

Trabocchetti nell'importazione di attributi

Claudio Rocchini
Istituto Geografico Militare

Introduzione

- L'importazione di attributi da un sistema cartografico ad un altro è, in generale, molto meno problematica di quella della geometria.
- Gli attributi sono di solito, gestiti in modo uniforme dai vari gestori di basi di dati (es. DB2 associati agli Shape, MS-Access nei personal DB, Oracle , etc.).
- Tuttavia ci possono essere in agguato alcune insidie che possono portare alla perdita di dati durante le operazioni di import/export.



Interpretazione delle impostazioni locali per i numeri

- Sistemi diversi possono interpretare in modo incoerente (anche sullo stesso computer) le impostazioni locali per i numeri (punto decimale al posto della virgola).
- Es: Postgis utilizza le proprie impostazioni interne per le impostazioni dei numeri, mentre il Personal DB access usa le impostazioni utente di Windows.
- La cosa si complica con l'utilizzo di server remoti (Oracle), che risiedono su di un'altra macchina, che può avere impostazioni diverse dalla macchina di lavoro.
- Es. Geomedia Intergraph non carica i file geotiff (TIFF+TFW) se le impostazioni di Windows non sono su Inglese-Stati Uniti.
- In ogni caso è necessario controllare la corretta importazione dei dati numerici con cifre decimali. In caso di errore bisogna controllare le impostazioni locali dei computer coinvolti.



Interpretazione delle impostazioni locali sulle date

- In alcuni casi i dati cartografici contengono dati temporali (orari, intervalli di tempo, date).
- Impostazioni internazionali incoerenti possono portare ad una cattiva interpretazione del formato di data.
- Es. il formato standard della data anglosassone è mese-giorno-anno.
- Anche in questo caso bisogna controllare la corretta importazione dei dati: es. finché il giorno del mese è minore di 12, l'errore è insidioso perché non viene rilevato da nessun controllo.



Piccole divergenze sul tipo di dati

- Altri errori insidiosi possono fiorire da piccole divergenze sulla definizione del tipo di dati:
 - Riduzione della massima lunghezza del campo testo: rilevato solo in presenza di un testo molto lungo.
 - Riduzione dei decimali significativi: perdita di precisione.
 - Riduzione della risoluzione dei numeri interi (caso reale IGM: popolazione di un comune PPL con 16 bit di risoluzione, max. numero 32767, non rilevato finché non sono stati trattati grossi centri abitati).
 - Importazione di dati numeri in dati testuali (caso reale IGM: quote dei punti quota, non rilevato finché non si sono utilizzate query sugli attributi con intervalli altimetrici, '91' risulta maggiore di '1200' (come testo in ordine lessicografico)).



Trattazione dei campi vuoti

- Gli attributi facoltativi di un db possono essere vuoti (valore NULL).
- Durante l'import/export il gestore del database può tentare di:
 - Sostituirli con un valore di default (se previsto).
 - Trasformare i campi testo in stringhe di zero caratteri (che non corrispondono a NULL).
 - Eseguire il "padding" sui campi testo (= riempimento con caratteri spazio).
- Questi comportamenti possono generare informazioni fantasma, non realmente presenti nei dati di origine.

