

Potrzeba przepływu informacji w systemach informacji geograficznej (GIS) wymusza konstruowanie modeli informacyjnych w kategoriach ogólnych, niezależnych od środowisk sprzętowo-programowych. Celowi temu służy metodyka modelowania pojęciowego, która jest przedmiotem normalizacji w skali międzynarodowej, europejskiej i krajowej. O uwarunkowaniach i potrzebach rynkowych takiej normalizacji była mowa w części I tego materiału (GEODETA 11/2002). Dziś przedstawione zostaną poglądowo (na przykładzie fikcyjnego modelu katastru) formalizmy UML i EXPRESS stosowane w powstających Polskich Normach. Natomiast za miesiąc zajmiemy się działalnością krajowej Normalizacyjnej Komisji Problemowej 297 i rolą środowiska geodezyjnego w kreowaniu i stymulowaniu rozwoju GIS w Polsce. Tezy artykułu odzwierciedlają oficjalne stanowisko Głównego Urzędu Geodezji i Kartografii w stosunku do tych koncepcji rozwojowych GIS w Polsce, które oparte są na normalizacji informacji geograficznej.

Poszukując rozwiązań umożliwiających przepływ informacji pomiędzy zróżnicowanymi środowiskami, zwłaszcza sprzętowo-programowymi, można rozważać m.in. następujące koncepcje:

- 1) Jednostkowa i w pełni scentralizowana realizacja kompletnego systemu informacyjnego (np. ogólnokrajowy kataster) na pojedynczej platformie sprzętowo-programowej.
- 2) Realizacja GIS w sposób rozproszony w ośrodkach terenowych (np. w powiatach), ale przy pełnej unifikacji platform sprzętowo-programowych.
- 3) Jak poprzednio, z dopuszczeniem różnorodności platform sprzętowo-programowych, ale za to z wymaganiami ścisłego przestrzegania określonych instrukcji i przepisów regulujących budowę GIS w kategoriach prawno-administracyjnych.

Działalność normalizacyjna w dziedzinie informacji

Metodyka i for

WOJCIECH PACHELSKI

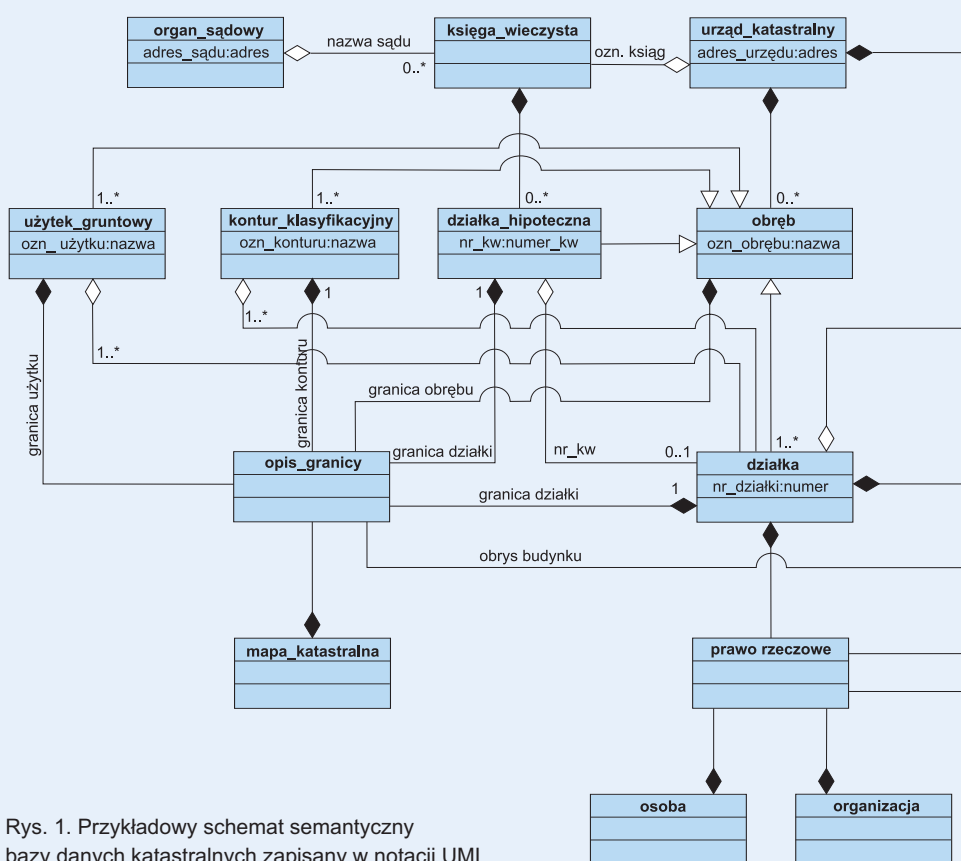
4) Budowa GIS w ośrodkach terenowych z uwzględnieniem tzw. standardów definiujących stałe struktury baz danych i plików, formaty i kody zapisu, procedury transferu danych itp.

5) Spójne realizacje GIS w ośrodkach terenowych według wcześniej skonstruowanych modeli pojęciowych.

Konceptje 1) i 2) zakładają unikalność bądź pełną jednolitość sprzętu i oprogramowania użytych dla celów danej dziedziny przedmiotowej. Łatwo przy tym zauważyć, że realizacje takie i tak musiałyby się kontaktować z wieloma innymi systemami informacyjnymi, zarówno krajowymi, jak i zagranicznymi, co oznaczałoby istnienie nadal problemu przepływu informacji pomiędzy nimi.

W obu przypadkach bądź to system byłby wysoce niefunkcjonalny wskutek konieczności wielokrotnego „dostrajania” go do częstych (w miarę postępu technicznego) zmian technologicznych sprzętu i oprogramowania, bądź też postępowałaby stagnacja technologiczna systemu w wyniku zaniechania takiego „dostrajania”.

Natomiast koncepcję 3), która dominuje w obecnej praktyce GIS w Polsce, cechuje nieprzystawalność, wieloznaczność i niekompletność naturalnego języka instrukcji i przepisów do informatycznego charakteru postulowanych aspektów, cech, rozwiązań itp., zarówno co do ogólnokrajowej koncepcji GIS, jak i każdej jej realizacji. Te cechy języka naturalnego¹, zawarte siłą rzeczy w tradycyjnych



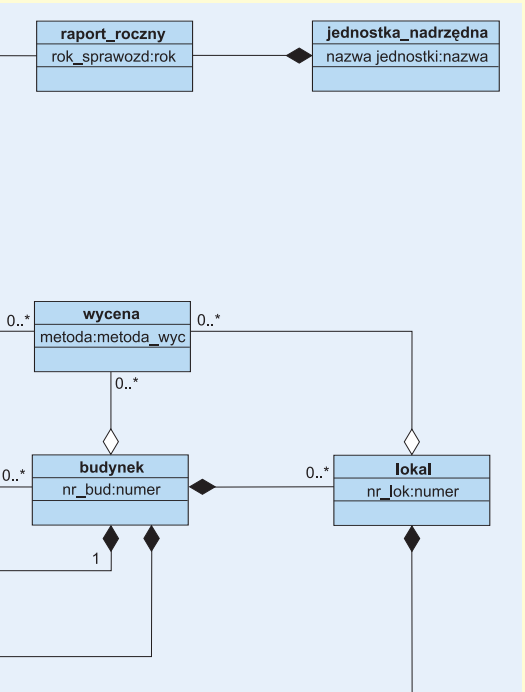
Rys. 1. Przykładowy schemat semantyczny bazy danych katastralnych zapisany w notacji UML

geograficznej, cz. II

malizm

instrukcjach i przepisach, łatwo mogą powodować, że np. powiatowe systemy katastralne, ściśle zgodne z tymi przepisami, nie będą zgodne między sobą i tym samym nie zapewnią pożądanego efektywnego przepływu informacji.

W ramach koncepcji 4) prace podjęte w wielu krajach i organizacjach doprowadziły do powstania różnych narodowych i instytucjonalnych standardów dla transferu danych², które jednakże umożliwiają przenoszenie danych jedynie w ograniczonym zakresie zastosowań. Ogólnie, standardy te definiują pewne stałe struktury, formaty i kody jako pośrednie pomiędzy formami stosowanymi przez konkretne platformy narzędziowe, pozostawiając na ogół użytkownikom prob-



```

SCHEMA schemat_opisu_dzialki;
.....
ENTITY punkt;
  x, y: REAL;
END_ENTITY;

ENTITY znak_graniczny;
  w_punkcie: punkt;
  numer_zg: numer_znaku;
  typ_zg: typ_znaku;
END_ENTITY;

ENTITY linia_graniczna;
  SUPERTYPE OF
    (ONEOF(granica_nat,
            granica_prawna));
  od_znaku, do_znaku:
    znak_graniczny;
  INVERSE
    wystepuje_w: SET[1:?] OF dzialka
    FOR granica;
WHERE
  w1: od_znaku :<>: do_znaku;
END_ENTITY;
ENTITY dzialka;
  punkt_odniesienia: punkt;
  granica: LIST[2:?]
    OF UNIQUE linia_graniczna;
  powierzchnia_rej: powierzchnia;
  numer: numer_dzialki;
  zabudowania: uzytek_w_dzialce;
  grunty_orne: uzytek_w_dzialce;
  lasy: uzytek_w_dzialce;
  nieuzytki: uzytek_w_dzialce;
  DERIVE
    powierzchnia_obl: powierzchnia :=
      oblicz_pow_dzialki(SELF)
WHERE
  w1: granice_tworza_wielobok
    (granica);
  w2: (ABS
    (powierzchnia_rej-
     powierzchnia_obl)<=
     powierzchnia_rej*0.05);
  w3: (zabudowania+grunty_orne+
    lasy+nieuzytki)
    = powierzchnia_rej
END_ENTITY;
.....
END_SCHEMA;
  
```

Przykładowy zapis fragmentu schematu pojęciowego katastru (opis działki) w języku EXPRESS

lem przekształcenia danych do i z tej formy pośredniej. Krzysztof Miksa (1997) wykazał tymczasem, że taki typ transferu danych może być co najwyżej pewnym rozwiązaniem szczególnym problemu przenoszenia danych o nader ograniczonej stosowalności, nie zaś – ogólnym, uniwersalnym. Może on bowiem dotyczyć jedynie pewnego z góry ustalonego

Początek deklaracji schematu.

Deklaracja encji o nazwie punkt, której atrybutami są współrzędne x i y, obie typu REAL. Inne pominięte tu deklaracje winny szczegółowo opisać te współrzędne, np. co do układu odniesienia.

Ta encja zawiera trzy atrybuty: w_punkcie, numer_zg i typ_zg, opisujące lokalizację znaku granicznego oraz jego numer i typ. Typy tych atrybutów winny być zadeklarowane oddzielnie (tutaj pominięto deklaracje typów numer_znaku i typ_znaku).

Encja linia_graniczna jest nadrzędna względem encji granica_nat i granica_prawna, tutaj pominiętych. Encja ta posiada dwa atrybuty, które odpowiadają znakowi początkowemu i końcowemu linii, oba typu znak_graniczny. Jej dziedzinę ogranicza się do atrybutu granica w zbiorze encji dzialka (patrz niżej) oraz dodatkowym warunkiem, że znaki początkowy i końcowy linii nie mogą być tym samym znakiem (symbol :<>:).

Główna encja niniejszego opisu, zawierająca wymienione atrybuty jawne o podanych i niezależnie zadeklarowanych typach, przy czym atrybut granica jest zadeklarowany jako lista nie powtarzających się danych typu encji linia_graniczna.

Oprócz atrybutów jawnych występuje atrybut pochodny powierzchnia_obl typu powierzchnia, który ma być wyliczony za pomocą zadeklarowanej funkcji oblicz_pow_dzialki.

Dziedzina ważności atrybutów jest ograniczona warunkami w1, w2 i w3, które przewidują: wywołanie funkcji standardowej ABS, wywołanie innych funkcji zadeklarowanych w schemacie oraz wykonanie pewnych operacji arytmetycznych i operacji relacji. Wyliczona wartość każdego z tych warunków musi być TRUE, by konkretne wartości atrybutów mogły należeć do tak opisanej bazy danych.

Inne specyfikacje.
Koniec deklaracji schematu.

grona użytkowników, pewnych zadanych warstw przedmiotowych, określonych instytucji i organizacji oraz konkretnych realizacji sprzętowo-programowych. Standardy te nie umożliwiają również transferu danych w warunkach coraz wyraźniejszych potrzeb przepływu informacji geograficznej w skali międzynarodowej i globalnej.

Koncepcję 5) można traktować jako logiczne połączenie zmodyfikowanych koncepcji 3) i 4), gdzie opis modelu informacyjnego w języku naturalnym zastąpiono ścisłym, jednoznacznym i kompletnym opisem za pomocą informatycznych środków formalnych, w miejsce zaś sztywnych i stałych struktur, formatów i kodów umożliwiono ich definiowanie w sposób dowolny³. Nowoczesne środki formalne, jak na przykład języki UML i EXPRESS, a także cała metodologia modelowania informacji geograficznej w postaci schematów pojęciowych oraz pewne typowe schematy, stały się przedmiotem normalizacji, a tym samym przedmiotem *powszechnego i wielokrotnego stosowania*⁴. Tym samym normalizacja informacji geograficznej w ramach tej koncepcji, oznacza przede wszystkim normalizację metodologii modelowania informacji oraz środków formalnych dla opisu modeli pojęciowych, nie zaś normalizację samych modeli pojęciowych w poszczególnych działach przedmiotowych (w niczym nie ogranicza to jednak możliwości normalizacji tych modeli w ramach osobnych działań).

Koncepcja ta jest zilustrowana na rys. 1 (s. 34-35) na przykładzie pewnego fikcyjnego modelu informacyjnego katastru zapisanego w notacji UML, jak również w tabeli (s. 35) jako zapisany w języku EXPRESS fragment pewnego fikcyjnego schematu pojęciowego katastru, opisujący model działki gruntu. W obu przypadkach opisane w taki sposób kompletne modele pojęciowe (schematy pojęciowe) pełnią wielorakie funkcje, a mianowicie:

- formy dokumentacyjnej reguł interpretacji danych, niezbędnych dla powszechnego i poprawnego ich rozumienia,
- czytelnej dla zróżnicowanych mediów komputerowych formy opisu danych, umożliwiającej zgodne stosowanie tych mediów do zarządzania danymi,
- uzgodnionego pomiędzy użytkownikami środka zapewniającego poprawne komunikowanie informacji geograficznej poprzez efektywny transfer danych z uwzględnieniem ich aspektów znaczeniowych,
- wspólnej podstawy dla zgodnych realizacji GIS w zróżnicowanych środowiskach narzędziowych.

Budowa modelu informacji geograficznej, której wynikiem jest pojęciowy schemat aplikacyjny (zilustrowany przykładowo na rysunku i w tabeli na poprzedniej stronie), jest procesem, na który składają się następujące etapy:

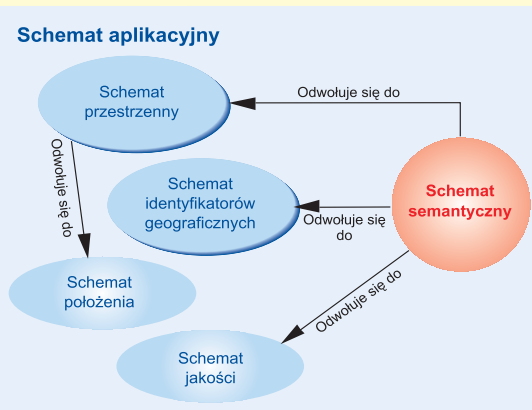
- ścisłe sformułowanie przedmiotu i zakresu modelowania oraz przegląd uwarun-

kowań zewnętrznych i wymagań użytkowników w danej dziedzinie zastosowań;

- modelowanie pojęciowe, polegające na sformułowaniu głównych encji (ew. klas), związków, atrybutów i ograniczeń w ustalonym zakresie przedmiotowym w postaci schematu semantycznego;

- powiązanie (czyli integracja) tak powstałego modelu ze schematami pojęciowymi zawartymi w normach do postaci kompletnego schematu aplikacyjnego.

Budowę takiego modelu z wykorzystaniem niektórych norm będących przedmiotem działania NKP 297 ilustruje rys. 2.



Rys. 2. Integracja schematu semantycznego ze schematami znormalizowanymi

Powyższa koncepcja modelowania pojęciowego informacji geograficznej, będąca przedmiotem normalizacji krajowej opartej na normach europejskich CEN i normach międzynarodowych ISO, będzie miała zastosowanie do typowych zadań związanych z racjonalnym rozwojem GIS w Polsce, jak:

- 1) realizacja pewnego wzorcowego (ew. standardowego lub znormalizowanego) schematu pojęciowego za pomocą różnych narzędzi sprzętowo-programowych i w różnych środowiskach instytucjonalnych;
 - 2) realizacja schematów pojęciowych pochodzących z różnych źródeł na tej samej platformie,
 - 3) harmonizacja istniejących aplikacji GIS.
- W pierwszym przypadku chodzi o wypracowanie i uzgodnienie wzorcowych schematów pojęciowych, przeznaczonych do wielorakich niezależnych realizacji w odmiennych środowiskach. Do grupy tej należą m.in. części składowe państwowego zasobu geodezyjnego i kartograficznego, jak: kataster, mapa zasadnicza, geodezyjna ewidencja sieci uzbrojenia terenu, osnowy geodezyjne i inne.

Drugi przypadek obejmuje natomiast sytuacje, kiedy wypełnienie określonego zadania bądź funkcji systemu informacji geo-

graficznej wymaga skojarzenia i łącznego opracowania danych pochodzących z co najmniej dwóch odmiennych przedmiotowo, narzędziowo i instytucjonalnie baz danych. Realizacja schematu pojęciowego (bądź jego fragmentu) „obcej” bazy danych na „własnej” platformie narzędziowej pozwala najpierw dokonać „importu” samych danych bez zmiany ich składni (struktur, formatów, kodów itp.), a następnie wykorzystać „własne” środki programowe (z uwzględnieniem aspektów znaczeniowych danych) do odpowiedniego przekształcenia tych danych do „własnej” składni.

Trzecia sytuacja sprowadza się właściwie do tego, by każda z istniejących już aplikacji, jakie mają podlegać uzgodnieniu, udokumentować niejako „wstecz” za pomocą schematu pojęciowego, przy wykorzystaniu znormalizowanych metod i środków formalnych (czyli Polskich Norm i języka EXPRESS). Dokumentacja taka jest warunkiem koniecznym i wystarczającym do tego, by w odpowiednich ośrodkach realizacyjnych były dostępne i rozumiane informacje o strukturach „obcych” baz danych, co w konsekwencji winno umożli-

wić efektywny transfer danych, tak jak w przypadku drugim.

cdn.

Prof. Wojciech Pachelski jest przewodniczącym Normalizacyjnej Komisji Problemowej Nr 297 ds. *Informacji geograficznej*. Opracowanie na podstawie referatu wygłoszonego podczas konferencji „Jakość, standaryzacja, normalizacja w geodezji i kartografii”, Pogorzela, 3-5 października 2002 r. Spis literatury został umieszczony w cz. I (GEODETA 11/2002).

¹ Jedną z takich cech jest to, że w języku naturalnym można poprawnie wyrazić zdanie błędne znaczeniowo, podczas gdy w języku formalnym jest to niemożliwe. Fakt ten umożliwia m.in. kontrolę wewnętrznej (formalnej) poprawności modeli informacyjnych zapisanych w językach formalnych.

² Przykładami takich standardów są: NTF (Wielka Brytania), EDIGeO (Francja), DIGEST (NATO), SDTS (USA), CCOGIF, SAIF (Kanada), DX90 (Międzynarodowa Organizacja Hydrologiczna – IHO), ATKIS (Niemcy).

³ Do grupy tej zalicza się opublikowany w 1995 r. krajowy Standard Wymiany Informacji Geodezyjnej SWING, jak również Standard Wymiany Danych Ewidencyjnych SWDE, wprowadzony rozporządzeniem ministra rozwoju regionalnego i budownictwa z 29 marca 2001 w sprawie ewidencji gruntów i budynków (DzU nr 38, poz. 454).

⁴ Cytat z definicji normy według PN-N-02000:1994 (Chowańska-Szwoch, 2000).