The Role of Metadata and Semantic Web in SDI - selected issues

Adam Iwaniak

Wroclaw University of Environmental and Life Sciences

Kon-dor s.c.

How to make GIS systems more intelligent ?
How to make their interface simpler and more intuitive in use?

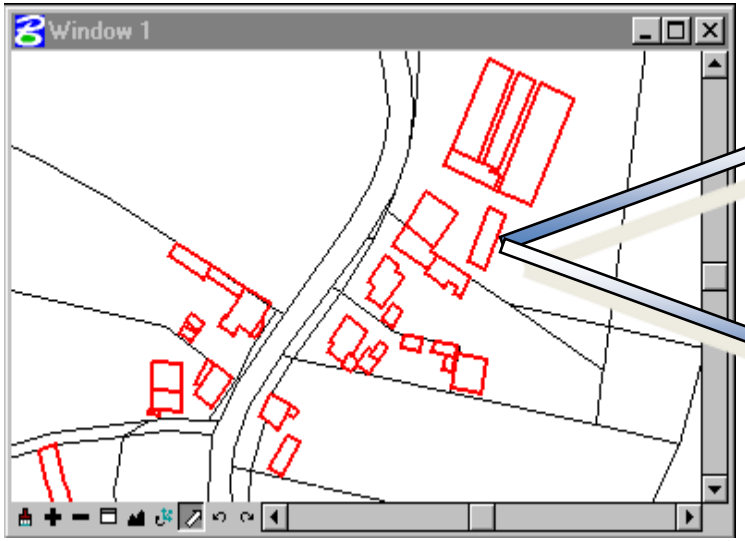
Add mouse to CAD –the ’80s

Line segment is represented by a pair of points P1 and P2 with coordinates (x1,y1) and (x2,y2)

Add RDB – ’90s

How to find buildings that have two floors?

- graphics in CAD system,
- attributes in relational databases (Access, Oracle, ...),
- example: MGE, GeoGraphics, AutoCad Map.



Features

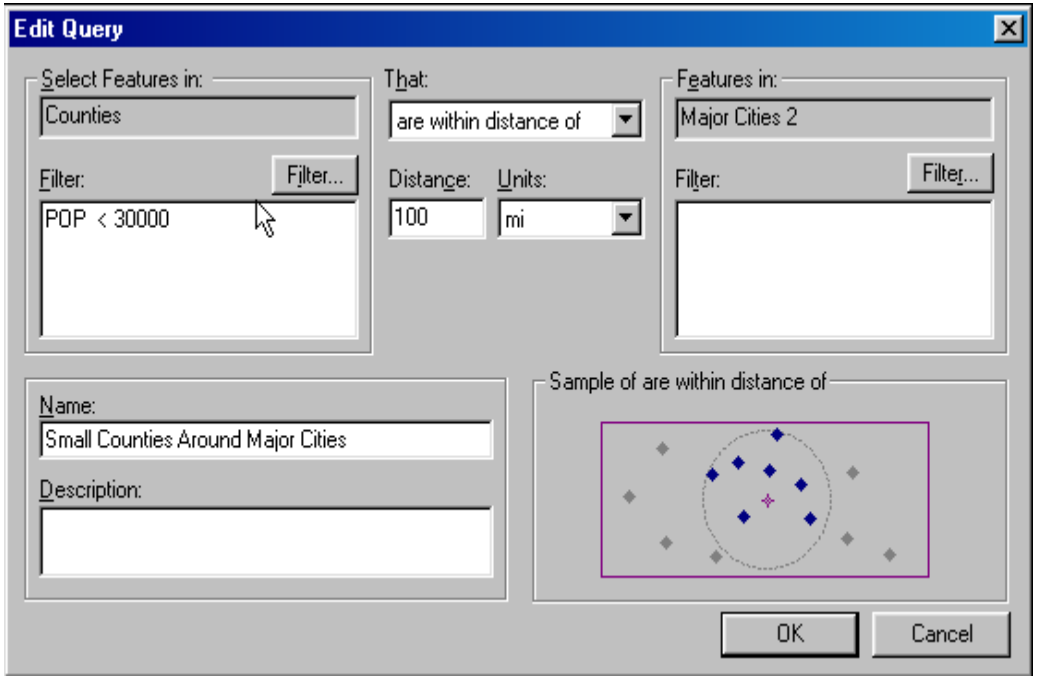
Name_of_feature
parcels
buidlings
roads

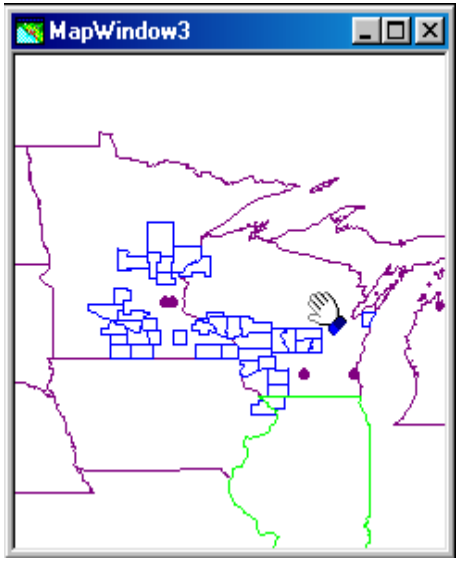
Buildings

Function	Nr_storeys	id
16	2	1011
17	3	1012
18	1	1013

Add topology – ’90s

Spatial queries





Select * from Counties where Counties.pop < 30000 and Counties *are within distance* *100 mi of* Major Cities;

Spatial operator

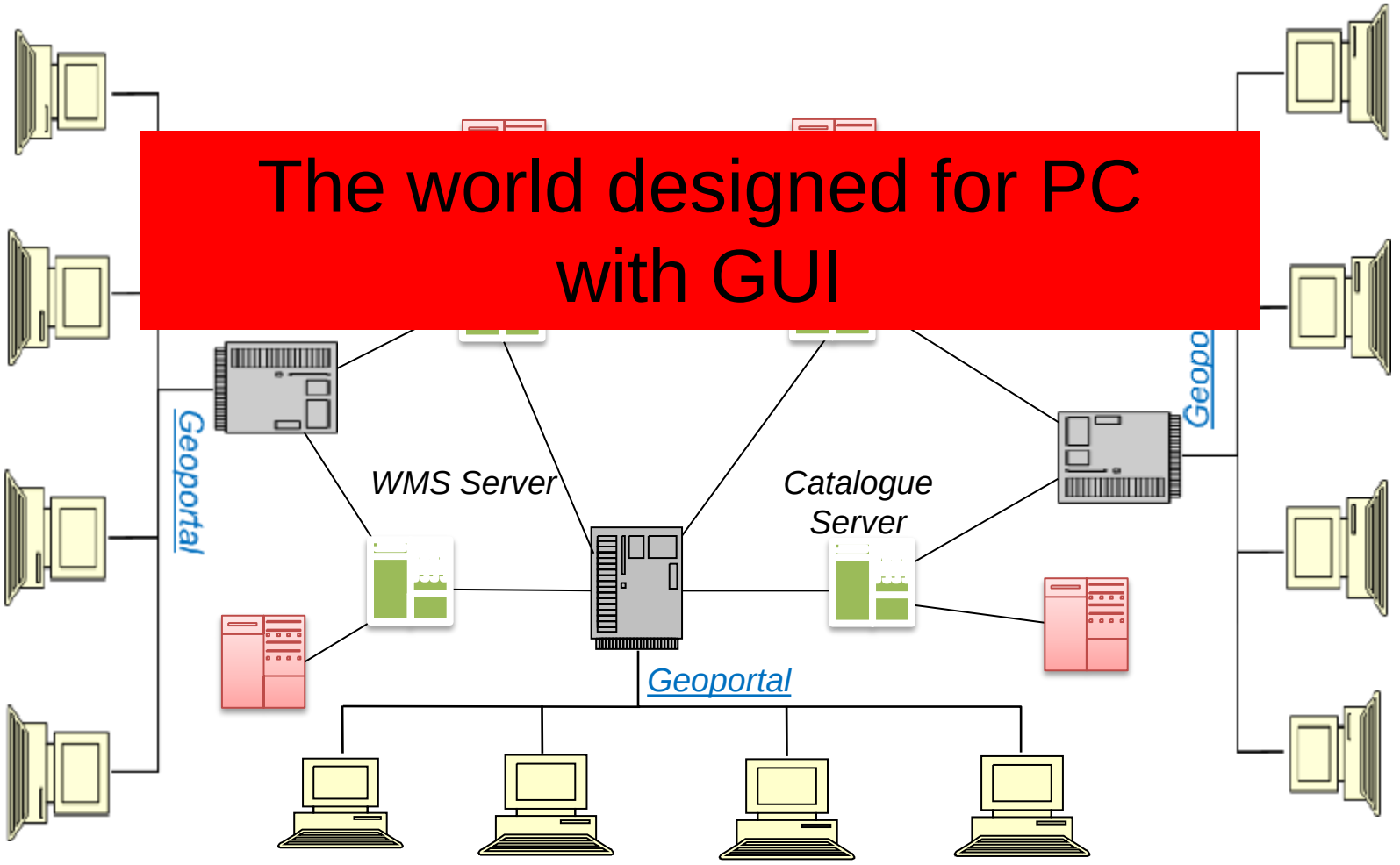
Add spatial option – ’90s

The development of technology which stores geometry and attributes into a single database allows to use the mechanisms of RDBMS (Relational Database Management System in GIS. This resolve problems with:

- multiaccess
- data protection
- versioning
- transactions
- backup performance
- data integrity.

Use SOA architecture

Collect spatial data once and use it many times



The new era of the Internet - smartphones

- Today is a big day for Android... 500 million devices activated globally, and over 1.3 million added every single day," *Hugo Barra, director of product management, Android, notes in a post on Google+.*
- More than half of Internet users in Poland use the Internet via a smartphone, it's four times more than a year ago. 45% of Internet users as a reason for buying indicate the use of maps and GPS, *MEC Analytics & Insight*
- People have easier access to smartphones than toilets, *Zet Radio*

Example 1

1. Find the nearest gas station
2. Find the shortest route between point A and point B
3. Find international hotel in Zagreb

Task does not seem to be difficult to implement when the questions are predefined and the data is stored in one database which model well known.

Example 2

Find building in Wroclaw that are owned by Adam Iwaniak

There are three databases available: administrative boundaries, a map of the parcels and there owners and topographic database

Classic GIS

1. Find all parcels in Wroclaw, owned by Adam Iwaniak
2. Find buildings which are located in these parcels
3. Show a list of buildings and their addresses

Key issues in „smartphone SDI”

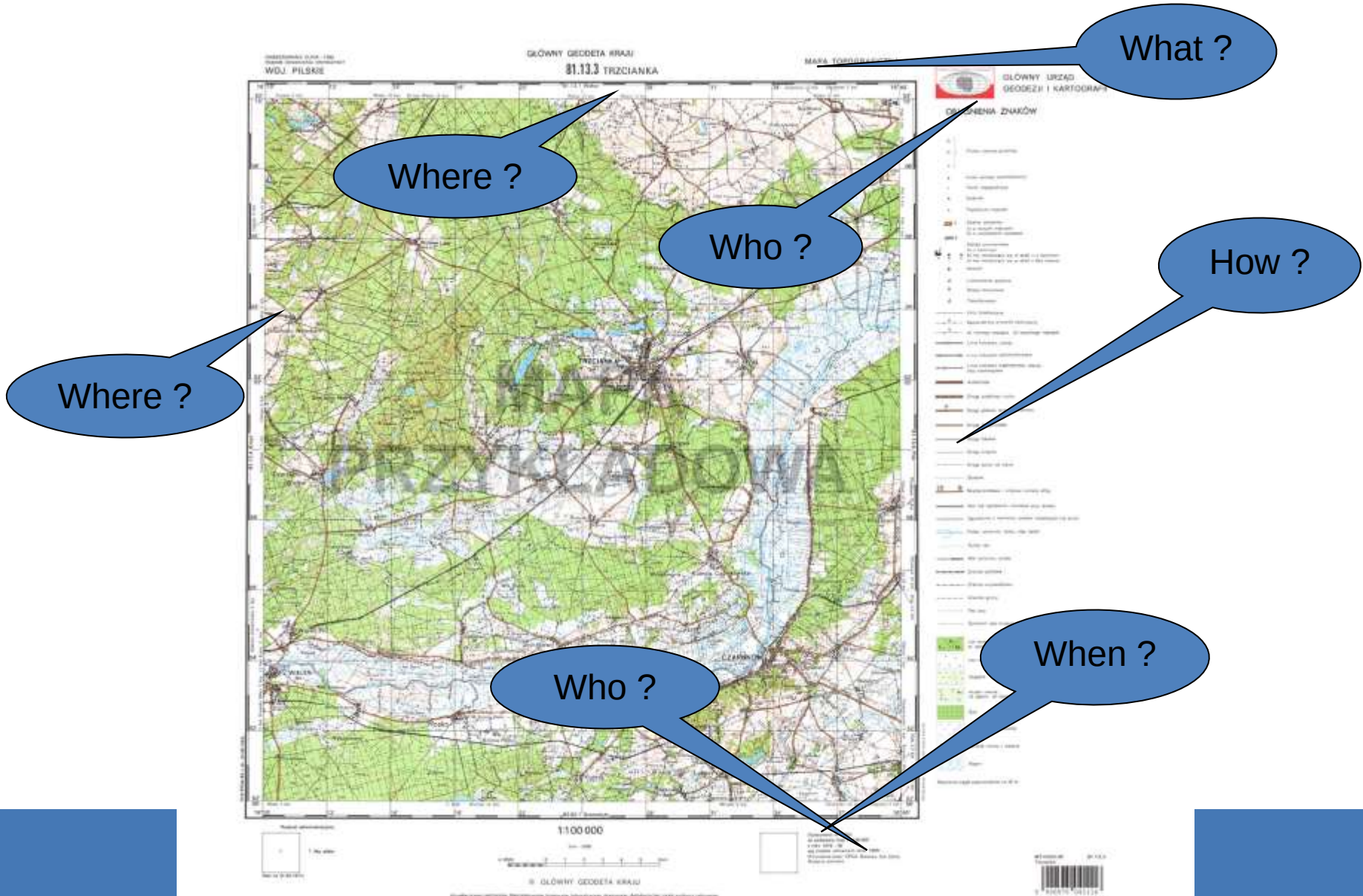
- Context-sensitive query analysis
- Discovering spatial data sets and services
- Understanding the semantics of discovered resources
- Integration resources
- Data analysis - attribute and spatial queries but how to ask the right question?

The main problem is the distributed and heterogeneity nature of datasets and services in SDI.

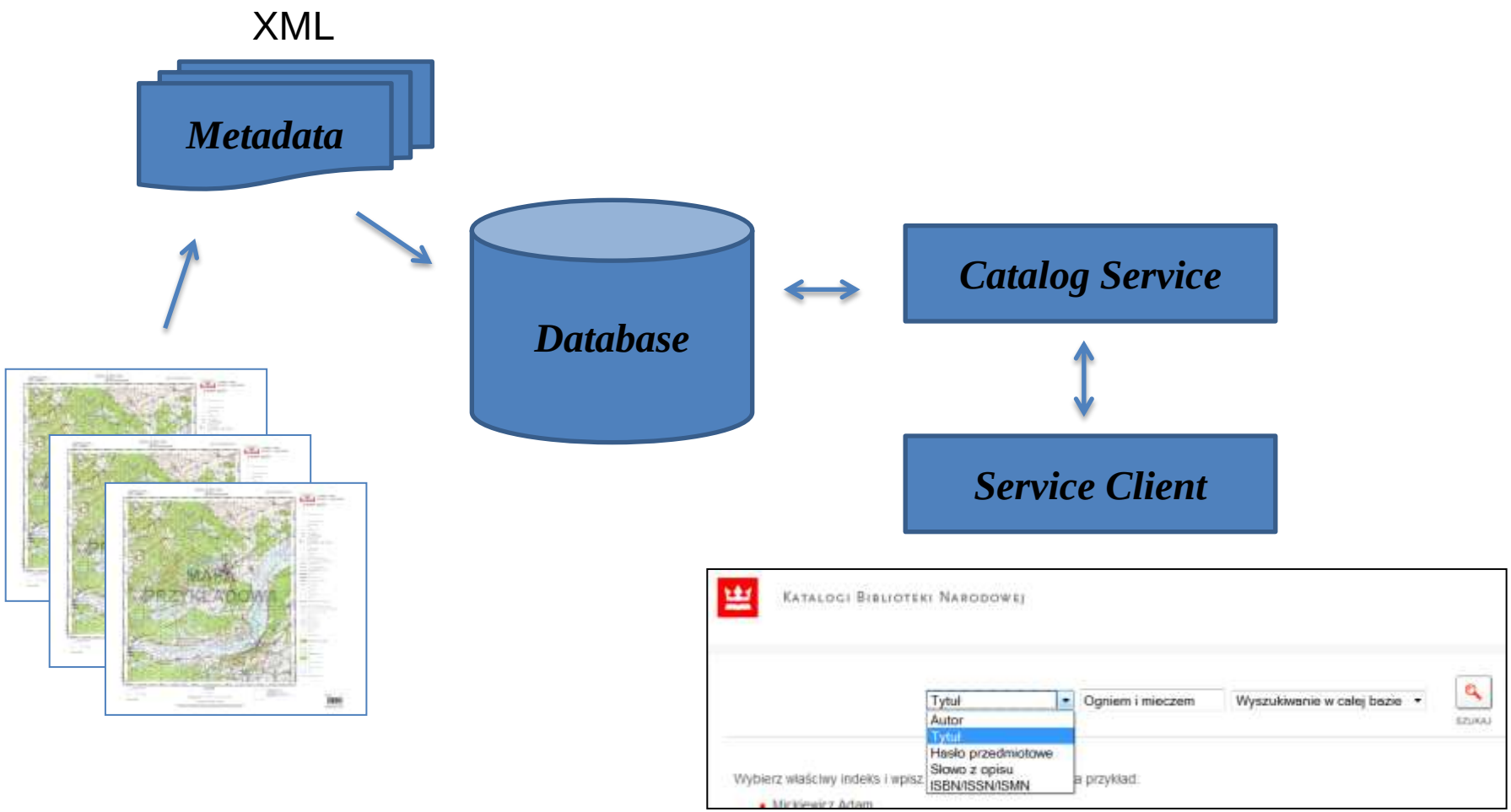
Possible solutions

- Creation of specialized software e.g. for android to search buildings in Polish cities
- Tagging buildings with a name of the owner - Google
- The use of semantic web technologies
- …
- Siri is an intelligent personal assistant and knowledge navigator which works as an application for Apple's iOS. The application uses a natural language user interface to answer questions, make recommendations, and perform actions by delegating requests to a set of Web services /wikipedia/

Discovering spatial data sets and services



Architecture of catalog services



Metadata of a book in XML

Title: With Fire and Sword

Author: Henryk Sienkiewicz

Publisher: Polish Publisher

Year of publication: 2010

Different ways of recording the metadata **– *XML***

```
<książka>
   <tytuł>Ogniem i mieczem</tytuł>
   <autor>Henryk Sienkiewicz</autor>
   <wydawnictwo> Polskie Wydawnictwo
Literackie</wydawnictwo>
   <rok_wydania>2010</rok_wydania>
</książka>
```

```
<pozycja_ks>
   <tytuł>Ogniem i mieczem</tytuł>
   <autor_ks>Henryk Sienkiewicz</autor_ks>
   <wyd> Polskie Wydawnictwo
Literackie</wyd>
   <data_wyd>2010-06-16</data_wyd>
</pozycja_ks>
```

ISO 19115

- Developed by ISO/TC 211 over the approximate period 1994-2002. The release of ISO 19115 "**Geographic Information - Metadata**" in 2003
- About 400 elements
- Serialization in XML – ISO 19139
- Excellent job but complicated

INSPIRE

COMMISSION REGULATION (EC) No 1205/2008

of 3 December 2008

implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata

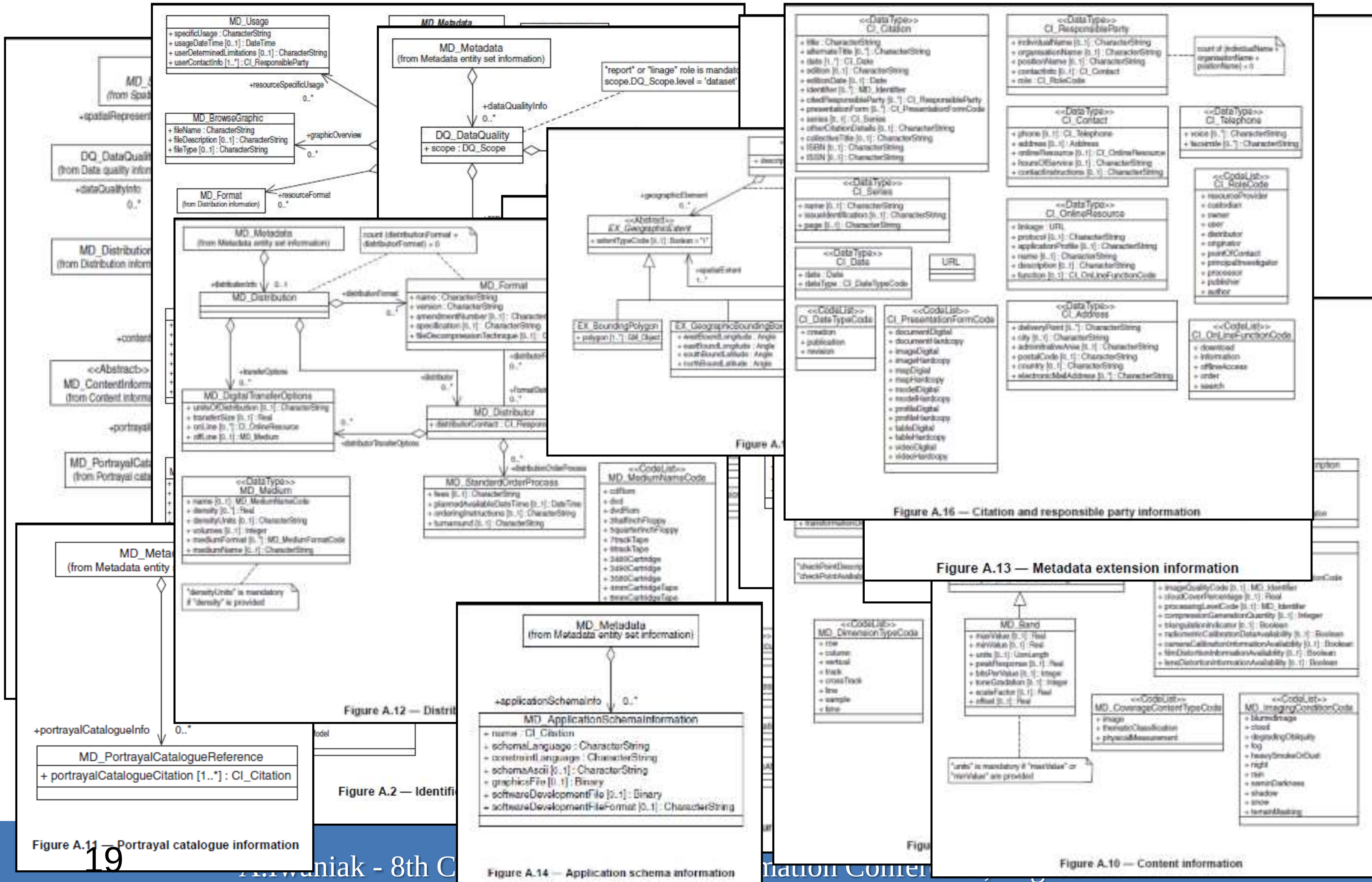
INSPIRE
Infrastructure for Spatial Information in Europe

INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119

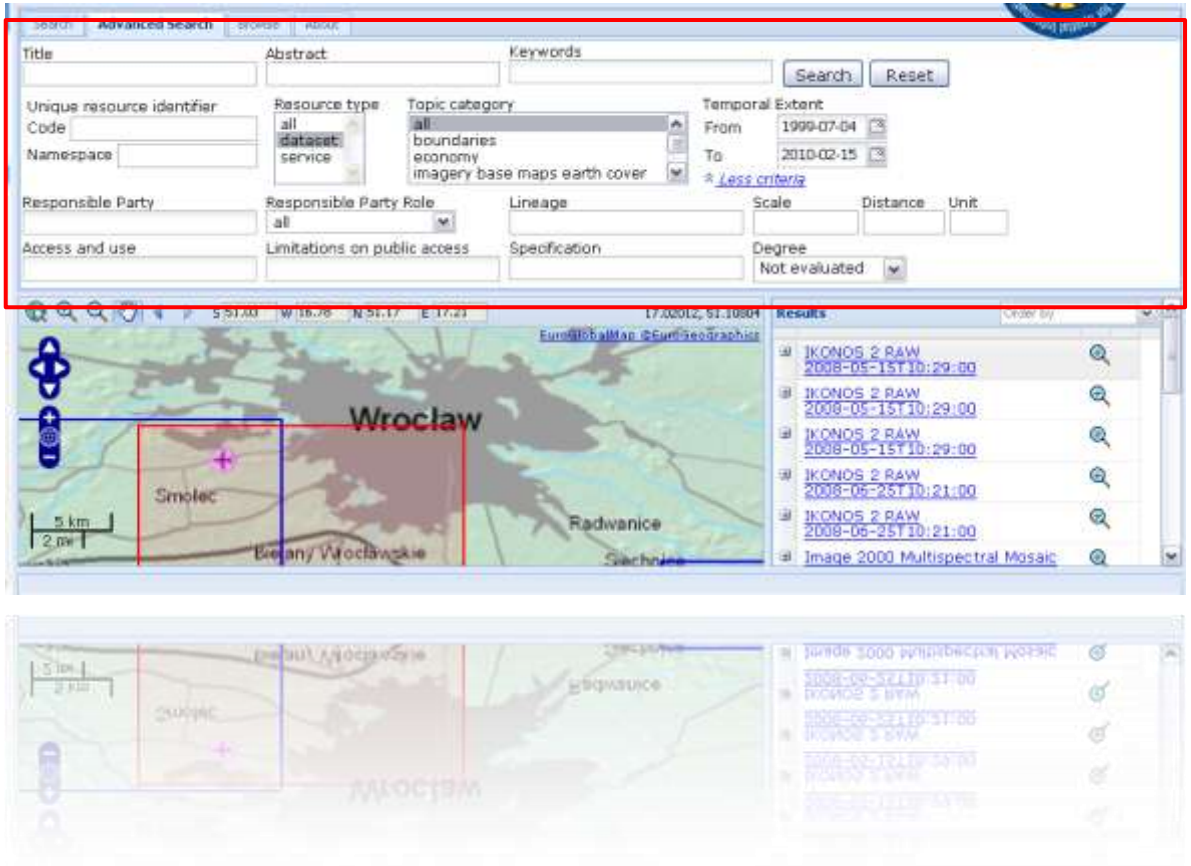
INSPIRE Metadata IR

- Simple file 15Kb-20Kb
- Complicated structure in XML

ISO 19115


MD_Metadata
(from Metadata entity set information)
+dataQualityInfo
DQ_DataQuality
+ scope : DQ_Scope
+applicationSchemaInfo
MD_ApplicationSchemaInformation
+ name : CI_Citation
+ schemaLanguage : CharacterString
+ constraintLanguage : CharacterString
+ schemaAscii [0..1] : CharacterString
+ graphicsFile [0..1] : Binary
+ softwareDevelopmentFile [0..1] : Binary
+ softwareDevelopmentFileFormat [0..1] : CharacterString
+portrayalCatalogueInfo
MD_PortrayalCatalogueReference
+ portrayalCatalogueCitation [1..*] : CI_Citation
Figure A.11 — Portrayal catalogue information
Figure A.12 — Distrib
Figure A.2 — Identifi
Figure A.14 — Application schema information
Figure A.16 — Citation and responsible party information
Figure A.13 — Metadata extension information
Figure A.10 — Content information

Functionality of metadata discovery portals



Limited functionality of metadata discovery portals

Discovery Geoportals search capabilities

- title
- abstract
- keywords
- etc.

Metadata files that meet the criteria

Discovery Geoportals search inabilities

to find all datasets and services that contain specified objects like: bridges, border points, plots

Semantics
Keywords: water
Search: rivers channels
Plots - parcels

The implementation of metadata in INSIRE is beautiful story of how to complicate a simple problem and increase its complexity a lot

Semantic Web

„The Web of meanings”

The extension of the current Web in which information is given well-defined meanings, better enabling computers and people to work in cooperation.

Scientific American, Berners-Lee at al.

The Semantic Web provides a common framework that allows data to be shared and reused accross application, enterprise, and community boundaries,

W3C

Semantic Web

Technologies and standards that allow machines to understand the semantics of information on the Web:

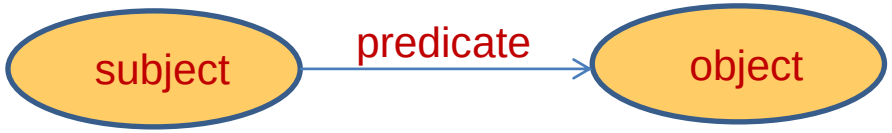
- Ontologies and thesauri
- Reasoning based on ontologies
- RDF
- RDFa
- RDFS, OWL, OWL2
- SPARQL
- SWRL

Thesaurus

- Can help to formally organize knowledge for a given domain
- Controlled vocabulary with hierarchical, equivalency or associative relations between concepts
- In Semantic Web encoded in SKOS standard (Simple Knowledge Organization System)
- Every concept in SKOS is indetified by a URI and expressed in RDF – machine readable

RDF – the basis of SW

- RDF – Resource Description Framework, initially a standard for encoding metadata
- Since 2004 W3C standard - a language used for describing any resources and their relations in the Web
- Information in RDF is expressed as a list of *statements (Subject-Predicate-Object)*



- The name of the resource and predicate can be identified by URI (*Uniform Resource Identifier*)

RDF

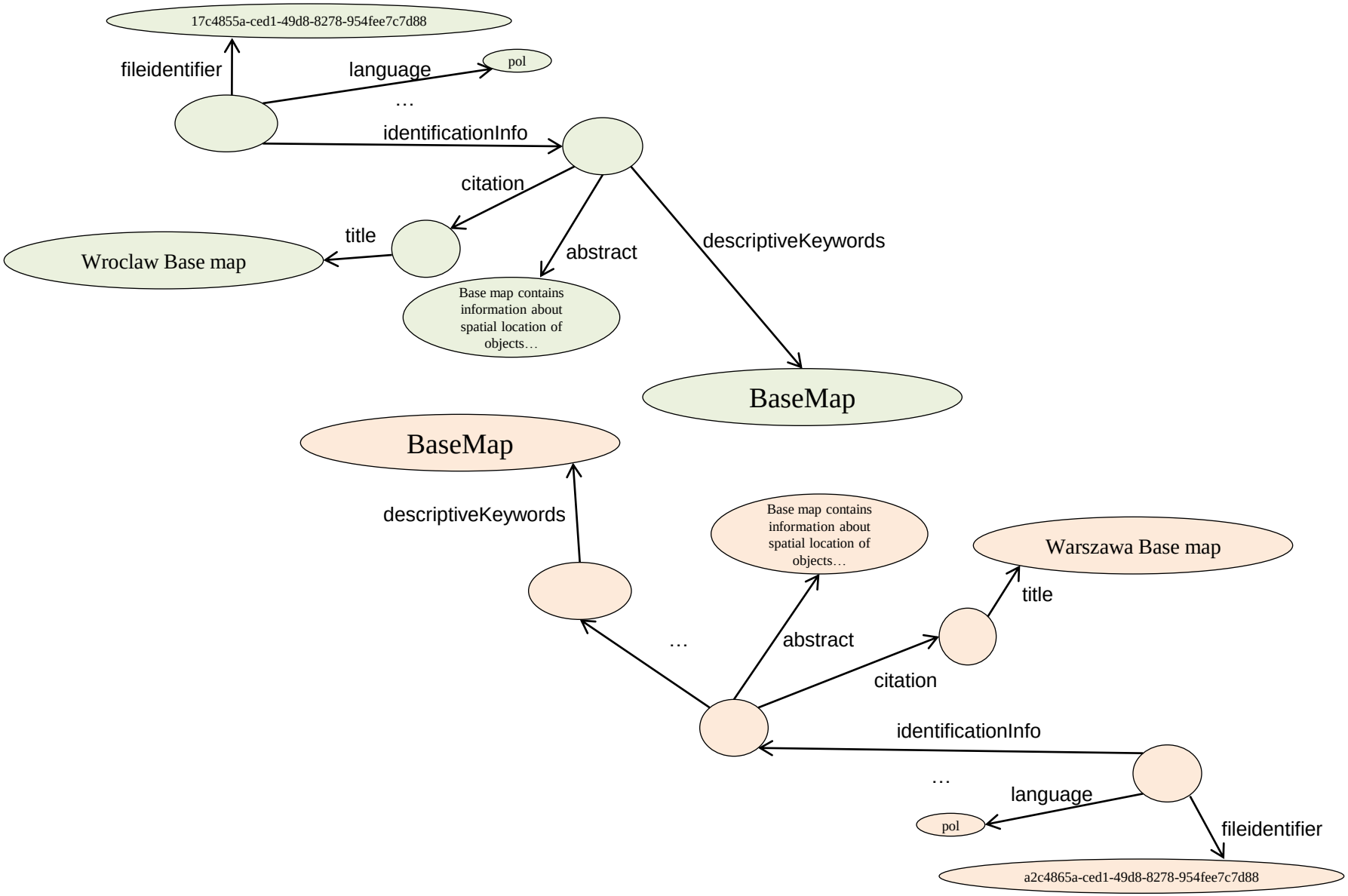
subject - predicate – object

<http://dbpedia.org/resource/Category:Constantinople>
<http://xmlns.com/foaf/0.1/page>
<http://pl.wikipedia.org/wiki/Kategoria:Konstantynopol>

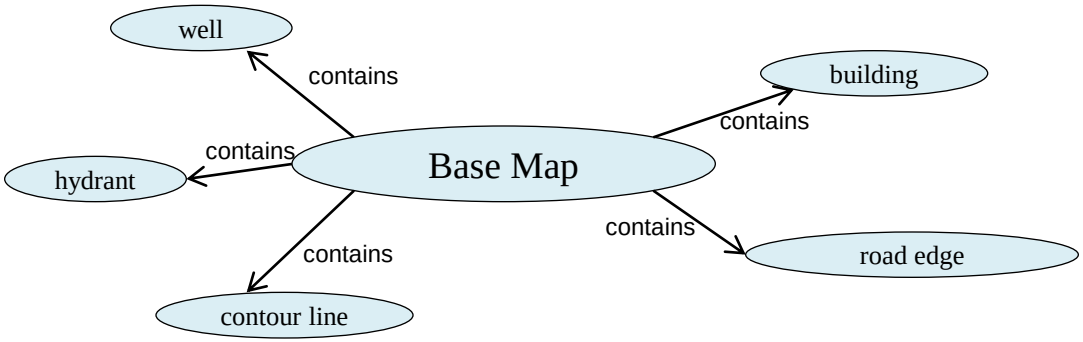
http://xmlns.com/foaf/0.1/page

http://dbpedia.org/resource/Category:Constantinople → http://pl.wikipedia.org/wiki/Kategoria:Konstantynopol

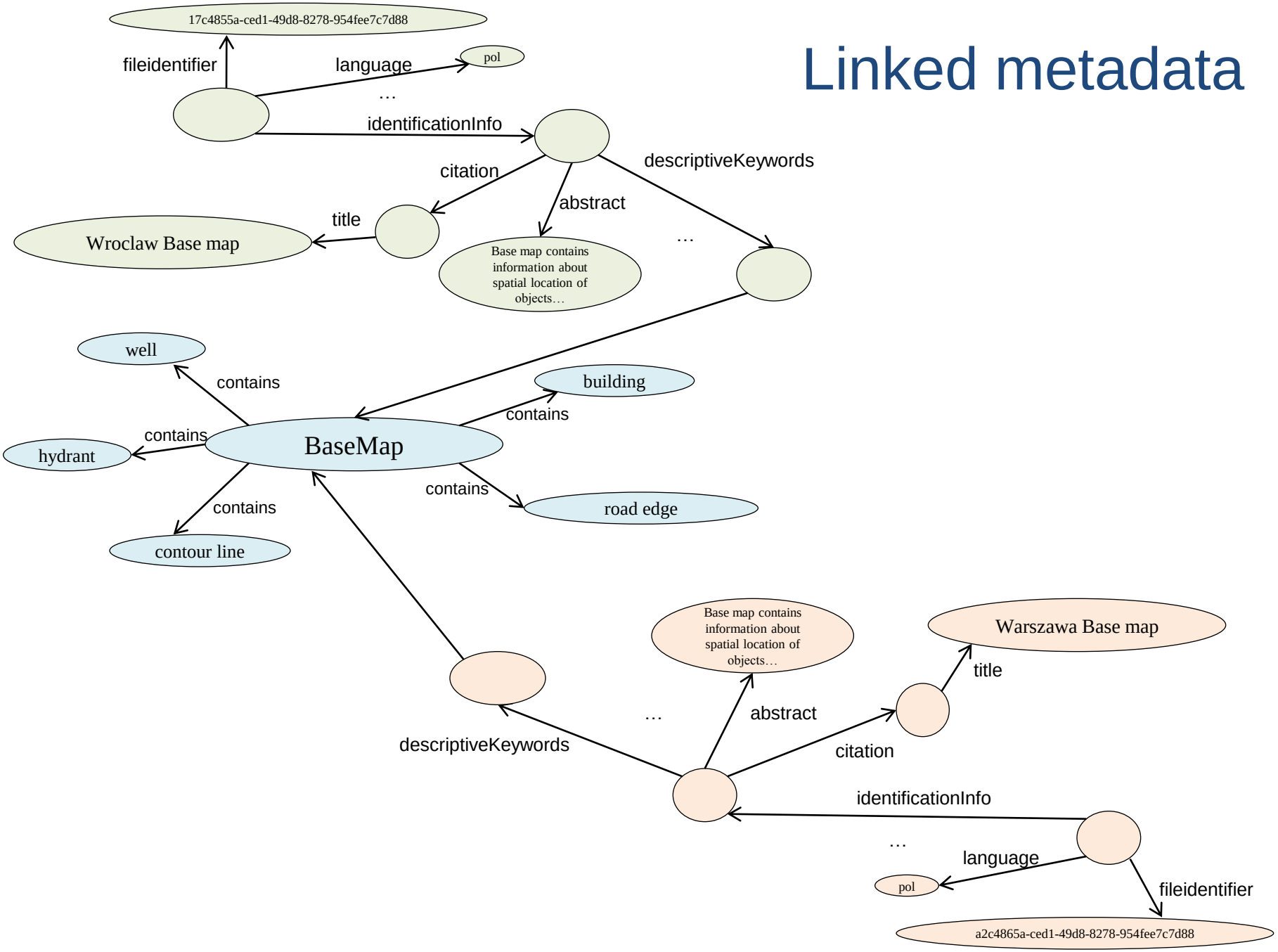
Metadata in RDF



Base Map in RDF



Linked metadata



Ontology

- Pillar of the Semantic Web
- The most popular definition:

  An ontology is a specification of a conceptualization. *(Gruber, 92)*

- More friendly:

  An ontology formally defines a common set of terms that are used to describe and represent an area of knowledge.

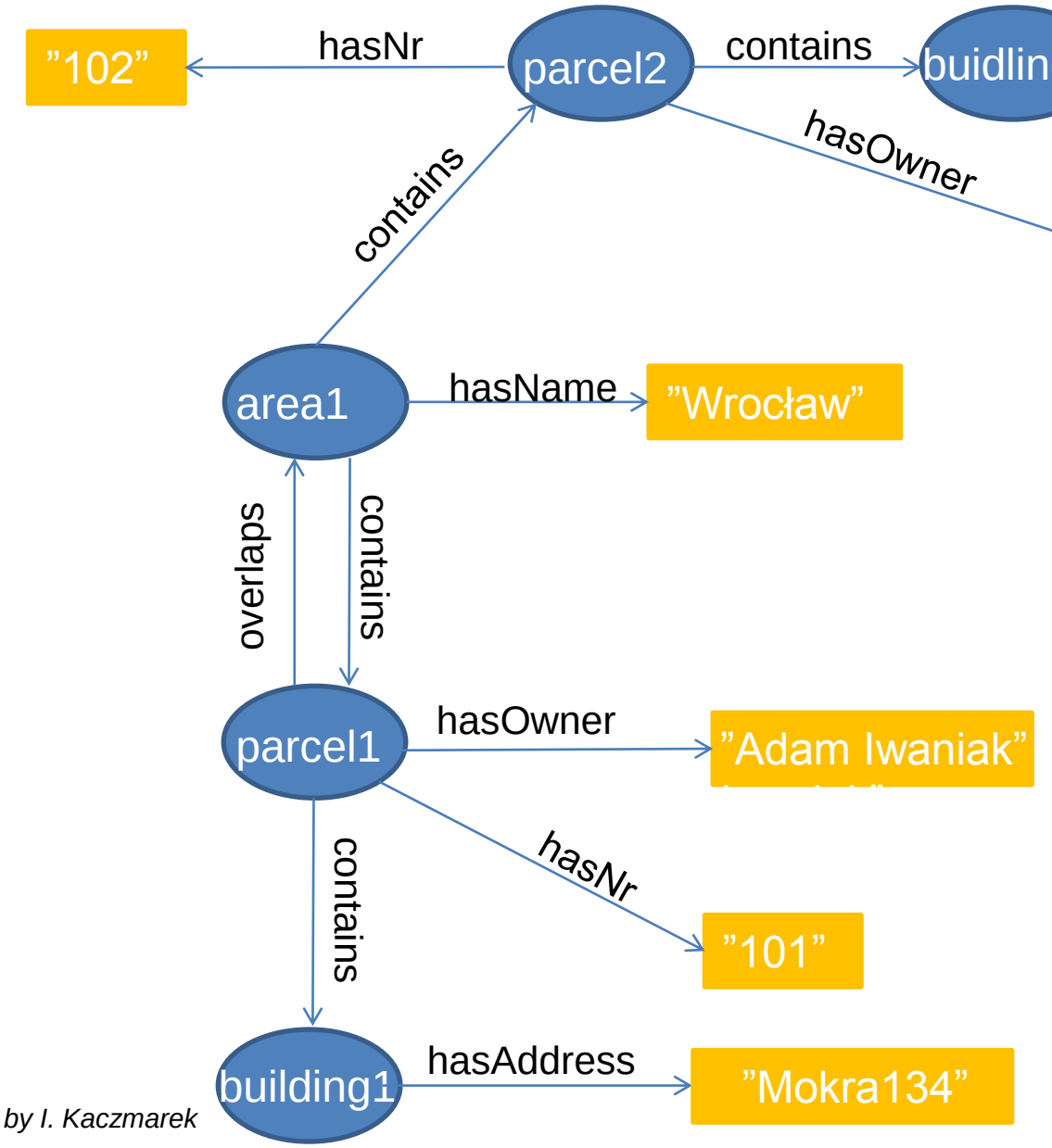
  (W3C, OWL Use Cases and Requirements)

- **Basic idea of ontology** – encode knowledge of the domain in such way it can be understood by computers

OWL

- Web Ontology Language
- The most popular when creating ontologies
- Three sublanguages:
  - OWL DL – maximum expressiveness with guaranteed computational completeness and decidability, based on Description Logic (DL)
  - OWL Lite – restricted subset of OWL DL, classification hierarchy and simple constraints, lower formal complexity than OWL DL (most used)
  - OWL Full - maximum expressiveness with no computational guarantees
- Each of languages have their advantages and disadvantages – compromise between expressiveness and the efficiency of the reasoning

RDF



by I. Kaczmarek

```
cd:area1
        cd:hasName ”Wrocław”
        spatial:contains :parcel1
        spatial:contains :parcel2

cd:parcel1
        spatial:contains :building1
        cd:hasOwner ” Adam Iwaniak”
        cd:hasNr ” 101”
        spatial:overlaps :area1

cd:parcel2
        spatial:contains :building2
        cd:hasOwner ”Jan Kowalski”
        cd:hasNr ” 102”

cd:building1
        cd:hasAddress ” Mokra 134”
cd:building2
        cd:hasAddress ” Sienkiewicza”
```

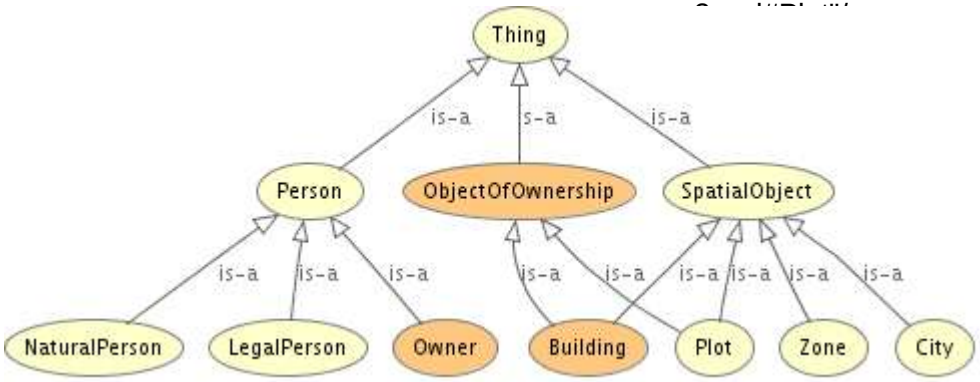
SPARQL query

Select all the buildings in Wroclaw which belong to Adam Iwaniak

```
PREFIX cd: http://rdf.parcelontology.com/
       spatial: http://geovocab.org/spatial/
SELECT ?building ?address
WHERE {?parcel cd:hasOwner „Adam Iwaniak”.
       ?parcel  spatial:contains ?building.
       ?parcel  spatial:overlaps ?area.
       ?area    cd:hasName „Wrocław”.
       ?building cd:hasAddress ?address}.
```

Answer: Building: budynek1, Address: Mokra 134

Building ontology in Protege



<Ontology1348457921722:isOwnedBy
rdf:resource="http://www.semanticweb.org/ontologies/2012/8/Ontology1348457921722.owl#JanKowalski"/>
<rdf:type
rdf:resource="http://www.semanticweb.org/ontologies/2012/8/Ontology1348457921722

<Ontology1348457921722:overlaps
w.semanticweb.org/ontologies/2012/8/Ontology1348457921722

Ontology1348457921722:SpatialObject>
)ntology1348457921722:owns>
ntology1348457921722:owns
w.semanticweb.org/ontologies/2012/8/Ontology1348457921722

lf:type
w.w3.org/2002/07/owl#NamedIndividual"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>
<Ontology1348457921722:owns
rdf:resource="http://www.semanticweb.org/ontologies/2012/8/Ontology1348457921722.owl#plot_05"/>
<Ontology1348457921722:owns
rdf:resource="http://www.semanticweb.org/ontologies/2012/8/Ontology1348457921722.owl#building_03"/>
<rdf:type
rdf:resource="http://www.semanticweb.org/ontologies/2012/8/Ontology1348457921722.owl#Owner"/>
<Ontology1348457921722:owns>
<owl:NamedIndividual
rdf:about="http://www.semanticweb.org/ontologies/2012/8/Ontology1348457921722.owl#plot_15">

by J.Łukowicz

Rule in SWRL

- Building(?building), OwnerOfPlot(?owner), Plot(?plot), isContainedBy(?building, ?plot), isOwnedBy(?plot, ?owner) -> owns(?owner, ?building)

SPARQL

- SELECT ?budynek FROM <http://www.semanticweb.org/ontologies/2012/8/Ontology1348457921722.owl#>
  WHERE {
  ?owner rdf:about myont:AdamIwaniak .
  ?owner myont:owns ?budynek .
  ?budynek myont:isLocatedIn myont:Wroclaw . }

WoGIS-Semantic Research Group

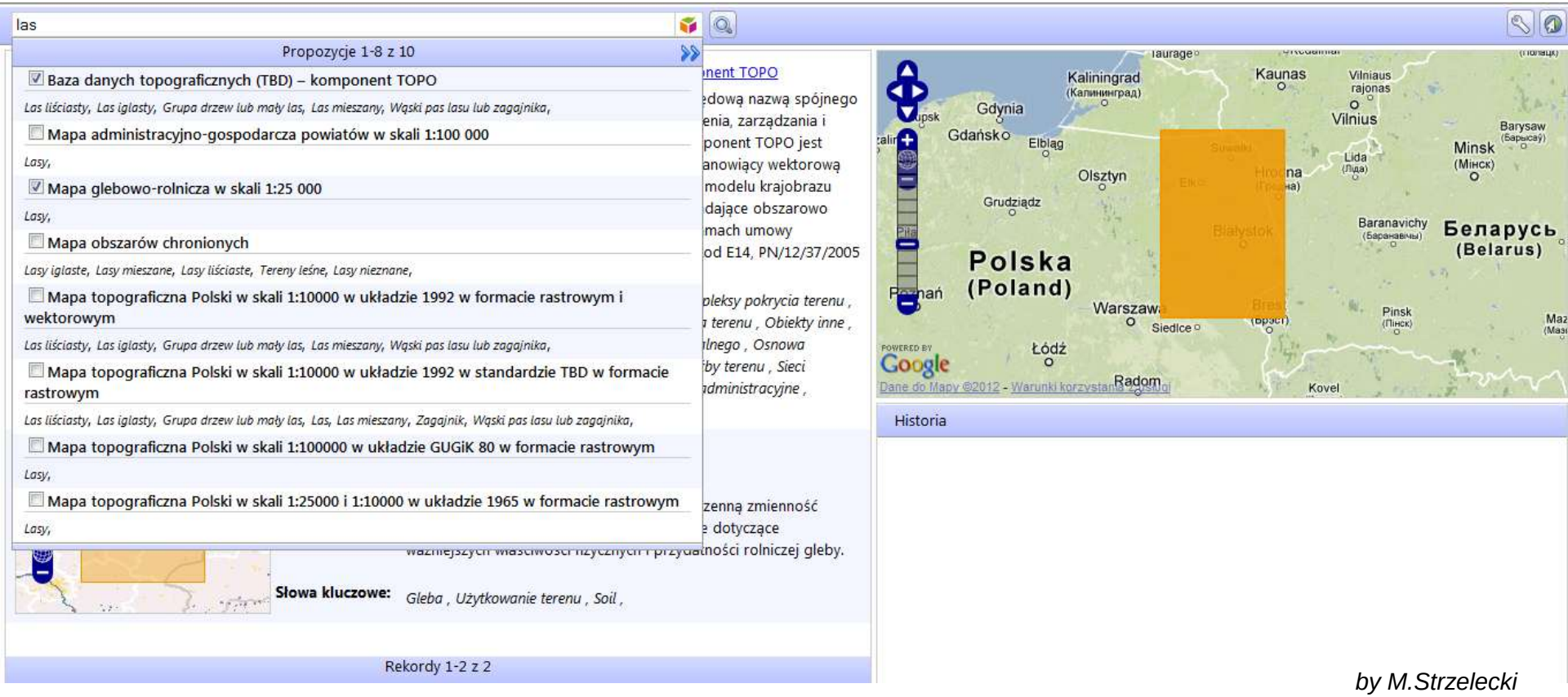
- Semantic metadata client
- Integration of spatial planning data
- Building spatial decision support systems based on SW technologies
- Creation the metadata for use
- SemGIS – environment for creating GIS systems that use ontologies



Iwona Kaczmarek
Jaromar Łukowicz
Marek Strzelecki

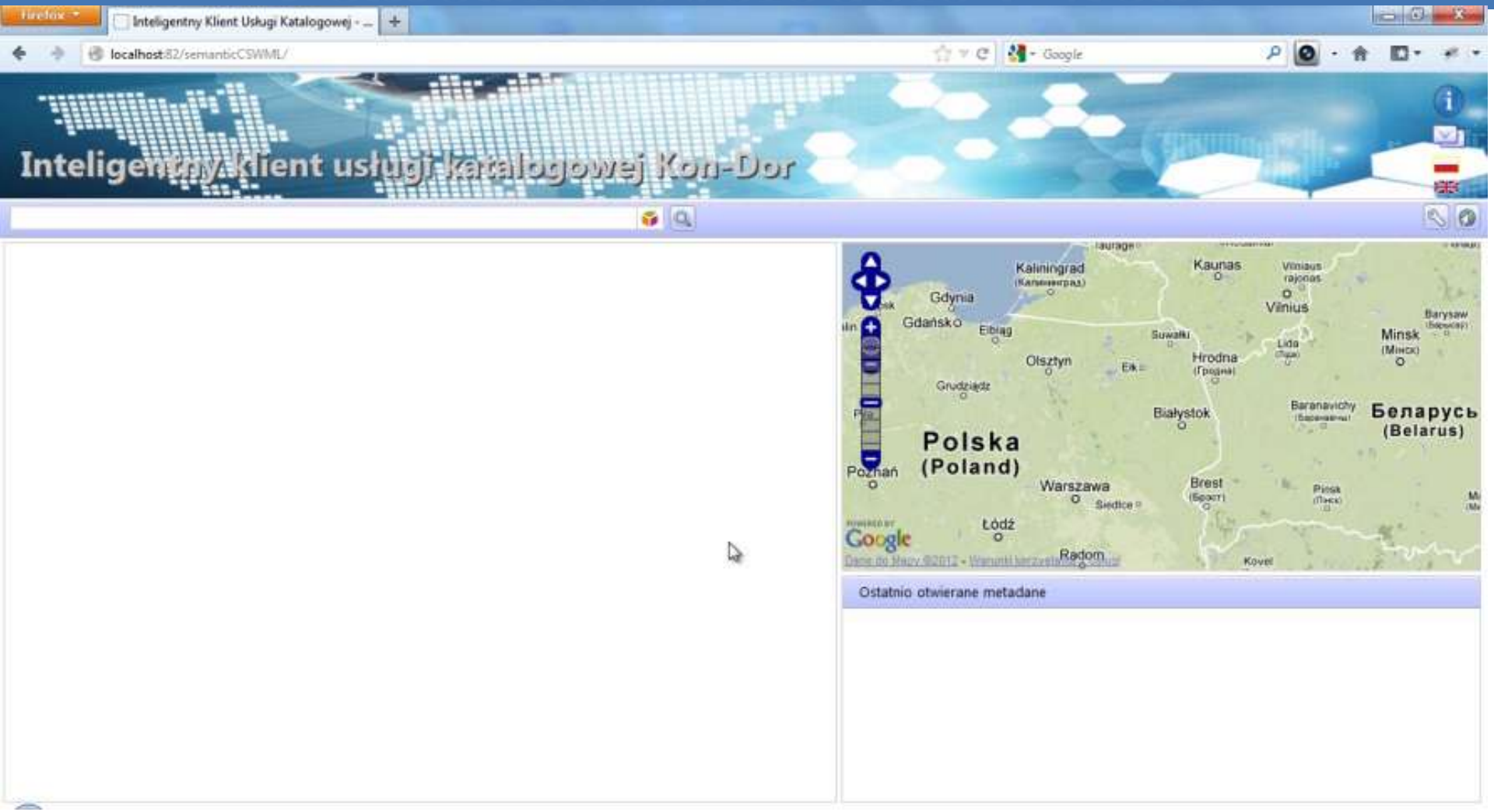
Semantic Metadata Client

http://www.psip.wrotapodlasia.pl/semantic_client



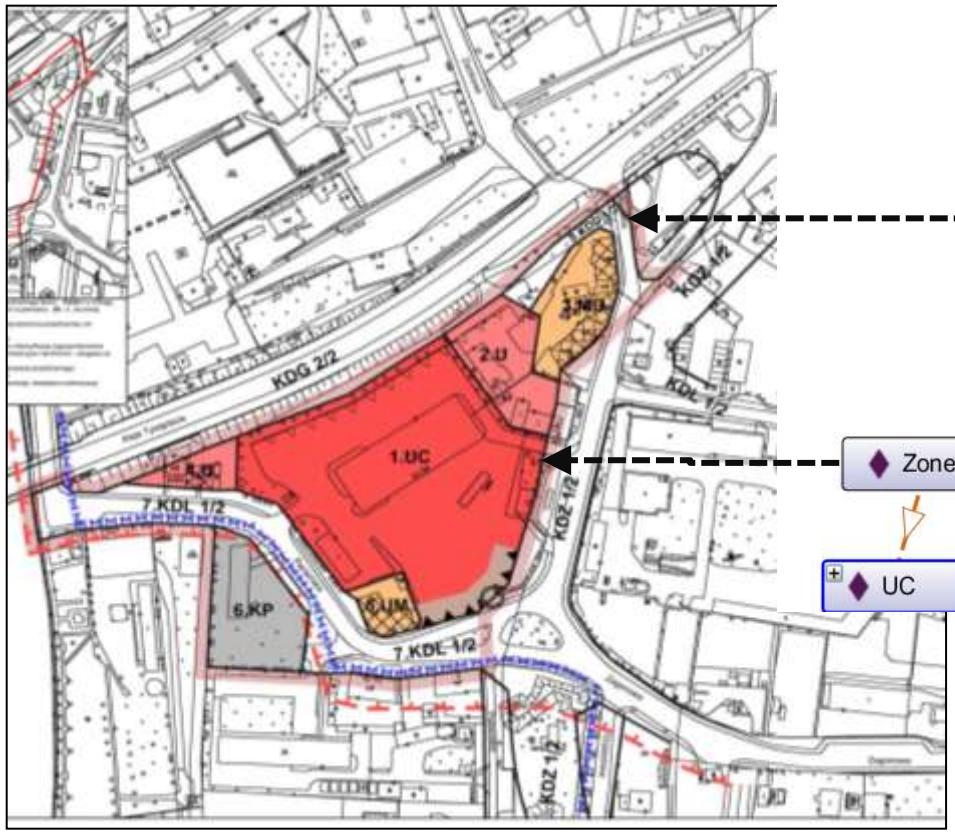
by M.Strzelecki

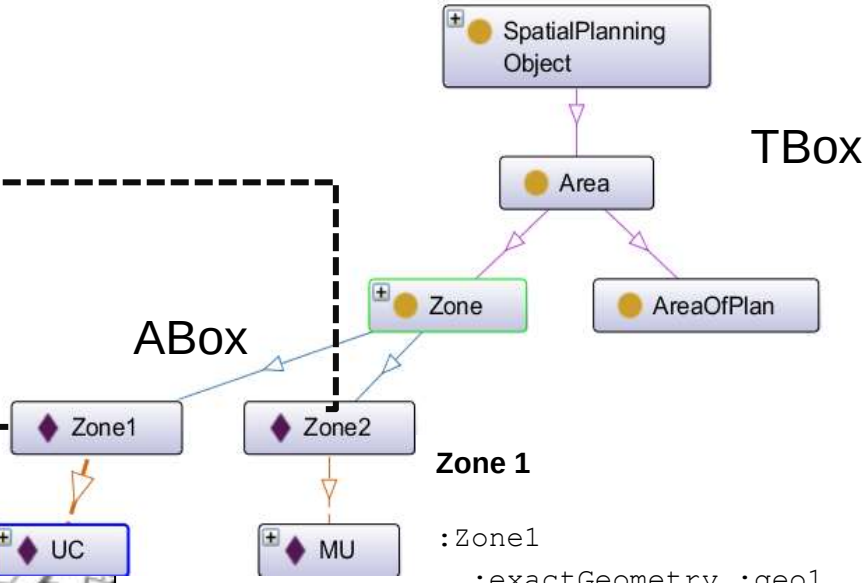
Semantic Metadata Client



by M.Strzelecki

Integration of spatial plannig data





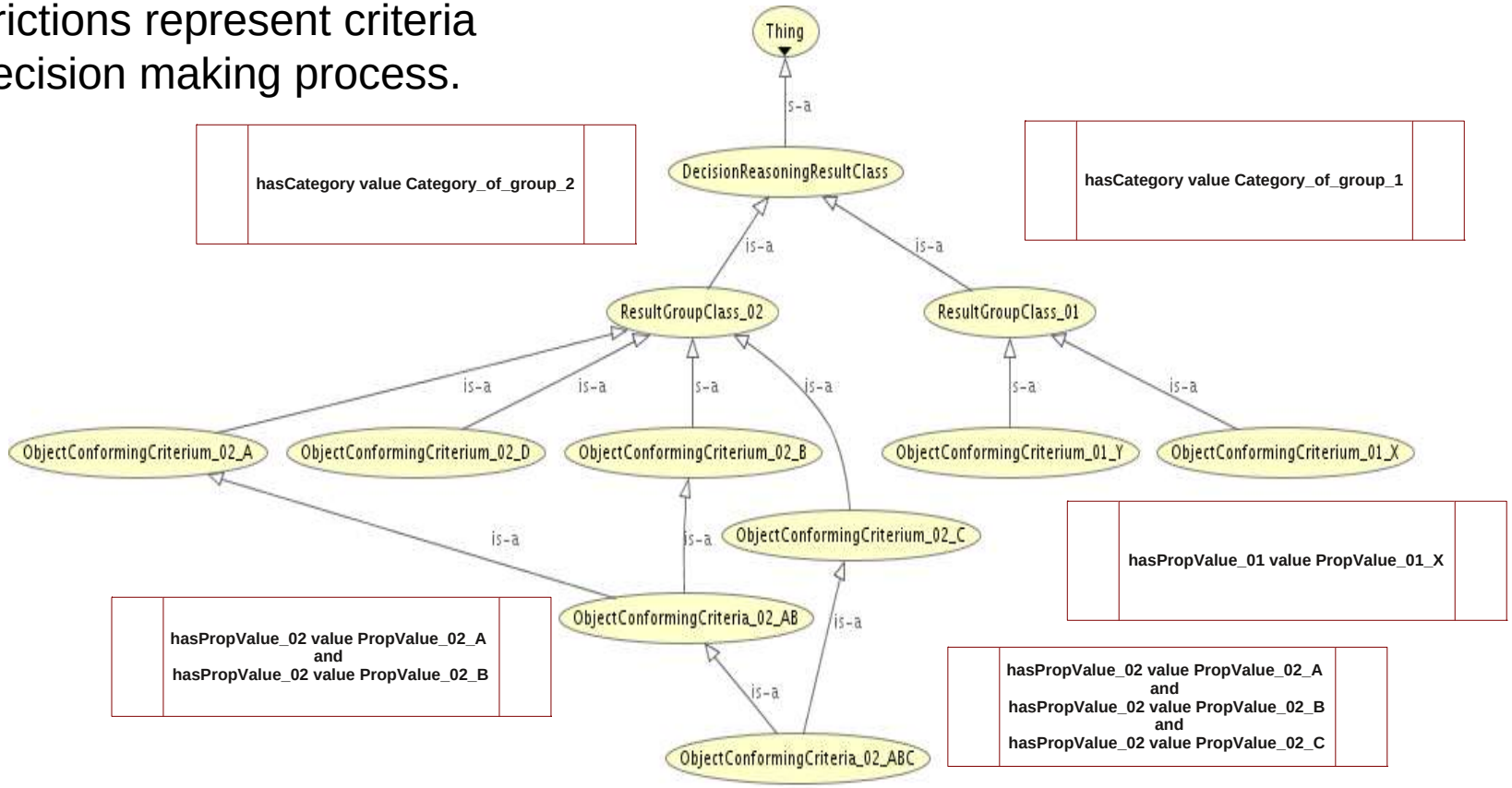
Zone 1

```
:Zone1
  :exactGeometry :geo1
  :geo1 rdf:type ogc:Polygon
  :geo1 ogc:asWKT "Polygon((…))
          "^^ogc:WKTLiteral
  :hasDesignation :UC
  :hasFormOfDevelopment :FD1
  :hasParcelArea „2000"
  :hasBiologSurface „70"
  :refersTo :Plan1
  :hasBuiltUpRatio :Ratio2
```

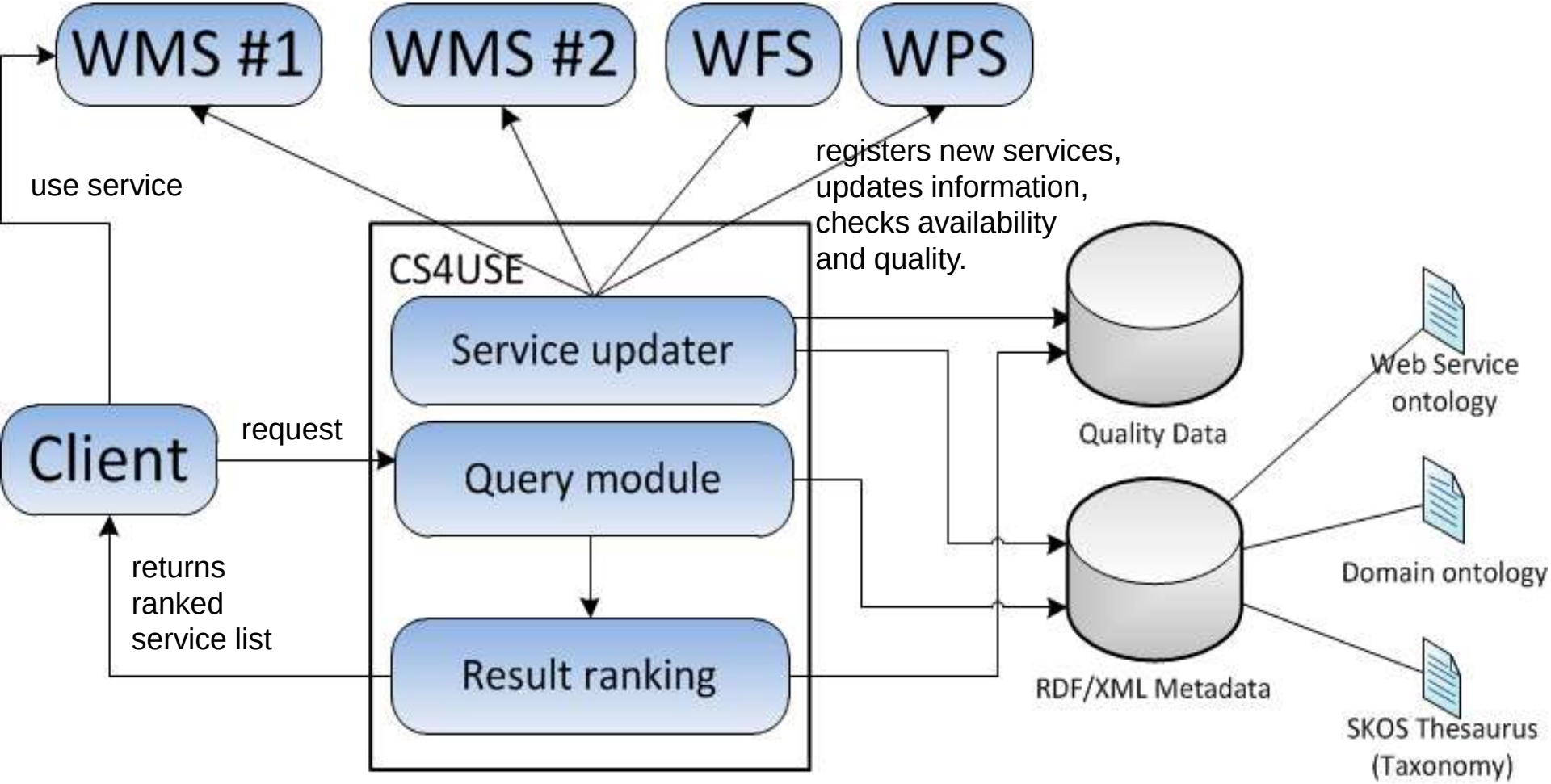
Decision Support Systems for zoning decisions

Decision criteria set:

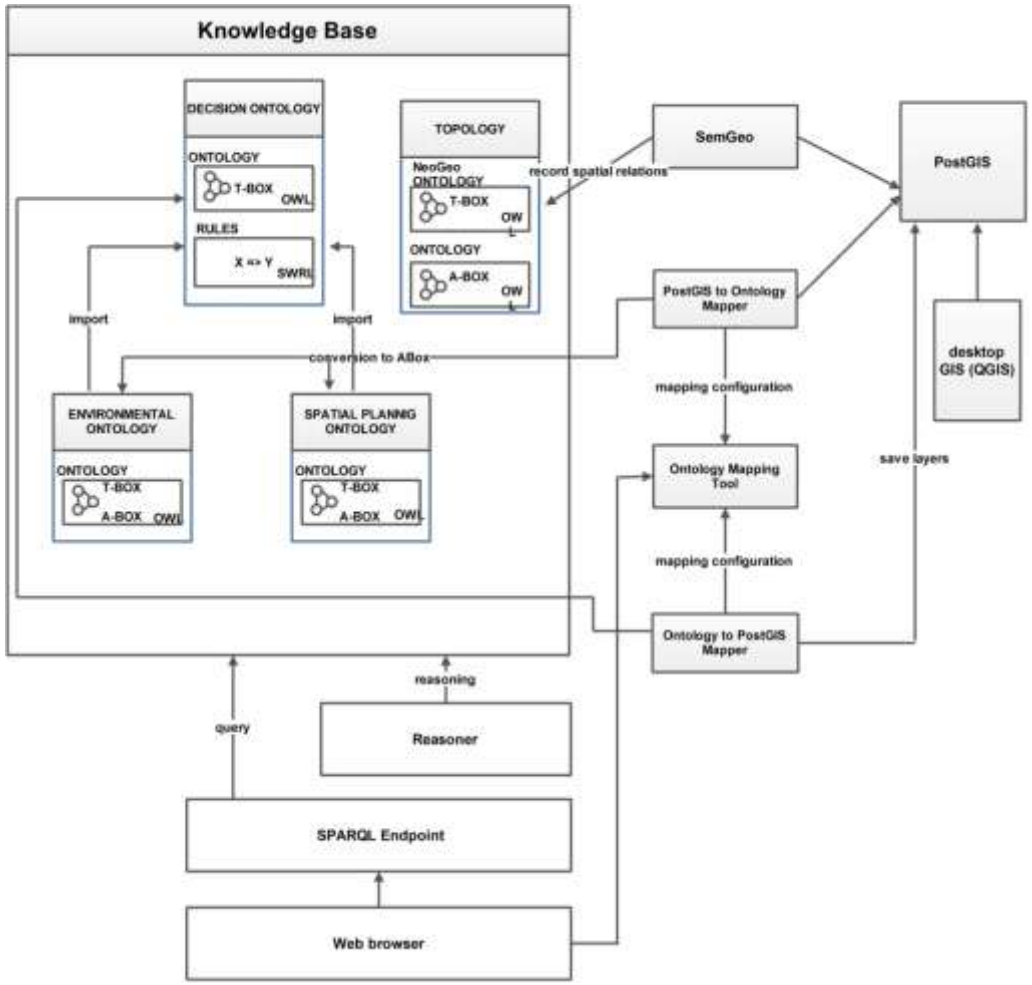
- Result classes hierarchy. Each result class is defined by restrictions on selected properties.
- Restrictions represent criteria for decision making process.



Metadata for use



Architecture of SemGIS



Intelligent Spatial Information Infrastructure

- **Spatial Data Infrastructure**

  retrieval, publication and access to spatial data
- **Spatial Information Infrastructure**

  Interpretation of data, analysis and processing
- **Intelligent Spatial Information Infrastructure**

  the ability to understand the data and services across systems significantly facilitating the information retrevial and integration, automated geospatial web services orchestration and inference new facts based on existing ones.

Summary

- Research on ontologies in SDI has great potential
- It is conducted by many leading academic centers, but the number of practical solutions is limited
- However there are attempts to extend existing SDI services with semantic layer, and use a new gespatial data modeling paradigm - GeoSPARQL
- The use of thesauri can significantly increase functionality of directory services (and support multilingualism)
- The popularity of smartphones could be one of the key drivers of research on ontologies.

Thank you for your attention

adam.iwaniak@gmail.com