Center for Geographic Analysis
Guide to Handling Confidential Information

Version 1.3, revised 1/28/2011

There are two types of confidential information currently recognized at the university[1]:

1) High Risk Confidential Information (HRCI)

This is data containing a person's name + state, federal or financial identifiers
Or research data containing private sensitive information about identifiable individuals.[2]

2) Harvard Confidential Information (HCI)

Business information specifically designated by the School as confidential
Or identifiable business information that puts individuals at risk if disclosed
Or research data containing private information about identifiable individuals
Or student records (such as collections of grades, correspondence)

Procedure for identifying confidential information:

1) Ask the researcher if their data is HRCI or HCI, provide the above definitions if they don't know what the terms are. Make it clear that we are not familiar with their data, and have them state the Harvard assigned sensitivity level of the data, and if there are any data use agreements associated with it, and who is the PI responsible for it.
2) If the data is HRCI or HCI, can the sensitive data be stripped out before sending to CGA? For example can the data be reduced to just a unique ID and address in the case of CGA doing geocoding.
3) If the sensitive information cannot be removed, can the data be supplied to CGA in a non-identified form, so that it cannot be directly or indirectly linked to an individual.
4) If the researcher does not know the sensitivity level of the data, and cannot strip any human identifying information, then CGA should not process the data.
5) Micah Altman can be consulted regarding questions about deidentifying geospatial data: Micah_Altman@harvard.edu

Summary for handling confidential information:

If the answer is no to #2 and #3 above, then IQSS/HMDC approved procedures must be used to transfer, store, access, and dispose of the data. These confidentiality procedures are described in detail here: http://support.hmdc.harvard.edu/kb-930/hmdc_policies
A summary of how to handle the different types of data is found below.

High Risk Confidential Information must be:

- Transported **only** over a encrypted/physically secured network connection, and **directly** to approved storage
- Stored **only** on HMDC/IQSS approved file service, and only in a directory **specifically approved for that HRCI data set**
- Accessed only by those who have been given **written permission by the PI or school security officer**

Harvard Confidential Information must be:

- Transported in encrypted form
- Stored on (a) approved HMDC file service; (b) encrypted portable storage; or (c) securely configured individually owned/managed laptop or desktop
- Accessed only by those who been given permission by the owner or school security officer

[1] Slide 11 of Micah Altman's "Confidential Information at Harvard – Staff Basics" presentation contains more ways to identify confidential information: **http://maltman.hmdc.harvard.edu/presentations/StaffSecurityPres.pdf**
The definitive descriptions of HCI and HRCI are contained in the Harvard Enterprise information security policy: http://www.security.harvard.edu

[2] Note that only an Internal Review Board (IRB) is empowered to determine whether private data involving human subjects is either identified or sensitive. In the absence of IRB documentation, private research data about human subjects should be assumed to be HRCI.