

Piotr BAJERSKI, Krzysztof LACH
Politechnika Śląska, Instytut Informatyki

PRZEGLĄD METOD INTERAKCYJNEGO UDOSTĘPNIANIA METADANYCH PRZESTRZENNYCH

Streszczenie. W artykule omówiono pojęcie metadanych przestrzennych oraz przedstawiono ich klasyfikację. Po krótkim przeglądzie stosowanych w praktyce mechanizmów dostępu do metadanych opisano i porównano funkcjonalność wiodących aplikacji udostępniania metadanych.

Słowa kluczowe: metadane przestrzenne, metody dostępu do metadanych.

REVIEW OF GEOSPATIAL METADATA INTERACTIVE ACCESS METHODS

Summary. The paper presents geospatial metadata concept and its taxonomy. After short review of metadata access methods used in practice, description and comparison of most important metadata applications is given.

Keywords: geospatial metadata, metadata access methods.

1. Wstęp

Rozwój systemów informacji przestrzennej (ang. Geographical Information Systems – GIS) doprowadził do powstania ogromnej liczby danych przestrzennych. Dostępne zasoby informacji przestrzennej mogą być użyteczne dla wielu różnych użytkowników o zróżnicowanych potrzebach i wykorzystujących różnorodne narzędzia. Cechami charakterystycznymi danych przestrzennych, wpływającymi na możliwość ich współdzielenia, jest ich złożona struktura, duży rozmiar, znaczna liczba formatów i geograficzne rozproszenie. Niemożność sprawdzenia, jakie dane przestrzenne istnieją, jaka jest ich jakość, jak można do nich uzyskać dostęp i na jakich warunkach oraz jak można uzyskać potrzebny format, prowadzi do kosztownego powielania prac. Aby przeciwdziałać tym problemom, wyodrębniono dane opisu-

jące dane przestrzenne oraz usługi pozwalające na przetwarzanie danych przestrzennych i podjęto prace nad serwisami umożliwiającymi gromadzenie, przeszukiwanie i prezentację tych opisów. Dane opisujące dane oraz usługi zostały nazwane metadanymi.

Metadane opisujące dane przestrzenne można podzielić z punktu widzenia faz dostępu do danych na trzy grupy, [2]:

1. metadane służące do odkrywania danych (ang. discovery metadata),
2. metadane służące do analizowania danych (ang. exploration metadata),
3. metadane służące do pozyskiwania danych (ang. exploitation metadata).

Z drugiej strony, atrybuty metadanych są dzielone na grupy tematyczne. Na ogół wyróżnia się następujące najważniejsze grupy tematyczne, [1, 2, 4 i 8]:

- Identyfikacja – nazwa i opis zbioru danych, słowa kluczowe, twórca, pokrywany obszar geograficzny.
- Jakość – dokładność określenia współrzędnych, kompletność danych, informacja, czy spójność danych była weryfikowana, na podstawie jakich danych zbiór został utworzony i jaki proces przetwarzania danych zastosowano.
- Organizacja przestrzenna – model przestrzenny, który został użyty do zakodowania danych.
- Informacja na temat atrybutów danych przestrzennych – jakie dane są dostępne (np. drogi, budynki, wysokość). Jakie są dziedziny tych atrybutów, jakie zostały użyte kody i jak należy je interpretować.
- Informacja dystrybucyjna – gdzie można uzyskać dane, w jakich formatach, za jaką cenę.
- Informacja o osobach i instytucjach odpowiedzialnych za metadane i dane.

Podstawową grupą atrybutów służącą do odkrywania danych jest grupa *Identyfikacja*. Możliwe jest jednak wyszukiwanie zasobów po atrybutach należących do pozostałych podanych grup tematycznych. Podczas analizowania przydatności danych najważniejsze są informacje dotyczące jakości, organizacji przestrzennej i atrybutów danych przestrzennych. Dwie ostatnie grupy tematyczne są kluczowe dla pozyskania danych.

Metadane ze względu na różnorodność tworzących i utrzymujących je instytucji powstawały w różnych formatach, co znacząco utrudniało ich wykorzystanie. Problemy te legły u podstaw przedsięwzięć zmierzających do opracowania międzynarodowych standardów metadanych oraz infrastruktury pozwalającej na udostępnianie metadanych szerokiemu gronu odbiorców. Głównymi normami dotyczącymi metadanych są opracowywana norma ISO 19115, [8] i amerykańska narodowa norma FGDC-STD-001, [4].

Zgodnie z definicją OpenGIS™ [10 i 11] termin katalog metadanych opisuje zbiór interfejsów serwisów wspomagających organizację, odkrywanie i dostęp do informacji przestrzennych. Serwisy katalogowe pomagają użytkownikowi lub innemu programowi w odnalezieniu informacji, która istnieje gdziekolwiek w rozproszonym środowisku kompute-

rowym. Katalog można traktować jak specjalizowaną bazę danych zawierającą informacje o geoprzestrzennych zasobach dostępnych dla pewnej grupy lub społeczności użytkowników.

W katalogu metadanych muszą być przechowywane tylko te atrybuty metadanych, po których można wyszukiwać metadane. Pozostałe atrybuty mogą być również przechowywane w katalogu lub w repozytorium metadanych.

Specyfika serwerów katalogów metadanych wynika z następujących wymagań:

- przechowywanie stosunkowo małej liczby obiektów o złożonej strukturze,
- udostępnianie metadanych szerokiemu gronu odbiorców,
- nastawienie na przetwarzanie zapytań, przy czym większość zapytań jest prosta (zawiera mało warunków),
- powstawanie metadanych w wielu różnych miejscach,
- rzadkie dodawanie nowych metadanych,
- sporadyczna aktualizacja metadanych,
- metadane są związane z danymi i nie powinny być od nich fizycznie separowane.

2. Przegląd stosowanych mechanizmów dostępu do metadanych

Metadane ze swej natury powinny być udostępniane szerokiemu gronu odbiorców. Związane są z tym dwa podstawowe problemy: uzyskanie zdalnego dostępu do katalogu metadanych i dostępność interfejsu pozwalającego na wygodne i efektywne przeszukanie tego katalogu. Obydwa te problemy można rozwiązać za pomocą infrastruktury Internetu, która rozwiązuje problem połączenia komputerów oraz umożliwia używanie interfejsów opartych na komponentach pobieranych z serwera (strony HTML, aplety itd.). Innym rozwiązaniem byłyby rozwiązania firmowe działające w oparciu o określone protokoły i wykorzystujące specjalizowane oprogramowanie. Wydaje się, że przyszłość dostępu do metadanych (przynajmniej najbliższa przyszłość) należy do rozwiązań opartych na przeglądarkach internetowych. Jak wynika z przeglądu istniejącej infrastruktury danych przestrzennych, wszystkie narodowe systemy metadanych udostępniają interfejsy użytkownika w Internecie. Również rozwiązania firmowe są tworzone w oparciu o to środowisko.

W literaturze wyróżnia się dwa typy interakcji między użytkownikiem a interfejsem w przeglądarce www służące do odnajdowania informacji [2 i 9]:

- przeglądanie – użytkownik nawiguje w hierarchii metadanych, przeglądając atrybuty wskazanych elementów,
- wyszukiwanie – użytkownik formułuje zapytanie, nakładając warunki na atrybuty metadanych, korzystając z interfejsów o różnym stopniu złożoności. W wyniku wykonania zapytania otrzymuje zbiór metadanych, których atrybuty może przeglądać.

Obydwe usługi z punktu widzenia interakcji z użytkownikiem i realizujących je mechanizmów są do siebie podobne. Wyszukiwanie jest szczególnie przydatne dla zaawansowanych użytkowników o sprecyzowanych potrzebach. Wymaga znajomości terminologii dziedziny, atrybutów metadanych oraz umiejętności zapisu zapytań. Przeglądanie jest użyteczne dla mniej zaawansowanych użytkowników. Z tym trybem wiąże się jednak poważny problem klasyfikacji metadanych w środowisku rozproszonym w oparciu o wspólne słownictwo. Dużym utrudnieniem klasyfikacji jest wymaganie, aby powstała jedna hierarchia. Przeglądanie i wyszukiwanie wzajemnie się uzupełniają i wydaje się celowe zastosowanie obydwu podejść w dostępie do rozproszonego katalogu metadanych. Wbrew pozorom przeglądanie metadanych w środowisku rozproszonym obejmującym katalogi metadanych różnych instytucji może być trudniejsze w implementacji niż wyszukiwanie. Złe utworzenie hierarchii metadanych może wprowadzać użytkownika w błąd i utrudniać mu odnalezienie danych.

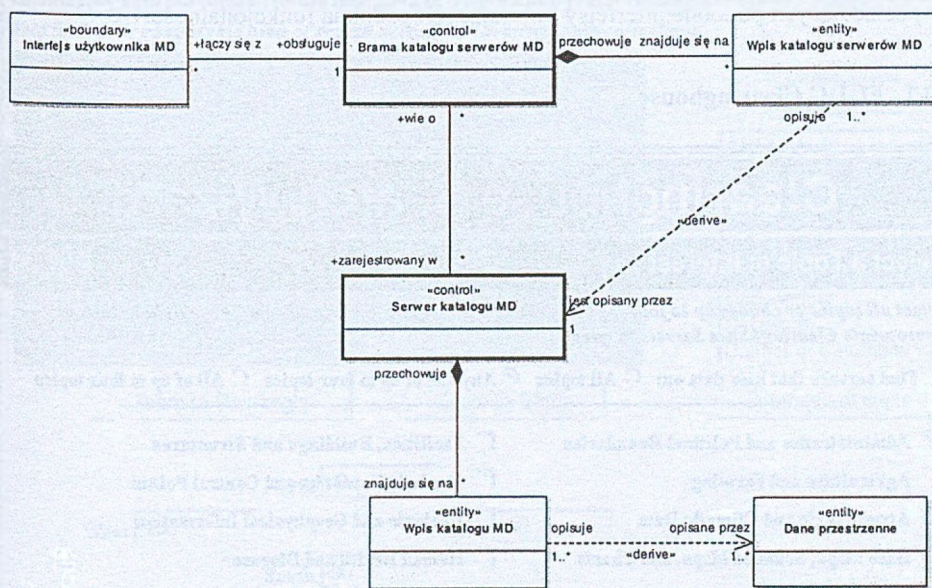
Z punktu widzenia użytkownika pragnącego wyszukać metadane mechanizmy zadawania zapytań dostępne w interfejsach internetowych można podzielić na:

1. Wyszukiwanie identyfikatora lub nazwy metadanych pasującej do wzorca – lokalizowany jest kolejny obiekt metadanych, począwszy od bieżącej pozycji w zbiorze metadanych, którego identyfikator lub nazwa pasuje do podanego wzorca (który może zawierać znaki wieloznaczne); ten rodzaj wyszukiwania może zostać użyty podczas przeglądania metadanych zorganizowanych hierarchicznie lub liniowo.
2. Wyszukiwanie z warunkami na atrybuty jednego obiektu metadanych (nie można specyfikować warunków odwołujących się do grupy obiektów) – w zależności od elementów, które mogą wystąpić w tych warunkach, można wprowadzić podział na:
 - 2.1. Zapytania, w których atrybuty mogą być porównywane tylko ze stałymi; szczególnym przypadkiem jest postać koniunkcyjna, w której wszystkie warunki są połączone iloczynem logicznym.
 - 2.2. Zapytania, w których atrybuty mogą być porównywane zarówno ze stałymi, jak i z innymi atrybutami.
3. Wyszukiwanie z warunkami na zbiory metadanych (może wystąpić grupowanie, operacje na zbiorach itd.).

Dostępność metadanych jest istotna dla szybkiego odnalezienia interesujących użytkowników danych. Jednakże samo istnienie metadanych nie jest wystarczające – niezbędna jest infrastruktura pozwalająca na przeszukiwanie wszystkich metadanych, które potencjalnie mogą opisywać przydatne zasoby danych. Przykładowo, umieszczanie katalogu metadanych na płytach CD z danymi przestrzennymi niewątpliwie ułatwia korzystanie z danych, wymaga jednak dostępu do każdej z płyt. Udostępnienie katalogów metadanych jako osobnych serwisów polepsza sytuację, ale problem wyszukiwania pozostaje, ponieważ należy odnaleźć każdy z tych serwisów i wyszukać na nim metadane. Praktyczne zastosowanie rozwiązania

opartego na modelu scentralizowanego serwera katalogu metadanych jest trudne ze względu na naturalne rozproszenie metadanych.

Rozwiązaniem stosowanym w praktyce jest model rozproszonego katalogu metadanych [2 i 10]. Pozwala on na przezwyciężenie problemów występujących w modelu z niezależnymi serwerami katalogów metadanych, przy jednoczesnym zachowaniu ich autonomii. Użytkownik poprzez interfejs łączy się z bramą rozproszonego katalogu metadanych i przesyła do niej zapytanie. Brama na podstawie metadanych opisujących serwery i/lub wskazówek od użytkownika określa, do których serwerów katalogu metadanych należy odesłać zapytanie. Zapytanie jest rozsyłane do wybranych serwerów i wykonywane na nich lokalnie. Następnie otrzymane odpowiedzi są scalane i odsyłane do interfejsu użytkownika. Na rys. 1 przedstawiono model klas dla rozproszonego katalogu metadanych.



Rys. 1. Model klas ilustrujący architekturę rozproszonego katalogu metadanych
Fig. 1. Class diagram showing distributed metadata catalog architecture

Brama umożliwia rejestrowanie nowych serwerów metadanych, aktualizację opisów zarejestrowanych oraz usuwanie serwerów z rejestru. Aby całość mogła funkcjonować, wszystkie serwery katalogu metadanych muszą rozumieć przesyłane do nich zapytania i odsyłać odpowiedzi w formacie zrozumiałym dla bramy. W praktyce wykorzystuje się standard wymiany informacji Z39.50 [3 i 6]. Przedstawiona koncepcja jest szczególnie przydatna, jeżeli metadane są udostępniane przez różne równoprawne organizacje, które chcą udostępniać dane w oparciu o ten sam profil.

Na podkreślenie zasługuje fakt, że w serwerze bramy nie są replikowane metadane zawarte w serwerach katalogów metadanych. Jednakże metadane przechowywane przez każdy

z serwerów katalogu metadanych znajdują odzwierciedlenie w jego opisie w serwerze bramy. Niewłaściwy opis serwera może więc powodować udzielanie niepełnej lub błędnej odpowiedzi. Z tego powodu w dostępnych rozwiązaniach pozwala się użytkownikowi na określanie, które serwery metadanych należy przeszukać.

3. Przegląd aplikacji udostępniających metadane przestrzenne

W rozdziale przedstawiono wybrane aplikacje udostępniające metadane w ramach narodowych infrastruktur danych przestrzennych oraz rozwiązanie komercyjnej firmy ESRI, będącej liderem rynku GIS. Z pośród wielu katalogów metadanych wybrano reprezentatywne dla grup posiadających podobne interfejsy i oferujących podobną funkcjonalność.

3.1. FGDC Clearinghouse

Clearinghouse Search Wizard

Select a Topic of Interest
Help...

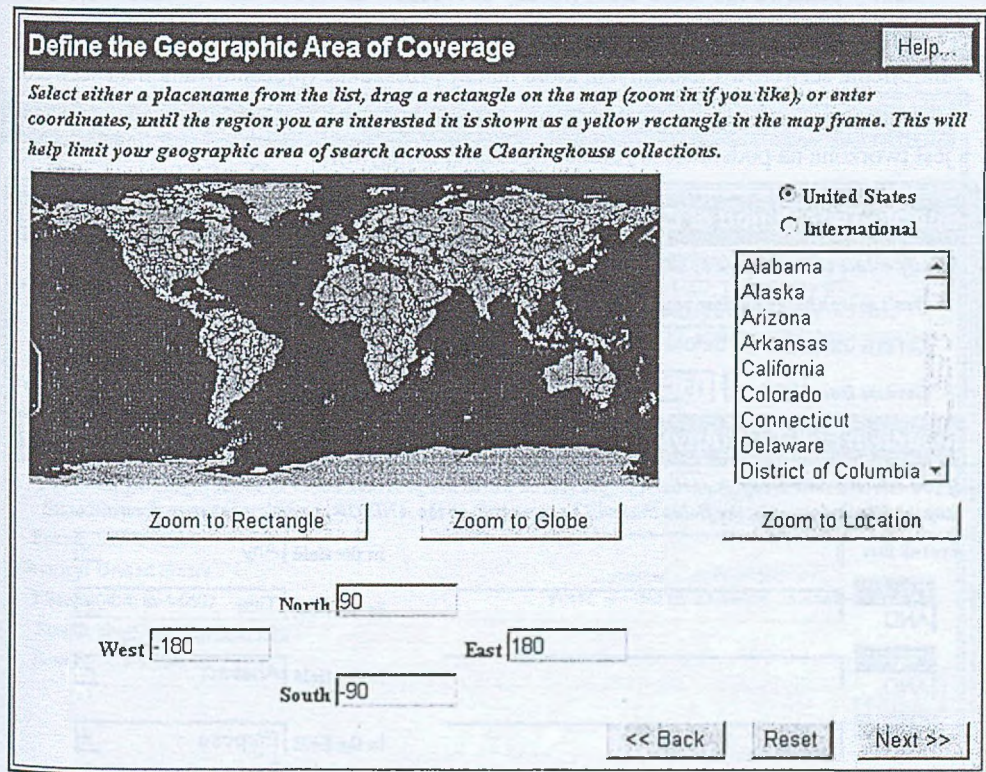
Select all topics or choose up to four topics of interest from the categories listed. This will help select appropriate Clearinghouse Servers to query.

Find servers that have data on: All topics Any one of up to four topics All of up to four topics

<input checked="" type="checkbox"/> Administrative and Political Boundaries	<input type="checkbox"/> Facilities, Buildings and Structures
<input type="checkbox"/> Agriculture and Farming	<input type="checkbox"/> Geodetic Networks and Control Points
<input type="checkbox"/> Atmospheric and Climatic Data	<input type="checkbox"/> Geologic and Geophysical Information
<input type="checkbox"/> Base Maps, Scanned Maps, and Charts	<input type="checkbox"/> Human Health and Disease
<input type="checkbox"/> Biologic and Ecologic Information	<input type="checkbox"/> Imagery and Aerial Photographs
<input type="checkbox"/> Business and Economic Information	<input type="checkbox"/> Inland Water Resources and Characteristics
<input type="checkbox"/> Cadastral and Legal Land Descriptions	<input type="checkbox"/> Ocean and Estuarine Resources and Characteristics
<input type="checkbox"/> Earth Surface Characteristics & Land Cover	<input type="checkbox"/> Society, Cultural and Demographic Information
<input type="checkbox"/> Elevation and Derived Products	<input type="checkbox"/> Tourism and Recreation
<input type="checkbox"/> Environmental Monitoring and Modeling	<input type="checkbox"/> Transportation Networks and Models
	<input type="checkbox"/> Utility Distribution Networks

Rys. 2. Wybór tematów w interfejsie typu kreator portalu FGDC Clearinghouse
 Fig. 2. Topic category selection in the FGDC Clearinghouse Wizard interface

Clearinghouse utworzony i utrzymywany przez *Federal Geographic Data Committee* (FGDC) jest bramą do rozproszonego katalogu metadanych [12]. Został on utworzony jako system narodowy, jednak obecnie jest w nim zarejestrowanych ponad 200 serwerów katalogu metadanych z całego świata i może być traktowany jako brama do globalnego katalogu metadanych. Rozpoczynając wyszukiwanie użytkownik wybiera na mapie jedną z kopii bramy katalogu [13]. Każda z kopii udostępnia tę samą listę serwerów i te same mechanizmy wyszukiwania. Clearinghouse działa w oparciu o normę metadanych FGDC-STD-001-1998 [4] i protokół Z39.50 [6].



Rys. 3. Selekcja przestrzenna w interfejsie typu kreator portalu FGDC Clearinghouse

Fig. 3. Spatial selection in the FGDC Clearinghouse Wizard interface

Brama katalogu metadanych udostępnia trzy różne interfejsy wyszukiwania metadanych:

- seria formatek do zadawania zapytania (interfejs typu kreator) – kolejne zakładki pozwalają na zadawanie powiązanych tematycznie warunków,
- pojedyncza formatka z mapą,
- pojedyncza formatka bez mapy.

Największe możliwości wyszukiwania udostępnia interfejs typu kreator, który na kolejnych formatkach pozwala na:

- wybór tematyki metadanych (rys. 2) w oparciu o kategorie tematyczne zdefiniowane w normie ISO 19115,
- określenie lokalizacji geograficznej (rys. 3), która może być zadawana za pomocą:
 - prostokąta rysowanego na mapie, której skalę można zmieniać,
 - czterech współrzędnych,
 - nazwy państwa lub stanu USA (należy podkreślić, że wybór nazwy aktualizuje zakres współrzędnych i prostokąt ograniczający),
- określenie serwerów metadanych, które należy przeszukać (prezentowana lista serwerów, na której użytkownik może wybrać serwery, do których ma zostać rozestane zapytanie, jest tworzona na podstawie wybranej wcześniej tematyki),

Specify Time Period of Content Help...

Specify a date or date range by selecting one of the methods below.

Don't search based on time period

Get data whose date is the date

Get data from through

Search in Full-Text (Any) or by Field Help...

If you selected search topics earlier, sample search words are provided below to assist in finding data. You may modify these words, the fields that will be searched, or the AND/OR condition between the words.

Search for: in the field

in the field

in the field

in the field

Additional Configuration Settings Help...

Maximum number of records to show on each page:

Rys. 4. Zadawanie warunków na atrybuty metadanych w interfejsie typu kreator portalu FGDC Clearinghouse

Fig. 4. Metadata attribute search criteria specification in the FGDC Clearinghouse Wizard interface

- specyfikację warunków na atrybuty metadanych (rys. 4); warunki są podzielone na:
 - sekcję okresu czasu, w której można określić warunek na datę lub przedział czasu (nie jest sprecyzowane, jakiego dokładnie atrybutu z normy FGDC te warunki dotyczą),
 - sekcję atrybutów opisowych, która pozwala nałożyć do trzech warunków określających, że słowo lub lista słów ma wystąpić w zadanym atrybucie metadanych (można używać znaku wieloznacznego), jeden atrybut może występować więcej niż w jednym warunku, warunki mogą być połączone sumą lub iloczynem logicznym, lista atrybutów, na które można nakładać warunki, obejmuje: dowolny atrybut (wyszukiwanie obejmuje wszystkie atrybuty), tytuł, dostawcę, abstrakt, cel, formę prezentacji i słowa kluczowe tematu.

Georgia GIS Data Clearinghouse

Title: Georgia GIS Data Clearinghouse [[Related website](#)]

Acronym: Georgia GIS

Abstract: Themes include administrative boundaries, labor statistics, census population and housing, labor statistics, population projections, economic development data, natural resources data and more

Keywords: farming 001; climatologyMeteorologyAtmosphere 004; environment 007; boundaries 003; earthCover 010 201; elevation 006; geoscientificInformation 008; society 016; imagery 010 202; inlandWaters 012; planningCadastre 015; utilitiesCommunication 019; structure 017; transportation 018; baseMaps 010 200;

Contact: Georgia GIS Data Clearinghouse

Person: Rama Sivakumar

Address: 245 Fourth Street, Bldg. 204

City: Atlanta

State: Georgia

Postal: 30332-0155


Country: United States

Phone: (404) 894-6607

Email: siva@arch.gatech.edu

Hours: M-F 9:00-5:00

Geographic Area of Coverage



West: -86 North: 35 South: 31 East: -81

[<< Back](#)

Rys. 5. Metaopis serwerów metadanych w portalu FGDC Clearinghouse
 Fig. 5. Metadescription of metadata servers in the FGDC Clearinghouse

- Serwery katalogu metadanych są opisywane za pomocą następujących atrybutów (rys. 5):
- pełna oficjalna nazwa serwera,
 - skrót nazwy serwera,
 - abstrakt opisujący zawartość serwera,

- słowa kluczowe opisujące zasoby udostępniane przez serwer; słowa te odpowiadają kategoriom tematycznym z normy ISO 19115 i służą do wstępnej selekcji serwerów na podstawie tematyki wybranej przez użytkownika w oknie kreatora pokazanym na rys. 2,
- informacje kontaktowe,
- obszar geograficzny obejmowany przez serwer (w postaci czterech współrzędnych).

W opisie serwera nie jest dostępny jego adres (URL). Zamiast niego istnieje element kontrolny (*Related website*), który umożliwia przejście do tego serwera.

Drugi interfejs (pojedyncza formatka z mapą) oferuje prawie te same możliwości formułowania zapytań co kreator. Nie jest dostępny wybór tematów, przez co cały czas jest dostępna pełna lista serwerów. Ostatni interfejs jest wersją interfejsu drugiego, z którego usunięto mapę, dzięki czemu nie jest wymagana obecność środowiska Javy. Interfejs oparty na kreatorze jest najnowszym produktem oferującym największe możliwości i obrazuje kierunek rozwoju oraz jedną z metod radzenia sobie z dużą liczbą serwerów.

W wyniku wykonania zapytania zwracana jest lista etykiet znalezionych metadanych, które następnie można oglądać w formie skróconej lub pełnej (jako plik sformatowany zgodnie ze standardem FGDC-STD-001-1998). Nie jest tworzony ranking zwracanych wyników.

Serwery dostępne w portalu *FGDC Clearinghouse* są serwerami ośrodków rządowych, serwerami stanowymi, ośrodków naukowych, dużych firm lub programów badawczych. Serwery zarejestrowane w Clearinghouse są również dostępne przez własne interfejsy. Część z nich udostępnia własne mniej lub bardziej ograniczone możliwości wyszukiwania, inne przekierowują obsługę zapytań do głównej strony Clearinghouse. Portal umożliwia sprawdzanie statusu zarejestrowanych serwerów.

3.2. GIGateway

Portal *GIGateway* [14] jest nową wersją bramy narodowego katalogu metadanych Wielkiej Brytanii (rys. 6). W połowie 2002 roku zastąpił on wcześniejszy portal o nazwie *askGIRaffe*. Funkcjonalność obydwu portali jest podobna. Główna różnica polega na przeniesieniu w portalu *GIGateway* zadawania warunków geograficznych do osobnego okna. Ponieważ opcja ta nie została jeszcze udostępniona, w rozdziale opisano możliwości udostępniane w tym zakresie przez wcześniejszą wersję (*askGIRaffe*).

GIGateway oferuje podobne możliwości zadawania warunków jak *FGDC Clearinghouse*:

- określenie tematów,
- wyszukiwanie słowa lub frazy w tytule, dostawcy, abstrakcie lub w dowolnym z tych atrybutów,
- warunek na zakres dat,

- warunki przestrzenne zadawane przy użyciu mapy, której skalę można zmieniać, warunek przestrzenny jest określony przez widoczną część mapy, nie można natomiast zadać warunku przez zakres współrzędnych,
- możliwość wyszukiwania według nazw geograficznych w następujących kategoriach: kraje, obszary administracyjne, nazwy miejsc lub kody pocztowe,
- określenie kartotek (odpowiadających pojęciu serwerów z *FGDC Clearinghouse*), które należy przeszukiwać.

21 November 2002

gigateway

Home | Data Locator | Data Directory | Area Search | More Information

the **data locator** helps you to find and use up-to-date and accurate UK geographic information from a range of sources.

There are 6* search methods. You can use any combination of these to retrieve the results you need.
*geographic search temporarily unavailable

1 free text search:

2 field to search:

3 keyword:

4 which directory would you like to search:

5 location*:

6 before date:
 after date:

to search for location using a map, click here.*

search →

agi | Data Providers | Terms & Conditions | Contact Us | About Us | Links | Site Map

search this site! →

Rys. 6. Zadawanie warunków na metadane w portalu GIGateway

Fig. 6. Metadata attribute search criteria specification in the GIGateway Data Locator

Interfejs użytkownika jest prosty, ergonomiczny i przyjazny, jednakże możliwości zadawania zapytań są ograniczone. Interfejs nie udostępnia przeglądania metadanych.

W trakcie wyszukiwania jest pokazywany status połączeń z wybranymi katalogami metadanych. Wyniki wyszukiwania są widoczne w postaci listy serwerów, informacji o połączeniu i liczby znalezionych rekordów. Metadane są prezentowane w postaci tekstowej jako strona HTML, której forma zależy od serwera katalogu metadanych udostępniającego daną pozycję odpowiedzi.

3.3. Australian Spatial Data Directory (ASDD)

Portal *Australian Spatial Data Directory* (ASDD) [14] jest bramą do rozproszonego katalogu metadanych w ramach australijskiej infrastruktury danych przestrzennych. Oferuje dwa interfejsy wyszukiwania danych: prosty i złożony. Interfejs prostych zapytań pozwala na wyszukiwanie podanych napisów w podstawowych atrybutach metadanych: tytuł, abstrakt, dostawca i słowa kluczowe według ANZLIC [19]. Istnieje również możliwość jednoczesnego wyszukiwania napisu we wszystkich wymienionych atrybutach. Warunki logiczne można łączyć spójnikami logicznymi *And*, *Or* lub *And-Not* (zgodnie ze standardem Z39-50 [6]).

Interfejs zadawania złożonego zapytania pozwala dodatkowo na nałożenie warunków przestrzennych oraz temporalnych. Zakres funkcjonalności tego interfejsu jest podobny do interfejsu z pojedynczą formatką z mapą w *FGDC Clearinghouse*.

Portal umożliwia wybór serwerów, do których ma zostać skierowane zapytanie. W bramie zarejestrowanych jest ponad dwadzieścia serwerów. Istnieje możliwość sprawdzenia ich statusu, czasu połączeń i liczby przechowywanych obiektów metadanych na poziomie zbiorów danych (łącznie jest zarejestrowanych ponad 32 tys. wpisów o metadanych).

3.4. GeoConnections Discovery Portal

Portal *GeoConnections Discovery Portal* [16] jest bramą do rozproszonego katalogu metadanych w ramach kanadyjskiej infrastruktury danych przestrzennych. Oferuje on osobne interfejsy do wyszukiwania danych, organizacji i usług. Funkcjonalność interfejsu wyszukiwania danych jest zbliżona do interfejsu AskGIRaffe, ale zamiast mapy ma warunki na cztery współrzędne geograficzne wyznaczające warunek w postaci prostokąta. Wyszukiwanie organizacji i usług jest podobne do wyszukiwania informacji w standardowych wyszukiwarkach internetowych.

3.5. U.S. Environmental Protection Agency (EPA)

Portal *U.S. Environmental Protection Agency (EPA) Node of the National Geospatial Data Clearinghouse* [17] jako jeden z nielicznych udostępnia przeglądanie metadanych. Są one udostępniane w postaci list uporządkowanych według tytułu, tematu lub obszaru. Dostępne jest proste wyszukiwanie ciągów znaków bez wskazywania, jakie atrybuty metadanych mają być przeszukiwane. W ostatnich miesiącach zrezygnowano ze złożonej składni zapytań pozwalającej na odwoływanie się do stref i pól dokumentów oraz na nakładanie warunków dotyczących położenia słów względem siebie (ang. proximity search).

Odwołania do innych serwerów są dostępne przez ich listę. Katalog metadanych nie jest rozproszony – główny serwer może nie wiedzieć o metadanych zawartych w serwerach, do

- przeglądanie metadanych, które jest realizowane przez odesłanie do serwera, który posiada metadane i prezentuje je w postaci listy, metadane są dostępne tylko na poziomie zbiorów danych.

3.7. ArcCatalog firmy ESRI

ArcCatalog jest komercyjnym narzędziem do obsługi metadanych dostępnym w ramach środowisk *ArcInfo* i *ArcView* firmy ESRI [5]. Umożliwia on tworzenie, edycję, przeglądanie i wyszukiwanie metadanych opisujących dowolne zasoby: dane geograficzne, pliki nieprze-strzenne, katalogi, bazy danych itd. Metadane są przechowywane w osobnych plikach w formacie XML. Podstawowym standardem metadanych jest standard FGDC, istnieje jednak możliwość używania innych standardów. *ArcCatalog* umożliwia utworzenie w prosty sposób węzła Clearinghouse.

W *ArcCatalog* dane i metadane są zorganizowane hierarchicznie w kartoteki, a interfejs do ich przeglądania przypomina eksplorator systemu Windows. W zintegrowanym środowisku firmy ESRI możliwy jest podgląd dowolnych danych zapisanych we wspieranych przez to środowisko formatach.

W programie *ArcCatalog* użytkownik zapisuje warunki przy użyciu zakładki:

- *Nazwa i lokalizacja* – możliwość nakładania warunków na podstawowe informacje o metadanych i danych,
- *Warunki geograficzne* – możliwość nakładania warunków przestrzennych przez zaznaczenie obszaru na mapie, wskazanie zakresu współrzędnych lub wybranie nazwy geograficznej lub administracyjnej z listy,
- *Daty* – możliwość nakładania warunków na wymiar temporalny metadanych i danych,
- *Warunki zaawansowane* – możliwość nakładania warunków na dowolne elementy metadanych w postaci trójek (nazwa elementu, porównanie, wartość), co odpowiada zapytaniom na poziomie 2.1.

4. Porównanie rozwiązań

Z przedstawionych w rozdziale 3 interfejsów katalogu metadanych wszystkie dostarczają mechanizmów wyszukiwania metadanych o różnym stopniu zaawansowania, a tylko nieliczne (*EPA*, *South African* i *Arc Catalog*) pozwalają na ich przeglądanie.

Wyszukiwanie metadanych dostępne w bramach rozproszonych katalogów metadanych charakteryzuje się podobną funkcjonalnością:

- rozdzielenie wyszukiwania na podstawowe i zaawansowane, z mapą i bez mapy,

- formułowanie warunków na poziomie 2.1, pozwalających tylko na porównywanie atrybutów ze statymi,
- wydzielenie zadawania warunków na tematy (kategorie); wydaje się, że kategorie stają się najważniejszym kryterium wyszukiwania metadanych (dostępne do wyboru kategorie są określone normą metadanych ISO 19115),
- nakładanie warunków tylko na mocno ograniczony zbiór atrybutów metadanych (tytuł, abstrakt, dostawcę, słowa kluczowe),
- ograniczenie liczby warunków do kilku,
- wyszukiwanie we wszystkich atrybutach (dla których wyszukiwanie jest udostępniane),
- łączenie warunków spójnikami logicznymi zdefiniowanym w Z39.50 (*And*, *Or* i *And-Not*),
- nakładanie warunków przestrzennych w postaci jednego prostokąta rysowanego na mapie lub określanego przez cztery współrzędne geograficzne,
- wyszukiwanie po nazwach geograficznych,
- wybór serwerów, do których zostanie rozesłane zapytanie.

Przeglądanie metadanych jest możliwe do zrealizowania tylko w przypadku wcześniejszego zorganizowania ich hierarchii. Pokazanie w interfejsie bramy listy wszystkich dostępnych metadanych (która w *FGDC Clearinghouse* może być liczona w dziesiątkach tysięcy wpisów) wydaje się niecelowe. Z drugiej strony utworzenie jednej hierarchii jest bardzo trudnym problemem [2]. Przeglądanie metadanych jest dostępne na serwerach prezentujących lokalne zasoby metadanych i danych. Serwery takie mogą być zarejestrowane w bramach rozproszonego katalogu metadanych, które bezpośrednio nie pozwalają na przeglądanie metadanych, ale zawierają odnośniki do nich.

Znalezione metadane są przesyłane jako strony HTML. Podczas prezentacji odpowiedzi można wybrać podstawowy lub pełny zakres widocznych atrybutów. Zazwyczaj metadane są formatowane zgodnie ze standardem FGDC. Często można wybrać publikowanie metadanych w postaci tekstu lub zgodnie z profilem GEO. Większość ośrodków ma w planach publikowanie metadanych w standardzie ISO 19115 [8].

Najbardziej zaawansowanym rozwiązaniem wydaje się brama *FGDC Clearinghouse*, która poza Stanami Zjednoczonymi zawiera odnośniki do serwerów z Kanady, Anglii, Australii, krajów Ameryki Łacińskiej, Japonii i Afryki.

Wyniki przedstawionego w artykule przeglądu aplikacji udostępniania metadanych przestrzennych zostały wykorzystane w analizie wymagań i projekcie prototypu portalu Regionalnego Systemu Informacji Przestrzennej w Województwie Śląskim [7]. Ze względu na specyfikę rozwiązania, polegającą na istnieniu jednej organizacji zarządzającej danymi, zastosowano zarówno przeglądanie metadanych, jak i ich wyszukiwanie.

LITERATURA

1. Content Standard for Digital Geospatial Metadata Workbook Version 2.0, <http://www.fgdc.gov> 2000
2. Douglas D. Nebert, The SDI Cookbook, Global Spatial Data Infrastructure, <http://www.gsdi.org> 2001
3. Douglas D. Nebert, Z39.50 Application Profile for Geospatial Metadata or "GEO", Version 2.2, U.S. Federal Geographic Data Committee, 590 National Center, U.S. Geological Survey <http://www.blueangeltech.com/Standards/GeoProfile/geo22.htm>
4. Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata FGDC-STD-001-1998 <http://www.fgdc.gov>
5. ESRI, <http://www.esri.com>
6. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification, ANSI/NISO Z39.50-1995, Z39.50 Maintenance Agency 1995
7. Instytut Systemów Przestrzennych i Katastralnych S.A., Wdrożenie Regionalnego Systemu Informacji Przestrzennej (RSIP) w Województwie Śląskim dla wsparcia planowania regionalnego i lokalnego, restrukturyzacji regionu oraz zarządzania w sytuacjach kryzysowych" Zadanie 3.2 „Zagadnienia tworzenia i dostępu do bazy metadanych RSIP", Gliwice 2002
8. Projekt normy ISO 19115.3, Geographic information – Metadata, 2000, ISO/TC211 <http://www.isotc211.org>
9. The OpenGIS™ Abstract Specification, Topic 11: Metadata, Version 4, Open GIS Consortium 1999 <http://www.opengis.org>
10. The OpenGIS™ Abstract Specification, Topic 12 OpenGIS™ Service Architecture Version 4.2, Open GIS Consortium 2001 <http://www.opengis.org>
11. The OpenGIS™ Abstract Specification, Topic 13: Catalog Services, Version 4, Open GIS Consortium 1999 <http://www.opengis.org>
12. FGDC Clearinghouse <http://www.fgdc.gov/clearinghouse/clearinghouse.html>
13. <http://130.11.52.184/>
14. GIGateway Data Locator <http://www.gigateway.org.uk/default.asp>
15. Australian Spatial Data Directory <http://www.ausliq.gov.au/asdd/>
16. GeoConnections Discovery Portal <http://ceonet.ac.ca/cs/en/index.html>
17. *South African National Spatial Information Framework* <http://www.nsisf.org.za/>
18. U.S. Environmental Protection Agency (EPA) <http://www.epa.gov/nsdi/>
19. <http://www.anzlic.org.au/>

Wpłynęło do Redakcji 23 stycznia 2003 r.

Abstract

Geospatial metadata play a key role in supporting the discovery, evaluation, and application of geographic data. The term *geospatial metadata* is used to describe the summary information or characteristics of a set of spatial data.

The paper starts with geospatial metadata definition and classification according to FGDC and ISO standards [4 and 8]. Next review of metadata access mechanisms is presented with particular attention given to the concept of a distributed metadata catalog architecture.

The paper proposes a classification of capabilities of query formulation in interfaces to spatial metadata available in the Internet. Some selected applications facilitating access to national spatial data infrastructures and a commercial application offered by ESRI, called *ArcCatalog*, are described. Among many national solutions a group of representative applications were chosen: FGDC Clearinghouse, GIGateway Data Locator, Australian Spatial Data Directory, GeoConnections Discovery Portal, U.S. Environmental Protection Agency (EPA) Node of the National Geospatial Data Clearinghouse, and South African National Spatial Information Framework.

Evaluated applications for geospatial metadata access offer similar search capability: thematic filtering, selection on a limited set of non-spatial attributes, and selection on geographical location. Some applications offer a basic and an advanced query interface. The most advanced solution, in the authors opinion, is FGDC Clearinghouse which contains links to geospatial resources in other national spatial data infrastructures and provides the most powerful interface for metadata searching.

Adresy

Piotr BAJERSKI: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16,
44-101 Gliwice, Polska, bajerski@zti.iinf.polsl.gliwice.pl .

Krzysztof LACH: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16,
44-101 Gliwice, Polska, lach@zti.iinf.polsl.gliwice.pl .