Data Curation for the Humanities

Geneva Henry, George Washington University

CNI Fall Meeting - December 8, 2014
Overview

• Where’s the data in humanities research?
• Creating sustainable digital content
• Teams and infrastructure
• Project case studies
Where’s the data in humanities research?

• Digital content enables structure to be added to otherwise unstructured resources
  – Metadata
  – OCR
  – Text mark-up
  – Georeferencing
  – 3D recreations of space

• Data enables new research never before possible
  – Social network analysis to discover historic relationships
  – Time and space analysis
  – Closer inspection of historic sites that no longer exist
  – Content analysis for frequency of terms (e.g. names, places) across many works
Creating Sustainable Digital Content

- Digital content is powerful, but enhancements must be maintained and reusable
- Balance in extreme mark-up/annotation vs. minimal additions of interpretation
- Managing elements associated with a work as separate objects
Teams and Infrastructure

• Domain expertise and technical expertise needed for success
  – Partnering between academic faculty and librarians/staff is powerful
• Opportunities for learning new skills
• Platforms that can handle varied content
Project Case Studies at Rice University

- Our Americas Archive Partnership (OAAP)
  http://oaap.rice.edu
- Travelers in the Middle East Archive (TIMEA)
  http://timea.rice.edu
- The Rice Ephemera Archive
  http://scholarship.rice.edu/handle/1911/39097
- The Shepherd School of Music Collection
  http://scholarship.rice.edu/handle/1911/43628
- The Houston Asian American Archives Oral Histories Collection (HAAA)
  http://scholarship.rice.edu/handle/1911/75168