Data Integration: A New Teenage

C. A. Curino, G. Orsi and L. Tanca

DEI ICT Institute
Politecnico di Milano

ADI Workshop (OTM 2008)
Monterrey (Mexico)

November 9th, 2008
State of the Art

The Past: 80s - 90s
- Distributed and Federated databases
- Multidatabase

The Present: 2000-2008
- Mediated Schema - DIS [2]
  - Reverse Engineering
  - Semantic Annotation and publishing
- P2P - Data Exchange [1]
  - Full P2P
  - Publish & Subscribe
- Data Rings and Data Spaces [3, 4]

Limitations
- Data sources known in advance
- Static, hand-made mappings.
The Future

Challenges

- Internet of Things
  - Autonomic data-sources
  - Data and Users Mobility
  - Embedded systems Databases
  - Uncertainty and Lineage Management
- Multimodality
- Ambient and body intelligence
- Information noise

Problems

- Past problems are not completely solved.
- Missing or expensive infrastructures.
- Sound and complete data integration in open world is unpractical.
- 85% of interesting data is unstructured!
## Context-ADDICT

(joint work with C. Bolchini, E. Quintarelli and F. A. Schreiber)

### Features

- Context-aware data/ontology tailoring [5]
- Ontology-driven, on-the-fly data integration of heterogeneous and dynamic data sources (HD-DIS)
- Multimodal access to resources
- Focus on small and mobile devices (sensors, mobile phones, custom embedded-systems)

### Applications

- Automotive
- e-Health
- Logistics
- Energy
- Production Automation
- Automated and Personalized Advertisement
- Personal Information Systems
Context-ADDICT
(Context-aware Data Design, Integration, Customization and Tailoring)
References

R. Fagin and P. G. Kolaitis and R J. Miller and L. Popa
Data exchange: Semantics and query answering.
In Proc. of 9th Intl Conf. on Database Theory (ICDT), 2003

M. Lenzerini
Data integration: a theoretical perspective.

S. Abiteboul and N. Polyzotis
The Data Ring: Community Content Sharing.
Proc. of the 3rd Biennial Conf. on Innovative Data Systems Research (CIDR), 2007

A. Halevy, M. Franklin and D. Maier
Principles of Dataspase Systems.
Proc. of the 25th ACM SIGMOD-SIGACT-SIGART Symp. on Principles of database systems (PODS), 2006

C. Bolchini, C. A. Curino, F. A. Schreiber and L. Tanca
Context Integration for Mobile Data Tailoring.
Proc. of the 7th Intl Conf. on Mobile Data Management (MDM), 2006