Geocoding at the Centers for Disease Control and Prevention

Andrew L. Dent, MBA, MA

adent@cdc.gov

Geographic Research, Analysis, and Services Program (GRASP)

Division of Health Studies, NCEH-ATSDR

SAFER · HEALTHIER · PEOPLE™

Welcome

- GIS at the CDC

- What is Geocoding?

- Geocoding at the CDC

- Geocoding and Surveillance Systems

- Geocoding and Privacy/Confidentiality

GIS in Public Health

- Roles of GIS in Public Health
  - **Characterize** populations
  - **Visualize** patterns
  - **Analyze** population, health outcome, and environmental data
  - **Integrate** disparate data sets on a single platform
  - **Communicate** findings and results

SAFER · HEALTHIER · PEOPLE™

Geographic Research, Analysis and Services Program (GRASP)

- Created in 1988 through EPA and DOE assistance

- Long-standing role in providing mapping and spatial analysis support for ATSDR and CDC

- Most GIS activities centered on protection of persons living near hazardous waste sites

- Other activities include cancer cluster investigations, review of health information in spatial format, and mapping data for presentation to the public.

CDC

What is Geocoding?

- Geocoding is the process of developing **coordinate information** based on geographically-referenced data.

- **Key:** Geocoding can be accomplished using many different mechanisms, but it always involves linking geographically-referenced data to spatial data that has a known position.

- **Definition:** Geographically-referenced data is data that contains an **identifier for spatial data** (ex: address, county FIPS Code, mile marker)

Geocoding to a Street Center Line

- **Process**
  - ◆ Match address record to line segment
  - ◆ Use linear interpolation to position point along the line segment
  - ◆ Derive and store coordinates

100 Oak Street

Name	RightLow	RightHi	LeftLow	LeftHi
Oak St	1	49	2	50
Oak St	51	151	50	150
Main St	21	135	22	136

50   150

51   Oak Street   151

CDC

Geocoding to a Mile Marker

- Process
  - Match record to mile marker point
  - Store point coordinates

Mile Marker #11 I-85

Name	Lat	Long
MM10	33.1235	-88.5678
MM11	33.1286	-88.5678
MM12	33.1299	-88.5678

I-85
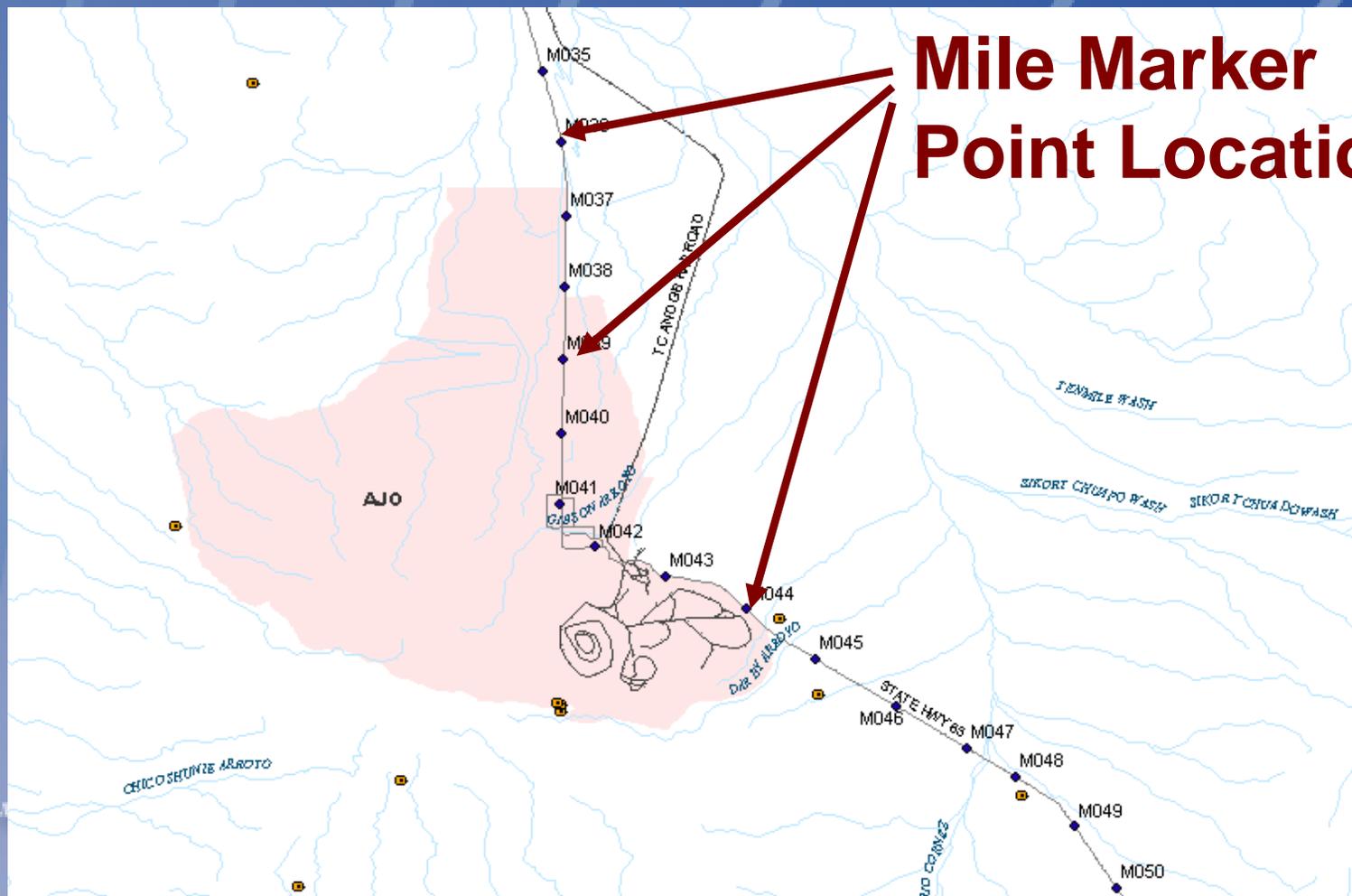
Geocoding to Areal Geographic Units

- Process
  - ◆ Match record to area unit centroid
  - ◆ Store point coordinates

Tax Parcel 06-327

Name	Centroid Lat	Centroid Long
06-326	33.1235	-88.5678
06-327	33.1286	-88.5678
0	33.1299	-88.5678

CDC
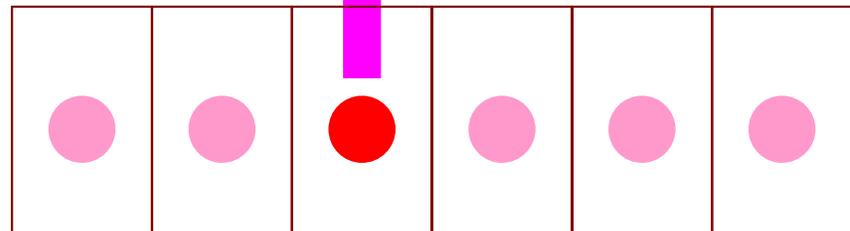
Geocoding to Areal Units
Metro Atlanta Birth Defects Project (Environmental Health Tracking Network)

- Purpose
  - Link air quality data to birth defect data
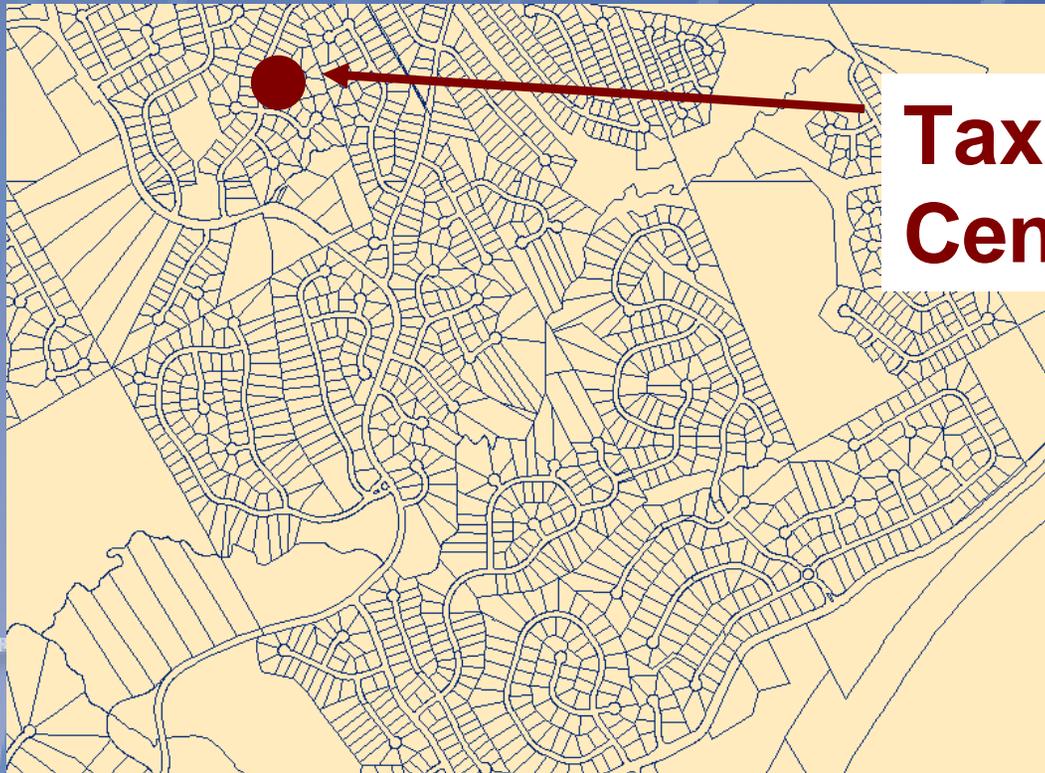
Tax Parcel Centroid

Geocoding and Surveillance Systems

- Surveillance systems at the CDC collect a wide range of information – from mortality/natality statistics to information on behavioral risk factors (BRFSS)

- Geocoding this data enables researchers **analyze the relationships** that these data have with other key environmental, social, and epidemiological factors.

Hazardous Substance Emergency Event Surveillance System (HSEES)

- **Mission** Capture and visualization of data pertaining to the release of hazardous contaminants into the environment.

- **Users** Officials in 18 states currently capture and manage hazardous release information using HSEES.

HSEES Inline Geocoding

Address | **2** | GRASP Geocoding WebService

3

Coordinates

Event Info

State Workstation

CDC HSEES Server

1

4

HSEES Record

CDC HSEES Database

SAFER · HEALTHIER · PEOPLE™
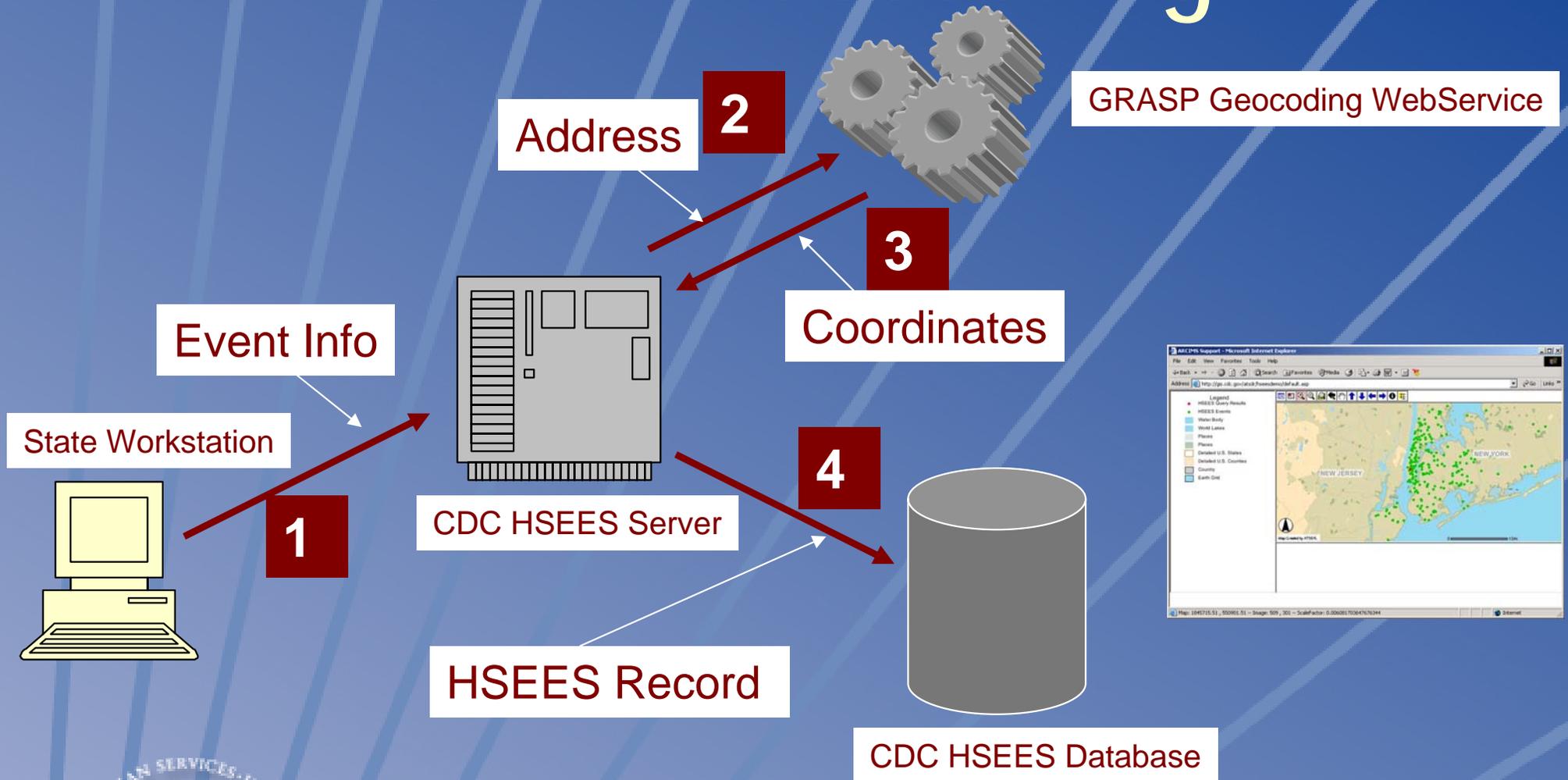
Privacy/Confidentiality in the Use of Geocoded Data

- HIPAA (Health Insurance Portability and Accountability Act of 1996
    - HIPPA Rules clarify that "identifiable" include information such as **address or latitude/longitude**
- **How do we conform to these standards?**

HIPAA Compliance

- When geocoding …

  - Process de-identified files only (with address or geo-id and NO name, SS#, etc)

- When mapping …

  - Randomize point locations

  - Aggregate cases to areal units (ZIP Code, census units, county, etc)

The End

- Questions / Comments?