
統合データベースセンターへの期待

ー産業界からー

協和発酵キリン株式会社

相談役 平田 正

1. 産業界の取り組み
2. 背景
3. わが国の現状と課題
4. DB公開の原則
5. ナショナルセンターの意義
6. その必要な機能
7. 統合DBの機能
8. 個人情報への取り扱い
9. バイオインフォマティクス研究と人材の育成
10. 提言(まとめ)

- バイオテクノロジー(BT)戦略会議での提言
 - ✓ 2002 BT戦略大綱
 - ✓ LS/BTのインフラ整備として課題提起
 - ✓ 2005 DBナショナルセンターの提言・・・連携施策群プロジェクトへ

- バイオ産業情報化コンソーシアム(JBIC)を代表して
 - ✓ 4省共管 100社参加 最大のコンソーシアム
 - ✓ FL-cDNAライブラリーの活用
 - ・「ゲノム情報統合プロジェクト」H-invitational
 - ・疾患遺伝子探索
 - ・蛋白構造解析
 - ・蛋白相互作用(ケミカルバイオロジーへ)
 - ・機能性RNA
 - ・トランスレーショナルリサーチ(TR)
 - ✓ BI人材育成
 - ・バイオインフォマティクス技術者認定制度 2級検定試験

- 現代のライフサイエンスの特徴＝“Database Driven Science”
 - ✓ 膨大なデータの蓄積とその利活用
 - ✓ 研究開発の競争力強化やスピードアップに統合DBは必須
 - ✓ DBの高度なインフォマテクスから新しい価値情報の創出
 - ✓ 個別化医療への必須な情報インフラ

- 大量情報時代への突入
 - ✓ 次世代シーケンサーの出現
 - ✓ 個体レベルのデータ量増大への対応
 - 大量データのストレージ
 - 大規模計算処理
 - 大量データの圧縮・転送技術
 - 臨床情報の円滑な活用(個人情報保護システムの整備)

- 恒久的なナショナルセンターがない
 - ✓ データが散在していて検索しにくい
 - ✓ データフォーマットが標準化されていない
 - ✓ プロジェクト終了や研究者の移動等でDBの維持管理が不十分
- 公的資金で取得したデータの公開に関する取り決めが不明確
 - ✓ 情報開示が不十分、抱え込み、死蔵
- 欧米への過度な依存 NCBI, EBI
 - ✓ 国際貢献度の低さ、「ただ乗り批判」 情報安全保障問題
- 臨床DBが未整備／アクセス困難
 - ✓ 創薬クリティカルパスや個別化医療への流れに対応できない
- バイオインフォマティクス研究の遅れ
 - ✓ 大量計算処理に対応できる研究環境の整備の遅れと人材不足

- 登録(公開)を推進・徹底する制度
 - ✓ 公費研究からの一次情報は、例外なくタイムリーに公開されるべき
 - ✓ 企業が参加したプロジェクトも例外にすべきでない (一定の猶予条件付きで)
- データの囲い込みに対するルール作りが必要
 - ✓ 研究契約の中に一次情報の公開に関する条項を盛り込む
 - ✓ 後継プロジェクトに継承時も、公開時期が引き延ばされるべきでない
 - ✓ 開示で研究の質の担保
- 産業化、特許化、倫理的配慮はデータ公開を妨げない
 - ✓ 一定の猶予期間を設ければ、特許取得を妨げることはない
 - ✓ 個人レベルの情報の研究目的の収集・保存・公開のための指針策定や差別禁止などの法整備が必要

“公費研究は税金から、アウトプットは国民共有の資産”

➤ 意義

- ✓ 省庁の壁を超えたオールジャパン体制
総合科学技術会議(内閣府)所管
「国立電子図書館」といったようなもの
- ✓ 恒久的な統合データベースサービス
競争的資金による時限プロジェクトでは継続性担保されない
- ✓ 国際貢献
米国、欧州に次ぐ第3のアジアの情報拠点となるもの
- ✓ 情報安全保障 国際競争力の要

➤ 必要な機能

1. データベースセンター（情報サービス業務）

- ✓ 大規模計算機センター
- ✓ データストレージセンター
- ✓ DB受け入れ、品質管理、公開ならびに維持管理・アップデート

2. インフォマティクスセンター（研究機能）

- ✓ 新たな価値を生み出すための高度なバイオインフォマティクス研究
- ✓ 情報を取り扱うための技術開発
- ✓ 人材育成 しっかりしたDBセンターこそ最良のインキュベーター

➤ 情報の網羅性

世界への情報発信を通じ、アジアの情報拠点として国際貢献にも繋がるもの

- ✓ 過去のプロジェクトや進行中のプロジェクトも含めた網羅性
- ✓ 独自にDB公開している研究機関の一次情報も受け入れ(取り込み)、一元的に検索できるようにする
- ・ 単独でDB作れないPJの受け皿 データデポ
- ✓ 民間資金で取られたデータの公開も受け入れる
- ✓ データの品質管理で信頼性確保

➤ 機能

Wet研究者、Dry解析専門家、医者、一般国民などの幅広いユーザーのニーズに合ったインターフェースや解析プログラム

- ✓ わかりやすい統一された操作感
- ✓ DB間の横断検索機能
- ✓ 強力な検索機能
- ✓ 日本語も含めた専門用語の同義語(シソーラス)辞書の整備
- ✓ ウェブインターフェースでの解析
- ✓ データ配布機能(FTPなど)
- ✓ アノテーションの充実

➤ 物理的存在場所

一つのDBとして検索やデータ取得ができれば、必ずしも物理的に一箇所にデータをまとめる必要はない

- 個人情報保護や倫理的配慮はデータ公開の前提
 - ✓ 研究目的の個人レベル情報の収集・保存・公開のため法律／指針整備 (cf. HIPPA)
 - ✓ 個人の遺伝子情報に基づく差別(保険・就学・就職)を防ぐ法整備 (cf. GINA)
- 匿名化と個人情報保護
- 二段階の公開ルール
 - ✓ 匿名化され集団として解析された結果は、ほぼ無制限に公開
 - ✓ 個人レベルの情報を含む情報は、契約を交わした特別の研究者にのみに閲覧や利用を限定

統合データベース推進本部に設置の部会で早急に方針を決定すべき

- **新たな価値を生み出すための基盤研究**
 - ✓ 新たな発見のための方法論やアルゴリズムの開発
 - ✓ 日本語を含む専門用語の同義語(シソーラス)辞書の整備
 - ✓ 異なる質や種類の情報を統合するための情報規格化法の開発
- **情報の統合による高付加価値の創出**
 - ✓ 新たな付加価値を生む二次・三次データベースの構築
- **情報を取り扱うための技術開発**
 - ✓ 統一された操作感や複雑なデータを統合するためのデータフォーマット
 - ✓ 初めてのユーザーにもわかりやすいインターフェース
 - ✓ 次世代DBのためのコンピュータ言語や強力な検索エンジンの開発
 - ✓ DB品質管理手法の開発
- **教育・広報活動**
 - ✓ オールジャパンでWetとDryの対等なコラボレーションができる環境作り

- 恒久的なナショナルセンターの設立
 - ✓ 省庁の壁を超えたオールジャパン体制の組織
 - ✓ 継続的なデータベース維持管理とアップデート
 - ✓ 予算の確保(例:研究予算の1%~を割り当てる)
 - ✓ 国際貢献できるアジアの拠点構想

- 公的資金によって取得された一次情報の公開義務 ルールの徹底

- 個人レベルの情報の利用に関する指針と差別禁止に関する法整備

- 高度バイオインフォマティクス研究と人材育成

**DBナショナルセンター設立スケジュールの前倒しを！
(あるべきオールジャパン体制へ最速で)**