

データベース受入に関するアンケート調査の実施について(プレリナリーな解析結果の紹介)

1. 目的

本調査は、分担機関や補完課題以外で整備・構築された国内データベースを対象として、データベースの内容、維持・更新状況、提供可能性及び提供時の問題点などをアンケートにより調査し、受入による統合化の方針策定に資する基礎情報を収集するとともに、受入対象の候補を抽出することを目的として実施した。

2. アンケート対象

アンケート送付先は、ライフサイエンス統合データベースセンターにて抽出しているデータベースリストに基づいて選定した(104件)。原則として以下に該当するデータベースは調査対象外とした。

ライフサイエンス統合データベースプロジェクトの分担機関及び補完課題機関 理化学研究所が運営する DB(別途統合の活動が進められているため) 産業技術総合研究所の糖鎖 DB 農林水産省傘下の研究機関が運営する DB 厚生労働省傘下の研究機関が運営する DB

3. アンケート内容

- 公開されているデータベースコンテンツと検索方法に関する質問
- データベースの経歴・管理状況
- データベース維持の将来計画
- 提供可能性及び条件
- 提供方法
- 受入時の連絡先

4. 集計結果の概要

1) 回収数(率) 60/104 (57.7%)

2) アンケート結果

・データベースの内容

ヒト・マウス・微生物などが多いが、生物種を限定せず生物種網羅的なデータベースも多い。主なデータソースの種類は、実験データ(約半分)、公共DB、文献データなど。

・データベースを提供する環境

データベースを提供するサーバの多くは1台のみから構成。サーバOSはLinux、WebサーバはApache、データベースシステムはMySQL、PostgreSQLがよく使われるが、ファイルのみの管理も多い。

・データベースの提供方法

大半はWeb上でのキーワード検索・固定条件による絞込み検索・分類別のハイパーリンクを提供。

・データベースの利用条件

大半は研究目的・商用目的に関わりなくデータベースの1次利用に対して制限を設けていない。データベースの2次利用や再配布に対しては、目的に関わりなく半数近くが制限を設けている。

・現在のデータベースの管理状況

データベース更新頻度は、年1回または2回、データ取得時に更新で半数を占める。維持・管理専任者は大半で1名以上。管理体制としては、研究責任者が研究メンバー・SE・学生等にデータベースの維持・管理に当たらせることが多い。追加データ取得費用、維持・管理費用は、大半の場合500万円/年未満。

・今後のデータベースの維持・管理計画

ほとんどの場合、現状維持・データの増強を検討。しかし、2008年度以降もプロジェクトが継続しているのは全体の30%程度にすぎない。残りの70%のデータベースでは、プロジェクト以外から維持・増強のための資金を賄う必要があると考えられる。

・統合DB事業との提携

半数以上が、統合DB事業に対して全データもしくは一部データを開示できると回答。提携の条件としては、現在のDB作成者が今後もデータを更新する権限や著作権などの留保、現在の利用条件の継承、統合DB事業による維持・管理・提携費用の負担を挙げる回答が多い。

3) アンケート結果をふまえた今後の検討課題

・本アンケートによって、個々のデータベースと統合DB事業との提携の際に懸案となり得ることとして、データベースの更新・維持・管理の問題とデータベースのライセンスの問題があることが判明した。統合データベースセンターに受け入れた後も、データベース作成者がデータ更新可能な仕組み作りの検討が必要。データベースの著作権については、作成者の権利を留保しつつ、一定条件下での2次利用・再配布を可能にするライセンス体系の策定の検討が必要。

・統合DB事業との連携が可能なデータベースの作成機関との間で、具体的な連携の方法について検討を開始する。