

— 報告 —

海洋観測データ公開システムの持続的发展を目指して

佐藤 孝子^{1*}, 角田 晋也¹, 富山 隆将¹, 小熊 幸子¹, 市山 祐司¹, 長谷 英昭¹, 福田 和代¹, 齋藤 秀亮¹,
華房 康憲¹, 園田 朗¹, 坪井 誠司¹

独立行政法人海洋研究開発機構（JAMSTEC）は海洋観測のため毎年約百航海を実施している。地球情報研究センターではこれらの航海により得られた海洋観測を始めとし、陸域観測も含めた様々な地球環境データや、関連航海の多種サンプルを公開しており、サンプルのメタデータも含むデータ公開システムの使い易さ、すなわちユーザビリティの向上を目指している。ここでは主にJAMSTEC航海・潜航データ探索システム（Data Research System for Whole Cruise Information in JAMSTEC: DARWIN）および深海底岩石サンプルデータベース（Geochemistry and Archives of ocean floor rocks on Networks for Solid Earth Knowledge Integration: GANSEKI）で改善された直近の事項について解説する。

キーワード: 海洋観測, 地球環境, メタデータ, データ公開, ユーザビリティ

2013年10月4日受領；2013年12月10日改訂稿受理；2013年12月11日受理

1 独立行政法人海洋研究開発機構 地球情報研究センター

*代表執筆者：

佐藤 孝子

独立行政法人海洋研究開発機構 地球情報研究センター

〒236-0001 神奈川県横浜市金沢区昭和町3173-25

045-778-5405

satot@jamstec.go.jp

著作権：独立行政法人海洋研究開発機構

— Report —

The trial for the sustainable development of the system for publicizing of the ocean-observation data of JAMSTEC

Takako Sato^{1*}, Shinya Kakuta¹, Takayuki Tomiyama¹, Sachiko Oguma¹, Yuji Ichiyama¹, Hideaki Hase¹, Kazuyo Fukuda¹,
Hideaki Saito¹, Yasunori Hanafusa¹, Akira Sonoda¹, and Seiji Tsuboi¹

The Japan Agency for Marine-Earth Science and Technology (JAMSTEC) are operating about one hundred of research cruises every year. The collected data are published at multiple JAMSTEC data sites as part of the metadata-tagged geo-environmental data maintained by Data Research Center for Marine-Earth Sciences (DrC). At DrC, continuous efforts are being made to improve usability (user-friendliness) of the data sites and the related databases. In this article, we report some progresses made recently on the ocean-observation database “Data Research System for Whole Cruise Information in JAMSTEC (DARWIN)”, and the deep seafloor rock sample database” Geochemistry and Archives of ocean floor rocks on Networks for Solid Earth Knowledge Integration (GANSEKI)”.

Keywords : Ocean observation, Global environment, Metadata, Open data, Usability

Received 4 October 2013 ; Revised 10 December 2013 ; Accepted 11 December 2013

1 Data Research Center for Marine-Earth Sciences (DrC), Japan Agency for Marine-Earth Science and Technology (JAMSTEC)

*Corresponding author:

Takako Sato

Data Research Center for Marine-Earth Sciences (DrC), Japan Agency for Marine-Earth Science and Technology (JAMSTEC)

3173-25 Showa-machi, Kanazawa-ku, Yokohama 236-0001, JAPAN

Tel.+81-45-778-5405

satot@jamstec.go.jp

Copyright by Japan Agency for Marine-Earth Science and Technology

1. 背景

JAMSTEC（海洋研究開発機構）では、1980年代より研究体制を整え、1990年代には深海微生物プロジェクトを立ち上げた。その後、生物学だけでなく海洋学やシミュレーションを含む環境学、地震学などを含む地球科学など、多分野の、大きいプロジェクトが次々に立ち上がり、船舶や潜水船などの施設・設備を利用した研究活動を行ってきた。船舶「みらい」は1998年度から、船舶「なつしま」、「よこすか」、「かいかい」及び各種の潜水調査機器は2001年度から、これら研究船利用研究に関して公募されるようになった。課題が採択されれば、JAMSTECに所属しない研究者でも深海底を含む海洋地球観測の調査航海に参加できる仕組みである。現在では、毎年約百航海および二百潜航での調査が実施されている。これらの研究活動からは、幅広い分野の観測、分析、画像等のデータや岩石、コア、生物等のサンプルが取得され、これらを記述するクルーズレポートなどの報告書が作成されるが、JAMSTEC地球情報研究センター（以下、DrC）では、これら貴重な情報やサンプルを人類の財産と考え、情報公開サイトで公開することによって、他の研究、教育やアウトリーチなどを目的とする活動へのデータ、サンプルの広範囲な2次利用を促進し、有効活用されることを目指してきた（佐藤・他、2011）。また、国際的な関係機関による海洋観測データ公開に対しても、JODC（日本海洋データセンター；Japan Oceanographic Data Center）を通じてデータ提供を実施し、世界中の海洋研究に寄与してきた（Seta, et. al., 2013; Levitus, et. al., 2013）。

2008年度からは、JAMSTECの基本方針に基づき、データ・サンプル取り扱いに関する規程類を整備し、航海で取得されたデータ・サンプルを一元的に管理して、「観測航海データサイト」等による地球観測データの公開を進めてきた。その後、2012年10月より「観測航海データサイト」内のデータをデータベース化し、「JAMSTEC航海・潜航データ探索システム（Data Research System for Whole Cruise Information in JAMSTEC: DARWIN）」として新たに公開を始めた（佐藤・他、2013）（図1、2）。

この図2はまた、観測データの取得から公開までの流れも示している。すなわち、JAMSTECの船舶、潜水調査機器で取得された観測データは、輸送専用のルートによりDrCに届けられる。その後、DrCにおいて、受領、検品、データ処理も伴う品質管理などの過程を経て公開準備がなされ、オンラインやオフラインで観測データは公開される。岩石、堆積物コア、生物の各種サンプルには専門知識をもってこれらの適切な管理と提供を行うキュレータがあり、メタデータ（採取日時や場所などのサン

プルに関する情報のこと）や岩石、堆積物コアサンプルそのものの管理、公開を担当するというシステムになっている。

DrCが担う公開業務の対象は、5調査船「なつしま」「かいかい」「よこすか」「みらい」「かいかい」で実施された全航海、そして「ちきゅう」船舶航海のうち、試験や訓練などのJAMSTECの独自航海である。2008年度からの規程類の運用、データ・サンプル管理、各種データ公開サイトの運用に伴い、データ量が飛躍的に増大した。具体的には、公開対象となる年間航海数は、2007年度までは60以下であったのが2008年度以降は100近くに倍増した。これに伴い、データ管理の自動化や、公開サイトの検索機能の向上が求められてきた。そこで、「船舶観測メタデータ（観測項目や取得地点などのデータに関する情報のこと）管理システム（CMO）」によるメタデータ

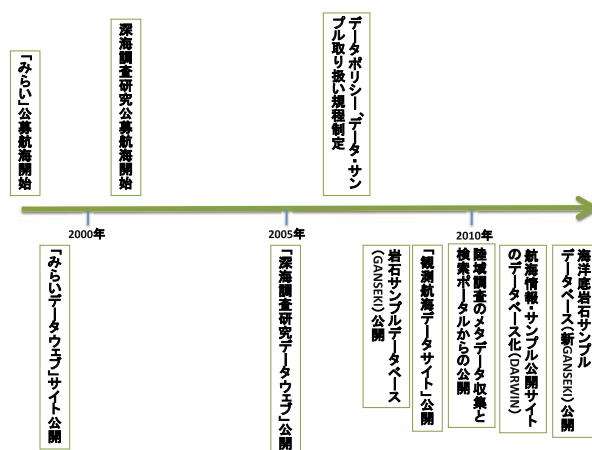


Fig. 1. The development history of uniform management and opening “data and sample” by DrC (mainly with DARWIN and GANSEKI database).

図1. DrCによるデータ・サンプルの一元的管理と公開システムの開発史（DARWINと新GANSEKIシステムを中心に）。

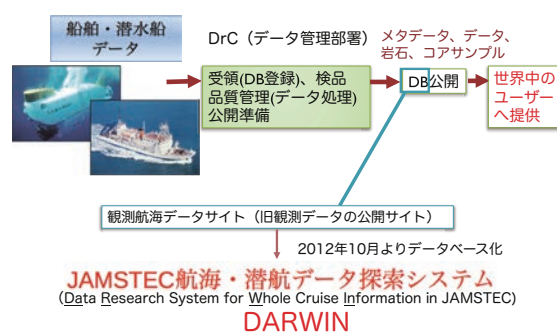


Fig. 2. Data flow from cruises and dives to DARWIN

図2. 航海・潜航からDARWINまでのデータ公開の流れ

一括管理システムを構築し、種々のデータベースとの情報の連携を図り、データ管理の自動化、効率化に寄与している。

これら各航海・潜航にて取得される主なデータは水温、塩分、流向、流速、栄養塩、海上気象、重力、磁力、地形等に関するものであり、主なサンプルは岩石、堆積物コア、生物、海水等である。2013年10月1日現在公開中の航海数は925（1987～2013年）および潜航数は2333（2001～2013年）となっている。DARWIN内に登録されている船舶で取得された主なデータ種（メタデータのみの提供も含む。また、特に探査機取得データはまとめて1項目とした）を表1に、データ種別ファイル数を表2に纏めた（DARWIN URL: <http://www.godac.jamstec.go.jp/darwin/j/>）。

これらのデータやサンプルをJAMSTECに提供する者は、調査航海を計画、実施する研究者である。一方、データの有効活用という概念から、視点を利用者（ユーザー）に移すと、それらは1次から3次までのユーザーに分類することができよう。1次ユーザーはこれらデータやサンプルを利用した研究を企画、実施する直接調査航海に関わる研究者が多く、航海メタデータなども含めた多項目のデータを有効に利用する傾向がある。また、2次ユーザーとはデータやサンプルを研究以外の用途に活用して行く利用者と定義する。教育や啓蒙活動の利用などもここに含まれる。3次ユーザーはさらにニーズを創成し、潜在的なユーザーから新規のユーザーを指す。これら多種ユーザーの発掘が公開システムの持続的発展に必要とさ

Table 1. Various kinds of data provided by DARWIN (as on 1st October 2013)

表1. DARWINで提供されている主なデータ種（2013年10月1日現在）

地球物理探査	航跡、海底地形（MBES）、重力、三成分磁力計、全磁力、サブボトムプロファイラ
海上気象	海上気象、雲底高度、ラジオゾンデ、気象ドップラーレーダー、全天画像、放射照度・輝度
海水連続観測	流向流速（ADCP, LADCP）、水温・塩分、全炭酸、溶存酸素、一次生産、クロロフィル、光合成色素、植物プランクトン量
サンプル	岩石、生物、堆積物コア、堆積物、海水
探査機取得	サブボトムプロファイラ、音響測位、サイドスキャンソナー、連続水温・塩分・深度計、静止画像、映像、航跡図
研究者提出 (PI)	補正等の処理済、サンプル分析

Table 2. Numbers of files of open data (as on 1st October 2013)

表2. 公開中の主なデータ種別ファイル数（2013年10月1日現在）

CTD	潜水船 CTD	XBT	XCTD	ボトル採水	連続 pCO ₂
7966	544	2799	3521	4369	80

れる。それには、これまでの公開機能に対してさらなる使い易さ、すなわちユーザビリティの向上が求められる。

2. 公開データのユーザビリティの向上

2008年から運用してきた「観測航海データサイト」での課題は次の通りであった。

- ・データベース化していないため、検索機能がない。キーワードや地図で、データやメタデータを検索できない。したがって、航海のID番号を知らないと航海、潜航、観測データページにたどりつけない。
- ・データは個別にダウンロードする仕組みなのでデータ種の全体像もわからない。
- ・HTMLページが膨大になり、管理が煩雑である。

そこで、2011年度よりDARWIN開発の準備を開始し、データ公開の仕組みの大幅な改善に取り組んだ。

その結果、上記の課題は、DARWINでは次の通り解決された。

- ・（用語を含め）統合データベース化を行ったため、地図やキーワード検索が可能となった（図3）。したがって、多様なそれぞれのキーワードでの絞り込みが可能となった。
- ・全体を俯瞰することも可能となった（船舶や潜水船、観測データ種等）。したがって、どのようなデータを持つデータベースであるかという全体像の把握が容易となった。
- ・DrCでの実現目標であった、「One stop data shop（1つの公開システム内で、関連データを一度に検索、入手）」の構築により、バスケット機能（お買い物カート）による一括ダウンロードが可能となった（Hanafusa, et. al., 2013）。



Fig. 3. Japanese top page of DARWIN. The page can be switched to English; Overlooking search or search by keywords or on maps was realized.

図3. DARWINの日本語トップページ。英語ページも併せ持つ；俯瞰しながらの検索またはキーワードや地図による検索を実現した。

2.1.3 JAMSTEC内の他のデータサイトとの連携

DARWINを中心軸として、JAMSTEC内の各サンプル・映像画像サイトとの相互リンクを張り、用語を含め統合データベース化を行った。これにより、例えばDARWINとサンプル、画像データベースとの領域横断検索が可能となった（図8）。

2.2 GANSEKI

2006年より本格運用されていた旧GANSEKI（市山・他，2011）は、サンプルの取得情報や保管状況、化学分析データや画像データを表示する機能を備えていたが、検索機能としては簡単な地図検索のほか、メタデータによる絞り込み検索があるのみで、閲覧性や外部データベースとの連携機能についても改善の余地があった。2013年度に公開された新GANSEKI（富山・他，2013a; 2013b）では、システムとインターフェースの一新により、自由度の高い検索と快適な閲覧環境を実現した（図9）。検索機能としては、従来機能に加え全文キーワード検索、分析データの数値検索などの機能が追加され、ユーザーの目的により、簡便な検索にも詳細な検索にも対応できるよ



Fig. 8. Integration of data sites in JAMSTEC

図8. JAMSTEC内のデータサイト間の連携



Fig. 9. Search and reading functions of GANSEKI

図9. GANSEKIの検索・閲覧機能

うになった。閲覧機能としては、薄片写真やサンプル写真を一覧できるサムネイル表示機能や、検索したデータを整理して表示するマイリスト機能の実装により、ユーザーは快適にサンプルやデータを比較検討することが可能になった。また、新GANSEKIではJAMSTEC内外で進む各種データベース構築の取り組みに対応し、各サンプルに複数のURLを登録する領域を設け、容易に関連データを参照できるようになった。DARWINと同様、アクセス解析にてサイトのJAMSTEC外からの訪問数を調査すると、新GANSEKI公開以前は、毎月200～300であったが、2013年5月には学会発表等の周知活動とともに更新された新GANSEKIシステムの公開により、これまでの訪問数で最も多い400を記録した。今後の定着度合は、本サイトの適切なメンテナンスと更新によると思われる。

3. 今後の方向性および展望

海洋観測データを公開している仕組みは数多くあるが、DrCが提供するDARWINシステムとそのリンクで繋がれたデータサイト群は国際的な視点からも大変ユニークである。航海を一つの単位とし、そこで取得された全ての観測データと、これに付属するタグのように、データを記述する「メタデータ」を一つのデータベースに纏める。これでユーザーは自分の求める条件のデータを検索、抽出することが可能となる。このシステム自体は汎用的であるが、各航海に関連した種々の情報を纏めたデータサイト群の種類（特に生物サンプル、画像情報など）とその連携は、世界的にも大変希有な例である。すなわち、ユーザーが関心を持つ航海を軸に、DARWINからは主に観測データを、そしてリンクを介して、岩石・コア・生物各サンプルのカテゴリ毎のデータやメタデータを集めたそれぞれの公開サイトへ、そして同時に膨大な「潜水船による映像画像情報」をまとめたデータベースに繋がるので、その調査航海について科学的データ・サンプルも含めて再現、解析することが可能である。この総合システムの質の高さは世界に類をみない。さらに、航海関係者ではない研究者でも、2次利用を通して、追って解析することができるので、貴重なデータ・サンプルの有効利用に効果が高いと期待される。実際、これらのサイトのユーザーからは、アンケート等を通して「観測データは高品質であり、信頼性が高い」、「公開サイトが良くまとまっているので、さらに使用可能なデータ・サンプルを収集し、充実させて欲しい」等の反応が得られている。

データ管理に関して、常に技術向上を目指すためには、それぞれのシステムの効率化を尊重すると同時に、お互

いのシステムの大枠を把握する必要がある。さらにそのためには、時間とともに変化する共有すべき情報の明文化と効率的な更新が必要となる。なお、観測データ取得、輸送、データ管理、品質管理など、各現場では予想外の事象も発生するので、データ管理を円滑に行うには、常に関係者間の速やかな連絡が重要と考えられる。

データ・サンプルの有効活用には、さらにユーザーを増やす、ニーズを創設することが必須であろう。これに対しては、既に実施しているユーザー登録の管理およびアンケート、アクセスの解析等やヒアリング等によるユーザーニーズの解析が重要である。一方で、ユーザビリティの高いシステムによるデータ提供を行う、ユーザーへのサービスも大変重要なポイントである。我々はその考えに基づき、「One stop data shop」を構想し、実現してきた。更なるサービスとしては、2つの方向性が考えられるだろう。1つは可視化である。グラフによる海洋観測データの可視化は、DARWINに実装すべきサービスの課題として、準備が進んでいる。これにより、ユーザーは最も必要なデータを選択し、入手することが容易になると考えられる。このインタラクティブな可視化の方向性をさらに進めると、2次元の地図選択から、3次元の地球空間によるデータ選択への遷移が考えられる。さらに進めれば、時間軸を導入したインタラクティブな全地球4次元データ選択の仕組みが必要であろう。もう1つは時間軸の未来方向である。すなわち、過去の地球環境観測データ項目に加えて、同じ項目の未来への予測とリンクしていくならば、高いユーザビリティを維持しながら、より多くのユーザーの多様なニーズに 대응していくことができると考えられる。4次元データ選択のできる全球データベースに、観測データやサンプルのメタデータ、潜水船映像画像や生物出現予測データなどが連続的に可視化できるシステムが実現すれば、その魅力は科学的価値のみならず、社会に対して多様なニーズを生むと考えられる。さらにこの情報プラットフォームにJAMSTEC内外のデータを混在させていく方向性は議論の余地もあろうが、ユーザーの立場で考えれば、ユーザビリティの高いシステムのデータ種・数は多い方が良く考えられる。

このように、データ公開システムの持続的発展を目指すことは、科学研究推進のみならず、重要な社会貢献にも繋がると考えられる。今後、さらに大容量のデータをリアルタイムに処理して可視化、ダウンロード、保存、活用など、それぞれの技術革新、機能向上と共に多様な発展が期待される。我々はこれに向けて、刻々と変化する技術などを判断しながら、その時々で最適なシステムを構築していく必要があろう。

謝 辞

本稿は平成25年6月25日にJAMSTEC横浜研究所にて開催されたDrC海洋地球観測データワークショップでの口頭報告を基に作成したものである。また英語要旨作成においては、DrCの若松 剛博士にご協力頂いた。ここに感謝を捧げる。

データ・サンプルの収集公開は、各船の乗組員、観測技術員、研究者、機関関連部署など、多数の方の協力によって支えられている。これら多大な支援へも感謝を捧げる。

参考文献

- Hanafusa, Y., H. Saito, and Y. Abe (2013), THE JAMSTEC METADATA PUBLICATION AND SEARCH SYSTEM, *Advance Publication, Data Science Journal*, 10 April, 1–4.
- 市山祐司, 華房康憲, 相馬伸介 (2011), 深海底岩石試料データベース「GANSEKI」の紹介, *地質学雑誌*, 117, 579–584, doi:10.5575/geosoc.117.579.
- Levitus, S., J. I. Antonov, O. K. Baranova, T. P. Boyer, C. L. Coleman, H. E. Garcia, A. I. Grodsky, D. R. Johnson, R. A. Locarnini, A. V. Mishonov, J. R. Reagan, C. L. Sazama, D. Seidov, I. Smolyar, E. S. Yarosh, and M. M. Zweng (2013), THE WORLD OCEAN DATABASE, *Advance Publication, Data Science Journal*, 2 May, 1–6.
- 佐藤孝子, 長谷英昭, 畑山隆紀, 齋藤秀亮, 華房康憲, 園田朗, 坪井誠司 (2011), JAMSTEC船舶により取得された、データ・サンプルの有効活用を目指して, *BlueEarth'11 要旨集*, 88pp.
- 佐藤孝子, 市山祐司, 長谷英昭, 福田和代, 齋藤秀亮, 華房康憲, 園田朗, 坪井誠司 (2013), JAMSTEC航海・潜航データ探索システム(略称DARWIN)構築による、利便性向上への試み, *BlueEarth'13 要旨集*, 78–79pp.
- Sellars, S., P. Nguyen, C. Wei, G. Xiaogang, H. Kuo-lin, and S. Soroosh (2006), Computational Earth Science: Big Data Transformed Into Insight, *Eos*, 94, 277–288.
- Seta, A., S. Wakamatsu, T. Miyake, and Y. Iwabuchi (2013), JAPANESE CONTRIBUTION TO THE WORLD DATA CENTER FOR OCEANOGRAPHY, *Data Science Journal*, 12, 89–91.
- 富山隆將, 市山祐司, 堀川博紀, 佐藤悠介 (2013a), 新しいGANSEKI: システム更新と機能強化, *日本地球惑星科学連合2013年大会, SCG67-P03*.
- 富山隆將, 市山祐司, 堀川博紀, 相馬伸介, 佐藤悠介 (2013b), GANSEKIが便利-JAMSTEC深海底岩石データベースの改修と多元的データベース網の整備, *2013年度日本地球化学会年会*, 1D09.