

ディレクトリを用いた情報提供に関する一手法 ～地図情報への適用～

星野 隆, 長谷川 靖, 池田哲夫, 星 隆司

NTT サイバースペース研究所

{hoshino, hasegawa, ikeda, hoshi}@dq.isl.ntt.co.jp

近年, GIS (Geographic Information System) 環境を容易に構築・利用できるようになってきた。それに伴い, 地図とそれ以外の情報とを組み合わせ情報提供を行うシステムが望まれている。このような情報提供システムの構築において, 既存の GIS やデータベースなども利用できることが望まれる。さらに, ユーザの検索の利便性を高めるため, 地図そのものを検索インタフェースとすることも望まれている。

そこで, 本稿では, 地図とそれ以外の情報を関連づけ, 既存 GIS や既存のデータベースや Web ページなど複数の異種情報源の利用可能とし, 地図を検索のインタフェースとして利用可能とする情報提供方式を提案する。本方式は, 分散情報資源の管理を分類木で行う, 具体的には LDAP をインタフェースとするディレクトリを用いる方式とし, 既存の分類木の統合には筆者らの提案したディレクトリビュー機構を用い, また, 地図をディレクトリとして管理し, 地図以外の情報源との連携を可能とする。

An Information Sharing Method using Directory -- For Geographic Information --

Takashi Hoshino, Yasushi Hasegawa, Tetsuo Ikeda, Takashi Hoshi

NTT Cyber Space Laboratories

{hoshino, hasegawa, ikeda, hoshi}@dq.isl.ntt.co.jp

Recently, the GIS (Geographic Information System) environment has come to be able to be constructed and used easily. The information sharing system that combines GIS with information sources is needed. In general, the GIS and those sources have already existed. In addition, because the benefit and convenience of the retrieval of the user is improved, the retrieval system by which the classified tree and the map are made a retrieval interface is hoped.

In this paper, we propose an information sharing method, whose retrieval interface is classified trees and maps, using directories that manage information sources.

Our method manages distributed information resources using classified tree and adopts a directory, whose interface is LDAP, as a classified tree. A directory view mechanism integrates classified trees. A map is also managed as a directory.

1 はじめに

近年, GIS (Geographic Information System) 環境を容易に構築・利用できるようになってきた。これまでの限られたユーザだけではなく, 広く一般にも GIS や地図を利用したシステムが構築されつつある。従来 GIS は, 施設管理や地籍管理など, 空間データとそれに付随する属性データを管理するシステムとして構築されていた。現在はこのような情報管理システムにとどまらず, GIS, 地図とさまざまな情報を組み合わせた, 情報提供システムが要望され, 実現されつつある。

GIS の普及にともない, 複数の箇所で GIS システムが利用され, それらを連携した利用が望まれている。自治体においても, 情報の管理主体が各部署毎でもあり, 個別に GIS システムが導入されており, これら複数の GIS システム間の情報流通が問題となっている [1]。

ここで, 情報提供システムという視点から流通したい情報は空間データではなく, むしろ属性データである場合が多いと考える。空間データである地図は, 情報の背景または情報を探索する手段となるためである。

さらに GIS 以外のデータベースや Web ページに代表される既存の情報源でも, 地図と対応づけ情報を提供したいという要望がある。すでに作成されたデータベースを活用し, 効率的な情報提供を行うためである。

一方, 情報提供システムにおいては, ユーザが容易に情報を検索できることが必要である。ユーザの検索インタフェースとして分類木がよく利用されている。分類木検索は, 多様なジャンルにわたる大量の情報の中から, 容易に能率よく必要な情報を検索するための有力な 1 手法である [2]。Internet で利用されている地図情報検索サイトにおいても数多く利用されている。さらに, Web サーチエンジンにおい

ても, コンテンツの質の高さや, 初心者でも利用が容易なことから, 多くのロボット型サーチサイトでも, 分類木を用いた仕組みを合わせて提供している [3]。このことから, 情報提供において, 分類木の利用は有効であるといえる。

地図と関連づけ情報提供を行う場合, ユーザによる検索を容易にするためには, 分類木だけではなく, 検索インタフェースとして地図を利用できると望ましい。この場合, 表示されている地図に対して, 分類木を利用したり, 地図の表示範囲を移動させたりして情報を検索する。

このように, 地図と情報を関連づけ情報提供するにあたり, その情報源も既存の GIS や既存のデータベースや Web ページなど, 複数の異種情報源の利用が望まれている。さらに, ユーザの検索を容易にするため, 分類木の利用や, 地図を検索のためのインタフェースとして利用することも望まれている。そこで本論文では, 複数の情報源を分類木により管理・検索可能とし, 地図と対応づけ, 分類木や地図からそれら情報の検索を可能とする方式を提案する。

以下, 2 章で地図を用いた情報提供システムにおける, 複数情報源の管理・検索に関する問題点をあげ, 3 章で解決へのアプローチを述べる。4 章では提案する方式を示し, 5 章で方式の一実装を述べ, 6 章ではまとめと今後の課題を示す。

2 問題点

分類木と地図を利用した, 複数の情報源で構成する情報提供システムを構築する際の問題点を以下に述べる。

2.1 複数情報源管理

分類木から情報を検索する場合, 分類に対応する情報源を検索するため, 情報源はそれがもつ情報の種別毎に分類木の部分木と対応づけ管理されている。このとき, 1 つの分類木で複

数の情報源を扱うことになる。同様に、1つの分類木ですべての範囲を網羅できないときには、複数の分類木を組み合わせる必要がある。そのため、情報源レベルでの統合と分類木レベルでの統合の問題がある。

2.1.1 情報源

情報源の統合検索では、異種性の解消が解決すべき問題である。[4]で分類されているように、情報源の異種性は、多種、多様である。

2.1.2 分類木レベル

既存の分類木は独自の検索インターフェースで作成されている。そのため、共用化がこれまで困難であった。

それぞれの分類木は独自の分類体系に基づき作成される場合が多い。情報源と深く関連付けられているため、自分の情報源がもつ情報の範囲の分類だけが作成されている。たとえば、特定の地域に関する住所の分類木や、目的に応じた施設の分類のみが作成されている。逆に、共用しようとした場合、不要な分類やレイヤを持つものなどがある。

2.2 検索インタフェースとしての地図

表示する地図はこれまでラスタ形式の地図が用いられる場合が多く、地図から情報を検索するとき、表示領域の周辺を検索する場合に地図が途切れてしまうなど、検索が容易ではなかった。また、用意されている画像に制限があり、所望の領域や縮尺での表示が困難であった。さらに、従来のラスタ形式では、地図の縮尺、範囲が実際に表示している地図と一意に対応づけられていた。

ラスタ形式に代り、ベクトル形式の地図をインタフェースとすることが考えられる。この場合、連続的な表示領域の変更が可能であり、ユ

ーザが地図を移動させて情報検索を行う場合有利である。ベクトル形式の地図を検索インタフェースとして用いるには、縮尺や表示領域の変動に対して、地名などが追従しないという問題を解消する必要がある。

3 アプローチ

3.1 分類木と情報源

分類木を検索インタフェースとすることに対応し、複数の情報源を分類木により管理する。このとき複数の情報源に対する検索が必要であり、複数の情報源間の異種性を解消する必要がある。そのため情報源の仮想的統合を行う。

既存の分類木を用いる場合、複数の分類木を利用することが必要となる。そこで、共用しようとする複数の分類木の仮想的統合を行う。

また、既存の分類木がそのまま利用できる場合は少ないと考える。既存の分類木には不要な分類項目や不要な階層が存在すると考えられる。そこで、複数の分類木を統合する際には、既存の分類木の一部分を切り出すことができる必要がある。

3.2 地図インタフェース

ユーザの検索の便宜を図るため、ベクトル形式の地図を利用する。このとき縮尺や表示領域の連続的な変化に応じて、表示している位置や地名を明確にすることが必要である。そこで、住所も地図もある領域を持ち、階層構造で管理できるという類似性に着目し、これら縮尺、地域、地名の情報を住所をあらわす分類木により管理することとする。

4 提案方式

4.1 分類木と情報源

情報源、分類木を共用化するため、それぞれのレベルでの共用化が必要となる。さらに、分

類木と情報源との対応付けを行うことも必要である。これらを実現するため、それぞれのメタ情報を管理する。以下、それぞれの方式について説明する。

4.1.1 情報源の共用化

複数の情報源を利用するためには、既存技術であるメディアータを利用する。本方式では、これまで我々のグループで研究開発してきた DBSENA[6]を利用する。DBSENA は普遍関係インタフェース[6]をその検索インタフェースとしており、データ項目の所在を指定する必要がないところに特色がある。さらに DBSENA では、RDB だけではなく、表や箇条書きのように構造化された HTML 文書も検索することが可能である[8][9]。

4.1.2 分類木の共用化

(A) ディレクトリ

分類木の共用化のため、ディレクトリを分類木に適用する。

ディレクトリは、木構造による階層構造管理を行っているため、分類木との構造的親和性も高いと言える。加えて、複数情報源に対する分類木を考えたとき、ディレクトリの機能により、管理範囲の限定が可能である。さらに、各社のディレクトリサーバやツール類が充実、普及してきており、利用しやすい環境が整いつつある。

具体的には、ディレクトリのインタフェースとして急速に普及しつつある LDAP[10]を利用するディレクトリを採用する。そして、このインタフェースを通し情報のやりとりを行う。

本方式では、複数情報源に対する検索については DBSENA を用いる。そこで、ディレクトリでは、分類木の分類と DBSENA との連携のための情報を管理することとする。ディレクトリはネットワークサービス管理を目的に作成

されたが、この特徴をいかしつつ、別の応用である情報提供への展開を図る。ディレクトリの別の用途への応用として、複数情報源検索を行う CORDS がある[11]。CORDS ではスキーマ情報をディレクトリで管理している。

(B) ディレクトリの View

ディレクトリの共用化を行うために、ディレクトリに View を適用する。ディレクトリ View とはディレクトリの一部分の切り出し、複数のディレクトリの仮想的統合により、ユーザがディレクトリをたどる際の利便性を高めることをねらっている[12]。本方式では、[13]において集合演算による View として整理したもののうち、ディレクトリの一部分を切り出す、射影、選択と、ディレクトリを統合する和を利用する。

射影 View とは、ディレクトリの不要な階層を取り除くものである。あるディレクトリの上位部分、もしくは下位部分のみが必要な場合に適用する。図 1 に示すように、全国の住所ディレクトリのうち、兵庫県のみが必要であれば、上位にあたる日本、関西などのエントリを表示せず、神戸市、中央区など神戸市エントリ配下のエントリのみを射影する。これにより、ユーザはディレクトリをたどる際に、不要なエントリをたどる操作が不要となる。

選択 View では、ディレクトリの不要な広がりを取り除く。あるディレクトリの興味のあるエントリを指定し、そのエントリにたどり着くまでの上位エントリのパスと、そのエントリ配下のエントリのみを表示する。図 2 に示すように、兵庫県内のみが必要であれば、兵庫県エントリを選択し、日本、関西、兵庫県、など、兵庫県までたどる上位のエントリと、姫路市、神戸市、中央区など兵庫県配下のエントリのみを表示する。これによりユーザはディレクトリをたどる際に、選択すべきエントリの数が減る

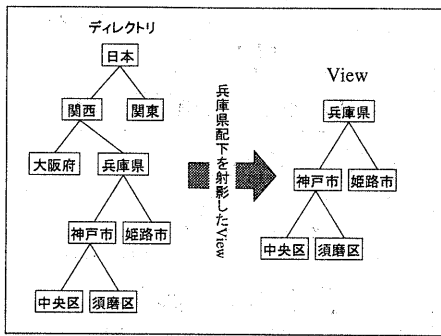


図 1：射影 View

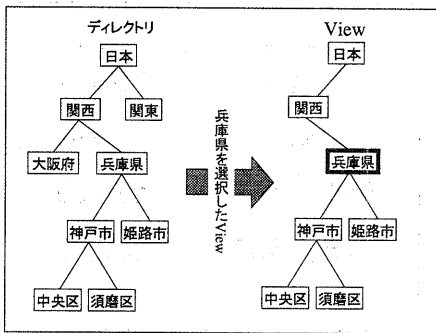


図 2：選択 View

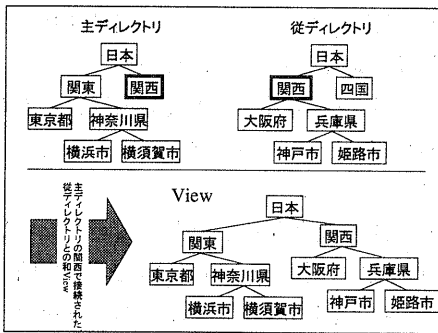


図 3：和 View

ため、情報検索が効率的になる。

和 View は、異なる複数のディレクトリを統合する。あるディレクトリを主とし、主ディレクトリのあるエントリに対し、従となるディレクトリのエントリを対応づけ、そのエントリ配下を結合するものである。図 3に示すように、東日本のディレクトリと西日本のディレクト

リがあった場合、関西というエントリで和をとれば、主である東日本のディレクトリから関西配下が検索可能となる。これにより、管理する分類体系が異なるディレクトリを 1 つのディレクトリとしてまとめることが可能となる。なお、現在のところ、和をとるエントリの上位となるエントリの構成が、主ディレクトリと従ディレクトリで同一であることが必要である。但し、射影 View により、仮想的に同一にみなせれば定義可能である。

4.1.3 ディレクトリと情報源の対応づけ

分類木、すなわちディレクトリにより情報源を管理するとき、あるディレクトリが管理する情報源を定義する。このとき 1 つのディレクトリに対し複数の情報源を対応づけ可能とする。データベースなどの情報源にとどまらず、複数の DBSENA も利用可能とする。

詳細は次節で述べる。

4.1.4 メタ情報管理

上記のような方式を実現するため、情報源、分類木の情報の管理構造や、情報源と分類木の対応関係をメタ情報として定義する。

情報源に関しては、その検索インタフェースを管理する。このインタフェースを通して、それぞれの情報源に検索を依頼する。普遍関係インタフェースを検索インタフェースとする DBSENA の場合、検索可能な項目、検索条件として指定可能な項目を管理する。この検索項目、検索条件などの他、情報源との接続情報などを情報源クラスとして定義する。

分類木に関しては、各エントリの情報とその構造を管理する。

ディレクトリの各エントリで保持する情報をクラスとして管理する。ディレクトリのエントリは、それ自身の型を示すオブジェクトクラ

ス名、型に対応する属性およびその値を保持している。エントリのメタ情報としてこのオブジェクトクラスを利用する。ディレクトリは階層構造を持っており、かつ複数の階層で同じオブジェクトクラスを利用することが可能となっている。そのため、図 4 に示すように、同一のオブジェクトクラスがあらゆる階層、あらゆる情報に定義可能となり、該エントリがどんな情報を持っているか、これだけでは規定できない。そこで CDIR クラスという概念を新規に導入し、オブジェクトクラスとクラス階層中の位置により該エントリの情報を規定する。

ディレクトリの構造は、CDIR クラスを利用し、クラスツリーという階層構造で管理する。図 5 のように、CDIR クラスを階層構造で組合せ、ディレクトリの階層構造を定義する。同一階層に異なる CDIR クラスが存在する場合には、クラスツリー自身が木構造となる。

ディレクトリと情報源を対応づけを行うため、ディレクトリのクラスツリーの最下端に情報源クラスを定義する。ディレクトリを利用して検索する場合、そのディレクトリに対応するクラスツリーのうち、選択したエントリの CDIR クラスが存在するパスに接続された情報源を検索することとなる。

また、クラスのエントリと情報源の検索インタフェースとの対応づけは、検索項目と比較演算子、それに CDIR クラスの属性名を対応づけたものを検索パターンと呼ぶこととする。これにより、ある CDIR クラスのエントリを指定した場合、そのエントリの属性値を検索条件として、情報源を検索する。

定義内容、検索方式の例を5章で述べる。

4.2 地図ディレクトリ

地図を管理するため、住所ディレクトリを利用する。そこで、地図データと表示する地図の

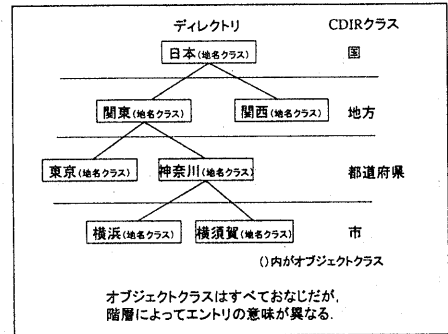


図 4：オブジェクトクラスと CDIR クラス

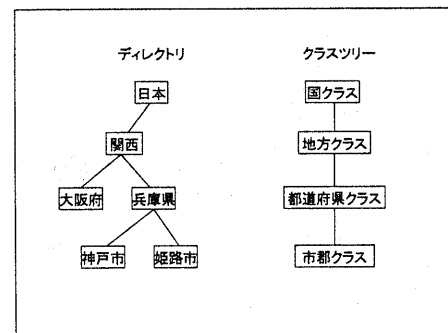


図 5：クラスツリー

縮尺、住所ディレクトリなどと連携させるための情報を定義する。なお、住所ディレクトリの階層は都道府県、市町村など行政区分に限られるものではなく、地域、エリアなど利用目的により変更可能である。

本方式では地図を、階層構造を持つディレクトリとして管理する。地図の階層は、地図の縮尺や、表示領域が対応する。このようなディレクトリを住所ディレクトリと対応づけ管理する。すなわちディレクトリの階層を縮尺に対応づけ、各エントリで対応する領域の外接四角形の緯度経度を管理する。これにより、住所ディレクトリでエントリを指定すれば、所望の領域の緯度経度を指定することが可能となり、それに対応した地図、情報を検索することが可能となる。逆に、表示している地図の緯度経度がわかれば、どの住所ディレクトリのエントリに対

応するかが、ディレクトリを検索することでわかる。これにより、表示している領域の地名を検索することも可能となる。

5 プロトタイプ

前章において提案した方式を確認するプロトタイプとして「MM 情報マップディレクトリ」を作成した。図 6 に構成図を示す。

MM 情報マップディレクトリは Windows NT/98 上で動作する。データベースは MS SQL Server を用い、地図ソフトとして Increment P 社 MapDKIII を用いた。現時点では、異種性のほとんどない複数の情報源に対する検索、複数のディレクトリの管理、ディレクトリの View、地図ディレクトリを実現している。

以下、ディレクトリと DBSENA を連携させた検索方式の一例を述べる。このプロトタイプでは、ディレクトリとして住所ディレクトリを用いる。情報源は施設が分類と緯度経度で管理されている施設情報のデータベースを用意し、DBSENA の検索項目として、「施設名、緯度、経度、分類」とする。本プロトタイプでは、住所ディレクトリのあるエントリを指定すると、その領域に含まれる施設情報を検索する。

5.1 定義内容

メタ情報の定義を図 7 に示す。

5.2 検索処理

住所ディレクトリの神奈川県配下にある「横須賀」エントリを選択した検索を考える。検索処理の流れを以下に示す。概略を図 8 に示す。

- i) 「横須賀」エントリの CDIR クラス「市クラス」に対応する検索パターンを参照する。
- ii) CDIR クラス「市クラス」を含むクラスツリーを参照し、それより下位にある情報源クラス「施設クラス」を得る。
- iii) 施設クラスと検索パターンを対応づけ、

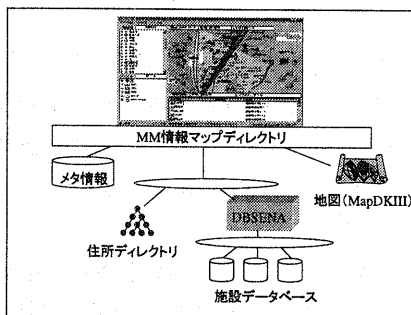


図 6：MM 情報マップディレクトリ

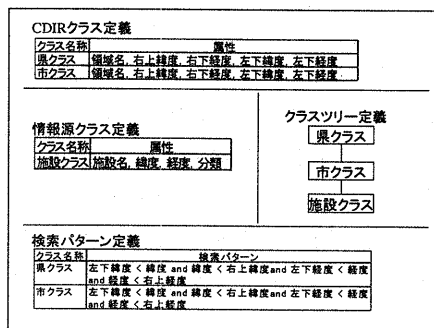


図 7：メタ情報

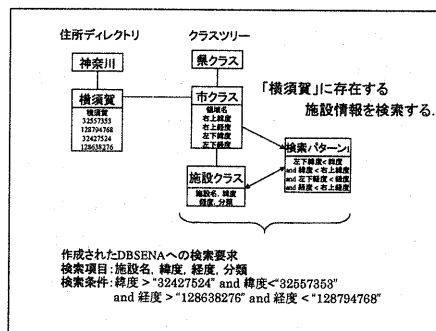


図 8：ディレクトリ連携検索例

検索パターンの検索条件項目と、DBSENA の検索項目の網羅性を判定する。

- iv) 検索要求として、「検索項目：施設名、緯度、経度、分類、検索条件：緯度 > "32427524" and 緯度 < "32557353" and 経度 > "128638276" and 経度 < "128794768"」を DBSENA へ送信する。

6 おわりに

地図と情報を関連付け、既存 GIS や既存のデータベースや Web ページなど、複数の異種情報源を利用可能とし、分類木や地図を検索のインタフェースとして利用可能とする情報提供方式を提案した。

本方式では、i)情報源の異種性の問題を筆者らが従来提案した DBSENA を用いて解決し、ii)分類木としてディレクトリを採用し、iii)CDIR クラスなどのメタ情報を管理し情報検索を可能とした。また、ディレクトリ View 方式を用いて部分木の抽出、統合を可能とした。さらに、地図を検索インタフェースとすることで、ユーザの利便性を高めた。また、プロトタイプを作成し、方式の一部の実証を行った。

今後は、異種分散環境におけるプロトタイプ作成、および、方式の評価を行う予定である。

参考文献

- [1] 桜井博行, “GIS 電子地図革命”, 東洋経済新報社, 1997.
- [2] 石田順一, “情報検索のための多次元分類体系と GUI”, ADBS, Dec., 1994.
- [3] S.Jenkins, M.Jackson, P.Burden, J.Wallis, “Searching the World Wide Web: an evaluation of available tools and methodologies”, Information and Software Technology, Vol39, 1998.
- [4] W.Kim, J.Seo, “Classifying Schematic and Data Heterogeneity in Multidatabase Systems”, IEEE Computer, Dec., 1991.
- [5] L.Raschid, Y-H.Chang, “Interoperable Query Processing from Object and Relational Schemas Based on a Parameterized Canonical Representation”, International Journal of Cooperative Information Systems”, Vol.4, No.1, 1995.
- [6] 星野 隆, 綱川光明, 町原宏毅, “DBSENA: マルチデータベース環境における情報資源管理と検索方式”, 情報処理学会第 114 回 DBS 研究会, Jan., 1998.
- [7] R.Fagin, A.O.Mendelzon, J.D.Ullman, “A Simplified Universal Relation Assumption and its Properties”, ACM Transaction on Database Systems, Vol.7, No.3, Sept., 1992.
- [8] 永末寿宏, 綱川光明, 町原宏毅, “WWW 環境における情報資源管理と検索方式に関する検討”, 1998 年電子情報通信学会情報・システムソサイエティ大会, 1998.
- [9] Y.Iizuka, M.Tsunakawa, S.Seo, T.Ikeda, “An Approach to Integration of Web Information Source Search and Web Information Retrieval”, to be appeared in ACM Symposium on Applied Computing 2000.
- [10] T.Howes, M.Smith, “LDAP: Programming Directory Enabled Application with Lightweight Directory Access Protocol”, Macmillan Technical Publishing, 1997.
- [11] M.Bauer, N.Coburn, P-A.Larson, and T.P.Martin, “Managing Global Information in the CORDS Multidatabase System”, Proc. of the 1994 CoopIS Conference, Toronto, Canada, 1994.
- [12] 長谷川靖, 加藤謹詞, 林剛, 岸本義一, “コンテンツ・ディレクトリ管理システムにおけるビュー機能”, 情報処理学会第 57 回全国大会, Sept., 1998.
- [13] 長谷川靖, 田辺弘実, 岸本義一, 武田英昭, “集合演算によるディレクトリツリーのビュー定義”, 情報処理学会第 117 回 DBS 研究会, Jan., 1999.