

オープンサイエンスを巡る世界の最新動向 極域データ分野

矢吹裕伯

国立極地研究所 国際北極環境研究センター

情報・システム研究機構 データサイエンス共同利用基盤施設

本日の講演

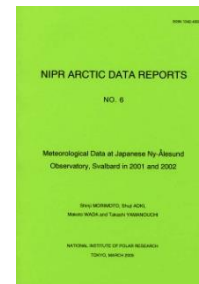
- 極地研のこれまでの取り組み
- 北極域におけるデータの取り組み
- 国立極地研究所によるオープンサイエンスの取り組み
- 今後の検討

国立極地研究所のデータに関するこれまでの取り組み

年月	出来事
1957年1月	「昭和基地」開設
1957年12月	「南極資料」創刊
1968年8月	「JARE Data Reports」創刊
1991年9月	「ニーオルスン基地」開設(北極観測拠点)
1996年12月	「NIPR Arctic Data Reports」創刊
2007年8月	学術データベース(後の極域メタデータベース)のサービス開始
2012年8月	北極域データアーカイブ(Arctic Data archive System)のサービス開始
2014年10月	JaLC DOI実験プロジェクトに参加
2014年10月	極地研学術情報リポジトリのサービス開始
2015年7月	データレポート将来構想WGの活動開始(データ出版)
2015年9月	極地研第1番目のデータDOI付与(doi:10.17592/001.2015093001)
2017年1月	「Polar Data Journal」創刊

国立極地研究所の取り組み

- 現在：
 - 観測データ報告 (編集・査読)
 - JARE-data reports
 - NIPR arctic data reports

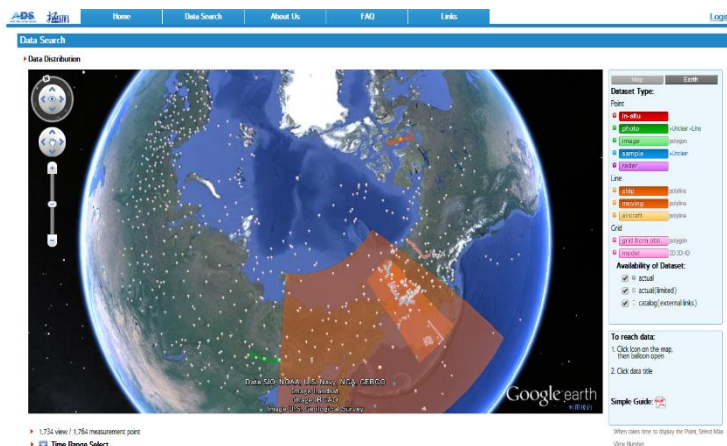


- 検討中 (データ出版)
 - Data Reportsを発展させ査読制度を設ける
 - 品質を確保する
 - これらの査読を通過したデータに関してDOIの付与を行う。
 - 極地研独自のデータ品質管理
 - 極地研のデータ出版としてdata reportsの活用 (CrossReff ?)
 - データリポジトリとしてADSの活用 (DataCite)
- 学術論文に利用されたデータの登録 (DOI付与)
- 他データジャーナルに投稿されたデータ (DOI付与)

ADS Webアプリケーション群

極KIWA

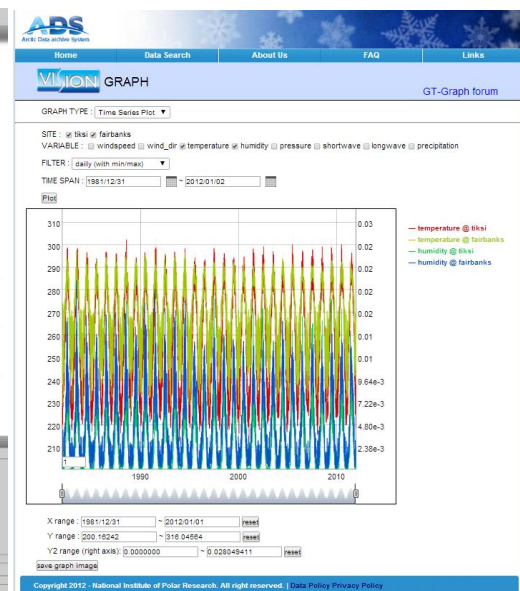
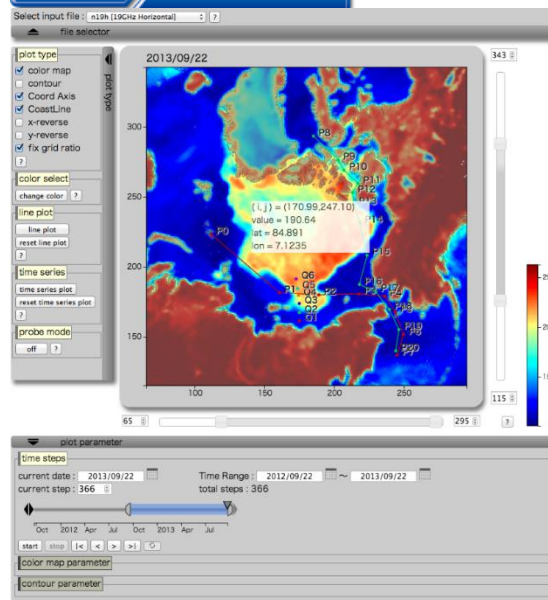
研究データ検索・
管理システム



データ登録数: 320

VISION

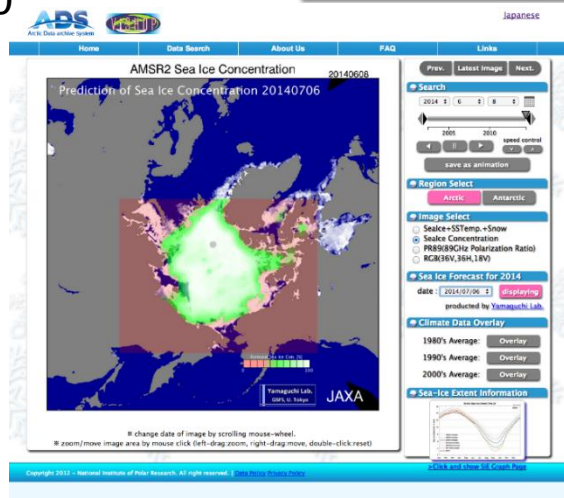
オンライン可視化アプリケーション



ViSHOP

極域環境監視モニター

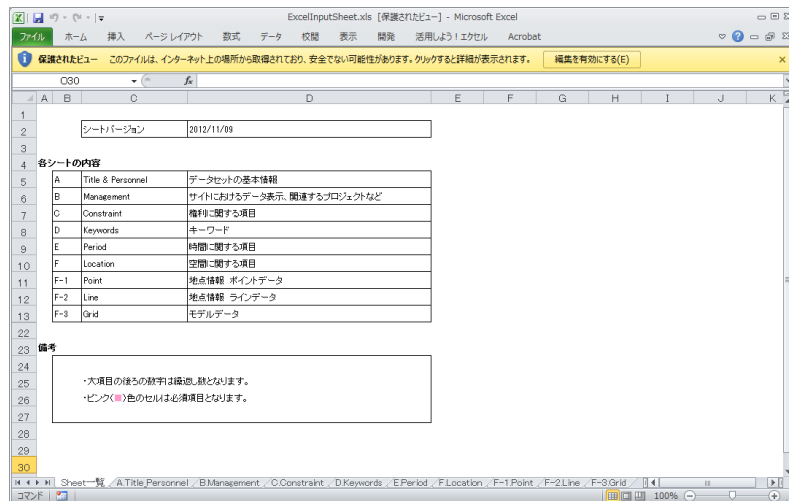
- ・可視化された衛星データを準リアルタイムでWebで提供する



ブラウザ上でグリッドデータや
時系列データを可視化し、解
析を可能にするオンラインで可
視化するWebアプリケーション

<https://ads.nipr.ac.jp/>

ADSメタデータ



メタ情報規格

- **国際規格:** **ISO19139**
- **関連プロジェクトの規格:**
DIAS, GCMD DIF, DataCite
⇒各々の必須項目は必ず含み、マッピングに対応
- **ADS独自規格:** **ADS**

関連する規格	カテゴリ	要素数
DIAS	地球科学全般	85
DIF (Directory Interchange Format)	GCMDの定義	55
DataCite	科学データDOI関連	18

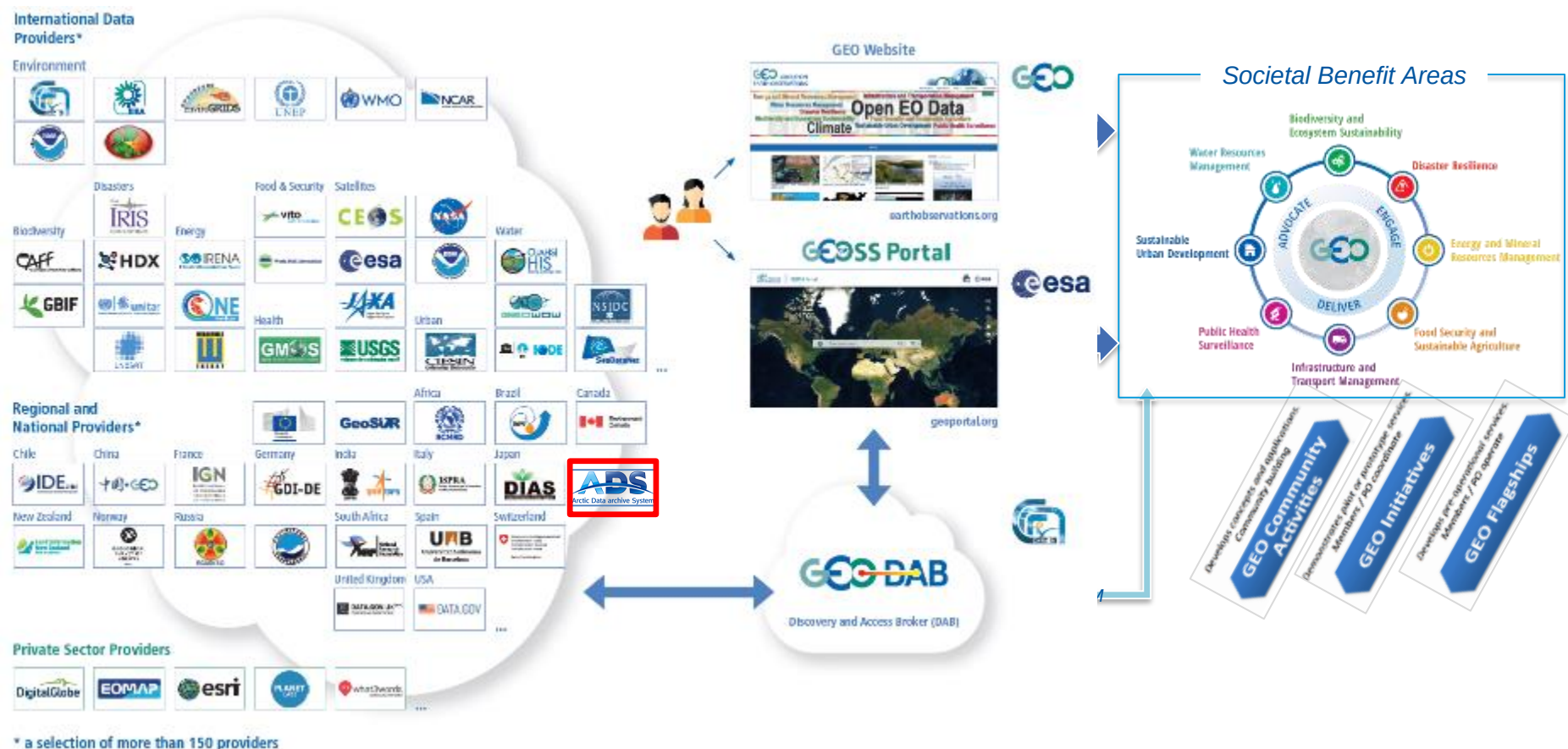
シート	シート名	項目	要素数(必須項目)
A	Title & Personnel	データセットの基本情報	32(22)
B	Management	サイトにおけるデータ表示、関連するプロジェクトなど	22(5)
C	Constraint	権利に関する項目	16(5)
D	Keywords	キーワード	14(5)
E	Period	時間に関する項目	11(2)
F	Location	空間に関する項目	17(4)
F-1	Point	地点情報 ポイントデータ	20
F-2	Line	地点情報 ラインデータ	8
F-3	Grid	モデルデータ	17



この部分: データの種類により入れ替え(粒度の細かいデータセットを記述できる)

メタデータの記載は全てエクセルシートを用いて行う。エクセルシートではマクロが組んであり、入力補助を行っている。

GEOSS (Global Earth Observation System of Systems : 全球地球観測システム) Portal

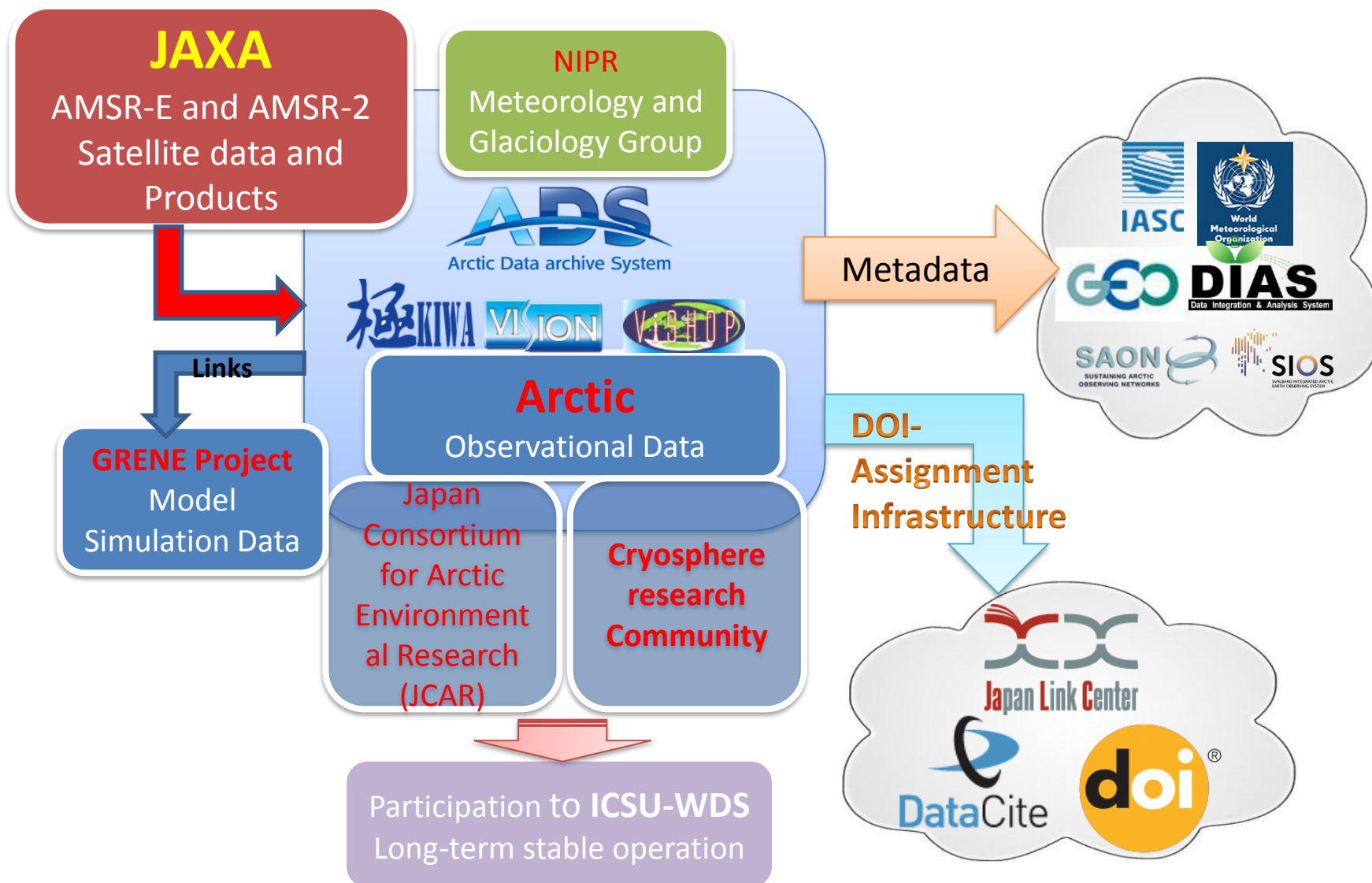


GEOの連携データプロバイダー
現在150以上のデータセンターとの連携

- ・国際的なデータセンター
- ・地域的なデータプロバイダー
- ・民間のデータセンター

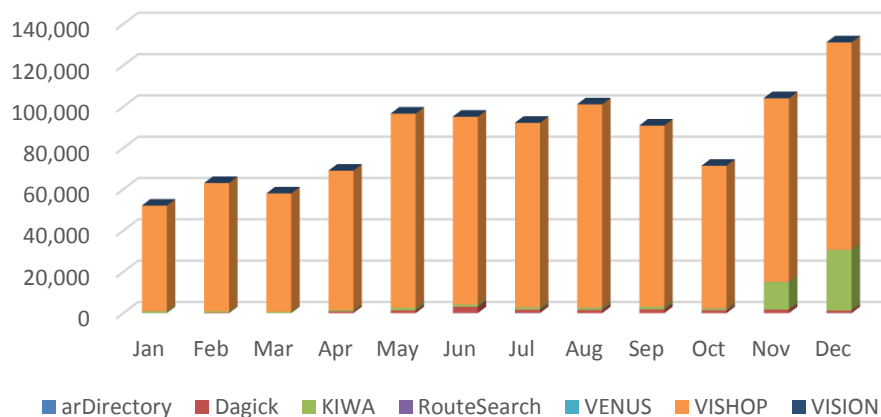
GEOへのメタデータの連携のメリット

1. 新たなる共同研究
2. Funding

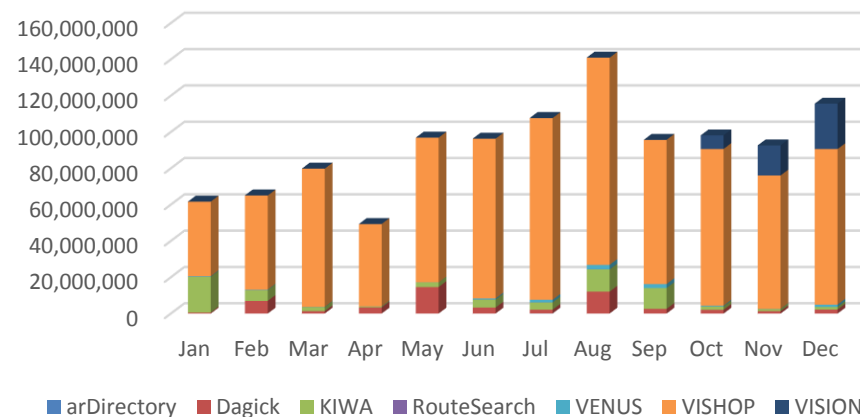


ADSのアクセス数(訪問者数・転送量) 2016年1月-12月

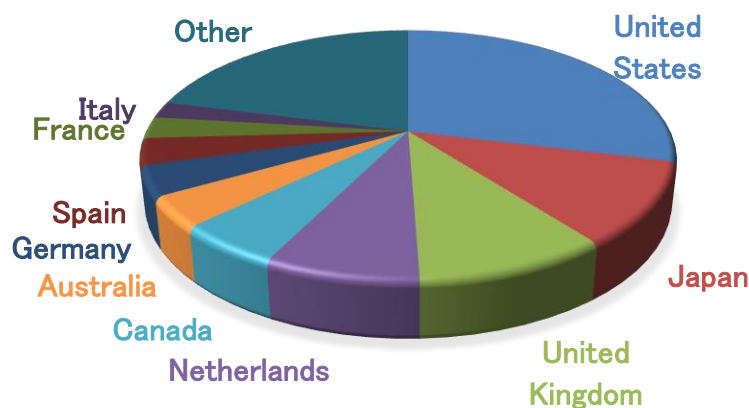
訪問者数



データ転送量



国別訪問者数



- 世界中からアクセスされている
- 1年間のデータ転送量が約1.3TB
- 1日当たり百万ファイルを転送

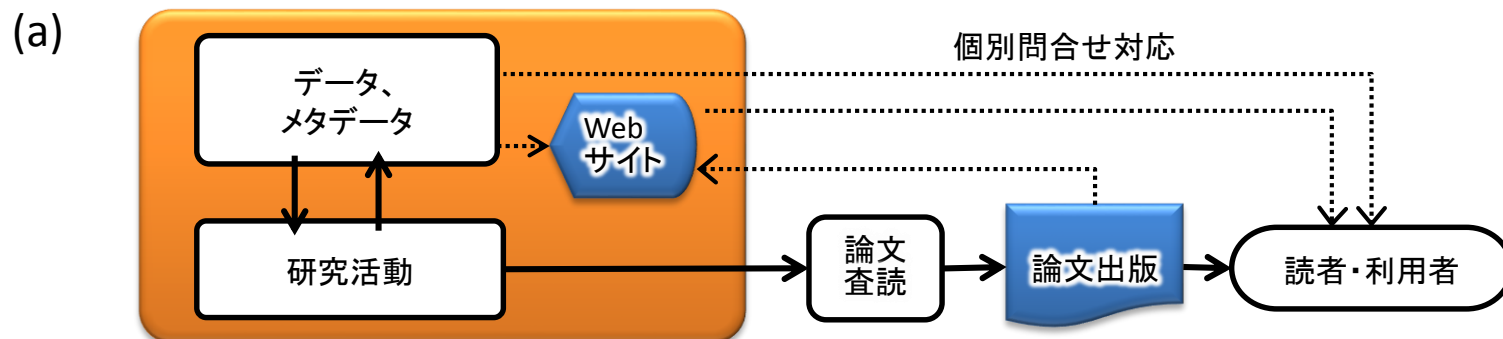
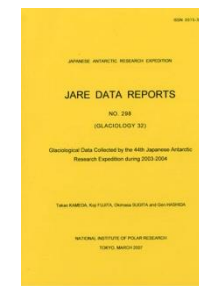
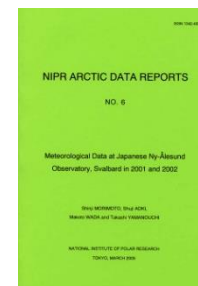
極地研によるオープンサイエンスの推進

- データセンターの使命
 - 観測・研究データの保管・管理・公開
- データセンターとしてのニーズ
 - 登録データの増加および利用者の増加
 - データ提供者へのメリット提供

データ出版

- データ利用者のメリット
 - 論文に引用されるデータにアクセスができる。
 - 論文の引用から新たなデータへのアクセスが可能
- データ提供者のメリット＝>
 - データ公開が業績・評価の対象
 - データの利用数、被引用数を調べられる=>被引用論文数と同じ
- データセンターのメリット
 - データの整形等を通して2次利用が可能になり、データベースへの利用が可能になる。(統合的データベースの構築)
 - データ提供者へのインセンティブの確保

- これまで:
 - 研究成果及び公式記録(査読)
 - 南極資料
 - 観測データ報告(査読)
 - JARE-data reports(371編)
 - NIPR arctic data reports(7編)
 - 2017年3月現在



RDA-WDS data publishing workflow WG output (RDA, 2017)
南山他 (2017)

論文出版の概念図

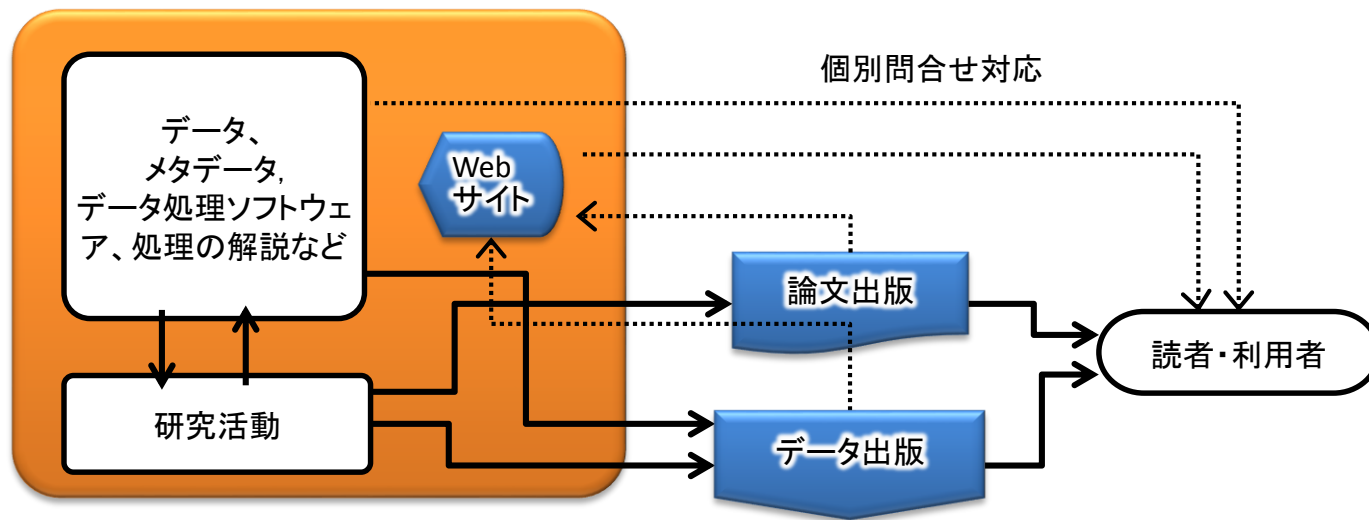
国立極地研の検討

- **データ出版**
- 「JARE Data Reports」及び「NIPR Arctic Data Reports」の役割の再考
- 実データの再利用促進
- 公開されたデータにDOIを付与する際のデータの品質保証

データDOIの付与

- DOIは単なる恒久的なパーマントリンクを確保する“しくみ”。
- データセンター
 - **研究・観測データの品質が確保**されることが重要
 - データセンター自体の品質にかかわる
- DOI発行者の責務
 - 公開・出版されるデータの**品質の確保**
 - DOIを発行するということは研究・観測データの品質の確保が**最重要課題**⇒データセンターにとって重要なアクション

(b)



RDA-WDS data publishing workflow WG output (RDA, 2017)
南山他 (2017)

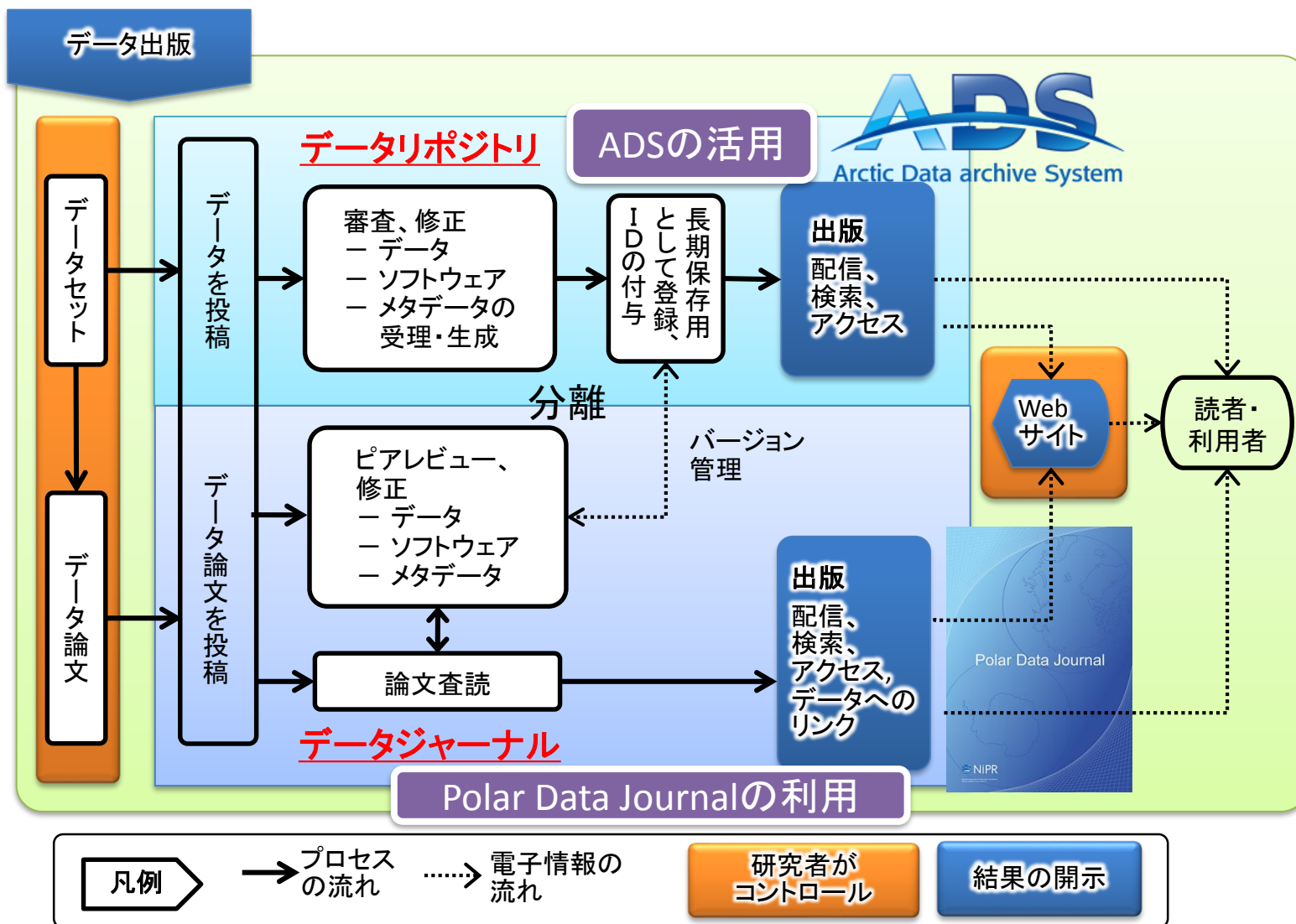
データ出版の概念図

Polar Data Journalの創刊(2017年1月)

- ・ 実データに対する詳細な記述を提供
- ・ 論文という媒体を通じて対象となる実データの可視性を高める
- ・ データの品質確保
- ・ 科学的に新たな知見の記述を行わない

<https://pdr.repo.nii.ac.jp/>

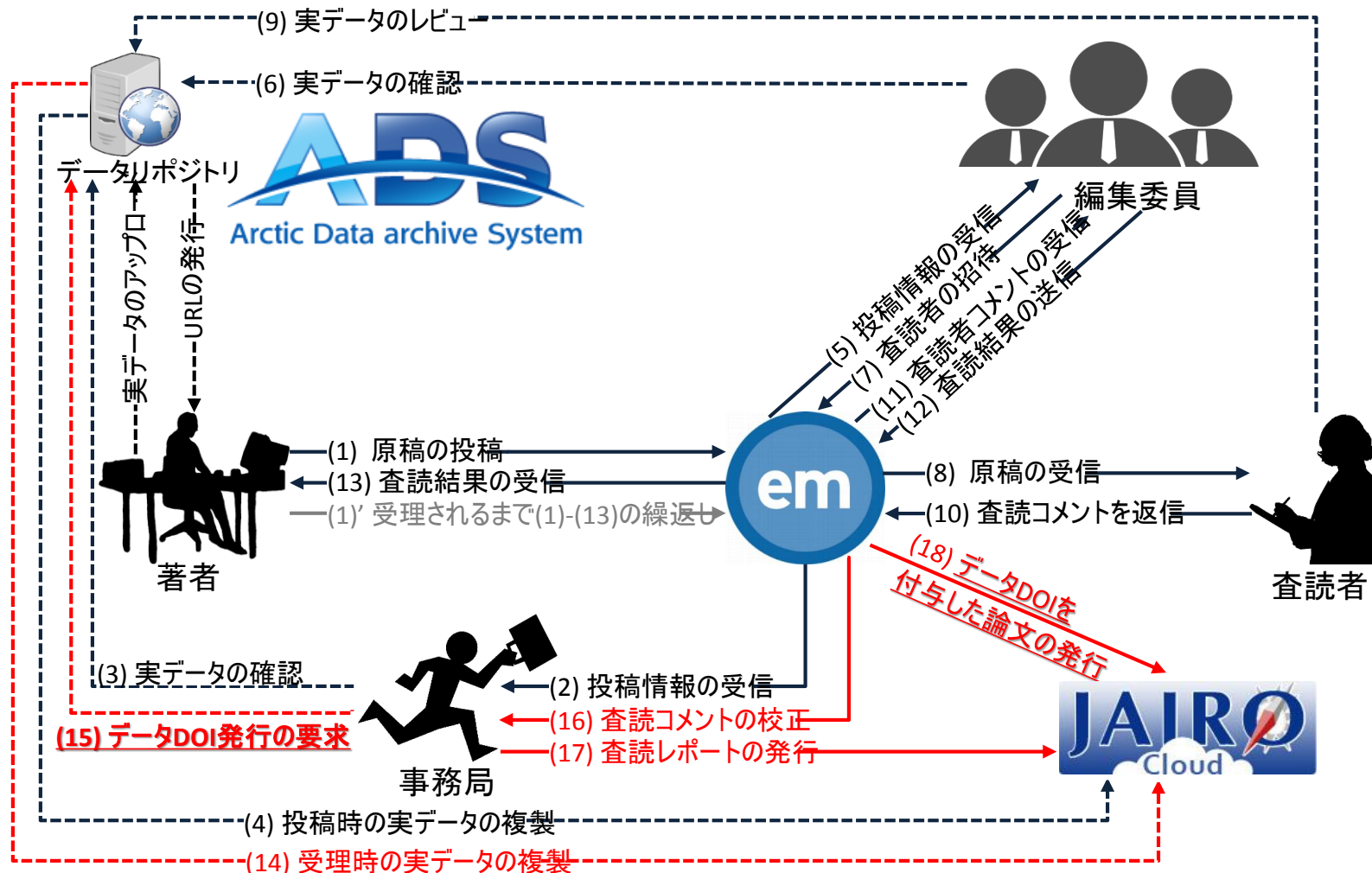




RDA-WDS data publishing workflow WG output (RDA, 2017)
南山他 (2017)

データ出版の概念図

Polar data journal の投稿プロセス



- 査読中のワークフロー: (1) ~ (13)
- 受理後のワークフロー: (14) ~ (18)

南山他(2017)

データジャーナル『Polar Data Journal』 創刊の取り組み

極域科学データの新たな公開体制構築に向けて

Launching a new data journal "Polar Data Journal"

Toward a new data publishing framework for polar science

南山 泰之¹ 照井 健志¹ 村山 泰啓² 矢吹 裕伯¹
MINAMIYAMA Yasuyuki¹ TERUI Takeshi¹ MURAYAMA Yasuhiro² YABUKI Hironori¹
山地 一禎³ 金尾 政紀¹
YAMAJI Kazutsuna³ KANAO Masaki¹

¹ 情報・システム研究機構 国立極地研究所

¹ Research Organization of Information and Systems, National Institute of Polar Research

² 情報通信研究機構

² National Institute of Information and Communications Technology

³ 情報・システム研究機構 国立情報学研究所

³ Research Organization of Information and Systems, National Institute of Informatics

「データジャーナル」は従来の「出版」の枠組みを活用しながら、データの保存や再利用を促すために有効な手法と考えられているが、国際的にも概念形成の途上である。本稿では、国内におけるデータジャーナル構築、運用に向けた具体的な検討材料を提供するため、国内での実践例として2017年1月に創刊されたデータジャーナル『Polar Data Journal』を紹介する。具体的には、背景としての国立極地研究所におけるデータ公開の歴史、構想から創刊に至るまでの合意形成プロセス、「データジャーナル」を通じてデータを適切に取り扱うために必要な要件やワークフロー、およびそれらと連携する基盤構築について示す。

データジャーナル、データ出版、研究データ、オープンサイエンス、データリポジトリ、Polar Data Journal, JAIRO Cloud, 極域科学

原稿受理 (2017-03-21)

情報管理. 2017, vol. 60, no. 3, p. 147-156. doi: <http://doi.org/10.1241/johokanri.60.147>

極地研によるデータ出版の取り組み

情報管理. 2017, vol. 60, no. 3, p. 147-156. doi: <http://doi.org/10.1241/johokanri.60.147>

今後の検討

国立極地研究所で管理する標本・サンプルの管理

極域動物標本データベース <http://animal.nipr.ac.jp/>

2687レコード

極域動物標本データベース

整列 登録番号 学名 和名

登録番号 A00001 - 0001 一般分類/Eng. 鳍脚類 seals

学名 Leptonychotes weddellii 命名者名 (Lesson)

和名 ウェッデルアザラシ 英名 Weddell seal

門/Phylum 綱/Class 目/Order 科/Family
脊索動物門 哺乳綱 食肉目 アザラシ科
CHORDATA MAMMALIA CARNIVORA PHOCIDAE

採集地域 East Antarctica

採集場所 昭和基地周辺
/Eng. Syowa Station area

採集者
/Eng.

採集日 1961/12/22

備考 成体
/Eng. adult

標本写真



写真番号 ph00001

国立極地研究所 生物資料室

データDOI付与の検討

前例、何かいいアイデアはありますか？

ありがとうございました。

URL: <https://ads.nipr.ac.jp>

Contact : ads-info@nipr.ac.jp

Twitter: ADS_NIPR

YouTube: NIPR_ADS

Facebook:ArcticDataarchiveSystem