PUBLISHING DATA FROM THE SMITHSONIAN AMERICAN ART MUSEUM AS LINKED OPEN DATA

USC **Viterbi**
School of Engineering

INFORMATION SCIENCES INSTITUTE
· · · agent of innovation · · ·

Craig A. Knoblock, Pedro Szekely, Shubham Gupta, Animesh Manglik, Ruben Verborgh, Fengyu Yang, Rik Van de Walle
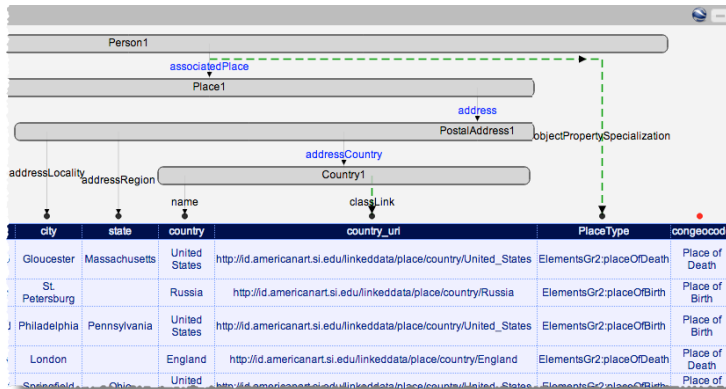
THE PROBLEM

- Museums have data about their art work that they want to publish and share
- But it is stored in proprietary formats without explicit connections
- The challenge is to relate this data to a cultural heritage ontology, create the RDF, and link the data to related sources
- Previous work on cultural heritage data required manual rules to model the data and had few links to other sources

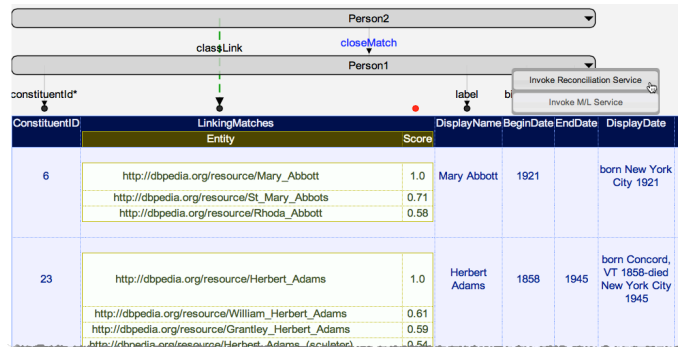## Example Smithsonian Data

Example Source Model in Karma

SEMI-AUTOMATIC MODELING

- We developed an approach to automatically map data to a domain ontology
- Supports complex cultural heritage ontologies including EDM and CDOC-CRM
- Handles data in a wide range of formats including relational and hierarchical data
- Learns from previous mappings to improve the system over time
- Allows the user to interactively refine the mappings if the system gets it wrong
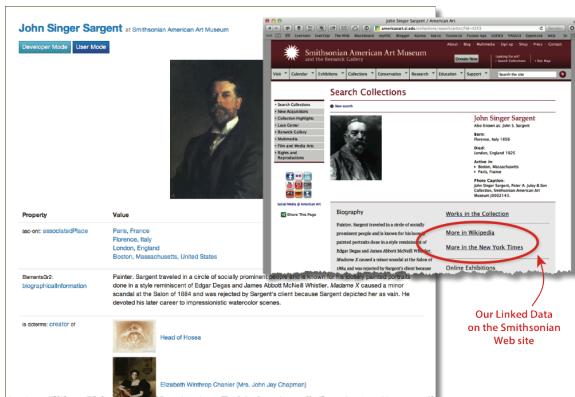
INTERACTIVE LINKING

- We integrated a linking services that allows a user to reconcile the entities (e.g., artists, locations, etc.) with other sources
- We used this service to link artists to Dbpedia and locations to Geonames
- Allows the user to view the results and curate the links to ensure the final results are correct

Example Linking in Karma

Example Visualization of Results

RESULTS AND FUTURE WORK

- We mapped the data from the Smithsonian American Art Museum on their 44,000 pieces of art and linked the data to DBpedia and Geonames
- The tools can easily be applied to new datasets from other museums
- We developed a user-friendly browser to view the data
- We are now working on tools to
  - curate data using other linked data sources
  - build virtual museums
  - build multimedia stories from the linked data