

Ewa Dębińska\*, Piotr Cichociński\*\*

ZASTOSOWANIE NARZĘDZI CASE DO PROJEKTOWANIA BAZ DANYCH  
SYSTEMÓW INFORMACJI GEOGRAFICZNEJ\*\*\*

---

## 1. Wprowadzenie

Podstawą efektywnego korzystania z systemów informacji geograficznej (SIG) jest poprawnie skonstruowana baza danych, pozwalająca na szybkie i sprawne uzyskiwanie informacji na interesujący temat. Ze względu na specyfikę SIG, konieczne jest, by taka baza oprócz danych opisowych obiektów przechowywała również informacje dotyczące położenia i kształtu tych obiektów. Jeszcze kilka lat temu z powodu ograniczonej wydajności sprzętu komputerowego i oprogramowania baz danych powszechnie wykorzystywany był tak zwany georelacyjny model danych. Dane przestrzenne (informacja o kształcie i położeniu obiektów) były gromadzone w indeksowanych plikach binarnych, zoptymalizowanych pod kątem szybkości wyświetlania i dostępu. Atrybuty (dane opisowe) zapisywano w tabeli o liczbie wierszy odpowiadającej liczbie obiektów w pliku binarnym, dodatkowo były one powiązane z obiektami poprzez wspólny identyfikator. Pomimo podziału danych przestrzennych i atrybutowych, georelacyjny model danych dominował przez długie lata. Główny powód był jeden – możliwość efektywnej obsługi złożonych zbiorów danych. Jednakże model ten miał również istotne ograniczenia. Jednym z głównych był fakt grupowania obiektów w jednorodne zbiory punktów, linii i obszarów. Spowodowane to było między innymi zapisem topologii, która na przykład nie dopuszczała przecinających się linii.

---

\* Doktorantka, Akademia Górniczo-Hutnicza, Wydział Geodezji Górniczej i Inżynierii Środowiska, Katedra Informacji o Terenie

\*\* Akademia Górniczo-Hutnicza, Wydział Geodezji Górniczej i Inżynierii Środowiska, Katedra Informacji o Terenie

\*\*\* Praca naukowa finansowana ze środków na naukę w latach 2005–2008 jako projekt badawczy nr 4 T12E 016 29

Ponadto wzbogacenie tych prostych obiektów o możliwości zachowań wymagało pisania specjalistycznych aplikacji.

Jednakże postęp, jaki dokonał się w ostatnich latach w zakresie wydajności sprzętu komputerowego oraz możliwości oprogramowania baz danych, pozwolił w końcu na zapisanie danych przestrzennych bezpośrednio w relacyjnej bazie danych. Bazę taką, zawierającą informacje geograficzne, co ją w sposób szczególny wyróżnia spośród innych baz danych, można nazwać geobazą. Podstawowym zastosowaniem geobazy jest manipulowanie złożonymi danymi geograficznymi z jednolitym modelem danych niezależnym od konkretnej relacyjnej bazy danych, na której jest oparty. Zaletą gromadzenia danych przestrzennych w komercyjnych systemach bazach danych jest bardziej zaawansowane zarządzanie danymi, wykorzystanie usług dostępu do danych i lepsza integracja z innymi bazami prowadzonymi w ramach instytucji.

## 2. Geobaza

Geobaza jest fizyczną reprezentacją obiektów istniejących w rzeczywistym świecie. Umożliwia przechowywanie danych przestrzennych (geometrycznych, opisowych, rastrowych oraz modeli TIN) w systemie zarządzania bazą danych.

Atutem geobazy, jako formatu przechowywania danych, jest swobodny dostęp do danych, który umożliwia użytkownikom tworzenie, wykorzystywanie i operowanie danymi geograficznymi. Dzięki otwartości technologii i ogólnie dostępnej dokumentacji praca z geobazą nie wymaga wykorzystania oprogramowania konkretnej firmy.

W geobazie dane modeluje się z wykorzystaniem metodologii obiektowej [8], co pozwala na opisanie rzeczywistości w sposób bardziej naturalny. Takie podejście do danych sprzyja lepszemu zrozumieniu, jak obiekty oddziałują wzajemnie na siebie.

Do zalet geobazy można zaliczyć m.in.:

- przechowywanie i zarządzanie wszystkimi danymi dotyczącymi jednego projektu centralnie w jednej geobazie;
- możliwość pracy z intuicyjnymi obiektami; poprawnie zaprojektowana geobaza zawiera obiekty, które nawiązują do modelu danych użytkownika; wówczas użytkownik pracuje na obiektach jego zainteresowań takich jak graniczniki, drogi i działki zamiast obiektów ogólnych takich jak punkty, linie czy poligony;
- szersze możliwości definiowania kształtów obiektów – oprócz linii prostych dostępne są również krzywe eliptyczne i Beziera;
- możliwość przechowywania zbiorów danych dużych rozmiarów, bez konieczności dzielenia ich na mniejsze części;
- możliwość opracowywania jednocześnie tych samych danych przez kilku użytkowników;
- rozbudowane mechanizmy wprowadzania, edycji, kontroli i aktualizacji danych.

W odróżnieniu od wcześniejszych relacyjnych modeli danych, w których każdy obiekt i jego atrybuty były zapisywane jako wiersz w tabeli, geobazy przechowują informacje w relacyjno-obiektowych bazach danych, w których zachowane są sprawdzone technologie relacyjne (w szczególności język zapytań SQL – *Structured Query Language*), przy jednoczesnym wprowadzaniu koncepcji obiektów w celu rozszerzenia pojęcia tabeli oraz typów danych wykorzystywanych w modelu relacyjnym. Relacyjne bazy danych dominują na rynku, ponieważ oparte są o prostą, elegancką i dobrze zrozumiałą teorię. Ta prostota jest jednocześnie zaletą i wadą – jest pojęciowo prosto zbudować relacyjne bazy danych, lecz trudno modelować złożone dane.

Geograficzne bazy danych zawierają niestety dane złożone. Kształty obiektów liniowych i powierzchniowych są zestrukturyzowanymi zbiorami współrzędnych, które nie zapisują się dobrze w polach standardowych typów, takich jak: całkowity, rzeczywisty czy tekstowy. Ponadto obiekty są łączone w systemy, które mają jawne związki topologiczne, niejawne związki przestrzenne i inne związki natury ogólnej. Jednak konstruktorom baz danych udało się obejść to ograniczenie, co pozwoliło na zastosowanie metod projektowania relacyjnych baz danych również do informacji geograficznej.

### 3. Projektowanie baz danych

Można wyróżnić trzy etapy projektowania baz danych [2]:

1. zbudowanie pojęciowego modelu danych obejmującego zgromadzenie informacji na temat obiektów, relacji i atrybutów;
2. stworzenie logicznego modelu bazy danych, czyli przekształcenie pojęciowego modelu danych na logiczną strukturę bazy danych;
3. fizyczna implementacja logicznego modelu danych.

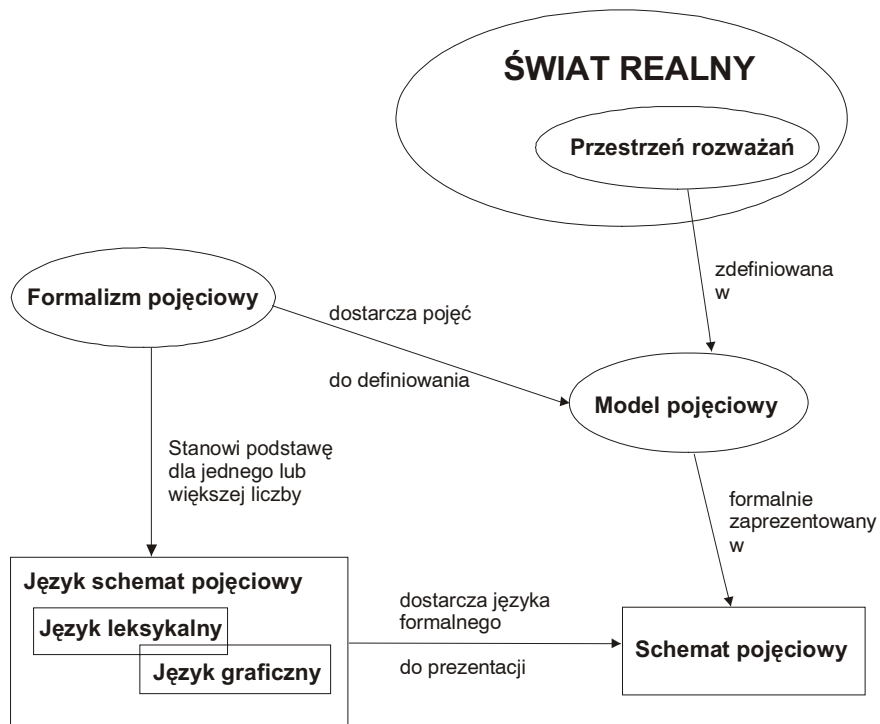
Modele: pojęciowy i logiczny odpowiadają spojrzeniu na dane przez użytkownika, a model bazy danych implementuje model danych w ramach technologii relacyjnej bazy danych.

#### 3.1. Modelowanie pojęciowe

Budowanie modelu pojęciowego obejmuje szereg procesów myślowych i wyobrażeń dotyczących projektu. Projektant musi dokładnie wyobrazić sobie problem oraz metodę jego rozwiązania. Zasadnicze procesy tworzenia bazy danych zachodzą na początku w ludzkim umyśle i nie są powiązane ani z konkretnym językiem programowania ani też z jakimkolwiek narzędziem.

Rysunek 1 prezentuje relacje pomiędzy światem rzeczywistym a schematem pojęciowym. Definiowanie pojęć z przestrzeni rozważań, będącej częścią świata realnego, odbywa się w modelu pojęciowym. Wspomnianych pojęć dostarcza formalizm pojęciowy, który

można traktować jako zbiór pojęć z zakresu modelowania, stosowanych w opisie modelu pojęciowego. Jeden konkretny formalizm pojęciowy może być zastosowany w różnych językach schematów pojęciowych. Wyróżnia się dwa języki schematów pojęciowych: leksykalny oraz graficzny. Oba języki są językami formalnymi korzystającymi z formalizmu pojęciowego i przeznaczone są do przedstawiania schematów pojęciowych.



Rys. 1. Od świata rzeczywistego do schematu pojęciowego

Źródło: [6]

Głównym zadaniem w procesie budowania pojęciowego modelu danych jest precyzyjna definicja obiektów zainteresowania i zidentyfikowanie związków pomiędzy nimi. Jako przykłady obiektów można podać: ulice, działki, właścicieli i budynki. Przykładami związków pomiędzy nimi mogą być „znajdujący się na”, „będący własnością”, „jest częścią”.

### 3.2. Modelowanie logiczne

Logiczne projektowanie bazy danych polega na przekształceniu modelu pojęciowego w model logiczny z uwzględnieniem modelu danych w docelowej bazie danych (np. model relacyjny). Na tym etapie określany jest sposób zapisu atrybutów opisowych i właściwości przestrzennych obiektów oraz relacji między nimi.

Logiczny model danych jest abstrakcją obiektów występujących w konkretnym zastosowaniu. Ta abstrakcja jest zamieniana na elementy bazy danych. Obiekt reprezentuje takie encje jak budynek, działka czy właściciel i zapisywany jest jako wiersz. Obiekt posiada zestaw atrybutów. Atrybuty charakteryzują cechy obiektu, takie jak jego nazwa, miara, klasyfikacja lub identyfikator (klucz) do innego obiektu. Atrybuty są zapisywane w bazie danych w kolumnach (zwanymi także polami). Klasa jest zbiorem podobnych obiektów. Każdy obiekt w klasie ma ten sam zestaw atrybutów. Klasa jest zapisywana w bazie danych jako tabela.

W przeszłości logiczne modele danych były często rysowane w postaci tak zwanych diagramów encja – relacja. Ostatnio jednak największą popularność zdobył zuniifikowany język modelowania – UML (*Unified Modeling Language*), który jest standardowym zapisem do wyrażania modeli obiektowych, popieranym przez największych producentów oprogramowania i baz danych.

### 3.3. Fizyczna implementacja modelu bazy danych

Na podstawie logicznego modelu budowany jest fizyczny model bazy danych. Najczęściej odbywało się to w taki sposób, że specjalista od relacyjnych baz danych otrzymywał logiczny model danych i używał narzędzi administracyjnych bazy danych do zdefiniowania schematu bazy danych i utworzenia nowych baz danych gotowych do transferu i wprowadzania danych.

Wprowadzenie modelu geobazy pozwoliło na zastosowanie metod projektowania relacyjnych baz danych również do informacji geograficznej. Obecnie najpopularniejszym sposobem tworzenia baz danych jest użycie narzędzi CASE (*Computer Aided Software Engineering* – komputerowo wspomaganą inżynierię oprogramowania), które umożliwiają budowanie modeli baz danych. Logiczny model bazy danych zapisany z wykorzystaniem języka UML może być zastosowany do automatycznego wygenerowania schematu bazy danych zgodnego z założoną specyfikacją.

## 4. *Unified Modeling Language*

UML (*Unified Modeling Language*) jest to graficzny język modelowania pozwalający na obrazowanie, opisywanie, specyfikowanie, analizowanie oraz dokumentowanie świata realnego w ujęciu obiektowym. Ideą UML jest umożliwienie zastosowania prostych środków do utworzenia graficznego modelu prezentującego dowolny fragment rzeczywistości. Dzięki czemu stała się możliwa komunikacja w jednym języku pomiędzy klientami, analitykami, projektantami oraz programistami. Upraszczając, można powiedzieć, iż UML jest to zbiór czytelnych symboli i oznaczeń, których zrozumienie nie jest problemem nawet dla osób z minimalną wiedzą z zakresu informatyki. UML został przyjęty jako formalny język zapisu modeli pojęciowych w serii norm ISO 19100 poświęconych informacji geograficznej [6].

UML ma wiele zastosowań, przede wszystkim wykorzystywany jest do opisu systemów informatycznych oraz podczas realizacji systemów biznesowych w branżach typu: produkcja, bankowość itp., jak również podczas projektowania bazy danych. Utworzony w ten sposób schemat może być dowolnie modyfikowany i udoskonalany.

Modelowanie pojęciowe z użyciem UML wymaga znajomości kilku podstawowych definicji z zakresu projektowania obiektowego [7]:

- klasa** – byt semantyczny rozumiany jako miejsce przechowywania takich cech grupy podobnych obiektów, które są dla nich niezmiennie (np. zestawu atrybutów, nazwy, metod);
- dziedziczenie** – związek pomiędzy klasami obiektów określający przekazywanie cech (definicji atrybutów, metod) z nadklasy do jej podklas, jest podstawowym mechanizmem sprzyjającym ponownemu użyciu;
- agregacja** – związek pomiędzy klasami obiektów, modelujący stosunek całości do jej części (np. stosunek działki do budynku); obiekty są powiązane związkiem agregacji, jeżeli jeden z nich można uważać za część drugiego.

## 5. Model bazy danych na potrzeby wyceny nieruchomości

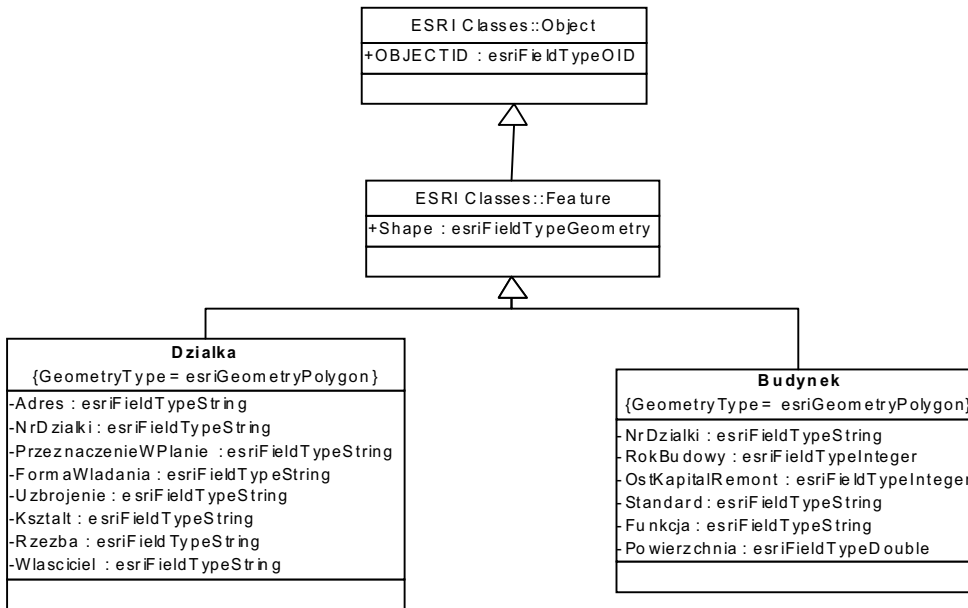
W poniższym przykładzie przedstawiono elementarny model danych katastralnych. Związany on jest z pracami autorów nad wykorzystaniem systemów informacji geograficznej do wspomagania wyceny nieruchomości [1]. Projektowanie modelu bazy danych rozpoczęto od zdefiniowania dwu podstawowych klas obiektów: działek i budynków. Po analizie możliwych atrybutów każdej z klas, wybrano najistotniejsze dla procesu wyceny. Dla działek są to: adres, numer działki, przeznaczenie w miejscowym planie zagospodarowania przestrzennego, właściciel, stan prawny oraz forma władania, zaś dla budynków to numer działki, rok budowy, ostatni kapitalny remont, standard wykończenia, funkcja oraz powierzchnia. Budynek jest w „trwałym związku” z działką, każde istnienie budynku wymaga istnienia działki, dlatego w projektowanym modelu należy to uwzględnić.

## 6. Zastosowanie narzędzi CASE w modelowaniu logicznym

Do zapisania schematu UML posłużono się oprogramowaniem Microsoft Office Visio 2003, gotowy schemat stał się podstawą do automatycznego wygenerowania struktury bazy danych w systemie ArcGIS firmy ESRI [4].

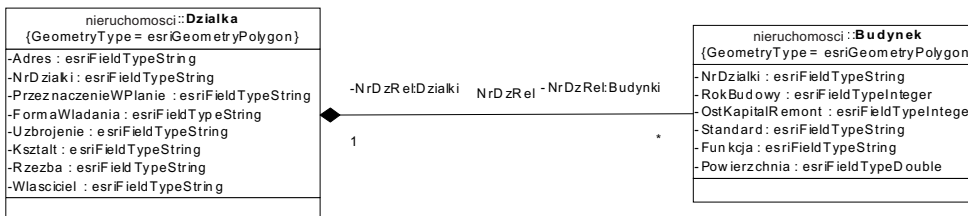
Tworzenie schematu odbywa się poprzez umieszczanie na stronie rysunku kształtów z gotowych wzorników. Visio dostarcza ośmiu wzorników, liczba ta jest ściśle związana ze specyfikacją języka UML, w którym definiuje się osiem podstawowych diagramów, m.in. diagram aktywności, diagram przypadków użycia, diagram wdrożenia.

W diagramie struktury statycznej, który przeznaczony jest do projektowania baz danych, zdefiniowano dwie klasy obiektów (rys. 2): **Działka** i **Budynek**. Atrybuty mają określone typy zmiennych. Z modelu danych ArcInfo udostępnionego przez **ESRI**, klasy obiektów **Działka** oraz **Budynek** dziedziczą następujące atrybuty: **ObjectID** – indywidualny identyfikator dla każdego obiektu oraz **Shape** – definiujący geometrię (w obu klasach jest to poligon).

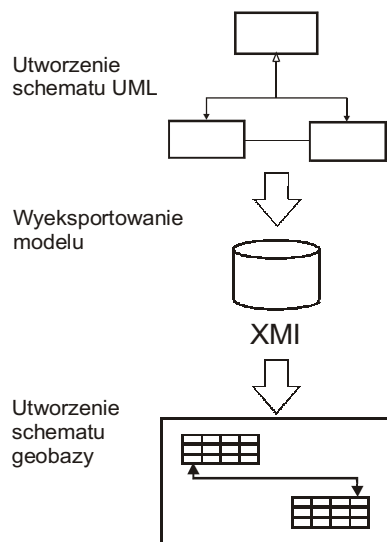


Rys. 2. Schemat UML – klasy obiektów i ich atrybuty

Dodatkowo poprzez agregację zdefiniowano relację pomiędzy klasami **Działka** a **Budynek** (rys. 3), w której klasą nadrzędną jest **Działka**. Liczebność relacji ustalono na jeden do wielu, co oznacza, iż na jednej działce może stać wiele budynków. Dla uproszczenia pominięto przypadek szczególny tej relacji, a mianowicie sytuację, gdy granica pomiędzy działkami przebiega pod budynkiem.



Rys. 3. Relacje pomiędzy klasami w geobazie są reprezentowane poprzez agregację w schemacie UML



Rys. 4. Schemat przejścia z modelu UML do geobazy w ArcCatalog  
Źródło: [5]

Gotowy schemat bazy danych (rys. 4) za pomocą załączonego makra ESRI XMI Export wyeksportowano do pliku XML [3]. Poprawność utworzonego pliku XML sprawdzono za pomocą makra Schematic Checker. Ostatni etap to import poprawnego pliku XML za pomocą Schema Wizard w aplikacji ArcCatalog do geobazy. Proces ten dokumentowany jest poprzez automatycznie generowany raport. W wyniku powstała pusta geobaza, o strukturze zgodnej z projektem, gotowa do wypełnienia danymi.

## 7. Wnioski

Od pewnego czasu jesteśmy świadkami ciągłego rozwoju technologii w branży informatycznej, z którą systemy informacji geograficznej są w dużej mierze związane. Czym bowiem byłyby SIG bez komputera, bazy danych i odpowiedniego systemu jej obsługi. Dlatego też możliwość wprowadzenia nowoczesnych technologii informatycznych już na etapie projektowania przestrzennych baz danych daje wiele korzyści. Przede wszystkim tworzenie baz danych z pomocą UML jest szybsze, co oznacza, że i mniej kosztowne. Tworzone schematy są czytelne, przejrzyste, a ponadto istnieje możliwość ich wielokrotnego użycia oraz modyfikacji. Zaś sam model geobazy pozwala na efektywne korzystanie z danych w niej zawartych.

## Literatura

- [1] Cichociński P., Dębińska E., Parzych P.: *Zastosowanie systemów informacji geograficznej do wspomagania wyceny nieruchomości*. XIX Jesienna Szkoła Geodezji „Geoinformacja dla wszystkich”, Politechnika Wroclawska, Polanica-Zdrój 2005



- [2] Connolly T., Belg C.: *Systemy baz danych. Praktyczne metody projektowania, implementacji i zarządzania – tom 1*. Warszawa, Wydawnictwo RM 2004
- [3] Gajc B.: *UML w akcji*. Geodeta, nr 10 (113) 2004, Warszawa, 2004
- [4] Perencsik A., Idolyantes E., Booth B., Andrade J.: *ArcGIS 9. Designing Geodatabase With Visio*. Redlands, ESRI Press 2004
- [5] Perencsik A., Idolyantes E., Booth B., Andrade J.: *ArcGIS 9. Introduction to CASE Tools*. Redlands, ESRI Press 2004
- [6] PN-EN ISO 19101: *Informacja geograficzna – Model tworzenia norm (Geographic information – Reference model)*. 2002
- [7] Subieta K.: *Słownik terminów z zakresu obiektowości*. Warszawa, Akademicka Oficyna Wydawnicza PLJ 1999
- [8] Zeiler M.: *Modeling our World. The ESRI Guide to Geodatabase Design*. Redlands, ESRI Press 1999