

ANNO SECONDO
NUMERO DUE

dicembre 1994

contributi di:
BUCCIANTI
CAPACCIONI
CARNIEL
MANSINI
PEVERIERI
PRINCIPE
ROMANO
ROMEO
SARACCO
SAROCCHI
SIGNORE
SPAZIANI
SPERONI
TACCETTI
VANNOZZI

Rivista del



Gruppo per l'Informatica
Applicata alle Scienze della Terra



WWW: la panacea?

Oreste Signore

CNR-CNUCE
Via S. Maria, 36
56126 Pisa (Italy)

GEOINFORMATICA, N. 2, pp. 99 - 103, 1994

1. I dati di Scienze della Terra

I dati delle Scienze della Terra sono dispersi tra le tante istituzioni, scientifiche e non, che li producono e li utilizzano per fini scientifici o strumentali. Questa situazione è un fatto storico: ognuno ha raccolto e organizzato i dati secondo le sue finalità, tenendo conto solo in parte delle esperienze condotte da altri. L'ovvia conseguenza è che mentre in alcuni casi sono stati ripetuti degli errori già commessi, in altri casi non è stato possibile rendere patrimonio comune alcune brillanti soluzioni. Ovviamente, tra le varie istituzioni esistono anche incompatibilità di hardware e di software. Di converso, la comunità scientifica avverte la necessità di poter accedere, manipolare e scambiare dati, indipendentemente dalla loro locazione fisica e dalle restrizioni imposte dall'hardware e dal software, chiaramente artificiose e ingiustificate.

1.1. Uno schema concettuale globale?

Come già accennato in ([Signore93]), il processo di rappresentazione dei dati di scienze della terra parte da una fase di raccolta di dati oggettivi, e costituisce (concettualmente) una base di dati in cui convivono i dati grezzi, i dati interpolati, i dati geometrici e gli elementi cartografici, mediante un processo al quale contribuiscono la conoscenza specifica del geologo e i modelli. Il linguaggio è adeguatamente normalizzato grazie all'esistenza di thesauri e lessici, e le forme di rappresentazione più utilizzate sono diagrammi, tabelle e, soprattutto, carte.

I problemi essenziali, ad alcuni dei quali si è peraltro già fatto cenno, sono costituiti dall'integrazione di dati puramente alfanumerici con dati relativi ad oggetti con ben definite caratteristiche spaziali (punti, linee, poligonali, aree) sui quali è necessario poter operare con opportuni operatori. Inoltre, l'opportunità, segnalata da più autori, di operare in un contesto a quattro dimensioni (x,y,z,t) aggiunge un ulteriore fattore di complessità. Infine, i problemi dello scambio dei dati e del reperimento delle informazioni rendono evidente la necessità di disporre di un linguaggio normalizzato e di uno schema di riferimento univoco, che consenta l'identificazione precisa di tutti i concetti.

La risposta a queste esigenze dovrebbe essere fornita dalla definizione di uno schema concettuale dei dati di Scienze della Terra. Esso consente di giungere ad una visione comune dei dati, distinguendo in modo chiaro tra dati grezzi, base per future analisi, eventualmente

fondate su ipotesi scientifiche diverse, e dati interpretati, che costituiscono la vera ricchezza delle carte geologiche, ma possono venir messi in discussione alla luce di nuove conoscenze scientifiche. Ovviamente, la definizione di uno schema concettuale comune non limitata, ma esalta la possibilità di visualizzare i dati in maniera diversa a seconda delle specifiche necessità; esso ha il compito fondamentale di offrire una rappresentazione semiformale della realtà di interesse, e consente ad utenti e progettisti di applicazioni e di basi di dati di comunicare con un linguaggio non ambiguo. La visualizzazione dei dati è affidata al software applicativo, da sviluppare in base alle specifiche esigenze.

Il logico corollario di un approccio di questo tipo è costituito dalla possibilità di accedere all'informazione in qualunque posto essa sia localizzata, e, inoltre, consente di definire in modo semplice e immediato un formato di scambio dei dati, perchè garantisce l'equivalenza dei dati a livello semantico, mentre è evidentemente elementare definire successivamente il livello logico (cioè la struttura dei record) e il livello fisico (il supporto o il mezzo di trasmissione) di un formato di scambio.

Di conseguenza, il modello concettuale costituisce il vero e unico quadro di riferimento per tutte le attività svolte nel settore. Tuttavia è evidente che la definizione di uno schema concettuale integrato costituisce un obiettivo estremamente ambizioso, che richiede un enorme dispendio di energie umane e finanziarie, senza poter garantire un immediato ritorno economico, e passa attraverso una fase di raffinamenti successivi che possono dare concretamente l'impressione di una tela di Penelope.

2. L' iniziativa del GIAST

Il GIAST ha avviato alcuni anni fa un' iniziativa, finanziata dal CNR su fondi del Comitato per le Scienze Geologiche e Minerarie e del Comitato per le Tecnologie dell'Informazione, che si proponeva di giungere alla definizione di uno schema concettuale comune e alla identificazione dei paradigmi di interazione dell'utente ritenuti più utili e significativi.

Il primo passo è stato costituito dall'analisi e omogeneizzazione degli schemi concettuali prodotti indipendentemente da Servizio Geologico nazionale e dalla Regione Emilia-Romagna. La metodologia adottata è stata quella classica dei sistemi informativi, e prevedeva una definizione precisa di ogni Entità, con una separazione netta degli aspetti puramente geometrici.

I risultati ottenuti, in verità, non sono stati particolarmente esaltanti: i due schemi, per quanto sviluppati con lo stesso formalismo, sono risultati di fatto inconciliabili, perchè basati su due presupposti diversi. Il problema fondamentale, però, è stato costituito dalla mancanza di un adeguato apporto da parte degli specialisti del settore, a differenza di quanto avvenuto in altri contesti stranieri, in cui il gruppo di progetto era costituito essenzialmente da studiosi di Scienze della Terra. In conclusione, pur non esistendo dubbi sulla validità dell'idea di base, bisogna riconoscere che in questo settore, come del resto in

tanti altri, sono necessari il coinvolgimento degli esperti ma, soprattutto, una loro forte motivazione.

3. Una soluzione?

In realtà, e forzando un po' i termini della questione, esiste una dicotomia tra la visione del problema da parte degli esperti che affrontano il problema della carta geologica e della gestione dei dati di Scienze della Terra, a seconda che provengano dall'area culturale GIS o database.

Nel primo caso, l'enfasi è posta essenzialmente sulla riproduzione cartografica, e il database viene visto come un elemento destinato a contenere i dati di supporto. Molto spesso, inoltre, vengono utilizzate solo marginalmente le funzionalità più interessanti dei GIS, che vengono quindi ridotti a meri strumenti per la produzione di carte, senza mettere a punto dei modelli. Infine, fa parte dell'esperienza comune la difficoltà, e talvolta l'impossibilità, di trasferire i dati da un sistema GIS all'altro.

Nell'approccio più spiccatamente database, i dati vengono visti come la risorsa principale, e si suppone che sia possibile rispondere a qualunque domanda posta dall'utente formulando una o più query; la mappa viene vista come uno dei meccanismi per la presentazione dei dati. Anche in questo caso, lo scambio di informazioni non sempre è semplice, perchè i dati sono organizzati in maniera diversa, non sempre adeguatamente normalizzata, e campi apparentemente equivalenti possono presentare differenze semantiche rilevanti. In realtà, l'utente vede il problema con una prospettiva affatto diversa, sostanzialmente indipendente dalle soluzioni tecnologiche: richiede rappresentazioni mappali, dati di dettaglio, ricontestualizzazione spazio-temporale delle informazioni, riferimenti bibliografici, possibilità di "annotare" osservazioni personali.

Di conseguenza, nè i GIS, con le loro limitazioni nella gestione dei dati tridimensionali e temporali, nè i DBMS, caratterizzati da una elevata parcellizzazione dell'informazione, possono da soli offrire una risposta soddisfacente. Inoltre, nessuno dei due ambienti supporta in modo immediato la possibilità di "annotare", e quindi catturare la conoscenza dell'utente. Entrambi, infine, richiedono opportune estensioni per integrare i dati provenienti da fonti esterne (filmati, foto aeree, documenti, etc.).

Un esempio tipico di interazione potrebbe essere quello nel quale l'utilizzatore del sistema, partendo da una rappresentazione geografica:

o
individua una zona di interesse
o,
interroga il sistema per avere le caratteristiche geologiche
e una lista di pubblicazioni riguardanti la zona
o,
ottiene indicazioni per accedere ai dati grezzi relativi alla zona

selezionata,
accede a quelli di interesse,
li consulta ed importa sulla propria workstation quelli di interesse
specifico
o,
applica un modello per ottenere una rappresentazione
delle informazioni secondo una chiave di lettura personale
o,
a partire dalle caratteristiche geologiche della zona
ne cerca altre che presentino degli elementi comuni
o,
a partire dalla lista delle pubblicazioni ricerca altre pubblicazioni
prodotte da qualcuno degli autori, oppure altre pubblicazioni
che trattino argomenti analoghi
o,
da una delle pubblicazioni ritrovate individua
una nuova zona geografica di interesse
o,
in un qualunque punto di questa navigazione "annota"
le sue osservazioni personali.

Ovviamente, le azioni delineate precedentemente possono essere ripetute più volte e in sequenza arbitraria: esse sono dettate unicamente dal particolare meccanismo associativo che l'utente intende privilegiare in quello specifico momento, sulla base dei suoi interessi scientifici contingenti.

4. L'ambiente ipermediale

Allo stato attuale della tecnologia, l'ambiente ipermediale, che consente la rappresentazione contestuale di informazioni grafiche e alfanumeriche, ma soprattutto, permette di realizzare applicazioni nelle quali l'utente si comporta in modo "attivo", e può accedere ai dati utilizzando i meccanismi associativi che più gli sono congeniali, sembra costituire il contesto ottimale per soddisfare esigenze sofisticate come quelle delineate precedentemente.

Tuttavia, se è vero che la possibilità di utilizzare molteplici paradigmi di interazione (vicinanza spaziale o temporale, legame di classificazione, etc.) e di utilità e potenza straordinarie, non si deve trascurare l'intrinseca complessità del problema di fondo. La progettazione e realizzazione di una applicazione ipermediale soddisfacente richiede una comprensione approfondita dei tipi di legame che possono sussistere tra i vari nodi, e quindi una rappresentazione della realtà di interesse, ossia uno schema concettuale. L'implementazione di paradigmi di interazione basati sull'associazione dei concetti attribuibili ai singoli nodi richiede evidentemente la definizione di standard di linguaggio e di classificazione, cioè, in ultima analisi, la competenza disciplinare dello studioso.

Di conseguenza, le nuove tecnologie possono essere senza alcun dubbio di eccezionale aiuto, ma le loro potenzialità restano inutili, e, spesso, fastidiosi orpelli se non vengono affrontati e risolti problemi sul tappeto ormai da diversi anni.

È evidente, a questo punto, la realizzabilità di un ipermedia distribu-

ito, grazie al quale l'utente, grazie alle "autostrade dell'informazione" può accedere ai dati dispersi tra varie località geograficamente distribuite, utilizzando meccanismi di interazione vari e intercambiabili, che seguono legami espliciti o impliciti.

Vorrei tuttavia temperare l'entusiasmo suscitato in molti ricercatori e studiosi dalla possibilità di accedere via rete a server di informazione situati in luoghi geograficamente distribuiti. Se si accede alle "autostrade dell'informazione" senza avere una idea precisa di cosa andare a trovare e dove, si può cadere vittime di quella che, per analogia, potremmo chiamare la "sindrome di Francoforte", nota a qualunque automobilista che si trovi ad attraversare la Germania e debba scegliere il percorso migliore per raggiungere o superare Francoforte. Spesso, la mancanza di informazioni precise, e la differente organizzazione concettuale dei dati, trasformano la navigazione su Internet in una inutile e frustrante perdita di tempo.

5. Conclusioni

La gestione integrata dei dati di Scienze della Terra costituisce un problema di complessità enorme, che può essere affrontato solo grazie ad un consistente impegno intellettuale da parte degli studiosi. Le tecnologie ipermediali offrono un valido supporto per poter rappresentare in modo integrato vari tipi di informazione. La disponibilità di reti e di applicazioni che consentono l'accesso amichevole ai dati permette di realizzare sistemi molto sofisticati, che però possono ridursi a mere esercitazioni tecnologiche se non vengono comunque affrontati e risolti alcuni dei problemi tradizionali legati alla formalizzazione e rappresentazione dell'informazione.

Riferimenti

- [SIGNORE93] O. SIGNORE (1993) *Verso un sistema unificato per la gestione dei dati di Scienze della Terra?*, Geoinformatica, Anno 1, n. 1 (pp. 161-169).
- [SIGNORE94] O. SIGNORE, R. BARTOLI, G. SEVERINO (1994) *A Unified Approach for the Management of earth Sciences Data*, 1st European Congress on Regional Geological Cartography and Information Systems, Bologna, June 13-16, 1994, Volume 4 (extended abstract n. 132).
- [SIGNORE95] O. SIGNORE (1995) *Issues on Hypertext Design*. Proceedings of DEXA '95 - 6th International Conference and Workshop on Database and Expert Systems Applications, London, UK, Sept. 4-8, 1995 - Springer-Verlag, (in pubblicazione)