



Introduzione al corso

Università degli Studi del Sannio
Facoltà di Ingegneria
Corso di Laurea in Ingegneria Informatica

Corso di Basi di Dati
Anno Accademico 2006/2007
docente: ing. Corrado Aaron Visaggio

email: visaggio@unisannio.it

ricevimento: mercoledì 11.00-13.00.

Corrado Aaron Visaggio

1



It' s me

Laureato in Ingegneria Elettronica al Politecnico di Bari, 2001.

Ricercatore presso il Fraunhofer Institute of Empirical Software Engineering, 2001.

Dottorato in Ingegneria dell'Informazione presso l'Università degli Studi del Sannio, 2005.

Presidente della commissione del corso di Basi di Dati, 2005.

Ricercatore (Ing-INF/05) presso l'Università degli Studi del Sannio, 2006.

Ricercatore presso il Research Centre on Software Technology-Benevento.

Visiting Reseracher presso il gruppo Alarcos del prof. Piattini, dell'Escuela Superiora de Informatica – Universidad Castilla La Mancha (Spagna).

Amenità: lettura (giallistica, psicologia, filosofia) e palestra.

Corrado Aaron Visaggio

2

What I do

Autore di circa 40 pubblicazioni su riviste e conferenze internazionali, nazionali, e libri.

Ambiti di Ricerca:

- Empirical Software Engineering
- Metodi e Tecniche di Sviluppo Software
- Sicurezza informatica

E' stato membro di commissioni d'esame per il corso di Ingegneria Informatica di:

- Software Testing
- Fondamenti di Informatica
- Tecnologie per la Produzione del Software
- Gestione dei Processi Software

Docente presso master universitari e formazione post-universitaria di:

- Programmazione Object-Oriented
- Software Testing
- Ingegneria del Software

Il corso...

Gli Obiettivi:

- imparare a **progettare** un DB
 - usabilità
 - interoperabilità
 - efficienza
 - efficacia
- imparare ad **manutenere** un DB
 - evolvibilità
 - modificabilità

METRICHE

...Il corso...

Il programma:

- Algebra Relazionale
- SQL
- Progettazione di DB
- Tecnologia di un DB Server
- MySQL
- Elementi di DB ad Oggetti
- Elementi di Sicurezza dei DB (Seminario)
- JDBC
- Persistenza dei dati - Hibernate (Seminario)
- Qualità dei Dati (Seminario)

...Il corso

L'esame:

- verifica sui contenuti;
- verifica sulla capacità di progettazione ed implementazione;
- verifica sul laboratorio.

Durante l'anno:

- lezioni frontali (Marzo-Aprile-Giugno);
- esercitazioni in aula (Marzo-Aprile-Giugno);
- sedute di laboratorio (Aprile-Maggio);
- sviluppo di un progetto (Aprile-Maggio-Giugno);

II Materiale

P. Atzeni, S. Ceri, S. Paraboschi, R. Torlone
Basi di dati: modelli e linguaggi di interrogazione
 Seconda edizione
 McGraw-Hill Italia, 2006



P. Atzeni, S. Ceri, P. Fraternali, S. Paraboschi, R. Torlone
Basi di dati: architetture e linee di evoluzione
 McGraw-Hill Italia, 2003



Corrado Aaron Visaggio

Sistemi Informativi

Un sistema informativo è un componente di una organizzazione il cui scopo è gestire le informazioni di business.

GESTIRE= acquisire, elaborare, conservare, produrre, distribuire.

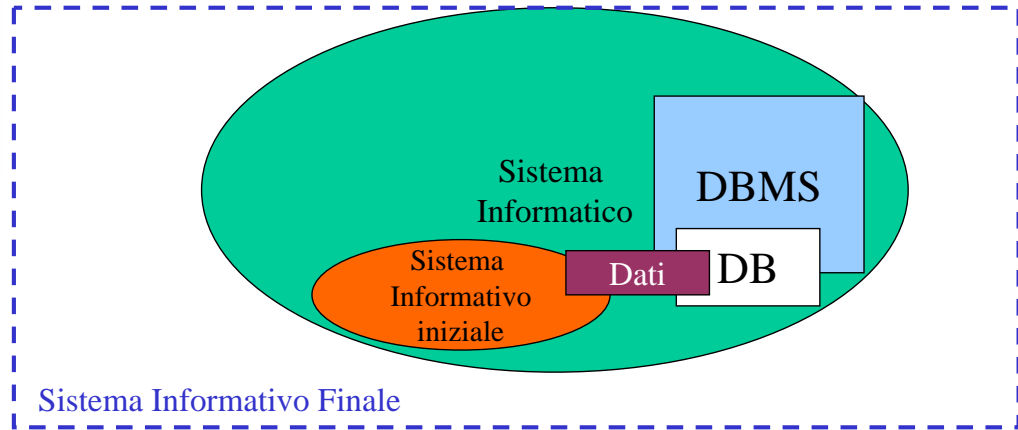
I processi organizzativi trasformano le materie prime ed i semilavorati in prodotti finiti; i sistemi informativi trasformano dati in informazione e conoscenza.



Corrado Aaron Visaggio

Sistemi Informatici

L'esistenza dei sistemi informativi è indipendente dalla sua automazione, che è il sistema informatico. Il primo esiste senza il secondo. Non è vero il contrario.



Una Base di Dati è una collezione di dati (DB). Un DB è gestito tramite un Data Base Management System (DBMS)

Corrado Aaron Visaggio

9

Data, Information, Knowledge...

"Data are not information. Information is data endowed with relevance and purpose. Knowledge is information endowed with application. Wisdom is knowledge endowed with age and experience."

Ken Davenport, Northeast Iowa Regional Library System, in a letter printed in the May 1, 2002, Library Journal (v. 127, no. 8, p. 10).



...Data, Information, Knowledge...

Secondo **Russell Ackoff**, professore di organizational change, Chairman of Interact, the Institute for Interactive Management, and Professor Emeritus of the Wharton School, University of Pennsylvania.

il contenuto della mente umana può dividersi in cinque categorie:

- **Data**: symbols
- **Information**: data that are processed to be useful; provides answers to "who", "what", "where", and "when" questions
- **Knowledge**: application of data and information; answers "how" questions
- **Understanding**: appreciation of "why"
- **Wisdom**: evaluated understanding.



...Data, Information, Knowledge...

Data: data is **raw**. It simply exists and has no significance beyond its existence (in and of itself). It can exist in any form, usable or not. It does not have meaning of itself. "Mario" "Rossi" "1256"

Information: information is data that has been given meaning by way of **relational connection**. This "meaning" can be useful, but does not have to be. In **computer parlance**, a **relational database** makes information from the data stored within it.

"Mario Rossi con numero di matricola 1256"

Knowledge: knowledge is the appropriate **collection of information**, such that it's intent is to be useful. Knowledge is a deterministic and volunteer process. Knowledge has **useful meaning** to them, but it does not provide for, in and of itself, an integration such as would infer further knowledge. In **computer parlance**, most of the applications we use (**modeling, simulation**, etc.) exercise some type of stored knowledge.



...Data, Information, Knowledge...

Understanding: understanding is an interpolative and probabilistic process. It is cognitive and analytical. It is the process by which I can take knowledge and **synthesize new knowledge** from the previously held knowledge. The difference between understanding and knowledge is the difference between "learning" and "memorizing". In **computer parlance**, **AI systems** possess understanding in the sense that they are able to synthesise new knowledge from previously stored information and knowledge.



...Data, Information, Knowledge

Wisdom: wisdom is an extrapolative and non-deterministic, non-probabilistic process. It calls upon all the previous levels of consciousness, and specifically upon special types of human programming (moral, ethical codes, etc.). It beckons to give us understanding about which there has previously been no understanding, and in doing so, goes far beyond understanding itself. It is the essence of philosophical probing. Unlike the previous four levels, it asks questions to which there is no (easily-achievable) answer, and in some cases, to which there can be no humanly-known answer period. Wisdom is therefore, the process by which we also **discern**, or **judge**, between right and wrong, good and bad. I personally believe that **computers do not have**, and will never have the ability to possess wisdom. **Wisdom is a uniquely human state**, or as I see it, wisdom requires one to have a **soul**, for it resides as much in the heart as in the mind.



Platone

La Scuola di Atene

Aristotele



Eraclito

Euclide

Pitagora

La scuola di Atene è un affresco di cm 772 di base realizzato tra il 1509 ed il 1511 dal pittore Raffaello Sanzio.

È conservato nella Stanza delle Segnature nei Palazzi Vaticani di Città del Vaticano.

Questo dipinto è uno dei massimi capolavori del Rinascimento.

Sotto le magnifiche volte a botte di un immaginario edificio classico appaiono celebri filosofi dell'antichità mentre discutono le loro scuole di pensiero.

Venne commissionato da papa Giulio II come parte di una serie di dipinti per la sua biblioteca privata nel Palazzo Vaticano. La sala venne in seguito utilizzata per la firma di documenti ed è per questo motivo che viene chiamata la Stanza della Segnatura.



Evoluzione della teoria della conoscenza

Platone (Atene, 428 a.C.) -> conoscenza come anamnesi

San Tommaso (Roccasecca, 1225 o 1226) -> cognitum est in cognoscente per modum cognoscentis / il conoscere è un atto di astrazione dell'universale dal particolare.

Kant (Konigsberg, 1724) -> Il metodo della ragione fondante ritiene fondato un concetto quando se ne possa dimostrare la mancanza di contraddizioni interiori / l'impossibilità della conoscenza di trascendere i limiti dell'esperienza diventa la base dell'effettiva validità della conoscenza.

Russell (UK, 1872) -> L'esperienza come punto di partenza della conoscenza / Il principio di inferenza garantisce l'identità di struttura tra le esperienze sensibili e le loro cause fisiche.

Putnam -> (la prospettiva funzionale) la mente è un insieme di programmi che girano su quella complicata macchina biologica che è il cervello.

I DBMS

Un sistema per la gestione di basi di dati gestisce basi di dati

- **grandi** (oltre i Giga, Tera Byte),
- **condivise** (applicazioni ed utenti diversi),
- **persistenti** (i dati hanno un tempo di vita non limitato a quello delle singole esecuzioni),

assicurando la loro

- **affidabilità** (conservare integro il contenuto nonostante i malfunzionamenti),
- **privatezza** (autorizzazione).

Un DBMS deve essere **efficace** ed **efficiente**.

Modello dei dati...

Un **modello dei dati** è un insieme di concetti utilizzati per organizzare i dati di interesse e fornisce meccanismi di strutturazione per relazionarli.

Il **modello relazionale** dei dati permette di definire dei tipi attraverso il costruttore **relazione**, che organizza i dati in record a struttura fissa.

Titolo Film	Regista	Anno
La dolce Vita	Federico Fellini	1960
The addiction	Abel Ferrara	1994
Roma, città aperta	Roberto Rossellini	1945



...Modelli di dati

Gerarchico: si basa sull'uso di strutture ad albero (anni 60).

Reticolare: si basa sull'uso di grafi (anni 70).

Oggetti: riprende il paradigma di programmazione orientato agli oggetti (anni 80).

Uno **schema logico** è una descrizione della base di dati per mezzo del modello di dati di riferimento.

Uno **schema interno** è la rappresentazione dello schema logico per mezzo di strutture fisiche.

Uno **schema esterno** è la descrizione di una porzione della base di dati di interesse, per mezzo del modello logico.



Indipendenza dei dati

L'**Indipendenza fisica** consente di interagire con il DB in modo indipendente dalla struttura fisica dei dati.

L'**Indipendenza logica** consente di interagire con il livello esterno della base di dati in modo indipendente dal livello logico.

I **Data Definition Language** (DDL) sono usati per definire gli schemi logici, interni ed esterni.

I **Data Manipulation Language** (DML) sono usati per l'interrogazione e l'aggiornamento delle basi di dati.

L'**accesso ai dati** può avvenire tramite:

- linguaggi testuali interattivi (SQL);
- tramite comandi immersi in linguaggi di programmazione tradizionali;
- tramite linguaggi di programmazione ad hoc
- tramite interfacce amichevoli.

Utenti e Progettisti

Amministratore della base di dati

- gestire le autorizzazioni di accesso al DB
- garantire sufficienti prestazioni
- garantire controllo centralizzato dei dati ed il soddisfacimento di tutti gli utenti

Progettisti

- usano il DB tramite interfacce (non solo grafiche) e DML

Utenti

- gli **utenti finali** utilizzano le transazioni
- gli **utenti casuali** impiegano i linguaggi interattivi per l'accesso al DB

Il mercato dei DB

Close Source:

- IBM (DB2)
- Microsoft (SQL Server)
- OpenLink (Virtuoso)
- Oracle (Oracle 9i)
- Progress software (Progress)
- Software AG (Tamino)
- Sybase (ASE)

Open Source:

- Informix (<http://www-306.ibm.com/software/data/informix/>)
- MySQL (<http://www.mysql.com/>)
- PostgreSQL (<http://www.postgresql.org/>)

Il fatturato del mercato dei DBMS si aggira intorno ai 8 Miliardi di US\$, di cui l'80% imputabile ai RDBMS e cresce del 10% ogni anno.



Object-Oriented Database

- Birdstep Technology, Inc. (Birdstep RDM Mobile)
- db4objects, Inc. (db4o - database for objects - open source)
- GemStone Systems, Inc. (GemStone/S and GemStone Facets)
- GLinkNET Technology (Link.NET)
- InterSystems Corporation (Cache `)
- Jade Software Corporation (JADE)
- JYD Software Engineering Pty Ltd. (JYD Object Database)
- Logic Arts, Ltd. (VOSS)
- Matisse Software, Inc. (Matisse)
- Micro Data Base Systems, Inc. (TITANIUM)
- Moscow State University (GOODS - open source)
- ObjectDB Software (ObjectDB for Java/JDO)
- Objectivity, Inc. (Objectivity/DB)
- Orient Technologies (Orient Enterprise Edition and Orient Just Edition)
- The Ozone Database Project (ozone - open source)
- Progress Software Corporation (ObjectStore, PSE Pro, and ObjectStore Event Engine)
- run Software-Werkstatt GmbH (ODABA)
- Sorceforge hosted open source (Prevayler - open source)
- Sysra (EyeDB)
- Versant Corporation (Versant Developer Suite, Versant enJin) -- Versant Corporation has merged with POET Holdings (FastObjects, Inc.).

Corrado Aaron Visaggio

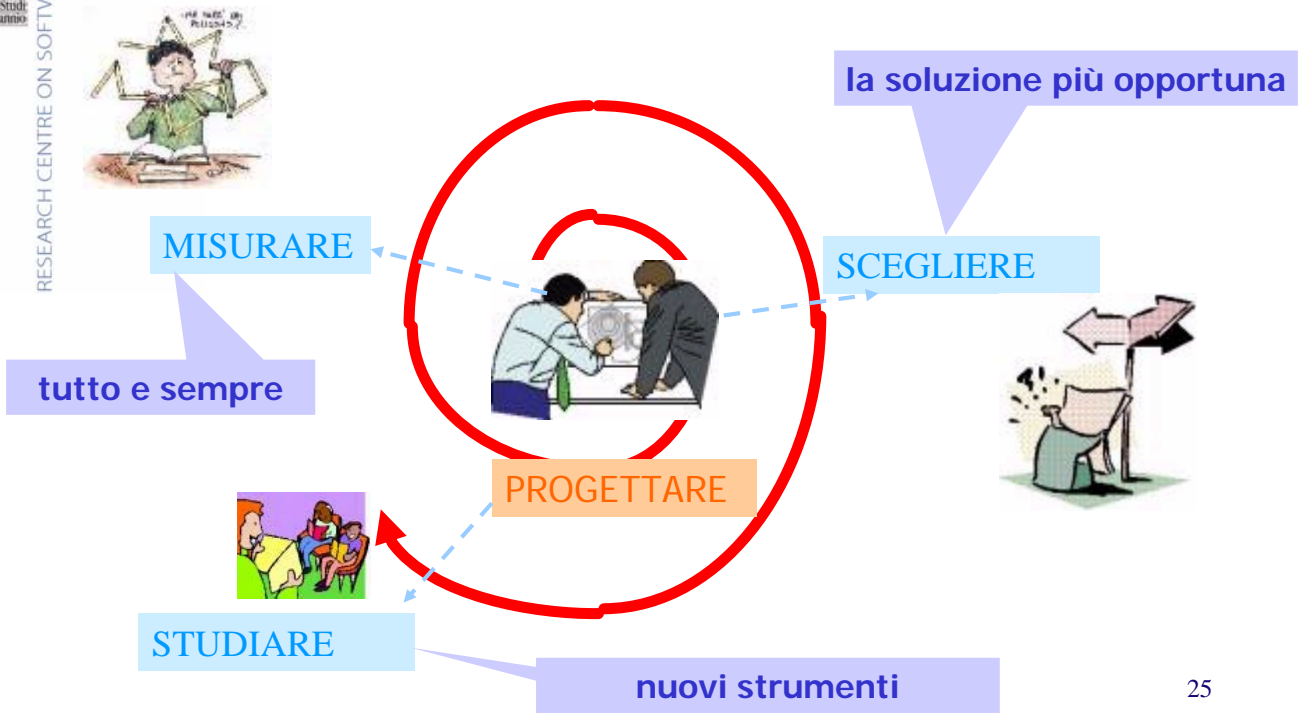
23



Applicazioni dei DB

- Distributed Computing** (Branch Office Automation, Distributed Business, Sales Force Automation);
- Decision Support** (Data Mining, Analytical Processing, Data Warehouse);
- OLTP/Electronic Commerce** (Financial Services, Ticketing, Retailing, Reservation Systems);
- Internet/Intranet** (Application Development; Web Page Generation; Remote Access Data);

Il Lavoro dell'Ingegnere



Any Question?

