

## **DATABASE TERRITORIALI (DBT) INTESAGIS: UNA PROPOSTA PROGETTUALE PER IL PASSAGGIO DAL DB10K A SCALE DI SINTESI**

Ugo FALCHI

Università di Napoli "Parthenope", Centro Direzionale isola C4, tel: +39.081.5476593,  
ugo.falchi@uniparthenope.it

### **Introduzione**

Le specifiche tecniche dell'IntesaGIS prima e le attività del gruppo di lavoro costituito in seno al CNIPA poi, non hanno ancora affrontato il problema della rappresentazione dei dati dei DBT alle scale di sintesi quali, per esempio, 100k o 250k, né la eventuale derivazione dei DBT a tali scale.

L'obiettivo di questa ricerca consiste nella redazione degli elementi essenziali di una proposta progettuale per una strutturazione e integrazione di un database, a scale di sintesi, all'interno di una base dati alla scala 1:10.000, DB10k.

Questo obiettivo sarà perseguito, per quanto possibile, nel rispetto della struttura del database originale, a partire dalle prescrizioni tecniche, "Specifiche di contenuto dei Database Topografici di interesse generale", redatto dal gruppo di lavoro Intesa GIS WG01 che ha operato nell'ambito dell'Intesa Stato-Regioni, utilizzate per la sua implementazione.

Il documento propone, di fatto, alcuni suggerimenti per una strutturazione originale di un modello logico per la convivenza di scale di sintesi, nel rispetto dei dati territoriali contenuti nel DBT alla scala di partenza.

Durante questo passaggio, si cercherà di preservare l'architettura del database geografico: in particolare la strutturazione delle classi di oggetti, con i loro attributi, e il loro raggruppamento in temi e questi in strati.

### **Abstract**

IntesaGIS and CNIPA have never dealt with database data representation at 100k e 250k. The objective of this research concerns with summary information structuring and integration within a database at a scale 1:10000, DB10k.

This goal will be pursued while respecting main database structuring, from the requirements contained in "Specifiche di contenuto dei Database Topografici di interesse generale", IntesaGIS WG01, and it offers some suggestions for a method of generating a logical data model.

To the extent possibile, one tries to preserve the architecture of geographic information, in particular, class, themes and layers categories.

### **Scale di sintesi**

La disponibilità di informazione geografica a scale di sintesi, è una esigenza sentita oramai da tutte le amministrazioni locali e nazionali che si occupano sia di pianificazione sia di gestione del territorio, in particolare, per quanto riguarda le emergenze ambientali e di protezione civile.

In un primo tempo sembrava che la strada maestra fosse quella di derivare i dati necessari, a partire da supporti rilevati a scala maggiore.

Andando in questa direzione, però, è necessario stabilire una serie di vincoli, relativi alla generalizzazione delle geometrie e le relazioni tra le classi, necessarie a garantire la congruenza topologica tra gli oggetti.

Sulla falsariga della fortunata sperimentazione condotta dal Centro interregionale nell'anno 2006, nel passaggio dal 10k a scale minori, 100K e 250k, è necessario effettuare una semplificazione delle geometrie, sia per esigenze di rappresentazione, si pensi, ad esempio, alla presentazione a video o alla stampa, le geometrie acquisite per una scala di riferimento originale sono, infatti, di difficile lettura ad una scala minore, sia per esigenze di lettura dei dati. E' prevista, infatti, una diversa primitiva geometrica per le diverse scale; per esempio alcuni manufatti che nel DB10k hanno una geometria poligonale, collassano a punti nel DB100k e nel DB250k; inoltre nel DB10k sono presenti oggetti non solo con geometria più dettagliata, ma anche con suddivisioni interne, portatrici, cioè, di un dettaglio tale che non è di interesse a scale di sintesi.

Per questo motivo, le operazioni che devono essere previste, devono riferirsi a:

- indicare il valore della superficie al di sotto della quale gli oggetti di geometria poligonale devono essere eliminati o degenerati in punti o linee;
- indicare se gli oggetti interessati sono, prima della eliminazione o degenerazione, inclusi in altri oggetti e se gli oggetti che li contengono ne acquisiscono la superficie;
- indicare il valore di lunghezza, al di sotto della quale gli oggetti con geometria lineare che non facciano parte di un grafo, vengono eliminati o degenerati in punti;
- indicare la distanza reciproca al di sotto della quale gli oggetti della classe dell'edificato, vengono resi contigui mantenendone però l'individualità;
- tenere conto del fatto che tutti gli oggetti, adiacenti a poligoni, che hanno subito la degenerazione a linea, devono essere modificati nella zona interessata dalla contiguità.

Per preservare l'integrità del dato, bisogna stabilire un costante e puntuale riscontro dell'informazione attraverso la verifica dei seguenti casi:

1. *elementi nuovi*: si applicano le procedure di semplificazione geometrica e si verifica se i risultati ottenuti, nel DB100k e nel DB250k, soddisfano tutte le relazioni ed i vincoli indicati per la classe di appartenenza; in caso positivo gli oggetti vengono introdotti automaticamente, in caso negativo invece vengono sottoposti ad una fase di *editing* manuale.
2. *elementi con diversa geometria*: si elimina l'oggetto nel DB100k e nel DB250k, e si trasferiscono gli attributi in quello nuovo proveniente dal DB10k; si applica poi quanto previsto per il caso 1.
3. *elementi mancanti nel DB10k*: si cancellano gli oggetti in questione dal DB100k e DB250k.
4. *elementi presenti in tutti i database con identica geometria*: si mantiene l'oggetto esistente senza modifiche.

Sarà necessario, infine, implementare e definire tre tipi fondamentali di vincoli, all'interno della gestione del *database*. Questi sono i vincoli del modello logico, i vincoli degli oggetti e i vincoli delle relazioni logiche.

I primi si riferiscono alla conservazione del contesto logico degli oggetti e del livello di dettaglio; i restanti due, hanno a che fare, sostanzialmente, con la conservazione delle forme, dei modelli e degli allineamenti tra gli oggetti, laddove esistono interazioni tra di essi.

Tuttavia, le numerose esperienze condotte in Europa e negli Stati Uniti hanno evidenziato come la produzione di carte generalizzate su larga scala sia afflitta dal livello tecnologico raggiunto dalla attuale generazione di programmi e macchine, incapaci per strutturazione, ad affrontare le complesse tematiche legate all'approccio semantico e topologico.

Inoltre, anche i prototipi sviluppati, hanno messo in luce come sia necessario un tempo molto alto di elaborazione per ottenere risultati rappresentativi di coperture cartografiche molto grandi.

Pragmatismo ha voluto che la ricerca si indirizzasse verso soluzioni alternative, capaci di interessare al massimo la presenza di supporti già realizzati e di coinvolgerli, insieme con le informazioni originali, nel processo di integrazione e aggiornamento.

In questo caso, in assenza di specifiche tecniche elaborate in seno ai gruppi di lavoro su scala nazionale, che ha operato riguardo a 1k, 2k, 5k, e 10k, sarà necessario, a partire dalla documentazione

già redatta per le scale maggiori, elaborare una struttura originale per le informazioni geografiche con dettaglio minore. Tale intervento sarà da applicarsi anche laddove i supporti cartografici non siano già stati strutturati all'interno di banche dati, rispettose di questi standard.

### **Rappresentazioni multiple**

L'obiettivo principale della realizzazione di un sistema informativo o di una banca dati è quello di fornire una descrizione, la più verosimile possibile, del mondo reale ivi rappresentato.

Spesso, vista l'elevata specializzazione di tali sistemi, il punto di vista che si vuole veicolare, non rispetta dei criteri generalisti, ma punta a contenere tutte le informazioni che utenti specifici del sistema necessitano per lo svolgimento delle proprie attività.

In questo modo, tenuto conto della complessità del sistema in cui viviamo e della moltitudine degli attori e delle amministrazioni a cui appartengono, deputate in diversa misura, alla raccolta, implementazione, gestione e aggiornamento di informazioni geografiche, territoriali e ambientali, si è assistito a una vorticosa moltiplicazione dei database utilizzati, spesso e volentieri all'interno della stessa struttura.

Differenti approcci all'uso delle informazioni geografiche richiedono, infatti, differenti set di dati, sebbene questi insistano sulla stessa realtà geografica.

Il diverso punto di vista dell'utilizzatore, che influisce in maniera determinante sull'architettura della base dati, ovvero su quale tipo di visione essa avrà del mondo reale, determinerà una particolare struttura dei dati e, di essi, una particolare scelta dei valori e degli attributi che dovranno veicolare. Così come differenti utenti possono condividere la stessa visione ma necessitano di una scala diversa del dato informativo.

E' ragionevole pensare che punti di vista e risoluzioni debbano essere integrati nel sistema come indipendenti tra di loro. Il primo, caratterizzato dallo sguardo di chi utilizzerà le informazioni, avrà influenza su come i dati saranno conservati e rappresentati, gli oggetti che non ne faranno parte, quali attributi essi avranno, quali relazioni li legheranno e come gli attributi saranno valutati. Il secondo, sarà attinente al livello di dettaglio scelto e ai criteri geometrico dimensionali degli oggetti in esso veicolati; i criteri semantici determineranno, invece, la granularità dell'acquisizione delle informazioni, ovvero quanto sarà spinta la gerarchizzazione degli oggetti rappresentati.

Basti pensare alle diverse sensibilità mostrate nell'utilizzo della classe di oggetti relativa all'edificato. In alcuni ambiti è necessario che ogni *feature* porti con se specifiche informazioni legate, per esempio, alla sua volumetria, alle quote in gronda e al piede, alla definizione delle estrusioni e dei cavetti e alle pendenze della copertura, laddove presenti. Contemporaneamente, per le stesse entità potrebbe essere necessario conoscere esclusivamente l'ingombro rispetto agli oggetti adiacenti, quali la rete stradale o le vie pedonali.

Il costo elevato, legato al mantenimento di strutture indipendenti tra di loro e le enormi difficoltà relative alla propagazione degli aggiornamenti, ha reso necessario indirizzare gli sforzi verso soluzioni di integrazione con lo sviluppo di applicazioni *ex novo*; questo perché gli attuali sistemi di gestione, (GIS o database), non possiedono funzionalità native adeguate al coordinamento di rappresentazioni multiple dei fenomeni geografici che ci circondano.

Sono stati sviluppati diversi approcci per l'unificazione di informazione geografica nei database, principalmente riconducibili a due modelli: il primo, chiamato *data warehousing approach*, consiste nella sostituzione dei database iniziali con uno centralizzato (Elmagarmid et al. 1999), come indicato in figura 1 con un esempio molto banale; il secondo, chiamato *federated* o *mediation approach*, mira a mantenere separate le basi dati e costruire uno schema comune che li connetta (Sheth and Larson 1990), (Parent and Spaccapietre 2000), o semplicemente implementare particolari relazioni tra i modelli logici esistenti (Kilpelainen 1998).

In tutti i casi, è necessario sviluppare un modello dati innovativo, capace di descrivere opportunamente la visione globale connessa con le diverse scale.

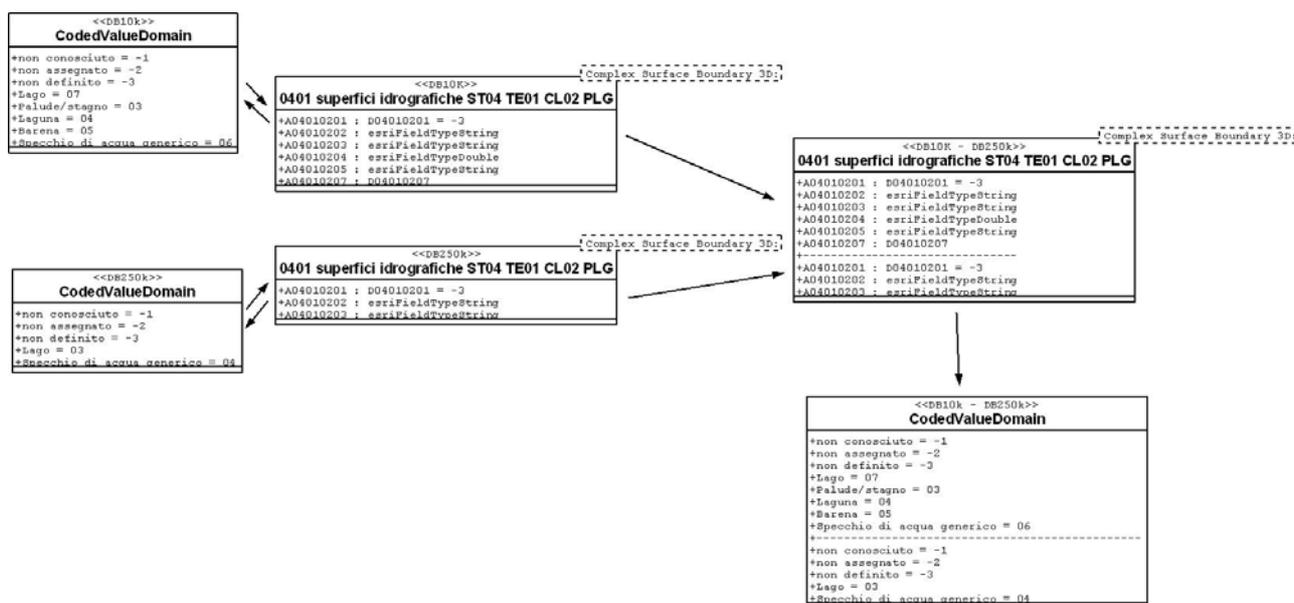


Figura 1 – integrazione di due classi, DB10k e DB250k

### Lo schema concettuale

Sulla scia dei risultati raggiunti dal progetto europeo MurMur, (Multi Representations - Multi Resolutions) e dallo sviluppo di MADS, un modello concettuale per lo sviluppo di applicazioni tradizionali e spazio – temporali, (un progetto diretto dal Prof. Christine Parent (Université de Lausanne) e Stefano Spaccapietra (EPFL), con la cooperazione del Prof. Esteban Zimanyi, Université Libre de Bruxelles), è stato possibile iniziare ad ipotizzare la progettazione di uno schema globale, capace di supportare rappresentazioni multiple, dalla gestione del singolo oggetto fino ai suoi attributi ed ai valori da loro posseduti.

Un tipo – oggetto, infatti, così come un tipo – relazione, può veicolare un diverso set di dati geografici, legati alla diversa scala di acquisizione o al diverso “punto di vista” della realtà rappresentata.

Il modello dati, rappresentato dall’esperienza di MADS, è sostenuto dai seguenti concetti:

1. **Oggetti:** sono organizzati in gerarchie dinamiche, gli oggetti possono cambiare la loro appartenenza alle classi, spostandosi da una ad un’altra e sono *multi-instantiation*, (creazione di classi multiple di oggetti), le entità del mondo reale possono essere descritte da due o più istanze appartenenti a due classi diverse della gerarchia; per esempio, un tratto di corsia può essere descritto come pedonale e ciclabile, entrambi tipi-oggetto, sotto-tipi di un super-tipo corsia;
2. **Relazioni:** collegano due o più oggetti. Come i tipi-oggetto, possiedono attributi e metodi e possono essere organizzate in gerarchie di classificazione. Sono supportati diversi tipi di relazioni, le semplici, senza vincoli o semantica o le aggregazioni. Esistono molti tipi di aggregazioni come nel mondo reale; nessuna aggregazione porta con se automaticamente dei vincoli, per esempio, le componenti degli oggetti possono essere condivise da molti oggetti composti, altrimenti, un vincolo può essere esplicitato da una particolare gerarchia dei valori delle sue componenti. L’equivalenza è un altro tipo di relazione binaria con uno specifico significato (due istanze dell’oggetto rappresentano la stessa l’entità del mondo reale); è una relazione simmetrica con un vincolo intrinseco; la sua cardinalità, infatti, è 1:1, ovvero ogni istanza di un tipo- oggetto collegata da una relazione di equivalenza deve essere collegata al massimo ad una istanza dell’altro tipo-oggetto.
3. **Attributi:** possono essere semplici o composti, per esempio, composti da altri attributi e possiedono valori singoli o un *set* di valori multipli.

E' stato implementato, inoltre, un nuovo tipo di dato, relativo al tempo, che affiancandosi alla definizione spaziale delle entità rappresentate, amplia notevolmente le capacità di questo modello logico, in particolare è composto da:

- un set di dati spaziali che include, (vedi OpenGIS consortium):
  - o semplici: *point*, *line* e *surface*
  - o compositi: *pointset*, *lineset* e *complexsurface*
  - o generici: come *Geo*, il super-tipo di tutti i dati spaziali;
- un set di dati temporali che includono:
  - o semplici: *Instant* e *TimeInterval*
  - o compositi: *InstantSet* e *IntervalSet*
  - o generici: *Time*, il super-tipo di tutti i dati temporali;
- un tipo di dato specifico, *Varying(D, S/T)* che definisce, per ogni dominio D, un nuovo dominio di valori che varia nello spazio, nel tempo o in entrambi. Ogni valore di *Varying(D, S/T)*, è una funzione che definisce, per ogni punto nello spazio, (o ogni istante nel tempo o entrambi) un valore di D;
- dei tipi-oggetto; essi possono essere definiti come:
  - o spaziali, essi possiedono un particolare attributo chiamato geometria che definisce l'estensione spaziale dell'oggetto; il suo dominio è uno dei tipi di dati spaziali;
  - o temporali, essi possiedono un particolare attributo chiamato *lifecycle* che definisce l'estensione temporale di un oggetto; il suo dominio è uno dei tipi di dati temporali;
- ogni oggetto o relazione può avere un certo numero di attributi di tipo tematico, spaziale, temporale o *varying*.

Per unificare, all'interno della stessa banca dati, informazioni geografiche a differente livello di dettaglio, è necessario, però, definire un ulteriore elemento che ci possa permettere di qualificare e riconoscere un determinato set di dati da un altro, appartenente a una rappresentazione diversa, per scala o per contenuti veicolati.

Questo elemento, chiamato *DRstamp* nel progetto MADS, ha il compito di materializzare, per una data rappresentazione, la sua appartenenza a uno schema preciso, contenuto all'interno del database. Grazie a ciò, è possibile per un utente che interroga la base dati, visualizzare direttamente il *set* di informazioni cercate, semplicemente richiamando la "marca" che li contraddistingue.

Gli attributi posseduti da ogni oggetto o relazione, possono, allora, possedere diverse "marche" che determinano per quali rappresentazioni, (per scala o punto di vista) esse sono valide.

In questo modo, che si decida di integrare le diverse informazioni geografiche all'interno dello stesso "contenitore" o si decida di mantenerle fisicamente separate ma collegate da relazioni binarie, (*relationship*), la presenza di un indicatore importante come la "marca" offre a chi progetta la base dati, insieme con il bagaglio di concetti sopra riportato, la possibilità di implementare uno schema multi – rappresentazione per descrivere i fenomeni del mondo reale.

Una volta effettuata la ristrutturazione delle informazioni a scala di sintesi, 100k e 250k, secondo le specifiche tecniche proprie del DB10k, sarà possibile lavorare sul modello logico al fine di introdurre le innovazioni presentate.

### **Il modello logico unificato**

Il rispetto della congruenza delle informazioni, nonché la necessità di garantire la propagazione degli aggiornamenti, rende necessario effettuare un confronto approfondito tra le diverse istanze che caratterizzano tutti gli oggetti contenuti nella base dati, alle diverse scale o ai diversi "punti di vista".

Queste operazioni, pianificate in modalità manuale o automatica, secondo il tipo di tecnologia impiegata e dell'estensione delle informazioni, hanno come obiettivo la conoscenza dell'andamento delle geometrie delle *feature*, la diversa complessità e il grado di sovrapposizione tra le stesse.

Per esempio, la rappresentazione di un tratto stradale, alla scala di acquisizione 1:10000, 1:250000 e 1:2000, presenterà un andamento diverso, caratterizzato dalla presenza del doppio ciglio piuttosto che della linea singola, la definizione degli incroci piuttosto che il loro rilievo come puntuali e la diversa densità di acquisizione dei percorsi.

La consapevolezza dei diversi sviluppi degli oggetti, permetterà di definire quale tipo di rapporto intercorre tra di essi e, grazie ad esso, stabilire con precisione una tavola di alias geometrici, necessaria a stabilire quale sia la migliore struttura dei dati da costruire, sebbene essa debba essere redatta nel rispetto, per quanto possibile, delle specifiche tecniche che sono alla base dell'implementazione del *database* originale, come indicato nei presupposti di partenza.

Le possibili combinazioni, per esempio tra uguali sezioni stradali alle diverse scale, possono essere così sintetizzate:

- a. 1 : 1, quando ad un tratto stradale ne corrisponde uno e uno soltanto alle altre scale;
- b. 1 : n, quando ad un elemento delle vie di trasporto corrisponde un numero maggiore di istanze, alle altre scale di interesse, per esempio, è il caso di un incrocio a rotonda rappresentato da un elemento puntuale alle scale di sintesi e dal suo reale andamento geometrico alle scale di maggior dettaglio;
- c. n : m, quando la granularità dell'informazione, comunque complessa, si sviluppa in un numero non corrispondente di elementi, è il caso, per esempio, di un sovrappasso autostradale.

Le diverse combinazioni sopra esposte guideranno la mano del progettista nello stabilire se quelle istanze debbano confluire in una unica classe multi – rappresentazione o se sia necessario ricorrere a soluzione più complesse, quali, ad esempio, il collegamento delle classi originali tramite una relazione o la creazione di una nuova classe capace di mantenere la relazione con i diversi oggetti appartenenti alle scale utilizzate.

### **Conclusioni**

Questo lavoro rappresenta il primo passo di una ricerca più ampia che ha come obiettivo, l'applicazione delle metodiche sopra accennate, ad una banca dati istituzionale di grande estensione. La continua evoluzione nell'impiego di banche dati, richiede la progettazione di strutturazioni sempre più efficienti, in grado di sostenere tutte quelle applicazioni mirate all'incrocio delle informazioni territoriali con dati fiscali e catastali. Basti pensare, per esempio, al decentramento catastale, all'informatizzazione galoppante delle procedure di concessione edilizia, della pianificazione attuativa e generale e delle trasformazioni edilizie, alla necessità per gli enti locali di recupero dell'evasione e dell'elusione fiscale.

La necessità di estrazione di indici sintetici e di aggregazione delle informazioni, spesso e volentieri acquisite con diverso dettaglio o con diverse finalità, unitamente all'esigenza di mantenere tutti questi dati aggiornati, richiede, infatti, il superamento della visione fin qui seguita sia dagli ambienti software commerciali e *open source* sia dalle indicazioni tecniche a livello locale e nazionale.

Lungi da considerarsi una soluzione definitiva, la traccia seguita in questo lavoro cerca di fornire una soluzione adeguata, in linea con i risultati delle più importanti esperienze condotte da alcuni gruppi di lavoro europei.

### **Riferimenti Bibliografici**

Regione Autonoma della Sardegna (2005), "Implementazione e realizzazione di data-base topografici comunali "

IntesaGIS, "Specifiche per la realizzazione dei data base topografici di interesse generale", 1n1007(1,2,3,4,5,6)

Balley, S., Parent, C. and Spaccapietra S. (2006), "Modeling geographic data with multiple representations", International Journal on GIS (IJGIS). v18 iJune. 329-354.

MurMur, *Multi-representations and multiple resolutions in geographic databases*, "Project 10723 - 1.1.2000 to 31.12.2002"