

Corso di Aggiornamento in DB Topografici

Progettazione e costruzione del DB secondo le specifiche Intesa

Claudio Rocchini

Istituto Geografico Militare

Introduzione

- Questa parte dell'esercitazione riguarda la costruzione del geo-db in formato finale.
- Si potrebbe suddividere in due sottoparti:
 - La progettazione delle specifiche.
 - La realizzazione fisica a partire dalle specifiche.
- Nel nostro caso la progettazione è (quasi del tutto) già stata eseguita dall'Intesa, per cui non ci resta che realizzare le specifiche.
- In realtà vedremo come ci siano ancora delle scelte da prendere durante la realizzazione (es. nomi attributi).
- Inoltre vedremo alcuni punti chiave della progettazione, ed alcuni esempi di progetti di specifiche.



Progettazione 1: Feature Class

- Il primo passo è quello di classificare tutte le possibili entità della carta in classi di oggetti omogenee, che spesso vengono chiamate “feature-class” (= classi di caratteristiche), abbreviato poi nella sola parola “Features”, anche se invece il concetto fondamentale è quello di classe.
- Esempi di feature-class sono: ponti, aree boschive, fiumi, etc.
- L’intera classe (es. tutti i ponti) non deve essere confusa con il singolo oggetto o istanza della classe (es. il ponte a Verrazano), che spesso viene anche questo chiamato feature (per non sbagliare, in questi lucidi feature-class è sempre scritto esteso).
- A livello fisico, di solito, la feature-class corrisponde esattamente ad una struttura informatica precisa: il file shape, una tabella del personal geodb, eccetera.
- La suddivisione in feature-class degli oggetti della carta è un passo molto importante, perché comporta una forte interazione con l’utilizzo del dato (ad esempio le regole topologiche si impostano per feature).
- Il livello di raggruppamento delle feature-class (quante classi definisco?) è un fattore importante, anche se i sistemi più evoluti (come i db cartografici) possono simulare l’esistenza di sotto-feature (query con filtri sugli attributi), o di super-feature (query unione di tabelle, più rara).



Feature-Class: elementi identificativi

- Un elemento importante della feature-class è il suo nome, esempi:
 - CODICI INTERNI IGM “LAB”: L307 (mulattiera), X431C (sinagoga), P354C (ponte in ferro per strade di categoria A3/A4, con luce minore di 25 metri).
 - DB25-VPF: vegeta, tubol, pontep, stradal, ...
 - DIGEST: AP030(strade), BH140 (fiumi) AL015 (edifici), ...
 - REGIONE TOSCANA 10k: AV, AF, EA, PQ, ...
 - INTESA: 010100, 090301, 070405, ...
- Molto spesso i sistemi cartografici impongono di associare in modo univoco alla feature-class un tipo geometrico (punto,linea,area); così non è però per le specifiche intesa, in cui il tipo geometrico dovrebbe essere un attributo: per realizzarlo nei sistemi reali è necessario a volte creare più feature per la stessa classe, con nomi simili e geometrie diverse (F020301L = ponte lineare, F020301P = ponte puntuale).



Progettazione 2: Livelli di gerarchia delle feature-class

- Molto spesso, la collezione delle feature-class è raggruppata in un qualche sistema gerarchico.
- La strutturazione gerarchica può essere realizzata esplicitamente nel supporto, se lo prevede il software corrispondente, (impossibile con file shape, feature-dataset e sottoclassi nei personal-geodb), oppure implicitamente nelle specifiche.
- Non bisogna confondere la strutturazione possibile in fase di progetto e di visualizzazione (es. layer di Arcgis), dalla strutturazione presente nei dati (es. sottoclassi): solo quest'ultima deve essere presa in considerazione perché esportabile senza ambiguità.
- I vari livelli di gerarchia, a parità di funzionalità hanno purtroppo vari nomi, a seconda dell'utilizzatore, dall'estrazione culturale di quest'ultimo, dal software utilizzato (es. Tematismo, Strato Tematico, Coverage, Layer, Livello, etc.)
- La regione toscana (shape 10k) ad esempio, non ha una struttura gerarchica.



Esempi di Gerarchie: DB25 IGM (Internal)

- La gerarchia delle specifiche prevede 3 livelli:
 - Tematismo (o Coverage o Layer)
 - Classi (feature-class)
 - Sottoclassi (sub feature-class)
- I tematismi sono: *Altimetria, Confini, Geomorfologia, Idrografia, Industrie, Insedimenti, Servizi, Toponomastica, Trasporti, Vegetazione, Simboli* (curiosamente la toponomastica è un tematismo che contiene tutte le classi toponimo degli altri tematismi, il tematismo simboli colleziona tutti gli oggetti geometrici che non corrispondono ad oggetti del db vero e proprio, ma sono stati creati ai fini di stampa: bordo carta, maschere grasse per simboli di croce in abitato, etc.).
- Il dato interno IGM è rappresentato di solito con file CAD (+ database associato) o personal geodatabase, per questo non è possibile rappresentare esplicitamente i tematismi. Le sotto-classi sono definite utilizzando gli attributi (attributo codificato “LAB”): ogni sottoclasse corrisponde esattamente ad uno stile di stampa (quindi nasce dalla stampa, non dalle esigenze della base di dati). Solo le classi sono quindi rappresentate in modo diretto.

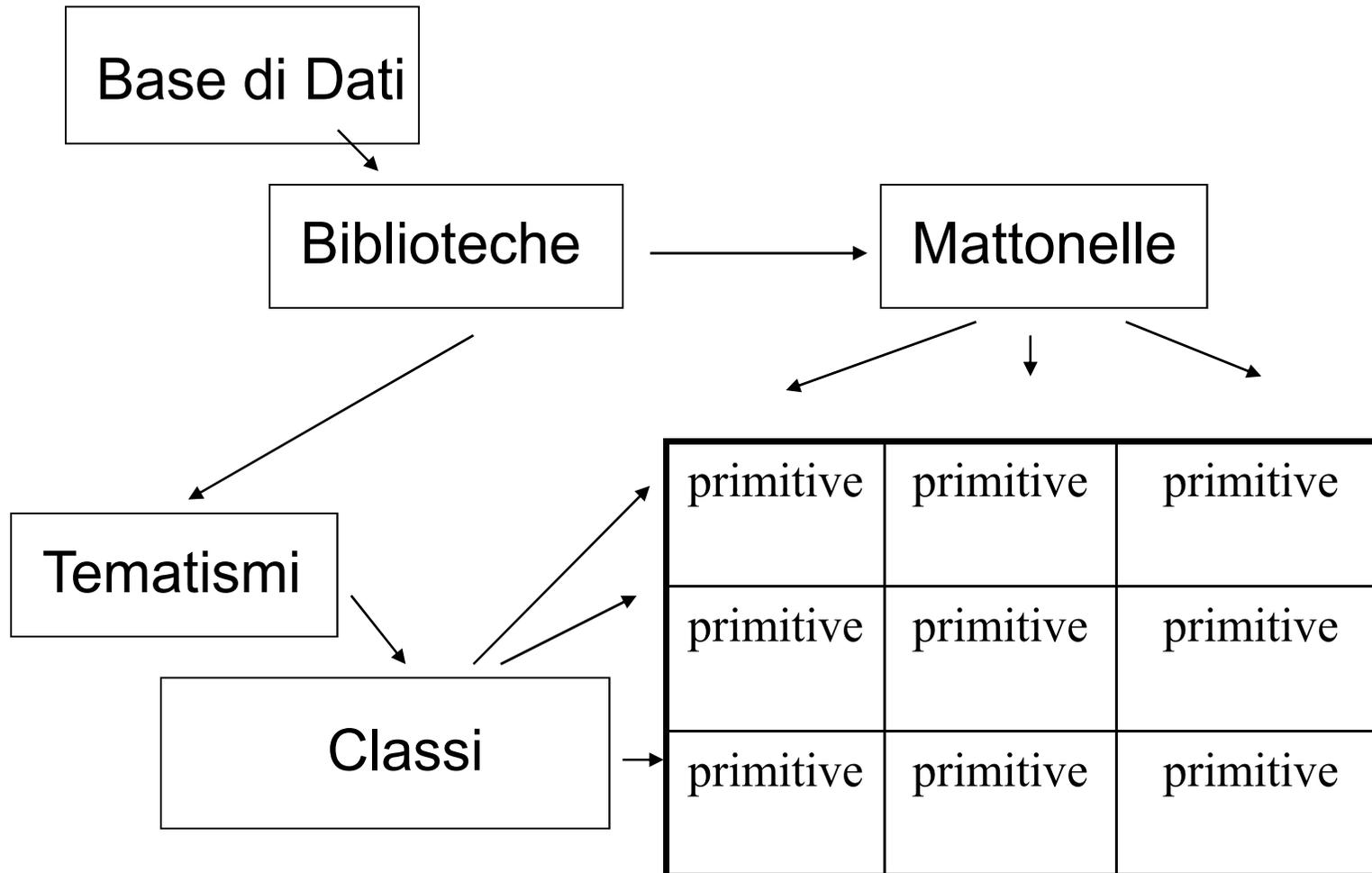


Esempi: DB25 VPF (1/2)

- Nota: il dato finale IGM è venduto in formato VPF.
- Questo formato permette di creare una gerarchia esplicita di notevole complessità.
- La gerarchia è sia:
 - spaziale (divisione in mattonelle e librerie) che
 - gerarchica per contenuto (Tematismi e Classi).



Inciso: Gerarchia VPF



Esempi: DB25 VPF (2/2)

- L'unico database prodotto, contiene l'intera Italia.
- Le librerie corrispondono al taglio del 1:50K (ma in WGS non in ED50, come la cartografia cartacea dell'ultima versione).
- Le mattonelle (tiles) corrispondono al taglio del 1:25K.
- I Tematismi sono: *Altimetria, Confini, Geomorfologia, Idrografia, Industrie, Insedimenti, Quadro di Unione, Qualità, Servizi, Trasporti, Vegetazione, LIBREF, TILEREF* (gli ultimi due sono tematismi di sistema del VPF). Si noti il tematismo qualità, il quale contiene aree che descrivono i meta-dati di acquisizione.



Esempi: Specifiche Intesa

- Le specifiche intesa prevedono una gerarchia a 3 livelli:
 - Strati
 - Temi
 - Classi
- Le specifiche non si appoggiano (giustamente) ad una particolare struttura fisica, quindi non si può parlare di rappresentazione esplicita della gerarchia.
- I nomi codificati degli strati, dei temi, delle classi e anche degli attributi, seguono però uno schema gerarchico, per cui in questo caso si può dire che la gerarchia è esplicitamente codificata nei nomi degli oggetti.
- Ad esempio 04 è l'idrografia, 0404 è il reticolo idrografico, 040404 è un corso d'acqua naturale.



Effetti collaterali della gerarchia esplicita sull'utilizzo dei dati.

- Spesso la realizzazione esplicita di una certa gerarchia, per mezzo di uno strumento software, comporta dei vincoli sull'utilizzo dei dati.
- Esempio: sia nel VPF che nel personal geodatabase di ArcGis, le regole topologiche possono essere definite solo all'interno dello stesso feature-dataset (ArcGis) o dello stesso tematismo (VPF).
- In particolare il livello topologico del VPF (spaghetti, grafo, grafo planare, full) è selezionabile solo per tematismo (tutte le classi del tematismo hanno lo stesso livello topologico).



Progettazione 3: Attributi

- Il terzo passo fondamentale per la costituzione del gdb è la definizione degli attributi.
- Gli attributi sono definiti, come minimo dal nome e dal tipo di dato (stringa, numero intero, numero con virgola, etc.).
- Gli attributi sono spesso sottovalutati: la loro stabilità nelle specifiche è invece fondamentale. Un cambio di nome di un attributo, ad esempio, può distruggere il funzionamento di tutti i vostri modelli costruiti con Model-Builder o degli script di manipolazione.
- Piccole discrepanze nei nomi o nei tipi degli attributi possono inficiare anche operazioni semplici come l'”Append” di due feature.



Progettazione 4: domini e vincoli sugli attributi

- Definire il nome ed il tipo di un attributo è il minimo che si possa fare.
- Un passo più evoluto è la definizione di vincoli di pertinenza degli attributi: questi vincoli, se implementati esplicitamente nel sistema, permettono un controllo a priori della correttezza dei dati.
- I database professionali (come Oracle Spatial), permettono una notevole espressività per la definizione dei vincoli sui dati degli attributi.
- Una possibilità prevista in molti sistemi è quella di definire dei domini di validità degli attributi (es. Domini di ArcGis), che possono essere codificati (un'enumerazione di valori) o compresi in intervalli.
- Sia le specifiche Digest che le specifiche Intesa-Stato prevedono la definizione di domini codificati per quasi tutti gli attributi.



Esempio di trigger (Dal sistema amministrativo IGM)

```
CREATE TRIGGER assenze_trigger AFTER INSERT OR UPDATE ON ASSENZE
SELECT INTO x EXISTS( SELECT *
  -- guardo che non esiste un intervallo...
  FROM assenze
  WHERE dipendente = NEW.dipendente
        -- Del dipendente
        AND id      <> NEW.id
        -- non conto quello che sto' modificando!
        AND (fine  >= NEW.inizio OR fine IS NULL)
        -- che toccano quello inserito
        AND (inizio <= NEW.fine  OR NEW.fine IS NULL) );
IF x THEN
  RAISE EXCEPTION ;
END IF; END;
```

- Che in italiano significa: posso assegnare un periodo di ferie, se il dipendente in questione, per quel periodo, non è già in ferie, altrimenti gli segno due giorni di ferie per lo stesso giorno e lo frego.



Note sui vincoli degli attributi

- La specifica dei vincoli sugli attributi può essere molto complessa, ad esempio un vincolo di un attributo può dipendere dal valore di un altro attributo (linea ferroviaria: attributo tipo = non elettrificata, l'eventuale attributo voltaggio deve essere NULL).
- Se il sistema in uso non prevede la specifica di vincoli complessi a priori, si possono fare sempre dei controlli a posteriori tramite query SQL sui dati. Esempio di query IGM sulle curve di livello:
- **lab='L902' AND (val(crv) mod 25 <> 0 OR val(crv) mod 100=0);**
- Che in italiano vuol dire: le curve di livello intermedie devono avere un'altezza che è un multiplo di 25 ma non un multiplo di 100 (altrimenti sarebbero curve principali). La query è scritta al contrario, vale a dire che come risultato ritorna i valori errati.



Progettazione 4: Metadati

- La progettazione della base di dati e la sua realizzazione devono prevedere la memorizzazione di metadati (dati sui dati).
- I metadati, nella loro realizzazione esplicita possono essere:
 - Codificati con gli strumenti previsti dagli strumenti software (Es. i metadati di ArcGigs, a cui vanno aggiunti gli alias delle feature e degli attributi, le note testo o descrizioni che si possono aggiungere ai vari oggetti).
 - Memorizzati con speciali feature-class geometriche (tipicamente aree) appositamente previste. In questo caso sono realizzate praticamente con gli strumenti standard di tutti i software (ad esempio con file shape standard, in file CAD, etc.). I metadati quindi diventano attributi delle feature geometriche speciali. E' questo il caso delle feature areale Intesa, (Area di definizione della scala, aree di pertinenza), e anche delle specifiche DB25 VPF (tematismi LIBREF, TILEREF, feature-class QUALITA').



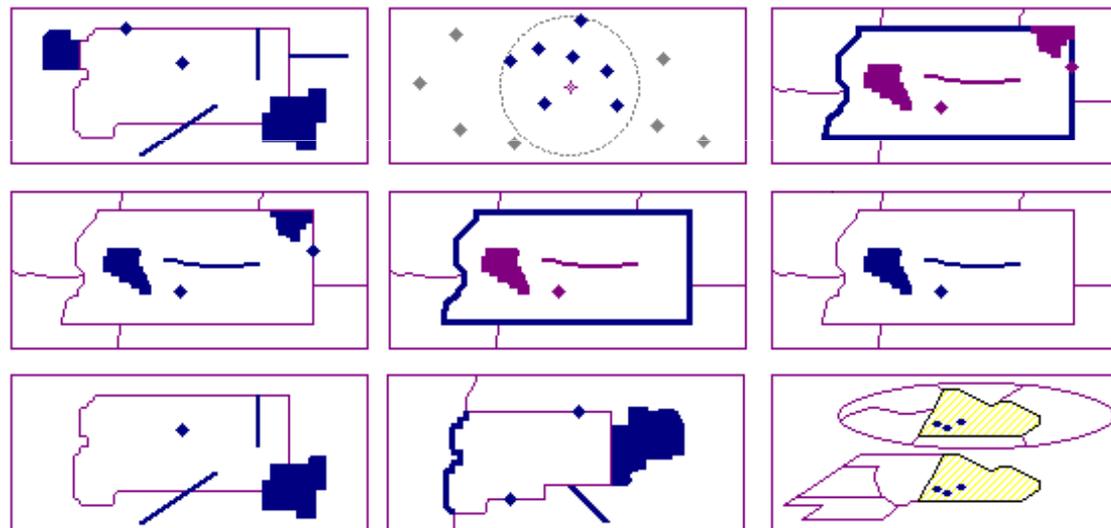
Progettazione 5: Vincoli topologici

- La realizzazione di un gdb fisico può prevedere la realizzazione fisica di vincoli topologici.
- La definizione dei vincoli può essere molto dipendente dall'utilizzo dei dati, quindi può essere lasciata non specificata (es. Intesa).
- In questa campo ancora una volta, il terreno non è ben assestato: i vari sistemi gestiscono i controlli in modo molto diverso l'uno dall'altro (oppure non li gestiscono affatto).
- I controlli possono essere definiti esplicitamente e quindi controllati a priori, oppure possono essere impostati solo in fase di controllo e quindi a posteriori.
- Vedremo in dettaglio la gestione dei vincoli topologici di ArcGis, durante l'esercitazione.



Esempio: Software Intergraph

- E' possibile definire dei controlli a posteriori, utilizzando i seguenti costrutti.



Tocca

Entro la distanza

Contiene

È Contenuto

Contiene Strettamente

E' contenuto strettamente

Si sovrappone

Incontra

Equivale



Vincoli Topologici in IGM

- I capitolati di appalto della cartografia non citano esplicitamente le regole topologiche da rispettare (se non con brevi frasi testuali).
- La restituzione fotogrammetrica e la produzione dei dati di stampa non prevedere l'applicazione di regole topologiche.
- In realtà questa parte della progettazione è ancora in fase semi-sperimentale.
- La produzione del VPF invece impone forzatamente il rispetto del livello topologico impostato all'interno del tematismo (per la sua struttura interna, è impossibile realizzare un dato VPF scorretto topologicamente).
- Il servizio GIS ha steso una serie di vincoli topologici da controllare a posteriori. I vincoli sono specifici di ogni singola feature, ed entrano nel merito della semantica dei dati. Una prima lista approssimativa contiene 3.000 regole topologiche (Autore: Cartografo Fortini Claudio).



Costruzione e Rappresentazione delle specifiche

- Un modo di rappresentare le specifiche è quello di scriverle su di un pezzo di carta.
- Poi ci potrebbero essere una serie di strumenti software che potrebbero essere d'ausilio alla stesura delle specifiche:
 - Visio: progettazione e realizzazione di database standard (senza parte cartografica).
 - ArgoUML ed altri: progettazione di specifiche UML.
- Tutti i condizionali del punto precedente derivano dalla ricerca effettuata per la realizzazione di questo corso: al momento attuale gli strumenti sono in fase di sviluppo e non ancora pronti all'uso, che permettano sia la modellazione del progetto, sia la sua realizzazione automatica.
- Uno strumento importante è comunque il Model-Builder di ArcGis, che sebbene chiuso nell'ambiente ESRI, permette di creare dei modelli di specifica, per la creazione del DB, compresi i domini ed i vincoli topologici.

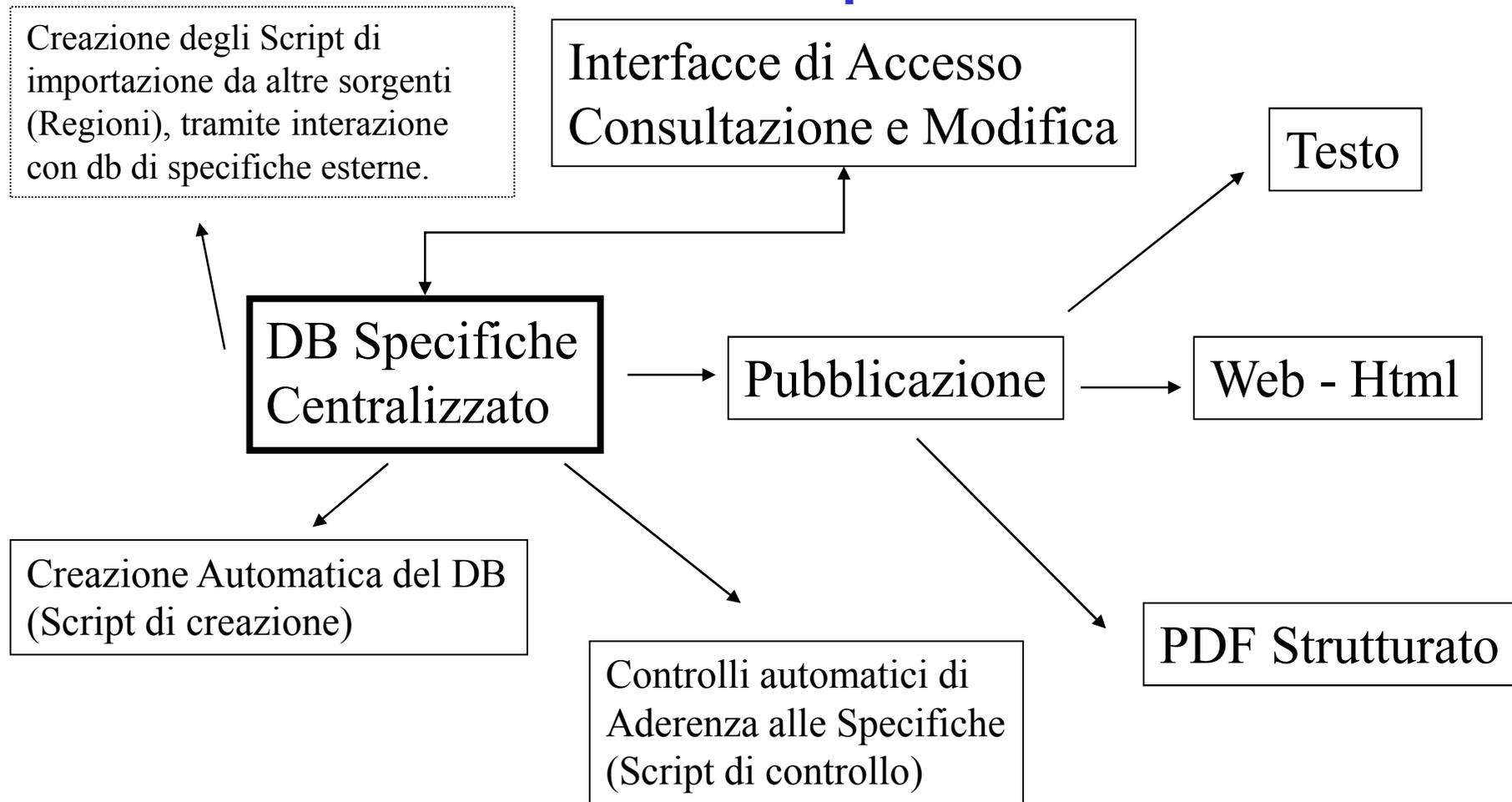


Un esempio di approccio: Il database delle specifiche IGM

- Vediamo adesso l'approccio utilizzato in IGM, per la stesura delle specifiche per il geodb.
- I principali autori del lavoro sono i Cartografi Perugi, Di Rienzo, Ponziani e Fortini.
- L'idea è quella di descrivere le specifiche del db geografico tramite un altro DB (questa volta una semplice base di dati convenzionale).
- Il meta-database (il database che descrive il database cartografico), memorizza tutte le informazioni di specifica, utili per la realizzazione fisica del geodb.



Schema utilizzo specifiche IGM

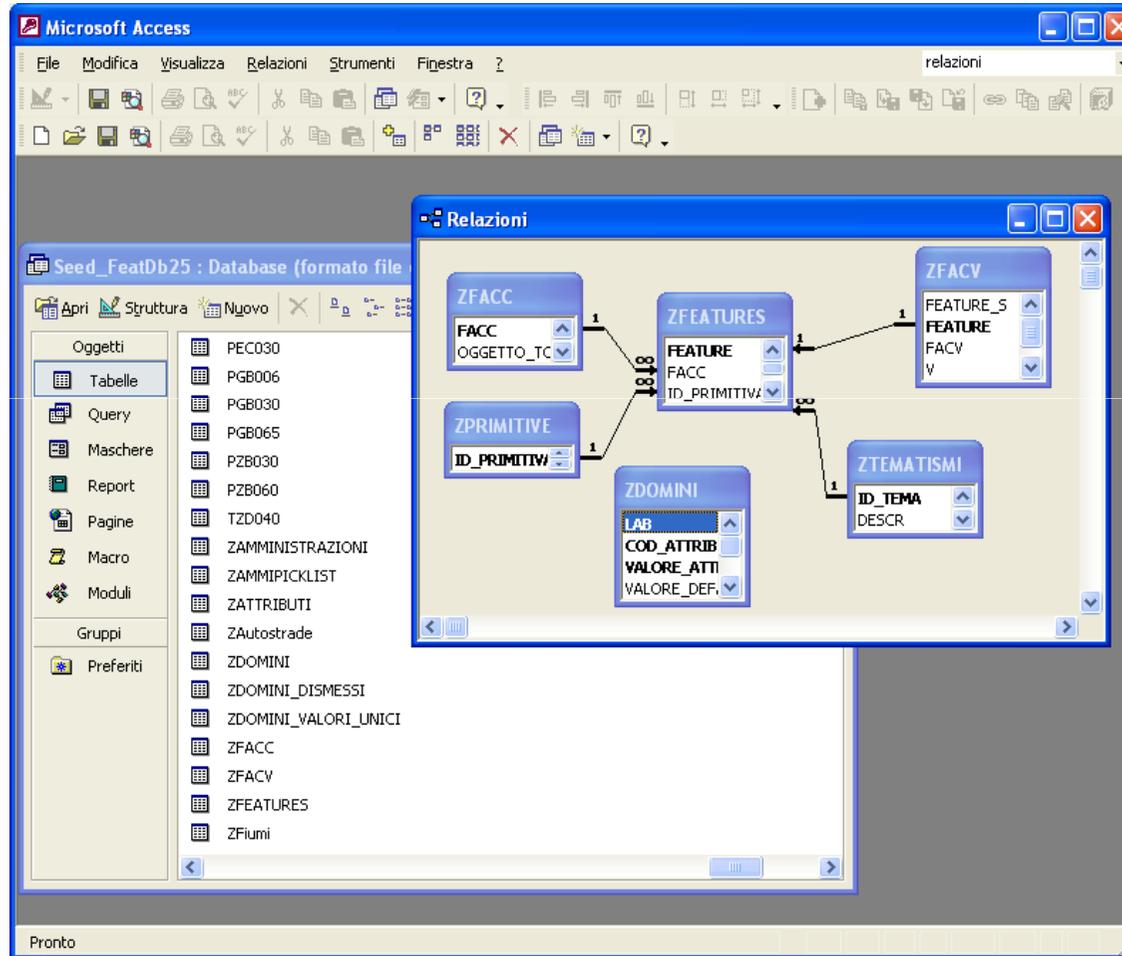


Note implementative

- Disegnare lo schemino del lucido precedente è (abbastanza) facile.
- Poi però il sistema lo abbiamo realizzato davvero (almeno il prototipo).
- Il db utilizzato per le specifiche è Access.
- Il sistema di Accesso e Modifica è in Visual Basic.
- I generatori automatici di script sono in C++ con accesso ODBC.
- Il PDF viene prodotto con l'ausilio di Latex (versione Miktex).



Il database Access



Interfaccia Accesso e Modifica

db25

TEMATISMO FACC

DESCRIZIONE FACC

FEATURE

OGGETTO

Sotto_Descrizione

Vestizione (i valori sono espressi in mm)

Relazioni logiche e geometriche

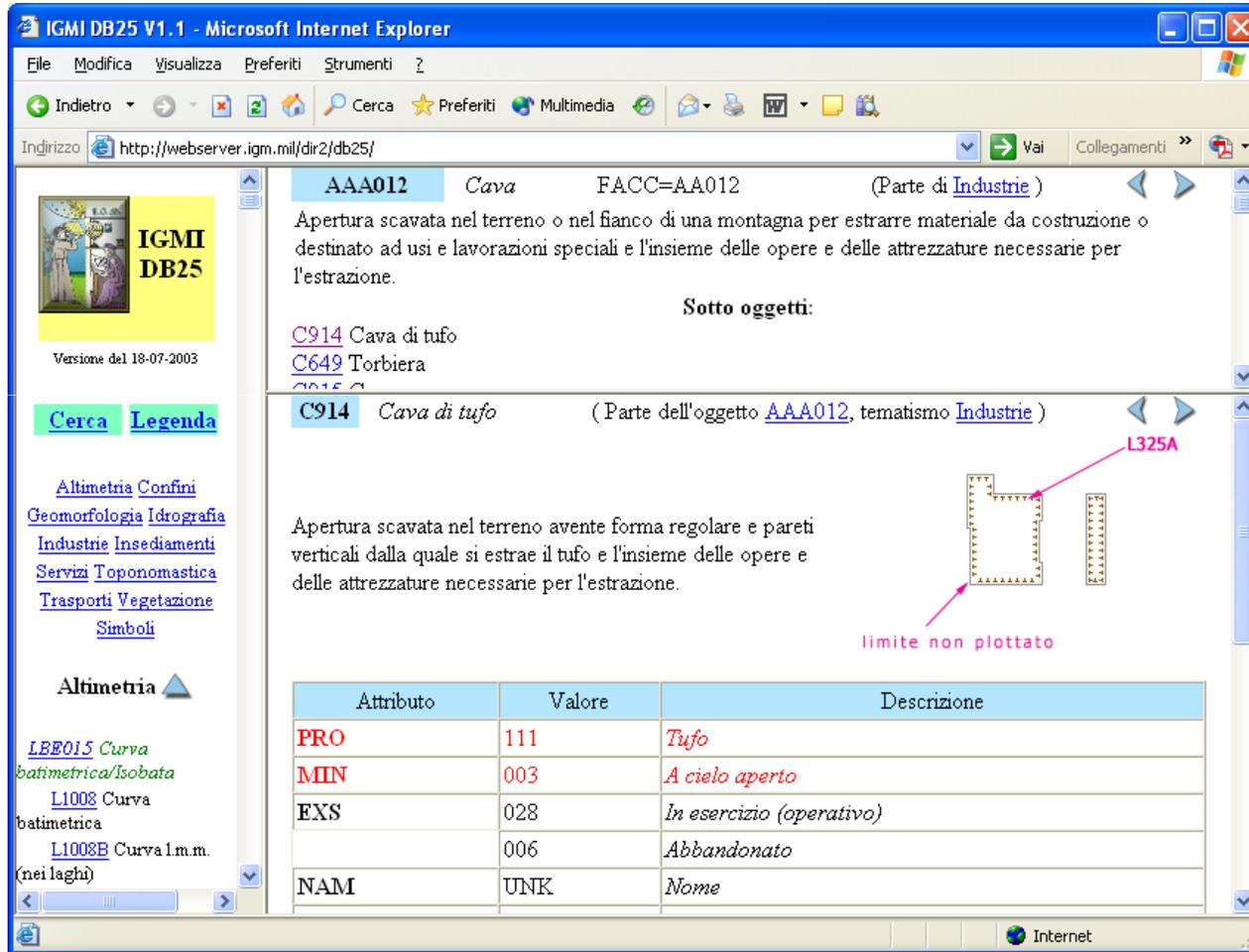
Fonti di acquisizione

Criteri di selezione

db25 9/23/2005 10:57 AM



Esempio Pubblicazione Web



IGMI DB25 V1.1 - Microsoft Internet Explorer

Indirizzo: <http://webserver.igm.it/dir2/db25/>

AAA012 Cava FACC=AA012 (Parte di [Industrie](#))

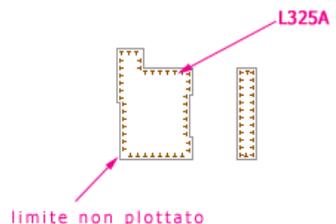
Apertura scavata nel terreno o nel fianco di una montagna per estrarre materiale da costruzione o destinato ad usi e lavorazioni speciali e l'insieme delle opere e delle attrezzature necessarie per l'estrazione.

Sotto oggetti:

- [C914](#) Cava di tufo
- [C649](#) Torbiera
- [C915](#) Cava di tufo

C914 Cava di tufo (Parte dell'oggetto [AAA012](#), tematismo [Industrie](#))

Apertura scavata nel terreno avente forma regolare e pareti verticali dalla quale si estrae il tufo e l'insieme delle opere e delle attrezzature necessarie per l'estrazione.



Attributo	Valore	Descrizione
PRO	111	<i>Tufo</i>
MIN	003	<i>A cielo aperto</i>
EXS	028	<i>In esercizio (operativo)</i>
	006	<i>Abbandonato</i>
NAM	UNK	<i>Nome</i>



Database Intesa: Scelte implementative

- Per la realizzazione fisica del DB Intesa, abbiamo effettuato le seguenti scelte:
 - Realizzazione di un personal geo-database, contenente un feature-dataset (per poter inserire i vincoli topologici).
 - Le feature-class corrispondono alle feature Intesa, il nome è quello numerico preceduto da una F (che sta per feature), eventualmente è presente l'alias con il nome testuale.
 - Il sistema di riferimento è ovviamente UTM fuso 32N, WGS84, l'estensione del dato è stata importata dall'unione delle maschere di ritaglio (che coprono l'intera superficie).
 - Per quasi tutte le feature si è creato un unico tipo geometrico rappresentativo, per alcune feature (ponti) è stato necessario inserire più feature con tipi geometrici diversi, in questo caso i nomi delle feature terminano con P,L,A (punto, linea, area).
 - I nomi delle feature sono quelli numerici (preceduti da A=attribute), lo standard SQL vieta di creare attributi che iniziano con cifre, ArcGis non è in grado di gestirli.
 - Tutti i domini codificati sono stati creati con il loro nome simbolico e l'alias testo, quindi inseriti nel personal geodatabase. Ovviamente tutti gli attributi associati ai domini sono stati collegati ai rispettivi domini dove previsto.
 - Il tipo degli attributi codificati è sempre testo, questo perché potrebbe essere necessario distinguere i codici con numero di cifre diverso (2 o quattro). Con attributo di tipo numerico il codice 04 risulta uguale al codice 0004. Nello spirito delle specifiche Intesa sembra che gli zero all'inizio del codice siano significativi (si scrive 0104 non 104).
 - Non c'è una gerarchia esplicita (non si sono sottoclassi).

