

# Big Data



**KAPITAŁ LUDZKI**  
NARODOWA STRATEGIA SPÓJNOŚCI



Celownik – najfiszce inwestycji  
Projekt nr POKL.04.01.00-00-151/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



**UNIA EUROPEJSKA**  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPŁOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNE

str. 1

## Agenda

1. Co to jest Big Data?
2. Źródła Big Data
3. Model 3V
4. Typy Big Data
5. Big Data w biznesie
6. Platforma IBM Big Data



**KAPITAŁ LUDZKI**  
NARODOWA STRATEGIA SPÓJNOŚCI



Celownik – najfiszce inwestycji  
Projekt nr POKL.04.01.00-00-151/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



**UNIA EUROPEJSKA**  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPŁOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNE

## *Co oznacza Big Data?*

Zbiór danych tak duży, że jego przetwarzanie w poniższym zakresie jest trudne bez specjalistycznych narzędzi:

- gromadzenie danych,
- przechowywanie danych,
- przeszukiwanie danych,
- przesyłanie danych,
- analiza danych,
- wizualizacja danych.



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



Celownik – najfiszce inwestycji  
Projekt nr POKL.04.01.00-00-151/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
wspofinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPŁOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNE

## *Ile to jest petabajt danych?*

- 1PB
  - tyle danych gromadzi dziennie Facebook
- 2.5 PB
  - taka jest pojemność ludzkiego mózgu
- 422 EB
  - tyle danych stworzono na świecie w 2008
- 1.8 ZB
  - tyle danych stworzono na świecie w 2011



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



Celownik – najfiszce inwestycji  
Projekt nr POKL.04.01.00-00-151/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
wspofinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPŁOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNE

# 3V

- variety,
- velocity,
- volume



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



Celownik – najfiszce inwestycji  
Projekt nr POKL.04.01.00-00-151/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
wspofinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY

MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPŁOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNE

## Big Data a hurtowanie danych

### Hurtownia danych

### Big Data

Variety

Dane  
ustrukturalizowane



Dane **nieustrukturalizowane**  
Dane ustrukturalizowane

Velocity

Dane  
w spoczynku



Dane w spoczynku  
Dane **w ruchu**

Volume

Średni rozmiar  
hurtowni danych  
Netezza: 10TB



TB ( $10^{12}$ ) do ZB ( $10^{21}$ )



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



Celownik – najfiszce inwestycji  
Projekt nr POKL.04.01.00-00-151/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
wspofinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY

MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPŁOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNE

## Typy Big Data



strumienie informacji

oceany informacji



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



Celownik – najfiszce inwestycji  
Projekt nr POKL.04.01.00-00-151/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



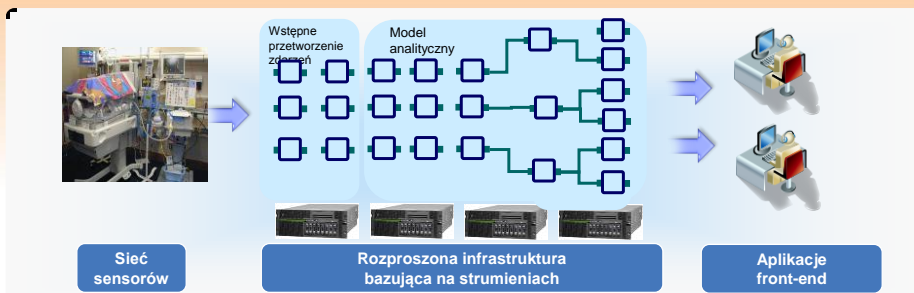
UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPLOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNIE

## University of Ontario

- Monitorowanie wcześniaków
- Przewidywanie infekcji wymagającej OIOM z 24-godzinnym wyprzedzeniem
- 120 dzieci pod ciągłą obserwacją: 120 tysięcy odczytów na sekundę, miliard odczytów dziennie
- Próby zdalnego włączenia do programu szpitali w USA i Chinach



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



Celownik – najfiszce inwestycji  
Projekt nr POKL.04.01.00-00-151/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



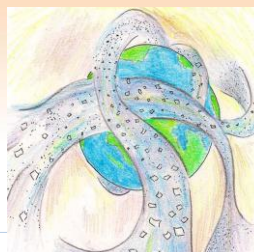
UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPLOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNIE

## Urząd Miasta Sztokholm

- Zbieranie i analiza w czasie rzeczywistym danych
  - GPS wysyłanych przez pojazdy,
  - sensorów natężenia ruchu,
  - środków komunikacji publicznej,
  - stacji monitorowania zanieczyszczeń powietrza,
  - stacji meteorologicznych
- Udostępnienie władzom miasta i mieszkańcom informacji o
  - obecnym natężeniu ruchu,
  - czasach przejazdu,
  - najlepszych (najszybszych) alternatywach dostarcia w pożądane miejsce
- Wcześniejszy projekt (Congestion Tax)
  - ograniczył ruch w Sztokholmie o 20%
  - zredukował średnie czasy podróży o 50%



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



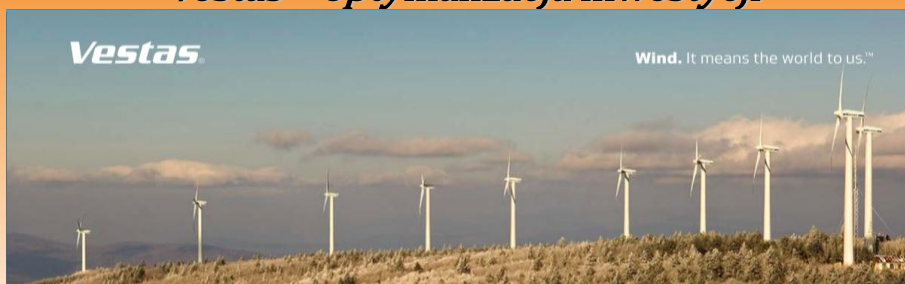
Celownik – najlepsze inwestycje  
Projekt nr POKL.04.01.00-00-131/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
wspofinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY

MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPŁOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNIE

## Vestas – optymalizacja inwestycji



- optymalne ulokowanie turbiny wiatrowej – stabilne i maksymalne możliwości generowania prądu,
- czas podjęcia decyzji o ulokowaniu turbiny skrócono z 3 tygodni do 3 dni,
- modelowanie bazujące na siatce 1km x 1km z setkami parametrów,
- modele budowane w oparciu o dane dostarczane z pracujących turbin.



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



Celownik – najlepsze inwestycje  
Projekt nr POKL.04.01.00-00-131/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
wspofinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY

MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPŁOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNIE

## Catalina Marketing

- Zindywidualizowane oferty promocyjne
  - Wykorzystanie kuponów wzrosło o 30%
  - Zwrot z inwestycji dla kampanii marketingowych wyższy o 30%
- Przewidywanie oparte o dane
  - 450 milionów rekordów dziennie – zbieranych, przetwarzanych i ładowanych do IBM Netezza Data Warehouse Appliance
  - 130 milionów klientów (gospodarstw domowych)
  - 400 miliardów transakcji (co wylądowało w koszyku sklepowym)



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



Celownik – najfiszce inwestycji  
Projekt nr POKL.04.01.00-00-151/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
wspofinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego

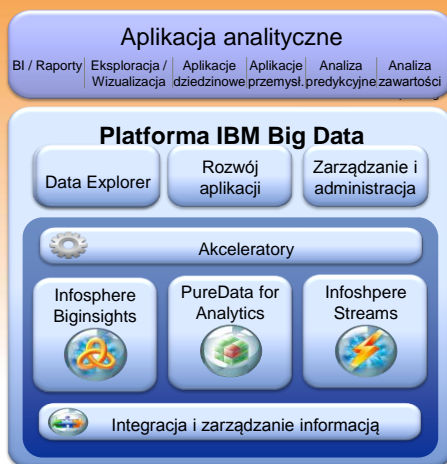


UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPLOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNE

## Platforma IBM Big Data



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



Celownik – najfiszce inwestycji  
Projekt nr POKL.04.01.00-00-151/12 „Inżynieria wiedzy dla inteligentnego rozwoju”  
wspofinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



MATERIAŁY DYDAKTYCZNE I SZKOLENIOWE NA STUDIA PODYPLOMOWE ORAZ NA SZKOLENIA DYSTRYBUOWANE SĄ BEZPŁATNE