

INVESTIGATING DATA PROVENANCE IN THE CONTEXT OF NEW PRODUCT DESIGN AND DEVELOPMENT

Sudha Ram
Eller Professor of MIS
Department of MIS
University of Arizona
Tucson, AZ
Email: ram@eller.arizona.edu
URL: <http://adrg.eller.arizona.edu>

May 17, 2005

What is Provenance?

- Lineage, Pedigree, Origin
- Enables correct interpretation
- Includes:
 - who
 - what
 - where
 - how
 - assumptions

Research Objective

- Investigate the semantics of Data Provenance
 - ❖ Develop an ontology to represent the semantics of Provenance
 - ❖ Develop ways to automate the capture of Provenance

New Product Design

- New Product Design Scenario
- Importance of Provenance
 - ❖ Many types of Data
 - ❖ Many places of origin
 - ❖ Many users
 - ❖ Reusability of data

Previous Work on Provenance

- Investigated in many domains: e.g. Bioinformatics, Physics, Astronomy
- Focus on “Why” and “Where”
- Closely related to Data Quality, Reliability and Metadata
- Gap: Comprehensive definition of Provenance

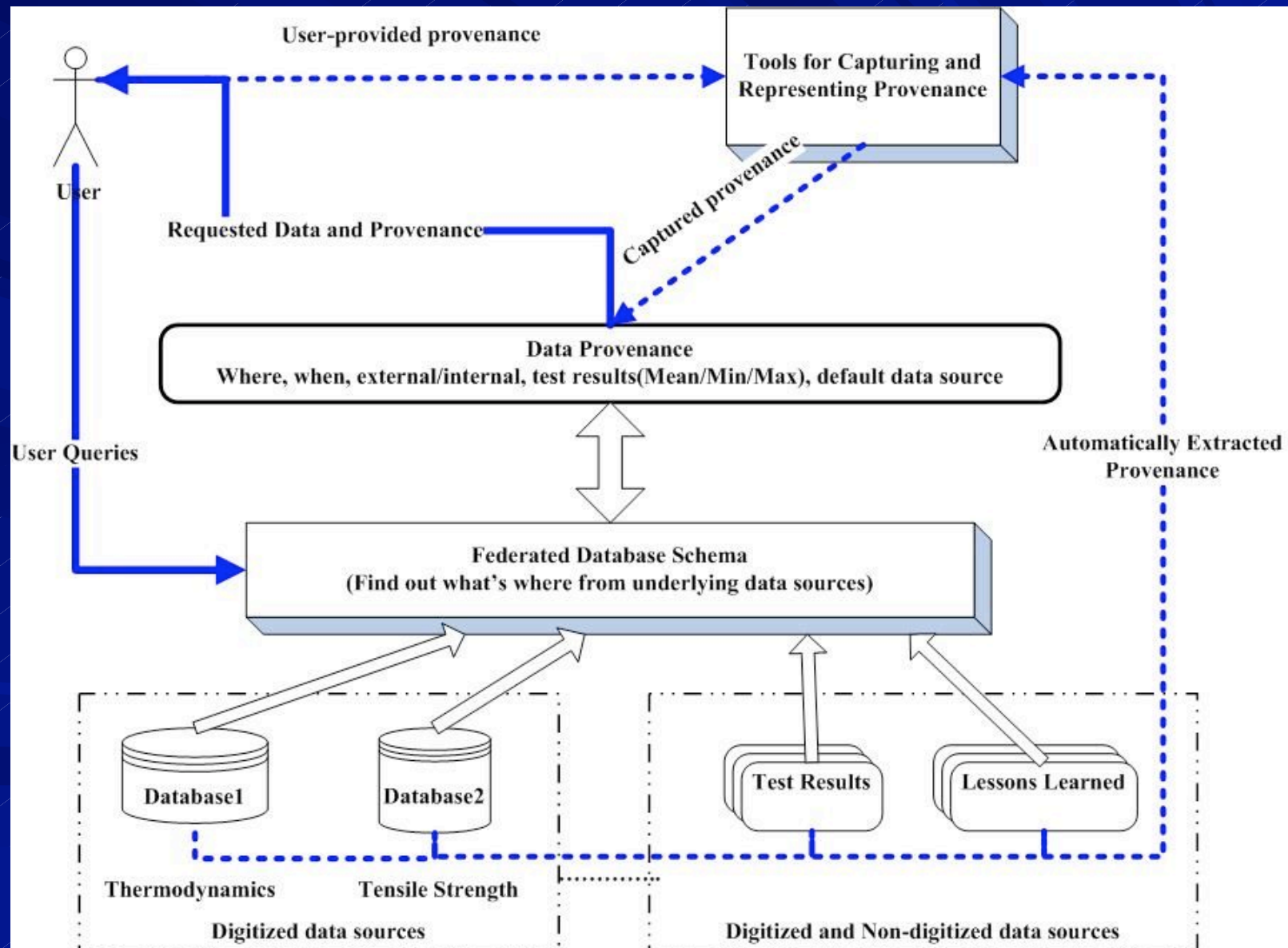
Specific Research Questions

- Understanding Semantics of Provenance
- Representing and Capturing Provenance
- Implementing and Using Provenance

Methodology

- Project Partner: Raytheon Missile Systems
- New Product Design Teams
- Understand Provenance
- Develop a prototype system: PROMS

Architecture of PROMS



Project Plan

- Understand the New Product Design Process, Define Metrics
- Develop Provenance Ontology
- Implement PROMS
- Evaluate using metrics
- Time Frame: 12 months

Conclusion

- Understand how to represent Provenance
- Autonomous harvesting of Provenance
- Extend to other domains
- Impact on long term digital archiving and preservation

QUESTIONS?

