

# 地球環境データベース： 地球環境データの利活用に向けて

Global environmental database : supporting/promoting open data

白井 知子\*  
Tomoko SHIRAI\*

国立研究開発法人 国立環境研究所 地球システム領域  
Earth System Division, National Institute for Environmental Studies

## 摘 要

国立環境研究所 (National Institute for Environmental Studies; NIES) は、発足以来、時代の動きに合わせて組織を編成しながら広範囲の環境研究を学際的かつ総合的に進めている。地球環境研究センター (Center for Global Environmental Research; CGER) は、地球環境研究の中核的機関を目指し 1990 年に発足したが、その事業の一環として「地球環境データベース」が立ち上がり、CGER のデータ基盤としての役割を果たしてきた。本稿では、情報通信技術 (ICT) が飛躍的に発展・普及したこの 30 年余、気候変動をはじめとする地球環境研究の必要性が高まる中で、地球環境データベースが、そのデータ収集・管理・公開・利活用促進等のために、どのように変遷・発展してきたかを振り返るとともに、近年のオープンサイエンスの流れに対応した現在の取組、課題や将来展望について述べる。

**キーワード：** オープンサイエンス、研究データ基盤、地球環境研究、データベース、DOI

**Key words :** open science, research data infrastructure, global environmental research, database, DOI,

## 1. はじめに

18 世紀後半に始まった産業革命以来、世界各地で工業化・都市化が進み、大気や水の汚染をはじめ、自然環境に大きな負荷を与えて来た。限られた地域の環境に影響を及ぼすだけの頃は公害と呼ばれ、原因も突き止めやすかったが、徐々に影響範囲が広がり、地球全体の環境を脅かすようになった。気候変動に代表されるような地球環境問題は、様々な要因が複雑に絡み合っているため、その現象解明や予測は容易ではない。データ収集だけを取っても、1 人の研究者、1 つの研究グループ、1 つの組織、1 つの国だけでは、とても追い付かないため、世界中が協力して取り組む必要がある。その意味で、地球環境問題に取り組む上で、世界中のデータをお互いに使い合えるようにする重要性は高い。このように、地球規模で起きる環境変動の現状把握、現象解明、将来予測や影響評価などに資するデータを本稿では「地球環境データ」と呼ぶ。地球環境データの利用価値が高まる中、データの公開・相互利用を促進するためには、データ公開基盤・デー

タ管理基盤の整備が必要である。国立環境研究所 (NIES) 地球環境研究センター (CGER) では、Web ベースのデータ公開基盤として、地球環境データベース (Global Environmental Database, 以下, GED) を構築、2014 年から公開している。本稿では、今から約 30 年前の CGER 発足時からあった、広義の「地球環境データベース」構想を元に、世界的なデータ流通の変遷の中で形を変えながら発展してきた「地球環境データベース」のこれまでを振り返るとともに、近年のオープンサイエンスに向けた動きや現在進行形の取組、そして今後の展望について紹介する。

## 2. 地球環境データベースのこれまで

まずは CGER 発足前の国立公害研究所時代からの情報発信の取組を振り返る。1990 年の改組以前は環境情報部・統合解析部に研究者が配属され研究活動も行われていたが、本稿では実務部分に特化して記載した。過去の組織や活動内容については、主に国立環境研究所年報 (国立環境研究所, 1976–2021) を参

受付：2021 年 11 月 5 日、受理：2021 年 12 月 22 日

\* 〒 305-8506 茨城県つくば市小野川 16-2, E-mail : tshirai@nies.go.jp

考にした。

## 2.1 公害研時代

NIESは、1973年に国立公害研究所として発足した。研究所の設立以来、図書資料等の研究情報、数値情報(主に大気質・水質)、文献情報、情報源情報等の収集、それらの情報に基づく環境データベースの構築と運用、また、1975年に指定された国連環境計画(UNEP)のINFOTERRA(International Referral System for Sources of Environmental Information: 国際環境情報源照会システム)の代表機関(ナショナルフォーカルポイント)としての業務を環境情報部が担当していた(藤原 1977; 土屋ほか, 1980; 土屋, 1982; 土屋, 1983)。INFOTERRAは、環境情報の内容そのものではなく、情報源に関する情報(どこでどのような情報が手に入るかという情報)を案内するために作られたデータベースで、1980年時点で66か国から約7,700の情報源が登録され、日本からも政府行政機関・国立試験研究機関・特殊法人・公益法人など計169の情報源が登録されていた(土屋ほか, 1980)。INFOTERRAによる情報源照会は、いったん情報源を調べて、それから情報の提供を依頼するという2段階の手順を踏む上、通信手段も郵便が主であったため、実際の情報入手までに相当な手間と時間がかかった(春山・細山, 1987)。当時は、データ入手の手前の段階で、欲しい情報の所在や入手方法を調べるのも容易ではなかったことが推察できる。

## 2.2 環境研への改組：CGER 立ち上げ期

1990年7月、NIESは大幅な組織改編を行い、名称も「国立環境研究所」に改めた。この際、研究組織として従来の専門分野別の部室別のみの体制を改め、社会ニーズに対応したプロジェクト研究を行う総合研究部門(地球環境研究グループ及び地域環境研究グループ)と、シーズ創出や総合研究部門の支援のための研究を行う基盤研究部門(6部)が設置された。CGERは、地球的規模の環境変化とその影響を調べる地球環境研究グループと連携しながら地球環境研究を促進する組織として、同年10月に発足した。

CGERの設置目的は、地球環境の保全に関し、国際的な協力のもと学際的、省際的な地球環境研究の総合化を図るとともに、データベース等の研究支援体制を充実させ、また、地球環境の長期的モニタリングを行うなど、環境問題の解明と解決のために、幅広く貢献していくことであった。「地球環境データベースの整備・提供」は、CGER発足当時から「地球環境研究の支援」業務の一環として、「地球規模で生じる環境変動に関する研究及び環境行政施策遂行に必要とされる地球環境問題に係る独自性と国際性を備えた各種データを収集、蓄積し、それぞれのニーズに応じてその提供を図る」と定義されていた(地球環境研究センター, 1990)。

立ち上げ期の主な活動は以下のようであった。

### ①地球環境データベースシステムの構築

#### ②個別分野のデータセット整備

#### ③地球環境研究を行うための基礎データの収集

①は、地球環境に関するデータについて、数値データ、データ概要、データ情報源、データ作成機関に関する情報をリンクさせて取り出すことが可能で、データの検索、表示(メッシュ、国単位)、情報源情報の検索が系統的に行え、また、データ管理、ユーザ管理といったシステム運用の機能ももたせたデータベースを目指していた。②については次節で詳しく紹介する。③としては、国連の人口、貿易データ等国際機関の地球環境関連データ及び地球環境研究を行うための基礎データの収集を行い、情報源情報については、「国際研究計画・機関情報」としてまとめ、出版した(地球環境研究センター, 1992)。

また、データベース整備の一環として、UNEPの環境情報利用システムである地球資源情報データベース(GRID)との提携を図るべく交渉し、1991年5月にCGERはGRIDの第8番目のセンター(GRID-つくば)に指名された(原沢, 1993)。組織体制としては、CGERの総括研究管理官、研究管理官、業務係長がそれぞれGRID-つくばのディレクター、施設担当、事務担当を務め、研究者が対外窓口及び研究協力員として参画、データ提供はNIESの環境情報センターの協力を得ていた。GRIDはデータの提供だけでなく、オリジナルデータの作成も奨励しており、CGERでも衛星データを元にしたアジア植生指数や二酸化炭素排出量予測値、全球土壌水分データなどをGRIDネットワークに登録、提供した(地球環境研究センター, 1994)。

当初のハードウェアは、1台のワークステーションとNIESの大型計算機の磁気テープ装置のみであったが、1992年に所内のパソコン、ミニコン、1991年度末に導入されたスーパーコンピュータ等とネットワーク接続して、データの蓄積・管理・提供及び画像解析技術の開発が行われていた。1993年度にはGISを搭載したGRID専用ワークステーションが導入された(地球環境研究センター, 1995)。

## 2.3 オリジナルデータセットの作成・公開

1990年代はインターネットが急速に普及した。NIESでも1992年度にスーパーコンピュータシステムが新規導入されたことに伴い、ローカルエリアネットワークとして、FDDI(Fiber Distributed Data Interface)を利用した国立環境研究所ネットワーク(NIESNET)が構築された。1996年3月末にはWWWサーバによる国立環境研究所ホームページの運用が開始され、1997年1月に起きた日本海におけるナホトカ号難破による油汚染では、事故の翌月には調査結果を公開するなど、ホームページによる情報発信の速報性は事故・災害時にも威力を発揮した(国立環境研究所, 1998)。インターネットにより情報収集

が容易になったため、INFOTERRAのような情報源情報データベースを介さなくても直接ネットワークを通じてデータそのものにアクセスできるようになっていった。ただし、回線はまだ遅く、2000年以前は、フロッピーやCD-ROMを利用したデータ提供が中心であった。

地球温暖化に関しては、解決に向けた国際的な機運も高まり、気候変動枠組み条約(UNFCCC)が1992年に締結され、1997年に京都で開かれた第3回締約国会議(COP3)で温室効果ガスの排出削減目標が設定される(京都議定書)など、国民の認知も広がった。NIESにおける地球規模の環境問題に関する研究活動が活発化したこの時期には、各分野の研

究者を中心にオリジナルデータベース(温室効果ガス排出シナリオ、排出インベントリ、炭素吸収源、森林生態系、温暖化影響・対策、海面上昇、土地利用、マテリアルフロー等)の構築が進み、CGERではこのような地球温暖化の科学的根拠を支えるデータ収集・提供を積極的に進めた。表1にこの頃までにCGERから提供していたデータセット等のリストを示す。

#### 2.4 法人化：地球環境モニタリングとの連携強化

2001年度よりNIESは独立行政法人化し、CGERの活動は「知的研究基盤」として位置づけられることとなった。独立行政法人化に際して策定された中長期計画に沿って事業体制を強化するために、従来

表1 CGER 提供データベース一覧(2000年当時).

地球環境研究所センター(2000)より引用.

\*データベースの名称の後の( )は、提供形式を示す。( )のないものは印刷物.

	データベース名称	CGER No.
データセット		
1	東アジア海域海洋汚染モニタリングデータ(CD-ROM)	CGER-D007(CD)-'95 CGER-D012(CD)-'97 CGER-D021(CD)-'99
2	Nimbus7-CZCS月別複合画像による北西太平洋海域における植物プランクトン分布の衛星画像時系列データベース(CD-ROM)	CGER-D015(CD)-'97
3	産業連関表による二酸化炭素排出原単位(FD5枚添付)	CGER-D016-'97
4	IGAC/APARE/PEACAMPOT 航空機・地上観測データ('91~'95集成版)(CD-ROM)	CGER-D018(CD)-'97
5	全国第1次酸性雨調査データセット(FD2枚)	-
6	東京23区の人工排熱(エネルギー消費)時空間分布(CD-ROM)	CGER-D019(CD)-'97
7	東アジア植生指数月別モザイク図1996年(CD-ROM)	CGER-D020(CD)-'98
8	東アジア植生指数月別モザイク図1997年(CD-ROM)	CGER-D023(CD)-2000
9	Data of IGAC/APARE/PEACAMPOT Aircraft and Ground-based Observations ('96-'98 collective Volume) (CD-ROM)	CGER-D026(CD)-2000
データブック		
1	海面上昇データブック2000 (Data Book of Sea-Level Rise 2000)	CGER-D025-2000
2	砂漠化/土地荒廃データブック (Data Book of Desertification/Land Degradation)	CGER-D013-'97
3	国際研究計画・機関情報II	CGER-D017-'97
4	Data Book of Information about International Research Institutions/Programmes	CGER-D024-2000
5	'94 IGAC/APARE/PEACAMPOT 航空機・地上観測データ集	CGER-D010-'96
6	'95 IGAC/APARE/PEACAMPOT 航空機・地上観測データ集	CGER-D011-'96
7	マテリアルフローデータブック～日本を取りまく世界の資源のフロー～(Material Flow Data Book-World Resource Flows around Japan-)	CGER-D022-'99
研究支援ソフトウェア		
1	Acclaim (環境資源国際収支地図表示システム) (FD 4枚)	-
2	RDF (環境資源データフロー図表示システム) (FD 1枚)	-
GRIDデータ解説書など		
1	PGRID (パソコン版GRIDデータ表示プログラム: FD 3枚)	-
2	GRID全球データセットユーザズガイド	CGER-D004-'94
3	GRID DATA BOOK	CGER-D006-'94
4	GRID-Tsukuba (パンフレット)	CGER-D008-'95
インターネットによる情報提供		
1	NOAA/AVHRR画像検索システム	-
2	IPCC Scenario Database	-

の業務分担にとらわれず、分野横断的な体制で事業を推進することとなった。2006年度からの第二期中期計画では大幅な組織再編と研究の重点化のため、全ての分野でセンター化され、各センターが中核研究プログラムを担うこととなった。

CGERでは、モニタリング・データベースの位置づけを明確にするため、大気海洋モニタリング推進室、陸域モニタリング推進室と共に地球環境データベース推進室(以下、DB室)の3つの推進室が設置された。各推進室には研究者が配属され、研究管理官を中心に運営されていた第一期までの体制と比べ、地球温暖化研究プログラムとの連携を強めた運用体制となった。それに伴い、地球環境データベース事業は、社会経済的な地球環境研究の成果のフォローアップに比重が置かれていた初期の頃と比べ、地球環境モニタリング事業と連携して、自然科学的な地球環境データのとりまとめ(データベース化)とデータ提供の推進、研究支援ツールの整備等に力を入れるようになった。

この時期、地球環境データベース事業では、情報関係のセキュリティ強化、個人情報保護の観点配慮して、地球環境研究センター基幹サーバ/情報提供サーバ群の更新を実施した。これにより、安全性・安定性を高めた環境から多様かつ多量の情報を所内外に提供する枠組みが整備された。DB室が管理するサーバ室が確保され、システム開発や予備システムも含めると10台以上のサーバや大容量ストレージから構成されるシステムを運用するようになった。また複数のデータベースサーバからの情報発信を <http://db.cger.nies.go.jp/xxx>(xxxは各事業の短縮名等)という統一されたURLから行うことが出来るようになった。体制・システム整備が進み、DB室は、オープンサイエンスの流れに乗り、研究データの発信等に力を入れられるようになった。具体的な取組について、次章で紹介する。

### 3. 地球環境データベースの提供サービス

多種多様な環境問題の中でも地球環境分野は、地域限定的な環境問題と比べ、全球的なデータを集めるためにデータの相互利用の必要性が高く、また、個人情報や倫理面に配慮すべきデータが少ないこともあり、国際的にもオープンデータが進みやすい性格を持つ。逆に言えば、地球環境データを速やかに公開し、研究コミュニティ等に提供することの価値、メリットは大きい。地球環境データベース事業では、地上ステーション、船舶、航空機、人工衛星など、多様なプラットフォームを用いた、日本最大の温室効果ガスの観測網である地球環境モニタリング事業のデータをはじめとする多様な地球環境データの公開・利活用促進に向けたサービスを提供している。

### 3.1 地球環境データベースの構築

地球環境データベース事業のホームページでは、データセットへのアクセスビリティを向上させるため、包括的検索ができるポータル機能や、デジタルデータ提供機能をもつデータベースの構築を進めてきた。第二期中期計画期間中(2006年~2010年)に行ったホームページの再構築では、コンテンツの内容に応じた分類・メニュー化を行い、トップページからデータ検索を行うことができるよう改良した。

当時、CGERが取得している温室効果ガスの大気中濃度や、海や陸域におけるフラックスといったデータは、次項で紹介する速報値提供を除くと、研究者への個別リクエストを通じて、もしくは、NOAA(National Oceanic and Atmospheric Administration; 米国大気海洋庁)のGlobal Monitoring Laboratoryによる大気観測データベースをはじめ、WDCGG(World Data Centre for Greenhouse Gases)、表層海洋CO<sub>2</sub>データベース(Surface Ocean Carbon dioxide Atlas: SOCAT)、二酸化炭素情報分析センター(Carbon Dioxide Information Analysis Center; CDIAC)、全球のCO<sub>2</sub>フラックス観測ネットワーク(FLUXNET)等の国際的データベースを通して提供されていた。2011年度からは、文部科学省の予算による「地球環境情報統融合プログラム(DIAS-P)」(DIAS II期)に参加し、CGERが取得してきた地上・海上・上空の温室効果ガス観測データやフラックス観測データについて、DIASフォーマットでのメタデータの作成とDIASへの登録を行ったが、DIASから直接のデータセット提供は行っていなかった。

これらの国際的・統合的なデータベースからデータ公開すること自体は有効なのだが、品質管理(Quality Control: QC)を行ったデータを、各データベース用のフォーマットに変換し、各データベース用のメタデータを揃えて提出する研究者の手間と、それが各データベースから公開されるまでの時間の遅れにより、データをタイムリーに公開することは難しかった。そこで、より簡便・迅速にCGERから直接データを公開・提供するためのデータベースとしてGEDを構築し、2014年5月より一般に公開した。

### 3.2 地球環境データベースの特長

GEDのデータ提供サイトの特長として、データ利用者側からは以下の3つが挙げられる。1)フォーマットが統一されたデジタルデータを直接ダウンロードできる、2)クイックプロットツールで簡単にデータを可視化できる、3)観測値を準リアルタイムで見ることができる。一方、データ提供者側からも以下のようなメリットがある。1)データ公開の支援が得られる、2)データにデジタルオブジェクト識別子(DOI)を付与して公開できる、3)アクセス解析によるデータダウンロード状況の把握が可能である。

GEDのグローバルナビゲーションからは、「デー

データベース「速報値」「解析支援」「関連データ」「データ検索」の各メニューが利用できる。「データベース」では、各データ提供者によるオリジナルフォーマットまたはメタ情報付のテキスト形式(NASA Ames フォーマット)での QC 済のデータセットをダウンロードできる。また、データをダウンロードしなくても、自動グラフ作成(クイックプロット)機能によりデータの中身を確認することができる。一部、ダウンロードするのにユーザ登録が必要なデータがあるが、これはユーザの利用目的等を把握したいというデータ提供者からの要望に応じている。これまで地球環境モニタリング事業による観測データを中心に「大気海洋モニタリング」「陸域モニタリング」というカテゴリから公開してきたが、最近、モデルを用いたデータセットの公開が急増しているため、2021年度から「モデルシミュレーション」というカテゴリを追加した。

「速報値」では、CGERの地上ステーションで観測しているCO<sub>2</sub>濃度の毎時アップデートと共に、日平均、月平均、年平均のほか、週、月、年間の時間変化を見ることができる。この他にもUVインデックス、ビタミンD生成・紅斑紫外線量、温暖化影響モニタリング(高山帯)、成層圏局渦予測などのパラメータについて、ほぼリアルタイムでのデータ公開を行っている。速報値取得用のWeb APIも提供しており、EICネット(<https://www.eic.or.jp/>)のHP上ではCO<sub>2</sub>濃度(日平均)とUVインデックス(時間値)が準リアルタイム表示されている。

「解析支援」では、ユーザがデータを解析するのに役立つオンラインアプリケーション等を提供している。中でも大気観測データの解析・検証に使われる、気塊の移動経路、すなわち流跡線(トラジェクトリ)の算出ツールは歴史が長く、広く使われている。1987年にNIESで開発されたトラジェクトリ解析モデル(Hayashida-Amano *et al.*, 1991)を元に改良を重ね、2001年には、研究者が自分のPCにインストールし、入手しやすい気象データを用いてトラジェクトリの算出や気象場の表示ができるツールとして、ソースコードの提供を開始した(Zeng *et al.*, 2010)。この流跡線解析・気象場表示システムMETEX(Meteorological Data Explorer)は、2002年からオンライン計算サービスを開始し、産官学交えた様々な分野から広く利用されるようになった。さらに、ある地点に到達する粒子の集合体や、ある地点から放出された粒子の集合体について、どの場所にとどの程度の時間滞留したかを計算できるラグランジュ粒子散布プログラムFLEXCPPも開発し、オンライン計算ツールを公開するとともに、パッケージの提供も行っている。この計算はフットプリント計算とも呼ばれ、観測地点に到達するまでの大気塊が輸送中に地上フラックスから受けた寄与の見積もりや、ある地点で放出された物質がどのように空間分

布を拡げるか等のシミュレーションに使われている(Zeng *et al.*, 2012)。CGERが開発・提供している解析支援ツールは、DIASからも公開されている。

「関連データ」では、これまでCGERがかかわってきた幅広い分野にわたる研究成果・データをテーマ別に一覧できる。これらの中には温暖化影響や対策、自然・生物・水環境等、NIESの他の分野との協力で作られているコンテンツや、データに関連した研究報告(CGERレポート等)も含まれる。「データ検索」では、サイトに含まれる幅広いデータをプロジェクト横断的・分野横断的に条件絞り込み検索できるほか、観測地点マップからの検索やメタデータ検索も可能である。

GEDの現時点での規模として、サイト全体から提供しているデータ総量は約4.2TBで、そのうちNIES DOIを付与して公開しているデータが3.5TBを占める。(ただし、ストレージ容量の制限から、提供データセットの容量が、研究データと補足資料を合わせて100GBを超える場合は、提供者側でデータ保管サーバを確保してもらっている。)なお、日本長期生態学研究ネットワーク(Japan Long-Term Ecological Research Network; JaLTER)やアジア地域のフラックス観測ネットワークAsiafluxのデータベースへのサーバ提供及び管理・運用支援も行っているが、それらのデータ量は含めていない。また、現在観測中の、温室効果ガス観測技術衛星GOSAT(Greenhouse gases Observing SATellite)及びGOSAT-2については、プロジェクトとしてデータ提供を行っているため、そのデータ量も含めていない。

#### 4. オープンサイエンスに向けた研究データ基盤の整備

近年、研究環境のICT化とともに、世界中で研究データの電子化が急速に進みつつある。その中で、科学研究への新しいアプローチとして、デジタルテクノロジーを用いて、研究成果やデータを共有・公開し、相互協力しながら進めていくオープンサイエンスという概念が広まった。特に2013年のG8科学技術大臣会合で研究データのオープン化が合意されて以降、政策にもオープンサイエンスが取り入れられるようになった。NIESにおいては2016年からの第四期中長期計画からは、DB室は地球環境データ統合解析推進室と名を変え、データ発信にとどまらず、データの管理支援・利活用促進を通じて研究活動を支援する基盤作りを視野に入れ業務を展開している。国際的には、研究データの流通や活用を推進する国際イニシアティブFORCE11(The Future of Research Communications and e-Scholarship 11)によりFAIR原則が提案された(FORCE11, 2016)。FAIRとはFindable(見つけられる)、Accessible(アクセスできる)、Interoperable(相互運用できる)、Re-usable

(再利用できる)の略であり、データの適切な公開方法を表す原則である。GED も、FAIR 原則に沿ったデータ公開を目指し、時代に即したオープンサイエンスに向けた取組を実施している。

#### 4.1 研究データへの DOI 付与

DOI とは、デジタルオブジェクトのアクセスを永続的に保証する(置き場(URL)が変わっても変更後の URL にリンクする)識別子で、1998 年頃から出版業界でオンライン化された学術論文を対象に広まった。研究データへの DOI 付与は遅れて開始され、2014 年 10 月から「研究データへの DOI 登録実験プロジェクト」(武田ら, 2016)がジャパンリンクセンター(JaLC)により実施されるまで国内では付与例はなかった。CGER では 2015 年はじめに、複数のデータ提供者から「研究データに識別子(DOI)を付与したい」という要望があったのをきっかけに上記プロジェクトに参加し、JaLC のメタデータ登録システムのプロトタイプを用いた DOI 登録実験や「研究データへの DOI 登録ガイドライン」(ジャパンリンクセンター運営委員会, 2015)の作成に貢献した。

上記プロジェクトへの参加により、研究データに DOI を付与する仕組みを理解し、実務的な体制整備を進めていった。また、DOI の持続性を保つため、DOI 付与の主体は CGER ではなく、より組織として安定な NIES とするべく、所内において、関連する環境情報部・企画部・理事室の合意を得て、2016 年 6 月、NIES として JaLC の正会員となった。アクセス持続性を重んじるため、ランディングページの作成及び XML ファイルを用いた DOI 登録申請は、環境情報部に依頼した。CGER の作業としては、データ提供者と協力の上、DOI 付与データの選定とフォーマット統一、メタデータの作成支援、ランディングページ案の作成、データ登録用 XML ファイルの作成等を行うワークフローの整備を行なった。NIES として初めて DOI を付与するデータセットとして、波照間・落石の両ステーションで観測された CO<sub>2</sub> 及び CH<sub>4</sub> を選び、2016 年 9 月 1 日に公開した。登録したデータの DOI は、データ DOI 登録のための国際的なコンソーシアムである DataCite に登録され、DataCite を通じて検索できる。

データに DOI を付与することで、データ引用はもちろん、バージョン管理等、公開データの管理もしやすくなるため、研究者からの需要も多い。特に近年は、オープンデータの流れて、論文投稿時に根拠データの公開を要請してくる学術雑誌や出版社が増えているが、その多くで、データへの DOI 付与も条件とされている。GED は論文投稿・公開前の研究者のデータ公開をスムーズに行うことのできるリポジトリとして、直接的に研究活動を支援できるようになったため、これが、GED からのデータ公開のインセンティブにも繋がっている。GED からの DOI 登録は加速度的に増え続けており、2021 年

10 月現在、DOI を付与して公開中の合計 36 データセットのうち半数が最近 2 年間に公開された(表 2)。また、既に公開済の DOI の下で、定期的にデータを追加・更新しバージョンアップを続けているデータセットも増えており、データ公開のアクティビティは確実に高まっている。

#### 4.2 研究データ管理システム(RDMS)の開発

近年、オープンサイエンス推進の潮流、研究不正対策、研究資金の助成機関の要請などから研究データを適正に管理する必要性が高まっている。これを実現するための基盤として、2018 年度に Research Data Management System (RDMS) の設計・開発を開始した。研究データ管理とは、研究活動を通じて収集・生成したデータを、データのライフサイクルに沿って管理する一連の流れを指し、多様な要素が含まれる。一口に RDMS といってもその担う範囲は、システムごとに異なっていることが多い。DB 室が開発中の RDMS である GERDaMS (Global Environmental Research Data Management System) は、研究者、研究チームが効率的に研究データを管理・共有できること、特に、データ公開までの流れをスムーズにするところに主眼を置いている。研究者のデータ公開を支援する中で、データ公開に必要な作業はデータセット生成時から準備できるものが多いこと、支援担当者とのメールやファイルのやりとりよりも、データ提供側に Web アプリケーションベースで作業してもらった方が効率的であることがわかり、その効率化を図った。

GERDaMS の具体的な機能として、ユーザ登録の上、データセットを登録すると、メタデータの作成、利用規約の作成、データのバージョン管理、DOI 付与、データファイルや補足データのアップロード及び各ファイルの公開・非公開設定等が行える。また、各データセットに対し、複数のアクセス可能なユーザを登録でき、メンバーで協働しながらデータ管理を行うことができる。データ・メタデータとも、データの種類や研究のフェーズに応じて公開・非公開を設定できる。特に GED からのデータ公開は、共通の統合データベースを通じて連動させているため、スムーズに行うことができる。DB 室側でも、GERDaMS の管理者機能により、データ提供者やデータセットを一元的に把握できるようになれば、ユーザ数が拡大しても必要な支援を提供しやすくなる。研究公正の観点からも、特に論文のエビデンスデータは共通基盤で安定して管理することが望ましく、今後、研究提案時に作成を求められることが増えるであろう DMP (Data Management Plan) 作成の上でもデータの管理体制を整備しておくことは重要である。GERDaMS は 2021 年現在、プロトタイプ試験運用を行いつつ、機能の実装・改良を進めており、年度内の本格運用開始を目指している。RDMS のリリースと同時に、GED のリニューアル

表2 CGER から DOI を付与して公開したデータセット一覧(2016年9月~2021年11月).  
 背景白が観測値, 背景グレーがモデル計算値. DOI suffix の最初の8桁は初回公開年月日を示す.  
 公開後, データの追加・更新を行ったデータセットについては, これまでのバージョン更新回数を記した.

データセット名	DOI	更新回数
波照間ステーションにおける大気 CO <sub>2</sub> 濃度の連続観測データセット	DOI:10.17595/20160901.001	2
落石岬ステーションにおける大気 CO <sub>2</sub> 濃度の連続観測データセット	DOI:10.17595/20160901.002	0
波照間ステーションにおける大気 CH <sub>4</sub> 濃度の連続観測データセット	DOI:10.17595/20160901.003	2
落石岬ステーションにおける大気 CH <sub>4</sub> 濃度の連続観測データセット	DOI:10.17595/20160901.004	0
波照間ステーションにおける大気 N <sub>2</sub> O 濃度の連続観測データセット	DOI:10.17595/20170113.001	0
落石岬ステーションにおける大気 N <sub>2</sub> O 濃度の連続観測データセット	DOI:10.17595/20170113.002	0
ODIAC 化石燃料燃焼による二酸化炭素排出量のデータセット	DOI:10.17595/20170411.001	5
富士山頂の大気中 CO <sub>2</sub> 濃度データセット	DOI:10.17595/20170616.001	1
民間航空機観測(CONTRAIL)の CO <sub>2</sub> 濃度連続観測データセット	DOI:10.17595/20180208.001	2
Improved Limb Atmospheric Spectrometer (ILAS), Version 6.1	DOI:10.17595/20180628.001	0
Improved Limb Atmospheric Spectrometer (ILAS), Version 8.0	DOI:10.17595/20180628.002	0
Improved Limb Atmospheric Spectrometer-II (ILAS-II), Version 2	DOI:10.17595/20180628.003	0
Improved Limb Atmospheric Spectrometer-II (ILAS-II), Version 3.0	DOI:10.17595/20180628.004	0
CONTRAIL フラスコサンプリングによる太平洋上空の大気微量気体データ	DOI:10.17595/20190828.001	2
CONTRAIL フラスコサンプリングによるユーラシア大陸上空の大気微量気体データ	DOI:10.17595/20190828.002	2
南極昭和基地におけるフーリエ変換赤外分光(FTIR)による O <sub>3</sub> , HNO <sub>3</sub> , HCl の高度分布及び気柱全量データ(2007年)	DOI:10.17595/20190911.001	0
南極昭和基地におけるフーリエ変換赤外分光(FTIR)による O <sub>3</sub> , HNO <sub>3</sub> , HCl の高度分布及び気柱全量データ(2011年)	DOI:10.17595/20190911.002	0
データ駆動型アップスケーリングモデルによって推定された全球の GPP(総一次生産), NEE(純生態系交換), ER(生態系呼吸)	DOI:10.17595/20200227.001	0
CMIP5 をベースにした CDFDM 手法による日本域バイアス補正気候シナリオデータ	DOI:10.17595/20200415.001	1
ニューラルネットワークを用いて推定した全球海洋表層 CO <sub>2</sub> 濃度と吸収量(英語)	DOI:10.17595/20201020.001	0
逆解析システム NISMON-CO <sub>2</sub> による長期全球 CO <sub>2</sub> フラックスデータ	DOI:10.17595/20201127.001	0
東京都における化石燃料起源二酸化炭素排出量推定値の 1x1 km データ	DOI:10.17595/20210129.001	0
CMIP6 をベースにした CDFDM 手法による日本域バイアス補正気候シナリオデータ	DOI:10.17595/20210501.001	0
東京・代々木における大気中 CO <sub>2</sub> 濃度観測データ(NIES 分析)(英語)	DOI:10.17595/20210510.001	0
東京・代々木における大気中 CO 濃度観測データ(英語)	DOI:10.17595/20210510.002	0
東京・代々木における大気中 CO <sub>2</sub> 濃度観測データ(AIST 分析)(英語)	DOI:10.17595/20210510.003	0
東京・代々木における大気中 O <sub>2</sub> 濃度観測データ(英語)	DOI:10.17595/20210510.004	0
東京・代々木における大気中 CO <sub>2</sub> フラックス観測データ(英語)	DOI:10.17595/20210510.005	0
陸域生態系モデル VISIT による温室効果ガス収支及び炭素循環シミュレーション出力データ	DOI:10.17595/20210521.001	0
苫小牧フラックス観測サイトにおける微気象学的 CO <sub>2</sub> フラックス観測データ	DOI:10.17595/20210611.001	0
国設八方尾根酸性雨測定所における大気オゾン濃度の観測データセット	DOI:10.17595/20210709.001	0
富士北麓フラックス観測サイトにおける微気象学的 CO <sub>2</sub> フラックス観測データ	DOI:10.17595/20210730.001	0
1980~2020年の全球海洋 CO <sub>2</sub> フラックスの推定値(英語)	DOI:10.17595/20210806.001	0
国立環境研究所化学気候モデルを使った将来の HFC 増加のオゾン層への影響を調べるための 100 アンサンブルシミュレーションのアウトプット	DOI:10.17595/20210806.002	0
CONTRAIL-CME による大気 CO <sub>2</sub> 濃度データ(最新データ)	DOI:10.17595/20210827.001	0
CONTRAIL フラスコサンプリングによるユーラシア大陸上空の大気微量気体データ(最新データ)	DOI:10.17595/20210827.002	0



図1 CGER データ基盤.

も計画しており、連携してメタデータ・データの管理・公開・検索・利活用支援ができるようになる予定である(図1)。

## 5. おわりに

地球環境データというと、広くは地理データや気象データも含まれるが、本稿では、NIES CGERが推進してきた地球環境データベースに注目して、その展開を追った。この数十年の間に情報媒体の変化、通信技術の発達等により、データ流通が大きく変化し、データ基盤に求められる役割も変わった。NIESは、国内の環境研究の中核的機関として、多分野にまたがる研究データを生み出し、集積してきたが、2017年に「積極的にデータ公開し、国の研究開発成果の最大化に資する」という、研究所のデータポリシー(国立環境研究所, 2017)を策定した。これは、国としての動き(内閣府, 2018)に先んじた取り組みだったが、このデータポリシーは理念的なものであり、現場のインセンティブ向上には、より実質的な支援が必要である。

地球環境データベース事業は、立ち上げ時には具体的なデータベースサイトが存在したわけではなかったが、インターネットの普及やオープンサイエンスの潮流に押され、GEDやRDMS等の研究データ基盤を開発・運用するようになった。DOIを付与して公開するだけであれば、近年、整備が進みつつある機関リポジトリや、分野や機関を特定しない汎用リポジトリでも可能だが、地球環境モニタリングのように、長期間継続して系統的に取得しているデータの場合、独自リポジトリから公開することで、データセットの履歴や、関連データを一箇所から公開できるメリットは無視できない。

地球環境モニタリングが産み出す多様なプラットフォームを用いた長期観測データの数々はその独自性と信頼性から国際的にも広く研究コミュニティで評価されている。GEDからのデータ公開により、データが利用しやすくなったとはいえ、見つけれやすさ(FAIR原則のFindable)のためには、今後も分野の国際データベースにも登録することが推奨される。ただ、その際、NIESのDOIと、国際データベースのDOIのどちらを引用すべきか、両方なのか等について、まだ決まったルールはない。このように、データ引用の手法が国際的にも統一されていないこと、さらにデータ利用者の間でもデータ引用が習慣づいていないことは、データ提供側がデータの利用状況を把握する際の大きな障害となる。データ公開へのインセンティブ向上には、データ公開をデータ提供者の成果として評価することが望まれるが、そのためにはデータ相互利用の方法を標準化し、データ引用の文化を浸透させることが課題である。この他にも、研究現場においてはデータ公開までの障害がまだ多い(白井, 2021)。

2019年12月頃に拡大を始めた新型コロナウイルス感染症の蔓延により、テレワークやオンライン授業など、社会のデジタル化の必要性が急速に認知された。第6期科学技術・イノベーション基本計画(2021年3月)でも、デジタルトランスフォーメーションやデータ利活用の重要性が強調され、2021年9月にはデジタル庁が発足するなど、国を挙げて社会のデジタル化が進められる情勢である。社会が急激に変化する中、今後の研究データ管理・公開は、学術分野にとどまらず、民間やシチズン・サイエンス(市民科学)との連携も視野に入れて進める必要があるだろう。

地球環境問題は、現実社会で起きている、全人類

の未来にかかわる課題であり、解決には社会全体の関心・協力が必要である。つまり今の動きは地球環境研究の立場としては追い風であり、この機を利用して、分野・セクター・国を超えて、問題の周知から政策・行動変容までを加速できるチャンスでもある。例えば、NIESが第5期中長期計画(2021~2025年度)において実施中の気候変動・大気質研究プログラムでは、気候変動問題対策として2020年から運用されているパリ協定に貢献するため、国際社会全体の温室効果ガスの排出量削減の達成度を5年ごとに評価するグローバルストックテイクへの情報提供を大きな目標の1つとしている。また、2050年までのカーボンニュートラル実現に向け、所内の気候変動適応センターや社会システム領域等とも連携して、率先して情報発信や社会への働きかけを行う。GEDは、そのデータ収集・管理・公開を支援し、各国の各分野の研究者、データ・情報センターと連携し、社会との接点を意識しながら最新の情報を積極的に発信していく。貴重な研究データをしっかり管理するという基本を大切にしつつ、時代に合わせた情報発信・データ利活用促進を実施できるようこれからも発展を続けていきたい。

## 謝 辞

これまでNIESからのデータ発信及び地球環境データベースの運用に携わって来られた皆様に敬意を表す。三枝信子氏をはじめとするCGERの皆様の日常的な協力・支援に感謝する。METEX等の支援ツールを開発・運用している曾維業氏、GEDの改良・RDMS開発の主担当を務める福田陽子氏、インフラ管理の塚田康弘氏にはDB室の活動を支えて頂き感謝する。この原稿作成に当たりご助言頂いた西岡秀三氏及び笹野康弘氏に謹んで感謝する。

## 引用文献

- 地球環境研究センター(1990)地球環境研究センター紹介. 地球環境研究センターニュース 1990年10月号. <http://www.cger.nies.go.jp/publications/news/vol1/vol1-1.pdf>. (2021年9月28日確認)
- 地球環境研究センター(1992)国際研究計画・機関情報. 国立環境研究所地球環境研究センター報告 CGER-D002-'92. <https://www.cger.nies.go.jp/publications/report/d002/D002.zip>. (2021年9月28日確認)
- 地球環境研究センター(1994)GRID 全球データセットユーザーズガイド. 国立環境研究所地球環境研究センター報告 CGER-D004-'94. <https://www.cger.nies.go.jp/publications/report/d008/gridpamp.html>. (2021年9月28日確認)
- 地球環境研究センター(1995)GRID-TSUKUBA(パン

- フレット). 国立環境研究所地球環境研究センター報告 CGER-D008-'95. <http://www.cger.nies.go.jp/publications/report/d004/D004.zip>. (2021年9月28日確認)
- 地球環境研究センター(2000)地球環境研究センターってどんなところ? 地球環境研究センターニュース 2000年4月号. <http://www.cger.nies.go.jp/publications/news/vol11/11-1.pdf> (2021年9月28日確認).
- FORCE11 (2016) The Fair Data Principles. <https://www.force11.org/group/fairgroup/fairprinciples> (2021年9月28日確認)
- 藤原正弘(1977)国立公害研究所における情報サービス. 情報管理, 19(12), 947-954. <https://doi.org/10.1241/johokanri.19.947>
- 原沢英夫(1993)地球資源情報データベース: GRID. 水文・水資源学会誌 6(1), 66-70. <https://doi.org/10.3178/jjshwr.6.66>
- 春山 暁・細山美樹(1987)JOISによるINFOTERRA環境情報源ファイル. 情報管理 30(4), 335-344. <https://doi.org/10.1241/johokanri.30.335>
- Hayashida-Amano, S., Sasano, Y. and Iikura, Y. (1991) Volcanic disturbances in the stratospheric aerosol layer over Tsukuba, Japan, observed by the National Institute for Environmental Studies lidar from 1982 through 1986. *Journal of Geophysical Research: Atmospheres*, 96(D8), 15469-15478. <https://doi.org/10.1029/91JD01392>
- ジャパンリンクセンター運営委員会(2015)研究データへのDOI登録ガイドライン. [https://doi.org/10.11502/rd\\_guideline\\_ja](https://doi.org/10.11502/rd_guideline_ja).
- 国立環境研究所(1976-2021)国立環境研究所年報, ISSN (print) 1341-3619, ISSN (online) 2187-8919. <https://www.nies.go.jp/kanko/nenpo/index.html> (2021年9月28日確認)
- 国立環境研究所(1998)日本海重油汚染事故調査資料. 環境庁 国立環境研究所業務報 F-111-'98. <https://www.nies.go.jp/kanko/gyomu/pdf/f111-1998.pdf>. (2021年9月28日確認)
- 国立環境研究所(2017)国立研究開発法人国立環境研究所データの公開に関する基本方針(データポリシー). [https://www.nies.go.jp/kihon/kitei/kt\\_datapolicy.pdf](https://www.nies.go.jp/kihon/kitei/kt_datapolicy.pdf). (2021年9月28日確認)
- 内閣府(2018)国立研究開発法人におけるデータポリシー策定のためのガイドライン. <https://www8.cao.go.jp/cstp/stsonota/datapolicy/datapolicy.pdf> (2021年9月28日確認)
- 白井知子(2021)地球環境データベース: 30年の歩みとこれから. 情報処理学会論文誌デジタルプラクティス (DP), 2(2). <https://www.ipsj.or.jp/dp/contents/publication/46/TR0202-01.html> (2021年9月28日確認)

- 武田英明・村山泰啓・中島律子(2016)研究データへのDOI登録実験. 情報管理, 58(10), 763-770. <https://doi.org/10.1241/johokanri.58.763>
- 土屋 巖(1982)国立公害研究所における文献情報のオンライン検索について. 情報管理 25(2), 137-146. <https://doi.org/10.1241/johokanri.25.137>
- 土屋 巖(1983)国際機関における情報活動〔第9回〕UNEP(国連環境計画). 情報管理, 26(9), 745-759. <https://doi.org/10.1241/johokanri.26.745>
- 土屋 巖・春山 暁・広崎昭太(1980)国立公害研究所の情報システムの現状と将来構想. 情報管理, 23, 548-559. <https://doi.org/10.1241/johokanri.23.548>
- Zeng, J., Matsunaga, T. and Mukai, H. (2010) METEX: A flexible tool for air trajectory calculation. *Environmental Modelling & Software*, 25(4), 607-608. <https://doi.org/10.1016/j.envsoft.2008.10.015>.
- Zeng, J., Nakajima, H., Matsunaga, T., Mukai, H., Hiraki, K. and Yokota, Y. (2012) Linking Carbon Dioxide Variability at Hateruma Station to East Asia Emissions by Bayesian Inversion. In: Lin, J., Brunner, D., Gerbig, C., et al. (eds.) *Lagrangian Modeling of the Atmosphere*, 200, 163-172. American Geophysical Union, Washington, DC. <https://doi.org/10.1029/2012GM001245>



白井 知子 / Tomoko SHIRAI

国立環境研究所 地球システム領域 地球環境データ統合解析推進室 室長 / グローバル・カーボン・プロジェクト つくば国際オフィス代表 / 企画部 主席企画連携主幹。

東京大学理学部化学科卒業。東京大学大学院理学系研究科修了，博士(理学)。宇宙航空研究開発機構開発部員，米カリフォルニア大学アーバイン校研究員を経て，2004年，国立環境研究所入所。大気化学・炭素循環研究とならび，2011年より地球環境データのデータベース化をはじめとする研究データ管理公開基盤の構築・運用やデータ利活用の推進を行う。