Navigating HIPAA in a Geospatial World

Este Geraghty, Esri

Nicholas Poorte, Spatialitics

Lucas Green, GISi

* This presentation is informational only and does not constitute legal advice.

A Conundrum

"Data can either be useful or perfectly anonymous but never both" (Ohm, 2009)

Why We Care

Population health improvement and community development require quality data to inform policy creation, planning, programming, evaluation and other critical decision-making processes.

Additional Value

1. If limited demographic and socioeconomic variables are available on study subject, their location can provide proxy variables
2. Insight into other variables which may be related to health outcomes, like travel time to nearest health facility or distance to pollution sources


HIPAA
REQUIREMENTS

Health Insurance Portability and Accountability Act

- Federally enacted in 1996
- Is really about continuity of health insurance coverage for workers and their families in the event of a job loss or change of employment
- Includes a Privacy Rule – to protect individually identifiable health information (aka Protected Health Information) that is recorded, stored, or exchanged in any form or medium (e.g. on paper, orally, or electronically)

Protected Health Information

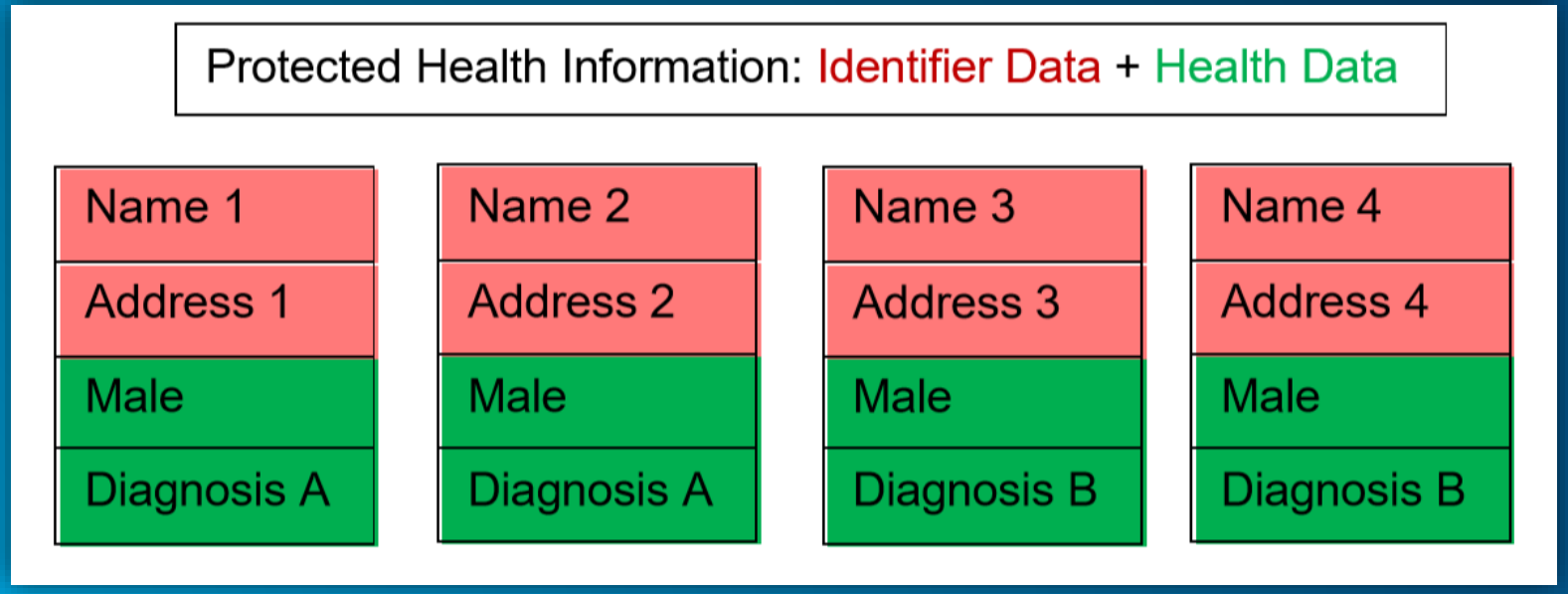
- Any individually identifiable health information including demographic information, that is created, used, disclosed, or received by a health care provider, health plan, or health care clearinghouse. PHI includes data about past, present, and future health conditions, health care provision, and health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

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

(HHS, 2015)

The Two Components of PHI

- PHI is record-level (individual) data that includes identifiers combined with health information. Aggregation involves grouping of information about individuals



Aggregation example: two male individuals with Diagnosis A and two male individuals with Diagnosis B

Who is affected by the Privacy Rule?

- The Privacy Rule applies only to covered entities. Many organizations that use, collect, access, and disclose individually identifiable health information will not be covered entities, and thus, will not have to comply with the Privacy Rule.
- The Privacy Rule does not apply to research; it applies to covered entities, which researchers may or may not be. The Rule 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.
- Not all health data is PHI!

Covered Entities

Covered entities are defined in the HIPAA rules as (1) health plans, (2) health care clearinghouses, and (3) health care providers who electronically transmit any health information in connection with transactions for which HHS has adopted standards. Generally, these transactions concern billing and payment for services or insurance coverage. For example, hospitals, academic medical centers, physicians, and other health care providers who electronically transmit claims transaction information directly or through an intermediary to a health plan are covered entities. Covered entities can be institutions, organizations, or persons.

Definitions of Covered Entity Types

- **Health Plan** – With certain exceptions, an individual or group plan that provides or pays the cost of medical care (as defined in section 2791(a)(2) of the PHS Act, 42 U.S.C. 300gg-91(a)(2)). The law specifically includes many types of organizations and government programs as health plans.
- **Health Care Clearinghouse** – A public or private entity, including a billing service, repricing company, community health management information system or community health information system, and “value added” networks and switches that either process or facilitate the processing of health information received from another entity in a nonstandard format or containing nonstandard data content into standard data elements or a standard transaction, or receive a standard transaction from another entity and process or facilitate the processing of health information into a nonstandard format or nonstandard data content for the receiving entity.

Definitions of Covered Entity Types

- **Health Care Provider** – A provider of services (as defined in section 1861(u) of the Act, 42 U.S.C. 1395x(u)), a provider of medical or health services (as defined in section 1861(s) of the Act, 42 U.S.C. 1395x(s)), and any other person or organization who furnishes, bills, or is paid for health care in the normal course of business.
- **Health Care** – Care, services, or supplies related to the health of an individual, including (1) preventive, diagnostic, therapeutic, rehabilitative, maintenance, or palliative care, and counseling, service, assessment, or procedure with respect to the physical or mental condition, or functional status, of an individual that affects the structure or function of the body; and (2) sale or dispensing of a drug, device, equipment, or other item in accordance with a prescription.

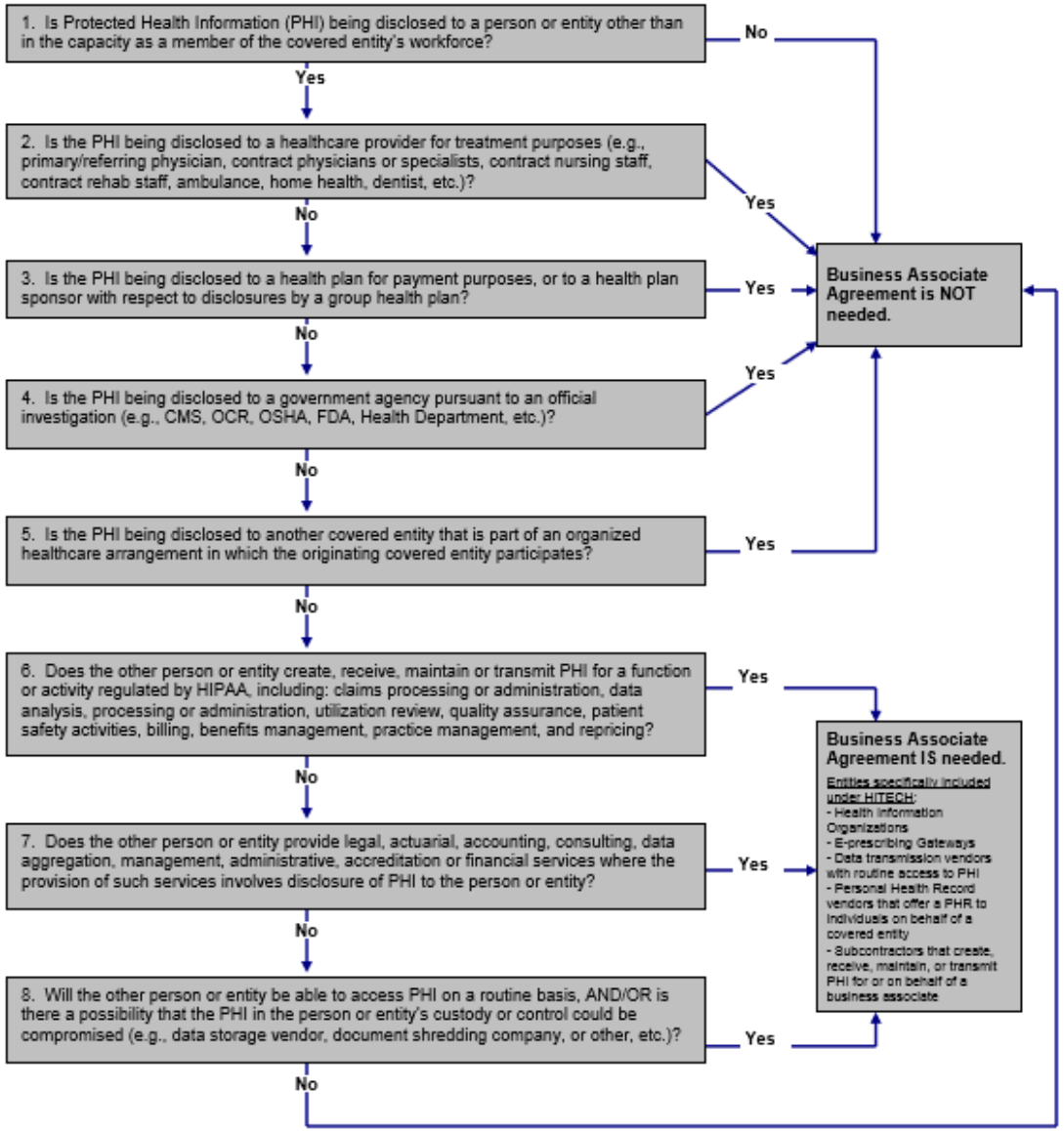
Geocoding - DOs

- If you are geocoding any data that is protected health information you must take care!
- You can geocode using your ArcGIS desktop software if you have reference data (like 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

Geocoding – DON’Ts

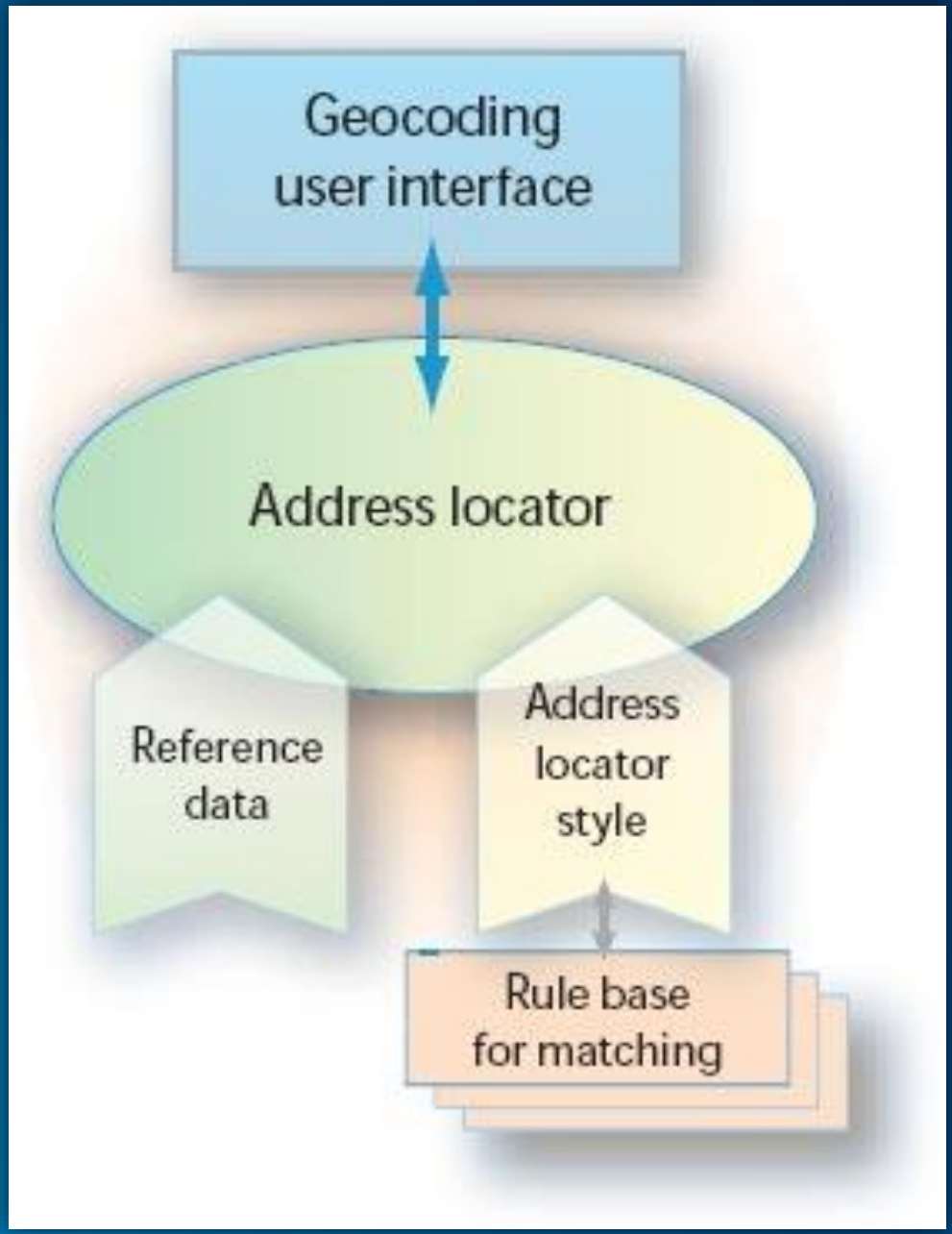
- You may NOT give your data to someone else to geocode for you if they are not listed on your IRB protocol or given permission to handle your PHI (such as a BAA)
- You may not geocode PHI in ArcGIS Online

HIPAA/HITECH
Business Associate Decision Tree

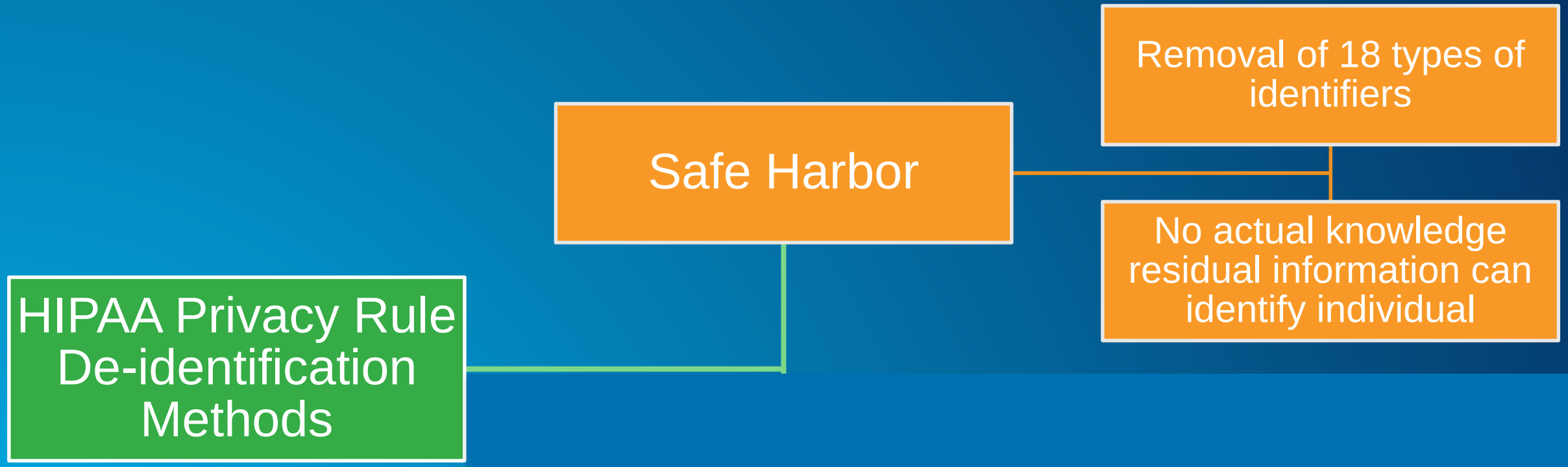


Geocoding Accuracy

- Potential to unintentionally obfuscate your results
  - Be sure you have high quality reference data
  - Adjust your matching rules to suit the situation



Data De-Identification Under HIPAA



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

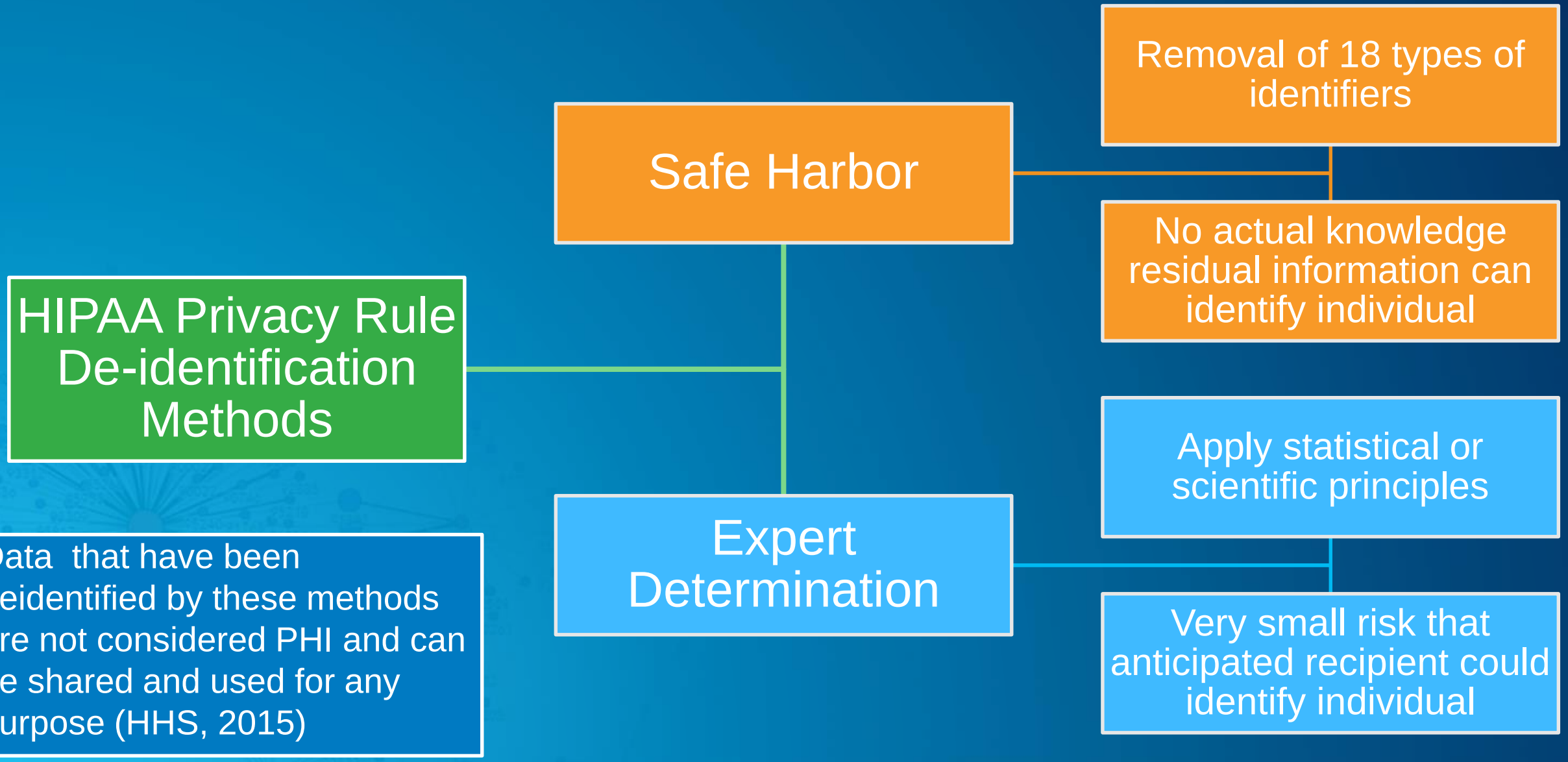
Geographies as defined by Safe Harbor

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:

(1) The geographic unit formed by combining all ZIP codes with the same three initial digits contains more than 20,000 people; and

(2) The initial three digits of a ZIP code for all such geographic units containing 20,000 or fewer people is changed to 000

Data De-Identification Under HIPAA



Mapping County-level Data May Violate HIPAA

- 10% of all U.S. counties have more tan 200,000 residents
- More than 40% of counties have fewer than 20,000 residents

County Population Variability in the U.S. and Four Sample States
(Source: 2015 ACS 5-Year Estimates)

State	Region	# of Counties	Min. County Population	Max. County Population	Mean County Population	Std. Deviation
California	West	58	1,131	10,038,388	662,439	1,442,050
Kansas	Midwest	105	1,224	566,814	27,552	76,464
New Jersey	Northeast	21	65,120	926,330	424,020	252,530
South Carolina	South	46	9,838	474,903	103,860	112,986
United States		**3,142**	**85**	**10,038,388**	**100,737**	**322,983**

Focus on the Expert Determination method

From Section 164.514(b)(1) of the Privacy Rule (HHS, 2015)

(1) A person with appropriate knowledge of an experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable:

i. Applying such principles and methods, determines that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information; and

ii. Documents the methods and results of the analysis that justify such determination.

EXPERT DE-IDENTIFICATION PROCESS

(Source: HHS, 2015)

1. Expert works with covered entity to determine appropriate statistical or scientific methods to mitigate risk of identification

2. Expert applies method to mitigate risk

3. Expert assesses risk

*Risk is **not** very small*

*Risk **is** very small*

Risk mitigation complete
Expert documents methods and results to justify determination

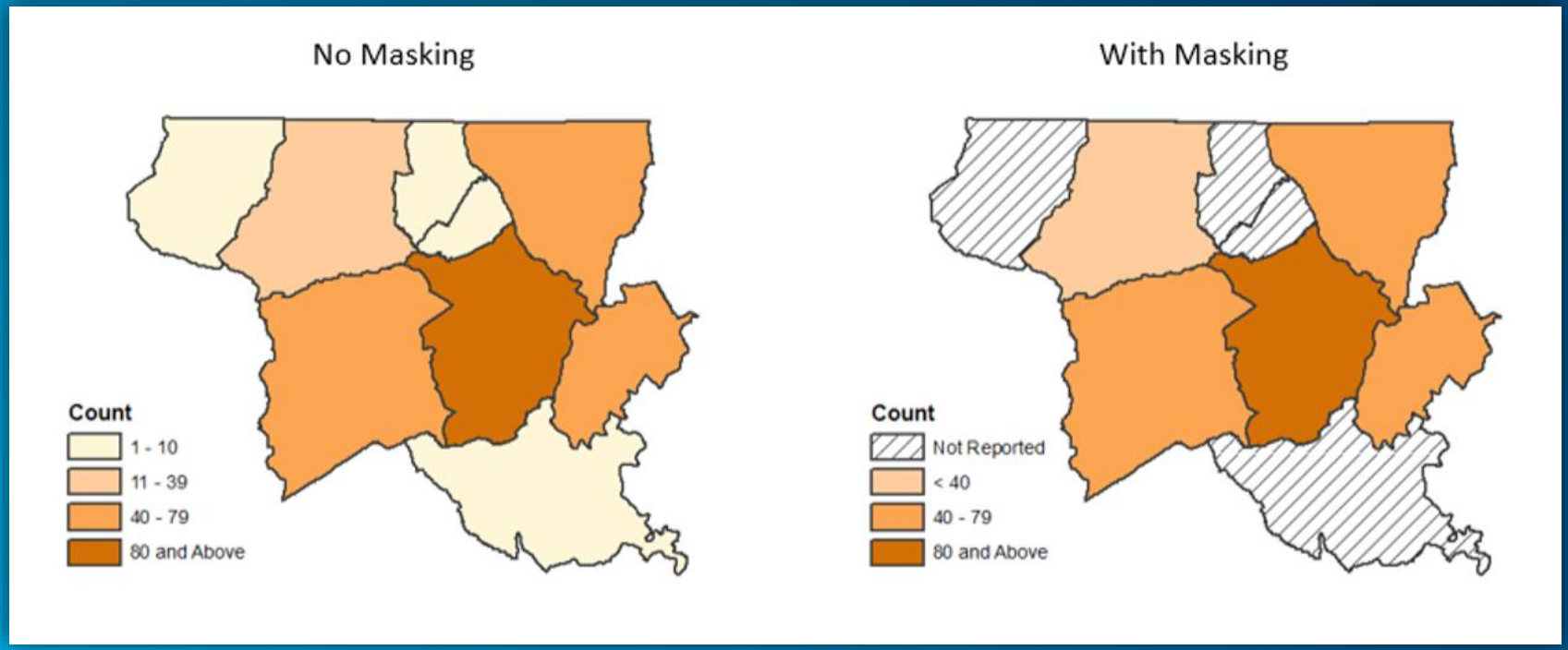
Examples & Methods

Numerator Rule:
Suppression of Small Cells – the experts…

- 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

Suppression of Small Cells

- Geographic masking to suppress cell counts ≤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 – the experts…

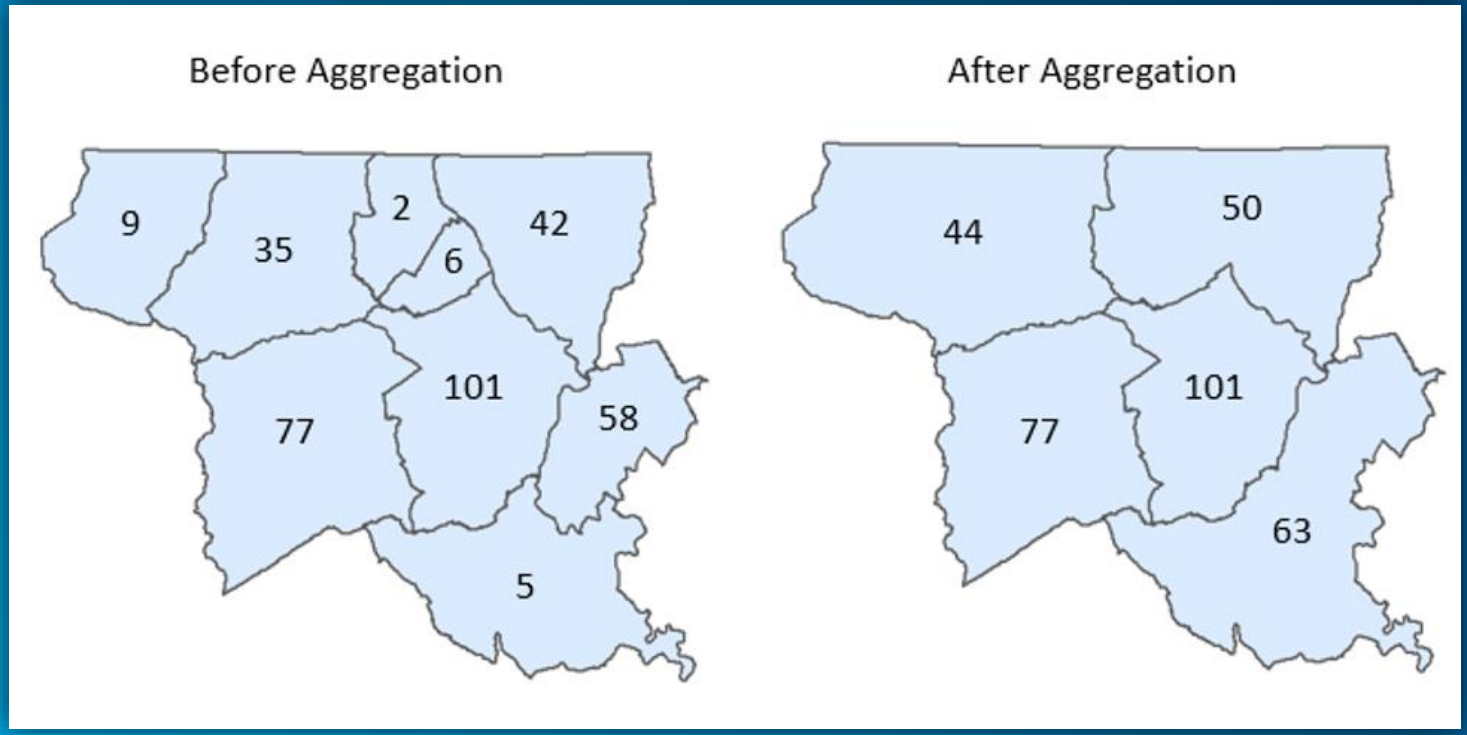
- National Center for Health Statistics – 250,000
- National Environmental Health Tracking Network – 100,000
- Maine Integrated Youth Health Survey – 5,000

Denominator Rule - Geography

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

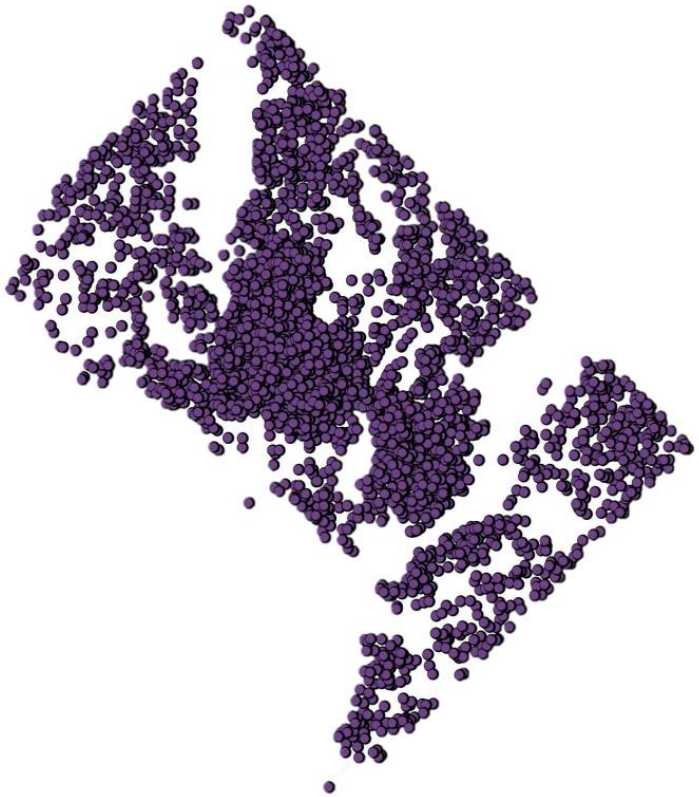
Aggregation

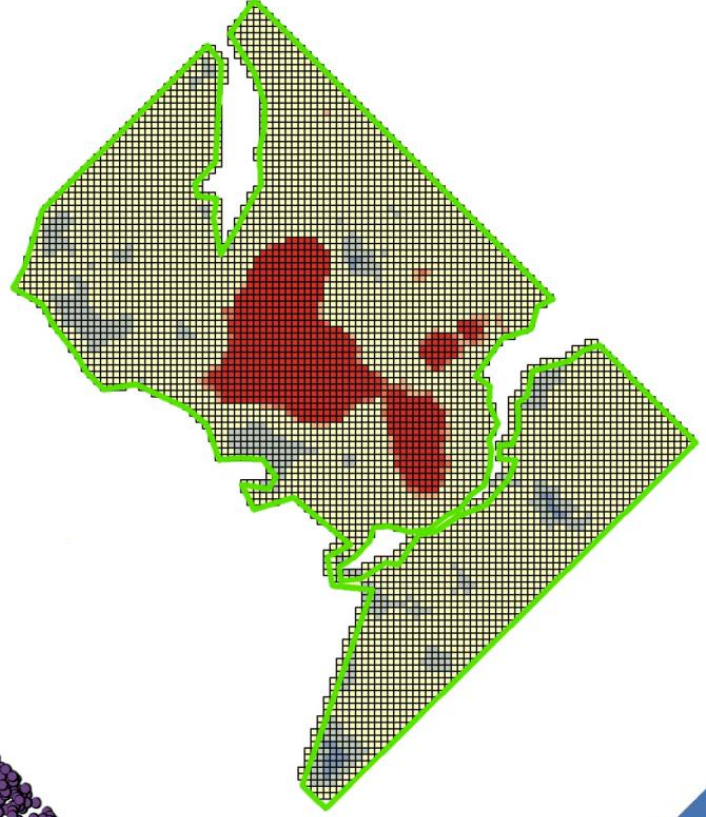
You can aggregate on any of your cells: age, racial/ethnic categories, time periods (months, years), geography



Methods: combine polygons, use alternate boundary types, manipulate latitude/longitude

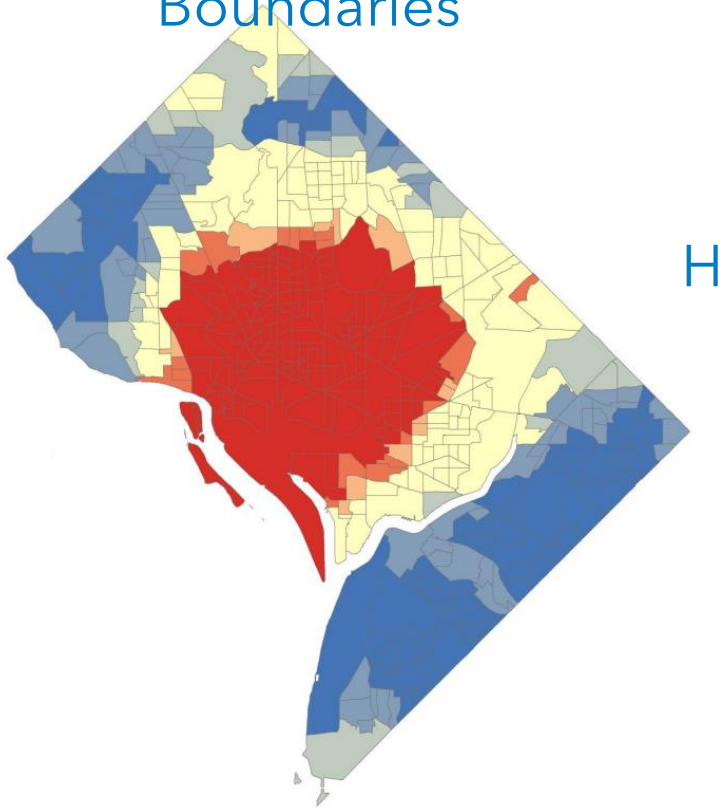
Raw Data

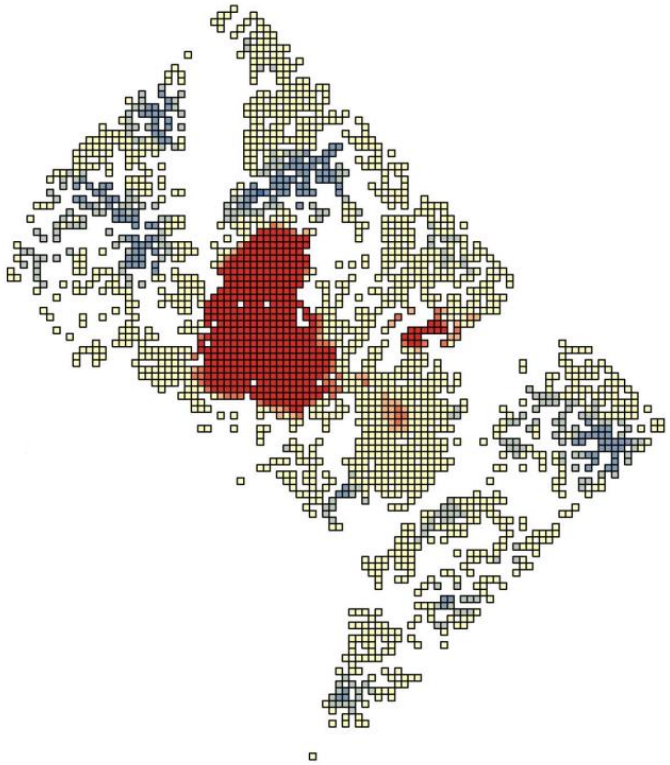




Fishnet grid

Political Boundaries

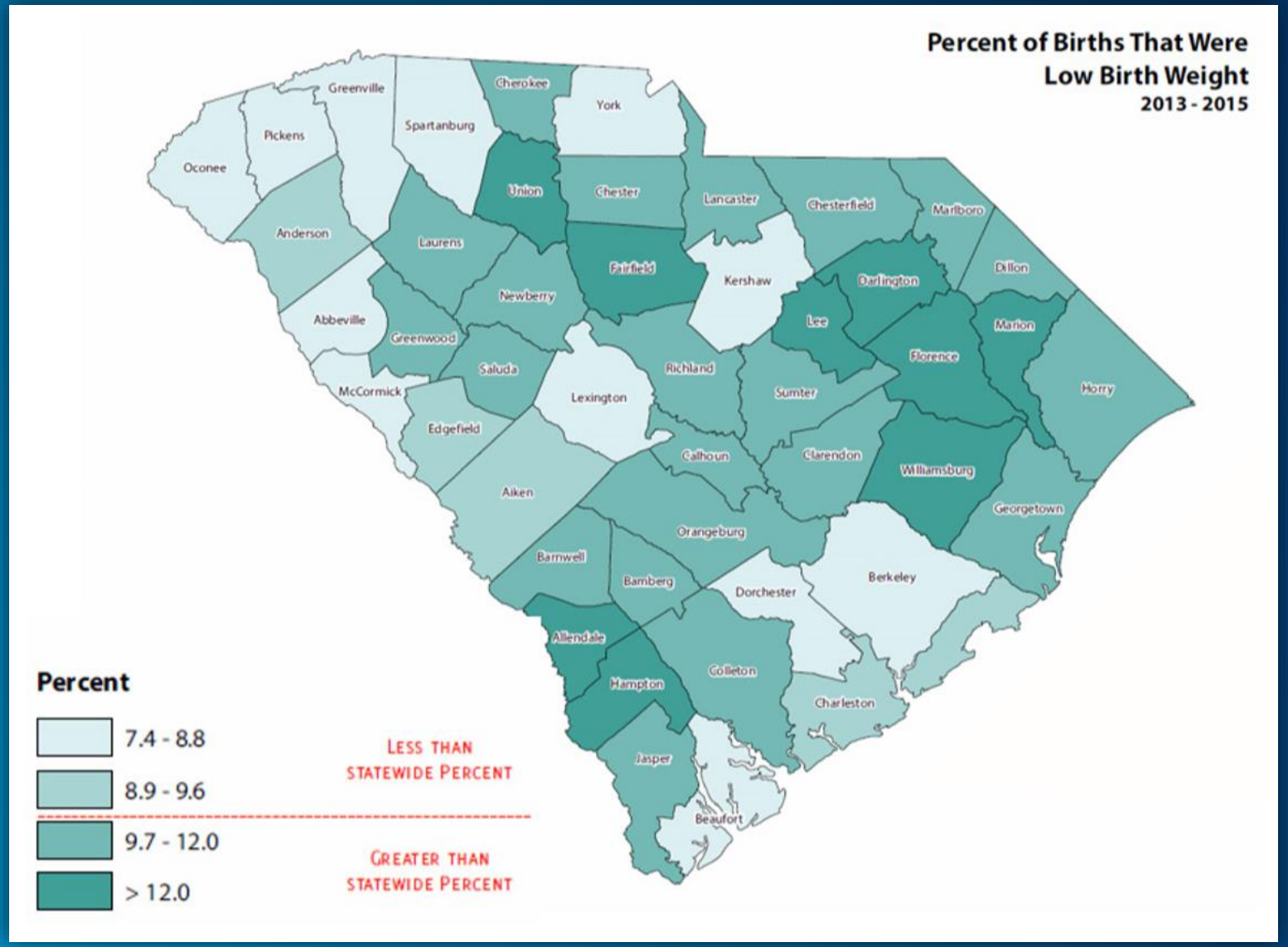




Hexagon grid

Blurring

- Replace specific data values with data ranges (e.g. mapped classes in choropleth maps)
- Transform continuous data into ordered categories (e.g. low, medium, high)
- Additional blurring achieved here by combining 3 years of data

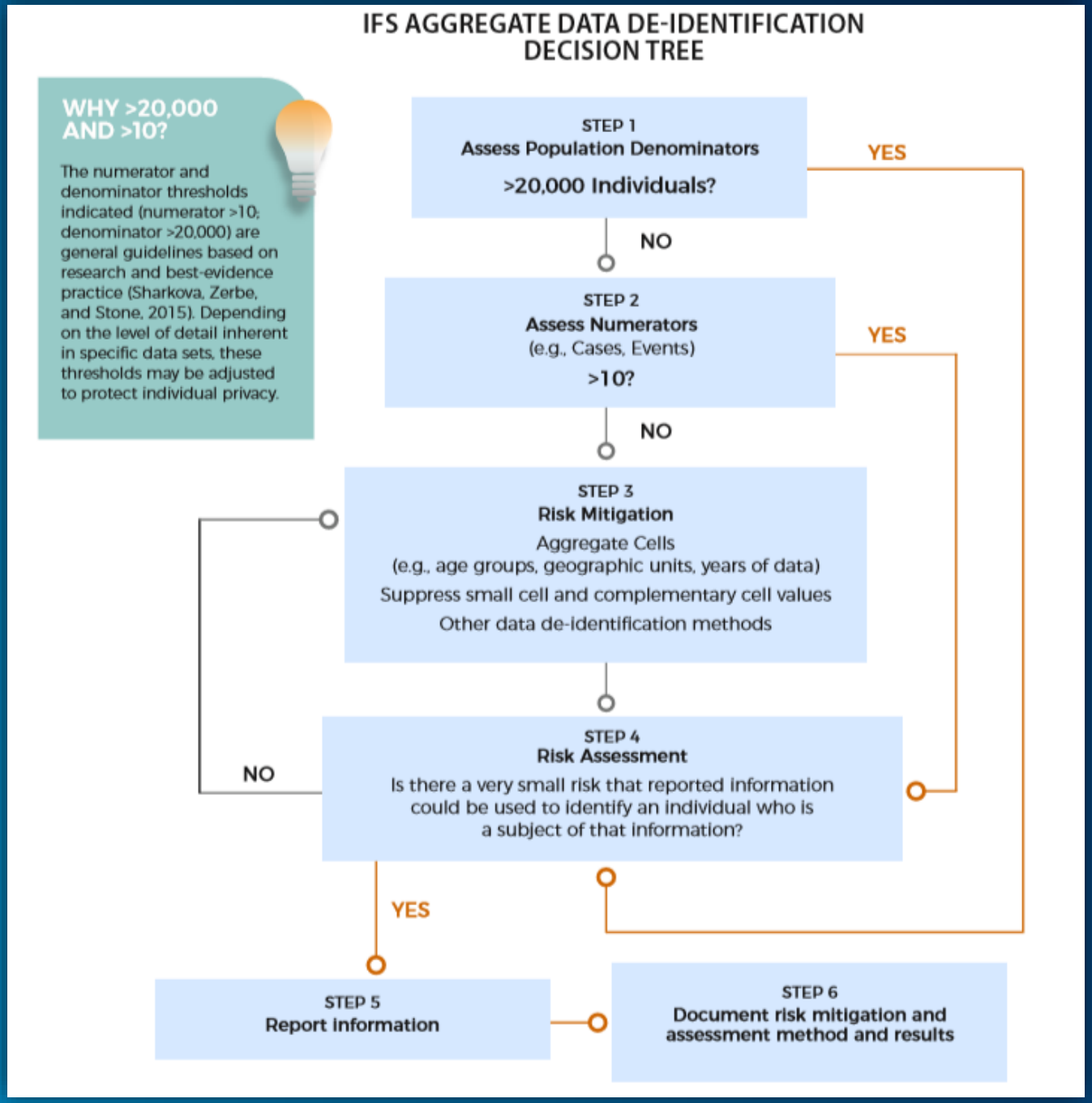


Perturbation

Randomly offsetting point data in a way that obscures actual address locations (in blue), while preserving the overall spatial pattern.

Institute for Families in Society (South Carolina) Decision Tree

Numerator OR Denominator rule is okay.



California Department of Healthcare Services (DHCS)

Assessment of risk for data release takes into account assessment of de-identification as well as legal, programmatic, and policy risks.

DCHS Decision Tree

- Must meet numerator AND denominator condition

OR

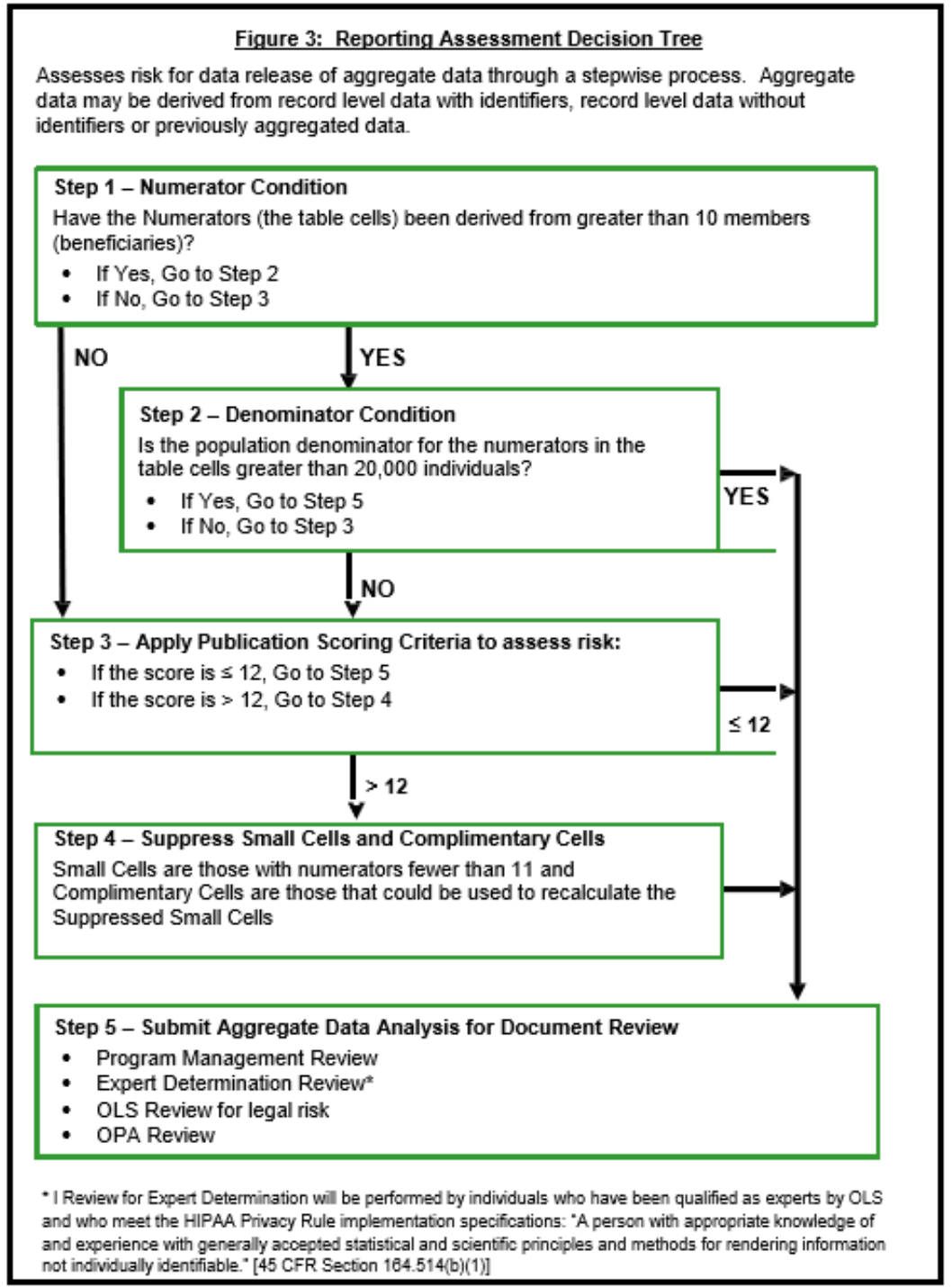
- Apply scoring criteria

IF Score ≤12 (go to risk assessment)

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

THEN

- Assess legal, policy and programmatic risk



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)

Variable	Characteristics	Score
Sex	Male or Female	+1
Age Range	>10-year age range	+2
	6-10 year age range	+3
	3-5 year age range	+5
	1-2 year age range	+7
Race Group	White, Asian, Black	+3
	Detailed Race	+5
Hispanic Ethnicity	yes or no	+2
	Detailed ethnicity	+3
Language Spoken	English, Spanish, Other Language	+2
Events	1000+ events in a specified population	+2
	100-999 events	+3
	11-99 events	+5
	<11 events	+8
Geography	State or geography with population >2,000,000	-5
	Population 560,001 - 2,000,000	-3
	Population 20,001 - 560,000	0
	Population ≤ 20,000	+5
Data Year	5 years aggregated	-5
	2-4 years aggregated	0
	1 year (e.g., 2001)	+3
	Bi-Annual	+4
	Quarterly	+5
	Monthly	+7

42 CFR Part 2

- Privacy law focused on drug and alcohol programs – also subject to HIPAA
- Both laws cover much of the same material
- There are points of difference – (notably 42 CRF Part 2 – has more prohibitions, consents and disclosures)
- Research: Researchers who receive patient identifying information are prohibited from redisclosing the patient-identifying information to anyone except back to the program [42 CFR § 2.52(b)].

References

- “Guidelines and Methods for De-Identifying Protected Health Information.” Institute for Families in Society, University of South Carolina, October 2017.
- “Department of Health Care Services Policies Regarding Public Records Act Requests and Public Aggregate Data Reporting.” California Department of Health Care Services, May 14, 2015.
- Zandbergen, Paul A. “Ensuring Confidentiality of Geocoding Health Data: Assessing Geographic Masking Strategies for Individual-Level Data.” *Advances in Medicine*, 29 April 2014.
- “HIPAA/HITECH Business Associate Decision Tree.” WEDI Privacy & Security Workgroup, Business Associate Sub-Workgroup. Updated: July 2013
- “To Whom Does the Privacy Rule Apply and Whom Will It Affect?” NIH. https://privacyruleandresearch.nih.gov/pr_06.asp


esri
THE
SCIENCE
OF
WHERE

spatialitics™ Health
POWERED BY esri

Navigating HIPAA in a Geospatial World

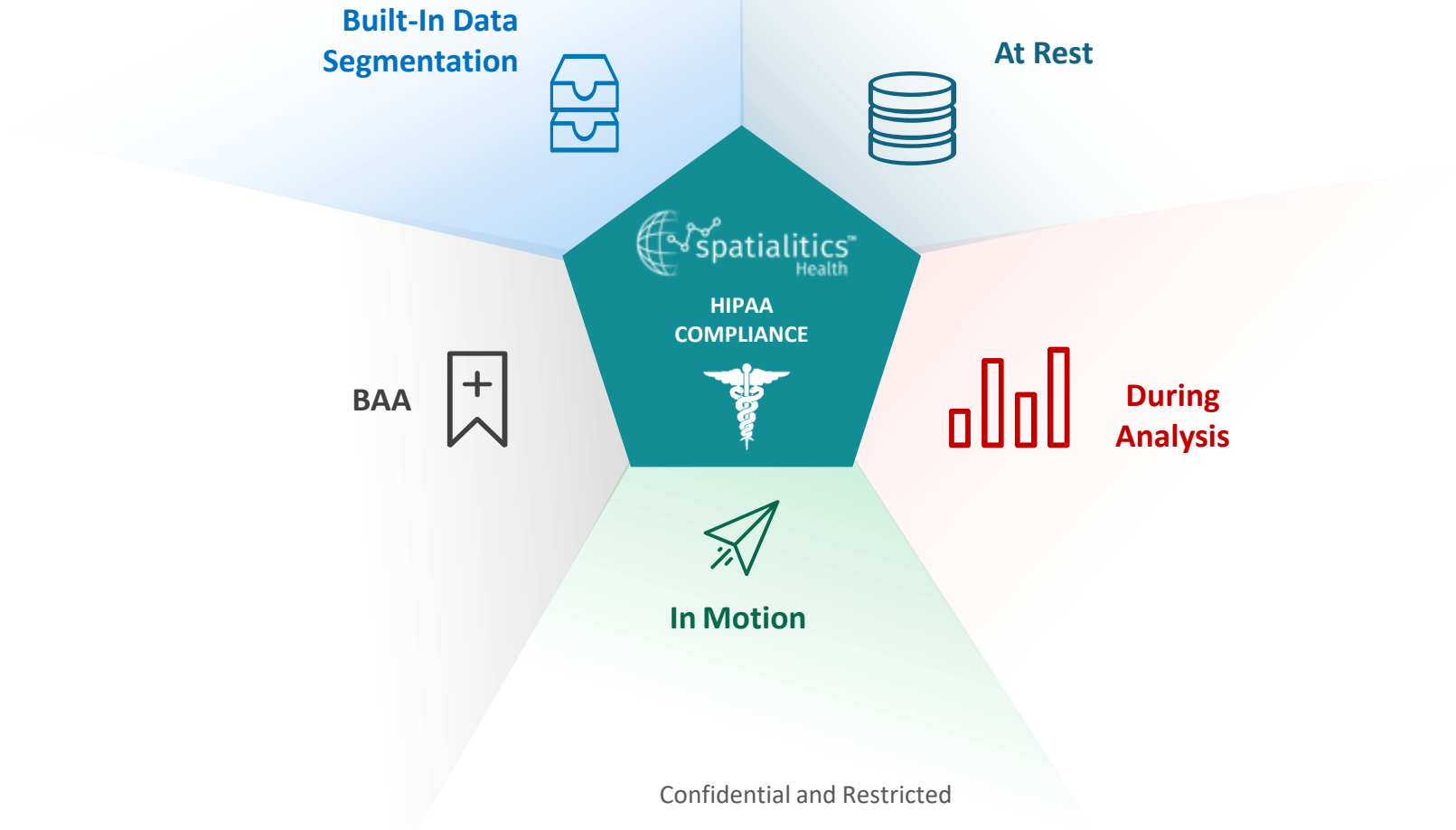
HIPAA Compliant, cloud based geocoding

Geocoding in a HIPAA Regulated World

- Compliance
  - Features
  - Security Layers
  - Operations
- Security Architecture
- Geocoding Options
  - Examples
  - Output / Results
- Benefits

Spatialitics Health Overview

Spatialitics™ Health is built with foundational HIPAA-Compliant data management policies and rules. Applications dealing with ePHI and PII adhere to HIPAA-Compliant Policies and Rules



GEOCODING HIPAA-COMPLIANT FEATURES

spatialitics™ Health
POWERED BY esri

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 processes, following protected data SOPs and policies

Regular Audit

Consistent auditing and testing for internal and external vulnerability assessment

SPATIALITICS HIPAA-COMPLIANT OPERATIONS

spatialitics™ Health
POWERED BY esri

HIPAA Operations Approach

Rules for personnel involved with the monitoring, maintaining, supporting and access of data

- Environmental safeguards in place to protect PHI
- Usage & sharing of PHI Highly Regulated
- Agreements in place with any service providers involved
- Procedures in place to limit access to PHI
- Training program for employees about protecting PHI

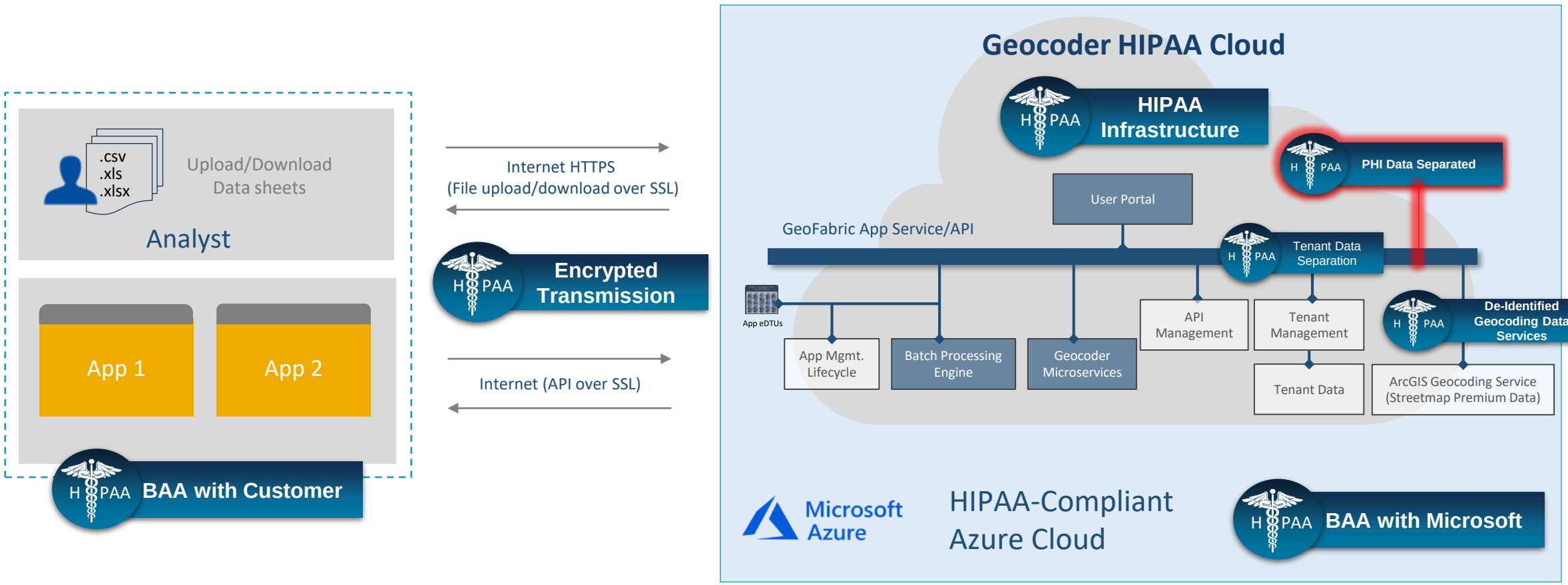
HIPAA Operations Policies

Spatialitics Operations Policies

- All PHI data will only be accessed and managed by authorized personnel
- BAAs in place to help Customers and Business Partners understand policy & procedures
- Strict compliance SOPs and “Rules and Regulation” materials are provided to all personnel for handling PHI and HIPAA Services
- Authorized personnel undergo yearly HIPAA and Data Security Training
- All HIPAA PHI Data is isolated and partitioned with “No-Cross-Contamination” Policy, and supporting infrastructure and storage

Sources: Aug 2018 - Microsoft BAA, HIPAA Compliance and Enforcement & De-Identification @ HHS.gov

HIPAA-COMPLIANT ARCHITECTURE & DEPLOYMENT MODEL



Thank You

Nicholas Poorte
Operations Manager, Client Officer
nicholas.poorte@spatialitics.com
(m) 1 801 882 9060

Vince A. Rosales
Vice President
vince.rosales@spatialitics.com
(m) 1 303 884 4268

www.spatialitics.com

spatialitics™

Unleash the Power of Spatial Analytics