

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』



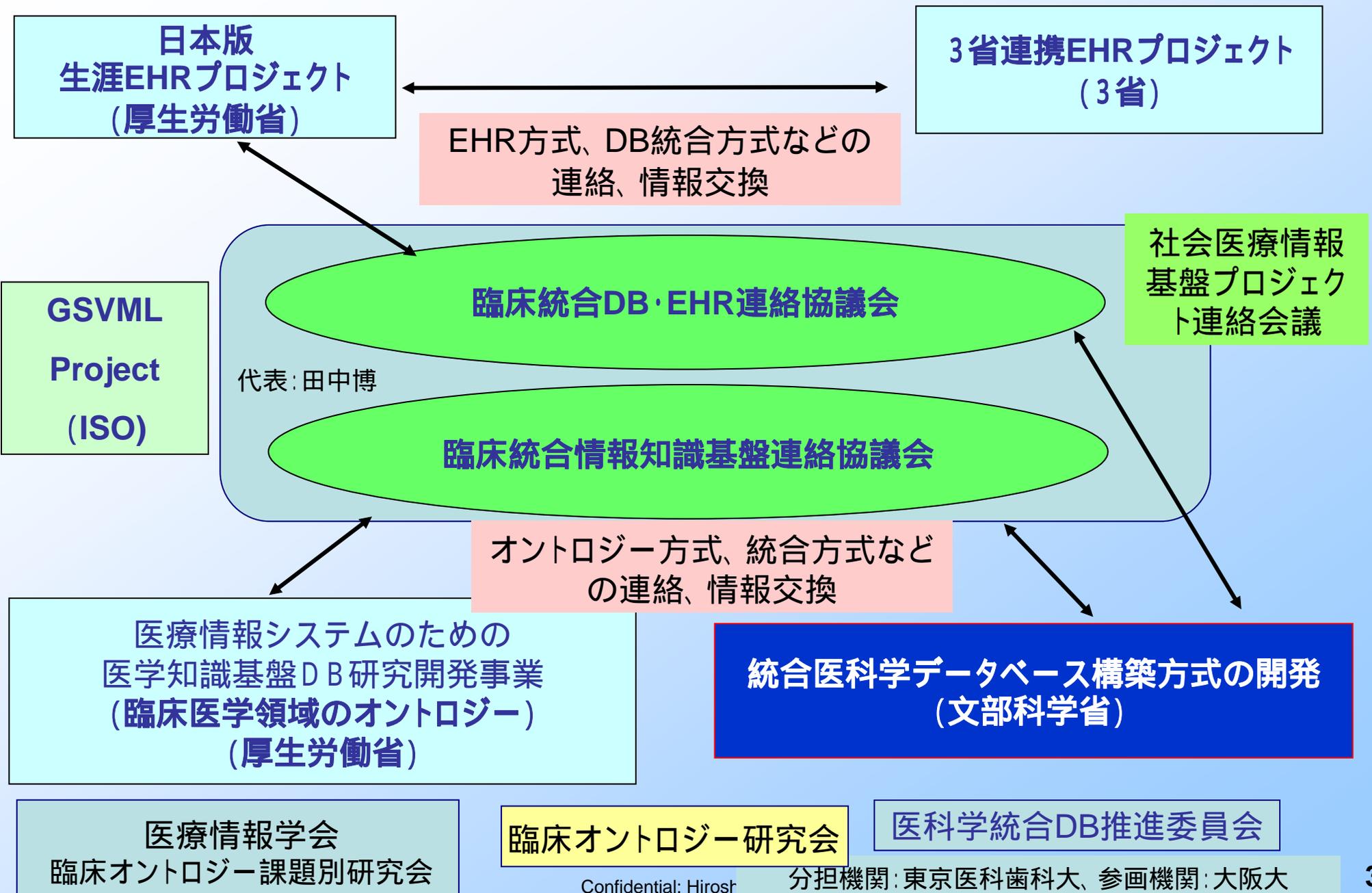
東京医科歯科大学 田中 博

平成20年1月8日

事業目標

- **わが国の医科学分野**のデータベースの統合化を行う
 - 2大疾患(がん・神経難病)から**全疾患**への**対象拡大の試み**
 - 臨床DBの全国調査
- **疾患DB統合利用システムの構築**
 - わが国に存在する臨床各分野の疾患データベースを統合的に利用可能とする情報環境を構築する。そのための方式を開発する
 - **他の関連プロジェクト(厚労省、総務省など)との連携**
 - 医科学DB統合化推進委員会の設置(疾患DBの担当者の連絡会)
 - 疾患臨床オントロジーの開発連携
- **医科学データベースの高度化**
 - オミックスと臨床・病理情報を統合した新しい疾患区分による統合
- **ヒト症例データベース公開における社会倫理課題**の検討

統合医科学データベースプロジェクト体制と関連主要プロジェクトとの連携



平成19年度 成果目標

- 臨床疾患領域における、**統合DB利用システムの基本設計**を確立する。
- 日本において多施設からの**臨床情報の所在**を公開の社会的制約に関する**倫理的、法的問題**を整理する。
- 疾患統合DBシステムの概念を提示した**事例DB**を公表する

具体的な
公開数値目標

倫理案の作成
着手

国内臨床DBの調査

平成19年度中に試験公開を行う
がん症例(100例前後)

パーキンソン病症例(100例前後, 現在倫理委員会審査中)

国内臨床疾患DBの調査方法

目的: 要素DBの把握、要素倫理基準の把握、国内技術傾向の把握と要素標準技術の選定

外部公開情報の収集

方法

- ・インターネットによる検索
- ・関連機関からの資料入手

対象

- ・各研究機関のHP及びDB
- ・各省庁の科学研究費成果報告一覧DB
例) 文部科学省科学研究費補助金採択DB公開状況一覧
- ・研究機関が発行する調査報告書
例) 学術情報DB実態調査報告書(国立情報学研究所)
- ・関連ポータルサイト
- ・その他、個別調査により判明したDB

実態調査のまとめ

方法

1. 調査項目に沿って、各データベースについて外部公開情報を収集する
2. 非公開情報のアンケート回答を調査項目に沿って収集する
3. 両者をまとめた上で総括し、実態と今後の活動課題に即した活動方針を策定する

非公開情報の独自調査

方法

- ・アンケート郵送調査
- ・アンケートWEB調査
- ・アンケートFAX調査
- ・ヒアリング調査(一部)

対象

- ・医学系学会
- ・大学付属病院
- ・国立病院

調査項目

概要

- DB概要 (データ件数、内容、方式、データ形式、アクセス件数)
- ポリシー (公開基準、倫理規定、個人情報保護基準)
- 匿名化 (上記個人情報保護基準に準じた具体的方法)
- 拡張性 (多施設連携性、標準化基準(ISO,HL7等))
- オミックスデータ連携性(オミックスデータの有無)
- データの中身 (内容、構造、項目)

- (1) ライフサイエンス分野の統合データベース整備事業(以下本事業)に対する認知状況
- (2) 本事業を知った情報経路
- (3) 回答者の所属する組織の臨床・疾患データベース(以下DB)の保有の有無とその個数
- (4) 保有する臨床・疾患DBの種別(対象疾患・オミックスデータ保有の有無)
- (5) 保有する臨床・疾患DBの概要
- (6) 保有する臨床・疾患DBのデータボリュームと関連する事項
 - (6-1) 登録症例数
 - (6-2) 症例の項目数
 - (6-3) 症例追跡期間
 - (6-4) 登録遺伝子数
 - (6-5) 遺伝子の項目数
 - (6-6) データ容量
 - (6-7) アクセス数
 - (6-8~10) その他任意記載項目

- (7) 保有する臨床・疾患DBのデータ保存フォーマット
- (8) 保有する臨床・疾患DB利用時のアクセス手段
- (9) 保有する臨床・疾患DBの公開有無とそのURL
- (10) 保有する臨床・疾患DBデータの更新頻度
- (11) 保有する臨床・疾患DBの該当する法令又は運用指針の有無および生命倫理面・安全面又は個人情報の保護のために講じている措置
- (12) 本事業によるDB参照の許諾の可否
- (13) 上記(12)で参照不可となった場合の理由
- (14) 他の組織とのDB連携の有無とその方法
- (15) 保有する臨床・疾患DBの医科学オントロジーへの準拠の有無とその内容
- (16) 回答者の所属する組織以外の臨床・疾患DBの紹介依頼項目
- (17) 本事業への期待やご意見、ご質問等の自由記述欄

実際に用いられた Web調査票(一部)

Mozilla Firefox

https://www.axiohelix.net/lifescience/

11. 貴組織が実運用中または計画中のデータベースについて、該当する法令又は指針がある場合には、該当するものすべてにチェックをつけてください。法令、指針等に定められる手続以外にも貴組織が必要と認めて講じる生命倫理面・安全面または個人情報等の保護についての具体的措置がございましたら、「18. その他」の欄にご記入ください。

- 1. 医療法
- 2. 医師法
- 3. ヒトゲノム・遺伝子解析研究に関する倫理指針
- 4. ヒトに関するクローン技術等の規制に関する法律
- 5. ヒトに関するクローン技術等の規制に関する法律施行規則
- 6. 特定胚の取扱いに関する指針
- 7. ヒトES細胞の樹立及び使用に関する指針
- 8. 遺伝子組換え生物等の使用等の規制による生物の多様性の確保に関する法律
- 9. 遺伝子治療臨床研究に関する倫理指針
- 10. 疫学研究に関する倫理指針
- 11. 臨床研究に関する倫理指針
- 12. ヒト幹細胞を用いる臨床研究に関する指針
- 13. 動物の愛護及び管理に関する法律
- 14. 実験動物の飼養及び保管並びに苦痛の軽減に関する基準
- 15. 研究機関等における動物実験等の実施に関する基本指針
- 16. 医薬品の臨床試験の実施の基準に関する省令
- 17. 医薬品の臨床試験の実施の基準に関する省令の一部を改正する省令

臨床・疾患データベースの統合化に関するアンケート調査票

東京医科歯科大学 統合データベース整備事業「疾患DB実情調査」
文部省統合データベースプロジェクト
ライフサイエンス分野の統合データベース整備事業

平成19年10月吉日

平成19年度文部科学省「ライフサイエンス分野の統合データベース整備事業」
アンケートへの回答のお願い

文部科学省「ライフサイエンス分野の統合データベース整備事業」分担機関
東京医科歯科大学 情報医科学センター長 田中 博

拝啓 時下ますますご清栄のこととお慶び申し上げます。

東京医科歯科大学では我が国のライフサイエンス関係のデータベースの統合的な利用環境の構築を図るため文部科学省の委託を受け、平成19年度より「ライフサイエンス分野の統合データベース整備事業」を実施しています。(http://lifesciencedb.mext.go.jp/)

東京医科歯科大学は本事業の「臨床疾患統合データベース開発」を担う分担機関として、散在する臨床・疾患データベース(DB)を統合的に利用できる環境を開発して、分子から個体にいたる多階層的で疾患横断的な臨床医学データの有効利用を広く可能にする総合的な利用システムの開発を進めています。

本アンケートは、我が国の臨床・疾患データベースの現状把握と統合化のための意見収集を目

受信... Micr... (無題... 2 S... Adob... あ般 CAPS KANA

非公開情報の調査対象(詳細)

外部公開情報の収集

方法
・インタ
検索
・関連機
入手

医学系学会

・調査対象...日本医学会(1件)及びその分科会として加盟している**医学系学会102件**

・回収数: 現在のところ、**50件(49%)**

・調査期間...郵送によるアンケート期間
平成19年10月31日～11月12日

WWW上でのアンケート期間
平成19年11月13日～11月21日

DB
開状況一覧
(研究所)

データのまとめ

方法
1. 調査項目に沿って、各データベースについて(二次データ)を収集しまとめる

非公開情報の独自調査

方法
・アンケート郵送調査
・アンケートWEB調査
・アンケートFAX調査
・ヒアリング調査(一部)

対象
・医学系学会
・大学付属病院
・国立病院

国公立病院/国公立大学付属病院

・調査対象...WWW上にホームページを公開している国公立病院/国公立大学付属病院**300件中**、メールでコンタクトできた**139件**

・回収数...現在のところ、**10件(7.2%)**
(診療科長宛で再調査予定)

・調査期間...WWW上アンケート期間
平成19年11月22日～12月4日

| 国内の臨牀・疾患関連データベースの実態調査 | | | | | | | | | | | | | | | 東京医科大学 | | | | | | | | | | |
|-----------------------|-------------------|-----------------|-------------------|--|---|------------------------|-----------------|-----|-------|------|--------------|-------|----------|---------|-------------------|------------------------------------|------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|-------|-------|-------------|-----------|-----------------|
| | | | | | | | | | | | | | | | 最終更新日: 2023年12月1日 | | | | | | | | | | |
| No. | 研究機関名称 | 部門/学部 | DB名称(別名) | DB名称(本名) | URL | 提供機関 | 作成経緯 | 備考 | 調査方法 | 件数 | 内容 | 方式 | データベース形式 | システム | アクセス手段 | 公開範囲 (注: 公開範囲が不明なものは「非公開」として記載) | 権限管理 (注: 権限管理が不明なものは「非公開」として記載) | 個人情報 (注: 個人情報管理が不明なものは「非公開」として記載) | 多国籍連携 (注: 多国籍連携が不明なものは「非公開」として記載) | 標準化連携 (注: 標準化連携が不明なものは「非公開」として記載) | 互換性 | 匿名化 | クラウドサービスの有無 | データベースの中核 | 備考 |
| N030 | 国立中央総合研究所 | メイズ研究センター | 国立中央総合データベース | NCDB | https://ncdb.nih.go.jp/index.html | 国立中央総合研究所 | 研究機関主導の共同データベース | | Web調査 | 3000 | 1999年2月21日現在 | クラウド型 | Oracle | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N071 | 国立がん研究センター | 国立がん研究センター | 国立がん研究センターデータベース | COG Cancer Gene Expression Database | http://biocancer.jp/ | 国立がん研究センター | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N075 | 北海道大学 | | | | | | | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N077 | 大阪府立成人病センター研究センター | | | | | | | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N078 | 東京大学 | | | | | | | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N079 | 名古屋大学 | | | | | | | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N084 | 九州大学 | | | | | | | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N086 | 名古屋大学 | | | | | | | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N083 | 国立がんセンター研究所 | がんゲノム・ゲノム情報センター | がんゲノム・ゲノム情報データベース | Genome Medicine Database of Japan (GMDJ) | https://gmdj.nih.go.jp/ | 国立がんセンター研究所 | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N132 | 京都大学 | 医学部 | ヒト胎児発育データベース | Human Fetal Development Database | 公開予定(予定2019年10月) | 京都大学 医学部 | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N050 | 大阪大学 | 医学部 | 日本全国がんデータベース | Japanese National Cancer Database | http://www.jncdb.jp/ | 大阪大学 医学部 | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N136 | 北里医科大学 | 医学部 | 調査中 | 調査中 | 調査中 | 調査中 | 調査中 | 調査中 | 調査中 | 調査中 | 調査中 | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N064 | 日本循環器学会 | | 心臓病診療ガイドラインデータベース | Cardiac Disease Registration (CADR) | https://www.ncc.or.jp/CADR/ | 主任研究員 国立循環器病センター心臓内科 | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N083 | 京都大学 | 医学部 | 遺伝子発現データベース | | http://www.kuh.kyoto-u.ac.jp/genome/DB/index.html | 京都大学 医学部附属病院 遺伝子発現センター | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N061 | 日本脳神経科学会 | | 日本人脳血管中枢神経系データベース | Japan Adult Cardiovascular Stroke Database | http://www.jacsd.univ.jp | 国立 | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N147 | 日本脳神経科学会 | | 脳血管中枢神経系データベース | | http://accsdb.univ.jp/ | 研究機関主導 | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N152 | 国立がんセンター研究所 | がんゲノム・ゲノム情報センター | がんゲノム・ゲノム情報データベース | Genome Medicine Database of Japan (GMDJ) | http://gmdj.nih.go.jp | 国立がんセンター研究所 | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N143 | 京都大学 | 医学部 | 遺伝子発現データベース | | http://www.kuh.kyoto-u.ac.jp/genome/DB/index.html | 京都大学 医学部附属病院 遺伝子発現センター | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N144 | 日本脳神経科学会 | 脳血管 | 脳血管中枢神経系データベース | | | 研究機関主導 | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |
| N153 | 国立循環器病研究センター | 循環器病研究センター | 循環器病研究センターデータベース | | | 研究機関主導 | 研究機関主導の共同データベース | | Web調査 | | | クラウド型 | MySQL | Webブラウザ | 2020年10月14日現在 | 一部公開 | 権限管理あり | 個人情報管理あり | 多国籍連携あり | 標準化連携あり | 互換性あり | 匿名化あり | クラウドサービスの有無 | データベースの中核 | 研究機関主導の共同データベース |

• 内容まで良く見て臨床DBかどうか判断した

• 各DBの中の内容を可能な限りよく見て吟味した

• アカウントの取れるものは、実際に使用してみた

• 相手に十分説明し相手との意識合わせをしたうえで、質問に回答していただいた

間違いなくこれから発展してゆく分野なので期待している(数件)

要素疾患DBの疾患データベースへの 統合化の条件

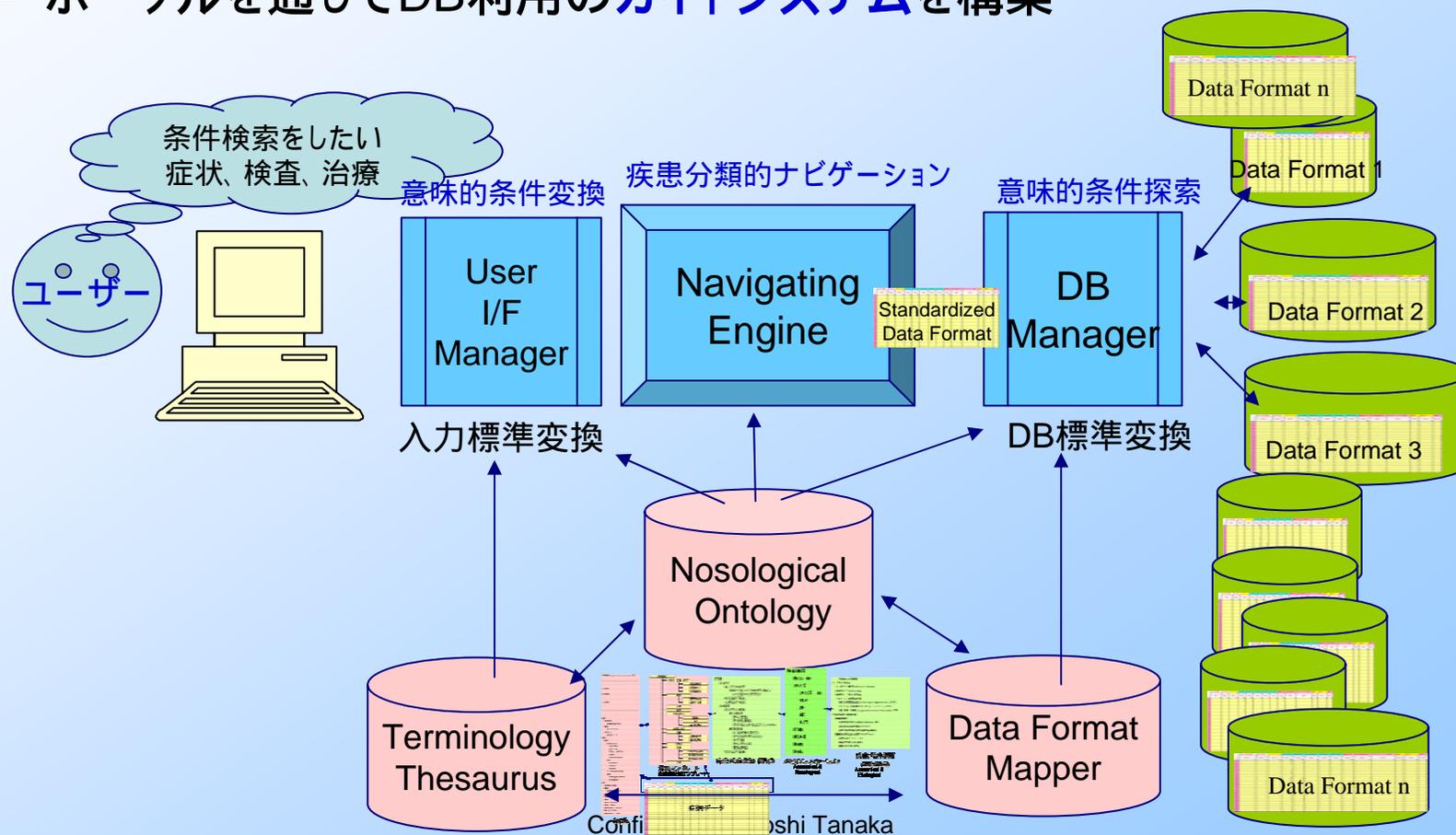
- 疾患データの妥当性 (Validity)
 - 疾患データベース内の情報は有効か
- 疾患データの更新性 (Renewality)
 - 構築後も更新されているか
- 倫理的保証 (Ethical Admission)
 - 患者の公開ICは取れているか

情報倫理面について

- 国内臨床疾患DB調査の中で、倫理面も調査
 - ポリシー（公開基準、倫理規定、個人情報保護基準）
 - 匿名化（連結不可能性、個人情報保護基準具体的方法）
- まずは、2DB統合（医科歯科、阪大）における倫理案を検討中
 - 統合公開時（3年後）までに完全案作成予定
 - 関係法令の動向とも歩調を合わせる
 - 第三者利用の法整備
 - HIPPA法 など

統合医科学DB利用ガイドシステムの基本設計

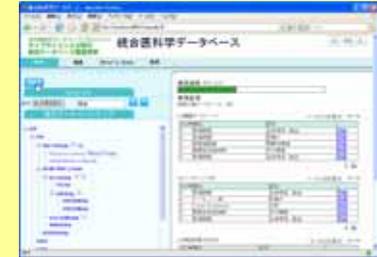
- コア疾患DB (2個 + (数10個程度)を想定 : 全国調査により選定)
 - 統一的な検索を可能とする**統合利用システム**を構築
- 周辺疾患DB (数10個を想定)
 - ポータルを通じてDB利用の**ガイドシステム**を構築



検索条件指定: **症状**
 キーワード入力: **貧血**



標準的な表現での結果表示



2

裏処理: ターミノロジーの標準化

- 疾患オントロジーの「**症状**」項目の中に、「**貧血**」とある疾患を検索。
- ターミノロジー、シソーラスおよび疾患オントロジーより「**貧血**」の別の表現を検索。」

他の項目に「**貧血**」がある場合は無視する。

3

- 「**疾患:再生不良性貧血**」 or 「**疾患:アレルギー性胃腸症**」 or 「**疾患:潰瘍性大腸炎**」などがヒットする。
- 「**Hbが一定値以下**」がヒットする

Navigating Engine

Nosological Ontology

4

裏処理: 標準検索条件を各DB毎の検索条件に変換

- 疾患オントロジーの検索結果より標準検索条件を作成:
 ダイレクト検索条件: 「**症状:貧血**」
 セマンティック検索条件: 「**疾患:再生不良性貧血**」 or 「**疾患:アレルギー性胃腸症**」 or 「**疾患:潰瘍性大腸炎**」 or 「**Hbが一定値以下**」
 etc

裏処理: 検索結果の集計

裏処理: 対象となっているDBを選定し各要素DBの検索

統合医科学データベース - Mozilla Firefox

ファイル(E) 編集(E) 表示(V) 履歴(S) ブックマーク(B) ツール(T) ヘルプ(H)

http://localhost:8080/tougoudb/

文科省統合データベースプロジェクト
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

S M L

検索 概要 What's New 管理

1 検索条件の指定

Search

Search

病名 条件項目表示

統合データベースマップ

2 疾患オントロジー分類におけるDB表示

血液

- 貧血
 - 再生不良性貧血
 - Fanconi's anemia
 - dyskeratosis congenita
 - 鉄代謝の異常による貧血
 - 鉄欠乏性貧血
- 乳児の生理的貧血
- 鉄芽球性貧血
- 巨赤芽球性貧血

赤血球

白血球

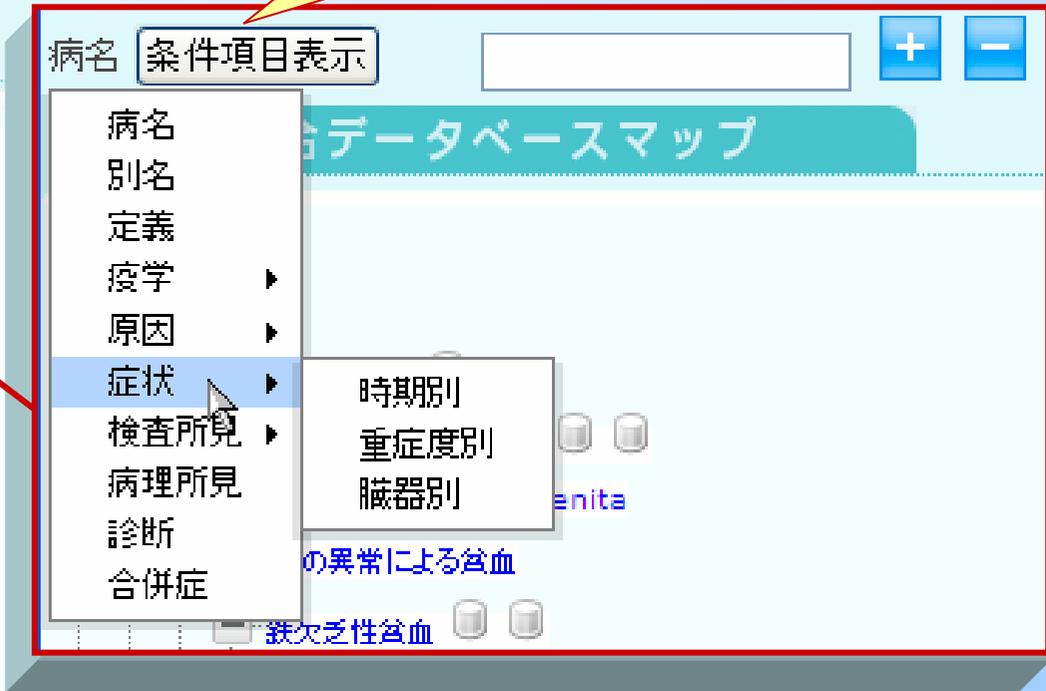
完了

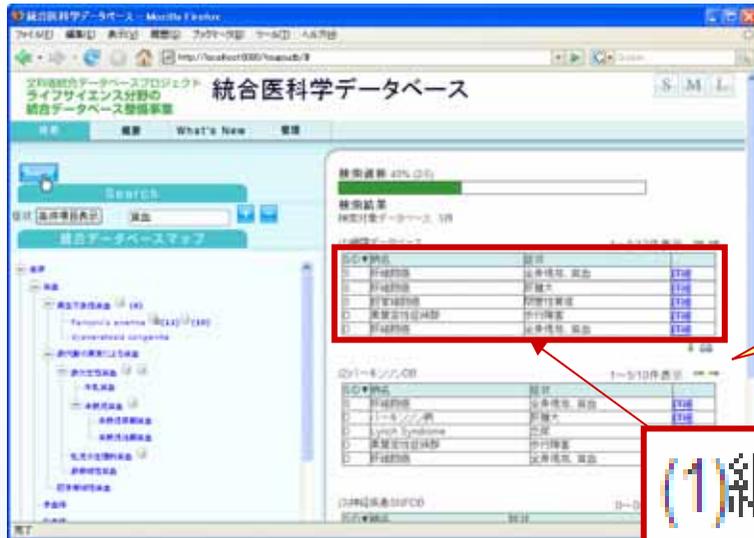
3 結果表示フレーム

検索進捗 0% (0/0)

検索結果
検索対象データベース 0件

検索条件項目の表示と指定





検索対象となったDB毎に結果が表示される。

(1) 網羅データベース

| S/D▼ | 病名 | 症状 |
|------|---------|------|
| S | 肝細胞癌 | 全身倦怠 |
| S | 肝細胞癌 | 肝腫大 |
| S | 肝細胞癌 | 肝腫大 |
| D | 黒質変性症候群 | 歩行障害 |
| D | 肝細胞癌 | 全身倦怠 |

Semanticな検索結果の場合、その理由が表示される。

Semanticな検索が実行されたかどうかがわかる。
S : semanticな検索
D: Directな検索

オントロジーから生成した検索条件を利用
例: Hbが一定値以下となっている症例

ダイレクト検索の結果

症状として貧血のある症例

計画の基本方針

(中核機関との協議・検討結果に基づく)

- 19年度の公開は、各機関ごと
 - 本年度の倫理案は、各施設ごとの倫理基準を基にする
 - 入り口は東京医科歯科大学の統合医科学DBサーバーの予定
- 課題拡大への対応の必要性について双方認識した
 - コア疾患DBの統合的利用システム構築対象の拡大
 - 対象疾患: 2疾患系統 多疾患系統(26系統)
 - 統合化対象DB数: 2DB 多DB(45DB)
 - 統合化参画施設: 2施設 多施設(45施設)
 - 事業規模の拡大への対応(要素DB統合化追加予算)
- 統合方式の基本方針について、了解を得た
 - 3レベルメソッド・2系統の統合アプローチ
 - 統合医科学DB利用ガイドシステム

ありがとうございました