Geocoding Public Health Data: The Next Generation

Stephanie Foster, Olayiwola Taylor, Andrew Dent

ESRI Health GIS Conference
October 19, 2010
Presentation Overview

- Project Background
- Data Acquisition
- Methodological Issues
- Next Steps
2500 BC
Clay Tablet Map from Ga-Sur

http://www.henry-davis.com/MAPS/AncientWebPages/100mono.html
Frames of the Greek world
Homer, Hecateus, Eratosthenes, and Ptolemy
Definitions

- Geocoding is “the process by which an entity on the earth’s surface, a household, for example, is given a label identifying its location with respect to some common point or frame of reference”\(^1\)

- Address matching is “A specialized case of feature matching, strictly dealing with matching postal street addresses to features in the reference data source, usually TIGER type street segments or areal unit delineations (Census delineations, USPS ZIP Code de-lineations, etc.).”\(^2\)

---

1 (Goodchild, 1984, p. 33).
Geocoding Systems Options

- Geocoding options
  - Commercial
  - Open source
  - Build your own
What option is best?

- Number of records geocoding
- Level of expertise and technology support
- Cost
Previous Geocoding Systems Evaluations and Experiences

- Previous research describing variability by geocoding software
- Limited nationally representative data
- Limited evaluation of match and error codes
## Data Acquisition

<table>
<thead>
<tr>
<th></th>
<th>Geocoder 1</th>
<th>Geocoder 2</th>
<th>Geocoder 3</th>
<th>Geocoder 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>WholeDw dataframe</td>
<td>TeleAtlas</td>
<td>TeleAtlas</td>
<td>USPS Zip +4</td>
<td>TIGER/line files</td>
</tr>
<tr>
<td>Addressrange</td>
<td>Dsghvfudgrj</td>
<td>Dsghvfudgrj</td>
<td>Wh JH U2dhyUuv</td>
<td>Dsghvfudgrj</td>
</tr>
<tr>
<td>QDYWH Twhhw 5343</td>
<td>Dsghvfudgrj</td>
<td>Dsghvfudgrj</td>
<td>Dsghvfudgrj</td>
<td>Dsghvfudgrj</td>
</tr>
<tr>
<td>XVSV</td>
<td>Dsghvfudgrj</td>
<td>Dsghvfudgrj</td>
<td>Dsghvfudgrj</td>
<td>Dsghvfudgrj</td>
</tr>
</tbody>
</table>

- 2 Commercial
- 2 Open-source
- TeleAtlas 2008
Data Acquisition

- **Data for evaluation**
  - Public health focus
  - Address level data
  - EMS emergency responses to cardiac arrest events
  - Residential and business addresses
  - Cities across the United States

- **Element for comparison**
  - Tax parcel polygons
Issues with Geocoding Methodology
Issues with Geocoding Methodology
Issues with Geocoding Methodology
Issues with Geocoding Methodology
Issues with Geocoding Methodology
Issues with Geocoding Methodology
Issues with Geocoding Methodology
Issues with Geocoding Methodology
Issues with Geocoding Methodology
Solutions for Geocoding Methodology
Next Steps

- Larger dataset with greater tax parcel polygon and metadata
- Explore matching tax parcels with street addresses
- Evaluation of geocoded points to parcel centroids and/or centerline offsets, address ranges
- Utilization of orthoimagery
CARES Project Team

- Comilla Sasson, MD, PhD
- Linda Schieb, MSPH

GIS Offices

- Anchorage
- Columbus
- Denver
- Durham
- Fulton
- Houston
- Kansas City
- Las Vegas
- Richmond
- San Diego
- Sioux Falls
- Sterling Heights
- Suffolk
- Travis
- Wake
- Williamson
- Wilmington

Geospatial Research, Analysis, and Services Program, ATSDR

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.