

Trendy i możliwości prowadzenia publicznych zbiorów danych

Marek Niezgódka

ICM, Uniwersytet Warszawski

marekn@icm.edu.pl

Konferencja, Mazowiecki Urząd Marszałkowski, 20.04.2017

Zasoby danych

- Procesy integracyjne - od odrębnych zasobów do zunifikowanych e-infrastruktur operacyjnych:
 - Zbiory danych (nieustrukturyzowanych)
 - Funkcjonalności:
 - strukturyzacja danych
 - kuracja danych
 - narzędzia analityki i modelowania w oparciu o dane (wtórne użycie)
 - Platformy cyfrowe, repozytoryjne, dziedzinowe i instytucjonalne, coraz częściej sfederowane z wysokim poziomem interoperacyjności na wszystkich poziomach funkcjonalnych

Infrastruktury danych: uwarunkowania

- Zakres i zasady dostępu oraz użytkowania publicznych zasobów danych:
 - istota rozwoju e-państwa i gospodarki cyfrowej
 - znaczenie aspektów legislacyjnych:
 - poziom krajowy
 - harmonizacja międzynarodowa

Perspektywy rozwoju

- Powstawanie wielkich systemów operacyjnych:
 - inteligentne miasta i regiony (*Smart Cities/Regions*)
 - nowa generacja inteligentnych sieci energetycznych (*Smart Grids*)
 - sieci społecznościowe, Internet rzeczy

Standardy międzynarodowe: UE

- Dyrektywa **PSI (Re-Use of Public Sector Information)**, 2003, rozszerzona 2013:
 - Open Government data
 - Dziedzictwo kulturowe
- Rekomendacja Komisji Europejskiej nt. otwartości treści i wyników badań naukowych (2012)
- Otwarte infrastruktury danych:
 - Europejska Infrastruktura Danych
 - EOSC (*European Open Science Cloud*) – wirtualna infrastruktura
 - OpenAIRE (Infrastruktura publikowanych wyników badań naukowych, obligatoryjna dla beneficjentów H2020)
- Dyrektywa nt. Copyright (w przygotowaniu):
 - Nowe zasady dot. TDM (*Text and Data Mining*)

Otwartość danych publicznych

- Zasada a nie wyjątek
- Klauzula wyjątkowa: ograniczenia otwartości
- Konsekwencje: możliwości wtórnego użycia

UE: Otwarta Nauka

- Treści upublicznianych wyników badań realizowanych ze środków publicznych – OTWARTE
- Rozwój kultury dzielenia się
- Procesy demokratyzacji nauki
- Znaczenie usuwania barier:
 - międzyinstytucjonalnych
 - międzysektorowych
 - transgranicznych

Otwarte infrastruktury danych: wyzwania

- Trwałość
- Integralność
- Bezpieczeństwo
- Spójność:
 - Interoperacyjność
 - Standardy otwarte

Big Data: nowe problemy

- Charakteryzacja - **4x V**:
Volume, Variety, Variability, Veracity
- Wyzwania technologiczne:
 - zdolność przetwarzania
 - umożliwienie głębokich analiz i użycia w modelowaniu komputerowym
- Wyzwania strukturalne:
 - interoperacyjność
 - operacje rozproszone lokalizacyjnie
 - wirtualne organizacje

Otwarte infrastruktury danych: Polska

- Poziom deklaratywny:
 - Programy Operacyjne poprzedniej i obecnej perspektywy
- Sytuacja rzeczywista:
 - duży zakres redundancji
 - niski poziom integracji
 - silosowe bariery międzysektorowe
- Ciągłe jeszcze niski poziom świadomości

Polska: wyzwania

- Rzeczywista implementacja zasad wtórnego użycia danych publicznych:
 - Zakres otwartości danych
- Partnerstwa dla tworzenia i usług związanych z danymi publicznymi
- Rola i szansa samorządów terytorialnych:
 - Inteligentny region, miasto, ...
 - Interoperacyjność na poziomie międzyregionalnym:
 - Rola wspólnych standardów
 - Programy edukacyjne
 - Rola instytucji nauki

Marek Niezgódka

marekn@icm.edu.pl