

第2章 経済産業省の取り組み

1 経済産業省商務情報政策局実施事業について

1 - 1 G-XML プラットフォーム構築等事業（クリアリングハウス事業）

（1）事業の目的と課題

1）事業の背景

地理情報システム（GIS）は、これまでハードウェア、ソフトウェアの低価格化が進み、簡易なGISの導入が可能となる一方で、地図データなどが電子化されていない、標準がなくデータやシステムの相互接続性、相互運用性が確保されていないなどにより、GISを導入する各主体がそれぞれ独自に電子地図などを整備しなければならず、社会的に二重三重の投資になるなどの問題が生じていた。

またインターネットの爆発的な普及とともに、インターネット上での相互接続性、相互運用性を実現するグローバルスタンダードの確立が不可欠であり、地理情報の分野では民間や国際標準化機構（ISO）の第211技術委員会（TC211）においてデータやシステムの相互接続性、相互運用性のためのグローバルスタンダードの審議がなされている。

一方、国内では地理情報の所在案内システムであるクリアリングハウスが「空間データの相互利用を促進し、空間データ整備の重複投資を回避するために不可欠」なものとして「国土空間データ基盤標準及び整備計画」（GIS関係省庁連絡会議平成11年3月30日決定）において位置付けられ、「経済構造の変革と創造のための行動計画（第3回フォローアップ）～新たな経済成長に向けての新行動計画～」（平成12年12月1日閣議決定）により、「政府全体の地理情報クリアリングハウス（地理情報の所在案内システム）を国際標準を踏まえて整備し、今年度中に運用開始するほか、各省庁が保有する地理情報について地理情報標準に準拠したメタデータを整備する。」とされた。

しかしながら、当該国際標準（ISO23950）を基にしたJIS X0806「情報検索（Z39.50）応用サービス定義及びプロトコル仕様」においては、情報を検索する際に端末とサーバ間で地理情報標準に従ったメタデータをやり取りするために必要なパラメータ属性とその属性間関係を規定した国内向け地理情報プロファイルが制定されていないため、問題となっている。

2）事業の目的

上記背景の下、経済産業省は、「平成13年度G-XMLプラットフォーム構築等事業（クリアリングハウス事業）」を実施し、米国政府における地理情報プロファイルの策定動向、

ISO/TC211 でのメタデータ標準の審議動向調査に併せ、相互接続性、相互運用性を確保した国内用地理情報プロファイル（GEO-J プロファイル）案の作成、及び、当該 GEO-J プロファイル案を用いたクリアリングハウスの構築実証を行った。

3) 事業の内容と課題

本調査は、昨年度に経済産業省が実施した「平成 12 年度地理情報システム標準化等推進事業」による成果のうち、以下の主要な成果を基に実施した。事業の実施フローは図 2-1-1 のとおりである。

GEO-J プロファイル案

3 モデル地区（豊中地区、大分県地区、臼杵地区）の地理情報メタデータ

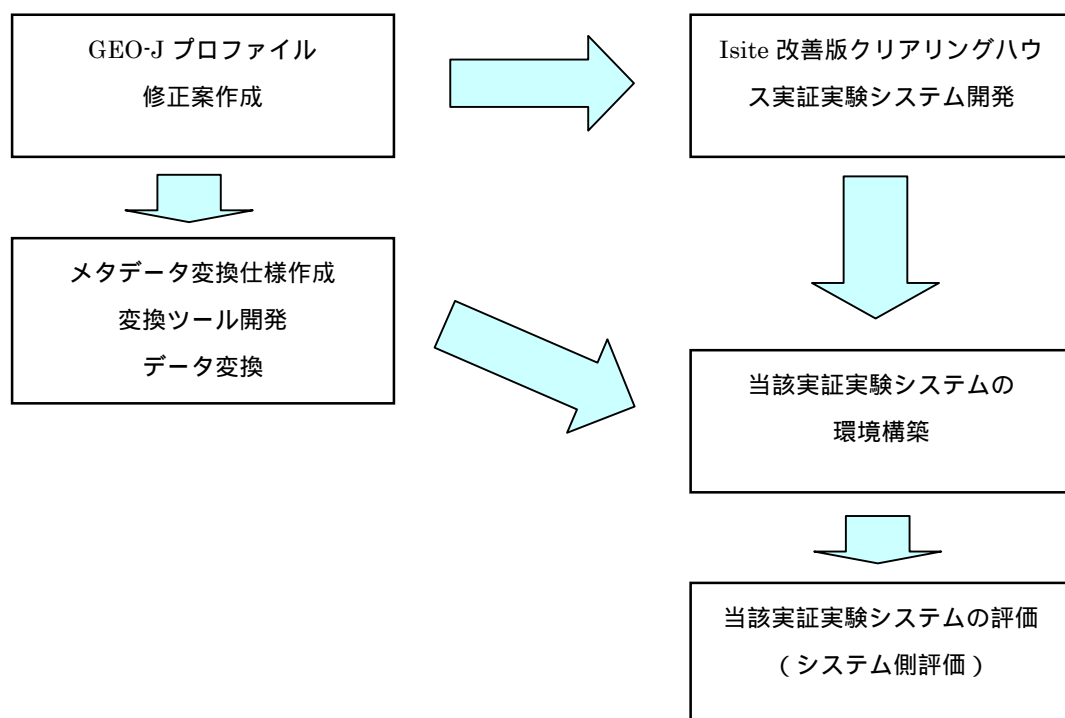


図 2 - 1 - 1 事業実施フロー

GEO-J プロファイル案については、平成 12 年度事業の中で、米国の地理情報プロファイルである GEO Profile Version 2.2 (GEO2.2) 及び日本国内用メタデータ標準である JMP1.1 を基に作成した。今年度は、ISO/TC211 のメタデータ標準 (ISO19115) が DIS (Draft International Standard) となり、それに基づいて GEO2.2 並びに JMP の改定が予定されていたため、それに歩調を併せて GEO-J プロファイル案の修正を行った。

3 モデル地区のメタデータについては、平成 12 年度事業の中で、豊中市、大分県、臼杵市において地理情報を収集し、JMP1.1 をベースにメタデータを作成した。今年度は、

ISO19115DIS 並びに JMP の改定動向に併せて、昨年度に収集したメタデータの調査分析・編集を行った。

今年度は、実証実験用地理情報クリアリングハウスを、Z39.50-Web ゲートウェイサーバ1台とメタデータサーバ3台で構成した。ゲートウェイサーバには昨年度事業で開発した METIS を使用し、メタデータサーバには、Z39.50 用フリーソフトウェアである Isite を改善してシステム構築を行った。

なお、実証実験では、上記システムに、GEO-J プロファイル修正案を採用し、今年度事業用に編集・加工したメタデータをメタデータサーバに蓄積し、ユーザはインターネットを経由して当該メタデータの検索並びにサンプル地理情報画像の閲覧を行った。これにより、本事業で構築した実証実験用地理情報クリアリングハウスのシステム側実証実験を行うとともに、GEO-J プロファイル修正案の確実性、システム運用面、データ整備面における課題等について検証した。

GEO-J プロファイル案の修正にあたっては、ISO/TC211 における標準化動向及び米国における GEO プロファイルの改定動向を調査し、修正案に反映させるとともに、国内標準化に向けた検討を行った。具体的調査内容及び課題は以下のとおりである。

- a) 昨年度成果を基に国内向け地理情報プロファイル案の修正及び関連調査を行う。
- b) 地理情報メタデータプロファイル改定動向に併せ、昨年度成果である3モデル地区（豊中地区、大分県地区、臼杵地区）の地理情報メタデータの調査分析・編集を行う。
- c) Z39.50 情報検索プロトコル用のメタデータサーバソフトである Isite を改善し、クリアリングハウスを再構築する。
- d) 3モデル地区の編集済みメタデータのクリアリングハウスへの蓄積を行う。
- e) 当該クリアリングハウスに係るシステム側実証実験を行う。

4) 実施概要

実施期間及び実施体制

本調査は、平成13年5月11日に開始し、平成14年3月29日に終了した。本調査実施にあたっては、財団法人データベース振興センターに、Z39.50 情報検索プロトコル、地理情報、GIS 等に関する有識者からなる「クリアリングハウス調査委員会」を設置し、調査方針、調査内容・課題、調査結果等について審議・検討を行った。

また、調査作業の実施においては、同センターにプロジェクトチームを設置し、3モデル地区における地理情報メタデータの調査分析・編集、地理情報クリアリングハウスの再構築、当該クリアリングハウスを利用したシステム側実証実験を実施した。

実証実験環境

実証実験環境として、富山サイト及び東京サイトそれぞれに、地理情報メタデータを格納する Z39.50 メタデータサーバを分散配備するとともに、ユーザがクライアント PC からインターネットを経由して、2つのメタデータサーバにアクセスするための Web サーバ及び Z39.50-Web ゲートウェイを富山サイトに配備した。

3モデル地区の地理情報メタデータ及びサンプル地理情報画像データについては、富山サイトに豊中地区データ、東京サイトに大分県地区及び臼杵地区データをそれぞれ独立したサーバに格納した。詳細は、図 2-1-2 に示すとおりである。これにより、ユーザから Web ブラウザによりアクセス可能な分散環境での実験用 Z39.50 地理情報クリアリングハウスを実現した。

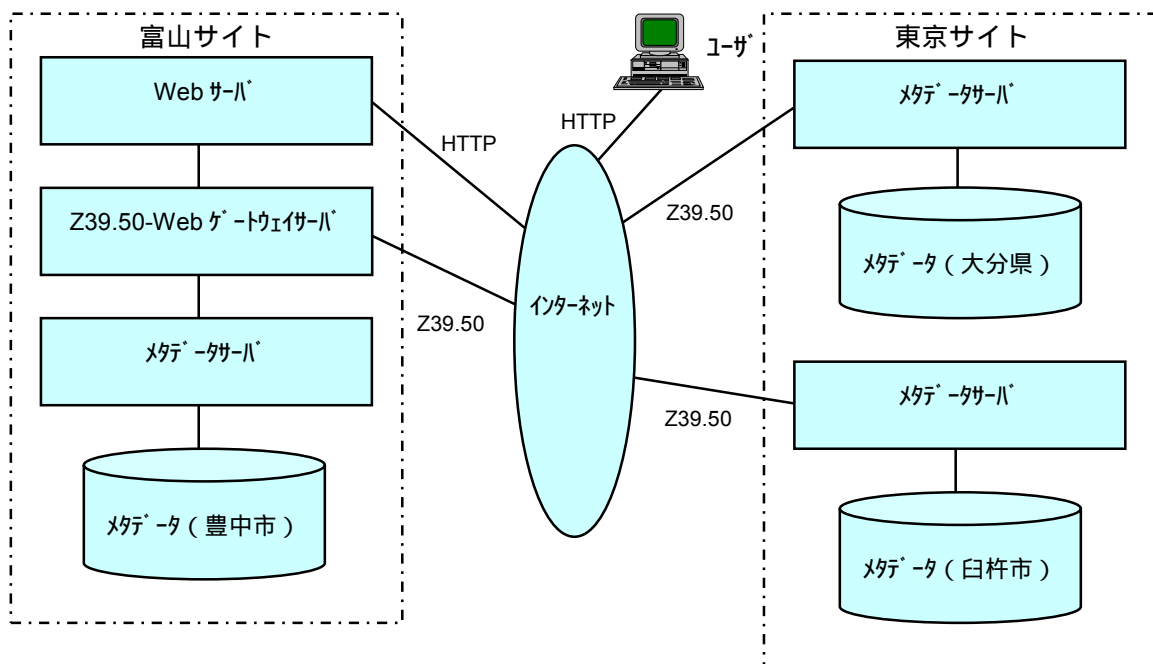


図 2 - 1 - 2 実証実験システムの構成

実証実験及び評価

上記実験環境下において、当該クリアリングハウスの実験的ユーザである本事業プロジェクトチームメンバー及びクリアリングハウス調査委員会メンバーを対象に、インターネット上での、当該地区のメタデータ及びサンプル画像の検索・返戻・参照の実証実験を行った。

なお、本実証実験では、昨年度事業で実施した実ユーザによる実証は行わず、システム側実証及びユーザを想定したシステム検証のみを実施した。システム側実証においては、（今年度事業で改善した Isite により）構築したメタデータサーバ、及び、昨年度事業で開発した METIS のゲートウェイ機能を用いた Z39.50-Web ゲートウェイを対象に、分散処理環境下におけるシステム評価を行った。また、GEO-J プロファイル修正案及びメタデータプロファイル改定想定案に基づいて編集されたモデル地区のメタデータ、サンプル画像データを使用して、ユーザを想定したシステム検証を行った。

またシステム評価の一部として、今年度事業で改善した Isite によるメタデータサーバと、昨年度事業で開発した METIS によるメタデータサーバとの応答時間特性における比較評価を行った。

米国における GEO 改定動向に関する調査

国内向け地理情報プロファイル（GEO-J プロファイル）の修正案を作成するため、米国 FGDC（地理情報委員会）において、日本国内における状況及び本事業の実施について説明するとともに、ISO19115 を中心とした TC211 における標準化の動向について意見交換し、本調査実施の一助とした。

（２）GEO-J プロファイル修正案

１）国際標準化動向の確認

平成 12 年度事業において作成した、国内向け地理情報プロファイル案（GEO-J プロファイル案）の修正のため、米国 FGDC（連邦地理情報委員会）において、日本国内における状況及び本事業の実施について説明するとともに、地理情報メタデータ標準である ISO19115 を中心とした TC211 における標準化の動向について意見交換し、本調査実施の一助とした。

FGDC との意見交換

ISO23950 を採用した米国の地理情報クリアリングハウスについて、平成 13 年 6 月 7 日～6 月 13 日の期間に FGDC に調査団を派遣し以下の事項について意見交換を行い、国内における同システムの在り方を検討するための参考意見とした。

- a) 平成 13 年度の経済産業省クリアリングハウス事業の概要について
- b) 平成 13 年度事業で採用予定の Isite 改善版について
- c) GEO プロファイル Version 3（現行 V2.2 の改定版）の改定動向について
- d) OGC Catalog Services について
- e) FGDC におけるクリアリングハウス事業の支援方針について
- f) ISO19115 を中心とした TC211 における標準化の動向について

意見交換の結果

平成 13 年 6 月の時点で想定していた FGDC の標準改定計画の進捗は大幅に遅れていた。原因の一つに、ISO19115 の DIS 発行が遅れ、平成 13 年 9 月になってしまったという事情がある。これに伴い、FGDC が外部企業に委託している移行ツールの開発も遅れていた。このような状況の中で、FGDC は平成 14 年の早い段階で、FGDC のメタデータ仕様の次期版と GEO プロファイル V3.0 の策定をしたいということであった。また、FGDC は整備済みの CSDGM(Content Standard for Digital Geospatial Metadata)に基づくメタデータと ISO19115 DIS のメタデータを、各メタデータ仕様の Harmonization Project の一環として共存する形で運用することを計画している。なお、ISO19115 ベースの新しいメタデータ標準は、FGDC プロファイルという位置付けになる。

f) について、FGDC は、ISO 規格の制定が間近に迫っていることと、国際的な地理情報メタデータの流通の観点から、従来の米国の国内規格である CSGDM から ISO 規格への移行作業に取り組んでおり、ISO 規格を前提とした日本の取り組みについて関心を寄せているようであった。

NSDI(National Spatial Data Infrastructure)は毎年、GIS 関連の事業に対して補助金による支援を行っている。このうち、地理情報メタデータの作成とクリアリングハウスと Web サービスとの統合に関する事業が占める割合が大きいということであった。地理情報メタデータの整備により、GIS が継続的に利用できるという点が重要視されている印象を受けた。

2) GEO-J プロファイル修正案

修正方針

FGDC により策定された GEO は、地理情報メタデータに関する米国の国内規格である Content Standard for Digital Geospatial Metadata (CSDGM) を取り込む形式で策定されている。一方、GEO-J(案)は、CSDGM ではなく ISO19115 を取り込む形式で作成する。この理由は、日本国内の地理情報標準に含まれる地理情報メタデータ規格が策定中の ISO19115 をベースに開発されており、今後も ISO19115 に準拠する形式で国内の地理情報メタデータ規格が策定されるためである。

今年度事業では、平成 12 年度事業において作成した国内向け地理情報プロファイル案 (GEO-J プロファイル案)について、FGDC が 1994 年 6 月に発行した電子的地理空間メタデータのための内容規格[2]の応用プロファイル GEO V2.2(現行バージョン)を参考とし、更に現在 ISO/TC211 で策定作業中である地理情報メタデータの国際標準規格 ISO19115(現時点で DIS)をベースに修正した。

GEO-J プロファイルの JIS 化案

GEO-J プロファイル修正案をもとに、JIS 化を想定した原案を作成した。原案作成にあたっては、JIS の様式及び表記・表現規則に一致させることを中心に作業し、内容については、日本固有の部分（例：国内地理情報標準）などはそのまま表記し、その他の部分については技術的に一致させた。

JIS 化案の構成を次に示す。

本文

序文

- 1．適用範囲
- 2．引用規格
- 3．定義
- 4．適合性
- 5．GEO-J のための (Z39.50) プロトコル仕様

附属書 1 (規定) GEO-J 属性集合

附属書 2 (規定) GEO-J スキーマ

上記構成のうち、本文「3.定義」については、この規格案作成の基となった GEO で記載されている用語から、GEO-J で利用していない用語を省いて定義した。しかしながら規格上に定義されている用語は米国と日本では異なるため、他の JIS 規格で定義されていない、あるいは異なる意味で使用されている用語については、別途定義を追加する必要がある。これらについては、プロトコル仕様が最終的に固まった段階で、定義の追加を行う必要がある。なお、上記作成方針に基づき作成した本案は、今後本格的な原案作成を進めるための事前準備作業の成果である点について付記する。

GEO-J プロファイル修正案と今後の課題

平成 12 年度事業で作成した GEO-J プロファイル案は、ISO19115 CD2 を基に策定された国内地理情報標準の JMP1.1 を採用しているのに対し、平成 13 年 9 月の ISO19115 の DIS 化に伴い、従来の ISO19115 CD2 が大きく変更されたため、近い将来に正式に IS になることを考慮して、ISO19115 DIS をベースに GEO-J プロファイル案を修正することとした。

一方、FGDC において策定されている GEO プロファイルについても ISO 規格を採用する形式が検討されており、本事業において策定した GEO-J プロファイル修正案と整合することが必要になると考えられる。

本事業では、ISO/TC211 で策定中の地理情報メタデータの国際標準規格 ISO19115 を DIS の段階から先行的に取り込む形で GEO-J プロファイル案を修正することにより、近い将来、国内はもとより、国際的な地理情報メタデータのクリアリングハウスの相互連携を可能に

することを目指している。

このためにも、今後、GEO-J プロファイル修正案の JIS 化に向けた作業に取り組み、今回の修正作業で作成された仕様や、明らかになった問題点を解決しつつ、最終的な原案における技術的な完全性を追求することが必要である。同時に TC211 などの国際機関の動向をフォローしながら、国際的な整合性を図る必要がある。また、今後の普及に向けて、実装者の便宜を図るため、本規格の技術的な解説、応用例を附属書として記載する必要があると考えられる。

(3) メタデータの調査分析・編集

1) 収集済みメタデータの整備方針

昨年度調査では、豊中市、大分県、臼杵市の3モデル地区が保有する地理情報を調査し、メタデータを収集・作成した。当該メタデータの作成のために採用したメタデータプロファイル標準に関しては、米国 FGDC が採用している CSDGM を採用する案、国土交通省国土地理院による JMP1.1 を採用する案、両者から必要なデータを抽出する案等が検討されたが、最終的には JMP1.1 をベースとしたメタデータプロファイルを採用することとした。

ただし、JMP1.1 では定義されていないデータ項目として、参照画像情報を追加し、当該調査においてメタデータを整備する際、それぞれのメタデータに対応する地図等のサンプル画像データも同時に整備した。これによりユーザが、クリアリングハウスを利用する際、メタデータを検索し、対応する実データ（ここではサンプル画像データ）にシームレスにアクセスし、閲覧できるようにした。

今年度は、ISO/TC211 のメタデータ標準 (ISO19115) が DIS (Draft International Standard) となり、それに基づいて JMP の改定が予定されている。それに歩調を併せて3モデル地区の収集済みメタデータの調査分析・編集を行い、実証実験用のメタデータとして、対応するサンプル画像とともに実証実験システムに蓄積し検索・閲覧実験を行った。

収集済みのメタデータ編集・加工作業の具体的手順としては、JMP の改定に併せメタデータプロファイルの修正版を作成し、これに対応したメタデータの XML-DTD を作成した。更に昨年度使用したメタデータの XML-DTD から今年度改定した XML-DTD への変換プログラムを開発し、収集済みメタデータを XML-DTD 改定版対応のメタデータに変換した。

2) メタデータプロファイルの改定

地理情報メタデータ標準の動向

ISO/TC211 において検討されているメタデータ (Geographic information – Metadata ISO19115) は、平成 13 年 9 月に DIS が提出され、FDIS のための投票が行われている。この投票については、日本からも、ドラフトに対する 100 以上の問題点を挙げ、かつ幾つかの代替案を提案しており、賛成票については、その代替案の採用を条件にしている。最新の国際規格案 (ISO19115 DIS) では、メタデータのうちクリアリングハウスの運用に主要な要素をコアメタデータとして定めている。コアメタデータにはデータ集合の題名、範囲、要約、解像度のための情報が含まれている。

一方、国内では、国土地理院が主宰する官民連帯共同研究「地理情報標準の運用に関する研究」の中に設置された地理情報標準検討部会 WG3 において、地理情報メタデータ標準が策定されている。策定されたメタデータ標準は、日本版メタデータプロファイルとして公開されており、現在の最新バージョンは JMP1.1a である。

地理情報標準検討部会 WG3 の今年度作業において、現行の JMP1.1a が JMP2.0 に改定される予定であったが、平成 13 年 9 月に ISO/TC211 よりメタデータに関する DIS が提出されたため、JMP2.0 への改定を行わず、DIS にあるコアメタデータを採用する方向で検討を行っている。

メタデータプロファイル

今年度事業の当初予定では、事業実施期間内に国内地理情報メタデータ標準が現行の JMP1.1a から JMP2.0 に改定されることを想定していたが、上記事情により、本事業のメタデータ編集・加工作業で使用するメタデータプロファイルのベースを、地理情報標準検討部会 WG3 が採用する予定である ISO19115DIS のコアメタデータとすることとした。

3) メタデータの XML-DTD

今回のメタデータ変換作業においては、メタデータ変換を 2 段階に分けて行うため、それぞれの段階に対応する 2 種類の DTD を作成し利用した。すなわち、JMP1.1a に参照画像情報を追加したメタデータに適用する DTD (JMP1.1 ベースのメタデータから JMP1.1a ベースへの変換用) と、ISO19115 DIS に適用する DTD (JMP1.1a ベースのメタデータから ISO19115 DIS ベースへの変換用) である。

ISO19115 DIS ベースのメタデータの XML-DTD は、ISO19115 DIS で定義されている Comprehensive metadata を採用した。ただし、ISO19115 DIS の中で定義されている Comprehensive metadata の DTD には、若干の文法的な不具合があるため、XML パーサで文法チェックをした場合、“妥当でない (not validated)” エラーが発生することから、文法

的な不具合を解消するための修正を加えた。

4) 収集済みメタデータの変換ツールの開発

変換ツールは、JMP1.1 にサンプル画像情報を付加した仕様の XML ファイルを ISO 19115DIS におけるコアメタデータへ変換するものである。変換ツールの作成は、以下の手順で行った。

- a) 変換用 XSLT (eXtensible Stylesheet Language Transformations) スタイルシート作成
- b) XSLT プロセッサ、XML パーサプログラム、文字コード変換プログラムを統合した変換用プログラム作成

変換作業は、以下の2段階に分割して行うこととした。

- 作業1 JMP1.1 形式のメタデータから JMP1.1a 形式のメタデータへの変換
- 作業2 JMP1.1a 形式のメタデータから 19115DIS コアメタデータへの変換

したがって、XSLT スタイルシートもそれぞれ作成することとした。また、変換対象となる XML ファイルの数が多い(約 2,000 ファイル)ことから、バッチ処理を可能とする変換ツールを開発した。

メタデータの変換処理自体は、変換用 BAT ファイルの作成を除き、全てをメタデータ変換ツールによって自動的に行うことができた。メタデータ変換ツールについては、特に高いスペックのパソコンを必要とするわけではないが、IE 等のバージョンが低い場合には IE のバージョンを上げる、或いは動作に必要なアプリケーションの導入が必要になる。しかしながら、今回開発したメタデータ変換ツールは、IE については、現行の最新バージョン及び1つ前のメジャーバージョンまでをサポートしている。従って、1つ前のメジャーバージョンであれば IE 以外のアプリケーションの導入は不要である。変換作業においては、特に変換エラー等もなく良好に動作した。

(4) クリアリングハウスの構築及びシステム側実証実験

1) クリアリングハウスの構築

システムの構築方針

本事業では、地理情報メタデータの国際標準規格として策定中である ISO19115 の動向及び国内メタデータ標準仕様である JMP の改定動向を考慮しつつ、昨年度に作成した GEO-J プロファイル案の修正及びそれに準拠する地理情報クリアリングハウスシステムを構築し実証実験を行った。また、本事業ではクリアリングハウス用ノードサーバソフトとして Isite を改善して使用した。Isite とは、CNIDR(Clearinghouse for Networked Information Discovery and Retrieval) が米国 NSF (National Science Foundation) の助成を受けて

開発した汎用的な無償 ISO23950 パッケージシステムであり、FGDC はクリアリングハウス用のノードサーバで採用している。今回の実証実験システムのノードサーバ部分には、Isite に内在する問題点を解決し、GEO-J プロファイル案修正版に対応したシステム(Isite 改善版)を使用した。

なお、今年度事業で構築する実証実験用の地理情報クリアリングハウスは、ゲートウェイサーバとして昨年度事業で開発した METIS のゲートウェイサーバ機能を使用し、クリアリングハウス用ノードサーバには、上記 Isite を改善してシステムを構築した。さらに、当該システムに、GEO-J プロファイル修正案及び今年度事業用に編集・加工した 3 モデル地区のメタデータを格納し、ユーザ側からはインターネットを經由して当該メタデータの検索及びサンプル地理情報画像の閲覧を行った。これにより、本事業で構築した実証実験用地理情報クリアリングハウスのシステム側実証実験を行うとともに、GEO-J プロファイル修正案の確実性、システム運用面、データ整備面における課題等を検証した。

Isite の問題点と改善版の特長

現在、Isite にはシステム上の問題がいくつか存在しており、改善版ではこれらに対して表 2-1-1 に示すとおり対応した。

表 2 - 1 - 1 Isite の問題と対応

問題点	対応
ISO23950 処理部の不完全実装	ISO23950 処理部を別モジュールに変更
ISO23950 規格違反の実装	新規にノードサーバを作成
プログラム上のバグ	修正
検索処理部のパフォーマンス低下	検索効率の改善
検索処理部の不完全及び未実装機能	近接検索、検索結果の再検索、シソーラス検索、検索結果管理の各機能の正式な実装

上記に加え、GEO-J プロファイル案及び日本語処理として表 2-1-2 の対応を行った。

表 2 - 1 - 2 Isite 改善版における GEO-J プロファイル案と日本語対応

項目	内容
GEO-J プロファイル案の対応	処理モジュールを開発
日本語解析	日本語形態素解析システム「茶筌」を導入
日本語文字コード	JIS, Shift-JIS, EUC, UTF-8 に対応
日本語文字コードのクライアント・ノードサーバ間交渉	ISO23950 に追加承認されている規約を実装

機能構成、稼働環境及び使用ツールは表 2-1-3 に示す通りである。なお、今回の事業において改善対象である Isite は、オリジナル版 (Version 2.07 i) である。

表 2 - 1 - 3 Isite 改善版における機能構成、稼動環境及び使用ツール

項目		内容
Isite オリジナル版 (CNIDR)		Version 2.07i
データベース作成機能		Isite が用意する機能及び GEO-J プロファイル案に対応する。
ISO23950 ノードサーバ機能		Isite の問題点を解決し、適切な ISO23950 機能及び改良した検索機能を提供する。
ISO23950 クライアント機能		ノードサーバの機能検証を目的とする。
稼動環境 (推奨)	CPU	Intel PentiumII 300MHz 以上
	メモリ	256MB 以上
	HD	20MB。この他にデータベースに別途必要。
	OS	Windows 2000 SP1, Windows NT 4.0 SP6 Solaris 8 FreeBSD 4.2-RELEASE 以上
ツール	Perl	プログラミング言語。Version 5.005_03 以上
	Java	J2SE Version 1.3.1 以上
	XSLT	W3C 勧告 XSLT スタイルシート処理系。Version 1.0 準拠であること。Xalan-Java 2.1.0 以上等。

上記対応により、Isite 改善版は、改善前の Isite と比較してシステム面の信頼性が向上し、運用性の高いシステムとなることが期待される。なお、2001 年 6 月に上記 Isite の後継版である Isite2 が公開されているが、上記現行版と比較して、処理効率と運用性が低下（投入データの量的特性において急速に破綻傾向を示す）する傾向がある。従って、今回は Isite 現行版を改善対象とした。

実証実験システムの機能要件

今年度の事業では、クリアリングハウスのノードサーバソフト用には Isite 改善版を開発し導入した。またゲートウェイサーバには、昨年度事業において開発した METIS システムのゲートウェイサーバ機能を導入した。なお、GEO-J プロファイル修正案を適用した上記クリアリングハウスの構築にあたり、実証実験システムの機能要件を以下のサーバ機能別に定義した。

GEO-J プロファイル修正案を適用するノードサーバ (Isite 改善版)

GEO-J プロファイル修正案を適用するゲートウェイサーバ (METIS システム)

a) ノードサーバの機能要件

ノードサーバは、GEO-J プロファイル修正案で規定される仕様に対応することが求められる。なお、Z39.50 情報検索プロトコルの各機能は、GEO-J プロファイル修正案において規定され、Isite 改善版により実現する。

- ・ 検索機能と連携するデータベース作成機能を有すること。
- ・ ISO19115 DIS で規定される XML 形式のメタデータを扱えること。

- ・文字列型、数値型、日付型の検索機能を有すること。
- ・項目別及び全文の検索機能を有すること。
- ・GEO-J プロファイル修正案に規定される Z39.50 機能に対応すること。
 - ・Init、Search、Present の各機能に対応すること。
 - ・属性集合に対応すること。
 - ・XML、HTML の各レコード構文に対応すること。
 - ・B、S、F、A の各要素集合名に対応すること。
- ・UTF-8 日本語文字コード系に対応すること。
- ・ノードサーバ運用時の接続情報をロギングすること。

b) ゲートウェイサーバの機能要件

ゲートウェイサーバは、GEO-J プロファイル修正案で規定される仕様に対応することが求められる。なお、Z39.50 情報検索プロトコルの各機能は、GEO-J プロファイル修正案において規定され、METIS システムのゲートウェイサーバ機能により実現する。

- ・GEO-J プロファイル修正案に規定される Z39.50 機能に対応すること。
 - ・Init、Search、Present の各機能に対応すること。
 - ・属性集合に対応すること。
 - ・XML、HTML の各レコード構文に対応すること。
 - ・B、S、F、A の各要素集合名に対応すること。
- ・UTF-8 日本語文字コード系に対応すること。
- ・ゲートウェイサーバ運用時の接続情報をロギングすること。

c) 機能・データ連携

実証実験システムの機能及びデータの連携を、図 2-1-3 “機能・データ連携” に示す。

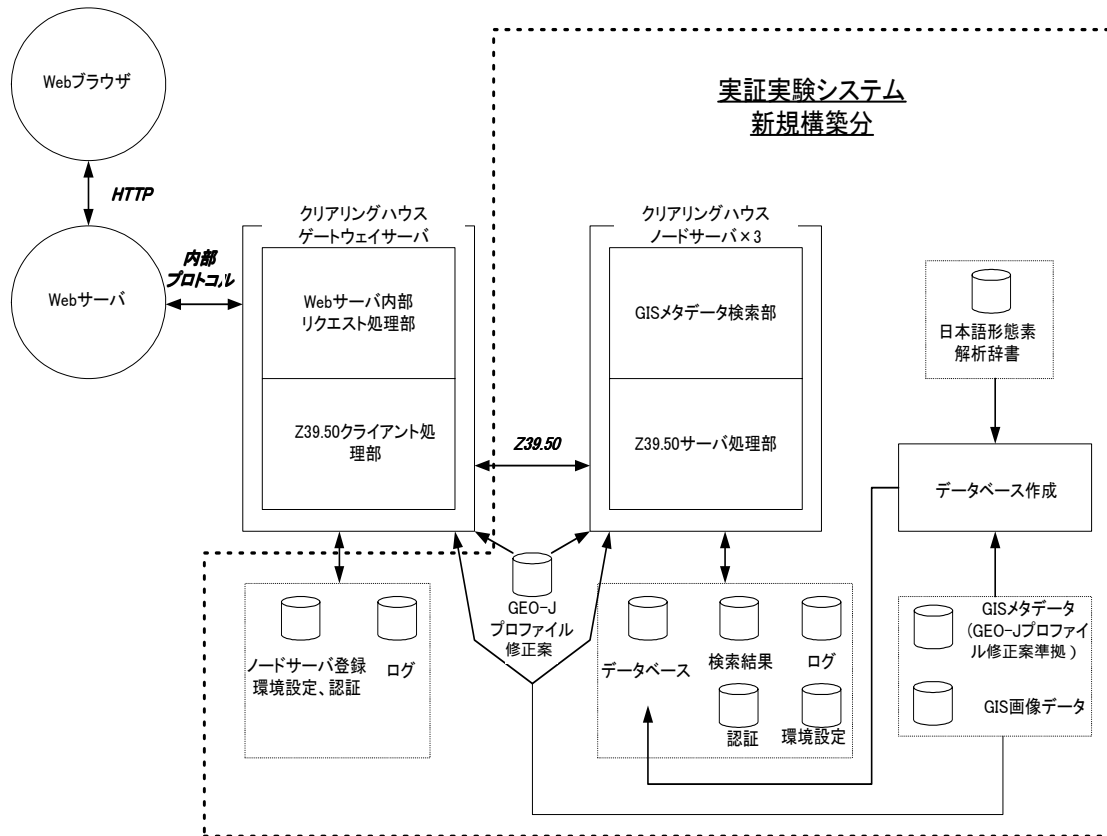


図 2 - 1 - 3 機能・データ連携

d) システム構成とネットワーク環境

本実証実験のシステム環境を構築するにあたり、3 台のクリアリングハウス用ノードサーバを東京サイト（財団法人データベース振興センター内に 2 台）と、富山サイト（インテック・ウェブ・アンド・ゲノム・インフォマティクス株式会社内に 1 台）の 2 つのサイトに分散配備した。また、Web サーバ及びゲートウェイサーバは富山サイトに配備した。なお、クライアント環境には、3 モデル地区の地方公共団体等において職員が日常業務で使用しているパソコンを想定しており、ブラウザソフトウェアであるインターネットエクスプローラ、ネットスケープによりインターネットにアクセスできるマシンを前提とした。実証実験システムの構成は、図 2-1-2 “(1) 事業の目的と課題” に示すとおりである。

e) データ環境

本実証実験において、検索対象として使用する地理情報メタデータは、平成 12 年度事業においてモデル地区である大分県、臼杵市、豊中市の 3 つの地方公共団体から収集し作成したものである。当該メタデータは国内メタデータ標準 JMP1.1 準拠のメタデータであるが、本事業では最新の標準化動向に対応するため、ISO19115 DIS（平成 14 年 3 月現在で最新）準拠のものに変換して使用した。なお、昨年度事業では、実証実験期間中に当該クリアリ

ングハウスへの不正アクセスを排除するためユーザ ID 及びパスワードによる認証の設定を行ったが、今年度の実証実験期間中においても、同様に認証を設定した。この認証は、ノードサーバ及びゲートウェイサーバの両方に対して設定した。

本実証実験環境に投入したメタデータ及びサンプル画像データのモデル地区毎の格納件数を表 2-1-4 に示す。

表 2 - 1 - 4 データ格納件数

地区名	メタデータ件数	画像データ件数
大分県	62	18
臼杵市	92	72
豊中市	1,958	848

データベースと格納データの構成を表 2-1-5 に示す。東京サイトに 1 台及び富山サイトに 2 台のノードサーバに、モデル地区毎に独立して作成した 3 つのデータベースを配置した。

表 2 - 1 - 5 データベース配置とデータ構成

地域名	データベース名	東京サイト	富山サイト
大分県	oita		
臼杵市	usuki		
豊中市	toyonaka		

注) 印は、地区及びデータベース名と対応するサイトを示す。

f) 日本語形態素解析

日本語は、英語に見られるような分かち書きスタイルではない。この点を考慮して、本実証実験システムでは、XML に記述されるメタデータを日本語形態素解析技術を用いて、品詞及び品詞細分類のレベルで自動的に区切る手法を導入している。これにより抽出された索引語を、その出現位置の情報を含めてインデックスとしてデータベースに格納することで、クリアリングハウスサーバの検索機能において、文字の位置情報を活用したフレーズ検索を可能としている。

日本語形態素解析技術には、奈良先端科学技術大学院大学で開発されたソフトウェアシステム「茶筌第 1.51 版」を利用した。

なお、本実証実験システムにおいては、全ての品詞及び品詞細分類について索引語に採用する設定とした。茶筌品詞分類ファイルの内容を図 2-1-4 に示す。

;;; 形態素解析システム茶筌：品詞分類ファイル
;;;
;;; 1990/11/14/Wed Yutaka MYOKI
;;; 1991/07/26/Fri Yuji MATSUMOTO
;;;

((特殊)

((句点)

(読点)

(空白)

(括弧開)

(括弧閉)

(記号))

((動詞 %))

((形容詞 %))

((判定詞 %))

((助動詞 %))

((名詞)

((普通名詞)

(サ変名詞)

(固有名詞)

(地名)

(人名)

(数詞)

(形式名詞)

(副詞的名詞)

(時相名詞))

((指示詞)

((名詞形態指示詞)

(連体詞形態指示詞)

(副詞形態指示詞))

((副詞)

((様態副詞)

(程度副詞)

図 2 - 1 - 4 茶筌品詞分類ファイルの内容(1/2)

(時制相副詞) ; テンス・アスペクト副詞 (陳述副詞) (評価副詞) (発言副詞)) ((助詞) ((格助詞) (副助詞) (引用助詞) (名詞接続助詞) (述語接続助詞) (終助詞)) ((接続詞) ((連体詞) ((感動詞) ((接頭辞) ((名詞接頭辞) (動詞接頭辞) (イ形容詞接頭辞) (ナ形容詞接頭辞)) ((接尾辞) ((名詞性述語接尾辞) (名詞性名詞接尾辞) (名詞性名詞助数辞) (形容詞性述語接尾辞 %) (形容詞性名詞接尾辞 %) (動詞性接尾辞 %)))

図 2 - 1 - 4 茶筌品詞分類ファイルの内容(2/2)

2) システム側実証実験

実証実験期間中のアクセス総数

本実証実験は、平成 14 年 3 月 4 日～平成 14 年 3 月 8 日の期間で実施した。想定ユーザとしての実証実験参加者は、(財)データベース振興センター並びに本事業への協力企業であるインテック・ウェブ・アンド・ゲノム・インフォマティクス(株)、国際航業(株)、神鋼リサーチ(株)、(株)システムソフトを含むクリアリングハウス調査委員会関係者であった。本実証実験におけるゲートウェイサーバへのアクセス状況を表 2-1-6 に示す。

なお、ゲートウェイと連携するクリアリングハウスサーバは、Z39.50 認証機能を使用することで、実証実験評価時にはゲートウェイ経由でのみ使用する形態とした。

表 2 - 1 - 6 実証実験期間中のアクセス件数

プロトコル機能	データベース	アクセス件数
Z39.50 Search リクエスト数 (検索要求件数)	リクエスト総数	3,435
	大分県	1,111
	臼杵市	1,040
	豊中市	1,284
Z39.50 Present リクエスト数 (メタデータ取得件数)	リクエスト総数	1,633
	大分県	469
	臼杵市	396
	豊中市	768
HTTP による画像データ取得件数	リクエスト総数	129
	大分県	37
	臼杵市	51
	豊中市	41
アクセス総数		5,197

システム運用側のシステム評価

昨年度事業において、システムの応答時間等のパフォーマンスに関するシステム側の評価を行ったが、インターネットを経由したネットワークトラフィックに依存したシステムパフォーマンスへの影響があまりに大きく、システム単体の応答特性の計測以外は、測定結果にあまり意味が無いとの指摘が、委員会においてなされていた。

そこで今回は、システム側の評価として、昨年度事業の成果の METIS システムと今年度事業の成果である Isite 改善版システムとの応答特性に関する比較を行うこととした。

この場合、有意な比較として対象となるものは、実装している各 Z39.50 機能の応答特性である。なお、Z39.50 機能を担当するプログラムは、Isite 改善版システムと METIS システムで共通プログラムを使用しているため、Init (初期化) 機能及び Present (返戻) 機能については応答特性に差が生じず、応答特性に差が生じる機能は Search (検索) 機能ということになり、検索機能と連携している Isite 改善版と METIS の各検索エンジンの性能差が大きく反映されることになる。

そこで本実証実験では、上記視点に基づくシステム側の応答特性に関する評価を行うこととした。これはノードサーバの検索機能を実現する、それぞれの検索エンジンの応答特性を比較評価することになる。この時、ネットワークトラフィックによる測定値への影響を排除するため、Isite 改善版システムと METIS システムの両方を同一のマシンで稼働させ、同一のローカルマシン上で計測用の Z39.50 テストクライアントを用いて測定し比較評

価することとした。

この値の比較において有意と考えられる評価項目として、Isite 改善版システムと METIS システムの各ノードサーバ検索機能の応答時間と、1 検索式の個数との関連に着目することとし、評価にあたっては、3 モデル地区のうち、メタデータ数が最も多い豊中地区のデータベースを対象とした。

評価の結果、1 検索式（例えば、全文 = 大阪府）の個数とサーバ応答の遷移は、Isite 改善版システムと METIS システムのそれぞれについて、表 2-1-7 に示すとおりとなった。

表 2 - 1 - 7 ノードサーバ検索機能の応答時間と 1 検索式の個数との関連

1 検索式の個数	1	2	4	8	16	32
Isite 改善版システム 平均検索機能の応答時間(秒)	0.25	0.49	1.04	2.26	5.02	11.12
METIS システム 平均検索機能の応答時間(秒)	0.28	0.49	0.93	1.81	3.57	7.02

これは、論理演算子 (AND,OR,NOT) による差異を受けずに、概ね上記のとおり推移を見せている。ただ、メタデータ件数が 1,958 件と少ないこともあり、各システム毎に 1 検索式の個数と検索機能の応答時間に明瞭な関連を見出すには到っていない。また、Isite 改善版システムと METIS システムの検索エンジンの特性から、検索機能の応答時間は METIS システムの方が短いという比較結果となった。

その原因の一つとして、Isite 改善版システムの検索エンジンは Isite オリジナルの検索エンジン及びデータベース仕様を受け継いでおり、検索処理時に常にキーワードの出現位置を調査し、メタデータファイル内を seek して照合処理を行っていることから、この部分の効率面の差が、性能差 (Isite 改善版システムは METIS システムより若干長い応答時間であった) として現れたものと考えられる。

3) ユーザ側によるシステム検証

ユーザ側によるシステム検証では、メタデータ変換ツールを用いて変換を行った新規格のメタデータについて、実際にクリアリングハウスに登録し使用できるかどうかを系統的に検証した。具体的評価方法としては、検索用に構築したユーザ側の環境から、任意に抽出したメタデータを基に様々なキーワードにより実際に検索し、ヒットの可否とその検索結果を詳細に分析した。これにより、ユーザ側環境構築時の問題点の把握、各種検索条件での検索結果の妥当性評価、応答速度や検索上の問題点の把握等、ユーザ側から見たシステム側、つまり総合的なクリアリングハウス機能を確認した。

上記のシステム検証の結果、ブラウザ～ノードサーバへの一連の検索過程で一部 UTF-8 文字コード系による処理に問題が見受けられたが、検索機能そのものは概ね良好な精度を示した。この文字コード系による処理の問題とは、具体的には、電話番号、住所、値に全角のハイフン“ - ”が含まれるキーワードに対しては正常な検索が行われなかった、などである。今後のメタデータ整備において、全角・半角の取り扱いに関し、統一的な整備方針をマニュアルとして作成するなどの対応が必要である。

4) 今後の期待

今回の事業において、ISO19115DIS に基づく GEO-J プロファイル修正案を作成し、これに準拠するクリアリングハウス環境を先行的に構築し実証実験を実施したことは、近い将来、ISO19115 に準拠する国内外の地理情報クリアリングハウスの広域連携を図る際に、有意な経験となることが期待される。