Data Jigsaw: Do We Understand It?

Mirela Danubianu – Stefan cel Mare
University of Suceava , Romania
Petre Dini- Concordia University, Canada/
China Space Agency, China

ICCGI 2012 , June 24-29, Venice Italy

Data, data… data is everywhere!

- the explosive use of ICT in all areas of life has led to the collection of huge data volumes

- we could say that this also led to an increase of the available information as decision support
  - **Wrong!!!**
    - **Why?**
      - different organizations collect data for their own informational necessities
      - for this reason:
        - data has a certain model imposed by the owner
          - data cover a limited area of observations
          - data is placed in various locations

Data, data… data is everywhere!

- ## The result -
  - history comes around and we can say that we are in a situation similar to that existing at the beginning of the 90's,even though at an extensive scale where we have huge data volumes where information is still hidden.
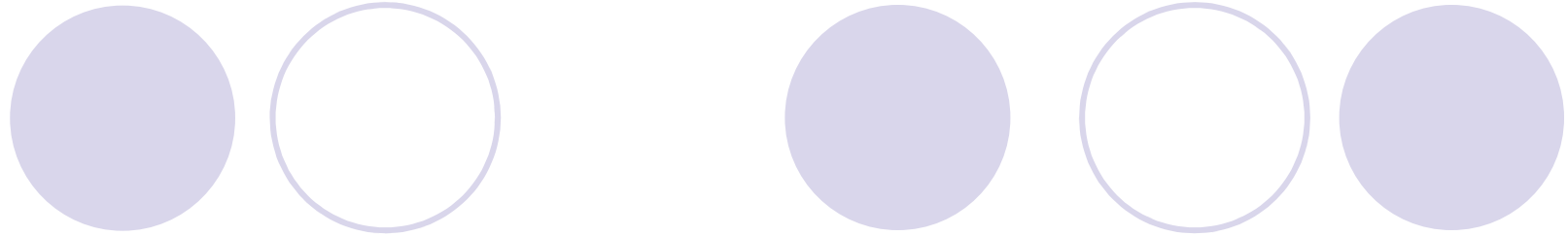
- ## Solution -
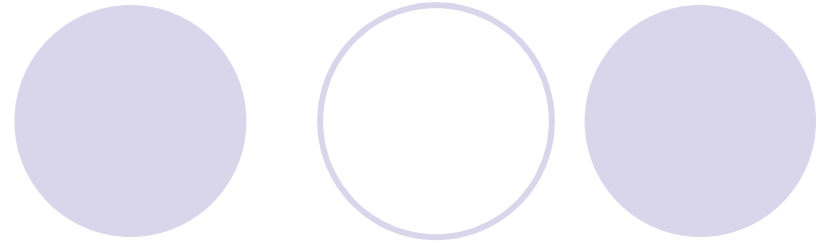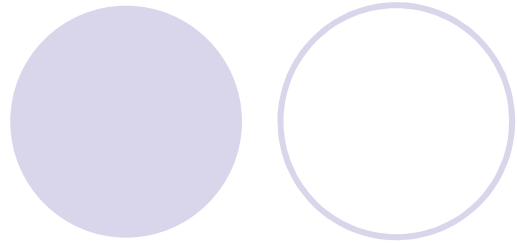 - data integration should match like puzzle pieces, so that they give us the image of a whole universe

… this panel aims to discuss:

- some aspects regarding data Jigsaw
- some ways to make data more efficient
- and why not – some existing solutions which could be used in order to do this

Prof. Petre Dini

- syntactic and semantics of data
- data localization
- volume of data and media storage
- models
- generalized categories yet computable

Mirela Danubianu

- data as a service (DaaS)
- Jigsaw Data

Data as a Service

- Based on the concept that data can be provided on demand
  - regardless of geographic or organizational separation of provider and consumer
- Brings the idea that:
  - data quality can happen in a centralized place
  - data is offered to different systems, application or users wherever they were in the organization or on the network

Data as a Service - Benefits

- Agility
  - customers can move quickly due to the simplicity of the data access
  - they don't need extensive knowledge of the underlying data
  - If they require different data structures the implementation is easy because the changes are minimal
- Cost effectiveness
  - Providers can build the base with data experts and outsource the presentation layer
    - Cost effective user interfaces
- Data quality
  - Access to the data is controlled through data service which tend to improve data quality because there is a single point for updates

Jigsaw Data

- a web directory of business card and company records
  - founded on January 1st, 2004
  - built on crowdsourcing
  - around this, it has developed a community of more than 1 million members who built and maintain the database
    - mainly sales, marketing and human resources professionals