

「ライフサイエンス分野の統合データベース整備事業」 第2回研究運営委員会 議事要旨

【日 時】 平成18年12月25日(月) 15:30～17:30

【場 所】 情報・システム研究機構 事務局会議室(秀和神谷町ビル2階)

【出席者】 堀田機構長(座長)、秋山委員、大久保委員、大倉委員、勝木委員、久原委員、五條堀委員、小原委員、高木委員、辻井委員、中村桂子委員、松原委員(座長以外は50音順)

【陪 席】

文部科学省 : 坂下課長補佐、松永調査員

経済産業省 : 牧係長

内閣府 : 柴田主監補佐

(独)科学技術振興機構 : 酒井主任調査員

【事務局】 高野事務局長、河野総務課長、能住財務課長、笹島総務課課長補佐、加藤財務課課長補佐、

【議 事】

(1) 第1回研究運営委員会・統合DB整備戦略作業部会議事要旨の確認

堀田委員長から、11月8日に開催された合同会議の議事要旨(資料1)に関して、事前に各委員には確認をしているが、意見があれば会議の終了までに事務局まで連絡して欲しい旨の発言があった。特に意見はなく、議事要旨は承認された。

(2) ライフサイエンス分野の統合データベース整備事業の全体構想について

最初に、11月8日に開催された合同部会における委員各位からの意見と12月18日に開催された戦略作業部会での議論に基づき、資料2-1～資料2-4を作成したとの説明が堀田委員長よりあった。

続いて、高木委員から資料2-1を用いて、「事業推進の基本的考え方(案)」について説明があった。さらに、高木委員から資料2-2及び資料2-3に基づいて、「ライフサイエンス委員会報告書(H18.5)指摘課題への対応策(案)」、「事業化計画(案)」について説明があった。次に、大久保委員より、資料2-4を用いて、「5年後の利用イメージ(統合検索結果)」の説明があった。

これらの説明についての質疑応答は、次項の「(3)平成18年度の実施計画、成果(案)及び現在の実施状況について」の質疑応答と合わせて、まとめた形で後述する。

(3) 平成18年度の実施計画、成果(案)及び現在の実施状況について

大久保委員から、資料3-1、資料3-2、資料3-3、及び資料3-4を用いて、「18年度成果の位置づけ(案)」、及び「個別タスクの実施計画及び成果イメージ(案)」について説明があった。

以下、(2)項と(3)項に関する意見及び質疑応答の概要は下記のとおりである。

全体構想に関して

・ 資料2-3は過不足なく書かれている。ヒトとその他の生物のDBは資料2-3では別の項目になっているが、対等の扱いにすべきである。

⇒ 資料2-3の2)事務局、及び連絡調整機能の中に「下記のプロジェクトチームの連絡調整統合機能」という項目があり、これが仕様の共通化を担当する。

・ 目標はユーザーから見たナレッジマネジメントであるというスタンスを明確に出すべきである。

・ 戦略として、最初に優先的に実施することを重点化したらどうか。運営委員全員が議題の全体にわたって知識が充分であるとは限らないので、各委員自身が各々の専門に応じて適切な情報を集めて提供するという機能を委員会に持たせる方が良い。

⇒ 指摘された機能については、検討の上、付け加えていきたい。

・ 実施イメージの見せ方には検討がまだ必要だが、全体的にはよくまとまっている。

- ・ 研究立案は個人でやるものであり、統合 DB による支援までは考えなくても良いのではないかと。DB 自体はブリミティブなほうが良いと考える。個々の DB を全ての人がアクセス可能にすることが第1の作業である。

5年後のユーザーイメージに関して

[日本語文献検索の重点化について]

- ・ PubMed の価値は非常に高い。文献の中に秘められた価値に対してわが国はもっと注目をすべきであった。
- ・ これまでのスコープから若干飛躍しすぎている感がある。日本語は重要であるが、それ以外に膨大なデータがある。
- ⇒ 飛躍はなく、文科省ライフサイエンス委員会報告の課題に対応した実施事項として説明している。分子データと文献部分の連携により、利用者にとってより使いやすい DB を目指している。
- ・ 日本語文献検索の重点化自体については賛成だが、コストと目標の達成度が明確でない。
- ・ JST において、日本語文献の DB 化を実施してきている。
- ⇒ JST の DB は有料であり、他の分子 DB との連携がなく、索引付けも一般的用語を用いて実施しており、全文検索にも対応していない点が、ライフサイエンス統合 DB の観点からは課題であると考えている。ここで目標は、他の分子データ DB との連携構築とライフサイエンス分野に特化した最適な索引付け、及び全文検索の実現により、研究企画の支援が可能なナレッジマネジメントシステムの構築を目指すことである。
- ⇒ 日本語文献の統合 DB への活用に関しては、JST からの方針提案という方向もある。

[ユーザータイプに合わせた仕組みの構築について]

- ・ ユーザーニーズの収集を遅滞なく行うにはどうしたらよいか。
- ⇒ 様々なユーザーに対応するユーザーモデルを構築し、それに合わせたたたき台を作成する。
- ・ 提示された、「5年後の利用イメージ(統合検索結果)」は、ユーザータイプに応じた仕組みの観点では不十分である。異なるユーザータイプが多く、タイプ毎に視点の異なる検索をする可能性があり、それに対してどういう機能を準備するかという議論が必要である。具体的な問題を列挙して解決することをいわないと、google と同じだといわれる。
- ⇒ ユーザーのタイプに応じた検索のイメージを複数パターン作成して提示したいが、十分にはそこまでいたっていない。しかし、18年度の成果の資料が、具体的な問題への最初の対応を表しているので紹介する。
- ・ バイオのユーザーは多様性が大きいので、辞書も含めてツールパッケージを用意し、ユーザーごとに組み合わせを変えて使うツールキット的なものを構築し、個別のユーザー向けのものが簡単に作れるようにすることが重要と考える。
- ⇒ これについては本事業においても実施を計画しており、資料2-1に記載している。
- ・ ツールキットありきでは失敗する(例:NCBI のツールキット)ので、具体的なニーズを良く踏まえてツールキットを開発することが重要である。
- ・ 大部分のユーザーの情報処理のレベルは高くないので、使いやすいツールキットが必要である。またツールを使いやすいデータが準備されている状況も必要である。
- ・ どう仕様がかわっても対応可能なツールの研究開発者を育成することも重要だと考える。

ソフト等の知財の公開について

- ・ 有用でアクセスの多いツールの作成にあたっては、仕様を公開することが重要である。
- ・ ソフトの知財に関してはイギリスでも検討中であり、まだ方針が定まっていない。
- ・ 重要なことは公開可否の判断をメリット/デメリットに基づいて行うことであり、判断は一律にはできない。
- ・ 知財に関しては、文科省の委託事業には一定のルールがある
- ・ ソフトの知財ははっきりしているが、データベースの著作権やテキストマイニングされたものの知財等に関してははっきりしない面がある
- ・ 良く分からないから非公開にするということが生じないようにしたい。
- ・ 日本では、知財の扱いについて制度的に十分整理されていない。
- ・ 日本語文献のオープン化を促進するために、研究費委託の条件として「日本語の報告提出」を義務付ける

ことも考えられる。

平成 18 年度成果に関する質疑応答

【教育に関して】

- ・ 現在アノテーションを実施中のグループと連携してそのノウハウを生かす形でやっているのか？
- ⇒ アノテーションに関しては、実際にやってきた人たちと連携して、アノテーションフローの共通部分の抽出とマニュアル化、及び教材作成をお願いしている。
- ・ アノテーションを実施中のグループに対するメリットも必要ではないか？
- ⇒ プロジェクトと大学と連携して教育を行い、各種プロジェクトで雇用できるような人を養成する予定である。
- ・ キュレーターによって読み取る箇所が異なるので、マニュアル化には文献から実際に読み取っている箇所のデータを蓄積することが必須である。そのためのツールを各キュレーショングループに与えてデータを集める必要がある。
- ⇒ キュレーションに関しては、ターゲットを絞って調査を実施している。
- ・ 本年度は第一バージョンができるのか？
- ⇒ そうである。

【リソースポータル構築に関して】

- ・ リソースポータル構築は、日進月歩の分野である。ここでは、先進的なものか、あるいは実用的なものをやるのか？またどんなグループでやるのか？
- ⇒ 中心は東大医科研の片山氏が実施している。先進的なところは、片山氏が中心に実施しており、それとともに、実用的なものの開発も平行してやっている。
- ・ NLP に関しては、主要なグループが作成したツール群を共通プラットフォーム上で動作できるような環境を構築する動きがあり、各サイトのツールをネットで連携して使用可能にすることも 1~2 年で実現の可能性はある。
- ⇒ 現在調査中で、どこまでデファクト化するのかを見極める予定である。

【調査に関して】

- ・ ユーザーからの視点の調査は、本年度の調査項目に含まれているのか？
- ⇒ 本年度は、よりマクロな調査を実施中である。方向が見えてきた段階で、個々のユーザーのニーズ調査へ移行の予定である。本年度はそこまでは実施しない予定である。

以上の議論を踏まえた上で、ライフサイエンス分野の統合データベース整備事業の全体構想、及び平成 18 年度の実施計画、成果(案)及び現在の実施状況について了承された。

(4) その他

坂下文部科学省課長補佐から、以下の 3 点に関して報告があった。

- 1) 統合データベース整備事業の予算に関して閣議決定があり、平成 19 年度は 16 億円の予算案が認められた。平成 19 年度以降、フィージビリティスタディから本格的な実施へと移行することになる。来月早々からの公募開始に向けて準備中であり、4 月当初からの事業開始を想定している。
- 2) 公募においては、中核機関のみではなく、分担機関も合わせて公募する考えである。分担機関は、中核機関の戦略に沿って、統合化技術開発を行うだけでなく、リソースとなる DB (コンテンツ) を持つことを条件とする。また、中核機関は、プロジェクト費用でまかなうコア部分以外も含めたわが国の統合 DB に関する戦略立案を行い、JST や大型 PJ 等への提言も行う。
- 3) 内閣府連携施策群の省庁横断的な DB の検討に関して、来年度はより緊密に連携し、統合 DB 運営委員会と内閣府連携施策群運営委員会を合同で開催することも考えている。

堀田委員長から、次回、今年度最後の運営委員会を 3 月の中旬に開催することを予定しており、日程の調整を後日事務局から行いたい旨の発言があり、了承された。

以上