Chapter 11a: CULTURAL AND DEMOGRAPHIC DATA

Coordinator: Tim McMonagle
U.S. Census Bureau, Los Angeles Regional Office
timothy.william.mcmonagle@census.gov

Theme Description:

These geospatially referenced data describe the characteristics of people, the nature of the structures in which they live and work, the economic and other activities they pursue, the facilities they use to support their health, recreational and other needs, the environmental consequences of their presence, and the boundaries, names and numeric codes of geographic entities used to report the information collected. (OMB Circular No. A-16)

Status:

There are enormous amounts of statistical information available from federal, state and local government and nonprofit and private organizations. There are a number of issues that affect the availability and usefulness of this data. These issues, while not unique to cultural and demographic data, are particularly significant to this data theme include:

- Finding Data
- Geospatial Referencing
- Temporal
- Confidentiality
- Data Sharing Strategies
- Security

Finding Data:

Cultural and Demographic data is the largest and most diverse data theme. In many ways finding the appropriate data has changed little in these days of electronic databases and the internet. Data users still have to search agency by agency to find out what data is available, who produces it, and whom they need to contact. The sheer volume of data available can make this task difficult, time consuming, and costly. The expansion of the data available over the internet can helped make the search process faster, but a good search strategy is still needed.

A result of the federal E-Government Initiative has been the establishment of data clearinghouses that include cultural and demographic data. FedStats, for example, provides data and links to the data generating agencies. There is no comparable clearinghouse for state and county data in Hawai'i. A few state agencies, such as DBEDT, provide links to other data providers along with their own data.

Geospatial Referencing:

1. Data tabulated at a levels of geography:

Many federal, state, local, and private organizations aggregate and publish data by political and/or administrative units such as counties, zip codes, cities, etc. The aggregation of the data allows for the protection of the anonymity of the individual, while still allowing the release of data.

Data aggregation imposes some limitations on the usefulness of data particularly where the levels of geography used do not match data users needs. For instance, a great deal of information is available at the county level, but this would not meet the needs of someone requiring sub-county data. An agency may compile data at a level that is not readily compatible with the geography that is used by another agency (e.g. zip codes vs. census tracts). The geography used for tabulation may be uniquely tailored to meet the needs of that particular agency (e.g. police beat areas). Some of these tabulation areas can and do change over time to meet the current needs of an agency thus compromising the historical comparability of the resulting data.

The aggregation of data also homogenizes the data to the geography employed and this can have a negative impact on the quality of the data. For example, auto thefts data in Honolulu are released by police beat areas. This generalizes the data to the whole beat area when the thefts may be occurring in specific locations within the aggregation area. Another problem can occur if the aggregation area includes highly diverse characteristics. What if areas of high and low income were included in the same aggregation area? The median income for that area would not reflect the true diversity present. The severity of these problems is usually proportional to the size of the geographic areas used to aggregate the data.

2. Data available with addresses:

Addresses provide a means of locating people, structures and other spatial objects. Addresses are a key part of most data sets involving people and organizations as well as their activities, and many agencies maintain data keyed to addresses. Most organizations maintain address lists or have databases or datasets that contain addresses. Agencies out of necessity have developed their own address standards in response to their own needs. As a result existing agency address standards do not provide common semantic definitions, which can make the exchange of address data at best difficult. At a minimum, data users must have knowledge of both the structure and content of an agency's address standard to successfully share the address-based data. The use of a common addressing standard would enable an easier exchange of address-based data.

The Federal Geographic Data Committee's (FGDC), Subcommittee on Cultural and Demographic Data has worked on developing a national address standard for a number of years. In the spring of 2005 the FGDC accepted a proposal submitted by Urban and Regional Information Systems Association (URISA) to create a Street

Address Standard. A draft version of their proposed Street Address Data Standard is available on URISA's website (http://www.urisa.org/address_data_standard.htm). The URISA's proposal will be submitted through the FGDC's formal standards approval process. (Federal Geographic Data Committee, 2005)

There does not appear to be an address standard employed throughout the state of Hawai'i. There appear to be situations where not all the agencies in a county government employ the same addressing standard.

3. Data with no geospatial reference:

Most cultural and demographic data has or had some form of geospatial referencing. There are situations were the geospatial information has been stripped from the data set. In some instances this is done to ensure confidentiality. In other instances an agency may not have any use for the geospatial information.

Temporal:

Most cultural and demographic data have an important temporal element. Most human related data is in a constant state of change. It is therefore critical to know when the data was collected and the time period that it represents. These temporal components must be included in the datasets metadata. (OFIT,?)

Confidentiality:

Confidentiality is a significant issue for Cultural and Demographic Data. Much of the potential cultural and demographic data public agencies and non-profit and private organizations collect and maintain are by their very nature confidential. One of the primary goals of confidentiality is to protect privacy by not allowing an individual observation to be identifiable. This can profoundly impact the release and potential analyses related to many data sets. Confidentiality is generally ensured by either data aggregation techniques or by means of constraining the access and/or use of confidential data. Federal and state agencies in particular need to protect confidentiality. This creates a complex problem for the data steward: Balancing the "right to know" versus "right to privacy".

It is important to be able to use, display and document data while not violating individual's rights and privacy. Some confidential data is related to specific human activities such as employment, taxes, mental health, welfare programs, and other social programs. Having access to this data is becoming more critical to making effective data-driven decisions and reducing cost in governmental organizations. These costs are associated with building and maintaining databases, replicating efforts, and using objective information to make quick and defendable decisions.

Two major sets of laws exist nationally and in states related to dispersing data. The first set of laws is intended to maintain confidentiality and protect an individual's rights. The

second set, the Freedom of Information laws, deal with the public dispersal of information. In many cases, there is no clear dividing line between the opposing mandates. (Wartell and McEwen, 2001)

In Hawai'i all government agencies at the state and county level are subject to The Uniform Information Practices Act (Modified), Chapter 92F, Hawai'i Revised Statutes ("UIPA"). UIPA is an informational practices statute, which encompasses both freedom of information principles and privacy principles. The statute also addresses the secondary uses of personal information and data sharing between governmental agencies. The Office of Information Practices (OIP) has oversight for the state's information practices (http://www.hawaii.gov/oip/index.html).

Data Sharing Strategies:

 Data sharing falls into three general levels:

1. Publicly shared and distributed data-This is often aggregated data and can often be accessed on the internet.
2. Restricted access to, use of, or redistribution of data- Share limited amount of data based on specific request and applications. Some data can only be released for research purposes or to other agencies. In some cases Memorandums of Understanding for data sharing and exchange need to be worked out.
3. Data not shared outside the agency-These restrictions are often mandated by law.

The Census Bureau, for example, cannot by law (Title 13 of the US Code) share the information from individuals or specific housing units with anyone, but data aggregated at various levels of geography is available to all via the internet or various other medias.

Some states have data sharing strategies between agencies. Some of these strategies are specifically related to GIS activities in order to minimize costs among and between state programs and to ensure high quality data that is transferable between applications. For example, Maryland also has an established interagency sharing agreement that specifically outline data flow to specifically promote fast and effective geo-coding between agencies. Several agencies share building a master database and thereby reducing costs between programs.

In Hawai'i there does not appear to be any general data sharing strategies. UIPA mandates that state and local government agencies must share public information with any one who requests it. Each state department helps data requesters as much their resources permit, with staff time often being the critical resource.

There can be some technical issues that make data sharing difficult; incompatibility of operating systems, data layout, or software. However, the major data sharing issues are mostly due to politics and personalities. (Wartell and McEwen, 2001)

Security:

A small portion of Cultural and Demographic data could pose risks to security and may therefore require safeguarding. The FGDC has come up with guidelines that provide a procedure consisting of a sequence of decisions that an originating organization should make about geospatial data. (FGDC, 2005)

Safeguarding is justified only for data that contain sensitive information that are the unique source of the sensitive information and for which the security risk outweighs the societal benefit of dissemination. The guidelines offer two safeguard options:
1. Change the data: Change the data to remove or modify the sensitive information.
2. Restrict the data: Establish restrictions on the data that are commensurate with the assessed risk.

If your organization did not originate the geospatial data you should not make decisions about safeguarding the data. You should honor any instructions that accompany the data. If no instructions accompany the data you may presume that no restrictions apply to the data. Instructions, terms, and conditions may be found in the accompanying metadata and/or in licenses, signed agreements (including non-disclosure agreements), or other instruments that accompany data. The data user is responsible for knowing and honoring restrictions that accompany the data. (Wartel and McEwen, 2001)

Data Sources:

There is an enormous amount of geospatially referenced cultural and demographic data that is produced by a multitude of federal, state and local agencies and non-profit and private organizations. The challenge is to know what data is available, how to get it, and quite often who to contact. At the federal level there are clearinghouses that provide links to the agencies that produce the information. There are no comparable state or county clearinghouses for Hawaiian cultural and demographic data. A data user needs to contact individual agencies find out what data is available. A few state sites do provide links to other agencies that have data available. The following list of sources is a sample of clearinghouses and data sites that are available on the internet.

Theme	Source	Location*	Sector	Description
Clearinghouse	Geodata.gov	geodata.gov	Federal	Includes links and metadata for a number of categories of Cultural and Demographic Data. Agricultural and Farming; Business and Economic; Cultural, Society and Demographic; and Human Health and Disease

Clearinghouse	FGDC Subcommittee on Cultural and Demographic Data: Links to Cultural and Demographic Data	Links page	Federal	Links to web sites maintained by subcommittee members
Clearinghouse	FedStats	fedstats.gov	Federal	Gateway to statistics from over 100 U.S. Federal Agencies. Provides links to statistics and to statistical agencies. MapStat- has statistical profiles of states, counties, cities, congressional districts, and Federal judicial districts.
Population, Housing, and Economic	U.S. Census Bureau	census.gov	Federal	Wide range of population, housing, and economic data gathered during the Decennial Census, American Community Survey and other surveys, and the Economic Census. Data tabulated at various layers of geography. TIGER.
Agriculture	USDA: National Agricultural Statistic Service	Hawaii agriculture statistics	Federal	Produces official estimates of agriculture for the nation and states.
Crime and Justice	Hawaii State Department of the Attorney General: Research and Statistics	AG	State	Crime and justice statistics for Hawaii
Crime	Honolulu Police Department	HPD	City/County	Crime statistics-Reported offense in seven crime categories by police beat areas.
Economic	Bank of Hawaii	Bank of Hawaii	Commercial	Information about the economy of Hawaii and the Pacific region.
Economic	Department of Business, Economic Development, and Tourism	DBEDT	State	Statistical and economic information.
Economic	University of Hawaii Economic Research Organization	UHERO	Education	Provides information on economic, demographic, and business trends in Hawaii and the Asia-Pacific region.

Education	Hawaii Department of Education: Reports	DOE	State	Education statistics for Hawaii
Health	Hawaii State Department of Health	DOH	State	Health and vital statistics for Hawaii
Historical	Hawaii State Archives	Archives	State	Collects, appraises, preserves, and make available Hawaiian government records.
General	County of Hawaii: Department of Research and Development	County of Hawaii	County	General statistics for the County of Hawaii

* See Appendix for Complete URLs for these sites.

Standards:

"Street Address Data Standard (Working Draft 2.0)" Address Standards Working Group, The Subcommittee on Cultural and Demographic Data, Federal Geographic Data Committee, 2005 (http://www.urisa.org/address_data_standard.htm).

"Privacy In The Information Age: A guide for sharing crime maps and spatial data", Julie Wartell and J. Thomas McEwen, 2001, US Department of Justice, Institute for Law and Justice, The authors propose a set of guidelines to promote and encourage the proper use and distribution of maps and data that will assist an agency in addressing confidentiality issues. They do not feel that a universal standard was possible because agencies at different levels of government are subject to a wide variety of laws concerning the confidentiality of data. While these guidelines were developed specifically for crime data, they are equally relevant for other forms of Cultural and Demographic Data. (http://www.ncjrs.gov/pdffiles1/nij/188739.pdf)

"Guidelines for Providing Appropriate Access to Geospatial Data in Response to Security Concerns." Federal Geographic Data Committee, 2005. A link to a PDF version of this document is located at: http://www.fgdc.gov/policyandplanning/fgdc-guidelines

"Content Standard for Digital Geospatial Metadata" Federal Geographic Data Committee, 1998, STD-001-1998, Http://www.fgdc.gov/standards/standards_publications/   General metadata standard that should be applied to all data themes.

Priority:

Low. The data already exists and is available.

Estimated total investment in this theme:

This would be difficult to calculate, the data already is being generated by agencies in the performance of their duties.

Estimated current state and local contributions:

There is a tremendous amount of information gathered by agencies while planning and implementing their assigned duties.

What is needed:

What is being proposed is an effort to catalog information about existing data sets and actions that might make the exchange of information easier. Listed below are some suggestions on what can be done to make Culture and Demographic data more available to data users.

1. Identify key cultural and demographic geospatial data sets that meet critical planning and management needs. Work to improve access to these data within the restrictions in applicable confidentiality and privacy regulations.

2. Add a page to the Hawai'i GIS Clearinghouse node for Cultural and Demographic Data. There are a number of potential implementation levels for posting cultural and demographic data at the clearinghouse:

- Post identified critical and/or commonly used data sets, metadata, and links to existing data sets and data clearinghouses. Be able to sort by subject, agency, and location.
- Post metadata and links to known data sets and data clearinghouses. "Geodata.gov" at http://www.geodata.gov/gos provides a good, if somewhat elaborate model. This web site provides metadata and links to sites maintained by the data stewards, it does not host the data itself.
- Post links to existing data sets and data clearinghouses.

3. Set up a metadata repository for the information about the data sets. Ideally the information would be provided by the agency that generates the data. In California, state agencies and other contributors can provide information about their geospatial data sets by posting metadata to an online directory (**http://gis.ca.gov/catalog/contrib1.html**). A metadata template is employed to ensure uniformity in the submissions. A potential option could be to allow data users who find a useful data source to provide the metadata.

There would need to be guidelines for what metadata would be acceptable to be posted to the repository.

4. Promote the adoption an Address Data Content Standard in Hawai'i.

5. Promote the retention of geospatial referencing information in cultural and demographic datasets.

What is the likely source:

Numerous federal, state and county agencies and non-profit and private organizations produce Cultural and Demographic data.

Estimated total investment needed to complete this theme:

Costs would be dependent on the choice of what to post to the geo-data clearinghouse.

Estimated current allocation of funding:

Unknown.

Estimated budget shortfall:

Unknown.

Possible ways to overcome this gap:

Unknown.

Most appropriate data steward:

The agency or organization that generates the geospatially referenced data would be the data stewards.

Maintenance Process:

Maintenance would involve periodic updates of existing links and the addition of new links.

Estimated Maintenance cost:

Unknown.

Bibliography:

Federal Geographic Data Committee, 2005, "Street Address Data Standard (Working Draft 2.0)" Address Standards Working Group, The Subcommittee on Cultural and Demographic Data, Federal Geographic Data Committee (http://www.urisa.org/address_data_standard.htm).

FGDC, 2005, "Guidelines for Providing Appropriate Access to Geospatial Data in Response to Security Concerns" http://www.fgdc.gov/policyandplanning/fgdc-guidelines

Julie Wartell and J. Thomas McEwen, 2001, "Privacy In The Information Age: A guide for sharing crime maps and spatial data", US Department of Justice, Institute for Law and Justice (http://www.ncjrs.gov/pdffiles1/nij/188739.pdf

OFIT (Oregon Framework Implementation Team: Oregon Preparedness Framework Group), ?, " Draft Cultural/Demographic GIS Implementation Plan for Oregon" http://www.oregon.gov/DAS/IRMD/GEO/docs/prepfit/Draft_Cultural_Implementation_Plan.pdf

Appendix: URLs for Data Sources Table

Geodata.gov	http://www.geodata.gov/gos
FGDC Subcommittee On Cultural and Demographic Data: Links to Cultural and Demographic Data	http://www.census.gov/geo/www/standards/scdd/CDDLinks.html
Fedstats	http://www.fedstats.gov/
US Census Bureau	http://www.census.gov/
USDA National Agricultural Statistics: Hawaii	http://www.nass.usda.gov/hi/homepage.htm
Hawaii State Department of The Attorney General: Research and Statistics	http://www.cpja.ag.state.hi.us/rs/index.shtml
Honolulu Police Department	http://www.honolulupd.org/info/statistics.htm

Bank of Hawaii	http://www.boh.com/econ/index.asp
Department of Business, Economic Development, and Tourism (DBEDT)	http://www3.hawaii.gov/DBEDT/index.cfm
University of Hawaii Economic Research Organization (UHERO)	http://www.uhero.hawaii.edu/
Hawaii Department of Education: Reports (DOE)	http://doe.k12.hi.us/reports/index.htm
Hawaii State Department of Health (DOH)	http://www.state.hi.us/doh/stats
Hawaii State Archives	http://www.state.hi.us/dags/archives/
County of Hawaii: Department of Research and Development	http://www.hawaii-county.com/info/stats.html