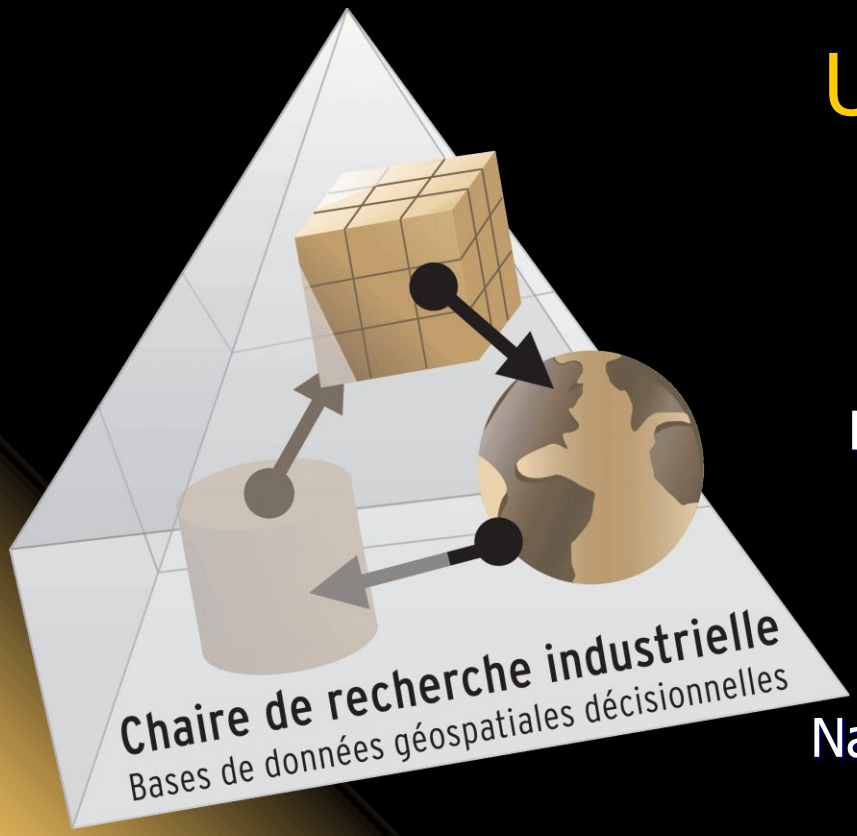
Integrating GIS and BI: a Powerful Way to Unlock Geospatial Data for Decision-Making



Professor Yvan Bedard, PhD, P.Eng.
Centre for Research in Geomatics
Laval Univ., Quebec, Canada

National Technical University of Athens
Mai 27th, 2011

UNIVERSITÉ LAVAL

Origins

- Organisations worldwide invest hundreds of millions of dollars annually to acquire large amounts of data about the land, its resources and uses
- These data however prove difficult to use by managers who need:
  - aggregated information - trends analysis
  - spatial comparisons - space-time correlations
  - fast synthesis over time - unexpected queries
  - interactive exploration - crosstab analysis
  - geogr. knowledge discovery - hypothesis dev.
  - etc.

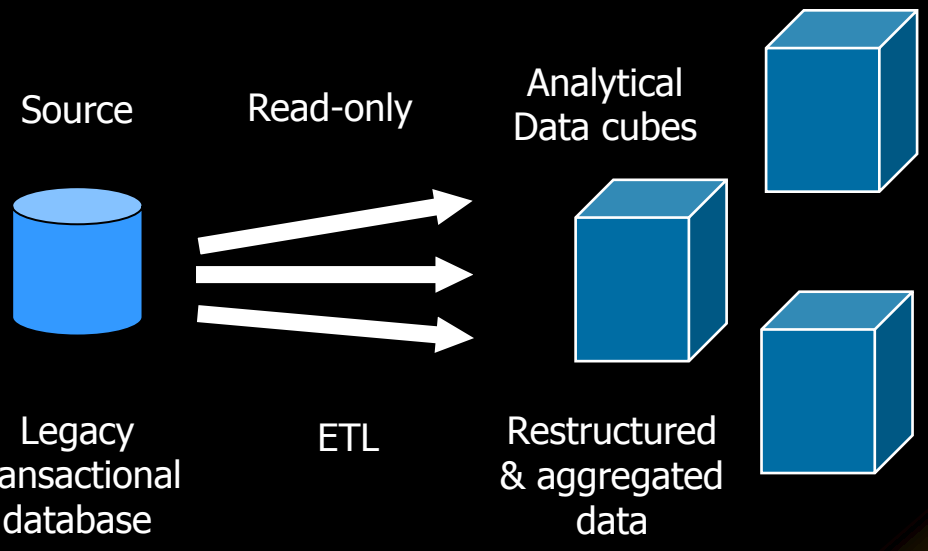
Barriers to make analysis with transactional systems

- GIS and DBMS design are transactional by nature
  - Oriented towards data acquisition, storing, updating, integrity checking, simple querying
- Transactional databases are usually normalized so duplication of data is kept to a minimum :
  - To preserve data integrity and simplify data update
- A strong normalization makes the analysis of data more complex :
  - High number of tables, therefore high number of joins between tables (less efficient).
  - Long processing time
  - Development of complex queries

Analytical approach vs transactional approach

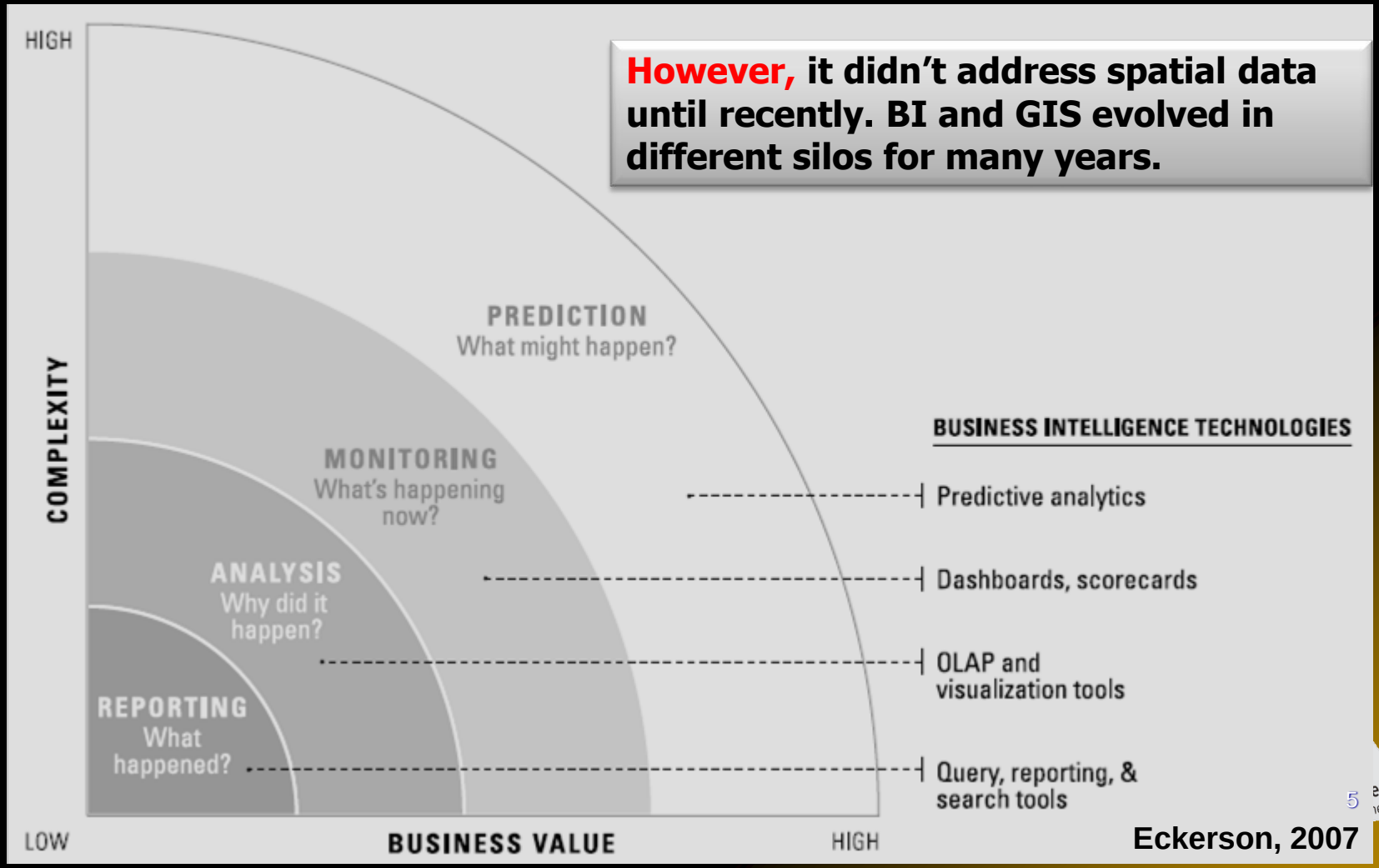
No unique data structure is good for BOTH managing transactions and supporting complex queries. Therefore, two categories of databases must co-exist: transactional and analytical (E.F. Codd).

Example of co-existence: one source -> several datacubes



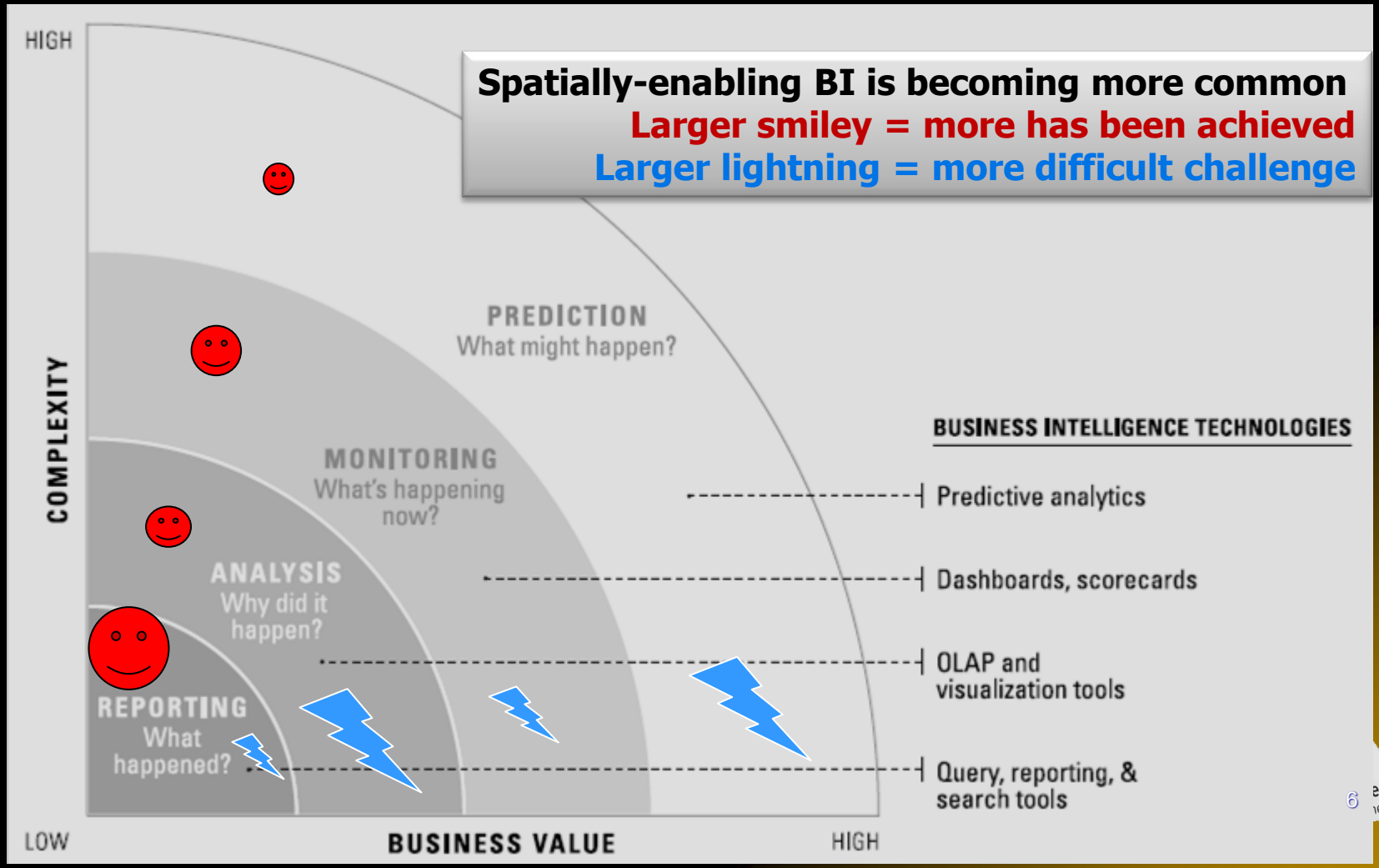
BI Market

- Business Intelligence exists since the early 1990s and its market is larger that the GIS market.



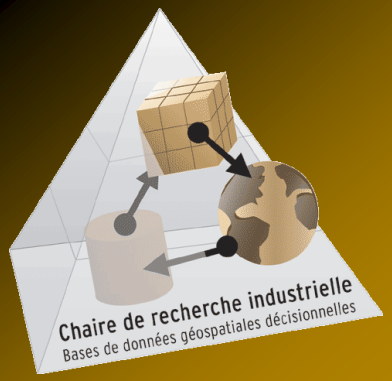
Today's Level of Integration

- Integrating GIS and BI is a recent field with a lot of potential



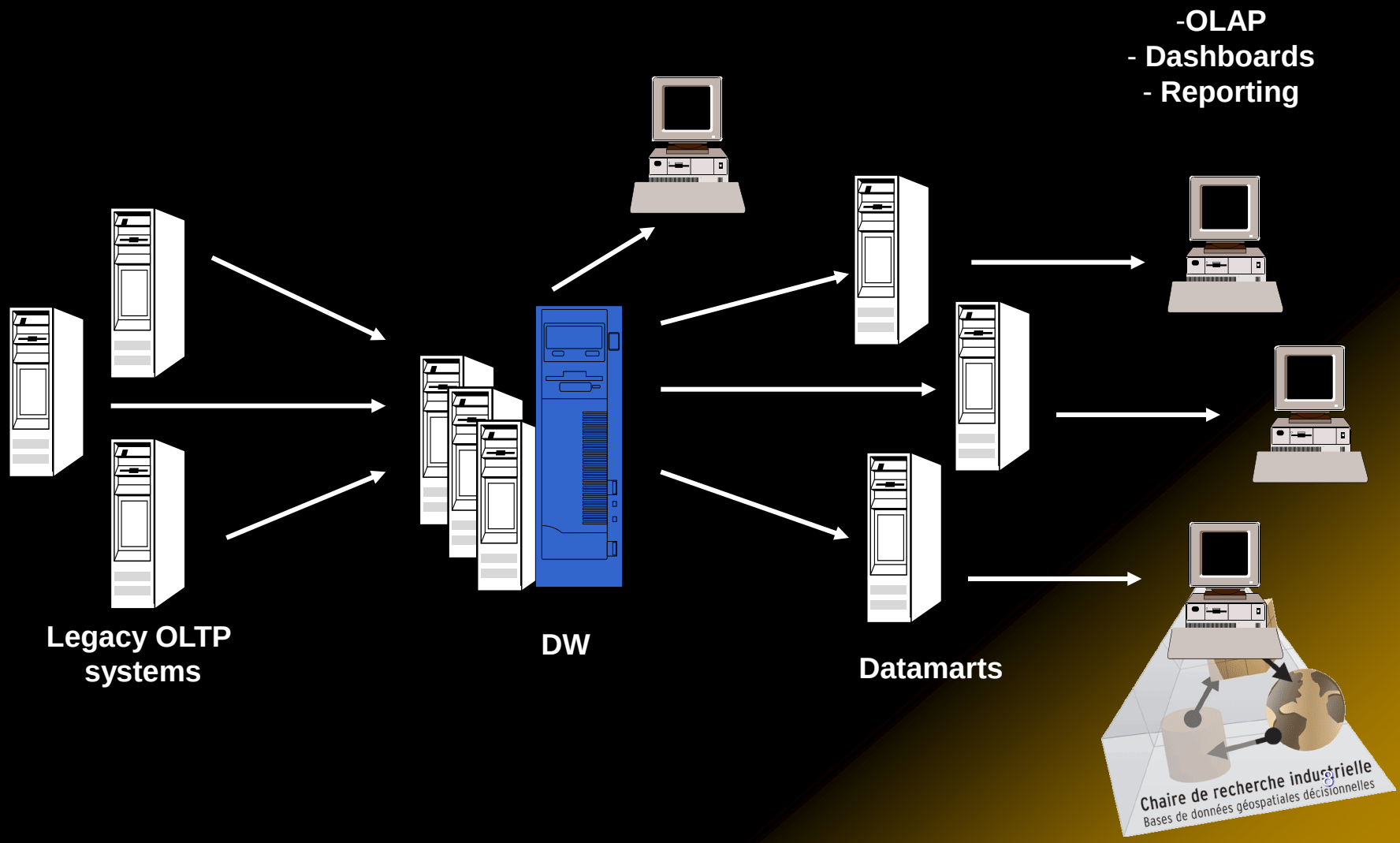
Historical Epochs

- 1996-2000: pionneering
  - early prototypes in universities
    - Laval U. - Simon Fraser U. - U. Minnesota
- 2001-2004: early adopters
  - advanced prototypes in universities
  - first applications in industry
- 2005-... : maturing
  - larger number of ad hoc applications
  - SOLAP technologies to facilitate the development of SOLAP applications
- 2010-...: wide adoption
  - 0ver 30 commercial products



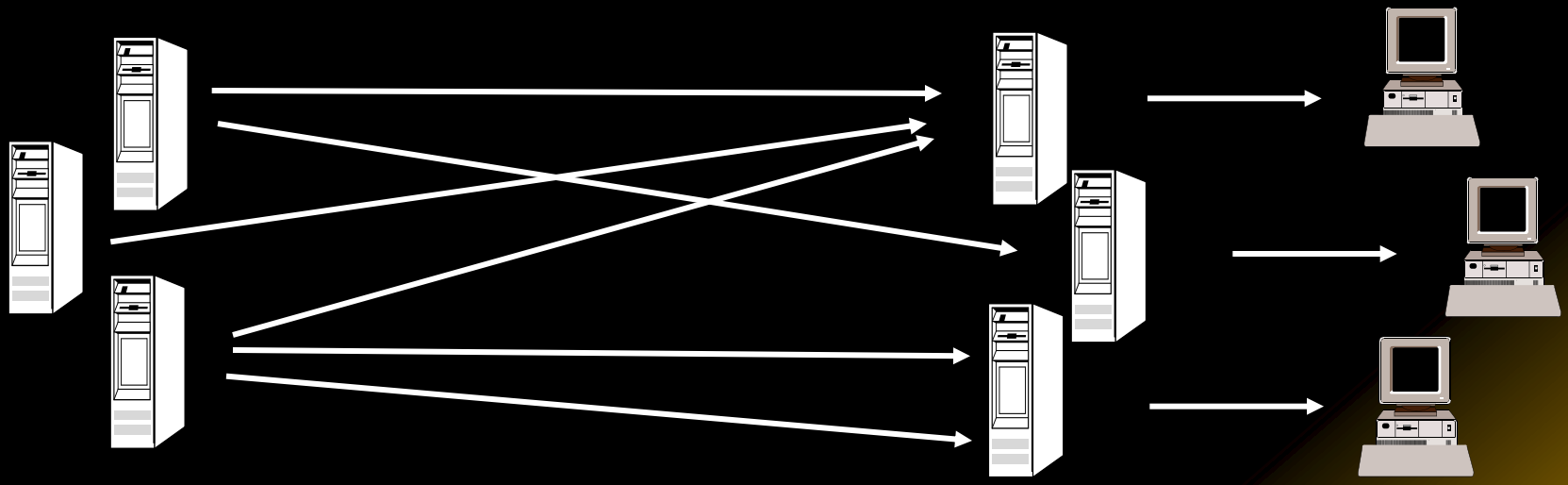
Analytical System Architectures

(ex. standard data warehouse)



Analytical System Architectures

(ex. without data warehouse)

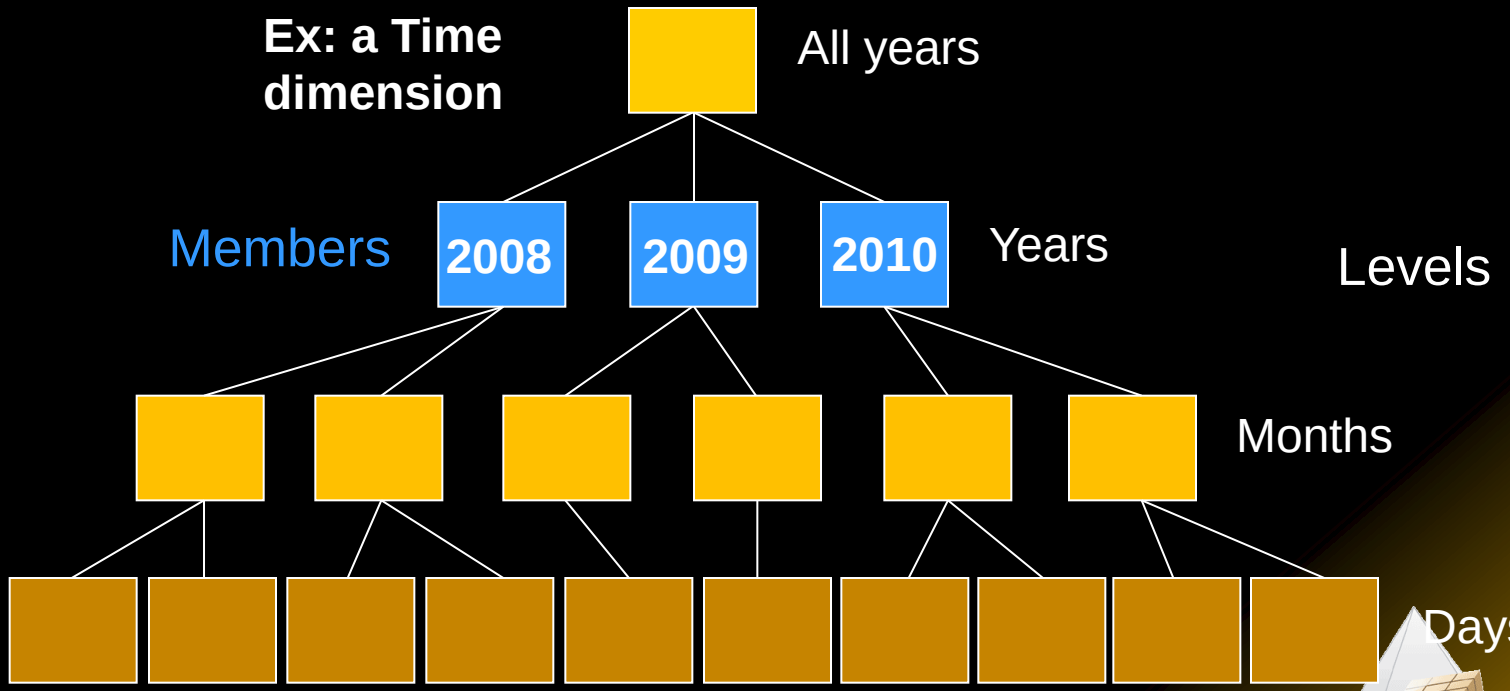


Legacy transactional databases

Datacubes

Datacube Concepts

Dimension = axis of analysis organized hierarchically

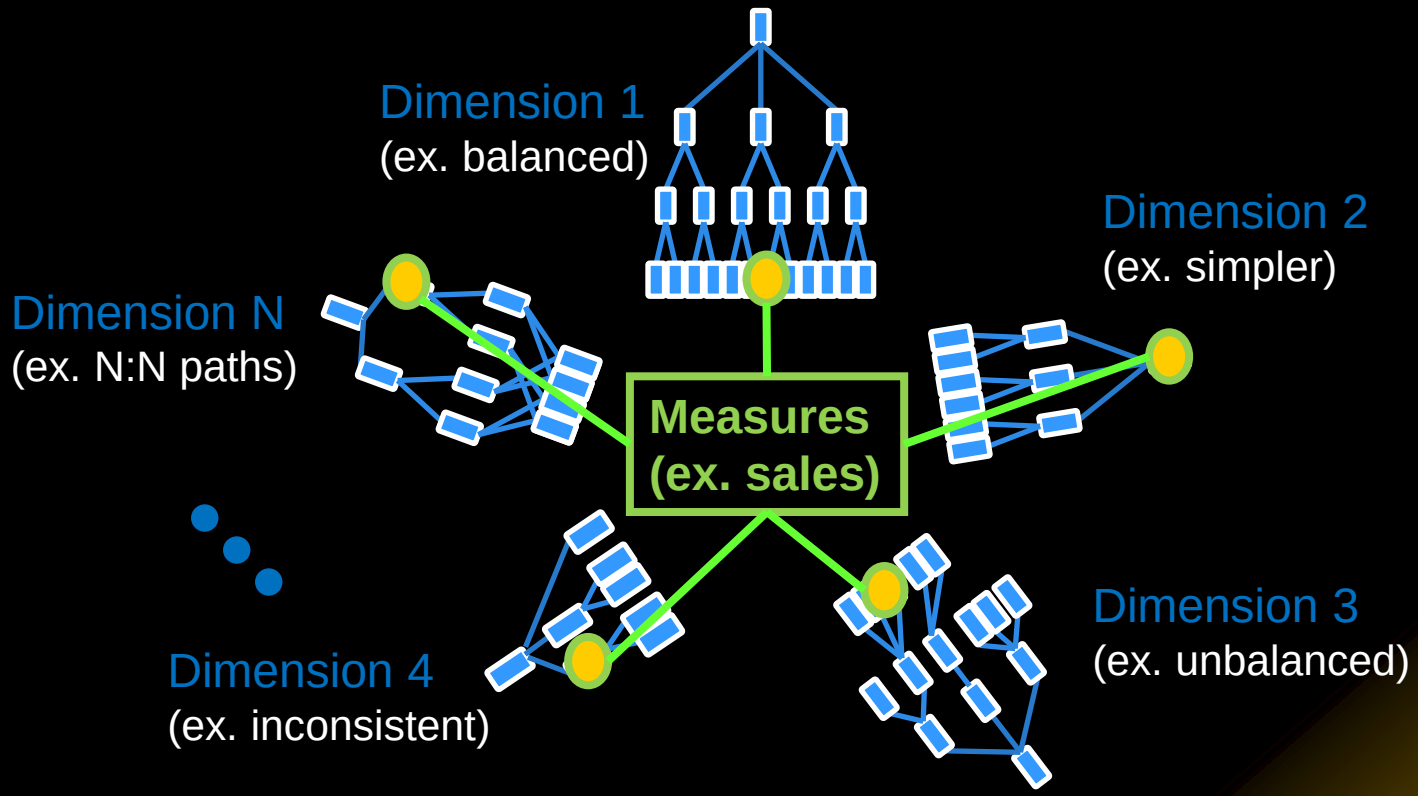


Hypercube = N dimensions

Cube = casual name = hypercube

Chaire de recherche industrielle
Bases de données géospatiales décisionnelles

Datacube Concepts



Members = independant variables

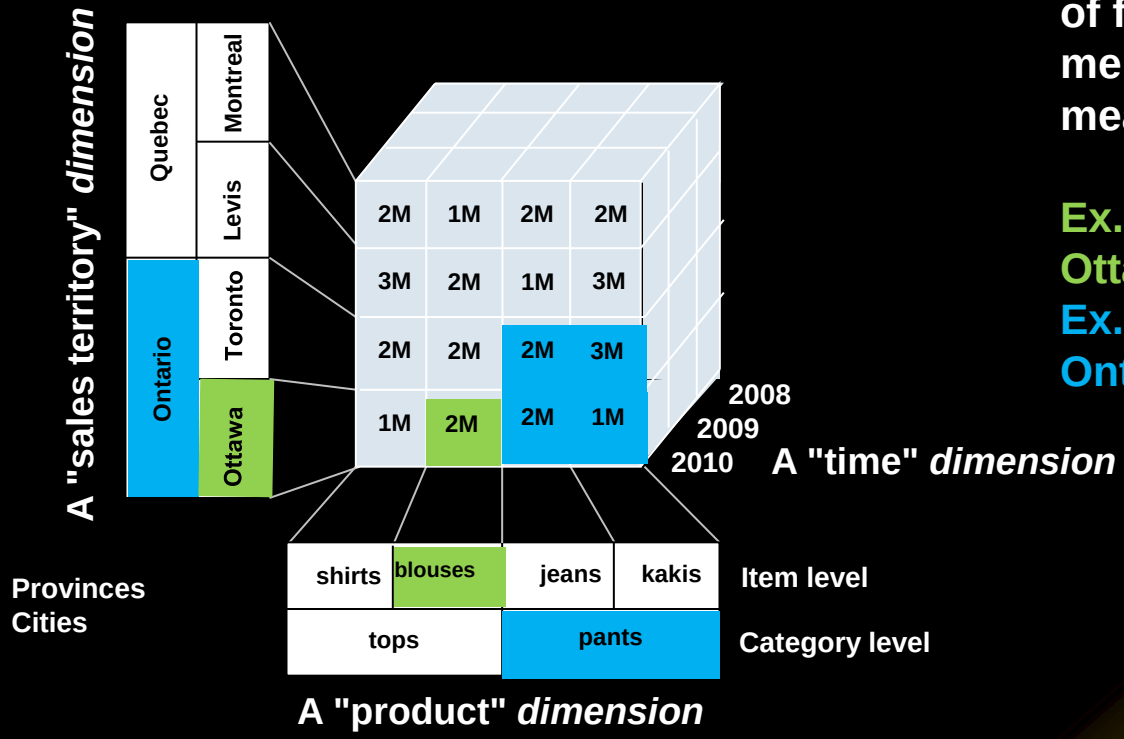
Measures = dependant variables

Chaire de recherche industrielle
Bases de données géospatiales décisionnelles

Datacube Concepts

Cube (hypercube) = all facts

A "sales" *data cube*



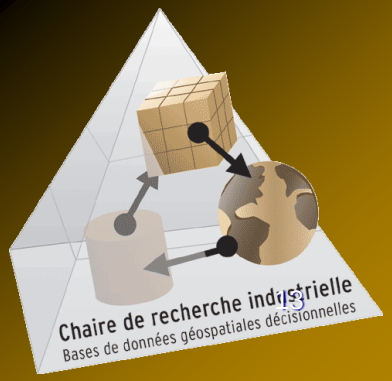
Fact: **each unique combination of fine-grained or aggregated members and of their resulting measures**

Ex.: sold for 2M$ of blouses in Ottawa in 2010

Ex. : sold for 8M$ of pants in Ontario in 2010

Datacube Concepts

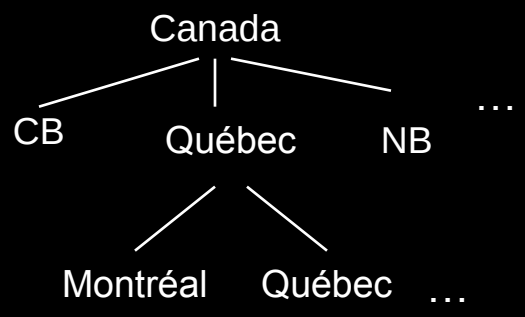
- Data structures (MOLAP, ROLAP, HOLAP):
  - Multidimensional (proprietary)
  - Relational implementation of datacubes
    - Client tool provides the multidimensional view
      - Star schemas, snowflake schemas, constellation schemas
  - Hybrid solutions
- Query languages:
  - SQL = standard for transactional database
  - MDX = standard for datacubes



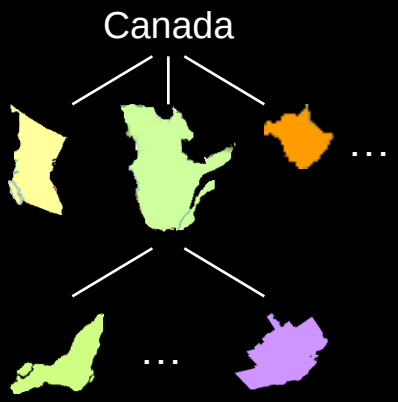
Spatial Datacube Concepts

Spatial dimensions

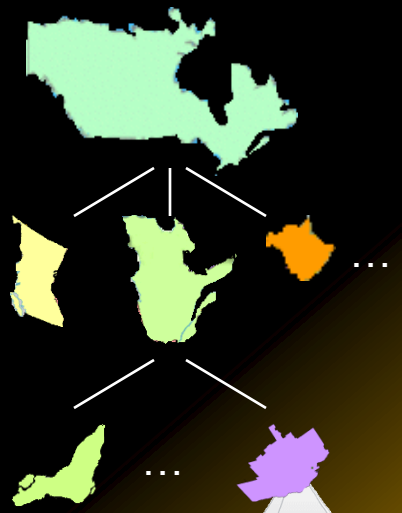
Non-geometric spatial dimension



Mixed spatial dimension



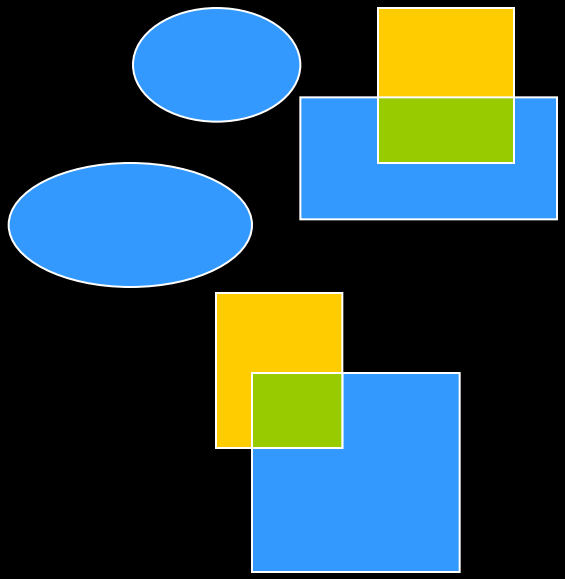
Geometric spatial dimension



N.B. more concepts exist

Spatial Datacube Concepts

Spatial measures



Spatial dimension 1

Spatial dimension 2

Metric operators

- Distance
- Area
- Perimeter
- …

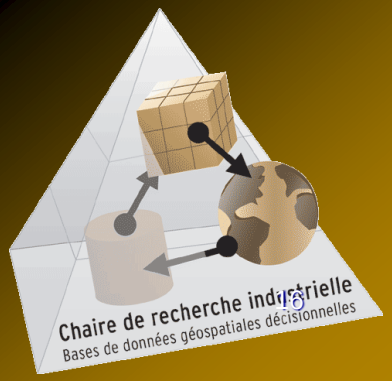
Topological operators

- Adjacent
- Within
- Intersect
- …

N.B. more concepts exist

Spatial Datacube and SOLAP

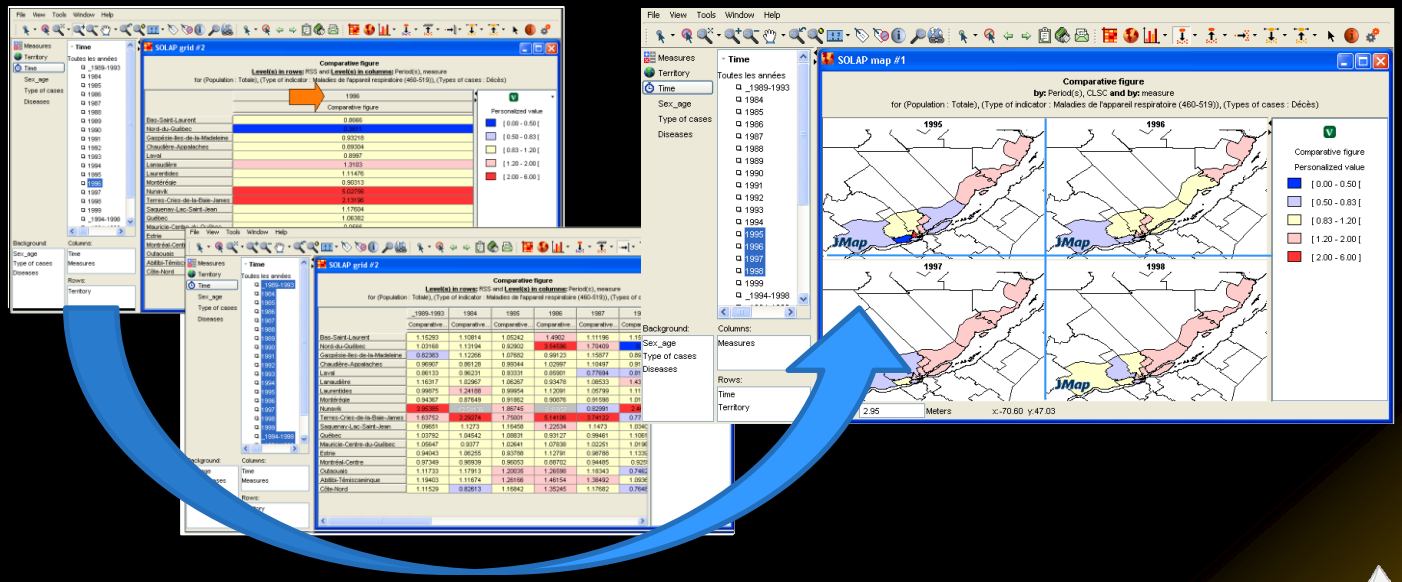
- Spatial OLAP (On-Line Analytical Processing)
- SOLAP is the most widely used tool to harness the power of spatial datacubes
  - It provides operators that don't exist in GIS
- SOLAP = generic software supporting rapid and easy navigation within spatial datacubes for the interactive exploration of spatio-temporal data having many levels of information granularity, themes, epochs and display modes which are synchronized or not: maps, tables and diagrams



Characteristics of SOLAP

- Provides a high level of interactivity
  - response times < 10 seconds independently of
    - the level of data aggregation
    - today's vs historic or future data
    - measured vs simulated data
- Ease-of-use and intuitiveness
  - requires no SQL-type query language
  - no need to know the underlying data structure
- Supports intuitive, interactive and synchronized exploration of spatio-temporal data for different levels of granularity in maps, tables and charts that are synchronized at will

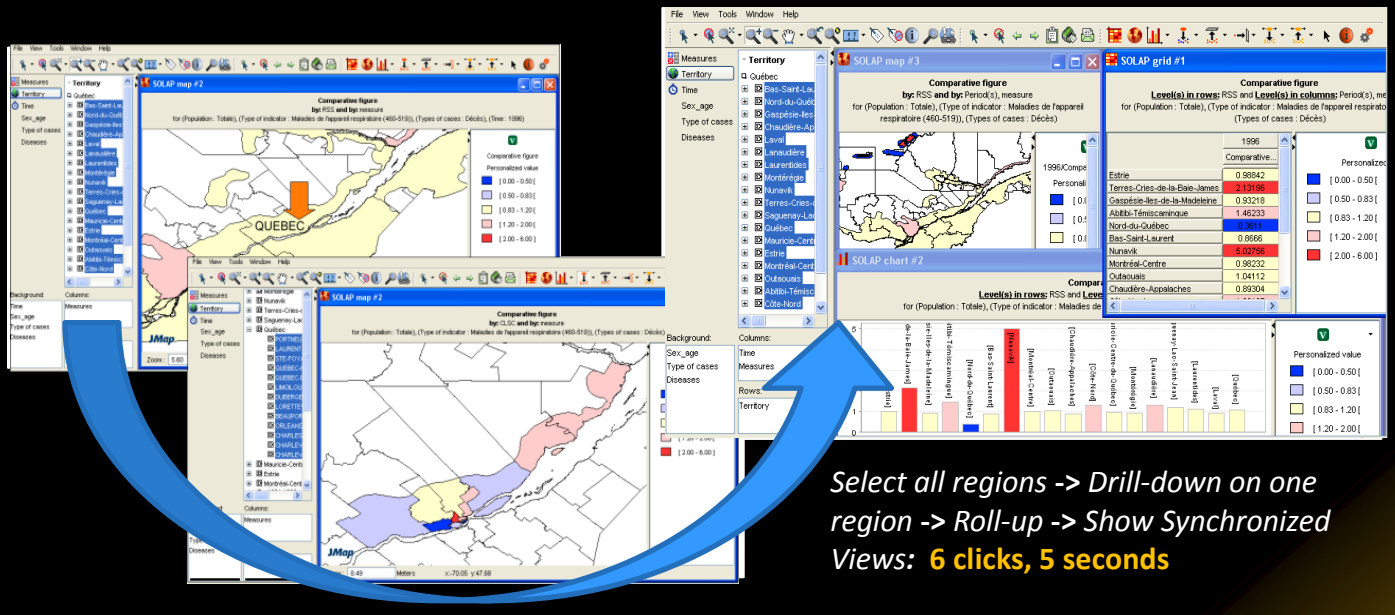
The Power of SOLAP Lies on its Capability to Support Fast and Easy Interactive Exploration of Spatial Data



Select 1 year -> Select all years ->
Select 4 years -> Multimap View:
7 clicks, 5 seconds

Chaire de recherche industrielle
Bases de données géospatiales décisionnelles

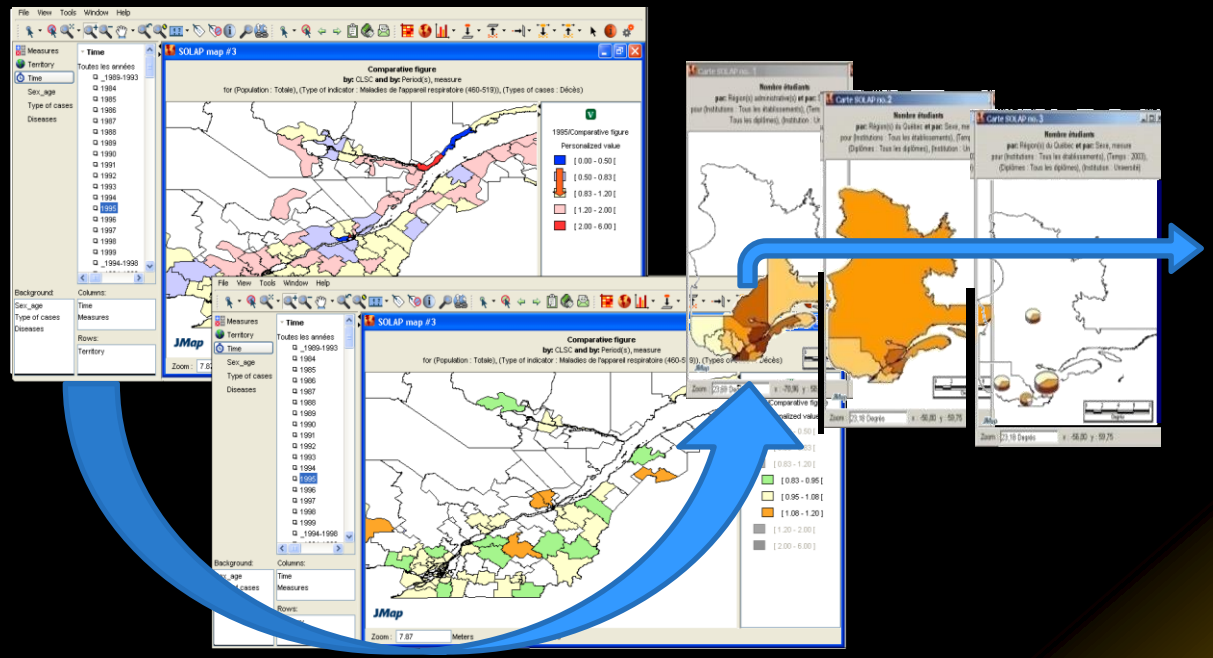
The Power of SOLAP Lies on its Capability to Support Fast and Easy Interactive Exploration of Spatial Data



Select all regions -> Drill-down on one region -> Roll-up -> Show Synchronized Views: **6 clicks, 5 seconds**

Chaire de recherche industrielle
Bases de données géospatiales décisionnelles

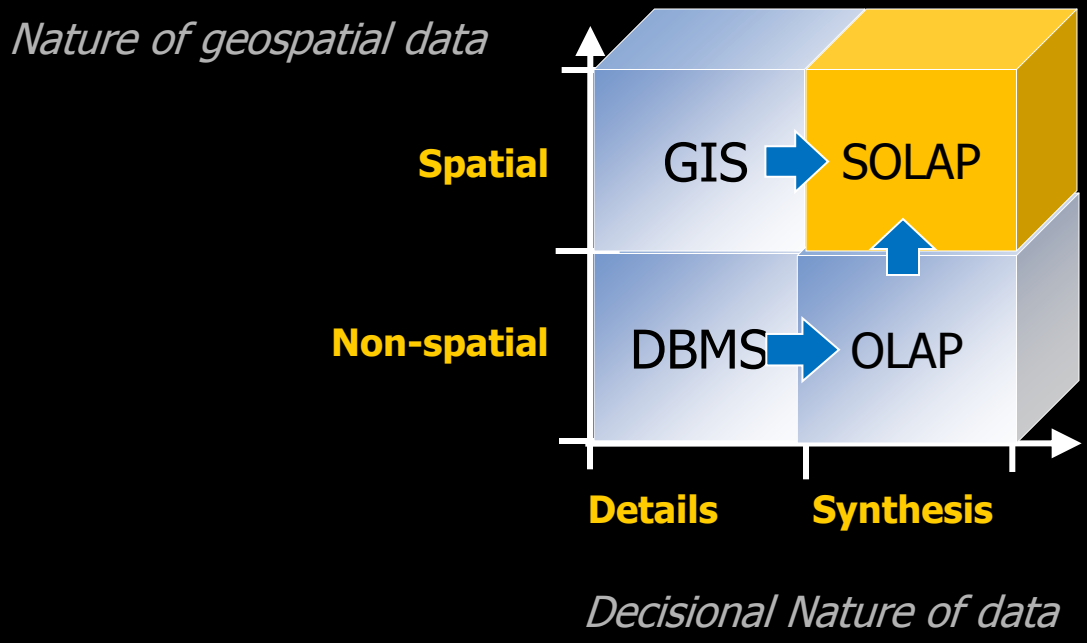
The Power of SOLAP Lies on its Capability to Support Fast and Easy Interactive Exploration of Spatial Data



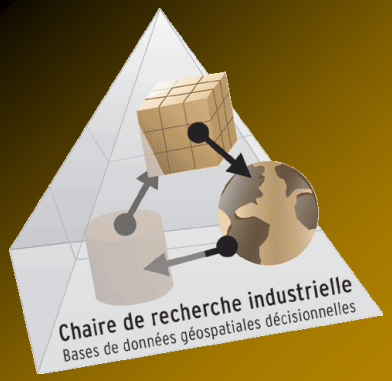
Change data -> Roll-up -> Roll-up -> Pivot ... : **6 click, 5 seconds**

Chaire de recherche industrielle
Bases de données géospatiales décisionnelles

A Natural Evolution

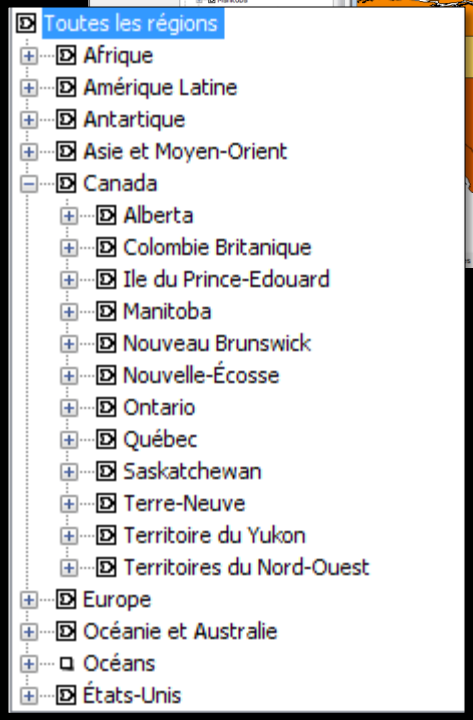


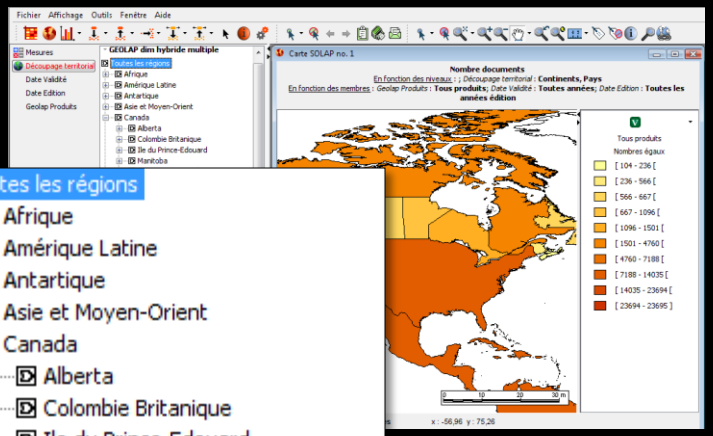
- Add capabilities to existing systems, don't aim at replacing them
- Add value to existing data, no attempt to manage these data

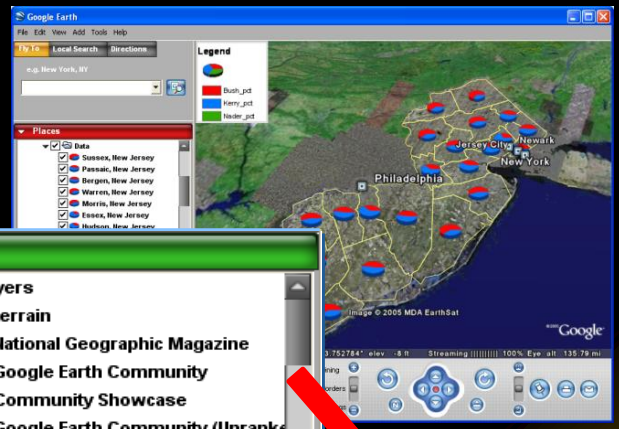


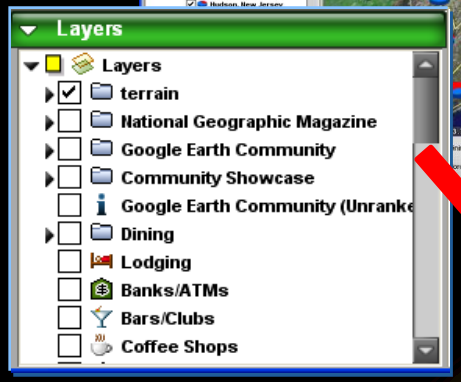
Functionalities: *Spatial datacube structure*

- ✓ Contains spatial dimension organized in a hierarchy of spatial members with their geometry









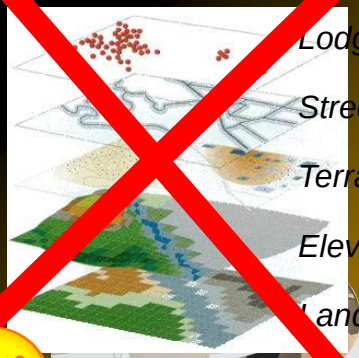
World

Countries

Provinces

Abstraction layers

Map layers



Lodging

Streets

Terrain

Elevation

Land use

Chaire de recherche industrielle
Bases de données géospatiales décisionnelles

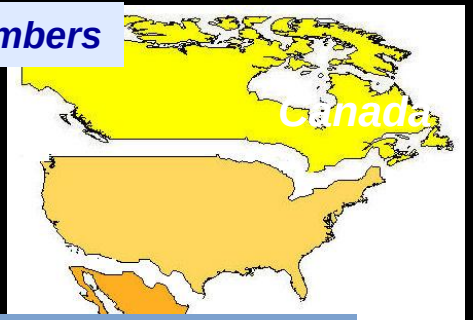
Functionalities: *Interactive Exploration*

supporting the true spirit of data drilling

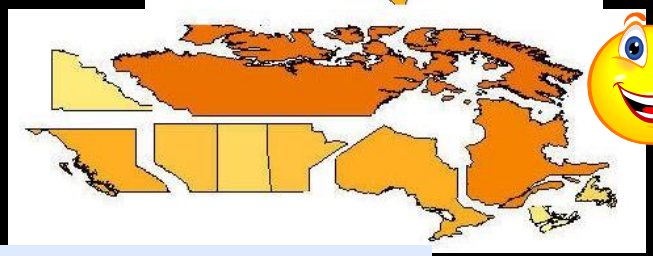
- ✓ Drilling is executed on a member or on a selection of several members (at the same or at different levels of detail of a dimension).

Drilling by members
at different level of details

Country members



Spatial drill on Canada member

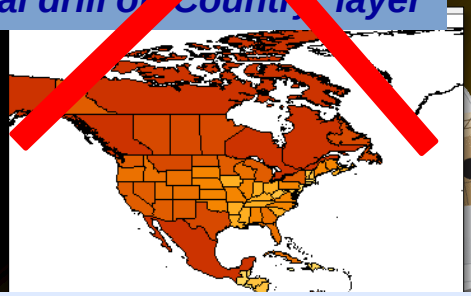


Result in province members of Canada only

This is not OLAP drilling !

Country layer

Spatial drill on Country layer



Result in opening the entire Province/states layer

Functionalities: *Interactive Exploration*

Pivot on map

- Changes the orientation of the dimensions to produce a new display.
- Applied to a map = a different type of map
- Built-in rules must exist to produce the pivoting map corresponding to dimensions selection instantaneously without SQL

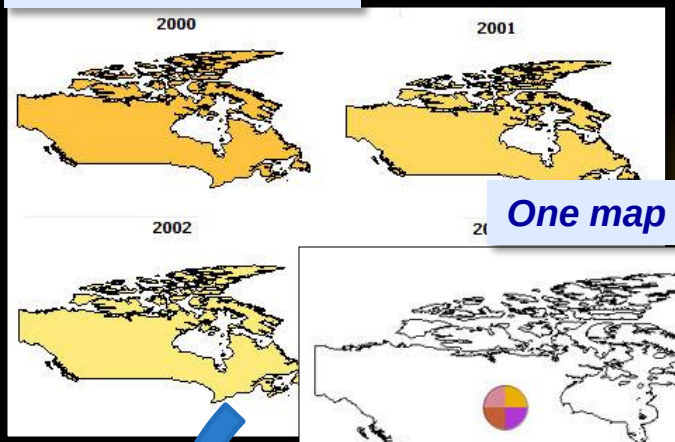
One row per year

	Canada
2000	
2001	
2002	
2003	

One row for all years

	2000	2001	2002	2003
Canada				

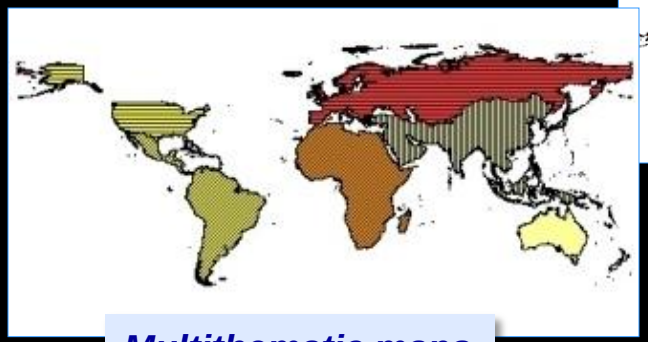
One map per year



One map for all years

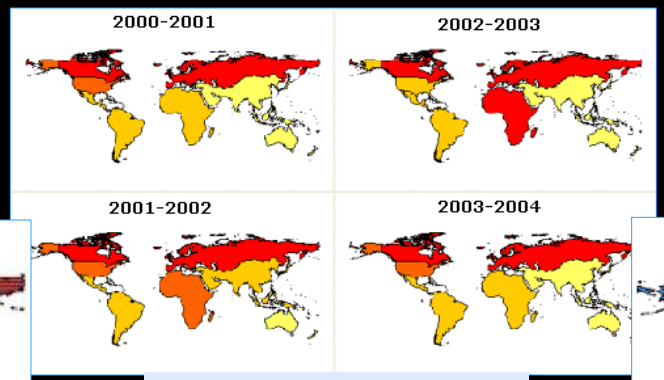
Functionalities: *Exploration-Oriented Visualization and a*dvanced maps

- ✓ Must support various types of maps (not only choropleth map)
- ✓ Advanced maps are used to represent many dimensions or many measures on a map.
- ✓ Built-in rules must exist to produce advanced maps instantaneously without SQL.



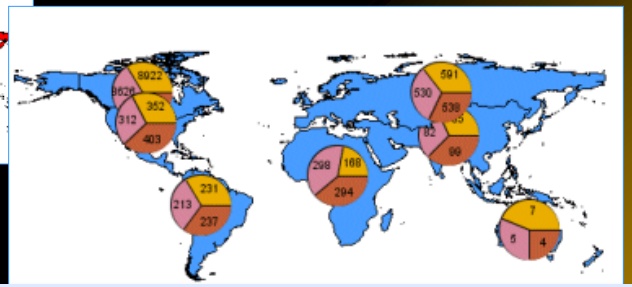
Multithematic maps

(one measure/color + one measure/pattern)



Temporal multimap

(one map per year)

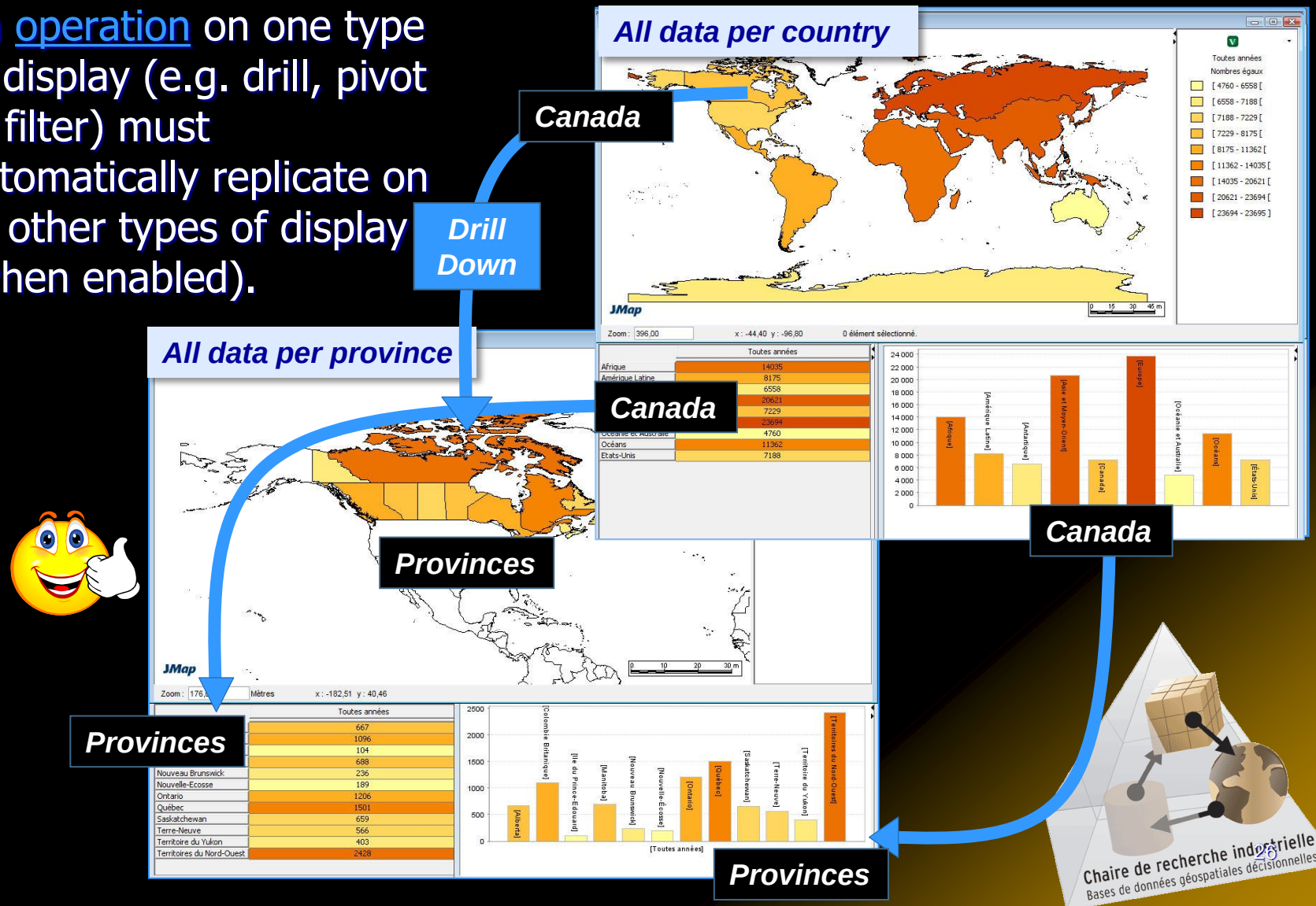


Map with superimposed diagrams

(many members in pie charts)

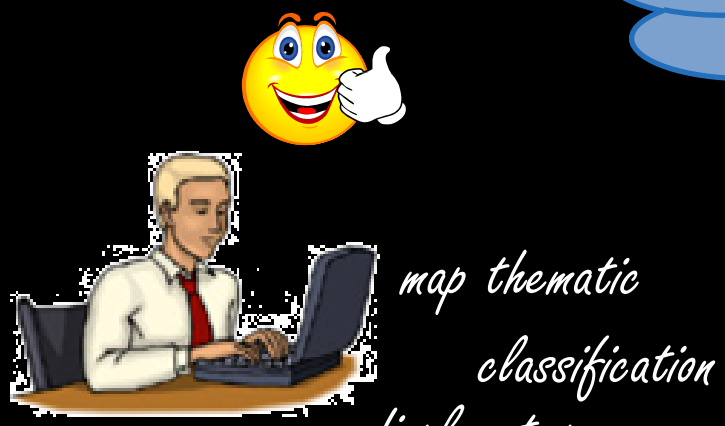
Functionalities: *Exploration-oriented Visualization and synchronized displays*

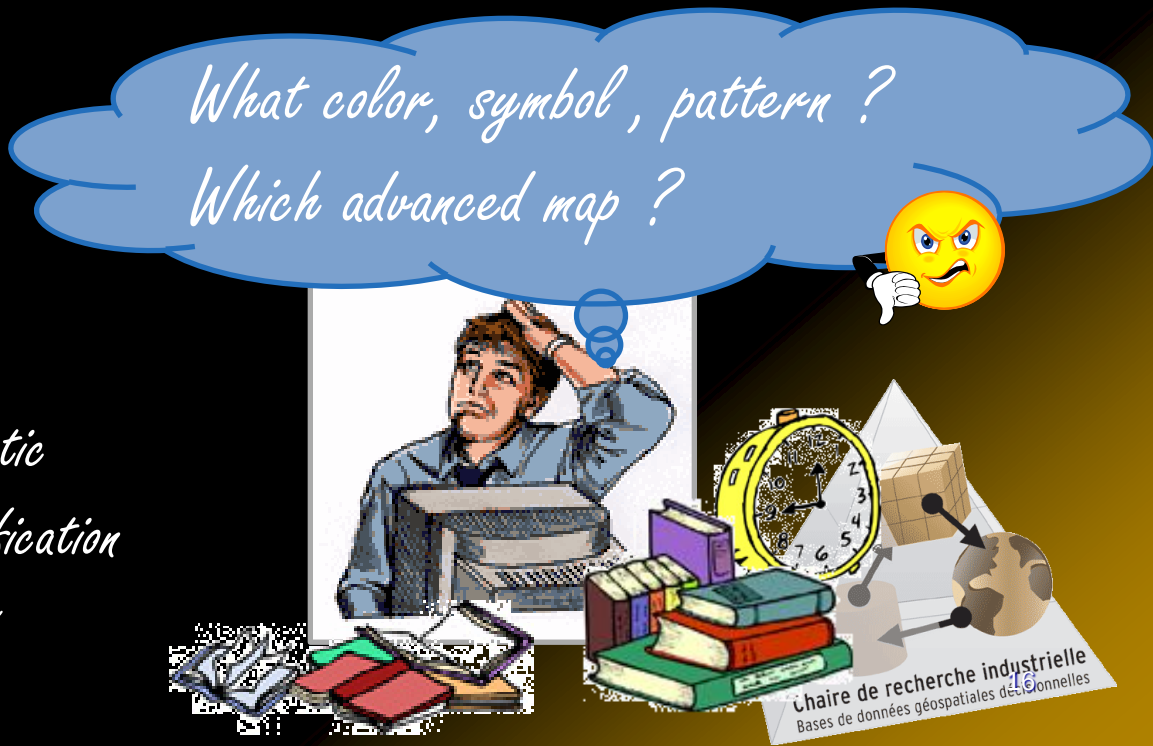
- An operation on one type of display (e.g. drill, pivot or filter) must automatically replicate on all other types of display (when enabled).



Functionalities: *Exploration-oriented Visualization and intelligent automatic mapping*

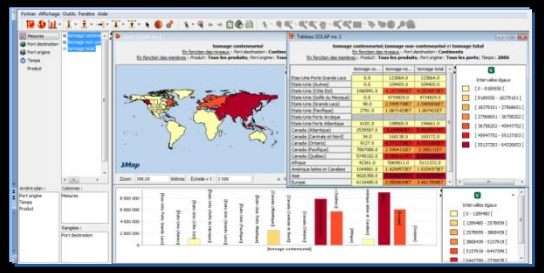
- ✓ Intelligent automatic mapping:
  - ✓ Supports user's knowledge
  - ✓ Generates coherent maps by using predefined display rules in accordance to the user's selection
  - ✓ Instantaneous display
  - ✓ No SQL involved
- ✓ Manual processing:
  - ✓ Involve specific knowledge by the user (database, semiology, mapping)
  - ✓ Is time-consuming





Example of Measured Benefits in a Project for Transport Quebec

M.J.Proulx, Intelli3 (2009)



Annual Report :

150 maps and tables
Static data

Analysis & page editing (3 months-person)
Updating (1 month-person)

Ad hoc queries continuously
Delays to produce outputs

Depend upon an expert in cartography

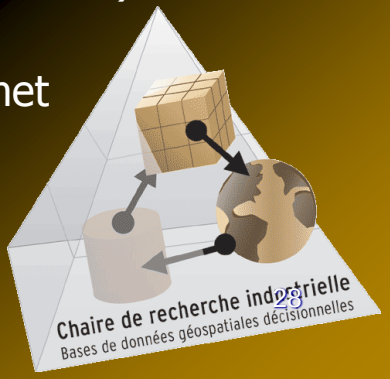
Solution géodécisionnelle :

200 000 maps and tables
Dynamic applications

Data structuration (15 days-person)
Updating (5 dayx-person)

Application in intranet
Fast response

Easy user interface

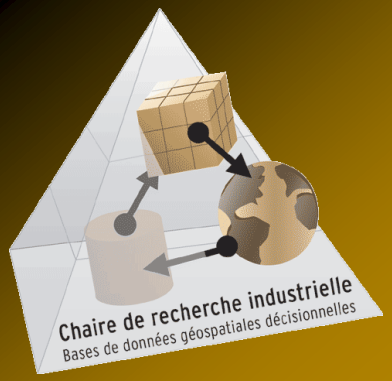


Approaches to Develop SOLAP Applications

- *Ad hoc*, proprietary programming specific to one application
- Combining GIS + OLAP capabilities
  - GIS-centric
  - OLAP-centric

  } *-The dominant tool offers its full capabilities but gets minimal capabilities from the other tool*
  -GUI provided by the dominant tool

  - Integrated SOLAP
    - Ad hoc programming (ex. using diverse open-source softwares)
    - SOLAP technology (the most efficient)



Off-the-Shelf Integrated SOLAP

Facilitates the deployment of a SOLAP application by offering built-in elements (e.g. Framework, operators, unique GUI)

GIS or OLAP Dominant

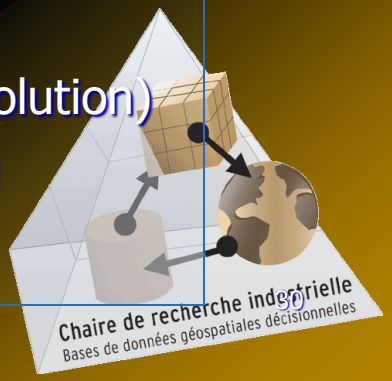
Total-integration capable but requiring programming

Total integration (COTS)

Loosely coupled

Strongly coupled

- ✓ 2 GUI *vs* common and unique GUI
- ✓ Built-in integration framework (no need to program the solution)
- ✓ Offers built-in functionalities to visualize and explore data
- ✓ No dominant component



Video of an Example of SOLAP

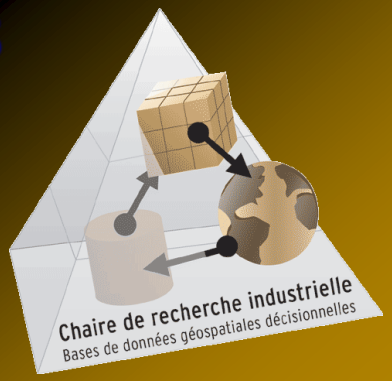
- Map4Decision (www.intelli3.com)
- 100% Java
- Reads MOLAP and ROLAP datacubes
- Reads all popular GIS files
- The 1st SOLAP on the market
- More than 30 SOLAP-like products exist
  - Map4Decision still leads with regards to
    - Variety of spatial datacubes
    - Multi-platform support
    - Integration of BI and GIS concepts
    - Ease of installation

VIDEOS


Chaire de recherche industrielle
Bases de données géospatiales décisionnelles

Conclusion

- GIS and BI have evolved in silos for many years
- R&D bridging both universes started mid-90s
- Market is reaching maturity
- A scientific community exists
- Different application development approaches
  - ad hoc
  - BI-centric
  - GIS-centric
  - Integrated
- More R&D will bring even better solutions



Thank you !

More info at these web sites:

http://sirs.scg.ulaval.ca/yvanbedard/

http://www.spatialbi.com/

http://mdspatialdb.chair.scg.ulaval.ca/english/Eindex.asp

Technology transfer =
Map4Decision (www.intelli3.com)