



Granice personalizacji w świecie BIG DATA

**„VIII Dzień Ochrony Danych”
„Prywatność w cyfrowym świecie”**



BIG DATA – reguła 4 V's – w ZUS

Big Data to szereg pojęć i działań związanych z pozyskiwaniem, utrzymywaniem i operowaniem na danych które definiują poniższe pojęcia:



BIG DATA – co przetwarza ZUS, kanały zasilania danymi

Dokumenty

- ok 30 000 000 dokumentów zgłoszeniowych Płatnika
- ok 450 000 000 dokumentów rozliczeniowych Płatnika
- blisko 600 000 000 dokumentów zgłoszeniowych Ubezpieczonego
- ponad 10 000 000 000 dokumentów rozliczeniowych Ubezpieczonego

System zasilania EWD i PUE

- System pracujący w rygorze 24/7/365
- Obsługujący ok 2 100 000 przesyłek miesięcznie (Program Płatnika i ePłatnik)
- Pomiędzy 5-15 dniem każdego miesiąca obsługa ponad 60% wszystkich przesyłek



BIG DATA – przetwarzanie danych, analityka w ZUS

System Aktuarialny

Prognozowanie indywidualnych ścieżek kariery zawodowej dla całej populacji Polski
i w efekcie prognozowanie wpływów i wydatków
FUS oraz FEP

Zasilenie danymi o ubezpieczonych i świadczeniobiorcach, które są pozyskiwane
ze specjalnie przygotowanych ekstraktów danych
z systemów ZUS

Pozyskiwanie zagregowanych danych o liczbie
ubezpieczonych i świadczeniobiorców w podziale
na kohorty wiekowo-płciowe z:

- Kasy Rolniczego Ubezpieczenia Społecznego
- Ministerstwa Spraw Wewnętrznych
- Ministerstwa Sprawiedliwości
- Ministerstwa Obrony Narodowej

Modelowanie życia poszczególnych osób w modelu
na dwóch płaszczyznach:

- zawodowej
- niezawodowej

BIG DATA – udostępnianie danych przez ZUS



MINISTERSTWO NAUKI
I SZKOLNICTWA WYŻSZEGO

„Badanie losów absolwentów”

Ekstrakt danych o absolwentach (wraz z historią ich ubezpieczenia) jako wynik zapytań do utrzymywanej w Zakładzie bazy Kompleksowego Systemu Informatycznego (KSI).

Proces połączenia informacji z historią edukacji absolwentów oraz jej statystyczną charakterystyką.

Zanonimizowane dane przekazywane będą do MNiSW.

Nawet w przypadku zagregowanych danych statystycznych należy zachować zgodność z ustawą o tajemnicy statystycznej, aby nie dopuścić do sytuacji w której na podstawie danych zagregowanych będzie możliwość zdobycia informacji o jednostce.

BIG DATA – udostępnianie danych przez ZUS

Udostępnianie danych gromadzonych przez ZUS Komornikom Sądowym

Wniosek komornika sądowego o udostępnienie informacji ze zbiorów ZUS jest nadzorowany i realizowany przez system EPWD powiązany z aplikacjami dziedzinowymi z obszaru rozliczeń, świadczeń emerytalno – rentowych i zasiłków oraz z systemem SAP.

W tym zakresie rozwiązanie pozwala na:

- Rejestrację odpisu wniosku komornika sądowego o udostępnienie danych
- Koordynację obsługi wniosku
- Zlecenie zadania wyjaśniającego kancelaryjnego,
- Przekazanie wniosku do innej jednostki,
- Identyfikację podmiotów z wniosku,
- Pozostawienie wniosku bez rozpatrzenia,
- Utworzenie i dołączenie do sprawy pism
- Zlecanie realizacji zapytań z wniosku
- Obsługę realizacji zapytań z wniosku
- Rejestrację informacji dotyczącej wysłania pisma do klienta lub odbioru osobistego.

Ponad 250 000
zapytań
miesięcznie

Udostępnianie informacji Komornikom Sądowym z uwagi na wrażliwość danych musi być skrupulatnie nadzorowane i kontrolowane.

BIG DATA – udostępnianie danych przez ZUS



MINISTERSTWO PRACY
I POLITYKI SPOŁECZNEJ

Udostępnianie danych Ośrodkom Pomocy Społecznej i Powiatowym Centrom Pomocy Rodzinie.

Zakres:

Obszar zasiłków - okres i kwota wypłaconego świadczenia, podanie czy pobiera/ł zasiłki, w jakich okresach, informacja o wypłacanych świadczeniach; okres i kwota świadczenia; wskazanie formy pobierania zasiłku (konto/przekaz); informacja o ubezpieczeniu i prawie do zasiłku; informacja o niezdolności do pracy - okres, kwota netto świadczenia rehabilitacyjnego.

Obszar świadczeń emerytalno-rentowych - podanie czy pobiera/ł świadczenia, ich wysokość, prawo do dodatku pielęgnacyjnego, podanie informacji z dokumentów (bez ich załączenia), informacje o pobieranych świadczeniach; kopie decyzji, przysługujące dodatki, rodzaj potrąceń, wskazanie okresu w którym osoba była uprawniona do dodatku pielęgnacyjnego lub odpowiedź, że nie była, informacja o zasiłku pielęgnacyjnym.

Obszar dochodów - podanie przez jaki podmiot jest zgłoszony, czy podlega ubezpieczeniom i z jakiego tytułu, okresy podlegania ubezpieczeniom, aktualne zgłoszone do ubezpieczeń, informacja o przekazanych składkach, podstawy wymiaru składek, dane płatnika składek, numer rachunku bankowego, dane adresowe, informacja o pobieranych świadczeniach.

Inne dane - informacja o pobieraniu lub nie świadczenia emerytalno-rentowego, odpisy orzeczeń lekarskich.

BIG DATA – plany związane z danymi gromadzonymi w ZUS

Korporacyjna hurtownia danych ZUS (Platforma Analityczno-Raportująca ZUS)

Cel podstawowy projektu:

- konieczność ujednolicenia i standaryzacji w obszarze platform informacyjnych eksploatowanych przez ZUS
- zapewnienie spójności informacji zarządczej wykorzystywanej w organizacji (tzw. „jedno źródło prawdy”)
- dostarczanie informacji szybko i efektywnie pod względem finansowym

Cele dodatkowe:

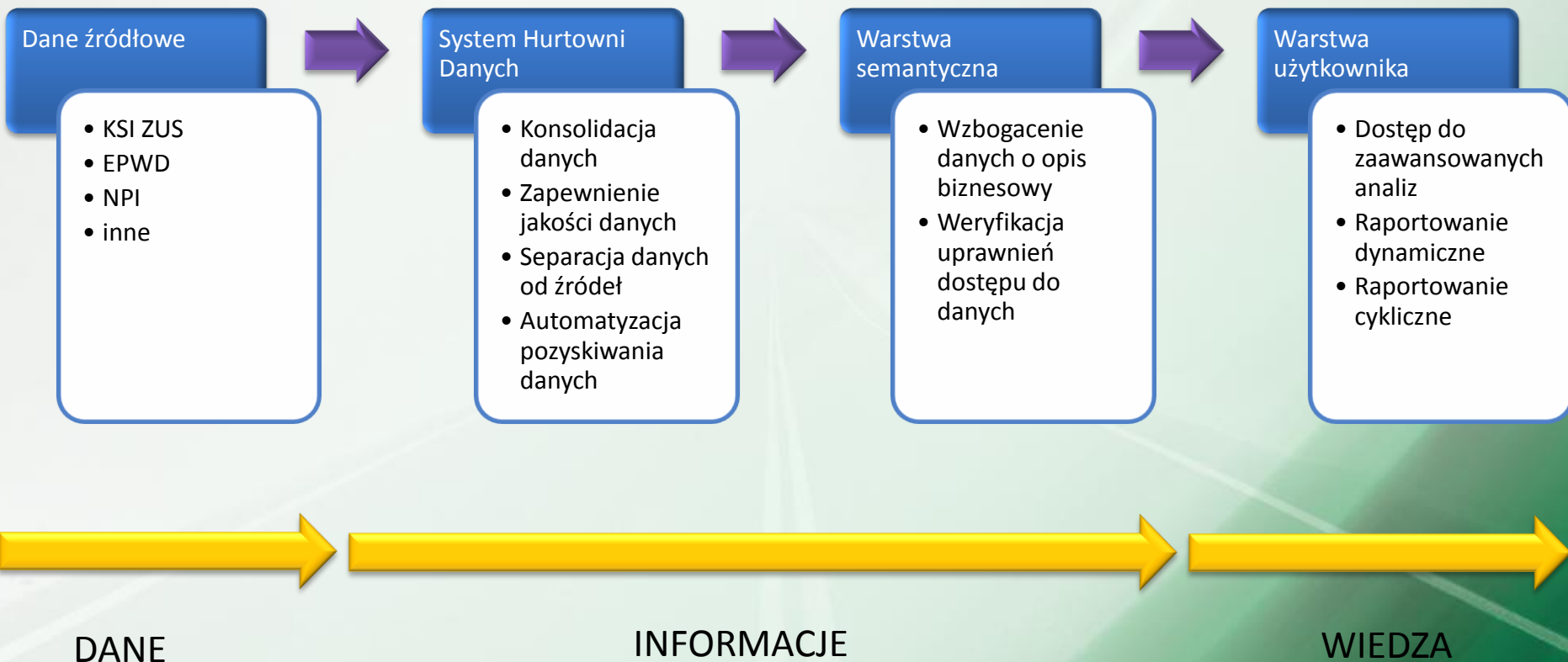
Dostarczenie mechanizmów zarządzania jakością danych przetwarzanych w systemie hurtowni danych oraz wzbogacania danych z wykorzystaniem reguł decyzyjnych i algorytmów statystycznych.

Dostarczenie użytkownikom platformy hurtowni danych narzędzi pozwalających na samodzielne budowanie analiz i raportów na podstawie zgromadzonych danych oraz informacji posiadanych przez użytkowników.

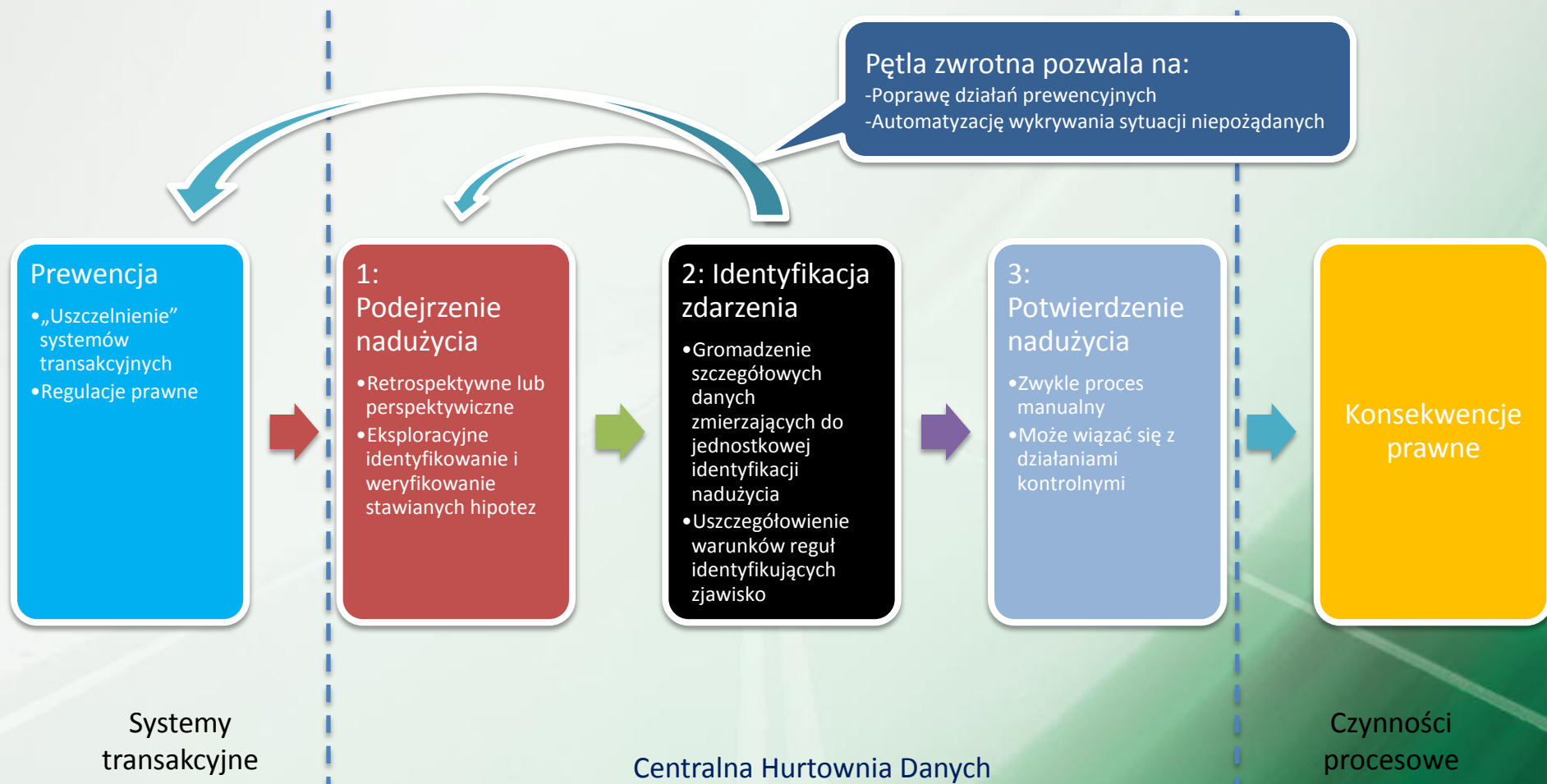
Dostarczenie platformy hurtowni danych umożliwiającej zasilanie systemów dziedzinowych w dane o znanej jakości i monitorowanej metodzie pozyskania i transformacji, w tym w pierwszej kolejności dostarczenie danych dla **systemu aktuarialnego**.



BIG DATA – Perspektywa transformacji danych gromadzonych w ZUS w informację i wiedzę



BIG DATA – perspektywa wykrywania nadużyć przez ZUS



BIG DATA – zagrożenia, aspekty prawne

Granica między budową nowych usług, a Orwellowskim światem?

Zakład Ubezpieczeń Społecznych jest w posiadaniu ogromnych zbiorów danych o Obywatelach. Często są to informacje niejawne, a nawet wrażliwe. Dane o miejscu zatrudnienia, wysokości wynagrodzenia, ilości i częstotliwości zwolnień chorobowych, informacje o zasiłkach i dodatkowych świadczeniach – zbierane przez wiele lat, poprzez swoją zmienność obrazują pewne trendy i umożliwiają metodami predykcyjnymi przewidywanie przyszłość poszczególnych osób (kariera zawodowa, awans itd.).

Dodając do tego informacje z pozostałych rejestrów państwowych (NFZ, CEPIK, pl.ID) i stosując zaawansowane mechanizmy przetwarzania można bardzo dokładnie spersonalizować profil dowolnego obywatela, ingerując w jego prywatność. **Jest to zagrożenie.**

Ale te same techniki dają też szanse wykorzystania tych danych w innym aspekcie – np. w systemie aktuarialnym już wykorzystywanym i w budowanej korporacyjnej hurtowni danych (zaawansowana analityka, informacja zarządcza, itd.)

Dużo zależy od ilości i jakości danych oraz zastosowanych mechanizmów (w tym filtrów odsiewających informacje zbędne).

Korelacja danych z systemu P1 w CSIOZ z danymi ZUS o świadczeniach dałaby zapewne lepsze wyniki aktuarialne, ale czy byłoby to zgodne z prawem i nie naruszałoby naszej prywatności?

[...] W innych krajach UE (np. Niemcy) obowiązuje zakaz łączenia danych o ubezpieczeniach społecznych z danymi z innych rejestrów publicznych. Podobna konstrukcja obowiązuje w Czechach. W krajach tych profilowanie obywatela na podstawie danych pochodzących z różnych baz publicznych jest niemal niemożliwe. [...]"

Dariusz Śpiewak

Zakład Ubezpieczeń Społecznych

