



**POLSKIE
TOWARZYSTWO
INFORMACJI
PRZESTRZENNEJ**

ROCZNIKI 2012 **GEOMATYKI**

**Modele danych przestrzennych w UML
i ich transformacja do schematów GML
i struktur baz danych**

**Tom X
Zeszyt 1(51)
Warszawa**

Janusz Michalak

11. Schematy aplikacyjne tematów aneksów II i III dyrektywy INSPIRE

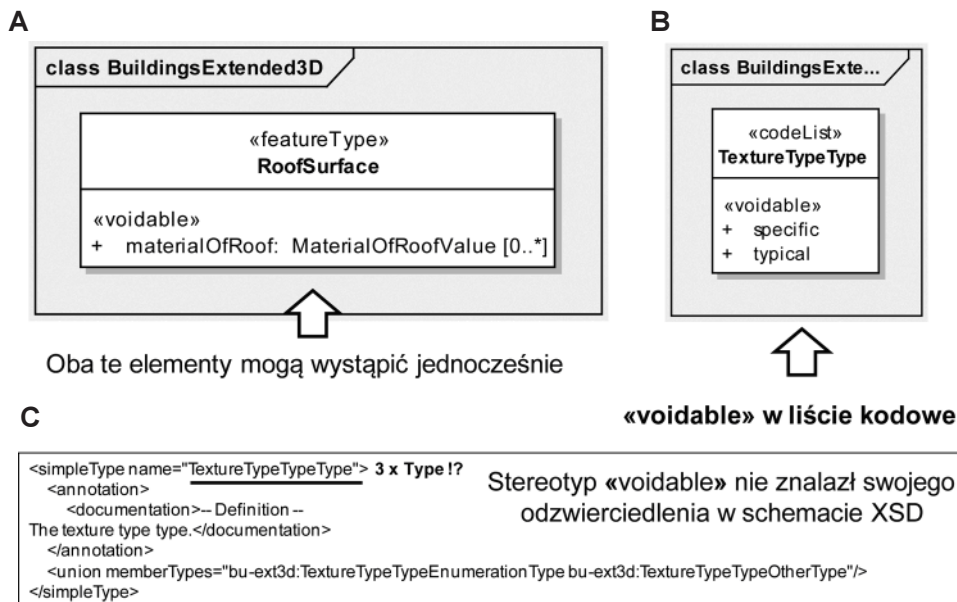
Okres opracowywania specyfikacji danych dla aneksów II i III obejmuje czas od listopada 2010 do maja 2012. W tym okresie w 19 tematycznych zespołach roboczych TWG (*Thematic Working Group*) prowadzono prace dotyczące 24 specyfikacji danych z zakresu 25 tematów (dane dwóch tematów dotyczących atmosfery są ujęte w jednej specyfikacji). W czerwcu 2011 zostały udostępnione publiczne wstępne robocze wersje 2.0 tych specyfikacji. Do października tego roku w poszczególnych krajach członkowskich Unii Europejskiej zespoły stron zainteresowanych i zarejestrowanych w INSPIRE jako LMO (*Legally Mandated Organisation*) lub SDIC (*Spatial Data Interest Community*) prowadziły prace z zakresu testowania i oceny zawartych w specyfikacjach modeli danych w języku UML i utworzonych na ich podstawie schematów XSD. Wyniki prac testowych wykazały potrzebę istotnych zmian. W konsekwencji okazał się potrzebny jeszcze jeden dodatkowy etap prac w zespołach tematycznych nad przejściowymi wersjami oznaczonymi numerem wersji 2.9 (Michalak, 2012). Te przejściowe wersje także zostały poddane ocenie, jednak już w znacznie ograniczonym niepublicznym zakresie. Termin zakończenia prac i opublikowania końcowych wersji 3.0 to maj 2012.

Proces prac nad tematami aneksów II i III był bardziej złożony niż nad tematami aneksu I, a rezultat nie jest w pełni do końca zadawalający. Złożyły się na to trzy przyczyny:

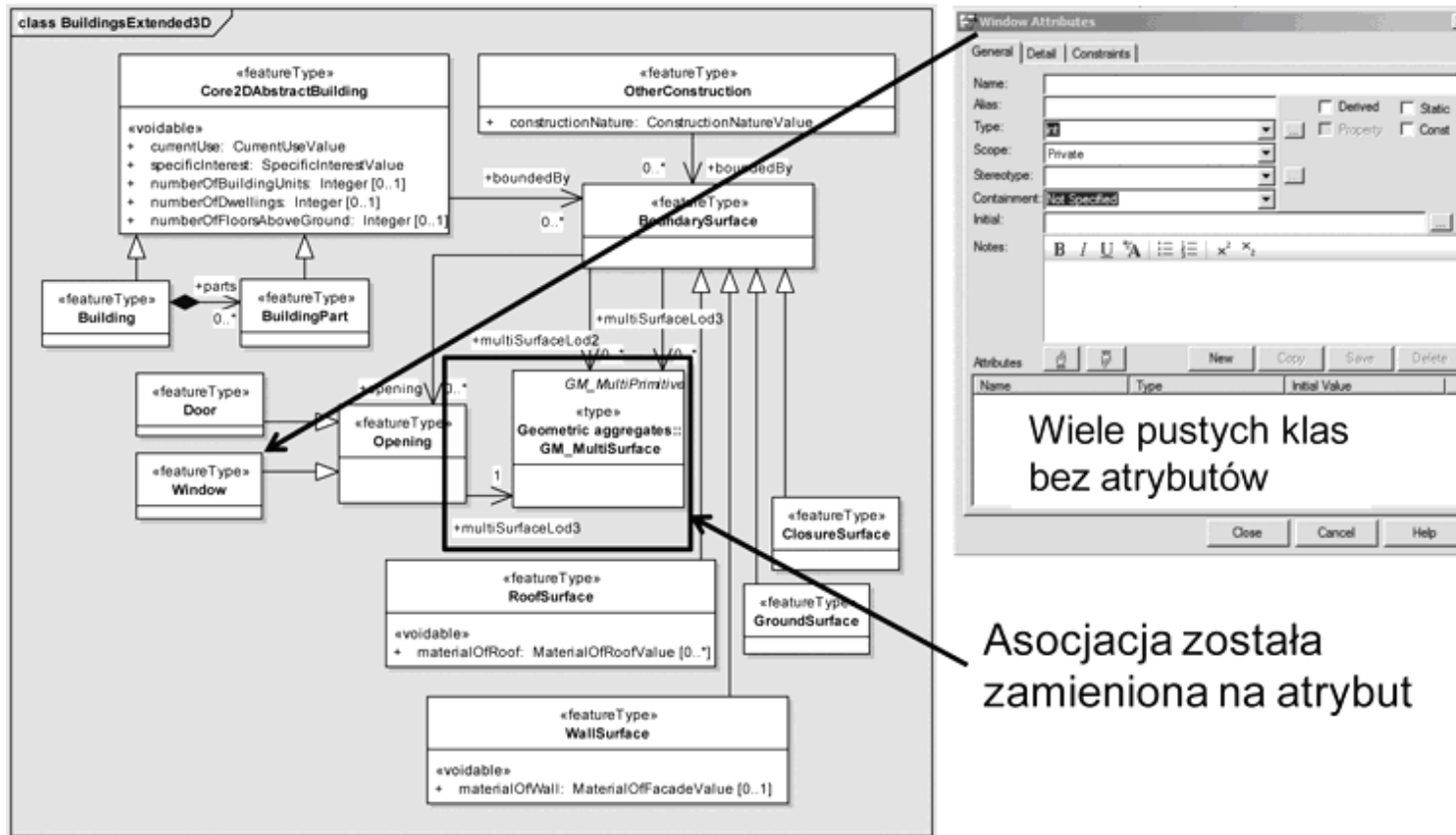
1. Problematyka tych tematów jest bardziej skomplikowana, ponieważ dotyczy złożonych zjawisk przyrodniczych i społecznych. Zjawiska te są bardzo często dynamiczne i silnie wzajemnie powiązane – nie tylko w obrębie poszczególnych tematów, ale także pomiędzy różnymi tematami. Dane w wielu przypadkach nie są prostymi rejestrami zbiorów jednoznacznie zdefiniowanych obiektów, jak to było najczęściej w specyfikacjach danych dla tematów I aneksu. Często dotyczą przestrzennych lub czasowych tendencji zmian określonych właściwości środowiska lub zachodzących w nim procesów. Wymownym przykładem była konieczność zastosowania danych typu pokrycie (*coverage*), czego nie było w tematach aneksu I.
2. W przeciwieństwie do tematów aneksu I, w tych tematach nie było do tego czasu takich modeli danych lub były to modele w bardzo wczesnym etapie opracowania – były jeszcze dalekie od stanu dojrzałości. W wielu przypadkach poszczególne tematyczne zespoły robocze (TWG) musiały rozpocząć prace nad modelami od przysłowiowego „zera”, gdy dla tematów aneksu I często problem sprowadzał się do wyboru już istniejących dojrzałych modeli opracowanych w poszczególnych krajach członkowskich i połączenia ich w jeden spójny model europejski.

3. Uczestnicy prac nad specyfikacjami dla tematów aneksów II i III mieli znacznie mniejsze doświadczenie i mniejszą wiedzę z zakresu opracowywania modeli danych w języku UML i ich transformacji do schematów XSD. Dotyczy to zarówno poszczególnych grup roboczych (TWD), jak i zespołów testujących i opiniujących w krajach członkowskich (SDIC i LMO). Często był to ich pierwszy kontakt z tą metodyką i z tymi technologiami. Jak już wcześniej wspomniano wiele wskazuje na to, że ostateczne rezultaty (wersje 3.0) nie są w pełni zadawalające. W dalszej części tego rozdziału przedstawione będzie jedynie kilka wybranych przykładów nie do końca poprawnych rozwiązań zawartych w wersjach 2.0 i nie wszystkie z nich zostały poprawione w wersjach 2.9.

Pierwszy przykład pochodzi z modelu danych tematu **Budynki (Buildings)** i dotyczy stosowania stereotypu **«voidable»** z licznnością zawierającą **[0]**, co jest opisane w rozdziale 9.2 (rys. 9.6). Zastosowano ten stereotyp do atrybutu klasy `RoofSurface` z jednoczesnym podaniem licznności `[0..*]`. Jest to formalnie poprawne i można przypuszczać, że gdy jest podana wartość atrybutu (stereotyp nie został wykorzystany do pominięcia wartości atrybutu) to w przypadku licznności zerowej ten budynek nie ma dachu, co jest możliwe i w konsekwencji takie dane mogą być poprawne (rys. 11.1A). W drugim jednak przypadku zastosowano ten stereotyp do elementów listy kodowej `TextureTypeType` (rys. 11.1B). Trudno tu jest zrozumieć, jaki sens ma możliwość pomijania poszczególnych pozycji listy kodowej. Lista kodowa podobnie do typu wyliczeniowego (klasy ze stereotypem **«enumeration»**) jest wyjątkowym typem klasy – jest to klasa jednej instancji (jednego obiektu) zawierającego listę elementów, które nie są oddzielnymi atrybutami, lecz pozycjami, z których można wybrać tylko jedną z nich. W konsekwencji wszystkie pozostałe pozycje listy są pomijane i nie są potrzebne im stereotypy **«voidable»**, a nawet nie mogą być uwzględnione, bo narzuca to obowiązek obecności wszystkich pozycji listy z podaniem przyczyny, dlaczego są puste (rys. 9.6).



Rys. 11.1. Przykłady wątpliwego (A) i niepoprawnego (B) użycia stereotypu **«voidable»**, a także niezbyt fortunnnej nazwy klasy (C) w modelu danych tematu Budynki.



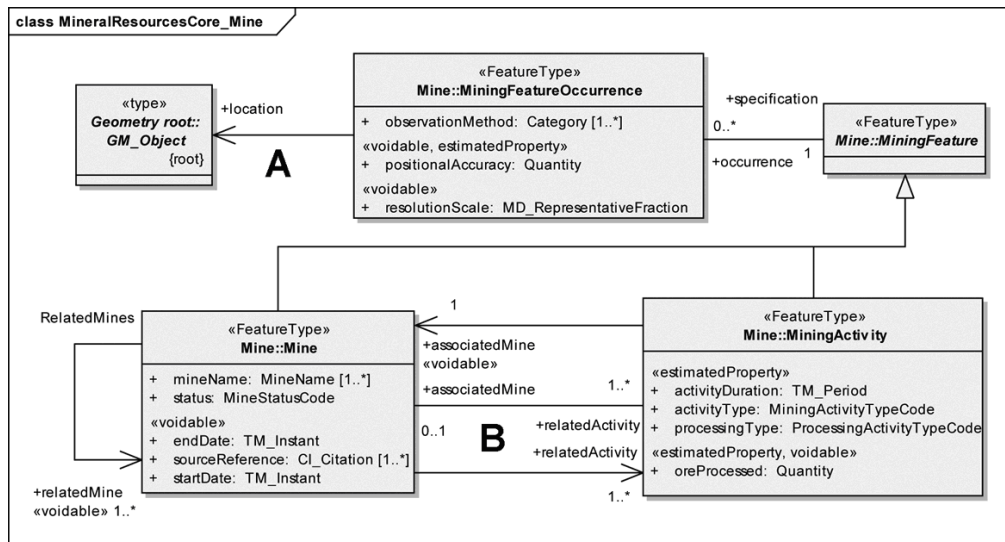
Rys. 11.2. Kolejne dwa przykłady z modelu danych tematu Budynki. Pierwszy dotyczy „pustych” klas bez atrybutów i powiązań, a drugi asocjacyjnego przypisania geometrii do klasy ze stereotypem «featureType».

Rysunek 11.1 jest także przykładem nazwy klasy `TextureTypeType`, która po transformacji do XSD ma aż 3 identyczne przyrostki `TextureTypeTypeType`. Bardziej odpowiednią nazwą tej klasy jest `TypeOfTexture`, co w XSD przełoży się na `TypeOfTextureType`. Oba przypadki przedstawione na tym rysunku pochodzą z wersji 2.0, jednak pozostały w wersji 2.9 i na tej podstawie można przypuszczać, że będą także w wersji 3.0.

Kolejne dwa przykłady też pochodzą ze specyfikacji tematu Budynki w wersji 2.0. Pierwszy dotyczy „pustych” klas `Door` i `Window` wyprowadzonych z klasy `Opening`. Klasy te, jako specjalizacje klasy `Opening`, nie mają żadnych własnych atrybutów ani powiązań i w konsekwencji różnią się od klasy macierzystej tylko nazwą. W takim przypadku poprawnym rozwiązaniem jest dodanie do klasy `Opening` atrybutu w postaci typu wyliczeniowego (ze stereotypem `«enumeration»`) zawierającego dwie pozycje: `Door` i `Window`.

Rysunek 11.2 przedstawia także drugi inny problem, jaki wystąpił w tym modelu. Dotyczy on przypisania geometrii do klasom ze stereotypem `«featureType»` za pomocą jednokierunkowych asocjacji (czarna ramka na rysunku). W takich przypadkach geometria jest atrybutem klasy i w modelu tym podczas prac nad jego transformacją do schematów XSD ktoś musiał dokonać poprawki, a w rezultacie w schematach asocjacja została zamieniona na atrybut. Jednak takie nieudokumentowane zmiany prowadzą do istotnych różnic pomiędzy specyfikacją danych i jej modelem UML a schematami XSD, które stanowią podstawę poprawnego zapisu zbiorów danych tego tematu. W wersji 2.9 ten błąd został poprawiony, jednak przypadek „pustych” klas nadal pozostał.

Kolejne dwa przykłady dotyczące błędów metodyki modelowania danych w języku UML pochodzą z tematu Zasoby mineralne (*Mineral resources*) i są przedstawione na rysunku 11.3. Pierwszy z nich (A) jest analogiczny do przypadku z tematu Budynki



Rys. 11.3. Dwa przykłady niepoprawnych elementów modeli specyfikacji danych tematu Zasoby mineralne (*Mineral resources*).

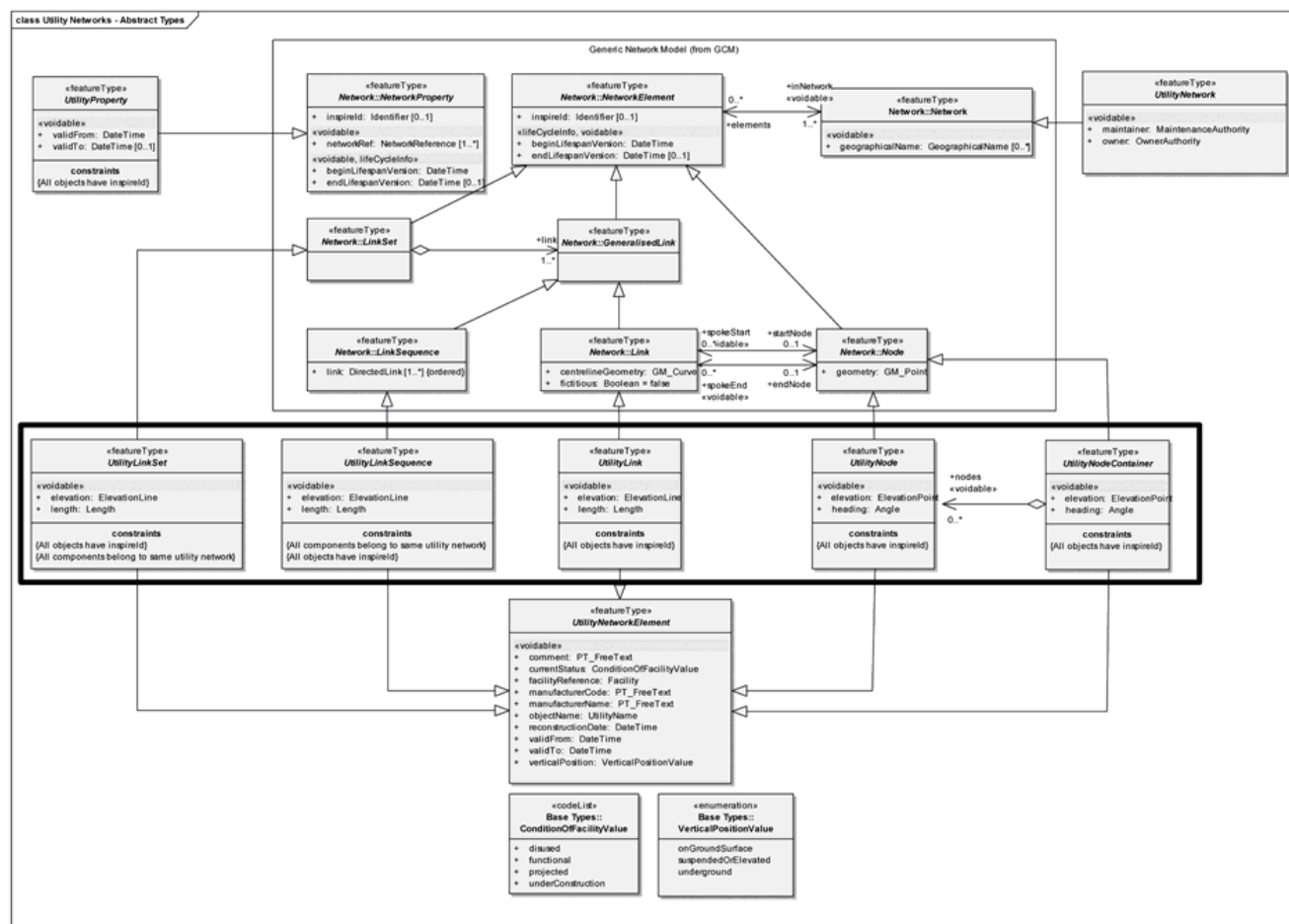
(rys. 11.2) i dotyczy przypisania geometrii przy użyciu asocjacji, jednak w tym przypadku nie zostało to poprawione w wersji 2.9, ale ta asocjacja przestała być widoczna na diagramach tej specyfikacji. Drugi przypadek to podwójne zdefiniowanie asocjacji pomiędzy klasami `Mine` i `MiningActivity`. Jest to częsty błąd popełniany przez osoby niedoświadczone i prawdopodobnie wynika z nałożenia się linii reprezentujących graficznie te asocjacje na pierwotnym diagramie. Na rysunku 11.3, w części B linie te zostały specjalnie rozsunęte tak, aby wszystkie trzy były dobrze widoczne wraz z ich elementami tekstowymi. Błąd ten został usunięty w wersji 2.9.

Rysunek 11.4 przedstawia często powtarzający się błąd dwukrotnego dziedziczenia występujący w specyfikacjach dla różnych tematów, w których korzysta się z bazowego modelu dotyczącego różnych sieci i zdefiniowanego w GCM (*Generic Conceptual Model*) jako `GenericNetworkModel`. Wszystkie pięć klas dotyczących sieci użytkowanych publicznie (na rysunku zaznaczone czarną ramką) ma podwójne dziedziczenie, od klasy `UtilityNetworkElement` i od klas należących do `GenericNetworkModel`. Z czysto pojęciowego podejścia do modeli podwójne (wielokrotne) dziedziczenie jest dopuszczalne, co można zobaczyć na rysunku 2.17. Jednak w modelach przeznaczonych do transformacji do schematów XSD taka konstrukcja nie może wystąpić i jest to kolejny przykład sytuacji, w której ktoś musi dokonać nieudokumentowanej poprawki doprowadzającej do niezgodności schematów XSD ze oficjalną specyfikacją danych.

Problem wyeliminowania wielokrotnego dziedziczenia przedstawionego powyżej (rys. 11.5A) można rozwiązać na trzy sposoby (przyjmując, że obie klasy bazowe w wersji początkowej były abstrakcyjne) i we wszystkich tych trzech przypadkach poprawność modelu pojęciowego zostaje dostatecznie zachowana:

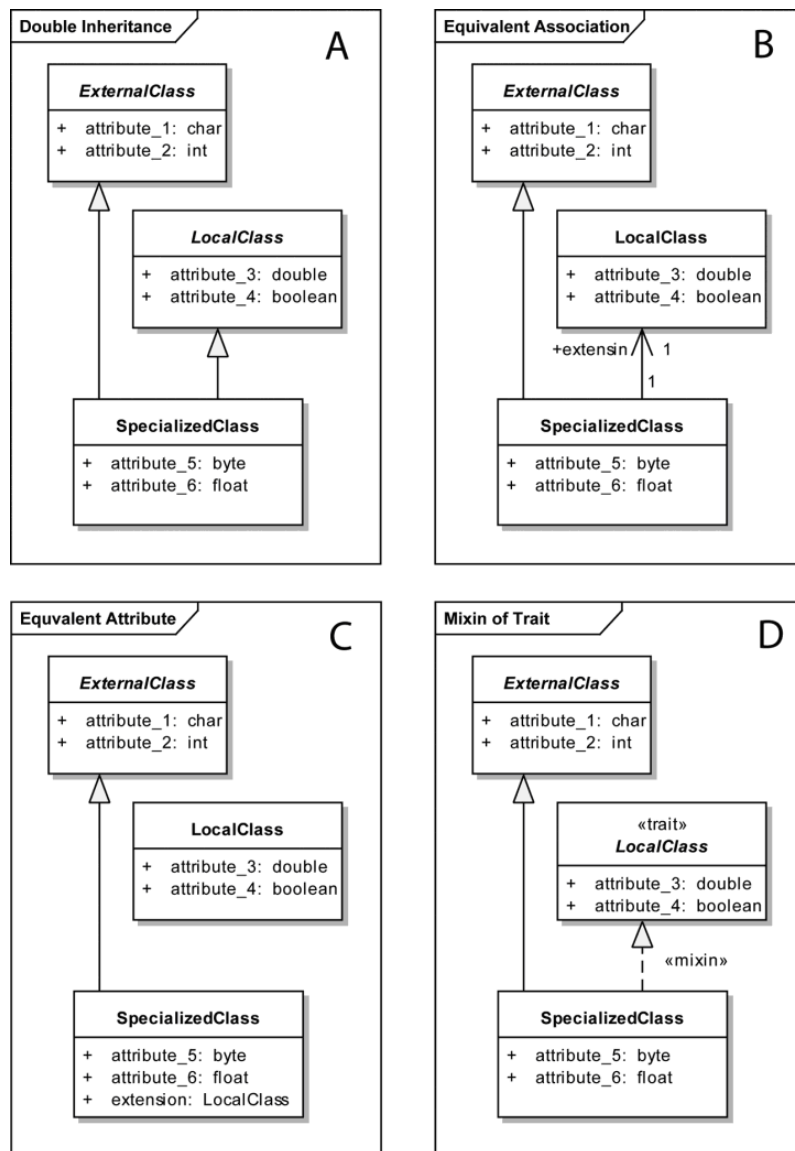
1. Zastąpienie związku dziedziczenia od klasy lokalnej nawigacyjną asocjacją jednokierunkową do tej klasy. Asocjacja ta pozwoli na udostępnienie atrybutów i innych asocjacji tej klasy lokalnej klasie specjalizacyjnej pozbawionej tych własności przez usunięcie dziedziczenia (rys. 11.5B). W tym przypadku klasa lokalna nie może być klasą abstrakcyjną.
2. Wstawienie atrybutu do klasy specjalizowanej typu klasy lokalnej (rys. 11.5C), podobnie jak to można zrobić w przypadku zastąpienia kompozycji przez atrybut (rys. 9.7 w rozdziale 9.2). Również w tym przypadku klasa lokalna nie może być klasą abstrakcyjną.
3. Użycie powiązania do klasy lokalnej typu realizacja ze stereotypem «`mixIn`» (rys. 11.5D). Klasa lokalna ma stereotyp «`trait`», co zakłada, że musi być abstrakcyjna.

Konstrukcje ze stereotypami «`mixIn`» (domieszka) i «`trait`» (cecha) nie są standardowe dla języka UML i w różnych językach programowania są implementowane różnie lub w nich nie występują. Często obie te konstrukcje są traktowane jako konkurencyjne, lecz o różnych właściwościach. W wytycznych kodowania danych przestrzennych INSPIRE (DT_DS, 2010b) problem eliminacji wielokrotnego dziedziczenia jest oparty na konstrukcji *mixIn*, jednak bez wyjaśnienia mechanizmu zamiany klas bazowych na klasy typu domieszki i w konsekwencji model nadal zawiera wielokrotne dziedziczenie. Implementacja tego w schemacie XSD jest uzależniona od wartości metki (*taged value*) `gmlMixIn`, która powinna być w takim przypadku ustawiona na „true”. Z załączonego tam przykładu wynika, że przy takim ustawieniu wartości metki, dziedziczenie jest traktowane jak realizacja ze stereotypem «`mixIn`». Jest to jednak zastosowanie w języku GML rozwiązania niestandardowego dla

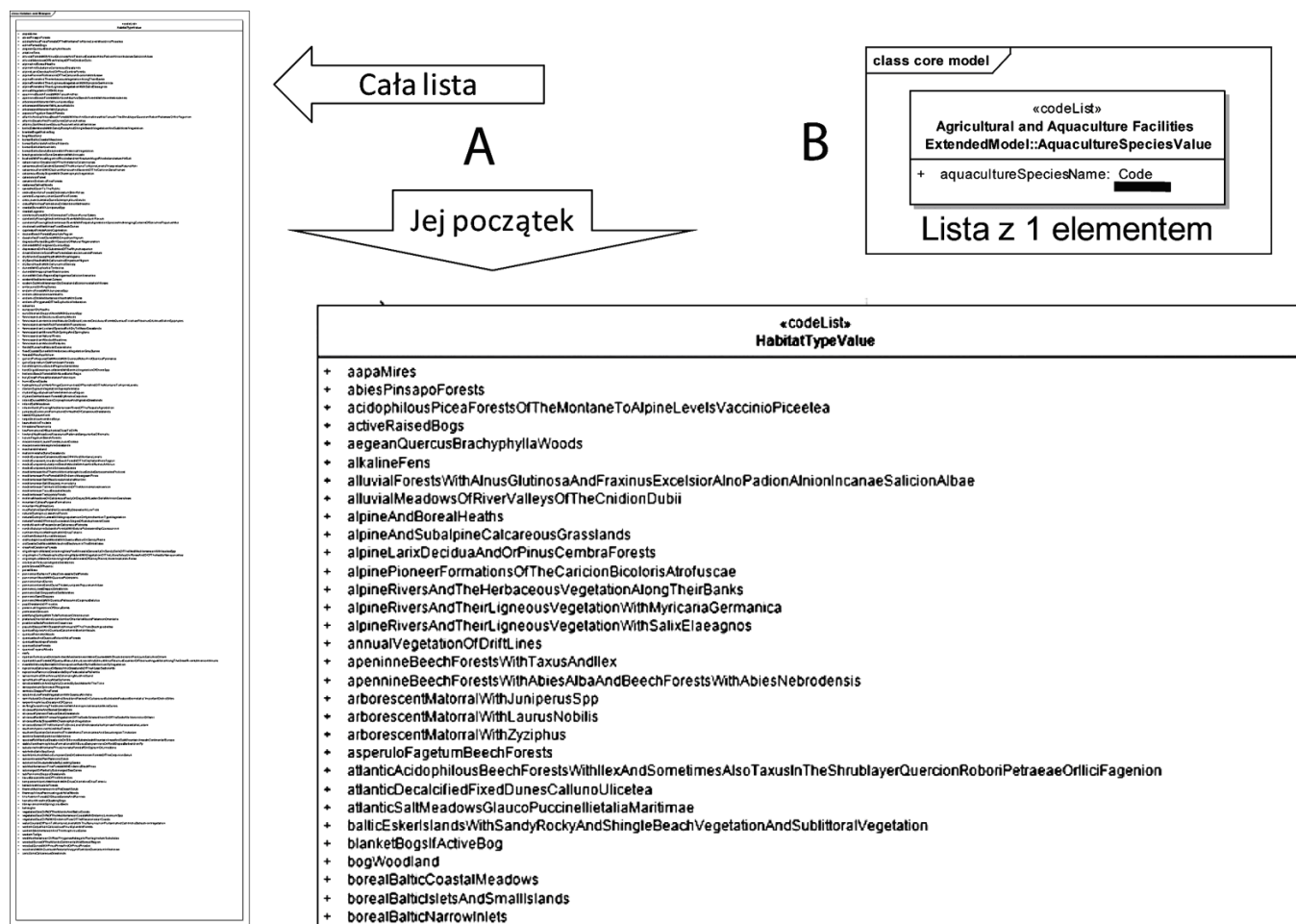


Rys. 11.4. Przykład dwukrotnego dziedziczenia klas w modelu specyfikacji danych dla tematu Usługi użyteczności publicznej i służby państwowej (Utility and governmental services).

aplikacji języka XML i jest obsługiwane jedynie przez niedostępną publicznie wersję oprogramowania *ShapeChange* firmy Interactive Instruments dedykowaną modelom INSPIRE (Woolf, 2009). Z tego względu bezpośrednia transformacja modeli UML tematów INSPIRE z wielokrotnym dziedziczeniem do schematu XSD za pomocą innego oprogramowania (na przykład oprogramowania *FullMoon* lub starszych publicznie dostępnych wersji oprogramowania *ShapeChange*) nie jest możliwa.



Rys. 11.5. Trzy sposoby rozwiązania problemu wielokrotnego dziedziczenia w modelach UML dedykowanych schematom XSD. Objasnienia w tekście.



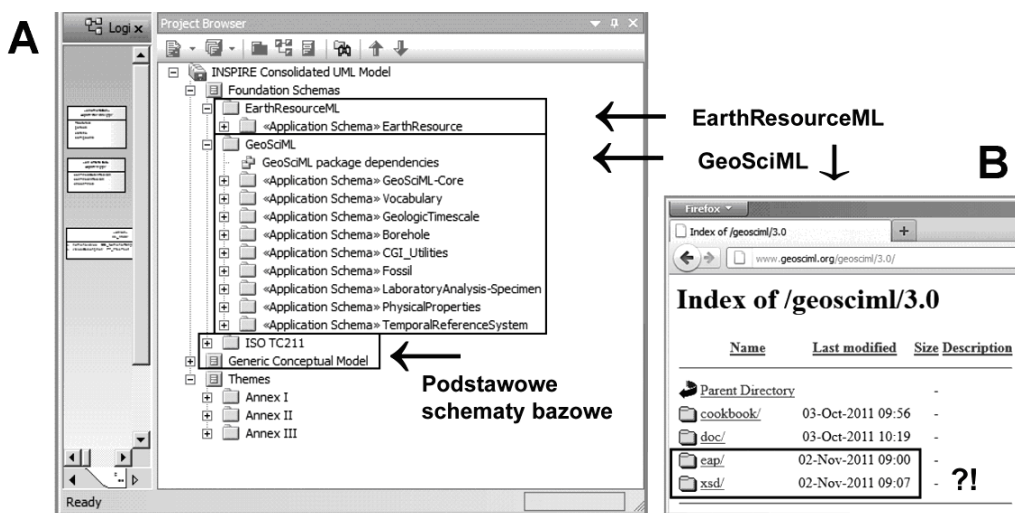
Rys. 11.6. Dwa przypadki nietypowych list kodowych: A – olbrzymia lista typów siedlisk w temacie Siedliska i obszary przyrodniczo jednorodne (*Habitats and biotopes*), B – lista z jedną pozycją w temacie Obiekty rolnicze oraz akwakultury (*Agricultural and aquaculture facilities*).

Inny kłopotliwy problem występujący w modelach danych INSPIRE to nietypowe listy kodowe. Rysunek 11.6 przedstawia dwa skrajne przypadki. Pierwszy to olbrzymia lista kodowa *HabitatTypeValue* w modelu dla tematu Siedliska i obszary przyrodniczo jednorodne (*Habitats and biotopes*). Stosowanie takiej listy w zapisach danych jest wyjątkowo trudne i z tego względu tak długie wyliczenia powinny mieć budowę hierarchiczną, jednak taka konstrukcja w modelach INSPIRE nie jest stosowana. W wersji 2.9 tego modelu lista ta została zastąpiona trzema podziałami siedlisk pod względem: rodzaju pokrycia, formy wegetacji i występujących tam gatunków.

Drugi przypadek to potraktowanie listy kodowej jak zwykłej klasy z jednym atrybutem typu *Code*, który może przybierać wartości odpowiadające poszczególnym pozycjom tej listy. Lista ta była zdefiniowana w modelu tematu Obiekty rolnicze oraz akwakultury (*Agricultural and aquaculture facilities*) i w wersji 2.9 została poprawiona.

11.1. Nietypowy przypadek – temat Geologia

Wśród wielu tematów aneksów II i III dyrektywy INSPIRE na szczególną uwagę zasługuje temat Geologia (*Geology*) i nie dlatego, że autor tego rozdziału jest geologiem, lecz z powodu dużego nagromadzenia ciekawych przypadków nieprawidłowości. Tematem tym, a także tematem Zasoby mineralne (*Mineral resources*) zajmował się jeden roboczy zespół tematyczny (TWG GE-MR) złożony głównie z osób bardzo powiązanych z projektami języków GeoSciML i EarthResourceML. W konsekwencji tego do grupy schematów podstawowych, obok schematów ISO/TC 211, w tym GML, zostały wprowadzone oba te języki, jako baza dla opracowywania modeli dedykowanych poszczególnym tematom (rys. 11.7A).



Rys. 11.7. A – Okno przeglądarki projektu programu *Enterprise Architect* przedstawiające pakiety UML stanowiące podstawę dla modeli tematycznych, B – Fragment okna przeglądarki Firefox przedstawiającego zawartość katalogu serwera z najnowszą wersją 3.0 języka GeoSciML. Objaśnienia w tekście.

Pomysł użycia obu tych języków, jako bazy dla modeli tematycznych, zaowocował poważnym zamieszaniem. Zbudowane na tej bazie modele okazały się metodycznie błędne, ponieważ z wyjątkiem bardzo nielicznych klas własnych składały się z klas i innych elementów importowanych z tych języków. Sytuację tą można rozpatrzeć w dwóch aspektach.

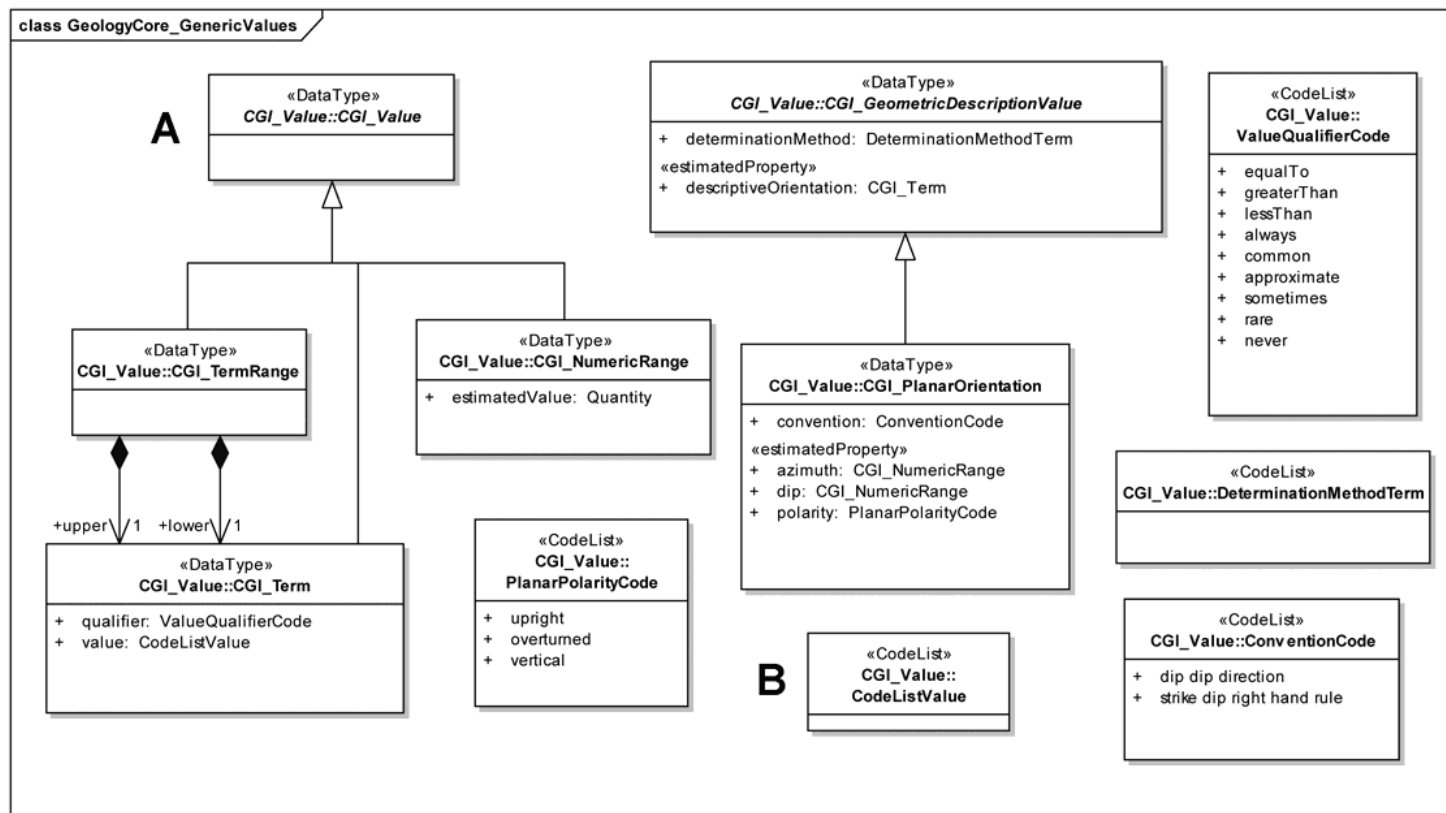
Pierwszy aspekt to zasadnicza różnica pomiędzy językiem dziedzinowym, jakim jest GeoSciML, i schematem aplikacyjnym dla określonego ograniczonego zastosowania, jakim jest temat INSPIRE Geologia. Czym różni się konkretna wąska tematyczna aplikacja od ogólnego języka dziedzinowego jest opisane w rozdziale 2. Tu można przedstawić kilka szczegółów dotyczących tego konkretnego przypadku. Rysunek 11.8 przedstawia diagram klas zawierający podstawowe typy proste (*GenericValues*) zdefiniowane w GeoSciML traktowane jako elementy do ewentualnego wykorzystania w szczegółowych schematach aplikacyjnych lub elementy pełniące role szablonów wymagających dostosowania do konkretnej wąskiej dziedziny, na przykład `CGI_Value::CodeListValue` (rys. 11.8B). Przykładami typów prostych są typy danych wywodzących się z klasy bazowej `CGI_Value::CGI_Value` (rys. 11.8A). Elementy przedstawione na tym rysunku nie zostały zmienione w celu dostosowania ich do potrzeb tej aplikacji – są dokładnie takie, jak je zdefiniowano w języku bazowym.

Drugi aspekt to reguły budowania tematycznych modeli aplikacyjnych na bazie języka będącego dziedzinowym rozwinięciem języka GML. W tym przypadku problem sprowadza się do jednoznacznego określenia, jakie elementy modelu są zapożyczone (importowane) z języka bazowego, a jakie są zdefiniowane lokalnie. Diagram klas przedstawiony na rysunku 11.8 jest częścią modelu „Geology→GeologyMain→GeologyCore” (TWG GE, 2011), jednak klasy na nim przedstawione nie są elementami tego modelu, ponieważ należą do pakietu „GeoSciML→GeoSciML-Core→CGI_Utilites→CGI_Value”, co można rozpoznać po przedrostkach nazw klas.

Przyczyną przedstawionych powyżej błędów metodycznych jest przyjęcie, że te elementy, które widać na diagramach modelu są częściami tego modelu. Błąd ten wynika z braku doświadczenia autora modelu, czego rezultatem jest przekonanie, że część graficzna modelu jest najważniejsza. Zgodnie z tym, co jest opisane w rozdziale 2.6 (rys. 2.19) wiele ważnych szczegółów modelu nie jest widocznych na diagramach, a inne są trudno zauważalne. Edytor języka UML, jak na przykład *Enterprise Architect*, pozwalają na dostosowanie zakresu widocznych na diagramie różnych kategorii elementów przy pomocy funkcji *Feature Visibility*, jednak i tak nie wszystkie mogą być tam widoczne. W takim przypadku dostęp do niewidocznych danych modelu jest realizowany przy pomocy dodatkowych okienek (rys. 2.19, 11.2 i 11.7)

Na tę sytuację nałożył się fakt, że jako języka bazowego użyto najnowszej wersji 3.0 języka GeoSciML, która w tym czasie nie była jeszcze w pełni opracowana i równoległe z pracami nad modelem lokalnym tematu Geologia trwały nadal prace nad językiem GeoSciML. Powstała sytuacja, w której język bazowy zmieniał się w czasie jednocześnie z opracowywaniem na jego podstawie modelu tematycznego. Ostatnia ”sub-wersja” wersji 3.0 została opublikowana w listopadzie 2011 (rys. 11.17B), gdy model tematu Geologia był już ukończony i zakończyły się także prace nad jego testowaniem (Michalak, 2012).

Ostateczny efekt opisanych powyżej błędów modelu UML uwidocznił się dopiero w schematach XSD, które zostały wygenerowane na tej podstawie (rys. 11.19). W schematach tych tylko dwa elementy (*AnthropogenicGeomorphologicFeatureType* i *NaturalGeomorphologicFeature*) ze stereotypem «*featureType*» jako wyróżnienia prze-



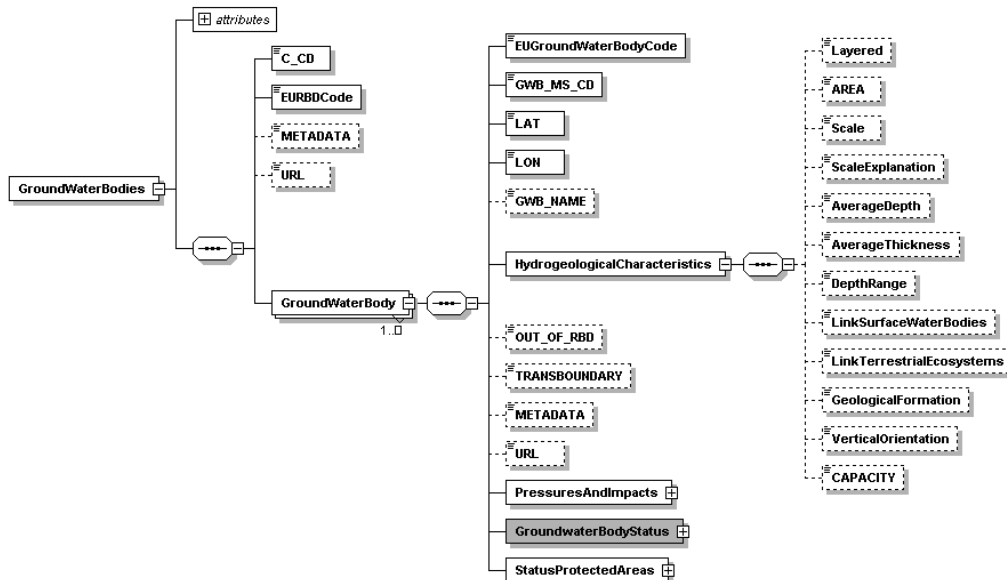
Rys. 11.8. Podstawowe elementy proste (A) i szablony (B) języka GeoSciML przeznaczone do wykorzystania w aplikacjach tego języka. Objasnienia w tekście.

strzenne mogą być wykorzystane do zapisu danych podtematu „GeologyMain”. Wszystkie pozostałe klasy modelu zostały pominięte podczas transformacji wykonanej oprogramowaniem *ShapeChange*, jako nie należące do tego modelu, pomimo że są w tym modelu widoczne i w konsekwencji tego szczegółowo opisane w specyfikacji tekstowej (TWG GE, 2011). Prace testowe dotyczące specyfikacji danych opisane we wstępie do rozdziału 11 wykazały niepoprawność tego modelu i w rezultacie w kolejnej wersji 2.9 model „GeologyMain” został całkowicie zmieniony. Zmiana ta polegała głównie na usunięciu z części bazowej (*Foundation Schemas*) obu języków GeoSciML i EarthResourceML, a w zamian za to umieszczenia ich wybranych fragmentów w postaci lokalnej kopii w obrębie modelu tematu Geologia.

import	loc:http://schemas.opengis.net/gml/3.2.1/gml.xsd	
import	loc:http://inspire-twg.jrc.ec.europa.eu/annexII+III/schemas/geosciml-core/3.0/geosciml-core.xsd	
comment	XML Schema document created by ShapeChange	
element	AnthropogenicGeomorphologicFeatureType	← Element rzeczywisty
complexType	AnthropogenicGeomorphologicFeatureTypeType	
complexType	AnthropogenicGeomorphologicFeatureTypePropertyType	
simpleType	NaturalGeomorphologicFeatureTypeTermType	
simpleType	NaturalGeomorphologicFeatureTypeTermEnumerationType	
simpleType	NaturalGeomorphologicFeatureTypeTermOtherType	
simpleType	AnthropogenicGeomorphologicFeatureTypeTermType	
simpleType	AnthropogenicGeomorphologicFeatureTypeTermEnumerationType	
simpleType	AnthropogenicGeomorphologicFeatureTypeTermOtherType	
element	NaturalGeomorphologicFeature	← Element rzeczywisty
complexType	NaturalGeomorphologicFeatureType	
complexType	NaturalGeomorphologicFeaturePropertyType	
element	GeomorphologicFeature	← Element abstrakcyjny
complexType	GeomorphologicFeatureType	
complexType	GeomorphologicFeaturePropertyType	

Rys. 11.9. Lista elementów schematu aplikacyjnego GeologyCore w oknie programu *XML Spy*. Tylko dwa elementy z tej listy mogą być użyte do zapisu danych. Objaśnienia w tekście.

W obrębie tematu Geologia jest także podtemat Hydrogeologia (*Hydrogeology*) i w początkowych planach miał on stanowić jedynie rozszerzenie podtematu „GeologyMain”. Zamiar ten został zmieniony i w opublikowanej wersji 2.0 (TWG GE, 2011) był w dużym stopniu niezależny. Pozytywną cechą modelu podtematu Hydrogeologia jest jego rozszerzalność pozwalająca na uwzględnienie powiązań z innymi tematami, jak na przykład: Hydrografia, Obszary chronione, Gospodarowanie obszarem, strefy ograniczone i regulacyjne oraz jednostki sprawozdawcze, Urządzenia do monitorowania środowiska i inne. Inne istotne w tym przypadku i możliwe do realizacji rozszerzenia podtematu Hydrogeologia to powiązanie tych danych z innymi danymi hydrogeologicznymi o zasięgu europejskim, jak na przykład dane związane z realizacją postanowień Ramowej Dyrektywy Wodnej (WFD – *Water Framework Directive*) (rys. 11.10). Dane systemu WISE będącego realizacją dyrektywy WFD zawierają wiele cennych informacji (WFD-WG-GIS, 2003; EC, 2009), jednak ich przypisanie przestrzenne jest ograniczone do centroidu – punktu położonego wewnątrz obszaru objętego określoną JCWPd (Jednolite Części Wód Podziemnych), którego położenie jest określone współrzędnymi geograficznymi (układu WGS84).



Rys. 11.10. Diagram elementów schematu XML danych opisujących JCWPd w przyjętym modelu danych systemu WISE w ramach implementacji dyrektywy WFD (Michalak i inni, 2011).

Zagadnienie wzajemnego wiązania różnych zasobów danych przestrzennych i nieprzestrzennych o zasięgu ogólnoeuropejskim, których gromadzenie i udostępnianie wynika z aktów prawnych Unii Europejskiej jest często podkreślane w dokumentach określających rolę i zastosowanie praktyczne infrastruktury INSPIRE.

Problematyka niniejszej monografii stanowi przedmiot szerokiego zainteresowania środowisk współtworzących i współużytkujących infrastrukturę informacji przestrzennej budowaną w Polsce zgodnie z przepisami krajowymi i unijnymi. Zainteresowanie to znalazło swój wyraz w warsztatach „Modele danych przestrzennych w UML i ich transformacja do schematów GML i struktur baz danych”, które odbyły się w ramach konferencji Polskiego Towarzystwa Informatyki Przestrzennej na temat „Informacja przestrzenna dla Polski i Europy”, Warszawa, 7–9 listopada 2011 roku. Odpowiadając na ujawnione wówczas zapotrzebowanie, zespół wykładowców podjął trud zawarcia zaprezentowanych przez siebie treści w opracowaniu o charakterze monograficznym. W rezultacie powstała publikacja, która przedstawia w sposób uporządkowany bogaty zasób wiadomości określonych tytułem warsztatów i dotyczących wybranych metod i technologii geoprzestrzennych.

Godne uznania jest, że zespół autorski w składzie: dr inż. A. Chojka, dr inż. A. Zwirowicz-Rutkowska, dr inż. Z. Parzyński i dr hab. J. Michalak, pełniący rolę redaktora naukowego, zrealizował podjęte przedsięwzięcie w stosunkowo krótkim terminie z niewątpliwą korzyścią dla potencjalnych Czytelników.

Jerzy Gaździcki

Warszawa, maj 2012 r.

Autorzy

dr hab. Janusz Michalak

Wydział Geologii, Uniwersytet Warszawski

J.Michalak@uw.edu.pl

Redakcja naukowa i rozdziały:

1. Wstęp
 2. Różnice pomiędzy językiem zapisu danych i jego dziedzinową aplikacją
 9. Najczęściej popełniane błędy w modelach UML dla schematów aplikacyjnych GML
 11. Schematy aplikacyjne tematów aneksów II i III Dyrektywy INSPIRE
 12. Podsumowanie
- Słownik podstawowych terminów stosowanych w tekście

dr inż. Agnieszka Chojka

Wydział Geodezji i Gospodarki Przestrzennej, Uniwersytet Warmińsko-Mazurski

agnieszka.chojka@uwm.edu.pl

Rozdziały:

3. Wprowadzenie do modelowania informacji przestrzennej – metodyka MDA i diagramy klas UML
6. Budowa schematu aplikacyjnego GML – reguły budowy, narzędzia i przykłady
7. Transformacja schematu aplikacyjnego UML do schematu aplikacyjnego GML – wymagania, ograniczenia i wybrane narzędzia

dr inż. Agnieszka Zwirowicz-Rutkowska

Wydział Geodezji i Gospodarki Przestrzennej, Uniwersytet Warmińsko-Mazurski

agnieszka.zwirowicz@uwm.edu.pl

Rozdziały:

4. Przegląd standardów i narzędzi wykorzystywanych do modelowania informacji geograficznej
5. Schematy aplikacyjne UML – reguły budowy i przykłady
10. Zastosowanie metodyki MDA – wybrane zagadnienia transformacji schematów aplikacyjnych UML do struktur relacyjnych baz danych

dr inż. Zenon Parzyński

Wydział Geodezji i Kartografii, Politechnika Warszawska

z.parzynski@gik.pw.edu.pl

8. Przykład zastosowania metod modelowania danych z zakresu Służby Geodezyjno-Kartograficznej

MODELE DANYCH PRZESTRZENNYCH W UML I ICH TRANSFORMACJA DO SCHEMATÓW GML I STRUKTUR BAZ DANYCH

Słowa kluczowe: geoinformacja, informacja geograficzna, model pojęciowy, UML, schemat aplikacyjny, GML, model relacyjny, transformacja

Streszczenie

Celem monografii jest przedstawienie czytelnikom podstawowych metodyk, technik i narzędzi przeznaczonych do budowy modeli pojęciowych danych przestrzennych na poziomie pojęciowym i implementacyjnym, a następnie do transformacji tych modeli do schematów XSD bazujących na języku GML i do zapisów struktur baz danych w języku DDL. Całość składa się z dwunastu rozdziałów dotyczących poszczególnych aspektów budowy modeli i ich transformacji. Wstęp wprowadza czytelników w całą przedstawianą problematykę i naświetla szerszy teoretyczny kontekst z zakresu modelowania i wykorzystania modeli w zastosowaniach praktycznych. Rozdział drugi poświęcony jest nowym metodom zapisu danych przestrzennych opartego na językach znacznikowych, a w szczególności na języku GML, objaśnia zasady takiego zapisu, zawiera krótką historię języka GML i przedstawia inne języki znacznikowe z nim powiązane. Rozdziały trzeci i czwarty stanowią wprowadzenie do modelowania informacji przestrzennej opartego o metodykę MDA z wykorzystaniem języka UML i zawierają przegląd standardów i narzędzi służących temu modelowaniu. W rozdziałach piątym i szóstym przedstawione są podstawowe zasady budowy tematycznych schematów aplikacyjnych w metodyce języka UML i języka GML zilustrowane przykładami. Rozdział siódmy poświęcony jest zagadnieniom transformacji schematów aplikacyjnych z UML do GML, a w szczególności wymaganiom i ograniczeniom, jakie muszą być spełnione, a także dostępnym narzędziom. Kolejny ósmy rozdział dotyczy modeli UML dedykowanych komponentowi infrastruktury krajowej, przeznaczonym dla Służby Geodezyjnej i Kartograficznej. W rozdziale dziewiątym dokonany jest przegląd najczęściej popełnianych błędów w budowie modeli UML przeznaczonych do utworzenia schematów bazujących na języku GML. Tematem rozdziału dziesiątego jest zastosowanie metodyki MDA do transformacji modeli UML do struktur relacyjnych baz danych. Rozdział jedenasty zawiera metodyczną analizę różnych przypadków występujących w modelach danych tematów aneksów II i III dyrektywy INSPIRE, w tym porównanie z modelami tematów aneksu I, analizę różnych typów i form danych, jakie tam występują. Dwunasty rozdział to podsumowanie, w którym zwraca się szczególną uwagę na dynamiczny rozwój metod z tego zakresu, zmiany zachodzące w zakresie stosowanej terminologii i skutki, jakie te zmiany za sobą pociągają.

UML GEOSPATIAL DATA MODELS AND THEIR TRANSFORMATION INTO GML SCHEMAS AND DATABASE STRUCTURES

Keywords: geoinformation, geographic information, conceptual model, UML, application schema, GML, relational model, transformation

Abstract

The main objective of the monograph is to present essential methodologies, technologies and software tools dedicated to building conceptual models of geospatial data on conceptual level, and implementation level, and then to be transformed into XSD schemas based on GML language and to encode data bases structures in DDL language. The whole monograph consists of twelve chapters concerning different aspects of models development and their transformation. The introduction familiarizes readers with all issues presented and clarifies broader theoretical context with regard to modeling and exploitation of models in practical applications. The second chapter is dedicated to modern methods of encoding spatial data based on markup languages, in particular on GML language; rules for that encoding are also explained. This chapter contains a short history of GML language and presents other markup languages associated with it. Chapters three and four provide an introduction to spatial information modeling based on MDA methodology with application of UML language and it contains a review of standards and tools dedicated to such modeling. In chapters five and six, essential rules of development of thematic application schemas are presented in the methodology of UML and GML languages. Examples to illustrate them are provided. Chapter seven is dedicated to issues of transformation application schemas from UML to GML, in particular to the requirements and constraints that must be fulfilled and also to available tools. The next chapter eight concerns UML models dedicated to components of the national infrastructure designated for Geodetic and Cartographic Service. In chapter nine, a review of most frequent mistakes committed in drawing up UML models dedicated to generating of schemas based on GML language are presented. The subject of chapter ten is the application of MDA methodology for transformation of UML models into relational databases structures. Chapter eleven contains methodological analysis of various cases occurring in data models for the themes defined in Annex II and III of INSPIRE Directive as well as a comparison with the models for themes defined in Annex I and an analysis of various data forms occurring there. In chapter twelve, the recapitulation is presented, in which dynamic development of methods in this area is taken in consideration. In addition, significant changes in the terminology and the effects of these changes are discussed.

Spis treści

1. Wstęp	11
2. Różnice pomiędzy językiem zapisu danych i jego dziedzinową aplikacją	15
2.1. Podstawy zapisu znacznikowego na bazie języka XML	15
2.2. Wprowadzenie do języka GML	18
2.3. Krótka historia zapisu geoinformacji	24
2.4. Języki oparte na GML i z nim powiązane	25
2.5. Przyszłość języka GML	27
2.6. Modele UML dedykowane zapisom w języku GML	30
3. Wprowadzenie do modelowania informacji przestrzennej – metodyka MDA i diagramy klas UML	35
3.1 Wprowadzenie	35
3.2. Reguły budowy schematów aplikacyjnych w UML	39
4. Przegląd standardów i narzędzi stosowanych do modelowania informacji geograficznej	43
4.1. Model dziedzinowy informacji geograficznej	44
4.2. Funkcjonalność narzędzi do modelowania pojęciowego	45
5. Schematy aplikacyjne UML – reguły budowy i przykłady	49
5.1. Pojęcie schematu aplikacyjnego, jego rola i znaczenie	49
5.2. Proces budowy schematów aplikacyjnych	50
5.3. Przykłady schematów aplikacyjnych UML	52
6. Budowa schematu aplikacyjnego GML – reguły budowy, narzędzia i przykłady	55
6.1. Reguły budowy schematów aplikacyjnych GML	55
6.2. Przykład przekształcenia schematu aplikacyjnego UML na GML	66

7. Transformacja schematu aplikacyjnego UML do schematu aplikacyjnego GML – wymagania, ograniczenia i wybrane narzędzia	69
7.1. Metody transformacji UML do GML	69
7.2. Metoda ręczna	70
7.3. Metoda automatyczna	71
7.4. Podsumowanie	76
8. Przykład zastosowania metod modelowania danych z zakresu Służby Geodezyjno-Kartograficznej	79
8.1. Założenia przyjęte w GUGiK przy opracowywaniu projektów rozporządzeń	79
8.2. Realizacja założeń	80
8.3. Przykłady schematów aplikacyjnych do projektów rozporządzeń	83
9. Najczęściej popełniane błędy w modelach UML dla schematów aplikacyjnych GML	87
9.1. UML jest cierpliwy jak papier	87
9.2. Wymagania dotyczące modeli UML dla INSPIRE	90
10. Zastosowanie metodyki MDA – wybrane zagadnienia transformacji schematów aplikacyjnych UML do struktur relacyjnych baz danych	95
10.1. Transformacja w ujęciu metodyki MDA	95
10.2. Ogólne zasady mapowania pomiędzy modelem obiektowym i modelem relacyjnym	97
10.3. Transformacja schematu aplikacyjnego UML do logicznej struktury relacyjnej bazy danych	101
11. Schematy aplikacyjne tematów aneksów II i III dyrektywy INSPIRE	107
11.1. Nietypowy przypadek – temat Geologia	115
12. Podsumowanie	121
Słownik podstawowych terminów stosowanych w tekście	125
Literatura	131

Janusz Michalak

Słownik podstawowych terminów stosowanych w tekście

Abstrakcyjny – obiekt, atrybut, typ, klasa (*abstract – object, attribute, type, class*) – **1:** Określony ogólnie, bez szczegółów związanych z określoną implementacją (uwarunkowaniami technologicznymi) lub z określoną aplikacją (uwarunkowaniami wynikającymi z dziedziny zastosowania). Na przykład wynik pomiaru w znaczeniu ogólnym jako atrybut w modelu pojęciowym nie musi mieć określonego typu. Typ będzie zależał od fizycznego charakteru mierzonego elementu i od typu przyrządu pomiarowego. **2:** Klasa abstrakcyjna w modelu danych to klasa, która nie ma własnych obiektów, a jedynie służy jako klasa bazowa dla innych klas. Użycie takiej klasy jest uzasadnione tylko gdy są (lub mogą być) wyprowadzone z nie przynajmniej dwie klasy.

Atrybut (*attribute*) – Właściwość wyróżnienia lub obiektu określona przez nazwę tej właściwości i zakres wartości, jakie mogą być przypisane tej nazwie dla określenia tej właściwości.

Atrybut geoprzestrzenny (*geospatial attribute*) – Właściwość (cecha) wynikająca z faktu, że wyróżnienie zajmuje pewne miejsca w rzeczywistości w sensie geoprzestrzennym. Najczęściej przez domniemanie przyjmuje się, że określenie geoprzestrzenny obejmuje również czas, czyli jest równoznaczne z określeniem czaso-geoprzestrzenny. Przykładami takich atrybutów są: wielkość, kształt, położenie, przynależność geoprzestrzenna (np. leży w obrębie), relacje geoprzestrzenne względem innych wyróżnień (np. odległość lub rodzaj sąsiedztwa).

Atrybut niegeoprzestrzenny (*non-geospatial attribute*) – Wszystkie pozostałe atrybuty niezwiązane z odniesieniem przestrzennym. Atrybuty te mogą należeć zarówno do wyróżnień geoprzestrzenne jak i do innych obiektów i wystąpień niegeoprzestrzennych.

Cecha (*trait*) – Kategoria klasy, której zadaniem jest (w przypadku modeli danych) dostarczenie innej klasie określonych własności (atrybutów i powiązań z innymi klasami). W tym przypadku klasa ma stereotyp «*trait*». Podobnym mechanizmem pozwalającym na uniknięcie problemów wielokrotnego dziedziczenia jest **domieszka**.

Dane (*data*, w liczbie pojedynczej: *datum*) – **1:** Jednostki informacji, czyli pojedyncze fragmenty informacji. Dane niezorganizowane nie stanowią informacji i często są bezużyteczne. Dane zorganizowane stanowią elementy informacji. Zorganizowanie danych może być jawne, na przykład w językach znacznikowych lub niejawne, na przykład miejsce umieszczenia adresu na kopercie decyduje, czy jest to adres nadawcy czy odbiorcy. **2:** Fakty, statystyki, opinie i przewidywania zebrane z różnych wewnętrznych i zewnętrznych źródeł. Dane bez kontekstu są szumem (Nowicki i Staniszkis, 2002).

Dane geoprzestrzenne (*geospatial data*) – **1:** Dane w sensie zdefiniowanym przez informatykę, ale w odróżnieniu od innych rodzajów danych są one odniesione do określonego miejsca (fragmentu przestrzeni) i w rezultacie niezbędnymi ich składnikami są dane określające położenie tego miejsca względem Ziemi. **2:** Dane przestrzenne dotyczące Ziemi i wszystkich obiektów przestrzennych z nią związanych (Gaździcki, 2004).

Domieszka (*mixin*) – **1:** Kategoria klasy, której zadaniem jest (w przypadku modeli danych) dostarczenie innej klasie określonych własności (atrybutów i powiązań z innymi klasami). Taka klasa nie ma własnych obiektów, czyli musi być abstrakcyjna. Stosowanie tego rodzaju klasy jest uzasadnione tylko w przypadkach, gdy przynajmniej dwie zwykłe klasy otrzymują w ten sposób własności. Jest to sposób na uniknięcie problemów z implementacją wielokrotnego dziedziczenia. Jedyny przypadek zastosowania klasy *mixin* do języka GML to modele dla niektórych tematów INSPIRE. **2:** Ograniczony sposób dziedziczenia ma pozwalający również na uniknięcie problemów z implementacją wielokrotnego dziedziczenia. W takim przypadku powiązanie dziedziczenia ma stereotyp «*mixin*». Porównaj: **cecha**.

Encja (*entity*) – Pojęcie z modelu encja-związek, oznaczające konkretny lub abstrakcyjny byt wyróżnialny w modelowanej rzeczywistości. W odróżnieniu od obiektu, encja nie jest kojarzona z metodami (Subieta, 1999a).

GML (*Geography Markup Language*) – Język znaczników geograficznych, aplikacja języka (metajęzyka) XML przeznaczona do zapisu geoinformacji w celu przesyłania jej pomiędzy różnymi systemami – on-line, niezależnie od platformy sprzętowo-systemowej i niezależnie od charakteru i technologii systemu geoinformacyjnego (Gaździcki, 2004).

Informacja (*information*) – **1:** Dane komputerowe, które są zorganizowane i przedstawione w usystematyzowanej formie dla zrozumiałości ich podstawowego znaczenia. Związki pomiędzy informacją i danymi wyjaśnia definicja danych. **2:** Dane interpretowane w kontekście określonego celu (Nowicki i Staniszkis, 2002). **3:** Wiedza uzyskiwana w drodze interpretacji danych, która w ustalonym kontekście ma określone znaczenie i dotyczy obiektów, takich jak fakty, zdarzenia, przedmioty, zjawiska, procesy i idee (Gaździcki, 2004).

Informacja geograficzna – patrz: **informacja geoprzestrzenna**.

Informacja geoprzestrzenna (*geospatial information*) – **1:** Informacja w sensie zdefiniowanym przez informatykę, ale w odróżnieniu od innych rodzajów informacji jest ona odniesiona do określonego miejsca (fragmentu przestrzeni) i w rezultacie niezbędnymi jej składnikami są dane określające położenie tego miejsca względem Ziemi. **2:** Informacja uzyskiwana w drodze interpretacji danych geoprzestrzennych (Gaździcki, 2004).

Instancja (*instance*) – Synonim egzemplarza stosowany w normach PN-EN ISO 19100 (Gaździcki, 2011).

Klasa (*class*) – Pojęcie klasy jest używane w trzech dość bliskich znaczeniach: **(1)** zbiór obiektów o zbliżonych własnościach; **(2)** byt semantyczny rozumiany, jako miejsce przechowywania takich cech grupy podobnych obiektów, które są dla nich niezmiennie (np. zestawu atrybutów, nazwy, metod, ograniczeń dostępu); **(3)** wyrażenie językowe specyfikujące budowę obiektów, dozwolone operacje na obiektach, ograniczenia dostępu, wyjątki, itd. Zwykle klasy wiąże się ze sobą poprzez hierarchię (lub inną strukturę) dziedziczenia (Subieta, 1999a).

MDA – 1: (*Model Driven Approach*) Podejście oparte na modelu: pojęciowym, logicznym i fizycznym. Niezależny od implementacji schemat aplikacyjny zostaje odwzorowany na różne specyfikacje (wykorzystujące różne technologie, np. usługi sieciowe, relacyjne bazy danych, XML), a te z kolei mogą zostać zaimplementowane (wdrożone) na różnych platformach sprzętowo-programowych (CEN, 2006). **2:** (*Model Driven Architecture*) Zbiór metod porządkujących proces tworzenia systemów informatycznych opartych na budowie modeli i ich transformacji. Koncepcja MDA została opracowana przez międzynarodową organizację OMG, której celem jest rozwiązywanie problemów związanych z integracją systemów informatycznych pochodzących od różnych dostawców oraz działających na różnych platformach informatycznych (OMG, 2003).

Metamodel (*metamodel*) – W założeniu, model definiujący składnię, semantykę i pragmatykę wprowadzonego modelu, notacji lub diagramu. Metamodel proponowany przez autorów UML ustala pewne elementy składni diagramów, ograniczenia typologiczne, klasyfikację pojęć oraz związki pomiędzy pojęciami (Subieta, 1999a).

Metka (*tagged value*) – Inaczej wartość etykietowana. Obok stereotypów i ograniczeń, to jeden z mechanizmów rozszerzenia semantyki języka UML. Pozwala dołączyć do elementu modelu UML dodatkowe właściwości. Metka to para *klucz=wartość*.

Metodyka (*methodology*) – Zestaw pojęć, notacji, modeli formalnych, języków i sposobów postępowania służący do analizy rzeczywistości (stanowiącej przedmiot projektowanego systemu informatycznego) oraz do projektowania pojęciowego, logicznego i/lub fizycznego. Zwykle metodyka jest powiązana z odpowiednią notacją (diagramami) służącymi do zapisywania wyniku poszczególnych faz projektu, jako środek wspomagający ludzką pamięć i wyobraźnię i jako środek komunikacji w zespołach oraz pomiędzy projektantami i klientem (Subieta, 1999a).

Model pojęciowy (*conceptual model*) – Model procesów lub model struktury danych odwołujący się do ludzkiej percepcji i wyobraźni, mający za zadanie zrozumienie problemu, udokumentowanie wyniku analizy lub projektu w czytelnej i abstrakcyjnej formie językowej oraz ułatwienie komunikacji w zespołach ludzkich (Subieta, 1999a).

Model semantyczny (*semantic model*) – Zestaw pojęć, technik i notacji mający na celu odwzorowanie semantyki danych, czyli ich znaczenia w świecie zewnętrznym. Modele semantyczne wprowadzają w tym celu pojęcia, takie jak: generalizacja, specjalizacja, asocjacja, agregacja, klasyfikacja, własności temporalne, zdarzenia, własności behawioralne, itd. Przykładem prostego modelu semantycznego jest model encja-związek. Niekiedy terminem “model semantyczny” określa się również konkretny diagram (lub inną formę językowo-graficzną) odwzorowującą rzeczywistość opisywaną przez dane (Subieta, 1999a).

Norma (*standard*) – **1:** Dokument przyjęty na zasadzie konsensu i zatwierdzony przez upoważnioną jednostkę organizacyjną, ustalający – do powszechnego i wielokrotnego stosowania – zasady, wytyczne lub charakterystyki odnoszące się do różnych rodzajów działalności lub ich wyników i zmierzający do uzyskania optymalnego stopnia uporządkowania w określonym zakresie (Ustawa, 2002). **2:** Polska Norma – jest normą o zasięgu krajowym, przyjętą w drodze konsensu i zatwierdzoną przez krajową jednostkę normalizacyjną (Polski Komitet Normalizacyjny), powszechnie dostępną, oznaczoną – na zasadzie wyłączności – symbolem PN (PKN, 2010). Zobacz: **normy ISO serii 19100, standard, standardy OGC.**

Normy ISO serii 19100 (*ISO 19100 series of International Standards*) – Rodzina norm ISO w dziedzinie informacji geograficznej. Wynik prac Komitetu Technicznego ISO/TC211, który pracuje nad wieloma projektami standaryzacji informacji przestrzennej w bardzo szerokim zakresie tej problematyki. Zobacz: **norma, standard, standardy OGC**.

Obiekt (*object*) – **1:** W teorii informacji – konkretny lub abstrakcyjny byt (wystąpienie) wyróżnialny w modelowanej rzeczywistości, posiadający nazwę, jednoznaczną identyfikację, wyraźnie określone granice, atrybuty i inne właściwości takie jak rodzaj struktury wewnętrznej lub struktury danych z nim związanych. Te składniki obiektu określają: jego stan (poprzez wartości atrybutów i powiązania) i jego zachowanie się (poprzez operatory i funkcje, czyli metody) (Subieta, 1999a). **2:** W geomatyce przyjmuje się, że obiekt jest wystąpieniem klasy i jest to oparte na paradygmacie obiektowości wywodzącym się z języka UML, który jest przyjęty do opisu modeli pojęciowych (OMG, 2001). **3:** Termin stosowany w różnych znaczeniach; dla uniknięcia wątpliwości, jeśli jego znaczenie nie wynika z kontekstu, powinien być uzupełniony dodatkowym określeniem (Gaździcki, 2004).

Rola (*role*) – W języku UML jedna z możliwości opisu powiązania. Pozostałe to nazwa powiązania oraz krotność. Każda klasa biorąca udział w powiązaniu ogrywa w nim określoną rolę. Inaczej jest to „oblicze”, które klasa przy jednym końcu powiązania prezentuje klasie przy drugim jego końcu.

Schemat (*schema*) – **1:** Opis logicznej struktury bazy danych lub innego systemu związanego z danymi, np. interfejsu wymiany danych (XML Schema). **2:** Opis atrybutów wyróżnień (*feature*), lub bardziej dokładnie – specyficzny model atrybutów dla wyróżnień określony za pomocą elementarnych typów danych i ograniczeń dotyczących tych typów (Buehler, McKee, 1996).

Schemat aplikacyjny (*application schema*) – Schemat przeznaczony dla konkretnego systemu lub dla konkretnej dziedziny zastosowań.

Schemat implementacyjny (*implementation schema*) – Schemat uwzględniający technologiczne środowisko, w którym będzie realizowana jego aplikacja. Na przykład zapisany w formie schematu XML.

Specyfikacja (*specification*) – **1:** Abstrakcyjny opis bytu programistycznego (procedury, modułu, klasy, obiektu, bazy danych, itp.) określający reguły użycia lub ustalający podstawowe założenia jego implementacji (Subieta, 1999a). **2:** Dokument lub opis, który określa w sposób kompletny, precyzyjny i sprawdzalny wymagania, projekt lub charakterystykę systemu lub jego fragmentu, a często także procedury dla określenia czy te wymagania są spełnione.

Standard (*standard*) – Wzorzec rozwiązania sprzętowego lub programowego zatwierdzony przez instytucję normalizacyjną lub przyjęty nieformalnie wskutek dużego upowszechnienia, w przypadku standardów informatycznych najczęściej o zasięgu światowym. Do najważniejszych instytucji opracowujących standardy należą: ISO, IEEE, ANSI. Przykładami standardów są: RS-232-C (fabryczny standard interfejsu sprzętowego), ANSI C++ (oficjalny standard języka programowania), POSIX (standard IEEE przenośnego systemu uniksowego), CORBA (standard obiektowych systemów rozproszonych) (Płoski, 1999). Zobacz: **standardy OGC, norma, normy ISO serii 19100**.

Standardy OGC (*OGC standards*) – Techniczne dokumenty specyfikujące interfejsy i reguły zapisu danych geoprzestrzennych. Stanowią one główne rezultaty działalności OGC (*Open Geospatial Consortium*) i są opracowywane przez zespoły złożone z członków OGC dla rozwiązywania różnorodnych problemów dotyczących interoperacyjności. Wszystkie publiczne dokumenty OGC są łatwo dostępne bez żadnych opłat. OGC ma ponad 400 członków, w tym ponad połowa to wyższe uczelnie i instytucje naukowe, także prawie połowę stanowią członkowie europejscy. Standardy OGC dzielą się na specyfikacje abstrakcyjne i standardy implementacyjne. Wiele z tych standardów zostało przyjęte przez komitet ISO/TC 211 jako normy ISO, na przykład: 19107, 19115, 19119, 19123, 19125, 19128, 19136, 19139, 19142, 19143, i 19156. Ze standardami OGC powiązane są inne oficjalne dokumenty OGC, na przykład: *OGC Reference Model (ORM)*, *Engineering Reports* lub nieoficjalne, na przykład *Best Practices Documents* i *Discussion Papers*. Zobacz: **standard, norma, normy ISO serii 19100**.

Stereotyp (*stereotype*) – W terminologii UML, klasyfikacja elementu modelu posiadająca semantyczne konsekwencje. Stereotypy mogą być predefiniowane lub zdefiniowane przez użytkownika (Subieta, 1999a).

Struktura (*structure*) – Termin w C++ (także w innych językach) na oznaczenie zestawu nazwanych wartości, w innych językach odpowiada jej zapis lub rekord (Subieta, 1999a).

Tabela (*table*) – Struktura danych implementowana w relacyjnych bazach danych, często nazywana relacją. Tabela składa się z wierszy lub inaczej krotek. Należy zwrócić uwagę, że pomiędzy relacją (w sensie matematycznym) i tabelą występują dość istotne różnice koncepcyjne. Tabela jest wyposażona w nazwy kolumn (które niosą informację semantyczną) (Subieta, 1999a).

Tożsamość (*identity*) – Tożsamość obiektu oznacza, że obiekt istnieje i jest odróżnialny niezależnie od jego aktualnego stanu (wartości atrybutów), który może się zmieniać; możliwe są dwa różne obiekty o identycznych wartościach atrybutów. Praktycznie, tożsamość oznacza istnienie unikalnego wewnętrznego (nieczytelnego dla użytkownika) identyfikatora obiektu, który nie ulega zmianie podczas życia obiektu (Subieta, 1999a).

Unia (*union*) – Typ struktury, rekordu lub obiektu, który może mieć alternatywnie dwa lub więcej zestawów atrybutów. Przykładowo, jeżeli właścicielem samochodu może być osoba lub firma, to obiekt *Samochód* może posiadać alternatywnie albo atrybut *Nazwisko Właściciela* albo atrybut *WłasnośćFirmy*. Unia może mieć związany dyskryminator (*discriminator*), tj. atrybut, którego wartość określa, z którym wariantem mamy do czynienia. Może też nie mieć dyskryminatora; wówczas odpowiedzialność za rozróżnianie wariantów spada na programistę (tak jest np. w C i C++). Brak dyskryminatora w unii podkopuje koncepcję mocnej kontroli typów i stwarza okazję do bardzo trudnych błędów (Subieta, 1999a).

Walidator (*validator*) – Program komputerowy sprawdzający poprawność dokumentu (np. XML) o określonej składni.

Wyróżnienie geoprzestrzenne (*geospatial feature*) (w literaturze polskiej termin *feature* jest często tłumaczony jako obiekt) – **1**: Podstawowy fragment (atom) informacji geoprzestrzennej. Posiada atrybuty geoprzestrzenne (geometryczne i topologiczne) np. kształt, rozciągłość, położenie, relacje z innymi wyróżnieniami. Często pojęcie wyróżnienie jest my-

lone z pojęciem obiekt, jednak wyróżnienie może być obiektem, ale też może nim nie być (Mark i in., 2001). Ponieważ w geomatyce wszystkie wyróżnienia są geoprzestrzenne, przymiotnik geoprzestrzenny jest na ogół pomijany i używa się krótszego terminu wyróżnienie.

2: Cyfrowa reprezentacja zjawiska (bytu) świata rzeczywistego lub jego abstrakcja w modelu pojęciowym. Wyróżnienie ma określone miejsce w przestrzeni i czasie jako jego atrybuty (Buehler, McKee, 1996). Przykładem wyróżnienia może być prawie wszystko co może być umieszczone w przestrzeni i czasie: stół, budynek, miasto, drzewo, fragment lasu, ekosystem, trasa przejazdu lub wyż atmosferyczny jako obszar wysokiego ciśnienia powietrza.

3: Abstrakcja zjawiska świata rzeczywistego. Termin wyróżnienie może odnosić się do typu zjawiska lub jego konkretnego wystąpienia (ISO/TC 211, 2002a), np. „rzeka” i „Wisła”.

Związek (*relationship*) – **1:** W języku UML i w konsekwencji także w normach grupy ISO 19100 – semantyczne połączenie pomiędzy elementami modelu. Przykładami związków są agregacje, kompozycje (agregacje całkowite), powiązania i uogólnienia. **2:** W modelu encji-relacji – powiązanie pomiędzy encjami (Michalak, 2005a).

Literatura

- AB ORMSC (Architecture Board ORMS), 2001: Model Driven architecture (MDA). Document number ormsc/2001-07-01.
URL: <http://www.enterprise-architecture.info/Images/MDA/MDA%20Technical.pdf>
- Altova, XMLSpy. URL: <http://www.altova.com/xml-editor>
- BGWM (Biuro Geodety Województwa Mazowieckiego), 2009: Opis koncepcji identyfikatorów, wersjonowania zmian, stosowania reguły *nil reason*.
URL: <http://www.geointegracja.gov.pl/download/file.php?id=80&sid=f2c1f79e942a2cf12ed12a99aee5eec0>
- Biron P. V., Permanente K., Malhotra A. (W3C), 2004: XML Schema Part 2: Datatypes. Second Edition. W3C Recommendation 28 October 2004. URL: <http://www.w3.org/TR/xmlschema-2>
- Boisvert E., Brodaric B., 2008: GroundWater Markup Language Specification v. 1.0.
URL: http://ngwd-bdnes.cits.nrcan.gc.ca/service/api_ngwds/en/gwml.html
- Booch G., Rumbaugh J., Jacobson I., 2002: UML – przewodnik użytkownika. Z serii: Inżynieria oprogramowania. Wydanie polskie. Wydawnictwa Naukowo-Techniczne, Warszawa.
- Brink L., Portele C., Vretanos P. A. (OGC), 2011: Geography Markup Language (GML) simple features profile (with Corrigendum). OpenGIS Implementation Standard Profile.
URL: http://portal.opengeospatial.org/files/?artifact_id=42729
- Buechler K., McKee L. (ed.), 1996: The OpenGIS Guide – Introduction to Interoperable Geoprocessing – Part I of the Open Geodata Interoperability Specification (OGIS). OGIS TC Document 96-001, Open GIS Consortium, Wayland.
- Burggraf D., 2011: Input to the GML 4 workshop.
URL: http://external.opengeospatial.org/twiki_public/GML/Gml4WorkshopInput
- CEN, 2006: prCEN/TR 15449, Geographic information – Standards, specifications, technical reports and guidelines, required to implement Spatial Data Infrastructure.
- Chojka A., 2006: Przegląd metod, środków formalnych i narzędzi programowych wspomagających modelowanie pojęciowe informacji geograficznej. Część I – Modelowanie pojęciowe. *Magazyn Geoinformacyjny Geodeta*, nr 5 (132).
- Cox S. (ed.) (OGC), 2010: Geographic Information: Observations and Measurements – OGC Abstract Specification Topic 20. URL: http://portal.opengeospatial.org/files/?artifact_id=41579
- Cox S., 2011: Hollow World: a GML application schema template. Solid Earth and Environment GRID (SEE GRID community website). URL: <https://www.seegrid.csiro.au/wiki/AppSchemas/HollowWorld>
- CTWG-O&M (INSPIRE Cross Thematic Working Group on Observations & Measurements), 2011: D2.9 Guidelines for the use of Observations & Measurements and Sensor Web Enablement-related standards in INSPIRE Annex II and III data specification development, Version 1.0.
URL: http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.9_O&M_Guidelines_V1.0.pdf
- DT_DS (INSPIRE Drafting Team „Data Specifications”), 2008: D2.6: Methodology for the development of data specifications, Version 3.0.
URL: http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.6_v3.0.pdf
- DT_DS (INSPIRE Drafting Team „Data Specifications”), 2010a: D2.5: Generic Conceptual Model, Version 3.3. URL: http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.5_v3_3.pdf

- DT_DS (INSPIRE Drafting Team „Data Specifications”), 2010b: D2.7: Guidelines for the encoding of spatial data, Version 3.2. URL: http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.7_v3.2.pdf
- EC (European Commission), 2009: Guidance Document No. 22 – Updated Guidance on Implementing the Geographical Information System (GIS) Elements of the EU Water policy. Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Technical Report – 2009 – 028.
URL: http://circa.europa.eu/Public/irc/env/wfd/library?l=/framework_directive/guidance_documents/guidance-no22-_nov08pdf_1/_EN_1.0_&a=d
- EP&CEU (European Parliament and Council of the European Union), 2007: Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).
- Fowler M., Scott K., 2002: UML w kropelce. Oficyna wydawnicza LTP, Warszawa.
- Gaździcki J., 2004: Leksykon geomatyczny – Lexicon of geomatics. Polskie Towarzystwo Informatyki Przemysłowej, Warszawa.
- Gaździcki J., 2011: [W:] (red.) Gaździcki J. Internetowy leksykon geomatyczny.
URL: <http://www.ptip.org.pl/auto.php?page=Encyclopedia&enc=1>
- Githaiga J., 2010: Project Overview – FullMoon. Solid Earth and Environment GRID (SEE GRID community website). URL: <https://www.seegrid.csiro.au/wiki/Siss/FullMoon>
- Huang C-H., Chuang T-R., Deng D-P., Lee H-M., 2009: Building GML-native web-based geographic information systems. *Computers&Geosciences*, no 35, 1802-1816.
URL: <http://www.iis.sinica.edu.tw/papers/trc/8843-F.pdf>
- IGW-CGI-IUGS (Commission for the Management and Application of Geoscience Information – CGI, Interoperability Working Group – IWG, International Union of Geological Sciences – IUGS), 2008: GeoSciML Cookbook – How To Map Data to GeoSciML, Version 2.
URL: http://www.geosci.ml.org/geosci.ml/2.0/cookbook/GeoSciML_Data_CookBook_V2.pdf
- ISO/TC 211 (Geographic Information/Geomatics), 2002a: ISO 19101: Geographic information – Reference model. URL: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26002
- ISO/TC 211 (Geographic Information/Geomatics), 2002b: ISO 19108:2002 Geographic information – Temporal schema.
URL: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26013
- ISO/TC 211 (Geographic Information/Geomatics), 2003: ISO 19107:2003 – Geographic information – Spatial schema. URL: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26012
- ISO/TC 211 (Geographic Information/Geomatics), 2005a: ISO 19103 Technical Specification, Geographic information – Conceptual schema language.
URL: http://www.iso.org/iso/catalogue_detail.htm?csnumber=37800
- ISO/TC 211 (Geographic Information/Geomatics), 2005b: ISO 19109:2005 Geographic information – Rules for application schema.
URL: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39891
- ISO/TC 211 (Geographic Information/Geomatics), 2006: ISO 19110 – Geographic information – Methodology for feature cataloguing.
URL: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39965
- ISO/TC 211 (Geographic Information/Geomatics), 2007a: ISO 19136:2007 – Geographic information – Geography Markup Language (GML).
URL: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=32554
- ISO/TC 211 (Geographic Information/Geomatics), 2007b: ISO 19139 Technical Specification, Geographic information – Metadata – XML schema implementation.
URL: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=32557
- ISO/TC 211 (Geographic Information/Geomatics), 2009: ISO 19104:2009 Technical Specification, Geographic information – Terminology
URL: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=32557
- ISO/TC 211 (Geographic Information/Geomatics), 2011: ISO 19118, Geographic information – Encoding.
URL: http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=44212
- ISO/TC 211 (Geographic Information/Geomatics), 2012: Introduction: Welcome to the ISO/TC 211 Harmonized Model Web server. URL: <http://www.isotc211.org/hmmg/HTML/root.html>

- Lake R., Burggraf D., Trninc M., Rae L., 2004: Geography Markup Language: Foundation for the Geo-Web. Wiley (w znacznej części dostępna bezpłatnie)
URL: http://media.wiley.com/product_ancillary/47/04708715/DOWNLOAD/Lake.zip
- Mark D. M., Skupin A., Smith B., 2001: Features, Objects, and other Things: Ontological Distinctions in the Geographic Domain. Spatial Information Theory, Proceedings of COSIT 2001, Springer.
URL: <http://wings.buffalo.edu/philosophy/faculty/smith/articles/COSIT01MSS.pdf>
- Michalak J., 2003a: Modele pojęciowe hydrogeologicznych danych geoprzestrzennych – podstawy metodyczne. Biuletyn PIG – *Hydrogeologia*, z. V, nr 406, monografia.
- Michalak J., 2003b: Podstawy metodyczne i technologiczne infrastruktur geoinformacyjnych. *Roczniki Geomatyki*, t. 1, z. 2. PTIP, Warszawa, monografia.
- Michalak J., 2003c: Geomatics in hydrogeology. *Geological Quarterly*, 47(1): 69-76.
- Michalak J., 2005a: Terminologia polska w zakresie technologii interoperacyjnych w geomatyce. [W:] (red.) Gaździcki J. Internetowy Leksykon Geomatyczny.
URL: <http://www.ptip.org.pl/auto.php?page=Encyclopedia&enc=1>
- Michalak J., 2005b: HGLML – HydroGeoLogical Markup Language – znacznikowy język wymiany geoinformacji hydrogeologicznej. Współczesne Problemy Hydrogeologii, t. XII: 499-504.
- Michalak J., Nawalany M., Sadurski A., (red.) 2011: Schematyzacja warunków hydrogeologicznych na potrzeby numerycznego modelowania przepływu w JCWPd. Wyd. PIG – PIB, Warszawa.
URL: http://www.psh.gov.pl/plik/id,6091,v,artykul_4556.pdf
- Michalak J., 2012: Testowanie roboczych wersji specyfikacji danych tematów załączników II i III INSPIRE. *Roczniki Geomatyki*, t. 10, z. 2, PTIP, Warszawa.
- Nowicki B., Staniszkis W., 2002: Inteligentny system zarządzania wiedzą – prezentacja projektu. [W:] Mat. Konferencji eDemocracy, VI Konf. Miasta w Internecie, Zakopane.
- OMG (Object Management Group), 2001: OMG Unified Modeling Language Specification, version 1.4. OMG Document Repository. URL: <http://cgi.omg.org/docs/formal/01-09-67.pdf>
- OMG (Object Management Group), 2003: Object Management Group, Model Driven Architecture Guide Version 1.0.1 URL: <http://www.omg.org/cgi-bin/doc?omg/03-06-01.pdf>
- OMG (Object Management Group), 2010: Object Constraint Language.
URL: <http://www.omg.org/spec/OCL/2.2>
- Pachelski W., Parzyński Z., 2007: Aspekty metodyczne wykorzystania norm serii ISO 19100 do budowy geodezyjnych składników krajowej infrastruktury danych przestrzennych. *Roczniki Geomatyki*, t.5, z.3, PTIP, Warszawa.
- Peng Z. R., Zhang C., 2004: The roles of geography markup language (GML), scalable vector graphics (SVG), and Web feature service (WFS) specifications in the development of Internet geographic information systems (GIS).
URL: <http://gis.geog.uconn.edu/personal/paper1/journal%20paper/3%202004%20GeographicalSystem1.pdf>
- PKN (Polski Komitet Normalizacyjny), 2010: Informacje podstawowe o PN.
URL: <http://www.pkn.pl/informacje-podstawowe-o-pn>
- Płoski Z., 1999: Słownik Encyklopedyczny – Informatyka. Wyd. Europa, Warszawa.
- Portele C., 2007: OpenGIS Geography Markup Language (GML) Encoding Standard. OpenGIS Standard.
URL: http://portal.opengeospatial.org/files/?artifact_id=20509
- Portele C., 2008a: Mapping UML to GML Application Schemas. Guidelines and Encoding Rules. Interactive Instruments GmbH.
URL: <http://www.interactive-instruments.de/ugas/UGAS-Guidelines-and-Encoding-Rules.pdf>
- Portele C., 2008b: Mapping UML to GML Application Schemas. ShapeChange – Architecture and Description. Interactive Instruments GmbH. URL: <http://www.interactive-instruments.de/ugas/ShapeChange.pdf>
- Portele C., 2012: OGC Geography Markup Language (GML) – Extended schemas and encoding rules. OpenGIS Implementation Standard. URL: https://portal.opengeospatial.org/files/?artifact_id=46568
- Refsgaard J. C., Henriksen H. J., 2004: Modelling guidelines – terminology and guiding principles. *Advances in Water Resources* 27 (2004): 71-82.
- Schmuller J., 2003: UML dla każdego. Helion, Gliwice.
- Ustawa z dnia 12 września 2002 r. o normalizacji.
URL: http://www.pkn.pl/sites/default/files/ustawa_o_normalizacji_2.pdf

- Ustawa z dnia 4 marca 2010 r. o infrastrukturze informacji przestrzennej, Dz.U. 2010 Nr 76, poz. 489.
- Skogan D., 1999: UML as a Schema Language for XML based data Interchange. Materiały konferencji UML'99. URL: <http://xml.coverpages.org/skoganUMLpaper-pdf.gz>
- Sparx System, Enterprise Architect. URL: <http://www.sparxsystems.com.au>
- Subieta K., 1998: Obiektowość w projektowaniu i bazach danych. Akademicka Oficyna Wydawnicza PLJ, Warszawa.
- Subieta K., 1999a: Słownik terminów z zakresu obiektowości. Akademicka Oficyna Wyd. PLJ, Warszawa. URL: http://www.ipipan.waw.pl/~subieta/artykuly/sloownik_obiektowosci/hasla_sloownika.html
- Subieta K., 1999b: Wprowadzenie do obiektowych metodyk projektowania i notacji UML. Jedenasta Górńska Szkoła PTI Szczyrk.
- Tennakoon W., 2003: Visualization of GML data using XSLT. URL: http://www.itc.nl/library/Papers_2003/msc/gim/tennakoon.pdf
- TWG GE (INSPIRE Thematic Working Group – Geology), 2011: D2.8.II.4 INSPIRE Data Specification on Geology – Draft Guidelines. Version 2.9.1. URL: http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_GE_v2.01.pdf
- TWG-CP (INSPIRE Thematic Working Group – Cadastral Parcels), 2010: D2.8.I.6 INSPIRE Data Specification on Cadastral Parcels – Guidelines, version: 3.0.1. URL: http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_CP_v3.0.1.pdf
- WFD WG GIS (Working Group 3.1 – GIS), 2003: Guidance Document No 9 – Implementing the Geographical Information System Elements (GIS) of the Water Framework Directive. Water Framework Directive (WFD) – Common Implementation Strategy. URL: <http://www.ec-gis.org/docs/F2305/GIS-GD.PDF>
- Woolf A., 2009: Enterprise Architect instructions, STFC Rutherford Appleton Laboratory. URL: http://wiki.services.eportal.org/tiki-download_wiki_attachment.php?attId=732
- Zhang C., Peng Z-R., Li W., Day M. J., 2003: GML-Based Interoperable Geographical Databases. URL: <http://www.ucgis.org/summer03/studentpapers/chuanrongzhang.pdf>