Federated Data Model to Improve Accessibility of Distributed Cadastral Databases in Land Administration

A. M. Tuladhar, M. Radwan (the Netherlands),
Fatma Abdel Kader and Samir El Roby (Egypt)

FIG Working Week 2005 and GSDI-8
From Pharaoh to GeoInformatics
Intercontinental Samiramis, Cairo, Egypt, 16-21 April 2005

Content

- Introduction
- Cadastral data management
- Federated database management and three Key dimensions of FDBM
- Architecture and schema mapping
- Alternative approach using UML/OCL
- Preliminary findings
- Conclusions

Experimenting data models at Egyptian Survey Authority, Cairo, Egypt (photography by Tuladhar)

Introduction

- GIS manages a large volume of cadastral data and other related data such as data about market value, tax and use of land
- Many agencies involved in a distributed and heterogeneous processing environment
- Accessing data and integrating them in a single synchronous, consistent dataset using GIS software
  - Tremendous challenges
  - High resource, time and money
- Federated data model - a solution?
  - Integrating layers on the top of existing database systems
  - Reduce complex tasks at the users' desks

Cadastral data management

- Many developing countries organize their land data in separate agencies
  - Cadastral Parcel data
  - Topographic data
  - Ownership data
  - Tax data
- Managed at different location at central, provincial and local levels
- Egypt
  - Egyptian Survey authority (Ministry of public work and Irrigation)
    - Cadastral data
    - Topographic data
  - Real Estate department (Ministry of Justice)
    - Ownership data and registration data
  - State owned Land Agency
    - State owned land data
  - Tax department (Ministry of Finance
    - Revenue Tax Collection

Federated Database system (FDBS)

- FDBS - a collection of cooperating but autonomous component databases
  - A component db can participate in more than one federation
  - Integrated to various degree
    - Depends on the needs of federation users and desires of the administrators of the component dbs to participate in the federation and share their databases

Three key dimensions of FDBS

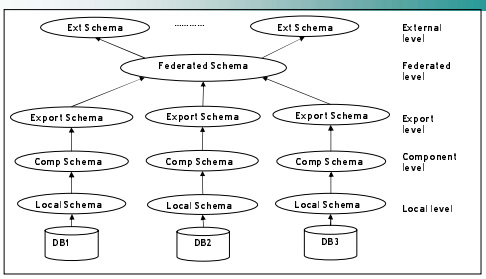
- Distribution
- Heterogeneity
- Autonomy

- Distribution
  - Multiple databases on a single/multiple computer systems - increased availability and reliability as well as improved access time
  - Existence of multiple Databases before FDBS

Three key dimensions of FDBS

- Heterogeneity
  - Database management systems
    - Difference in data models
    - Direct impacts in data structure, constraints and query languages
  - Semantic heterogeneity (Bishr, 1997)
    - Difference in hierarchies
    - Differences in classes
    - Differences in geometry
    - Differences in attributes lists and domains
- Autonomy
  - Organisations that manage databases are often autonomous - DBs are under separate and independent control
  - Three types
    - Design autonomy - ability to choose its own design. This contributes to the problem of heterogeneity
    - Execution autonomy - ability to execute local operations without interference from external operations
    - Association autonomy - ability to decide whether or how much to share its functionality and resources with others.

Federated Database Architecture

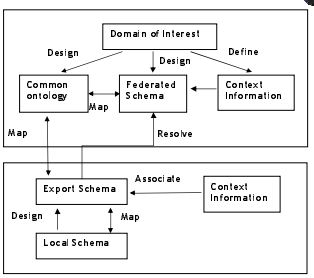


Federated database architecture (Benchikha, Boufaida and Seinturier, 2001)

- Critical issue is to map export schema to federated schema
- resolving semantic heterogeneity

Mapping Export Schema to Federated Schema

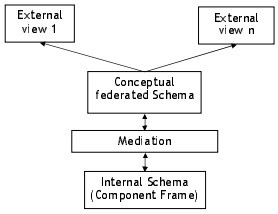
- Definition of common ontology for context to share
- Mapping its elements on federated schema
- Resolve semantic heterogeneity



Mapping between Export schema and Federated schema (Bishr, 1997)

Alternative approach for Land Administration (1)

- The above approach requires common ontology and precise design of domain within context
- Alternatively we can use simple approach for land administration based on mediation using three-tier architecture

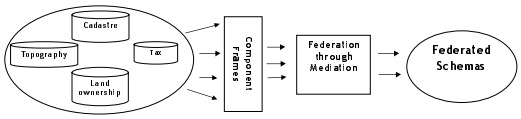


Three-level Federation Architecture based onMediation

Alternative approach for Land Administration (2)

- UML/OCL - the notion of derived class facilitates as a means to integrate certain classes and constraints which need to be shared in a federated context
- A mediator class is used for deriving federated schema
- Conflicts can be resolved within a mediator class (Balsters, 2003a and 2003b)
- Resolving conflicts
  - Renaming (homonyms and synonyms)
  - Data conversion (different data types for related attributes)
  - Default values (adding default values for new attributes)
  - Missing attributes (adding new attributes in order to discriminate between certain class objects)
  - Sub classing (creating a common super class and subclasses

Federating distributed databases



Provincial databases

- Component frame is specified in UML
- Semantic heterogeneity is analyzed to detect conflicts
- Introduce an Class mediator
- Construct an integrated schema
- Resolve conflicts via mediator class using suitable conversion functions

Preliminary findings

- Integration process does not affect individual databases
- Local people can work independently to satisfy local users
- Component frame means that local administration maintains the control over their systems and yet provides access to their data by global users
- Three level concept is possible in client/server environment

Conclusions

- We discuss various approaches for data access using federated schema (i.e. federated data model)
- Three-level federation architecture is promising, but detail investigation is recommended.
- Resolving conflicts through a mediation class require more research using real case data in land administration.
- Emerging concept in spatial data infrastructure

FIG Working Week and GSDI-8, From Pharaohs to GeoInformatics, Intercontinental Samiramis, Cairo, Egypt, 16 - 21 April 2005

Thank you for your attention