

事例

## DIASにおけるデータ公開と課題

### Data Release and Issues in the DIAS

絹谷弘子<sup>1\*</sup>, 清水敏之<sup>2</sup>, 吉川正俊<sup>2</sup>, 喜連川優<sup>3</sup>, 小池俊雄<sup>4</sup>  
Hiroko KINUTANI<sup>1\*</sup>, Toshiyuki SHIMIZU<sup>2</sup>, Masatoshi YOSHIKAWA<sup>2</sup>,  
Masaru KITSUREGAWA<sup>3</sup>, Toshio KOIKE<sup>4</sup>

1 東京大学地球観測データ統融合連携研究機構

Earth Observation Data Integration & Fusion Research Initiative, The University of Tokyo

〒153-8505 東京都目黒区駒場4-6-1

E-mail: kinutani@tkl.iis.u-tokyo.ac.jp

2 京都大学情報学研究科社会情報学専攻

Department of Social Informatics, Graduate School of Informatics, Kyoto University

〒606-8501 京都市左京区吉田本町

E-mail: {tshimizu,yoshikawa}@i.kyoto-u.ac.jp

3 国立情報学研究所 & 東京大学生産技術研究所

National Institute of Informatics & Institute of Industry Science, The University of Tokyo

〒101-8430 東京都千代田区一ツ橋2-1-2 & 〒153-8505 東京都目黒区駒場4-6-1

E-mail: kitsure@tkl.iis.u-tokyo.ac.jp

4 東京大学大学院工学系研究科社会基盤学専攻

Department of Civil Engineering, School of Engineering, The University of Tokyo

〒113-8656 東京都文京区本郷7-3-1

E-mail: tkoike@hydra.t.u-tokyo.ac.jp

\*連絡先著者 Corresponding Author

研究者が収集したデータ, 作成したプロダクトデータ分野を超えた利活用を目的としてデータ統合・解析システム(DIAS: Data Integration and Analysis System)の開発は2006年から開始された。2010年10月からDIASが保有する研究データと研究成果データの公開を始めた。4年間の公開データセット数の増加により現在ではプロジェクト以外の研究者に研究データを提供する環境が整ってきた。有期プロジェクトが持続的に研究データを公開し続けることは困難であり, 今後DIASが長期運用体制に移行することを見据えて研究データの長期的な管理体制を整備する必要がある。本稿ではDIASデータ公開の現状と持続的なデータ, メタデータの維持管理に関するDIASが抱える問題点を

述べる.

Our project, the DIAS (Data Integration and Analysis System) started in 2006, has a purpose of constructing data infrastructure that can integrate earth observation data, numerical model outputs, and socio-economic data effectively. From October 2010, we have released research data and research result data in the DIAS. The 4-years increased amount of released data have improved the environment that we provide research data to the people not related to our project. Limited term project is difficult to continue to release research data in a sustained manner. It is necessary to establish a management system for long-term research data in anticipation that the DIAS is transferred to the long-term operational system in the future. In this paper, we describe the current status of the DIAS data release, and discuss the issues for sustainable maintenance of data and metadata in the DIAS.

キーワード: DIAS, 地球観測データ, データ公開, メタデータ  
DIAS, Earth Observation data, Data Release, metadata

## 1 はじめに

「データ統合解析システム(DIAS: Data Integration and Analysis System)」[注1]は、第3期科学技術基本計画の国家基幹技術「海洋地球観測探査システム」の一翼を担うプロジェクトとして2006年から開発が開始された。DIASは、分野を超えて共有できる知の創造、世界で共有できる知の創造、体感できるデータと情報の提供を目標としている。2006年からの5年間でプロトタイプの開発と実証を行い、DIASの長期的・安定的サービス機能の基礎開発を行った。2011年からは地球環境情報統融合プログラム(DIAS-P)が後継プロジェクトとしてDIASを開発している。現在はプロトタイプの高度化・拡張を行い、DIASの長期的・安定的サービス機能の応用開発を行っている。このプロジェクトの特徴は、多分野研究者が協力してデータ中心の新たな情報基盤を作る環境が整備されている点である。DIASではシステムの高度化・拡張を行いつつ複数の分野の研究者が協働して課題を解決するワークベンチのプロトタイプを実装している。そのためプロジェクトに参加している研究者のニーズを踏まえた地球環境データを整備するとともに、研究者が研究上作成したデータ等様々な地球環境データをDIAS内に蓄積してきた。

本稿では、DIASにおける研究データ公開の現状を報告し、研究データの公開とデータ・メタデータの持続可能な管理を行うための問題点を論じる。本稿の構成は次のとおりである。DIASについて2節で紹介し、DIASにおけるデータ公開の現状を3節で述べ、4節では公開データに関するメタデータの品質管理について、5節では持続的なデータ・メタデータの維持管理について、DIASの問題点と課

題を論じる。また最近の研究データに関する動きについて6節で述べ、最後に7節でまとめを述べる。

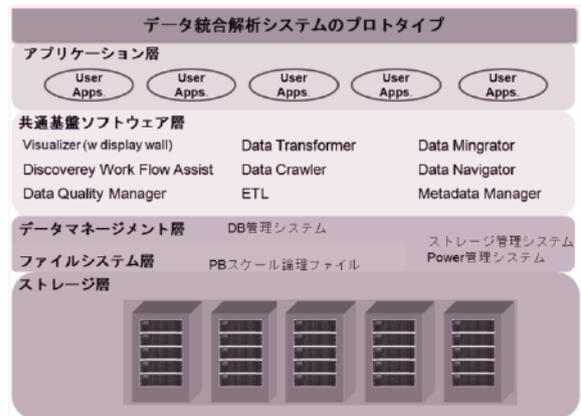


図1 DIAS システムアーキテクチャ階層構造

## 2 データ統合解析システム

本節ではDIASについて紹介する。DIASが開発するサーバ群は、2006年度から毎年増強を繰り返してきた。現在のハードウェア構成は複数の10ギガビットイーサネットのネットワーク上に接続された60以上のクラスタノード、解析サーバ、ファイルサーバとストレージを組合せたシステムである。ストレージは数十ペタバイトサイズのファイルシステムとしてディスクアレイとテープライブラリが接続されている[注2]。さらにサーバとストレージのミラーを国内に配置しシステム全体の冗長化を図っている。

また2016年からのDIAS運用開始を見据えて安定的なサーバ運用のための準備を進めている。図1はDIASシステムのアーキテクチャ階層構造を示した図である。ストレージ層、ファイルシステム層、データマネジメント層の上に共通基盤ソフトウェア層、アプリケーション層が構築される。共通基盤ソフトウェア層は主に情報科学研究者が担い、アプリケーション層は地球科学研究者と情報科学研

研究者の協働で研究開発を行っている。

DIASでは超大容量データストレージシステムを効果的に用いて、データのライフサイクル管理、データ品質管理、データ検索、情報の検索、科学的解析、データの部分取得などの支援機能を開発している[1, 2, 3]. DIASの特徴は、ハードウェアの構成変更を共通基盤ソフトウェア層、アプリケーション層と分離していることである。その結果DIAS内部のハードウェア構成変更を気にせずに地球環境科学の研究に必要なデータを共有し、各研究者が個別にデータをダウンロードすることを回避し、さらに研究者のプログラムをDIAS上で実行することができる。ストレージの共有と解析空間の共有が分野を超えた研究を支援している。

## 2.1 DIASにおけるユーザ

DIASを利用するユーザを管理するアカウントは1. 解析目的アカウント, 2. 共通認証基盤アカウント, 3. 特定の利用者向けアカウントの3種類に分けられる。

### 2.1.1 解析目的アカウント

地球科学研究者の研究のために設置されている解析サーバにログインし、DIASが保有するデータとユーザのデータを組合せてユーザ作成プログラムを利用して解析を行うため、あるいは大量のデータをDIASに転送し、フォーマット変換を実施するデータ提供者、さらに地球科学研究者が開発し、他の研究者の利用性が高いプログラムをDIASに提供するユーザに対するアカウントである。提供されたプログラムは、データ統融合のために必要な共通プログラムとして利用可能な環境を構築している。提供されたデータは、利用者に関してのアクセス制限を付与して

DIAS-Pと関連するプロジェクトの研究者が利用している。DIASデータ公開の対象となるデータの識別は、2.2節で述べるDIASメタデータ基盤システムが行う。

### 2.1.2 共通認証アカウント

DIASが提供するサービスは、Webブラウザを利用したアプリケーションとして開発する方針である。この共通認証アカウントを利用者がWebブラウザを介して利用するアプリケーション（今後Webアプリケーションと呼ぶ）用に利用する。DIASは様々なWebアプリケーションを開発しているが、公開時に統一的なログイン環境を利用者に提供するために共通認証基盤を構築している。この共通認証アカウントがDIAS利用者のシングルサインオンを可能にしている。

### 2.1.3 特定の利用者向けアカウント

DIASが提供するWebアプリケーションにおいて特定の利用者を対象としている場合の個別管理アカウントである。DIASが関与する特定の課題解決を目的としたWebアプリケーションを利用することが主な目的となっている。このアカウントは開発段階のWebアプリケーションのためのものである。Webアプリケーションが一般に公開できる状況になれば、共通認証アカウントによる利用に移行する。

DIASを利用する目的に応じてこれら3種類のアカウントを管理している。現在 解析目的アカウントの利用者が約200名、共通認証アカウントの利用者数は約700名、特定の利用者向けアカウントの利用者が約100名である。DIASデータ公開が対象としているユーザは共通認証アカウントを保有するユーザである。

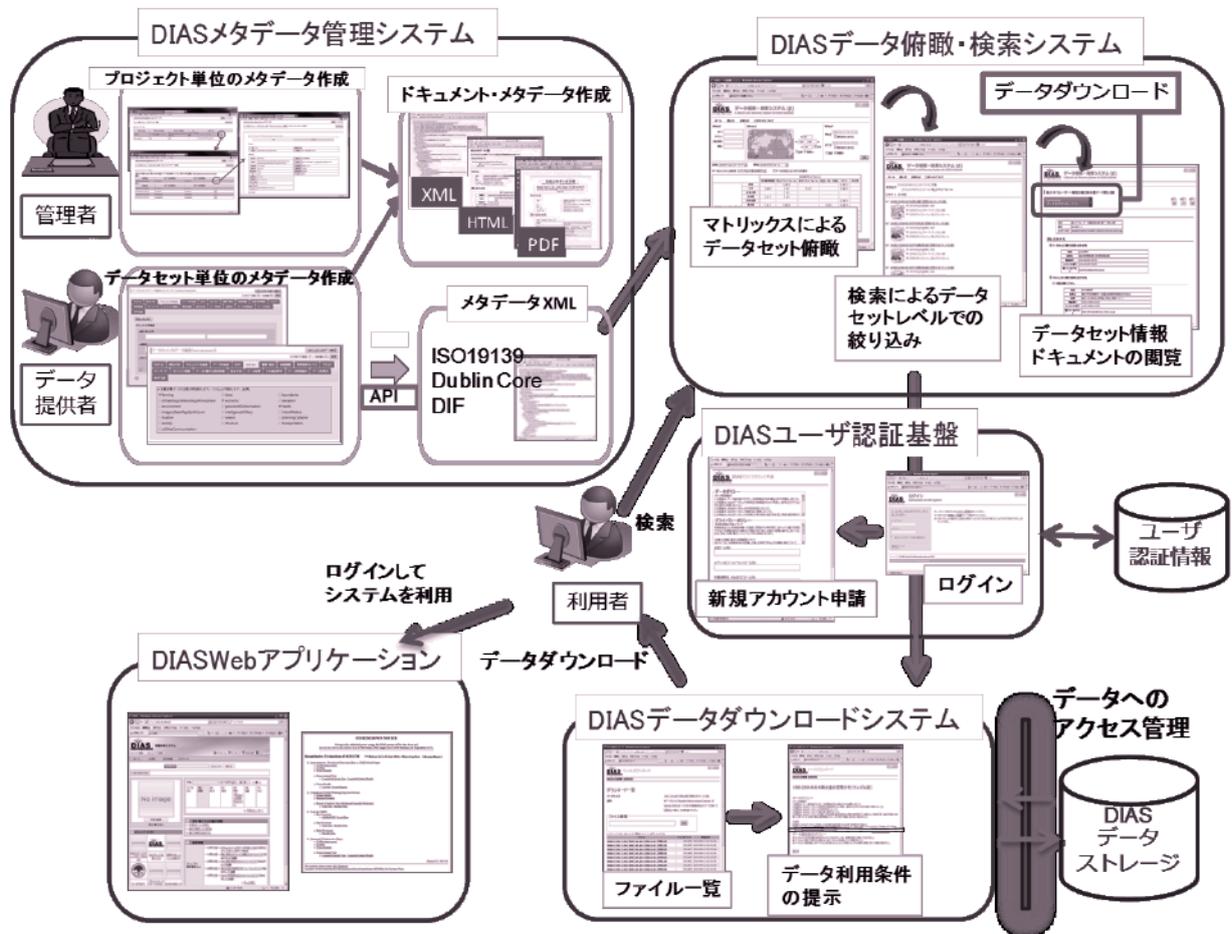


図 2 DIAS データ公開に関連する Web アプリケーションシステムの構成

## 2.1.4 DIAS-P 関係者

DIAS-P関係者は、DIASサーバの解析空間とストレージを利用して地球環境に関する気象、気候、生物多様性、水循環、炭素循環、農業、健康、災害や都市問題が相互に関連する課題を設定して研究を進めている。DIAS-P関係者は約100名である。DIAS-Pにおける最近の研究成果数は表1の通りである。

またDIAS-Pと大学発グリーンイノベーション創出事業「グリーン・ネットワーク・オブエクセレンス (GRENE)」環境情報分野[注3]との協働による課題解決に向けたワークベンチがDIASを利用して開発されている。研究内容はDIAS-GRENE 環境情報統融合フォーラムの発表資料として公開されている[注4, 5]。

表 1 DIAS-P の研究成果

	論文	口頭発表	ポスター発表
平成 24 年度	31	35	15
平成 25 年度	31	67	30

## 2.2 DIASメタデータ基盤システム

DIASはプロジェクト参加者が研究上必要とする地球環境データを集約しデータの統融合利用を促進する場を提供している。さらに研究に必要なデータや研究上作成されたデータプロダクトを研究プロジェクト以外の研究者とも広く共有するため、それらを公開する環境を整備してきた。DIASが保有する

データの公開を2010年10月に開始し4年が経過した。データ公開までの協働については先行研究で紹介した[4]。DIASデータ公開においては、データ提供者の所属機関、研究分野、データ容量やファイル数が多様であることを踏まえる必要がある。そのためデータ公開にあたり、データ提供者が提供するファイル形式に制約を設けずに受け入れることとし、データに対するメタデータを共通化することで多様な研究分野のデータに適用可能な公開方法とした。DIASが開発したメタデータ管理システムが管理するメタデータ項目は、多様な分野に適用可能な汎用的な項目に絞ったため、分野によって異なってくるデータファイルや解析ソフトウェア等とは独立して設計された。同時期に開発を進めている超高層大気長期変動の全球地上ネットワーク観測研究(IUGONET: Inter-university Upper atmosphere Global Observation NETwork) [注6]のメタデータデータベースの場合、アーカイブされているデータと解析ソフトウェアとの連動を考慮したデータベースとして設計されているのとは、設計の方針が異なる。

公開のために(1)保有データの把握、(2)データ提供者とデータ利用者の特定、(3)アーカイブデータの管理方法、(4)データを説明するメタデータ項目、を検討した。データ公開のために我々はメタデータ管理システムに加え、DIASデータ俯瞰・検索システム、DIASユーザ認証基盤、DIASデータダウンロードシステムを開発した。図2はDIASデータ公開に関連するWebアプリケーションとその基盤となるデータベースのシステム構成図である。

公開するデータのメタデータを維持管理するDIASメタデータ管理システム、公開デー

タの発見と利用のためのDIASデータ俯瞰・検索システム、公開データファイルをダウンロードするためのDIASデータダウンロードシステムに加え、これらのWebアプリケーション共通のアカウントを維持管理するためのDIASユーザ認証基盤、さらに特定のデータセット向けのインタフェースを開発したWebアプリケーションで構成される。DIASでは研究プロジェクト以外の利用者へ提供するアプリケーションは、Webアプリケーションとして開発している。表1の研究成果には、これらのシステム開発に関する成果が含まれる。

## 2.3 メタデータを介した他データセンターとの連携

我々は2012年6月から文部科学省地球観測事業を対象として事業で収集したデータや作成されたデータに関するメタデータの調査とメタデータ作成・収集を行っている[注7]。この調査結果を踏まえ、メタデータや公開データベースが既に作成されていて、しかも他のデータセンターとDIASの同時利用によりデータ利用者への相乗効果が見込まれるデータセンターを選定し、データの相互利用をめざし連携を始めた。データの所在を気にすることなくデータ利用者がデータを利用する環境を構築することが理想であるが、連携するデータセンターに新たな開発が発生しない状況を前提として実現可能な環境構築を行っている。その結果、それぞれのデータセンターが管理するデータファイルの所在までDIASが把握するような密な連携ではなく、それぞれのデータセンターが管理しているデータに関するメタデータへのアクセス性による連携を行うことにした。

### 2.3.1 メタデータ受信

現在、海洋開発研究機構(JAMSTEC)データ

カタログ[注8], 国立環境研究所日本長期生態学研究ネットワーク(JaLTER) データベースのデータ目録[注9], 国立極地研究所学術データベースのメタデータ[注10] をDIASが取得し, DIASデータ俯瞰・検索システム[注11] に検索対象として登録している. これら連携先のデータセットが検索結果としてDIASデータセットと同時に表示され, 連携データセンターのメタデータ表示サイトへの誘導が可能となっている. 連携データセンターから取得するメタデータの書式はDIASが作成しているメタデータの書式と異なり, 様々な形式のメタデータ項目間の対応付けを行いながら連携先を増やしているところである. 連携データセンターのメタデータ更新へは未対応である. 定期的なメタデータ取得の仕組みについては, 今後連携データセンターと個々に協議して進める必要がある. 現在国土地理院地理空間情報ライブラリー[注12] のメタデータを取得し, 具体的な連携方法を検討中である. このメタデータは, データファイルと対になっているため, DIASのメタデータと国土地理院のメタデータではメタデータを付与するデータの粒度が異なる. したがって取得したメタデータをDIASデータ俯瞰・検索システムへ登録するためには, 適切な粒度のメタデータへの変換が重要である. 現在変換方法を検討中である.

### 2.3.2 メタデータ配信

2014年1月からDIAS公開データに関するメタデータをGEOSSポータル[注13] に提供し, GEOSSポータルにおいてDIASのデータを検索することが可能になっている. GEOSSポータルへのDIASメタデータ提供は, 年4回実施する予定である. DIASが公開しているデータセットに対応したISO19139フォーマットのメタデータXMLファイル(英語表記)をGEOSS

ポータル向けに加工し, GEOSSポータルへ提供している.

メタデータの自動配信については, 国際バーチャル天文台(International Virtual Observatory Alliance) のIVOA標準仕様[5] が採用しているOAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) [注14, 15] によるメタデータの自動配信や, STARSプロジェクトにおけるRDD/RDF を利用したメタデータの配信を検討した先行事例がある[6]. メタデータの更新をリアルタイムに通知する際には有効な手法である. DIASではデータセット単位のメタデータの配信方法についての検討は行っているが, メタデータの更新を外部に自動的に通知する機能は未開発である. OAI-PMH機能などメタデータ更新部分を配信する機能による高度化を検討中である.

### 2.3.3 メタデータ仲介システムの検討

連携先データセンターの増加によりDIASが取得するメタデータの書式の数も増加する. 連携先のメタデータの内容と書式についてDIASメタデータとの対応付けをするのは, 連携先データセンター数が増加することでより複雑になる. そのためDIASでは多様なメタデータの書式を統一的なメタデータの書式に変換するためのメタデータ仲介システムを構築する予定である. イタリアESSI Lab(Earth and Space Science Informatics Laboratory) [注16] が開発し, GEOSSポータルが採用しているGCI(GEOS Common Interface) [注17] に利用されているブローカーDAB(Discovery and Access Broker) 機能を採用する予定である. DAB では多様なデータセンターのメタデータフォーマットをGCIの共通メタデータフォーマットに変換する機能を保有している[7]. この機能をDIAS

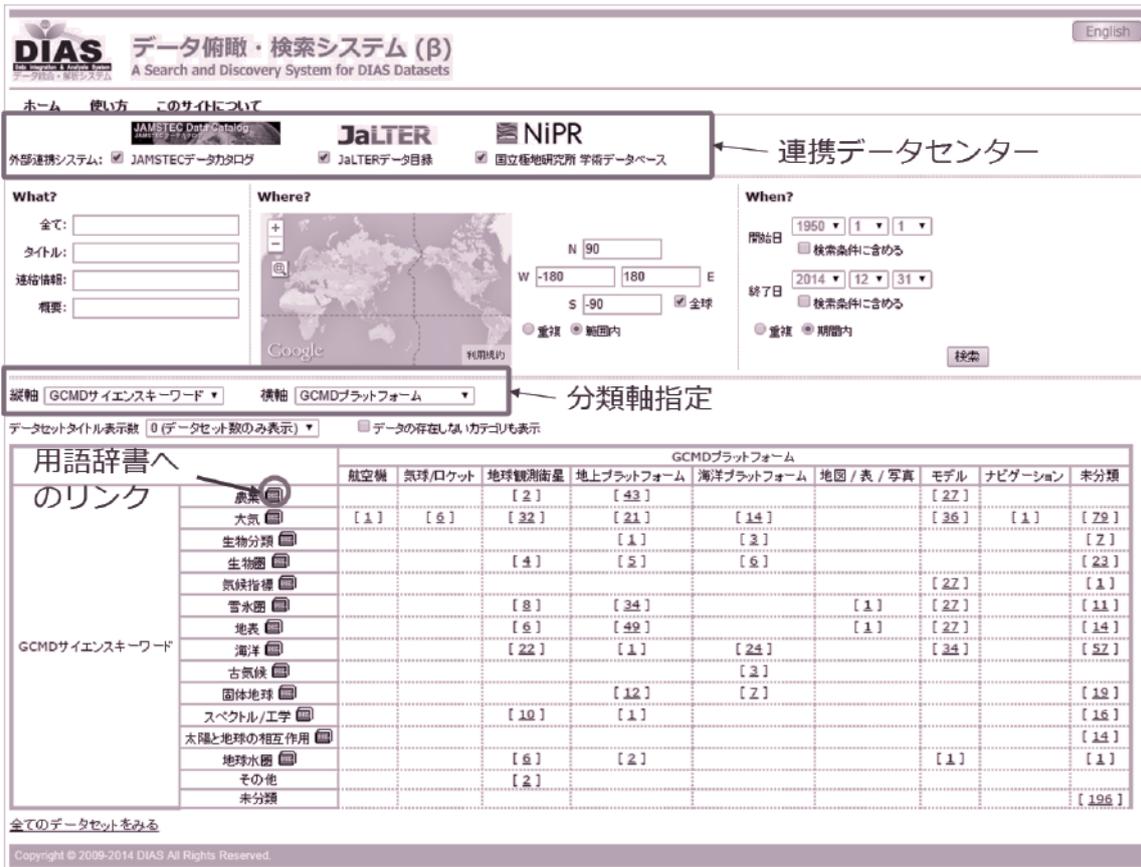


図3 DIAS データ俯瞰・検索システム

に導入し、多様なメタデータの書式を一元化してDAB が実装しているOAI-PMH 機能を利用することを検討中である。

### 3 データ公開の現状

本節ではDIASデータ公開の現状について述べる。DIAS公開データへの入口として我々はDIASデータ俯瞰・検索システムを開発してきた。DIASデータ俯瞰・検索システムからDIASが公開しているデータのドキュメント・メタデータの閲覧とデータダウンロードサイトへのアクセスが可能である。次にDIASデータ俯瞰・検索について述べる。

#### 3.1 DIASデータ俯瞰・検索システム

図3はデータの俯瞰・検索システムサイトである。一般的なカタログ検索サイト同様検索内容をいつ(When), どこで(Where), 何を

(What) に関して利用者が対象データを検索することができる。図3の上半分は一般的な検索条件指定項目である。下半分は、分野や作成手段等によりデータセットを表形式で分類しており、セル内のリンクをクリックすることで対応するデータセット一覧を取得することができる。分類軸としてGCMD サイエンスキーワード[注18], GCMDプラットフォームとGEOSS社会利益分野[注19] を指定し、切り替えることができる。GCMDサイエンスキーワードについては、専門用語辞書への参照を設け用語に関する理解を支援している。

DIASデータ俯瞰・検索システムではDIASが公開しているデータに加え2.3.1節で述べたように、国内のデータセンターとメタデータを通じた連携を行っている。DIASデータ俯瞰・検索システムではDIASメタデータと同時に連携データセンターのメタデータを検索

することができる。

DIASデータ公開についての特徴は以下の4点である。

1. DIASが公開するデータの俯瞰，検索，データダウンロードをシームレスに行えること。
2. ゲストユーザはユーザ登録をすればデータダウンロードが行えること。さらに
3. データ提供者がデータへのアクセスについてユーザごとにデータへの許可・不許可を決めることができること。
4. DIASからデータ提供者へのデータダウンロード記録の報告が可能なことである。

検索対象となるデータセット数の増加に従い検索結果として表示されるデータセット数も多くなっている。開発当初は、DIASが保有するデータセットを分類して一覧にすることが分野間連携のために最優先であったが、順位づけられた検索結果の表示機能、検索結果に対して再分類して絞り込む機能など検索機能の高度化は未開発である。さら

に連携データセンターのメタデータから対象となるデータセットをDIASデータ俯瞰・検索システムに表示される該当するセルに自動的に分類する手法の開発も課題である。

約4年にわたりデータ公開を行い公開対象データ数、ユーザ登録数は徐々に増加してきた。次に公開データについて述べる。

### 3.2 公開データ

DIASが公開しているデータには、地上観測データ、衛星観測データ、気象予測モデル計算結果データ、気候変動モデル計算結果データとその他データがある。2014年8月末で202データセット、2ペタバイト、約4,000万ファイルを公開している。図4が公開データセット数の推移である。また2014年8月末で公開データのダウンロードを行えるユーザ数は約700、ユーザの所属機関数は約180である。平成25年末で公開データのダウンロードを行えるユーザ数は約700、ユーザの所属機

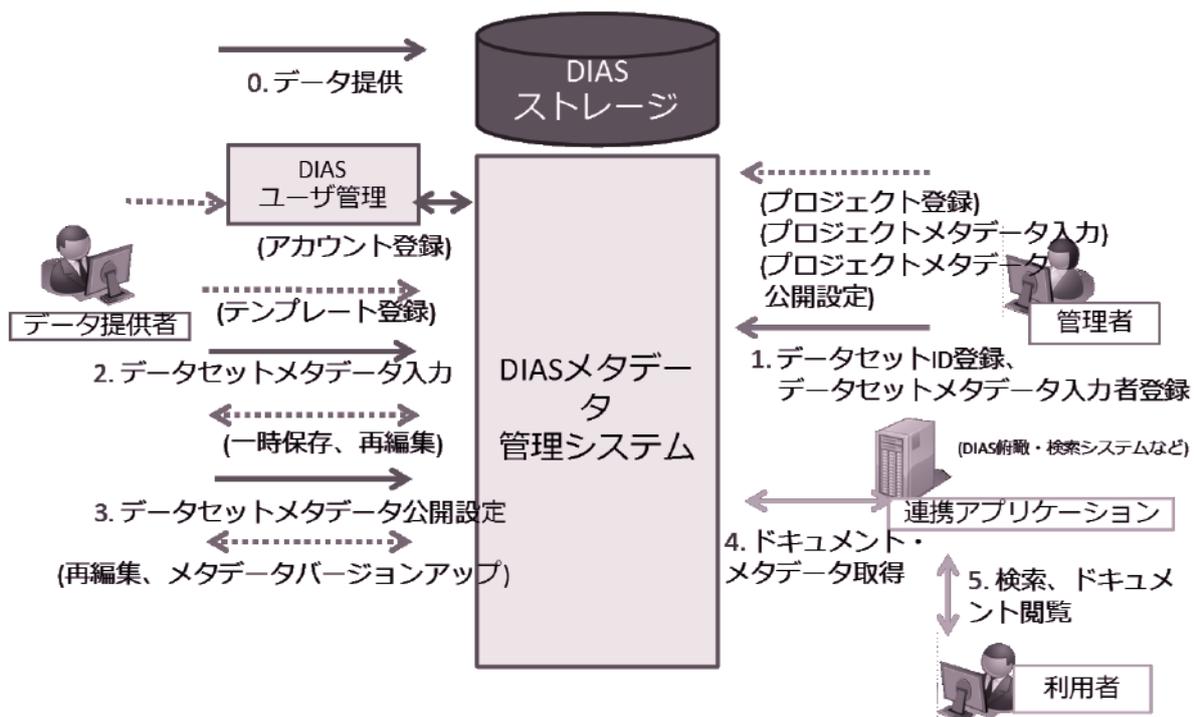


図5 DIAS データ公開プロセス

関数は約180 である。データダウンロード回数は、平成25年度(2013年度) 約5,000回、平成26年度(2014年) 4月から7月までで約6,000回である。



図4 公開データセット数推移

### 3.3 DIASデータ公開手順

4年あまりの運用で、データ公開に関する手順が整ってきた。当初はデータファイルとドキュメント・メタデータを利用者が閲覧する環境の構築が重要だとシステム開発に力点を置いていたが、運用によりアクセス権やデータポリシーの記述の重要性が高まって

きた

図5はデータ提供者がDIASからデータを公開するまでの手順である。データ公開までは7段階の過程を経る。

この7段階の流れはデータ公開のために順番に実施されることが理想だが、現状ではメタデータ作成の労力が大きくすべてのメタデータ項目の記入が完了するためには、データ提供者では不明なデータ利用規約の確認などが必要なため時間を要する。さらに研究の進捗によってはメタデータを先行して作成し、公開時期まで保留する場合もある。

#### 3.3.1 データ提供の連絡とデータセットIDの登録

データ提供者がDIAS管理者にデータ提供の予定を伝え、データファイルの送付方法をデータ提供者とDIAS管理者が協議する。データ送付後DIAS管理者がデータファイルをDIASストレージに配置する。

DIAS管理者がデータ提供者にデータセットIDとメタデータ入力担当者を決めるよう要

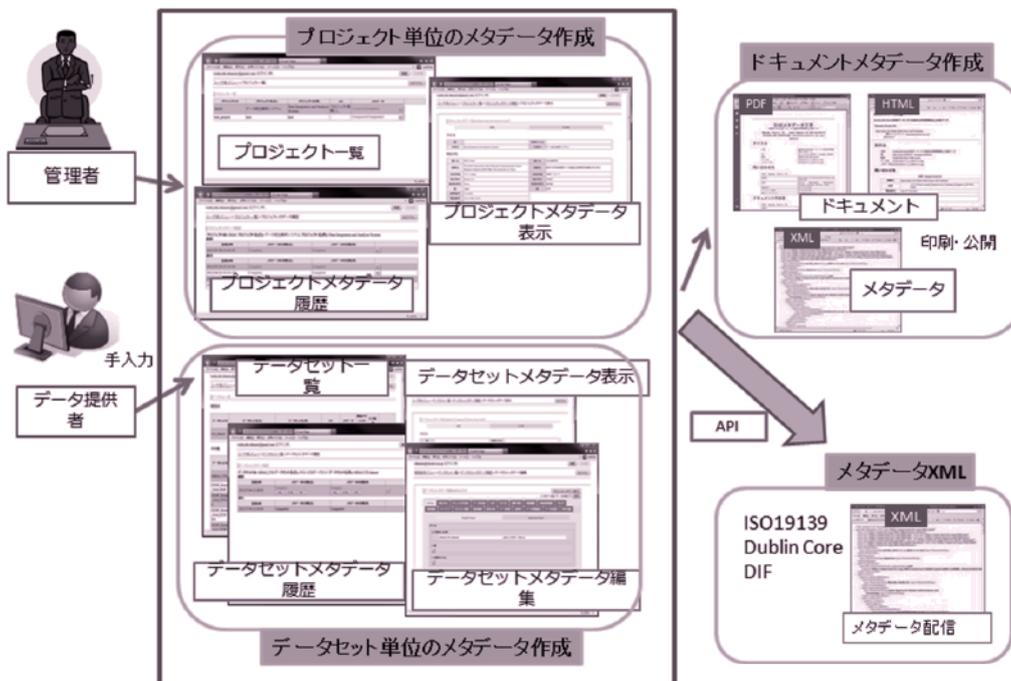


図6 DIAS ドキュメント・メタデータ作成概要

請し、データセットID とメタデータ入力担当者をDIASメタデータ管理システムに登録する。このデータセットID はDIASが公開するデータセットを識別するために利用する識別子となっている。将来的にはデータセットに対する永続的な識別子を導入する場合を考慮している。メタデータ入力担当者のDIAS共通認証アカウントを確認し、未登録の場合はアカウントを登録する。

また提供データセットが何らかのプロジェクトの研究成果であり、プロジェクトに関するデータ利用規約や連絡先をメタデータとして関連付けておくことが必要な場合は、プロジェクトに関するメタデータをDIAS管理者が作成する。

なおデータセットの粒度は常に検討課題となっている。原則としてデータ提供者の判断でデータセットID を付与するデータセットの粒度を決めることにしている。運用上は同一データセット内のファイルについては、すべて同一のデータ利用規約が適用されるという条件を満たすことを求めている。

### 3.3.2 データセットメタデータ入力

メタデータ入力担当者はDIASメタデータ管理システムを利用してデータセットID に対応したメタデータを登録する。入力が完了した時点でドキュメント・メタデータ (HTML, PDF, XML) が作成され、Webサイトからドキュメント・メタデータにアクセスできるようになる。メタデータ入力者が作成するメタデータはデータセットに関するメタデータである。プロジェクトに関するメタデータが関連付けされている場合は、それも含めてドキュメント・メタデータが作成される。図6 がドキュメント・メタデータ作成の概要である。

DIASメタデータ管理システムにはプロジェクトメタデータ作成機能とデータセット

メタデータ作成機能がある。プロジェクトメタデータ作成とデータセットメタデータ作成には一覧表示機能、編集履歴表示機能、表示機能と編集機能がある。このシステムでは英語と日本語のメタデータを同時に作成するインタフェースを提供している。データセットメタデータ入力者が入力を完了した時点で、プロジェクトメタデータとデータセットメタデータの内容からドキュメント・メタデータ (HTML, PDF, XML) が生成され、Webでアクセス可能となる。またこのDIASメタデータ管理システムにはメタデータ入力者、作成データセットに関するメタデータの管理に加え、類似メタデータへの内容コピーが容易に行えるようテンプレートの保存と利用機能がある。さらOAI-PMHの一部機能を実装しており、この機能を利用してメタデータ作成が完了したデータセットを確認している。

### 3.3.3 データセット公開設定

DIASからデータを公開するに当たり、データファイルへの利用者のアクセスに許可申請が必要か否かをデータ提供者が指示する。DIASではデータをダウンロードする際、データ提供者のデータ利用規約を考慮したデータへのアクセス権を次の4段階で設定している。

- (1) DIASユーザ登録が必要
  - (2) (1) に加えデータ利用規約への同意が必要
  - (3) (2) に加えデータダウンロードシステムを通しての利用申請とその許可が必要
  - (4) (2) に加えデータ提供者との個別交渉による承認が必要
- (3), (4) の場合は、データ利用者からの利用申請の許可・不許可を判断するための担当者を決める必要がある。

アクセス権の指定を含め、DIAS公開サーバ



図7 DIASドキュメント例

に格納されているデータセットに対応したファイルリストと作成されたドキュメント・メタデータをDIASデータダウンロードシステムに登録する。DIASデータダウンロードシステムがドキュメント・メタデータとDIASストレージのファイルとの関連付けを管理する。詳しくは3.5節で述べる。

### 3.3.4 連携アプリケーションによるドキュメント・メタデータ取得

DIASが開発しているDIASデータ俯瞰・検索システムなどのWebアプリケーションがドク

ュメント・メタデータを取得し、システムに登録することで検索ができるように設定する。その結果このデータセットが利用者の検索対象となる。

### 3.3.5 データ利用者による検索

利用者はDIASデータ俯瞰・検索システムなどで公開データを検索し、検索結果としてドキュメント・メタデータを閲覧することができる。

### 3.3.6 データダウンロード

利用者はドキュメント・メタデータを閲覧

し、ドキュメント・メタデータ中のデータダウンロードサイトへのリンクを通してデータのダウンロードができる。

図2中のDIASデータ俯瞰・検索システムにおいて、データセットの検索とドキュメントの閲覧を行い、データダウンロードが可能なデータの場合は、「データダウンロード」ページへのリンクが表示される。DIASデータダウンロードシステムへのアクセスには共通認証アカウントによるログインが必要である。データダウンロードサイトでは同一データセットID に登録されているファイルの一覧が表示される。ダウンロードしたいファイルを選択し、ダウンロードボタンを押すとデータ利用条件やデータ利用申請手続きを経てデータのダウンロードが可能となる。利用申請が必要なデータセットの場合は、利用申請フォームが表示され申請内容についてデータ提供者の担当者の判断の結果許可された場合は、データのダウンロードができる。

### 3.4 DIASドキュメント・メタデータ

DIASでは多種多様なデータに対してメタデータを記述する必要があり、データの概要を共通な記述方法で提供することが分野を超えたデータ利用を促進させることに寄与するという観点から、データセットに対するメタデータ作成のためのフレームワークをドキュメントメタファを利用したメタデータ作成として設計し、メタデータ登録システムを構築した[8]。このシステムによりデータ提供者には統一された形式のドキュメントを、そして我々はメタデータXML ファイルを取得することが可能となった。図7 は生成されるドキュメントの例である。データに関する問合せ先やデータダウンロードサイトへのリンク、データ利用規約を重視している。

表2 ISOコア項目とDIASドキュメント項目の対応

	ISO19115コアデータ項目	DIASドキュメント項目
1	Dataset title (必須)	1. TITLE
2	Dataset reference date (必須)	6. DATE OF DATASET
3	Dataset responsible party	2. CONTACT
4	Geographic location	7.4 Geographic Bounding Box
5	Dataset Language (必須)	非表示
6	Dataset character set	非表示
7	Dataset topic category	7.2 Topic Category(ISO19139)
8	Spatial resolution	7.5 Grid (必須)
9	Abstract	7.1 Abstract (必須)
10	Distribution format	7.10 Distribution Information
11	Vertical, temporal extent	7.3 Temporal Extent
12	Spatial representation type	非表示
13	Reference System	非表示
14	Lineage	8. DATA PROCESSING
15	On-line resource	7.8 Online Resource
16	Metadata file identifier	1. TITLE (Metadata Identifier)
17	Metadata standard name	非表示
18	Metadata standard version	非表示
19	Metadata language	非表示
20	Metadata character set	非表示
21	Metadata point of contact (必須)	3. DOCUMENT AUTHOR
22	Metadata date stamp	5. DATE OF THIS DOCUMENT

データのディレクトリ構造やファイル構造についての記述欄は特には設けていない。

我々が作成するメタデータは、地理空間情報に対するメタデータ標準である ISO19115[9], ISO19139[10] を基本とし、ISO19115 の主要な項目とデータセットドキュメント項目との対応付けをする方法を採用している。表2はISO19115コア項目とドキュメント項目との対応である。

### 3.5 DIASデータダウンロードシステムへのデータファイル登録

本節ではDIASデータダウンロードシステムへのデータファイル登録に関する手続きについて述べる。メタデータ管理システムとデータダウンロードシステム(データダウンロード機能)を独立したシステムとして開発したため、公開データセットのメタデータとデータセットID、アクセス権、データファイ

ルの保存場所などの関連付けは人手で管理している。

データファイルの登録からデータ公開までの流れは以下の3段階の手順となる。

1. データセットが公開可能な状況，すなわち(1)DIAS公開サーバにデータファイルが格納されていること，(2) データセットに関するドキュメント・メタデータの作成が完了していること，を確認する。
2. DIASデータダウンロードシステムに新規データセットを登録する。登録内容は，データセットID，データセット名，概要，データ利用規約，データ引用規約，公開されているドキュメント・メタデータのWebサイトのアドレス，利用許可申請が必要か否か，公開するファイルリストである。
3. DIASデータダウンロードシステムではデ

ータセットに対応したダウンロードのWebサイトアドレスが割り当てられる。そのアドレスをDIASドキュメント・メタデータのオンライン情報(図7の7.8節部分)に追加する。

将来的には，データ公開プロセスを含んだDIASデータ登録・メタデータ管理・データ公開のワークフローを管理するシステムを構築する必要がある。現状では，その基本的な部分のプロトタイプを作成し，改良を加えているところである。図8はデータ公開状況の一覧表示画面とデータセットごとのデータ公開状況入力画面である。

#### 4 公開データに関するメタデータの品質管理

2010年10月にデータ公開を開始した後，多様なデータおよびメタデータに対するシス

DIASデータ公開状況一覧

no	Dataset ID	データ公開	メタデータ作成者	日本語データセット名	英語データセット名	メタデータ日本語	メタデータ英語	登録日	詳細
207	AGURAM_GI	公開	heromiya@ccis.u-tokyo.ac.jp	AGURAM全球都市域マップ地上情報データベース	AGURAM Ground Information Database	日本語	英語	2014-02-06	
196	AlligatorHope	公開	noji@nriez.go.jp	日本-アメリカ西海神間の定期貨物船(Alligator Hope)を利用する大気海洋二酸化炭素モニタリング	Atmosphere and ocean carbon dioxide monitoring using volunteer observing ship (Alligator Hope) servicing between Japan and West Coast	日本語	英語	2013-02-21	
216	AMY_CTCZ_Pilot_dataset	公開	lhsk@cees.iiso.ernet.in	AMY Project CTCZ-Pilot dataset	AMY Project CTCZ-Pilot dataset	日本語	英語	2014-04-04	
208	AMY_TRITON_buoy	公開	tritong@amstec.go.jp	AMY Project TRITON buoy dataset	AMY Project TRITON buoy dataset	日本語	英語	2014-02-21	
184	AWCI_Eang				Global Earth Observation System of Systems				
186	AWCI_Ehut								

データセットごとの公開状況入力画面

Dataset description

No: 207

id: AGURAM\_GI

Dataset released:  公開  非公開

連絡先メールアドレス: heromiya@ccis.u-tokyo.ac.jp

Dataset Description: 全球の都市域のデータ

Dataset Title: Japanese AGURAM全球都市域マップ地上情報データベース  
English AGURAM Ground Information Database

DIAS download system URL: http://dias-dss.tkl.iis.u-tokyo.ac.jp/d/storages/llists/dataset/207

格納ディレクトリ: /dias/data/aguram-gi

登録日時: 2014-02-06

改訂日時: 日時 0000-01-01 merno

メタデータ: Japanese http://dias-dmg.tkl.iis.u-tokyo.ac.jp/dmm/doc/AGURAM\_GI-DIAS-ja.html  
English http://dias-dmg.tkl.iis.u-tokyo.ac.jp/dmm/doc/AGURAM\_GI-DIAS-en.html

履歴メモ1

データセット保存 保存せずにデータセット一覧に戻る

図8 DIAS データ公開状況の管理画面

テムの対応可能性を高めるために、継続的にメタデータ書式の改良、メタデータ管理システム、俯瞰・検索システムの高度化を行い、優先度の高い事項を各システムに実装してきた。さらに公開されているメタデータの入力内容を確認し、過不足部分についてメタデータの改訂を行ってきた。

データ提供者がそれぞれ担当のデータセットに関するメタデータの作成を行った後、DIAS管理者は各データセットのドキュメント・メタデータを閲覧して作成された内容を横断的に精査してメタデータの品質管理を行っている。2013年度に行った内容の確認は146箇所であったが、そのほとんどはDIAS管理者が表記等の訂正を行うことで解決されている。品質検査を行った主要項目は次の5項目である。

#### 4.1 データ・メタデータの連絡先や担当者の変更

ドキュメント・メタデータ項目の中に記述されている連絡先や担当者変更は常に発生するものであり、これらの項目の維持管理はメタデータの品質維持のため重要であるが、可能な限りDIAS管理者が変更の有無の把握に努め関係者と連絡をとることで解決している。

#### 4.2 データダウンロード先のアドレスの記入

DIASが保有しているデータセットについては、公開非公開にかかわらずドキュメント・メタデータを公開する方針である。そのため非公開のデータではデータダウンロード先のアドレスが記入されていない場合がある。またデータセットによってはDIASからデータを公開している場合に加え、自サイト

でも公開している場合がある。データ利用者のデータへのアクセス性を高めるため、データダウンロード先のアドレスが遺漏なく正確に記述されているように確認を行っている。

#### 4.3 データセットの収録地理範囲の記述方法

地理範囲の緯度経度指定の表記方法についてメタデータ入力の際に表記ゆれが多く発生していた。どのような表記ゆれがあるかを確認し、表記ゆれを解消するような入力チェック機能改善を行っている。この項目は地理範囲を指定した検索で重要な役割を担っている。

#### 4.4 データセットに関するカテゴリやキーワードの選択や記述方法

データ発見のためには多様な分類軸でデータを分類しておきたいが、データ提供者には各分類軸についてなじみがない場合が多い。横断的な見直しによりデータセットが該当するカテゴリや該当するキーワードを複数の専門用語辞書に関連付けるため現状の入力内容を確認している。この項目については、データ提供者に対してキーワード入力支援機能を提供する必要がある。データ提供者が入力するキーワードから類似したキーワードを複数の専門用語辞書からキーワード候補を提案し、容易に選択できるような枠組みを検討中である[11]。

#### 4.5 データセットに関する参考文献の記述方法

データ提供者が掲載可能な文献情報をDIASではメタデータ項目として入力欄を設けている。この項目はISO19115メタデータ項

目には明示的な定義がないため、現状では表記についてはデータ提供者に任せている。DIAS管理者が可能な限り表記統一を行っている。またこの項目にデータ利用者の文献をデータ提供者が入力することは現実的ではない。今後データと参考文献の関連性はこのドキュメント・メタデータ内部で維持管理するのではなく別な方法で行う方向で検討している。

## 5 持続的なデータ・メタデータの維持管理

DIASを運営しているプロジェクトは2015年度に終了し、DIASは地球観測データ統融合に関する実運用組織に引き継がれる予定である。実運用組織では「地球環境情報の統合的アーカイブ」、「超大容量・多様なデータの統合解析環境」、「多様な協働を形成し展開・流通させる情報基盤」、「持続可能な社会・国土形成のための予測情報」を提供することが目標となっている。研究・開発段階のDIASではデータ公開についてもプロトタイプを開発し、できるところから解決方法を提案してきた。実運用組織への移行においては、正確で確実なデータ公開が求められる。そのためにはシステムのセキュリティ管理、ユーザの個人情報保護、データのポリシー管理、アクセス権管理、データ不正利用に対する対策などが必要になる。持続的なデータ・メタデータの維持と管理を行う組織は持続的な組織が期待されている。実運用組織について現時点では未確定である。

### 5.1 DIASの問題点

DIASの開発は、毎年度の業務計画と研究課題への対応をチェックポイントとして進んできた。2015年度で終了するDIAS-Pで今まで

に挙げた課題をすべて解決するのは困難である。本節ではDIASの問題点について述べる。

DIASは大学が開発しているシステムである。第一の目的はDIAS-P参加者が研究成果を出すことである。そのためおおぜいの人にとって役立つ共通基盤づくりに対する制度設計は大学内プロジェクトとして行うには、人材や時間が限られている。また具体的な課題解決に結びつく開発が優先されるため、長期的な制度設計担当者と100名以上の研究者が今後のDIASの方向性に関し議論をする場を持たないまま時間が経過している。実運用組織に移行する時点では、現在の開発者が運用組織で運用に携わる担当者にデータやシステムに関する内容を伝える必要がある。残り時間は少なくなっているが、可能な限りスムーズな移行作業が行える環境が必要である。

次にDIASデータ・メタデータに関する問題点と課題をまとめる。

### 5.2 DIASデータ・メタデータに関する問題点と課題

今までに開発してきたDIASデータ俯瞰・検索システム、メタデータ管理システム、データダウンロードシステムに関する課題はすでに2.3.2節、2.3.3節、3.1節、3.5節、4節で述べてきた。DIASが採用したメタデータの標準ISO19115, ISO19139は地理空間情報を対象として設計されたものであるがこのメタデータ標準を採用しているデータセンターは比較的多い。しかし実際に連携データセンターから取得したメタデータの記述には英語、日本語だけのものが存在する。DIASでは英語と日本語のメタデータを作成しているが、DIASの利用者が研究者に加え行政担当者や一般市民に拡大する場合は提供する

メタデータの言語での制約が大きな問題である。また2.3.1節で述べたメタデータを付与するデータとして何を選ぶかは引き続き問題である。

### 5.3 高品質なデータの維持管理

我々は現地観測データを収集する仕組みを構築してきた。現地観測データファイルを保有するデータ提供者がWebインタフェースを利用してデータをアップロードし、品質管理を行い品質に関するフラグを付与しうえ、観測項目に関する詳細情報を含むメタデータを作成する。さらにデータ公開のためのポリシーを確認してDIASから公開する流れをアジア水循環イニシアチブ、アジアモンスーン年参加プロジェクトについて構築してきた[12, 13]。研究に利用する場合、研究利用可能なデータの品質情報が利用可否の判断尺度として求められる。

DIASでは多種多様な大量のデータの受け入れを優先してきたため、データセット単位のメタデータ記述を比較できることに力点を置いて開発を行ってきた。そのためデータセット内のディレクトリ構造、ファイルフォーマット、観測項目名、シミュレーション変数名、欠損値や観測機器の不具合などの記述については、DIASが関与することがなかった。

例えば、研究者がよく利用しているnetCDF (Network Common Data Form) データフォーマット[注20]のような機種独立なフォーマットで記述されたデータは、共通プログラムで解析を行うことが可能でありデータ利用者へ提供する場合に適している。DIASが公開しているデータセットの一部はこのフォーマットのファイルである。またnetCDFには分野ごとに共通の変数名等を推奨する規約を定めている。DIASが公開するデータについてもデータ提供者にフォーマットに関する

ガイドラインを提案し、より利便性を高めるべきだとしてガイドライン導入に向けての協議を始めている。

## 6 研究データ管理計画とオープンデータ

本稿では、DIASにおけるデータ公開と課題について述べてきた。研究データの公開に関して世界の動きが加速している。本節では研究データ管理計画とオープンデータの動きとDIASの今後の課題について述べる。

「科学的プロセスは研究データの管理と共有によって向上する。正しいデータ管理は、信頼できる結果検証と既存情報を利用した新しく革新的な研究を可能にする。研究データが適切に体系付けられ、文書化され、保存され、アクセス可能な状態であり、その精度と妥当性が常に管理されていれば、結果として高品質なデータ、効率的な研究、確固たる証拠に基づく発見、そして時間と資源の節約がもたらされる。」という考え方が広く認められるようになってきた[14]。研究データの共有基盤の開発の重要性が高まっている。研究データは従来から研究者の手元に紙ベースで、あるいはパソコンの中に保存されているものが多い。今後はそれらを共有の資産として利用する基盤を作ることも重要である。

米国では研究資金を得るためには研究データ管理計画を提出することが義務化され、さらに公的資金で行った研究成果のデータは公開が義務付けられるようになってきた。英国、米国をはじめとして研究資金提供者は、研究費申請段階におけるデータ管理計画作成を義務づけるようになってきた。研究データの管理と共有の実施計画を各国のデータセンターや大学図書館がサポートする必要にせまられている。この流れは我が国にも近

い将来及ぶことが見込まれる。

図書館情報学の研究者は論文のデジタル化に加えデジタルデータのアーカイブについての検討を始めている[注21][15]。持続的なデータ・メタデータの維持管理を行う場合、データ保有・管理を行うためにはデータアーカイブに関する人材と知識が必要になる。池内は研究データマネジメントとデータキュレーションについて図書館情報学の立場からの最近の動向を調査している[16]。研究データアーカイブにおいては、従来の図書や論文のアーカイブとは異なった技術やサービスを必要としている。また真板は研究者の立場から研究データ公開に関する人材や体制の問題点を論じている[17]。研究データを扱う人材が必要としている背景知識や技術は、現在の図書館員とは異なる。また研究者も研究データ管理に対して従来よりも注意を払う必要がでてくる。

この研究データに対する管理と共有を誰が担うべきか、どんな人材が必要かという議論は国内では始まったばかりである。DIAS開発の中では、データ公開の方法やメタデータの維持管理については具体的な事例に当たるであろう。しかしデータ共有に関する倫理的および法的問題、データの著作権や知的財産権など地球科学や情報科学研究者の範疇では解決できない問題が山積している。これらの問題に対して、国際的な組織RDA (Research Data Alliance)[注22]やベルモントフォーラムE-インフラストラクチャ[注23]によって研究分野を超えた共通合意に向けた議論がなされている。その議論を踏まえて2016年度からのDIAS実運用組織において問題解決への方向性を見出す必要がある。

特に、分野を超えたデータの共有促進のためにもデータ公開に関する様々な議論の場

が重要である。データ公開においては「オープンデータ」や「オープンガバメント」が推進されている。総務省における「オープンデータ戦略の推進」によって公共データの二次利用可能な形式での公開を5段階で推進する方針を示している[注24]。また日本に国際プログラムオフィスが置かれている国際科学会議(ICSU: International Council for Science)の科学データシステム(WDS: World Data System)ではオープンデータを推進している[注25]。

公的な研究資金の成果として作成されたデータのオープンデータとしての扱いについても今後議論が必要である。DIASでは、メタデータを介してデータへのアクセス性を高めている点では、オープンデータの方向性と一致している。しかし、DIASが公開するデータのデータ利用規約には二次利用を認めていない。データの由来や来歴の保証、著作権や知的財産権などの問題点を解決しながらオープンデータとして公開が可能か検討する必要がある。DIASに関係するすべての研究者に研究データ管理を徹底するには、早い段階からの教育や啓蒙が重要になるであろう。

研究者がデータを作成し、公開するための努力を評価する機会がないため、研究者のデータ公開に関する関心は決して高くなかった。WDSでは研究データを論文公開と同様に行うためのデータパブリケーションについても推進している。研究論文の引用回数が研究者を評価するのと同じように研究データの引用回数で研究者を評価することを可能にするデジタルデータに対する識別子としてDOI (Digital Object Identifier)[注26]の導入が始まっている。WDSに参加するメンバーは、持続的に同一のデータにアクセスで

きる環境を整える為に、研究データへのDOIの導入を進めている。DIASが保有するデータや公開しているデータへのDOI導入を検討する必要がある。生物多様性・生体系分野では公開データを評価するデータペーパーの出版も行われている[注27, 28]。DIASが公開するデータを広めデータに関する評価に結びつけるための手法の開発は、今後の課題である。

## 7 まとめ

本稿ではDIASにおけるデータ公開について述べた。DIASデータ公開に関する作業は、今後の研究データ共有に向けての先進事例のひとつである。

DIASは複数分野の研究者と情報科学研究者の協働によって地球環境が抱える問題を解決する目的で開発されてきた。データ量、データの多様性の増加による限界を超えて課題解決をするためには、情報科学者と情報科学研究者の綿密なコミュニケーションがますます必要である。システム開発と課題解決はハードウェア・ソフトウェアの進展と利用可能なデータや知見の投入が相互に組み合わせられてこそ生きる。研究者の研究成果を地球環境が抱える問題解決に利用するためにも地球科学研究者と情報科学研究者の協働は重要である。研究分野の異なる研究者が同じ課題解決に向けてお互いの研究分野についての情報共有をすることは重要である。今までのDIAS開発では、他分野が保有しているデータの利用から類似点や相違点を確認し、自分分野研究への利用を模索することが多かった。さらにある地域での課題解決の成功事例を他地域へ展開する方向での研究の進展がみられた。地球科学研究者にとって研究成果データを他の研究者の利用に供したり、

アーカイブすることはまだまだ負担となっている。研究データ管理計画の作成支援や研究成果データ管理支援をDIASが担うことで研究者の負担軽減をすることに加え、研究者のデータの持続的な維持管理方法を研究者間の情報共有を密にしながら開発することも課題である。

## 謝辞

本研究は、文部科学省「地球観測データ統融合連携研究機構 地球環境情報統融合プログラム(DIAS-P)」の支援を受けたものである。

## 注記

[注1]

<http://www.editoria.u-tokyo.ac.jp/projects/dias/>

[注2] DIAS 利用ワークショップ「システム構成とユーザ管理」

[http://www.editoria.u-tokyo.ac.jp/projects/dias/pdf/dias-workshop/20131224/2\\_UT-nemoto-DIASWS.pdf](http://www.editoria.u-tokyo.ac.jp/projects/dias/pdf/dias-workshop/20131224/2_UT-nemoto-DIASWS.pdf)

[注3] <http://grene.jp/>

[注4] 第4回DIAS-GRENE環境情報統融合フォーラム」発表資料

[http://www.editoria.u-tokyo.ac.jp/projects/dias/meetings.php?q=dias-forum-2013&locale=ja\\_jp](http://www.editoria.u-tokyo.ac.jp/projects/dias/meetings.php?q=dias-forum-2013&locale=ja_jp)

[注5] 第5回DIAS-GRENE環境情報統融合フォーラム」発表資料

[http://www.editoria.u-tokyo.ac.jp/projects/dias/meetings.php?q=dias-forum-2014&locale=ja\\_jp](http://www.editoria.u-tokyo.ac.jp/projects/dias/meetings.php?q=dias-forum-2014&locale=ja_jp)

[注6] <http://www.iugonet.org/>

[注7]

<http://dias-d.tkl.iis.u-tokyo.ac.jp/dia>

s-report/top/

[注8]

[http://www.godac.jamstec.go.jp/catalog/data\\_catalog/](http://www.godac.jamstec.go.jp/catalog/data_catalog/)

[注9]

<http://db.cger.nies.go.jp/JaLTER/metacat/>

[注10] <http://scidbase.nipr.ac.jp/>

[注11]

<http://dias-dss.tkl.iis.u-tokyo.ac.jp/dcd/>

[注12] <http://geolib.gsi.go.jp/>

[注13] <http://www.geoportal.org/>

[注14] OAI-PMH

<http://www.openarchives.org/OAI/openarchivesprotocol.html>

[注15] Open Archives Initiative メタデータ・ハーベスティング・プロトコル

<http://www.nii.ac.jp/irp/archive/translation/oai-pmh2.0/OpenArchivesProtocol.html>

[注16] <http://www.essi-lab.eu/joomla/>

[注17]

[https://www.earthobservations.org/gci\\_gci.shtml](https://www.earthobservations.org/gci_gci.shtml)

[注18]

[http://gcmd.nasa.gov/learn/keyword\\_list.html](http://gcmd.nasa.gov/learn/keyword_list.html)

[注19]

[http://www.mext.go.jp/a\\_menu/kaihatu/kankyou/suishin/detail/1285011.htm](http://www.mext.go.jp/a_menu/kaihatu/kankyou/suishin/detail/1285011.htm)

[注20]

<http://www.unidata.ucar.edu/software/netcdf/>

[注21]

[http://www.ndl.go.jp/jp/aboutus/tech/bigdata\\_sympto.html](http://www.ndl.go.jp/jp/aboutus/tech/bigdata_sympto.html)

[注22] <https://rd-alliance.org/>

[注23] <http://www.bfe-inf.org/>

[注24]

[http://www.soumu.go.jp/menu\\_seisaku/ict\\_seisaku/ictriyu/opendata/opendata01.html](http://www.soumu.go.jp/menu_seisaku/ict_seisaku/ictriyu/opendata/opendata01.html)

[注25] <https://www.icsu-wds.org/>

[注26] <http://www.doi.org/>

[注27]

[http://www.gbif.org/publishingdata/data\\_papers](http://www.gbif.org/publishingdata/data_papers)

[注28] <http://www.emaita.jp/erdp.html>

## 参考文献

- [1] 渋尾欣弘; 生駒栄司; Saavedra, Valeriano Oliver Cristian; Wang, Lei; Koudelova, Petra; 喜連川優; 小池俊雄: 「リアルタイムアンサンブル洪水予測実用化システムの開発. 水工学論文集土木学会水工学委員会編, 第58巻, pp. 397-402, 2014.
- [2] 山本昭夫; Thanda, Nyunt Cho; 喜連川優: 「GCM 日降水量の統計的バイアス補正の並列化」. 情報処理学会第76回全国大会講演論文集1B-2, 東京, 2014.
- [3] 長井正彦; 小野雅史; 柴崎亮介; 坂路和也; 岡田泰征: 「地球観測データ相互流通のためのデータモデルレジストリの開発」. GIS理論と応用, Vol. 20, No. 2, pp. 71-81, 2012.
- [4] 絹谷弘子; 清水敏之; 吉川正俊; 喜連川優: 「DIAS における多分野研究者連携による地球科学データ公開に向けた協働」. 電子情報通信学会信学技報Vol. 110 No. 328, pp. 45-50, 2010.
- [5] Benson, Kevin; Plante, Ray; Auden, Elizabeth; Graham, Matthew; Greene,

- Gretchen; Hill, Martin; Linde, Tony; Morris, Dave; O' Mullane, Wil; Rixon, Guy; Stébé, Aurélien; Andrews, Kona: "IVOA Recommendation: IVOA Registry Interfaces Version 1.0". *ArXiv e-prints*, 2011.
- [6] 木村映善; 村田健史: 「RSS/RDF を利用した太陽地球系物理観測データのメタデータ配信の検討」. 情報処理学会論文誌, Vol. 47, No. 4, pp. 1051-1062, 2006.
- [7] Nativi, Stefano; Bigagli, Lorenzo: "Discovery, Mediation, and Access Services for Earth Observation Data". *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, Vol. 2, No. 4, pp. 233-240, 2009.
- [8] 絹谷弘子; 清水敏之; 吉川正俊; 喜連川優: 「地球観測データに対するメタデータ登録システムの構築-データセットドキュメントメタファを利用したメタデータ登録-」. 情報処理学会第72 回全国大会講演論文集 1K-5, 東京, 2010.
- [9] ISO: "ISO 19115:2003 Geographic information - Metadata", 2003.
- [10] ISO: "ISO/TS 19139:2007 Geographic information - Metadata - XML schema implementation", 2007.
- [11] 石田陽一; 清水敏之; 吉川正俊: 「地球科学データに対するタグと検索語の推薦手法」. 電子情報通信学会第6 回データ工学と情報マネジメントに関するフォーラム/第12 回日本データベース学会年次大会 (DEIM2014) B9-5, 兵庫, 2014.
- [12] 玉川勝徳; 太田哲; 生駒栄司; 絹谷弘子; 大柳美佐; 金内志津; 喜連川優; 松本淳; 小池俊雄: 「DIAS におけるGEOSS/AWCI, AfWCCI, GEWEX/AMY in-situ データアーカイブ」. 日本地球惑星科学連合大会 2013 (JpGU2013) ACG40-03, 幕張, 2013.
- [13] 玉川勝徳; 太田哲; 生駒栄司; 絹谷弘子; 大柳美佐; 松本淳; 喜連川優; 小池俊雄: 「DIAS におけるGEWEX/AMY データアーカイブと公開」. 日本地球惑星科学連合大会 2014 (JpGU2014) ACG38-03, 横浜, 2014.
- [14] 社会調査・データアーカイブ共同利用・協同研究拠点: 「データの管理と共有: 研究者向け最良事例 第3版 (日本語版)」 <http://ssjda.iss.u-tokyo.ac.jp/ukdaguide.pdf> (2014-09-05 参照). 2013.
- [15] 国立国会図書館: 「ビッグデータ時代の図書館の挑戦{研究データの保存と共有}」. 国立国会図書館月報639号 (2014年6月), pp. 12-16, 2014.
- [16] 池内有為: 「研究データ共有時代における図書館の新たな役割: 研究データマネジメントとデータキュレーション」. カレントアウェアネスNo. 319 2014年3月20日 <http://current.ndl.go.jp/ca1818> (2014-09-05 参照), pp. 21-26, 2014.
- [17] 真板英一: 「「研究データ公開」における人材と体制の問題-研究図書館の可能性-」. 日本生態学会誌, Vol. 1, No. 64, pp. 81-86, 2014.

(2014年5月23日受付)  
(2014年9月9日採択)