Repurposing Data Across Disciplines:
A Study of Data Reuse Issues Between Climate Science & Social Science
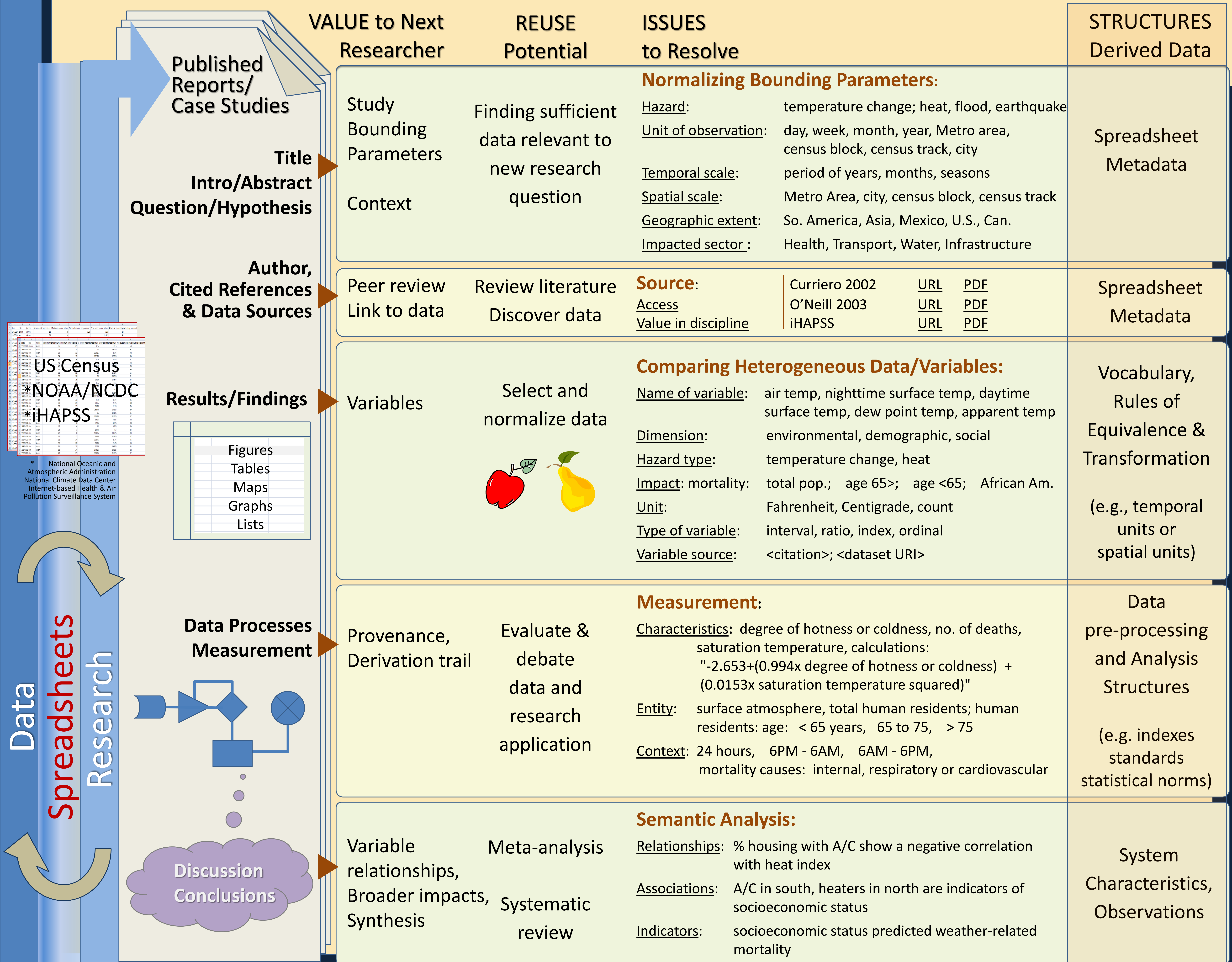*'What factors will improve urban resilience to impacts of climate change?"*

Lynne Davis, University Corporation for Atmospheric Research
Peter Alston, University of Colorado at Boulder
John D'Ignazio, Syracuse University

	VALUE to Next Researcher	REUSE Potential	ISSUES to Resolve	STRUCTURES Derived Data
Published Reports/ Case Studies			**Normalizing Bounding Parameters:**	
Title Intro/Abstract Question/Hypothesis	Study Bounding Parameters Context	Finding sufficient data relevant to new research question	Hazard: temperature change; heat, flood, earthquake Unit of observation: day, week, month, year, Metro area, census block, census track, city Temporal scale: period of years, months, seasons Spatial scale: Metro Area, city, census block, census track Geographic extent: So. America, Asia, Mexico, U.S., Can. Impacted sector: Health, Transport, Water, Infrastructure	Spreadsheet Metadata
Author, Cited References & Data Sources	Peer review Link to data	Review literature Discover data	**Source:** Curriero 2002 — URL — PDF Access — O'Neill 2003 — URL — PDF Value in discipline — iHAPSS — URL — PDF	Spreadsheet Metadata
Results/Findings	Variables	Select and normalize data	**Comparing Heterogeneous Data/Variables:** Name of variable: air temp, nighttime surface temp, daytime surface temp, dew point temp, apparent temp Dimension: environmental, demographic, social Hazard type: temperature change, heat Impact: mortality: total pop.; age 65>; age <65; African Am. Unit: Fahrenheit, Centigrade, count Type of variable: interval, ratio, index, ordinal Variable source: <citation>; <dataset URI>	Vocabulary, Rules of Equivalence & Transformation (e.g., temporal units or spatial units)
Data Processes Measurement	Provenance, Derivation trail	Evaluate & debate data and research application	**Measurement:** Characteristics: degree of hotness or coldness, no. of deaths, saturation temperature, calculations: "-2.653+(0.994x degree of hotness or coldness) + (0.0153x saturation temperature squared)" Entity: surface atmosphere, total human residents; human residents: age: < 65 years, 65 to 75, > 75 Context: 24 hours, 6PM - 6AM, 6AM - 6PM, mortality causes: internal, respiratory or cardiovascular	Data pre-processing and Analysis Structures (e.g. indexes standards statistical norms)
Discussion Conclusions	Variable relationships, Broader impacts, Synthesis	Meta-analysis Systematic review	**Semantic Analysis:** Relationships: % housing with A/C show a negative correlation with heat index Associations: A/C in south, heaters in north are indicators of socioeconomic status Indicators: socioeconomic status predicted weather-related mortality	System Characteristics, Observations

US Census
*NOAA/NCDC
*iHAPSS

Figures
Tables
Maps
Graphs
Lists

* National Oceanic and Atmospheric Administration National Climate Data Center Internet-based Health & Air Pollution Surveillance System

Data
Spreadsheets
Research

Deposit/curate derived data (spreadsheets) and metadata

Date | Metro Area | air temp | dew point temp | Apparent Temp | mortality: total population | mortality: age less than 65

Dataset 1: Curriero 2002
Dataset 2: O'Neill 2003
Dataset 3: iHAPSS

Combined from datasets 1 and 2

Transformed temperature via rules

Aggregated from similar data with different spatial scales (City to Metro area)

EXAMPLE: What is the relationship between air temperature and mortality?