

Sviluppi e semantica nel Portale del Servizio Geologico d'Italia

Valentina Campo, Carlo Cipolloni, Maria Pia Congi, Daniela Delogu

ISPRA – Istituto Superiore per la Protezione e la Ricerca Ambientale, Via V. Brancati 48, Roma;
valentina.campo@isprambiente.it, carlo.cipolloni@isprambiente.it

Riassunto

Il portale del Servizio Geologico d'Italia (Portale SGI) ha esteso le sue funzionalità, implementando un modulo specifico per sfruttare l'interoperabilità semantica nella ricerca di dati e metadati e pubblicando come LinkData il Thesaurus Italiano di Scienza della Terra, sviluppato e predisposto dalla biblioteca ISPRA insieme all'unità di gestione e sviluppo del portale e delle ontologie del Servizio Geologico anch'esso parte di ISPRA.

Abstract

The Geological Survey of Italy Portal (Geoportal SGI) has developed an extensional module that is able to use the semantic interoperability in the data and metadata searching, it's also published as LinkData the Italian Thesaurus of Earth Science, based and developed on the ISPRA library unit together with Database and ontologies management unit of Geological Survey unit.

Introduzione

Da diversi anni il Servizio Geologico d'Italia ha realizzato un portale con cui permettere l'accesso a specifici dati spaziali. La realizzazione dell'infrastruttura (SDI) era partita già nel 1990 e, allo stato attuale, è composta da oltre 15 geodatabase relazionali contenenti dati geologici e geotematici.

Il portale del Servizio Geologico d'Italia (Portale SGI) disponibile al seguente indirizzo <http://sgi.isprambiente.it/geoportal>, è basato sulla piattaforma tecnologica ESRI open-source Geoportal Server 1.2.2 e, nel corso dell'ultimo anno, un gran lavoro è stato fatto per migliorare le capacità di ricerca ed interoperabilità di questo strumento, inserendo per esempio nel motore di ricerca un Thesaurus multi-lingua specifico per le Scienze della Terra denominato ThIST (Thesaurus Italiano di Scienze della Terra).

In questo modo utilizzando l'interoperabilità semantica anche attraverso l'uso di strumenti quali i LinkOpenData, sono stati impiegati specifici thesauri di definizione al fine di migliorare l'interfaccia di ricerca, facilitando così l'utente nell'individuare l'oggetto della sua ricerca. Tale implementazione ha dimostrato efficacemente come l'uso dei thesauri e dei servizi ontologici collegati a metadati e dati, migliori la qualità della ricerca mettendo in evidenza un maggior numero di record di risposta e arricchendo l'informazione ad essi associata.

Il ThIST grazie al progetto europeo eENVplus (<http://www.eenvplus.eu>) è oramai interlinkato ad altri thesauri e vocabolari specifici sui temi ambientali e in particolare risulta un Link Open Data disponibile nella piattaforma semantica del progetto eENVplus nota come LusTRE (<http://linkeddata.ge.imati.cnr.it:2020/directory/ThIST>).

All'interno del portale SGI l'uso di questo thesaurus può essere combinato con altri servizi ontologici come il GEMET aumentando in modo esponenziale le capacità esplorative dello stesso motore di ricerca.

Inoltre il Geoportale, sviluppato dal Servizio Geologico d'Italia, supporta anche termini multilingua, facilita la ricerca delle informazioni peraltro in ambito della direttiva INSPIRE volta a mettere in condivisione informazioni a livello europeo.

Nello stesso tempo sono stati messi a punto ulteriori strumenti di consultazione dei dati via web per agevolare l'accesso ad un più completo ambito di dati geologici e relativi metadati, considerando anche il grande patrimonio storico di carte geologiche e geotematiche, che costituiscono la base per la conoscenza dei fenomeni territoriali.

Funzione sviluppata

La funzione implementata ha lo scopo di trovare i termini collegati ad un termine dato e di comporre una query che possa essere utilizzata dal motore di ricerca metadati per l'interrogazione del catalogo; tale query si interfaccia con il Thesaurus che è pubblicato come Link Data nel formato RDF/Skos di cui si presenta un esempio nella sezione seguente.

```
<rdf:Description
rdf:about="http://linkeddata.ge.imati.cnr.it:2020/resource/ThIST/acidoiodidrico">
<rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
<skos:prefLabel xml:lang="it">acido iodidrico</skos:prefLabel>
<skos:inScheme
rdf:resource="http://linkeddata.ge.imati.cnr.it:2020/resource/SkosConceptScheme/7"/>
</rdf:Description>
```

Il vocabolario viene considerato e pubblicato come un servizio di Web Service REST, dove si possono selezionare tramite una semplice interfaccia o nella stringa di chiamata alcuni parametri che affinano la ricerca nel thesaurus.

I parametri sono:

- **term** parametro obbligatorio. E' l'elemento da ricerca nel Thesaurus, ad esempi "Anidride Carbonica". Sono ammessi solo caratteri alfanumerici.
- **language** parametro opzionale. Indica il codice della lingua basato su due caratteri. Attualmente sono implementate le lingue "it" e "en". Il default è configurato nel file di properties (installazione base:"en").
- **outputType** parametro opzionale, di default è settato a "lucene" e indica il tipo di output da restituire.
- **flagCaseSensitive**: parametro opzionale , di default settato a 0. Può valere 0, cioè il termine può essere scritto indifferentemente maiuscolo o minuscolo, oppure 1.
- **flagWordMatch**: parametro opzionale , di default settato a 1. Può valere 0, nel qual caso viene fatto un match dell'intera stringa, oppure 1, nel qual caso anche un solo termine presente fa ritornare la sezione.

La risposta ha questa richiesta, sarà una stringa composta dal termine dato e dai termini collegati con un ranking diverso a seconda del tipo di legame.

Esempio di stringa di richiesta:

```
http://sgi.isprambiente.it/OnThist/servlet/onthist?term=Rossiya&language=it&outputType=xml
```

Esempio di risposta:

```
"Rossiya" "Russia" "eurasia"^0.5 "asia"^0.5
"sodruzhestvonezavisimykhgosudarstv"^0.5 "siberia"^0.7 "lagobaikal"^0.7
"moskva"^0.7 "kamchatka"^0.7 "mared%27azov"^0.3 "urali"^0.3
"caucaso"^0.3 "cccp1917-1992"^0.3
```

La ricerca mostrata nell'esempio a sua volta può essere potenziata concatenando i vocabolari e i servizi di ontologia permettendo così una ricerca ramificata più efficace e completa.

Nel portale sono inoltre presenti una serie di servizi WMS e WFS che sfruttano l'interoperabilità semantica interlinkando gli attributi a termini specifici di vocabolari geologi, anche in quest'ottica le attività in corso ora sono orientate alla messa in linea dei vocabolari geologici nazionali da affiancare ai già presenti internazionali per facilitare l'utente nella lettura e ricerca delle informazioni.