UCLA
Presentations

Title
The Data Conservancy: Science-driven Information Science

Permalink
https://escholarship.org/uc/item/9d30934x

Authors
Borgman, Christine L.
Palmer, Carole L.

Publication Date
2010-06-08

Copyright Information
This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, availalbe at https://creativecommons.org/licenses/by-nc-nd/4.0/

eScholarship.org Powered by the California Digital Library
University of California

UCLA Department of Information Studies

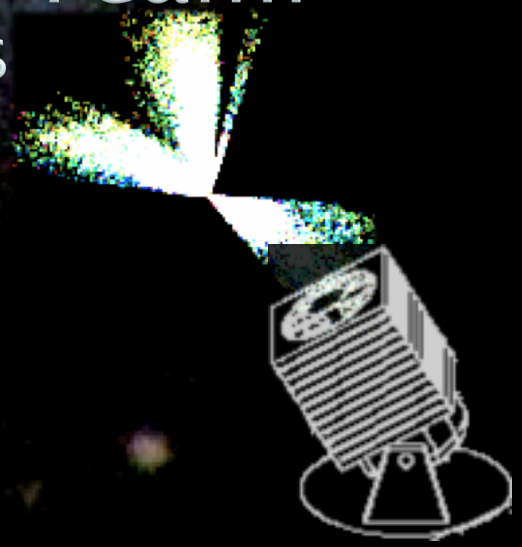
Curators to the Stars!

Berliner Bibliothekswissenschaftliches Kolloquium

UCLA Data Conservancy Team: Vision, Methods, Early Findings

8 June 2010

Christine Borgman,
Sharon Traweek, David Fearon, Laura Wynolds



The Data Conservancy

Data repository

- Infrastructure for multiple sciences
- Responsive to scientists' needs

Social science: finding contexts for data curation

JOHNS HOPKINS UNIVERSITY

UCAR
UNIVERSITY CORPORATION FOR ATMOSPHERIC RESEARCH

User-based design

Data Requirements

CIRSS UIUC

Comparative analysis

Life Sciences	Earth Sciences	Social Sciences

UCLA

Deep Case Analysis

Astronomy

Deep Case Analysis: Astronomy

Following projects and people

Sloan Digital Sky Survey (SDSS) → Pan-STARRS → Large Synoptic Survey Telescope

history, archives, data practices → people/knowledge transfer, development history, curation plans

Sloan Digital Sky Survey
http://www.sdss.org/photos/98_628.web1.jpg

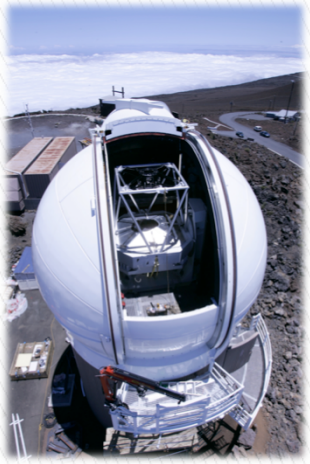
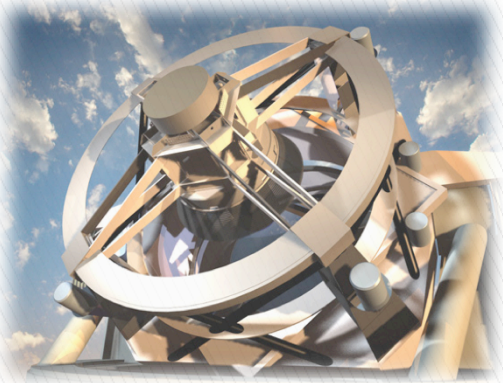


Photo by Brett Simison
http://pan-starrs.ifa.hawaii.edu



Todd Mason, Mason Productions Inc. / LSST Corporation

Following data practices & standards

Infrared Processing & Analysis Center

"Mature" data center organization

Faculty astronomers & small-scale projects

International Virtual Observatory Alliance (IVOA)

Space Telescope Science Institute Archives (STScI)

Curation & metadata standards, policies

Research Questions

1. Data practices

What are the data management, curation, and sharing practices?

2. Social networks

Who uses what data when, with whom, and why?

3. Curation

What data are most important to curate, how, and for whom?

Methods

Project analysis

- History
- Documents
- People, social networks

Data practices

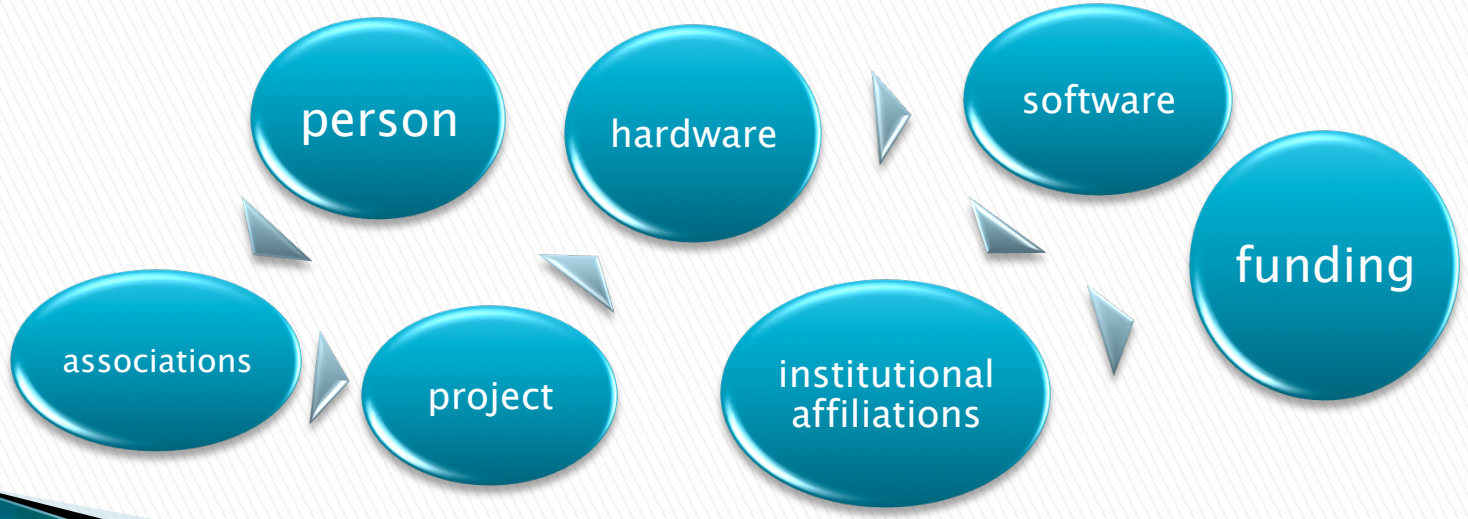
- Interviews, oral histories
- Data center observations
- Ethnography

Research repository

- Comparative analysis
- Best practices for curation
- Reflexive design and development

Data Collection & Analysis

- For data practices:
  - Open ended interview question set developed
  - Draft of code book in development
- For project and document analysis:
  - Data structure for entity mapping developed
  - Analysis of relationships over time among:

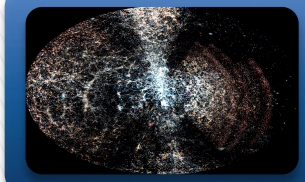


Data Collected to date

- Interviews at 3 sites
  - JHU
  - Caltech/IPAC
  - UCLA
- Documentation archive of project websites
  - SDSS
  - Pan-STARRS
  - LSST
- Archive of project publications and reports
- Bibliography of astronomy data practices

Our impressions so far…

Changes to astronomy field



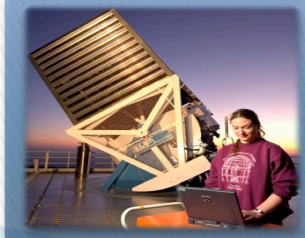
Early adopters of

- Digital Data
- Large-scale data sets

Established Data Centers

Data sharing practices

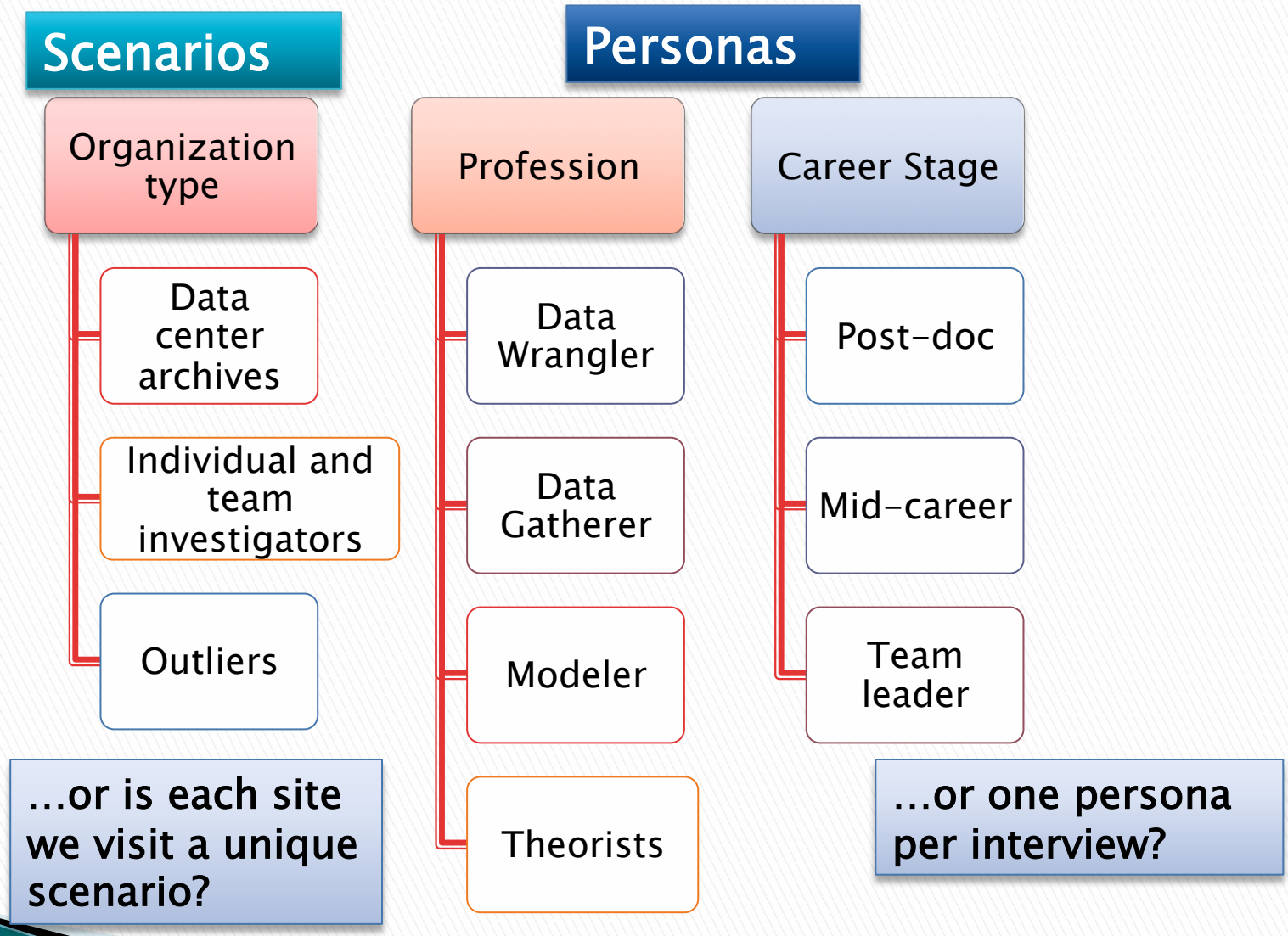
- Collaborative use of research instruments
- Data standards exist; some need for non-std
- Shared and local tools for data analysis



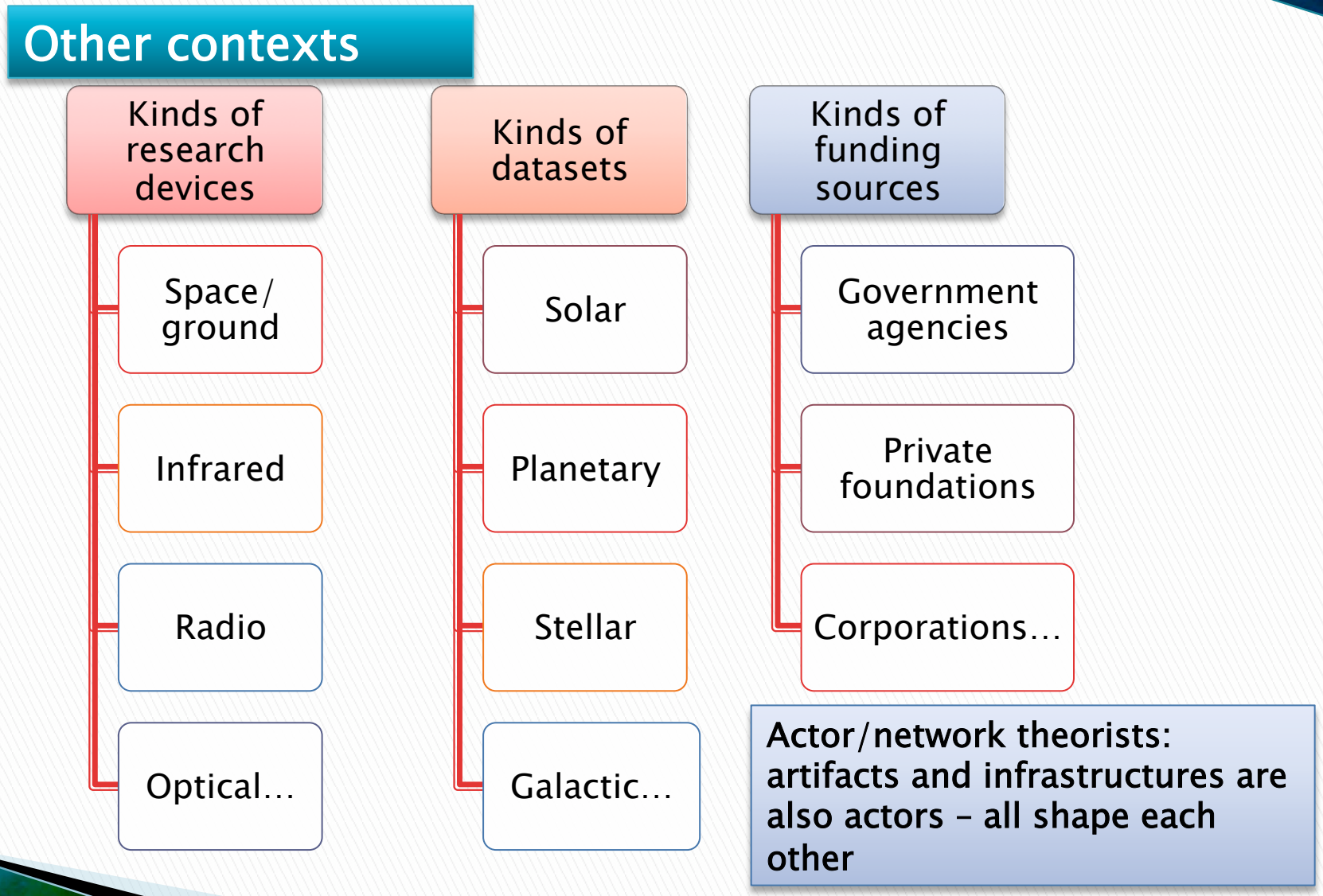
Evolving data management professions

- Data management roles
- Disparate relationships to data
- Data managers move with the data

Early observations



Early observations



UCLA Department of Information Studies

Curators to the Stars!

Acknowledgments:
NSF Data Conservancy Funding
Microsoft Research Partners:
Catherine van Ingen
Catherine Marshall
CENS Data Practices Team:
Sharon Traweek
David Fearon
Laura Wynolds
Alberto Pepe
Jillian Wallis
Matthew Mayernik
Katie Shilton