

DIASにおける取組み

絹谷弘子

東京大学 地球観測データ統融合連携研究機構

2015/07/03

ジャパンリンクセンター 研究データへのDOI登録
実験プロジェクト 中間報告会

概要

- ❖ DIASとJaLC 研究データDOI登録実験
- ❖ DIASにおけるDOI付与
- ❖ JaLCプロジェクトへの期待

データ統合解析システム(DIAS)

気候変動予測モデル
気候データベース構築

学術・研究界との連携
による科学知の創出

- ▶ 研究コミュニティ
- ▶ 研究開発機関
- ▶ 国際衛星観測機関
JAXA NASA
- ▶ 国際共同研究
GEO



- ▶ 衛星観測データ
- ▶ 地球温暖化データ
- ▶ 地上観測データ
- ▶ 大気観測データ
- ▶ 海洋データ

- ◀ 社会・経済統計
- ◀ デモグラフィ
- ◀ 健康・公衆衛生情報
- ◀ 地理情報

- ◀ 仮説検証支援
- ◀ 国際連携支援
- ◀ 遠隔協働支援
- ◀ 大規模可視化データ
- ◀ 能力開発のための知識ベース

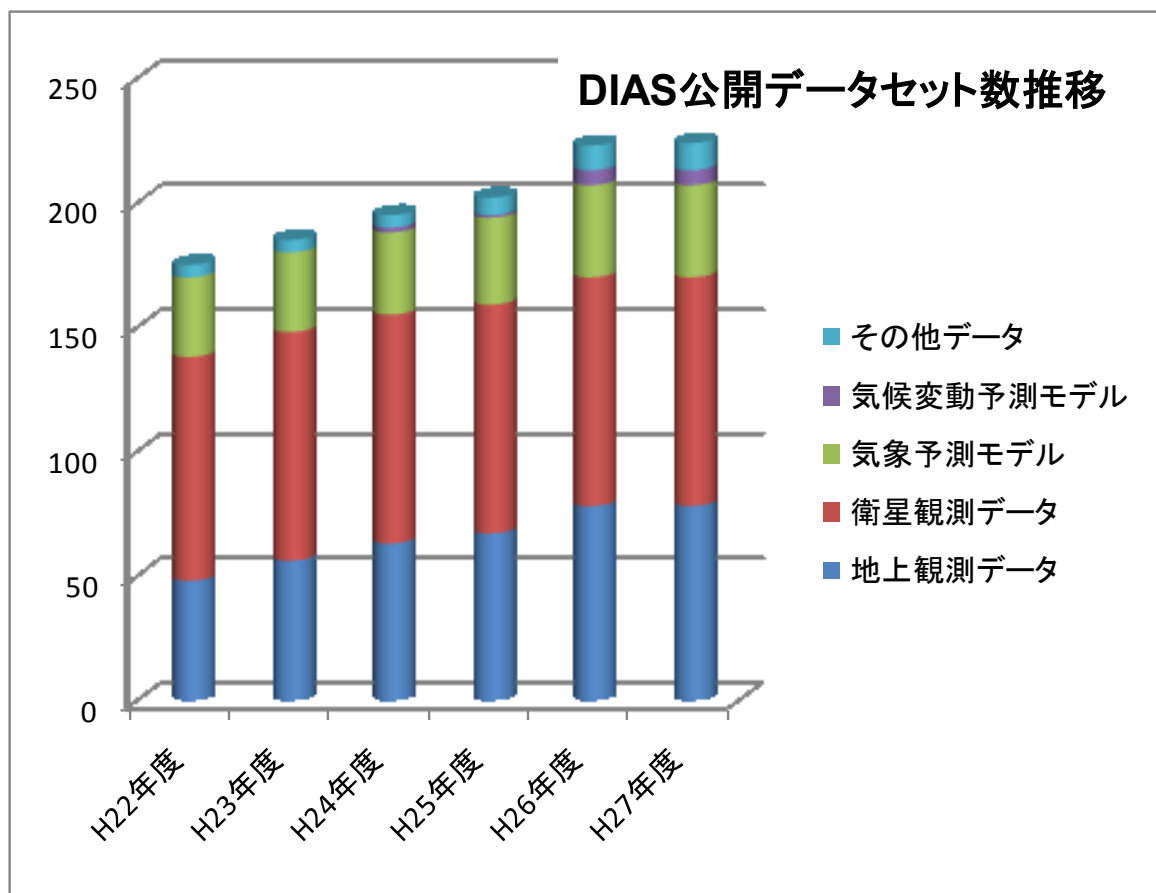
- ▶ 政策決定支援
- ▶ 研究者とのコミュニケーション支援
- ▶ 社会経済シナリオ支援
- ▶ 防災支援
- ▶ 都市計画支援

社会実装による
公共的価値の創出

- 民間企業
TEPCO
- 開発援助機関
jica
- 行政・自治体
長野県
- 政策立案者
都市計画者
- 非政府組織
市民コミュニティ

登録実験対象データについて

DIASデータ公開データセットが対象



研究領域

- 気象・気候
- 水
- 都市
- 防災
- 農業・水産
- 生物多様性
- 健康
- 経済

連携するプロジェクトの研究におけるプライオリティによりデータをアーカイブしてきた。

登録実験

DIASは永続的な組織体制が未確定なためこのプロジェクト期間にはDOI実登録は実施しない

- テスト方針
 - 識別子を付与する粒度・スライス・コレクション・多言語等にかかわる問題を検討
 - 既存のメタデータからの変換やプロジェクト内での運用方式などを検証
- テスト方法
 - 識別子付与規則決定
 - 対象データセット選定
 - DIASメタデータとの対応項目確認
 - Webインタフェース、XMLファイルアップロード、API利用の3つの方法を実験し、いずれも問題なく登録できた
- 実験結果
 - JaLC, DataCiteへの登録方法は理解した
 - DIASメタデータとJaLCメタデータの項目対応も可能(必須項目について確認した)
- 今後の課題
 - 必要なランディングページの設計と現状の改訂方針作成
 - DIASが公開するデータのDOI付与の手続きとワークフロー作成
 - DIASにメタデータがあり、DIAS以外が公開するデータへのDOI付与についての調査

DIASデータ公開プロセス(現状)

- ❖ 公開データは平均10件/年
- ❖ メタデータ作成重点項目
 - ❖ 問合せ先(個人、組織やプロジェクトの継続を重要視)
 - ❖ 収録期間と地理空間範囲、キーワード(シソーラスの利用推進)
 - ❖ データに関するオンライン情報のリンク
 - ❖ データ利用規約とデータ引用規約
 - ❖ これらは、データ提供者の編集に加えDIASメタデータ管理者が確認して編集する
- ❖ メタデータの品質チェック
 - ❖ 問合せ先やオンライン情報を可能な限りリンク切れしないように保守している
 - ❖ 表記ゆれや緯度経度、グリッドの表記等の統一化に努める
- ❖ データ訂正時等にデータ利用者情報をデータ提供者に通知できるように利用記録を管理している

DIASデータ公開プロセス

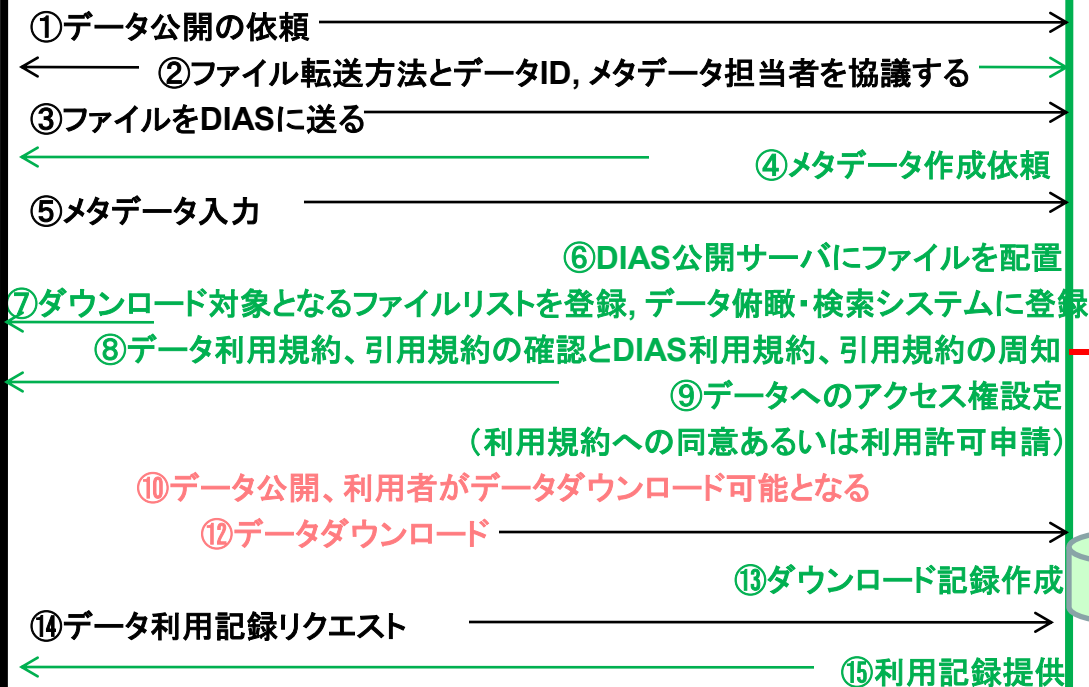
DOI導入

データ提供者

DIAS

Web

JaLC/DataCite



ドキュメント・メタデータ

DOI登録



ランディングページ



ログ



ダウンロードページ

DOIを導入することで、DIAS公開データへのアクセスルートの多様化が見込まれる

DIASオープンサイエンス勉強会の活動

- ❖ JaLCへの参加を含め、オープンサイエンスのトレンドにDIASがどう応えるべきかを検討する会
- ❖ メタデータ, 運用体制設計チームを中心としたDIASメンバーに国立環境研メンバーが加わり、すでに8回開催した
 - ❖ 対話と共創の場をきっかけとして国立環境研が参加
- ❖ 勉強会での議論をベースに、JaLCに対しても不具合の報告やガイドラインの方向性に対して、積極的に提言を行ってきた

ガイドラインへの要望(1)

- ❖ ガイドラインを継続的にアップデートする仕組みを構築して欲しい
 - ❖ データDOIと論文DOIの違いについては未だ明確ではない部分がある
 - ❖ 世界的な状況の変化や研究環境の変化に合わせてガイドラインの見直しが必要となる
 - ❖ 研究データの現場からの意見を集約する場も引き続き必要
- ❖ 多様な分野のデータ提供者が見てわかる「用語集」を追加してほしい
 - ❖ 例:ドメインメタデータ、JaLCメタデータ、データリポジトリ、ランディングページ、データのライフサイクル、ワークフロー

ガイドラインへの要望(2)

- ❖ JaLC会員が主語となる行為に義務・許可・推奨の違いがわかるようにしてほしい
 - ❖ 先行事例に基づいて推奨している部分が大部分ではあるが、shall(強い強制)、will(弱い強制)、may(許可)などの区別が文章でわかるとよい
 - ❖ ランディングページに記載する内容に義務・許可・推奨項目を明確に示してほしい
- ❖ JaLCがDOI登録機関に対して提供するサービスの範囲と責任を明らかにしてほしい

国内の研究データを国内でアーカイブし、世界をリードするためにJaLCのDOI活動に期待したい