

Małgorzata Leszczyńska  
Artur Janowski  
Jacek Górski

## ANALIZA PORÓWNAWCZA STANDARDÓW METADANYCH

### Streszczenie

W referacie przedstawiono ogólne spojrzenie na standardy metadanych z uwzględnieniem przydatności ich do transferu informacji geograficznych. Wzięto pod uwagę potrzebę stworzenia standardów, koncentrując się na rozwiązaniach odpowiednich dla polskich warunków. Uwzględniono zakres informacji ujętych w metadanych oraz ich wewnętrzną strukturę. Jako podstawę rozważań przyjęto dwa standardy metadanych: europejski i amerykański.

### 1. Podstawowe zagadnienia

Początki GIS-u sięgają schyłku lat sześćdziesiątych. Jednak w Polsce jego gwałtowny rozwój nastąpił dopiero z początkiem lat dziewięćdziesiątych. Z powodu tej różnicy czasowej nadal brakuje polskiego standardu metadanych, tymczasem czołówka światowa opracowania takie już posiada.

Nowa rzeczywistość rynku geoinformacyjnego wymusiła konieczność zunifikowania całości gromadzonych danych tak, aby stały się one dostępne i możliwe do wykorzystania przez wszystkie podmioty ich potrzebujące. Stan dzisiejszy jawi się niestety jako chaos mnogości danych pozyskanych z różnymi poziomami dokładności, aktualności i kompleksowości wykonania, wykluczających wspólne ich użycie w jednym konkretnym opracowaniu, co zaprzecza podstawowej cesze współczesnego GIS-u.

Dzisiejsze negatywne zjawisko „rozchwiania” rynku bazodanowego jest konsekwencją partykularnych potrzeb konkretnych inwestorów-odbiorców, na potrzeby których były one opracowywane. Każdy, kto cokolwiek wykonywał w tej dziedzinie, podejmował niemalże pracę od podstaw, bez możliwości pełnego wykorzystania istniejących już zbiorów danych. Wynikiem takiego postępowania była zawsze wysoka pracochłonność i, co z tym się wiąże, wysokie koszty przedsięwzięć.

Warunkiem sprawnej wymiany danych jest precyzyjne uzgodnienie potrzeb odbiorcy z ofertą przedstawioną przez dystrybutora. Związła i sformułowana według określonych reguł

(dla zachowania jednolitości) charakterystyka zbioru danych stanowi rodzaj „wizytówki” bądź „karty katalogowej”, umożliwiającej stwierdzenie przydatności lub nieprzydatności do konkretnego zastosowania. Unifikacja opisu danych nabiera wielkiej wagi wobec szybkiej rozbudowy dostępnego zasobu danych, obejmującego najróżniejsze dziedziny, między innymi dotyczące systemów informacji przestrzennych.

Doskonalenie techniki komputerowej w zakresie gromadzenia i przesyłania danych w postaci cyfrowej sukcesywnie rozwiązuje problem dostępności. Trudno się oprzeć wrażeniu, że dzięki wykorzystaniu Internetu wszystko leży w zasięgu ręki.

Znalezienie nawet niezłożonej informacji jest jednak trudne bez użycia słów kluczowych, a odszukanie zbioru danych przestrzennych, spełniającego ściśle określone, kryteria wymaga posłużenia się obszernym zestawem parametrów. Nieporozumienie prowadzi do chaosu informacyjnego i może poważnie zakłócić realizację podjętego przedsięwzięcia. Należy przy tym pamiętać, że decyzja o sprowadzeniu jakichkolwiek materiałów wiąże się z poniesieniem kosztów zakupu, często bardzo wysokich.

Zbiór nie opatrzony ujętą w znormalizowany sposób charakterystyką zawartości staje się po prostu bezwartościowy.

## 2. Standardy metadanych i ich zestawienie

Zestaw danych wyższego rzędu, czyli zawierający opis zbioru danych, przyjęto nazywać metadanymi. Istotne jest ściśle ustalenie zarówno zasobu treści, jak i wymogów formalnych. W tym celu opracowuje się standardy metadanych.

Metadane są grupowane pod względem rangi: wyróżnia się np. składniki niezbędne i uzupełniające (dodawane w określonych warunkach). Przejrzystość metadanych poprawia wyodrębnienie grup tematycznych, których przykładami mogą być: zasięg przestrzenny, rodowód i stan danych, parametry jakości, zasady dystrybucji.

Niniejsza praca jest głosem w dyskusji nad potrzebą zdefiniowania standardu metadanych dostosowanego do polskich realiów. Oparto ją na analizie dwóch różnych opracowań:

- *Geographic information – Data description – Metadata. European Prestandard ENV 12657, October 1998.* European Committee for Standardization (CEN)
- *Content Standards for Digital Geospatial Metadata. FGDC-STD-001-1998, June 1994.* Federal Geographic Data Committee (FGDC).

Porównanie dotyczy elementów uznanych przez oba Komitety za niezbędne, pominięto natomiast znacznie rozbudowaną (w obu przypadkach) część fakultatywną. Zestawienie ujęto w formie tabelarycznej (tab. 1.) z zachowaniem podziałów przewidzianych w dokumentacji źródłowej.

Tabela 1.

	ENV 12657	FGDC-STD-001-1998
<b>Identyfikacja zbioru danych</b>	1. Tytuł zbioru danych	1. Informacja używana jako odwołanie do zbioru danych
<b>Przegląd zbioru danych</b>	1. Streszczenie 2. Nazwa jednostki tworzącej 3. Sposób prezentacji przetrzeźni 4. Język zbioru danych (według ISO 639) 5. Strona kodowa zbioru danych (np. ISO 8859-10)	2. Opis (przewidywane zastosowania i ograniczenia) 3. Czasowy zakres aktualności 4. Status (stan utrzymania informacji, bądź zaawansowania prac) 5. Dziedzina przestrzenna
<b>Zasięg – zakres</b>	1. Stan zaawansowania prac oraz data ustalenia 2. Zasięg na płaszczyźnie (prostokąt lub obszar ograniczający lub nazwa własna) 3. Zasięg w pionie (maks. i min. wartość wysokości) 4. Zasięg w czasie (opis ogólny, data początkowa i końcowa)	6. Słowa kluczowe 7. Dostępność (ograniczenia) 8. Użyteczność (zastrzeżenia)
<b>Informacja o jakości danych</b>	Przynajmniej jeden z następujących elementów: 1. Rodowód danych 2. Parametry jakości 3. Dotychczasowe zastosowania 4. Jednolitość jakości w zbiorze	1. Dokładność atrybutu 2. Poprawność logiczna 3. Kompletność 4. Dokładność położenia przestrzennego 5. Rodowód danych
<b>Informacje o metadanych</b>	1. Data wprowadzenia 2. Data aktualizacji bądź kontroli	1. Data stworzenia, aktualizacji 2. Zarządca metadanych 3. Nazwa standardu metadanych 4. Wersja standardu 5. Rachuba czasu (np. lokalny)
<b>Przestrzenny system odniesienia</b>	1. Bezpośredni – zdefiniowany matematycznie lub 2. Pośredni – zdefiniowany przy pomocy nazw własnych	1. Definicja układu współrzędnych poziomych 2. Definicja układu współrzędnych wysokościowych
<b>Organizacja danych przestrzennych</b>		1. Odniesienie pośrednie (typy obiektów geograficznych, schematy adresowe i inne) 2. Odniesienie bezpośrednie (system obiektów użytych do opisu przestrzeni: punkt, wektor, raster)
<b>Definicja danych</b>	1. Opis przeglądowy 2. Typ obiektu (nazwa i opis) 3. Typ atrybutu (nazwa i opis) 4. Typ powiązania (nazwa występujących relacji)	1. Opis szczegółowy obiektów, atrybutów i ich wartości itp. i / lub 2. Opis ogólny (zawartość opisu szczegółowego i odniesienie do źródeł)
<b>Informacja o obiektach i ich atrybutach</b>		
<b>Klasyfikacja</b>	1. Nazwa dziedziny elementów lub 2. Zestaw terminów i definicji.	
<b>Metadane administracyjne</b>	1. Organizacja i jej rola (przynajmniej jeden z następujących elementów: nazwa, skrót nazwy, adres, rola). 2. Punkt kontaktowy i jego rola (nazwa, adres, rola) 3. Dystrybucja (ograniczenia używania, właściciele praw autorskich, jednostka dystrybucji)	1. Dystrybutor 2. Sygnatura zasobu 3. Odpowiedzialność dystrybutora 4. Sposób zamawiania (standardowy i niestandardowy)
<b>Informacja o dystrybucji</b>		

### 3. Funkcje metadanych

Bliższe spojrzenie na dwa przedstawione uprzednio standardy pozwala zauważyć, że nie w pełni odpowiadają one potrzebom polskich użytkowników.

W odniesieniu do polskich warunków wymagają one zbyt dużo informacji, które nie są istotne w naszym kraju. Przykładem może być tu punkt „rachuba czasu, np. czas lokalny”. W Polsce przecież nie ma stref czasowych, taka informacja byłaby więc zbędna. Należy zrezygnować z informacji nie mających zastosowania w Polsce, aby nie zabierać klientom czasu na wypełnianie zbędnych rubryk. Zbyt rozbudowane formularze zniechęcają odbiorców, przyczyniają się do błędów, a nawet mogą prowadzić do niezrozumienia potrzeby wprowadzania standardów metadanych. Adaptacja standardów nie polega jednak tylko na odrzucaniu zbędnych punktów, gdyż wtedy pracę można byłoby oprzeć tylko na jednym ze standardów. Istnieją też naszym zdaniem bardzo ważne informacje, które zostały ujęte tylko w jednym ze standardów, czego przykładem mogą być punkty: „zasięg w pionie”, „zasięg w czasie”, „zasięg na płaszczyźnie”.

Żaden ze standardów w swojej oryginalnej postaci nie zdołałby samodzielnie sprostać stawianym mu zadaniom. Z tego powodu bezpośrednia adaptacja jednego z nich nie jest wskazana. Rozwiązania problemu należy raczej szukać w drodze analizy, a następnie kompilacji wybranych elementów.

Metadane znajdują wykorzystanie w praktyce jako elementy wspomagające i usprawniające szeroko rozumianą wymianę informacji w postaci numerycznej, będąc inicjatorem ich transferu. W ścisłym ujęciu taką rolę spełniają metadane w trakcie poszukiwań danych o określonej charakterystyce i uzgodnień między użytkownikami (zastosowanie wyższego poziomu – zewnętrzne) oraz w procesie zautomatyzowanej komunikacji między konkretnymi aplikacjami tj. uwzględnienie protokołów i technologii wymiany danych na poziomie oprogramowania (zastosowanie niższego poziomu – wewnętrzne). Dlatego też każda próba tworzenia standardu metadanych powinna uwzględniać wspomnianą dwoistość zastosowania. Wynikiem takiego podejścia jest poniżej przedstawiona tabela 2. Przydatność poszczególnych elementów do każdej z wymienionych grup została oznaczona graficznie.

- ▣ - wymiana zewnętrzna;
- - wymiana wewnętrzna;
- ◼ - wymiana zewnętrzna i wewnętrzna.



Tabela 2.

<b>Przegląd zbioru danych</b>	<ol style="list-style-type: none"> <li>1. Nazwa jednostki tworzącej</li> <li>2. Sposób prezentacji przestrzeni</li> <li>3. Język zbioru danych (według ISO 639)</li> <li>4. Strona kodowa zbioru danych (np. ISO 8859-10)</li> <li>5. Opis (przewidywane zastosowania i ograniczenia)</li> <li>6. Status (stan utrzymania informacji, bądź zaawansowania prac)</li> <li>7. Słowa kluczowe</li> <li>8. Dostępność (ograniczenia)</li> </ol>	<input type="checkbox"/> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<b>Zasięg – zakres</b>	<ol style="list-style-type: none"> <li>1. Zasięg na płaszczyźnie (prostokąt lub obszar ograniczający lub nazwa własna)</li> <li>2. Zasięg w pionie (maks. i min. wartość wysokości)</li> <li>3. Zasięg w czasie (opis ogólny, data początkowa i końcowa)</li> </ol>	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>
<b>Informacja o jakości danych</b>	<ol style="list-style-type: none"> <li>1. Rodowód danych</li> <li>2. Parametry jakości</li> <li>3. Dokładność atrybutu</li> <li>4. Poprawność logiczna</li> <li>5. Kompletność</li> </ol>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>
<b>Informacje o metadanych</b>	<ol style="list-style-type: none"> <li>1. Data stworzenia, aktualizacji</li> <li>2. Zarządca metadanych</li> <li>3. Nazwa standardu metadanych</li> </ol>	<input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/>
<b>Przestrzenny system odniesienia</b>	<ol style="list-style-type: none"> <li>1. Definicja układu współrzędnych poziomych</li> <li>2. Definicja układu współrzędnych wysokościowych</li> </ol>	<input type="checkbox"/> <input type="checkbox"/>
<b>Informacja o obiektach i ich atrybutach</b>	<ol style="list-style-type: none"> <li>1. Typ obiektu (nazwa i opis)</li> <li>2. Typ atrybutu (nazwa i opis)</li> <li>3. Typ powiązania (nazwa występujących relacji)</li> </ol>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<b>Informacja o dystrybucji</b>	<ol style="list-style-type: none"> <li>1. Dystrybutor</li> <li>2. Sygnatura zasobu</li> <li>3. Odpowiedzialność dystrybutora</li> <li>4. Sposób zamawiania (standardowy i niestandardowy)</li> </ol>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

□ - wymiana zewnętrzna; ● - wymiana wewnętrzna; ◻ - wymiana zewnętrzna i wewnętrzna.

#### 4. Wnioski

Powstanie standardu metadanych zobowiązałoby wszystkich „twórców” danych do opracowania w jednolitym i przejrzystym systemie, ułatwiającym korzystanie z nich szerokiej rzeszy odbiorców. Opracowany i wprowadzony obligatoryjnie normatyw wyeliminowałby występującą do dziś dowolność akwizycji danych, tworząc tym samym, wzorem krajów przodujących w technologii GIS, podwaliny dla pełniejszego jej wykorzystania na terenie naszego kraju.

W świetle przedstawionych faktów nieuchronne wydaje się wykonanie i zatwierdzenie standardu metadanych, przy opracowaniu którego zasadną wydaje się analiza i uwzględnienie doświadczeń krajów stanowiących światową czołówkę posługujących się technologią GIS.

Recenzował: dr inż. Krystian Pyka