Trentino government
linked open geodata: a case study

Pavel Shvaiko, Informatica Trentina
Feroz Farazi, University of Trento
Vincenzo Maltese, University of Trento
Alexander Ivanyukovich, Trient Consulting Group
Veronica Rizzi, University of Trento
Daniela Ferrari, Segreteria SIAT, PAT
Giuliana Ucelli, Segreteria SIAT, PAT

Boston, ISWC-2012
Trentino landscape

Sources:
• Coutesy of Autonomous Province of Trento
• World Wide Web
Outline

- The application setting
  - Semantic geo-catalogue
  - Linked open government data
- Case study
  - Publishing Pipeline
  - Lessons learned
- Conclusions
Application domain: eGovernment

By eGovernment we mean an area of application for ICT to modernize public administration by optimizing the work of various public institutions and by providing citizens and businesses with better (e.g., more efficient) services as well as with the new services (that did not exit before).
Semantic geo-catalogue

- The geo-catalogue offers a standard mechanism to classify, describe and search (also at the conceptual level) information on geo-data and geo-services.

- This is a web-based instrument that extends the existing portfolio of services of the Trentino geo-portal.
Guidelines: european, national, local

- European level: The INSPIRE Directive aims at creating the framework for sharing of spatial information

- National level: DigitPA produced the so-called Repertorio Nazionale Dati Territoriali

- Local level: various PAT deliberations orienting the development and evolution of SIAT towards INSPIRE
Typical scenario: data view

Q: bicycle tracks, Riva del Garda, 2012

Data

MetaData

Catalog(s)

INSPIRE

Spatial data themes (34)
- Annex-I (9 themes)
- Annex-II (4 themes)
- Annex-III (21 themes)

ISO topic categories (19)
- Environment
- Planning and Cadastre
- ...
Architectural view (INSPIRE)
Discovery services

- OGC
- SOAP/WSDL
- SA-WSDL

Metadata
(ISO 19119)

Metadata
(ISO 19115)

Discovery Services
- OGC CSW 2.0.2
- UDDI

Catalogs

Applications

Geocartografic Portal
(BEA ALUI)

Data
Outline

- The application setting
  - Semantic geo-catalogue
  - Linked open government data
- Case study
  - Publishing Pipeline
  - Lessons learned
- Conclusions
Open Government Data (OGD)

- Data owned by the government authorities
- Making some data available in an open (from both legal and technical perspectives) manner
OGD: benefits

- Increased transparency and simplification for the public administration
- Potential economic growth through data reuse
- Increased participation of citizens in public administration
Linked open data

- Equal access for all
- Licensing
- Transparency
- Changes PA

Open data

Easier to find
Easier to integrate
Comply with standards
Linked Data

Data publishing principles:
1. Resources are identified by URIs
2. URIs are dereferenceable
3. When a URI is dereferenced, a description of the identified resource should be returned
4. Published web data sets must contain links to other web data sets

Five star rating:
★ publish data on the web in any format
★ ★ use structured data format
★ ★ ★ use non-proprietary format
★ ★ ★ ★ use universal format to represent data, that encapsulates both syntax and semantics
★ ★ ★ ★ ★ link data to other data sets on the web to provide context
Outline

- The application setting
  - Semantic geo-catalogue
  - Linked open government data
- Case study
  - Publishing Pipeline
  - Lessons learned
- Conclusions
Case study: an overview

Context
- Datasets used in the semantic geo-catalogue application
- The datasets consist of both data (shape files) and metadata (xml)

Production of linked open geo-data
- by following Open Government Data paradigm
- by following Linked Open Data paradigm

Goals
- To obtain insights on how the services of PA can be improved
- To create evidence on the expected benefits with reduced costs
- To create a critical mass
Deliberation n195 of 17/02/2012

Authorization to release some core (160) geographic datasets following the OGD paradigm
Publishing pipeline

Conversion  Linking  Sharing  Evaluation
Conversion

Data in shp files ➔ GeoTools ➔ Data in XML ➔ SAX parser ➔ Jena

MetaData in XML
Geo-metadata conversion

<rdf:Description rdf:about =
   "http://www.territorio.provincia.tn.it/geodati/p_tn:piste_ciclabili">
   <dc:language>it</dc:language>
   <dcmibox:westlimit>10.41</dcmibox:westlimit>
   <dc:identifier>http://www.naturambiente.provincia.tn.it/</dc:identifier>
   <dc:format>shp</dc:format>
   <dc:rights>Dato pubblico</dc:rights>
   <dc:title>Piste ciclabili</dc:title>
   <dc:creator>Dipartimento Risorse Forestali e Montane</dc:creator>
</rdf:Description>
Geo-data conversion

Class

Instance

Geometric Shape Polyline

```
<rdf:Description rdf:about="http://www.territorio.provincia.tn.it/geodati/resource/piste_ciclabili"/>
<owl:sameAs rdf:resource = "http://rdf.freebase.com/ns/guid.9202a8c04000641f8000000000428308"/>
</rdf:Description>

<rdf:Description rdf:about = "http://www.territorio.provincia.tn.it/geodati/resource/piste_ciclabili/529">
<geontology:length rdf:datatype = "http://www.w3.org/2001/XMLSchema#double">1445.8484810675</geontology:length>
<rdfs:label xml:lang = "it">Mori - torbole</rdfs:label>
</rdf:Description>

<rdf:Description rdf:about = "http://www.territorio.provincia.tn.it/geodati/resource/piste_ciclabili_529">
</rdf:Description>

<rdf:Description rdf:about = "http://www.territorio.provincia.tn.it/geodati/resource/piste_ciclabili_529">
<geontology:polyline>646339.346896746,5082179.74045936 645575.851739799,5081173.68539361</geontology:polyline>
</rdf:Description>
```
Links

- Links were established to external data sources
- High quality of the links were guaranteed through validating them manually

Diagram:
- DBpedia
  - Segregated_cycle_facilities
  - Drainage_basin
  - 592728
  - 3744753
- TrentinoGeoData
  - piste_ciclabili
  - bacini_idrografici_principali
  - Adige
  - Sarca
- Freebase
  - segregated_cycle_facilities
  - drainage_basin
  - adige
  - sarca
Sharing

Portale Geocartografico Trentino

S.I.A.T.
Cartografia di base
Interoperabilità - Servizi WMS
Interoperabilità - Servizi WFS
Lidar
Ricerca nel Geo-catalogo
Carta Tecnica Provinciale
Beni librari ed archivistici
Edilizia scolastica
Geologia
Meteotrentino
Minerario
Organizzazione e qualità attività sanitarie
Recupero ambientale e urbanistico aree industriali
Risorse Forestali e Montane
Sistema Informativo Ambientale - APPA
Territorio rurale
Urbanistica e tutela del paesaggio
Utilizzazione Acque Pubbliche
Valutazioni ambientali

Home | S.I.A.T. | Ricerca nel Geo-catalogo

Ricerca
Ricerca per contenuto

piste ciclabili

Inizia la ricerca

Il geo-catalogo permette di classificare, descrivere e ricercare informazioni relative a geo-dati e geo-servizi secondo le specifiche tecniche del Repertorio Nazionale dei Dati Territoriali del DigitPA. I servizi di accesso ai tematismi Open Government Data, con licenza Creative Commons Zero - CCZero e in formato RDF, sono attualmente in fase sperimentale.

Risultati: 1-1 (1)

PISTE CICLABILI

Contatto: Dipartimento Risorse Forestali e Montane
Parole chiave: SIAT, piste, ciclabili, RNDT

DESCRIZIONE
Lo strato informativo rappresenta il reticolato delle piste ciclabili del Trentino

[ ] Scarica dati
[ ] Scarica RDF

[ ] Metadato
[ ] Scarica XML
Evaluation: mash-up example
Evaluation: regional workshop

- Good participation from PA, industry, and research
- A local start-up company SpazioDati was asked to use the released data sets
- In one week:
  - TINDES – a naturalistic index
  - 32 PAT datasets
  - 9 OpenStreetMap datasets
Outline

- The application setting
  - Semantic geo-catalogue
  - Linked open government data
- Case study
  - Publishing Pipeline
  - Lessons learned
- Conclusions
Lessons learned

Conversion. URIs, which patterns to adopt: self-explanatory vs. GUIDs?

Linking. Research tools are still not flexible and precise enough, hence manual process is often preferred

Sharing. Adoption of the CreativeCommonsZero license as well as the use of the geo-portal, being a single point of access for geo-data were well received

Evaluation. A mash-up and workshop indicated that the approach adopted was a useful tactic. Feedback received on what would be useful to have:

- APIs with SLAs to give access to the fresh data to built services on top of it
- Feedback loop with citizens to improve data
Conclusions

- Experimental work on releasing some (40%) core geographic datasets as linked open government data
- Representation language and vocabulary
  - RDF for representing both the data and metadata
  - WGS84 vocabulary for data and Dublin Core for metadata
  - OWL for linking to external sources: DBPedia and Freebase
- Vertical tactical experimentation to gain momentum and engagement with the stakeholders. This was a success and it prepared and has opened the road for a larger transversal initiative
Thank you for your attention!

Pavel Shvaiko, PhD.
Innovation and Research Manager
TasLab, Informatica Trentina S.p.A.
Trento, Italy
Pavel.Shvaiko@infotn.it