

Rapporti tecnici

INGV

Trasferimento di dati tabellari da un database MySQL a un geodatabase di ArcGis® utilizzando lo standard XML

91



Direttore

Enzo Boschi

Editorial Board

Raffaele Azzaro (CT)

Sara Barsotti (PI)

Mario Castellano (NA)

Viviana Castelli (BO)

Anna Grazia Chiodetti (AC)

Rosa Anna Corsaro (CT)

Luigi Cucci (RM1)

Mauro Di Vito (NA)

Marcello Liotta (PA)

Lucia Margheriti (CNT)

Simona Masina (BO)

Nicola Pagliuca (RM1)

Salvatore Stramondo (CNT)

Andrea Tertulliani - coordinatore (RM1)

Aldo Winkler (RM2)

Gaetano Zonno (MI)

Segreteria di Redazione

Francesca Di Stefano - coordinatore

Tel. +39 06 51860068

Fax +39 06 36915617

Rossella Celi

Tel. +39 06 51860055

Fax +39 06 36915617

redazionecen@ingv.it



Rapporti tecnici INGV

TRASFERIMENTO DI DATI TABELLARI DAL DATABASE MYSQL EURO-MEDITERRANEAN PALEOTSUNAMI A UN GEODATABASE DI ARCGIS® UTILIZZANDO LO STANDARD XML

Antonio Patera e Manuela Sbarra

INGV (Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Sismologia e Tettonofisica)

91

Indice

	Introduzione	5
1.	Organizzazione dei dati	7
1.1	Attributi descrittivi delle tabelle e relazioni	9
2.	Architettura del geodatabase	11
3.	Schema XML del geodatabase	12
4.	Tipi di documenti XML del geodatabase	13
4.1	Documento Workspace	13
4.1.1	WorkspaceDefinition e WorkspaceData	13
4.1.1.1	Data Elements: DataElement - DEDataset - DETable	14
4.1.1.2	RecordSet	17
4.2	Passaggio dal database MySQL allo Schema XML di un geodatabase	18
4.2.1	Tabelle – Attributi	18
4.2.2	Valori degli Attributi	21
5.	Procedura di importazione dei dati dal file XML al geodatabase	23
	Ringraziamenti	29
	Bibliografia	29

Introduzione

Un sistema informativo geografico o GIS (dall'inglese *Geographical Information System*) è l'integrazione di sei componenti di base (utenti, dati, hardware, software, procedure e reti) che permettono nel loro insieme di catturare, archiviare, interrogare, analizzare, visualizzare e restituire l'informazione geografica (ESRI, 2006a). I dati geografici vengono comunemente archiviati in un GIS sotto forma di livelli informativi (in inglese *layer*) contenenti ciascuno uno specifico tipo di informazione geografica (Fig. 1).

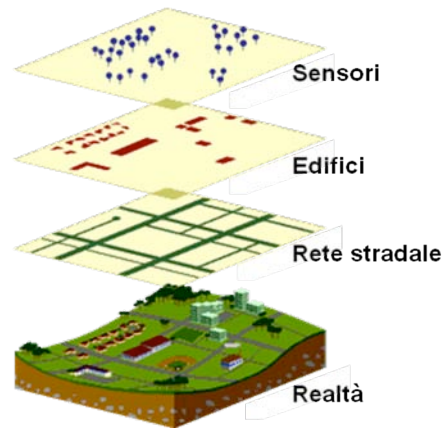


Figura 1 Un esempio di archiviazione di dati geografici in un sistema informativo geografico.

ArcGIS®, prodotto software sviluppato dalla ESRI, è un software GIS che offre la possibilità di archiviare dati geografici e altri tipi di oggetti all'interno di database relazionali chiamati *geodatabase*, contenitori forniti di funzionalità specifiche per la gestione e il mantenimento dell'integrità dei dati

Una delle caratteristiche principali di questo software è l'abilità di accedere a dati archiviati in numerosi formati (Fig. 2) e provenienti da più fonti, quali differenti tipi di database, dati archiviati in file su disco e servizi GIS su Web. L'applicazione fornisce inoltre gli strumenti per la conversione dei dati tra i diversi formati supportati.

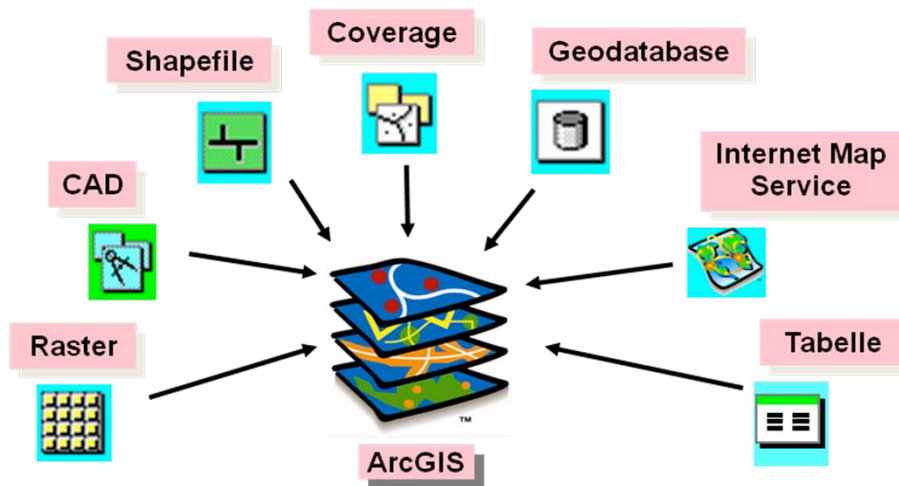


Figura 2 Tipi di formati di dati utilizzabili in ArcGIS.

Il geodatabase è la struttura di archiviazione di ArcGIS ed è il formato nativo di dati utilizzato per l'editing, la memorizzazione e la gestione delle informazioni geografiche. Un geodatabase è una collezione di insiemi di oggetti geografici e di altro tipo (Fig. 3) contenuta nella stessa cartella su disco (*file geodatabase*), o in un database Microsoft® Access (*personal geodatabase*) o in database relazionale

multiutente quale Oracle®, Microsoft® SQL Server™, IBM® DB2™ o Informix™ (*enterprise geodatabase*).

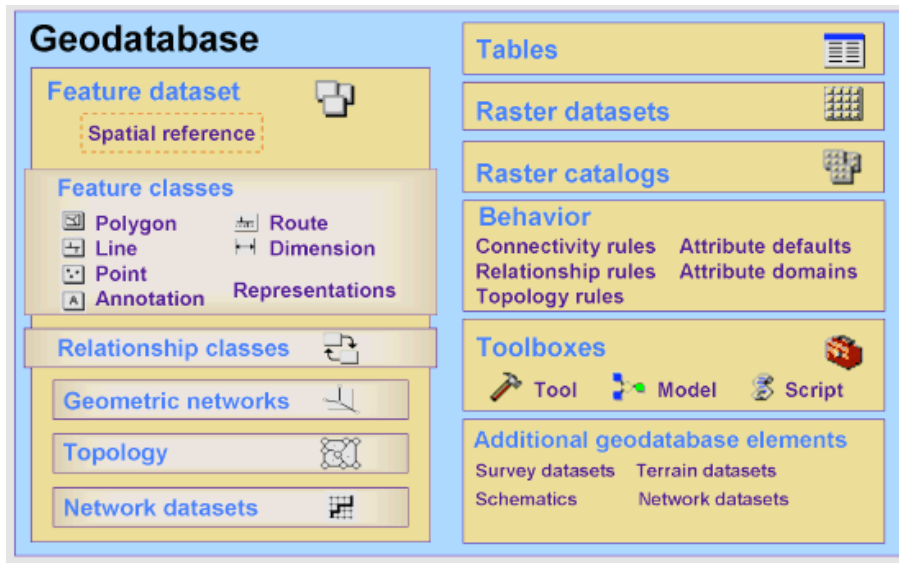


Figura 3 Tipi di oggetti archiviabili in un geodatabase.

Il geodatabase può avere dimensioni variabili dai piccoli file di alcuni Megabyte utilizzabili da un singolo utente, fino ai grandi sistemi aziendali di vari Terabyte cui possono avere accesso contemporaneamente più utenti. Essendo operativo su sistemi DBMS risponde a tutta la serie di funzioni e di istruzioni SQL (*Structured o Standard Query Language*) tipiche di tali sistemi.

Per il lavoro qui presentato è stato scelto come formato di archiviazione il *personal geodatabase*, nel quale i dati sono memorizzati all'interno di un database di Microsoft Access. Questo tipo di database può avere dimensione massima di 2GB, anche se il limite massimo reale della dimensione del file prima che le prestazioni subiscano un drastico degrado è di 250-500MB. Il personal geodatabase non è così veloce, efficiente e scalabile come gli altri tipi di geodatabase, però permette di editare e visualizzare gli attributi direttamente utilizzando Microsoft Access. In questo formato le transazioni lunghe e il *versioning* non sono supportati.

1. Organizzazione dei dati del geodatabase

Il geodatabase è un contenitore di oggetti di vario tipo, tra i quali quelli che rivestono la maggiore importanza dal punto di vista geografico sono le feature class e i raster e le relative tabelle associate:

- **Feature class:** livello informativo contenente collezioni ordinate di oggetti (elementi) geografici vettoriali aventi la stessa rappresentazione spaziale (punti, linee o poligoni) e gli stessi attributi descrittivi. I tre tipi di feature class più comunemente utilizzati nel geodatabase contengono punti, o linee o poligoni (Fig. 4). Un altro tipo particolare di feature class sono le annotazioni (termine utilizzato nel geodatabase per il testo georeferenziato).

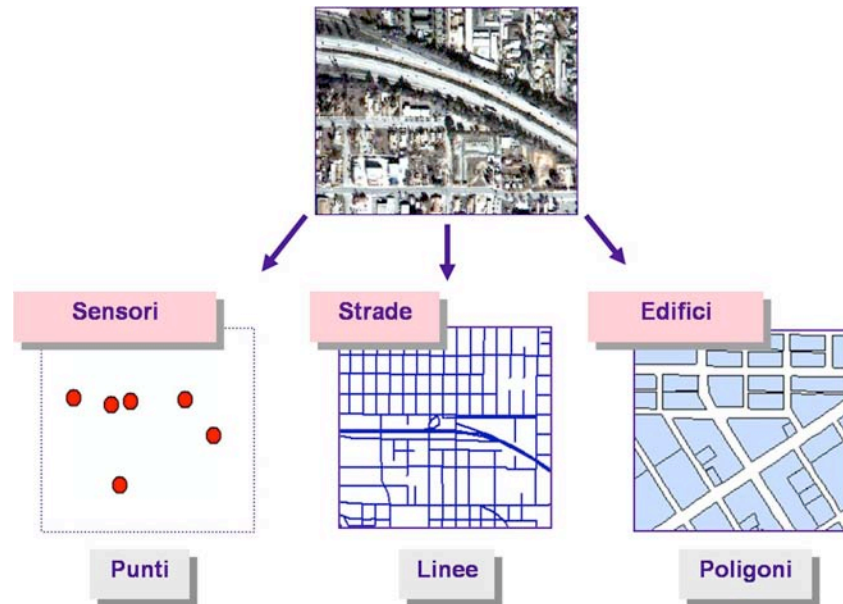


Figura 4 I tre tipi principali di feature class archiviabili all'interno di un geodatabase.

- **Raster:** modello di dati spaziali che rappresenta un oggetto o un fenomeno attraverso una matrice di celle uguali di forma quadrata, ordinate in righe e colonne. Ogni cella è identificata con un valore di un attributo quantitativo o qualitativo e dalle coordinate della posizione. Un'immagine da satellite, un modello digitale del terreno o il prodotto di una scansione sono un tipico esempio di dato raster (Fig. 5).

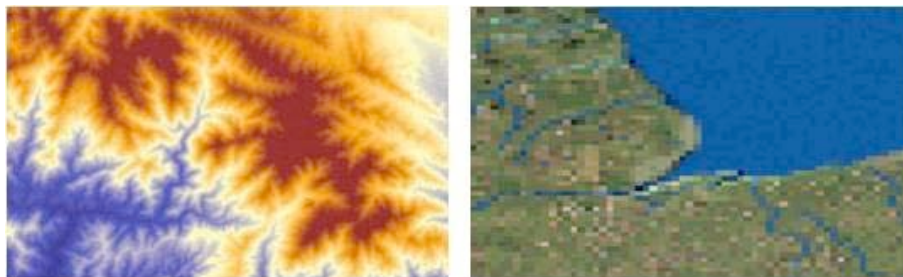


Figura 5 Due esempi di dati raster. A sinistra, un modello digitale del terreno. A destra un'immagine da satellite.

- **Tabelle:** in un database relazionale rappresentano dati organizzati orizzontalmente in righe e verticalmente in colonne o campi (Fig. 6). Una tabella ha un numero definito di colonne, ma può avere un qualsiasi numero di righe. Le tabelle associate alle feature class o ai raster contengono le informazioni descrittive degli oggetti geografici e sono dette tabelle degli attributi.

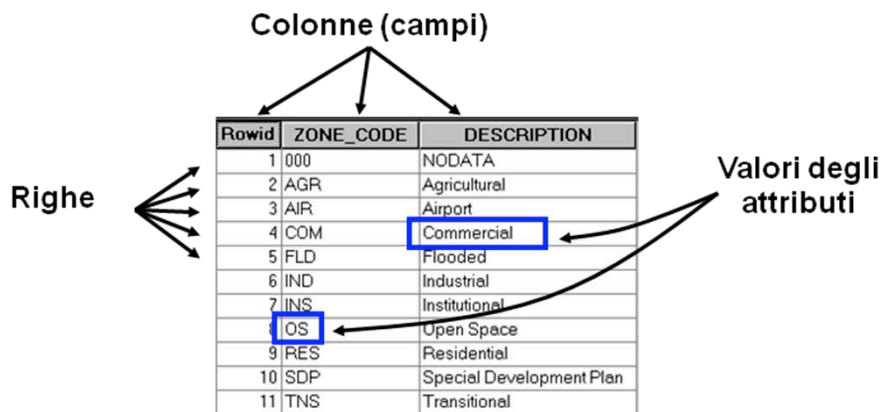


Figura 6 Un esempio di tabella.

Il primo passo nella costruzione di un GIS è la raccolta dei dati geografici e la loro archiviazione nel geodatabase nel formato più idoneo al tipo di dato raccolto. È inoltre possibile aggiungere al geodatabase funzionalità avanzate come la topologia, le relazioni tra le tabelle, le reti, i sottotipi e i domini. Queste funzionalità, insieme alle definizioni delle feature class e delle tabelle, rappresentano lo *schema del geodatabase* (Fig. 7) e sono fondamentali nelle applicazioni GIS perché permettono il mantenimento dell'integrità dei dati e delle relazioni spaziali tra gli oggetti archiviati nel geodatabase (per esempio, ogni riga ha sempre le stesse colonne e un dominio di valori indica i valori validi o gli intervalli di valori validi per una colonna).

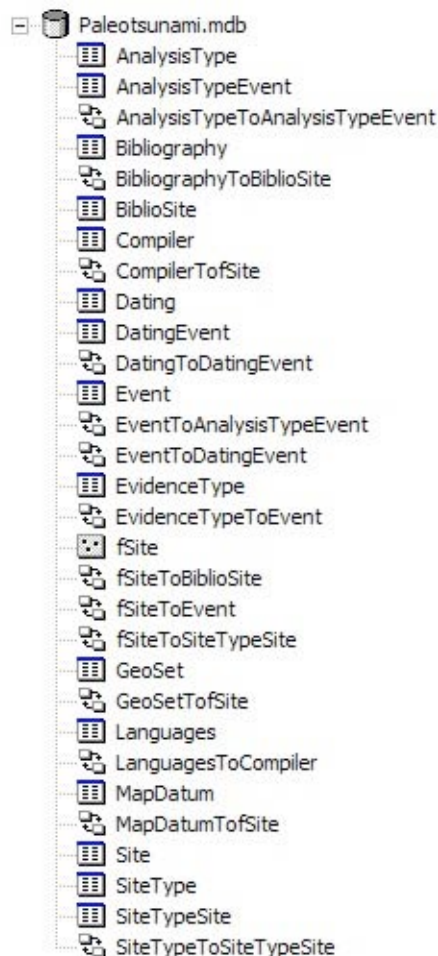


Figura 7 Lo schema del geodatabase *Euro-Mediterranean Paleotsunami*, con la feature class *fSite*, le tabelle a essa associate e le classi di relazione.

1.1 Attributi descrittivi delle tabelle e relazioni

L'organizzazione degli attributi descrittivi degli oggetti geografici in tabelle (Fig. 8) gioca un ruolo chiave nei modelli di dati GIS, così come nelle tradizionali applicazioni per i database.

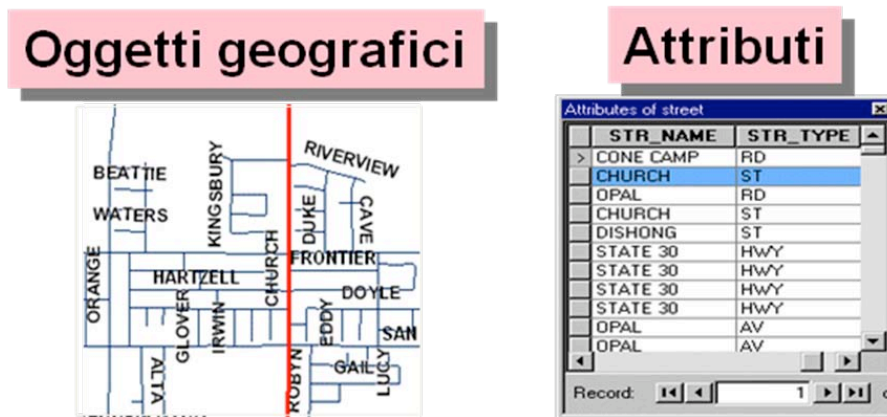


Figura 8 In un GIS a ogni oggetto geografico sono collegate le informazioni descrittive contenute nella tabella degli attributi.

Inoltre, più tabelle possono essere associate tra di loro o alle tabelle degli attributi degli elementi geografici attraverso un campo in comune, la *chiave*, con un'operazione che viene detta di relazione e che permette di associare le righe di una tabella con le righe di un'altra tabella (Fig. 9). Esistono vari tipi di relazioni, in base al numero degli oggetti di una tabella che possono essere associati ai corrispondenti oggetti di un'altra tabella. Esistono quindi relazioni 1:1 (uno a uno), dove a ogni riga di una tabella è associata una sola riga di un'altra tabella, 1:M (uno a molti), dove a una riga di una tabella sono associate più righe di un'altra tabella, M:1 (molti a uno), dove a più righe di una tabella è associata una riga di un'altra tabella e M:M (molti a molti), dove a più righe di una tabella sono associate più righe di un'altra tabella.

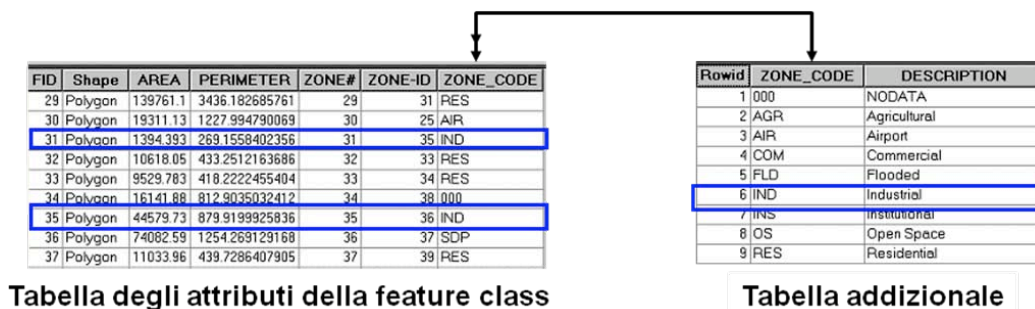


Figura 9 In una relazione le righe di una tabella sono associate alle righe di una seconda tabella tramite un campo in comune detto *chiave*.

Il modello di archiviazione del geodatabase è basato sulla tecnologia dei DBMS che fornisce un modello semplice e formale per l'archiviazione dei dati basata sui principi dei database relazionali e che permette l'interazione con le informazioni contenute nelle tabelle attraverso strumenti standard.

Nel presente lavoro, la feature class *fSite*, contenente la localizzazione dei siti dei paleotsunami, è archiviata sotto forma di tabella di un DBMS. Ogni riga rappresenta un oggetto geografico (feature), cioè un sito, visualizzato nella mappa sotto forma di oggetto puntuale. Le colonne contengono le varie caratteristiche o proprietà dell'oggetto e in particolare la colonna *Shape* archivia le informazioni sulla geometria dell'elemento geografico. La figura seguente (Fig. 10) mostra la tabella della feature class *fSite* nella quale il campo *Shape* contiene la geometria degli oggetti puntuali che rappresentano i siti.

Nelle tabelle dei DBMS sono utilizzati vari tipi di colonne per archiviare le informazioni sulla geometria degli oggetti. Generalmente si utilizzano campi BLOB (Binary Long Object), oppure un tipo spaziale specifico supportato da alcuni DBMS.

OBJECTID *	Id *	Name #	Country	Region	Province	TypeSite	MaxDistance	Lat	Lon	IdDatum *	Elev	Distance	IdGeoSet *	Shape *
1	1	Siponto	Italia	Puglia	Foggia	0	200	41.6078	15.8989	2	4	40		3 Point
2	2	Torre Degli Inglesi	Italy	Sicily	Messina	1	<Null>	38.251301	15.7004	2	6	40		7 Point
3	3	Lesina Est	Italy	Puglia	Foggia	0	350	41.903301	15.5758	2	5	1000		2 Point
4	4	Augusta	Italy	Sicily	Siracusa	0	100	37.144501	15.1357	4	6	400		8 Point
5	5	Priolo	Italy	Sicily	Siracusa	0	150	37.146301	15.2177	4	1	500		4 Point
6	6	Torre Squillace	Italy	Puglia	Lecce	0	70	41.903301	12.4924	4	2	40		8 Point
7	7	Torre Santa Sabina	Italy	Puglia	Brindisi	0	40	40.757999	17.7069	4	2	30		8 Point
8	8	Magnisi Peninsula	Italy	Sicily	Siracusa	0	350	37.1586	15.234	4	5	35		8 Point
9	9	Maddalena Peninsula	Italy	Sicily	Siracusa	0	400	37.036999	15.3143	4	5	70		8 Point
10	10	Ognina	Italy	Sicily	Siracusa	0	80	36.986688	15.2689	4	3	35		8 Point
11	12	Sarkoy	Turkey	Marmara Region	Tekirdag	1	<Null>	0	0	1	3	200		5 Point
12	13	Stromboli	Italy	Sicily	Messina	1	<Null>	36.980999	12.4924	5	15	50		8 Point
13	14	Didim	Turkey	Southwestern Turkey	Aydin	0	70	37.361401	27.2043	5	2	40		3 Point
14	15	Dalaman	Turkey	Southwestern Turkey	Mugla	0	19	36.698799	28.7623	5	2	250		5 Point
15	16	Fethiye	Turkey	Southwestern Turkey	Mugla	1	<Null>	36.6768	29.097799	1	2	140		5 Point
16	17	Taskopru	Turkey	Eastern Marmara	Yalova	0	400	40.674198	29.386299	1	20	2000		4 Point
17	18	Kucukcekmece	Turkey	Marmara	Istanbul	1	<Null>	40.9856	28.7609	1	-14	500		4 Point
18	19	Aiki	Greece	southern Corinth Gulf	Egion	0	250	38.262199	22.2561	5	2	220		4 Point
19	20	Cap De Ses Salines	Spain	Balearic Islands	SE Mallorca	0	1000	39.264999	3.0526	4	4	30		8 Point
20	21	Artà Peninsula	Spain	Balearic Islands	NE Mallorca	0	1500	39.752399	3.4134	4	7	30		8 Point
21	22	Torre S. Emiliano	Italy	Puglia	Otranto	0	1200	40.087399	18.493401	4	10	90		8 Point
22	23	Kirra	Greece	northern Corinth Gulf	Delphi	0	250	38.4305	22.4522	4	2	250		5 Point
23	24	Salinas De Cabo De Gata	Spain	Andalucia	Almeria	1	<Null>	36.7775	-2.2319	4	4	800		2 Point

Figura 10 La tabella degli attributi della feature class fSite con i principali campi.

In un geodatabase, il linguaggio SQL fornisce tutti gli strumenti per operare sullo schema del database, sulle righe, sulle colonne e sui tipi di dati (Fig. 11). I tipi di colonna (numeri, caratteri, date, BLOB, tipi spaziali, e così via) sono oggetti nell'algebra SQL. Il DBMS gestisce questi tipi di dati semplici e le tabelle, mentre la logica aggiuntiva dell'applicazione implementa il comportamento degli oggetti e le regole di integrità.

The screenshot shows the 'Attributes of fSite' window with a 'Select by Attributes' dialog box. The dialog has a 'Method' dropdown set to 'Create a new selection'. Below it, a list of fields is shown: [PElev], [Province], [Region], [SiteAdded], [SiteDesc], [Timeall]. The 'Region' field is selected. Below the list, there are operators: =, <>, Like, >, >=, And, <, <=, Or, ? (wildcard), () (parentheses), and Not. The 'Sicily' value is entered in the text box. The SQL query at the bottom reads: 'SELECT * FROM fSite WHERE: [Region] = 'Sicily''. Buttons for 'Clear', 'Verify', 'Help', 'Load...', 'Save...', 'Apply', and 'Close' are visible at the bottom of the dialog.

Figura 11 Un esempio di interrogazione sui dati attributo tramite SQL.

2. Architettura del geodatabase

Il geodatabase è implementato utilizzando un'architettura a più livelli: archiviazione e applicazione (Fig. 12). Gli aspetti riguardanti il livello di archiviazione (memorizzazione, definizione dei tipi di attributo, elaborazione delle interrogazioni processi di transazione tra più utenti) sono delegati ai DBMS, mentre il compito di costruire lo schema specifico del DBMS (definizione, regole di integrità dei dati, comportamento degli oggetti geografici) e le funzioni di elaborazione dell'informazione sono mantenute nell'applicazione e nel dominio del software (ArcGIS). In dettaglio, questo livello applicativo comprende le proprietà per le feature class, le regole topologiche, le reti, i cataloghi raster, le relazioni, i domini, ecc. Tutte le applicazioni di ArcGIS interagiscono con questo modello generico di oggetti GIS per i geodatabase e non con le istanze SQL del DBMS. I componenti software del geodatabase implementano il comportamento e le regole di integrità implicite nel modello generico e traducono le richieste di dati nell'appropriato disegno fisico del geodatabase. Il software ArcSDE fornisce la porta di accesso che permette ad ArcGIS di interagire con i DBMS (ESRI, 2006b).

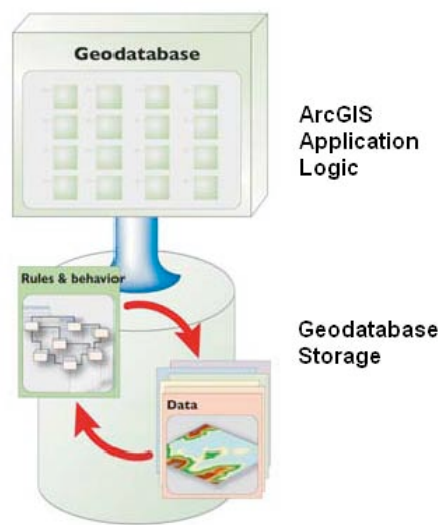


Figura 12 L'architettura del geodatabase è basata sulla memorizzazione relazionale e sulla logica dell'applicazione.

La separazione della logica del geodatabase dall'archiviazione dei dati permette di gestire numerosi tipi di file, DBMS e XML (Fig. 13). Per esempio, la capacità di accedere a quasi tutti i formati di dati geografici e tabellari è fornita dall'estensione ArcGIS Data Interoperability. Questa estensione fornisce la via d'accesso per leggere e interagire con decine di formati di dati utilizzando la logica del geodatabase.

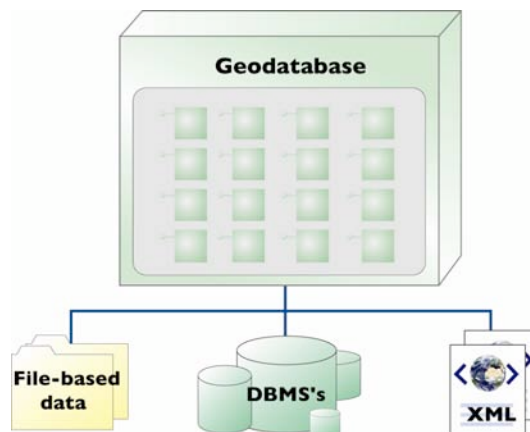


Figura 13 La separazione della logica del geodatabase dall'archiviazione dei dati permette di gestire numerosi tipi di file, DBMS e XML.

3. Schema xml del geodatabase

Il geodatabase XML è il principale meccanismo di scambio per la condivisione dei dati tra gli utenti di ArcGIS e delle altre applicazioni (ESRI, 2006b). Utilizzando lo schema XML e il formato ArcGIS XML è possibile importare ed esportare in un geodatabase tutti gli oggetti archiviabili. Con questo sistema ESRI pubblica e mantiene, in modo aperto, lo schema completo e i contenuti del geodatabase sotto forma di specifica XML e fornisce esempi di implementazioni per spiegare come gli utenti possono condividere l'aggiornamento dei dati tra sistemi eterogenei. Le altre applicazioni possono ricevere i flussi di dati XML come:

- Scambio e condivisione completa o parziale degli schemi dei geodatabase tra gli utenti di ArcGIS
- Scambio di dataset senza perdita di dati
- Scambio di elementi geografici semplici, simile alla scambio di shapefile
- Scambio dei soli record aggiornati (delta) usando i flussi XML per trasferire gli aggiornamenti e le modifiche tra i geodatabase e le altre applicazioni.

Nel seguente documento sarà descritto lo schema XML utilizzato per la conversione del database MySQL *Euro-Mediterranean Paleotsunami* in un geodatabase di ArcGIS. Il database, il cui schema è rappresentato nella Figura 14, è stato sviluppato all'interno del progetto TRANSFER finanziato dalla Commissione Europea e raccoglie i dati sulle inondazioni causate da tsunami nel passato. I dati contenuti nel database devono essere trasferiti all'interno del geodatabase in modo da poter essere analizzati con strumenti GIS per lo sviluppo di scenari e per il calcolo del rischio a essi collegato.

Ulteriori informazioni sul progetto TRANSFER sono reperibili sul sito www.transferproject.eu, mentre le informazioni relative al database Euro-Mediterranean Paleotsunami sono disponibile sul sito <http://paleotsunami.rm.ingv.it>.

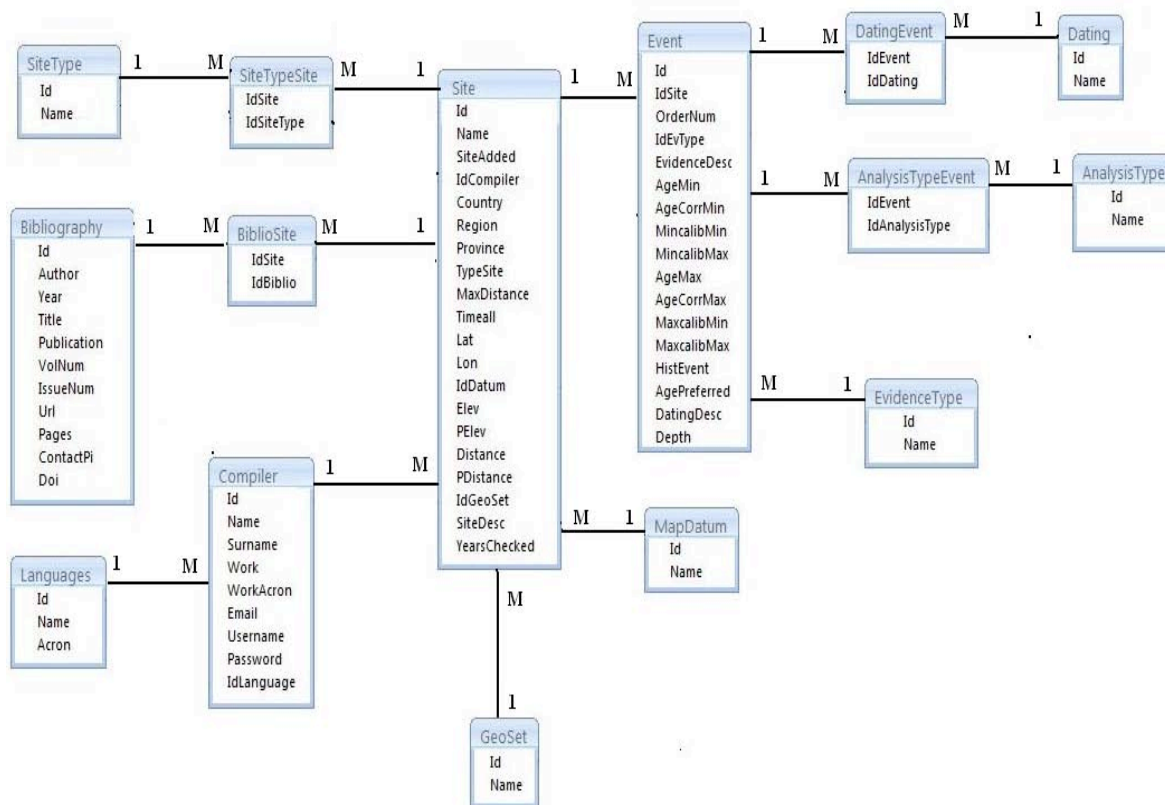


Figura 14 Schema del database *Euro-Mediterranean Paleotsunami* implementato in MySQL e accessibile via Internet all'URL <http://paleotsunami.rm.ingv.it>.

4. Tipi di documenti xml del geodatabase

In ArcGIS è possibile creare tre tipi di documenti XML: il documento *Workspace*, il documento *RecordSet* e il documento *DataChanges*.

In questo lavoro sarà trattato in dettaglio solo il documento *Workspace*, in quanto utilizzato per convertire il database Euro-Mediterranean Paleotsunami, implementato in MySQL, in un geodatabase di ArcGIS. La descrizione specifica di tale documento è stata ricavata dai rapporti tecnici di riferimento prodotti dalla ESRI (ESRI, 2004).

4.1 Documento Workspace

L'esportazione di un geodatabase in XML genera un documento *Workspace* che oltre a possedere tutti i contenuti e tutte le informazioni sullo schema del geodatabase può anche contenere i dati. XML *Workspace* contiene due elementi figli: *WorkspaceDefinition* e *WorkspaceData*.

Nello schema seguente (Fig.15) vengono rappresentati gli elementi che compongono il documento.

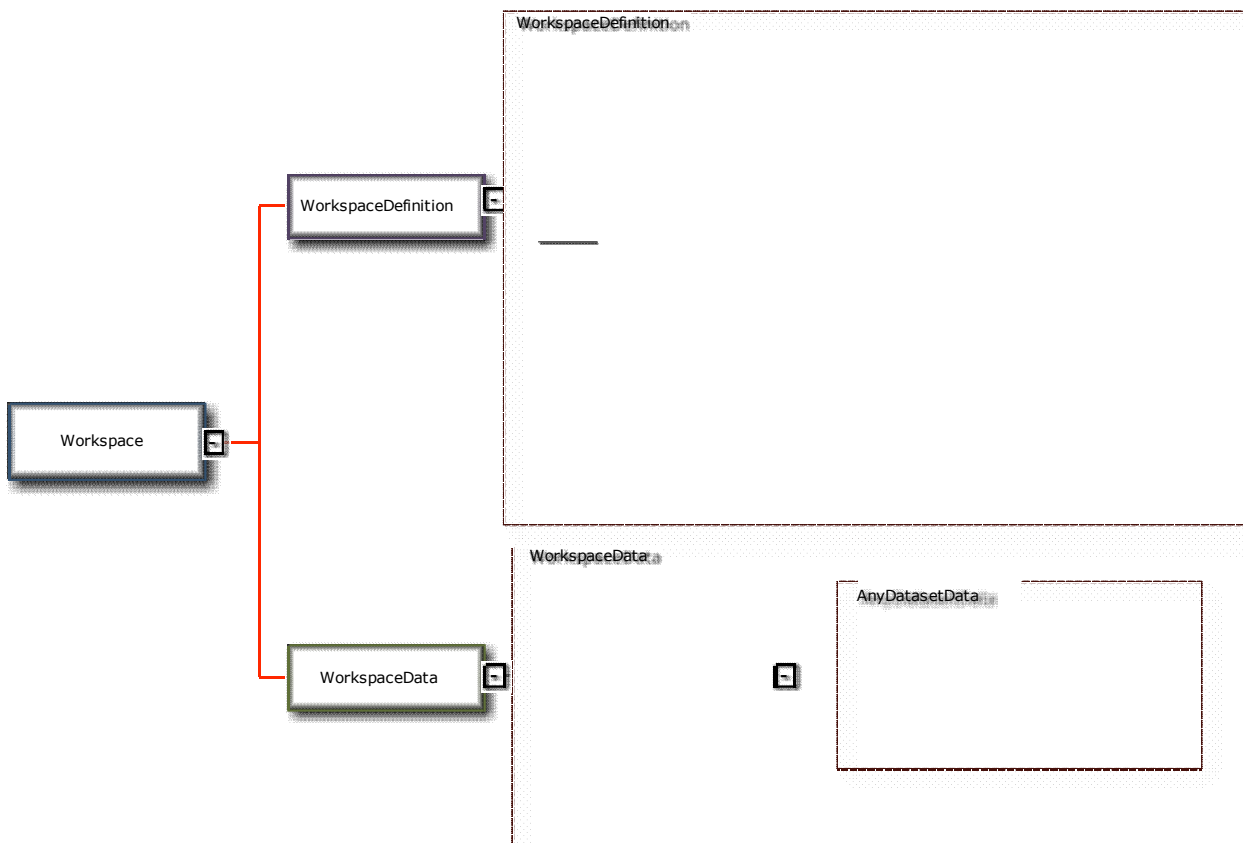


Figura 15 Lo schema del documento *Workspace* con gli elementi che lo compongono.

4.1.1 *WorkspaceDefinition* e *WorkspaceData*

WorkspaceDefinition esporta lo schema, i rapporti e le informazioni sul comportamento del geodatabase, comprendendo tutte le caratteristiche semplici e personalizzate dei dati, la partecipazione a reti e a topologie, la connettività di rete e le regole topologiche, relazioni semplici e composte e tutte le altre informazioni associate con il set di dati del geodatabase. Pertanto, tutti i comportamenti connessi con il geodatabase sono conservati e possono essere ricreati quando il documento XML viene importato.

In particolare il *WorkspaceDefinition* contiene il tipo di workspace, la versione, un array dei domini presenti e un array di *DataElement* (vedi sezione specifica *DataElement*).

WorkspaceData può avere da zero a molti elementi figlio *DatasetData*, del tipo *AnyDatasetData*. Dal tipo *AnyDatasetData* derivano due tipi XML: *TableData* e *RasterDatasetData*.

TableData è derivato da *AnyDatasetData* aggiungendo un elemento chiamato *Data* di tipo *RecordSet* (Fig. 16). A sua volta *RecordSet* contiene i record della tabella (vedi sezione specifica *RecordSet*). Nel caso di tabelle e feature class, il nome del *data element* nella sezione della definizione corrisponde al nome della tabella di dati nella sezione dei dati.

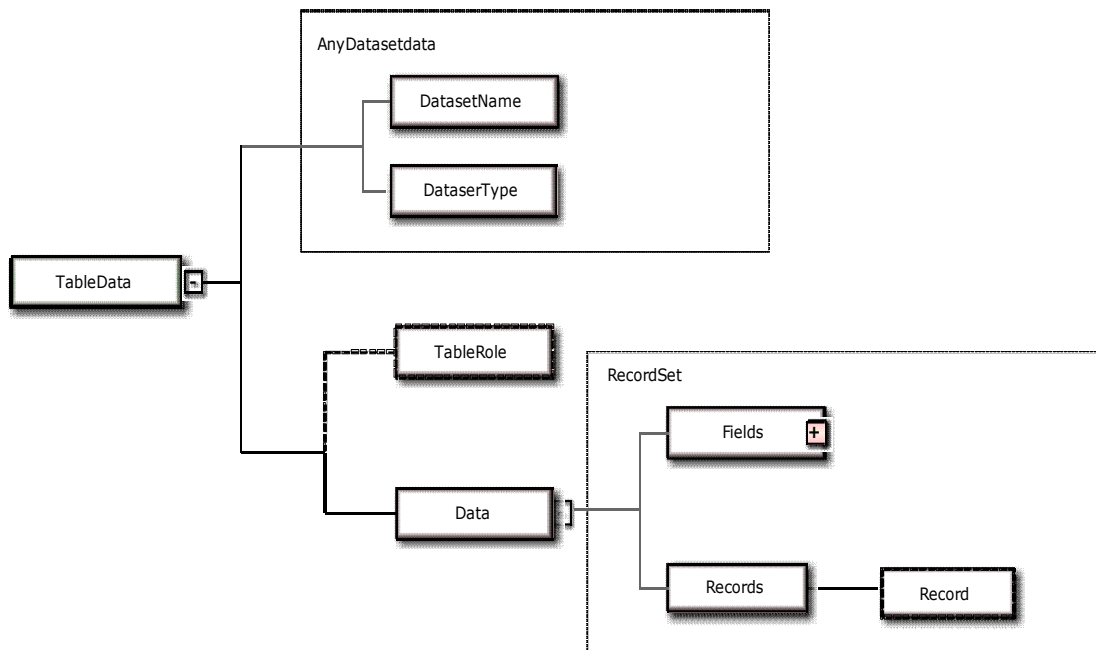


Figura 16 Espansione dell'elemento *TableData* di tipo *AnyDatasetData* del documento *WorkspaceData*.

RasterDatasetData non è stato utilizzato nel presente lavoro.

Utilizzando *WorkspaceDefinition* e *WorkspaceData*, un documento *Workspace* contiene sia lo schema del geodatabase sia i dati delle sue tabelle. Avere la definizione separata dei dati offre alcuni benefici, per esempio è possibile leggere solo la sezione del documento che contiene lo schema oppure leggere l'intero documento (schema e dati).

4.1.1.1 Data Elements: *DataElement* - *DEDataset* - *DETable*

I *DataElements* sono un insieme di classi *ArcObjects* che descrivono tutti gli aspetti di un dataset in un geodatabase. Per esempio, il data element contiene il nome, il tipo di elemento geometrico, i campi, gli indici e le altre caratteristiche di una feature class.

Ci sono data element per la maggior parte degli oggetti del geodatabase. Tutte le informazioni relative allo schema del geodatabase sono importate o esportate in XML utilizzando questi data element *ArcObject*.

Di seguito viene riportata una parte della gerarchia che corrisponde ai data element *ArcObjects*. La freccia indica il tipo genitore e i relativi tipi astratti (Fig. 17).

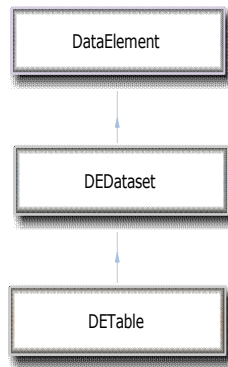


Figura 17 Espansione dei Data Elements (DataElement-DEDataset-DETable).

DataElement è un tipo astratto utilizzato come tipo base per tutti i data elements; pertanto tutti i data elements derivano da DataElement ereditandone gli elementi e aggiungendone altri. Per espandere il concetto di DataElement vengono presi in esame gli elementi che lo compongono (Fig. 18).

Il valore dell'elemento **CatalogPath** è una stringa che descrive il percorso del dataset. L'elemento **Name** contiene il nome del data element.

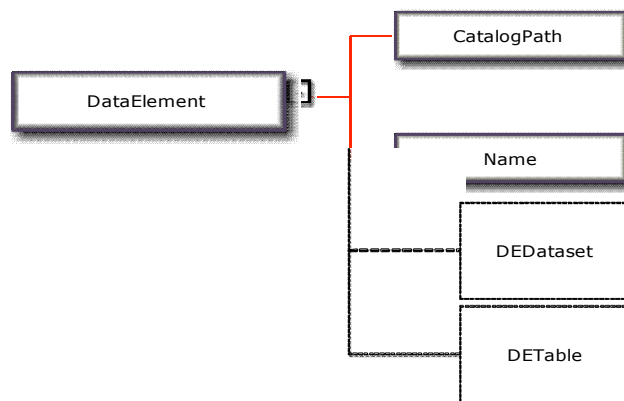


Figura 18 Espansione dell'elemento DataElement.

Un DEDataset viene utilizzato come tipo di base astratta per la rappresentazione delle caratteristiche comuni a tutti dataset. Il DEDataset estende il DataElement aggiungendo elementi che contengono l'ID del dataset, il suo tipo, se può avere versioni ed eventualmente la versione corrente (Fig. 19).

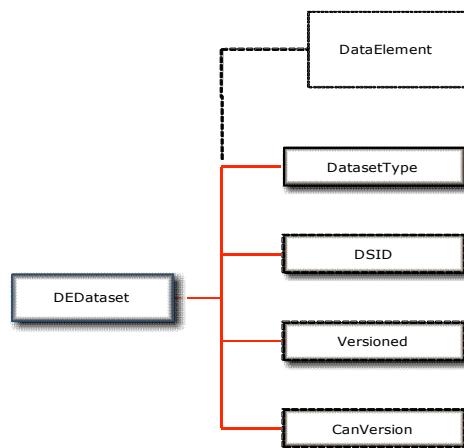


Figura 19 Espansione dell'elemento DEDataset.

DETable deriva dal DEDataset al quale aggiunge altri elementi al tipo di base, quali per esempio il campo OBJECTID field, un array di Fields, gli indici e i sottotipi (Fig. 20).

L'elemento **CLSID** contiene l'identificativo del tipo di classe di oggetti. Se la classe è del tipo feature class, gli elementi **EXTCLSID** e ClassExtensionProperties descrivono rispettivamente le informazioni sull'estensione del comportamento della classe e le sue proprietà.

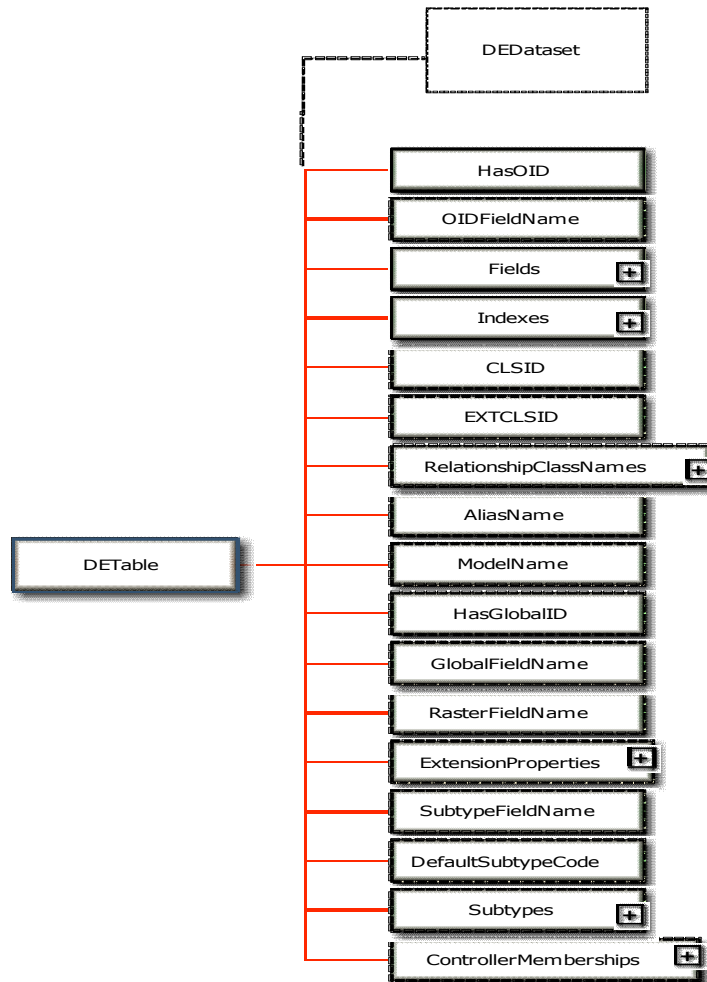


Figura 20 Espansione dell'elemento DETable.

Tabelle, feature class e altri dataset contengono collezioni di campi e di indici. Il tipo XML **Fields** contiene un elemento il cui nome è FieldArray. Come indica lo stesso nome, il tipo XML FieldArray è un array di campi che può essere composto da zero a un numero infinito di campi. Il tipo XML Field contiene le caratteristiche di un campo come per esempio il nome, il tipo, se permette valori nulli, l'alias e altre proprietà (Fig. 21).

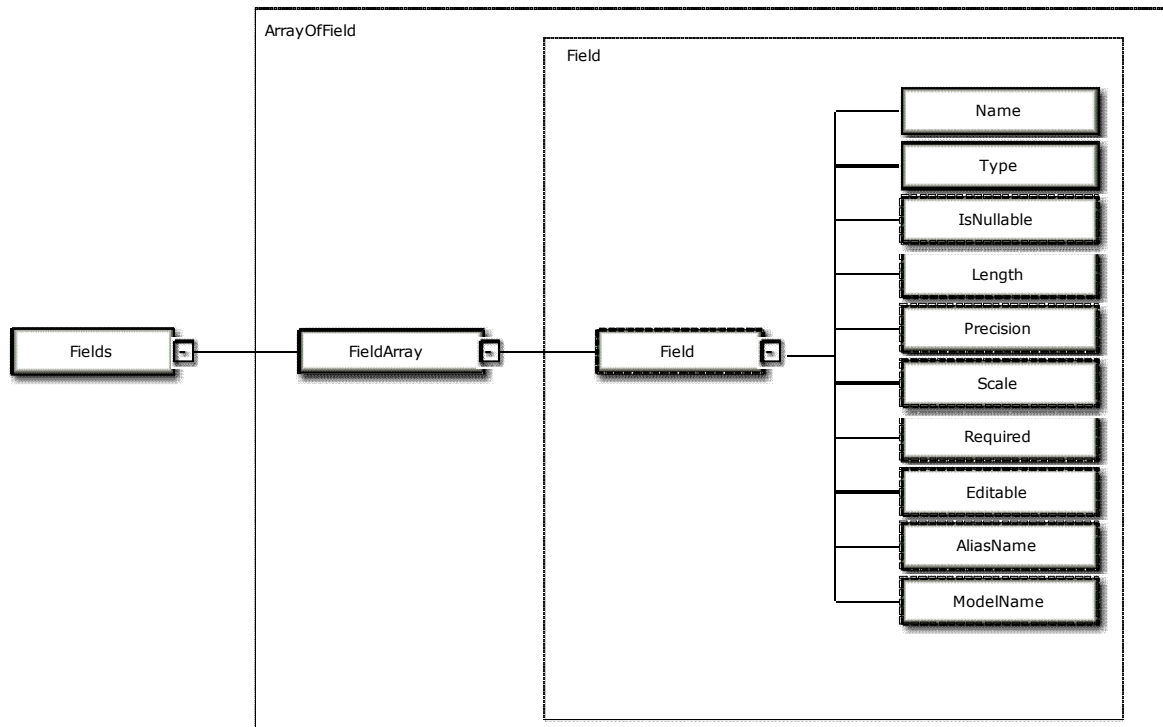


Figura 21 Espansione dell'elemento Fields.

Come per il caso di Fields, esistono i tipi XML **Indexes** e **IndexArray**. **Index** descrive l'indice del geodatabase. Contiene gli elementi per il nome dell'indice, se è univoco o ascendente e l'insieme dei campi di cui è composto l'indice (Fig. 22).

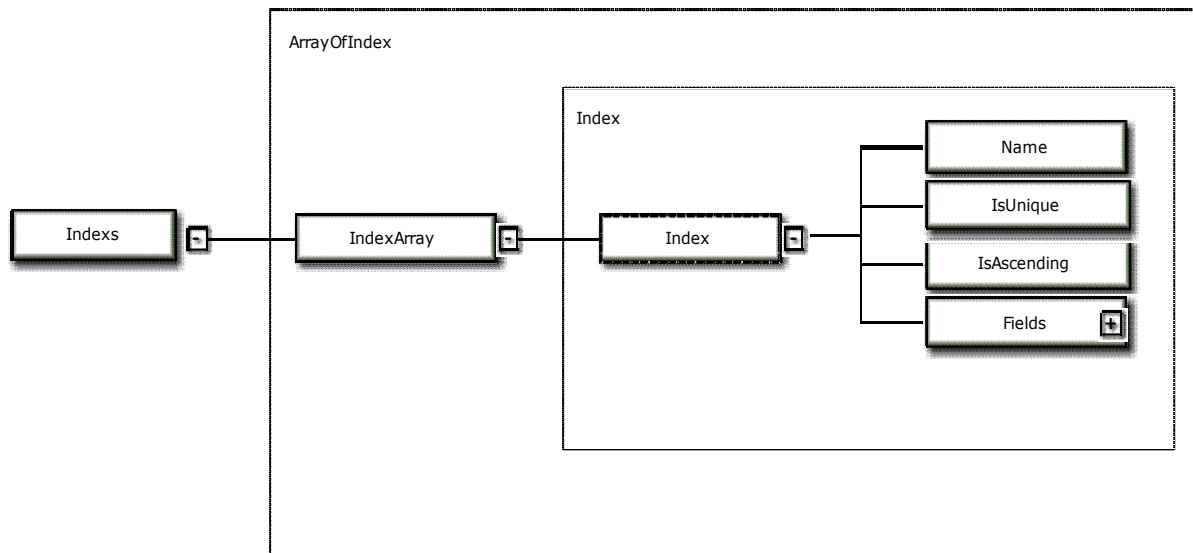


Figura 22 Espansione dell'elemento Fields.

4.1.1.2 RecordSet

Un **RecordSet** rappresenta i dati tabellari in un database. È formato da un insieme di campi e un insieme di righe. Un record è un insieme di valori, uno per ogni campo del **RecordSet**. L'insieme dei campi e dei valori di un record sono ordinati e devono corrispondere (il primo valore di ogni record deve

corrispondere al primo campo e così via). Il tipo XML **Fields** è analogo a quello descritto in precedenza per il documento WorkspaceDefinition (Fig. 23).

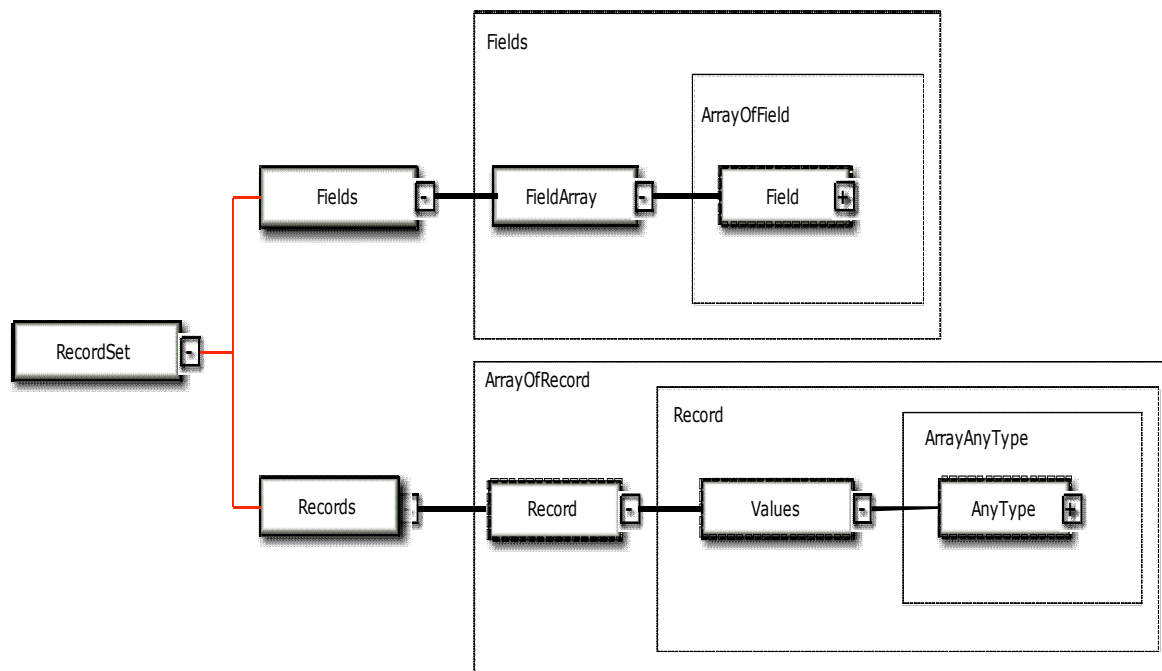


Figura 23 Espansione dell'elemento RecordSet.

4.2 Passaggio dal database MySQL allo Schema XML di un geodatabase

Un database relazionale è caratterizzato da entità e loro relazioni, dai rispettivi attributi, quali, nome, tipo di dato, possibilità di contenere valori nulli, lunghezza e dai vincoli di integrità referenziale. Il database MySQL a cui fa riferimento il presente lavoro è rappresentato nello schema ER (Entità - Relazioni) della Figura 14 ed è stato sviluppato nell'ambito del progetto *Euro-Mediterranean Paleotsunami*.

Il database è organizzato in 15 tabelle (entità) legate tra di loro attraverso relazioni con cardinalità 1:M. I concetti fondamentali sono contenuti nelle seguenti entità: *Siti*, *Eventi*, *Compilatori* e *Bibliografia*; mentre le altre entità sono di supporto alla struttura del database. Il *Sito* è l'entità principale e rappresenta un'area di studio. Per ciascun *Sito* possono essere presenti più *Eventi* a ognuno dei quali è associata la specifica *Bibliografia*. Ciascun *Compilatore* può inserire le informazioni relative a più *Siti*.

Per ulteriori informazioni particolareggiate sullo schema del database si rimanda alla specifica documentazione prodotta per il progetto.

Per ciascun elemento che compone il database verrà di seguito descritta la corrispondenza con gli elementi dello schema XML. Inoltre saranno inseriti dei frammenti di codice XML che ne mostrano la struttura e che permetteranno di creare un template per la generazione del documento.

4.2.1 Tabelle – Attributi

Per ciascuna tabella del database deve essere inserito all'interno dello schema XML un elemento DataElement, DEDataset e DETable elencando i campi con i parametri corretti, per ciascun attributo un elemento Field nell'array di Fields e per ciascun indice un elemento Index nell'array di Indexes. Di seguito vengono riportati dei frammenti di codice XML per la creazione di una tabella riportati nel documento WorkspaceDefinition e la descrizione dei campi più interessanti.

In particolare: il campo **CatalogPath** contiene la posizione della tabella all'interno del geodatabase mentre il campo **Name** contiene il nome della tabella.

Il campo **DSID** è il DatasetID e dovrà contenere un codice univoco per ciascuna tabella.

Il campo **HasOID** indica che esiste un campo di tipo contatore che in ArcGIS per default è chiamato **OBJECTID** ed è indicato nel campo **OIDFieldName**. Questo campo è indicizzato e contiene un codice univoco per ogni oggetto contenuto in una tabella.

```
<!-- DataElement -->
<DataElement xsi:type="esri:DETable">
  <CatalogPath>/OC=NomeTabella</CatalogPath>
  <Name>NomeTabella</Name>
<!-- DEDataset -->
  <DatasetType>esriDTable</DatasetType>
  <DSID>1</DSID>
  <Versioned>false</Versioned>
  <CanVersion>false</CanVersion>
<!-- DETable -->
  <HasOID>true</HasOID>
  <OIDFieldName>OBJECTID</OIDFieldName>

```

..segue

Il campo **Fields** contiene la struttura della tabella: un array di campi **Field**, uno per ciascun attributo, e il campo **OBJECTID** con le proprietà associate (nome, tipo di dato, lunghezza...).

```
..continua
<!-- Fields-->
<Fields xsi:type="esri:Fields">
  <FieldArray xsi:type="esri:ArrayOfField">
    <!-- OBJECTID -->
    <Field xsi:type="esri:Field">
      <Name>OBJECTID</Name>
      <Type>esriFieldTypeOID</Type>
      <IsNullable>false</IsNullable>
      <Length>4</Length>
      <Precision>0</Precision>
      <Scale>0</Scale>
      <Required>true</Required>
      <Editable>false</Editable>
      <AliasName>OBJECTID</AliasName>
      <ModelName>OBJECTID</ModelName>
    </Field>
  </FieldArray>
</Fields>

```

```

<!--Nome Attributo -->
<Field xsi:type="esri:Field">
  <Name>NomeAttributo</Name>
  <Type>esriFieldTypeString</Type>
  <IsNullable>true</IsNullable>
  <Length>100</Length>
  <Precision>0</Precision>
  <Scale>0</Scale>
</Field>
</FieldArray>
</Fields>

```

..segue

Il campo **OBJECTID** deve essere necessariamente indicizzato e viene indicato nella sezione Indexes, dove vengono indicati anche gli altri campi sui quali è stato creato un indice.

..continua

```

<!-- Indexes-->
<Indexes xsi:type="esri:Indexes">
  <IndexArray xsi:type="esri:ArrayOfIndex">
    <Index xsi:type="esri:Index">
      <Name>FDO_OBJECTID</Name>
      <IsUnique>true</IsUnique>
      <Fields xsi:type="esri:Fields">
        <FieldArray xsi:type="esri:ArrayOfField">
          <Field xsi:type="esri:Field">
            <Name>OBJECTID</Name>
            <Type>esriFieldTypeOID</Type>
            <IsNullable>false</IsNullable>
            <Length>4</Length>
            <Precision>0</Precision>
            <Scale>0</Scale>
            <Required>true</Required>
            <Editable>false</Editable>
            <AliasName>OBJECTID</AliasName>
            <ModelName>OBJECTID</ModelName>
          </Field>
        </FieldArray>
      </Fields>
    </Index>

```

```
</IndexArray>
</Indexes>
```

..segue

Nel campo **CLSID** deve essere indicato il valore relativo al tipo di classe di oggetti. Se la classe di oggetti è una feature class, nei campi **EXTCLID** e **ExtensionProperties** devono essere indicate rispettivamente l'estensione del comportamento e le proprietà della classe. Il frammento seguente è relativo a una semplice classe di oggetti.

..continua

```
<CLSID>{7A566981-C114-11D2-8A28-006097AFF44E}</CLSID>
<EXTCLSID/>
<RelationshipClassNames xsi:type="esri:Names"/>
<AliasName>AnalysisType</AliasName>
<ModelName/>
<HasGlobalID>false</HasGlobalID>
<GlobalIDFieldName/>
<RasterFieldName/>
<ExtensionProperties xsi:type="esri:PropertySet">
  <PropertyArray xsi:type="esri:ArrayOfPropertySetProperty"/>
</ExtensionProperties>
<ControllerMemberships xsi:type="esri:ArrayOfControllerMembership"/>
</DataElement>
```

L'elemento Fields (contenente l'array di Field) deve essere indicato sia nell'area WorkstationDefinition sia nell'area WorkstationData per consentire il disaccoppiamento dello schema dai dati, in modo che sia possibile importare solo lo schema o solo i dati.

4.2.2 Valori degli Attributi

Nell'area WorkstationData, dopo aver effettuato la definizione della struttura della tabella, sarà inserito per ciascun record della stessa un array di Record contenente un array di valori corrispondenti agli attributi della tabella.

```
<Records xsi:type="esri:ArrayOfRecord">
  <!-- array Of Record -->
  <Record xsi:type="esri:Record">
    <Values xsi:type="esri:ArrayOfValue">
      <!--array Of Value -->
      <Value xsi:type="xs:int"/>
      <Value xsi:type="xs:float"/>
      <Value xsi:type="xs:short"/>
      <Value xsi:type="xs:string"/>
    </Values>
  </Record>
</Records>
```


5. Procedura di importazione dei dati dal file xml al geodatabase

La procedura di importazione dei dati dal file XML al geodatabase viene eseguita attraverso l'utilizzo delle funzionalità presenti nell'interfaccia ArcCatalog di ArcGIS, porta di accesso alla gestione ed esplorazione dei dati geografici (Fig. 24).

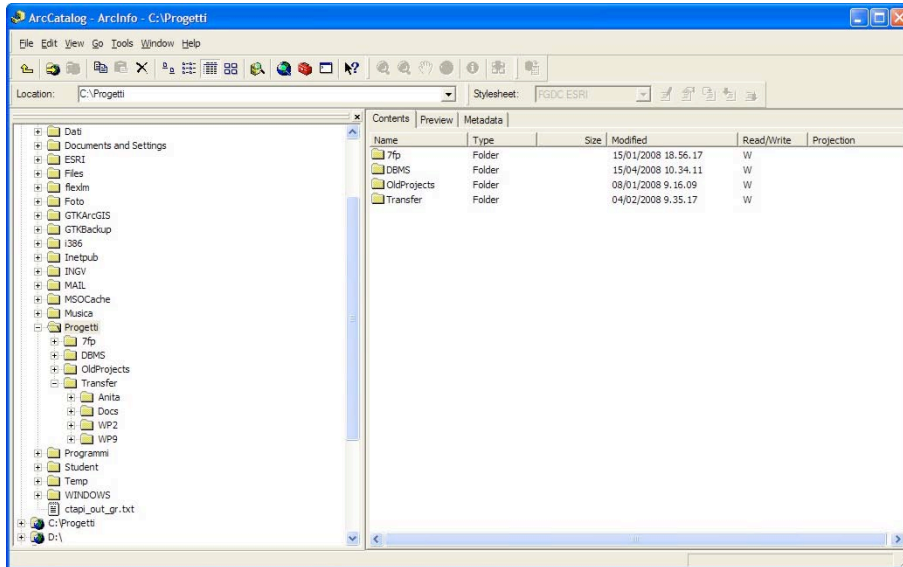


Figura 24 L'interfaccia ArcCatalog di ArcGIS.

La prima operazione da effettuare è la creazione di un nuovo geodatabase vuoto, che può essere di uno dei tre tipi gestiti da ArcGIS. Nel nostro caso, creeremo un *personal geodatabase* di Microsoft Access, ma la procedura di importazione si applica in modo analogo agli altri tipi di geodatabase (Fig. 25).

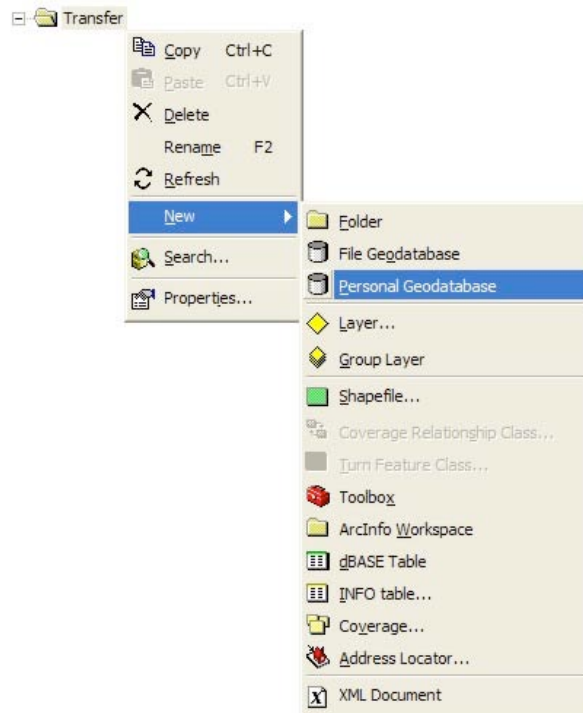


Figura 25 Creazione di un personal geodatabase da ArcCatalog.

Dopo avere creato un geodatabase vuoto, facendo clic con il pulsante destro del mouse sul geodatabase si accede al menu contestuale. Puntando a Import, selezionare *XML Workspace Document* per avviare il wizard per l'importazione dello schema e dei dati dal file XML (Fig. 26).

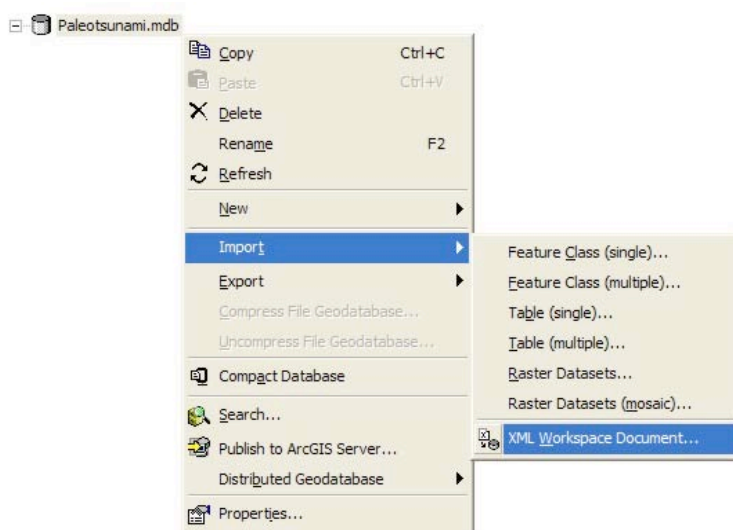


Figura 26 Importazione dei dati dal documento XML.

Il wizard di importazione dei dati permette di importare il solo schema del database (cioè la struttura delle tabelle) oppure sia lo schema sia i dati (Fig. 27).

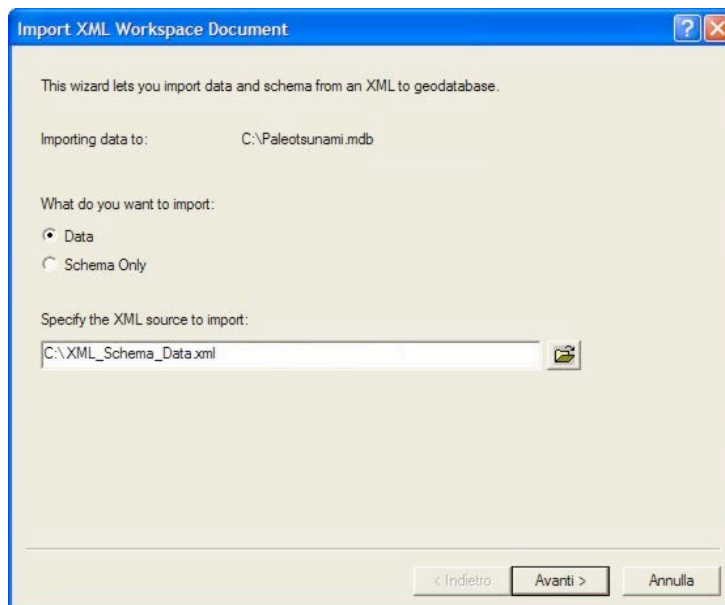


Figura 27 La prima schermata del wizard di importazione dei dati.

La schermata successiva del wizard visualizza lo schema del database XML e permette di modificare il nome delle tabelle di output nella colonna Target Name (Fig. 28).

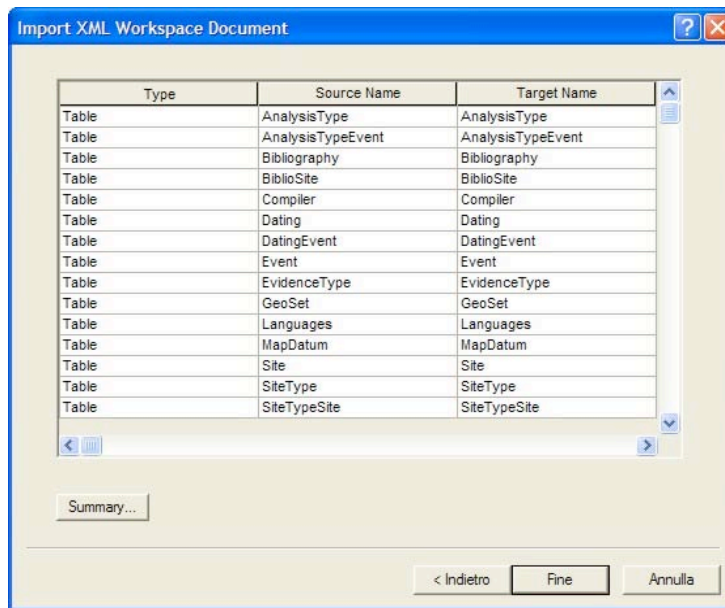


Figura 28 La schermata del wizard che permette la modifica del nome delle tabelle di output.

Il risultato dell'esecuzione del wizard è la creazione delle tabelle e l'importazione dei dati (Fig. 29).

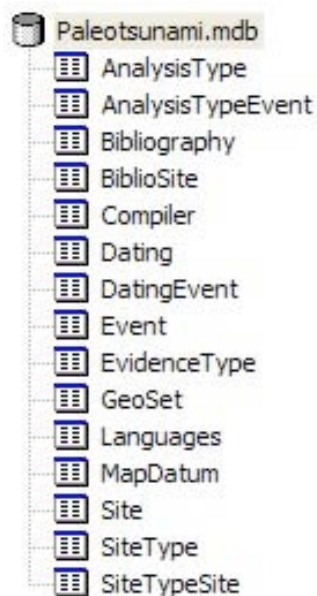


Figura 29 Il geodatabase dopo l'esecuzione del wizard per l'importazione dei dati.

Dopo avere importato i dati nel geodatabase, è possibile creare la feature class puntuale *fSite* utilizzando le coppie di coordinate latitudine/longitudine archiviate nella tabella *Site* (Fig. 30).

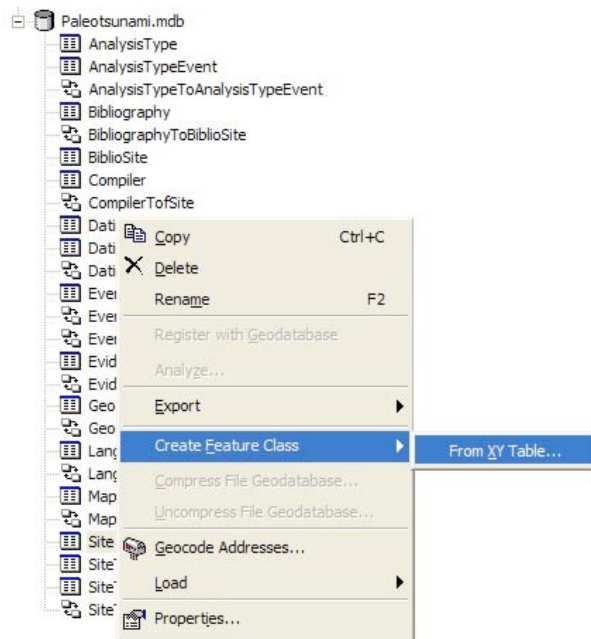


Figura 30 Creazione della feature class puntuale fSite a partire dai dati di latitudine e longitudine archiviati nella tabella Site.

Dopo avere creato la feature class contenente i punti che indicano la posizione dei paleotsunami è possibile creare le relazioni tra la feature class e le altre tabelle (Fig. 31), in base allo schema del geodatabase.

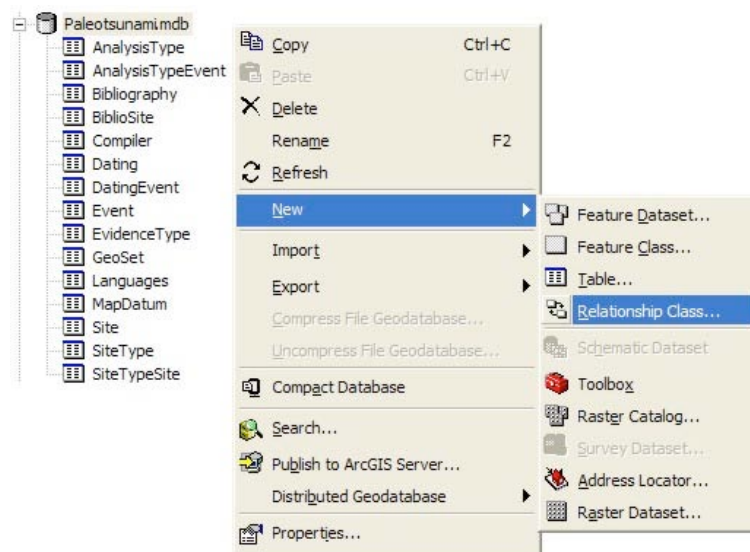


Figura 31 Creazione di una classe di relazione da ArcCatalog.

Lo schema del geodatabase *Euro-Mediterranean Paleotsunami* e le relazioni esistenti tra la feature class fSite, contenente i dati geometrici, la tabella originaria Site e le altre tabelle presenti nel geodatabase *Euro-Mediterranean Paleotsunami* sono mostrati nella Figura 32 e nella Figura 33.

Schema del geodatabase

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					10
Name	Strg	Yes					

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
IEvent	ShortInt	No					
IAntisType	ShortInt	No					

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					
Author	Strg	Yes					255
Year	LongInt	Yes					255
Title	Strg	Yes					255
Publizin	Strg	Yes					255
VidNum	ShortInt	Yes					
IssNum	ShortInt	Yes					
Url	Strg	Yes					255
Pages	Strg	Yes					255
CoverPI	Strg	Yes					255
Di	Strg	Yes					255

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ISb	ShortInt	No					
IBib	ShortInt	No					

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					
Name	Strg	Yes					30
Surname	Strg	Yes					30
Work	Strg	Yes					100
WorkAcron	Strg	Yes					10
Email	Strg	Yes					200
Username	Strg	Yes					20
Password	Strg	Yes					255
Language	ShortInt	No					

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					
Name	Strg	Yes					100

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
IEvent	ShortInt	No					
IDatg	ShortInt	No					

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					
ISb	ShortInt	No					
OrderNum	ShortInt	No					
IEVtype	ShortInt	No					
EvidenceDesc	Strg	Yes					255
AggMh	LongInt	Yes					
AggCorrMh	LongInt	Yes					
MedMh	LongInt	Yes					
MaxMh	LongInt	Yes					
AggMx	LongInt	Yes					
AggCorrMx	LongInt	Yes					
MaxMx	LongInt	Yes					
MaxMh	LongInt	Yes					
MaxMx	LongInt	Yes					
HEvent	Strg	Yes					100
AggPreferred	Strg	Yes					255
DatgDesc	Strg	Yes					255
Datg	Int	Yes			0	0	

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					
Name	Strg	Yes					100

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					
Name	Strg	Yes					100
SEvent	Date	Yes			0	0	8
IDCont	ShortInt	No					
Conty	Strg	Yes					50
Regn	Strg	Yes					50
Pronc	Strg	Yes					50
TypeSb	ShortInt	No					
MacDate	ShortInt	Yes					
Trial	ShortInt	No					
Lat	Real	Yes					0
Lon	Real	Yes					0
IDatm	ShortInt	No					
Elev	LongInt	Yes					
PElev	LongInt	Yes					
Distm	LongInt	Yes					
PDistm	LongInt	Yes					
IDGeoSet	ShortInt	No					
SIDesc	Strg	Yes					255
YearsChecked	Strg	Yes					255
Shape	Geomty	No					

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					
Name	Strg	Yes					100

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					
Name	Strg	Yes					20
Acron	Strg	Yes					2

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					
Name	Strg	Yes					100

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					
Name	Strg	Yes					100
SEvent	Date	Yes			0	0	8
IDCont	ShortInt	No					
Conty	Strg	Yes					50
Regn	Strg	Yes					50
Pronc	Strg	Yes					50
TypeSb	ShortInt	No					
MacDate	ShortInt	Yes					
Trial	ShortInt	No					
Lat	Real	Yes					0
Lon	Real	Yes					0
IDatm	ShortInt	No					
Elev	LongInt	Yes					
PElev	LongInt	Yes					
Distm	LongInt	Yes					
PDistm	LongInt	Yes					
IDGeoSet	ShortInt	No					
SIDesc	Strg	Yes					255
YearsChecked	Strg	Yes					255

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ID	ShortInt	No					
Name	Strg	Yes					100

Field name	Datatype	Allow nulls	Default value	Domain	Prec- ision	Scale	Length
OBJECTID	ObjID				0		
ISb	ShortInt	No					
ISbType	ShortInt	No					

Figura 32 Schema del geodatabase Euro-Mediterranean Paleotsunami. È mostrata la struttura delle tabelle presenti nel geodatabase e della feature class fSite.

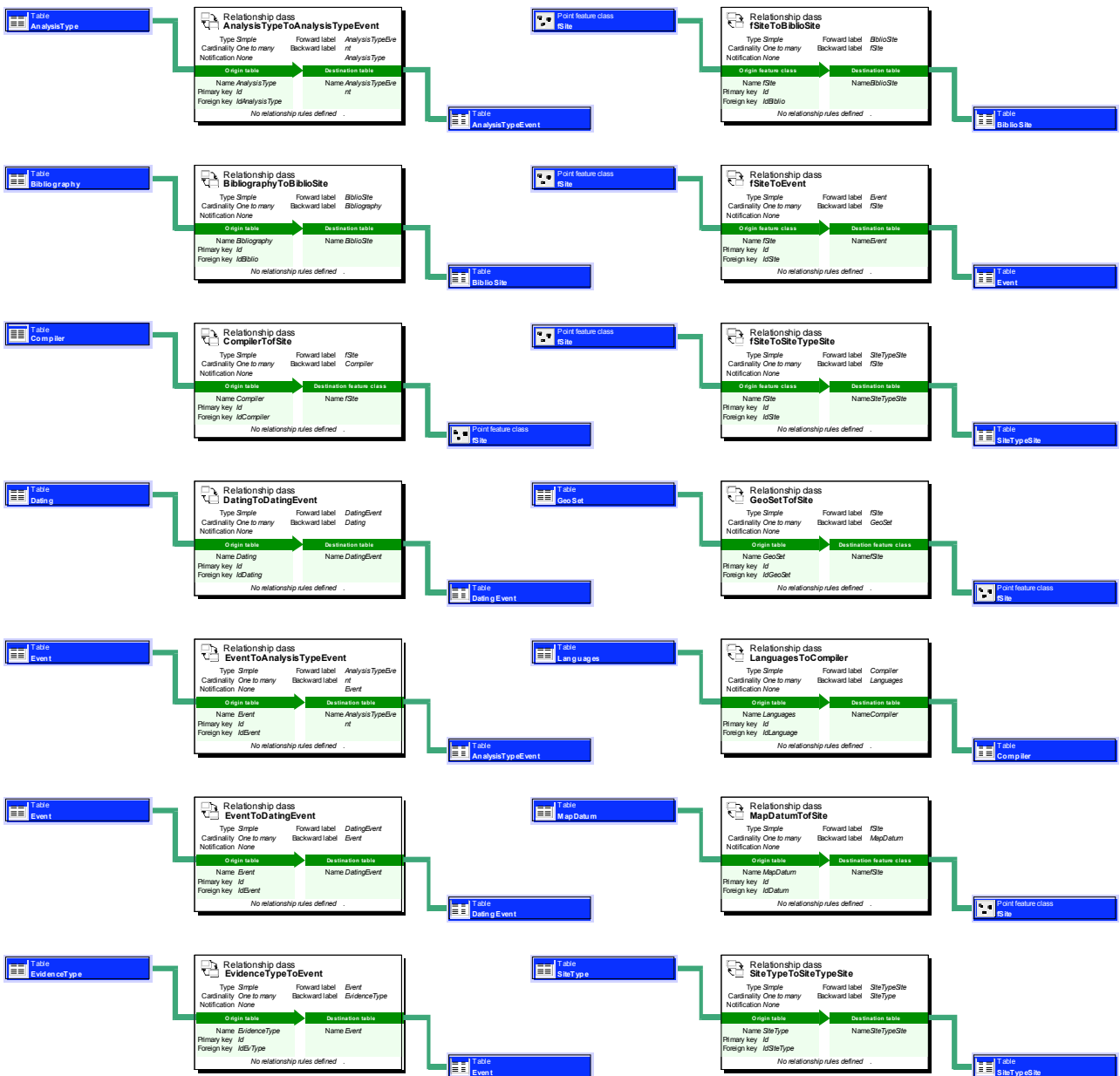


Figura 33 Schema del geodatabase. Sono mostrate le classi di relazione esistenti tra le varie tabelle.

Ringraziamenti

Si ringrazia il Sig. Diego Sorrentino per il prezioso supporto e per l'attiva collaborazione all'esportazione dei dati nel formato XML dal database MySQL *Euro-Mediterranean Paleotsunami*.

Si ringrazia inoltre ESRI Italia per la gentile concessione delle Figure 1, 2, 3, 4, 5, 6, 8, 9, 10 e 11 estratte dalla manualistica tecnica e dalle brochure informative.

Bibliografia

ESRI (2004). *XML Schema of the Geodatabase*. ESRI, ArcGIS SDK Technical Documents, http://edndoc.esri.com/arcobjects/9.1/ArcGISDevHelp/TechnicalDocuments/Geodatabase/XML_Schema/XMLSchema.htm

ESRI (2006a). *What is ArcGIS 9.2*. ESRI Press, Redlands, 126 pp.

ESRI (2006b). *Elements of the Geodatabase*. ESRI, <http://www.esri.com/software/arcgis/geodatabase/about/elements.html>

Coordinamento editoriale e impaginazione

Centro Editoriale Nazionale | INGV

Progetto grafico e redazionale

Laboratorio Grafica e Immagini | INGV Roma

© 2008 INGV Istituto Nazionale di Geofisica e Vulcanologia

Via di Vigna Murata, 605

00143 Roma

Tel. +39 06518601 Fax +39 065041181

<http://www.ingv.it>



Istituto Nazionale di Geofisica e Vulcanologia