基調講演 ④

海洋生物情報のデータベース化と統合：JAMSTECの取り組み

園田 朗
海洋研究開発機構地球情報研究センター

海洋研究開発機構(JAMSTEC)における観測データ管理・公開業務は、1998年に共同利用を開始した「みらい」の運航に伴って始まった[1]。その後、有人潜水調査船「しんかい6500」などを用いた深海調査研究についても共同利用となり、2007年にJAMSTEC全体の調査研究を対象としたデータポリシー等が整備されて本格的なデータ・サンプルの公開・提供が行われるようになった[2]。この間、データ管理・公開体制の整備を進めるとともに、1998年の「みらいデータWeb」による海洋地球研究船「みらい」の観測データ公開をはじめとして、2001年の「深海映像データベース」や、「文書情報データベース」など、各種データ公開システムを構築し、データ提供してきている[2][3]。その後、乗船研究者以外の二次利用者の増加とともに、各種データ公開システムを横断的に検索できる「JAMSTECデータ検索ポータル」や、他部署で個別プロジェクトに特化して運用・管理されているデータベースやデータセット等についてカテゴリーの整理をしてカタログ化した「JAMSTECデータカタログ」などのサービスも提供している[2]。また、観測データだけでなく、採取された各種サンプルの所在情報を管理・公開するシステムも「岩石サンプルデータベース」を始めとして順次構築し、前述の各種データ公開システムと併せて、一元的に国際海洋環境情報センター(GODAC)で運営するWebサイトの「海と地球のデータベース」から公開している[4]。

そして海洋生物情報については、「生物サンプルデータベース」[5]の整備と併せて、その生物サンプル取得のメタ情報、並びに「深海映像・画像アーカイブス(J-EDI)」[6]に格納されている深海記録映像に映っている生物とその映像メタ情報を対象とした海洋生物情報統合システム「BISMaL」を構築[7]し、海洋生物生態研究や海洋生物多様性研究ユーザへの情報サービスを2009年より運用を開始した[8]。

これらの各種データ公開システムのいくつかは、国際的な枠組みで運用されているデータベースや、データポータルサービスと連携を進めており、ユネスコ政府間海洋学委員会の国際海洋データ・情報交換システム(IODE)や、その配下にある海洋生物地理情報システム「OBIS」、コロンビア大学ラモント・ドハティ地球観測所（LDEO）の「EarthChem」、NASAの地球変動マスターディレクトリ(GCMD)等へデータ提供を行っている。特に「OBIS」との連携においては、その地域ノードとしてOBIS日本ノード（J-RON）としての活動も開始[9]しており、J-RON運営委員会を組織してJAMSTECのデータだけでなく国内の他機関から海洋生物データセットの提供を受け、「BISMaL」経由でデータ提供を開始し、日本周辺の海洋生物情報の提供に貢献している[7]。

海洋生物情報のデータベース構築においては、外洋の生物を扱う場合、船舶調査が主体となるため、船舶運航情報が基本情報としてメタ情報に付帯するので採取や観察時の位置情報や深度情報も同時に付与されることになる。これと生物分類情報とを併せて検索基盤を構築していくことになるが、特に中深層では生物多様性が高いため未記載の新種が多々出てくるので生物種を同定できない場合がある。また生物分類系統そのものについても、形態分類と遺伝子分類とで相違がある場合があり、これらの課題へのシステム的な対応が必要となってくる。後者については以前に比べると、両者の連携が取られるようになってきているが、完全に融合することは難しく、生物分類が細分化していく傾向もある。狭いコミュニティではあまり問題にはならないが、広くデータ統合を進めようとすると、生物分類系統情報の同義語・類義語の整理と管理等も課題となる。また、中層の生物の場合、クラゲのような船上に採取してしまうと形態判別不明（単なるゼラチン質の塊？）になるため、生息現場での映像や画像記録が必要

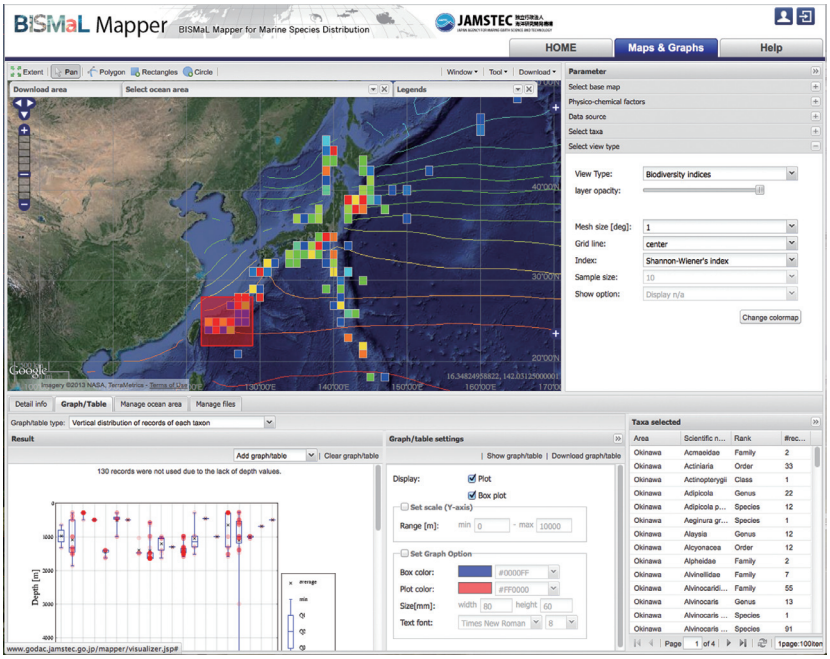

図　BISMaL 可視化・解析サブシステム「BISMaL－Mapper」の画面例

となるが、最近では映像・画像の高精細化が進んだことである程度分類できるレベルになってきており、それら映像データを統合したデータベース化も重要と考えている。

海洋生態研究にとっても、生息現場での映像や環境情報と統合したデータベースは有効で、最近では遺伝子解析が容易になってきたことから、サンプルの遺伝子情報データベースとの連携も重要な要素となっている。加えて、生物出現データの母数を増やすことが重要であるが、観測データ提供・公開やデータセット作成の文化が十分に醸成されていないため、遺伝子情報等を除けば、海洋生物情報の集約化は効率的に行われているとは言い難い状況にある。

「BISMaL」では、データ取得者の提供したデータセット単位での管理機能や、データ可視化・解析支援機能[10]（図）などを提供することで、データ提供のインセンティブを少しでも出すように努めている。

2012年1月より開始した「東北マリンサイエンス拠点形成事業（海洋生態系の調査研究）」（文部科学省補助事業）においては、複数機関が実施する調査の結果を地元自治体等へも提供していくことを目的としており、その実現のため同事業の観測データ等を共有・公開する機能の整備も併せて計画に盛り込まれており、現在システムを構築中である[11]。

これらのデータ公開に際しては、漁業復興との関係を強く意識しながら実施する必要があり、風評被害などへの影響を配慮しつつ、適切かつ複雑な公開管理が必要とされている。

本発表では、これらに加えてユーザアクセス分析や将来的な展開等についても紹介したい。

関連論文等

[1] 土屋、他(2005)独立行政法人海洋研究開発機構(JAMSTEC)における観測データの管理と公開,情報管理,48,163-170.
[2] 園田朗(2010)JAMSTECにおけるデータ公開,月刊 海洋/号外,53,123-132.
[3] 園田朗、ほか(2005) 国際海洋環境情報センター(GODAC)における情報管理について,情報管理,48(4),214-219.
[4] http://www.godac.jp/top/
[5] http://www.godac.jamstec.go.jp/bio-sample/
[6] http://www.godac.jamstec.go.jp/jedi/j/index.html
[7] 田中克彦、ほか(2009)JAMSTECにおける海洋生物の統合データシステム構築,日本プランクトン学会報,56(2),178-183.
[8] http://www.godac.jamstec.go.jp/bismal/j/
[9] http://www.jamstec.go.jp/obis_j/
[10] http://www.godac.jamstec.go.jp/mapper/index.jsp
[11] http://www.i-teams.jp/