



Community Engagement

Mary Marlino and Carole Palmer



Overview

- Strategic goals for community engagement
 - Engaging diverse users and designing for a science-driven infrastructure
 - Libraries as a cornerstone of sustainable infrastructure
 - Capacity building for current and future professionals



Designing for a Science-Driven Infrastructure

- Engagement with disciplinary, data, and information scientists informs DC of the data issues within the Astronomical, Earth, Life, and Social Sciences
 - User-centered approaches promotes direct engagement with various communities
 - * Ethnographic research
 - * Multi-method approach studying practices across a variety of disciplines
 - * Action research
- Reflects both the expertise of the partners and our belief that a mixed method will provide a broader view of the challenges and complexities of data-rich science

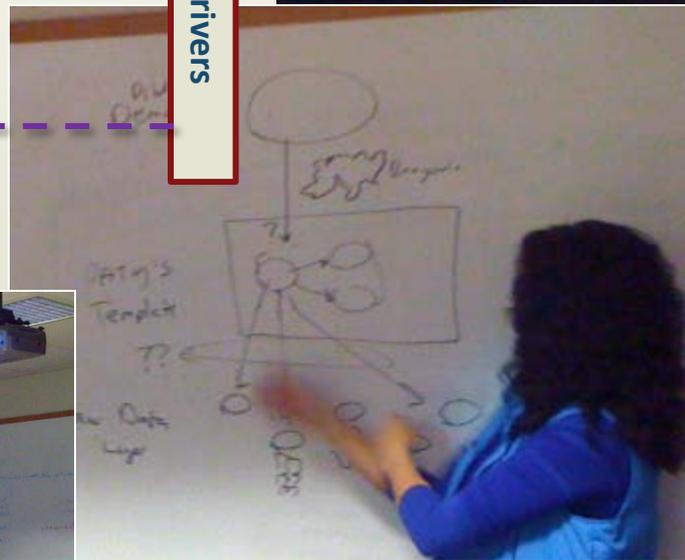
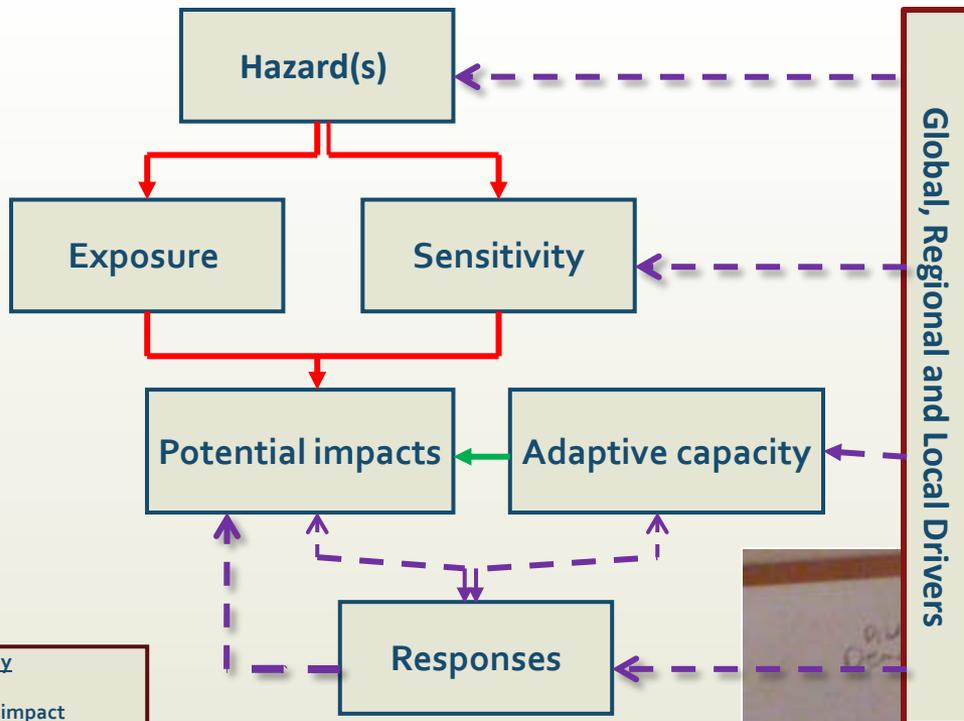


Understanding Scientific and User Needs

Multi-site user research methods are a blend of:

- Case study and domain comparisons
- Depth and breadth
- Local and global

	Astronomy	Life Sciences	Earth Sciences	Social Sciences	
UCAR		Task-based design and usability testing ⇒ Use cases, data requirements, system recommendations			UCAR
UCLA	Ethnography, virtual ethnography, oral histories ⇒ Use cases, data requirements	Interviews, Surveys, Worksheets, Content analysis ⇒ Curation requirements, taxonomy, metadata/provenance framework			UIUC



Key

- Positive impact
- Negative impact
- - - Positive or negative Impact





Resilient-Sustainable Cities (RS-Cities)

- Exploring how urban areas worldwide can more effectively respond to climate change
 - Vulnerable populations (and environments) are at risk to events exacerbated by climate change (heat waves, floods, droughts, sea level rise)
- Urban areas account for 50 – 70% of global carbon emissions
 - Urban centers play a key role in mitigation and adaptation of climate change

Currently little research has focused specifically on the relationship between urban vulnerability and climate change





RS-Cities Research Questions

- A primary anthropogenic driver of climate change is carbon emissions in the form of greenhouse gases
 - What are the common **dimensions and determinants** of carbon emissions, vulnerabilities and required capacities?
 - What **information is needed** by cities to effectively **understand** the complexities *unique to their socio-environmental conditions*?
 - What information is needed to **mitigate** the effects of carbon emissions in **their locale**?
- * Urban vulnerability is characterized and determined by thresholds, tipping points, 2nd and 3rd order impacts, and responses
 - Katrina: Lack of Adaptive Capacity: infrastructure, emergency response, collective civil action





Climate 2.0 - Usable Science for Society

The fundamental question that society is asking of climate science has dramatically changed.

Climate 1.0

Is anthropogenic climate change occurring?

- Classic, low-resolution, global climate modeling (past 40 years)

After IPCC AR₄ findings, the question is now....

Climate 2.0

What is the impact of this climate change on our coupled human & natural systems?

- Sustainable Systems: Energy, Food, Water, Health, Cities, Ecosystems
- Societal Impacts: GIS, extremes, climate services, policy responses
- Unit of analysis is local in a global context

Addressing these more complex questions require new approaches:

- Data provenance and preservation
- Transparency
- Interdisciplinary data
- Integration and synthesis



Methodologies and Goals

- Meta-analysis (RS-Cities)
 - Goal: identifying dominant...
 - * Theories
 - * Research methodologies
 - * Data analysis techniques
 - Analysis from 217 case studies of urban vulnerability
- Action Research (DC)
 - Goal: understanding data and data practices
 - * Expert interviews of case study authors
 - * Analysis of data use in case studies
 - * Survey of case study authors



Data Conservancy-Class Problems

- Data provenance and preservation problem
 - RS-Cities Meta-analysis e.g.: 217 case studies including hundreds of datasets, many are no longer accessible
- Transparency problem
 - With missing data, science is not reproducible
- Interdisciplinary data problem
 - Few consistent protocols and analysis tools due partially to multi-disciplinary nature of the data
- Integration and synthesis problem
 - Variations in data practices
 - Semantics problem
 - * Definitions, e.g.: population density, city boundaries poorly defined
 - Data organization
 - * Schemas, structures, metadata, accessibility



Engaging Allied Professional Communities

- Co-development
 - PBS&J (pbsj.com)
 - * Co-developed use case for determining the location of endangered species in Alaska for the purpose of ice road planning
 - * Resulted in a proof-of-concept cross data set query tool now available for use
- Leveraging expertise
 - BioSharing (biosharing.org)
 - * Consortium dedicated to building standards and policies to encourage sharing of biological data.
 - Earth Science Information Partners (esipfed.org)
 - * Network of scientists, data stewards and technology developers that improves access to Earth science data and information.
 - Integrated Earth Data Applications (iedadata.org)
 - * Lamont Doherty Observatory



Libraries for Sustainable Infrastructure

- New responsibilities and leadership in data curation
- Source of information for science and collection curation expertise
- Key component of academic infrastructure
- DC in national level initiatives
 - DC participant in ARL-sponsored institutes on strategic planning and data curation activities
 - CLIR and DLF-requested data curation “boot camps” for future post docs
 - Outreach to professional library groups
- UCAR is window to 150 academic institutions, internationally



Capacity Building for Current and Future Professionals

New Professionals

- Data curation specialization in MSLIS
- Doctoral research track

In-Service Professionals

Summer Institute in Data Curation

**Informed by
Data Conservancy
R & D**

Educators in the Information Professions

- Workforce Summit
- Curriculum database & Clearinghouse

Educators in Science

- New summer institute model
- Curation & reuse curriculum



Preparing new professionals and researchers

- New and revised courses
 - Data, Data Practices, & Data Curation at UCLA
 - Foundations of Data Curation
 - Identifier Semantics
 - Information Transfer & Collaboration in Science
 - Systems Analysis & Design
- Advances in master's specialization in curation of research data
 - Fellowships targeted for under-represented groups
 - Fieldwork opportunities (practicums and internships)
- Doctoral research to build knowledge base (Illinois, UCLA, Cornell)
 - Apprenticing and building own research programs
 - Invaluable project & collaboration, course development experiences
 - Building cohort across schools to lead next generation of faculty



Preparing current professionals for new roles

- Summer Institute on Data Curation - 4-day workshops
 - 2010: Earth Sciences – 35 participants
Major contributions from DC & DO contributions
 - 2011: Life Sciences
Scope and expertise expanded through DC
 - 2012: New model – information professional / scientist pairs



Building the Education Community

Data Curation Workforce Summit, 6 December 2010 – 31 participants

- DOE, ORNL, NOAA, CIESIN, NCAR, i-schools and data science programs
- build awareness and synergies among educational activities
- perspectives from employment sectors
- current and projected workforce needs and practices

Clearinghouse

- Summit and SIDC materials
- Database of curriculum / courses across universities
 - 63 universities with 594 courses
- UCLA syllabus collection

Teaching materials for scientists' undergraduate and graduate courses

- principles and practices of data curation
- working with archived data – hypothesis development and reuse issues



Community Engagement – Summing Up

- Engaging diverse users and designing for a science-driven infrastructure
 - Informing development of infrastructure to ensure that it is responsive to communities we serve
 - data expertise needed at all stages of data life cycle from planning and production stages of scientific process, to repository to reuse.
- Libraries as a cornerstone of sustainable infrastructure
 - Helping libraries to become key providers of curatorial services
- Capacity building for current and future professionals
 - Building capacity through workforce development
 - Building a significant cadre of professionals who will curate data and data products