CHAPTER 15: DATA DISTRIBUTION/PUBLISHING

Coordinator: Joan Esposo
State DBEDT/Office of Planning/GIS
jesposo@dbedt.hawaii.gov

Theme Description:

Data Distribution/Publishing:
Data Distribution and Publishing refers to the mechanisms used to share and publish geospatial data. In Hawai`i, this includes several GIS data download sites and one National Spatial Data Infrastructure (NSDI) Metadata Clearinghouse Node. In addition, much geospatial data is shared informally, via email attachments, CDs, etc.

Metadata:
No discussion of data distribution is complete without discussing metadata. Metadata is commonly referred to as "data about data." In relation to geospatial data sets, metadata should provide information as to a data set's origin, content, quality, condition and availability.

The two most common metadata standards for geospatial data are the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM) Version 2, and the International Organization for Standardization (ISO) International Standard 19115 for Geographic Information - Metadata. The FGDC Standard is a widely used national standard, while the ISO standard is an international standard which is likely to eventually be adopted by the FGDC. The HIGICC will have to choose which of these standards it will adopt.

In Hawai`i, there is a wide range of adherence to the practice of creating metadata. Some agencies don't create it at all, others create metadata that is not compliant with any standard, and still others create FGDC compliant metadata.

Status:

Data Distribution/Publishing:

Many of the most commonly used data sets in Hawai`i are already distributed on line, either via FTP or via web browser download. For example, both the State of Hawai`i GIS Program and the City and County's Honolulu Land Information System (HOLIS) have over 50 data layers each available for download. Another way that geospatial data is made available in Hawai`i is through web services, for example using ESRI's ArcIMS to allow queries of the State's or the County's GIS databases.

In addition to the publicly distributed databases, there are numerous data sets that are not available, for a variety of reasons. Some data sets have distribution restrictions either because they are proprietary (e.g., GDSI parcel data) or because they contain sensitive information (e.g. archaeological sites, rare species). Other data sets are not available because they were created for a specific project, because they haven't been quality checked, or because there is no easy mechanism to distribute the data. Another common reason for not distributing data sets is that they haven't been properly documented (i.e., metadata has not been created).

Metadata:

Data publishers often don't create metadata for their geospatial data – data is frequently exchanged with little or no documentation. When metadata is created, it is often missing key information – for example metadata may not contain information about the data's datum or projection, or doesn't contain information describing attributes. Much of the metadata available, although it may contain all of the relevant necessary information for proper use of the data, may not be FGDC compliant (e.g., the State of Hawai`i uses .txt files which describe source, projection and attribute information, but are not in a compliant form). Other metadata may be FGDC compliant, but still doesn't contain the information necessary to properly use the data (e.g., the City and County of Honolulu has a great deal of FGDC compliant metadata that lists identification information, but does not describe attribute values, source information, dates, etc.).

Currently, there is one NSDI Metadata Clearinghouse Node in Hawai`i , hosted by the Office of Planning In addition, geospatial data users can query the National Biological Information Infrastructure (NBII) clearinghouse node.

Data Sources:

Data Distribution/Publishing:

There are multiple sources of on-line geospatial data in Hawai`i, as listed in Appendix E. The key data distribution/publishing sites are:

State of Hawai`i GIS Program Download Site: http://www.state.hi.us/dbedt/gis/
State of Hawai`i GIS Program IMS Site: http://gis.state.hi.us/website/OPMap
Honolulu Land Information System (HOLIS) Download Site: ftp://gisftp.hicentral.com/
Honolulu Land Information System (HOLIS) Interactive Mapping Sites:
http://gis.hicentral.com/website/parcelzoning/viewer.htm,
http://gis.hicentral.com/website/ecodev/ed.asp
Pacific Basin Information Node (PBIN): http://pbin.nbii.gov/
Pacific Disaster Center (PDC) Interactive Mapping Site: http://www3.pdc.org/iweb/
Maui High Performance Computing Center (MHPCC) Data Download FTP Site:
ftp://sync.mhpcc.edu (Currently hosting USDA DOQQs for download).

Metadata:

Currently, there is one Z39.50 NSDI Metadata Clearinghouse Node using I-Site in Hawai`i , hosted by the Office of Planning and accessed through the FGDC metadata clearinghouse: http://www.fgdc.gov/clearinghouse/clearinghouse.html

In addition, Hawai`i GIS users access the National Biological Information Infrastructure (NBII) Metadata Clearinghouse: http://www.nbii.gov/datainfo/metadata/clearinghouse/

Standards:

Data Distribution/Publishing:

There are no standards in Hawai`i for data distribution or publishing at this time. Data is available for download at the various download sites listed above in either http or ftp format. Many of the interactive mapping sites listed above use ESRI's ArcIMS, though the use of this software is not required, nor is it an established standard.

Metadata:

Hawai`i's NSDI metadata clearinghouse node uses the ANSI standard Z39.50 (http://www.blueangeltech.com/Standards/GeoProfile/geo22.htm) for the query, search, and presentation of search results to web clients. Currently, the metadata stored in the clearinghouse node uses the FGDC Content Standard for Digital Geospatial Metadata, Version 2.0 (FGDC-STD-001-1998; http://www.fgdc.gov/metadata/contstan.html).

As mentioned above, the two common standards used for geospatial metadata are the FGDC standard and the ISO standard. There are plans to merge the two standards into one in 2003. Specifically, "The US will adopt ISO 19115 as FGDC Version 3 and will expect and support the new XML structure as the primary service exchange. The existing required search fields will be mapped against the new ISO field targets, so the Z39.50 search "will go through" using the same old tags." (See http://clearinghouse4.fgdc.gov/fgdcfaq/showquestion.asp?faq=6&fldAuto=148)

Once the FGDC/ISO metadata standard harmonization activity is completed, Hawai`i's Clearinghouse node should be converted to use the FGDC Version 3 / ISO 19115 standard.

The NBII Clearinghouse Node uses the NBII Metadata Standard (FGDC-STD-001.1-1999), which is an enhancement of FGDC metadata standard (http://www.fgdc.gov/standards/status/sub5_2.html).

Priority:

Data Distribution/Publishing:

The establishment of downloading or interactive web sites should have a fairly high priority, as it makes data sharing easier.

Metadata:

The maintenance of Hawai`i's metadata clearinghouse node should be given a high priority, as the node could be the best mechanism for geospatial data discovery and to reduce duplication of effort in the creation and maintenance of GIS data.

Estimated total investment in this theme:

Data Distribution/Publishing:

Unknown. The various download and interactive mapping sites are maintained by a variety of governmental entities in and outside of Hawai`i, with varying degrees of investment in both dollars and staff time.

Metadata:

Unknown. In the case of the NSDI Clearinghouse Node maintained by the Office of Planning, there has been no monetary investment. The site was set up on an existing web server using free I-Site software. Investment of staff time has been minimal as well, due to lack of resources, which has resulted in some down time for the node, as well as a very minimal set of metadata records being housed on the node.

Estimated current state and local contributions:

Data Distribution/Publishing:

Almost all data distribution and publishing is done at the state and local level in Hawai`i, with some download/interactive mapping sites being housed and maintained by federal agencies.

Metadata:

The State of Hawai`i Clearinghouse Node is maintained entirely by State government. The NBII Clearinghouse is a partnership between local, state, federal and private agencies and organizations.

What is needed:

Data Distribution/Publishing:

There are no pressing needs at this time.

Metadata:

There should be one metadata clearinghouse for data developed in Hawai`i , to which all data developers/holders would submit their metadata for publication.

In order to have a successful metadata clearinghouse, GIS data developers/publishers must be encouraged to develop useful, compliant metadata.

A decision must be made as to whether the Office of Planning will continue to host the node, and whether this will be the only Hawai`i clearinghouse node, housing all of the metadata from various GIS data holders in the State (i.e., State, County, Private, Non-profit). Other possible hosts for a single, comprehensive node would be the Pacific Disaster Center (PDC), the Maui High Performance Computing Center (MHPCC) the University of Hawai`i (UH), and the Pacific Basin Information Node (PBIN). Another option would be for each county to host their own node. Issues to consider in making the decision would include staffing and funding to maintain the node, as well as the future organizational placement and structure of the Office of Planning.

What is the likely source:

Unknown.

Estimated total investment needed to complete this theme:

More staff time, possible server upgrade or new server ($7,000).

Estimated current allocation of funding:

$0

Estimated budget shortfall:

$0 - $7,000

Possible ways to overcome this gap:

Unknown

Most appropriate data steward:

State of Hawai`i, Office of Planning.

Maintenance Process:

A mechanism needs to be established for the host of the node to receive metadata from participants. Most likely, this would require that data holders submit compliant XML files to whatever agency is hosting the node, which would then publish the file.