

Database geografici

GIANFRANCO AMADIO

Istituto Geografico Militare

Il fruitore delle informazioni geografiche è abituato alla consultazione della carta ben conoscendo le regole e le metodologie attraverso cui il mondo reale viene in essa rappresentato.

L'automazione cartografica e la cartografia numerica prima ed i *database* geografici poi hanno modificato questo stato di cose e adesso anche il mondo delle informazioni geografiche è legato a quello dell'informatica, della quale utilizza le regole ed il supporto di rappresentazione.

Le tecnologie informatiche sono pertanto divenute strumenti necessari, che bisogna conoscere, ma che, com'è ovvio, non si sostituiscono alla tradizionale cultura cartografica e geografica.

Sempre più le informazioni geografiche sono disponibili solo su supporto informatizzato (basti pensare alle *warehouse* geografiche accessibili tramite Internet), e sempre meno le «informazioni geospaziali» hanno anche una rappresentazione grafica tradizionale su carta.

Nel contesto di questa opera, la presente scheda si prefigge l'obiettivo di introdurre il lettore al mondo dei *database* geografici e di fornire strumenti conoscitivi attraverso i quali i «tipi geografici», individuabili tradizionalmente tramite i segni convenzionali cartografici, possano essere riconosciuti ed analizzati disponendo di un *database* geografico e di un relativo *software* di interrogazione.

Database generico

È una raccolta organizzata di dati, in accordo con una definita struttura concettuale, relativi a caratteristiche di oggetti (entità) appartenenti ad una certa area di applicazione.

Sono *database*, per citare alcuni esempi, l'anagrafe comunale (dati relativi ai residenti in un comune), il magazzino prodotti (caratteristiche degli articoli di un supermercato), i dati relativi ai dipendenti di una società (situazione anagrafica, qualifica, stipendio, ecc.).

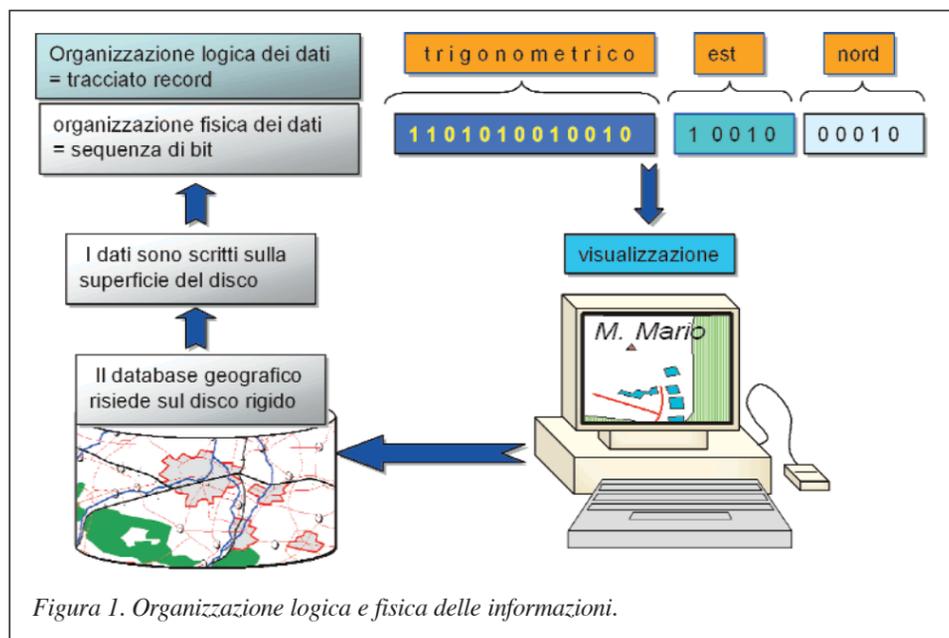
Il termine *database* è comunemente associato a dati residenti su supporto informatico; gli stessi dati costituiscono comunque un *database* se riportati su supporto cartaceo in forma di tabulati, tabelle o grafici.

Nel primo caso la gestione del contenuto del *database* (creazione, aggiunta, modifica, cancellazione, ricerca) avviene tramite un *software* detto *Data Base Management System* (DBMS), nel secondo caso la gestione avviene manualmente.

In termini informatici, le singole caratteristiche degli oggetti costituiscono i campi (*field*); l'insieme dei campi forma il *record* descrittivo dell'oggetto; l'insieme dei *record* relativi a più oggetti della stessa raccolta costituisce l'archivio (*file*). Questa organizzazione logica dei dati è descritta con il cosiddetto «tracciato *record*» che definisce il contenuto informativo del *database* (figura 1).

Database geografico: GeoDB

È una raccolta organizzata di dati relativi a caratteristiche e posizione di oggetti (entità) geografici. Un oggetto geografico è un particolare del mondo reale, di origine naturale o generato dall'opera dell'uomo, di tipo fisico o frutto di convenzione, non ulteriormente suddivisibile in accordo con il modello dei dati adottato (es. casa, strada, serbatoio, stazione ferroviaria, limite amministrativo comunale etc.).



Dati geografici sono quelli relativi alla localizzazione e descrizione di oggetti geografici e comprendono posizione, forma, caratteristiche e relazioni con altri oggetti limitrofi.

La presenza del dato di posizione comporta che un *database* geografico contenga oggetti «georeferenziati» rispetto ad un determinato «sistema di riferimento»; la localizzazione spaziale e la forma di oggetti geografici sono registrati come punti, linee, poligoni (rappresentazione vettoriale), *pixels* (rappresentazione *raster*), grigliati di celle regolari (rappresentazione *matrix*) o triangolari (superfici tridimensionali) con le loro caratteristiche, gli attributi e le relazioni reciproche.

L'attributo esprime una caratteristica, qualitativa o quantitativa, dell'oggetto geografico; ad esempio sono attributi dell'oggetto geografico «strada» il numero di corsie, la presenza di spartitraffico, il nome, ecc.

Caratteristica fondamentale del GeoDB è quella di non lasciare «buchi», ovvero aree non qualificate, all'interno della estensione geografica di copertura; la superficie del terreno è pertanto completamente codificata, senza soluzione di continuità.

Modello di astrazione

Partendo dal mondo reale percepito attraverso i sensi, misurato tramite gli strumenti ed interpretato tramite la sua conoscenza, il modello di astrazione consiste nell'individuare ed isolare un certo numero di fenomeni ed oggetti geografici che hanno una loro individualità o funzione, attraverso i quali si ritiene di esemplificare la realtà geografica in relazione ai propri fini.

Questo processo, similmente a quello utilizzato nella cartografia tradizionale, nel campo dei DB identifica le «entità geografiche» da rappresentare, con le loro caratteristiche, proprietà e mutue relazioni.

Schema e modello concettuale

Uno schema definisce le proprietà generali che potranno possedere le singole istanze di informazione che verranno caricate nel *database*; pertanto, un modello per definire lo schema deve permettere di identificare le classi e le entità, quali i fiumi e le strade, e le proprietà che le singole istanze, cioè la strada o il fiume, potranno possedere. Tali proprietà sono: gli attributi con i relativi domini, inclusi gli attributi geometrici, le classi di associazioni tra entità e gli identificatori. Questo modello, orientato a definire le proprietà intrinseche dell'informazione, è chiamato «modello concettuale», e gli schemi che con esso si realizzano sono detti «schemi concettuali». Il modello ER (entità/relazioni) è il più diffuso nei *database*, con una tendenza ad essere sostituito dalla parte più *database oriented* del modello ad oggetti.

Nel complesso, il modello di astrazione e il modello concettuale portano alla definizione del contenuto informativo del DB.

Modello di rappresentazione

Gli oggetti o fenomeni (entità), individuati dal modello di astrazione e qualificati dal modello concettuale, vanno a costituire il *database* geografico attraverso il modello di rappresentazione.

Questa operazione può avvenire con un processo che utilizza la forma implicita o esplicita della realtà. La prima tende a rappresentare l'oggetto nella sua dimensione sostanziale, la seconda tende invece a rappresentarlo così come i nostri sensi lo percepiscono.

Vector

La forma (implicita) *vector* o vettoriale rappresenta gli oggetti definiti dal modello di astrazione in modo schematico ma fedele; per ognuno di essi, a seconda della sintesi e risoluzione di rappresentazione scelte, utilizza una appropriata primitiva geometrica che è descritta, per ogni punto o vertice, da una coppia (2D) o da una tripletta (3D) di coordinate riferite ad un certo sistema di riferimento geodetico-cartografico.

Le primitive geometriche elementari utilizzate sono: punto, linea e poligono. Con la «primitiva punto» si rappresentano oggetti puntuali (trigonometrici, ecc.) a dimensione zero o altri ad essi assimilabili (pozzi, tralicci, ecc.). La «primitiva linea» viene utilizzata per rappresentare oggetti lineari (limite amministrativo, limite di bacino, ecc.) a dimensione uno o altri ad essi assimilabili (strada, fiume, curva di livello, ecc.). Con la «primitiva poligono» si rappresentano oggetti ad estensione areale (bosco, lago, fiume, ecc.) a dimensione due.

Nella figura 2.a sono riportate, per alcuni oggetti geografici, la rappresentazione vettoriale e quella corrispondente all'interno dell'archivio.

Raster

È una forma esplicita di rappresentazione bidimensionale con la quale il dato geografico è discretizzato ed espresso come matrice di celle o *pixels*. Il termine *pixel* deriva da *picture element* e indica un'immagine divisa in modo regolare in piccoli elementi (tassellazione), in modo da poter esprimere le quantità associate a ciascuno di essi. Esiste anche una forma tridimensionale, il *voxel*, che però non ha una significativa applicazione.

Con la forma *raster*, il dato geografico è rappresentato in modo discontinuo e diviso in unità discrete; questo rende i dati *raster* particolarmente adatti per certi tipi di operazioni, come ad esempio *overlay* e calcoli di aree. Ad ogni *pixel* corrisponde una definita area di territorio in base alla sua dimensione.

Nell'archivio *raster*, la componente di posizione è esplicitata per un solo *pixel*, ad esempio per quello in basso a sinistra; data la uniformità e geometricità della matrice, la posizione degli altri *pixels* rimane definita conseguentemente (figura 2.b).

Si distinguono due tipologie di *raster*: *image* e *graphic*. Alla prima categoria appartengono le ortofotografie digitali, alla seconda la scansione e georeferenziazione di cartografia tradizionale.

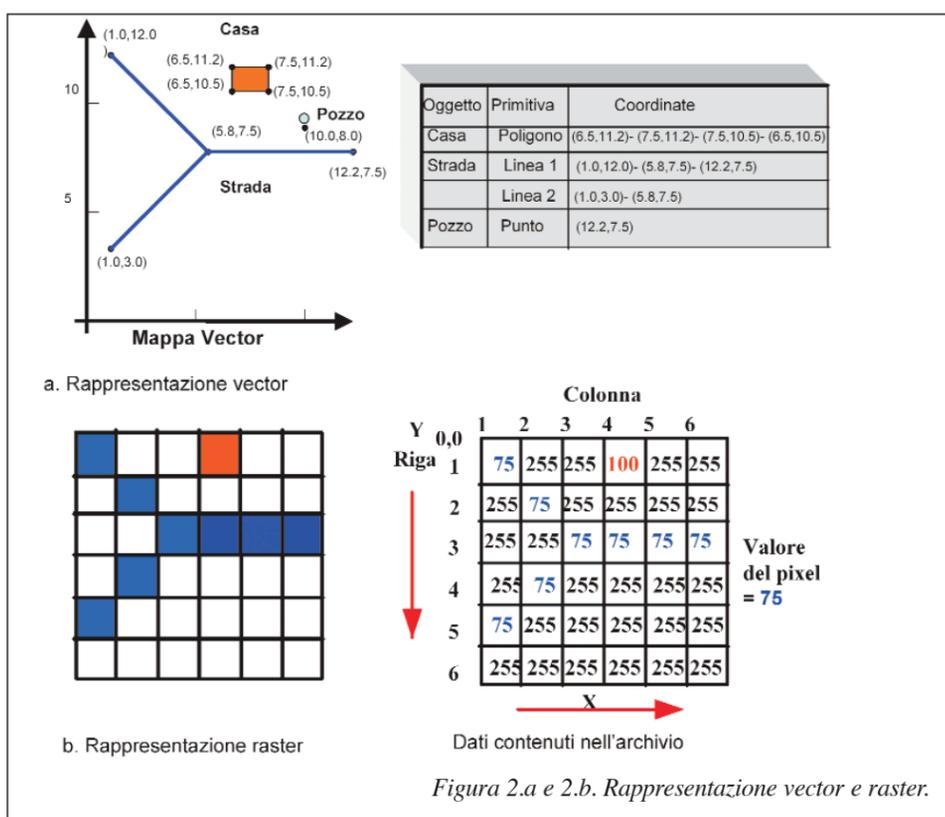


Figura 2.a e 2.b. Rappresentazione vector e raster.

Matrix

La *matrix* è una particolare forma di *raster* da cui eredita gran parte delle caratteristiche, ma se ne differenzia in quanto il contenuto dei *pixels* non è riferito a valori radiometrici riflessi o emessi, ma consta di valori misurati o calcolati di un certo fenomeno geografico. Le procedure di *remote sensing* utilizzano in ingresso immagini multispettrali, quindi *raster images* e, dopo le procedure di classificazione che sfruttano le firme spettrali degli oggetti, generano una matrice di codici, cioè un *matrix*. In quanto matrice di valori o codici, la *matrix* è facilmente utilizzabile, ma non è immediatamente visualizzabile. Perciò si ricorre normalmente all'espedito di tradurlo in immagini *raster* (che però non è più possibile utilizzare numericamente) associando tonalità di colore o sfumature di grigio ai valori che identificano il fenomeno (figura 2.c). Tipico uso della forma *matrix* è quello di contenere i dati del modello numerico del terreno (DTM), delle mappe di pendenza o di copertura del suolo. Nelle realizzazioni pratiche la base dati geografica è composta da tutte queste forme di rappresentazione (figura 3).

Modello fisico

Il modello fisico è l'insieme dei formati e delle strutture (tracciati *record*) mediante i quali le varie componenti dell'informazione geografica vengono immagazzinate sulle memorie di massa del sistema. Quindi un DB geografico realizzato con un certo modello di astrazione, concettuale e di rappresentazione, può essere implementato su un sistema *hardware/software* con diversi modelli fisici. Questi dipendono dai prodotti *software* commerciali utilizzati per gestire il GeoDB.

Componenti d'informazione

In generale, le componenti d'informazione di una entità geografica sono quattro: geometrica, di attributo, di relazione e di informazione.

Nella forma vettoriale sono rappresentate rispettivamente dalle primitive geometriche, dagli attributi, dalle relazioni e dai *metadata*:

- Primitive geometriche: punto, linea e poligono. In fase di applicazione del modello di astrazione, per ogni entità geografica viene scelta la relativa primitiva geometrica da utilizzare per la sua rappresentazione.

- Attributi: sono caratteristiche descrittive degli oggetti; ad esempio, per un edificio la destinazione d'uso. Per ogni attributo vengono anche definiti i relativi domini che costituiscono i possibili valori che esso può assumere. Per un ponte, ad esempio, il materiale di costruzione potrà essere legno, ferro, calcestruzzo.

- Relazioni: costituiscono le proprietà spaziali degli oggetti geografici. Sono esplicitate tramite la topologia ed identificano i vincoli spaziali ammessi fra categorie di oggetti. Le relazioni topologiche utilizzate sono: connessione, adiacenza, appartenenza; i vincoli sono: *disjoint*, *touch*, *in* e *overlap* (figura 4). Questa componente, finora spesso implicita e nascosta in strutture dati specializzate, nei moderni GeoDB deve essere indicata in modo esplicito. La gestione delle informazioni territoriali mediante classi di oggetti comporta che vengano esplicitate sia le relazioni che intercorrono tra loro sia i vincoli che devono essere rispettati per garantire la validità del contenuto informativo nel suo complesso. Nelle forme *raster* e *matrix* l'unica primitiva geometrica è il *pixel*, che può possedere, come unico attributo, il valore, e non è in grado di esplicitare relazioni o vincoli.

- *Metadata*: oltre ai dati geografici veri e propri, i *database* geografici, quali essi siano, sono corredati di «dati sui dati» denominati *metadata*, ed ancora di dati sui *metadata*; ciò costituisce la «metainformazione».

I *metadata* rivestono enorme importanza per la ricerca, l'utilizzo ed il riutilizzo dei dati per scopi diversi da quelli per i quali sono stati prodotti. Appartengono ai *metadata* i dati relativi al sistema di riferimento geodetico-cartografico, il produttore, l'epoca di produzione ed aggiornamento, ecc.

Topologia

La topologia in matematica studia l'invarianza di alcune proprietà delle figure geometriche, quando esse sono soggette a deformazioni di tipo continuo. La topologia applicata alle informazioni geografiche esprime il modo con cui gli oggetti geografici si connettono fra di loro ed il loro mutuo rapporto spaziale, utilizzando le relazioni di adiacenza, appartenenza e connessione. In topologia le primitive geometriche punto, linea e poligono divengono nodo, bordo e faccia (figura 5). La topologia può essere 2D e 3D. La prima trova da molti anni larga applicazione nel settore geografico, mentre la seconda stenta ad essere implementata a causa della difficoltà a produrre strumenti *software* efficienti per la sua gestione. Il dato *raster*, contrariamente a quello vettoriale, non contiene relazioni topologiche esplicite.

Terza dimensione

Nella descrizione della superficie terrestre, la quota è un dato irrinunciabile in quanto componente della collocazione spaziale degli oggetti ed entità fisica che ha un notevole impatto notevolmente sulle attività umane; come superficie di riferimento per la sua misura si è scelta quella equipotenziale della gravità (geoide). Nei *database* geografici, per rappresentare gli oggetti, si adoperano primitive geometriche 3D con i valori di est, nord e quota; ciò non è attuabile nella cartografia, in cui la terza dimensione si rappresenta con curve di livello (2D1/2) e punti quota (3D). In tal modo otteniamo una rappresentazione tridimensionale di tutti gli oggetti che popolano il *database*. Con queste metodologie da sole non riusciamo però a fornire informazione altimetrica in ogni punto e per ogni oggetto del DB. È per questo che si adottano forme complementari all'utilizzo della geometria 3D, con le quali si costruiscono superfici nello spazio che rappresentino nel modo più fedele possibile quella terrestre, fornendo informazione altimetrica in ogni punto, anche dove non siano presenti oggetti specifici; si tratta del *Digital Terrain Model* (DTM) e del *Triangular Irregular Network* (TIN) (figure 6.a, 6.b e 6.c).

Nei DB a grande risoluzione (grande scala) le primitive geometriche possono essere ampliate, ed oltre alle linee 3D possiamo avere superfici piane, oggetti solidi rappresentati mediante la *Constructive Solid Geometry* (CSG) (figura 6.d), che utilizza primitive geometriche spaziali a tre dimensioni quali cubi, parallelepipedi, coni, piramidi, cilindri, ecc., e solidi matematici ottenuti mediante funzioni matematiche. Si potrebbe in questo modo anche riuscire a rappresentare correttamente le unità immobiliari all'interno di un edificio.

Il DTM utilizza la forma *matrix* per esprimere un grigliato regolare di quote generato o manualmente da fotogrammetria, o da correlazione automatica di immagini, oppure frutto di interpolazioni di valori di quota derivanti da semine irregolari come profili altimetrici, punti quota, curve di livello o loro com-

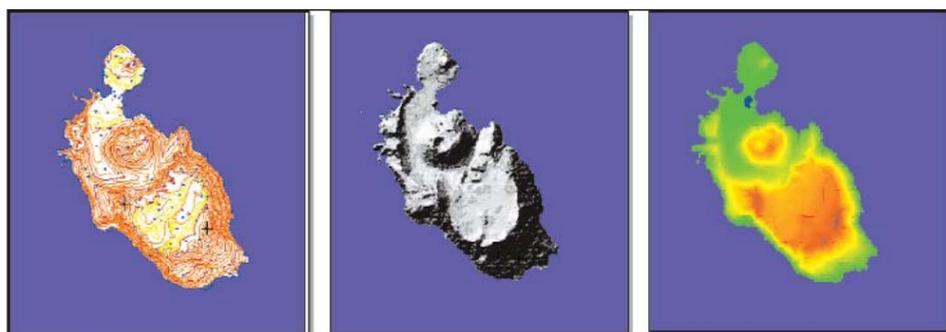


Figura 2.c. Altimetria: curve, DTM a sfumo e tinte ipsometriche (Isola di Vulcano).

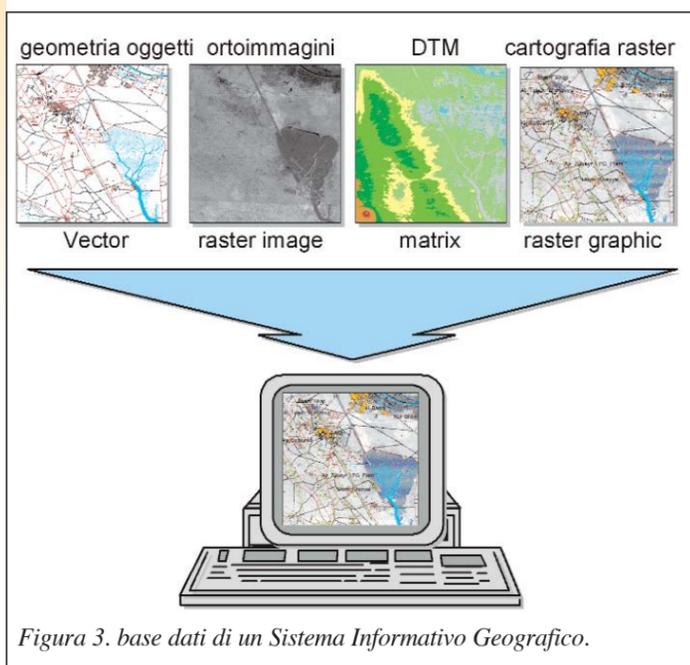


Figura 3. base dati di un Sistema Informativo Geografico.

binazione.

La rappresentazione altimetrica del terreno avviene quindi in forma discreta, paragonabile ad una serie di mattonelle contigue piane ed orizzontali di determinato lato (passo). Ogni mattonella rappresenta una porzione di superficie di terreno ed il valore di quota associato è comunemente la quota media di detta

superficie (figura 7). La quota di un qualsiasi punto della superficie terrestre è quella della mattonella nella quale esso planimetricamente ricade (metodo del «prossimo più vicino»), oppure viene individuata dalla intersezione della verticale per il punto con la superficie locale passante per un certo numero di punti di DTM contigui (con interpolazione). Il TIN è una ulteriore forma vettoriale per descrivere la superficie terrestre e si realizza tramite la costituzione di una serie discreta e continua di superfici, piane e non, a delimitazione triangolare, che si appoggiano su punti di quota nota.

Dal TIN la quota di un punto, di cui si conoscono solo le coordinate piane, si ottiene individuando la coordinata Z del punto di intersezione della normale passante per esso con la relativa superficie del triangolo di competenza.

Quarta dimensione

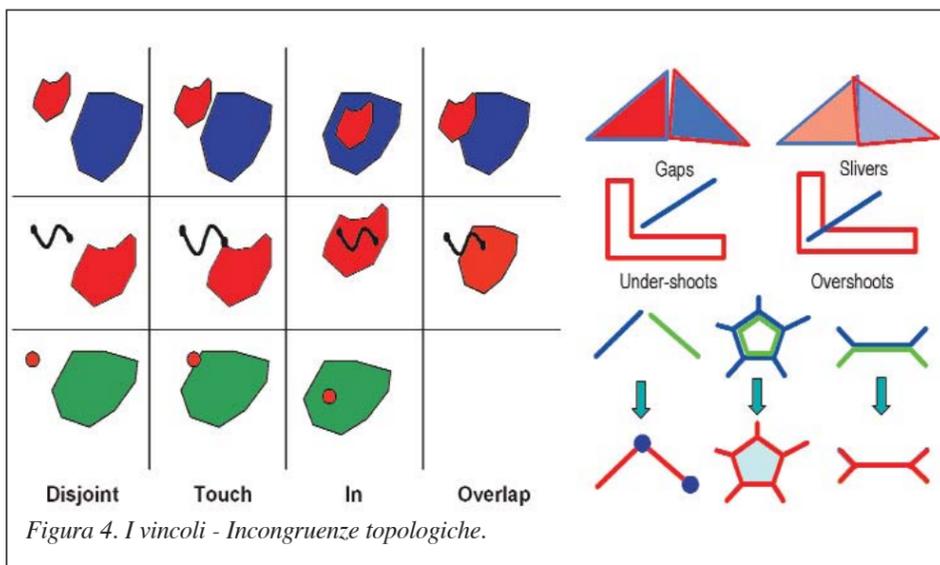


Figura 4. I vincoli - Incongruenze topologiche.

Le variazioni temporali dei fenomeni naturali ed antropici che insistono sulla superficie della Terra sono tradizionalmente documentate da cartografie redatte in epoche successive.

Gli *Spatial Temporal Database*, costituiti per gestire dati spazio-temporali, sono DB che contengono informazioni indicizzate in rapporto alla variabile tempo. Il tempo può essere rappresentato come una variabile a *step* o, meno comunemente, continua.

Attualmente, la maggior parte dei Sistemi Informativi Geografici (GIS) hanno una limitata capacità di immagazzinare e manipolare dati temporali.

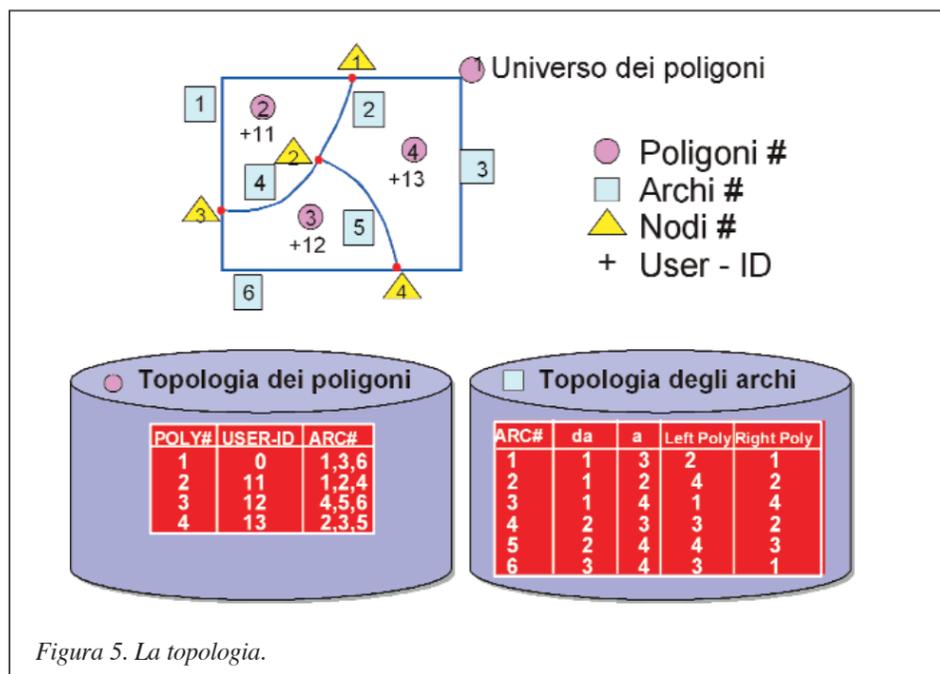


Figura 5. La topologia.

Toponomastica

Le denominazioni degli oggetti geografici (centri abitati, regioni geografiche, fiumi, case, ecc.) nel GeoDB sono contenute nell'attributo «nome» e sono interrogabili per effettuare ricerche. In fase di *display* di un DB, su monitor o nel plottaggio su carta, per facilitare l'analisi visiva, alcune denominazioni importanti sono presenti anche sotto forma di primitiva geometrica «testo» e quindi immediatamente leggibili.

Scala nei DB

Nei GeoDB il concetto di scala non ha un senso proprio e, quando è utilizzato, sta per «contenuto informativo ed accuratezza della corrispondente carta alla scala...».

Nella cartografia la scala sintetizza l'accuratezza posizionale (graficismo), la forma ed il livello di dettaglio e la densità informativa. Nei GeoDB, il parametro simile è la «risoluzione», che nel suo complesso qualifica il contenuto informativo del DB in termini quantitativi e qualitativi. Essa si distingue in risoluzione «geometrica» e «di contenuto».

Nel vettoriale la risoluzione geometrica è il grado di discriminazione fra oggetti limitrofi, o distanza minima distinguibile, e relativa soglia minima di cattura; quella di contenuto identifica la numerosità delle classi, e relativi attributi, delle entità geografiche rappresentate.

Nelle forme *raster* e *matrix*, la risoluzione geometrica si identifica con la dimensione del *pixel*. La risoluzione di contenuto nel *matrix* è data dalla numerosità delle classi utilizzate per discriminare gli oggetti; nel *raster* invece si distinguono quella radiometrica (numero di *bit* dedicati a descrivere il valore di radianza) e quella spettrale (larghezza di banda rilevata dal sensore).

Mentre la carta è nel contempo contenitore e supporto di presentazione delle informazioni geografiche, nel supporto informatizzato questi aspetti non sono più coincidenti ed il contenuto informativo non è più limitato, in termini di quantità e dovizia di dettagli sugli oggetti descritti, dall'impossibilità di poterli disegnare sulla carta.

I GeoDB, a qualsiasi risoluzione, sono considerati alla scala 1:1, in quanto il dato è riportato nel supporto di memorizzazione con l'accuratezza che la corrispondente metodologia di rilievo ha consentito. Nel «plottaggio» della componente geometrica il concetto di scala riprende il suo significato originale, e l'identificazione del suo valore corrispondente deve essere compatibile con la sua risoluzione geometrica. Ovvero non avrebbe senso «plottare» un dato geografico, solo perché numerico, ad una scala più grande di quella corrispondente alla risoluzione di acquisizione.

DB multiscala

Un GeoDB è multiscala, se comprende zone acquisite a risoluzione diversa. È tipico il caso di DB in cui le aree urbane e quelle extra urbane vengono rilevate in modo continuo ma a risoluzioni differenti, rispettivamente corrispondenti alla cartografia 1:2 000 e 1:5 000; possibilità, questa che la cartografia tradizionale non consente. I vantaggi di questo approccio evitano una doppia produzione e le incongruenze fra le rappresentazioni alle due risoluzioni.

GeoDB derivati

La necessità di rappresentare il territorio con diversa sintesi è stata soddisfatta con l'utilizzo della scale cartografiche, distinguendo le relative cartografie in «rilevate» e «derivate»: le prime ottenute con rilevamento diretto o telerilevate tramite fotogrammetria; le seconde prodotte utilizzando i dati di quelle rilevate o di quelle ottenute da precedenti derivazioni. Per i GeoDB si procede analogamente ed il concetto di «sintesi di rappresentazione» si traduce in «risoluzione».

a. Forma vettoriale: l'operazione di generalizzazione per portare il DB da una risoluzione ad un'altra più bassa può essere guidata da varie strategie: geometrica, attributo (appartenenza di classe), funzionalità, struttura.

Geometrica. L'operazione è effettuata per filtrare gli oggetti al di sotto di certe dimensioni e semplificare la geometria di quelli che transitano dalla risoluzione più alta a quella più bassa.

Attributo. Il criterio di generalizzazione può essere guidato dalla necessità di ridurre l'eccessiva risoluzione tematica che si ha quando gli oggetti geografici sono descritti con un dominio troppo ampio per la stessa classe di attributi.

Generalizzazione per funzione. La generalizzazione per funzione comporta la generazione di nuovi oggetti con un livello di aggregazione gerarchicamente più alto.

Generalizzazione per struttura. Prevede una semplificazione della struttura dell'oggetto geografico, lasciando inalterata la maggior parte della sua geometria e soprattutto la sua funzionalità, ma eliminando piccole parti inutili per la descrizione sintetica.

Normalmente, la generalizzazione geometrica è preceduta da quella di attributo. Nella derivazione dei *database* vettoriali di norma si adottano i seguenti operandi:

- selezione (filtro «passa non passa»): si individuano e si eliminano gli oggetti che non sono contemplati nel contenuto informativo del *database* derivato;
- inglobamento: gli oggetti con dimensioni sotto il limite di soglia di cattura vengono inglobati da quelli limitrofi più grandi;
- sfoltimento: un certo numero di oggetti di uguale tipologia viene eliminato per limitarne la densità;
- fusione: più oggetti di uguale tipologia, ma con diverso attributo, si uniscono a formarne di nuovi;
- aggregazione: più oggetti semplici di uguale o diversa tipologia perdono individualità e si uniscono a formare un oggetto semplice o composto;
- riclassificazione: agli oggetti sottoposti a fusione e/o aggregazione vengono ridefiniti, mediante calcolo, gli attributi;
- semplificazione: la componente geometrica (forma dell'oggetto geografico) viene semplificata eliminando un certo numero di vertici nelle poligonali, nei poligoni e nelle curve (*squaring*, *smoothing* etc.), fino a giungere al collasso di alcuni oggetti, che da areali si trasformano in lineari o puntuali e da lineari in puntuali.
- ricostruzione delle congruenze (geometriche e logiche) e della topologia: la tessitura della componente geometrica degli oggetti superstiti dopo la derivazione viene ricostruita laddove le operazioni di spoglio, aggregazione e generalizzazione hanno lasciato vuoti e sconessioni.

b. Forma *raster*: operare la generalizzazione significa diminuire la risoluzione aumentando la dimensione del *pixel*, e calcolare i valori *radiometrici* da associare ai nuovi *pixels*. Questo processo avviene con il ricampionamento e si attua molto agevolmente attraverso vari algoritmi (convoluzione cubica, prossimo più vicino, ecc.).

c. Forma *matrix*: si realizza con l'aumento della dimensione del *pixel* e/o diminuendo il numero di *classi* che descrivono il fenomeno (DTM, uso e copertura del suolo, ecc.).

Mentre le operazioni di derivazione di DB *raster* e *matrix* sono essenzialmente automatiche, una volta definito l'algoritmo di ricampionamento, in quelle vettoriali la definizione e l'applicazione degli algoritmi di generalizzazione sono regolate ancora da forti interventi interattivi da parte dell'operatore, data la vastità e complessità delle regole e delle casistiche. Importanti studi sono in corso in molti centri di ricerca sparsi in tutto il mondo per applicare a questo settore, con l'aiuto di potenti calcolatori, l'intelligenza artificiale.

Dalla cartografia numerica al GeoDB

Il passaggio dalla cartografia analogico/digitale al GeoDB è stato un processo evolutivo, durato molti anni, che ci ha consentito la costituzione dei Sistemi Informativi Geografici con capacità di elaborazione spaziale dei dati oltre che di rappresentazione cartografica.

a. Automazione cartografica

Una delle prime applicazioni dell'informatica nel campo della cartografia ha visto l'utilizzo di *software* di editor grafici (CAD - Computer Aided Drafting / Design) per la fase di compilazione/disegno della carta, e di *plotters* (disegnatori automatici) per la restituzione; si è così realizzata la cosiddetta «automazione cartografica». In pratica, ci si è serviti di strumenti informatizzati per produrre la cartografia a tratto su carta nella forma classica. I dati geografici numerici così acquisiti hanno una struttura semplice, finalizzata principalmente alla sola riproduzione grafica automatizzata.

b. Cartografia numerica

La dicitura «cartografia numerica» viene utilizzata per identificare una raccolta di dati geografici in forma digitale, finalizzata al disegno automatico ed alla gestione interattiva, attuata con metodologie informatiche, in cui è generalmente presente solamente un *file* grafico dove le informazioni associate agli oggetti rappresentati (es. larghezza, numero di corsie, fondo delle strade) vengono esplicitate attraverso l'uso degli attributi grafici (livello, colore, spessore, tipo di linea, ecc.). Nella cartografia numerica non è presente la topologia.

c. Database geografici

Con il passare del tempo, alla cartografia numerica è stata richiesta la possibilità di supportare applicazioni tipiche di un sistema informativo, si è giunti in tal modo alla creazione del Sistema Informativo Geografico. Transitoriamente, in attesa della generazione diretta dei DB, si è proceduto, con pesanti rielaborazioni, alla conversione della vecchia cartografia numerica.

I primi database geografici si basavano sulla struttura dei *File system*, ma le limitazioni di questo approccio hanno portato a pensare a strutture vere e proprie di *database* corredati di schemi per la definizione del contenuto informativo (*Definizione dei Dati* - DLL), di linguaggi di interrogazione standardizzati (*Structured Query Language* - SQL) e *software* di gestione *Data Base Management Systems* (DBMS) con strutture diverse (gerarchica, reticolare, relazionale). Sono stati realizzati *database* con file grafici per la componente geometrica e *database* per la componente delle informazioni associate (attributi) o *database* basati sulla filosofia *Object Oriented* (OO), in cui le due componenti sono integrate. La struttura attualmente più utilizzata è quella relazionale (*Relational Data Base* - RDB), gestita dal relativo *software Relational Data Base Management Systems* (RDBMS). Esistono due tipologie

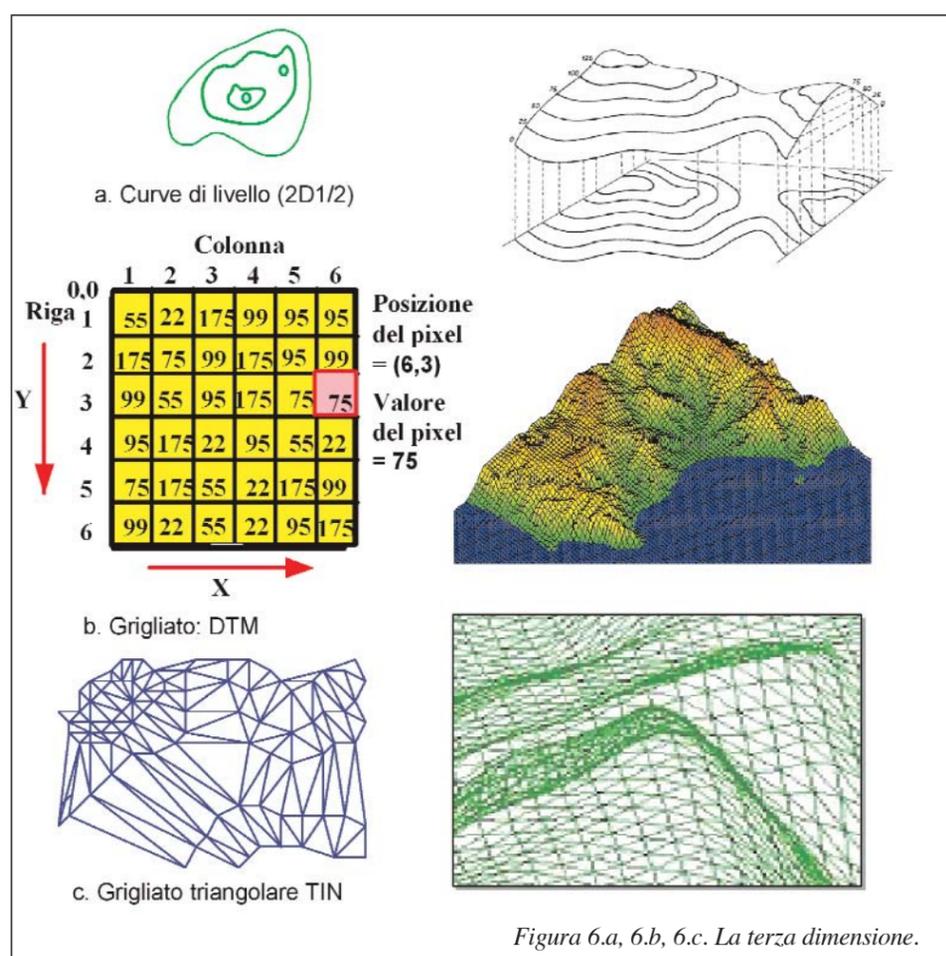


Figura 6.a, 6.b, 6.c. La terza dimensione.

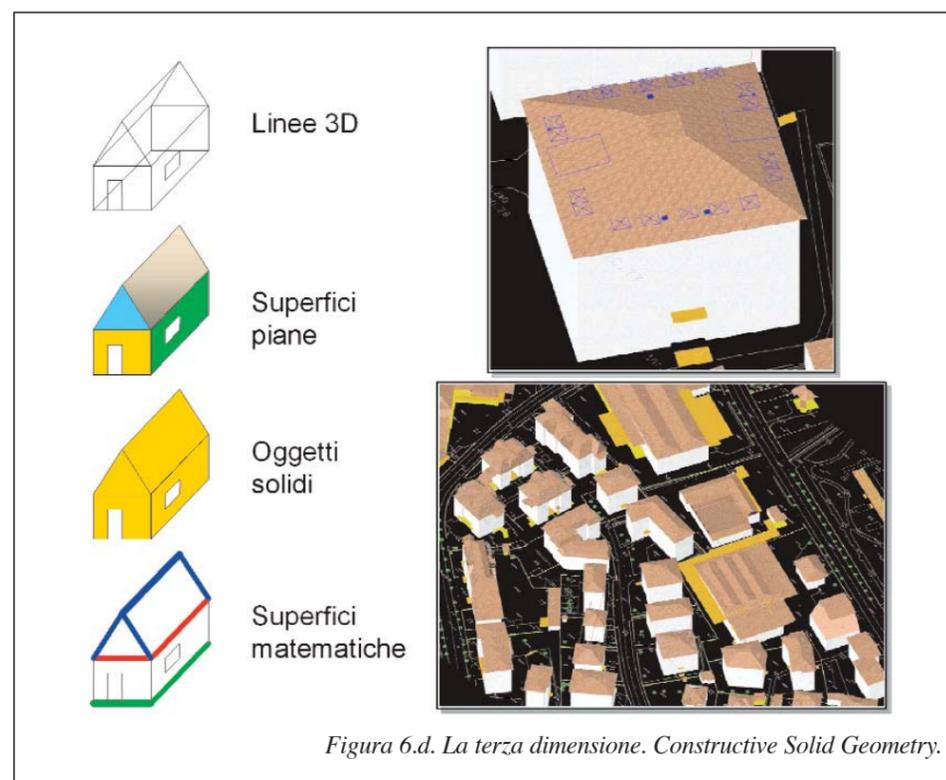


Figura 6.d. La terza dimensione. Constructive Solid Geometry.

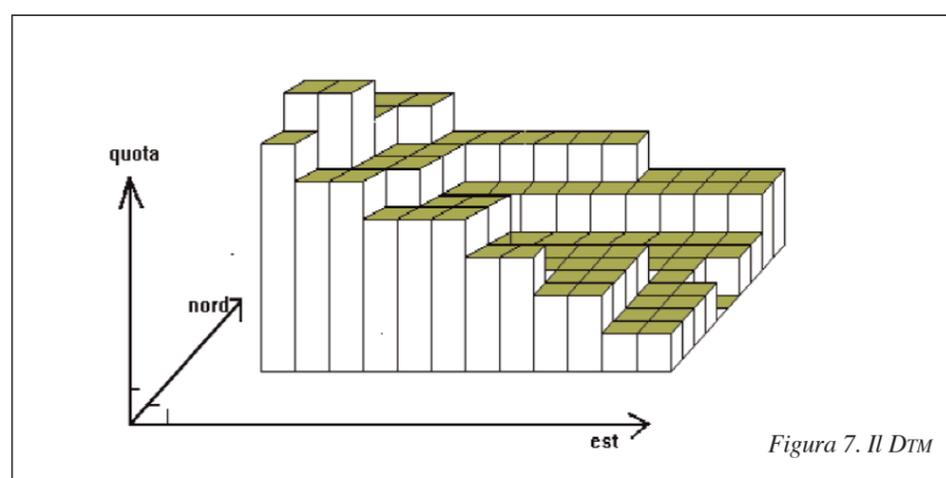


Figura 7. Il DTM

di RDB: relazionali puri, con incluso il campo coordinate, e relazionali con campi particolari denominati *Binary Large Object* (BLOB), contenenti la componente spaziale (coordinate).

Qualità

I dati geografici in forma digitale, per le loro peculiari caratteristiche e modalità di utilizzo, necessitano di una descrizione della qualità più puntuale e sistematica rispetto a quella utilizzata per la cartografia analogica. L'aumento della diffusione dei dati geografici su supporto informatizzato ne ha esteso l'utilizzo in numerose discipline ed accade quindi che essi vengano impiegati in usi diversi da quelli per i quali erano stati acquisiti. In modo particolare, per i dati ufficiali utilizzati nella Pubblica Amministrazione è indispensabile che le indicazioni di qualità, riportate nel *metadata*, diventino parte integrante del dato stesso e vi siano delle procedure standardizzate per la loro determinazione e riporto. Di qui la necessità di

descrivere la qualità di un *dataset* in modo da permettere la sua valutazione a soggetti diversi.

I parametri di valutazione della qualità sono:

- completezza: descrive la presenza o assenza di oggetti, attributi, relazioni;
- accuratezza posizionale: quantifica la differenza fra la posizione plano-altimetrica del particolare topografico descritto nel *dataset* e la sua posizione vera;
- accuratezza temporale: descrive l'accuratezza degli attributi e delle relazioni temporali;
- accuratezza tematica: descrive l'accuratezza degli attributi qualitativi e quantitativi e la correttezza delle classificazioni;
- congruenza: descrive il grado di aderenza alle regole di attribuzione, di relazione e di struttura dei dati; la struttura dei dati riguarda la componente concettuale, fisica e logica;
- fedeltà dei testi: descrive la correttezza dei testi cartografici (toponomastica).

Nel caso si debba valutare la qualità di un *dataset* già prodotto, in assenza dei parametri di qualità di progetto, si ricorre alla valutazione indiretta della qualità tramite il suo «lignaggio», termine con il quale si indica la genealogia e la storia dei dati costituenti il *database*. Anche se non è un vero e proprio parametro di qualità, fornisce una valutazione globale del *dataset*; esso prevede di esaminare:

- la fonte dei dati (da rilievo diretto, da processo fotogrammetrico, da digitalizzazione di cartografia preesistente, da derivazione da *database* a più grande risoluzione);
- il processo di produzione;
- le specifiche tecniche di acquisizione;
- l'epoca di produzione;
- lo scopo originale dei dati;
- il produttore;
- la storia del trattamento ed aggiornamento dei dati (chi, come, quando).

Standard e organismi di standardizzazione

Scopo fondamentale degli *standards* nel settore geografico è quello di trasportare un GeoDB da un sistema, formato da una determinata configurazione *hardware/software*, ad un altro sistema, comunque configurato, senza perdita di informazione; ovvero di accedere a *database* geografici ed utilizzarli via rete telematica in modo *System and Media Independent*, per permettere l'interoperabilità. Ciò richiede l'utilizzo di formati *standard* o linguaggi *standard* di comunicazione.

Tra le numerose entità, nazionali e sopranazionali, che hanno sviluppato tali *standards* per i GeoDB si segnalano:

- il Comitato Europeo di Normalizzazione 287 (CEN/TC 287), che ha prodotto gli *standard ENV*;
- l'*International Standardization Organization/Technical Committee Geographic Information-Geomatics* (ISO/TC 211), che ha realizzato i documenti della serie 191xx;
- l'*International Hydrographic Organization* (IHO), che ha prodotto lo *standard S-57* per la realizzazione dell'*Electronic Chart Display and Information System* (ECDIS);
- il *Digital Geographic Information Working Group* (DGIWG), che ha prodotto il *Digital Geographic Exchange Standard* (DIGEST). Il DGIWG è stato istituito nel 1983 per sviluppare uno *standard* per facilitare lo scambio di dati geografici e sostenere la interoperabilità fra nazioni, e condivide progetti di produzione dei dati digitali sulle aree di comune interesse.

Una menzione particolare richiede l'ISO/TC 211, che ha affrontato in modo completo tutti gli aspetti connessi con l'informazione geografica digitale e costituisce un riferimento per tutti gli ambiti che questo settore abbraccia. Il lavoro è stato svolto in cooperazione con i principali enti internazionali di standardizzazione del settore geografico e con l'*Open GIS Consortium* (OGC), formato da aziende operanti nel settore geomatico.

Formati standard di trasferimento

I più importanti sono:

- Vector Product Format* (VPF), *Feature Attribute Coding Catalogue* (FACC), *Arc Digital Raster Graphic* (ADRG), *UTM Standard Raster Graphics* (USRG), *Digital Terrain Elevation Model* (DTED), appartenenti alla famiglia DIGEST;
- National Transfer Format* (NTF), *Geographic Tagged Image File Format* (GEO TIFF);
- Geographic Data File* (GDF). È uno *standard* nato per lo scambio di *database* digitali sulla viabilità orientato alle problematiche della *car-navigazione*; esso però è divenuto più di un semplice *standard* per lo scambio dei dati, in quanto permette anche una loro interrogazione diretta. Lo sviluppo del GDF è cominciato nel 1984 nel progetto *Demeter*, è continuato in diversi altri progetti europei, ed è stato portato avanti nel 1994 anche dal CEN/TC 278 e dal *Road Transport and Traffic Telematics* (RTTT).

Linguaggi standard di comunicazione

Negli ultimi anni, sotto la spinta evolutiva di Internet, si stanno afferman-

do i linguaggi *standard* di comunicazione; infatti, tutte le transazioni (richieste e risposte) che un *software* di navigazione (*browser*) esegue sono formulate unicamente mediante il *Transmission Control Protocol/Internet Protocol* (TCP/IP), un protocollo di dialogo fra *computers* per lo scambio di dati, utilizzato nella rete delle reti Internet.

Gli *standards* di comunicazione più diffusi sono:

- Hypertext*: è un termine coniato nel 1965 per indicare una collezione di documenti contenenti riferimenti incrociati (*cross-reference*) o *link* che permettono ad un programma interattivo (*browser*) di leggere e muoversi da un documento all'altro;
- Hypertext Markup Language* (HTML): è l'ipertesto *standard* utilizzato nel *World-Wide Web* (Www);
- Extensible Markup Language* (XML): è un *markup language* per documenti che contiene informazioni strutturate, cioè un meccanismo per definire strutture in un documento. Le informazioni strutturate contengono indicazioni sul ruolo del contenuto. È un linguaggio di Internet e dell'interscambio in generale. L'XML non è una replica dell'HTML; infatti il primo è orientato alla descrizione delle informazioni e il secondo alla loro visualizzazione;
- Unified Modeling Language* (UML): è un linguaggio «non proprietario» di terza generazione, usato per specificare e visualizzare costrutti e documenti. L'UML rappresenta un compendio delle migliori pratiche dell'ingegneria, che ha avuto successo nella modellazione di sistemi grandi e complessi;
- Geo Unified Modeling Language* (GEOUML): è l'applicazione alle informazioni geografiche del «*General Feature Model*» dell'ISO/TC 211;
- Geography Markup Language* (GML): è una grammatica XML scritta, come uno schema XML, per la modellazione, il trasferimento e la memorizzazione di informazioni spaziali. Il GML è l'XML geografico. È possibile implementare il GML o per immagazzinare dati geografici con il relativo schema applicativo, o per trasferirli da un sistema ad un altro. Il GML potrebbe utilmente essere utilizzato come sistema per lo scambio di dati fra produttori ed utilizzatori forniti entrambi di interprete GML.

La versione 3.0 del GML (GML3) applicata ai dati geografici è un linguaggio in grado di comunicare geometria e topologia 3D, sistema di riferimento, unità di misura, dizionario, componente temporale, modello dei dati. In GML3 è implementabile un modello ad oggetti flessibile ed estensibile, che è in grado di trattare le relazioni arbitrarie fra oggetti, le proprietà multiple della geometria e quelle della topologia.

Standard di Interrogazione

Il *Query Language* è un modo di comunicare i comandi di manipolazione ad un *database*. Il linguaggio standardizzato comunemente utilizzato per effettuare manipolazioni ed interrogazioni nei *database* relazionali è lo *Structured Query Language* (SQL). Sviluppato dall'IBM negli anni Settanta, è divenuto uno *standard* industriale utilizzato universalmente nelle strutture (sistemi) relazionali di DB.

Dati di progetto di un GeoDB

Le caratteristiche di un GeoDB sono: la forma, il contenuto informativo, il sistema di riferimento geodetico-cartografico, la strutturazione dei particolari geografici e le loro proprietà, gli indicatori della qualità, il sistema di codifica, il formato di scambio. Le fasi di progetto sono: la definizione del modello di astrazione, del modello di rappresentazione, dello schema e modello concettuale, del modello di comunicazione/trasferimento.

Operazioni sui Database geografici

Tipiche operazioni che si eseguono sui *database* geografici sono le «selezioni» (*Spatial Query*) e le «analisi spaziali» (*Spatial Analysis*), in cui gli oggetti del *database* vengono individuati per caratteristica, posizione o mutua relazione. Le possibili strutture di *database* geografici sono: gerarchica, reticolare e relazionale. Quest'ultima è la più diffusa.

Geomatica

Il connubio cartografia-informatica ha generato la «geomatica», definita come l'approccio sistematico integrato multidisciplinare per selezionare gli strumenti e le tecniche appropriate per acquisire, integrare, trattare, analizzare, archiviare e distribuire dati spaziali georeferenziati, relativi alla superficie terrestre, per la formazione di Sistemi Informativi Territoriali.

Infrastruttura di dati geografici

Si compone di GeoDB collegati a servizi ed applicazioni, disponibili sempre ed ovunque, come strumenti per scoprire, utilizzare, integrare e condividere dati georeferenziati. È sempre più diffusa la disponibilità di varie tipologie di *database* geografici (di base e tematici) distribuiti geograficamente (*data server*), detenuti, accessibili ed utilizzabili via rete telematica, aggiornati e gestiti da soggetti istituzionali e da aziende specializzate.