

## サブテーマ 2 : 地球・生命などの巨大システム解明のための 統合的情報基盤の形成と活用手法の確立

### 研究代表者

[国立情報学研究所] 藤山 秋佐夫

### 共同研究者

[国立情報学研究所] 武田 英明、市瀬 龍太郎、北本 朝展、佐藤 真一、孟 洋

[国立遺伝学研究所] 菅原 秀明、城石 俊彦

[国立極地研究所] 神田 啓史、野木 義史、土井 浩一郎、宮岡 宏、本山 秀明

[統計数理研究所] 長谷川 政美、樋口 知之、上野 玄太、尾形 良彦、村田 泰章、  
種村 正美、石黒 真木夫、島谷 健一郎

[東京理科大学] 宮崎 智

[Johns Hopkins 大学] 大谷 晋一

### 1 . 研究目標

本研究項目は、地球・生命システム解明に向けた情報・統計統合処理と意味抽出の情報基盤形成、地球環境データの情報空間統合に基づく横断型情報基盤とポータルサイトの構築、の二つの研究項目を中心に取り組む。

生命・地球システムの解明に向けて、生命と地球環境との相互作用にまで視野を広げて、生命体全体のコントロールシステムの全貌を理解することを目指す。

具体的には、まず、ゲノム上の遺伝子情報と、これに基づいて具体化された生命体の関連性を明らかにすることが重要となる。しかし、具体的な観測・計測の対象となる生命体の表現型(形態や行動等)は、遺伝要因と環境要因との相互作用の結果として生じるものであり、本質的に多様性を持つという問題が存在する。したがって、遺伝子の個別的な機能をさまざまな生命現象との関連性を明らかにすることが必要であり、計測対象の要素選定、計測値の統計処理、情報の統合化と意味抽出などの地球・生命・情報の融合的統計情報処理手法の開発による異分野共通の研究基盤を構築する。このため、実験動物(マウス)を通じてのヒトの特異性解明、および、地球環境と生命システムの総合作用という2つの観点から研究を進める。また、国立情報学研究所と国立遺伝学研究所が16年度に共同研究開発したバイオポータルの枠組みを利用して、成果として得られた情報を公開、活用するための情報発信ツールの開発を行う。

また、地球システムの総合的な理解のために必要な地球観測データの統合システムを構築する。具体的には、地球システムの層(圏など)別に分散している多くの観測データをつなぎ合わせ、これらを鉛直方向に統合する方法を提案する。地球観測データ統合システムを構築するための情報統合モデルとして、地理情報システム(GIS: Geographic Information Systems)におけるレイヤを基本単位とした構造、および、鉛直方向に異種データを重ね合わせ、これらを縦断

した検索や可視化する概念を参照する。また、地理情報システムで利用されているWeb Map Service (WMS)等のウェブサービスの利用を図る。このように、レイヤを用いた鉛直的統合と、ウェブサービスを用いた水平的統合との組み合わせた地球観測データ統合システムの構想の実現を目指す。

具体的には、まず、国立情報学研究所が保有する気象衛星データや気象観測データ、台風経路データなど、主に対流圏に関する地球観測データを対象とした研究を進める。次に、国立極地研究所が保有する南極地質データなど、主に地圏に関する地球観測データ、および、統計数理研究所が保有する地球環境データなどを取り扱う。さらに、データベース技術やデータマイニング技術、信号処理技術によって、多様な視点での分析や可視化が可能な地球環境の総合的研究環境を構築することを目標とする。

## 2 . 年次研究計画

研究項目 では、ライフサイエンスメタデータベースシステム構築と極限環境生物統合データベースの構築を具体的課題として取り組んでいる。

ライフサイエンスメタデータベースシステム構築研究に関しては、ライフサイエンス研究全体の効率化を目指した利用者主導型メタデータ DB システムの構築と運用が第一の目的である。このため、18年度には、データ収集を精力的に進めると共に、レコードの整理を進め、データベースの所在情報に加えて、機能情報とその評価とを連結させる。その過程においては、さまざまな分野の研究者の参画が必要になるため、共同研究者グループを中核とした評価体制を確立する。また、日本語バイオポータル(<http://www.biportal.jp>)からの公開・運用システムについての検討を行う。

極限環境生物統合データベースの構築研究では17年度に続き、統合データベース構築のために必要な基礎データの一層の充実を図る。精度の高いデータ解析に必要な量のデータが集積した時点で、画像と遺伝子情報を基に、極域限定種と汎地球種についての地球規模での多様性分布の解明を目指す。

研究項目 では、平成18年度は、ウェブサービスを用いて国立情報学研究所と国立極地研究所のシステムを接続するとともに、「鉛直地球ポータル」および「南極GISポータルサイト」を一般向けに利用可能な形で公開する。また、気象シミュレータを稼働させ、主にメソスケールの気象シミュレーションを行い、「鉛直地球ポータル」における蓄積・処理データの定常的な生成を行う。地球環境情報をさまざまなフォーマットに変換して、地球情報を効果的にユーザに提示するための多様な情報ブラウザの開発を目指す。オープンソースの地理情報システム (MapServer) を導入し、地球環境情報を表示するためのシステムを開発する。紙地図データのデジタル化および座標系の統一化を引き続き進める。

平成19年度以降については、18年度の進捗状況と成果を検討し、必要な組織の再編と計画の調整を行い、これらの検討結果に基づいて、最終目的実現のための研究の推進を行う。

### 3. 平成17年度の研究進捗

分野横断型融合研究を実効的に推進するための情報空間・情報基盤を構築向けて、主に地球・生命システム解明に向けた情報・統計統合処理と意味抽出の情報基盤形成とポータルサイトの構築を目指して、以下の研究項目を推進した。

ライフサイエンスメタデータベースシステム構築研究では、ライフサイエンス研究全体の効率化を目指した利用者主導型メタデータ DB システム構築を、「新世代バイオポータル」プロジェクトの成果を発展させる形で推進している。平成17年度は、データ量（約29,000件のレコードを収集済）の増大、多様化に継続的に対処可能なシステムの基本設計、および、これらに必要なハードウェアの整備を実施した。また、収集済レコードの整理を進め、データベースの所在情報と機能情報を連結させる目的で約40件について精査し、新たにPubMedとのリンクを設定して日本語バイオポータル(<http://www.bioportal.jp>)から公開した。

極限環境生物統合データベースの構築研究では、国立極地研究所が南極大陸で採集した生物標本を対象に、高精度3次元画像、ゲノム情報、分子進化情報等を基礎情報に加えて統合データベース化し、極地を中心とした生物多様性についての新しい研究基盤の構築を進めている。平成17年度は画像処理技術の開発を進め、被写体を中心とする半球を画像上で3次元表示させるためのソフトウェア開発を行うと共に、撮影用機材の整備と開発を実施した。また、ゲノムと分子進化解析に関しては、極域限定種と、極地および世界各地で採集した汎地球種について葉緑体のゲノム解析を行うこととし、約6,000件のコケ葉緑体 DNA 配列情報の整理を行った。

研究項目では、まず、地球観測データを鉛直的に統合することが、地球システムを総合的に理解するために必要であるという視点に立ち、地球環境ポータルの構想を「鉛直地球ポータル (Vertical Earth)」として具体化した。従来より地球システムの層構造（圏など）毎に蓄積されてきた地球観測データについて、複数の層をまたいだ重ね合わせや統合の実現を目指す。また、地理的に分散した機関をまたいだデータ統合を実現するためにウェブサービスを用いる。具体的には、地理情報システム(GIS)において広く使われつつある Web Map Service (WMS)等のウェブサービスを利用する。この基礎調査、および、利用のための試験的なアプリケーションの開発を行った。

また、地球環境データの一つとして気象シミュレーションデータのための基盤システムを整備した。鉛直データ統合の実験とデモンストレーションに適したデータを生成する気象シミュレーションのためのLinuxクラスタ購入、および、気象シミュレータに関する基礎調査を行った。その結果、気象シミュレータとして著名な米国のMM5、その後継プロジェクトであるWRF、あるいは気象庁が開発する非静力学モデルNHMなどのソフトウェアをインストールして利用することに決定した。

さらに、地理情報システム(GIS)の導入を図り、そのフロントエンド(クライアント)とバックエンド(データベース)を整備することにより、南極に関する地理情報を整備・閲覧するための環境を構築した。構築した地理情報のウェブブラウザ閲覧を可能にするため、ポータルサイト(専用ウェブサイト)を立ち上げた。平成17年8月30日に国立極地研究所において、南極

GIS ポータルサイト利用講習会を開催し、同日から9月22日にかけてテスト公開を実施した。さらに、この期間に実際にポータルサイトを利用した所員からアンケートも回収し、今後のシステム構築に役立つための意見を収集した。地図データの整備に関しても、ADD(Antarctic Digital Database)やラスタ化された地形図を取り込んだほか、紙の地形図の一部もデジタル化した。その結果、地形図と地質図や航空写真、衛星写真などを重ね合わせて表示することが可能となった。

#### 4. 平成17年度研究成果

##### (1) 知見・成果物・知的財産権等

- 日本語バイオポータル バイオデータベースショウケース  
<http://www.ps.noda.tus.ac.jp/biometadb/db-j/index>  
(上記アドレス、名称については変更の可能性有り)
- 地球観測データの提供
  - 台風中心位置を中心に切り取った台風画像(過去11年分、約64000件)
  - 台風経路データ(ベストトラックデータ(過去55年分、約1470件))
  - アメダス気象観測データ(過去30年分、約3億2000万件)
  - 台風関連の気象災害データ(過去34年分、約3600件)
  - 台風関連のニュース記事(過去3年分、約7000件)
  - ウェブログからのトラックバック(過去2年分、約400件)
- 過去に作成された地図をデジタル化と南極GISへの格納  
1/25000の南極DEM(デジタル標高モデル)5mを約200枚デジタル化し、地理情報システム上で表示・検索可能。

##### (2) 成果発表及び著書執筆等

- 藤山秋佐夫:「生命の秘密に挑戦するゲノムインフォマティクス~人間とチンパンジーの違いを読み解く」、情報通信ジャーナル, 24, 30-31 (2006)
- 藤山秋佐夫、菅原秀明ら:「ごぞんじですか?日本語バイオポータルサイト」, 図書館情報, 印刷中, (2006)
- 藤山秋佐夫:「科学コミュニケーションと日本語バイオポータル」 国立遺伝学研究所共同研究会「生物情報資源の相互運用性」, 2006.3.6
- 北本 朝展: "デジタル台風:異種データストリームを融合した準リアルタイム台風データベース", 画像の認識・理解シンポジウム (MIRU 2005), pp. DE-14, 2005年07月
- 北本 朝展: "デジタル台風 - 網羅的な台風データベースはどう「使える」のか? -", 東京工芸大学「21世紀COEプログラム」オープンセミナー, No.37, 2005年11月(招待講演)
- 北本 朝展: "Vertical Earth: 地球の鉛直構造をまたぐデータベース/ポータルサイトへ

の構想", 「宇宙地球系情報科学研究会」「巨大データベース構築に関する研究集会」合同研究会, 2005年12月(招待講演)

- 北本 朝展: "地理情報資源のデータベースと相互運用の試み", 国立遺伝学研究所研究会「生物情報資源の相互運用性」, 2006年03月
- 北本 朝展: "デジタル台風:地上と宇宙の多様なデータを網羅的に統合した台風データベース", 第15回生研フォーラム「宇宙からの地球環境モニタリング」, pp. 49-50, 2006年03月
- Asanobu KITAMOTO: "Digital Typhoon: Toward the Hub of Typhoon Emergency Information Aggregated from Various Sources", NII International Workshop on Cyber Science Infrastructure, 2006年03月 (in English)

## 5. その他

- 藤山秋佐夫:「科学コミュニケーションと日本語バイオポータル」 国立遺伝学研究所共同研究会「生物情報資源の相互運用性」, 2006.3.6(招待講演)
- 「鉛直地球ポータル(Vertical Earth)」プロジェクト <http://earth.nii.ac.jp/>

## 〔参考文献〕

- 日本語バイオポータル <http://www.biportal.jp>
- 国立極地研究所生物資料  
[http://www.nipr.ac.jp/japan/facilities/t03\\_biodata.html](http://www.nipr.ac.jp/japan/facilities/t03_biodata.html)
- Open Geospatial Consortium, <http://www.opengeospatial.org/>
- Web Mapping Illustrated, T. Mitchell, O Reilly, 2005.