



ISSN 1731-5522

POLSKIE
TOWARZYSTWO
INFORMACJI
PRZESTRZENNEJ

ROCZNIKI **2016**
G E O M A T Y K I

Tom XIV
Zeszyt 1(71)
Warszawa

ROCZNIKI GEOMATYKI



POLISH
ASSOCIATION
FOR SPATIAL
INFORMATION

ANNALS 2016
OF GEOMATICS

**Methodology of transformation
of environmental geospatial data
to the INSPIRE schemas**

(The case of hydrogeology)

**Volume XIV
Number 1(71)
Warsaw**



**POLSKIE
TOWARZYSTWO
INFORMACJI
PRZESTRZENNEJ**

ROCZNIKI 2016

GEOMATYKI

**Metodyka transformacji
środowiskowych danych geoprzestrzennych
do schematów INSPIRE**

(Na przykładzie hydrogeologii)

**Tom XIV
Zeszyt 1(71)
Warszawa**

ZESPÓŁ REDAKCYJNY EDITORIAL STAFF

Redaktor Naczelny	Editor-in-Chief	Prof. dr hab. inż. Jerzy Gaździcki
Sekretarz Redakcji	Managing Editor	Mgr inż. Ewa Musiał
Redaktor	Editor	Dr hab. Marek Baranowski, prof. IGIK
Redaktor	Editor	Dr hab. Janusz Michałak
Redaktor statystyczny	Statistical Editor	Dr hab. inż. Anna Barańska
Redaktor językowy	Language Editor	Mgr Maria Wojciechowska
Tłumaczenie	Translation	Mgr inż. Jacek Domański

Redaktorzy tematyczni Thematic editors

Dr hab. inż. Jerzy Balcerzak, prof. em. Politechniki Warszawskiej
Dr hab. inż. Joanna Bac-Bronowicz, Politechnika Wroclawska
Prof. dr hab. inż. Konrad Eckes, Akademia Górniczo-Hutnicza
Prof. dr hab. Marek Graniczny, Państwowy Instytut Geologiczny
Prof. dr hab. inż. Józef Jachimski, prof. em. Akademii Górniczo-Hutniczej
Dr inż. Dariusz Korpetta, Szkoła Główna Gospodarstwa Wiejskiego
Dr hab. Jacek Kozak, prof. Uniwersytetu Jagiellońskiego
Prof. dr hab. inż. Andrzej Stateczny, Akademia Morska w Szczecinie

RADA NAUKOWA EDITORIAL ADVISORY PANEL

Alessandro Annoni, Head of the Digital Earth and Reference Data Unit, JRC, Ispra
Prof. dr hab. inż. Stanisław Białousz, Politechnika Warszawska
Prof. dr ir. Theo Bogaerts, prof. em., Delft University of Technology
Lukasz Grus, Ph.D., Wageningen University
Francis Harvey, Ph.D. University of Minnesota
Prof. Mike Jackson, Ph.D., University of Nottingham
Dr hab. Adam Linsenbarth, prof. em. Instytutu Geodezji i Kartografii, Warszawa
Prof. dr hab. inż. Bogdan Ney, Polska Akademia Nauk
Prof. dr hab. inż. Heronim Olenderek, Szkoła Główna Gospodarstwa Wiejskiego
Prof. dr hab. inż. Stanisław Oszczak, prof. em. Uniwersytetu Warmińsko-Mazurskiego
Prof. dr Erik Stubkjær, prof. em., Aalborg University



Published under the Creative Commons Attribution 3.0 Unported License:
<https://creativecommons.org/licenses/by/3.0/>

Kwartalnik Naukowy. Wersją pierwotną jest wersja drukowana p-ISSN 1731-5522.
Dodatkowo artykuły i monografie udostępniane są nieodpłatnie w wersji elektronicznej
e-ISSN 2449-8963, rg.ptip.org.pl

Ogólne informacje o publikowaniu w Rocznikach Geomatyki:
<http://rg.ptip.org.pl/index.php/rg/about/submissions#authorGuidelines>

Polskie Towarzystwo Informacji Przestrzennej
Polish Association for Spatial Information
ul. rtm. W. Pileckiego 112/5, 02-781 Warszawa
e-mail: ptip@ptip.org.pl <http://www.ptip.org.pl>

Wydawca: WIEŚ JUTRA Sp. z o.o.
Redakcja techniczna i projekt układu graficznego: Barbara Werbanowska
Nakład: 100 egz. Objętość: ark. wyd. 17,5 (str. 1–176) + 16 stron wkładki barwnej

Niniejsza monografia powstała jako rezultat projektu badawczego Ministerstwa Nauki i Szkolnictwa Wyższego oraz Narodowego Centrum Nauki pod tytułem „Metodyka transformacji polskich zbiorów geoprzestrzennych danych hydrogeologicznych do modeli określonych w specyfikacji danych Dyrektywy INSPIRE”. Przedstawiona w niej problematyka jest obecnie przedmiotem żywego zainteresowania instytucji, firm i specjalistów zajmujących się tworzeniem, utrzymaniem i użytkowaniem infrastruktury informacji przestrzennej w Unii Europejskiej, a co za tym idzie, również w Polsce. Zgodnie z obowiązującym prawem wspólnotowym i krajowym w określonych terminach mają być wykonane prace, wymagające specjalistycznej, trudno dostępnej wiedzy, mającej często charakter pionierski.

Tematyka monografii dra hab. Janusza Michalaka jest zgodna z aktualnymi potrzebami i wyzwaniem wiążącymi się z infrastrukturą informacji przestrzennej, a jej ujęcie ma walory zarówno naukowe, jak też w znacznym stopniu dydaktyczne. Na uwagę zasługuje podanie bogatej literatury przedmiotu.

Jerzy Gaździcki

Dr hab. Janusz Michalak
Laboratorium Geomatyki
janusz.michalak@onet.eu

Monografia przedstawia wyniki prac badawczych projektu sfinansowanego przez
Ministerstwo Nauki i Szkolnictwa Wyższego i Narodowe Centrum Nauki
o numerze N N525 253640 pod tytułem:
„Metodyka transformacji polskich zbiorów geoprzestrzennych danych hydrogeologicznych
do modeli określonych w specyfikacji danych Dyrektywy INSPIRE”.

Metodyka transformacji środowiskowych danych geoprzestrzennych do schematów INSPIRE

(Na przykładzie hydrogeologii)

Słowa kluczowe: dane geoprzestrzenne dane hydrogeologiczne, transformacja, XSLT, infrastruktura INSPIRE

Streszczenie

Obecnie w Polsce trwają prace nad budową polskiej części europejskiej infrastruktury informacji geoprzestrzennej w ramach inicjatywy INSPIRE. Jednym z kluczowych i jednocześnie bardzo trudnym zadaniem jest utworzenie zbiorów danych spełniających wymagania dokumentów INSPIRE jako rezultat przekształcenia zbiorów krajowych w ich obecnej postaci. Problem ten dotyczy wszystkich dziedzinowych tematów wyszczególnionych w załącznikach dyrektywy, w tym także danych hydrogeologicznych.

Monografia ta przedstawia wyniki prac wykonanych w ramach projektu badawczego, którego celem było opracowanie technologii i metodyki transformacji krajowych danych hydrogeologicznych do formy i struktury określonej w specyfikacji INSPIRE dotyczącej tego tematu. Podstawą tej transformacji jest koncepcja technologiczna określana akronimem ETL: Extract – Transform – Load, co w skrócie sprowadza się do procesu złożonego z trzech faz: 1 – pobierz dane ze źródła i zapisz je w formie znacznikowej (XML), 2 – przekształć je do określonej nowej treści i formy (również XML) za pomocą procesora XSLT (Extensible Stylesheet Language Transformations), 3 – umieść uzyskane wyniki w repozytorium lub bazie danych.

Przyjęte tu rozwiązania technologiczne i metodyczne nie ograniczają się jedynie do danych hydrogeologicznych, które były głównym przedmiotem analiz i testów. Uzyskane wyniki mogą bez istotnych modyfikacji być zastosowane do geoinformacji z innych dziedzin, w tym szczególnie z zakresu dyscyplin związanych ze środowiskiem przyrodniczym.

Przedstawiana w monografii problematyka składa się z wielu aspektów, a do najważniejszych z nich należą:

- Stan obecny krajowych danych hydrogeologicznych – ich zawartość, struktura, forma zapisu, sposób przechowywania i zasady udostępniania.
- Wymagania określone w dokumentach INSPIRE, w tym w aktach prawnych, specyfikacjach tematycznych i technicznych, a także przyjęte w infrastrukturze reguły organizacyjne.
- Przyjęte międzynarodowe normy i standardy dotyczące geoinformacji, interoperacyjności systemów geoinformatycznych ze szczególnym uwzględnieniem modeli danych, które mają bezpośrednie zastosowanie do danych hydrogeologicznych.
- Standardowe rozwiązania technologiczne dotyczące przetwarzania danych zapisanych w formie znacznikowej (XML), a w szczególności języków XPointer, XPath, XLink i XQuery, ponieważ są bezpośrednio związane z technologią XSLT.
- Metodyka i służące jej narzędzia przechowywania i pobierania danych w formie znacznikowej, a w tym bazy danych dedykowane takiej formie zapisu.
- Specjalistyczne systemy narzędziowe przeznaczone do przekształcania danych geoprzestrzennych, zarówno komercyjne jak i typu Open Source.
- Oprogramowanie wspomagające procesy transformacji, jak na przykład edytory XML i przeglądarki danych geoprzestrzennych zapisanych w języku GML.

Wykonane w ramach tego projektu prace analityczne i testowe wykazały, że realizacja zadań z zakresu przedstawionej tu transformacji danych hydrogeologicznych jest w pełni wykonalna, jednak nie wszystkie szczegółowe operacje mogą na tym etapie być wykonane automatycznie bez interwencji manualnej. Z tego względu potrzebne są dalsze prace badawcze, które pozwolą w pełni zautomatyzować proces transformacji, a bezpośredni udział człowieka będzie sprowadzał się do wyznaczenia zadań przetwarzania wsadowego i do weryfikacji uzyskanych wyników.

Methodology of transformation of environmental geospatial data to the INSPIRE schemas

(The case of hydrogeology)

Keywords: geospatial data, hydrogeological data, transformation, XSLT, INSPIRE infrastructure

Abstract

At present, teams of experts are carrying out works on development of the Polish part of the European geospatial information infrastructure under the INSPIRE initiative. One of the key and very difficult tasks is to create datasets that meet the requirements of INSPIRE documents as a result of transformation of national data sets from their present form. This problem applies to all domain-specific themes listed in the Annexes of the Directive, including hydrogeological data.

This monograph presents the results of works performed within the framework of a research project which aim was to develop technology and methodology of the national hydrogeological data transformation to the form and structure specified in INSPIRE guideline documents related to this theme. The basis for this transformation is the technological concept referred to as the ETL: Extract – Transform – Load, which basically boils down to a process consisting of three phases: 1 – download the data from the source and save it in a tagged form (XML), 2 – convert it to a given new content and forms (including XML) using an XSLT processor (Extensible Stylesheet Language Transformations), 3 – put the results in a repository or a database.

Assumed methodological and technological solutions are not limited to hydrogeological data, being the main subject of analysis and testing. The results can – without significant modifications – be applied to geospatial data from other fields, particularly in the scope of disciplines related with the natural environment.

Problems presented in this monograph concern a number of aspects; the most important of them include:

- The present state of national hydrogeological data – their content, structure, forms of encoding, the way of storage and the rules of sharing.
- The requirements of the INSPIRE documents, including legal acts, thematic and technical specifications, as well as the adopted organizational rules in the infrastructure.
- Accepted international standards for geoinformation, geospatial systems, their interoperability with particular emphasis on data models, which are directly applicable to the hydrogeology.
- The standard technological solutions for data processing in the tagged encoding (XML), and particularly the use of XPointer, XPath, XLink, and XQuery languages because they are directly related to the XSLT technology.
- The methodology and its tools for storing and retrieving data in the form of tagged encoding, and databases dedicated to this form of data storage.
- Specialized tool systems, both commercial and Open Source, designed to transformation of geospatial data.
- Software tools for supporting transformation processes, such as XML editors and GML data viewers.

Analytical and test works performed in the frame of this project have shown that transformation of presented hydrogeological data is fully feasible, but not all the detailed operations can be automatically performed at this stage without manual intervention. Therefore, the need for further research exists, that would fully automate the process of transformation and direct human intervention would be required only to determine the tasks for batch processing and to check the results.

Spis treści

1. Wstęp	11
2. Zadania w zakresie hydrogeologii wynikające z dyrektywy INSPIRE	15
2.1. Międzynarodowa Mapa Hydrogeologiczna Europy	19
2.2. Udostępnianie krajowych danych hydrogeologicznych	21
2.3. Treść i forma udostępnianych danych	21
2.4. Krajowe dane hydrogeologiczne w środowisku infrastruktury	23
2.4.1. Testowanie serwera e-PSH	24
2.4.2. Testowanie serwera SPD-PSH	26
2.4.3. Porównanie wyników testów	29
3. Podstawy technologiczne infrastruktury INSPIRE	33
3.1. Usługi geoinformacyjne infrastruktury	33
3.2. Proces opracowania modelu danych hydrogeologicznych INSPIRE	36
4. Krajowe dane hydrogeologiczne	39
4.1. Mapa hydrogeologiczna Polski – MHP	39
4.1.1. Użytkowy Poziom Wodonośny	40
4.1.2. Pierwszy Poziom Wodonośny – Występowanie i Hydrodynamika	40
4.1.3. Pierwszy Poziom Wodonośny – Wrażliwość na Zanieczyszczenie i Jakość Wód	41
4.2. Bank HYDRO	42
4.3. Baza danych Monitoring Wód Podziemnych	44
4.4. Główne Zbiorniki Wód Podziemnych	45
4.5. Jednolite Części Wód Podziemnych	46
4.6. Zasoby dyspozycyjne	47
4.7. Pobory	48
4.8. Obszary zagrożone podtopieniami	48
4.9. Wody podziemne zaliczone do kopalin	48
4.10. Metadane	49
4.11. Inne dane wykorzystywane w hydrogeologii	50
5. Przegląd modeli danych hydrogeologicznych	55
5.1. Porównanie modelu INSPIRE z modelami GWML i WaterML	56
5.1.1. Język GWML 1.0	56
5.1.2. Prace nad językiem GWML 2.0	58
5.1.3. Wody podziemne w języku WaterML – moduł <i>HY_Features</i>	63
5.2. Projekt języka HydroGeoML	66
6. Specyfikacja i model danych INSPIRE dla podtematu Hydrogeologia	71
6.1. Elementy modelu INSPIRE	72
6.2. Główne problemy technologiczne z zakresu danych hydrogeologicznych	76
6.3. Modele danych krajowych jako rozszerzenia modelu INSPIRE	79

7. Analiza różnic pomiędzy modelem krajowym i INSPIRE	81
7.1. Mapowanie elementów modeli	85
8. Podstawy metodyczne transformacji	87
8.1. Przegląd metod	90
9. Transformacje XSLT	93
9.1. Podstawy metodyczne transformacji	93
9.1.1. Technologia XSLT	94
9.1.2. Język XPath	96
9.1.3. Języki XPointer i XLink	97
9.1.4. Język XQuery	99
9.2. Procesory XSLT	100
9.3. Składnia i semantyka arkuszy XSLT	101
10. Bazy danych XML	103
10.1. BaseX	105
10.2. Podsumowanie	109
11. Przegląd narzędzi do transformacji	111
11.1. Oprogramowanie FME (Feature Manipulation Engine)	111
11.2. Oprogramowanie HUMBOLDT Alignment Editor	119
11.3. Oprogramowanie GeoKettle	126
11.4. Oprogramowanie Stetl	128
11.5. Porównanie narzędzi do transformacji	134
11.6. Oprogramowanie narzędziowe wspomagające transformacje	135
12. Transformacja danych hydrogeologicznych	137
12.1. Dane wykorzystane do testowania	139
12.2. Przekształcenie do płaskiego zapisu znacznikowego	140
12.3. Testowe arkusze XSLT	142
12.4. Analiza uzyskanych wyników	150
12.5. Główne problemy transformacji danych hydrogeologicznych	157
13. Podsumowanie	165
Literatura	167

1. Wstęp

Monografia została napisana po zakończeniu prac projektu badawczego, którego celem było opracowanie praktycznej metodyki przekształcenia obecnych krajowych danych geo-przestrzennych z zakresu nauk o środowisku do formy odpowiedniej dla powszechnego udostępniania ich z zastosowaniem nowoczesnych technologii teleinformatycznych. Inspiracją do podjęcia tego tematu badawczego była pilna potrzeba opracowania poprawnych i efektywnych metod transformowania danych pomiędzy różnymi ich modelami, ze szczególnym uwzględnieniem wymagań specyfikacji danych i usług zawartych w dokumentach implementacyjnych dyrektywy INSPIRE (EP&CEU, 2007). Dyrektywa ta została uchwalona przez Parlament Europejski i zatwierdzona przez Radę Unii Europejskiej w roku 2007 i jej celem jest stworzenie Europejskiej Infrastruktury Informacji Przestrzennej (ESDI – *European Spatial Data Infrastructure*) głównie dedykowanej problematyce środowiska. Przez kolejne lata opracowywano szczegółowe akty prawne niższego rzędu (przepisy wykonawcze – *implementing rules*) zatwierdzane przez Przewodniczącego Komisji Europejskiej, a także szczegółowe specyfikacje techniczne, których zadaniem jest precyzyjne określenie zakresu udostępnianych danych, ich formy i rozwiązań technologicznych do ich rozpowszechniania. W dokumentach tych zawarty jest między innymi harmonogram czasowy określający terminy wykonania poszczególnych zadań budowy tej infrastruktury. Wśród tych terminów są w tym przypadku szczególnie ważne dwa określające:

- 1) dostępność nowoutworzonych i znacznie przeorganizowanych danych z zakresu tematów aneksów II i III – wyznaczona na 21 października 2015 roku.
- 2) dostępność pozostałych danych, utworzonych wcześniej, z tego samego zakresu tematycznego – wyznaczona na 21 października 2020 roku.

Jeszcze nie dawno, gdy planowano prace tego projektu, terminy te wydawały się odległe, jednak pierwszy z tych terminów minął już po roku od daty zakończenia prac projektu. Z tego względu autor ma nadzieję, że monografia pomoże zespołom zajmującym się transformacją danych wykonać spoczywające na nich zadania poprawnie i terminowo i że w rezultacie organy wiodące dla tych poszczególnych tematów będą mogły przekazać informacje do Komisji Europejskiej o poprawnym i terminowym wykonaniu tego etapu.

Inną dyrektywą, która obejmuje także problematykę informacji przestrzennej, w tym przypadku dotyczącej wód powierzchniowych i podziemnych, jest Ramowa Dyrektywa Wodna, jej przepisy wykonawcze i specyfikacje techniczne (EC, 2009; WFD WG GIS, 2003). W zakresie udostępniania informacji przestrzennej obok zadań wynikających z prawodawstwa unijnego istnieje wiele aktów prawnych krajowych, które nakładają różne obowiązki na ośrodki krajowe. Ministerstwo Środowiska opublikowało istotne w tym przypadku dwa dokumenty, jeden w zakresie hydrogeologii (MŚ, 2008a) i drugi w zakresie kartografii geologicznej (MŚ, 2008b), które określają politykę resortu na lata 2008-2015. W dokumentach tych sprawa geoinformacji geologicznej i hydrogeologicznej jest po raz pierwszy wyraźnie wyeksponowana.

Realizacja powyższych wymagań określonych w prawodawstwie unijnym i krajowym stawia poważne wyzwania technologiczne, zarówno w zakresie struktury (modeli) danych, jak i metod ich udostępniania (usług sieciowych). Jednym z trudniejszych zadań jest przekształcenie istniejących i nowo tworzonych danych w różnych modelach i w różnych formach zapisu do modeli i form standardowych. Monografia jest poświęcona właśnie tym zagadnieniom. Opisuje skuteczne zastosowanie metodyki opartej na technologii ETL-XSLT (*Extract, Transform, Load – Extensible Stylesheet Language Transformations*) (Van den Broecke, 2011; 2013) do krajowych zbiorów danych hydrogeologicznych. Trzeba tu jednak zaznaczyć, że jest to metodyka uniwersalna, która bez większych problemów i z niewielkimi modyfikacjami może być zastosowana do danych z innych dziedzin związanych ze środowiskiem. W ramach prac projektu przeprowadzono wiele testów z zastosowaniem specjalistycznego oprogramowania zarówno komercyjnego, jak i dostępnego na prawach *Open Source*. Szczegółowy tego opis znajduje się w dalszych rozdziałach monografii. Prace nad metodyką transformacji były realizowane w powiązaniu z innymi zbliżonymi problemowo pracami, do których trzeba zaliczyć opracowanie modelu danych hydrogeologicznych w języku UML (*Unified Modeling Language*) (Booch, Rumbaugh, Jackbson, 2002) stanowiącego podstawę specyfikacji INSPIRE dla tego podtematu, a także ich transformację do struktur języka XML (*Extensible Markup Language*) zapisanych z zastosowaniem języka schematów XSD (*XML Schema Definition*) (Carlson, 2001; Skogan, 1999). Inne zagadnienie związane z tematem monografii to opracowanie podstaw ontologicznych i semantycznych (Collins, Clark, 2004; Ma, 2011; Goczyła, 2011; Mark, Skupin, Smith, 2001; Smith, 2001) nowej wersji języka GWML (*GroundWater Markup Language*) (Boisvert, Brodaric, 2007; 2012). Odrebnym zagadnieniem związanym z INSPIRE jest określenie w jakim stopniu i w jaki sposób krajowe dane hydrogeologiczne spełniają wymagania specyfikacji INSPIRE jako źródło, bez względu na to w jakiej formie obecnie istnieją. Prace nad tym zagadnieniem były prowadzone wspólnie z zespołem Państwowego Instytutu Geologicznego, a wyniki zostały przekazane w postaci obszernego opracowania do Ministerstwa Środowiska (Nowakowska, Michalak i inni, 2011; Nowakowska, Hordejuk i inni, 2012).

Specyfikacje INSPIRE przewidują dwa rodzaje transformacji:

- transformację *on-line*, gdy dane pierwotne zawarte w bazie danych lub repozytorium są dynamicznie („w locie”) przekształcane w trakcie realizacji polecenia *get* w ramach usługi pobierania (*download service*) opartej na standardzie WFS (*Web Feature Service*) (INSPIRE IOCTF-NS, 2013); w takim przypadku dla pojedynczego polecenia *get* transformacji poddane zostają tylko te dane, które mają być przesłane jako odpowiedź serwera na to polecenie;
- transformację *off-line*, gdy zawartość bazy lub repozytorium danych pierwotnych zostaje poddana transformacji jako oddzielnej procedurze i wynik przetwarzania zostaje zapisany w innej bazie dedykowanej danym wynikowym, odpowiadający wymaganiom specyfikacji danych INSPIRE; gdy serwer odbierze polecenie *get* to w wyniku zostanie wysłana porcja danych w nowej formie pobrana z tej drugiej bazy.

Pierwszy sposób może być z sukcesem stosowany w przypadkach, gdy różnice pomiędzy danymi pierwotnymi i wynikowymi są względnie nieduże. W przypadku krajowych danych hydrogeologicznych (a także polskich danych objętych innymi tematami) różnice pomiędzy modelem źródłowym a modelem wynikowym są tak duże, że tryb *on-line* nie może być praktycznie zastosowany i pozostaje jedynie tryb *off-line*. Z tego względu przeprowadzone prace badawcze opisane w tej monografii są ograniczone do transformacji *off-line*.

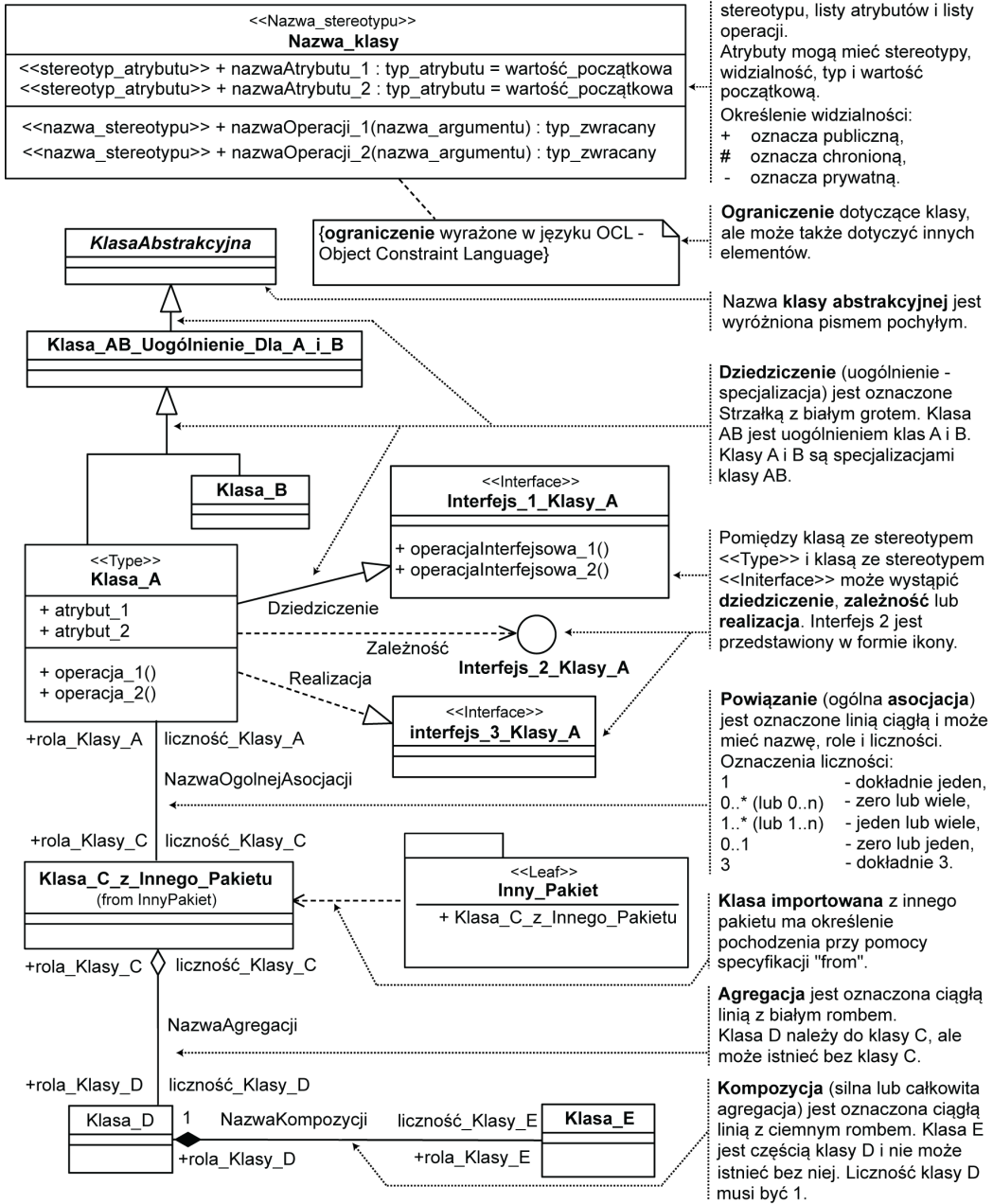
W kolejnych rozdziałach monografii przedstawione są szczegółowo wszystkie aspekty technologiczne, a także organizacyjne, poruszanego tu problemu transformacji.

- Rozdział 2. przedstawia zadania polskiego środowiska hydrogeologów jakie wynikają z dyrektywy INSPIRE. Prace te wymagają ścisłej współpracy przedstawicieli różnych subdyscyplin hydrogeologicznych, lecz także znaczącego udziału informatyków zajmujących się geoinformacją.
- W rozdziale 3. opisane są podstawy technologiczne infrastruktury INSPIRE, a w szczególności istota usług sieciowych i modele danych określone w specyfikacjach.
- Przegląd krajowych danych hydrogeologicznych jako źródła danych INSPIRE jest zawarty w rozdziale 4.
- W rozdziale 5. opisano i porównano najbardziej znane typy modeli i ich praktyczne instancje ze zwróceniem uwagi na wymagania standaryzacyjne dotyczące takich modeli.
- Rozdział 6. przedstawia model danych INSPIRE dla podtematu Hydrogeologia. Zawarta jest tam także analiza porównawcza pomiędzy tym modelem a modelami dotyczącymi danych krajowych.
- W rozdziale 7. zawarta jest analiza różnic pomiędzy modelami krajowymi i modelem INSPIRE, a także problematyka mapowania pomiędzy tymi różnymi modelami.
- Podstawy metodyczne transformacji danych geoprzestrzennych przedstawione są w rozdziale 8. Jako najbardziej efektywna metoda transformacji została tam przedstawiona transformacja zapisów znacznikowych w języku XML (W3C, 2008).
- Dla transformacji dokumentów (zbiorów danych) zapisanych w języku XML stosowana jest technologia oparta na XSLT (W3C, 1999). Szczegółowy opis tej technologii zawiera rozdział 9.
- Rozdział 10. zawiera przegląd baz danych XML i analizę możliwości ich wykorzystania do przechowywania zbiorów danych w języku XML, filtrowania zawartych w nich danych i zmiany ich struktury.
- W rozdziale 11. opisane zostały metody i odpowiadające im narzędzia transformacji, a w szczególności komercyjny system FME (*Feature Manipulation Engine*) i oprogramowanie otwarte (*Open Source*): HALE (*Humboldt Alignment Editor*), GeoKettle i Stetl.
- Opis procedury transformacji krajowych danych hydrogeologicznych do wymagań INSPIRE zawarty jest w rozdziale 12. Opisano tam wybór metodyki i narzędzi, opracowanie próbek danych i kolejne etapy prac, od zapisu płaskiego do końcowego wyniku w języku GML.
- Ostatni rozdział 13. zawiera podsumowanie.

Ponieważ struktury danych opisywanych w poszczególnych rozdziałach są przedstawiane graficznie za pomocą diagramów klas lub diagramów pakietów języka UML (*Unified Modeling Language*) stosowanego powszechnie do opracowywania, zapisywania i udostępniania, konieczne jest objaśnienie tej formy zapisu i występujących tam symboli graficznych. Notacja ta w zakresie profilu ISO dla modeli pojęciowych w geomatyce jest objaśniona na rysunku 1.1. Jednak sam diagram w formie graficznej nie zawiera wszystkich informacji dotyczących poszczególnych elementów modelu. Z tego względu dla pełnego poznania szczegółów modelu potrzebne jest dysponowanie jego zapisem w formie cyfrowej, na przykład w postaci plików XMI (*XML Metadata Interchange*) i odpowiedniego oprogramowania narzędziowego (na przykład Rational Rose lub Enterprise Architect). Szczegółowy opis modelu w formie tekstowej z ilustracjami przedstawiającymi poszczególne diagramy można także uzyskać za pomocą tego oprogramowania i jest to najczęściej stosowana droga do opracowania szczegółowej specyfikacji.

Symbole i ich elementy:

Objaśnienia:



Rysunek 1.1. Notacja graficzna diagramów klas języka UML (*Unified Modeling Language*) w zakresie profilu ISO dla modeli pojęciowych w geomatyce.

2. Zadania w zakresie hydrogeologii wynikające z dyrektywy INSPIRE

Dyrektywa INSPIRE określa w sposób zwięzły 34 zakresy tematyczne danych, jakie mają być udostępniane w Europejskiej Infrastrukturze Informacji Przestrzennej (EP&CEU, 2007). Szczegółowy opis podstawowych elementów tych danych określony jest w rozporządzeniach Komisji Europejskiej (EC, 2010; 2011; 2013a; 2013b) i w specyfikacjach danych dla poszczególnych tematów. Ustawa o infrastrukturze informacji przestrzennej w Polsce przypisuje poszczególne tematy organom wiodącym, którymi są ministrowie właściwi poszczególnym resortom lub prezesi odpowiednich urzędów centralnych (tab. 2.1). W przypadku hydrogeologii rolę organu wiodącego pełni Główny Geolog Kraju, a za realizację wynikających z tego zadań Państwowa Służba Hydrogeologiczna stanowiąca część Państwowego Instytutu Geologicznego.

Tabela 2.1. Zbiorcze zestawienie 34 tematów załączników I, II i III wraz z organami wiodącymi, którym te tematy podlegają na podstawie ustawy o IIP.
Pogrubieniem wyróżniono tematy dotyczące środowiska.

Organ wiodący	Pozycja (w ustawie o IIP)	Temat	Nazwa tematu w języku angielskim
Minister właściwy do spraw budownictwa, gospodarki przestrzennej i mieszkaniowej	Rozdz. 3. pkt. 4.	Zagospodarowanie przestrzenne	Land use
Minister właściwy do spraw gospodarki morskiej	Rozdz. 3. pkt. 15.	Warunki oceanograficzno-geograficzne	Oceanographic geographical features
	Rozdz. 3. pkt. 16.	Obszary morskie	Sea regions
Minister właściwy do spraw rolnictwa	Rozdz. 3. pkt. 9.	Obiekty rolnicze oraz akwakultury	Agricultural and aquaculture facilities
Minister właściwy do spraw środowiska	Rozdz. 1. pkt. 8.	Hydrografia	Hydrography
	Rozdz. 3. pkt. 12.	Strefy zagrożenia naturalnego	Natural risk zones
	Rozdz. 3. pkt. 13.	Warunki atmosferyczne	Atmospheric conditions
	Rozdz. 3. pkt. 14.	Warunki meteorologiczno-geograficzne	Meteorological geographical features
	Rozdz. 3. pkt. 19.	Rozmieszczenie gatunków	Species distribution
Minister właściwy do spraw zdrowia	Rozdz. 3. pkt. 5.	Zdrowie i bezpieczeństwo ludności	Human health and safety

cd. tabeli 2.1

Główny Geodeta Kraju	Rozdz. 1. pkt. 1.	Systemy odniesienia za pomocą współrzędnych	Coordinate reference systems
	Rozdz. 1. pkt. 2.	Systemy siatek georeferencyjnych	Geographical grid systems
	Rozdz. 1. pkt. 3.	Nazwy geograficzne	Geographical names
	Rozdz. 1. pkt. 4.	Jednostki administracyjne	Administrative units
	Rozdz. 1. pkt. 5.	Adresy	Addresses
	Rozdz. 1. pkt. 6.	Działki ewidencyjne	Cadastral parcels
	Rozdz. 1. pkt. 7.	Sieci transportowe	Transport networks
	Rozdz. 2. pkt. 1.	Ukształtowanie terenu	Elevation
	Rozdz. 2. pkt. 2.	Użytkowanie ziemi	Land cover
	Rozdz. 2. pkt. 3.	Ortoobrazy	Orthoimagery
	Rozdz. 3. pkt. 2.	Budynki	Buildings
	Rozdz. 3. pkt. 3.	Gleba	Soil
	Rozdz. 3. pkt. 6.	Usługi użyteczności publicznej i służby państwowe	Utility and governmental services
	Rozdz. 3. pkt. 8.	Obiekty produkcyjne i przemysłowe	Production and industrial facilities
	Rozdz. 3. pkt. 11.	Gospodarowanie obszarem, strefy ograniczone i regulacyjne oraz jednostki sprawozdawcze	Area management/restriction /regulation zones and reporting units
Główny Geolog Kraju	Rozdz. 2. pkt. 4.	Geologia	Geology
	Rozdz. 3. pkt. 20.	Zasoby energetyczne	Energy resources
	Rozdz. 3. pkt. 21.	Zasoby mineralne	Mineral resources
Główny Inspektor Ochrony Środowiska	Rozdz. 1. pkt. 9.	Obszary chronione	Protected sites
	Rozdz. 3. pkt. 7.	Urządzenia do monitorowania środowiska	Environmental monitoring facilities
Główny Konserwator Przyrody	Rozdz. 3 pkt. 17.	Regiony biogeograficzne	Bio-geographical regions
	Rozdz. 3. pkt. 18.	Siedliska i obszary przyrodniczo jednorodne	Habitats and biotopes
Prezes Głównego Urzędu Statystycznego	Rozdz. 3. pkt. 1.	Jednostki statystyczne	Statistical units
	Rozdz. 3. pkt. 10.	Rozmieszczenie ludności (demografia)	Population distribution – demography

Wśród 34 zakresów tematycznych wymienionych w aneksach dyrektywy INSPIRE (Michalak, 2009), których dane będą udostępniane, znajdują się także dane z zakresu tematu Geologia, a w konsekwencji także dane z podtematu Hydrogeologia (INSPIRE TWG-GE-MR et al., 2013). W ograniczonym stopniu dane hydrogeologiczne znajdują się także w innych tematach, co będzie przedstawione w dalszej części. Wiele innych danych z tych 34 zakresów jest także potrzebne w badaniach i pracach hydrogeologicznych. W tej sytuacji udział

polskiego środowiska hydrogeologicznego należy rozpatrywać w dwóch komplementarnych aspektach:

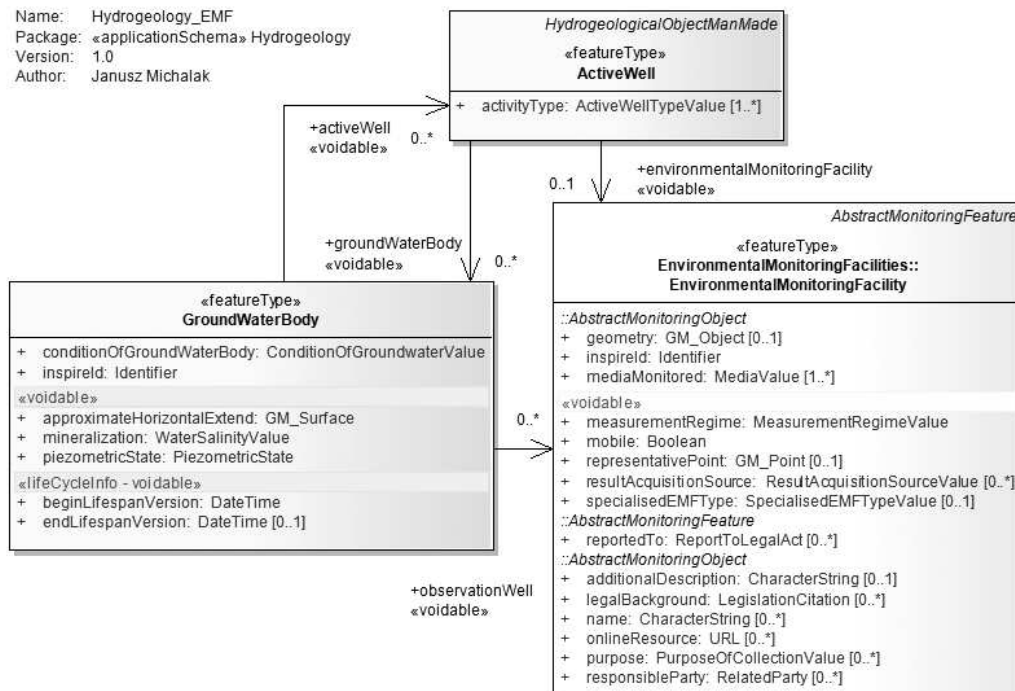
- korzyści z dostępu do danych udostępnianych przez ESDI z zakresu innych tematów,
- obowiązków dostarczenia do infrastruktury wiarygodnych, kompletnych i dobrze zorganizowanych danych z zakresu hydrogeologii.

Należy zaznaczyć, że dyrektywa INSPIRE nakłada obowiązek udostępniania tych danych, które istnieją w poszczególnych krajach członkowskich lub będą tworzone na potrzeby krajowe. Nie przewiduje się oddzielnych funduszy Unii Europejskiej na pozyskiwanie danych brakujących lub nowych. Wszelkie prace dotyczące danych hydrogeologicznych będą ukierunkowane i w konsekwencji jednocześnie ograniczone do transformacji danych z ich obecnych formatów, modeli i struktur do modeli i struktur określonych w specyfikacjach INSPIRE. Specyfikacje te bazują na modelu obiektowym z wykorzystaniem schematów aplikacyjnych języka GML (*Geography Markup Language*) (INSPIRE DT-DS, 2008). Z tym jest ściśle powiązana problematyka harmonizacji poszczególnych zbiorów danych hydrogeologicznych ze zbiorami należącymi do innych zakresów tematycznych (Nałęcz, Michalak, 2011).

Biorąc pod uwagę znaczne różnice w zakresie formatów, modeli i struktur pomiędzy polskimi danymi hydrogeologicznymi a rozwiązaniami proponowanymi w przepisach wykonawczych, brak harmonizacji tych danych z innymi polskimi danymi geoprzestrzennymi i także stosunkowo krótkie terminy, potrzebne są pilne prace badawcze z zakresu geomatyki hydrogeologicznej, których celem powinno być opracowanie metod transformacji i harmonizacji naszych danych. Na przeciw tym potrzebom wychodzi opisana tu inicjatywa badawcza i uzyskane wyniki powinny istotnie przyczynić się do zrealizowania tych zadań z sukcesem.

Dane hydrogeologiczne są wymienione *explicite* jako wydzielony podtemat tematu Geologia. Jednak poza tym podtematem (w specyfikacjach innych tematów) występują także elementy informacji hydrogeologicznej. Do tych tematów należą:

- Temat Gospodarowanie obszarem, strefy ograniczone i regulacyjne oraz jednostki sprawozdawcze (*Area management/restriction/regulation zones and reporting units*) zawiera typ wyróżnienia przestrzennego *ManagementRestrictionOrRegulationZone*. Zgodnie ze zdefiniowaną tam listą kodową *ZoneTypeCode* (Kod typu strefy) definiującą atrybut *zoneType* tego wyróżnienia, strefą tą może być: *drinkingWaterProtectionArea* (obszar ochrony wody pitnej), *prospectingAndMiningPermitArea* (obszar objęty wymaganiem uzyskania zezwolenia na wydobywanie) w przypadku wód mineralnych lub obszar *WFDWaterBody* (część wód zgodnie z ramową dyrektywą wodną (2000/60/WE)) bez wyszczególnienia. W przypadku WFD nie ma znaczenia czy dotyczy to wód powierzchniowych czy podziemnych (INSPIRE TWG-AM, 2013). W specyfikacji tego tematu obiekty należące do tej klasy nie mają powiązań z obiektami z innych tematów, w tym z podtematem Hydrogeologia.
- Temat Hydrografia (*Hydrography*) dotyczący wód powierzchniowych w ostatecznej wersji specyfikacji nie zawiera powiązań z elementami modelu dla wód podziemnych.
- Temat Urządzenia do monitorowania środowiska (*Environmental Monitoring Facilities*) zawiera typ wyróżnienia przestrzennego *EnvironmentalMonitoringFacility* (Urządzenie do monitorowania środowiska), który może odnosić się do studni obserwacyjnej (piezometru) zdefiniowanej w podtemacie Hydrogeologia (INSPIRE TWG-EM, 2013). Powiązania asocjacyjne pomiędzy tymi elementami należącymi do różnych tematów przedstawia rysunek 2.1.



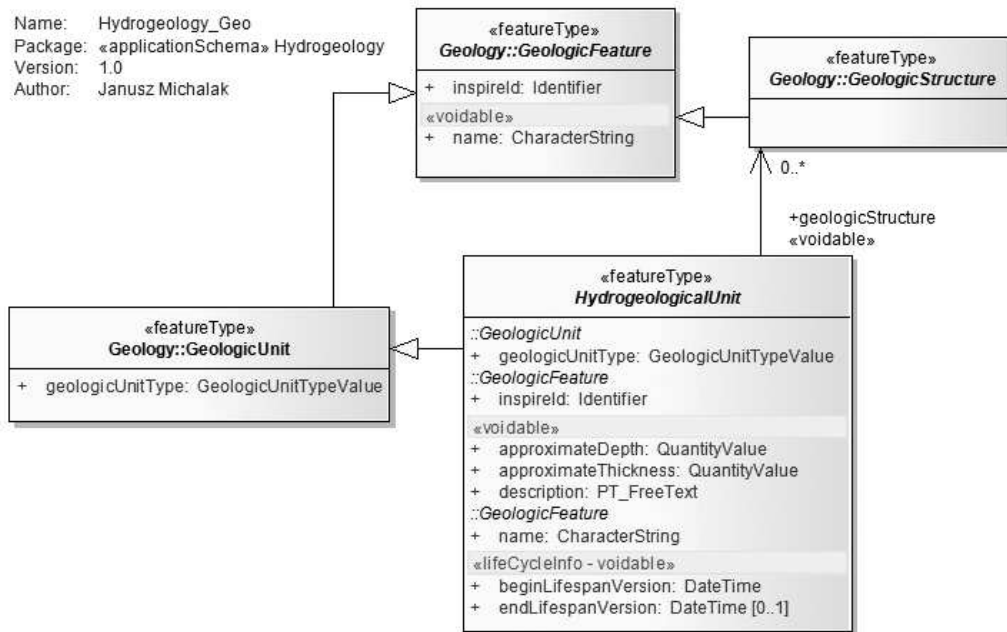
Rysunek 2.1. Diagram klas UML przedstawiający powiązania asocjacyjne pomiędzy klasami modelu danych hydrogeologicznych i klasami tematu Urządzenia do monitorowania środowiska.

- Ze zrozumiałych względów najwięcej powiązań zewnętrznych występuje pomiędzy podtematem Hydrogeologia i podtematem Geologia (*Geology*). Powiązania te są przedstawione na rysunkach 2.2 i 2.3.

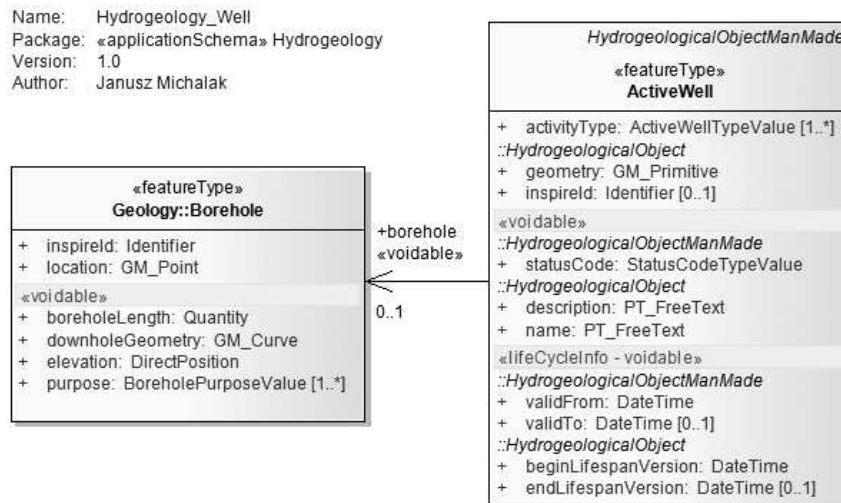
2.1. Międzynarodowa Mapa Hydrogeologiczna Europy

Jednym z nielicznych wspólnych europejskich przedsięwzięć z zakresu geoinformacji hydrogeologicznej jest kończący się w grudniu 2015 roku projekt IHME1500 (*International Hydrogeological Map of Europe 1:1 500 000*) (Gilbrich, 2000; Duscher, 2014a; 2014b). Jednak bliższe zapoznanie się z założeniami, koncepcją metodyczną i wynikami tego projektu nie pozwala patrzeć optymistycznie na efektywny udział europejskich hydrogeologów w pracach nad implementacją specyfikacji INSPIRE. Wyjaśnieniem tej sytuacji jest fakt, że projekt ten rozpoczął się w roku 1960, czyli czas jego realizacji wynosi 56 lat (BGR, 2008). Podstawowa koncepcja powstała przed jego rozpoczęciem, czyli była oparta o stan wiedzy z lat pięćdziesiątych. Legenda tej mapy jest oparta na publikacji z zakresu standaryzacji legend map hydrogeologicznych (Struckmeier, Margat, 1995). Obecnie mapa ta jest dostępna w kilku wersjach:

- w formacie PDF ze zredukowaną skalą do 1:5 00 000 i w konsekwencji zredukowaną szczegółowością i dokładnością (PlanetEarth, 2008);
- w postaci rastrowej zapisanej w języku KML (*Keyhole Markup Language*) do wykorzystania za pomocą oprogramowania Google Earth; są to jedynie niskiej jakości obrazy powstałe przez skanowane dawnych map papierowych.

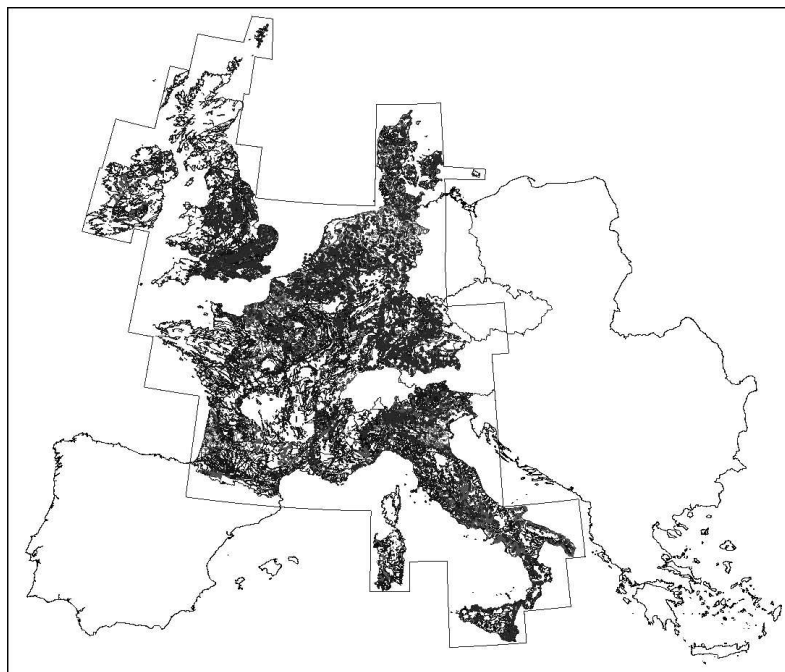


Rysunek 2.2. Diagram klas UML przedstawiający powiązania asocjacyjne i hierarchiczne pomiędzy abstrakcyjną klasą *HydrogeologicalUnit* modelu danych hydrogeologicznych i klasami tematu Geologia: *GeologicStructure* i *GeologicUnit*. Klasy *Aquifer*, *Aquitard* i *Aquiclude* dziedziczą zarówno te powiązania jak i atrybuty od klasy *HydrogeologicalUnit*.



Rysunek 2.3. Diagram klas UML przedstawiający powiązanie asocjacyjne pomiędzy klasą *ActiveWell* modelu hydrogeologicznego i klasą *Borehole* modelu geologicznego zawierającą dane o otworze wiertniczym, który posłużył do wybudowania studni.

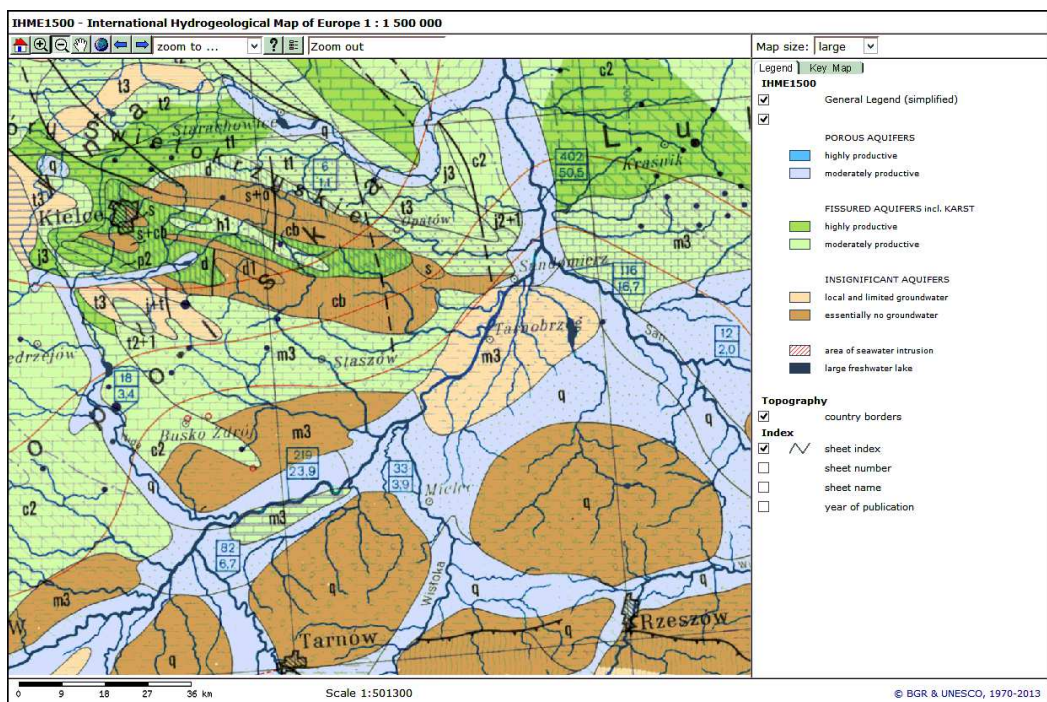
- przez usługę dla map w Internecie bazująca na niestandardowych rozwiązaniach w oprogramowaniu w języku Java; ta wersja ma dostatecznie dobrą dokładność i szczegółowość (rys. 2.5.);



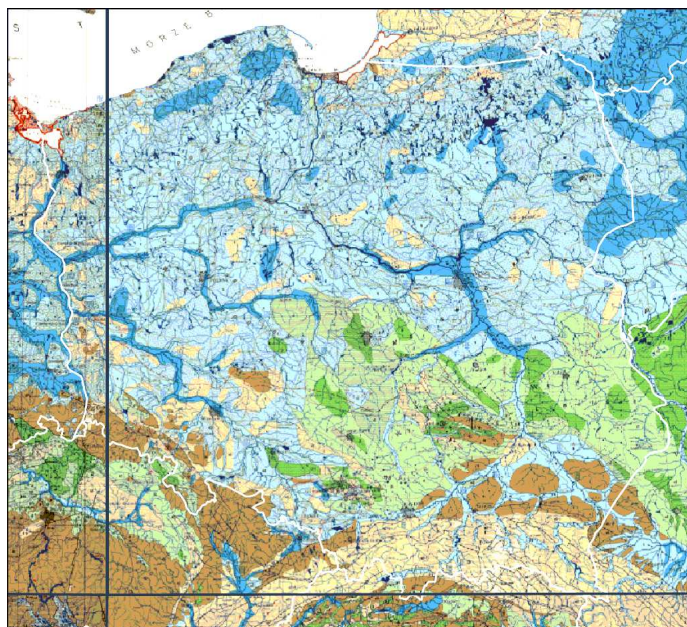
Rysunek 2.4. Wersja wektorowa mapy IHME 1500 w formacie ESRI Shape udostępniana w Internecie przez protokół FTP. Wersja ta obejmuje tylko część krajów europejskich.

- przez standardową usługę OGC WMS w wersji 1.1.1; przekazanie polecenia GetMap: http://www.bgr.de/Service/groundwater/ihme1500/?REQUEST=GetMap&VERSION=1.1.1&SERVICE=WMS&SRS=EPSG:4326&WIDTH=1000&HEIGHT=1000&BBOX=13.5,48.5,24.5,56.5&LAYERS=ihme_sheets&FORMAT=image/png zwraca z serwera obraz rastrowy przedstawiony na rysunku 2.6; jednak jakość tego obrazu jest dość niska i widać wyraźnie różnice pomiędzy poszczególnymi arkuszami mapy dotyczące jej treści i formy przedstawianych wydziałów kartograficznych (rys. 2.7.);
- przez protokół FTP (*File Transfer Protocol*); wersja wektorowa tej mapy zapisana w formacie Esri Shape, udostępniona w roku 2014 (Duscher, Günther, 2014); dla tej wersji mapy opracowano także metadane, które są dostępne w Internecie (BGR, 2014); jednak ta wersja obejmuje tylko część krajów Europy Zachodniej (rys. 2.4.); najbardziej szczegółowe dane tej wersji obejmują wyspy brytyjskie, co przedstawione jest na rysunku 2.9. dla południowego fragmentu Anglii.

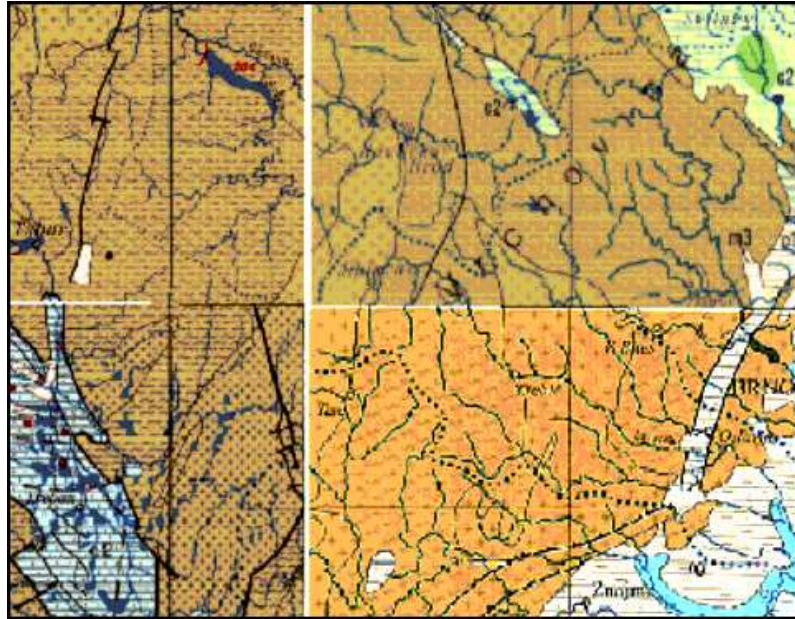
Dostępne obecnie wyniki tego projektu w formie elektronicznej (Hollis i in., 2002; Duscher, 2014a) świadczą o wprowadzonych ostatnio pewnych modyfikacjach w zakresie technologicznym, jednak w porównaniu z technologiami proponowanymi w ESDI INSPIRE są to raczej zabytki muzealne. Analizując treść i formę danych hydrogeologicznych tego projektu można łatwo zauważyć, że struktura danych jest całkowicie niezgodna zapisami w języku GML (*Geography Markup Language*) przyjętym jako podstawa kodowania danych INSPIRE. Obecnie język GML jest już powszechnie uznaną formą zapisu danych i literatura dotycząca tego języka jest bogata – od podstawowych specyfikacji do wielu przykładów



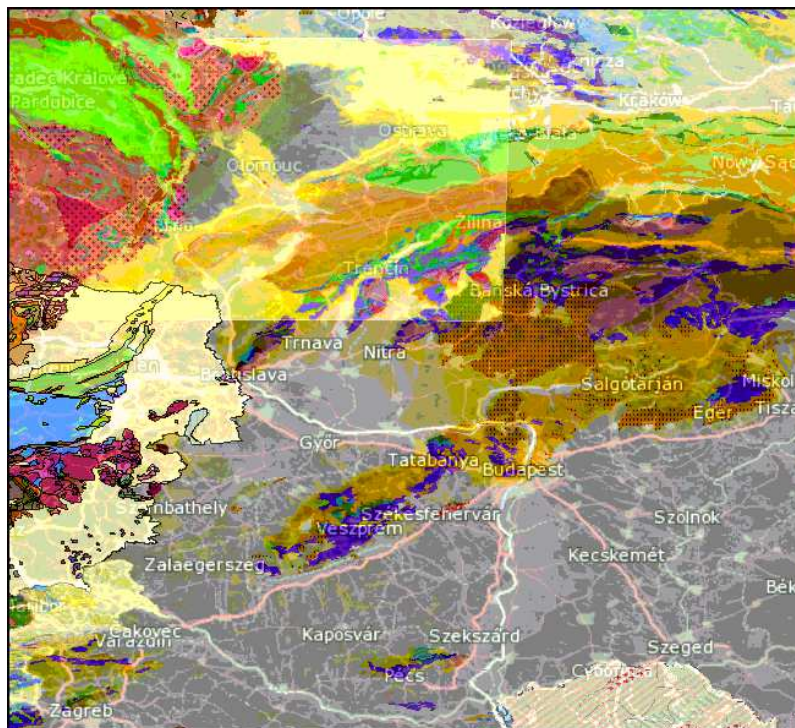
Rysunek 2.5. Fragment mapy IHME 1500 wraz z legendą udostępnianej w Internecie w technologii opartej na języku Java. Fragment ten obejmuje okolice Kielc, Tarnowa i Rzeszowa.



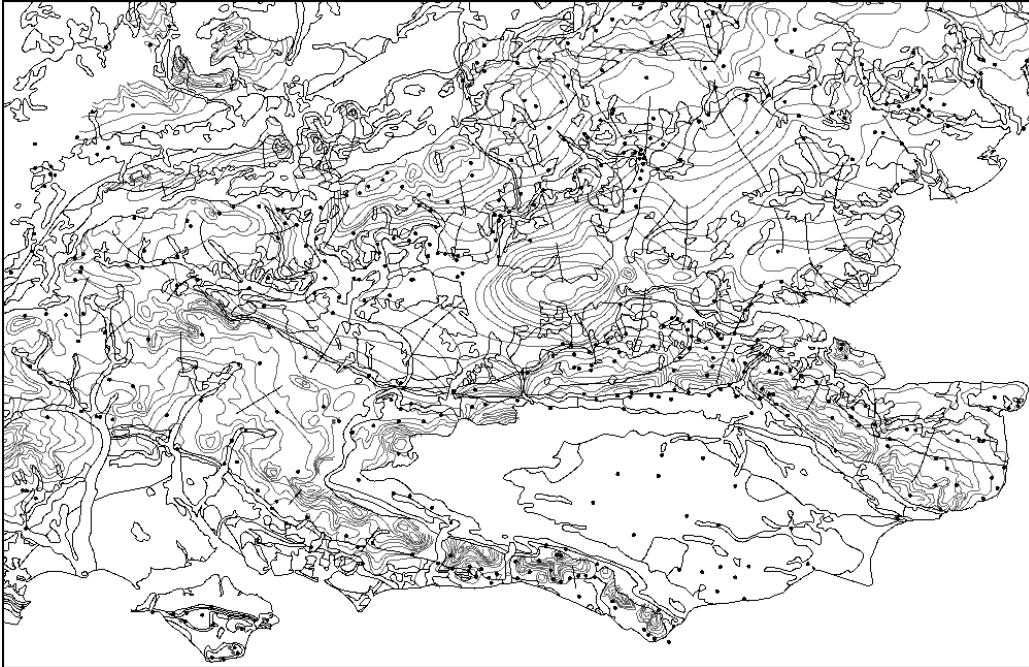
Rysunek 2.6. Fragment mapy IHME 1500 dla obszaru Polski. Wersja udostępniana przy pomocy interfejsu WMS zgodnego ze standardem OGC w wersji 1.1.1. Objaśnienia w tekście.



Rysunek 2.7. Fragment mapy IHME 1500 w wersji KML dla Google Earth jako zeskanowane arkusze papierowe. Widoczne arkusze różnią się częściowo treścią i formą



Rysunek 2.8. Dla porównania fragment mapy One Geology obszaru Europy centralnej. Tu również poszczególne fragmenty różnią się częściowo treścią



Rysunek 2.9. Fragment mapy IHME 1500 w wersji wektorowej w formacie ESRI Shape. Fragment ten obejmuje południową Anglię i mapa ta jest dla tego obszaru najbardziej szczegółowa.

zastosowań (Lake, 1999; Lake, Burggraf, Trninic, Rae, 2004; Portele, 2007; Wang, Wang, Qian, Zhang, 2011; ISO/TC211, 2007a; OGC, 2008a; Peng, Zhang, 2004; Portale, Van den Broecke, Cox, 2009; Tennakoon, 2003). Jednak nie tylko zbiory danych IHME1500 odbiegają od stawianych obecnie wymagań – dla porównania rysunek 2.8. przedstawia dane projektu One Geology dla fragmentu Europy Centralnej – tam także są widoczne rozbieżności pod względem treści i formy poszczególnych arkuszy mapy. Zbiory danych związane z tym projektem, a także z innymi projektami zbliżonymi do niego tematycznie są dostępne w repozytorium JRC (*Joint Research Centre*) (JRC, 2009).

2.2. Udostępnianie krajowych danych hydrogeologicznych

Jednym ze statutowych zadań Państwowej Służby Hydrogeologicznej (PSH) jest udostępnianie danych hydrogeologicznych zgromadzonych w bazach utrzymywanych przez tę służbę. Przykładem, jak zadania w tym zakresie są określone, jest opracowanie pt. „Zadania Państwowej Służby Hydrogeologicznej w 2013 roku – Synteza”. Opracowanie to wykonano w ramach realizacji przedsięwzięcia „Zadania państwowej służby hydrogeologicznej w latach 2012-2014” (Woźnicka, Skrzypczyk, Nawrocki, 2013). Procedura udostępniania tych danych jest jednak objęta szeregiem ograniczeń, między innymi wynikających z faktu, że dane są własnością Skarbu Państwa. Innym istotnym utrudnieniem jest forma udostępnianych danych – najczęściej są to papierowe kserokopie lub pliki PDF. Zbiory danych w formie cyfrowej zgodnej z przyjętymi obecnie standardami wymiany geoinformacji mają bardzo ograniczony zakres tematyczny i przestrzenny, a także najczęściej są zawężone pod

względem treści (liczby elementów lub atrybutów) z powodu niekompatybilności pomiędzy ich wewnętrzną formą (strukturą) a wymaganiami standardów wymiany danych pomiędzy systemami zewnętrznymi. Najczęściej spotykanym formatem wektorowym w jakim dane są udostępniane jest format Esri Shape, co wymusza stosowanie bardzo prostego płaskiego modelu danych i w konsekwencji utraty znacznej części informacji. Potencjalny użytkownik danych PSH może w witrynie Narodowego Archiwum Geologicznego znaleźć wskazówki jak przeglądać i wyszukiwać dane hydrogeologiczne (PSH, 2012b) i jak te dane zamówić (PSH, 2013c). Wiele szczegółowych informacji z zakresu udostępniania danych hydrogeologicznych przez PSH można znaleźć w publikacji Ferta, Mordzonka i Węglarza (2011).

Stan jaki obecnie istnieje, odbiega bardzo daleko od koncepcji przyjętej w podstawowych dokumentach określających wizję europejskiej infrastruktury INSPIRE, jej zadania, założenia prawne i przyjęte rozwiązania technologiczne. Z tego powodu można przypuszczać, że implementacja dyrektywy INSPIRE zdecydowanie zmieni na lepsze warunki i sposoby dostępu do krajowych danych hydrogeologicznych.

2.3. Treść i forma udostępnianych danych

Szczegółowy opis danych hydrogeologicznych pozyskiwanych i utrzymywanych przez Państwową Służbę Hydrogeologiczną jest zawarty w rozdziale 4. Z tego względu tu jest przedstawiony jedynie aspekt udostępniania tych danych. Ponieważ zasady udostępniania danych, a w konsekwencji ich zakres i forma ulegają częstym zmianom, poniższe informacje służą jedynie przedstawieniu ogólnego stanu w tym zakresie, w celu porównania z zakresem i formą danych jakie mają być udostępniane przez infrastrukturę INSPIRE. Głównym źródłem informacji na ten temat jest witryna internetowa PSH (PSH, 2012a; 2012b; 2013a; 2013b; 2013c; 2013d; 2013e; 2013f; 2013i; 2014a; 2014b; 2014d). Niepełna lista kategorii udostępnianych danych w postaci elektronicznej z przypisaniem geoprzestrzennym, zarówno ściśle hydrogeologicznych jak i geologicznych wykorzystywanych w hydrogeologii, jest podzielona na dane wektorowe i na dane rastrowe.

Dane wektorowe:

- MHP (Mapa Hydrogeologiczna Polski) – Główny Użytkowy Poziom Wodonośny (GUPW) – arkusze mapy w formacie Esri Shape;
- MHP (Mapa Hydrogeologiczna Polski) – Pierwszy Poziom Wodonośny (PPW) – Występowanie i Hydrodynamika (WH) – arkusze mapy w formacie Esri Shape;
- MHP (Mapa Hydrogeologiczna Polski) – Pierwszy Poziom Wodonośny (PPW) – Wrażliwość na zanieczyszczenie i Jakość wód (WJ) – arkusze mapy w formacie Esri Shape;
- dodatkowe dane, które z przyczyn technologicznych nie mogą być zapisane w tym formacie i dotyczą map MHP – GUPW, PPW-WH i PPW-WJ są udostępniane w formie geobazy oprogramowania GeoMedia firmy Intergraph; dodatkowe dane zapisane w geobazie dotyczą:
 - analiz chemicznych w otworach, studniach kopanych, ze źródeł i wsięków, a także innych punktów pomiarowych,
 - klasyfikację na jednostki i piętra,
 - dane o otworach i innych punktach dokumentowanych;
- MGP (Mapa Geośrodowiskowa Polski) – arkusze mapy w formacie Esri Shape;
- SMGP (Szczegółowa Mapa Geologiczna Polski) – arkusze mapy w formacie Esri Shape.

Dane rastrowe:

- Mapa Lito-genetyczna – arkusze mapy w postaci zapisu rastrowego TIFF (*Tagged Image File Format*);
- MGGP (Mapy Geologiczno-Gospodarczej Polski) – arkusze mapy w postaci zapisu rastrowego TIFF.

Wiele innych danych o charakterze źródłowym znajduje się w bazie Centralnego Banku Danych Hydrogeologicznych oraz jednostkach prowadzących bieżący monitoring wód podziemnych. Jednak w jakim zakresie i w jakiej formie są one udostępniane można się dowiedzieć jedynie przez bezpośredni kontakt z pracownikami PSH. Podane w tym rozdziale informacje o formie i zakresie udostępnianych przez Internet danych hydrogeologicznych odnoszą się do stanu jaki był w lipcu 2015 roku. Należy oczekiwać, że z upływem czasu zakres tych danych będzie się zwiększał, a forma będzie w znacznie większym stopniu odpowiadała dedykowanym do tego celu nowym technologiom.

2.4. Krajowe dane hydrogeologiczne w środowisku infrastruktury

Państwowa Służba Hydrogeologiczna podejmowała i nadal podejmuje szereg działań mających na celu udostępnianie danych, w tym geoprzestrzennych, z zastosowaniem technologii internetowych opartych na standardach OGC (*Open Geospatial Consortium*) (Buechler, McKee, 1996) i ISO TC 211 (*Geographic information/Geomatics*) (Rossa, Gogołek, Łukasiewicz, 2009).

Na początku roku 2015 czynne były trzy serwery danych geoprzestrzennych z zakresu hydrogeologii:

- wcześniejszy epsh.pgi.gov.pl z usługami opartymi na technologii GeoMedia SDI Portal firmy Intergraph;
- późniejszy spdpsh.pgi.gov.pl z usługami realizowanymi przez oprogramowanie *Open-Source* Geoserver i OpenLayers;
- serwer Centralnej Bazy Danych Geologicznych, który udostępnia część danych z zakresu hydrogeologii: JCWP w 2 kategoriach, obszary zagrożone podtopieniami i otwory hydrogeologiczne (obserwacyjne i eksploatacyjne). Serwer ten ma usługi w standardzie WMS wersji 1.1.1 i 1.3.0, a także interfejs WebMap dla urządzeń mobilnych.

Dwa pierwsze serwery mają interfejsy WebMap umożliwiające korzystanie z przeglądark dla oglądania kompozycji mapowych utworzonych z udostępnianych danych.

Interfejs serwera [epsh](http://epsh.pgi.gov.pl) (PSH, 2013g) przedstawiony jest na rysunku 2.12, a serwera [spdpsh](http://spdpsh.pgi.gov.pl) (PSH, 2014c) na rysunku 2.15. Szczegółowy opis interfejsu [epsh](http://epsh.pgi.gov.pl) zawiera instrukcja systemu GeoMedia SDI Portal (IP, 2011), a opis interfejsu [spdpsh](http://spdpsh.pgi.gov.pl) także zawarty jest w instrukcji (PSH, 2013h). Dla obu tych serwerów przeprowadzone zostały testy pod kątem spełniania standardów i norm, a także wymagań INSPIRE. Część wyników tych testów przedstawiona jest w rozdziale 2.4.3.

Testowanie serwerów zostało przeprowadzone w przygotowanym do tego środowisku testowym przedstawionym na rysunku 2.10. Śledzenie komunikacji pomiędzy serwerem a klientem dotyczyło pakietów TCP/IP, jako najniższy poziom komunikacji i dwustronnego przepływu bloków danych w protokole HTTP – wysyłanych do serwera poleceń i odbieranych przez klienta odpowiedzi na te polecenia. W pierwszym przypadku zastosowano oprogramowanie narzędziowe przechwytyjące pakiety TCP/IP, między innymi program Snoop

systemu operacyjnego Solaris. W drugim przypadku został użyty moduł FireBug zaawansowanej technologicznie przeglądarki FireFox. Metodyka testowania serwerów geoinformacyjnych opisana jest także w innej monografii (Michalak, 2010a).

Moduł FireBug jest zaawansowanym rozbudowanym narzędziem do analizy wielu elementów składających się na języki opisu stron (HTML, CSS, JavaScript) i na funkcjonowanie WWW oparte na protokole HTTP. Funkcje tego programu są pogrupowane w 6 kategorii:

- Konsola: Błędy, Ostrzeżenia, Informacje, Ciasteczka,
- HTML – szczegółowa analiza kodu strony z możliwością edycji,
- CSS – szczegółowa analiza kodu arkuszy stylów z możliwością edycji,
- Skrypty – analiza kodu JavaScript,
- DOM – analiza obiektowego modelu dokumentu (W3C, 2005),
- Sieć: HTML, CSS, JavaScript, XMLHttpRequest, Obrazki, Wtyczki, Media, Czcionki.

Przykłady użycia tego programu przedstawiają rysunki: 2.11., 2.12., 2.16., 2.17. i 2.18.

2.4.1. Testowanie serwera e-PSH

Serwer e-PSH jest pierwszym serwerem krajowym dedykowanym danym hydrogeologicznym. Serwer ten został opracowany przez Intergraph Polska w technologii GeoMedia SDI Portal i jego podstawowe dane są zawarte w tabeli 2.2. W trybie domyślnym pracuje w standardzie WMS 1.3.0 z układem odniesienia przestrzennego PUGW 1992 i oferuje 15 warstw tematycznych. W porównaniu z jego następcą – serwerem SPD-PSH (rozdział 2.4.2.), dobrze spełnia wymagania standardów i norm z tego zakresu. Jest typowym przykładem serwera WMS, ponieważ poprzez jedno polecenie GetMap można uzyskać obraz mapy w postaci jednego pliku dla wielu warstw, w oficjalnie obowiązującym układzie odniesienia i w dokładnie i dowolnie podanej skali odpowiedniej do szczegółowości i przestrzennego zasięgu danych. Tak jak zdecydowana większość takich serwerów ma interfejs WebMap dedykowany przeglądarkom. Okno przeglądarki wyświetlającej ten interfejs jest przedstawione na rysunku 2.12., a szczegółowy opis zawiera instrukcja użytkownika (PSH, 2013g).

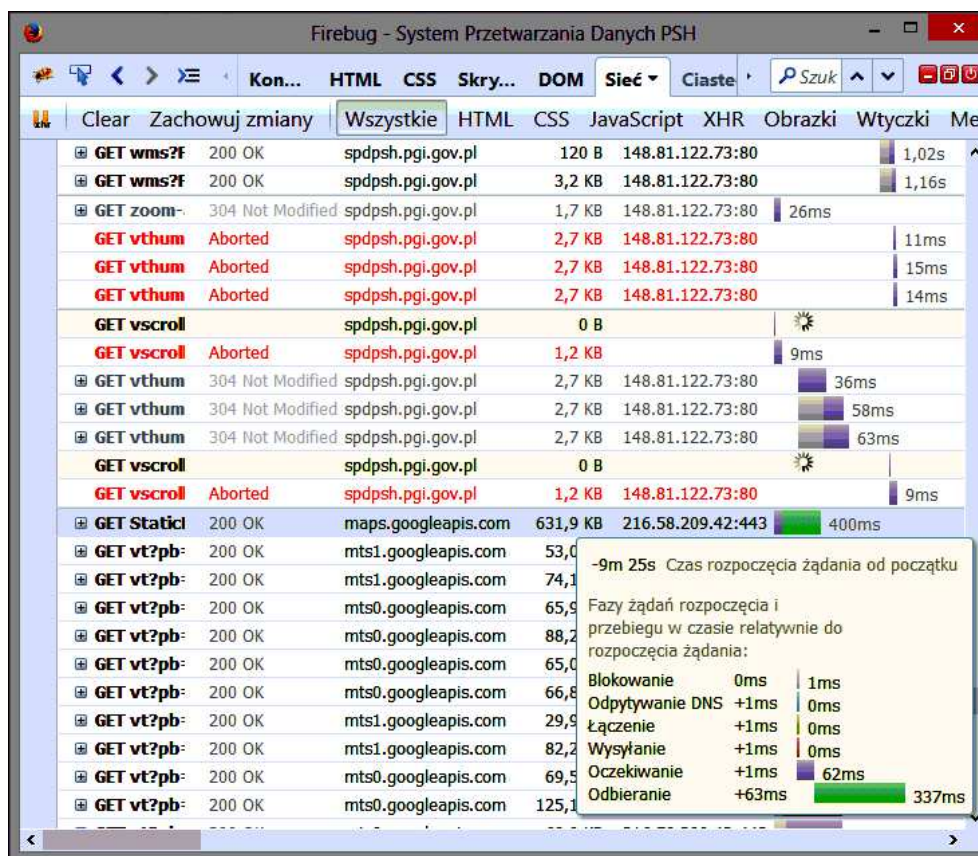
Pierwszym etapem prac testowych jest identyfikacja elementów oprogramowania agenta, w tym kodu HTML i skryptów JavaScript. Początkowy fragment kodu HTML przesyłanego od tego serwera jest przedstawiony na rysunku 2.13. Jest on napisany w języku XHTML wersji 1.0. Kod ten nie ma zdefiniowanego sposobu kodowania znaków (na przykład jako UTF-8) i z tego względu mogą nie być poprawnie przetwarzane polskie znaki diakrytyczne (na przykład dni tygodnia: [”poniedziałek”, ”Łroda”, ”piątek”). Jeżeli pochodzenie zastosowanej technologii nie jest widoczne, to można je ustalić na podstawie nazw występujących tam elementów kodu, na przykład nazw funkcji JavaScript. W tym przypadku znajduje się tam napis „<div id=’powered_by’>Powered by Intergraph</div>”, co upraszcza analizę.

Dalszą analizę przeprowadza się za pomocą modułu FireBug. Dla wyszukania poleceń WMS GetMap w programie FireBug należy w grupie „Sieć” wybrać kategorię „Obrazki” i przełączyć okno przeglądarki połączonej z serwerem. Z długiej listy poleceń wyświetlonych w oknie modułu FireBug można znaleźć polecenie zgodne ze standardem WMS, co w przypadku tego serwera jest przedstawione na rysunku 2.14.

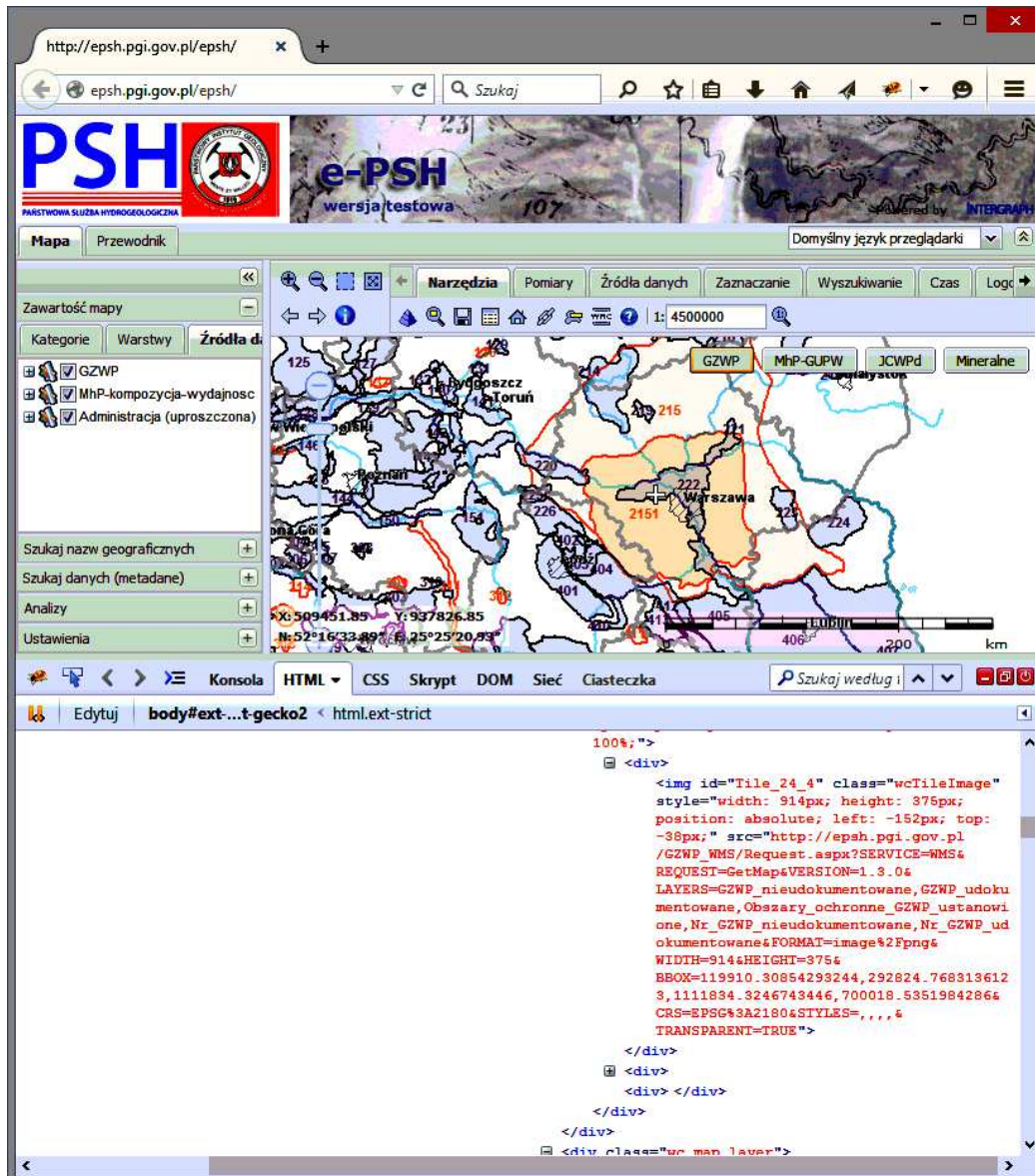
Szczególne uwagę przy testowaniu przykładano do zgodności usług WMS ze standardami OGC dotyczącymi tych usług (OGC, 2006) i normami ISO (ISO, 2005a; 2005c). Spełnianie tych wymagań jest warunkiem poprawnej interoperacyjności serwerów danych geo-



Rysunek 2.10. Schematyczne przedstawienie procesu inicjowania sesji WebMap i jej trwania, a także lokalizacja oprogramowania testującego ten proces. Opis w tekście.



Rysunek 2.11. Okno modułu FireBug podczas inicjowania sesji WebMap pomiędzy serwerem SPD-PSH i przeglądarką FireFox. Tekst czerwony – polecenia odrzucone, żółte pasy – polecenia niezrealizowane. Opis w tekście.



Rysunek 2.12. Okno przeglądarki Firefox przedstawiające interfejs WebMap serwera e-PSH. W dolnej części okna mieści się panel modułu FireBug z fragmentem kodu XHTML zawierającym polecenie GetMap w wersji WMS 1.3.0 dla wybranego obszaru określonego parametrem BBOX i dla 5 warstw określonych parametrem LAYERS: – GZWP_nieudokumentowane, – GZWP_udokumentowane, – Obszary_ochronne_GZWP_ustanowione, – Nr_GZWP_nieudokumentowane, – Nr_GZWP_udokumentowane. Z treści tego polecenia wynika, że zwracany obraz mapy w formacie PNG ma wymiary 912 pikseli szerokości i 375 pikseli wysokości. Mapa jest odwzorowana w układzie EPSG:2180, co odpowiada układowi współrzędnych 1992. Bardziej szczegółowy opis w rozdziale 2.4.1.

```

3 <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
4
5 <html xmlns="http://www.w3.org/1999/xhtml" xmlns:v="urn:schemas-microsoft-
  com:vml">
6 <head><meta http-equiv="X-UA-Compatible" content="IE=EmulateIE7" /><title>
7
8 </title><link href="App Themes/Default/Default.css" type="text/css"
  rel="stylesheet" /><script type="text/javascript"
  src="http://www.google.com/jsapi?key="></script><script
  type="text/javascript" src="http://ecn.dev.virtualearth.net/mapcontrol
  /mapcontrol.ashx?v=6.2"></script><link href="/EPSH
  /WebResource.axd?d=k5cTP1QH5yxThewpA2aumSshKJPFoRwlstuiSLBU2nq23h6ykfesUQ3
  xPid4TtjXty896owOtmJHkQ2uVFEwDOnckNeATVmGRWGGpJeQINH3S5AUypNacEGZPA9L9Bf4e
  wJCUNI77XkPK3xk6wJQ3PpsbvOTE3ro4PErJYKtJbVgeRq50&
  amp;t=634332686518676250" type="text/css" rel="stylesheet" /><link
  href="/EPSH/WebResource.axd?d=YH0BwvPjvU iG7j4ErY5q9i2-
  VOyGw81dBmNWXcFEHpMdVqvIBTO4wggq2geRgmFJzYagi2ZKhFmLnU9QZFwy-
  tiBiiGRcdjUDRC1N1DnQ-cWcuDeQoU3Zu1WC2fKij-dh2c73vE3PFO-
  Q2MjhgFQaE6lQ4uJacJ38gXgn8zPK3f6Hole0&amp;t=634332686518676250"
  type="text/css" rel="stylesheet" /><link
  href="Compositor.WebClient.ashx?p=%2fepsh%2fdefault.aspx&e=css&
  amp;t=635357303947711677" type="text/css" rel="stylesheet" /></head>
9 <body>
10 <form name="form1" method="post" action="default.aspx" id="form1">
11 <div id="powered_by">
12
13 Powered by <span>Intergraph</span>
14
15 </div><div>
16 <input type="hidden" name="__wc_selected_features"
  id="__wc_selected_features" value="" />
17 (...)
18 </div>
19 (...)
20 <div class="wc_loading_indicator" id="ctl105Mask">
21
22 <div id="splash">
23 <div id="splash_title"></div>
24 <div id="splash_wait">Wczytywanie aplikacji. ProszÄ
  czekaÄ...</div>
25 </img>
26 </div>
27 </div>
28 (...)
29 <span id="ctl105" style="display:none;"></span>
30 <div id="ctl106">
31 <div id="ctl106_Banner" style="height:85px;">

```

Rysunek 2.13. Początkowy fragment kodu XHTML agenta przesłanego z serwera e-PSH do przeglądarki. Opis w tekście.

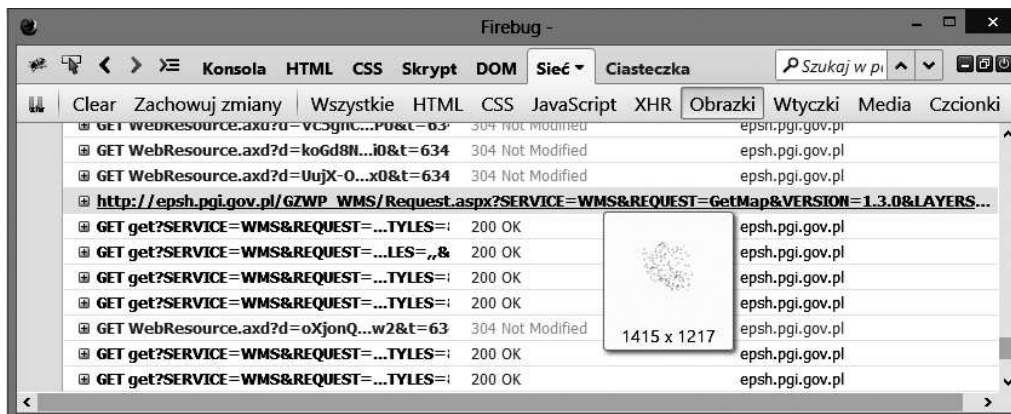
przestrzennych. W wielu przypadkach, szczególnie wynikających z koncepcji infrastruktury INSPIRE, serwery muszą mieć możliwość współdziałania w taki sposób, aby ostateczny wynik był złożoną kompozycją z danych pochodzących z różnych baz, co sprowadza się do konieczności współdziałania różnych serwerów.

Przykład 2.1. Postać polecenia wysłanego przez agenta do serwera e-PSH (wiersz podświetlony na rysunku 2.13) z podziałem na poszczególne parametry:

```

http://epsh.pgi.gov.pl/GZWP_WMS/Request.aspx?
SERVICE=WMS&
REQUEST=GetMap&
VERSION=1.3.0&
LAYERS=
    GZWP_nieudokumentowane,
    GZWP_udokumentowane,
    Obszary_ochronne_GZWP_ustanowione,
    Nr_GZWP_nieudokumentowane,
    Nr_GZWP_udokumentowane&
FORMAT=image%2Fpng&
WIDTH=1415&
HEIGHT=1217&
BBOX=
    -251178.9543452313,
    -206394.79848778632,
    1284755.954345231,
    1114541.7984877862&
CRS=EPSG%3A2180&
STYLES=,,,&
TRANSPARENT=TRUE

```



Rysunek 2.14. W oknie programu FireBug obok poleceń „GET WebResource...” widocznych jest szereg standardowych dla WMS poleceń typu „GET get?SERVICE=WMS...” wysyłanych do serwera epsh.pgi.gov.pl. Jedno z nich jest rozwinięte i wyświetlony jest obrazek mapy o wymiarach 1415×1217 przesłany z tego serwera jako odpowiedź na to polecenie. Opis w tekście.

Wyniki przeprowadzonych testów tego serwera, zarówno pod względem zgodności ze standardem WMS, jak i poprawności funkcjonowania interfejsu WebMap są pozytywne. Bardziej szczegółowe dane, stanowiące zestawienie wyników testów są podane w tabeli 2.2. Wymagania wynikające ze specyfikacji INSPIRE są w dużym stopniu spełnione, jednak serwis ten w aspekcie wymagań INSPIRE ma kilka mankamentów. Najważniejsze z nich to:

- brak wielojęzyczności interfejsu WebMap, w INSPIRE wymagane są przynajmniej dwa języki – język danego kraju i angielski,
- brak usług pobierania danych w standardzie WFS, WMS, jest jedynie usługą przeglądania,
- obawa, że ograniczona wydajność sprzętu może być przyczyną niespełnienia wymogów Rozporządzenia Komisji Europejskiej w zakresie jakości usług sieciowych (EC, 2010),
- serwer ten nie obsługuje wszystkich układów odniesienia wymaganych przez specyfikacją INSPIRE dotyczącą tego zagadnienia: *Data Specification on Coordinate Reference Systems – Technical Guidelines* (INSPIRE TWG-CRS&GGS, 2014a).

2.4.2. Testowanie serwera SPD-PSH

Drugim serwerem PSH poddanym testowaniu był nowszy serwer o nazwie System Przetwarzania Danych (SPD-PSH) i adresie <http://spdpsh.pgi.gov.pl/GeoPSHv7/> (PSH, 2014c). Oprogramowanie tego serwera jest oparte o kod *OpenSource* – Geoserver i OpenLayers. Aplikację dla potrzeb PSH wykonała firma SMT Software S.A. Baza danych udostępnianych przez ten serwer jest zarządzana jednak przez system komercyjny Oracle 11 (PSH, 2013h). Metodyka i procedura testowania była taka sama jak w przypadku pierwszego serwera e-PSH (rozdział 2.4.1) i z tego względu nie trzeba jej tu opisywać. Bardziej szczegółowe dane uzyskane z wyników testów są podane w tabeli 2.2., opis interfejsu WebMap zawarty jest w instrukcji użytkownika (PSH, 2013h), a tu w skrócie są przedstawione najważniejsze wyniki przeprowadzonych testów:

- początkowy fragment kodu HTML przesyłanego od tego serwera (rys. 2.15); jest napisany w języku podstawowej wersji HTML i ma zdefiniowany sposób kodowania znaków jako UTF-8,
- w tekście kodu (rys. 2.15) są zdefiniowane układy odniesienia obsługiwane przez oprogramowanie agenta: EPSG: 2180 (PUWG 1992), EPSG: 3785 (Popularny wizualizacyjny Merkator dla całego globu), EPSG: 900913 – nieprawidłowe określenie układu EPSG:3857 (WGS84 Web Pseudo-Mercator), EPSG:4326 (WGS 84 – współrzędne geograficzne), EPSG:4179 (Pułkowo 1942(58)), EPSG:4284 (Pułkowo 1942), EPSG:3333 (Pułkowo 1942(58)/Gauss-Kruger strefa 3), EPSG:3334 (Pułkowo 1942(58)/Gauss-Kruger strefa 4), EPSG:3335 (Pułkowo 1942(58)/Gauss-Kruger strefa 5), EPSG:3857 (WGS84 Web Pseudo-Mercator, podany wcześniej, jako EPSG:900913); tak jak w przypadku serwera e-PSH, nie ma jednak na tej liście układów odniesienia wymaganych przez specyfikacją INSPIRE dotyczącą tego zagadnienia: *Data Specification on Coordinate Reference Systems – Technical Guidelines* (INSPIRE TWG-CRS&GGS, 2014a),
- funkcje interfejsu WebMap obrazuje okno przeglądarki FireFox z uruchomionym modulem FireBug (rys. 2.16),
- w tym przypadku również brak jest wielojęzyczności interfejsu WebMap, wymagane są przynajmniej dwa języki – polski i angielski,
- domyślnym układem odniesienia interfejsu WebMap jest wymieniony wcześniej EPSG: 900913 (WGS84 Web Pseudo-Mercator),
- obraz mapy na ekranie jest niskiej jakości, a napisy są często nieczytelne.

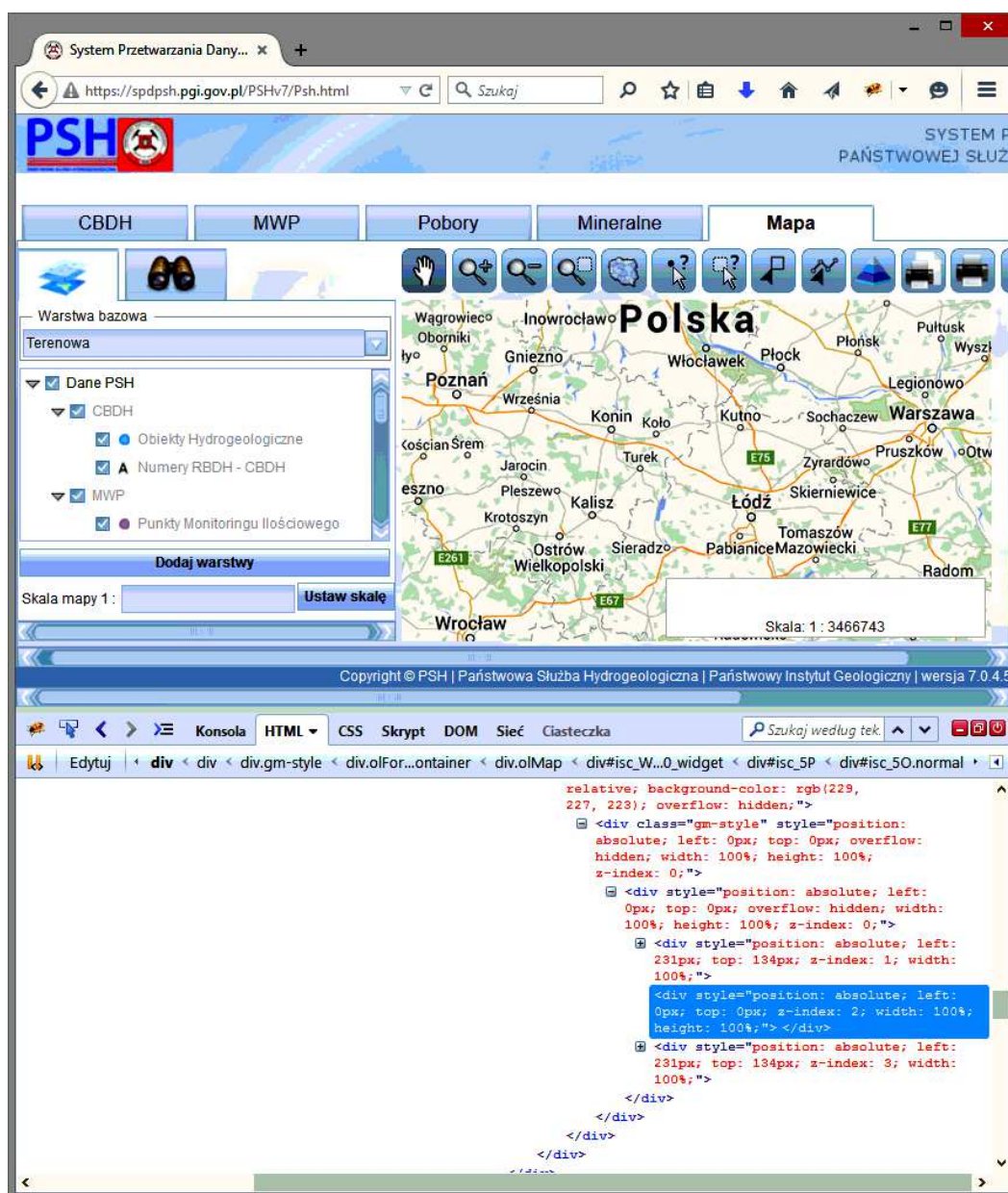
```

1 <!doctype html>
2 <html>
3 <head>
4 <meta http-equiv="X-UA-Compatible" content="IE=EmulateIE7" />
5 <meta http-equiv="content-type" content="text/html; charset=UTF-8">
6 <meta name="author" content="SMT Software S.A.">
7 <meta name="copyright" content="PSH">
8 <link rel="shortcut icon" type="image/x-icon"
9 href="./images/icons/pig.ico">
10 <meta http-equiv="content-script-type" content="text/javascript">
11 <meta http-equiv="content-style-type" content="text/css">
12 <meta name="gtw:property" content="locale=pl">
13 <link type="text/css" rel="stylesheet" href="Psh.css">
14 <title>System Przetwarzania Danych PSH</title>
15 <script type="text/javascript" language="javascript"
16 src="main/main.nocache.js"></script>
17 </head>
18
19 <!-- <script type="text/javascript"
20 src="http://proj4js.org/lib/proj4js.js"></script> -->
21 <script type="text/javascript"
22 src="script/proj4js/lib/proj4js-combined.js"></script>
23 <script type="text/javascript"
24 src="script/OpenLayers.js"></script>
25
26 <script type="text/javascript">
27 /*<![CDATA[*]
28 Proj4js.defs["EPSG:2180"]="+proj=tmerc +lat_0=0 +lon_0=19 +k=0.9993 +x_0=
29 Proj4js.defs["EPSG:3785"]="+proj=merc +a=6378137 +b=6378137 +lat_ts=0.0
30 Proj4js.defs["EPSG:900913"] = "+proj=merc +a=6378137 +b=6378137 +lat_ts=0
31 Proj4js.defs["EPSG:4326"]="+proj=longlat +ellps=WGS84 +datum=WGS84 +no_defs
32 Proj4js.defs["EPSG:4179"] = "+proj=longlat +ellps=krass +towgs84=33.4,-1
33 Proj4js.defs["EPSG:4284"] = "+proj=longlat +ellps=krass +no_defs";
34 Proj4js.defs["EPSG:3333"] = "+proj=tmerc +lat_0=0 +lon_0=15 +k=1 +x_0=350
35 Proj4js.defs["EPSG:3334"] = "+proj=tmerc +lat_0=0 +lon_0=21 +k=1 +x_0=450
36 Proj4js.defs["EPSG:3335"] = "+proj=tmerc +lat_0=0 +lon_0=27 +k=1 +x_0=550
37 Proj4js.defs["EPSG:3857"] = "+proj=merc +a=6378137 +b=6378137 +lat_ts=0.0
38 /*]]>*/
39 </script>
40
41 <script type="text/javascript">
42 var _gaq = _gaq || [];
43 _gaq.push(['_setAccount', 'UA-31548086-1']);
44 _gaq.push(['_trackPageview']);

```

Rysunek 2.15. Początkowy fragment kodu HTML wraz ze skryptami JavaScript agenta przesłanego z serwera SPD-PSH do przeglądarki. Widoczny jest skrypt definiujący układy odniesienia przestrzennego, w tym niepoprawnie nazwany układ EPSG:900913. Opis w tekście.

- obraz wyświetlanej mapy można pobrać i zapisać jako plik graficzny, jednak niestety tylko w nie odpowiednim do tego celu formacie JPEG o niskiej jakości i bez georeferencji,
- nie ma możliwości wyboru dowolnej skali i w takim przypadku pojawia się komunikat „Podana skala nie jest obsługiwana przez podkłady Google Maps. Skala zostanie dostosowana do najbliższej wartości zgodnej z daną warstwą bazową”; takie rozwiązanie wynika z niestandardowej technologii opartej na piramidzie segmentów map



Rysunek 2.16. Okno przeglądarki Firefox przedstawiające interfejs WebMap serwera SPD-PSH.

W dolnej części okna mieści się panel modułu FireBug z fragmentem kodu HTML określającego graficzne elementy tego interfejsu i ich rozmieszczenie. Dla uzyskania przedstawionej w tym oknie mapy interfejs wysyła do różnych serwerów WMS w wersji 1.1.1 aż 1440 poleceń GetMap, z czego tylko 360 uzyskanych odpowiedzi w formie fragmentu mapy jest wykorzystane do utworzenia widocznego w przeglądarce obrazu.

Obraz mapy składa się z 36 warstw, a każda z nich składa się z 40 segmentów (fragmentów mapy).

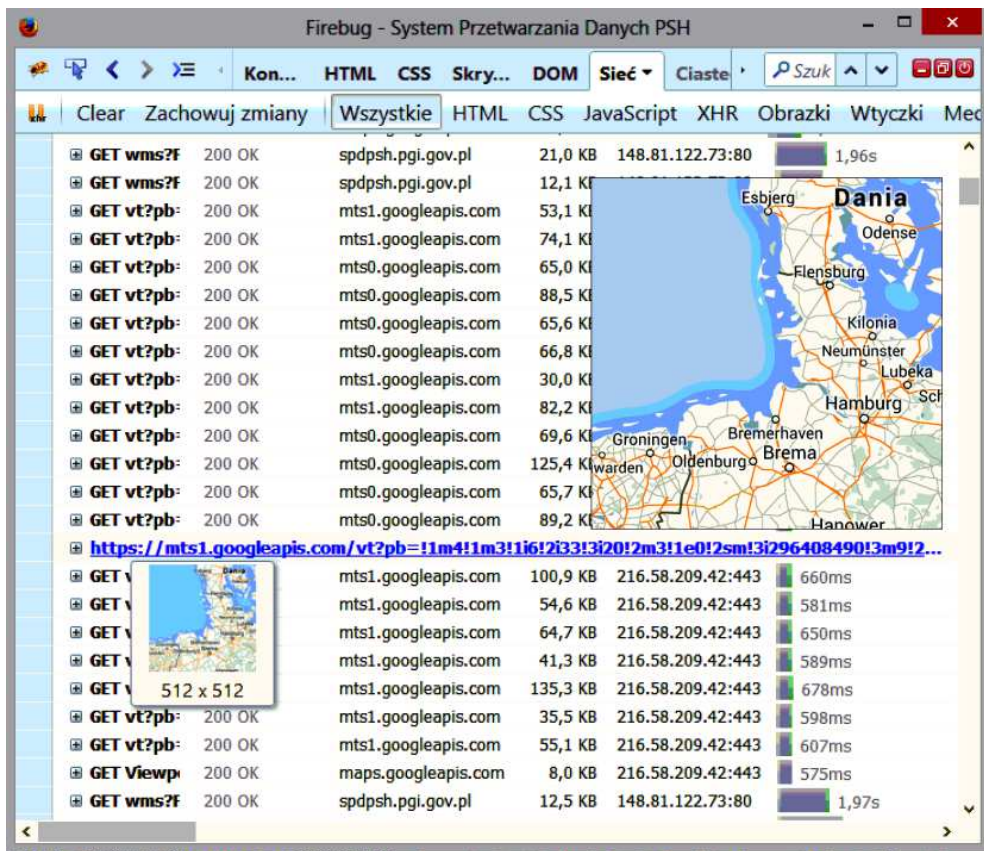
Liczby warstw pozyskiwanych z poszczególnych serwerów:

maps.googleapis.com – 1 warstwa, mts0.googleapis.com – 1 warstwa,

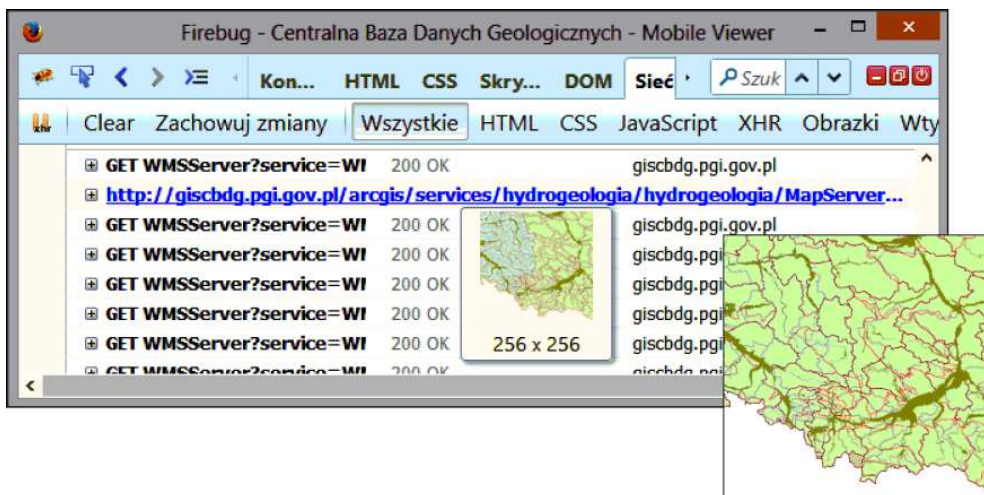
mapy.geoportal.gov.pl – 1 warstwa, spdpsch.pgi.gov.pl – 24 warstwy,

wms.gdos.gov.pl – 6 warstw, epsh.pgi.gov.pl – 3 warstwy.

Bardziej szczegółowy opis zawarty jest w rozdziale 2.4.2.



Rysunek 2.17. Okno modułu FireBug podczas wykonywania przez serwer SPD-PSH poleceń GetMap. Niebieski tekst to niezbędne polecenie dla mapy fragmentu Danii. Opis w tekście.



Rysunek 2.18. Okno modułu FireBug podczas wykonywania przez serwer geologiczny CBDG poleceń GetMap dla JCWP. W tym przypadku również pobierane są segmenty, a układ odniesienia to EPSG: 3857 (WGS 84/Pseudo-Mercator). Opis w tekście.

z określonymi sztywno wartościami skali; to rozwiązanie nie jest poprawne według wymagań określonych w dokumentach INSPIRE; aby było możliwe wybranie dowolnej, w wymaganej przez użytkownika skali, konieczne jest wyłączenie mapy podkładowej, jednak uzyskany obraz traci wiele na wartości i może być w wielu przypadkach bezużyteczny.

Przykład 2.2. Postać polecenia wysłanego przez agenta do serwera spdpsh (odpowiada to jednemu wierszowi typu „GET wms...” na rysunku 2.17) z podziałem na poszczególne parametry:

```
http://spdpsh.pgi.gov.pl/GeoPSHv7/GeoPSHv7/wms?  
FORMAT=image%2Fpng8&  
LAYERS=GeoPSHv7%3AJCWP_GWB&  
TRANSPARENT=TRUE&  
SERVICE=WMS&  
VERSION=1.1.1&  
REQUEST=GetMap&  
STYLES=&  
SRS=EPSG%3A900913  
&BBOX=  
    2522135.1848899,  
    6717643.7926316,  
    2678678.2187962,  
    6874186.8265379&  
WIDTH=256&  
HEIGHT=256
```

- Pomimo, że usługi WMS tego serwera spełniają standardy OGC i normy ISO, a także wymagania INSPIRE, to interfejs użytkownika (WebMap) nie jest zgodny z tymi wymaganiami, bo jest niestandardowym pomieszczeniem dwóch różnych technologii – WMS i WMTS (*Web Map Tile Service*).
- Konsekwencją zastosowania pomieszczenia dwóch technologii jest wyjątkowo duża liczba poleceń GetMap dla wyświetlenia w przeglądarce jednej mapy – często liczba tych poleceń wynosi 1440, a tylko 25% otrzymanych od serwera segmentów mapy jest wykorzystywana do zbudowania obrazu mapy. Przykład zbędnych segmentów jest przedstawiony na rysunku 2.17. W tym przypadku obszar mapy obejmuje jedynie fragment Polski, a z serwera są pobierane odległe segmenty, na przykład przedstawiające fragment Danii.
- Moc sprzętu tego serwera jest zdecydowanie większa niż serwera e-PSH. Z tego powodu tak duże obciążenie poleceniami GetMap nie przejawia się wyraźnie zauważalnym jego przeciążeniem. Z pewnością wpływa na to także stosunkowo nieduża liczba nawiązywanych sesji.
- Przykład 2.2. pokazuje wyraźnie, że poszczególne polecenia GetMap dotyczą małych segmentów mapy (kwadratów o wymiarach 256 pikseli na 256 pikseli) i w dodatku tylko dla jednej warstwy. Wyjaśnia to dlaczego dla jednego obrazu mapy potrzebne jest tak wielka liczba poleceń WMS. Trzeba tu jeszcze zaznaczyć, że segmenty pobierane z Google Maps mają wymiar 4 razy większy: 512 pikseli na 512 pikseli (rys. 2.17) i przez to jest ich 4 razy mniej.

- Segmentacja obrazu mapy na małe fragmenty wynika z zastosowania obrazu podkładu pobieranego z Google Maps, który stosuje tego rodzaju segmentację i jest to rozwiązanie niezgodne ze standardami i normami.
 - W przypadku, gdy podkład pochodzi z innego serwera lub jest pominięty, segmentacja nie jest potrzebna i można by przysłać jednym poleceniem GetMap cały obraz mapy, tak jak to jest realizowane w serwerze e-PSH. Jednak, prawdopodobnie struktura kodu agenta jest ukierunkowana na segmentację i w rezultacie nie pozwala to na zmianę sposobu budowania obrazu mapy.
- Wnioski wynikające z tej analizy przedstawione są w rozdziale 2.4.3.

2.4.3. Porównanie wyników testów

Zbiorne zestawienie najważniejszych wyników prac testowych zawiera tabela 2.2. W podsumowaniu należy postawić dwa podstawowe pytania:

1. W jakim stopniu usługi świadczone przez te serwery są użyteczne?
Odpowiedź: To zależy dla kogo: dla amatora może to być wystarczające, ale mało atrakcyjne, dla profesjonalisty jest to wstępne zapoznanie się z dostępnymi danymi, które do pracy zawodowej powinien móc pobrać. Taka jest podstawowa rola usługi przeglądania (WMS) i w ślad za nią powinna iść usługa pobierania (WFS), której jeszcze nie ma.
 2. Jak należy ocenić oba serwery pod względem spełniania standardów i norm?
Odpowiedź: Serwery te znacznie się różnią i przez to wymagają oddzielnej oceny:
 - serwer e-PSH jest oparty na oprogramowaniu komercyjnym i nieco starszym od drugiego serwera, jednak spełnia on te wymagania wyraźnie lepiej, chociaż ma mniejszy zakres możliwości,
 - serwer SPD-PSH jest wyraźnie nowszy i częściowo oparty na oprogramowaniu *Open Source*. Jego usługi WMS i interfejs WebMap to tylko część większego systemu o nazwie System Przetwarzania Danych. Informacje o jego innych modułach nie są publikowane. System ten w znacznie mniejszym stopniu spełnia wymagania standardów i spowodowane jest to niefortunnym połączeniem dwóch odmiennych technologii – WMS i WMTS. To niestandardowe połączenie sprawiło, że system ten jest obciążony wszystkimi wadami obu technologii i nie wykorzystuje zalet, które posiada każda z nich oddzielnie.
- Zalety technologii WMS:
- możliwość utworzenia i przesłania obrazu mapy jako złożonej kompozycji wielu warstw w wybranej dowolnie skali, w wybranym prostokącie ograniczającym i wybranym układzie odniesienia z jednoczesnym wyborem sposobu zobrazowania,
 - możliwość kaskadowania usług różnych serwerów pracujących w tym standardzie, serwer kaskadowy występuje w roli klienta wobec innych serwerów i innemu klientowi po przetworzeniu może przesłać obraz mapy utworzony z danych pochodzących z innych źródeł.
- Wady technologii WMS:
- każdorazowo obraz mapy jest od podstaw budowany dla każdego oddzielnego polecenia GetMap, ponieważ praktycznie każde takie polecenie jest niepowtarzalne, stanowi to istotne obciążenie serwera.

Zalety technologii WMTS:

- bardzo małe obciążenie serwera, ponieważ na polecenie GetTile wysyłany jest gotowy, wcześniej wygenerowany, obrazek segmentu o z góry ustalonej skali, zasięgu przestrzennym i układzie odniesienia, a także określonej kompozycji warstw i sposobie zobrazowania. Najczęściej dla utworzenia obrazu mapy potrzebne jest kilkanaście segmentów, czyli kilkanaście poleceń GetTile.

Wady technologii WMTS:

- poszczególne segmenty przeznaczone do przesłania do klienta są wcześniej przygotowane, z tego wynika wiele ograniczeń: z góry ustalone skale, ustalony układ odniesienia, ustalona treść mapy i jej zobrazowanie,
- małe obciążenie serwera przekłada się na duże obciążenie klienta (przeglądarki z załadowanym do niej agentem), kod agenta jest znacznie większy i bardziej złożony,
- usługa ta jest mało elastyczna – z powodu wielu sztywnych ograniczeń interoperacyjność serwerów we wspólnej realizacji zadań jest bardzo ograniczona.

3. Kolejne pytanie, które trzeba postawić to, jak obecnie jest daleko do spełnienia wymagań określonych w obowiązujących dokumentach INSPIRE?

Odpowiedź można zawrzeć w kilku punktach:

- serwer e-PSH pomimo że jest starszy jest bliższy spełnienia wymagań INSPIRE,
- serwer SPD-PSH z powodu przyjęcia dziwnej koncepcji połączenia dwóch różnych technologii (WMS i WMTS) spełnia wymagania INSPIRE w niewielkim stopniu,
- dokument INSPIRE *Data Specification on Coordinate Reference Systems – Technical Guidelines* (INSPIRE TWG-CRS&GGS, 2014a) określa niżej wymienione układy odniesienia przestrzennego, które powinny być dostępne w usługach sieciowych INSPIRE, a które niestety nie są uwzględnione w obu serwerach:
 - EPSG:4936 (ETRS89, współrzędne geograficzne, zasięg: cała Europa),
 - EPSG:4937 (ETRS89, współrzędne geograficzne, zasięg: cała Europa),
 - EPSG:4258 (ETRS89, współrzędne geograficzne, poziomy komponent 3D, zasięg: cała Europa),
 - EPSG:3035 (ETRS89 / ETRS-LAEA, odwzorowanie azymutalne równopowierzchniowe, zasięg: cała Europa),
 - EPSG:3034 (ETRS89 / ETRS-LCC, odwzorowanie stożkowe, zasięg: cała Europa),
 - EPSG:3045 (ETRS89 / ETRS-TM33, Transverse Mercator – strefa 33, zasięg: Europa od 12°E do 18°E),
 - EPSG:3046 (ETRS89 / ETRS-TM34, Transverse Mercator – strefa 34, zasięg: Europa od 18°E do 24°E),
- obok wymagań dotyczących zakresu i formy udostępnianych danych istotne znaczenie mają wymagania dotyczące jakości usług (EC, 2010) dotyczące między innymi wydajności, przepustowości i dostępności usługi. System komputerowy realizujący usługi sieciowe INSPIRE, aby te wymagania mógł spełnić, powinien być dedykowany jedynie tym usługom i powinien mieć dostatecznie dużą moc przetwarzania i przesyłania drogą internetową. Wykonywanie równoczesne innych zadań znacznie ogranicza jego możliwości jako serwera usług INSPIRE,
- usługi przeglądania danych (obok usług wyszukiwania) to dopiero pierwszy krok. Kolejnym znacznie trudniejszym krokiem jest utworzenie usług udostępniania danych. W tym przypadku wymagania INSPIRE są ściśle określone i znacznie trudniejsze do

spełnienia, między innymi ze względu na określony w specyfikacjach tematycznych zakres danych, a także ich strukturę i formę.

W dalszych rozdziałach monografii przedstawiona będzie metodyka pozwalająca przekształcić krajowe dane hydrogeologiczne do struktury i formy wymaganej przez dokumenty INSPIRE. Jest to jeden z kluczowych elementów niezbędnych do utworzenia usług udostępniania opartych na technologii WFS (*Web Feature Service*) (ISO/TC211, 2010; Peng, Zhang, 2004; Portale, Van den Broecke, Cox, 2009).

Tabela 2.2. Zestawienie wyników testów przeprowadzonych na obu serwerach danych geoprzestrzennych z zakresu hydrogeologii – e-PSH i SPD-PSH. Opis w tekście.

Rodzaj testu	Serwer e-PSH	Serwer SPD-PSH
Adres interfejsu WebMap	http://epsh.pgi.gov.pl/epsh/	http://spdps.pgi.gov.pl/PSHv7/
Technologia	GeoMedia SDI Portal Intergraph	OpenSource Geoserver i OpenLayers
Wykonawca	Intergraph Polska Sp. z o.o.	SMT Software S. A.
Oprogramowanie serwerowe	Brak danych	Brak danych, system bazy Oracle 11
Zgodność serwera WMS ze standardem	Zgodny	Zgodny
Oprogramowanie agenta	GeoMedia SDI Portal	OpenLayers
Zgodność agenta ze standardem	Zgodny	Niezgodny, niestandardowe pomieszczenie technologii WMS z WMTS
Interfejs serwera	WMS 1.1.1, WMS 1.3.0	WMS 1.1.1, WMS 1.3.0
Interfejs agenta	WMS 1.3.0	WMS 1.1.1
Język kodu agenta	XHTML	HTML
Układ odniesienia przestrzennego w środowisku agenta	EPSG:2180 (PUWG 1992)	EPSG:3857 (WGS 84/Pseudo-Mercator) błędnie określony jako EPSG:900913
Wybór skali	Można ustalić dowolną skalę odpowiednią do zasięgu	Ograniczony tylko do skal piramidy GoogleMaps
Objętość danych ładowanych na startcie	około 60 KB	około 1,6 MB
Średni czas pobierania i ładowania na startcie (5 testów dla ETL)	6,17 s (11,13 s onload)	5,71 s (23,49 s onload)
Liczba wysłanych poleceń GetMap dla 1 kompozycji	15	1440
Liczba wykorzystanych poleceń GetMap dla 1 kompozycji	15	360 (1080 poleceń jest niepotrzebnych)
Liczba warstw	15	36
Liczba pobranych elementów mapy (poleceń GetMap) dla warstwy	1	40
Zaobserwowane błędy	– 6 usług nieczynnych – błąd w skrypcie Java po stronie serwera	– brak 13 elementów grafiki – brak arkusza stylu CSS

3. Podstawy technologiczne infrastruktury INSPIRE

Problematyka technologii infrastruktury INSPIRE była opisana w kilku wcześniejszych publikacjach autora i z tego względu można tu przedstawić jedynie najważniejsze jej aspekty i to w znacznym skrócie odwołując się jednocześnie do tamtych publikacji. Najważniejsze z nich to dwie monografie: *Podstawy metodyczne i technologiczne infrastruktury geoinformacyjnych* (Michalak, 2003c) i *Metodyka i technologia budowy geoserwera tematycznego jako komponentu INSPIRE* (Michalak, 2010a).

Trudno zdefiniować w jednym zdaniu, czym jest infrastruktura informacji geoprzestrzennej. Można jedynie wyliczyć, co nie jest infrastrukturą, z czego się składa infrastruktura i jakie ma główne cechy:

- infrastrukturą nie jest pewna liczba serwerów geoinformacyjnych spełniających standardy, ale nie współdziałających ze sobą na ściśle określonych zasadach interoperacyjności,
- infrastrukturą nie są bazy danych geoprzestrzennych, bez względu na ich formę organizacyjną i zastosowane w nich technologie, nawet gdy możliwy jest pomiędzy nimi przepływ danych,
- infrastrukturą nie są także poszczególne mapowe portale internetowe, takie jak Google Maps lub Geoportal.gov.pl,
- podstawowymi elementami logicznymi infrastruktury informacji geoprzestrzennej są węzły, które wyposażone są w interfejsy usług sieciowych dotyczących danych geoprzestrzennych,
- najogólniej usługi sieciowe dzielą się na cztery kategorie usług: wyszukiwania na podstawie metadanych, przeglądania danych, pobierania i przetwarzania (przekształcania).

Infrastruktura INSPIRE jest przypadkiem wyjątkowym – wynika to z faktu, że obejmuje wiele krajów będących członkami Unii Europejskiej, ale nie tylko. Konsekwencją tego jest wiele języków stosowanych do zapisu danych tekstowych, różne w różnych krajach zakresy tematyczne, różne modele danych i różne formy ich zapisu. Wyjątkowo trudnym zadaniem było znalezienie „wspólnego mianownika” pozwalającego na ujednoczenie zakresu, struktury i formy zapisu, tak aby objąć ją jak najwięcej danych z zachowaniem ich użyteczności, przy jednoczesnym spełnieniu wymagań aktualnych standardów umożliwiających interoperacyjność (Michalak, 2003d). Cel ten został osiągnięty w znacznym stopniu, jednak nie do końca i z pewnością potrzebne są dalsze prace, które doprowadzą do opublikowania nowych ulepszonych wersji specyfikacji technicznych INSPIRE.

3.1. Usługi geoinformacyjne infrastruktury

Istotą infrastruktury są standardowe sieciowe usługi danych geoprzestrzennych realizowane przez jej węzły i zorganizowane w łańcuchy: usługi wyszukiwania, przeglądania, pobierania i przetwarzania danych geoprzestrzennych. Pojęcie węzła nie należy do kategorii pojęć z zakresu technologii. Jest to raczej pojęcie należące do sfery organizacyjnej – węzeł jest budowany w trybie projektu o charakterze inwestycyjnym, jest utrzymywany i nadzoro-

wany przez określoną instytucję. Najczęściej węzeł jest elementem przestrzennie skupionym – ma jedną lokalizację, może to być tylko jeden system komputerowy w sensie fizycznym. Jednak z punktu widzenia technologicznego węzeł może być rozproszony przestrzennie – jego poszczególne elementy mogą się znajdować w różnych miejscach, także w różnych heterogenicznych systemach w znaczeniu sprzętowym i programowym.

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 opierając się o protokół obsługiwany przez dany interfejs – w tym przypadku system-klient również musi także posiadać odpowiedni dla tego protokołu interfejs. Usługa może być też realizowana pośrednio z zastosowaniem technologii WebMap, poprzez geoinformacyjną witrynę WWW (nazywaną często geoportalem).

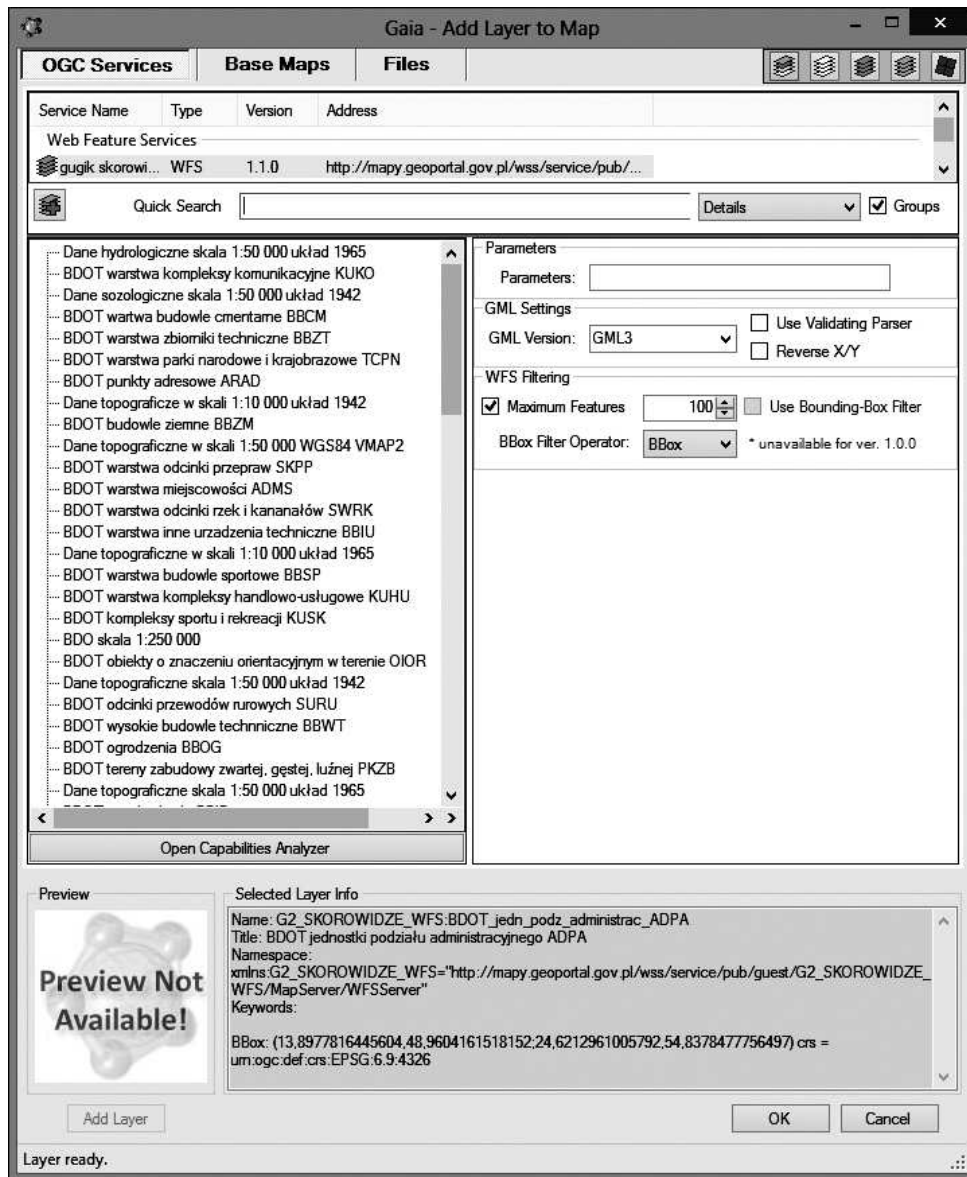
W koncepcji organizacyjnej infrastruktury INSPIRE przyjęto pewien schemat postępowania jej użytkownika w procesie wykorzystania udostępniania danych. Według tej koncepcji proces ten najczęściej dzieli się na cztery etapy: 1 – wyszukiwanie danych, 2 – ich przeglądanie, 3 – pobieranie danych wybranych na etapie przeglądania i 4 – ewentualne ich przetwarzanie. Procesem tym może użytkownik sterować „ręcznie” przez podejmowanie kolejnych decyzji, ale może to być także proces do pewnego stopnia „zautomatyzowany”, w takim przypadku występuje jeszcze jedna usługa: 5 – wywoływania innych usług. Poniżej przedstawiono krótką charakterystykę poszczególnych usług.

Usługa wyszukiwania (usługa katalogowa) oparta jest na standardzie CSW (*Catalog Service for the Web*) (OGC, 2007e) a przetwarzanymi i udostępnianymi informacjami są metadane (dane opisujące zbiory danych i usługi). W przypadku tej usługi interoperacyjność serwerów katalogowych jest szczególnie ważna, ponieważ meta-informacje potrzebne w jednym miejscu są najczęściej rozproszone po wielu serwerach. W takich przypadkach stosowany jest harvesting – przeszukiwanie innych serwerów CSW i pobieranie rekordów metadanych dla umożliwienia przeszukiwania lokalnego. Innym rozwiązaniem jest przeszukiwanie rozproszone, w którym polecenie znalezienia metadanych jest interoperacyjnie przekazywane do innych serwerów a do klienta zwracany jest wynik zbiorczy.

Usługa przeglądania danych może być realizowana za pomocą dwóch różnych standardów: WMS (*Web Map Service*) (OGC, 2006) o wysokiej jakości wyników lub WMTS (*Web Map Tile Service*) (OGC, 2010b) o uproszczonym sposobie przedstawiania wyników. Porównanie tych dwóch technologii jest przedstawione w rozdziale 2. W przypadku przeglądania danych również bardzo często uczestniczą różne serwery, ponieważ do oceny przeglądanych danych potrzebne jest porównanie z danymi z innych węzłów (serwerów) infrastruktury, między innymi poprzez kaskadowanie tej usługi, co również wymaga interoperacyjności.

Do realizacji usługi pobierania stosuje się dwa standardy technologiczne: WFS (*Web Feature Service*) dla danych wektorowych (OGC, 2009c) udostępnianych w zapisie GML (*Geography Markup Language*) (OGC, 2007c) i WCS (*Web Coverage Service*) dla danych typu pokrycie (OGC, 2009b) w formie siatek (*matrix data*). Usługa pobierania WFS praktycznie w Polsce nie funkcjonuje, z jednym wyjątkiem geoportalu GUGiK, gdzie jest zainstalowanych 17 interfejsów tej usługi i 8 z nich jest zadeklarowanych jako dostępne publicznie. Jednak dla tych ostatnich można jedynie pobrać plik *Capabilities* (rysunek 3.1.), a próba pobrania danych kończy się komunikatem „Serwer zdalny zwrócił błąd: (401) Bez autoryzacji” (przykład 3.1.). Pozostałe 9 interfejsów tej usługi jest zabezpieczonych hasłem i w rezultacie publiczny dostęp jest praktycznie niemożliwy.

Druga usługa przeznaczona do pobierania danych macierzowych (WCS) w Polsce jeszcze nie była stosowana, podobnie jak opisane dalej usługi związane z przetwarzaniem danych, a z dostępnej literatury dotyczącej tego zagadnienia nie jest to także w planach. Usługi pobierania danych i przetwarzania (przekształcania) są najczęściej realizowane przez oddzielne pojedyncze serwery, lecz i w tych przypadkach zdarzają się sytuacje wymagające współdziałania.



Rysunek 3.1. Okno analizatora przeglądarki Gaia WFS przedstawiające wynik otrzymany po wysłaniu polecenia `GetCapabilities` na adres `http://mapy.geoportal.gov.pl/wss/service/pub/guest/G2_skorowidze_WFS/MapServer/WFSServer?service=wfs&request=getcapabilities`.

Przykład 3.1. Komunikat przeglądarki WFS Gaia o odmowie odpowiedzi na wysłane do serwera mapy.geoportal.gov.pl polecenie `GetFeatures`.

```
-----[ Message #1: ]-----
CarbonTools.CarbonToolsException: Reading from WFS failed
[Source URL =
http://mapy.geoportal.gov.pl/wss/service/pub/guest/G2_SKOROWIDZE_WFS/MapServer/WFSServer]
[Layer name = G2_SKOROWIDZE_WFS:BDOT] ---> System.Net.WebException:
Serwer zdalny zwrócił błąd: (401) Bez autoryzacji.
w System.Net.HttpWebRequest.EndGetResponse(IAsyncResult asyncResult)
w CarbonTools.Content.OCG.WFS.HandlerWFS.AsyncReader(IAsyncResult result)
--- Koniec śladu stosu wyjątków wewnętrznych ---
```

Usługi przetwarzania danych geoprzestrzennych (OGC, 2010a) dzielą się na kilka kategorii, a do najczęściej stosowanych należą:

- CTS (*Coordinate Transformation Service*) – przeznaczony do przeliczania danych wektorowych pomiędzy różnymi układami odniesienia przestrzennego (OGC, 2001),
- WPS (*Web Processing Service*) – otwarty ramowy standard przeznaczony do przetwarzania zarówno w zakresie prostych algorytmów obliczeniowych, jak również do specyficznych zadań z zakresu modelowania i symulacji przestrzenno-czasowych WPS (OGC, 2007d),
- WCPS (*Web Coverage Processing Service*) (OGC, 2009a; 2009b; 2009d) – specyfikacja standardowego interfejsu definiującego język niezależny od protokołu, przeznaczony do pobierania, przetwarzania i analizy wielowymiarowych siatkowych wyróżnień przedstawiających dane pochodzące z sensorów, zawierające obrazy lub statystyki. Usługi wykorzystujące ten język mają zapewnić dostęp do oryginalnych lub przetworzonych zbiorów danych geoprzestrzennych typu pokrycia w formie odpowiedniej do zobrazowania po stronie klienta, także jako dane wejściowe do modeli naukowych i innych specjalistycznych aplikacji klienckich.

Usługi przetwarzania danych są wymienione w specyfikacjach INSPIRE, jednak nie ma tam żadnych szczegółowych ustaleń i w konsekwencji tego należy je traktować jako rozszerzenia funkcji infrastruktury w przyszłości.

Z punktu widzenia krajowych zadań, dotyczących danych hydrogeologicznych, istotnym zagadnieniem jest dbanie o to aby poszczególne aplikacje (oprogramowanie interfejsów tych usług) było w pełni zgodne z wymaganiami określonymi w wyszczególnionych powyżej standardach.

3.2. Proces opracowania modelu danych hydrogeologicznych INSPIRE

We wczesnym etapie prac dotyczących infrastruktury INSPIRE, tematyka hydrogeologiczna miała stanowić część tematu Hydrografia wymienionego w aneksie pierwszym dyrektywy. Wynikało to ze związków wód podziemnych z wodami powierzchniowymi i należy przypuszczać, że takie rozwiązanie byłoby korzystniejsze. Jednak ostatecznie zdecydowano, że problematyka hydrogeologiczna będzie podtematem tematu Geologia, razem z dwoma innymi podtematami – Geologia (właściwa) i Geofizyka.

Konsekwencją tej zmiany jest fakt, że model hydrogeologiczny nie ma żadnych powiązań z modelem hydrograficznym i w rezultacie nie ma możliwości opierając się o dane INSPIRE ustalenia, jakie istnieją związki pomiędzy wyróżnieniami wód podziemnych z wyróżnieniami wód powierzchniowych, występujących w bezpośrednim sąsiedztwie. Inną niekorzystną

konsekwencją jest uzależnienie modelu hydrogeologicznego od modelu geologicznego, co niesie ze sobą konieczność zapisywania w zbiorach danych hydrogeologicznych wiele najczęściej niepotrzebnych elementów. Bardziej szczegółowy opis tego modelu i jego składników: zawartych w nim klasy, ich atrybutów, powiązań pomiędzy klasami, a także enumeratorów i list kodowych, jest zawarty w rozdziale 6.

Opracowanie modelu danych hydrogeologicznych INSPIRE przebiegało z wieloma perturbacjami i przyczyną tego był brak specjalistów z zakresu hydrogeologii w zespole pracującym nad tematem Geologia (*INSPIRE Thematic Working Group Geology and Mineral Resources*). W początkowym okresie prac nad modelem hydrogeologicznym przyjęto model danych bazujący na modelu języka GWML v. 1.0 (*GroundWater Markup Language version 1.0*) opracowany przez zespół kanadyjski, w którym również nie było kompetentnych hydrogeologów. Zarówno model kanadyjski, jak i pierwsza wersja modelu INSPIRE, były rozwinięciem dla wód podziemnych modelu języka GeoSciML (*Geoscience Markup Language*) dedykowanego podstawowej geologii, a w szczególności stratygrafii i kartografii geologicznej. Takie podejście z uzasadnionych powodów spotkało się z krytyką środowiska hydrogeologów. W rezultacie utworzono oddzielną trzyosobową grupę składającą się z jednego członka zespołu tematycznego INSPIRE i był nim Tomasz Nałęcz z Polski i zaproszonych dwóch hydrogeologów spoza zespołu – Bernharda Wagnera z Niemiec i Janusza Michalaka (autora tej monografii) z Polski. W wyniku wielomiesięcznej dyskusji powstała nowa wersja modelu hydrogeologicznego INSPIRE, uwzględniająca większość potrzeb w zakresie danych hydrogeologicznych. Model ten był niezależny od modelu dotyczącego danych geologicznych, jak to było na wstępie, a jedynie odwoływał się do elementów modelu geologicznego w przypadkach gdy to było uzasadnione, na przykład do opisu geologicznego jednostek hydrogeologicznych. Kierownictwo zespołu tematu Geologia uznało jednak, że model ten zawiera zbyt wiele elementów i w konsekwencji musi być znacznie ograniczony. Powstała nowa koncepcja polegająca na podziale modelu ogólnego na dwie części – część podstawową (*core*) jako obowiązkową i część rozszerzoną (*extension*) jako nieobowiązkową. Dalsze prace nad tym modelem były skoncentrowane nad modelem podstawowym, ponieważ wymagały tego napięte terminy, a model rozszerzony miał być opracowany w drugiej kolejności. Tak się jednak nie stało ponieważ, po zakończeniu prac nad częścią podstawową i opracowaniu dokumentu specyfikacji, zespół tematu Geologia jako całość zawiesił swoją działalność. Dodatkowo na koniec, kierownictwo zespołu reprezentowane przez geologów dokonało istotnych zmian w modelu hydrogeologicznym, polegających na uzależnieniu go od modelu geologicznego (rys. 2.2.) i usunięciu istotnych elementów, a także usunięciu ważnych powiązań z modelami innych tematów. Jedynie dwa z tych powiązań zostało zachowanych i są przedstawione na rysunkach 2.1. i 2.3. Ostateczna wersja modelu jest opisana w specyfikacji (INSPIRE TWG-GE-MR et al., 2013) i w krajowej publikacji (Nałęcz, Michalak, 2011).

Z tego co zostało przedstawione wcześniej trzeba uznać, że prace nad hydrogeologicznym modelem INSPIRE nie zostały zakończone i istnieje potrzeba prowadzenia dalszych prac nad częścią rozszerzoną, która powinna zawierać w szczególności następujące elementy:

- wyróżnienia ogólne: GH_HydrogeologicFeatureAppearance, HGA_HydrogeologicFeature, HGA_HydraulicBoundary, HGA_HydrogeologicEnvironment, HGA_GroundWaterFlowSystem,
- wyróżnienia jako typy ośrodków wodonośnych: HGE_PorousAquifer, HGE_FracturedAquifer, HGE_KarsticAquifer, HGE_CompoundAquifer,

- rozszerzoną listę typów danych dla hydrogeologii: `HG_GridCoverageGeometry`, `HG_Geometry`, `HG_VulnerabilityToPollution`, `HG_HorizontalExtent`, `HG_DynamicCategory`, `HG_DynamicQuantity`, `HG_DynamicQuantityRange`, `HG_DynamicBoolean`, `HG_DynamicCount`, `HG_Identifier`, `HGW_HydrogeologicalProfile`,
- rozszerzoną listę obiektów sztucznych: `HGO_Well`, `HGO_AbstractionWell`, `HGO_InjectionWell`, `HGO_ObservationWell`, `GW_MonitoringStation`, `GW_MonitoringStationDescription`, `GW_MonitoringInternationalNetwork`, `GW_Intake`,
- obiekty naturalne: `Spring`, `SwallowHole`,
- strefy związane z ujęciami: `ZoneOfContribution`, `ZoneOfInfluence`, `ProtectionZone`, `ConeOfDepression`, `GW_ProtectedAreaDetails`,
- pakiet modelu dla systemów krasowych: `KarsticSystem`, `KarsticAquifer`, `KarsticSpring`, `SwallowHole`, `KarsticChannel`, `Cave`, `KarsticGroundWaterBody`, `KarsticStream`,
- pakiet modelu dla wód mineralnych: `MineralWaterAquifer`, `MineralSpring`, `MineralGroundWaterBody`, `MineralWaterPurpose`, `MineralWaterType`, `MineralWaterAbstractionWell`,
- pakiet modelu dla wód termalnych: `ThermalWaterAquifer`, `ThermalGroundWaterBody`, `ThermalSpring`, `ThermalWaterAbstractionWell`, `ThermalWaterPurpose`,
- pakiet modelu dla danych związanych z Ramową Dyrektywą Wodną: `WFD_ProtectedArea`, `WFD_GroundWaterBody`, `WFD_GW_MonitoringStation`.

W podsumowaniu trzeba tu zaproponować opracowanie nowego modelu dla danych hydrogeologicznych – bardziej kompletnego i szczegółowego, a także uwzględniającego różne aspekty prac i badań, w wyniku których tworzone są dane przeznaczone do ponownego wykorzystania. W nawiązaniu do tego w rozdziale 5. porównane zostały różne już istniejące modele możliwe do zastosowania w hydrogeologii, a w rozdziale 5.2. przedstawiony jest wstępny zarys projektu modelu dla języka HydroGeoML, który mógłby pełnić również rolę rozszerzenia podstawowego modelu INSPIRE (*core*) jako dodatkowy moduł *extension*, który był planowany, ale ostatecznie nie został opracowany.

4. Krajowe dane hydrogeologiczne

Pojęcie krajowe dane hydrogeologiczne jest bardzo ogólne i przez to niejednoznaczne. Można tu wyodrębnić trzy kategorie:

- dane pozyskiwane bezpośrednio z prac hydrogeologicznych (terenowych, laboratoryjnych i kameralnych); do nich należą dane o odwiertach, pomierzone na posterunkach obserwacyjnych, dane kartograficzne i wiele innych pochodzących z różnych obserwacji i pomiarów, a także przetwarzania danych pierwotnych w systemach komputerowych, w tym modelowania hydrogeologicznego;
- dane z innych dziedzin zajmujących się skorupą ziemską i zjawiskami w niej i na niej zachodzącymi, które stanowią niezbędny komponent danych przetwarzanych przez hydrogeologów; przykładem może być model powierzchni terenu (DEM – *Digital Elevation Model*) niezbędny do przetworzenia danych o głębokości od powierzchni terenu w dane w postaci rzędnych nad poziom morza;
- dane z innych dziedzin wykorzystywane przez hydrogeologów do analiz i ocen oddziaływania otoczenia na wody podziemne i odwrotnie oddziaływania wód podziemnych i zmian warunków hydrogeologicznych na środowisko; prostym przykładem może być mapa topograficzna.

Trzeba też zwrócić uwagę, że istotą infrastruktury INSPIRE jest wzajemne dzielenie się informacją przestrzenną pomiędzy różnymi środowiskami, które z tą informacją mają do czynienia. Twórcy danych (instytucje, organizacje, przedsiębiorstwa) z danego zakresu tematycznego przechowują i udostępniają dane użytkownikom z innych środowisk, a w zamian mają dostęp do danych tworzonych przez te inne środowiska. Dane są przechowywane w jednym miejscu, tam gdzie powstały i nie tworzy się ich kopii wtórnych gdzie indziej. Takie reguły muszą być przestrzegane dla zachowania ich integralności i spójności wzajemnych powiązań, w szczególności pomiędzy zbiorami danych należących do różnych tematów.

Nie wszystkie dane z zakresu hydrogeologii są udostępniane na podanych powyżej zasadach. Znaczna ich część nie może być udostępniana, ponieważ jest to ograniczone wymogami formalno-prawnymi lub organizacyjnymi, na przykład mają formę nieodpowiednią do udostępniania lub są surowe – niedostatecznie przetworzone.

Na temat danych hydrogeologicznych dostępnych jest wiele publikacji, głównie pochodzących z Państwowego Instytutu Geologicznego PIB (z PSH) i z tego względu nie ma potrzeby opisywania ich tu szczegółowo. Na stronach witryny Instytutu można dowiedzieć się o dostępnych zbiorach danych, a także o sposobie i warunkach ich udostępniania (PSH, 2012b, 2013c, 2014a) W dalszej części wymienione zostaną poszczególne kategorie danych i ich instancje, ze szczególnym uwzględnieniem formy w jakiej występują.

4.1. Mapa hydrogeologiczna Polski – MHP

Mapa Hydrogeologiczna Polski (MHP) w skali 1:50 000 jest najważniejszym i największym projektem z zakresu kartografii hydrogeologicznej. Jej koncepcja powstała na początku lat dziewięćdziesiątych, a realizacja pierwszej wersji tej mapy przypadła na lata 1996 do

2004 (PSH, 2013d). Była to mapa w podziale arkuszowym liczącym 1069 arkuszy i obejmowała obszar całej Polski (Fert, Mordzonek, Węglarz, 2011). Ta pierwsza edycja mapy była ukierunkowana na rozpowszechnianie jej w formie papierowej, a ze względu na trudne do przewidzenia zapotrzebowanie na poszczególne arkusze, była drukowana na ploterze.

Opracowana w latach dziewięćdziesiątych koncepcja mapy zakładała jej maksymalną przydatność praktyczną, co zaowocowało tym, że mapa ta dotyczyła jedynie głównego użytkowego poziomu wodonośnego i zawierała podstawowe informacje potrzebne do projektowania studni:

- zasięgi i głębokości występowania tego poziomu oraz jego miąższości i przewodności,
- dane o jakości wód pod kątem źródła zaopatrzenia ludności w wodę,
- informacje o stopniu zagrożenia wód zanieczyszczeniami z powierzchni terenu,
- przewidywaną wydajność z typowej studni,
- przybliżone położenie zwierciadła i kierunek przepływu,
- informacje o odnawialności zasobów wód podziemnych.

Tak wąsko określony cel tej mapy ogranicza poważnie jej zastosowanie jako źródła danych dla opracowywania modeli przepływu, co obecnie jest już powszechnie stosowane przy projektowaniu studni.

Pierwsza edycja MHP była opracowywana w technologii MGE/Microstation firmy Intergraph z przechowywaniem danych w bazie zarządzanej systemem Oracle, a układem odniesienia przestrzennego był stary układ Pułkowo 1942.

Przez wiele kolejnych lat projekt MHP był rozwijany i obecnie składa się z trzech modułów:

- Użytkowy Poziom Wodonośny,
- Pierwszy Poziom Wodonośny – Występowanie i Hydrodynamika,
- Pierwszy Poziom Wodonośny – Wrażliwość na Zanieczyszczenie i Jakość Wód.

4.1.1. Użytkowy Poziom Wodonośny

Moduł ten, określane obecnie akronimem MHP-UPW, wywodzący się z pierwszej wersji, został poddany istotnej reorganizacji i rozbudowie (PSH, 2013d) polegającej na:

- przejściu z formy arkuszowej (wiele oddzielnych baz danych lub plików) na formę ciągłą (cały obszar Polski w jednej bazie),
- zmianie technologii z MGE/Microstation na GeoMedia Professional, również firmy Intergraph, co pozwala na wybieranie i przetwarzanie danych dla dowolnie określonych obszarów ze zbioru danych znajdującego się w jednej bazie,
- dostosowaniu całego zasobu do przyjętych innych rozwiązań, w tym do modelu MPHP (Mapy Podziału Hydrograficznego Polski) i klasyfikacji prezentowania stanu wód,
- prowadzeniu aktualizacji zasobu danych w miarę napływu nowych danych,
- prowadzeniu prac nad rozwojem oprogramowania do wprowadzania danych, ich edycji i prezentacji w postaci wybranych warstw informacyjnych.

4.1.2. Pierwszy Poziom Wodonośny – Występowanie i Hydrodynamika

Dane geoprzestrzenne dotyczące pierwszego poziomu wodonośnego podzielone są na dwa moduły. Pierwszy (MHP PPW-WH) obejmuje występowanie wód podziemnych w pierwszej od powierzchni terenu warstwie wodonośnej i informacje o hydrodynamice tych wód (PSH, 2013e). Drugi moduł opisany w następnym rozdziale 4.1.3. dotyczy jakości wód i wrażliwości na zanieczyszczenia.

Pojęcie „pierwszy poziom wodonośny” jest w tym przypadku zdefiniowany następująco: „warstwa lub zespołu warstw wodonośnych wykazujących dobrą łączność hydrauliczną, osiągających łączną miąższość co najmniej 2 m przy średnim stanie retencji wód podziemnych”.

Tak jak poprzedni moduł mapy MHP (MHP-UPW) moduł ten (MHP PPW-WH) jest opracowywany w cięciu arkuszowym i w starym układzie odniesienia przestrzennego Pułkowo 1942. W latach 2004–2011 opracowano 684 arkusze, co odpowiada pokryciu 64% całego obszaru kraju.

Z instrukcji opracowywania arkuszy tego modułu wynika, że jego treść jest ukierunkowana na określenie związków hydraulicznych wód tego poziomu z wodami powierzchniowymi, wpływem działalności gospodarczej na te wody, a także oddziaływaniem poziomu wód na ekosystemy od niego zależne. Główne komponenty tematyczne to:

- wyodrębnienie pierwszego poziomu wodonośnego, w tym określenie jego rodzaju i stratygrafii,
- podział na podjednostki hydrodynamiczne i geomorfologiczne,
- hydrodynamika wód tego poziomu i charakterystyka zwierciadła wody,
- głębokość występowania pierwszego poziomu,
- granice obszarów objętych zasięgiem znaczących zmian zwierciadła wywołanych czynnikami gospodarczymi,
- związki hydrauliczne z wodami powierzchniowymi.

4.1.3. Pierwszy Poziom Wodonośny – Wrażliwość na Zanieczyszczenie i Jakość Wód

Tak jak poprzednie moduły mapy MHP moduł ten (MHP-PPW-WJ) opracowywany jest w cięciu arkuszowym i w układzie odniesienia przestrzennego Pułkowo 1942. W latach 2006–2013 opracowano 390 arkuszy, co odpowiada pokryciu 36% całego obszaru kraju (PSH, 2013e). Prace nad arkuszami tego modułu stanowią kontynuację prac dotyczących rozpoznania i charakterystyki pierwszego poziomu wodonośnego.

W tym przypadku treść jest ukierunkowana na potrzebę oceny podatności płytkich wód podziemnych na zanieczyszczenia z powierzchni oraz na ocenę stanu jakościowego tych wód, stanowiących najczęściej źródło zaopatrzenia w wodę ludności wiejskiej. Wody te są bezpośrednio związane z ekosystemami wód powierzchniowych, a także z ekosystemami lądowymi – w szczególności z siedliskami wyróżnionymi przez sieć obszarów chronionych, w tym także programu Natura 2000.

Główne komponenty tematyczne z zakresu dotyczącego wrażliwości na zanieczyszczenia to:

- zasięgi przestrzenne jednostek hydrogeologicznych z określeniem warunków występowania pierwszego poziomu wodonośnego,
- wskaźniki wrażliwości wód tego poziomu wodonośnego na zanieczyszczenie,
- obiekty przemysłowe, rolnicze i komunalne oraz inna działalność gospodarcza pogarszająca stan wód tego poziomu wodonośnego,
- obszary znaczącego obniżenia bądź podniesienia zwierciadła wód.

Główne komponenty tematyczne dotyczące jakości wód to:

- zawartość związków azotu w wodach tego poziomu,
- najważniejsze wskaźniki jakości wód w punktach pobrania prób wody.

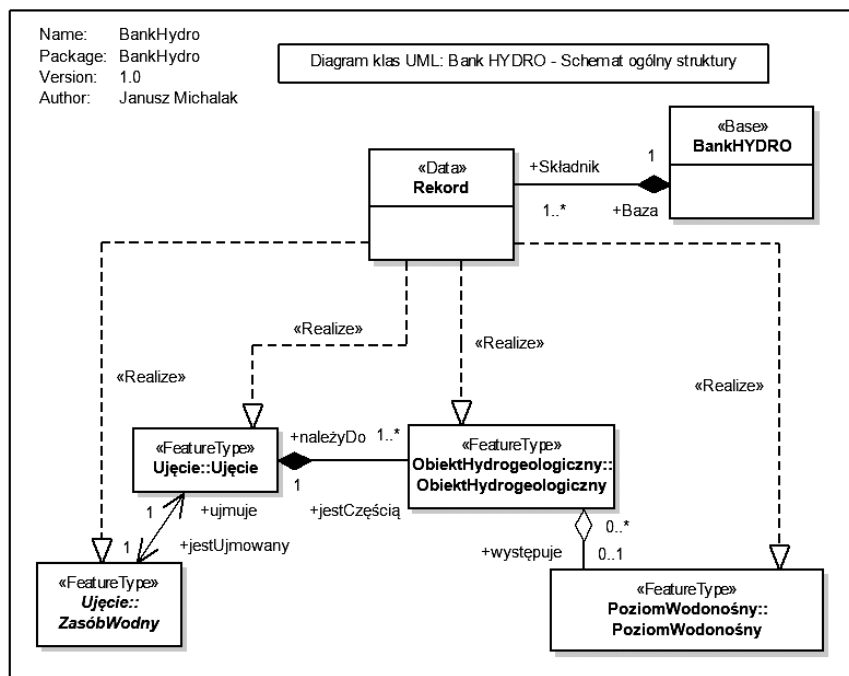
4.2. Bank HYDRO

Bank HYDRO (obecnie ma nową nazwę: Centralny Bank Danych Hydrogeologicznych HYDRO – CBDH) to relacyjna baza danych zarządzana systemem Oracle, przechowująca dane o studniach, ujęciach i innych odwiertach, a także źródłach wód podziemnych zwykłych, mineralnych i termalnych z obszaru kraju. Historia bazy danych o nazwie Bank HYDRO sięga lat siedemdziesiątych i pierwsze jej wcielenie to specjalistyczne dedykowane tej tematyce oprogramowanie w języku Fortran zrealizowane na maszynie cyfrowej Odra. Do danych tych należą (PSH, 2013a):

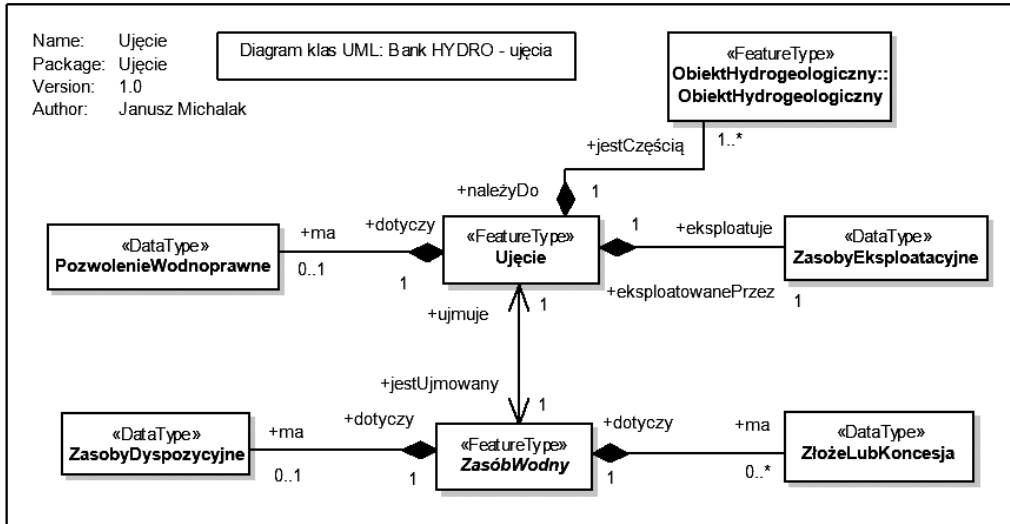
- punktowe odniesienie przestrzenne obiektu hydrogeologicznego (studni, odwiertu, źródła),
- dane hydrogeologiczne pochodzące z pomiarów lub obliczeń,
- dane dotyczące profilu geologicznego i hydrogeologicznego,
- wyniki analiz chemicznych i pomiarów fizycznych pobranych prób wody.

Bank HYDRO zawiera dane o ponad 140 000 obiektach hydrogeologicznych z obszaru całego kraju. Baza jest ciągle aktualizowana, rocznie jest wprowadzanych do niej prawie 2 tysiące nowych rekordów. Podobnie jak inne zbiory danych gromadzonych przez PSH dane Banku HYDRO mają zastosowanie do celów projektowych i dokumentacyjnych w dziedzinie hydrogeologii, geologii, gospodarki wodnej i ochrony środowiska, służą także administracji rządowej, samorządowej oraz służbom państwowym jako system wspomaganie decyzji w zarządzaniu środowiskiem.

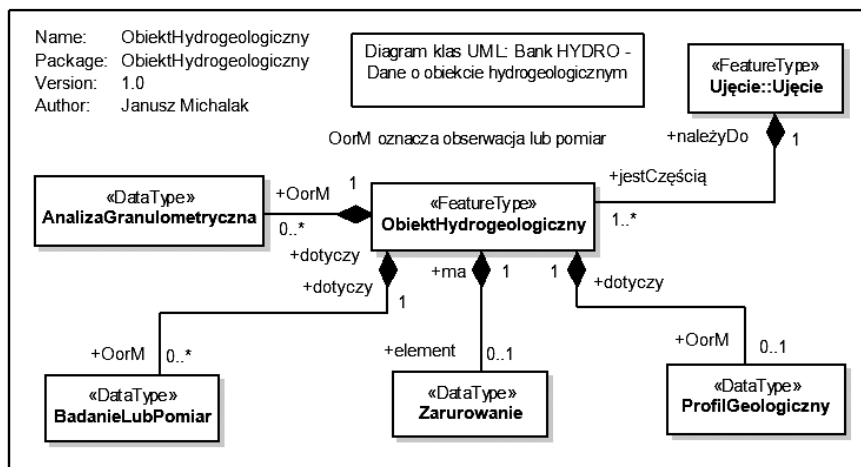
Zarys struktury banku HYDRO przedstawiają diagramy klas UML zawarte na rysunkach 4.1., 4.2., 4.3. i 4.4.



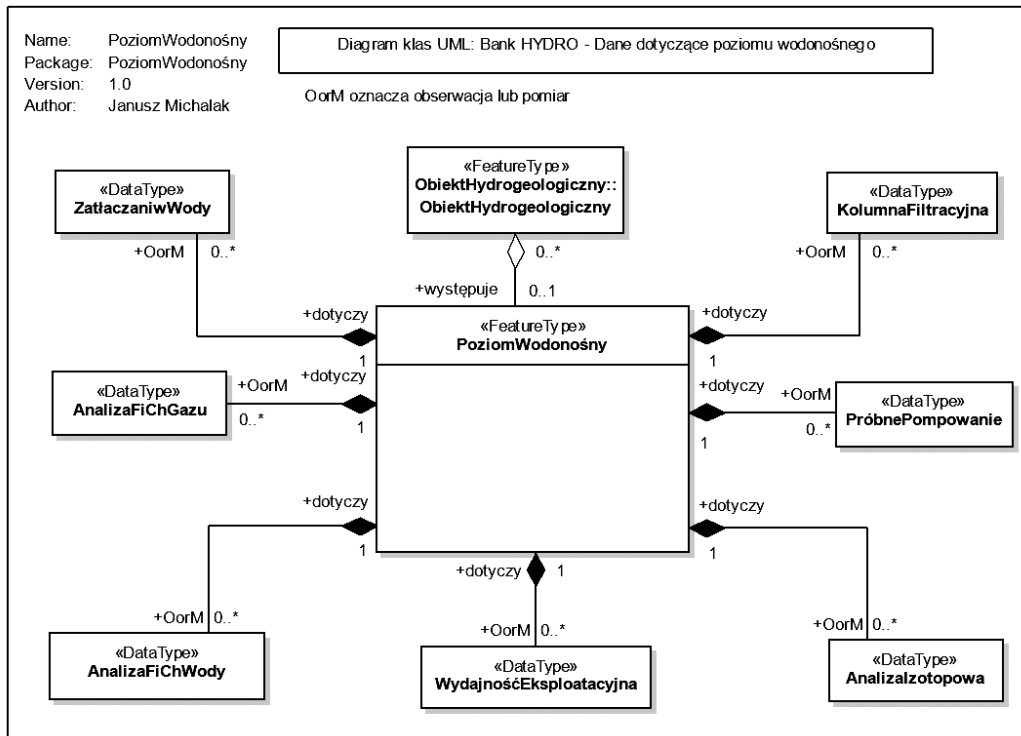
Rysunek 4.1. Diagram klas UML przedstawiający ogólną strukturę bazy danych Bank HYDRO. Opracowano na podstawie materiałów PSH.



Rysunek 4.2. Diagram klas UML przedstawiający strukturę danych bazy Banku HYDRO dotyczących ujęć. Opracowano na podstawie materiałów PSH.



Rysunek 4.3. Diagram klas UML przedstawiający strukturę danych bazy Banku HYDRO dotyczących obiektów hydrogeologicznych. Opracowano na podstawie materiałów PSH.



Rysunek 4.4. Diagram klas UML przedstawiający strukturę danych bazy Banku HYDRO dotyczących poziomów wodonośnych. Opracowano na podstawie materiałów PSH.

4.3. Baza danych Monitoring Wód Podziemnych

Kolejną bazą mającą istotne znaczenie badawcze i praktyczne jest baza danych Monitoringu Wód Podziemnych (Baza MWP) (Kazimierski, Kulczyńska, Sadurski, Skrzypczyk, 2011). Z materiałów zawartych na stronie poświęconej tej bazie w witrynie PSH (PSH, 2012a) dowiadujemy się, że jest to baza „zawierająca charakterystykę punktów badawczych, wyniki pomiarów wahań zwierciadła wód podziemnych i wydajności źródeł (monitoring ilościowy) oraz wyniki analiz chemicznych wód podziemnych (monitoring chemiczny)”. Główne zadania bazy Monitoring Wód Podziemnych to gromadzenie i udostępnianie danych dotyczących sieci i punktów badawczych monitoringu, pomiarów zwierciadła, wyników analiz chemicznych oraz wspomaganie oceny stanu wód podziemnych w zakresie ich ilości i jakości.

Baza MWP została utworzona w roku 2004 z połączenia dwóch istniejących wcześniej baz – bazy SOH (System Obserwacji Hydrogeologicznych) z bazą MONBADA – MONitoringowa Baza Danych. Pierwsza z nich przechowywała dane dotyczące punktów obserwacyjnych oraz wyników obserwacji wahań zwierciadła wody. Były to pomiary głębokości zwierciadła w otworach obserwacyjnych lub wielkości wypływu w przypadkach źródeł, a także analizy fizykochemiczne wykonywane okresowo.

Analizując metadane udostępniane przez usługi katalogowe PIG-PSH można ustalić, że baza MWP składa się z kilku modułów:

- podstawowej bazy danych monitoringu wód podziemnych,
- „automatycznej” bazy danych monitoringu wód podziemnych,
- bazy danych monitoringu diagnostycznego,
- bazy danych monitoringu operacyjnego,
- bazy danych monitoringu granicznego,
- bazy danych lokalnego monitoringu badawczego,
- bazy danych monitoringu ilościowego,
- bazy danych archiwalnych monitoringu chemicznego,
- bazy danych archiwalnych monitoringu ilościowego.

Bardziej wyczerpujące informacje na temat gromadzonych przez PSH danych monitoringu można znaleźć w publikacjach (Kazimierski, Kulczyńska, Sadurski, Skrzypczyk, 2011; PSH, 2012a).

4.4. Główne Zbiorniki Wód Podziemnych

Ze stron internetowych PSH (PSH, 2013b) dowiadujemy się że: „Baza Danych Głównych Zbiorników Wód Podziemnych (GZWP) zawiera klasyfikację GZWP według wykorzystania zasobów, stopnia przeobrażeń antropogenicznych, odporności na zanieczyszczenia, ekonomicznego aspektu zaleceń ochronnych oraz wskaźników opłat wodnych”. W formie kartograficznej dane o GZWP stanowią mapę w skali 1:500 000, i należy przypuszczać, że takiej skali odpowiada dokładność i szczegółowość tych danych.

W obecnej postaci zbiory danych dotyczące GZWP dzielą się na dwie kategorie:

- dla GZWP udokumentowanych szczegółowo i z odniesieniem przestrzennym odpowiadającym skali 1:50 000; zbiorniki należące do tej kategorii zaliczają się głównie do grupy zbiorników wymagających pilnego ustanowienia obszarów ochronnych,
- dla GZWP udokumentowanych przeglądowo i z odniesieniem przestrzennym odpowiadającym skali 1:500 000.

Baza Danych Głównych Zbiorników Wód Podziemnych zawiera szereg klasyfikacji GZWP z uwzględnieniem kryteriów:

- według stopnia wykorzystania zasobów tych zbiorników,
- stopnia przeobrażeń antropogenicznych, głównie spowodowanych poborem wody,
- przeciętnego wskaźnika odporności na zanieczyszczenia,
- ekonomicznego i gospodarczego aspektu zaleceń dotyczących ich ochrony,
- przyjętego wskaźnika opłat za pobór wody.

Zbiory danych dotyczące atrybutów geoprzestrzennych tej bazy podzielone są na kilka podzbiorów, do których należą:

- wydzielenia obszarowe zamknięte granicami udokumentowanych GZWP,
- wydzielenia obszarowe zamknięte granicami nieudokumentowanych GZWP,
- wydzielenia obszarowe zamknięte granicami obszarów ochronnych GZWP,
- wydzielenia określające podział administracyjny kraju,
- wydzielenia liniowe i obszarowe hydrografii.

Bardziej szczegółowe informacje na temat Baza Danych Głównych Zbiorników Wód Podziemnych można znaleźć w publikacjach (PSH, 2013b; Herbich, Kapuściński, Nowicki i inni, 2008).

4.5. Jednolite Części Wód Podziemnych

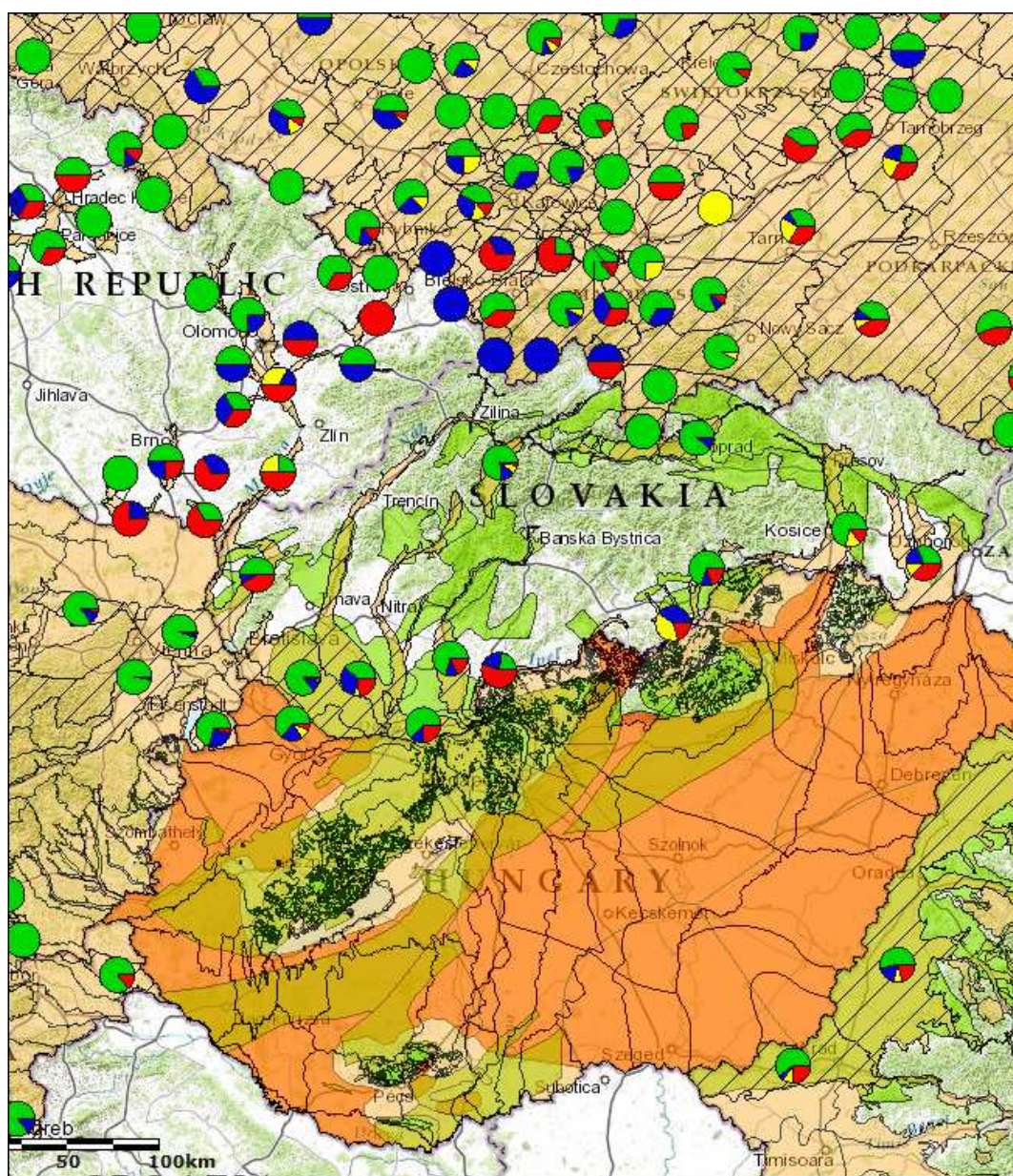
Jednolite części wód Podziemnych (JCWPd – GWB – *GroundWater Body*) to wyróżnienia obszarowe określające rejony występowania wód podziemnych objętych ochroną na zasadach określonych w Ramowej Dyrektywie Wodnej (RDW – WFD – *Water Framework Directive*) Unii Europejskiej. Dyrektywa ta nakłada na kraje członkowskie obowiązek ochrony i poprawy stanu wód podziemnych oraz ekosystemów bezpośrednio od nich zależnych, w szczególności, gdy wody występujące w tych obszarach mają duże znaczenie w zaopatrzeniu ludności w wodę (PSH, 2014d). Wyodrębnienie tych obszarów ma za zadanie umożliwienie indywidualnej oceny jakości i ilości wody występującej w tych rejonach, a także indywidualne opracowywanie i realizowanie programów działań zmierzających do poprawy jakości wód lub zapobiegania procesom powodującym jej pogorszenie (Nowicki i inni, 2009).

Dyrektywa nie określa w jaki sposób należy wyznaczyć granice tych wyróżnień. W różnych krajach członkowskich podejście metodyczne do wyodrębniania JCWPd jest różne (rys. 4.5. i 4.6.) Część państw wybrała metodę niewyodrębniania różnych horyzontów głębokościowych i niepozostawiania obszarów bez przypisania do określonego GWB. Do nich należą między innymi Polska (rys. 4.5.) i Francja (rys. 4.6.). W takim przypadku dla danego kraju wyróżnienia GWB pokrywają całkowicie obszar kraju, tak jak ma to miejsce w przypadku podziału administracyjnego. W innych krajach zastosowano metodę podziału na horyzonty głębokościowe i ograniczenie zasięgu GWB do rzeczywistych zasięgów zbiorników wód podziemnych. Do nich należą między innymi Czechy, Słowacja (rys. 4.5.) i Włochy (rys. 4.6.). W tym przypadku występują obszary, w których brak jest wyróżnień GWB, a w innych może wystąpić ich kilka, ponieważ są położone na różnych głębokościach. Rezultat takiego podejścia w przybliżeniu podobny jest do wyróżnień dla Głównych Zbiorników Wód Podziemnych (GZWP) opracowanych przez polską Państwową Służbę Hydrogeologiczną (rozdział 4.4.).

Bardziej szczegółowy opis zastosowanej w Polsce metodyki wyodrębniania JCWPd jest przedstawiony w innej publikacji monograficznej (Michalak, Nawalany, Sadurski i inni, 2011). Atrybuty opisujące charakterystyki poszczególnych GWB (JCWPd), w tym szczególnie ich stan aktualny pod względem ilości i jakości wód, a także tendencje zmian w tym zakresie, są we wszystkich krajach członkowskich jednakowe, ponieważ jest to jednoznacznie określone w dokumentach dyrektywy WFD.

Rysunki 4.5. i 4.6. przedstawiają fragmenty obszaru Unii Europejskiej z danymi pozyskanymi przez systemy monitoringu, funkcjonujące zgodnie z zasadami określonymi w tej dyrektywie. Wizualizacja tych danych jest udostępniana przez serwisy mapowe bazy WISE-WFD prowadzonej przez EEA (*European Environment Agency*), w tym przypadku obok przestrzennego rozmieszczenia w poszczególnych krajach członkowskich GWB (JCWPd) przedstawione są wyniki pomiarów zawartości amoniaku (jonu amonowego) w wodzie dla poszczególnych GWB w postaci diagramów kołowych (EEA, 2015).

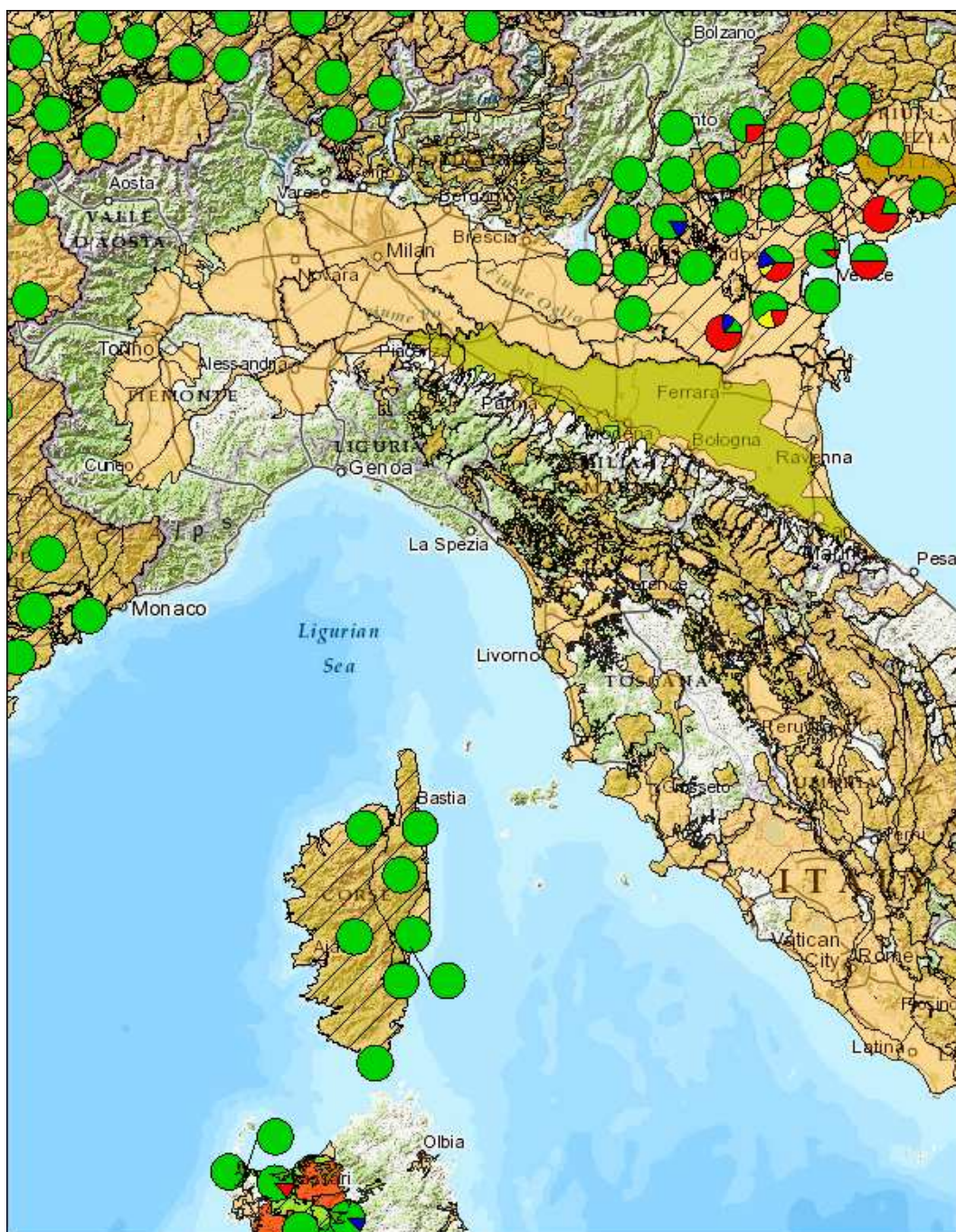
Mapa ta pokazuje średnie roczne stężenie jonu amonowego (NH_4^+) wyrażone w miligramach $\text{NO}_2\text{-N}$ na litr wody, obserwowane w stacjach monitoringu wód podziemnych i zgłoszonych przez kraje członkowskie do EEA poprzez system WISE. Celem mapy jest przedstawienie przeglądowe średniej rocznej wartości amoniaku w wodach podziemnych w całej Europie, a także umożliwienie użytkownikom porównania wartości pomiędzy poszczególnymi krajami lub pomiędzy poszczególnymi GWB (JCWPd) w zależności od skali wizualizacji.



Rysunek 4.5. Fragment mapy przedstawiający zasięgi przestrzenne GWB (JCWPd) w różnych krajach członkowskich Unii Europejskiej (w Polsce, Czechach, na Słowacji, Węgrzech, w Austrii i Rumunii) i widać tu różne podejście metodyczne określania granic. Diagramy kołowe przedstawiają zawartość amoniaku (jonu amonowego) w wodzie dla poszczególnych GWB (JCWPd), gdzie (EEA, 2015):

- | | |
|--|--|
| ■ Klasa 1: $\leq 0,1 \text{ mg NH}_4^+/\text{l}$ | ■ Klasa 2: $> 0,1 \leq 0,3 \text{ mg NH}_4^+/\text{l}$ |
| ■ Klasa 3: $> 0,3 \leq 0,5 \text{ mg NH}_4^+/\text{l}$ | ■ Klasa 4: $> 0,5 \text{ mg NH}_4^+/\text{l}$ |

Objaśnienia w tekście.



Rysunek 4.6. Fragment mapy przedstawiający zasięgi przestrzenne GWB (JCWPd) w różnych krajach członkowskich Unii Europejskiej (we Włoszech, Francji i Austrii), a także w Szwajcarii (EEA, 2015). W porównaniu z rysunkiem 4.5. zwraca tu uwagę znikomy stopień zanieczyszczenia wód podziemnych związkami amonowymi. Objaśnienia jak na rysunku 4.5, a także w tekście.

Wydzielone do tego celu jednostki powierzchniowe pokrywają się częściowo z JCWPd, ale w wielu przypadkach tej zgodności nie ma. Jednak zdaniem autorów koncepcji przyjęty podział pozwala na wyznaczenie zasobów dyspozycyjnych drogą bezpośrednią lub pośrednią dla wszystkich JCWPd.

Od roku 2013 PSH prowadzi prace nad sześcioletnim programem „Wykonanie programów prac i dokumentacji hydrogeologicznych ustalających zasoby dyspozycyjne wód podziemnych dla potrzeb przeprowadzania bilansów wodno-gospodarczych oraz opracowania warunków korzystania z wód regionu wodnego i zlewni”. Program ten jest podzielony na dwa etapy – pierwszy kończy się w roku 2016, a drugi w roku 2018. Po zakończeniu obu etapów cały obszar kraju ma mieć wyznaczone zasoby dyspozycyjne przypisane wyznaczonym wydzieleniom przestrzennym. O ile dla wcześniej opisanych kategorii wydzieleni – GZWP i JCWPd – istnieje dostęp do usługi przeglądania, to w przypadku obszarów (jednostek) bilansowych takiego dostępu (według stanu na początek roku 2015) jeszcze nie ma.

4.7. Pobory

Baza danych Pobory zawiera informacje uzyskiwane z urzędów marszałkowskich o rocznej sumie poboru z ujęcia wód podziemnych. Dane są opracowywane corocznie począwszy od roku 2008, ale baza zawiera także dane archiwalne z lat 2000–2005 i dane pozyskane z własnych prac PSH w latach 2006 i 2007. Obok samych wielkości poborów jest tam wiele innych informacji o charakterze administracyjnym i wybrane dane o ujęciach pochodzące z bazy CBDH. Serwis mapowy PSH umożliwia przeglądanie tych danych za pomocą przeglądarki internetowej. Jednak tą drogą można uzyskać jedynie informacje o lokalizacji ujęcia i jego nazwie bez informacji o wielkości poboru. Więcej informacji o tej bazie danych zawiera strona PSH – Pobory (PSH, 2014b) w oficjalnej witrynie PSH.

4.8. Obszary zagrożone podtopieniami

Mapa obszarów o wysokim ryzyku podtopień przedstawia maksymalne możliwe zasięgi występowania podtopień rozumianych jako miejsca, gdzie zwierciadło wody podziemnej może wystąpić tuż pod powierzchnią terenu, co przejawia się podmokłościami i występuje najczęściej w obrębie dolin rzecznych. Przestrzenna dokładność tych danych odpowiada skali 1:50 000, a układem odniesienia jest EPSG:2180 (PUGW 1992). Dane te są dostępne do przeglądania przez serwis WMS lub przez przeglądarkę internetową za pośrednictwem serwisu [geoportal.gov.pl](http://mapy.geoportal.gov.pl) (<http://mapy.geoportal.gov.pl/imap/?locale=pl&gui=new&sessionID=370445>) prowadzonego przez Główny Urząd Geodezji i Kartografii (GUGIK). Więcej informacji zawiera strona PSH – Obszary zagrożone podtopieniami (PSH, 2013f) w oficjalnej witrynie PSH.

4.9. Wody podziemne zaliczone do kopalin

Kolejną przedstawianą tu bazą danych prowadzoną przez PSH jest Bank Danych Wód Podziemnych Zaliczonych do Kopalin. Według informacji PSH baza ta zawiera dane geometryczne i opisowe dotyczące wód podziemnych zaliczonych do kopalin: wód leczniczych, termalnych lub solanek oraz dane o warunkach eksploatacji i ochrony ich złóż. Jak wynika z powyższego jest to baza o ujęciach tych wód, a dane o ujęciach znajdują się także w Banku HYDRO (w CBDH – Centralny Bank Danych Hydrogeologicznych HYDRO). Na podstawie

dostępnych publikacji trudno ustalić w jakim stopniu dane w obu bazach się dublują lub czy są jakoś wzajemnie powiązane. Z serwisu mapowego WebMap serwera SPD-PSH (rozdział 2.4.2.) można wyczytać, że identyfikatory obiektów (ujęć i studni) są zbudowane na zasadach przyjętych w Banku HYDRO, ale to nie wyjaśnia możliwych powiązań pomiędzy tymi dwoma bazami. Porównując ogólną strukturę danych Banku HYDRO (PSH, 2013a) z ogólną strukturą bazy o wodach podziemnych jako kopaliny (PSH, 2011) można stwierdzić, że są prawie identyczne i podstawowa różnica to dwa dodatkowe elementy dodane do klasy obiekt hydrogeologiczny: zwierciadła wód podziemnych i profil termiczny (rys. 4.3.).

4.10. Metadane

Metadane o zbiorach i seriach zbiorów danych z zakresu hydrogeologii, a także usług dotyczących tych metadanych są utrzymywane i udostępniane przez serwer usług wyszukiwania, bazujący na standardowych rekordach metadanych (Myciuk, 2012; PSH, 2013j). Serwer ten jest prowadzony przez Państwowy Instytut Geologiczny PIB i dostęp do niego przez interfejs dla przeglądarek znajduje się pod adresem <http://metadane.pgi.gov.pl/>.

Znajduje się tam 17 rekordów metadanych opisujących dane hydrogeologiczne należące do trzech kategorii:

- dane utrzymywane w CBDH (Centralny Bank Danych Hydrogeologicznych),
- dane pochodzące z monitoringu wód podziemnych,
- dane dotyczące wód podziemnych zaliczonych do kopalin.

CBDH (dawniej Bank HYDRO) jest bazą danych hydrogeologicznych, w której gromadzone są dane dokumentacyjne o odwiertach, ujęciach i źródłach wód podziemnych, mineralnych i termalnych z obszaru Polski. Dla opisu zawartości tej bazy za pomocą metadanych zostało w niej wyodrębnione osiem podzbiorów i dla nich opracowano oddzielne rekordy metadanych:

- podzbiór danych o sztolniach,
- podzbiór danych o szybach,
- podzbiór danych o otworach,
- podzbiór danych o otworach badawczych,
- podzbiór danych o źródłach,
- podzbiór danych o źródłach „podwierconych”,
- podzbiór danych o abisynkach,
- podzbiór danych o studniach kopanych.

Rekordy metadanych dotyczące Bazy Danych Monitoringu Wód Podziemnych:

- Monitoringu Granicznego (zawiera dane o punktach obserwujących wody podziemne w rejonach przygranicznych),
- Monitoringu Diagnostycznego (pomiarów ilościowej wody podziemnej w obrębach JCWPd wynikające z wymagań dyrektywy RDW (WFD)),
- Danych Archiwalnych Monitoringu Ilościowego (archiwalne wyniki pomiarów głębokości do zwierciadła wody lub wydajności źródeł i wyniki analiz chemicznych wody),
- Danych Archiwalnych Monitoringu Chemicznego (pomiarów jakości wody podziemnej w obrębach JCWPd wynikające z wymagań dyrektywy RDW (WFD)),
- Monitoringu Wód Podziemnych (wyniki aktualnych pomiarów głębokości do zwierciadła wody lub wydajności źródeł i wyniki analiz chemicznych wody),
- Danych o Wodach Podziemnych Pozyskiwanych Automatycznie (wyniki pomiarów głębokości do zwierciadła wody i wyniki pomiarów temperatury wody),

- Monitoringu Ilościowego (dane z punktów badawczych – stacji hydrogeologicznych pierwszego i drugiego rzędu),
- Lokalnego Monitoringu Badawczego (wody pod wpływem silnej presji górnictwa, dużych zakładów przemysłowych oraz aglomeracji miejskich),
- Bank Danych Wód Podziemnych Zaliczonych do Kopalni (dane dotyczące otworów i źródeł, w których ujęto wody zaliczone do leczniczych, termalnych lub solanek).

Brak jest tam rekordów opisujących inne zbiory danych hydrogeologicznych, a także serie zbiorów i usługi z tymi danymi związane (stan na początek roku 2015). Dotyczy to zbiorów opisanych tu wcześniej: Mapa hydrogeologiczna Polski – MHP, w tym: Użytkowy Poziom Wodonośny, Pierwszy Poziom Wodonośny – Występowanie i Hydrodynamika, Pierwszy Poziom Wodonośny – Wrażliwość na Zanieczyszczenie i Jakość Wód, a także Główne Zbiorniki Wód Podziemnych, Jednolite Części Wód Podziemnych, Zasoby dyspozycyjne, Pobory i Obszary zagrożone podtopieniami.

4.11. Inne dane wykorzystywane w hydrogeologii

Jak to jest przedstawione na początku rozdziału 2., środowisko hydrogeologów tworzy dane hydrogeologiczne na własne potrzeby i dla innych środowisk dziedzinowych. Problem wzajemnej wymiany danych ma jednak charakter symetryczny – hydrogeolodzy potrzebują do swoich analiz także danych z innych dyscyplin (innych tematów INSPIRE). Potrzeba ta jest szczególnie widoczna w przypadku zaawansowanych analiz z zastosowaniem metodyki modelowania matematycznego przepływu wód podziemnych i transportu zawartych w nich składników. Szczegółowe opisy metod zbierania danych geoprzestrzennych i ich opracowywania dla potrzeb budowy takich modeli są zawarte w innej monografii (Michalak, Nawalany, Sadurski i inni, 2011). Tu trzeba wymienić jedynie kategorie danych z innych dziedzin, które są często niezbędne dla weryfikacji pozyskanych danych hydrogeologicznych, zbudowania dostatecznie złożonego modelu numerycznego i przeprowadzenia analiz uzyskanych wyników, także w aspekcie różnych oddziaływań pomiędzy wodami podziemnymi i otaczającym je środowiskiem.

W pierwszej kolejności trzeba tu wymienić dane geologiczne i geofizyczne, ale w analizach i modelach hydrogeologicznych wykorzystuje się dane przestrzenne także z wielu innych dziedzin. Syntetyczne zestawienie najczęściej używanych danych zawiera tabela 4.1.

Tabela 4.1. Zestawienie danych pochodzących z innych dziedzin i najczęściej używanych w analizach i modelach hydrogeologicznych.

Kategoria danych	Element (klasa) modelu INSPIRE (nazwaPakietu::nazwaKlasy)	źródło krajowe
Dane geologiczne		
Profile otworów geologicznych	Geology::Borehole Geology::MappedInterval	PIG CBDG: – Otwory wiertnicze
Geologiczne opracowania kartograficzne	Geology::GeologicCollection Geology::MappedFeature Geology::GeologicUnit Geology::GeologicStructure Geology:: ShearDisplacementStructure Geology::Fold Geology::Anthropogenic- GeomorphologicFeature Geology::Natural- GeomorphologicFeature	PIG: – Szczegółowa mapa geologiczna Polski w skali 1:50 000 – Mapa geologiczno-gospodarcza Polski w skali 1:50 000 – Mapa geosrodowiskowa Polski w skali 1:50 000.
Dane geofizyczne Sejsmika 2D Sejsmika 3D Sondowania geoelektryczne	Geophysics:: GeophMeasurement Geophysics::GeophProfile Geophysics::GeophObject Geophysics::GeophObjectSet	PIG CBDG: – Geofizyka
Dane o ukształtowaniu terenu		
Warstwyce Numeryczny model terenu (DEM)	Elevation::ContourLine Elevation:: ElevationVectorObject Elevation:: ElevationGridCoverage	Centralny Ośrodek Dokumentacji Geodezyjnej i Kartograficznej (CODGIK): – Bezpośrednie dane pomiarowe – Warstwyce DGN / DXF – Dane NMT (Numeryczny Model Terenu) – Dane NMPT (Numeryczny Model Pokrycia Terenu)
Obszary chronione		
Parki narodowe Rezerваты przyrody Parki krajobrazowe Obszary chronionego krajobrazu Obszary Natura 2000 Pomniki przyrody Zespoły przyrodniczo-krajobrazowe	Protected Sites Full:: ProtectedSite	Generalna Dyrekcja Ochrony Środowiska (GDOŚ): – Baza Centralnego Rejestru Form Ochrony Przyrody PIG CBDG: – Środowisko – geostanowiska – Środowisko – jaskinie
Pokrycie terenu		
CORINE Land Cover 2000 CORINE Land Cover 2006	LandCoverRaster:: LandCoverGridCoverage LandCoverVector:: LandCoverDataset LandCoverVector:: LandCoverUnit	Główny Inspektorat Ochrony Środowiska (GIOŚ): – Krajowa baza danych pokrycia terenu CLC2006 – Krajowa baza danych zmian pokrycia terenu CLC 2000-2006 – Krajowa baza danych pokrycia terenu CLC2000

cd. tabeli 4.1

Dane hydrograficzne		
Wody powierzchniowe stojące i płynące	Hydro-PhysicalWaters:: ManMadeObject Hydro-PhysicalWaters:: DamOrWeir Hydro-PhysicalWaters:: RiverBasin Hydro-PhysicalWaters:: HydroObject Hydro-PhysicalWaters:: SurfaceWater Hydro-PhysicalWaters:: StandingWater Hydro-PhysicalWaters:: Watercourse Hydro-PhysicalWaters:: LandWaterBoundary	Krajowy Zarząd Gospodarki Wodnej (ZGW): Mapa Podziału Hydrograficznego Polski w skali 1 : 10 000 – Zlewnie elementarne – Zlewnie jednostek hierarchicznych – Działy wodne – Węzły na działach wodnych – Cieki wyróżnione (odcinki cieków) – Cieki wyróżnione (route) – Szerokie rzeki (szer. powyżej 30 m) – Węzły hydrograficzne cieków – Jeziora wyróżnione – Cieki niewyróżnione (odcinki cieków) – Jeziora niewyróżnione Instytut Meteorologii i Gospodarki Wodnej (IMGW)
Użytkowanie ziemi – obecne i planowane		
Dane topograficzne Planowanie przestrzenne: – w skali kraju, – województwa, – gminy	Existing Land Use:: ExistingLandUseDataSet Existing Land Use:: ExistingLandUseObject Planned Land Use:: SpatialPlan Planned Land Use:: ZoningElement	CODGIK: – Bazy danych obiektów topograficznych (BDOT10k) Ministerstwo Infrastruktury i Rozwoju (MIIR): – KPZK (koncepcja przestrzennego zagospodarowania kraju) – PZPK (plan zagospodarowania przestrzennego kraju) Jednostki samorządu terytorialnego (JST): – PZPW (plan zagospodarowania przestrzennego województwa) – SUiKZP (studium uwarunkowań i kierunków zagospodarowania przestrzennego gminy) – MPZP (miejscowy plan zagospodarowania przestrzennego)
Zasoby mineralne		
Złóża kopalin Obszary eksploatacji	MineralResources:: EarthResource MineralResources::Mine MineralResources:: MineralOccurrence	PIG CBDG: – Bloki koncesyjne dla węglowodorów – MIDAS – obszary górnicze – MIDAS – tereny górnicze – MIDAS – złoża kopalin
Gleby		
Struktura gleb Parametry hydrauliczne i chemiczne gleb Podatność na zanieczyszczenia	Soil::ObservedSoilProfile Soil::SoilBody Soil::SoilHorizon Soil::SoilLayer Soil::SoilSite Soil::SoilThemeCoverage	Instytutem Uprawy Nawożenia i Gleboznawstwa (IUNG): – Mapa glebowo-rolnicza w skali 1 : 25 000 dla obszaru całego kraju

cd. tabeli 4.1

Regiony morskie		
Linia brzegowa Bałtyku i jej warunki drenażu Granice morskich wód wewnętrznych Stan wód przejściowych i przybrzeżnych	Sea Regions::Coastline Sea Regions::Shoreline	Urzędy Morskie i Biuro Hydrograficzne Marynarki Wojennej (BHMW): – wewnętrzne bazy danych przestrzennych PIG CBDG: – Bałtyk – punkty badawcze
Ortoobrazy		
Ortofotomapy	Orthoimagery:: OrthoimageCoverag	Centralny Ośrodek Dokumentacji Geodezyjnej i Kartograficznej (CODGIK): – Ortofotomapy o rozdzielczości 1,00 lub 0,50 m w układzie 1992, w barwach RGB lub CIR
Gospodarowanie obszarem, strefy ograniczone i regulacyjne oraz jednostki sprawozdawcze		
Strefy ochronne różnego rodzaju Obszary koncesyjne	AreaManagementRestriction- AndRegulationZones:: ManagementRestriction- OrRegulationZone	PIG CBDG: – Bloki koncesyjne dla węglowodorów Regionalne Zarządy Gospodarki Wodnej (RZGW): – strefy ochronne ujęć wód
Urządzenia do monitorowania środowiska		
Dane z zakresu zjawisk atmosferycznych i hydrologicznych	EnvironmentalMonitoring- Facilities:: AbstractMonitoringFeature EnvironmentalMonitoring- Facilities:: EnvironmentalMonitoring- Facility EnvironmentalMonitoring- Facilities:: EnvironmentalMonitoring- Network	Instytut Meteorologii i Gospodarki Wodnej (IMGW): – Wewnętrzne bazy danych aktualnych i historycznych z dziedziny meteorologii i hydrologii
Zasoby energetyczne		
Surowce energetyczne (zasoby nieodnawialne – złoża) i ich wydobycie	Energy Resources Vector:: FossilFuelResource	PIG CBDG: – MIDAS – obszary górnicze – MIDAS – tereny górnicze – MIDAS – złoża kopalin

5. Przegląd modeli danych hydrogeologicznych

Początek prac nad modelami danych hydrogeologicznych sięga lat 70. XX wieku. Jedną z pierwszych opracowanych i uruchomionych baz danych hydrogeologicznych był Bank HYDRO Państwowego Instytutu Geologicznego o stosunkowo prostej strukturze rekordów przetwarzanych przez dedykowane oprogramowanie w języku Fortran dla komputera Odra. Głównym autorem koncepcji i oprogramowania był P. Stenzel (Stenzel, Berestka, 1979; Skrzypczyk, 1997), obecnie jednak ta koncepcja ma znaczenie tylko historyczne. Innym historycznym przykładem jest baza danych hydrogeologicznych opracowana w roku 1984 przez służbę hydrogeologiczną Republiki Południowej Afryki (Cogho, Kirchner, Morris, 1989; Howden, Mather, 2013). Późniejszym, ale znacznie bardziej dojrzałym projektem była relacyjna baza danych hydrogeologicznych opracowana przez służbę australijską (Rice, 1999; Brodie, 1999). Analiza modelu danych tej bazy jest przedstawiona w innej monografii (Michalak, 2003b).

Od roku 2000 obserwuje się lawinowy wzrost liczby różnych baz danych w większości krajów. Są to głównie rozwiązania oparte na technologii relacyjnych baz danych, jednak ich wspólną cechą jest hermetyczność polegająca na bardzo różnych modelach danych, braku możliwości działania interoperacyjnego z innymi systemami i w konsekwencji uzyskiwane dane mają formę wydruku na papierze.

Przegląd wielu rozwiązań dotyczących hydrogeologicznych baz danych zawiera rozprawa doktorska P. Wojdy *Hydrogeological data modelling in groundwater studies* (Wojda, 2008), a także jego publikacja we współautorstwie (Wojda, Brouyère, Dassargues, 2010). W obecnym czasie wymagania stawiane wobec takich systemów znacznie wzrosły i niezbędnym warunkiem jest interoperacyjna wymiana danych pomiędzy różnymi systemami. Osiągnięcie tego celu jest możliwe tylko pod warunkiem standaryzacji modeli danych, która powinna być jak najbardziej powszechna i o zasięgu międzynarodowym.

Obecnie istnieją dwa standardy o zasięgu ponad narodowym (ich porównanie przedstawione jest w następnym rozdziale):

- model danych języka *GroundWater Markup Language* (GWML) opracowany przez kanadyjską służbę hydrogeologiczną (wersja 1.0) (Boisvert, Brodaric, 2007) i nadal rozwijany jako program międzynarodowy (wersja 2.0) (Boisvert, Brodaric, 2012; Nathaniel, Booth, Brodaric, Lucido, Kuo, Boisvert, Cunningham, 2011; Brodaric, Booth, Boisvert, Lucido, 2014),
- model danych hydrogeologicznych opracowany dla potrzeb europejskiej infrastruktury INSPIRE (Michalak, 2009; Nałęcz, Michalak, 2011; INSPIRE TWG-GE-MR et al., 2013).

Opublikowany jest także stosunkowo prosty model pojęciowy, dotyczący danych hydrogeologicznych zawarty w ogólnym modelu pojęciowym, dotyczącym wyróżnień hydrogeologicznych przedstawiony w dokumencie OGC *HY_Features: a Common Hydrologic Feature Model Discussion Paper* (Dornblut, Atkinson, 2014). Znaczenie tego modelu dla hydrogeolo-

gii wynika z jego powiązania z językiem WaterML, a w szczególności jego pierwszą częścią *OGC WaterML 2.0: Part 1–Timeseries*, będącą standardem międzynarodowym OGC (OGC, 2014) i z pewnością wnoszącą wiele do uporządkowania zagadnień dotyczących obserwacji hydrogeologicznych. Z tego względu hydrologiczny model pojęciowy dla wód podziemnych jest w zarysie przedstawiony w rozdziale 5.1.3.

5.1. Porównanie modelu INSPIRE z modelami GWML i WaterML

W ramach wstępnych prac nad wersją 2.0 języka GWML przeprowadzono porównanie wersji 1.0 tego języka z modelem danych hydrogeologicznych dla INSPIRE i modelem hydrologicznym. Wyniki tego porównania są przedstawione w prezentacji umieszczonej w repozytorium projektu GWIE2 (Boisvert, 2012), a tu jest przedstawiona ogólna analiza podstawowych klas dotyczących jednostek hydrogeologicznych i znajdujących się w nich wód podziemnych.

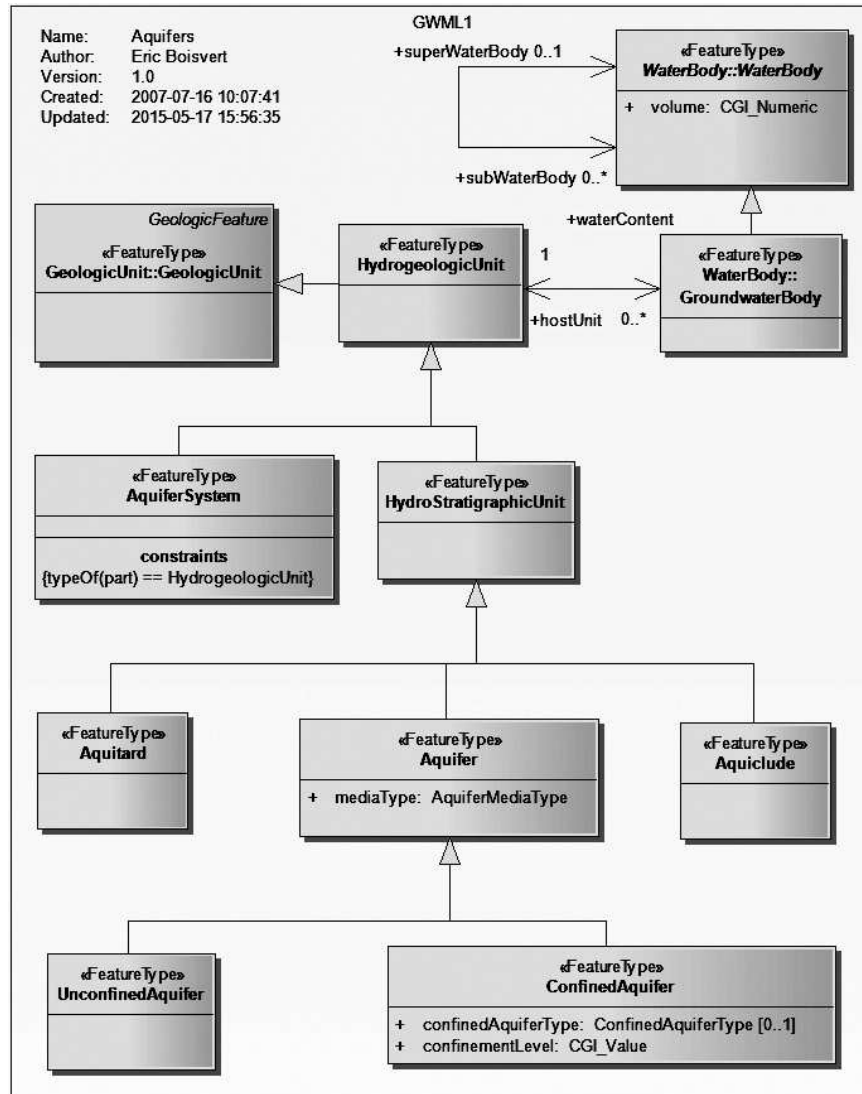
5.1.1. Język GWML 1.0

Język GWML 1.0, a także opracowywany jego następca GWML 2.0, jest rozszerzeniem języka GeoSciML w wersji 2.0 (IGW-CGI-IUGS, 2008) przeznaczonego dla zapisu informacji geologicznej. Konsekwencją tego jest podejście do zagadnień hydrogeologicznych z punktu widzenia geologów, którzy często zajmują się także hydrogeologią.

Na rysunku 5.1. Przedstawione są klasy języka GWML 1.0 dotyczące głównych elementów środowiska wód podziemnych. Klasa `HydrogeologicUnit` i klasy od niej pochodne (`HydroStratigraphicUnit`, `AquiferSystem`, `Aquitard`, `Aquifer`, `Aquiclude`, `UnconfinedAquifer` i `ConfinedAquifer`) są podtypami klasy `GeologicUnit` i w konsekwencji dziedziczą od niej wszystkie właściwości (atrybuty i powiązania typu asocjacja, agregacja kompozycja i inne). To sprawia, że wszystkie obiekty tych klas dotyczące podstawowych wyróżnień hydrogeologicznych muszą zawierać wiele składników istotnych dla geologii, ale nieistotnych dla hydrogeologii (przykład 5.1.) (Michalak, Boisvert, 2013).

Przykład 5.1. Fragment zapisu danych dotyczących wyróżnienia typu `HydrogeologicUnit` z wyekspozowaniem obowiązkowych elementów geologicznych nie istotnych dla hydrogeologii.

```
<gsm1gu:HydrogeologicUnit gml:id="xtest" (...) >
  (...)
  <gsm1gu:geologicUnitType xlink:href="http://. / HydrogeologicUnit "/>
  <gsm1gu:bodyMorphology xsi:nil="true" nilReason="missing"/>
  <gsm1gu:exposureColor xsi:nil="true" nilReason="missing"/>
  <gsm1gu:outcropCharacter xsi:nil="true" nilReason="missing"/>
  <gsm1gu:rank xsi:nil="true" nilReason="missing"/>
  <gsm1gu:metamorphicCharacter xsi:nil="true" nilReason="missing"/>
  <gsm1gu:part xsi:nil="true" nilReason="missing"/>
  <gsm1gu:alterationCharacter xsi:nil="true" nilReason="missing"/>
  <gsm1gu:bedding xsi:nil="true" nilReason="missing"/>
  <gsm1gu:geochemistry xsi:nil="true" nilReason="missing"/>
</gsm1gu:HydrogeologicUnit>
```



Rysunek 5.1. Diagram UML przedstawiający podstawowe klasy języka GWML 1.0 dotyczące głównych elementów środowiska wód podziemnych.

To tylko jeden z wielu przykładów niedostosowania tego modelu do potrzeb hydrogeologii. Innym przykładem może być określanie zasobów wody podziemnej w metrach sześciennych (m^3) jako całkowitej objętościowej zawartości wody w określonej jednostce hydrogeologicznej, co pewnie może mieć zastosowanie na przykład do złóż ropy naftowej, ale jest zupełnie nieracjonalne, a nawet bezsensowne, w przypadku wód podziemnych (przykład 5.2.).

Przykład 5.2. Atrybut `volume` (objętość w m³) w klasie `GroundwaterBody` odziedziczony od klasy `WaterBody`.

```

Class: WaterBody
  Stereotype: FeatureType
  SubClass: GroundwaterBody
  Stereotype: FeatureType
  Attribute: volume
  Definitions: Volume of water present in a water body
  AttributeType: CGI_Numeric

```

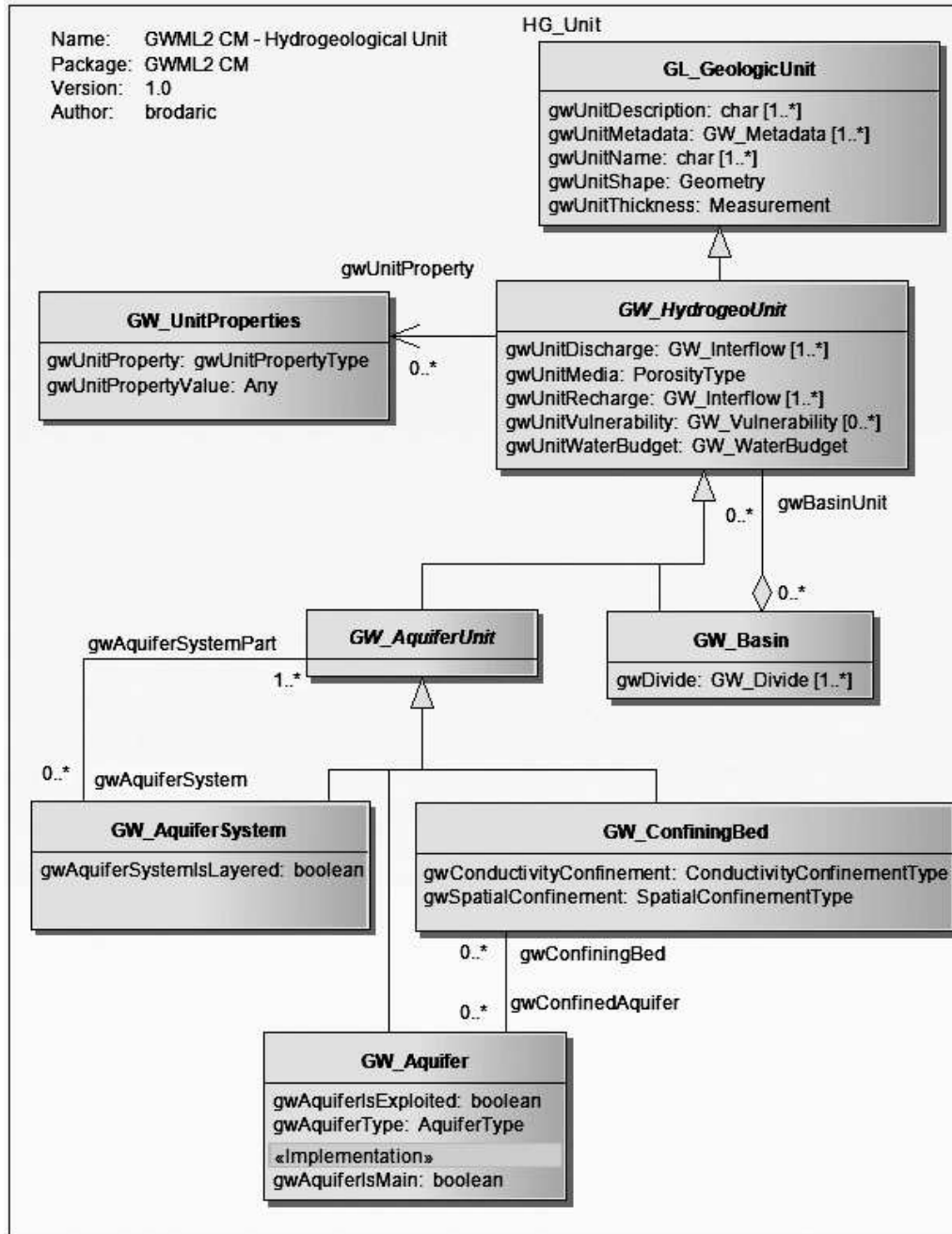
Obecnie (początek roku 2015) nadal trwają prace nad nową wersją języka GWML i z przeprowadzonych dyskusji wynika, że wiele wymienionych powyżej nieprawidłowości powinno zostać wyeliminowanych. Jest to jednak uzależnione od prac nad nową wersją języka GeoSciML 4.0. (Michalak, Boisvert, 2013), od którego GWML jest uzależniony. Wersja 1.0 języka GWML jest uzależniona od wersji 2.0 języka GeoSciML.

5.1.2. Prace nad językiem GWML 2.0

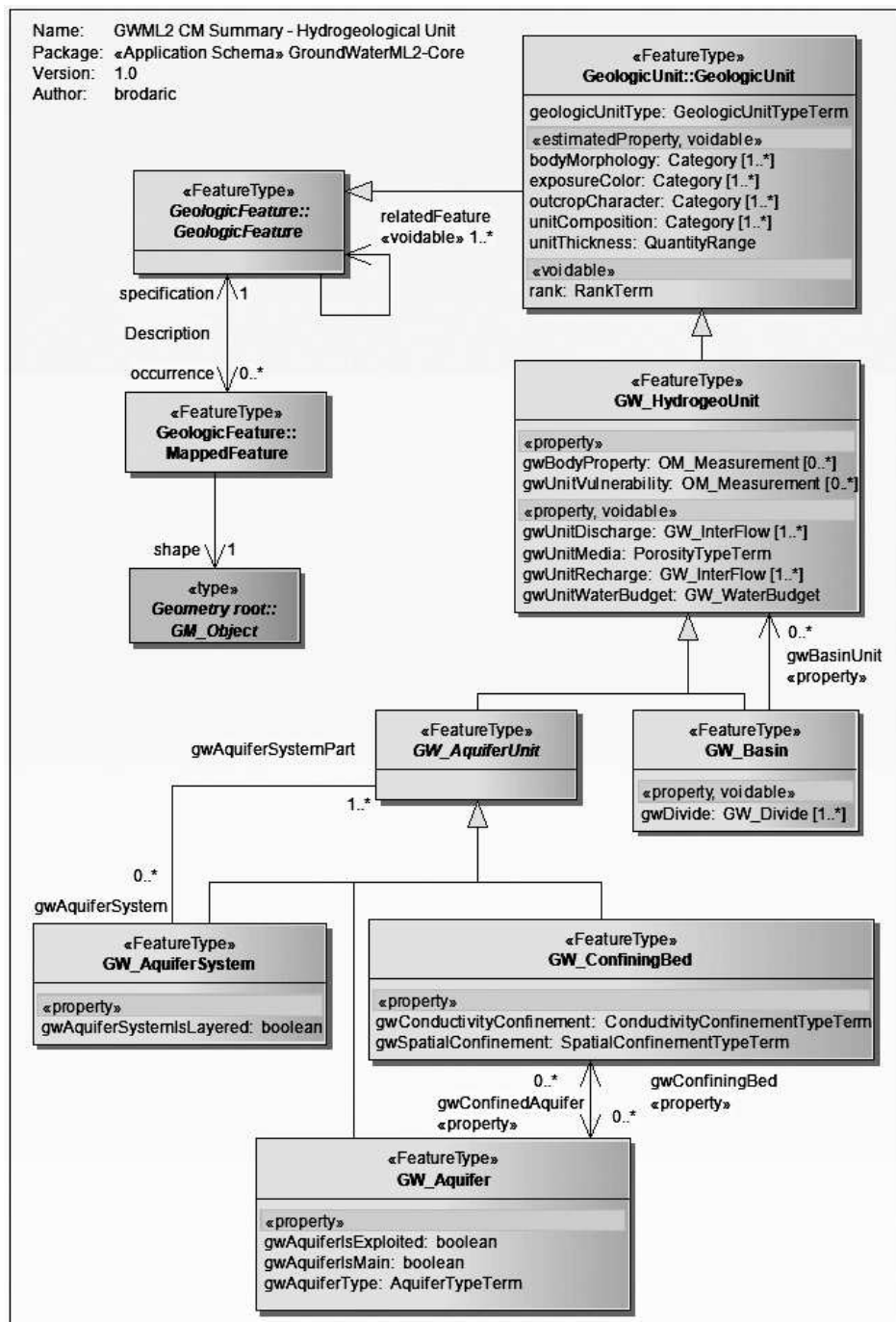
Obecnie prowadzone prace (początek roku 2015) nad rozwojem języka GWML w wersji 2.0 dotyczą logicznej wersji modelu UML (Simons, Brodaric, Grellet, 2015). Wersja ta zgodnie z przyjętą tam metodyką jest opracowywana na bazie wcześniejszej wersji koncepcyjnej opracowanej bez uwzględnienia środowiska implementacji, którym jest język GML. Rysunki 5.2. i 5.4. przedstawiają odpowiednio diagramy klas dla otoczenia klasy i otoczenia klasy `GW_Flow`. Analogiczne diagramy klas wersji logicznej tego modelu są przedstawione na rysunkach 5.3. i 5.5.

Porównanie obu wersji modelu wykazuje, że nie wszystkie założenia koncepcyjne mogą być zrealizowane w modelu logicznym. Na przykład założono, że klasa `GW_HydrogeoUnit` i wszystkie klasy od niej pochodne będą miały atrybut `gwUnitShape` typu `Geometry` odziedziczony od klasy `GeologicUnit` (rys. 5.2.), co pozwoli na określenie lokalizacji i zasięgu przestrzennego tych wyróżnień. Jednak klasa `GeologicUnit` w języku GeoSciML wersji 3.2. nie posiada geometrii (lokalizacji i kształtu), dziedziczy jedynie od klasy `GeologicFeature` dwukierunkowe powiązanie asocjacyjne z klasą `MappedFeature`, która z kolei przez asocjację z `GM_Object` pozwala na udostępnienie danych geometrycznych. Jako pozytywną zmianę trzeba uznać przyjęcie asocjacji dwukierunkowe pomiędzy `GeologicFeature` i `MappedFeature`, ponieważ we wcześniejszych wersjach języka GeoSciML nie było nawigacji z `GeologicFeature` do `MappedFeature`, co uniemożliwiało określenia geometrii wyróżnień należących do wszystkich pochodnych klas, w tym geometrii obiektów klasy `HydrogeoUnit` i jej podtypów.

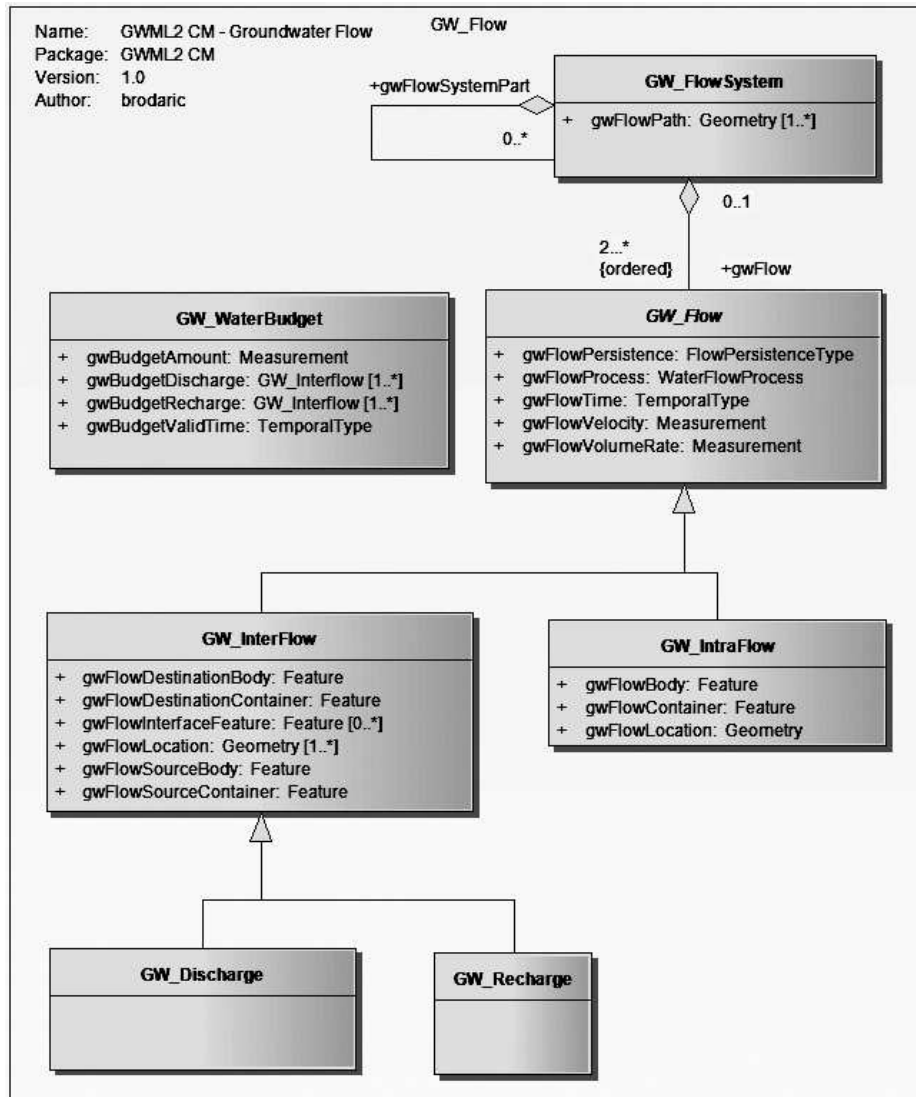
Konsekwencją przyjęcia założenia, że język GWML (w obu wersjach) jest rozszerzeniem aplikacyjnym języka GeoSciML (IGW-CGI-IUGS, 2008) dla wód podziemnych jest ograniczenie występujących tam odniesień przestrzennych tylko do formy wektorowej 1D, 2D lub 3D. Wszystkie dopuszczalne typy geometrii są podtypami abstrakcyjnej klasy `GM_Object`, co stwarza ryzyko możliwości użycia typu geometrii nieodpowiedniej do natury fizycznej opisywanego wyróżnienia przestrzennego. Niepokojącym faktem jest zastosowanie w części modelu dotyczącego przepływu wody abstrakcyjnej klasy `GFI_Feature` (rysunek 5.5.).



Rysunek 5.2. Diagram UML modelu pojęciowego przedstawiający podstawowe klasy opracowywanego języka GWML 2.0 dotyczące głównych elementów środowiska wód podziemnych.
 Na rysunku 5.3. przedstawiona jest wersja logiczna tego fragmentu modelu.

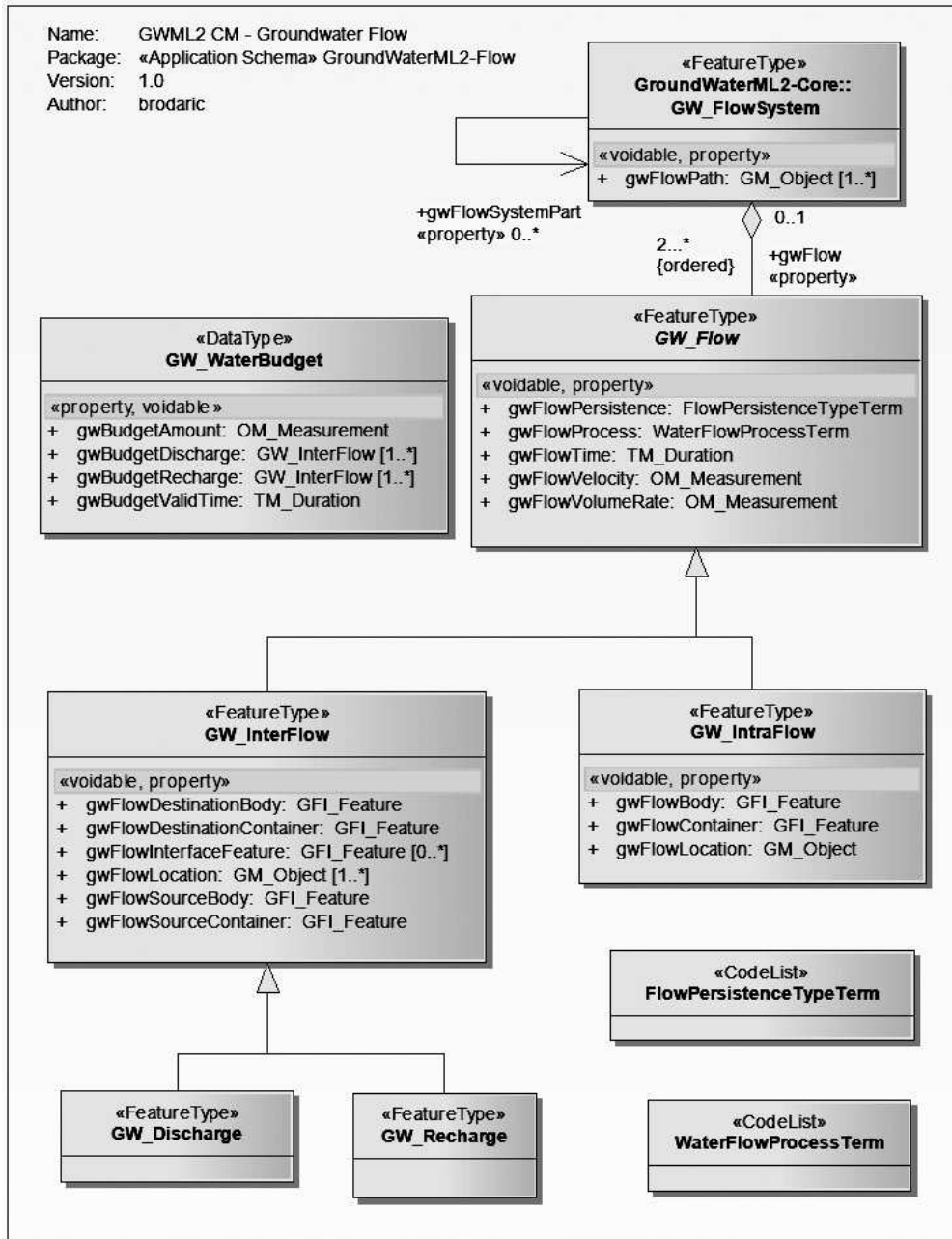


Rysunek 5.3. Diagram UML modelu logicznego przedstawiający podstawowe klasy opracowywanego języka GWML 2.0 dotyczące głównych elementów środowiska wód podziemnych. Na rysunku 5.2. przedstawiona jest wersja koncepcyjna tego fragmentu modelu.



Rysunek 5.4. Diagram UML modelu koncepcyjnego przedstawiający podstawowe klasy języka GWML 2.0 dotyczące głównych elementów związane z przepływem wód podziemnych.

Na rysunku 5.5. przedstawiona jest wersja logiczna tego fragmentu modelu.



Rysunek 5.5. Diagram UML modelu logicznego przedstawiający podstawowe klasy języka GWML 2.0 dotyczące głównych elementów związane z przepływem wód podziemnych. Na rysunku 5.4. przedstawiona jest wersja koncepcyjna tego fragmentu modelu.

Specyfikacja abstrakcyjna OGC (*topic 20*) dotycząca obserwacji i pomiarów definiuje abstrakcyjną klasę `GFI_Feature` (OGC, 2013). W myśl tej definicji klasa `GFI_Feature` jest instancją metaklasz `GF_FeatureType`. Stanowi ona zbiór wszystkich instancji typu wyróżnienie (*feature*) na poziomie klas (rys. 5.6.). Klasa `GFI_Feature` jest realizowana w GML przez element GML:AbstractFeature typu GML:AbstractFeatureType. W implementacji ta klasa abstrakcyjna zostaje zastąpiona przez rzeczywiste klasy reprezentujące typy obiektów w schemacie aplikacyjnym związanym z określoną dziedziną zastosowań. W tym przypadku akronim GFI oznacza *General Feature Instance* – Ogólna Instancja Wyróżnienia (na poziomie metaklasz).

Zastosowanie klas abstrakcyjnych w modelu logicznym przeznaczonym do utworzenia schematu aplikacyjnego pociąga za sobą wiele niejasności i pozwala na dowolne stosowanie wyróżnień całkowicie nie odpowiadających opisywanym zjawiskom.

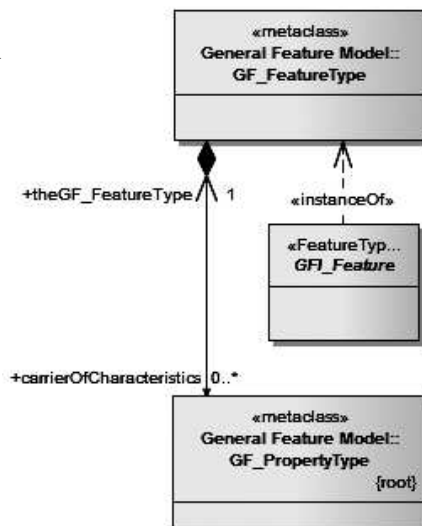
Szczególnym wymownym przykładem jest klasa `GW_InterFlow`, która zawiera wyłącznie 6 abstrakcyjnych atrybutów, z czego 5 są typu `GFI_Feature`: `gwFlowDestinationBody`, `gwFlowDestinationContainer`, `gwFlowInterfaceFeature`, `gwFlowSourceBody` i `gwFlowSourceContainer`. Nie ma tu możliwości sprecyzowania jakimi typami wyróżnień rzeczywistych można się w tych przypadkach posłużyć. Ponieważ prace nad tym modelem nie są jeszcze zakończone i prawdopodobnie jego praktyczne wykorzystanie będzie poddane testom, należy się spodziewać, że w przyszłości będą w tym przypadku zastosowane bardziej odpowiednie typy atrybutów.

5.1.3. Wody podziemne w języku WaterML – moduł *HY_Features*

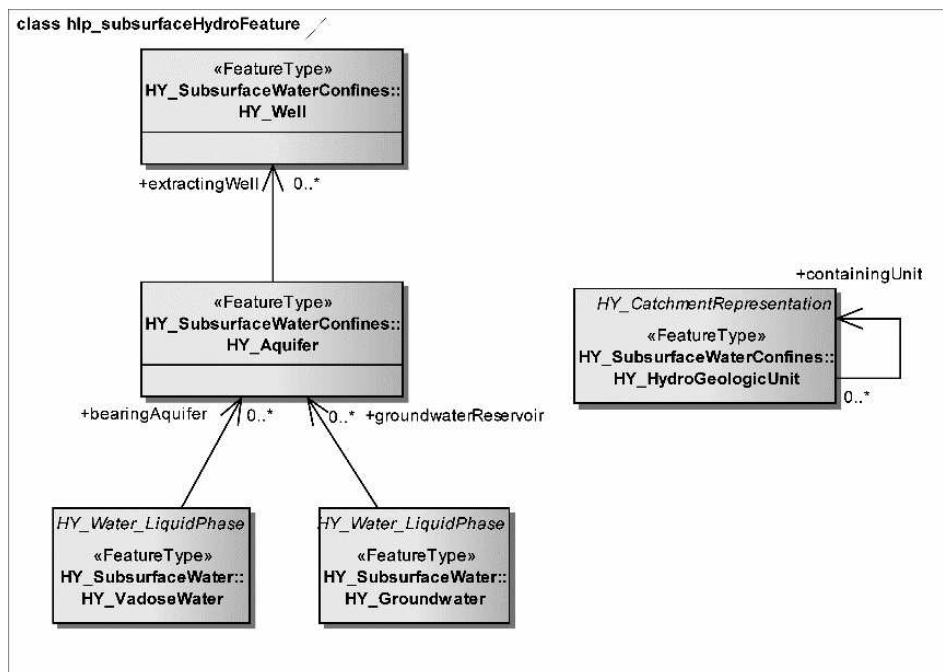
Trzecim z kolei modelem dotyczącym także wód podziemnych jest koncepcyjny model pod nazwą *HY_Features* (Dornblut, Atkinson, 2014), opisujący wyróżnienia hydrologiczne. Ma on z założenia stanowić podstawę nowego modułu w języku WaterML (OGC, 2014). Pakiet tego modelu dotyczący wód podziemnych (`HY_SubsurfaceWaterConfines`) zawiera tylko pięć klasy (rys. 5.7.):

- `HY_Well` z rolą asocjacyjną `extractingWell` dla klasy `HY_Aquifer`,
- `HY_Aquifer` z rolą asocjacyjną `bearingAquifer` dla klasy `HY_VadoseWater` z rolą asocjacyjną `groundwaterReservoir` dla klasy `HY_Groundwater`,
- `HY_VadoseWater` jako podtyp abstrakcyjnej klasy `HY_WaterLiquidPhase`,
- `HY_Groundwater` również jako podtyp abstrakcyjnej klasy `HY_WaterLiquidPhase`,
- `HY_HydroGeologicUnit` – podtyp abstrakcyjnej klasy `HY_CatchmentRepresentation`.

Klasy te nie posiadają atrybutów i z tego względu model ten należy traktować jako wstępny zarys koncepcyjny dla dalszych prac nad jego rozwijaniem.



Rysunek 5.6. Model instancji wyróżnienia (OGC, 2013)

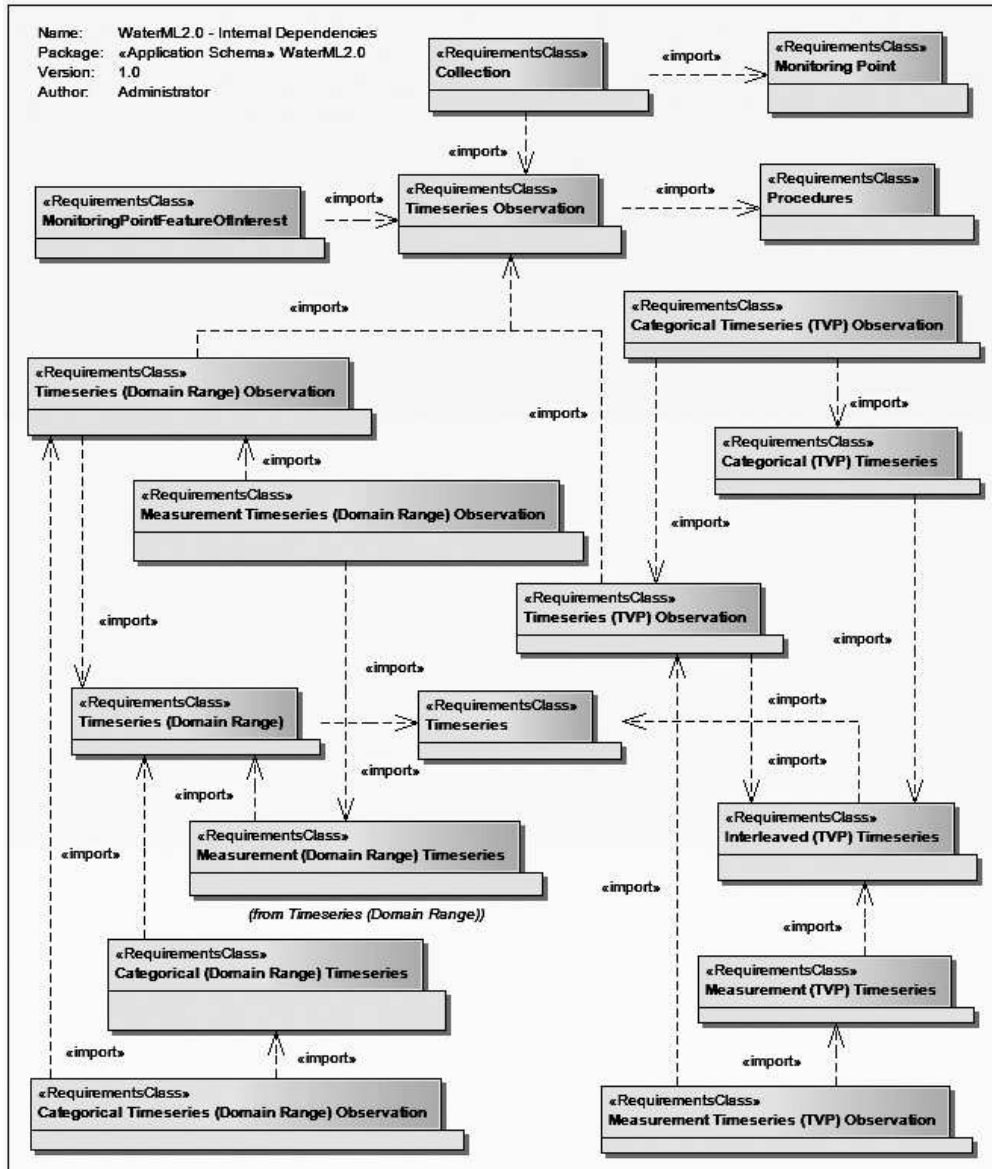


Rysunek 5.7. Diagram UML przedstawiający podstawowe klasy dotyczące wód podziemnych modelu *FY_Features* związanego z językiem WaterML 2.0 (Dornblut, Atkinson, 2014).

Jak to było przedstawione wcześniej, związek modelu *HY_Feature* z językiem WaterML ma istotne znaczenie dla hydrogeologii, ponieważ jeden z już opracowanych jego modułów o nazwie *Part 1 – Timeseries* dotyczy cyklicznych obserwacji i pomiarów wód powierzchniowych. Ogólną strukturę tego modułu podzielonego na pakiety przedstawiono na rysunku 5.8. Moduł ten jest jednak opracowany na tyle ogólnie, że bez trudu może być dostosowany do potrzeb hydrogeologicznych. Takie podejście pozwala na zapis wyników wielu różnych cyklicznych pomiarów i obserwacji w punktach monitoringu wód podziemnych, takich jak piezometry i bardziej rozbudowane stacje pomiarowe. Ponieważ WaterML jest uznanym standardem międzynarodowym dostępnych jest wiele systemów informatycznych przeznaczonych do przechowywania, przesyłania, przetwarzania i wizualizacji danych zapisanych w tym języku. Obecnie język ten jest standardem przyjętym przez OGC i WMO-CHy (*World Meteorological Organization – Commission for Hydrology*) i jest także zgłoszony do komitetu ISO TC 211 w celu przyjęcia go jako normy.

Przy opracowywaniu modułu *Timeseries* (ciągi czasowe) dołożono wiele starań aby mógł służyć bardzo różnorodnym pomiarom i metodom dalszego przetwarzania uzyskanych wyników. Zbiory pomiarów i obserwacji są tam traktowane jako dane typu pokrycie (*coverage*) analogiczne jak model powierzchni tereny, z tą tylko różnicą, że w przypadku modelu tereny występują dwie współrzędna przestrzenne X,Y, a w ciągach czasowych jest tylko jedna – czas T. Między innymi są tam dwie podstawowe kategorie:

- zbiory danych jako ciągi czasowe w postaci par liczb „czas i wartość” – TVP (*Time-Value Pair*),



Rysunek 5.8. Diagram pakietów UML przedstawiający podstawowe komponenty modelu *Part 1 – Timeseries* języka WaterML 2.0 (OGC, 2014).

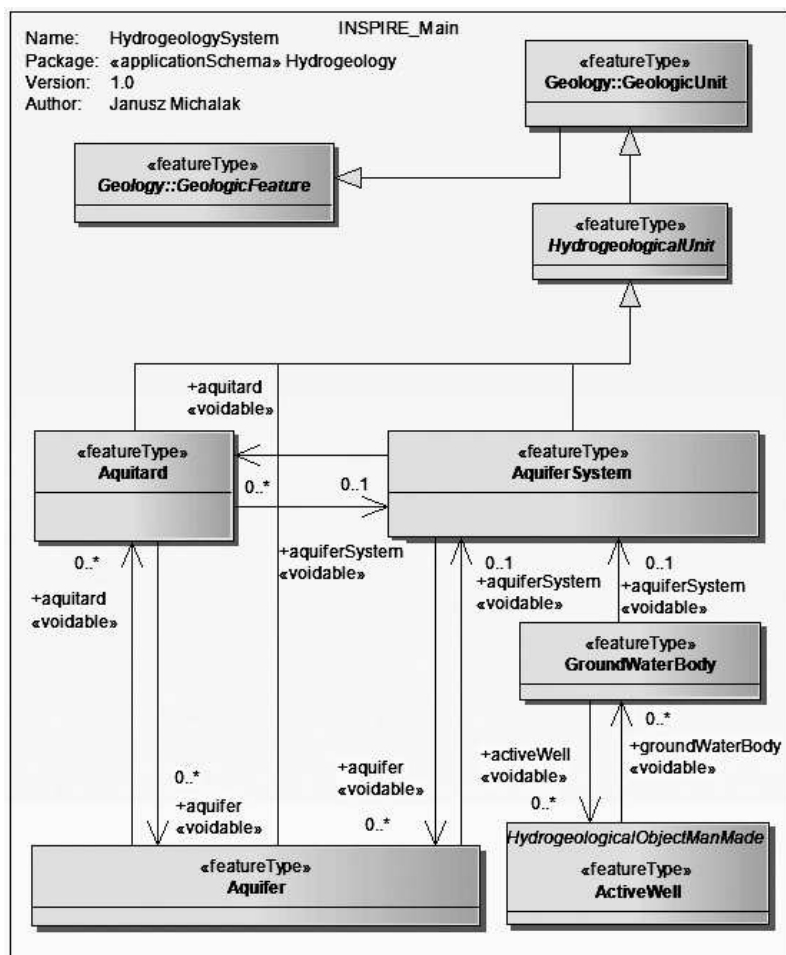
- zbiory danych jako ciągi czasowe, w których wartości czasu i wartości pomiaru są od siebie oddzielone (*domain – range*), na przykład lista terminów lub pierwszy termin i wielkość odstępu pomiędzy kolejnymi stanowią zbiór *domain*, a wyniki stanowią drugi zbiór *range*. Takie podejście wymaga określenia reguł, przypisanie wartości z jednego zbioru wartościom drugiego. Jest to podejście zbliżone do modelu terenu.

Drugi podział na dwie oddzielne kategorie obserwacji polega na wyróżnieniu pomiarów ilości i obserwacji jakości: *MeasurementTimeseries* i *CategoricalTimeseries*. Dzięki tym dwóm ortogonalnym podziałom otrzymujemy cztery podstawowe typy ciągów (serii) czasowych:

- *MeasurementDomainRangeTimeseries* – serie rozdzielonych danych pomiarowych,
- *CategoricalDomainRangeTimeseries* – serie rozdzielonych danych kategoryzacyjnych,
- *MeasurementTimeseriesTVP* – serie par danych pomiarowych,
- *CategoricalTimeseriesTVP* – serie par danych kategoryzacyjnych.

5.2. Projekt języka HydroGeoML

Na rysunku 5.9. diagram UML przedstawia podstawowe klasy modelu Hydrogeology INSPIRE dotyczące głównych elementów środowiska wód podziemnych (INSPIRE TWG-GE-MR et al., 2013).



Rysunek 5.9. Diagram UML modelu Hydrogeology INSPIRE przedstawiający podstawowe klasy dotyczące głównych elementów środowiska wód podziemnych (INSPIRE TWG-GE-MR et al., 2013).

Porównanie tego diagramu (rys. 5.9.) z diagramami modeli GWML1, GWML2 (rys. 5.3 i 5.5) i *HY_Feature* (rys. 5.7.), a także z modelami języka WaterML (rys. 5.8.), pełną wersją roboczą modelu INSPIRE i językiem GeoSciML przedstawione jest w tabeli 5.1. Porównanie pełnych specyfikacji tych modeli pozwala na ocenę ich poprawności i dojrzałości, a także ich braków z punktu widzenia hydrogeologii i na wyciągnięcie wniosków dotyczących dalszych prac w zakresie modelowania danych hydrogeologicznych. Pierwszy wniosek jaki się nasuwa z tej analizy porównawczej to to, że żaden z nich nie osiągnął wystarczającego poziomu dojrzałości (tab. 5.1.)

Tabela 5.1. Zbiorcze zestawienie cech modeli danych hydrogeologicznych i danych z nimi powiązanych, a także ocen przydatności tych modeli dla dalszych prac w tym zakresie

Model	Złożoność i zakres	Metodyka	Ocena i uwagi
GWML1	<ul style="list-style-type: none"> – Istotnie rozbudowany – 118 klas, w tym – 52 dotyczą wyróżnień 	<ul style="list-style-type: none"> – Wyłącznie geometria wektorowa – Listy kodowe lokalne – Brak wielojęzyczności 	<ul style="list-style-type: none"> – Uzależniony od GeoSciML (jest to jego aplikacja dla wód podziemnych) – Ukierunkowany na zastosowania praktyczne – Słabe podstawy hydrogeologiczne
GWML2 (wersja robocza)	<ul style="list-style-type: none"> – Istotnie rozbudowany – 88 klas, w tym – 25 dotyczy wyróżnień 	<ul style="list-style-type: none"> – Wyłącznie geometria wektorowa – Repozytorium list kodowych – Brak wielojęzyczności 	<ul style="list-style-type: none"> – Uzależniony od GeoSciML (jest to jego aplikacja dla wód podziemnych) – Szereg modułów zbyt teoretycznych, np. moduł <i>Flow</i> – Istotne problemy w zastosowaniach praktycznych – Słabe podstawy hydrogeologiczne
<i>HY_Features</i> (WaterML)	<ul style="list-style-type: none"> – Jedynie wstępny zarys – Tylko 5 podstawowych klas bez atrybutów 	<ul style="list-style-type: none"> – Dobre podstawy – Brak geometrii – Brak wielojęzyczności 	<ul style="list-style-type: none"> – Pretenduje do roli łącznika pomiędzy modelem hydrogeologicznym i WaterML
INSPIRE Hydro-geology (wersja końcowa)	<ul style="list-style-type: none"> – Zawężony do podstawowych klas – 22 klasy, w tym – 10 dotyczy wyróżnień 	<ul style="list-style-type: none"> – Geometria wektorowa i pokrycia siatkowe X,Y – Repozytorium list kodowych – Jest wielojęzyczność 	<ul style="list-style-type: none"> – Uzależniony od GeoSciML (jest to jego rozszerzenie dla wód podziemnych) – Dobre podstawy hydrogeologiczne
Dla porównania			
WaterML 2.0 (całość)	<ul style="list-style-type: none"> – 45 klas i w tym – 15 dotyczy wyróżnień 	<ul style="list-style-type: none"> – Geometria wektorowa i pokrycia w funkcji czasu T – Repozytorium list kodowych – Brak wielojęzyczności 	<ul style="list-style-type: none"> – Standard międzynarodowy – Zaawansowany metodycznie – Możliwość łatwej adaptacji do potrzeb obserwacji i pomiarów hydrogeologicznych
INSPIRE Hydro-geology (pełna wersja robocza)	<ul style="list-style-type: none"> – Istotnie rozbudowany – 170 klas, w tym – 75 dotyczy wyróżnień 	<ul style="list-style-type: none"> – Geometria wektorowa i pokrycia siatkowe X,Y – Brak repozytorium list kodowych – Jest wielojęzyczność 	<ul style="list-style-type: none"> – Niezależny od GeoSciML – Dobre podstawy hydrogeologiczne – Zawiera moduł dla dyrektywy WFD – Dobra podstawa dla dalszych prac
GeoSciML 3.2	<ul style="list-style-type: none"> – Istotnie rozbudowany – 131 klas, w tym – 36 dotyczy wyróżnień 	<ul style="list-style-type: none"> – Wyłącznie geometria wektorowa – Repozytorium list kodowych – Brak wielojęzyczności 	<ul style="list-style-type: none"> – Zaawansowany metodycznie – Ma być kandydatem do standardu międzynarodowego – Potrzebny w hydrogeologii dla opisu aspektu geologicznego

Przedstawione analizy i porównania istniejących modeli pojęciowych dotyczące danych hydrogeologicznych prowadzą do wniosku, że potrzebne są dalsze intensywne prace nad tym zagadnieniem. Powinny one doprowadzić do utworzenia modelu pełnego, spełniającego zadania praktyczne i badawcze, opartego na najnowszych osiągnięciach metodycznych i technologicznych – wykorzystującego wszelkie inne dojrzałe rozwiązania bazowe, a także z dziedzin powiązanych z hydrogeologią (Michalak, 2005).

Rysunek 5.10. przedstawia syntetyczne zestawienie procesów przetwarzania danych hydrogeologicznych, a także powiązanych z nimi danych geologicznych i hydrologicznych, od pomiarów i obserwacji terenowych i laboratoryjnych do złożonych analiz i symulacji komputerowych. Kolejno od lewej do prawej przedstawione są kolejne fazy przetwarzania danych od obserwacji, pomiarów, opróbowania i badań laboratoryjnych, poprzez wstępne przetwarzanie i kalibrację, następnie segregowanie danych i ich agregację, integrację danych z różnych obszarów dziedzinowych (na przykład dołączenie danych geologicznych i hydrologicznych), kończąc na wieloaspektowej analizie i zaawansowanych symulacjach komputerowych. Tworzy to logiczny łańcuch postępujących po sobie procesów, których końcowym wynikiem jest uzyskana tą drogą wiedza o badanych zjawiskach i procesach hydrogeologicznych określonej przestrzeni podziemnej. W dolnej części rysunku 5.10. przedstawiony jest przykład łańcucha operacji dotyczący potencjału hydraulicznego w punkcie o współrzędnych X,Y,Z pomierzonego sensorem umieszczonym w studni obserwacyjnej. Jednym z założeń języka HydroGeoML jest umożliwienie przeprowadzenia tych wszystkich operacji z jego zastosowaniem do zapisu danych w formie znacznikowej opartej na XML.

Hydrogeologia jako niezależna dyscyplina naukowa i praktyczna ma własny dobrze zdefiniowany i rozbudowany paradygmat. Z tego względu próby przeniesienia paradygmatów innych pokrewnych dziedzin na grunt modeli danych hydrogeologicznych nie mogą przynieść pozytywnych rezultatów. Dla potrzeb modelu pojęciowego niezbędne jest przedstawienie podstawowych elementów tego paradygmatu (pojęć i występujących między nimi powiązań) w formie ściśle sformalizowanej (Brodaric, Hahmann, 2014; Refsgaard, Henriksen, 2004). Niezbędnymi narzędziami do tego są:

- dla opisu ontologii: *Web Ontology Language* (OWL) lub *Ontology Inference Layer* (OIL),
- dla opisu taksonomii: *RDF Schema* (RDFS),
- dla zapisu i wymiany danych: *Resource Description Framework* (RDF).

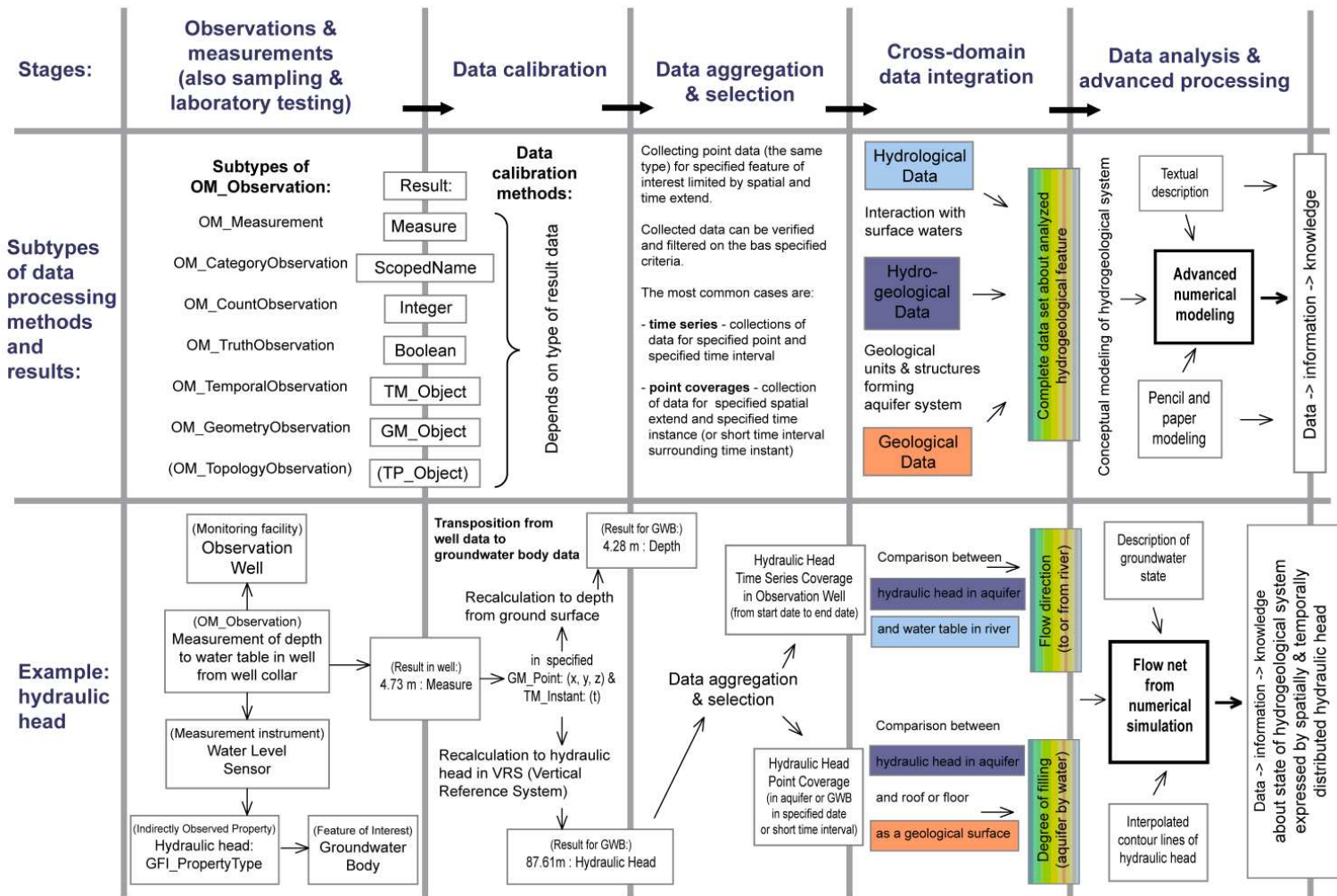
Szczegółowy i obszerny opis tych zagadnień w odniesieniu do danych hydrogeologicznych zawiera inna monografia (Michalak, 2003b).

Rysunek 5.11. przedstawia przyjęty w projekcie języka HydroGeoML fundament pojęciowy jako podstawę dla opracowania koncepcyjnej wersji modelu pojęciowego:

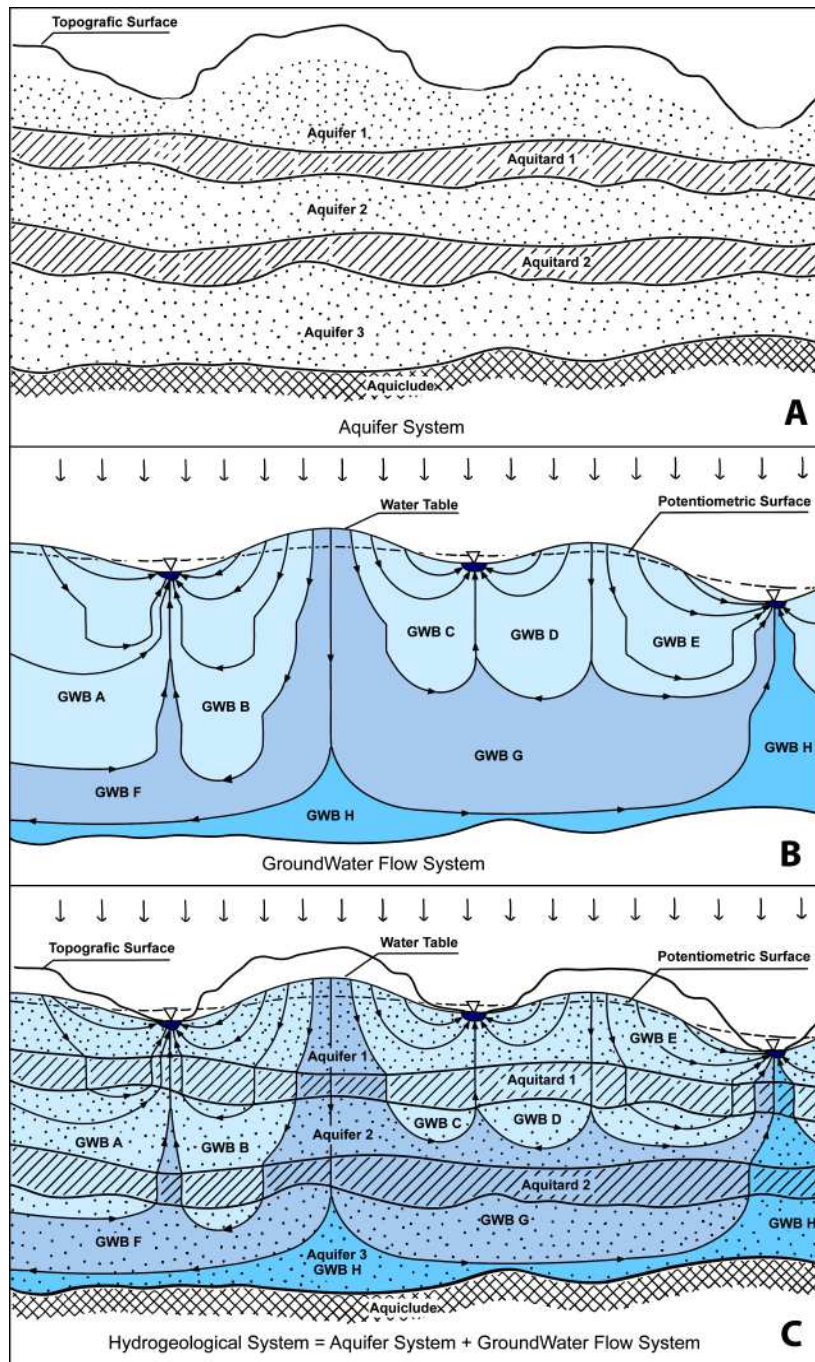
- system hydrogeologiczny (C) to wzajemnie powiązane dwa systemy,
- pierwszy z nich niezmienny w czasie system wodonośny (A),
- drugi to dynamiczny system przepływu wody podziemnej (B),
- pozostałe pojęcia modelu (na przykład studnia) są z tymi dwoma systemami powiązane.

Dotychczasowe studia i analizy pozwalają sformułować założenia i wymagania stawiane wobec nowej wersji języka dla danych hydrogeologicznych. Na obecnym etapie można w sposób ogólny podać je w punktach.

- Język dla danych hydrogeologicznych ma być aplikacją języka GML (*Geography Markup Language*) (Michalak, 2000; 2008) w wersji 3.3 lub 4.0, jeżeli taka będzie już opracowana.

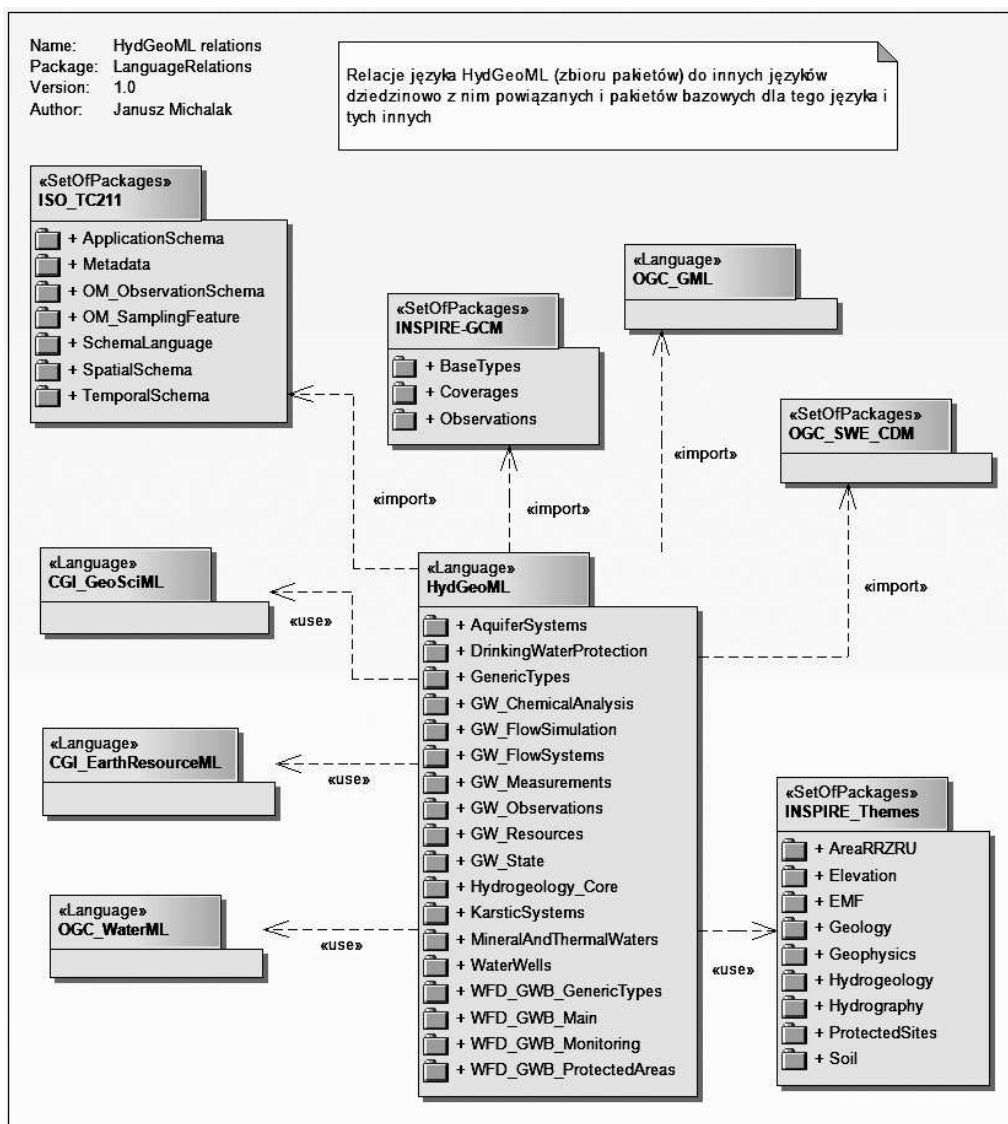


Rysunek 5.10. Syntetyczne przedstawienie procesów przetwarzania danych hydrogeologicznych, a także powiązanych z nimi danych geologicznych i hydrologicznych, od pomiarów i obserwacji terenowych i laboratoryjnych do złożonych analiz i symulacji komputerowych. Objaśnienia w tekście



Rysunek 5.11. Przyjęty w projekcie języka HydroGeoML fundament pojęciowy jako podstawa dla opracowania koncepcyjnej wersji modelu pojęciowego: system hydrogeologiczny (C) to wzajemnie powiązane dwa systemy – niezmienny w czasie system wodonośny (A) i dynamiczny system przepływu wody podziemnej (B). Pozostałe pojęcia modelu (na przykład studnia) są z tymi dwoma systemami powiązane (INSPIRE TWG-GE-MR et al., 2013).

- W miarę możliwości powinno to być rozszerzenie obecnego oficjalnego modelu INSPIRE dla podtematu Hydrogeologia.
- Wiele elementów pełnej wersji roboczej modelu INSPIRE dla podtematu Hydrogeologia powinno być wykorzystane na zasadzie ponownego użycia (*reuse*).
- Model danych powinien być oparty na dostatecznie dojrzałych modułach uznanych za standardy międzynarodowe, a w szczególności na zbiorze pakietów ISO/TC211, INSPIRE-GCM, języku GML i standardzie OGC SWE/CDM (rys. 5.12). Powinien wykorzystywać ich elementy zapisane w schematach XSD na zasadzie importu.



Rysunek 5.12. Diagram pakietów UML przedstawiający wstępny zarys struktury języka HydroGeoML i jego powiązań z innymi językami i schematami bazowymi.

- W stosunku do innych modeli z dziedzin pokrewnych, takich jak języki GeoSciML, EarthResourceML, WaterML, a także modele innych tematów INSPIRE, model HydroGeoML powinien być niezależny – powinien móc funkcjonować bez odwoływania się do ich własnych schematów XSD.
- Współdziałanie z tymi pokrewnymi modelami jest jednak jak najbardziej wskazane, ale głównie przez nieobowiązkowe powiązania asocjacyjne. Na przykład informacje geologiczne o formacji wodonosnej, jeżeli są potrzebne, powinny być zapisane w języku GeoSciML i powinien być odpowiedni odsyłacz z zapisu danych w języku HydroGeoML.
- Najnowsze prace z zakresu budowania modeli danych opartych na języku GML wykazują, że obecne wersje tego języka są zbyt monolityczne i przez to zbyt „ciężkie”. Prowadzi to do wniosku, że niezbędne jest jak najszersze stosowanie struktury modułowej, co znajduje swój wyraz w opracowywanej 4. wersji języka GML (Burggraf, 2011), ale także już częściowo w opublikowanej wersji 3.3.
- Model danych powinien być użyteczny nie tylko do utworzenia schematów XSD dla przesyłania danych w zapisach znacznikowych XML, ale także dla utworzenia struktur relacyjnych baz danych za pomocą zapisów w języku DDL (*Data Definition Language*). Zagadnienie to jest opisane we wcześniejszej monografii (Michalak, Chojka, Zwirowicz-Rutkowska, Parzyński, 2012).
- Język HydroGeoML powinien używać typów geometrii dostosowanych do przyrodniczej natury opisywanych zjawisk, jednak z zachowaniem przyjętych rygorów ograniczających pełną dowolność wyboru geometrii, jak to ma miejsce w przypadku geometrii abstrakcyjnej klasy GM_Object. W miarę możliwości powinny być stosowane tylko typy określone w profilu Simple Feature języka GML (Brink, Portele, Vretanos, 2011).
- Dla zapisu danych hydrogeologicznych dotyczących parametrów przestrzennie rozłożonych w układzie X,Y, takich jak zwierciadło wody lub strop i spąg warstwy, potrzebne jest stosowanie struktur danych typu pokrycie siatkowe *RectifiedGridCoverage* lub *ReferenceableGridCoverage* zdefiniowane w INSPIRE *Generic Conceptual Model* (GCM) i skutecznie zastosowane w modelu INSPIRE Hydrogeology. Analiza zastosowania takich form zapisu danych hydrogeologicznych jest przedstawiona w innej monografii (Michalak, Nawalany, Sadurski i inni, 2011).

6. Specyfikacja i model danych INSPIRE dla podtematu Hydrogeologia

Proces opracowania modelu INSPIRE podtematu Hydrogeologia został opisany w rozdziale 3.2. Teraz trzeba przedstawić jego zakres tematyczny i elementy, takie jak: klasy, ich atrybuty, powiązania pomiędzy klasami, a także enumeratory i listy kodowe. Ponieważ wszystkie dokumenty definiujące ten model – specyfikacja tekstowa, rozporządzenie Komisji Europejskiej i pliki elektroniczne – są dostępne w internetowych repozytoriach INSPIRE, tu przedstawione będą jedynie ogólne uwagi i główne problemy technologiczne dotyczące modelu. Lokalizację źródeł dotyczących modelu, schematów XSD i list kodowych podano w punktach.

- Obowiązujący w krajach członkowskich akt prawny: Rozporządzenie Komisji (UE) Nr 1089/2010 z dnia 23 listopada 2010 r. w sprawie wykonania dyrektywy 2007/2/WE Parlamentu Europejskiego i Rady w zakresie interoperacyjności zbiorów i usług danych przestrzennych (Dz.U. L 323 z 8.12.2010, s. 11), zmienione przez: Rozporządzenie Komisji (UE) nr 102/2011 z dnia 4 lutego 2011 r. i Rozporządzenie Komisji (UE) nr 1253/2013 z dnia 21 października 2013 r. – str. 201-213
URL: <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:02010R1089-20131230&from=PL> (EC, 2013b).
- Szczegółowy opis modelu jest zawarty w specyfikacji danych tematu Geologia (INSPIRE TWG-GE-MR et al., 2013), str. 42-63.
URL: http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_GE_v3.0.pdf
- Pełen model UML w wersji r4618 dla wszystkich tematów wraz z innymi modelami bazowymi:
 - INSPIRE Consolidated UML Model – ANNEX I-II-III – complete.xml – w formacie XML,
 - INSPIRE Consolidated UML Model – ANNEX I-II-III.EAP – jako plik binarny projektu programu Enterprise Architect.URL: <http://inspire.ec.europa.eu/data-model/approved/r4618/ea+xmi/EAXMI.zip>
Dokumentacja tego modelu w formacie HTML jest dostępna pod adresem
URL: <http://inspire.ec.europa.eu/data-model/approved/r4618/html/>
- Katalog wyróżnień wyszczególnionych w modelu hydrogeologicznym jest dostępny pod adresem:
URL: http://inspire.ec.europa.eu/data-model/approved/r4618/fc/#_P3519
- Robocza wersja schematu w postaci zapisu XSD (wersja 3.0rc3) dla modelu hydrogeologicznego jest dostępna pod adresem:
URL: http://inspire.ec.europa.eu/draft-schemas/ge_hg/3.0rc3/HydrogeologyCore.xsd
- Repozytorium list kodowych i enumeratorów – URL: <http://inspire.ec.europa.eu/codelist/>.
Dla podtematu hydrogeologia jest tam 9 list kodowych:
 - Hydrogeochemiczny rodzaj skały – Wartości opisujące hydrogeochemiczny stan środowiska wód podziemnych (Lista kodowa rozszerzalna).
URL: <http://inspire.ec.europa.eu/codelist/HydroGeochemicalRockTypeValue/>

- Okresowość wód – Rodzaje hydrologicznej okresowości wód (Lista nierozszerzalna – enumerator). URL: <http://inspire.ec.europa.eu/codelist/WaterPersistenceValue/>
- Rodzaj obiektu naturalnego – Rodzaje naturalnych obiektów hydrogeologicznych (Lista kodowa rozszerzalna). URL: <http://inspire.ec.europa.eu/codelist/NaturalObjectTypeValue/>
- Rodzaj ośrodków warstwy wodonośnej – Wartości opisujące właściwości ośrodka warstwy wodonośnej (Lista kodowa rozszerzalna). URL: <http://inspire.ec.europa.eu/codelist/AquiferMediaTypeValue/>
- Rodzaj studni – Typy studni aktywnych (Lista kodowa rozszerzalna). URL: <http://inspire.ec.europa.eu/codelist/ActiveWellTypeValue/>
- Rodzaj warstwy wodonośnej – Typy warstw wodonośnych (Lista nierozszerzalna – enumerator). URL: <http://inspire.ec.europa.eu/codelist/AquiferTypeValue/>
- Stan wód podziemnych – Wartości wskazujące przybliżony stopień zmian, które zaszły w naturalnym stanie wód podziemnych (Lista nierozszerzalna – enumerator). URL: <http://inspire.ec.europa.eu/codelist/ConditionOfGroundwaterValue/>
- Typ kodu statusu – Wartości opisujące status sztucznych obiektów hydrogeologicznych (Lista kodowa rozszerzalna). URL: <http://inspire.ec.europa.eu/codelist/StatusCodeTypeValue/>
- Zasolenie wody – Lista kodowa wskazująca klasy zasolenia wody (Lista kodowa rozszerzalna). URL: <http://inspire.ec.europa.eu/codelist/WaterSalinityValue/>

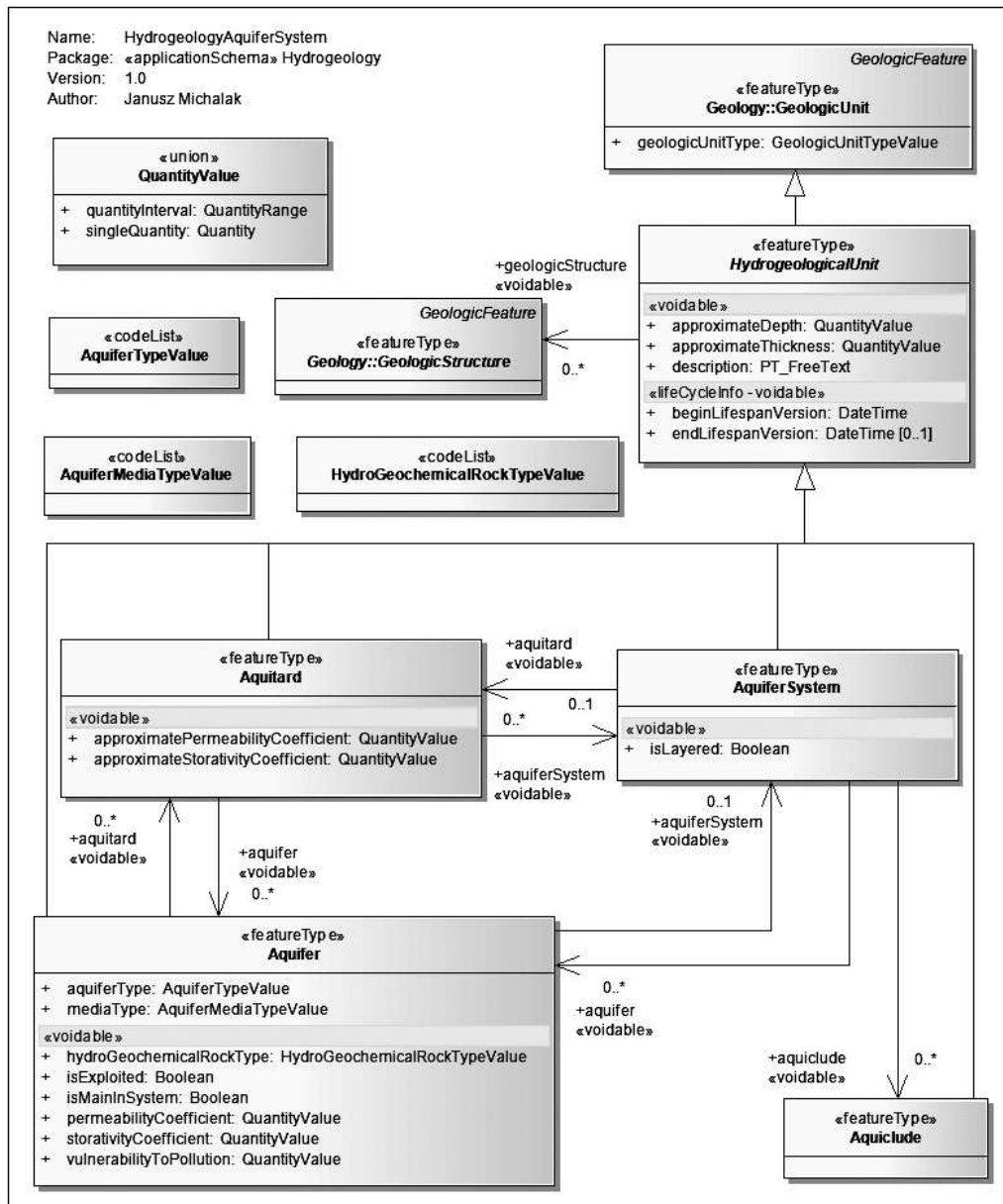
6.1. Elementy modelu INSPIRE

Przebieg prac nad utworzeniem modelu danych INSPIRE dla podtematu Hydrogeologia opisany jest w rozdziale 3.2. Końcowy rezultat tych prac zawarty jest w wyliczonych powyżej dokumentach i repozytoriach INSPIRE. Tu jest przedstawiony ogólny jego zarys i problemy technologiczne z nim związane. Rysunek 5.9. przedstawia podstawową część modelu hydrogeologicznego, a rysunki 2.1., 2.2. i 2.3. przedstawiają powiązania tego modelu z innymi tematami opisanymi w rozdziale 2. Dla zilustrowania pełnego obrazu modelu na kolejnych rysunkach (rys. 6.1., 6.2., 6.3.) pokazane są pozostałe elementy modelu:

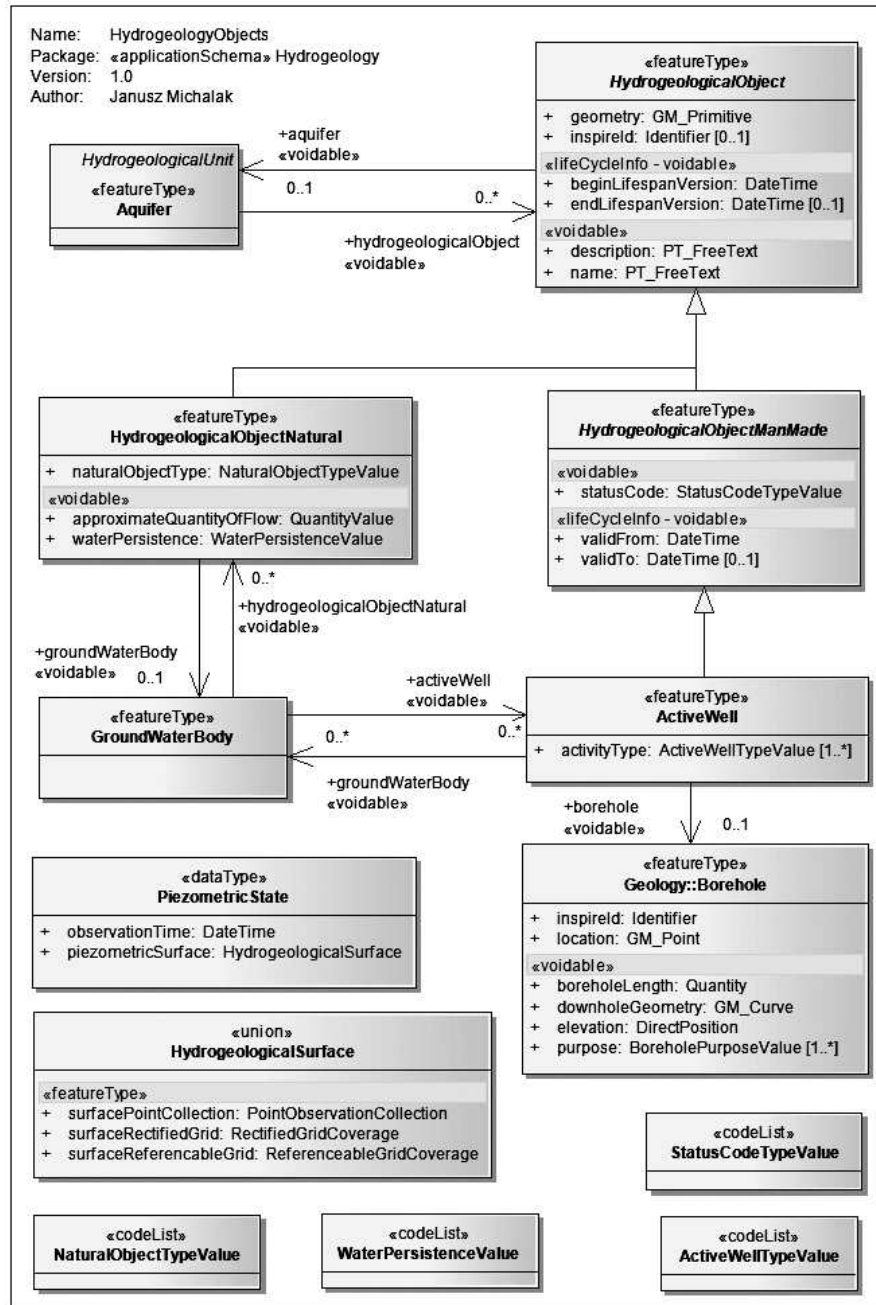
- system wodonośny (*AquiferSystem*), jego listy kodowa i powiązania z jednostką geologiczną (*GeologicUnit*) i strukturą tektoniczną (*GeologicStructure*),
- obiekty hydrogeologiczne (*HydrogeologicalObject*), a w szczególności studnia aktywna (*ActiveWell*) i związany z nią otwór wiertniczy (*Borehole*). Diagram ten zawiera także typy danych opisujących powierzchnie hydrogeologiczne (*HydrogeologicalSurface*) i stan piezometryczny (*PiezometricState*),
- powiązania pomiędzy systemem wodonośnym (*AquiferSystem*), częściami wód podziemnych (*GroundWaterBody*) a obiektami hydrogeologicznymi naturalnymi (*HydrogeologicalObjectNaturam*) i sztucznymi (*HydrogeologicalObjectManMade*).

Model ten nie zawiera szeregu ważnych w hydrogeologii modułów, które były w jego wersji roboczej, ale niestety ze względu na konieczność jego ograniczenia zostały usunięte. Do modułów tych należą schematy aplikacyjne (ze stereotypem «*applicationSchema*»): *WFD_GWB*, *MineralAndThermalWaters*, *DrinkingGroundWaterProtection* i *KarsticSystem*. Ponadto, model nie zawiera wielu istotnych dla zagadnień praktycznych klas, jak na przykład: *ObservationWell*, *Well* (jako typ ogólny), *CompoudSystem*,

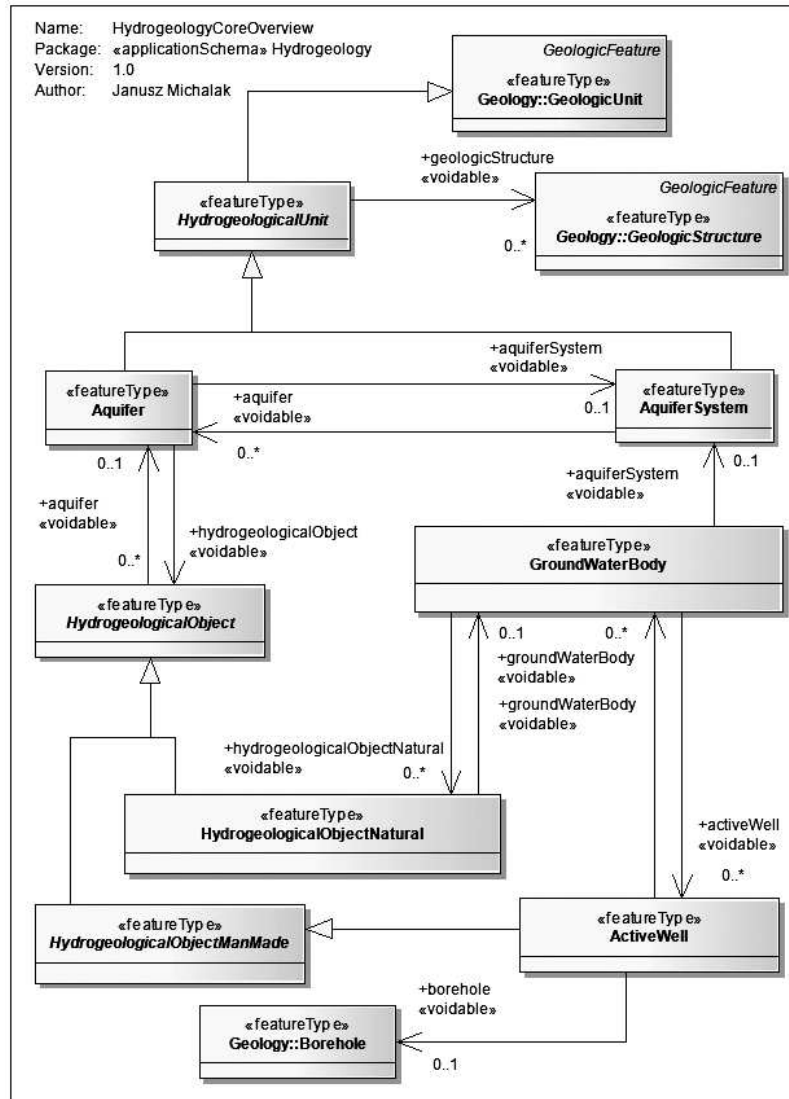
FracturedSystem, GroundWaterFlowSystem, HydrogeologicalProfile, GroundWaterIntake, GroundWaterMonitoringStation, KarsticSystem, PorousSystem, Seep, Spring i SwallowHole.



Rysunek 6.1. Diagram UML modelu INSPIRE przedstawiający system wodonośny (AquiferSystem), jego listy kodowa (stereotyp «CodeList») i powiązania z jednostką geologiczną (GeologicUnit) i strukturą tektoniczną (GeologicStructure).



Rysunek 6.2. Diagram UML przedstawiający obiekty hydrogeologiczne (HydrogeologicalObject), a w szczególności studnia aktywna (ActiveWell) i związany z nią otwór wiertniczy (Borehole). Diagram ten zawiera także typy danych opisujących powierzchnie hydrogeologiczne (HydrogeologicalSurface) i stan piezometryczny (PiezometricState).



Rysunek 6.3. Diagram UML przedstawiający powiązania pomiędzy systemem wodonośnym (AquiferSystem), częściami wód podziemnych (GroundWaterBody), a obiektami hydrogeologicznymi naturalnymi (HydrogeologicalObjectNatural) i sztucznymi (HydrogeologicalObjectManMade).

Konsekwencją zawężenia modelu INSPIRE jest jego ograniczona przydatność dla zadań, jakie spoczywają na środowiskach hydrogeologicznych krajów członkowskich, w tym Polski, gdzie funkcjonuje Państwowa Służba Hydrogeologiczna. Z tego względu dla wielu zastosowań praktycznych model ten wymaga rozszerzenia a metodyka opracowania modelu rozszerzonego przedstawiona jest w rozdziale 6.3.

6.2. Główne problemy technologiczne z zakresu danych hydrogeologicznych

W większości przypadków modele i schematy INSPIRE nie uwzględniają aspektu wielojęzyczności tekstów (nazw, opisów) zawartych w danych przez powszechne użycie najprostszego typu `CharacterString`, który nie daje możliwości określenia języka. Problem wielojęzyczności jest w Europie zagadnieniem bardzo ważnym i uzasadnienie tego znajduje się w publikacji *Multilingual European Subset of Unicode in geospatial data encoding* (Michalak, 2004). W modelu UML i schemacie XSD dla danych hydrogeologicznych zostało to uwzględnione przez zastosowanie typu `PF_FreeText` zdefiniowanego w normie ISO 1939 dotyczącej metadanych. Typ ten pozwala na użycie kilku elementów `LocalizedCharacterString`, w których umieszcza się teksty w wybranych językach jako `CharacterString` i w atrybucie określa się w jakim języku jest ten tekst napisany. Poniżej w przykładzie 6.1. zacytowany jest fragment schematu XSD dla danych hydrogeologicznych (obiektu hydrogeologicznego – `HydrogeologicalObject`) z elementami `nane` i `description` jako typu `PF_FreeText`. Sposób zapisu tekstów w różnych językach przy użyciu tego typu elementu przedstawiony jest w rozdziale 12.5. (przykład 12.5.).

Inny problem dotyczy określania typu geometrii dla wyróżnień geoprzestrzennych. Język GML, w którym są zapisywane dane tematów INSPIRE określa bardzo wiele różnych typów geometrii i o różnym stopniu ogólności (INSPIRE DT-DS, 2014a). Który z tych typów i na jakim poziomie ogólności zostanie przyjęty dla określonego wyróżnienia zależy od autora modelu. W różnych tematach jest to przyjmowane różnie, jednak najczęściej bez zastanowienia przyjmowany jest typ `GM_Object` jako najogólniejszy abstrakcyjny typ, co pozwala na użycie każdej dowolnej geometrii.

W modelu danych hydrogeologicznych najogólniejszy abstrakcyjny typ geometrii (`GM_Object`) jest przypisany do przestrzennej reprezentacji (`MappedFeature`) jednostki hydrogeologicznej (`HydrogeologicalUnit`) jako jej geologiczna interpretacja, na przykład w formie poszczególnych wychodni lub innych śladów powierzchniowych lub podziemnych. W rezultacie tego, jednostka hydrogeologiczna jako taka nie posiada geometrii, czyli nie posiada przypisania przestrzennego. Brak jest jakiegokolwiek określonej lokalizacji, nawet w sposób przybliżony. Jest to konsekwencją niefortunnej zmiany koncepcji modelu hydrogeologicznego, w rezultacie której jednostka hydrogeologiczna (`HydrogeologicalUnit`) stała się podtypem jednostki geologicznej (`GeologicUnit`) z całymi konsekwencjami przeniesienia w dół wszystkich specyficznych właściwości i relacji geologicznych. W wyniku tego inne wyróżnienia hydrogeologiczne: `Aquifer`, `Aquitard`, `AquiferSystem` i `Aquiclude`, będące podtypami jednostki hydrogeologicznej, nie mają przypisania przestrzennego, czyli nie mają lokalizacji i nie mogą być na przykład wybrane ze zbioru takich jednostek poprzez określenie zasięgu w zapytaniu geoprzestrzennym.

Przykład 6.1. Specyfikacja elementów `nane` i `description` jako typu `PF_FreeText` w schemacie XSD dotyczącym danych hydrogeologicznych,
źródło: URL: http://inspire.ec.europa.eu/schemas/ge_hg/3.0/HydrogeologyCore.xsd/

Na szczęście wyróżnienie `GroundWaterBody` w wyniku tych zmian modelu nie zostało zaliczone do podtypów jednostki geologicznej (`GeologicUnit`) i zachowało swój atrybut przestrzenny `approximateHorizontalExtend` typu `GM_Surface`, co odpowiada powierzchni 2D (płaskiej i zamkniętej konturem granicznym). Również wyróżnienia pochodne

```

<element name="name" type="gmd:PT_FreeText_PropertyType"
nillable="true">
  <annotation>
    <documentation>
      Name: Name
      Definition: The name or code of the HydrogeologicalObject.
    </documentation>
  </annotation>
</element>
<element name="description" type="gmd:PT_FreeText_PropertyType"
nillable="true">
  <annotation>
    <documentation>
      Name: Description
      Definition: The description of the HydrogeologicalObject.
    </documentation>
  </annotation>
</element>

```

od HydrogeologicalObject mają przypisanie przestrzenne poprzez atrybut geometryczny typu GM_Primitive, co pozwala na stosowanie geometrii punktowej 0D (GM_Point), liniowej 1D (GM_Curve), powierzchniowej 2D (GM_Surface), a także innych prostych typów geometrycznych będących podtypami GM_Primitive. Specyfikowanie tego typu geometrii jest przedstawione w przykładzie 6.2.

Przykład 6.2. Specyfikacja elementu geometry w schemacie XSD dla danych hydrogeologicznych,

źródło: URL: http://inspire.ec.europa.eu/schemas/ge_hg/3.0/HydrogeologyCore.xsd

```

<element name="geometry" type="gml:GeometricPrimitivePropertyType">
  <annotation>
    <documentation>
      Name: Geometry
      Definition: The geometry defining the spatial location
      of the HydrogeologicalObject.
    </documentation>
  </annotation>
</element>

```

Kolejnym problemem technologicznym jest wzajemne wiązanie obiektów należących do klas ze sobą powiązanych poprzez asocjacje, agregacje i kompozycje. Aby móc takie wiązania zrealizować w zapisie języka GML potrzebne są niepowtarzalne (unikalne) identyfikatory dostępne zewnętrznie. W specyfikacjach danych INSPIRE rolę tę pełnią elementy *inspireId* typu *Identifier* (INSPIRE DT-DS, 2014b). Opisy jak można tworzyć i wykorzystywać te identyfikatory można znaleźć w krajowych opracowaniach, na przykład w publikacji monograficznej z tego zakresu (Michalak, Chojka, Zwirowicz-Rutkowska, Parzyński, 2012) lub w opracowaniu BGWM (2009). Zastosowanie identyfikatora zewnętrznego w schemacie hydrogeologicznym przedstawia przykład 6.3.

Przykład 6.3. Specyfikacja elementów *inspireId* w schemacie XSD dotyczącym danych hydrogeologicznych.

źródło: URL: http://inspire.ec.europa.eu/schemas/ge_hg/3.0/HydrogeologyCore.xsd/

```
<element name="inspireId" type="base:IdentifierPropertyType"
  minOccurs="0">
  <annotation>
    <documentation>
      Name: INSPIRE Identifier
      Definition: External object identifier of the spatial object.
      Description: NOTE: An external object identifier is a unique
        object identifier published by the responsible body, which may
        be used by external applications to reference the spatial
        object. The identifier is an identifier of the spatial object,
        not an identifier of the real-world phenomenon.
    </documentation>
  </annotation>
</element>
```

Deklaracja i definicja identyfikatora INSPIRE jest zawarta w specyfikacji Ogólnego Modelu Pojęciowego GCM (*Generic Conceptual Model*) (INSPIRE DT-DS, 2014b) w formie zalecenia nr 33:

- określeniem wyróżnialności (tożsamości) powinien być atrybut *inspireId*,
- atrybut *inspireId* jest definiowany jako: zewnętrzny identyfikator obiektu przestrzennego,
- definicja ta wymaga uzupełnienia – jest to unikalny identyfikator obiektu opublikowany przez odpowiedzialną za to instytucję, który może być używany przez zewnętrzne aplikacje jako odniesienie do obiektu przestrzennego,
- identyfikator jest identyfikatorem obiektu przestrzennego (agregatu danych), a nie identyfikatorem zjawiska świata rzeczywistego.

Zgodnie z zasadami tworzenia identyfikatorów *inspireId* składa się on z dwóch obowiązkowych elementów (przykład dotyczy Głównego Zbiornika Wód Podziemnych):

- *namespace* – przestrzeń nazw: kraj + organizacja + zestaw danych, przykład: PL.PSH.GZWP,
- *localId* – skrót od typu obiektu przestrzennego + wewnętrzny identyfikator, przykład: GZ000138.

Można tu stosować także trzeci element nieobowiązkowy *versionId* typu *DateTime* dla określenia wersji systemu identyfikacji (*namespace*) i jest to najczęściej czas wprowadzenia tej wersji, na przykład: 2014-12-31T23:00:00Z. Zapis tego identyfikatora w języku GML przedstawia przykład 6.4.

Przykład 6.4. Zapis identyfikatora *inspireId* w języku GML dla obiektu przestrzennego opisującego Główny Zbiornik Wód Podziemnych nr 138.

Identyfikatory „nieunikalne” obiektów są stosowane zazwyczaj dla typów obiektów przestrzennych, dla których podstawowe zbiory danych nie zawierają takich identyfikatorów i gdy wiadomo, że ze znanych powodów nie wymaga się stosowania takich identyfikatorów. Przykładem jest cytat ze specyfikacji tematu Geologia: „Jest możliwe, że typy obiektów przestrzennych w temacie Geologia nie wymagają stosowania identyfikatorów unikalnych”.

```
<ge_hg:inspireId>
  <base:Identifier>
    <base:localId>GZ000138</base:localId>
    <base:namespace>PL.PSH.GZWP</base:namespace>
    <base:versionId>2014-12-31T23:00:00Z</base:versionId>
  </base:Identifier>
</ge_hg:inspireId>
```

Przedstawione tu problemy technologiczne to tylko kilka wybranych najważniejszych spośród wielu. Więcej przykładów można znaleźć w innej publikacji monograficznej dotyczącej tej problematyki (Michalak, Chojka, Zwirowicz-Rutkowska, Parzyński, 2012).

6.3. Modele danych krajowych jako rozszerzenia modelu INSPIRE

W trakcie opracowywania specyfikacji danych INSPIRE w roboczych grupach tematycznych przywiązywano dużą wagę, aby modele zawarte w tych specyfikacjach nie były zbyt rozbudowane i skomplikowane. Przyjęto założenie, że jeżeli modele te mają być stosowane we wszystkich krajach członkowskich, gdzie dotychczasowe praktyki były różne, modele danych INSPIRE powinny stanowić pewnego rodzaju „wspólny mianownik”, czyli powinny w możliwie poprawny sposób zawierać te elementy, które są dla wszystkich wspólne. Dodatkowymi mechanizmami ułatwiającymi stosowanie tych modeli w poszczególnych krajach wspólnoty jest duża liczba elementów nieobowiązkowych (ze stereotypem «voidable») i wariantowość modeli pozwalająca na wybór najbardziej odpowiadającej w danym kraju wersji fragmentu modelu.

Z przedstawionych powyżej powodów, jest rzeczą zrozumiałą, że modele INSPIRE nie mogą zaspokoić wszystkich potrzeb wynikających ze specyfiki przyjętych rozwiązań w danym kraju. Mając do wyboru stosowanie innych modeli danych na potrzeby krajowe i innych dla infrastruktury INSPIRE lub rozszerzenie modeli INSPIRE do potrzeb krajowych, ten drugi wariant jest bez wątpienia korzystniejszy. Postulat takiego podejścia był zgłaszany przez autora tej monografii wielokrotnie dla różnych tematów, jednak z trudnych do zrozumienia powodów nie zyskał akceptacji. W konsekwencji tego, obserwuje się wiele przypadków opracowywania zupełnie nowych modeli w różnych zastosowaniach, znacznie odbiegających od koncepcji przyjętych w INSPIRE z pominięciem faktu, że członkowie roboczych tematycznych zespołów INSPIRE dysponowali znacznie większą wiedzą i doświadczeniem niż często autorzy modeli krajowych.

Nie ma tu miejsca na szczegółowe rozpatrywanie, w jaki sposób należy rozszerzyć model danych INSPIRE podtematu Hydrogeologia, aby spełniał wszystkie potrzeby krajowe. Jest jedynie miejsce na wskazanie drogi metodycznej, w jaki sposób można to wykonać. Poniżej w punktach podane są kolejne etapy prac nad opracowaniem modelu rozszerzonego tą metodą.

- Modele dla wszystkich tematów są zapisane w języku UML i są dostępne w repozytorium modeli INSPIRE, skąd można je pobrać.
- Za pomocą oprogramowania narzędziowego dla języka UML, w tym przypadku programu Enterprise Architect (Woolf, 2009), można na bazie modelu INSPIRE opracować nowy model krajowy poprzez tworzenie nowych klas (na przykład z przedrostkiem „PL_”) będących podtypami klas modelu INSPIRE (rys. 6.4.). Powiązania pomiędzy istniejącymi klasami INSPIRE (w tym przypadku klas: *Aquifer*, *Ground*

WaterBody, ActiveWell i HydrogeologicalObjectManMade) a nowymi klasami modelu krajowego (w tym przypadku PL_Aquifer, PL_GroundWaterBody, PL_ObservationWell i PL_ExtractionWell) są typu dziedziczenia (strzałka z białym trójkątnym grotem) i w konsekwencji nowe klasy dziedziczą wszystkie elementy klas nadrzędnych (atributy i powiązania asocjacyjne) i mogą być uzupełnione nowymi dodatkowymi elementami własnymi. Na tym rysunku (6.4.) kolorem niebieskim są wyróżnione istniejące elementy modelu INSPIRE, kolorem żółtym i brązowym nowe elementy dodane, a kolorem fioletowym odziedziczone asocjacje. Przykład takiego rozszerzenia zbioru elementów przedstawia rysunek 6.5. Do każdej z czterech rozszerzanych klas dodano dla przykładu po trzy nowe atrybuty z przedrostkiem „PL_” i stereotypem «PL_ext», tak aby nowy model mógł spełnić wymagania krajowe.

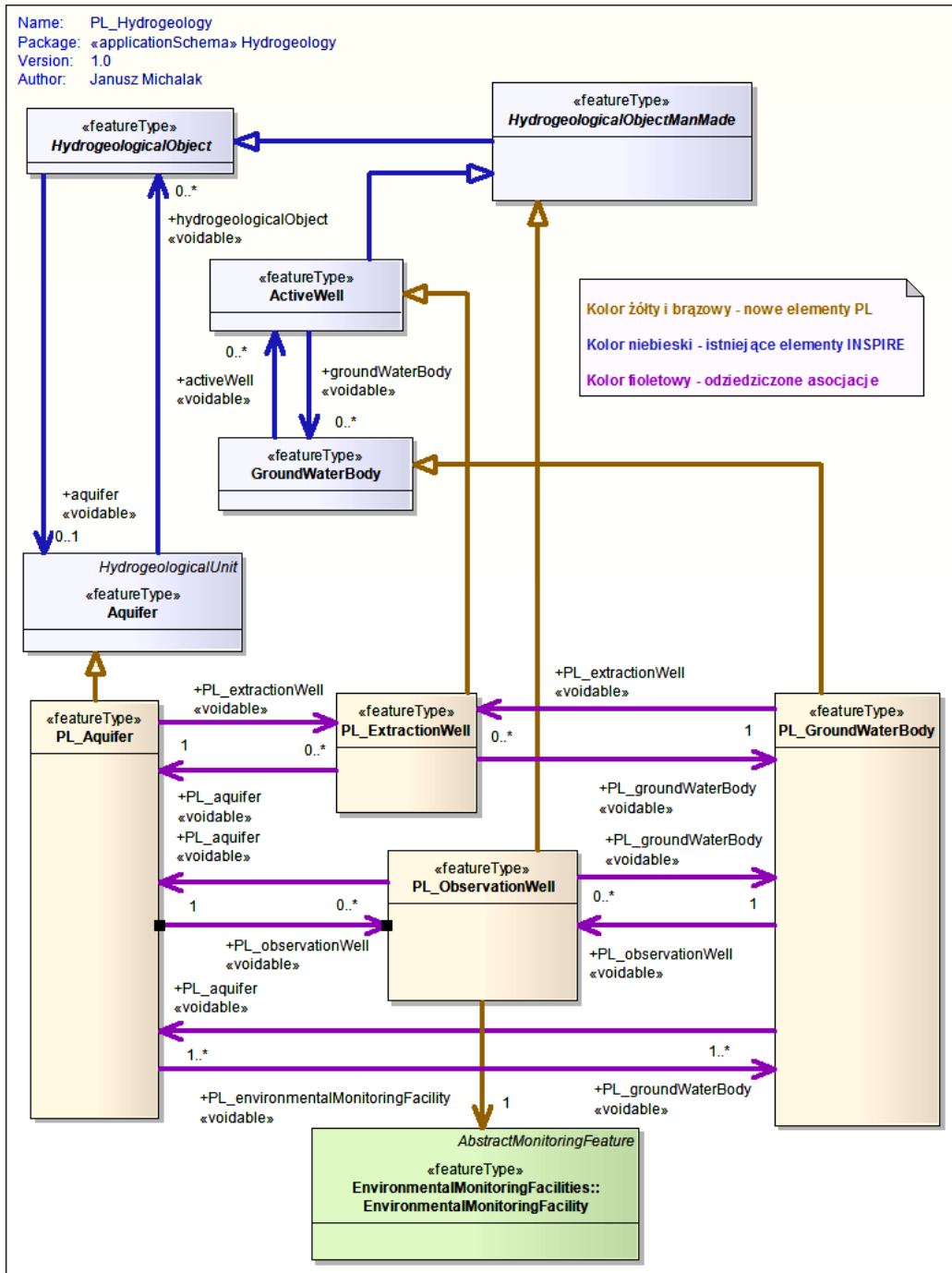
- Opracowany w ten sposób model krajowy (będący rozszerzeniem modelu INSPIRE) może być przekonwertowany do schematów aplikacyjnych XSD dla języka GML. Schematy te, bazujące na schematach INSPIRE i w konsekwencji na schematach GML, stanowią podstawę dla poprawnych zapisów danych przestrzennych z zakresu hydrogeologii.
- Zapisy w języku GML są jedynie zewnętrzną formą zapisów danych przestrzennych. Dane są najczęściej przechowywane w relacyjnych bazach danych. Aby była możliwa jak największa zgodność zapisów zewnętrznych w GML z zapisami wewnętrznymi (w bazie danych) można model w języku UML przekonwertować do języka DDL (*Data Definition Language*) definiującego wewnętrzną strukturę bazy danych. Takie podejście jest także zaproponowane w uwagach metodycznych dotyczących opracowania nowego języka HydroGeoML w rozdziale 5.2.

Przedstawione tu w punktach procedury są szczegółowo opisane w innej monografii dotyczącej tej problematyki (Michalak, Chojka, Zwirowicz-Rutkowska, Parzyński, 2012).

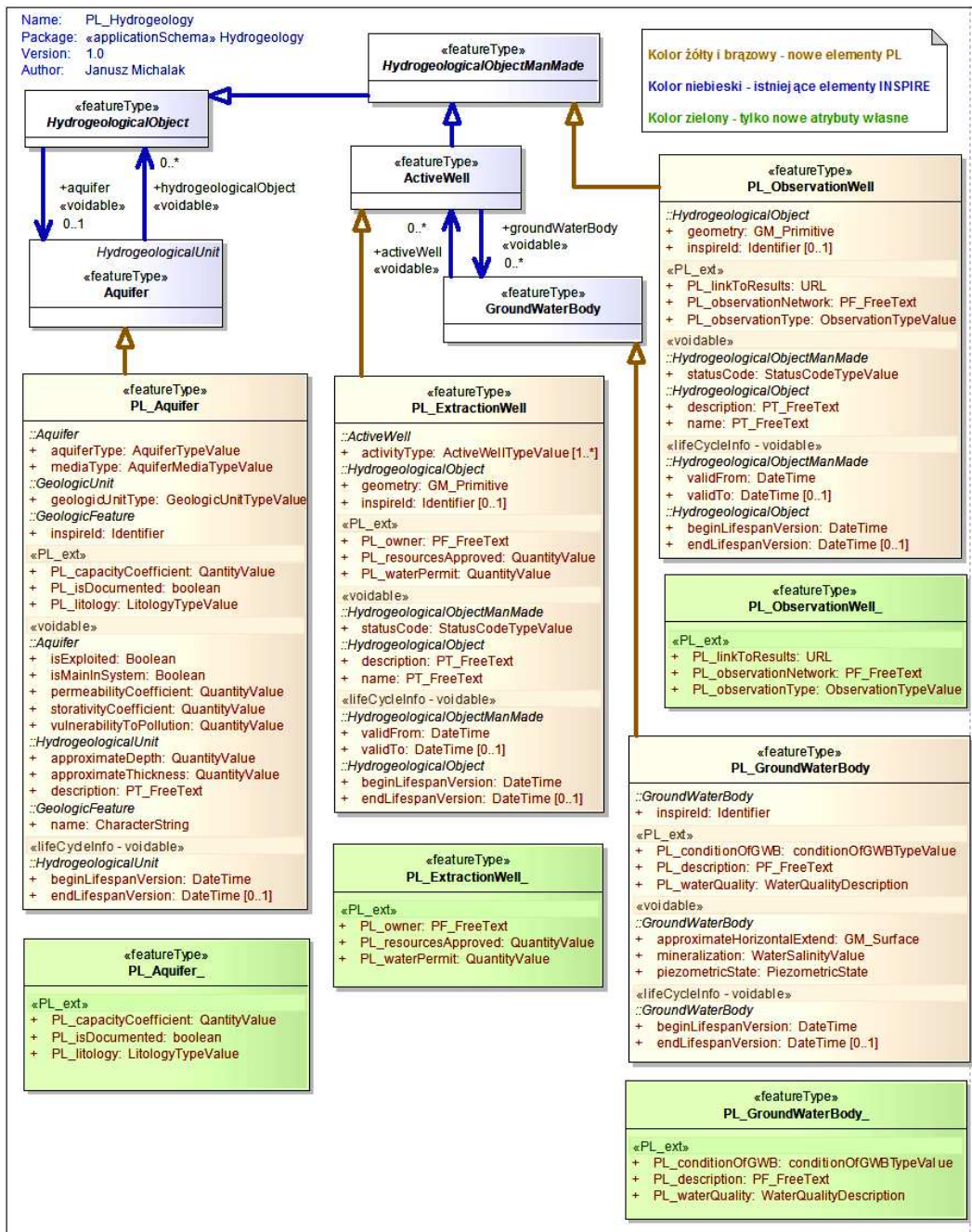
Zastosowanie do potrzeb krajowych modelu danych będącego rozszerzeniem modelu INSPIRE znacznie usprawni korzystanie z danych hydrogeologicznych, z jednoczesnym obniżeniem kosztów w porównaniu do zastosowania dwóch różnych modeli. Takie podejście pozwoli na wykorzystanie infrastruktury tego tematu do dwóch kategorii zadań: tych które wynikają z dyrektywy INSPIRE i tych które mają specyfiką krajową. W takim przypadku architektura infrastruktury składa się z dwóch poziomów:

- poziomu podstawowego (bazowego) dedykowanego danym określonym w dokumentach INSPIRE – model tych danych w podstawowej wersji wektorowej jest stosunkowo prosty,
- poziomu bazowego modelu INSPIRE nadbudowanego przez dodanie nowych typów wyróżnień (ze stereotypem «FeatureType») i nowych atrybutów, zarówno w typach bazowych, jak i w typach dodanych – niezbędnego dla spełnienia bardziej szczegółowych wymagań krajowych.

Za pomocą prostych stosunkowo transformacji można oddzielić od pełnego zbioru danych dane dedykowane potrzebom krajowym, a pozostała część zbioru będzie zawierała jedynie elementy wymagane przez specyfikacje INSPIRE. W takim rozwiązaniu usługi INSPIRE będą udostępniały tylko elementy INSPIRE, a usługi dla potrzeb krajowych będą mogły udostępniać wszystkie elementy określone w pełnym modelu danych lub odrzucić część elementów INSPIRE, które w jakimś zastosowaniu nie są potrzebne.



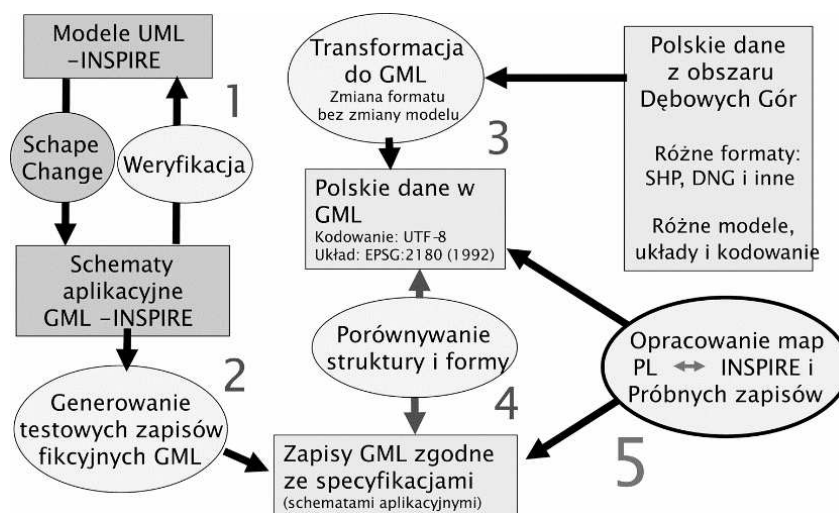
Rysunek 6.4. Diagram UML przedstawiający rozszerzenie podstawowego modelu INSPIRE dla potrzeb krajowych przez dodanie nowych klas z przedrostkiem "PL " na zasadzie dziedziczenia, na przykład nowa klasa PL_Aquifer jest podtypem istniejącej klasy Aquifer. Nowe klasy dziedziczą także powiązania asocjacyjne (strzałki fioletowe).



Rysunek 6.5. Diagram UML przedstawiający nowe klasy utworzone w sposób przedstawiony na rysunku 6.4. Do nowych klas zostały dodane nowe atrybuty z przedrostkiem "PL_" i stereotypem «PL_ext». Objasnienia w tekście.

7. Analiza różnic pomiędzy modelem krajowym a INSPIRE

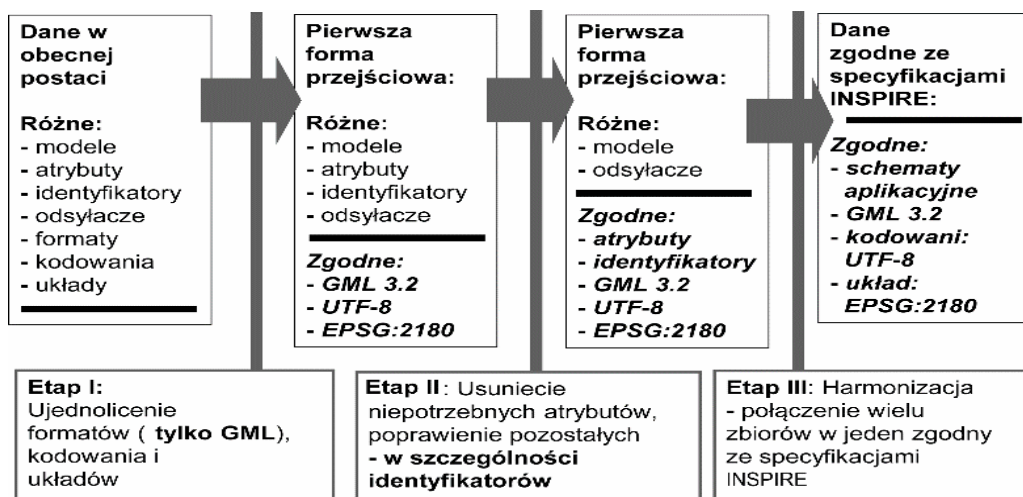
W rozdziale 6. opisany został podstawowy (*core*) model INSPIRE dla danych hydrogeologicznych. Z przeprowadzonej tam analizy wynika, że daleko posunięta prostota tego modelu stwarza istotne trudności w praktycznym zastosowaniu danych zapisanych zgodnie z tym modelem i potrzebne jest rozszerzenie modelu INSPIRE o wiele nowych elementów. Istnieje również poważny problem w relacji odwrotnej. Polega on na trudnościach przekształcenia (transformacji) danych krajowych do struktur i form określonych w specyfikacjach INSPIRE. Temu zagadnieniu poświęcone były prace przeprowadzone w roku 2011 w instytucjach pełniących role dysponentów danych. Prace te były przeprowadzone przez Państwową Służbę Hydrogeologiczną wraz z Państwową Służbą Geologiczną. Autor tej monografii pełnił wtedy rolę krajowego koordynatora z ramienia Rady Infrastruktury Informacji Przestrzennej. Wyniki, a także wnioski wynikające z tych prac, obejmujących wszystkie organy wiodące zostały opublikowane (Michalak, 2012), a oddzielne publikacje (Nowakowska, Michalak i inni, 2011; Nowakowska, Hordejuk i inni, 2012) dotyczą tematów, dla których organem wiodącym jest Główny Geolog Kraju: Geologia, Geofizyka, Hydrogeologia, Zasoby energetyczne (INSPIRE TWG-ER, 2014) i Zasoby mineralne (INSPIRE TWG-GE-MR, 2013). Między innymi opracowana została metodyka prac testowych, w której proces testowania dzieli się na 5 etapów (rys. 7.1.) (Michalak, 2011).



Rysunek 7.1. Ogólny algorytm prac testowych – podział na 5 etapów. Opracowany dla tematów aneksu I i wykorzystany w tematach aneksów II i III.

- Etap 1. Analiza modeli danych dla poszczególnych tematów. Na drodze inżynierii odwrotnej ze schematów XSD dla INSPIRE utworzonych za pomocą oprogramowania SchapeChange i przechowywanych w repozytorium INSPIRE odtworzone zostały modele UML, a wyniki porównano z odpowiednimi modelami w wersji pierwotnej dostępnymi w repozytorium.
- Etap 2. Na podstawie schematów XSD za pomocą oprogramowania narzędziowego – edytora XML Spy i edytora Oxygen zostały wygenerowane testowe zapisy danych w GML zgodne w pełni z wymaganiami tych schematów, jednak zapisana w nich treść z konieczności była całkowicie fikcyjna. Te testy miały za zadanie wykazanie jak wygląda poprawna struktura zapisu.
- Etap 3. Transformacja wybranych próbek danych krajowych do zapisu znacznikowego (XML) zgodnego ze schematami języka GML. Jest to tylko zmiana formatu bez zmiany struktury (modelu) danych. Do tego celu zebrane zostały różnorodne dane obejmujące obszar Dębowych Gór (północne Mazowsze).
- Etap 4. Porównanie struktury i formy zbiorów danych uzyskanych w etapie 2. (poprawne formalnie z niepoprawną treścią) ze zbiorami uzyskanymi w etapie 3. (poprawna treść, ale niepoprawna forma i struktura).
- Etap 5. W tym końcowym etapie zostaje opracowane mapowanie elementów danych: jakim elementom struktur INSPIRE odpowiadają elementy danych krajowych. Mapy takie są między innymi podstawą do opracowania algorytmów transformacji zapisywanych w tak zwanych arkuszach stylu, które następnie można wykorzystać do sterowania procesorem XSLT (*XSL Transformations*).

Przy okazji opisanego powyżej procesu testowania przeprowadzono także inne eksperymenty badawcze ukierunkowane na bardziej szczegółową metodykę transformacji danych. W tym przypadku proces transformacji dzieli się na trzy etapy (rys. 7.2.).



Rysunek 7.2. Ogólny algorytm transformacji danych krajowych do formy i treści zgodnej z wymaganiami specyfikacji danych INSPIRE.

W tym przypadku procedura została podzielona na trzy etapy.

- Przed etapem pierwszym dane krajowe mają pierwotną postać i nie spełniają wymagań INSPIRE w zakresie: modeli (struktur), atrybutów poszczególnych wyróżnień, identyfikatorów wyróżnień, powiązań pomiędzy wyróżnieniami w formie odsyłaczy, formatu zapisu danych, kodowania danych tekstowych i układu odniesienia przestrzennego.
- Etap pierwszy polega na ujednoczeniu kodowania, formy zapisu w postaci GML i transformacji współrzędnych do jednolitego układu. Aby zapisać dane krajowe w ich własnej strukturze za pomocą języka GML potrzebne jest opracowanie odpowiednich dla nich schematów XSD. Schemat taki z dużymi uproszczeniami dla danych dotyczących jednolitych części wód podziemnych (JCWPd) jest przedstawiony w przykładzie 7.1.
- W rezultacie powstaje pierwsza forma przejściowa z układem odniesienia EPSG:2180 (układ 1992), kodowana w Unicode UTF-8 i zapisana w języku GML. Jeżeli jest to możliwe, to wskazana jest wersja 3.3. Bardzo istotne jest, aby zmiana kodowania była przeprowadzona przed transformacją do zapisu znacznikowego (XML), ponieważ zapis znacznikowy z różnym kodowaniem jest praktycznie bezużyteczny. Nadal struktura modelu, atrybuty, identyfikatory i odsyłacze odbiegają od wymagań INSPIRE.
- W etapie drugim usuwa się niepotrzebne atrybuty wyróżnień, pozostałe należy poprawić tak, aby były zgodne z wymaganiami, w szczególności dotyczy to identyfikatorów.
- Rezultat uzyskany w drugim etapie odbiega od wymagań pod względem struktury danych i często elementy danych dla jednego zbioru wynikowego są rozproszone w kilku zbiorach uzyskanych w tym etapie. W przeciwieństwie do zagnieżdżonych struktur INSPIRE, zbiory krajowe mają nadal struktury płaskie – dają się przedstawić w formie prostych pojedynczych tabel. Również powiązania pomiędzy elementami są niepoprawne lub najczęściej w ogóle ich nie ma.
- Etap trzeci – ostatni, sprowadza się do wzajemnej harmonizacji danych krajowych uzyskanych w etapie drugim. W etapie tym następuje przeniesienie elementów wymaganych przez specyfikację INSPIRE z wielu plików pierwotnych (uzyskanych w etapie drugim) do jednego pliku wynikowego z wykorzystaniem mapowań opracowanych wcześniej i zapisanych w arkuszach XSL. W tym etapie buduje się także powiązania odpowiadające asocjacji jednokierunkowym w modelu UML i są one przeznaczone do nawigacji pomiędzy elementami powiązanymi ze sobą semantycznie.
- Wynikiem trzeciego etapu są zbiory danych całkowicie zgodne ze specyfikacjami INSPIRE. Pozostaje jedynie problem braku tych elementów danych, które są wymagane przez INSPIRE, ale których nie ma w źródłowych zbiorach danych krajowych.

Przykład 7.1. Plik JCWPd.xsd – uproszczony i skrócony schemat XSD dla danych krajowych dotyczących jednolitych części wód podziemnych (JCWPd):

```
<?xml version="1.0" encoding="UTF-8"?>
<schema xmlns:hg="http://www.geomlab.pl/gml/hg"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:gml="http://www.opengis.net/gml/3.2"
xmlns="http://www.w3.org/2001/XMLSchema"
targetNamespace="http://www.geomlab.pl/gml/hg"
elementFormDefault="qualified">
  <import namespace="http://www.opengis.net/gml/3.2"
schemaLocation="../../../ogc/gml/3.2.1/gml.xsd"/>
  <element name="jcwpd" type="hg:jcwpdType"
substitutionGroup="gml:AbstractFeature"/>
  <complexType name="jcwpdType">
    <complexContent>
      <extension base="gml:AbstractFeatureType">
        <sequence>
          <element name="nr_JCWPd" minOccurs="1" maxOccurs="1">
            <simpleType>
              <restriction base="decimal">
                <totalDigits value="10"/>
                <fractionDigits value="0"/>
              </restriction>
            </simpleType>
          </element>
          <element name="powierzchnia_km2" minOccurs="1"
maxOccurs="1">
            <simpleType>
              <restriction base="decimal">
                <totalDigits value="19"/>
                <fractionDigits value="8"/>
              </restriction>
            </simpleType>
          </element>
          <element name="dorzecze" minOccurs="1" maxOccurs="1">
            <simpleType>
              <restriction base="string">
                <maxLength value="250"/>
              </restriction>
            </simpleType>
          </element>
          <element name="kod_UE" minOccurs="1" maxOccurs="1">
            <simpleType>
              <restriction base="string">
                <maxLength value="50"/>
              </restriction>
            </simpleType>
          </element>
          <!-- Element określający geometrię dla JCWPd -->
          <element ref="gml:surfaceProperty" minOccurs="1"
maxOccurs="1"/>
          <!-- (...) -->
        </sequence>
      </extension>
    </complexContent>
  </complexType>
</schema>
```

7.1. Mapowanie elementów modeli

Kluczowym, ale jednocześnie bardzo żmudnym procesem jest opracowywanie mapowania pomiędzy elementami modelu INSPIRE dla hydrogeologii a elementami krajowych danych hydrogeologicznych. Mapowanie to, jak przedstawiono to wcześniej, jest podstawowym składnikiem niezbędnego zestawu informacji, jakie są potrzebne dla przeprowadzenia poprawnej transformacji danych. W ramach prac wykonanych w Państwowym Instytucie Geologicznym w latach 2011–2013 z zakresu testowania specyfikacji opracowane zostały tabele mapowania dla danych geologicznych i hydrogeologicznych (Nowakowska, Michałak i inni, 2011; Nowakowska, Hordejuk i inni, 2012). Jedną z nich – tabelę mapowania dla podstawowych danych hydrogeologicznych zawiera ponad 210 pozycji. Jednak w zdecydowanej większości przypadków brak jest krajowego odpowiednika dla elementu występującego w modelu INSPIRE. Tabela 7.1. przedstawia kilkanaście wybranych przykładów wyniku mapowania z zakresu hydrogeologii.

Sam fakt znalezienia odpowiednika krajowego dla elementu INSPIRE nie rozwiązuje całkowicie problemu. Kolejną trudność to niezgodność typów danych powiązanych wzajemnie przez proces mapowania. Najbardziej jest to widoczne w przypadku list kodowych, gdzie poszczególne pozycje po obu stronach nie pokrywają się swoim znaczeniem lub nawet w listach krajowych nie ma pozycji odpowiadających pozycjom w listach INSPIRE. Częściowo ten problem można rozwiązać dla list otwartych INSPIRE, gdzie istnieje możliwość dopisywania nowych pozycji. Do takich list w modelu hydrogeologicznym opisanym w specyfikacji INSPIRE tematu Hydrogeologia należą: Hydrogeochemiczny rodzaj skały (`HydroGeochemicalRockTypeValue`), Rodzaj obiektu naturalnego (`NaturalObjectTypeValue`), Rodzaj ośrodka warstwy wodonośnej (`AquiferMediaTypeValue`), Rodzaj studni aktywnej (`ActiveWellTypeValue`), Typ kodu statusu (obiektu hydrogeologicznego) (`StatusCodeTypeValue`) i Zasolenie wody (`Water-SalinityValue`). Listy nierozszerzalne (enumeratory) to: Okresowość wód (`Water-PersistenceValue`), Rodzaj warstwy wodonośnej (`AquiferTypeValue`) i Stan wód podziemnych (`ConditionOfGroundwaterValue`). W tym przypadku, gdy występuje niezgodność, jedynym rozwiązaniem jest zastosowanie mapowania, tak jak to się robi w przypadku atrybutów. Inne problemy, to na przykład konieczność przeliczania wartości liczbowych z powodu użycia różnych jednostek miar.

Tabela 7.1. Przykład mapowania pomiędzy elementami modelu INSPIRE dla hydrogeologii a elementami krajowych danych hydrogeologicznych.

Atrybut modelu INSPIRE ← Nazwa klasy	Typ atrybutu i wymaganie	Element danych krajowych
activityType ← ActiveWell	ActiveWellTypeValue (code list), non-voidable	PSHv7 – CBDH – CELE_UZYTEKOWANIA_OTWORU
approximateHorizontalExtend ← GroundWaterBody	GM_Surface, voidable	PSHv7 – JCWP_GWB – GEOMETRY
approximatePermeabilityCoefficient ← Aquitard	QuantityValue (union), voidable	Brak
approximateQuantityOfFlow ← HydrogeologicalObjectNatural	QuantityValue (union), voidable	PSHv7 – CBDH – WYDAJNOSC_EKSP_OTW – WYDAJNOSC
conditionOfGroundWaterBody ← GroundWaterBody	ConditionOfGroundwaterValue (code list), non-voidable	PSHv7 – GEOMETRIE – JCWP_GWB – ZAGROZONE
description ← HydrogeologicalUnit (Aquiclude)	PT_FreeText, voidable	MHP_GIS – MHP_GIS – JEDN_HYDROGEOL – UWAGI
elevation ← BoreholeCollar	DirectPosition, non-voidable	CBDG – Podsystem OTWORY – OTWORY
EU_GWB_code ← WFD_GroundWaterBody	FeatureUniqueEUCodeType, non-voidable	PSHv7 – GEOMETRIE – JCWP_GWB
geometry ← HydrogeologicalObject	GM_Primitive, non-voidable	PSHv7 – GEOMETRIE – GEOM_PUNKTY
geometry ← HydrogeologicalObject	GM_Primitive, non-voidable	PSHv7 – GEOMETRIE – GEOM_PUNKTY
GWB_MS_code ← WFD_GroundWaterBody	FeatureUniqueCodeType, non-voidable	PSHv7 – GEOMETRIE – JCWP_GWB
GWB_name ← WFD_GroundWaterBody	PT_FreeText, voidable	PSHv7 – GEOMETRIE – JCWP_GWB
isConfined ← Aquifer	Boolean, non-voidable	MHP_GIS – JEDN_HYDROGEOL – SL_STOP_IZOLACJI
isExploited ← Aquifer	Boolean, voidable	MHP_GIS
isMainInSystem ← Aquifer	Boolean, voidable	MHP_GIS
location ← BoreholeCollar	GM_Point, non-voidable	CBDG – Podsystem OTWORY
name ← HydrogeologicalObject	PT_FreeText, voidable	PSHv7 – CBDH – OTWORY – NAZWA
name ← HydrogeologicalObject	PT_FreeText, voidable	PSHv7 – CBDH – OTWORY
naturalObjectType ← HydrogeologicalObjectNatural	NaturalObjectTypeValue (code list), non-voidable	PSHv7 – CBDH – OTWORY – SL_TYP_OTWORU_ID
statusCode ← HydrogeologicalObjectManMade	StatusCodeTypeValue (code list), voidable	PSHv7 – CBDH – OTWORY – L_STANY_EKSPLOATACJI
thickness ← Aquifer	QuantityValue (union), voidable	MHP_GIS – MIAZSZOSC_PWOD
URL ← WFD_GroundWaterBody	URL, voidable	PSHv7 – GEOMETRIE – JCWP_GWB
validFrom ← HydrogeologicalObjectManMade	DateTime, non-voidable	PSHv7 – CBDH – OTWORY – DATA_WYKONANIA
vulnerabilityToPollution ← Aquifer	QuantityValue (union), voidable	MHP_GIS – STOPIEN_ZAGROZ_PWOD

8. Podstawy metodyczne transformacji

Przyjęta koncepcja technologiczna infrastruktury INSPIRE zakłada, że nie jest istotna forma i struktura danych INSPIRE przechowywanych w jakiejś bazie, lecz forma i struktura danych, które z tej bazy wychodzą na zewnątrz. Specyfikacje danych dla poszczególnych tematów określają zawartość i strukturę danych wychodzących na zewnątrz w celu przesyłania do innego systemu lub udostępniania użytkownikowi. Koncepcja ta określa, że wewnętrzną formą zapisu danych jest język GML w wersji 3.2. Z tego wynika, że proces transformacji powinien być ukierunkowany na uzyskanie wyniku końcowego w postaci zapisu w języku GML 3.2 i zgodnego z wymaganiami poszczególnych specyfikacji tematycznych (Curtis, 2009; Latvala, Lehto, 2011; Michalak, 2010b; Murray, 2010).

W jaki sposób dane po transformacji mają być zapisane w bazie nie jest tu nigdzie określone i poszczególni dysponenti danych mogą w tym zakresie postępować dowolnie. Jednak zbyt daleko posunięta dowolność może się okazać poważnym obciążeniem i w efekcie może doprowadzić do znacznych trudności w przetwarzaniu danych z formy wewnętrznej do formy zewnętrznej w trakcie ich wyprowadzania na zewnątrz. Taka sytuacja ma miejsce, gdy wewnętrzna struktura i zawartość znacznie się różni od struktury i zawartości wymaganej przez specyfikacje INSPIRE. Z tego względu optymalnym rozwiązaniem jest przechowywanie danych w formie i strukturze jak najbardziej zbliżonej do wymaganej formy i struktury zewnętrznej. Koncepcja technologiczna INSPIRE przyjmuje dwa tryby transformacji opisane już częściowo we wstępie.

- Transformacja *off-line*. Niezależny proces, w którym dane pierwotne ze swoją formą i strukturą są w całości przekształcane do formy i struktury wynikowej, a następnie zapisywane w statycznych plikach lub w przygotowanej wcześniej nowej bazie danych spełniającej wymagania INSPIRE. W takim przypadku przetwarzanie danych towarzyszące ich wyprowadzaniu na zewnątrz jest ograniczone do minimum. Jednym z problemów jakie się w tym przypadku ujawniają jest zagadnienie aktualizacji bazy danych i może to być rozwiązanie za pomocą technologii przyrostowego zasilania bazy. W takim przypadku, aktualizacja jest wykonywana za pomocą odpowiednio zaprojektowanych plików, zawierających zapis zmian jakie mają być dokonane w zapisach danych w bazie. Transformacja *off-line* jest efektywnym rozwiązaniem w przypadkach gdy forma, treść i struktura danych w dotychczasowych rozwiązaniach znacznie się różni od tego co określają specyfikacje INSPIRE. Takie przypadki najczęściej występują w krajach, gdzie dominują technologie tradycyjne i gdzie zastosowanie języka GML ciągle jest na etapie eksperymentów. Do takich krajów należy także Polska.
- Transformacja *on-line*. Przekształcanie danych wykonywane jest na bieżąco przy każdej odpowiedzi serwera na polecenie przekazania określonej porcji danych znajdujących się w systemie. Jest to proste i efektywne rozwiązanie pod warunkiem, że forma i struktura danych wewnątrz systemu niewiele się różni od formy i struktury zewnętrznej. Aby ten warunek mógł być spełniony struktura i forma zbioru wewnętrzny (bazy danych) musi odpowiadać przynajmniej schematom aplikacyjnym języka

GML i mogą to być inne jego wersje niż przyjęta w INSPIRE wersja 3.2. Sytuacja taka ma miejsce w krajach, w których zastosowanie języka GML jest już powszechnie przyjęte.

Specyfiką zapisu danych w GML, jako aplikacji języka XML dla informacji geoprzestrzennej, jest jego złożoność strukturalna. Złożoność ta sprowadza się między innymi do możliwości wielokrotnego zagnieżdżania elementów, hierarchicznego sposobu definiowania elementów na bazie innych już zdefiniowanych, nawigacyjnego wiązania elementów odległych za pomocą odsyłaczy. Język XML jest językiem bardzo prostym, lecz duża swoboda w definiowaniu poszczególnych elementów składowych i całej struktury stwarza sytuację, w której można opracować schemat niepotrzebnie zbyt skomplikowany, niepoprawny lub bezużyteczny. Dla wstępnego pokazania możliwości tego języka są tu podane podstawowe jego reguły syntaktyczne w postaci przedstawionych poniżej bardzo prostych przykładów (8.1. do 8.5.).

Przykład 8.1. Podstawowym fragmentem informacji zapisanej w XML jest „element” ograniczony znacznikiem początku i końca:

```
<jakiśElement> Wnętrze_jakiegośElementu </jakiśElement>
      ↑                               ↑
      znacznik początku             znacznik końca
```

Przykład 8.2. Znacznik początku może zawierać atrybuty, na przykład nazwę:

```
<jakiśElement nazwa="NazwaJakiegośElementu">
```

Przykład 8.3. Element złożony zawiera w sobie inne elementy – proste lub również złożone:

```
<elementZłożony><elementProsty>"coś"</elementProsty></elementZłożony>
```

Przykład 8.4. Bardziej czytelny zapis wymaga rozłożenia pionowego i wcięć dla elementów wewnętrznych (zapis zagnieżdżony):

```
<elementZłożony>
  <elementProsty>
    "coś"
  </elementProsty>
  <innyElementProsty>
    "coś innego"
  </innyElementProsty>
</elementZłożony>
```

Zarówno wewnątrz elementu, jak i jego atrybuty zawierają informację. Jednak oba sposoby zapisu informacji nie są równoważne – wewnątrz zawiera treść informacji, a atrybuty zawierają informację o informacji zawartej wewnątrz elementu, czyli w pewnym sensie „metadane lokalne”.

Przykład 8.5. Jeżeli element zawiera informację tylko w formie argumentów, to stosuje się skróconą formę zapisu ograniczoną tylko do jednego znacznika:

```
<odwołaniaDoInnegoElementu xlink:href="#innyElement"/>
```

Informacja zawarta we wnętrzu (treść) i w argumentach ma zdefiniowany typ. Jest wiele prostych typów zdefiniowanych wstępnie – łącznie ponad 40. Oto niektóre z nich: *string*, *boolean*, *ENTITY*, *ID*, *IDREF*, *QName*, *binary*, *decimal*, *float*, *uriReference*, *timeDuration* i *recurringDuration*. Na ich podstawie definiowane są własne typy proste, jeżeli jest taka potrzeba. Wszystkie elementy, jakie mogą wystąpić w określonym miejscu zapisu z zastosowaniem aplikacji XML i ograniczenia dotyczące tych elementów są zdefiniowane w zapisie (dokumencie) zdefiniowanym przez *XML Schema* (XSD) (Bray, Paoli, Sperberg-McQueen, 1998; Mercer, 2000). W starszych aplikacjach posługiwano się inną specyfikacją – XML DTD (*Document Type Definition*). Bardzo prosty schemat jest przedstawiony w przykładzie 8.6., a schemat zgodny z wymaganiami języka GML i odwołujący się do niego jest przedstawiony w poprzednim rozdziale w przykładzie 7.1.

Przykład 8.6. Przykład prostego schematu XSD:

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  elementFormDefault="qualified"
  attributeFormDefault="unqualified">
  <xs:element name="OkresCzasu">
    <xs:annotation>
      <xs:documentation>
        Prosty przykład schematu
      </xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:sequence>
        <xs:element name="Nazwa"/>
        <xs:element name="Początek"/>
        <xs:element name="Koniec"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

Przykład 8.7. Zapis odpowiadający schematowi z przykładu 8.6.:

```
<?xml version="1.0" encoding="UTF-8"?>
<OkresCzasu
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://laboratorium-geomatyki.pl/schemas/czas.xsd">
  <Nazwa>Jakiś okres czasu</Nazwa>
  <Początek>1.1.1996</Początek>
  <Koniec>31.12.2003</Koniec>
</OkresCzasu>
<!-- ... -->
```

Stosowane w schematach XML przestrzenie nazw pozwalają na strukturalną organizację wielu schematów ze sobą powiązanych lub od siebie zależnych. Jest to konieczne w przypadkach, gdy jeden zapis opiera się na wielu schematach, na przykład GML3 zawiera 27 schematów posługujących się czterema przestrzeniami nazw. Trzeba tu także zaznaczyć, że różne języki podrzędne (aplikacje XML) mogą być ze sobą mieszane lub jeden język może być rozwinięciem innego – wykorzystywać schematy drugiego języka. Jednak takie rozwiązania wymagają przestrzegania reguł opracowań standardowych – podziału kompetencji i kompletności w zakresie własnego obszaru tematycznego.

W rozpatrywanych tu przypadkach zbiór wyników jest zawsze zapisywany w formie znacznikowej, jako aplikacja języka GML. Jednak zbiór źródłowy nie jest najczęściej zapisem znacznikowym, chociaż taka forma jest najbardziej pożądana, ponieważ pozwala na zastosowanie standardowych metod transformacji z zastosowaniem technologii XSLT opisanej bardziej szczegółowo w rozdziale 9. W przypadkach, gdy dane źródłowe są zapisane w inny sposób, a jest to najczęściej bardzo prosty format Esri Shape, dla zastosowania XSLT konieczna jest zmiana formy zapisu danych źródłowych do formy znacznikowej opartej na XML. Jednak pomimo zastosowania tej formy struktura danych jest płaska, to znaczy daje się przedstawić w formie prostej tabeli prostokątnej bez żadnych zagnieżdżeń i bez powiązań nawigacyjnych. Bardzo często zastosowanie tak prostej formy zapisu prowadzi do utraty części informacji. W wielu przypadkach można uniknąć tej straty części informacji przez umieszczenie danych źródłowych w kilku uzupełniających się plikach płaskich. Konsekwencją tego jest fakt, że pojedynczy znacznikowy plik wyników będzie musiał być utworzony z kilku plików źródłowych.

Są jednak metody pozwalające na przynajmniej częściowe zachowanie informacji o zagnieżdżonych strukturach zapisywanych w modelu płaskim. Jedną z nich jest notacja kropkowa zilustrowana przykładem 8.8. Jest to zapis płaski odpowiadający danym zagnieżdżonym z przykładu 8.7., w którym element `OkresCzasu` zawiera trzy zagnieżdżone elementy wewnętrzne `Nazwa`, `Początek` i `Koniec` wraz z zapisanymi tam danymi. Dzięki zapisowi kropkowemu tę strukturę można odtworzyć.

Przykład 8.8. Zapis kropkowy w strukturze płaskiej zachowujący strukturę zagnieżdżoną przedstawioną w przykładzie 8.7.

```
OkresCzasu.Nazwa="Jakiś okres czasu"  
OkresCzasu.Początek="1.1.1996"  
OkresCzasu.Koniec="31.12.2003"
```

8.1. Przegląd metod

W literaturze dotyczącej tej problematyki można znaleźć wiele metod przekształcania jednych zbiorów danych w inne o innej organizacji, lecz na ogół nie są to rozwiązania standardowe i są związane tylko ze ściśle określonym oprogramowaniem narzędziowym. Do nich należy między innymi metodyka szablonów (*template*) stosowana w oprogramowaniu (FME) firmy Safe Software. Oprogramowanie to jest opisane w rozdziale 11.1. W tym przypadku szablon to wzorec zapisu danych w języku GML zgodny ze schematami XSD dla określonego tematu INSPIRE. W szablonie takim są odpowiednie pola przewidziane do wstawienia tam określonych elementów danych z jakiegoś zbioru źródłowego. Operacja wstawiania

tych danych jest realizowana za pomocą funkcji `{fme:get-attribute („NAZWA_ELEMENTU_W_ZBIORZE_ZRÓDŁOWYM“)}`. Przykład 8.8. przedstawia fragment takiego szablonu dla tematu Obszary chronione.

Przykład 8.9. Fragment szablonu dla transformacji oprogramowaniem FME dla tematu Obszary chronione:

```
<!-- (...) -->
<gmd:CI_ResponsibleParty>
  <gmd:organisationName>
    <gco:CharacterString>
      {fme:get-attribute("INSPIRE_ORGANISATIONNAME")}
    </gco:CharacterString>
  </gmd:organisationName>
  <gmd:contactInfo>
    <gmd:CI_Contact>
      <!-- Phone, etc can be added -->
      <!-- We can add more information later -->
      <gmd:address>
        <gmd:CI_Address>
          <gmd:electronicMailAddress>
            <gco:CharacterString>
              {fme:get-attribute("INSPIRE_ELECTRONICMAILADDRESS")}
            </gco:CharacterString>
          </gmd:electronicMailAddress>
        </gmd:CI_Address>
      </gmd:address>
    </gmd:CI_Contact>
  </gmd:contactInfo>
  <gmd:role>
    <gmd:CI_RoleCode codeList="http://standards.iso.org/ittf/
      PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/
      gmxCodelists.xml#CI_RoleCode" codeListValue="owner">
      {fme:get-attribute("INSPIRE_ROLE")}
    </gmd:CI_RoleCode>
  </gmd:role>
</gmd:CI_ResponsibleParty>
<!-- (...) -->
```

Inna niestandardowa metoda transformacji jest zastosowana w oprogramowaniu HALE (Humboldt Alignment Editor) opisanym w rozdziale 11.2. W tym przypadku mapowanie zapisywane jest w dokumencie XML, którego fragment jest przedstawiony w przykładzie 8.10. W dokumencie tym główny element `alignment` stanowi pojemnik dla wielu wewnętrznych elementów `cell` definiujących powiązanie atrybutu obiektu źródłowego (tu: `uso_pol_part2`) z atrybutem obiektu wynikowego (tu: `ZoningElementType`). Wartości atrybutów `type` wskazują, że przykład ten dotyczy transformacji zapisu w formacie Esri Shape danych z zakresu planowania przestrzennego do języka GML zgodnie ze schematem PLU (*Planned Land Use*) w wersji 3.0rc3.

Przykład 8.10. Fragment zapisu alignment w dokumencie XML dla transformacji oprogramowaniem HALE dla tematu Zagospodarowanie przestrzenne (*Land use*):

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<alignment xmlns="http://www.esdi-humboldt.eu/hale/alignment">
  <cell relation="eu.esdihumboldt.hale.align.retype"
    id="C4205a1e9-e95e-4274-98a1-ff06add58808" priority="normal">
<!-- Określenie elementu w zbiorze źródłowym -->
    <source>
      <class>
        <type name="uso_pol_part2"
          ns="http://www.esdi-humboldt.eu/hale/shp"/>
      </class>
    </source>
<!-- Określenie elementu w zbiorze wynikowym -->
    <target>
      <class>
        <type name="ZoningElementType"
          ns="http://inspire.ec.europa.eu/schemas/plu/3.0rc3"/>
      </class>
    </target>
  </cell>
<!-- (...) -->
</alignment>
```

W obu przedstawionych powyżej przypadkach mamy do czynienia z zapisem XML, czyli zwykłym tekstem, który jest czytelny i można w nim dokonywać modyfikacji. Jednak oba te rozwiązania są niestandardowe i związane z konkretnym oprogramowaniem. W pełni standardowym rozwiązaniem dla tego typu transformacji jest jedynie wspomniana wcześniej technologia XSLT, która jest przedstawiona bardziej szczegółowo w rozdziale 9.

9. Transformacje XSLT

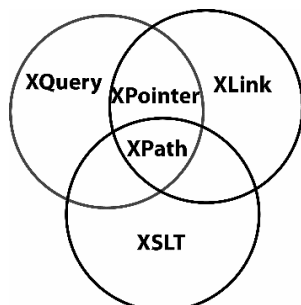
Akronim XSLT oznacza przekształcenia XSL (*XSL Transformations*). Technologia ta służy do transformacji dokumentów XML z jednych schematów do innych schematów, takich jak na przykład XHTML. XSL (*Extensible Stylesheet Language*) jest językiem arkuszy stylów dla dokumentów XML. XSLT jest rekomendacją Konsorcjum W3C, a pełen szczegółowy opis zawiera dokumentacja dostępna w witrynie tego konsorcjum. XSLT zawiera ponad 100 wbudowanych funkcji. Istnieją funkcje dla tekstów (ciągów znaków), wartości liczbowych, porównywania daty i czasu, manipulacji węzłami i kanonicznymi nazwami, manipulacji ciągami, wartościami logicznymi i wiele innych.

W początkowym okresie technologia XSLT była dedykowana przetwarzaniu stron hipertekstowych WWW zapisanych w języku HTML (Holzner, 2002; Tennakoon, 2003). Język ten, a w szczególności jego nowsza wersja XHTML, jako język zapisu stron internetowych jest aplikacją XML i ma określony zbiór predefiniowanych znaczników („tagów”). Arkusze stylów dla języka HTML są określane akronimem CSS (*Cascading Style Sheets*) i ich zadaniem jest nadanie odpowiedniego wyglądu wyświetlanej w oknie przeglądarki stronie zapisanej w HTML. Wygląd ten może być różny w zależności od tego jaki arkusz zostanie użyty. Obecnie każda zaawansowana przeglądarka internetowa ma wbudowane narzędzia do transformacji XSLT. W takich zastosowaniach główny element dokumentu XLS ma nazwę *style sheet*, jednak specyfikacja tego języka dopuszcza równoważny synonim *transform*. W dalszej części opisu ta druga nazwa będzie stosowana, ponieważ znacznie bardziej odpowiada opisywanym zastosowaniom.

Zastosowanie technologii XSLT jest bardzo szerokie i znacznie wykracza poza zastosowania do transformacji danych geoprzestrzennych. Z tego względu mamy wielką liczbę różnych opracowań i w tym obszernych podręczników dotyczących albo zagadnień ogólnych albo zastosowań dziedzinowych niezwiązanych z geoinformacją. Głębsze poznanie tej problematyki wymaga przestudiowania wielu opracowań, a tu można przedstawić jedynie jej bardzo ogólny zarys. Do najważniejszych opracowań i podręczników można zaliczyć: *XSLT Cookbook* (Mangano, 2006), *XSLT Receptury* (Mangano, 2007), *XSLT Interview – Questions, Answers and Explanations* (Sanchez-Clark, 2007), *Beginning XSLT and XPath: Transforming XML Documents and Data* (Williams, 2009), *Learning XSLT* (Fitzgerald, 2004), *Sams – Teach Yourself XSLT in 21 Days* (Van Otegem, 2002), *Beginning XSLT 2.0: From Novice to Professional* (Tennison, 2005), *XSLT* (Tidwell, 2008) i *XSLT 2.0 and XPath 2.0 Programmer's Reference* (Kay, 2008). Te dziewięć pozycji literatury ma łączną objętość ponad 3700 stron.

9.1. Podstawy metodyczne transformacji

Dalsze prace nad rozwojem tej technologii prowadzone w Konsorcjum W3C doprowadziły do jej uogólnienia i rozszerzenia jej zastosowań do dowolnych zapisów aplikacyjnych języka XML. Ponieważ w języku XML nie są używane predefiniowane znaczniki („tagi”) tak jak HTML, każdy może definiować własne, co prowadzi do sytuacji, w której znaczenie

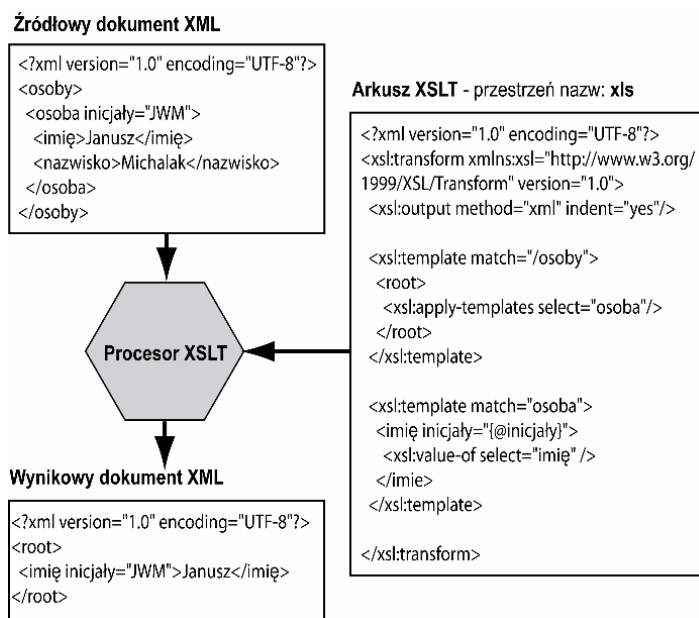


Rysunek 9.1. XSLT jako technologia posługuje się całą grupą standardów ze sobą powiązanych (W3Schools, 2000).

każdego znacznika nie musi być powszechnie znane. Na obecnym etapie rozwoju jako XLS rozumie się całą grupę standardów ściśle ze sobą powiązanych (rys. 9.1.). Podstawą dla pozostałych jest język XPath przeznaczony do nawigacji po dokumentach XML, XSLT służy do transformacji zapisów XML, XPointer (rozszerzenie języka XLink) pozwala na wiązanie różnych miejsc w różnych dokumentach, a XQuery jest językiem zapytań XML (w dokumentach i bazach) analogicznym do SQL w bazach relacyjnych.

9.1.1. Technologia XSLT

Jak było przedstawione wcześniej, XSLT służy do przekształcania (transformacji) struktury dokumentów XML z jednego schematu XSD do struktury określonej w innym schemacie XSD. Proces ten jest realizowany przez oprogramowanie narzędziowe nazywane procesorem XSLT, a przegląd tych narzędzi zawiera rozdział 9.2. Reguły tego przekształcania zapisane są w dokumencie XML nazywanym arkuszem XSLT z przestrzenią nazw `xsl`. W rozpatrywanych tu zastosowaniach element główny (*root*) ma nazwę `xsl:transform`, ale najczęściej w innych zastosowaniach stosuje się jego synonim `xsl:stylesheet`. Rysunek 9.2. przedstawia główne elementy uczestniczące w procesie transformacji: procesor XSLT, źródłowy dokument XML, arkusz XSLT kierujący procesem i dokument wynikowy, który powstaje w wyniku tego procesu.



Rysunek 9.2. Schemat przedstawiający cztery główne elementy uczestniczące w procesie transformacji opartej o technologię XSLT.

Zarówno dokument źródłowy, jak i arkusz XSLT są na wstępie przetwarzane przez procesor do formy drzewa złożonego z węzłów, którymi są poszczególne elementy tam występujące, a także ich atrybuty. Dokument wynikowy w trakcie przetwarzania też jest budowanym drzewem, które w końcowej fazie jest zapisywane do dokumentu XML.

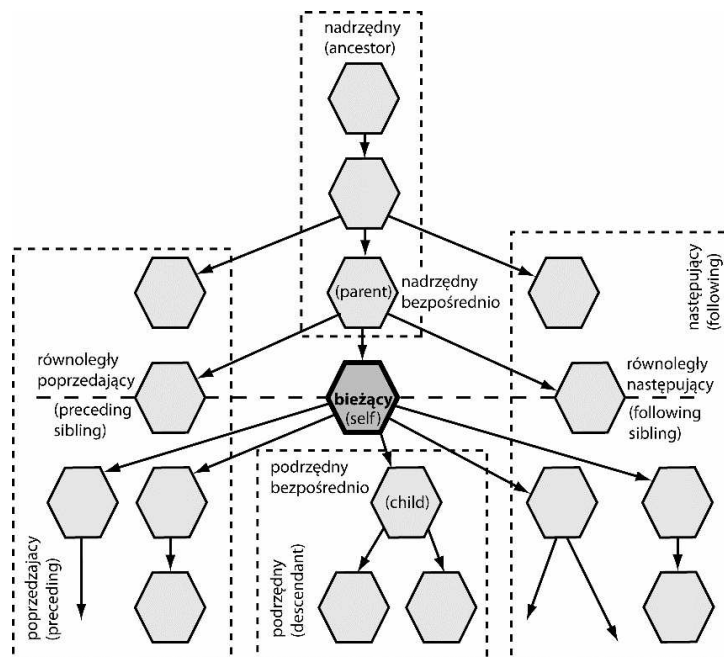
Najbardziej istotnym elementem jest XPath, który służy do nawigacji (wędrówki) po tych drzewach z węzła do węzła i przy jego pomocy ustalane są reguły budowy drzewa wynikowego. XPath jest opisany bardziej szczegółowo w następnym rozdziale (9.1.2.).

Procesor XSLT pracuje ze stałym, z góry określonym algorytmem, który składa się z szeregu etapów.

1. Wstępna analiza przed właściwą transformacją:
 - 1.1. Procesor XSLT przetwarza jeden lub więcej dokumentów źródłowych XML i jeden lub więcej arkuszy stylów XSLT i na ich podstawie będzie tworzył dokument wynikowy.
 - 1.2. Etap przeszukiwania i analizy wejściowego dokumentu XML i arkuszy XSLT.
 - 1.3. Po wczytaniu arkusza (arkuszy) stylów procesor tworzy drzewo źródłowe na podstawie dokumentu źródłowego.
 - 1.4. Jeden arkusz może odwoływać się do innych arkuszy i wszystkie odwołania w postaci *include* lub *import* muszą być uwzględnione.
 - 1.5. Z dokumentów usuwane są zbędne białe znaki (spacje, znaki tabulacji i tym podobne).
 - 1.6. Do drzewa arkusza XSLT dołączane są reguły, zarówno standardowe jak i zdefiniowane lokalnie.
2. Sam proces transformacji:
 - 2.1. Procesor przetwarza węzeł główny drzewa źródłowego, znajduje najlepiej pasujący szablon dla tego węzła w arkuszu stylów i analizuje zawartość tego szablonu.
 - 2.2. Dla drzewa wyjściowego powstaje element główny (*root node*).
 - 2.3. Zasadniczy etap transformacji to przetwarzanie elementów drzewa wejściowego.
 - 2.3.1. Podstawą tego przetwarzania jest poszukiwanie pasujących wzorców.
 - 2.3.2. Nawigacja po drzewie wejściowym odbywa się z zastosowaniem XPath.
 - 2.3.2.1. Dla danego węzła szukane są szablony spełniający określone warunki.
 - 2.3.2.2. Reguły przetwarzania zawarte w szablonach określają jak należy potraktować dany węzeł pasujący do wzorca ustalonego przez XPath.
 - 2.3.2.3. Spośród tych szablonów wybierany jest jeden o najwyższym priorytecie.
 - 2.3.3. Ten szablon, który ma najwyższy priorytet zostaje użyty w drzewie wynikowym.
 - 2.3.3.1. Instrukcje zawarte w szablonie najczęściej wskazują jak utworzyć nowy węzeł w drzewie wynikowym lub jak przetwarzać kolejne węzły w drzewie źródłowym.
 - 2.3.3.2. Te elementy szablonu, które mają prefiks *xsl* (należą do przestrzeni nazw XSLT) są instrukcjami określającymi działania na drzewie wynikowym.
 - 2.3.3.3. Pozostałe elementy są bezpośrednio zapisywane w drzewie wynikowym.
 - 2.3.4. Kolejny etap to przetwarzanie instrukcji *apply-templates* dotyczących węzłów bezpośrednio podrzędnych (typu *child*).
 - 2.4. Po ustaleniu reguł dla poszczególnych węzłów drzewa źródłowego, zgodnie z tymi regułami tworzy się węzły drzewa wynikowego.
3. Na zakończenie procesu przetwarzania drzewo wynikowe zostaje zapisane w dokumencie wynikowym.

9.1.2. Język XPath

Język XPath jest podstawowym komponentem w standardzie XSLT W3C, inne komponenty tego standardu – XQuery i XPointer wykorzystują XPath do realizowania swoich zadań (W3Schools, 2000). Język XPath służy do poruszania się po drzewie dokumentu XML w celu znalezienia odpowiednich poszukiwanych elementów i atrybutów w tym dokumencie (Almendros-Jimenez, Becerra-Terón, Garcia-Garcia, 2010).



Rysunek 9.3. Nazwy węzłów stosowane do określania ścieżek względnych (kontekstowych). W takim przypadku ścieżka do określonego węzła rozpoczyna się od węzła bieżącego (*self*).

Drzewo takie swoją strukturą bardzo przypomina strukturę systemu plików w komputerze. W systemie plików katalog główny zawiera wiele podkatalogów i może zawierać także poszczególne pliki. Katalogi mogą się dowolnie zagnieżdżać, a każdy z nich może zawierać katalogi niższego rzędu i równoległe obok nich poszczególne pliki. Po takiej strukturze można wędrować posługując się ścieżkami względnymi lub bezwzględnymi. Przykłady ścieżek katalogowych w systemach Unix:

- ścieżka bezwzględna rozpoczyna się od katalogu głównego (*root*) oznaczonego „/”, na przykład /opt/dane/rok2014/listopad/dnia07/pomiary.dat,
- jeżeli katalogiem bieżącym jest katalog /opt/dane/2013, to ścieżka względna do tego samego pliku ma postać ../rok2014/listopad/dnia07/pomiary.dat, gdzie „.” (kropka) określa katalog bieżący, a dwie kropki („..”) to katalog nadrzędny.

Podobne zasady obowiązują w nawigacji z zastosowaniem XPath, jednak w takich przypadkach najczęściej stosowane są ścieżki względne i określanie grup węzłów (elementów i atrybutów) odbywa się na takich samych zasadach. Rysunek 9.3. przedstawia zasady względnego adresowania węzłów otaczających węzeł bieżący (*self*). W różnych opracowaniach dotyczących względnego określenia położenia węzłów stosowane są różne nazwy. Aby móc porównać te określenia z różnych źródeł są one zestawione w tabeli 9.1.

Tabela 9.1. Zestawienie nazw określających względne położenie węzłów lub ich grup stosowanych w różnych opracowaniach

Polska terminologia topologiczna	Angielska terminologia genealogiczna	Polska terminologia genealogiczna
Nadrzędny	ancestor	przodek
Nadrzędny lub bieżący	ancestor-or-self	przodek lub bieżący
Nadrzędny bezpośrednio	parent	rodzic
Bieżący	current (self)	bieżący
Podrzędny bezpośrednio	child	dziecko
Podrzędny	descendant	potomek
Podrzędny lub bieżący	descendant-or-self	potomek lub bieżący
Poprzedzający	preceding	poprzedzający
Następujący	following	następujący
Równoległy	sibling	rodzeństwo
Równoległy poprzedzający	preceding sibling	rodzeństwo poprzedzające
Równoległy następujący	following sibling	rodzeństwo następujące

Podobnie jak XSLT, język XPath zawiera ponad 100 wbudowanych funkcji. Istnieją funkcje dla tekstów (ciągów znaków), wartości liczbowych, porównywania daty i czasu, manipulacji węzłami i kanonicznymi nazwami, manipulacji ciągami, wartościami logicznymi i wiele innych.

Zastosowanie języka XPath wykracza znacznie poza XSLT, jest on podstawą innych języków z tej rodziny jak XQuery 1.0 i XPath 2.0. Oba te języki mają ten sam model danych i obsługują te same funkcje i operatory. XPath może być także zastosowany w dowolnym innym oprogramowaniu narzędziowym przeznaczonym do dokumentów XML.

9.1.3. Języki XPointer i XLink

Język XLink (*XML Linking Language*) jest przeznaczony do posługiwania się w dokumencie XML konstrukcjami opisującymi powiązania między zasobami. Konstrukcje te używają składni XML do tworzenia struktur, które opisują proste lub rozszerzone jednokierunkowe hiperłącza podobne do tych, które są stosowane w HTML do nawigacji, jak również do tworzenia powiązań bardziej zaawansowanych. Język XLink jest koncepcją wcześniejszą i znacznie prostszą w porównaniu z XPointer. Za pomocą XLink można wiązać albo całe dokumenty albo ich fragmenty oznaczone etykietami.

Język XPointer (*XML Pointer Language*) jest rozszerzeniem języka XLink i na nim bazuje. Język XPointer jest przeznaczony do wskazywania dowolnego fragmentu dokumentu XML przez identyfikator URI, które lokalizuje zasób medialny w Internecie typu `text/xml` lub `application/xml`.

Język XPointer został podzielony na 4 moduły: pierwszy to schemat służący do określania lokalizacji, a trzy pozostałe schematy to: `element()`, `xmlns()` i `xpointer()`. Schemat `xpointer()`, który jest oparty na XPath (*XML Path Language*), jest wciąż w fazie rozwoju. Obsługuje on adresowanie do wewnętrznych struktur dokumentów XML. Pozwala na wędrówkę

po drzewie i wybór jego wewnętrznych elementów w oparciu o różne właściwości, takie jak typy elementów dokumentu, wartości atrybutów, zawartości tekstów, a także wzajemne położenie atrybutów. Język ten nie jest używany w przeglądarkach internetowych, tak jak XSLT i XPath, ale jest szeroko stosowany w innych aplikacjach XML (językach dziedzinowych opartych na XML). Proste przykłady odsyłaczy z zastosowaniem XPointer i identyfikatorów przedstawione są poniżej (9.1., 9.2., 9.3.). W zapisach języka GML elementy opisujące wyróżnienia («FeatureType») mają identyfikatory (gml:id). Podobne identyfikatory są też zadeklarowane w schematach aplikacyjnych INSPIRE i one też mogą być wykorzystane do realizacji odwołań opartych na XPointer.

Przykład 9.1. Fragment zapisu danych dla JCWPd (Jednolitej Części Wód Podziemnych) nr 68 w pliku o adresie <http://geomatyka.pl/gml/jcwpds.gml>.

```
<gml:featureMember>
  <hg:jcwpd gml:id="JCWPd_68">
    <hg:nr_JCWPd>68</hg:nr_JCWPd>
    <hg:powierzchnia_km2>138</hg:powierzchnia_km2>
    <hg:dorzecze>Wisła</hg:dorzecze>
    <hg:kod_UE>PLGW200068</hg:kod_UE>
    <gml:surfaceProperty>
      <gml:Surface gml:id="JCWPd_68_Surface"
        srsName="EPSG:2180" srsDimension="2">
        (...)
      </gml:Surface>
    </gml:surfaceProperty>
  </hg:jcwpd>
</gml:featureMember>
```

Odwołanie zewnętrzne (z dokumentu zewnętrznego znajdującego się gdzieś w Internecie) do elementu hg:jcwpd opisującego JCWPd nr 68 ma postać przedstawioną w przykładzie 9.2., a wewnętrzne (z tego samego dokumentu) w przykładzie 9.3. Przykłady te zawierają także odwołania do elementu gml:Surface określającego zasięg przestrzenny tego JCWPd. Schematyczne przedstawienie URI z użyciem odsyłacza wykorzystującego język XPointer dla tych przykładów jest przedstawione na rysunku 9.4.

Przykład 9.2. Odwołanie zewnętrzne za pomocą XPointer do elementu zapisu danych dla JCWPd nr 68 i do elementu opisu zasięgu JCWPd nr 68.

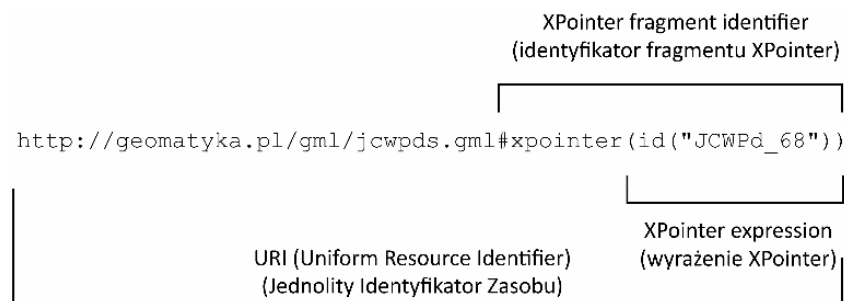
```
Odwołanie zewnętrzne do JCWPd nr 68:
http://geomatyka.pl/gml/jcwpds.gml#xpointer(id("JCWPd_68"))

Odwołanie zewnętrzne do zasięgu JCWPd nr 68:
http://geomatyka.pl/gml/jcwpds.gml#xpointer(id("JCWPd_68_Surface"))
```

Przykład 9.3. Odwołanie wewnętrzne za pomocą XPointer do elementu zapisu danych dla JCWPd nr 68 i do elementu opisu zasięgu JCWPd nr 68.

```
Odwołanie wewnętrzne do JCWPd nr 68:
xpointer(id("JCWPd_68"))

Odwołanie wewnętrzne do zasięgu JCWPd nr 68:
xpointer(id("JCWPd_68_Surface"))
```



Rysunek 9.4. Schematyczne przedstawienie URI (*Uniform Resource Identifier* – Jednolity Identyfikator Zasobu) z użyciem odsyłacza wykorzystującego język XPointer.

9.1.4. Język XQuery

Kolejnym językiem z rodziny XSL jest język XQuery przeznaczony do wyszukiwania danych w zapisach XML w formie zapytań analogicznych do języka SQL i dotyczy to nie tylko prostych dokumentów ale wszelkich zapisów w XML, w tym także baz danych. XQuery 1.0 i XPath 2.0 wykorzystują wspólnie ten sam model danych i obsługują te same funkcje i operatory. XQuery może być użyty do wydobywania informacji z dokumentów XML, może być wykorzystany w usługach sieciowych (w tym INSPIRE), wspomagać generowanie raportów, przekształcanie danych XML do XHTML, przeszukiwanie dokumentów w sieci dla pozyskania istotnych informacji i ma dużo innych zastosowań. XQuery jest kompatybilny z wieloma standardami W3C, takich jak: XML, XSLT, Namespaces, XPath i XML Schema. Przykład (9.4.) przedstawia proste przypadki użycia tego języka.

Przykład 9.4. Proste przypadki użycia języka XQuery do pliku, którego fragment jest przedstawiony w przykładzie 9.1.

```
Polecenie otwarcia pliku do przeszukiwania i pobierania danych:
doc("jcwps.gml")

Wybranie wszystkich elementów typu: hg:nr_JCWPD
doc("books.xml")/
  gml:featureCollection/gml:featureMember/hg:jcwpd/hg:nr_JCWPD

W wybranych elementach znajdzie się także ten z przykładu 9.1.:
<hg:nr_JCWpd>68</hg:nr_JCWpd>

Wybranie elementów typu hg:jcwpd których powierzchnia jest większa niż 100 km2
doc("jcwps.gml")/gml:featureCollection/gml:featureMember/hg:jcwpd/
  [hg:powierzchnia_km2>100]

W wybranych elementach znajdzie się także ten z przykładu 9.1., ale nie znajdują się te, których
powierzchnia jest mniejsza lub równa 100 km2:
<hg:jcwpd gml:id="JCWPd_68">
  (...)
<hg:powierzchnia_km2>138</hg:powierzchnia_km2>
  (...)
</hg:jcwpd>
```

Język XQuery jest jednym z podstawowych modułów baz danych opartych na języku XML. Do takich baz należy także opisana w rozdziale 10. baza BaseX i tam zapytania XQuery są opisane bardziej szczegółowo, także z rozszerzeniami dotyczącymi zapytań geoprzestrzennych.

9.2. Procesory XSLT

Jak wspomniano wcześniej, wszystkie zaawansowane przeglądarki internetowe mają wbudowane procesory XSLT, jednak ich rola jest ściśle ograniczona do przekształcania dokumentów typu HTML (w tym XHTML – aplikacji XML) w celu wyświetlenia ich treści w oknie przeglądarki na ekranie komputera. W zastosowaniach tu opisywanych wykorzystywane są procesory samodzielne, chociaż często takie samodzielne narzędzia są wbudowywane w bardziej złożone systemy modułowe. Najbardziej popularne samodzielne procesory przedstawiono poniżej.

- Procesory funkcjonujące w środowisku MS Windows. Należą do nich MSXML, .NET.MSXML i XslCompiledTransform, a także uruchamiany z linii komend msxml.exe. Ze względu na ograniczenie ich zastosowania tylko do tego środowiska systemowego nie będą tu omawiane.
- Procesor Saxon jest systemem dostępnym zarówno w wersji *Open Source*, jak i w profesjonalnej wersji komercyjnej. Ma interfejsy dla środowisk: Java, .NET i JavaScript, a także może pracować w pełni samodzielnie. Obsługuje standardy: XSLT 2.0 i częściowo 3.0 i XQuery 3.0.
- Procesory firmy Altova, w tym AltovaXML 2010 – dostępny bezpłatnie silnik XSLT 1.0/2.0, pracujący niezależnie jako uruchamiany z linii komend lub przez interfejsy COM, Java i .NET. Inny procesor tej firmy to Altova RaptorXML Server – wielopłatformowy silnik obsługujący XSLT 1.0 i 2.0, w większości XPath 3.0, a także częściowo XSLT 3.0 w wersji roboczej, także obsługuje XQuery. Ma interfejsy do COM, Java i .NET. Może także pracować samodzielnie z linii komend i ponadto ma własny wbudowany serwer HTTP.
- TransformiX jest procesorem *Open Source* rozwijanym przez społeczność Mozilla, dedykowanym do pracy w przeglądarce Firefox i innych systemach rozwijanych przez tę społeczność. Na podstawie kodu źródłowego i dostępnych skryptów można zbudować wersję działającą samodzielnie na różnych platformach systemowych.
- Xalan to kolejny procesor *Open Source* rozwijany przez niekomercyjne konsorcjum Apache Software Foundation. Procesor ten obsługuje XSLT 1.0 i XPath 1.0, a jego interfejsy to Java i C++, a także Perl. Xalan jest częścią rozbudowanego pakietu o nazwie Xerces, który zawiera także inne moduły dedykowane zapisom XML. Xerces to zbiór modułów i bibliotek przeznaczonych do analizowania, weryfikacji, serializowania zapisów XML i manipulowania dokumentami XML, umożliwia on zarówno analizę jak i tworzenie dokumentów XML. Jest to obecnie najbardziej zaawansowany *Open Source* system dedykowany językowi XML.
- Ostatnią pozycją na tej liście jest biblioteka w języku C rozwijana w ramach projektu GNOME o nazwie libxslt. Biblioteka bazuje na popularnej w systemie Linux bibliotece libxml i obsługuje XSLT 1.0. Jest to także oprogramowanie *Open Source*, które może pracować w środowiskach systemowych Linux, Unix i Windows (poprzez Cygwin), a także może być uruchamiane z linii komend poprzez xsltproc. Procesor ten posiada interfejsy do języków: Python, Perl, Ruby, PHP, Common Lisp, Tcl i C++.

Na podstawie analizy budowy systemów programowych stosowanych do transformacji danych geoprzestrzennych można uznać, że najczęściej stosowanym do tego celu jest procesor Xalan, także wraz z innymi modułami pakietu Xerces. Jest on wykorzystywany do transformacji w systemie FME opisanym w rozdziale 11.1., w HALE będącym aplikacją zbudowaną na bazie Eclipse (rozdział 11.2), a także w oprogramowaniu bazy danych BaseX (rozdział 10.1.). Odstępstwem od tej reguły jest zestaw narzędzi ETL pod nazwą Stetl (rozdział 11.4.), w którym do transformacji używane są komponenty biblioteki libxslt i libxml2. System zarządzania bazą danych XML BazeX, obok procesora Xalan, może także wykorzystywać procesor Saxon. Opisane w rozdziale 11.3. oprogramowanie narzędziowe do transformacji danych geoprzestrzennych GeoKettle może wykorzystywać oba te procesory – Xalan i Saxon.

9.3. Składnia i semantyka arkuszy XSLT

Arkusze XSLT (lub krócej XSL) służą do sterowania procesem transformacji. Są to dokumenty XML zawierające elementy o nazwach określonych przestrzenią nazw xsl. Najczęściej używanymi elementami są:

- `xsl:stylesheet` lub `xsl:transform` (jest to bardziej odpowiedni w tym przypadku jego równoważnik) – jest to element główny (*root*) zawierający wszystkie inne elementy jakie występują w tym arkuszu,
- `xsl:output` – element definiujący format dokumentu wynikowego,
- `xsl:strip-space` – element definiujący elementy dokumentu źródłowego, w których mają być usunięte niepotrzebne białe znaki (spacje, znaki tabulacji i inne),
- `xsl:template` – element definiujący szablon, czyli zestaw reguł, które mają być zastosowane gdy określony węzeł drzewa źródłowego zostanie znaleziony.

Element `xsl:template` definiujący szablon jest elementem złożonym i w jego obrębie mogą wystąpić inne elementy. Najczęściej są to:

- `xsl:if` – element używany do wykonania testu warunkowego dla fragmentu zapisu zawartego w pliku XML,
- `xsl:call-template` – element ten wywołuje inny nazwany szablon po jego nazwie,
- `xsl:with-param` – element definiuje wartość parametru, która ma być przekazana do innego szablonu.

Poza wymienionymi powyżej, jest wiele innych, rzadziej używanych elementów, opisanych w literaturze wymienionej na początku rozdziału 9. lub w elektronicznych dokumentach konsorcjum W3C na temat reguł XSLT (W3School, 2000): <http://www.w3schools.com/xsl/>. Dla zilustrowania, jak te elementy są używane w dokumencie XSL jest przedstawiony bardzo prosty przykład (9.5.) oparty na materiałach opublikowanych przez Justa van den Broeckera i przystosowanych do krajowych (polskich) danych hydrogeologicznych. Transformowane dane zawierają pomiary zwierciadła wody w kilku otworach.

Przykład 9.5. Prosty zapis arkusza XSL ilustrujący użycie elementów schematu XSL definiujących przebieg procesu transformacji.

```

<?xml version="1.0" encoding="UTF-8"?>
<!--
  Znacznie uproszczony przykład arkusza opracowany na podstawie
  przykładu Just van den Broecka
  Transformowane będą dane z pomiarów zwierciadła wody podziemnej
  Przestrzeń nazw danych źródłowych: hge
  Przestrzeń nazw danych wynikowych: gmlhge -->
<!-- Początek elementu root arkusza -->
<xsl:transform version="1.0"
  xmlns:hge="http://www.laboratorium-geomatyki/hge"
  xmlns:gmlhge="http://www.laboratorium-geomatyki/gmlhge"
  xmlns:gml="http://www.opengis.net/gml"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<!-- Deklaracja formatu dokumentu wynikowego -->
  <xsl:output method="xml" omit-xml-declaration="no" indent="yes"/>
<!-- Polecenie usunięcia niepotrzebnych białych znaków -->
  <xsl:strip-space elements="+"/>
<!-- Szablon dla węzła głównego (root) oznaczonego jako "/" -->
  <xsl:template match="/">
<!-- Poniższe wpisać jako element (węzeł) root drzewa wynikowego -->
    <hge:FeatureCollection
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xmlns:gmlge="http://www.laboratorium-geomatyki/gmlhge"
      xmlns:gml="http://www.opengis.net/gml"
      xsi:schemaLocation="
        http://www.laboratorium-geomatyki/gmlhge
        ./gmlhge.xsd <!-- Lokalna kopia schematu wynikowego -->
        http://www.opengis.net/gml
        http://schemas.opengis.net/gml/3.2.1/gml.xsd"
      gml:id="PL-784-2015-45">
      <gml:boundedBy>
<!-- Układ odniesienia przestrzennego EPSG:2180 = PUWG 1992 -->
        <gml:Envelope srsName="EPSG:2180" srsDimension="2">
          <gml:lowerCorner>442550.7 583254.7</gml:lowerCorner>
          <gml:upperCorner>538542.1 704229.6</gml:upperCorner>
        </gml:Envelope>
      </gml:boundedBy>
<!-- Dla wszystkich węzłów podrzędnych (oś descendant) drzewa źródłowego
      zastosować szablony w odpowiednich węzłach drzewa wynikowego -->
        <xsl:apply-templates/>
      </hge:FeatureCollection>
    </xsl:template>
<!-- Koniec szablonu dla węzła głównego (root) -->
    <!-- Każdy pomiar zwierciadła ("hge:zwierciadło")
      ma być elementem typu gml:featureMember -->
    <xsl:template match="hge:zwierciadło">
<!-- Poniższe wpisać jako węzeł drzewa wynikowego odpowiadający
      węzłowi "hge:zwierciadło" w drzewie źródłowym -->
      <gml:featureMember>
        <gmlhge:zwierciadło>
          <gmlhge:pomiar>
            <xsl:value-of select="pomiar"/>
          </gmlhge:pomiar>
          <gmlhge:położenie>
            <gml:Point srsName="urn:ogc:def:crs:EPSG:2180">
              <gml:coordinates>
                <xsl:value-of select="x"/>,
                <xsl:value-of select="y"/>
              </gml:coordinates>
            </gml:Point>
          </hge:położenie>
        </hge:zwierciadło>
      </gml:featureMember>
    </xsl:template>
<!-- Koniec szablonu dla każdego pomiaru zwierciadła -->
  </xsl:transform>
<!-- Koniec elementu root arkusza (koniec arkusza)-->

```

10. Bazy danych XML

W procesie transformacji danych do postaci wymaganej przez dokumenty INSPIRE powstają zapisy nowych danych w tekstowym formacie znacznikowym XML. Również zapisy danych w formie przejściowej, w tym dane przygotowane do transformacji, mają formę XML. Prowadzenie tych prac w trybie operacyjnym (nie eksperymentalnym), dla dużych i licznych zbiorów danych, prowadzi do okazałego zasobu zapisów XML, który trzeba przechowywać i którym trzeba zarządzać. Racjonalnym rozwiązaniem jest zastosowanie w takich przypadkach systemu zarządzania bazą danych, lecz powstaje pytanie jakiego rodzaju baza danych będzie tu najbardziej odpowiednia – relacyjna, obiektowa czy baza XML? (Zhang, Peng, Li, Day, 2003; Huang, Chuang, Deng, Lee, 2009). Od wielu lat utrzymuje się przekonanie, że zadanie przechowywania danych powinno być realizowane przy użyciu relacyjnych baz danych, ponieważ ich wydajność i stabilność są bez porównania wyższe niż w przypadku innych rozwiązań (Connolly, Begg, 2004). Takie stanowisko było prawdziwe wiele lat temu, gdy bazy obiektowe nie były jeszcze dostatecznie dojrzałe, a bazy XML jeszcze nie istniały. Obecnie sytuacja jest inna, bo jest do dyspozycji dużo wydajnych i bezpiecznych baz obiektowych, a także są już dostatecznie wydajne i bezpieczne bazy XML (Graves, 2002; Lekkerkerk, 2012; Newell, Loren, Hopkins, Bedient, 1990; Wojda, 2008).

Wydajność tradycyjnych baz relacyjnych bierze się z zastosowanych w nich rozwiązań technicznych jako bufora w bardzo szybkiej pamięci operacyjnej, ale w pierwszej kolejności z bardzo prostego modelu logicznego struktury danych. Są to regularne prostokątne tabele powiązane prostymi relacjami, gdzie podstawowym elementem logicznym jest rekord odpowiadający jednemu wierszowi takiej tabeli. Jest to płaski model danych bez możliwości zapisywania danych o strukturach wielostopniowo zagnieżdżonych, jakie występują w zapisach XML. Konsekwencją tego, przy zapisie danych XML w bazie relacyjnej, jest konieczność całkowitej zmiany struktury danych, co jest operacją pracochłonną i często zmuszającą do rezygnacji z części informacji.

Wady modelu relacyjnego nie występują w bazach obiektowych, w których obiekty mogą tworzyć zagnieżdżone struktury, a model danych XML może być stosunkowo łatwo odwzorowany bez strat. Jednak popularność baz obiektowych jest obecnie nadal stosunkowo niewielka. Wynika to z nieznaności ich zalet, czyli z braku wiedzy, ale także z faktu, że większość zbiorów danych ma stosunkowo prostą budowę, dającą się sprowadzić do regularnych tabel bazy relacyjnej. Jednak w przypadku baz obiektowych również wymagana jest zmiana formy zapisu i dane tracą swoją formę znacznikową, czego nie ma w przypadku baz XML.

Język XML i jego aplikacje skutecznie podbijają współczesny świat gwałtownie rozwijających się technologii teleinformatycznych. Początkowo jego rolę widziano w wymianie informacji pomiędzy systemami informatycznymi, jednak w ostatnich latach stał się wszechobecny – od zapisów konfiguracji i sterowania systemami po zapisy grafiki komputerowej. Przykładem może być jego zastosowanie do własnych potrzeb – zastosowanie XML do XML. Aby zapisać jakieś dane w pliku XML potrzebne jest zdefiniowanie struktury tego zapisu. Do tego służą schematy XML (XSD) opracowywane dla różnych dziedzinowych zastosowań. Te schematy obecnie również są zapisane w XML. Drugi przykład to opisywa-

na tu transformacja XSLT – arkusz XSL, który decyduje jak ta transformacja ma być wykonana, jest również zapisany w formacie XML. Rozwój zastosowań języka GML do różnorodnych danych geoprzestrzennych sprawia, że niedługo potężne zasoby kartograficzne będą miały prawie wyłącznie formę zapisów XML, a ściślej jego aplikacji dziedzinowej GML. Jeżeli przyjmie się, że taki zapis służy wyłącznie do przesyłania danych, to jednak aby móc te dane przesłać trzeba tymi danymi dysponować. Jeżeli się takimi danymi dysponuje to trzeba je w jakiejś formie przechowywać. Coraz częściej podkreśla się, że bazy XML stają się niezbędne i w ostatnich latach trwają intensywne prace, głównie w konsorcjum W3C nad technologiami z zakresu XML, które są stosowane w bazach tego typu, w tym nad opisanymi wcześniej językami XPath, XPointer i XQuery – większość z nich jest już rekomendowanymi standardami międzynarodowymi.

Wszystkie przedstawione powyżej działania i ich rezultaty doprowadziły do sytuacji, w której mogą powstać i powstają zaawansowane systemy zarządzania bazami XML. Wydajność i bezpieczeństwo stopniowo przestaje być problemem. Wykorzystanie bardzo szybkich pamięci masowych typu SSD (*Solid-State Drive*), brak konieczności przetwarzania danych na wejściu i na wyjściu dla zmiany modelu i skuteczne metody indeksacji sprawiają, że wydajność tych baz jest w większości przypadków całkowicie wystarczająca. Problemy bezpieczeństwa danych w niewielkim stopniu zależą od typu systemu bazy, ponieważ za to jest głównie odpowiedzialny system operacyjny i inne komponenty teleinformatyczne. O ile rozwój baz relacyjnych jest już w zasadzie procesem zamkniętym, to rozwój baz XML dopiero się niedawno rozpoczął i trudno jest obecnie przewidzieć, jakie rozwiązania przyniesie przyszłość.

Analizując wymagania stawiane bazom danych XML dla potrzeb procesu transformacji i przechowywania wyników tej transformacji rozpatrzone zostały następujące cztery systemy:

- **Sedna** (Rekouts, 2004; Sedna Team, 2011a; 2011b; 2011c) jest systemem bazy danych rodzimym dla XML dostępnym na prawach *Open Source*. Zapewnia pełen zakres podstawowych usług wymaganych dla baz danych: trwałe składowanie danych, transakcje ACID (*Atomicity, Consistency, Isolation, Durability* – atomowość, spójność, izolacja, trwałość), bezpieczeństwo, indeksowanie na bazie B-tree, Hot Backup, obsługa Unicode (utf-8). System Sedna jest opracowany w językach C i C++. Zapewnia elastyczne przetwarzanie danych XML obejmujące zastosowanie języka XQuery, ścisłą integrację XQuery z możliwością pełnotekstowego (*Full Text*) wyszukiwania, przetwarzania i aktualizacji na poziomie poszczególnych węzłów drzewa. System ten zapewnia bezpieczeństwo bazy danych z uwzględnieniem kategorii: użytkownicy, role i uprawnienia. Posiada następujące programistyczne interfejsy aplikacyjne: połączenie SQL z XQuery, funkcje zewnętrzne realizowane w C dla XQuery, API dla Java, C, PHP, Python, Ruby, Perl, Delphi i C#. Ma także moduł integracji z serwerem Apache.
- **eXist** jest systemem *Open Source* zarządzania bazą danych NoSQL i platformą aplikacyjną zbudowaną w technologii XML (eXist Solutions, 2014). System napisany w języku Java jest klasyfikowany jako baza NoSQL dla dokumentów, a także jako baza danych rodzima dla XML. Zapewnia wsparcie dla XML, JSON, HTML i dokumentów binarnych. W przeciwieństwie do większości systemów relacyjnych i systemów baz danych NoSQL, eXist zapewnia pełne wsparcie dla XQuery i XSLT. Zaimplementowane podstawowe standardy to: XPath, XQuery, XSLT, XSL-FO (formatowanie obiektów XSL), WebDAV (*Web Distributed Authoring and Versioning*), REST (*Representational State Transfer*), RESTXQ (adnotacje REST dla XQuery), SOAP

(*Simple Object Access Protocol*), XACML (język kontroli dostępu XML), XML-RPC (protokół zdalnego wywoływania procedur dla XML), a także XQuery API dla języka Java. eXist pozwala twórcom oprogramowania na tworzenie procedur trwałego przechowywania dokumentów XML, JSON i binarnych bez pisania obszernych kodów pośrednich. eXist automatycznie indeksuje dokumenty za pomocą systemu indeksowania słów kluczowych.

- **BaseX** – system ten ze względu na szczególną właściwość jaką jest moduł funkcji, operacji i zapytań geoprzestrzennych został opisany oddzielnie w następnym rozdziale (10.1.).
- **MarkLogic** jest systemem komercyjnym i cena jego używania przez rok w najprostszej wersji wynosi około 8,5 tys. dolarów. Jest to zaawansowany system typu NoSQL (nie bazujący wyłącznie na zapytaniach SQL stosowanych w bazach relacyjnych), ale ze względu na cenę jego zastosowanie może być uzasadnione tylko w przypadkach wielkich zbiorów danych XML o krytycznym znaczeniu. Porównanie podstawowych właściwości tego systemu z systemem BaseX zawarte jest w tabeli 10.1.

Tabela 10.1. Porównanie głównych cech komercyjnej bazy MarkLogic i bazy BaseX typu *Open Source*

	MarkLogic	BaseX
Licencja	Komercyjna	<i>Open Source</i>
Język	C++	Java
System operacyjny	Linux, OS X, Solaris, Windows	Linux, OS X, Windows
Rodzaj bazy	Enterprise NoSQL DBMS dla JSON, XML i RDF	Lekki rodzimy system dla XML ze wsparciem XQuery 3.0 i interakcyjnym GUI.
Model bazy	Repozytorium dokumentów, rodzimy system dla XML, repozytorium RDF, wyszukiwarka	Rodzimy system dla XML
Programistyczne interfejsy aplikacyjne	Java API, Node.js API, ODBC, RESTful HTTP API, WebDAV, XDBC, XQuery, XSLT	Java API, RESTful, HTTP API, RESTXQ, WebDAV, XML.DB, XQJ
Wspierane języki programowania	C, C#, C++, Java, JavaScript (Node.js), Perl, PHP, Python, Ru	Actionscript, C, C#, Haskell, Java, JavaScript, Lisp, Perl, PHP, Python, Qt, Rebol, Ruby, Scala, Visual Basic

10.1. BaseX

BaseX jest otwartym systemem zarządzania zbiorami (bazami) danych XML na licencji otwartej BSD (*Open Source*) (BaseX Team, 2014). Interfejs graficzny tego systemu jest przedstawiony na rysunkach 10.1., 10.2., 10.3. i 10.4. Ma on zaimplementowanych wiele technologii opartych na standardach W3C, dotyczących przetwarzania dokumentów XML, w tym języki: XQuery, XSLT, XPath i XLink, a także rozszerzenie *Full Text* (Mahlow, Grün, Holupireky, Scholly, 2012) pozwalające na przetwarzanie ciągów znaków stanowiących teksty zawarte w elementach lub ich atrybutach. Dzięki tym rozwiązaniom nie tylko przechowuje i zarządza danymi XML, ale także je skutecznie przetwarza. Również w zakresie operacji geoprzestrzennych. Z tego względu zasługuje na szczególną uwagę, ponieważ operacje te

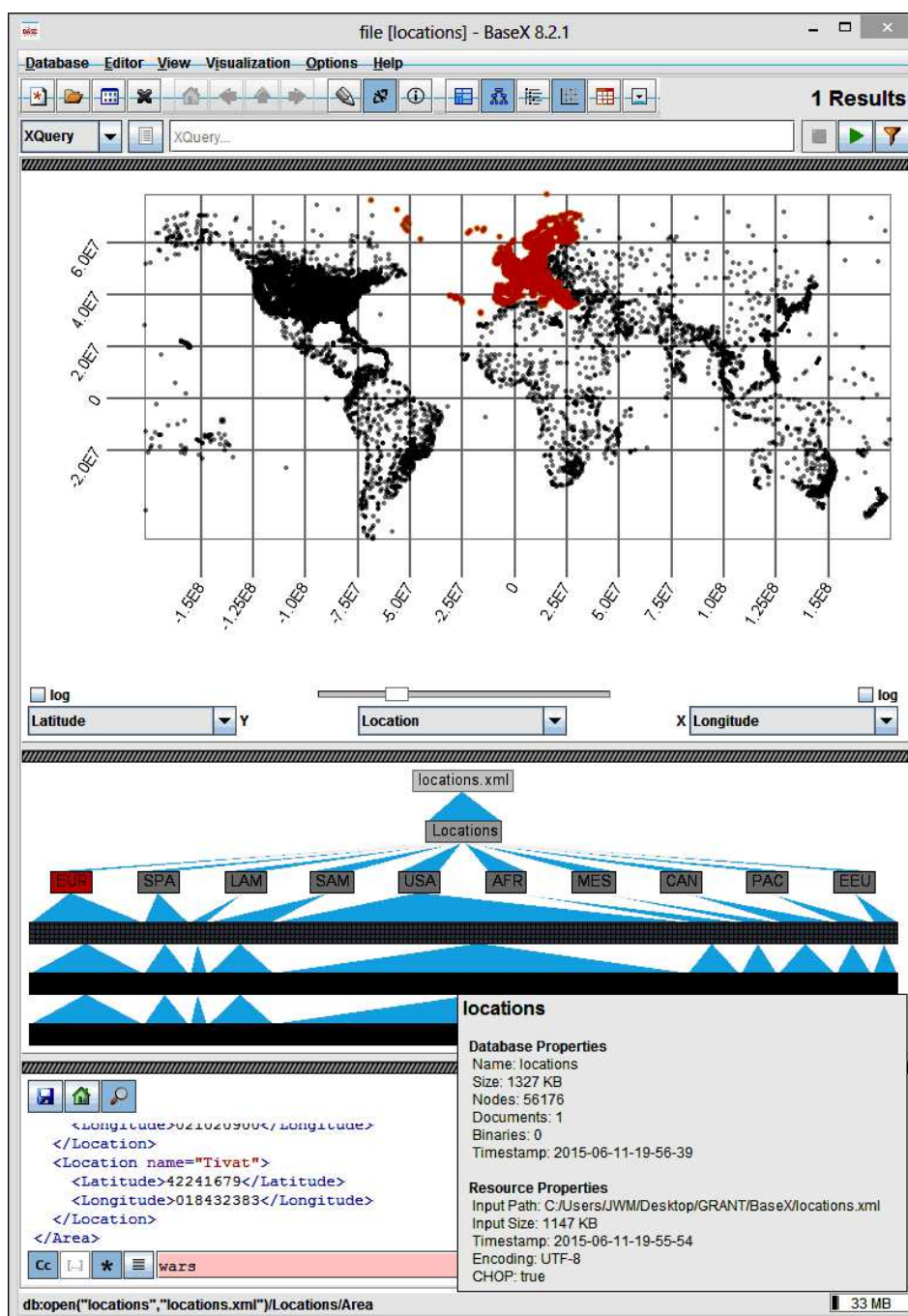
są dedykowane zapisom danych przestrzennych w języku GML (Seydi-Gheranghiyeh, 2013). W dalszej części tego rozdziału przedstawiona jest pełna lista funkcji geoprzestrzennych, które mogą być wykorzystane w zapytaniach XQuery lub w przetwarzaniu zapisów XML.

Jednym z bardziej interesujących rozwiązań zastosowanych w BaseX jest koncepcja DeepFS dotycząca rozmieszczenia danych w systemie plików (Holupirek, Grün, Scholl, 2008; 2009). Samo przechowywanie danych w zaawansowanych technologicznie systemach plików współczesnych systemów operacyjnych nie nastręcza istotnych problemów. Jednak wygodny do nich dostęp i możliwość wyszukiwania informacji ze zgromadzonych tam danych jest często istotnym problemem. Stosowanie języka zapytań bazy danych zgodnego z jej charakterem dzięki koncepcji DeepFS może znacznie polepszyć efektywność tej bazy. Traktując tę strukturę systemu plików jako warstwę przechowywania danych dla procesora XQuery bazy danych BaseX, pomimo że podstawą jej jest typowy system plików, przeważającą część jego zawartości można przeszukiwać za pomocą XPath i XQuery.

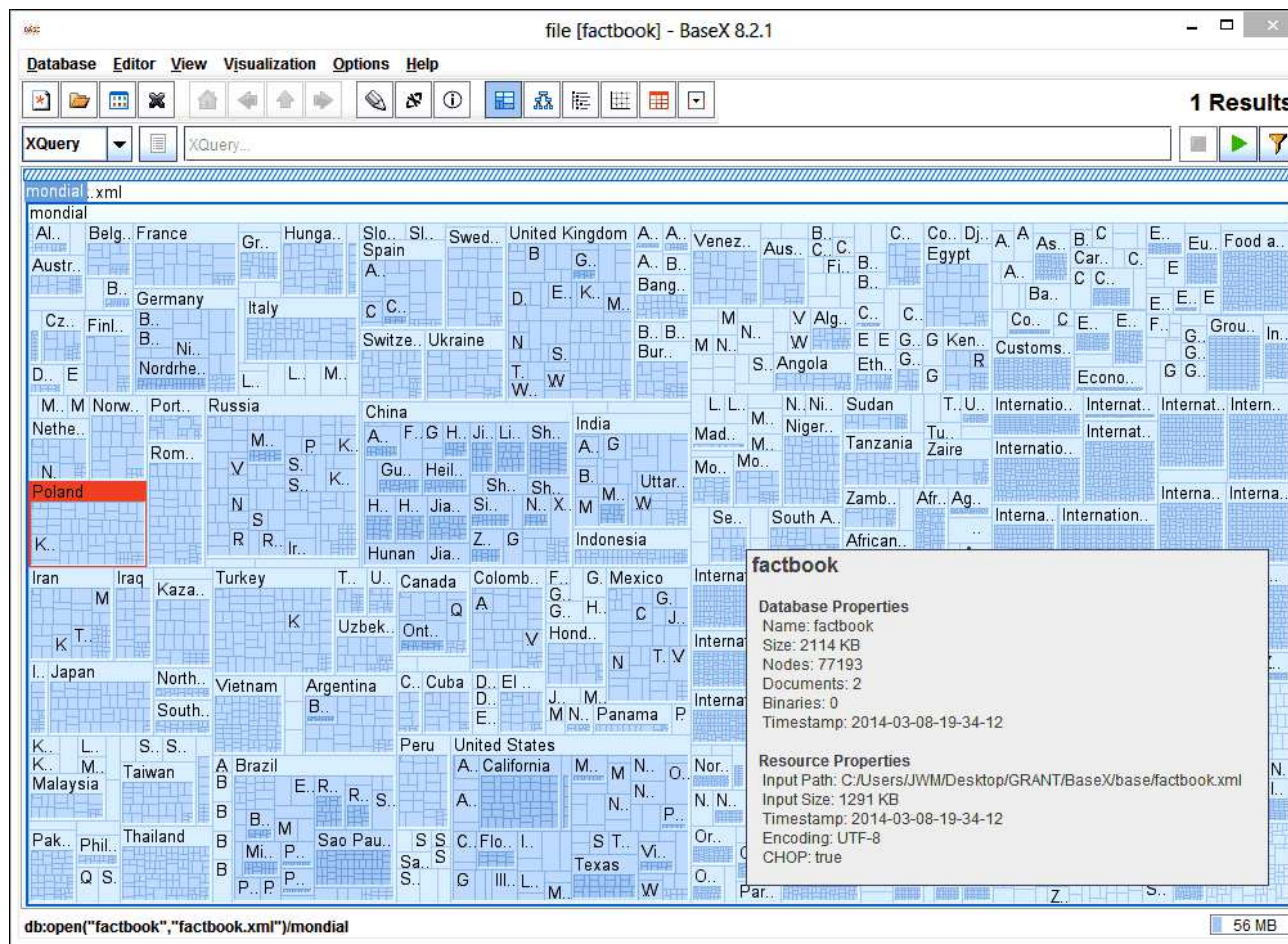
Oprócz standardowych funkcji XQuery, system BaseX oferuje jako rozszerzenia dużo dodatkowych modułów, zawierających funkcje dedykowane specyficznym dziedzinowym zastosowaniom, wśród których jest także moduł Geo (Seydi-Gheranghiyeh, 2013). Moduł ten zawiera rozszerzone funkcje XQuery, które mogą być stosowane do geometrii danych zgodnych ze specyfikacją OGC (*Open Geospatial Consortium*) dotyczącą profilu SF (*Simple Feature*) (Chen, He, Wang, 2010). Jest oparty na module EXPath Geo i wykorzystuje bibliotekę JST (*Java Topology Suite*). Moduł EXPath Geo definiuje funkcje, które mogą być stosowane w odniesieniu do poprawnie określonych elementów geometrycznych zgodnych ze specyfikacją Simple Feature OGC – są w tym przypadku zapisywane za pomocą języka GML (Córcoles, González, 2003; Li, 2006). Geometrie przyjęte profilu SF pochodzą z języka GML w wersji 2 i są to: Point (punkt), LineString (krzywa), LinearRing (pierścień), Polygon (wielokąt), MultiPoint (grupa punktów), MultiLineString (grupa krzywych), MultiPolygon (grupa wielokątów) i MultiGeometry (grupa wcześniej wymienionych geometrii). Wszystkie elementy geometryczne przeszukiwane zapytaniami w BaseX powinny być formalnie poprawne. Jedynym typem geometrii, która nie jest teraz obsługiwana przez BaseX jest MultiGeometry. Funkcje modułu Geo podzielone są na cztery kategorie: funkcje ogólne, funkcje relacji przestrzennych, funkcje analityczne i funkcje specyficzna dla określonych typów geometrii.

Funkcje ogólne:

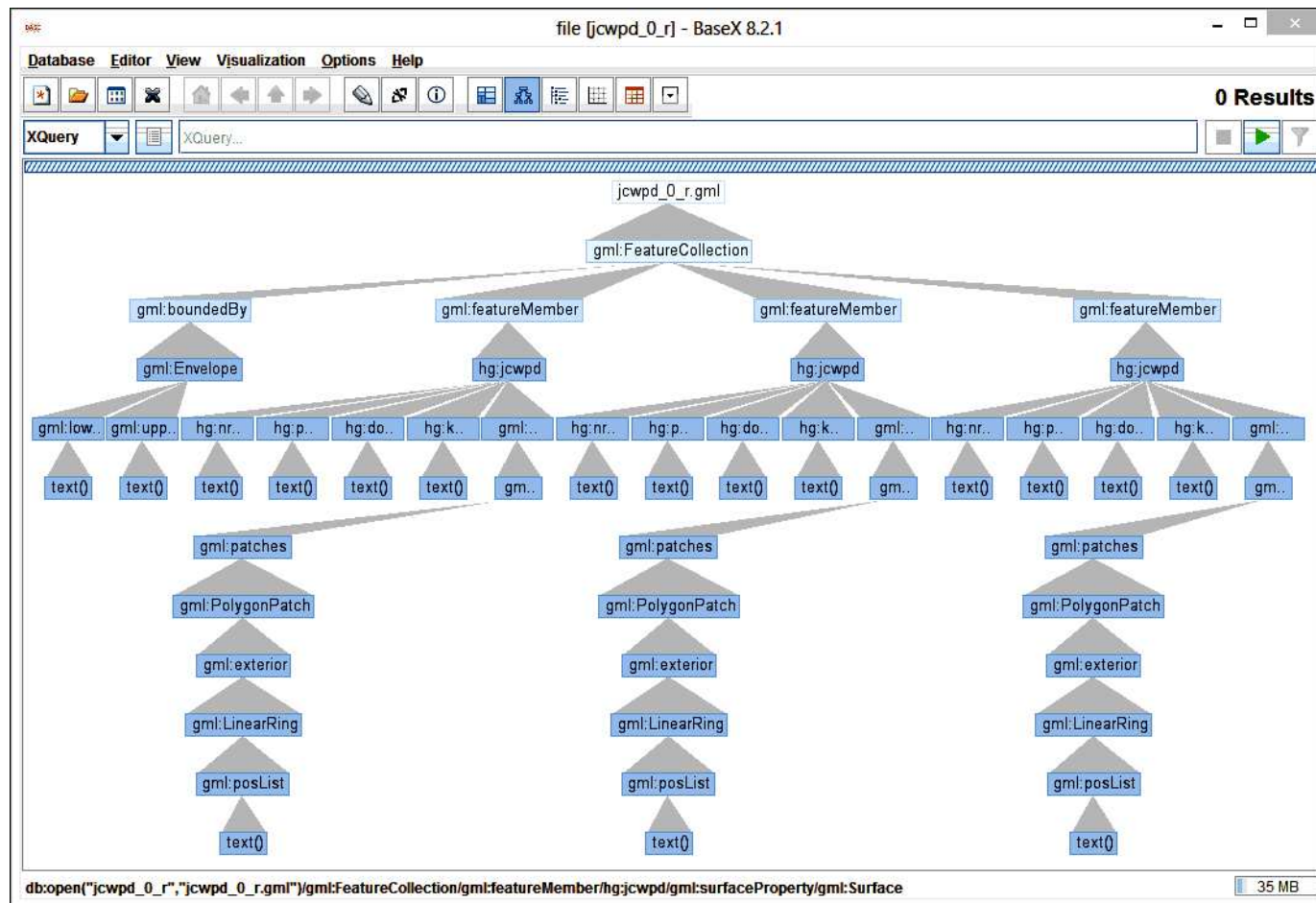
- geo:dimension – zwraca wymiar analizowanej geometrii,
- geo:geometry-type – zwraca nazwę typu analizowanej geometrii,
- geo:srid – zwraca identyfikator (ID) układu odniesienia przestrzennego (SRS – *Spatial Reference System*) analizowanej geometrii,
- geo:envelope – zwraca najmniejszy prostokąt ograniczający analizowanej geometrii w postaci elementu gml:Envelope,
- geo:as-text – zwraca reprezentację tekstową w formie WKT (*Well-known Text*) analizowanej geometrii,
- geo:as-binary – zwraca reprezentację binarną w formie WKB (*Well-known Binary*) analizowanej geometrii,
- geo:is-simple – zwraca logiczne 1 (*truth*) gdy analizowana geometria należy do kategorii prostych i gdy nie posiada punktów anomalnych (sama siebie nie przecina i nie przechodzi przez określony punkt więcej niż raz), w przeciwnym razie odpowiedzią jest logiczne 0 (*false*),



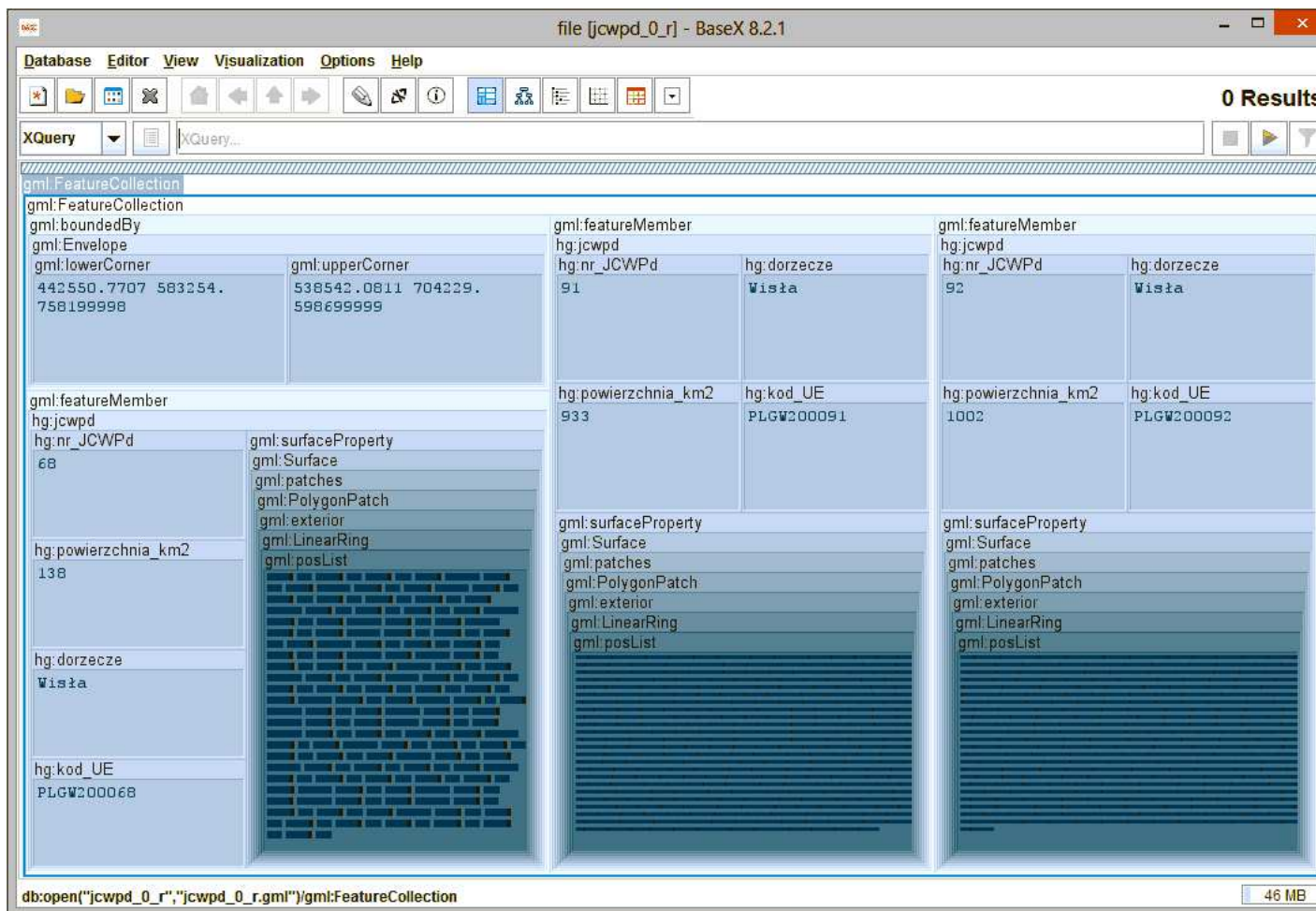
Rysunek 10.1. Okno graficznego interfejsu użytkownika (GUI) systemu zarządzania danymi BaseX. W górnej części przedstawiona jest graficzna interpretacja danych, w tym przypadku są to współrzędne geograficzne punktów (X = długość geograficzna, Y = szerokość geograficzna). Środkowa część prezentuje graficznie drzewo struktury danych – wybrany jest element "Europa Zachodnia" (EUR) na drugim poziomie tego drzewa, co ma swoje odzwierciedlenie na wykresie powyżej. Na dole przedstawiony jest fragment kodu XML (po lewej) i podstawowe parametry przetwarzanego zbioru danych „locations” (po prawej).



Rysunek 10.2. Okno graficznego interfejsu użytkownika (GUI) systemu zarządzania danymi BaseX przedstawiające strukturę zbioru danych XML w postaci zagnieżdżających się prostokątów. Element główny (root) to "mondial". poziom niższy tego zagnieżdżenia to kraje i międzynarodowe organizacje, a najniższy to poszczególne miasta.



Rysunek 10.3. Okno graficznego interfejsu użytkownika (GUI) systemu zarządzania danymi BaseX przedstawiające drzewo zbioru danych dotyczących jednolitych części wód podziemnych (JCWPd). Są to dane testowe zawierający tylko cztery przykładowe wyróżnienia, jednak struktura zagnieżdżeń tego zbioru ma 12 poziomów. Najniższy poziom „text()” zawiera teksty z liczbami opisującymi geometrię tych JCWPd.



Rysunek 10.4. Okno graficznego interfejsu użytkownika systemu BaseX przedstawiające strukturę tego samego zbioru danych XML co na rysunku 10.3. w postaci zagnieżdżających się prostokątów. Im niższy poziom zagnieżdżenia tym ciemniejszy kolor prostokąta.

- geo:boundary – zwraca granicę analizowanej geometrii w formie elementu GML2, zwracany wynik jest sekwencją elementów typu gml:Point lub gml:LinearRing jako GeometryCollection,
- geo:num-geometries – zwraca wartość określającą liczbę geometrii w analizowanym zbiorze typu GeometryCollection,
- geo:geometry-n – zwraca n -tą geometrię zawartą w analizowanej GeometryCollection w postaci zapisu GML. Jeżeli nie jest to GeometryCollection to wynik jest 1,
- geo:length – zwraca długość analizowanej geometrii, jeżeli jest to punkt to zwracana wartość wynosi 0,
- geo:num-points – zwraca wartość określającą liczbę punktów, z których jest zbudowana analizowana geometria, poprawny wynik otrzymuje się nie tylko dla geometrii typu linia, ale także dla innych typów, jak na przykład MultiPolygon, w przypadku punktu zwracana jest wartość 1,
- geo:area – zwraca powierzchnię analizowanej geometrii, dla punktu zwracana jest wartość 0,
- geo:centroid – zwraca matematyczny centroid dla analizowanej geometrii w postaci elementu gml:Point, ten punkt nie musi się znajdować wewnątrz tej geometrii,
- geo:point-on-surface – zwraca punkt położony we wnętrzu analizowanej geometrii w postaci elementu gml:Point.

Funkcje relacji przestrzennych:

- geo>equals – zwraca logiczne 1 (*truth*) gdy jedna geometria jest dokładnie identyczna z drugą geometrią, w przeciwnym razie odpowiedzią jest logiczne 0 (*false*),
- geo:disjoint – zwraca logiczne 1 (*truth*) gdy jedna geometria jest rozłączna z drugą geometrią (nie mają wspólnych punktów), w przeciwnym razie odpowiedzią jest logiczne 0 (*false*),
- geo:intersects – zwraca logiczne 1 (*truth*) gdy dwie geometrie się przecinają, jest to przypadek odwrotny niż w funkcji geo:disjoint, w przeciwnym razie odpowiedzią jest logiczne 0 (*false*),
- geo:touches – zwraca logiczne 1 (*truth*) gdy dwie geometrie się stykają, czyli mają wyłącznie wspólne punkty brzegowe, w przeciwnym razie odpowiedzią jest logiczne 0 (*false*),
- geo:crosses – zwraca logiczne 1 (*truth*) gdy dwie geometrie się przecinają, czyli mają część punktów wspólnych ale nie wszystkie każdej z nich, warunkiem jest aby punkty wspólne tworzyły geometrię o wymiarze o jeden niższym, w przeciwnym razie odpowiedzią jest logiczne 0 (*false*),
- geo:within – zwraca logiczne 1 (*truth*) gdy wszystkie punkty pierwszej geometrii są także punktami drugiej geometrii, ale nie wszystkie tej drugiej należą także do pierwszej, w przeciwnym razie odpowiedzią jest logiczne 0 (*false*),
- geo:contains – zwraca logiczne 1 (*truth*) gdy wszystkie punkty drugiej geometrii są także punktami pierwszej geometrii, ale nie wszystkie tej pierwszej należą także do drugiej, w przeciwnym razie odpowiedzią jest logiczne 0 (*false*),
- geo:overlaps – zwraca logiczne 1 (*truth*) gdy obie geometrie mają punkty wspólne, ale każda z nich ma także punkty, które nie należą do drugiej, w przeciwnym razie odpowiedzią jest logiczne 0 (*false*),
- geo:relate – zwraca logiczne 1 (*truth*) gdy zależności pomiędzy granicami, wnętrzami i zewnętrznymi dwóch geometrii spełniają wymagania określonej podanej macierzy prze-

cinania (*Intersection Matrix*), w przeciwnym razie odpowiedzią jest logiczne 0 (*false*) – jest to macierz o wymiarze 3×3 i określa relacje pomiędzy granicami, wnętrzami i zewnętrznymi obu geometrii.

Funkcje analityczne:

- `geo:distance` – zwraca najkrótszą odległość pomiędzy geometriami w jednostkach przyjętego układu odniesienia przestrzennego dla geometrii pierwszej,
- `geo:buffer` – zwraca geometrię typu wielokąt reprezentującą bufor o podanym promieniu odległości w jednostkach przyjętego układu odniesienia przestrzennego,
- `geo:convex-hull` – zwraca otoczenie niewklęsłe w postaci elementu GML – najmniejszy geometryczny zbiór niewklęsły zawierający daną geometrię,
- `geo:intersection` – zwraca rezultat przecięcia (*intersection*) w postaci elementu GML jako geometrię zawierającą wszystkie punkty wspólne – jeżeli obie geometrie się nie przecinają to wynik jest pusty,
- `geo:union` – zwraca geometrię w postaci elementu GML stanowiącą zbiór wszystkich punktów jakie mają łącznie obie geometrie,
- `geo:difference` – zwraca geometrię w postaci elementu GML stanowiącą zbiór wszystkich punktów jakie ma geometria pierwsza a nie ma ich geometria druga – jeżeli nie ma takich punktów to wynik jest pusty,
- `geo:sym-difference` – zwraca geometrię w postaci elementu GML stanowiącą połączenie dwóch zbiorów: pierwszy zbiór to wszystkie punkty jakie ma geometria pierwsza a nie ma ich geometria druga, a drugi to wszystkie punkty drugiej geometrii, których nie ma geometria pierwsza – jeżeli nie ma takich punktów to wynik jest pusty.

Funkcje specyficzna dla określonych typów geometrii:

- `geo:x` – zwraca współrzędną x punktu, punkt musi mieć tę współrzędną,
- `geo:y` – zwraca współrzędną y punktu, jeżeli jej nie ma to zwracane jest 0,
- `geo:z` – zwraca współrzędną z punktu, jeżeli jej nie ma to zwracane jest 0,
- `geo:start-point` – zwraca punkt początkowy linii (krzywej), może to być pojedyncza linia (krzywa), element typu `LineString` lub `LinearRing`.
- `geo:end-point` – zwraca punkt końcowy linii (krzywej), może to być pojedyncza linia (krzywa), element typu `LineString` lub `LinearRing`,
- `geo:is-closed` – zwraca logiczne 1 (*truth*) gdy linia (krzywa) jest zamknięta, czyli punkt początkowy jest jednocześnie punktem końcowym, może to być pojedyncza linia (krzywa), element typu `LineString`, `LinearRing` lub `MultiLineString`, w przeciwnym razie odpowiedzią jest logiczne 0 (*false*),
- `geo:is-ring` – zwraca logiczne 1 (*truth*) gdy linia (krzywa) jest pierścieniem (prosta zamknięta pętla), może to być pojedyncza linia (krzywa), element typu `LineString` lub `LinearRing`, w przeciwnym razie odpowiedzią jest logiczne 0 (*false*),
- `geo:point-n` – zwraca *n*-ty punkt linii (krzywej), może to być pojedyncza linia (krzywa), element typu `LineString` lub `LinearRing`,
- `geo:exterior-ring` – zwraca zewnętrzny pierścień określonej podanej geometrii typu wielobok (*polygon*), wynikiem jest element typu `gml:LineString`,
- `geo:num-interior-ring` – zwraca podaną liczbę pierścieni wewnętrznych określonej podanej geometrii typu wielobok (*polygon*), wynikiem jest zbiór elementów typu `gml:LineString`,
- `geo:interior-ring-n` – zwraca *n*-ty pierścień wewnętrzny określonej podanej geometrii typu wielobok (*polygon*), wynikiem jest element typu `gml:LineString`.

10.2. Podsumowanie

We wcześniejszych rozdziałach (10. i 10.1.) zostały przedstawione wyniki przeprowadzonych analiz i testów dla czterech systemów baz XML: Sedna, eXist, BaseX i MarkLogic. Systemy eXist i MarkLogic nie są czystymi rodzimymi bazami dla XML, mają jedynie status NoSQL, czyli są innymi rozwiązaniami niż tradycyjne bazy relacyjne z zapytaniami w języku SQL, ale także mogą obsługiwać dane (dokumenty) XML. MarkLogic jako bardzo drogi system komercyjny może znaleźć zastosowanie tylko w szczególnych przypadkach, gdy objętość danych jest olbrzymia, a dane te mają krytyczne znaczenie. Z tego względu bardziej dokładnej analizie zostały poddane dwa pozostałe systemy: Sedna i BaseX.

Oba te systemy są interesujące, typu *Open Source* i rodzime dla XML. Pierwszy ma przewagę polegającą na tym, że jest opracowany w językach C i C++, co sprawia że jest bardzo wydajny. Drugi nie jest tak wydajny ze względu na zastosowanie niekompilowanego do końca języka Java, co pociąga konieczność realizacji w środowisku wirtualnej maszyny Java. Jego bardzo ważnym atutem jest możliwość wykonywania operacji geoprzestrzennych za pomocą modułu Geo. Na podstawie przeprowadzonych prac można uznać, że system Sedna jest odpowiedni do przechowywania, aktualizowania i udostępniania już gotowych przetransformowanych danych, a system BaseX jest bardziej odpowiedni do przechowywania i przetwarzania danych roboczych na poszczególnych etapach ich transformacji. Z tego względu ten system był użyty w pracach testowych dotyczących opracowania metody transformacji.

11. Przegląd narzędzi do transformacji

Proces transformacji danych składa się z kilku etapów opisanych w rozdziałach 7, 8 i 9. Na poszczególnych etapach prac niezbędne jest specjalistyczne oprogramowanie narzędziowe, które służy nie tylko do samej transformacji XSLT, ale także do wielu innych prac. Problematyka przechowywania i zmiany ogólnej organizacji danych XML z wykorzystaniem systemu BaseX była przedstawiona w rozdziale 8. Do innych prac niż transformacje XSLT można zaliczyć:

- przeniesienie danych źródłowych z baz i repozytoriów plików do plików roboczych w środowisku, gdzie jest przeprowadzana transformacja,
- zmianę układu odniesienia przestrzennego danych źródłowych, do przyjętego jako wspólny, poprzedzająca przekształcenie tych danych w zapisy XML,
- zmianę sposobu kodowania tekstów w danych źródłowych do Unicode UTF-8 odpowiedniego dla danych w zapisy XML,
- weryfikację i ewentualne poprawienie geometrii elementów typu *FeatureType*,
- przekształcenie danych zapisanych w innych formatach niż język XML do takiego zapisu znacznikowego, na przykład, zapisy w formacie Esri Shape do GML w wersji 3.2.1., uzyskane z takiego przekształcenia dane mają nadal swoją pierwotną strukturę,
- analizę poprawności uzyskanych wcześniej zapisów XML pod względem syntaktycznym, semantycznym i zgodności ze schematami XSD definiującymi ich strukturę.

Po transformacji, do postaci częściowo zgodnej pod względem struktury z wymaganiami specyfikacji INSPIRE, na tych danych trzeba jeszcze przeprowadzić dużo operacji zależnych od treści transformowanych danych, a w szczególności:

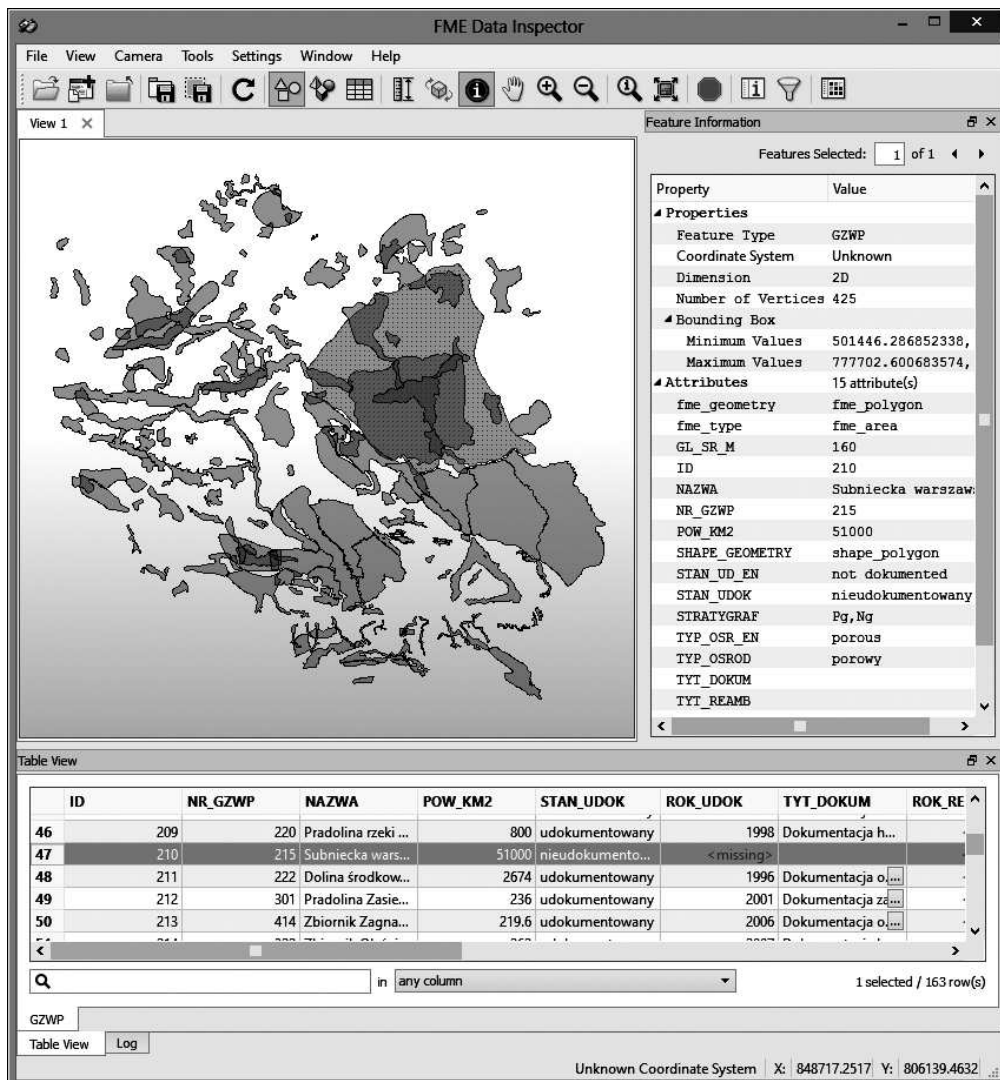
- przypisanie poszczególnym elementom typu *FeatureType* identyfikatorów INSPIRE, jeżeli są wymagane przed specyfikacją i nie zostały wcześniej przypisane,
- weryfikację i poprawienie lub uzupełnienie powiązań asocjacyjnych pomiędzy elementami w wynikowych zapisach GML, dane źródłowe z reguły takich powiązań nie mają.

Główny proces transformacji z zastosowaniem XSLT jest realizowany za pomocą procesora XSLT, co było opisane w rozdziale 9.2., są tam wymienione podstawowe narzędzia dedykowane geoinformacji, wyposażone w takie procesory lub z nimi współpracujące: FME, HALE, GeoKettle i Stetl. To specjalistyczne oprogramowanie narzędziowe będzie opisane szczegółowo w następujących rozdziałach 11.1–11.4.

Przegląd narzędzi wspomagających transformację jest przedstawiony w rozdziale 11.6.

11.1. Oprogramowanie FME (Feature Manipulation Engine)

Oprogramowanie FME jest pakietem czterech narzędzi do przetwarzania danych geoprzestrzennych, w tym do ich transformacji (Bragg, 2014; Safe Software, 2014a; 2014b; 2014c). Jest to oprogramowanie komercyjne rozwijane i rozpowszechniane przez kanadyjską firmę Safe Software inc. Pakiet ten składa się z programów FME Workbench – główne środowisko pracy, FME Quick Translator – do prostego ale szybkiego przekształcania danych, FME DataInspector (rys. 11.1.) – dla weryfikacji danych przed translacją i po

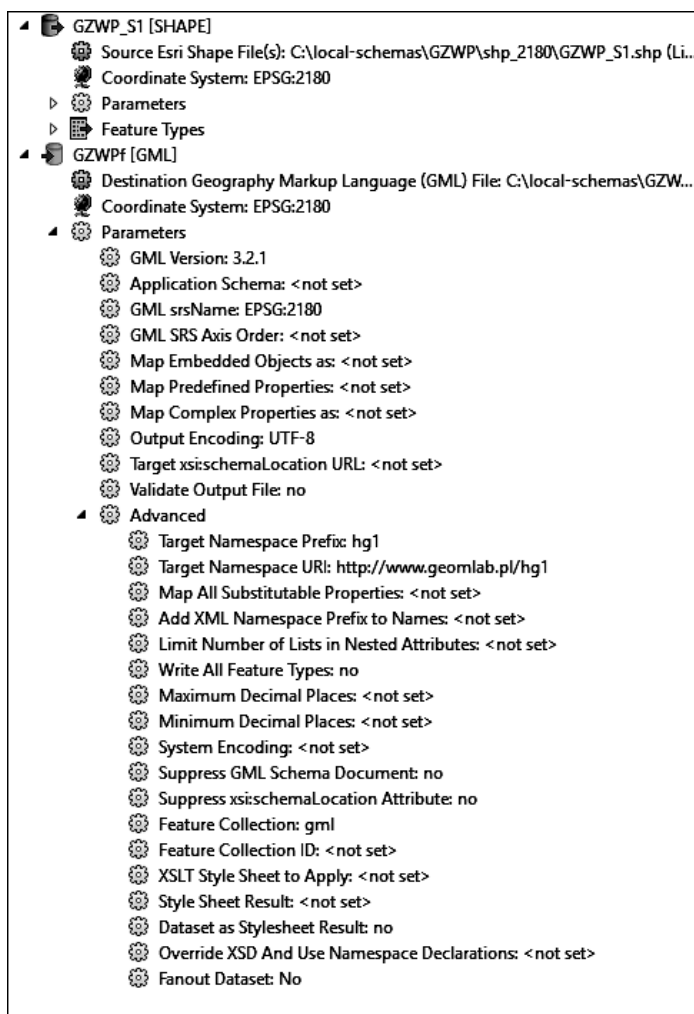


Rysunek 11.1. Okno interfejsu graficznego programu FME Data Inspector podczas analizy danych dotyczących GZWP (Głównych Zbiorników Wód Podziemnych). Lewa część na górze – graficzny obraz zasięgów poszczególnych zbiorników. Prawa część na górze – szczegółowe dane w postaci atrybutów (elementów XML) dotyczące jednego wybranego zbiornika (szary obszar regularnie zakropkowany) o nazwie „Subniecka warszawska”. Dolna część okna zawiera listę wszystkich zbiorników opisanych w tym zbiorze danych.

translacji, ostatnim narzędziem jest FME Universal Viewer – do szybkiego przeglądania graficzne zobrazowanych danych.

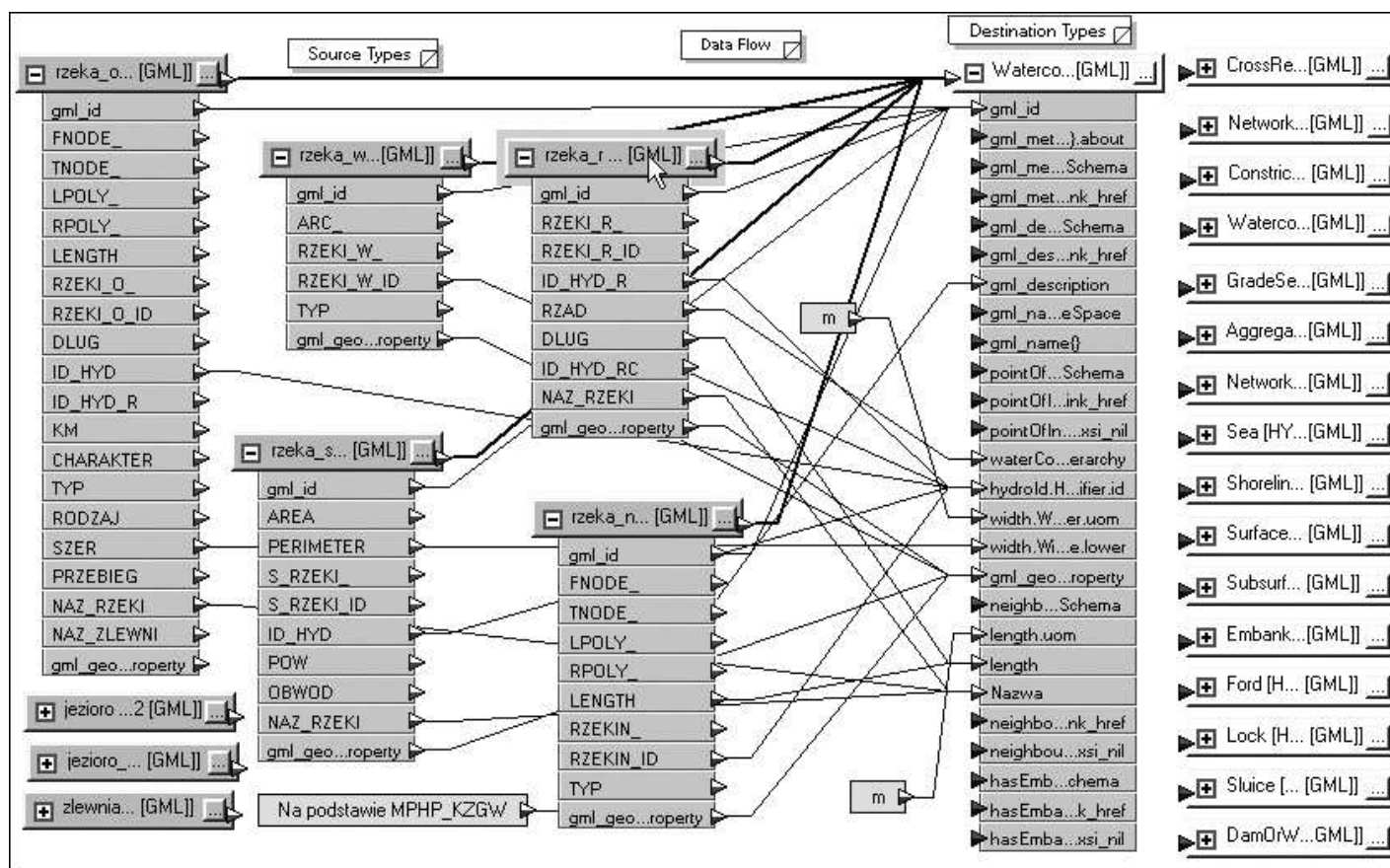
Transformacja danych do zapisów znacznikowych, w tym do języka GML, oparta jest o technologię szablonów opisaną w rozdziale 8.1. (przykład 8.9.). Oprogramowanie FME nie ma możliwości generowania arkuszy XSLT, tak jak to jest w niektórych innych systemach,

ale ma procesor Xerces (rozdział 9.2.). W rezultacie można w nim użyć przygotowany w innym oprogramowaniu odpowiedni arkusz sterujący transformacją – pozycja „*XSLT Style Sheet to Apply*” w panelu ustawień parametrów transformacji (rys. 11.2.).



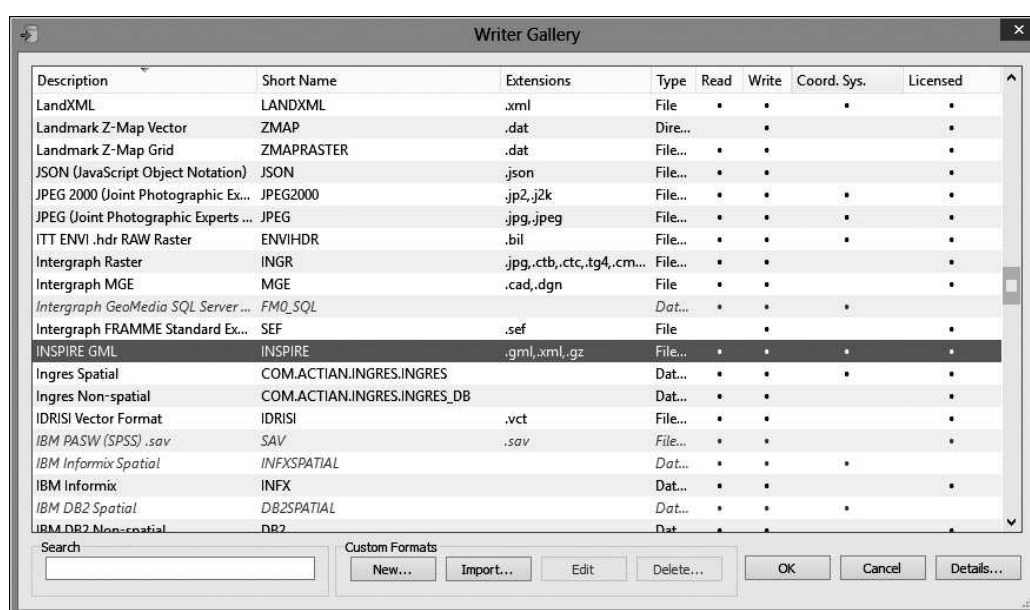
Rysunek 11.2. Fragment okna programu FME Workbench z panelem ustawień parametrów transformacji. Na samej górze są wyszczególnione parametry zbioru źródłowego GZWP_S1 w formacie Esri Shape. Wszystkie pozostałe parametry poniżej tego fragmentu dotyczą zbioru wynikowego GZWPf, który ma być zapisany w języku GML wersji 3.2.1. Przykład dotyczy danych o Głównych Zbiornikach Wód Podziemnych (GZWP).

Jedną z głównych zalet systemu FME jest obsługa ponad 300 różnych formatów, języków i interfejsów do baz danych, w tym także języka GML w wersji 3.2.1. dostosowanego do wymagań INSPIRE (rys. 11.4.). Operacje na danych geoprzestrzennych są realizowane przez łańcuchy powiązanych ze sobą, tak zwanych transformerów, z których każdy wykonuje pewną operację cząstkową (rys. 11.3.). Takich modułów przeznaczonych do różnorodnych operacji jest około 500, a ponadto można konstruować własne transformery.



Rysunek 11.3. Przykład mapowania elementów modelu danych źródłowych z elementami modelu danych wynikowych w programie FME Workbench. Dane z kilku zbiorów źródłowych są łączone w jeden zbiór wynikowy. Podstawione są też dwie stałe tekstowe „m” jako jednostki miary. Przykład dotyczy danych z tematu Hydrografia.

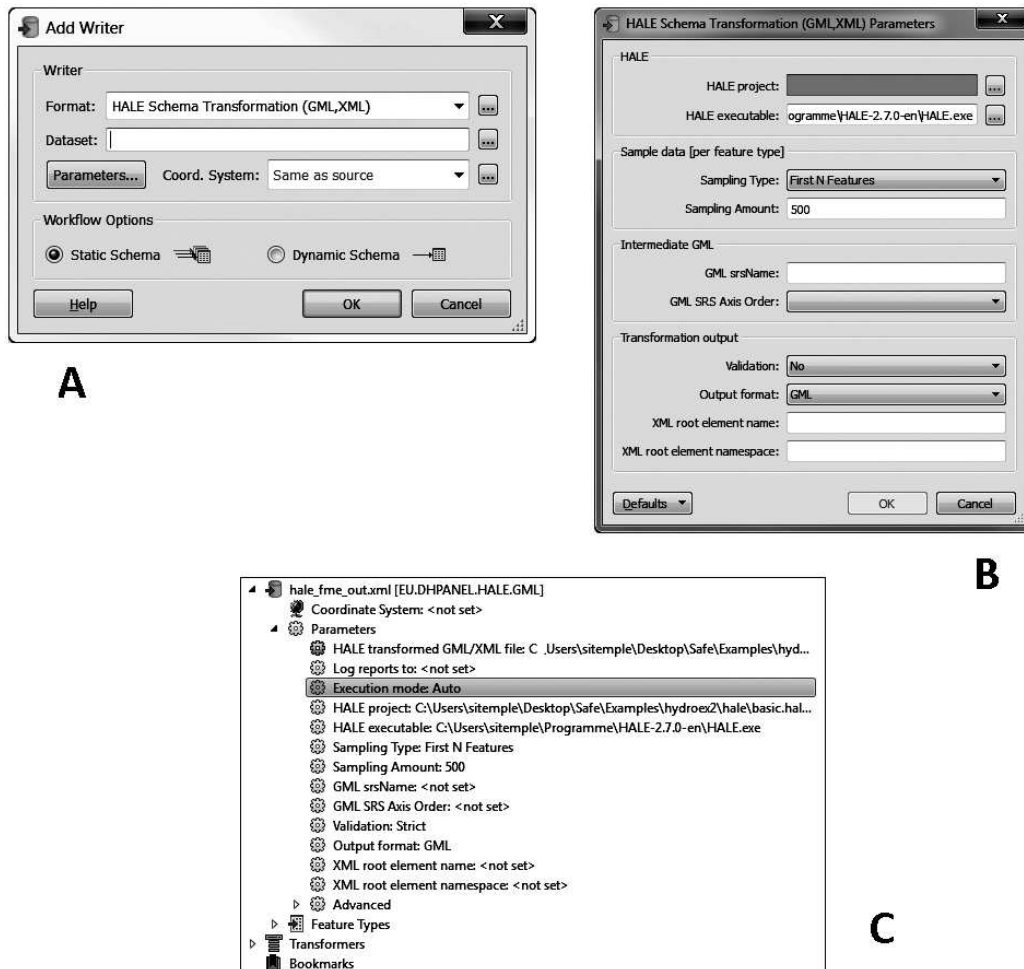
Operacje odczytu danych w systemie FME są realizowane za pomocą modułów kategorii *reader*, a zapis jest wykonywany za pomocą modułów kategorii *writer*. Liczba dostępnych modułów obu tych kategorii zależy od liczby obsługiwanych formatów. Między innymi, z powodu wymienionych tu właściwości system FME ma dominującą pozycję w obszarze oprogramowania komercyjnego dedykowanego tym zastosowaniom. Takie systemy jak ArcGIS firmy Esri i Geomedia firmy Intergraph mają wbudowane pewne elementy oprogramowania FME do translacji danych pomiędzy różnymi formatami. Jednak są one bardzo ograniczone i ten fakt nie jest bezpośrednio widoczny, zaś ustalenie tego wymaga głębszego wniknięcia w struktury wewnętrzne obu tych systemów.



Rysunek 11.4. Okno panelu Writer Gallery programu FME Workbench pokazujące listę formatów i języków jakie mogą być wczytywane (kolumna *Read*) i zapisywane (kolumna *Write*) przez ten program.

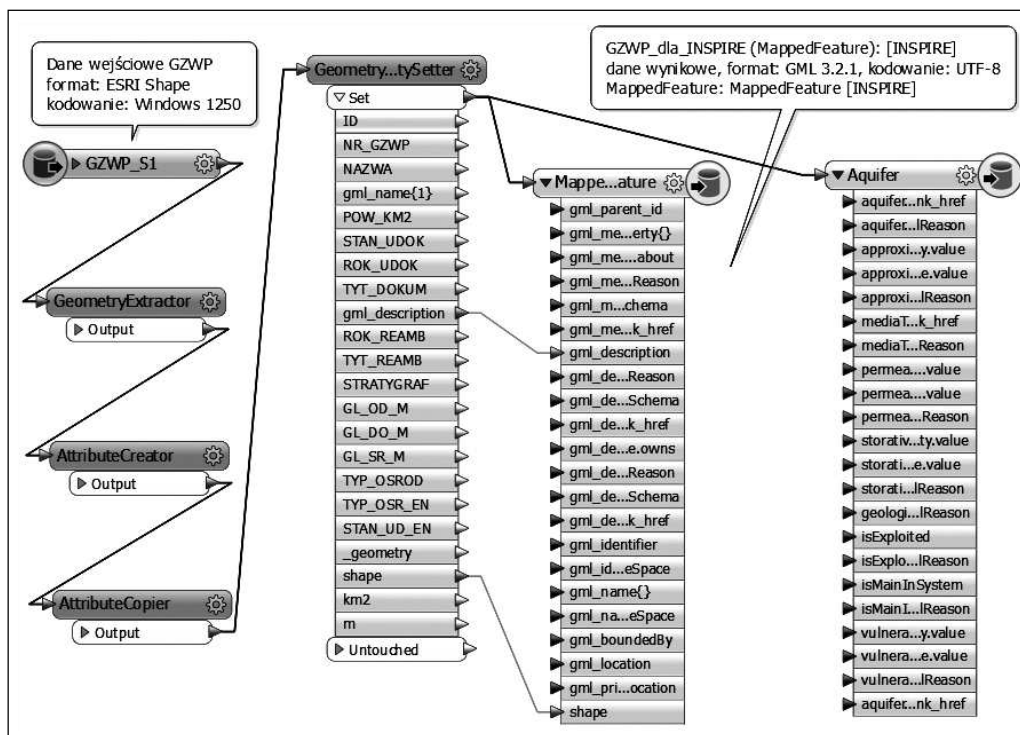
Interesującą i obiecującą koncepcją jest połączenie komercyjnego systemu FME z otwartym (*Open Source*) oprogramowaniem HALE (Reitz, 2013). Połączenie obu tych systemów polega na zastosowaniu interfejsu, dzięki któremu HALE występuje w roli modułu typu *writer*. Powstaje w ten sposób łańcuch złożony z dwóch ogniw – FME wczytuje dane, przetwarza je według określonego dla niego algorytmu, przekazuje wynik na wyjście (typu *writer*), które jest jednocześnie wejściem do środowiska systemu HALE, gdzie są kontynuowane dalsze operacje transformacji. Uruchamianie i konfigurację takiego połączenia ilustruje rysunek 11.5.

Kolejność uczestniczących w tym łańcuchu programów nie jest przypadkowa. Wyróżniającą cechą systemu FME jest bogactwo różnych przetwarzanych formatów, w tym także z zapisem binarnym, których nie można użyć bezpośrednio do transformacji XSLT. Z tego względu system FME jest użyty na początku i po wykonaniu swojego zadania, najczęściej wynikiem tego jest zapis GML lub przynajmniej XML, przekazuje ten wynik edytorowi HALE. Ponieważ cechą wyróżniającą edytor HALE jest jego ukierunkowanie na wymagania INSPIRE, zostaje użyty do końcowej fazy transformacji, po której uzyskuje się pełną zgodność zapisu z wymaganiami specyfikacji INSPIRE.



Rysunek 11.5. Uruchamianie i konfiguracja połączenia systemu HALE z systemem FME: A – dołączenie interfejsu HALE w roli modułu typu *writer*, B – ustawienie parametrów transformacji wykonywanej przez HALE w środowisku FME, C – fragment panelu ustawień parametrów transformacji FME z widocznym modulem HALE.

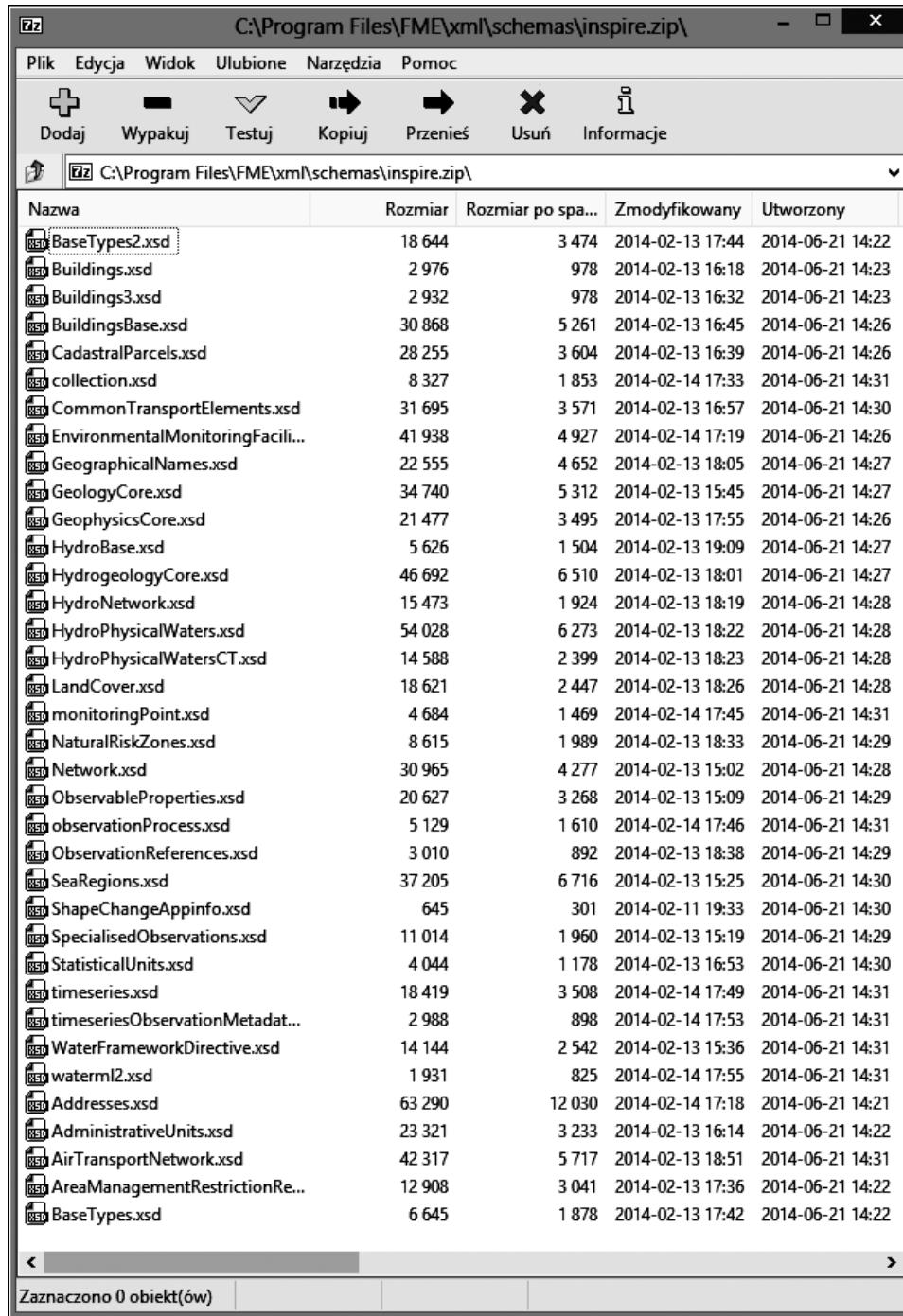
W ramach przeprowadzonych prac testowane były zapisy danych w GML 3.2.1, INSPIRE GML i wyniki uzyskane z oprogramowania HALE (rozdział 11.2.). Uzyskane wyniki testowania oprogramowania FME wskazują, że jest to narzędzie przydatne głównie w pierwszych etapach transformacji. Jednak nie można pominąć faktu, że są problemy z atrybutami geometrycznymi. Dla wyróżnień powierzchniowych (*Aera*) FME nie daje możliwości wyboru formy zapisu geometrii. Przyjmowana jest domyślnie bez możliwości zmiany opcja *All types are allowed*. Program automatycznie sam wybiera typ geometrii dla zapisu wynikowego opierając się o typ geometrii w pliku źródłowym. Nie ma możliwości zapisu geometrii powierzchniowej w formie *Location*, która występuje w GML 3.2 i 3.1, ale jest uznana za nieaktualną (*deprecated*). Dla prostych geometrii powierzchniowych FME stosują atrybut geometryczny tylko w postaci *surfaceProperty* typu *Surface*.



Rysunek 11.6. Mapowanie elementów modeli danych źródłowych GZWP_S1 (format Esri Shape) do modelu danych INSPIRE dla hydrogeologii (GZWPf w zapisie GML 3.2.1). Dane z jednego zbioru są przesyłane do dwóch zbiorów zawierających oddzielenia wyróżnienia typu MappedFeature i Aquifer. W dalszym etapie te dwa zbiory wynikowe będą połączone w jeden. Więcej objaśnień w tekście.

W ramach przeprowadzonych prac testowych (rys. 11.6.) były zapisy danych w GML 3.2.1, INSPIRE GML i wyniki uzyskane z oprogramowania HALE (rozdział 11.2.). Rysunek 11.6. ilustruje problem jaki wystąpił w wyniku uzależnienia klasy *HydrogeologicalUnit* od klasy *GeologicUnit*, co zostało opisane w rozdziale 6.2. W konsekwencji tego, wyróżnienia hydrogeologiczne są obciążone wieloma niepotrzebnymi atrybutami (wewnętrznymi elementami GML), ale także nie posiadają geometrii, czyli lokalizacji geoprzestrzennej. Lokalizacja ta w modelu geologicznym INSPIRE jest określana za pomocą klasy *MappedFeature*, co ma odpowiadać wychodniom utworów geologicznych. Nie jest to rozwiązanie odpowiednie z punktu widzenia hydrogeologii.

Inny problem jaki wystąpił podczas prac testowych to aktualność schematów aplikacyjnych XSD dla tematów INSPIRE. Oprogramowanie FME jest w dużym stopniu dostosowane do transformacji danych dedykowanych specyfikacjom INSPIRE. Jednak schematy aplikacyjne dotyczące danych INSPIRE wbudowane w FME Workbench w czasie prowadzenia opisanych tu prac testowych były bardzo nieaktualne, ponieważ miały datę 26.04.2010. Z tego powodu konieczne było dokonanie niezbędnych aktualizacji, zarówno samych schematów (rys. 11.7.), jak i wielu funkcji tego oprogramowania, między innymi zapisu definicji schematów. Fragmenty wersji oryginalnej pliku „inspire.fmf” definiującego moduł *INSPIRE Writer* jest przedstawiony w przykładzie 11.1.



Nazwa	Rozmiar	Rozmiar po spa...	Zmodyfikowany	Utworzony
BaseTypes2.xsd	18 644	3 474	2014-02-13 17:44	2014-06-21 14:22
Buildings.xsd	2 976	978	2014-02-13 16:18	2014-06-21 14:23
Buildings3.xsd	2 932	978	2014-02-13 16:32	2014-06-21 14:23
BuildingsBase.xsd	30 868	5 261	2014-02-13 16:45	2014-06-21 14:26
CadastralParcels.xsd	28 255	3 604	2014-02-13 16:39	2014-06-21 14:26
collection.xsd	8 327	1 853	2014-02-14 17:33	2014-06-21 14:31
CommonTransportElements.xsd	31 695	3 571	2014-02-13 16:57	2014-06-21 14:30
EnvironmentalMonitoringFacili...	41 938	4 927	2014-02-14 17:19	2014-06-21 14:26
GeographicalNames.xsd	22 555	4 652	2014-02-13 18:05	2014-06-21 14:27
GeologyCore.xsd	34 740	5 312	2014-02-13 15:45	2014-06-21 14:27
GeophysicsCore.xsd	21 477	3 495	2014-02-13 17:55	2014-06-21 14:26
HydroBase.xsd	5 626	1 504	2014-02-13 19:09	2014-06-21 14:27
HydrogeologyCore.xsd	46 692	6 510	2014-02-13 18:01	2014-06-21 14:27
HydroNetwork.xsd	15 473	1 924	2014-02-13 18:19	2014-06-21 14:28
HydroPhysicalWaters.xsd	54 028	6 273	2014-02-13 18:22	2014-06-21 14:28
HydroPhysicalWatersCT.xsd	14 588	2 399	2014-02-13 18:23	2014-06-21 14:28
LandCover.xsd	18 621	2 447	2014-02-13 18:26	2014-06-21 14:28
monitoringPoint.xsd	4 684	1 469	2014-02-14 17:45	2014-06-21 14:31
NaturalRiskZones.xsd	8 615	1 989	2014-02-13 18:33	2014-06-21 14:29
Network.xsd	30 965	4 277	2014-02-13 15:02	2014-06-21 14:28
ObservableProperties.xsd	20 627	3 268	2014-02-13 15:09	2014-06-21 14:29
observationProcess.xsd	5 129	1 610	2014-02-14 17:46	2014-06-21 14:31
ObservationReferences.xsd	3 010	892	2014-02-13 18:38	2014-06-21 14:29
SeaRegions.xsd	37 205	6 716	2014-02-13 15:25	2014-06-21 14:30
ShapeChangeAppinfo.xsd	645	301	2014-02-11 19:33	2014-06-21 14:30
SpecialisedObservations.xsd	11 014	1 960	2014-02-13 15:19	2014-06-21 14:29
StatisticalUnits.xsd	4 044	1 178	2014-02-13 16:53	2014-06-21 14:30
timeseries.xsd	18 419	3 508	2014-02-14 17:49	2014-06-21 14:31
timeseriesObservationMetadat...	2 988	898	2014-02-14 17:53	2014-06-21 14:31
WaterFrameworkDirective.xsd	14 144	2 542	2014-02-13 15:36	2014-06-21 14:31
waterml2.xsd	1 931	825	2014-02-14 17:55	2014-06-21 14:31
Addresses.xsd	63 290	12 030	2014-02-14 17:18	2014-06-21 14:21
AdministrativeUnits.xsd	23 321	3 233	2014-02-13 16:14	2014-06-21 14:22
AirTransportNetwork.xsd	42 317	5 717	2014-02-13 18:51	2014-06-21 14:31
AreaManagementRestrictionRe...	12 908	3 041	2014-02-13 17:36	2014-06-21 14:22
BaseTypes.xsd	6 645	1 878	2014-02-13 17:42	2014-06-21 14:22

Rysunek 11.7. Lista zaktualizowanych schematów XSD dotyczących wszystkich tematów INSPIRE zaimportowana do systemu FME dla umożliwienia prowadzenia testów transformacji.

Przykład 11.1. Fragment pliku konfiguracyjnego „inspire.fmf” dla modułu *INSPIRE Writer* zawierający nieaktualne deklaracje przestrzeni nazw i odsyłacze dla tematu Geologia, w tym Hydrogeologia:

```
! INSPIRE format metafile
(...)
! Revision: $Id: inspire.fmf 82700 2011-10-11 20:16:59Z
(...)
DEFAULT_MACRO INSPIRE_NS_URI_THEME_LIST_DEFAULT
(...)
urn:x-inspire:specification:gmlas:Geology:0.0
(...)
DEFAULT_MACRO INSPIRE_NS_URI_THEME_LIST_LOOKUP
(...)
%Geology:0.0,urn:x-inspire:specification:gmlas:Geology:0.0.
```

Aby móc wykorzystać to oprogramowanie do testów transformacji, w pliku tym dwie linie definiujące schemat tematu Geology zostały zastąpione dwoma grupami po 6 linii definiujących odpowiednio sześć aktualnych schematów dotyczących danych tego tematu (przykład 11.2.)

Przykład 11.2. Dwie grupy linii pliku konfiguracyjnego z przykładu 11.1., którymi zastąpiono nieaktualne i przez to niepoprawne wpisy w tym pliku:

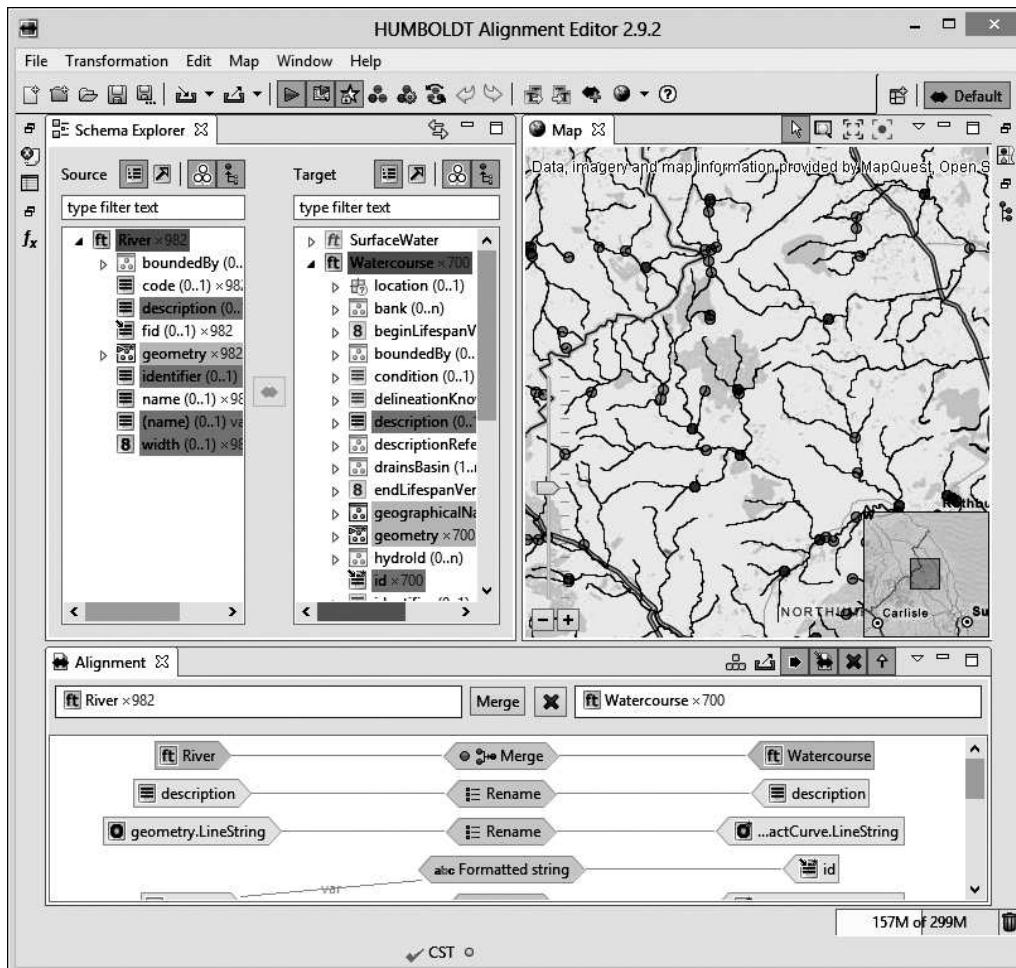
```
urn:x-inspire:specification:gmlas:Geology:2.0
urn:x-inspire:specification:gmlas:GeologyCore:3.0rc3
urn:x-inspire:specification:gmlas:GeologyExtension:2.0
urn:x-inspire:specification:gmlas:GeophysicsCore:3.0rc3
urn:x-inspire:specification:gmlas:GeophysicsExtension:2.0
urn:x-inspire:specification:gmlas:Hydrogeology:3.0rc3

%Geology:2.0,urn:x-inspire:specification:gmlas:Geology:2.0
%GeologyCore:3.0rc3,urn:x-inspire:specification:gmlas:GeologyCore:3.0rc3
%GeologyExtension:2.0,urn:x-inspire:specification:gmlas:GeologyExtension:2.0
%GeophysicsCore:3.0rc3,urn:x-inspire:specification:gmlas:GeophysicsCore:3.0rc3
%GeophysicsExtension:2.0,urn:x-inspire:specification:gmlas:GeophysicsExtension:2.0
%Hydrogeology:3.0rc3,urn:x-inspire:specification:gmlas:Hydrogeology:3.0rc3
```

Bez tych modyfikacji konfiguracji oprogramowania i podstawienia aktualnych plików XSD dla tematu Geology użycie modułu *INSPIRE Writer* dla zapisu danych dedykowanych tematom INSPIRE nie jest możliwe.

11.2. Oprogramowanie HUMBOLDT Alignment Editor

Humboldt Alignment Editor (HALE) (rys. 11.8.) to oprogramowanie *Open Source* zbudowane na platformie Eclipse (DHP, 2014a). Jest to bardzo zaawansowane technologicznie narzędzie do transformacji danych geoprzestrzennych, głównie ukierunkowane na tworzenie zapisów zgodnych ze specyfikacjami INSPIRE. Zastosowanie platformy Eclipse przesądza



Rysunek 11.8. Okno interfejsu graficznego (GUI) edytora HALE przedstawiające proces mapowania modeli danych z zakresu hydrografii. Panel na górze po stronie lewej pokazuje elementy modelu źródłowego i modelu wynikowego. Pozycje na ciemnym tle to elementy już zmapowane.

Panel na górze po stronie prawej przedstawia interpretacją graficzną danych i pozwala sprawdzić zgodność geometryczną danych wynikowych z danymi źródłowymi. Na panelu na dole widoczny jest graficzny obraz poszczególnych powiązań utworzonych w procesie mapowania modeli.

o tym, że wszystko jest napisane w języku Java i wykonanie tego oprogramowania odbywa się w środowisku wirtualnej maszyny Java (*Java Virtual Machine – JVM*), co wyraźnie obniża wydajność tego systemu.

Nazwa „Alignment Editor” wywodzi się z problematyki chemii molekularnej i genetyki, gdzie istnieje potrzeba wyszukiwania sekwencji cząstek organicznych, porównywania ich i dopasowywania pomiędzy różnymi, często bardzo długimi, ciągami zapisów. Tu podobne podejście zastosowano do struktur danych geoprzestrzennych i ich modeli.

Pierwsza wersja systemu HALE została opracowana w latach 2006–2011 w ramach zintegrowanego europejskiego projektu pod nazwą *Humboldt Project* (DHP, 2009; 2011; 2014b).

Zadaniem tego projektu było ułatwienie i wsparcie prac nad harmonizacją danych przestrzennych w skali międzynarodowej. Od roku 2011, po zakończeniu projektu Humboldt, prace nad rozwojem HALE są koordynowane pod auspicjami organizacji typu asocjacja o nazwie *Data Harmonisation Panel* (DHP). DHP został założony z celem wspierania międzynarodowej społeczności ekspertów i organizacji, które mają do czynienia z harmonizacją danych przestrzennych, w ramach prac dotyczących INSPIRE. Cel ten jest realizowany przez rozpowszechnianie, wykorzystywanie i pogłębianie metodyki i technologii harmonizacji, a w szczególności dalszego rozwoju narzędzi i usług opracowanych w projekcie Humboldt. DHP jest strukturą otwartą i każdy może uczestniczyć w jej pracach nad zadaniami określonymi w programie. Należą do nich między innymi:

- rozwój współpracy między stronami zainteresowanymi harmonizacją danych przestrzennych, w tym działań normalizacyjnych,
- badania nad aktualną sytuacją i tendencjami w zakresie harmonizacji danych, w tym analizy rynku,
- kontynuowanie utrzymania, rozwoju i eksploatacji narzędzi i usług opracowanych w ramach projektu Humboldt, w tym dbanie o możliwie jak najwyższy ich poziom,
- wsparcie rozwoju europejskiej infrastruktury danych przestrzennych zgodnie z planem przyjętym w INSPIRE,
- transfer wiedzy z zakresu danych geoprzestrzennych w różnych dziedzinach zastosowań, takich jak: bezpieczeństwo, środowisko i zarządzanie w sytuacjach zagrożeń.

Mapowanie, a w dalszej kolejności transformowanie elementów, do których należą typy obiektów i atrybuty jednego schematu pojęciowego (na przykład schematy aplikacyjne GML, schematy bazy danych lub modele UML) do drugiego schematu jest podstawą harmonizacji danych. Edytor HALE jest w pierwszej kolejności narzędziem do definiowania i oceny poprawności pojęciowej mapowania pomiędzy różnymi schematami. HALE pozwala specjalistom z danej dziedziny stworzyć spójne logicznie i semantycznie mapowania i przekształcać dane przestrzenne na podstawie tych mapowań. Co więcej, główny nacisk kładzie się tu na rejestrację wszystkich faz i operacji procesu transformacji i jego rezultatów w postaci zapisu wynikowego. W procesie mapowania HALE używa języka wysokiego poziomu do zapisu wyników mapowania. Zapisy te mogą być później wykorzystane przez komponent przetwarzania pod nazwą *Conceptual Schema Transformer* do przeprowadzania procesu transformacji, jako następny etap prac, na przykład z niezharmonizowanych danych źródłowych do formy zgodnej z wymaganiami specyfikacji INSPIRE. Aby ten złożony proces przebiegał bardziej efektywnie, z pełną kontrolą i aby jakość transformacji była dostatecznie wysoka, HALE umożliwia wstępną pracę z próbkami danych geoprzestrzennych i ich graficzną wizualizację. Główne zalety tego oprogramowania można wyrazić w kilku punktach:

- HALE umożliwia wyjątkowo przyjazne i deklaratywne podejście do interaktywnego budowania struktury mapowania porównywanych schematów danych,
- w oparciu o silny paradygmat na poziomie pojęciowym dotyczącym reguł mapowania, edytor ten stwarza warunki, w których praca z nim pozwala na lepsze zrozumienie tego procesu i lepsze nim zarządzanie,
- system HALE wykorzystuje zarówno informacje zawarte w różnych zaimplementowanych schematach pojęciowych, jak i w poszczególnych wyróżnieniach geoprzestrzennych w celu zapewnienia wysokiej jakości struktur mapowania,
- edytor ten posiada bogaty interfejs, zarówno tekstowy jak i graficzny przystosowany do pracy nad danymi geoprzestrzennymi,

- HALE daje na bieżąco informację zwrotną o przebiegu procesu mapowania danych i transformacji z jednego schematu do drugiego, dzięki czemu proces mapowania i transformacji jest stale nadzorowany,
- jedną z zalet tego edytora jest możliwość opracowywania mapowań we współpracy, w ramach zespołu osób zajmujących się jednym projektem,
- HALE daje możliwość dokumentowania przyjętych ograniczeń dotyczących mapowania i zapisy te są tworzone za pomocą języka opisu niezgodności MDL (*Mismatch Description Language*).

Edytor HALE jest jednym z kilku przedsięwzięć zrealizowanych w ramach projektu *Humboldt Project*. Inne cele jakie postawiono to przede wszystkim:

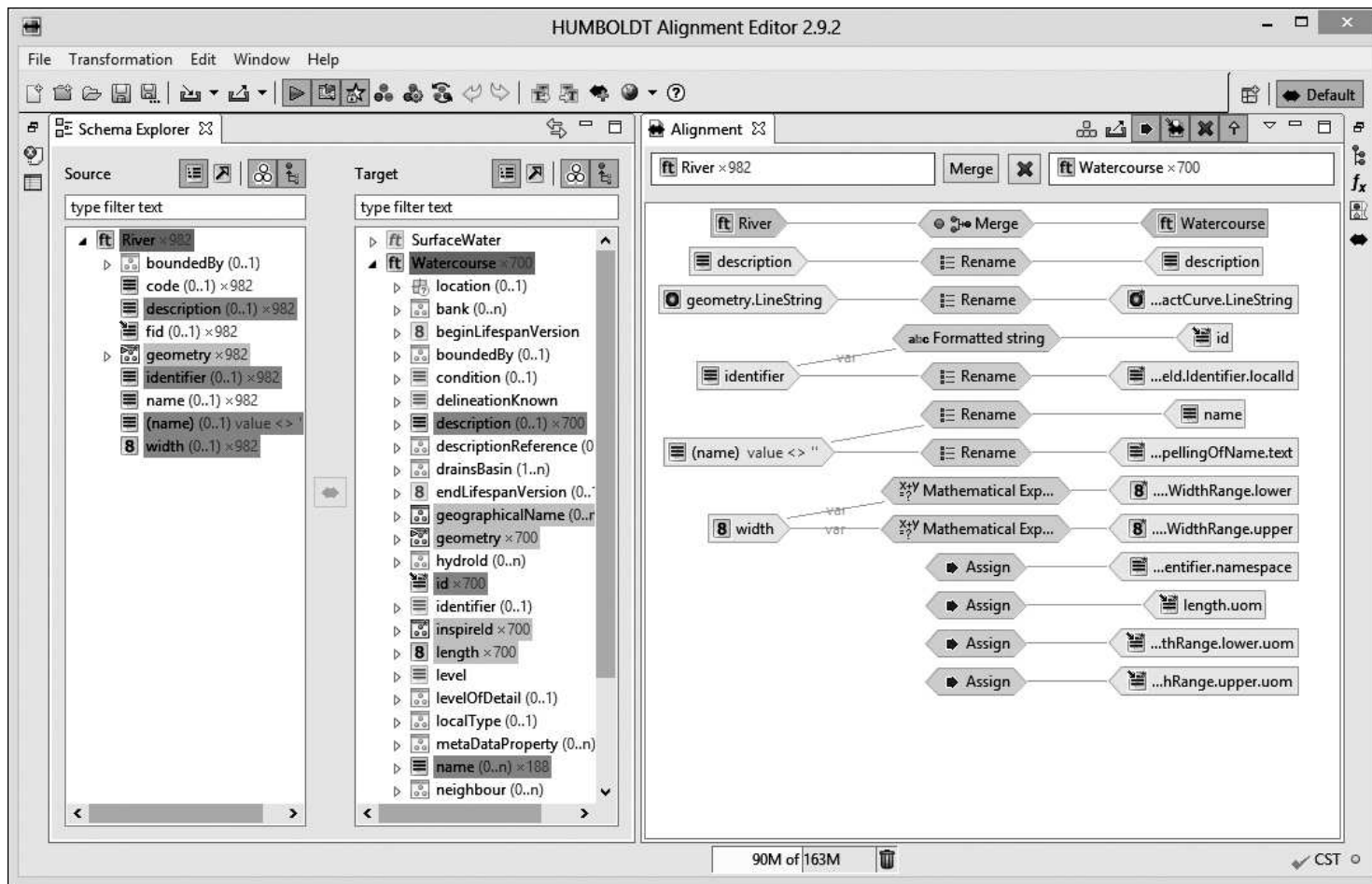
- zgromadzenie wiedzy dziedzinowej i specyficznej dla określonych zastosowań niezbędnej do opracowania metodyki z uwzględnieniem zaawansowanych formalizmów dotyczących procesu transformacji z jednego schematu pojęciowego do drugiego,
- wsparcie procesu definiowania różnych produktów informatycznych z zakresu geomatyki, jak na przykład: schemat źródłowy i wynikowy, SRS (*Spatial Reference System* – układ odniesienia przestrzennego), zasięg przestrzenny i inne, elementy te są wymagane przy specyfikowaniu wyniku transformacji,
- przeprowadzanie transformacji może być traktowane jako część ogólniejszego procesu przetwarzania informacji w trybie wsadowym na polecenia, także często zdalne,
- usprawnienie automatyzacji procesów harmonizacji danych, co jest uwarunkowane możliwością zdobycia wymaganej wiedzy oraz dostępnością tej wiedzy.

Dla osiągnięcia tych celów, w ramach projektu Humboldta zostały zaprojektowane i wdrożone trzy główne elementy, których podstawowym zadaniem jest pomoc w zdobywaniu wiedzy użytkowników w zakresie modelowania i harmonizacji danych, jak również późniejszego wykonywania transformacji pomiędzy schematami pojęciowymi.

- *Humboldt GeoModel Editor* – oprogramowanie narzędziowe do opracowywania modeli pojęciowych dotyczących geoinformacji. W procesie prac dotyczących mapowania i transformacji edytor ten pomaga w opracowaniu specyfikacji modeli danych źródłowych i wynikowych.
- *Humboldt Alignment Editor* (HALE) – tu opisane bardziej szczegółowo oprogramowanie narzędziowe do mapowania różnych modeli pojęciowych w formie schematów aplikacyjnych. Jego zadaniem jest ułatwienie specyfikowania mapowania między modelem danych źródłowych i modelem wynikowym. Edytor ten umożliwia także realizację transformacji, między innymi dla sprawdzenia poprawności opracowanych algorytmów mapowań.
- *Humboldt Conceptual Schema Transformer* – oprogramowanie narzędziowe przeznaczone do wykonywania procesów transformacji w trybie operacyjnym na podstawie wcześniej opracowanych algorytmów mapowań, na przykład za pomocą edytora HALE.

Ze względu na eksperymentalny charakter opisywanych tu prac, zarówno *Humboldt GeoModel Editor*, jak i *Humboldt Conceptual Schema Transformer* nie były w pracach testowych potrzebne i z tego względu przedmiotem bardziej szczegółowych analiz był tylko edytor HALE.

Testy zastosowania tego narzędzia wypadły całkowicie pomyślnie. Rysunek 11.9. przedstawia wynik mapowania i przeprowadzonej na jego podstawie transformacji danych testowych. W graficznym panelu mapowania widocznym na tym rysunku po prawej stronie



Rysunek 11.9. Okno interfejsu graficznego (GUI) edytora HALE przedstawiające ten sam proces mapowania modeli danych z zakresu hydrografii, który jest pokazany na rysunku 11.8. W tym przypadku inna konfiguracja paneli pozwala na pokazanie wszystkich powiązań tego schematu mapowania. Objasnienia w tekście.

przedstawione są wszystkie powiązania, jakie można było w tym przypadku wykonać. Ponieważ poszczególne transformowane elementy nie mają takiej samej formy w danych źródłowych i wynikowych, wymagają różnych przekształceń – jest to realizowane za pomocą zbioru zdefiniowanych funkcji. Powiązanie dwóch elementów wraz z przypisaną odpowiednią funkcją jest tam nazywane komórką mapowania (*mapping cell*). Funkcje są w pewnym sensie odpowiednikami transformerów w systemie FME. Jest ich 29 i są podzielone na 6 grup: *General* – funkcje podstawowe, *Geometric* – operacje geometryczne, *Groovy* – funkcje zdefiniowane w języku Groovy z możliwością rozbudowy, *Numeric* – funkcje matematyczne, *XSLT* – jedna funkcja dedykowana generowaniu schematu mapowania w postaci arkusza XSLT. Do oddzielnej grupy funkcji o nazwie *Inspire* należy funkcja generowania identyfikatorów INSPIRE (funkcja *Inspire Identifier*) i funkcja *Geographical Name* dla nazw geograficznych zdefiniowanych w specyfikacji tego tematu, a używanych w wielu innych tematach.

Na rysunku 11.9. w panelu mapowania widoczne są przykłady zastosowanie niektórych funkcji należących do grupy podstawowych i matematycznych:

- *Retype* – mapowanie jednego typu źródłowego w inny typ wynikowy,
- *Merge* – łączenie wielu instancji typu źródłowego w jedną instancję typu wynikowego na podstawie jednej lub kilku właściwości dopasowania,
- *Rename* – kopiowanie wybranego elementu źródłowego do elementu wynikowego ze zmianą nazwy,
- *Assign* – przypisanie wartości stałej do elementu wynikowego,
- *Formatted string* – tworzenie sformatowanego tekstu w oparciu o wzorzec i zmienne wejściowe,
- *Mathematical Expression* – wyliczenie wartości wynikowej za pomocą wyrażenia matematycznego z użyciem podanych zmiennych.

Ze względu na znaczne objętości tworzonych zapisów i wynikającą z tego potrzebę ich ograniczania przedstawione tu przykłady są bardzo uproszczone. Poniżej (przykład 11.3.) przedstawiony jest przykładowy wynik transformacji przeprowadzonej edytorem HALE, zapisany w języku GML 3.2.1. zgodnie z wymaganiami specyfikacji INSPIRE. Fragment zapisu schematu mapowania (*alignment*) w dokumencie XML dla transformacji oprogramowaniem HALE dla tematu Zagospodarowanie przestrzenne (*Land use*) jest przedstawiony w rozdziale 8. (przykład 8.10.).

Przykład 11.3. Wynik transformacji przeprowadzonej edytorem HALE i zapisany w języku GML 3.2.1. zgodnie z wymaganiami specyfikacji INSPIRE:

```
<?xml version="1.0" ?>
<base:SpatialDataSet
  xmlns:base="urn:x-inspire:specification:gmlas:BaseTypes:3.2"
  xmlns:gts="http://www.isotc211.org/2005/gts"
  xmlns:wfd="urn:x-inspire:specification:gmlas:WaterFrameworkDirective:0.0"
  xmlns:gco="http://www.isotc211.org/2005/gco"
  xmlns:net="urn:x-inspire:specification:gmlas:Network:3.2"
  xmlns:hfp="http://www.w3.org/2001/XMLSchema-hasFacetAndProperty"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:hy="urn:x-inspire:specification:gmlas:HydroBase:3.0"
  xmlns:hy-p="urn:x-inspire:specification:gmlas:HydroPhysicalWaters:3.0"
  xmlns:gml="http://www.opengis.net/gml/3.2"
  xmlns:gss="http://www.isotc211.org/2005/gss"
```



```

xmlns:gsr="http://www.isotc211.org/2005/gsr"
xmlns:gmd="http://www.isotc211.org/2005/gmd"
<!-- (...) -->
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
gml:id="_bf35d7c2-791f-49a1-9573-c27297d18132"

xsi:schemaLocation=
"urn:x-inspire:specification:gmlas:HydroPhysicalWaters:3.0
http://inspire.ec.europa.eu/schemas/hy-p/3.0/HydroPhysicalWaters.xsd">

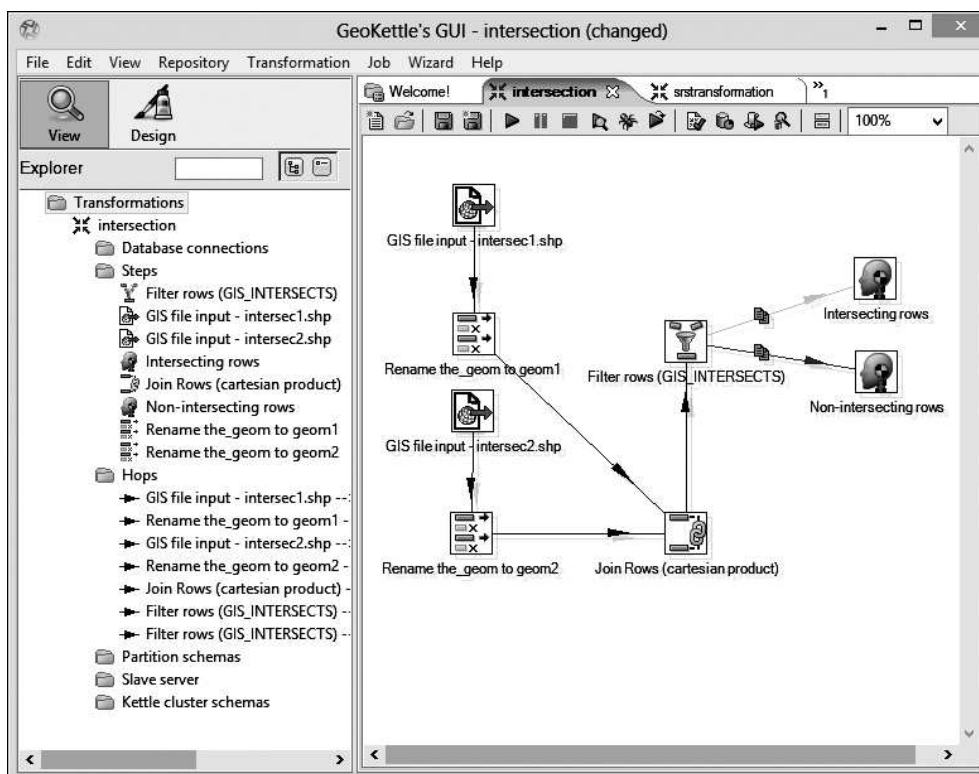
<base:identifier>
  <base:Identifier>
    <base:localId>HydroGeo</base:localId>
    <base:namespace>ge-h</base:namespace>
  </base:Identifier>
</base:identifier>
<base:metadata xsi:nil="true"></base:metadata>
<base:member>
  <hy-p:Watercourse gml:id="_d1a6215e-3185-429f-b9b1-f160437e4f99">
    <gml:description> Dane testowe dotyczące hydrografii pochodzą
      z przykładu dostarczonego przez autorów edytora HALE
    </gml:description>
    <hy-p:beginLifespanVersion xsi:nil="true"/>
    <hy-p:geometry>
      <gml:Polygon gml:id="_4be88abe-fd12-49ec-a60e-bfeec8a4c941"
        srsName="EPSG:27700">
        <gml:exterior>
          <gml:LinearRing>
            <gml:posList>389051.4 603715.4 388912.4 603626.1
              388911.4 603625.6 388701.4 603555.6
              388700.6 603555.4</gml:posList>
          </gml:LinearRing>
        </gml:exterior>
      </gml:Polygon>
    </hy-p:geometry>
    <hy-p:inspireId>
      <base:Identifier>
        <base:localId>
          _d1a6215e-3185-429f-b9b1-f160437e4f99
        </base:localId>
        <base:namespace>_example</base:namespace>
      </base:Identifier>
    </hy-p:inspireId>
    <hy-p:origin xsi:nil="true"/>
    <hy-p:persistence xsi:nil="true"/>
    <hy-p:tidal xsi:nil="true"/>
    <hy-p:drainsBasin xsi:nil="true"/>
    <hy-p:delineationKnown xsi:nil="true"/>
    <hy-p:length uom="m">848.2</hy-p:length>
    <hy-p:level xsi:nil="true"/>
    <hy-p:width>
      <hy-p:WidthRange>
        <hy-p:lower uom="m">3.5</hy-p:lower>
        <hy-p:upper uom="m">4.7</hy-p:upper>
      </hy-p:WidthRange>
    </hy-p:width>
  </hy-p:Watercourse>
</base:member>
<!-- (...) -->
</base:SpatialDataSet>

```

11.3. Oprogramowanie GeoKettle

Kolejnym zaawansowanym narzędziem przeznaczonym do transformacji danych geoprzestrzennych jest GeoKettle (OSGeo-Live, 2014). Jest to odmiana standardowego uniwersalnego narzędzia ETL, znanego pod nazwą *Pentaho Data Integration* (dawna nazwa to *Kettle*) dedykowana geoinformacji (rys. 11.10.). W aspekcie geoprzestrzennym GeoKettle korzysta również z wielu innych dojrzałych, solidnych i dobrze znanych narzędzi geoprzestrzennych, takich jak: WST, GeoTools, deegree, ogr, a także za pomocą interfejsu z Sextante. GeoKettle jest narzędziem ETL o dużych możliwościach, między innymi posługuje się metadanymi przestrzennymi i jest dedykowane integracji różnych źródłowych danych przestrzennych dla tworzenia i aktualizowania hurtowni danych. Zadania systemu GeoKettle jako narzędzia ETL sprowadzają się do:

- pobierania danych z plików lub baz źródłowych,
- ich przekształcania w celu skorygowania błędów, weryfikacji poprawności i usuwania błędów, transformacji w celu zmiany struktury danych tak, aby stały się zgodne z określonymi standardami,
- następnie ładowania przekształconych danych do docelowego systemu bazy danych (DBMS) w technologii OLTP lub SOLAP-OLAP, do pliku o określonym formacie



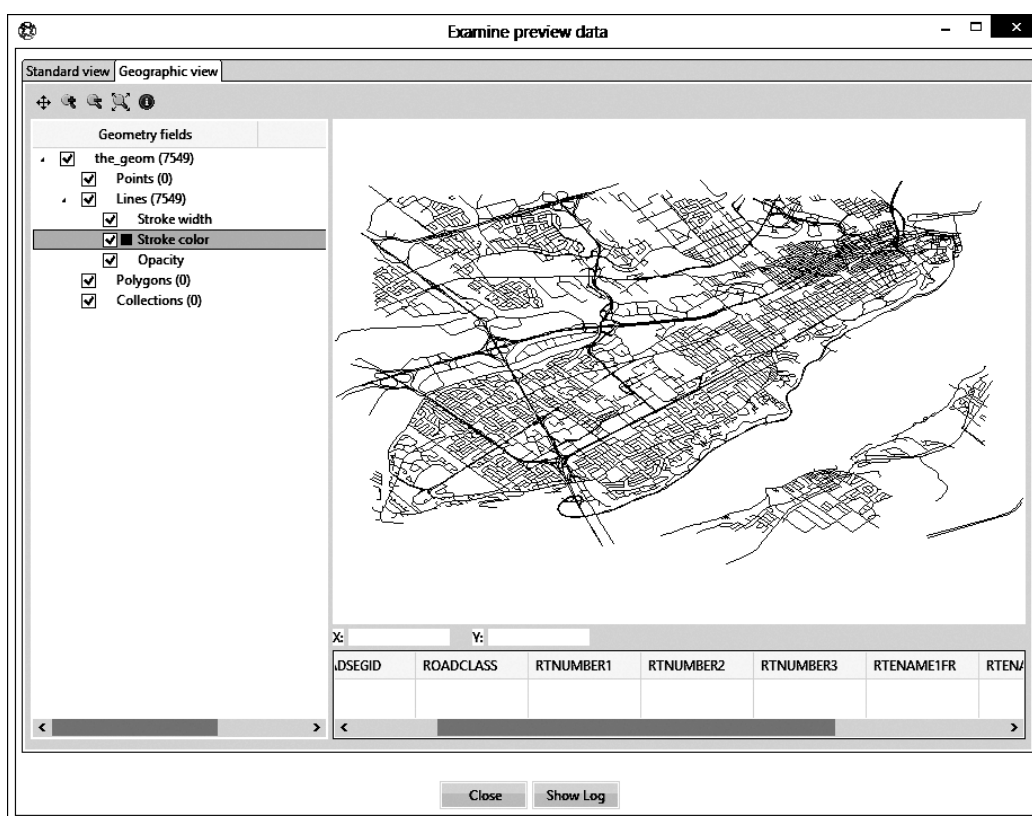
Rysunek 11.10. Okno interfejsu graficznego (GUI) oprogramowania narzędziowego GeoKettle typu ETL. Po stronie lewej – lista elementów (*steps* i *hops*) uczestniczących w procesie transformacji. Po prawej – graficzny obraz powiązań tych elementów.

geoprzestrzennym (na przykład GML) lub do udostępniania za pomocą GWS (*Geospatial Web Service*).

Transformacje danych w technologii XSLT mogą być w tym oprogramowaniu realizowane za pomocą procesora Xsalan lub procesora Saxon.

O mocy tego systemu i bogatych możliwościach operacji typu ETL świadczy liczba różnorodnych modułów, wynosząca w aktualnej wersji 148. Moduły te są przeznaczone do różnych zadań i są pogrupowane w następujące kategorie: operacje wejścia dla różnych źródeł danych – 30, operacje wyjścia dla różnych form zapisu – 19, transformacje – 25, funkcje pomocnicze – 8, operacje przepływu danych – 8, obsługa skryptów w różnych językach – 6, operacje przeszukiwania – 12, operacje łączenia danych – 5, dedykowane hurtowniom danych – 2, sprawdzanie poprawności (walidacja) – 4, definiowanie zadań do wykonania – 6, mapowanie – 5, funkcje operacji *in-line* – 3, funkcje eksperymentalne – 6 oraz masowe ładowanie danych do baz – 4.

Interfejs graficzny tego oprogramowania jest przystosowany do prezentacji graficznej danych zawierających geometrię w formie wektorowej (rys. 11.11.). Funkcje przeznaczone do operacji na danych geoprzestrzennych zawarte w oprogramowaniu GeoKettle są podobne do tych jakie zawiera moduł Geo w systemie BaseX. Funkcje te można podzielić na dwie grupy:



Rysunek 11.11. Graficzny obraz danych geoprzestrzennych przedstawiony w panelu *Geographic View* interfejsu graficznego oprogramowania narzędziowego GeoKettle.

- pierwsza grupa to funkcje obliczeniowe: wyznaczanie buforów, wyznaczanie centroidów, wyznaczanie losowych punktów wewnątrz obszaru, obliczanie powierzchni, obliczanie długości, obliczanie odległości, wyznaczanie części wspólnej, wyznaczanie unii, wyznaczanie prostokąta ograniczającego, wyznaczanie granicy, wyznaczanie otoczenia niewklęsłego, wyznaczanie różnicy logicznej, wyznaczanie symetrycznej różnicy logicznej, wyznaczanie geometrii odwrotnej,
- druga grupa to funkcje przetwarzania geoprzestrzennego:
 - ogólne: wycinanie, wycinanie za pomocą prostokąta, dzielenie (rozkładanie na elementy) wyróżnień złożonych,
 - dla punktów –: algorytm Delaunaya,
 - dla linii – zamiana linii złożonej na wielokąt, uproszczenie linii, wygładzenie linii, dzielenie (rozkładanie) linii złożonej do pojedynczych segmentów, dzielenie linii złożonej w węzłach, dzielenie linii złożonej w punktach,
 - dla wielokątów – uproszczenie wielokątów, usunięcie granic wewnętrznych, zamiana wielokąta na linię złożoną.

W pracach analitycznych i testowych system GeoKettle był wykorzystany w ograniczonym zakresie, ponieważ pomimo jego niewątpliwych zalet, nie jest ukierunkowany na rozwiązania technologiczne przyjęte w INSPIRE.

11.4. Oprogramowanie Stetl

Oprogramowanie Stetl znacznie się różni od trzech systemów opisanych wcześniej (FME, HALE i GeoKettle). Jest to zbiór skryptów języka Python i skryptów powłoki (*shell scripts*) systemu typu Unix (Nemeth, Snyder, Hein, Whaley, 2011; Sobell, 1995). Z tego względu Stetl może być używany wyłącznie w tym środowisku, na przykład w systemie Solaris lub Open Solaris, we wszelkich odmianach Linuxa (Hill, Harris, Vyas, 2006; Hudson, Hudson, 2008; Negus, 2005; Sobell, 2012), a także w systemach operacyjnych OS X w komputerach firmy Apple. Istnieje możliwość stworzenia takiego środowiska w systemach operacyjnych typu Windows firmy Microsoft za pomocą oprogramowania Cygwin, ale taka konfiguracja nie była testowana – z tego powodu nie ma gwarancji, że będzie poprawnie funkcjonowała. Efektywna praca z tym oprogramowaniem wymaga nie tylko dobrej znajomości systemów typu Unix, ale także wiedzy i umiejętności w posługiwaniu się i w pisaniu takich skryptów. Trzeba tu podkreślić, że Stetl nie ma interfejsu graficznego (GUI) tak jak to miały systemy opisane wcześniej, a sterowanie realizacją procesów odbywa się za pomocą konsoli tekstowej systemu operacyjnego, uzupełnionej edytorem tekstu dla języków skryptowych.

Stetl powstał w ramach prac nad projektem INSPIRE-FOSS, jego autorem jest holender Just van den Broecke (Van den Broecke, 2014). W latach następnych Stetl został przekształcony w narzędzie o szerszym zastosowaniu, głównie do transformacji do GML danych holenderskich, takich jak: Top10NL, IMGEO/BGT (wielkoskalowe mapy topograficzne) i IMKAD/BRK (dane katastralne) (Dutch Kadaster, 2011). Ostatnio obserwuje się duże zainteresowanie tym oprogramowaniem przez zespoły zajmujące się harmonizacją danych dla potrzeb INSPIRE w innych krajach europejskich. W tym celu zostało zorganizowane powszechnie dostępne repozytorium dla tego oprogramowania na serwerze GitHub (<https://github.com/justb4/stetl>).

Nazwa Stetl jest skrótem od „Streaming ETL”. Jest to narzędzie do transformacji danych geoprzestrzennych kategorii *Open Source*, które wykorzystuje inne, dalej wymienione narzędzia *Open Source* związane z procesem transformacji.

- Biblioteki **GDAL/OGR** – a w szczególności część wektorowa OGR dla zmiany różnych formatów danych wektorowych na zapis znacznikowy XML, w tym GML lub zmiany układu współrzędnych. Z bibliotekami tymi związane są programy dla przekształcania danych uruchamiane z linii komend, na przykład często tu stosowany program ogr2ogr bazujący na bibliotece OGR.
- Niezbędny w tych zagadnieniach procesor XSLT – większość systemów operacyjnych typu Unix ma wbudowany procesor XSLT uruchamiany z linii poleceń, na przykład program xsltproc.
- Programy narzędziowe deegree-Tools – a w szczególności program FSLoader (Feature Store Loader). Program ten umożliwia ładowanie plików z zapisami GML do modułu typu deegree Store i jest to często baza danych typu PostGIS.
- Programy narzędziowe i skrypty systemu operacyjnego typu Unix – programy narzędziowe są „sklejane” ze sobą przez wykorzystanie skryptów powłoki systemu Unix dla tworzenia większych modułów obliczeniowych. Często potrzebne jest użycie także kilku programów narzędziowych systemu Unix uruchamianych z wiersza poleceń. Są one wykorzystywane do różnych zadań, takich jak konwersja zestawu znaków (na przykład z kodu MS Windows CP1250 na Unicode UTF-8) za pomocą programu iconv, a także różnych przekształceń łańcuchów znaków w tekstach przy użyciu programu sed.

Dla zilustrowania działania oprogramowania Stetl poniżej przedstawiony jest bardzo prosty przykład transformacji danych zapisanych w prostym modelu XML do modelu opartego na języku GML (przykłady od 11.4. do 11.9.). Autorem tego przykładu jest Just van den Broecke – twórca tego oprogramowania, a cytowane tu pliki znajdują się pod adresem https://github.com/justb4/stetl/blob/master/examples/basics/2_xslt.

Przykład 11.4. Bardzo prosty skrypt języka shell uruchamiający procedurę ETL:

```
#!/bin/sh
# Shortcut to call Stetl main.py with etl config.
# Author: Just van den Broecke
stetl -c etl.cf
```

Przykład 11.5. Dane źródłowe opisujące lokalizacje trzech miast europejskich zapisane w XML o bardzo prostym modelu – nie jest to zapis GML:

```
<?xml version='1.0' encoding='utf-8'?>
<cities>
  <city>
    <name>Amsterdam</name>
    <lat>52.4</lat>
    <lon>4.9</lon>
  </city>
  <city>
    <name>Bonn</name>
    <lat>50.7</lat>
    <lon>7.1</lon>
  </city>
  <city>
    <name>Rome</name>
    <lat>41.9</lat>
    <lon>12.5</lon>
  </city>
</cities>
```

Przykład 11.6. Schemat XSD definiujący strukturę danych zapisu wynikowego w języku GML, który powstanie w wyniku transformacji:

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema targetNamespace="http://ogr.maptools.org/"
xmlns:ogr="http://ogr.maptools.org/"
xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:gml="http://www.opengis.net/gml"
elementFormDefault="qualified" version="1.0">
  <xs:import namespace="http://www.opengis.net/gml"
  schemaLocation=
    "http://schemas.opengis.net/gml/2.1.2/feature.xsd"/>
  <xs:element name="FeatureCollection"
    type="ogr:FeatureCollectionType"
    substitutionGroup="gml:_FeatureCollection"/>
  <xs:complexType name="FeatureCollectionType">
    <xs:complexContent>
      <xs:extension base="gml:AbstractFeatureCollectionType">
        <xs:attribute name="lockId"
          type="xs:string" use="optional"/>
        <xs:attribute name="scope"
          type="xs:string" use="optional"/>
      </xs:extension>
    </xs:complexContent>
  </xs:complexType>
  <xs:element name="city" type="ogr:City_Type"
    substitutionGroup="gml:_Feature"/>
  <xs:complexType name="City_Type">
    <xs:complexContent>
      <xs:extension base="gml:AbstractFeatureType">
        <xs:sequence>
          <xs:element name="name" nillable="false"
            minOccurs="1" maxOccurs="1">
            <xs:simpleType>
              <xs:restriction base="xs:string">
                <xs:maxLength value="42"/>
              </xs:restriction>
            </xs:simpleType>
          </xs:element>
          <xs:element name="geometry"
            type="gml:PointPropertyType" nillable="false"
            minOccurs="1" maxOccurs="1"/>
        </xs:sequence>
      </xs:extension>
    </xs:complexContent>
  </xs:complexType>
</xs:schema>
```

Przykład 11.7. Aby dane źródłowe w XML z przykładu 11.5. mogły być przetransformowane do zapisu zdefiniowanego w pliku XSD w przykładzie 11.6. potrzebne jest zastosowanie arkusza stylów XSL przedstawionego poniżej. Na podstawie tego arkusza został opracowany arkusz dla pomiarów zwierciadła wody podziemnej przedstawiony w rozdziale 9. (przykład 9.5.).

```

<?xml version="1.0" encoding="UTF-8"?>
<!-- Transform plain XML cities XML to valid GML.
Author: Just van den Broecke, Just Objects B.V. -->
<!-- Nazwa elementu root została zmieniona z "stylesheet"
na „transform” -->
<xsl:transform version="1.0"
  xmlns:ogr="http://ogr.maptools.org/"
  xmlns:gml="http://www.opengis.net/gml"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
  <xsl:output method="xml" omit-xml-declaration="no"
    indent="yes"/>
  <xsl:strip-space elements="*/>

  <xsl:template match="/">
    <ogr:FeatureCollection
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xmlns:ogr="http://ogr.maptools.org/"
      xmlns:gml="http://www.opengis.net/gml"
      xsi:schemaLocation="http://ogr.maptools.org/
        ../gmlcities.xsd http://www.opengis.net/gml
        http://schemas.opengis.net/gml/2.1.2/feature.xsd">
      <gml:boundedBy>
        <gml:Box>
          <gml:coord>
            <gml:X>-180.0</gml:X><gml:Y>-90.0</gml:Y>
          </gml:coord>
          <gml:coord>
            <gml:X>180.0</gml:X><gml:Y>90.0</gml:Y>
          </gml:coord>
        </gml:Box>
      </gml:boundedBy>
      <!-- Loop through all cities. -->
      <xsl:apply-templates/>
    </ogr:FeatureCollection>
  </xsl:template>

  <!-- Make each city an ogr:featureMember. -->
  <xsl:template match="city">
    <gml:featureMember>
      <ogr:city>
        <ogr:name>
          <xsl:value-of select="name"/>
        </ogr:name>
        <ogr:geometry>
          <gml:Point srsName="urn:ogc:def:crs:EPSG:4326">
            <gml:coordinates>
              <xsl:value-of select="lat"/>,
              <xsl:value-of select="lon"/>
            </gml:coordinates>
          </gml:Point>
        </ogr:geometry>
      </ogr:city>
    </gml:featureMember>
  </xsl:template>
</xsl:transform>
<!--
Nazwa elementu root została zmieniona z "stylesheet"
na „transform”
-->

```

Przykład 11.8. Plik konfiguracyjny „etl.cfg” oprogramowania Stetl określający poszczególne kroki transformacji i zastosowane w niej narzędzia i lokalizację plików, które mają być poddane przetwarzaniu:

```
# Transform input xml to valid GML file using an XSLT filter.

[etl]
chains = input_xml_file|transformer_xslt|output_file

[input_xml_file]
class = inputs.fileinput.XmlFileInput
file_path = input/cities.xml

[transformer_xslt]
class = filters.xsltfilter.XsltFilter
script = cities2gml.xsl

[output_file]
class = outputs.fileoutput.FileOutput
file_path = output/gmlcities.gml
```

Przykład 11.9. Plik wynikowy zawierający dane geoprzestrzenne opisujące lokalizację trzech miast zapisany w języku GML:

```
<?xml version='1.0' encoding='utf-8'?>
<ogr:FeatureCollection
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:ogr="http://ogr.maptools.org/"
xmlns:gml="http://www.opengis.net/gml"
xsi:schemaLocation="http://ogr.maptools.org/ ../gmlcities.xsd
http://www.opengis.net/gml
http://schemas.opengis.net/gml/2.1.2/feature.xsd">
  <gml:boundedBy>
    <gml:Box>
      <gml:coord>
        <gml:X>-180.0</gml:X>
        <gml:Y>-90.0</gml:Y>
      </gml:coord>
      <gml:coord>
        <gml:X>180.0</gml:X>
        <gml:Y>90.0</gml:Y>
      </gml:coord>
    </gml:Box>
  </gml:boundedBy>
  <gml:featureMember>
    <ogr:City>
      <ogr:name>Amsterdam</ogr:name>
      <ogr:geometry>
        <gml:Point srsName="urn:ogc:def:crs:EPSG:4326">
          <gml:coordinates>52.4,4.9</gml:coordinates>
        </gml:Point>
      </ogr:geometry>
    </ogr:City>
```

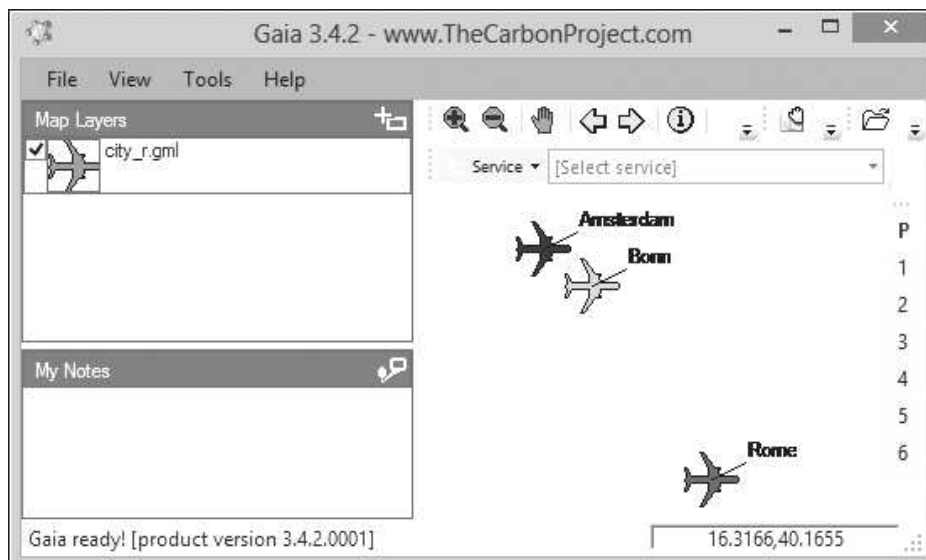


```

</gml:featureMember>
<gml:featureMember>
  <ogr:City>
    <ogr:name>Bonn</ogr:name>
    <ogr:geometry>
      <gml:Point srsName="urn:ogc:def:crs:EPSG:4326">
        <gml:coordinates>50.7,7.1</gml:coordinates>
      </gml:Point>
    </ogr:geometry>
  </ogr:City>
</gml:featureMember>
<gml:featureMember>
  <ogr:City>
    <ogr:name>Rome</ogr:name>
    <ogr:geometry>
      <gml:Point srsName="urn:ogc:def:crs:EPSG:4326">
        <gml:coordinates>41.9,12.5</gml:coordinates>
      </gml:Point>
    </ogr:geometry>
  </ogr:City>
</gml:featureMember>
</ogr:FeatureCollection>

```

Dla sprawdzenia poprawności uzyskanego wyniku zostało użyte oprogramowanie Gaia firmy Carbon Project przeznaczone do wizualizacji i weryfikacji (walidacji) poprawności zapisów GML. Oprogramowanie to jest często stosowane do tego celu w weryfikacji takich danych w różnych projektach prowadzonych przez *Open Geospatial Consortium*. Wynik pozytywnej weryfikacji jest przedstawiony na rysunku 11.12.



Rysunek 11.12. Graficzny obraz danych geoprzestrzennych w języku GML (przykład 11.9.) uzyskanych w wyniku transformacji wykonanej przy użyciu oprogramowania Stetl. Do weryfikacji poprawności zapisu użyto programu Gaia.

Analiza i przeprowadzone testy wykazały dużą przydatność tego oprogramowania w pracach dotyczących transformacji danych hydrogeologicznych do modeli INSPIRE. Pomimo braku interfejsu graficznego, oprogramowanie to ma istotną przewagę przy przetwarzaniu dużych ilości danych. Sterowanie pracą oprogramowania za pomocą skryptów systemu operacyjnego sprawia, że jest to nieporównywalnie szybsze niż praca programów sterowanych za pomocą GUI, szczególnie gdy są napisane w języku Java i wymagają użycia wirtualnej maszyny Java, a często wymagają także wsparcia środowiska programistycznego jakim jest system Eclipse. Jednak dla osiągnięcia dobrych wyników niezbędna jest znajomość systemu Unix, umiejętność pisania skryptów systemowych, a także znajomość języka Python.

11.5. Porównanie narzędzi do transformacji

Jest rzeczą zrozumiałą, że opisane cztery systemy przeznaczone do transformacji danych geoprzestrzennych znacznie się różnią, ponieważ powstały w różnych warunkach, dla różnych zadań i są rozwijane przez różne zespoły o różnych doświadczeniach. Główną linią podziału jest status oprogramowania – czy jest to oprogramowanie komercyjne, czy *Open Source*. Z tych czterech jedynie FME jest oprogramowaniem komercyjnym, a jego podstawy zostały opracowane gdy jeszcze nie stosowano do geoinformacji zapisu znacznikowego i nie było jeszcze języka GML. Rozwijaniem systemu FME zajmuje się duży zespół pracowników firmy Safe Software. Równie liczny zespół zajmuje się jego promocją i wsparciem dla użytkowników. Zupełnym przeciwieństwem do niego jest oprogramowanie Stetl jednego autora, w którym z założenia stosuje się technologię ETL z zastosowaniem procesora XSLT, jako centralnego elementu. Koncepcja ta powstała w momencie gdy język GML w wersji 3.2.1 już istniał i brakowało narzędzi pozwalających zapisywać dane w tym języku.

Dwa pozostałe systemy przeznaczone do transformacji lokują się pośrodku. GeoKettle jest systemem *Open Source*. Prace nad nim rozpoczęto w kanadyjskim francuskojęzycznym uniwersytecie „Université Laval” na bazie systemu Kettle, również *Open Source*, ale opracowanym i rozwijanym przez firmę Pentaho Corporation. Obecnie system Kettle ma inną nazwę – Pentaho Data Integration i stał się oprogramowaniem komercyjnym w wersji Enterprise Edition.

Czwarty z opisywanych systemów – HALE jest rezultatem europejskiego projektu badawczego zorientowanego na potrzeby implementacji dyrektywy INSPIRE. Jest rozwijany głównie przez niemieckie środowiska uniwersyteckie.

Każdy z tych czterech systemów ma swoje cenne właściwości, ale także często istotne ograniczenia. Pewnym mankamentem rozwiązań *Open Source* jest brak pełnej implementacji języka GML. Wynika to z posługiwania się przez te wszystkie programy tym samym silnikiem obsługi GML w postaci bibliotek GDAL/OGR, które obsługują GML głównie w wersji 2.1.2 lub w wersji 3.1.1 w zakresie SF-0. Zakres SF-0 oznacza profil *Simple Feature* na poziomie najniższym (ISO/TC211, 2004; 2007b). Aby takie zapisy mogły być przekształcone do tych, które są określone w specyfikacjach INSPIRE potrzebne są dodatkowe transformacje z zastosowaniem arkuszy XSL dedykowanych tym danym. W wielu przypadkach potrzebne jest transformowanie z zastosowaniem technologii *Full Text*, czyli przekształcanie łańcuchów znaków znajdujących się wewnątrz elementów XML lub w ich atrybutach.

Przeprowadzone analizy i testy wykazały, że żaden z tych programów narzędziowych nie jest w stanie samodzielnie wykonać pełnej transformacji danych źródłowych (na przykład płaskiego zapisu w formacie Esri Shape) do w pełni poprawnego w INSPIRE zapisu

w języku GML wersji 3.2.1. Z tego względu transformacja musi być podzielona przynajmniej na trzy fazy. Najbardziej odpowiednim systemem do wykonania zadań fazy pierwszej jest system FME. Dalsze przetwarzanie może być realizowane za pomocą edytora HALE lub skryptów Stetl. Jeżeli w końcowej fazie będzie niezbędne przetworzenie danych zapisanych w GML z wersji niższej do wersji 3.2.1, to najbardziej odpowiednim narzędziem jest Stetl, ale może być także użyty system FME.

11.6. Oprogramowanie narzędziowe wspomagające transformacje

Obok przedstawionych czterech systemów przeznaczonych głównie do transformacji, potrzebne są także inne narzędzia, między innymi do sprawdzania poprawności formalnej (walidacji) różnych zapisów w języku XML, w tym zapisów danych w różnych strukturach określonych za pomocą dokumentów XSD, samych dokumentów XSD bo też są zapisane z zastosowaniem XML, zapisów danych w języku GML i definiujących je schematów aplikacyjnych w XSD, a także szablonów XSL sterujących procesem transformacji XSLT. Część tych zapisów, w szczególności zapisy w językach XSD i XSL są pisane przez programistę lub przynajmniej konieczne jest wprowadzenie różnych korekt i modyfikacji. Poprawne wykonanie tej pracy jest możliwe tylko za pomocą wyspecjalizowanych edytorów języka XML.

W przeprowadzonych pracach testowych wykorzystano dwa komercyjne edytory: *XML Spy* firmy Altova i *Oxygen XML Editor* firmy SyncRO Soft. Nie są to narzędzia specjalnie dedykowane geoinformacji, bo są przeznaczone do zastosowań ogólnych dotyczących języka XML, czyli są użyteczne także we wszystkich zastosowaniach dziedzinowych tego języka, w tym także do geoinformacji. Oba programy narzędziowe spełniają swoje zadanie poprawnie, w szczególności do analizowania wyników transformacji, jak również do weryfikacji ich poprawności syntaktycznej i semantycznej na bazie dokumentów XSD.

Drugą kategorią oprogramowania wspomagającego proces transformacji jest oprogramowanie dedykowane specjalnie danym geoprzestrzennym różnego typu i w różnej postaci. Do zadań tej kategorii oprogramowania należy przechowywanie i obsługa danych zapisanych w innej formie niż XML, udostępnianie tych danych do środowiska programowego transformacji, sprawdzanie geoprzestrzennej poprawności wyników końcowych, a także zapisów przejściowych pomiędzy poszczególnymi etapami transformacji. W przeprowadzonych pracach testowych do tych zadań wykorzystano głównie oprogramowanie kategorii *Open Source*, a w szczególności systemy geoinformatyczne GRASS GIS i QGIS. Do prezentacji graficznej w celu oceny poprawności geometrii wykorzystane zostały programy: GML Viewer firmy Snowflake Software, Gaia Viewer firmy Carbon Project, a także wielu innych, w tym programy udostępniane bezpłatnie przez firmy komercyjne, na przykład Geomedia Viewer lub instytucje państwowe, na przykład FalconView. Nie można tu polecić konkretnych produktów ponieważ dobór narzędzi wspomagających prace nad transformacją zależy w dużym stopniu od specyfiki przetwarzanych danych.

12. Transformacja danych hydrogeologicznych

W rozdziałach od 4. do 7. przedstawione zostało otoczenie problematyki transformacji krajowych danych hydrogeologicznych, a w szczególności przegląd danych krajowych i ich modeli pojęciowych, a także ontologiczne, semantyczne i strukturalne różnice pomiędzy danymi krajowymi i modelami określonymi w INSPIRE. W rozdziałach od 8. do 11. poruszane były zagadnienia dotyczące metodyki i technologii transformacji w ujęciu ogólnym, także z uwzględnieniem wymagań specyfikacji INSPIRE. Ten rozdział jest poświęcony przede wszystkim samej transformacji dziedzinowych danych hydrogeologicznych i problemom technologicznym i metodycznym, jakie w tym przypadku występują.

Pod pojęciem transformacji zbiorów danych rozumie się tu przekształcenie tych danych z postaci obecnej – bardzo różnorodnej i niespełniającej przyjętych w tym zakresie standardów i norm – do postaci standardowej określonej w specyfikacjach INSPIRE. Specyfikacje te odwołują się między innymi do standardów OGC i norm ISO grupy 19100, będących też polskimi normami przyjętymi przez PKN. Jako krajowe zbiory geoprzestrzennych danych hydrogeologicznych rozumie się tu wszelkie dane, które mają bezpośrednie lub pośrednie odniesienie do określonych miejsc na Ziemi i dotyczą wód podziemnych. Praktycznie wszystkie dane hydrogeologiczne, podobnie jak w innych naukach o Ziemi, mają takie odniesienie.

Zgodnie z tym, co jest przedstawione w rozdziale 2. (rys. 2.1., 2.2. i 2.3.), aby krajowe dane hydrogeologiczne po transformacji spełniały całkowicie wymagania specyfikacji INSPIRE muszą być powiązane z danymi z innych tematów:

- temat Geologia (podstawowa) – przez powiązanie dziedziczenia pomiędzy `GeologicUnit` i `HydrogeologicalUnit`, dzięki któremu dziedziczony jest atrybut geometryczny `MappedFeature`. `HydrogeologicalUnit` powinien być także powiązany asocjacją z `GeologicStructure`, a `ActiveWell` również powiązana przez asocjację z `Borehole`,
- temat Urządzenia do monitorowania środowiska – `ActiveWell` pełniąc zadanie urządzenia obserwacyjnego, a także obserwowane `GroundWaterBody` powinny być powiązane asocjacyjnie z `EnvironmentalMonitoringFacility`,
- temat Gospodarowanie obszarem/strefy ograniczone/regulacyjne oraz jednostki sprawozdawcze – `GroundWaterBody` podlegając sprawozdawczości wynikającej z Ramowej Dyrektywy Wodnej powinien być dwukierunkowo powiązany asocjacyjnie z `WFDGroundWaterBody` należącego do tamtego tematu.

Cel ten jednak nie mógł być osiągnięty ponieważ wymaga to istnienia odpowiednich zbiorów danych z zakresu tych tematów zawierających elementy, do których można by utworzyć powiązania. Jedyny wyjątek stanowi wyróżnienie `GeologicUnit` dzięki powiązaniu przez dziedziczenie. Korzystny jest fakt, że prawie wszystkie powiązania asocjacyjne w modelu tematu Hydrogeologia mają stereotyp «voidable», co oznacza, że nie są obowiązujące i w efekcie nie muszą występować w zbiorach danych.

Analizowana i testowana metodyka transformacji danych hydrogeologicznych jest wieloetapowa.

- Pierwszy etap polegający na transformacji wszystkich elementów zbioru danych z formatu źródłowego do zapisu znacznikowego XML z zachowaniem źródłowego modelu danych. W tym etapie zmienia się układ odniesienia na EPSG:2180, kodowanie na UTF-8, aby w dalszej kolejności móc łączyć elementy danych z różnych zbiorów. Dostępne do testowania zbiory danych hydrogeologicznych prawie zawsze były w formacie Esri Shape i z tego powodu ich modele były zawsze płaskie. Głównym narzędziem, jakie było na tym etapie stosowane to FME (*Feature Manipulation Engine*).
- Kolejną operacją jest poprawianie znaków diakrytycznych w danych tekstowych i nazwach atrybutów za pomocą tabel odnośników (*lookup tables*), a także zamiana nazw atrybutów na zgodne z przyjętymi zasadami w GML i jednocześnie zrozumiałymi. W zapisach XML, w tym także w GML, dla nazw stosuje się notację „*camelCase*” (lub „*UpperCamelCase*”) polegającą na łączeniu wyrazów pisanych jak nazwy własne, na przykład „toJestNazwaAtrybutu” lub „AToJestNazwaKlasy”. Operacje te nie mogą najczęściej być wykonane w pełni automatycznie i potrzebna jest w takim przypadku żmudna praca edycyjna. Tego typu prace były wykonywane za pomocą edytorów XML Spy i Oxygen.
- Jeżeli elementy danych dotyczące tych samych wyróżnień (*features*) znajdują się w różnych zbiorach źródłowych, to potrzebne jest ich połączenie w zbiorze wynikowym. Można to wykonać automatycznie, gdy w zbiorach źródłowych występują takie same identyfikatory wyróżnień. W przeciwnym przypadku konieczne jest uzupełnienie jednolitych identyfikatorów na drodze pracochłonnej edycji. Uzyskany tą drogą zapis danych w wersji roboczej po wstępnej transformacji z formatu Esri Shape do języka GML co jest przedstawione w przykładzie 12.1.
- Kolejny etap to zmiana modelu danych ze źródłowego na model docelowy (dla danych wynikowych), którym w tym przypadku jest schemat aplikacyjny *HydrogeologyCore.xsd*. To zadanie może być zrealizowane zarówno przez oprogramowanie FME, jak i przez HALE, jednak najbardziej zaawansowanym podejściem jest transformacja oparta na XSLT. Wymaga to utworzenia odpowiedniego dokumentu XML (XSL), w którym zapisane są reguły transformacji i użycia wybranego procesora XSLT. Oba programy FME i HALE mają takie procesory wbudowane, jednak w przypadku FME nie ma możliwości za jego pomocą utworzenia dokumentu XSL, można jedynie dołączyć go z zewnątrz. W pracach testowych wykorzystano edytor HALE do utworzenia arkusza XSL, a następnie wykonano wiele modyfikacji edytorem XML Spy. Utworzony w ten sposób arkusz XSL z głównym elementem `xsl:transform` i z zapisem reguł transformacji danych uzyskanych ze wstępnej transformacji do postaci w znacznym stopniu zgodnej z wymaganiami INSPIRE jest przedstawiony w przykładzie 12.2.
- Kolejną operacją jest utworzenie powiązań (odsylaczy) zewnętrznych za pomocą elementu `xlink:href` do repozytoriów. Dotyczy to między innymi list kodowych i wyliczeń (enumeratorów), definicji układów odniesienia i siatek, a także definicji jednostek miar, które są umieszczone w zdalnych repozytoriach internetowych. Operację tę można przeprowadzić automatycznie za pomocą tabel odnośników, jednak pod warunkiem, że w zbiorach źródłowych zachowany był porządek i konsekwencja w określaniu tych elementów w sposób niestandardowy. Na przykład gdy zapis „m n. p. m.” nie jest stosowany konsekwentnie wymaga najczęściej ręcznego podziału na dwa atrybuty *UoM (UnitOfMeasure)*: „m” i *baseFrame*: „aboveSeaLevel”, które z kolei zostaną zamienione na odsylacze `xlink:href` do odpowiednich pozycji w

repozytoriach. Jeżeli brakuje konsekwencji w tych zapisach źródłowych, to również w tym przypadku konieczne jest edytowanie ręczne. Utworzenie odpowiednich repozytoriów dla jednostek miar jest dopiero w planach i z tego powodu w pracach testowych jednostki miary były określane tradycyjnie. Uproszczony i z konieczności skrócony zapis danych w wersji w znacznym stopniu zgodnej z wymaganiami INSPIRE jest przedstawiony w przykładzie 12.3.

- Utworzenie odsyłaczy wewnętrznych (w obrębie danych hydrogeologicznych) `xlink:href` w ramach poszczególnych zbiorów nie stwarza istotnych problemów, ponieważ decyduje o tym tylko jeden schemat aplikacyjny (*HydrogeologyCore.xsd*). Jednak, gdy zbiory danych w takim zapisie są rozproszone po różnych bazach, odsyłacze muszą to uwzględniać, co powoduje dodatkowe trudności.
- Podobny problem występuje w przypadku powiązań zewnętrznych pomiędzy różnymi tematami INSPIRE. Jest to zjawisko pozytywne i potrzebne, jednak stwarza problemy, głównie natury organizacyjnej, ponieważ dotyczy zbiorów danych w różnych instytucjach z różnych resortów. W pracach testowych nie było możliwości utworzenia odsyłaczy do zewnętrznych zbiorów danych ponieważ nie są one jeszcze utworzone lub publicznie dostępne.
- Oddzielnym trudnym zagadnieniem jest poprawne zapisanie tych elementów i atrybutów, których typami są wyliczenia (enumeratory) i listy kodowe. W zbiorach źródłowych występują one jako wartości tekstowe, a poprawne zapisy zgodne ze specyfikacjami INSPIRE wymagają stosowania w takich przypadkach URI (*Uniform Resource Identifier*) w formie URL (*Uniform Resource Locator*). Problematyka transformacji takich elementów nie jest jeszcze do końca rozwiązana – uwagi na ten temat są przedstawione w rozdziale 13. (Podsumowanie).
- Ostatnim etapem jest kompletne i szczegółowe sprawdzenie utworzonych nowych zapisów. W tym przypadku na szczęście, całkowita automatyzacja w zakresie atrybutów nieprzestrzennych jest w pełni możliwa. Atrybuty geometryczne dotyczące kształtu i położenia powinny być jednak dodatkowo zweryfikowane wzrokowo na ekranie komputera. Do tego celu wykorzystano wiele przeglądarek danych geoprzestrzennych. Przeglądarki te są opisane w rozdziale 11.6.

12.1. Dane wykorzystane do testowania

Dane hydrogeologiczne tworzone głównie przez Państwową Służbę Hydrogeologiczną (PSH) są własnością Skarbu Państwa. Dostęp do nich jest ograniczony wieloma warunkami, a ponadto są one przechowywane w bazach danych PSH, a bazy te są zarządzane systemami komercyjnymi, głównie firmy Oracle. Z tych, a także z wielu innych powodów, zestaw danych przeznaczony do prac analitycznych i testowych był bardzo ograniczony. Szczegółowej analizie poddane zostały następujące kategorie danych:

- Mapa Hydrogeologiczna Polski w skali 1:50 000 – głównego poziomu wodonośnego (MHP), próbki danych zapisane w formacie Esri Shape i Geomedia MDB,
- Mapa Hydrogeologiczna Polski w skali 1:50 000 – pierwszy poziom wodonośny – występowanie i hydrodynamika (PPW-WH), próbki danych zapisane w formacie Esri Shape i Geomedia MDB,
- Mapa Hydrogeologiczna Polski w skali 1:50 000 – pierwszy poziom wodonośny – wrażliwość na zanieczyszczenie i jakość wód (PPW-WJ), próbki danych zapisane w formacie Esri Shape i Geomedia MDB.

- Centralna Baza Danych Hydrogeologicznych (dawniej Bank HYDRO) – dane o otworach hydrogeologicznych, próbki danych zapisane w formacie Microsoft XLS i DBF,
- Baza danych o poborach wód podziemnych – archiwalne wartości rocznych poborów, próbki danych zapisane tabelarycznie,
- Mapa Głównych Zbiorników Wód Podziemnych w formie cyfrowej, dane zapisane w formacie Esri Shape,
- SOH – System Obserwacji Hydrogeologicznych, próbki danych zapisane tabelarycznie,
- Jednolite Części Wód Podziemnych, dane zapisane w formacie Esri Shape.

Dla potrzeb testowania przygotowano próbki danych zawierające najwięcej elementów kandydujących do zbiorów INSPIRE i jednocześnie mających największą różnorodność tych elementów. Przetestowane szczegółowo dane hydrogeologiczne:

- wybrane zbiory danych Mapy Głównych Zbiorników Wód Podziemnych w formie cyfrowej, dane zapisane w formacie Esri Shape,
- wybrane zbiory danych dotyczące Jednolitych Części Wód Podziemnych, dane zapisane w formacie ESRI Shape.

W rozdziałach 12.2., 12.3. i 12.4 przedstawione są przykłady zapisów XML (a ściślej GML i XSL) jakie powstawały w kolejnych fazach procesu transformacji. Są to przykłady bardzo skrócone i często pozbawione elementów, które nie są istotne dla opisywanej tu problematyki. Jednak skróty i uproszczenia są tu niezbędne, ponieważ zapisy w pełnej postaci są tak obszerne, że nie nadają się do publikacji.

12.2. Przekształcenie do płaskiego zapisu znacznikowego

Przedstawione powyżej dane, głównie dotyczące GZWP (Głównych Zbiorników Wód Podziemnych) i JCWPd (Jednolitych Części Wód Podziemnych), po szczegółowej analizie zostały przekształcone za pomocą oprogramowania FME do formy zapisu znacznikowego. Dane te mają strukturę pierwotnych danych krajowych i są przygotowane do transformacji do modelu INSPIRE. Przykład zapisu danych w wersji roboczej po wstępnym przekształceniu z formatu Esri Shape do języka GML przedstawia przykład 12.1.

Przykład 12.1. Zapis danych w wersji roboczej po wstępnej transformacji z formatu Esri Shape do języka GML po zmianie kodowania tekstów na UTF-8 i przeliczeniu współrzędnych do układu EPSG:2180:

```
<!-- Początek zapisu danych. Zapis jest zgodny z GML ale nie
      jest zgodny w modelem schematu aplikacyjnego INSPIRE dla
      tematu Hydrogeologia. -->
<?xml version="1.0" encoding="UTF-8"?>
<gml:FeatureCollection
  xmlns:gsr="http://www.isotc211.org/2005/gsr"
  xmlns:gts="http://www.isotc211.org/2005/gts"
  xmlns:gss="http://www.isotc211.org/2005/gss"
  <!-- Definicja przestrzeni nazw schematu roboczego -->
  xmlns:hg1="http://www.geomlab.pl/hg1"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:gco="http://www.isotc211.org/2005/gco"
  xmlns:gmd="http://www.isotc211.org/2005/gmd"
  xmlns:xlink="http://www.w3.org/1999/xlink"
```

```

<!-- Definicja przestrzeni nazw języka GML 3.2.1 -->
xmlns:gml="http://www.opengis.net/gml/3.2"
gml:id="ide86394de-6af2-48b0-b942-8c0256780e23">
<gml:boundedBy>
  <gml:Envelope srsName="EPSG:2180" srsDimension="2">
    <gml:lowerCorner>239699.9 440507.8</gml:lowerCorner>
    <gml:upperCorner>550698.0 707693.8</gml:upperCorner>
  </gml:Envelope>
</gml:boundedBy>
<!--
Początek elementu z danymi dla Zbiornik Koluszki-Tomaszów
-->
  <gml:featureMember>
    <hg1:Aquifer_hg1
      gml:id="idc1a9095e-dd91-4120-b362-0a27d8506113">
      <hg1:id>311</hg1:id>
      <hg1:nrGZWP>404</hg1:nrGZWP>
      <hg1:nazwa>Zbiornik Koluszki-Tomaszów</hg1:nazwa>
      <hg1:powierzchnia>1109</hg1:powierzchnia>
      <hg1:powierzchniaUoM>km2</hg1:powierzchniaUoM>
      <hg1:stanUdokumentowania>nieudokumentowany
        </hg1:stanUdokumentowania>
      <hg1:rokUdokumentowania>0</hg1:rokUdokumentowania>
      <hg1:tytułDokumentacji></hg1:tytułDokumentacji>
      <hg1:rokReambulacji>0</hg1:rokReambulacji>
      <hg1:tytułReambulacji></hg1:tytułReambulacji>
      <hg1:stratygrafia>J3</hg1:stratygrafia>
      <hg1:zakresGłębokościOd>0</hg1:zakresGłębokościOd>
      <hg1:zakresGłębokościDo>0</hg1:zakresGłębokościDo>
      <hg1:głębokośćŚrednia>200</hg1:głębokośćŚrednia>
      <hg1:głębokośćUoM>m</hg1:głębokośćUoM>
      <hg1:typOśrodka>krasowo-szczelinowy</hg1:typOśrodka>
      <hg1:typOśrodkaEN>karst-fractured</hg1:typOśrodkaEN>
      <hg1:stanUdokumentowaniaEN>not documented
        </hg1:stanUdokumentowaniaEN>
    <gml:LinearRing>
      <!-- Lista współrzędnych punktów opisujących
        granicę jest znacznie skrócona -->
      <gml:posList>
        457178.8 559462.1 (...) 457178.8 559462.1
      </gml:posList>
    </gml:LinearRing>
  </hg1:Aquifer_hg1>
</gml:featureMember>
<!--
Koniec elementu z danymi dla Zbiornik Koluszki-Tomaszów
-->
<!--
Początek elementu z danymi Niecki Miechowskiej (NW)
-->
  <gml:featureMember>
    <hg1:Aquifer_hg1
      gml:id="id5471675e-384b-4d9d-9bd7-40b35c9e3fa8">
      <hg1:id>219</hg1:id>
      <hg1:nrGZWP>408</hg1:nrGZWP>

```



```

<hgl:nazwa>Niecka Miechowska (NW)</hgl:nazwa>
<hgl:powierzchnia>3200.44</hgl:powierzchnia>
<hgl:powierzchniaUoM>km2</hgl:powierzchniaUoM>
<hgl:stanUdokumentowania>udokumentowany
  </hgl:stanUdokumentowania>
<hgl:rokUdokumentowania>1999</hgl:rokUdokumentowania>
<hgl:tytułDokumentacji>Dokumentacja hydrogeologiczna
  Głównego Zbiornika Wód Podziemnych (GZWP) - nr 408
  Niecka Miechowska (część NW)</hgl:tytułDokumentacji>
<hgl:rokReambulacji>2011</hgl:rokReambulacji>
<hgl:tytułReambulacji>Dodatek do dokumentacji
  hydrogeologicznej Głównego Zbiornika Wód Podziemny
  (GZWP) - nr 408 Niecka Miechowska (część NW)
  </hgl:tytułReambulacji>
<hgl:stratygrafia>T</hgl:stratygrafia>
<hgl:zakresGłębokościOd>0</hgl:zakresGłębokościOd>
<hgl:zakresGłębokościDo>200</hgl:zakresGłębokościDo>
<hgl:głębokośćŚrednia>20</hgl:głębokośćŚrednia>
<hgl:głębokośćUoM>m</hgl:głębokośćUoM>
<hgl:typOśrodka>krasowo-szczelinowy</hgl:typOśrodka>
<hgl:typOśrodkaEN>karst-fractured</hgl:typOśrodkaEN>
<hgl:stanUdokumentowaniaEN>dokumented
</hgl:stanUdokumentowaniaEN>
<gml:LinearRing>
  <!-- Lista współrzędnych punktów opisujących
    granicę jest znacznie skrócona -->
  <gml:posList>
    415455.5 480621.7 (...) 415455.5 480621.7
  </gml:posList>
</gml:LinearRing>
</hgl:Aquifer_hgl>
</gml:featureMember>
<!--
  Koniec elementu z danymi Niecki Miechowskiej (NW)
-->
<!-- Tu powinny być kolejne elementy z danymi o innych GZWP
</gml:FeatureCollection>
<!-- Koniec zapisu danych -->

```

12.3. Testowe arkusze XSLT

W ramach prac testowych wygenerowano lub opracowano wiele arkuszy XSL, głównie za pomocą edytora HALE, a także edytora XML_Spy z późniejszymi modyfikacjami, w celu ich dostosowania do specyfiki danych źródłowych. Przykład arkusza XSL z głównym elementem `xsl:transform` i z zapisem reguł `xsl:template` transformacji danych uzyskanych ze wstępnej transformacji do postaci zgodnej z wymaganiami INSPIRE jest przedstawiony w przykładzie 12.2. Zawarte w tym arkuszu szablony tworzą element `ge:Mapped Feature` w drzewie wynikowym dla każdego elementu `hgl:Aquifer_hgl` w drzewie źródłowym, a następnie wypełniają te elementy w drzewie wynikowym przetransformowanymi elementami i atrybutami odpowiedniego elementu `hgl:Aquifer_hgl` z drzewa źródłowego. Arkusz ten służył do testowania transformacji danych dotyczących GZWP (Głównych Zbiorników Wód Podziemnych).

Przykład 12.2. Arkusz XSL z głównym elementem `xsl:transform` i z zapisem reguł `xsl:template` transformacji danych uzyskanych ze wstępnej transformacji do postaci w znacznym stopniu zgodnej z wymaganiami INSPIRE:

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:transform
  <!--
    Deklaracje przestrzeni nazw dla danych INSPIRE i transformacji
  -->
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:def="http://www.esdi-humboldt.eu/hale/xslt/definitions"
  xmlns:gmd="http://www.isotc211.org/2005/gmd"
  xmlns:ge="http://inspire.ec.europa.eu/schemas/ge-core/3.0rc3"
  xmlns:swe="http://www.opengis.net/swe/2.0"
  xmlns:gts="http://www.isotc211.org/2005/gts"
  xmlns:sam="http://www.opengis.net/sampling/2.0"
  xmlns:om="http://www.opengis.net/om/2.0"
  xmlns:hg1="http://www.geomlab.pl/hg1"
  xmlns:sams="http://www.opengis.net/samplingSpatial/2.0"
  xmlns:sr="http://inspire.ec.europa.eu/schemas/sr/3.0rc3"
  xmlns:sc="http://www.interactive-instruments.de/ShapeChange/AppInfo"
  xmlns:gsr="http://www.isotc211.org/2005/gsr"
  xmlns:gss="http://www.isotc211.org/2005/gss"
  xmlns:ge_hg="http://inspire.ec.europa.eu/schemas/ge_hg/3.0rc3"
  xmlns:omop="http://inspire.ec.europa.eu/schemas/omop/2.0rc3"
  xmlns:omso="http://inspire.ec.europa.eu/schemas/omso/2.0rc3"
  xmlns:hfp="http://www.w3.org/2001/XMLSchema-hasFacetAndProperty"
  xmlns:wml2="http://www.opengis.net/waterml/2.0"
  xmlns:wfd="http://inspire.ec.europa.eu/schemas/wfd/3.0rc3"
  xmlns:omor="http://inspire.ec.europa.eu/schemas/omor/2.0rc3"
  xmlns:base2="http://inspire.ec.europa.eu/schemas/base2/1.0rc3"
  xmlns:gmlcov="http://www.opengis.net/gmlcov/1.0"
  xmlns:gco="http://www.isotc211.org/2005/gco"
  xmlns:ns2="urn:x-inspire:specification:gmlas:BaseTypes:3.2"
  xmlns:ef="http://inspire.ec.europa.eu/schemas/ef/3.0rc3"
  xmlns:ns3="urn:x-inspire:specification:gmlas:WaterFrameworkDirective:0.0"
  xmlns:am="http://inspire.ec.europa.eu/schemas/am/3.0rc3"
  xmlns:base="http://inspire.ec.europa.eu/schemas/base/3.3rc3/"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:gml="http://www.opengis.net/gml/3.2"

  <!-- Określenie lokalizacji schematu dla tego arkusza transformacji
    z przestrzenią nazw "xsl" -->
  xsi:schemaLocation="http://www.w3.org/1999/XSL/Transform
  http://www.w3.org/2007/schema-for-xslt20.xsd">

  <!-- Określenie formatu dokumentu wynikowego -->
  <xsl:output method="xml" encoding="UTF-8" indent="yes"/>

  <!-- Arkusz XSL do transformacji XSLT opracowany na podstawie
    wyników uzyskanych z edytora HALE -->
  <!-- Nazwa projektu: GZWP_2_INSPIRE -->
```

```

<!-- Autor projektu: Janusz Michalak -->
<!-- Opis projektu: Transformacja testowa danych dotyczących GZWP
      (Głównych Zbiorników Wód Podziemnych) do schematu INSPIRE
      'HydrogeologyCore' - wyróżnienie 'Aquifer'. Geometria tego
      wyróżnienia zgodnie z wymaganiami specyfikacji jest określona
      przez element 'MappedFeature' -->

<!-- Szablon ten tworzy element główny (root - "/") -->
<xsl:template match="/">
  <xsl:variable name="MappedFeatureType">
    <!-- Transformuj elementy do drzewa tymczasowego -->
    <xsl:call-template name="MappedFeatureType"/>
  </xsl:variable>
  <!-- Umieszczaj przetransformowane elementy w drzewie
        wynikowym -->
  <gml:FeatureCollection
    gml:id="_cfc8b671-8cb2-4f30-a6b0-2694ee59a1cb">
    <xsl:for-each select="$MappedFeatureType/*">
      <gml:featureMember>
        <xsl:copy-of select="."/>
      </gml:featureMember>
    </xsl:for-each>
  </gml:FeatureCollection>
</xsl:template>

<!-- Szablon ten tworzy element 'MappedFeature' w drzewie
      wynikowym dla każdego elementu 'Aquifer_hg1' w drzewie
      źródłowym -->
<xsl:template name="MappedFeatureType">
  <xsl:for-each select="//hg1:Aquifer_hg1">
    <ge:MappedFeature>
      <xsl:call-template name="MappedFeatureType_inline"/>
    </ge:MappedFeature>
  </xsl:for-each>
</xsl:template>

<!-- Szablon ten wypełnia element 'MappedFeature' w drzewie
      wynikowym przetransformowanymi elementami i atrybutami
      odpowiedniego elementu 'Aquifer_hg1' z drzewa źródłowego
      -->
<xsl:template name="MappedFeatureType_inline">
  <xsl:attribute name="gml:id"
    namespace="http://www.opengis.net/gml/3.2">
    <xsl:value-of select="concat(generate-id(),
      '_e2e4861e-e990-4225-a63c-94d9a2c2d276')"/>
  </xsl:attribute>
  <xsl:for-each select="gml:surfaceProperty">
    <xsl:variable name="children">
      <children>
        <xsl:variable name="children">
          <children>
            <xsl:for-each select="gml:Surface/@srsDimension">
              <xsl:variable name="results">
                <result>

```

```

<!-- Dla każdej wartości źródłowej atrybutu
'srsDimension' w elemencie 'Surface' wpisz
taką samą wartość do atrybutu 'srsDimension'
w elemencie 'Surface' w drzewie wynikowym.
W razie potrzeby stosuje się konwersję. -->
<xsl:choose>
  <xsl:when test=".">
    <!-- Jeżeli jest wartość źródłowa,
    to należy ją przekopiować -->
    <xsl:value-of select="."/>
  </xsl:when>
  <xsl:otherwise>
    <!-- Gdy niema wartości źródłowej
    to niema kopiowania -->
    <def:null/>
  </xsl:otherwise>
</xsl:choose>
</result>
</xsl:variable>
<xsl:if test="$results/result[not(def:null)
and (node() or @* or . = '')]">
  <xsl:attribute name="srsDimension">
    <xsl:if test="$results/result[not(def:null)
and (node() or @* or . = '')]">
      <!-- Kopiuj zawartość wyniku z najwyższym
      priorytetem -->
      <xsl:variable name="firstNonNullResult"
        select="$results/result[not(def:null)
and (node() or @* or . = '')][1]">
      <xsl:copy-of select="$firstNonNullResult/@*,
        $firstNonNullResult/child::node()"/>
    </xsl:if>
  </xsl:attribute>
</xsl:if>
</xsl:for-each>
<xsl:for-each select="gml:Surface/@srsName">
  <xsl:variable name="results">
    <result>
<!-- Dla każdej wartości źródłowej atrybutu 'srsName'
w elemencie 'Surface' wpisz taką samą wartość do
atrybutu 'srsName' w elemencie 'Surface' w drzewie
wynikowym. W razie potrzeby stosuje się konwersję.
-->
    <xsl:choose>
      <xsl:when test=".">
        <!-- Jeżeli jest wartość źródłowa,
        to należy ją przekopiować -->
        <xsl:value-of select="."/>
      </xsl:when>
      <xsl:otherwise>
        <!-- Gdy niema wartości źródłowej
        to niema kopiowania -->
        <def:null/>
      </xsl:otherwise>
    </xsl:choose>
  </result>
</xsl:for-each>

```

```

</xsl:choose>
</result>
</xsl:variable>
<xsl:if test="$results/result[not(def:null) and
(node() or @* or . = ')]">
  <xsl:attribute name="srsName">
    <xsl:if test="$results/result[not(def:null) and
(node() or @* or . = ')]">
      <!-- Kopiuj zawartość wyniku z najwyższym
priorytetem -->
      <xsl:variable name="firstNonNullResult"
select="$results/result[not(def:null) and
(node() or @* or . = ')] [1]"/>
      <xsl:copy-of select="$firstNonNullResult/@*,
$firstNonNullResult/child::node()"/>
    </xsl:if>
  </xsl:attribute>
</xsl:if>
</xsl:for-each>
<xsl:for-each select="gml:Surface/@gml:id">
  <xsl:variable name="results">
    <result>
      <!-- Dla każdej wartości źródłowej elementu 'id'
w elemencie ' Surface ' dodaje taką samą
wartość do elementu 'id' w elemencie 'Surface'
w drzewie wynikowym. W razie potrzeby stosuje
się konwersję.
-->
      -->
      <xsl:choose>
        <xsl:when test=".">
          <!-- Jeżeli jest wartość źródłowa,
to należy ją przekopiować -->
          <xsl:value-of select="."/>
        </xsl:when>
        <xsl:otherwise>
          <!-- Gdy niema wartości źródłowej
to niema kopiowania -->
          <def:null/>
        </xsl:otherwise>
      </xsl:choose>
    </result>
  </xsl:variable>
  <xsl:if test="$results/result[not(def:null) and
(node() or @* or . = ')]">
    <xsl:attribute name="gml:id"
namespace="http://www.opengis.net/gml/3.2">
      <xsl:if test="$results/result[not(def:null) and
(node() or @* or . = ')]">
        <!-- Kopiuj zawartość wyniku z najwyższym
priorytetem -->
        <xsl:variable name="firstNonNullResult"
select="$results/result[not(def:null) and
(node() or @* or . = ')] [1]"/>
        <xsl:copy-of select="$firstNonNullResult/@*,
$firstNonNullResult/child::node()"/>

```

```

        </xsl:if>
        </xsl:attribute>
    </xsl:if>
</xsl:for-each>
</children>
</xsl:variable>
<xsl:variable name="results">
    <result>
<!-- Dla każdej zawartości źródłowej elementu 'Surface'
przekopiuj taką samą zawartość do elementu 'Surface'
w drzewie wynikowym. W razie potrzeby stosuje się
konwersję. -->
    <xsl:choose>
        <xsl:when test="gml:Surface">
            <!-- Jeżeli jest wartość źródłowa,
to należy ją przekopiarować -->
            <xsl:value-of select="gml:Surface"/>
        </xsl:when>
        <xsl:otherwise>
            <!-- Gdy nie ma wartości źródłowej
to nie ma kopiowania -->
            <def:null/>
        </xsl:otherwise>
    </xsl:choose>
</result>
</xsl:variable>
<xsl:if test="$children/children[node() or @*] or
    $results/result[not(def:null) and
        (node() or @* or . = '')]">
    <xsl:element name="gml:Surface"
        namespace="http://www.opengis.net/gml/3.2">
    <xsl:if test="$children/children[node() or @*]">
<!-- Kopiuj elementy podrzędne (wewnętrzne - children) -->
        <xsl:copy-of select="$children/children/@*,
            $children/children/*"/>
    </xsl:if>
    <xsl:if test="$results/result[not(def:null) and
        (node() or @* or . = '')]">
<!-- Kopiuj zawartość wyniku z najwyższym priorytetem -->
        <xsl:variable name="firstNonNullResult"
            select="$results/result[not(def:null) and
                (node() or @* or . = '')][1]">
        <xsl:copy-of select="$firstNonNullResult/@*,
            $firstNonNullResult/child::node()"/>
    </xsl:if>
    </xsl:element>
</xsl:if>
</children>
</xsl:variable>
<xsl:if test="$children/children[node() or @*]">
    <xsl:element name="ge:shape" namespace=
        "http://inspire.ec.europa.eu/schemas/ge-core/3.0rc3">
    <xsl:if test="$children/children[node() or @*]">
<!-- Kopiuj elementy podrzędne (wewnętrzne - children) -->

```

```

        <xsl:copy-of select="$children/children/@*,
            $children/children/*"/>
    </xsl:if>
</xsl:element>
</xsl:if>
</xsl:for-each>
<xsl:variable name="children">
    <children>
        <xsl:variable name="children">
            <children>
                <xsl:for-each select="@gml:id">
                    <xsl:variable name="results">
                        <result>
                            <!-- Dla każdej wartości źródłowej elementu 'id'
                                w elemencie 'Surface' dodaje taką samą wartość
                                do elementu 'id' w elemencie 'shape' w drzewie
                                wynikowym. W razie potrzeby stosuje się
                                konwersję. -->
                            <xsl:choose>
                                <xsl:when test=".">
                                    <!-- Jeżeli jest wartość źródłowa,
                                        to należy ją przekopiować -->
                                    <xsl:value-of select="."/>
                                </xsl:when>
                                <xsl:otherwise>
                                    <!-- Gdy niema wartości źródłowej
                                        to niema kopiowania -->
                                    <def:null/>
                                </xsl:otherwise>
                            </xsl:choose>
                        </result>
                    </xsl:variable>
                    <xsl:if test="$results/result[not(def:null) and
                        (node() or @* or . = '')]">
                        <xsl:attribute name="gml:id"
                            namespace="http://www.opengis.net/gml/3.2">
                            <xsl:if test="$results/result[not(def:null) and
                                (node() or @* or . = '')]">
                                <!-- Kopiuj zawartość wyniku z najwyższym
                                    priorytetem -->
                                <xsl:variable name="firstNonNullResult"
                                    select="$results/result[not(def:null) and
                                        (node() or @* or . = '')][1]">
                                <xsl:copy-of select="$firstNonNullResult/@*,
                                    $firstNonNullResult/child::node()"/>
                                </xsl:if>
                            </xsl:attribute>
                        </xsl:if>
                    </xsl:for-each>
                    <gname xsi:nil="true"/>
                    <geologicHistory xsi:nil="true"/>
                    <gcomposition xsi:nil="true"/>
                    <ge_hgdescription xsi:nil="true"/>
                    <ge_hgapproximateDepth xsi:nil="true"/>
                </xsl:for-each>
            </children>
        </xsl:variable>
    </children>
</xsl:variable>

```

```

<ge_hgapproximateThickness xsi:nil="true"/>
<xsl:for-each select="hg1:typOśrodka">
  <xsl:variable name="children">
    <children/>
  </xsl:variable>
  <xsl:variable name="results">
    <result>
      <!-- Dla każdej wartości źródłowej 'typOśrodka' wpisuje
           się taką samą wartość do 'aquiferType' w drzewie
           wynikowym. W razie potrzeby stosuje się konwersję.
      -->
      <xsl:choose>
        <xsl:when test=".">
          <!-- Jeżeli jest wartość źródłowa,
                to należy ją przekopiować -->
          <xsl:value-of select="."/>
        </xsl:when>
        <xsl:otherwise>
          <!-- Gdy niema wartości źródłowej
                to niema kopiowania -->
          <def:null/>
        </xsl:otherwise>
      </xsl:choose>
    </result>
  </xsl:variable>
  <xsl:if test="$children/children[node() or @*] or
    $results/result[not(def:null) and (node() or
    @* or . = '')]">
    <xsl:element name="ge_hg:aquiferType" namespace=
      "http://inspire.ec.europa.eu/schemas/ge_hg/3.0rc3">
      <xsl:if test="$children/children[node() or @*]">
        <!-- Kopiuj elementy podrzędne
              (wewnętrzne - children) -->
        <xsl:copy-of select="$children/children/@*,
          $children/children/*"/>
      </xsl:if>
      <xsl:if test="$results/result[not(def:null) and
        (node() or @* or . = '')]">
        <!-- Kopiuj zawartość wyniku
              z najwyższym priorytetem -->
        <xsl:variable name="firstNonNullResult"
          select="$results/result[not(def:null) and
            (node() or @* or . = '')][1]"/>
        <xsl:copy-of select="$firstNonNullResult/@*,
          $firstNonNullResult/child::node()"/>
      </xsl:if>
    </xsl:element>
  </xsl:if>
</xsl:for-each>
<ge_hgisExploited xsi:nil="true"/>
<ge_hgisMainInSystem xsi:nil="true"/>
<ge_hgvulnerabilityToPollution xsi:nil="true"/>
<ge_hgpermeabilityCoefficient xsi:nil="true"/>
<ge_hgstorativityCoefficient xsi:nil="true"/>

```



```

        <ge_hghydroGeochemicalRockType xsi:nil="true"/>
    </children>
</xsl:variable>
<xsl:if test="$children/children[node() or @*]">
    <xsl:element name="ge_hg:Aquifer" namespace=
        "http://inspire.ec.europa.eu/schemas/ge_hg/3.0rc3">
        <xsl:if test="$children/children[node() or @*]">
            <!-- Kopiaj elementy podrzędne
                 (wewnętrzne - children) -->
            <xsl:copy-of select="$children/children/@*,
                $children/children/*"/>
        </xsl:if>
    </xsl:element>
</xsl:if>
</children>
</xsl:variable>
<xsl:if test="$children/children[node() or @*]">
    <xsl:element name="ge:specification" namespace=
        "http://inspire.ec.europa.eu/schemas/ge-core/3.0rc3">
        <xsl:if test="$children/children[node() or @*]">
            <!-- Kopiaj elementy podrzędne (wewnętrzne - children) -->
            <xsl:copy-of
                select="$children/children/@*, $children/children/*"/>
        </xsl:if>
    </xsl:element>
</xsl:if>
</xsl:template>
</xsl:transform>
<!-- Koniec schematu XSL -->

```

12.4. Analiza uzyskanych wyników

Zastosowane do transformacji oprogramowanie narzędziowe, pozwoliło na generowanie plików zapisu GML o dużej zgodności z wymaganiami specyfikacji danych tematu Hydrogeologia. Jednak nie udało się uzyskać pełnej zgodności i potrzebne było ręczne poprawianie arkuszy XSL, a także ręczne poprawianie samych zapisów wynikowych. Głównie dotyczyło to identyfikatorów GML i INSPIRE, ale także odsyłaczy do zewnętrznych repozytoriów list kodowych i innych elementów niezbędnych do uzyskania pełnej zgodności. Ręczne poprawianie arkuszy XSL jest metodą całkowicie poprawną i opłacalną, ponieważ taki dopracowany arkusz może być następnie wielokrotnie wykorzystywany dla całej grupy zbiorów danych o takiej samej strukturze i rodzaju treści. Jednak ręczne poprawianie zbiorów wynikowych jest dopuszczalne jedynie w pracach testowych, a w zastosowaniach masowych musi być zastąpione procedurą całkowicie zautomatyzowaną. Rozwiązanie tego problemu jest rozpatrywane w rozdziale 13. (Podsumowanie). Przykład zapisu danych w wersji w zgodnej z wymaganiami INSPIRE po niewielkich ręcznych modyfikacjach jest zawarty w przykładzie 12.3. Należy zwrócić uwagę, że wartości atrybutów, których typem jest lista kodowa (lub enumerator), są wyrażone w postaci odwołań do określonych pozycji tych list znajdujących się w zdalnym repozytorium. Odwołania takie występują w tym przykładzie i dotyczą określenia wieku geologicznego (elementy `ge:olderNamedAge` i `ge:youngerNamed`

Age) i rodzaju ośrodka wodonośnego (element `ge_hg:mediaType`). Odpowiedź, jaką przesyła repozytorium list kodowych INSPIRE, gdy otrzyma polecenie zapisanie w elemencie `ge_hg:mediaType`, gdy ma to być środowisko krasowo-szczelinowe przedstawia przykład 12.7.

Przykład 12.3. Zapis danych w wersji w zgodnej z wymaganiami INSPIRE (zapis ten zawiera konieczne skróty i uproszczenia):

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Początek zapisu danych. Zapis jest zgodny z GML 3.2.1. i
      jest zgodny w modelem schematu aplikacyjnego INSPIRE dla
      tematu Hydrogeologia. -->
<gml:FeatureCollection
  xmlns:ns2="urn:x-inspire:specification:gmlas:BaseTypes:3.2"
  xmlns:base="http://inspire.ec.europa.eu/schemas/base/3.3rc3/"
  xmlns:gco="http://www.isotc211.org/2005/gco"
  xmlns:wfd="http://inspire.ec.europa.eu/schemas/wfd/3.0rc3"
  xmlns:hfp="http://www.w3.org/2001/XMLSchema-hasFacetAndProperty"
  xmlns:ge="http://inspire.ec.europa.eu/schemas/ge-core/3.0rc3"
<!-- Definicja przestrzeni nazw schematu roboczego -->
  xmlns:ge_hg="http://inspire.ec.europa.eu/schemas/ge_hg/3.0rc3"
  xmlns:base2="http://inspire.ec.europa.eu/schemas/base2/1.0rc3"
  xmlns:ge_gp="http://inspire.ec.europa.eu/schemas/ge_gp/3.0rc3"
  xmlns:sr="http://inspire.ec.europa.eu/schemas/sr/3.0rc3"
<!-- Definicja przestrzeni nazw języka GML 3.2.1 -->
  xmlns:gml="http://www.opengis.net/gml/3.2"
  xmlns:sam="http://www.opengis.net/sampling/2.0"
  xmlns:om="http://www.opengis.net/om/2.0"
  xmlns:omop="http://inspire.ec.europa.eu/schemas/omop/2.0rc3"
  xmlns:omso="http://inspire.ec.europa.eu/schemas/omso/2.0rc3"
  xmlns:gts="http://www.isotc211.org/2005/gts"
  xmlns:sams="http://www.opengis.net/samplingSpatial/2.0"
  xmlns:omor="http://inspire.ec.europa.eu/schemas/omor/2.0rc3"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:ef="http://inspire.ec.europa.eu/schemas/ef/3.0rc3"
  xmlns:wml2="http://www.opengis.net/waterml/2.0"
  xmlns:am="http://inspire.ec.europa.eu/schemas/am/3.0rc3"
  xmlns:gmlcov="http://www.opengis.net/gmlcov/1.0"
  xmlns:sc="http://www.interactive-instruments.de/ShCh/AppInfo"
  xmlns:gss="http://www.isotc211.org/2005/gss"
  xmlns:gssr="http://www.isotc211.org/2005/gssr"
  xmlns:swe="http://www.opengis.net/swe/2.0"
  xmlns:gmd="http://www.isotc211.org/2005/gmd"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  gml:id="ID_PL_000"
  xsi:schemaLocation=
    "http://inspire.ec.europa.eu/schemas/ge_hg/3.0rc3
    file://C:/local-schemas/inspire/ge_hg/HydrogeologyCore.xsd">
  <gml:boundedBy>
    <gml:Envelope srsName=
      "http://www.opengis.net/def/crs/EPSS/0/2180"
      srsDimension="2">
```



```

<!-- Odwołanie z elementu „ge: youngerNamedAge” przy pomocy
„xlink:href” do pozycji listy kodowej w odległym repozytorium -->
  <ge:youngerNamedAge
    xlink:href="http://inspire.ec.europa.eu/
    codelist/GeochronologicEraValue/lateJurassic/
    lateJurassic.pl.iso19135.xml"/>
  </ge:GeologicEvent>
</ge:geologicHistory>
<ge:geologicUnitType/>
<ge:composition>
  <ge:CompositionPart>
    <ge:material/>
    <ge:proportion xsi:nil="true"/>
    <ge:role/>
  </ge:CompositionPart>
</ge:composition>
<ge_hg:description>
  <gmd:PT_FreeText>
    <gmd:textGroup>
      <gmd:LocalisedCharacterString
        locale="http://www.ethnologue.com/language/pol">
        [hg_pl:nrGZWP] 404 [/hg_pl:nrGZWP]
        [hg_pl:powierzchnia] 1109 [/hg_pl:powierzchnia]
        [hg_pl:powierzchniaUoM] km2 [/hg_pl:powierzchniaUoM]
        [hg_pl:stanUdokumentowania] nieudokumentowany
        [/hg_pl:stanUdokumentowania]
        [hg_pl:rokUdokumentowania/]
        [hg_pl:tytułDokumentacji/]
        [hg_pl:rokReambulacji/]
        [hg_pl:tytułReambulacji/]
      </gmd:LocalisedCharacterString>
      <gmd:LocalisedCharacterString
        locale="http://www.ethnologue.com/language/eng">
        [hg_pl:typOśrodkaEN] karst-fractured
        [/hg_pl:typOśrodkaEN]
        [hg_pl:stanUdokumentowaniaEN] not documented
        [/hg_pl:stanUdokumentowaniaEN]
      </gmd:LocalisedCharacterString>
    </gmd:textGroup>
  </gmd:PT_FreeText>
</ge_hg:description>
<ge_hg:approximateDepth>
  <ge_hg:QuantityValue>
    <ge_hg:singleQuantity uom="m">200
  </ge_hg:singleQuantity>
</ge_hg:QuantityValue>
</ge_hg:approximateDepth>
<ge_hg:approximateThickness>
  <ge_hg:QuantityValue>
    <ge_hg:quantityInterval/>
  </ge_hg:QuantityValue>
</ge_hg:approximateThickness>
<ge_hg:beginLifespanVersion/>

```

```

    <ge_hg:aquiferType/>
<!-- Odwołanie z elementu „ge_hg:mediaType” przy pomocy „xlink:href”
do pozycji listy kodowej w odległym repozytorium -->
    <ge_hg:mediaType xlink:href="http://inspire.ec.europa.eu/
    codelist/AquiferMediaTypeValue/karsticAndFractured/
    karsticAndFractured.pl.iso19135xml"/>
    <ge_hg:isExploited>true</ge_hg:isExploited>
    <ge_hg:isMainInSystem>true</ge_hg:isMainInSystem>
    <ge_hg:vulnerabilityToPollution>
        <ge_hg:QuantityValue>
            <ge_hg:singleQuantity/>
        </ge_hg:QuantityValue>
    </ge_hg:vulnerabilityToPollution>
    <ge_hg:permeabilityCoefficient>
        <ge_hg:QuantityValue>
            <ge_hg:singleQuantity/>
        </ge_hg:QuantityValue>
    </ge_hg:permeabilityCoefficient>
    <ge_hg:storativityCoefficient>
        <ge_hg:QuantityValue>
            <ge_hg:quantityInterval/>
        </ge_hg:QuantityValue>
    </ge_hg:storativityCoefficient>
    <ge_hg:hydroGeochemicalRockType/>
</ge_hg:Aquifer>
</ge:specification>
</ge:MappedFeature>
</gml:featureMember>
<!--
Koniec elementu z danymi dla Zbiornik Koluszki-Tomaszów
-->
<!--
Początek elementu z danymi Niecki Miechowskiej (NW)
-->
    <gml:featureMember>
        <ge:MappedFeature>
            <gml:description/>
            <gml:descriptionReference/>
            <gml:identifier>IDENT_PL_002</gml:identifier>
            <gml:name/>
            <ge:shape>
                <gml:Surface gml:id="ID_PL_002_1" srsName=
                "http://www.opengis.net/def/crs/EPSSG/0/2180"
                srsDimension="2">
                    <gml:patches>
                        <gml:PolygonPatch>
                            <gml:exterior>
                                <gml:LinearRing>
                                    <!-- Lista współrzędnych punktów opisujących
                                    granicę jest znacznie skrócona -->
                                    <gml:posList>
                                        415455.5 480621.7 (...) 415455.5 480621.7
                                    </gml:posList>
                                </gml:LinearRing>
                            </gml:exterior>
                        </gml:PolygonPatch>
                    </gml:patches>
                </gml:Surface>
            </ge:shape>
        </ge:MappedFeature>
    </gml:featureMember>

```

```

        </gml:LinearRing>
    </gml:exterior>
    </gml:PolygonPatch>
</gml:patches>
</gml:Surface>
</ge:shape>
<ge:mappingFrame/>
<ge:specification>
    <ge_hg:Aquifer gml:id="ID_PL_002_2">
        <ge:inspireId>
            <base:Identifier>
                <base:localId>219</base:localId>
                <base:namespace>PL_PSH</base:namespace>
                <base:versionId>1.0.</base:versionId>
            </base:Identifier>
        </ge:inspireId>
        <ge:name>Niecka Miechowska (NW)</ge:name>
        <ge:geologicHistory>
            <ge:GeologicEvent gml:id=" ID_PL_002_3">
                <ge:eventProcess>creation</ge:eventProcess>
                <!-- Odwołanie z elementu „ge:olderNamedAge” przy pomocy „xlink:href”
                do pozycji listy kodowej w odległym repozytorium -->
                <ge:olderNamedAge
                    xlink:href="http://inspire.ec.europa.eu/
                    codelist/GeochronologicEraValue/triassic/
                    triassic.pl.iso19135xml"/>
                <!-- Odwołanie z elementu „ge:youngerNamedAge” przy pomocy
                „xlink:href” do pozycji listy kodowej w odległym repozytorium -->
                <ge:youngerNamedAge
                    xlink:href="http://inspire.ec.europa.eu/
                    codelist/GeochronologicEraValue/triassic/
                    triassic.pl.iso19135xml"/>
            </ge:GeologicEvent>
        </ge:geologicHistory>
        <ge:geologicUnitType/>
        <ge:composition>
            <ge:CompositionPart>
                <ge:material/>
                <ge:proportion xsi:nil="true"/>
                <ge:role/>
            </ge:CompositionPart>
        </ge:composition>
        <ge_hg:description>
            <gmd:PT_FreeText>
                <gmd:textGroup>
                    <gmd:LocalisedCharacterString
                        locale="http://www.ethnologue.com/language/pol">
                        [hg_pl:nrGZWP] 408 [/hg_pl:nrGZWP]
                        [hg_pl:powierzchnia] 3200.44 [/hg_pl:powierzchnia]
                        [hg_pl:powierzchniaUoM] km2 [/hg_pl:powierzchniaUoM]
                        [hg_pl:stanUdokumentowania] udokumentowany
                        [/hg_pl:stanUdokumentowania]
                        [hg_pl:rokUdokumentowania] 1999
                    </gmd:LocalisedCharacterString>
                </gmd:textGroup>
            </gmd:PT_FreeText>
        </ge_hg:description>
    </ge:Aquifer>
</ge:specification>
</ge:mappingFrame>
</gml:Surface>
</gml:PolygonPatch>
</gml:exterior>
</gml:LinearRing>

```

```

[/hg_pl:rokUdokumentowania]
[hg_pl:tytułDokumentacji]Dokumentacja hydro-
geologiczna Głównego Zbiornika Wód Podziemnych
(GZWP) - nr 408 Niecka Miechowska (część NW)
[/hg_pl:tytułDokumentacji]
[hg_pl:rokReambulacji]2011[/hg_pl:rokReambulacji]
[hg_pl:tytułReambulacji]Dodatek do dokumentacji
hydrogeologicznej Głównego Zbiornika Wód
Podziemnych (GZWP) - nr 408 Niecka Miechowska
(część NW)
[/hg_pl:tytułReambulacji]
</gmd:LocalisedCharacterString>
<gmd:LocalisedCharacterString
  locale="http://www.ethnologue.com/language/eng">
  [hg_pl:typOśrodkaEN]karst-fractured
  [/hg_pl:typOśrodkaEN]
  [hg_pl:stanUdokumentowaniaEN]dokumented
  [/hg_pl:stanUdokumentowaniaEN]
</gmd:LocalisedCharacterString>
</gmd:textGroup>
</gmd:PT_FreeText>
</ge_hg:description>
<ge_hg:approximateDepth>
  <ge_hg:QuantityValue>
    <ge_hg:singleQuantity uom="m">20
  </ge_hg:singleQuantity>
</ge_hg:QuantityValue>
</ge_hg:approximateDepth>
<ge_hg:approximateThickness>
  <ge_hg:QuantityValue>
    <ge_hg:quantityInterval uom="m">0,200
  </ge_hg:quantityInterval>
</ge_hg:QuantityValue>
</ge_hg:approximateThickness>
<ge_hg:beginLifespanVersion/>
<ge_hg:aquiferType/>
<!-- Odwołanie z elementu „ge_hg:mediaType” przy pomocy „xlink:href”
do pozycji listy kodowej w odległym repozytorium -->
<ge_hg:mediaType xlink:href="http://inspire.ec.europa.eu/
  codelist/AquiferMediaTypeValue/karsticAndFractured/
  karsticAndFractured.pl.iso19135xml"/>
<ge_hg:isExploited>true</ge_hg:isExploited>
<ge_hg:isMainInSystem>true</ge_hg:isMainInSystem>
<ge_hg:vulnerabilityToPollution>
  <ge_hg:QuantityValue>
    <ge_hg:singleQuantity uom="degee">2.5
  </ge_hg:singleQuantity>
</ge_hg:QuantityValue>
</ge_hg:vulnerabilityToPollution>
<ge_hg:permeabilityCoefficient>
  <ge_hg:QuantityValue>
    <ge_hg:singleQuantity uom="m/d">20
  </ge_hg:singleQuantity>

```

```

        </ge_hg:QuantityValue>
    </ge_hg:permeabilityCoefficient>
    <ge_hg:storativityCoefficient>
        <ge_hg:QuantityValue>
            <ge_hg:quantityInterval uom="m3/m3">0.12
        </ge_hg:quantityInterval>
        </ge_hg:QuantityValue>
    </ge_hg:storativityCoefficient>
    <ge_hg:hydroGeochemicalRockType/>
</ge_hg:Aquifer>
</ge:specification>
</ge:MappedFeature>
</gml:featureMember>
<!--
Koniec elementu z danymi Niecki Miechowskiej (NW)
-->
<!-- Tu powinny być kolejne elementy z danymi o innych GZWP -->
<gml:FeatureCollection>
<!-- Koniec zapisu danych -->

```

12.5. Główne problemy transformacji danych hydrogeologicznych

Przykład przedstawiony w rozdziale 12.4. spełnia wymagania specyfikacji danych tematu Hydrogeologia tylko dlatego, że dokonano w nim kilka niewielkich poprawek ręcznych. Do nich można zaliczyć różnice w zestawie danych pomiędzy danymi krajowymi i strukturami danych INSPIRE. Na przykład dla danych dotyczących GZWP:

- brak jest w modelu INSPIRE miejsca na numery lub identyfikatory (krajowe i europejskie) dla poszczególnych GZWP, na dane o dokumentacji jednostki hydrogeologicznej GZWP i na dane o wielkości zajmowanej powierzchni,
- informacja o stratygrafii utworów jednostki jest zupełnie innego typu w modelu geologicznym – występuje w elemencie `ge:geologicHistory` i jest odsyłaczem do określonej pozycji w repozytorium, w danych krajowych jest to symbol danego okresu geologicznego.

Dla danych dotyczących JCWPd brak w modelu INSPIRE miejsca na numery lub identyfikatory krajowe i kody UE dla poszczególnych JCWPd, na dane o wielkości zajmowanej powierzchni i nazwę związanego z nią dorzecza. Jeżeli w krajowych danych hydrogeologicznych są takie składniki, dla których nie ma odpowiednich elementów w schematach INSPIRE i w języku GML, to racjonalnym podejściem jest umieszczenie ich w elemencie tekstowym o zdefiniowanej jednoznacznie strukturze i odpowiednim takim elementem jest element `ge_hg:description` zdefiniowany w elemencie `ge_hg:HydrogeologicalUnit`. Dla zachowania struktury tych danych może to być zapis zbliżony do znacznikowego, na przykład z użyciem nawiasów kwadratowych, które w zapisach XML nie mają specjalnego znaczenia (przykład 12.4.). Taki sposób umieszczenia danych został zastosowany także w zbiorze danych przedstawionym w przykładzie 12.3.

Przykład 12.4. Przykład użycia bloku danych w elemencie `ge_hg:description` dla zapisu danych krajowych, dla których nie ma miejsca w strukturze danych INSPIRE. Blok ten ma strukturę zapisaną zgodnie z wymaganiami języka XML, ale zamiana nawiasów ostrych („<”, „>”) na kwadratowe („[”, „]”) sprawia, że jest traktowany jako zwykły tekst.

```

<!--
Element ge_hg:description jest zdefiniowany w schemacie INSPIRE
i może zawierać dowolny tekst w dowolnym języku. Pozwala to na
umieszczenie w tym elemencie obok opisu także dowolnych innych
danych w formie struktury zbliżonej do XML.
-->

<ge_hg:description>
  Tu można umieścić opis jednostki hydrogeologicznej,
  a także dane dotyczące tej jednostki, dla których nie ma
  miejsca w schemacie INSPIRE. Przykład takich danych jest
  umieszczony poniżej i zastosowano notacją zbliżoną do XML.

<!--Blok danych dotyczących tej jednostki hydrogeologicznej-->
[hg_pl:id]219[/hg_pl:id]
[hg_pl:nrGZWP]408[/hg_pl:nrGZWP]
[hg_pl:nazwa]Niecka Miechowska (NW)[/hg_pl:nazwa]
[hg_pl:powierzchnia]3200.44[/hg_pl:powierzchnia]
[hg_pl:powierzchniaUoM]km2[/hg_pl:powierzchniaUoM]
[hg_pl:stanUdokumentowania]udokumentowany
[/hg_pl:stanUdokumentowania]
[hg_pl:rokUdokumentowania]1999[/hg_pl:rokUdokumentowania]
[hg_pl:tytułDokumentacji]Dokumentacja hydrogeologiczna
  Głównego Zbiornika Wód Podziemnych (GZWP) - nr 408
  Niecka Miechowska (część NW)[/hg_pl:tytułDokumentacji]
[hg_pl:rokReambulacji]2011[/hg_pl:rokReambulacji]
[hg_pl:tytułReambulacji]Dodatek do dokumentacji
  hydrogeologicznej Głównego Zbiornika Wód Podziemnych
  (GZWP) - nr 408 Niecka Miechowska (część NW)
[/hg_pl:tytułReambulacji]
[hg_pl:stratygrafia]T[/hg_pl:stratygrafia]
[hg_pl:zakresGłębokościOd]0[/hg_pl:zakresGłębokościOd]
[hg_pl:zakresGłębokościDo]200[/hg_pl:zakresGłębokościDo]
[hg_pl:głębokośćŚrednia]20[/hg_pl:głębokośćŚrednia]
[hg_pl:głębokośćUoM]m[/hg_pl:głębokośćUoM]
[hg_pl:typOśrodka]krasowo-szczelinowy[/hg_pl:typOśrodka]
[hg_pl:typOśrodkaEN]karst-fractured[/hg_pl:typOśrodkaEN]
[g_pl:stanUdokumentowaniaEN]dokumented
[/hg_pl:stanUdokumentowaniaEN]
</ge_hg:description>

```

Jak wcześniej przedstawiono w rozdziale 6.2., europejska infrastruktura geoinformacyjna INSPIRE jest wielkim przedsięwzięciem międzynarodowym i dotyczy informacji geoprzestrzennej, jednak często występują tam dane tekstowe. W Europie jest obecnie w użyciu ponad 205 języków w 46 krajach (Michalak, 2004). Z tego względu ogromne znaczenie ma wielojęzyczność udostępnianej geoinformacji. Niestety, pomimo formalnego doceniania tego problemu, w praktyce nie jest to realizowane. Powszechną regułą jest zapisywanie informa-

cji tekstowych przy użyciu najprostszego typu `CharacterString` bez możliwości określenia w jakim języku jest ten tekst. W modelu danych INSPIRE tematu Hydrogeologia problem ten jest rozwiązany zgodnie z obowiązującymi w tym zakresie standardami. Wszystkie typy wyróżnień pochodzących od `HydrogeologicalUnit` i od `HydrogeologicalObject` mają atrybut `description` typu `PT_FreeText`, który pozwala na dowolne stosowanie tekstów w różnych językach. Zastosowanie tego typu jest zilustrowane w przykładzie 12.5. Jest to jednak wyjątek w skali wszystkich tematów – występujące już w modelu hydrogeologicznym elementy z innych tematów (na przykład nadrzędne klasy `GeologicFeature` i `GeologicUnit`) posługują się wyłącznie typem `CharacterString`.

Przykład 12.5. Zastosowanie elementu `ge_hg:description`, który pozwala na umieszczenie w nim dowolnego tekstu w dowolnym języku:

```
<!-- Ponieważ znaki "<" i ">" mają w zapisie XML specjalne
znaczenie i nie mogą wystąpić w tekstach. Istnieje możliwość
zastąpienia ich przez "&lt;" i "&gt;", co zostało wykorzystane
poniżej -->
<ge_hg:description>
  <gmd:PT_FreeText>
    <gmd:textGroup>
      <gmd:LocalisedCharacterString
        locale="http://www.ethnologue.com/language/pol">
          Tu można umieścić dowolny tekst w języku polskim
          ale bez znaków „&lt;" i „&gt;"
      </gmd:LocalisedCharacterString>
      <gmd:LocalisedCharacterString
        locale="http://www.ethnologue.com/language/eng">
          Here you can put any text in English
          but without the characters "&lt;" and "&gt;"
      </gmd:LocalisedCharacterString>
      <gmd:LocalisedCharacterString
        locale="http://www.ethnologue.com/language/deu">
          Hier können Sie einen beliebigen Text in Deutsch
          gestellt aber ohne die Zeichen "&lt;" und "&gt;"
      </gmd:LocalisedCharacterString>
      <gmd:LocalisedCharacterString
        locale="http://www.ethnologue.com/language/fra">
          Ici vous pouvez mettre n'importe quel texte
          en français mais sans les caractères "&lt;" et "&gt;"
      </gmd:LocalisedCharacterString>
    </gmd:textGroup>
  </gmd:PT_FreeText>
</ge_hg:description>
```

Metodyki i technologie dotyczące infrastruktur geoinformacyjnych są w ciągłym szybkim rozwoju. Często to, co było dwa lub trzy lata temu nowym zalecanym rozwiązaniem, obecnie jest już uważane za nieaktualne (*deprecated* – przestarzałe). Z tego względu ujmowanie zmieniających się szybko rozwiązań technologicznych w sztywne i trwałe ramy aktów prawnych, jakimi są rozporządzenia Komisji Europejskiej, nie sprzyja rozwojowi. Przykładem jest określanie systemu odniesienia przestrzennego w zapisach danych przestrzennych.

- Dawniej wystarczyło podać „układ 92” lub „1992” i było to wystarczające.
- Język GML wymaga jednak ujednoczenia kodów w oparciu o przynajmniej nieformalny standard w postaci `srsName="EPSG:2180"`.
- Po pewnym czasie powyższe rozwiązanie uznano za nieformalne i zalecano stosowanie formy URN (*Uniform Resource Name*) w postaci `srsName="urn:ogc:def:crs:EPSG::2180"`. Takie zalecenie obowiązuje nadal w dokumentach technicznych INSPIRE.
- Obecnie wszelkie zbiory typu listy różnorodnych standardowych określeń są przechowywane w repozytoriach internetowych i udostępniane publicznie. Przykładem może być repozytorium list kodowych INSPIRE. W przypadku nazw układów odniesienia przestrzennego takie repozytorium prowadził EPSG (European Petroleum Survey Group), a obecnie IOGP (International Association of Oil & Gas Producers). W rezultacie atrybut określający SRS ma postać URL (*Uniform Resource Locator*) kierującego do odpowiedniej pozycji w repozytorium `srsName="http://www.epsg-registry.org/export.htm?gml=urn:ogc:def:crs:EPSG::2180"`. Odpowiedź, jaką przesyła to repozytorium jest przedstawiona w przykładzie 12.6.

Przykład 12.6. Odpowiedź, jaką przesyła repozytorium definicji układów odniesienia przestrzennego EPSG na polecenie `export.htm?gml=urn:ogc:def:crs:EPSG::2180`:

```
<gml:ProjectedCRS gml:id="epsg-crs-2180">
  <gml:metaDataProperty>
    <epsg:CommonMetaData>
      <epsg:type>projected</epsg:type>
      <epsg:alias alias="EUREF89 / CS92" code="1254"
        codeSpace="urn:ogc:def:naming-system:EPSG::7301"/>
      <epsg:alias alias="PL_EUREF89 / 1992" code="2951"
        codeSpace="urn:ogc:def:naming-system:EPSG::7307"/>
      <epsg:alias alias="System 1992" code="3084"
        codeSpace="urn:ogc:def:naming-system:EPSG::7314"/>
      <epsg:alias alias="ETRF89 / Poland CS92" code="3653"
        codeSpace="urn:ogc:def:naming-system:EPSG::7301"/>
      <epsg:informationSource>
        Glówny Urząd Geodezji i Kartografii via EuroGeographics;
        http://crs.bkg.bund.de/crs-eu/
      </epsg:informationSource>
      <epsg:revisionDate>2007-02-02</epsg:revisionDate>
      <epsg:changes>
        <epsg:changeID xlink:href="
          urn:ogc:def:change-request:EPSG::2005.180"/>
        <epsg:changeID xlink:href="
          urn:ogc:def:change-request:EPSG::2007.014"/>
      </epsg:changes>
      <epsg:show>true</epsg:show>
      <epsg:isDeprecated>false</epsg:isDeprecated>
    </epsg:CommonMetaData>
  </gml:metaDataProperty>
  <gml:identifier codeSpace="OGP">
    urn:ogc:def:crs:EPSG::2180
  </gml:identifier>
  <gml:name>
```

```

    ETRS89 / Poland CS92</gml:name>
  <gml:remarks>
    See ETRS89 / Poland CS2000 zones 5 - 8 (codes 2176-79)
    for large scale purposes.
  </gml:remarks>
  <gml:domainOfValidity xlink:href=
    "urn:ogc:def:area:EPSG::1192"/>
  <gml:scope>
    Medium and small scale topographic mapping
    (1:10,000 and smaller).
  </gml:scope>
  <gml:conversion xlink:href=
    "urn:ogc:def:coordinateOperation:EPSG::18300"/>
  <gml:baseGeodeticCRS xlink:href="urn:ogc:def:crs:EPSG::4258"/>
  <gml:cartesianCS xlink:href="urn:ogc:def:cs:EPSG::4531"/>
</gml:ProjectedCRS>

```

- Ostatecznie uznano, że odwoływanie się do nieformalnego repozytorium nie jest zgodne ze standardami i w OGC utworzone zostało repozytorium spełniające wszystkie wymogi formalne. W konsekwencji teraz należy ten atrybut w języku GML określać następująco: `srsName="http://www.opengis.net/def/crs/EPSSG/0/2180"`. Jednak informacje o tym układzie odniesienia przekazywane przez to repozytorium są tak obszerne, że zajmują ponad 800 linii tekstu i z tego powodu nie mogą tu być zacytowane.

Kolejnym problemem transformacji danych, nie tylko hydrogeologicznych, są listy kodowe często błędnie nazywane słownikami. W dawniejszych rozwiązaniach technologicznych poszczególne kody (nazwy) umieszczane były *in-line* bezpośrednio w zapisie danych analogicznie jak przedstawiony powyżej atrybut `srsName="EPSG:2180"`. Obecnie zgodnie z wymogami dokumentów technicznych INSPIRE listy kodowe są umieszczone w repozytorium INSPIRE, a w zapisach danych przy użyciu języka GML umieszcza się odwołanie do odpowiedniej pozycji w repozytorium. Zgodnie z tą zasadą w przykładzie 12.3. w elementach, których typem jest określona lista kodowa umieszcza się odsyłacz w formie `xlink:href`. Na przykład dla określenia rodzaju środowiska wodonośnego (typu *Aquifer*) jako krasowo-szczelinowe za pomocą elementu `mediaType` trzeba użyć następującego zapisu URL: `<ge_hg:mediaType xlink:href="http://inspire.ec.europa.eu/codelist/AquiferMediaTypeValue/karsticAndFractured/karsticAndFractured.pl.iso19135xml"/>`. Pod tą wartością URL znajduje się w repozytorium zapis XML definiujący znaczenie i parametry wskazanego elementu listy kodowej (przykład 12.7.). W źródłowych zbiorach danych informacje tego typu występują wyłącznie jako proste krótkie teksty, dlatego w procesie transformacji danych konieczna jest zamiana tego tekstu na odpowiednią wartość URL. Aby nie trzeba było wykonywać tych prac ręcznie potrzebne jest opracowanie słowników, gdzie po jednej stronie są teksty użyte w zbiorach danych źródłowych, a po drugiej odpowiadający mu wartości URL. Automatyzacja takich podstawień wymaga zapisu słownika w formie XML i opracowania odpowiedniego skryptu w języku zależnym od przyjętego środowiska systemowego, w którym przeprowadzana jest transformacja.

Przykład 12.7. Odpowiedź, jaką przesyła repozytorium list kodowych INSPIRE, gdy otrzyma polecenie zapisanie w elemencie `ge_hg:mediaType`, gdy ma to być środowisko krasowo-szczelinowe:

```
<?xml version="1.0" encoding="UTF-8"?>
<RE_RegisterItem
  xmlns:gmd="http://www.isotc211.org/2005/gmd"
  xmlns:gco="http://www.isotc211.org/2005/gco"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="http://www.isotc211.org/2005/grg"
  xsi:schemaLocation="http://www.isotc211.org/2005/grg
  http://standards.iso.org/iso/19135/-2/register.xsd">
  <itemIdentifier gco:nilReason="inapplicable"/>
  <name>
    <gco:CharacterString>
      krasowy i szczelinowaty
    </gco:CharacterString>
  </name>
  <status>
    <RE_ItemStatus>valid</RE_ItemStatus>
  </status>
  <definition>
    <gco:CharacterString>
      Połączenie warstwy wodonośnej o charakterze zarówno
      krasowym, jak i szczelinowatym.
    </gco:CharacterString>
  </definition>
  <description gco:nilReason="missing"/>
  <additionInformation
    xlink:href="http://inspire.ec.europa.eu/codelist/
AquiferMediaTypeValue/karsticAndFractured" />
  <itemClass>
    <RE_ItemClass>
      <name>
        <gco:CharacterString>CodeListValue</gco:CharacterString>
      </name>
      <technicalStandard gco:nilReason="inapplicable"/>
      <alternativeNames gco:nilReason="inapplicable"/>
      <describedItem gco:nilReason="inapplicable"/>
    </RE_ItemClass>
  </itemClass>
</RE_RegisterItem>
```

Podobny problem, również przedstawiony w przykładzie 12.3. dotyczy określania geologicznego wieku utworów, które tworzą środowisko wodonośne. W zbiorze źródłowym występuje tylko jedna litera „T” symbolizująca okres Trias, a w zbiorze wynikowym potrzebne jest odwołanie do odpowiedniej pozycji w repozytorium list kodowych INSPIRE w postaci URL: <http://inspire.ec.europa.eu/codelist/GeochronologicEraValue/triassic/triassic.pl.iso19135xml>. Pod tą wartością URL znajduje się w repozytorium zapis XML definiujący znaczenie i parametry wskazanego elementu listy kodowej (przykład 12.8.). W tym przypadku problem może być także rozwiązany za pomocą słownika.

Przykład 12.8. Odpowiedź, jaką przesyła repozytorium list kodowych INSPIRE, gdy otrzyma polecenie wskazania wieku utworów wodonośnych, gdy ma to być Trias:

```
<?xml version="1.0" encoding="UTF-8"?>
<RE_RegisterItem
  xmlns:gmd="http://www.isotc211.org/2005/gmd"
  xmlns:gco="http://www.isotc211.org/2005/gco"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="http://www.isotc211.org/2005/grg"
  xsi:schemaLocation="http://www.isotc211.org/2005/grg
    http://standards.iso.org/iso/19135/-2/register.xsd">
  <itemIdentifier gco:nilReason="inapplicable"/>
  <name>
    <gco:CharacterString>Triassic</gco:CharacterString>
  </name>
  <status>
    <RE_ItemStatus>valid</RE_ItemStatus>
  </status>
  <definition>
    <gco:CharacterString>
      Triassic (older bound-251 +/-0.4 Ma,
      younger bound-199.6 +/-0.6 Ma)
    </gco:CharacterString>
  </definition>
  <description>
    <gco:CharacterString>IUGS /CGI value</gco:CharacterString>
  </description>
  <additionInformation xlink:href="http://inspire.ec.europa.eu/
    codelist/GeochronologicEraValue/triassic"/>
  <itemClass>
    <RE_ItemClass>
      <name>
        <gco:CharacterString>CodeListValue</gco:CharacterString>
      </name>
      <technicalStandard gco:nilReason="inapplicable"/>
      <alternativeNames gco:nilReason="inapplicable"/>
      <describedItem gco:nilReason="inapplicable"/>
    </RE_ItemClass>
  </itemClass>
</RE_RegisterItem>
```

13. Podsumowanie

Przedstawiona problematyka, dotycząca metodyki i technologii transformacji krajowych danych hydrogeologicznych do schematów INSPIRE tematu Hydrogeologia, może się czytelnikowi wydawać zbyt rozległa i zbyt zawiła. Zdaniem autora, są to jednak tylko fragmenty dużej złożonej dziedziny, którą można nazwać „informatyzacją hydrogeologii”. To, co jest przedstawione w monografii to tylko metaforyczny „czubek góry lodowej”, która w dodatku jest zjawiskiem dynamicznym – przemieszcza się i zmienia swój kształt. Wiele innych zagadnień dotyczących teleinformatycznego przetwarzania, przechowywania i udostępniania informacji hydrogeologicznej nie zostało w niej poruszone lub zasygnalizowane bardzo ogólnie. Dotyczy to na przykład wymaganych przez INSPIRE usług sieciowych, takich jak: CSW, WMS, WFS i WPS dla wyszukiwania, przeglądania, pobierania i przetwarzania danych hydrogeologicznych. Odrębna problematyka w monografii nieporuszona to dane hydrogeologiczne w postaci pokryć (*coverage*) (Herring, Kottman, 1997; ISO/TC211, 2000; 2005b, 2009; INSPIRE DT-DS, 2013), które w hydrogeologii mają często kluczowe znaczenie (Michalak, Leśniak, 2003; Michalak, Nawalany, Sadurski i inni, 2011), na przykład przy zapisie danych do modelowania przepływu i transportu masy, a także do zapisu i udostępniania wyników uzyskanych z tych modeli. Specyfikacja INSPIRE określa definicje siatek dla danych typu pokrycie (INSPIRE TWG-CRS&GGS, 2014b). Model danych INSPIRE dla tematu Hydrogeologia przewiduje zastosowanie tego typu danych, to dla nich są przeznaczone klasy *HydrogeologicalSurface* ze stereotypem «union» i *PiezometricState* ze stereotypem «dataType». W pierwszej klasie można wymiennie używać trzech typów danych:

- *PointObservationCollection* jako zbiór danych hydrogeologicznych w elemencie *surfacePointCollection*,
- *RectifiedGridCoverage* jako zbiór danych hydrogeologicznych w elemencie *surfaceRectifiedGrid*,
- *ReferenceableGridCoverage* jako zbiór danych hydrogeologicznych w elemencie *surfaceReferenceableGrid*.

Trzeba podkreślić, że ten oryginalny i w wielu przypadkach niezbędny typ danych jest w innych tematach rzadko spotykany. Na przykład schematy tematu Geologia posługują się tylko formami wektorowymi typu *point*, *line* lub *surface*.

Kolejnym problemem związanym z danymi hydrogeologicznymi są standardy dotyczące obserwacji i pomiarów oraz przyjęta w INSPIRE norma ISO 19156:2011 *Geographic information – Observations and measurements* (ISO/TC211, 2011; OGC, 2013; Cox, 2010; OGC, 2007a; 2007b; INSPIRE CTWG-O&M, 2014), a także język WaterML, również przyjęty w INSPIRE. Standardy te powinny znaleźć szerokie zastosowanie w obserwacjach i pomiarach hydrogeologicznych.

Z pewnością może budzić zastanowienie fakt, że w przedstawionych w rozdziale 12. testach dane wynikowe (przykład 12.3.) mają około trzy razy większą objętość w porównaniu z danymi źródłowymi (przykład 12.1.), pomimo że zawierają praktycznie prawie te same informacje. Wytłumaczeniem jest inna struktura tych danych – trzeba podkreślić, że struktu-

ra też niesie ze sobą informacje i w rezultacie dane wynikowe mają znacznie więcej informacji niż źródłowe. Informacje, jakie niesie ze sobą struktura danych, mają znaczenie semantyczne i można na ich bazie przeprowadzać wnioskowanie, co jest w tym przypadku bardzo istotne, nie musi tego robić człowiek, ale odpowiednie algorytmy wnioskowania semantycznego, czyli oprogramowanie komputerowe. Jest to podstawa nowo rozwijającej się dziedziny określanej nazwą *Semantic Web* – pojęcie to jest najczęściej wiązane z Internetem trzeciej generacji (*Web 3.0* lub *Intelligent Web*). Podstawą takiego Internetu są technologie semantyczne Web, do których między innymi należą: RDF, OWL, SWRL, sparql, aplikacje *Semantic platform*, magazyny danych oparte na *tuplestores* i asocjacyjne bazy danych, a także technologie związane z przetwarzaniem języka naturalnego, maszynami uczącymi się i maszynami rozumującymi. Można przyjąć, że problematyka będąca przedmiotem tej monografii w jakimś niewielkim stopniu też przyczynia się do rozwoju koncepcji Web 3.0 na gruncie hydrogeologii. Znaczna objętość zapisów danych w mocno ustrukturalizowanych językach, takich jak GML 3.2.1 w tej sytuacji nie jest problemem. Szczególnie gdy się weźmie pod uwagę, że moc obliczeniowa komputerów i pojemność ich pamięci masowych rośnie z każdym rokiem wykładniczo.

W trakcie prowadzenia prac badawczych, jedną z kluczowych spraw było poszukiwanie informacji o podobnych pracach na gruncie hydrogeologii w innych ośrodkach w Polsce i na świecie. Rezultat tych poszukiwań był negatywny i taki jest nadal. Jedynym wyjątkiem są prace zespołu kanadyjskich geologów we współpracy z geologami australijskimi nad rozszerzeniem języka geologicznego GeoSciML do zastosowań w zakresie wód podziemnych, jednak prace te są prowadzone bez udziału hydrogeologów i z tego względu ich przydatność w hydrogeologii będzie ograniczona.

Z przedstawionych w niniejszej monografii prac wynika pilna potrzeba utworzenia zespołu hydrogeologów, najlepiej międzynarodowego, dla opracowania własnego języka znacznikowego dla danych hydrogeologicznych. Język ten prawdopodobnie bazowałby na modelu INSPIRE *HydrogeologyCore* i mógłby się stać wobec niego rozszerzeniem jako *HydrogeologyExtension*. Wzorem mogą być dwa już działające zespoły: zespół projektu GeoSciML, który składa się z kilkudziesięciu osób, w tym kilkunastu bardzo aktywnych i zespół projektu WaterML, który liczy kilkanaście również aktywnych osób. Na zakończenie pozostaje autorowi pozostać w głębokim przekonaniu, że wkrótce powstanie równie aktywny i liczny zespół hydrogeologów i rozpoczną się prace nad językiem danych hydrogeologicznych na poziomie porównywalnym z dwoma wyżej przytoczonymi.

Literatura

- Almendros-Jimenez J. M., Becerra-Terón A., Garcia-Garcia F., 2010: Development of a Query Language for GML based on XPath. URL: <http://www.complang.tuwien.ac.at/WWV2010/proc/regular1.pdf>
- BaseX Team, 2014: BaseX Documentation, Version 7.9. URL: <http://files.basex.org/releases/7.9/BaseX79.pdf>
- BGR (Bundesanstalt für Geowissenschaften und Rohstoffe), 2008: IHME 1500 – International Hydrogeological Map of Europe 1:1 500 000. URL: http://www.bgr.bund.de/nn_336918/EN/Themen/Wasser/Projekte/Berat_Info/Ihme1500/ihme1500_projektbeschr_en.html
- BGR (Federal Institute for Geosciences and Natural Resources), 2014: Metadata of the International Hydrogeological Map of Europe 1:1,500,000 in Vector Format (Version IHME1500 v1.1): 6 p.; Hannover. (PDF, 196 KB). URL: http://www.bgr.bund.de/EN/Themen/Wasser/Projekte/laufend/Beratung/Ihme1500/ihme1500-v11_metadata.pdf
- 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=f2c1f79e942a2cf12ed12a99ace5eec0>
- 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., 2012: Summary of models. URL: http://external.openeospatial.org/twiki_public/pub/HydrologyDWG/GWIE2-ModelComparison/Boisvert_summary_of_models_21June12.ppt
- Boisvert E., Brodaric B., 2007: GroundWater Markup Language (GWML): Extending GeoSciML for Groundwater. AGU Fall Meeting Abstracts 12/2007. URL: http://www.researchgate.net/publication/252950240_GroundWater_Markup_Language_%28GWML%29_Extending_GeoSciML_for_Groundwater
- Boisvert E., Brodaric B., 2012: GroundWater Markup Language (GWML) – enabling groundwater data interoperability in spatial data infrastructures. *Journal of Hydroinformatics*. International Water Association Publishing. URL: <http://www.iwaponline.com/jh/014/0093/0140093.pdf>
- Booch G., Rumbaugh J., Jacobson I., 2002: UML – przewodnik użytkownika. Z serii: Inżynieria oprogramowania. Wydanie polskie. Wydawnictwa Naukowo-Techniczne, Warszawa.
- Bragg K., 2014: INSPIRE Directive: 5 Tips for Providing Data to Europe’s SDI. Safe Software. URL: <http://blog.safe.com/2014/08/inspire-directive-europe-sdi/>
- Bray T., Paoli J., Sperberg-McQueen C.M., 1998: Extensible Markup Language (XML) 1.0, W3C Recommendation. URL: <http://www.w3.org/TR/1998/REC-xml-19980210>
- 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.openeospatial.org/files/?artifact_id=42729
- Brodaric B., Booth N., Boisvert E., Lucido J., 2014: Levels of data interoperability in the emerging North American groundwater data network. 11th International Conference on Hydroinformatics. URL: http://acwi.gov/sogw/pubs/papers/DataInterop_HIC2014.pdf
- Brodaric B., Hahmann T., 2014: Toward a foundational hydro ontology for water data interoperability. 11th International Conference on Hydroinformatics. URL: http://www.spatial.maine.edu/~torsten/publications/BBrodaric_HIC-2014.pdf
- Brodie R. (ed.), Working Group on National Groundwater Data Standards, 1999: The Australian National Groundwater Data Transfer Standard, Release 1.0. National Groundwater Committee, Bureau of Rural Sciences (BRS) Arch., Canberra. URL: <http://www.brs.gov.au/land&water/groundwater/>
- 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/wiki_public/GML/Gml4WorkshopInput
- Carlson D., 2001: Modeling XML Applications with UML: Practical e-Business Applications. Addison-Wesley, Boston.
- Chen J., He B., Wang W., 2010: GXQuery: A GML Spatial Data Query Language.
URL: <http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.ieee-000005478020>
- Cogho V.E., Kirchner J., Morris J.W., 1989: A National Groundwater Database for South Africa: Development of the database. Pretoria, Report Water Research Commission 150/1/89.
- Collins J.B., Clark D., 2004: Towards an Ontology of Physics. URL: http://www.nrl.navy.mil/itd/imda/sites/www.nrl.navy.mil.itd.imda/files/pdfs/04E-SIW-044_final.pdf
- Connolly T., Begg C., 2004: Systemy baz danych. Tom 1 i 2. Wyd. RM, Warszawa.
- Córcoles J.E., González P., 2003: A Specification of a Spatial Query Language over GML.
URL: <http://www.dsi.uclm.es/personal/JoseEduardoCorcoles/pdf/ACM2001Corcoles.pdf>
- Cox S. (ed.) (OGC), 2010: Geographic Information: Observations and Measurements – OGC Abstract Specification Topic 20. URL: http://portal.opengeospatial.org/files/?artifact_id=41579
- Curtis E., 2009: Tools and Experiences in Implementing INSPIRE Data Specifications IV: Practical Schema Translation for INSPIRE. URL: <http://www.gsdi.org/gsdiconf/gsd11/wrkshpslides/w4.3a.pdf>
- DHP (Data Harmonisation Panel), 2009: HUMBOLDT Harmonisation Toolkit – HUMBOLDT GeoModel Editor. URL: <http://www.esdi-community.eu/projects/geomodel>
- DHP (Data Harmonisation Panel), 2011: Conceptual Schema Translation (CST) – Web Processing Service. URL: <http://www.esdi-community.eu/projects/cst>
- DHP (Data Harmonisation Panel), 2014a: HUMBOLDT Harmonisation Toolkit – HUMBOLDT Alignment Editor. URL: <http://www.esdi-community.eu/projects/hale>
- DHP (Data Harmonisation Panel), 2014b: The HUMBOLDT Service Integration Framework. URL: <http://www.esdi-community.eu/projects/framework>
- Dornblut I., Atkinson R., 2014: HY_Features: a Common Hydrologic Feature Model Discussion Paper. Open Geospatial Consortium, OGC Repository, Wayland.
URL: https://portal.opengeospatial.org/files/?artifact_id=55157
- Duscher K., 2014a: IHME1500 – International Hydrogeological Map of Europe 1:1,500,000. URL: http://www.bgr.bund.de/EN/Themen/Wasser/Projekte/laufend/Beratung/Ihme1500/ihme1500_projekt_beschr_en.html
- Duscher K., 2014b: International Hydrogeological Map of Europe 1:1,500,000 – Extract from the INSPIRE-conform metadata set. URL: http://www.bgr.bund.de/EN/Themen/Wasser/Projekte/laufend/Beratung/Ihme1500/ihme1500-v11_metadata.pdf?__blob=publicationFile&v=3
- Duscher K., Günther A., 2014: Spatial Data of the International Hydrogeological Map of Europe 1:1,500,000 in Vector Format (Version IHME1500 v1.1). – Technical note, 3 p, BGR, Hannover. (PDF, 53 KB). URL: http://www.bgr.bund.de/EN/Themen/Wasser/Projekte/laufend/Beratung/Ihme1500/ihme1500-v11_technical-note.pdf
- Dutch Kadaster (GEORZ-lab & Research), 2011: INSPIRE Prototypes (Phase 2) – Dutch Kadaster, Release 2.0. URL: <http://inspire.kademo.nl/doc/pdf/inspire-prototypes-phase2.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
- EC (European Commission), 2010: Rozporządzenie Komisji (UE) nr 1089/2010 z dnia 23 listopada 2010 r. w sprawie wykonania dyrektywy 2007/2/WE Parlamentu Europejskiego i Rady w zakresie interoperacyjności zbiorów i usług danych przestrzennych.
URL: <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=OJ:L:2010:323:FULL&from=EN>
- EC (European Commission), 2011: Rozporządzenie Komisji (UE) nr 102/2011 z dnia 4 lutego 2011 r. zmieniające rozporządzenie (UE) nr 1089/2010 w sprawie wykonania dyrektywy 2007/2/WE Parlamentu Europejskiego i Rady w zakresie interoperacyjności zbiorów i usług danych przestrzennych.
URL: <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:32011R0102&from=EN>

- EC (European Commission), 2013a: Rozporządzenie Komisji (UE) nr 1253/2013 z dnia 21 października 2013 r. zmieniające rozporządzenie (UE) nr 1089/2010 w sprawie wykonania dyrektywy 2007/2/WE w zakresie interoperacyjności zbiorów i usług danych przestrzennych.
URL: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:331:0001:0267:PL:PDF>
- EC (European Commission), 2013b: Rozporządzenie Komisji (UE) nr 089/2010 z dnia 23 listopada 2010 r. w sprawie wykonania dyrektywy 2007/2/WE Parlamentu Europejskiego i Rady w zakresie interoperacyjności zbiorów i usług danych przestrzennych. (Tekst ujednolicony). URL:
<http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:02010R1089-20131230&from=EN>
- EEA (European Environment Agency), 2015: WISE SoE – Ammonium in groundwater by countries and WFD gw bodies. URL: <http://www.eea.europa.eu/data-and-maps/explore-interactive-maps/ammonium-in-groundwater-by-countries>
- 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). Official Journal of the E. L 108/1-14.
- eXist Solutions, 2014: Vitamins for your Applications.
URL: <http://exist-db.org/exist/apps/homepage/index.html>
- Fert M., Mordzonek G., Węglarz D., 2011: Aktualny stan i formy udostępniania bazy danych GIS Mapy Hydrogeologicznej Polski. Biuletyn PIG, nr 445: 87-96.
URL: http://www.pgi.gov.pl/pl/component/docman/doc_download/378-fert-i-in.html
- Fitzgerald M., 2004: Learning XSLT. Wyd. O'Reilly Media Inc.
- Gilbrich W.H., 2000: International Hydrogeological Map of Europe. – Feature Article, Waterway No. 19, Paris. URL: <http://webworld.unesco.org/water/ihp/publications/waterway/wat19/W19Feature.html>
- Goczyła K., 2011: Ontologie w systemach informatycznych. Wyd. AOW EXIT.
- Graves M., 2002: Projektowanie baz danych XML, Vademecum profesjonalisty. Wyd. Helion.
- Herbich P., Mordzonek G., Przytuła E., 2011: Stan rozpoznania i stopień wykorzystania dostępnych do zagospodarowania zasobów wód podziemnych w Polsce. Biuletyn PIG. Nr 445, Hydrogeologia z. 12/1: 193-202. URL: http://www.pgi.gov.pl/en/component/docman/doc_view/382-herbich-i-in
- Herbich P., Kapuściński J., Nowicki K., Prażak J., Skrzypczyk L., 2008: Metodyka wyznaczania obszarów ochronnych Głównych Zbiorników Wód Podziemnych dla potrzeb planowania i gospodarowania wodami w obszarach dorzeczy. URL: http://www.pgi.gov.pl/pl/component/docman/doc_download/248-metodyka-wyznaczania-obszarow-ochronnych-wod-podziemnych
- Herbich P., Kapuściński J., Nowicki K., Rodzoch A., 2013: Metodyka określania zasobów dyspozycyjnych wód podziemnych w obszarach bilansowych.
URL: https://www.mos.gov.pl/g2/big/2013_11/8fb9ee316d33d4561f4fdbd169b5f14d.pdf
- Herring J., Kottman C., 1997: Features and Coverages for the Layman. OpenGIS Newsletter, 2, 4: 6-7.
- Hill B.M., Harris D.B., Vyas J., 2006: Debian GNU/Linux 3.1. Wyd. Helion.
- Hollis J.M., Holmani. P., Burton R.G.O., Willers C.M., Pavelley D., 2002: A Digital Dataset of European Groundwater Resources at 1:500,000. (v. 1.0).
URL: http://eusoiils.jrc.ec.europa.eu/ESDB_Archive/groundwater/docs/EUGWFR.pdf
- Holupirek A., Grün C., Scholl M.H., 2009: BaseX & DeepFS – Joint Storage for Filesystem and Database.
URL: http://kops.uni-konstanz.de/bitstream/handle/123456789/6449/p1108_holupirek.pdf?sequence=1
- Holupirek A., Grün C., Scholl M.H., 2008: BaseX & DeepFS, Joint Storage for Filesystem and Database.
URL: http://www.inf.uni-konstanz.de/dbis/publications/download/joint_storage.pdf
- Holzner S., 2002: XSLT. Vademecum profesjonalisty. Wydawnictwo Helion.
URL: <http://helion.pl/ksiazki/xslt-vademecum-profesjonalisty-steven-holzner,xsltpv.htm>
- Howden N., Mather J., 2013: History of Hydrogeology. CRC Press.
- 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>
- Hudson A., Hudson P., 2008: Fedora 7, księga eksperta. Wyd. Helion.
- 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.geosciml.org/geosciml/2.0/cookbook/GeoSciML_Data_CookBook_V2.pdf

- INSPIRE CTWG-O&M (INSPIRE Cross Thematic Working Group on Observations & Measurements), 2014: D2.9 Guidelines for the use of Observations & Measurements and Sensor Web Enablement – related standards in INSPIRE Annex II and III data specification development.
URL: http://inspire.ec.europa.eu/documents/Data_Specifications/D2.9_O&M_Guidelines_v2.0.pdf
- INSPIRE 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
- INSPIRE DT-DS (INSPIRE Drafting Team „Data Specifications”), 2013: D2.10.2: INSPIRE Data Specifications – Base Models – Coverage Types, Version 1.0rc3.
URL: http://inspire.ec.europa.eu/documents/Data_Specifications/D2.10.2_CoverageTypes_v1.0rc3.pdf
- INSPIRE DT-DS (INSPIRE Drafting Team „Data Specifications”), 2014a: D2.7: Guidelines for the encoding of spatial data, Version 3.3. URL: http://inspire.ec.europa.eu/documents/Data_Specifications/D2.7_v3.3.pdf
- INSPIRE DT-DS (INSPIRE Drafting Team „Data Specifications”), 2014b: D2.5: Generic Conceptual Model, Version 3.4. URL: http://inspire.ec.europa.eu/documents/Data_Specifications/D2.5_v3.4.pdf
- INSPIRE IOCTF-NS (INSPIRE Initial Operating Capability Task Force for Network Services), 2013: Technical Guidance for the implementation of INSPIRE Download Services.
URL: http://inspire.ec.europa.eu/documents/Network_Services/Technical_Guidance_Download_Services_v3.1.pdf
- INSPIRE TWG-AM (INSPIRE Thematic Working Group Area Management/Restriction/Regulation Zones and Reporting Units), 2013: D2.8.III.11 INSPIRE Data Specification on Area Management/Restriction/Regulation Zones and Reporting Units – Technical Guidelines. URL: http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AM_v3.0.pdf
- INSPIRE TWG-CRS&GGS (INSPIRE Thematic Working Group Coordinate Reference Systems & Geographical Grid Systems), 2014a: D2.8.I.1 Data Specification on Coordinate Reference Systems – Technical Guidelines.
URL: http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_RS_v3.2.pdf
- INSPIRE TWG-CRS&GGS (INSPIRE Thematic Working Group Coordinate Reference Systems & Geographical Grid Systems), 2014b: D2.8.I.2 Data Specification on Geographical Grid Systems – Technical Guidelines. URL: http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_GG_v3.1.pdf
- INSPIRE TWG-EM (INSPIRE Thematic Working Group Environmental Monitoring Facilities), 2013: D2.8.II/III.7 INSPIRE Data Specification on Environmental Monitoring Facilities – Technical Guidelines. URL: http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_EF_v3.0.pdf
- INSPIRE TWG-ER (INSPIRE Thematic Working Group Energy Resources), 2014: D2.8.III.20 INSPIRE Data Specification on Energy Resources – Technical Guidelines.
URL: http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_ER_v3.0.pdf
- INSPIRE TWG-GE-MR (INSPIRE Thematic Working Group Geology and Mineral Resources), Wagner B., Michalak J., Robida F., Schubert C., 2013: D2.8.II.4 INSPIRE Data Specification on Geology – Technical Guidelines. URL: http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_GE_v3.0.pdf
- INSPIRE TWG-GE-MR (INSPIRE Thematic Working Group Geology and Mineral Resources), 2013: D2.8.III.21 INSPIRE Data Specification on Mineral Resources – Technical Guidelines.
URL: http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_MR_v3.0.pdf
- IP (Intergraph Polska), 2011: GeoMedia SDI Portal – wersja 06.01.00.23 – Instrukcja użytkownika.
URL: <http://epsh.pgi.gov.pl/epsh/doc/instrukcja.pdf>
- ISO/TC211 (Geographic Information/Geomatics), 2000: ISO/TR 19121:2000 – Geographic information – Imagery and gridded data. ISO/TC211 Doc. Repository, Lysaker, Norway.
- ISO/TC211 (Geographic Information/Geomatics), 2004: ISO 19106:2004 – Geographic information – Profiles. ISO/TC211 Doc. Repository, Lysaker, Norway.
- ISO/TC211 (Geographic Information/Geomatics), 2005a: ISO 19119:2005 – Geographic information – Servi-

- ces. ISO/TC211 Doc. Repository, Lysaker, Norway.
- ISO/TC211 (Geographic Information/Geomatics), 2005b: ISO 19123:2005 – Geographic information – Schema for coverage geometry and functions. ISO/TC211 Doc. Repository, Lysaker, Norway.
- ISO/TC211 (Geographic Information/Geomatics), 2005c: ISO 19128:2005 – Geographic information – Web map server interface. ISO/TC211 Doc. Repository, Lysaker, Norway.
- ISO/TC211 (Geographic Information/Geomatics), 2007a: ISO 19136:2007 – Geographic information – Geography Markup Language (GML). ISO/TC211 Doc. Repository, Lysaker, Norway.
URL: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=32554
- ISO/TC211 (Geographic Information/Geomatics), 2007b: ISO 19137:2007 – Geographic information – Core profile of the spatial schema. ISO/TC211 Doc. Repository, Lysaker, Norway.
- ISO/TC211 (Geographic Information/Geomatics), 2009: ISO/TS 19129:2009 – Geographic information – Imagery, gridded and coverage data framework. ISO/TC211 Doc. Repository, Lysaker, Norway.
- ISO/TC211 (Geographic Information/Geomatics), 2011: ISO/CD 19156 – Geographic information – Observations and measurements. ISO/TC211 Doc. Repository, Lysaker, Norway.
- ISO/TC211 (Geographic Information/Geomatics), 2010: ISO/DIS 19142 – Geographic information – Web Feature Service. ISO/TC211 Doc. Repository, Lysaker, Norway.
- JRC (European Commission, DG Joint Research Centre, Institute for Environment and Sustainability), 2009: Digital Dataset of European Groundwater Resources, version 1.0 (Overview).
URL: http://eusoiils.jrc.ec.europa.eu/ESDB_Archive/groundwater/gw.html
- Kay M., 2008: XSLT 2.0 and XPath 2.0 Programmer's Reference. 4th Edition. Wyd. Wiley Publishing Inc.
- Kazimierski B., Kulczyńska A., Sadurski A., Skrzypczyk L., 2011: Założenia do modernizacji Monitoringu Wód Podziemnych. Biuletyn PIG, nr 445: 279-290.
URL: http://www.pgi.gov.pl/pl/component/docman/doc_download/389-kazimierski-sadurski.html
- Lake R., 1999: Introduction to GML – Geography Markup Language, Galdos Systems Inc. Arch.
URL: <http://www.focalpoint.org/galdos/GMLIntroduction.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
- Latvala P., Lehto L., 2011: Implementation of INSPIRE Schema Transformation Service.
URL: <http://ojs.tsv.fi/index.php/njs/article/view/4306/5275>
- Lekkerkerk H.-J., 2012: Integrating data: a water quality case study. GIMA – Geographical Information Management and Application.
URL: http://www.gdmc.nl/publications/2012/Integrating_data_water_quality_case_study.pdf
- Li X., 2006: XQuery As a Spatial Query Language.
URL: http://www.researchgate.net/publication/220968590_XQuery_as_a_Spatial_Query_Language
- Ma X., 2011: Ontology Spectrum for Geological Data Interoperability. Doctoral dissertation.
URL: http://www.itc.nl/library/papers_2011/phd/ma.pdf
- Mahlow C., Grüny C., Holupireky A., Scholly M.H., 2012: A Framework for Retrieval and Annotation in Digital Humanities using XQuery Full Text and Update in BaseX.
URL: <http://www.inf.uni-konstanz.de/gk/pubsys/publishedFiles/MaGrHo12.pdf>
- Mangano S., 2006: XSLT Cookbook. Second Edition. Wyd. O'Reilly Media Inc.
- Mangano S., 2007: XSLT Receptury. Wydanie 2. Wyd. Helion. Polskie wydanie pozycji Mangano S., 2006.
- 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>
- Mercer D., 2000: XML – kurs podstawowy. Wyd. Edition 2000, Kraków.
- Michalak J. (red.), Chojka A., Zwirowicz-Rutkowska A., Parzyński Z., 2012: Modele danych przestrzennych w UML i ich transformacja do schematów GML i struktur baz danych. *Roczniki Geomatyki* t. 10, z. 1 (51), 134 s., PTIP, Warszawa, monografia. URL: <http://rg.ptip.org.pl/index.php/rg/issue/view/238>
- Michalak J., 2000: GML – język zapisu geoinformacji. [W:] Systemy Informacji Przestrzennej, tom X: 189-198, PTIP, Warszawa. URL: <https://pbn.nauka.gov.pl/works/30481>
- Michalak J., 2003a: Geomatics in hydrogeology. *Geological Quarterly* 47(1): 69-76.

- Michalak J., 2003b: Modele pojęciowe hydrogeologicznych danych geoprzestrzennych – podstawy metodyczne. *Biuletyn PIG – Hydrogeologia*, z. V, nr 406: 154 s., monografia. URL: <http://www.pgi.gov.pl/pl/wydawnictwa-geologiczne-pig-pib/seryjne/biuletyn-pig/138-biuletyn-pig-nr-406.html>
- Michalak J., 2003c: Podstawy metodyczne i technologiczne infrastruktur geoinformacyjnych. *Roczniki Geomatyki* t. 1, z. 2: 140 s., PTIP, Warszawa, monografia.
URL: <http://rg.ptip.org.pl/index.php/rg/issue/view/189>
- Michalak J., 2003d: Standardy ISO 19100 i OpenGIS jako podstawa państwowej infrastruktury geoinformacyjnej w zakresie geologii. *Przegląd Geologiczny* v. 51, nr 4: 311-315.
- Michalak J., 2004: Multilingual European Subset of Unicode in geospatial data encoding. *Annals of Geomatics* v. 2, no. 1: 53-66, PTIP, Warszawa.
- Michalak J., 2005: HGLML – HydroGeoLogical Markup Language – znacznikowy język wymiany geoinformacji hydrogeologicznej. *Współczesne Problemy Hydrogeologii*, t. XII: 499-504.
URL: <https://pbn.nauka.gov.pl/works/56426>
- Michalak J., 2008: Języki pochodne od GML i z nim powiązane. *Roczniki Geomatyki* t. 6, z. 5: 75-84, PTIP, Warszawa. URL: <http://rg.ptip.org.pl/index.php/rg/article/view/RG2008-5-Michalak2>
- Michalak J., 2009: Zadania środowiska hydrogeologów w budowie infrastruktury INSPIRE. *Biuletyn PIG – Hydrogeologia*, nr 436: 329-334.
URL: http://geoshale.plnwww.geoshale.pl/images/stories/NW/Biuletyny_PIG/436/b436_michalak.pdf
- Michalak J., 2010a: Metodyka i technologia budowy geoservera tematycznego jako komponentu INSPIRE. *Roczniki Geomatyki* t. 8, z. 3 (39): 98 str., PTIP, Warszawa, monografia.
URL: <http://rg.ptip.org.pl/index.php/rg/issue/view/226>
- Michalak J., 2010b: Transformacja polskich danych przestrzennych do modeli INSPIRE. [W:] Baranowski M., Bielecka E. (red.) Modelowanie danych przestrzennych. *Roczniki Geomatyki* t. 8, z. 4 (40): 27-51, PTIP, Warszawa. URL: <http://rg.ptip.org.pl/index.php/rg/article/view/RG2012-2-Michalak>
- Michalak J., 2011: Informacja i wskazówki dotyczące testowania projektów specyfikacji danych II i III grupy tematycznej. Rada Infrastruktury Informacji Przestrzennej.
URL: http://www.radaiiip.gov.pl/_data/assets/pdf_file/0006/34494/Testowanie_specyfikacji_II_III.pdf
- Michalak J., 2012: Testowanie roboczych wersji specyfikacji danych tematów załączników II i III INSPIRE. *Roczniki Geomatyki*, t. 10, z. 2 (52): 51-63 PTIP, Warszawa.
URL: <http://rg.ptip.org.pl/index.php/rg/article/view/RG2010-4-Michalak>
- Michalak J., Boisvert E., 2013: Discussion Page for GWML 2 Consolidated Entity List . GWIE2 Repository.
URL: http://external.opengeospatial.org/twiki_public/HydrologyDWG/Gwml2ConsolFeatureListDiscussionPage
- Michalak J., Leśniak P. M., 2003: Features and coverages in hydrogeological information. *Acta Geologica Polonica* vol. 53, nr 3: 247-255.
- Michalak J., Nawalany M., Sadurski A., (red.) i inni, 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
- MŚ (Ministerstwo Środowiska, Departament Geologii i Koncesji Geologicznych), 2008a: Polityka resortu w dziedzinie hydrogeologii (na lata 2008-2015).
URL: https://www.mos.gov.pl/g2/kategoriaPliki/2009_04/a8dd0a8d9e7e4634bc79dc63fe715cde.pdf
- MŚ (Ministerstwo Środowiska, Departament Geologii i Koncesji Geologicznych), 2008b: Kierunki badań w dziedzinie kartografii geologicznej (na lata 2008-2015).
URL: https://www.mos.gov.pl/g2/kategoriaPliki/2009_04/01_kartografia.pdf
- Murray D., 2010: INSPIRE-d: This is the Year of Transformation.
URL: <http://blog.safe.com/2010/06/inspire-d-this-is-the-year-of-transformation/>
- Myciuk K., 2012: Metadane hydrogeologiczne w Polsce.
URL: http://www.psh.gov.pl/artykuly_i_publicacje/artykuly1/metadane-hydrogeologiczne-w-polsce.html
- Nalęcz T., Michalak J., 2011: Modele pojęciowe INSPIRE dla tematu Geologia jako wstęp do opisu i schematyzacji zjawisk hydrogeologicznych. *Roczniki Geomatyki*, t. 9, z. 4 (48): 117-128, PTIP, Warszawa. URL: <http://rg.ptip.org.pl/index.php/rg/article/view/RG2011-4-NaleczMichalak>
- Nathaniel L., Booth N.L., Brodaric B., Lucido J. M., Kuo I-L, Boisvert E., Cunningham W.L., 2011: Develop-

- ment of an Interoperable Groundwater Data Exchange Network between the United States and Canada. URL: http://acwi.gov/sogw/pubs/tr/pilot_results/Booth_etal_GeoHydro2011.pdf
- Negus C., 2005: Fedora Core 3., biblia. Wyd. Helion.
- Nemeth E., Snyder G., Hein T.R., Whaley B., 2011: UNIX i Linux, przewodnik administratora systemów. Wydanie IV. Wyd. Helion.
- Newell C.J., Loren P., Hopkins L.P., Bedient P.B., 1990: A Hydrogeologic Database for Ground-Water Modeling. *Groundwater*, vol. 28, issue 5, pages 703–714.
URL: <http://onlinelibrary.wiley.com/doi/10.1111/j.1745-6584.1990.tb01986.x/abstract>
- Nowakowska P., Hordejuk M., Józwiak K., Myciuk K., Przasnyska J., Sadłowska K., Tekielska A., 2012: Testowanie specyfikacji danych II i III grupy tematycznej INSPIRE w zakresie tematów przyporządkowanych Głównemu Geologowi Kraju. *Roczniki Geomatyki* t. 10, z. 2(52):71-79, PTIP, Warszawa.
URL: <http://rg.ptip.org.pl/index.php/rg/article/view/RG2012-2-Nowakowska-inni>
- Nowakowska P., Michalak J., Hordejuk M., Józwiak K., Myciuk K., Przasnyska J., Tekielska A., 2011: Testowanie specyfikacji danych II i III grupy tematycznej dyrektywy INSPIRE w zakresie tematów, za które odpowiada Główny Geolog Kraju jako organ wiodący. Raport dla Ministerstwa Środowiska.
- Nowicki Z. i inni, 2009: Jednolite części wód podziemnych w Polsce – Charakterystyka geologiczna i hydrogeologiczna – Uwagi wprowadzające. URL: http://www.psh.gov.pl/plik/id,4787,v,artykul_5576.pdf
- OGC (Open Geospatial Consortium), 2001: OpenGIS Coordinate Transformation Service Implementation Specification. Ver. 1.0, Doc. Nr: 01-009, Ed.: Daly M., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2006: OpenGIS Web Map Service (WMS) Implementation Specification. Ver. 1.3.0, Doc. Nr: 06-042, Ed.: de La Beaujardiere J., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2007a: Observations and Measurements – Part 1 – Observation schema. Ver. 1.0, Doc. Nr: 07-022r1, Ed.: Cox S., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2007b: Observations and Measurements – Part 2 – Sampling Features. Ver. 1.0, Doc. Nr: 07-002r3, Ed.: Cox S., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2007c: OpenGIS Geography Markup Language (GML) Encoding Standard. Ver. 3.2.1, Doc. Nr: 07-036, Ed.: Portele C., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2007d: Web Processing Service. Ver. 1.0.0, Doc. Nr: 05-007r7, Ed: Schut P., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2007e: OpenGIS Catalogue Services Specification. Version 2.0.2. Eds: Nebert D., Whiteside A., Vretanos P., document: OGC 07-006r1. OGC Repository, Wayland.
URL: http://portal.opengeospatial.org/files/?artifact_id=20555
- OGC (Open Geospatial Consortium), 2008a: OpenGIS City Geography Markup Language (CityGML) Encoding Standard. Ver. 1.0.0, Doc. Nr: 08-007r1, Eds.: Gröger G., Kolbe T.H., Czerwinski A., Nagel C., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2008b: Web Coverage Service (WCS) Implementation Standard. Ver. 1.1.2, Doc. Nr: 07-067r5, Ed.: Whiteside A., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2009a: OpenGIS Web Coverage Processing Service (WCPS) Language Interface Standard. Ver. 1.0.0, Doc. Nr: 08-068r2, Ed.: Baumann P., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2009b: OpenGIS Web Coverage Service (WCS) – Processing Extension (WCPS). Ver. 1.0.0, Doc. Nr: 08-059r3, Ed.: Baumann P., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2009c: OpenGIS Web Feature Service (WFS) Implementation Specification with XLinks. Ver. 1.1.0, Doc. Nr: 04-094, Ed.: Vretanos P.A., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2009d: Web Coverage Processing Service (WCPS) Language Interface Standard. Version: 1.0.0, Document: OGC 08-068r2, Ed.: Baumann P., OGC Repository, Wayland. URL: http://portal.opengeospatial.org/files/?artifact_id=32319
- OGC (Open Geospatial Consortium), 2010a: OGC Web Service Common Implementation Specification. Ver. 2.0.0, Doc. Nr: 06-121r9, Eds.: Whiteside A., Greenwood J., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2010b: OpenGIS Web Map Tile Service Implementation Standard. Ver. 1.0.0, Doc. Nr: 07-057r7, Eds.: Masó J., Pomakis K., Juliá N., OGC Repository, Wayland.
- OGC (Open Geospatial Consortium), 2013: Geographic information – Observations and measurements. OGC Abstract Specification – Topic 20. URL: http://portal.opengeospatial.org/files/?artifact_id=41579
- OGC (Open Geospatial Consortium), 2014: OGC WaterML 2.0: Part 1 – Timeseries. OGC Repository,

- Wayland. URL: https://portal.opengeospatial.org/files/?artifact_id=57222
- OSGeo-Live, 2014: GeoKettle – Integracja danych.
URL: http://live.osgeo.org/pl/overview/geokettle_overview.html
- 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>
- PlanetEarth, 2008: International Hydrogeological Map of Europe. Reduced to the scale 1:5 000 000. Special Map Mosaic printed on the occasion of the International Year of Planet Earth 2008.
URL: http://www.bgr.de/app/fishy/ihme1500/pdf/ihme_150dpi.pdf
- Portale C., Van den Broecke J., Cox S. and others, 2009: Annex I Schemas – GML 3.2.1 with WFS ? INSPIRE Forum.
URL: <http://inspire-forum.jrc.ec.europa.eu/pg/forum/topic/3309/annex-i-schemas-gml-321-with-wfs/>
- Portale C., 2007: OpenGIS Geography Markup Language (GML) Encoding Standard. OpenGIS Standard.
URL: http://portal.opengeospatial.org/files/?artifact_id=20509
- PSH (Państwowa Służba Hydrogeologiczna), 2011: Wody lecznicze, wody termalne i solanki – baza danych.
URL: <http://mineralne.pgi.gov.pl/o-nas.html>
- PSH (Państwowa Służba Hydrogeologiczna), 2012a: Baza danych Monitoring Wód Podziemnych.
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/bazy_danych_mapy/monitoring-wod-podziemnych.html
- PSH (Państwowa Służba Hydrogeologiczna), 2012b: Jak Przeglądać dane?
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/jak_przegladowac_dane/
- PSH (Państwowa Służba Hydrogeologiczna), 2013a: Bank HYDRO.
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/bazy_danych_mapy/bankhydro.html
- PSH (Państwowa Służba Hydrogeologiczna), 2013b: Główne Zbiorniki Wód Podziemnych (GZWP).
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/bazy_danych_mapy/gzwp.html
- PSH (Państwowa Służba Hydrogeologiczna), 2013c: Jak zamówić dane?
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/jak_zamowic_dane/
- PSH (Państwowa Służba Hydrogeologiczna), 2013d: Mapa hydrogeologiczna Polski (MhP).
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/bazy_danych_mapy/mhp.html
- PSH (Państwowa Służba Hydrogeologiczna), 2013e: Mapa hydrogeologiczna Polski – Pierwszy Poziom Wodonośny (MhP PPW).
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/bazy_danych_mapy/mhp-ppw.html
- PSH (Państwowa Służba Hydrogeologiczna), 2013f: Obszary zagrożone podtopieniami.
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/bazy_danych_mapy/obszary-zagrozone-podtopieniami.html
- PSH (Państwowa Służba Hydrogeologiczna), 2013g: Przeglądarka mapowa e-PSH.
URL: <http://epsh.pgi.gov.pl/epsh/>
- PSH (Państwowa Służba Hydrogeologiczna), 2013h: System przetwarzania danych Państwowej Służby Hydrogeologicznej – Instrukcja użytkownika.
URL: http://spdps.pgi.gov.pl/PSHv7/Instrukcja_uzytkownika.pdf
- PSH (Państwowa Służba Hydrogeologiczna), 2013i: Zasoby dyspozycyjne.
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/bazy_danych_mapy/zasoby-dyspozycyjne.html
- PSH (Państwowa Służba Hydrogeologiczna), 2013j: Metadane.
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/metadane/
- PSH (Państwowa Służba Hydrogeologiczna), 2014a: Bazy danych i mapy.
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/bazy_danych_mapy/
- PSH (Państwowa Służba Hydrogeologiczna), 2014b: Pobory.
URL: http://www.psh.gov.pl/bazy_danych_mapy_i_aplikacje/bazy_danych_mapy/pobory.html
- PSH (Państwowa Służba Hydrogeologiczna), 2014c: System Przetwarzania Danych PSH.
URL: <http://spdps.pgi.gov.pl/PSHv7/>
- PSH (Państwowa Służba Hydrogeologiczna), 2014d: Jednolite części wód podziemnych (JCWPd) – cha-

- rakterystyka geologiczna i hydrogeologiczna. URL: http://www.psh.gov.pl/artykuly_i_publicacje/publikacje/jednolite-czesci-wod-podziemnych-charakterystyka-geologiczna-i-hydrogeologiczna.html
- Refsgaard J.C., Henriksen H.J., 2004: Modelling guidelines – terminology and guiding principles. *Advances in Water Resources* 27 (2004): 71–82.
- Reitz T., 2013: HALE 2.6.0 brings integration with FME.
URL: <http://blog.dhpanel.eu/2013/06/27/hale-2-6-0-release-integrates-with-fme/>
- Rekouts M., 2004: Application Programming Interface for XML DBMS: design and implementation proposal. (Sedna). URL: <http://modis.ispras.ru/downloads/publications/XMLAPI.pdf>
- Rice P. [ed.], The NGC Groundwater Data Standards Working Group, 1999 – The Australian National Groundwater Data Transfer Standard. National Groundwater Committee Arch., Canberra.
URL: <https://publications.csiro.au/rpr/pub?list=BRO&pid=procite:e473d507-228e-42b8-baae-e7c889011f85>
- Rossa M., Gogolek W., Łukasiewicz A., 2009: Geostandardy, metadane i dyrektywa INSPIRE. Poradnik metodyczny Zintegrowanego Systemu Kartografii Geologicznej IKAR. URL: http://www.pgi.gov.pl/pl/dokumenty-in-edycja/doc_view/243-geostandardy-metadane-i-dyrektywa-inspire.html
- Safe Software, 2014a: Achieve Harmonization with INSPIRE. URL: http://www.safe.com/solutions/for-initiatives/inspire/?utm_source=safe&utm_medium=blog&utm_campaign=inspire-summer-2014.
- Safe Software, 2014b: How to Create and Share INSPIRE-Compliant Data. URL: http://www.safe.com/webinars/how-to-create-and-share-inspire-compliant-data-with-fme/?utm_source=safe&utm_medium=blog&utm_campaign=inspire-summer-2014
- Safe Software, 2014c: INSPIRE Success Stories with FME.
URL: http://www.safe.com/webinars/inspire-success-stories-withfme/?utm_source=safe&utm_medium=blog&utm_campaign=inspire-summer-2014
- Sanchez-Clark T., 2007: XSLT Interview – Questions, Answers and Explanations. Wyd. Equity Press.
- Sedna Team, 2011a: Sedna Administration Guide. URL: <http://www.sedna.org/one-page/AdminGuide.html>
- Sedna Team, 2011b: Sedna Native XML Database Client/Server Protocol.
URL: <http://www.sedna.org/one-page/ClientServerProtocol.html>
- Sedna Team, 2011c: Sedna Programmer's Guide. URL: <http://www.sedna.org/progguide/ProgGuide.html>
- Seydi-Gheranghiyeh M., 2013: GeoModule. (in Base X Documentation).
URL: http://docs.basex.org/wiki/Geo_Module
- Simons B., Brodaric B., Grellet S., 2015: GWML2 GML Logical Model. GWIE2 Repository.
URL: http://external.opengeospatial.org/twiki_public/HydrologyDWG/GWML2-UML-Model
- 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>
- Skrzypczyk L., 1997: Zastosowanie banków danych hydrogeologicznych jako źródła informacji wyjściowej na potrzeby opracowań kartograficznych i dokumentacyjnych. *Przegląd Geologiczny* vol. 45, nr 9.
- Smith B., 2001: Fiat Objects. Topoi.
URL: <http://wings.buffalo.edu/philosophy/faculty/smith/articles/fatobjects.pdf>
- Sobell M.G., 1995: A Practical Guide to the UNIX System. Third Edition. Wyd. Benjamin/Cummings.
- Sobell M.G., 2012: Fedora i Red Hat Enterprise Linux, praktyczny przewodnik. Wydanie VI. Wyd. Helion.
- Stenzel P., Berestka A., 1979: System Hydro – regionalne banki danych hydrogeologicznych. Wydawnictwo Geologiczne, Warszawa.
- Struckmeier W.F., Margat J., 1995: Hydrogeological Maps – A Guide and a Standard Legend. International Association of Hydrogeologists (IAH). International Contributions to Hydrogeology. Vol. 17, Heise (Hannover), (PDF, 10 MB). URL: http://www.bgr.bund.de/EN/Themen/Wasser/Projekte/laufend/Beratung/Ihme1500/standard_legend_hydro_maps.pdf?__blob=publicationFile&v=2
- Tennakoon W., 2003: Visualization of GML data using XSLT.
URL: http://www.itc.nl/library/Papers_2003/msc/gim/tennakoon.pdf
- Tennison J., 2005: Beginning XSLT 2.0: From Novice to Professional. Wyd. Apress.
- Tidwell D., 2008: XSLT. Second Edition. Wyd. O'Reilly Media Inc.
- Van den Broecke J., 2011: INSPIRE FOSS Workshop. URL: <https://inspire-foss.googlecode.com/svn-history/r347/trunk/doc/workshop/workshop-nl-11030-v1.pdf>
- Van den Broecke J., 2013: Taming Rich GML with Stetl – A lightweight Python Framework for Geospatial

- ETL. URL: <http://2013.foss4g.org/conf/programme/presentations/156/>
- Van den Broecke J., 2014: Welcome to Stetl. URL: <http://stetl.readthedocs.org/en/latest/>
- Van Otegem M., 2002: SAMS Teach Yourself XSLT in 21 Days, Wyd. Sams.
- W3C (World Wide Web Consortium), 1999: XSL Transformations (XSLT) Version 1.0.
URL: <http://www.w3.org/TR/xslt>
- W3C (World Wide Web Consortium), 2005: Document Object Model (DOM).
URL: <http://www.w3.org/DOM/>
- W3C (World Wide Web Consortium), 2008: Extensible Markup Language (XML) 1.0 (Fifth Edition).
URL: <http://www.w3.org/TR/2008/REC-xml-20081126/>
- W3Schools (Refsnes Data), 2000: XPath Tutorial. URL: <http://www.w3schools.com/xpath/>
- Wang W., Wang F., Qian Z., Zhang L., 2011: GML Data Management: Framework and Prototype. In DASFAA Workshops (2011): 101-111.
URL: http://link.springer.com/chapter/10.1007%2F978-3-642-20244-5_10#page-1
- 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>
- Williams I., 2009: Beginning XSLT and XPath: Transforming XML Documents and Data. Wyd. Wiley Publishing Inc.
- Wojda P., 2008: Hydrogeological data modelling in groundwater studies. Doctoral dissertation. URL: http://bictel.ulg.ac.be/ETD-db/collection/available/ULgetd-02262009-181604/unrestricted/PhDThesis_Piotr_WOJDA.pdf
- Wojda P., Brouyere S., Dassargues A., 2010: Geospatial information in hydrogeological studies.
URL: http://orbi.ulg.ac.be/bitstream/2268/6168/1/hsa304_PostprintAuthor.pdf
- Woolf A., 2009: Enterprise Architect instructions, STFC Rutherford Appleton Laboratory.
URL: http://http://wiki.services.eportal.org/tiki-download_wiki_attachment.php?attId=732
- Woźnicka M., Skrzypczyk L., Nawrocki J., 2013: Zadania Państwowej Służby Hydrogeologicznej w 2013 roku – Synteza. Wykonano w ramach realizacji przedsięwzięcia: Zadania państwowej służby hydrogeologicznej w latach 2012-2014. URL: <http://www.psh.gov.pl/plik/id,7070.pdf>
- 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>
- Zhanga S., Gana J., Xua J., Lva G., 2008: Study on Native XML Database Based GML Storage Model.
URL: http://www.isprs.org/proceedings/XXXVII/congress/4_pdf/09.pdf
- Zhu F., Chen H., Guan J., 2010: Querying GML Documents: An XQuery based Approach.
URL: <http://www.complang.tuwien.ac.at/WWV2010/proc/regular1.pdf>