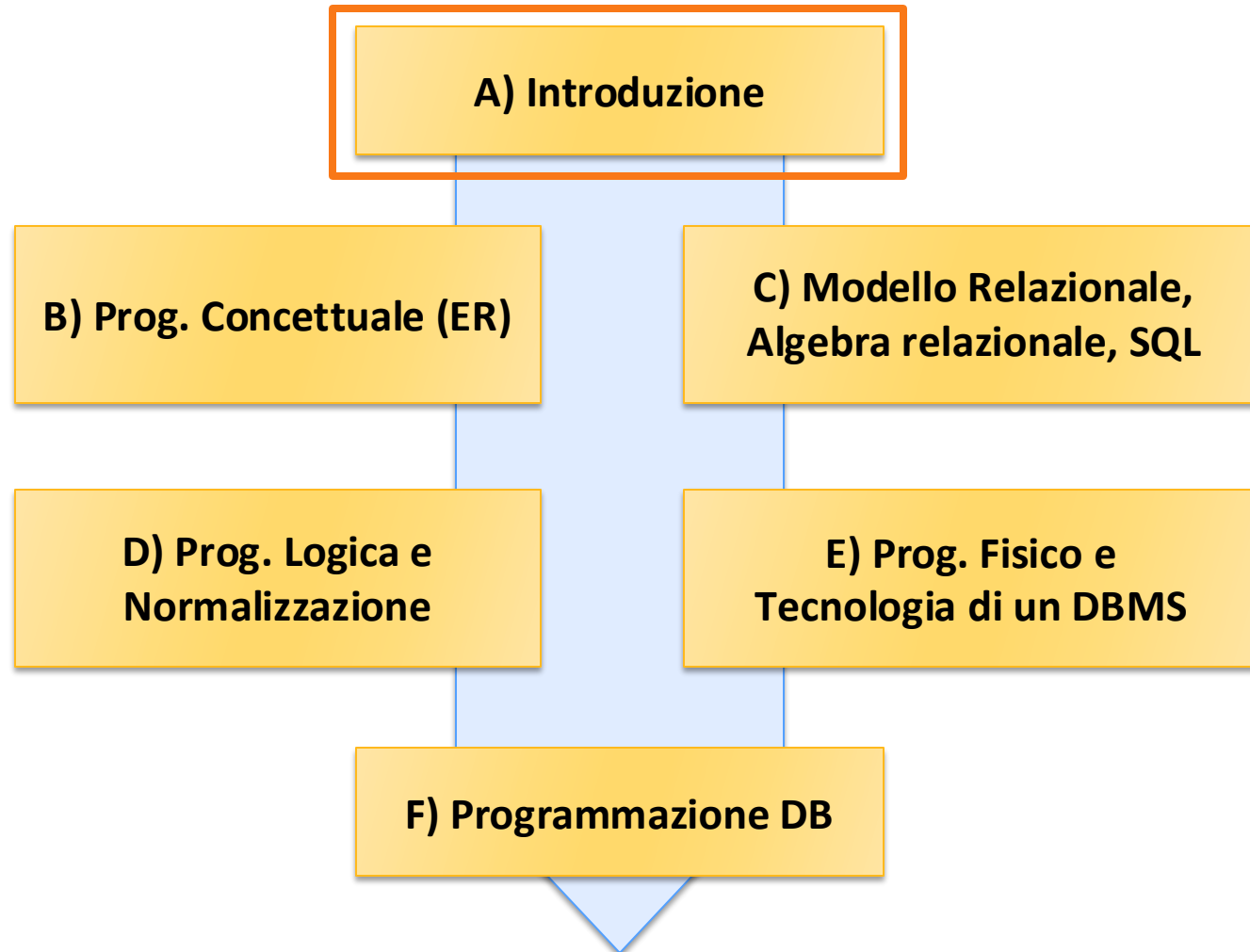


Basi di Dati

Introduzione ai sistemi informativi

Basi di Dati – Dove ci troviamo?



**“It is a capital mistake to theorize before
one has data”**

– Arthur Conan Doyle

Contenuto

- Sistemi informativi
- Data Base Management Systems (DBMS)
- Schemi, istanze e modelli dei dati
- Architettura di un DBMS



Sistema informativo

- ❑ Componente di una organizzazione che gestisce le informazioni di interesse (cioé utilizzate per il perseguimento degli scopi dell'organizzazione)
- ❑ Ogni organizzazione ha un sistema informativo, eventualmente non esplicitato nella struttura
- ❑ Il sistema informativo è di supporto ad altri sottosistemi, e va quindi studiato nel contesto in cui è inserito



Gestione delle informazioni

- Raccolta, acquisizione
- Archiviazione, conservazione
- Elaborazione, trasformazione, produzione
- Distribuzione, comunicazione, scambio



Sistemi informativi e automazione

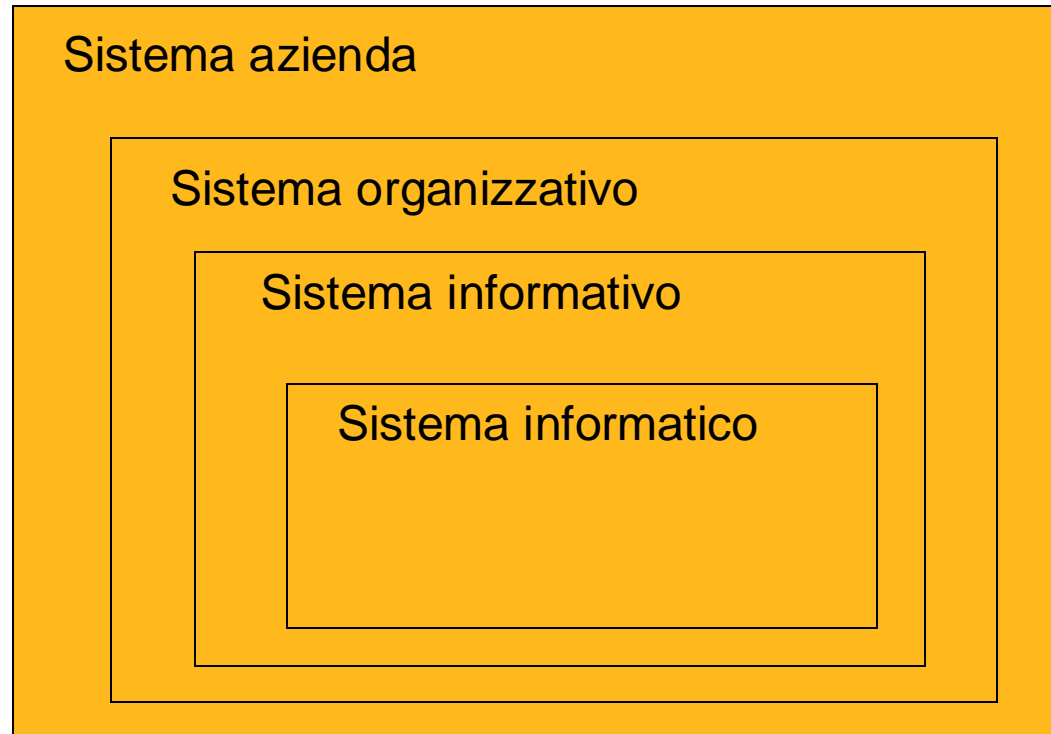
- Il concetto di “sistema informativo” è indipendente da qualsiasi *automatizzazione*:
 - esistono organizzazioni la cui ragion d’essere è la gestione di informazioni (p.es. servizi anagrafici e banche) e che operano da secoli

Sistema informatico

Porzione automatizzata del sistema informativo:

- la parte del sistema informativo che gestisce informazioni con tecnologia informatica

Sistema Informatico



Gestione delle informazioni

- Nelle attività umane, le informazioni vengono gestite in forme diverse:
 - idee informali
 - linguaggio naturale (scritto o parlato, formale o colloquiale, in varie lingue)
 - disegni, grafici, schemi
 - numeri e codici
- e su vari supporti
 - mente umana, carta, dispositivi elettronici

Informazioni e dati

- Nei sistemi informatici (e non solo), le **informazioni** vengono **rappresentate** in modo essenziale, spartano: attraverso i **dati**

Informazione*: notizia, dato o elemento che consente di avere conoscenza più o meno esatta di fatti, situazioni, modi di essere.

Dato*: ciò che è immediatamente presente alla conoscenza, prima di ogni elaborazione; (in informatica) elementi di informazione costituiti da simboli che debbono essere elaborati

* dal Vocabolario della lingua

Dati e informazioni



Lun-Ven



Sabato



Festivo

- che cosa significano questi numeri?
- cartelli stradali, in Finlandia; sono orari
- ma la differenza?
- senza "interpretazione," il dato serve a ben poco

Gestione delle informazioni

- I dati sono spesso il risultato di forme di organizzazione e codifica delle informazioni
- Ad esempio, nei servizi anagrafici e nel riferimento a persone
 - descrizioni discorsive
 - nome e cognome
 - estremi anagrafici
 - codice fiscale

Perché i dati?

- La rappresentazione precisa di forme più ricche di informazione e conoscenza è difficile
- I dati costituiscono spesso una risorsa strategica, perché più stabili nel tempo di altre componenti (processi, tecnologie, ruoli umani):
 - ✓ ad esempio, i dati delle banche o delle anagrafi

Base di dati

(accezione generica, metodologica)

- insieme organizzato di dati utilizzati per il supporto allo svolgimento delle attività di un ente (azienda, ufficio, persona)

(accezione specifica, metodologica e tecnologica)

- insieme di dati gestito da un DBMS

Contenuto

- Sistemi informativi
- Data Base Management Systems (DBMS)
- Schemi, istanze e modelli dei dati
- Architettura di un DBMS

Sistema di gestione di basi di dati

DataBase Management System (DBMS)

- Sistema che gestisce collezioni di dati:

- grandi
- persistenti
- condivise

... garantendo

- privatezza
- affidabilità
- efficienza
- efficacia

DBMS

- Prodotti software (complessi) disponibili sul mercato; esempi:
 - ❖ DB2
 - ❖ Oracle
 - ❖ SQLServer
 - ❖ MySQL
 - ❖ PostgreSQL
 - ❖ Access



Le basi di dati sono ... grandi

- dimensioni (molto) maggiori della memoria centrale dei sistemi di calcolo utilizzati
- il limite deve essere solo quello fisico dei dispositivi
- esempi di dimensioni molto grandi
 - 5 Terabyte (dati transazionali)
 - Multipli di 10 Terabyte (dati decisionali)
 - Multipli di 100 Terabyte (dati scientifici)
 - 100 miliardi di record



Le basi di dati sono ... persistenti

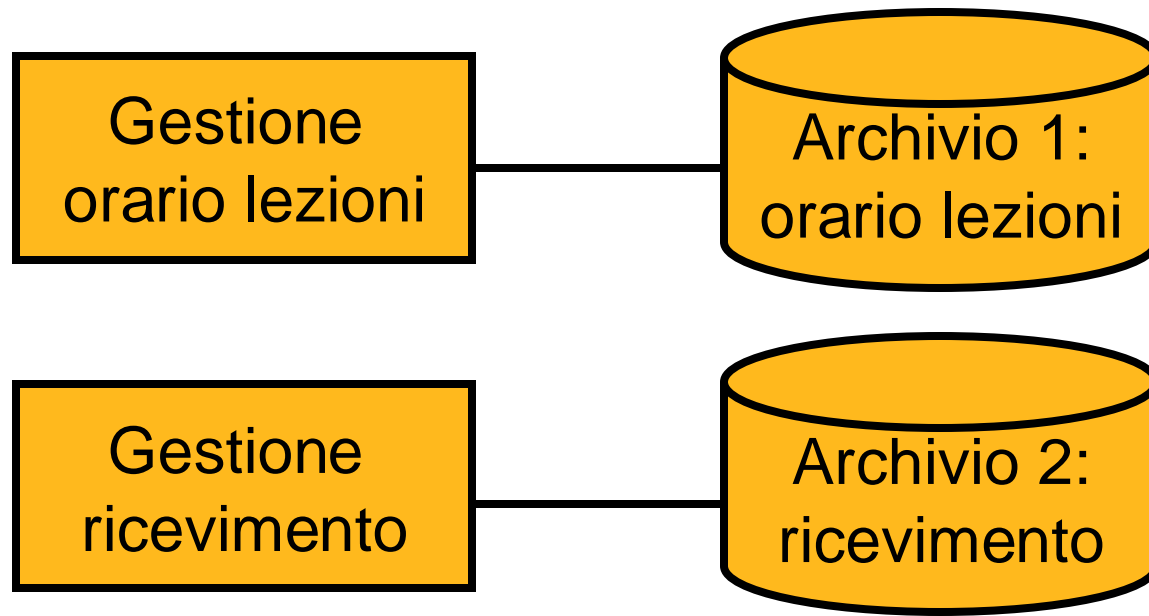
- Hanno un tempo di vita indipendente dalle singole esecuzioni dei programmi che le utilizzano



Le basi di dati sono ... condivise

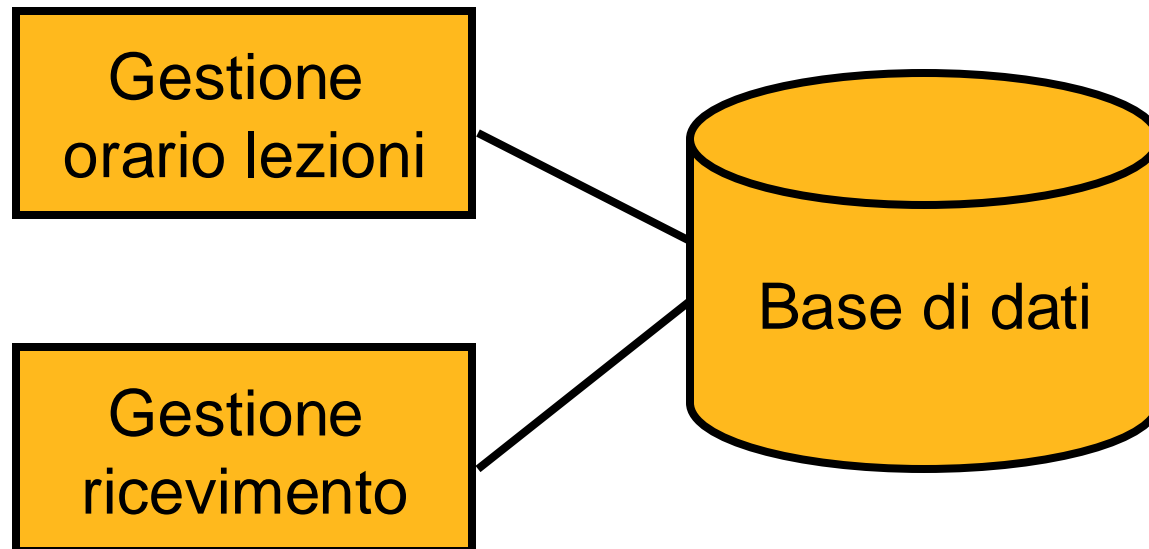
- Ogni organizzazione (specie se grande) è divisa in settori o comunque svolge diverse attività
- Ciascun settore/attività ha un (sotto)sistema informativo (non necessariamente disgiunto)

Archivi e basi di dati



- **Ridondanza:**
 - informazioni ripetute
- **Rischio di incoerenza:**
 - le versioni possono non coincidere

Archivi e basi di dati



Le basi di dati sono condivise

- Una base di dati è una risorsa **integrata, condivisa** fra applicazioni
- conseguenze
 - Attività diverse su dati condivisi:
 - meccanismi di **autorizzazione**
 - Accessi di più utenti ai dati condivisi:
 - controllo della **concorrenza**



I DBMS garantiscono ... privacy

- Si possono definire **meccanismi di autorizzazione**
 - l'utente A è autorizzato a leggere tutti i dati e a modificare X
 - l'utente B è autorizzato a leggere dati X e a modificare Y



I DBMS garantiscono... affidabilità

- **Affidabilità** (per le basi di dati):
 - resistenza a malfunzionamenti hardware e software
- Una base di dati è una risorsa pregiata e quindi deve essere conservata a lungo termine
- Tecnica fondamentale:
 - gestione delle transazioni

Transazione

Insieme di operazioni da considerare indivisibile ("atomico"), corretto anche in presenza di concorrenza e con effetti definitivi



Le transazioni sono ... atomiche

- Una sequenza di operazioni correlate:
 - trasferimento di fondi da un conto A ad un conto B: o si fanno il prelevamento da A e il versamento su B o nessuno dei due
- ... deve essere eseguita per intero o per niente:
 - o si fanno il prelevamento da A e il versamento su B o nessuno dei due



Le transazioni sono ... concorrenti

- L'effetto di transazioni concorrenti deve essere coerente
 - se due assegni emessi sullo stesso conto corrente vengono incassati contemporaneamente
 - ... si deve evitare di trascurarne uno

- se due agenzie rischiedono lo stesso posto (libero) su un treno
- ... si deve evitare di assegnarlo due volte



I risultati delle transazioni ... sono permanenti

- La conclusione positiva di una transazione corrisponde ad un impegno (in inglese **commit**) a mantenere traccia del risultato in modo definitivo, anche in presenza di guasti e di esecuzione concorrente



I DBMS debbono essere...efficienti

- Cercano di utilizzare al meglio le risorse di spazio di memoria (principale e secondaria) e tempo (di esecuzione e di risposta)
- I DBMS, con tante funzioni, rischiano l'inefficienza e per questo ci sono grandi investimenti e competizione
- L'efficienza è anche il risultato della qualità delle applicazioni



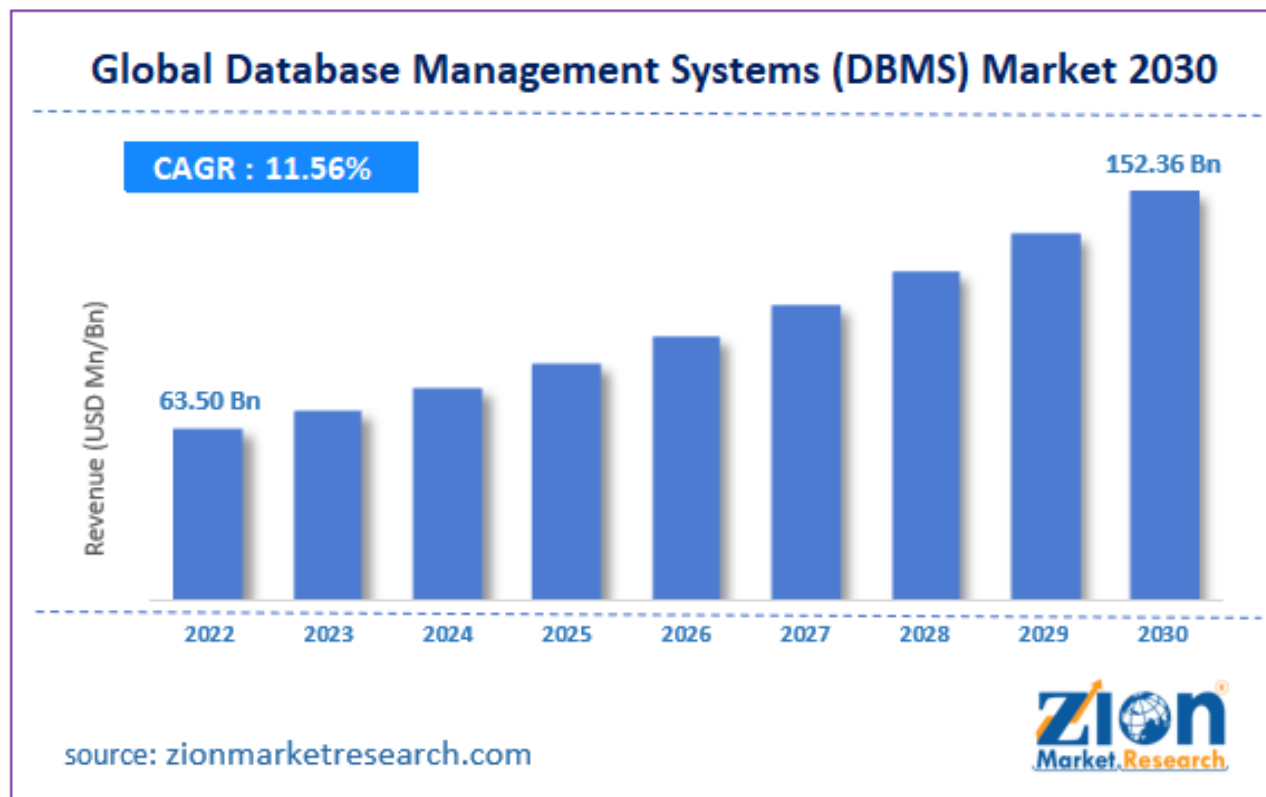
I DBMS debbono essere...efficaci

- Cercano di rendere produttive le attività dei loro utilizzatori, offrendo funzionalità articolate, potenti e flessibili:

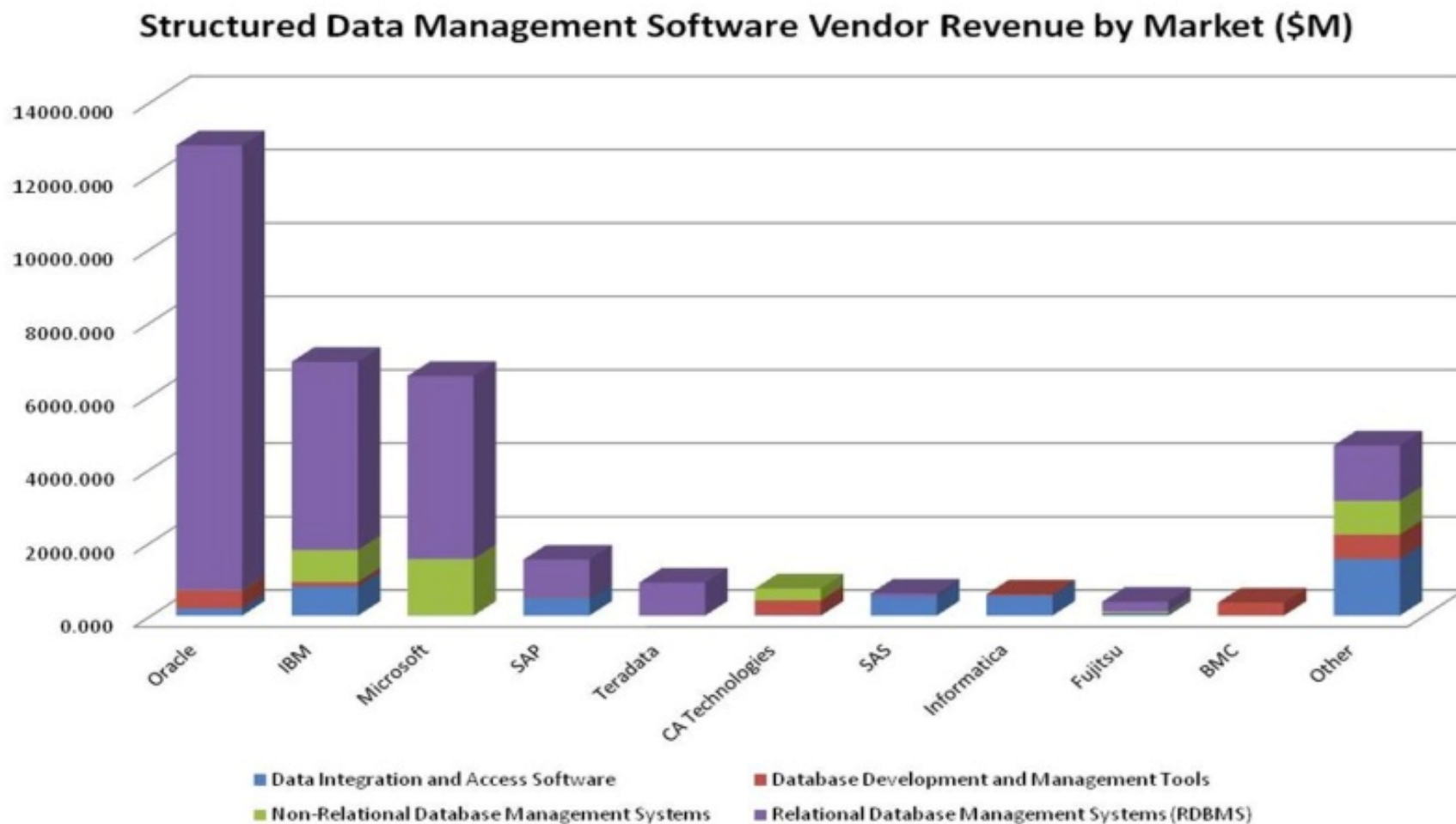
Il corso è in buona parte dedicato ad illustrare come i DBMS perseguono l'efficacia

Uno sguardo al mercato dei DBMS

- Fatturato globale 2017: 23×10^9 US\$
- Fatturato globale 2022: 34×10^9 US\$

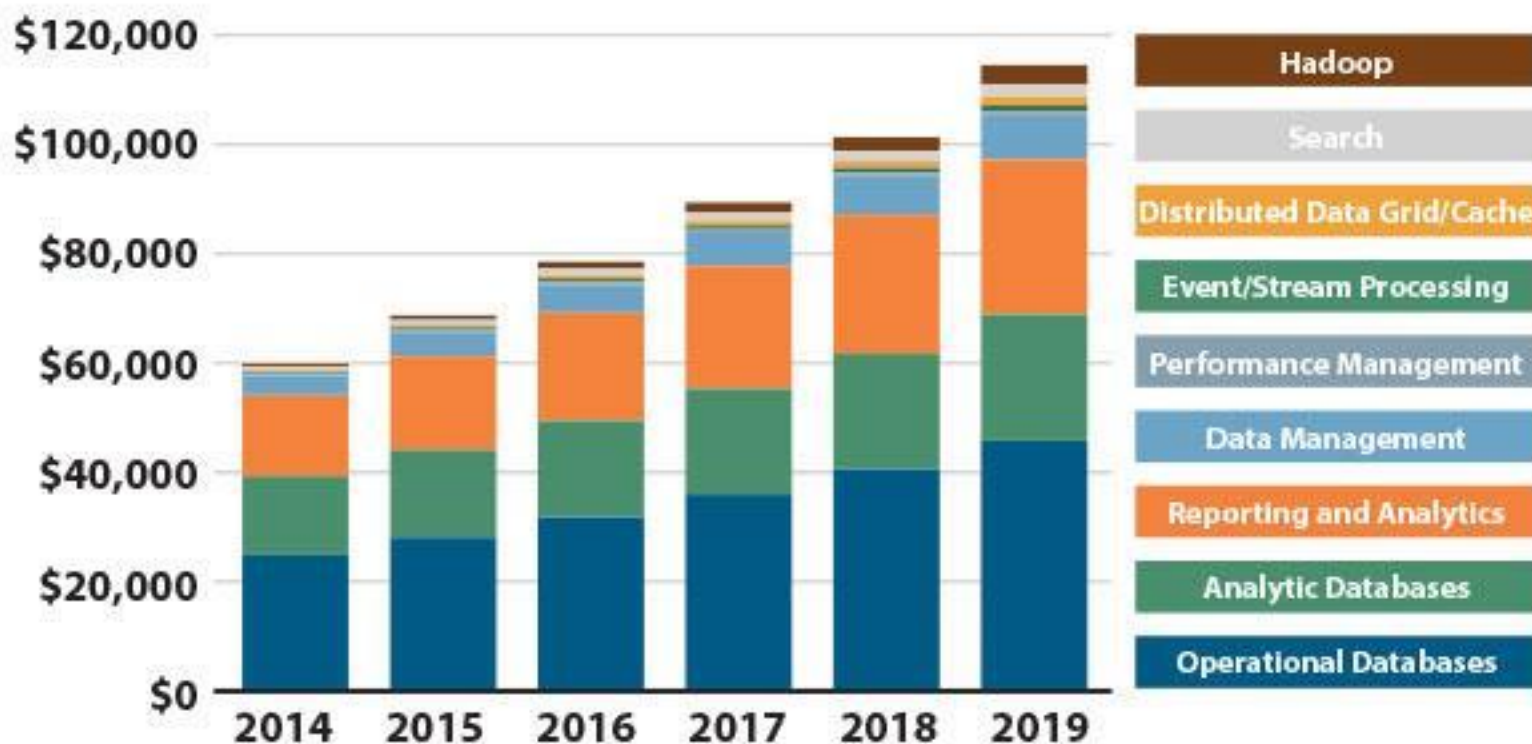


Uno sguardo al mercato dei DBMS



Uno sguardo al mercato dei DBMS

Worldwide Total Data Revenue by Segment (\$M) 2014-2019

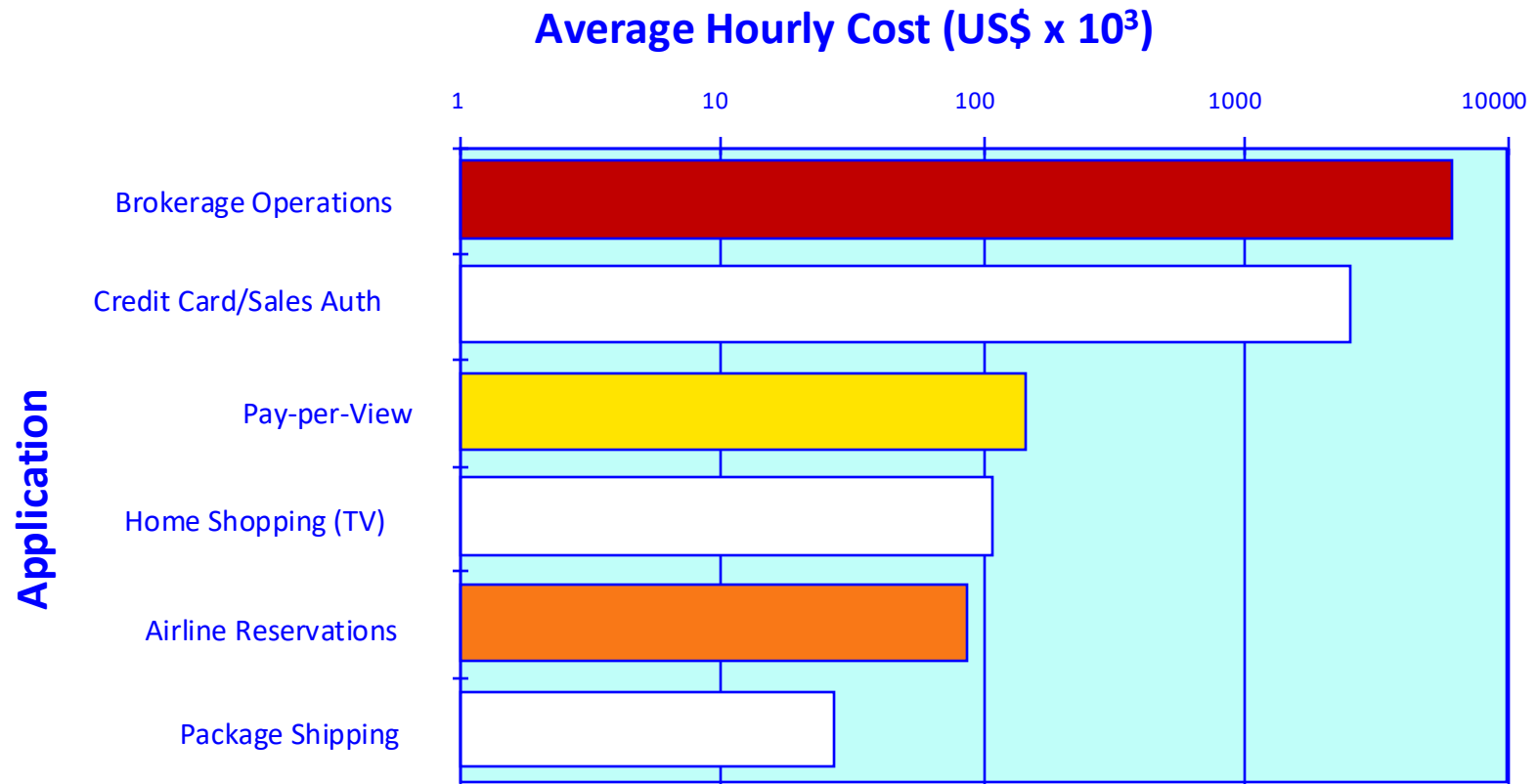


Source: 451 Research Market Monitor



Il valore di una buona gestione dati

Financial Impact of System Failure



Contenuto

- Sistemi informativi
- Data Base Management Systems (DBMS)
- Schemi, istanze e modelli dei dati
- Architettura di un DBMS

Organizzazione dei dati in una base di dati

Orario

Insegnamento	Docente	Aula	Ora
Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	9:45
Chimica	Nicola Mori	N1	9:45
Fisica I	Mario Bruni	N1	11:45
Fisica II	Mario Bruni	N3	9:45
Sistemi inform.	Piero Rossi	N3	8:00

Basi di dati: schema e istanza

Orario

Insegnamento	Docente	Aula	Ora
Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	9:45
Chimica	Nicola Mori	N1	9:45
Fisica I	Mario Bruni	N1	11:45
Fisica II	Mario Bruni	N3	9:45
Sistemi inform.	Piero Rossi	N3	8:00

Lo
schema
della base
di dati

L'**istanza** della base di dati

Schema e istanza

- In ogni base di dati esistono:
 - lo **schema**, sostanzialmente invariante nel tempo, che ne descrive la struttura (aspetto intensionale)
 - es.: le intestazioni delle tabelle
 - l'**istanza**, i valori attuali, che possono cambiare anche molto rapidamente (aspetto estensionale)
 - es.: il “corpo” di ciascuna tabella
- Lo schema e le istanze di una base di dati si basano su un **modello dei dati**

Modello dei dati

- Insieme di costrutti utilizzati per organizzare i dati di interesse e descriverne la dinamica
- Componente fondamentale: **meccanismi di strutturazione** (o **costruttori di tipo**)
- Come nei linguaggi di programmazione esistono meccanismi che permettono di definire nuovi tipi, così ogni modello dei dati prevede alcuni costruttori
- Esempio: il **modello relazionale** prevede il costruttore **relazione**, che permette di definire insiemi di record omogenei
- Due principali tipologie di modelli dei dati:
 - **modelli logici**
 - **modelli concettuali**

Modelli logici

- Adottati nei DBMS esistenti per l'organizzazione dei dati
 - utilizzati dai programmi
 - indipendenti dalle strutture fisiche
- Esempi: **relazionale**, reticolare, gerarchico, a oggetti, basato su XML

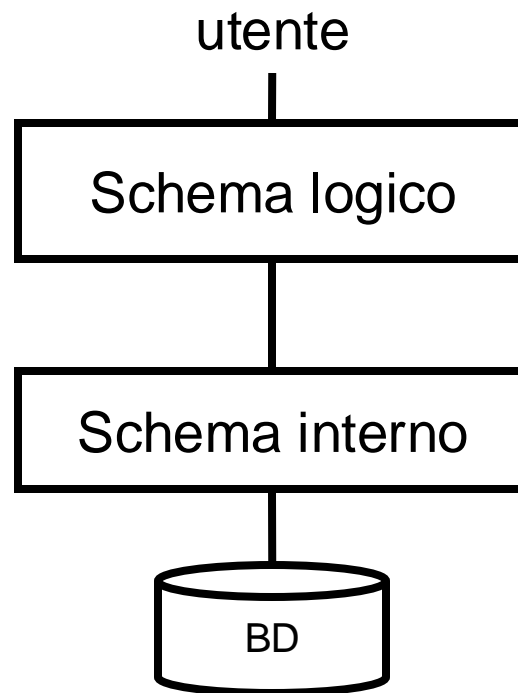
Modelli concettuali

- Permettono di rappresentare i dati in modo indipendente da ogni sistema
 - ▣ cercano di descrivere i concetti del mondo reale
 - ▣ sono utilizzati nelle fasi preliminari di progettazione
- Il più diffuso è il modello **Entity-Relationship**

Contenuto

- Sistemi informativi
- Data Base Management Systems (DBMS)
- Schemi, istanze e modelli dei dati
- Architettura di un DBMS

Architettura (semplificata) di un DBMS



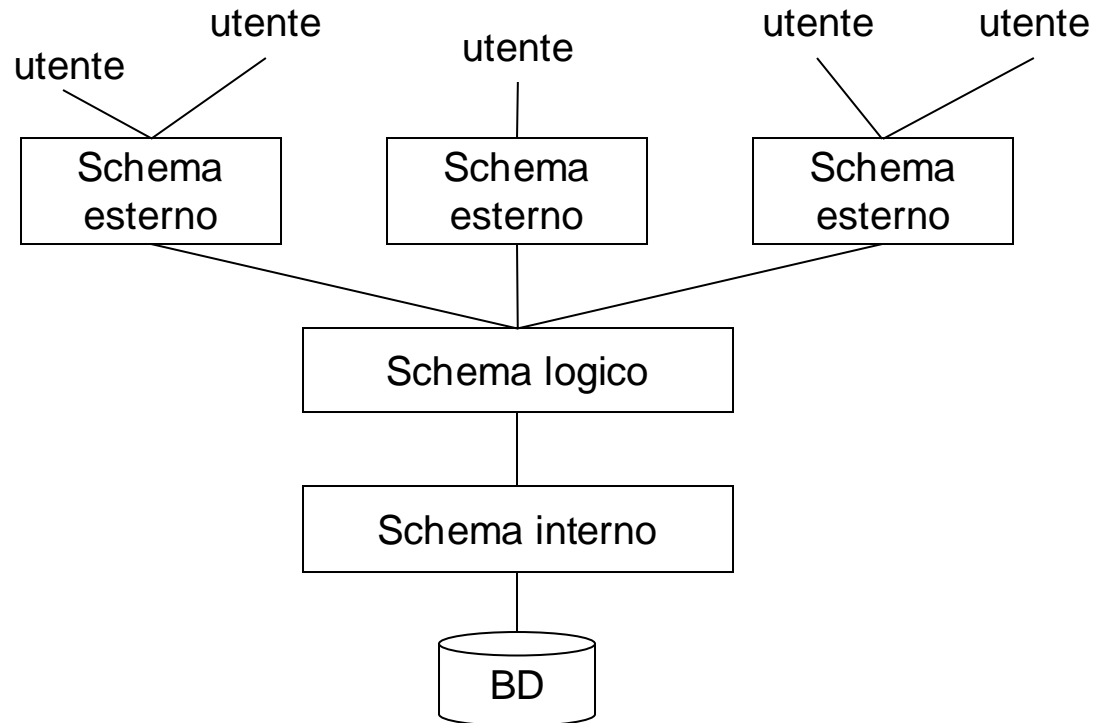
Architettura semplificata di un DBMS: schemi

- **schema logico**: descrizione della base di dati nel modello logico (ad esempio, la struttura della tabella)
- **schema interno** (o **fisico**): rappresentazione dello schema logico per mezzo di strutture memorizzazione (file; ad esempio, record con puntatori, ordinati in un certo modo)

Indipendenza dei dati

- Il livello logico è indipendente da quello fisico:
 - una tabella è utilizzata nello stesso modo qualunque sia la sua realizzazione fisica (che può anche cambiare nel tempo)
- Perciò in questo corso vedremo solo il livello logico e non quello fisico

Architettura standard (ANSI/SPARC) a tre livelli per DBMS



Architettura ANSI/SPARC: schemi

Schema logico: descrizione dell'intera base di dati nel modello logico “principale” del DBMS

Schema interno (o fisico): rappresentazione dello schema logico per mezzo di strutture fisiche di memorizzazione

Schema esterno: descrizione di parte della base di dati in un modello logico (“viste” parziali, derivate, anche in modelli diversi)

Una vista

Corsi

Corso	Docente	Aula
Basi di dati	Rossi	DS3
Sistemi	Neri	N3
Reti	Bruni	N3
Controlli	Bruni	G

Aule

Nome	Edificio	Piano
DS1	OMI	Terra
N3	OMI	Terra
G	Pincherle	Primo

Corsi Sedi

Corso	Aula	Edificio	Piano
Sistemi	N3	OMI	Terra
Reti	N3	OMI	Terra
Controlli	G	Pincherle	Primo

Indipendenza dei dati

- conseguenza della articolazione in livelli
- l'accesso avviene solo tramite il livello esterno (che può coincidere con il livello logico)
- due forme:
 - indipendenza fisica
 - indipendenza logica

Indipendenza fisica

- il livello logico e quello esterno sono indipendenti da quello fisico
 - una relazione è utilizzata nello stesso modo qualunque sia la sua realizzazione fisica
 - la realizzazione fisica può cambiare senza che debbano essere modificati i programmi



Indipendenza logica

- il livello esterno è indipendente da quello logico
- aggiunte o modifiche alle viste non richiedono modifiche al livello logico
- modifiche allo schema logico che lascino inalterato lo schema esterno sono trasparenti

I linguaggi del DBMS

Structured Query Language (SQL)

DATA DEFINITION LANGUAGE (DDL)

es: CREATE, DROP, ALTER

DATA MANIPULATION LANGUAGE (DML)

**es: SELECT, INSERT, UPDATE,
DELETE**

SQL, un linguaggio interattivo

- "Trovare i corsi tenuti in aule a piano terra"

Corsi

Corso	Docente	Aula
Basi di dati	Rossi	DS3
Sistemi	Neri	N3
Reti	Bruni	N3
Controlli	Bruni	G

Aule

Nome	Edificio	Piano
DS1	OMI	Terra
N3	OMI	Terra
G	Pincherle	Primo

SQL, un linguaggio interattivo

```
SELECT Corso, Aula, Piano  
FROM Aule, Corsi  
WHERE Nome = Aula  
      AND Piano = 'Terra'
```

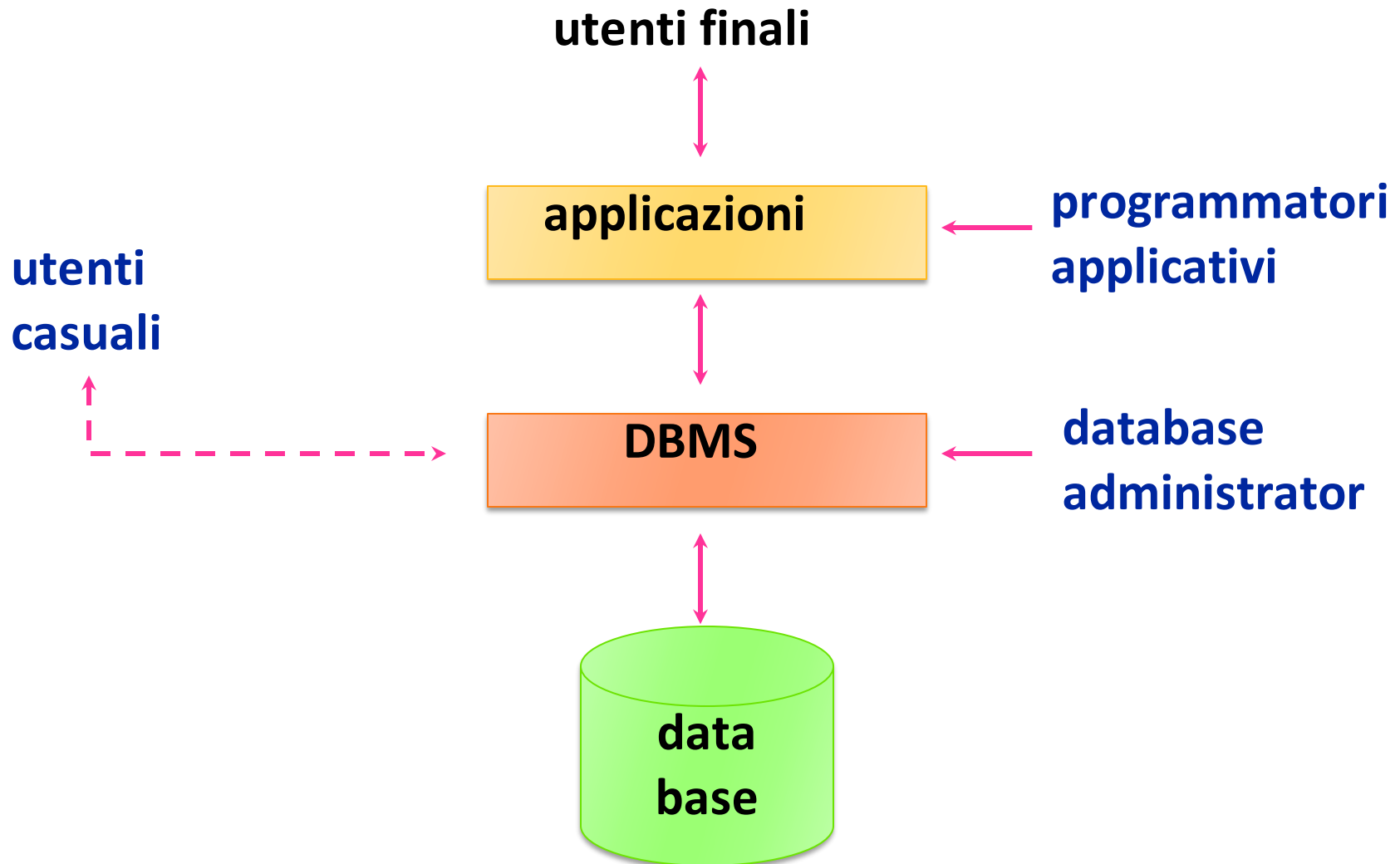
Corso	Aula	Piano
Sistemi	N3	Terra
Reti	N3	Terra

Gli “utenti” di un DBMS

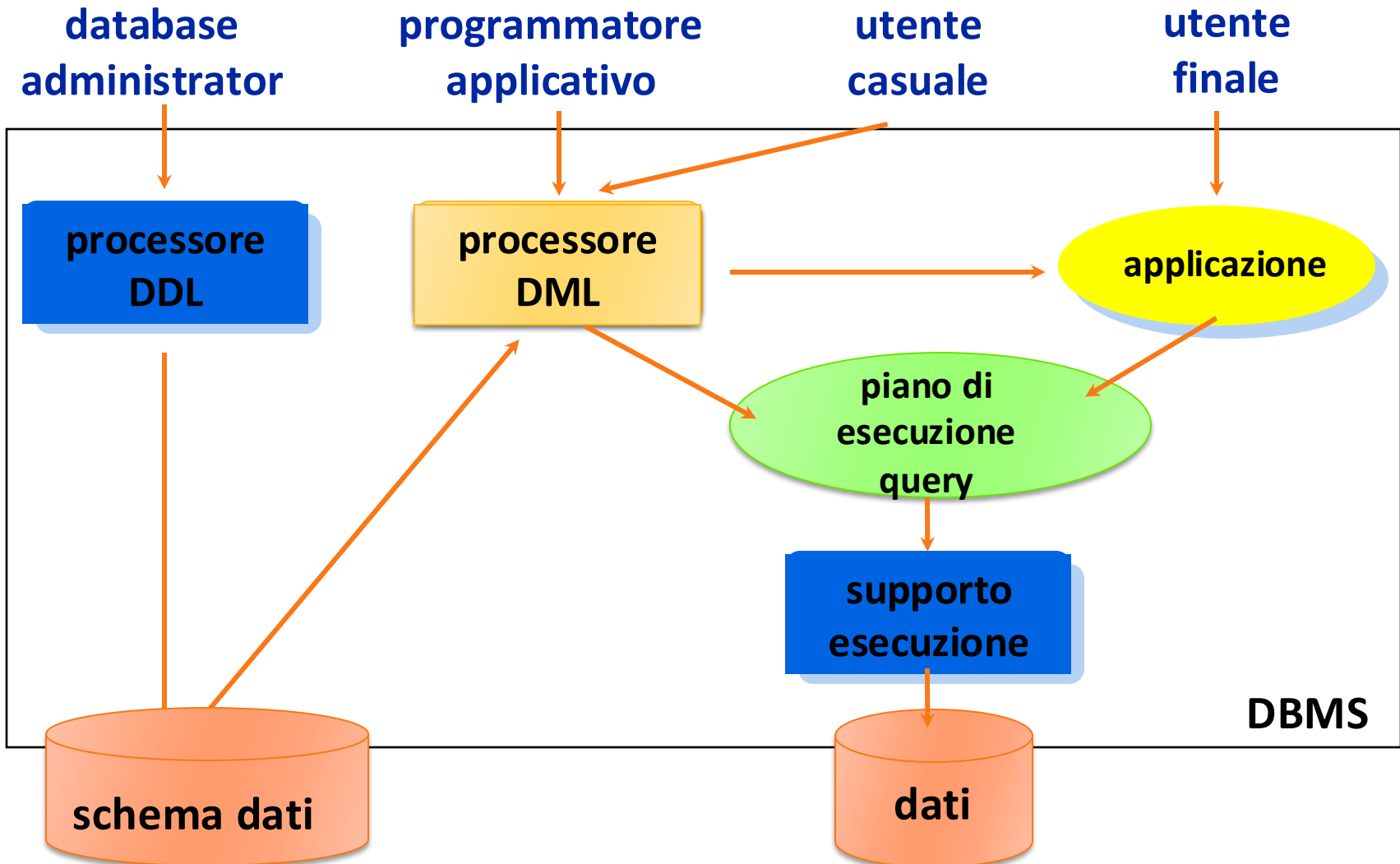
- progettisti e realizzatori di DBMS
- progettisti della base di dati e amministratori della base di dati (DBA)
- progettisti e programmatori di applicazioni
- utenti
 - ▣ utenti finali (terminalisti): eseguono applicazioni predefinite (transazioni)
 - ▣ utenti casuali: eseguono operazioni non previste a priori, usando linguaggi interattivi



Gli utenti del DBMS



I moduli del DBMS



DBMS in un contesto di rete

