

## **MODELOWANIE DANYCH W PROCESIE ODWZOROWYWANIA INFORMATYCZNEGO SYSTEMÓW EMPIRYCZNYCH STANOWIĄCYCH PRZEDMIOT INŻYNIERII ROLNICZEJ**

Wojciech Mueller, Piotr Boniecki, Jerzy Weres, Krzysztof Nowakowski  
*Instytut Inżynierii Rolniczej, Akademia Rolnicza w Poznaniu*

**Streszczenie.** Systemy empiryczne, leżące w obszarze zainteresowania inżynierii rolniczej, charakteryzują się wyjątkową złożonością. Próbą radzenia sobie z sygnalizowaną komplikacją w procesie poznawania jest wielopoziomowe modelowanie dziedziny problemowej. Istotną kwestią w procesie modelowania dziedziny problemowej, jak i projektowania systemów informatycznych, które są wykorzystywane do badania systemów empirycznych jest prawidłowe zrealizowanie fazy modelowania danych. Zaistnienie tego etapu jest konsekwencją podejmowanych wysiłków zmierzających do poznania coraz bardziej złożonych systemów empirycznych, opisywanych coraz większą porcją informacji wzajemnie ze sobą powiązanych.

**Słowa kluczowe:** modelowane danych, bazy danych

### **Wprowadzenie**

Złożoność systemów empirycznych, w tym również systemów tworzonych przez człowieka, wymaga w procesie ich dalszego pełniejszego poznawania, jak i doskonalenia istotnego wsparcia ze strony specjalistycznych narzędzi informatycznych, jakimi są różnorodne aplikacje. Operują i dostarczają one nam coraz większej ilości informacji o zróżnicowanym charakterze, które generalnie są gromadzone w systemach zarządzania bazami danych (SZBD). Przyjęcie takiego typu rozwiązania przy tworzeniu systemu informatycznego, który ma charakter niepisanego standardu, narzuca już częściowo zarys przyszłej struktury aplikacji, co w połączeniu z złożonością dziedziny problemowej z reguły wymusza przeprowadzenie wielopłaszczyznowego modelowania. Efektem końcowym podjętego działania jest poprawne odwzorowanie interesującego nas systemu empirycznego w formach informatycznych. Proces ten przebiega wieloetapowo i towarzyszy mu ciągła transformacja uzyskiwanych modeli, aż do otrzymania właściwych struktur informatycznych. W tym obszarze działania mieści się również modelowanie danych, których istnienie jest trwałe i, które stanowią jednocześnie przedmiot przetwarzania powstających systemów informatycznych. Proces ten, z uwagi na złożoność danych i związków pomiędzy nimi, przebiega także wieloetapowo i zawiera sobie właściwe fazy [Allen 2006]:

- Modelowanie koncepcyjne,
- Modelowanie logiczne,
- Modelowanie fizyczne (implementacyjne).

Realizacja poszczególnych etapów wymaga wykorzystania specyficznych języków na każdym poziomie modelowania. Aktualnie dysponujemy szerokim wachlarzem języków, do których zaliczamy głównie języki bazujące na grafach. Ostatnio pojawiły się propozycje wykorzystania w procesie modelowania, na wszystkich jego etapach, języka XML [Graves 2002]. Obecnie standardem w projektowaniu systemów informatycznych jest język obiektowy UML 2.0, który zgodnie z założeniami jego twórców ma służyć modelowaniu dziedziny problemowej w różnych jej aspektach. Zatem obszar jego wykorzystania jest zdecydowanie szerszy, wykraczający poza ramy modelowania danych. Powoduje to, iż w pewnych przypadkach, to narzędzie modelowania nie jest zbyt efektywne. Przesądzają o tym kolejne etapy procesu modelowania danych, gdzie zmuszeni jesteśmy podjąć decyzję o zastosowaniu architektonicznego modelu danych. Przebieg modelowania danych, polegający na realizacji szeregu transformacji uzyskiwanych modeli, charakteryzują się zróżnicowanym poziomem komplikacji. Złożoność tych operacji jest konsekwencją przyjętego języka modelowania, jak i architektonicznego modelu danych, a ten wybór jest nierozdzielnie związany z aktualnie dostępnymi na rynku informatycznym SZBD.

Istnienie standardu modelowania UML 2.0 przy równoległym występowaniu innych języków, które nie wprost związane są z określonym architektonicznym modelem danych powoduje, iż obserwuje się znaczną dowolność w stosowaniu narzędzi w procesie projektowania danych. Oznacza, to również, iż często decyzje implementacyjne o wykorzystanym SZBD zapadają już na początku etapu modelowania koncepcyjnego, co z reguły przesądza o wykorzystanych narzędziach. Tak wczesne decyzje implementacyjne są sprzeczne z zasadami modelowania.

### **Modelowanie danych**

Modelowanie koncepcyjne, dostarczające nam charakterystykę dziedziny problemowej z perspektywy opisujących ją danych, traktujemy z reguły, jako zbiór zidentyfikowanych wymagań zaprezentowanych w postaci sformalizowanego zapisu. Użyty w tym pierwszym etapie projektowania danych język powinien być elastyczny, posiadający odpowiednią notację graficzną, a zarazem powinien stanowić kompromis pomiędzy łatwością jego zrozumienia przez przyszłych użytkowników, a prostotą dalszej transformacji uzyskanych modeli w trakcie kolejnej fazy modelowania danych. Oczywiście lista sygnalizowanych postulatów pod adresem tego języka jest zdecydowanie dłuższa.

Aktualnie projektant na tym poziomie modelowania ma do dyspozycji następujące notacje graficzne:

- język UML w wersji 2.0,
- diagramy związków encji i rozszerzone diagramy związków encji,
- grafy.

Efekty jego pracy, czyli zidentyfikowane w trakcie tego etapu dane i związki między nimi są odwzorowywane w postaci właściwych struktur graficznych. Ich forma zależy od przyjętego przez projektanta języka modelowania.

W przypadku, gdy obszar danych jest wyjątkowo szeroki, obejmujący oddzielne zakresy tematyczne, wzbogacamy modelowanie koncepcyjne o dwa podprocesy obejmujące modelowanie poszczególnych schematów i ich późniejszą integrację.

Modelowanie koncepcyjne, przebiegające z wykorzystaniem języka UML 2.0, owocuje z reguły diagramami klas. W przypadku dużej złożoności dziedziny problemowej zostaje ono wzbogacone o diagramy pakietów, które wraz z poprzednimi stanowią diagramy struktury. Podstawowe pojęcia, wykorzystywane na tym etapie modelowania, posiadające właściwą sobie notację graficzną, to [Wrycza i in. 2005]:

- pakiety, klasy, obiekty
- związki – asocjacje (agregacje), uogólnienia, zależności.

Różną perspektywę szczegółowości modelowanych danych, opisujących rozpatrywaną dziedzinę problemową, osiągamy tworząc najpierw diagramy poziomu konceptualnego, a w dalszej kolejności diagramy poziomu implementacyjnego [Muller 2000].

Uzyskany model lub modele stanowią punkt wejściowy kolejnej fazy projektowania, czyli modelowania logicznego. Przebieg tego procesu w skład, którego wchodzi transformacja wcześniejszego modelu jest zdeterminowany przyjętym na tym etapie prac modelem architektonicznym danych, który zostanie wykorzystany. Model ten to nic innego jak zbiór zasad określający sposób posługiwania się danymi. Reguły te w zależności od obszaru, którego dotyczą możemy podzielić na trzy grupy [Beynon-Davies 2003]:

- definiowania danych,
- operowania danymi,
- zapewniające integralność danych.

Aktualnie projektant może wybierać następujące modele danych:

- relacyjny,
- obiektowy,
- postrelacyjny.

Aczkolwiek mamy do dyspozycji jeszcze inne architektoniczne modele danych, to jednak z punktu widzenia użytecznego nie znajdują one większego zastosowania w aktualnych SZBD.

Proces transformacji zmierzający do uzyskania modelu logicznego, zgodnego ze zbiorem reguł definiowania danych wynikających z dokonanego wyboru architektonicznego modelu danych jest nie tylko konsekwencją tego wyboru, ale również wynika z zastosowanego języka w kroku poprzedzającym tę fazę modelowania. W przypadku utworzenia modelu koncepcyjnego w notacji UML 2.0 i następnie podjęciu decyzji o bazowaniu na relacyjnym architektonicznym modelu danych, przekształcanie jednego modelu w drugi przebiega z reguły według poniższego algorytmu [Muller 2000]:

- dla zidentyfikowanych pakietów tworzone są schematy czyli tzw. przestrzenie nazw,
- dla każdej klasy tworzona jest tabela, jednak musi to być poprzedzone zdefiniowaniem odpowiednich typów dla kolumn (czyli domen),
  - atrybuty klas odwzorowywane są w postaci kolumn,
  - operacje reprezentujące dynamikę klasy podlegają procesowi partycjonowania, co oznacza, że część funkcji przejmuje aplikacja, a część SZBD,
- zależności pomiędzy klasami są wbudowywane w tabele poprzez wykorzystanie kolumny klucza podstawowego i tworzonej na te potrzeby kolumny klucza obcego.

Ten proces w rzeczywistości ma jednak bardziej złożony charakter, wynikający ze specyfiki realizowanego systemu informatycznego i niejednokrotnie obejmuje swym zasię-

giem kolejny etap modelowania, który określamy mianem modelowania fizycznego lub implementacyjnego. Przystępujemy do niego po podjęciu kolejnej decyzji, która określa precyzyjnie użyte narzędzie jakim jest SZBD. Tworzony w tym kroku model jest budowany przy użyciu języka SQL, właściwego wybranemu SZBD, ale z wykorzystaniem tych elementów języka, które są odpowiedzialne za definiowanie struktur.

### **Systemy informatyczne wspomagające badania i nauczanie**

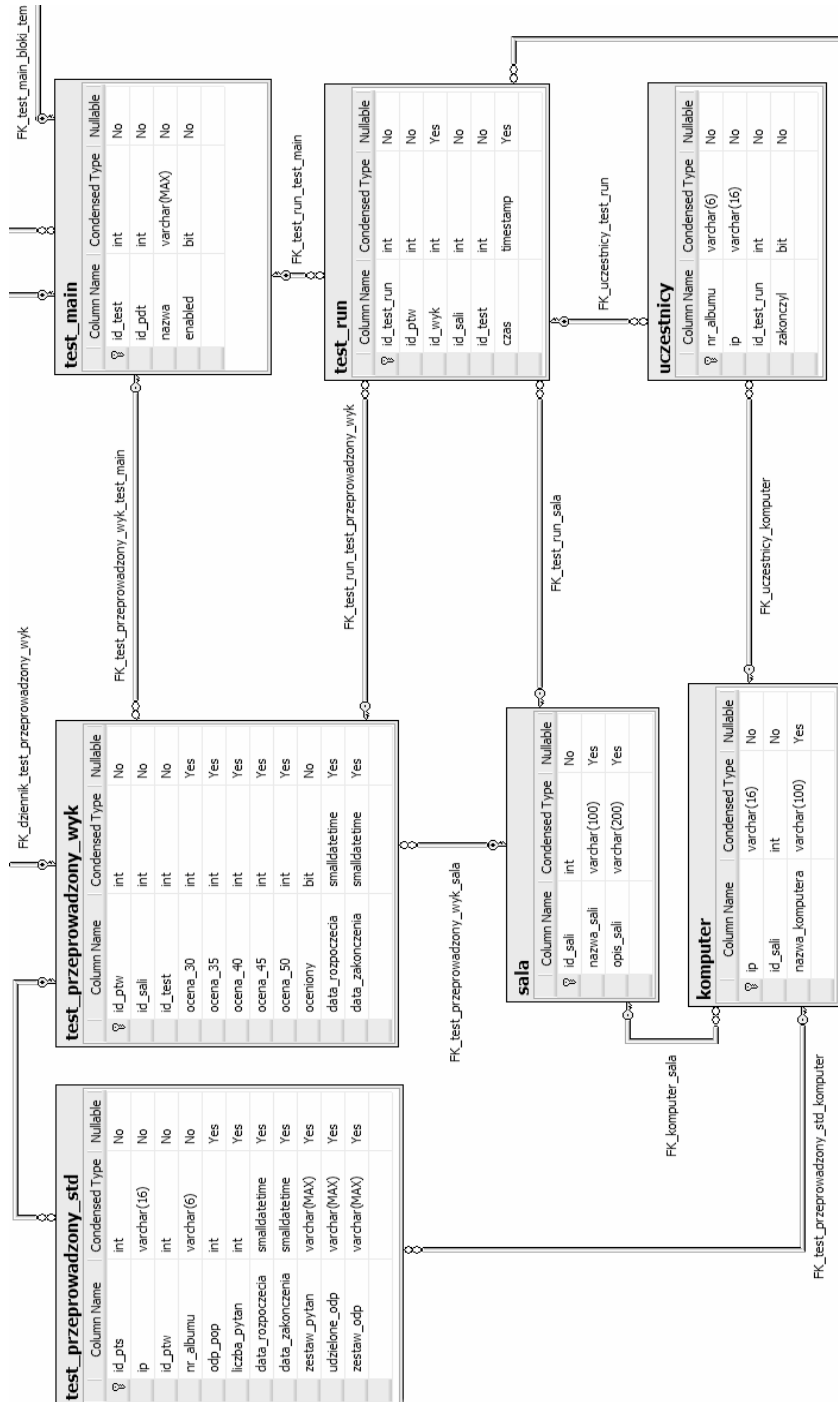
Powstające, jak i już wytworzone systemy informatyczne wspierające działalność Instytutu Inżynierii Rolniczej zarówno w procesie kształcenia, jak i prowadzenia oraz dokumentowania badań mają charakter internetowy lub intranetowy. Istotnym elementem tych aplikacji jest SZBD, na którym osadzono zaprojektowane bazy danych. Charakter deklaratywny etapów projektowania danych, faz modelowania i ich notacji pozostawia projektantowi dużą swobodę, która z jednej strony może przyczynić się do szybszego powstania systemu informatycznego, a z drugiej utrudniać późniejszą migrację powstałego produktu na inną platformę bazodanową.

Jedną z bardziej złożonych aplikacji, już funkcjonujących i wspomagających proces kształcenia, jest EDUIR sukcesywnie doskonalony przez autorów. Wysoki stopień komplikacji danych, niezbędność tego systemu wraz z wcześniej podjętą decyzją o wykorzystaniu relacyjnego SZBD, jakim jest SQL Server 2005, z uwagi na oferowaną przez niego obiektość, spowodowały modyfikacje w procesie modelowania danych. Wykorzystując w procesie modelowania środowisko SQL Sever Managment Studio scalono te trzy fazy w jedną całość, a uzyskany diagram związków encji, za którymi kryją się określone tabele powiązane ze sobą przedstawiono częściowo na rys. 1.

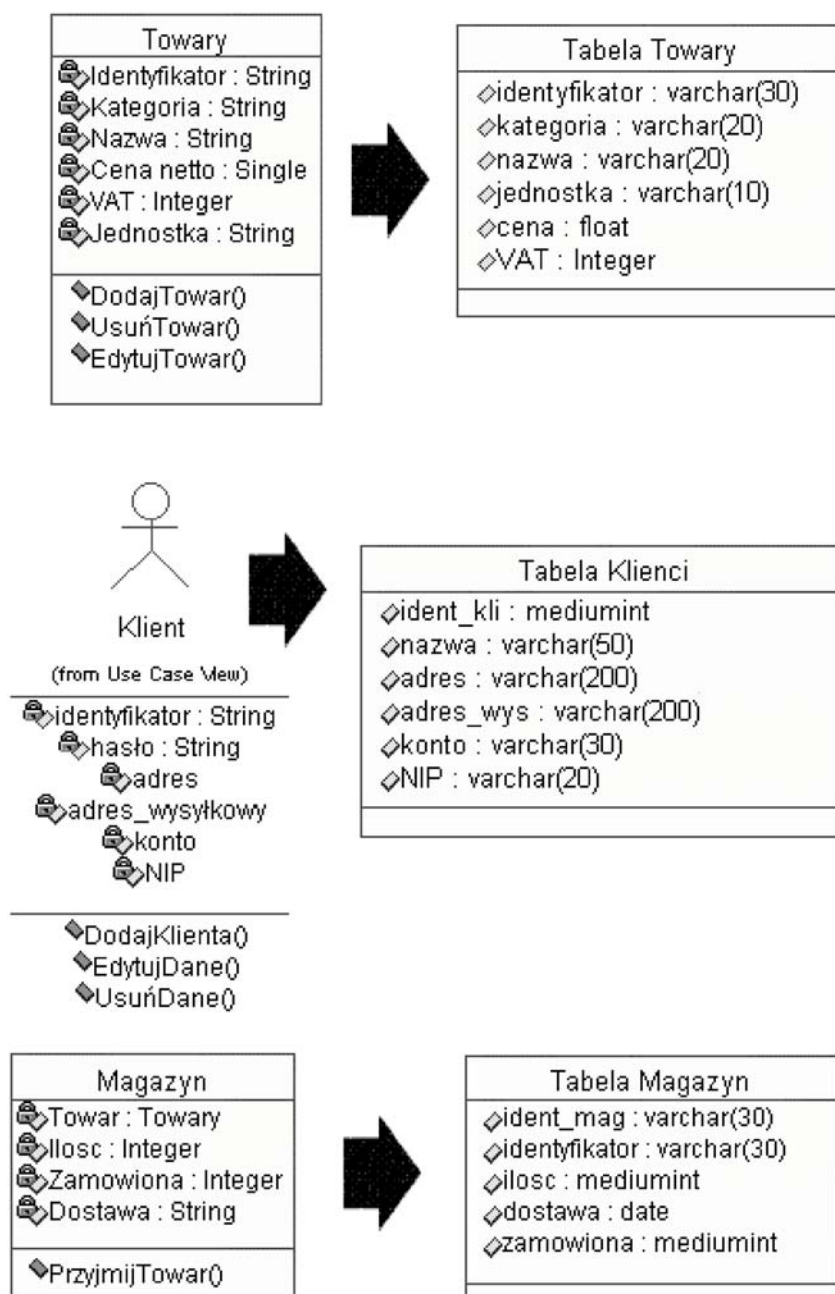
Generalnie większość wspomnianych wyżej systemów informatycznych, jak i aplikacji wspomagających sektor rolniczy, przechowujących dane w SZBD, była projektowana z zachowaniem sygnalizowanych trzech faz modelowania danych. Z uwagi na uniwersalny charakter notacji UML 2.0, język ten stanowił główne narzędzie modelowania, nie tylko aplikacji ale również modelowania koncepcyjnego danych. Zatem kolejna faza modelowania zawierała między innymi transformacje otrzymanych wcześniej modeli, albowiem powszechnie dokonywanym wyborem architektonicznego modelu danych, który musiał być dokonany na tym etapie przez projektanta, był model relacyjny. Efekt tak realizowanego procesu przekształcania modelu koncepcyjnego w logiczny ilustruje rys. 2. W wyniku tych operacji klasy wraz z atrybutami zostały przekształcone w tabele relacyjne.

Oczywiście te działania nie kończyły fazy logicznego modelowania danych, gdyż do rozwiązania pozostawał problem partycjonowania rozpoznanych operacji.

Uporanie się z tym zagadnieniem pozwalało przejść do ostatniego etapu modelowania implementacyjnego, czego efektem było utworzenie struktur gotowych na przejście danych w wybranym SZBD.



Rys. 1. Fragment diagramów związków encji związanych z testowaniem wiedzy systemu informatycznego EDUIR  
Fig. 1. Entity relationship diagram for the EDUIR information system



Rys. 2. Przekształcanie klas w relacje  
Fig. 2. Transformation of classes into relations

## Podsumowanie i wnioski

Podjęte przez autorów prace związane z wytwarzaniem różnorodnych systemów informatycznych opartych na SZBD, wspomagających funkcjonowanie instytutu oraz sektora rolniczego, nasunęły im szereg uwag i spostrzeżeń, które zaprezentowano poniżej:

1. Modelowanie obiektowe UML, w stosunku do innych metod stosowanych na poziomie tworzenia modelu koncepcyjnego baz danych, pełniej odwzorowuje dane i ich dynamikę. Dostarcza ponadto pojęć, wraz z odpowiadającą im notacją graficzną, umożliwiając stosowanie zstępującej reguły modelowania od ogółu do szczegółu.
2. Wybór metody modelowania koncepcyjnego i logicznego baz danych uzależniony jest od jej złożoności, lokalizacji struktur odpowiedzialnych za przetwarzanie danych oraz posiadanych informacji o zastosowanym SZBD.
3. Hybrydowy charakter najnowszych systemów bazodanowych, udostępniających równocześnie różnorodne architektoniczne modele danych, komplikuje modelowanie danych, wprowadzając kolejne fazy transformacji modeli. Jednak ta proponowana różnorodność, która z reguły ma charakter dualizmu, pozwala na dokładniejsze odwzorowanie dziedziny problemowej.

## Bibliografia

- Allen S.** 2006. Modelowanie danych. Helion, Gliwice. ISBN 83-246-0186-8
- Beynon-Davies P.** 2003. Systemy baz danych. WNT, Warszawa. ISBN 83-204-2993-5
- Graves M.** 2002. Projektowanie baz danych XML. Helion, Gliwice. ISBN 83-7197-667-4.
- Muller R.J.** 2000. Bazy danych język UML w modelowaniu danych. MOKOM, Warszawa. ISBN 83-7279-000-0.
- Wrycza S., Marcinkowski B., Wyrzykowski K.** 2005. Język UML 2.0 w modelowaniu systemów informatycznych. Helion, Gliwice. ISBN 83-7361-892-9.

## **DATA MODELING IN A PROCESS OF MAPPING EMPIRICAL SYSTEMS OF AGRICULTURAL ENGINEERING ONTO INFORMATION SYSTEMS**

**Summary.** Empirical systems investigated within agricultural engineering are extremely complex. A multi-level modeling of a problem domain, extended by mapping developed operational structures onto information systems is a solution to deal with the complexity. A crucial issue, becoming more and more pronounced in modeling a problem domain and designing information systems oriented at investigation of empirical systems, is to complete correctly data modeling phase. Such analysis is a result of substantial attempts made to better understand the empirical systems, the systems which are more and more complex, requiring more and more interrelated information. The attempts have been favored by appearance of new information technologies dedicated to representation of data.

**Key words:** data modeling, databases

**Adres do korespondencji:**

Wojciech Mueller; e-mail: muellerw@au.poznan.pl  
Instytut Inżynierii Rolniczej w Poznaniu  
ul. Wojska Polskiego 50  
60-627 Poznań