

1222 • 2022
8000
ANNI



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Overall Organisation

Basi di Dati

Bachelor's Degree in Computer Engineering
Academic Year 2024/2025



DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Stefano Marchesin

Intelligent Interactive Information Access (IIA) Hub
Department of Information Engineering
University of Padua



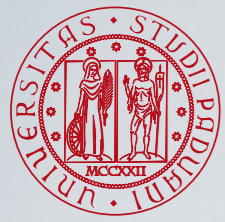


Outline



- Objectives
- Contents
- Organisation
- Exam modalities
- Some logistics

Objectives and Contents



Objectives



To learn how to **design and develop a distributed application for managing and persisting structured data over time**

- Building strong computer science competences on the management of structured data
- Building competences on how to design and develop a “real” database application, by carrying out a project based on a relational database management system [RDBMS]



Contents (1/2)



- Introduction to databases
- Requirement analysis
- The Entity-Relationship Model
 - conceptual design
- The Relational Model
 - logical design
 - relational algebra

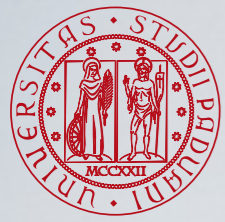


Contents (2/2)

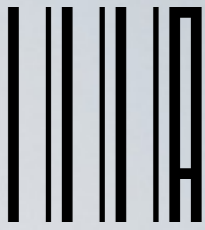


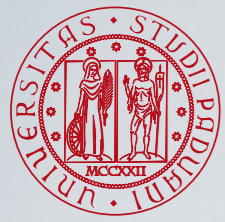
- The SQL language
 - data definition language
 - data manipulation language

- “Advanced” topics
 - Indexes
 - Views
 - Stored procedures

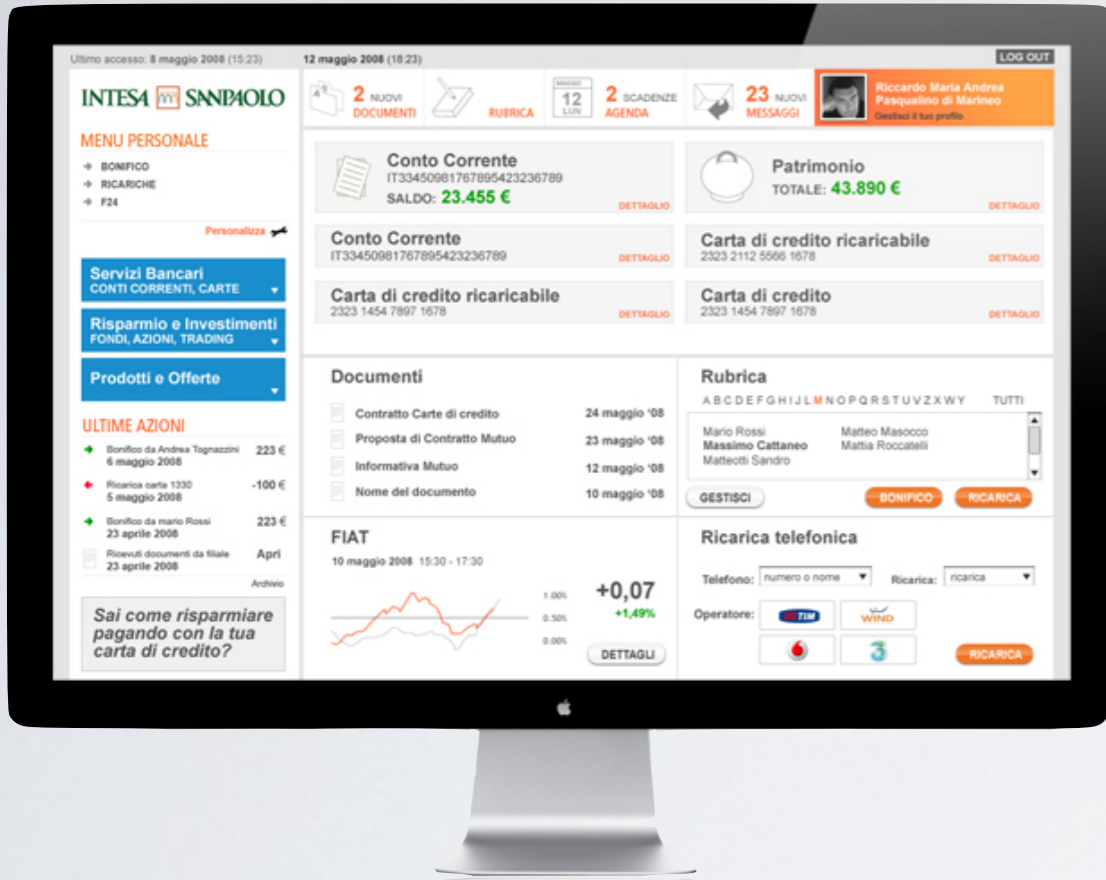


Traditional Applications for a Database



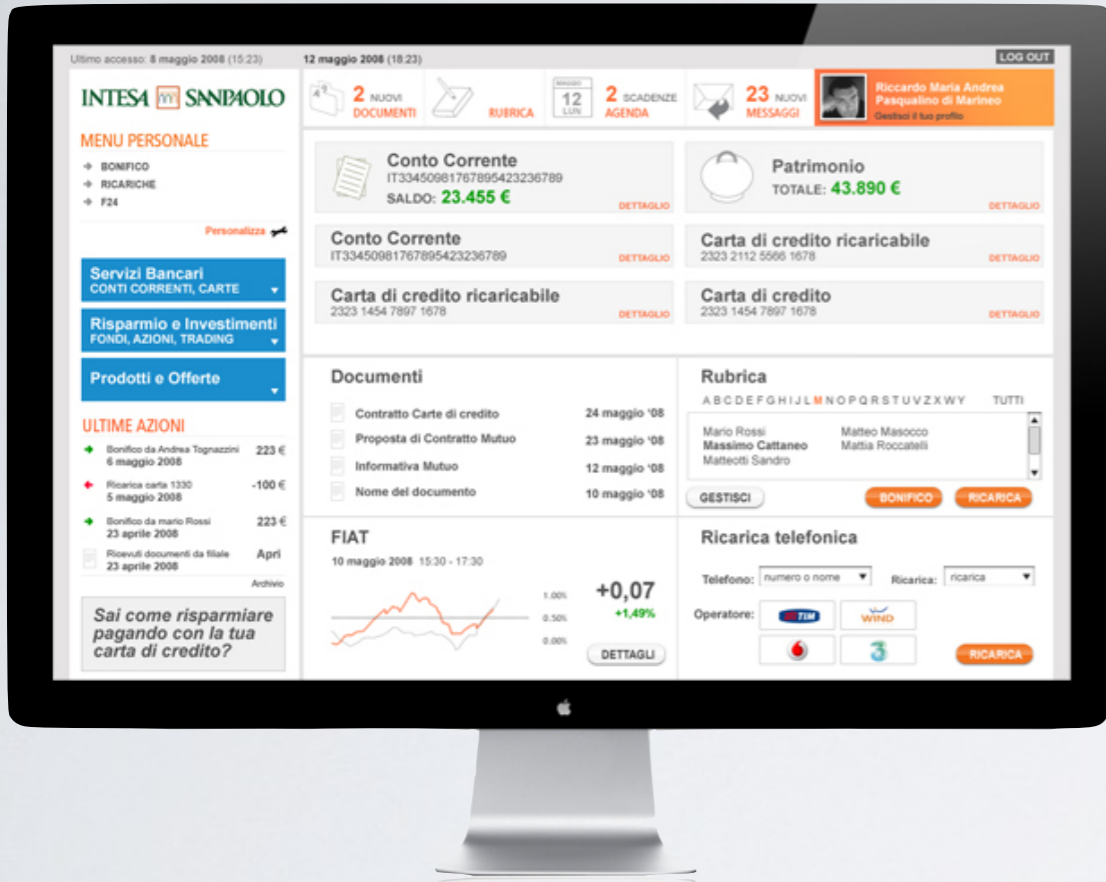


Traditional Applications for a Database



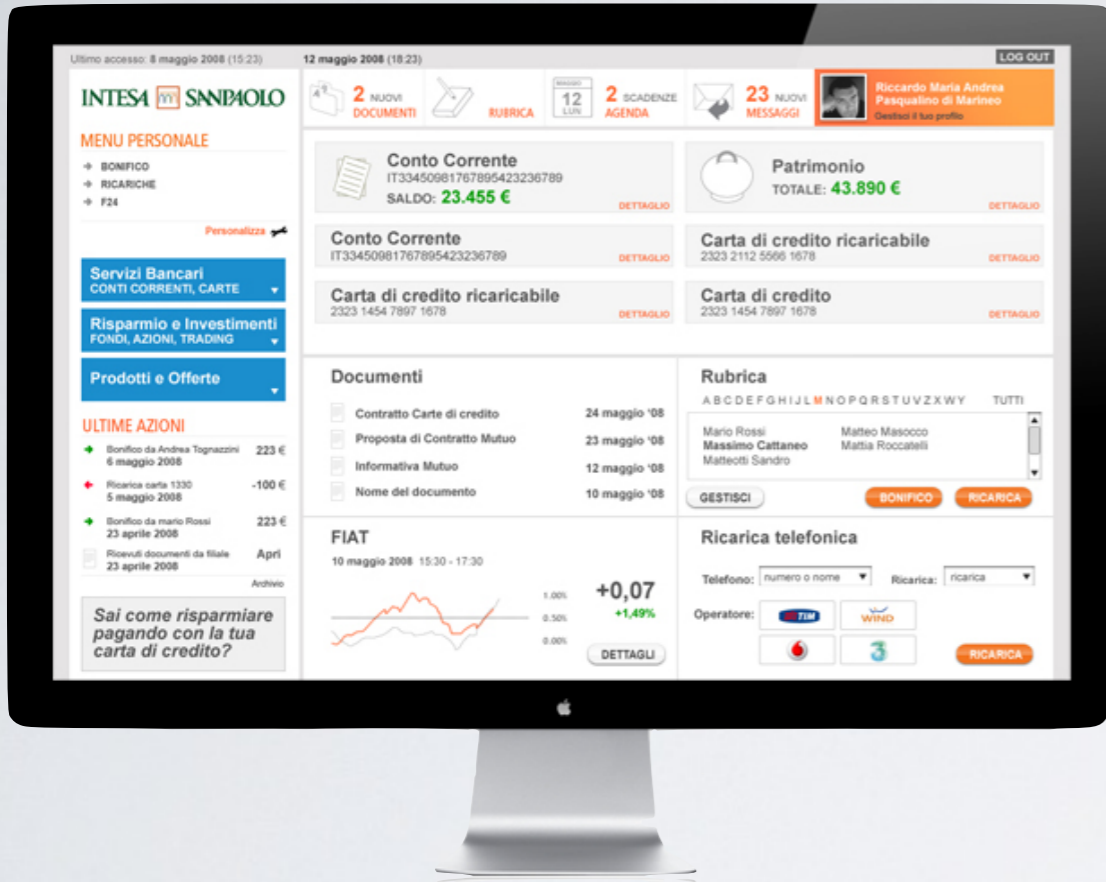


Traditional Applications for a Database



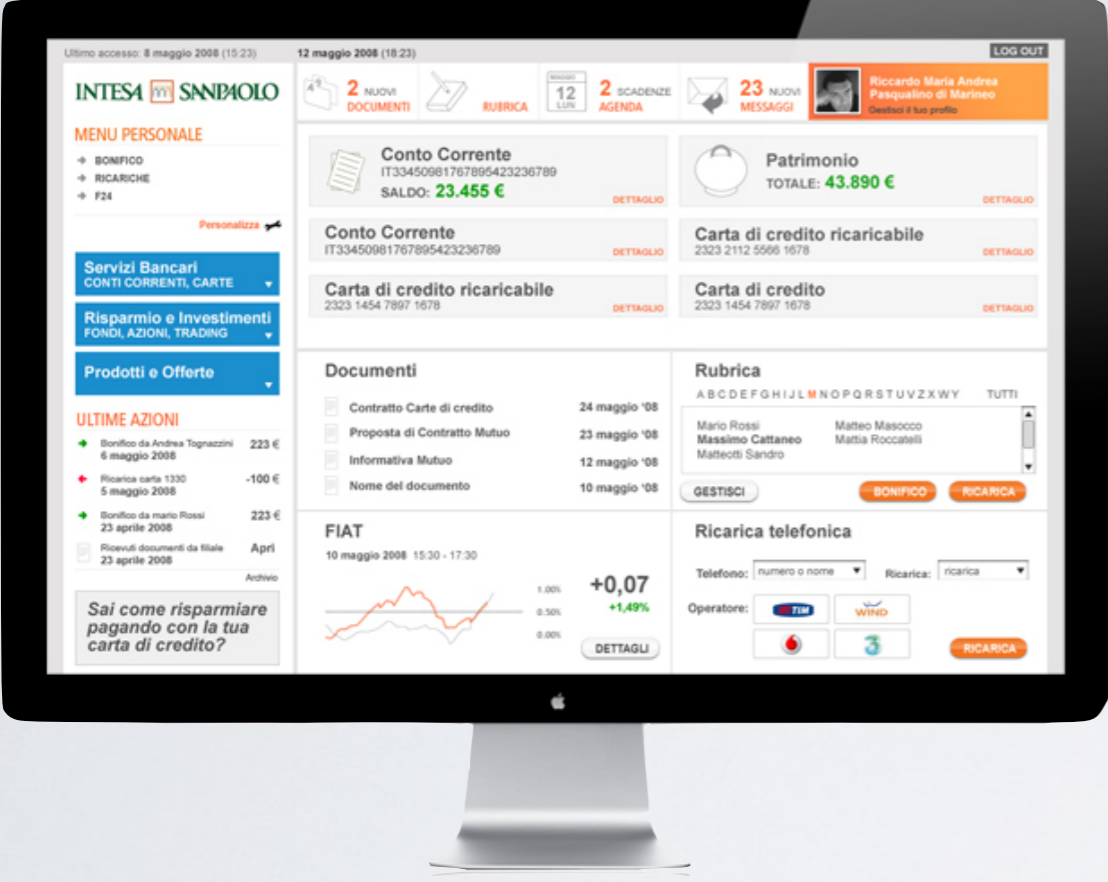


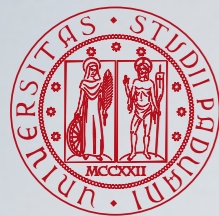
Traditional Applications for a Database



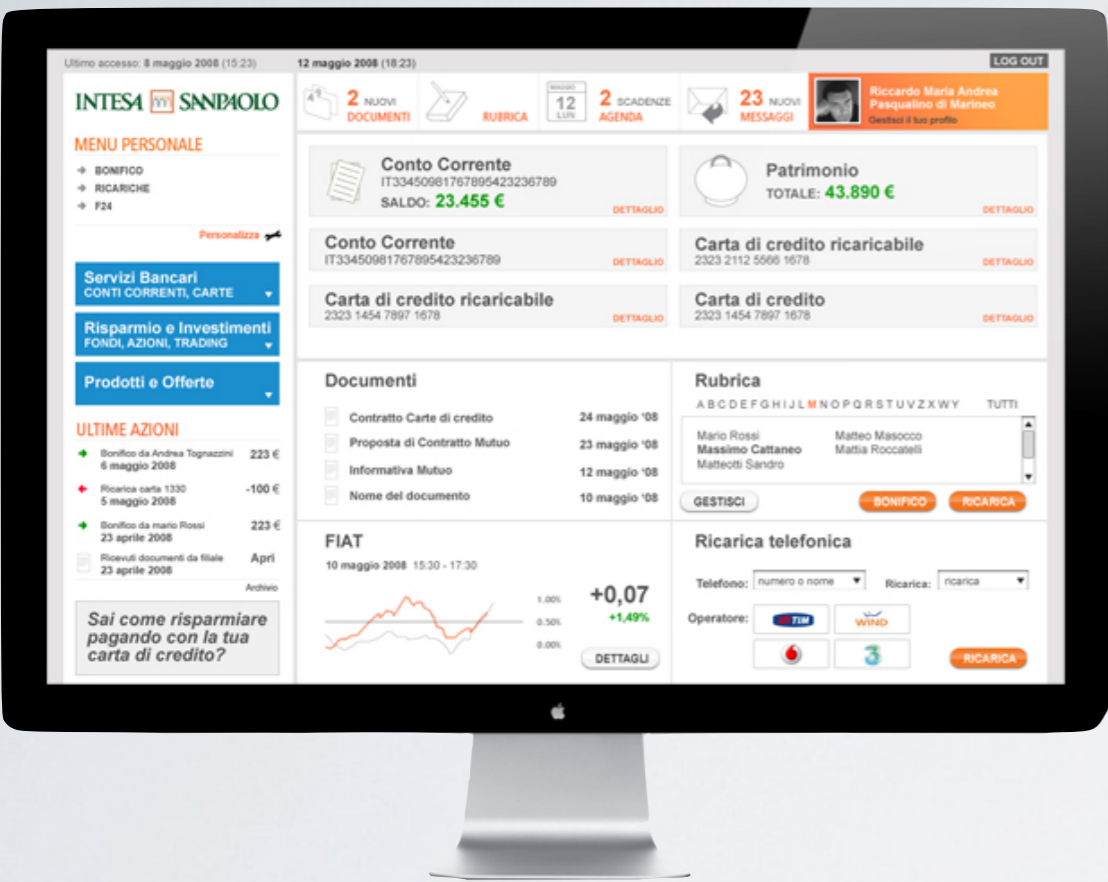


Traditional Applications for a Database





Traditional Applications for a Database





24 Hours with Databases



- **07:30** Wake up and have breakfast, listening music on Last.fm
 - 40 millions of unique visitors each month and more than 500 millions visited pages
 - 63 millions songs, 60 billions scrobble
 - PostgreSQL to manage the catalog, Hadoop to manage scrobble



- **08:30** Check email and Web surf
 - McAfee GroupShield uses PostgreSQL to manage quarantined items



- Trend Micro InterScan Web Security Suite uses PostgreSQL to store data on user activities to prevent phishing, to filter URLs and so on



- **09:30** Call on Skype
 - Skype uses PostgreSQL to store call data and chats



- **10:30** HTTP and DNS [<http://tools.ietf.org/rfc/rfc2616.txt>] e W3C [<http://www.w3.org/standards/techs/http>]



- The Public Interest Registry [PIR] manages more than 10 millions .org domains relying on PostgreSQL to manage DNS tables



- **15:00** Chatting with friends on Facebook
 - 800 million total users, 500 million active users each day [350 million mobile users]
 - 60 million queries per second, 1.5 billion reads second e 4 million writes per second [4 ms to read, 5 ms to write], 90 GByte di network traffic per second
 - MySQL to manage data, Memcached to speed-up querying



- **21:00** Share the party on Twitter
 - 140 million users, 5000 tweets per seconds [400 millions per day], 3 millions writes per second
 - MySQL to manage data



Information Management Systems



Information Retrieval

Databases



What Goes Around Comes Around... And Around...

Michael Stonebraker
Massachusetts Institute of Technology
stonebraker@csail.mit.edu

Andrew Pavlo
Carnegie Mellon University
pavlo@cs.cmu.edu

ABSTRACT

Two decades ago, one of us co-authored a paper commenting on the previous 40 years of data modelling research and development [188]. That paper demonstrated that the relational model (RM) and SQL are the prevailing choice for database management systems (DBMSs), despite efforts to replace either them. Instead, SQL absorbed the best ideas from these alternative approaches.

We revisit this issue and argue that this same evolution has continued since 2005. Once again there have been repeated efforts to replace either SQL or the RM. But the RM continues to be the dominant data model and SQL has been extended to capture the good ideas from others. As such, we expect more of the same in the future, namely the continued evolution of SQL and relational DBMSs (RDBMSs). We also discuss DBMS implementations and argue that the major advancements have been in the RM systems, primarily driven by changing hardware characteristics.

rather than the lasting power of these systems. In other words, there still are many IBM IMS databases running today because it is expensive and risky to switch them to use a modern DBMS. But no start-up would willingly choose to build a new application on IMS.

A lot has happened in the world of databases since our 2005 survey. During this time, DBMSs have expanded from their roots in business data processing and are now used for almost every kind of data. This led to the “Big Data” era of the early 2010s and the current trend of integrating machine learning (ML) with DBMS technology.

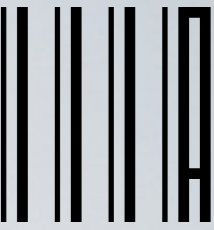
In this paper, we analyze the last 20 years of data model and query language activity in databases. We structure our commentary into the following areas: (1) **MapReduce Systems**, (2) **Key-value Stores**, (3) **Document Databases**, (4) **Column Family / Wide-Column**, (5) **Text Search Engines**, (6) **Array Databases**, (7) **Vector Databases**, and (8) **Graph Databases**.

SIGMOD Record, June 2024. Authors version [here](#)

Organisation



Overall Organisation (1/2)



● Course

- 9 CFU, 72 hours of lectures
- Expected end of the course: 19 December 2024

● Exam Modality

- 20 screening questions (30 min): 20/20 (2 extra points), 19/20 (1 extra points), 18/20 (0 extra points), $\leq 17/20$ **FAILED**
- Written exam (120 min) on database design and modeling: ER schema, Relational Schema, and SQL

● Exams:

- I exam: 22 January 2025 at 09:00 (Room Ve)
- II exam: 10 February 2025 at 14:00 (Room Ve)
- III exam: 30 June 2025 at 09:00 (Room Be)
- IV exam: TBD



Overall Organisation (2/2)



● Office hours

- [send an e-mail] and/or just after lecture

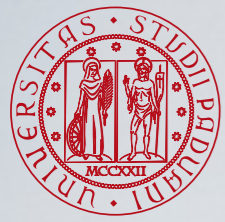
- stefano.marchesin@unipd.it — antonio.giunta@unipd.it

● Suggested textbooks

- Paolo Atzeni, Stefano Ceri, Piero Fraternali, Stefano Paraboschi, Riccardo Torlone, **Basi di dati**, 6th Edition, McGraw Hill, 2023

- Giorgio M. Di Nunzio, Emanuele Di Buccio, **Basi di dati. Progettazione Concettuale, Logica e SQL**, 1st Edition, Esculapio, 2017

- Ramez Elmasri, Shamkant B. Navathe, **Fundamentals of Database Systems**, 7th Edition. Pearson, 2016



Scheduling



Monday

● 10:30 - 12:30, Room Ke

Wednesday

● 10:30 - 12:30, Room Ke

Thursday

● 12:30 - 14:30, Room Ke



Teaching Material



BASI DI DATI (B) 2024-2025 - IN06100770

Corso Partecipanti Valutazioni Competenze Media Gallery Altro ▾

▾ IN06100770 - BASI DI DATI (B) 2024-2025 - PROF. STEFANO MARCHESIN Minimizza tutto

Annunci

Pagina dell'offerta Formativa

Organizzazione Corso

Introduzione & Analisi dei Requisiti

Materiale Didattico

Lezioni

Progettazione Concettuale

Materiale Didattico

Lezioni

> IN06100770 - BASI DI DATI 1 (B) 2024-2025 - PROF. ANTONIO GIUNTA

- Elearning is the central reference for teaching material, official announcements, and so on. If not done already, please, register asap.
- <https://stem.elearning.unipd.it/>
- Please, do NOT use Elearning internal forum/messages to get in contact. Please, use
 - email

Questions?

