

---

Volume 4  
Number 2  
*Government Information: Past, Present and  
Future (Summer 1998)*

---

July 2014

## OSU's government information sharing project: Future directions

Carolyn Ottow  
*Oregon State University*

### Recommended Citation

Ottow, C. (2014). OSU's government information sharing project: Future directions. *OLA Quarterly*, 4(2), 15-16. <http://dx.doi.org/10.7710/1093-7374.1483>

© 2014 by the author(s).

*OLA Quarterly* is an official publication of the Oregon Library Association | ISSN 1093-7374

---

# OSU's Government Information Sharing Project: Future Directions

by Carolyn Ottow  
Government Publications Librarian  
Oregon State University

---

In 1995 the OSU Library began providing census data over the Internet on its Government Information Sharing Project (GISP) Web site (<http://govinfo.library.orst.edu>). At the time, the library was receiving hundreds of CDs from the 1990 census and, as a land grant institution and a depository library, wanted to be able to share the data with remote users in Oregon. With funds from the U.S. Dept. of Education, staff on the project created an easy-to-use Web site to provide access to many of these CDs. The GISP Web site proved to be a resounding success. Not only has it made data available to users throughout Oregon but throughout the world. It currently receives over 90,000 hits per week from over 10,000 unique hosts. Over 3,000 Web pages link to it. It started small, with only data from the Northwest states, but has since grown to include demographic, economic and educational statistics for all of the U.S. Today there are 12 databases online:

- 1990 Census of Population and Housing
- USA Counties, 1996
- Population Estimates, 1990-1994
- Equal Employment Opportunity File, 1990
- Regional Economic Information System, 1969-1995
- 1992 Economic Census
- Census of Agriculture, 1982, 1987, 1992
- U.S. Imports/Exports History, 1992-1996
- Consolidated Federal Funds Reports, 1987-1996
- Earnings by Occupation and Education, 1990
- School District Data Book Profiles, 1989-1990
- Oregon Population Survey, 1990-1996

When the project began, people were thrilled to find such a wealth of information among the chaos on the Internet. A few other sites provided census data, but they were intended for researchers and academic users, and their interfaces were less than intuitive. The GISP was often complimented for the straightforward, easy-to-use interface. Users who have submitted comments to the site have called it "tremendously easy to use" and "one of the best designed Web pages I have seen." But as the Internet grows, users are becoming more sophisticated and demanding more features. Initial glowing comments changed to questions about how to get just the right information out of the database. Many people complained that the data was too old and asked when it would be updated. One user wrote, "What a dog ... I have tried to use your data base ... but I am able to pick only one county at a time." It's no longer enough to put information online; people want more flexibility and more choices for displaying and saving the information.



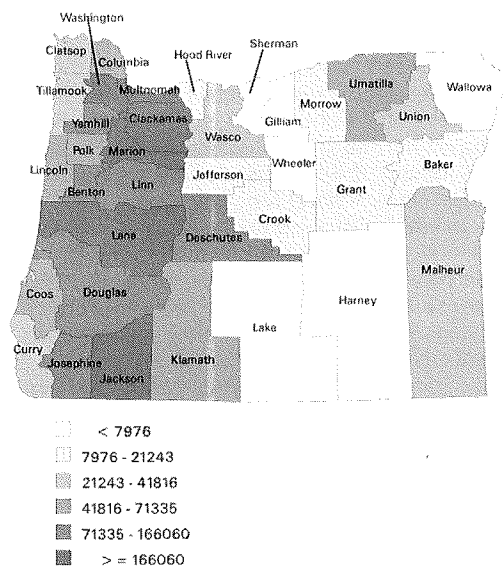
New funding from an LSTA grant through the State Library will help the GISP to respond to the growing needs and expectations of the public. Several enhancements are planned for the site. Already implemented is the long-asked-for ability to download data. Most reports are now available in semi-colon delimited files, which makes them easy to import and use in spreadsheets. Future plans also include providing more display options, a mapping function and more data at greater levels of detail.

When the site first went online, basic summary statistical tables were available through an easy graphical interface. Users first selected the geographic entity (a particular state, county or city) and then chose from a list of topical reports available. While simple, this gave users access to thousands of reports for the entire U.S. If, however, someone wanted to compare the population of all the counties in a state, they would have to display and print out a separate report for each county. New developments will allow users to compare variables across geographic areas in "area comparison reports." Instead of choosing a state, county, or city and then displaying a statistical report, users will be able to display a list of variables available for a given database and then generate a report on that variable for all cities or counties in a state or all states in the U.S. The results can be displayed alphabetically or sorted by rank. This feature is already available for the USA Counties database and will be applied to more databases in the future. Next up will be the 1990 Census of Population and Housing.

The ability to create area comparison reports has enabled project staff to develop another new feature: online mapping. One of the site's most popular databases is USA Counties, a compendium of statis-

tical tables on dozens of topics at the county level. The CD version contains mapping software that allows users to create maps of the statistical data contained on the CD. GISP research assistant Ron Stillinger and graduate research assistant Matt Gregory have developed a way to allow users to access this feature from the Web. With ArcInfo running on the Web server, the area comparison reports from USA Counties can be displayed as thematic maps. An example is the report listing the population for all the counties in Oregon, shown below.

One of the problems with creating maps over the Internet is speed. Whereas text data reports can be generated within a few seconds, maps often take minutes to display. The main time lag GISP programmers discovered was that the ArcInfo program created the map in Postscript format, which then had to be converted into GIF format to be displayed on the Web. By testing several different programs for converting the images, they were able to select the quickest one and pare down the time it takes to generate a map. The current prototype takes 25 seconds



Map showing the population of Oregon counties, as derived from USA Counties.

to generate a map on the site. Times may vary for remote users, however, due to Internet traffic and local connection speeds. The mapping function for USA Counties should be available this summer.

Another area of development is in the expansion of data available in the 1990 census database. The GISP currently contains reports only for states, counties and places. The data on the CDs, however, go down to the tract, block numbering area (BNA), and block group level. Tracts and BNAs are small statistical subdivisions within counties. Tracts are used for metropolitan counties and usually contain from 2,500 to 8,000 people; BNAs are used in less densely populated counties that have not yet established

tract boundaries. Block groups are even smaller subdivisions within tracts and BNAs. Information at this level of detail is invaluable for providing a picture of small areas, which is not usually reported in the standard statistical compendia. Users can compare such variables as income, race, or age across neighborhoods within a city or areas within a county. Qing Yang, the former graduate research assistant at the GISP, has written the programs necessary to provide reports at the tract, BNA, and block group level for the 1990 census. Only more hardware is needed to be able to mount all the CDs necessary to provide the information on the Web site. But because tracts and blocks lack any locally recognizable names (they are designated only by numbers), they are usually meaningless to most people unless they have a way of locating them. Another enhancement being considered for the future is to provide image maps which show the location of tracts and block groups. Users could then simply click on the tract or block group and display the data for it.

The GISP began with one goal in mind: to demonstrate that providing a user-friendly interface for remote users of the library's statistical information on CD-ROM was possible. It has become increasingly clear that maintaining such a service is a continual challenge. Many of the CDs are updated each year and often file formatting changes require extensive reprogramming. This, in addition to the desire to enhance the features and interface design, means that providing Internet access to data is not a simple affair.

The state of Web publishing is still very much in its infancy. Certainly, however, the Web provides a great opportunity for increasing access and usability of government statistics. In the past, census data in machine readable format was only available on computer tapes and was used only by researchers at universities or large organizations that could afford to acquire them and provide the computers to analyze them. Now the Internet and the prevalence of personal computers make the data available to a much wider audience. Not only is the data more accessible, it is easier to use. Searching for the data on a Web site is often much quicker and more convenient than trying to sort through the hundreds of printed reports published by the Census Bureau.

More and more, Internet access is taken for granted. The Census Bureau itself has begun publishing its data directly on the Web and is planning to use the Web as its main avenue for disseminating information from the 2000 census, drastically reducing the number of printed reports produced. But Web applications, while an improvement over many of the traditional sources for census data, still have a ways to go. The experience of the GISP bears this out. The 11 databases and millions of reports now available on the site are just a subset of the data contained on the CD-ROMs. Making Web applications for statistical information more complete, flexible, robust, and innovative will certainly continue to be a challenge for the future. ☐