジンと IBMD 要素 DB インタフェースの間のやり取りを、要件定義にそって実装したものであり、その詳細は次頁の図7,8で示す。この要求・応答の枠組みは、パーキンソン病 DBのみならず、他の DB とも共通して用いられる。

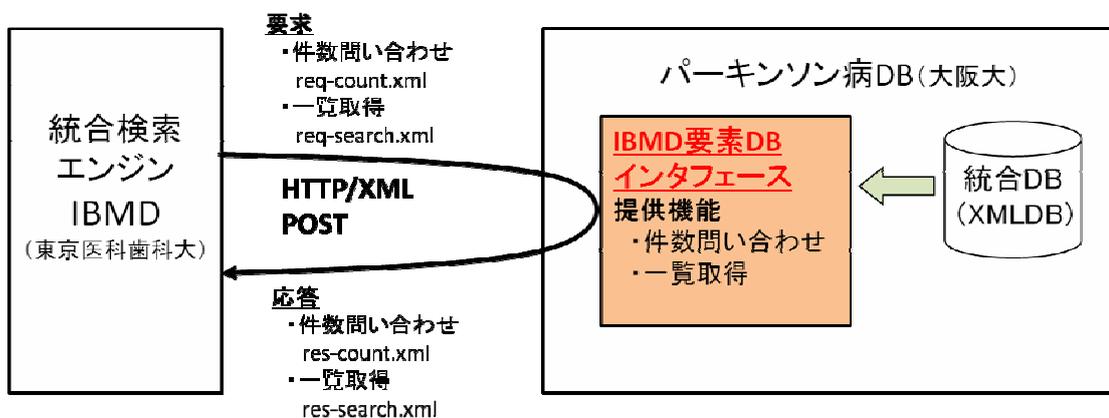


図6 IBMD 要素インタフェースの概要

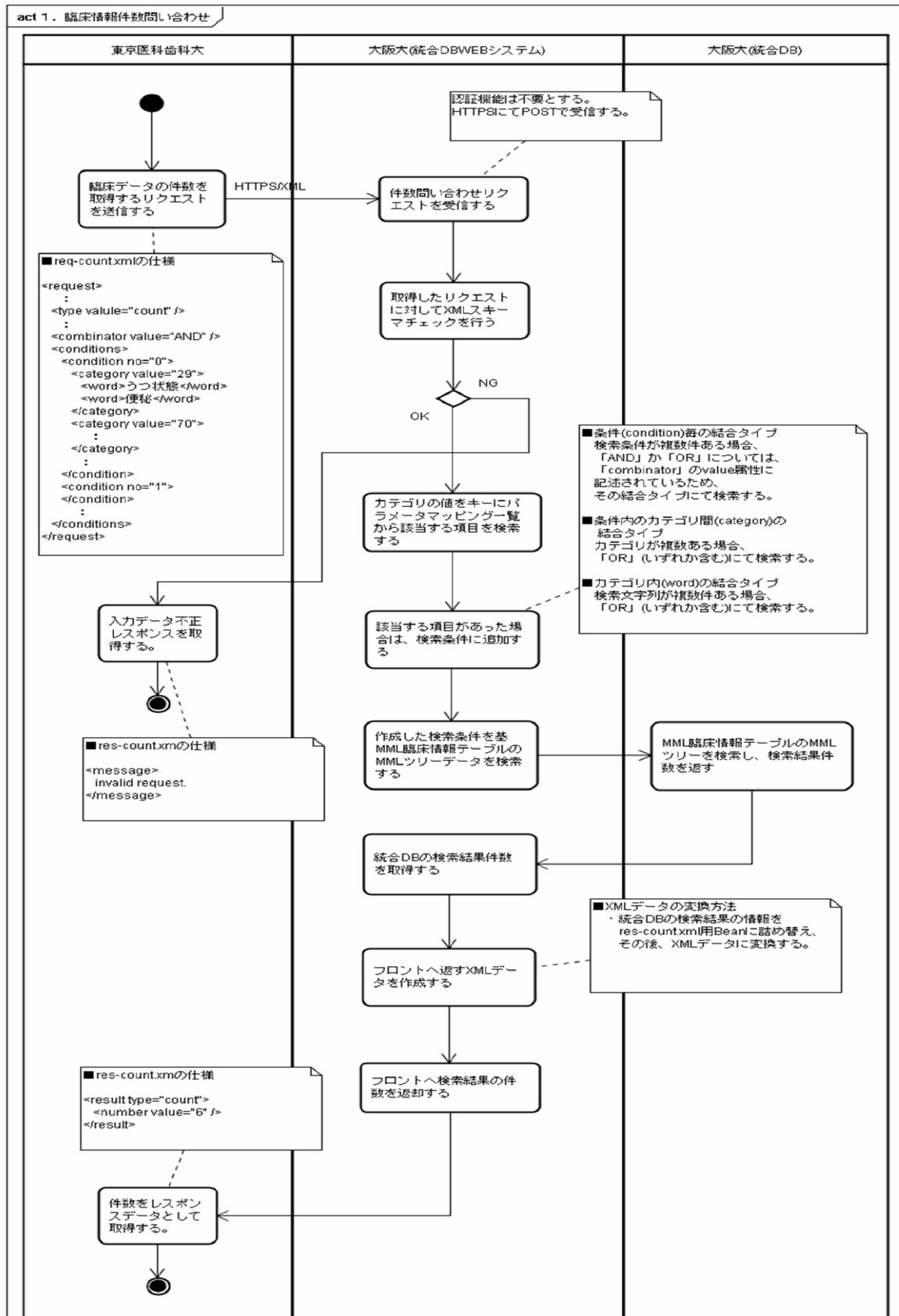


図7 IBMD 要素インタフェースの件数取得機能の処理の詳細

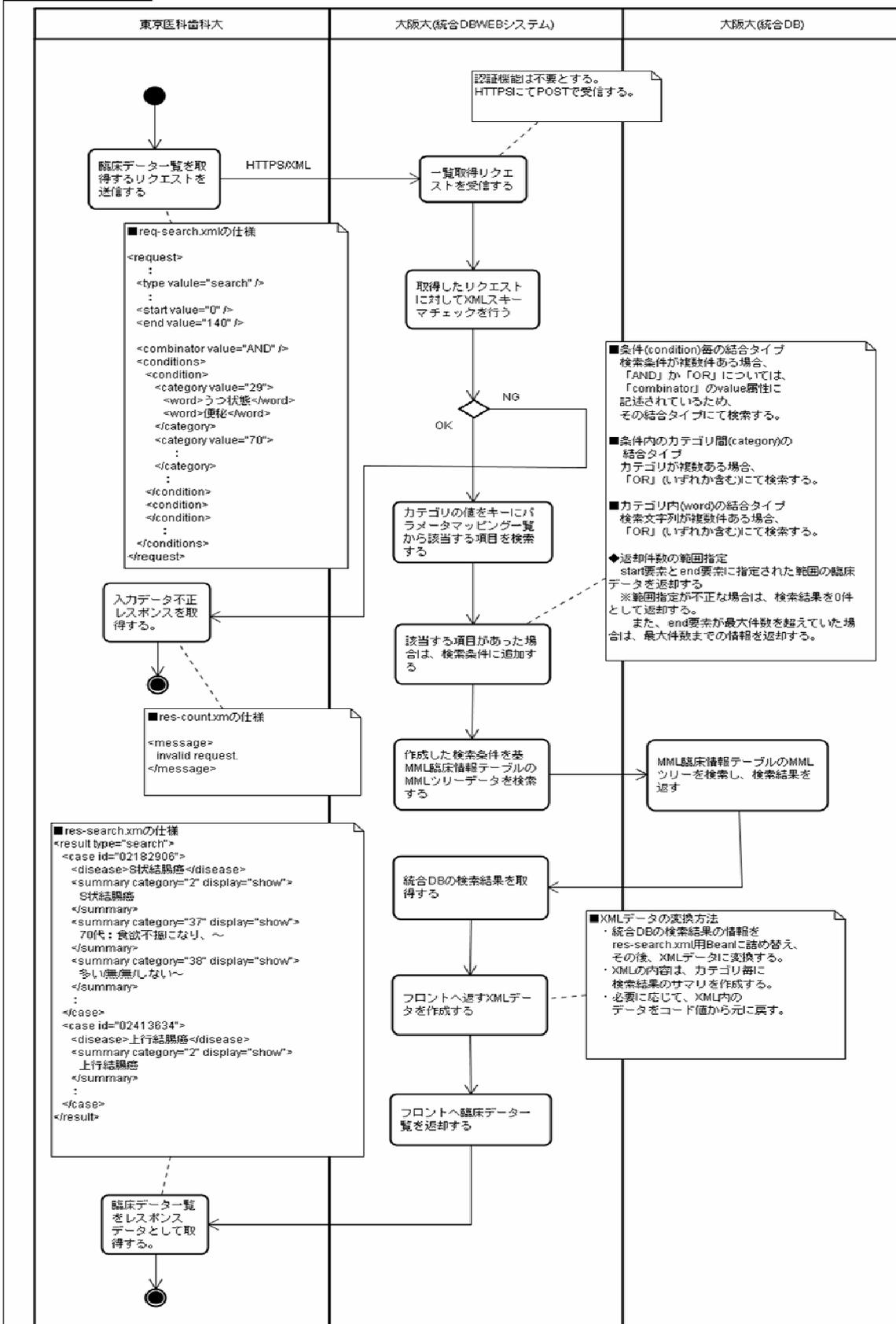


図 8 IBMD 要素インタフェースの一覧取得機能の処理の詳細

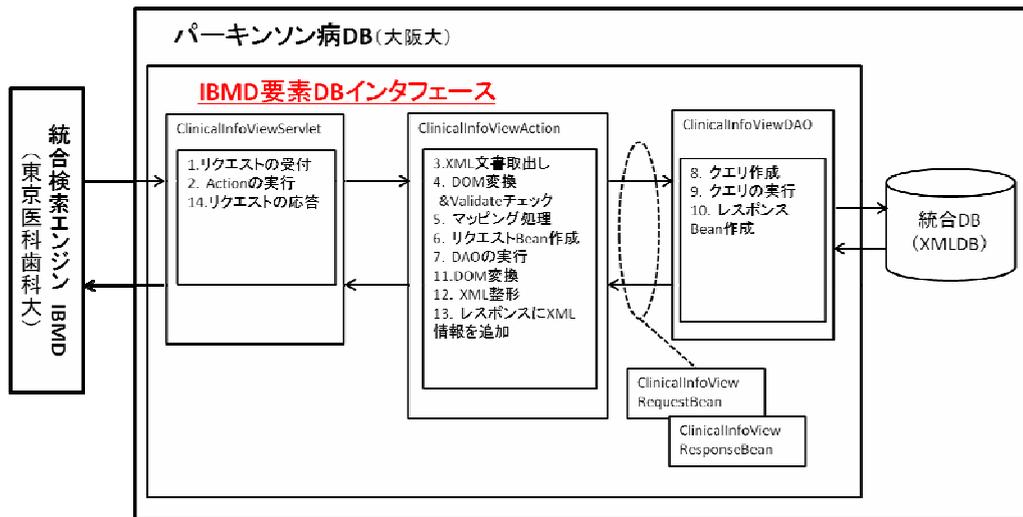


図7 IBMD 要素インタフェースの内部処理の概要

前述の統合検索エンジンと IBMD 要素インタフェースの質問・応答に対し、パーキンソン病 DB の内部処理としては、データを格納している XMLDB(統合 DB)と IBMD 要素 DB インタフェースとの間で質問・応答処理が実装されている。図7に示すように、その処理は14のステップを経て実現される。この処理は、IBMD との統合を実現する個々のデータベースで設計・実装を行う必要がある。

本年度は、上述の仕様に沿った IBMD 要素インタフェースをパーキンソン病 DB に追加し、医科歯科大に設置された IBMD の統合検索エンジンからの統合検索を可能とした。

② 疾患データベースの高度化と実証的統合に関する研究

大阪大学の神経疾患データベースについて、症例に基づいた検索等の具体的な要求事項とその結果について検討した。まず、データベースの公開方法として、昨年度に検討した結果、データ収集用の DB から公開用 DB にデータを移行する段階で匿名化の処理が施されていることと、検索結果はすべて統計情報としてのみ表示されるといった点から、ユーザ認証なしでの公開で倫理的に問題ないという結論に至ったので、パーキンソン病 DB のユーザ認証を外してフリーアクセスでの公開とした。

また、ユーザの使い勝手を向上するための工夫として、検索条件の入力方法および結果の提示方法の改良について検討した。その結果として、

- 1) 従来の検索条件をリストから選択して検索する方法に加えて、自由入力キーワードとパーキンソン病に特化したシソーラスを組み合わせた「シノニム検索」
- 2) 複数条件を用いたクロス検索の結果を、従来の一覧表示に加えて、グラフ表示やテーブル表示などの表示形式の導入

といった仕様案を作成した。これらの実装については、来年度を予定している。

なお、神経疾患（パーキンソン病）データベースの公開データ数については引き続き拡充作業を進め、400 データから 500 データまで増加させた。

2.3 成果の外部への発表

2.3.1 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
○Systems Pathology - A new concept for understanding disease -, 口頭	Tanaka H.	<i>CBI-KSBSB Joint Conference, Haeundae Grand Hotel, Busan, Korea</i>	Nov 4-6, 2009	国外
○Systems Pathology、口頭	Tanaka H.	<i>Personalized medicine 2009, Bioinformatics Center, Hangzhou, China</i>	May 29-31, 2009	国外
○Present Situation and Future Trends of Omics-based medicine and systems pathology、 口頭	Tanaka H.	<i>Conf. on Biomedical Electronics and Biomedical Informatics, Moscow, Russia</i>	Aug, 2009	国外
○Systems pathology- A new concept for understanding disease -, 口頭	Tanaka H.	<i>Chinese Association of Biophysics, China</i>	Jul, 2009	国外
○Systems evolutionary analysis on the putative Hox transcriptional networks、口頭	Ogishima S, Tanaka H.	<i>European Society for Evolutionary Biology, Trin</i>	Aug 2009	国外
○Preliminary analysis on systems evolution of the putative Hox transcriptional	Ogishima S, Tanaka H.	<i>OIST Evolution of Complex Systems, Okinawa</i>	Dec 2009	国外

networks、口頭				
○Within-host coevolution of gag p453l and protease d30n/n88d demonstrates virological advantage in a highly protease inhibitor-exposed hiv-1 case、口頭	Shibata J, Ren F, Iwatani Y, Tsang H, Matsuda M, Hasegawa N, Tanaka H, Sugiura W.	<i>10th Annual Symposium on Antiviral Drug Resistance, Virginia USA</i>	Nov 2009	国外
○「病気をシステムで解くーオミックス医療の未来」、口頭	田中 博	サイエンスフォーラム講演	2010年3月28日	国内
○「次世代オミックス医療とシステム創薬」、口頭	田中 博	東京医科歯科大学バイオ医療オミックス情報学人材養成プログラム成果発表会ー国際シンポジウム「次世代疾患オミックスの新展開」、みらいホール(桜木町)	2010年2月2日	国内
○「Omics-based Medicine and Systems Drug Discovery」、口頭	田中 博	横浜市大Palaverセミナー	2010年2月1日	国内
○「基礎研究成果を医療ニーズに活かすBioinformatics」、口頭	田中 博	横浜市立大学	2009年12月16日	国内
○「次世代疾患オミックスとシステム創薬」、口頭	田中 博	並列生物情報処理イニシアティブ IPAB 創立10周年記念シンポジウム、東京 SGI ホール	2009年11月27日	国内
○「オミックス医療とシステム創薬」、口頭	田中 博	オミックス医療情報学寄附講座開設記念講演会、歯学部附属病院特別講堂	2009年11月19日	国内

○「病気を「システムで解く」オミックス医療の可能性」、口頭	田中 博	かわさきサイエンス&テクノロジーフォーラム 2009、かながわサイエンスパーク西棟 (川崎市)	2009年11月17日	国内
○「Structure of protein interaction networks and their implication on drug design」、口頭	田中 博	Harvard Medical School セミナー	2009年11月16日	国内
○「デジタル・オミックスとシステム創薬」、口頭	田中 博	第12回シンポジウム ゲノム創薬フォーラム、日本薬学会館長井記念ホール	2009年11月9日	国内
○「疾病解明と治療・創薬のためのシステムバイオロジー」、口頭	田中 博	神奈川科学技術アカデミー講演、東京	2009年10月27日	国内
○「次世代疾患オミックスシステムパソロジー」、口頭	田中 博	第37回日本磁気共鳴医学会大会、パンパシフィック横浜	2009年10月2日	国内
○「オミックス医療とシステム創薬—パーソナルゲノム時代の医療—」、口頭	田中 博	ゲノム創薬フォーラム 講演	2009年7月21日	国内
○「生命のシステム論的解析」、口頭	田中 博	九州工業大学特別講義、福岡	2009年6月5日	国内
○「疾患を「システムで解く」オミックス医療の将来」、口頭	田中 博	大学院生命情報科学教育部公開講演会「病気に挑む生命科学」、歯学部特別講堂	2009年6月2日	国内
○ Ontological Models and Approaches in the Integrated BioMedical Database Project、口頭	Hiroshi Tanaka and Jun Nakaya	Proceedings of International Health Level Seven Interoperability Conference 2009 HL7 Japan in Press	2009年5月9日	国外

Existing content Existing sources e. g. Integrated BioMedical Database Project、口頭	Jun Nakaya	WHO ICD11 Revision Internal Medicine -TAG International Conference Tokyo	April. 2009	国外
◎統合医科学データベース Integrated BioMedical DataBase (IBMDB)、ポスター	中谷純, 井戸 敬介, 佐古田 三郎, 溝口理 一郎, 古崎晃 司, 下川和 郎, 田中博	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」、東京	2009	国内
○ Transdisease Omics analysis of Clinical data by SAGE、ポスター	Kazuro Shimokawa, Yasunori Sasaki, Hiroshi Tanaka	CBI-KSBSB joint conference 2009	2009	国外
◎ iCOD : an integrated clinical omics database based on the systems-pathology view of disease	Kazuro Shimokawa, Kaoru Mogushi, Satoshi Shoji, Atsuko Hiraishi, Hiroshi Mizushima, Hiroshi Tanaka	日本分子生物学会 (2009)	2009	国内
◎パーキンソン病データベース構築の試み	須貝 分宣, 山本 洋一, 佐古田 三 郎	第 50 回日本神経学会 総会 (仙台国際センター)	2009 年 5 月 25 日	国内
◎パーキンソン病の臨床データベースの	山本 洋一, 古崎 晃司,	統合データベースプロ ジェクトシンポジウム		国内

構築と課題	溝口 理一郎, 野崎 一徳, 下條 真司, 濱崎 俊光, 佐古田 三郎	2009 「データベースが拓く これからのライフサイ エンス」 東京大学本郷キャンパ ス浅野地区武田ホール	2009年6 月12日	
-------	-------------------------------------	--	----------------	--

2.3.2 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
○Gene expression phenotypes for vascular invasiveness of hepatocellular carcinomas.	Tanaka S., Mogushi K., Yasen M., Noguchi N., Kubo A., Nakamura N., Ito K., Miki Y., Inazawa J., Tanaka H., Arii S.	<i>Surgery,</i> <i>147:405-14</i>	2010	国外
○Mutations in the interferon sensitivity determining region and virological response to combination therapy with Pegylated-interferon alpha 2b plus ribavirin in patients with chronic hepatitis C-1b infection.	Nakagawa M., Sakamoto N., Ueyama M., Mogushi K., Nagaie S., Itsui Y., Azuma S., Kakinuma S., Tanaka H., Enomoto N., Watanabe M.	<i>J Gastroenterol</i>	2009 (accepted)	国外
○IQGAP1 and vimentin	Tsubota A.,	<i>Carcinogenesis</i>	2009	国外

<p>may be key regulator genes in naturally occurring hepatotumorigenesis induced by oxidative stress.</p>	<p>Mogushi K., Nariai K., Tanaka H., Tada N.</p>		<p>(accepted)</p>	
<p>○Comparison of HCV-associated gene expression and cell signaling pathways in cells without HCV replicon and in replicon-cured cell.</p>	<p>Nishimura-Sakurai Y., Sakamoto N., Mogushi K., Nagaie S., Nakagawa M., Itsui Y., Sekine-Osajima Y., Tasaka-Fujita M., Onuki-Karakama Y., Suda G., Mishima K., Yamamoto M., Ueyama M., Funaoka Y., Watanabe T., Chen CH., Kakinuma S., Tsuchiya K., Tanaka H., Enomoto N., Watanabe M.</p>	<p><i>J Gastroenterol</i></p>	<p>2009 (in press)</p>	<p>国外</p>

Inferring within-patient HIV-1 evolutionary dynamics under anti-HIV therapy using serial virus samples with vSPA.	Hasegawa N., Sugiura W., Shibata J., Matsuda M., Ren F., Tanaka H.	<i>BMC Bioinformatics</i>	2009 (accepted)	国外
○Structures of protein interaction network and their implications on drug design.	Hase T., Tanaka H., Suzuki Y., Nakagawa S., Kitano H.	<i>PLoS Comput Biol</i>	2009, 5(10) e1000550	国外
○Prognostic value of matrix Gla protein in breast cancer.	Yoshimura K., Takeuchi K., Nagasaki K., Ogishima S., Tanaka H., Iwase T., Akiyama F., Kuroda Y., Miki Y.	<i>Molecular Medicine Reports</i>	2009 (in press)	国外
○Mutation of epidermal growth factor receptor is associated with MIG6 expression.	Nagashima T., Ushikoshi-Nakayam R., Suenaga A., Ide K., Yumoto N., Naruo Y., Takahashi K., Saeki Y., Taiji M., Tanaka H., Tasai	<i>FEBS Journal</i> , 276:5239-5251	2009	国外

	SF., Hatakeyama M.			
○Expression of Aurora B and their Alternative Variant Forms in Hepatocellular Carcinoma and the Adjacent Tissue.	Yasen M., Mizushima H., Mogushi K., Obulhasim G., Miyaguchi K., Inoue K., Makahara I., Ohta T., Aihara A., Tanaka S., Arii S., Tanaka H.	<i>Cancer Science</i> , 100(3):472-480	2009	国外
○In silico and in vitro analyses reveal role of Hes1 in taste cell differentiation.	Ota MS., Kaneko Y., Kondo K., Ogishima S., Tanaka H., Eto K., Kondo T.	<i>PLoS Genetics</i> , 5(4): e1000443	2009	国外
○Surgical contribution to recurrence-free survival in patients with macrovascular invasionnegative hepatocellular carcinoma.	Tanaka S., Mogushi K., Yasen M., Noguchi N., Kubo A., Kurokawa T., Nakamura N., Inazawa J., Tanaka H., Arii S.	<i>Journal of the American College of Surgeons</i> . 208(3):368-374	2009	国外

A likelihood look at the supermatrix-supertree controversy.	Ren F., Tanaka H., Yang Z.	<i>Gene</i> , 441:119-125	2009	国外
○ Development of a database and ontology for pathogenic pathways in periodontitis.	Suzuki A, Takai-igarashi T, Numabe Y, Tanaka H.	<i>In Silico Biol.</i> 9:0020	2009	国外
○ Rich can get poor: conversion of hub to non-hub proteins.	Tun K., Rao RK., Samavedham L., Tanaka H., Dhar P. K.	<i>Systems and Synthetic Biology</i> , DOI 10.1007/s11693-09-9024-9	2009	国外
○ Benefits of pharmacogenomics in drug development - earlier launch of drugs and less adverse events	Ohashi W., Tanaka H.	<i>Journal of Medical Systems</i> , DOI 10.1007/s10916-09-9284-7	2009	国外
○ 「分担執筆」 第3章 創薬に向けた生命情報の統合 3、オミックス情報を基盤にした創薬	田中 博	遺伝子医学 MOOK14 次世代創薬テクノロジー 実践：インシリコ創薬の最前線、180-185	2009	国内
○ 「一般誌」 広範囲 血液・尿化学検査 免疫学的検査-その数値をどう読むか-[第7版](1) I 総論 オミックス医療と臨床検査	西堀 眞弘、 田中 博	日本臨牀、Vol. 67、 増刊号 8、47-51	2009	国内
○ Genomic Sequence Variation Markup Language (GSVML)	Nakaya, J., Kimura, M., Hiroi, K., Ido, K., Yang, W., Tanaka, H.	"International Journal of Medical Informatics (Int J Med Inform.)" Volume 79, Issue 2,	February 2010	国外

		130-142 PMID: 19969503		
○ A Content Model for the ICD-11 Revision	Samson W. Tu, Olivier Bodenreide r, Can Celik, Christophe r G. Chute, Sam Heard, Robert Jakob, Guoquian Jiang, Sukil Kim, Eric Miller, Mark M. Musen, Jun Nakaya, Jon Patrick, Alan Rector, Guillermo Reynoso, Jean Marie Rodrigues, Harold Solbrig, Kent A Spackman, Tania Tudorache, Stefanie Weber, Tevfik	Journal of American Medical Informatics Association accepted	2010	国外

	Bedirhan Ustun			
○"A Content Model and An Information Model for Disease Description"	Jun Nakaya, Keisuke Ido, Kaei Hiroi, Hiroshi Tanaka.	"Interdisciplinary Ontology Proceedings of the Second Interdisciplinary Ontology Meeting" Vol. 2, Keio University Press. Edit by M. Okada and B. Smith 103 - 113	2009	国外

2.4 活動（運営委員会等の活動等）

分担機関内統合化要素技術検討会

平成21年 8月 7日（場所：東京医科歯科大学 3階会議室）

医科歯科大関係出席者：中谷純、下川和郎、井戸敬介、荒木栄一、古江基樹、森下尚浩、水元清

阪大関係出席者：古崎晃司、須藤広明、大黒敏弘

平成21年 12月 24日（場所：東京医科歯科大学 3階会議室）

医科歯科大関係出席者：田中博、下川和郎、井戸敬介、荒木栄一、古江基樹、森下尚浩、水元清

阪大関係出席者：古崎晃司、須藤広明、大黒敏弘

（大阪大学内）

平成21年 6月 3日（場所：大阪大学）

阪大関係出席者：佐古田三郎、山本洋一、古崎晃司、濱崎俊光、安井涼子
システム開発関係者：須藤 広明、大黒敏弘、濱田 茂樹

平成21年 9月 16日（場所：大阪大学）

阪大関係出席者：佐古田三郎、山本洋一、古崎晃司、濱崎俊光、安井涼子

システム開発関係者：須藤 広明、大黒敏弘、阪本和洋

平成21年 3月 11日（場所：大阪大学）

阪大関係出席者：佐古田三郎、山本洋一、古崎晃司、安井涼子

システム開発関係者：須藤 広明、阪本和洋

倫理基準検討会

平成21年 11月 14日（場所：東京医科歯科大学 3階会議室）

医科歯科大関係出席者：中谷純、井戸敬介

阪大関係出席者：山本洋一

2.5 実施体制

2.5.1 東京医科歯科大学

研究項目	担当機関等	研究担当者
疾患・臨床医科学データベースに特化した統合技術開発に関する研究	東京医科歯科大学 情報医科学センター	○中谷 純 野川 裕記 高井 貴子 下川 和郎 荻島 創一 井戸 敬介 茂櫛 薫 広井 嘉栄 武若 真樹子
	東京医科歯科大学 難治疾患研究所	三木 義男
	東京医科歯科大学 大学院医歯学総合研究科	新村 芳人
疾患データベースの高度化と実証的統合に関する研究	東京医科歯科大学 情報医科学センター	○中谷 純 水島 洋 下川 和郎 荻島 創一 井戸 敬介 茂櫛 薫 広井 嘉栄
	東京医科歯科大学 大学院医歯学総合研究科	有井 滋樹 天笠 光雄 新村 芳人
	東京医科歯科大学 難治疾患研究所	稲澤 譲治
プロジェクトの総合的推進	東京医科歯科大学 情報医科学センター	◎田中 博 中谷 純 下川 和郎

2.5.2 大阪大学

氏名	所属	役職	役割	業務内容
佐古田三郎	大阪大学 臨床医工学融合研究教育センター (大学院医学系研究科)	教授	業務主任者	研究総括
山本洋一	大阪大学医学部附属病院 臨床試験部 自主臨床研究部門(大学院医学系研究科)	特任准教授 (常勤)	担当責任者	臨床データベース公開
溝口理一郎	大阪大学 産業科学研究所	教授	実施担当	臨床データベース高度化
下條真司	大阪大学 サイバーメディアセンター	教授	実施担当	臨床データベース高度化
濱崎俊光	大阪大学 大学院医学系研究科	准教授	実施担当	臨床データベースバリデーション
古崎晃司	大阪大学 産業科学研究所	准教授	実施担当	臨床データベース高度化
安井涼子	大阪大学臨床医工学融合研究教育センター(大学院医学系研究科)	技術補佐員	実施担当	研究補助

2.6 整備実績一覧【代表機関名：東京医科歯科大学】

(1) 保有データ情報

(1-1) データの種類

①生物種	(1) 東京医科歯科大学：ヒト 1 大阪大学：ヒト
②試料・ライブラリー等の種類、数	(1) 東京医科歯科大学：＜試料＞肝臓癌、大腸癌、口腔癌の手術または生検検体（約350） ＜データ＞肝臓癌、大腸癌、口腔癌の症例情報（約400）、 肝臓癌、大腸癌、口腔癌検体の遺伝子発現解析結果（約200） (2) 大阪大学：＜データ＞神経難病の症例情報（約500）
③測定方法	(1) 東京医科歯科大学：診療情報収集、DNA マイクロアレイによる遺伝子発現解析 1 大阪大学：診療情報収集大阪大学：診療情報収集
④データの内容	(1) 東京医科歯科大学：臨床情報（基本情報、病歴・生活歴、臨床検査、画像診断、治療、予後等） 分子情報（マイクロアレイ遺伝子発現情報） (2) 大阪大学：臨床情報（基本情報、症状、治療等）
⑤その他、特記事項	

(1-2) データソース

①現在のデータ量	(1) 東京医科歯科大学：試験公開可能なデータは、約250症例分の症例情報。 非公開データを含めると、約400のがん症例情報を有する。 1 大阪大学：試験公開可能なデータは、約500症例情報による統計情報。 非公開データを含めると、約600の症例情報を有する。
②データ区分	■自前 □第三者 □文献データ ■計算結果等の二次データ □その他（下欄に詳細を記述） 倫理審査委員会が承諾する範囲内で、今後も継続的に症例情報を追加していく。
③将来の増加の見込み	所有者（東京医科歯科大学、大阪大学） 公開（□可 □否 ■その他 [各機関の倫理審査委員会が承諾する範囲での公開]）
④権利関係	所有者（東京医科歯科大学、大阪大学） 公開（□可 □否 ■その他 [各機関の倫理審査委員会が承諾する範囲での公開]）
⑤その他、特記事項	個別に試験公開可能なデータの公開をしている。 (http://ibmd.tmd.ac.jp) 各データベースの倫理規定に基づいての公開を行うとともに、統合医

	科学データベースの倫理規定案の策定を推進する。
--	-------------------------

(1-3) データの管理状況

①更新頻度等の管理状況、体制	症例情報の収集及び検体からの遺伝子発現解析は継続的に実施しており、集積データをクレンジングし、公開用データベースへのデータ更新は年1回～2回を予定している。
②その他、特記事項	

(1-4) データベース関係

①DB 管理者数	(1) 東京医科歯科大学：3名 1 大阪大学：1名
②キュレータ・アナレータ数	(1) 東京医科歯科大学：3名 (CRC他) 1 大阪大学：1名
③データ構造	各データベースで独自のデータ構造
④DB 管理ソフト	PostgresSQL
⑤サーバのOS	Linux
⑥サーバ規模	
⑦DB へのアクセス数	
⑧独立 IP 数	1
⑨その他、特記事項	DB への検索メニューは報告書を参照、オントロジーについては東京医科歯科大学が独自に作成したものを使用。

(2) データ (又はDB) の連結、統合化整備 (※試験的、限定的公開済みのものも含む。)

通番	データ (又はDB) の名称	公開 / 未公開	概要 (データの種類 (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	統合的医療データベース http://ibmd.tmd.ac.jp	試験公開	臨床、病理、分子情報 (遺伝子発現情報) を統合化したデータベース。 疾患共通のテンプレートに情報を適応し、250 症例のデータを公開した。 最終年度までに 300 症例のデータ公開を行う予定。
2	パーキンソンデータベース http://ibmd.tmd.ac.jp	試験公開	神経難病特の、臨床情報データベース。 500 症例情報の臨床 10 項目に対して、統計情報を取りこれを公開した。 最終年度までに 700 症例情報のデータ公開を行う予定。

(3) DB 基盤システム、ツール等開発成果物の整備 (※試験的、限定的公開済みのものも含む。)

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	要件定義書	プロジェクト内部公開	統合医科学データベース構築に関わる統合化技術（情報モデル、オントロジー、セマンティクス）、標準化、公開倫理等に関して、全国の全疾患DBを対象とした調査に基づき、要件定義書を完成した。
2	統合検索システムプロトタイプ http://ibmd.tmd.ac.jp/	試験公開	東京医科歯科大学の統合的医療データベース、大阪大学のパーキンソンデータベース、国立がんセンター研究所のGeMDBJに対しセマンティクス検索技術、ユーザインターフェースを検証するためのプロトタイプシステムを構築した。 機能限定版のプロトタイプシステムを評価用に試験公開する予定。

(4) その他の成果物（(2)、(3)に該当しないもの）

通番	名称	公開／未公開	概要