

Sistemi Informativi

Alessio Blascovich

June 28, 2023

Contents

1	Introduzione e principi generali	6
1.1	Società della conoscenza	6
1.1.1	La strategia di Lisbona	6
1.2	Digitalizzazione e convergenza tecnologica	6
1.3	Nuove leggi	7
1.3.1	Leggi di Moore	7
1.3.2	Leggi di Sarnoff, MetCalfe, Reed	8
1.3.3	Hype Cycle di Gartner	9
1.3.4	Magic Quadrant di Gartner	10
1.3.5	Long tail di Anderson	11
2	Concetti generali sull'informatica aziendale	11
2.1	Sistema Informativo Aziendale (SIA) prima definizione	11
2.2	Fenomeni che guidano alla costruzione di un SIA	11
2.3	Elementi che compongono un SIA	12
2.4	Sistema Informativo Aziendale (SIA) seconda definizione	13
2.4.1	Sistema delle informazioni aziendali	13
2.5	Fattori di evoluzione di un SIA	13
2.6	L'impatto dei SIA	13
2.6.1	Conoscenza dei fenomeni aziendali	13
2.6.2	Processi classici	14
2.6.3	Nuovi processi	15
2.6.4	Business Process Rengineering (BPR)	15
2.7	Struttura aziendale italiana	16
2.7.1	Conseguenze per i SIA	16
2.8	Cambiamenti organizzativi aziendali	16
2.8.1	Organizzazione interna	16
2.8.2	Organizzazione esterna	17

3	Struttura dell'azienda e del suo sistema informativo	17
3.1	Esigenza informativa	17
3.2	Schema di Anthony	18
3.3	Scomposizione del sistema informativo	18
3.3.1	Sistema operazionale	18
3.3.2	Sistema informazionale	19
3.4	Comparazione tra i sistemi informativi	20
4	Scelte organizzative	20
4.1	Opzione make	20
4.2	Opzione buy	21
4.3	Opzione outsource	22
4.4	Figure professionali nell'informatica aziendale	23
4.4.1	Livello 1	23
4.4.2	Livello 2	23
4.4.3	Livello 3	24
4.4.4	Livello 4	24
4.5	Posizionamento nell'organigramma	25
4.5.1	Supporto amministrativo	25
4.5.2	Servizio altre direzioni aziendali	26
4.5.3	Servizio all'organizzazione	26
4.6	Infrastruttura tecnologica dei SIA	27
4.6.1	Da server locale a cloud (outsource)	27
4.7	Interrompibilità del servizio	27
4.7.1	Guasti Hardware	27
4.7.2	Guasti software	28
4.7.3	Guasti dolosi	28
5	Sistemi operazionali	28
5.1	Finalità dei sistemi operazionali	28
5.1.1	Registrazione delle transizioni	29
5.1.2	Pianificazione e controllo delle operazioni	29
5.1.3	Organizzazione della conoscenza	30
5.1.4	Elaborazione della situazione aziendale	30
5.2	Informazioni operative	30
5.2.1	Qualità dei dati	31
5.2.2	Caratteristiche strutturali	32
5.2.3	Caratteristiche funzionali	32
5.3	Potenzialità informatica	33
5.3.1	Intensità informativa	33

5.3.2	Intensità informatica	34
5.4	Composizione dei sistemi informativi operazionali	35
5.4.1	Portafoglio operativo	35
5.4.2	Portafoglio istituzionale	36
5.4.3	Dai sistemi tradizionali agli ERP	36
5.4.4	Ambiti degli ERP	37
5.4.5	Sistemi operazionali complementari	37
6	ERP: l'area amministrativa	37
6.1	Obbiettivi	37
6.2	Struttura di base	38
6.2.1	Piano dei conti	38
6.2.2	Strutture anagrafiche speciali	38
6.2.3	Movimenti contabili	39
6.3	Procedure di base	40
6.3.1	Procedure di alimentazione	40
6.4	Flussi evolutivi	40
6.4.1	Contabilità analitica	41
6.4.2	Budget	41
6.4.3	Controllo di gestione	41
7	ERP: l'area logistica	42
7.1	Obbiettivi	42
7.2	Struttura di base	43
7.2.1	Nominazione articoli	43
7.2.2	Anagrafiche dei prodotti	43
7.2.3	Layout aziendale	43
7.2.4	Movimentazione logistica	44
7.3	Procedure di base	44
7.4	Flussi evolutivi	44
7.4.1	Magazzino a lotti	45
7.4.2	Magazzino a matricole	45
7.4.3	Magazzino a celle	46
7.4.4	Magazzini automatici	46
8	ERP: l'area vendite	46
8.1	Obiettivi	46
8.2	Strutture di base	47
8.2.1	Anagrafica commerciale	47
8.2.2	Condizioni commerciali	47

8.2.3	Documenti di processo	49
8.2.4	Piani di processo	50
8.3	Procedure di base	50
8.3.1	Procedure di flusso	50
8.3.2	Procedure di analisi	50
8.4	Flussi evolutivi	50
8.4.1	Condizioni commerciali	51
8.4.2	Processi	51
8.4.3	Estensioni degli ERP	51
9	ERP: l'area acquisti	52
9.1	Obiettivi	52
9.2	Strutture di base	52
9.2.1	Anagrafiche commerciali	52
9.2.2	Condizioni commerciali	52
9.2.3	Documenti di processo	52
9.3	Procedure di base	53
9.3.1	Procedure di flusso	53
9.3.2	Procedure di analisi	53
9.4	Flussi evolutivi	53
9.4.1	Condizioni commerciali	54
9.4.2	Processi	54
9.5	Estensioni dell'ERP	54
9.5.1	SCM	54
9.5.2	E-procurement	54
10	ERP: l'area produttiva	55
10.1	Obiettivi	55
10.2	Concetti generali	55
10.2.1	Criteri di classificazione	55
10.2.2	Logiche organizzative dell'azienda	56
10.3	Strutture di base	57
10.3.1	Strutture anagrafiche	57
10.3.2	Strutture costi	58
10.3.3	Dati dinamici	58
10.4	Procedure di base	59
10.4.1	Procedure di trattamento materiali	59
10.4.2	Procedure di trattamento lavorazioni	59
10.4.3	Procedure di manutenzione	60
10.5	Flussi evolutivi	60

10.5.1	Pianificazione a medio/lungo termine	60
10.5.2	Schedulatori di produzione	60
10.5.3	Rilevazione automatizzata dei dati da campo	60
10.5.4	Commesse cliente/impianto	60
10.5.5	Sistemi tecnici	61
11	Sistemi informativi	61
11.1	Obiettivi	61
11.2	Concetti generali	62
11.3	Caratteristiche dei dati informativi strutturati	63
11.3.1	Modello multidimensionale	63
11.3.2	Caratteristiche strutturali	64
11.3.3	Caratteristiche funzionali	65
11.3.4	Data warehouse e data mart	65
12	Data warehousing	66
12.1	Data warehousing e OLAP	66
12.2	Architettura dei sistemi di data warehousing	66
12.3	Modelli concettuali per il data warehousing	67
12.3.1	DFM (Dimensional Fact Model)	67
12.4	Modelli logici per il data warehouse	68
12.4.1	ROLAP	69
12.4.2	MOLAP	69
12.4.3	HOLAP	69
12.4.4	Schemi multidimensionali su db relazionali	69
12.5	Ciclo di vita del DWH	71
12.5.1	Costruzione dei data mart	71
12.6	Popolamento del data warehouse	72
12.6.1	Fasi di popolamento	72
12.7	Tecniche di analisi dei dati	73
12.7.1	Drill down	73
12.7.2	Roll up	74
12.7.3	Slice	74
12.7.4	Dice	74
12.7.5	Pivot	75
13	Data mining	75
13.1	Limiti di OLAP	75
13.2	Fasi del processo di mining	76
13.3	Da OLAP a OLAM	76

13.4 Architettura dei sistemi di Data Mining	76
13.5 I 4 principi di analisi	77
13.6 Staticstiche elementari e analisi relative	77
13.6.1 Generalizzazione	77
13.7 Analisi associative	78
13.8 Classificazione e predizione	78
13.8.1 Classificazione	78
13.9 Meccanismi di clustering	79
13.9.1 Ricerca degli outlier	80

1 Introduzione e principi generali

1.1 Società della conoscenza

Una società nella quale l'informazione assume una centralità fondamentale nella vita quotidiana, e che quindi fonda la propria crescita sul sapere, l'innovazione e la ricerca.

Questa idea mette le basi per una società basata sull'informazione che soppianta la precedente era, l'era industriale basata su dispositivi meccanici.

1.1.1 La strategia di Lisbona

Nel 2000 l'UE si è prefissato un piano strategico per il decennio successivo, **diventare l'economia basata sulla conoscenza più competitiva del pianeta**, quindi realizzando uno sviluppo sostenibile con nuovi posti di lavoro e più coesione sociale.

E' qui di richiesta una strategia mondiale volta a:

- predisporre il passaggio verso una società dell'informazione e di ricerca & sviluppo.
- modernizzare il modello sociale.
- sostenere uno sviluppo sano.

1.2 Digitalizzazione e convergenza tecnologica

Sempre più i beni materiali verranno astratti per essere trasposti in modo digitale, e diventeranno interscambiabili in tempo reale.

Questo cambiamento porterà a delle conseguenze.

- **Veloità** Ogni aspetto della vità umana e ogni istituzione opererà sempre di più in tempo reale.
- **Connettività** Ogni cosa viene interconnessa in modo sempre maggiore, e.g. frighi smart.
- **Intangibilità** Molti oggetti e processi diventano intangibili per l'essere umano, e.g. e-commerce e cartelle cliniche.

1.3 Nuove leggi

Queste nuove leggi regolano il nostro nuovo modello di società, ed una loro infrazione porta alla rottura nel nostro modello sociale.

1.3.1 Leggi di Moore

1. 1° legge Gordon Moore disse : “Le prestazioni dei processori, e il numero di transistor ad esso relativo, raddoppiano ogni 18 mesi.”.
I limiti di questa legge sono solo i raggiungimenti di vari limiti fisici imposti dalla dimensione sempre minore dei transistor, però le nuove capacità raggiunte forniscono per uno stesso prezzo un raddoppio delle prestazioni.
2. Conseguenze La crescita esponenziale mette le basi per la *singularità tecnologica*, ovvero entro una x data (2045) un singolo computer potrebbe avere capacità computazionale pari a tutta l'umanità.
3. 2° legge “L'investimento per realizzare una nuova tecnologia di microprocessori cresce in maniera esponenziale con il tempo.”
Tutte le innovazioni che rendono possibile dimension i sempre ridotte hanno però un costo esponenziale, infatti si ha sempre maggior bisogno di studi, ricerche e test per garantire l'affidabilità dei nuovi componenti. Questo incremento dei costi ha come richiesta un aumento sempre maggiore del fatturato e quindi delle conseguenze:
 - Riduzione del numero di società che possono permettersi di investire.
 - Fenomeni di fusione di società.
 - Crescita dei rischi legati ad un investimento.

Si viene quindi a formare un consolidamento del mercato che arriva ad avere barriere di entrata altissime, con una gruppo ristretto di aziende

che controlla il mercato.

Diventa problematico anche avere un ritorno economico per questi nuovi investimenti, visto che i mercati di massa sono un numero finito si potrebbe arrivare in un punto in cui il benefit non giustifica l'investimento iniziale.

1.3.2 Leggi di Sarnoff, MetCalfe, Reed

1. Legge di Sarnoff Legata al concetto di broadcasting, “Il valore V di una rete di broadcasting è direttamente proporzionale al numero N degli utenti”.

$$V = N$$

Quindi il valore della rete cresce in base agli utenti connessi.

2. Legge di MetCalfe “Il valore V di un sistema di comunicazione cresce con il quadrato del numero N di persone collegate”.

$$V = N^2 - N$$

E.g. il fax da solo non vale nulla, acquista valore solo se diversi utenti iniziano ad usare il Fax.

Questa legge implica che la connessione di reti tra loro indipendenti è maggiore rispetto alla somma delle singole reti.

Questa legge si rifà molto allo sviluppo di internet.

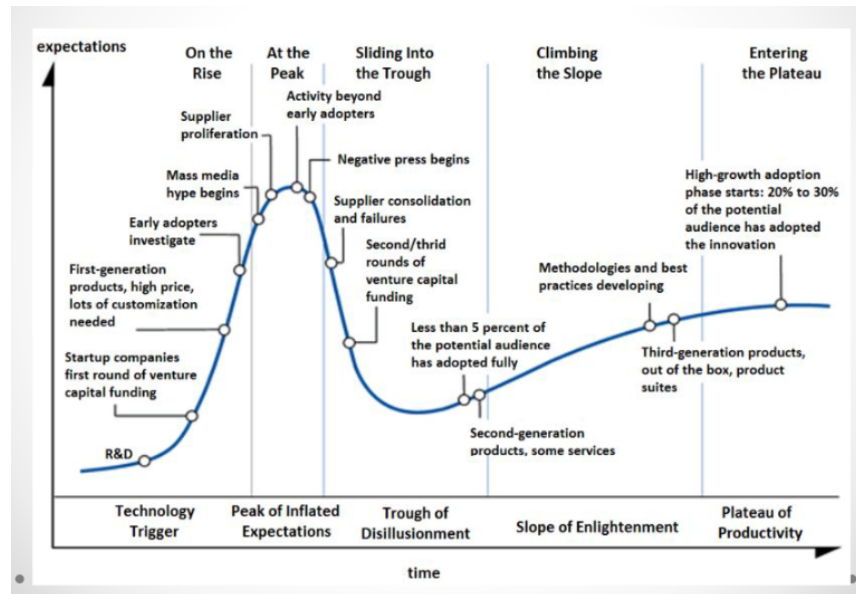
3. Legge di Reed “L'utilità delle grandi reti, formate da reti di reti (con particolare riferimento alle reti di relazione sociale) cresce esponenzialmente con la dimensione della rete”. Con n numero di nodi ho $2^n - n - 1$ possibili sottogruppi.

4. Conseguenze

- Distribuzione di solo contenuto, valore lineare. 1
- Attivazioni di comunicazioni 1 a 1, valore quadratico. 2
- Comunità, social network, valore esponenziale. 3

Risulta quindi molto conveniente puntare sulla costruzione di comunità piuttosto che su sistemi di broadcast

1.3.3 Hype Cycle di Gartner



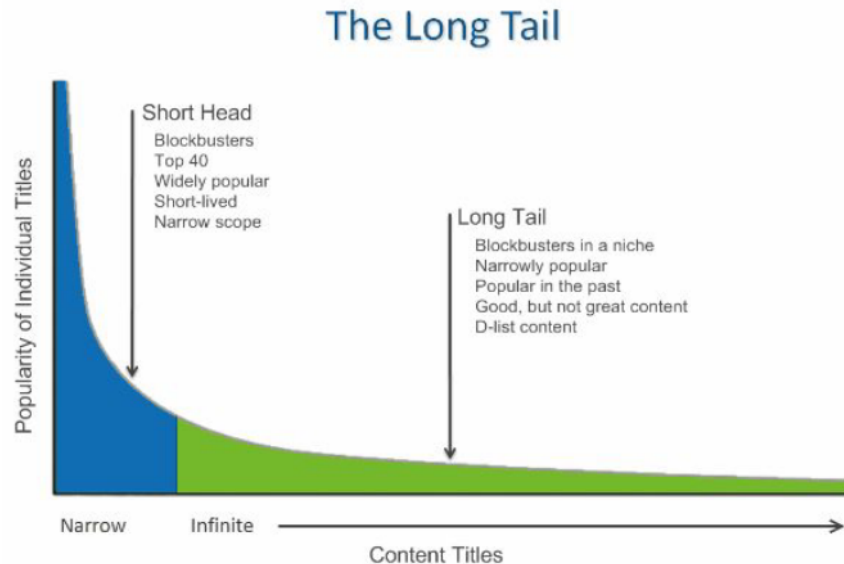
Il ciclo di vita di un progetto è diviso in fasi.

- **Technology trigger** ovvero i primi concept, l'interesse dei media con annessa pubblicità.
Spesso in questa fase non esistono prodotti utilizzabili.
- **Picco di aspettativa** molta pubblicità anticipata produce molto hype, molte aziende cavalcano questa onda e ne escono vittoriose, ma esistono decine di casi in cui questo non è vero.
- **Picco di disillusione** l'interesse svanisce quando i test e le implementazioni non soddisfano le aspettative, gli investimenti continuano solo se l'azienda riesce a riprendersi.
- **Risalita dall'imbarazzo** molti esempi dimostrano come questa tecnologia sia utile.
- **Altopiano della produttività** il mercato si assesta e la tecnologia diventa di uso comune.

1.3.4 Magic Quadrant di Gartner



1.3.5 Long tail di Anderson



Le grandi aziende possono arrivare a ricavare la maggior parte del loro fatturato dalla long tail, e.g. Amazon il 57%.

Nei nuovi mercati la coda ha un valore molto più alto, mentre i rivenditori sono focalizzati nella parte sinistra della distribuzione.

2 Concetti generali sull'informatica aziendale

2.1 Sistema Informativo Aziendale (SIA) prima definizione

Applicazione dell'informatica nella gestione dell'organizzazione di impresa. Ha come scopo la definizione di processi aziendali che inviano informazioni all'interno ed all'esterno dell'azienda nel momento in cui l'informazione è necessaria.

Il modo in cui il SIA interpreta le dinamiche interne e si interfaccia con il mondo esterno è scelto dall'azienda in modo tale da rappresentare in modo preciso i fenomeni che si vuole trattare.

Un SIA raccoglie dati in archivi organizzati, estrae informazioni tramite l'elaborazione di informazioni e distribuisce informazioni agli utenti.

2.2 Fenomeni che guidano alla costruzione di un SIA

La costruzione di un SIA è guidata da:

- L'identificazione di quali fenomeni, interni ed esterni, l'azienda vuole rappresentare.
- La modalità attraverso la quale dovranno essere rappresentati i fenomeni, quindi che tipo di memorizzazione e il grado di precisione.
- Dalla natura delle informazioni che l'azienda si aspetta di ottenere.

2.3 Elementi che compongono un SIA

I componenti di un SIA sono molteplici e legati ai fenomeni che un'azienda vuole rappresentare, possono però essere divisi in tre classi.

- **Dati** fungono come rappresentazione, attraverso i processi, di fenomeni aziendali elementari.
Vengono raccolti per essere memorizzati in modo ordinato, elaborati e usati per generare nuovi processi.
I dati possono essere divisi in varie categorie:
 - di configurazione, e.g. la valuta corrente o il nome dell'azienda
 - operativi, e.g. le informazioni relative allo stato di un progetto
 - di supporto, e.g. il nome del server web attraverso il quale un cliente ha effettuato una richiesta
 - di stato, e.g. il fatturato aziendale
- **Procedure** sono la parte dinamica di un sistema informativo, sono le azioni che un sistema informativo compie per aderire alla realtà.
Sono azioni che operano sui dati categorizzabili come:
 - Acquisizione
 - Controllo ed elaborazione
 - Pianificazione
- **Mezzi e strumenti per il trattamento delle informazioni** l'apparecchiatura fisica attraverso la quale il SIA opera per raccogliere/memorizzare/elaborare/restituire dati.

Esiste anche una parte organica che deve occuparsi del mantenimento del SIA, che è per sua natura dinamico e perciò soggetto ad evoluzioni.

2.4 Sistema Informativo Aziendale (SIA) seconda definizione

L'insieme dei dati, delle procedure, dei modelli organizzativi e dei mezzi adottati per utilizzare l'informatica all'interno dell'azienda.

2.4.1 Sistema delle informazioni aziendali

Insieme delle informazioni quantitative e qualitative ottenibili sullo stato passato/presente/futuro dei fenomeni aziendali.

Per cui la struttura del SIA definisce e vincola il sistema delle informazioni aziendali.

2.5 Fattori di evoluzione di un SIA

Come già detto i SIA sono dinamici e per ciò soggetti ad evoluzioni, ma cosa scatuisce questi cambiamenti? I fattori sono diversi e categorizzabili in due tipi di eventi:

- **Interni** miglioramento delle prestazioni, miglioramento della rappresentazione aziendale.
Quindi modifiche che portano alla creazione di nuovi processi.
- **Esterni** come i vincoli imposti dal mercato/dallo Stato/da un fornitore, quindi necessario solo all'agente esterno ma non all'azienda in se.

Spesso si tende a migliorare il SIA dove si vede un vantaggio o nel breve periodo, questo porta svantaggio in settori in cui il vantaggio è difficilmente quantificabile oppure sul medio/lungo periodo.

E' perciò fondamentale che lo sviluppo avvenga nel modo più armonico possibile e completare a tutti i bisogni dell'azienda.

2.6 L'impatto dei SIA

2.6.1 Conoscenza dei fenomeni aziendali

I SIA diffondono informazioni all'interno di un'azienda, ma ogni persona all'interno dell'azienda ha bisogno di conoscere diversi fenomeni, le differenze di necessità sono:

- **Livello di astrazione** le informazioni possono essere analitiche o sintetiche quindi ottenute elaborando diversi dati.

- **Tempestività** la velocità con la quale le informazioni viaggiano all'interno dell'azienda.
- **Livello di copertura** rappresenta quanti fenomeni vengono analizzati e da quanto tempo.

Il SIA deve quindi assicurare di raggiungere tutti i destinatari di un informazione in un tempo accettabile e senza errori.

Nelle aziende le informazioni viaggiano attraverso le due seguenti direzioni:

- **Orizzontale** quindi tra le varie aree aziendali definendo e sincronizzando i processi.
- **Verticale** attraverso l'organigramma aziendale, sintetizzando i dati all'occorrenza.

2.6.2 Processi classici

Questa è la categoria dei processi aziendali che hanno la maggior attrattiva dal punto di vista informatico, infatti da sempre i sistemi informativi vengono utilizzati per i seguenti scopi.

- **Sviluppo di funzioni operative**, infatti il primo scopo che un SIA ha è quello di automatizzare diverse azioni per:
 - Ridurre il costo del lavoro automatizzando alcuni processi.
 - Migliorare i processi definendoli meglio.
 - Aumentare la mole e la qualità dei dati raccolti.

La prima area ad essere tipicamente automatizzata è l'area amministrativa perchè soggetta a molte standardizzazioni dal punto di vista legislativo.

- **Pianificazione** la grande mole di dati che un SIA può raccogliere viene usata per pianificare delle future strategie ed avere una visione d'insieme di come l'azienda sta procedendo.
- **Controllo** l'immediatezza nella disponibilità dei dati porta ad una possibilità di intervento estremamente rapida.

2.6.3 Nuovi processi

L'introduzione di nuove tecnologie spinge le aziende a dover generare dei nuovi processi:

- **Organizzazione** magari proprio l'organizzazione interna dell'azienda viene ridefinita.
- **Flussi** delle informazioni viene modificato.
- **Modalità di trattamento dei dati** i dati raccolti dal SIA devono rispettare alcune legislazioni a livello statale o superiore.
- **Sistemi di comunicazione**

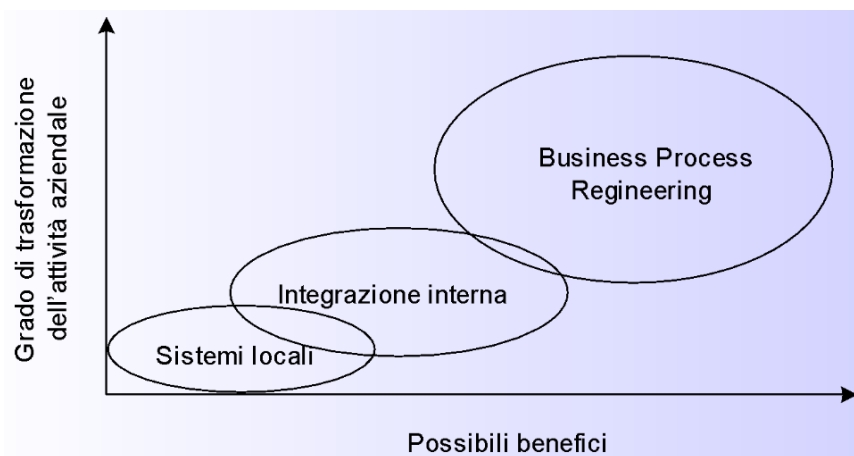
2.6.4 Business Process Rengineering (BPR)

La spinta nel ricercare nuove forme organizzative dopo aver adottato una tecnologia innovativa.

La tecnologia adottata serve ad esempio ad aumentare il numero e la qualità dei servizi offerti oppure offrire servizi di corredo rispetto al core business.

Un esempio di BPR potrebbe essere l'introduzione degli ATM, le banche non hanno aumentato il numero di sportelli ma introdotto il servizio di distribuzione contanti automatico.

Il servizio risulta quindi sempre attivo, e con il tempo si è arricchito di servizi come la possibilità di pagare le bollette.



2.7 Struttura aziendale italiana

Il tessuto sociale italiano ha favorito un maggiore sviluppo delle PMI che rappresentano il 99.9% delle imprese in Italia, mentre le grandi imprese sono solo lo 0.1%.

Sul totale della forza lavoro del paese il 78% lavora in aziende medio o piccole.

Nelle grandi aziende, necessariamente più strutturate, si è affrontata da tempo la necessità di migrare verso un SIA adeguato a loro, mentre nelle PMI questo cambio è iniziato dopo e in molti casi non ancora terminato.

Una PMI ha molte più difficoltà nell'adottare un SIA per svariati motivi, docuti soprattutto alla grande flessibilità e alla mancanza di una struttura ben definita.

Bisogna anche considerare che una PMI ha una struttura molto dipendente dal mercato in cui opera, è quindi necessario un SIA che si sappia interfacciare al meglio con la maggior parte delle casistiche.

2.7.1 Conseguenze per i SIA

I SIA per le grandi aziende si adattano molto male alle PMI, perchè necessitano di soluzioni più snelle e maggiormente configurabili.

I grandi produttori di sistemi informativi hanno prodotti pensati appositamente per le PMI.

- SAP: S/4HANA → Business One
- Microsoft: Dynamics-AX → Dynamics-NAV

I concetti generali e principali basici sono uguali, ma differiscono implementazioni e strutturazione diversa.

2.8 Cambiamenti organizzativi aziendali

La disponibilità di strumenti come i SIA ha favorito la standardizzazione e la velocizzazione dei processi classici.

Dall'altra parte la maggior quantità e qualità dei dati ha favorito un'enorme innovazione.

2.8.1 Organizzazione interna

L'evoluzione dei SIA e l'uso sempre maggiore di essi li ha portati ad essere una leva strategica non indifferente.

- Riduzione di personale amministrativo.
- Riqualifica dei ruoli in tutta l'azienda.
- Riduzione dei ruoli di supporto, come la segreteria e il mantenimento di archivi.
- Revisione dei processi di front office, e.g. le video conferenze e le mail.
- Revisione del modello organizzativo, si è passati da un'organizzazione per funzioni ad una per processi.
- Reperimento sempre più dettagliato di info sul mercato e sui clienti.

2.8.2 Organizzazione esterna

Permette di ripartizionare le risorse dell'azienda in luoghi molto distanti l'uno dall'altro.

Le grandi aziende hanno iniziato in processo chiamato “terziarizzazione” ovvero una delegazione ad aziende più piccole e molto specializzate, che a volte diventano delle vere e proprie succursali.

Per questi motivi una struttura che prima era monolitica diventa reticolare.

3 Struttura dell'azienda e del suo sistema informativo

3.1 Esigenza informativa

La funzione primaria di un sistema informativo è interfacciarsi con che fa funzionare l'azienda, o parte di essa, attraverso la propria attività.

In questa definizione ricadono praticamente tutti i componenti di un'azienda dal dirigente al semplice operaio, il sistema si deve quindi adattare a molte necessità differenti con livelli di astrazione differente.

Si conclude che i livelli operativi (operai, impiegati, ...) hanno bisogno di informazioni dettagliate e attuali, mentre man mano che si sale con i livelli decisionali sale anche il livello di astrazione e di sintesi.

3.2 Schema di Anthony



Questo modello distingue tra tre diversi livelli di ruolo, ognuno di questi livelli comunica con quelli adiacenti con cicli di pianificazione e controlli.

- **Alta direzione** identifica gli obiettivi primari a medio-lungo termine.
- **Direzione funzionale** si occupa di analisi economiche e previsioni a medio termine.
- **Personale esecutivo** attua i piani definiti.

	Frequenza	Dati	Provenienze (rispetto all'azienda)	Volumi
Alta direzione	sporadica	molto sintetici	interna ed esterna	bassi
Direzione funzionale	prefissata	sintetica	interna	medi
Personale esecutivo	continua	analitica	interna	elevati

3.3 Scomposizione del sistema informativo

Come si vede nell'immagine della piramide di Anthony il sistema informativo usato deve essere di due tipi differenti, ovvero operativo per le fasce più in basso e informativo per le fasce più in alto.

3.3.1 Sistema operativa

- Funzioni principali
 - Automazione di attività procedurali

- Definizione nuovi processi
- Aiuto nelle attività aziendali
- Raccolta di dati
- Guida per l'operatore
- Azioni sui dati
 - Accesso interattivo: lettura, scrittura dei dati puntuali
 - Aggiornamento puntuale dei dati
 - Elaborazione di eventi
 - Aggregazione per il calcolo di stato
- Componenti fondamentali
 - Vase di dati operativa
 - Funzioni operative

3.3.2 Sistema informazionale

- Funzioni principali
 - Informazioni sui risultati rispetto agli obbiettivi prefissati
 - Confronto tra indicatori aziendali e parametri esterni
 - Facilitare il processo decisionale
- Azioni sui dati
 - Aggregazione di grandi moli di dati
 - Analisi con profondità temporale
 - Ricerca per argomento anzichè per evento
 - Analisi multidimensionale

Il punto fondamentale dei sistemi informativi è la base di dati che in questa sezione si chiama “Data Warehouse”.

- Componenti fondamentali
 - Warehouse costantemente alimentato
 - Strumenti di analisi

3.4 Comparazione tra i sistemi informativi

Presentano ovviamente delle differenze, pertanto nel tempo si sono evoluti non come due sottocategorie di un unico sistema, ma bensì come due tipi di sistemi assestanti.

- **OLTP (On Line Transaction Processing)** pensati per supportare grandi moli di transizioni con molti utenti collegati, sono ottimizzati per gestire i processi operazionali di un'azienda.
- **OLAP (On Line Analytical Processing)** utili per analizzare in modo interattivo i dati, ottimizzati per l'elaborazione di grandi moli di dati e con molta flessibilità nelle interrogazioni.

	OLTP	OLAP
Finalità	Supporto all'operatività	Supporto al processo di decisione
Utenti	Molti, operai	Pochi, direzione
Dati	Analitici	Sintetici, solitamente numerici
Modalità di utilizzo	Guidata, passo passo	Integrazioni ad hoc
Dati per attività elementari	Bassa	Alta
Orientamento	Per processo/applicazione	Per area/tema
Aggiornamento dei dati	Continuo	Sporadico
Copertura temporale	Dati correnti	Storica
Ottimizzazione	Per accessi lettura e scrittura	Per accessi in lettura ed elaborazioni

4 Scelte organizzative

Ancora prima di poter iniziare ad organizzare un sistema informativo è necessario comprarne uno, ci sono 3 opzioni:

- **Make** costruire il proprio
- **Buy** comprarlo da un'azienda esterna
- **Outsource** farlo gestire completamente da un'altra azienda

4.1 Opzione make

Questa scelta organizzativa prevede la costruzione integrale di tutto il sistema informativo (software), questa opzione prevede l'instaurazione di un team di lavoro apposito.

Sebbene sia una scelta di nicchia rimane un'opzione nelle grandi aziende.

- **Costi fissi notevoli** infatti l'azienda deve formare il team per lo sviluppo.
- **Investimenti consistenti** l'azienda deve acquistare comunque tutto l'hardware e il software supplementare.
- **Struttura chiusa** il sistema così costruito non si dovrà mai interfacciare con il mercato per cui potrebbe diventare obsoleto in poco tempo.
- **Tempi di soluzione** il team per lo sviluppo potrebbe affrontare delle fasi critiche nella ricerca della soluzione migliore.
- **Mantenimento interno del know-how** viene mantenuto tutto all'interno dell'azienda quindi la riservatezza aumenta.
- **Soluzioni ad-hoc** tutte le soluzioni presenti nel SI saranno puntuali e fatte ad hoc per gli utilizzatori.

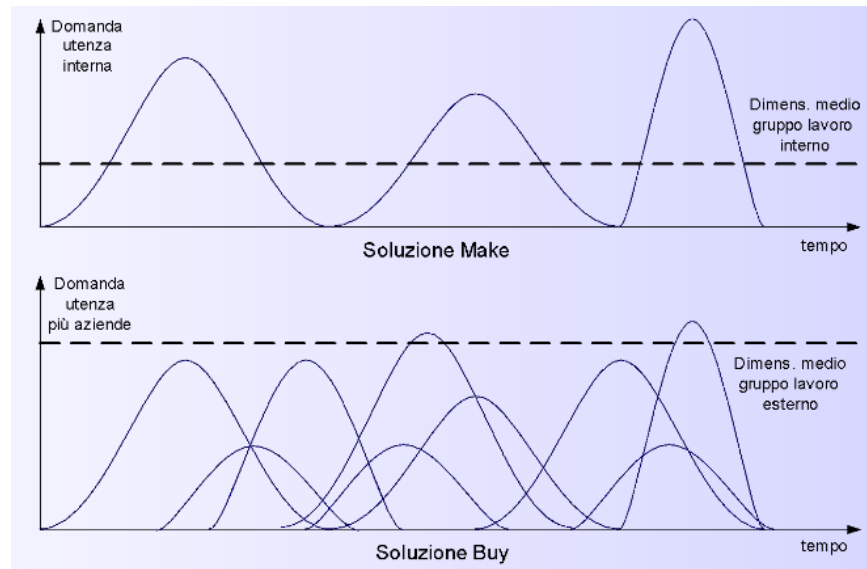
4.2 Opzione buy

Acquisto del proprio SI dall'esterno e costruzione di un piccolo team per il mantenimento ed il dialogo con la casa madre.

Costituisce la scelta tipica delle PMI, anche se gli ERP stanno diventando sempre più comuni nelle grandi imprese.

- **Costi fissi** seppur ridotto dobbiamo comunque costruire un team per mantenere il SI.
- **Parziale smobilitazione** l'azienda si deve dotare solo dell'infrastruttura necessaria.
- **Concentrazione sul core business** l'azienda deve occuparsi di una cosa in meno e quindi può dedicarsi al suo core business.
- **Dipendenza dall'esterno** per cambiare fornitore ci vuole molto quindi il potere della software house è elevato.
- **Maggior flessibilità rispetto a make** il fornitore essendo sul mercato dovrà essere sempre competitivo.
- **Fuorisce parte del know-how** perchè la software house deve conoscere i processi ed i dati interni all'azienda.
- **Non proprietà del software** gli investimenti fatti sul software non portano alla proprietà di esso.

- **Modello del SI mediato** non essendo l'unico cliente la soluzione proposta dalla software house non sarà ad-hoc.



4.3 Opzione outsource

Con questa opzione si delega completamente all'esterno la gestione e la creazione del SIA, tipicamente va pagato un canone annuale.

E' bene ricordare che outsourcing è diverso da:

- **Hosting** si affida all'esterno solo l'hardware ma il software rimane proprietario.
- **Body rental** utilizzo on demand di esterni per trasformare i costi fissi in variabili.

L'abbassamento dei costi di connessione, e la crescente esigenza di sicurezza posizione questa scelta come favorita tra le PMI.

- **Costi variabili** non c'è personale ma bisogna pagare un canone
- **Totale smobilitazione degli investimenti** non ci sono investimenti iniziali per l'hardware ed il software.
- **Completo vincolo con il fornitore**
- **Maggior flessibilità** rispetto alle altre opzioni questa è la più flessibile data la possibilità di scalare in ogni momento.

- **Fuoriuscita di tutto il know-how** il know-how dei processi e dei dati è completamente affidato a terzi.
- **Perdita di controllo** il fornitore ha un potere contrattuale elevatissimo.
- **Impossibilità di interventi diretti.**
- **Aderenza totale al mercato.**
- **Modelli mediati** essendo una soluzione ancora più generale l'acquirente ha poco spazio per soluzioni ad hoc.

4.4 Figure professionali nell'informatica aziendale

Le figure professionali “informatiche” che si trovano all'interno di un'azienda dipendono dal grado di maturità informatica dell'azienda stessa.

4.4.1 Livello 1

Il team informatico è costituito da un insieme di persone molto ristretto, con competenze molto diversificate e solo parzialmente dedicate al campo IT dell'azienda.

Si occupano, in modo indistinto, di quasi tutto, questo è tipico delle fasi iniziali di automazione dell'azienda.

In questa fase il team è caratterizzato da una struttura puramente orizzontale, con l'assenza di un budget ben definito e degli obbiettivi specifici.

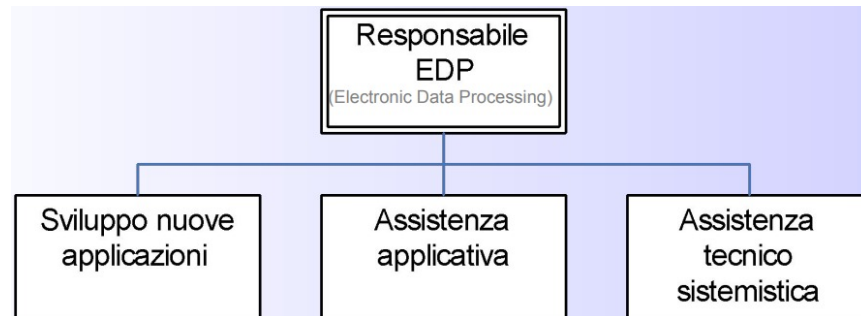
Purtroppo il team ha una visione miope dell'organizzazione e del ruolo del IT nell'azienda.

4.4.2 Livello 2

Comincia a formarsi un'organizzazione all'interno del team, e quindi con delle funzioni di responsabilità, il responsabile EDP (Electronic Data Processing). Si delineano anche altri ruoli, sottostanti ad esso.

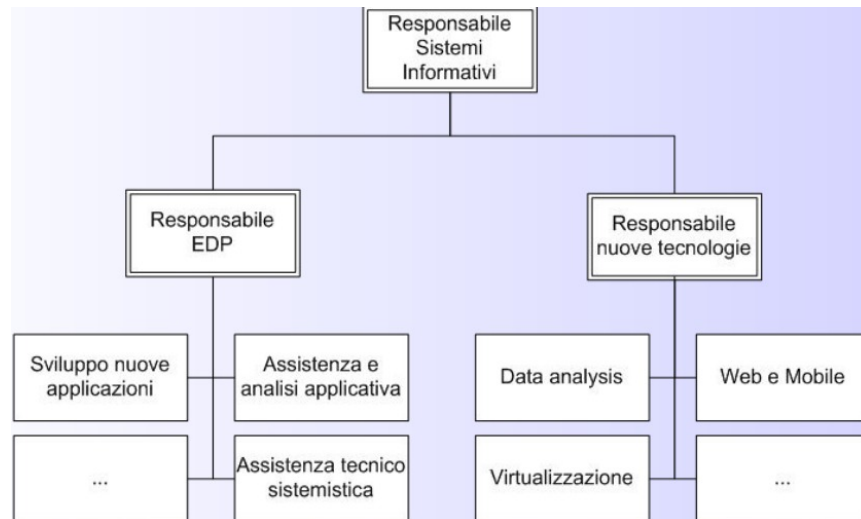
- Il sistemista che gestisce l'infrastruttura tecnologica.
- L'analista, che supporta gli utenti e fornisce indicazioni per nuove funzionalità e implementazioni.
- Il programmatore che si occupa della scrittura del codice (non presente nel modello buy).

Può succedere che i ruoli di programmatore e analista siano fusi in un unico individuo, l'ingegnere del software.



4.4.3 Livello 3

In questo livello la struttura del team diventa ancora più definita, viene riconosciuta una vera e propria direzione, quindi il team assume rilevanza a livello strategico per i progetti dell'azienda.

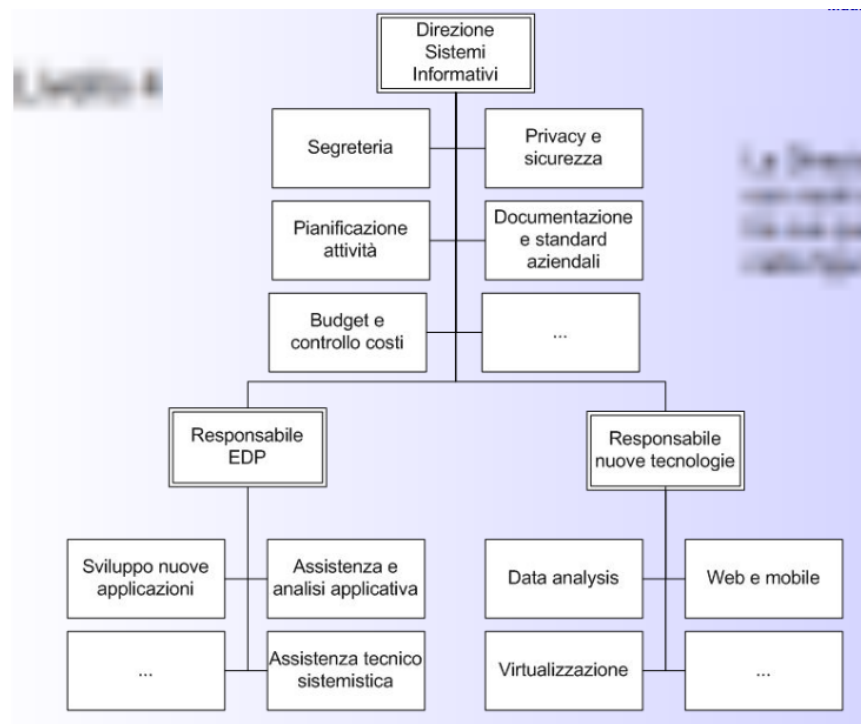


4.4.4 Livello 4

Questo rappresenta lo strutturamento più complesso, la direzione si struttura con ruoli di supporto quali:

- Segretaria

- Pianificatore
- Privacy e sicurezza
- Documentazione e standard
- Gestione budget



4.5 Posizionamento nell'organigramma

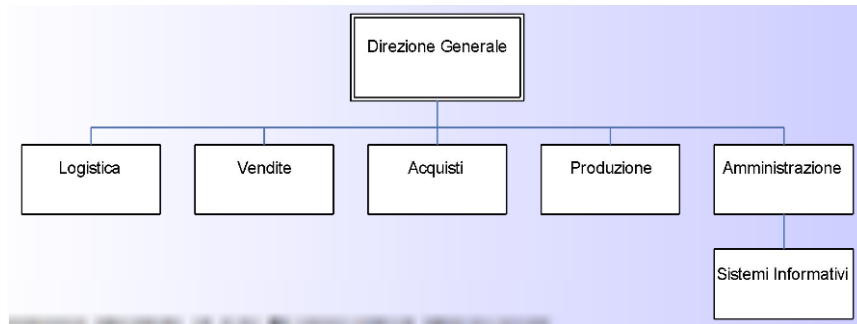
Oltre all'identificazione è importante saper posizionare i vari settori all'interno dell'organigramma aziendale.

Si possono identificare tre macro categorie principali, anche se spesso si trovano soluzioni miste.

4.5.1 Supporto amministrativo

Analogo all'attuale amministrazione, ha una visione obsoleta del SI.

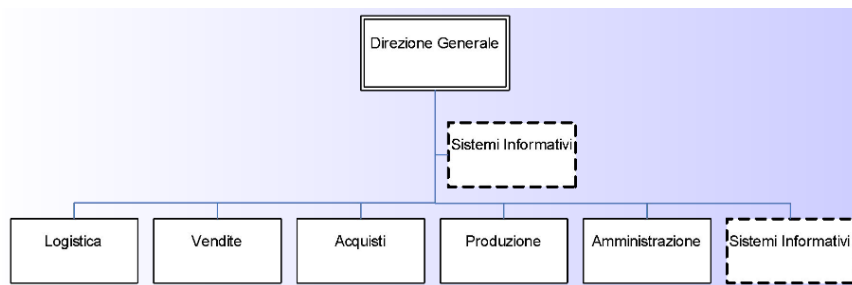
Il SI viene considerato solo come una banca dati in chiave quasi esclusivamente amministrativa per misurazione e controllo.



4.5.2 Servizio altre direzioni aziendali

Evoluzione in cui l'IT viene visto come a servizio di tutti, il servizio viene visto come in staff alla direzione generale.

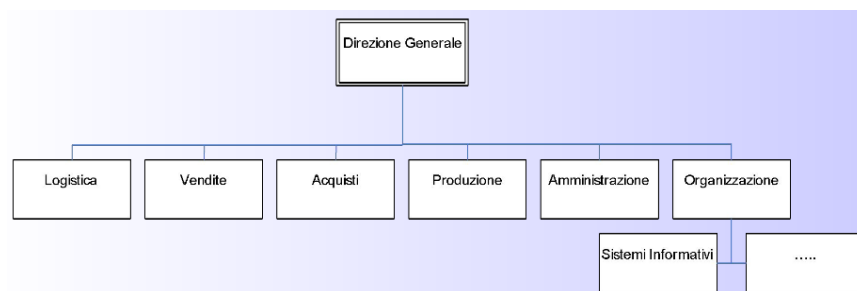
Serve a supporto della direzione nella definizione di processi e nelle analisi dei risultati ottenuti.



4.5.3 Servizio all'organizzazione

Il livello più elevato, viene identificata un'area specifica che si occupa dell'organizzazione complessiva di un'azienda.

Questo tipo di modello è tipico nelle grandi aziende, le quali sono sensibili alla formazione dei processi organizzativi.



4.6 Infrastruttura tecnologica dei SIA

Si è passati da un modello con unità centrale e molti terminali, ad uno client-server.

Questo passaggio è stato reso possibile grazie alla progressiva apertura verso internet che ha permesso la continua virtualizzazione.

Un SI è formato da diversi componenti generatori ma anche fruitori, come:

- Stazioni di lavoro
- Dispositivi mobili
- Dispositivi IoT

La normalizzazione di queste tecnologie con l'avanzamento del wireless ha reso possibile il collegamento da un qualsiasi punto geografico.

Questo però non ha portato solo pro, ma anche molti problemi come la difficoltà nella gestione delle varie infrastrutture e il crescente bisogno di security in ambiente informatico.

4.6.1 Da server locale a cloud (outsourcing)

Con la crescente disponibilità di internet si è passati ad un modello in cui è rimosso l'investimento in hw e sw per favorire un investimento a canone fisso.

Si va in outsourcing del SI, viene tutto delegato ad una ditta esterna, questo fa abbassare i costi iniziali permettendo di fatto anche alle PMI di accedere a certi servizi.

C'è un enorme problema però, si è costantemente dipendenti dalla connettività internet.

4.7 Interrompibilità del servizio

Il servizio del SI può essere interrotto per svariati motivi, sia hw che sw.

4.7.1 Guasti Hardware

Per i vari guasti hardware il tempo di ripristino può diventare troppo lungo, per cui esistono svariate misure di prevenzione:

- Sistema ridondato con RAID
- Hot swap dei dischi

- Sistema fault tolerant per sistemi su cui girano applicazioni “mission critical”
- Back-up

4.7.2 Guasti software

Ovviamente, neanche il codice di un SI è perfetto e può essere soggetto a malfunzionamenti.

A questo tipo di errori è imputabile la maggior parte dei malfunzionamenti, comunque questi sono anche i guasti meno gravi perchè raramente bloccano tutto il sistema ma solo un processo o un insieme di attività.

- Il fornitore deve fornire un servizio di assistenza.
- L'azienda dovrebbe usare un servizio di back-up con storicizzazione.

4.7.3 Guasti dolosi

Sono guasti particolarmente relativi all'ambiente (bancari e sanitari), possono arrecare interruzioni prolungate del servizio ma anche tipi di danni: furto monetario, di proprietà intellettuale e di segreti aziendali.

Gran parte di questi danni è causata da personale interno o da terzi con accesso alla struttura, è per cui importante installare sistemi di riconoscimento come : password, OTP e lettori biometrici.

5 Sistemi operazionali

5.1 Finalità dei sistemi operazionali

Sono SI orientati alla processazione di attività quotidiane, le loro funzionalità basiche sono:

- Registrare transazioni.
- Pianificare e controllare operazioni.
- Acquisizione e organizzazione della conoscenza.
- Elaborazione della situazione aziendale.

Il sistema operativo si deve quindi essere costruito attorno ad una base di dati strutturata e ben ordinata.

5.1.1 Registrazione delle transizioni

Una transazione è un'azione atomica, che rappresenta una o più azioni basiliche che l'azienda ha interesse nel tracciare.

Esistono due tipi di transizioni:

- **Semplici** registrate all'interno del SI come un singolo dato.
- **Complesse** composte da una serie di registrazioni elementari collegate tra di loro (spesso correlate a documenti fisici).

Una transizione non è sempre autocontenuta e può, a sua volta, generare altre transazioni a cascata.

Ogni azienda genera transazioni ma non è obbligatorio registrarle tutte nel SI, dipende anche dal volume di azioni che devono essere registrate.

Il volume dei dati (legati ad una transizione) è direttamente influenzato dalla struttura dell'azienda stessa, perchè più passaggi vengono compiuti più le informazioni legate ad essa crescono, di fatto ogni passaggio è una transazione.

5.1.2 Pianificazione e controllo delle operazioni



Come visto i processi sono concatenati tra di loro, quindi potremmo usare i dati dei processi a monte per pianificare le attività a valle.

Entra quindi in gioco l'uso dei SI, che rende possibile l'utilizzo di modelli di pianificazione complessi con un monitoraggio continuo dello stato di avanzamento.

Le funzionalità che i processi di organizzazione devono avere sono piuttosto articolate:

- Elaborare dei piani (ottimizzazione risorse e sincronizzazione dei passaggi).
- Registrare l'avanzamento delle operazioni.
- Misurare gli scostamenti rispetto agli obiettivi.

5.1.3 Organizzazione della conoscenza

Lo scopo dei sistemi operazionali è l'archiviazione di dati, questi dati, o informazioni di supporto, vengono trattati in modo centralizzato così da fornire a diversi settori aziendali i dati sempre aggiornati.

Le basi di conoscenza aziendali sono:

- Registrazione delle attività.
- L'anagrafica (fornitori, clienti e prodotti).
- Informazioni che correlano le anagrafiche.

Le informazioni mantenute in modo centralizzato devono pertanto essere strutturate, cioè riconducibili ad un insieme di caratteristiche predefinite e correlate.

5.1.4 Elaborazione della situazione aziendale

Dobbiamo pensare ad un SIA come ad un sistema dinamico, gli archivi ne sono la struttura, gli eventi che generano transazioni gli input e le pianificazioni la parte dinamica.

L'ultima parte modifica lo stato del sistema in base allo stato attuale e agli input.

Conoscere lo stato corrente permette di agire su stime aziendali, tramite appositi eventi, che le guidano nella direzione appropriata.

Gli indicatori di stato possono essere:

- La giacenza degli articoli.
- L'elenco degli ordini non evasi.

5.2 Informazioni operative

Le informazioni operative vengono mantenute in modo ordinato all'interno di un archivio, che seppur fisicamente separato, virtualmente unitario.

Al suo interno si distinguono alcune informazioni omogenee in struttura, gestione e destinazione d'uso:

- **Movimenti:** sono transazioni semplici e finali come i movimenti in magazzino.
- **Documenti di processo:** transazioni complesse che tipicamente sono collegate a liste di oggetti utilizzate per definire flussi di azioni.
 - Testa della coda: informazioni generali della transazione.
 - Righe: dettagli riferiti ai singoli prodotti.
- **Informazioni di stato:** descrivono l'attuale situazione del SI e per riflesso la situazione dell'azienda, possono essere dati puntuali o derivati da elaborazioni di dati anagrafici e movimenti.
- **Informazioni anagrafiche:** sono la base della conoscenza organizzativa, descrivono le proprietà fisiche di un'entità fissa nel tempo o molto stabile, e.g. prodotti, macchinari e materiali.

5.2.1 Qualità dei dati

Lo standard ISO 8402-1995 definisce la qualità dei dati come “Il possesso della totalità delle caratteristiche che portano al soddisfacimento delle esigenze, esplicite o implicite, dell'utente”.

Gli utenti andranno ad interagire con i dati restituiti dal sistema, la qualità è tanto più alta quanto la rappresentazione dei dati è vicina alla percezione della realtà.

In misura ridotta la qualità è influenzata anche dalla progettazione del SI, in modo più specifico da come sono stati strutturati i dati e come il SI viene alimentato.

La qualità può essere inficiata da sottoinsiemi autonomi e non integrati e la mancanza di un inventario aggiornato nella base di dati.

La scarsa qualità dei dati è legata ad un impatto economico e organizzativo nefasto per l'azienda:

- Difficoltà nell'introduzione di nuove tecnologie e nella modifica di processi.
- Difficoltà nella costruzione di un data warehouse.
- Bassa soddisfazione di operatori e utenti finali.

5.2.2 Caratteristiche strutturali

Ogni tipo di informazione operativa ha caratteristiche proprie.

- **Aggregazione** cioè il grado di sintesi di un'informazione rispetto agli eventi registrati.
 - **Analitica** quando una registrazione corrisponde ad un solo oggetto.
 - **Aggregata** quando è ottenuta dall'elaborazione di più informazioni.
- **Tempificazione** fa riferimento all'arco temporale coperto dall'informazione.
 - **Puntuale** se fa riferimento ad un istante preciso.
 - **Cumulativa** si riferisce ad un periodo (anche breve come poche ore).
- **Dimensionalità** ovvero il numero minimo di parametri per estrarre un'informazione.

	Aggregazione	Tempificazione	Dimensionalità
Anagrafiche	Analitica	Puntuali	Unitaria
Movimenti/Documenti	Analitica	Puntuali	Bassa
Informazioni di stato	Analitica o aggregativa	Puntuale o cumulativa	≥ 2

5.2.3 Caratteristiche funzionali

E' l'insieme delle caratteristiche che inficiano sull'utilizzo dei dati all'interno del SI.

- **Completezza:** l'estensione con cui vengono raccolte e memorizzate le informazioni, si riferisce sia ad oggetti singoli sia ad un insieme di essi.
- **Correttezza:** è la corrispondenza tra dato e realtà, è condizionata dalla possibilità di inserire dati errati.
- **Precisione:** indica l'approssimazione con la quale sono inseriti i dati, spesso si parla di misure fisiche.
- **Omogeneità:** richiede che i dati all'interno di una struttura vengano trattati con le stesse funzioni di accesso.
- **Fruibilità:** è la semplicità con cui un utente riesce a recuperare e comprendere le informazioni di cui ha bisogno.

5.3 Potenzialità informatica

Le aziende affrontano lo sviluppo in modo graduale ed in base al rapporto tra costi e benefici.

La potenzialità di un'azienda è data da due fattori:

1. **L'intensità informativa** con la quale si indica la necessità di informazioni proprie dell'azienda, dipenda dal mercato in cui si opera.
2. **L'attrattiva informatica** ovvero la facilità ed efficacia dell'informatizzazione nei processi aziendali.

Bisogna considerare anche quanto è disposto il manager in questo settore.

5.3.1 Intensità informativa

Le aziende hanno tre principali fattori indicatori della necessità di adottare un sistema informativo.

1. Complessità

- **Dimensione** più è elevato il numero delle persone che lavorano maggiore è la complessità.
- **Area geografica** nella quale opera l'azienda.
- **L'appartenenza ad un gruppo** la difficoltà nel coordinare diverse aziende.
- **Diversificazione dei prodotti**

Le aziende con una complessità elevata devono elaborare molti dati e potrebbero anche decidere di adottare soluzioni organizzative, anche se questo si dovesse tradurre in un incremento dell'organico.

1. **Intensità informativa del prodotto:** ovvero il grado di guadagno che si avrebbe informaticizzando un prodotto/servizio, ci sono prodotti che si prestano meglio e prodotti che si prestano peggio. e.g. i numeri telefonici hanno un'intensità informativa elevata.
2. **Intensità informativa del processo:** più è articolato un processo più ha necessità di essere automatizzato, processi lineari hanno un'intensità bassa.

Un buon esempio di diagramma per comprendere l'intensità informativa di un'azienda è il diagramma di Porter-Miller, in cui le ascisse sono l'intensità dei prodotti mentre le ordinate l'intensità dei processi.



5.3.2 Intensità informatica

L'intensità informativa non è sufficiente, per questo abbiamo bisogno di un ulteriore indicatore che ci suggerisce quanto sarebbe automatizzabile un processo. I parametri da valutare sono:

- **Proceduralità:** un processo ha diversi gradi di strutturazione, ovvero una sequenza di azioni elementari che servono a completarlo.
alta proceduralità → elevata attrattiva informatica
- **Complessità:** indica la complessità delle azioni da compiere per eseguire un processo.

alta complessità → elevata attrattiva informatica

- **Ripetitività:** se un processo ha una proceduralità elevata è anche molto ripetitivo, ovvero viene ripetuto svariate volte in un breve lasso di tempo.

alta ripetitività → elevata attrattiva informatica

- **Volume:** la quantità di dati da elaborare.
alto volume → elevata attrattiva informatica

I processi con attrattiva maggiore sono quelli molto procedurali con complessità elevata, che si ripetono nel tempo e con un volume di dati da analizzare molto elevato.

5.4 Composizione dei sistemi informativi operazionali

Per sistema operativo si indica l'insieme di tutti i sottosistemi informativi (che trattano l'attività operativa) di un'azienda.

Questi sottosistemi posso essere più o meno integrati tra di loro, questo alimenta la mancanza di uno standard.

Di fatto i sistemi informativi operazionali sono differenti in base al settore in cui opera l'azienda.

Oltre ad una divisione dei sottosistemi i criteri di classificazione variano in base all'aspetto che si vuole considerare: per funzione, per processo, per architettura, ecc.

In passato si usava una classificazione in base al supporto fornito all'area funzionale.

Questo apporccio venne abbandonato quando è emersa la necessità di integrare i diversi sottoinsiemi.

5.4.1 Portafoglio operativo

L'insieme di tutte le applicazioni informatiche legate ai processi *market driven*, ovvero legati al core business..

Tutte le attività coinvolte nella produzione e nella vendita dei prodotti/servizi. Sono molto variabili al variare del settore di cui ci si occupa, ma anche all'interno dello stesso settore c'è molta varianza tra aziende di dimensioni diverse.

Si deve fare una distinzione anche sul processo produttivo di un'azienda.

E' molto frequente la presenza di software molto specializzato a livello settoriale per quanto riguarda la sua implementazione.

5.4.2 Portafoglio istituzionale

Contiene tutte le applicazioni informatiche realizzate per le attività di sostegno: contabilità, risorse umane, ecc.

Molte di queste hanno un'elevata attrazione informatica, perchè standardizzate da leggi, molto procedurali e ripetitive caratterizzate da semplicità di elaborazione.

Essendo normate dalla legge risulta parecchio simili al variare del settore e per questo il mercato offre molti software che sono standard.

5.4.3 Dai sistemi tradizionali agli ERP

Fino ai primi anni 90 le aziende erano divise in isole informatiche, ognuna delle quali soddisfaceva una particolare funzione aziendale.

Questa situazione era data da:

- Uno sviluppo incrementale e per aree del SI.
- Rigidità nella struttura aziendale.
- I produttori di SW producevano soluzioni molto mirate.

Sistemi così articolati avevano parecchi problemi nell'ottica dell'intera azienda.

- I diversi sistemi erano spesso prodotti da aziende diverse con diversi contratti.
- La divisione in insiemi implicava uno sforzo notevole per mettere in comunicazione gli stessi insiemi.
- L'autonomia che ogni insieme doveva avere produceva molta duplicazione di dati.
- Le differenze tra gli insiemi rendeva difficile avere una visione di insieme.

Da questi disagi nascono gli ERP (Enterprise Resource Planning), il cui principale obiettivo è la gestione ottimale delle risorse che un'azienda utilizza per svolgere le proprie attività.

I nuovi ERP sono sviluppati da una singola software house e condividono una base di dati comune a tutte le sezioni.

Gli ERP sono formati da sottoinsiemi integrati tra di loro dal punto di vista architetturale ma anche da quello virtuale (UI simili).

Gli ERP sono sviluppati a moduli così da dare la possibilità alle aziende di richiedere solo quelli necessari a loro.

5.4.4 Ambiti degli ERP

Al variare del settore nel quale oper l'azienda restano invariati gli impianti generali degli ERP e le procedure di supporto all'azienda.

I sistemi ERP hanno i seguenti flussi di base:

- **Amministrativo:** grazie ai vincoli legislativi e agli standard è uno dei primi ambiti in cui gli ERP hanno operato.
- **Logistico:** ovvero il trattamento dei materiali e la loro movimentazione.
- **Attivo o vendite:** la catena di processi che porta l'azienda ad interagire con i clienti.
- **Passivo o acquisti:** la catena di processi che fa interagire azienda e fornitori.
- **Produttivo:** flusso opzionale perchè dipendente dal settore in cui si opera, per ridurre complessità gli ERP vendono software specifico per il tipo di mercato.

5.4.5 Sistemi operazionali complementari

Gli ERP non esauriscono gli aspetti operazionali necessari ad un'azienda.

- Sistemi di **supporto primario** agli ERP servono a far sviluppare flussi di dati standardizzati alle aziende che non hanno un ERP.
- Le **estensioni** degli ERP che permettono agli ERP di scambiare dati con l'esterno.
- I sistemi **tecnici** che risolvono alcuni processi di natura specifica.
- I sistemi di **dematerializzazione dei documenti**.

6 ERP: l'area amministrativa

6.1 Obiettivi

Gli obiettivi che ha un ERP in questa sezione dell'azienda sono i seguenti:

- Rispettare le norme fiscali e civilistiche, adempiere a tutti gli obblighi di legge.

- Valutazione consuntiva dell'andamento aziendale (contabilità ordinaria/finanziaria/analitica).
- Valutazione presuntiva degli andamenti aziendali (budget e controllo della gestione).

Questi punti sono in ordine dal più usato al meno, questo perchè ci sono poche aziende con un livello di informatizzazione così alto.

6.2 Struttura di base

Le strutture di base su cui si articola il sistema informativo amministrativo sono le seguenti:

- Parte anagrafica:
 - Piano dei conti.
 - Clienti.
 - Istituto di credito.
- Parte transazionale:
 - Movimenti contabili.
 - Movimenti finanziari.
 - Movimentazione IVA.

6.2.1 Piano dei conti

Il piano dei conti è una struttura dati che rappresenta l'impianto logico del sistema contabile, è una struttura gerarchica divisa ad aree di interesse.

6.2.2 Strutture anagrafiche speciali

In ogni SI alcune voci contabili hanno informazioni aggiuntive, sviluppate con l'ausilio di strutture particolari.

Le principali sono i clienti, i fornitori e gli istituti di credito.

1. Clienti Le informazioni anagrafiche di un cliente sono raggruppate in sottoinsiemi.
 - **Anagrafica di base:** ovvero tutte le informazioni di un cliente non legate a flussi aziendali.

- **Anagrafica finanziaria:** i dati usati per gestire un cliente dal punto di vista amministrativo (modalità di pagamento/codice cliente/ecc.).
 - **Anagrafica commerciale:** ovvero le informazioni che regolano il rapporto commerciale (sconti/spedizione/ecc.).
2. Fornitori Presentano informazioni analoghe ma simmetriche, adatte ad un flusso più passivo anziché attivo.
 3. Istituti di credito Le informazioni anagrafiche legate alle banche si dividono in due categorie:
 - (a) Istituti dei clienti/fornitori, per questi è sufficiente avere il BIC/SWIFT/IBAN.
 - (b) Istituti propri, per i quali è importante avere riferimenti contabili, remote banking, indicazioni sul tipo valuta e commissioni.

Solitamente i sistemi ERP per le PMI si fermano qua, lasciando a moduli speciali la gestione finanziaria.

6.2.3 Movimenti contabili

1. Struttura dei movimenti contabili I movimenti contabili hanno quattro parti, due delle quali principali.
 - **Testa del movimento contabile:** contenente le informazioni che valgono per tutto il pagamento come: data, causale e tipo di movimento.
 - **Righe del movimento:** in cui sono presenti le informazioni proprie di ogni riga come: valuta, importo e segno della registrazione.

Nei sistemi più raffinati ci sono anche le ulteriori due.

- **Competenze:** indica il periodo nel quale il movimento è stato effettuato.
 - **Riferimento:** ovvero l'indicazione della provenienza di un'operazione.
2. Struttura dei movimenti finanziari Sono previste solo la testa e le righe.
 - **Testa:** come il precedente, contiene: soggetto, debito o credito, tipo di movimento, ecc.
 - **Righe:** come le precedenti, contengono: scadenza, tipo di pagamento, valore riga, ecc.

3. Movimenti IVA Dovendo rispettare le norme vigenti i casi particolari si sprecano.
 - esenzioni
 - Iva intracomunitaria
 - Plafond IVA
 - ecc.

6.3 Procedure di base

6.3.1 Procedure di alimentazione

I sistemi amministrativi forniscono due metodologie di alimentazione.

1. Alimentazione interattiva Consiste nell'inserire a mano i dati all'interno di un form guidato.

I sistemi sono molto intelligenti e fanno molti controlli:

- **Congruenza dei dati:** le transizioni sono fatte in modo atomico.
- **Correttezza dei dati:** i sistemi fanno una serie di controlli all'inserimento.
- **Persistenza:** una volta fatto il submit la movimentazione non è più annullabile o modificabile.

Ad oggi la maggior parte delle alimentazioni non sono interattive.

2. Alimentazione automatica Le operazioni contabili sono generate automaticamente dai flussi, in particolare da:

- Fatture attive e passive.
- Operazioni finanziarie.

I SIA offrono procedure automatiche anche per:

- **Pagamenti/Incassi** sia da istituti che da privati.
- **Strumento di riconciliazione**, ovvero un controllo di congruenza con i dati in mano alle banche.

6.4 Flussi evolutivi

Sono delle procedure che solo i sistemi più complessi e reffinati hanno.

6.4.1 Contabilità analitica

Fornisce un quadro più preciso sull'andamento dell'azienda, solitamente opera solo nella parte economica della contabilità ordinaria.

Scompone la contabilità ordinaria in più voci analitiche dette *centri di costo*.

1. Procedure di alimentazione Il sistema può essere alimentato in modo interattivo in modo autonomo o direttamente connesso al data-entry della contabilità ordinaria.
I sistemi di alimentazione automatici creano, i movimenti di contabilità analitica, alla creazione dei movimenti contabili.
Un altro modo di alimentazione automatica è il ribaltamento.
2. Procedure di controllo Il sistema fornisce un insieme di procedure autonome per analizzare e controllare i dati, solitamente elaborano la movimentazione elementare per intervalli di tempo.

6.4.2 Budget

Il budget definisce delle previsioni sugli andamenti aziendali, viene sviluppato sulla parte economica del piano dei conti ordinari o nel piano dei conti di analitica.

1. Procedure di alimentazione Solitamente il sistema di budget può essere alimentato in maniera manuale o semiautomatica.
La modalità manuale consiste solamente nell'inserimento manuale di entry, mentre in modalità semiautomatica si ha un maggior aiuto da parte del sistema.
I sistemi più raffinati hanno una fase di alimentazione più complessa, si parte da un piano vendite storico e in base a dei meccanismi correttivi si passa ad un calcolo di costi diretti/indiretti.
2. Procedure di analisi e controllo Sono incluse delle procedure di analisi e controllo, solitamente l'elaborazione dei movimenti elementari è divisa in intervalli di tempo

6.4.3 Controllo di gestione

E' un argomento molto ampio e complesso, spesso diviso in sottoinsiemi autonomi e dedicati.

L'obiettivo è offrire una visione dell'andamento aziendale prima che i documenti contabili vengano chiusi.

Questo controllo è spesso trattato con due metodi, spesso integrati:

1. **Conto economico organizzativo:** confronta redditività del singolo cliente e prodotto.
 2. **Confronto budget e consuntivi:** confronta per ogni periodo di analisi i valori previsti nel budget con quelli consuntivi.
-
1. Cespiti Il costo per l'acquisto dei beni materiali o immateriali deve essere scaricato sul conto economico al raggiungimento degli obiettivi aziendali.
Questi periodo di scarico sono generalmente più lunghi e che idealmente rispecchiano la vita del prodotto.
I SI supportano la maggior parte delle procedure per il trattamento dei cespiti e della documentazione fiscale connessa.
 2. Compensi a terzi Quando un'azienda richiede servizi a terzi, questi addebitano all'azienda delle fatture.
Le procedure sono complicate da particolarità legate a normative o dal punto di vista operativo.

7 ERP: l'area logistica

7.1 Obiettivi

Questa parte del SI si occupa principalmente di:

- Definire caratteristiche gestionali.
- Controllare la movimentazione degli articoli.
- Analisi di disponibilità.
- Fornire valutazioni inventoriali.

Nei sistemi più evoluti

- Identificare le ubicazioni fisiche dei prodotti.
- Tracciare le origini dei prodotti.
- Operare con movimentazioni automatiche.

7.2 Struttura di base

La struttura di base su cui si basa la logistica sono:

- Le anagrafiche dei prodotti.
- Il layout aziendale.
- La movimentazione.

7.2.1 Nominazione articoli

Le aziende devono passare per un piano di codifica, nel quale vengono definite un insieme di caratteristiche che (nell'universo aziendale) permettono di identificare un articolo.

La codifica semplice può essere fatta in due modi:

1. **Lineare:** una semplice stringa di lunghezza fissa con caratteristiche comuni a tutti gli articoli.
2. **Condizionata:** una stringa di lunghezza variabile in cui l'attributo x è influenzato dall'attributo $x - 1$.

Le codifiche possono essere parlanti (con codici umanamente capibili) oppure strutturata (concisa ma capibile solo dalla macchina).

7.2.2 Anagrafiche dei prodotti

Il SI raggruppa le informazioni dei singoli prodotti, le informazioni possono essere:

- **Base:** codice, descrizione, imballaggio e gestione.
- **Approvvigionamento e produttive:** lead time, scorta minima e livello di riordino.

7.2.3 Layout aziendale

Per il SI che si occupa di questa area è molto importante conoscere il layout aziendale.

Può così categorizzare l'ubicazione degli articoli, in magazzini interni o esterni, di proprietà o di terzi.

L'azienda potrebbe anche avere un insieme di politiche da adottare per ogni categoria di oggetto, come codice, descrizione e valorizzazione.

7.2.4 Movimentazione logistica

Il SI deve gestire tutti i movimenti che vanno ad alterare lo stato dei magazzini, questi comprendono ingressi e uscite per svariate ragioni.

7.3 Procedure di base

I SI offrono procedure di base per il controllo anagrafico e procedure di alimentazione.

Le procedure di alimentazione sono:

- **Interattive**
- **Basate sui flussi operativi**
 - Ingressi/uscite merci
 - Produzione
- **Basate sui flussi di controllo**
 - Verifiche inventoriali
 - Inventari a rotazione

Il SI fornisce anche procedure di analisi e controllo:

- Analisi di giacenza di articoli.
- Disponibilità degli articoli
 - G = giacenza
 - R = richieste
 - C = copertura

$$D = G - R + C$$

- Generazione dell'inventario.
- Generazione della scheda degli articoli.

7.4 Flussi evolutivi

I SI più evoluti offrono molte funzioni per il trattamento della logistica, ma in casi particolari si preferisce affidarsi a software specializzato.

Nel seguito si vedranno tre tipologie di magazzini gestibili con un ERP.

7.4.1 Magazzino a lotti

I lotti suddividono gli articoli per informazioni comuni legate alla movimentazione.

Questa tecnica si usa per prodotti molto impattanti sul mercato, questo perchè è possibile risalire a tutta la filiera produttiva.

Sono necessarie:

- Informazioni di nominazione.
- Informazioni logistiche.
- Informazioni di stato.
- Informazioni di tracciabilità.
- Informazioni fisiche.

Esistono delle procedure per l'alimentazione del magazzino a lotti:

- Ricezione dei materiali e quindi creazione lotti.
- Spedizione e quindi chiusura totale/parziale dei lotti.
- Controllo qualità e quindi movimentazione dei lotti.
- Creazione di punti intermedi con chiusura e successiva riapertura dei lotti.
- Movimentazione interna dei lotti.

Procedure di analisi e controllo:

- Lotti in scadenza.
- Stati dei lotti.
- Tracciamento.

7.4.2 Magazzino a matricole

I SI fanno differenza tra numeri di serie e matricole.

- **Numero di serie:** intervalli numerici collegati univocamente a insiemi di articoli.

- **Matricola:** numero che collega articolo a movimentazione, è indispensabile per esplicitare ogni singolo elemento mosso. Svincola quindi dall'uso di lotti.

Le strutture di riferimento e le funzioni di controllo sono analoghe a quelle dei magazzini a lotti.

7.4.3 Magazzino a celle

Normalmente la posizione di un articolo all'interno di un magazzino è statica e quindi definita dalla sua ubicazione.

In un magazzino a celle questa staticità non vale, può quindi cambiare nel tempo e ammettere dispersioni degli articoli in posizioni diverse del magazzino.

Il modello del magazzino deve perciò essere definito come un insieme di coordinate spaziali, le quali definiscono la posizione della cella di un oggetto. Per trattare magazzini di questo tipo i SI devono mettere a disposizione funzioni per la movimentazione che ottimizzano lo spazio percorso.

7.4.4 Magazzini automatici

Sono gestiti da sistemi specializzati che interagiscono con la parte logistica del ERP.

La movimentazione può essere totalmente automatica oppure avere delle parti manuali nello scarico e carico.

8 ERP: l'area vendite

8.1 Obiettivi

Gestisce le caratteristiche commerciali dei prodotti e dei clienti e di fornire un supporto informativo per le funzioni operative.

Nei sistemi più evoluti sono presenti funzioni per gestire:

- Gruppi d'acquisto
- Ordini aperti
- ecc.

8.2 Strutture di base

Le strutture di quale si avvale sono:

- Le anagrafiche di clienti e prodotti
- Le strutture dati come listini, scontistiche e rapporti
- I documenti di processo
- I piani di processo come i piani di spedizione e di fatturazione

8.2.1 Anagrafica commerciale

1. Clienti e contatti commerciali Sono le informazioni di natura commerciale erlative ai clienti.
2. Prodotti in vendita I prodotti in vendita, si dividono in due categorie: prodotti fisici e servizi.

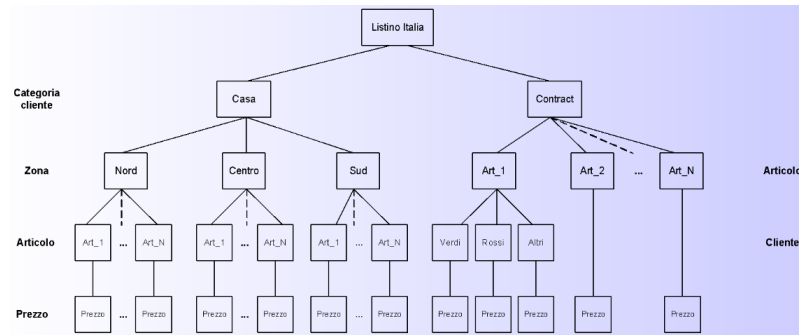
8.2.2 Condizioni commerciali

Sono tutte le relazioni che regolano i rapporti commerciali, i listini, le condizioni di scontistiche e le condizioni di agenzia.

1. Listini di vendita Sono la collezione dei prezzi degli articoli e degli elementi che ne concorrono alla modifica, come:
 - Il cliente
 - L'articolo
 - La valuta
 - La data nella quale viene effettuata la valutazione

I SI più semplici permettono una definizione tabellare del listino, mentre quelli più evoluti permettono di definire un algoritmo per inserire le scelte operative.

In generale questi algoritmi scendono un albero decisionale, oppure ad un livello superiore anche un grafo orientato.



I listini sono soggetti a non poche complicazioni.

- Il prezzo espresso come una maggiorazione/diminuzione rispetto ad un altro anzichè come valore puro.
- Definizione a scaglioni.
- Alta varianza in termini temporali.

2. Condizioni di scontistica Molte aziende preferiscono operare con listini semplici ma condizioni di sconto complesse, difficilmente si vedono sistemi misti.

Gli elementi che influenzano lo sconto:

- Cliente
- L'articolo
- la data in cui ci si trova

La lista delle scontistiche può incorrere in complessità se:

- Si verificano più scontistiche a cascata
- Gli sconto sono definiti a scaglioni
- Si utilizzano sconti composti

3. Condizioni di agenzia Definiscono i rapporti con chi fa da intermediario tra il cliente e l'azienda nelle vendite.

La logica di base è simile a quella dei listini e delle scontistiche, però qua gli elementi in gioco sono più numerosi.

Le condizioni sono legate a: clienti, articoli e modalità di pagamento. Come nei precedenti casi ci possono essere degli elementi che fanno aumentare la complessità:

- Applicazione a cascata di provvigioni.
- La definizioni di scaglioni per le provvigioni.

8.2.3 Documenti di processo

Tutti i documenti di processo hanno una struttura simile, si differenziano solitamente da dati non utilizzati o non definiti.

1. Testa dei documenti I dati fondamentali dei documenti possono essere scomposti in base alla collocazione logica
 - **Identificazione documento:** protocollo, data di inserimento e conferma, stato, ecc.
 - **Identificazione cliente:** cliente di riferimento, indirizzi aggiuntivi, id del cliente
 - **Condizioni commerciali:** cliente da fatturare, valuta, modalità di pagamento, ecc.
 - **Spedizione:** tipo di trasporto, aspetto esteriore del bene, ecc.
 - **Dati generali:** tipo di movimentazione logistica, tipo di IVA, ecc.
2. Righe dei documenti Sono i dati ripetuti e quindi più numerosi negli archivi.
 - **Identificatori riga:** numero di riga, stato della riga, ecc.
 - **Identificatori del prodotto in vendita:** codice articolo, codice commerciale, ecc.
 - **Spedizione:** colli, data di consegna, ecc.
 - ecc..
3. Piedi del documento Sono i dati totalizzanti, che possono essere estrapolati solo a valle della definizione.
 - Testi ed oggetti allegati
 - Condizioni finali
 - Sconto complessivo
 - Sconto finanziario
 - ecc.

8.2.4 Piani di processo

I piani di processo sono raggruppamenti in righe dei documenti usati per definire in blocco altri documenti utili.

1. Piano di spedizione Definisce una spedizione di prodotti che vanno mobilitati in un unico passaggio.
2. Piano di fatturazione Raggruppa tutte le righe dei DDT in un unico passaggio, anche in questo caso gli ERP mettono a disposizione delle funzioni apposite per la creazione.

8.3 Procedure di base

Le procedure di base al supporto delle vendite sono di due tipi, quelle che supportano il flusso esecutivo e quelle che forniscono le principali analisi.

8.3.1 Procedure di flusso

- **Preventivazione**
- **Raccolta ordini:** via diretta, differita o online
- **Spedizione:**
- **Fatturazione:** concettualmente automatico ma altamente critico.

8.3.2 Procedure di analisi

Tutti i sistemi ERP mettono a disposizione strumenti di reporting per analizzare il flusso attivo.

Solitamente gli elementi analizzati sono : ordinato, fatturato e marginalità.

8.4 Flussi evolutivi

In molti casi i processi di base non sono sufficienti a supportare le esigenze di un'azienda, soprattutto se l'azienda opera nella grande distribuzione.

Quando le esigenze informative sono spinte l'azienda si muove verso soluzioni verticali.

8.4.1 Condizioni commerciali

- **Gruppi d'acquisto**
- **Accordi quadro** Vengono stabiliti dei prezzi, condizioni di sconti, premi per vendite.
- **Orini aprti** Sono ordini che si trovano a monte, articoli che il cliente si è impegnato a comprare.

8.4.2 Processi

- Controllo e validazione
- Outsourcing delle spedizioni
- Meccanismi alternativi di emissione dei documenti

8.4.3 Estensioni degli ERP

Negli ultimi tempi si è creato il termine ERP II.

Riconoscendo una nuova fase, nella quale gli ERP dialogano sempre di più con servizi esterni.

1. CRM Trattano e organizzano i dati dei clienti e prospect.

- Allargare il portafoglio dei clienti.
- Conoscere il mercato e i bisogni.
- Mantenere traccia delle attività.
- Fidelizzazione clienti.

I sistemi CRM interagiscono con gli ERP nei punti estremi del flusso attivo.

I dati raccolti sono estremamente strategici per le aziende.

2. E-commerce E' un canale aggiuntivo e parallelo alla vendita, svincola dalla presenza di un operatore intermedio.

Spesso gli e-commerce fungono sia da B2B a B2C.

9 ERP: l'area acquisti

9.1 Obiettivi

Si occupano principalmente di definire le caratteristiche di prodotti, fornitori e servizi.

Nei sistemi più evoluti oltre a queste funzioni di base esistono procedure per trattare quanto avviene nel settore vendite (flusso attivo).

9.2 Strutture di base

- Anagrafiche commerciali dei prodotti e fornitori
- Le strutture per listini e sconti
- I documenti che registrano i fenomeni di transazione

9.2.1 Anagrafiche commerciali

1. Fornitori I fornitori si distinguono in tre categorie: di materiali, di lavorazioni e di trasporti.

Ad ogni fornitore sono associate le informazioni commerciali oltre a quelle anagrafiche di base.

9.2.2 Condizioni commerciali

Sono molto più semplici di quelle nei flussi attivi, devono solo contenere prezzi d'acquisto e sconti associati.

1. Listini d'acquisto Un obiettivo è registrare l'evoluzione dei prezzi nel tempo, l'unica complessità è data dall'esistenza di prezzi a scaglioni analoghi a quelli del flusso attivo.
2. Sconti Sono simili a quelli del flusso attivo, sono modellizzabili tramite un albero decisionale.

9.2.3 Documenti di processo

1. Struttura dei documenti di processo Come nel flusso attivo i documenti sono formati da una testa delle righe e un piede.
 - **Testa:** i dati comuni all'intero documento
 - Identificazione: protocollo, data inserimento, data conferma.

- Identificazione fornitore
- Condizioni commerciali: valuta , modalità di pagamento
- Spedizione: tipo di trasporto
- Dati generali: tipo di movimentazione logistica, movimentazione contabile e IVA
- **Righe:**
 - Identificatori della riga
 - Identificatori prodotto o servizio
 - Spedizione
 - Logistica
 - Amministrazione
- **Piede:**
 - Condizioni commerciali: i vari sconti
 - Totalizzatori: totale documento, totale IVA e scadenza del pagamento

9.3 Procedure di base

Le procedure di base sono divise in due blocchi principali: quelle di supporto al flusso esecutivo e quelle che forniscono le analisi operative.

9.3.1 Procedure di flusso

Sono elencate nell'ordine cronologico in cui vengono eseguite: richiesta d'acquisto, preventivazione, emissione ordini, ricezione materiali e controllo fatture.

9.3.2 Procedure di analisi

Le analisi che i sistemi ERP offrono sono analoghe a quelle sui flussi attivi, reporting ordinato per vari campi.

Date le richieste molto mutevoli, vengono utilizzati dei sistemi informativi con metodologie di data warehousing.

9.4 Flussi evolutivi

Il flusso acquisti (passivo) ha problematiche simmetriche rispetto al flusso attivo, soprattutto se l'azienda ha attività prettamente commerciali o se opera nella grande distribuzione.

9.4.1 Condizioni commerciali

1. Gruppi d'acquisto Se l'azienda aderisce ad un gruppo d'acquisto parte degli acquisti dovrà seguire le direttive date dal capo gruppo.
2. Accordi quadro Che siano essi legati ad un gruppo di acquisto o trattati autonomamente.
3. Ordini aperti In modo simmetrico alle vendite, questi ordini definiscono quello che l'azienda si è impegnata a comprare.

9.4.2 Processi

1. Validazione Legata a due lemenete: la tipologia di acquisti e l'entità di essi.
Ci sono gerarchie a buyer autorizzati e firme digitali.
2. Controllo qualità Dopo aver superato un controllo quantitativo si va in un deposito "accettazione" nel quale vengono fatti controlli qualitativi approfonditi.
Gli ERP spesso estrapolano delle valutazioni sui fornitori, ovvero una stima sulla loro qualità considerando la puntualità e la conformità degli articoli.

9.5 Estensioni dell'ERP

9.5.1 SCM

La sincronizzazione tra fabbisogno dell'azienda e disponibilità dei fornitori è un punto delicato.

Esistono pertanto le SCM (Supply Chain Management), ovvero una catena di fornitura composta da più aziende coinvolta in un unico processo virtuale.

9.5.2 E-procurement

Un esempio è il CONSIP per gli appalti pubblici, solitamente gli ERP si interfacciano ad essi tramite servizi web.

Sono simmetrici agli e-commerce, forniscono funzioni per:

- Analizzare il fabbisogno di acquisto
- Gestire le gare
- Interagire con i fornitori

Mentre ai fornitori permettono di:

- Gestire le proprie offerte
- Aggiornare lo stato dell'ordine

10 ERP: l'area produttiva

10.1 Obiettivi

- Il trattamento dei processi di definizione del prodotto
- La pianificazione della produzione
- Il trattamento dei processi di produzione
- La prevenziopne e la consunticazione dei costi di produzione

10.2 Concetti generali

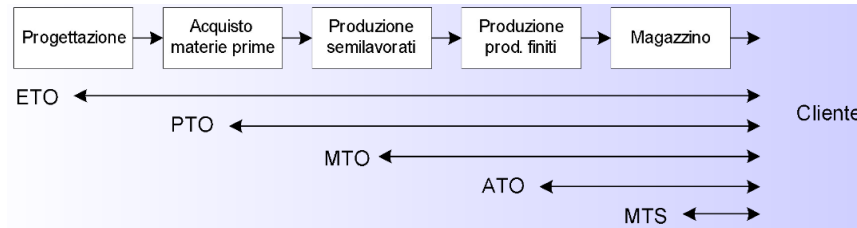
In azienda il termine *sistema produttivo* indica l'insieme delle risorse coinvolte nella produzione.

10.2.1 Criteri di classificazione

Esistono diversi criteri per classificare un sistema di produzione.

- Tipo di processo
 - Processi diretti
 - Processi continui
 - Processi ibridi
- Elementi prodotti
 - Elementi unitari
 - Bassi volumi
 - Medi volumi
 - Alti volumi
- Tempo di risposta dell'azienda alla domanda
 - Produzione per magazzino (Make to Stock MTS)

- Assemblaggio su ordine (Assemble to Order ATO)
- Produzione su Ordine (Make to Order MTO)
- Acquisto su ordine (Purchase to Order PTO)
- Progettazione su ordine (Engineer to Order ETO)



1. Sistemi di pianificazione delle risorse La crescente competitività ha creato la necessità di organizzare sempre meglio l'azienda.
Si sono creati due metodi:

- (a) MRP (Material Requirement Planning)
- (b) MRP II (Manufacturing Resource Planning)
Non ci limita a considerare la sola richiesta di materiali ma tutte le risorse del sistema aziendale.
Alcuni blocchi di MRP II sono informatizzati (MRP, CRP, ecc.) mentre altri, come la pianificazione strategica possono non esserlo.

10.2.2 Logiche organizzative dell'azienda

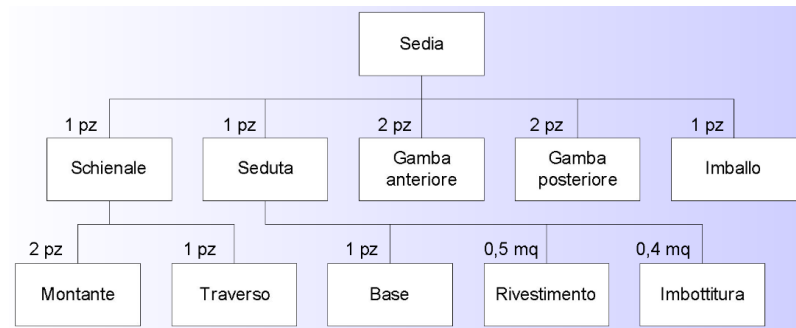
- **Decisioni a lungo termine:** strategie prese dall'alta direzione, definisce le strategie nel lungo termine da un punto di vista generale, di marketing e di produzione.
- **Decisione a breve / medio termine:** di competenza del management intermedio, si analizza la domanda (Demand Management), si pianifica l'aggregazione dei carichi e si stila un piano principale di produzione (Master Production Schedule MPS)
- **Decisioni operative:** si occupano del breve periodo e vengono prese dalla fascia operativa dell'azienda.
Il primo passo è l'analisi del fabbisogno dei materiali (Material Requirement Planning MRP), analisi delle risorse produttive (Capacity Requirement Planning CRP) ed infine la stesura del piano operativo di produzione con richiesta di aiuto o produzione.

10.3 Strutture di base

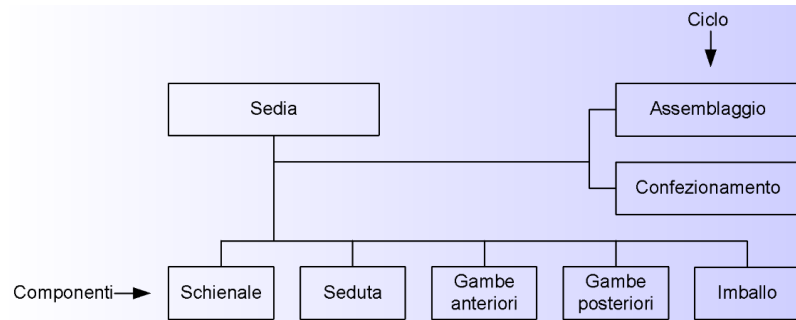
- Le strutture di tipo anagrafico che identificano l'organizzazione del prodotto
- Le strutture che ospitano i costi preventivi e consuntivi della produzione
- Le strutture di tipo dinamico che mappano l'evoluzione del processo

10.3.1 Strutture anagrafiche

1. Distinta base Il prodotto è rappresentato come uno schema gerarchico ad albero, la radice è formata assemblata con i nodi ad essa collagata, si prosegue così in modo ricorsivo fino alle foglie. Le foglie sono i materiali da acquistare.



- **Testa della distinta base:** codice e descrizione del padre, ecc.
 - **Righe della distinta base:** codice e descrizione del componente, ecc.
2. Ciclo produttivo Associato alla distinta base, ogni sistema informativo permette di definire il ciclo produttivo per trasformare i componenti nella radice della base distinta.



Di base sono presenti tutte le informazioni basilari: lavorazione, risorse impegnate, tempi di utilizzo, ecc.

3. Layout di fabbrica Si intende la descrizione della struttura organizzativa delle risorse.
4. Calendari I calendari indicano i giorni lavorativi e i turni a loro associati, permettono di calcolare disponibilità di ogni risorsa.

10.3.2 Strutture costi

Lo schema generale a cui si fa riferimento è il *roll-up* dei costi, ovvero si sommano in modo ricorsivo i costi dalle foglie alla radice.

Molti SI implementano anche il *roll-over* dei costi, in cui il costo si ottiene da dati esterni rispetto alla distinta base.

10.3.3 Dati dinamici

Sotto questo nome ricadono le strutture che tracciano il comportamento della produzione.

1. Ordini di produzione Gli ordini di produzione (OP) hanno una struttura simile alle operazioni relative ai materiali, quindi con testa e righe.
 - **Testa:** id e tipo di OP, tipo di processo, ecc.
 - **Righe:** Articolo, data prelievo, ecc.
2. Ordini di lavorazione L'ordine di lavorazione (OL) è la struttura che contiene le operazioni di produzione relative alle fasi operative, con una struttura testa-righe.

10.4 Procedure di base

- Trattamento anagrafico di base
- La generazione, manutenzione e consunzione delle commesse produttive per materiali e lavorazioni
- L'analisi delle commesse e dei costi collegati

10.4.1 Procedure di trattamento materiali

Il modello ottimale inizia con l'analisi della domanda e la funzione MPS, l'analisi MPR attinge le richieste direttamente dagli ordini opportunamente filtrati.

1. Definizione MPR Il sistema MPR genera un piano di produzione che è l'insieme degli OP necessario a coprire le richieste indipendenti. L'operatore può:
 - Spostare alcune richieste indipendenti.
 - Modificare i lead time.
 - Modificare i valori di scorta.

10.4.2 Procedure di trattamento lavorazioni

1. Lavorazioni interne Le lavorazioni interne vengono pianificate attraverso la procedura CRP nei seguenti modi:
 - CRP a capacità infinita basato su lead-time
 - CRP a capacità infinita basato sui tempi tecnici
 - CRP a capacità finita
 - CRP a capacità finita con ottimizzazioni

Le lavorazioni interne proseguono con i seguenti passaggi:

- Emissione della bolla di lavorazione
 - Emissioni di altra documentazione associata alle singole fasi
 - Avanzamento delle fasi
2. Lavorazioni esterne
 - C/lavoro di livello

- Si emette un documento di trasporto ricavato dall'OCL
- Il materiale esce e arriva al terzista
- Il terzista esegue OCL
- Il materiale lavorato rientra
- **C/lovaoro di fase**
 Situazione più complessa che tratta un ciclo con un'unica fase all'esterno.
 Nelle fasi intermedie il materiale che entra/esce non è determinato a livello OP e non è ben definito in anagrafica.
 I processi in entrata/uscita sono complicati dal tentativo di identificazione corretta dei materiali.

10.4.3 Procedure di manutenzione

1. Procedure di analisi L'effettivo avanzamento di OP e OL con valutazione di risorse e materiali.
2. Procedure di controllo Controllo costi con *roll-up* su OP e OL.

10.5 Flussi evolutivi

10.5.1 Pianificazione a medio/lungo termine

Necessaria nelle organizzazioni più complesse, molti sistemi utilizzano moduli specifici per gestire l'analisi della domanda oppure usano l'analisi aggregata dei carichi.

10.5.2 Schedulatori di produzione

- Pianificazione di produzione giornaliera
- Definire le sequenze ottimali di lavorazione
- Valutare i piani elaboratori
- ecc.

10.5.3 Rilevazione automatizzata dei dati da campo

10.5.4 Commesse cliente/impianto

- Anagrafica

- Dati generali
- Dati amministrativi
- Dati tecnici
 - Articoli non codificati
 - Lavorazioni non codificate
 - Costi generici
 - ecc.
- Sottocommesse
- Commesse produttive associate
- Valori associati

10.5.5 Sistemi tecnici

I sistemi, che a vari livelli, interagiscono con gli ERP.

- Sistemi per il trattamento e manutenzione degli impianti.
- Sistemi per la progettazione come CAD.
- Sistemi per l'automazione.

11 Sistemi informativi

11.1 Obiettivi

I sistemi operazionali hanno fatto nascere l'esigenza di nuovi strumenti per i livelli dirigenziali dell'azienda.

I sistemi operazionali hanno dei grandi limiti per quanto riguarda la reportistica:

- I dati estratti sono molto statici.
- La difficoltà e la lentezza dei processi di interrogazione
- L'incompletezza

I vantaggi dell'essere altamente customizzabili permettono di ospitare molti tipi di dato, porta molto overhead nei fogli di calcolo:

- La macchinosità dell'estrazione dei dati
- Lo scarso controllo sui dati
- Molti strumenti di calcolo personalizzati e incontrollati
- L'enorme complessità strutturale

Oltre a tutto questo i sistemi operazionali non sono ottimizzati per analizzare enormi quantità di dati e ottimizzano le procedure interattive.

I sistemi informativi devono pertanto mettere a disposizione dell'utente un ambiente in cui analizzare i dati sia facile.

- La base di dati deve comprendere l'intero insieme dei dati interessanti per l'azienda.
- I dati devono essere riorganizzati per supportare al meglio le attività di analisi.
- La base di dati deve essere strutturata in modo semplice
- Le fonti dei dati devono essere integrate
- Devono essere disponibili strumenti di analisi
- Gli strumenti di analisi devono supportare grandi moli di dati

11.2 Concetti generali

- **Data warehouse:** è la base dei dati informativi
- **Data warehousing:** è l'insieme delle attività che definiscono, costruiscono e mantengono la struttura
- **Decision Support System (DSS):** i sistemi informatici che aiutano nel processo decisionale come il supporto all'estrazione delle informazioni
- **Data Mining:** l'insieme degli strumenti e delle tecniche usate per estrarre dati "nascosti"
- **Business intelligence:** l'insieme delle attività volte a estrarre informazioni dai dati di business come quelli generati dai processi operativi
- **Knowledge management:** i sistemi di gestione della conoscenza aziendale che non è mantenuta solo dai dati strutturati/strutturizzabili

- **Big Data:** sono i flussi continui di informazioni non generate dai processi, e che non richiedono costi aggiuntivi, per essere generate. Il trattamento di questi dati ammette soluzioni diverse con elaborazione distribuita e algoritmi complessi. Si è passati dal dato che permette di osservare gli andamenti al dato dato che può essere usato come materiale.

11.3 Caratteristiche dei dati informativi strutturati

Ovviamente le informazioni operazionali differiscono di molto da quelle informative, come per esempio in:

- **Finalità:** fornire al management il substrato di informazioni per conoscere l'azienda. Descrivere il passato per aiutare a identificare i problemi e le loro cause, suggerire i cambiamenti da apportare per evitarli.
- **Struttura:** i dati sono articolati attorno a funzioni, procedure e eventi.
- **Utenti e strumenti:** sono utilizzati prevalentemente da personale esecutivo e manager.
- **Storicità dei dati:** vengono mantenute le informazioni con un arco temporale maggiore.
- **Dettaglio:** servono informazioni aggregate, è preferibile che esistano diversi livelli di aggregazione, dalla sintesi massima al dettaglio massimo.
- **Tipo di accesso:** è solitamente in sola lettura con aggiornamenti applicati nei periodi di inattività.

11.3.1 Modello multidimensionale

Uno stesso evento può essere analizzato mettendolo in relazione con diversi soggetti, da questa visione si intuisce che un dato può essere rappresentato come un punto in un *ipercubo*.

All'interno di questi ipercubi i punti verranno chiamati *fatti* mentre le coordinate *misure*, inoltre le coordinate di ciascun elemento costituiscono le *dimensioni* di analisi dei fatti.

1. **Fatti** Sono gli eventi che accadono all'interno dell'azienda e si ha interesse a misurare. Ogni fatto ha:

- Attributi che lo collocano nel tempo
- Almeno un dato numerico
- Informazioni descrittive

Un fatto è quindi identificato univocamente da una ennupla di coordinate.

2. Misure Sono caratteristiche numeriche del fatto che ne descrive aspetti quantitativi rilevanti per l'analisi. Le misure possono essere:

- **Effettive:** contenute nella base di dati informazionale
- **Calcolate:** a run time utilizzando i valori già presenti

I fatti elementari possono essere aggregati, ovvero si può ottenere un insieme di dati sintetici raggruppando le misure con caratteristiche comuni.

Il processo di aggregazione è fatto con appositi operatori com: somma, media, max, min, ...

Gli operatori devono essere scelti con attenzione, esistono infatti operatori aggregabili in alcune dimensioni e non su altre.

Il caso più comune è l'aggregabilità con l'operatore somma detta *additività*.

Le misure possono quindi essere divise in:

- **Misure di livello:** che esprimono valori propri del fatto nel momento in cui è registrato. Non sono mai additive rispetto alla dimensione del tempo.
- **Misure unitarie:** ovvero i valori relativi a dei soggetti, valido nel momento in cui viene registrato. Non sono mai additive.
- **Misure di flusso:** valori propri del fatto rapportati ad un intervallo temporale di riferimento.
Additive lungo qualsiasi dimensione.

Esiste infine una categoria speciale di misure, le *misure proprie*.

Queste sono tutte le misure di cui ci interessa è che siano verificate o meno, quindi la presenza/assenza del fatto.

Solitamente si esprimono con valori booleani oppure 0/1.

11.3.2 Caratteristiche strutturali

- **Multidimensionalità:** ovvero il numero di parametri che occorre fornire per poter estrapolare un'informazione specifica.

- **Granularità:** il grado di aggregazione delle informazioni, se l'ipercubo ha granularità minima significa che ogni fatto corrisponde ad un solo dato di origine.
Spesso per limitare le dimensioni del warehouse i dati hanno già un'aggregazione di base.
- **Arco temporale:** il periodo temporale coperto dai dati presenti nel warehouse.
- **Profondità storica:** i dati possono variare nel tempo, questo parametro identifica quante informazioni storiche su un parametro si posseggono.

11.3.3 Caratteristiche funzionali

- **Integrazione dei dati:** la coerenza che hanno dati provenienti da diverse applicazioni e da differenti basi di dati.
- **Accessibilità:** facilità d'uso e velocità di risposta.
- **Flessibilità e sintesi:** capacità di assecondare l'utente nel fare le richieste, mettere in relazione misure con criteri non prefissati.
- **Correttezza:** aderenza con la realtà.
- **Completezza:** la capacità di rappresentare tutti gli eventi importanti, si riferisce sia alle istanze dei dati sia al modello.

11.3.4 Data warehouse e data mart

Il data warehouse è definito come il magazzino unico, completo e consistente delle informazioni aziendali.

Contiene articoli di interesse aziendali dai quali è possibile estrarre gli ipercubi, per questa sua utilità può raggiungere dimensioni molto elevate.

Operare sull'intero data warehouse sarebbe impensabile, vengono così introdotti i *data mart*, ovvero dei piccoli data warehouse tematici limitati ai fatti che riguardano un'area.

I data mart possono essere definiti con particolari proprietà:

- Estensioni temporali ridotte.
- Granularità minore.

12 Data warehousing

12.1 Data warehousing e OLAP

OLAP (On Line Analytic Processing) identifica l'insieme degli strumenti atti ad aiutare il processo decisionale all'interno di un'azienda.

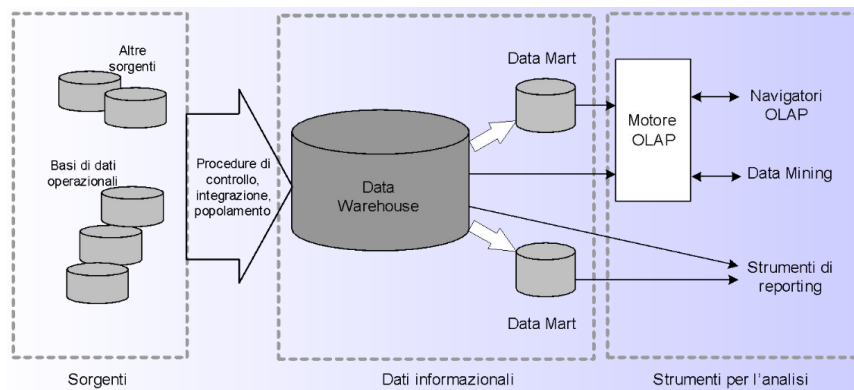
Esistono alcune regole per la definizione dei sistemi OLAP, la cosiddetta *FASMI*:

- Fast
- Analytical
- Shared
- Multidimensional
- Informatic

12.2 Architettura dei sistemi di data warehousing

I sistemi di DWH sono costituiti da DB posti a diversi livelli.

1. **Sorgenti:** db di origine dei dati, possono essere esterni o operazionali
2. **Staging Area (opzionale):** area intermedia usata per la trasformazione dei dati.
3. **Data warehouse:** db centrale, contiene tutti i dati necessari per le analisi.
4. **Data mart:** db multidimensionali su cui si appoggia l'analisi.



Esistono architetture con un numero variabile di livelli, quelle a 2 livelli non comprendono la staging area mentre le architetture a 3 livelli sì. Le soluzioni a 3 livelli sono spesso usate aziende più complesse, i sistemi a 2 livelli presentano elementi come:

- Un primo livello costituito dalle sorgenti dei dati.
- Il secondo livello contenente i dati informativi quindi dal DWH in poi.

12.3 Modelli concettuali per il data warehousing

I Sistemi informativi sono soggetti a molte evoluzioni nel corso della loro vita.

Solitamente un'azienda costruisce un nucleo contenente i dati di maggior interesse, i quali verranno poi ampliati.

Non esiste un metodo evolutivo standard, vedremo solo DFM.

12.3.1 DFM (Dimensional Fact Model)

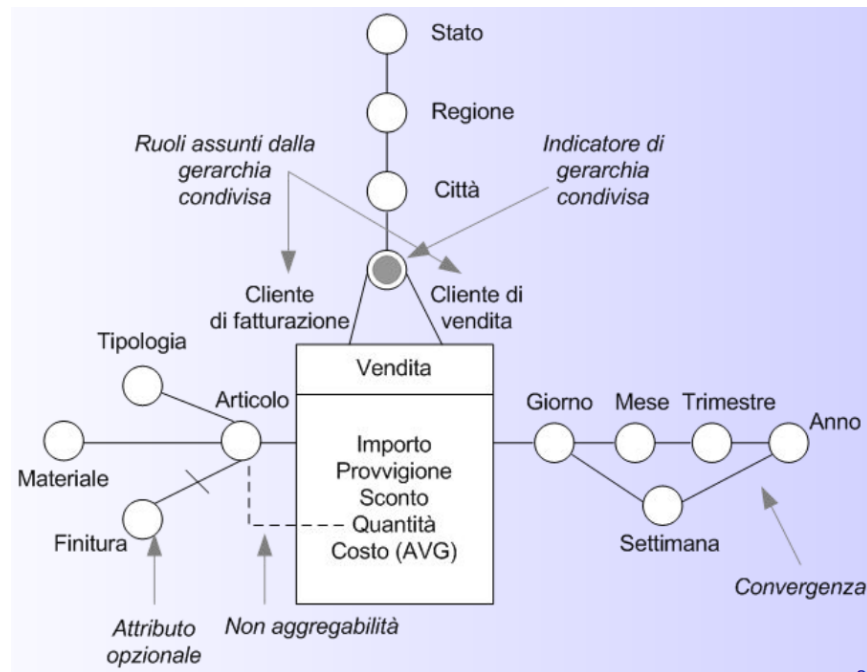
Fornisce una visione ad alto livello descrivendo graficamente i fatti attorno a cui si struttura un warehouse.

Ogni fatto è rappresentato tramite uno schema di fatto:

- **Fatto:** rettangolo contenente il nome del fatto e le sue misure
- **Dimensioni:** circoletti etichettati, vengono collegati ai vari fatti

Le gerarchie dimensionali sono alberi con radice nelle dimensioni di base, mentre i nodi sono gli attributi su cui la gerarchia è costruita. DFM permette di rappresentare alcune caratteristiche dei fatti:

- L'opzionalità di una o più dimensioni
- La presenza di gerarchie
- La convergenza
- Non aggregabilità



12.4 Modelli logici per il data warehouse

Nel momento in cui bisogna realizzare un warehouse si deve scegliere quale DBMS usare.

I dati possono essere memorizzati in db relazionali oppure in db multidimensionali come gli ipercubi.

Dobbiamo scegliere anche il tipo di interrogazione da fare:

- Motori di db relazionali come SQL.
- Motori multidimensionali tramite linguaggi come MDX di Microsoft.
- Elaborazione delegata ai client tramite linguaggi proprietari.

Dalla combinazione delle caratteristiche sopra citate nascono tre tipo di modelli:

1. Relational OLAP (ROLAP)
2. Multidimensional OLAP (MOLAP)
3. Hybrid OLAP (HOLAP)

12.4.1 ROLAP

Si basa su una struttura a db puramente relazionali interrogati tramite query SQL.

Risultano quindi molto compatte e con un diffuso know-how, bisogna però anche considerare la ridotta velocità per query con molte dimensioni e che le soluzioni (denormalizzazione e materializzazione) fanno aumentare la complessità di gestione e le dimensioni.

12.4.2 MOLAP

Con questo approccio i dati sono memorizzati come strutture multidimensionali, basta pensare a dei vettori.

Questa soluzione non ha molta popolarità dato il tasso di spazio occupato in cui solo il 20% è spazio utile, la mancanza di standard e il grande successo dei db relazionali.

12.4.3 HOLAP

Soluzione intermedia che combina i vantaggi delle presenti.

Il warehouse viene realizzato con un db relazionale così da essere mantenibile e scalabile facilmente.

Viene poi fatta una distinzione nei data mart, in cui i dati sono realizzati con db multidimensionali per avere una maggior efficienza nelle query e con un overhead dimensionale minore.

12.4.4 Schemi multidimensionali su db relazionali

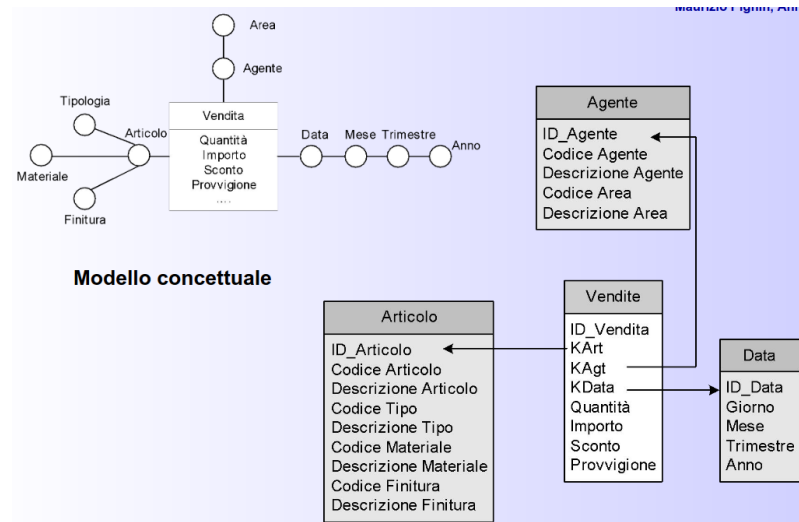
1. Schema a stella Nelle soluzioni ROLAP e HOLAP la modellazione logica segue lo schema a stella e le sue varianti.

Viene usata una tabella dei fatti in cui ogni elemento è un fatto elementare, per ogni misura propria del fatto viene inserito un campo di tipo numerico.

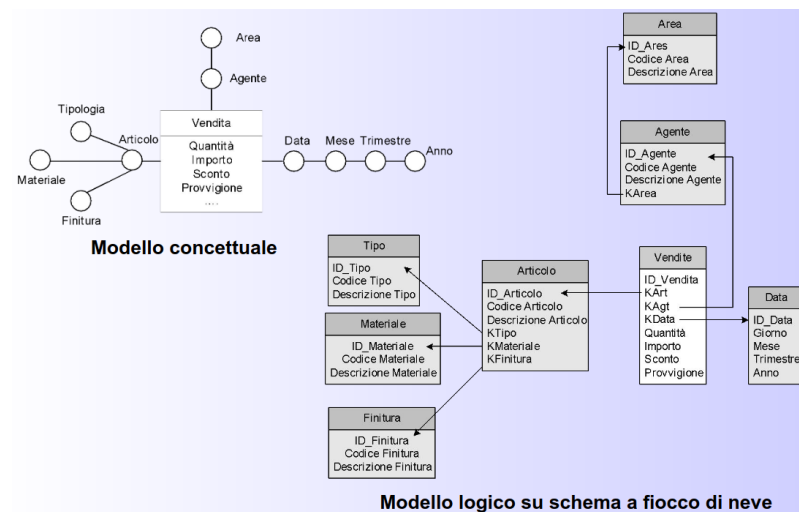
Vengono anche definite le tabelle delle dimensioni per ogni dimensione di base, queste tabelle sono soggette ad una denormalizzazione completa.

L'elevata denormalizzazione permette di fare un unico join per avere tutti i dati relativi ad un'unica dimensione, questo massimizza la velocità.

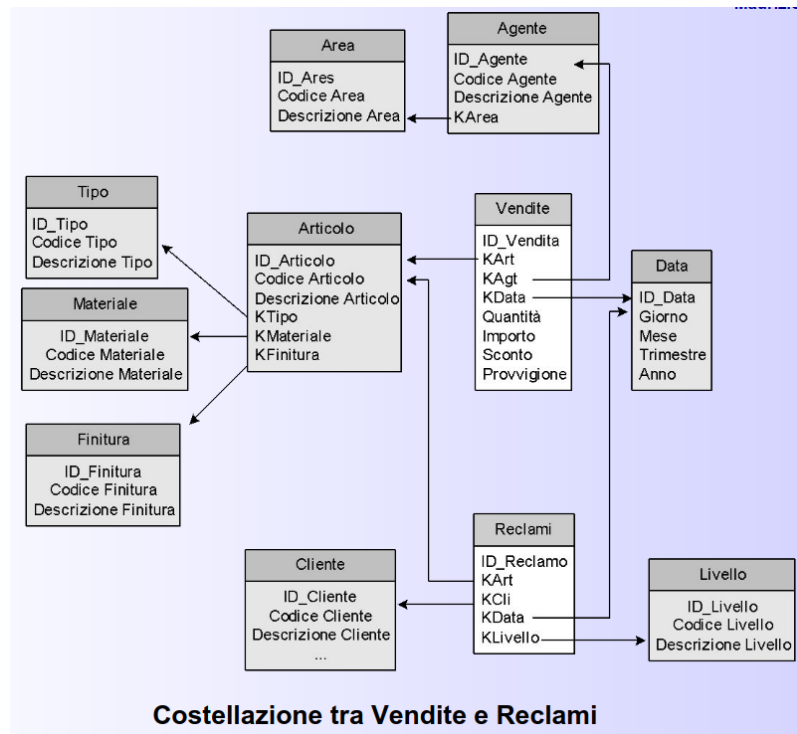
La denormalizzazione porta anche molti svantaggi come la scarsa intuitività e lo spazio occupato da gerarchie profonde.



- Schema a fiocco di neve Questo schema riduce la denormalizzazione esplicitando delle dipendenze funzionali. Questo permette di chiarificare la separazione tra i soggetti, migliora le prestazioni e riduce la sensibilità alle variazioni logiche. Ne risente però la velocità di risposta alle richieste.



- Schema a costellazione Se diverse tabelle dei fatti condividono delle tabelle dimensionali, risulta essere il miglior approccio da seguire quando più fatti coinvolgono gli stessi soggetti.



12.5 Ciclo di vita del DWH

La costruzione di un data warehouse è un processo che avviene, solitamente, in modalità iterativa.

Viene prima definito e popolato un ipercubo principale e man mano vengono aggiunti gli altri fatti, una volta rilasciati tutti i fatti di uno specifico interesse aziendale è possibile rilasciare il corrispettivo data mart.

Vantaggi:

- Premi risultati disponibili subito
- Investimenti obbligatoriamente diluiti
- Sviluppare il modello in base alle necessità

12.5.1 Costruzione dei data mart

E' costituita dai seguenti passaggi:

- **Analisi delle sorgenti:** capire quali dati sono disponibili e verificare che siano compatibili con i requisiti lato utente.

- **Progettazione concettuale degli schemi:** identificare misure, dimensioni ed eventuali limiti di aggregabilità.
- **Progettazione logica:** decisione su schemi a stella/fiocco di neve e la necessità di costruire viste materializzate o ipercubi con molta aggregazione.
- **Alimentazione:** le procedure che straggono i dati dalle sorgenti e li processano per prepararli al DWH.

12.6 Popolamento del data warehouse

12.6.1 Fasi di popolamento

1. Estrazione dei dati I dati vengono estratti dalle diverse fonti, viene definito quali dati devono essere acquisiti (tabelle e campi) e anche come devono essere trattati gli eventi di origine (aggregazione o massimo dettaglio).

Esistono due tipi di estrazione, la prima è tipicamente usata per la popolazione iniziale:

- (a) **Statica:** vengono prelevati tutti i dati presenti nella sorgente.
- (b) **Dinamica:** vengono estratti solo i dati modificati o prodotti alla fonte.
 - L'estrazione avviene in modo automatico, guidata da apposite funzioni trigger.
 - L'estrazione avviene in modo periodico, tipica delle PMI.
 - Viene fatto un confronto diretto con la sorgente.

Spesso i dati estratti non vengono subito scritti nel DWH ma messi nella staging area per essere modificati.

2. Integrazione e trasformazione Prima di essere scritti i dati devono essere resi omogenei.
 - **Riconciliazione:** i dati che riguardano lo stesso soggetto ma provengono da fonti diverse sono messi in relazione.
 - **Riconoscimento dei duplicati.**
 - **Trasformazione dei valori continui:** vengono parametrizzati in valori discreti.
 - **Standardizzazione dei formati.**

3. Pulizia Questa fase può essere fatta anche prima dell'integrazione ma anche in modo parallelo ad essa.
Devono essere eseguite analisi per rilevare, possibilmente anche riparare, le incorrette presenti nei dati.
 - **Dati incompleti:** si può inserire un codice per indicare la mancanza del dato, oppure nei sistemi più complessi si fanno inferenze sugli altri dati per ricavare quello mancante.
 - **Dati errati o incomprensibili:** vengono confrontati con dizionari di dati o valori limite, per stabilire la loro ammissibilità.
 - **Dati inconsistenti:** vengono applicate delle regole per inferire quale risultato sia corretto.
4. Caricamento dei dati Questa è la fase in cui i dati vengono effettivamente caricati nel DWH, l'aggiornamento avviene dalle dimensioni verso i fatti con l'applicazione delle politiche di aggiornamento agli elementi già esistenti.
 - Non fare nulla, ogni fatto usa gli attributi dimensionali validi all'inserimento della dimensione.
 - Aggiornare l'elemento sovrascrivendolo.
 - Creare una nuova istanza che verrà associata ai nuovi fatti che verranno inseriti.
 - Creare una nuova istanza con marcatori temporali.

12.7 Tecniche di analisi dei dati

L'analisi OLAP è la principale modalità di interrogazione interattiva del warehouse.

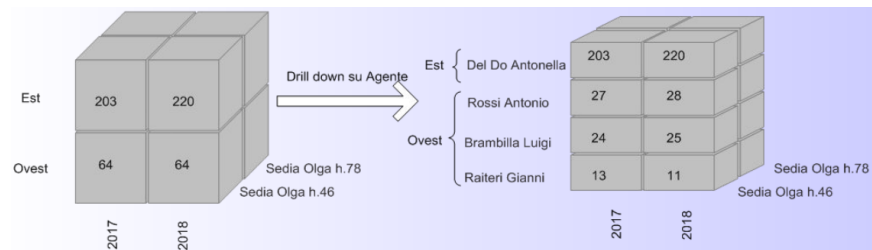
Il paradigma che si usa è quello dell'esplorazione guidata da ipotesi, l'utente formula un'ipotesi e inoltra la richiesta per verificarla.

Durante le interrogazioni viene costruita una sessione di analisi, ciascun passo diventa conseguenza dei risultati precedentemente ottenuti.

Ogni passaggio di navigazione è costituito da un operatore OLAP, se non diversamente detto viene applicato all'ultimo risultato ottenuto.

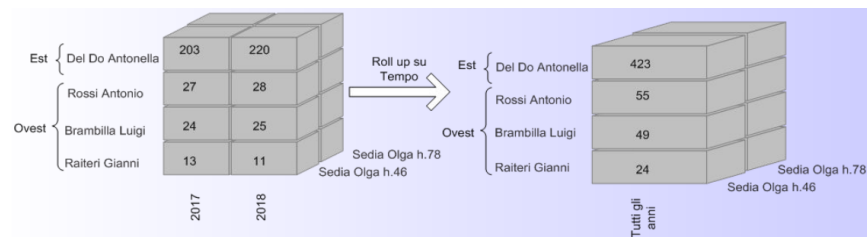
12.7.1 Drill down

Permette di partire da un livello generale e approfondire i dettagli passo passo, si scende lungo una gerarchia o aggiungendo una dimensione di analisi.



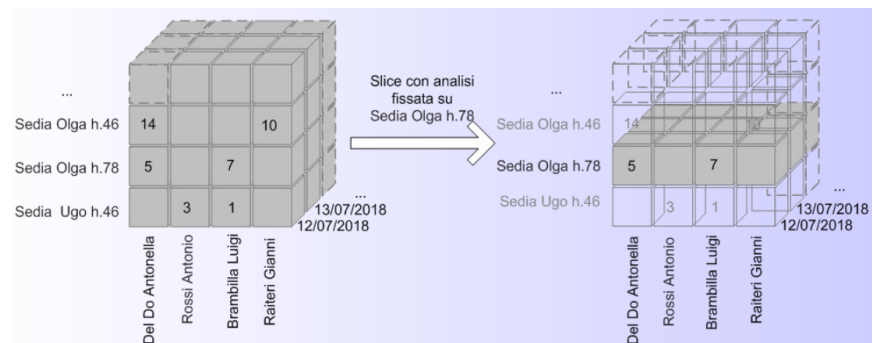
12.7.2 Roll up

E' il reciproco del drill down, si procede risalendo una gerarchia oppure eliminando una dimensione di analisi.



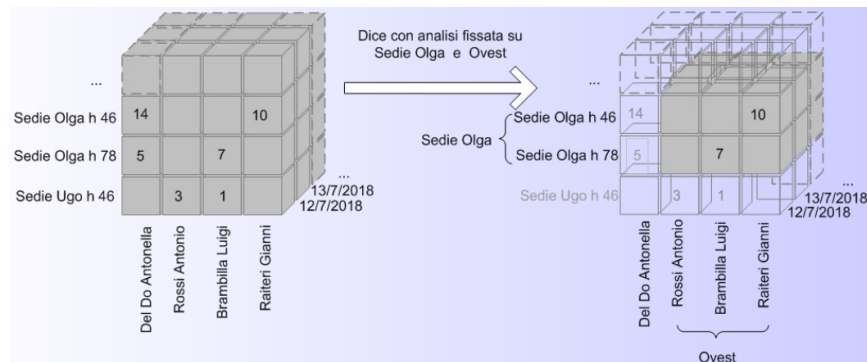
12.7.3 Slice

Viene fissato il valore di una dimensione e vengono analizzati i dati così ottenuti.



12.7.4 Dice

Simile allo slice ma può operare su più dimensioni fissandone il valore, si può applicare a dimensioni di qualsiasi livello.



12.7.5 Pivot

(Trasposizione di matrice nel caso 2D) Le dimensioni della matrice vengono invertite.

Prodotto	Area/Anno	2017	2018
Sedia Olga h.46	Est	203	220
	Ovest	64	64

↓

Prodotto	Anno/Area	Est	Ovest
Sedia Olga h.46	2017	203	64
	2018	220	64

Asse rotazione

13 Data minig

13.1 Limiti di OLAP

Data l'elevata quantità e la complessità delle relazioni, le informazioni non sono completamente identificabili.

Si predispone una situazione in cui i dati rimangono inutilizzati o sotto-utilizzati (molti dati ma poche informazioni).

Gli strumenti OLAP non sono più sufficienti perchè operano per supportare

processi decisionali, quindi sviluppano percorsi di analisi da ipotesi che è l'utente a formulare.

13.2 Fasi del processo di mining

Il data mining è spesso definito KDD (Knowledge Discovery in Databases) e viene diviso nelle seguenti fasi elementari:

- **Pulizia dei dati:** vengono eliminate le incorrettezze.
- **Integrazione dei dati:** uniformare i dati.
- **Selezione dei dati.**
- **Trasformazione dei dati:** riorganizzazione dei dati.
- **Data mining:** il vero e proprio processo di analisi.
- **Valutazione dei pattern:** l'insieme delle condizioni viene ridotto a quelle interessanti.
- **Presentazione delle conoscenze.**

Le prime fasi coincidono con il popolamento del DWH, può essere considerato un'evoluzione delle indagini OLAP.

13.3 Da OLAP a OLAM

OLAM (On Line Analytical Mining)

Partendo dai DWH abbiamo dati ben strutturati, puliti e completi. Ciononostante il processo di mining non può essere interamente automatico, infatti i pattern rilevati potrebbero essere troppi e non interessanti, il processo di data mining deve quindi essere interattivo con gli utenti che specificano la direzione in cui indagare.

Un processo interattivo permette di affinare le ricerche.

13.4 Architettura dei sistemi di Data Mining

L'architettura dei sistemi di data mining si appoggia ai seguenti componenti:

- **Data warehouse**
- **Base di conoscenza (knowledge base):** l'insieme di regole e conoscenze note, verranno utilizzate per guidare le ricerche.

- **Motore di data mining (data mining engine):** l'insieme delle funzioni di analisi dei dati.
- **Valutazione delle condizioni (pattern evaluation):** i moduli che fanno focalizzare la ricerca sulle condizioni interessanti.
- **Sistema di presentazione:** l'interfaccia con la quale l'utente fa le ricerche.

13.5 I 4 principi di analisi



13.6 Statistiche elementari e analisi relative

13.6.1 Generalizzazione

Deve fornire una visione ad alto livello tramite l'accorpazione di concetti e riassumendo caratteristiche di base.

Il principio di base è che gli elementi che un utente può analizzare devono essere un numero limitato.

Un diffuso tipo di generalizzazione è l'aggregazione dei sistemi OLAP, i sistemi di data mining amplificano il potenziale mettendo a disposizione anche delle metodologie di induzione.

1. **Caratterizzazione** Serve a comprendere le caratteristiche di una classe, che siano queste (caratteristiche) di tendenza o di dispersione. Viene spesso rappresentata con tabelle, grafici e boxplot.

2. **Discriminazione** Con questa modalità invece, le caratteristiche di una classe vengono messe a confronto con quelle di un'altra classe ad essa paragonabile.
Viene quindi eseguito un confronto diretto sulle tabelle o sui grafici.

13.7 Analisi associative

Meccanismi che permettono di identificare situazione che ne implicano altre con un'elevata frequenza.

Devono essere individuati pattern che rappresentano implicazioni logiche come $A \rightarrow B$.

La significatività di un'associazione viene definita con due parametri:

1. **Confidenza:** misura la certezza di un pattern, è definita come $P(A|B)$.
2. **Supporto:** la frequenza con cui il pattern è stato verificato nel DB, è definito come la percentuale degli elementi che verifica la regola.

13.8 Classificazione e predizione

Prevediamo degli eventi futuri oppure facciamo delle inferenze su dei valori mancanti.

- **Classificazione:** definiamo i criteri che permettono di assegnare un soggetto ad una classe.
- **Predizione:** calcolo di funzioni di tendenza interpolando dati noti.

13.8.1 Classificazione

Specifica quali sono le classi obiettivo della classificazione, quali sono i dati su cui costruire il modello e a quale classe appartengono.

Le tecniche per costruire i classificatori sono diverse:

- Funzioni matematiche
- Regole associative
- Alberi decisionali
- Reti della verità bayesiane
- Reti neurali

1. Alberi decisionali Un albero di decisione è una struttura semi alberi in cui:

- I nodi interni sono attributi del soggetto.
- Gli archi in uscita da un nodo sono i valori che l'attributo può avere.
- Le foglie sono le classi.

L'albero è di fatto una struttura *if-then-else* che va letta dalla radice alle foglie.

2. Predizione La predizione permette di prevedere (ma pensa un po') valori non ancora noti di un dominio continuo.

La costruzione delle funzioni di tendenza avviene tramite interpolazione dei punti noti (regressione), esistono diversi modelli di regressione:

- **Lineare semplice:** $Y = q + mX$
- **Multilineare:** $Y = q + m_1X_1 + m_2X_2 + m_3X_3$
- **Non lineare:** $Y = q + m_1X + m_2X^2 + m_3X^3$

13.9 Meccanismi di clustering

I meccanismi di clustering, come i classificatori, ripartiscono i dati in classi differenti senza però conoscere le classi, solo sulle affinità che gli elementi hanno.

Esistono diverse tecniche:

- **Partizionamento:** l'utente indica quante classi esistono.
- **Classificazione gerarchica:**
 - Aggregativa quando iterativamente aggrego gli elementi in base alle similitudini.
 - Divisiva quando spezzo gli insiemi in sottoinsiemi di elementi caratterizzati.
- **Valutazione della densità:** gli elementi vengono divisi in base alla loro posizione nell'iperpiano.

13.9.1 Ricerca degli outlier

I metodi di clustering hanno come effetto secondario l'identificazione degli outlier, ovvero degli elementi che si discostano dai raggruppamenti.

L'analisi degli outlier è simile alla procedura di clustering ma si concentra sull'identificazione degli elementi che si discostano maggiormente. La ricerca si avvale di:

- **Metodi statistici:** utilizzabili solo quando si conosce la distribuzione (statistica) dei dati.
- **Metodi basati sulla distanza:** ricerca degli elementi con distanza maggiore.
- **Metodi basati sulla deviazione:** ricerca di elementi che distano dal gruppo.