

I principali formati digitali di codifica dell'informazione geografica

fabio.lucchesi_02.2005

I formati digitali di registrazione dell'informazione geografica

- L'informazione geografica digitale può essere archiviata secondo diversi formati di codifica digitale
- La natura tecnica dei diversi formati dipende
 - Dal tipo di modello spaziale attraverso il quale l'informazione è strutturata (raster vs vettoriale)
 - Dalle scelte operate dalle società produttrici di software (soprattutto nel caso di formati "proprietary")
- Verranno esaminati di seguito
 - formati vettoriali ESRI (ESRI coverage, ESRI shapefile, ESRI personal geodatabase)
 - altri formati vettoriali GIS (MapInfo Tab)
 - formati vettoriali ASCII (formato RT)
 - formati vettoriali CAD (Autodesk dwg, Autodesk dxf)
 - formati raster GIS (ESRI GRID, ASCII grid)
 - formati raster immagine (tiff, jpg)

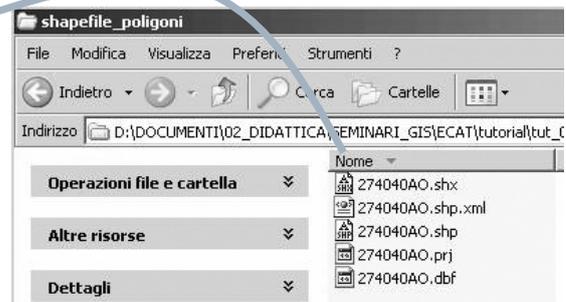
Il formato ESRI shapefile

- Una definizione di shapefile:**
formato di archiviazione di dati **vettoriali** capace di registrare localizzazione, forma e attributi di entità spaziali
- Uno shapefile è composto da un set di files relazionati e contiene una sola classe di oggetti (punti, linee o poligoni)**
 - formato creato con la prima versione di **ESRI ArcView**
 - è **completamente modificabile** in ArcGis, sia per quanto riguarda l'informazione geometrica, sia per quanto riguarda l'informazione tabellare
 - **non usa una struttura arco-nodo per registrare informazione topologica, dunque**
 - è **visualizzato velocemente** in ArcGis
 - linee "connesse" **non** devono condividere necessariamente un nodo comune
 - poligoni adiacenti **non** devono condividere necessariamente uno (o più) archi comuni

fabio.lucchesi_02.2005

Come appare uno shapefile in Windows Explorer

- In Windows Explorer appare **tutto il set di file che compongono il formato shapefile**
- Tutti i file devono avere lo stesso nome (salvo, ovviamente, l'estensione)**
- Un set di file deve comprendere almeno tre files**
 - *nomefile* .shp : è il file che contiene le **informazioni geometriche** (database di coordinate)
 - *nomefile* .dbf : è il file che contiene l'**informazione tabellare** (dati attributo)
 - *nomefile* .shx : è il **file indice** che permette di ricostruire il raccordo tra geometria e informazione tabellare
- **Il dataset può inoltre contenere files accessori, destinati a registrare**
 - indici spaziali** (.sbn, .shx, .fbn, fbx)
 - indici di attributo** (.ain, .aih)
 - sistemi di riferimento** (.prj)
 - metadati** (.xml)



fabio.lucchesi_02.2005

Lavorare con gli shapefile

- Quando si lavora con shapefile, e si vuole spostare un dataset da una directory a un'altra, p.e. scrivendo i dati su un CD, occorre ricordarsi di **copiare tutti i file** che compongono il dataset

- Se è necessario rinominare uno shapefile utilizzando Windows Explorer, occorre ricordarsi di **rinominare tutti i file** che compongono il dataset

fabio.lucchesi_02.2005

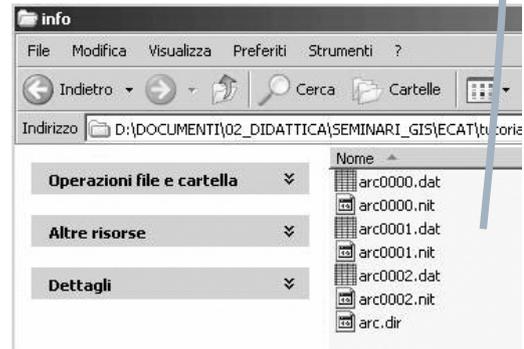
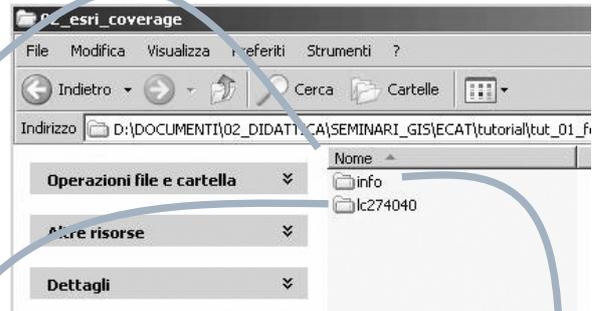
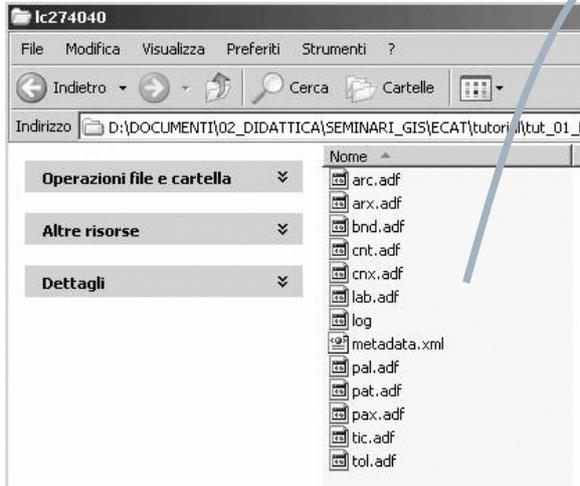
Il formato ESRI coverage per informazione vettoriale

- Una definizione del formato **coverage**:
formato di archiviazione di dati **vettoriali** capace di registrare localizzazione, forma, **relazioni topologiche** e attributi di entità spaziali
- Una coverage è composta da un set di files relazionati e contiene una o più classi di oggetti (il formato coverage è un formato **polimorfico**: nella stessa coverage possono convivere punti, linee, poligoni, annotazioni ...)
 - Si tratta di un formato legato alla diffusione di **ESRI ArcInfo**
 - Usa una articolazione complessa basata sulla **struttura arco-nodo**
 - supporta l'informazione topologica
 - la visualizzazione in ArcGis è più lenta
 - I dati geometrici non sono modificabili in ArcGis/ArcView

fabio.lucchesi_02.2005

Come appare una coverage in Windows Explorer

- in Windows Explorer la struttura del set di file è molto complessa e prevede una **articolazione in cartelle**
- quando si lavora con formati coverage è opportuno archiviare i files in uno spazio di lavoro dedicato (**workspace**)
 - Si consideri che ciascun workspace contiene una subdirectory **INFO** che archivia i dati attributo per tutte le coverage contenute



fabio.lucchesi_02.2005

Caratteristiche del formato ESRI coverage

- **Una ESRI coverage è un dataset di file organizzati in due directory**
 - **Una directory con lo stesso nome della coverage** contenente un numero variabile di file con diverse estensioni
 - **arc.adf**: contiene la definizione degli archi e dei loro vertici (ha un indice nominato arx.adf)
 - **pal.adf**: contiene la definizione dei poligoni (ha un indice nominato pax.adf)
 - **lab.adf**: contiene informazione sulle etichette dei record di punti
 - **cnt.adf**: contiene informazione sui centroidi (ha un indice nominato cnx.adf)
 - **prj.adf**: contiene informazione sul sistema di coordinate in uso nella copertura
 - ...
 - **La directory INFO**, che è condivisa con le altre GRID e coverage che sono collocate allo stesso livello nella struttura delle directory, e che contiene ugualmente diversi file
 - **arc.dir**: contiene l'elenco dei record
 - **arc.dat**: contiene i dati tabulari
 - **arc.nit**: contiene le definizioni dei campi
 - ...

fabio.lucchesi_02.2005

Lavorare con le coverage

- Quando si lavora con coverage è opportuno organizzare il proprio lavoro sistemando **una directory workspace** per contenere tutti gli elementi che compongono il dataset

- Non è opportuno utilizzare Windows Explorer** per operazioni di spostamento o per rinominare i file; ArcGis offre a questo scopo le funzionalità di **ArcCatalog**

fabio.lucchesi_02.2005

Personal geodatabase

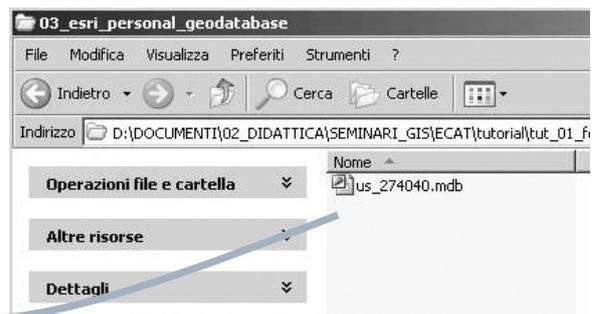
- Una definizione del formato geodatabase:**
database relazionale “object-oriented” che fornisce servizi di gestione di dati geografici
- Questi servizi includono regole di validazione, relazioni e associazioni topologiche
- Un geodatabase contiene un dataset polimorfico di entità ed è ospitato in un sistema di gestione di database relazionali**
 - è un formato molto efficiente e si sta imponendo come **formato preferenziale** di registrazione di informazione GIS
 - viene **visualizzato molto velocemente**
 - è **completamente modificabile** in ArcGis

- Si consideri che ArcGis/ArcInfo supporta **anche** il formato **ArcSDE geodatabase**, che è progettato per database relazionali di grandi dimensioni e può essere gestito attraverso diversi RDBMS (Oracle, SQL Server, ...)

fabio.lucchesi_02.2005

Come appare un geodatabase in Windows Explorer

- Il formato **ESRI Personal geodatabase** è un file registrato nel formato Microsoft Access 2000 e ha dunque estensione **.mdb**



fabio.lucchesi_02.2005

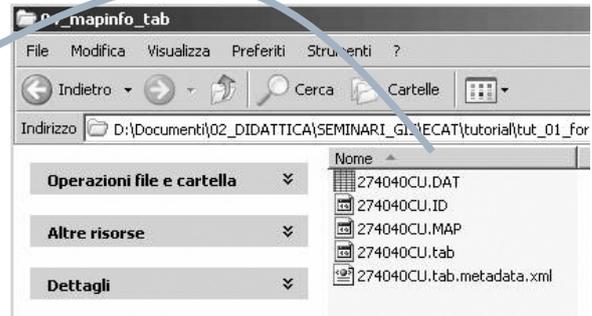
Il formato MapInfo TAB

- Il formato **TAB** è un formato proprietario MapInfo utilizzato per la codifica di informazione spaziale vettoriale e di informazione attributo associata
- È un formato polimorfico e consente a punti, linee e poligoni di convivere in una stessa copertura

fabio.lucchesi_02.2005

Come appare un file TAB MapInfo in Windows Explorer

- In Windows Explorer appare **tutto il set di file che compongono il formato TAB**
- Tutti i file devono avere lo stesso nome (salvo, ovviamente, l'estensione)**
- Un set deve comprendere almeno quattro file**
 - *nomefile.tab* : è il file che descrive la **struttura della tabella**
 - *nomefile.map* : è il file che contiene l'**informazione spaziale**
 - *nomefile.dat* : è il file che contiene l'**informazione attributo**
 - *nomefile.id* : è il **file indice** che permette l'associazione tra informazione spaziale e informazione non spaziale



- Un dataset TAB può inoltre contenere file ausiliari**
 - **.ind** : registra indici di attributo e velocizza le operazioni di ricerca
 - **.tab.metadata.xml** : registra metadati

fabio.lucchesi_02.2005

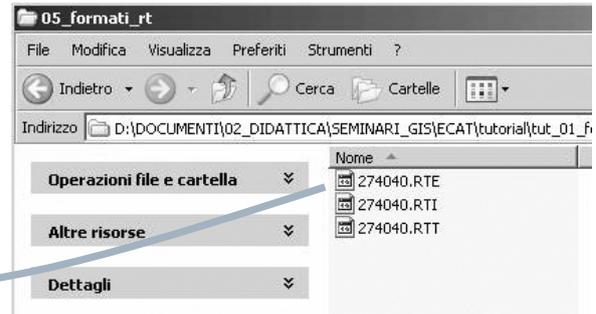
Il formato RT

- Il formato RT è un formato ASCII definito dalla Regione Toscana per una prima codifica della Cartografia Tecnica Regionale numerica**
- Per ciascun foglio di CTR vengono definiti tre file:**
 - **File delle informazioni (.RTI)**
 - Descrive le informazioni generali relative al foglio: metadati e coordinate geometriche dei limiti dell'area cartografata
 - **File delle entità (.RTE)**
 - Descrive la geometria delle entità lineari e areali associandole a un codice descrittivo
 - **File della toponomastica e simboli (.RTT)**
 - Descrive la geometria delle entità puntuali, la toponomastica e i punti di applicazione delle stringhe di testo

fabio.lucchesi_02.2005

Come appare il formato RT in Windows Explorer

- Ogni elemento di un dataset RT appare in Windows Explorer come un **singolo file**



fabio.lucchesi_02.2005

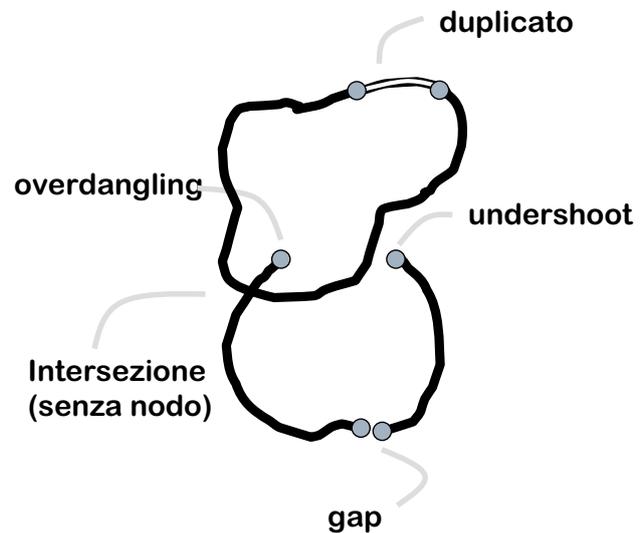
Formati CAD

- Una definizione generica di un formato CAD
formato di archiviazione di dati **vettoriali** capace di registrare localizzazione e forma degli oggetti rappresentati; l'informazione geometrica può essere **organizzata in strati distinti (layers)**
 - Si consideri che per essere utilizzato utilmente in un sistema GIS, il **tematismo CAD deve contenere informazione di localizzazione georeferenziata** rispetto a un sistema di riferimento
- Un **tematismo CAD** è costituito da un solo file, che all'interno del software GIS può essere **decostruito** separando l'informazione relativa alle **geometrie puntuali**, alle **geometrie lineari**, alle **geometrie poligonali**, alle **annotazioni testuali**
- Formati CAD più diffusi**
 - ***.dwg**, formato Autodesk
 - ***.dxf**, formato Autodesk di interscambio
 - ***.dgn**, formato Bentley/Integraph

fabio.lucchesi_02.2005

Caratteristiche dei formati CAD

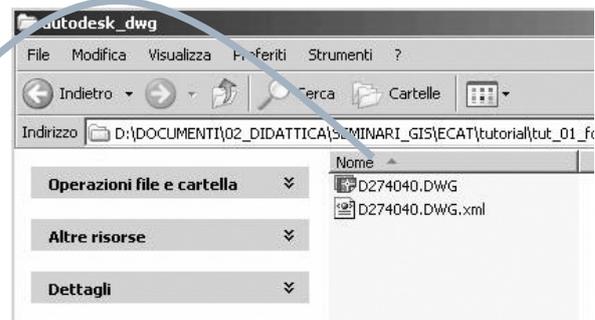
- I formati CAD sono **molto comuni** negli ambiti della progettazione architettonica e ingegnerile
- **Non sono connessi a informazione tabellare** (salvo la considerazione delle proprietà grafiche degli elementi (colore, spessore delle linee, ...) e della articolazione interna in layer)
- **Non supportano informazione topologica** e hanno quindi protocolli di costruzione meno rigidi; ciò può comportare **interruzioni di continuità** della descrizione geometrica degli elementi
 - duplicazioni
 - undershoots
 - overdanglings
 - duplicazioni
 - Gaps
 - ...



fabio.lucchesi_02.2005

Come appare un formato CAD in Windows Explorer

- Ogni tematismo CAD appare in Windows Explorer come un **singolo file**, al quale può essere associato un file ausiliario che registra i **metadati**



fabio.lucchesi_02.2005

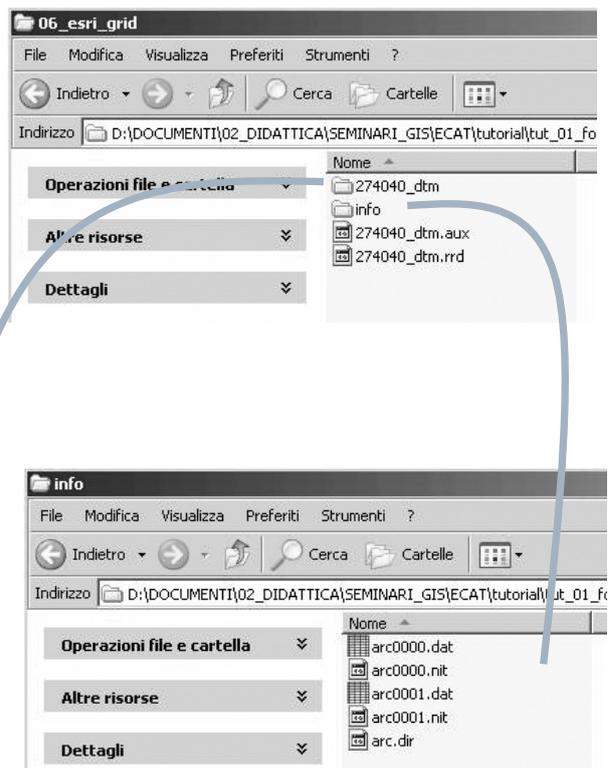
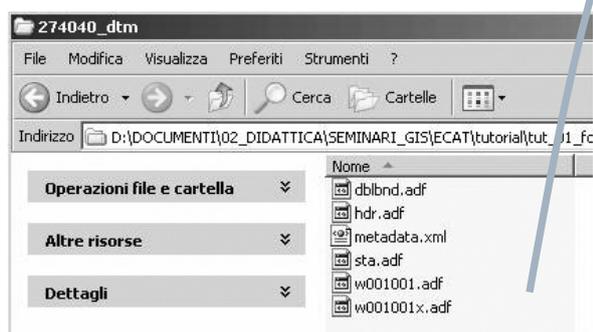
ESRI GRID

- Una definizione di una **ESRI GRID**:
formato di archiviazione di un **dataset raster** basato su una griglia di celle quadrate organizzate in righe e colonne, in cui ciascuna cella è riferita a una collocazione geografica attraverso coordinate x, y
- Una grid è **costituita da un set di files relazionati** che contengono l'informazione descrittiva di ciascuna cella

fabio.lucchesi_02.2005

Come appare una grid in Windows Explorer

- In Windows Explorer la struttura del dataset è molto complessa e prevede una **articolazione in cartelle**
- quando si lavora con formati grid è opportuno archiviare i files in uno spazio di lavoro dedicato (**workspace**)



fabio.lucchesi_02.2005

Caratteristiche del formato ESRI GRID

- Una ESRI GRID è un dataset di file organizzati in due directory**
 - **Una directory con lo stesso nome della GRID** contenente un numero variabile di file con estensione *.adf (Arc Data File)
 - dblbnd.adf**: registra il confine rettangolare (l'estensione spaziale) della GRID
 - hdr.adf**: registra il tipo di griglia (integer or floating point), così come il passo spaziale della griglia
 - prj.adf**: registra le informazioni relative al sistema di coordinate in uso
 - sta.adf**: registra informazioni statistiche sulla griglia (valori minimi e massimi, valori medi, deviazione standard)
 - vat.adf**: registra i valori delle celle di una integer grid e il conteggio (count) del numero di celle contenente ogni singolo valore
 - w001001.adf**: registra i valori delle celle di una floating point grid
 - w001001x.adf**: è un indice dei valori contenuti w001001.adf
 - log**: il file log archivia la sequenza dei comandi di elaborazione applicati alla griglia
 - **La directory INFO**, che è condivisa con le altre GRID e coverage che sono collocate allo stesso livello nella struttura delle directory

fabio.lucchesi_02.2005

Pyramids

- La risoluzione di una griglia raster tipicamente è superiore alla risoluzione del monitor in uso: ogni pixel del monitor contiene più di una cella della griglia**
- Per velocizzare la visualizzazione della griglia raster può essere conveniente creare delle copie a bassa risoluzione della griglia chiamate pyramids**
- I file che costituiscono le pyramids hanno lo stesso nome della griglia ed estensione .rrd (reduced resolution dataset)**
- I file pyramids sono in genere associati a file con estensione .aux (auxiliary file) in cui sono contenute anche informazioni statistiche relative al dataset raster**

fabio.lucchesi_02.2005

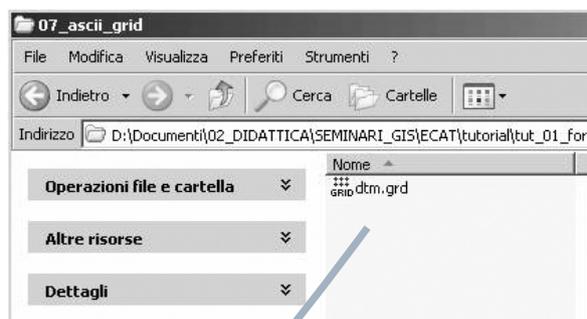
Lavorare con le ESRI grid

- Quando si lavora con ESRI grid è opportuno organizzare il proprio lavoro sistemando **una directory workspace** per contenere tutti gli elementi che compongono il dataset
- **Non è opportuno utilizzare Windows Explorer** per operazioni di spostamento o per rinominare i file; ArcGis offre a questo scopo le funzionalità di **ArcCatalog**

fabio.lucchesi_02.2005

Il formato ASCII grid

- Il formato **ASCII grid** codifica un modello spaziale raster in un unico file che descrive l'estensione della griglia, i parametri di georeferenziazione in un sistema di riferimento e i valori di ogni singola cella
- In windows Explorer appare come un **singolo file**



fabio.lucchesi_02.2005

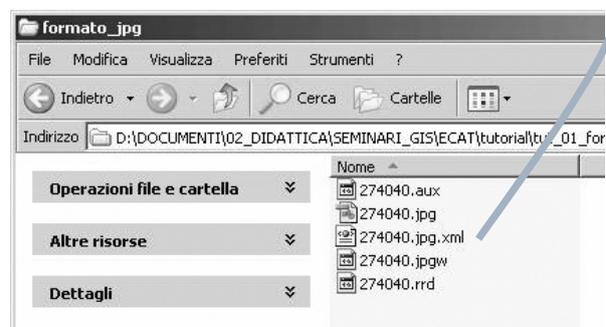
Immagini

- Una definizione generica dei **formati immagine** in un sistema GIS:
una immagine è una rappresentazione della superficie della Terra, tipicamente prodotta da una apparecchiatura ottica o elettronica, e **organizzata attraverso una griglia raster georeferenziata**, ossia di cui è registrata la posizione di ogni pixel rispetto a un sistema di riferimento
- Una immagine è archiviata come un **dataset raster** che registra valori che rappresentano l'intensità della luce riflessa, del calore, o un qualche **intervallo di valori nello spettro elettromagnetico**
- Esempi di immagini:
 - Rilevazioni satellitari
 - Fotografie
 - Scansioni di cartografie

fabio.lucchesi_02.2005

Come appare una immagine in windows explorer

- Il dataset di una immagine georeferenziata contiene sempre almeno il **file immagine** potenzialmente registrato in **formati diversi**
- Alcuni formati (**ECW, Geotiff, ...**) incorporano nella intestazione le informazioni necessarie per la georeferenziazione, altri formati (**Tiff, Jpeg, MRsid, ...**) hanno bisogno di essere associati a un **file testuale che riporta le informazioni essenziali per la georeferenziazione della griglia raster** (di norma: coordinate del primo pixel in alto a sinistra, rotazione della griglia, passo dimensionale della griglia)
- Il file di georeferenziazione **deve avere lo stesso nome del file immagine**, e una estensione corrispondente al formato del file immagine
 - **.tfw** (per immagine.tif)
 - **.jpgw** (per immagine.jpg)
 - **.sdw** (per immagine.sid)
- Il dataset può inoltre contenere **files accessori, destinati a registrare**
 - **informazioni che velocizzano la visualizzazione** (pyramids: **.aux, .rrd**)
 - **metadati** (**.xml**)



fabio.lucchesi_02.2005

Lavorare con immagini georeferenziate

- I formati immagine più diffusi (p.e. tiff e jpg) non registrano l'informazione di georeferenziazione nella intestazione del file; debbono perciò essere **accompagnati da un file ausiliario** che riporta le informazioni necessarie per la **georeferenziazione**
- Altri file ausiliari possono essere utili per registrare **metadati o copie a bassa risoluzione delle immagini (pyramids)**
- Quando si lavora con immagini georeferenziate occorre conservare i due (o più) file che compongono il dataset **nella stessa directory**
- Per operazioni di spostamento, o per rinominare i file, è opportuno utilizzare **l'ambiente ArcCatalog**

fabio.lucchesi_02.2005

Esercizio

- Esplorare attraverso Windows Explorer il contenuto della cartella tut_01_formati; la cartella contiene:**
 - tre shapefile (uno puntuale, uno lineare, uno poligonale)
 - una coverage
 - un personal geodatabase
 - una copertura MapInfo TAB
 - un set di file RT
 - una ESRI Grid
 - una ASCII grid
 - due temi CAD (rispettivamente nei formati .dwg e .dxf)
 - due immagini georeferenziate (rispettivamente in formato .jpg e .tif)
- Senza salvare eventuali modifiche,**
 - Aprire i file del dataset shapefile con estensione .dbf in **Microsoft Excel**
 - Aprire il geodatabase in **Microsoft Access**
 - Aprire ciascun file del dataset RT nel **Blocco Note** di Windows
 - Aprire il file **ASCII grid** nel **Blocco Note** di Windows
 - Aprire i file CAD in **Autodesk Autocad**
 - Aprire i file immagine .tif e .jpg in **Adobe Photoshop**
 - Aprire i file di georeferenziazione .tfw e .jpgw nel **Blocco Note** di Windows

fabio.lucchesi_02.2005