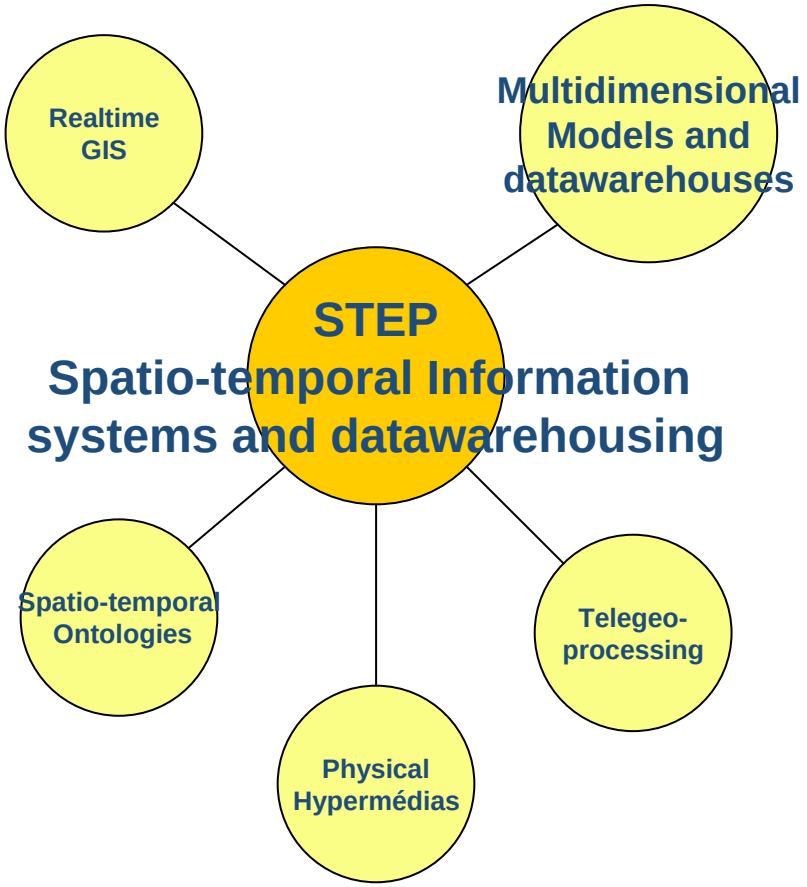
# Spatial On Line Analytical Processing per Dati Ambientali

*S. Bimonte, A. Tchounikine, M. Miquel, R. Laurini*

*LIRIS, INSA Lyon, France*

# LIRIS / STEP

**Staff** : 21 full professors ; 54 associate professors ; 110 Ph.D. students ; 4 full time researchers ; 7 Post-docs (total ≈ 200)




4 axis :

# 1 : Data, Documents and Knowledge

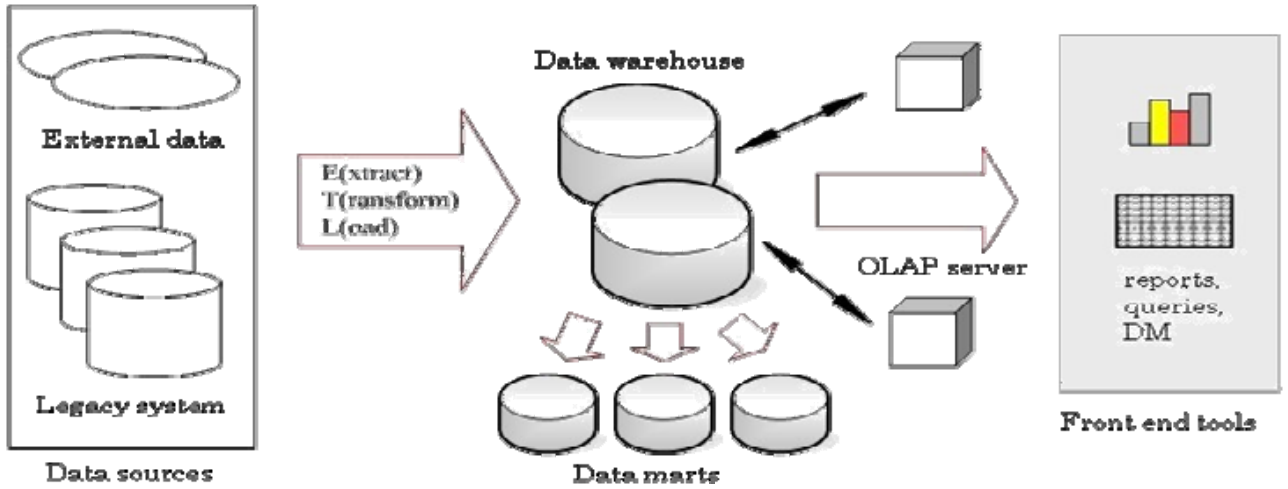
# 2 : Image and Videos: Segmentation and Information Extraction

# 3 : Virtual Reality

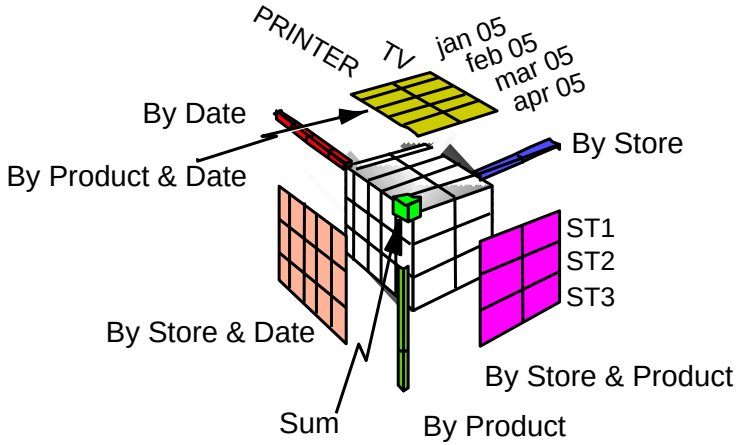
**# 4 : Information and Communication Systems**

# Data warehousing and OLAP

"A data warehouse is a subject-oriented, integrated, time-variant, and nonvolatile collection of data in support of management's decision-making process." (Inmon)

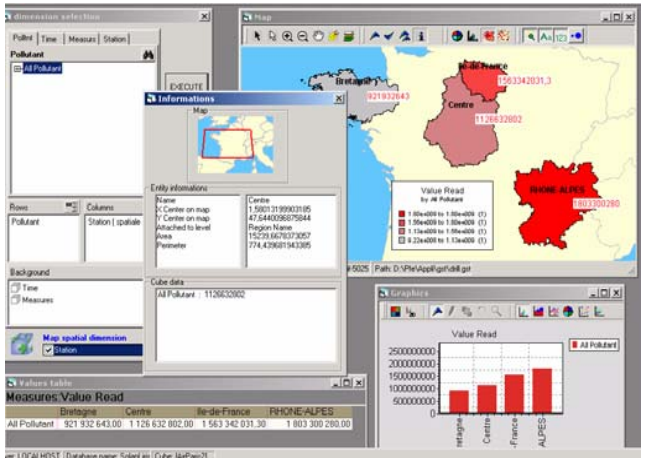



- gather and homogenize data
- guarantee the quality of data
- fast access to summarized data
- interactive navigation through data at different levels of detail

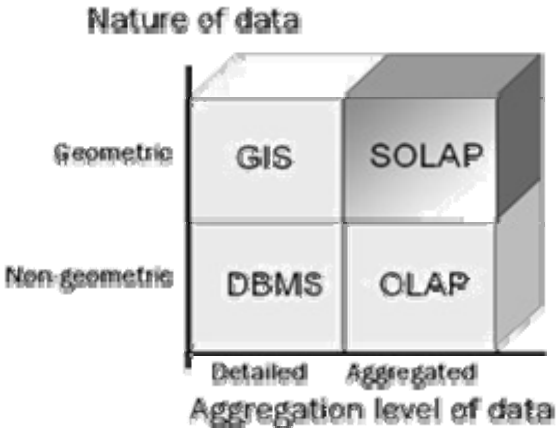


# SOLAP

- Motivation
  - 80% ~ of transactionnal data contains spatial information
  - Enhance decision making
    - visual information
    - discover relationships, correlation beetween phenomena

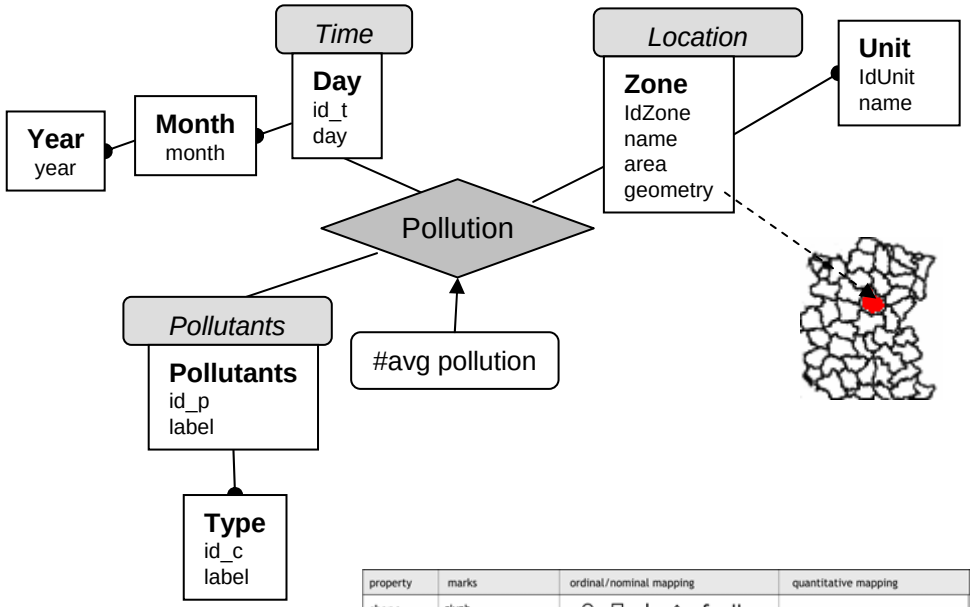


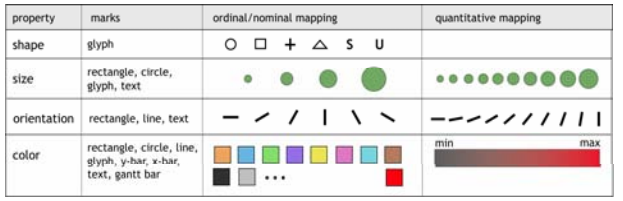
- Definition
  - "A visual platform built especially to support rapid and easy spatio-temporal analysis and exploration of data following a multidimensional approach comprised of aggregation levels available in cartographic displays as well as in tabular and diagram displays" [Bédard 97]
  - SOLAP > GIS + OLAP
- Applications
  - Health (Supervisation of epidemics, environmental health...)
  - Road network management, physical planning
  - Environment (pollution, ...)
  - Population flow...



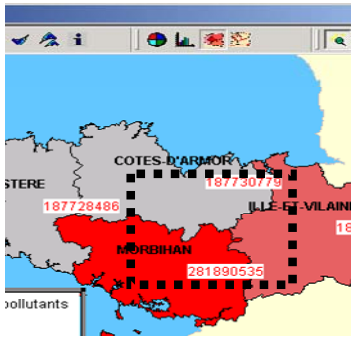
# Spatial dimension

- Definition/usage
  - *"Avg polllution in dec. 2000 with pollutant = "FERRO" in location FIUME DESE ?"*
- Some issues
  - Modelization and storage:
    - integrate a geometric information
    - spatial indexes
  - Visualization:
    - spatial members at different levels of granularity
    - visual display for measure(s) and non spatial dimension(s)
  - Algebra:
    - integrate spatial OLAP operators in OLAP queries



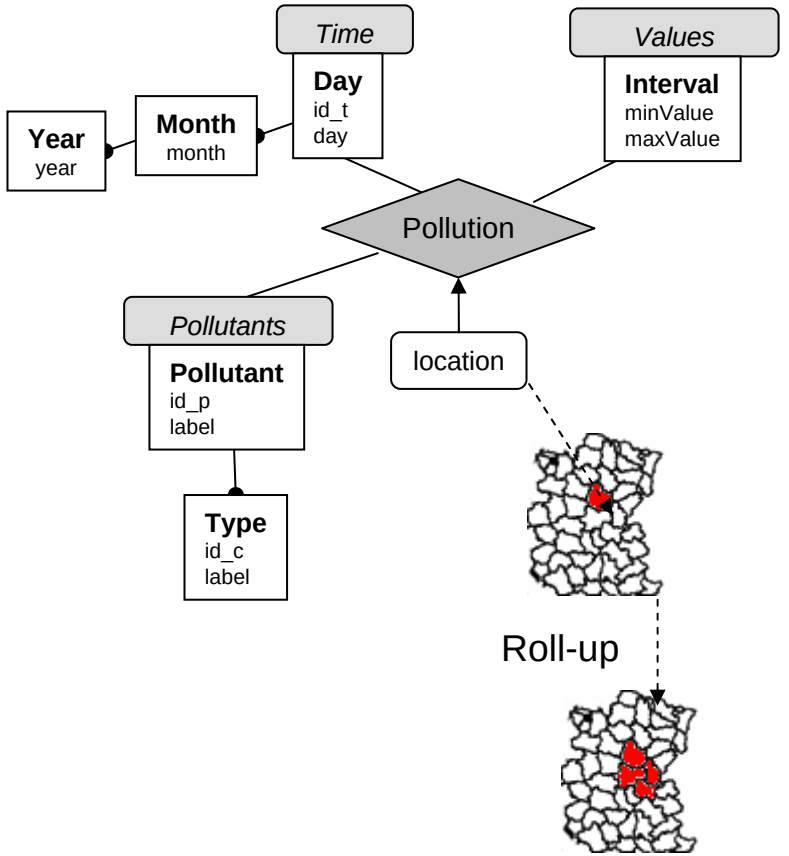




[Polaris]



# Spatial measure

- Definition/usage
  - "*Where did FERRO exceeded 0.5 in dec. 2000 ?*"
- Some issues
  - Computation of hypercube:
    - pre-aggregation with un-classical and time-consumming aggregation operators [Stefanovic 00]
  - Visualization:
    - (multiple) cartographic cells into crosstabs (visual abstraction/data abstraction)
  - Modelization:
    - integration of the descriptive and metric attributes of the spatial measure that can be useful to the decision process



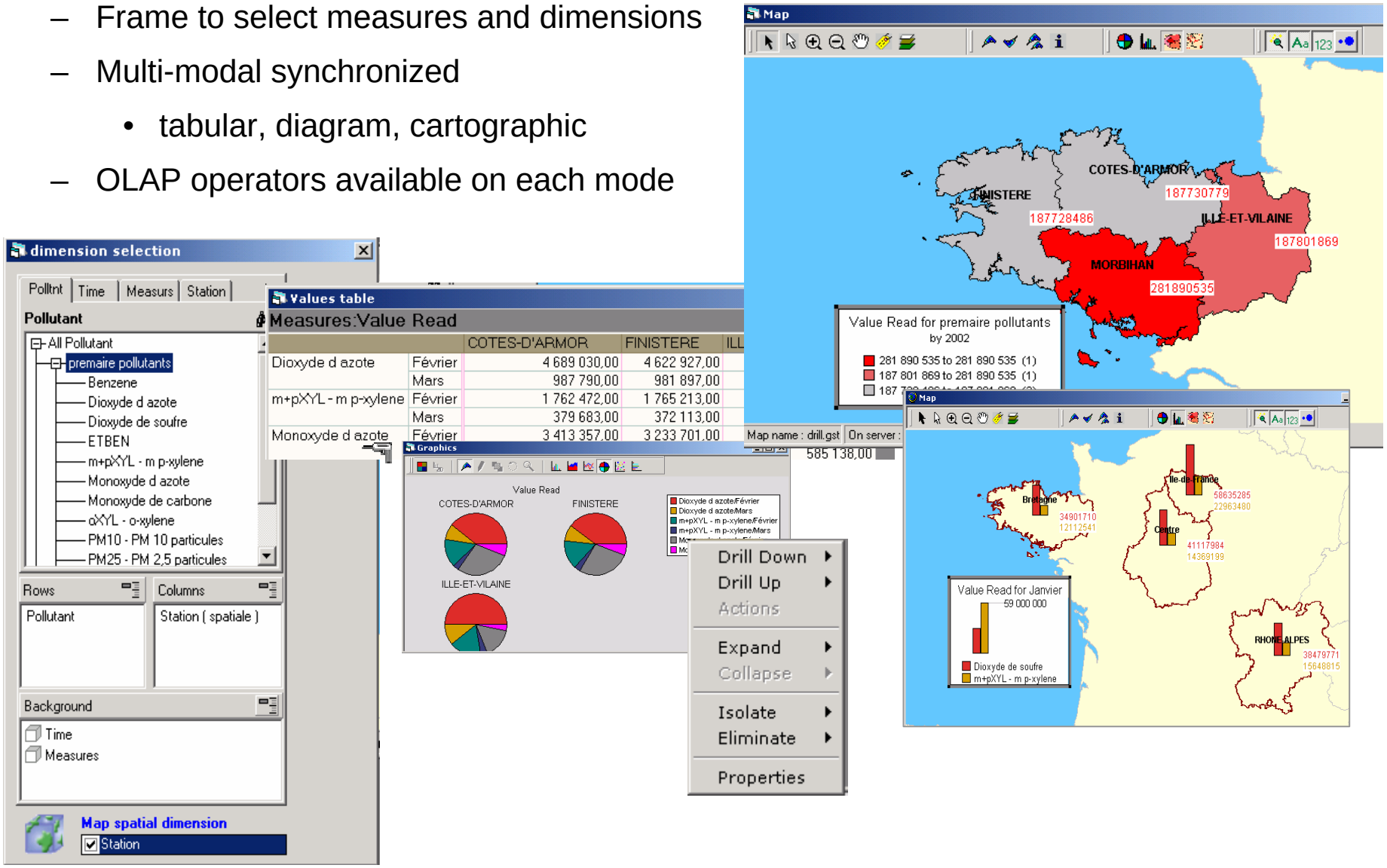
# Formal models

- Model for spatial dimension and/or spatial measure
  - provide a symmetric treatment of measure and dimension
  - allow the definition of spatial dimension and/or spatial measure taking in account all spatial, metric and descriptive attributes : towards a geographic measure and a geographic dimension
  - define an aggregation mode in order to perform aggregation on geographic complex measure
- Integration of the continuous nature of spatio-temporal data
  - Recuperation of hidden and missing data
  - Analysis at very low level of detail
  - simulate spatio-temporal continuity during navigation thanks to on-the-fly interpolation of the data

# Prototype

GéOlap

- Functionnalities
  - Frame to select measures and dimensions
  - Multi-modal synchronized
    - tabular, diagram, cartographic
  - OLAP operators available on each mode



# Drain Project

- 8 data files
  - Drain1, Drain2, … Drain8  in excel and csv format
    - columns :
      - location : 10 locations (FIUME OSELLINO, FIUME DESE…)
      - date : 2826 timestamps (3 years, 28 months, 207 days)
      - pollutants : 100 pollutants
    - pollution values : 294425
- Problems
  - Heterogeneity (types and values)
  - Multiple inputs
  - Missing data (sparcity)
  - Redundant data

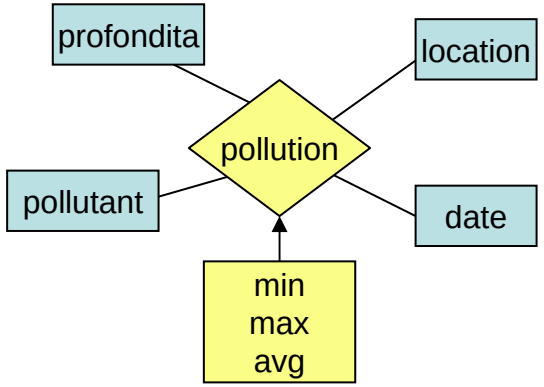
# Drain Project

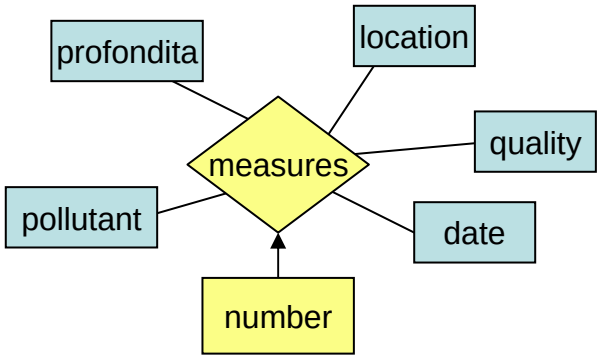
- Migrate the files to SQL Server
  - add one table for the pollutants and turn column pollutant to foreign key
  - add one attribute to qualify the pollution value

- Build 2 multidimensionnal cubes

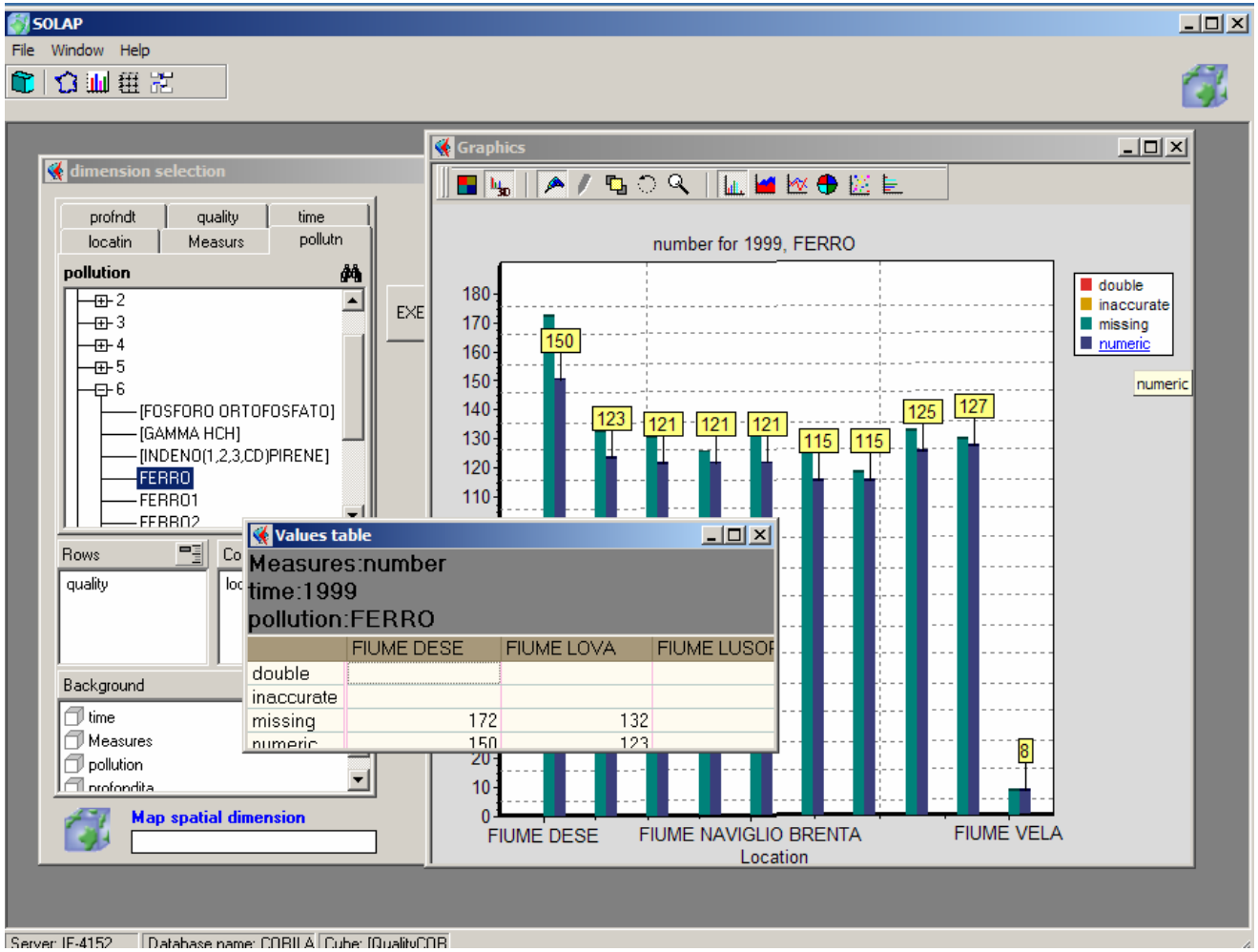
  - CubeCorila : with selected "clean" data
    - 26527 numeric pollution values
    - litteral values for "profondita"

  - CubeAudit
    - all pollution values (294425)
    - quality measure (numeric, double, inacurrate, missing)

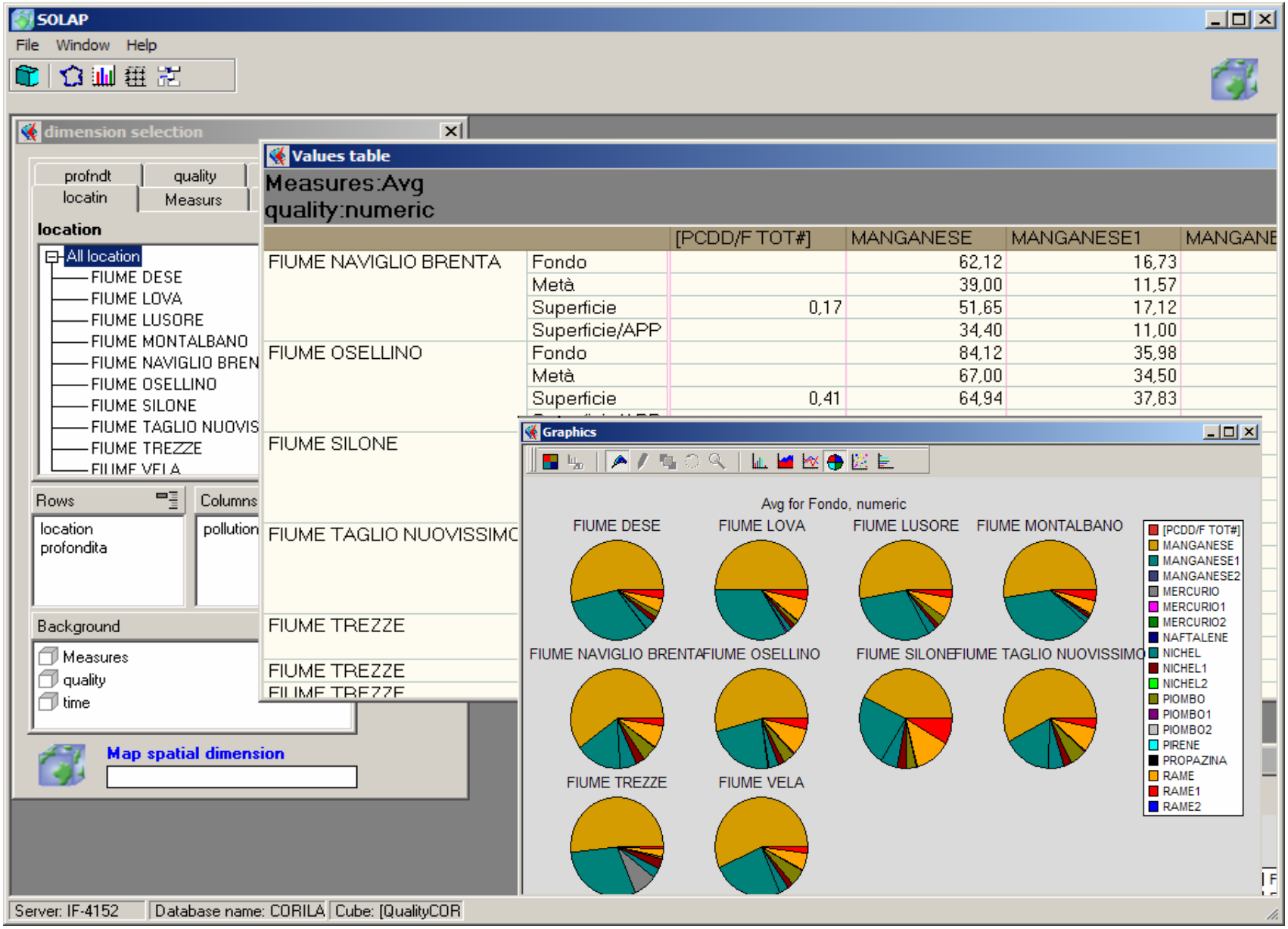


# Drain Project



# Drain Project



# On going works

- Spatial visualization of the laguna for the drain data
- Web version of GéOlap
- SOLAP operators