Talking points for acquiring local data for the Geographic Support System (GSS) Initiative

Please note that this is not intended to serve as a "script" for making phone calls, rather a concise set of talking points and/or guidelines to use as questions arise.

Why are we asking the regions to call local governments?

The Census Bureau must enhance the quality, coverage, and currency of the Master Address File (MAF) and TIGER Database, as well as maintain the improved quality achieved over the past decade. To achieve this goal, the Census Bureau is particularly interested in expanding our partnership activities by incorporating local geospatial and address datasets into our research and development activities.

The Census Bureau wants to establish data sharing partnerships with local governments. We would like the local governments to provide the Census Bureau with copies of their address/feature/parcel data, so we can then utilize their data to analyze the accuracy of the spatial features (roads, boundaries, etc.), addresses, and geocodes (address locations) within the MAF/TIGER Database and identify areas requiring corrections or updates.

Simultaneously, we will be analyzing the accuracy and completeness of the local data (by comparing both address and spatial data to the MTDB, commercial datasets, and possibly conducting field work), and will provide partners with information on the results of our analysis. Where we determine that the data is usable for MTDB updates, we will be seeking to initiate a regular exchange with the data provider in order to capture change and/or growth.

What is the GSS Initiative? What are its goals?

The GSS Initiative is an integrated program of improved address coverage, continual address and spatial feature updates, and enhanced quality assessment and measurement.

One of the goals is to develop programs to implement continual updates of the MTDB throughout the decade to support current surveys (including the American Community Survey). This differs from the MAF/TIGER Accuracy Improvement Project (MTAIP) in that we would be seeking to acquire suitable local data on a regular, recurring basis, in order to assure that the data in the MTDB is as accurate and current as possible.

Another benefit of continual updates is that where we are able to ensure that the MTDB is up-to-date, we may be able to conduct a targeted, rather than complete, Address Canvassing operation during 2019 in preparation for the 2020 Decennial Census.

Some of the more specific goals of the GSS Initiative are:

- ***Address improvement****: explore methodologies to achieve complete coverage and a current address list, concentrating on rural areas, Puerto Rico, and group quarters, and improving geocoding of all addresses to their location*
    - Initiate programs with partners to continually receive addresses throughout the decade
- ***Feature improvement:*** *continual update of the street network and attributes to improve the matching of addresses to their correct geography*
    - Broaden participation in existing programs for receiving partner Geographic Information System (GIS) files and imagery
    - Research change detection techniques
- ***Quality improvement:*** *broaden quality assessments and provide quantitative measures*
    - Research effective methods for evaluating existing data, including the use of local data as a benchmark for comparison

- ***Improved Partnerships:*** *strengthen existing and develop new partnerships*
    - Research methodologies and develop pilot programs for working with partners in acquiring address and spatial data in the most efficient and least intrusive ways

How will we use the partner's data?

Initially, the Census Bureau will be using the data for research and development purposes only (i.e. no updates of the MAF/TIGER DB). At a minimum, we will use the data for change detection and completeness/coverage testing of the MTDB. We will run comparisons between the local dataset(s) and the MTDB to identify areas where we are missing features and/or addresses, as well as areas where our data appears to be inconsistent relative to the local's data. In the process, we will be learning more about the ways local governments create and store their data, evaluating possible data exchange mechanisms, and seeking the partner's input as to data/feedback they would like to receive from us.

Feedback on spatial data may include information similar to that provided during the MAF/TIGER Accuracy Improvement Project (MTAIP), including overall accuracy measurements, but possibly additional information that was not included in the MTAIP feedback – content to be determined.

Note that while Title 13 of the US Code prohibits the Census Bureau from sharing information on individuals, including individual addresses in the census address list, we would be able to discuss with partners the results of our research and evaluation of their address data in a general sense.

What specifically are we looking for?

We want addresses, structure points, features, parcels, boundaries, **and** imagery.

If the local would prefer to provide their complete dataset without extracting any of the specific layers that we asked for, we will gladly accept the offer.

NOTE: If the local states that these items are available on line, ask for the web address.

1) Address Data: either as a list, or as structure coordinates or building footprints, either or both of these with attribute data (including address) – including any data linking E911 addresses to mailing addresses.
2) Cadastral or Tax Parcels - including address attribute data (location preferable, mailing acceptable). Specific householder information and assessed value is not necessary.
3) Street Centerline file with attribute data - Feature identification (street names), road classification codes, and address attribute data (address ranges). ZIP Code information is helpful as well.
4) Other feature layers that may be useful include:
    - Hydrography - rivers, streams, lakes, ponds.
    - Rail Features - heavy and light rail, subways.
    - Boundaries - County, Municipal (Incorporated Place, Township), School District, Voting District, Congressional, Emergency Service Zone (Fire, Police, Ambulance), etc.
    - Land Use layers and/or boundaries - parks, cemeteries, Federal land, military installations, etc.
    - ZIP Code Boundaries.
    - Census Statistical Entity boundaries - tracts, block groups and blocks (spatially rectified to your own more accurate layers).
    - Key Geographic Locations - government centers, schools, colleges, hospitals, etc.

- Natural Resource Pipelines - oil, natural gas, etc.
- Electric Power Transmission Lines.

For any of the above that are available we would be very interested in acquiring the metadata, especially:

- measurements of spatial accuracy,
- source information,
- update frequency, and
- quality checks.

Also, if there is a data dictionary available for any or all of the above layers, we want to acquire it as well. In short, if there is documentation of any type that will help us understand/evaluate the data, it should be part of the acquisition.

We can accept digital geographic files in any format, but a Geodatabase or shape files are preferred. If layers are linked to other databases, or (for example) the data is in a tabular format, we can accept data base information in a variety of formats (ASCII text files, Comma Delimited, Excel, Access, etc.).

For now, the preferable medium of transfer is FTP transfer, direct download from provider websites, or CD-ROM/DVD. Additional transfer mechanisms will be available in the future, such as a secure web exchange system (currently in the requirements development phase).

Other good things to ask (Note that much of this information may be available online, please gather any information you can prior to making contact then ask only the questions where you are unable to find answers in advance):

- Are there any restrictions on the use of this data, for example licensing agreements? Note that during the research and development phase, we WILL NOT be using this data to update the MTDB, but it would be good to know this info for future reference (as we will be seeking to use data, where suitable, to update the MTDB in the future).
- How frequently (and via what methods) are the datasets updated? As part of the "continuous update" concept, we will be seeking to acquire and use the data, where suitable, to make regular updates to the MTDB.
- Do you offer a web mapping or web feature service? If yes, what is the address of the service?
- Do you know what office houses your addressing authority? Again, for future reference as an additional data source if necessary.
- Do you acquire imagery on a regular basis and if so is it available via ftp or other download method?
- What type of address data does the file contain?
    - Are the addresses residential, commercial, or both (if both, how are they differentiated in the file)?
    - Are the addresses owner addresses and not location/situs addresses (if both, how are they differentiated in the file)?
    - Are the addresses mailing addresses or location/situs addresses (if both and they are different, how are they differentiated in the file)?
    - If the ZIP Code is provided for an address that is not the mailing address, how was the ZIP Code determined?
    - Does the file contain any Rural Route, or Highway Contract Route (HCR) type addresses?
- What type of location data is the address data linked to?

    - If addresses are linked to parcel data, are the addresses situs/location addresses, owner (tax) addresses, or both? Are there addresses for parcels with structures only, or for all parcels?
    - If addresses are linked to building footprints, how was the building footprint derived (from imagery, etc.)? Are all building footprints, including sheds, barns, etc. included, or only structures where people can live or work?
    - In the situation where there are multiple housing units within a single building footprint, how are those represented in your file? (For example, a multi unit structure, do you provide only the BSA? For a townhouse structure do you subdivide the building footprints, or only provide one townhouse address for the building footprint?)
- If the data contains structure coordinates, how are the coordinates calculated (from imagery, GPS, mailbox location, etc.)?
- Are you aware of the National States Geographic Information Council (NSGIC) GIS Inventory (also know as RAMONA)? If so, have you created an account and provided information about your data? If not, would you allow the Census Bureau to create a record of your data within the inventory?

What if the local government asks for compensation in exchange for their data?

Please notify GEO so we can evaluate the feasibility of making a purchase.

"What's in it for me? How can we both benefit from sharing data"

- Feedback on what we learn – **NOTE THAT WE ARE SOLICITING IDEAS FROM POTENTIAL PARTNERS – WHAT WOULD THEY LIKE TO SEE?** The exact format of the "feedback" is to be determined, and will likely evolve as we continue meeting with partners and evaluating partner-supplied data. Note that Title 13 restrictions apply to any potential address-level feedback, so please do not make promises of that nature.
- Improved address and feature coverage to support current survey samples, including the American Community Survey.
- Taxpayer savings – by establishing these partnerships, and developing the means to continually update the Census Bureau's Address List and spatial data, we avoid the necessity of having address canvassers walk every street in the U.S. looking for updates prior to the 2020 Decennial Census.
- A more accurate 2020 Census, with all the benefits therein (increased funding, etc.)

Other common questions?