



図4 時間軸

## 8. 検討の経緯

### 8.1 有識者インタビュー

以下の有識者から提案の構想段階で有益なご意見を伺った。

- ・開原成允（国際医療福祉大学 大学院長）
- ・池田俊也（国際医療福祉大学 教授）
- ・堀田凱樹（情報・システム研究機構 機構長）
- ・高木利久（ライフサイエンス統合DBセンター センター長）
- ・井原茂男（東京大学 先端科学技術研究センター システム生物学ラボラトリ 教授）
- ・田中博（東京医科歯科大学 大学院生命科学研究部 教授）
- ・大江和彦（東京大学 大学院医学系研究科 医療情報経済分野 教授）
- ・真鍋一郎（東京大学 大学院医学系研究科 疾患生命工学センター 特任講師）
- ・興紹貴英（東京大学 大学院医学系研究科 22世紀医療センター 助教）
- ・中村祐輔（東京大学 医科学研究所 ヒトゲノム解析センター長 教授）
- ・豊田哲郎（理化学研究所 生命情報基盤研究部門 部門長）
- ・安達淳（国立情報学研究所 学術基盤推進部長 教授）

### 8.2 ワークショップ

科学技術未来戦略ワークショップ「トランスレーショナル・インフォマティクス・ベースの展開」を2010年1月18日（月）13:30～18:30 研究開発戦略センター2F大会議室にて開催した。少人数の有識者を招聘し、提案の内容を補強する目的で実施した。

<ワークショップ参加者>

- ・安達淳 国立情報学研究所 学術基盤推進部長 教授
- ・大江和彦 東京大学 大学院医学系研究科 医療情報経済分野 教授
- ・田中博 東京医科歯科大学 大学院生命科学研究部 教授
- ・豊田哲郎 理化学研究所 生命情報基盤研究部門 部門長
- ・永井良三 東京大学 大学院医学系研究科 教授

<WSプログラム>

13:30～14:00 オープニング

- 連絡事項・資料確認など（CRDS）

- 提案趣旨説明 (CRDS)

14:00~15:00 セッション1 TIB<sup>4</sup>に求められること (情報循環に関して、医療の現場から)

- 永井良三 (医学・医療における知の循環)
- 大江和彦 (医療データの標準化)
- 田中博 (オミックス医療とデータベース)

(休憩)

15:20~16:00 セッション2 TIBの方向性

- 安達淳 (学術共有クラウド)
- 豊田哲郎 (理研サイネス)

16:00~17:30 セッション3 総合討論 (CRDS)

- データ収集、解析に関する課題
- 制度的課題
- 人材育成
- ファンディング
- 研究推進の方法、時間軸
- etc.

17:30~18:30 まとめ

<WSで主な議論>

#### 1. TIBの目指す方向

- 所有するITから利用するITへという観点からクラウド技術が注目を集めている。研究者はデータを所有したいわけではなく、データを利用して研究したい訳であるから、研究に専念できる環境という意味でもクラウド化していく方向であろう。
- バーチャルラボやバーチャルコミュニティの実現により、グローバルな研究環境を構築でき、世界の知を集めることも可能となる。

#### 2. より効率的、持続的なデータ収集、保管、活用のあり方

- データ提供者、病院およびそのデータを利用する研究者、研究機関など多様なステークホルダーのインセンティブ (メリット、デメリットなど) を整理し、課題を見つけ出し、そこを解決するシステム設計をする必要がある。
- データの標準化に関しては、メタデータの標準化がむしろ重要である。

#### 3. 研究課題の推進体制ならびに資金配分のあり方

- インフラとしての研究環境へのファンディングのあり方に関して、期

<sup>4</sup>本文中では、THIBとなっているが、ワークショップ開催時点では、TIB (Translational Informatics Base) という呼称で議論した。

限のある競争的資金での継続的運用は難しく、競争的資金ではないファンディングの枠組みを検討する必要がある。

- 推進体制に関しては、疾患別プロジェクト、病院での疾患収集など、いくつかの方法が考えられる。
4. 研究課題の実施に必要な人材育成のあり方
- トランスレーショナルインフォマティクスの人材育成は既存の大学組織の中ではサポートを受けにくい。現状では、競争的資金に依存せざるを得ない。境界領域的人材の養成、雇用、活動を支援し、講座に定着させる仕組みが不十分。プロジェクトの切れ目におきるキャリアパスの中断は大きな問題である。
  - 米国の大学では競争的資金が切れても一定期間研究者をPI (Principal Investigator) が雇うことができる資金制度がある
5. T I Bを恒久的に維持、発展させるための国民の理解、評価のあり方
- データベースがどのように役立っているか、ということをおある程度定量的に呈示しないと、T I Bを維持管理する妥当性や資金の必要性が説明しにくいと考えられる。
  - T I Bに必要な研究領域を包括して学術領域（新学術領域）として確立させることが重要である。メカニズム解明にとどまらない研究領域としての設定が必要。この分野はうまくいけば「自己収入」が可能であることもアピールのポイントなる。



## 9. 国内外の状況

### 9.1 国内の状況

ライフサイエンスデータの統合化の動きは、第3期科学技術基本計画の分野別推進戦略「ライフサイエンス研究全体を支える基礎・基盤課題」の中でも重要なテーマとして認識され、大量データ時代の到来にあわせ、施策として加速している。その中の主要プロジェクトである、統合データベースプロジェクト（平成18年度～平成22年度；平成21年度政府予算額9億円 平成20年度11億円）は、我が国のライフサイエンス関係のデータベースの利便性の向上を図るため、データベースの統合化及び利活用のための基盤技術開発、人材育成等を行い、データベースの統合的活用システムを構築することを目的とし、現在、(独)情報・システム機構を中心とした関連中核機関が積極的なデータ統合化を進めている。また、バイオインフォマティクス研究の推進（平成21年度政府予算額18億円、平成20年度17億円）が(独)科学技術振興機構で実施されており、解析の効率化・省力化、利用の高度化等を実現するための革新的なゲノム解析ツールの研究開発等、ファンディング機能も活用してバイオインフォマティクス研究を行っている。上記2つのプロジェクトは平成23年度より統合化されることが決まっており、現在その準備段階にある。

公共的なデータ提供ソースとしては、(独)理化学研究所を中心に、バイオリソース事業（平成21年度政府予算額32億円+補正33億円 平成20年度32+5億円）が我が国の中核的拠点として、生物遺伝資源（バイオリソース）の提供を行うとともに、基礎データの蓄積も含めた基盤技術の開発を進めている。データとそれに伴うバイオリソース両者を保有する機関である、ライフサイエンス研究全体の基盤拠点として期待される。

さらに医療分野に特化した大型プロジェクトとして、バイオバンクジャパン計画（個人の遺伝情報に応じた医療実現化プロジェクト）(第1期 2002-2007年度 2002年度は補正、計220億円 第2期 2008年—2012年（予定）平成21年度予算額27億円）が進捗している。通称「オーダーメイド医療実現化プロジェクト」として、2期目に入っており、東京大学医科学研究所に設置されたバイオバンクに収集された患者DNAサンプルを活用し、(独)理化学研究所がSNP（遺伝子の個人差）と薬剤の効果、副作用などの関係や疾患との関係を明らかにする研究等を実施、バイオバンクに収集された疾患について発症に関与する遺伝子や発症メカニズムの解明を目指し疾患関連遺伝子研究の加速を目指している。

いずれも、信頼性、継続性及び先導性の確保と共に戦略的かつ効率的な整備の促進が重要であり、データの活用、研究成果や基盤技術の普及、人材育成、国内外の関連機関との有機的な連携等が必須であり、このために

も、本戦略プログラムで提案のTHIBの構築が重要となる。

## 9.2 海外の状況

バイオインフォマティクス関連研究に対する公的資金サポートの状況について以下、米国、英国、欧州、アジアに関する調査結果を述べる。

### <米国>

米国では、国立衛生研究所 (National Institutes of Health, NIH) 傘下の国立研究資源センター (National Center for Research Resources, NCRR)内において、2001年からBiomedical Informatics Research Network (BIRN)という研究プログラムが開始されている<sup>5</sup>。

BIRNは全米各地の提携機関の協力を得て医科学研究者向けにソフトウェアツールとインフラリソースを提供している。現在は脳科学に関連した、ヒト脳のMRI画像と医療情報の連携検索を可能にするFunctional Imaging BIRN Tools、マウス脳解剖図のデジタルアトラスを扱う Mouse BIRN Tools、そしてヒト脳の形態計測画像の管理、保管、検索を行うMorphometry BIRN Toolsを提供している。Functional Imaging BIRN Toolsにおいては、被験者(患者) 各種のMRI計測画像と医療情報をURLのリンクによって結びつけることによって、アルツハイマー病、躁鬱病、多動性注意欠陥障害、自閉症などの神経疾患をもつ脳の機能的な特徴の抽出、分析が可能になっている。これらのツールを利用するのは脳科学研究従事者に限らず、他の疾患研究に携わる研究者、技術者、医療従事者も多いことがわかっている。

米国に拠点を置くデータベース系企業では、診療データを元に患者の治療フローを分析し、実際の医療プロセスとそのアウトカムを比較することで医療の質の向上、医療経済的な評価に役立てている。

また、2009年のAmerica Recovery and Reinvestment Act (アメリカ再生・再投資法) のなかにはHealth Information Technology for Economic and Clinical Health (HITECH) Actが含まれ、約20億ドルが医療情報のインフラ整備や導入に費やされることが決定している。

<sup>5</sup> [http://www.nbirn.net/index\\_ie6.shtml](http://www.nbirn.net/index_ie6.shtml)

新規プログラム	内容	投資額 (百万ドル)
地域普及センター	医療情報技術の取り扱い等を医療提供者に普及させるための機関を全米で70機関設立	643
健康情報交流 人材育成プログラム	地域間での健康情報のやり取りを支援 健康情報の専門化の育成など。4万5千人の育成を目標としている	564 118
健康情報技術の高度化に向けた研究	情報の機密性、ネットワークプラットフォームの構築など健康情報技術の高度化を目指した研究開発の支	60
全米健康情報ネットワークなど	健康情報基盤の構築、整備のための投資	64
その他含めた総計		2000

(New England Journal of Medicine 362:382, February 4, 2010より抜粋)

この法律によって、健康情報の電子化がなされ、アメリカ国民の健康やアメリカの医療制度を改善することが目標とされている。また、こうした目標の達成には国民の参加や医療従事者による情報の適切な活用が必要とされており、これに向けた技術開発やインセンティブの付与の仕組みづくりが進められている。

我が国でもEHR (Electronic Health Record) の整備など、THIB構築とともに実施していく必要がある。

#### <欧州>

欧州の主な資金配分組織は欧州委員会のSeventh Framework Program for Research and Technological Development (FP7)である。FP7Infrastructure枠において、FP6期間中に推進された欧州分子生物学研究所 (European Molecular Biology Laboratory : EMBL) のバイオインフォマティクス事業など、国・地域縦断的に関連した事業を統合・拡充して、インフォマティクス研究プロジェクトのうち急速に増加、拡大する生物医学情報を活用する次世代のバイオインフォマティクス構築の計画を纏めるために、欧州ライフサイエンス情報基盤計画コンソーシアム (European life-science infrastructure for biological information ELIXIR) が設立されている<sup>6</sup>。ELIXIRは2011年からの本格稼働を目指した基盤整備のために、2007年から2010年の間に€4.5millionが投資されている。2011年以降も総額€470millionの基盤整備費用と、毎年€100millionのコンソーシアム運営費用が投じられる予定であり、欧州のバイオインフォマティクスの中核を担う組織になると予想される。その他、欧州委員会の科学技術政策に関する情報検索サービスCORDIS (Community Research and Development Information Service) によると、医科学研究、健康科学など、本戦略プログラムに関連したFP7Cooperation (EU加盟国以外の国の研究者も参加可能な目標達成型国際共同研究) Health枠

<sup>6</sup> <http://www.elixir-europe.org/page.php?page=home>

において11件採択されている。実施期間は3～5年、配分額は実施期間が3年のものは€2.8 million前後、5年のものは€10 million前後である。研究代表者の所属国を見ると英国が4件と最も多く、次いでイタリア、スペインが2件ずつ、残りは、オランダ、デンマーク、ハンガリーの研究者を代表とするチームに配分されている。

EU各国でも医療情報とライフサイエンス研究の連係を推進している。デンマークでは、病院における患者情報と、分子レベルのオミックスデータを統合する事によって、疾患の病態と遺伝子発現の関係解明を進めている。

特に、欧州での基盤整備プロジェクトは、本プロポーザルのTHIBと関連しており、国際連携を視野に入れたTHIBの構築を目指す必要がある。

## 付録1 参考文献等

- (1) 内閣府 健康研究推進会議 「健康研究推進戦略」 平成21年7月
- (2) ELIXIR <http://www.elixir-europe.org/page.php>
- (3) 統合データベースプロジェクト <http://lifesciencedb.mext.go.jp/>
- (4) 中村裕輔著 「これからのゲノム医療を知る」 羊土社 2009年
- (5) 森下真一、阿久津達也編、「バイオデータベースとソフトウェア最前線」 実験科学増刊 Vol.26-No7 2008 羊土社

## 付録2 専門用語説明

### アクセス制御

情報セキュリティの用語で、主体（Subject）が対象（Object）に対してアクセス（読み／書き／実行）することを許可したり、拒否したりすること

### XML (Extended markup Language)

XMLは文書やデータの意味や構造を記述するためのマークアップ言語の一つ。マークアップ言語とは、「タグ」と呼ばれる特定の文字列で地の文に情報の意味や構造、装飾などを埋め込んでいく言語である。

### オントロジー

文書の内容を説明する意味情報（メタデータ）を各文書に付加し、メタデータを記述する用語を定義する構造を構築する。この構造がオントロジーとなる。オントロジーを導入することにより、検索対象となる文書が単なる単語の集まりとしてではなく、文書全体で大きな意味を持ったデータとして扱われ、各文書について統一的な付加情報をもたせることができる

### コホート研究

コホート研究は、疫学におけるひとつの研究手法である。特定の因子に暴露した集団と暴露していない集団について、研究対象となる疾患への罹患率を調査し比較することで、因子と疾患の関連を検討する研究手法。

### クラウドコンピューティング

クラウドは、ネットワーク（インターネット）をさし、ネットワーク上のサーバで各種の処理が行われる形態をさす。

### ゲノム疫学

ゲノムなどの遺伝子情報をもとに、集団における疾患と特定の集団における健康に関連する状況あるいは事象の、分布あるいは規定因子に関する研究を実施すること。

### ストリーミング処理

並列処理の1つで、入力したデータに対して、処理を実行しながら、出力を行う。

### セマンティックWeb

Webページの意味を扱うことを可能とするために考案され、W3Cにおいて現在標準化が進められている。データの意味を記述したタグが文書の含む意味を形式化し、コンピュータによる自動的な情報の収集や分析へのアプローチが可能となると期待されて

いる。

### パスウェイ

パスウェイとは細胞の機能を実現する分子メカニズムを書き下した知識のことで、ヘテロな構成要素からなり、強く（ネットワーク状に）構造化されている特徴がある。

### マッシュアップ

Web上に提供されている情報やサービスなどを組み合わせて、新しいソフトウェアやサービス、データベースなどを作ること

### メタデータ

データについての情報を記述したデータ。あるデータそのものではなく、そのデータに関する情報のことである。例えば、データの作成日時や、作成者、データ形式などである。