

Przemysław Loranty

Rola metadanych w systemach informacji przestrzennej

Acta Scientifica Academiae Ostroviensis nr 30, 45-59

2008

Artykuł został zdigitalizowany i opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

Przemysław Loranty

ROLA METADANYCH W SYSTEMACH INFORMACJI PRZESTRZENNEJ

Rozwój informacji geoprzestrzennych – tworzenie danych dla wielu zastosowań

W czasach tradycyjnej geodezji i kartografii pozyskiwanie i dystrybucja geoinformacji były procesami wysoce scentralizowanymi i zmonopolizowanymi przez władze państwowe. Wynikało to głównie z wysokich kosztów i długiego czasu realizacji przedsięwzięć geodezyjnych i kartograficznych, których końcowym rezultatem były mapy o treści dostosowanej do podstawowych zadań państwa, zwłaszcza militarnych i fiskalnych. Potrzeby obywateli miały ograniczone znaczenie, a mapy traktowano przede wszystkim jako dobro państwowe, w mniejszym zaś stopniu jako produkty komercyjne. Państwo określało zatem rodzaje informacji prezentowanych na mapach. Z reguły były to informacje katastralne, topograficzne i ogólnogeograficzne, które stawały się warstwami podkładowymi dla tematycznych opracowań kartograficznych. W minionych latach sytuacja uległa radykalnej zmianie pod wpływem rozwoju metod i technik pozyskiwania i przetwarzania danych geoprzestrzennych. Niemal każdy może obecnie tworzyć swoje własne mapy i stosować geoinformację zgodnie z potrzebami, korzystając z komputerów osobistych, Internetu, GIS, GPS, obrazów satelitarnych, technik skanowania i inteligentnego oprogramowania. W ten sposób monopolistyczna pozycja państwa została poważnie ograniczona. Postęp technologiczny ułatwił znacznie przetwarzanie i wizualizację danych geoprzestrzennych, nie wpłynął jednak równie pozytywnie na możliwości uzyskiwania tych danych dla wykonywania konkretnych zadań użytkowych. Różne są tego przyczyny, np. użytkownik może stwierdzić, że potrzebne mu dane nie istnieją, ale nawet jeśli istnieją, to:

- często trudno do nich dotrzeć i sprawdzić ich użyteczność,
- ich jakość i szczegółowość jest nieodpowiednia lub też reprezentują one nieodpowiedni model,
- obowiązujące przepisy zbyt utrudniają uzyskanie dostępu do danych,
- opłaty za korzystanie z danych są za wysokie,
- użycie danych jest zbyt trudne pod względem technicznym, np. ze względu na różnice standardów i formatów.

W efekcie ponoszone są straty w skali społeczeństwa: decydenci i wszelkiego rodzaju użytkownicy geoinformacji mają utrudniony dostęp do danych, a prace w zakresie ich pozyskiwania, w tym prace pomiarowe, są niepotrzebnie powtarzane. W obecnej, nowej fazie rozwojowej GIS wymienione powyżej niekorzystne zjawiska nasiliły się w sposób znaczący ze względu na to, że:

- zwiększa się liczba i zróżnicowanie podmiotów zajmujących się pozyskiwaniem, przetwarzaniem i dystrybucją danych geoprzestrzennych,
- wzrasta liczba aplikacji¹ geoinformacyjnych, produktów i usług geoinformacyjnych oraz stosowanych formatów,
- coraz trudniejsze jest korzystanie z danych zgromadzonych przez różne podmioty,
- rosną koszty integrowania danych pochodzących z różnych źródeł oraz koszty nieuzasadnionego, powtórnego pozyskiwania danych.

Wobec tej narastającej różnorodności istotne jest, aby pewne najważniejsze dane geoprzestrzenne były jednolite, kompletne i utrzymywane w stanie aktualności. Dane takie, nazywane podstawowymi (*base, framework, fundamental*), z reguły odpowiadają dwóm kryteriom:

- stosowane są przez niemal wszystkich użytkowników, stanowiąc zasób głównych, wspólnych danych (*core data*),
- umożliwiają identyfikowanie innych danych lub obiektów, stanowiąc dla nich dane odniesienia zwane również referencyjnymi (*reference data*).

Przykładem danych podstawowych mogą być dane opisujące jednostkę podziału terytorialnego: są one potrzebne wielu użytkownikom i umożliwiają przyporządkowanie tej jednostce różnego rodzaju danych tematycznych.

Dane podstawowe stanowią ważny komponent każdej infrastruktury danych przestrzennych. Utworzenie tego komponentu wymaga m.in.:

- określenia warstw danych podstawowych oraz ustalenia specyfikacji ich treści,
- opracowania procedur, technik i wytycznych niezbędnych dla integrowania, udostępniania i stosowania tych danych,
- ukształtowania relacji instytucjonalnych i praktyk komercyjnych sprzyjających tworzeniu i utrzymywaniu danych oraz ich powszechnemu użytkowaniu.

¹ W znaczeniu programu użytkowego.

W wyniku tych działań oraz związanych z nimi prac wdrożeniowych, powstaje baza informacyjna, na podstawie której mogą być budowane tematyczne zasoby danych różnych organizacji. Podejście tego rodzaju przyczynia się do usprawnienia wymiany danych między zainteresowanymi stronami.

Za tworzenie zasobów danych podstawowych oraz zarządzanie nimi odpowiedzialne są przede wszystkim organy rządowe i samorządowe, które udostępniają te dane nieodpłatnie, za odpłatnością równą kosztom powielenia lub na zasadach komercyjnych². Dla osiągnięcia pełnego powodzenia w tym zakresie niezbędne jest jednak zapewnienie należytego współudziału innych producentów i użytkowników tych danych, jak też dostawców odpowiednich usług informacyjnych, systemów, sprzętu i oprogramowania. Realizacja całości przedsięwzięcia wymaga wielu lat pracy, jednakże wyniki w pełni użyteczne mogą być uzyskiwane już po zakończeniu poszczególnych zadań i etapów wdrożeniowych.

W jednym z pierwszych etapów prac powinny być opracowane specyfikacje obiektów (*features*³) i modeli zawartości (*content*) dla danych podstawowych, a w dalszej kolejności – dla pozostałych danych SDI. Służą do tego celu międzynarodowe normy Komitetu Technicznego ISO/TC211 *Informacja geograficzna/Geomatyka*⁴ tworzące rodzinę norm 19100. Wśród nich najważniejsze są tu dwie:

² J. Gaździcki, *Ochrona i udostępnianie baz danych geoprzestrzennych: płacić czy nie?*, „Geodeta” 6/2002.

³ Termin *feature* odnosi się do jednego z najważniejszych pojęć geomatematyki. Norma ISO 19101 podaje następującą definicję: *feature – abstraction od real world phenomena*, czyli jest to abstrakcja zjawisk świata realnego, gdzie *abstrakcją* jest wytwór abstrahowania (wyodrębniania elementów i właściwości uznanych za istotne z pominięciem innych), a *zjawisko* należy traktować szeroko jako przedmiot postrzegania i rozważania. Termin *feature* może być stosowany dla typu (w sensie klasy) lub pojedynczego egzemplarza, np. dom jako typ, czyli zbiór domów, lub dom jako konkretny budynek. Termin ten bywa również stosowany w znaczeniu informacji geoprzestrzennej, reprezentującej rozpatrywaną abstrakcję zjawisk. Nie ma, niestety, powszechnie przyjętego polskiego odpowiednika *feature*, chociaż proponowano już użycie słów *cecha*, *wyróżnienie* oraz *element*. W kompendium przyjęto zatem słowo *obiekt*, które wydaje się być łatwiejsze do zaakceptowania, przyporządkowując mu znaczenie wynikające z interpretacji podanej wyżej definicji ISO: *obiekt – abstrakcja zjawisk świata realnego, które są rozpatrywane w określonej przestrzeni*. Użycie słowa *obiekt* w innych znaczeniach będzie odpowiednio komentowane w tekście.

⁴ W Polsce problematyką tą zajmuje się Komisja Problemowa nr 297 Polskiego Komitetu Normalizacyjnego.

- standard 19109 *Rules for application schema* (*Reguły dla schematu aplikacyjnego*),
- standard 19110 *Feature cataloguing methodology* (*Metodyka katalogowania obiektów*).

Pierwsza z tych norm zajmuje się schematem aplikacyjnym jako formalnym opisem modelu pojęciowego, który dotyczy danych występujących w określonej aplikacji lub grupie aplikacji (gdzie przez aplikację rozumie się zastosowanie tych danych zgodnie z wymaganiami użytkownika). Schemat aplikacyjny określa:

- zawartość i strukturę (*structure*) danych,
- specyfikacje operacji manipulowania danymi oraz ich przetwarzania w ramach aplikacji.

Schematy aplikacyjne spełniają bardzo ważną rolę, służąc do:

- opisywania danych, z uwzględnieniem ich struktur, w sposób czytelny dla komputera, co pozwala na używanie zautomatyzowanych mechanizmów zarządzania danymi,
- dokumentowania treści danych, co pozwala na jednoznaczne i poprawne rozumienie danych oraz ich interpretowanie, celem uzyskania informacji.

Z normą 19109 blisko związana jest druga wspomniana norma –19110, która dotyczy katalogowania obiektów, tj. tworzenia katalogu mającego charakter słownika i zawierającego definicje i opisy typów obiektów, atrybutów obiektów oraz powiązań między obiektami, łącznie z dopuszczalnymi operacjami odnoszonymi do skatalogowanych obiektów. Katalog obiektów (np. topograficznych) opracowywany jest na podstawie schematu aplikacyjnego, stanowiąc przejrzysty i wyczerpujący dokument usprawniający współpracę użytkowników oraz producentów danych i oprogramowania.

Katalogowanie obiektów, które są przecież abstrakcjami o określonych atrybutach, reprezentacjach geometrycznych i operacjach, wiąże się z problemem stosowania jednolitych, trwałych identyfikatorów dla odpowiadających tym abstrakcjom obiektów świata rzeczywistego. Uzyskano już pozytywne wyniki w stosowaniu tego rodzaju identyfikatorów.

Jak pokazuje doświadczenie, w państwowych infrastrukturach danych przestrzennych jako dane podstawowe⁵ przyjmuje się na ogół dane należące do następujących warstw:

- kataster,
- osnowa geodezyjna,
- nazwy geograficzne,
- ortoobrazy,
- rzeźba terenu,
- transport,
- hydrografia,
- podział terytorialny.

Znaczący postęp uzyskano w zakresie określenia danych, zwłaszcza podstawowych, dla infrastruktur regionalnych⁶ i infrastruktury globalnej.

Zalecenia

Opracowanie specyfikacji danych podstawowych jest zadaniem żmudnym i trudnym do wykonania przez pojedynczą organizację. Odnosi się to do każdej infrastruktury danych przestrzennych, w tym do infrastruktury globalnej. Przyjęto dwa przedstawione poniżej zalecenia.

1. Inicjatywy i prace dotyczące danych podstawowych infrastruktur różnych poziomów powinny być podejmowane i prowadzone z możliwie szerokim udziałem zainteresowanych partnerów. Specyfikacje przyjęte dla Global Map mogą być z powodzeniem adaptowane i rozszerzane w projektach o zasięgu międzynarodowym.

2. W projektowaniu państwowych⁷ SDI do modelowania danych należy stosować normy ISO 19109 oraz 19110⁸.

⁵ Ogólnie rzecz biorąc, można przyjąć, że w Polsce geoprzestrzenne dane podstawowe są objęte krajowym systemem informacji o terenie, określonym rozporządzeniem Ministra Rozwoju Regionalnego i Budownictwa z roku 2001.

⁶ Dane określone powyższą listą warstw, po nieznacznej modyfikacji tej listy, uznane zostały za podstawowe w dokumencie *INSPIRE Reference Data and Metadata Position Paper*, October 2002. Nazwano je ogólnie danymi referencyjnymi (*reference data*). Zmodyfikowana lista warstw (komponentów) przedstawia się, jak następuje: 1) geodezyjne dane odniesienia, 2) jednostki administracyjne, 3) jednostki praw własności (działki, budynki), 4) adresy, 5) wybrane tematy topograficzne (hydrografia, transport, wysokości), 6) ortoobrazy, 7) nazwy geograficzne. Przyporządkowano im siedem aspektów: 1) geodezyjny system odniesienia, 2) jakość, 3) utrzymanie, 4) zdolność do współdziałania, 5) rozdzielczość/skala i priorytety wdrożeniowe, 6) język i kultura, 7) metadane.

Standaryzacja w tym zakresie jest nieodzowna, wpływając na lepsze zaspokajanie potrzeb użytkowników infrastruktur, ułatwiając dostęp do rozproszonych zasobów geoinformacyjnych oraz umożliwiając osiągnięcie wymiernych korzyści ekonomicznych.

Metadane i ich rola

Metadane (*metadata*) są danymi o danych, a ujmując to bardziej konkretnie - danymi zawierającymi informacje o zbiorach danych. Koncepcja metadanych nie jest nowa i od dawna znajduje swój wyraz w katalogowaniu zbiorów bibliotecznych oraz sporządzaniu legend map. W związku z rozwojem technologii informacyjnych i telekomunikacyjnych, stosowanie metadanych dla zbiorów danych cyfrowych stało się nieodzowne. Odnosi się to zwłaszcza do zbiorów cyfrowych danych geoprzestrzennych. Każdy zbiór tych danych, aby był użyteczny, musi być scharakteryzowany przez dane zawierające odpowiedzi na pytania typu: **co, kto, gdzie, kiedy, jak**. Stosowanie metadanych dla zbiorów danych geoprzestrzennych przynosi liczne korzyści, a mianowicie:

- ułatwia zarządzanie zasobami danych w ramach organizacji odpowiedzialnej za te dane,
- ułatwia korzystanie z nagromadzonych zasobów zgodnie z aktualnymi potrzebami, a także stwarza możliwości korzystania z nich w przyszłości, gdy będą stanowiły materiały historyczne,
- pozwala na lepsze planowanie przedsięwzięć dotyczących pozyskiwania i aktualizacji danych,
- rozszerza krąg użytkowników danych geoprzestrzennych,
- umożliwia realizację istotnych usług w ramach infrastruktur danych przestrzennych.

Zbiory danych, które nie są zaopatrzone w odpowiednie metadane, mają znacznie mniejszą wartość, a nawet mogą stawać się całkowicie bezużyteczne. Jest rzeczą oczywistą, że metadane zależą od charakterystycznych cech zbioru, a zwłaszcza od typu danych (np. od tego, czy są one wektorowe, czy też rastrowe).

Wyróżnia się trzy poziomy stosowania metadanych i związane z tymi poziomami trzy niżej wymienione rodzaje metadanych.

⁷ W sposób oczywisty zalecenie to dotyczy także infrastruktur na poziomie niższym od krajowego, np. w Polsce na poziomie miasta lub województwa.

⁸ Stosowanie tych norm zalecane jest również przez INSPIRE.

1. **Metadane wyszukiwania** (*Discovery metadata*) służą do wybrania zbiorów, które mogą być przedmiotem zainteresowania użytkownika o określonych wymaganiach. Metadane te obejmują:

- nazwę i opis zbioru danych,
- podstawowe przeznaczenie i zakres stosowania danych,
- datę pozyskania danych i ich aktualizacji,
- producenta, dostawcę i głównych użytkowników danych,
- obszar, do którego dane się odnoszą, określony przez współrzędne, nazwy geograficzne lub jednostki podziału administracyjnego,
- strukturę zbioru i sposób dostępu do danych.

2. **Metadane rozpoznania** (*exploration metadata*) zawierają bardziej szczegółowe informacje o zbiorze, które umożliwiają:

- ocenę właściwości danych zbioru,
- określenie przydatności danych zbioru pod względem wymagań użytkownika,
- nawiązanie kontaktu z dysponentem danych celem uzyskania dalszych informacji, w szczególności informacji na temat warunków korzystania z danych.

3. **Metadane stosowania** (*exploitation metadata*) określają te właściwości zbioru, które są potrzebne do:

- odczytania danych oraz ich transferu,
- interpretacji danych i praktycznego korzystania z nich w aplikacji użytkownika.

Ten ostatni rodzaj metadanych zawiera często dane objęte katalogiem obiektów, dane dotyczące struktur danych i odwzorowania kartograficznego oraz wszelkie parametry ułatwiające człowiekowi i maszynie właściwe korzystanie z danych geoprzestrzennych. Trzy poziomy stosowania metadanych i odpowiadające im rodzaje metadanych tworzą hierarchiczną strukturę wyborów (decyzji) dokonywanych przez użytkownika i umożliwiających ustalenie, jakie zbiory danych znajdują się w zakresie jego zainteresowania, które z nich odpowiadają jego wymaganiom, jak do nich dotrzeć, a także jak przetransferować wyselekcjonowane dane oraz zastosować je we właściwy sposób, odpowiadający potrzebom użytkownika.

Ogólne stwierdzenie, że metadane opisują zbiory danych, nie jest wystarczające. Np. metadane mogą dotyczyć zbioru jednorodnych map cyfrowych, pojedynczej mapy, grupy obiektów danego typu lub pojedynczego obiektu.

Świat realny jest skomplikowany i może być przedstawiany za pomocą modeli cyfrowych o różnym stopniu szczegółowości, z uwzględnieniem różnych jego cech. Aby przezwyciężyć wiążące się z tym trudności, należy wprowadzić odpowiednie, dostatecznie ogólne standardy metadanych.

Standaryzacja metadanych jest niewątpliwie bardzo potrzebna. Stosowanie powszechnie akceptowanych standardów w tym zakresie znacznie usprawnia procesy korzystania z metadanych, a zwłaszcza ułatwia porównywanie i ocenianie zbiorów danych pochodzących z różnych źródeł. Standaryzacja wpływa również korzystnie na powstawanie zunifikowanego oprogramowania dotyczącego metadanych. Konsekwencją zrozumienia ważności tej problematyki były stosunkowo liczne inicjatywy opracowania i wdrażania norm metadanych. Znaczenie międzynarodowe mają normy niżej wymienione:

- Norma Federalnego Komitetu Danych Geograficznych (*Federal Geographic Data Committee – FGDC*) Stanów Zjednoczonych wprowadzona w roku 1994 do użycia w Państwowej Infrastrukturze Danych Przestrzennych (*National Spatial Data Infrastructure – NSDI*), a w roku 1998 znowelizowana. Norma ta została zastosowana także poza Stanami Zjednoczonymi, m.in. w Kanadzie, Wielkiej Brytanii, Republice Południowej Afryki oraz w niektórych krajach Ameryki Łacińskiej i Azji.
- Norma 12657 Europejskiego Komitetu Normalizacyjnego (*Comité Européen de Normalisation – CEN*) opracowana przez Komitet Techniczny TC 287 i przyjęta w roku 1998 jako ENV (*Euro-Norme Voluntaire*).
- Norma 19115 Międzynarodowej Organizacji Normalizacyjnej (*International Standardization Organization – ISO*) opracowana przez Komitet Techniczny TC 211 i zatwierdzona w roku 2001 jako DIS (*Draft International Standard*).

Niezależnie od wymienionych wyżej inicjatyw opracowano normy o zasięgu krajowym lub regionalnym, zwracając szczególną uwagę na metadane podstawowe (*core, discovery*), najważniejsze pod względem użytkowym.

Przykładami są tu standardy zarządzane przez Australijską i Nowozelandzką Radę Informacji o Terenie (*The Australian and New Zealand Land Information Council – ANZLIC*) oraz standardy powstałe w wyniku projektów europejskich LaCLEF i ESMI.

Spośród wymienionych norm na wyróżnienie zasługuje norma ISO 19115 jako jedyna powszechnie akceptowana. Jej znaczenie uzasadniają przedstawione poniżej fakty:

- Norma ISO 19115 stanowi rezultat intensywnej i szerokiej współpracy międzynarodowej, którą prowadzono z uwzględnieniem bogatych doświadczeń zebranych przy opracowaniu i stosowaniu wszystkich pozostałych norm metadanych. Jest więc ona swego rodzaju syntezą uzasadnionych teoretycznie i sprawdzonych praktycznie (wcześniej zastosowanych) rozwiązań w tym zakresie.
- Implementacja tej normy obejmuje rozwój związanego z nią oprogramowania komercyjnego i jest objęta programem prac Open GIS Consortium.
- Norma ta będzie wprowadzana do użycia w powiązaniu z całą rodziną norm geomatycznych ISO/TC 211, co doprowadzi do integracji standardów treści i przekazu dla danych i metadanych. Poza normami dotyczącymi wyłącznie metadanych geoprzestrzennych, istnieją ogólne normy metadanych. Do takich należy międzynarodowy standard *Dublin Core* ułatwiający wyszukiwanie zasobów elektronicznych, zwłaszcza w sieci WWW. Standard ten stosowany jest w szczególności do dokumentowania i wyszukiwania zbiorów bibliotecznych, archiwalnych i muzealnych. W geomatyce znajduje zastosowanie jako standard pomocniczy służący do łączenia zasobów przestrzennych z nieprzestrzennymi.

Tworzenie metadanych przypomina katalogowanie książek w bibliotece, z tą różnicą, że wymagana jest znacznie głębsza wiedza i doświadczenie w dziedzinie, której dotyczą metadane. Nie każdy, kto ma przygotowanie zawodowe w tym kierunku, nadaje się do tej pracy. Potrzebne jest zrozumienie jej celu, solidność i dociekliwość. Zbiór metadanych może być prowadzony w różny sposób, zależnie od:

- wielkości zasobów danych,
- wielkości organizacji,
- zasad i technik zarządzania tymi zasobami w organizacji.

Jeśli zasób danych jest niewielki, wówczas metadane przechowuje się w oddzielnych dokumentach, stosując dowolne, dostatecznie dogodne oprogramowanie w rodzaju edytora tekstu, arkusza kalkulacyjnego lub prostej bazy danych. Dla dużych zasobów, zwłaszcza rozproszonych, powinny być stosowane środki bardziej zaawansowane w postaci specjalnego oprogramowania, najczęściej stanowiącego produkty komercyjne.

Oprogramowanie tego rodzaju może być powiązane z systemami geoinformacyjnymi, dostarczającymi pewnych metadanych. Przed przystąpieniem do opracowania metadanych należy ustalić, w jaki sposób będący przedmiotem zainteresowania zasób danych ma być podzielony na zbiory danych podlegające opisaniu metadanymi. Zależnie od charakteru zasobu i jego przeznaczenia można się tu kierować podziałami wynikającymi z organizacji i techniki pozyskiwania danych, ze sposobu funkcjonowania zasobu lub też z określenia produktów geoinformacyjnych przechowywanych w zasobie.

W procesie sporządzania metadanych dla jednoznacznie określonego zbioru danych geoprzestrzennych wyodrębnić można pewne typowe działania:

- zebranie informacji o zbiorze danych,
- utworzenie pliku zawierającego odpowiednio uporządkowane metadane,
- sprawdzenie i poprawienie pliku pod względem syntaktycznym,
- dokonanie końcowego przeglądu metadanych, celem sprawdzenia, że opisują one zbiór danych w sposób poprawny i kompletny.

Z reguły proces ten jest wspomagany przez odpowiednie oprogramowanie. Uzyskane metadane muszą być czytelne dla komputera oraz – po odpowiedniej prezentacji – dla człowieka. Przy sporządzaniu metadanych i korzystaniu z nich istotną rolę odgrywa tezaurs⁹ oraz słownik nazw geograficznych. Jest sprawą pierwszorzędnej wagi, aby urzędy i instytucje odpowiedzialne za zasoby geoinformacyjne stworzyły odpowiednie warunki dla rozwoju standaryzowanych metadanych. Wchodzą tu w grę odpowiednie środki prawne i ekonomiczne z uwzględnieniem pomocy technicznej, np. w postaci szkoleń i bezpłatnego oprogramowania. Godne naśladowania pod tym względem są działania wspomnianego powyżej Federalnego Komitetu Danych Geograficznych Stanów Zjednoczonych.

Zalecenia:

1. Zamiast tworzyć własną normę dotyczącą metadanych, należy przyjąć normę ISO 19115 lub utworzyć i zastosować jej krajowy profil¹⁰.

⁹ Tezauresem nazywa się tu uporządkowany zestaw terminów, służący do precyzowania ich wzajemnych powiązań; jest on wykorzystywany do określania treści zbiorów danych geoprzestrzennych lub ogólniej – do określania zawartości dokumentów.

¹⁰ Przez krajowy profil należy rozumieć podzbiór normy międzynarodowej, wynikający z potrzeb danego kraju. Z pełnej definicji podanej w ISO 19101 wynika, że profil normy serii 19100 powstaje przez określenie tych postanowień, klas, opcji i parametrów normy, które są niezbędne dla realizacji pewnej funkcji. Profil może być również tworzony na podstawie grupy norm serii 19100.

Uzyskuje się w ten sposób ewidentne korzyści, zarówno na poziomie krajowym (oszczędności czasu i środków, ułatwione wdrażanie), jak i międzynarodowym (wymiana informacji, ułatwiona współpraca)¹¹.

2. Program opracowania metadanych powinien uwzględniać priorytety wynikające ze znaczenia zbiorów danych geoprzestrzennych w zastosowaniach obecnych i przyszłych. W pierwszej kolejności należy zająć się danymi podstawowymi.

3. Metadane powinny być rejestrowane w trakcie pozyskiwania danych, tworzenia zbiorów danych, ich przetwarzania i aktualizacji.

4. Postuluje się opracowanie:

- systemu identyfikacji produktów geoinformacyjnych,
- systemu klasyfikacji tematycznej danych geoprzestrzennych.

Metadane i ułatwienie wyszukiwania danych

Dzięki metadanom nastąpiło ułatwienie wyszukiwania danych. Usługi¹² tego rodzaju nazywane są przez Open GIS Consortium katalogowymi, ponieważ ich podstawą jest katalog metadanych, lub też nadawana jest im ogólna nazwa *clearinghouse*, używana np. przez US FGDC. Usługi katalogowe mogą być łączone z innymi usługami dotyczącymi danych geoprzestrzennych, a polegającymi m.in. na integrowaniu, przetwarzaniu i wizualizacji tych danych¹³.

¹¹ W dokumencie *INSPIRE Reference Data and Metadata Position Paper* podano analogiczne zalecenia, a w szczególności następujące dwa: - należy opracować dla metadanych profil zgodny z ISO 19115 i wprowadzić go jako obowiązujący we wszystkich państwach członkowskich Unii Europejskiej ; - w każdym państwie członkowskim powinien być wyznaczony urząd odpowiedzialny za produkcję podstawowych danych geoprzestrzennych oraz zarządzanie systemami informacyjnymi metadanych. W *INSPIRE Architecture and Standards Position Paper* zaleca się dodatkowo stosowanie ogólnego standardu metadanych *Dublin Core* dla wyszukiwania informacji.

¹² W normie ISO 19101 podaje się następująca definicję: *service – capability which a service provider entity makes available to a service user entity at the interface between those entities*. Wynika stąd, że usługa (*service*) jest zdolnością działania, którą jednostka dostarczająca usługę udostępnia jednostce korzystającej z usługi; jednostki te kontaktują się ze sobą przez interfejs między nimi. Przez jednostkę (encję) rozumie się tu człowieka lub zautomatyzowany system.

¹³ *INSPIRE Architecture and Standards Position Paper* określa następujące rodzaje usług geoprzetwarzania (*geoprocessing*), jako aplikacji dostępnych w sieci WWW: 1) usługi zarządzania, administracji i koordynacji, 2) usługi katalogowe, dotyczące wyszukiwania danych, 3) usługi katalogowe dotyczące wyszukiwania usług, 4) usługi dotyczące map (*Web Map Server – WMS*), 5) usługi dotyczące pokryć (*Web Coverage Server – WCS*), patrz przypis 8, 6) usługi dotyczące obiektów (*Web Feature Server – WFS*), 7) usługi dotyczące nazw geograficznych, 8) usługi dotyczące transformacji współrzędnych, 9) usługi autoryzacyjne, 10) usługi w zakresie analiz i łączenia danych geoprzestrzennych, 11) usługi dotyczące zamawiania i rozliczania.

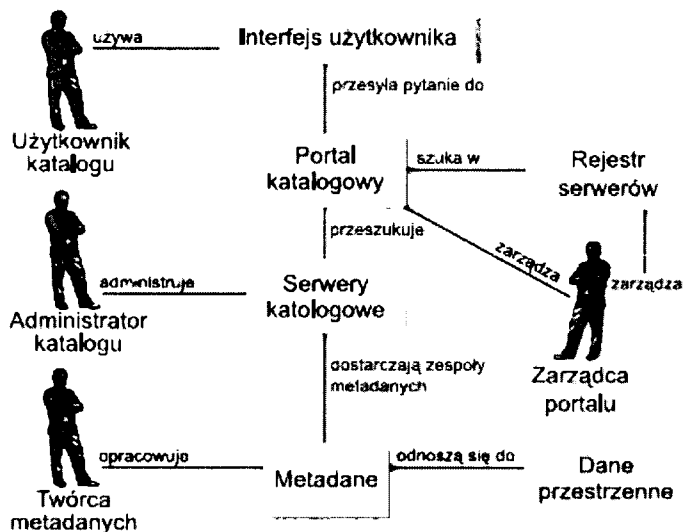


Diagram interakcji przedstawiający podstawowe użycie usług katalogowych w powiązaniu z odpowiednimi elementami SDI

W infrastrukturach danych przestrzennych znajduje zastosowanie koncepcja katalogu rozproszonego. Można sobie wyobrazić, że korzystając z odpowiedniego interfejsu, użytkownik zainteresowany odnalezieniem pewnych danych przestrzennych określa swoje pytania i życzenia, które są przekazywane do portalu katalogowego (*Catalogue Gateway, Catalogue Portal*).

W rezultacie działania portalu niektóre z zarejestrowanych serwerów, zwane katalogowymi, przeszukują swoje zasoby metadanych, odnajdując te odpowiadające życzeniom użytkownika.

U podstaw koncepcji katalogu rozproszonego leży pojęcie interoperacyjności lub też interoperowalności (*interoperability*¹⁴), czyli zdolności systemów do współdziałania.

¹⁴ *US Institute of Electrical and Electronics Engineers (IEEE)* podał następującą definicję: *interoperability* czyli dosłownie jest to zdolność dwóch lub więcej systemów lub komponentów do wymiany informacji oraz użycia informacji, która została wymieniona. Definicja IEEE została przyjęta przez INSPIRE oraz zastosowana w *The New Oxford Dictionary of English*, 2001. Jednowyrazowy polski odpowiednik słowa *interoperability* może być utworzony na podstawie przymiotnika *operacyjny* (*operating*): *interoperacyjność* albo *operowalny* (*operable*): *interoperowalność*. Druga z tych form jest bliższa znaczeniu angielskiego *interoperability* (zdolność do współdziałania), ale pierwsza jest już upowszechniona, i to w różnych znaczeniach, m.in. w wojskowości i transporcie.

Współdziałanie to ma miejsce na poziomie serwerów katalogowych, które mogą należeć do różnych SDI, ale posługują się wspólnym standardem metadanych, wspólnymi protokołami wyszukiwania i wspólnym rejestrem serwerów.

Koncepcję rozproszonego katalogu metadanych ilustruje powyższy rysunek, który stanowi diagram interakcji. Obejmuje on następujących aktorów¹⁵:

- użytkownika katalogu,
- administratora katalogu,
- zarządcę portalu,
- twórcę metadanych.

Treść diagramu należy interpretować, stosując następujące definicje:

- zespół metadanych – metadane opisujące wyodrębniony zbiór danych geoprzestrzennych, np. reprezentujących kolekcję obiektów (*feature collection*), obraz (*image*) lub pokrycie (*coverage*¹⁶),
- katalog – zbiór zespołów metadanych,
- usługa katalogowa – usługa wykonywana w odpowiedzi na zgłoszone pytanie lub zapotrzebowanie dotyczące metadanych objętych katalogiem.

Twórca zespołu metadanych jest odpowiedzialny za prawidłowe i zgodne z przyjętym standardem przedstawienie istotnych cech charakteryzujących opisywany zbiór danych geoprzestrzennych. Zależnie od instytucji i konkretnej sytuacji może to być naukowiec, który uczestniczył w pracach dotyczących utworzenia zbioru, wykonawca zleconych prac w zakresie pozyskiwania danych lub duża organizacja produkcyjna, która zgodnie z obowiązującymi przepisami zaopatruje swoje produkty informacyjne w odpowiednie opisy stanowiące metadane. Zebrane i uporządkowane metadane dotyczące opisywanego zbioru danych, zapisywane są w odpowiednich formatach za pomocą specjalnego oprogramowania, tworząc w rezultacie zespół metadanych tego zbioru danych.

¹⁵ ISO 19103 zawiera następującą definicję: *actor – coherent set of roles that users of an entity can play when interacting with the entity*, czyli aktor jest spójnym zbiorem ról, które użytkownicy pewnej jednostki (encji) mogą odgrywać, współdziałając z nią. Pojęcie to występuje w metodykach analizy i projektowania, odnosząc się do zewnętrznego obiektu modelującego rolę zewnętrznego użytkownika systemu.

¹⁶ Zgodnie z materiałami INSPIRE *pokrycie* można określić prosto, jako zbiór położeń przestrzennych (*spatial locations*) o znanych wartościach pewnych atrybutów.

Zarządca portalu odpowiada za rozwój i utrzymanie zdolności wyszukiwania w rozproszonych zasobach środowiska użytkowników. Jego zadania obejmują zarządzanie rejestrem serwerów, które stanowią elementy różnego rodzaju infrastruktur danych przestrzennych.

Administrator katalogu akceptuje wprowadzanie i usuwanie zespołów metadanych oraz udostępnia zgromadzone zasoby użytkownikom, przydzielając im odpowiednie uprawnienia. Zazwyczaj wykonuje on dodatkowo pewne funkcje kontroli jakości metadanych oraz inne funkcje administracyjne.

Podstawowym zadaniem **użytkownika katalogu** jest określenie kryteriów, które powinny umożliwić zlokalizowanie i zastosowanie potrzebnej mu informacji. Znając te kryteria, użytkownik wyszukuje informację poprzez:

- przeglądanie udostępnianych przez system list zawierających kategorie informacji, lub
- formułowanie zapytania według ustalonych zasad i wzorów.

W zakresie katalogów rozproszonych i związanych z nimi usług znajdują zastosowanie dokumenty standaryzacyjne, powstające w ramach ISO, Open GIS Consortium oraz World Wide Web Consortium.

Usługi w zakresie rozproszonego katalogu metadanych zostały rozwinięte m.in. w Stanach Zjednoczonych, Kanadzie, Meksyku, Australii i Republice Południowej Afryki.

Systemy oprogramowania opracowano, stosując normę ISO 23950 oraz standardy WWW. Prace były finansowane głównie ze środków rządowych.

Zastosowano trójpoziomą architekturę oprogramowania, obejmującą:

- poziom klienta (*client tier*), w którym występować może przeglądarka WWW lub aplikacja wyszukująca, powiązana bezpośrednio z wydzielonym interfejsem użytkownika,
- poziom pośredni (*middleware tier*), zawierający portal katalogowy z interfejsem użytkownika dla przeglądarek WWW oraz rejestr serwerów,
- poziom serwerów katalogowych (*server tier*).

Zalecenia

1. Zaleca się, aby metadane były publikowane i udostępniane przy zastosowaniu specyfikacji usług katalogowych, które są wynikiem prac Open GIS Consortium⁹. Prace wdrożeniowe w tym zakresie ułatwia istniejące już oprogramowanie.
2. Postuluje się, aby serwery katalogowe były zgłaszane do GSDI celem ich rejestracji.

Literatura:

1. red. Nebert D. D., *Kompendium infrastruktur danych przestrzennych*, oprac. wersji polskiej Jerzy Gaździcki.
www.gsdi.org/pubs/cookbook/kompendium-SDI.pdf
2. www.Google.pl
3. www.Geoforum.pl