



POLSKIE
TOWARZYSTWO
INFORMACJI
PRZESTRZENNEJ

ROCZNIKI 2010 GEOMATYKI

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

Tom VIII
Zeszyt 3(39)
Warszawa

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

Jerzy Gaździcki

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

Spis treści

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

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

9. Testowanie geoserwera

Jednym z kluczowych etapów opisywanego tu prac było testowanie geoserwera OakHills w różnych warunkach i pod różnymi kryteriami. Do testowania geoserwera wykorzystano, obok danych hydrogeologicznych i geologicznych, różnorodne dane wchodzące w zakres tematyczny załącznika I dyrektywy INSPIRE. Ponieważ dla tych danych są już opracowane specyfikacje INSPIRE, testowanie obejmowało także metodykę i technologię transformacji polskich zbiorów do struktur określonych w tych specyfikacjach.

9.1. Zbiory danych testowych

W ramach prac testowych pozyskano dane dla obszaru poligonu testowego Dębowe Góry (*OakHills*) położonego na północy Mazowsza (rys. 25) opisanego we wstępie rozdz. 8. Obszar ten obejmuje około 273 km², a w przypadku mniej szczegółowych danych 4 arkusze mapy w skali 1:50 000, co odpowiada w przybliżeniu powierzchni 1600 km². W przypadku braku odpowiednich danych testowych z tego poligonu, zostały pozyskane dane z innych miejsc. Pozyskano następujące zbiory danych:

Dane typu pokrycie (*coverage*) w formie rastrowej:

- ortofotomapa satelitarna Landsat 7 ETM+,
- dane modelu powierzchni terenu SRTM-3 DTED Level 1 opracowane przez NASA i USGS (Michalak, 2004a),
- szczegółowe dane wysokościowe pochodzące z opracowań fotogrametrycznych i innych.
- dane CORINE Land Cover 2001 w wersji rastrowej.

Dane typu wyróżnienia (*feature*) w formie wektorowej:

- dane cyfrowe dotyczące obszarów chronionych w formacie SHP (0,1 MB) z Generalnej Dyrekcji Ochrony Środowiska,
- dane cyfrowe Mapy Głównych Zbiorników Wód Podziemnych w skali 1:50 000 w formacie SHP (0,5 MB) z Państwowego Instytutu Geologicznego,
- dane cyfrowe Mapy Jednolitych Części Wód Podziemnych w skali 1:50 000 w formacie SHP (0,1 MB) z Państwowego Instytutu Geologicznego,
- dane cyfrowe MHP (Mapy Hydrogeologiczna Polski) w skali 1:50 000 w formacie SHP (5 MB) z Państwowego Instytutu Geologicznego,
- dane cyfrowe MHP (Mapy Hydrogeologiczna Polski) w skali 1:50 000 w formacie MDB (4 MB) z Państwowego Instytutu Geologicznego,
- dane cyfrowe MPHP (Mapy Podziału Hydrograficznego Polski) w skali 1:50 000 w formacie SHP (9 MB) z Krajowego Zarządu Gospodarki Wodnej,
- granice działek gruntowych w formacie DXF (53 MB) z Powiatowego Zasobu Danych Geodezyjnych i Kartograficznych w Przasnyszu,
- LMN (leśna mapa numeryczna) w formacie SHP (85 MB) z Nadleśnictwa Przasnysz.

W przypadku braku odpowiednich danych z obszaru poligonu badawczego Dębowe Góry, pozyskano dane dla innych obszarów:

- dane cyfrowe Mapy Ostoi Przyrody w Polsce w skali 1:50 000 w formacie MDB (18 MB) z Generalnej Dyrekcji Ochrony Środowiska,
- dane cyfrowe Mapy Natura 2000 w Polsce w skali 1:50 000 w formacie SHP (4 MB) z Generalnej Dyrekcji Ochrony Środowiska,
- dane cyfrowe MGŚP (Mapy Geośrodowiskowej Polski) w skali 1:50 000 w formacie SHP (2 MB) z Państwowego Instytutu Geologicznego.

Pozyskanie tak dużej liczby różnorodnych zbiorów danych było możliwe dzięki współpracy z innymi instytucjami w ramach wspólnych prac testowych nad transformacją polskich danych do specyfikacji określonych w dokumentach dyrektywy INSPIRE. Prace wykonane w ramach projektu stanowiły podstawę do opracowania raportów dotyczących możliwości transformacji polskich danych do schematów INSPIRE.

Powyższe zbiory danych, zarówno w formie pokryć (rastrowej), jak i w formie wyróżnień (wektorowej), po dokładnej analizie zostały przetransformowane do jak najbardziej jednolitej formy pod względem układu odniesienia i formy zapisu. Przyjęto, że najbardziej odpowiednią formą dla danych wektorowych będzie język GML, układ odniesienia EPSG:2180 (PUWG 1992) i kod danych tekstowych UTF-8. W tym przypadku kryterium wyboru była zgodność z normami ISO 19100, specyfikacjami OGC i specyfikacjami INSPIRE. W przypadkach, gdy nie można było utworzyć poprawnych zapisów w języku GML, dane zostały zapisane w popularnym i prostym formacie ESRI Shapefile, z zachowaniem układu EPSG 2180 i kodowaniem UTF-8. Jednak format Shapefile pozwala na zapis danych atrybutowych jedynie w prostym płaskim modelu tabeli dBASE. Dodatkowym problemem jest brak informacji o sposobie kodowania atrybutów tekstowych, co w przypadku polskich danych geoprzestrzennych jest informacją konieczną, ze względu na stosowanie różnych sposobów kodowania.

9.2. Transformacja danych testowych do języka GML, w tym do modeli INSPIRE

Z powodów szczegółowo przedstawionych w rozdz. 7.1. nie można było danych testowych przetransformować do postaci w pełni zgodnych ze specyfikacjami danych INSPIRE. Z tych samych powodów brak jest jakichkolwiek przykładów danych INSPIRE, które mogłyby być użyte do testowania. Wymagania INSPIRE dotyczące struktur i treści danych są określone w specyfikacjach danych w formie tekstowej, ich precyzyjny zapis jest zawarty w modelach danych zapisanych w diagramach klas języka UML i schematach aplikacyjnych XSD języka GML. Wykorzystując te schematy, dane testowe z zakresu poszczególnych tematów zostały przetransformowane do postaci możliwie jak najbardziej zbliżonej do wymaganej, jednak stopień tego zbliżenia dla poszczególnych rodzajów danych był różny. Dane z zakresu tematów, dla których obecnie nie ma specyfikacji, zostały zapisane w języku GML jedynie z niewielkimi zmianami ich modeli i z uwzględnieniem, że zmiana modeli może być przeprowadzona później, gdy będą już znane odpowiednie specyfikacje. Ze względu na wymagania geoserwera i specyfikę przeprowadzanych testów, dane były także przetransformowane do formatu ESRI Shapefile. Bardziej szczegółowy opis problematyki transformacji polskich danych do modeli INSPIRE znajduje się w monografii pt. „Modelowanie danych przestrzennych” w rozdz. „Transformacja polskich danych przestrzennych do modeli INSPIRE” (Michalak, 2010).

9.3. Testowanie poprawności funkcji geoserwera

Wymienione wcześniej dane geoprzestrzenne zostały wykorzystane do prac testowych w zakresie poprawności funkcji budowanego systemu. W wyniku prac testowych wybrano najbardziej odpowiedni zestaw wszystkich elementów geoserwera. Końcowa wersja oprogramowania systemu geoserwera oparta jest na platformie sprzętowej z procesorami typu Intel i SUN Sparc, a platforma systemowa na systemie operacyjnym GNU/Linux Debian 5.0 dla obu tych platform sprzętowych. Oprogramowanie geoserwera w końcowej testowej wersji składało się z następujących głównych elementów:

- serwer komunikacji HTTP – Apache 2,
- oprogramowanie dynamicznego generowania odpowiedzi – PHP Engine v. 5.2.6,
- oprogramowanie dla odbioru danych po stronie klienta – przeglądarki WWW Firefox.
- podstawowy komponent geoserwera – MapServer v. 5.6.1,
- system wspomagania komunikacji klient-serwer – moduł PHP/MapScript,
- środowisko budowy aplikacji tego modułu – p.mapper – Framework v. 4.0.0,

Dla poprawnego funkcjonowania powyższych komponentów systemy operacyjne były uzupełnione następującymi rozszerzeniami aplikacyjnymi:

- biblioteki oprogramowania graficznego: libpng, libtiff, libgeotiff, libjpeg, libgif i pdflib,
- biblioteki wspomagające: freetype dla tekstów, GD dla generowania obrazów wynikowych, zlib dla kompresji plików, AGG w celu uzyskania obrazów o wysokiej jakości rozdzielczej, libpq dla komunikacji z bazą danych PostgreSQL z modułem geoprzestrzennych PostGIS.

Dla osiągnięcia pełnych możliwości funkcjonalnych i pełnej wydajności budowanego systemu kluczowe były następujące komponenty:

- biblioteka OGR dla obsługi 18 różnych formatów wektorowych danych geoprzestrzennych,
- biblioteka GDAL dla obsługi 42 formatów danych rastrowych (macierzowych),
- biblioteka Proj4 dla transformacji danych przestrzennych (wektorowych i rastrowych) pomiędzy różnymi układami odniesienia i odwzorowania,
- biblioteka cURL dla obsługi wielu protokołów internetowych umożliwiającą zastosowanie podstawowych standardów wymiany danych przestrzennych w infrastrukturach geoinformacyjnych, takich jak: WMS, WFS i WCS.

W części E literatury podano odwołania do dokumentacji wszystkich składników oprogramowania systemu geoserwera OakHills. Lista ta zawiera 39 pozycji, co wymownie świadczy o złożoności tego systemu.

W ramach tego etapu prac przetestowano poprawność wzajemnej współpracy dużej liczby wymienionych powyżej komponentów, w różnych sytuacjach i dla różnych danych. Testy wykazały wiele usterek i błędów – różnych w różnych środowiskach systemów operacyjnych. W wyniku tych testów, między innymi, zrezygnowano z platformy systemowej SUN Solaris 10 z powodu poważnych problemów programistycznych dotyczących bardzo istotnej w tym przypadku biblioteki GDAL, co opisano w rozdz. 8.10.

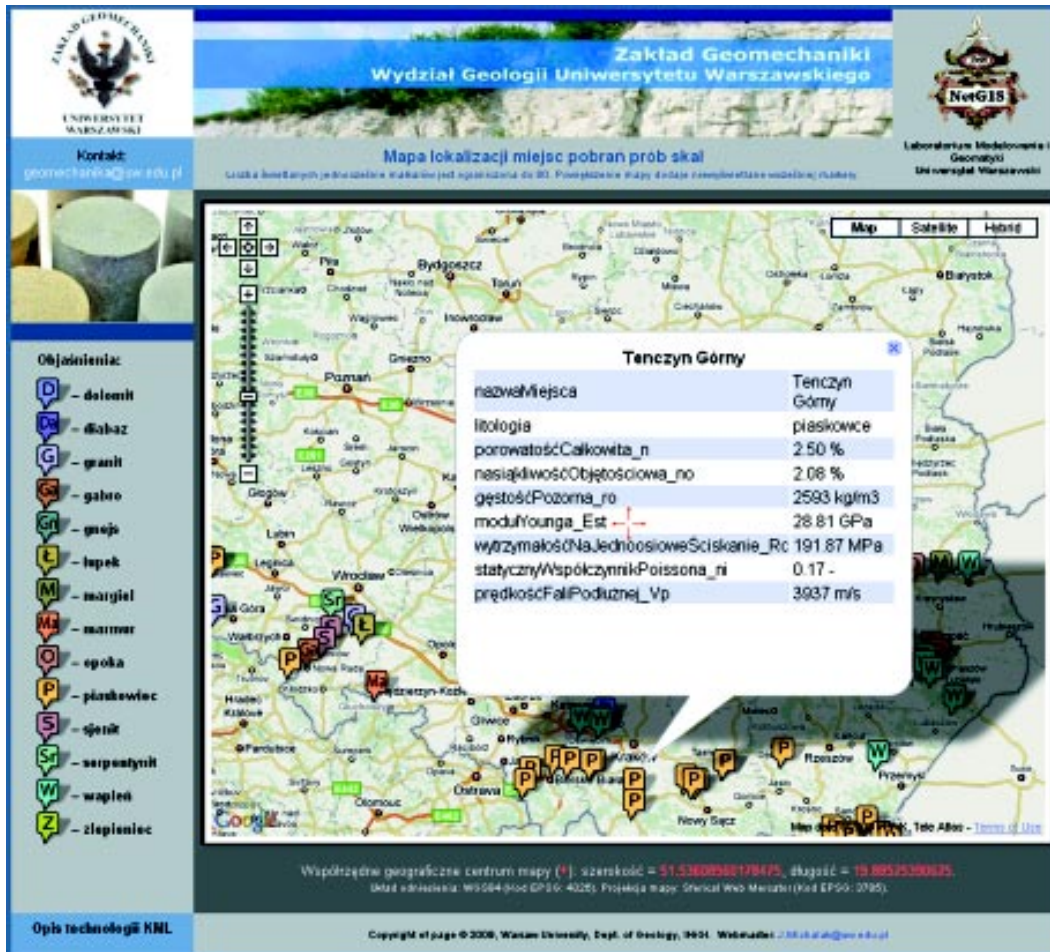
9.4. Testowanie wydajności w różnych konfiguracjach

Rozporządzenia wykonawcze Komisji Europejskiej dotyczące realizacji postanowień dyrektywy INSPIRE określają ściśle wymagania dotyczące jakości usług sieciowych, jakie mają spełniać geoserwery działające w węzłach tej infrastruktury:

- w zakresie dostępności usług – 99% czasu w ciągu roku,
- w zakresie wydajności usługi WMS – mniej niż 5 sekund oczekiwania na odpowiedź w postaci obrazu o wielkości 470 KB, z pominięciem obciążeń szczytowych, które stanowią nie więcej niż 10% czasu pracy w ciągu roku,
- w zakresie pojemności serwera dla usługi WMS – jednoczesne realizowanie przynajmniej 5 poleceń równolegle.

Z uwagi na eksperymentalny charakter opisywanego tu geoserwera, jego obciążenie robocze było bardzo niewielkie w porównaniu z serwerem pracującym w trybie w pełni operacyjnym. Z tego względu testowanie wydajności nie mogło być przeprowadzone całkowicie zgodnie z wymaganiami INSPIRE. Testy miały za zadanie sprawdzić różne sposoby pomiaru parametrów określających jakość usług. Na tej podstawie została opracowana koncepcja pracy systemu, którego zadaniem jest zdalne testowanie jakości usług, która będzie w przyszłości wykorzystana w projekcie takiego systemu testującego.

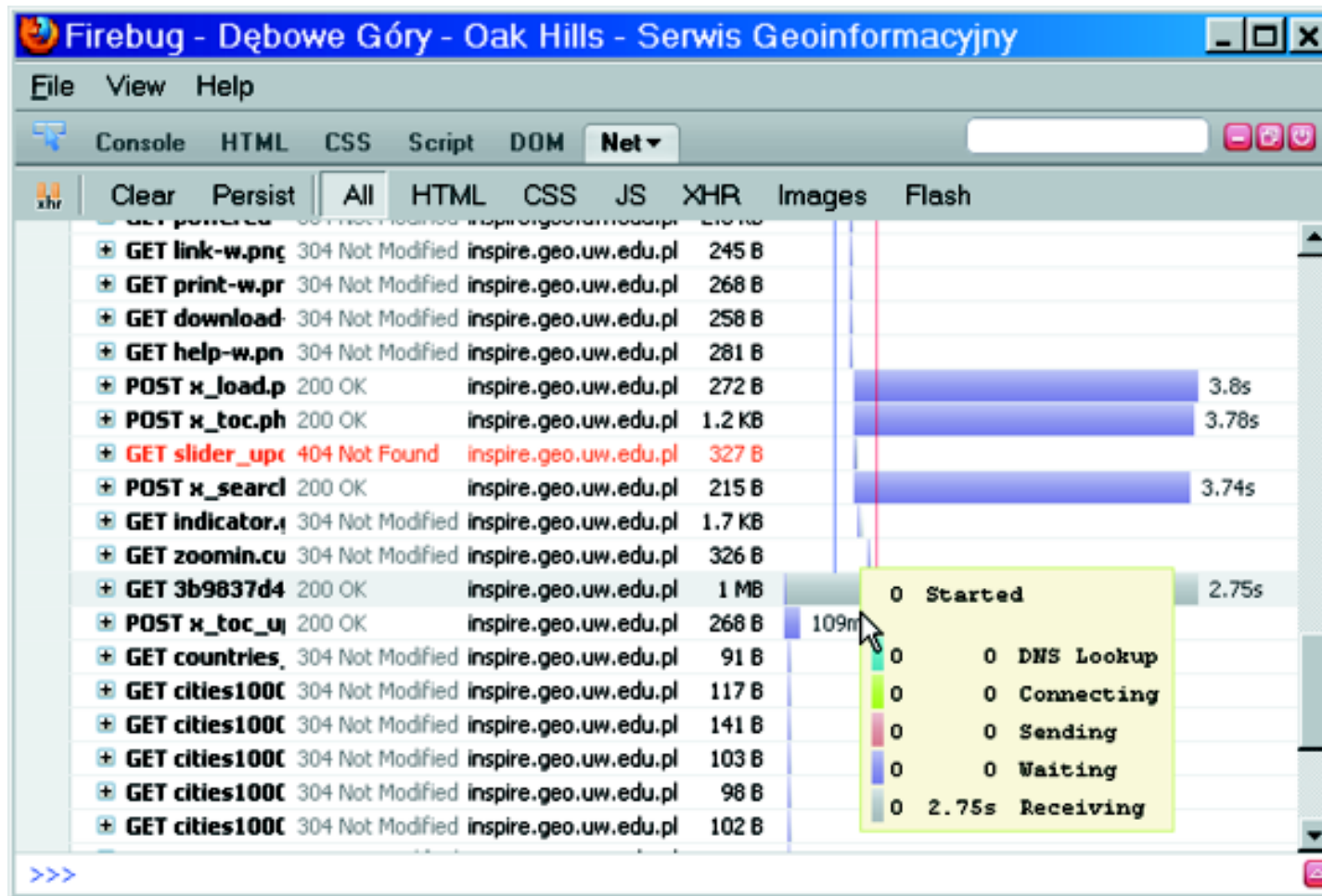
W przypadku opisywanego tu geoserwera eksperymentalnego można było zastosować jedynie proste metody testowe, polegające na „ręcznym” mierzeniu pojedynczych odpowiedzi serwera na przesyłane do niego polecenia. Testy były wykonywane przy pomocy programu narzędziowego Mozilla FireBug, współpracującego z przeglądarką WWW Mozilla Firefox. Przykład takiego testu przedstawiono na rysunku 42. Wynik testu wydajności geoserwera pokazał, że czas odpowiedzi na polecenie przesłania map w postaci obrazu o wielkości 1 MB wyniósł 2,75 sekundy. Jednocześnie test wykazał brak pewnego elementu graficznego w interfejsie WWW (komunikat w kolorze czerwonym), co po jego wykryciu mogło być łatwo poprawione. Test ten dotyczył jedynie danych pobranych z zasobów własnych geoserwera. W przypadku przesyłania obrazu, który zawiera także dane kaskadowane z innych geoserwerów, czas ten jest dłuższy i w wielu przypadkach oczekiwanie testowanego systemu na odpowiedź z innego geoserwera przekracza ustalony limit czasu, co jest traktowane jako brak odpowiedzi.



Rys. 41. Wyniki badań laboratoryjnych parametrów geomechanicznych próbek skał pobranych z różnych kamieniołomów na obszarze Polski południowej.

Przykład powiązania technologii geoserwera OakHills z technologią opartą na języku KML.

Wyświetlone na dole współrzędne odnoszą się do środka czerwonego krzyżyka



Rys. 42. Okno programu narzędziowego FireBug współpracującego z przeglądarką WWW FireFox. Program ten był wykorzystywany do testowania szybkości odpowiedzi geoserwera na wysłane do niego polecenia. Zaznaczone na czerwono polecenie GET slide_updown nie zostało zrealizowane poprawnie, ponieważ po stronie serwera nie było tego elementu. Pobranie obrazu o wielkości 1 MB trwało 2,75 sekundy