

## L'organizzazione di metadati e dati relativi a piattaforme aeree e satellitari per il telerilevamento

Paolo PLINI, Rosamaria SALVATORI, Sabina DI FRANCO, Valentina DE SANTIS

Consiglio Nazionale delle Ricerche - Istituto sull'Inquinamento Atmosferico,  
via Salaria km 29,300, tel. 0690672975/451, fax 0690672660, e-mail {plini; salvatori; difranco; vds}@iia.cnr.it

### Riassunto

Il progetto si propone di sviluppare un sistema terminologico, controllato e strutturato, per la gestione e la consultazione di dati e metadati relativi alle caratteristiche delle piattaforme aeree e satellitari per il telerilevamento.

Questa iniziativa si inserisce nel più ampio progetto di organizzazione della conoscenza nel settore ambientale volto alla realizzazione di un thesaurus di riferimento per l'ambiente (EARTH) e di un microthesaurus specifico per il GIS e telerilevamento, entrambi gestiti da EKOLab.

### Abstract

The project aims to develop a controlled, structured terminological system for the handling of data and metadata dealing with airborne and satellite remote sensing platforms.

These activities are part of a larger effort on environmental knowledge organisation through the development of a reference thesaurus for the environment (EARTH) and a microthesaurus on remote sensing and GIS.

### Introduzione

Gli archivi elettronici costituiscono importanti supporti operativi per la realizzazioni di molteplici attività, sia in ambito scientifico che istituzionale. Queste raccolte di dati costituiscono dei depositi potenziali di informazione, ma necessitano di alcuni accorgimenti, non solo tecnologici, per poter essere utilizzate al meglio come risorse effettive di informazione.

La quantità, l'eterogeneità e la complessità dei dati, la dispersione su fonti multiple, la velocità di diffusione e di variazione delle informazioni, rendono il processo di recupero delle informazioni tutt'altro che agevole.

Uno degli aspetti cruciali da risolvere, inoltre, è quello linguistico-concettuale. La descrizione del contenuto delle banche dati avviene, principalmente, mediante l'utilizzo di parole-chiave, spesso insufficienti a garantire un'efficiente gestione dell'informazione: la presenza di ambiguità lessicali, gli equivoci che si possono generare esprimendo in diversi modi lo stesso concetto e il cosiddetto *information noise*, costituiscono ostacoli al raggiungimento di livelli adeguati di precisione ed esaustività.

Per superare tale criticità, spesso, le banche dati si avvalgono di strumenti terminologici, come i thesauri, che tramite un lessico controllato, strutturato e finalizzato all'organizzazione dei concetti chiave e al successivo recupero dell'informazione, razionalizzano il lavoro di documentazione.

Uno degli aspetti più importanti dell'utilizzo di questi strumenti è la possibilità, per gli utenti, di adottare un vocabolario comune di riferimento che imponga una corrispondenza univoca tra oggetto o fenomeno descritto e termine utilizzato per descriverlo. Ciò consente, inoltre, di risolvere problematiche di interoperabilità tra sistemi.

Il settore del telerilevamento non è esente da questi problemi: l'eterogeneità delle informazioni (per forma e per contenuto), la vastità dell'ambito disciplinare che richiede l'utilizzo di terminologia

specialistica, l'ampiezza del bacino di utenza, rendono necessaria l'adozione di strumenti per l'organizzazione della conoscenza che superino i limiti delle tradizionali liste di parole chiave. Le informazioni inerenti alle caratteristiche tecniche delle piattaforme aeree o satellitari per il telerilevamento sono rivolte sia ad un pubblico specialistico, quale quello della comunità scientifica, sia ad un pubblico di professionisti che utilizzano le immagini telelerivate prevalentemente come supporto cartografico. Ogni descrizione degli elementi strutturali, tecnologici e di funzionamento viene solitamente realizzata mediante un testo, un elenco di parole chiave corrispondenti alle caratteristiche dell'apparecchiatura in oggetto, oppure tramite una o più tabelle riepilogative. Ad ogni caratteristica descritta sono, inoltre, associati i relativi valori e attributi numerici. Se le tabelle, pur mancando di informazioni accessorie, forniscono una buona sintesi delle informazioni inerenti le caratteristiche tecniche delle piattaforme aeree o satellitari, i testi, redatti da persone diverse, presentano generalmente le informazioni utilizzando un linguaggio non organizzato, tanto meno armonizzato e strutturato.

La necessità di condivisione dei dati e le problematiche di interoperabilità a livello semantico evidenziano l'utilità di realizzare un sistema che associ ad ogni piattaforma aerea o satellitare un *set* di informazioni, che tramite un linguaggio comune e condiviso migliori la comprensibilità semantica dei dati e dei metadati.

### Metodologia utilizzata

Per la realizzazione del sistema come primo passo è stato redatto un elenco delle piattaforme aeree e satellitari. Per ogni piattaforma, sono state poi identificate le caratteristiche essenziali (tipologia del sensore, risoluzione, numero di bande, ecc.). Si è poi proceduto alla raccolta, selezione e armonizzazione delle parole chiave.

I termini sono stati classificati su base semantica; ciascun termine è stato collocato in un albero di relazioni gerarchiche, a partire dal concetto più generale di riferimento fino a quelli più specifici. (fig. 1)

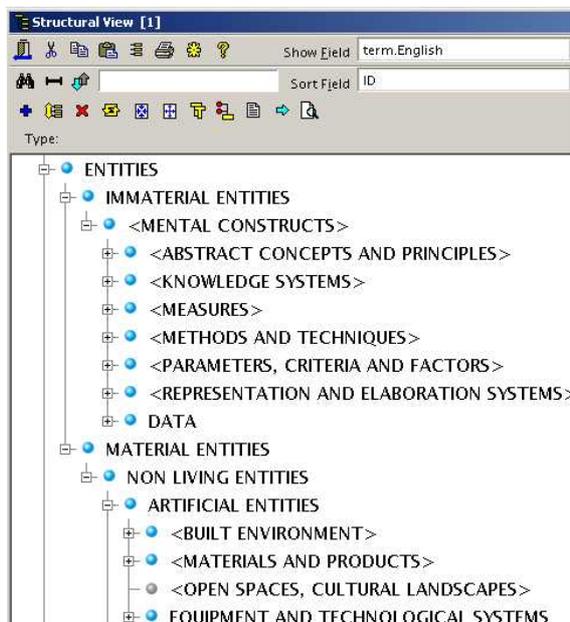


Figura 1 – Lo schema di classificazione e le macrocategorie

Sono state poi aggiunte le informazioni relative alle caratteristiche delle diverse piattaforme e sono state raggruppate in “etichette di snodo”, sviluppate ad hoc. Le etichette di snodo costituiscono i “serbatoi” dove raccogliere termini omogenei e/o coerenti per tipologia. Ad es.: <satellites by

name> è l'etichetta di snodo che raccoglie i nomi dei vari satelliti; <radiometric resolution values> è l'etichetta di snodo dei valori di risoluzione raccolti.

In seguito, è stata realizzata una rete di associazioni tra termini dell'albero gerarchico, creando un insieme di sottotipi della tradizionale relazione RT (related term), normalmente utilizzata nei thesauri per collegare termini distanti tra loro nell'albero ma concettualmente connessi.

Ad esempio, "TK350" e "10 x 10 m", rispettivamente classificati nell'albero come "entità materiali artificiali – apparecchiature" e "entità immateriali – dati – valori – valori di risoluzione spaziale", sono stati collegati secondo una relazione associativa reciproca ma asimmetrica, di tipologia "sensore - risoluzione spaziale" (fig. 2).

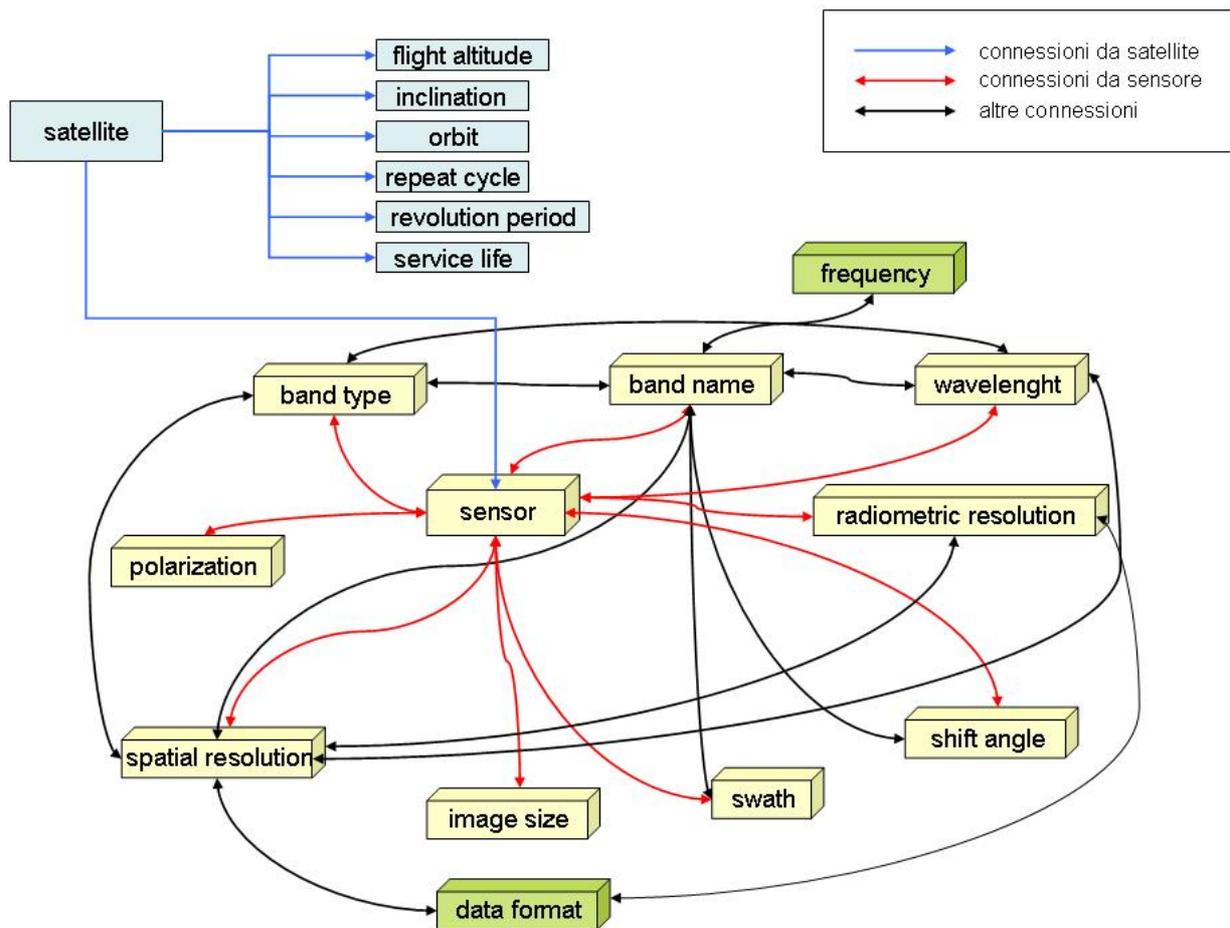


Figura 2 – Dettaglio sulla rete semantica con evidenziate le connessioni tra i vari elementi utilizzati per lo sviluppo del sistema

Il database terminologico è stato strutturato a matrice, organizzato su base concettuale e gerarchica prendendo in considerazione parametri e caratteristiche di ciascuna piattaforma. Al tempo stesso è stata implementata la possibilità di una lettura trasversale dei parametri e delle caratteristiche, riaggregabili per ciascuna delle piattaforme. La consultazione permette di accedere al sistema mediante vari punti di accesso. Partendo dai satelliti, il sistema visualizza, per ogni voce, le principali caratteristiche (altitudine, orbita, inclinazione, sensori montati, periodo di operatività, ecc., cfr. fig. 3).

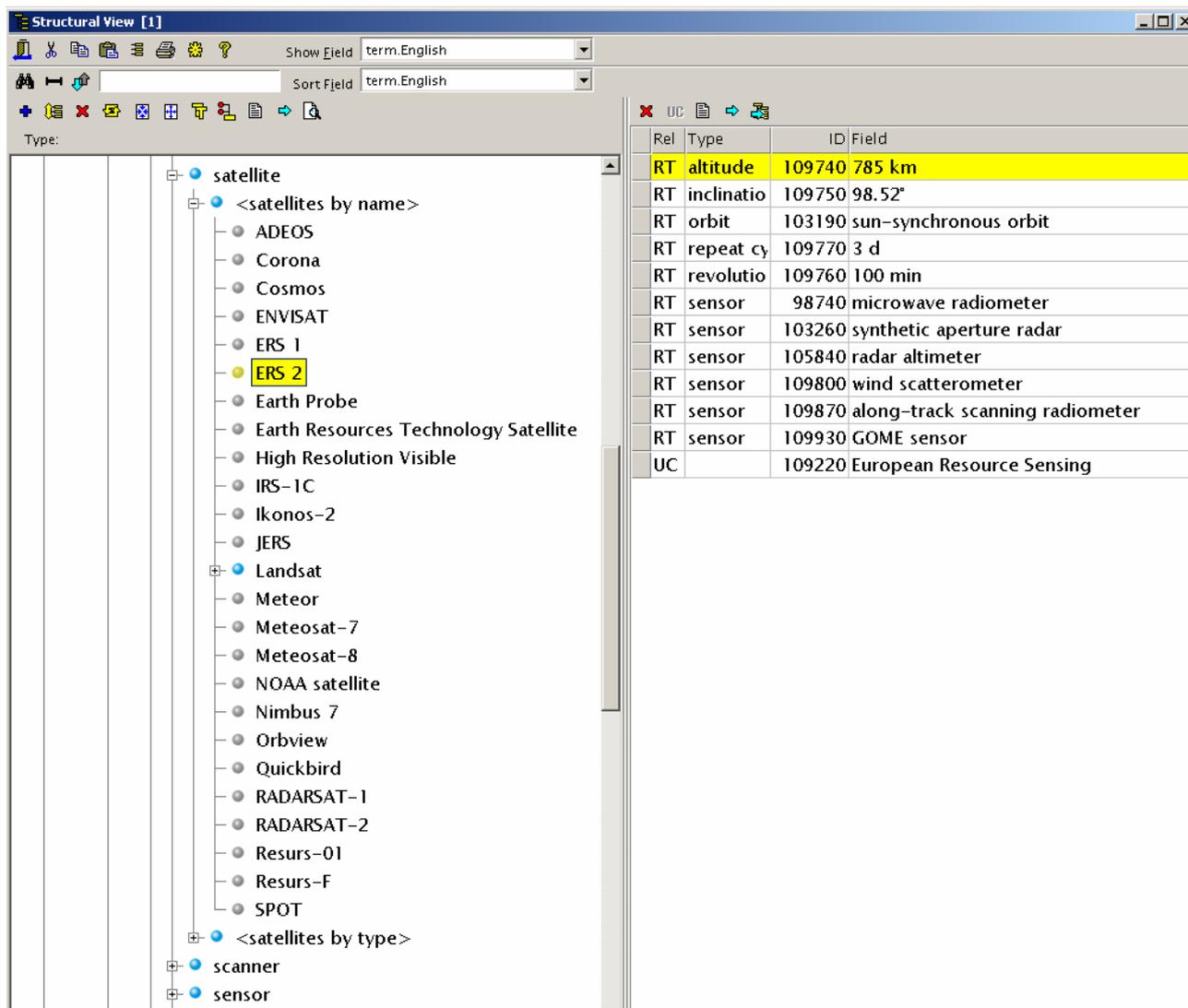


Figura 3 – Dettaglio sulla visualizzazione delle caratteristiche di ogni satellite

Selezionando un sensore e aggiornando la visualizzazione, si può proseguire la navigazione nel sistema evidenziando, per ogni sensore, le caratteristiche strutturali. E' quindi possibile ottenere informazioni su bande, lunghezza d'onda, frequenze, risoluzione spaziale e radiometrica, ecc. (fig. 4) Naturalmente per ogni sensore viene anche segnalato il satellite su cui il sensore è montato. La navigazione è quindi "reversibile", cioè bidirezionale: il sistema realizzato risulta essere particolarmente flessibile e, allo stesso tempo, assicura la coerenza interna del dato fornito. E' importante notare come le informazioni vengano gestite e presentate secondo un formato unico e tramite un linguaggio "normalizzato": all'interno del sistema i dati sono confrontabili tra di loro.

The screenshot shows the 'Structural View [1]' window. On the left, a tree view lists sensor types under 'sensor', with 'LISS-III sensor' highlighted. On the right, a table displays related terms (RT) for the selected sensor. The table has columns for 'Rel', 'Type', 'ID', and 'Field'.

Rel	Type	ID	Field
RT	band name	110230 8	(LISS-III)
RT	band name	110510 2	(LISS-III)
RT	band name	110520 3	(LISS-III)
RT	band name	110560 4	(LISS-III)
RT	satellite	97070	IRS-1C
RT	spat_resolution	110170	23 x 23 m
RT	spat_resolution	110180	70 x 70 m
RT	swath	110970	141 km
RT	swath	110980	148 km
RT	wavelength	110190	1.55 – 1.70 μm
RT	wavelength	110200	0.77 – 0.86 μm
RT	wavelength	110210	0.62 – 0.68 μm
RT	wavelength	110220	0.52 – 0.59 μm
RT		110160	optical sensor

Figura 4 – Le caratteristiche di ogni sensore distinte per tipologia

Nel caso in cui l'utente volesse compiere una ricerca per identificare un satellite che risponda a determinati prerequisiti a partire da un dato specifico, sarà sufficiente identificare nell'albero gerarchico la voce di interesse. Per ogni voce compariranno le informazioni correlate distinte per tipologia.

Nell'esempio di fig. 5 viene illustrato come, selezionando uno specifico intervallo di lunghezza d'onda (1.55 – 1.70 μm), vengono mostrate nella finestra adiacente tutte le informazioni associate (band name, band type, sensor, spat\_resolution). A questo punto selezionando il sensore (LISS-III sensor) e aggiornando la visualizzazione si otterrà come risultato l'elenco dei satelliti che montano quel particolare sensore in grado di lavorare sulla lunghezza d'onda di interesse.

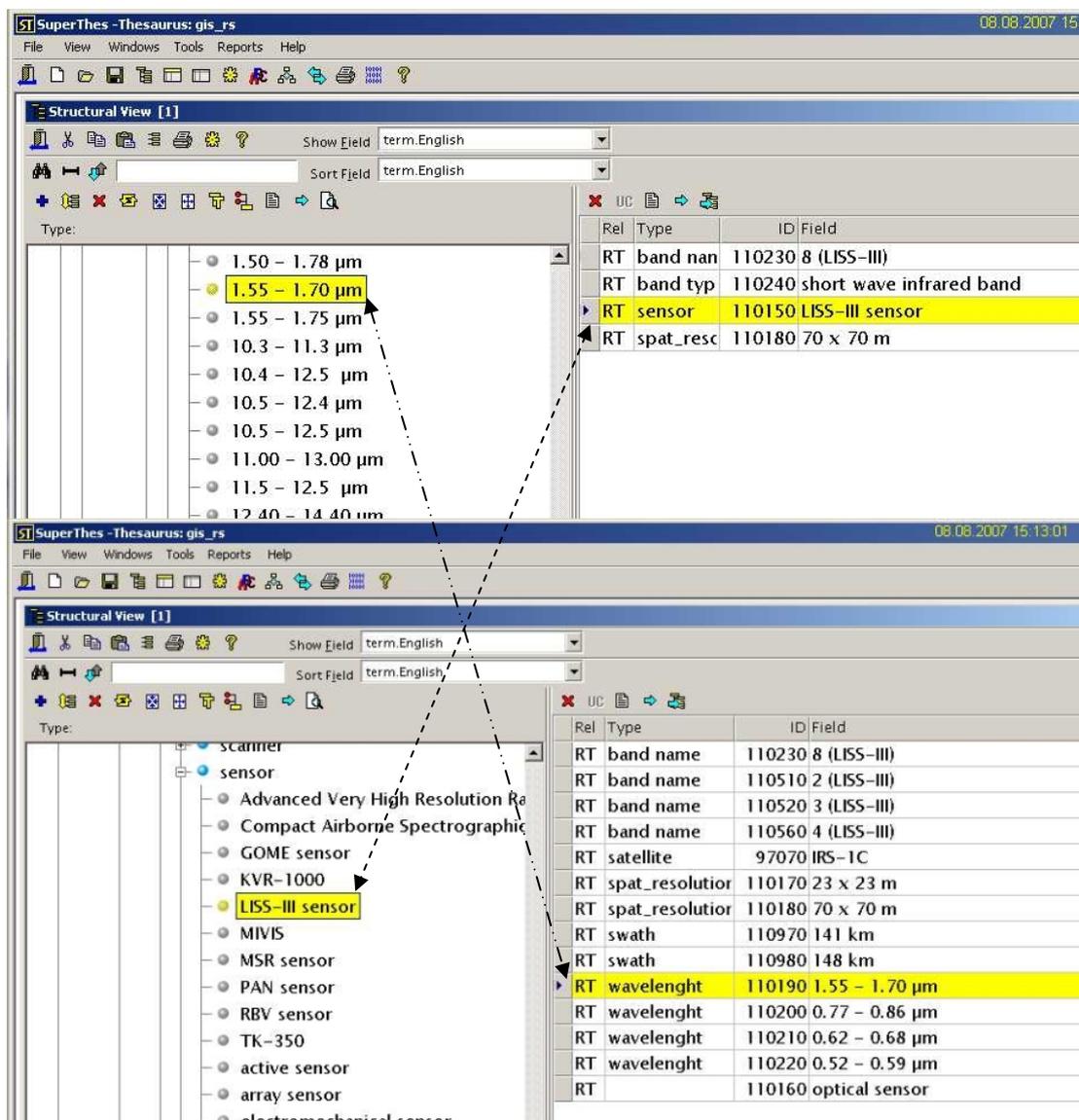


Figura 5 – Dettaglio sulle informazioni ottenute accedendo al sistema tramite intervalli di lunghezza d'onda.

## Conclusioni

Il prodotto realizzato, pur suscettibile di integrazioni e aggiornamenti, si propone di consentire una gestione ottimizzata dell'interscambio di dati e della comunicazione, sia in virtù dell'adozione di un linguaggio comune che per la maggiore comprensibilità semantica dei dati.

In termini operativi, sarà possibile per la comunità degli utenti, interrogare il sistema identificando la tecnologia migliore in funzione dell'obiettivo da raggiungere. E' infatti previsto, una volta realizzata l'interfaccia web attualmente in corso di sviluppo, di mettere a disposizione la banca dati su internet.

## Riferimenti bibliografici

AA.VV. (2007), *Satellites and sensors*. EO Edu, Belgian Science Policy. <http://telsat.belspo.be/bEO/>  
 ITC (2006), ITC's database of Satellites and Sensors.

<http://www.itc.nl/research/products/sensordb/searchsat.aspx>. (last modified: March 06, 2006)

Grignetti A, Plini P., Mazzocchi F., De Santis V. (2004), Terminologia e interoperabilità semantica per il Telerilevamento e il GIS. 8<sup>a</sup> Conferenza Nazionale ASITA. Roma, 14-17.XII.2004. Pp. 1255-1259.