

Appunti di Sistemi Informativi

Luca Facchini

Matricola: 245965

Corso tenuto dal prof. Casari Paolo

Università degli Studi di Trento

A.A. 2024/2025

Sommario

Questo documento contiene gli appunti del corso di Sistemi Informativi, tenuto dal prof. Casari Paolo presso l'Università degli Studi di Trento. Il corso è stato seguito nell'anno accademico 2024/2025.¹

¹Dove non specificato diversamente eventuali immagini provengono dalle slide del corso (o da materiale didattico fornito dal docente) le quali sono fornite e tratte dal libro: "Sistemi Informativi Aziendali" di Pighin M. e Mazona A. solo per le immagini possono essere soggette a copyright e comunque usabili solo per fini didattici e in base alla legge italiana sul diritto d'autore (al massimo 15%).

Indice

1	La società della conoscenza	4
1.1	Introduzione	4
1.2	Nuove leggi della società della conoscenza	4
1.2.1	Leggi di Moore	4
1.2.2	Leggi di Sarnoff, MetCalfe, Reed	5
1.3	Hype Cycle di Gartner	6
1.3.1	La curva della domanda	6
1.3.2	Hype Cycle	6
1.3.3	Magic Quadrant di Gartner	8
1.4	Long Tail di Anderson	8
2	Concetti Generali	10
2.1	Introduzione e definizioni	10
2.2	Sistema informativo aziendale	10
2.3	Impatto dell'informatica nelle azienda	11
2.3.1	Conoscenza dei fenomeni aziendali	11
2.3.2	Processi classi di un sistema informativo	12
2.3.3	Nuovi processi	12
2.4	I SI nelle aziende Italiane e relazione ICT - azienda	12
2.4.1	Le aziende in italia	12
2.4.2	Cambiamenti dei SI	13
3	Struttura aziendale e del suo SI	14
3.1	Concetto di esigenza informativa	14
3.2	Scomposizione dei sistemi informativi	15
3.2.1	Sistemi operazionali	15
3.2.2	Sistemi informazionali	16
3.2.3	Comparazione	16
4	Le scelte organizzative	17
4.1	Costruzione del SI	17
4.1.1	Opzione <i>make</i>	17
4.1.2	Opzione <i>buy</i>	17
4.1.3	Opzione <i>outsource</i>	18
4.2	Le figure professionali	18
4.2.1	Sviluppo del reparto	18
4.2.2	Posizione all'interno dell'organigramma	20
4.3	Infrastruttura Tecnologica	20
4.4	Interrompibilità del servizio informatico	21
5	I sistemi operazionali	22
5.1	Finalità dei sistemi operazionali	22
5.1.1	Transazioni - Definizione e registrazione	22
5.1.2	Pianificazione e controllo delle operazioni	23
5.1.3	Organizzazione della conoscenza aziendale	23
5.1.4	Elaborazione delle situazioni aziendali	23
5.2	Informazione Operativa	23

5.2.1	Qualità dei dati	24
5.2.2	Caratteristiche strutturali	24
5.2.3	Caratteristiche funzionali	25
5.3	Potenzialità informatica	25
5.3.1	Intensità informativa	25
5.3.2	Attrattiva Informatica	26
5.4	Composizione dei SI Operazionali	26
6	ERP	29
6.1	Area Amministrativa	29
6.2	L'area logistica	30
6.3	L'area vendite	30
6.4	L'area acquisti	30
6.5	L'area produttiva	30
6.6	Sistemi Operazionali complementari	30
7	I sistemi Informazionali	31
7.1	Gli Obbiettivi	31
7.2	Concetti Generali	32
7.2.1	Terminologia	32
7.2.2	Le caratteristiche	32
7.3	Modello multidimensionale	32
7.3.1	Caratteristiche	33
7.3.2	Aggregabilità	33
7.4	Caratteristiche strutturali e funzionali	34
7.4.1	Caratteristiche strutturali	34
7.4.2	Caratteristiche funzionali	34
7.4.3	<i>Data warehouse</i> e <i>Data mart</i>	35
8	<i>Data Warehousing</i>	36
8.1	<i>Data Warehouse</i> e metodologia OLAP	36
8.2	Architettura dei sistemi di <i>data warehousing</i>	36
8.3	Modelli concettuali per il <i>data warehouse</i>	37
8.3.1	<i>Dimensional Fact Model - DFM</i>	37
8.4	Modelli logici per il <i>data warehouse</i>	38
8.4.1	ROLAP	38
8.4.2	MOLAP	38
8.4.3	HOLAP	38
8.4.4	Schemi multidimensionali su basi di dati relazionali	39
8.5	Ciclo di vita di DWH e Popolazione del DWH	40
8.5.1	Ciclo di vita di un DWH	40
8.5.2	Popolazione della DWH	40
8.6	L'analisi OLAP e principali operatori OLAP	40
9	Introduzione al <i>Data Mining</i>	42
9.1	Limiti analisi OLAP	42
9.2	Architettura e tipi di analisi con <i>data mining</i>	42

Capitolo 1

La società della conoscenza

In questo capitolo andremo ad analizzare il concetto di “società della conoscenza” e le nuove leggi che regolano questa società. Inoltre andremo ad analizzare il *Hype Cycle* di Gartner e il *Long Tail* di Anderson.

1.1 Introduzione

Cos'è una “società della conoscenza” Il concetto di “società della conoscenza” si riferisce ad un modello di società nel quale la **conoscenza**, l'**informazione** e l'**innovazione** diventano i principali **mezzi della crescita economica e dello sviluppo sociale**. In questo tipo di società viene “premiato” chi ha la **capacità di apprendere, di innovare e di adattarsi ai rapidi cambiamenti tecnologici**.

Perché è importante? La **conoscenza** è importante perché in una economia basata su questa la ricchezza e il potere sono determinati dalla **capacità di creare**, inoltre la **innovazione e competitività** è essenziale e la nuova idea domina i mercati. Altro asset importante è il **capitale umano e l'istruzione** in quanto le persone ben formate sono in grado di creare valore. La **globalizzazione e interconnessione** sono altri fattori importanti in quanto la conoscenza è un bene che si diffonde rapidamente e facilmente. Infine anche lo **sviluppo sostenibile** è importante in quanto la conoscenza permette di trovare soluzioni a problemi ambientali.

1.2 Nuove leggi della società della conoscenza

In questa sezione sono analizzate le nuove leggi che regolano la società della conoscenza, ovvero la **legge di Moore**, la **legge di Sarnoff**, la **legge di MetCalfe** e la **legge di Reed**.

1.2.1 Leggi di Moore

Di seguito sono analizzate le due leggi di Moore collegate al mondo dell'informatica.

Prima legge di Moore

La **legge di Moore** è un'osservazione fatta da **Gordon Moore** nel 1965

Definizione 1.1 (Prima legge di Moore). *La **densità di transistor** nei circuiti integrati raddoppia ogni 18 mesi.*

Limiti I limiti della prima legge di Moore riguardano il raggiungimento dei limiti fisici della materia.

Conseguenze Questo comporta che la potenza di calcolo dei computer raddoppia ogni 18 mesi. Come conseguenza tale legge suggerisce un **aumento esponenziale** della potenza di calcolo dei computer, portando ad un **aumento della velocità** e della **capacità di memorizzazione** dei computer. Questo ha portato ad un **aumento della diffusione** dei computer e ad una **riduzione dei costi**.

Seconda legge di Moore

La **seconda legge di Moore** è una osservazione eseguita da **Gordon Moore** nel 1975, successivamente integrata da altri autori.

Definizione 1.2 (Seconda legge di Moore). *L'investimento per realizzare una nuova tecnologia di microprocessori cresce in maniera esponenziale con il tempo.*

Conseguenze Questa legge implica che per aumentare la potenza di calcolo dei computer è necessario un **investimento sempre maggiore**. Questo ha portato ad un **aumento dei costi** per la realizzazione di nuove tecnologie.

Questi effetti fanno sì che le società che si possono permettere di investire sono sempre meno e quelle che da sole non riescono ad investire si uniscono. Effetto economico di questa legge è l'aumento del rischio degli investimenti. Difatti la probabilità di fallimento di una nuova azienda si incrementa con il tempo.

1.2.2 Leggi di Sarnoff, MetCalfe, Reed

Dopo aver analizzato le leggi di Moore riguardanti il mondo dell'informatica, analizziamo le leggi di Sarnoff, MetCalfe e Reed che riguardano il mondo delle reti nella società della conoscenza.

Legge di Sarnoff

La **legge di Sarnoff** è un'osservazione fatta da **David Sarnoff** nel 1950 la quale interessa il valore di un sistema di comunicazione del tipo *broadcast*.

Definizione 1.3 (Legge di Sarnoff). *Il valore V di una rete di broadcasting è direttamente proporzionale al numero N di utenti della rete.*

$$V = N$$

Conseguenze Il valore della rete aumenta con il numero di utenti.

Legge di MetCalfe

La **legge di MetCalfe** è un'osservazione fatta da **Robert MetCalfe** nel 1980 la quale interessa il valore di un sistema di comunicazione del tipo *peer-to-peer* ovvero una rete di relazione sociale. Esempi di reti di relazione sociale sono la rete telefonica o il sistema di fax per le quali è possibile una comunicazione bidirezionale ma limitata ad 1-1.

Definizione 1.4 (Legge di MetCalfe). *Il valore V di un sistema di comunicazione cresce con il quadrato del numero di utenti N della rete.*

$$V = N^2 - N$$

Implicazione La connessione di reti indipendenti crea un valore più elevato rispetto alla somma del valore delle singole reti. Questa legge è molto legata al successo di internet quando questa si basava sulla esistente rete telefonica.

Legge di Reed

La **legge di Reed** è un'osservazione fatta da **David Reed** nel 1999 la quale interessa il valore di un sistema di comunicazione nella quale è possibile la comunicazione tra più di due utenti o gruppi di utenti. Esempi di reti di relazione sociale sono i social network o i forum.

Definizione 1.5 (Legge di Reed). *L'utilità delle grandi reti, formate da reti di reti (con particolare riferimento alle reti di relazione sociale) cresce esponenzialmente con il numero di nodi.*

$$V = 2^N - N - 1$$

Conseguenze Questa legge implica che il valore di una rete sociale cresce esponenzialmente con il numero di nodi. Questo è dovuto al fatto che con ogni nodo si possono creare nuovi sottogruppi di nodi.

Conclusione leggi di Sarnoff, MetCalfe, Reed

Se si distribuisce un solo contenuto allora si ha una crescita lineare del valore, se si attivano transizione per il commercio elettronico si ha una crescita quadratica del valore, se si crea una comunità si ha una crescita esponenziale del valore.

Per chi investe è meglio puntare sulla legge di Reed in quanto il valore cresce esponenzialmente con il numero di nodi.

1.3 Hype Cycle di Gartner

L'**Hype Cycle** è un modello sviluppato da **Gartner** (società) che rappresenta la **maturità, l'adozione e l'applicazione delle tecnologie emergenti**. Il modello è rappresentato da un grafico che mostra la **curva di hype** di una tecnologia, ovvero la **fase di crescita e declino** di una tecnologia.

1.3.1 La curva della domanda

Prima di analizzare l'**Hype Cycle** è importante conoscere la **curva della domanda** la quale rappresenta come un determinato prodotto e/o tecnologia viene adottato dal mercato.

Componenti della curva La curva è composta da tre picchi di acquisto: i **pionieri**, la **maggioranza** e i **ritardatari**.

I **pionieri** sono quella cerchia ristretta di persone che sono disposte a comprare un prodotto appena uscito anche a costi elevati. La **maggioranza** è quella cerchia di persone che comprano un prodotto quando il prezzo è ragionevole ma pretendono tecnologie semplici e facili da usare e sono sensibili ai trend creati dai pionieri. Infine i **ritardatari** sono coloro che comprano un prodotto quando ormai è diventato un bene di uso comune e non si possono esimere dall'acquisto per rimanere competitivi sul mercato, questi sono molto prudenti e non amano rischiare inoltre la loro preoccupazione principale è il costo.

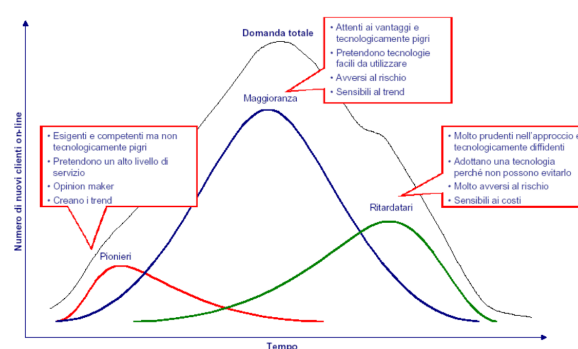


Figura 1.1: Curva della domanda

1.3.2 Hype Cycle

Il grafico di seguito illustra la **curva di hype** di una tecnologia, ovvero la **fase di crescita e declino** di una tecnologia. Questo grafico è importante per capire come una nuova tecnologia viene richiesta ed adottato dal mercato.

Le fasi

Questa curva ha diversi alti e bassi:

1. **Innovazione** è la fase in cui la tecnologia viene presentata al mercato e le prime startup iniziano a sviluppare prodotti basati su questa tecnologia.
2. **Crescita esponenziale** è la fase in cui la tecnologia inizia a venire seguita da un numero sempre maggiore di persone e il mercato potenziale dietro a questa tecnologia inizia a crescere. I mass media iniziano a parlare di questa tecnologia. I primi prodotti basati su questa tecnologia iniziano ad essere venduti ad un costo molto elevato.

3. **Picco delle aspettative inflazionate** è la fase in cui la tecnologia raggiunge il suo picco di interesse e le aspettative sono molto alte. In questa fase la tecnologia è vista come la soluzione a tutti i problemi.
4. **Declino** è la fase in cui la tecnologia non riesce a soddisfare le aspettative e il mercato inizia a perdere interesse, in quanto la tecnologia non è in grado di risolvere i problemi del suo punto di massimo.
5. **Risalita della produttività** è la fase in vengono realizzati la seconda/terza generazione dei prodotti basati su questa tecnologia e il mercato inizia a capire come utilizzare questa tecnologia in modo efficace. In questa fase la tecnologia inizia a diventare un bene di uso comune.
6. **Piatto di produttività** è la fase in cui la tecnologia è diventata un bene di uso comune e il mercato inizia a saturarsi. Qui la tecnologia viene usata per risolvere problemi specifici ma non è più vista come la soluzione a tutti i problemi (20/30 % del mercato rispetto al picco).

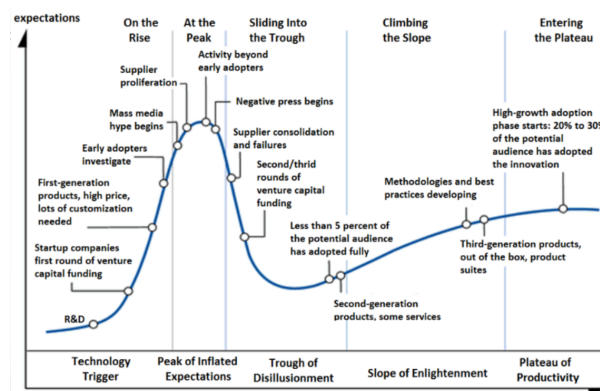


Figura 1.2: Curva di hype

Qualche volta però la tecnologia non riesce a riprendersi dal declino, vengono quindi proposte altre due curve con fasi che si sostituiscono alle fasi 5 e 6:

5. **Cimitero delle tecnologie** è la fase in cui la tecnologia non riesce a riprendersi dal declino e viene abbandonata.
5. **Palude di uso comune** è la fase in cui la tecnologia è diventata un bene di uso comune ma non riesce a trovare nuove applicazioni e il mercato è molto più ristretto, il prodotto viene usato solo da una cerchia ristretta di persone.

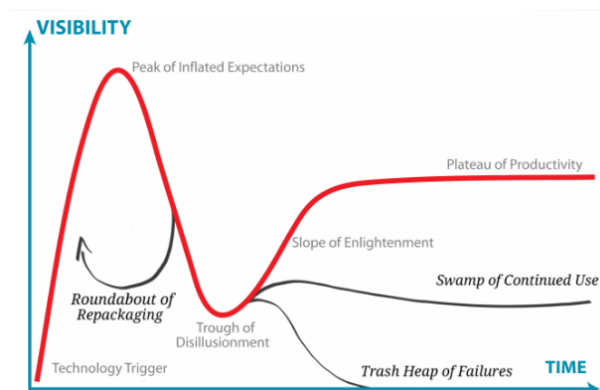


Figura 1.3: Curva di hype con fasi alternative

1.3.3 Magic Quadrant di Gartner

Il **Magic Quadrant** è un modello sviluppato da **Gartner** che permette di valutare le **aziende** in base alla loro **completezza della visione** e alla loro **abilità di esecuzione**. Questo modello è importante per capire come un'azienda si posiziona rispetto ai suoi concorrenti e per capire quali sono i punti di forza e di debolezza di un'azienda.

I quadranti

Il modello è composto da quattro quadranti:

- **Leader** sono le aziende che hanno una **completa visione** del mercato e che sono in grado di **eseguire** in modo efficace.
- **Challenger** sono le aziende che hanno una **completa visione** del mercato ma che non sono in grado di **eseguire** in modo efficace.
- **Visionary** sono le aziende che hanno una **visione parziale** del mercato ma che sono in grado di **eseguire** in modo efficace.
- **Niche Player** sono le aziende che hanno una **visione parziale** del mercato e che non sono in grado di **eseguire** in modo efficace.



Figura 1.4: Magic Quadrant di Gartner. Source: (Maggio 2021) Gartner

1.4 Long Tail di Anderson

La **Long Tail** è un modello sviluppato da **Chris Anderson** che rappresenta la **distribuzione delle vendite** di un prodotto (non per forza informatico o di nuova tecnologia). Questo modello descrive come la vendita di una certa categoria di prodotti è distribuita tra i prodotti più venduti e i prodotti meno venduti. L'effettiva “**coda lunga**” è la parte della distribuzione delle vendite che si trova dopo i prodotti più venduti. Questa parte della distribuzione è composta da molti prodotti che vendono poche copie ciascuno.

Conseguenze Questo modello ha diverse conseguenze:

- **Aumento della varietà** di prodotti disponibili sul mercato.
- **Aumento della disponibilità** di prodotti di nicchia.

- **Aumento della vendita** di prodotti di nicchia.
- **Aumento della vendita** di prodotti meno popolari.
- **Aumento della vendita** di prodotti meno venduti.

Conseguenza fondamentale riguarda soprattutto i motori online, in quanto questi possono permettersi di vendere prodotti di nicchia in quanto non hanno i costi di un negozio fisico.

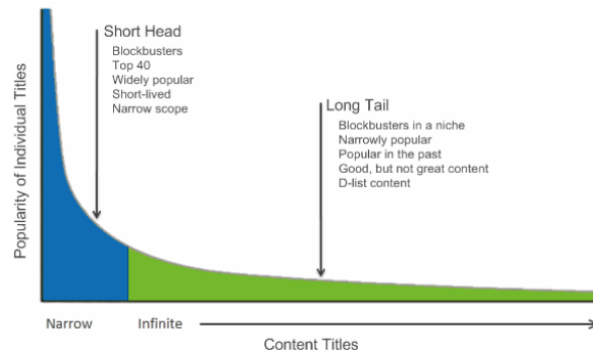


Figura 1.5: Long Tail

Capitolo 2

Concetti Generali sull'Informatica Aziendale

Nel presente capitolo verranno trattati i concetti generali dell'informatica aziendale, in particolare verranno trattati i seguenti argomenti: definizioni di informatica aziendale, di sistema informativo aziendale ed altro, verrà inoltre trattato l'impatto dell'informatica nelle aziende e come le aziende italiane si stanno adattando a queste nuove tecnologie.

2.1 Introduzione e definizioni

Definiamo innanzitutto alcuni concetti chiave per la comprensione del corso.

Informatica Aziendale L'**informatica aziendale** è la disciplina che studia l'applicazione dell'informatica nelle aziende, studia inoltre l'influenza di questa nelle diverse categorie di un sistema aziendale. Esistono diversi settori di applicazione tra i quali:

- **Aiuto e guida operativa** - Assistenza agli operatori a seguire le corrette procedure di lavoro con un costante controllo iterativo sui dati. Facilitazione di ricerca e recupero di informazioni.
- **Organizzativa** - Automazione di processi da un lato, richiesta di **competenze e risorse** differenti dall'altro.
- **Controllo** - Rilevazione di caratteristiche e comportamenti di un sistema, possibilità di **analisi quantitative e qualitative**.
- **Strategia** - Supporto ai processi di trasformazione e innovazione, supporto alle decisioni strategiche.

Sistemi Informativi aziendali I **sistemi informativi aziendali** sono l'insieme delle procedure e delle infrastrutture che definiscono e supportano l'elaborazione, la distribuzione e l'utilizzo delle informazioni all'interno di una azienda. Molto spesso ci si basa su una infrastruttura elettronica. È importante non confondere i sistemi informativi con i sistemi informatici, infatti è vero che ogni sistema informatico è un sistema informativo, ma non è vero il contrario.

Risorsa Una **risorsa** "è tutto ciò con cui l'organizzazione opera" sia che questo possa essere un bene fisico o che questo sia un bene immateriale

Processo Un **processo** è un insieme di attività atte a gestire una risorsa nel suo ciclo di vita.

2.2 Sistema informativo aziendale

Un **sistema informativo aziendale** è un sistema che permette di raccogliere, elaborare, memorizzare e distribuire informazioni all'interno di un'organizzazione. Questo sistema si compone di:

- **Dati:**

- di Configurazione - Dati che descrivono la struttura dell'organizzazione
- operativi - Dati che descrivono le attività dell'organizzazione
- di supporto - Dati che supportano le attività dell'organizzazione
- di stato - Dati che descrivono lo stato dell'organizzazione

- **Procedure:**

- acquisizione - Raccolta di dati
- controllo ed elaborazione - Controllo e manipolazione dei dati
- pianificazione

- **Mezzi e strumenti:**

- Hardware - server e periferiche
- Stazioni di lavoro
- ...

Da notare come in questa definizione non si parli di software, ma di mezzi e strumenti, questo perché il sistema informatico può essere o meno una parte del sistema informativo aziendale, questo perché il sistema informativo aziendale può consistere in un sistema informatico, ma non è detto che debba essere così. Questo rende di fatto il sistema informatico un “sottoinsieme” del sistema informativo aziendale.

2.3 Impatto dell'informatica nelle aziende

L'informatica ha avuto un impatto molto forte nelle aziende, infatti ha portato a una serie di cambiamenti che hanno rivoluzionato il modo di lavorare delle aziende.

2.3.1 Conoscenza dei fenomeni aziendali

Ogni sistema **informativo** aziendale è lo strumento per diffondere la conoscenza all'interno dell'azienda. Per adempiere al suo compito questo deve rispettare alcuni criteri, che sono poi gli stessi che permettono di dividere i fenomeni aziendali:

Livello di astrazione Il sistema deve essere in grado di rappresentare la realtà aziendale in modo corretto, sintetico ma completo. In alcuni livelli più “alti” si devono avere informazioni più sintetiche, mentre in livelli più “bassi” si devono avere informazioni più dettagliate¹.

Tempestività Il sistema deve essere in grado di fornire le informazioni in tempo utile ed appropriato al contesto dell'operazione e della mole di dati. Anche in questo caso si ha una differenziazione tra i vari livelli, infatti in livelli più “alti” le informazioni possono essere meno tempestive con attese più lunghe, mentre in livelli più “bassi” le informazioni devono disponibili immediatamente.

Livello di copertura Il sistema deve essere in grado di coprire tutte le aree aziendali e tutti i processi aziendali nei vari livelli di dettaglio. Questo racchiude entrambi i concetti di **orizzontalità** e **verticalità** del flusso informativo.

Allo stesso tempo il sistema informativo deve **garantire**: Accessibilità dei dati e Correttezza del flusso, flusso che si divide in:

Orizzontale tra le varie aree aziendali

Verticale tra i vari livelli gerarchici

¹Più avanti nella sezione 3.1 - Schema di Anthony si approfondirà questo concetto

2.3.2 Processi classi di un sistema informativo

Tra tutte le attività di una azienda possiamo distinguere tre processi i quali sono solitamente i primi ad essere informatizzati per la loro elevata “attrattiva informatica”. Questi processi sono:

Sviluppo funzioni operative Processo che si occupa di automatizzare dei processi che sono già presenti andando a ridurre i tempi e la mano d’opera necessaria.

Pianificazione Processo che prende i dati inseriti nel SI e li elabora per automatizzare processi di pianificazione.

Controllo Processo che renda automatico il controllo dei dati i inseriti nel SI e li confronta con criteri e dati di riferimento segnalando eventuali anomalie.

2.3.3 Nuovi processi

Introduzione dell’informatica L’introduzione dell’informatica in azienda non si occupa semplicemente di supporto a processi già esistenti, come nel caso dei processi classici, ma introduce nuovi processi e ne modifica altri. Questi processi particolarmente informatizzati sono impossibili o molto difficili da realizzare senza l’ausilio di un sistema informatico.

BRP Nasce da questa idea il concetto di **Business Process Re-engineering** o **Reingegnerizzazione dei processi aziendali** che consiste nel ripensare e ridisegnare i processi aziendali per sfruttare al meglio le nuove tecnologie informatiche.

Contatto col cliente al tempo di internet La spinta verso il processo è generata dalla vasta adozione delle reti informatiche come supporto alla comunicazione e alla collaborazione tra le persone. Con l’avvento di internet, infatti, il contatto con il cliente assume delle modalità completamente nuove, si passa da un contatto diretto a un contatto mediato da un sistema informatico, che per certi versi può essere più efficiente e più efficace. Questa trasformazione ha portato ad una enorme riduzione delle tempistiche di contatto e di risposta, ma ha anche portato ad una interazione diretta tra cliente e sistema informativo, il quale è in grado di fornire al cliente informazioni in tempo reale e di rispondere alle sue richieste in modo automatico.

2.4 I SI nelle aziende Italiane e relazione ICT - azienda

Analizziamo ora come i sistemi informativi si sono evoluti per adattarsi alle esigenze delle aziende italiane ed in che modo i servizi ICT hanno influenzato l’organizzazione interna ed esterna delle aziende.

2.4.1 Le aziende in italia

Le aziende in italia Le aziende in Italia assumono una conformazione molto differente rispetto al panorama europeo, infatti il 99,9% delle aziende italiane sono **PMI** (Piccole e Medie Imprese) e solo lo 0,1% sono grandi aziende.

PMI e SI Le PMI sono aziende che hanno una struttura molto semplice e che spesso l’investire in un sistema informativo non è una priorità visto che i processi sono molto semplici e non richiedono un sistema informativo complesso.

Spesso quindi un SI potrebbe essere visto come un costo inutile, ma con l’avvento di internet e delle nuove tecnologie, anche le PMI stanno iniziando ad adottare un sistema informativo in piccola misura, ovviamente non adotteranno SI di grandi dimensioni, ma sistemi informativi, spesso italiani in quanto più vicini alla realtà delle PMI, più piccoli e adeguati alle loro esigenze.

Evoluzione dei SI I sistemi informativi delle grandi aziende che in un primo luogo, come già discusso, erano molto distanti dalle PMI, si sono evoluti e si sono adattati alle esigenze delle PMI creando linee di SI adattati alla struttura flessibile e semplice delle PMI. Questo ha portato ad una maggiore diffusione dei sistemi informativi anche nelle PMI e ad una maggiore diffusione delle nuove tecnologie informatiche.

2.4.2 Cambiamenti dei SI

I sistemi informativi aziendali stanno entrando molto più facilmente all'interno delle aziende, questo è dovuto ad una evoluzione del mondo ICT il quale ha favorito lo sviluppo di sistemi informativi più flessibili e adattabili alle esigenze delle aziende.

Oltre a questo lo sviluppo dei servizi ICT hanno portato a cambiamenti nell'organizzazione interna ed esterna delle aziende.

Organizzazione Interna L'evoluzione dei servizi ICT ha portato alla riduzione dei ruoli impiegatizi, ovvero ruoli che non sono direttamente legati alla produzione, ma che sono necessari per il funzionamento dell'azienda. Oltre alla necessaria riqualificazione dei ruoli aziendali ed alla riduzione dei ruoli di supporto alla produzione (controllo, amministrazione, ecc...). Il processo di *front-office*, quali la vendita e il marketing, sono stati completamente rivoluzionati. Tutto ciò ha scaturito una revisione del modello organizzativo aziendale passando da una organizzazione "per funzioni" ad una organizzazione "per processi".

Organizzazione Esterna L'evoluzione dei servizi ICT ha portato ad una maggiore collaborazione tra aziende, infatti la rete ha permesso di creare nuove forme di collaborazione tra aziende, come ad esempio la semplificazione del processo di *outsourcing* e la creazione di nuove forme di collaborazione tra aziende. Questo accade al discapito della dimensione dell'azienda che non è più un fattore determinante per il successo dell'azienda stessa.

Capitolo 3

La struttura dell'azienda e del suo sistema informativo

In questo capitolo verrà analizzato il concetto di **esigenza informativa** e verranno trattati i sistemi operazionali e informativi.

3.1 Concetto di esigenza informativa

Funzione SI La funzione primaria del **sistema informativo** è quella di aiutare e guidare chi svolge mansioni che mandano avanti l'azienda attraverso queste. Inoltre il **SI** deve essere di aiuto e guida in modo diverso per aree diverse, ciò tramite il **livello d'astrazione** che sale man mano che si sale di livello gerarchico. L'**esigenza informativa** dipende dal tipo di attività svolta e dal livello gerarchico dell'utente. (es. i livelli operativi hanno bisogno di informazioni attuali ed precise spesso il singolo dato, mentre i livelli direzionali hanno bisogno di informazioni sintetizzate anche su periodi più lunghi).

Schema di Anthony

L'organizzazione aziendale è vista a forma piramidale con i livelli operativi alla base, i livelli intermedi al centro e i livelli direzionali in cima. Ogni livello ha bisogno di informazioni diverse e quindi il **SI** deve essere in grado di fornire informazioni adeguate a ciascun livello.



Figura 3.1: Schema di Anthony

Profili Informativi

Di seguito è riportata una tabella con i profili informativi di ciascun livello, si può notare come i livelli operativi abbiano bisogno di poche informazioni ma molto dettagliate, precise e in modo continuo, il livello direzionale tattico ha accesso ai dati con frequenza minore ma prefissata e con un livello di dettaglio minore produce quindi un volume medio di informazioni, il livello strategico ha bisogno di

informazioni molto sintetizzate e con una frequenza molto bassa se non sporadica ma ha bisogno anche di informazioni esterne all'azienda.

	Frequenza	Dati	Provenienza dati	Volume
Livello direzionale strategico	Sporadica	molto sintetizzati	interni ed esterni	basso
Livello direzionale tattico	Prefissata	sintetici e analitici	interni	medio
Livello operativo	Continua	analitici	interni	elevati

3.2 Scomposizione dei sistemi informativi

Le diverse esigenze informative, dei vari livelli gerarchici, hanno portato ad una separazione delle funzioni decisionali e operative in due tipi di sistemi informativi: quelli orientati al supporto operativo e quelli orientati alle decisioni. Spesso i sistemi orientati al supporto operativo hanno delle funzioni di *reporting* che si avvicinano a quelle dei sistemi orientati alle decisioni ma vista la necessità di rapidità e precisione delle informazioni per i livelli decisionali e tattici è sorta la necessità di sviluppare sistemi informativi direzionali.

3.2.1 Sistemi operazionali

I sistemi operazionali sono sistemi che permettono di raccogliere, elaborare e presentare informazioni utili per il supporto delle attività operative dell'azienda. Questi sistemi sono orientati a supportare le attività di esecuzione e controllo delle operazioni aziendali e dunque sono usati dai livelli operazionali e dalla direzione funzionale. *On Line Transaction Processing (OLTP)* è un termine usato per descrivere un'attività di gestione dei dati che permette di svolgere attività operative. Questi sistemi permettono di gestire i dati in modo operativo e di automatizzare le attività aziendali.

Funzioni principali

- Automazione di attività procedurali - In questo caso il SI è un supporto all'operatore
- Definizione di nuovi processi - come visto sottosezione 2.3.3
- Aiuto nelle attività aziendali
- Raccolta di dati - gli operatori inseriscono i dati nel sistema in modo continuo
- Guida per l'operatore - il sistema guida l'operatore nelle attività da svolgere in questo modo si riducono gli errori

Azioni sui dati

- Accesso interattivo in inserimento, lettura, modifica - l'operatore può interagire con il sistema e modificare i dati nei limiti imposti
- Trattamento di dati - il sistema tratta i dati in modo automatico e li presenta all'operatore in modo chiaro
- Descrizione di eventi - il sistema descrive le transazioni e le attività svolte in modo da poterle ripetere in caso di necessità
- Valutazione e trattamento di informazioni utili - il sistema valuta i dati se sussistono errori e li segnala all'operatore
- Aggregazione per il calcolo di indicatori di stato - il sistema aggrega i dati per calcolare indicatori di stato dell'azienda

Componenti fondamentali

- Base di dati operativa - contiene i dati operativi dell'azienda
- Funzioni operative - funzioni che permettono di svolgere le attività operative

3.2.2 Sistemi informativi

I sistemi informativi (o direzionali) sono sistemi che permettono di raccogliere, elaborare e presentare informazioni utili per il processo decisionale. Questi sistemi sono orientati a supportare le attività di controllo e di pianificazione dell'azienda e dunque sono usati dai livelli dell'alta direzione e dalla direzione funzionale. *On Line Analytical Processing (OLAP)* è un termine usato per descrivere un'attività di analisi dei dati che permette di estrarre informazioni utili per il processo decisionale. Questi sistemi permettono di analizzare i dati in modo multidimensionale e di presentarli in modo chiaro e comprensibile.

Funzioni principali

- Facilitazione del processo decisionale
- Presentazione dei dati secondo diverse aggregazioni e viste
- Confronto tra indicatori aziendali e indicatori esterni

Azioni sui dati

- Accesso in lettura
- Aggregazione dei dati
- Descrizione di aree/temi
- Profondità temporale
- Multi-dimensionalità

Componenti fondamentali

- Base dati informativa
- Strumenti di analisi
- Procedure di alimentazione (dati)

3.2.3 Comparazione

	Operazionale (OLTP)	Informativa (OLAP)
Finalità	Supporto all'operatività	Supporto alle decisioni
Utenti	Molti, operativi	Pochi, direzionali
Dati	Dettagliati, transazionali	Sintetici, aggregati
Modo d'uso	Guidata (Operazioni predeterminate)	Interrogazioni ad hoc
Quantità dati per attività elementare	Pochi record	Molti record
Orientamento	Per processo	Per area
Frequenza aggiornamento	Continuo	Periodica se non sporadica
Copertura temporale	Breve	Lunga
Ottimizzazione	Per accessi in scrittura	Per accessi in lettura

Capitolo 4

Le scelte organizzative

Andiamo ora ad analizzare come il sistema informativo possa essere costruito e gestito all'interno di un'azienda. In particolare andremo ad analizzare le scelte che un'azienda può fare riguardo alla costruzione e gestione del sistema informativo, quali sono le figure professionali che si occupano di questo sulla base della dimensione dell'azienda. Andiamo inoltre ad affrontare come il sistema informativo si posiziona nell'organigramma aziendale e come l'infrastruttura tecnologica può essere organizzata e gestita.

4.1 Costruzione del SI

Al momento quando si parla di adottare un nuovo SI si possono fare tre scelte principali:

Make costruire il SI internamente

Buy acquistare il SI da uno o più fornitori esterni

Outsource far gestire ad una azienda esterna il SI

4.1.1 Opzione *make*

Con l'opzione **make** ovvero di **costruzione interna** si intende la costruzione del SI all'interno dell'azienda da parte di un gruppo di lavoro incaricato della realizzazione, manutenzione e gestione del SI. Questa opzione è raramente scelta e solitamente si limita a funzioni marginali rispetto ad un SI completo. Il tutto però rispecchia in pieno i modelli organizzativi e si ha un controllo totale del SI.

Vantaggi & Svantaggi

Questa opzione comporta dei costi fissi molto elevati usati sia per il personale che è incaricato di sviluppare e mantenere l'SI che per l'infrastruttura sulla quale l'SI è installato. Inoltre quando vi è necessità di un aggiornamento importante bisogna investire molte risorse economiche. Questa soluzione ha lo svantaggio di non confrontarsi con il mercato attuale e che dunque alcune funzionalità non sono le più efficienti o efficaci. Il primo vantaggio lo si riscontra nella situazione nella quale si dovessero verificare dei problemi, allora i tempi di risoluzione saranno molto brevi per questioni banali, ma lunghi per difficoltà più complesse. Altro vantaggio importante di questa soluzione è il mantenimento interno del “*know-how*”.

4.1.2 Opzione *buy*

L'opzione *buy* consiste nell'acquisto di un SI da parte di uno o più fornitori esterni, alla quale si aggiunge l'instaurazione di un piccolo gruppo di lavoro interno per la gestione del SI e per la gestione dei rapporti con i fornitori del SI. Questa è una scelta tipica nell'economia italiana delle PMI.

Vantaggi & Svantaggi

Questa opzione favorisce una struttura interna umana e tecnologica di molto ridotta rispetto all'Opzione *make* insieme ai costi fissi. L'azienda si concentra sul proprio core business senza perdere tempo e risorse nella gestione del SI, questo al prezzo di una stretta dipendenza dai fornitori esterni ed una perdita parziale del *know-how* aziendale oltre alla mancata proprietà del *software*. Come ulteriori vantaggi si

ha la maggior flessibilità del SI e il confronto con il mercato attuale. Il tutto però con un modello organizzativo non su misura per l'azienda e la difficoltà nel passaggio da un fornitore all'altro.

4.1.3 Opzione *outsource*

Con l'opzione di **outsourcing** ovvero di **esternalizzazione** si delega ad una terza parte la gestione e l'organizzazione del SI dopo pagamento di canone periodico.

Questa opzione al giorno d'oggi sta prendendo piede tra le PMI in quanto permette, seppur ad un costo variabile elevato, di avere un SI completo e funzionante senza dover investire in *hardware* e *software* e senza dover adibire personale interno alla gestione del SI.

Hosting

Con questa opzione si affida solo la parte di **infrastruttura** tecnologica, non il *software* e altri servizi che solitamente vengono gestiti in ottica *buy*. Con questa opzione si può noleggiare un server fisico o una macchina virtuale.

Body rental

Con questo termine intendiamo l'uso di personale specialistico di una azienda esterna per trasformare costi fissi in costi variabili.

Vantaggi & Svantaggi

Nell'opzione *outsource* i costi sono variabili ma abbastanza alti, inoltre in caso di necessità si può aumentare o diminuire le risorse in base alle esigenze. Si è però vincolati al fornitore della soluzione utilizzata. È presente inoltre una maggiore flessibilità rispetto all'Opzione *make*, come al livello dell'Opzione *buy*. Questa opzione però non permette di avere alcun *know-how* interno e quindi non si ha la possibilità di personalizzare la soluzione in base alle proprie esigenze. Inoltre si ha una maggiore dipendenza dal fornitore del servizio. Possiamo dunque definire l'opzione *outsource* come una Opzione *buy* con costi variabili e senza controllo su nessun aspetto del SI.

4.2 Le figure professionali

Maturità Informatica La **maturità informatica** delle aziende è un indicatore rispetto alla diversa organizzazione e alla collocazione nell'organigramma aziendale del reparto IT, ovvero il reparto che si occupa della gestione del SI.

4.2.1 Sviluppo del reparto

Livello 1

Spesso le aziende a questo livello sono alle fasi iniziali dell'automazione. Il *team* è composto in modo da pochi elementi, spesso non specializzati ma altamente flessibili e con competenze trasversali. Questo *team* ha una struttura completamente **orizzontale** e non vi è una figura di riferimento. Può capitare che i componenti del *team* siano anche responsabili di altre attività aziendali e che la parte informatica sia solo una parte del loro lavoro.

Livello 2

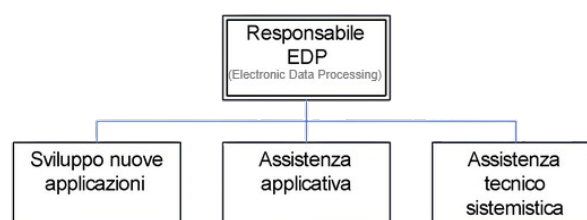


Figura 4.1: Schema di organizzazione del reparto IT a livello 2

In questo livello si inizia ad osservare una organizzazione gerarchica dove il **responsabile EDP** (Electronic Data Processing) è responsabile del reparto IT, sotto di lui ci sono i **sistemisti** che si occupano della gestione e dell'assistenza sull'infrastruttura. Poi ci sono gli **analisti** che supportano gli utenti e analizzano i requisiti. Infine ci sono i **programmatore** che si occupano dello sviluppo del software (non presenti nell'organizzazione *buy*)

Livello 3

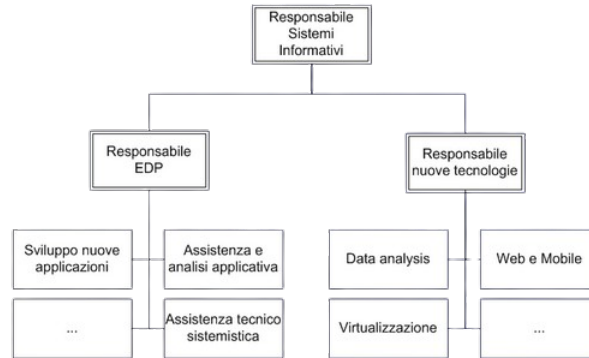


Figura 4.2: Schema di organizzazione del reparto IT a livello 3

In questo livello è presente una vera e propria Direzione, sotto questa sono presenti i reparti di EDP e il reparto per la ricerca su nuove tecnologie. Come mostrato nella figura il primo si occupa di sviluppo di nuove applicazioni, assistenza e analisi e assistenza tecnico-sistemistica mentre nel secondo ci si occupa di Analisi dei dati, ricerca web e mobile e virtualizzazione della infrastruttura.

Livello 4

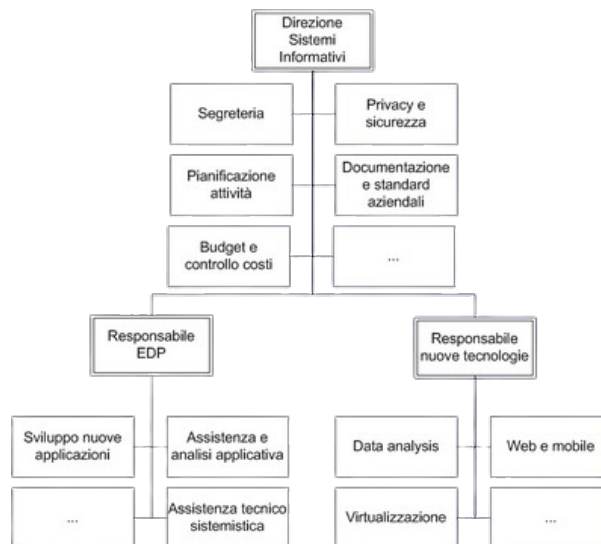


Figura 4.3: Schema di organizzazione del reparto IT a livello 4

Livello nel quale è presente un vero e proprio **dirigente del sistema informativo** con una organizzazione gerarchica molto più complessa rispetto ai livelli precedenti. È presente una divisione **EDP** con relativo responsabile che è isolata dal reparto **nuove tecnologie** che si occupa di sviluppare nuove tecnologie e di supportare il reparto **EDP**. Inoltre sopra di questi è presente una sezione comune del sistema informativo che si occupa di coordinare i due reparti e gestire anche **privacy** e **sicurezza** o anche **pianificazione attività**. Questo livello dispone di un vero e proprio **budget** per il reparto IT.

4.2.2 Posizione all'interno dell'organigramma

Avendo definito il livello di maturità del reparto IT come struttura interna andiamo ora a valutare come il reparto stesso si posiziona all'interno dell'azienda, anche questo aspetto contribuisce a definire il livello di maturità dell'azienda stessa rispetto all'informatica.

Supporto amministrativo

In questo caso il reparto IT è visto come un supporto amministrativo, questa è una visione obsoleta

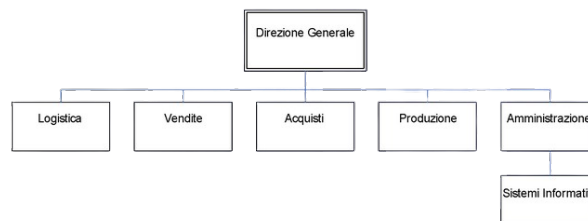


Figura 4.4: Posizione del reparto IT come supporto amministrativo

Servizio ad altre direzioni generali

In questo caso il reparto SI è o al pari degli altri reparti o a supporto della direzione generale. Se il reparto IT è al pari degli altri reparti allora questo è a supporto di tutti i reparti dell'azienda, se invece è a supporto della DG allora questo si posiziona più in alto rispetto agli altri reparti.

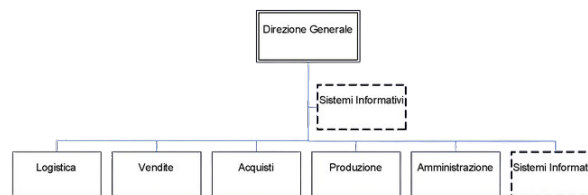


Figura 4.5: Posizione del reparto IT come supporto a tutte le direzioni

Organizzazione autonoma

In questo caso è vero che il reparto SI viene messo alle dipendenze del reparto **organizzazione** ma è anche vero che in questo modello il reparto SI è autonomo. Il suo ruolo è quello di supportare tutte le altre direzioni e di coordinare le varie aree operative dell'azienda, solitamente questo è il modello delle grandi aziende.

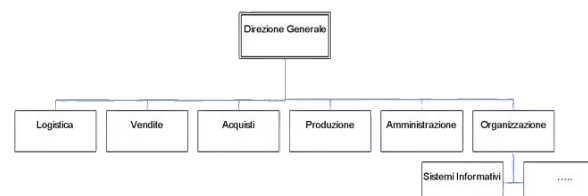


Figura 4.6: Posizione del reparto IT come organizzazione autonoma

4.3 Infrastruttura Tecnologica

Negli anni l'infrastruttura tecnologica si è evoluta siamo passati da server con vari terminali connessi fino alla virtualizzazione a archiviazione in cloud, il tutto passando dall'architettura client-server. Il SI deve tenersi sempre aggiornato con le tecnologie in uso.

Il passaggio da server locali a soluzioni cloud Recentemente molte aziende stanno eseguendo il passaggio da una infrastruttura locale ad una in cloud, in questo modo sussistono meno investimenti lato *hardware* e *software* che ai giorni nostri invecchiano ancora prima di essere ammortizzati. Tutto il lavoro di gestione dell'infrastruttura è spostato su una azienda esterna senza che venga a mancare l'accessibilità degli strumenti informativi aziendali. In questo modo però si è **dipendenti dalla connettività** ovvero il lavoro non può essere svolto se non si è connessi a internet.

4.4 Interrompibilità del servizio informatico

In ogni caso bisogna sempre valutare che danno può causare l'interruzione del servizio informatico. In alcuni casi l'interruzione di esso può portare al “blocco totale” del lavoro, questo rischio viene spesso sottovalutato dai *top manger*. Come conseguenza è importante garantire la “continuità operativa” ovvero la riduzione o il completo annullamento che un “blocco totale” possa avvenire.

Problematiche legate all'hardware

L'infrastruttura dei SI è soggetta a guasti, bisogna dunque prevenire questi andando a implementare il concetto di “*sistema ridondato*”¹ (esempio per i dischi di archiviazione: **RAID**).

Hot Swap Importante per i supporti di archiviazione è anche il concetto di *Hot Swap* che consiste nella possibilità di sostituzione di questi senza dover spegnere il sistema e quindi dover interrompere l'operabilità del SI.

Failure tolerancy L'ideale infrastruttura di un SI dovrebbe essere *fault tolerant*, ovvero non deve esistere un oggetto² che in caso di fallimento comporta al fallimento di tutto il sistema, ciò per quelle parti del sistema che sono ritenute essenziali.

Backup In ogni ambiente si rende necessaria l'implementazione di una strategia di *backup*, queste copie devono essere conservate in ambienti protetti (casseforti ignifughe, locazioni remote...). Una teoria suggerisce che bisogna avere sempre a disposizione tre copie dei dati su tre apparati diversi: uso, *backup* interno, *backup* remoto.

Problematiche legate al software

Maggiormente le problematiche di un SI sono legate al *software*. I problemi legati a consistono nella “indimostrabilità” che un qualsiasi programma cerchi di computare un risultato senza “ciclare” all'infinito. D'altra parte questo genere di malfunzionamenti raramente comporta un blocco di tutto il SI ma più spesso interferiscono con un processo specifico o con un gruppo di attività.

Problematiche legate ad azioni dolose

Molto più frequenti rispetto agli altri generi di criticità sono quelle problematiche legate ad azioni intente a ledere l'operabilità o la segretezza dei dati. Molto spesso questo genere di problematiche è causato da vulnerabilità dell'SI o dell'*hardware*. La causa principale però rimane l'errore umano che può essere solamente mitigato.³

¹La ridondanza prevede che in caso di fallimento di un singolo sistema ne esista un'altro che possa sostituirsi ad esso andando a “coprire” il punto difettoso

²server, apparato di rete, griglia elettrica, etc...

³Più informazioni sulla sicurezza informatica in “Appunti di Introduction to Computer and Network Security” di Luca Facchini, capitoli 1/2

Capitolo 5

I sistemi operazionali

In questo capitolo verranno esposti concetti base sui “sistemi operazionali” parlando dunque delle finalità di questi, dell’organizzazione della informazione operativa, della potenzialità informatiche di una azienda ed infine degli accenni alla composizione di un sistema operativo.

5.1 Finalità dei sistemi operazionali

Le principali finalità dei **sistemi operazionali** riguardano:

Registrazione delle transazioni Il processo di acquisizione e memorizzazione delle informazioni relative alle transazioni aziendali.

Pianificazione e controllo La possibilità di pianificare le operazioni aziendali e controllarne l’effettiva esecuzione.

Acquisizione ed organizzazione della conoscenza La possibilità di acquisire e organizzare la conoscenza aziendale.

Elaborazione delle situazioni aziendali La possibilità di elaborare le informazioni aziendali per ottenere una visione complessiva della situazione aziendale.

Per raggiungere questa finalità il sistema operativo si compone di due sottosistemi principali:

Base di dati operativa Contiene le informazioni operative in forma organizzata.

Funzioni operative Sono le funzioni che permettono di acquisire, memorizzare, elaborare e trasmettere le informazioni.

5.1.1 Transazioni - Definizione e registrazione

Cos’è una transazione

Definizione Per definizione una **transazione** è una operazione detta **atomica** (ovvero indivisibile) che si manifesta in un certo e conosciuto momento ed è una informazione che l’azienda è interessata a registrare.

Esempi Alcuni esempi di transazioni sono: gli ordini tra cliente e fornitori, prelievi da magazzino, spedizioni, pagamenti, ecc. . .

Registrazione delle transazioni

Le transazioni da dover registrare, possono essere sostanzialmente di due tipi:

Semplici Si deve registrare nel sistema solo un singolo dato.

Esempio di registrazione di transazione semplice è la registrazione di una movimentazione del magazzino

Complesse Si devono registrare più operazioni elementari connesse in senso logico e spesso correlate a documenti fisici, quali ad esempio una spedizione che è correlata ad una *bolla di spedizione*.

Esempio di registrazione di transazione complessa è la registrazione di una spedizione dove sono coinvolte più operazioni elementari quali la raccolta degli attributi (destinatario, prodotti, data e ora...) ed genera una bolla di spedizione.

Inoltre una transazione può generarle delle altre e quindi si parla di **transazioni a cascata**.

Volume dei dati Ogni transazione produce un volume di dati dipendentemente dalla natura dell'attività e dell'organizzazione aziendale.

5.1.2 Pianificazione e controllo delle operazioni

Alcuni processi aziendali sono dipendenti da altri, si rende quindi necessario usare i dati dei processi "a monte" per pianificare e controllare i processi "a valle". Tramite l'uso di SI è possibile adottare modelli più complessi di pianificazione e monitorare continuamente l'andamento dello stato dei processi aziendali.

Perché pianificare e controllare Pianificare e controllare i processi aziendali ha diversi vantaggi per l'azienda, sia per il passato che per il presente fino ad avere anche una utilità per i processi futuri. Questi vantaggi sono raggiunti tramite: La possibilità di elaborare piani e strategie di produzione, registrare e monitorare l'avanzamento delle operazioni ed infine la possibilità di misurare se e quanto i piani sono stati rispettati rispetto agli obiettivi prefissati.

Come pianificare e controllare Il SI deve essere dotato di funzioni molto articolate e specifiche per l'azienda alla quale si riferisce, ad esempio quando parliamo di "Elaborazione di piani" il SI di riferimento deve essere in grado di: ottimizzare le risorse disponibili, sincronizzare le operazioni ed essere coerente con lo stato degli indicatori aziendali.

5.1.3 Organizzazione della conoscenza aziendale

La conoscenza aziendale (*Knowledge Base* - KB) è l'insieme delle informazioni a supporto dell'attività di produzione aziendale, come ad esempio le informazioni sui prodotti, sui processi, sui clienti, sui fornitori, ecc... Queste informazioni sono essenziali per la gestione aziendale e per la pianificazione delle attività future, avere dunque all'interno del SI operativo la loro versione più aggiornata sempre a disposizione è fondamentale per l'azienda.

Queste informazioni devono essere strutturate, ovvero riconducibili ad un insieme di caratteristiche predefinite, e devono essere correlate, ovvero devono essere collegate ad articoli, clienti, fornitori, ecc...

5.1.4 Elaborazione delle situazioni aziendali

Il SI è un sistema dinamico che serve per modellare la realtà aziendale e per fornire informazioni utili per la gestione aziendale. La conoscenza dello stato corrente, oltre che di quello passato, è fondamentale per la gestione aziendale, questa conoscenza permette di pilotare l'azienda grazie a determinati eventi. Alcuni indicatori di stato sono ad esempio: le giacenze di magazzino, i tempi di consegna, i tempi di produzione, ecc...

Gli indicatori dunque non rappresentano una situazione statica, ma una situazione dinamica che cambia nel tempo. Questi indicatori sono utili per la gestione aziendale e per la pianificazione delle attività future. Tutti gli indicatori di stato sono calcolati a partire dai dati inseriti, modificati e cancellati dalle transazioni aziendali e sono utili per la gestione aziendale.

5.2 Informazione Operativa

L'informazione operativa è costituita principalmente da archivi nei quali sono presenti relazioni che coinvolgono diverse entità, questi archivi solitamente li classifichiamo in:

Movimenti Contengono le informazioni relative alle transazioni semplici, relative ad un singolo oggetto.

Documenti Contengono le informazioni su transazioni complesse che riguardano una lista di oggetti (classica tabella) dove in testa troviamo anche una serie di informazioni comune a tutte le righe.

Informazioni di stato Ovvero un insieme di indicatori di stato che permettono di avere una visione complessiva della situazione aziendale. Questi possono essere *de-materializzati* e quindi calcolati al momento della richiesta o *materializzati* e quindi calcolati e memorizzati in un archivio.

Informazioni Anagrafiche Contengono le informazioni relative alle entità che partecipano alle transazioni, questi non si limitano a contenere solo dati di anagrafica di persone fisiche, ma anche di oggetti, di entità giuridiche, ecc. . .

5.2.1 Qualità dei dati

Per qualità dei dati si fa riferimento allo standard ISO 8402-1995:

Il possesso della totalità delle caratteristiche che portano al soddisfacimento delle esigenze espresse o implicite, dell'utente.

(*International Organization for Standardization* — ISO 8402-1995)

La qualità dei dati Per stabilire un indice di qualità dei dati si possono utilizzare diversi parametri quali:

- Tanto più elevata quanto più il sistema fornisce rappresentazioni degli eventi vicine alla percezione diretta della realtà
- La dipendenza dalla struttura del SI è minore quanto più i dati sono indipendenti dalla struttura del sistema
- La qualità è diminuita da sottosistemi non integrati e da dati ridondanti

In sostanza un dato per essere di qualità non deve essere ridondante, deve essere coerente con la realtà e deve essere indipendente dalla struttura del sistema.

Impatto della qualità dei dati Se all'interno del proprio SI si ha una bassa qualità dei dati, allora si avrà un forte impatto economico/organizzativo tra cui: la difficoltà nell'introduzione di innovazioni tecnologiche (adozione di una nuova tecnologia) e di processo (modificare un processo produttivo), la difficoltà nell'avvio di processi del tipo *data warehousing*, inoltre dal lato umano avere una bassa qualità dei dati può portare a una scarsa soddisfazione degli utenti finali del SI (ovvero quelle persone che utilizzano il SI per svolgere il proprio lavoro).

5.2.2 Caratteristiche strutturali

L'informazione operativa è l'informazione che serve per svolgere le attività operative dell'azienda, questa informazione è costituita da dati che vengono acquisiti, memorizzati, elaborati e trasferiti all'interno dell'azienda. Questi dati sono utili per la gestione aziendale e per la pianificazione delle attività future. L'informazione operativa è caratterizzata da:

Aggregazione I dati sono aggregati in base alle esigenze dell'utente, possono essere:

Analitici Se si vuole avere una visione dettagliata di un singolo evento.

Analitici Se si vuole avere una visione complessiva di un insieme di eventi, ottenuto aggregando i dati.

Tempificazione I dati possono essere temporizzati in base alle esigenze dell'utente, possono essere:

Puntuale Se si vuole avere una visione istantanea della situazione aziendale.

Cumulata Se si vuole avere una visione della situazione aziendale in un certo periodo di tempo.

Dimensionalità Intendiamo come dimensionalità il numero minimo di parametri necessari per estrarre una specifica informazione.

Esempio delle caratteristiche dell'informazione operativa:

	Aggregazione	Tempificazione	Dimensionalità
Anagrafe	Analitica	Puntuale	unitaria
Movimenti & Documenti	Analitica	Puntuale	Contenuta
Informazioni di stato	Analitica o aggregata	Puntuale o cumulata	Contenuta

5.2.3 Caratteristiche funzionali

I dati operativi oltre ad essere strutturati in maniera particolare, devono anche avere delle caratteristiche funzionali che permettano di svolgere le attività operative dell'azienda. Queste caratteristiche sono:

Completezza Estensione con cui i dati vengono registrati

Correttezza Quanto quel dato si avvicina alla realtà

Precisione Quanto quel dato è vicino alla realtà

Omogeneità Se tra tutti i dati della stessa natura sono rappresentati con una stessa struttura

Fruibilità Facilità con cui i dati possono essere reperiti, acquisiti e compresi dall'utente in relazione alle sue esigenze

1

5.3 Potenzialità informatica

Definizione 5.1 (Potenzialità informatica). *La **potenzialità informatica** è costituita principalmente da due indicatori:*

Intensità informativa *Misura la quantità di informazioni della quale l'azienda necessita, sia che queste vengano da fonti interne o esterne.*

Attrattiva informatica *Ovvero la propensione, sulla base dei processi aziendali, dell'azienda ad utilizzare un SI per la gestione delle informazioni.*

Inoltre la propensione del management verso l'investimento sulla infrastruttura informatica è un indicatore di potenzialità informatica.

5.3.1 Intensità informativa

Definizione 5.2 (Intensità informativa). *L'intensità informativa è costituita da un insieme di fattori che concorrono a determinare la quantità di informazioni di cui l'azienda necessita.*

Questi fattori sono:

- La complessità dell'attività aziendale: oltre alla dimensione dell'azienda va presa in considerazione l'area geografica l'eventuale appartenenza ad un "gruppo", il livello di diversificazione dei prodotti, dei mercati e delle tecnologie. Tutti questi fattori influenzano la quantità di informazioni necessarie.
- L'intensità informativa

Definizione 5.3 (Intensità informativa). *L'intensità informativa è costituita dal prodotto dell'intensità informativa di prodotto e dell'intensità informativa di processo.*

Definizione 5.4 (Intensità informativa di prodotto). *L'intensità informativa di prodotto è la quantità di informazioni necessarie per la progettazione, la produzione e la commercializzazione di un prodotto.*

Definizione 5.5 (Intensità informativa di processo). *L'intensità informativa di processo è la quantità di informazioni necessarie per l'avanzamento dei processi aziendali. Viene presa in considerazione anche la mole di dati che emergono dal processo e la complessità delle operazioni elementari previste dal processo.*

¹La sezione 4.3 del libro "Rappresentazione della realtà" è stata omessa in quanto trattata marginalmente a lezione ed approfondita al corso di *Basi di Dati*.

5.3.2 Attrattiva Informatica

L'intensità informativa non è sufficiente a determinare se l'adozione di un SI possa essere vantaggiosa per l'azienda, infatti è necessario anche valutare l'attrattiva informatica.

Definizione 5.6 (Attrattiva informatica). *L'attrattiva informatica è un indicatore che misura la propensione dell'azienda ad utilizzare un SI per la gestione delle informazioni. Questo indicatore è costituito dall'insieme dell'attrattiva informatica dei processi aziendali.*

Definizione 5.7 (Attrattiva informatica di processo). *L'attrattiva informatica di processo è costituita dai seguenti fattori:*

Proceduralità *Grado di formalizzazione dei processi aziendali. Più un processo è formalizzato, più è attrattivo.*

Complessità *Grado di difficoltà delle azioni elementari previste dal processo. Meno un processo è complesso e più è attrattivo.*

Ripetitività *Frequenza con cui un processo viene ripetuto. Più un processo è ripetuto, più è attrattivo.*

Volume *Quantità di dati e informazioni elaborate dal processo. Alti volumi di dati rendono un processo attrattivo*

5.4 Composizione dei sistemi informativi operazionali

I SI operazionali sono composti da diversi sotto-sistemi che si occupano di diverse funzioni. Al momento non esiste una classificazione standard dei sotto-sistemi in quanto varia in base all'azienda e al settore di appartenenza. I criteri principali usati per la distinzione tra diversi sotto-sistemi sono: la funzione svolta, il processo aziendale coinvolto, l'architettura tecnologica, ecc. . .

Portafoglio Applicativo

Definiamo come **portafoglio applicativo** l'insieme delle applicazioni software che costituiscono il SI operativo, possono essere individuate due aree principali, ovvero il **portafoglio operativo** e il **portafoglio istituzionale**.

Portafoglio Operativo È costituito da applicazioni informatiche che trattano di processi legati al *core-business* dell'azienda. Questo genere di portafoglio è caratterizzato da una elevata specializzazione ad un settore specifico oltre ad una elevata variabilità tra aziende dello stesso settore di appartenenza. Inoltre questo portafoglio è caratterizzato da una forte verticalizzazione e da una elevata specializzazione delle funzioni implementate.

Portafoglio istituzionale È costituito da applicazioni informatiche che trattano di processi a sostegno delle principali attività aziendali. Questo genere di portafoglio è caratterizzato da una elevata attrattiva informatica e da una alta proceduralità (e ripetitività) dei processi. Inoltre questo genere di portafoglio è caratterizzato da una elevata omogeneità tra aziende anche di settori diversi e la poca variabilità tra servizi e prodotti offerti.

Sistema gestionale classico

Nel modello del sistema gestionale classico solitamente sono presenti delle isole informatiche autonome e molto specializzate, questo genere di sviluppo è causato principalmente da uno sviluppo incrementale (ovvero lo sviluppo viene eseguito a compartimenti "stagni" uno per volta), da una rigidità delle organizzazioni aziendali, dalla specializzazione dei produttori di software ecc. . .

Problematiche Questo genere di sviluppo porta a diverse problematiche quando si discute sulla gestione dei dati, in questo sistema infatti i dati sono ridondanti, disomogenei e spesso incoerenti. Ad aggiungersi a ciò questo genere di gestionale rende difficile avere una visione complessiva dell'azienda.

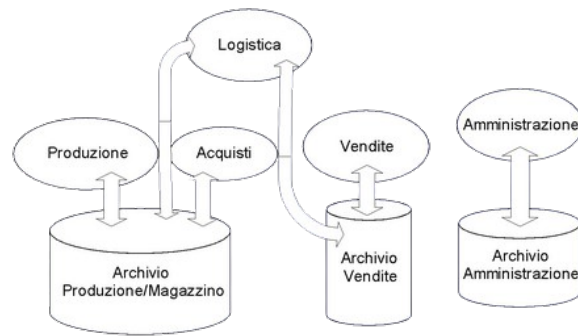


Figura 5.1: Schema dei settori di un sistema gestionale classico

Sistema ERP - *Enterprise Resource Planning*

Un sistema ERP è un sistema informativo aziendale integrato, ovvero un sistema che permette di gestire in maniera integrata e coordinata tutte le informazioni aziendali. Questo genere di sistema, grazie ad una basi di dati unica ed a processi integranti e cooperanti, punta a trattare i dati in modo ottimale e ottimizzato, oltre a gestire il controllo dei processi aziendali.

Vantaggi Questo genere di sistema così implementato è molto flessibile ed in grado di assecondare l'azienda in ogni sua esigenza, inoltre permette di avere una visione complessiva dell'azienda e di avere una gestione ottimale dei processi aziendali. Inoltre questo si adatta molto rispetto all'organizzazione aziendale e all'architettura tecnologica.

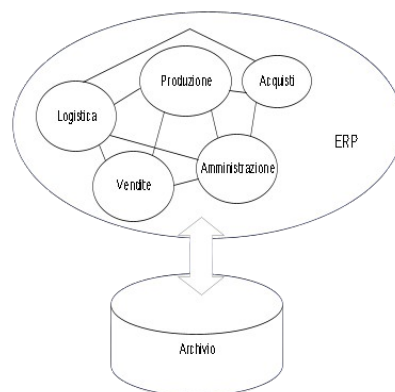


Figura 5.2: Schema dei settori di un sistema gestionale ERP

Ambiti applicativi

La presenza di moduli indipendenti presente negli ERP rendono questi sistemi molto flessibile anche a tipologie diverse di aziende, infatti un ERP può essere utilizzato in diversi ambiti applicativi, anche molto diversi tra loro. Tra questi ambiti troviamo: Servizi Finanziari, Produzione, Distribuzione, Commercio, Servizi, ecc. . .

Flussi di base I flussi di base di un ERP sono:

Amministrativo Flusso di prima applicazione che riguarda la gestione amministrativa dell'azienda.

Logistico Flusso che riguarda la gestione dei processi logistici dell'azienda.

Attivo (vendite) Flusso che riguarda la gestione delle vendite dell'azienda.

Passivo (acquisti) Flusso che riguarda la gestione degli acquisti dell'azienda.

Produttivo Uno tra i flussi più complessi, riguarda la gestione della produzione dell'azienda, questo può variare anche di molto da azienda ad azienda.

Sistemi operazionali complementari

I sistemi operazionali complementari sono sistemi che vengono aggiunti al sistema operativo principale per coprire delle funzioni che non sono presenti nel sistema operativo principale. Questi sistemi sono caratterizzati da una forte specializzazione e da una forte integrazione con il sistema operativo principale.

Esempi di sistemi operazionali complementari Alcuni esempi di sistemi operazionali complementari sono: i sistemi di supporto alle decisioni, i sistemi di gestione della qualità, i sistemi di gestione ambientale, le estensioni dell'ERP ecc. . .

Capitolo 6

ERP

In questo capitolo affrontiamo le varie aree di un ERP e le funzionalità che queste offrono.

6.1 Area Amministrativa

In questo capitolo si analizzeranno le funzionalità dell'area amministrativa di un sistema ERP. Questa area è fondamentale per la gestione delle risorse finanziarie e contabili dell'azienda. In particolare, si analizzeranno le seguenti funzionalità: Piano dei conti, gestione di anagrafiche e movimenti contabili, finanziari ed IVA.

La parte amministrativa di un SI si basa su delle strutture base, queste possono essere o per la gestione dell'anagrafica, o per la gestione dei movimenti contabili. In particolare nella parte anagrafica ritroviamo:

- Piano dei conti
- Clienti
- Fornitori
- Istituti bancari

Nella parte contabile invece troviamo:

- Movimenti contabili
- Movimenti finanziari
- Movimenti IVA

Questa separazione di vari aspetti ci permette di avere una visione più chiara e ordinata delle informazioni.

Piano dei conti Il piano dei conti è una struttura gerarchica che permette di classificare i conti contabili in modo ordinato. In questa struttura troviamo sia conti di attivo che di passivo, e conti di costo e ricavo. Solitamente questa è organizzata in tre livelli ma in moduli più avanzati si possono trovare anche più (o meno) livelli. Questa struttura permette di avere una visione chiara e ordinata delle informazioni contabili andando a raggruppare i conti in base alla loro natura. Il piano dei conti è rappresentato Solitamente tramite una tabella, i cui campi principali sono: Codice, Descrizione, Livello, Classe e Tipo. Il codice, anche questo organizzato in modo gerarchico permette di identificare in modo univoco il conto, es: 1.1.001 significa il conto 001 di livello 3, sotto il conto 1.1 di livello 2, sotto il conto 1 di livello 1. La descrizione è il nome del conto come "Crediti", "Acquisto merci", "Vendita merci" ecc... Il livello indica il livello gerarchico del conto, la classe indica se il conto è un mastro, un conto o un sotto/conto, infine il tipo indica se il conto è patrimoniale o economico.

Anagrafiche Le anagrafiche sono una raccolta delle informazioni riguardanti i clienti, i fornitori, ecc... Queste informazioni sono fondamentali per la gestione delle attività dell'azienda. In particolare possiamo raccogliere informazioni proprie del cliente quali il nome, la ragione sociale, il codice fiscale,

la partita IVA, ecc. . . queste non riguardano dei flussi in particolare ma sono proprie del cliente e difficilmente subiranno modifiche. Inoltre possiamo raccogliere informazioni di natura contabile e/o finanziaria che riguardano a pieno i flussi di denaro, come il limite di credito, la valuta, il pagamento, ecc. . . oltre ai dati in se per sè possono essere presenti informazioni riguardo ad accordi commerciali e/o sospensione del servizio.

Movimentazione contabile La movimentazione contabile è la parte più importante dell'area amministrativa di un ERP. Mentre la parte di anagrafiche permette di raccogliere informazioni riguardanti i clienti, i fornitori, ecc. . . la movimentazione contabile permette di raccogliere informazioni riguardanti i flussi di denaro. La struttura dei movimenti contabili si divide in: voce contabile (quale azienda, l'erario IVA, la cassa, ecc. . .), il Dare (uscite), l'avere (entrate) il Saldo ed il Segno (positivo o negativo). In questa struttura andiamo a registrare tutte le operazioni che vengono eseguite dall'azienda, in modo da avere un quadro chiaro e preciso della situazione finanziaria dell'azienda. Questa struttura è fondamentale per la gestione delle risorse finanziarie dell'azienda, in quanto permette di avere un quadro chiaro e preciso della situazione finanziaria dell'azienda.

Movimentazione finanziaria La movimentazione finanziaria traccia i debiti e i crediti rateizzati, i pagamenti e le scadenze. In generale quando si deve ricevere/effettuare un pagamento si crea un movimento finanziario, questo rimane con stato "Aperto" fino a quando non viene saldato e/o viene emessa una nota di credito o di debito da parte della stessa controparte che vada a saldare tutta o parte della somma dovuta. In sostanza la parte di movimentazione finanziaria deve tenere traccia del debito/credito di un'azienda nei confronti di un'altra e aggiustare di conseguenza i saldi contabili.

Movimentazione IVA La movimentazione IVA è una parte molto complessa ed in continuo aggiornamento in quanto deve tenere traccia della attuale normativa fiscale. In generale la movimentazione IVA deve tenere conto dell'imponibile e dell'aliquota IVA applicata, la quale varia in base alla natura del prodotto e se il cliente rientra nella categoria di "non soggetto" o "esente". Inoltre la movimentazione IVA deve tenere conto delle fatture emesse e ricevute, e delle note di credito e debito emesse e ricevute. In generale la movimentazione IVA deve tenere traccia di tutte le operazioni che riguardano l'IVA, in modo da poter calcolare in modo corretto l'IVA da versare all'erario. Tutte le righe di movimentazione IVA devono essere collegate ad una riga di movimentazione contabile.

6.2 L'area logistica

6.3 L'area vendite

6.4 L'area acquisti

6.5 L'area produttiva

6.6 Sistemi Operazionali complementari

Capitolo 7

I sistemi Informazionali

In questo capitolo si affrontano i sistemi informazionali, ovvero quei sistemi che permettono di estrarre informazioni utili dai dati aziendali. Questi sistemi sono fondamentali per la gestione delle risorse aziendali e per la pianificazione delle attività aziendali.

7.1 Gli Obiettivi

L'obiettivo generale che un buon sistema informazionale dovrebbe avere è quello di poter sfruttare tutti i dati raccolti tramite i flussi operazionali per poterli usare a supporto delle attività operative e per identificare informazioni utili per la gestione dell'azienda. In particolare un sistema informazionale segue i seguenti obiettivi:

- Il sistema informazionale come strumento a supporto delle decisioni superando sistemi di reporting e fogli di calcolo
- Interrogazione dei dati in maniera puntuale e complessa
- Base di dati
- Strumenti di analisi dei dati

Sistema informazionale come supporto alle decisioni Il sistema informazionale deve essere uno strumento al supporto delle decisioni tramite l'elaborazione di dati, prima dei sistemi informazionali questi venivano forniti tramite *report* i quali però presentano *bias* basati sulla persona che lo prepara (come dati omessi o punti di enfasi), inoltre risultano statici e difficili nella loro creazione. Altri strumenti superati grazie ai sistemi informazionali sono i fogli di calcolo, i quali sono soggetti a errori, è difficile collaborare e sono difficili da mantenere in quanto se si vuole modificare una formula bisogna modificarla in tutti i fogli di calcolo distribuiti.

Sistema informazionale come strumento di interrogazione dei dati Le interrogazioni si suddivono principalmente in: interrogazioni puntuali e interrogazioni complesse. Esempio di interrogazione puntuale è la ricerca della modalità di pagamento di un dato cliente, mentre un esempio di interrogazione complessa è l'analisi dell'aumento del margine operativo per una serie di prodotti rispetto all'anno precedente.

Sistema informazionale come strumento di accesso alla base di dati Il sistema informazionale deve offrire agli utenti finali un modello intuitivo ed efficiente per l'analisi, inoltre deve garantire la possibilità di integrare da fonti diverse i dati ed offrire processi di aggiornamento di questi.

Sistema informazionale come strumento di analisi dei dati In questo caso il sistema informazionale deve offrire strumenti di analisi dei dati, come *reporting* di dati, oppure offre dei sistemi per l'analisi interattiva da una ipotesi iniziale oppure offre sistemi di *Data Mining*

7.2 Concetti Generali

7.2.1 Terminologia

Data Warehouse Il *Data Warehouse* è quell'insieme di tecniche per definire, costruire e mantenere un *Data Base* che sia orientato all'analisi dei dati. Questo *Data Base* deve essere molto strutturato e i dati organizzati in maniera efficiente per permettere l'analisi dei dati.

Decision Support System - DSS Il DSS è una parte del sistema informazionale che integrato nel processo decisionale dell'azienda, permette di visualizzare ed estrarre le informazioni da basi di dati ben organizzate

Data mining Il *Data Mining* è una tecnica che permette di estrarre informazioni utili da un grande insieme di dati, questa tecnica è utilizzata per scoprire relazioni tra i dati e per fare previsioni.

Buisness Intelligence Ovvero tutte le attività di estrazione di informazioni dai dati di *business* generati dai processi operativi aziendali.

Knowledge Management Ovvero l'insieme della conoscenze che ogni individuo possiede i quali dovrebbero essere distribuiti in maniera efficiente all'interno del sistema informazionale. Spesso questa conoscenza è derivata da esperienze passate, da informazioni acquisite e da competenze acquisite, alcune volte questa conoscenza è difficile da trasferire ed è difficile da codificare.

Big Data Estensione del concetto di *Data Warehouse* in cui si considerano anche dati provenienti da flussi continui e dati non strutturati. Vengono considerati, sotto il concetto di *Big Data*, anche tutti quei dati la cui produzione non ha costi aggiuntivi (stringhe di un motore di ricerca, tempo medio di visita ad una pagina web, ecc...).

7.2.2 Le caratteristiche

Finalità Il sistema informazionale deve essere in grado di fornire un substrato informativo per la conoscenza dell'azienda, inoltre deve essere in grado di descrivere il passato ed aiutare ad identificare i problemi e le loro cause. Inoltre deve poter suggerire i cambiamenti da apportare e fornire anticipazioni sui scenari futuri.

Struttura I dati devono essere articolati intorno ai soggetti di cui si vuol conoscere l'apporto alla vita aziendale. Si organizzano dunque le informazioni per "tema" e non per "funzione" come nei sistemi operazionali.

Utenza I principali utenti del sistema informazionale sono i manager e i decisori che devono avere una visione e conoscenza ampia dell'azienda

Storicità Il sistema deve mantenere e fornire uno storico dei dati con un arco temporale adeguato e molto più esteso rispetto a quello dei sistemi operazionali, inoltre deve fornire l'evoluzione storica dei soggetti di interesse.

Dettaglio I dati sono quasi esclusivamente in forma aggregata e devono essere disponibili a diversi livelli di aggregazione.

Accesso L'accesso ai dati è principalmente solo in lettura, eventuali aggiornamenti sono solo periodici e in momenti nei quali l'attività aziendale è ferma.

7.3 Modello multidimensionale

Il modello multidimensionale è un modello di rappresentazione dei dati che permette di rappresentare i dati in maniera più intuitiva rispetto al modello relazionale. In questo modello il processo di analisi

viene posto al centro del sistema e non più il processo di inserimento dei dati (i quali vengono inseriti non in maniera diretta ma vengono estratti da un SI Operazionale¹).

In questo modello lo spazio delle informazioni viene rappresentato come insieme di matrici multidimensionali, dove ogni matrice rappresenta un tipo di evento (quale ad esempio l'immatricolazioni degli studenti), ogni elemento della matrice rappresenta un singolo evento (la singola immatricolazione) e ogni coordinata della matrice rappresenta una dimensione dell'evento (come ad esempio il corso di laurea, l'anno di immatricolazione, ecc...).

7.3.1 Caratteristiche

Ipercubo Il modello multidimensionale è rappresentato da un ipercubo, ovvero una matrice multidimensionale con n dimensioni, dove ogni cella rappresenta un singolo **fatto elementare**, ogni **dimensione** costituisce una coordinata del fatto e ogni **misura** è il valore numerico del fatto.

Fatti Un fatto è un evento che si vuole analizzare e misurare, quale una vendita, un acquisto, ecc... I fatti sono caratterizzati da un insieme di dimensioni che lo collocano nel tempo (quando è avvenuto), nello spazio aziendale (dove è avvenuto) e inoltre sono presenti altre dimensioni che lo quantificano e altre informazioni descrittive.

Misure La misura è una caratteristica (e non dimensione) del fatto che si vuole analizzare, e ne descrive un aspetto quantitativo. Ogni fatto può contenere una o più misure, le quali possono essere: **effettive** se memorizzate in maniera diretta, **calcolate a run-time** usando i valori delle misure effettive o **implicite** ovvero che indicano la presenza o meno di un fatto.

7.3.2 Aggregabilità

A partire dai **fatti** elementari si possono ricavare dei **fatti** sintetici quando si procede all'eliminazione di una o più dimensioni, in questo caso si parla di **aggregazione** dei dati. Questo processo di aggregazione viene fatto tramite opportuni operatori sui dati, in base a cosa rappresenta la dimensione aggregata, ad esempio se aggregiamo per mese non ha "senso" avere a fine anno la somma del totale dei prodotti in magazzino a fine mese, ma ha senso avere la media mensile.

Operatori di aggregazione Per ogni coppia costituita da (misura, aggregazione) possono essere definiti operatori e regole di aggregazione, ad esempio può essere presente una misura non aggregabile lungo una dimensione (vedi esempio precedente), oppure un operatore può essere utilizzato per aggregare lungo determinate dimensioni ma non in altre. Definiamo quindi con il termine **aggregabilità**

Definizione 7.1 (Aggregabilità). *La possibilità di usare un operatore di aggregazione su una misura o su una coppia (misura, dimensione).*

Nel caso più specifico definiamo anche il termine **additività** ovvero

Definizione 7.2 (Additività). *La possibilità di usare l'operatore di aggregazione "somma" su una misura o su una coppia (misura, dimensione).*

Tipi di misura Le misure possono essere di diversi tipi, tra i quali troviamo:

di Livello si prende in considerazione il valore proprio del fatto, ed il momento nel quale è stato registrato. Non viene mai usata l'aggregazione additiva sulla dimensione temporale.

Unitaria si prende in considerazione il valore di uno dei soggetti della misura, e non è mai aggregabile con l'additività.

di Flusso si prende in considerazione il valore proprio del fatto ed un intervallo di riferimento, questa misura è aggregabile con l'additività lungo una qualunque dimensione.

¹Si rimanda al capitolo 5: I sistemi operazionali per ulteriori informazioni sui sistemi operazionali

Articolo		Deposito	Data	Misura (quantità)
PP1007015	Pannello di polistirolo 100x70x15	1	01/01/2019	100
PP1007015	Pannello di polistirolo 100x70x15	2	01/01/2019	200
VA1010	Vite autofilettante 10x10	1	01/01/2019	1000

PP1007015	Pannello di polistirolo 100x70x15	1	01/02/2019	150
PP1007015	Pannello di polistirolo 100x70x15	2	01/02/2019	250
VA1010	Vite autofilettante 10x10	1	01/02/2019	1100

Tabella 7.1: Esempio di tabella di fatti

Esempio In questo esempio abbiamo una tabella di fatti, dove ogni riga rappresenta un singolo fatto, ogni colonna rappresenta una dimensione del fatto e l'ultima colonna rappresenta la misura del fatto. In questo caso la misura è la quantità di prodotto in magazzino. Questa è additiva rispetto alla dimensione "Deposito" ma non lo è rispetto alla dimensione "Data", inoltre non è aggregabile rispetto alla dimensione "Articolo".

7.4 Caratteristiche strutturali e funzionali

Andiamo ad analizzare come sono strutturati i dati all'interno di un sistema informazionale, in particolare analizziamo le caratteristiche strutturali e funzionali di un sistema informazionale.

7.4.1 Caratteristiche strutturali

Le caratteristiche appena descritte sono le fondamenta dell'organizzazione dei dati in un sistema informazionale. Esistono però altre caratteristiche che sono fondamentali per la corretta organizzazione dei dati in un sistema informazionale.

Multidimensionalità Per dimensionalità, in generale, si intende il numero di dimensioni necessarie per identificare un fatto, mentre nei sistemi operazionali abbiamo una dimensionalità puntuale (ad esempio il codice fiscale di un cliente), nei sistemi informazionali, vista la loro natura, per identificare un fatto servono tutte le dimensioni che lo caratterizzano.

Granularità La granularità misura quanto l'informazione proposta è sintetica rispetto agli eventi sulla quale si basa, nel caso dell'ipercubo questa è una granularità minima in quanto ogni fatto corrisponde alla cella di origine dell'ipercubo. Avere diversi gradi di granularità permette di stilare analisi più efficienti ed avere una visione più chiara dei dati.

Arco temporale L'arco temporale rappresenta il periodo di tempo che si vuole analizzare, in generale questo arco temporale è molto più ampio rispetto a quello dei sistemi operazionali, inoltre deve essere possibile analizzare i dati in maniera retrospettiva. Deve essere possibile avere almeno uno storico di 10-15 anni.

Profondità storica Mentre i sistemi operazionali non tengono traccia solitamente di come un particolare dato e/o anagrafica si è evoluto nel tempo. Questa caratteristica è fondamentale nei sistemi informazionali in quanto anche l'evoluzione del fatto nel tempo è una dimensione del fatto stesso.

7.4.2 Caratteristiche funzionali

Oltre ad avere una struttura ben definita i sistemi informazionali devono avere delle caratteristiche funzionali ben definite, queste sono fondamentali per avere una analisi completa e coerente dei dati rilevati.

Integrazione dei dati I dati proposti al SI informazionale spesso provengono da fonti diverse con diverse strutture e formati, è quindi fondamentale avere un sistema che permetta di integrare questi dati in maniera efficiente e coerente in modo da evitare errori, duplicazioni, ecc...

Accessibilità In quanto i sistemi a supporto delle decisioni spesso vengono usati da persone con poca esperienza informatica è fondamentale che l'accesso ai dati sia il più semplice ed intuitivo possibile. Inoltre è fondamentale che il tempo di accesso ai dati sia in tempo utile per l'attività decisionale.

Flessibilità Nel contesto dei sistemi informazionali intendiamo con flessibilità la capacità del sistema di adattarsi alle diverse interrogazioni possibili sui dati. Quindi il sistema deve essere in grado di articolare le richieste, aggregare dati a più livelli e con criteri non predefiniti, dare la possibilità di mettere in relazione misure diverse, ecc. . . . Questo è fondamentale perché ogni azienda ha esigenze e dati diversi, e visto che è compito dell'SI informazionale supportare le decisioni ed suggerire cambiamenti è fondamentale che questo possa rappresentare i dati come richiesto dall'utente.

Correttezza I dati proposti dal sistema informazionale devono essere necessariamente corretti, ciò visto lo scopo del sistema informazionale è fondamentale in quanto decisioni prese su dati errati possono portare a conseguenze disastrose. I dati necessari all'analisi non sono spesso di utilità operativa, quindi non sono controllati come i dati operativi, inoltre i dati necessari all'analisi provengono da più fonti ed è quindi possibile che la stessa entità sia memorizzata in maniera diversa in due fonti diverse.

Completezza I dati proposti dal sistema informazionale devono essere completi, ovvero non devono mancare dati utili al processo decisionale, per garantire questo la *data warehouse* deve essere popolata con tutti i dati necessari all'analisi.

7.4.3 *Data warehouse e Data mart*

Come già affrontato la *data warehouse* è un insieme di dati raccolti su decenni di informazioni aziendali, organizzati in maniera efficiente all'analisi. Questi dati provengono da fonti relazionali diverse (SI operazionali, fonti esterne, ecc. . .) e vengono integrati in un unico sistema. Dato che la *data warehouse* può raggiungere dimensioni molto grandi quando si deve andare ad analizzare un particolare aspetto dell'azienda l'utente può creare un *data mart*, ovvero una porzione della *data warehouse* che contiene solo i dati necessari all'analisi. Le *data mart* possono essere "telematiche" ovvero possono essere create in tempo reale selezionando aspetti specifici della *data warehouse* oppure possono essere "fisiche" ovvero possono essere create in maniera fisica e mantenute in maniera separata dalla *data warehouse*.

Capitolo 8

Data Warehousing

I sistemi di *data warehousing* sono alla base di quasi tutti i DSS (*Decision Support System*) e BI (*Business Intelligence*) sono progettati per gestire grandi quantità di dati, per fornire un accesso rapido e per supportare le operazioni di analisi e di reporting.

8.1 *Data Warehouse* e metodologia OLAP

Prima della metodologia OLAP nasce la metodologia OLTP (*On-Line Transaction Processing*) che è una tecnica di gestione dei dati per sistemi operazionali basata su 12 criteri. La metodologia OLAP (*On-Line Analytical Processing*) è una tecnica di analisi dei dati iterativa. Negli anni i criteri per definire se un prodotto è OLAP o meno sono stati semplificati in 5 punti caratterizzati dall'acronimo FASMI:

F *Fast Analysis*: l'analisi deve essere veloce. (Tempo di risposta medio inferiore a 5 secondi)

A *Analytical* (Analitico): l'analisi deve essere analitica e deve:

Dare la possibilità di eseguire nuovi calcoli a partire dai calcoli già eseguiti.

Fornire risposte a richieste specifiche e non solo a domande predefinite.

Fornire i dati in rappresentazioni molteplici.

S *Shared* (Condiviso): i dati devono essere utilizzabili da più utenti contemporaneamente che condividono la stessa base di dati di analisi.

M *Multidimensional* (Multidimensionale): i dati devono essere organizzati in modo multidimensionale.

I *Informational* (Informativo): il sistema deve contenere tutte le informazioni necessarie per l'analisi indipendentemente da dove e come sono memorizzate.

8.2 Architettura dei sistemi di *data warehousing*

Un sistema di *data warehousing* è costituito da basi di dati poste a livelli diversi, ognuno di questi ha una finalità, una struttura ed una tipologia di dati contenuti. Questi livelli sono:

Sorgenti Sono le basi di dati operative (o esterne) da cui si estraggono i dati.

Staging Area È una area intermedia in cui i dati estratti dalle sorgenti vengono trasformati e caricati in una forma adatta per l'immagazzinamento presso la *Data Warehouse*.

Data Warehouse È il livello in cui i dati vengono immagazzinati in modo da poter essere utilizzati per l'analisi.

Data Mart È un sottoinsieme del *Data Warehouse* che contiene dati specifici per un determinato settore o per un determinato gruppo di utenti.

Elementi propri di un sistema sono oltre alle varie basi di dati:

- Procedure che permettono di estrarre, trasformare e caricare i dati.

- Strumenti di analisi

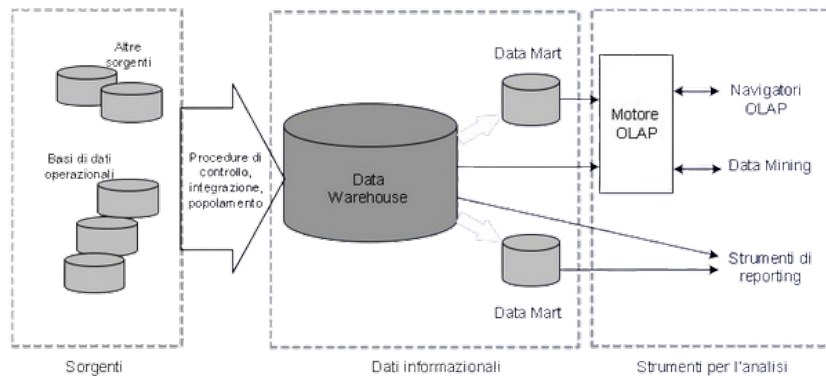


Figura 8.1: Livelli di un sistema di *data warehousing*

8.3 Modelli concettuali per il *data warehouse*

I sistemi informativi vista la quantità di informazioni contenute all'interno necessitano di modelli dati a descrivere la composizione del sistema e della sua base di dati. Viene dunque definito uno schema dei fatti, il DFM (*Dimensional Fact Model*).

8.3.1 *Dimensional Fact Model* - DFM

Il modello DFM viene usato per descrivere un fatto, tutte le sue misure e le dimensioni usabili per l'analisi (quelle sulle quali il fatto è aggregabile). Un DFM è composto da:

- Un fatto rappresentato tramite un rettangolo
- Le misure contenute nel fatto (sia quelle proprie che quelle derivate)
- Le dimensioni base collegate al fatto (le coordinate del fatto) rappresentate da cerchi
- Gli attributi descrittivi collegati o al fatto o alla dimensione.
- Le gerarchie tra le dimensioni sono rappresentate come un albero con la radice in alto.

Se una gerarchia è condivisa tra più dimensioni si rappresenta con un cerchio pieno con n collegamenti al fatto e alle dimensioni

- È possibile inoltre rappresentare le relazioni di aggregabilità tra misure e dimensioni. Queste relazioni sono rappresentate con una linea tratteggiata (se non è possibile aggregare) o continua (se è possibile aggregare).

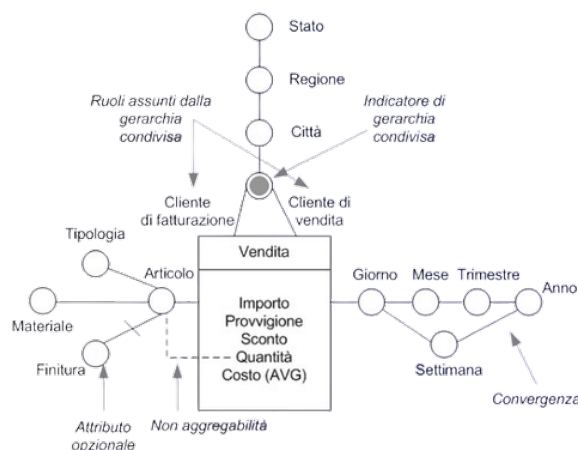


Figura 8.2: Esempio di DFM con tutte le caratteristiche descritte

C'è da aggiungere che il presente modello è riconducibile ad un modello **E-R** (*Entity-Relationship*) in cui le dimensioni sono le entità e le gerarchie sono le relazioni tra le entità ed il fatto è l'entità principale.

8.4 Modelli logici per il *data warehouse*

Mentre la modellazione concettuale riguarda come i fatti e le dimensioni sono collegati tra loro, la modellazione logica riguarda come i fatti e le dimensioni sono effettivamente memorizzati nel *Data Warehouse*. I modelli logici si differenziano sulla base della scelta del DBMS e della struttura di memorizzazione dei dati. Oltre a definire come i dati sono memorizzati, i modelli logici definiscono anche come i dati sono interrogati e analizzati.

8.4.1 ROLAP

Il modello ROLAP (*Relational OLAP*) è un modello logico che prevede l'utilizzo di un DBMS relazionale e la memorizzazione viene eseguita tramite tabelle. Questo modello inoltre prevede che le interrogazioni avvengano tramite *query SQL* con l'uso di viste e/o funzioni di aggregazione.

Vantaggi Ridotto uso di spazi di memorizzazione e la maggiore conoscenza degli strumenti relazionali portano ad una maggiore facilità di gestione.

Svantaggi Le prestazioni delle interrogazioni possono essere peggiori rispetto ad altri modelli e la complessità delle interrogazioni può aumentare se lavoriamo con diverse aggregazioni ed analisi.

8.4.2 MOLAP

Il modello MOLAP (*Multidimensional OLAP*) è un modello logico che prevede l'utilizzo di strutture dati intrinsecamente multidimensionali. Questo modello prevede che i dati siano memorizzati in strutture multidimensionali (quali vettori, matrici, ...) i sistemi allocano spazio per tutte le possibili combinazioni di dimensioni e misure.

Vantaggi Le prestazioni delle interrogazioni sono migliori rispetto ad altri modelli in quanto il fatto ricercato viene trovato in un'unica posizione e non è necessario "simulare" le dimensioni. Inoltre le operazioni di aggregazione sono più veloci in quanto si considera solo il livello di aggregazione richiesto.

Svantaggi L'uso di spazio di memorizzazione è maggiore rispetto ad altri modelli, solitamente solo il 20% dello spazio allocato viene utilizzato. Inoltre non esiste uno standard, tutte le implementazioni sono proprietarie e non inter compatibili.

8.4.3 HOLAP

Il modello HOLAP (*Hybrid OLAP*) è un modello logico che prevede l'utilizzo di un DBMS relazionale e di strutture multidimensionali. Questo modello prevede che i dati siano memorizzati in strutture multidimensionali (per le aggregazioni di livello più alto) e in tabelle relazionali (per le aggregazioni di livello più basso). Le interrogazioni possono essere eseguite sia tramite *query SQL* che tramite *query* multidimensionali.

Vantaggi Questo modello permette di sfruttare i vantaggi di entrambi i modelli, inoltre permette di utilizzare le strutture multidimensionali per le aggregazioni di livello più alto e le tabelle relazionali per le aggregazioni di livello più basso. Infatti le *data mart* possono essere memorizzati in modo multidimensionale mentre il *data warehouse* può essere memorizzato in modo relazionale.

Svantaggi La complessità di gestione è maggiore rispetto ad altri modelli e le prestazioni delle interrogazioni possono essere peggiori rispetto ad altri modelli.

8.4.4 Schemi multidimensionali su basi di dati relazionali

Visto che la maggior parte delle *data warehouse* vengono memorizzati in basi di dati relazionali, vediamo come è possibile implementare uno schema multidimensionale su una base di dati relazionale.

Schema a stella

Lo schema a stella è un modello multidimensionale in cui il fatto, le sue misure semplici e elementi “chiave” riguardanti le dimensioni sono memorizzati in una tabella centrale (il fatto), le dimensioni vere e proprie sono memorizzate in tabelle separate collegate al fatto tramite una chiave esterna. Questo modello è molto semplice e facile da implementare, inoltre è molto efficiente per le interrogazioni.

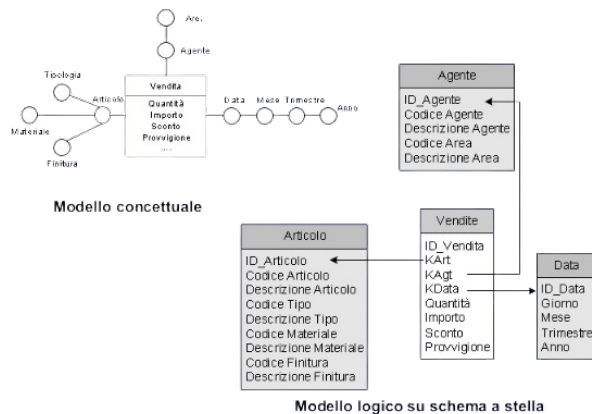


Figura 8.3: Schema a stella

Schema a fiocco di neve

Lo schema a fiocco di neve è un'estensione dello schema a stella in cui le tabelle delle dimensioni sono normalizzate. Questo modello permette di risparmiare spazio di memorizzazione in quanto le tabelle delle dimensioni sono più piccole, inoltre permette di ridurre la ridondanza dei dati. Tuttavia le interrogazioni possono essere più complesse e le prestazioni possono essere peggiori rispetto allo schema a stella.

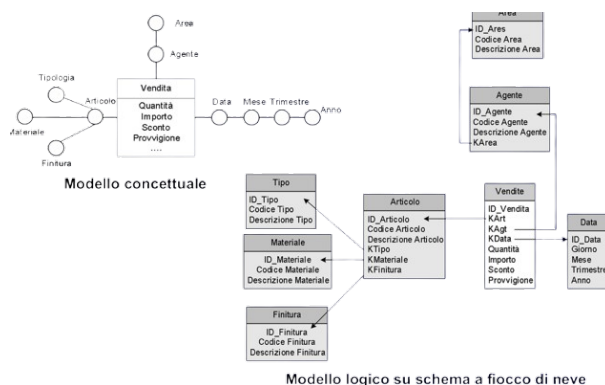


Figura 8.4: Schema a fiocco di neve

Costellazione di fatti

Lo schema a costellazione di fatti è un'estensione dello schema a stella in cui sono presenti più fatti collegati tra loro. Questo modello permette di analizzare più fatti contemporaneamente e di effettuare analisi più complesse. Tuttavia le interrogazioni possono essere più complesse e le prestazioni possono essere peggiori rispetto allo schema a stella.

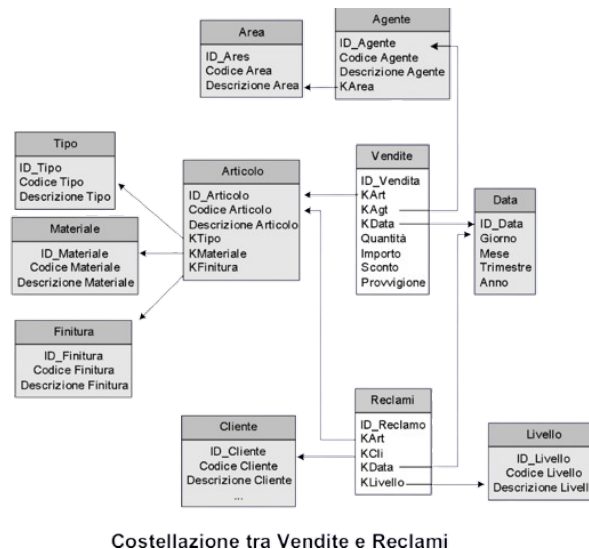


Figura 8.5: Schema a costellazione di fatti

8.5 Ciclo di vita di DWH e Popolazione del DWH

Andiamo ora a vedere come viene costruito un DWH e come vengono popolati i vari livelli.

8.5.1 Ciclo di vita di un DWH

La costruzione di una DWH è un processo iterativo che prevede diverse fasi:

1. Costruzione del primo ipercubo multidimensionale sul fatto più importante.
2. Integrazione progressiva degli altri fatti
3. Rilascio di *data mart* per i vari settori aziendali

Il vantaggio principale di questo approccio è che i primi risultati sono visibili in tempi brevi, gli investimenti sulla DWH sono diluiti nel tempo ed si può tarare il modello sulla base delle esigenze degli utenti.

8.5.2 Popolazione della DWH

La popolazione della DWH prevede diverse fasi:

1. Estrazione dei dati dalle sorgenti - i dati vengono estratti dalle sorgenti e trasferiti nella DWH
2. Integrazione e trasformazione dei dati - associazione dei dati estratti con i dati già presenti nella DWH
3. Pulizia dei dati - aumento della qualità
4. Caricamento dei dati nella DWH - i dati vengono caricati nella DWH

8.6 L'analisi OLAP e principali operatori OLAP

L'analisi OLAP è un'analisi iterativa che si basa sul paradigma dell'"esplorazione guidata delle ipotesi". In sostanza una analisi OLAP prevede che in una stessa sessione di analisi ciascun passo è conseguenza dei risultati ottenuti al passo precedente, inoltre tutte le interrogazioni operano per differenza rispetto alla sessione precedente, ovvero il risultato di una interrogazione è la differenza tra il risultato dell'interrogazione corrente e il risultato dell'interrogazione precedente. I risultati finali sono presentati sotto forma di tabelle, grafici o mappe.

Operatori OLAP Gli operatori OLAP sono operatori che permettono di effettuare analisi sui dati. Gli operatori principali sono:

Drill Down Consente di passare da un livello di aggregazione ad un livello di dettaglio inferiore.

Roll Up Consente di passare da un livello di dettaglio ad un livello di aggregazione superiore.

Slice Consente di selezionare un sottoinsieme di dati su una dimensione.

Dice Consente di selezionare un sottoinsieme di dati su più dimensioni.

Pivot Consente di scambiare le righe con le colonne di una tabella oppure ruotare una rappresentazione grafica su un asse.

Drill Down

L'operatore *Drill Down* consente di passare da un livello di aggregazione ad un livello di dettaglio inferiore. Questo operatore permette di visualizzare i dati in modo più dettagliato. Si può aggiungere una dimensione oppure si scende lungo una gerarchia. Graficamente andiamo a “separare” i dati in più sottoinsiemi, esempio possiamo passare da una tabella che mostra per la dimensione zona Nord Centro Sud il totale delle vendite per il 2020 ad una tabella che mostra per la dimensione delle regioni il totale delle vendite per il 2020. Siamo andati a “separare” i dati per regione e non più per zona.

Roll Up

L'operatore *Roll Up* è l'inverso dell'operatore *Drill Down*, consente di passare da un livello di dettaglio ad un livello di aggregazione superiore. Questo operatore permette di visualizzare i dati in modo più aggregato. Si può rimuovere una dimensione oppure si sale lungo una gerarchia. Graficamente andiamo a “unire” i dati in un unico insieme, esempio possiamo passare da una tabella che mostra per la dimensione delle regioni il totale delle vendite per il 2020 ad una tabella che mostra per la dimensione zona Nord Centro Sud il totale delle vendite per il 2020. Siamo andati a “unire” i dati per zona e non più per regione.

Slice

L'operatore *Slice* consiste nel selezionare un sottoinsieme di dati su una dimensione. Questo significa che una volta determinato il valore che vogliamo selezionare, andiamo a “tagliare” i dati in modo che vengano visualizzati solo i dati che rispettano il valore selezionato. Tutte le altre dimensioni rimangono invariate.

Dice

L'operatore *Dice* è di base una combinazione di più operatori *Slice*. Consente di selezionare un sottoinsieme di dati su più dimensioni. Questo significa che una volta determinati i valori che vogliamo selezionare, andiamo a “tagliare” i dati in un cubo in modo che vengano visualizzati solo i dati che per ogni dimensione rispettano il valore selezionato. Tutte le altre dimensioni rimangono invariate.

Pivot

L'operatore *Pivot* inverte la relazione tra le dimensioni andando a ruotare il cubo dell'analisi su un asse. Questo operatore consente di visualizzare i dati in modo differente, ad esempio possiamo passare da una tabella che mostra per la dimensione zona Nord Centro Sud il totale delle vendite per il 2020 ad una tabella che mostra per la dimensione delle regioni il totale delle vendite per il 2020. Siamo andati a “ruotare” i dati su un asse.

Capitolo 9

Introduzione al *Data Mining*

Nota dell'autore: Questo capitolo è stato tagliato in quanto durante il corso dell'anno accademico 2024/2025 il prof. Bouquet per mancanza di tempo non ha potuto trattare interamente l'argomento, si riporta dunque solo la parte che è stata trattata.

In questo capitolo verranno introdotti i concetti fondamentali del Data Mining, ovvero l'insieme di tecniche e metodologie che permettono di estrarre informazioni utili da grandi quantità di dati in modo automatico.

9.1 Limiti analisi OLAP

I sistemi di *data warehouse* con analisi OLAP permettono di analizzare i dati in modo interattivo, tuttavia presentano queste analisi sono basate su supposizioni che l'utente fa sui dati, e non permettono di trovare informazioni che l'umano non è in grado di trovare. Inoltre, l'analisi di grandi quantità di dati può essere molto dispendiosa in termini di tempo e risorse. Questo accade perché i sistemi OLAP per la loro natura sono progettati per essere a supporto delle decisioni umane e non si basa sui dati oggettivi per trovare correlazioni e *pattern*.

Il *Data mining* Il *Data Mining* è stato introdotto per rispondere alle problematiche dei sistemi OLAP, permettendo di trovare informazioni riguardanti correlazioni nascoste e supporta modelli descrittivi e predittivi.

Il *data mining* permette dopo aver pulito, integrato, selezionato e trasformato i dati viene applicato un algoritmo di *data mining* che permette di trovare *pattern* e relazioni nascoste nei dati. A questo punto i risultati vengono valutati ed presentati a chi deve prendere decisioni. Possiamo notare come i primi due passaggi combacino con quelli del popolamento del *data warehouse*, infatti il *data mining* può essere visto come un ampliamento del *data warehouse* ed in alcuni casi un suo completamento.

Da OLAP a OLAM Partendo da un *data warehouse* possiamo estrarre i dati da sottoporre a *data mining* anche se il processo di *data mining* non deve essere completamente automatico in quanto potremmo incorrere in *pattern* non significativi. Lavorando con uno strumento iterativo possiamo ottenere risultati migliori.

9.2 Architettura e tipi di analisi con *data mining*

Architettura

L'architettura di un sistema di *data mining* è composta da:

***data warehouse* - sorgente dati** Il *data warehouse* è la sorgente dei dati da analizzare, i dati vengono estratti e trasformati in modo da essere pronti per l'analisi.

***Knowledge Base* - base di conoscenza** La base di conoscenza contiene i modelli e le regole che vengono utilizzate sia per l'analisi dei dati che per la valutazione dei risultati.

Data Mining Engine* - motore di *data mining Il motore di *data mining* è il cuore del sistema, contiene gli algoritmi che permettono di trovare i *pattern* nei dati.

***pattern evaluation* - valutazione delle condizioni** Questo modulo interagisce coi moduli di *mining* per focalizzare la ricerca sui *pattern* più interessanti.

***User Interface* - sistema di presentazione** Questo modulo permette all'utente di interagire con il sistema, visualizzando i risultati e permettendo di modificare i parametri di ricerca.

Tipi di analisi

Le attività che possono essere eseguite sono molteplici, troviamo due macro-categorie:

***Mining* descrittivo** estrae informazioni che descrivono le proprietà dei dati.

***Mining* predittivo** determina regole che permettono di fare previsioni sui dati.

Oltre a questi due tipi di analisi possiamo avere anche analisi diagnostiche ovvero che permettono di capire le cause di un determinato fenomeno, ma anche prescrittiva che permette di suggerire azioni da intraprendere per ottenere un determinato risultato.