

Il database di sintesi nazionale (DBSN) dell'Istituto Geografico Militare (IGM)

Carlo Perugi^(a) Cinzia Tafi^(b) Nicola Nozzoli^(c)

- (a) Istituto Geografico Militare, Direzione Acquisizione Dati, via Cesare Battisti 10, 50122 Firenze vdiracdati@geomil.esercito.difesa.it
- (b) Istituto Geografico Militare, Direzione Acquisizione Dati, via Cesare Battisti 10, 50122 Firenze caservgdbni@geomil.esercito.difesa.it
- (c) Istituto Geografico Militare, Reparto Informazioni Geografiche, via Cesare Battisti 10, 50122 Firenze noznoz@libero.it

PREMESSA

L'informazione geografica, grazie alle sempre più sofisticate tecnologie impiegate nei cellulari, tablet e navigatori satellitari, è entrata ormai a far parte della quotidianità di ognuno di noi e diamo ormai per scontato l'aiuto pratico che ne deriva. Possiamo, almeno per il momento, distinguere l'utilizzo "al dettaglio", con gli strumenti sopraccitati, delle informazioni finalizzate ad un'esigenza contingente come trovare il più vicino ristorante di pesce, trovare il più rapido itinerario stradale oppure un uso di più ampio respiro, fruibile con SW e HW diversi, dove le informazioni geografiche rilevate in modi diversi e riunite in un unico contenitore possono dare luogo a relazioni utili per la creazione di dati statistici. In questa seconda categoria di utilizzo si possono annoverare semplici interrogazioni come il numero di edifici scolastici esistenti in un determinato comune o operazioni più complesse di cosiddetto "data mining" con le quali si possono elaborare, incrociando varie tipologie di dati, ipotesi sulle ragioni di un certo fenomeno. Forse il primo esempio conosciuto in tal senso lo possiamo attribuire all'opera di un medico, John Snow che nel 1854 trovò la relazione fra i morti di colera e l'uso da parte della popolazione dell'acqua di una fontana situata in una strada di Londra, mappando la distribuzione dei casi della malattia nel quartiere di Soho rispetto all'esatta posizione della fontana.



Figura 1 – La mappa di Snow

IGM, in qualità di Organo Cartografico di Stato ha compreso le nuove esigenze dettate dalla necessità di elaborare dati sempre più completi e disponibili in tempo reale e, andando oltre l'usuale concetto di scala al quale era legata la sua produzione, ha focalizzato l'attenzione su ciò che è più importante nella creazione di una Banca Dati (DB) geografica, l'informazione completa, omogenea, accurata nella localizzazione e utile nell'ambito di

competenza che per un Ente statale non può essere che il territorio nazionale.

IL CONTENITORE

IGM sta quindi lavorando alla creazione di un database topografico nazionale (DBSN). La prima fase ormai conclusa ha comportato la raccolta ed omogeneizzazione dei dati provenienti da DB e CAD (Computer Aided Drafting) regionali in un unico modello di database. Tale modello deriva dalla Specifica di contenuto dei database geotopografici (Catalogo dei dati territoriali allegato al DPCM 11/11/2011) e ne costituisce un sottoinsieme di classi e attributi. Infatti da 159 classi divise in dieci categorie tematiche sono state selezionate 91 classi appartenenti a nove tematismi.

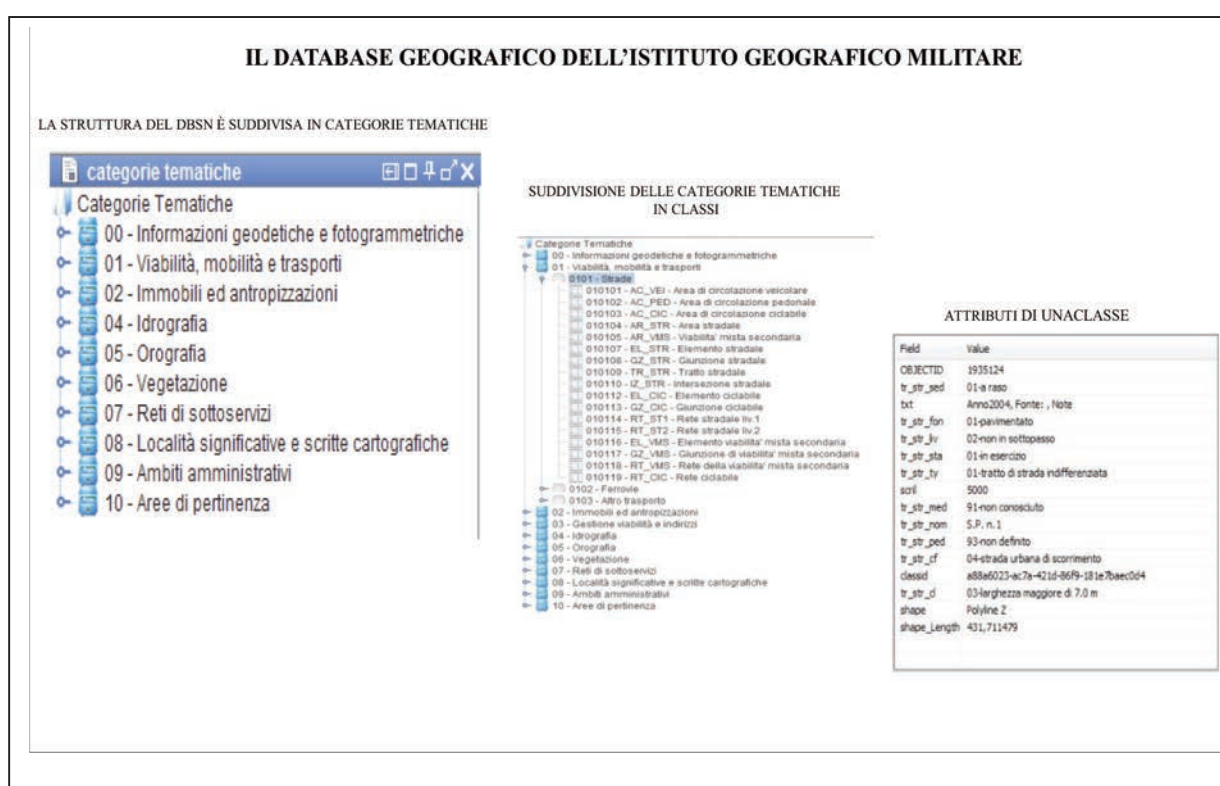


Figura 2 – Tematismi classi e attributi presenti nel Database

La selezione operata per le classi del Catalogo è stata effettuata proprio con la finalità di poter rappresentare e collocare le informazioni più significative nel database concepito per un ambito nazionale. DBSN sta per DataBase di Sintesi Nazionale. Il termine sintesi si riferisce alla selezione delle informazioni, cioè di classi e attributi, e all'elaborazione geometrica e topologica e alla fusione degli oggetti basata sulla identità dei valori degli attributi. Si riferisce inoltre alla possibilità di integrare i propri dati da più dataset di diversa origine. Il modello di dati deve quindi essere adeguato per un livello di scala nazionale. Un modello in cui la risoluzione informativa può, specialmente per certe classi, rimanere ad una scala nominale più grande ma che deve trovare il compromesso migliore per adeguare tale risoluzione all'estensione dei fenomeni descritti attraverso le classi di oggetti selezionate.

CARATTERISTICHE DEL PROGETTO

I dati appartenenti ad ognuna delle Regioni italiane presentano diversità di formato, di schema, di interpretazione e di aggiornamento. Nella prima fase il lavoro si è quindi concentrato nel rendere omogenei i dati nel formato DBSN utilizzando strumenti efficaci come FME (Feature Manipulation Engine). I dati regionali sono stati inclusi con la loro geometria originaria. Solo nei casi in cui lo schema DBSN ne prevede una diversa, il dato originario è stato adeguato. Non sono state quindi portate a termine operazioni di generalizzazione sui dati vettoriali atte a ridimensionarli ad una diversa ed unica scala nominale. La topologia è quella originaria dei dati stessi. La stessa filosofia viene applicata per gli altri dati che vengono inseriti, se necessario, ad integrazione della base regionale.

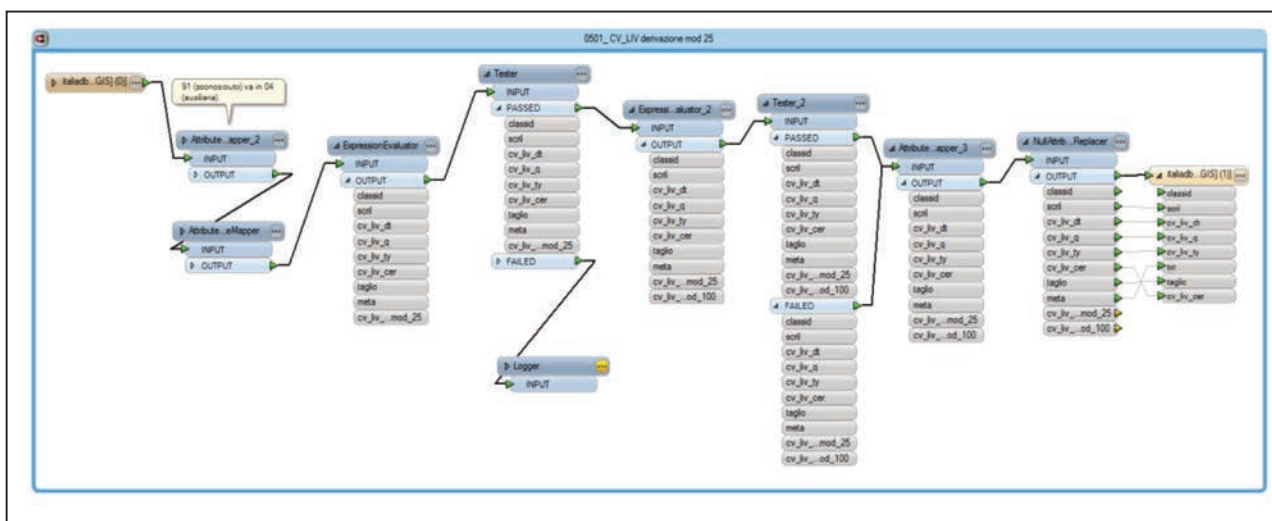


Figura 3 – Un diagramma di conversione dati fra uno schema regionale e quello DBSN

La scala nominale dei vari dati che vanno a popolare il DBSN varia a seconda della provenienza, per i dati regionali tipicamente è il 5K ma si possono trovare anche dati al 10K. Il catasto ha digitalizzato le informazioni partendo dalle canapine, storicamente alle scale 1K e 2K ed altre fonti di informazione raccolgono e assemblano vettori rilevati con metodi satellitari diretti. Vi è quindi una variabilità nell'accuratezza posizionale degli oggetti che comunque rimane in un ambito conforme alla risoluzione dei dati.

La fase attuale consiste nell'aggiornamento di alcune delle principali classi presenti nel DBSN con l'ausilio di fonti informative georiferite ufficiali e non ufficiali. Ogni record porta il metadata relativo alla sua provenienza e la data di acquisizione. Il piano degli interventi di aggiornamento prevede di concentrare l'attenzione, in un primo momento, su un certo numero di classi significative intervenendo successivamente su altre classi appartenenti ad altri tematismi. Quindi a regime si innesterà un ciclo continuo che dovrà mantenere aggiornato il Database intervenendo su un certo numero di classi alla volta. Attualmente siamo in una fase nella quale gli interventi di aggiornamento sono rilevanti, per la mancanza di omogeneità dei dati, e ciò comporta l'allungamento del periodo

di un ciclo completo mentre i prossimi cicli saranno ovviamente via via più brevi. L'unità di lavoro rimane per ora il territorio regionale salvo che per alcune classi per le quali ci può essere un approccio diverso.

AGGIORNAMENTO TRAMITE INTEGRAZIONE DA VARIE FONTI

Fonti di interesse per l'integrazione dei dati sono quelle ufficiali, come ad esempio l'ISTAT, le banche dati georiferite e non come per esempio l'elenco dei plessi scolastici di ogni ordine e grado e delle università, compilato dal MIUR (Ministero dell'Istruzione dell'Università e della Ricerca). In questo caso l'elemento di georeferenziazione consiste nell'esatto indirizzo, comune strada e numero civico, riportati nella tabella a fianco della denominazione dell'edificio scolastico. L'informazione è puntuale ma se esatta ed insistente su di un oggetto areale privo di attributi, serve a completare l'informazione. Fra le fonti non ufficiali, open source, costituisce un riferimento preminente Open Street Map (OSM), sicuramente il più aggiornato e completo specialmente per quanto riguarda i dati relativi ai Trasporti ma non solo. Infatti, per esempio, sono presenti molte informazioni puntuali relative alla destinazione d'uso di edifici di interesse: scuole, ospedali, edifici amministrativi etc.. Nella figura 4 il confronto fra strade ed edifici dove in verde sono i dati regionali ed in



rosso i dati appartenenti ad OSM. In questo caso vi è una certa congruenza per la viabilità mentre vi sono edifici non presenti in OSM. Per contro in questo dataset si possono riscontrare un maggior numero di informazioni puntuali riguardanti proprio gli edifici che nel dato regionale possono mancare. Per quanto riguarda i trasporti, viene lanciata una routine con la quale si confrontano i diversi dataset, mettendo in grado di agire in modo semiautomatico.

Figura 4 – Esempio di confronto fra dato regionale e dato OSM

L'integrazione avviene anche per il tramite della più recente edizione di ortofotografie a disposizione mediante le quali vengono effettuate le validazioni

su alcuni attributi delle classi da aggiornare.

CLASSI PRIORITARIE E VERIFICHE DI QUALITA'

La pianificazione del lavoro di integrazione prevede l'aggiornamento di alcune classi significative, per questo designate come prioritarie, appartenenti ai tematismi dei trasporti, dell'edificato, dell'idrografia, degli ambiti amministrativi e delle pertinenze. Si è scelto quindi di aggiornare in questa prima versione del DBSN le classi di cui sopra e rimandare a fasi successive l'aggiornamento delle rimanenti.

<p>GRAFI LINEARI DELLA VIABILITA' STRADALE E FERROVIARIA</p> <ul style="list-style-type: none"> • TRATTO STRADALE (TR_STR) • INTERSEZIONE STRADALE (IZ_STR) • ELEMENTO FERROVIARIO (EL_FER)
<p>IDROGRAFIA</p> <ul style="list-style-type: none"> • LINEA DI COSTA MARINA CARTOGRAFICA (CS_MAR) • ELEMENTO IDRICO (EL_IDR) • SPECCHIO D'ACQUA (SP_ACQ) • INVASO ARTIFICIALE (INVASO) • AREA BAGNATA DI CORSO D'ACQUA (AB_CDA) • AREA DI MARE (AR_MAR) • GHIACCIAIO-NEVAIO PERENNE (GHI_NV)
<p>AMBITI AMMINISTRATIVI</p> <ul style="list-style-type: none"> • COMUNE (COMUNE) • PROVINCIA (PROVIN) • REGIONE (REGION) • STATO (STATO)
<p>EDIFICATO E MANUFATTI</p> <ul style="list-style-type: none"> • EDIFICIO (EDIFC) (municipio, chiesa, stadio, stabilimento industriale, centrale elettrica, sede di scuola e università, sede di ospedale, stazione ferroviaria, museo) • MANUFATTO DI INFRASTRUTTURA DI TRASPORTO (MAN_TR) • DIGA (DIGA) • TRATTO DI LINEA DELLA RETE ELETTRICA (TR_ELE)
<p>PERTINENZE</p> <ul style="list-style-type: none"> • AREA A SERVIZIO STRADALE (SV_STR) • AREA A SERVIZIO DEL TRASPORTO SU FERRO (SV_FER) • AREA A SERVIZIO PORTUALE (SV_POR) • AREA A SERVIZIO AEROPORTUALE (SV_AER) • UNITA' INSEDIATIVA (PE_UIINS)

Nella figura a fianco l'elenco delle suddette classi. A parte l'idrografia, le classi appartenenti a gli altri tematismi sono certamente quelle più suscettibili a frequenti aggiornamenti.

Completate le operazioni di integrazione, vengono effettuate validazioni sui dati. I parametri di qualità presi in esame sono: accuratezza posizionale, completezza, accuratezza tematica. Una serie di controlli vengono eseguite su tutta la popolazione dei dati, altri a campione. Per il controllo dei dati del DBSN vengono utilizzate principalmente le ortofoto AGEA, e tutti i dati ancillari disponibili. L'analisi dei dati tende ad accertare che i

Figura 5 – Classi prioritarie

risultati ottenuti sul campione siano per deduzione significativi per l'intera regione. Confrontando le geometrie con le analoghe interpretate nell'ortofoto, si elaborano dei valori di accuratezza relativa che possono essere utilizzati per determinare valori di accuratezza assoluta conoscendo i parametri dell'ortofoto stessa. Per la completezza di ogni singola classe si determinano le percentuali degli oggetti errati/mancanti sui totali. Il controllo dell'accuratezza tematica interessa una selezione di attributi fra quelli obbligatori e ritenuti più importanti.

Un errore tematico è presente quando, in un singolo record sono stati inseriti uno o più attributi errati fra quelli esaminati.

PRODOTTI E SERVIZI

I dati del DBSN vengono vestiti con i segni convenzionali IGM per la scala 1:25000 in modo da predisporre l'allestimento cartografico che sarà visualizzato in un continuo geografico su apposito portale mediante il quale l'utente (ad oggi solo personale IGM) potrà scegliere e ritagliarsi l'area di interesse, corredata da informazioni marginali, in un formato immagine scaricabile.

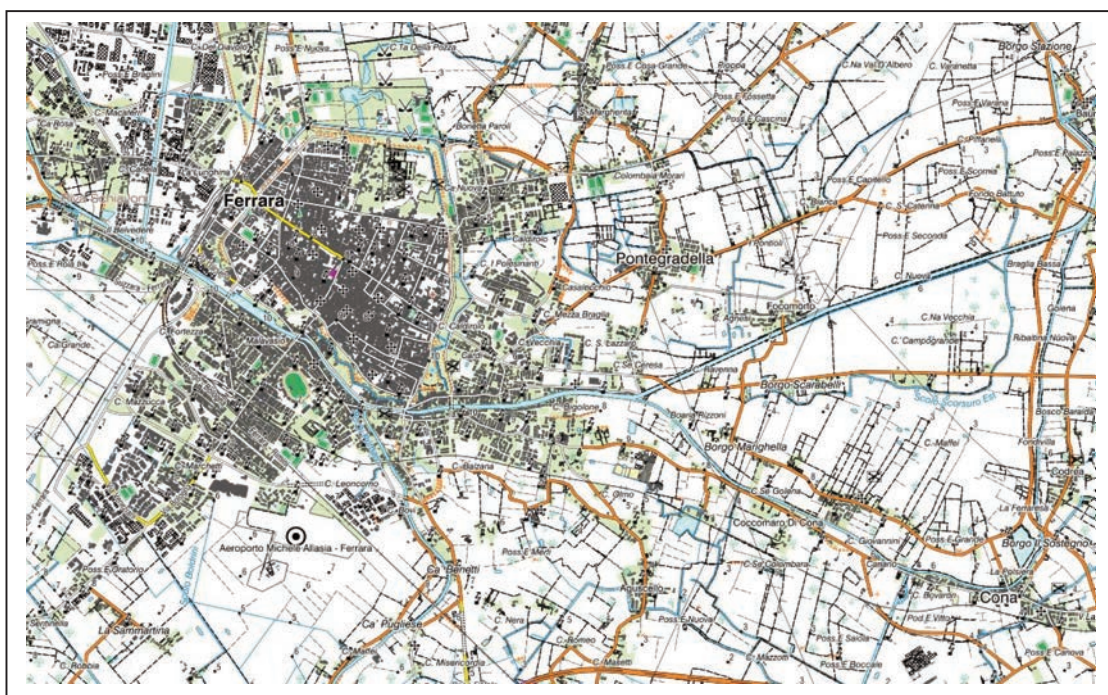


Figura 6 – Spezzone cartografico visualizzabile da WEB

Il taglio dello spezzone cartografico è piano quindi coincide con valori interi del reticolo cartografico e la sua ampiezza è di 18 Km in est e 14 Km in nord (252 Km quadrati). Il reticolato è il prodotto dalla trasformazione cartografica Trasversa Mercatore di meridiani e paralleli, appartenenti al Sistema di Riferimento ETRF2008, applicata ad un unico fuso con origine in corrispondenza del meridiano centrale dei 12° (Fuso Italia).

Nel progetto del servizio proposto, il layout cartografico è la visualizzazione dei dati contenuti nel DBSN, almeno di quelli rappresentabili cartograficamente. Considerando che nel database le informazioni devono essere raccolte costantemente perché, in special modo alcuni tematismi, soggette a continuo aggiornamento, le operazioni di vestizione cartografica devono essere automatiche. In effetti sono applicate alcune procedure di generalizzazione cartografica necessarie per la leggibilità della carta, facilitate dal fatto che il passaggio da una risoluzione informativa alla scala nominale 1:10000 a quella ad una scala 1:25000 non necessita di operazioni onerose e complicate.

Sono oggetto di studio l'istituzione di analoghi servizi per la costruzione di layout dalla scala 1:50000 a scale minori.

CONCLUSIONI

Il progetto tende a favorire la partecipazione e concertazione fra Enti produttori di dati geografici mettendo a sistema le informazioni di interesse nazionale. Il futuro del DBSN consiste nel favorire il processo di aggiornamento dei dati geografici per tutti i possibili utilizzatori a partire da quelli della PA (Pubblica Amministrazione) sia per l'ordinaria amministrazione che per gli episodi di emergenza.

BIBLIOGRAFIA

Steven Johnson, *The Ghost Map: The Story of London's Most Terrifying Epidemic – and How it Changed Science, Cities and the Modern World*, Riverhead Books, 2006, pp. 195-196, ISBN 1-59448-925-4.

Dulli Susi; Furini Sara; Peron Edmondo. *Data Mining*, Springer Verlag, 2009

DECRETO 10 novembre 2011 Regole tecniche per la definizione delle specifiche di contenuto dei database geotopografici. (Gazzetta Ufficiale n. 48 del 27/02/2012 - Supplemento ordinario n. 37). Allegato 1 - Catalogo dei dati territoriali – Specifiche di contenuto per i DataBase Geotopografici. Allegato 2 - Il modello GeoUML – Regole di interpretazione delle specifiche di contenuto per i DataBase Geotopografici

