L – 21: De-Identifying Protected Information and Limited Data Sets

POLICY: See L-21 De-Identifying Protected Information and Limited Data Sets, at www.cosdcompliance.org.

DEFINITIONS: See HHSA Policy L-30 Privacy Definitions.

PROCEDURES:

A. De-Identified Data: When an appropriate business need has been determined, HHSA Programs may disclose data that does not identify an individual and for which there is no reasonable basis to believe that it could be used to identify an individual ('de-identified data'). HHSA Programs must use one of the two methods below to ensure data is de-identified:

1. Safe Harbor Method: Program may assume data is de-identified if Program does not have knowledge that the information could be used alone or in combination with other information to identify an individual AND the following identifiers are removed:
    a. Names (including first or last or initials)
    b. All geographic subdivisions smaller than a State, including street address, city, precinct, zip code, and their equivalent geocodes, except for the initial three digits of a zip code if, according to the current 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.
    c. All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death
    d. All ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older
    e. Telephone and fax numbers
    f. Email addresses, URLs, and IP address
    g. Social security (even last 4 digits), medical record, or health plan beneficiary numbers
    h. Any type of account numbers
    i. Certificate or license numbers
    j. Vehicle identifiers and serial numbers, including license plate numbers
    k. Biometric identifiers, including finger and voice prints, as well as device identifiers and device serial numbers
    l. Full face photographic images and any comparable images AND
    m. Any other unique identifying number, characteristic, or code, except as permitted by section (C) below, Record Identifiers

2. Expert Determination: If Program wishes to include some fields that are excluded from Safe Harbor method above, Program may request an Expert Determination from the Agency Compliance Office (ACO)
    a. Program will provide list of fields, sample data, or actual data set to ACO, as well as any additional information or context requested by ACO.
    b. ACO will evaluate the extent to which the information can (or cannot) be identified.
        1) If risk is very small, then data may be considered de-identified.
        2) If risk is greater than very small, then:
            a) ACO will provide guidance as to what, if any, statistical or scientific methods can be applied to mitigate the anticipated risk, such as to remove, roll-up, redact, or mask data.
            b) After Program has made any changes as suggested by ACO, Program will resubmit data and ACO will re-evaluate the resulting data to confirm the risk is no more than very small. If risk is greater than very small, ACO will provide additional guidance.
    c. ACO will document the methods and results of the analysis that justify the determination.

B. ***Limited Data Sets***: HHSA may use or disclose certain specific data (a 'limited data set') for the purposes of research or public health if ALL of the following are met:
1. HHSA enters into a Data Use Agreement (DUA), Business Associate Agreement (BAA), or similar Agreement with the recipient. The DUA, BAA, or Agreement must be approved by ACO.
2. The data disclosed **excludes** all of the following identifiers:
    a. Names
    b. Postal address information (i.e. street or PO Box address)
    c. Telephone and fax numbers
    d. Email addresses, URLs, and IP address
    e. Social security, medical record, or health plan beneficiary numbers
    f. Account numbers
    g. Certificate or license numbers
    h. Vehicle identifiers and serial numbers, including license plate numbers
    i. Biometric identifiers, including finger and voice prints, as well as device identifiers and device serial numbers
    j. Full face photographic images and any comparable images.
3. The data disclosed may include the following identifiers:
    a. All elements of dates; and
    b. Geographic elements of town, city, and full zip code
    c. Record Identifiers as described below in section C

C. ***Record Identifiers***: HHSA Programs may assign a code or other means of record identification ('record identifier') to allow HHSA Programs a means of re-identifying data, provided that:
1. Derivation: The record identifier is not derived from or related to information about the individual and is not otherwise capable of being translated so as to identify the individual; and
2. Security: HHSA does not use or disclose the record identifier for any other purpose, and does not disclose the mechanism for re-identification

QUESTIONS/INFORMATION: HHSA Privacy Officer at 619-338-2808