

地域情報学の構築と時空間情報解析ツール

原 正一郎

京都大学地域研究統合情報センター

柴山 守

京都大学東南アジア研究所

地域研究は歴史・文学・宗教・政治・経済学・民族学・民俗学・農業・環境などの学問分野を包摂する学際的研究分野の1つであり、特定地域の総合的な理解を目的としている。しかし地域研究者は自らの専門分野に立脚した特定地域の分析や解明を主たる研究テーマとしており、専門領域を超えた学際的研究の実現は容易ではない。地域情報学は、地域に関する多様なデータや情報を統合して地域の解明や地域に関する研究知の創出を支援する、情報学ないしは情報工学分野からの研究パラダイムである。本論文では地域情報学の概要と、その基盤をなす時空間情報処理技術の開発状況について述べる。

Development of Area Informatics and Spatiotemporal Analysis Tools

Shoichiro HARA

Center for Integrated Area Studies

Kyoto University

shara@cias.kyoto-u.ac.jp

Mamoru SHIBAYAMA

Center for Southeast Asian Studies

Kyoto University

sibayama@cseas.kyoto-u.ac.jp

Area Study is one of the interdisciplinary sciences consisting of many disciplines such as history, literature, religions, politics, economics, ethnology, folklore, agriculture, environmental sciences, and so on, and aims to understand specified areas comprehensively. However, each researcher engages in a specific area and discipline, and it is difficult to achieve interdisciplinary study. “Area Informatics” is the new paradigm from information science/technology field to integrate various data/information on area studies and to support researchers in finding new knowledge on areas. This paper will describe the notion of Area Informatics and present status of spatiotemporal information processing technologies that we have been developed.

1. まえがき

地域研究 (Area Studies) は多様な地域 (アジア、アフリカ、オセアニアなど) を歴史・文学・宗教・政治・経済学・民族学・民俗学・農業・環境などの専門領域の立場から解明し理解する学際的研究分野の1つであり、特定地域の社会・文化構造や特質を人間活動とともに様々な視点から総合的に理解することを目的としている。しかし地域により言語が異なる、学問領域により研究手法や学術語彙が異なるなどの理由により、多くの地域研究者は特定の地域と主題に特化した研究を展開しており、専門分野を超えた学際的研究の実現は容易ではない。一方グローバル化の進展のもと、地域間の比較や地域横断的な課題設定による地域研究が求められている。これを地域相関型研究と呼んでいる[1]。地域情報学 (Area Informatics) は、特定の地域や地域相関研究に対する情報学ないしは情報工学分野からのアプローチであり、地域および地域研究に関する多様なデータ・情報の統合と地域の解明と理解および地域研究知の創造支援を目的とした新しい研究パラダイムである[2,3]。

本稿では、最初に地域情報学の概要と基盤をなす情報処理技術について概観する。特にデータベースと時空間属性の重要性について述べる。続いて筆者らの HGIS (Humanities GIS [4]) 研究グループが開発中の時空間情報解析ツールから、空間に注目した情報処理システム (地図システム)、時間に注目した情報処理システム (年表システム)、およびこれら

を利用する上で必要不可欠なオントロジー辞書 (地名辞書および暦日テーブル) について説明する。最後に、これらの時空間情報解析ツールを利用するためのデータ作成法、特に歴史史料を利用した事例について述べる。

2. 地域情報学の概要

地域研究のプロセスを、データ・情報・知識¹のフローという視点からモデル化したものを図1に示す。地域研究は、①フィールドにおけるインタビューや資料収集などに始まり、②収集されたデータ・資料はデータモデルの検討を通じて、③個別データベースに組織化され、最終的には資源共有化システム (2.3 参照) に統合される。次に、④オントロジー辞書を用いた検索を資源共有化システムに適用して関連性の高い多様なデータ・情報を抽出し、⑤これらに GPS (Global Positioning System: 全地球測位システム) や GIS (GIS: Geographic Information System: 地理情報システム) に代表される時空間分析や統計分析などを基礎としたデータマイニング (data mining) を適用し、仮説の生成や知識の発見を導く。これらの成果は再びフィールドへフィードバックされ、新しい研究プロセスが始まる。地域研究におけ

¹ 本稿ではデータ・情報・知識を厳密には区別していない。大まかには、研究対象から収集されてあまり加工されていない比較的「生」の数値や記述などをデータ、データに統計処理や類別化などの知的作業を加えた結果を情報、データや情報の間の構造・関連あるいは利用法を知識としている。

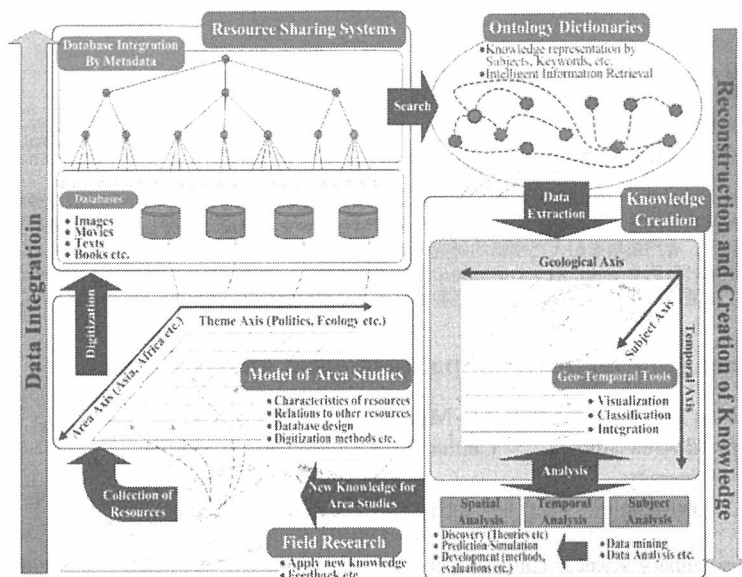


図 1. 地域情報学の概要

るこのようなデータ・情報・知識の循環を支援する情報工学的体系が、地域情報学モデルの一つとして考えられる。このモデルでは、図 1 に示す①フィールド支援、②地域情報モデル構築、③資源共有化システム、④オントロジー辞書、⑤知識生成支援の各プロセスに適した様々な情報処理技術の研究・開発が要件となる。本稿では、この立場に立ち、以降の議論を展開する。

ところで、これらの情報処理技術には「地域研究からのアプローチ」と「情報工学からのアプロ

チ」の二つの側面が考えられる。地域研究からのアプローチとは、地域研究において従来から採用されている手法に情報学や情報技術を援用し、その高度化を目指すものである。これに対して情報工学からのアプローチとは、必ずしも地域研究には馴染みのない、あるいは最新の情報工学的手法を地域研究に適用し、地域の解明や地域研究知の創造を目指すものである。図 1 には両者が含まれているが、筆者らの主たる関心は後者にある。

2.1 フィールド支援

フィールドは地域研究者が地域研究を展開する場であり、インタビューや資料収集などが実施され、その過程で多量の静

止画像・動画・音声・メモなどが生成される。地域研究の黎明期において、これらを記録するメディアはカメラ、テープレコーダ、ノート、スケッチなどのアナログ機器であったが、最近では GPS 機能付きデジタルカメラやノートパソコンなどのデジタル機器が主流となりつつある。

地域情報学からのフィールド支援として、データ入力支援が考えられる。つまりデジタル機器の可搬性と操作性の向上、遠隔地サーバとのネットワーク連携によるデータ交換、センサを利用した計測機能の実現などである。これは地域研究からのアプローチに相当すると考えられ、収集したデータ・資料などのデータベース化を効率化するための前処理として重要である。HGIS 研究グループでは、デジタルカメラ、GPS、レーザ距離計、加速度センサをタブレット PC と組み合わせた「データ入力支援システム」の開発を進めている[5,6]。このシステムでは、①カメラの位置ではなく被写体の位置と高度を算出する（GPS にレーザ距離計と加速度センサを組み合わせることにより実現）、②写真データをリアルタイムに地図上に表示するとともに簡単な GIS 処理を行う、③アノテーションなどの参考情報を追加する、④ネットワークを経由してデータベースサーバと連携するなどの機能の実現を目指している。計測機能の試験を行っている段階で、可搬性や操作性の向上は今後の課題である。なおデータ入力支援システムについては文献[6]を参照。

2.2 地域情報モデル構築

地域情報モデルの構築では、研究資料であるモノ・画像・文書などの物理的・内容的特性などを検

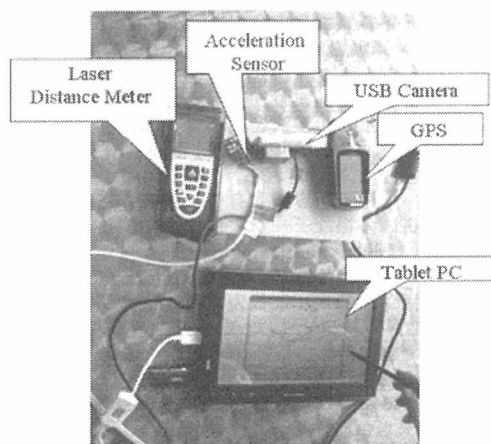


図 2. データ入力支援システム(Ishikawa)[6]

討し、データベース構築に必要なデータ項目選定・項目名称・記述の粒度・言語・記述法・符号化法などの設計を行う。次に述べる資源共有化を考慮した語彙の統制、データ項目間の関連づけなどの設計も、ここに含まれる。データベースの設計も地域研究からのアプローチに相当すると考えられるが、あらゆる情報システムにとって、データベースはまさに基盤（ベース）であり、その設計と実装のいかんは情報処理の速度や精度に大きな影響を与える。

2.3 資源共有化システム

個別のデータベースの設計は、上述の地域情報モデル構築の検討を経て決定される。つまりデータ項目・項目名・検索手順・ユーザインタフェースなどは、データベースごとに異なったものとなる。このままではデータベース検索が複雑になってしまうので、データベースの統合が必要となる。データベースの統合とは、異なるデータベースをシームレス（seamless）に接合することである。つまり利用者からは、データ項目・項目名・検索手順・システムの所在などが一つに見える、言い換えるとネットワーク上に分散する複数のデータベースが単一データベースであるかのように見えるシ情報システムを構築することである。このような分散型データベースシステムを資源共有化システム（resource sharing system）と呼ぶ。資源共有化システムの事例としては、人間文化研究機構の資源共有化システムなどがある[7,8]。資源共有化システムは情報工学からのアプローチに相当する。

資源共有化システムを実現するためには、データベースごとに異なっているデータ構造を統一する必要がある²。各データベースは固有のデータ構造を持っているものの、意味的に共通したデータ項目が多く含まれていることも経験的な事実である。そこで多くのデータベースに共通な意味を持つデータ項目からなるデータベースを定義する。これがメタデータ（metadata）と呼ばれるものである。データベースの各データ項目を、意味的に対応するメタデータのデータ項目にマッピング（mapping）することにより、データ構造の統一を実現する。つまり資源共有化システムではメタデータの設計が重要となる。

人間文化研究機構の資源共有化システムでは、メタデータとして Dublin Core[9]を採用している。しかし Dublin Core は最少のデータ項目定義であり、詳細なデータ記述が困難であるため、使い勝手のよいメタデータとは言い難い。そこで HGIS 研究グループで開発中の資源共有化システムでは、資料一点ごとの記述用に、Dublin Core よりデータ項目が豊富な、Metadata Object Description Schema (MODS) [10]を採

用している。ところが膨大な写真資料群や退職教授の研究資料群などを組織化するためには、資料一点ごとの記述に加え、資料間の依存関係も記述する必要がある。そこでアーカイブズ用のメタデータである Encoded Archival Description (EAD) [11]を導入した。さらに MODS と EAD を関係づけるために、多様なオブジェクトを記述できる電子資料館用メタデータ Metadata Encoding and Transmission Standard (METS) [12]の導入を試みている。HGIS 研究グループで開発中の資源共有化システムの特徴は、3種類のメタデータを導入していることに加え、時空間情報処理を想定して、位置と時間に関するデータ記述法を定義している点にある。なお、この資源共有化システムは後述する知識生成支援システムのクリアリングハウス機能も提供する。

2.4 オントロジー辞書

言語や語彙の違いを乗り越えた検索の実現、さらに知識を体系化するための枠組みとして利用する。HGIS 研究グループでは、日本の歴史地名および暦日に関するオントロジー辞書を構築中である。これも情報工学からのアプローチに相当する。詳細は 4.3 および 4.4 で述べる。

2.5 知識生成支援

知識生成の支援では、資源共有化システムに蓄積された多様なデータ・情報を検索し、それら进行分析・統合・解析して新しい知識生成の支援を試みる。具体的には可視化・語彙分析・画像解析・統計処理・時空間解析・シミュレーションなどを行い、データ間の関連付け・分類・情報縮約・因果関係の抽出・将来予測・評価などを行う。地域研究における知識生成では、データ・情報の時空間的な把握が特に重要となる。実際、地域研究においても GIS を利用した空間データの可視化は一般的な手法として定着しつつある。しかしシミュレーションや時空間統計解析などは未だ馴染みが薄く、知識生成支援過程で利用する殆どの手法は、情報工学からのアプローチに相当するものと考えられる。この点については次節で述べる。

前記のように、地域情報学を具体的に推進するためにはフィールド支援、地域情報モデル構築、資源共有化システム、オントロジー辞書、知識生成支援の各プロセスに対応した情報処理技術を確認しなければならない。ただし以降では、地域研究において特に重要と考えられる、「時間」および「場所」の情報処理に焦点を絞る。

3. 時空間情報の重要性

情報を整理する観点の一つとして古くから 5W1H がある。5W は対象に関する Who（誰が）、What（何を）、When（いつ）、Where（どこで）、Why（どのような目的で）情報であり、1H は How（どのようにして）情報である。我々が報告書などを作成する場合、殆ど無意識のうちに、報告内容を 5W1H の観点に分解している。

²例えば A と B という 2 つのデータベースに人名を登録する場合を考える。データ項目が、データベース A では「Name」、B では「名前」である場合、同一の検索命令で二つのデータベースを同時に検索するは困難である。さらにデータベース A では姓と名が別々のデータ項目となっているのに対して、データベース B では姓と名が一つのデータ項目にまとめられている場合、同時検索はより複雑となる。

データベースの設計においても 5W1H は重要な観点であるが、観点の重みには差があるように思われる。例えば書誌目録の場合、題名 (What) と作者 (Who) の情報が多用される。時間 (When) と場所 (Where) に関する情報も重要であるが、一般的な書誌目録では資料の内容とは関係の薄い情報が記述されているので (When は出版年、Where は出版地)、これらを使う機会は少ない。主題は 5W1H の全てを含みうる。しかし日本の場合、主題の整備が遅れているうえに、主題の中身が区別されていないことが多い³。したがって、例えば「近衛」について検索しようとしても、それが人名なのか、地名なのか、あるいは職名であるのか区別できない。このように、現状の書誌目録における When と Where はあまり重宝しない。しかし内容に関する When と Where 情報が整備されれば、情報検索のための重要な手がかりとなることは疑いがない。

ところで、非文字資料、例えば考古的遺物あるいは地震などの自然現象に著者はいないので、これらが Who に関する情報を持つことは希である (遺物の発見者などはあり得る)。また What に関する情報についても、識別性の高い記述を与えることが困難な場合も多い (単に「壺」など)。これに対して、When (「いつ」対象物が作られたのか、使われていたのかなど) と Where (「どこで」対象物が作られたのか、発見されたのか、あるいは「どこに」存在していたのかなど) は、殆どの事象・事物に共通の情報である。実際、When と Where が唯一の検索手段あるいは比較の指針となる場合も多い。

このように、When と Where は殆どの事象・事物に共通な情報である (以降では When と Where を併せて時空間 (spatiotemporal) 情報と呼ぶ)。したがって、多様な情報資源を統合する資源共有化システムの検索において、時空間情報は Who および What とともに不可欠な観点となる。また資料分析の過程で、事象の発生場所を地図上にプロットしたり、複数のプロットを重ね合わせて分布を比較したり、事象を時間順序に並べ替えて年表を作成したり、因果関係を見いだすために複数の年表を並べるような操作は、地域研究においても基礎的な作業である。これらは時空間情報を基盤としている。以上の考察から、地域情報学のあらゆる場面において時空間情報は重要な役割を果たすと考えられる。

4. 時空間情報解析ツール

時空間情報処理に関する情報モデルを図 3 に示す。この図では時空間情報 (Spatiotemporal Data) を主題軸 (Theme/Subject Axis)、空間軸 (Spatial Axis)、

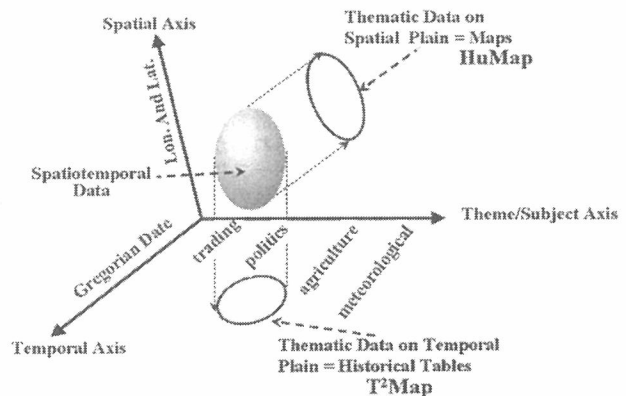


図 3. 時空間情報モデル(Kubo and Sekino)[14]

時間軸 (Temporal Axis) の 3 次元に分解して表現している (空間情報を緯度と経度で表せば全体として 4 次元となる)。主題軸は、その名の通り、対象の主題・内容を語彙情報に従って配列するもので、従来の目録に相当する。空間軸は対象を位置情報に従って配列するもので、従来の地図に相当する。標準的な位置情報の記述法は緯度と経度の対であるが、住所や郵便番号などでも構わない (後述する地名辞書などを利用して緯度・経度の対に変換される)。主題軸と空間軸が形成する平面には、様々な主題図 (ベース地図、道路地図、土地利用図、汚染分布図、患者発生地図など) が分布する。GIS とは、これらの地図をコンピュータ上に表示したり、重ね合わせたり、演算処理を行う情報処理システムである。時間軸は対象を時間情報に従って配列するもので、従来の年表に相当する。標準的な時間情報の記述はグレゴリオ暦であるが、和暦などでも構わない (後述する暦日テーブルなどを利用してグレゴリオ暦に変換される)。主題軸と時間軸が形成する平面には、様々な年表 (日本史、朝鮮史、中国史、タイ史など) が分布する。GIS に相当し、複数の年表をコンピュータ上に表示したり、重ね合わせたりする情報処理システムは、あまり例がない (例えば MIT による SMILE プロジェクトの TimeLine[13])。

以下では HGIS 研究グループが中心となり開発を進めている時空間情報解析ツールについて述べる。主題軸と空間軸が形成する地図平面に対応する GIS ツールとしての HuMap (Humanities Map)、および主題軸と時間軸が形成する年表平面に対応するツールとしての T²Map (Theme and Time Map) が基盤となる時空間情報解析ツールである。さらに時空間情報解析を支援するツールとして、デジタル歴史地名辞書と暦日テーブルの開発を進めている。デジタル歴史地名辞書は、歴史地名の属性、現在地名との対応、緯度・経度などを登録した地名辞書である。暦日テーブルは、グレゴリオ暦と和暦など、異なる暦日間の相互参照表である。以下、詳しく述べる。

³ 例えば LC MARC (米国議会図書館の機械可読目録: Machine Readable Cataloging) の場合、tag 150 \$y に時間に関するトピックを記述することができる。また tag 651 には地理名称を記述することができる。

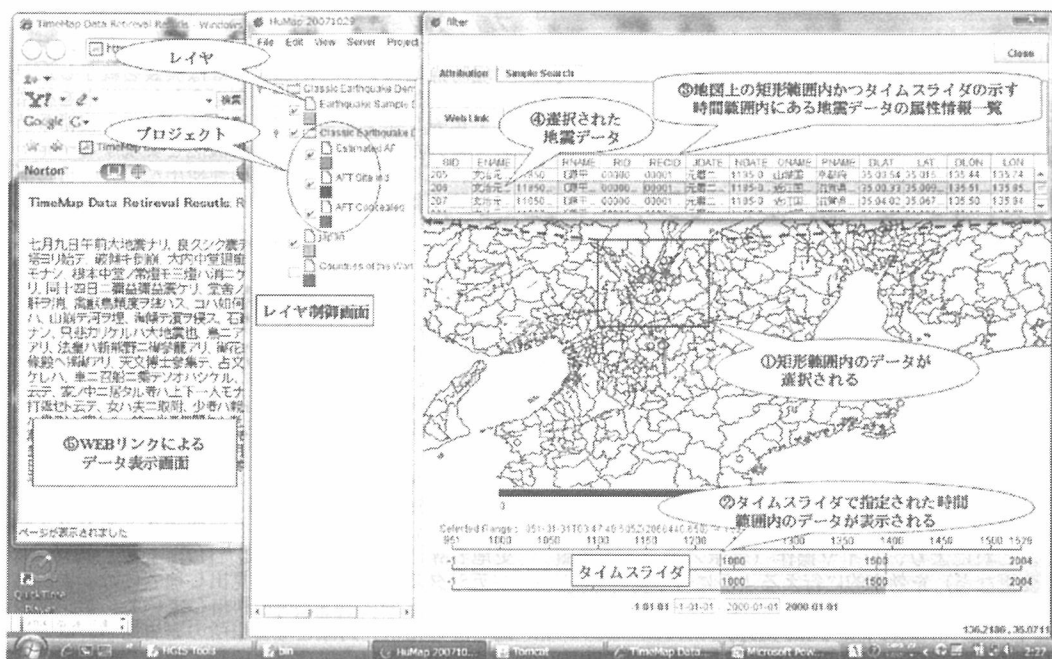


図4. HuMapによる時空間データの表示例

4.1 HuMap (Humanities Map)

HuMapは主題軸と空間軸が形成する地図平面上の情報を処理するGISツールである。ただし時間情報も処理できる特徴を持っている。当初はシドニー大学のTimeMap Project[15]で開発されたTMWinをベースとしていたが、最新バージョンの全ソースコードはHIGS研究グループのオリジナルとなっている。HuMapはGISツールであるため、

- ① ベース地図の上に複数のレイヤ (layer: 主題図、分布図、衛星写真、航空写真など空間データの単位) を重ね合わせて可視化する
- ② 空間領域におけるズームイン・ズームアウト
- ③ フィーチャ (feature あるいは地物: 道路、病院、神社、河川、行政界など実世界に存在するものを抽象化した概念。レイヤ上では点、線、多角形などの図形として表現される) の主題属性を対象とした検索などのデータ処理
- ④ 空間属性あるいは時間属性を対象とした検索などのデータ処理
- ⑤ レイヤ間の比較・論理演算などの基本的なGIS機能を実現している。さらにHuMapの特徴として、
- ⑥ フィーチャに関する情報 (写真、参考文献など) へのWebリンクを介したアクセス
- ⑦ 複数のレイヤをプロジェクト (project) という単位にまとめることができ、レイヤの表示・非表示などの操作が容易
- ⑧ 時間情報に従ってフィーチャを時系列的に表示するアニメーション機能

⑨ アノテーションの付加

⑩ JAVAで作成されているGISツールであるためプラットフォームへの依存性が低い

⑪ JAVAのPlug-in機能により、ツール外で作成された高度な時空間情報処理プログラムをHuMapの時空間処理ツールとして利用できる

⑫ クリアリングハウス⁴と連携したデータおよび時空間処理ツール検索とインポート、およびクリアリングハウスへのエクスポートができるなどの機能をあげることができる (一部は実装中)。HuMapではESRI ShapeFile、CSV (Comma Separated Values) ファイル、XML ファイルを扱うことができる。これらに加えて、画像データとしてJPEG、JPEG2000、GIF、TIFF、PNGも利用可能である。

図4にHuMapの表示例を示す。これは近畿付近における歴史地震の分布例 (5章参照) を示している。ここでは地図上の矩形 (図中①) とタイムスライダの時間 (図中②) で指定された時空間範囲にある地震データとその属性情報一覧を表形式 (図中③) で

⁴ 本来のクリアリングハウス (clearinghouse) は、資料やデータを収集・保存して利用できるようにする機関のことである。GIS分野のリアリングハウスは、インターネット上に分散している地理情報を交換する仕組みである。複数のデータベースで蓄積されている地理情報に関するメタデータを管理し、メタデータ検索を介して空間データの所在情報や空間データへのアクセス機能などを提供する。クリアリングハウスによっては、ダウンロードだけではなく、独自に作成した空間データを公開する機能も提供している。HIGS研究グループのクリアリングハウスは空間データのみならず、空間解析ツールの登録・公開も可能となっている。



図5. T2Mapによる時間データ表示例

提示している。この表から特定の地震データを選択すると（図中④）、地図上では色の反転で位置を示し、WEBリンクを介して詳細なデータを表示することができる（図中⑤）。またHuMapでは複数のレイヤをプロジェクトという単位にまとめることができる。これにより、レイヤ操作（表示／表示や表示順の変更など）を効率的に行えるようになった。

4.2 T²Map (Theme and Time Map)

T²Mapは主題軸と時間軸が形成する年表平面上の情報を処理する新しい時空間情報処理ツールである。GISツールでは、位置情報に従ってフィーチャを配列した地図をレイヤとし、複数のレイヤを重ねて可視化する。T²Mapでは、時間情報に従って事象を配列した年表をレイヤとし、複数のレイヤを並べて可視化する。T²Mapにおける時間的事象はテキスト、数値、画像により表現される。

T²Mapの機能は、複数のレイヤの可視化、レイヤの重ね合わせ、時間領域におけるズームイン・ズームアウト、レイヤ間演算など、HuMapの空間処理を時間処理に置き換えたものである（一部は実装中）。T²Mapが扱えるデータファイルはXMLファイルである。これに加えて画像データとしてJPEG、JPEG2000、GIF、TIFF、PNGが利用可能である。HuMapと同様にT²MapもJAVAで作成されており、

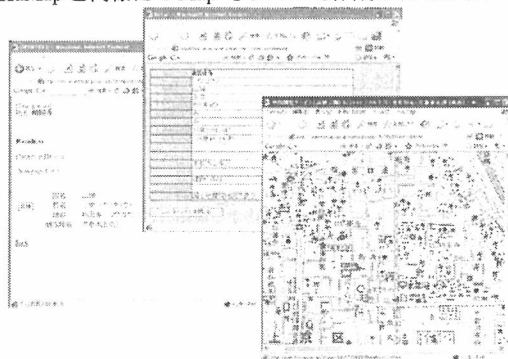


図6. デジタル地名辞書の表示例

プラットフォーム依存性が低い。図5にT²Mapの表示例を示す。

現時点において表示機能はほぼ実装を終了したが、データ解析機能は特定文字列を含むレコードを検索して年表レイヤを生成する程度である。今後は数値型を含めた複雑な問い合わせ機能を実現する予定である。さらに移動平均や周期性検出などの解析機能も実装する予定である。T²Mapの詳細については文献[17]を参照。

4.3 オントロジー辞書

空間データを作成する場合、地名を緯度・経度に変換しなければならない。また地名は時代により変化しているため、その変遷を辿れなければ検索は困難である。さらに同名異地などの問題もある。これらを解決するためには、地名に関するシソーラスと地理参照機能を備えたデジタル地名辞書（Digital Gazetteer）が必要である。海外には大規模かつフリーのデジタル地名辞書がいくつか存在するが（例えば[17,18]）、残念ながら日本にはフリーのデジタル地名辞書がない。そこで歴史地名に関する「デジタル歴史地名辞書」の構築に着手した[19]。

デジタル歴史地名辞書の見出し地名には、吉田東伍の大日本地名辞書[20]を利用している。デジタル地名辞書には、対象となる場所の歴史地名、ヨミ、包含関係（対象の場所はその郡に属し、その郡はどの国に属すかなど）を示す地名とヨミ、対象の本来の形状（点、線、多角形など）、対象となる場所の現在地名、ヨミ、包含関係、緯度、経度、地名属性（町、山、川などの区別）などが登録されている。現在、平成大合併を考慮した地名データの追加作業を進めている。デジタル地名辞書試作システムの実行例を図6に示す。なおデジタル地名辞書の詳細については文献[19]を参照。

4.4 暦日テーブル

デジタル歴史地名辞書が、地名と緯度・経度の参照機能を提供していると同様に、暦日テーブルは異なる年表間の日付の参照機能を実現する。暦日テーブルではユリウス通日をキーとし、ユリウス通日の一日ごとに、各暦の日付を対応させている。現在はグレゴリオ暦と和暦（北朝と南朝）のみが登録されているが、海外の暦日を追加する予定である。

5. 歴史データのHuMapへの適用例

史料から時空間情報を抽出し、HuMapを適用するまでの過程について「歴史地震データベースの構築」を事例として紹介する[21]。

5.1 歴史地震資料のデジタル化

我が国の歴史地震記事は、田山実編『大日本地震史料』[22]や武者金吉編『増訂大日本地震史料』[23]などに網羅されている。これらの歴史地震史料集には正史のみならず伝聞・噂あるいは後世に書き加えられたテキストなど多種多様な記事が収集されている

```

<Earthquake page="228"><Header>明應七年八月二十五日（西暦 1498,9,20）</Header>
<E.ID>14980920</E.ID><J.Date>明應七年八月二十五日</J.Date>
<S.Date type="Gregorian">14980920</S.Date>
<E.Description><section>伊勢、<gaiji set="mojikyō" code="067322">紀</gaiji>伊、<gaiji set="daikanwa" code="039047">遠
</gaiji>江、三河、駿河、甲斐、相模、伊豆諸國、地大ニ震ヒ、顔<gaiji set="daikanwa" code="017503">海</gaiji>ノ國ハ
津浪ノ害ヲ<gaiji set="mojikyō" code="075258">蒙</gaiji>リ、就中伊勢國大湊ニテハ家千軒押シ流サレ五千人<gaiji
set="daikanwa" code="017990">溺</gaiji>死ス、マタ鎌倉由比浜ニテハ水勢大佛殿ニ及ビ二百人<gaiji set="daikanwa"
code="017990">溺</gaiji>死セリ、是日、京都、奈良及ビ陸奥國<gaiji set="mojikyō" code="066797">曾</gaiji>津モ強ク震
ヒ、<gaiji set="mojikyō" code="054722">餘</gaiji>動月ヲ重ネタリ、</section></E.Description>
<E.Sources page="228"><Header>〔御湯殿の上の日記〕</Header> <E.ID>14980920</E.ID>
<Source.ID>00000348</Source.ID><Source.Record page="228"><Record.ID>00003733</Record.ID><E.ID>14980920</E.ID>
<Source.ID>00000348</Source.ID><J.Date>明應七年八月廿五日</J.Date><S.Date type="Gregorian">
<Record.Description> <section><ruby><rb>廿五</rb></ruby></section>（明應七年八月）</ruby>日、けさ<ruby><rb>ちしん
</rb></ruby>（地震）</ruby>けう／＼／しうゆる、しよくゆへの御いのりの事、ふ行におほせいたさるゝ、
</section></Record.Description></Source.Record>

```

図 7. 歴史地震史料の電子化テキスト例

る。そのため信頼性の低い記事も混在しており、誤った結論を導く原因ともなっている。

「古代・中世の全地震史料の校訂・電子化と国際標準震度データベース構築に関する研究」は、歴史地震史料集が抱える上記の問題解決を目指した研究プロジェクトである（以下、歴史地震研究）[24]。歴史地震研究では、史料の信頼性・異本関係・校訂などの本文批判（text critic）を通じて、信頼性の高い地震記事の選別・修正を行い、精度の高い地震史料データベースの構築を目指している。歴史地震研究では史料のデジタル化から着手した。ここでの特色は XML を全面的に導入している点にある。デジタル化はほぼ終了し、作成された電子テキストを利用した本文批判の作業が進行中である。XML マークアップを施した電子テキストの例を図 7 に示す。

5.2 時空間情報の抽出と利用

電子化テキストを眺めてみると、各地震記事には必ず時間情報（図 7 の<J.Date>など）が付与されており、主な内容は地震に関連した事象（建物の倒壊など）と場所に関する記述である。そこで地震に伴う事象の発生時刻と位置の同定を試みた。サンプルとして 9 つの地震を選択した。

実験では、まず地震記事から建物・市街・遺跡などの地名を抽出した。次に抽出した地名に対応する現在地名を可能な限り推定した。最後に現在地名から緯度と経度を求めた。市街のように広がりを持つ地名の場合、役場など、その区域を代表するフィチャーの緯度・経度（地点）を採用した。和暦は年月日までをグレゴリオ暦に換算した。これらの作業は歴史専攻の大学院生に依頼した。

作成された時空間データの一部を表 1 に示す。ここで NDATE は地震に伴う事象の発生時刻、ONAME は事象が発生した場所、PNAME はその現在地名、DLAT はその緯度、DLON は経度である。このデータを HuMap で利用するには ESRI Shapefile 形式あるいは CSV ファイル形式に変換しなければならない。ESRI Shapefile 形式への変換には特別なツールが必要である。HuMap は CSV ファイルを利用できるので、特別なツールは不要である。

ところで GIS ツールではレイヤ制御などの情報を設定する必要がある。HuMap ではレイヤ用（tmm）とプロジェクト用（tms）の 2 種類の制御ファイルが必要である。両者とも XML 形式のパラメータ設定ファイルであるが、かなり煩雑である。図 4 に示した HuMap 表示例における Earthquake Sample Data レイヤ用の tmm ファイルを図 8 に示す（説明は省略する）。これを手作業で作成するのは簡単ではない。支援ツールの構築が必要である。

歴史地震研究は継続中であり、ここで示したデータも作成途上であるが、この地震発生分布地図に震度分布や災害地図などを重ね合わせれば、基本的な震度データベースを構築できると考えている。本格的な解析作業はこれからである。

5. あとがき

地域研究における地域情報学の構築と、それを実現する情報システムの開発に着手した。試作レベルの情報システムが形になりつつあり、今年度中の試験公開を目指している。

時空間情報処理は地域研究における新しい研究パラダイムであり、Google Earth などの既存ツールに

表 1. 地震記事の時空間データ例

SID	PNAME	EID	RNAME	RID	RECID	JDATE	SIDATE	ONAME	PNAME	DLAT	DLON	NOTE
1	貞元元年六月十八日	09760722	〔日本紀略〕	00000022	00001151	貞元元年六月十八日癸丑	0976-07-22	山城国/京都/宮城	京都府/京都市/上京区/小山西	33.00.54	135.44.43	千本丸太町交差点近くに大徳院基壇跡の石碑(平成6年)あり
2	貞元元年六月十八日	09760722	〔日本紀略〕	00000022	00001151	貞元元年六月十八日癸丑	0976-07-22	山城国/京都/宮城/八省院	京都府/京都市/上京区/小山西	35.00.54	135.44.43	千本丸太町交差点近くに朝堂院〔＝八省院〕跡のプレートあり
3	貞元元年六月十八日	09760722	〔日本紀略〕	00000022	00001151	貞元元年六月十八日癸丑	0976-07-22	山城国/京都/宮城/豊樂院	京都府/京都市/中京区/東樂遊西町	35.00.54	135.44.43	平安宮豊樂院跡
4	貞元元年六月十八日	09760722	〔日本紀略〕	00000022	00001151	貞元元年六月十八日癸丑	0976-07-22	山城国/京都/東寺	京都府/京都市/南区/九条東	34.58.40	135.44.58	左は京寺〔教王護国寺〕の現在地
5	貞元元年六月十八日	09760722	〔日本紀略〕	00000022	00001151	貞元元年六月十八日癸丑	0976-07-22	山城国/京都/西寺	京都府/京都市/南区/唐橋平垣町	34.58.43	135.44.19	左は西寺の現在地

```

<?xml version="1.0" encoding="UTF-8"?>
<METADATA>
<META NAME="dc.title" TMCODE="8" SCHEMECODE="7"
  CONTENT="Earthquake Sample Data"/>
<META NAME="dc.subject.specifc" TMCODE="10"
  SCHEMECODE="7" CONTENT=""/>
. . . . .
<META NAME="dc.coverage.x.min" TMCODE="4"
  SCHEMECODE="27" CONTENT="130.308"/>
<META NAME="dc.coverage.x.max" TMCODE="5"
  SCHEMECODE="27" CONTENT="139.933"/>
<META NAME="dc.coverage.y.min" TMCODE="6"
  SCHEMECODE="27" CONTENT="31.760"/>
<META NAME="dc.coverage.y.max" TMCODE="7"
  SCHEMECODE="27" CONTENT="37.573"/>
<META NAME="dc.coverage.t.early" TMCODE="38"
  SCHEMECODE="14" CONTENT="0"/>
<META NAME="dc.coverage.t.late" TMCODE="39"
  SCHEMECODE="14" CONTENT="2003"/>
. . . . .
<META NAME="tm.Filter.AccessibleFields" TMCODE="80"
  SCHEMECODE="7"
  CONTENT="SID,ENAME,EID,RNAME,RID,RECID,JDATE,
  NDATE,ONAME,PNAME,DLAT,LAT,DLO,LO,ON"/>
<META NAME="tm.MapWebLinks.URL.Mask"
  TMCODE="63" SCHEMECODE="7"
  CONTENT="http://localhost:8080/ClassicEarthquake/servlet/
  EarthquakeRecordWriterServlet?recid={RECID}"/>
<META NAME="tm.sys.Local.GISType"
  TMCODE="34" SCHEMECODE="8" CONTENT="csv"/>
<META NAME="tm.Spatial.MapObjType"
  TMCODE="36" SCHEMECODE="17" CONTENT="Point"/>
<META NAME="tm.Spatial.PointXfield"
  TMCODE="75" SCHEMECODE="7" CONTENT="DLON"/>
<META NAME="tm.Spatial.PointYfield"
  TMCODE="76" SCHEMECODE="7" CONTENT="DLAT"/>
</METADATA>

```

図8. tmm 設定ファイル例(一部)

拘泥されるよりも、自由な発想で研究開発を進めた方がよいと考えている。実際 HuMap や T²Map などは我々自身が開発しているツールであるため、研究の進展に応じた機能の変更・拡張を臨機応変に行うことができる。例えば HuMap と T²Map を相互に連携できるツールに変更する予定である。つまり、特定の事象を空間的な視点から俯瞰・分析する際には HuMap を利用し、同じ事象を時間的な視点から俯瞰・分析したければ T²Map に切り替える。これを実現するため、HuMap と T²Map が同じメタデータファイルを共有できるように仕様を変更する⁵。時間と空間を同時に処理できるようにするため、TimeMap や TimeLine あるいは Google Earth より優れた解析ツールになると言える。これは進展が著しい研究領域における支援ツールとして重要な要件であると考える。

謝辞

本稿に掲載したシステムおよびデータは HGIS 研究グループの研究成果に基づいている。HGIS 研究グループのメンバー、特に桶谷猪久夫（大阪国際大学教授）、久保正敏（国立民族学博物館教授）、貴志俊彦（神奈川大学教授）、関野樹（総合地球環境学研究所準教授）、相田満（国文学研究資料館助教）、石川正敏（東京農工大学助教）の各氏に感謝いたします。

参考文献

[1] 地域研究統合情報センター: http://www.cias.kyoto-u.ac.jp/index.php/research_activities.

- [2] シンポジウム地域研究と情報学: 新たな地平を拓く (2007/2/9-10), http://www.cias.kyoto-u.ac.jp/index.php/news_detail/id/48.
- [3] 柴山守: 地域情報学の創出, <http://gissv2.cseas.kyoto-u.ac.jp/kiban-s/>.
- [4] HGIS 研究会: 京都大学地域研究統合情報センター 地域情報資源共有化プロジェクト「地域情報学の創出」, <http://www.cias.kyoto-u.ac.jp/index.php/kyodo#pub02>.
- [5] 石川正敏, 原正一郎, 柴山守: 位置情報作成システムの設計, 日本情報考古学会講演論文集, Vol.3, pp.53-58, 2007.3.
- [6] 石川正敏, 原正一郎, 柴山守: 人文科学のための現地調査支援システムの試作, 人文科学とコンピュータシンポジウム論文集 2007, (編集中).
- [7] 原正一郎: 国文学支援のための SGML/XML データシステム, 情報知識学会論文誌, Vol.11, No.4, pp.17-35, 2002.
- [8] 原正一郎: 人間文化研究機構資源共有化システムについて, シンポジウム 地域研究と情報学: 新たな地平を拓く 講演論文集, pp.107-136, 2007.
- [9] Dublin Core Metadata Initiative: Dublin Core Metadata Element Set Version 1.1 Reference Description, <http://dublincore.org/documents/dces/>, 2004.
- [10] Library of Congress: MODS Metadata Object Description Schema Official Web Site, <http://www.loc.gov/standards/mods/>.
- [11] Library of Congress: Encoded Archival Description Version 2002 Official Site: <http://www.loc.gov/ead/eadd.html>.
- [12] Library of Congress: Metadata Encoding & Transmission Standard Official Site, <http://www.loc.gov/standards/mets/mets-schemadocs.html>.
- [13] MIT: "SMILE Project TimeLine", <http://simile.mit.edu/timeline/>.
- [14] Tatsuki SEKINO: Temporal Based Information System (T2Map), PNC2007 Abstract, http://pnclink.org/pnc2007/pdf/20_0940_AreaInformaticI.pdf.
- [15] TimeMap: <http://www.timemap.net/>.
- [16] 関野 樹, 久保 正敏: T2Map - 時間情報に特化した解析ツール, 人文科学とコンピュータシンポジウム論文集, 2007, (編集中).
- [17] Alexandria Digital Library Project Gazetteer Development, <http://www.alexandria.ucsb.edu/gazetteer/>, <http://middleware.alexandria.ucsb.edu/client/gaz/adl/index.jsp>.
- [18] Getty Thesaurus of Geographic Names Online: http://www.getty.edu/research/conducting_vocabularies/tgn/.
- [19] 桶谷猪久夫: 人文分野における日本地名辞書の構築と地名属性の特徴分析, 人文科学とコンピュータシンポジウム論文集, 2007, (編集中).
- [20] 吉田東伍: 大日本地名辞書, 富山房.
- [21] 原正一郎: 地震史料の XML データ作成, 月刊 地球, 総特集 地震史料の校訂とデータベース化 - 日本の古代・中世を中心に -, 海洋出版 vol.27, No.11, pp.853-860.
- [22] 田山実: 大日本地震史料, 震災予防調査会報告, No.46, 1904.
- [23] 武者金吉編: 増訂大日本地震史料, 文部省震災予防防議, 第1巻, 1941, 第2巻, 1943, 第3巻, 1943.
- [24] 石橋克彦: 古代・中世の全地震史料の校訂・電子化と国際標準震度データベース構築に関する研究, 科学研究費補助金基盤研究 (A)(1), <http://historical.seismology.jp/erice/>.

⁵ 本稿執筆時点において、HuMap と T²Map の開発進捗状態が異なっているため、相互利用は実現していない。