

Provenance

What a Concept!

Report on the iTRUST
Interdisciplinary Workshop on
Provenance
May 14, 2015



Team Members

- Lucie Burgess, Bodleian Library, Oxford University
- Ken Cavalier, UBC iTRUST
- Adrian Cunningham, Queensland State Archives, AUS
- David Dubin, Research Data Alliance
- Luciana Duranti, UBC iTRUST
- Larry Lannom , CNRI/Research Data Alliance
- Victoria Lemieux, UBC iTRUST/World Bank
- Bertram Ludaescher, University of Illinois
- Giovanni Michetti, iTRUST, University of Rome – Sapienza, ITALY
- Paolo Missier, University of Newcastle
- Corinne Rogers, UBC iTRUST
- Joe Tennis, University of Washington
- Ken Thibodeau, iTRUST
- Sharon Thibodeau, Independent
- Margaret Varga, University of Oxford/NATO
- Ashley Wheat, Middlesex University
- Kai Xu, Middlesex University



Motivation

How do we capture and represent provenance in increasingly distributed and networked information processing environments?



Disciplines/Perspectives Represented

- Archival Science
- Computer Science
 - Information Management
 - Data Management
 - Human-Computer Interaction
- Law
- Library Science
 - Classification Theory
- Visual Analytics



Why Does Provenance Matter?

- Establishing Trust
- Semantic interpretability
- Assessing value
- Transparency and accountability
- Reproduceability
- Attribution and protection of rights
- Sense-making
- Digital Preservation
- Digital Records Preservation
- Archival Arrangement & Description
- E-Science (e.g., Research Data Alliance)
- Digital Humanities
- Open Data
- Decision-Support
- Intelligence Analysis



Perspectives on Provenance

A dynamic and constantly reformulating concept even within disciplines and cognate areas

- Provenance as creator (AS)
- Provenance as original order (AS)
- Provenance as chain of custody/preservation (AS/LA)
- Provenance as relationships (AS)
- Provenance as graphs (AS)
- Provenance as context (AS)
- **Parallel provenance (AS)**
- **Provenance as process (W3C PROV standard; IM)**
- **Analytic Provenance (VA)**
- **Reasoning Provenance (VA)**
- Agent-centred provenance (IM)
- Object-oriented provenance (IM)
- What, where, why provenance (DM)
- **Differential provenance (*NEW)**



Research Challenges

Conceptual . . .

A deeper and agreed understanding of provenance is needed in order to cope with new forms of documentation and new modes of communicating and processing information.

-Michetti



Research Challenges

Interoperability . . .

W3C PROV Standard

&

Archival Arrangement and Descriptive Standards

InterPARES Chain of Preservation Model

Etc.



Research Challenges

Technical. . .

Capture

Storage

Security

Analytics/Exploitation



Next Steps

- Development of a synthesis paper and publication of the workshop proceedings (including an interdisciplinary bibliography on Provenance)
- Collaboration on iTRUST Project Archival Arrangement & Description in the Cloud (Michetti)
- Preparation of a new research SSHRC Insight Grant proposal and exploring other opportunities with other funding bodies e.g., NSF, JISC
- Monthly conference call and September research meeting (to be hosted by RDA)



Old Wine in New Bottles?



New Wine in New Bottles!

