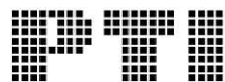


Wykorzystanie technologii integracji danych, raportowania, analiz i prezentacji w spisach

Seminarium „Wykorzystanie nowoczesnych technologii informatycznych w spisach powszechnych”



Główny Urząd Statystyczny, 2 marca 2011

Zenon Dogoński



THE
POWER
TO KNOW

Adam Bartos – Dyrektor Technologii SAS Institute Polska
Piotr Skarzyński – Project Manager
Łukasz Leszewski – Architekt Danych, Analityk

Świat analityczny

Świat operacyjny

...zi statystycznych, integracji
...nalitycznych.

...cji w 109 krajach.

- Zespół liczy ponad 10 000 osób, w Polsce blisko 180 specjalistów.

Wsparcie służb statystyki publicznej

More than 190
independent a

EMEA

- Statistik Austr
- Belgium's Stat
- Bulgaria's Nat
- Croatian Bure
- Cyprus' Minis
- Services
- Czech Statisti
- Danmarks Sta
- Estonia's Stati
- Finland's Tilas
- Germany's Sta
- Hungarian Ce
- Ireland's Cent
- Central Burea
- Italy's ISTAT Is
- Kuwaiti Minist
- Central Statist
- Statistics Lithu
- Luxemburg's s



- Dutch Central Bureau voor de Statistiek
- Norway's Statistisk Setral Byraa, (SSB)
- Poland's Central Statistical Office (GUS)

The Americas

- Argentina's INDEC
- Brazilian Census Bureau (IBGE)

Accounting & Statistics

- Thailand National Statistics Office

Free public sector *sascom* magazine at: www.sas.com/govedu/emea_sascom_15sep04.pdf

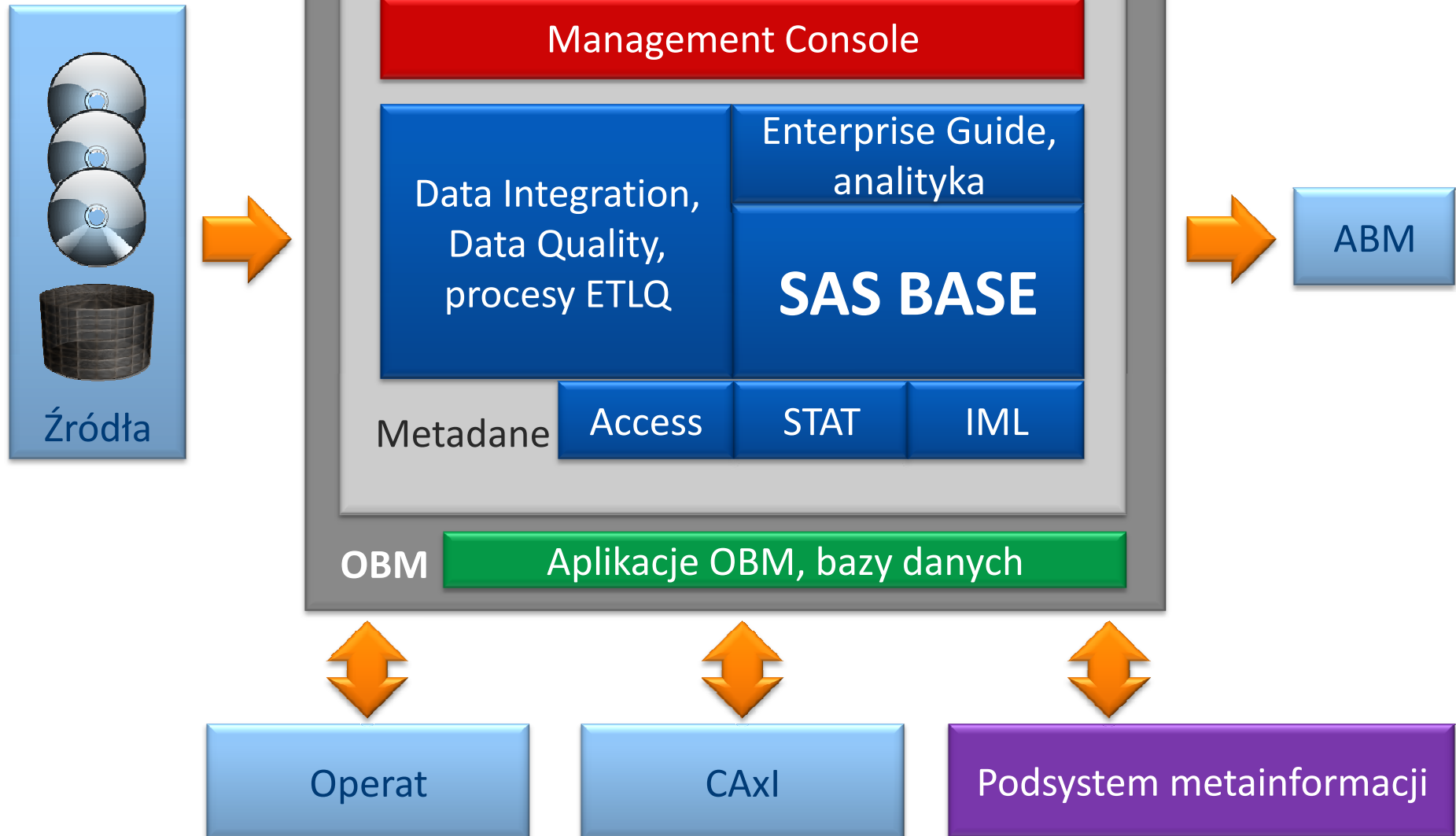
Spisy powszechne z punktu widzenia SAS Institute

- Nowatorskie podejście w skali międzynarodowej – doświadczenia, m.in.:
 - U.S. Census Bureau,
 - Urząd Statystyczny Danii
- Bezpieczeństwo
- Wydajność
- Skala przedsięwzięcia
- Różnorodność technologiczna
- Zagadnienia naukowe
- Terminy – znane i niezmiennie
- Dane – Czy będą na czas? Jakiej będą jakości?

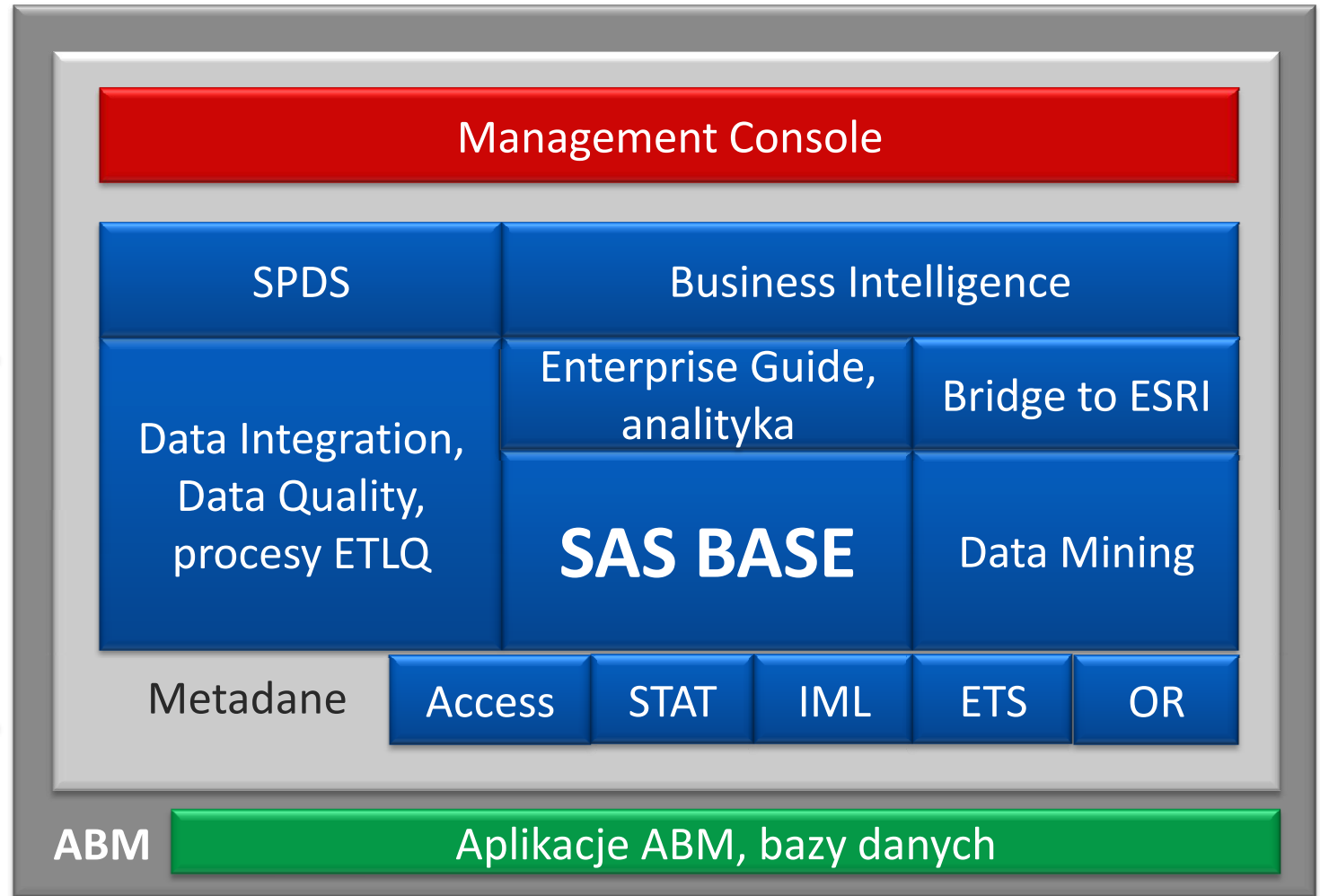
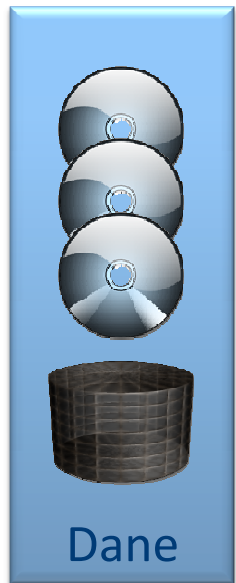
Wymagania

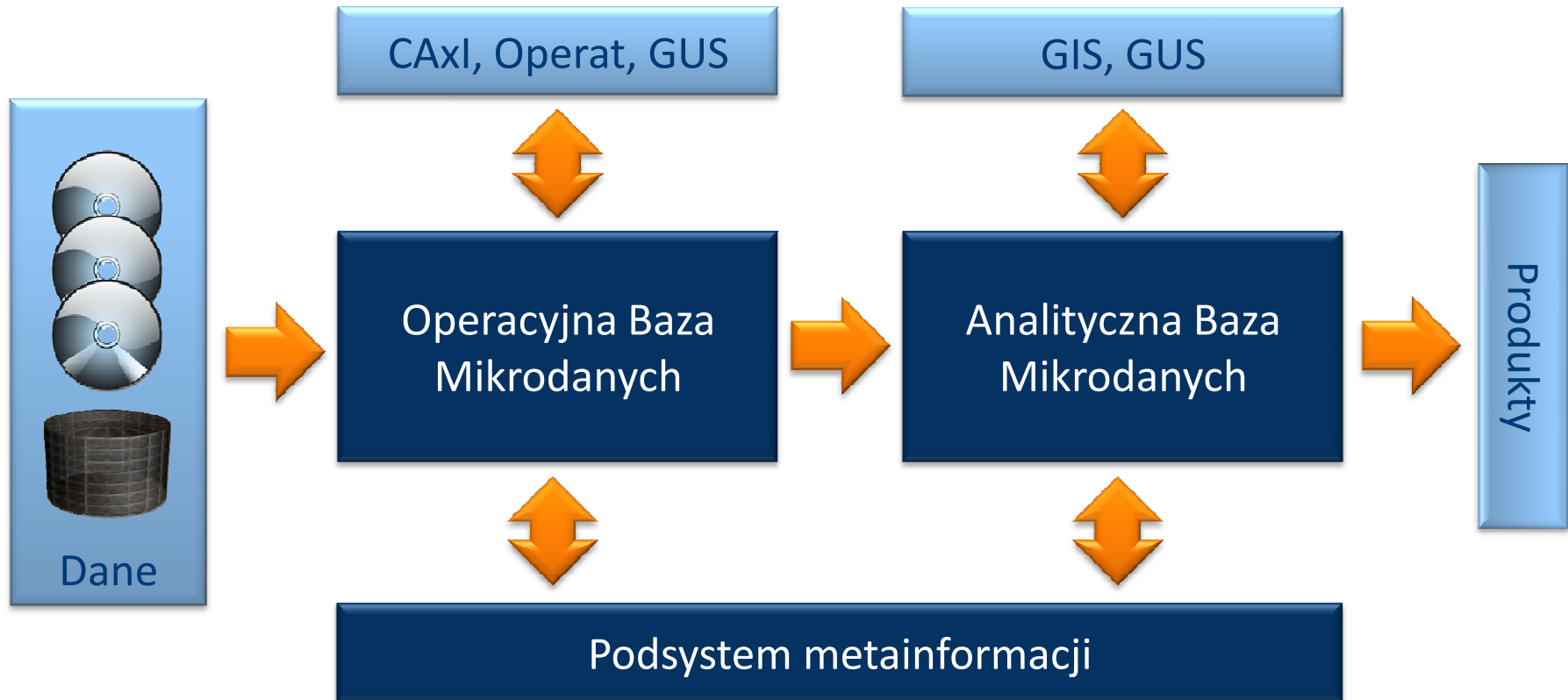
- Dostęp do danych w zewnętrznych bazach danych oraz w plikach.
- [Analiza jakości danych](#), procesy podwyższania jakości danych.
- Integracja danych, procesy przetwarzania danych wraz z procesami podwyższania jakości, [imputacje](#) (proste, zaawansowane, met. statystyczne, met. eksperckie).
- Składowanie danych – struktury przejściowe (tymczasowe), struktury docelowe.
- Nietypowe struktury danych – „[szerokie tabele](#)”.
- [Funkcjonalności ekonometryczne](#) – statystyka, data mining, implementacja „nowych” metod statystycznych (pakiet R).
- Raportowanie – przeglądarka internetowa, MS Office.
- Analityka – aplikacja analityczna.
- Dystrybucja i rozpowszechnianie – portal, mail.
- Współpraca z systemami klasy GIS.
- [Bezpieczeństwo](#).
- Ergonomia – użytkowników (wymiana wyników, praca grupowa), administratorów.

Technologia (OBM)



Technologia (ABM)





Podsumowanie

- Zbudowano uniwersalną platformę analityczno-informacyjną, która:
 - teza SAS – jest najnowocześniejszym w kraju środowiskiem analitycznym, zarówno od strony funkcjonalnej jak i wydajnościowej.
- Wyzwania:
 - „transfer wiedzy” – przeszkolenie zespołu odpowiedzialnego za przygotowanie danych w środowisku OBM, przekazanie najlepszych praktyk w zakresie przetwarzania danych, jakości danych,
 - technologia – zastosowano najnowsze narzędzia, systemy operacyjne oraz elementy infrastruktury (m.in. serwery, macierze), nie było wielu doświadczeń (globalnie) w integracji i strojeniu użytych komponentów, stąd niezależnie w laboratoriach firmy HP stworzono „poligon” gdzie inżynierowie prowadzili badania i testy, których celem było optymalnie zestrojone środowisko,
 - metody matematyczne – implementacja w środowisku SAS zaawansowanych metod matematycznych wykorzystywanych przez statystyków pracujących na rzecz spisów powszechnych, czasem przełożenie kodów utworzonych w środowisku R do środowiska SAS.



Dziękuję



www.sas.com/poland
zenon.dogonski@sas.com