



POLSKIE  
TOWARZYSTWO  
INFORMACJI  
PRZESTRZENNEJ

# ROCZNIKI 2010 GEOMATYKI

**Metodyka i technologia  
budowy geoserwera tematycznego  
jako komponentu INSPIRE**

Tom VIII  
Zeszyt 3(39)  
Warszawa

*Niniejszy zeszyt specjalny jest opracowaniem monograficznym dr hab. Janusza Michalaka przedstawiającym wyniki badań wykonanych przez Autora w ramach tematu „Eksperymentalny interoperacyjny system geoinformatyczny udostępniania danych dla potrzeb hydrogeologii i spełniający wymagania norm grupy ISO 19100, norm polskich i specyfikacji Open Geospatial Consortium” sfinansowanego przez Ministerstwo Nauki i Szkolnictwa Wyższego. Zawarty w tym opracowaniu wszechstronny przegląd dostępnych technologii budowy geoserwerów zainteresuje z pewnością specjalistów geomatyków. Zaproponowane koncepcje i konkretne rozwiązania stanowią wkład Autora do dyskusji na temat projektowania infrastruktury informacji przestrzennej w Polsce.*

*Jerzy Gaździcki*

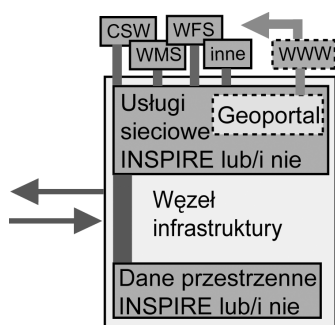
**Dr hab. Janusz Michalak**  
**Uniwersytet Warszawski**  
**Wydział Geologii**  
**J.Michalak@uw.edu.pl**  
**<http://netgis.geo.uw.edu.pl>**

## Spis treści

1. Wstęp .....	11
2. Ogólna koncepcja architektury polskiej IIP .....	14
2.1. Cztery aspekty problematyki architektury polskiej części infrastruktury INSPIRE .....	16
2.2. Podstawy prawne wymagań architektonicznych .....	19
2.3. Przyjęte standardy i specyfikacje techniczne .....	19
2.4. Problematyka harmonizacji standardów .....	21
2.5. Schemat architektoniczny IIP .....	22
2.6. Technologiczne uwarunkowania polskiej części infrastruktury INSPIRE .....	23
2.7. Dwa pozornie oddzielne aspekty technologiczne – dane i usługi .....	24
2.8. Próba syntetycznego ujęcia relacji polskiej IIP do infrastruktury INSPIRE ....	25
3. Założenia technologiczne infrastruktury INSPIRE .....	28
3.1. Relacje dokumentów INSPIRE do normy ISO i specyfikacji OGC .....	28
4. Podstawowe węzły polskiej IIP .....	30
4.1. Model pojęciowy węzła architektury .....	30
4.2. Podstawowe trzy kategorie węzłów .....	32
4.3. Interoperacyjność w zakresie usługi wyszukiwania .....	34
4.4. Geoserwer jako podstawowy element węzła .....	36
4.5. Komponenty geoserwera .....	36
4.6. Komunikacja i interfejsy wewnątrz geoserwera .....	37
4.7. Interfejsy zewnętrzne geoserwera .....	38
4.8. Role konwerterów transformujących dane przestrzenne .....	39
5. Oprogramowanie stosowane w infrastrukturach .....	41
5.1. Kategorie i role systemów geoinformacyjnych .....	41
5.2. Otwarte oprogramowanie .....	43
5.3. Systemy operacyjne .....	45
5.4. Kluczowe systemy przetwarzania geoinformacji .....	47
5.5. Ogólnoinformatyczne oprogramowanie geoserwera .....	48
5.6. Oprogramowanie systemów klienckich .....	49
6. Kryteria doboru oprogramowania geoserwera .....	52
6.1. Heterogeniczność i rozproszenie w infrastrukturach .....	53
6.2. Przenośność i skalowalność oprogramowania .....	53
7. Specyfika geoserwera tematycznego .....	54
7.1. Uwarunkowania wynikające ze specyfikacji danych .....	56
7.2. Dziedziczne modele danych .....	57
7.3. Harmonizacja zobrazowania danych dziedzinowych .....	57
7.4. Szczegółowość i aktualność i danych dziedzinowych .....	58

8. Eksperymentalny geoserwer OakHills .....	59
8.1. Przyjęte założenia wstępne .....	59
8.2. Platforma sprzętowo-systemowa .....	60
8.3. Analiza potrzeb w zakresie oprogramowania serwerowego .....	61
8.4. Architektura geoserwera OakHills .....	62
8.5. Zewnętrzne interfejsy geoserwera .....	64
8.6. Interfejs WWW usługi przeglądania .....	64
8.7. Funkcjonalność interfejsu WWW .....	64
8.8. Wielojęzyczność interfejsów geoserwera .....	69
8.9. Obsługa wielu układów odniesienia .....	69
8.10. Problemy modyfikacji oprogramowania .....	70
9. Testowanie geoserwera .....	71
9.1. Zbiory danych testowych .....	71
9.2. Transformacja danych testowych do języka GML i w tym do modeli INSPIRE .....	72
9.3. Testowanie poprawności funkcji geoserwera .....	73
9.4. Testowanie wydajności w różnych konfiguracjach .....	73
10. Podsumowanie .....	75
Literatura .....	77
A. Cytowane publikacje z czasopism i książki .....	77
B. Specyfikacje i standardy Open Geospatial Consortium .....	80
C. Normy i raporty Komitetu Technicznego ISO/TC211 .....	81
D. Przepisy wykonawcze, specyfikacje i instrukcje techniczne INSPIRE .....	84
E. Dokumentacje i materiały źródłowe komponentów zastosowanego oprogramowania .....	87
Dodatek – Przykład pliku konfiguracyjnego geoserwera OakHills .....	89

## 4. Podstawowe węzły polskiej IIP



**Rys. 5.** Schematyczne przedstawienie elementów pojedynczego węzła i powiązań pomiędzy tymi elementami (objaśnienia w tekście)

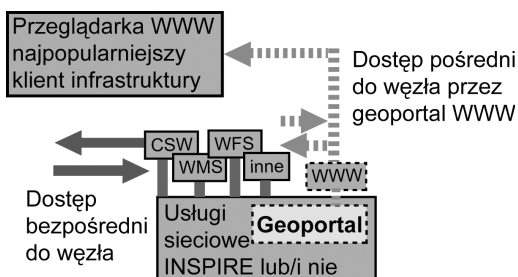
Podstawowymi elementami infrastruktury są węzły (elementy: 1a, 3a, 3b, 3c, 3d, 4a, 4b, 5a i 5b na rysunku 2), często o różnych rolach i w konsekwencji o różnych architekturach. Z uwagi na ogólny wstępny charakter przedstawianego tu schematu architektury IIP budowa poszczególnych węzłów przedstawiona jest w sposób uproszczony (rys. 5).

### 4.1. Model pojęciowy węzła architektury

Z punktu widzenia interoperacyjnego funkcjonowania infrastruktury najistotniejszymi elementami węzła są interfejsy, realizujące usługi udostępniane przez ten węzeł lub realizujące komunikację wzajemną pomiędzy węzłami. Usługa może być realizowana bezpośrednio, czyli wyłącznie w oparciu o protokół obsługiwany przez dany interfejs. W tym przypadku system-klient również musi posiadać odpowiedni dla tego protokołu interfejs. Usługa może być także realizowana pośrednio, poprzez geoinformacyjną witrynę WWW (nazywaną często geoportalem).

Interfejsy usług bezpośrednich to: standardowe CSW, WMS WMTS, WFS WFST, WCS, WPS i inne określone w specyfikacjach OGC i normach ISO. W dyrektywie INSPIRE usługi tego rodzaju odpowiadają w przybliżeniu kategorii usług sieciowych, do których należą usługi wyszukiwania, przeglądania, pobierania i przekształcania. Specyfikacje INSPIRE nakładają dodatkowe wymagania na interfejsy tych usług, np. wielojęzyczność przesyłanych komunikatów lub wymóg stosowania określonych układów odniesienia.

Interfejsy usług pośrednich mogą być ukryte i mogą być dostępne jedynie pośrednio – poprzez geoinformacyjną witrynę WWW (rys. 5). Interfejs ten pełni w węźle szczególną rolę (rys. 6), ponieważ jest on najpopularniejszym sposobem dostępu (pośrednim) do usług udostępnianych przez ten węzeł. Jest to metoda dostępu bardzo wygodna dla szerokiej rzeszy użytkowników, którym wystarczy przeglądanie map w Internecie. W takim przypadku oprogramowaniem-klientem jest zwykła przeglądarka WWW, a interfejsem klienta jest skrypt przesyłany z serwera w trakcie inicjowania połączenia.



**Rys. 6.** Schematyczne przedstawienie zewnętrznych interfejsów węzła infrastruktury. Linie ciągłe – usługi bezpośrednie, linie przerywane – pośrednia realizacja usług poprzez witrynę WWW (pozostałe objaśnienia w tekście)

czenia (często określane jako klient-agent). Bardzo często jest to jedyny sposób dostępu do tych usług, ponieważ pozostałe usługi bezpośrednie są ukryte, zablokowane lub brak jest informacji o bezpośrednim sposobie dostępu – informacje te często nie są opublikowane w serwisie katalogowym usług. Dla zastosowań profesjonalnych dostęp poprzez witrynę WWW nie jest wystarczający – wiele obecnie stosowanych systemów GIS, pełniąc rolę klienta, wykorzystuje dostęp bezpośredni do interfejsów usług i w takim przypadku geoportal jest przeszkodą. Interfejs WWW (geoportal) nie jest w węźle elementem koniecznym do realizacji usług – w takim przypadku można korzystać z usług bezpośrednich, pod warunkiem, że są one opublikowane w katalogu innego węzła lub użytkownik korzysta z klienta (aplikacji) pozwalającego na bezpośredni dostęp do interfejsu CSW lub przez polecenie *GetCapabilities*.

Zewnętrzne interfejsy usług węzłów infrastruktury informacji przestrzennej (rys. 6) są do siebie bardzo podobne. Składają się na nie:

- interfejs serwera katalogowego (CSW) – bezpośrednia usługa wyszukiwania,
- interfejs serwera map (WMS) – bezpośrednia usługa przeglądania,
- interfejs serwera map (WMTS) – uproszczona bezpośrednia usługa przeglądania wykorzystująca technikę segmentacji (*tiling*) i przechowywania segmentów (*cache*),
- interfejs serwera wyróżnień (w tym obiektów przestrzennych – WFS) – bezpośrednia usługa pobierania danych geoprzestrzennych w postaci wektorowej,
- transakcyjny interfejs serwera wyróżnień (w tym obiektów przestrzennych – WFST) – bezpośrednia usługa wymiany danych geoprzestrzennych pomiędzy węzłami lub zdalnej aktualizacji bazy danych węzła,
- interfejsy innych serwerów realizujące usługi bezpośrednie (w tym serwera pokryć WCS) – usługa pobierania danych typu pokrycie (*coverage*), serwera przetwarzania – WPS i innych pod wspólną nazwą OWS (*Open Web Services*),
- interfejs serwera WWW dedykowanego przeglądarkom WWW (nazywanych popularnie przeglądarkami internetowymi) i realizujący usługi pośrednie wyszukiwania, oglądania i pobierania.

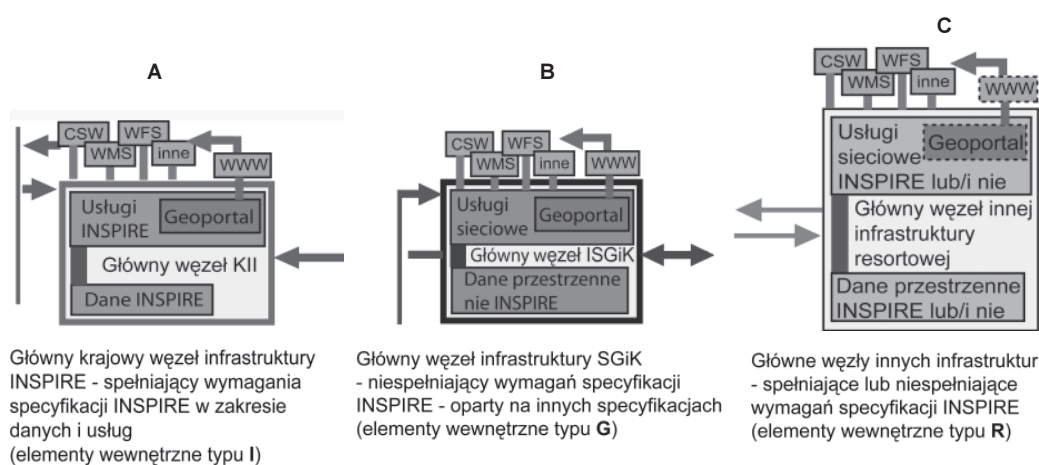
Do wewnętrznych elementów węzła zaliczane są repozytoria zbiorów, bazy danych i metadanych (zbiory danych zarządzane przez systemy zarządzania), systemy przetwarzania danych i metadanych, np. transformujące dane do innych układów odniesienia lub do innych struktur (modeli) lub łączące dane z różnych zbiorów i baz, także uzyskane z innych węzłów, np. w trybie kaskadowym, gdy węzeł występuje w roli klienta wobec innego węzła.

W polskiej infrastrukturze informacji przestrzennej (rys. 2 i 4 – obszar 2, patrz rozdz. 2.8) można wyróżnić trzy podstawowe typy węzłów: węzły INSPIRE, węzły Służby Geodezyjnej i Kartograficznej (SGiK) oraz węzły innych resortów – węzły dziedziczne. Bardziej szczegółowy opis tych trzech typów węzłów znajduje się w następnym rozdziale. Węzły dziedziczne (rys. 7C – typ R) nienależące do infrastruktury INSPIRE są dedykowane danym przestrzennym, dla których instytucjami wiodącymi są instytucje nienależące do SGiK, które mogą utrzymywać dane i usługi zarówno zgodne ze specyfikacjami INSPIRE jak i z nimi niezgodne. Zależy to w głównej mierze od obowiązków i decyzji określonych instytucji wiodących. W przypadku węzła należącego do infrastruktury INSPIRE, zarówno usługi jak i dane geoprzestrzenne muszą być w pełni zgodne ze specyfikacjami INSPIRE (rys. 7A – typ I). Zgodność ta dotyczy obu grup specyfikacji – dla usług i dla danych. Węzły infrastruktury SGiK (rys. 7B – typ G), ze względu na ich szczególną w kraju rolę, przechowują, przetwarzają i udostępniają dane zgodnie z obowiązującymi i projektowanymi w kraju przepisami wykonawczymi. Zakres tych danych określa ustawa o IIP w części zmieniającej Prawo geodezyj-

ne i kartograficzne. Ze względu na specyfikę zadań SGiK wymagania dotyczące zarówno danych jak i usług mogą znacznie odbiegać od wymagań INSPIRE.

#### 4.2. Podstawowe trzy kategorie węzłów

Uwzględniając różne potrzeby i różne wymagania dotyczące poszczególnych trzech podstawowych kategorii węzłów w IIP lub węzłów z nią powiązanych można przedstawić ogólną odpowiednio dobraną strukturę tych węzłów – zarówno ich elementy wewnętrzne, jak i interfejsy zewnętrzne – na przykładzie głównych węzłów poszczególnych infrastruktur krajowych (rys. 7).



**Rys. 7.** Schematyczna budowa trzech kategorii węzłów w IIP lub węzłów z nią powiązanych, na przykładzie głównych węzłów poszczególnych fragmentów infrastruktury krajowej (objaśnienia w tekście)

Poszczególne węzły należące do tych trzech podstawowych kategorii występują na syntetycznym schemacie ogólnym (rys. 2) w różnych rolach i na różnych poziomach hierarchii organizacyjnej. Uwzględniając te role i poziomy organizacyjne w obrębie trzech podstawowych kategorii można sobie wyobrazić bardziej szczegółowe typy węzłów.

- **Węzły infrastruktury INSPIRE nienależące do IIP**, w tym główny europejski węzeł infrastruktury INSPIRE (1a), określany geoportalem INSPIRE z uwagi na pośredni dostęp do jego usług poprzez witrynę WWW. Węzeł taki jest tylko jeden, stanowi on główny punkt dostępu do danych i usług udostępnianych przez całą infrastrukturę. Jego główną i podstawową rolą jest pośredniczenie w wyszukiwaniu danych i usług – jest w pewnym sensie brokerem katalogowym dla innych serwerów katalogowych w głównych węzłach poszczególnych krajów członkowskich.
- **Węzły infrastruktury INSPIRE należące do IIP**, w tym główny węzeł polskiej części infrastruktury INSPIRE (3a), określonej w ustawie o IIP. Według ustaleń IOC-TF (Zespołu Roboczego INSPIRE do Spraw Początkowej Zdolności Operacyjnej), jest to węzeł przeznaczony do bezpośredniej komunikacji głównego europejskiego węzła INSPIRE z polską częścią infrastruktury INSPIRE w zakresie usług katalogowych, z wykorzystaniem polskich zbiorów metadanych INSPIRE.



- **Węzły polskiej części infrastruktury INSPIRE innych resortów niż SGIK (3b).** Ich rola jest podobna do węzła głównego, z tym wyjątkiem, że w zakresie usług katalogowych nie komunikują się bezpośrednio z węzłem europejskim – metadane są przesyłane do głównego węzła krajowego, a ten z kolei komunikuje się z europejskim serwerem katalogowym, w trybie wybranym z trzech przyjętych w dokumentach INSPIRE (*real time, full cached i partially cached*).
- **Inne dowolne węzły polskiej części infrastruktury INSPIRE (3c)** utrzymywane przez osoby trzecie, którym umożliwiono włączenie się do infrastruktury (Ustawa o IIP, art. 4 ust. 1 pkt 3 lit. b). W tym przypadku, w myśl ustawy jako osobą trzecią rozumie się osobę fizyczną, osobę prawną lub jednostkę organizacyjną nieposiadającą osobowości prawnej, niebędącą organem administracji. Rola tych węzłów jest podobna do roli węzłów innych resortów, jednak ze względu na konieczne w tym przypadku dwustronne porozumienie, zasady włączenia się do infrastruktury mogą być określone bardziej szczegółowo w tym porozumieniu. Węzły tej kategorii nie są obowiązkowe.
- **Regionalne i lokalne węzły polskiej części infrastruktury INSPIRE (3d).** Węzły tej kategorii nie są obowiązkowe. Obecnie w Polsce istnieje lub jest tworzonych wiele regionalnych i lokalnych geoportali, głównie udostępniających usługi przeglądania danych przestrzennych z zakresu geodezji i kartografii. Istnieje możliwość przekształcenia ich w przyszłości w węzły polskiej części infrastruktury INSPIRE, pod warunkiem spełnienia wymogów INSPIRE w zakresie udostępnianych przez nie usług. W tym przypadku istotnym problemem może stać się niezgodność tych danych z danymi udostępnianymi przez główny krajowy węzeł INSPIRE. Kłopotliwa może być również redundancja różnych wersji danych.
- **Węzły należące do IIP i nienależące do infrastruktury INSPIRE,** w tym główny węzeł (4a) infrastruktury SGIK. Z uwagi na zadania tej infrastruktury wynikające z polskich przepisów prawnych, ma ona charakter częściowo publiczny i częściowo wewnętrzny. W zakresie roli publicznej udostępnia ona, podobnie jak infrastruktura INSPIRE, dane przestrzenne i związane z nimi usługi, jednak w sposób określony przez inne specyfikacje i instrukcje techniczne niż INSPIRE. Zakres tematyczny danych przestrzennych infrastruktury SGIK w ogólnym zarysie określony jest w ustawie o IIP (art. 23 pkt 2 lit. b). Szczegółowy podział ról węzłów na poszczególnych szczeblach: centralnym, wojewódzkim i powiatowym będzie określony w przepisach wykonawczych do tej ustawy. Rola głównego węzła infrastruktury SGIK w przypadku usługi wyszukiwania jest podobna do roli głównego węzła polskiej części infrastruktury INSPIRE, jednak w przypadku danych powinien on ściśle współpracować z innymi węzłami tej infrastruktury ustawionymi na niższych szczeblach hierarchii – z węzłami wojewódzkimi i prawdopodobnie w pewnych przypadkach z węzłami powiatowymi. Współpraca tego węzła z głównym węzłem polskiej części infrastruktury INSPIRE sprowadza się głównie do dostarczenia danych dla transformacji *off-line* w trybie okresowym (okresowa transformacja wszystkich danych) lub przyrostowym (częste transformacje jedynie w zakresie uzupełnień i poprawek poszczególnych zbiorów danych).
- **Wojewódzkie i powiatowe węzły infrastruktury SGIK (4b).** Zadania tych węzłów są podobne do zadań węzła głównego infrastruktury SGIK, z tą różnicą, że zasięg terytorialny utrzymywanych w nich danych obejmuje powierzchnie odpowiadających im jednostek administracji terenowej. Szczegółowy podział ról węzłów na poszcze-

gólnych szczeblach: centralnym, wojewódzkim i powiatowym będzie określony w przepisach wykonawczych do ustawy o IIP. Węzły te mogą stanowić źródło danych dla transformacji *off-line* w celu zasilenia węzłów regionalnych lub lokalnych polskiej części infrastruktury INSPIRE, jeżeli takie zostaną utworzone. Węzły wojewódzkie współpracują z węzłem głównym infrastruktury SGiK i z węzłami powiatowymi tej infrastruktury w relacjach określonych przez podział terytorialny kraju. Węzły powiatowe współpracują z odpowiadającymi im terytorialnie węzłami wojewódzkimi. Z uwagi na znaczną liczbę powiatów, a także problemy organizacyjne i ekonomiczne budowy węzłów powiatowych, w początkowym okresie tworzenia infrastruktury SGiK należy uwzględnić możliwość *hosting-u* dla węzłów powiatowych w innych węzłach tej infrastruktury, np. w węzłach sąsiednich powiatów lub w odpowiednich dla tych powiatów węzłach wojewódzkich.

- **Główny węzeł innej infrastruktury resortowej (5a)** – nienależący do infrastruktury INSPIRE, ale należący do IIP na zasadach podobnych jak węzły infrastruktury SGiK. Węzeł taki może mieć geoportal, lecz nie jest to konieczne. Dane i usługi tego węzła mogą spełniać wymagania specyfikacji INSPIRE, lecz to również nie jest konieczne.
- **Inne węzły innej infrastruktury resortowej (5b)** – nienależące do infrastruktury INSPIRE, ale należące do IIP na zasadach podobnych jak główny węzeł tej infrastruktury. Ponieważ często inne infrastruktury resortowe zawierają stosunkowo niewielkie zasoby danych przestrzennych, mogą składać się z tylko jednego węzła. Z tego względu inne węzły, poza węzłem głównym, nie są konieczne. Tak samo jak węzeł główny, mogą one mieć geoportal lub nie, a także ich dane i usługi mogą spełniać wymagania specyfikacji INSPIRE, lecz to również nie jest konieczne.
- **Węzły nienależące do IIP i nienależące do infrastruktury INSPIRE (6)**, np. węzły innych infrastruktur resortowych – węzły innych służb publicznych – mogą utrzymywać dane przestrzenne lub nie, ale w pewnych przypadkach może być potrzebne ich współdziałanie z węzłami IIP. Obszar tych zagadnień jest obecnie słabo określony, ale można oczekiwać, że w przyszłości problematyka ta stanie się bardziej istotna.

### 4.3. Interoperacyjność w zakresie usługi wyszukiwania

Usługa wyszukiwania geoinformacji realizowana przez serwery katalogowe jest jedną z podstawowych usług infrastruktury geoinformacyjnej i często jest usługą kluczową. Obecnie tworzy się lub aktualizuje wiele dużych zasobów różnorodnych danych przestrzennych. Wyszukanie przez użytkownika infrastruktury tych danych, które mu są potrzebne i jednocześnie dostatecznie szczegółowe, dokładne i aktualne jest zadaniem bardzo trudnym. Podstawą wyszukiwania w tym przypadku są zasoby metadanych o zbiorach i usługach infrastruktury. Zasoby te są rozproszone po wielu bazach metadanych i ich zawartość zmienia się dynamicznie. Z tego względu interoperacyjność serwerów katalogowych jest zagadnieniem bardzo ważnym i bardzo złożonym. W infrastrukturze INSPIRE przyjęto hierarchiczną organizację tych serwerów, w której najwyższą pozycję zajmuje europejski geoportal INSPIRE, a o jeden poziom niżej znajdują się główne serwery katalogowe poszczególnych krajów członkowskich. W Polsce krajowy serwer katalogowy jest w gestii Głównego Urzędu Geodezji i Kartografii i tylko ten serwer komunikuje się bezpośrednio z serwerem europejskim.

Dostęp europejskiego serwera katalogowego INSPIRE do krajowego serwera katalogowego może być realizowany na trzy sposoby:

- **w czasie rzeczywistym** (*real time*) – każde polecenie dotyczące metadanych o danych (*metadata*) lub metadanych o usługach (*capabilities*) dla obszaru Polski wysyłane przez dowolnego w Europie klienta (programu użytkownika) jest przekierowywane do głównego polskiego serwera katalogowego, a uzyskana odpowiedź przez serwer europejski jest następnie wysyłana przez niego dalej do klienta. Sposób ten gwarantuje aktualność metadanych;
- **z pełnym przechowywaniem** (*full cached*) – wszystkie metadane (*metadata* i *capabilities*) w pełnej postaci są okresowo pobierane przez serwer katalogowy INSPIRE z głównego serwera krajowego. Klient otrzymuje od serwera INSPIRE metadane, które on posiada w swojej bazie i mogą to być dane nie w pełni aktualne;
- **z częściowym przechowywaniem** (*partially cached*) – jest to kombinacja dwóch poprzednich przypadków – europejski serwer katalogowy przechowuje jedynie wybrane elementy metadanych określone w przepisach wykonawczych INSPIRE (*INSPIRE metadata regulation*), które są w tym przypadku wykorzystywane do wyszukiwania danych i usług. Gdy klient potrzebuje metadanych bardziej szczegółowych, których nie ma w serwerze europejskim, uruchamiany jest tryb czasu rzeczywistego, jak to ma miejsce w pierwszym przypadku.

Aby krajowy serwer katalogowy, będący podstawowym elementem głównego węzła polskiej części infrastruktury INSPIRE, mógł poprawnie komunikować się z europejskim serwerem katalogowym, musi mieć zaimplementowany przynajmniej jeden z tych trzech sposobów komunikacji.

Inną bardzo ważną funkcją głównego węzła polskiej części infrastruktury INSPIRE jest przechowywanie i udostępnianie polskich danych przestrzennych w postaci zgodnej ze specyfikacjami danych INSPIRE. Ze względu na konieczność w obecnych krajowych warunkach stosowania transformacji danych *off-line*, dane po transformacji muszą być gdzieś przechowywane w celu realizacji usług przeglądania, pobierania, a także dla innych pozostałych usług INSPIRE. Przechowywanie danych przez ten węzeł jest konieczne, gdy nie ma innych węzłów, które mogłyby pełnić tę funkcję. Dotyczy to następujących przypadków:

- dane przestrzenne w węzłach infrastruktury SGiK, których modele danych (w zakresie struktury i zawartości) różnią się od danych zgodnych ze specyfikacjami INSPIRE;
- analogicznie jak powyżej, ale w odniesieniu do innych infrastruktur lub pojedynczych węzłów dedykowanych danym przestrzennym innym resortów, gdy dla tych danych nie ma odpowiednich węzłów spełniających wymagania specyfikacji INSPIRE;
- z różnych powodów jakaś instytucja nie ma zamiaru tworzenia węzła dla swoich danych z zakresu tematycznego INSPIRE, zarówno zgodnych z wymaganiami INSPIRE jak i niezgodnych.

W tym ostatnim przypadku główny węzeł polskiej części infrastruktury INSPIRE może przejąć obowiązki wynikające z przepisów wykonawczych dyrektywy na zasadach tzw. *hosting-u*. Jednak do obowiązków dysponenta tych danych należy dostarczenie ich w ustalonej formie i następnie ich stała aktualizacja. Należy jednak podkreślić, że funkcję *hosting-u* może pełnić dowolny inny węzeł należący do polskiej części infrastruktury INSPIRE spełniający określone wymagania.

#### 4.4. Geoserwer jako podstawowy element węzła

Pojęcie „węzeł infrastruktury” ma charakter bardziej organizacyjny niż technologiczny. W jednym w sensie fizycznym, systemie komputerowym może znajdować się kilka węzłów, jak również jeden węzeł może być rozmieszczony w kilku systemach fizycznych. Dla określenia organizacyjnego sensu tego terminu można go porównać z definicją terminu „system informacji przestrzennej”, która w Internetowym Leksykonie Geomatycznym PTIP (Gaździcki, 2004) jest określona jako *system pozyskiwania, gromadzenia, weryfikowania, integrowania, analizowania, transferowania i udostępniania danych przestrzennych, w szerokim rozumieniu obejmuje on metody, środki techniczne, w tym sprzęt i oprogramowanie, bazę danych przestrzennych, organizację, zasoby finansowe oraz ludzi zainteresowanych jego funkcjonowaniem*. W tak rozumianym węzle infrastruktury geoserwer jest jego częścią zasadniczą i centralną, składającą się ze sprzętu komputerowego, oprogramowania i zbiorów danych zgromadzonych w bazie lub repozytorium. Z tego względu, zgodnie z zakresem opisywanych prac, problematyka przedstawiona w dalszych rozdziałach będzie dotyczyła wyłącznie geoserwera – jego komponentów wewnętrznych i zewnętrznych powiązań, a także zasad funkcjonowania.

#### 4.5. Komponenty geoserwera

Geoserwer jako główny element węzła infrastruktury, ma za zadanie realizowanie podstawowych jego usług, a w szczególności przeglądania i pobierania geoinformacji. Usługi te są realizowane poprzez łącza internetowe działające w oparciu o podstawową technologię sieciową warstwy transportu TCP/IP (*Transmission Control Protocol/Internet Protocol*) i technologię wyższej warstwy – aplikacji bazującą na protokole HTTP (*HyperText Transfer Protocol*). Z tego względu serwer HTTP jest jednym z podstawowych komponentów geoserwera. Rysunki 27 i 28 (patrz rozdz. 8.4) przedstawiają konkretny przykład podziału geoserwera na komponenty i ich wzajemne wewnętrzne powiązania, jakie tam występują w przypadku systemu opartego na oprogramowaniu MapServer (często także nazywanego „Minnesota Mapserver” lub „UMN Mapserver”). Oprogramowanie to stanowi drugi podstawowy komponent geoserwera. Jego zadaniem jest przetwarzanie danych geoprzestrzennych zgromadzonych w bazie lub w repozytorium do postaci odpowiedniej dla serwera HTTP, zgodnie z otrzymanym od klienta poleceniem protokołu HTTP w formie GET lub POST. Trzecim kluczowym komponentem geoserwera jest system przechowywania danych. W tym przypadku, w zależności od formy i objętości tych danych, stosuje się różne rozwiązania. Może to być zwyczajna baza danych typu relacyjnego (Beynon-Davies, 2003; DuBois, 2004; Connolly, Begg, 2004a, 2004b; Pelikant, 2009), baza obiektowa lub obiektowo-relacyjna (Subieta, 1998; 1999) lub baza z rozszerzeniami geoprzestrzennymi, podnoszącymi istotnie jej efektywność w zastosowaniach geoinformatycznych. Baza danych jest zarządzana przez system, którego zadaniem jest umożliwienie gromadzenia danych, ich utrzymywania w sposób bezpieczny, aktualizacji i selektywnego dostępu do danych. Gdy zbiory danych udostępnianych przez geoserwer nie są wielkie dla ich przechowywania i zarządzania nimi stosuje się repozytoria plików, najczęściej obsługiwane przez system zarządzania plikami systemu operacyjnego.

Wymienione powyżej trzy główne komponenty geoserwera są we wszystkich przypadkach niezbędne. Obok nich mogą występować inne komponenty, w zależności od specyfiki

geoserwera i realizowanych przez niego usług. Usługa wyszukiwania wymaga komponentu w postaci oprogramowania serwerowego dla katalogu metadanych, a usługa przekształcania potrzebuje oprogramowania przetwarzającego dane geoprzestrzenne dedykowanego rodzajom przekształceń jakie mają być przez geoserwer realizowane. W przypadku ograniczenia dostępu do danych i usług, konieczny jest komponent, który będzie realizował procedury identyfikacji, uwierzytelniania i autoryzacji wraz z bazą danych o użytkownikach i ich prawach dostępu. Kolejnym dodatkowym komponentem może być system pobierania opłat za udostępniane dane i usługi. Te wszystkie dodatkowe funkcje geoserwera, łącznie z podstawowymi, tworzą pewnego rodzaju proces złożony z łańcucha usług. Za zainicjowanie tego procesu jest odpowiedzialna usługa wywoływania innych usług określona w dyrektywie INSPIRE (*invoking service*).

Przedstawiona w dalszej części instalacja eksperymentalna geoserwera ma za zadanie realizowanie usług podstawowych – przeglądania i udostępniania danych geoprzestrzennych bez ograniczeń i bez opłat. Z tego względu dodatkowe komponenty nie były uwzględnione w architekturze węzła. Główny wysiłek badawczy był skierowany na możliwie jak najbardziej efektywne i wydajne realizowanie usług podstawowych, a także na funkcjonalny i łatwy w obsłudze interfejs WWW w postaci witryny, dedykowany mniej doświadczonym użytkownikom. Szczegółowy opis funkcjonalności interfejsu WWW zawarty jest w rozdziale 8.6.

#### 4.6. Komunikacja i interfejsy wewnątrz geoserwera

Opisane w poprzednim rozdziale komponenty geoserwera ściśle ze sobą współpracują. Współpraca ta wymaga przepływu poleceń, potwierdzeń i danych. Również w tym przypadku stosowane są interfejsy standardowe – wewnątrz geoserwera. Do podstawowych należy zaliczyć dwa z nich.

Pierwszy to interfejs współdziałania serwera HTTP (realizującego usługi zewnętrzne) z systemem przetwarzania i renderowania lub pakowania danych. W tym przypadku podstawowym interfejsem jest CGI (*Common Gateway Interface*) – znormalizowany interfejs, umożliwiający komunikację pomiędzy oprogramowaniem serwera HTTP, a współdziałającym z nim w tle oprogramowaniem tworzącym dynamiczne elementy udostępnianych stron WWW. Najważniejsze cechy tego interfejsu to niezależność od platformy sprzętowo-systemowej, skalowalność i niezależność od języka programowania. Wadą tej technologii jest konieczność tworzenia nowego oddzielnego procesu na każde polecenie przekazywane przez serwer HTTP, co powoduje duże obciążenie całego systemu geoserwera. Rozwiązaniem tego problemu jest nowa zmodyfikowana wersja tego interfejsu pod nazwą FastCGI. Rozwiązanie to zwiększa szybkość oraz zapewnia lepszą skalowalność aplikacji internetowych opartych o technologię CGI. Różnica samego protokołu w porównaniu z prostym CGI jest niewielkie, jednak FastCGI lepiej zapewnia ciągłość funkcjonowania serwera i obsługę ruchu. Istota tego rozwiązania leży w tym, że w FastCGI uruchomiony jest jeden proces (lub wiele procesów) pracujący w sposób ciągły. Po obsłużeniu polecenia oczekują one na kolejne, nie kończąc pracy, to pozwala na znacznie lepsze gospodarowanie zasobami serwera. Inna ważna zaletą FastCGI jest komunikacja oparta na protokole transportu TCP/IP i dzięki temu oba współpracujące komponenty mogą znajdować się w różnych fizycznych systemach, nawet znacznie oddalonych. Umożliwia to budowanie rozproszonych węzłów infrastruktury, nie tylko na odrębne platformy sprzętowe, ale także na odległe ośrodki komputerowe.

Drugim istotnym interfejsem jest połączenie systemu przetwarzania i renderowania lub pakowania danych w systemem przechowywania tych danych. W przypadku bazy danych rodzaj zastosowanego interfejsu i protokołu zależy od typu tej bazy. Na przykład komunikacja systemu zarządzania bazą danych PostgreSQL z systemami geoinformatycznymi może odbywać się na dwa sposoby. Pierwszy to interfejs zgodny ze standardem OGC i zrealizowany w module PostGIS, który jest dołączany do systemu przetwarzania na etapie kompilacji kodu źródłowego. Interfejs ten obsługuje polecenia bardzo efektywnie, ale jedynie w zakresie prostych wyróżnień zdefiniowanych w zawężonym profilu OGC (*Simple Feature Specification*). Drugi interfejs systemu PostgreSQL jest uniwersalny i przez to nie ma rozszerzeń dotyczących danych geoprzestrzennych. Oba te interfejsy są oparte na protokole transportu TCP/IP, dzięki czemu również umożliwiają rozproszenie węzła infrastruktury przez ulokowanie bazy danych z dala od systemu przetwarzania. W przypadku repozytorium plików komunikacja jest realizowana przez oprogramowanie systemowe, gdy jest to pojedynczy system komputerowy lub w przypadku rozproszenia może być realizowana za pomocą interfejsu opartego na protokole przesyłania plików FTP (*File Transfer Protocol*) działającego na bazie TCP/IP.

Dla interfejsu zewnętrznego w postaci witryny WWW, istotnym elementem pośredniczącym pomiędzy serwerem HTTP i systemem przetwarzania danych i renderowania obrazów jest procesor dynamicznego generowania stron HTML. Spośród różnych rozwiązań najbardziej uniwersalną i efektywną technologią jest PHP (*PHP Hypertext Preprocessor*). Silnik PHP działając po stronie serwera interpretuje skrypty napisane w języku o tej samej nazwie (Melon, 2009; Hayder, 2009), za których pomocą tworzone są zarówno strony HTML, jak i skrypty w języku JavaScript wysyłane następnie do klienta – przeglądarki WWW. Skrypty JavaScript są wykonywane po stronie klienta i m.in. obsługują interakcję użytkownika z serwerem (Powers, 2007; McFarland, 2010), także za pomocą biblioteki jQuery i technologii AJAX (*Asynchronous JavaScript and XML*) (Darie, Brinzarea, Chereches-Tosa, Bucica, 2006; Negrino, Smith, 2010). Połączenie tych czterech technologii (HPH, JavaScript, jQuery i AJAX) pozwala na prowadzenie sesji komunikacji użytkownika z serwerem w trybie asynchronicznym, bez przeładowywania stron i skryptów podczas tej sesji.

#### 4.7. Interfejsy zewnętrzne geoserwera

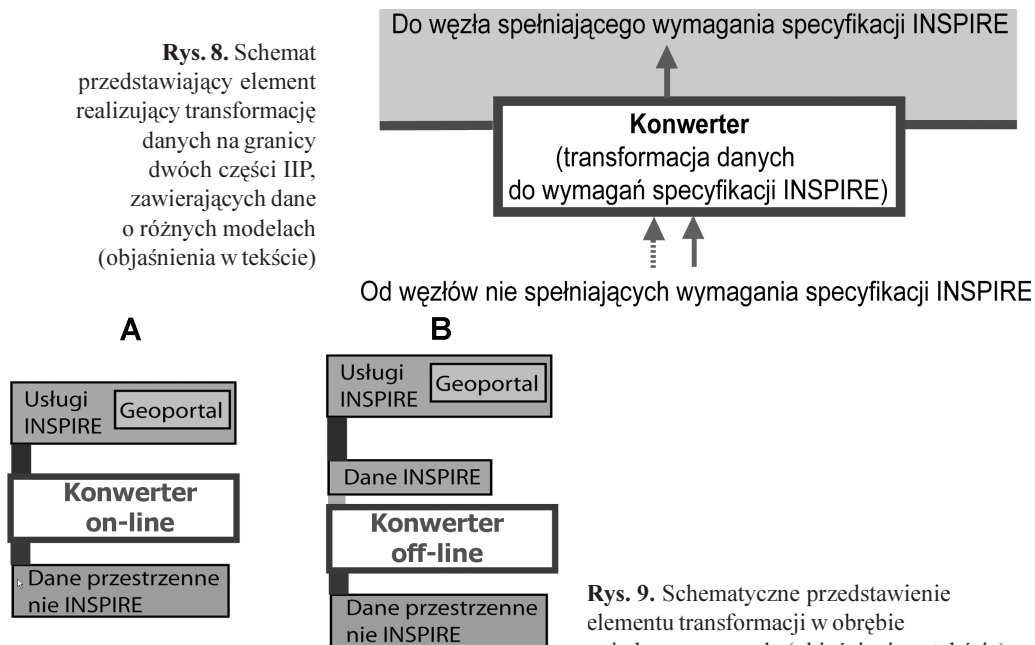
Zadaniem interfejsów zewnętrznych geoserwera jest umożliwienie jego komponentom komunikowanie się z innymi geoserwerami lub innymi elementami infrastruktury, a także realizowanie jego podstawowych zadań z zakresu usług sieciowych, w celu spełnienia potrzeb użytkowników. Problematyka interfejsów zewnętrznych była już przedstawiona w rozdziałach 2.1. i 4.1. W tym miejscu pozostaje do wyjaśnienia jedynie ich aspekt technologiczny. Komunikacja w infrastrukturze pomiędzy jej elementami odbywa się za pomocą protokołu aplikacyjnego HTTP działającego na bazie niższej warstwy transportu TCP/IP. Systemem, który realizuje ten protokół aplikacyjny dla poleceń przychodzących z zewnątrz jest serwer HTTP. Dotyczy to wszystkich usług realizowanych przez geoserwer, np. takich jak witryna WWW i usługi bezpośrednio WMS, WFS, WCS i inne. Jeżeli geoserwer jest serwerem kaskadowym, to pełni również rolę klienta wobec innych elementów infrastruktury. Rola klienta wymaga możliwości wysyłania poleceń protokołu HTTP i to jest najczęściej realizo-

wane za pomocą biblioteki libcURL, którą należy dołączyć na etapie kompilacji oprogramowania przetwarzającego dane geoprzestrzenne.

#### 4.8. Role konwerterów transformujących dane przestrzenne

Szczególną rolę w przedstawianej architekturze infrastruktury informacji przestrzennej pełnią elementy transformacji danych, nazywane także konwerterami. Jeżeli infrastruktura składa się z różnych części zawierających dane o różnych modelach, tak jak to jest w przypadku polskiej IIP, to konwertery muszą wystąpić na wspólnych granicach tych części. Na syntetycznym ogólnym diagramie (rys. 2), a także na rysunku 8, elementy te są przedstawione bardzo ogólnie – jedynie w postaci prostokąta na granicy dwóch części infrastruktury. Wynika to z faktu, że problem transformacji danych jest bardzo skomplikowany, ciągle nie jest rozwiązany, a oprogramowanie konwerterów jest ciągle na etapie eksperymentów. Uwzględniając uwarunkowania przedstawione w rozdziale 2.6 i na rysunku 3 można tu zaproponować kilka opisanych poniżej rozwiązań.

- Transformacja określana w specyfikacjach INSPIRE jako *on-line* (transformacja dynamiczna) realizowana w typowym węźle (rys. 9 – A). Takie rozwiązanie, jak było to przedstawione wcześniej w rozdziale 2.6, jest możliwe w przypadku stosunkowo małej różnicy modeli danych (rys. 3) i tylko wtedy, gdy pojedynczy zbiór danych lub pojedyncza baza danych jest transformowana do pojedynczego zbioru.
- Transformacja określana w specyfikacjach INSPIRE jako *off-line* realizowana okresowo w typowym węźle (rys. 9 – B) w trybie automatycznym, tzn. z minimalnym udziałem operatora. Możliwość realizacji takiego rozwiązania jest tak samo ograniczona jak w przypadku poprzednim.
- Transformacja *off-line* realizowana okresowo w typowym węźle (rys. 9 – B) w try-



bie nadzorowanym – jednak ze znacznym udziałem operatora. Takie rozwiązanie może okazać się bardzo pracochłonne, szczególnie, gdy potrzebna jest ręczna edycja dużych zbiorów danych.

- Transformacja określana w specyfikacjach INSPIRE jako *on-line* (transformacja dynamiczna) realizowana w wyspecjalizowanym do tego celu węźle. W takim przypadku prostokąt symbolizujący konwerter (rys. 8) powinien być zastąpiony serwerem usług zgodnych ze standardem WPS. Możliwość realizacji takiego rozwiązania jest tak samo ograniczona jak w przypadku pierwszym.
- Transformacja określana w specyfikacjach INSPIRE jako *off-line* realizowana okresowo na zasadach *outsourcing-u*. Pod tym określeniem rozumie się wykorzystywanie zasobów zewnętrznych – zlecenie wyspecjalizowanym podmiotom zewnętrznym procesów niezbędnych dla przekształcenia zawartości i struktury zbiorów danych przestrzennych, stanowiących zasób określonego węzła infrastruktury lub konkretnej bazy danych albo zestawu zbiorów danych. Na obecnym etapie rozwoju polskiej IIP takie rozwiązanie jest najbardziej racjonalne i prawdopodobnie stan taki będzie trwał jeszcze przez kilka najbliższych lat.

Jak wspomniano wcześniej, problem transformacji danych jest bardzo skomplikowany, ciągle nie jest dostatecznie poprawnie rozwiązany. Zagadnienia te stanowią główną przeszkodę, na jaką napotyka się podczas prac projektowych i wykonawczych dotyczących polskiej IIP. Z tego względu potrzebne są pilne prace badawcze – studialne, eksperymentalne, prototypowe i wdrożeniowe, bo bez nich dalszy rozwój polskiej IIP, zgodny z obowiązującymi i projektowanymi regulacjami prawnymi, krajowymi i europejskimi, może okazać się niezwykle trudny.

W przypadku, gdy transformacja danych w trybie *off-line* dokonywana jest w konkretnym węźle i w węźle tym struktura podstawowych danych nie jest zgodna ze specyfikacjami INSPIRE, a węzeł ten ma udostępniać usługi INSPIRE (rys. 9 – B), czyli zwracając dane zgodne ze specyfikacjami INSPIRE, wymagane są podwojone elementy dla danych (bazy, repozytoria i inne). Jedne z nich są przeznaczone dla danych podstawowych, a drugie dla danych po transformacji i w rezultacie zgodnych ze specyfikacjami INSPIRE.