

「地球観測データ及びメタデータのアーカイブと連携利用の促進」
に関する今後の対応（案）

平成 30 年 8 月 8 日

地球観測推進部会

1. 目的

「我が国における地球観測の実施計画」に登録されている事業で収集されているデータの内容を国内外で広く共有し、これらのデータへのアクセスビリティを高め、分野を超えた統合的利用を推進する。

2. 背景及び現状

「地球観測の推進戦略（平成 16 年 12 月 27 日 総合科学技術会議）」では、「関係府省・機関の特徴や強みを活かしながら、データ収集から情報提供にいたる段階が適切に統合された地球観測システムの構築に向けて、連携・協調する必要がある。」としている。

また、「今後 10 年の我が国の地球観測の実施方針（平成 27 年 8 月 25 日 地球観測推進部会）」では、「課題解決型の地球観測の推進には、観測・予測データの体系的な収集、合理的な管理、データの統合・標準化や情報の融合が重要である。」、「我が国では、「データ統合・解析システム（D I A S）」を世界に先駆けて開発してきたことに鑑み、今後は D I A S のような永続性・堅ろう性のあるシステムを中核とした地球環境情報プラットフォームを国として整備し、それを十二分に活用していくことが必要である。」としている。

第 4 期地球観測推進部会第 5 回会合（平成 24 年 5 月 30 日 別添参照）において、地球観測データ及びメタデータのアーカイブと連携利用の促進に関する議論が行われた後、関係府省・機関と D I A S 開発関係者の協力により、相当数のメタデータが D I A S に登録されるとともに、フォーマットの変換機能の開発によって多数の既存のメタデータが国際的にも閲覧可能になり、近年のアクセス数が増加傾向にある。また、メタデータ登録の講習会を継続的に開催することによって、生物多様性の分野等におけるメタデータ登録が加速されるなど、データ利用の利便性が高まってきている。

しかしながら、「我が国における地球観測の実施計画」に登録されている事業で収集される全てのデータに対するメタデータが、D I A S に登録されるには至っていない。

3. 今後の対応

平成 30 年度「我が国における地球観測の実施計画」に登録されている事業で収集される全ての地球観測データに対するメタデータを D I A S に登録することによって、データへのアクセスビリティを高め、分野を超えた統合的利用を加速する。

【参考】 D I A S に登録された地球観測データ（メタデータ）へのアクセス数

	H27	H28	H29
アクセス数	16,107	20,637	36,781

地球観測データおよびメタデータのアーカイブと連携利用の促進

目的：

わが国の地球観測の実施計画に登録されている事業で収集されているデータの内容を、国内外で広く共有し、これらのデータへのアクセスビリティを高め、分野を超えた統合的利用を推進する。

対象：

地球観測の実施計画に登録されている事業で収集されるデータ

目標：

- 第1段階：ドキュメントメタデータ*¹の登録と検索サービスの提供
- 第2段階：データおよびデータプロダクトへのアクセスの実現
- 第3段階：データ統融合機能の提供

基本方針：

第1段階は、わが国の地球観測の実施計画を取りまとめる地球観測推進部会のリーダシップで実施され、DIAS はその実施を支援する。一方、各省庁・研究コミュニティ共同で第3段階による科学的理解の深化と公共的利益の発現を推進し、メリットを共有することにより、第1段階、第2段階が持続的、発展的に実施されることを目指す。

方法：

1. ドキュメントメタデータの登録と検索サービスの提供

A: データセンター機能(メタデータ整備、オンライン検索)を有さない場合

方針：DIAS で開発したドキュメントメタデータ登録システムを利用して、ドキュメントメタデータアーカイブし、DIAS による検索サービスを提供する。

- 1) 合同説明会 (DIAS-データ提供者)
- 2) データ内容、ドキュメントメタデータ利用のヒアリング (DIAS-データ提供者、グルーピングを考慮)
- 3) データセットドキュメント共通項目とそのメタデータ要素について ISO19115 に基づいて最低限必要な項目案を作成 (DIAS)
- 4) 必要な項目案の確認 (データ提供者)
- 5) ドキュメントメタデータ登録システムの高度化・改良/試験利用(DIAS/データ提供者)
- 6) ドキュメントメタデータ入力 (データ提供者)

- 7) DIAS 検索システム検索機能（オントロジー検索を含む）による検索機能の高度化・改良／試験利用(DIAS／データ提供者)
- 8) 検索サービスの公開(DIAS)

B: データセンター機能(メタデータ整備、オンライン検索)を有する場合

方針：各データセンターのメタデータ構造を踏まえて、DIAS よりメタデータをクロール、アーカイブする機能（ブローカ機能）を開発して、メタデータをアーカイブし、DIAS による検索サービスを提供する。

- 1) 合同説明会（DIAS－データ提供グループ）
- 2) メタデータ構造のヒアリング（DIAS－データ提供者）
- 3) メタデータのクローリングと DIAS でのアーカイブ(DIAS)
- 4) DIAS 検索システム検索機能（オントロジー検索を含む）による検索機能の高度化・改良／試験利用(DIAS／データ提供者)
- 5) 検索サービスの公開(DIAS)

2. データおよびデータプロダクツへのアクセスの実現

A: データセンター機能(インターネットによるデータへのアクセス)を有さない場合

方針：データ公開を希望する場合は、DIAS で開発したデータ登録、品質管理、メタデータ登録システムを利用して、データおよびメタデータアーカイブし、DIAS による検索・データ利用サービスを提供する。なお、この段階は、研究開発のプロトタイプとして実施する

- 1) DIAS データ登録、品質管理、メタデータ登録システム講習会（DIAS－データ提供者、グルーピングを考慮）
- 2) データアーカイブ（データ提供者）
- 3) DIAS 検索システム検索機能（オントロジー検索を含む）による検索機能及びデータ取得機能の高度化・改良／試験利用(DIAS／データ提供者)
- 4) 検索、データ取得サービスの公開(DIAS)

B: データセンター機能(インターネットによるデータへのアクセス)を有する場合

方針：DIAS と各データセンターの協議に基づき、DIAS でデータアーカイブする場合は、そのデータ構造を踏まえて、DIAS よりデータをクロール、アーカイブし、DIAS による検索・利用サービスを提供する。DIAS でデータアーカイブしない場合は、DIAS とデータセンター間での連携運用を協議し、連携利用のためのシステムを開発し、相互のユーザが相手側のデータを利用できる環境を整備する。なお、この段階は、研究開発のプロトタイプとして実施する。

- 1) データ構造のヒアリング（DIAS－データ提供者）

- 2) DIAS でアーカイブする場合：データクローリング (DIAS)
DIAS でアーカイブしない場合：データセンター間連携運用の協議と、連携利用システムの構築 (DIAS-データ提供者)
- 3) DIAS 検索システム検索機能 (オントロジー検索を含む) による検索機能及びデータ取得機能の高度化・改良/試験利用(DIAS/データ提供者)
- 4) 検索、データ取得サービスの公開(DIAS)

3. データ統融合機能の提供

方針：DIAS 上にアーカイブされているデータおよび各データセンターと連携利用できるデータを、DIAS 上で統融合することによって、科学的理解の深化と公共的利益の創出を目指すワークベンチのプロトタイプをプロジェクト的に各省庁、研究コミュニティ (GRENE 環境情報や RECCA などとの協力) と開発する。

スケジュール：

1. ドキュメントメタデータの登録と検索サービスの提供

2012年度	データセンター機能なし	データセンター機能あり
4月	事前準備	事前準備
5月	合同説明会・グルーピング	合同説明会・グルーピング
6月	ヒアリング	ヒアリング
7月	ヒアリング、項目案を作成と確認	ヒアリング
8月	システムの高度化・改良	クローリングとアーカイブ
9月	システムの試験利用・改良	クローリングとアーカイブ
10月	ドキュメントメタデータ入力	クローリングとアーカイブ
12月	ドキュメントメタデータ入力	クローリングとアーカイブ
1月	検索機能の高度化・試験利用・改良	検索機能の高度化・試験利用・改良
2月	検索機能の高度化・試験利用・改良	検索機能の高度化・試験利用・改良
3月	検索サービスの公開	検索サービスの公開

2. データおよびデータプロダクトへのアクセスの実現

2012年度は地球環境情報統融合プログラム参加機関を中心に試行、2013年度より独法等研究機関を中心に研究開発を行った上で、現業機関との連携運用を目指す。

3. データ統融合機能の提供：

地球環境情報統融合プログラムにおけるワークベンチのプロトタイプ、GRENE 環境情報事業として、2011年度より実施中。

*1 ドキュメントメタデータとは

「良質なメタデータがほしい」というユーザニーズに対して、「メタデータは馴染みがなく、入力が大変だが、データセットのドキュメントは作成する」という一般的なデータ提供者の姿勢に鑑み、各分野でつくられるドキュメント中の章立て、内容を分析し、メタデータ項目との対応付けを行い、ドキュメントに含まれているメタデータの収集し、ISO19115（地理空間メタデータ標準）の項目との対応づけや検索に役立つキーワードを付加したもの。下記の利点、欠点を有する。

利点：

- 1) 分散管理されたシステム（クリアリングハウス）での利用が想定されているため、複数のシステムとの連携などが可能。
- 2) データの主題（what）・データが対象とする場所（where）・データの作成日（when）・データの作成者（who）といった基本的な情報をおさえている。
- 3) 標準化された様式のため、その知識がある人間にとっては理解が容易。つまり、理解の相互流通性が確保される。
- 4) 実際にいくつかのアプリケーションで運用されているため、アプリケーションレベルの相互流通性が（原理的には）可能。

欠点：

- 1) 個々の分野やデータフォーマット等で特徴的な内容（例えば、衛星では FOV などのセンサの詳細な仕様、流域ではモデルの情報、気象・海洋では詳細な時間特性など）を記述するには向いていない。
- 2) つまり、個別のデータセットに注目して、詳細な解析を行う等の用途で使うには不向き。

ドキュメントメタデータ入力項目の例

1. TITLE
2. CONTACT
3. DOCUMENT AUTHOR
4. DATASET CREATOR
5. DATE OF THIS DOCUMENT
6. DATE OF DATASET
7. DATASET OVERVIEW
 - 1) Abstract
 - 2) Temporal Extent
 - 3) Begin Date-End Date
 - 4) Temporal Characteristics
 - 5) Geographic Bounding Box

- 6) Grid
 - 7) Geographic Description
 - 8) Keywords
 - 9) Online Resource
 - 10) Data Environmental Information
 - 11) Distribution Information
 - 12) Distribution Format
8. DATA PROCESSING
- 1) General Explanation of the data producer's knowledge about the lineage of a dataset
 - 2) Data Processing
 - 3) Data Source
 - 4) Data Source Citation Name
 - 5) Description of derived parameters and processing techniques used
9. DATA REMARKS
10. USE CONSTRAINTS
- 1) [Data Policy]
 - 2) [Reference Requirements]
11. REFERENCES