PUBLIC HEALTH GIS AND THE INTERNET

Charles M. Croner
Office of Research and Methodology, National Center for Health Statistics,
Centers for Disease control and Prevention, Hyattsville, MD 20782; e-mail:
ccroner@cdc.gov

KEY WORDS: National Spatial Data Infrastructure, geospatial, metadata, Department of Health and Human Services

ABSTRACT

Internet access and use of georeferenced public health information for GIS application will be an important and exciting development for the nation's Department of Health and Human Services and other health agencies in this new millennium. Technological progress towards public health geospatial data integration, analysis, and visualization of space-time events using the Web portends eventual robust use of GIS by public health and other sectors of the economy. Increasing Web resources from distributed spatial data portals and global geospatial libraries, and a growing suite of Web integration tools, will provide new opportunities to advance disease surveillance, control and prevention, and insure public access and community empowerment in public health decision making. Emerging supercomputing, data mining, compression and transmission technologies will play increasingly critical roles in national emergency, catastrophic

planning and response, and risk management. Web-enabled public health GIS will be guided by Federal Geographic Data Committee spatial metadata, OpenGIS Web interoperability, and GML/XML geospatial Web content standards. Public health will become a responsive and integral part of the National Spatial Data Infrastructure.

INTRODUCTION

GIS deployment through the Internet (also World Wide Web, Web) is a relatively new technological development. The remarkable increase in use of the Internet is creating new standards, and challenges, for the efficient use of Web-based geospatial applications (56). Challenges include spatial scale, size of data files, data compression and transmission, and other conditions for the extensive use of GIS functionality. For public health applications, geospatial databases created for the Web will have the added requirements to a) meet confidentiality safeguards to insure the anonymity of the individual from disclosure (20, 52) and b) insure Section 508 compliance with the Rehabilitation Act Amendments to make complex graphical and mapping files accessible to visually-impaired users (28). GIS and Web technologies offer emerging opportunities to analyze complex geospatial data, solve problems, and present data in a graphical format that public health decision makers and the public can easily see and understand (65).

Prior to the Internet, data "stovepipes" or data storage silos were characteristic of public health and other government agencies (103). Gaining access to data holdings, especially geographic or geospatial data, was a rigorous process for all but the most knowledgeable users. The Internet is breaking down this stovepipe legacy. Accelerating the transformation at the highest levels of policy making is the growing recognition that cost-effective development of, and access to, geospatial information is essential to the successful operation of government and the nation (5).

The federal government now supports the premise that digital geospatial data constitute a capital asset (34). That is, the return on geospatial investment can be highly cost effective through the one-time development of geospatial data, and the subsequent sharing of that data among many users, at all levels of government and all sectors, over time. This concept has the support of the federal government's Office of Management and Budget (OMB)(82). Through OMB, the Federal Geographic Data Committee (FGDC), within which the Department of Health and Human Services (DHHS) is a member, has lead responsibility for the orderly deployment of the nation's National Spatial Data Infrastructure (NSDI). According to Executive Order 12906 and OMB Circular A-16, Federal agencies are required to make geospatial data comply with existing FGDC-endorsed metadata standards and make that data available to the public (32).

Additionally, one of the most recent OMB E-Government initiatives, Geospatial One-Stop, is intended to revolutionize electronic or E-government by providing a geographic component for use in all Internet-based government activities across local, state, tribal and Federal governments (31, 33). For DHHS data users, identification of existing geospatial coverages from an online inventory of DHHS databases will provide immediate discovery and access to geospatial metadata information. In essence, Geospatial One-Stop will spatially enable the delivery of government services and eliminate the redundancies of costs associated with geospatial data collection, production and dissemination.

Collectively, these developments provide a significant catalyst for the advancement of public health GIS and the use of geospatial data through the Internet. They provide timely stimulus for the delivery of public health geospatial information for community, state and national uses. They portend important changes, based on new emerging geospatial and Web technologies, for the nation's public health geospatial infrastructure in the Twenty-First century.

WHY WEB-ENABLED GIS FOR PUBLIC HEALTH?

The Internet is becoming the most efficient means available for electronic communication of information and data (14) and the technology is evolving rapidly. Accommodation of geospatial data is no exception, even though constraints persist on bandwidth, transmission speed and integration (44). The Internet can now provide timely access to geospatial information. It's estimated that at least 80-90 percent of all government databases contain georeferenced information (35) meaning data can be tied to a specific location or place such as area code, latitude and longitude, street address, and many other Census and political boundaries. This suggests the Internet will, in fact, revolutionize our perception and use of geographic and georeferenced information.

While the nation has experienced a variety of catastrophic events over time, the terrorist attacks on the World Trade Center and Pentagon, September 11, 2001 (including related airplane hijacking over Pennsylvania), and subsequent anthrax bioterrorism, dramatized the urgent need in all sectors of the economy for comprehensive geospatial preparedness for emergency planning and management. Because public health data are so integral to surveillance, response, management, mitigation and prevention of adverse human health events and conditions, Internet strategies that can assure secure timely geospatial health data delivery and analysis, for routine and crisis circumstances, now must be viewed as high public health and national priorities. Preparations for readiness and Web deployment of public health geospatial data resources, in 2002, are in a beginning stage.

MOBILIZATION OF PUBLIC HEALTH FROM BOTTOM UP

There are several tasks required in order to reach the goal of comprehensive Internet geospatial readiness in public health. Perhaps the starting point is to empower all public state and local health departments (LHDs) with basic geospatial technology, tools and expertise. Insuring that state and LHDs have, or have access to, the needed technology and training is a key investment for developing the nation's public health geospatial infrastructure-- from the bottom up. Health database holdings containing geographic or spatially referenced information then become essential infrastructure content, standardized for Web interoperability, and cost-effectively shared.

Public health is beginning to engage geospatial infrastructure-building tasks. Geocoding is one of the most essential (3, 17) and the translation of address information into corresponding latitudes and longitudes in health databases is a national public health goal (69, 108). Findings from a 1997 survey confirm the growing level of interest in geocoding activities among the nation's State Vital Statistics Offices where 21 of 49 project directors conduct some type of automated geocoding of vital statistics data (58). Respondents from 93 percent of state vital statistics registration areas that did not geocode expressed interest in obtaining this capability.

In contrast to states, geocoding activities may be less developed among LHDs. Many LHDs today still may lack resources for GIS hardware and software as well as other related GIS investments e.g., training, staff expertise, budget, etc., in support of geocoding and other basic geospatial data activities (24, 89, 97).

Geospatial readiness for state and LHDs also requires that geospatial data holdings be documented, described and Web-searchable in a standardized manner. Only then will there be a truly navigable national inventory of existing public health geospatial data resources. DHHS can be expected to play a key role in this development by defining and requiring inclusion of a spatial metadata component searchable in all public health databases.

As a model, the federal government has provided leadership by publishing metadata standards that document and describe geospatial data holdings (32). The OMB requires that federal agencies, in FY2003, meet FGDC standards for all geospatial data collection activities. Moreover, federal agencies will insure that these data as defined in Executive Order 12906 (April 1994) be collected, produced and disseminated in accord with FGDC standards, before any funds are obligated for activities related to these data.

Adaptation to federal geospatial metadata standards by state and LHDs will go a long way towards building a uniform nationally-integrated geospatial data resource. Once Web-enabled and metadata searchable, state and LHDs will possess a new potential for readiness and response. From OMB's perspective, it is critical that geospatial data assets be created, well maintained, readily available to those who need them, and interoperable (40).

In a 2002 survey of New York State's GIS Coordination Program, members (of which the majority were state, county and local representatives) were asked to suggest up to 3 ideas for low cost GIS initiatives for implementation in the coming year, keeping in mind limited funding. The four leading response categories, in order of magnitude, were: (1) promote the development of data sets and standards, (2) continue training and education programs, (3) collaborate with all sectors to share data, applications and expertise, and (4) expand Internet data and services (73).

The foundation for the NSDI is being built at state and local levels. An NSDI guiding principle, "build once, use many times," is crucial at these grassroots levels of government. Here geospatial data form the basis for truly integrated applications both within and across agencies and as a means of leveraging scarce resources. Because public health budgets can be especially limited, partnerships and external agreements for data sharing will help support many agencies and health departments create an Internet GIS presence.

CONFIDENTIALITY OF PUBLIC HEALTH GEOSPATIAL DATA

For public health, a key constraint to the release of geospatial data on the Web has been data confidentiality and the protection from any unauthorized disclosure, through location, of an individual's identity. All health agencies, including the Federal government (48), are highly sensitive to any possible public release of data containing geographic identifiers that could lead to the identification of an individual, without some protective and thorough prerelease screening of the data. Public trust is essential to the conduct of government-supported data collection, analysis and dissemination.

Preparing and sharing data for GIS mapping creates an additional level of complexity to these concerns. GIS tools easily can layer, parse and spatially reduce geospatial information from an unlimited number of databases and potentially uncover unique geographic locations on a map. Usual approaches to safeguard data against disclosure include temporal or spatial aggregation, smoothing, and other masking techniques (91).

However, new cautions for public health researchers must be exercised as in the case of using geocoded records with Zip Codes. ZIP Codes and newly created Census ZIP Code Tabulation Areas (ZCTAs) sharing the same 5-digit code may not necessarily cover identical areas. The potential for spatiotemporal mismatches and privacy disclosures stemming from the replacement of Zip Codes with ZCTAs in the 2000 census will require thorough examination and new confidentiality guidelines for data release and GIS analysis (15, 52, 53, 94, 109).

There are few, if any, exceptions permitting public identifiability of individuals in public health geospatial databases. One recent departure exists in the related field of law enforcement and public safety. Under the 1996 “Megan’s Law” (8), and within the legal interpretation of the U.S. Justice Department, residential information on registered child sex offenders can be released to the public by local government (70).

This information can be displayed with GIS static proximity maps and posted on the Internet (Figure 1)(42). Although cautions appear that misuse of the information may subject persons to civil damages, this use of GIS and the Internet raises potential community concerns and impacts about mistaken identity, status, and location of the offender over time, safety of related family members or other household occupants, vigilantism, and other possible adverse community reactions. In public health, any public identification of an individual’s health status and residence, regardless of level of contagion or risk, is prohibited (20).

----------------------- INSERT FIGURE 1 About Here------------------------------

Nevertheless, and with possibly the exception of public safety under Megan’s Law, the need for efficient methods to protect individual disclosure, but not inhibit research from geospatial information on the Web, occupies an important area of public health GIS study. While data aggregation, smoothing and other data masking techniques normally are employed to help protect and insure the individual's confidentiality in publicly accessible geospatial databases (61, 62, 68), these often can create limitations of spatial scale for effective small area analysis (51) and, in instances, still have an affect on individuals. For example, aggregated environmental risk data can indirectly lead to elevated community costs for individuals if an exposure identified for that community is considered harmful (16).

In the state of Maryland, the Governors's Office of Crime Control and Prevention, Department of Juvenile Justice, is using Internet GIS to provide a statewide service to identify and assist high-crime and at-risk neighborhoods (63). This approach is intended to help in the fight against crime. However, disseminating crime maps to the public could contribute to informal redlining e.g., identification and exclusion of high risk neighborhoods, once used by insurance and banking companies, and result in residential flight and further consequence to problematic areas (104)

In the absence of any one standardized approach to geospatial data confidentiality, many agencies are contributing a wide range of geospatial products for Web use carefully designed for both private and public access. For example, one new Internet GIS tool designed primarily for researchers is the Geographical Information System for Health (GIS-H), National Cancer Institute. It is being developed as part of the Long Island Breast Cancer Study Project (LIBCSP) and addresses data access confidentiality issues (67).

Epidemiologists and other researchers may gain access to the GIS-H Data Warehouse, software and tools through an approval and coordination process. The "Researchers" area of the LIBCSP website provides applications necessary for access and use of the GIS-H non-public resources which are subject to privacy and licensing restrictions. For all others, anyone has access to data approved for public dissemination, online interactive maps, software not subject to licensing restrictions, and other information incorporated into the GIS-H that is not restricted by privacy limitations. Similarly, the Primary Care Service Area (PCSA) project, Health Resources and Services Administration, screens online users for approved access to viewing PCSAs e.g., one or more area ZIP Codes grouped according to utilization patterns for Medicare beneficiaries receiving primary care services, and their attributes.

The Washington State Health Department's online developmental Epidemiologic Query and Mapping System (EpiQMS) incorporates three levels of security in order to accommodate citizens, public health and medical practitioners, and public health agency investigators access to state and regional health data. This security model allows different levels of access to the data depending on the likelihood that an individual's privacy could be compromised. For example, citizens can query the application to see what the breast cancer rate is in their county, ZIP Code or census tract; medical practitioners can determine the sexually transmitted disease rate in their community; and a health agency epidemiologist can look at the historical trends of *E coli* infections in a county or town, or school district (106). Ordinary rates and confidence intervals, and adjustments for small numbers, are included. This online application in the future will offer basic cartographic displays and other analytic data methods for users such as Bayesian smoothing.

Finally, new special national data screening precautions related to privacy of patient medical records applies to public health care providers (27). The Health Insurance Portability and Accountability Act (HIPAA) of 1996, and its new DHHS privacy provisions (De-Identification of Protected Health Information - § 164.514(a)), apply to health information created or maintained by health care providers who engage in certain electronic transactions, health plans, and health care clearinghouses (78). This standard pertains to “de-identification” of protected health information requiring that health information not identify an individual and that no reasonable basis exists to believe that the information can be used to identify an individual. HIPAA privacy rules contain extensive exemptions if the identification information is used for treatment, payment, research, or national priority activities that are carried out in the interest of public health and safety (25). There are no studies, currently, to suggest that DHHS privacy provisions would have any adverse effect on geospatial and GIS data activities.

It is important for public health agencies to continue to develop confidentiality guidelines for their statistical database holdings (29) and the sharing and use of geospatial data. These and other guidelines, including accessibility restrictions to the public and research community, can be Web documented in metadata that describe essential elements of the database. Through a searchable Web environment, these metadata are the building blocks of information dissemination. Metadata provide all public health agencies the opportunity to inform others of their geospatial data holdings and any limitations associated with their use. DHHS can play an important lead role for state and LHDs by providing standardized guidelines and rules for data confidentiality in Web accessible geospatial databases.

PUBLIC HEALTH GEOSPATIAL DATA ON THE INTERNET

The availability of public health geospatial data on the Web is growing. Most of these databases are available as either static or dynamic mapping products. Static public health related maps do not contain GIS functionality. However, they often combine small area boundaries (e.g., Census tracts, ZIP Codes) and other information layers (e.g., income and race, disease) making useful displays for a wide public audience (4). The online cancer incidence maps of New York State

Cancer Surveillance Improvement Initiative provide one of the more detailed static Web GIS displays of geographic area and disease outcome. These maps compare individual ZIP Codes with expected cancer incidence (68). Where downloadable, static display data from a source geospatial database could be prepared for use in a GIS.

In contrast to static presentations, geospatial maps become dynamic when users are allowed to access, or interact with, the database from their own computer (10, 12, 17, 66). Users can customize maps and tables and interactively query the database to search for features based partly on their own criteria. These allow for a wider, but predetermined, selection of parameters and tools for geospatial analysis.

Web interactive cancer mortality maps, designed with MapInfo GIS, have been developed by the National Cancer Institute (NCI), National Institutes of Health (NIH) and present users choices about type of cancer (41 sites), age, race (black and white), sex, geography (state, State Economic Area or county), and selection of class intervals, color shading and scaling (71). Future plans may incorporate a module that calculates rates on the fly, combining user selected cancer codes with other variables for mapping or summaries (D Grauman, personal communication). Charts and graphs associated with the maps translate graphical data into a comparison form accessible by screen readers and are Section 508 compliant for those with visual or manual impairment.

At the US Environmental Protection Agency, EPA Enviromapper allows the user to include map features, including hospitals, which may be layered or spatially associated with other EPA environmental exposure databases. For example, EPA's National Priorities List database provides the identification of pollutant remediation sites within a wide range of geographic boundaries or by latitude and longitude (99). Similar to NCI, EPA graphics are considered to be Section 508 compliant.

While Web-based public health mapping products that offer an expanded range of GIS functions are emerging, several conditions, some technological, currently influence and challenge robust use of Web GIS functionality. Web geospatial interoperability, geospatial data transmission and automated conflation of geospatial databases are some of the key conditions.

Interoperability

Today, the promise of interoperability whereby geospatial data distributed anywhere on the Web can be searched, located, retrieved and compiled, either by a Web GIS service provider or at an individual's desktop, is becoming reality. This is a significant accomplishment given the long-standing lack of industry consensus about hardware platforms, operating systems, network protocols and programming languages in support of Web GIS use (93). In fact, specifications are now being developed to accommodate these operational differences and allow Web GIS clients and desktop users to fully integrate Web accessible geospatial data resources.

The effort to "geoenable" the Web is being led by the OpenGIS Consortium (OGC). The development of common ground for integrated geospatial mapping applications (87) depends on interface specifications designed to enable GIS interoperability of geospatial information regardless of operational differences in the vendor environment. Geography Markup Language (GML) is the base language developed by OGC, and GML is becoming the world standard for eXtensible Markup Language (XML) encoding of geographic features and geoprocessing service requests (80). OGC standards for interoperability will advance Web-based use of geospatial data.

Data Transmission

Timely data transmission is another condition influencing efficient use of Web geospatial resources. Expeditious access to geospatial data resources depends on transmission capabilities. Bandwidth is a key component of the transmission process. Over the next few years, as the amount of unused optical fiber infrastructure increases, bandwidth is expected to become a commodity (46). Commoditization implies a tremendous amount of flexibility, a high rate of change in topology in the way things are connected together, and requires a very robust infrastructure (13). This development should result in increased bandwidth for many locations in the U.S. and improved geospatial data transmission speed.

Increased bandwidth will benefit public health emergency response situations. In the emergency response to the September 11, 2001 World Trade Center attack, lack of bandwidth in some areas of New York City (NYC) resulted in delays in providing processed and urgently needed orthophotography for the Emergency Mapping and Data Center (EMDC). Because of low bandwidth Internet connections, large data files had to be written to CD-ROM and driven by state police twice daily from Albany to NYC for delivery to the fire department, Federal Emergency Management Agency (FEMA), and EMDC (6). For public health, a variety of similar rapid developing emergency related events, including floods, fires, chemical spills and earthquakes, necessitate timely Web delivery of large geospatial databases for responsive disaster intervention and control.

Conflation

In addition to Web geospatial data interoperability and transmission, a key property of Web GIS enablement involves conflation, or the ability to precisely georeference variant data layers compiled into one view, which can be crucial in emergency situations such as terrorist and bioterrorist attacks. The need currently exists for the development of automated conflation techniques transparent to the user.

Once again in the case of NYC, lack of automated conflation methods following the November 11, 2001 terrorist attack on the World Trade Center resulted in time-consuming problems for emergency response teams. This omission created significant delays in the ability of the NYC GIS response team to provide integrated GIS map views for precision excavation, below ground entry, and other search and rescue functions dependent on above and below ground composite views (55). NYC will build automated conflation capability this year by modifying all city planning geospatial databases to include standardized “hooks” such as street centerlines, building footprints, and address and parcel locations for common identification and seamless linkage.

Other more basic challenges exist for state and LHDs. These span the acquisition of software tools and expertise necessary to support GIS functionality to investment considerations related to Web-based GIS system design, transaction and administration (41). For most of our nation’s 3,200 or so health departments (90) the need to leverage scarce resources with partners may be essential to developing a geospatial data presence on the Web.

PUBLIC HEALTH GIS FUNCTIONALITY ON THE INTERNET

The availability of public health geospatial data, in a robust GIS functional Web environment, is in a nascent state. By comparison, many non-public health agencies have an extensive Internet presence with digital geospatial data. Many of these agencies, including EPA, USGS, FEMA, Bureau of Land Management (BLM), U.S. Forest Service (USFS), National Oceanic and

Atmospheric Administration (NOAA), National Aeronautics and Space Administration (NASA), Department of Transportation (DOT) and Census Bureau have long-established geospatial products and inventories related to georeferenced themes, as part of their institutional mission.

Intranet GIS Functionality

The Intranet is helping many of these agencies to develop enterprise GIS activities. The Intranet provides advantages for secure sharing and analysis of restricted geospatial information within respective agencies. These activities may provide operational models for public health where scarce GIS resources can be similarly leveraged. Enterprise GIS may become a workable strategy, as evidenced in a variety of non-public health settings, to unify GIS services for state and LHDs.

For example, in Oregon and Washington states, new generation client-server based GIS architecture will be installed for BLM employees to support a spatially-enabled workforce and statewide enterprise GIS. In addition to secure internal agency use, it eventually will allow accessibility to Web-based GIS data and services to all statewide employees by placing simple Web-based GIS applications on any employee's desktop through a web browser. The BLM enterprise GIS utilizes a new data storage structure, called a "geodatabase," which will allow BLM to centrally store and serve GIS application data to employees (1). It also is intended to facilitate establishment of uniform data standards, metadata documentation and database security throughout the agency.

As state and local public health departments build towards a Web presence capable of utilizing GIS services, the Intranet may provide a startup and more affordable scalable approach. In partnership with other state and local agencies, secure public health Intranet services might offer a cost-effective means for geospatial data sharing and application through resident GIS technology and expertise. State and local planning departments, especially, appear to be more GIS operational at this time than others (S Warnecke, personal communication) and may offer partnering opportunities for public health.

In the case of the District of Columbia (DC), the DC GIS Atlas initiative was designed for employees at all levels of District government to have an Intranet access point to standardized, updated GIS data, eliminating the need to search disparate resources for critical information (84). The atlas consists of a collection of thematic mapping modules (e.g., public safety, transportation, socioeconomic, environmental, and others) that can pull information from a central GIS server as well as from the databases of all district agencies, depending on what kind of queries employees request. The atlas now offers over 130 map layers for the "average" district employee and another 16 layers (available via a password-protected emergency management module) for authorized personnel. Several more layers are under development.

A point-notification tool exists that allows DC Atlas users to encircle/polygon areas of interest and receive all address/owner information for these areas from the district's tax and revenue database. This enables users to easily generate mailing lists and a variety of other notification services. Another powerful tool is a reporting capability that allows users to compile all available information for a particular area (point or polygon) as well as link to the various agency Intranet sites from which the information originated. And, a mapping tool allows users to plot data from any district database and perform "on-the-fly" geocoding (43).

Part of the DC Atlas will become Internet-enabled sometime this year to provide full customer service with access to citizens. This extension to the Internet will comprise a more limited selection of data layers from the centralized GIS database but help serve routine citizen needs concerning public health, maintenance, city services and crime (M Sherman, personal communication).

Internet GIS Functionality

Partnering with other agencies also may be a worthwhile approach to assure the public's access to public health geospatial data on the Internet. Several government Internet GIS developments are noteworthy because of the GIS functionality they offer and their accommodation of public health data.

The online USGS National Atlas of the United States (or National Atlas) provides an interactive GIS Web mapping environment that allows users to choose and layer a wide variety of earth science geospatial databases. It contains several public health data layers accessible to the public. These include a map layer of U.S. mortality (68) with information for eleven leading causes of death, eight subset causes, and all mortality causes combined for Health Service Areas in the U.S. Another mortality map layer is included which contains 1970-1994 cancer mortality information for the United States (100). Much of this latter information is county based and includes death rates, number of deaths, confidence intervals, and expected number of deaths for white males and females for 20 cancer sites.

Additionally, U.S. West Nile Virus (WNV) activity for 2000, reported by CDC to USGS, is posted in several atlas locations, each with distinct GIS mapping presentations (101). The 2000 surveillance data offers five surveillance categories including humans, wild birds, mosquito, veterinary and sentinel flock e.g., chickens placed in cages exposed to mosquitoes. Users can track these events interactively, by week of reported occurrence and cumulatively, for county locations (Figure 2). This WNV space-time surveillance and GIS mapping component of the

USGS atlas serves as a useful model for dynamic representation of disease surveillance data on the Internet.

------------------------INSERT FIG. 2 ABOUT HERE

Cumulative 2001 WNV surveillance data will be included in the atlas when released by CDC to USGS (J Donnelly, personal communication). Other related public health data, such as arsenic exposure and contamination of groundwater, are new themes also being prepared for inclusion. The USGS site offers a Map Data Layers Warehouse where users may download National Atlas geospatial map layers at no cost via file transfer protocol.

The U.S. Environmental Protection Agency (EPA) offers an online GIS "Window to My Environment" (WME), designed to provide public accessibility to a wide range of federal, state, and local geospatial data about environmental conditions and features in any U.S. location (Figure 3). It is a broad reaching application in partnership with federal, state and local government and other organizations representing a collection of Internet services including the Census Bureau, USGS, FEMA, NOAA, and EPA.

--------------------INSERT Figure 3 About Here------------------------

The data one finds in the WME application are distributed and reside at their respective agency servers (18). Thus each participating agency manages its own data and its timeliness, which can be current and even real-time, as in the case of USGS water monitoring information from selected surface-water, ground-water, and water-quality sites (102). EPA software with Environmental Systems Research Institute, Inc. (ESRI) ArcIMS and Oracle database functionality drive the WME interactive mapping and associated statistical reports by transparent data extraction from the partnering agency servers.

There is no limit on the number of WME partners and one can conceive of any agency as a participant. With EPA technical assistance, participating partners can create a reciprocal interface on their home server with WME connectivity. Public health databases are not yet included in the WME system and there are no specific barriers to inclusion. Many of the geospatial data layers within WME are linked with the USGS National Atlas. WME potentially offers wide national visibility, at minimal cost, for both large and small scale public health geospatial applications.

One notable WME GIS feature is that the user defines the dimensions of the spatial window that appears on the opening screen. The window is associated with a variety of Census statistics. The summary demographic report that accompanies the spatial window changes with window size using an area weighted algorithm to allocate the proportion of truncated Census boundaries to the estimated residential population count for that window. An updated WME is planned that incorporates a variety of user-defined GIS metric functions including distance, radii, and buffers (D Wolf, personal communication).

In a related development, the Department of Housing and Urban Development (HUD) is building an Enterprise Geographic Information System (EGIS) with ESRI. The EGIS will provide users with easy access to mapping tools and HUD geospatial data to support housing and community development programs at the state, county, city, and neighborhood levels (49). Users will be able to enter an address or click on a map and have the application take them to a map of that location and display any combination of HUD Housing and Community Development, FEMA, EPA, and Census data for the location and surrounding area (J Sperling, personal communication). It also is expected that HUD data will link to EPA's WME, providing the WME application a key housing data layer, maintained and updated at HUD.

In England, the Multi-Agency Internet Geographic Information Service (MAIGIS) project is a 3-yr pilot effort funded by the Public Health Development Fund to establish an interactive map-based web site for sharing health and health-related data for the West Midlands Region (95). Data sets within the MAIGIS project follow three broad themes of health, socioeconomic and environmental information and are made available by different organizations and shared using geography as the linking theme. The web site has an interactive mapping capability that includes selected cancer mortality and infectious disease rates (e.g., cumulative HIV, TB and meningococcal infection) (85).

PERSPECTIVE ON FUTURE DEVELOPMENTS

With the nation now focused on the growing urgency to build the National Spatial Data Infrastructure (NSDI), the Internet can be expected to play an increasingly central role in its deployment. There will be new opportunities and initiatives to build a geospatial infrastructure in public health. A truly dynamic system of public health readiness will depend, in part, on key developments in technology that make use of the Internet in support of a geospatial infrastructure.

Robust public health GIS Internet interoperability will depend on consensus approaches to dynamic exploration of occurring statistically complex space-time data, satellite data delivery and mobile geoprocessing (107), distributed data mining (98), advanced forms of wireless communications (30, 47, 50), and data sharing and publishing within accepted standards of data security and privacy. Data intensive applications will be designed to exploit the high bandwidth provided by emerging domestic and international networks so that multi gigabyte, even terabyte, data sets can be remotely explored in real time (21).

Geospatial Enablement of the Internet

As suggested earlier, developing the World Wide Web for geospatial information may well depend on applying data format independent eXtensible Markup Language (XML) to geographic information. XML encoding of geodata, using OGC's GML and Open Web Services (OWS) standards, makes it possible to display, overlay, and analyze geodata on any Web browser, even if the browser obtains views of different map layers from different remote map servers (88). XML and XML Web Services are designed for integrating diverse types of content in a distributed enterprise. XML is emerging as the universal standard for preserving and communicating electronic information for structured documents and data on the Web (38, 76). Essentially, XML is data tagged with metadata about its own organization.

XML has important bearing on recent major government initiatives. FirstGov.gov will rely on Web search engine-based technology and a host of integration services (37). OGC will issue a request this year to fund solutions for the implementation of a Unified Modeling Language (UML) to GML tool that will be available for later use in future Geospatial One-Stop framework standards efforts (86).

Homeland Security will benefit in the case of emergencies common to adjoining but politically independent geographic areas. XML can integrate independent data silos to respond to questions through the layering of Web services (57). EPA Regions and Program Offices have begun to pilot the use of XML for geospatial metadata and the use of GML/OWS for the integration of diverse and distributed geospatial data. The incorporation of XML and XML schema into the design of EPA's National Environmental Information Exchange and CDC and ATSDR's National Environmental Health Tracking Network is essential to insure interoperability and not just more silos of replicated, warehoused, and/or disconnected information repositories (75, 77).

The Census Bureau will work with the U.S. Army Topographic Engineering Center, to develop and test OGC interoperability specifications for mapping and geospatial data exchange on the Web. This pilot activity will allow the Census Bureau to encode the Topologically Integrated Geographic Encoding and Referencing (TIGER) data in GML and implement a Web-based data update capability for TIGER (2). This activity will serve as a major stepping stone towards building a mature street centerline spatial database for the nation (59, 60) and is one of many designed to support the GeoSpatial One Stop initiative.

In support of GML geospatial databases, agencies will insure a new era of Web-enabled GIS readiness and fully accessible information communities. The "Geo-Web" concept may become a reality where the Web and Geo-Web fuse into one seamless, completely integrated application (54).

Geospatial One-Stop

OMB's "Geospatial One Stop" E-Government initiative will impact on DHHS. NSDI provides the groundwork for a geographic information component for E-Government. It provides a geographic component for use in all Internet-based government activities and an online point of entry to geospatial data. Geospatial One Stop is designed to support "one-stop" access e.g., citizens will have only to go to one Internet location, for local, state, tribal and federal geospatial data assets, and industry web-mapping portals and other private sector businesses (33). This information includes geographic or spatial data that references a specific location, such as latitude and longitude, county, street address, ZIP codes or place name.

The project will accelerate the development and implementation of the NSDI. Because state and local governments, combined, are estimated to collect more than twice the geospatial information as the federal government, they will be asked to play a leading role in the initiative. Improved collection, sharing and use of geospatial information in programs at all levels of government have the potential for improved decision making, greater efficiency, and cost savings.

Several recommended Geospatial One Stop tasks are timely for DHHS in support of public health geospatial mobilization and readiness. These include publication of metadata records of geospatial database holdings and reports, which also would populate the Geospatial One Stop Clearinghouse Network (e.g., a distributed network of data producers, managers, and users linked electronically over the Internet). Through the Clearinghouse, public users benefit from a single interface to search and access metadata and data for themes of interest. The Clearinghouse will include: those sites across the country where the metadata and data are stored, usually at the site of the producer or intermediary; development and deployment of prototypes for enhanced data access and web mapping services for geospatial data; and establishment of a comprehensive electronic “Portal” as a logical extension to the NSDI Clearinghouse Network (82).

A recently developed Web portal serves as a gateway to DHHS data and statistics (26). It has been described as a DHHS meta directory containing a comprehensive listing, with descriptions of the statistical and surveillance systems supported by DHHS agencies. It also includes links to state health agency data sites. One of these, FedStats, contains geospatial data in a MapStats mapping component with links to EPA’s hazardous waste sites, ATSDR’s health assessment locations, and other health related sites (36). Information is included about geospatial data activity, such as the development and maintenance of a comprehensive inventory of telemedicine projects funded by the Federal government (39). Although neither a geospatial metadata element or GIS map server is included in this beta version of the DHHS portal, a geospatial metadata information/data element will be proposed for inclusion in the next iteration (R Poulson, personal communication).

Emergency Planning and Response

Geospatial readiness in emergency planning and response is one of the current driving forces behind geoenablement of the Internet. The recent terrorist events at home, and subsequent proposed establishment (at this writing) of a Department of Homeland Security, will result in key initiatives and motivate--with restrictions on public access to selected spatial data--a national blueprint for a geospatially integrated and interoperable Internet.

One development integral to public health is the Multi-Sector Crisis Management Consortium’s (MSCMC) goal of managing the threat of harm to humans from natural and man-made disasters, including terrorist acts. Beginning in 1999, the MSCMC was formed through the

Alliance Center for Collaboration, Education, Science and Software (ACCESS), an entity of the National Center for Supercomputing Applications (NCSA), in collaboration with the National Response Center, an entity of the United States Coast Guard (USCG). USCG (which could become part of the proposed Department of Homeland Security) serves as the sole national point of contact for reporting all oil, chemical, radiological, biological, and etiological discharges into the environment anywhere in the United States and its territories (64).

Uses of advanced integrated Internet technologies are consistent with MSCMC policies, goals and objectives identified in the recently released White House Executive Order, which calls for the establishment of a Global Disaster Information Network (GDIN) (45). In the U.S., GDIN will enable disaster managers, relief workers, and others to access information systematically and rapidly to help mitigate natural and man-made crises. MSCMC supports development of a computational grid and high bandwidth environment over which supercomputing and data mining technologies will operate (22).

Another related public health development is the Multi-Hazard Mapping Initiative (MMI), sponsored by FEMA. The MMI is a pilot study to develop an OGC standards-based framework of interoperable services in support of FEMA multi-hazard mitigation, response and recovery functions. It includes federal, state, and local agencies and other organizations supporting hazard mitigation efforts (81). Similarly, EPA's Superfund Program is building a content network of emergency response and chemical and biological counter-terrorism information to support a physical and virtual "situation room" for both emergency and day-to-day management of EPA operations for protecting human health and safeguarding the environment (77).

Web GIS services are being mobilized in many cities in the aftermath of September 11, 2001. New York City's Office of Emergency Management (OEM) is preparing online delivery of the City's Geographic Information System (72). The GIS program involves putting every street in the City into a computer program, and creating both aerial and flat grid maps of any neighborhood in the five boroughs. With the computer, OEM can highlight any area of any size and then access demographic information about the people who live in the vicinity. Available information includes population, gender, age, marital status, ethnicity, and income. The GIS program also can create a "plume model", which shows the direction in which a gas leak might flow. This model can help determine those areas requiring evacuation during an emergency.

Based on the September 11, 2001 events, there is a need for a national coordinated policy concerning Web or public availability of emergency planning and response related geospatial databases. Some of these data fall within the realm of public health. Following the September events, many federal and local geospatial databases, including local water intake locations (H Rossmeissl, USGS, personal communication), natural gas and oil pipelines (C Brandt, Department of Transportation (DOT) and D Morehouse, Department of Energy (DOE), personal communication), environmental facility databases, including Toxic Release Inventory sites (I DeLoatch, EPA, personal communication), Landview V databases (F Broome, Census Bureau personal communication), and National Imagery and Mapping Agency (NIMA) topographic maps (83) and others, were assessed by individual agencies as a potential liability to national security and withdrawn from the Internet or public dissemination.

Many publicly available geospatial databases have been restored but a few remain permanently withdrawn, at least for the foreseeable future, such as DOT gas and oil transmission and distribution points and DOE nuclear facility sites. The FGDC Working Group on Homeland

Security will address the issues of defining and establishing uniform policies on the public availability and security of “critical infrastructure” geospatial data (M Domaratz, personal communication). The current concern is to find an appropriate balance between public access to geospatial information and protection of information considered a priority for national security.

CHALLENGES TO DHHS AND STATE AND LOCAL HEALTH DEPARTMENTS

Public health needs to be an integral part of a larger structural whole where currently many other agencies are involved in geospatial mobilization and readiness. This has become even more evident following the recent terrorist attacks at home. There is a sense of urgency to bring state, city and local governments together to build, integrate, leverage through sharing and partnerships, and optimize geospatial information, both vertically within and horizontally across agencies, for comprehensive routine as well as emergency planning and response services. Intranet and Internet environments can serve to facilitate public health geospatial data accessibility and integration at both local and national levels.

In spite of the restrictions governing confidentiality in many public health geospatial databases, access to confidential data can be accommodated for qualified users in secure Intranet or Internet settings. One DHHS precedent for sharing the use of confidential geospatial public health data, in a non-Web but secure data environment, exists at the National Center for Health Statistics, CDC, Research Data Center (11, 108). An extension or variant of this model could be developed for the Intranet to provide a secure enterprise approach to data integration within federal, state and local health departments. DHHS requirements for secure enterprise geospatial data accessibility and integration will require definition. Similar rules might be established with external DHHS partners.

As a point of departure, each public health department in the U.S. will require assessment of its geospatial potential and state of readiness to respond to normal and emergency community health needs. Those that have well-developed geospatial and GIS Web accessible technologies need to be identified to the larger public health community as resources for guidance and other forms of assistance to those agencies and departments not yet or in early formative stages of involvement. More Web-enabled integration at the state health department level would help to build critical connectivity, with state and LHDs sharing geospatial information through a secure Intranet environment (TB Richards, personal communication). Scarce resources for LHDs dictate that Web-supported geospatial and GIS activities not be independently recreated by all. The development of strategies leading to the cost-effective leveraging and sharing of geospatial resources through partnerships and sharing is essential to building a responsive public health infrastructure.

Regional GIS collaborative activities offer still another cost-effective means of leveraging geospatial resources. A variety of existing partnerships can serve as potential models for public health. For example, the California GIS Council has been formed to improve coordination within California State government with input from the California GIS community. The Council works together to develop, maintain and share accurate, consistent place-based information that is accessible to all Californians and the agencies and organizations that serve them. The California GIS Council cites a variety of benefits for participation that include opportunities for local and regional government to participate as equal partners with state and federal agencies, reduced agency costs of data development, increased productivity and other benefits (7).

In San Diego County, the San Diego Association of Governments (SANDAG) serves as the technical and information regional planning agency for 18 incorporated cities and county governments (92). SANDAG makes many GIS layers available to the public on the Internet. Many of the databases are obtained through cost-effective partnerships with state, federal, and local government agencies, academic institutions, and the private sector enabling costs to be spread among the participating agencies. Satellite imagery and digital aerial photography are used in a cost effective way to update vector land cover, street networks, and other GIS databases. A SANDAG report, "Guidelines for Data Development Partnership Success," is based on many years of GIS partnering experience and cites guidelines that may help other agencies model similar partnership activities (9). Other examples of cost-effective use of geospatial databases and GIS through regional and metropolitan partnerships exist throughout the country (23, 79, 105) and in other nations (95).

Perhaps the ultimate responsibility for a public health geospatial and Web-enabled environment lies with DHHS. As the lead U.S. public health agency, DHHS will need to increase its role of overall vision and leadership--at all levels of government—to assure national public health geospatial mobilization and readiness. The agenda to advance the use of Web-enabled GIS and geographic analysis within public health practice, while extensive, needs to begin to address several noteworthy issues. These include (1) the establishment of policy and funding for the creation of local GIS to be used by LHDs in those communities that lack the resources to join in a common coherent national initiative and (2) the creation of best designs and standards for Internet-enabled GIS that will be used by state, regional and local public health agencies so that standards, definitions, and look-and-feel of the data and Web-based technology are the same throughout the nation. (96).

Guidelines for DHHS-wide uniform best designs for geospatial metadata templates, geocoding standardization, Web-mapping Section 508 compliancy for visually and manually impaired data users, and the terms, conditions and uses regarding privacy and the release of geospatial databases, are essential to building a responsive GIS public health geospatial and Web-enabled infrastructure. As these evolve, DHHS will be positioned to better support the future needs of public health (19).

CONCLUSIONS

Under the guidance of DHHS, state and LHDs can be expected to evolve into the critical foundation of a geospatial Web-enabled U.S. public health infrastructure and provide public health full partnership in NSDI. This accomplishment will generate new opportunities to advance the mission of disease surveillance, understanding and prevention, and the well being of the nation. The overriding process to document, make accessible and share geospatial information and data, in a Web-enabled environment, is perhaps the key condition of NSDI.

In order to build the foundation, every U.S. public health department and agency needs to inventory its respective geospatial data holdings and, using the appropriate identifying metadata, render these discoverable through the Internet. DHHS has begun development of a corporate or enterprise metadata repository much like those developed at the Census Bureau and the Environmental Protection Agency, and that planned at the Bureau of Labor Statistics (94). Geospatial Web-searchable metadata will help public health agencies to communicate and make known their geospatial resources to internal and external users. Secure Intranet and Internet data sharing solutions, that uphold all database safeguards of an individual's anonymity and confidentiality, can be adapted to public health.

All public health agencies must become Web enabled and have access to basic geospatial tools and training in order for public health to ultimately become an integral part of NSDI. Investment in, and by, state and local health departments through partnerships and other cost-effective data sharing mechanisms is crucial to this process. Building a comprehensive and responsive geospatial Web-enabled public health infrastructure is clearly an exciting, and achievable, challenge for U.S. public health.

ACKNOWLEDGEMENTS

Appreciation is extended to Terry A. Drizd and Jimmie D. Givens, National Center for Health Statistics, CDC, Donna F. Stroup, National Center for Chronic Disease Prevention and Health Promotion, CDC, and Thomas C. Ricketts, University of North Carolina at Chapel Hill, for their valuable peer reviews of this manuscript.

LITERATURE CITED

1. BLM. 2002. *ArcGIS Implementation Plan: Executive Summary, Bureau of Land Management, Oregon/Washington*. http://www.blm.gov/nhp/efoia/or/FY2002/IB/ib-or-2002-086Att1.htm

2. Broome FR. 2002. Geospatial One-Stop Commitments Summary-Department of Commerce, unpublished data

3. Broome FR, Meixler DB. 1990. The TIGER Data Base Structure. In Marx RW ed. Cartography and Geographic Information Systems, J of Amer Congress on Surveying and Mapping 17(1): 39-47

4. Bowen W. 2000. *American Cities Atlas Project*. http://geography.miningco.com/library/maps/blusdc.htm

5. Cahan B. 2000. Financing the NSDI: Aligning Federal and Non-Federal Investments. http://www.fgdc.gov/publications/reports.html

6. Cahan B, Ball M. 2002. *GIS at Ground Zero: Spatial Technology Bolsters World Trade Center Response and Recovery*. http://www.geoplace.com/gw/2002/0201/0201wtc.asp

7. California GIS Council. 2001. *Regional GIS Members*. http://www.gis.ca.gov/council/index.epl

8. California Office of the Attorney General. 2002. *Registered Sex Offenders (Megan's Law).* http://caag.state.ca.us/megan/index.htm

9. Carnevale S, Parrott B, Tayman J. 2001. Data Development Partnerships in the San Diego Region. http://www.gis.com/trends/sandag.html

10. Centers for Disease control and Prevention. 2002. *Injury Maps*. http://www.cdc.gov/ncipc/maps/default.htm

11. Centers for Disease Control and Prevention. 2001. *NCHS Research Data Center*. 8 pp.

12. Centers for Disease Control and Prevention. 2002. *Women and Heart Disease: An Atlas of Racial and Ethnic Disparities in Mortality.* Second Edition. http://www.cdc.gov/nccdphp/cvd/womensatlas/atlas.htm

13. Chiappa N. 2001. *The COOK Report on Internet*. http://wireless.oldcolo.com/biology/progress2000/cookjul.txt

14. Clarke KJC, McLafferty SL, Tempalski BJ. 1996. On Epidemiology and Geographic Information Systems: A Review and Discussion of Future Directions. *Emerging Infectious Dis*. 2 (2): 85-92

15. Cox, LH. 2002. Bounds on Entries in 3-Dimensional Contingency Tables Subject to Given Marginal Totals. In Inference Control in Statistical Databases. *Lecture Notes in Computer Science 2316*, ed. J Domingo-Ferrer, pp. 21-33. Berlin: Springer-Verlag

16. Cox LH 1996. Protecting confidentiality in small population health and environmental statistics. *Stat. in Med.* 15:1895-1905.

17. Cromley EK, McLafferty SL. 2002. *GIS and Public Health*. New York: The Guilford Press.

18. Croner CM. 2002. Building an Environmental Health Tracking Network: Focus on Chronic Diseases. In *Public Health GIS News and Information*, ed. CM Croner 3(45): 21-24. http://www.cdc.gov/nchs/gis.htm

19. Croner CM. 2002. *Geographic Information Science: Exciting Spatial Tools for Detecting Inequalities in Public Health*. Presented at DHHS National Leadership Summit on Eliminating Racial and Ethnic Disparities in Health, Washington DC

20. Croner CM. 1996. Geographic Information Systems (GIS): New Perspectives in Understanding Human Health and Environmental Relationships. *Stat. in Med.* 15:1961-1977

21. Croner CM. (2001). "Integrating Geographical Information with Statistical Programs: Challenges and Opportunities." *Statistical Policy Working Paper 32*, Office of Management and Budget, Federal Committee on Statistical Methodology, 244-251

22. Croner CM. 2000. The National Spatial Data Infrastructure: New Developments in Crisis Management and Response. In Public Health GIS News and Information, ed. CM Croner 11(37): 22-24. http://www.cdc.gov/nchs/gis.htm

23. Croner CM. 1999. The Washington Geographic Information System Consortium. In *Public Health GIS News and Information*, ed. CM Croner 11(31): 17-20. http://www.cdc.gov/nchs/gis.htm

24. Croner CM, Stroup DF. 1999. GIS and EIS: Geographic Information Systems and the Epidemic Intelligence Service. *EIS Bulletin* Fall/Winter 10-13

25 Davenhall WF 2002. *Protecting Geocoded Health Data*. http://www.esri.com/industries/health/hipaa.html

26. Department of Health and Human Services. 2002. *HHS Data Council: Gateway to Data and Statistics*. http://aspe.hhs.gov

27. Department of Health and Human Services. 2002. *National Standards to Protect the Privacy of Personal Health Information*. http://www.hhs.gov/ocr/hipaa

28. Dizard WP. 2002. *GIS files elude efforts to meet 508 standards*. http://www.gcn.com/21_28/news/19989-1.html

29. Doyle P, Lane JI, Theeuwes JJM, Zayatz LV, ed. 2001. Confidentiality, Disclosure, and Data Access: Theory and practical Applications for Statistical Agencies. Amsterdam: Elsevier

30. Federal Communications Commission. 2002. *Enhanced 911*. http://www.fcc.gov/911/enhanced

31. Federal Geographic Data Committee . 2002. Geospatial Information One-Stop To Spatially Enable Delivery of Government Services. http://www.fgdc.gov/geo-one-stop/docs/factsheet.doc

32. Federal Geographic Data Committee . 1998. *Geospatial Metadata Standards*. http://www.fgdc.gov/metadata/meta_stand.html.

33. Federal Geographic Data Committee. 2002. Geospatial One-Stop: Best Practices White Paper. http://www.fgdc.gov/geo-one-stop/docs

34. Federal Geographic Data Committee. Geospatial One-Stop Office of Management and Budget Capital Asset Plan (Exhibit 300). http://www.fgdc.gov/geo-one-stop

35. Federal Geographic Data Committee. 2002. *Homeland Security and Geographic Information Systems*. http://www.fgdc.gov/publications/homeland.html

36. FedStats: 2002. *MapStats*. http://www.fedstats.gov/mapstats/whatsnew.html

37. FirstGov. 2002. *Official Government Gateways*. http://www.firstgov.gov

38. Flynn P. 2002. *The XML FAQ*. http://www.ucc.ie/xml/#import

39. Food and Drug Administration. 2001. Telemedicine Related Activities. http://www.fda.gov/cdrh/telemed.html

40. Foreman MA. 2002. *Statement before the Committee on Government Reform*, Subcommittee on Technology and Procurement Policy. Work. Pap. U.S. House of Representatives

41. Foresman TA. 1999. Spatial analysis and mapping on the Internet. In *J of Public Health Manage Prac*, ed. TB Richards, CM Croner, 5(4): 57-64, Frederick, MD: Aspen Publishers.

42. Fresno Police Department. 2002. Megan's Law Map site. http://www.ci.fresno.ca.us/fpd/meganlaw/index.html

43. Geoplace. 2002. *Nation's Capital Puts Powerful GIS Tool in Employees' Hands*. http://www.geoplace.com

44. Geoworld. 2002. *Interview with Jack Dangermond, President and Founder, ESRI*. http://www.geoplace.com/gr/interview/dangermond.asp

45. Green K. 2001. *TOPOFF, Preparing for the Worst*. http://archive.ncsa.uiuc.edu/mscmc/TOPOFF.htm

46. Grossman RL. 2001 *DataSpace Fact Sheet*. http://www.dataspaceweb.net

47. Grubesic TH, Murray AT. 2002. Constructing the divide: Spatial disparities in broadband access. Papers in Reg. Sci. 81(2): 197-221

48. HHS Data Council: Introduction. 1997. *Confidentiality of Individually-Identifiable Health Information: Recommendations of the Secretary of Health and Human Services*. http://aspe.hhs.gov/datacncl/intro.htm#dcagenda

49. HUD. 2002. Enterprise Geographic Information System (EGIS). http://hud.esri.com/egis

50. Hughes D. 2002. Challenges of extending scalable wireless to remote sensor locations. January 16 meeting of the Multi-Sector Crisis Management Consortium (MSCMC)

51. Jacquez GM, Greiling D. 2002. The Geographic Distribution of Breast, Lung and Colorectal Cancer in Long Island, New York. http://www.terraseer.com/casestudies/longisland/default.html

52. Karr A, Lee J, Sanil P. 2002. Web-based Systems that Disseminate Information from Data but Protect Confidentiality. *Advances in Digital Government*. Elmagarmid AK, McIver WM, ed. Amsterdam: Kluwer Press

53. Krieger N, Waterman P, Chen JT, Soobader MJ, Subramanian SV, Carson R. 2002. Zip Code Caveat: Bias due to Spatiotemporal Mismatches Between Zip Codes and US Census-Defined Geographic Areas: The Public Health Disparities Geocoding Project. *AJPH*. 92(7): 1100-1102

54. Lake A. 2002. *Will GML Enable an Accessible Geo-Web? GeoWorld* 15(7): 42-44

55. Leidner A, Homeland Security Working Group, Federal Geographic Data Committee, unpublished data, January 31, 2002

56. Longley PA, Goodchild MF, Maguire DJ, Rhind DW, ed. 1999. *Geographical Information Systems* Second Edition. New York: John Wiley & Sons, Inc.

57. Lowe JW. 2002. *Homeland Homework: Reconfiguring for Wider Spatial Integration. Geospatial Solutions* 12:42-45.

58. MacDorman MF, Gay GA. 1999. State Initiatives in Geocoding Vital Statistics Data. In *J of Pub Health Manage Prac*, eds. TB Richards, CM Croner 5(2): 91-93

59. Mapping Science Committee. 1995. *A Data Foundation for the National Spatial Data Infrastructure*. Washington, DC: National Academy Press.

60. Mapping Science Committee. 1997. *The Future of Spatial Data and Society*. Washington, DC: National Academy Press

61. Meador M, Ruggles AJ. 2000. Steps involved in Randomizing the Coordinates of Address-Matched Locations. In Croner CM ed. *Public Health GIS News and Information* 35:11-12. http://www.cdc.gov/nchs/gis.htm

62. Melnick AL. 2002. *Introduction to Geographic Information Systems in Public Health*. Gaithersburg: Aspen Publishers, Inc.

63. Maryland. Department of Planning. 2002. *Hot Spot Communities & Spot Light Schools*. http://www.mdp.state.md.us/GIS/mdmaps.htm

64. Multi-Sector Crisis Management Research Center. 2002. *MSCMC*. http://archive.ncsa.uiuc.edu/mscmc/mscmrc.html

65. National Association of County and City Health Officials. 2002. *Geographic Information systems*. http://www.naccho.org/

66. National Cancer Institute. 1999. Cancer Mortality Maps & Graph Web Site. http://www3.cancer.gov/atlasplus/

67. National Cancer Institute. 2002. *Geographical Information System for Health* (GIS-H). http://www.healthgis-li.com/researchers/researchers.htm

68. National Center for Health Statistics. 1996. *Atlas of United States Mortality* Hyattsville, MD

69. National Center for Health Statistics. 2002. *Healthy People 2010*. http://www.cdc.gov/nchs/hphome.htm

70. National Institute of Justice. 2002. *Mapping and Analysis for Public Safety.* http://www.ojp.usdoj.gov/nij/maps/welcome.html

71. National Institutes of Health. 2002. *Cancer Mortality Maps & Graph Web Site*. http://www3.cancer.gov/atlasplus/index.html

72. New York City Office of Emergency Management OEM. 2002. *OEM Mission Statement*. http://www.nyc.gov/html/oem/html/about.html

73. New York State GIS Clearinghouse. 2002. *Results from the 2002 NYSGIS survey*. http://www.nysgis.state.ny.us/index.html

74. New York State. 2002. *Cancer Surveillance Improvement Initi*ative. http://www.health.state.ny.us/nysdoh/cancer/csii/nyscsii.htm

75. Niemann B, 2002. *Building Peer-to-Peer XML Content Networks of Web Services for Federal Scientific and Statistical Data and Information: FedStats.Net and Beyond*. http://www.gsa.gov/attachments/GSA_PUBLICATIONS/extpub/11-BLNiemann-EPA_7.htm

76. Niemann B. 2002. *XML Web Services: Support for the Geospatial Information One-Stop*. Unpublished data

77. Niemann B. 2002. *XML Web Services: Virtual Centralization of Distributed Content*. Unpublished data

78 Office for Civil Rights. 2002. Medical Privacy-National Standards to Protect the Privacy of Personal Health Information: Final Modifications to the Privacy Rule, Federal Register, August 14, 2002. http://www.hhs.gov/ocr/hipaa/finalreg.html

79. OGETA, Inc. 2002. *About OGETA Consortium*. http://www.ogeta.com/consortium/about.html

80. OpenGIS Consortium. 2001. *Geography Markup Language* (GML) 2.0. http://opengis.net/gml/01-029/GML2.html

81. OpenGIS Consortium. 2002. *Multi-Hazard Mapping Initiative (MMI)* http://ip.opengis.org/mmi

82. Office of Management and Budget. 2002. Circular A-16, revised draft, unpublished data

83. Peckenpaugh J. 2001. Mapping agency blocks access, postpones outsourcing pact. http://www.govexec.com/news

84. Potomac Tech Journal. 2002. *D.C. Mapping Ahead of Pack.* http://www.potomactechjournal.com/displayarticledetail.asp?art_id=57707

85. Public Health Development Fund. 2002. *Multi-agency Internet Geographic Information Service*. http://maigis.wmpho.org.uk/

86. Reichardt ME. 2002.*OGC Seeks Interested Parties for Geospatial One-Stop*. http://www.opengis.org/ogcInterop.htm

87. Reichardt ME. 2001. Press announcement. http://www.opengis.org/ogcInterop.htm

88. Reichardt ME. 2002. *XML's Role in the Geospatial Information Revolution*. http://www.gsa.gov/attachments/GSA_PUBLICATIONS/extpub/11-MReichardt-OGC_3.htm

89. Richards TB, Croner CM, Rushton G, Brown CK, Fowler L. 1999.Geographic Information Systems and Public Health: Mapping the Future. *Pub Health Rep*. 114 (4): 359-373

90. Richards TB, Henriques WD, Croner CM, Brown CK, Saccenti JC, Berry P. Toward a GIS Sampling Frame for Surveys of Local Health Departments and Local Boards of Health. In *J of Public Health Manage Prac*, eds. TB Richards, CM Croner 5(4): 65-74

91. Rushton G, Lolonis P. 1996. Exploratory spatial analysis of birth defect rates in an urban population. *Statistics in Medicine* 15:717-26

92. San Diego Association of Governments. 2000. *Regional Information System Overview*. http://www.sandag.org/uploads/publicationid/publicationid_216_559.pdf

93. Soley MS, OMG Staff Strategy Group. 2000. *Model Driven Architecture*. http://www.omg.org/mda/presentations.htm

94. Steel P, Sperling, J. 2001.The Impact of Multiple Geographies and Geographic Detail on Disclosure Risk: Interactions between Census Tract and ZIP Code Tabulation Geography", Proceedings of Survey Research Methods Section, American Statistical Association

95. Theseira M. 2002. Using Internet GIS Technology for Sharing Health and Health Related Data for the West Midlands Region. Health & Place. 8 (1): 37-46. http://maigis.wmpho.org.uk

96. Thrall GI. 1999. The Future of GIS in Public Health Management and Practice. In *J of Pub Health Manage Prac*, eds. TB Richards, CM Croner 5(4): 75-82

97. Thrall SE. 1999. Geographic Information System (GIS) Hardware and Software. In *J of Pub Health Manage Prac*, eds. TB Richards, CM Croner 5(2): 82-90

98. University of Illinois at Chicago. 2002. *National Center for Data Mining*. http://www.ncdm.uic.edu

99. US Environmental Protection Agency. 2002. *Enviromapper*. http://epa.gov/superfund/sites/locate/index.htm

100. US Geological Survey. 2001. *Cancer Mortality in the United States: 1970-1994: Map Layer Description File*. http://www.nationalatlas.gov/cancerm.html

101. US Geological Survey. 2002. *National Atlas of the United States*. http://www.nationalatlas.gov; http://www.nationalatlas.gov/virusmap.html; http://www.nationalatlas.gov/virusprint.html

102. US Geological Survey. 2002. *NWISWeb Data for the Nation*. http://waterdata.usgs.gov/nwis/

103. Wallace M, Sperling J. 2000. Integrated Statistical Solutions, URISA Jn. 12(4,) http://www.urisa.org/Journal/protect/Vol12%20No4/sperling/Pages%20from%20coverfinal.pdf

104. Wartell J, McEwen JT. 2000. *Sharing Crime Maps and Spatial Data: Meeting the Challenges*. National Institute of Justice

105. Washington Geographic Information Consortium. 2002. Washington GIS Consortium. http://www.wgis.org

106. Washington State Department of Health. 2002. EpiQMS. http://www5.doh.wa.gov/EpiQMS

107. Xue Y, Cracknell AP, Guo HD. 2002. Telegeoprocessing: the Integration of Remote Sensing, Geographic Information System (GIS), Global Positioning System (GPS) and Telecommunication. Internl Jn of Remote Sens. 23(9): 1851-1893

108. Yasnoff WA, Sondik EJ. 1999. Geographic Information Systems (GIS) in Public Health Practice in the New Millennium. In *J of Public Health Manage Prac*, eds. TB Richards, CM Croner 5(4): ix-xii

109. Zayatz LY, Steel P, Rowland S. 2000. Disclosure Limitation for Census 2000. *Proceedings of the Section on Government Statistics*, American Statistical Association, Indianapolis

Glossary of Terms

Cadastral- data that describe the geographic extent of past, current, and future right, title, and interest in real property, and the framework to support the description of that geographic extent

Conflation- A process by which two or more digital maps of the same area, may be matched and merged into one; a computational process of converting an image or map from one co-ordinate system to another

Digital Orthoimagery- a dataset that contains georeferenced images of the Earth's surface, collected by a sensor in which image object displacement has been removed for sensor distortions and orientation, and terrain relief

Empirical Bayes (EB) mapping- a parametric statistical procedure to stabilize spatial statistical analysis, prior to mapping, by Bayesian modeling which “shrinks” the statistics from areas with a small population toward an overall mean

E911 program- Wireless E911 is a vital step toward improving public safety. Automatic Location Identification (ALI) capability permits rapid response in situations where callers are disabled or do not know their location by allowing for the immediate dispatch of emergency assistance to the location of the emergency. Phase II E911 (by December 31, 2005) requires wireless carriers to provide far more precise location information, within 50 to 100 meters in most cases

Executive Order 12906- "Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure," signed on April 11, 1994, by President William Clinton; requires Federal agencies to ensure that data comply with FGDC-endorsed standards prior to obligating funds

Extensible Markup Language (XML)- a specification developed by the World Wide Web Consortium (W3C) designed especially for Web documents. It allows designers to create their own customized tags (how formatted), enabling the definition, transmission, validation, and interpretation of data between applications and between organizations

FGDC (Federal Geographic Data Committee)- an Office of Management and Budget interagency committee, including DHHS, with responsibility for facilitating Circular A-l6 related activities and implementation of the NSDI

Geocoding- digital procedure for finding map coordinates that correspond to data attributes of features; the TIGER system provides a national computer-readable map database for geocoding operations

Geodetic Control- a network of surveyed and monumented points on the earth’s surface that provides a common reference system for establishing coordinates for all geographic data. A state plane coordinate system comprises a system of x,y coordinates for each state and is commonly used in GIS. All NSDI framework data require geodetic control to accurately register spatial data

Geographic Information System (GIS)- computer system for the input, editing, storage, retrieval, analysis, synthesis, and output of location-based information. GIS may refer to hardware and software, or include data

Geography Markup Language (GML)- an XML encoding for the transport and storage of geographic information, including both the spatial and non-spatial properties of geographic features

Georeference- to reference spatial data to a geodetic reference system or the process of tying spatial data to a geodetic reference system; spatial data may be derived from remote sensing, mapping, charting, surveying technologies, GPS, or statistical data, and other sources

Geospatial data or information- identifies the geographic location and characteristics of natural or constructed features and boundaries on the earth; Coordinate and attribute data for location-based features, usually in the categories of points, lines, polygons, and cells; geographic information that has a spatial component or the spatial component of geographic information

Global Positioning System (GPS)- a satellite-based system deployed to determine locations on the Earth's surface

Internet (and related terms)- a worldwide system for linking together, through a particular set of communication standards known as TCP/IP, smaller computer networks; World Wide Web is a hypertext system where everything presented to the user is a hypertext object in HTML format

Interoperability: the standards based interoperable framework of Web services that allows the services to be accessed anywhere at any time, using any computer platform

Metadata- "data about data" that describe the content, source, vintage, spatial scale, accuracy, quality, condition, and other characteristics of data; metadata are critical to document, preserve and protect the spatial data assets of agencies

National Spatial Data Clearinghouse- a network of geospatial data producers, managers, and users linked electronically that provides access to documented spatial data and metadata from distributed data sources

National Spatial Data Infrastructure (NSDI)- the technology, policies, standards, and human resources necessary to acquire, process, store, distribute, and improve utilization of geospatial data

NSDI Framework Data Themes- includes geodetic control, orthoimagery, elevation and bathymetry, transportation, hydrography, cadastral, and governmental units; Data associated with these themes include an encoding of the geographic extent of the features and a minimal number of attributes needed to identify and describe the features

NSDI Geographic Information Standards- common and repeated rules, conditions, guidelines or characteristics for data, and related processes, technology and organization; NSDI standards are developed and promulgated by the FGDC in accordance with OMB Circular A-16

OMB Circular No. A-16- establishes the FGDC as the interagency coordinating body for NSDI-related activities chaired by the Secretary of the Interior with the Deputy Director for Management, Office of Management and Budget (OMB) as Vice-Chair, revised August 2002

Open GIS Consortium, Inc. (OGC)- an international industry consortium of companies, government agencies and universities participating in a consensus process to develop publicly available geoprocessing specifications in support of interoperable solutions that "geo-enable" the Web

OpenGIS- defined as transparent access to heterogeneous geodata and geoprocessing resources in a networked environment

Public Health Spatial Data Theme- NSDI data theme that relates to the protection, improvement and promotion of the health and safety of all people; public health databases include spatial data on mortality and natality events, infectious and notifiable diseases, incident cancer cases, behavioral risk factor and tuberculosis surveillance, hazardous substance releases and health effects, hospital statistics and other similar data; DHHS has lead responsibility for public health data theme

Scale- the ratio of the distance measured on a map to that measured on the ground between the same two points; the larger the ratio, the smaller the map scale. Thus, a map of the world would have a very small scale compared to a map of a town, with a large scale

Spatial Data- any information about the location and shape of, and relationships among, geographic features; this includes remotely sensed data as well as map data

TIGER Database- Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) computer database containing the position of geographic information including roads, rivers, railroads, feature attributes such as name, address range and code, boundaries and other map features. TIGER represents all U.S. space as a single seamless data inventory

ZIP (Zone Improvement Plan) Code- a three, five, seven, or nine–digit code assigned by the U.S. Postal Service (USPS) to a section of a street, a collection of streets, an establishment, structure, or group of post office boxes, for the express purpose of mail delivery. ZIP Code "areas" are not polygons but actually entities of lines

ZIP Code Tabulation Area (ZCTA)- new statistical area, based on grouping whole census blocks, created by Census Bureau for Census 2000 to address the problems associated with tabulating data by ZIP Codes; ZCTAs are generalized representations of USPS ZIP Code areas and thus based on USPS delivery criteria. ZIP Codes for the majority of addresses in an area provide the basis for ZCTA codes

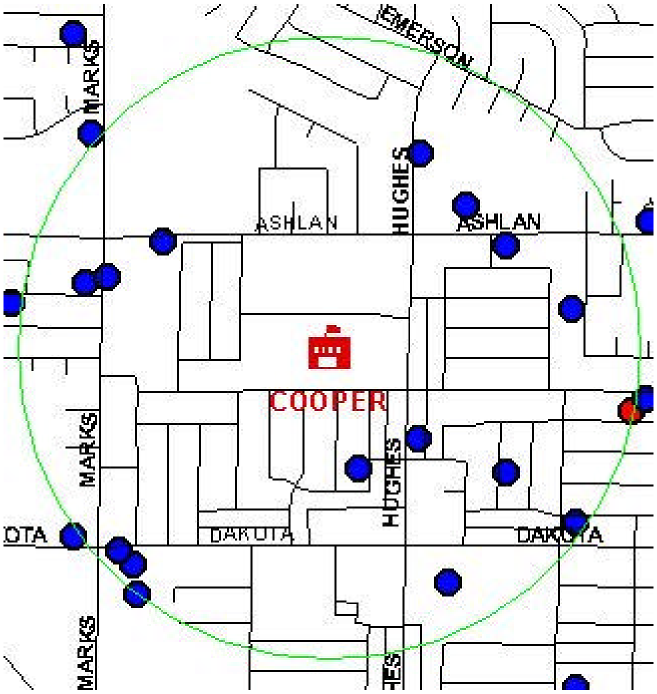

Figure 1. Green circle is ½ mile radius around Cooper Junior H.S. in Fresno, CA. It contains residential locations of 14 “serious” (blue dot locations) and one “high risk” (red dot) sex offenders. “Megan's Law" permits the release of this geocoded information to the Internet. (Fresno Police Department, Megan’s Law web site http://www.ci.fresno.ca.us/fpd/meganlaw/index.html)

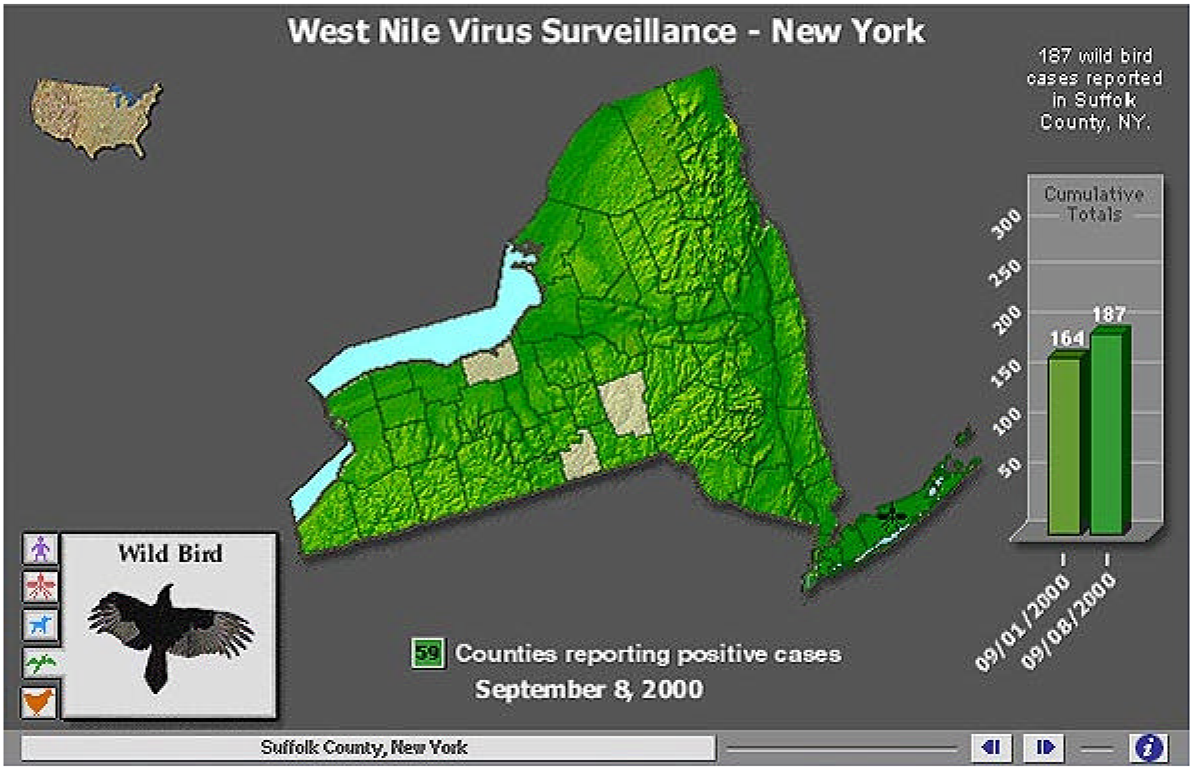

Figure 2. USGS National Atlas West Nile Virus surveillance data offers users interactive tools for space-temporal analysis. Infected wild bird counts are shown for first week September, 2000, in Suffolk County, NY. Users may change surveillance categories, counties, and weekly occurrence (USGS National Atlas of the U.S. website http://www.nationalatlas.gov/virusmap.html)

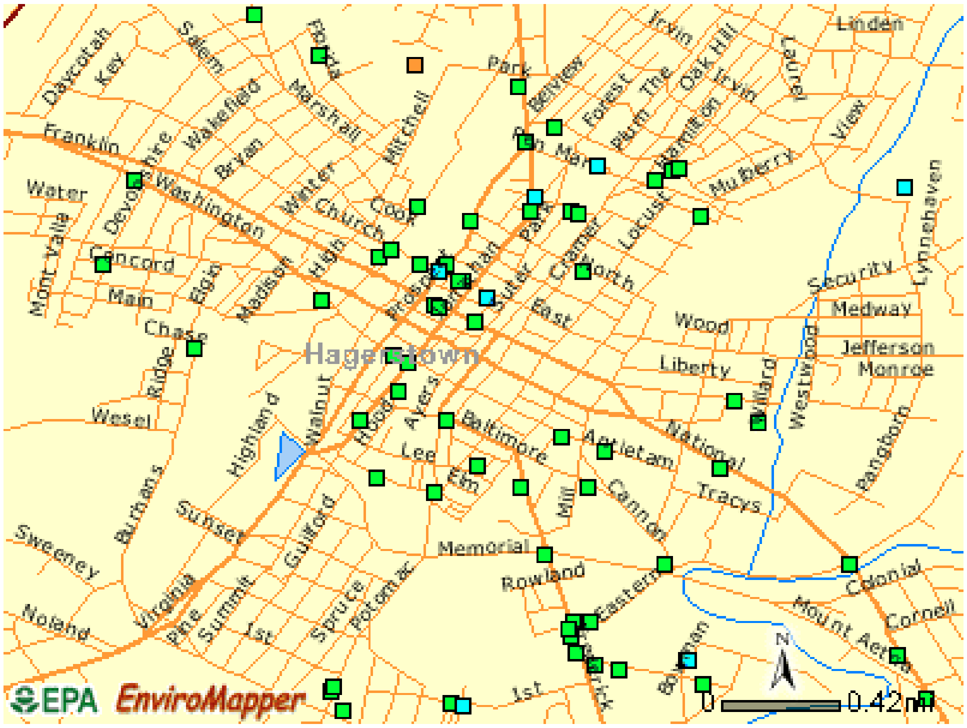

Figure 3. EPA's "Window to My Environment" allows user to identify sites in many locations. In Hagerstown, MD, a Superfund Site (Central Chemical) is shown by orange icon to north. Company onetime was involved in the manufacture of pesticides and some material was placed in a dump adjacent to the plant. Toxic release facilities are shown in blue and hazardous waste in green. Many cities can be queried for these and other environmental conditions. (EPA WME website http://www.epa.gov/enviro/wme)