Navigating HIPAA in a Geospatial World

Este Geraghty, MD, MS, MPH, GISP
Chief Medical Officer & Health Solutions Director, Esri
@EsteGeraghty
October 4, 2019
FOR IMMEDIATE RELEASE
May 6, 2013
PH13-017

CONTACT: Anita Gore
Heather Bourreau
(916) 440-7259

California Birth Records Found in Unsecure Location

SACRAMENTO - A California Department of Public Health (CDPH) investigation has found that a reel containing images of 2,000 State of California Birth Records for 1974 was found in an unsecure location. These birth certificates contain names, addresses, Social Security numbers and some medical information.

CDPH has no evidence that this incident has resulted in any unauthorized use of this information. CDPH continues to collaborate with the California Highway Patrol in this investigation.
We responded with…

- Increased training for staff
- No more data ‘check outs’
- Enhanced physical security for reels
- Sped up digitization plans
- Notifications to those impacted
- And more…
3 major concerns:

1) Guard against data breaches

2) Avoid misinterpretations of data

3) Don’t embarrass the administration
3 major concerns:

1) Don’t embarrass the administration!

2) Don’t embarrass the administration!

3) Don’t embarrass the administration!
Geraghty says the birth of California’s first Health and Human Services (HHS) open data portal is much like the birth of a child: It took about nine months to create and is the object of great affection.
Overview

• We can overcome perceived barriers
• Digital transformation creates challenges as well as opportunities
• GIS is a valuable part of digital transformation
• HIPAA is a perceived barrier to using GIS
• We can make progress and reap the rewards
Technology is playing a fundamental role . . .

. . . Integrated Into everything we do

Digital Transformation Is Just Beginning . . .

We Are Living in an Era of Exponential Technological Advancement

Pervasive Understanding of Our World

Artificial Intelligence

Design Thinking

Autonomous Vehicles

Sensed Data

Smart Devices

Automation

Immersive Experiences

Geospatial Solutions

Real-Time Monitoring

Data-Driven Analytics

Predictive Analytics

Advanced Modeling

Pervasive Mapping

Cloud Computing

Robotics

Efficiency and Collaboration

IoT

Open Data Access and Transparency

Location Intelligence

Synthetic Biology

Data Linkages

Persistent Surveillance

Technology is playing a fundamental role . . .

. . . Integrated Into everything we do
How Do We Take the Next Step?

Embrace Digital Transformation . . .

. . . and Leverage Technology Thoughtfully

Envision a Better Future

Create Solutions

Innovate

Societal Good

Inform & Educate

Understand the Possibilities

Learn Continuously

Take the Initiative

Participate & Take Action

Embrace Technology

Work Together

Create Enabling Technologies Foundation for Our Future
We’re still human

• Everybody loves to innovate, but nobody likes to change

• Status quo is the enemy of innovation

• What can we unlearn today?
A Conundrum:

“Data can either be useful or perfectly anonymous but never both” (Ohm, 2009)
Why We Care

High quality data is needed to:

- Inform policy creation
- Plan & prioritize initiatives
- Target interventions
- Evaluate outcomes
Why GIS?
aka
Location Intelligence
For centuries we’ve focused on the microscopic
The world of the small could be tricky!

Joseph Lister, the Father of Antiseptic Surgery
Our faith in the small flourishes

Watson & Crick discover the structure of DNA in 1953
The power of genetics is undeniable
We are unique

Our Genes

Shape us

Compel us

Make us distinctive
Shifting Focus

Appreciate the big picture with a telescopic view
Excluding genetics, 80% of our health outcomes are context driven.
Place Matters for Health!

Our communities have a powerful impact.
Make the healthy choice the easy choice

Our environments influence our personal behaviors
Our perception is based on our context

Our environments shape how we see our world.
Which is more important?

YOUR ZIP CODE > GENETIC CODE
The first step for a diabetic patient to be able to take their medications...
...is having the transportation to go and get them
Maybe it's not the will to eat better that's scarce...
...but rather its access to grocery stores that’s out of reach.
Not every community will have a fancy gym...
...but we know that the population can flourish when simple sidewalks are available.
Place Matters for Health!
Place Matters for Health

WHERE people live and the environments around them, has a significant impact on their personal health and outcomes

Hippocrates, 460 BC
THE SCIENCE OF WHERE
The Science of Where

Helps us to answer the fundamental questions of where . . .

- Where is it
- How do I get there
- What’s nearby
- Where are we going
- Where’s the problem
- Where is it changing
- Where is the issue
- Where is it suitable
- Where should we locate
What is where?
Why is it there?
Why do I care?
How can I prepare?
Leveraging the Power of Geography . . . to Make Better Decisions

A Framework & Process  Applied to the complete health workflow

Geographic Knowledge

Measuring

Data Management & Integration
Visualization & Mapping
Analysis & Modeling
Planning & Design
Decision-Making

Analyzing

Action & Outcomes

Understanding

Collaborating

Leveraging the Power of Geography . . . to Make Better Decisions
Many elements for success

ArcGIS Software Platform

- Data & Solutions
- Business Partners
- Education & Non-Profits
- Training
- Disaster Response
- Professional Services
- Publishing
Homelessness
Census tracts with homeless Veterans

Placing Veteran Beds in Shelters
Homeless Shelters in LA, per HUD

Placing Veteran Beds in Shelters

[Map of Los Angeles area showing homeless shelters and veteran beds]
Veteran beds within a 15 minute walk time

Placing Veteran Beds in Shelters
Vector Control
No more children with microcephaly

Mosquito Traps

Adulticiding
Suitable *Aedes* habitats

- **Locations where Aedes may be present year-round**
- **Seasonal locations where Aedes may be present**
Biosurveillance through Social Media

At the Subdivision by Basin 115

There were so many mosquitoes at the field we could not fly there. So we will check out the Basin for Mosquitoes on May 23rd to see how many. We moved to an abandoned subdivision at time there. Carl, Bobbi, Kevin, Ken and 2 other widows John.

#Mosquitoes

Icy lane closed for standing water in #Phoenix on Lp 101 EB at 51st Ave — Exit 22 #phxtraffic traffic https://t.co/ULr7Ze9nk

10:59am - Aug 10, 2016

Twitter (1 of 2)
Preparedness
California plans for medical surge & connects citizens to healthcare services.
CDC's Social Vulnerability Index
Socioeconomic Theme

Vermont Social Vulnerability Index (SVI)

This map shows the number of socioeconomic vulnerability measures above the 90th percentile for Vermont census tracts. For each of the vulnerability measures, census tracts in the 90th percentile of vulnerability were assigned a flag. These flags were then summed for each theme. There are five socioeconomic measures: Poverty, Unemployment, Income per capita, Education, and Health Insurance.

For More Information (PDF) About the SVI...
For More Information (Video) About the SVI...

The darker blue categories on this map are census tracts where there are more flagged socioeconomic variables, while the lighter yellow categories have fewer flags.

Census tracts marked with a yellow triangle have one or more of their state flags from estimates with high relative standard errors. These flags may be less accurate than the others.

The underlying population in each Vermont census tract may affect the SVI measures. For more information on Vermont's population density and college-aged population: Click Here

Source: U.S. Census ACS 2011-2015
Download the SVI Data
Compare to Other ACS 2009-2013 data

View the New Hampshire SVI by Tract map (90th percentile breaks)
View the U.S. Natives SVI by Counties map (U.S. percentiles breaks)

For more information contact: AVS-DOHVS@Vermont.gov
Housing & Transportation Theme

Vermont Social Vulnerability Index (SVI)

Housing/Transportation Theme

This map shows: the number of housing/transportation vulnerability measures above the 50th percentile for Vermont census tracts. For each of the vulnerability measures, census tracts in the 90th percentile of vulnerability were assigned a flag. These flags were then summed for each theme. There are five housing/transportation measures: Large Apartment Buildings, Mobile Homes, Crowding, No Vehicle, and Group Quarters.

View maps showing individual housing/transportation measures

For More Information (PDF) About the SVI...
For More Information (Video) About the SVI...

The darker blue categories on this map are census tracts where there are more flagged socioeconomic variables, while the lighter yellow categories have fewer flags.

Census tracts marked with a yellow triangle have 50% or more of their SVI flags from estimates with high Relative Standard Errors. These flags may be less accurate than the others.

The underlying population in each Vermont census tract may affect the SVI measures. For more information on Vermont’s population density and college-aged population: Click Here.

Source: U.S. Census ACS 2011-2015
Download the SVI Data
Compare to Older ACS 2008-2013 Data
View the New Hampshire SVI Tracts map (90% percentile breed)
View the U.S. Nationwide SVI Data by Counties (Map, U.S. percentile breed)

For more information, contact: 4VSOH@Vermont.gov
Opioid Epidemic
The world is aware
Celebrating Lost Loved Ones to the Opioid Epidemic - see new link at https://losttooopioids.nsc.org

Mark McManus, Marthas Vineyard

Mark is our son, brother, nephew, cousin, uncle and friend. We lost this big smiling, kind and gentle man on October 8, 2014. Mark was 23 years old. In life, Mark's presence lit up every room with a smile, a joke and a hug. He spent his free time by the ocean, surfing, fishing, and relaxing with his friends and family. Mark had a strong work ethic with a Bachelor of Arts in Business. Later in life addiction shadowed these gifts. But this disease will never define him. Mark walks beside us today as proof that addiction knows no boundaries. The smartest, most beautiful and loving souls can still be taken. In his memory we educate all who will listen. We miss you Mark, every minute of every day. May you rest in Peace. yecforgmcmanus
Better Decisions – Better Health
Counties combat heart disease by improving the built environment

Klamath County, Oregon
Kids can be placed in better foster homes
Early Death

Complications of Diabetes

Motor Vehicle Accidents

Cancer

Early Death
Navigating HIPAA in a Geospatial World

* This presentation is informational only and does not constitute legal advice.
HIPAA REQUIREMENTS
Protected Health Information (PHI)

- Any individually identifiable health information including demographic information, that is created, used, disclosed, or received by a health care provider, health plan, or health care clearinghouse. PHI includes data about past, present, and future health conditions, health care provision, and health care payments.

- Examples of PHI include name, address, telephone number, birth date, and Social Security Number

(HHS, 2015)
Protected Health Information (PHI)

• Any individually identifiable health information including demographic information, that is created, used, disclosed, or received by a health care provider, health plan, or health care clearinghouse. PHI includes data about past, present, and future health conditions, health care provision, and health care payments.

• Examples of PHI include name, address, telephone number, birth date, and Social Security Number

(HHS, 2015)
Who is affected by the Privacy Rule?

- Covered entities (note: not all health data is PHI)
- The Privacy Rule does not directly apply to research but may affect researchers because it may affect their access to information, but it does not regulate them or research, per se.
- To gain access for research purposes to PHI created or maintained by covered entities, the researcher may have to provide supporting documentation on which the covered entity may rely in meeting the requirements, conditions, and limitations of the Privacy Rule.
Geocoding, a first step to using geography

- Take care when geocoding PHI
- Beware of unintentionally obfuscating your results
  - Be sure you have high quality reference data
  - Adjust your matching rules to suit the situation
Geocoding - DONTs

- You may NOT give your data to someone else to geocode for you if they are not listed on your IRB protocol or given permission to handle your PHI (such as a BAA)
- You may not geocode PHI in ArcGIS Online unless deployed in a private cloud
  - What about just submitting the addresses?
Geocoding - DOs

• You can geocode using your ArcGIS desktop software if you have reference data (like the Esri World Geocoder or StreetMap Premium) available on that machine.

• You can geocode in your enterprise implementation of your ArcGIS Platform if you have ‘portal’ installed / turned on AND you have reference data also within your portal environment (i.e. on premise).

• You can geocode in a HIPAA compliant cloud (e.g. Spatialitics Health).
Geo-Enabling the Healthcare Enterprise
An Overview
Spatialitics™ Health is built on a foundation of HIPAA-Compliant data management policies and rules and is supported by HIPAA-Compliant cloud infrastructure.
Encode

Simple and Secure Geocoding of your PHI/PII Data

HIPAA Cloud Environment

SUBMIT
Apps

PROCESS
Address to coordinate transaction

RETURN
Input returned with coordinates appended to each record
GEOCODING PROCESS FLOW

1. Upload files (e.g., .csv, .xls, .xlsx)
2. Write to API
3. Triggers run
4. Extract only addresses
5. Read from Azure Table Storage
6. Geocoding service (ArcGIS Server)
7. Write to Azure Table Storage
8. On complete, read from Azure Table Storage
9. Update file with result
10. Update file with result
**SPATIALITICS - HIPAA-COMPLIANCE OPERATIONAL LAYERS**

<table>
<thead>
<tr>
<th>Layer</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAA</td>
<td>Agreements with Microsoft Azure for Cloud Infrastructure and with each customer</td>
</tr>
<tr>
<td>OPERATIONS</td>
<td>HIPAA-trained, US-based Cloud Operations staff</td>
</tr>
</tbody>
</table>
| DATA                   | Partitioned cloud storage models in order to keep each customer’s data truly separated, leveraging Azure’s HIPAA-compliant storage infrastructure  
                         | All PII are kept out of the Geocoding or other Spatial Analysis processes     |
| SOFTWARE PROCESSES     | Configured APIs and Service Engines to ensure that multiple users are running in partitioned environments |
| NETWORK PARTITIONS      | Virtual cloud network configured to keep tenant services and data separated and compliant |
| SECURITY               | SAAS Service with a set of Administrative, Physical, and Technical Safeguards to ensure the confidentiality, integrity, and security of protected health information (PHI) |
GEOCODING HIPAA-COMPLIANT FEATURES

Data Security
Partitioned storage to keep each customer’s data separated with all PII out of the Geocoding or other Spatial Analysis processes

Secure Infrastructure
Hosted on Azure’s secure HIPAA-compliant cloud infrastructure

Complete Encryption
TLS 256-bit AES data encryption at server and in-transit with access only to authorized users

Trained Team & SOPs
Trained team and HIPAA-compliant helpdesk incident and change management, following protected data SOPs and policies

Regular Audit
Consistent auditing and testing for internal and external vulnerability assessment

Third-Party Certification
Yearly third-party certification to guarantee adherence with HIPAA compliance
## Organization Name

### Overview
- **Quick Select:** Last 30 days
- **From:** 02/06/2018
- **To:** 02/06/2018
- **Go**

### Usage
- Showing records from: Last 30 days

<table>
<thead>
<tr>
<th>User</th>
<th>Role</th>
<th>Geocodes</th>
<th>Usage Type</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steve</td>
<td>User</td>
<td>1002</td>
<td>File upload</td>
<td>Jun 05, 2018</td>
</tr>
<tr>
<td>Caroline</td>
<td>Administrator</td>
<td>12452</td>
<td>API</td>
<td>Jun 04, 2018</td>
</tr>
<tr>
<td>Bruce</td>
<td>User</td>
<td>800</td>
<td>File upload</td>
<td>Jun 03, 2018</td>
</tr>
</tbody>
</table>
Spatialitics Health Geocoder User Interface

Batch Geocode

Step 1 of 3: Upload file to geocode

Choose File  No file chosen

1. Supported file types .xlsx, .csv
2. First row must contain headers
3. First worksheet from Excel workbook will be used for geocoding
4. Support file size is up to 1 GB

[Cancel] [Next]
Spatialitics Health Geocoder User Interface

Batch Geocode

Step 2 of 3: Upload file to geocode

- **Street**: select column
- **City**: select column
- **State**: select column
- **Zip**: select column

Address information in single column?

- [ ] Cancel
- [ ] Back
- [ ] Next
Spatialitics Health Geocoder User Interface

Batch Geocode

Step 3 of 3: Upload file to geocode

<table>
<thead>
<tr>
<th>FileName</th>
<th>No. of rows</th>
<th>credit</th>
</tr>
</thead>
<tbody>
<tr>
<td>sample geocode.csv</td>
<td>346</td>
<td>xx</td>
</tr>
</tbody>
</table>

By clicking "Start" you confirm that you agree for to our Terms of use.
Spatialitics Health Geocoder User Interface

View geocodes on map
Spatialitics Health Website
https://spatialitics.com/health/

Videos

Spatialitics Health: Product Introduction
https://youtu.be/BOLbyU0hFtM

Spatialitics HIPAA-Compliant Geocoder
https://youtu.be/KYrM67TzAUc

Spatialitics Health: Clinical and Business Apps
https://youtu.be/2yyoUCm7xSA
Data De-Identification Under HIPAA

Safe Harbor

Removal of 18 types of identifiers

No actual knowledge residual information can identify individual

Many organizations select Safe Harbor due to its simplicity
Safe Harbor identifiers

1. Names
2. Dates (except year)
3. Telephone numbers
4. Vehicle identifiers
5. FAX numbers
6. Device identifiers & serial numbers
7. Email addresses
8. URL’s
9. Social Security Numbers
10. IP addresses
11. Medical record numbers
12. Biometric identifiers (incl finger and voice prints)
13. Health plan beneficiary numbers
14. Full-face photographs & comparable images
15. Account numbers
16. Certificate/license numbers
17. Any other unique identifying number or characteristic
18. Geography
Eliminate all geographic subdivisions smaller than a state, including street address, city, county, precinct, ZIP code, and their equivalent geocodes, except for the initial three digits of the ZIP code if, according to the current publicly available data from the Bureau of the Census that geography has < 20,000 people.
Mapping county-level data may violate HIPAA

• 10% of all U.S. counties have more than 200,000 residents

• More than 40% of counties have fewer than 20,000 residents
  - E.g. 8/58 California counties have < 20,000 people

Plumas  Trinity
Inyo    Modoc
Mariposa Sierra
Mono    Alpine
Data De-Identification Under HIPAA

- **Safe Harbor**
  - Removal of 18 types of identifiers
  - No actual knowledge residual information can identify individual

- **Expert Determination**
  - Apply statistical or scientific principles
  - Very small risk that anticipated recipient could identify individual

**HIPAA Privacy Rule De-identification Methods**
Numerator Rule

- CDC WONDER databases suppresses cells with numerators <10
- National Environmental Public Health Tracking Network suppresses cells that are greater than 0 but less than 6
- CMS suppresses values of <11
- Range is 3 to 40
- Most fall into the range of 10-15
Small cell suppression

- Geographic masking to suppress cell counts $\leq 10$ on a map
- It is often necessary to suppress complementary cells as well to avoid the calculation of actual cell values by subtraction or other mathematical operations
Denominator Rule

\[
\frac{3}{5}
\]

- National Center for Health Statistics – 250,000
- National Environmental Health Tracking Network – 100,000
- Maine Integrated Youth Health Survey – 5,000

Risk varies based on the level of ZIP code and how the ZIP code is combined with other variables.

Date of Birth + Sex + 5-Digit ZIP code is unique for over 50% of US residents, meaning that with these 3 pieces of information more than half of US residents could be identified.

Year of Birth + Sex + 3-Digit Zip code is unique for 0.04% of US residents.
Aggregation

You can aggregate on any of your cells: age, racial/ethnic categories, time periods (months, years), geography.
Randomly offsetting point data in a way that obscures actual address locations (in blue), while preserving the overall spatial pattern. Also known as geomasking. Focus on the ‘k’ statistic.
Numerator OR Denominator rule is okay.
Must meet numerator AND denominator condition

OR

Apply scoring criteria

IF Score $\leq 12$ (go to risk assessment)

IF Score $>12$ (Suppress small and complimentary cells first)

THEN

Assess legal, policy and programmatic risk

---

**Figure 3: Reporting Assessment Decision Tree**

Assesses risk for data release of aggregate data through a stepwise process. Aggregate data may be derived from record level data with identifiers, record level data without identifiers or previously aggregated data.

1. **Step 1 – Numerator Condition**
   - Have the Numerators (the table cells) been derived from greater than 10 members (beneficiaries)?
   - If Yes, Go to Step 2
   - If No, Go to Step 3

2. **Step 2 – Denominator Condition**
   - Is the population denominator for the numerators in the table cells greater than 20,000 individuals?
   - If Yes, Go to Step 5
   - If No, Go to Step 3

3. **Step 3 – Apply Publication Scoring Criteria to assess risk**
   - If the score is $\leq 12$, Go to Step 5
   - If the score is $>12$, Go to Step 4

4. **Step 4 – Suppress Small Cells and Complimentary Cells**
   - Small Cells are those with numerators fewer than 11 and Complimentary Cells are those that could be used to recalculate the Suppressed Small Cells

5. **Step 5 – Submit Aggregate Data Analysis for Document Review**
   - Program Management Review
   - Expert Determination Review*
   - OLS Review for legal risk
   - OPA Review

*Review for Expert Determination will be performed by individuals who have been qualified as experts by OLS and who meet the HIPAA Privacy Rule implementation specifications. “A person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable.” [45 CFR Section 164.514(b)(1)]
Publication Scoring Criteria
(*originally from Illinois Dept of Public Health)

- Focuses on most commonly used variables in health
- Two identification criteria assessed:
  - Variable specificity (generally numerator)
  - Size of potential population (generally denominator – like geography)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Characteristics</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Male or Female</td>
<td>+1</td>
</tr>
<tr>
<td>Age Range</td>
<td>&gt;10-year age range</td>
<td>+2</td>
</tr>
<tr>
<td></td>
<td>6-10 year age range</td>
<td>+3</td>
</tr>
<tr>
<td></td>
<td>3-5 year age range</td>
<td>+5</td>
</tr>
<tr>
<td></td>
<td>1-2 year age range</td>
<td>+7</td>
</tr>
<tr>
<td>Race Group</td>
<td>White, Asian, Black</td>
<td>+3</td>
</tr>
<tr>
<td></td>
<td>Detailed Race</td>
<td>+5</td>
</tr>
<tr>
<td>Hispanic Ethnicity</td>
<td>yes or no</td>
<td>+2</td>
</tr>
<tr>
<td></td>
<td>Detailed ethnicity</td>
<td>+3</td>
</tr>
<tr>
<td>Language Spoken</td>
<td>English, Spanish, Other Language</td>
<td>+2</td>
</tr>
<tr>
<td>Events</td>
<td>1000+ events in a specified population</td>
<td>+2</td>
</tr>
<tr>
<td></td>
<td>100-999 events</td>
<td>+3</td>
</tr>
<tr>
<td></td>
<td>11-99 events</td>
<td>+5</td>
</tr>
<tr>
<td></td>
<td>&lt;11 events</td>
<td>+8</td>
</tr>
<tr>
<td>Geography</td>
<td>State or geography with population &gt;2,000,000</td>
<td>-5</td>
</tr>
<tr>
<td></td>
<td>Population 560,001 - 2,000,000</td>
<td>-3</td>
</tr>
<tr>
<td></td>
<td>Population 20,001 - 560,000</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Population ≤ 20,000</td>
<td>+5</td>
</tr>
<tr>
<td>Data Year</td>
<td>5 years aggregated</td>
<td>-5</td>
</tr>
<tr>
<td></td>
<td>2-4 years aggregated</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1 year (e.g., 2001)</td>
<td>+3</td>
</tr>
<tr>
<td></td>
<td>Bi-Annual</td>
<td>+4</td>
</tr>
<tr>
<td></td>
<td>Quarterly</td>
<td>+5</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>+7</td>
</tr>
</tbody>
</table>
• “Guidelines and Methods for De-Identifying Protected Health Information.” Institute for Families in Society, University of South Carolina, October 2017.


• “To Whom Does the Privacy Rule Apply and Whom Will It Affect?” NIH. https://privacyruleandresearch.nih.gov/pr_06.asp

What Does MapMasq™ Do?

• Take discrete data points and mask the spatial location.
• Randomly distort data while attempting to preserve spatial patterns.
• Support for variable aggregation into “bins”
• Aggregation based on census populations
• MapMasq is the calculator, not an “easy button” to HIPAA compliance – depending on the dataset, locations, and sensitivity other methods may need to be applied. – “Expert Determination”
• Links to published studies at mapmasq.com in the resource section
Two Areas of Masking

- **Retains Point Data**
  - K-Mean Neighbor
  - "Jiggering"
  - Rounding/Truncating
  - Random Obfuscation

- **Aggregation**
  - Standard Bins
    - Triangles, Squares, Hexegons
  - Census Driven
    - Numerator Rule
    - Denominator Rule
      - Age, Sex, Combination
Triangle, Square, Hexagon Bins
Census Numerator Rule
Census Denominator Rule
Rounding/Truncating Coordinates (lat/long)

• Stacks the points
• Ability to “aggregate” without loss of individual points!
Coordinate Rounding & Truncation

- Two simple displacement algorithms: rounding and truncation
- Allows you to enter the number of spaces past the decimal point on your latitude and longitude to round or truncate to
- Tendency to aggregate data points to a single location depending on the number of spaces past the decimal you are working with
- A report back to the user will indicate how the truncation coarsens the accuracy
- Maintains attribute features – it's still a point dataset (stacked)
Donut Masking

- Extent of the masking can be set by drawing a box or assuming your extent is the data frame in view
- The blue dots are the original record location
- A circle is generated for the minimum displacement distance (blue) – *that record cannot be placed within that circle*
- The outer circle (orange) is the farthest displacement distance – *the record should not be placed outside of that circle*
K Mean Neighbors

• Primary donut masking algorithm
• K is the number of neighbors around the feature to consider
• Creates the inner and outer circles of the donut based on the value of K and the density / distance of its neighbors
• Each point shifts within the donut in a random compass direction
• Includes a K-Anonymity score
K Mean Angular Neighbors

- The direction a feature is moved is based on the location of its neighbors
- Builds a range in the centroid of the direction of the nearest neighbors
- Considers the angular displacement of its neighbors relative to North
- Increases the number of circles that overlap
- Includes a K-Anonymity score
K-Mean Neighbor, “jiggering”, “Donut”
Assessing Risk

Conservative Parameters <1% with Score of 0
Deployment Information

• Works as an extension to ArcGIS Pro
• Can be use separately for your data
• Incorporate into an enterprise platform
The Future

• Always exploring new techniques, such as differential privacy and principal components analysis
• Your ideas / needs are welcome
Geraghty says the birth of California’s first Health and Human Services (HHS) open data portal is much like the birth of a child: It took about nine months to create and is the object of great affection.
What Happened Next?

- The other 12 HHS Departments began contributing data
- The state gov ops group launched their open data portal
- The open data handbook is online and used as a guide for many organizations wanting to share data
- 4 new laws before the legislature
Increased sharing
• Intra-departmental
• Inter-departmental
• Fewer FOIA requests
Innovation

- Inter-departmental Open Data Working Group
- Hackathons
- Innovation Fellow
- Innovation Program
Undersecretary
Secretary
Governor appointee for Innovation & Digital Services
Summary

• We can overcome perceived barriers
• Digital transformation creates challenges as well as opportunities
• GIS is a valuable part of digital transformation
• HIPAA is a perceived barrier to using GIS
• We can make progress and reap the rewards