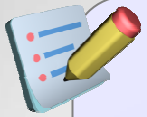


# Progettazione di un DW



I principali approcci per sviluppare un ambiente di data warehouse sono due:

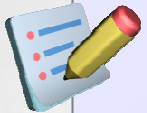
1. il primo è basato sulla creazione di un **data warehouse centrale**, usando dati dal sistema principale ed altre fonti. Questo data warehouse centrale può essere poi usato per creare/ aggiornare data warehouse dipartimentali o data mart locali.
2. Il secondo approccio è basato sulla creazione di **data mart indipendenti**, ognuno memorizzato direttamente dal sistema centrale e altre fonti dei dati.

# Data warehouse centrale

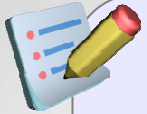
L'approccio di un data warehouse centrale può iniziare con un data warehouse semplice, ampliabile nel tempo per soddisfare utenti con richieste crescenti e diventare un ambiente che contenga sistemi di data warehouse interconnessi. In un ambiente di data warehouse semplificato bisogna organizzare tre aree:

- l'estrazione e la trasformazione dei dati dai sistemi operativi;**
- la base di dati del data warehouse;**
- gli strumenti per interpretare i dati.**

È necessario monitorare la rete che consente l'accesso agli utenti. Ci sono di solito almeno tre **repository** per i metadati e per le altre informazioni collegate: uno per **descrivere la struttura dei dati**, per la loro trasformazione e per l'estrazione dei dati; uno per il **database** del data warehouse; ed uno o più per gli **strumenti di navigazione**. Questi repository devono essere curati individualmente e complessivamente. I dati nell'ambiente del database del data warehouse dovrebbero essere maneggiati con la stessa cura. La complessità di questo compito dipende dalla complessità del database scelto, ma include copie di backup, recovery, riorganizzazioni, archiviazioni, operazioni di monitoraggio e tuning. Sono creati sub-set di dati dipartimentali o locali (data marts) per migliorare la performance delle consultazioni dell'utente e ridurre la dipendenza dal data warehouse. Questo livello aggiuntivo di dati aumenta la complessità di gestione dell'ambiente: aggiunge un altro livello di metadati e possibilmente un altro repository, richiede controllo e gestione della distribuzione dei dati dei data mart, e, a meno che l'amministrazione dei data mart sia completamente devoluta a livello locale, richiede anche la gestione di dati del database del data mart. La situazione diventa anche più difficile se l'ambiente continua ad evolvere tramite la creazione di data warehouse multipli. In alcuni di questi casi, le complessità di amministrazione diventano opprimenti.



# Data mart indipendenti/1

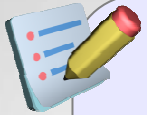


Nell'approccio con data mart indipendenti, la creazione di un solo data mart orientato a risolvere un particolare problema rappresenta una soluzione semplice. Le tre aree da amministrare sono:

- l'estrazione dei dati dalle fonti e la trasformazione nelle strutture dei dati corrette per il database del data mart;
- il database del data mart stesso;
- gli strumenti per interpretare i dati.

Poiché questo ambiente non contiene grandi volumi di data warehouse esso è più maneggevole. Nel caso si adotti una tale semplice soluzione di data mart nella realizzazione di data warehouse e nell'organizzazione, il compito dell'amministratore sarebbe relativamente facile. Questo approccio non si ferma di solito ad un data mart e, una volta che vengono aggiunti altri data mart, la situazione diventa più complicata. Il compito di portare numerosi data mart separati in un solo ambiente di data warehouse è estremamente difficile. Ogni data mart viene sviluppato di solito individualmente. Tali data mart hanno il potenziale di diventare parte del sistema centrale. In questo modo, possono porre il problema di discordanze nella definizione dei dati che il data warehouse è stato disegnato per risolvere. Questa situazione poco attraente si evita solamente se esiste un'architettura centralizzata di amministrazione dello sviluppo del sistema.

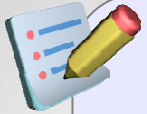
# Data mart indipendenti/2



Il data warehouse potrebbe arrivare a contenere volumi molto grandi di dati, non sempre interessanti per tutti gli utenti. Lavorare con questi volumi di dati non correlati può essere inefficiente e consumare molte risorse di calcolo. In questa situazione è possibile suddividere il data warehouse in aree di interesse specializzate.

Inoltre, molti tool per lo sfruttamento dei dati creano i loro primi ambienti, ognuno col proprio repository. Tale repository contiene le informazioni richieste per l'esplorazione dei dati. Se il data warehouse è amministrato centralmente, questi ambienti devono essere incorporati nella struttura di gestione centrale. Anche dove la responsabilità dell'amministrazione dei tool di sfruttamento dei dati è a livello locale, serve un collegamento tra il sistema di amministrazione centrale e gli ambienti distribuiti. Questo collegamento è necessario per assicurare che i cambiamenti dei tool degli ambienti distribuiti possano essere identificati anche centralmente.

# Altri aspetti progettuali



I livelli 'operativi' del data warehouse possono esistere sotto due condizioni fondamentali:

- l'esistenza di un'adeguata organizzazione di supporto al processo, con ruoli e responsabilità definiti. In modo analogo alle applicazioni transazionali, un sistema di decision support necessita di figure organizzative con la responsabilità di mantenerlo, soprattutto in chiave evolutiva, per far sì che esso sia costantemente allineato alle esigenze degli utenti di business, condizione necessaria e sufficiente perché continui ad esistere;
- il giusto rilievo alla tecnologia di supporto al processo, composta di scelte equilibrate e basate sulle esigenze funzionali del processo stesso. La tecnologia è cruciale per il data warehouse, date le problematiche di system integration che esso comporta. La gestione costante della variabile tecnologica è uno dei fattori di successo del data warehouse, a partire dalle scelte iniziali per arrivare alla gestione operativa degli aggiornamenti e degli ampliamenti della piattaforma.

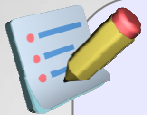
## Applicazioni del data warehouse

Il data warehouse è un sistema informativo dove i dati sono organizzati e strutturati per un facile accesso da parte dell'utente e per fornire supporto ai processi decisionali. I seguenti sistemi sono abilitati dal data warehouse:

**DSS** (Decisional Support System). Usato per risolvere problemi specifici

**EIS** (Executive/Enterprise Information System). Consente una continua circolazione dei dati non dipendente da problemi specifici.

# Strategia

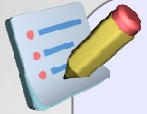


Nelle banche e in generale nelle istituzioni finanziarie gli ambiti di utilizzo sono molteplici, poiché tutte le aree gestionali di tali organizzazioni sono caratterizzate da volumi considerevoli di dati su cui devono essere prese decisioni strategiche. Poiché il data warehouse può avere un valore strategico, all'interno di tali tipi di organizzazioni è fondamentale per il management definire una strategia per il data warehouse. **La strategia per il data warehouse** è essenzialmente un percorso evolutivo che porta l'azienda da applicazioni DW non 'mission-critical' verso una situazione in cui il data warehouse è una componente fondamentale del sistema informativo aziendale.

La strategia di data warehousing di un'azienda può essere classificata in base a due dimensioni fondamentali:

- **utilizzo del DW esistente: livello di maturità degli utenti e delle funzioni di supporto del DW nell'utilizzo dell'esistente;**
- **utilizzo del DW in prospettiva: di utilizzo del DW come piattaforma di decision support.**

# Fasi nella storia dell'utilizzo del DW



**La prima fase**, chiamata supporto (basso utilizzo del DW esistente, basso utilizzo prospettico del DW), è la fase in cui si trovano le aziende che hanno fallito uno o più progetti di warehousing e non pensano di ampliarne l'utilizzo prospettico. In questa fase si possono trovare anche aziende che non hanno un DW e non pensano di realizzarlo;

**La seconda fase**, chiamata opportunità (basso utilizzo del DW esistente, alto utilizzo prospettico del DW), è la fase in cui si trovano le aziende che, pur avendo fallito uno o più progetti di warehousing o avendo semplicemente esplorato la tematica senza approfondirla, puntano a sviluppare le attività di decision support tramite il data warehouse.

**La terza fase** (alto utilizzo del DW esistente, alto utilizzo prospettico del DW), è quella fase in cui il data warehouse diviene strategico per i processi decisionali aziendali. In questa fase si trovano tutte quelle aziende che hanno intrapreso con successo un progetto di warehousing e che ne stanno sfruttando a pieno le potenzialità;

**La quarta fase**, chiamata factory (alto utilizzo del DW esistente, basso utilizzo prospettico del DW) è la fase in cui si trovano le aziende in cui il data warehouse è maturo, la metodologia di implementazione consolidata e le aree decisionali critiche sono presidiate. In questa fase l'imperativo principale è l'efficienza e il risparmio di costi derivanti dal data warehouse e nel suo utilizzo. Un processo di sclerotizzazione nell'uso del data warehouse può in alcuni casi far tornare l'azienda alla prima fase.

## Aree applicative più indicate per il data warehouse nel settore finanziario



### **Controllo di gestione**

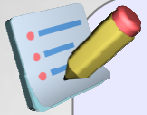
Questa può essere l'area applicativa di base per un sistema di data warehousing in qualunque organizzazione. In questo caso il data warehouse viene utilizzato sostanzialmente come piattaforma di reporting e analisi di redditività. È inutile e pericoloso ipotizzare di realizzare un data warehouse solo per il controllo di gestione. Tale iniziativa ha senso solo se questo è il primo passo evolutivo nella strategia di data warehousing dell'azienda. Infatti, costruire un data warehouse per il controllo di gestione consente di analizzare e risolvere rapidamente esigenze estremamente rilevanti ed il cui beneficio è immediatamente chiaro, affrontando problemi (a livello di struttura, validazione e calcolo dei dati) ben noti nella loro struttura.

### **Risk e Asset Management**

Un'altra area applicativa interessante è identificabile nelle attività di Risk e Asset Management (vedi Gestione del rischio), soprattutto in due attività ben specifiche: l'analisi e la simulazione dei portafogli e dei relativi rischi; il reporting.

Tali aree applicative sono di particolare importanza e strategicità ed il data warehouse è lo strumento appropriato per affrontarle, anche per la possibilità di integrare al suo interno dati provenienti da fonti esterne all'azienda. In questo caso il data warehouse va dotato di strumenti di analisi avanzati e basati su algoritmi statistici di analisi e simulazione.

# Sviluppo di sistemi per l'individuazione delle frodi/1

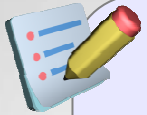


Un'altra sotto-area di grande interesse può essere lo sviluppo di sistemi per l'individuazione delle frodi. Anche in questo caso è necessario il ricorso a strumentazione di tipo statistico.

## **Supporto alle vendite**

Non necessariamente il data warehouse è appropriato per affrontare e risolvere questo tipo di esigenza, a meno che esista la necessità di immagazzinare e gestire rilevanti masse di dati. In molti casi il database di marketing è banalmente un'anagrafica clienti arricchita di alcune informazioni "non amministrative", in casi più avanzati diventa uno strumento fondamentale di supporto al "marketing one-to-one". In questo caso il database di marketing costituisce una base di informazioni fondamentale per indirizzare correttamente campagne e iniziative promozionali o per attivare servizi avanzati di "customer care". In questo caso, data la rilevante massa di dati da gestire, il data warehouse può diventare la piattaforma tecnologica ideale. Nel settore bancario il marketing one-to-one è ancora allo stadio embrionale, almeno dal punto di vista del marketing centrale, e questo è dovuto al fatto che molto spesso il marketing one-to-one viene fatto dalla filiale, l'unica struttura aziendale in grado storicamente di instaurare un rapporto fiduciario con il cliente finale, che identifica l'azienda nello 'sportello' e nel suo 'impiegato'.

# Sviluppo di sistemi per l'individuazione delle frodi/2



## **Sistema informativo di marketing**

Si tratta di utilizzare il data warehouse come una sorta di 'backbone' per supportare una serie di applicazioni integrate orientate alle analisi commerciali e di marketing. Gli aspetti fondamentali che caratterizzano questo tipo di architettura sono essenzialmente due:

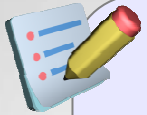
- la possibilità di integrare basi di dati transazionali diverse in un'unica base dati analitica e produrre quindi 'viste' integrate della clientela, del mercato e dei prodotti;
- la possibilità di effettuare analisi con strumenti e logiche diverse su una base unica.

L'idea di fondo del sistema informativo di marketing è quella di sviluppare un percorso evolutivo che parta dal reporting di base per arrivare ad analisi avanzate, passando attraverso sistemi di analisi del portafoglio prodotti e clienti e procedure di budgeting e simulazione.

## **Supporto al Call Center**

Anche in questo caso il data warehouse è un'opzione tecnologica, non l'unica praticabile e non necessariamente la più economica. Utilizzare un'architettura di data warehousing a supporto di un'attività di Call Center ha sicuramente senso nel caso in cui le richieste non sono necessariamente di tipo strutturato e quindi risolvibili con il classico "inquiry (interrogazione) da terminale". È evidente però che la tipologia di utente per questo tipo di sistema è più evoluto del normale operatore di Call Center.

# Sviluppo di sistemi per l'individuazione delle frodi/3



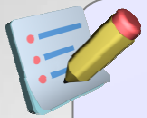
## **Knowledge Base**

Anche in questo caso valgono le considerazioni fatte per il Database di Marketing: non necessariamente il data warehouse è la tecnologia più idonea per questo tipo di esigenza, ma lo diventa nel momento in cui la conoscenza in oggetto è costituita prevalentemente da informazioni strutturate e preferibilmente numeriche. In questo caso, anche dal punto di vista tecnologico, un database relazionale è sicuramente la soluzione più idonea, efficiente ed economica. Non è così se invece le informazioni sono di tipo destrutturato, in questo caso la soluzione più adatta è una piattaforma di groupware. Si deve però fare attenzione a non confondersi con i cosiddetti database multimediali: il fatto che un database relazionale abbia funzionalità multimediali non significa che sia un data warehouse. Infatti, ciò che distingue un data warehouse da ciò che non lo è, non è la tecnologia utilizzata, ma l'architettura applicativa e il disegno della base di dati.

## **Engineering di prodotto**

Il data warehouse può essere una piattaforma decisionale per l'analisi e la concettualizzazione di nuovi prodotti da offrire alla clientela e/o per aggredire nuovi mercati o segmenti di mercato. Tale funzionalità è ovviamente supportata se il data warehouse è dotato non solo di strumenti di analisi dei risultati, ma anche di ambienti di simulazione che consentono la costruzione ed il testing 'in laboratorio' di nuove soluzioni da proporre ai clienti. In tali ambienti è possibile individuare alcuni importanti aspetti come la marginalità, il punto di pareggio economico, il segmento di clientela interessato, i meccanismi di cannibalizzazione, l'elasticità della domanda e l'impatto sull'equilibrio finanziario aziendale.

# Sviluppo di sistemi per l'individuazione delle frodi/4



## **e-business**

La diffusione del canale digitale nel settore finanziario pone una serie di problemi e di opportunità nuove. In primo luogo questo tipo di canale implica una velocità di cambiamento e quindi di reazione nettamente superiore. Il data warehouse può essere lo strumento analitico che consente di cogliere dinamiche all'interno di rilevanti masse di transazioni on-line. In secondo luogo l'informazione può essere uno strumento di supporto o l'oggetto stesso della transazione e in questo caso il data warehouse può essere la piattaforma utilizzata per coprire tale ambito applicativo.



Il data warehouse può essere quindi di supporto a sistemi di trading on-line sia dal punto di vista dell'analisi che dal punto di vista dell'architettura dati.

---

**Grazie per l'attenzione**

**Alessandro Di Benedetto**

**alessandro.dibenedetto@ghirbe.it**

**alessandro.dibenedetto@luxhora.it**

Via Leopardi, 3  
20123 Milano  
Tel 02 48517553  
Fax 02 48109189  
e-mail: [info@ghirbe.it](mailto:info@ghirbe.it)  
Web site : [www.ghirbe.it](http://www.ghirbe.it)