TerraPop Mission

Enable research, learning, and policy analysis by providing integrated spatiotemporal data describing people and their environment.

TerraPop Goals

- Lower barriers
  - Single location for data
  - Interoperability of data across scientific domains
  - Easier to acquire and use geographic data
  - Easier to conduct interdisciplinary human-environment interactions research
- Organization and technical framework
  - Preserve
  - Integrate
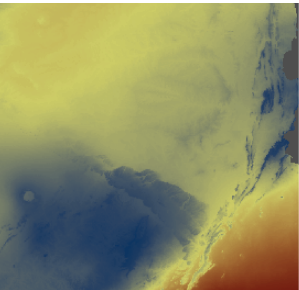 - Disseminate
  - Analyze

Data Formats

Microdata: Characteristics of individuals and households.
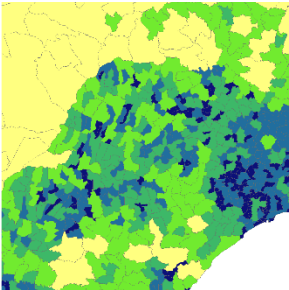
Area-level data: Characteristics of places defined by boundaries

Raster data: Values tied to spatial coordinates

Data Formats
Microdata Structure

Age **Sex** **Race** **Birthplace** **Occupation**

```
H910000240000000088001001000220100
P9100000201010321200100100100110011504
P9100000102010362200100100100110011999
P9102010003010112200600100100110011999
P9102010003010091200600100100110011999
P9102010003010071200600100100110011999
P9102010003010061200600100100110011999
P9102010003010042200600100100110011999
P9102010003010032200600100100110011999
P9102010003010022200600100100110011999
H910000240000000088001001000110100
P9100000201010301100102905105111310
P9100000102010212100102902901711999
P9102010003010011100600102902911999
H910000240000000088001001000220100
P9100000201010451200100100100110011100
P9100000102010252200100100100110011820
P9102010003010072200600100100110011999
H910000240000000088001001000220100
P9100000201010491200100100100110011100
P9100000102010492200100100100110011820
P9102010003010192200600100100110011820
P9102010003010152200600100100120820
```

Geographic location & housing characteristics

Household record (shaded) followed by a person record for each member of the household.

For each type of record, columns correspond to specific variables.

Data Formats
Area-level Data

Geographic Identifier	County ID	Label	Population
0500000US01001	1001	Autauga County, Alabama	54,571
0500000US01003	1003	Baldwin County, Alabama	182,265
0500000US01005	1005	Barbour County, Alabama	27,457
0500000US01007	1007	Bibb County, Alabama	22,915
0500000US01009	1009	Blount County, Alabama	57,322
0500000US01011	1011	Bullock County, Alabama	10,914
0500000US01013	1013	Butler County, Alabama	20,947
0500000US01015	1015	Calhoun County, Alabama	118,572
0500000US01017	1017	Chambers County, Alabama	34,215
0500000US01019	1019	Cherokee County, Alabama	25,989
0500000US01021	1021	Chilton County, Alabama	43,643

US Population 2010

County Population

Value

82 -	103,505
103,506 -	331,298
331,299 -	703,462
703,463 -	1,385,108
1,385,109 -	2,504,700
2,504,701 -	5,194,675
5,194,676 -	9,818,605
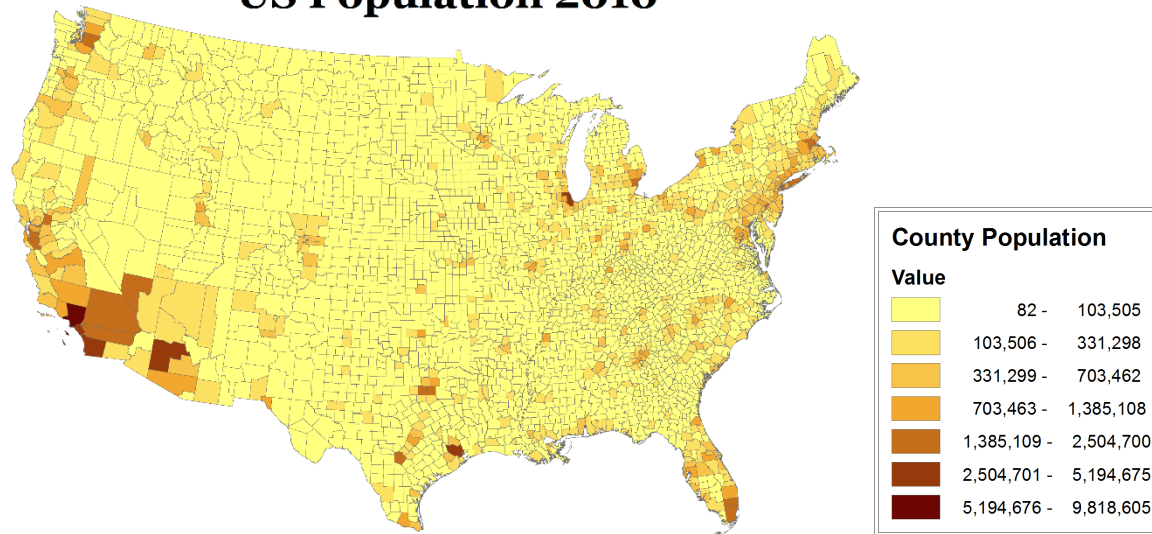
Data Formats
Raster Data
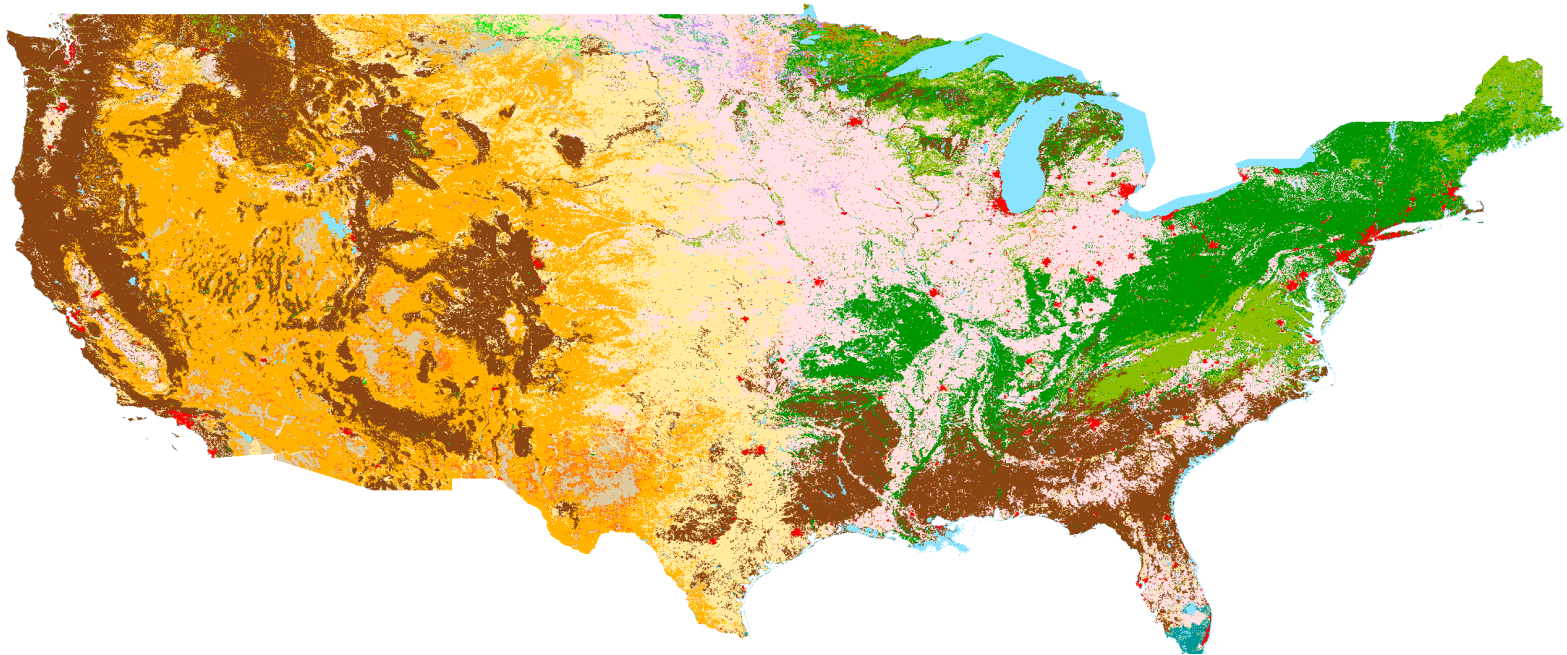
Global Landcover 2000

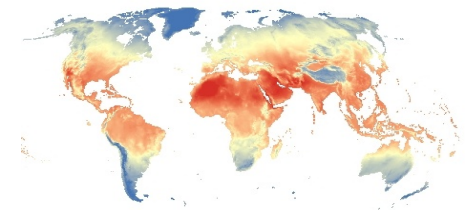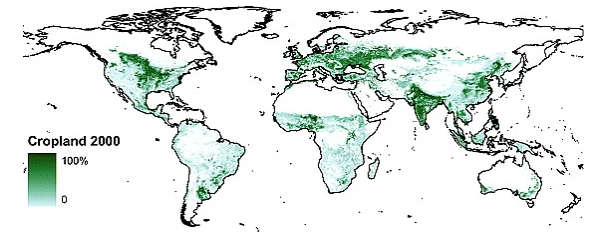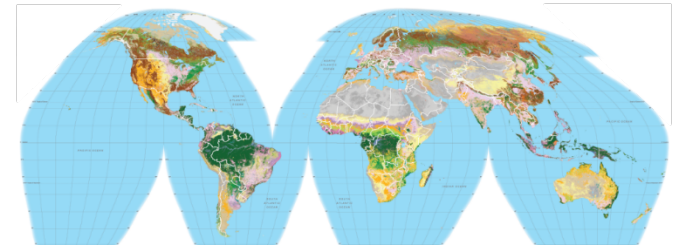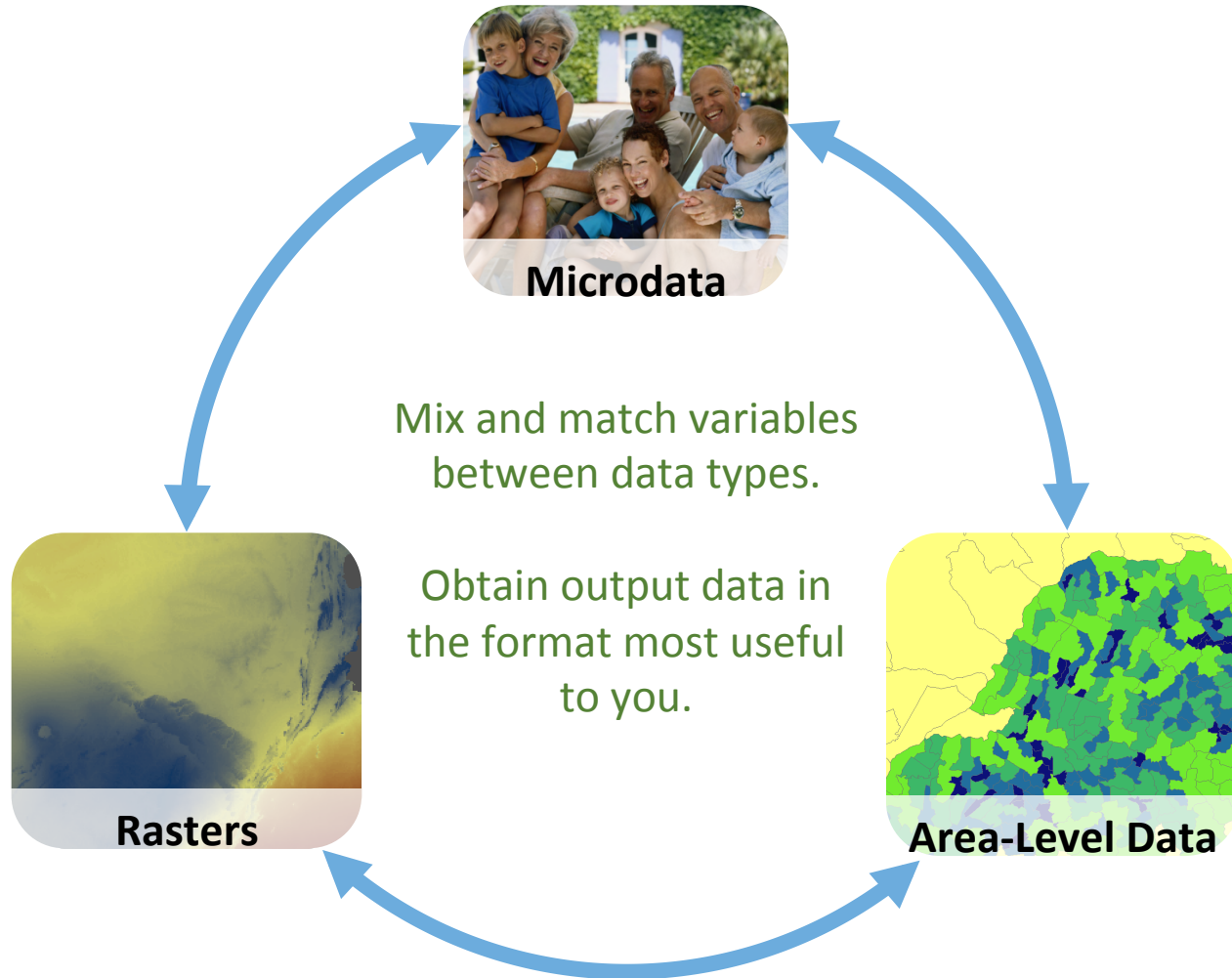|			
■ Artificial surfaces	■ Cropland / Tree Cover / Other natural vegetation	■ Sparse herbaceous or sparse shrub cover	■ Tree, needle-leaved, deciduous
■ Bare Areas	■ Tree Cover / Other natural vegetation	■ Tree, broadleaved, deciduous, closed	■ Tree, needle-leaved, evergreen
■ Cultivated and managed areas	■ Regularly flooded shrub and/or herbaceous cover	■ Tree, broadleaved, deciduous, open	■ Tree, regularly flooded, fresh water
■ Herbaceous Cover, closed-open	■ Shrub Cover, closed-open, deciduous	■ Tree, broadleaved, evergreen	■ Tree, regularly flooded, saline water
■ Irrigated Agriculture	■ Shrub Cover, closed-open, evergreen	■ Tree, burnt	■ Water Bodies
■ Cropland / Shrub and/or grass cover	■ Snow and Ice	■ Tree, mixed leaf type	

Available Data
Environmental Data

- Land Use & Land Cover
  - MODIS Land Cover 2001-2012
  - Global Land Cover 2000
  - Harvested Area and Yield for 175 crops
- Climate
  - Annual bioclimatic variables derived from CRU-TS
  - Long-term average temperature and precipitation
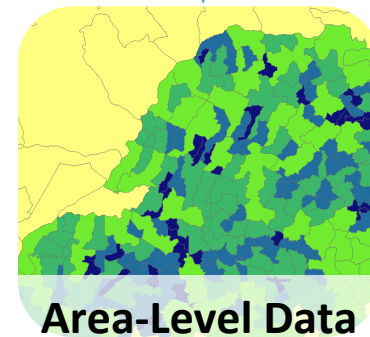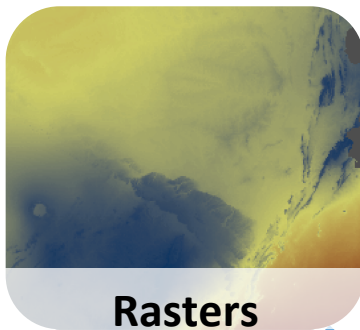
Location-Based Integration

Microdata

Rasters

Area-Level Data

Mix and match variables between data types.

Obtain output data in the format most useful to you.

Location-Based Integration

County ID	Avg. Ann. Temp.	Avg. Ann. Precip.	Rent, Rural	Rent, Urban	Own, Rural	Own, Urban
G17003100001	21.2	768	3129	1063	637	365
G17003100002	23.4	589	2949	1075	1469	717
G17003100003	24.3	867	3418	1589	1108	617
G17003100004	21.5	943	1882	425	202	142
G17003100005	24.1	867	2416	572	426	197
G17003100006	24.4	697	2560	934	950	563
G17003100007	25.6	701	2126	653	321	215

Output

Microdata

Summarized environmental and population characteristics for administrative districts.

Rasters

Area-Level Data

Location-Based Integration

AGE	SEX	LANDCOV	AVGTEMP
10	Male	Forest	21.20
27	Female	Forest	24.30
54	Female	Pasture	24.10
37	Male	Cropped	25.60
37	Female	Cropped	28.10
42	Female	Urban	26.70
20	Female	Forest	24.30

Output

Microdata

Individuals and households with their environmental and social context

Rasters

Area-Level Data

Location-Based Integration

Output

Microdata

Rasters of population and environment data
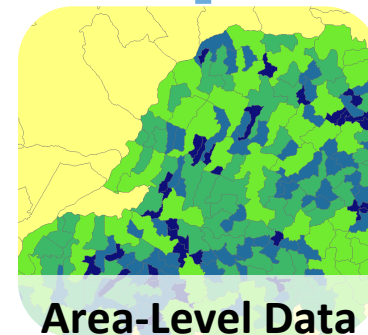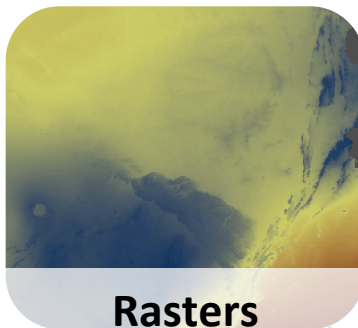
Rasters
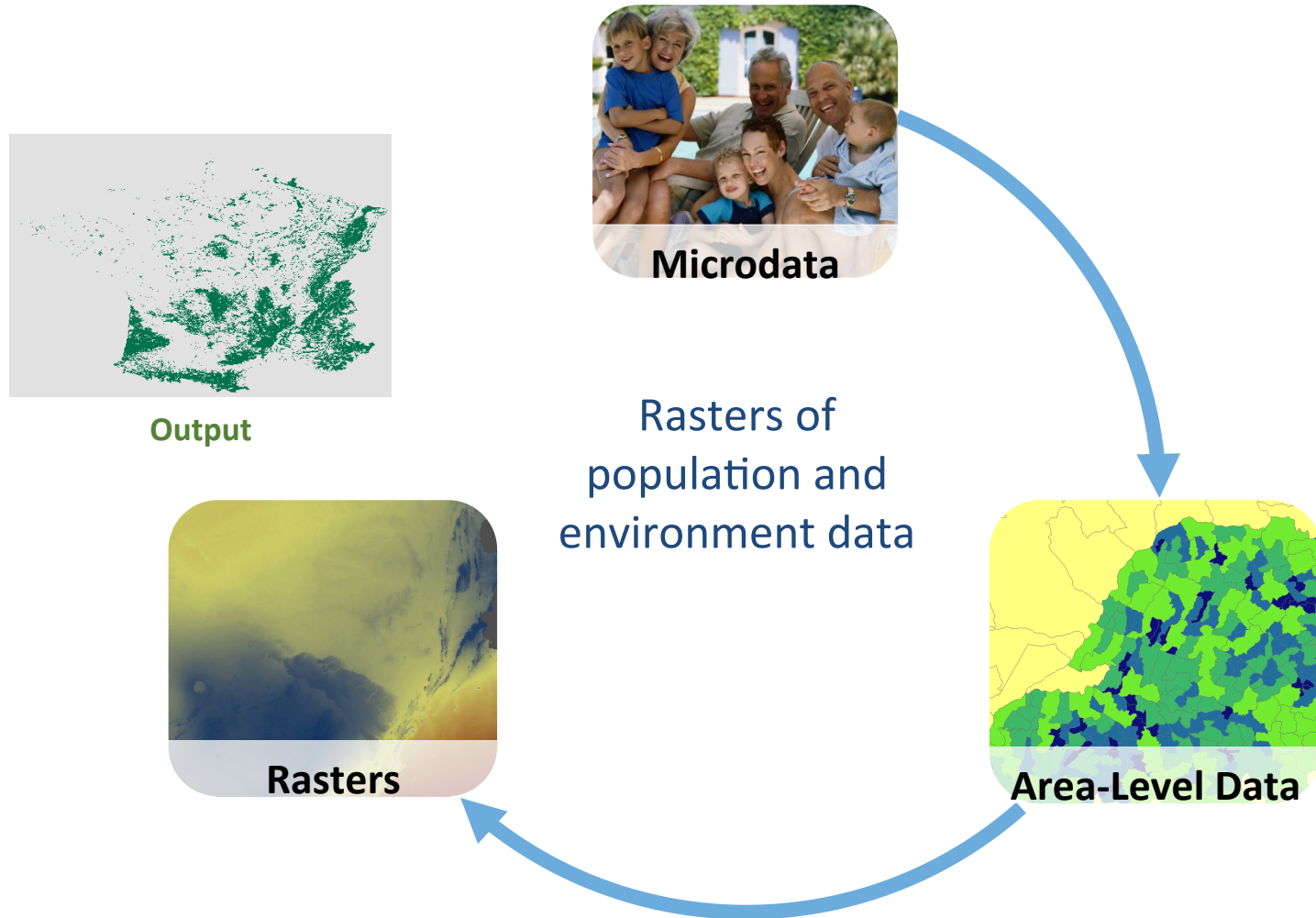
Area-Level Data

Location-Based Integration
Boundaries are Key

- Linkages across data formats rely on administrative unit boundaries
  - Containers for summarizing raster data to area-level data
  - Containers for distributing area-level data to raster cells
  - Codes link area-level and summarized raster data to microdata
- Sets of units and codes must match census data

TerraPop Boundaries
Boundary Processing

1. Source data

2. Align boundaries

3. Match codes

4. Historical adjustments

5. Harmonize and regionalize

Applications

- Data Extract Builder
  - Create an extract by selecting area-level variables and/or raster variables for a dataset
- TerraClip
  - Extract country-level subsets of global raster datasets
- TerraScope
  - Explore area-level and raster variable availability
  - Visualize the variables in a web map

Next Steps

- Data Releases
  - Environmental and Climate Data
    - GLI Crop Time Series
    - CRU monthly time series – precipitation and temperature
    - Vegetation characteristics – NDVI, greenness
    - Elevation and derived characteristics
    - Soils
    - Species distribution
  - Gridded Population of the World
  - Aggregate census data
    - Historical data (48 countries)
    - Variables in addition to population by sex (65 countries)

TerraPop Funding

National Science Foundation's (NSF) Office of Cyber Infrastructure. NSF Sustainable Digital Data Preservation and Access Network (DataNet) partner.