Research Data Alliance
Practical Policies for Data Management Adoption

Platform for Experimental, Collaborative Ethnography (PECE)
RPI, Troy, NY, USA

Luis Felipe R. Murillo
(Berkman Center for Internet and Society)

Research Data Alliance – 6th Plenary
Paris, September, 23, 2015
What is PECE?

• Primary Goal: Provide a platform for data sharing and collaborative, interpretive (hermeneutic) analysis of ethnographic and historical data

• Developed by an interdisciplinary team primarily located at Rensselaer Polytechnic Institute, Troy, New York

• Pilot Projects: The Asthma Files Disaster STS research network Astroanthro.net
PECE + RDA

• PECE researchers are also active RDA members:
  – Digital Practices in History and Ethnography Interest Group
  – Empirical Humanities Metadata Working Group (pending)

• We are working on the adoption of the following RDA WG outputs:
  – Practical Policy Working Group
  – PID Collections Working Group (planned)
PECE Design: Working Tensions

- **Collaboration:** support the different interpretations from diverse body of researchers, research collaborators, and the public of any given data set

- **Technical and Domain-Grounded Innovation:** Making the platform innovative while driven by domain practice and theory (enabling new forms of ethnographic work and dissemination)

- **UI and UX:** support ways of visualizing heterogeneous types of data, working with legacy interpretations as/of data and producing experimental literary forms
RDA Practical Policies WG

**Output:** 11 computer-actionable data management policies

- Contextual metadata extraction
- Data access control
- Data backup
- Data format control
- Data retention
- Disposition
- Integrity (including replication)
- Notification
- Restricted searching
- Storage cost reports
- User agreements
Data Management for Humanities: Implementation Challenges

• **Data Protection:** sharing ethnographic data involves complex ethical and privacy concerns

• **Enhancing Collaboration:** enabling collaborative analysis and discovery without having a fixed data model and over-structuring user experience on the platform

• **Multi-field expertise:** need for more scholars who bridge the domain of computing with research in humanities and social sciences
PECE Data Management

• **Contextual Metadata Extraction:** *PECE Open API*

• **Data Access Control:** *role-based permission system and permissions per digital object*

• **Data Format Control:** *only open formats and standards are allowed in the platform*

• **Backup:** *automated, redundant, and encrypted*

• **Restricted Searching:** *based on user roles, easy to remove a digital object from the search index, distributed ElasticSearch*

• **User Agreements:** *prepared with the Cyberlaw clinic of the Berkman Center at Harvard*

• And more...
Institutional Review Board (IRB) informed consent form questions translated into PECE permission system

1. Agree to be video/audio recorded
   NO: no digital recording is made
   YES

2. Agree to have excerpts in:
   Written online and offline files
   NO: non-uploadable material
   YES

Oral presentations only:
   non-uploadable material

3. Agree to have them uploaded online?
   NO: non-uploadable material
   YES

4. Allow for the open* utilization of the material online?
   NO: restricted content type
   YES: open content type

5. Agree to archive content permanently?
   NO: set expiration date on content
   YES: open content type

6. Waive anonymity requirements?
   NO: apply anonymization procedure
   YES

* Open means content released with the standard PECE license:
Creative Commons Attribution-Sharealike 4.0 International
Planned Features of PECE 2.0

• New features:
  – Real-time, collaborative writing
  – Embedded metadata for binary files
  – Cryptography for sensitive data
  – WebRTC support for audio/video interviews and group meetings
  – Online IRB consent form, permission settings by research subjects
  – Micro-attribution through content versioning
Contact us!

- Luis Felipe R. Murillo
  lfmurillo@cyber.law.harvard.edu

- Download the PECE Data Management Practical Policies report (v.0.9) at the RDA file repository

- Find us at the adopters' poster session

- We are looking for collaborators