

## Open DATA e INSPIRE

Gianni Campanile

Esri Italia s.p.a, Via Tiburtina 755 ROMA, +39 06 40696.1, gcampanile@esriitalia.it

### Riassunto

In questo articolo viene mostrato come il modello di metadati/dati/servizi proposto da INSPIRE risponde alle caratteristiche tipiche dei modelli Open Data. In particolare viene ripresa la ben nota classificazione proposta da Tim Berners Lee e confrontata con un'architettura INSPIRE completa di servizi di download.

Viene infine mostrato un caso pratico di implementazione che, utilizzando il software ArcGIS for INSPIRE, rende disponibili i propri Open Data attraverso una SDI compatibile con INSPIRE e quindi di rispondere ai requisiti richiesti dalla direttiva fino al 2017.

### Abstract

In this paper we will discuss the INSPIRE metadata/data/service model and show that this model fits well with Open Data pattern. Starting from the well known Tim Berners Lee star classification, we will make a comparison with an INSPIRE download model showing that data exposed through this model correspond to the highest classification.

We will also present a real use case based on Esri's ArcGIS for INSPIRE software that allows to expose available Open Data through an INSPIRE compatible SDI so to fulfill the directive's requirements up to 2017.

### Introduzione

L'esigenza di rendere disponibili i propri dati è sempre più sentita soprattutto da parte delle pubbliche amministrazioni sia per l'utilizzo diretto dei dati sia per creare grazie ad essi del valore aggiunto che possa poi essere fruito da tutti i cittadini.

In effetti molti siti già da tempo hanno sezioni dedicate agli Open Data, fornendo diverse modalità e licenze con cui scaricare i dati di interesse, spesso gratuitamente a volte attraverso la sola registrazione, a volte a pagamento.

Questo articolo parte dal presupposto che ci siano dei dati scaricabili, non entrando quindi nel merito delle modalità di accesso/uso né nel formato finale in cui i dati sono forniti (ad.es. Shapefile, CSV, etc.), facendo solo conto che questo sia un formato aperto in quanto mirato all'utilizzo comune.

L'obiettivo del lavoro è infatti di mostrare come gli stessi dati possano essere offerti secondo lo standard INSPIRE e come questo modello corrisponda alle 5 stelle della classificazione definita da Tim Berners Lee sugli Open Data.

### Linked Data

Secondo Wikipedia *Linked Data* □□.□ un termine utilizzato per descrivere una best practice raccomandata per esporre, condividere, e connettere porzioni di dati, informazioni e conoscenza sul Web Semantico utilizzando URI e RDF □ La definizione rimanda ad altri termini che è bene descrivere brevemente:

- *Semantic Web* è l'estensione del World Wide Web che si ottiene introducendo collegamenti tra i significati degli oggetti, espressi graficamente attraverso grafi orientati e archi che definiscono le relazioni fra gli oggetti.
- *URI (Uniform Resource Identifier)* è una stringa per identificare univocamente una risorsa
- *RDF (Resource Description Framework)* è una specifica del W3C (World Wide Web Consortium) che consiste in un modello concettuale (simile al modello Entità-Relazione o di Classi)

Parlando quindi di un sito o portale web che espone dei *LOD (Linked Open Data)*, si vuole indicare un sistema web che offre una modalità di navigazione e documentazione che ci permette di ricercare, trovare e ottenere i dati di nostro interesse.

Data la varietà di formati e metodi di accesso, non tutti i sistemi che offrono LOD sono uguali: per misurarne la "qualità", Tim Berners Lee ha proposto una classificazione basata su stelle illustrata nella seguente figura 1:

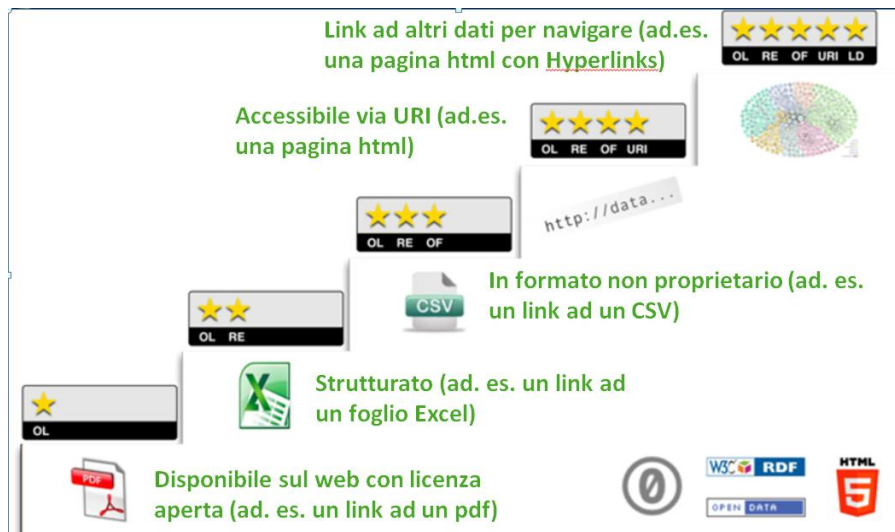


Figura 1 □ Tabella di classificazione degli Open Data.

Come si vede dalla figura, l'ultimo livello classificato con 5 stelle permette di avere dei dati accessibili tramite classici indirizzi http, e che contengono link ad altri dati collegati. E' ovvio che questo costituisce il livello ottimale per poter ottenere le informazioni di interesse.

E' possibile però notare che, anche al livello più alto, non vi sono specifiche su come *trovare* i dati di interesse, né su come ottenere una *descrizione* di come siano strutturati questi dati. In seguito vedremo come il modello INSPIRE fornisce anche le soluzioni a questi problemi.

### Architettura INSPIRE

Come è noto, l'architettura definita dalla direttiva INSPIRE definisce delle regole principalmente per tre componenti: metadati, dati e servizi. Sono infatti definite delle regole di compilazione dei metadati (profilo INSPIRE per serie, dataset, e servizi), delle specifiche per i dati e delle specifiche per i servizi.

Questo approccio però non implica che l'accesso ai dati e ai metadati possa essere effettuato *direttamente* perché INSPIRE stabilisce chiaramente che qualunque accesso a questi elementi è effettuato tramite i servizi definiti: il servizio di Discovery per i metadati, i servizi di View e di Download per i dati.

L'architettura è infatti quella visualizzata in figura 2:

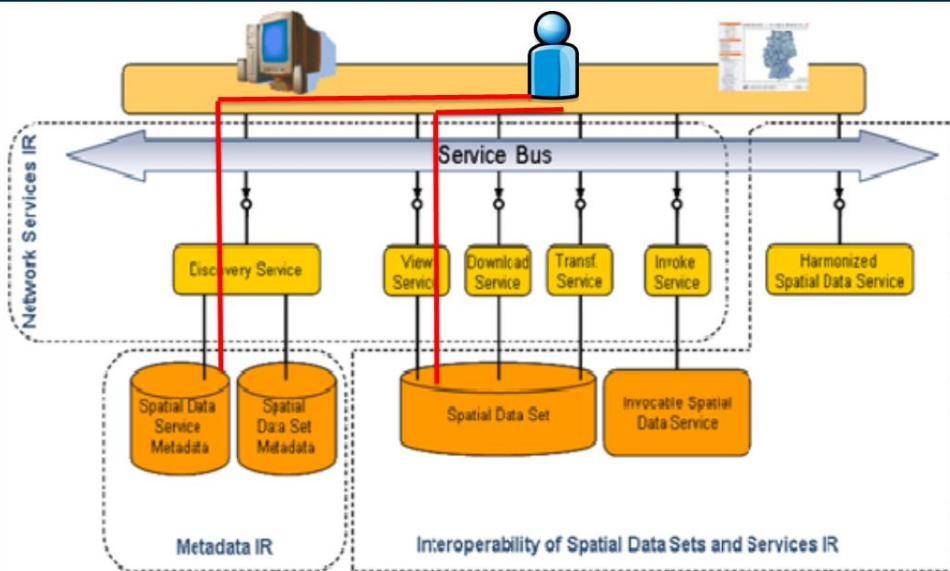


Figura 2 □ Architettura INSPIRE.

Come si vede dai collegamenti in rosso, l'accesso a metadati e dati è attraverso il *Service Bus*, cioè lo strato che espone i servizi. Questo implica che l'accesso a dei dati scaricabili, per essere compatibile con INSPIRE non può essere un accesso diretto ma deve utilizzare le specifiche del servizio di Download INSPIRE.

Tali specifiche definiscono due modalità operative, una minimale basata su *ATOM* ed un'altra basata sul protocollo *WFS* definito da *OGC*. La modalità minimale viene applicata di preferenza quando esistono già dei dataset scaricabili, e infatti viene denominata *pre-defined dataset download service*. Nel documento *Technical Guidance Download Services*, viene così descritta:

- Pre-defined dataset download service(s);  
*A pre-defined dataset download service provides for the simple download of pre-defined datasets (or pre-defined parts of a dataset) with no ability to query datasets or select user-defined subsets of datasets. A pre-defined dataset or a pre-defined part of a dataset could be (for example) a file stored in a dataset repository, which can be downloaded as a complete unity with no possibility to change content, whether encoding, the CRS of the coordinates, etc.*

Come si vede dal report di INSPIRE ([http://inspire-geoportal.ec.europa.eu/resources/INSPIREResourcesReports/resourcesReport\\_2014-02-18](http://inspire-geoportal.ec.europa.eu/resources/INSPIREResourcesReports/resourcesReport_2014-02-18)), la maggior parte di servizi di download in Europa di basa su *ATOM*:

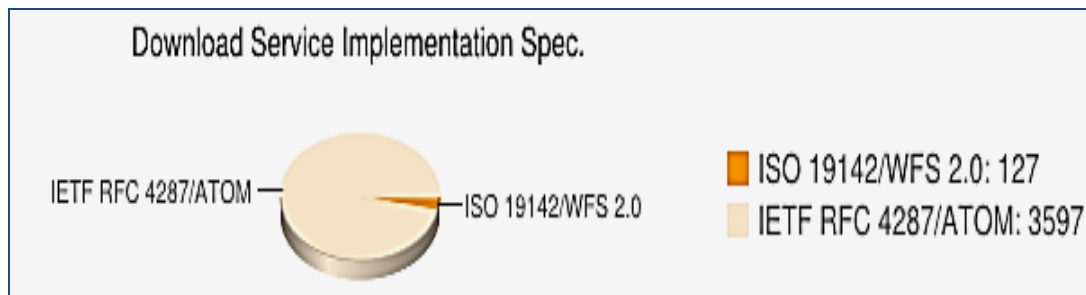


Figura 3 □ Report INSPIRE sui servizi di Download.

Tale specifica è quindi di elevata diffusione e si adatta perfettamente al nostro caso; nei capitoli seguenti vedremo come è possibile implementare tale servizio basandoci su un caso reale di uso.

### Implementazione

Il caso d'uso che verrà descritto è quello dell'implementazione del servizio di download INSPIRE presso l'ISPRA. Va notato comunque che questo è comunque un caso d'uso piuttosto generale che può essere applicato ad altri ambienti.

Nella figura seguente vediamo la pagina di download del servizio SINA di ISPRA:

**SINA**  
Rete del Sistema Informativo Nazionale Ambientale

IT

Tu sei qui: [Home](#) > [Sistemi informativi ambientali ISPRA](#) > Download MAIS

Home	Download MAIS
Rete SINAnet	<ul style="list-style-type: none"> <li>Corine Land Cover</li> <li>Complessi idrogeologici</li> </ul>
INSPIRE	<ul style="list-style-type: none"> <li>Reticolo idrografico (1:250.000)</li> <li>Bacini idrografici principali (1:250.000)</li> <li>Bacini idrografici secondari (1:250.000)</li> </ul>
Sistemi informativi ambientali ISPRA	<ul style="list-style-type: none"> <li>Impianti di monitoraggio ambientale</li> </ul>
Progetti	<ul style="list-style-type: none"> <li>Cartografia Gravimetrica Digitale</li> <li>DEM20</li> <li>DEM75</li> </ul>
Punto Focale Nazionale EIONET	<ul style="list-style-type: none"> <li>Idrografia INSPIRE</li> <li>Consumo di suolo nazionale e regionale</li> </ul>
Gruppiware	

Figura 4 □ Pagina di download del servizio SINA.

Come si vede è una pagina generica, nella quale sono presenti diversi dataset scaricabili liberamente. E' importante notare che questa pagina esiste già da diverso tempo, indipendentemente da INSPIRE ed è raggiungibile dal sito di ISPRA semplicemente navigando su opportuni link.

Ma rispetto alla classificazione degli Open Data definita in precedenza, i dataset scaricabili da questa pagina come sono classificati? Sicuramente sono disponibili su Web (1 stella), e, in quanto dataset geografici, sono strutturati (2 stelle); inoltre, come la maggior parte dei dataset geografici resi pubblici, sono in un formato non proprietario, in questo caso Shapefile (3 stelle). Per quanto riguarda l'accessibilità, anche se i singoli dataset sono accessibili tramite un URI, questo URI non

è individuabile se non quando si arriva sulla pagina vera e propria. Quindi la classificazione corretta è di 3 stelle.

A partire da questa pagina potremo costruire un'architettura compatibile con INSPIRE per i Download e View services, illustrata in figura 5:

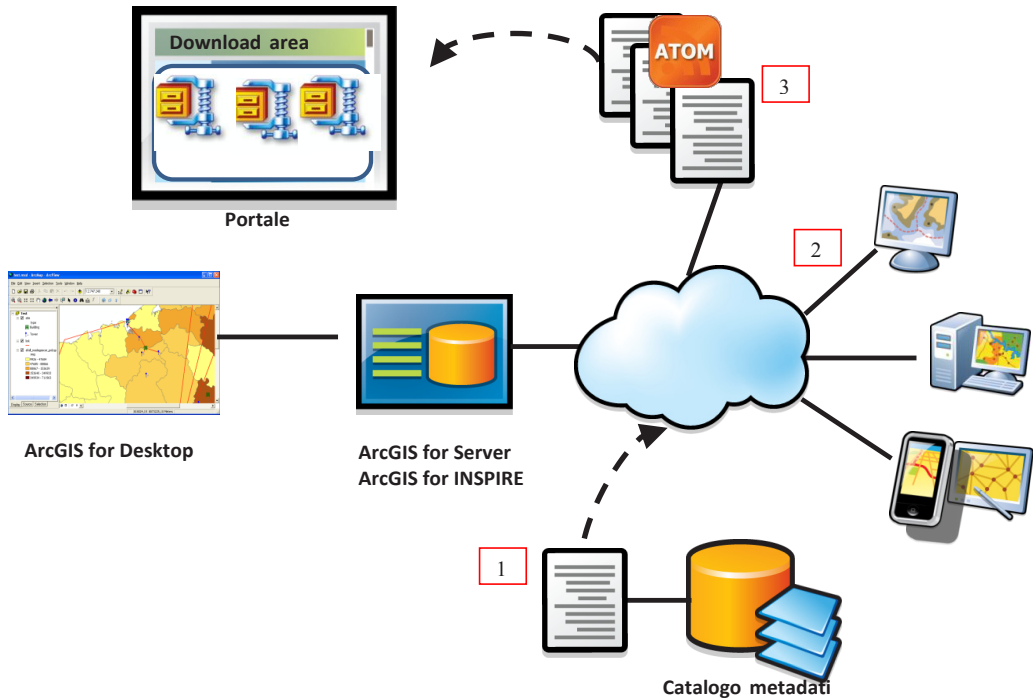


Figura 5 □ Architettura del servizio di Download basato su ATOM.

Come si vede dalla figura la pagina del portale rimane nella sua forma originale, ma vengono creati i seguenti elementi:

1. Metadati per il dataset, il servizio di View e il servizio di Download. I metadati dovranno essere opportunamente collegati
2. Servizio di INSPIRE View
3. Servizio di INSPIRE Download basato su ATOM (Pre-defined Dataset Download)

I metadati e un apposito catalogo che esponga un'interfaccia basata sul protocollo CSW di OGC permettono di soddisfare i requisiti del servizio di INSPIRE Discovery. Nel nostro caso questo servizio è fornito dal Esri Geoportal Server, un prodotto FOSS che Esri Italia ha modificato in modo da renderlo compatibile con le specifiche dei metadati nazionali (RNDT). Per i metadati è necessario documentare sia il dataset scaricabile, che il servizio di View e quello di download. Quest'ultimo avrà un riferimento all'endpoint del servizio vero e proprio.

Il servizio di INSPIRE View si crea pubblicando su un server i dataset e creando un servizio che risponda alle specifiche richieste. Da notare che per adesso non è necessario che i dati siano armonizzati, cioè che siano compatibile con le specifiche dati. Per la creazione del servizio di INSPIRE View, è stato utilizzato il client GIS ArcGIS for Desktop, il server GIS ArcGIS for Server con l'estensione ArcGIS for INSPIRE.

Per il servizio di Download è necessario definire dei file XML basati sul protocollo ATOM; questo definisce una serie di informazioni quali il nome e la descrizione del dataset, le modalità di accesso e uso, il punto di contatto etc. Così come per il servizio di View, per la creazione del servizio di INSPIRE View, è stato utilizzato il client GIS ArcGIS for Desktop, il server GIS ArcGIS for Server con l'estensione ArcGIS for INSPIRE

La creazione di questo servizio permette di collegare il suo indirizzo nei metadati relativi, in modo che tramite una ricerca è possibile reperire le necessarie informazioni ed il link per il download, come illustrato in figura 6 e figura 7:

The screenshot shows the ISPRAT portal search results for the service 'Idrografia - servizio (INSPIRE Download)'. The search results are displayed in a yellow box with the following text: 'Servizio di Download secondo la direttiva INSPIRE. Permette di accedere e scaricare la banca dati relativa al reticolo idrografico nazionale, i laghi, i nodi del reticolo ed i bacini. Scala 1:250.000.' Below this, there are several links: 'Apri nel Viewer', 'Anteprima', 'Metadati', 'Vedi Sito Web', 'Metadati-XML', 'Scarica', and 'Zoom a'. A red dashed circle highlights the 'Metadati-XML Scarica Zoom a' link. The interface includes navigation tabs like 'CSW', 'CERCA', 'SFOGLIA IL CATALOGO ISPRAT', and 'VISUALIZZATORE'. A sidebar on the left shows search filters and options.

Figura 6 □ Ricerca del metadato del servizio di Download.

The screenshot shows the 'Geodatabase del Reticolo Idrografico Nazionale' service page. The page title is 'Geodatabase del Reticolo Idrografico Nazionale strutturato secondo le Data Specifications di INSPIRE'. The page content includes a description of the dataset and a link to 'Reticolo Idrografico Nazionale - Dataset'. The page also features a feed subscription section at the top.

Figura 7 □ Servizio di Download basato su ATOM.

Quello che è stato fatto quindi è stato di rendere accessibile la pagina di download già esistente attraverso dei metadati e dei servizi compatibili con le specifiche INSPIRE. Ora possiamo riprendere l'analisi sulla classificazione di questi dataset e notare che abbiamo reso possibile accedere tramite un URI (esposto tramite il protocollo ATOM) e abbiamo collegato il servizio con i dataset su cui opera e questi con il servizio di View. In pratica abbiamo raggiunto le 5 stelle di Tim Berners Lee. In aggiunta, abbiamo la possibilità di ricercare questi dati tramite il catalogo di metadati, quindi abbiamo migliorato ulteriormente la raggiungibilità del dato.

### **Conclusioni**

L'implementazione di un'architettura INSPIRE permette di incapsulare facilmente una pagina di download preesistente creando sia un'architettura OpenData di alto livello sia un'architettura INSPIRE.

L'uso di tecnologia adeguata, quale ad esempio ArcGIS for INSPIRE, permette di creare automaticamente i servizi necessari e quindi di rispondere ai requisiti della direttiva INSPIRE fino alla prossima scadenza del 2017.