

**Il sistema di gestione degli archivi nello scavo di Poggio Imperiale a Poggibonsi
(Insegnamento di Archeologia Medievale dell'Università di Siena).
Una soluzione all'interno della "soluzione GIS".**

Vittorio Fronza

*Atti del I Convegno Nazionale di Archeologia Computazionale, Napoli 5-6 febbraio 1999,
"Archeologia e Calcolatori", 11, 2000, pp125-137*

1 – Cenni di Database Management e teoria relazionale

Com'è noto il *Database management* ha rappresentato uno dei primissimi obiettivi perseguiti nell'ambito delle scienze informatiche. Il modello di base per la gestione di un archivio si fonda sulla teoria matematica delle matrici e si esplicita in campo informatico con il concetto di tabella suddivisa in righe e in colonne. Una riga rappresenta un'unità di informazione all'interno dell'archivio ed è composta da un numero predefinito di colonne, variabile in base alla natura dell'archivio stesso. Correntemente un'unità di informazione, in altre parole la riga di una tabella, è detta record, mentre la colonna, ovvero una voce dell'intera informazione di un record, è detta campo (o attributo). In definitiva una tabella è un database semplice, bidimensionale; in questo senso è spesso definito come archivio lineare. Una simile organizzazione può risultare molto farraginoso con l'aumentare della complessità e della diversificazione delle informazioni; la soluzione è rappresentata da una gestione relazionale dei dati, attraverso la quale più archivi lineari sono messi in relazione per mezzo di una colonna comune. La teoria relazionale si basa su tre concetti fondamentali: le tabelle appena descritte, gli identificatori (ovvero i campi comuni a più tabelle che permettono di stabilire le relazioni), le relazioni.

Si generano in questo modo archivi complessi che da un punto di vista algebrico sono delle matrici multidimensionali. L'algebra però non è in grado di stabilire dei collegamenti logici fra i vari elementi di una matrice multidimensionale, se non quelli immediati rappresentati da un'identità di riga, di colonna, di "profondità", ecc. Si è dovuto perciò ricorrere alla teoria matematica degli *insiemi*. Questa prevede la manipolazione di elementi aventi almeno una caratteristica in comune, raggruppati appunto in insiemi; fra le operazioni principali previste ricordiamo l'addizione, la sottrazione, l'unione, l'intersezione. Le basi di questa teoria, opportunamente riviste e adattate, nonché ampliate in alcune sue caratteristiche soprattutto operative, sono risultate perfettamente aderenti alle necessità di esprimere la relazionalità fra archivi.

La teoria relazionale classica prevede tre categorie di relazioni:

1:1 si tratta della relazione che collega un record di un archivio ad uno ed un solo record di un altro archivio.

1:N; N:1 mette in relazione un record di un archivio (detto principale o master) con *N* record di un altro archivio (detto secondario), o viceversa. Per gli archivi di uno scavo archeologico ad si avrà, ad esempio, una relazione di tipo *1:N* dalla tabella *Unità stratigrafiche* verso la tabella *Reperti ceramici* (per uno ed un solo strato si possono avere *n* schede di reperti ceramici).

N:N si tratta delle relazioni più complesse da gestire e difficilmente si troveranno applicate in archivi di largo consumo. Tramite questa relazione vengono messi in collegamento *n* record di un archivio con *n* record di un altro archivio, secondo un criterio logico aderente all'architettura globale del database relazionale.

Questa succinta spiegazione della teoria relazionale, e l'organizzazione razionale del dato che ne deriva, dovrebbe bastare per intuire l'utilizzo che se ne può fare nella gestione dei dati prodotti dalle indagini archeologiche. Nel caso di uno scavo, per citare un esempio semplice, l'archivio delle US messo in relazione con gli archivi dei reperti snellisce notevolmente il database. Si evita infatti in questo modo la duplicazione dei dati relativi alla stratigrafia in ciascuna scheda dei reperti (in informatica detta ridondanza del dato); ciò consente di risparmiare spazio sulle memorie di massa del calcolatore e, soprattutto, tempo nella fase di immissione dei dati.

2 – Database di scavo: il panorama italiano

A livello italiano non si rintracciano molti esempi riferibili alla realizzazione di un database per la catastazione dei dati alfanumerici prodotti durante un'indagine stratigrafica. I primi tentativi, pubblicati a partire dalla metà degli anni '80, sono finalizzati soprattutto alla classificazione dei reperti; fra i più rilevanti si ricorda il progetto BDAR (MOSCATI, 1987, pp. 48-50; GUERMANDI, 1989, p. 30), realizzato dall'Università di Lecce che, pur prevedendo la schedatura dei dati stratigrafici generali, si concentra soprattutto sulla produzione di banche dati dei reperti al fine di effettuare analisi statistiche. Un lavoro simile è stato affrontato, nello stesso periodo, dall'Università di Pisa in collaborazione con il CNUCE ed ha riguardato la strutturazione di un modello relazionale dei dati di scavo (MOSCATI, 1987, p. 50). Altre due applicazioni che hanno previsto una gestione relazionale sono il programma CARTESIO utilizzato dalla Soprintendenza Archeologica della Lombardia e ADONIS sviluppato presso l'Ecole Francaise di Roma (GUERMANDI, 1989, pp. 28-29); in entrambi i casi si tratta di implementazioni in dbase III. Da segnalare anche l'esperienza dell'equipe (costituita da ricercatori universitari e personale della Soprintendenza archeologica del Lazio) che ha operato sul sito di Casale Nuovo (LT) dove, oltre a sperimentare l'applicazione della tecnologia informatica all'archeometria, sono state effettuate schedature e analisi statistiche dei reperti (GUERMANDI, 1989, p. 29).

Tutti i casi finora descritti sono riferibili ad esperienze comunque interessanti, soprattutto in relazione agli anni in cui furono effettuate, ma concentrate su problemi specifici e spesso legati indissolubilmente al contesto; si tratta in sostanza di soluzioni parziali. Il primo tentativo di gestione globale delle informazioni archeologiche è rappresentato dal progetto pisano denominato ARGO (RICCI, 1988; MOSCATI, 1987, p. 50; GUERMANDI, 1989, p. 29). Si è tentato, in questo caso, di realizzare una gestione integrata di tutti i dati prodotti dalle ricognizioni di superficie e dalle indagini stratigrafiche, prevedendo un complesso momento di analisi che rendesse il progetto indipendente dalle piattaforme hardware e software e minimizzasse l'uso di codici numerici per l'archiviazione dei dati, metodo fino ad allora dominante; la mancanza di finanziamenti adeguati ha causato il fallimento di questo interessante tentativo.

Un ulteriore passo avanti è stato compiuto attraverso la realizzazione del progetto Aladino (GUERMANDI, 1989), un database relazionale nato per la catastazione dei dati provenienti dallo scavo sul sito di Castelraimondo-Forgaria (Udine) e realizzato dall'IBC dell'Emilia Romagna in collaborazione con l'Istituto di Archeologia dell'Università di Bologna. Si tratta del primo (e forse unico) lavoro orientato soprattutto alle esigenze della ricerca; nell'intenzione dei curatori "il tentativo è stato insieme quello di fornire uno strumento che fosse una risposta concreta ad un problema scientifico molto preciso e di testare altresì una metodologia operativa che fosse applicabile anche ad altri ambiti, discipline o casi specifici" (GUERMANDI, 1989, p. 35). L'archivio relazionale che ne è derivato (costituito da sei tabelle: US, reperti ceramici, reperti metallici, reperti vitrei, altri reperti, reperti organici) prevede la

normalizzazione del linguaggio attraverso dizionari e un uso minimizzato di campi a carattere descrittivo e l'archiviazione di immagini e dati grafici oltre che alfanumerici. L'analisi effettuata ha portato allo sviluppo di un prodotto estremamente valido, soprattutto per l'indipendenza dalle piattaforme hardware/software, l'architettura aperta all'eventuale aggiunta o modifica dei tipi di informazioni gestite, la semplicità d'uso da parte di utenti di fascia medio-bassa attraverso l'implementazione di un'interfaccia a menù, la facilità d'esportazione dei dati.

In tempi più recenti la soluzione più completa per l'informatizzazione dei dati di scavo è rappresentata dal sistema ODOS, sviluppato dall'Università degli Studi di Lecce in collaborazione con il CNR (D'ANDRIA, 1997). Il database relazionale, inserito in un sistema globale di gestione dei dati (denominato appunto ODOS), prevede funzioni complete di data entry e funzioni di interrogazione diversificate (da semplice navigazione relazionale fra record correlati a programma di interrogazione interattivo e creazione di tabelle riassuntive). La struttura è composta dall'organizzazione gerarchica "classica" di un database relazionale partendo dal livello dei siti e arrivando agli archivi delle singole classi di reperti. L'utilizzo del database, gestito attraverso un'interfaccia grafica, prevede accessi diversificati (livello basso: solo interrogazione; livello medio: accesso ai dati; livello alto: accesso ai vocabolari e ai dati). Si può affermare di essere in presenza di un'analisi della massima chiarezza concettuale da un punto di vista informatico (notevoli in questo senso gli sforzi compiuti per una chiarificazione dell'architettura hardware/software del sistema e per una strutturazione efficiente dei dati), da implementarsi su una LAN attraverso l'utilizzo dell'applicativo Oracle, uno degli standard per la gestione relazionale di grosse banche dati. D'altro canto proprio la scarsa flessibilità, connaturata al tipo di progettazione logica e all'ambiente hardware e software utilizzato, rende il prodotto poco adattabile alle esigenze spesso variabili della ricerca archeologica (modificare un sistema chiuso di questo tipo comporterebbe sicuramente tempi piuttosto lunghi e costi elevati); essendo pensato insieme agli archeologi ma sostanzialmente plasmato dagli informatici, si rivela poco aperto ad analisi approfondite su problematiche specifiche, oltre a garantire una scarsa accessibilità ad utenti di fascia media o medio-bassa. Considerazioni simili, anche se applicate ad un complesso di dati più ridotto (la gestione dei dati provenienti dallo scavo di una necropoli), valgono per il progetto Eurialo incentrato sulla gestione integrata dei dati archeologici di Pontecagnano (MOSCATI, 1998, pp. 226-228).

3 – Le premesse e i principi fondamentali

La gestione dei dati alfanumerici provenienti da uno scavo ha rappresentato una delle prime applicazioni delle tecnologie informatiche all'archeologia sperimentate presso il Laboratorio di Informatica applicata all'Archeologia. Fin dalla fine degli anni '80 si sono imposte, utilizzando le prime versioni lineari di dell'applicazione FileMaker, tabelle lineari per la gestione dei dati pertinenti alle unità stratigrafiche ed alle principali classi di reperti; si trattava specificatamente di una semplice implementazione informatica delle schede cartacee ministeriali, primo e necessario approccio all'informatizzazione applicata ai dati di scavo. La scoperta delle possibilità offerte dal calcolatore, la cresciuta alfabetizzazione dei componenti del laboratorio e la commercializzazione della release 3.0 di FileMaker Pro hanno portato, nel 1995, alla messa a punto di una prima versione relazionale degli archivi alfanumerici di scavo. Si trattava di una base di dati basata su un'indice relazionale per US; l'architettura dei campi riprendeva ancora le schede cartacee ministeriali. I difetti principali di questa base di dati consistevano nella scarsità dei controlli durante la fase di *data entry*, nella limitata

implementazione dell'architettura relazionale e nell'incompletezza dell'interfaccia utente; in sostanza il database non si poteva considerare un prodotto finito.

Nel 1998, in concomitanza con la realizzazione della prima versione del sistema di gestione complessivo del dato archeologico denominato OpenArcheo si è reso necessario effettuare un aggiornamento sostanziale del database; in realtà si è trattato di un momento di riflessione che ha condotto ad una riformulazione delle esigenze, alla stesura di un'analisi, al rinnovo complessivo della struttura degli archivi e, di conseguenza, ad un'implementazione completamente nuova utilizzando la versione 4.1 di FileMaker Pro. Questo software, adottato fin dalle prime sperimentazioni presso il Dipartimento, pur non presentando un'implementazione estremamente rigorosa della teoria relazionale mette a disposizione funzioni di utilizzo immediato sufficienti a gestire il dato archeologico; altri pregi sono un'ottima gestione dei controlli sulla consistenza del dato, un linguaggio residente di discreta potenza, la possibilità di una facile immissione dei database su Internet (opzione in progetto da tempo presso il nostro Laboratorio), la possibilità di usufruire di una vasta libreria di comandi AppleScript. FileMaker, nella sua versione per server, prevede inoltre una gestione efficace e semplice degli accessi al database su una rete locale.

Le premesse alla base della nuova versione si possono riassumere nei punti che seguono:

- archivi progettati in funzione della ricerca, aperti alle esigenze che derivano dallo svolgimento di un progetto (necessità di gestire nuovi tipi di dati, realizzare analisi e letture interpretative, assecondare nuovi orientamenti); fondamentale in quest'ottica la presenza di livelli interpretativi differenziati;
- implementazione su personal computer;
- utilizzo esteso delle possibilità di gestione relazionale offerte dal software;
- creazione di un prodotto facilmente compilabile, che permetta la realizzazione di versioni *standalone* parziali o complete ai fini di un utilizzo mirato del database nelle diverse fasi dell'indagine;
- sviluppo di un'interfaccia utente di facile utilizzo.

Il database per la catastazione dei dati di scavo elaborato in base alle specifiche appena descritte si conforma inoltre ad un modello in corso di elaborazione per la gestione globale degli archivi relativi alle indagini archeologiche. Dal 1997, infatti, un'apposita commissione istituita dalla Regione Toscana sta lavorando alle *Linee Guida per la redazione della Carta Archeologica della Toscana* (LINEE GUIDA CARTA ARCHEOLOGICA, 1998). Il Laboratorio di Informatica applicata ha intrapreso, dall'inverno 1998, le operazioni di codificazione secondo l'analisi scaturita dal lavoro della commissione; ciò presenta problemi di diversa natura e riferibili soprattutto all'adeguamento al *software* scelto, alla riconversione degli archivi precedenti e al disegno di un'interfaccia utente per il *data entry* e per la consultazione. Si tratta di una mole di lavoro rilevante; in sostanza prevediamo la costruzione di una base di dati che contempra la gestione di qualsiasi tipo di dato archeologico alfanumerico. Le *Linee Guida*, estremamente complete per quanto riguarda il ventaglio dei dati prodotti da una carta archeologica, sono piuttosto schematiche nell'approfondimento dei dati oggettivi ed interpretativi connessi allo svolgimento di un progetto di ricerca; si renderà pertanto necessario integrarle con gli archivi di dettaglio, direttamente funzionali alle tematiche affrontate e già sviluppati presso il laboratorio; va inoltre prevista l'aggiunta di una serie di dati particolari relativi ad approfondimenti in corso di svolgimento presso il Dipartimento (archivio dei castelli italiani scavati, archivio degli insediamenti rurali e dell'edilizia in materiale deperibile a livello europeo). Fino ad ora abbiamo portato a compimento i soli archivi per le informazioni provenienti dalle indagini stratigrafiche, presentati in questa sede.

3 – L’architettura dei dati

L’organizzazione dei dati si fonda su un’architettura gerarchica dove il livello più alto è rappresentato dal sito (in relazione 1:1 o 1:N con tutti gli altri archivi), concetto di base per la gestione del dato in archeologia secondo la progettazione logica delle *Linee Guida*; sullo stesso piano della scheda di sito si colloca, nel nostro caso, la tabella relativa ai progetti di ricerca contenente i dati fondamentali inerenti le indagini stratigrafiche (archivio *Scavi*). Scendendo nell’albero gerarchico troviamo gli archivi relativi alle suddivisioni spaziali dello scavo, in particolare le aree, i settori ed i quadrati indagati. Allo stesso livello (rapporto diretto di N:1 con gli archivi *Siti* e *Scavi*) si pongono anche le tabelle inerenti le diverse fasi interpretative; si sono previste schede per la catastazione delle periodizzazioni dello scavo (con relative datazioni) e delle strutture scavate (edifici, strutture produttive, spazi aperti, viabilità, ecc.). I dati stratigrafici, ed in particolare le attività in rapporto 1:N con le schede US, sono posti al livello successivo. Al grado più basso dell’architettura si collocano infine tutte le tabelle dei reperti (ceramici, vitrei, metallici, numismatici, osteologici umani, osteologici animali, altri), con le eventuali tabelle secondarie (ad esempio l’archivio dei reperti dentali riferiti agli scheletri) e i file necessari per le quantificazioni. **FIGURA 1**

Un nodo importante è rappresentato dalla definizione degli identificatori relazionali. I frequenti interventi sui dati catastati e soprattutto la necessità di importazioni ed esportazioni continue durante la fase di data entry (nel nostro caso ad esempio non abbiamo a disposizione una rete locale sul cantiere di scavo, perciò ciascun responsabile di area inserisce la propria documentazione in un file che, al termine della campagna, dovrà confluire nel database completo) ha suggerito di evitare l’uso di numeri progressivi; si è invece optato per campi calcolati, che definissero univocamente i dati. In particolare gli identificatori si compongono di una stringa costituita dai diversi tipi di informazioni necessarie a creare i criteri di univocità, preceduti da una sigla di tre caratteri maiuscoli che identifica il tipo di informazione e separati da un carattere convenzionale neutro (nel nostro caso “%”). Ad esempio per le unità stratigrafiche, l’univocità è garantita dal numero di US (con eventuale bis), dal numero di area e dalla sigla dello scavo; l’identificatore relativo all’US 352 dell’area 2 dello scavo di Poggio Imperiale sarà quindi “SCAPI%ARE2%NUS352”, dove “SCA” sta per scavo, “PI” per Poggio Imperiale, “ARE” per area, “NUS” per numero US. L’identificatore per l’area di scavo è costituito, mantenendo lo stesso esempio, dai primi due elementi della stringa precedente (“SCAPI%ARE2”), mentre il solo elemento iniziale (“SCAPI”) rappresenta l’identificatore relazionale di scavo. Allo stesso modo si sono creati tutti gli identificatori necessari per la gestione del database; i più importanti sono: IDscavo, IDarea (scavo & area), IDus (scavo & area & US), IDattività (scavo & area & attività), IDsettore (scavo & area & settore), IDquadrato (scavo & area & settore & quadrato), IDstruttura (scavo & struttura), IDperiodo (scavo & periodo), IDfase (scavo e periodo e fase), IDdefstr (scavo & definizione stratigrafica US), IDdefint (scavo & definizione interpretata US).

Per quanto riguarda la struttura dei singoli archivi, ci siamo basati ancora una volta sulle schede ministeriali, spesso ritoccate per soddisfare le nuove esigenze derivate dall’implementazione del sistema di gestione informatizzato e dall’approfondimento dei livelli interpretativi della ricerca; in questa sede richiederebbe troppo spazio analizzare dettagliatamente, campo per campo, i singoli archivi del database. Va comunque sottolineato che le modifiche apportate si riferiscono principalmente alle necessità imposte dall’utilizzo della base GIS e del sistema OpenArcheo. In questo senso si è provveduto all’aggiunta di

campi specifici, rappresentanti stringhe di identificatori sulla quale si basano le relazioni del sistema e alcune ricerche sulla base GIS.

Ai fini dell'implementazione di OpenArcheo si è rivelato sufficiente creare un indice relazionale generico, in relazione con tutte le tabelle attraverso un identificatore da importarsi ad ogni collegamento fra i dati e costituito da una stringa che concatena tutte le chiavi di ricerca previste dal sistema. Ovviamente in ogni tabella si ritrova lo stesso identificatore (denominato *IDOpenArcheo*) che consente di rendere operativa la relazione. In sostanza se si vuole accedere alle schede US di un insieme di oggetti selezionati nella base GIS il sistema esporterà i relativi identificatori dalla tabella interna, effettuerà quindi un'importazione nell'indice di OpenArcheo e visualizzerà relazionalmente i dati riferiti alle US, catastati nell'apposito archivio.

La base GIS ha invece previsto la creazione di alcuni campi dedicati per tradurre le informazioni del database in stringhe agili allo scopo di ottimizzare la ricerca degli oggetti grafici senza incorrere nella perdita di informazioni utili. In particolare il problema si è presentato per i campi multipli, un'opzione di FileMaker Pro che consente di inserire più valori dello stesso tipo in un solo campo, evitando la creazione di tabelle da gestire relazionalmente nel caso di dati molto semplici. Nel nostro caso la struttura, una delle chiavi di ricerca indispensabili all'interno del GIS, è rappresentata da un campo multiplo nell'archivio delle unità stratigrafiche; ogni US può infatti appartenere a più di una struttura (ad esempio un piano di calpestio esterno può essere ricollegato diversi edifici). Si è pertanto reso necessario aggiungere un campo, utilizzato solamente in fase di esportazione dei dati per la base GIS, riempito attraverso una routine di programmazione con tutti i singoli valori del campo multiplo concatenati e separati da un carattere convenzionale. Un problema di maggiore rilevanza è rappresentato dall'associazione di una cronologia agli oggetti grafici. In realtà la messa in fase delle US secondo la teoria stratigrafica rende piuttosto complessa la rappresentazione della continuità nell'uso attraverso più periodi di frequentazione attestati su un sito; limitandoci ad un esempio semplice, l'arco cronologico relativo all'utilizzo di un muro (ma lo stesso concetto vale per un livello pavimentale, un piano di calpestio, una buca di palo, ecc.) costruito in un periodo ed in vita per uno o più periodi successivi si ricava solamente dalla datazione assegnata alla sua interfaccia di distruzione. Ciò crea qualche problema nell'output delle piante di fase all'interno del GIS. La soluzione è rappresentata dall'aggiunta di un codice relativo all'eventuale uso continuato di una particolare evidenza, attraverso la quale sia possibile visualizzare, all'interno del GIS, tutti gli oggetti effettivamente esistenti in un determinato periodo.

4 – La normalizzazione del linguaggio e i controlli sulla consistenza del dato

Un ulteriore aspetto che assume valore fondamentale nella costruzione di una base di dati relazionale è costituito dalla normalizzazione del linguaggio utilizzato durante l'immissione, soprattutto nei campi di sintesi delle informazioni. I linguaggi di comunicazione verbale e scritta utilizzati dall'uomo presentano infatti un'infinità di sfumature dipendenti dal contesto, dal tono della comunicazione, dalla specificità del soggetto, ecc. L'elaboratore non è ovviamente in grado di cogliere tali differenze. Anche se problematiche di questo genere sono affrontate con sempre maggiore successo in alcuni settori specialistici afferenti alla disciplina dell'intelligenza artificiale e della realizzazione di database evoluti non esiste finora una macchina in grado di comprendere ed interpretare, anche solo a livelli accettabili, il linguaggio verbale e scritto dell'uomo nel suo complesso. Si rende pertanto necessario, durante la progettazione di un DBMS, prevedere lo standard del linguaggio usato nella compilazione della base di dati attraverso la definizione di vocabolari e

thesaurus per i valori dei *campi*. Dalla chiarezza formale e dalla completezza di questi vocabolari dipende in gran parte la leggibilità e l'interpretabilità di una base di dati. In sostanza si tratta, insieme alla progettazione dell'architettura relazionale, dello sforzo maggiore implicato nella costruzione di un database relazionale efficiente. Nella creazione dei vocabolari si rende infatti necessario contemplare tutte le possibilità e le variabili che si potranno presentare durante la compilazione di un archivio, basandosi su criteri chiaramente formulati e sul grado di dettaglio che si intende raggiungere durante la schedatura informatizzata; nel caso di basi di dati complesse la definizione dei *thesaurus* rappresenta di per sé un grande momento di riflessione e sintesi dei processi che producono l'insieme di dati i quali confluiranno negli archivi; in questo senso la creazione di vocabolari è paragonabile al processo di tipologizzazione relativo ad un insieme di reperti provenienti da un'indagine archeologica.

Nella progettazione della base di dati per lo scavo archeologico abbiamo deciso di implementare liste valori vincolate, da assegnare solamente ai campi più rilevanti per lo scambio dei dati. Ad esempio per l'archivio US abbiamo previsto l'uso di liste valori vincolate sui campi *Definizione stratigrafica*, *Definizione interpretata*, *Colore*, *Consistenza*, *Stato di conservazione*, *Affidabilità*, *Formazione*, *Campionature*, *Inclusi*; per i reperti invece l'uso è solitamente ristretto ai campi *Classe*, *Forma*, *Tecnica di lavorazione*, *Decorazione*, oltre ad alcuni campi specificatamente pertinenti a ciascun tipo di reperto (*Zecca*, *Materia e Valore nominale* per le monete; *Specie*, *Razza*, *Identificatore anatomico* per i reperti osteologici animali, ecc.). In sostanza si tratta di liste valori immutabili dall'utente e aggiornate periodicamente nel corso di riunioni fra i responsabili delle varie indagini in corso. Tutte le altre liste valori, legate al singolo progetto di ricerca (aree, settori, quadrati, strutture, anno di scavo, periodi e fasi, tipologie dei reperti, impasti ceramici, ecc.) o pertinenti ai singoli archivi ma di importanza minore ai fini della normalizzazione sono direttamente modificabili dall'utente nel corso del data entry o attraverso l'apposito script dell'ambiente *Manutenzione* (si veda più avanti).

Parte integrante della struttura degli archivi sono anche i controlli sulla consistenza dei dati, cui è stata dedicata particolare attenzione in fase progettuale e realizzativa del database; in questo senso FileMaker permette di implementare in modo semplice accertamenti anche complessi sui valori immessi al fine di rendere il data entry più rapido ed accurato. Ovviamente i controlli coinvolgono tutti gli archivi del database, e prevedono messaggi personalizzati in caso di operazioni non consentite. La verifica più semplice, operata su tutti i campi, riguarda il tipo di dato (numero, testo, data, ecc.) cui ci si deve attenere per l'inserimento di un valore (ad esempio nel campo US è consentito inserire solo ed esclusivamente valori numerici). Per i campi che prevedono l'inserimento attraverso una lista valori sono stati realizzati appositi controlli (differenziati a seconda che si tratti di liste vincolate o libere). Un'attenzione particolare è stata ovviamente dedicata ai campi coinvolti nella definizione degli identificatori relazionali; in questi casi ai controlli finora descritti si aggiungono quello sull'inserimento obbligatorio e sull'eventuale unicità del dato. In casi particolari si è reso necessario stabilire un intervallo di valori consentiti per l'inserimento o prevedere la verifica in base a condizioni calcolate (ad esempio il numero di frammenti ceramici nell'archivio apposito deve essere maggiore di zero, l'anno di scavo di ciascuna US deve essere compreso fra l'anno iniziale del progetto e l'anno in corso, il grado di affidabilità interpretativa di un'US, secondo una convenzione da noi stabilita, deve essere espresso in valori numerici compresi fra uno e cinque, ecc.). In sostanza, nella definizione delle verifiche sulla consistenza dei dati, occorre trovare una soluzione di compromesso, in larga parte connessa alla definizione del grado di dettaglio dei dati, che permetta di effettuare un data

entry corretto senza tuttavia rendere i controlli talmente severi da rallentare anziché facilitare l'immissione dei dati.

5 – La personalizzazione del database: interfaccia utente e programmazione

L'aggiunta di routines di programmazione, appositamente create per l'espletamento di alcune funzioni specifiche legate ai singoli archivi e integrate nell'interfaccia utente personalizzata descritta più avanti, rappresenta una delle caratteristiche principali del nostro database. In particolare sono state previste funzioni per:

- la navigazione lineare fra le schede attraverso script che consentono di andare al record successivo e precedente, al cinquantesimo record successivo o precedente, al primo o all'ultimo record, ad un record specifico attraverso l'inserimento del suo numero progressivo;
- la navigazione relazionale fra archivi e record attraverso l'uso di routines che permettono di scegliere il tipo di dati correlati da visualizzare nell'indice, una navigazione incrociata sia fra i record dell'indice, sia fra i record correlati visualizzati all'interno di finestre apposite dette portali;
- l'automatizzazione delle operazioni di creazione, duplicazione, eliminazione e ordinamento dei record;
- l'automatizzazione delle *query* di ricerca, riproducendo peraltro un sistema quasi identico alle normali *richieste di ricerca* previste da FileMaker Pro; lo script che abbiamo sviluppato impedisce la modificazione delle liste valori vincolate e, allo stesso tempo, permette l'utilizzo di valori non compresi nelle liste come parametro di ricerca;
- le operazioni relative ai *task* di stampa per i quali sono stati previsti diversi formati tabellari ed estesi, una funzione di anteprima a video e la possibilità di impostare i parametri della stampante selezionata;
- una funzione di *marking* dei record (non prevista di serie su FileMaker Pro, a differenza di altre applicazioni simili); la combinazione delle potenzialità di scripting e delle capacità relazionali ha reso possibile l'implementazione di questa funzionalità, comprendente anche routines per la ricerca di tutte le schede segnate e l'eliminazione automatica dei marker;
- la quantificazione dei reperti attraverso la specificazione di parametri stratigrafici (per US, periodo, struttura, ecc.) e pertinenti al tipo di reperti (ad esempio nel caso della ceramica è possibile quantificare per classe, classe e forma, classe e forma e impasto, ecc.). Si tratta di routines abbastanza complesse che contemplano l'utilizzo di due file correlati appositamente dedicati al riassunto dei dati e alla loro quantificazione e prevedono anche l'esportazione dei risultati in RTF (*Rich Text Format*), pronti per un'eventuale pubblicazione o comunque per un utilizzo all'interno di documenti di testo. L'utilità di questa funzione, in termini di produttività del lavoro, è impressionante; operazioni che richiedevano anche settimane vengono effettuate in pochi minuti. Alcune prove effettuate sull'archivio ceramico di Poggio Imperiale a Poggibonsi, montato su un Power Macintosh G3 a 266MHz dimostrano la polverizzazione dei tempi di elaborazione : una quantificazione per US, classe, forma ed impasto su 202 schede ceramiche è stata completata in 26 secondi (per l'eventuale esportazione in RTF bisogna aggiungere 6 secondi), mentre la stessa routine lanciata su 4753 schede (relative a 32548 frammenti per un peso totale di oltre 499 kg) impiega 10 minuti e 42 secondi (cui si aggiungono 1':34" per l'eventuale esportazione in RTF);
- l'inserimento di caratteri speciali nel testo dei campi (ä, ö, ü, ß, ecc.).

Vi sono quindi una serie di script non accessibili direttamente dai singoli archivi o dagli indici relazionali, ma di pertinenza dell'ambiente *Manutenzione*, descritto più avanti:

- aggiornamento degli indici relazionali: aggiorna uno o tutti gli indici relazionali importando il campo identificatore dai singoli archivi;

- aggiornamento dei riferimenti relazionali: si tratta di campi di un archivio secondario i cui valori sono copiati dall'archivio master al momento del loro primo inserimento o dell'eventuale aggiornamento; sono necessari per il funzionamento di alcune relazioni in quanto FileMaker Pro non prevede la visibilità dei dati degli archivi non correlati direttamente. Ciò crea problemi nell'accesso da parte di archivi *master* ad archivi secondari collocati a livelli inferiori; ad esempio dall'archivio *US*, attraverso l'identificatore *IDus*, è possibile vedere tutti i record correlati dell'archivio *Reperti osteologici umani*, ma non quelli dell'archivio *Reperti dentali* (in relazione con il solo archivio dei resti scheletrici attraverso il campo *Idscheletro*);
 - recupero degli archivi danneggiati: recupera i dati e la struttura di archivi danneggiati in seguito ad un malfunzionamento del calcolatore o all'interruzione dell'alimentazione; la funzione recupero degli archivi compatta inoltre i file eliminando lo spazio inutilizzato che si crea durante l'uso degli archivi;
 - definizione e modifica delle liste valori: funzione tuttora in corso di elaborazione (esiste già una versione preliminare), relativa alla gestione delle liste valori per la normalizzazione del linguaggio; la versione finale dell'opzione *Aggiorna Liste Valori* nell'ambiente *Manutenzione* del database *Carta Archeologica* permetterà quindi di effettuare cambiamenti solo sulle liste non vincolate; per le liste vincolate invece sarà possibile registrare gli appunti e le proposte delle modifiche da concordare durante le riunioni di cui si è già detto.
- E' infine stata prevista una gestione degli eventuali errori con messaggi personalizzati, al fine di guidare l'utente nel compimento delle operazioni desiderate.

Per completare la descrizione della struttura della struttura del DBMS sono necessari alcuni cenni all'interfaccia utente, interamente realizzata presso il Laboratorio sia per la parte di analisi e progettazione, sia per la realizzazione pratica; si tratta di un'implementazione scritta nel linguaggio residente di FileMaker Pro che sfrutta appieno le potenti funzioni di scripting e di gestione della grafica nella creazione di sfondi e pulsantiere messe a disposizione dall'applicazione. I criteri di base cui ci siamo attenuti hanno riguardato la personalizzazione delle funzioni principali di FileMaker Pro, facilitando ulteriormente l'uso del software attraverso controlli tarati specificatamente sui singoli archivi e percorsi guidati ed obbligati nell'espletamento di alcune operazioni (già descritte sopra nella parte dedicata agli script). Un help in linea, costituito da un prodotto multimediale realizzato con Macromedia Director e contestualizzato alle funzioni del database attraverso routines di programmazione interne, è previsto per la versione finale del DBMS *Carta archeologica* e sarà allora realizzato anche per gli archivi oggetto di questa trattazione. L'elaborazione dell'interfaccia utente ha infine previsto la personalizzazione della veste grafica attraverso l'uso di sfondi e icone, e la creazione di menù e pulsantiere per l'uso specifico dei diversi archivi. Si sono resi operativi tre diversi ambienti, corrispondenti a tre modi di utilizzo della base di dati:

Singoli Archivi

Ambiente di lavoro, nel quale è possibile intervenire sui record dei singoli archivi che compongono la base di dati; dal menù principale si accede alla schermata dei singoli archivi e di qui, attraverso una lista a tendina, all'archivio che si intende adoperare. Nell'ambiente *Singoli Archivi* si possono creare nuovi record, duplicare record già esistenti, modificare i dati inseriti nell'archivio, eseguire richieste di ricerca che coinvolgono i campi dell'archivio attivo. In alcune occasioni è previsto l'accesso ai dati di alcuni archivi secondari, direttamente in relazione con la tabella sulla quale si sta operando. E' questo il caso dell'archivio *Reperti dentali* cui si accede direttamente (durante il *data entry*, oppure nel caso di richieste di ricerca, ordinamenti, eliminazione o modifica dei record) dall'archivio *Reperti*

osteologici umani; anche l'archivio *Misure* è visibile dalla tabella *Reperti osteologici animali*. In entrambi i casi l'accesso avviene attraverso l'uso dei *portali*, finestre apposite per la visualizzazione dei dati correlati in FileMaker Pro.**FIGURA 2**

Ambiente relazionale

Rappresenta l'ambiente privilegiato per la consultazione dell'intero DBMS; attraverso l'uso di indici relazionali tematici è possibile reperire in modo immediato le informazioni correlate. Abbiamo finora previsto indici per i singoli scavi, le aree di ciascuno scavo, le aree divise per periodo e fase, i settori, i quadrati, le strutture interpretate, le attività, le unità stratigrafiche, le definizioni US stratigrafiche, le definizioni US interpretate, i periodi, i periodi con le fasi. Oltre a questi è possibile crearne altri importando l'identificatore voluto in un apposito indice vuoto, facente parte integrante della base di dati.**FIGURA 3**

Manutenzione

Contiene tutte le routines di manutenzione riferite al database, realizzate nella maggior parte utilizzando il linguaggio di scripting residente (tranne alcuni casi nei quali si è dovuto ricorrere ad un'integrazione con AppleScript) e già descritte in precedenza.

6 – Il database di scavo. Considerazioni conclusive

Gli archivi che abbiamo descritto finora si configurano come uno strumento estremamente efficace; l'uso intensivo e pluriennale di sistemi per l'archiviazione relazionale dei dati di scavo ha rappresentato una delle prime e più evidenti dimostrazioni relative all'efficienza del mezzo informatico applicato all'archeologia. D'altronde va sottolineato come, per ottenere un prodotto veramente valido, non sia possibile rinunciare all'intelligenza archeologica. In questo senso riteniamo migliore un database forse meno coerente da un punto di vista informatico, ma aperto ed aderente alle mutevoli esigenze della ricerca; nel nostro caso siamo riusciti a ricavare maggiori vantaggi dal processo di informatizzazione del dato archeologico in quanto l'architettura del nostro database (e dell'intero sistema in generale), essendo stata realizzata direttamente da archeologi, si conforma molto più alla logica di scavo piuttosto che ai concetti di analisi e progettazione informatica.

Come più volte sottolineato gli archivi che abbiamo presentato in questa sede non sono da intendersi come un prodotto finito, ma piuttosto come uno stage della realizzazione che, per la coerenza interna e la possibilità di gestire tutti i tipi di dato alfanumerico prodotti da un'indagine stratigrafica, può essere considerata una release valida da testare su un numero rilevante di casi. In realtà lo sviluppo degli archivi continua a procedere, adeguandosi alle novità immesse sul mercato del software e dell'hardware e agli spunti suggeriti dall'avanzamento della ricerca.

L'uscita della versione 5 di FileMaker (avvenuta in occasione di SMAU 1999) che prevede il potenziamento delle capacità di scripting e l'implementazione dei driver ODBC e di funzioni SQL Server implicherà sicuramente una revisione parziale ed un'aggiunta di funzioni al database.

Un altro aspetto che prevediamo di potenziare riguarda la gestione unificata delle categorie interpretative per tutte le tabelle che compongono la base di dati; intendiamo in sostanza creare un archivio apposito che consenta di catastare tutti gli aspetti soggettivi della ricerca, in relazione a ciascun tipo di informazione presente negli archivi.

Un sistema di archivi aperto e progettato in funzione della ricerca va inteso quindi come un grande contenitore di dati alfanumerici in continua evoluzione. Una caratteristica fondamentale è comunque rappresentata dalla possibilità, in qualsiasi momento, di estrapolare notizie riassuntive con caratteri di standardizzazione (ad esempio adeguandosi alle schedature previste dall'ICCD) che siano fruibili dagli enti preposti alla salvaguardia del patrimonio

archeologico o permettano uno scambio di informazioni su basi paritetiche con altri progetti di ricerca; il database proposto in questa sede può rappresentare infatti un modello di base per l'organizzazione dei dati, facilmente esportabile ed adeguabile ad ogni tipo di scavo.

Bibliografia

- BIANCHIMANI A., PARRA M.C. 1991, *NIKE: progetto di una base di dati archeologica*, "Archeologia e Calcolatori", 2, pp. 179-203
- D'ANDRIA F. 1997, *Metodologie di catalogazione dei beni archeologici*, Lecce-Bari (Beni Archeologici - Conoscenza e Tecnologie, Quaderno 1.1)
- GUERMANDI M.P. 1989, *Aladino: uno strumento per la gestione dei dati di scavo*, "Bollettino d'informazioni del Centro di Elaborazione Automatica di Dati e Documenti Storico Artistici", X, pp. 21-56
- LINEE GUIDA CARTA ARCHEOLOGICA 1998, *Linee guida per la redazione della Carta Archeologica della Toscana*. Versione del 11/9/1998, Firenze, Regione Toscana
- MOSCATI P. 1987, *Archeologia e calcolatori*, Firenze
- MOSCATI P. 1998, *GIS applications in italian archaeology*, "Archeologia e Calcolatori", 9, pp. 191-236
- RICCI A. 1985, *Ricognizione di superficie e scavo:dalle schede cartacee ad un sistema automatico al servizio dell'archeologia sul campo: il protoripo Argo*, in *Archeologia e Informatica*, Atti del Convegno. Roma 3-4-5 marzo 1988, Roma, 77-83
- VALENTI M. (a cura di) 1999, *Carta Archeologica della Provincia di Siena. Vol. III, La Val d'Elsa (Colle di Val d'Elsa e Poggibonsi)*, Siena
- VALENTI M. 1998b, *Computer Science and the management of an archaeological excavation: the Poggio Imperiale Project*, "Archaeological Computing Newsletter", 50 (Spring), pp. 13-20
- VALENTI M., 1998a, *La gestione informatica del dato; percorsi ed evoluzioni nell'attività della cattedra di Archeologia Medievale del Dipartimento di Archeologia e Storia delle Arti- Sezione Archeologica dell'Università di Siena*, "Archeologia e Calcolatori", 9, pp.305-329