

Progetto Sistemi Informativi Geografici per la gestione e la modellazione di dati territoriali
10^a Edition - 2012

A hands-on introduction to databases and geodatabases

A two-sections foundation course with examples in MySQL, PostgreSQL/PostGIS, Oracle XE and ArcGIS/MsAccess for the design and development of Relational (I) and Spatial Databases (II)

Course scope

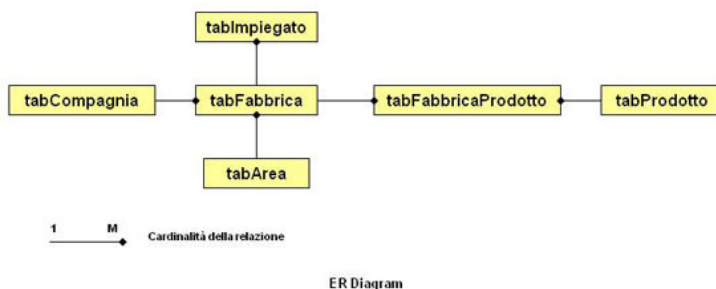
Data acquisition is critical for any decision-making processes. Effective data management is the key to maintain data value and data consistency over time, and it is an essential requirement in designing and administering successful information systems and applications. Improper practice in data management generally results in information and economic loss, leading to erroneous decisions as well as risks for human life in mission-critical applications.

Text files, spreadsheets and binary files, in specialized domains, are widely used for storing digital data. Storing data in separate files can be quick and easy but represents a problem when an application has to be scaled up, involving more data users and longer project life time. A more structured approach in data management must be adopted, the database approach.

A **first section**, following a brief introduction to current standards in textual data encoding, transfer and rules-binding (XML, DTDs and XML Schema), key technologies for the developments in Web 2.0 and Semantic Web, addresses comprehensively the issues mentioned above by introducing theory, design and implementation of relational database systems.

After introducing the fundamentals of Web **RDBMSs** (**R**elational **D**ata**B**ase **M**anagement **S**ystems), namely MySQL (and WAMP) and PostgreSQL/PostGIS from Open Source community and free industry standard Oracle XE, the course adopts a hands-on approach addressing design, development and deployment requirements of two sample database standards by providing examples (and data) from business management and environmental monitoring information systems,.

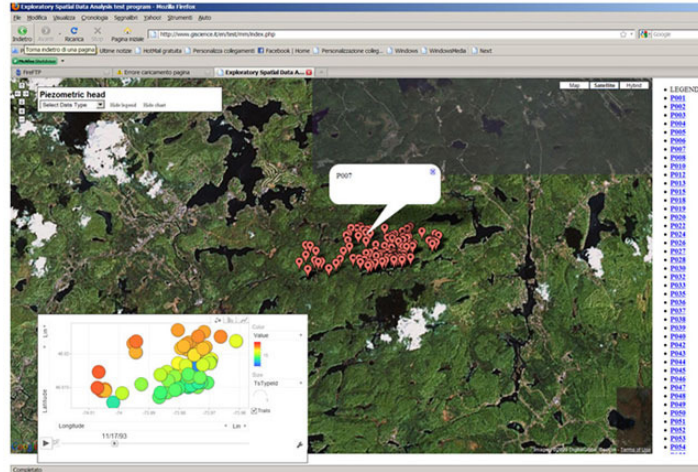
The databases creation will provide the opportunity to review basic and advanced concepts from a practical perspective, including an in-depth analysis of SQL, the standard foundation language to build and query relational database.



A **second section** covers advanced topics in spatial databases (or geodatabases), building on previous body-of-knowledge and addressing issues in design, development and deployment data management in geographical applications. Provided a general overview of spatial databases, different approaches are investigated and critically reviewed from a technical and economical perspective, including elementary approaches based on native

data types addressing simple point geometry, geographic data types in PostGIS, Oracle XE and ESRI geodatabase model. The course includes a brief introduction to and critical review of open-source and proprietary (web) database front-end solutions, namely uDig, QuantumGIS, OpenJump and ESRI ArcGIS, further to basic concepts in designing and development bespoke web mapping mashup solutions.

Disclaimer: Please be aware that the course is intended as a foundation course on theory and practice in RDBMS and geodatabase design, development and deployment, in order to build a consistent body-of-knowledge and awareness about key database issues, rather than a fully-featured introduction to any (geo)database platform. However, worked hands-on examples, course material and adopted textbooks are intended to provide a relevant gateway to specific platforms, supporting participants in their pathway toward development of professional and robust applications.



Case study: an ESDA (Exploratory Spatial Data Analysis) geodatabase-driven Google Maps mashup web application - E. Crestaz, 2009

Course program

SECTION I – Relational databases

Provided an introduction to fundamentals in data management and relational database management systems design and development, introduction to and installation of RDBMS platforms MySQL, PostgreSQL/PostGIS and Oracle XE, sample case studies review, the first part of the course will focus on RDBMS theory, SQL and non-geographic applications development.

Data management and database approach

- ❑ Issues in data management: data redundancy, integrity, sharing, maintenance costs, consistent interfaces, standards, security, promotion of corporate view. A socio-technical and economic perspective of different practices in data management.
- ❑ Text files, spreadsheets and binary files: when they work and when they do not. Brief introduction to **XML** (eXtensible Markup Language), the web lingua-franca for data encoding, **DTDs** (Document Type Definition) and XML Schemas for content-binding management. Fundamentals in ontology and semantic issues.
- ❑ The database way, definitions and **RDBMS** (Relational DataBase Management System) model. Data structures, records editing, data retrieving, security, integrity and applications building. Indexing systems. Design and implementation sequence: logical and physical design, testing, implementation and maintenance.

RDBMS platforms and course sample case studies

Desktop single-user database management systems (i.e. MsAccess): potentials and limitations. Fundamentals in web database management systems, client-server architecture, **HTTP** (HyperText Transfer

Protocol). Introduction to and installation of RDBMS software platforms, from Open Source community and from industry:

- ❑ [WAMP™](#) (Windows Apache MySQL and PHP) and [MySQL™](#) fundamentals in database development and querying in phpMyAdmin and MySQL console.
- ❑ [PostgreSQL](#) and spatial database extender [PostGIS](#) plus their web administration tool phpPgAdmin for integration in the WAMP stack.
- ❑ [Oracle XE](#):• Oracle XE: the free, also for commercial use, scalable professional entry point to Oracle enterprise solutions, addressing the requirements for advanced spatial information management (Locator features, a subset of Spatial Oracle).

Overview of the two course sample case studies: factory business data and environmental monitoring data management. Identification of initial entities and relationships, EAR (Entity Attribute Relationship) diagrams, normalisation.

Relational database model foundations and SQL

- ❑ Fundamentals: tables, columns and rows, primary and foreign keys, unique constraints, null values, relationships (one-to-one; one-to-many; many-to-many and junction tables)
- ❑ Normalization rules: atomic values and no repeating groups (1st), no partial functional dependencies (2nd) and no transitive dependencies (3rd)
- ❑ Database design: **EAR (Entity-Attribute-Relationship)** modelling, identification of entities, their attributes and relationships.
- ❑ SQL basics: syntax, standards, identifiers, data types (character strings, numeric, Boolean, datetime, **BLOBs – Binary Long Objects**)
- ❑ Data retrieving: SELECT and FROM, duplicate rows (DISTINCT), sorting (ORDER BY), filtering (WHERE), conditions combining and negating (AND, OR and NOT), patterns matching (LIKE), range and list filtering (BETWEEN, IN), testing for NULLs (IS NULL).
- ❑ Operators and functions: arithmetic, string operations (concatenation, extraction, case, trimming, length, searching), datetime, data types conversion, conditional values, nulls checking.
- ❑ Data summarizing and grouping. Aggregate functions (MIN, MAX, SUM, AVG, COUNT), grouping rows (GROUP BY), filtering groups (HAVING)
- ❑ Joins: table aliases (AS), joins building with JOIN or WHERE, cross, natural, inner, outer and self joins. Subqueries.
- ❑ Creating, altering and dropping tables. Inserting, updating and deleting rows. Indexes and views.

SECTION II – Spatial databases

Geographic data management and geodatabase approach

- ❑ Brief review of geographic data management and their complexity (geometries, projection systems, topology).
- ❑ **OGC (Open GIS Consortium)** and de-facto file-based industry standards (ESRI shape).
- ❑ The spatial database (geodatabase) way, definitions and extended RDBMSs model, including specialized geographic fields to store complex geometries and projection systems. Spatial indexes, as RTree++, for efficient access and retrieval of geographic component.

Geographic database platforms

Review and critical assessment of different solutions, other than simple point geometries management in native SQL data types:

- ❑ ESRI – the geodatabase paradigm, concepts of feature dataset, feature class, object class and relationship class, shape field and domains. Single user personal geodatabase in MsAccess vs. enterprise multi-user solutions, ArcGIS Server. ESRI licensing policies vs. geodatabase activated features. Available free industry-standard geodatabase models.
- ❑ Oracle XE: fundamentals in **SQL/MM (SQL/Multi-Media)** standard towards management of geographic data type. Conceptual boundaries between storing spatial data in a geodatabase and performing spatial analysis. Overview: data model, query, analysis and visualization. Oracle spatial technology: locator and spatial option. How to spatially enable applications: adding location information to tables and spatial metadata. SDO_GEOMETRY data type: spatial geometries, type, attributes and values (SDO_GTYPE, SDO_SRID, SDO_POINT, SDO_ELEM_INFO and SDO_ORDINATES), simple geometry examples. Spatial operators. Loading, transporting and validating spatial data. The way forward: basic review of enterprise solutions, few added built-in functionalities of Oracle Spatial (geocoding, spatial operators).
- ❑ PostGIS: system architecture, integration with PostgreSQL, pgAdmin III for system administration. Geometry and geographic data types. Data modelling strategies, functions, geographic reference systems and spatial analysis.

Geographic database front-end applications

Applications addressing geographic visualization requirements of course sample case databases: hands-on review of different options in geographically enabling the environmental monitoring geodatabase by storing point coordinates through:

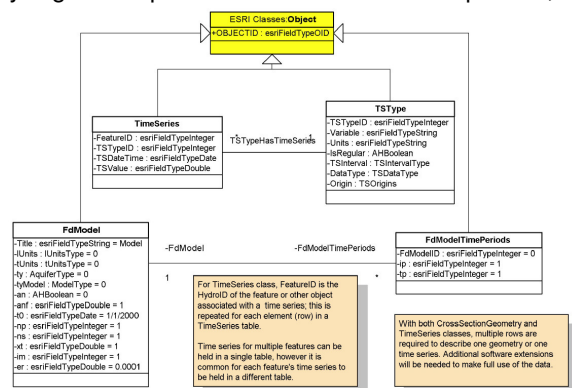
- ❑ native data types;
- ❑ geographic data types in PostgreSQL/PostGIS and Oracle XE;
- ❑ ESRI geodatabase, following the Hydro Data Model developed by ESRI and CRWR (Center for Research in Water Resources, Austin Un., Texas)

Introduction to and critical review of open-source and commercially available single user and web database front-end solutions, namely uDig, QuantumGIS, OpenJump and ESRI ArcGIS, further to basic concepts in designing and development bespoke web mapping mashup solutions, extensively addressed by the “Google Maps Mashups” course.

GIScience project

The course is part of a wider learning project in GIScience and GIS, applications and geographic mashup development, environmental applications. Supported by high-level professional and academic profiles, the project includes following courses:

- ❑ **Introduction to GIScience:** principle and techniques for application development, visualisation and spatial analysis
- ❑ **Design and development of GIS software applications:** object oriented model, software engineering, Python programming and its role in geoprocessing in ArcGIS
- ❑ **Design and development of Webmapping applications:** a course on theory and practice in GeoWeb 2.0 mashups design and development, integrating Google Maps and OpenLayer web services and spatio-temporal PostgreSQL/PostGIS database



- ❑ **Design and development of GIS software applications:** software engineering, OO and components models, programming in VBA/ArcObjects for ArcGIS 8.x/9
- ❑ **Object Oriented JAVA programming for GIS applications development:** a Open Source perspective to stand-alone and Web solutions
- ❑ **A hands-on introduction to databases and geodatabases:** A two-modules foundation course with examples in MySql, PostgreSQL/PostGIS, Oracle XE and ArcGIS/MsAccess for the design and development of Relational (I) and Spatial Databases (II)
- ❑ **Business Geographics & Geodemographics**
- ❑ **Google Maps Mashups**
- ❑ **Environmental project management**
- ❑ **Applied groundwater flow and transport modeling:** groundwater management at regional scale for water supply policies assessment and at local scale for environmental remediation of contaminated sites

Following a complex and coherent learning path, after traditional part-time and distance learning philosophy, the course has been based, since its first edition in 2003, on both formal classes and self-learning periods, making course accessibility easier and learning objectives realistic .

Software

WAMP/MySQL, PostgreSQL/PostGIS, Oracle XE, ArcGIS/MsAccess, uDig, QuantumGIS, OpenJump

Who i sit for?

The course is targeted to professional working in the public sector and young researchers that are interested in the foundations of database theory and in the practical issues related to efficient data management and database design and development using commercial and open source platforms. The course is structured in two sections. A first module covers the foundations of database design and development, with a focus on SQL and front end web applications. The second module more advance topics geodatabase theory and focus on geodatabase management for dynamic web mapping applications and WebGIS.

Location & dates

The course will take place at the Environmental Education Centre [Casa Archilei](#), located in Fano (PU) and easily accessible from major transport links. The course will have a duration of 40 hours, with 8 hours lesson. The course will have a duration of **48 hours**, with 8 hours lesson on **20-21 April and 4-5-18-19 May**.

Registration and payment options

Course fees are **1350 € + VAT (21%)**, to be paid in a unique installment before starting date. Fees are reduced to **1250 € + VAT (21%)** for participants with their own notebook.

Fees reduction

Registration before March 31th, 2011	5%
Partecipanti a precedenti corsi promossi da www.GIScience.it Students regularly enrolled in degree, post-degree and Phd courses at italian or foreign universities at the moment of registration Education Institutions, public and private companies enrolling more than one person	10%

Registration modules is available for download in [PDF](#) format or they can be requested at info@giscience.it

Certificate

At the end of the course, participants attending at least 70% of the lessons will be awarded a Completion certificate.

Didactic material

The course adopts the following test book, included in registration fees:

- ❑ Fehily, 2008. SQL. 3rd Ed., Visual Quickstart guide. Berkeley, CA
- ❑ Obe R. and Hsu L., 2010. PostGIS in Action. Manning

Basic web references

OracleXE	http://www.oracle.com/technetwork/database/express-edition/overview/index.html
WAMPServer	http://www.wampserver.com/en/
MySql	http://www.mysql.it/
PostgreSQL	http://www.postgresql.org/
PostGIS	http://postgis.refrations.net/
uDig	http://udig.refrations.net/
QuantumGIS	http://www.qgis.org/
OpenJump	http://www.openjump.org/
ESRI geodatabase models	http://www.esri.com/software/arcgis/geodatabase/data-models.html

Further recommended references

- ❑ Bobrowski S., 2006. Hands-on Oracle database 10g Express Edition for Windows. Osborne Oracle Press series, CA, USA
- ❑ Kothuri R., Godfring A. and Beinat E., 2004. Pro Oracle Spatial: the essential guide to developing spatially enabled business applications. Apress, Springer-Verlag, New York, USA
- ❑ Zeiler M., 1999. Modelling our world. ESRI Press, CA, USA

Books preview available at [Google Books](#)

Info

For further information, please refer to info@giscience.it Email address or refer to [contact](#) references at Contacts section of web site www.giscience.it .