

Geodata4all, le azioni di AgID per rendere disponibili i dati territoriali a tutti gli utenti

Gabriele Ciasullo ^(a), Antonio Rotundo ^(a)

^(a) Agenzia per l'Italia Digitale, Via Liszt 21 – 00144 Roma, tel. 06852641, e-mail {ciasullo}{antonio.rotundo}@agid.gov.it

L'Agenzia per l'Italia Digitale (AgID) è impegnata ad implementare una strategia nazionale per una infrastruttura per l'informazione del settore pubblico come base di conoscenza di tutte le risorse informative delle pubbliche amministrazioni, al fine di rendere fruibili tali risorse ad un pubblico sempre più vasto e diversificato.

Le componenti principali di tale infrastruttura sono rappresentate dai due cataloghi nazionali di metadati, il Repertorio Nazionale dei Dati Territoriali (RNDDT)¹ e il portale dei dati aperti², strumenti che consentono agli utenti di conoscere la dimensione e le caratteristiche del patrimonio informativo pubblico relativamente al dominio di pertinenza (dati territoriali, il primo, e dati aperti, il secondo) ai fini dell'accesso e dell'utilizzo.

I due cataloghi sono basati su specifiche tecniche definite anch'esse da AgID³ nell'ambito del più ampio *framework* di interoperabilità europeo e con riferimento ai pertinenti standard internazionali⁴.

La sempre crescente disponibilità di dati e servizi da parte della pubblica amministrazione ha, infatti, richiesto l'adozione di schemi, descrizioni e classificazioni condivisi e interoperabili. Ciò sia per agevolare il compito delle stesse amministrazioni che devono rendere disponibili le informazioni su dati e servizi sia per gli utenti che devono essere facilitati nella ricerca e nell'interpretazione e l'analisi dei risultati di tale ricerca.

Sulla base delle regole e degli strumenti messi a disposizione da AgID, le amministrazioni provvedono (ormai da anni per il RNDDT e più recentemente per dati.gov.it) a documentare le risorse informative di cui sono titolari per renderle disponibili agli utenti interessati (cittadini, professionisti, imprese, altre PA, ecc.).

¹ <https://geodati.gov.it>

² <https://www.dati.gov.it/>

³ Per il RNDDT, il decreto della Presidenza del Consiglio dei Ministri del 10 novembre 2011 recante "Regole tecniche per la definizione del contenuto del Repertorio nazionale dei dati territoriali, nonché delle modalità di prima costituzione e di aggiornamento dello stesso" pubblicato nella Gazzetta Ufficiale n. 48 del 27/02/2012 - Supplemento ordinario n. 37, in corso di revisione.

Per dati.gov.it, il profilo applicativo nazionale DCAT-AP_IT (<https://www.dati.gov.it/content/dcat-ap-it-v10-profilo-italiano-dcat-ap-0>).

⁴ Per il RNDDT la Direttiva INSPIRE (<https://inspire.ec.europa.eu/>) e gli Standard ISO e OGC. Per dati.gov.it, il profilo applicativo europeo DCAT-AP, DCAT Application Profile for data portals in Europe (<https://joinup.ec.europa.eu/solution/dcat-application-profile-data-portals-europe>).

Per i dati territoriali, in particolare, al fine di favorire questo processo di conoscibilità e disponibilità dei dati, con l'implementazione e messa in produzione della nuova versione del RNDT sono state realizzate nuove funzionalità per facilitare e migliorare la ricerca, l'accesso e l'utilizzo dei dati, anche in funzione dell'obiettivo più generale di allargare la platea di utenti.

Appare evidente, infatti, che uno dei fattori che influenzano e guidano l'esperienza degli utenti è la competenza specifica e la condizione degli utenti stessi. In tale ottica, le diverse funzionalità implementate sono finalizzate a soddisfare le esigenze delle differenti tipologie di utenti.

Nel prosieguo di questo articolo, si fornisce una sintesi degli strumenti disponibili per la ricerca e l'accesso dei dati in funzione della tipologia di utenza, a partire da quella più "naturale" rappresentata dagli esperti del settore geo.

Utenti esperti del dominio geo

Per gli esperti di dominio geo il portale mette a disposizione le funzionalità di ricerca avanzata e i servizi di rete.

Attraverso la ricerca dettagliata⁵ è possibile impostare specifici criteri, che, in conformità di quanto disposto nel contesto INSPIRE, sono i seguenti: tipologia di risorsa (tra dataset, serie, servizi), testo libero, parola chiave, localizzazione geografica (tramite limite amministrativo, *bounding box* disegnato su mappa o coordinate), amministrazione responsabile, estensione temporale. Si può poi scegliere l'ordine di visualizzazione, l'eventuale anteprima dei risultati o direttamente i risultati stessi. Saranno quindi visualizzati i metadati che soddisfano la combinazione dei criteri impostati.

The screenshot shows the 'RICERCA DETTAGLIATA' (Detailed Search) interface of the RNDT portal. On the left, there's a sidebar with 'SELEZIONA COSA CERCARE' (Select what to search) and radio buttons for 'TUTTI' (All), 'DATI' (Data), and 'OSERVIZI' (Services). The main search area has three sections: 'RICERCA TESTO LIBERO' (Free text search) with a text input and a 'RICERCA TRA' (Search in) dropdown; 'RICERCA PAROLA CHIAVE' (Keyword search) with a text input and a 'RICERCA PAROLA CHIAVE...' label; and 'DOVE' (Where) with radio buttons for 'Dovunque' (Anywhere), 'Intersecanti' (Intersecting), and 'Completamente contenuti' (Completely contained). Below these is a note 'Seleziona un'area geografica, disegna un box o inserisci le coordinate' (Select a geographical area, draw a box or enter the coordinates). At the bottom are two buttons: 'ANTEPRIMA' (Preview) and 'VISUALIZZA RISULTATI' (View results). On the right, a map of the Mediterranean region is shown, with a green box indicating the search area.

Figura 1 – La ricerca dettagliata nel RNDT

È disponibile anche una vista predefinita basata su un filtro preimpostato, rappresentato dai *cluster* tematici⁶ e i temi INSPIRE⁷. Si può scegliere uno dei 9 *cluster* tematici (definiti nell'ambito di INSPIRE) per visualizzare i temi che lo

⁵ <https://geodati.gov.it/geoportale/ricerca-dettagliata>

⁶ I 9 cluster tematici sono: "Biodiversità e Aree sottoposte a gestione", "Copertura del suolo e Utilizzo del Territorio", "Dati di base topografici e catastali", "Elevazione, Orto immagini, Sistemi di Riferimento, Griglie geografiche", "Impianti e Servizi di pubblica utilità", "Mare e Atmosfera", "Monitoraggio e Osservazioni ambientali", "Scienze della Terra", "Statistica".

⁷ <https://inspire.ec.europa.eu/Themes/Data-Specifications/2892>

compongono e, quindi, scegliendo uno di tali temi vengono visualizzati tutti i metadati relativi ai dati che fanno riferimento al tema stesso.

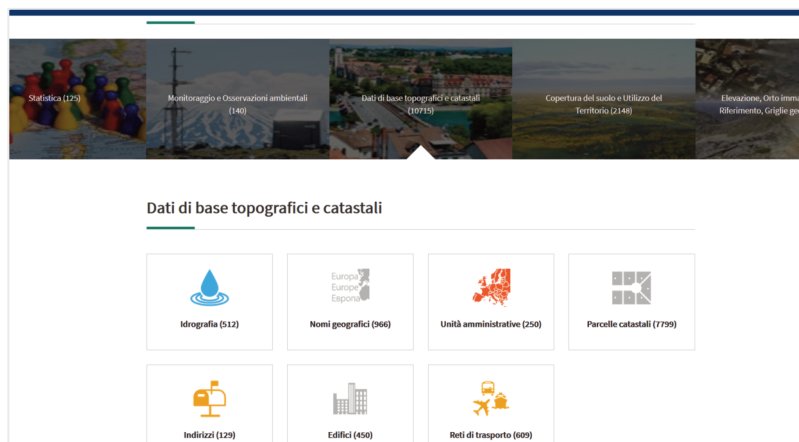


Figura 2 – La ricerca per cluster tematici e temi INSPIRE nel RNDT

A prescindere dal tipo di ricerca scelto, i risultati saranno visualizzati in un elenco che fornisce le informazioni minime su ciascuno: tipologia (dataset, serie o servizio), titolo, descrizione, tema INSPIRE di riferimento e informazioni di contatto del responsabile e del "punto di contatto".

Una volta individuati i metadati di interesse, si può visualizzare la scheda completa dei metadati o il file XML, accedere al sito eventualmente indicato o, in caso di servizio di rete (es. WMS o WFS), aprire il servizio nel visualizzatore o utilizzare l'URL del servizio o del documento di *capabilities* nel proprio ambiente di lavoro.

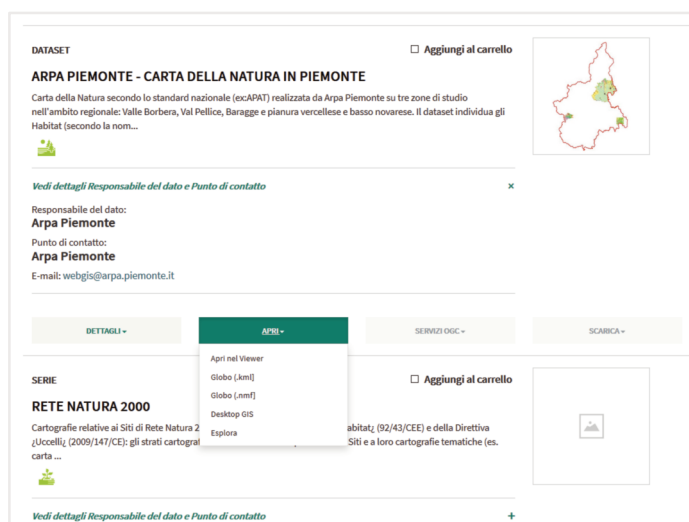


Figura 3 – I risultati della ricerca nel RNDT

Per quanto riguarda i servizi di rete, il RNDT fornisce il servizio di ricerca basato sullo standard CSW (*Catalogue Service for the Web*)⁸ definito dall'*Open Geospatial Consortium* (OGC), e coerente con le linee guida tecniche INSPIRE

⁸ <https://geodati.gov.it/geoportale/strumenti/2015-04-21-22-41-05>

sui *discovery services*⁹. Attraverso *client* esterni è possibile inviare richieste standard impostando specifici filtri.

Utenti sviluppatori

La ricerca dei dati nel RNDT è possibile anche tramite API REST. Per interrogare il Catalogo, l'URL da utilizzare è il seguente:

`http://geodati.gov.it/RNDT/rest/find/document?<parametro>&<parametro>&...`

dove *<parametro>* è uno dei parametri REST disponibili.

Il motore di ricerca è basato su Apache Lucene; per sfruttare al meglio la ricerca, può essere utile conoscere la sintassi delle *query* Lucene per le ricerche di testo. Nella pagina del sito RNDT dedicata¹⁰ sono indicati i parametri disponibili, gli elementi dei metadati interrogabili da utilizzare nelle ricerche nel RNDT e specifici esempi.

Le specifiche delle API di RNDT sono anche descritte secondo la specifica OpenAPI¹¹.

Per gli utenti che non hanno dimestichezza con la composizione delle richieste REST, il portale rende disponibile anche una funzione che, a partire dai risultati della ricerca dettagliata effettuata da interfaccia, restituisce l'URL della richiesta REST.

Utenti esperti del dominio open data

All'inizio dell'articolo si è detto dei due cataloghi: quello per i dati territoriali, RNDT, e quello per i dati aperti, dati.gov.it.

L'attività di documentazione nei due cataloghi, da parte delle amministrazioni, ha "naturalmente" prodotto un'area di sovrapposizione dei due ambiti di applicazione (dati territoriali e dati aperti), rappresentata appunto dai dati territoriali di tipo aperto: ciò ha comportato, ad esempio, una duplicazione delle informazioni descrittive nei due cataloghi di riferimento, con il rischio che esse non siano allineate e/o aggiornate, e in alcuni casi siano addirittura discordanti, compromettendo l'attendibilità delle informazioni stesse.

Per superare le criticità evidenziate e realizzare, quindi, l'integrazione dei due cataloghi, AgID ha definito una serie di regole organizzative e di indicazioni tecniche¹², basate sulla specifica GeoDCAT-AP. Tali regole, evitando alle amministrazioni un aggravio di adempimenti, mirano a garantire a tutti gli utenti l'accesso alle informazioni, qualunque sia la fonte di ricerca.

GeoDCAT-AP¹³, lo ricordiamo, è la specifica europea implementata proprio al fine di consentire la trasformazione dei metadati dei dati geografici nello standard utilizzato nei portali "generalisti" basato su DCAT-AP, per consentirne la ricerca, l'*harvesting* e il riuso.

⁹ <https://inspire.ec.europa.eu/documents/technical-guidance-implementation-inspire-discovery-services-0>

¹⁰ <https://geodati.gov.it/geoportale/strumenti/api-rest>

¹¹ <https://developers.italia.it/it/api/geodati>

¹² <https://geodati.gov.it/geoportale/documenti/12-documenti/277-linee-guida-nazionali-geodcat-ap>

¹³ <https://joinup.ec.europa.eu/solution/geodcat-application-profile-data-portals-europe>

Sulla base delle linee guida sull'implementazione di GeoDCAT-AP, è stata rilasciata l'API GeoDCAT-AP del RNDT¹⁴ che consente di trasformare i metadati dei dati documentati nel Repertorio, secondo il profilo italiano, dallo standard ISO TS 19139 allo standard DCAT-AP/DCAT-AP_IT (estensione italiana di DCAT-AP) utilizzato per i dati aperti.

L'API, che è basata ed estende quella sviluppata da JRC¹⁵ nell'ambito del programma europeo ISA/ISA², accetta sia richieste CSW (GET e POST) che richieste REST e restituisce i metadati in formato RDF/XML o JSON-LD.

Essa può essere utilizzata dagli utenti per avere una diversa struttura dei metadati, quella utilizzata per i dati aperti, ma è utilizzata anche dal portale dei dati aperti dati.gov.it per fare la raccolta (*harvesting*) dei dati territoriali aperti dal RNDT per una pubblicazione alternativa dei metadati.

Figura 4 – L'API GeoDCAT-AP del RNDT

Utenti non esperti

La ricerca, l'accesso e l'utilizzo dei dati territoriali possono risultare difficili per gli utenti non esperti. Ciò è dovuto soprattutto alle peculiarità di questa tipologia di dati e alle complessità proprie di questo dominio. Si pensi, per esempio, alla comprensione delle coordinate e dei sistemi di riferimento o ai formati che non sono facili da elaborare da parte di non esperti.

Per questo anche gli strumenti di ricerca devono essere adeguatamente commisurati affinché i dati territoriali resi disponibili dalle amministrazioni siano in qualche modo accessibili e utilizzabili anche da questa categoria di utenti.

A tale proposito, è interessante, per l'economia dell'articolo, il confronto tra i geoportali e i motori di ricerca riportato nella tabella¹⁶ che segue.

¹⁴ https://geodati.gov.it/geodcat-ap_it/

¹⁵ <https://github.com/SEMICEu/iso-19139-to-dcat-ap/tree/master/api>

¹⁶ Tratta e adattata dalla presentazione "Introduction to Google Dataset Search workshop" di Marcin Grudziński nell'ambito della Conferenza INSPIRE 2018 (<https://webgate.ec.europa.eu/fpfis/wikis/display/InspireMIG/Google+Dataset+Search+workshop+2018-09-19>)

Criteri	Geoportali	Motori di ricerca (Google)
Difficoltà	Difficile	Facile
Utenti	Soprattutto esperti di dominio	Tutti
Risultati della ricerca (da una prospettiva dell'utente e del fornitore di dati INSPIRE)	Buono	Scarso - Medio

La ricerca più immediata dal portale RNDT è attraverso il *box* in *homepage* in cui è possibile inserire una parola per effettuare una ricerca in tutti i campi di tipo testo dei metadati.

Per aumentarne la visibilità, da qualche mese, tali dati sono disponibili anche attraverso *Google Dataset Search* (<https://toolbox.google.com/datasetsearch/search?query=site:geodati.gov.it>).

Dataset Search consente agli utenti di trovare set di dati archiviati nel Web tramite una semplice ricerca per parola chiave. Lo strumento trova le informazioni relative ai set di dati ospitati in migliaia di repository sul Web e rende questi set di dati accessibili e utili a livello globale.

Per consentire a Google il rilevamento dei set di dati del RNDT, è stato utilizzato, in ciascuna *landing page* dei metadati il *markup* Dataset del vocabolario *schema.org*¹⁷, secondo le linee guida sui dati strutturati di Google stesso. Inoltre, per consentire a Google di trovare gli URL delle pagine dei metadati, sono stati utilizzati i file Sitemap.

Oltre a fornire i principali metadati del dataset, Dataset Search visualizza anche una mappa con l'inquadratura geografica (*bounding box*) del dataset stesso e, se disponibili, fornisce i *link* agli articoli accademici che citano il set di dati.

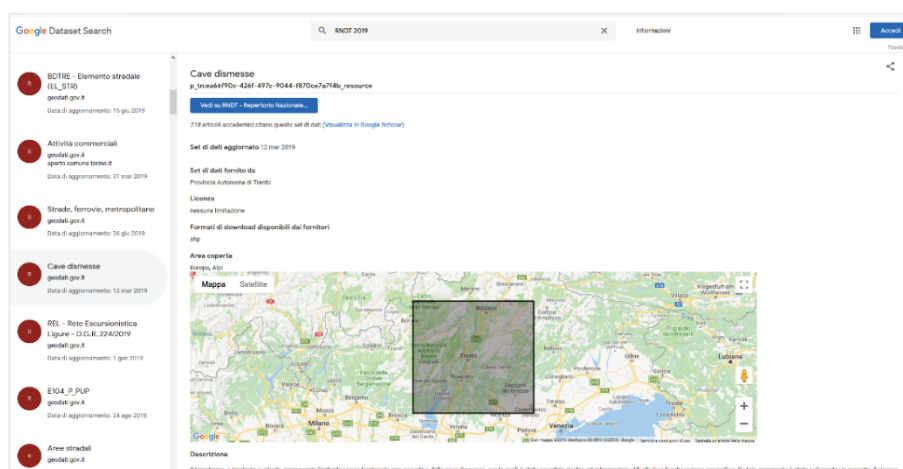


Figura 5 – I dati del RNDT nel Google Dataset Search

¹⁷ <https://schema.org/>

AgID ha lavorato anche per rendere identificabile ogni metadato presente nel catalogo attraverso un URI (*Uniform Resource Identifier*), stringa che identifica univocamente una risorsa in rete, con la seguente struttura

<https://geodati.gov.it/resource/id/{fileIdentifier}>

dove {fileIdentifier} è l'identificativo del record di metadati indicato nel record stesso.

Un esempio è il seguente:

https://geodati.gov.it/resource/id/p_tn:ea66f90c-426f-497c-9044-f870ce7a7f4b

che è l'URI dei metadati del dataset "Cave dismesse" della Provincia Autonoma di Trento (v. figura 5).

L'attività implementata è ricompresa all'interno delle buone pratiche pubblicate nel documento OGC e W3C "*Spatial Data on the Web Best Practices*"¹⁸ (Best practice 1, "*Use globally unique persistent HTTP URIs for Spatial Things*"¹⁹ e Best practice 2, "*Make your spatial data indexable by search engines*"²⁰) e riconosciuta come cruciale per l'implementazione di infrastrutture di dati territoriali nell'ambito dell'azione ELISE²¹ della Commissione Europea.

Tale azione è stata altresì valorizzata recentemente anche nell'ambito di INSPIRE, con il workshop "*Making spatial data discoverable through mainstream search engines*" del 3 e 4 luglio 2019²² nel corso del quale AgID ha presentato le proprie iniziative inerenti il tema.

¹⁸ <https://www.w3.org/TR/sdw-bp/>

¹⁹ <https://www.w3.org/TR/sdw-bp/#globally-unique-ids>

²⁰ <https://www.w3.org/TR/sdw-bp/#bp-indexable>

²¹ <https://joinup.ec.europa.eu/collection/elise-european-location-interoperability-solutions-e-government/about>

²² <https://webgate.ec.europa.eu/fpfis/wikis/display/InspireMIG/Workshop+Programme>

