



Seminarium:
‘Bussines intelligence’ a hydrogeologia

12 października 2010 r. – Warszawa



Główne zagadnienia seminarium

1. Oprogramowanie wykorzystywane w PSH
2. Obecne kierunki rozwoju
3. Rozwiązania na przyszłość
4. Business intelligence
5. Data mining
6. Sztuczne sieci neuronowe
7. Podsumowanie
8. Trendy

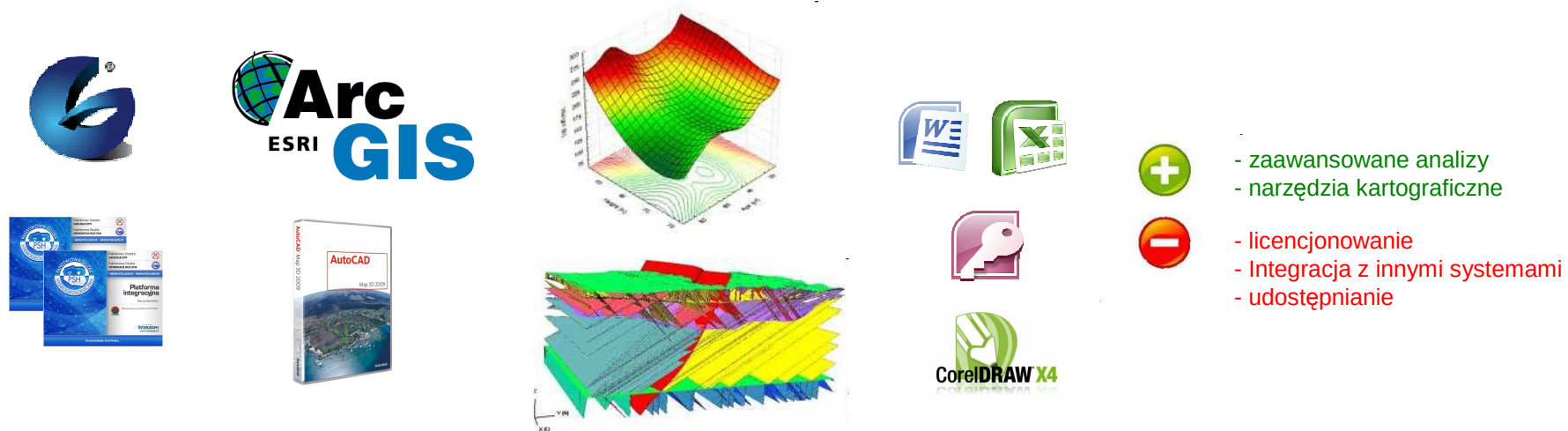
Geoinformacja

Systemy GIS

Eksploracja danych

Podstawowe narzędzia w pracy hydrogeologa

Grupa 1 : Narzędzia typu desktop



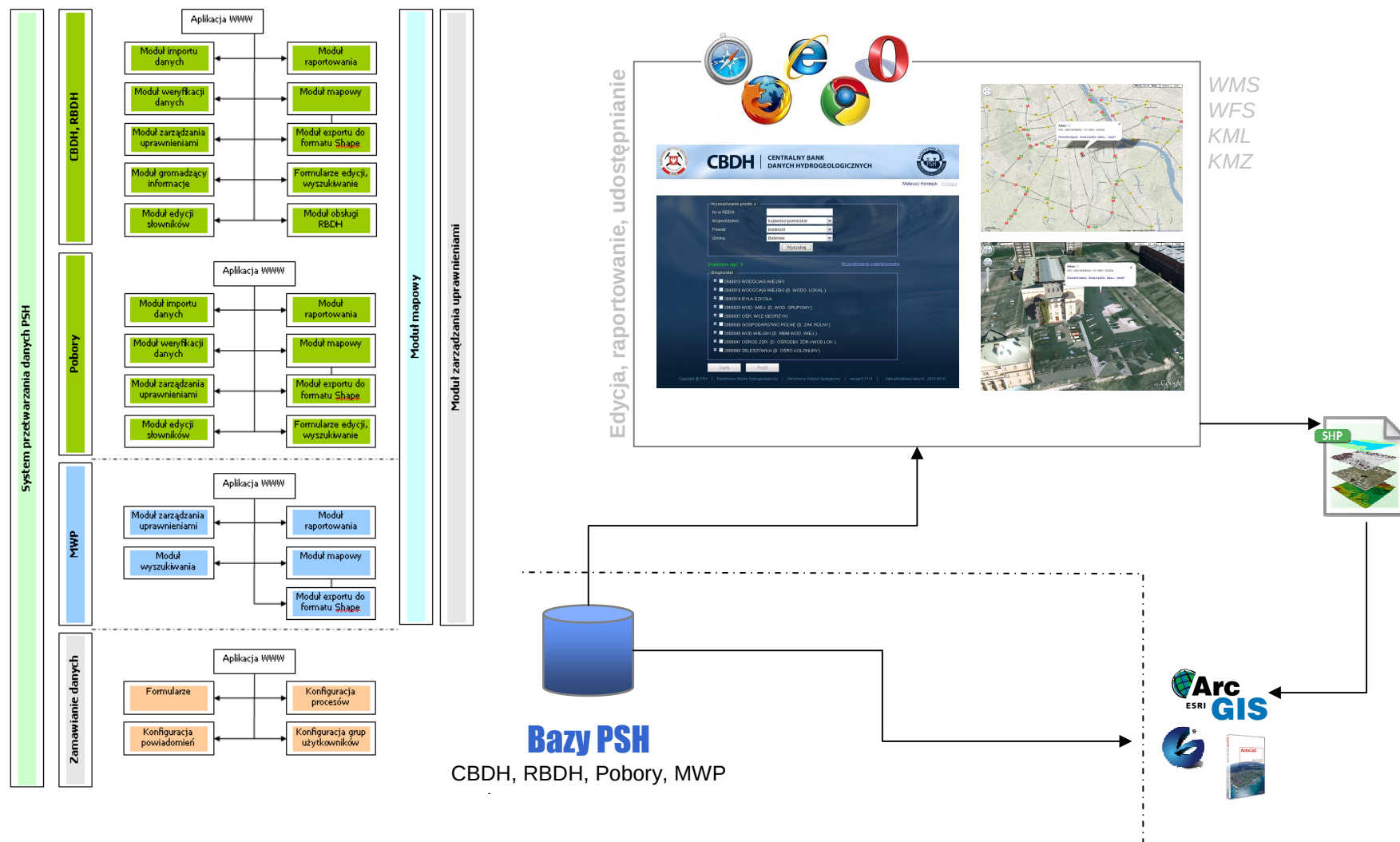
Grupa 2 : Narzędzia internetowe



Grupa 3 : Rozwiązania serwerowe



Kierunki rozwoju oprogramowania narzędziowego w PSH cz. 1



Web-mapping (przykładowe zastosowanie Google Maps)

Kontakt:
geomechanika@uw.edu.pl

Zakład Geomechaniki
Wydział Geologii Uniwersytetu Warszawskiego

Mapa lokalizacji miejsc pobrań skal

Liczba świetlnych jednocześnie makarów jest ograniczona do 80. Powiększenie mapy dodaje niewyświetlane wcześniej makary.

Objaśnienia:

- D – dolomit
- Da – diabaz
- G – granit
- Ga – gabra
- Gn – gnejs
- L – lupek
- M – margiel
- Ma – marmur
- Op – opoka
- P – piaskowiec
- S – sjenit
- Sr – serpentynit
- W – wapień
- Z – zlepieniec

Braszowice	
nazwaMiejsca	Braszowice
litologia	gabra
porowatośćCałkowita_n	3.79 %
nasiąkliwośćObjętościowa_no	0.99 %
gęstośćPozorna_ro	2934 kg/m3
modułYounga_Est	34.22 GPa
wytrzymałośćNaJednoosioweŚciskanie_Rc	132.48 MPa
statycznyWspółczynnikPoissona_ni	0.13 -
prędkośćFaliPodłużnej_Vp	5160 m/s

Współrzędne geograficzne centrum mapy (P): szerokość = 51.850051935216926, długość = 18.52234921075.
 Układ odniesienia: WGS84 (Kod EPSG: 4326). Projektcja mapy: Spherical Web Mercator (Kod EPSG: 3785).

Copyright of page © 2009, Warsaw University, Dept. of Geology, IHiG. Webmaster: J.Michalak@uw.edu.pl

Dane z zakresu geomechaniki

Shape



transformacja

GML

przetworzenie



zapis

KML

Dane zapisane w formacie KML przedstawione w postaci 'chmurki'

Informacja może być obrazem lub hiperłączem.

Dostępne reprezentacje : punkt, linia, wielobok.

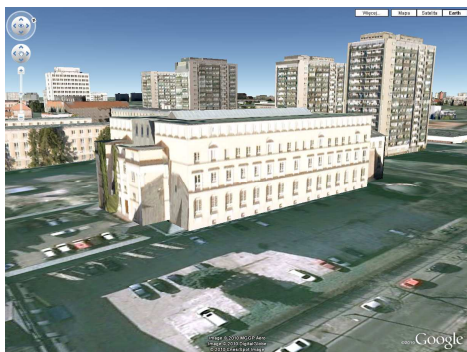
Źródło: <http://netgis.geo.uw.edu.pl/geomechanika/> (Laboratorium Modelowania i Geomatyki, Uniwersytet Warszawski)
 Uwarunkowania technologiczne w rozwoju zastosowań narzędzi Web-GIS i Web-mapping, Janusz Michalak, 2009.

Kierunki rozwoju - Następny krok

Obszary zainteresowania

Wizualizacja danych:

- 3D
- mapa/diagram/wykres



Źródło : <http://images.google.pl>

hydrogeologia ?

Business intelligence
Data mining



integracja i harmonizacja danych
usługi sieciowe/serwisy
zgodność ze standardami
model danych hydrogeologicznych

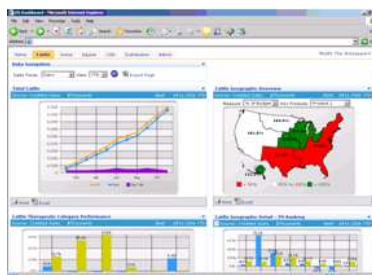
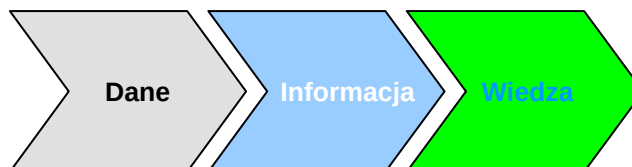


WMS (1997 r. – wstępna
specyfikacja OGC, eksperyment w
Cincinnati*), WFS, CSW

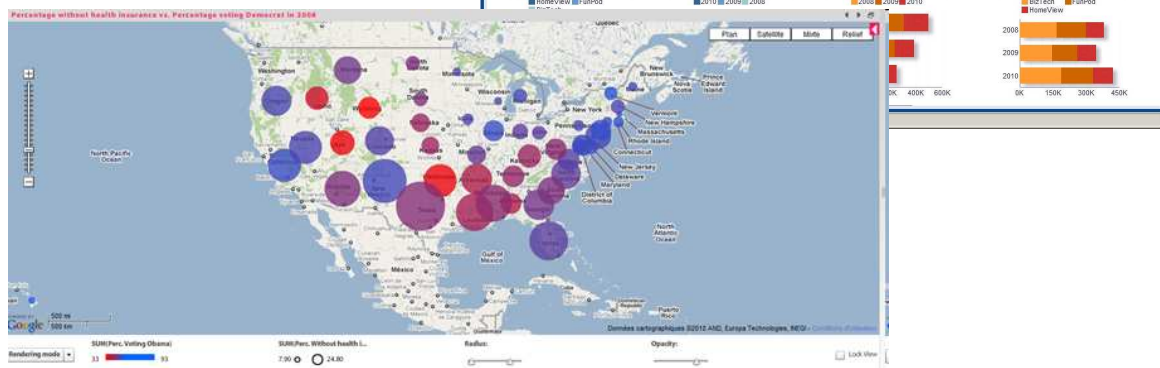
WPS (Web Processing Service)

* Uwarunkowania technologiczne w rozwoju zastosowań narzędzi Web-GIS i Web-mapping, Janusz Michalak, 2009.

Business Intelligence, czyli analityka biznesowa



Wizualizacja stanu aktualnego realizowana jest w postaci obrazkowej. **Kokpit menedżerski** (*management dashboard*) to atrakcyjny sposób prezentacji wyników – wizualizacja danych i raportów w postaci podobnej do pulpitów sterowniczych.

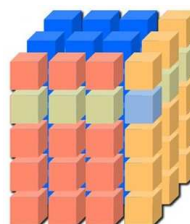


Business Intelligence, czyli analityka biznesowa



Technologie Business Intelligence służą do wspomagania decyzji biznesowych dzięki inteligentnemu wykorzystaniu dostępnych zasobów informacyjnych. Zadaniem nowoczesnych systemów jest umożliwienie łatwego i bezpiecznego dostępu do informacji, obsługa procesów ich analizy i dystrybucja raportów w obrębie przedsiębiorstwa, kooperantów bądź partnerów.

Koncepcja jest następująca: system BI generuje standardowe raporty lub wylicza kluczowe wskaźniki efektywności działania (Key Performance Indicators) na podstawie których stawia się hipotezy, po czym weryfikuje się je poprzez wykonywanie szczegółowych "przekrojów" danych. Do tego służą różnego rodzaju narzędzia analityczne (np. OLAP, data mining).



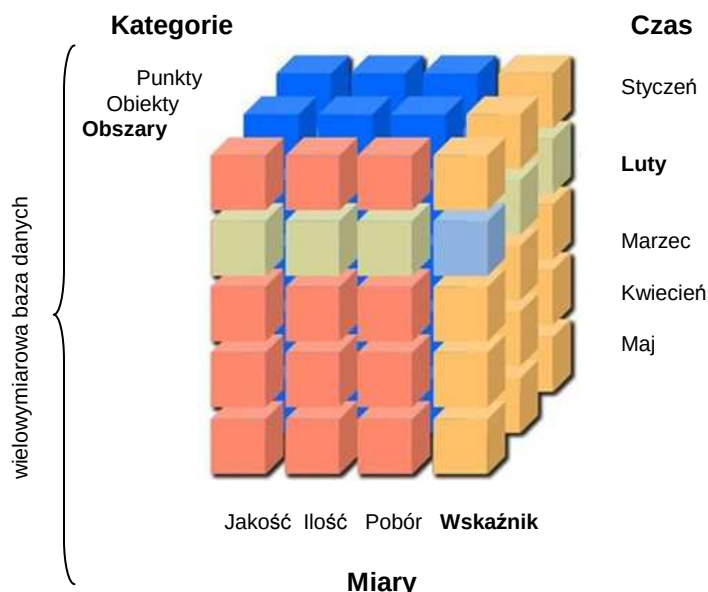
W celu odpowiedniej prezentacji danych użytkownikom stosuje się technikę zwaną jako Business Intelligence [BI], czyli zbiór środków pozwalających na dostarczanie społeczności użytkowników odpowiednich danych w odpowiedni sposób. BI mogą przyjąć rozmaite postaci, jak **OLAP (Online Analytical Processing)** i jego pochodne, które pośredniczą w dostępie do hurtowni danych za pomocą interfejsów instalowanych u użytkowników.

Do BI zalicza się również rozwiązania podejmujące decyzje na podstawie zadanych algorytmów postępowania.

Business Intelligence



OLAP (OnLine Analytical Processing)



- umożliwia wyświetlanie i oglądanie danych z różnych punktów widzenia
- do budowy potrzebny relacyjny model danych

Dotychczasowa technologia:

- implementacja (długa i skomplikowana)
- Długi czas oczekiwania na wynik
- mało elastyczne analizy (liczba wymiarów z góry określona)
- nieefektywna technologicznie

Przetwarzanie w pamięci (**in-memory**)

- Prosty model danych, wystarcza proste odwzorowanie relacji źródłowych
- Bardzo szybkie zasilanie modelu analitycznego
- Model asocjacyjny = ogromna kompresja danych + wizualizacja zależności
- Wszystkie agregaty i wyrażenia wyliczane "na żądanie"
- Sekundowe czasy odpowiedzi systemu
- Najlepszy stosunek kosztów do wydajności
- Radykalne skrócenie czasu wdrożenia
- Eliminacja większości ograniczeń starej technologii OLAP

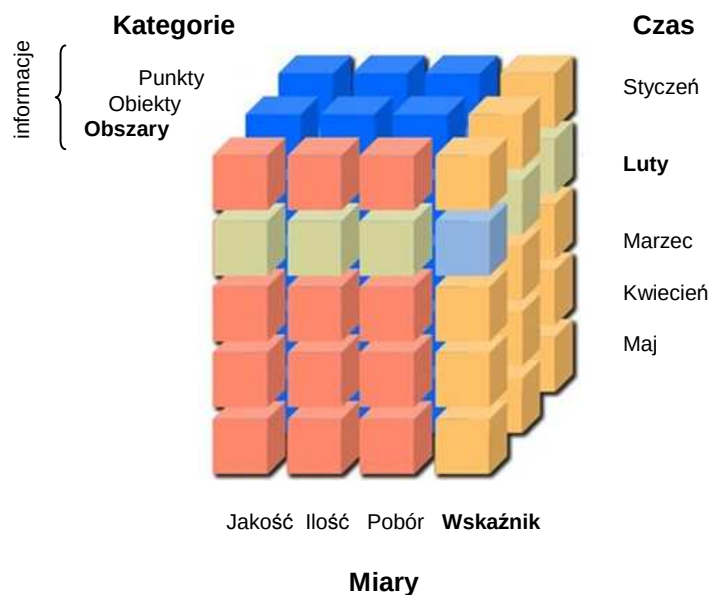
Źródło: Hogart

*Technologia **in-memory** zapowiada przełom w rozwiązaniach Business Intelligence. Według Gartnera, do 2012 roku 70% największych firm na świecie będzie wykorzystywać BI przetwarzających dane w pamięci operacyjnej. Przede wszystkim jednak nowa, tańsza technologia powinna spopularyzować rozwiązania BI w sektorze małych i średnich firm.*

Business Intelligence



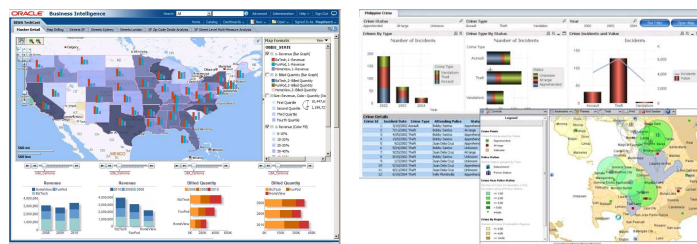
OLAP (OnLine Analytical Processing)



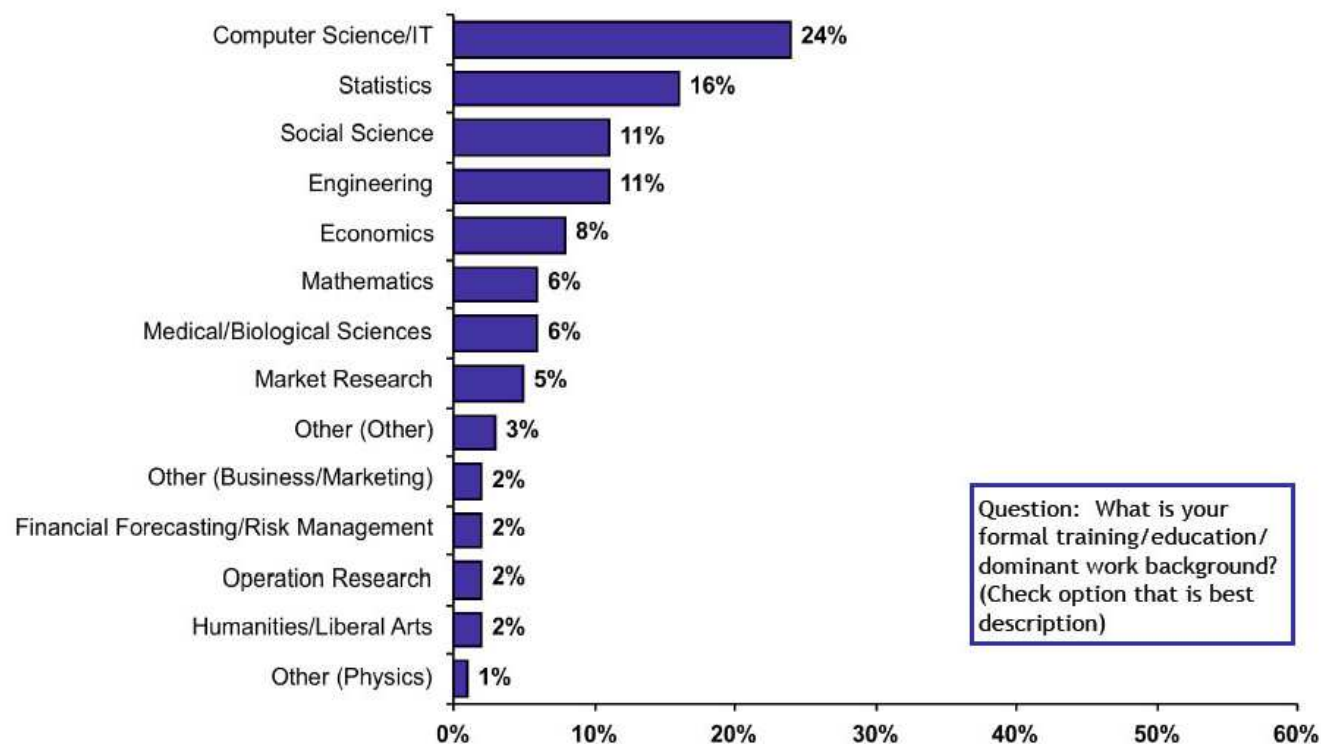
Zalety w raportowaniu nowych rozwiązań

- Możliwość analizy danych na najniższym poziomie szczegółowości;
- Czas załadowania danymi całego modelu (skrócenie o rząd wielkości);
- Szybkość reakcji systemu liczona w sekundach lub ułamkach sekund;
- Mniejsze wymagania sprzętowe;
- Krótsze szkolenia lub nawet brak konieczności przeprowadzenia szkoleń dla użytkowników końcowych;
- Brak ograniczeń na liczbę wymiarów podlegających jednoczesnej analizie (nawet 600 - 800 wymiarów bez spadku wydajności);
- Krótszy czas tworzenia aplikacji i wdrożenia;
- Mniejsze wymagania w zakresie znajomości SQL, bazy danych, języków programowania.

Źródło: Hogart

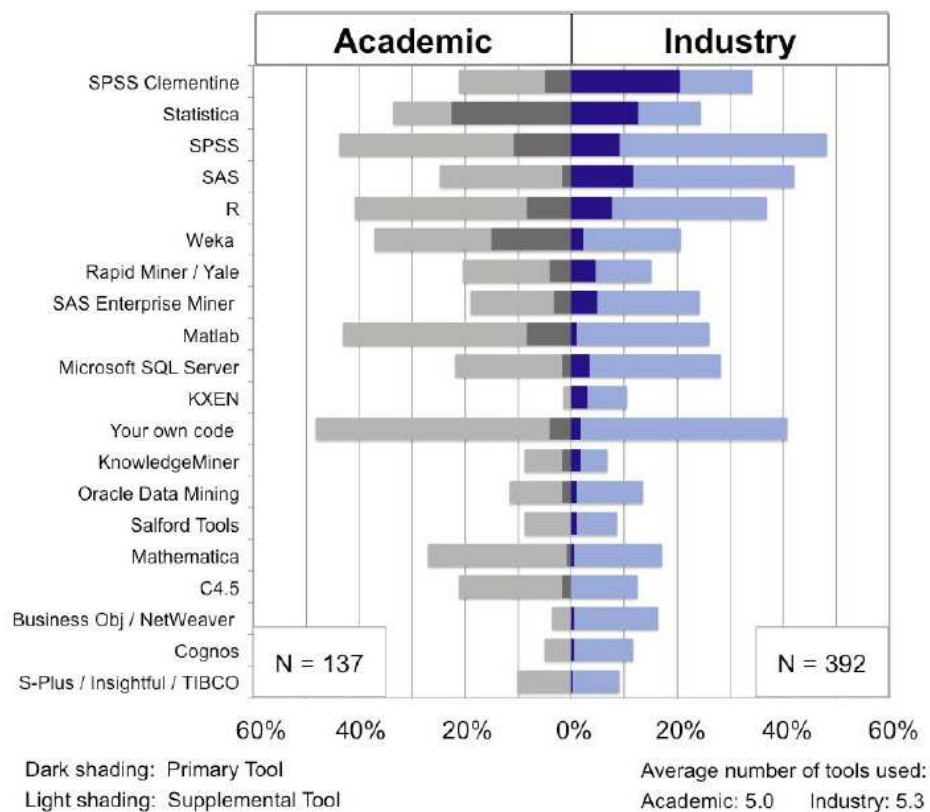


Business Intelligence okiem użytkowników



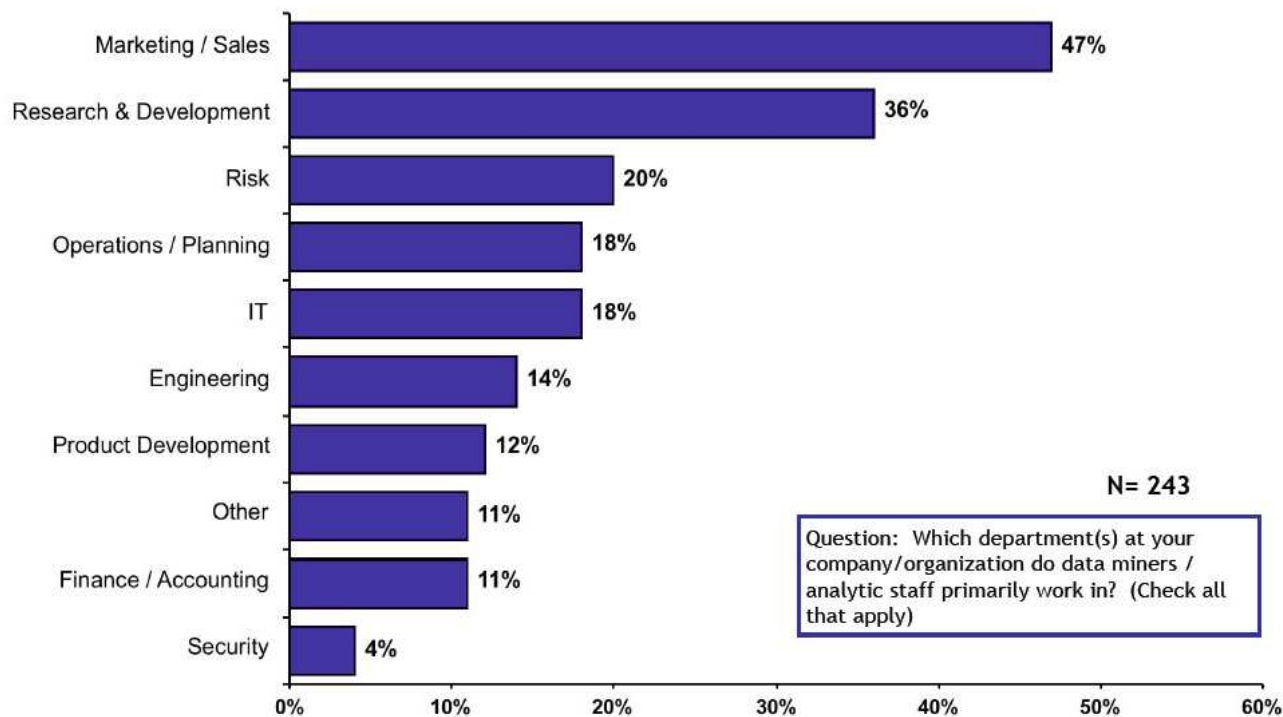
Źródło : 3rd Annual Data Miner Survey – 2009 Survey Summary Report, Rexer Analytics, 2010

Business Intelligence okiem użytkowników



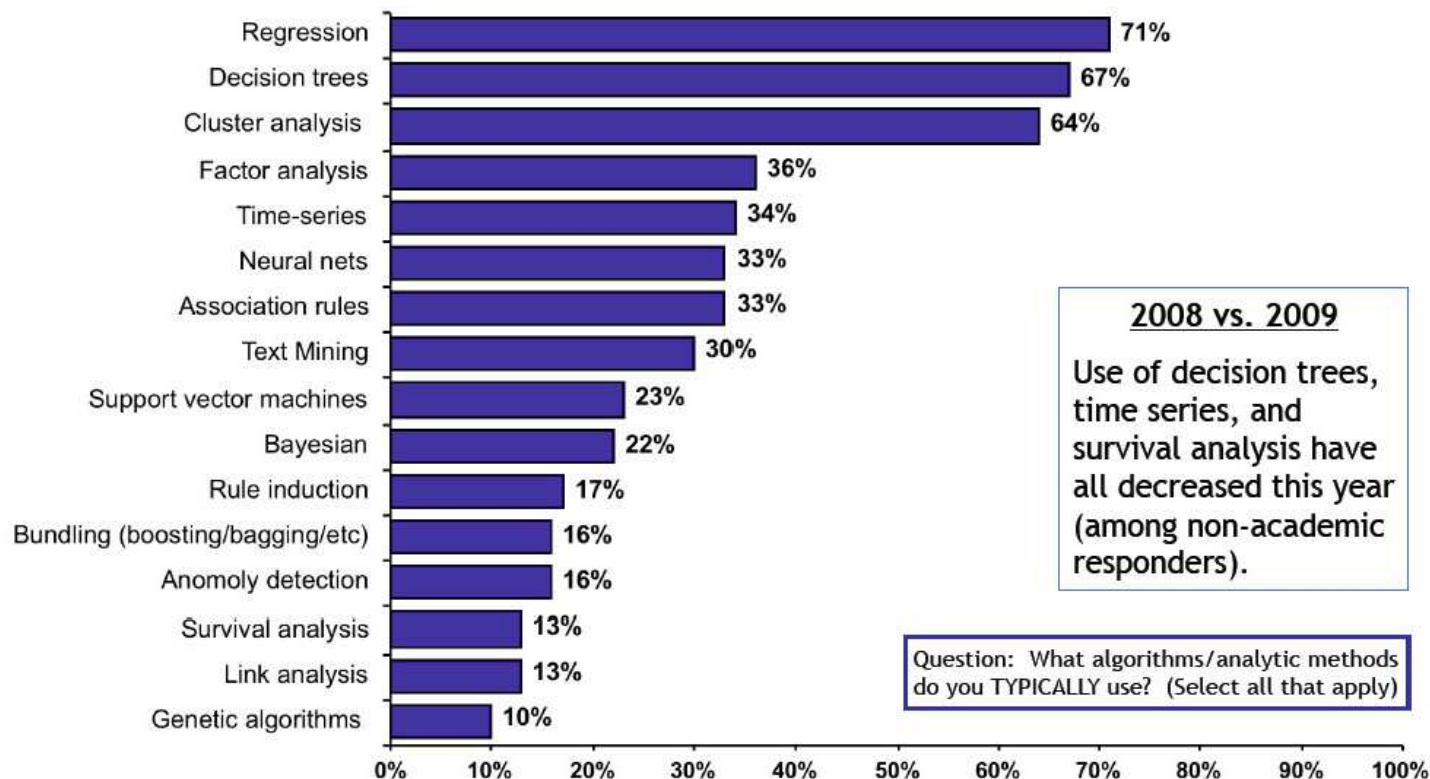
Źródło : 3rd Annual Data Miner Survey – 2009 Survey Summary Report, Rexer Analytics, 2010

Business Intelligence okiem użytkowników



Źródło : 3rd Annual Data Miner Survey – 2009 Survey Summary Report, Rexer Analytics, 2010

Business Intelligence okiem użytkowników



Źródło : 3rd Annual Data Miner Survey – 2009 Survey Summary Report, Rexer Analytics, 2010

Data mining



Eksploracja danych, zwana także odkrywaniem wiedzy w bazach danych (data mining, knowledge discovery in databases) to proces **odkrywania nowych, nieznanych, pożytecznych i zrozumiałych wzorców** w dużych wolumenach danych.

Idea eksploracji danych polega na wykorzystaniu szybkości komputera do znajdowania ukrytych dla człowieka (właśnie z uwagi na ograniczone możliwości czasowe) prawidłowości w danych zgromadzonych w hurtowniach danych.

- Prognozowanie metodami data mining
- Wykrywanie nadużyć (Fraud detection)
- Optymalizacja procesów (Process optimization)
- Badania innowacyjne (R&D)
- Monitorowanie pracy urządzeń

W recenzji najnowszej wersji ... dr. Johna Wassa można przeczytać: „Najnowsza wersja zawiera nie tylko bardzo szeroki zestaw narzędzi analitycznych, ulepszone procedury data mining oraz różnorodne możliwości wizualizacji i automatyzacji, ale również rozwiązania ułatwiające pracę wielu użytkowników. Zakres i szczegółowość dostępnych analiz są naprawdę zdumiewające”.

Scientific Computing World, 2010

Sztuczne sieci neuronowe (wspomaganie decyzji / SIP)



Sztuczna sieć neuronowa (SSN) – definicje:

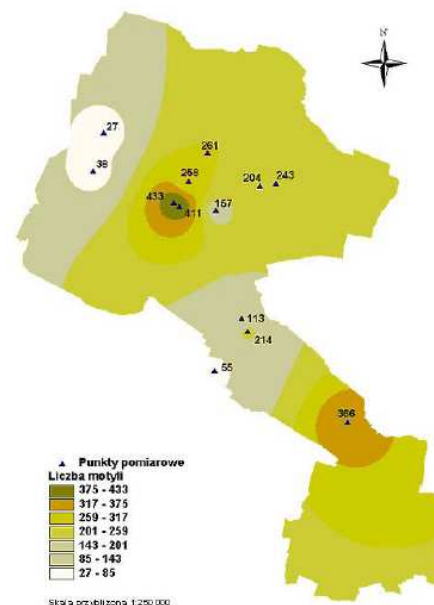
Zbiór prostych jednostek obliczeniowych przetwarzających dane, komunikujących się ze sobą i pracujących równolegle.

lub inaczej:

Zbiór połączonych ze sobą jednostek wyjściowo-wejściowych. Z każdym połączeniem skojarzona jest waga, która może zostać zmieniona w trakcie uczenia.

Dowolna sztuczna sieć neuronowa może być zdefiniowana poprzez określenie:

- modelu sztucznego neuronu
- topologii (sieci jednokierunkowe i rekurencyjne)
- reguły uczenia sieci



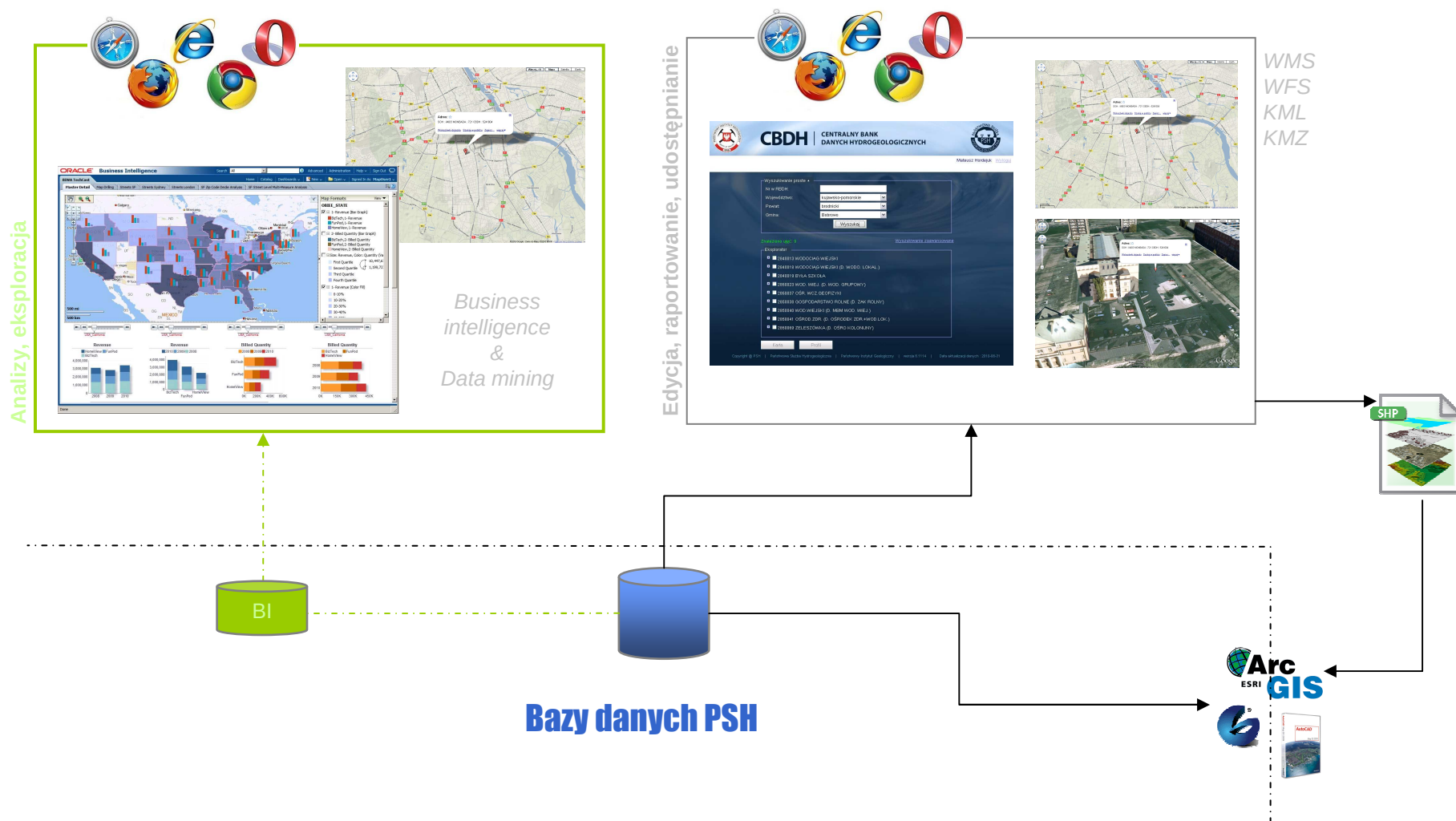
z wykorzystaniem
m.in. danych
meteorologicznych

błąd ~ 7%

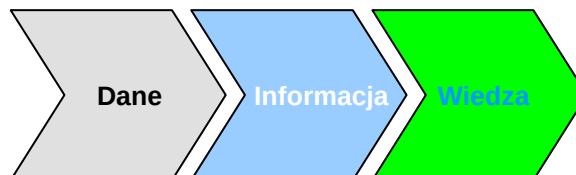
Rys. 2. Przestrzenna dystrybucja populacji motyli brudnicy mniszki – wynik interpolacji prognozowanych przez SN wartości.

Analiza obszarów leśnych w Systemach
Informacji Przestrzennej wspomaganych
Systemami
Ekspertowymi i Sztucznymi Sieciami
Neuronowymi.

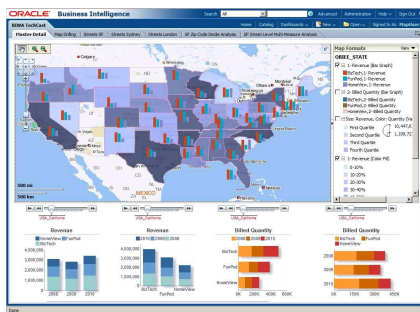
Kierunki rozwoju oprogramowania narzędziowego w PSH cz.2



Podsumowanie



badania jako element poznawczy

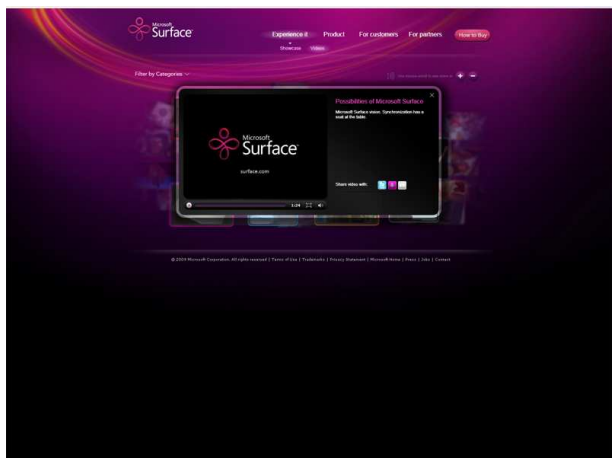


własny analityk



ostrożność interpretacji wyników

Trendy i nowe technologie



GaN –
azotek
galu
/
Ammono

"A little Polish company you've never heard of is beating the tech titans in a key technology of the 21st century"

Źródło : „The World's Best Gallium Nitride”, IEEE SPECTRUM MAGAZIN, July 2010



Źródło : Microsoft Office Labs



Państwowa służba hydrogeologiczna

www.psh.gov.pl

Mateusz Hordejuk

Zakład Analiz i Prognoz Hydrogeologicznych

tel. /22/ 849 53 51 wew. 353

tel. 607 701 790

e-mail: mateusz.hordejuk@pgi.gov.pl

Skype: mat_pgi



Państwowy Instytut Geologiczny
Państwowy Instytut Badawczy

www.pgi.gov.pl