

Dyrektywa INSPIRE – źródło zamieszania?

Głównym celem dyrektywy INSPIRE jest profesjonalna ochrona środowiska. By chronić je w sposób rozsądny, efektywny, na różne sposoby i z uwzględnieniem wielu wpływających na siebie elementów, niezbędne jest posiadanie danych o środowisku. Spójrzmy na ten problem tylko od strony zbiorów danych.

Zenon Parzyński

• Kolejne kroki do celu

W skład Unii Europejskiej wchodzi 28 państw o różnej kulturze technicznej, informatycznej, o różnym stanie środowiska. Środowisko jest zróżnicowane nie tylko na terenie Europy, ale również w Stanach Zjednoczonych, Chinach czy Rosji. Problemem w Europie jest jednak brak jednolitości informacji o środowisku wynikający z dużej liczby krajów, a w nich różnych sposobów opisu powierzchni ziemi i jej wykorzystania.

Jako przykład niech posłużą polska BDOT (baza danych obiektów topograficznych) i niemiecki ATKIS (Amtliches

Topographisch-Kartographisches Informationssystem, czyli Urzędowy Topograficzno-Kartograficzny System Informacyjny). W przypadku obu rozwiązań chodzi o topografię w skali 1:10 000, w obu występuje podział na trzy poziomy (tab. 1).

Zawartość tabeli powinna przekonać Państwa, że problem jest poważny – spojrzenie na powierzchnię ziemi z jednego punktu widzenia (bardzo ogólnego) w obu systemach nieco się różni. W aspekcie technicznym problemem jest osiągnięcie interoperacyjności, czyli możliwości łączenia danych z różnych źródeł (aplikacji, baz danych). By dane można było łączyć, muszą być zgodne ich typy – nie da się np. do rastra dodać liczby całkowitej.

• Zgodność typów

Musimy więc posługiwać się tymi samymi typami – czyli łączyć liczby całkowite z całkowitymi, rzeczywiste z rzeczywistymi itp. Ale to jest zdecydowanie za mało! Jeśli dodamy liczbę punktów granicznych jakiejś posiadłości (liczba całkowita) do liczby znajdujących się na niej pokoi (też liczba całkowita), to otrzymamy np. 157. I co z tego? Wynik nawet nie będzie oryginalny, za to na pewno bez sensu. Łączone dane muszą nieść ze sobą to samo znaczenie informacji – ten sam sens. Ale to też jest zdecydowanie za mało!

• Struktura

System obiektowy składa się z obiektów i łączących je relacji. Niektórzy krzyczą głośno, że mają systemy obiektowe, a tak naprawdę mają systemy, w których są zapisane dane o pojedynczych obiektach, bo brak w nich relacji pomiędzy obiektami. Te relacje łączące obiekty stwarzają kolejne problemy. Załóżmy, że chcemy połączyć dane z 28 krajów dotyczące np. działek i budynków. Ale pomiędzy działką i budynkiem zachodzi pewien związek, który w bazie można zapisać w różny sposób.

W górnej części rys. 1 związek jest zapisany w postaci atrybutu. Klasa budynek ma atrybut „nrDziałki”, w którym ma zostać zapisany numer działki, na terenie której budynek się znajduje. W części dolnej ten związek między działką i budynkiem jest zaznaczony znacznie wyraźniej – w postaci asocjacji łączącej

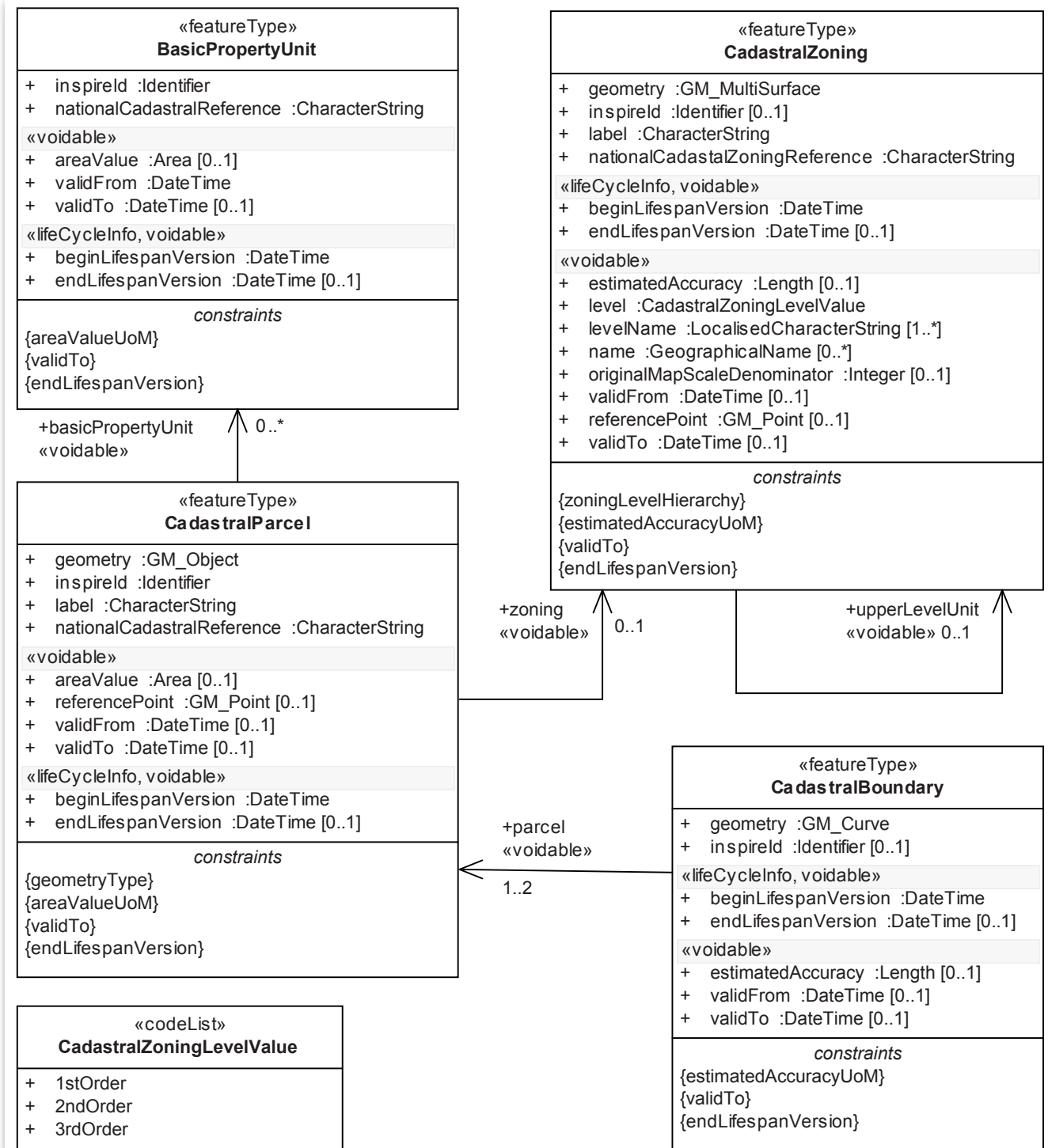
Tabela 1. Obiekty w systemach ATKIS i BDOT – I poziom (najwyższy)

ATKIS	BDOT
	Sieci cieków
	Sieci dróg i kolei
	Sieci uzbrojenia
	Kompleksy pokrycia terenu
Gebäude (Budowle)	Budowle i urzędnictwo
Tatsächliche Nutzung (Rzeczywiste wykorzystanie)	Kompleksy użytkowania terenu
Bauwerke, Einrichtungen und sonstige Angaben (Struktury, obiekty i inne informacje)	Obiekty inne
	Tereny chronione
Gesetzliche Festlegungen, Gebietseinheiten, Kataloge (Wymagania ustawowe, jednostki terenu, katalogi)	Podziały administracyjne i ewidencyjne
Relief (Rzeźba terenu)	Rzeźba terenu

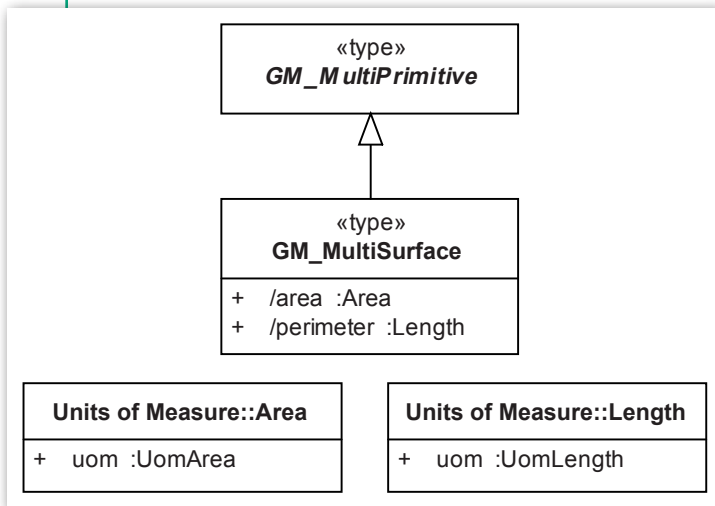
dwie klasy: EGB_Dzialka i EGB_Budynek. Taki zapis relacji oznacza, że będziemy mogli znaleźć działkę, na terenie której stoi dany budynek, ale też wszystkie budynki, które są na danej działce (relacja jest dwustronna).

Jeżeli chcemy na poziomie Unii gromadzić dane o działkach, to trzeba się zdecydować, czy o budynkach też. Jeśli chcemy

Rys. 1. Różny sposób zapisania w modelu związku (relacji) między obiektami (działka i budynek)



Rys. 2. Specyfikacja danych dla działek katastralnych



Rys. 3. Definicja GM_MultiSurface z normy PN-EN ISO 19107 (ISO 19107, 2010)

i o działkach, i o budynkach, to czeka nas kolejna decyzja – jak ma zostać zapisany związek między obiektami i jakie informacje o działkach i budynkach będziemy gromadzić. Jest to niezbędne, by była możliwość osiągnięcia interoperacyjności bez udziału operatora. Danych jest bowiem tak wiele, że jakakolwiek potrzeba ludzkiej interwencji podczas ich agregacji stawia wykonalność całego zadania pod dużym znakiem zapytania. Konieczne jest więc opracowanie ścisłych i dokładnych wytycznych, by uniknąć różnych interpretacji i w konsekwencji – kłopotów.

• Wytyczne (specyfikacje danych)

Przy okazji specyfikacji danych powstaje kilka pytań: w jakiej formie powinny być one ujęte, w jakim języku itp. Podjęto decyzję, że wytyczne zostaną opracowane w języku naturalnym (angielskim i potem będą tłumaczone) oraz formalnym języku „modelowania” – UML (*Unified Modelling Language*). Formalny język ma wadę – trzeba się go nauczyć. Zaletą jest to, że każdy znak ma określone znaczenie, co ogranicza do minimum różną interpretację zapisanego modelu. Użyłem słowa model, ponieważ częścią tych wytycznych (które zostały nazwane Specyfikacjami Danych – *Data Specification* – SD) są modele (w zasadzie minimodele). W UML, czyli formalnym języku, zostały zdefiniowane obiekty, ich właściwości, relacje między obiektami. Dane zapisane w tak zdefiniowanej strukturze będą gromadzone ze wszystkich państw człon-

kowskich, agregowane, zapewne też przekształcane.

W aneksach do dyrektywy INSPIRE określono 34 tematy danych, które mądrzy w Unii uznali za niezbędne do realizacji celu. Do każdego tematu zostały lub zostaną opracowane osobne SD.

Na rys. 2 są przedstawione SD dla działek katastralnych – *cadastral parcels* (SD Działki, 2010). Na

modelu pokazano 4 klasy:

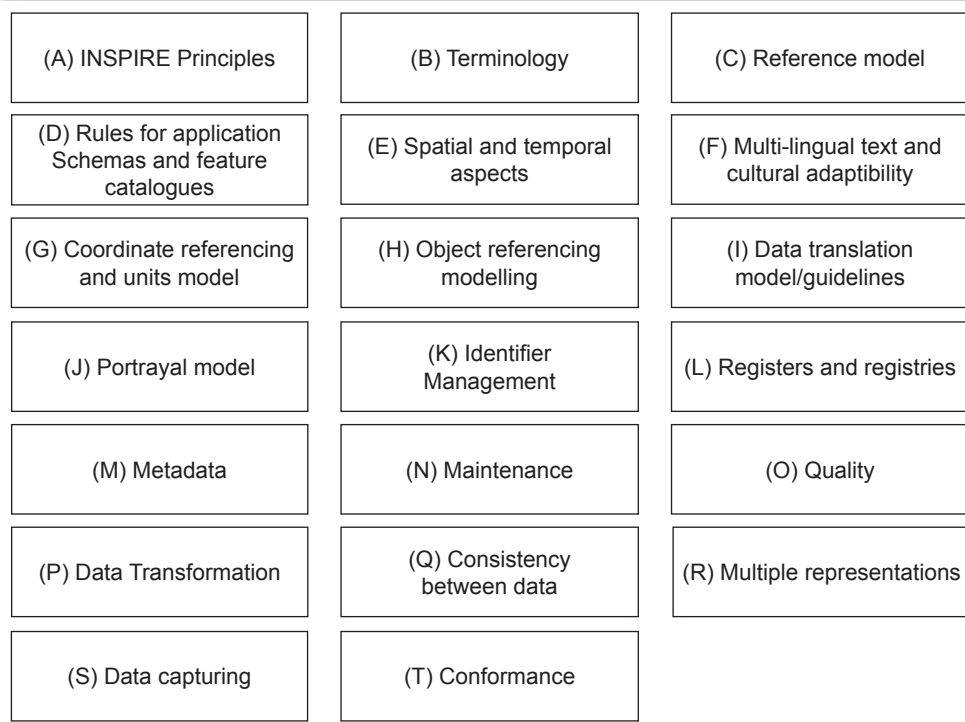
- *CadastralParcel* – reprezentuje działkę,
 - *CadastralBoundary* – reprezentuje granicę działki,
 - *CadastralZoning* – reprezentuje obiekty podziału ewidencyjnego kraju (w innych krajach Unii – katastralnego),
 - *BasicPropertyUnit* – reprezentuje działki definiowane poprzez prawa do nieruchomości (dla krajów, w których działki są określane głównie poprzez prawa, a nie geometrię; w Polsce ta klasa nie jest i raczej nie będzie wykorzystywana).
- Każde państwo członkowskie Unii będzie dostarczać dane opisujące działki zgodnie z modelem pokazanym na rys. 2, gdzie mamy zdefiniowane obiekty (w po-

staci klas), atrybuty, relacje. Ciekawostką jest to, że granica działki jest osobnym obiektem, na który nałożono warunek. Otóż można publikować dane o granicy działki pod warunkiem, że jest znana dokładność położenia obiektów (najczęściej punktów) tworzących tę granicę. Jeśli nie dysponujemy danymi o dokładności granicy, to granicy nie publikujemy. Dlatego relacja między działką a granicą jest opatrzona stereotypem «voidable». Ten stereotyp oznacza, że atrybut lub relacja muszą posiadać wartość (atrybut) lub być w bazie zrealizowane (relacja), a jeśli nie, to należy w bazie umieścić wyjaśnienie, dlaczego brak jest wartości atrybutu lub relacji. Specyfikacja jest zapisana w UML, każdy obiekt, atrybut i relacja są ściśle zdefiniowane. W efekcie łączenie danych zgodnych z modelem pokazanym na rys. 2 dostarczonych z 28 krajów Unii powinno przebiegać bez problemów. Innymi słowy, zbiory dostarczane do Brukseli osiągną interoperacyjność.

• Normy ISO serii 19100

Na rys. 2 jako typy atrybutów są widoczne klasy z norm ISO: *DateTime*, *Area*, *CharacterString*, *GM_Point* czy *GM_MultiSurface*. Są to typy zdefiniowane w normach w postaci klas w języku UML. Na rys. 3 jest przedstawiona definicja (w postaci modelu UML) klasy *GM_MultiSurface*, która składa się z n klas *GM_Surface*.

W jej atrybutach są przechowywane ogólna powierzchnia (*area*) i obwód (*perimeter*) obiektów wchodzących w jej skład.



Rys. 4. Komponenty harmonizacji

Znak „/” znajdujący się przed nazwą atrybutu oznacza, że wartość atrybutu jest wyliczalna (jest sumą powierzchni i objętości obiektów tworzących *GM_MultiSurface*). „Units of Measure” jest nazwą modelu, w którym zostały zdefiniowane wielkości (*Area*, *Length*) i ich jednostki (*UomArea* i *UomLength*).

Norm ISO serii 19100 jest kilkadziesiąt. Ta integracja SD z normami ISO także ma ułatwić osiągnięcie interoperacyjności. Im więcej typów atrybutów zostanie zdefiniowanych z wykorzystaniem klas z norm, tym wymiana danych będzie łatwiejsza. Integracja z normami ISO jest narzucona poprzez wykorzystanie klas z norm jako typy atrybutów w SD.

• Harmonizacja drogą do interoperacyjności

W dyrektywie i różnych dokumentach napisano, że drogą do osiągnięcia interoperacyjności jest harmonizacja, która w przeciwieństwie do interoperacyjności nie jest w zasadzie zdefiniowana. Jej pewne określenie pojawia się w ustawie o infrastrukturze informacji przestrzennej (Ustawa IIP, 2010) – są to różne działania mające doprowadzić do większej spójności zbiorów danych, która to spójność ma przystosować zbiory do wspólnego i łącznego wykorzystania.

A czym praktycznie jest harmonizacja? Według mnie harmonizacją możemy nazwać (w przypadku modelowania struktury bazy danych) proces odwołania się przy opracowywaniu modelu A do klas z innych modeli jako źródła potrzebnych informacji. Zgodnie z rozsądkiem należy oczekiwać, że dane będą gromadzone raz (postulat powtarzający się od lat i od lat aktualny!). Jeśli zatem miejscem gromadzenia, przechowywania, utrzymywania w stanie aktualnym nazw geograficznych jest państwowy rejestr nazw geograficznych (PRNG), a w ewidencji miejscowości, ulic i adresów (EMUiA) jest potrzebna nazwa miejscowości, to harmonizujemy EMUiA z PRNG. Odwołujemy się przy tym do klas/klasy z PRNG w modelu EMUiA (absolutnie nie tworzymy nowej klasy w EMUiA z nową definicją nazw miejscowości!) lub np. wykorzystujemy mechanizm referencji, który został opisany w lipcowym GEODECIE w artykule o modelu podstawowym (Parzyński Z., 2013), lub rozwiązujemy problem jeszcze jakoś inaczej.

Oczywiście proces czy procesy, które ja nazwałem harmonizacją, kiedyś muszą się zakończyć i wtedy można będzie mówić o osiągnięciu pewnego poziomu

mu harmonizacji, np. pomiędzy dwoma zbiorami.

W dokumencie o metodologii tworzenia specyfikacji danych (SD Metodologia, 2007) znajduje się rysunek przedstawiający komponenty harmonizacji (rys. 4). Jak łatwo zauważyć, proces harmonizacji obejmuje bardzo wiele elementów.

• Harmonizacja w GUGiK

Tworząc modele (schematy aplikacyjne) UML, staraliśmy się osiągnąć jak najwyższy stopień harmonizacji. Pytanie: Na ile nam się to udało? Na rys. 5 są pokazane odwołania między schematami UML. Każdy prostokąt reprezentuje jeden schemat aplikacyjny, np. odniesienia przestrzenne. I jeśli w jakimś innym schemacie, np. standardy syt.-wys. (model w rozporządzeniu o standardach technicznych wykonywania geodezyjnych pomiarów sytuacyjnych i wysokościowych...), występuje odwołanie do schematu odniesień przestrzennych, to takie dwa prostokąty (reprezentujące odpowiednie schematy) są na rys. 5 połączone relacją.

Na samym dole jest „osamotniony” schemat zasobu, chociaż w zasadzie powinien być połączony relacją z każdym innym schematem. Ale ponieważ bezpośrednio w schemacie opisującym zasób brak jest jasnych odwołań do innych schematów, to brak też jest odpowiednich relacji. Innymi słowy na rys. 5 jest przedstawiony poziom harmonizacji zdefiniowany tylko w schematach aplikacyjnych. Jest to minimum, ponieważ czasem te odwołania są ukryte. Jeśli w schemacie aplikacyjnym zostały wykorzystane do opisu geometrii klasy z normy PN-EN ISO 19107 (o geometrii i topologii), to automatycznie mamy odwołanie do odniesień przestrzennych, ponieważ w normie 19107 każda klasa jest połączona relacją z układem odniesienia (nie można zapisać współrzędnych bez informacji o układzie odniesienia, w jakim zostały określone). W zasadzie więc prostokąt „Odniesienia przestrzenne” powinien być połączony relacjami z większością pozostałych prostokątów. Musicie sobie Państwo sami odpowiedzieć na pytanie: Jaki stopień harmonizacji został osiągnięty? W sam

REKLAMA



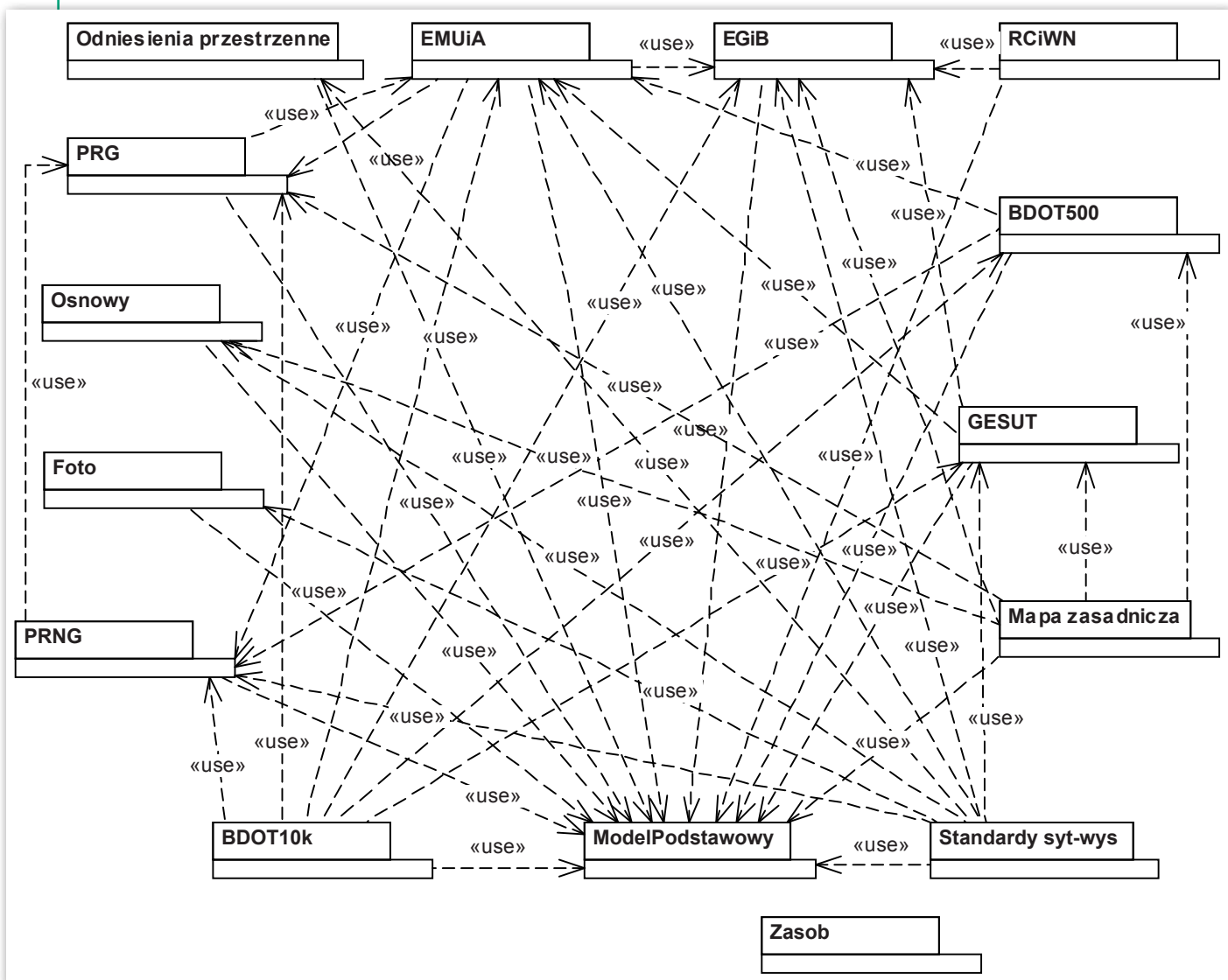
Chcesz czytać schematy UML?

Wykup prenumeratę GEODETY!

Z wrześniowym wydaniem GEODETY wszyscy prenumeratorzy otrzymają bezpłatnie książkę dr. Zenona Parzyńskiego i dr. Agnieszki Chojki „Infrastruktura informacji przestrzennej w UML”

- Ponad 140 stron
- Blisko 130 schematów
- UML krok po kroku





Rys. 5. Harmonizacja w GUGiK

raz, zdecydowanie za mały, za duży, a może jeszcze jakiś pośredni?

• Nie jesteśmy w ognie Europy

Powyższe rozważania podsumowałbym następująco. Po pierwsze, interoperacyjności sprzyja:

- zastosowanie modelowania pojęciowego i zapis tworzonych modeli w UML,
- harmonizacja modeli,
- integracja z normami ISO.

Po drugie, do uzyskania interoperacyjności niezbędny jest zapis i wysyłanie danych w ściśle określonym standardzie – zgodnym ze specyfikacjami danych i zapisanych w ściśle zdefiniowanym języku (dane mają być zapisane w GML).

Opracowując schematy aplikacyjne, staraliśmy się wypełnić przynajmniej w jakimś stopniu zalecenia dyrektywy INSPIRE i różnych dokumentów z nią związanych:

- zastosowaliśmy modelowanie pojęciowe jako metodologię tworzenia schematów aplikacyjnych,

- zapisać te schematy w UML,
- staraliśmy się zharmonizować schematy ze sobą, m.in. poprzez utworzenie i wykorzystanie Modelu Podstawowego,
- zintegrowaliśmy schematy z normami ISO serii 19100.

Twierdzą, że zrobiliśmy w Polsce dużo jak na pierwszy etap implementacji dyrektywy INSPIRE i na pewno nie jesteśmy pod tym względem w ognie Europy.

Czas odpowiedzieć na pytanie zadane w tytule: Czy dyrektywa jest źródłem zamieszania? Na tym etapie na pewno tak. Ale chyba nie jest możliwe przeprowadzenie takiej rewolucji bez początkowego zamieszania. Bo to, co się teraz dzieje, jest rewolucją. Jak bowiem inaczej nazwać wprowadzenie modelowania pojęciowego, UML, GML i wszelkie zmiany, jakie przy okazji zostały dokonane?

dr inż. Zenon Parzyński
pracownik Wydz. Geodezji i Kartografii
Politechniki Warszawskiej

oraz Głównego Urzędu Geodezji i Kartografii, jeden z twórców schematów UML do rozporządzeń z zakresu IIP

Literatura

- Dyrektywa INSPIRE, 2007: Dyrektywa 2007/2/WE Parlamentu Europejskiego i Rady z 14 marca 2007 r. ustanawiająca infrastrukturę informacji przestrzennej we Wspólnocie Europejskiej (INSPIRE), Official Journal of the European Union, L 108, vol. 50, http://www.gugik.gov.pl/gugik/w_pages/w_doc_idx.php?loc=69 (w jęz. polskim);
- ISO 19107, 2010: PN-EN ISO Informacja geograficzna: Schemat przestrzenny, PKN, Warszawa;
- Parzyński Z., 2013: Model pod modelami, GEODETA 7 (218);
- SD Działki, 2010: D2.8.1.6 INSPIRE Data Specification on Cadastral Parcels – Guidelines, http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_CP_v3.0.1.pdf;
- SD Metodologia, 2007: D2.6 Methodology for the development of data specifications, http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.6_v3.0.pdf;
- Ustawa o IIP, 2010: ustawa o infrastrukturze informacji przestrzennej, DzU nr 76, poz. 489.