

機関番号：13301

研究種目：基盤研究（C）

研究期間：2008～2010

課題番号：20510006

研究課題名（和文）地球環境観測データ活用のための相互参照可能な現象アーカイブシステム構築

研究課題名（英文）Construction of phenomenon archives system for a cross-reference for observation data of the global environment

研究代表者

高田 良宏（TAKATA YOSHIHIRO）

金沢大学・総合メディア基盤センター・助教

研究者番号：30251911

研究成果の概要（和文）：

地球規模の全体像を研究し、地球環境を理解するための総合的な解析支援環境の構築が望まれている。我々は、地球環境観測データを活用するための相互参照可能な現象アーカイブシステムの構築を目指した。そして、その実現のために不可欠な、大規模地球環境観測データベースの構築手法、分散データベースコンテンツを相互参照・配信する手法、データベースからテキストデータにおけるキーワード検索のように現象を検索する手法の開発を行った。

研究成果の概要（英文）：

In order to study various phenomena with global-scale, it is necessary to construct a comprehensive analysis system which supports to understand the global environment. We aim to construct a phenomenon archives system for a cross-reference with regard to observation data on the global environment. To construct the archives system, we have developed a method to build an enormous database on earth observation, a method of cross-reference and delivery of contents stored in distributed databases, and techniques to search phenomena such as keyword search on text data.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008 年度	1,400,000	420,000	1,820,000
2009 年度	1,200,000	360,000	1,560,000
2010 年度	900,000	270,000	1,170,000
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野：情報地球環境科学

科研費の分科・細目：環境学・環境動態解析

キーワード：地球環境観測、現象アーカイブス、類似現象検索、認証、Shibboleth、共通プラットフォーム

## 1. 研究開始当初の背景

地球環境観測データは、長期にわたって観測を持続することでその価値を飛躍的に増大させるものであるにも関わらず、取り扱う

データが一般的に非常に多種多様かつ容量が膨大であるため、十分に活用されることなく死蔵・破棄されるケースも少なくない。さらに、個々のデータを保有する各研究機関からは、インターネットなどを活用した容易な

相互参照システムの実現を求める要求は高いが、実現できていない。観測技術の急速な進歩により、高分解能データ化が進み、データ量が TByte, PByte へと爆発的に増大することが確実な今、必要な現象を示すデータをテキストデータにおける全文キーワード検索のように取り出す検索形態、そしてネットワークを利用したデータの相互参照・配信が行えるシステムが必要である。

我々は従来から、次に示す(1)(2)の研究開発を行っており、これらを発展させ、さらに組み合わせることにより、前述の検索形態およびデータの相互参照・配信形態を実現するという着想に至った。

(1) 従来の Web-DB 型検索システムは、データの構造化が不十分であり、データの再利用が困難である。データの解析を行うには検索結果を閲覧するだけでなく、その結果を容易に再利用できなければ、検索システムを使用するメリットが低い。これに対して、我々は XML/Web サービスを用いたデータ相互参照・配信を提案していた。また、非常に多種多様なデータを相互に理解する必要があるが、これに対して、XML/Web サービスに加え、自己記述型汎用データフォーマットと連携させる研究を行っていた。XML/Web サービスは分散プログラミング環境技術であり、自己記述型データフォーマットは地球環境観測データを相互に利用する場合に必要な技術である。我々の研究グループはこれらの技術を連携させるために必要な知見を有していた。

(2) 膨大なデータから、現象を選別し、興味ある現象だけを取り出せる必要があるが、我々は現象を自動分類するためのシステムを研究開発してきた。必要な現象を示すデータを取り出す技術は、様々な分野で研究が進められているが、地球環境観測データは、特性が既知の現象だけでなく、「未知」、「発見的な」特異現象が含まれる自然現象であるため、従来の方法では一部の既知現象の取得に留まる。これに対して、申請者らは、既知現象に加え、「未知」、あるいは、「発見的な」特異現象を取得するため、現象を統計学的に分類することで、代表的な現象の抽出を行い、さらに特異現象も抽出する手法を開発した。このような「発見科学」的な考え方は、遺伝子研究などの分野に浸透しつつあるが、地球環境分野を対象とした実用研究では、地球物理と電子情報工学双方の専門知識が要求されるため、両分野にまたがる我々のこれまでの取り組みが本研究課題の解決に必要な不可欠であった。

## 2. 研究の目的

当研究課題の目的は、地球規模の全体像を

研究し、地球環境を理解するための総合的な解析支援環境の実現を可能とするため、大規模地球環境観測データを自由に検索参照できる現象アーカイブシステムを構築することである。

具体的には、地球環境分野の中でもフォーマットが非常に多種多様かつ複雑である地球周辺電磁波環境観測データを対象にして、アーカイブシステム実現に必要な要素技術の研究開発を行う。以下の課題を設定した。

- (1) T (テラ) バイトオーダーの大規模地球環境観測データベースを構築し、その手法を確立する。
- (2) 世界中に分散して蓄積されているデータベースコンテンツを相互参照・配信するアクセスインターフェイスを開発する。
- (3) データベースから、テキストデータにおけるキーワード検索のように、特徴的な現象を検索する方式を開発する。

そして、(1)～(3)で確立した手法により、地球周辺電磁波環境観測データが自由に検索・参照できる現象アーカイブシステムの構築を目指す。

## 3. 研究の方法

研究目的で示した課題ごとに研究方法を述べる。

- (1) T (テラ) バイトオーダーの大規模地球環境観測データベースの構築

従来の研究成果より、特徴分析が進んでいるあけぼの衛星による VLF 帯電磁波動スペクトルデータに焦点を絞り、8 秒平均値をサンプルデータとして登録した(2.1GByte)。最終的には全観測データ(15TByte)を取扱うため、できる限り容量を抑え、さらに、観測データの特性を生かした上でデータの互換性を確保できるように、自己記述型汎用データフォーマット (CDF) を導入した。

観測データは膨大なため、巨大データベースとなる。開発にあたっては、できる限りデータベースの容量を抑え、さらに、検索の高速化を図るため、データの特性を考慮した圧縮登録法、高速索引テーブルの検討を行った。

- (2) 分散して蓄積されているコンテンツを相互参照・配信するアクセスインターフェイスの開発

### ①共通プラットフォームの提案

実験観測データおよび研究資料を公開するための公開用システムの共通機構化を図った。具体的には、実験観測データおよび研

研究資料等を公開するための指針となるべく、既存情報インフラ上で運用可能な公開手法を考案し、公開モデルを作成した。考案した公開手法を具象化し、共通プラットフォームとして提案した。共通プラットフォームとは、公開手法を汎用的に実装したシステムで、そのままの形で、あるいは、簡単な設定や簡単な追加実装で再利用できる、「使い回し」が可能なシステムである。共通プラットフォーム化することで、管理・公開のための手法の統一化が図られ、システムの再利用が可能となるので、公開までの準備期間の短縮、共通の運用保守体制の確立など、効率的な運用が期待できる。

## ②プロトタイプの開発

作成した公開モデルに基づき、次に示す3種類の共通プラットフォームのプロトタイプを作成した。

- 研究資料等のコンテンツ公開のための共通プラットフォーム（データリポジトリ）。
- Web-DB システムの一元的な管理・公開を可能とする共通プラットフォーム（多様なアクセス制限に対応可能なWeb-DB 管理システム）。
- 実験・観測データを配信するための共通プラットフォーム（汎用データフォーマットを用いたデータ配信システム）。

## ③公開システムの構築（実データへの適用）

共通プラットフォームを用いた実データを公開するための公開システムを構築し、学内外に公開し実証を行った。

## ④組織間での認証の連携

共通プラットフォームを利用した公開システムをはじめとする各種公開システムの認証機構の共通化（共通認証方式化）を図るために、Shibbolethによるシングルサインオン実証システムの研究開発を行った。

実装した一部の公開システムに、共通認証方式を組み合わせ、複数の大学間でデータの公開・共有が可能なシステムを実装し実証運用を行った。

## （3）特徴的な現象を検索する方式の開発

### ①現象の抽象的な特徴の定量化に基づく類似現象検索法の開発

観測された各種自然現象の抽象的な特徴を定量化し、検索キーとして類似現象をデータベースから抽出する方法を検討した。前述のCDFを適用したあけぼの衛星のVLFデータに対し、研究者が解釈に用いる時間・周波

数変化などの物理量の特微量データベースとして利用し、ユーザが指定した観測データと類似データを検索・閲覧できる手法を提案した。

さらに、膨大なデータに対し、同手法が適用できるように、検索アルゴリズムの高速化を図った。これらの検討に基づき、Webブラウザで利用できる類似現象検索システムを実装し、実証運用を行った。

## ②システムの一般化

抽象的特徴表現法の評価と類似現象の抽出のための分類指標の検討を行った。また、抽出アルゴリズムの一般化を図り、多様な観測データに対応できる汎用的なシステムに改良した。

さらに、複数種の観測データを登録し、改良した類似現象検索システムの実践的な評価を実施した。

## 4. 研究成果

本研究開発では、地球環境観測データ活用のための相互参照可能な現象アーカイブシステムの構築を目指し、その実現に必要な要素技術の研究開発を行った。

T（テラ）バイトオーダーの大規模地球環境観測データベースの構築では、我々は、自己記述型汎用データフォーマット（CDF）を導入したが、この方式は、月探査衛星「かぐや」の観測データにも適用され、その有用性を実証した。また、今回開発した類似現象検索システムは、月探査衛星「かぐや」の観測データにも適用され、その有用性を実証した。

提案した共通プラットフォームは、現在、複数のデータに適用し、図1に示す地球環境データベースシステムや図2に示すデータリポジトリという形で公開している。

今後は、今回の研究成果を発展させ、大規模地球環境観測データにおける自然現象の抽出・目録化と現象データベースの構築、および、組織間でデータを安全に公開/共有でき



図1 地球環境データベースシステム

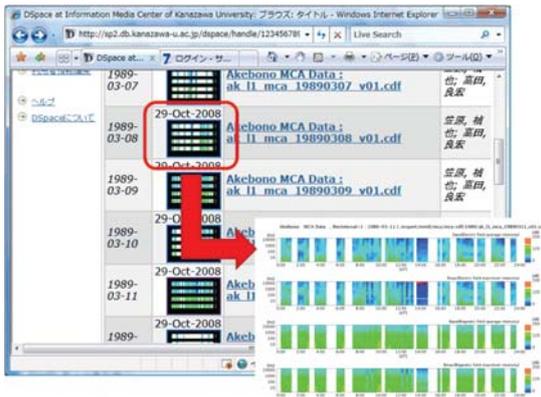


図2 データリポジトリ

る、組織間連携可能な分散型現象アーカイブシステムを実現する。そして、地球規模の全体像を研究し、地球環境の理解につながる総合的な解析支援のための e-サイエンス基盤の実現を目指したい。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 9 件)

- ① 高田 良宏, 笠原 禎也, 研究資料公開に向けての共通プラットフォーム構築の試み, 学術情報処理研究, 査読無, No.14, 214(2010), 153-157.
- ② Takuya Matsuhira, Yoshiya Kasahara, Yoshihiro Takata, Development of a file-sharing system for educational collaboration among higher-education institutions, *INTERNATIONAL JOURNAL OF EDUCATION AND INFORMATION TECHNOLOGIES*, 査読有, Issue 2, Vol.5, 359(2010), 149-156.
- ③ Y. Kasahara, A. Hirano and Y. Takata, Similar Data Retrieval from Enormous Datasets on ELF/VLF Wave Spectrum Observed by Akebono, *Data Science Journal*, 査読有, Vol.8, (2010), IGY66-IGY75, (<http://hdl.handle.net/2297/25293>).
- ④ Y. Takata, Y. Kasahara, and T. Matsuhira, Development of a Science Database System Applicable to Various Access Restrictions, *Data Science Journal*, 査読有, Vol.8, (2010), IGY32-IGY43, (<http://hdl.handle.net/2297/20470>).
- ⑤ 高田 良宏, 笠原 禎也, 西澤 滋人, 森雅 秀, 内島 秀樹, 非文献コンテンツのための可視性と保守性に優れた学術情報リポジトリの構築, 情報知識学会誌, 査読有, Vol.19, 58(2009), 251-263,

(<http://hdl.handle.net/2297/19715>).

- ⑥ 松平 拓也, 笠原 禎也, 高田 良宏, 井町 智彦, UPKI 認証連携基盤に基づく安全なデータ共有システム構築の試み, 学術情報処理研究, 査読有, 176(2009), 84-90.
- ⑦ 高田 良宏, 笠原 禎也, 西澤 滋人, 森雅 秀, Google Earth による学術情報の可視性向上, 学術情報処理研究, 査読無, No.12, 158(2008), 94-97.

[学会発表] (計 10 件)

- ① 松平 拓也, Shibboleth をベースとした異領域間でのデータ相互参照システムの開発, 情報知識学会 第 18 回(2010 年度)年次大会, 2010 年 5 月 24 日, 幕張メッセ国際会議場(千葉県).
- ② 高田 良宏, 大学における非文献コンテンツ公開のための共通プラットフォームの開発 ~非文献コンテンツのための可視性と保守性に優れた学術情報リポジトリの構築~, 情報知識学会 第 18 回(2010 年度)年次大会, 2010 年 5 月 16 日, 東京大学本郷キャンパス(東京都).
- ③ 松平 拓也, 安全・安心な大学間情報共有を実現するサービスプロバイダの構築と今後の展開, UPKI シンポジウム 2010, 2010 年 3 月 12 日, 学術総合センター一橋記念講堂(東京都).
- ④ 高田 良宏, 研究室に埋蔵されている研究・教育資料公開の試み~非文献コンテンツを対象とした学術情報リポジトリの構築~, DRF-KanNihonkai, 2010 年 2 月 17 日, 金沢大学角間キャンパス(石川県).
- ⑤ Y. Kasahara, Similar data retrieval from enormous dataset on plasma wave spectrum observed by Akebono satellite, International Symposium: Fifty Years after IGY - Modern Information Technologies and Earth and Solar Sciences, 2008.11.11, AIST Tsukuba (Ibaraki).
- ⑥ Y. Takata, Development of Science Database System Applicable to Various Access Restrictions, International Symposium: Fifty Years after IGY - Modern Information Technologies and Earth and Solar Sciences, 2008.11.11, AIST Tsukuba (Ibaraki).
- ⑦ 高田 良宏, デジタルコンテンツに適した学術情報リポジトリの構築, FIT2008 (第 7 回情報科学技術フォーラム), 2008 年 9 月 2 日, 慶應義塾大学湘南藤沢キャンパス(神奈川県).
- ⑧ 平野 晃, 大容量事前波動観測データからの類似現象検索システム, 地球惑星科

学連合2008年大会, 2008年5月26日, 幕  
張メッセ国際会議場 (千葉県) .

[その他]

ホームページ等

[https://wwfdb01.db.kanazawa-u.ac.jp/DB/  
pub/db/](https://wwfdb01.db.kanazawa-u.ac.jp/DB/pub/db/)

<https://sp2.db.kanazawa-u.ac.jp/dspace/>

## 6. 研究組織

### (1) 研究代表者

高田 良宏 (TAKATA YOSHIHIRO)

金沢大学・総合メディア基盤センター・助  
教

研究者番号 : 30251911

### (2) 研究分担者

笠原 禎也 (KASAHARA YOSHIYA)

金沢大学・総合メディア基盤センター・教  
授

研究者番号 : 50243051