

Appunti del corso di Progettazione Sistemi Informativi Territoriali

(versione preliminare per uso interno)

*Giovanni Conforti
Rodolfo Fabiani*

Si ringraziano per il prezioso contributo:

Giorgio Ghelli

Maria Vittoria Masserotti

Premessa

Le peculiarità dei Sistemi Informativi Territoriali (costo alto dell'acquisizione dei dati, richiesta di particolari strumenti hardware e software, progettazione concettuale dipendente dal modello geografico, ecc.) richiedono uno studio approfondito per estendere le tecniche standard di analisi, pianificazione e progettazione studiate nell'ambito dell'*Ingegneria del Software*.

Questo corso affronta il problema della progettazione applicato ai SIT (detti anche GIS), in particolare si descrivono procedure e metodologie per:

- 1) Pianificazione di SIT;
- 2) Valutazione delle potenziali sorgenti di dati;
- 3) Test, pianificazione dell'acquisizione e valutazione di hardware e software;
- 4) Progettazione (concettuale e fisica) del SIT Database;
- 5) Costruzione del SIT Database;
- 6) Sviluppo di applicazioni SIT;
- 7) Pianificazione a lungo termine per il mantenimento (e aggiornamento) del SIT.

Tali procedure sono applicabili (con i dovuti adattamenti) sia per una prima creazione di un Sistema Informativo Territoriale, che per la revisione e lo sviluppo ulteriore di un SIT esistente od anche per la ristrutturazione di un progetto SIT in corso di sviluppo.

Bisogna precisare che la maggior parte della documentazione relativa a questo corso è stata prodotta basandosi su [GdG], una guida alla progettazione e allo sviluppo di SIT orientata ad applicazioni per l'amministrazione locale (province, regioni, ecc.), per cui alcuni esempi e alcune argomentazioni potrebbero sembrare eccessivamente legate a quel campo. Speriamo comunque che il corso sia abbastanza generale per progettare anche applicazioni di tipologie diverse, vista la vastità di applicazioni che richiedono informazioni georeferenziate.

Questa dispensa è suddivisa in capitoli, uno per unità didattica presentata durante il corso, ad ogni capitolo è associato un insieme di lucidi che riassumono e focalizzano i punti più importanti dell'unità didattica. Al termine del capitolo vengono presentati alcuni esercizi per valutare l'apprendimento, per approfondire alcuni aspetti e per impraticarsi con le notazioni.

In dettaglio nel Capitolo 1 si presentano delle tecniche di pianificazione per SIT, il suo ciclo di vita e le fasi, il Capitolo 2 ci si concentra sull'analisi dei requisiti specializzata per SIT. Il Capitolo 3 affronta la fase, importante nei SIT, di scelta delle fonti di dati da adottare e degli strumenti. Il Capitolo 4 affronta il problema della modellazione concettuale, presentando la metodologia generale per la progettazione e specializzandola per SIT; inoltre si osservano alcuni esempi di modellazione, si descrivono alcuni standard di metadati esistenti e si spiega come passare, una volta scelto lo strumento SIT da adottare, dalla modellazione concettuale alla progettazione logica del geo-database. Il Capitolo 5 descrive la fase di pianificazione e costruzione del database, studiando le regole generali per l'acquisizione e la conversione dei dati. Nel Capitolo 6 si descrive come pianificare e realizzare un prototipo del sistema e come tale prototipo possa essere utile per le fasi successive e per la valutazione mediante benchmark degli strumenti (hardware e software) e, nel caso dei SIT, anche delle fonti dei dati. Nel Capitolo 7 si descrivono brevemente le restanti fasi del progetto e si forniscono delle regole per la valutazione dei risultati ottenuti.

Indice

Capitolo 1: Pianificazione e sviluppo di un SIT	6
Introduzione.....	6
Definizione di un sistema informativo territoriale	6
Peculiarità dei SIT – Perché è necessaria una fase di pianificazione	6
I SIT sono un’innovazione tecnologica.....	7
Problemi organizzativi e gestionali	7
Impatto sull’organizzazione	7
Gestione di un progetto SIT	8
Benefici attesi e investimenti necessari.....	8
Il ciclo di vita (di sviluppo) di un SIT	9
Pianificazione del ciclo di vita e stima dei costi	9
Descrizione del ciclo di sviluppo SIT	10
Le fasi di sviluppo del SIT	12
Studio di fattibilità.....	12
Analisi dei requisiti	12
Progettazione concettuale di un SIT.....	12
Rassegna dei dati disponibili	14
Rassegna dell’hardware e del software disponibile.....	14
Progettazione e pianificazione dettagliata del database	14
Prototipo e Benchmark.....	14
Costruzione del database	14
Integrazione del sistema SIT	15
Sviluppo di applicazioni SIT	15
Uso e manutenzione del SIT	15
Capitolo 2: Analisi dei Requisiti	16
Introduzione.....	16
Raccolta dei requisiti	17
Utilizzi di un SIT nella Pubblica Amministrazione	18
Browsing	19
Visualizzazione e stampa di mappe.....	19
Interrogazione e visualizzazione del risultato	19
Overlay Topologico.....	20
Buffering	20
Modellazione spaziale	20
I tipi di dati usati nella pubblica amministrazione	21
Documenti necessari per la raccolta dei requisiti	21
Documentazione dell’attività di un SIT	26
Le interviste	27
La specifica dei requisiti.....	28
Master Data List	32
Esempio di documentazione.....	33
Capitolo 3: Rassegna dati e valutazione strumenti.....	39
Introduzione.....	39
Rassegna dei dati disponibili	39
Possibili sorgenti di dati : formati	39
Possibili sorgenti di dati : tipi di dato.....	40

Descrizione e valutazione dati.....	40
Rassegna hardware e software.....	42
Capitolo 4: Modellazione concettuale e logica.....	43
Introduzione.....	43
La progettazione concettuale.....	43
Modellazione concettuale.....	45
Modello entità-relazioni	46
Specifica e progettazione.....	46
Peculiarità dei dati geografici.....	48
Modellare dati spaziali	49
Modello di Calkins	50
Identificazione delle entità	51
Identificazione e definizione della componente spaziale delle entità	51
Identificazione degli attributi	52
Identificazione e modellazione delle associazioni	53
Sottoclassi e superclassi	55
Modelli logici	56
Modello relazionale con layers.....	56
Modello UML ed estensione ESRI UML.....	58
Metadati e Standard di settore	59
Capitolo 5: Pianificazione e costruzione del database	60
Introduzione.....	60
La fase di pianificazione e progettazione nel ciclo di vita del SIT	60
Scelta delle sorgenti dati.....	61
Progetto logico e fisico del geodatabase	61
Pianificazione	62
Pianificare l'acquisizione e la conversione	63
Tecniche e strumenti d'acquisizione e conversione	63
Pianificare il processo di acquisizione	64
Pianificare la manutenzione	66
Capitolo 6: Il progetto pilota ed il benchmark	67
Introduzione.....	67
Il concetto di prototipo	67
Progettare un prototipo	67
Obiettivo di un prototipo	67
Pianificazione del prototipo.....	68
Scelta delle applicazioni	68
Scelta dei dati	69
Scelta dell'area di interesse	69
Specifiche preliminari di acquisizione dei dati	69
Scelta dell' hardware e del software.....	69
Scelta del fornitore dei servizi per l'acquisizione dei dati	70
Valutazione del prototipo	70
Esecuzione del progetto pilota	71
Preelaborazione dei dati (Scrubbing)	71
Gestione dei cambiamenti e del flusso di documenti	71
Preparare l'impatto del progetto pilota.....	71
Scelta del personale	72

Ottenere il feedback del personale	72
Valutazione del prototipo	72
Informazioni da ottenere utilizzando un prototipo	72
Il benchmark	73
Pianificazione	74
Obiettivo	74
Principi di base	74
Preparazione delle specifiche	75
Scelta del luogo e dei partecipanti	75
Preparazione dei dati	75
Pianificazione dei tempi del benchmark	75
Distribuzione delle specifiche dei dati ai partecipanti	75
Logistica	76
Cambiamenti delle specifiche	76
Definire i criteri del benchmark	76
Capitolo 7: Fasi finali	77
Introduzione	77
Acquisto dell'hardware e del software	77
Integrazione del sistema SIT	78
Sviluppo di applicazioni SIT	78
Uso e manutenzione del SIT	79
Appendice A: Glossario	80

Capitolo 1: Pianificazione e sviluppo di un SIT

Introduzione

Sviluppare un *Sistema Informativo Territoriale* (SIT o GIS) non è un problema semplice. Per spiegarne il motivo bisogna prima chiarire cosa è un SIT ed in cosa differisce dagli altri sistemi informativi. Per questo motivo in questo capitolo definiamo il SIT dal punto di vista ingegneristico e gestionale, presentandone le peculiarità e i problemi che si devono affrontare per svilupparlo (Sezioni 1 e 2).

Per progettare e pianificare lo sviluppo di un sistema complesso come un SIT è indispensabile una metodologia. L'idea di base è quella di scomporre il problema in fasi (analisi, progettazione del sistema, acquisizione dei dati, costruzione della base di dati, ecc.). Ogni fase risolve una parte del problema, coinvolgendo determinati ruoli e producendo un risultato (nella maggior parte dei casi documentazione per le fasi successive). La suddivisione del lavoro in fasi è guidata anche dalla rilevanza (e dal costo) di alcuni problemi nel sistema specifico; ad esempio, dato che per i SIT è importante la scelta dell'hardware e del software, viene introdotta una fase specifica per risolvere questo problema. Altri problemi (come l'acquisizione e l'immissione dei dati, cruciali per i SIT) pongono maggiore enfasi su fasi come la costruzione della base di dati, la rassegna dei dati disponibili e la progettazione del prototipo.

In questo capitolo si descrivono le fasi del ciclo di vita di un SIT, descrivendone le caratteristiche principali e studiando come lo sviluppo di tali fasi possa essere pianificato in modo da diminuire il più possibile i costi (sia in termini di tempo, che di personale coinvolto).

Definizione di un sistema informativo territoriale

Un *sistema informativo territoriale* (SIT o GIS) può essere definito come “...un sistema informativo basato su computer che cerca di catturare, immagazzinare, manipolare, analizzare e visualizzare dati spaziali con associati degli attributi, al fine di risolvere ricerche complesse, pianificare e gestire problemi” (Fischer and Nijkamp, 1992).

Dal nostro punto di vista il SIT non è altro che un particolare *sistema informativo* (cioè un insieme di hardware, software, dati, procedure aziendali, conoscenze, persone che raccolgono, gestiscono, elaborano, distribuiscono informazioni) in cui la maggior parte delle informazioni gestite hanno contenuto *geo-referenziato*.

I SIT sono molto utili come supporto alla pianificazione, alla gestione e alla ricerca e hanno innumerevoli applicazioni a partire dalla gestione del territorio fino ad arrivare all'archeologia e allo studio dell'ambiente.

Peculiarità dei SIT – Perché è necessaria una fase di pianificazione

I SIT appartengono alla categoria di sistemi informativi che per divenire realmente utili richiedono la presenza di un *sistema informatico* che gestisca una *base di dati*.

A differenza di molte applicazioni per personal computer, per le quali un utente può cominciare ad utilizzare il sistema subito dopo l'acquisizione dell'hardware e del software, l'uso del sistema informatico di un SIT richiede:

- la preventiva creazione di un grande database spaziale (o spazio/temporale);
- l'acquisto di hardware e software appropriato;
- lo sviluppo di alcune applicazioni;
- l'integrazione ed il test delle svariate componenti.

Queste fasi sono lunghe e complesse, così complesse da richiedere una pianificazione prima che qualsiasi dato, hardware e software sia acquisito. Lo scopo di questo corso è di descrivere il processo di pianificazione del SIT, in tutte le sue fasi, e di fornire alcuni esempi di come riuscire a svolgere tali fasi.

I SIT sono un'innovazione tecnologica

I SIT hanno tratto vantaggio nel rapido sviluppo della tecnologia degli ultimi decenni, per affrontare la sfida di memorizzare e analizzare dati spaziali. Attualmente i sistemi informatici che gestiscono SIT sono essi stessi un'innovazione tecnologica, e per questo motivo il loro uso non è semplice come altri sistemi che sono stati studiati e testati per molto tempo.

Si possono incontrare parecchi problemi adottando un sistema informatico per SIT:

- il personale non riesce a capire adeguatamente la tecnologia, se non dopo lunghi periodi di formazione;
- la stima del tempo necessario allo sviluppo differisce dal tempo realmente necessario;
- grande incertezza relativamente ai costi;
- buone probabilità che siano necessari cambiamenti programmatici durante le fasi di sviluppo, etc.

Il punto cruciale è che queste sono condizioni normali se si adotta una nuova tecnologia.

Bisogna evidenziare che, a causa enormi investimenti e tempi di start-up necessari, un progetto SIT mal pianificato ha un'alta probabilità di fallimento, per cui è fondamentale pianificare e stimare adeguatamente il progetto prima di realizzarlo.

Problemi organizzativi e gestionali

Impatto sull'organizzazione

L'adozione di una tecnologia all'interno di un sistema informativo può causare cambiamenti anche radicali nell'organizzazione di tale sistema. Questo è un problema generale che si riscontra quando la parte cruciale del sistema informatico è costituita da una base di dati.

Uno degli obiettivi dello sviluppo di una base di dati è l'eliminazione di collezioni di dati ridondanti in modo da ottimizzare le prestazioni e evitare i problemi di inconsistenza derivati dalla duplicazione dei dati. D'altra parte, nei sistemi informativi precedenti all'adozione di database i dati sono tipicamente immagazzinati in collezioni legate ad un singolo progetto o individuo con conseguente ridondanza. Per questo motivo, quando si introduce un database all'interno di un sistema informativo, si tende a modificare il flusso ed il mantenimento dei dati, fino ad ottenere un database centralizzato a cui gli utenti possano accedere, eventualmente avendo visibile solo la parte di loro interesse. Questo non soltanto riduce la ridondanza, ma migliora l'accuratezza dei dati e ci aiuta a capire quando un dato è usato da più utenti.

L'importanza di tale riorganizzazione è ancora maggiore nel caso dei SIT, in quanto l'enorme costo dell'acquisizione dei dati può portare allo sviluppo di un unico database per più unità dello stesso ente o per più enti, in modo da poter cooperare (sia in termini di spese che di conoscenze) per la creazione del sistema.

Si crea così il concetto di *corporate database*, cioè un'unica risorsa di dati per l'intera organizzazione (possibilmente con più enti).

I corporate database sono particolarmente importanti nell'ambito degli enti della pubblica amministrazione, per i quali si può creare un corporate database per gestire più sedi della stessa amministrazione o più entità governative della stessa regione. In particolare la condivisione dei

Capitolo 1: Pianificazione e sviluppo di un SIT

dati tra i diversi enti della pubblica amministrazione è virtualmente necessaria per garantire il successo di una progettazione a lungo termine.

Comunque decidere se e come creare un corporate database non è un problema tecnico, piuttosto una questione politica e di coordinazione e cooperazione tra più aziende (o enti).

Gestione di un progetto SIT

I SIT sono tuttora una tecnologia in via di sviluppo, per cui gli individui coinvolti nel progetto (utenti finali del SIT, tecnici e responsabili di progetto) potrebbero avere delle aspettative molto diverse tra loro, alcune basate su una conoscenza non approfondita dell'informatica.

Tali aspettative potrebbero essere eccessive e questo, assieme al lungo tempo che necessita lo sviluppo di un SIT, potrebbe portare il progetto ad un fallimento.

Per questo motivo è importante coinvolgere il personale appartenente ai diversi ruoli dell'organizzazione committente durante la progettazione, chiarendo sin dall'inizio quali sono i requisiti (ed i tempi di sviluppo) del sistema ed, allo stesso tempo, fornendo al personale la terminologia e le conoscenze che sono la base per la formazione all'uso di SIT.

È importante cogliere l'opportunità della progettazione del SIT, per cercare di migliorare il sistema esistente, chiarendo ruoli e responsabilità all'interno del sistema e fornendo gli strumenti adeguati.

Le regole per la progettazione di un SIT si possono riassumere in:

- evidenziare i vantaggi che i SIT porteranno agli utenti e all'intera organizzazione;
- richiedere un buon livello di competenza da parte di tutti i partecipanti;
- dare a tutti la possibilità di accedere in modo effettivo ai dati del SIT;
- ridimensionare le aspettative degli utenti;
- minimizzare il più possibile il lasso di tempo che intercorre tra la stesura dei requisiti e la disponibilità di un prodotto utilizzabile;
- sviluppare una attitudine positiva al cambiamento all'interno dell'organizzazione;
- assicurarsi che il livello di tecnologia sia adeguato per gli utenti coinvolti;
- progettare e realizzare un buon progetto pilota (prototipo).

Benefici attesi e investimenti necessari

Benefici

I benefici che un'organizzazione dovrebbe ottenere utilizzando un SIT sono: *efficienza e efficacia*. L'efficienza si riconduce ad una diminuzione di tempo e di personale necessario per l'esecuzione delle operazioni per cui è stato progettato il sistema (interrogazioni, produzione di mappe, etc.). Per valutare il potenziale guadagno in termini di efficienza si possono stimare i tempi medi di risposta ad una interrogazione eseguita manualmente o mediante SIT, e moltiplicare la differenza tra i due per il numero di interrogazioni attese.

L'efficacia è più difficile da misurare, ma può essere vista come la possibilità di poter effettuare nuove operazioni, che precedentemente erano ritenute troppo complesse o addirittura non erano neanche prese in considerazione (ad esempio analisi spaziali). Dato che queste operazioni sono nuove, un confronto con il sistema precedente non è possibile. Comunque, anche se non misurabile, il beneficio ottenuto potrebbe essere sostanziale; in generale si ottiene un miglioramento nella pianificazione delle attività dell'organizzazione e aumenta notevolmente il supporto alle decisioni (di conseguenza aumenta l'efficacia delle decisioni intraprese).

Capitolo 1: Pianificazione e sviluppo di un SIT

Investimenti necessari

Sviluppare un SIT comporta investimenti in cinque aree: hardware, software di base, dati geografici, software applicativo e formazione del personale. L'acquisto dell'hardware e del software sono visti spesso erroneamente come l'attività più dispendiosa nello sviluppo del SIT. La ricerca, condotta al "*National Center for Geographic Information and Analysis*" di SUNY-Buffalo, ha dimostrato che i costi maggiori nello sviluppo del SIT sono dovuti alla costruzione del geodatabase (dal 60% all'80% dell'intera spesa). Anche i costi successivi di mantenimento sono dominati dal costo relativo all'aggiornamento dei dati.

Requisiti per il personale di un SIT

La selezione del personale per il SIT è un aspetto critico. È necessario personale con esperienza nei seguenti campi:

- gestione del progetto SIT (ruolo responsabile di progetto);
- competenza nella gestione della base di dati geografica (ruolo amministratore del database);
- sviluppo delle applicazioni per utenti database (analisti e programmatori SIT).

La creazione iniziale del database SIT (con relativa digitalizzazione) richiederà una certa quantità di personale specializzato, dipendente dalla quantità di dati da acquisire. In alternativa all'espansione del personale per l'acquisizione, ci si può rivolgere a consulenti e fornitori specializzati in acquisizione di dati. La fase di acquisizione è molto settoriale e può essere facilmente delegata ad enti competenti, con tali enti si stipulerà un contratto specificando i requisiti di qualità richiesti.

Se possibile i tre campi di esperienza richiesti devono essere coperti dal personale, eventualmente assumendo nuovo personale o formando il personale esistente.

È importante che il responsabile di progetto sia coinvolto in ogni fase del progetto SIT.

Requisiti ulteriori sono: la formazione degli utenti in informatica generale, principi di basi di dati, e utilizzo del SIT. Questi requisiti si possono facilmente ottenere attraverso dei corsi di formazione offerti dalla maggior parte delle case di distribuzione di prodotti per SIT.

Il ciclo di vita (di sviluppo) di un SIT

Pianificazione del ciclo di vita e stima dei costi

Il ciclo di vita di un sistema informativo è suddiviso in fasi, ogni fase può essere svolta indipendentemente, coinvolge determinati ruoli e produce un risultato. Spesso alcune fasi necessitano di risultati di fasi precedenti, il che crea una dipendenza temporale tra le fasi imponendo un ordinamento tra di esse. Le dipendenze temporali limitano le possibilità di risolvere contemporaneamente le fasi che dipendono tra loro, aumentando così il tempo necessario allo sviluppo. Pianificare lo sviluppo SIT, vuol dire cercare di ridurre il più possibile tali dipendenze in modo da poter procedere parallelamente su più fasi, riducendo così il lasso di tempo che intercorre tra l'analisi e l'uso del sistema. Inoltre le fasi più complesse e dispendiose dovrebbero essere identificate in modo da pianificare adeguatamente l'uso delle risorse.

Come detto in precedenza il problema più difficile nello sviluppo di un SIT è costituito dalla costruzione del geodatabase. Questa fase è quella più dispendiosa in termini di tempo e di denaro, e richiede un notevole sforzo in termini di progettazione e pianificazione. Di conseguenza, nel ciclo di sviluppo da noi presentato riveste particolare importanza la pianificazione della base di dati.

È importante coinvolgere tutti gli utenti potenziali in tutte le fasi dello sviluppo del SIT. Il personale coinvolto gode dei seguenti benefici:

Capitolo 1: Pianificazione e sviluppo di un SIT

- può descrivere i propri requisiti all'analista SIT;
- può apprendere cosa il SIT sarà capace di produrre per loro;
- può capire la natura del ciclo di sviluppo del SIT, il tempo ed i costi necessari.

I potenziali utenti devono comprendere che ci sarà un lasso di tempo significativo tra la prima fase di analisi e l'istante in cui il SIT potrà realmente essere utilizzato. Questo è dovuto soprattutto alla lunghezza e complessità delle fasi di costruzione del database e di acquisizione dei dati.

Le fasi di *rassegna* dei dati, dell'hardware e del software disponibile possono influenzare la pianificazione del progetto SIT, ottenendo così un piano di sviluppo più dettagliato (con più o meno rilevanza alle fasi di acquisizione e scelta dell'hw/sw).

Stimare e pianificare il costo del SIT, è in qualche modo un compito difficile. Innanzitutto, è necessario tener presente che la costruzione del database sarà probabilmente la spesa più grande (se i dati vengono acquisiti da mappe e documenti interni, tale costo si aggira tra il 70% e l'80% del costo totale del sistema).

Quindi, importare (o acquistare) dati già digitalizzati da altri SIT, siano essi enti pubblici o privati, può portare ad un notevole risparmio (sia in termini di tempo che di denaro). Anche partecipare o organizzare progetti SIT che coinvolgono più enti con condivisione dei dati, può ridurre i costi. È importante includere nella pianificazione e nella stima dei costi, anche i costi a lungo termine di manutenzione, conservazione e archiviazione dei dati.

Descrizione del ciclo di sviluppo SIT

Il ciclo di sviluppo di un SIT parte dall'analisi dei requisiti e termina con l'uso e la manutenzione di un SIT funzionante. Le fasi sono logicamente progressive, cioè concettualmente ogni fase deve cominciare solo dopo che la precedente è terminata (ciclo di vita a cascata). Questa visione è soltanto astratta, perché nella realtà alcune attività del processo potrebbero avvenire contemporaneamente ad altre (pianificazione per diminuire il tempo di completamento), oppure si potrebbe seguire un approccio iterativo (a spirale) od infine si potrebbero apportare delle modifiche dipendenti da caratteristiche particolari del SIT in questione.

Il ciclo di sviluppo di un SIT si basa su una tecnica top-down, cioè prima si decide cosa il SIT dovrebbe fare, in seguito si decide in che modo ed alla fine si realizza il SIT.

Seguendo questa filosofia, per prima cosa si identificano i requisiti, in seguito si esegue una rassegna delle risorse disponibili (dati, hardware, software, personale, etc.), il terzo passo consiste nella creazione di uno schema preliminare adeguatamente testato, ed infine, si procede all'acquisizione dei dati ed alla creazione del database.

Capitolo 1: Pianificazione e sviluppo di un SIT

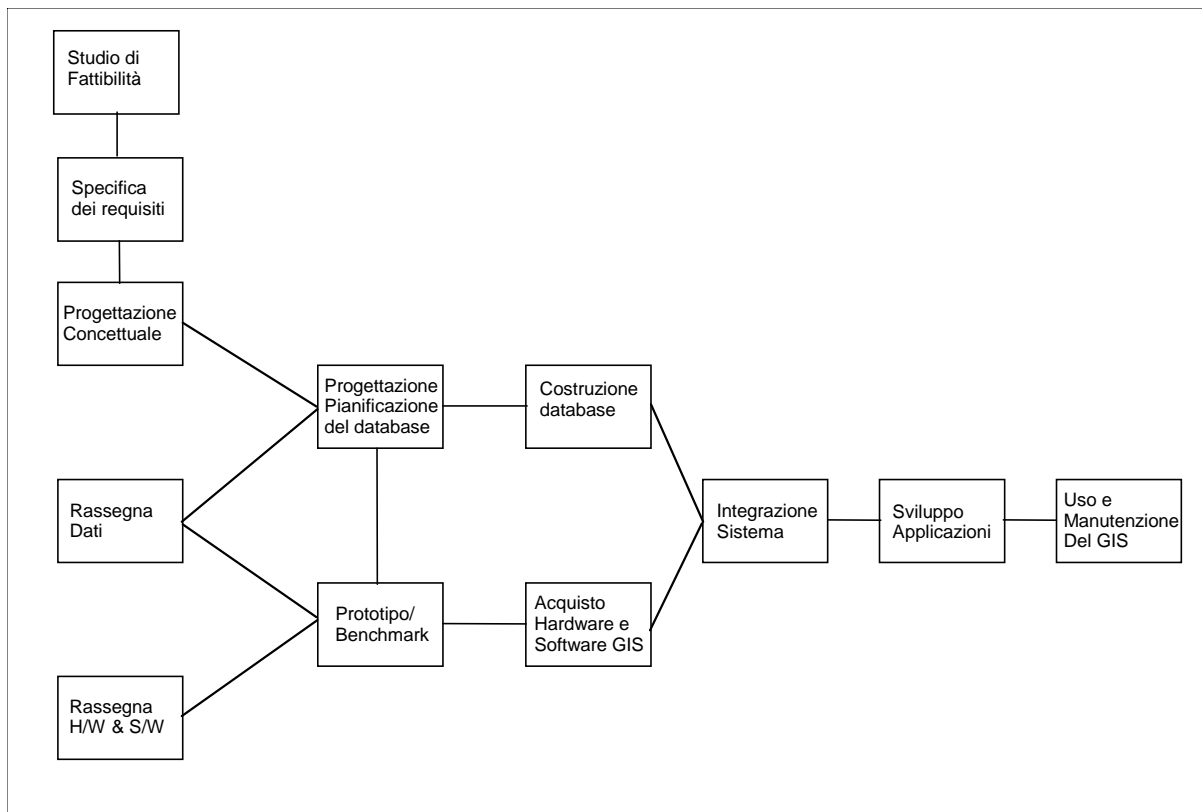


Figura 1 - Ciclo di sviluppo di un SIT

La Figura 1 mostra il ciclo di sviluppo di un SIT, che è descritto tramite dodici passi. Le dodici fasi del ciclo di sviluppo sono:

- 1) Studio di fattibilità
- 2) Specifica dei requisiti
- 3) Progettazione concettuale
- 4) Rassegna dei dati disponibili
- 5) Rassegna dell'hardware e software disponibile
- 6) Pianificazione e progettazione dettagliata del database
- 7) Costruzione del database
- 8) Prototipo/Benchmark
- 9) Acquisto e configurazione di hardware e software SIT
- 10) Integrazione del sistema
- 11) Sviluppo di applicazioni SIT
- 12) Uso e manutenzione del SIT

Questo è soltanto un possibile modo per dividere l'insieme di attività che devono essere realizzate per costruire un SIT con successo. Ci sono altri modi di esprimere ed organizzare queste attività, questa particolare struttura è stata scelta perché evidenzia lo sviluppo dei dati (definizione, modellazione, documentazione, raccolta, acquisizione, manutenzione e conservazione dei dati).

Ciò che realmente importa, non è il numero o l'ordine delle fasi, ma il fatto che tutte queste fasi devono necessariamente essere completate per ottenere un buon SIT.

Capitolo 1: Pianificazione e sviluppo di un SIT

In alcune situazioni, diversi metodi potrebbero essere più appropriati di quelli presentati in questo capitolo o potrebbe essere necessario un maggior livello di dettaglio. Non importa quanto sia semplice o complesso un ambiente SIT, tutte le fasi di cui sopra dovrebbero essere completate ad un appropriato livello di dettaglio.

In seguito forniremo alcuni esempi di progettazioni a diversi livelli di dettaglio.

Le fasi di sviluppo del SIT

Descriviamo brevemente ciascuna delle dodici fasi del ciclo di sviluppo. Le fasi più importanti saranno approfondite nei prossimi capitoli.

Studio di fattibilità

Data la delicatezza ed i costi elevati dei progetti SIT è importante che la prima fase di analisi del SIT riguardi lo *studio di fattibilità*, cioè prima di procedere con la progettazione e la realizzazione del sistema, bisogna accertarsi che i requisiti il sistema sia costruibile in tempi e costi ragionevoli. Un buon studio di fattibilità evita di investire in un progetto intrinsecamente fallimentare, ed, in alcuni casi, ridimensiona notevolmente i requisiti finali rendendoli “fattibili”. Lo studio di fattibilità richiede una prima stesura dei requisiti ed una prima valutazione (stima) dei costi.

Lo studio di fattibilità non verrà affrontato in questa dispensa in modo dettagliato. Le tecniche utilizzate sono simili a quelle adottate per l’analisi dei requisiti applicate ad un livello di astrazione più alto.

Analisi dei requisiti

L’analisi dei requisiti è finalizzata a definire:

- l’insieme delle operazioni che saranno necessarie agli utenti SIT;
- l’insieme dei dati geografici interessanti.

Le informazioni necessarie all’analisi dei requisiti sono ottenute attraverso interviste a potenziali utenti del SIT ed eventuali documentazioni del sistema esistente. Al termine di una serie di interviste si ottengono le descrizioni delle applicazioni del SIT, una serie di dati ritenuti importanti da uno o più utenti e una descrizione delle procedure di gestione del SIT; da questi dati bisogna estrarre le informazioni necessarie per la stesura dei requisiti. Per documentare i risultati delle interviste degli utenti coinvolti, si compilano dei moduli standard. Le informazioni ottenute durante la specifica dei requisiti sono il punto di partenza della fase di *progettazione concettuale*.

Progettazione concettuale di un SIT

La progettazione concettuale di un sistema è per la maggior parte un classico esercizio di progettazione di basi di dati. Esso include la modellazione formale (preparazione del schema concettuale) della base di dati del SIT e i passi iniziali dell’attività di pianificazione del database. La pianificazione del database è il passo più importante dello sviluppo del SIT in quanto comprende la gestione le diverse attività che riguardano i dati durante tutto il ciclo di vita del SIT, l’insieme di tali attività è definito: *ciclo di vita dei dati* (figura 2). Il ciclo di vita dei dati è costituito da:

- identificazione dei dati (che avviene durante la specifica dei requisiti);
- inclusione dei dati nello schema concettuale (progettazione concettuale);
- creazione dei metadati (progettazione di dettaglio);
- raccolta e inserimento dei dati nel geodatabase (costruzione del database);

Capitolo 1: Pianificazione e sviluppo di un SIT

- aggiornamento e mantenimento (uso e manutenzione).

La progettazione concettuale si conclude con una progettazione architeturale del sistema che identifica l'architettura SIT da adottare (centralizzata, distribuita, client/server, etc.) e fornisce un'idea della dimensione del SIT. Tutto questo viene realizzato con riferimento ai sistemi che dovranno essere interfacciati con il SIT.

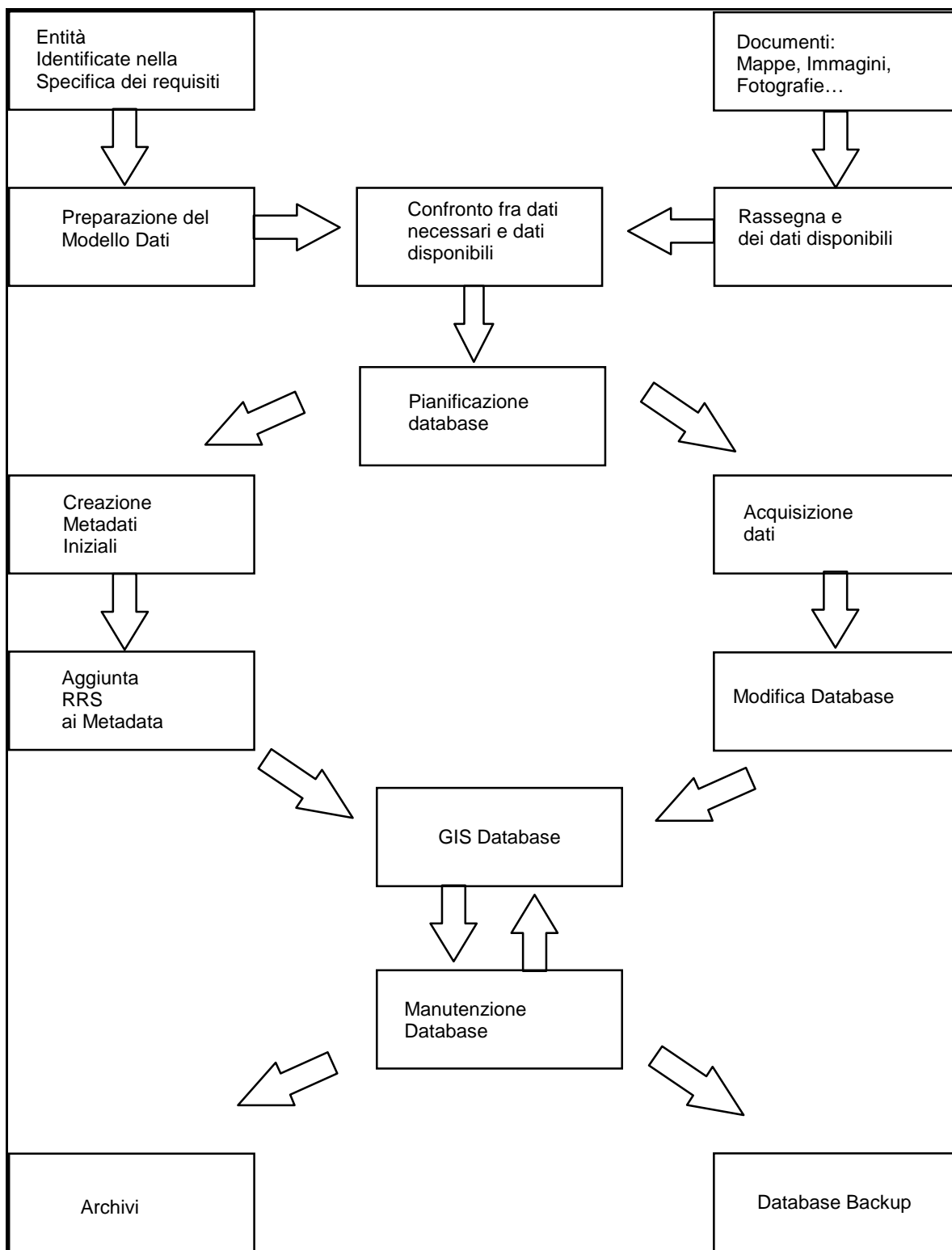


Figura 2 - Ciclo di vita dei dati

Capitolo 1: Pianificazione e sviluppo di un SIT

Rassegna dei dati disponibili

La rassegna dei dati disponibili può iniziare quando i dati necessari sono stati specificati durante l'analisi dei requisiti. In questa fase si dovranno passare in rassegna e documentare tutte le mappe, i dati digitali, le tabelle che si trovano all'interno dell'organizzazione, come pure dati disponibili attraverso altre fonti.

Potrebbero essere inclusi nell'inventario anche dati contenuti in altri SIT (eventualmente di altre organizzazioni) che possono essere sfruttati e importati. Questa è la fase in cui bisogna investigare la possibilità di costruire un corporate database o di sfruttare un meccanismo di condivisione di dati tra più organizzazioni. C'è anche la possibilità che alcune case distributrici di software SIT, potrebbero essere in grado di fornirci dati utili, per cui bisogna investigare in tal senso. La documentazione prodotta a questo punto dovrebbe essere sufficiente per valutare ogni possibile fonte di dati. L'informazione raccolta formerà anche parte dei metadati del geodatabase finale.

Rassegna dell'hardware e del software disponibile

Quasi tutti i SIT si baseranno su software SIT disponibile in commercio. Per questo motivo è necessario condurre una rassegna dei sistemi per SIT disponibili. Durante questa attività, le funzionalità di ogni sistema per SIT disponibile dovrebbero essere documentate per una valutazione futura.

Progettazione e pianificazione dettagliata del database

La fase di progettazione e pianificazione dettagliata del database include le seguenti attività: sviluppo di uno schema logico o fisico della base di dati basato sullo schema concettuale precedentemente costruito, valutazione delle potenziali fonti di dati, stima delle quantità di dati geografici, stima del costo di costruzione del database SIT e preparazione del piano di acquisizione dei dati. Parallelamente alla pianificazione dettagliata del database, può essere eseguita la costruzione del prototipo e/o il benchmark. L'informazione ottenuta dal prototipo e dai test potrebbe essere utile per avere un'idea delle risorse necessarie (spazio disco, memoria, etc.) e per determinare quanto sarà pesante la fase di sviluppo di applicazioni.

In seguito, i piani per l'acquisizione e la formazione di personale, per l'acquisto e l'installazione di strumenti dovranno essere completati. Dopo la preparazione di questi piani, l'intero costo del SIT sarà conosciuto e si potrà verificare se tali cifre rientrano nelle stime date durante lo studio di fattibilità, ed eventualmente riconsiderare la fattibilità del progetto.

Prototipo e Benchmark

Il prototipo ed i test di benchmark servono a dimostrare e verificare le funzionalità del software del SIT, evidenziando cosa i sistemi per SIT commerciali esistenti sul mercato possono fare. Questi test sono utili per dimostrare agli utenti potenziali ed al responsabile di progetto, cosa il SIT potrà fare per loro. Inoltre potranno essere valutate le performance del sistema.

Costruzione del database

La costruzione del database (detta anche talvolta *acquisizione del database*) è il processo di costruzione del database digitale a partire dai dati disponibili (mappe e tabulati). Questo processo dovrebbe essere pianificato durante le precedenti fasi mettendo in evidenza come gestire la costruzione e come assicurare e controllare la qualità dei dati acquisiti. Il processo di acquisizione è spesso demandato e coinvolge una grande quantità di documenti e mappe. Una gestione efficace è il fattore decisivo per una buona acquisizione dei dati.

Capitolo 1: Pianificazione e sviluppo di un SIT

Integrazione del sistema SIT

A differenza di molte altre applicazioni per computer, un SIT non appartiene alla categoria “plug and play”. I diversi componenti del SIT devono essere acquistati in accordo con le specifiche documentate. Il database deve essere creato in modo attento e organizzato. Una volta che tutte le componenti individuali sono state pronte, esse devono essere integrate e testate. Gli utenti devono essere informati sul sistema, devono essere formati se necessario e devono essere forniti di adeguata assistenza per cominciare ad usare il SIT. Le parti del SIT che potrebbero sembrare ben funzionanti individualmente, potrebbero non funzionare correttamente quando vengono integrate nel sistema. Il personale del SIT deve risolvere tutti i questi problemi prima che gli utenti possano accedere al SIT.

Sviluppo di applicazioni SIT

Per “Applicazione” intendiamo qualsiasi procedura che viene eseguita su di un SIT. Per prima cosa, ci sono le applicazioni del database. Queste sono tutte le funzioni necessarie per creare, modificare, costruire e mantenere la base di dati, tali funzioni sono normalmente svolte dal personale del sistema SIT. Alcuni utenti potrebbero avere l’accesso in scrittura solo a parte del database, comunque l’intera base di dati dovrebbe essere controllata dall’*amministratore del database*.

Le altre applicazioni sono dette *applicazioni utente*. Attualmente molti software base per SIT forniscono un certo numero di applicazioni semplici (es. Visualizzazione di mappe, interrogazione del database). Applicazioni più complesse, o non standard, devono essere sviluppate utilizzando un macro-linguaggio di programmazione.

La maggior parte dei sistemi per SIT hanno un macro-linguaggio per svolgere questo compito (es. Arc Macro Language (AML) in ARC/INFO™, Avenue in ArcView™). Tutte le applicazioni da sviluppare devono essere descritte durante la stesura dei requisiti in degli appositi *moduli per applicazioni SIT*.

Uso e manutenzione del SIT

Dopo aver descritto il piuttosto arduo compito di creazione del SIT, possiamo dire che l’uso e la manutenzione del SIT e del suo database richiederà tanta attenzione quanta ne era necessaria per costruirli.

La maggior parte dei database per SIT sono molto dinamici, cambiano quasi quotidianamente e gli utenti penseranno immediatamente ad applicazioni aggiuntive che vorrebbero fossero sviluppate. Bisogna creare e seguire delle procedure formali per tutte le attività di manutenzione e di aggiornamento. Tali procedure sono necessarie per garantire l’operatività e il buon funzionamento del SIT.

Capitolo 2: Analisi dei Requisiti

Introduzione

L'analisi dei requisiti di un SIT deve fornire una descrizione sistematica di come sono impiegati i dati spaziali nella normale attività lavorativa di un'organizzazione e di come operano ed interagiscono i diversi reparti. L'analisi dei requisiti termina con la stesura di un documento che definisce le esigenze dell'organizzazione e specifica formalmente gli obiettivi da raggiungere. Tale documento sarà il punto di riferimento per le fasi successive di progettazione e realizzazione del sistema.

Nel caso si voglia costruire un corporate database o si voglia condividere dati fra più enti, durante la fase di analisi bisogna porre particolare attenzione all'interazione dei vari reparti coinvolti. L'obiettivo, in questo caso, è di uniformare i vari sistemi di modo da permettere la condivisione e lo scambio di dati.

L'obiettivo dell'analisi dei requisiti è documentare tutte le informazioni necessarie per pianificare lo sviluppo del SIT. Queste informazioni sono descritte nel documento di specifica dei requisiti e possono essere raggruppate nelle seguenti categorie:

- **Descrizione del contesto.** Una breve descrizione dell'ambiente (del sistema informativo) in cui opera l'organizzazione che chiarisca in linea di massima quali sono le attività, i ruoli coinvolti e le interazioni col sistema.
- **Applicazioni da sviluppare e funzioni SIT richieste da queste.** Nella valutazione dei ruoli e del flusso di lavoro all'interno di un reparto è importante determinare le operazioni che possono essere realizzate "esclusivamente" o più efficientemente con l'ausilio del SIT. La specifica di tali operazioni è importante perché esse costituiscono la base dell'insieme di applicazioni (SIT e non) che dovranno essere implementate durante la fase di sviluppo del software. Per ogni applicazione identificata bisogna determinare le funzioni (funzionalità) SIT che sono necessarie per la sua implementazione. Queste funzioni includeranno operazioni standard quali: interrogazioni alla base di dati, visualizzazione di mappe, funzioni di analisi spaziale (es. il calcolo del cammino minimo (routing), l'overlay, il buffering); oppure, qualora si rendesse necessario, fanno riferimento ad operazioni personalizzate ad hoc secondo le esigenze del cliente.
- **Dati necessari per la creazione di una base di dati geografica.** I reparti dell'organizzazione usano spesso dei dati *georeferenziati* (cioè associati a delle coordinate geografiche in un qualche sistema di riferimento). Altri dati provenienti da cartografia o da tabelle di diverso formato hanno altri tipi di riferimento spaziale come ad esempio indirizzi e codici postali. L'analisi dei requisiti descrive tutti i dati facendo riferimento in particolare ai dati georeferenziati ed al ruolo che essi avranno nelle future applicazioni SIT.
- **Procedure per la generazione e la gestione dei dati.** Esaminando il flusso del lavoro, le operazioni richieste da ogni reparto e le responsabilità nella creazione dei dati si evince quali saranno gli eventuali aggiornamenti e le procedure di gestione del geodatabase che si sta sviluppando. Una volta che tutte queste informazioni sono raccolte ed analizzate per ogni reparto devono essere registrate in un documento in modo da poter essere utilizzate come punto di riferimento nelle fasi successive.

La specifica dei requisiti verrà in seguito usata per:

- progettare la base di dati geografica;

- determinare il software SIT che soddisferà le esigenze dell'organizzazione;
- pianificare l'esecuzione;
- stimare il rapporto costi-benefici.

Un errore comune è quello di ridurre l'analisi dei requisiti ad un semplice inventario delle mappe e dei dati geografici utilizzati in ogni reparto. Quest'approccio dà origine ai seguenti problemi:

- la mancanza totale di informazioni relative alle attività ed ai prodotti di ogni reparto non permette di valutare come il SIT possa essere impiegato per migliorare le capacità produttive e la qualità del lavoro di ogni reparto. Esaminando le attività e i prodotti di ogni reparto è più semplice capire il ruolo che il SIT potrà svolgere all'interno dell'organizzazione.
- l'inventario dei dati tende a focalizzare l'attenzione solo sui dati interni all'organizzazione. Invece in molti casi è necessario far riferimento ad enti esterni, pubblici e non, per acquisire i dati. L'esigenza di acquisire dati può essere valutata adeguatamente analizzando le applicazioni SIT e come i dati saranno utilizzati in esse; in questo modo risulta facile scegliere la fonte più adatta per l'acquisizione.

È importante coinvolgere attivamente personale di tutti i livelli dell'organizzazione durante la fase di analisi, questo serve a:

- comprendere le motivazioni strategiche del progetto;
- capire il modo di lavorare del personale;
- preparare un progetto che permetta di raggiungere meglio gli scopi dell'organizzazione, senza creare problemi al personale;
- coinvolgere e motivare il personale;
- aiutare dirigenza e personale ad avere una percezione corretta delle possibilità del sistema.

La fase di analisi dei requisiti è composta di due parti fondamentali:

- raccolta dei requisiti (che produce una documentazione relativa alle esigenze dei potenziali utenti SIT);
- stesura dei requisiti (che produce un documento di specifica analizzando il risultato della raccolta).

Raccolta dei requisiti

L'aspetto più rilevante dell'analisi dei requisiti consiste nel documentare il contesto operativo in maniera strutturata e standardizzata. Durante la raccolta devono essere identificate le operazioni da svolgere ed i dati coinvolti.

È molto importante adottare (o sviluppare internamente) un metodo standard da utilizzare per la descrizione di tutte le operazioni del SIT, dei processi e dei dati che saranno inclusi nell'analisi dei requisiti. Tale metodologia deve permettere di identificare:

- **Applicazioni SIT.** Queste saranno le operazioni da effettuarsi con l'ausilio di strumenti SIT, come ad esempio preparare una mappa, eseguire interrogazioni al geodatabase o condurre particolari analisi spaziali. Le applicazioni SIT possono essere descritte usando dei formalismi riportati in seguito.
- **Processi dell'organizzazione.** Questi fanno riferimento a processi complessi che necessitano un'adeguata archiviazione delle informazioni richieste; come ad esempio la pubblicazione dei permessi di costruzione, le ispezioni sanitarie, ecc. Il processo

Capitolo 2: Analisi dei Requisiti

può essere descritto attraverso un diagramma di flusso dati (DFD data flow diagram) che analizzeremo più avanti.

- **Dati SIT.** Dati la cui gestione e memorizzazione è necessaria; tali dati possono anche non essere coinvolti in nessuna delle applicazioni descritte precedentemente. Per l'archiviazione sistematica delle specifiche di questi particolari dati deve essere predisposto un metodo specifico.

La raccolta dei requisiti deve essere effettuata mediante:

- **Interviste personali.** Serie di domande finalizzate ad identificare le esigenze del committente e dei futuri utenti del SIT.
- **Raccolta dei documenti usati e prodotti.** Un buon punto di partenza per l'analisi è ottenuto raccogliendo i documenti usati e prodotti dal sistema informativo prima dello sviluppo del progetto.
- **Affiancamento al personale.** Seguire il personale durante lo svolgimento delle operazioni per apprendere direttamente sul campo.

Le interviste individuali devono coinvolgere tutti i livelli dell'organigramma. Una serie di interviste personali è il modo migliore per identificare le esigenze degli utenti. Durante l'intervista l'utente può fornire delle indicazioni che possono rivelarsi informazioni supplementari per chi effettua l'analisi del sistema SIT.

La documentazione prodotta come risultato della raccolta deve comprendere:

- lo scopo del progetto;
- la descrizione del dominio;
- la descrizione dei dati;
- la descrizione delle operazioni.

Utilizzi di un SIT nella Pubblica Amministrazione Questa sezione riporta in sintesi le principali attività che un SIT può svolgere (all'interno della pubblica amministrazione).

Di seguito sono riportati alcuni esempi di utilizzo dei SIT da parte di un ente pubblico:

- assistenza al pubblico;
- gestione di pratiche (es. pratiche edilizie);
- progettazioni politiche del territorio;
- gestione di servizi (bus, acqua, emergenze, etc.).

Queste sono attività tipiche degli enti pubblici, l'impiego della tecnologia SIT le rende più efficienti ed efficaci; lo sviluppo di un SIT faciliterà la gestione delle informazioni geografiche già esistenti, consentirà di realizzare delle applicazioni supplementari apportando notevoli benefici all'ente pubblico. Alcune problematiche (come la gestione delle emergenze, lo studio del traffico stradale e l'individuazione di locazioni ottimali per le stazioni di mezzi pubblici) sono abbastanza comuni e possono essere risolte efficacemente tramite l'uso di un SIT.

Ci sono due punti da tener presenti:

- buona parte dei dati gestiti dagli enti pubblici sono georeferenziati e quindi si gestiscono naturalmente attraverso l'uso di un SIT;
- altri sistemi che utilizzano dati georeferenziati o effettueranno operazioni spaziali, devono essere integrati con il SIT (influenzandone quindi la progettazione). Il SIT non deve essere assolutamente sviluppato separatamente.

Capitolo 2: Analisi dei Requisiti

La maggior parte delle organizzazioni che gestiscono dati geografici, in particolar modo la pubblica amministrazione, richiedono applicazioni che necessitano di funzionalità SIT. Tali funzionalità possono essere classificate e ordinate in base al livello di complessità delle applicazioni che le richiedono. Di seguito è riportata una lista, ordinata secondo questo criterio, delle principali categorie di funzionalità:

- browsing;
- visualizzazione/stampa di mappe;
- interrogazione e visualizzazione del risultato;
- overlay topologico;
- buffering;
- modellazione spaziale.

Browsing

Il browsing è una funzionalità presente in tutti i sistemi SIT e consente all'utente di navigare all'interno della rappresentazione grafica della cartografia alla ricerca di informazioni.

La navigazione all'interno della mappa necessita di diverse operazioni standard come lo zoom,

il pan (che equivale ad uno spostamento della mappa), la selezione interattiva, ecc.

Per le singole mappe, o per zone relativamente piccole, il cervello umano si presta meglio alla ricerca di informazione mediante browsing piuttosto che mediante un'analisi spaziale automatica. Tuttavia, quando il volume dei dati coinvolti aumenta, sono preferibili metodi automatici per la ricerca e l'estrazione di informazione.

Visualizzazione e stampa di mappe

Questa funzione SIT consente di generare (e visualizzare, stampare) una cartografia digitale tramite computer. Tale cartografia è spesso una riproduzione (eventualmente rielaborata in base a dati e corredata da informazioni statistiche) delle mappe cartacee che venivano utilizzate precedentemente all'interno del sistema informativo. La cartografia può essere stampata a diverse scale e quindi a diversi livelli di dettaglio (a scale diverse un fiume può apparire come una linea od un poligono a forma allungata tipo "striscia").

Alcuni esempi sono: una cartografia della città in scala 1:1000, il profilo geografico generale di un piano regolatore, le mappe che rappresentano graficamente dati relativi ai censimenti per ogni zona di interesse, ecc. Bisogna evidenziare che la stampa di mappe comporta delle problematiche aggiuntive rispetto alla visualizzazione mediante browsing. Infatti, nel primo caso lo strumento deve fornire alcune funzionalità per la corretta rappresentazione grafica degli elementi (es. disporre le etichette in modo da evitare sovrapposizioni, personalizzabilità della stampa di etichette e di attributi, gestione della legenda, ecc.).

Interrogazione e visualizzazione del risultato

Il geodatabase del SIT, in quanto base di dati, può essere interrogato opportunamente per estrarre delle informazioni dai dati memorizzati. In generale quello che accade nei SIT è che si vogliono cercare degli elementi geografici che soddisfano una determinata condizione, per permettere questo il sistema SIT deve supportare un meccanismo per la visualizzazione del risultato di interrogazioni geografiche. La capacità di soddisfare le diverse interrogazioni, soprattutto nella fase iniziale, è una delle caratteristiche più utili di un SIT. Ad esempio la query "restituisce le unità residenziali costruite prima del 1989" può essere utilizzata sia per

Capitolo 2: Analisi dei Requisiti

visualizzare a video il risultato (di solito evidenziando gli opportuni elementi della mappa), sia per costruire un nuovo strato informativo che contenga unicamente gli elementi del risultato.

Overlay Topologico

L'overlay topologico opera su strati informativi diversi, ne integra le informazioni e alla fine genera un nuovo strato informativo dove ogni elemento eredita tutte le caratteristiche provenienti dai precedenti. Nell'analisi di un problema di natura geografica spesso l'esame dei singoli strati informativi può fornire un'informazione parziale, ma sovrapponendo opportunamente uno strato informativo ad un altro si può determinare una conoscenza completa e una visione d'insieme del problema in analisi. Per esempio se si vuole determinare il potenziale rischio di inondazione in un'area residenziale occorre la sovrapposizione dello strato informativo dell'uso del suolo con quello delle aree inclini ad inondazioni. Gli enti pubblici, specialmente quelli che rivolgono la loro attenzione alla salvaguardia dell'ambiente, hanno la necessità di unire diversi strati informativi, ad esempio relativi: alla vegetazione, all'uso del suolo, alla geologia, all'idrologia, ecc. Le sovrapposizioni (overlay) possono essere suddivise in tre categorie principali: punti su poligoni, linee su poligoni e poligoni su poligoni.

Buffering

Il buffering consente la definizione di un *area di rispetto* che si estende intorno ad un elemento appartenente ad una certa primitiva geometrica (punto, linea, poligono). Per area di rispetto si intende, fissato un margine m , l'insieme di punti che distano al più m dal bordo dell'elemento considerato. La caratteristica peculiare di questa operazione è che trova riscontro nella pratica quotidiana, infatti, sul territorio sono presenti particolari elementi geografici, per i quali viene definita (per le loro caratteristiche o per normativa di legge) un'area di rispetto. Ad esempio:

- le linee elettriche ad alta tensione prevedono delle fasce di rispetto lunghe oltre 250 metri per parte;
- a norma di legge è prevista la non edificabilità entro una specifica distanza dalle rive di corsi d'acqua, laghi e mari;
- un antenna radio copre un'area con un determinato raggio.

Il margine di buffering può anche essere parametrico, ad esempio si potrebbero creare delle fasce di rispetto intorno alle linee elettriche usando come parametro discriminante la tensione di esercizio.

Modellazione spaziale

Questa applicazione consiste nell'impiego di modelli spaziali o altri metodi di analisi numerica allo scopo di determinare dei valori di interesse; ad esempio si realizzano dei modelli atti a rappresentare dei fenomeni fisici, reti di distribuzione, ecc. La modellazione spaziale è di difficile implementazione ma apporta notevoli benefici ad un SIT. Le operazioni derivanti dalla modellazione spaziale sono molto difficili da effettuare e solitamente possono essere realizzate solo disponendo di un sistema SIT. Un modello spaziale è in grado di rappresentare il risultato finale rivelandosi un valido supporto nel processo decisionale e consentendo ai progettisti una valutazione delle eventuali soluzioni alternative.

In un ente pubblico le operazioni più frequenti sono le interrogazioni alla base di dati e la visualizzazione di mappe.

Capitolo 2: Analisi dei Requisiti

I tipi di dati usati nella pubblica amministrazione

Molti dati usati in un ente pubblico possono essere gestiti con l'ausilio di tecnologia SIT. I dati gestibili attraverso un SIT possono essere di due tipologie: spaziali e non-spaziali. I primi sono georeferenziati e possono essere acquisiti ad esempio da carte topografiche e tematiche, da rilievi eseguiti direttamente sul terreno, da foto aeree, da immagini da satellite, ecc. I dati non-spaziali sono generalmente ricavati da tabelle o da altri documenti. Spesso le entità non spaziali sono associate a entità spaziali. La componente spaziale dei dati geografici è rappresentata comunemente da oggetti geometrici (punti, linee e poligoni). I dati non-spaziali che fanno riferimento ad elementi facilmente individuabili sul territorio (ad esempio indirizzo di un edificio) si possono considerare come geografici in senso lato.

Di seguito è riportata una lista dei principali tipi di dato spaziali che vengono utilizzati all'interno di una pubblica amministrazione:

- informazioni catastali;
- informazioni anagrafiche e censuali;
- reti elettriche, idriche, del gas, ecc.;
- reti di trasporto;
- uso del suolo;
- locazione di servizi (scuole, ospedali, ecc.);
- permessi edilizi.

Il SIT deve essere in grado di gestire i dati (spaziali e non) a seconda dei casi.

Documenti necessari per la raccolta dei requisiti

In questa sezione proponiamo alcuni moduli che possono essere utilizzati durante la raccolta dei requisiti. I moduli verranno compilati durante le interviste e costituiscono il punto di partenza per la stesura dei requisiti.

Identifichiamo i seguenti moduli principali:

- descrizione di un'applicazione SIT;
- formato di una mappa;
- formato di una stampa (report);
- diagramma di flusso;
- descrizione dei dati. Le figure seguenti illustrano il contenuto dei moduli che costituiscono la documentazione relativa alla raccolta dei requisiti.

Descrizione Applicazione GIS

Organizzazione Sistema Informativo Territoriale Analisi dei Requisiti

Codice dell'applicazione #ID:
Nome dell'applicazione:
Reparto:
Definita da:

Scopo e descrizione:

Tipo di applicazione:

Visualizzazione
Interrogazione
Interrogazione & Visualizzazione
Analisi spaziale
Modellazione spaziale

Scala:
Chiave di ricerca:
Tempo di risposta:
Frequenza:

Dati utilizzati:

Features (entità):

Attributi:

Preparato da:

Approvato da:

Data:

Formato Mappa

Organizzazione Sistema Informativo Territoriale Analisi dei Requisiti

Codice dell'applicazione #ID:
Nome dell'applicazione:
Reparto:
Definita da:

Esempio di Output:

Video

Stampante



Simboli/Legende

Preparato da:

Approvato da:

Data:

Formato Report

Organizzazione Sistema Informativo Territoriale Analisi dei Requisiti

Codice dell'applicazione #ID:
Nome dell'applicazione:
Reparto:
Definita da:

Rappresentazione report: Video Stampante

TITOLO				
SOTTOTITOLO				
TOTALI:				

Preparato da: Approvato da: Data:

Diagramma di Flusso

Organizzazione Sistema Informativo Territoriale Analisi dei Requisiti

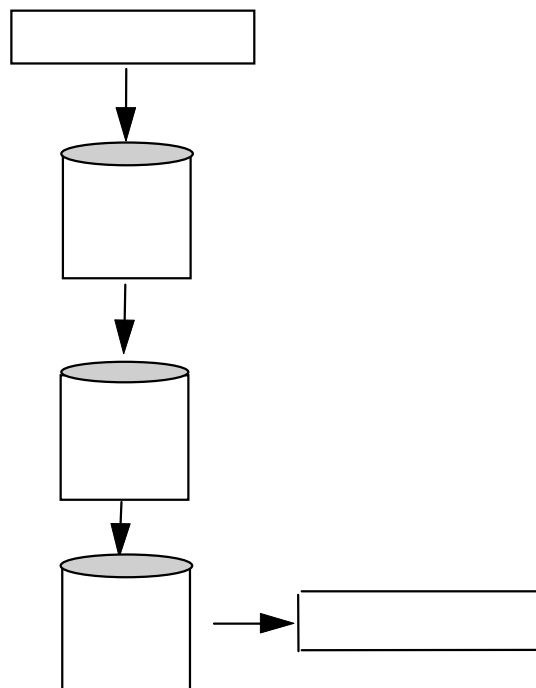
Codice dell'applicazione #ID:

Nome dell'applicazione:

Reparto:

Definita da:

Descrizione flusso di dati: diagramma di flusso



Preparato da:

Approvato da:

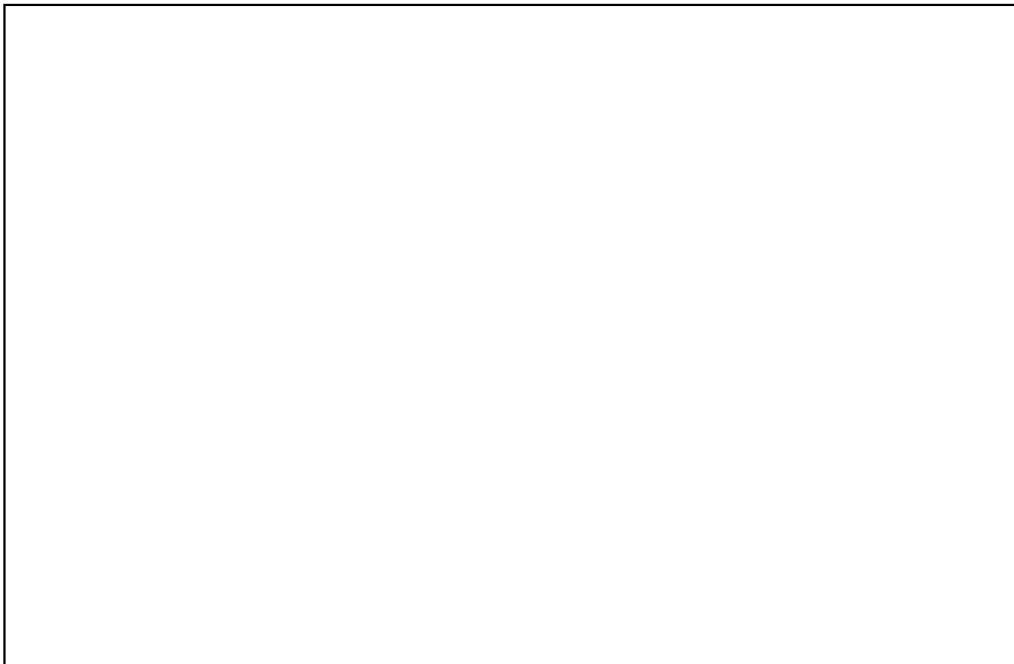
Data:

Descrizione Dati

Organizzazione Sistema Informativo Territoriale Analisi dei Requisiti

Codice dell'applicazione #ID:
Nome dell'applicazione:
Reparto:
Definita da:

Descrizione dei dati:


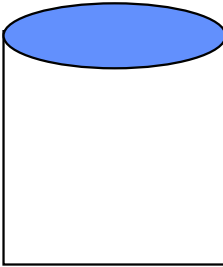




Preparato da: Approvato da: Data:

Documentazione dell'attività di un SIT

Per descrivere le attività SIT (e non) all'interno di un'organizzazione si utilizza generalmente il diagramma di flusso. Infatti, un diagramma di flusso è un metodo standardizzato che consente la rappresentazione di eventi, di attività e decisioni che trasformano input in output all'interno di un sistema.

Capitolo 2: Analisi dei Requisiti

Simbologia del diagramma di flusso	
Un quadrato rappresenta una persona, organizzazione, un concetto, una sorgente o una destinazione di dati o informazioni	
Una forma cilindrica rappresenta un processo o un'attività	
Un rettangolo aperto rappresenta un archivio dati	
Una freccia rappresenta il flusso dei dati. Si può usare un'annotazione sopra di essa a indicare il volume del flusso	

Le interviste

Le interviste individuali costituiscono il modo più efficace per individuare le caratteristiche delle applicazioni SIT che saranno impiegate all'interno dell'organizzazione. Prima di procedere con le interviste è opportuno organizzare un briefing con tutti i potenziali utenti. Nel corso di questa riunione, gli intervistatori devono illustrare a tutti i partecipanti, il modo in cui verrà condotta l'*analisi dei requisiti*. Le attività principali saranno:

- organizzare un seminario preliminare
- intervistare tutti i potenziali utenti
- compilare i moduli (descrizione applicazioni, formato mappe, ecc.)
- riesaminare i documenti ottenuti assieme all'utente
- ottenere l'approvazione degli utenti sulla documentazione.

Capitolo 2: Analisi dei Requisiti

Il seminario preliminare deve precedere le interviste perché esso garantisce il coinvolgimento di tutti i potenziali utenti e consente di preparare il contenuto delle interviste stesse. Infatti, può capitare che in questa fase preliminare della progettazione molti utenti non abbiano le idee chiare su cos'è un SIT e sul come sfruttarne le potenzialità. Inoltre, chi conduce l'intervista può non far parte dell'organizzazione e quindi non avere un'adeguata conoscenza delle necessità e di come è strutturata la stessa.

I contenuti del seminario preventivo sono riportati di seguito:

- definizione:
 - cos'è un SIT?
 - come viene usato all'interno dell'organizzazione? (applicazioni Tipiche)
- procedura per condurre un'intervista:
 - quale sarà il comportamento della persona intervistata?
 - che cosa ci si attende dall'intervistato?
 - chi approva la documentazione inerente la descrizioni di applicazione?
 - come saranno usate le informazioni relative alla descrizione delle applicazioni?

È preferibile intervistare gli utenti individualmente piuttosto che in gruppo. Ciò permette di valutare meglio le idee dei singoli e impedisce gli eventuali condizionamenti da parte di altri utenti. Le riunioni di gruppo tendono ad essere dispersive e quindi si prestano male a fornire le informazioni necessarie a descrivere adeguatamente le applicazioni SIT.

Condurre un'intervista non è un'operazione facile. Alcuni potenziali utenti possono avere una buona padronanza delle applicazioni SIT e possedere un background tale da permettergli di sapere come sfruttarne le potenzialità di questa tecnologia. In realtà, nella maggior parte dei casi i potenziali utenti non hanno una conoscenza completa sui SIT e quindi non sono in grado di identificare facilmente le applicazioni SIT. In questi casi, l'intervistatore (analista SIT) ha il compito di aiutare l'utente nell'analisi dell'attività lavorative che si svolgono all'interno della sua organizzazione, questo ha lo scopo di identificare come integrare le potenzialità di un SIT in questo sistema. L'analista SIT, solitamente prima di iniziare l'intervista, dovrebbe revisionare i criteri di lavoro dell'organizzazione per poi chiedere all'utente una descrizione delle potenziali applicazioni. Dopo questa discussione l'analista SIT dovrebbe essere in grado di definire le potenziali applicazioni SIT.

La specifica dei requisiti

Partendo dalla documentazione prodotta dalla raccolta dei requisiti, si procede con l'analisi e la stesura del documento di specifica dei requisiti.

Tale specifica dipende soprattutto dalla descrizione delle applicazioni e dei dati. La specifica dei requisiti è una fase importante, con una buona specifica dei requisiti si possono evitare molti problemi delle successive fasi di progettazione e di realizzazione. Il problema della stesura dei requisiti può essere affrontato con approcci metodologici diversi:

- **bottom up:** guidato dalle applicazioni
- **top down:** guidato dall'analisi preliminare dei dati
- **approccio misto:** validazione mutua dati / applicazioni

La specifica dei requisiti deve comprendere:

- descrizione del contesto
- descrizione delle applicazioni:
 - nome, tipo, frequenza di ogni applicazione

Capitolo 2: Analisi dei Requisiti

- elenco delle funzionalità SIT utilizzate (overlay, buffer, ecc.)
- descrizione dei dati
 - include la Master Data List
- procedure per la generazione e la gestione dei dati.

Nelle figure seguenti mostriamo uno schema della fase di analisi dei requisiti nelle sue sottofasi di raccolta e di stesura dei requisiti.

Capitolo 2: Analisi dei Requisiti

Intervista con l'utente

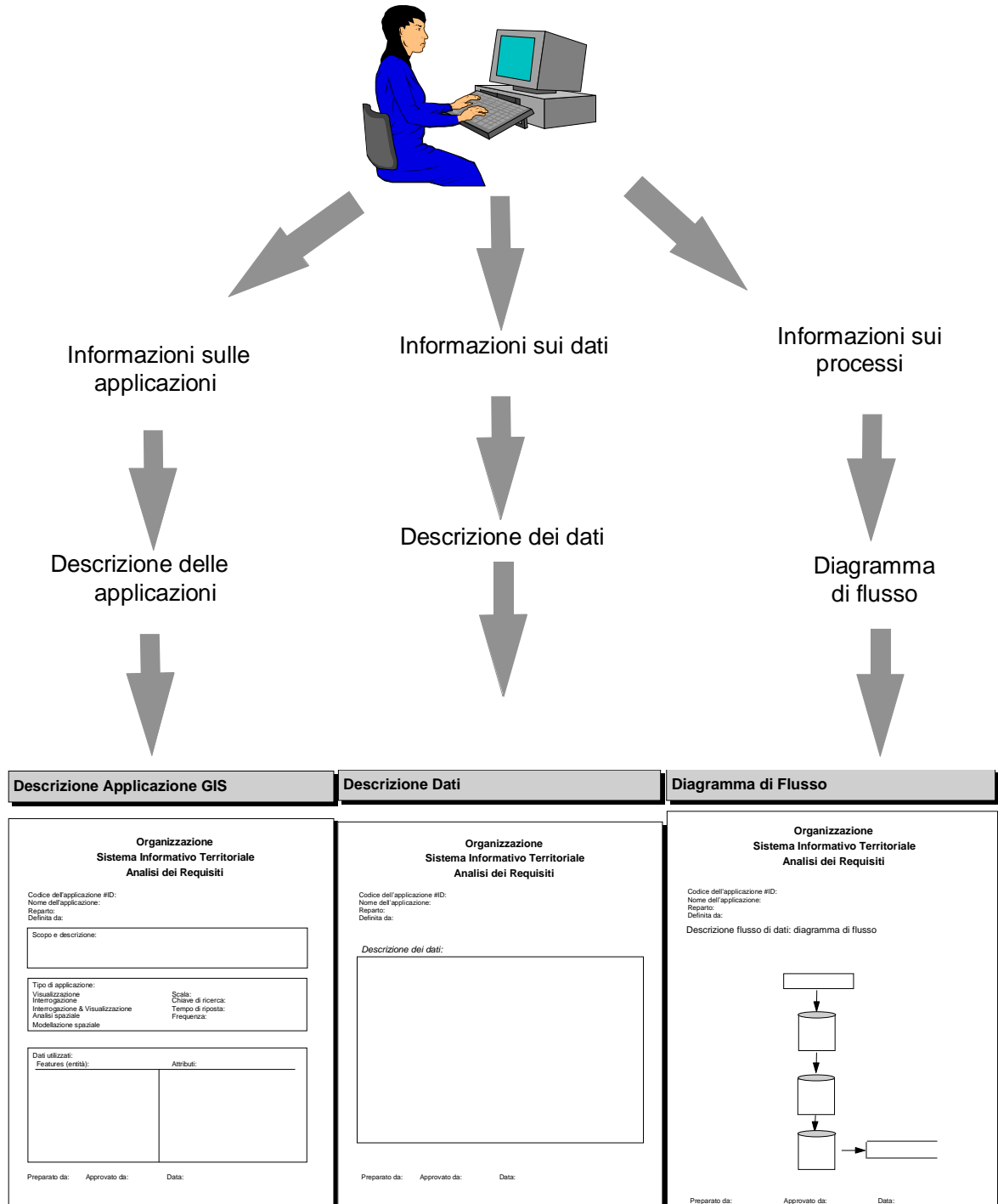


Figura 3 - Schema della raccolta dei requisiti

Capitolo 2: Analisi dei Requisiti

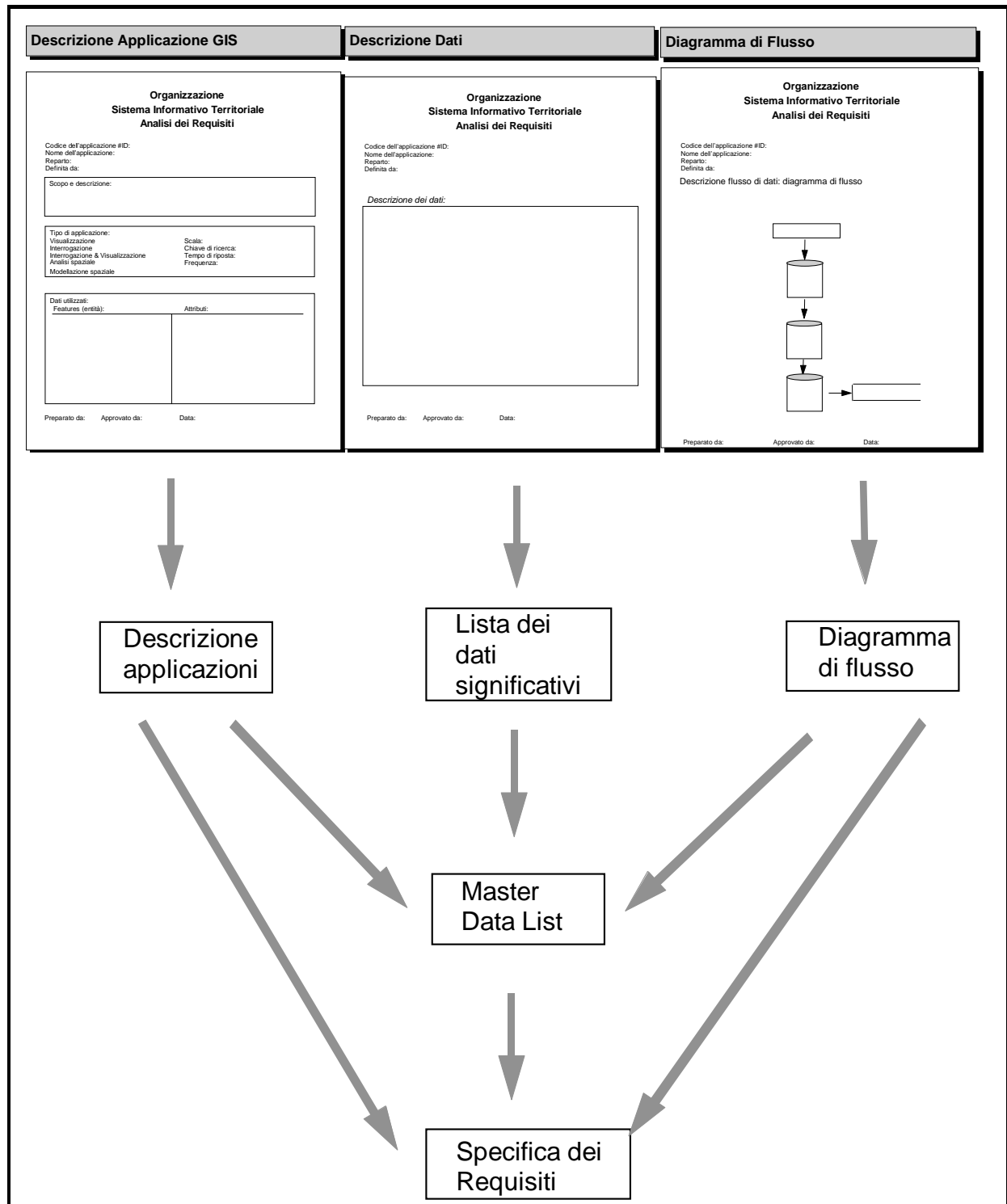


Figura 4 - Schema della stesura dei requisiti

Capitolo 2: Analisi dei Requisiti

Master Data List

La master data list è una descrizione preliminare del modello concettuale dei dati. Viene prodotta durante la stesura dei requisiti basandosi sulla documentazione raccolta (in particolare modo la descrizione dei dati). Per la costruzione della master data list si integrano le diverse descrizioni dei dati e delle applicazioni producendo un'unica lista contenente tutte le entità (features), i loro attributi necessari alla realizzazione dell'applicazione SIT e un'indicazione sul tipo di primitiva geometrica da utilizzare. La master data list sarà utilizzata durante la fase di progettazione concettuale per preparare il modello di dati (geografici e non). Di seguito è riportato un esempio generico di master data list.

Master data list		
Entità	Attributi	Oggetto spaziale
Segmento strada Intersezione strada Parcella catastale	Nome, indirizzo Nome strada Sezione_lotto#, Nome_proprietario, Indirizzo_proprietario, indirizzo_lotto, estensione, profondità	Linea Punto Poligono
Abitante ...	Nome, Cognome, Indirizzo ...	Nessuno

Capitolo 2: Analisi dei Requisiti

Esempio di documentazione

Esempio reale tratto da [GdG].

ECWA Geographic Information System

Erie County Water Authority Geographic Information System Requirements Analysis

Application Identification #: 19
Application Name: Customer Phone Inquiry
Department: Dispatch
Defined by: T. May

Purpose and Description:

To respond to phone inquiries of: 1) "no water;" or 2) to take requests for service. (Reference Donohue #18)

Type of Application:

Display

Query

Query & Display XX

Map Analysis

Spatial Model

Display/Map Scale: 1" = 100'

Query Key: Address

Response Time: 10 seconds

Frequency: xx/day

Data

Features

Attributes:

Row

Location (boundary), street

Pipe

Location (line),

Parcel

Location (boundary),

Building

Location

Services

Location(parcel), address, name,

Projects

Location (boundary),

Work_orders

Location (parcel_by_address),
date

Prepared by:

Approved by:

Date:

ECWA Geographic Information System

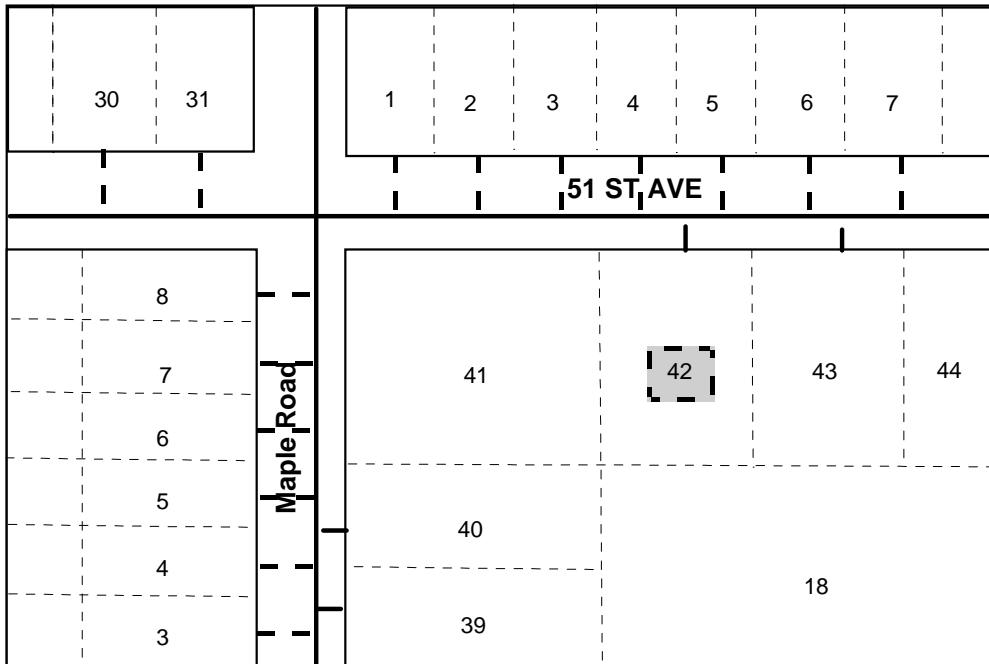
Erie County Water Authority Geographic Information System Requirements Analysis

Application Identification #: 19
 Application Name: Customer Phone Inquiry
 Department: Dispatch
 Defined by: T. May

Graphical Output Sample:

Screen: XX

Hard Copy:



Symbols/Legend

- Parcel
- Water Main
- Building
- ROW
- Service Connection

Prepared by: Approved by: Date:

ECWA Geographic Information System

Erie County Water Authority Geographic Information System Requirements Analysis

Application Identification #: 19
Application Name: Customer Phone Inquiry
Department: Dispatch
Defined by: T. May

Report Layout/Format: *Screen: XX* *Hard Copy:*

Customer Service Inquiry					
Name	Address	Status	Pipe Size	Project	Work Order
J.J. Jones	1551 51st Ave.	Active	4 in.	None	None

Prepared by: Approved by: Date:

Erie County Geographic Information Systems

County of Erie Geographic Information System Requirements Analysis

Application Identification #: 48
 Application Name: Job Training Site Location
 Department: Social Services
 Defined by: Jeff Embury (C/O Jim Kubacki)

Purpose and Description: To provide trainees with an adequate training site while minimizing the distance they must travel to reach that site.

Type of Application:	
Display	Display/Map Scale: Multiple
Query	Query Key: Trainee Address
Query & Display X	Response Time: < 1 Minute
Map Analysis	Frequency: Daily
Spatial Model	

Data Required:	
Features (entities):	Attributes:
ROADS	XY_Location, Name, Address_Range
TRAINEE	Trainee_Address, Trainee_Name
TRAINING SITE	Site_Name, Site_Address, Site_Phone #
SUBWAY	XY_Location, Subway_Stop_Name, Subway_Stop_Location
BUS ROUTES	XY_Location, Busroute_Number, Bus_Stop_Name, Bus_Stop_Location

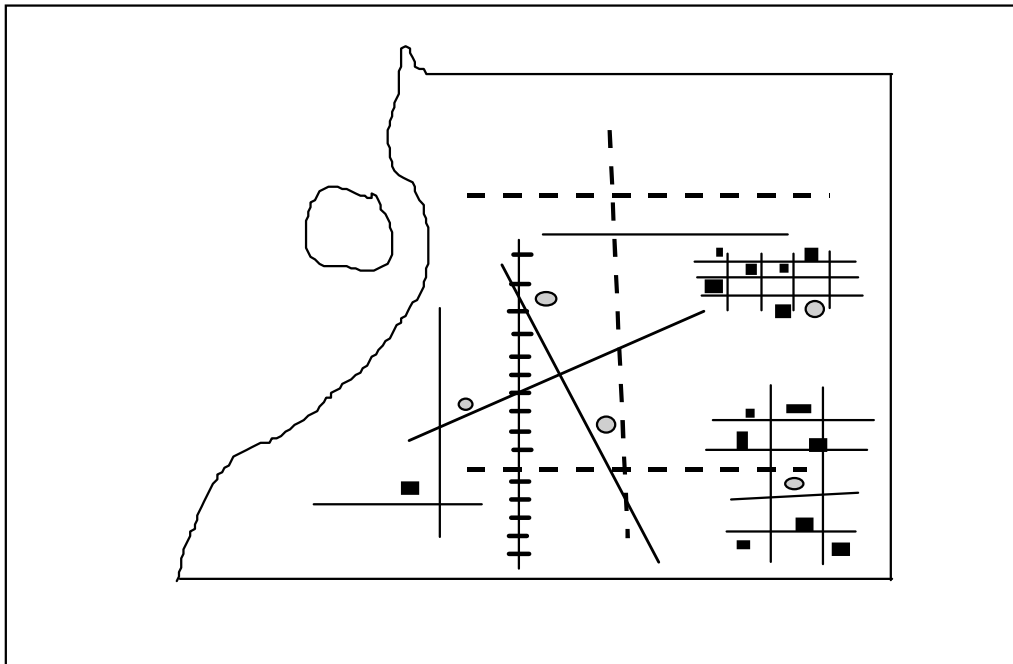
Prepared by: Approved by: Date:

Erie County Geographic Information System

County of Erie Geographic Information System Requirements Analysis

Application Identification #: 48
Application Name: Job Training Site Location
Department: Social Services
Defined by: Jeff Embury/Jim Kubacki

Graphical Output Sample: Screen: X Hard Copy:



Symbols/Legend

■ Trainee ○ Job Training Site — Road - - Bus ||| Subway

Prepared by: Eric Covino

Approved by:

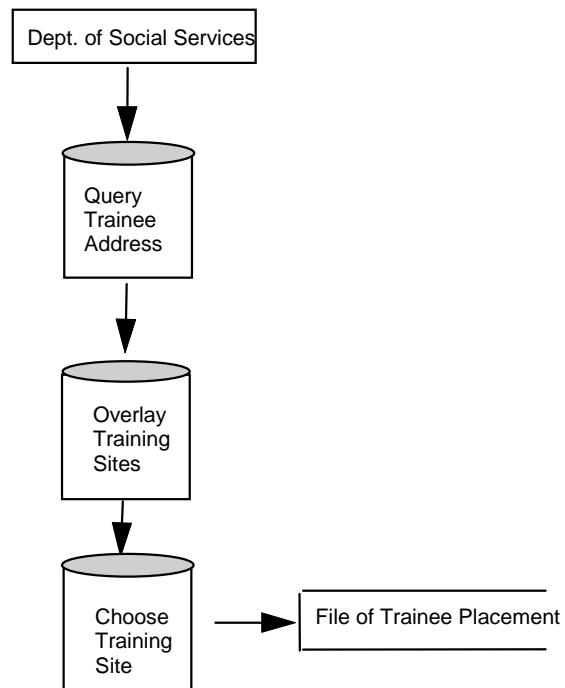
Date: 5/5/94

Data Flow Diagram

Name of Government Geographic Information System Requirements Analysis

Application Identification #: 48
Application Name: Job Training Site Location
Department: Social Services
Defined by: Jeff Embury (C/O JimKubacki)

Process Description: Data Flow Diagram or Flow Chart



Prepared by:

Approved by:

Date:

Capitolo 3: Rassegna dati e valutazione strumenti

Introduzione

In questo capitolo si descriveranno le fasi di rassegna delle sorgenti dati disponibili e di valutazione degli strumenti hardware e software da utilizzare nella progettazione di un SIT, evidenziandone le caratteristiche rilevanti.

Rassegna dei dati disponibili

Utilizzare al meglio i dati adeguati è uno degli elementi di vitale importanza nello sviluppo del progetto SIT. La qualità dei risultati prodotti da analisi geografiche effettuate con l'ausilio di strumenti SIT e da applicazioni SIT dipende strettamente dalla qualità dei dati usati. I dati possono essere ottenuti in vari formati da diverse fonti e il costo di acquisizione può variare in maniera sostanziale tra le varie fonti (es. digitali vs cartacee). Nonostante i dati debbano soddisfare una serie di requisiti specificati durante la fase di analisi, normalmente sono disponibili più sorgenti i cui dati soddisfano tali requisiti. La scelta tra queste è condizionata prevalentemente dal costo o da alcuni altri parametri quali: qualità, accuratezza, scala, diffusione, etc. Durante la fase di rassegna si raccolgono le informazioni sulle sorgenti dati disponibili, tali informazioni saranno utilizzate per determinare quali dati acquisire. Il prodotto finale della rassegna delle sorgenti è un documento che identifica i dati rilevanti e fornisce alcune informazioni sulle loro caratteristiche (metadati).

Possibili sorgenti di dati : formati

I dati interessanti per la rassegna possono essere di diversi formati (per formato intendiamo come il dato si presenta durante la fase di rassegna) che possiamo classificare come segue.

- **Formato “cartaceo/fisico”.** Prevalentemente comprende mappe cartacee che possono essere ottenute facilmente da varie organizzazioni pubbliche e private. In generale rientrano in questa categoria di formati anche le copie di mappe ottenute mediante microfilm o altro. Per essere utilizzati all'interno di un SIT i dati di questo formato devono subire un processo (abbastanza costoso) di acquisizione (che comprende una fase di digitalizzazione). L'acquisizione è soggetta a errori quali la distorsione del supporto cartaceo, errori di interpretazione umana, ecc. Bisogna tener presente che in alcuni casi è indispensabile utilizzare dei dati in questo formato soprattutto se si vogliono delle informazioni relative a periodi storici precedenti all'utilizzo ricorrente di strumenti informatici.
- **Fotografia (analogica).** Foto ottenute mediante macchine fotografiche analogiche, spesso provenienti da fotogrammetrie aeree. Queste raffigurano una porzione del territorio fornendo i dati “grezzi” necessari per svariate applicazioni planimetriche e topografiche. Tramite la rielaborazione di fotografie aeree si ottengono mappe ortofotografiche che sono molto utili per le applicazioni SIT. Un'altra tipologia di immagine che può essere utilizzata nei SIT, sono le fotografie descrittive (es. in un SIT turistico può essere utilizzata una foto raffigurante la Torre di Pisa come un hot-link di attrazione turistica della zona pisana). Le immagini analogiche possono essere facilmente scannerizzate ottenendo un formato digitale raster ad una determinata risoluzione.
- **Formato Digitale.** Il generico formato digitale comprende diverse varietà di dati. Questa varietà sta aumentando di pari passo con l'avanzamento tecnologico che produce sistematicamente nuovi formati. I tre principali formati digitali velocemente

integrabili con un SIT sono il formato *raster* (immagini scannerizzate, immagini satellitari, ortofotografia digitale, ecc.), il formato *vettoriale* (cartografia numerica, dati spaziali topologici, ecc.) ed il formato *tabelle digitali* (provenienti da basi di dati, fogli elettronici, file di testo, ecc.). Un formato che sta emergendo e per cui si stanno sviluppando software per l'integrazione con sistemi SIT è il TIN (Triangulated Irregular Network) per la rappresentazione di superfici tridimensionali utilizzato normalmente per interpolare le curve di livello dati i punti quotati.

Possibili sorgenti di dati : tipi di dato

In questa sezione classifichiamo i dati a seconda dell'uso che può esserne fatto all'interno di un SIT. Ci sono diversi tipi di dato ognuno dei quali ha caratteristiche e potenzialità proprie che contribuiscono in modi diversi ad arricchire la qualità e la completezza del geodatabase. Questi tipi di dato si possono classificare nel modo seguente.

- **Mappa.** È il tipo di dato caratteristico di un SIT. Essa rappresenta elementi geografici (o di altri fenomeni spaziali) identificandone la posizione, la forma e alcuni attributi. La cartografia può essere sia in formato cartaceo che digitale (numerico o topologico).
- **Tabella.** Forma leggibile di dati disponibile sotto forma di tabulati (tabelle, elenchi, registri, ecc.), questi dati possono essere associati a dei dati spaziali o rappresentare semplici informazioni generali di entità non spaziali. Anche le tabelle si possono trovare in fase di rassegna sia in formato cartaceo che digitale.
- **Immagini del territorio.** Le immagini costituiscono un'ottima sorgente di informazione per SIT. Principalmente questa tipologia comprende fotografie aeree del territorio e rilevazioni satellitari. Le fotografie aeree sono normalmente dati non digitali (analogici), mentre le rivelazioni satellitari hanno sempre un formato digitale (di tipo raster). Le immagini analogiche per essere utilizzate all'interno del geodatabase devono essere opportunamente digitalizzate, ottenendo formati raster o vettoriali a seconda dell'uso che se ne vuole fare. Per utilizzare un'immagine disponibile in formato digitale può essere necessaria una conversione od una rielaborazione dell'immagine (rettifica, geo-referenziazione, ecc.).
- **Altre immagini scannerizzate.** Le immagini scannerizzate si distinguono dalle immagini del territorio perché non sono adatte (per risoluzione o per contenuto) ad essere utilizzate per produrre cartografia, di solito queste immagini in origine erano dati cartacei (mappe, diagrammi, foto tematiche di luoghi o persone, ecc.). I dati di questa tipologia, pur essendo in formato digitale compatibile con il raster (matrici di pixel) non sono adatti per essere convertiti in dati vettoriali o per essere utilizzati in analisi raster. Essi vengono normalmente utilizzati solo per semplici visualizzazioni, e sono particolarmente adatti per fungere da sfondo ai vari strati informativi.

Descrizione e valutazione dati

I dati presi in rassegna devono essere identificati, descritti e valutati rispetto alle specifiche fornite durante la fase di analisi dei requisiti. Per descrivere i dati si raccolgono informazioni generali sui dati e sulla fonte degli stessi. Queste informazioni devono essere sufficienti per una valutazione ed una scelta della sorgente opportuna e possibilmente per una valutazione della qualità dei dati stessi. Tutte queste informazioni sono denominate metadati. Dovrebbe essere prodotto un documento preliminare per identificare quali sono i metadati interessanti da raccogliere durante la rassegna.

I dati più importanti da valutare sono i dati cartografici e le immagini (satellitari o aeree) perché la loro scelta costituisce normalmente una differenza non indifferente in termini di costo e di qualità.

Capitolo 3:Rassegna dati e valutazione strumenti

Di seguito riportiamo alcuni parametri di valutazione dei dati:

- scala,
- sistema di proiezione e di coordinate,
- disponibilità di punti di controllo geodetici,
- consistenza e completezza nell'area considerata,
- accuratezza e completezza dei simboli,
- qualità di linee e simboli,
- leggibilità (umana ed automatica),
- quantità di lavoro da effettuare per l'acquisizione,
- accuratezza di posizioni e di attributi.

Tutte questi parametri dovrebbero essere valutati e documentati per ogni potenziale sorgente dei dati. Le informazioni prodotte, nel caso la sorgente sia utilizzata realmente nel progetto, diventano dei metadati effettivi del geodatabase.

Informazioni da individuare

Di seguito riportiamo una proposta di identificazione dei metadati da raccogliere relative a documenti cartografici.

- nome documento
- provenienza
- id. documento
- tipo documento
- numero di fogli
- materiale del substrato
- sistema di proiezione
- sistema di coordinare
- data creazione
- accuratezza della mappa
- scala
- disponibilità
- analizzata da
- data di analisi
- area coperta
- formato del file
- data ultimo aggiornamento
- commenti

Una proposta di metadati relativamente alle entità e agli attributi (colonne della tabella) di dati del tipo tabulare è la seguente.

Entità

- id. documento
- nome entità
- tipo spaziale
- stima numero entità per documento
- simbolo
- accuratezza
- analizzata da
- data di analisi
- commenti

Attributo

- id. documento
- nome entità
- nome attributo
- descrizione attributo
- accuratezza
- analizzato da
- data di analisi
- commenti

Capitolo 3:Rassegna dati e valutazione strumenti

Altri aspetti da considerare

Per rispondere a dei quesiti rilevanti in fase di scelta della sorgente da utilizzare, sono necessarie ulteriori informazioni da prendere in considerazione e da documentare opportunamente durante la fase di rassegna. Alcune di queste domande cui bisognerebbe fornire risposta sono:

- il dato verrà mantenuto aggiornato e disponibile nel tempo?
- il dato è adatto per l'applicazione in questione, in particolare per ciò che riguarda la precisione?
- ha un buon rapporto qualità/prezzo?

Rassegna hardware e software

Un'altra scelta importante nella progettazione di un SIT riguarda gli strumenti hardware e software da utilizzare; tale scelta deve essere quindi supportata da un'attività di rassegna di tali strumenti. Durante tale attività si devono individuare le caratteristiche dell'hardware e del software presenti all'interno dell'organizzazione ed i possibili fornitori di strumenti software ed hardware per lo sviluppo del SIT. La documentazione prodotta deve essere di supporto per la decisione di quali strumenti acquistare, deve essere il più possibile imparziale ed eventualmente richiede una valutazione opportuna degli strumenti (benchmark test).

La scelta finale sarà guidata dai parametri di costo, prestazioni, qualità, ecc. Dato che la scelta degli strumenti dipende fortemente da ciò che si vuole ottenere e dal sistema di partenza di cui si disponeva, è conveniente far precedere alla valutazione dei diversi fornitori di strumenti SIT un'analisi di quali sono le funzionalità richieste e quali sono gli strumenti disponibili all'interno dell'organizzazione.

La fase di rassegna strumenti richiede uno studio preliminare delle tecnologie disponibili sul mercato. Fonti di informazione utili in tal senso sono libri, riviste, fiere e gruppi di utenti del settore. Buona norma è il confronto con altri utenti di sistemi SIT a cui chiedere consigli e delucidazioni.

Alcuni consigli generali riguardo la scelta del software sono:

- usare tecnologia consolidata dal punto di vista sia tecnico che di mercato;
- partire dalle funzionalità individuate nell'analisi dei requisiti;
- verificare il supporto a standard, in particolare per lo scambio di dati;
- verificare la performance;
- espandibilità (dimensioni, architettura, connettività): partire dall'assunto che il sistema crescerà;
- considerare le condizioni di vendita.

Per quanto riguarda l'hardware è importante verificare la compatibilità del sistema operativo con i piani di sviluppo delle strutture di calcolo dell'organizzazione. Bisogna sottolineare che per ottenere le migliori prestazioni è importante svolgere la valutazione del software indipendentemente dall'hardware scelto (e viceversa).

Capitolo 4: Modellazione concettuale e logica

Introduzione

Questo capitolo descrive la fase di progettazione concettuale del ciclo di vita di un SIT e tratta il problema della trasformazione dello schema concettuale in uno schema logico compatibile con la tecnologia da adottare. L'attività principale nella progettazione concettuale è la modellazione dei dati, in particolare nei SIT la modellazione dei dati spaziali; per cui in questo capitolo si descriveranno le tecniche ed i modelli per modellare entità spaziali e non.

Al termine del capitolo si presentano alcuni standard di dati spaziali e metadati.

La progettazione concettuale

La progettazione concettuale è la fase successiva all'analisi dei requisiti; essa si occupa di integrare i vari schemi di settore presenti nella specifica dei requisiti. Il prodotto della progettazione concettuale è lo schema (o modello) concettuale del geodatabase corredato da una specifica funzionale del sistema da implementare. Infatti, oltre a modellare i dati ed ad integrare gli schemi, la progettazione concettuale di un SIT comprende l'attività di identificare e specificare funzionalmente l'architettura del sistema informatico e di progettare l'inserimento del sistema informatico all'interno del sistema informativo.

La progettazione concettuale fornisce, inoltre, delle stime sull'uso e sulla dimensione del sistema risultante. Tali informazioni sono essenziali per le fasi successive di pianificazione del database e di costruzione del prototipo.

Capitolo 4: Modellazione concettuale e logica

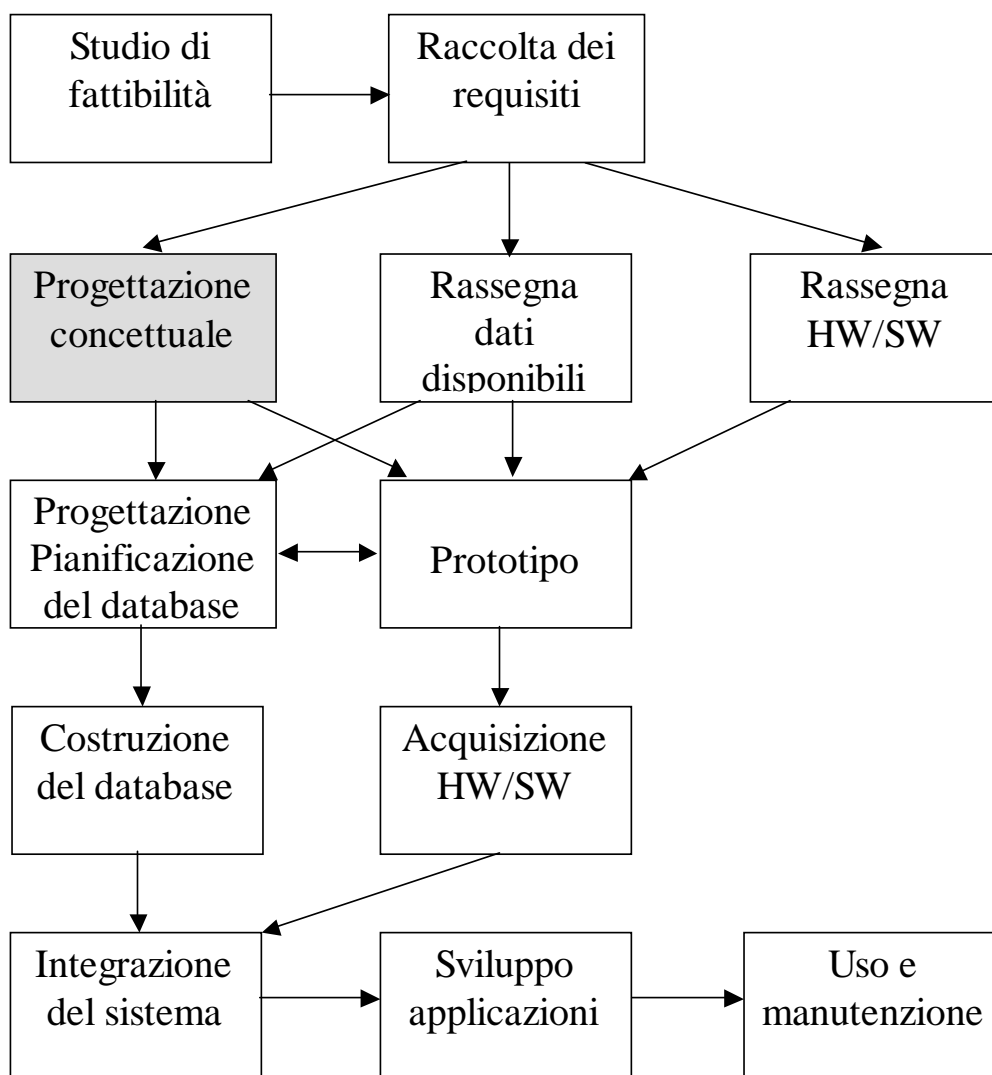


Figura 5 - Ciclo di vita del SIT

Dal punto di vista della progettazione di una base di dati, la modellazione concettuale è il primo passo in cui la base di dati da progettare viene identificata e descritta. La progettazione della base di dati viene normalmente suddivisa in tre attività principali:

- modellazione concettuale: identificazione e descrizione dei dati ad un livello astratto, concettuale; questo passo serve per specificare cosa il SIT deve fare e non si deve preoccupare di come implementarlo;
- modellazione logica: traduzione del modello concettuale nel modello dei dati del sistema da adottare;
- modellazione fisica: rappresentazione del modello logico nello schema fisico del sistema.

La progettazione concettuale si inquadra all'interno del ciclo di vita della base di dati come punto di partenza della pianificazione del database. Infatti, in seguito alla progettazione concettuale si può proseguire con:

- l'analisi dei dati disponibili: i dati disponibili (risultato della rassegna) vengono analizzati e confrontati con i dati necessari nello schema;
- progetto logico-fisico della base di dati;
- creazione dei metadati;

Capitolo 4: Modellazione concettuale e logica

- pianificazione dell'immissione, della validazione e della manutenzione dei dati.

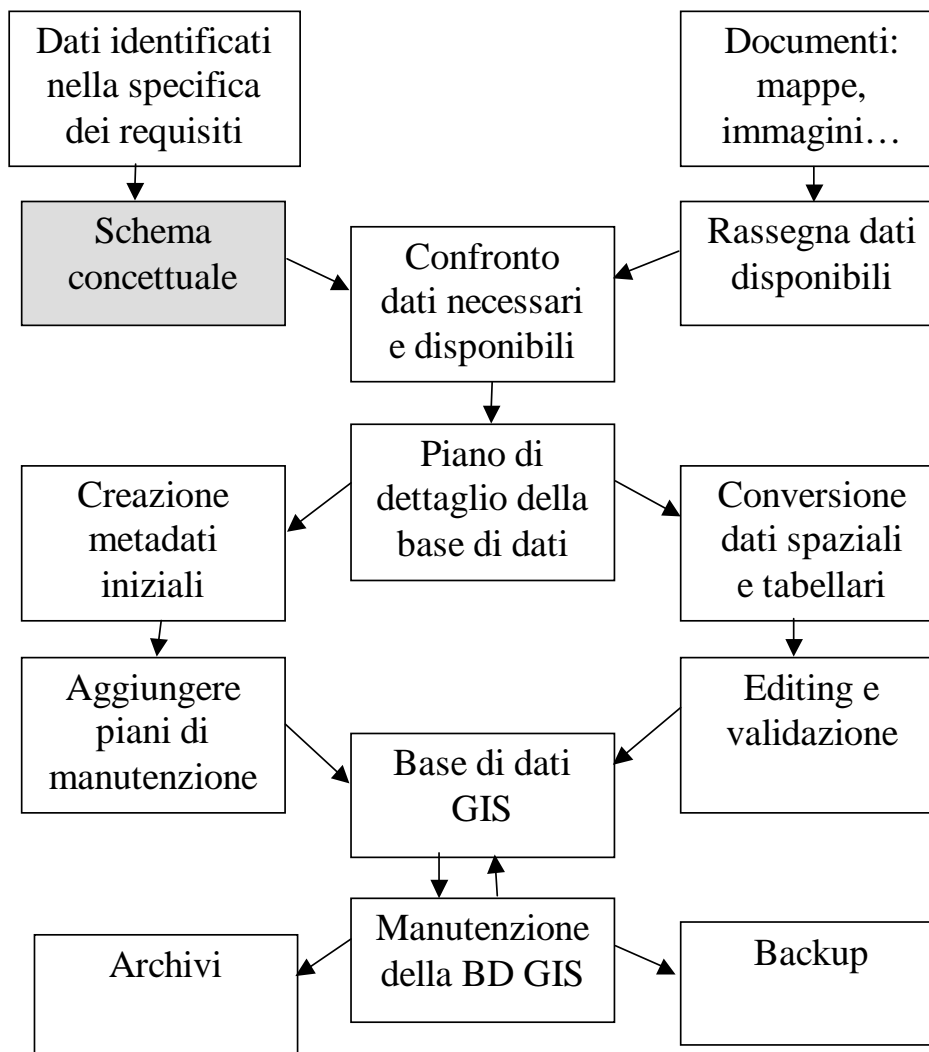


Figura 6 - Ciclo di vita del database

Modellazione concettuale

Un modello astratto o concettuale è la rappresentazione formale d'idee e conoscenze relative ad un fenomeno. Il modello può assumere diverse forme; le due forme più utilizzate sono la lista strutturata e il diagramma entità-relazioni.

Gli scopi del modello concettuale sono essenzialmente due:

- verificare formalmente e graficamente che il committente e l'analista concordano sulla struttura dei dati;
- fornire una specifica formale della struttura alle fasi successive di pianificazione e costruzione del database.

Difatti lo schema concettuale può essere utilizzato, oltre che come documento risultante dalla fase di progettazione concettuale, anche come strumento per la specifica dei requisiti. Nel capitolo relativo all'analisi dei requisiti avevamo introdotto il concetto di Master Data List, come lista strutturata dei dati che devono essere modellati. Tale struttura, specialmente se molto complessa, dovrebbe essere accompagnata, già in fase di specifica, da un diagramma entità-relazioni che specifichi esattamente le relazioni tra le varie entità.

Capitolo 4: Modellazione concettuale e logica

Modello entità-relazioni

Il modello entità-relazioni è il modello generalmente adottato per la modellazione concettuale di basi di dati. La descrizione formale del modello entità-relazioni non rientra negli scopi di questa dispensa, essendo argomento di altri corsi di basi di dati (vedi [AIGhO]). Ricordiamo comunque al lettore che nel modello entità-relazioni vengono graficamente identificate e descritte:

- le entità: cioè un'approssimazione degli oggetti della realtà che ci interessano
- gli attributi: cioè la rappresentazione di informazioni interessanti delle entità
- le relazioni tra le entità: insiemi di istanze di associazioni tra entità, cioè fatti che mettono in relazione più entità; anche le relazioni possono contenere attributi.

Un esempio di diagramma entità-relazioni con uso di gerarchia è il seguente:

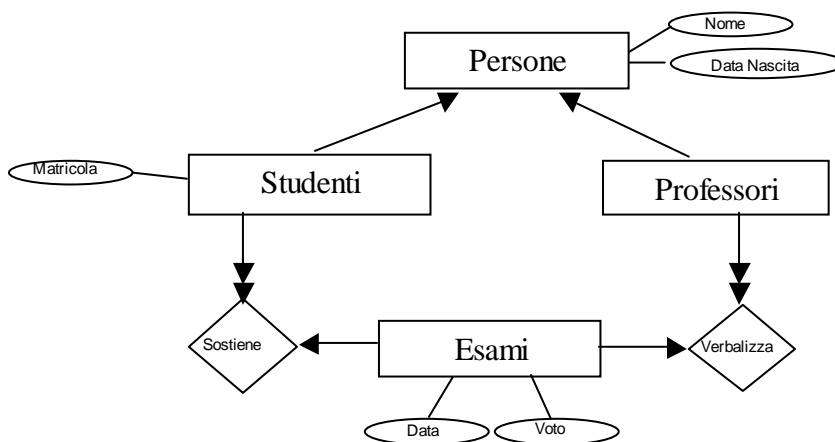


Figura 7 - Esempio diagramma E/R

Specifiche e progettazione

La specifica dei requisiti studia e descrive il sistema informativo ed i bisogni informativi di tutti i settori dell'organizzazione coinvolgendo tutti i livelli gerarchici. Essa quindi produce una definizione del comportamento del sistema dal punto di vista di un determinato settore.

Spesso la specifica dei requisiti, specie se fatta bene, copre parte della progettazione concettuale identificando e descrivendo le caratteristiche dei dati e del sistema. Bisogna comunque precisare che la specifica dei requisiti è strettamente dipendente dal settore ed è comunque necessaria una fase di progettazione concettuale almeno per integrare le varie specifiche.

Quindi è molto importante, ai fini della progettazione della base di dati, sia la specifica formale dei requisiti riguardanti i dati, sia l'integrazione delle specifiche e la progettazione di un unico modello dei dati. Esaminiamo questi due aspetti separatamente mettendo in evidenza il diverso uso del modello E/R.

La specifica. Si definisce il comportamento del sistema dal punto di vista di ciascun settore. In particolare per quanto riguarda i dati bisogna:

- descrivere le informazioni gestite e scambiate fissando la terminologia;
- modellare la conoscenza astratta: cioè identificare e descrivere la struttura della conoscenza concreta e le restrizioni e le regole su di essa;
- modellare la conoscenza procedurale: cioè identificare e descrivere le operazioni sui dati ed il modo in cui esse si svolgono;
- stimare la quantità di dati e la frequenza delle operazioni;

Capitolo 4: Modellazione concettuale e logica

- identificare e descrivere gli aspetti relativi alla sicurezza (ruoli coinvolti, politica di accesso ai dati, ecc.).

Il glossario è uno strumento indispensabile in questa fase per stabilire i termini e la loro semantica nel modello.

Una metodologia generale per la specifica dei requisiti è la seguente:

1. Raccolta dei requisiti: analizza il sistema informativo esistente e raccogli una prima versione dei requisiti, espressi in linguaggio naturale
2. Elimina ambiguità, imprecisioni, disuniformità linguistiche
3. Raggruppa le frasi relative a diverse categorie di dati, relative a vincoli, relative ad operazioni
4. Costruisci il glossario
5. Definisci lo schema preliminare di settore:
 - 5.1 Individua le classi
 - 5.2 Descrivi le associazioni
 - 5.3 Individua le sottoclassi
6. Specifica le operazioni degli utenti
7. Verifica la completezza e consistenza della specifica

Tale metodo deve essere applicato separatamente per ogni settore, possibilmente uniformando il glossario per semplificare la fase di integrazione.

Al termine della specifica si disporranno di informazioni sufficienti per descrivere la conoscenza astratta del settore mediante lista strutturata o mediante un diagramma E/R.

La progettazione. Lo scopo della progettazione concettuale è di integrare le descrizioni settoriali ed esprimerle in modo formale, producendo un progetto completo ed indipendente dall'efficienza delle applicazioni e dalla tecnologia adottata. Tale integrazione può avvenire seguendo due diversi approcci:

- approccio per particolarizzazione: prima si definiscono i dati comuni e poi si specializzano aggiungendo i dati particolari;
- approccio per integrazione: prima si definiscono gli schemi parziali e poi si integrano armonizzandoli.

Una possibile metodologia per la progettazione concettuale è la seguente:

1. Completa gli schemi di settore:
 - 1.1. Definisci i tipi oggetto delle classi
 - 1.2. Individua le generalizzazioni
 - 1.3. Tratta le dipendenze funzionali
 - 1.4. Completa la definizione delle associazioni
 - 1.5. Completa la definizione delle classi
2. Integra gli schemi di settore
 - 2.1. Risolvi i conflitti di nome, tipo, vincoli
 - 2.2. Fondi gli schemi
 - 2.3. Analizza le relazioni interschema (generalizzazioni, inclusioni, ...)
3. Ristruttura eventualmente lo schema finale
4. Definisci l'architettura delle operazioni
5. Controlla la completezza delle operazioni e la consistenza con i dati.

Capitolo 4: Modellazione concettuale e logica

La fase di integrazione degli schemi è spesso complessa in quanto i diversi schemi settoriali possono modellare diversamente entità e relazioni del mondo reale. Ovviamente ci sono problemi di conflitti di nome (entità diverse con lo stesso nome o due nomi diversi usati per la stessa entità reale), di tipo (domini diversi per gli stessi attributi) o di vincoli. Oltre a questi si aggiungono problemi di identificazione delle entità da modellare come nell'esempio della figura seguente.

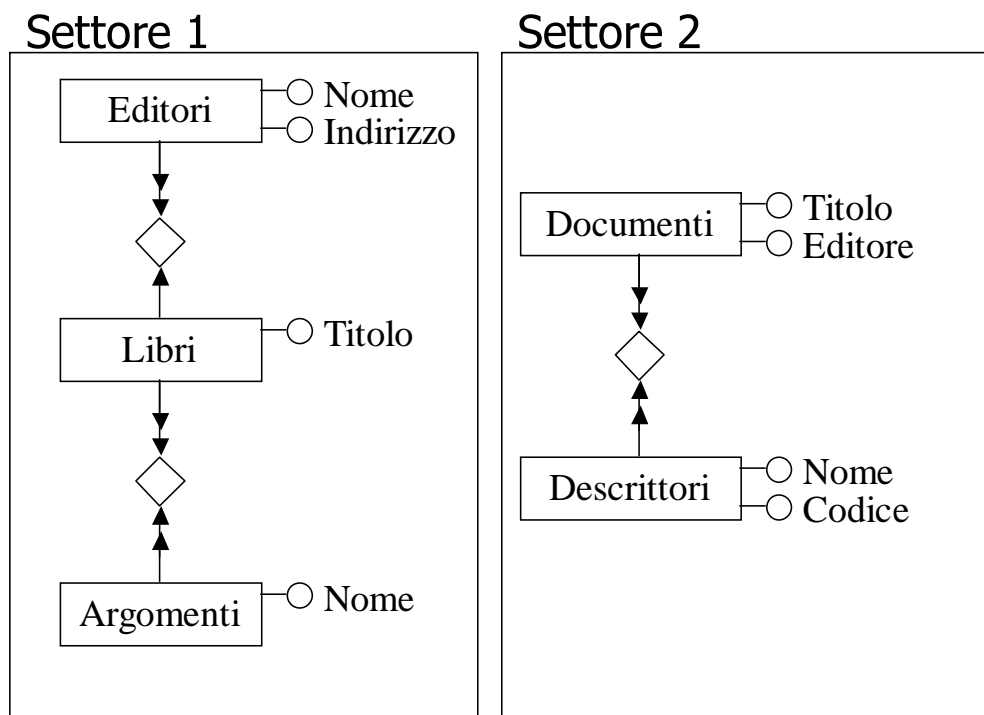


Figura 8 - Integrazione di schemi

Il modello E/R è lo strumento fondamentale della progettazione tramite il quale si presenta graficamente la struttura e le relazioni tra le entità, ma non è l'unico strumento utilizzato. Per studiare e descrivere i processi e l'architettura delle operazioni sono indispensabili diagrammi di flusso dati e diagrammi di stato. Ultimamente sono stati studiati e realizzati degli ambienti CASE per la gestione, sviluppo e manutenzione del progetto software. Tali ambienti sono un utile supporto alla progettazione e forniscono strumenti per l'analisi del sistema informativo, strumenti visuali per la progettazione mediante diagrammi di flusso e di stato, ecc.

Peculiarità dei dati geografici.

Possiamo pensare ai dati geografici come particolari entità georeferenziate di cui interessa la rappresentazione spaziale mediante primitive geometriche: linee, punti o aree. I SIT sono stati progettati per codificare digitalmente la componente (o le componenti) spaziale di questi oggetti associandola ad un certo numero di attributi alfanumerici. Esempi di entità spaziali sono:

- i confini territoriali, la cui componente spaziale è rappresentabile mediante poligoni;
- entità puntuali, associate ad un punto sul territorio, che identificano delle locazioni;
- entità lineari che identificano un percorso.

Un aspetto relativo ai dati geografici che merita di essere approfondito è la codifica topologica. Nella maggior parte di tecnologie per la gestione di SIT alcune relazioni tra entità spaziali vengono codificate topologicamente nel geodatabase. Tale codifica è basata sulla

Capitolo 4: Modellazione concettuale e logica

teoria dei grafi rappresentando le reti mediante insiemi di nodi e archi che collegano tali nodi. Applicato alla cartografia, questo concetto è utilizzato per astrarre dalla forma della figura e per rappresentare facilmente e correttamente concetti come connessione, adiacenza, coincidenza, ecc. I dati all'interno del database possono essere memorizzati in forma topologica (es. segmenti stradali come archi in un grafo i cui nodi sono gli incroci), oppure in forma non topologica, comunemente detta "a spaghetti", in cui le relazioni tra le coordinate di due entità non sono codificate formalmente all'interno di un database. Per esempio in una codifica a spaghetti il sistema non conserva informazione relativa all'adiacenza tra due poligoni, per cui l'unico modo per stabilire se due poligoni sono adiacenti è mediante il calcolo geometrico (con relativi problemi di efficienza e di correttezza dovuti all'imprecisione dei dati spaziali).

Modellare dati spaziali

La modellazione concettuale è essenzialmente un problema standard di basi di dati e tutto ciò che è stato descritto dall'inizio del capitolo finora sono degli aspetti standard di ingegneria del software e di progettazione di basi di dati. Tali metodologie sono applicabili nelle linee generali anche per la progettazione di SIT, ma devono essere estese per la modellazione di entità e relazioni spaziali. Infatti le entità di un database SIT sono per la maggior parte delle entità georeferenziate. Questo implica che nel modello devono essere previste e gestite le informazioni relative alla *forma* e alla *posizione* di un'entità sul territorio. Queste informazioni possono essere viste come degli attributi aggiuntivi dell'entità. Storicamente sono stati utilizzati diversi termini per descrivere i dati di un SIT. Tra questi i più usati sono:

- *oggetti* per riferirsi a entità non spaziali;
- *features* per riferirsi a entità di cui interessano anche la forma e la posizione.

Per essere più precisi, normalmente con il termine *feature* si identificano delle entità rappresentabili mediante il modello vettoriale, per cui oggetti rappresentabili con modelli TIN o raster non rientrano in questa categoria.

Il concetto di collezione di entità viene generalmente associato al termine *classe*. Tale concetto può avere diversi termini se riferito ad entità spaziali: *layer*, *coverage*, *base map*, *tema*, *framework*, ecc. La distinzione tra un termine e l'altro è rilevante a livello di progettazione logica o fisica del database in quanto ognuno di questi termini identifica un modo particolare di organizzare e memorizzare una collezione di *feature* all'interno del sistema. Il concetto di *framework* potrebbe essere comunque rilevante anche a livello di progettazione concettuale identificando insiemi di classi strettamente correlate tra loro (ad esempio un sistema che gestisce la viabilità potrebbe essere organizzato in un *framework* di classi che modellano la rete stradale ed un altro *framework* per modellare l'insieme di costruzioni ed edifici).

I dati geografici differiscono dai dati "regolari" nel modo in cui le entità sono definite e nelle possibili relazioni tra entità spaziali. L'identificazione di entità spaziali include la definizione per tali entità della componente alfanumerica (attributi semplici) e della componente spaziale, in particolare in fase concettuale bisogna opportunamente studiare ed identificare quale (o quali) primitiva geometrica è adatta per la rappresentazione spaziale dell'entità.

La modellazione concettuale deve tener traccia delle informazioni relative alla conoscenza concreta che identificano il tipo di relazione (topologica o non topologica) tra le entità. Per modellare correttamente dati spaziali bisogna conoscere quali sono i modelli dei dati spaziali utilizzati nei SIT attuali.

Capitolo 4: Modellazione concettuale e logica

- **Modello cartografico di dati:** si rappresentano punti, linee e poligoni (consistenti topologicamente) facendoli corrispondere, tramite un identificativo, con delle entità con associati degli attributi. Ad esempio lo strato informativo dell'uso del suolo è rappresentato tramite poligoni (spezzata chiusa di nodi e archi) dove ogni poligono ha un identificativo di tipologia di terreno e possibilmente altri attributi.
- **Modello a oggetti spaziali:** l'utente individua gli oggetti del mondo reale che ritiene di interesse per il proprio progetto. Ad ogni oggetto sono associati degli attributi che esprimono le caratteristiche dell'oggetto in questione. Alcuni degli oggetti presi in esame possono essere privi di riferimenti spaziali (coordinate geografiche). Di conseguenza le relazioni possono avvenire tra oggetti spaziali e non; ad esempio una relazione tra un oggetto spaziale e uno non spaziale può essere la seguente: un edificio (oggetto spaziale) è in relazione con il suo occupante (oggetto non spaziale). Ad un oggetto spaziale può corrispondere un'unica rappresentazione cartografica.
- **Modello a oggetti complessi:** come il precedente ma con rappresentazioni spaziali multiple. A seconda del livello di dettaglio con il quale si analizza un oggetto e le sue relazioni spaziali, queste due componenti possono avere un diverso tipo di rappresentazione. Ad esempio: una strada composta di diversi segmenti (spezzate) può essere rappresentata spazialmente sia da un insieme di linee sia da poligoni a seconda della scala.

Riassumendo, per modellare dati geografici è indispensabile introdurre nel modello concettuale alcuni nuovi concetti per tener traccia di informazioni della conoscenza concreta relative a:

- Componente spaziale delle entità (forma/e e posizione)
- Associazioni spaziali
- Vincoli topologici
- Strati informativi

A tal scopo sono stati studiati diversi modelli o estensioni di modelli pre-esistenti:

- modello di Calkins (estensione del modello E/R per modellare oggetti spaziali semplici e complessi);
- modello MADS (modello a oggetti esteso con attributi, entità, associazioni e gerarchie spaziali e temporali);
- modelli UML estesi (con introduzione di classi e tipi spaziali, es. ESRI UML);
- altri (GeoOM, POLLEN, MODUL-R, CONGOO, ecc.).

Questi modelli differiscono per i diversi gradi di libertà nell'associare informazioni spaziali a entità, attributi e/o relazioni.

Modello di Calkins

Il modello di Calkins sarà adottato nel resto della dispensa per la modellazione concettuale di SIT. Tale modello è un'estensione del modello E/R con entità e relazioni spaziali. Il modello di Calkins supporta una modellazione di oggetti spaziali complessi permettendo diverse rappresentazioni spaziali delle entità. Il modello di Calkins introduce nuovi concetti rispetto al modello E/R, di conseguenza vengono introdotte nuove notazioni e nuovi passi nel processo di modellazione. Esaminiamo il processo di modellazione in Calkins e le differenze notazionali.

Capitolo 4: Modellazione concettuale e logica

Identificazione delle entità

Il primo passo nel processo di modellazione è l'identificazione e la definizione corretta delle entità. Le entità facenti parte di una base di dati geografica sono definite attraverso oggetti discreti (come ad esempio: un edificio, un ponte, un nucleo familiare, un'attività commerciale, ecc.) oppure come oggetti astratti identificabili in termini dello spazio che occupano (per esempio una particella catastale, un edificio, un tipo del terreno, ecc.). In ognuno di questi casi le entità in questione vengono considerate come oggetti astratti a cui associare i rispettivi attributi (ed eventuali relazioni con altre entità).

Identificazione e definizione della componente spaziale delle entità

La componente spaziale di un'entità è esprimibile attraverso primitive geometriche georeferenziate che normalmente sono utilizzate nelle tecnologie SIT, come ad esempio: un punto, una linea, una zona, "un'unità volumetrica", ecc. Un'entità spaziale è quindi caratterizzata dall'insieme dei suoi attributi e dalla sua componente spaziale (eventualmente multipla). La componente spaziale è in sostanza del tutto simile ad uno o più oggetti vettoriali. Componenti spaziali multiple sono necessarie quando un singolo elemento vettoriale non riesce ad esprimere i diversi livelli di dettaglio (la scala) con cui si analizzano gli elementi geografici. Per esempio, una città ad un certo livello di dettaglio si rappresenta attraverso un'entità di tipo puntuale, mentre la stessa città potrebbe avere una rappresentazione poligonale

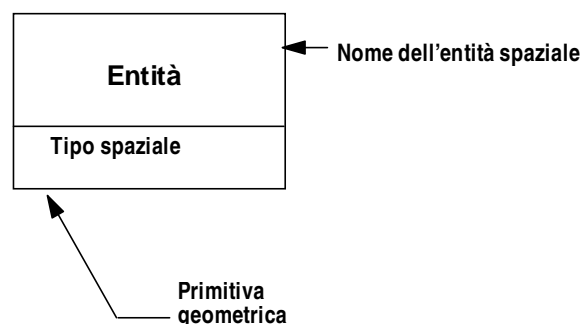


Figura 9 - Notazione di Calkins: entità

Le basi di dati geografiche multi-utente (o i corporate database) possono avere la necessità di disporre d'entità spaziali multiple o di trasformare quelle "normali" in multiple.

Per esempio, se prendiamo in considerazione la viabilità in un centro urbano (le vie sono espresse attraverso un insieme di segmenti "spezzate" che si intersecano) la singola via può essere rappresentata in un SIT in due modi:

- come un elemento lineare correlato alle proprie coordinate geografiche (ad esempio usata per modellare il trasporto nella rete urbana)
- come un elemento poligonale per individuare elementi interni alla via (come ad esempio: il marciapiedi, la linea della mezzeria, ecc.)

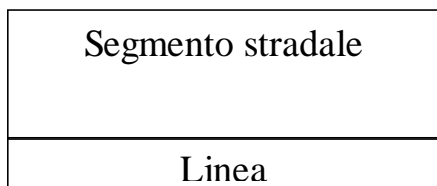
In entrambi i casi l'entità astratta è il segmento che rappresenta la via.

Inoltre, si può avere la necessità di gestire le entità in funzione delle variazioni temporali. Il caso più semplice è costituito dal fatto (quello da noi preso in considerazione) che la rappresentazione dell'entità spaziale non varia in funzione del tempo. Tuttavia, in ottica futura

Capitolo 4: Modellazione concettuale e logica

i sistemi SIT possono avere la necessità di rappresentare entità che variano la propria rappresentazione spaziale in funzione del tempo.

Entità georeferenziata



Entità con rappresentazioni spaziali multiple



Entità con versioni multiple

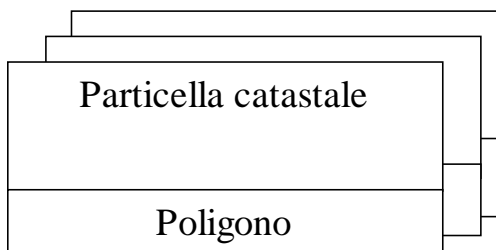


Figura 10 - Notazione di Calkins: entità spaziali multiple

Nel modello di Calkins sono previste tre simbologie per rappresentare l'entità (semplice, rappresentazione multipla, con versioni multiple). Come riportato nella figura seguente la simbologia per la rappresentazione dell'entità spaziale consta di due parti atte a:

- Identificare il nome dell'entità spaziale
- Identificare le componenti spaziali con la relativa primitiva geometrica

La codifica topologica ed il sistema di riferimento non vengono indicati perché sono in genere individuati dal contesto. Attualmente sono state studiate nuove codifiche topologiche che possono richiedere di essere gestite diversamente in un SIT, nel caso di introduzione di diverse tipi di topologia nello stesso modello è indispensabile indicare per ogni componente spaziale qual'è la sua codifica topologica.

Identificazione degli attributi

Per ogni entità identificata, sia essa normale o spaziale, si identificano gli attributi alfanumerici che ne rappresentano le informazioni di interesse. Nel modello di Calkins è previsto che tali attributi non possano assumere componenti spaziali; questa non è una grossa limitazione in quanto un attributo spaziale può essere visto come una nuova entità legata alla classe d'origine con un'associazione di appartenenza.

Identificazione e modellazione delle associazioni

Tra le entità modellate, come nel modello E/R tradizionale, devono essere identificate e definite le associazioni. Nel modello di Calkins, oltre alle associazioni regolari, possono sussistere associazioni spaziali. Tra queste associazioni bisogna distinguere quelle “forti” da quelle “deboli”. In generale ogni associazione spaziale tra entità (ad esempio contenimento di una città in una regione) può essere vista come un insieme di coppie di entità per vale un certo vincolo spaziale (in questo caso il vincolo che controlla se un punto si trova dentro un poligono).

La distinzione tra associazioni spaziali forti e deboli è sottile. Un’associazione debole tra due entità sussiste se e solo se è verificato il relativo vincolo topologico. In un’associazione forte il vincolo topologico è solo una condizione necessaria all’esistenza dell’associazione. In un certo senso è come se le associazioni calcolate siano associazioni implicite perché implicate dalla posizione sul territorio, mentre le associazioni forti devono essere espresse in modo esplicito.

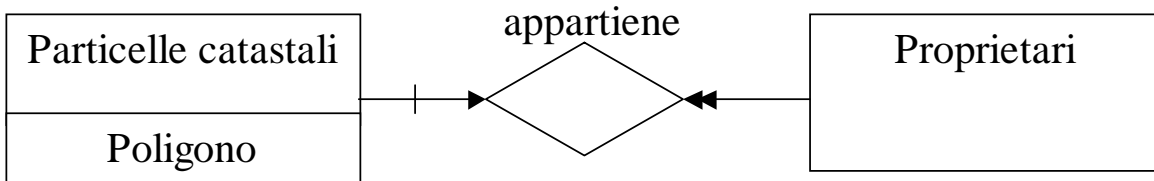
Un’associazione spaziale forte mette in relazione le componenti spaziali di modo che tra di esse sussista un vincolo topologico (es. nodi-archi) od un vincolo amministrativo (es. è deciso amministrativamente che una tale città si trova in un determinato comune, questo implica che effettivamente la componente spaziale di quella città deve trovarsi all’interno della componente spaziale di quel comune). Un’associazione spaziale debole può essere vista, invece, come un calcolo che viene effettuato e che coinvolge le due entità (es. strade prossime entro 10 metri da zone boschive).

La scelta tra le diverse tipologie di associazioni dipende strettamente dall’applicazione. Infatti se il SIT viene sviluppato da un ente amministrativo può essere interessante mantenere una relazione forte tra città e comuni, ma, se l’organizzazione si occupa di analisi del suolo, tale relazione può essere vista tranquillamente come un’associazione calcolata.

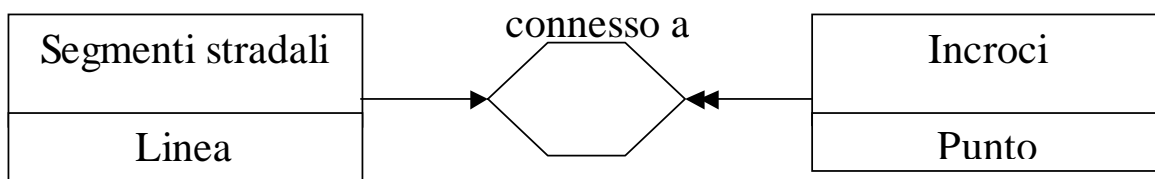
Bisogna comunque tener presente che, in generale, le associazioni spaziali forti (oltre ad essere più vincolanti) sono generalmente più prestanti in termini di tempo d’esecuzione di interrogazioni che le coinvolgono. Per cui la scelta tra associazioni forti o calcolate può anche dipendere dalla frequenza d’uso di tale associazione nel sistema risultante.

Nel modello di Calkins la notazione delle associazioni è la seguente:

Associazioni regolari



Associazioni spaziali



Associazioni calcolate

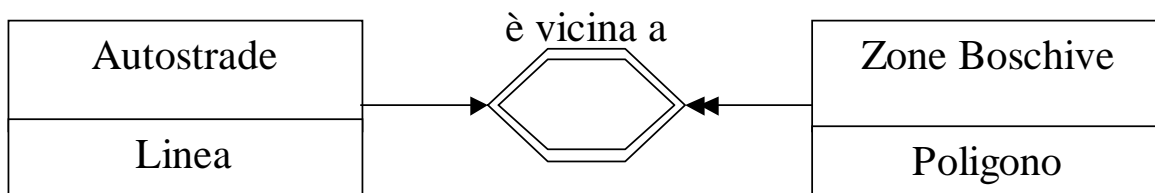


Figura 11 - Notazione di Calkins: associazioni

Le tipologie di associazioni spaziali tra entità possono essere classificate in base al vincolo tra le primitive geometriche che rappresentano la componente spaziale delle entità.

Queste primitive possono essere identificate come primitiva base B e primitiva confrontata C in base al loro ruolo nell'operazione. Ogni primitiva geometrica ha una dimensione nello spazio vettoriale, in particolare: il punto ha dimensione 0, la linea ha dimensione 1 e il poligono ha dimensione 2.

Data la funzione Dim che restituisce la dimensione di una primitiva geometrica, possiamo classificare le associazioni spaziali come segue:

Associazione spaziale tra B e C	Applicabile se
B Coincide con C	$dim(B)=dim(C)$
B Contiene C	$dim(B)\geq dim(C)$
B Attraversa C	$dim(B)=1$ or $dim(C)=0,2$
B È disgiunto da C	Sempre
B interseca C	$dim(B)\neq 0$ and $dim(C)\neq 0$
B È adiacente a C	$dim(B)\neq 0$ and $dim(C)\neq 0$
B È prossimo a C	Sempre

Capitolo 4: Modellazione concettuale e logica

Riportiamo di seguito una tabella riassuntiva relativa alla notazione di Calkins:

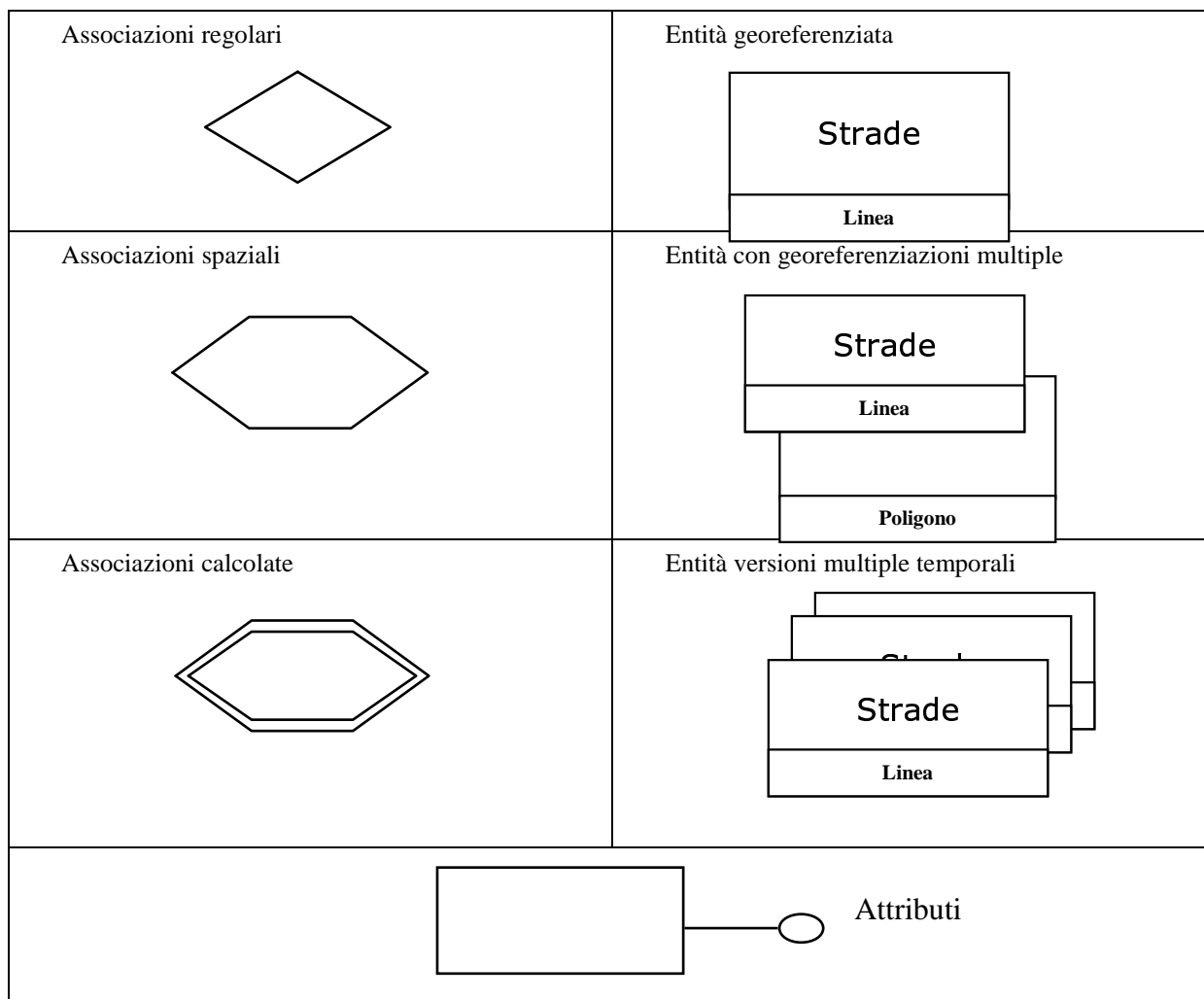


Figura 12 - Notazione di Calkins

Sottoclassi e superclassi

Il modello di Calkins estende la gerarchia del modello E/R anche per entità spaziali, la regola è che le sottoclassi ereditano (se presente) la componente spaziale delle superclassi. Ovviamente sarà possibile definire nuove entità spaziali a partire da entità astratte (senza componente spaziale).

Capitolo 4: Modellazione concettuale e logica

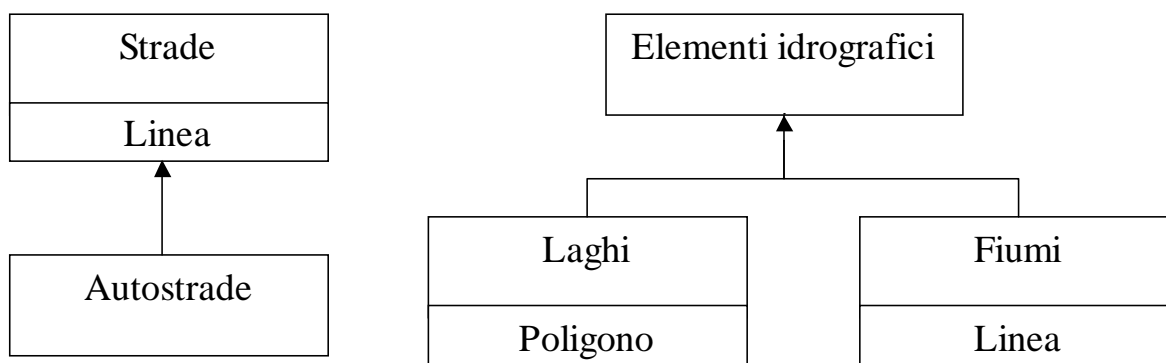


Figura 13 - Sottoclassi in Calkins

Modelli logici

Il modelli logici differiscono dai modelli concettuali perché introducono meccanismi dipendenti dal modello della tecnologia da adottare. Tali modelli possono essere (ad esempio) ad oggetti o relazionali nelle basi di dati standard. Nei geodatabase la scelta è più ampia a seconda dei meccanismi che ci mettono a disposizione le diverse tecnologie. In questa dispensa illustreremo due possibili scelte di modello logico:

- Modello relazionale esteso con Layers
- Modello ESRI UML (ad oggetti esteso con tipi parziali e domini)

Una volta scelto il modello logico da adottare, bisogna trovare una strategia per tradurre lo schema concettuale nello schema logico opportuno per la nostra applicazione. Questa traduzione può essere un processo automatico, ma alcune scelte fra le diverse strade devono essere intraprese dal progettista.

Modello relazionale con layers

Nel modello relazionale con layers viene introdotto il concetto di layer come struttura di memorizzazione topologica di componenti spaziali di una o più entità. Si assume che, se un layer è condiviso da più entità, tra le componenti spaziali di tali entità sussiste un vincolo topologico.

Il modello relazionale con layers ha essenzialmente tre meccanismi di strazione:

- chiavi esterne (per modellare associazioni)
- layer (per memorizzare componenti spaziali di entità)
- condivisione di chiave tra relazioni e layer (per modellare entità spaziali e associazioni topologiche)

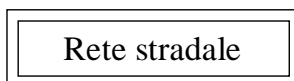
Questi meccanismi sono rappresentati con la seguente notazione:

Capitolo 4:Modellazione concettuale e logica

Chiave esterna



Layer



Condivisione di chiave

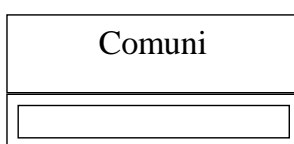


Figura 14 - Notazione nel modello logico

La traduzione tra modello concettuale di Calkins e modello logico relazionale con layers è sostanzialmente la traduzione da modello E/R a modello relazionale con l'aggiunta di alcune regole per tradurre entità e associazioni spaziali. In generale per ogni entità spaziale potrebbe essere introdotto un layer con il meccanismo di condivisione di chiave. Quando sussistono delle associazioni spaziali con vincolo topologico può risultare conveniente (se supportato dal modello logico della tecnologia da adottare) far condividere lo stesso layer tra più entità. L'esempio canonico è quello delle reti che vengono normalmente tradotte con un unico layer che mantiene i dati relativi alle componenti spaziali sia delle entità puntuali (nodi nel grafo) che delle entità lineari (archi). Sempre nel caso di associazione spaziale si può, in alternativa, decidere di mantenere le entità su layer separati. In questo caso il vincolo spaziale è rappresentato da una freccia che collega i due layer. La tipologia di vincolo è normalmente individuata dal contesto.

Di seguito alcuni esempi autoesplicativi di traduzione di entità e associazioni spaziali.

Capitolo 4: Modellazione concettuale e logica

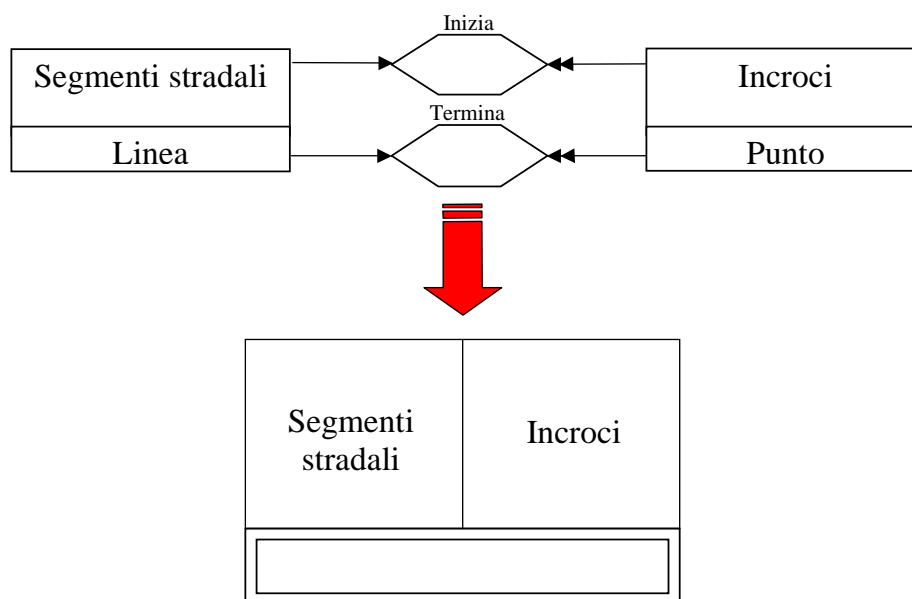


Figura 15 - Più entità su layer unico

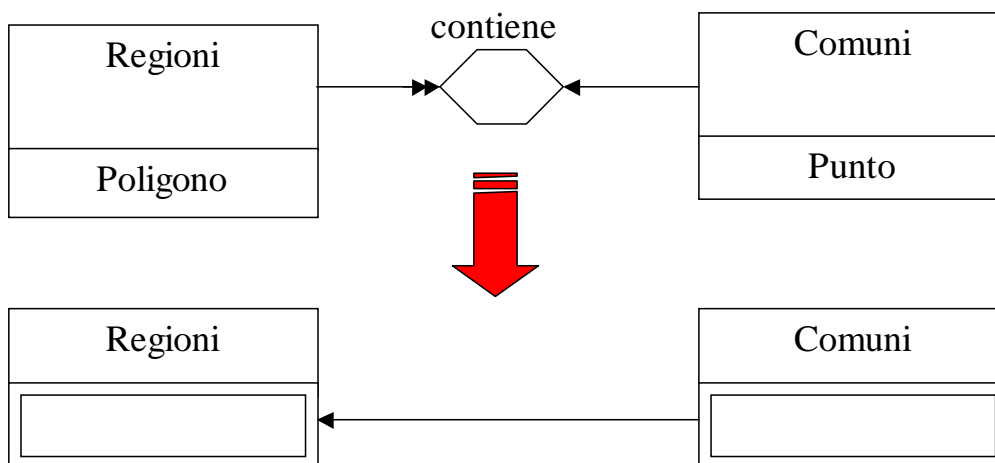


Figura 16 - Entità su layer separati

Modello UML ed estensione ESRI UML

Un'altra possibile scelta è di adottare come modello logico un modello ad oggetti. UML è la notazione standard per diagrammi ad oggetti, ma non è specializzato per rappresentare entità spaziali.

Negli ultimi anni è stata realizzata un'estensione del modello UML che consente di rappresentare entità spaziali e che si mappa completamente sul modello fisico della tecnologia SIT ESRI. Questo modello essenzialmente estende UML con i tipi oggetto (object, feature, network, ecc.) definiti nel modello fisico adottato dalla ESRI. In questo modo i diagrammi ESRI UML possono essere facilmente convertiti in schemi fisici reali degli strumenti GIS quali ArcView, ArcGis, ecc.

La traduzione da Calkins a ESRI UML non è un problema semplice in quanto il modello ESRI introduce nuovi meccanismi quali :

Capitolo 4: Modellazione concettuale e logica

- il concetto di framework (o workspace) per raccogliere entità dello stesso strato informativo
- associazioni tra entità di tipo non simmetrico (aggregazione e composizione)
- oggetti per entità spaziali non vettoriali (raster, TIN, ecc.)
- uso di stereotipi e valori predefiniti per identificare il tipo di implementazione della componente spaziale da adottare
- vincoli sul dominio.

Metadati e Standard di settore

Molti dei settori in cui la tecnologia SIT viene utilizzata, sono stati ampiamente studiati e modellati. Questa esperienza, in aggiunta alla grande necessità di condividere dati nel mondo dei SIT, ha portato alla produzione di meta-modelli standard in questi settori.

Per questo motivo, prima di affrontare il problema di modellazione concettuale o logica del proprio mondo, conviene controllare se esistono meta-modelli standard da cui poter partire. I vantaggi nell'uso di meta-modelli standard sono:

- i meta-modelli sono frutto di studi e di esperienze pluriennali
- è più semplice e meno costoso personalizzare un meta-modello piuttosto che produrne un modello concettuale partendo da zero
- i meta-modelli sono standard ed è quindi più semplice la condivisione dei dati
- normalmente i meta-modelli sono accompagnati da distribuzioni di dati, quindi è più facile acquisire ed acquistare tali dati se si usa il relativo meta-modello.

Gli standard di settore coprono, oltre allo schema concettuale/logico, anche altri aspetti dei dati. Ci riferiamo a questi aspetti con il termine di meta-dati. Alcuni esempi di meta-dati sono:

- natura dei dati
- formato
- sistemi di coordinate
- topologia...Gli standard di settore vengono attualmente distribuiti prevalentemente utilizzando il formato ESRI UML mantenendo informazioni relative sia al meta-schema che ai meta-dati.

Capitolo 5: Pianificazione e costruzione del database

Introduzione

Questo capitolo descrive due fasi cruciali del ciclo di vita del SIT: la pianificazione e la costruzione del database. La costruzione del database comprende le attività di acquisizione, immissione e gestione dei dati che sono le più dispendiose in termini di tempo e di risorse. Per questo motivo è importante la fase di pianificazione che deve studiare il problema, pianificare i processi e allocare le risorse. Bisogna evidenziare che una cattiva pianificazione può causare il fallimento del progetto per esaurimento fondi o per il superamento dei tempi limite. Inoltre è durante la fase di pianificazione e progettazione del database che viene scelto il modello logico e fisico da adottare, altro aspetto importante per le performance finali del sistema. Il capitolo inizialmente inquadra la fase di pianificazione come fase centrale del ciclo di vita del SIT. In seguito si affronta il problema del passaggio al modello logico e fisico del sistema, problema strettamente legato alla scelta dello strumento SIT da adottare. In seguito si esaminano e descrivono le attività da svolgere durante la fase di costruzione del database introducendo delle metodologie per la pianificazione di tali attività.

La fase di pianificazione e progettazione nel ciclo di vita del SIT

La fase di pianificazione e progettazione del database è, in un certo senso, il fulcro del ciclo di vita del SIT. Essa segue alle fasi di progettazione concettuale e di rassegna e, partendo dai risultati di queste fasi (schema concettuale, metadati, specifica funzionale del sistema, rassegna dei dati e delle risorse disponibili), deve decidere la struttura e l'architettura finale del database e si deve occupare di tutti quei "come" che non sono stati adeguatamente specificati nelle fasi precedenti di analisi dei requisiti e di progettazione concettuale

La pianificazione ha quindi come scopi principali:

- decidere e specificare *come* i dati saranno strutturati, sia dal punto di vista logico che fisico;
- specificare *come* le informazioni saranno presentate;
- specificare *come* i dati saranno gestiti.

Alcuni aspetti di problematiche da affrontare in questa fase sono: quale simbologia utilizzare nella visualizzazione delle mappe, quali sono i ruoli e i livelli d'accesso al sistema, come deve essere strutturato l'ambiente di lavoro degli utenti, ecc.

Di particolare importanza sono le specifiche del "come" gestire, acquisire e convertire i dati. Tali processi coprono normalmente buona parte delle risorse e del tempo allocati per la progettazione e costruzione di un SIT. Per questo motivo è fondamentale pianificare e gestire queste attività accuratamente. Oltre ad identificare e a specificare i "come", la fase di pianificazione si occupa anche di definire formalmente le procedure e allocare delle risorse.

Le attività principali da svolgere durante la pianificazione/progettazione del database sono:

- scelta delle sorgenti dati per ogni entità o attributo di interesse (nel diagramma risultato della progettazione concettuale);
- progettazione logica e fisica del (geo)database;
- definizione delle procedure di acquisizione e conversione dei dati;
- definizione di procedure di gestione e manutenzione dei dati.

La pianificazione e progettazione del database è svolta parallelamente allo studio e realizzazione di un prototipo del sistema (e agli eventuali test di verifica). La concorrenza di

queste due fasi è imposta dalla loro stretta correlazione. Infatti, il prototipo si basa sullo schema logico/fisico e sull'architettura del sistema, ma d'altra parte la fase di pianificazione necessita di informazioni sulle performance e sulla qualità del sistema che provengono dai test e dal prototipo. Perciò il modo più naturale per condurre queste due attività è di procedere in modo incrementale, parallelo ed iterativo.

Si comincia con un progetto logico semplice e si realizza un primo prototipo per testare sia il sistema che gli strumenti. Man mano si procede affinando il progetto ed il prototipo. Ad un certo punto si disporranno di informazioni sufficienti per:

- valutare e scegliere lo strumento adatto per il sistema;
- preparare una stima realistica dei tempi e dei costi su cui basare le restanti attività di pianificazione.

Bisogna tener presente che la scelta dello strumento da adottare non dipende unicamente dallo schema concettuale e dall'analisi dei requisiti. In molti casi la scelta dello strumento è vincolata da alcuni aspetti quali l'ambiente in cui il sistema verrà utilizzato e le interazioni con altri sistemi pre-esistenti. Ad esempio se si vuole costruire un corporate database bisognerà tener presente le esigenze ed i sistemi attualmente utilizzati delle organizzazioni coinvolte nel progetto. Sfortunatamente esistono ancora alcuni sistemi legacy utilizzati che sono molto vincolanti in termini di comunicazione e di interazione con altri sistemi. In questi casi è opportuno verificare se conviene mantenere il sistema precedente estendendolo con meccanismi per l'interazione oppure introdurre un sistema ex-novo.

Scelta delle sorgenti dati

Nelle fasi precedenti sono state raccolte le informazioni relative ad un certo numero di sorgenti dati disponibili ed è stato creato uno schema concettuale per i dati da inserire nel database. In quest'attività si fanno corrispondere le entità presenti nello schema concettuale con le informazioni raccolte da ogni sorgente. Lo scopo è determinare per ogni entità (o attributo) la lista di possibili sorgenti dati che soddisfano i requisiti. Se è possibile scegliere tra più fonti per una determinata entità i criteri di scelta generali da applicare sono:

- accuratezza del dato risultante;
- costo dei dati e della conversione;
- disponibilità di dati anche per il futuro mantenimento del database.

In alcuni casi le diverse fonti disponibili per la medesima entità potrebbero portare ad un formato di rappresentazione diverso del dato all'interno del SIT (es. raster vs vector). In questi casi bisogna valutare quanto questi formati soddisfino i requisiti delle applicazioni SIT da realizzare.

Una volta che la sorgente dei dati è stata scelta, si può procedere completando le tabelle dei metadati. In pratica per ogni entità dello schema concettuale si integrano i metadati scelti in fase di progettazione concettuale con i metadati della sorgente corrispondente. In realtà, alcuni metadati necessitano di ulteriori informazioni; ad esempio l'accuratezza aspettata e i dati relativi al controllo di qualità possono essere inseriti solo dopo un'adeguata stima (ottenibile mediante i test sul prototipo). I metadati sono molto utili, oltre che per la pianificazione e la costruzione del database, per la progettazione logica e fisica perché forniscono informazioni sulla struttura, sul dominio e sui vincoli dei dati.

Progetto logico e fisico del geodatabase

Quest'attività è stata descritta da un punto di vista teorico nel capitolo di progettazione concettuale come quella fase in cui lo schema concettuale viene tradotto in uno schema logico.

Capitolo 5:Pianificazione e costruzione del database

Ma prima di decidere come tradurre, vi è il problema di decidere quale modello logico adottare. In realtà alla base di tutto c'è una decisione ancora più drastica relativamente a quale tecnologia utilizzare. Infatti, il modello logico dipende strettamente dal sistema software che si vuole adottare. In particolare nella tecnologia SIT non vi è scelta all'interno del modello fisico (come nei DBMS) ed il modello logico caratterizza completamente tutti i meccanismi che possono essere utilizzati.

Una prima scelta da intraprendere è se utilizzare la tecnologia SIT solo per la componente spaziale delle entità (affiancando alla tecnologia SIT un DBMS relazionale o ad oggetti), oppure utilizzare i meccanismi di memorizzazione dei dati tabellari dei SIT per memorizzare anche le componenti non spaziali dei dati. Bisogna precisare che gli ultimi sistemi per la gestione di SIT supportano sia la memorizzazione interna dei dati che l'interfacciamento a DBMS esistenti. Ciò che accade normalmente è che se i dati alfanumerici sono "molti" o se l'architettura di accesso è complessa (client/server, con molti utenti, ecc.) la tecnologia SIT si poggia su DBMS esistenti, altrimenti la gestione dei dati tabellari può essere facilmente ed efficacemente gestita all'interno del software SIT.

A questo punto rimane da scegliere tra i vari strumenti di gestione SIT disponibili ed, in particolare, tra i vari meccanismi per la modellazione logica di dati spaziali.

Negli esempi fatti finora abbiamo supposto di utilizzare un modello logico relazionale esteso con meccanismi di layer a condivisione di chiave. Questi meccanismi sono forniti dalla maggior parte di sistemi per la gestione di SIT, ma esistono anche altri modelli logici con relativi meccanismi per la modellazione di dati spaziali (es. ESRI UML).

Durante questa fase, una volta scelto il modello logico da adottare, si procede con la traduzione dello schema relazionale nello schema logico e, nel caso si sia adottato anche un DBMS, alla modellazione fisica della componente relazionale (o a oggetti).

Noi supporremo nel seguito di utilizzare un modello che si basa su:

- layer: rappresentazione interna al SIT di punti, linee o aree (o pixel nel caso raster);
- tabelle: rappresentazione interna al SIT o immagazzinata in un DBMS esterno delle informazioni (attributi) delle entità.

Sarà inoltre possibile rappresentare in questo modello associazioni tra layer e tabelle, vincoli spaziali intra-layer e inter-layer. In particolare ogni layer avrà associato la sua (o le sue) componente tabellare. Tutti questi meccanismi sono standard nei sistemi per la gestione di SIT, ciò che varia è la quantità e la tipologia di vincoli topologici gestiti dal sistema e la possibilità di associare diversi tipi di rappresentazione ai dati spaziali.

Pianificazione

Le fasi di costruzione, integrazione e manutenzione del ciclo di vita del SIT devono essere adeguatamente pianificate per evitare sprechi di risorse e di tempo. La pianificazione consiste sostanzialmente nella definizione di procedure che devono essere seguite durante le fasi successive e nella schedulazione delle risorse (assegnazione delle risorse alle procedure nel tempo).

Le procedure da definire nel progetto di un SIT devono coprire le seguenti attività:

- raccolta delle fonti dei dati: può essere alternativamente
 - acquisto dei dati da fonti esterne con valutazione e controllo del materiale per completezza e qualità;
 - raccolta diretta dell'informazione (molto costosa e richiede esperienza nel campo della cartografia)

Capitolo 5: Pianificazione e costruzione del database

- correzione inesattezze e problemi dei dati: correzioni sulle fonti o sui dati digitali per garantire standard di qualità e precisione (map scrubbling);
- acquisizione e conversione dati nel formato richiesto: processo di trasformazione dal formato delle fonti dei dati al formato richiesto; il formato iniziale potrebbe essere anche non digitale, in questo caso si procederà alla digitalizzazione o alla scansione;
- costruzione del database: i dati acquisiti vengono organizzati ed inseriti nello schema, questo processo fa parte della fase di integrazione del database;
- gestione delle modifiche: molti dati (spaziali e non) necessitano un aggiornamento periodico, perciò è utile e importante definire procedure standard per la rilevazione di cambiamenti e l'aggiornamento della base di dati;
- uso e manutenzione: procedure standard di gestione dei dati quali backup, recovery, ecc.

Una volta definite le procedure da pianificare nel SIT, il processo di schedulazione si riduce a problematiche standard di ingegneria del software (che non tratteremo in questa dispensa).

Tra queste attività la più rilevante e costosa (in assenza di raccolta diretta dell'informazione sul territorio) è sicuramente la fase di acquisizione e conversione dei dati.

Pianificare l'acquisizione e la conversione

L'attività di acquisizione e conversione dei dati si occupa della trasformazione del formato dell'informazione dal formato di partenza (sia esso cartaceo, analogico o digitale) al formato, specificato in fase di progetto logico e fisico, che verrà utilizzato all'interno del SIT.

Tecniche e strumenti d'acquisizione e conversione

Per studiare quest'attività e fornire un'idea delle procedure da pianificare, bisogna conoscere quali sono gli strumenti e le tecniche utilizzate per svolgerla.

Le tecniche per l'acquisizione (con associati i relativi strumenti) sono:

- digitalizzazione manuale mediante digitizer: tramite una tavoletta di digitalizzazione si procede con l'acquisizione di carte geografiche punto per punto. Questo processo converte dati dal formato cartaceo/fisico a formato vettoriale. Dato che l'acquisizione è effettuata manualmente c'è un alto rischio di errore ed è richiesto un alto livello di coordinazione o di concentrazione. Più i dati devono essere precisi, più punti devono essere acquisiti e, quindi, vi è un maggior costo di acquisizione. Per ridurre i costi, normalmente, si procede con la selezione e l'acquisizione delle sole entità interessanti; ottenendo così una versione molto semplificata e specializzata della mappa.
- scannerizzazione: trasforma i dati da formato cartaceo/fisico (o analogico) al formato digitale raster. Il processo è svolto da una periferica denominata scanner ed è più veloce e meno costoso rispetto alla digitalizzazione. L'accuratezza del risultato dipende dalla risoluzione utilizzata (in punti per pollice). Lo svantaggio rispetto alla digitalizzazione è che questo processo non è "guidato dalla mano dell'uomo", per cui viene semplicemente riportato in un file una copia ad una certa risoluzione della mappa, senza poter selezionare o identificare le entità di interesse. Questo processo è ideale se il formato desiderato è raster, ma se il formato da utilizzare nel SIT è vettoriale bisogna procedere con un'ulteriore fase di conversione.
- conversione da raster a vettoriale: molti sistemi per la gestione di SIT contengono pacchetti per la conversione dei dati. Vettorializzare significa semplicemente eseguire un programma che trasforma il dato da formato digitale raster a formato vettoriale. Il programma si occuperà di tradurre ciò che era una tabella di pixel in una serie di punti o

Capitolo 5: Pianificazione e costruzione del database

di linee/aree (archi con nodi). Anche questo processo non è “guidato dalla mano dell’uomo”, perciò è spesso necessaria una fase di editing e controllo dopo la vettorializzazione.

- soluzione ibrida: i formati raster e vettoriale hanno i propri vantaggi e svantaggi. Per ottenere la soluzione ottima bisogna scegliere il formato opportuno per il tipo di dato e di uso che se ne vuole fare. Spesso entrambi i formati possono convivere nel medesimo SIT. Ad esempio la sovrapposizione di entità vettoriali ad un’immagine raster georeferenziata del territorio è un ottimo strumento per la visualizzazione.

Un discorso analogo, ma meno complesso, vale per i dati in formato tabellare. Tali dati potrebbero essere disponibili in origine in formato digitale (fogli elettronici, tabelle relazionali, file di testo, ecc.) oppure in formato cartaceo. Nel primo caso la routine di acquisizione si riduce ad una trasformazione del formato digitale (ci sono molti pacchetti software standard per la conversione di dati tabellari).

Nel secondo caso (dati in origine cartacei) sono possibili due scelte:

- immissione manuale;
- riconoscimento automatico di caratteri con successivo editing per eliminare gli errori. Un buon sistema SIT, comunque, contiene pacchetti per il supporto dell’acquisizione e della conversione di dati sia cartografici che tabellari. Per quanto riguarda l’hardware, se si è deciso di acquisire i dati internamente all’organizzazione, necessitano alcuni strumenti specializzati: tavolette digitalizzatrici, scanner ad alta risoluzione, plotter vettoriali o raster. In ogni caso la fase di acquisizione e conversione dei dati, sia essa manuale o semi-automatica, deve essere supportata da procedure per il controllo di qualità. Tali procedure sono necessarie per garantire correttezza, completezza e accuratezza dei dati inseriti.

Pianificare il processo di acquisizione

Il processo di acquisizione, abbiamo sottolineato più volte, è molto costoso. Per questo motivo è importante pianificarlo opportunamente e, se possibile, cercare di alleggerirlo. Il modo più semplice, ma non sempre possibile, di acquisire i dati è l’outsourcing; cioè rivolgersi a ditte specializzate che ci forniscano in un formato digitale opportuno i dati richiesti.

Le procedure da pianificare, relative ai dati spaziali, possono essere schematizzate come segue:

- **specifica di formati e accuratezza:** prima di iniziare la conversione, bisogna specificare accuratamente quali sono i formati ed i requisiti desiderati relativamente al risultato;
- **preparazione delle fonti:** serve quando i dati devono essere scannerizzati o digitalizzati, comprende la fase di scrubbing e di assegnazione dei codici sulla carta. Questi codici saranno utili all’addetto alla digitalizzazione per identificare l’entità immessa;
- **controllo del flusso dei documenti:** senza un monitoraggio del flusso dei documenti si possono introdurre problemi e ritardi per la ricerca delle informazioni, inoltre è opportuno schedare l’uso del documento se questo dovrà essere utilizzato da più utenti;
- **supervisione del processo:** per pianificare un qualsiasi processo bisogna identificare dei ruoli per la supervisione ed il controllo e distribuire le responsabilità;
- **gestione di problemi e inesattezze dei dati:** molti problemi potrebbero presentarsi più volte e per più utenti, per pianificare al meglio la soluzione di tali problemi basta

Capitolo 5: Pianificazione e costruzione del database

introdurre una serie di procedure formali da seguire per verificare e risolvere tali problemi.

- **conversione dei dati:** determinare formalmente quali procedure di conversione utilizzare, per eliminare qualsiasi ambiguità;
- **revisione dei dati digitali:** controllo del formato risultante dalla conversione;
- **procedure di controllo di qualità:** cruciali nei SIT; possono essere anche molto complesse e comprendono sia procedure manuali che automatiche per il controllo dell'accuratezza e della correttezza dei dati. Molti SIT hanno attualmente dei pacchetti software predisposti per il controllo della qualità dei dati, ad esempio il controllo dei vincoli topologici delle entità e dell'appartenenza al giusto layer;
- **correzione finale:** processo di correzione basato sul confronto di una mappa stampata dopo la conversione con la mappa inizialmente digitalizzata;
- **file matching:** combinazione di file (mappe diverse dello stesso strato informativo) con relativi problemi di trasformazione delle coordinate, connessione degli archi e delle aree;
- **valutazione dei criteri di accettazione finali:** tutte le fasi precedenti possono essere pensate come un'iterazione da eseguire finché i criteri di accettazione finali non vengono soddisfatti, questi criteri devono essere specificati preventivamente e sono normalmente dei limiti inferiori a parametri quali: l'accuratezza, la completezza, la consistenza topologica e la correttezza;
- **costruzione del database:** i dati acquisiti e convertiti devono essere inseriti all'interno del geodatabase, questa fase deve essere pianificata specificando struttura e nomi dei file da gestire. Questa specifica sarà utilizzata anche in fase di integrazione e manutenzione dei dati.

Un processo equivalente può essere definito per i dati alfanumerici. Per tali dati dobbiamo acquisire e inserire nel database gli attributi. Le procedure da pianificare per l'acquisizione degli attributi sono:

- **verifica e controllo delle fonti:** una revisione dei documenti esistenti è importante perché può rilevare delle inconsistenze o una cattiva organizzazione dei dati. Inoltre è fondamentale che i documenti in formato cartaceo siano leggibili e propriamente formattati;
- **gestione del flusso dei documenti:** come per i dati spaziali bisogna tener traccia dell'uso e dello spostamento delle fonti;
- **immissione dati:** può essere semplicemente un'importazione (con eventuale conversione) di tabelle esistenti oppure può (in alcuni casi) richiedere l'inserimento manuale o semi-automatico di dati provenienti da tabelle cartacee;
- **gestione dei file prodotti:** specifica dei nomi e delle locazioni (directory) dove memorizzare i file ed eventuale schedulazione del loro uso;
- **controllo di qualità:** per evitare errori di inserimento possono essere utilizzati dei meccanismi di restrizione del dominio (vincoli sui dati inseriti). Tali meccanismi non sono comunque sufficienti e si integrano a procedure standard (manuali o automatiche) di controllo di qualità;
- **gestione delle modifiche:** è quasi inevitabile che i dati subiscano delle modifiche strutturali durante la prima costruzione del database. Queste modifiche e altre dovute a dati mutevoli nel tempo devono essere adeguatamente pianificate;

Capitolo 5: Pianificazione e costruzione del database

- **criteri di accettazione del risultato:** come per i dati spaziali, anche per i dati alfanumerici si definiscono dei criteri di accettazione (correttezza, accuratezza, ecc.) che, se non vengono raggiunti, casuano la ri-esecuzione delle fasi precedenti;
- **integrazione nel database:** processo standard da pianificare per il collegamento e il relazionamento dei diversi file all'interno del database.

L'outsourcing ci permette di risparmiare in termini di risorse uomo e di strumenti da utilizzare. Inoltre disponendo di dati in formato digitale si evitano alcune fasi (quelle relative all'acquisizione), ma rimangono centrali i problemi di verifica della qualità. È buona norma valutare inizialmente le varie ditte che forniscono il servizio di outsourcing e questo può essere fatto con le tecniche di benchmark descritte nel capitolo relativo al prototipo.

Pianificare la manutenzione

La natura dei dati contenuti in un SIT è normalmente soggetta a mutamenti; questo è dovuto al fatto che i dati relativi al territorio riguardano entità del mondo fisico che cambiano costantemente. Se i dati non sono mantenuti aggiornati, l'utilità e la credibilità del sistema degradano rapidamente.

Mantenere i dati aggiornati non è un compito semplice perché, in primo luogo, bisogna identificare quali oggetti (o features) hanno subito un mutamento nel mondo reale. Questo può essere banalmente ottenuto rieseguendo l'acquisizione dei dati nel territorio, ma a costi altissimi. In alternativa se si dispone di informazioni relative ai cambiamenti (per esempio nuove concessioni edilizie per nuovi edifici) si può procedere con l'acquisizione dei soli dati che sono state modificati (o introdotti nel sistema). Ad ogni modo questa attività risulta pesante e necessita di personale specializzato e di procedure ben pianificate per l'identificazione dei cambiamenti e l'aggiornamento dei dati.

Per ridurre i costi si può ricorrere ad outsourcing o a condivisione di dati fra più organizzazioni. Quello che sta accadendo ultimamente è la creazione di cooperative specializzate per un determinato settore dello studio del territorio che condividono i dati geografici. Quest'opportunità è molto interessante, ma anche complessa da un punto di vista organizzativo e gestionale.

Anche le procedure di manutenzione possono (e dovrebbero) essere verificate sul prototipo.

Dato l'alto costo di questa fase, che potrebbe essere rapportabile alla fase di acquisizione iniziale, la pianificazione deve tener conto delle stime e della valutazione dei costi.

Capitolo 6: Il progetto pilota ed il benchmark

Introduzione

Prima di utilizzare una nuova tecnologia come quella SIT, è fondamentale testare il sistema su cui tale tecnologia si troverà ad agire.

Ciò può essere fatto realizzando un *prototipo* (detto anche *progetto pilota*) atto a stabilire se il SIT può essere utile nel normale svolgimento di una specifica attività lavorativa, e in caso affermativo, si può procedere con un test di “benchmark” per determinare il migliore hardware e software da utilizzare per soddisfare le esigenze specifiche del progetto.

Il concetto di prototipo

Progettare un prototipo

Un prototipo offre la possibilità di verificare l'integrabilità del SIT con il sistema esistente e di valutare il ruolo del SIT nelle sue applicazioni quotidiane.

Un prototipo fornisce una comprensione limitata ma sufficiente ad individuare quello che sarà il futuro ruolo del SIT all'interno dell'organizzazione.

Attraverso lo sviluppo del prototipo è possibile valutare le prestazioni e scoprire eventuali problemi,

se tali conoscenze vengono acquisite precedentemente o parallelamente alle fasi di progettazione di dettaglio del sistema e della base di dati, si ottiene un grande vantaggio in termini di progettazione più dettagliata e risparmio di eventuali tempi e costi di ristrutturazione del sistema.

In sintesi, l'uso di un prototipo permette di:

- Verificare l'utilità del SIT per il committente;
- Stimare l'efficienza;
- Stimare i costi e i tempi (conversione dati, formazione personale);
- Stimare il volume dati;
- Scoprire problemi;
- Raffinare la specifica dei requisiti;
- Verificare la qualità delle fonti;
- Verificare le procedure di costruzione e gestione dati.

Per aumentare l'efficienza di un prototipo, esso deve essere progettato e schematizzato in base all'organizzazione dell'attività lavorativa, e a quali sono le funzioni e gli obiettivi da raggiungere.

La realizzazione di un buon prototipo è dovuta anche al supporto e alla partecipazione del personale dell'organizzazione committente. Questa partecipazione fornirà l'occasione per valutare l'utilità e l'efficacia del SIT nell'ambiente di lavoro.

Obiettivo di un prototipo

Un prototipo si pone come obiettivo la verifica dell'utilità del SIT per il committente e la verifica delle procedure definite in fase di pianificazione. Non è una realizzazione completa del SIT né è semplicemente la dimostrazione di un SIT; piuttosto, è uno strumento che ci consente di valutare come un SIT può essere impiegato all'interno di un'organizzazione per migliorare l'efficienza operativa della stessa.

È la piattaforma per verificare l'evoluzione del progetto, le strategie di acquisizione dati e le applicazioni preliminari del sistema.

Un prototipo correttamente progettato ed eseguito dovrebbe permettere di:

- generare un campione della base di dati;
- verificare la qualità dei dati;
- verificare le applicazioni;
- verificare le procedure di gestione e manutenzione dei dati;
- stimare la quantità di dati;
- stimare i costi per l'acquisizione dei dati;
- stimare i costi per la formazione del personale.

Pianificazione del prototipo

Il prototipo deve essere limitato ad un piccolo numero di funzioni SIT e deve far riferimento ad una piccola zona geografica. La progettazione del prototipo dovrebbe essere orientata alle applicazioni, cioè l'obiettivo primario dovrebbe essere quello di verificare le funzionalità del SIT specificate nel documento di analisi dei requisiti. Inoltre lo schema concettuale della base di dati deve essere sviluppato precedentemente.

Lista delle procedure necessarie (fasi) per lo sviluppo del prototipo:

- scelta delle applicazioni a partire dalla valutazione dei requisiti;
- determinazione dell'area geografica d'interesse;
- revisione del progetto concettuale della base di dati;
- determinazione della strategia di acquisizione;
- sviluppare il progetto fisico della base di dati;
- procedure e servizi per lo sviluppo del lavoro di acquisizione;
- preparazione della fonte di dati;
- sviluppo criteri di accettazione e qualità;
- procedure di gestione, manutenzione e elaborazione dei dati;
- preparazione e prova delle applicazioni;
- valutazione della qualità dei risultati;
- preparazione delle stime dei costi.

Scelta delle applicazioni

Bisogna porre attenzione nel selezionare un appropriato insieme di applicazioni per verificare tutte le funzionalità del SIT e la struttura dell'intera base di dati. Una revisione del documento di specifica dei requisiti dovrebbe fornire una selezione delle applicazioni per venire incontro a questi requisiti. Tali applicazioni dovrebbero essere distribuite in modo da coinvolgere tutte le diverse tipologie di utenti (ruoli professionali), inoltre devono essere necessariamente introdotte le applicazioni critiche con potenziali problemi di efficienza.

È importante includere, oltre alle applicazioni per utenti finali del SIT, le applicazioni di amministrazione dei dati. Tali applicazioni (caricamento dei dati, backup, modifica e procedure di controllo di qualità) rappresentano funzioni importanti a cui l'organizzazione farà affidamento per un giornaliero aggiornamento e mantenimento del Geodatabase.

Capitolo 6:Il progetto pilota ed il benchmark

Scelta dei dati

I dati da inserire all'interno del prototipo possono essere sia acquistati da fonti esterne che acquisiti da cartografia, foto, disegni, documenti e basi di dati interni all'organizzazione. In ogni caso, tali dati dovrebbero rappresentare l'intero campo di variabilità dei dati che siano inclusi nel database finale. Se è previsto che documenti e relazioni di sistemi preesistenti saranno inclusi nel SIT, bisognerebbe includere esempi di tali dati. Tutti i potenziali formati e tipi di dato devono essere presi in considerazione. Ciò consente di verificare il processo di integrazione e di controllo dei dati, nonché l'utilità dei dati in ambiente SIT e i metodi differenti di acquisizione, conversione e compressione degli stessi, prima di intraprendere scelte definitive.

Scelta dell'area di interesse

La scelta dell'area geografica di interesse (riferimento spaziale) del prototipo dovrebbe essere fatta in funzione delle seguenti problematiche:

- densità dei dati;
- campione rappresentativo;
- tipologia di acquisizione dei dati.

L'area (o le aree) selezionata deve essere un buon campione, cioè deve rappresentare opportunamente il campo di variabilità della densità e complessità dei dati finali e deve comprendere tutte le entità di interesse. Questo dovrebbe bastare per ottenere un *dataset* rappresentativo e permettere di ricavare il volume ed il costo di acquisizione dei dati relativi all'intera zona d'interesse.

Per misurare le prestazioni dell'hardware, la zona d'interesse deve essere scelta tenendo conto delle mappe e dei file utilizzati normalmente dall'utente finale. L'utente finale deve avere la percezione della struttura finale del SIT (stratificazione di vari tematismi) anche se tali dati sono limitati.

Specifiche preliminari di acquisizione dei dati

Una serie di specifiche, per la conversione dei dati, deve essere definita per ogni strato informativo richiesto dal set di dati utilizzato per le verifiche. La necessità di definire le specifiche di acquisizione dei dati è relativa a:

- Accuratezza
- Affidabilità
- Copertura
- Convenienza
- Completezza
- Formato
- Tempo d'acquisizione
- Leggibilità
- Correttezza
- Priorità d'acquisizione
- Credibilità
- Facilità di manutenzione
- Validità
- Metadati

Scelta dell' hardware e del software

La scelta dell' hardware ed del software per lo sviluppo del prototipo, deve essere fatta in relazione alle specifiche di progetto della base di dati. Tali piattaforme (hardware+software) possono essere acquistate o prese in prestito dai fornitori cercando di minimizzare i costi.

Capitolo 6:Il progetto pilota ed il benchmark

Scelta del fornitore dei servizi per l'acquisizione dei dati

Il prototipo, sebbene non rappresenti uno sviluppo completo del sistema, è sufficiente per testare la bontà dei possibili fornitori dei servizi di acquisizione dei dati, del software e dell'hardware. Non dovrebbe importare il metodo usato dal fornitore per effettuare l'acquisizione dei dati, bensì la loro capacità di soddisfare le specifiche del progetto.

L'unico vincolo desiderabile riguardo ai dati è che il fornitore abbia esperienza nel settore acquisizione dati SIT per il sistema utilizzato e che possibilmente l'acquisizione avvenga direttamente utilizzando strumenti compatibili col formato desiderato, infatti c'è sempre la possibilità di modificare erroneamente gli attributi o di ereditare degli errori, quando si convertono i dei dati da un formato ad un altro.

Valutazione del prototipo

Il prototipo (progetto pilota) parte con uno studio il cui scopo è identificare le esigenze del sistema in via di progettazione. Tale studio si basa sull'analisi dei requisiti e su una serie di domande da porre agli utenti relative alle loro esigenze e richieste. Tramite queste domande si possono valutare alcuni parametri molto utili per la progettazione e la realizzazione del SIT. I principali criteri di valutazione si basano sulla misurazione di quanto il prototipo risponde alle esigenze e alle richieste degli utenti. Tali criteri si possono riassumere nelle seguenti categorie:

Applicazioni

- Integrazione delle applicazioni nel SIT
- Disponibilità nel SIT delle funzionalità richieste
- "Customizzabilità" del SIT
- Supporto tecnico **Database**
- qualità delle fonti
- necessità di preelaborazione delle fonti
- tempi di acquisizione
- problemi durante l'acquisizione
- necessità di dati aggiuntivi
- problemi nel progetto dei dati
- qualità dei dati

Procedure di gestione

- test delle procedure e delle applicazioni
- identificazione delle responsabilità
- verifica delle procedure di CQ
- verifica degli aspetti di sicurezza
- impatto del SIT sulle procedure dell'organizzazione

Infine il prototipo permette la stima di alcuni parametri fondamentali per la pianificazione e la progettazione del database:

- dimensione del database
- quantità di personale
- costo e tempi di acquisizione e conversione dati
- integrazione con i sistemi esistenti
- necessità di nuovo HW/SW

Capitolo 6:Il progetto pilota ed il benchmark

- necessità di addestramento del personale
- necessità di nuovo personale
- costo e tempi di sviluppo procedure

Esecuzione del progetto pilota

Preelaborazione dei dati (Scrubbing)

La documentazione relativa alla provenienza e alle diverse tipologie dei dati deve essere completata prima di iniziare il lavoro di acquisizione. Inoltre bisogna preparare i dati in modo da renderli chiari anche per persone che non hanno familiarità con le pratiche interne dell'organizzazione. Questa fase che precede l'acquisizione è detta di *scrubbing*.

Lo scrubbing serve a identificare ed evidenziare le entità cartografiche interessanti che saranno acquisite in forma digitale. Lo scrubbing è generalmente un processo operato internamente, ma può anche essere effettuato dall'organizzazione che si occupa dell'acquisizione. L'addetto all'acquisizione dei dati cartografici deve essere in grado di interpretare i dati forniti; per questo è importante illustrare i criteri di lettura relativi a tale documentazione, evidenziando le entità che ci interessano.

Allo stesso tempo, le mappe da acquisire possono contenere dei riferimenti ad attributi presenti nel database. In questi casi è importante associare ad ogni entità di interesse una *chiave* (un identificatore unico), di modo da consentirne l'inserimento in fase di acquisizione.

La chiave è cruciale in quanto ci consente di individuare in maniera univoca le entità, che in fase di acquisizione verranno collegate con i relativi attributi (che presumibilmente provengono da fonti tabellari). I valori assegnati alle chiavi dovrebbero essere numeri sequenziali senza un significato specifico, questa scelta è giustificata dal fatto che se il valore assegnato alla chiave fosse significativo risulterebbe più probabile una sua modifica (cosa da evitare vista il ruolo determinante della chiave).

Gestione dei cambiamenti e del flusso di documenti

Se durante l'acquisizione avvengono degli aggiornamenti ai dati od alle mappe, tali modifiche devono essere accuratamente documentate fino al termine dell'acquisizione effettiva. Questo è particolarmente importante se il prototipo verrà utilizzato come base per il SIT definitivo.

Bisogna inoltre tener traccia e pianificare adeguatamente la distribuzione dei documenti (mappe, tabelle, etc.) da acquisire, di modo da aver sempre presente lo stato del sistema (posizione e responsabile dei documenti).

Preparare l'impatto del progetto pilota

Un prototipo influisce sul lavoro quotidiano dell'utente, di conseguenza le persone che partecipano al suo sviluppo devono essere scelte in modo tale da non modificarne il volume di lavoro giornaliero. In caso contrario potrebbero esserci un impatto violento sulla produttività dell'organizzazione che può portare come estrema conseguenza al fallimento del progetto.

D'altra parte il progetto pilota può essere utilizzato come strumento di verifica sulla potenziale integrazione tra personale e tecnologia.

Perché il prototipo produca buoni risultati, è opportuno che le persone selezionate per lo sviluppo di questa fase siano coscienti dell'impatto che questo avrà sull'organizzazione, e che pertanto siano disponibili a fungere da "banco di prova" per lo sviluppo dello stesso.

In questo contesto diventa fondamentale fare dei seminari formativi orientati a informare sul ruolo della tecnologia SIT e sugli obiettivi del prototipo.

Capitolo 6:Il progetto pilota ed il benchmark

Scelta del personale

Per la valutazione del prototipo è necessario selezionare un gruppo ad hoc che funga da campione rilevante delle attività presenti nell'organizzazione, comprese quelle decisionali.

La scelta del personale deve essere condotta scrupolosamente, se possibile utilizzando il personale coinvolto nel processo di analisi dei requisiti del progetto stesso, in quanto in possesso di una conoscenza maggiore sul SIT.

Ottenere il feedback del personale

È fondamentale che tutto il personale che concorre alla valutazione del prototipo fornisca le proprie impressioni attraverso dei questionari o delle liste di controllo dall'inizio fino al completamento del prototipo stesso. È inoltre fondamentale che le domande siano studiate e poste in modo da consentire una facile valutazione delle corrispondenti risposte secondo dei criteri ben definiti.

Le risposte ai quesiti possono essere fornite in vari modi: lista di controllo (tipo si/no), sistemi di valutazione (come: soddisfacente, scarso, inaccettabile), etc.

Bisognerebbe cercare di motivare il personale di modo da coinvolgerlo attivamente, in questo modo il personale coinvolto potrebbe formulare importanti osservazioni e domande spontaneamente.

Valutazione del prototipo

Informazioni da ottenere utilizzando un prototipo

Il successo della fase di progettazione non significa necessariamente che il processo sia stato portato a termine senza intoppi. Infatti, un prototipo sviluppato correttamente può aver rilevato numerosi problemi e anche il fatto che il SIT non sia la tecnologia adeguata per quel determinato utente. Il successo del prototipo deve essere misurato in funzione delle informazioni che è riuscito a fornire. Sotto è riportata la lista delle informazioni analizzate tramite il prototipo.

Specifiche dei dati

Buona parte delle specifiche dei dati sono legate a caratteristiche come: qualità, quantità e costo d'acquisizione.

- **Qualità delle fonti di dati.** Generalmente gli utenti che per la prima volta utilizzano il software SIT rimangono impressionati dalla rappresentazione digitale della cartografia, rischiando di sopravvalutare la qualità e l'utilità dei dati acquisiti. Inoltre, può capitare che alcune mappe da digitalizzare siano così vecchie e sbiadite, da risultare inutilizzabili come sorgente per l'acquisizione dei dati.
- **Controlli di qualità.** Il rischio che può incombere sul progetto (e anche sul prototipo) è rappresentato dal fatto che l'acquisizione dei dati può essere stata condotta senza un adeguato controllo di qualità. Se il cliente non ha esperienza con le tecnologie SIT, può risultargli difficile determinare se tutti i dati sono presenti, se il contenuto informativo è stratificato correttamente o se tutti gli attributi sono consistenti. Poiché il SIT utilizza dati georeferenziati nella visualizzazione delle entità geografiche, attraverso delle opportune interrogazioni si può facilmente individuare la presenza di entità affette da errori. Gli eventuali errori dovrebbero essere segnalati immediatamente all'azienda che ha effettuato l'acquisizione per sollecitarne un intervento risolutivo. Inoltre il cliente può rivolgersi ad un'altra ditta per valutare la qualità dei dati e l'efficienza del software SIT.

Capitolo 6:Il progetto pilota ed il benchmark

- **Disponibilità dei dati.** Prima di aggiungere un attributo ad un'entità è necessario accertarsi se quel valore sia disponibile e sia rilevante ai fini del progetto, questo è infatti un aspetto centrale della fase di acquisizione dei dati. Nel valutare le informazioni da inserire nella base di dati dobbiamo tener conto che queste possono essere di difficile o costosa acquisizione.
- **Acquisizione.** È importante analizzare il monte ore del personale impiegato e i problemi incontrati nella fase di acquisizione (e di preparazione dell'acquisizione). Queste operazioni saranno ripetute durante l'acquisizione completa, perciò sono importanti per fornire una stima dei costi e dell'impatto sull'organizzazione.
- **Quantità di dati.** La quantità di dati e lo spazio disco disponibile sono parametri importanti da valutare nel prototipo. La quantità di dati non ha solo un'incidenza sullo spazio disco ma anche sul costo delle procedure di gestione. Infatti, gestire file di grosse dimensioni oltre ad un'onerosa e complessa acquisizione, comporta anche un impiego maggiore delle risorse del calcolatore (per esempio backup, restore).

Verifica dell'adeguatezza delle specifiche di acquisizione

Le specifiche di acquisizione sono un insieme di indicazioni sui i dati da utilizzare nel progetto e sono relative a: strati informativi, entità, attributi, precisione, livello di accuratezza e formato dei dati. Queste indicazioni sono utili sia per l'organizzazione si occupa dell'acquisizione, che per l'utente che l'ha richiesta.

I costi e le risorse impiegate nel soddisfare le specifiche di acquisizione devono poter essere stimati attraverso l'esecuzione del prototipo, che come detto, opera su un campione di dati significativo.

Prestazioni dell'hardware e del software.

La verifica delle performance vanno effettuate simulando le normali situazioni di lavoro. Ad esempio utilizzando le modalità: stand-alone, multi-utente e multi-task. Queste valutazioni possono essere fatte attraverso delle simulazioni che effettuano un grosso numero di accessi alla base di dati. Questo consente la verifica delle prestazioni del sistema e la stima delle difficoltà che avranno gli utenti nello svolgere il lavoro nel nuovo sistema.

Stima dei costi.

Documentando dettagliatamente sul prototipo i tempi necessari ad acquisire i dati (che dipendono strettamente dalla qualità delle fonti), è possibile stimare i costi per l'acquisizione completa. Il punto cruciale è che il prototipo dovrebbe fornire un'indicazione dei costi del sistema che funziona a pieno carico, tenendo in considerazione gli eventuali sviluppi futuri che le tecnologie SIT possono offrire.

Analisi del feedback

Si analizzano le risposte coinvolgendo tutti i partecipanti allo scopo di individuare il grado di soddisfazione dei requisiti raggiunto attraverso l'esecuzione del prototipo. Può essere necessario intervistare i partecipanti nel caso in cui le le risposte non fossero chiare. In base al valore assegnato alle risposte e al peso dato alle eventuali osservazioni è possibile stimare la soddisfazione dell'utente finale.

Il benchmark

Lo scopo del benchmark consiste nella valutazione e verifica competitiva di prestazioni e funzionalità delle procedure di acquisizione dei dati, dell'hardware e del software, utilizzando due approcci diversi (es: configurazioni hardware e software diverse, impiego di varie fonti dati) sotto condizioni analoghe.

Capitolo 6:Il progetto pilota ed il benchmark

Ogni prodotto software può essere confrontato con gli altri sulla stessa piattaforma (hardware + sistema operativo) oppure un determinato prodotto può essere valutato su piattaforme diverse; se si svolge una serie uniforme di operazioni da eseguire sull'insieme di dati predeterminato, si è in grado di individuare obiettivamente e imparzialmente i vantaggi e gli svantaggi delle diverse configurazioni.

Pianificazione

La fase di benchmark coinvolge tutti i partecipanti al progetto compreso le ditte esterne che forniscono i servizi di acquisizione, nonché l'hardware ed il software; essa può risultare molto onerosa per questo è necessario che sia adeguatamente pianificata.

In particolare è importante concentrarsi sulle questioni chiave in modo da produrre un risultato che sia immediatamente confrontabile, riducendo in questo modo tempi e costi del benchmark.

Obiettivo

Un benchmark consente di confrontare il rendimento di due diverse tecnologie e permette di scegliere, quindi, quale tecnologia utilizzare. Nell'ambito del progetto SIT bisogna valutare: le tecniche di acquisizione dei dati, l'hardware ed il software.

Le caratteristiche che il benchmark deve valutare devono essere definiti con chiarezza e comunicati a tutte le parti in causa. Di seguito è riportata una lista di obiettivi, ovvero le caratteristiche da confrontare:

- **Metodi di Acquisizione**
 - Rapporto qualità/prezzo
 - Qualità
 - Correttezza
- **Hardware**
 - Performance
 - Conformità agli Standard
 - Compatibilità ed Interoperabilità
- **Software**
 - Conformità agli Standard
 - Efficienza
 - Funzionalità SIT (standard ed avanzate)
 - Compatibilità con l'hardware esistente
 - Facilità di uso e di personalizzazione

Questa lista è solo una guida riferimento sui principali obiettivi da tenere in considerazione per la pianificazione di un benchmark, anche se chiaramente non li comprende tutti.

Principi di base

Sulla base degli obiettivi definiti, tutte le parti in causa dovrebbero essere informate sugli elementi da esaminare, sui criteri di valutazione del benchmark (cioè basso costo, rendimento elevato, bontà dei servizi, qualità, accuratezza, etc.). Le regole basilari per la pianificazione sono:

- i test devono essere eseguiti in modo imparziale;
- i test devono essere eseguiti con lo stesso insieme di dati e di operazioni

Capitolo 6:Il progetto pilota ed il benchmark

- pari opportunità per i fornitori
- i test devono essere misurabili quantitativamente
- le verifiche del software dovrebbero essere fatte sulle stesse piattaforme.

È necessario documentare le caratteristiche delle verifiche operate dai test di benchmark compreso il grado di soddisfazione degli utenti qualora si tratti di prove soggettive su cui è possibile un confronto, come ad esempio la facilità d'uso. Le caratteristiche da evidenziare per ogni prova sono ad esempio: le prestazioni dell'elaboratore, la velocità del clock, lo spazio su disco, i messaggi di errore, il tempo di reazione, ecc.

Preparazione delle specifiche

Le specifiche devono descrivere il tipo di prova da condurre (acquisizione dati, hardware e software), gli obiettivi della prova, una descrizione dettagliata della stessa, le misure per conformità e una stima dei tempi.

Scelta del luogo e dei partecipanti

Per sviluppare un benchmark, è necessario personale qualificato (sia interno che esterno). Il personale interno deve conoscere gli elementi da esaminare (acquisizione dati, hardware e software). La selezione dei partecipanti esterni dovrebbe essere fatta tenendo in considerazione fattori quali: il grado di conoscenza delle procedure interne dell'organizzazione dell'utente, esperienza nel settore, etc.

Il benchmark infatti può essere sviluppato presso il cliente o nella ditta del fornitore.

L'acquisizione dei dati può essere realizzata presso il fornitore per minimizzare i costi (senza dover acquistare gli strumenti per questa fase e preparare il personale).

Invece le prove hardware e software possono essere eseguite sia presso il fornitore che il cliente; in genere vengono effettuate prima presso il fornitore, e una volta raggiunti gli obiettivi fissati, l'ambiente di sviluppo viene ricreato e verificato presso il cliente.

Preparazione dei dati

Un benchmark condotto per l'acquisizione si svolge confrontando i risultati ottenuti da diversi fornitori di servizi d'acquisizione, ai quali è stato consegnato un insieme di fonti di dati (adeguatamente preelaborate); è necessario fornire anche un'idea di come i dati dovrebbero apparire nei risultati. Questo è ottenibile facilmente, se si dispone del prototipo, fornendo all'ente di acquisizione un insieme dei dati significativo utilizzato nel prototipo.

Per il benchmark del software o dell'hardware, si deve fornire un set di dati che contenga un campione di tutte le entità del SIT che stiamo sviluppando, un campione abbastanza grande da poter effettuare tutte le operazioni significative. È opportuno fornire una documentazione sufficiente sull'uso dei dati, e sui requisiti della struttura organizzativa.

Pianificazione dei tempi del benchmark

Una volta che il benchmark è stato definito ed approvato dai partecipanti, si deve pianificare la data d'inizio e la durata delle prove di verifica.

Distribuzione delle specifiche dei dati ai partecipanti

Prima di distribuire le specifiche (o eventuali dati come ad esempio la cartografia) ai fornitori, è opportuno farne una copia e archivarle. Chiaramente i dati devono essere in un supporto e in un formato tale che tutti i fornitori siano in grado di utilizzare.

Capitolo 6:Il progetto pilota ed il benchmark

Inoltre, va chiesto al fornitore di fare un inventario dei dati e della documentazione che ha ricevuto.

Logistica

Se le prove vengono effettuate presso il cliente, è necessario il consenso di tutti i partecipanti. Nel caso in cui le prove siano effettuate su hardware provenienti da fornitori diversi, questi devono essere presenti per valutare la prova comparativa.

Se le macchine dei diversi fornitori risiedono nella medesima stanza, si procede parallelamente su tutte le macchine per eseguire in tempo reale le stesse prove, ed avere un riscontro visivo immediato dei risultati.

Cambiamenti delle specifiche

Anche se le prove sono state condotte e definite con attenzione, può essere necessario effettuare delle prove supplementari durante il benchmark. Se i cambiamenti sono di lieve entità non necessitano di altre prove. Nel caso in cui i cambiamenti portino a modificare in parte gli obiettivi questa nuova situazione va comunicata a tutti i partecipanti.

Definire i criteri del benchmark

Specifiche di acquisizione

Per valutare i risultati di benchmark relativi all'acquisizione dei dati, deve essere effettuato un insieme standard di prove. Si devono controllare e analizzare le varie tipologie di dati da acquisire (cartografia, foto, immagini da satellite) per poi confrontare i risultati. Le indicazioni sulle metodologie di acquisizione sono fornite dal prototipo. In genere, attraverso un software SIT, che ci fornisce la visualizzazione dei vari strati informativi, è possibile evidenziare gli errori commessi in fase di acquisizione.

Software SIT

Le prove del software possono essere classificate in due gruppi: idoneità e performance. Le prove di idoneità esaminano se il software può effettuare un'operazione specifica (conversione file, registrare dati e immagini nei vari formati, accesso ai database esterni, etc.). Le prove di performance si occupano di come o di quanto velocemente il software effettua l'operazione che gli è stata assegnata. La velocità può essere misurata con un cronometro.

Performance hardware

L'obiettivo è di trovare l'hardware più veloce rientrando nel costo stimato. Vanno fatte anche le prove standard dell'hardware su altre applicazioni di uso comune quali: l'elaborazione di testi, i fogli elettronici ed i pacchetti grafici. I risultati di tali prove non saranno specifici dei SIT, ma mostreranno le prestazioni generali del calcolatore in questione. Talora due calcolatori con specifiche hardware apparentemente uguali (velocità di clock, RAM e spazio di disco) possono avere comportamenti diversi, in relazione a molti fattori (chipset, configurazioni interne del circuito integrato, etc.); pertanto, se necessario, si effettuano altre prove aggiuntive per chiarire i problemi e le discrepanze incontrate, con una nuova formulazione di domande ed osservazioni sul caso.

Valutazione dei risultati di un benchmark

Se le domande sono formulate chiaramente ed i risultati registrati correttamente, la valutazione dei risultati del benchmark dovrebbe essere un processo relativamente semplice. In caso contrario sarebbe necessario procedere ad una nuova formulazione di domande, di osservazioni sul caso, e ad altre prove aggiuntive atte a chiarire i problemi e le difficoltà incontrate.

Capitolo 7: Fasi finali

Introduzione

In questo capitolo si esaminano brevemente le restanti fasi del ciclo di vita del SIT. La figura seguente fa il punto della situazione evidenziando le fasi del ciclo di sviluppo di un SIT che sono già state affrontate nei capitoli precedenti (rettangoli bianchi) e quelle che devono essere ancora trattate (rettangoli grigi).

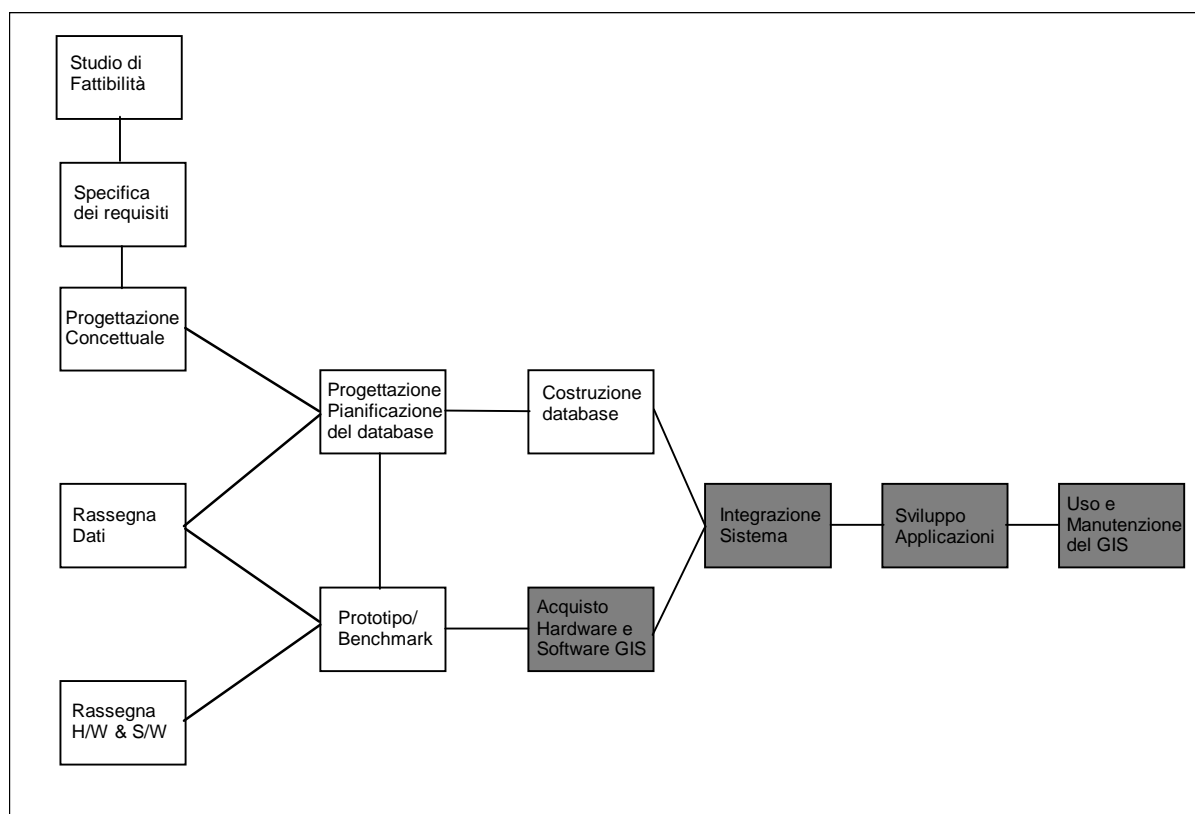


Figura 17- Il ciclo di sviluppo di un SIT: le ultime fasi

La pianificazione, progettazione e la verifica del progetto SIT sono stati ultimati nelle fasi precedenti, in particolare durante la fase di pianificazione sono state già identificate e definite delle procedure da seguire per ultimare correttamente il progetto.

Acquisto dell'hardware e del software

In questa fase si procede con l'acquisto dell'hardware e del software con la consapevolezza che queste componenti rispecchiano le specifiche del progetto. Questa affermazione deriva dal fatto che le fasi precedenti (il prototipo e il test di benchmark) ci consentono di determinare l'hardware e il software adeguato al nostro progetto attraverso una serie di test standard.

Un'aspetto importante relativamente all'acquisto dell'hardware e del software è la preparazione di un eventuale bando. Questa fase è vitale nel caso l'organizzazione coinvolta sia un ente pubblico o finanziato dallo stato. Il bando deve essere correttamente stilato e deve specificare con chiarezza:

- cosa serve;

- cosa e come specificare nella risposta;
- come saranno valutate le risposte.

È preferibile rivolgersi a consulenti specializzati per la compilazione e stesura del bando.

Altri aspetti relativi alla valutazione del prodotto (funzionalità, costi, ecc.) sono stati approfonditi adeguatamente nel capitolo relativo al prototipo.

Integrazione del sistema SIT

Nelle fasi precedenti la fase d'integrazione è stata pianificata (pianificazione e progettazione del database) e testata (prototipo/benchmark). Se le fasi precedenti sono state condotte in maniera oculata l'integrazione definitiva del SIT si dovrebbe limitare a creare il database progettato nell'hardware e software selezionati e all'inserimento e collegamento dei file acquisiti nel sistema.

L'integrazione delle componenti del SIT si svolge su tre livelli:

- livello hardware: le componenti hardware acquistate devono essere installate e collegate opportunamente;
- livello software:
 - il venditore deve installarlo sulle macchine e bisogna verificare che le promesse (ed i test) relative alle sue funzionalità corrispondano a verità;
 - integrare eventuali software diversi con il supporto del fornitore;
 - acquisire supporto tecnico;
- livello database: si seguono le procedure pianificate, in particolare
 - controllo di qualità dei dati acquisiti;
 - caricamento nell'area di lavoro con eventuali modifiche;
 - collegamento tra tabelle e dati spaziali;
 - integrare (unire/dividere) le immagini raster;
 - effettuare controlli finali;
 - immissione manuale dei dati nel database.

Sviluppo di applicazioni SIT

In questa fase si realizzano tutte le applicazioni SIT che nell'analisi dei requisiti sono state identificate come applicazioni non standard, cioè che richiedono programmazione aggiuntiva. La lunghezza e la complessità di questa fase dipendono dal livello di complessità della personalizzazione richiesta e da quanto tale applicazione possa sfruttare algoritmi esistenti. In generale questa in questa fase ci si limita ad inserire delle procedure di personalizzazione che fanno uso di macro-linguaggi o di altri linguaggi supportati dalla tecnologia SIT in uso.

Possiamo classificare le applicazioni da realizzare come segue:

- semplici personalizzazioni (query predefinite, maschere per l'utente, ecc.);
- analisi complesse che richiedono interazioni con utente o uso di altri pacchetti software (cammini minimi, visualizzazione grafica dello stato di una rete, ecc.);
- procedure personalizzate di verifica della qualità dei dati;
- procedure personalizzate di backup, archiviazione e caricamento dati.

Capitolo 7: Fasi finali

Uso e manutenzione del SIT

Una volta realizzato il SIT si passa al suo utilizzo a “pieno regime”. In questa fase vengono eseguite le procedure di gestione pianificate. Possiamo elencare alcune procedure di gestione standard:

- backup-restore
- gestione utenti e diritti
- procedure di logging delle modifiche
- archiviazione periodica
- controllo di qualità periodico: completezza, non ripetizione, posizione, attributi, errori topologici
- gestione dell’aggiornamento:
 - procedure per raccogliere le fonti, eseguire gli aggiornamenti, distribuire i dati aggiornati
 - effettuare le correzioni in area separata, eseguire procedure di CQ
- immissione di nuovi dati.

È ragionevole pensare che a breve o a lungo termine il sistema necessiti di manutenzione. Le cause che portano a richiedere la manutenzione di un sistema sono molteplici:

- necessità di nuovi requisiti da parte del committente;
- adattamento alle evoluzioni tecnologiche (hardware e/o software);
- a causa di un errato sviluppo del progetto la realizzazione del SIT non soddisfa a pieno (o in parte) la specifiche dei requisiti, nei casi più gravi questo può portare a una reingegnerizzazione dell’intero sistema.

Il sistema realizzato deve garantire un supporto di manutenzione ordinaria per la risoluzione di problemi ed errori, aggiornamento dei dati e/o dei vincoli.

Si identificano quindi due attività di manutenzione:

- attività di servizio e di supporto all’utenza;
- attività di manutenzione e aggiornamento (inizia logicamente non appena la raccolta dei requisiti termina, infatti i requisiti devono essere mantenuti aggiornati). Anche l’addestramento del personale deve essere mantenuto, nel senso che periodicamente potrebbe essere necessario fare nuovi corsi di formazione per
 - istruire il personale sulle modifiche apportate al sistema
 - mantenere il personale aggiornato sulle nuove tecnologie SIT.

Appendice A: Glossario

arco: elemento lineare archiviato in formato vettoriale costituito da una coppia di coordinate d'inizio (nodo iniziale), una coppia di coordinate finali (nodo finale) e da una serie di coppie di coordinate intermedie dette vertici. Ogni arco ha una direzione ed una lunghezza unendo più archi fra loro tramite i nodi, è possibile costruire dei grafi utilizzabili, ad esempio, per molti tipi d'analisi (stradale, idrografia, ecc.).

attributo: informazione descrittiva associata ad un oggetto geografico e che lo caratterizza. Generalmente nei GIS indica le caratteristiche non grafiche dell'elemento o quelle grafiche non rappresentabili nella scala d'acquisizione. Ad esempio potrebbero essere attributi di un elemento lineare (arco) che rappresenta una strada: il tipo di pavimentazione, il numero medio di veicoli in transito, la larghezza, ecc.

ArcInfo: software GIS dell'ESRI che gira sopra workstations UNIX o Windows NT. I dati geografici prodotti con ARCINFO possono essere utilizzati in un progetto d'ArcView.

ArcView: programma desktop GIS e di cartografia prodotto dall'ESRI. ArcView permette di visualizzare, analizzare e produrre dati geografici.

AutoCAD: programma desktop di Autodesk per applicazioni CAD. Vedi CAD

BIP (Band Interleaved by Pixel): è un formato standard di dati raster.

buffer, buffering: permette di creare all'intorno d'elementi predefiniti dei poligoni (aree di rispetto). Nel nostro caso anche l'area di impatto intorno all'evento incidentale può essere considerata una zona di buffering. Le dimensioni e quindi il raggio dell'area di rispetto possono essere definiti tramite attributi propri degli elementi bufferizzati: ad esempio per le strade (elementi lineari) la larghezza o l'intensità di traffico, per delle antenne (elementi puntuali) la potenza del segnale emesso, per delle zone industriali (elementi areali) il carico inquinante.

CAD (Computer Aided Design): è una famiglia di sistemi e programmi software per la progettazione assistita dal calcolatore, utilizzati per il disegno architettonico, il rilievo civile, il disegno meccanico e parametrico, il computo metrico, il disegno di schemi elettrici, ecc.

carta geografica: una carta geografica è una rappresentazione grafica di elementi geografici o altri fenomeni spaziali (laghi, fiumi, strade, città, ecc) che fornisce due tipi di informazioni circa l'area rappresentata: la posizione e l'aspetto. La posizione, consente di capire dov'è un particolare elemento geografico sulla superficie della terra e quali sono le sue relazioni con gli altri elementi. L'aspetto c'illustra sia le caratteristiche qualitative (nome, tipo, ecc) che quelle quantitative (area, lunghezza, ecc) dell'elemento rappresentato. Ciascuna carta geografica presenta, oltre alle informazioni spaziali e descrittive, alcune caratteristiche tecniche che la definiscono e ne determinano l'uso. Tali caratteristiche includono la scala, l'accuratezza, il sistema di proiezione, il sistema di coordinate.

cartografia digitale: nella versione digitale di una carta geografica tradizionale, la posizione e la descrizione degli elementi sono memorizzate in una serie di file su computer. Per individuare la posizione degli elementi geografici è utilizzato un sistema di coordinate cartesiane (x,y) o (x,y,z) : ogni punto è rappresentato con una singola coppia di coordinate (x,y) ; ogni linea è memorizzata come un serie ordinata di coordinate (x,y) ; ogni area è memorizzata come una serie ordinata di coordinate (x,y) che definiscono i segmenti perimetrali della figura chiusa. Con le coordinate (x,y) è possibile rappresentare punti, linee e poligoni come liste di coordinate, invece che come un disegno. Nella maggior parte dei casi la memorizzazione degli elementi geografici utilizza coordinate del mondo reale: queste coordinate rappresentano una localizzazione reale sulla superficie della terra, in uno dei tanti sistemi di coordinate.

cartografia di base: è l'insieme delle carte che riguardano un certo territorio, costruite con tecniche topografiche e fotogrammetriche, che rappresentano le planimetrie (corsi d'acqua, strade, ferrovie, ponti, case, ecc.) e l'altimetria composta da punti quotati o curve di livello.

connessione: è una proprietà topologica degli elementi di un GIS rappresentata dalla condivisione di un nodo tra due archi.

database: per database (base dati, banca dati o archivio) si intende un insieme di informazioni di diverso tipo. Questi dati sono organizzati secondo criteri ben precisi che permettono una rapida consultazione.

database geografico: database dedicato a informazioni di tipo territoriale.

database corporativo: database progettato per essere utilizzato da più organizzazioni. I database corporativi sono un'esempio di condivisione dei dati e sono molto importanti nel caso dei SIT in quanto permettono di diminuire il costo di acquisizione dei dati.

digitale: un dato digitale è un'informazione esprimibile mediante numeri interi che è possibile memorizzare in un supporto magnetico.

digitizer: dispositivo hardware per l'acquisizione manuale di informazioni grafiche (carte, documenti, ecc) in formato vettoriale. Consente ad un operatore di tracciare con un cursore, tutti gli elementi grafici che compongono il documento o la mappa da acquisire.

DTM (Digital Terrain Model): letteralmente modello digitale del terreno, raster o vettoriale, vedi DEM.

DXF (Drawing eXchange File format): formato di dati vettoriali non topologico tipico dei CAD utilizzato da AutoCAD, ma anche da molti GIS, per scambiare dati.

Framework: insieme di collezioni di entità che possono essere anche non omogenee, ma che sono correlate tra loro in quanto costituenti un sottosistema rilevante del SIT. (Ad esempio un SIT per la gestione idrica può essere suddiviso in due framework: uno relativo alle tubature e l'altro relativo alle abitazioni degli utenti).

georeferenziare: procedura software che consiste nel posizionare, mediante punti a coordinate note (punti di controllo), dati vettoriali o un'immagine raster nella rispettiva zona del territorio reale secondo un determinato sistema di riferimento.

GIS (Geographical Information System): questo termine, sebbene sia la traduzione letterale del SIT (sistema informativo territoriale), viene comunemente utilizzato per riferirsi alla componente informatica di un SIT. Con "strumento GIS" si intende il pacchetto di applicativi che permettono di acquisire, modellare, manipolare, restituire, analizzare e presentare dati georeferenziati.

GPS (Global Positioning System): è un sistema che consente, per mezzo di un ricevitore, un software dedicato e una costellazione di satelliti, di determinare la posizione al suolo e l'altimetria di un punto, con una precisione che varia da pochi millimetri ad alcuni metri in funzione del tipo d'apparecchiatura e delle procedure operative di rilievo.

GRID: è il nome dell'estensione di ARC/INFO della ESRI che elabora basi dati in formato raster ed è utilizzato anche per definire il corrispondente formato di dati.

GUI (Graphical User Interface): software per la creazione e gestione d'interfacce utente di tipo grafico. Permette all'utente di accedere alle funzioni del sistema anche per mezzo del mouse e di semplici oggetti grafici (pulsanti, liste a scorrimento, icone, immagini, ecc).

immagine raster: qualsiasi immagine può essere pensata come formata da un insieme di piccole aree uguali (pixel), ordinate secondo linee e colonne, tali da costituire una matrice. I valori associati ad ogni cella possono esprimere sia informazioni di tipo grafico (colore, tono di grigio, ecc) sia di tipo descrittivo, (temperatura, pendenza, ecc).

interfaccia utente: è l'insieme degli strumenti che costituiscono il mezzo di comunicazione tra l'utente e le procedure del sistema hardware e software utilizzato.

layer: struttura di memorizzazione topologica di componenti spaziali di una o più entità. Si assume che, se un layer è condiviso da più entità, tra le componenti spaziali di tali entità sussiste un vincolo topologico.

overlay, overlay topologico: procedura di analisi spaziale che consente di sovrapporre e intersecare gli strati informativi (layer) unendo così le informazioni associate a ciascuno di essi per produrre un nuovo strato di sintesi.

percorso minimo: è un algoritmo di analisi che consente di determinare, all'interno di un grafo, il percorso che ottimizza un parametro della rete, ad esempio su una rete stradale il tragitto che rende minima la distanza o il tempo tenendo conto dei limiti di velocità, di carreggiata, ecc.

pixel: contrazione di picture element, è il componente elementare di un'immagine raster caratterizzato da un valore associato.

poligono: un'area chiusa delimitata da archi.

proiezione geografica: è un sistema di proiezione globale, utilizzato storicamente per primo, basato sulla suddivisione della terra in meridiani e paralleli. Il suo sistema di riferimento, latitudine/longitudine, misura le distanze angolari partendo rispettivamente dal piano di riferimento dell'equatore e da quello del meridiano di Greenwich.

punto: elemento geometrico utilizzato per definire la posizione di elementi che nella scala di acquisizione sono puntiformi, definito da una coppia di coordinate (x,y) ed eventualmente da una quota (z).

raster: Vedi immagine raster.

scala: la scala di una carta geografica indica, in pratica, quante volte una porzione della superficie terrestre è stata ridotta per poter essere rappresentata su un foglio di carta. È espressa, in genere, come il rapporto tra una distanza sulla carta e la corrispondente distanza sul terreno. Il rapporto di scala utilizzato per una carta geografica determina il contenuto di informazioni e la dimensione dell'area che può ragionevolmente essere rappresentata.

scanner: dispositivo hardware per l'acquisizione automatica in formato digitale di elementi grafici (carte, foto, immagini, ecc). Le informazioni grafiche sono memorizzate in formato raster. È possibile tuttavia tramite appositi strumenti software convertire tali informazioni in formato vettoriale (vettorializzazione).

Sistema Informativo : insieme di hardware, software, dati, procedure aziendali, conoscenze, persone che raccolgono, gestiscono, elaborano, distribuiscono informazioni.

SIT (Sistema Informativo Territoriale): un sistema informativo basato su computer che cerca di catturare, immagazzinare, manipolare, analizzare e visualizzare dati spaziali con associati degli attributi, al fine di risolvere ricerche complesse, pianificare e gestire problemi.

strato informativo, strato geografico: collezione di entità di tipo omogenee. Si usa il termine strato in quanto una generica mappa può essere vista come la stratificazione di una serie di rappresentazioni geografiche di strati informativi.

tabella di attributi: le tabelle sono una parte integrante dello strato informativo. Ogni tabella è relativa ad un gruppo omogeneo d'elementi geografici della carta (le strade, i fiumi, le curve di livello, ecc) ed è costituita da un numero variabile di righe e colonne. Ogni riga (record) contiene la descrizione di un singolo elemento geografico ed ogni colonna (campo o attributo) memorizza uno specifico tipo di informazione.

topologia: la topologia è un insieme di regole per definire in maniera esplicita le relazioni, i rapporti di connessione e di contiguità tra gli elementi spaziali e per collegare tali elementi alle relative descrizioni (attributi). In un modello dati topologico, ad esempio, è possibile riconoscere le aree contigue e identificare le linee che delimitano ciascuna superficie (confini).

vettoriale: è un sistema di archiviazione di dati grafici, secondo il quale gli oggetti vengono memorizzati in base alle coordinate cartesiane dei punti e linee che li compongono.

vettorializzazione: operazione che consente in modo automatico o semiautomatico (cioè con l'assistenza di un operatore) di ricavare un insieme di vettori a partire da una immagine raster.

Bibliografia

- [GdG] – GIS Development Guide. Local Government Technology Services State Archives And Records Administration Albany, New York.
<http://www.archives.nysed.gov/pubs/gis/gisindex.htm>
- [AlGhO] – Basi di dati relazionali e a oggetti. A. Albano, G.Ghelli, R. Orsini. Zanichelli, 1997.
- [SpZi] – Modeling Spatial Data in the MADS Conceptual Model. S. Spaccapietra, E. Zimanyi et al. in Proc. of the International Symposium on Spatial Data Handling, SDH 98, Vancouver, Canada, 1998.
- [Zei99] – Modeling Our World. M. Zeiler. The ESRI Guide to Geodatabase Design. ESRI Press 1999.

