

# 「ライフサイエンス統合データベース基盤整備」プロジェクトの事業化計画(案)

情報・システム研究機構事務局

## (1) 統合化およびDB利便性向上のための動向調査・戦略立案支援チーム

- ・ライフサイエンス(医療、農業などの応用も含む)、DB、情報技術の動向調査
  - ・ライフサイエンスプロジェクトの現状、動向調査
  - ・上記に基づく報告書作成、戦略立案、情報提供(ホームページ、ポータルサイト他にて)
  - ・戦略立案委員会の事務局機能
  - ・各省庁および関係機関・団体との連携のための連絡調整機能
  - ・DB利用者、産業界との協議会の事務局機能
  - ・出版社や学会およびDBサービス業者との連携のための連絡調整機能
  - ・外国の関連機関との連携のための情報提供と連絡調整機能
  - ・下記のプロジェクトチームの連絡調整統合機能
  - ・医療データの収集と統合化のフィージビリティスタディ(外部機関と連携して)
  - ・その他の新興分野のデータの収集と統合化のフィージビリティスタディ(外部機関と連携して)
- (フィージビリティスタディ終了後は、下記の(2)あるいは(3)のチームで統合化実施)

## (2) ヒト統合DB開発チーム

- ・論文、教科書からの知識抽出システム開発とそれによるヒト知識の整理
- ・ヒトデータ統合化のためのインデックスや各種辞書の構築
- ・細胞、組織、器官、個体などの高次レベルの整理棚構築
- ・知識整理や辞書構築のための専門家(がん、免疫、脳など)集団からの意見集約と事務局機能
- ・文献情報(おもに日本語解説記事や特許情報)との連携による情報提供
- ・医療、医薬品に関するデータとの連携(将来的にはこれらも統合)
- ・(知識、機能を中心とした)ヒト統合DB構築とその運用・公開
- ・既存の(構造、測定データを中心とした)ヒト統合DBとの連携による相互運用性の向上
- ・ヒトゲノムブラウザの構築・運用・公開(一部外部委託、外国との連携)
- ・そのための高精度アノテーション実施(一部外部委託)
- ・医療への応用、創薬との連携のためのソフトウェア開発(一部外部委託)

## (3) モデル生物統合DB、産業応用生物統合DB開発チーム

- ・論文、教科書からの知識抽出システム開発とそれによる各モデル生物、産業応用生物知識の整理
- ・これらの生物の遺伝子名や機能の辞書の構築
- ・比較ゲノム解析、相互運用性向上のための整理棚の構築
- ・それぞれの生物毎の専門家集団(発生、微生物、植物、など)からの意見集約と事務局機能
- ・文献情報(おもに日本語解説記事や特許情報)との連携による情報提供
- ・(知識、機能を中心とした)モデル生物、産業応用生物それぞれの統合DB構築とその運用・公開
- ・生物種を横断した比較解析用統合DB構築とその運用・公開
- ・それぞれの生物毎のゲノムブラウザの構築・運用・公開(一部外部委託、外国との連携)
- ・そのための高精度アノテーション実施(一部外部委託)
- ・産業応用との連携のためのソフトウェア開発(一部外部委託)

#### (4)ポータルサイト運用・高度検索サービス提供チーム

- ・ポータルサイト、高度検索サービスの維持・管理・更新・運用・公開
- ・DBサービス・解析サービスサイトに関する最新情報の自動収集エンジンの開発、運用
- ・上記のサービスのために必要なインデックスや用語の収集とその自動化技術開発
- ・データベースやソフトウェアの利用法に関する情報の収集・編集・提供およびその自動化技術開発
- ・ポータルサイト構築のための専門家の意見集約システムの開発、運用
- ・利用者の意見集約(モニター・ランキングシステム)の開発、運用
- ・検索のパーソナル化(利用者の興味、レベルに応じた情報提供)とそのための技術開発
- ・情報提供の日本語化とそのための対訳辞書、技術開発
- ・データベース、ソフトウェアの利用法等に関する相談窓口
- ・他のポータルサイト、情報提供機関との連携

#### (5)情報技術開発チーム

- ・統合DB構築技術、DB標準化技術、DB相互運用性向上技術の開発
- ・文献検索・クラスタリング・ランキング・要約技術の開発、高速検索・情報収集エンジンの開発
- ・論文、教科書からの知識の表現、抽出、利用に関する技術の開発、オープンアクセス化への対応
- ・オントロジー(辞書、インデックス、専門用語)自動構築技術の開発
- ・画像、ポンチ絵、図表など文字以外のデータからの知識発見とアノテーション技術の開発
- ・異種データベースからの知識発見、機械学習技術の開発(一部外部委託)
- ・情報共有、情報交換のためのWEB技術開発(一部外部委託)
- ・情報の暗号化、個人情報の保護技術の開発(一部外部委託)
- ・新開発測定装置から出てくる新規データの解析・格納法の開発(一部外部委託)
- ・ワークフロー技術、並列処理技術、シミュレーション技術との連携
- ・DBの新たな利用法の開発の支援(現在JSTBIRD)
- ・他のプロジェクトチームやDBサービス業者への技術・ソフトウェア提供

#### (6)広報・教育・国際対応・産学連携チーム

- ・センター全体のホームページ(日本語、英語)の構築、運用
- ・我が国の活動や研究者に関する情報の外国への提供(英語)
- ・教育プログラム、教材の開発
- ・ニュース記事、コラムの作成(一部外部委託)
- ・講習会、シンポジウムの開催
- ・ニュースレター、パンフレットの発行
- ・国際対応、産学連携の窓口機能
- ・キュレーター・アノテータの養成(一部外部委託)
- ・DB管理者、DB構築者の養成(一部外部委託)
- ・DB高度利用者の養成(一部外部委託)

#### (7)支援チーム

- ・基盤的なDBの支援(現在JSTBIRD)とそれらとの相互運用(一部外部委託)
- ・プロジェクトDBの受け入れと相互運用可能DBへの変換と運用・公開
- ・新規DBの開発支援(現在JSTBIRD)とそれらの相互運用(一部外部委託)
- ・国内外の主要DBの更新維持管理(おもに内部利用が目的)
- ・知財(著作権や特許等)の管理

- ・外部DB受け入れに伴う権利関係などの処理
- ・民間との共同研究等の契約業務
- ・DBやライブラリの契約業務
- ・計算機やネットワークの導入・維持管理
- ・一般事務機能(経理、人事、庶務、契約、施設、安全衛生管理など)

(注)

・既存のサービス(BLAST 検索など)が確立しているものは本プロジェクトでは基本的に行わない。  
それに関してはそちらにサービスを投げ、その結果に付加価値をつけて返す。

・構造データの登録業務は行わない。機能データの整理は行う。機能データの登録業務は今後検討。