

地球研で実施されたプロジェクト群およびデータ

Data Management System at Research Institute for Humanity and Nature (RIHN)

関野 樹^{1*}

Tatsuki Sekino^{1*}

¹総合地球環境学研究所

¹RIHN

総合地球環境学研究所（地球研）は人間文化研究機構に属する大学共同利用機関の1つである。その目的は、環境問題を自然科学的な現象面だけでなく人間文化の問題として捉えることにより問題の根源に迫ることであり、現在14の研究プロジェクトが分野の壁を越えた学際的な手法により各々の研究課題に取り込んでいる。このような研究上の特色により、地球研には様々な学問的背景に基づく多様なデータが集まってくる。これには、気象、水文、水質、生物相や遺伝子情報などの自然科学的なデータだけでなく、アンケートなどの社会調査の結果、現地で収集した統計資料、文化人類学者が現地住民から聞き取った記録や歴史学者が集めた文書資料の目録といった人文社会学的な情報も含まれる。このため、データマネジメントという点においては、多様なデータをどのように蓄積し、それらをどのように活用するかという2つの課題が存在している。

（1）多様なデータをどのように蓄積するか

地球研の場合は、集めるデータが多様であるということに加えて、研究者の任期制に伴う問題がある。プロジェクト終了後は関連する研究者が地球研を出てしまい、誰が何の目的でとったデータなのか分からなくなるという事態に陥りやすい。

現在の地球研では、データを残すという点で「地球研アーカイブス」という、種類を問わずどのようなデータでも入れられる「箱」が用意されている。この「地球研アーカイブス」には、プロジェクトの履歴、メンバリスト、業績といったプロジェクトそのものに関する情報も収容されており、データはこれらのプロジェクトの情報と関連付けられて保管される。つまり、検索というよりは、データの来歴を残し、データが散逸すること防ぐことを主眼においたシステムである。現在は報告書や出版物などの2次資料の収容が中心であるが、観測データを含む1次資料の収容も順次進めてゆく予定である。

（2）多様なデータをどのように活用するか

地球研に蓄積されたデータを他の機関が持つデータと連携させる情報基盤の構築が進められている。これにより、データの利用範囲の拡大だけでなく、データ間の新たな関連性の発見が促進されることも期待されている。具体例として、地球研を中心に進められている「地域環境情報ネットワーク事業」では、各研究分野で標準的に用いられるメタデータ同士を関連付けることにより、分野間でデータを連携させる試みが行われている。例えば、地球研内に設置されたEML

(Ecological Metadata Language) による自然科学系のデータベースと京都大学地域研究統合情報センターに設置されたMODS (Metadata Object Description Schema) による人文社会系のデータベースを検索プロトコルの1つであるSRW/U (Search/Retrieve via the Web or URL) を用いて接続することで、両機関の資料の横断検索を実現した。さらに、同様のインターフェースを持つ人間文化研究機構の統合検索システム(研究資源共有化事業)とこれら接続することによ

り、EMLによるデータベースを博物館や資料館の100あまりのデータベースと連携させることにも成功している。このほかにも、図書館で用いられるMARC (MAchine-Readable Cataloguing) と地理情報科学 (GIS) の分野で用いられているISO-19115を連携させることにより、図書館が所蔵する近代以前の絵地図とGISで用いられる電子化された地図や衛星画像をシームレスに検索する仕組みの構築も進められている。

システム間の連携と同時にデータを利用するための基盤整備も人間文化研究機構の研究資源共有化事業を通じて行われている。例えば、主題間の関係を記したオントロジー辞書や地名と緯度経度の関係を記述した地名辞書などの基盤情報を整備することにより、これらをデータの表現方法の正規化や専門外の分野のデータを検索する際の検索語入力に活用することが可能になる。また、多様な情報の関連性を時間軸上で視覚的に解析するソフトウェアやトピックマップを使った主題間の関連性の視覚化などの研究開発も行われている。

キーワード:学際研究,横断検索,メタデータ,オントロジー

Keywords: interdisciplinary study, cross search, metadata, ontology