## **Washington State**

## Department of Social and Health Services

# **Geospatial Data Confidentiality Guidelines**

Last Revised: May 4, 2015

Irina Sharkova, PhD Jane Zerbe Howard Stone, JD, LLM

The Geospatial Data Confidentiality Workgroup, DSHS Privacy Officer, and DSHS GIS Steering Committee served as advisors and critically reviewed the document.

## **Table of Contents**

1.	Introduction: Why the agency-wide guidelines are needed	1
2.	Confidentiality Laws	
3.	Requirements for Disclosure and Use of DSHS Data	2
4.	DSHS Data Categories and Methods of Handling Geospatial Information	3
	Category 1 – Public Information	3
	Category 2 – Sensitive Information	4
	Category 3 – Confidential Information	5
	Category 4 – Confidential Information Requiring Special Handling	5
5.	Special Case: Emergency Response	8
6.	Remote Geospatial Services	9
7.	Glossary	10
8.	References	11
9.	Other Resources	12

## 1. Introduction: Why the agency-wide guidelines are needed

The need for the present guidelines is determined by (1) the continuing growth in the amount and complexity of the data that Department of Social and Health Services (DSHS) collects about its clients and service providers, (2) the increased complexity of federal, state and agency regulations governing collection and use of personal information, and (3) the implementation of the DSHS Enterprise GIS which brings a capacity to access client and other data to produce maps and other decision making information.

**DSHS Data**: DSHS serves every third resident of Washington, or 2.3 million people a year. The organization possesses vast amounts of administrative data about its clients, which includes services received, addresses, demographic, and other client information. Many data sources go as far back as 25 to 30 years; sometimes, they are combined into linked data systems to allow the agency to ask complex questions about client services and outcomes [3]. Additionally, DSHS collects information about its administrative offices, owned and leased facilities, service providers, and employees. This information is used in daily operations and long-term planning, and is often geographic in nature. The richness and specificity of DSHS data will continue to grow.

**Laws**<sup>1</sup>: The federal and Washington State regulations have long mandated privacy of individually identifiable information [1-2, 5-11, 13-14]. Recent federal rules bring into sharp focus the responsibility of data stewards to keep personal information secure and prevent risk of identification or unauthorized access.

**Enterprise GIS**: The implementation of the DSHS Enterprise GIS has led to a substantial increase in the access to GIS tools and geographic data across the agency. This opens opportunities to apply GIS technology to the data DSHS programs collect to assist day-to-day activities, long-range planning, and decision making. It also increases the urgency to develop a common understanding of how geographic data should be handled by DSHS GIS users given the laws safeguarding privacy of personal data.

**Does Your Data Need to be De-identified?** Here are some questions to help determine whether you need to protect your geospatial data using one of the methods described in these guidelines.

- Does your dataset contain information about individual clients or employees such as their name, Social Security number, date of birth, gender, home address, income, treatment of a medical condition, or other Protected Health Information?
- Does your data consist exclusively of clients who were in alcohol or substance abuse treatment, or mental health programs?
- Are you working with data about rare events or conditions?

<sup>&</sup>lt;sup>1</sup> The meaning of the term Law is intended broadly, and may include federal and state constitutional provisions, Congressional and other legislative enactment (public and private laws), federal and state statutes (codified enactments), federal and state regulations (often agency specific), local government (including Indian tribal) rules and ordinances, international laws and judicial or court rules cases or decisions. Additionally, many agreements have an effect of and are enforced as law, including contracts to share data or provide products and services, nondisclosure agreements, treaties or similar agreements between different entities or between entities and individuals.

- Do you have high precision location points for employees or clients (i.e., roof top or street-level address locations)?
- Do you need to display your client data at a fine geographic resolution (by Census block or zip code) or create a map of a neighborhood or a small town?

If you answered yes to any of these or similar types of questions, you may need to apply a special method for handling your geospatial information, or consult with someone with relevant scientific and/or de-identification expertise. [13]

## 2. Confidentiality Laws

Laws and regulations are adopted to reduce risks of disclosure of confidential information, especially as technology, data collection methods, analysis, and distribution become more sophisticated, faster, and cheaper. The necessity, principles, and methods to protect confidential data are defined by multiple federal and Washington State laws [1-2, 5-11, 13-14], and apply to any data, including geospatial data. A few select laws are provided below.

#### Federal level:

- Confidentiality of Alcohol and Drug Abuse Patient Records [1];
- Family Educational Rights and Privacy Act (FERPA) [2];
- Health Insurance Portability and Accountability Act of 1996 (HIPAA) through Final Rule: Modifications to the HIPAA Privacy, Security, Enforcement, and Breach Notification Rules under HITECH Act (January 25, 2013) [11, 13-14]
  - o HIPAA Privacy Rule: Title 45 of the U.S. Code of Federal Regulations, Part 164.

### Washington State level:

- Revised Code of Washington (RCW) Chapter 40.24. Address Confidentiality for Victims of Domestic Violence, Sexual Assault, and Stalking [5];
- RCW Chapter 42.48. Release of Records for Research [6];
- RCW 42.56.590. Personal information Notice of security breaches [7];
- RCW Chapter 70.02. Medical Records Health Care Information Access and Disclosure [8];
- RCW, Chapter 74.04, section 74.04.060(1)(a) (providing exceptions to confidentiality of public assistance records) [10].

## 3. Requirements for Disclosure and Use of DSHS Data

#### **Important Definitions**

#### **Identifiable Client Information:**

Client information that 1) reveals or can likely be associated with a client's identity, including for example explicit client identifiers such as name, SSN, date of birth, beneficiary number, client identification number, medical record number, or 2) could be used by the data recipient in combination with other reasonably available information to identify a client.

#### **Protected Health Information (PHI):**

Information about health status, provision of health care, or payment for health care that can be linked to an individual. This includes identifiable client information. Not all client records are PHI. Client information that is PHI may include variables that are not explicit identifiers.

**No** <u>identifiable personal (client/employee) information</u>, <u>PHI</u>, or <u>key that links geocodes to identifiable client information or PHI</u> may be disclosed **except** when:

- 1. In accordance with applicable Washington State laws (e.g., Revised Code of Washington, Chapter 74.04, section 74.04.060(1)(a) providing exceptions to confidentiality of public assistance records, including disclosures for purposes directly connected with public assistance program administration).
- 2. In accordance with applicable federal laws (e.g., Title 45 of the U.S. Code of Federal Regulations, Part 164 (referred to as the "HIPAA Privacy Rule")). For example, PHI may be disclosed to a DSHS health care component or business associate for treatment, payment or health care operations. PHI may also be disclosed if de-identified or if an appropriate expert determines and documents that the risk is very small that PHI could be used, alone or in combination with other reasonably available information, by a data recipient to identify any client. If not otherwise expressly permitted or accepted, PHI may only be disclosed pursuant to clients' written authorization.

A **De-identification Expert** with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable should be designated, with responsibility for making assessments of the risk of identification of information, as may be needed.<sup>2</sup> Any concerns or doubts about the identifiability of client information or the need to protect client information should be brought to the attention of the **De-identification Expert** specifically designated by your program area.

## 4. DSHS Data Categories and Methods of Handling Geospatial Information

DSHS and Washington State<sup>3</sup> classify data into 4 categories [ $\underline{16}$ ,  $\underline{17}$ ] according to the level of protection needed.

## Category 1 – Public Information

Public information is information that can be released to the public. It does not need protection from unauthorized disclosure, but does need protection from unauthorized change that may mislead the public or embarrass DSHS. [10]

A lot of geographic data are public. They include GIS data made available by federal, state, and local agencies.

<sup>&</sup>lt;sup>2</sup> [13] OCR Guidance Regarding Methods for De-identification of Protected Health Information. Pp 10-22.

<sup>&</sup>lt;sup>3</sup> The DSHS categories are based on the Office of Chief Information Officer's (OCIO) data classifications categories.

## **GIS Data Examples:**

- DSHS regional boundaries,
- Community Service Office locations,
- Licensed residential care facility locations,
- Geographic boundaries of Census tabulation areas and other Census data,
- City limits, US highways, state routes, and hydrography data from OFM, WSDOT, DNR, other state agencies.

## **Method of Handling:**

When manipulating or mapping Category 1 data in isolation, there is usually no concern about confidentiality. However, public geospatial data may become sensitive and require special handling when combined with non-public (sensitive or confidential) data. An example is when geocoded locations of persons with a particular type of cancer are aggregated in a GIS overlay to small areas such as Census blocks. If just a few people reside in a given Census block, an observer can determine who the person(s) with cancer may be.

## Category 2 - Sensitive Information

Sensitive information is not specifically protected by law, but should be limited to official use only, and protected against unauthorized access. [10]

a) Computer system documentation that is not classified as Confidential should be classified as Sensitive.

### **GIS Data Examples:**

- Licensed foster homes locations,
- In-home child care locations.

## Method of Handling:

The best methods to handle sensitive data depend on the type of data, the end product and the end-user. GIS specialists should become familiar with Category 4 methods of handling confidential information and use their best judgment in trying to limit disclosure of Category 2 individual-level data. For example, it may not be advisable to release for public use (outside DSHS) a detailed neighborhood map with precise locations and names of inhome child care providers. When developing public- facing mapping applications using sensitive data, it may be helpful to prohibit zooming into a very detailed (large) scale for geographic layers with sensitive information.

## Category 3 – Confidential Information

Confidential information is information that is specifically protected by law. It generally includes:

- a) Personal information about individual clients, regardless of how that information is obtained;
- b) Information concerning employee payroll and personnel records;
- Source code of certain applications programs that could jeopardize the integrity of department data or result in fraud or unauthorized disclosure of information if unauthorized modification occurred.

## **DSHS GIS Data examples:**

- DSHS employee residence locations,
- DVR client locations.
- Foster child locations.

#### Method of Handling:

Apply Category 4 methods of handling when manipulating or mapping Category 3 data.

## **Category 4 – Confidential Information Requiring Special Handling**

Confidential information requiring special handling is information for which:

- Especially strict handling requirements are dictated, e.g. by statutes, regulations, or agreements; or
- b) Serious consequences could arise from unauthorized disclosure, ranging from life threatening to legal sanctions.

Examples of confidential information requiring special handling include:

- a) Protected Health Information (PHI), as defined at Administrative Policy 5.01 Privacy Policy -- Safeguarding Confidential Information, and by the HIPAA Security Rule;
- b) Information that identifies a person as being or ever having been a client of an alcohol or substance abuse treatment, or mental health program;
- c) Federal wage data;
- d) Location of an abused spouse or addresses of those in the Secretary of State Address Confidentiality Program. [5]

Confidential information is usually associated with a person. However, some critical facilities, infrastructure, or natural resources may be confidential for safety and security reasons.

#### **DSHS GIS Data examples:**

- Locations of DSHS clients on dialysis,
- Behavioral Health and Service Integration Administration (BSHIA) client locations.
- Initial 3 digits of a ZIP code for de-identified BSHIA client records where the sum of

Census population for that 3-digit ZIP area contains 20,000 people or fewer.

## Method of Handling:

No standard rules exist which can be applied to any situation. The following are best practices summarized from published research and other guidelines.

### 1. Counts and numerators for rates

- a) No cell should contain all cases of any row or column. "Cell" could be a geographic area such as Census Tract, a population subgroup such as persons ages 0 to 17 with a disability, or both. "Cases" can be people (clients, services providers, staff), events (visits to a clinic), or other objects (count of licenses).
- b) Cell size: Fewer than 5 cases need to be suppressed, possibly as high as 10.
- c) Users should not be able to calculate the value of the suppressed cell by subtraction or other data manipulation.

#### 2. Population size of the area-unit of analysis and denominators for the rates

- a) Currently, thresholds of 20,000+ and 100,000+ are used when releasing deidentified survey micro-data, that is, individual-level data.
- b) Recent research calls for use of different population thresholds depending on data sensitivity, dimensions of the table, demographic and geographic detail, etc.
- c) US Census Bureau uses the threshold of at least 50 survey respondents when releasing American Community Survey tables which translates into the population size of a reporting area of 3,200 persons or more (roughly an equivalent of a Census tract).

## 3. For counts below threshold (small numbers), consider using aggregation to ensure that the number exceeds the minimum cell-size:

- a) Geographic aggregation such as coarsening spatial resolution: presenting data by county instead of Census tract, or by Census tract instead of Census block.
- b) Aggregation over time: combining several years of data (add 3 or 5 years of data instead of using 1 year).
- c) Categorizations of the data (reducing dimensions of the attribute table): aggregating sensitive information into broad categories (using age groups instead of individual years of age; using "any minority" category instead of individual racialethnic categories; using a general medical condition instead of a specific ICD10 code).

NOTE: Any references, in this manual, to "confidential information" apply to both category 3, "confidential information", and to category 4, "confidential information requiring special handling" unless stated otherwise.

Data stewards should examine their datasets with regard to DSHS' four data categories, specific uses, and potential risks to determine an appropriate method that protects confidentiality when releasing data in tables, maps, charts, or online query systems. There are several methods which can be applied to any data and some that are specific to geospatial data.

## **Common Data Protection Methods (Any Data)**

**Data** Generalization, grouping, coarsening, combining into larger categories

**Aggregation:** before its publication in tables, maps, or by other means.

**Suppression of** Unacceptable data (small Ns) in cells are suppressed (can be labeled as

**Small Numbers:** "SN" – "Small numbers"). When this is done, it is necessary also to

suppress other cells in the table to prevent determination of the unacceptable cell figure through subtraction. It is usually necessary to suppress four cells in a cross tabulation in order to avoid disclosure

through one cell.4

Modification of Data-swapping, addition of noise; applied before calculating summary

Individual-Level tables.

Data:

**Topping Off** Such as for very old age or very high income.

Ranged Values:

### **Geospatial Data Protection Methods**

Coarsening Geographic Scale of Analysis (Generalize data or "zoom out.")





After



**Spatial aggregation** (Assign values to a larger geographic area.)

Before



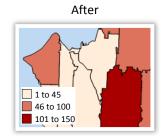
After



<sup>&</sup>lt;sup>4</sup> [12] National Center for Health Statistics Staff Manual on Confidentiality (pp. 16-17). See also [15] Washington State Department of Health (2012). Guidelines for Working with Small Numbers.

## **Display Ranges of Values** (Hide actual values.)



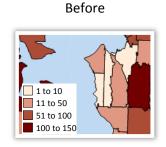


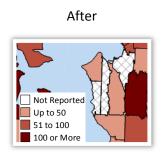
## Dithering or Random Offsetting (Hide actual locations.)





## Masking Geographic Areas (Hide small numbers.)





## 5. Special Case: Emergency Response

In the case of a natural or man-made disaster or other emergency, Protected Health Information may be disclosed when necessary to prevent or lessen a serious and imminent threat to a client's health or safety, and disclosure is to persons reasonably able to prevent or lessen the threat<sup>5</sup>. Additionally, RCW Chapter 70.02.050 stipulates that health care information may be disclosed if disclosure will avoid or minimize an imminent danger to the health or safety of the patient or any

<sup>&</sup>lt;sup>5</sup> 45 CFR 164.512(j)(1)(i), Standard: Uses and disclosures to avert a serious threat to health or safety. 45 CFR 164.512 - Uses and disclosures for which an authorization or opportunity to agree or object is not required. Online at <a href="http://www.law.cornell.edu/cfr/text/45/164.512">http://www.law.cornell.edu/cfr/text/45/164.512</a>, accessed 12-29-2014.

other individual [9]. In DSHS, disclosure of confidential information for disaster response is coordinated by the DSHS Office of Emergency Management.

## 6. Remote Geospatial Services

Certain IT security measures should be taken when using remote geospatial services, especially with regard to addresses. Address information by itself without a connection or nexus to any type of client identifier or status, independent of whether the information is covered by HIPAA is not protected health information according to Kathryn Ruckle, J.D., the DSHS Privacy Officer (2015).

The use of remote address services, such as USPS address correction and address geocoding, must meet DSHS security and confidentiality requirements:

## **Address Service Security and Confidentiality Requirements**

- Address must be transferred alone without a nexus of client information
   (e.g., No identifiable personal (client/employee) information, PHI, or key that links to identifiable client information or PHI).
- Network connection to address service must be encrypted (i.e., HTTPS).
- Origination of the address cannot be retained by the address service (i.e., from DSHS).
- Address service must undergo IT security review by CTS/DES and meet DSHS IT security requirements, which includes the Open Web Application Security Project's (OWASP) Top 10 list [4].

## 7. Glossary

#### **De-identification Expert:**

An individual with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable. No specific professional degree or certification program designates who is an expert at rendering personal information de-identified. Relevant expertise may be gained through various routes of education and experience. Experts may be found in statistical, mathematical, or other scientific domains. A de-identification expert at DSHS would typically be a subject matter expert with research and statistical training that has been designated by its division or program area to serve in this role. [13]

#### Geocoder:

Software or (web) service that helps in the geocoding process.

### **Geocoding:**

The process of finding associated geographic coordinates (often expressed as latitude and longitude) from other geographic data, such as street addresses, or ZIP codes (postal codes). With geographic coordinates the features can be mapped and entered into Geographic Information Systems. To assign a street address to a location.

#### GIS:

An acronym for *geospatial* or *geographic information system*. A collection of geospatial data, technology, and people used to analyze and portray geospatial information and relationships.

#### Geographic data:

Data associated with a location on Earth. It can be modeled in GIS as points, lines, polygons (areas), grids/surfaces; it includes attributes (characteristics of points, areas, etc.). Examples are: county boundaries, river and street networks, residential address points, and land elevation.

#### Geospatial:

Pertaining to the geographic location and characteristics of natural or constructed features and boundaries on, above, or below the earth's surface; especially referring to data that is geographic and spatial in nature. If the question contains a "where," geospatial tools will probably be the best way to answer it.

### **Geospatial Data or GIS Data:**

Data acquired, processed, managed, analyzed, displayed or shared/served using Geographic Information Systems such as Esri's ArcGIS or open-source GIS.

#### **Identifiable Personal Information:**

Client/employee information that 1) reveals or can likely be associated with a client's identity, including for example explicit client identifiers such as name, SSN, date of birth, beneficiary number, medical record number, or 2) could be used by the data recipient in combination with other reasonably available information to identify a client.

#### **Protected Health Information (PHI):**

Information about health status, provision of health care, or payment for health care that can be linked to an individual. This includes identifiable client information. Not all client records are PHI. Client information that is PHI may include variables that are not explicit identifiers.

### 8. References

- Confidentiality of Alcohol and Drug Abuse Patient Records, Title 42 Code of Federal Regulations Part 2 (E-CFR current as of April 11, 2014). <a href="http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr;sid=af45a7480ecfb95bc813ab4bbd37fb5b;rgn=div5;view=text;node=42%3A1.0.1.1.2;idno=42;cc=ecfr.">http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr;sid=af45a7480ecfb95bc813ab4bbd37fb5b;rgn=div5;view=text;node=42%3A1.0.1.1.2;idno=42;cc=ecfr.</a>
- Family Educational Rights and Privacy, Title 34 Code of Federal Regulations Part 99 (E-CFR current as of April 11, 2014). <a href="http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&sid=11975031b82001bed902b3e73f33e604&rgn=div5&view=text&node=34:1.1.1.1.33&idno=34.">http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&sid=11975031b82001bed902b3e73f33e604&rgn=div5&view=text&node=34:1.1.1.1.33&idno=34.</a>
- 3. Mancuso D. DSHS Integrated Client Database. Olympia, WA: Washington State Department of Social and Health Services Research and Data Analysis Division, 2014. RDA Report Number 11.205 (Accessed December 30, 2014, at <a href="http://www.dshs.wa.gov/sites/default/files/SESA/rda/documents/research-11-205.pdf">http://www.dshs.wa.gov/sites/default/files/SESA/rda/documents/research-11-205.pdf</a>).
- Open Web Application Security Project. OWASP Top 10 2013, The Ten Most Critical Web Application Security Risks. 2003-2013. <a href="http://owasptop10.googlecode.com/files/OWASP%20Top%2010%20-%202013.pdf">http://owasptop10.googlecode.com/files/OWASP%20Top%2010%20-%202013.pdf</a>.
- 5. Revised Code of Washington, Chapter 40.24 RCW. Address Confidentiality for Victims of Domestic Violence, Sexual Assault, and Stalking. http://apps.leg.wa.gov/RCW/default.aspx?cite=40.24.
- 6. Revised Code of Washington Chapter 42.48 RCW. Release of Records for Research. http://apps.leg.wa.gov/RCW/default.aspx?cite=42.48.
- 7. Revised Code of Washington, Chapter 42.56.590 RCW. Personal information Notice of Security Breaches. http://app.leg.wa.gov/RCW/default.aspx?cite=42.56.590.
- 8. Revised Code of Washington, Chapter 70.02 RCW. Medical Records Health Care Information Access and Disclosure. http://apps.leg.wa.gov/RCW/default.aspx?cite=70.02.
- 9. Revised Code of Washington, Chapter 70.02.050 RCW. Disclosure without patient's authorization Need-to-know basis. http://apps.leg.wa.gov/rcw/default.aspx?cite=70.02.050.
- 10. Revised Code of Washington, Chapter 74.04.060 RCW. Records, confidential Exceptions Penalty. http://apps.leg.wa.gov/rcw/default.aspx?cite=74.04.060.
- 11. U.S. Department of Health and Human Services (2013). 45 CFR Parts 160 and 164, Modifications to the HIPAA Privacy, Security, Enforcement, and Breach Notification Rules Under the Health Information Technology for Economic and Clinical Health Act and the Genetic Information Nondiscrimination Act; Other Modifications to the HIPAA Rules; Final Rule. Federal Register, Vol. 78, No. 17, January 25, 2013, Rules and Regulations. http://www.gpo.gov/fdsys/pkg/FR-2013-01-25/pdf/2013-01073.pdf.
- 12. U.S. Department of Health and Human Services, National Center for Health Statistics (2004). *NCHS Staff Manual on Confidentiality*. Washington, DC. <a href="http://www.cdc.gov/nchs/data/misc/staffmanual2004.pdf">http://www.cdc.gov/nchs/data/misc/staffmanual2004.pdf</a>.
- 13. U.S. Department of Health & Human Services, Office for Civil Rights (2012). *Guidance Regarding Methods* for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. <a href="http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/De-identification/guidance.html">http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/De-identification/guidance.html</a>.

- 14. U.S. Department of Health & Human Services, Office for Civil Rights (2013). HIPAA Administrative Simplification, Regulation Text 45 CFR Parts 160, 162, and 164 (Unofficial Version, as amended through March 26, 2013). http://www.hhs.gov/ocr/privacy/hipaa/administrative/combined/hipaa-simplification-201303.pdf.
- 15. Washington State Department of Health (2012). *Guidelines for Working with Small Numbers*. [Revised in Oct 2012]. <a href="http://www.doh.wa.gov/Portals/1/Documents/5500/SmallNumbers.pdf">http://www.doh.wa.gov/Portals/1/Documents/5500/SmallNumbers.pdf</a>. Accessed December 27, 2013.
- 16. Washington State Department of Social and Health Services, Information System Services Division. Information Technology Security Manual (Revision 13.1, March 1, 2013). http://ishare.dshs.wa.lcl/Security/Pages/Manual.aspx.
- 17. Washington State Office of the Chief Information Officer. 141- Securing Information Technology Assets Standards (Revision August 19, 2013). https://ocio.wa.gov/policies/141-securing-information-technology-assets/14110-securing-information-technology-assets.

#### 9. Other Resources

#### **Published research:**

- Benitez K, Malin B (2010). Evaluating Re-Identification Risks with Respect to the HIPAA Privacy Rule. Journal of the American Medical Informatics Association; 17(2):169–177.
- Dankar et al. (2012). *Estimating the Re-Identification Risk of Clinical Data Sets*. BMC Medical Informatics and Decision Making; 12:66.
- El Emam K, A. Brown, P. Abdel Malik (2009). *Evaluating Predictors of Geographic Area Population Size Cut-Offs to Manage Re-Identification Risk*. Journal of the American Medical Informatics Association; 16(2):256-266. <a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2649314/pdf/256.S1067502708002405.main.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2649314/pdf/256.S1067502708002405.main.pdf</a>.
- El Emam K, A. Brown, P. AbdelMalik, A. Neisa, M. Walker, J. Bottomley, and T. Roffey (2010). A Method for Managing Re-Identification Risk from Small Geographic Areas in Canada. BMC Medical Informatics and Decision Making; 10(18), <a href="http://www.biomedcentral.com/1472-6947/10/18">http://www.biomedcentral.com/1472-6947/10/18</a>.
- El Emam K, D. Buckeridge, R. Tamblyn, A. Neisa, E. Jonker, A. Verma (2011). The Re-identification Risk of Canadians from Longitudinal Demographics." BMC Medical Informatics and Decision Making; 11:46, <a href="http://www.biomedcentral.com/1472-6947/11/46">http://www.biomedcentral.com/1472-6947/11/46</a>.
- El Emam K, E. Jonker, L. Arbuckle, B. Malin (2011). A Systematic Review of Re-identification Attacks on Health Data. PLoS ONE 6:12: e28071.
   http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0028071.
- McGraw D. (2013). Building Public Trust in Uses of Health Insurance Portability and Accountability Act De-Identified Data. Journal of the American Medical Informatics Association; 20:29-34, http://jamia.bmj.com/content/20/1/29.full.pdf+html.

#### Other state and federal guidelines:

- Federal Committee on Statistical Methodology (1978). *Report on Statistical Disclosure and Disclosure-Avoidance Techniques*. (Statistical Policy Working Paper 2, NTIS PB86-211539/AS.) Washington, DC: U.S. Department of Commerce.
- Federal Committee on Statistical Methodology (1994). Report on Statistical Disclosure Limitation
  Methodology. (Statistical Policy Working Paper 22, NTIS PB94-165305.) Washington, DC: U.S. Office
  of Management and Budget. [Revised in 2005].
- Federal Committee on Statistical Methodology (1999). Checklist on Disclosure Potential of Proposed
  Data Releases. Washington, DC: U.S. Department of Commerce.
  <a href="http://fcsm.sites.usa.gov/files/2014/04/checklist\_799.doc.">http://fcsm.sites.usa.gov/files/2014/04/checklist\_799.doc.</a>
- Joint Task Forces from the Oregon Administrative Boundaries & Cultural/Demographic Framework
  Implementation Teams (2002). GIS and Confidentiality (Draft).
   http://dshsapoly80831/sites/rda/gis/GISDataConfidentiality/Shared%20Documents/OregonState\_GI\_SandConfidentialityDec2002.pdf.
- National Environmental Public Health Tracking Network (2008). Data Re-release Plan Version 2.5.
   Atlanta, GA: Centers for Disease Control and Prevention, National Center for Environmental Health,
   Division of Environmental Hazards and Health Effects, Environmental Health Tracking Branch.
   <a href="http://ephtracking.cdc.gov/docs/Tracking Re-Release Plan v2.5.pdf">http://ephtracking.cdc.gov/docs/Tracking Re-Release Plan v2.5.pdf</a>. Accessed December 27, 2013.