

「地域の知」の情報技術

関野 樹¹、山田 太造²、大向 一輝³、原正一郎⁴

¹総合地球環境学研究所, ²東京大学, ³国立情報学研究所, ⁴京都大学

地域研究は特定の地域を理解する知的体系として展開されてきた学際的研究分野である。しかし、グローバル化の急速な拡大と進展に伴い、地域や国境を越えて解決しなければならない課題が山積しつつある。このような中で、地域研究はグローバルとローカルをリンクしながら地域をデザインする学問への脱皮を迫られている。つまり地域研究は、諸課題を、個人・家族・集落・都市・国家・世界などのさまざまなサイズにおける、自然環境・歴史・文化・社会経済状況などの多様な要素の関わり合いとして理解し、解決を図る学問体系へと拡張される必要がある。そのため、地域研究に関わる多様な学問領域の情報を収集・共有・分析する情報学的パラダイムの確立が焦眉の急となっている。地域研究に関わる情報とは、伝承・習慣・技術など地域住民の振る舞いを規定する顕在的・暗黙的な知識、関連する資料・文献、研究成果として得られた発見・仮説などである。これらの情報と（情報学的）分析手法をあわせて「地域の知」と呼んでいる。また「地域の知」に関わる情報学的パラダイムを地域情報学と呼んでいる。

本企画セッションでは「地域の知」を展開するうえで期待される情報技術について議論する。地域研究を対象とするが、地域研究は本来的に学際的研究分野である。また「地域の知」は、テキスト・画像・動画・音声などさまざまなメディア形式をとり、かつ世界中のデータベースに分散している。つまり、「地域の知」の情報技術は、人文科学とコンピュータ研究会が対象としている人文科学分野へのコンピュータ応用の縮図とみなすことができ、したがって本セッションで議論される情報技術は、人文科学分野への汎用化が可能であると期待している。

本セッションでは 4 件の発表を予定している。最初に、地域情報学を具体的に展開している京都大学地域研究統合情報センターを事例として、データベースと資源共有化システムの実装、およびデータの可視化や時空間分析などのデータ利用について報告した後、現在および将来に想定される課題を解決する情報技術への期待について述べる。この報告を受けて、セマンティック Web と知識共有、テキストマイニングおよび時空間分析に関する情報技術の発表を行う。

セマンティック Web と知識共有

Web 上に分散している情報の共有化は「地域の知」を実現する第一歩である。しかし、地域研究統合情報センターや人間文化研究機構で稼働している資源共有化システムは Web 技術に立脚したデータ構造となっていないため、データの自在な追加・削除やコンテンツに注目した共有化は容易ではない。そこで従来の資源共有化システムに代わる基盤技術としてセマンティック Web が期待されている。本セッションでは、セマンティック Web の基礎技術と応用事例について発表し、次世代資源共有化システムへの応用可能性について議論する。

テキストマイニング

現在の資源共有化システムがリンクしている情報の粒度は資料レベルである。しかし、1990 年後半に資源共有化システムの議論が開始されたとき、リンクの対象は主題・地名・人名など語彙のレベルであった。ところが当時は、Web リソースを記述・連携・検索する技術的枠組みが未発達であったことと全文データベースが殆ど無かったため、資源共有化システムのリンクのレベルは資料となった。現在、全文データベースの数は増えつつあり、さらに地域研究情報の多くが Web を通じて発信されるようになってきた。これらのデータ

を「地域の知」として分析するためには、データを語彙レベルに分解し、さらに関連語彙をリンクする必要がある。その基礎技術としてテキストマイニングが期待されている。本セッションでは、テキストから特徴語を抽出し、テキストに潜在する話題に応じて特徴語を関連性に基づき分類することにより、特徴語間だけでなく特徴語-テキスト間、さらにテキスト間を関係づけていく手法について発表し、特にフィールドノートを対象とした応用について議論する。

時空間分析と可視化

時空間属性は、多様な情報を関連付けたり分析したりするための重要な計量的データであることから、地域研究情報の資源共有化システムにおける検索機能の一部として利用されている。また、時空間属性に基づいて可視化や分析を行う **HuTime / HuMap** を中心とした時空間情報解析ツールの開発とその利用について研究が推進されており、さらに、時空間属性を持つデータ構築や情報解析に必須となる基盤情報、つまり地名やベースマップなどの空間基盤情報、暦や基本的な出来事を記した年表などの時間基盤情報の構築も進められている。本セッションでは、地域情報学における時空間属性を使った情報処理の現状や問題点について報告するとともに、時間と空間の属性を相互に活用する統合分析ツールの開発、時空間情報解析機能の Web サービス化などについて議論する。

本セッションで取り上げた情報技術は「地域の知」の情報学パラダイムを構成する一部に過ぎないが、地域研究統合情報センターで議論や実装が進められている事例であり、具体的な議論が展開できる。また地域研究を対象としているが、人文科学分野への汎用化が可能な情報技術である。多くの参加者と活発な議論を期待する。