

GeoSmartCity Workshops, Katowice, 26th October 2016

## OPEN SPATIAL DATA AND METADATA ON THE WEB

#### Adam Iwaniak



Wrocław's Institute for Spatial Information and Artificial Intelligence Wroclaw University of Environmental and Life Sciences







#### www.wizipisi.pl

The Wroclaw Institute of Spatial Information and Artificial Intelligence

## Adam Iwaniak, PhD - CEO

From 2006 to 2008 vice-president of the Head Office of Geodesy and Cartography. Author of over 130 papers on the GIS and SDI.





## OUR COMPANY

- >20 employees
- 2 PhDs, 2 PhD candidates
- Scope of business
  - Intelligent spatial data infrastructures
  - Spatial BI, ETL
  - Semantic web / Linked data
- We truly believe our job is changing people's life for the better



OPEN DATA

T

计算机

16666

TIT



PRZESTRZENNE. TSZTUCZNEJ INTELIGENCJI



## OPEN DATA

- Public right of access to information has become a subject of public debate since the mid-XVIII century
- The pioneer in this field were Nordic countries. In 1776 in Sweden, the Act of press freedom was passed
- The next countries that have adapted the right to free access to information, were Denmark and Norway (1970), France and Finland (1978), Canada, New Zealand and Australia (1982), United States (1996) and the United Kingdom (2000).





## OPEN GOVERNMENT

- transparency of government activities
- availability of government services and information
- responding to new ideas, demands and needs,
- the ability to develop applications and stimulate the economy







## OPEN GOVERNMENT

• B.Obama, Open Government Directive, 2009

"To the extent practicable and subject to valid restrictions, agencies should publish information online in an **open format** that can be retrieved, downloaded, indexed, and **searched by commonly used web search applications**.

An open format is one that is platform independent, **machine readable**, and made available to the public **without restrictions** that would impede the re-use of that information"







# INFLUENCE of IT TRENDS on GI





Enables data published in Internet to be processed by humans.







# WEB OF DATA (3.0)

Enables data published in Internet to be integrated and processed by machines







# LINKED OPEN DATA

111

ROWN





## THE TECHNOLOGY STACK BEHIND THE SCENES

SPARQL – querying tools

. . .

OWL – ontology – semantic description

RDF – linked data interchange format

Uniform Resource Identifier (URI)





# STAR RATING SYSTEM



Available on the web (whatever format) but with an open licence, to be Open Data

Available as machine-readable structured data (e.g. excel instead of image scan of a table)

as (2) plus non-proprietary format (e.g. CSV instead of excel)

All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff

All the above, plus: Link your data to other people's data to provide context









# METADATA





## INSPIRE METADATA

- Bankruptcy of the ISO19139 metadata approach and CSW
- We need new and better data discovery technologies (instead of perfecting old methods for new challanges)
- We need more automation in Internet resources consumption.





## METADATA FOR REFERENCE OBJECTS

Q: How to make our geoportal more attractive, more discovered ?

A: Creating metadata for reference object in the HTML+RDFa





## METADATA FOR REFERENCE OBJECTS WITH LINKS TO GEOPORTALS





# (X)HTML + RDFa

#### N-34-83-D-B-3-2

Title:	N-34-83-D-b-3-2			
Identifier:	e9729928-37a5-4	e9729928-37a5-47ef-9939-3453A26b8b7e		
Abstract:	Ortofotomapa wy 2006. Zdjęcia zost opracowania wyk	/konana z bawnych zdjęć lotniczych w skali 1:26 0000 wykonanych w latach 2005- ały poddane procesowi skanowania z pikselem o rozmiarze 14 mikrometrów. Do corzystano także dane z aerotriangulacji i numeryczny model terenu.		
Keywords:	Sporządzanie ortoobrazów , Zdjęcie lotnicze , Obraz terenu , Fotografia terenu , Dane referencyjne , Ortofotogram , Ortofotomapa lotnicza , Rastry , Ortofotomapa			
Resource type:	dataset			
Format:	GeoTIFF 24 bity			
North Latitude:	53.79282			
South Latitude:	53.76968	<pre><div class="value"> </div></pre> <pre></pre> <pre></pre> <pre>/</pre> <pre>/ <pre>/<!--</th--></pre></pre>		
East Longitude:	23.4397	<pre></pre>		
West Longitude:	23.40407	<pre><div class="lobel"> <div class="lobel"> </div> </div>                                            </pre>		
		<pre></pre>		
4		<pre>/ <span class="" property="dc:subject">Zdjęcie lotnicze</span></pre>		
+		<pre><span class="" property="dc:subject">Obraz terenu</span></pre>		
-		<pre><span class="" property="dc:subject">Fotografia teren </span></pre>		
		<pre><span class="" property="dc:subject">Dane referencyjne /</span></pre>		
		<span class="" property="dc:subject">Ortofotogram,</span>		
	Lange and the second se	<span class="" property="dc:subject">Ortofotomapa lot: , <span class="" property="dc:subject">Rastry</span></span>		
		<pre><span class="">Resource type</span>:</pre>		



## INSPIRE metadata in HTML + RDFa

Zakliczyn ca. 218 m

## compliant with GeoDCAT-AP



ugust, 2016; https://pzgik.geoportal.gov.pl/semantic-metadata/

## CONVERSION OF INSPIRE METADATA







## ADVANTAGES

- With new representation, allows old metadata to be available for web search engines.
- New form /human-readable and machine-readable /
- Metadata is enriched with links to other resources: GEMET, GeoNames, DBpedia and metadata portals.

## HOWEVER,

citizens with their smartphones do not need metadata nor spatial data, they need answers for their questions.



# SPATIAL DATA ON THE WEB





## **KEY QUESTIONS**

- How to make spatial data more attractive ?
- How to make spatial data more available and accessible on mobile devices?
- It is already more than 2 200 000 mobile applications on Google Store and the next one will probably not solve the problem.





## INSPIRE DATA IN HTML + RDFa





SOol – Spatial Object of Interest /point, line and area/

## WROCŁAW – SOBÓTKA, CYCLE ROUTE

#### Wrocław – Sobótka (red trail) – CYCLE ROUTE

GENERAL INFORM	ATION		
Title:	Wrocław – Sobótka (red trail)		
Category:	http://inspire.ec.europa.eu/metadata-codelist/TopicC	ategory/transportation	
Abstract:	The cycling route is a route functionally connecting the includes: dedicated bicycle paths, cycle lanes, low traffi can be safely and comfortably used by cyclists	various parts of the city, th ic road and other roads that	at t
KEYWORDS			
Keywords:	Rower, Ścieżka rowerowa, Wypoczynek, Sieci transport Recreation, Transport networks	owe, Bicycle, Cycle path,	
REFERENCES TO TE	RMS OF KEYWORSDS OF GEMET DICTIONARY AND INSP	RESOURCE ADDRESS	
Keyword:	Rower	Resource address:	http:
URL:	http://www.eionet.europa.eu/gemet/concept?langcod		szlał
Keyword:	Ścieżka rowerowa		
URL:	http://www.eionet.europa.eu/gemet/concept?langcod	RESOURCE INFORMAT	ION
Keyword:	Wyporzynek	Coordinate system:	urn:o
URL:	http://www.eionet.europa.eu/gemet/concept?langcod	Route type:	Cycle
Keyword:	Sieci transportowe	Length :	40 k
URL:	http://inspire.ec.europa.eu/theme/tn	Route through:	Wro
Keyword:	Bicycle		Galo
URL:	http://www.eionet.europa.eu/gemet/concept?cp=788		nad
Keyword:	Cycle path	Elevation gain:	60 m
URI	http://www.eionet.europa.eu/gemet/concept?cp=108	Elevation profile:	
Keyword:	Recreation	1	180 -
URL:	http://www.eionet.europa.eu/gemet/concept?cp=700		170 =
Keyword:	Transport networks	1	160 -
URL:	http://inspire.ec.europa.eu/theme/tn		150
			140 -

	PLACES NEARBY (WITHI	N 1 KM)
	Restaurants:	Agawa
		Aroma
		Szczere Pole
		LOT Kitchen
		Pappatore Pizzeria Ristorante
		Pod Fura
		Pod Jeleniem
	Temples:	Church of the Immaculate Conception of the Blessed Virgin Mary
		Church. Holy Trinity
		Church. St. Padre Pio
		New Apostolic Church in Poland. The church Wroclaw
		Church. St. James the Apostle
		Pentecostal church
	Cemeteries:	The cemetery of Soviet Officers
	Castles:	Topacz Castle
	Green areas:	Ślęża Landscape Park
		South Park
	Grocery stores:	Odido
		Kiełcz Ryszard. Grocery store
		<u>Biedronka</u>
//w		Żabka
#/z1		Freshmarket
	Hotels:	Podkowa
		Pod Wieżycą
c:de		Ślęża Pension
rout		Topacz Castle
lout		Bacero Hotel
aw (		Pod Jeleniem
vice		
eża	METADATA	
4-4	ID:	fdsds6a-8aa2-43fd-8d92-20327dsvsvsd
	Encoding:	UTF-8
	Name of standard:	ISO 19115
	Version of standard:	2003/cor.1:2006
		· · · · · · · · · · · · · · · · · · ·
	n. ~~~	
	10,000 20,000 20	
	10 000 20 000 30	



SPATIALEATENT			
North Latitude:	51.0748		
West Longitude:	16.7467		
East Longitude:	17.0126		
South Latitude:	50.8862		

120 -

# WIZIPISI SEMANTIC COMPONENTS



## SPATIAL DATA TO RDF CONVERSION





## SEMANTIC COMPONENTS WIZIPISI





## USE CASE: NEW WAY OF DATA ANALYSIS

## AllegroGraph - Semantic Graph Database - Franz Inc.

Allows to represent knowledge with the use of ontology /RDFS++/ or rules /SPIN/

Questions:

1. Find parcels of Adam Iwaniak **family** 

2. Find property of Adam Iwaniak family







# THE END



adam.lwaniak@wizipisi.pl



## USE CASE: DATA INTEGRATION WITH THE WEB RESOURCES

#### select distinct ?church ?label str(?lat) AS ?lat str(?long) AS ?long







## CREATING LINKS

Task 5 stars:

1. Create an RDF link churches from topographic databases with their counterparts from DBpedia

2. Enrich topographic data with dereferenceable DBpedia URI





## LINKING SPATIAL OBJECTS

Example of connecting churches from topographic databases with their counterparts from Dbpedia with RDF links.







## DATA ENRICHMENT



Gene	ral Attributes	
	Name	Value
>	osm_id	137341515
	name	Riverside Church
	denominati	
	church	http://pl.dbpedia.org/resource/Riverside_Church
	label	Riverside Church
	MatchingType	automatic
	architekt	Henry C. Pelton, Charles Collens
	church_DB	http://pl.dbpedia.org/resource/Riverside_Church
	liczbaNaw	3
	liczbaWiernych	2400
	przeznaczenie	czynny kościół międzywyznaniowy
	styl_label	Neogotyk

## OpenStreetMap + Dbpedia





## **GRAPH ANALYSIS - SNA**

Social Network Analytics:

- Actor degrees
- Cliques
- Actor centrality
- Group centrality
- Group cohesiveness.



