

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

絹谷 弘子[†] 清水 敏之^{††} 吉川 正俊^{††} 喜連川 優^{†,†††}

[†] 東京大学地球観測データ統融合連携研究機構

^{††} 京都大学大学院情報学研究科

^{†††} 東京大学生産技術研究所

E-mail: [†]{kinutani,kitsure}@tkl.iis.u-tokyo.ac.jp, ^{††}{tshimizu,yoshikawa}@i.kyoto-u.ac.jp

あらまし DIAS「データ統合・解析システム」の開発開始から丸4年が経過した。この間様々な地球科学データが蓄積されている。我々はこれらのデータの科学的・社会的に有用な情報への変換，付加価値の創出を目的としてデータ提供者，研究者とシステム開発者が協議を重ね，本年10月データの公開に至った。本稿では，保有しているデータの把握，メタデータ作成，データ公開に至るまでのシステム開発における課題と連携・協働活動の経過について報告する。
キーワード データ中心科学，分野間連携，地球科学データ

A Multidisciplinary Researchers' Collaboration for Disclosure of Earth Science Data in DIAS

Hiroko KINUTANI[†], Toshiyuki SHIMIZU^{††}, Masatoshi YOSHIKAWA^{††}, and Masaru KITSUREGAWA^{†,†††}

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

^{††} Graduate School of Informatics, Kyoto University

^{†††} Institute of Industrial Science, The University of Tokyo

E-mail: [†]{kinutani,kitsure}@tkl.iis.u-tokyo.ac.jp, ^{††}{tshimizu,yoshikawa}@i.kyoto-u.ac.jp

Abstract Four years have elapsed from the start of the development of DIAS(Data Integration and Analysis System). Various kinds of Earth Science data has been stored in DIAS. And we, data providers, multiple disciplines' researchers and system developers, have discussed many times for converting to useful information for social science, or adding high value to data products, In this October, we have disclosed data to public. We report the current issues and collaborative developing activities for disclosing data in DIAS, understanding of the stored data, creating metadata and developing of search and data download systems.

Key words Data Centric Science, Multidisciplinary Researchers' Collaboration, Earth Science Data

1. はじめに

研究者の研究過程は一般的に仮説の設定，データの収集，加工，解析，可視化による仮説の検証，そして研究成果の論文作成と循環している。データ中心科学の時代となり地球科学データを利用する科学者が扱うデータは，衛星観測データやセンサーデータ，あるいはシミュレーションによって導出されたデータなどデータ量も膨大となっている。研究者が収集したデータを分野を超えて利活用するために我々のプロジェクト「データ統合・解析システム (DIAS : Data Integration and Analysis

System)」^(注1)は2006年から開発を始めた。DIAS開発の目的は，最先端の情報科学技術と地球環境に関わる様々な科学技術の連携によって，地球観測データや数値モデル，社会経済データを効果的に統合し，情報を融合するデータインフラを構築し，地球環境問題を解決に導く知を創造し，公共的利益を創出することである。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する。DIASプロジェクトの研究者は，東京大学に加え海洋研究開発機構，宇宙航空開発機構，京都大

(注1): <http://www.editoria.u-tokyo.ac.jp/dias/>

学，国立情報学研究所，慶應義塾大学，名古屋大学，農業環境技術研究所，農業・食品産業技術総合研究機構に所属し，水循環，気象，海洋，農業，生態系，情報科学など多分野にわたる。DIASに関連した研究者は150名を超える。DIAS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ところである。DIAS開発開始から4年が経過し，DIASが収集し保有するデータ量も増大している。また研究成果のデータプロダクトも増加し，これらを公開することで，高品質なデータを広く社会に提供することになった。DIASは論文による研究成果に加え，広く研究者が作成，利用しているデータの公開によってデータの配布，再利用による価値創出の機会を提供する。データ公開を目的として，我々が行ってきた協働について報告する。2.節ではDIASの現状について，3.節では地球科学データのキュレーションとメタデータについて，4.節ではデータ公開のために行った活動について述べ，5.節でまとめと今後の課題について述べる。

2. DIASの現状

2.1 保有データ

DIASに所属する研究者は，DIASストレージに膨大な量の研究データを収集し，利用している。データの利用者として個々のデータの詳細についてはとても詳しい。また保有しているデータの質や価値についての知識も豊富である。DIASによるデータ公開は，単に収集したデータを外部からアクセスするだけではなく研究者がデータに関して保有している知識も一緒に公開することが可能となる。まずDIASの現状について紹介する。

DIASが保有するデータ量は数百テラバイトである。それらは表1が示すように以下のように分類される。

(1) DIASプロダクト：主にDIASでの研究成果が該当する。研究者が作成したデータセットである。DIASプロジェクトでは毎年応用課題として研究課題を設定し，その研究に必要なデータを収集して分析を行っている。これらのデータは収集データ，中間データとともにDIASストレージに保存されている。プロダクトとして公開されるのは最終成果物となったデータである。

(2) 他プロジェクトプロダクト：DIASでは「科学技術振興調整費 地球環境データ統合・情報融合基盤技術の開発」のような前身プロジェクトで収集しアーカイブしたデータを引継ぎ継続的に保有している。また関連プロジェクト，アジア水循環イニシアチブ(AWCI)^(注2)や統合地球エネルギー水循環観測プロジェクト(CEOP)^(注3)などに，観測データアップロード機能，データ品質検査機能やメタデータ作成機能を提供してデータのプロダクト化を支援している。

(3) DIAS関係者一時データ：DIASでは研究者が保有不可能な超大容量データストレージシステムを提供し，大容量の解析ができる計算空間を提供している。研究に利用する元デー

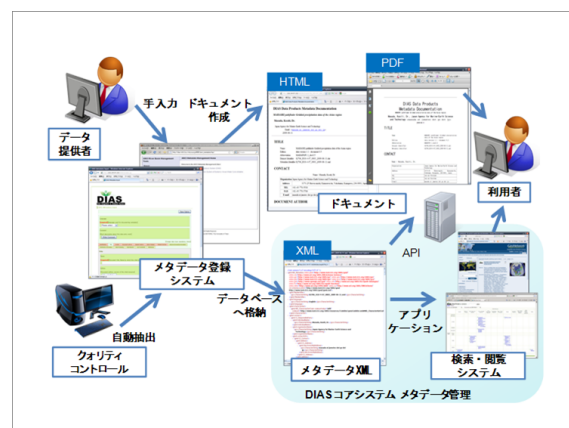


図1 メタデータレジストリフレームワーク

タ，研究用ワークデータ，研究成果データが該当する。これらは，公開の対象にはならない。

(4) ミラーデータ：DIASでは大容量のデータを共有するため，利便性の高いミラーデータを保有している。研究向けに切り出された衛星データ，研究向けモデル出力データなどが該当する。原則としてデータは公開するが，利用制約条項や管理条項などの制約から非公開のものもある。

表1ではDIAS関係者がデータ提供または利用している研究分野ごとに上記分類のデータの利用箇所を「」で示している。この表が示すようにデータの共同利用が盛んであることがわかる。

2.2 データ提供者とデータ利用者

DIASへの主なデータ提供者はDIASに所属する機関とDIAS研究者がデータ提供を要請し，受諾した機関である。また一部については他機関で公開されているデータもある。一方，DIASデータ利用者は，DIAS関係者に限られ，今まではDIAS外の研究者は特別の要請がある場合を除き利用できなかった。またDIAS関係者は昨年度開発した情報共有システムに全員アカウントを保持しているため，このアカウントが流用できる状況にある。従ってデータ公開のためには，DIAS外の利用者がデータにアクセスする機能を開発する必要があった。

2.3 アーカイブデータの管理

大規模ストレージシステムのハードウェア導入とデータ投入を継続して行っているところであり，階層ファイルシステムですべてのファイルに対して格納されているストレージシステムを意識することがない一元管理を行っている。またDIAS研究者は，アーカイブデータに直接アクセスすることができ，アクセスできるデータをストレージ中のディレクトリ単位で許可する方式を採用している。

2.4 DIASメタデータ・ドキュメント

我々は昨年度までに地球観測データに対するメタデータ処理システムを設計し，メタデータレジストリフレームワークについての検討を行い，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システムを構築した[1]~[3]。

当初，DIASが扱うデータセットのメタデータの構造と仕様を作成することに約1年間をかけた。まずDIASプロダクト

(注2)：<http://monsoon.t.u-tokyo.ac.jp/AWCI/index.htm>

(注3)：<http://www.ceop.net/>

DIAS 保有データセット分類		DIAS 関係者がデータを提供または利用している分野								公開/非公開
		農業	生物 多様性	気候	災害	生態系	エネルギー	水	気象	
DIAS プロダクト	DIAS プロジェクトにおける研究成果									公開
他プロジェクト プロダクト	DIAS 前身プロジェクト成果									公開
	DIAS サポートプロジェクト成果									公開
DIAS 関係者 一時データ	研究用素材データ									非公開
	研究用ワークデータ									非公開
	研究結果データ									非公開
ミラーデータ	研究用衛星データ									公開, 非公開
	研究用モデル出力データ									公開, 非公開
	研究用気象データ									非公開

表 1 DIAS が保有するデータの分類

データを対象とし、以前から研究者が習慣として作成しているこれらのプロダクトを説明するドキュメントを対象としてドキュメント中から地理空間情報に対するメタデータ標準である ISO19115 [4] の項目との対応付けを行った。この作業を通して、一般的なメタデータエディタより利便性の高いメタデータ登録システムのプロトタイプを作成し、他のデータに適用可能かどうか、メタデータとドキュメントの試作を続けることになった。DIAS が扱う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、メタデータ形式とメタデータ入出力インタフェースをデータセットごとに設計する方法ではデータ種類分の時間を要することになり現状に対応できないため、試作を繰返し入力内容とメタデータ項目の対応付けを改良した。この作業は、主に情報科学のメタデータ担当研究者と海洋開発機構の地球情報研究センターの担当者中心のメタデータワーキンググループが行った。

さらにメタデータの中で検索を目的として利用されるキーワードをあらかじめ選定しやすくするため、GCMD サイエンスキーワード [5] をはじめ公開されているキーワードリストを収集した。これらの検討結果を反映させる DIAS ドキュメント・メタデータ作成システムを構築した [3]。図 1 が構築したメタデータ登録システムを含むメタデータレジストリフレームワークである。このシステムでは、データを説明するドキュメントの記入方法をとっているため、メタデータの詳細知識を必要とせず、多数の関係者への入力依頼が可能となった。さらに ISO19115 の XML 仕様である ISO19139 形式のメタデータ [6] も同時に作成され、すべての分野のデータを同じ形式のメタデータで記述できるようになった。

3. 地球科学データのキュレーションとメタデータ

データ公開のためには、DIAS ストレージに保有しているデータのキュレーションを考慮する必要がある。キュレーションは情報を収集し、選別し、意味づけを与えて、それを共有することである。キュレーションの対象は、博物館や美術館等の作品や資料に始まり、電子図書館におけるデジタルドキュメントに広がっている。さらにデータ中心科学となり、デジタル化されたすべてのオブジェクトに広がっている。

Jim Gray は第 4 の科学といわれる E サイエンスでは、「研

究論文と科学データは相互参照することが可能となり、研究論文を読んでその根拠となるデータを取得し、再検証を行うこと。あるいは科学データを手がかりとしてそのデータを利用したすべての論文情報を取得することが可能となる」と述べている [7]。また「データ中心科学において、データは自己記述的である必要がある」とも述べている [8]。

情報爆発とデータへのオープンアクセスに対処するためには、適切なデータ発見、データ再利用、統合のために質の高いメタデータが必要となっている。

英国デジタルキュレーションセンター^(注4)は、科学データの長期保存を目的とした研究を行っている。標準化メタデータを記述メタデータ、技術メタデータ、管理メタデータ、利用メタデータと保存メタデータに分類している。またメタデータの機能としてリソース発見と抽出、リソース管理、アーカイブ記録の管理、データ共有と再利用の促進を挙げている [9]。

データは単なる数値や文字、バイナリの並びでしかない。このデータの意味や説明の記述はすべてメタデータによる。そのため、アーカイブデータの管理においてデータと共にメタデータの管理が重要である。先行研究では、DIAS が保有するデータの内容を明らかにすることを主な目的としていた。2.4 節で述べたメタデータレジストリフレームワークによって作成されるメタデータは上記の記述メタデータが中心であり、データのキュレーションを管理するためには、記述メタデータに加え他のメタデータについても検討する必要がある。

さらに DIAS 関係者の一時データを除きアーカイブデータのライフサイクルを考慮する必要がある。データのライフサイクルは (1) データ作成過程 (2) データアーカイブへのデータ投入過程 (3) データ長期保管のための維持管理過程 (4) データアーカイブから 3 次記憶装置へのデータ移行あるいはデータ廃棄過程に分類される。

アーカイブデータの長期保存、あるいはキュレーション管理を考慮したデータモデルとして ISO の標準規格である Open Archival Information System(OAIS) の参照モデル [10] がある。このモデルは「人とシステムの一組織で構成される、指定されたコミュニティのために情報を保存し、それを利用できる

(注4): <http://www.dcc.ac.uk/>

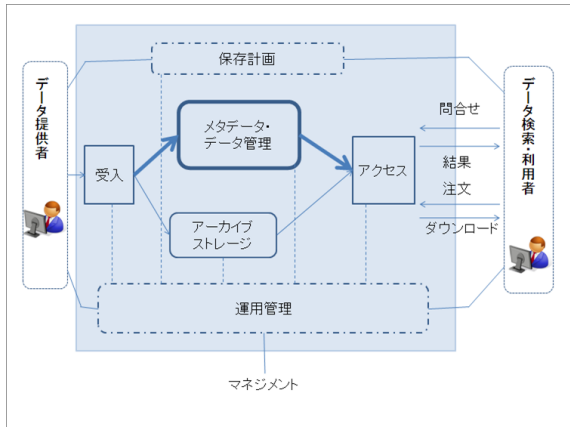


図 2 DIAS アーカイブデータを対象とした OAIIS 参照モデル

ようにする責任を負った「アーカイブ」を対象としたモデルである [11]。OAIIS は電子図書館分野では早くから議論されている。

OAIIS では対象とする情報と関連するメタデータを合わせてパッケージとして扱う。パッケージにはデータ提供者がデータを提出する提出情報パッケージ、アーカイブデータの保管情報パッケージと利用者に提供する配布情報パッケージがあり、各パッケージは内容情報（データオブジェクト、表現情報）と保存記述情報（来歴情報、コンテキスト情報、参照情報と不変性情報）、パッケージ化情報で構成される。OAIIS 参照モデルでは、発見・探索のためのメタデータはパッケージに含まれない。OAIIS 参照モデルは科学データの長期的な管理にも適用されるべきである。そのためには、DIAS ストレージに保有されているデータだけではなく、各パッケージの保存記述情報をどのように管理するか検討する必要がある。

この OAIIS 参照モデルに DIAS アーカイブデータを当てはめて図示したものが図 2 である。DIAS アーカイブデータを外部に公開し、DIAS 関係者以外の研究者や利用者にデータを提供するためには、少なくとも図 2 中の「受入、メタデータ・データ管理、アクセス」部分を確立する必要がある。

DIAS プロジェクトでは、「DIAS が保有するデータには必ずデータを記述するメタデータとドキュメントを作成する」という方針も打ち出された。この方針によって、DIAS 関係者に情報パッケージとしてメタデータが必要であることが共通認識となり、メタデータ作成への協力体制ができた。

4. データ公開プロジェクト

DIAS プロジェクトでは「アーカイブされたデータに関して、データ提供研究者、データ利用研究者らと密に連携をとり、データ公開機能の充実を図る」が本年度の目標となり、そのための運営体制を整えることとなった。メタデータ登録システム構築に携わった関係者を中心としたメタデータワーキンググループが運営にあたることとした。トップマネジメントは、月 1 回程度開催される DIAS テレビ会議に委ねることで、必要に応じて関係者への協力を仰ぐことになった。

4.1 プロジェクトの目標

データ公開にあたり、実現可能性を重視し、データの多様性

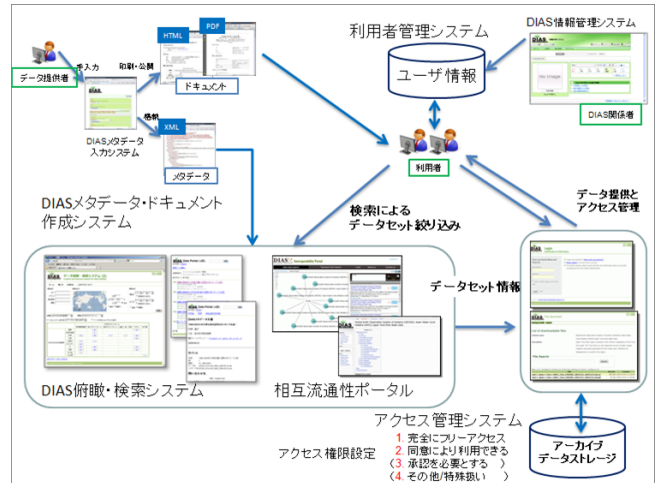


図 3 DIAS データ公開のためのシステム構成

を吸収するシステムを構築することを目標とした。長期的開発計画案を構築してからのプロジェクトであれば、DIAS アーカイブデータとメタデータ管理の設計を行って、データ公開機能の位置づけと開発計画を作成してプロジェクトを開始する。しかしこのプロジェクト開始時点において、現状分析が不足しているため、最低限必要な機能を実現することに焦点を絞った。このプロジェクトで実現する機能は次の 4 点であった。

- DIAS が保有するデータの俯瞰，検索，データダウンロードがシームレスに行えること
- ゲストユーザはユーザ登録をしてデータポリシーに同意してデータダウンロードが行えること
- データ，ユーザごとにデータへのアクセス管理を行うための基盤を作ること
- データ提供者へのデータダウンロードアクセス報告が行えること

これらの目標を達成するために次のシステムを開発することにした。

- (1) ユーザ管理システム：DIAS 関係者とゲストユーザを管理する。
- (2) DIAS 俯瞰・検索システム：DIAS が保有するデータを様々な切り口で分類，検索を多量・専門分野から離れば離れるほど横断検索の意義がでてくるため，検索することであるゆる分野の情報を俯瞰することができるシステムを目指す。
- (3) 相互流通性ポータルとの連携：専門用語の理解と検索を行う。専門用語辞書を中心としたオントロジーを可視化するシステムとの連携によりデータ内容理解の支援を行う。
- (4) アクセス管理システム：公開データの登録，アクセス制約の登録とデータのダウンロードを行う。

図 3 が DIAS データ公開のためのシステムである。DIAS メタデータ・ドキュメント作成システムで作成されたメタデータを利用して DIAS 俯瞰・検索システムと相互流通性ポータルの検索とデータセットに関する説明を行う。対象となるデータセットを特定し，詳細なファイルへのアクセスはアクセス管理システムが行う。利用者ごとに各データセットへのアクセス権

限があるかどうかを確認し、アクセスレベルごとに利用者に利用申請機能を提供する。またユーザ登録が必要な場合は、ユーザ登録システムを利用する。これらの一連の機能が稼動するシステムを構築することになった。

4.2 プロジェクト計画

我々は、データ公開に向けて特別なプロジェクトチームを編成せず、DIAS 会議とメタデータワーキンググループが中心となりプロジェクト計画を推進した。

DIAS データ公開に当たり、短期的に実現可能な開発に絞ったため、DIAS 全体のメタデータ管理の見地からのトップダウンでのシステム開発ではなく、試行錯誤をしながらのボトムアップの開発を行った。開発に当たって留意した点は次のとおりである。

- メタデータとドキュメント作成には手作業での入力を伴い、データ提供者への負担が大きくなる。現実的に一人が入力できるメタデータには限りがあるため、データセットごとにメタデータ作成を行うことにした。サブデータセット単位、ファイル単位のメタデータの作成は見送り、すでに作成済みのデータファイル単位のメタデータは、データと同様公開対象データとし、利用者がダウンロードできるようにした。
- 検索システムでは、データセットを特定する機能を持ち、アクセス管理システムでは特定のデータセットのファイルにアクセスする機能を持つよう役割を分けた。
- データへのアクセスが海外からも容易になるよう、インタフェースとメタデータは日本語と英語両方の表記で作成することにした。
- 公開するデータセットの識別子は、データ公開において重要と認識していたが、識別子の命名規則をあらかじめ作成せず、公開準備が整ったものから順番に割り振ることとした。
- データの利用制約条項、データ参照条項はデータを提供した機関ごとに取り決めがあるが、詳細な条項について確認を行い、データ利用者にこれらの条項の許諾をとるようにした。また、DIAS としての条項を作成し、データ提供者のデータ利用制約条項を優先し、データ参照条項は両者を順守する旨を記載することにした。
- データセットによっては数百万のファイルで構成されるものから数個のファイルで構成されるものまで存在するが、今回はすべて同じ方式とし、データセットごとの対応は行わないことにした。

4.3 プロジェクトの進捗

図 4 が実際にデータ公開までに行った各プロジェクト計画の進捗一覧である。点線で示されている期間は、システム開発者の作業期間である。実線で示されている期間は、DIAS 関係者にとどまらず、データについて熟知した研究者と連携して作業を進めた期間である。8 月から DIAS 関係者に試験公開し、10 月から DIAS 外の利用者にも俯瞰・検索システム、相互流通性ポータルを介しメタデータとデータを公開した。

次に各プロジェクト計画の内容について述べる。

アーカイブデータセットの確認

表 1 に示した DIAS が保有するデータセットを明らかにする

2010年	~3月	4月	5月	6月	7月	8月	9月
アーカイブデータセットの確認	←-----→						
メタデータ作成(日本語, 英語)	←-----→						
俯瞰・検索システム開発→						
ユーザ管理システム開発	←-----→						
データアクセス管理システム開発	←-----→						
公開データセットの確認	←-----→						
DIASデータ使用許諾書, 個人情報利用許諾書作成	←-----→						
各データセット利用制約条項確認	←-----→						
各データセットのカテゴリ確認	←-----→						
各データセットディレクトリ確認	←-----→						
各データセット公開ファイルリスト登録	←-----→						
オントロジー検索システム連携	←-----→						

図 4 各プロジェクト計画の進捗一覧

作業。DIAS 会議で毎回データセット名と DIAS ストレージ内のディレクトリを示しながら確認を行った。さらに何をデータセットとして扱うのが妥当か協議した。メタデータと DIAS ストレージ中にディレクトリが整理されている場合にデータセットに識別子を割り当てた。

日本語と英語のメタデータとドキュメント作成

データセットごとに 2.4 節で示したドキュメント・メタデータ作成システムを利用して日本語と英語のメタデータとドキュメントの作成を行った。メタデータ作成は、データセットの内容をよく把握している人に個別に依頼した。メタデータ中にデータに関する問合せ先、詳細な説明が掲載されている Web サイト、データ利用制約条件、データ参照方法を記入する必要がある。メタデータの完成のために数回の改訂作業も必要であった。この作業に 20 名強が当たった。

俯瞰・検索システムの開発

各データセットのメタデータの内容、メタデータとデータセットとの対応関係を利用してシステムを構築した。当初プロトタイプを公開し、DIAS 会議を通して機能を増強し、インタフェースを洗練した。

ユーザ管理システムの開発

ユーザ認証に CAS(Central Authentication Service) 認証を導入し、DIAS 関係者とゲストユーザを管理するシステムを構築した。

データアクセス管理システムの開発

公開するデータセットについて対応するメタデータ、DIAS ストレージ中のファイルリストを登録し、ユーザ指定ファイルのダウンロードを行うシステムを開発した。ユーザのアクセス権限設定は 4 段階を想定しているが、今回はフリーアクセスとユーザ同意によるアクセスの 2 段階を実装している。

公開データセットの確認

DIAS 会議に公開候補となるデータセットリスト、公開準備の整ったデータセットリストを配布し、個々のデータセットごとにデータ公開のために必要なアクションを確認した。7 月には準備の整ったデータセットをデータアクセス管理システムに登録し、DIAS 関係者が試用してシステムの改良を行った。

DIAS データ使用許諾書，個人情報利用許諾書作成

公開データセットのほとんどが DIAS 以外の機関が提供したデータ，あるいはそのデータを加工したものである。データを公開するためには DIAS としての方針を文書化する必要がある。組織としてデータ公開を行うために，他のデータ公開サイトの例を参考として DIAS 会議で協議した。

各データセット利用制約条項確認

各データセットの利用制約条項は，メタデータ・ドキュメント中に記述する欄がある。しかしメタデータ作成者はデータの利用内容に比べ，再配布，再利用のための条件内容を知らない場合が多かった。そのため再度公開予定のデータセットすべてについて利用制約条項の確認を行う必要があった。

各データセットのカテゴリ確認

DIAS 公開データが利用可能な分野について，データセット全体とのバランスのとれた分類を行うためには，個々のメタデータ作成者がメタデータ・ドキュメント中に記入したキーワードだけでは情報が足りない。そのため公開予定のデータセットリストを配布し，GCMD プラットフォーム，GCMD サイエンスキーワードと GEOSS 社会利益領域におけるカテゴリを再度確認した。

各データセットが保存されているディレクトリの確認

DIAS ストレージ中でのデータセットの配置はドライブの容量に依存し，格納の効率を考慮して配置されている。しかし，公開データセットに対して論理的なディレクトリ構造をとることで，非公開のデータセットと分けることができる。そのためデータセットごとに分かりやすいディレクトリ名に変更しデータセット内のデータの内容が理解しやすいようにした。

各データセットの公開ファイルリストの登録

今回開発したデータアクセス管理システムでは，同一データセット内のファイル一覧を表示してファイルごとにデータをダウンロードする方式をとっている。DIAS ストレージにおいてデータセット内に非公開のファイルの公開を避けるため，ディレクトリ内の公開ファイル一覧をリストとし，一時ファイルなどを除外して登録する。現在 173 のデータセットに関してファイルリストを登録している。

相互流通性ポータルとの連携

専門用語辞書を複数保持したオントロジ検索システムにメタデータ・ドキュメント内の専門用語を組み込み，データアクセス管理システムに接続した。

5. まとめと今後の課題

本稿では，DIAS が保有するデータを公開するために行った多分野研究者の協働について述べた。研究者にとってメタデータ作成や利用制約条項確認作業など個々の研究では必要のない作業ではあるが，積極的な協力が得られた。情報共有化によってデータの統合利用が進み新たな研究成果が生まれている。

今回構築したシステムは，分野を超えたデータ発見に重点を置いており，各々のデータセットに依存したデータの利活用支援に関するシステムとは分けた。特定のデータセットを対象とした可視化システム，データ解析システムとの連携は今後の課

題である。

また，アーカイブデータセットの確認，メタデータ作成，公開データセットの確認等の作業が同時進行で行われたため，これらの関連付けをシステム化することができなかった。今後，図 2 で示した DIAS アーカイブデータを対象とした OAIS 参照モデル中の「受入，メタデータ・データ管理，アクセス」部分のデータ，メタデータの流れに即したシステムを確立する必要がある。

さらに，DIAS ストレージに保有しているデータのキュレーション全体の設計，データ利用，アプリケーションの管理も含めた技術メタデータ，データ更新やメタデータ更新を考慮した保存メタデータ，そしてシステム全体の管理のための管理メタデータについての分析，詳細化が今後の課題である。

地球科学データを公開するプロジェクトは多数存在する。データ公開機関へのメタデータ提供も今後の検討課題である。

謝辞 本研究は，文部科学省による委託研究費「データ統合・解析システム」からの支援を受けて行われた。

また DIAS データ公開にあたり海洋研究開発機構 福田和代氏，市野美夏氏，川本温子氏，赤坂 郁美氏，東京大学地球観測データ統合連携研究機構 小野雅史氏を始め DIAS 関係者の皆様にご協力いただいた。ここに記して謝意を表す。

文 献

- [1] 絹谷，生駒，高橋，吉川，喜連川：“地球観測データに対するメタデータ処理システムの設計”，電子情報通信学会第 19 回データ工学ワークショップ，第 6 回日本データベース学会年次大会 (DEWS2008) 論文集 (2008). C9-6.
- [2] 立床，高橋，絹谷，吉川：“メタデータレジストリフレームワークの構築”，電子情報通信学会 第 13 回 Web インテリジェンスとインタラクション研究会 (2008).
- [3] 絹谷，清水，吉川，喜連川：“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 回全国大会予稿集 1K-5 (2010).
- [4] ISO: “ISO 19115:2003 Geographic information – Metadata” (2003).
- [5] L. G. M. Olsen, K. Shein, J. Scialdone, R. Vogel, S. Leicester, H. Weir, S. Ritz, T. Stevens, M. Meaux, C. Solomon, R. Bilodeau, M. Holland, T. Northcutt and R. A. Restrepo: “NASA/Global Change Master Directory (GCMD) Earth Science Keywords”, version 6.0.0.0.0 edition (2007).
- [6] ISO: “ISO/TS 19139:2007 Geographic information – Metadata - XML schema implementation” (2007).
- [7] T. Hey, S. Tansley and K. Tolle Eds.: “The Fourth Paradigm: Data-Intensive Scientific Discovery”, Microsoft Research, Redmond, Washington (2009).
- [8] J. Gray, D. T. Liu, M. Nieto-Santesteban, A. Szalay, D. J. DeWitt and G. Heber: “Scientific data management in the coming decade”, SIGMOD Record., **34**, pp. 34-41 (2005).
- [9] M. Day: “Dcc Digital Curation Manual”, <http://www.dcc.ac.uk/sites/default/files/documents/resource/curation-manual/chapters/metadata/metadata.pdf> (2005).
- [10] ISO: “ISO 14721:2003 Space data and information transfer systems – Open Archival Information System – Reference model” (2003).
- [11] 栗山：“長期保存型電子図書館と OAIS 参照モデル”，筑波大学・図書館情報大学統合記念 公開シンポジウム電子図書館シンポジウム報文集，<http://www.tulips.tsukuba.ac.jp/pub/dlsympo/hobunshu/> (2003).