

SGEPSS  
(地球電磁気地球惑星圏学会)  
将来構想における  
データ関連活動

村山泰啓、堀 智昭, 篠原 育  
家森俊彦

# もくじ

---

- 科学データ・情報の活用とサイエンス
- SGEPPSS・関連分野の活動例紹介
- 将来像の議論

# 過去のSGEPSS将来構想における「データ」

- 平成3年「地球電磁気学の発展的将来」、全67ページ。  
(全67ページ; 日本学術会議地球電磁気研究連絡委員会)
  - データネットワークのセクションを設けて、提言。
- 平成17年「21世紀の地球電磁気学」全14ページ。(日本学術会議地球電磁気研究連絡委員会)
  - データのセクションはないが、DB重要性の記述が各分野に埋め込まれていた。

## 第5章 研究推進のために必要な施策 ……………

- 1 宇宙空間科学および中層・超高層大気科学
  - 1. 1 惑星研究の推進 ……………
  - 1. 2 国際協力の推進 ……………
  - 1. 3 海外拠点観測所 ……………
  - 1. 4 航空機観測体制の確立 ……………
  - 1. 5 理論・シミュレーション ……………
  - 1. 6 データネットワーク・総合解析 ……………
  - 1. 7 研究・教育体制 ……………

(地球電磁気学の発展的将来、1991)

# サイエンスの展開とデータ・情報

---

- SGEPSSは新たな技術で新たな科学を切り開いてきた
  - 宇宙技術⇔衛星・ロケットによる宇宙空間観測
  - 電子技術⇔MULレーダー等によるレーダー観測
- いま：  
情報技術は、新しい科学の突破口になりうるのか？
- 情報爆発
  - 巨大な情報量からこれまでにないサイエンスが可能か？
  - 科学者は保有するデータを100%生かしているか？
  - ビッグデータ戦略(米政府) . . . etc.

## 注意点：

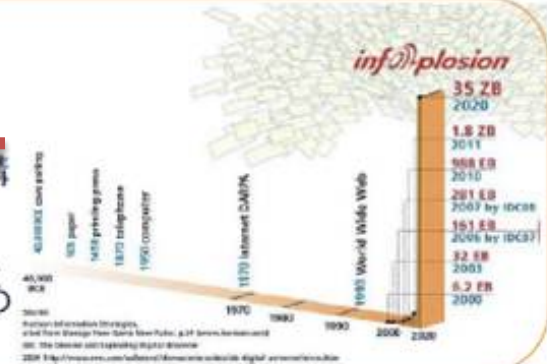
宇宙技術も、MULレーダーも、最先端の工学的学術研究というより、応用技術の最先端であった。いま情報・データについて必要なのは何か？  
は難しい問い。

# ビッグデータを巡る状況

## 背景

○デジタル情報は、2020年までには約35ゼタバイト(2010年時の約35倍、DVDメディア約140億枚)へ拡大する見込みである。また、その膨大な情報(ビッグデータ)の中には、新たな知識や洞察が埋もれているといわれているが、その多くが整理・構造化されておらず、有効に活用できていない状況。

○例えば、ライフサイエンス分野においては次世代シーケンサーの高性能化により、大量のデータが得られることになるが、大量データを活用して研究開発を進めるための情報分野からのさらなる対応が必須。



## 国内外の状況

○米国では、平成24年3月29日にBigDataイニシアチブに関する公告を発表。このイニシアチブには6機関(NSF、NIH、DOD、DARPA、DOE、USGS)が総額2億ドル(185億円相当)を投資し、データへのアクセス、体系化、知見を集める技術を改善強化することが挙げられている。その他、欧州、アジアにおいても、ビッグデータに対する研究投資が行われている。

○我が国は、各種センサー情報が発達していること、ハイパフォーマンスコンピューティング・自然言語処理などの領域において世界的に高い研究水準を保持していること、遺伝子情報など地域単位での研究が必要な必要が存在すること等から、大規模データの活用において、これらの強みが幅広い分野・領域に展開することで、我が国の科学技術の共通基盤の強化や産業競争力の強化が可能な環境である。

## さらなる研究が必要な項目

- 大規模データの蓄積・構造化・マイニング等のための研究開発
  - 新たな洞察を導くデータの解析可視化技術の開発
  - 文脈、意味・意図理解など高度なヒューマンインタラクションを実現するための研究開発
  - シミュレーションとデータ科学の融合(データ同化)、先進数理的手法や機械学習アルゴリズムの開発
  - アカデミッククラウド基盤構築のためのシステム研究
  - アプリケーション側(バイオインフォマティクス、地球観測、防災等)と連携したプロジェクト研究 等
- ※異分野の研究者による研究チームがアプリケーションデータを共有し、異なる方向性から、ビッグデータに迫っていく研究アプローチ  
※異分野の研究者同士のインタラクションや協働が可能となるよう、機動的な研究チームを編成する研究体制  
※上記の取り組みを通じて、ビッグデータに関係する研究者に流動的なネットワークを生み出し、新たな人材育成スキームを構築

# 改めて、「いま、なぜ、データか？」

- 近年、社会と科学者の関わりが問われている
  - 「社会との契約」「国民国家が支える科学」「科学者と社会との関わり」
- 科学者が生産し、社会にもたらすものは何か？
  - 科学的知の生成、原著論文による記録と保存：  
データ(情報)→分析→知見→原著論文・知の資産
    - (短期的には実利用価値が求められがちだが。)
    - 近代科学の指導原理の1つ: 追試験可能な系統的知識
    - 地球・宇宙科学では、データは論文中の真理の根拠。
      - まったく同じ現象は二度とない。
  - 研究者の自己満足でなく、後世の人類に残せるものは
    - 科学的知見: 原著論文
    - 科学的真理の証拠: 「データ」。
      - 論文の内容が「第三者によって追試験可能である」ことを担保する。
  - →「データの引用」・「データの出版」、  
という国際活動が進行中。

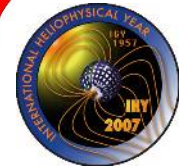


Helping you to find,  
access, and reuse data

DataCite

# 国際的な科学データ・情報に 関連する活動の例

- IPY2007-2008
  - Polar Information Commons
- eGY (Electronic Geophysical Year)
- IHY(Intl. Heliospheric Year) . . .
- WDS (World Data System)
  - WDC and FAGS から再編(2008～)
- CODATA (Committee of Science and Technology DATA)
  - 物理定数の国際標準値の策定などをしてきた。
  - 近年、「データベースIDの引用」、「データの出版」という考え方→国際活動が進展

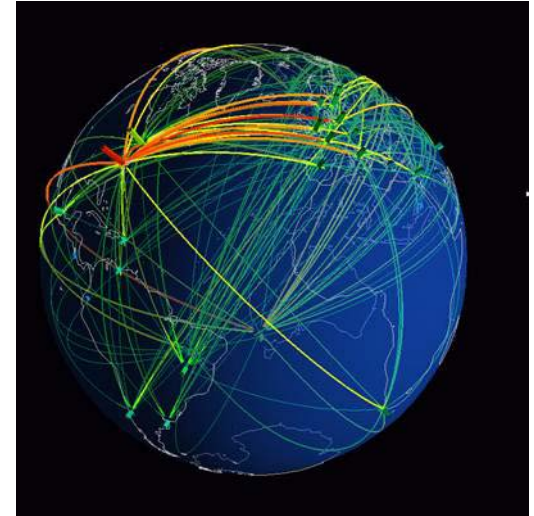




# drivers

## Modern capabilities

## Modern problems



## Government attitudes





# 国内の活動例から

- 京大地磁気センター、NIPR、宇宙研、京大生存研、NICT、名大、九大、地磁気観測所、東大、北大...
- SGEPPSS内でも種々の活動がありますが、その一部の例

# SGEPSSからの提言

- 定常的・連続的観測データの品質管理、維持・保全の重要性
  - － 特定の研究者によるデータ管理業務を解消すべき。(分業体制の未確立に起因する)
- 「連携データセンター」が学会として重要
  - － バーチャルな一体組織として機能するように
  - － 各機関のデータ専門スタッフの連携を図る
  - － 各機関に分散しているデータベースをネットワークで結合し運用
- 共通データ解析ツールの開発・供給
  - － 分散データセンターと研究者の研究活動を直接的に結びつける手段を提供
- データ維持・管理体制の強化の必要性
  - － 各機関にデータ専門スタッフを配置が急務。
  - － メタデータやデータファイルのデータベース化
  - － 関係組織との相互交換体制をつくることを条件づけた予算・人員の手当など  
⇒オープンなデータ流通機構の普及を推進していくことが必要不可欠である。
- 国際的潮流に対して、積極的に参加していくべき。
  - － 「データ・パブリケーション(data publication)」
  - － 「データ・サイテーション(data citation)」
    - サイテーションインデックスのように被引用頻度を通じた業績評価を行う

# 「連携データセンター」の発想

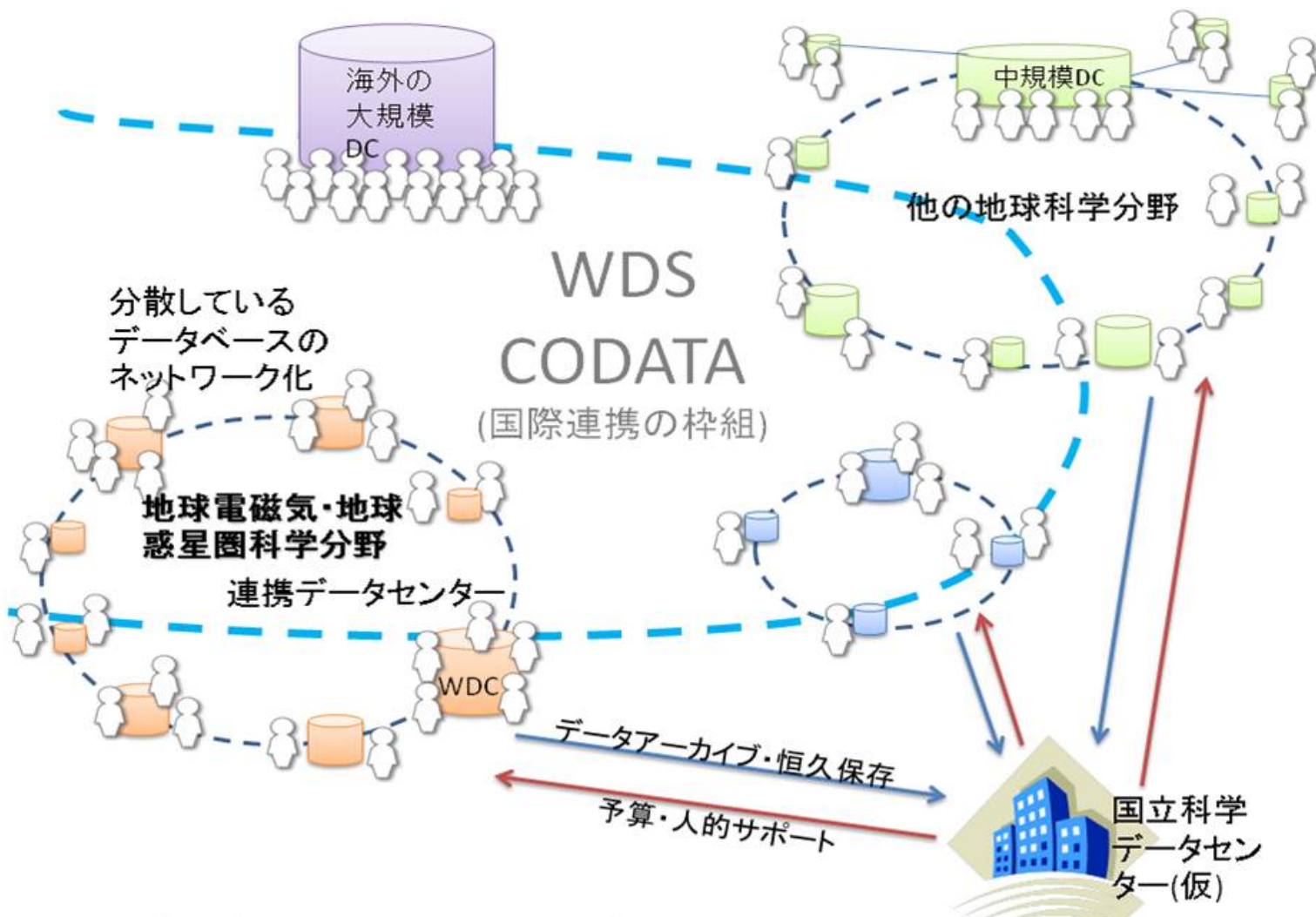


図4.3.1 連携データセンターと国立科学データセンターによる科学データサポート体制

# データパブリケーション、データ・サイテーションとは？

## ■ データパブリケーション

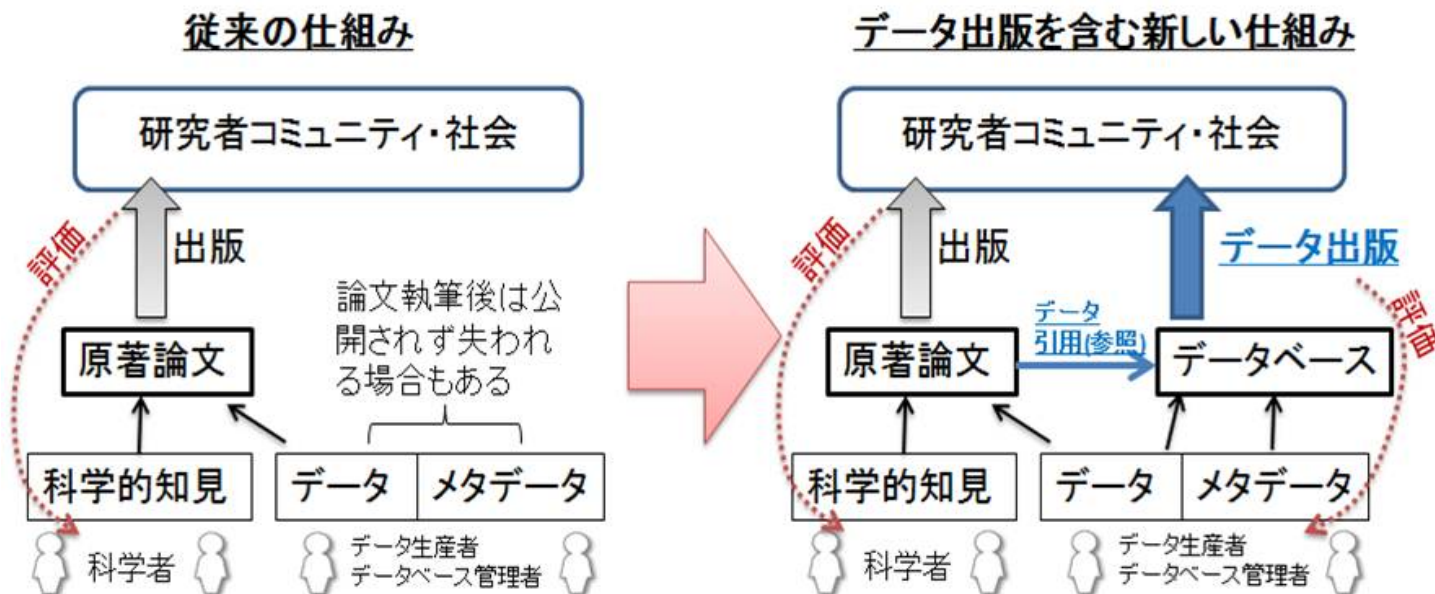
- データを「出版」する仕組み:
- 課題: データの「査読」「固定」「公表」等をどうするか。

## ■ データサイテーション

- データを文献のように「引用」「参照」する仕組み
- 課題: ID標準化、引用ルール確立、評価手法など国際団体等で模索中

## データを引用・参照すると

- 論文・書籍は知的生産力の基準⇒研究職・教育職の採択評価にも。
- 信頼できるデータ生成・提供は現代では科学者の仕事。←評価



# (参考) データサイテーションの活用施策例

## Building a Culture of Data Citation



出典「ands(※)」サイト

※オーストラリアの研究データ活用促進プログラム

<http://www.ands.org.au/cite-data/index.html>

**FIN.**