Ryerson University
Digital Commons @ Ryerson

Geography Publications and Research | Geography

1-1-2011

Exploring the Boundaries of Web Map Services: The Example of the Online Injury Atlas for Ontario

Claus Rinner
Ryerson University, crinner@ryerson.ca

others

Follow this and additional works at: http://digitalcommons.ryerson.ca/geography

Recommended Citation
C. Rinner, B. Moldofsky, M.D. Cusimano, S.P. Marshall, T. Hernandez (2011) Exploring the Boundaries of Web Map Services: The Example of the Online Injury Atlas for Ontario. Transactions in GIS 15(2): 129-145

This Article is brought to you for free and open access by the Geography at Digital Commons @ Ryerson. It has been accepted for inclusion in Geography Publications and Research by an authorized administrator of Digital Commons @ Ryerson. For more information, please contact bcameron@ryerson.ca.

Exploring the Boundaries of Web Map Services: The Example of the Online Injury Atlas for Ontario

Claus Rinner[1], Byron Moldofsky[2], Michael D. Cusimano[3], Sean Marshall[3], Tony Hernandez[1]

[1]Department of Geography, Ryerson University

[2]GIS and Cartography Office, Department of Geography and Program in Planning, University of Toronto

[3]Injury Prevention Research Office, Keenan Research Centre, Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto

Abstract

Injury from causes such as falls, traffic accidents, or violence is a major public health issue globally. Injury prevention research aims to identify vulnerable populations and places by analyzing the spatial patterns of demographic and socio-economic risk factors associated with elevated injury rates. The stakeholders in injury prevention and control are often distributed across government and public health institutions, non-profits, and even the private sector (e.g. insurance firms). While this situation calls for distributed, online research tools, their implementation may conflict with health data confidentiality and license limitations for socio-economic data. In this paper, we present the Online Injury Atlas for Ontario, which was designed with the explicit goal to make use of, and contribute to, the Canadian Geospatial Data Infrastructure. We propose a service-based architecture that integrates publicly accessible map services with protected data layers. Thereby, we demonstrate the benefits of using spatial data infrastructures alongside private data at different levels of protection. In addition, we discuss the extensive data processing needs and specific cartographic design requirements of a Web atlas in the health and social sciences domain.

1. Introduction

Injuries are the greatest single contributor to potential years of life lost for Canadians before 65 years of age. Annually, approximately 200,000 people are hospitalized for injuries in Canada (CIHI 2008), more than 13,000 lives are lost, and the economic burden exceeds $19.8 billion (Smartrisk 2009). In most industrialized countries, injuries account for more deaths among children and adolescents than all other causes combined (Peden et al. 2008, UNICEF 2001). Injury rates vary widely between different countries and world regions. Within and across countries, the rates of injuries caused by different mechanisms (e.g. violence, self-inflicted, traffic-related, or falls) also vary greatly. The World Health Organization has projected that injury will be the third leading cause of death worldwide by the year 2020 (Peden et al. 2008).

Injuries are often a result of a complex set of factors (e.g. Kim et al. 2007). Therefore, interdisciplinary, multisectoral, cross-jurisdictional, and multifaceted research and prevention approaches are required (Cusimano et al. 2007). Ultimately, the successful reduction of the burden of injury will come about through understanding the complex interrelationships that influence the phenomenon at a range of scales from the individual to the societal level. Through their capability of data integration, location-specific analysis, and cartographic visualization, geographic information systems (GIS) are increasingly used in injury prevention research (Cusimano et al. 2007, Edelman 2007, Bell and Schuurman 2010).

Comprehensive datasets on injury and its risk factors are available in Ontario, and the Province's public health units are mandated to devise injury prevention programs. However, there are no custom tools available that enable the use of the existing datasets for location-specific injury prevention programming. Therefore, we aimed to develop a map-centred decision support tool for health analysts, lobbyists, and policy-makers in government and non-government organizations that deal with injury prevention. We also wanted to design a tool that would allow for remote access by geographically distributed stakeholders. Therefore, our research objective was to assess the use of Web map services for injury prevention. The objective was addressed by developing an online atlas architecture on the basis of GIS interoperability standards and cartographic principles, and implementing a prototype atlas with stakeholder input.

The structure of this paper is as follows. In Section 2, we describe the project background and technological context of GIS in injury prevention research, including Web mapping of health data and the principles of spatial data infrastructures. In Section 3, we describe the data sources, characteristics, and preprocessing; the architecture of the online atlas; an end-user consultation process; and the final cartographic and functional design decisions. Section 4 summarizes and discusses the results of this project. Lastly, Section 5 concludes the paper with an outlook on future research.

2. Background and Technological Context

2.1. Geographic Information Systems in Injury Prevention Research

During the 1980s and 1990s, geographic information systems (GIS) matured as essential decision support tools in government and industry. Today, GIS are used in public health and epidemiology to analyze spatial patterns of population health, social determinants of health, and the spread of disease. In a special issue of the Annual Review of Public Health, Cromley (2003) reviews GIS applications to disease surveillance, while McLafferty (2003) assesses GIS use for health services planning. In the same issue, Rushton (2003) explores processes for geocoding and spatial analysis of health data. An important issue in health applications of GIS is the protection of individual privacy. Starting with the seminal paper by Armstrong et al. (1999), researchers have devised a number of algorithms for geographically masking (i.e. modifying) patient locations from health records. Acknowledging this possibility, Wartenberg and Thompson (2010) are concerned that the additional processing and the lack of an agreed-upon masking approach may inhibit public health research. In the work presented here, we use the traditional approach of aggregating individual records to geographic units of analysis, and not reporting outcomes below a threshold number, e.g. five cases.

Specific to injury, Bell and Schuurman (2010) review a number of studies where GIS have been used to examine spatial patterns of injury and its associated socio-demographic and environmental risk factors. For example, Haynes et al. (2003) found neighbourhood effects on school-aged child injuries (along with individual and family characteristics); Bell et al. (2008) observed a "social gradient" in some injury types but not others; and Cusimano et al. (2010) found clusters of violent injury in deprived neighbourhoods. Population health determinants broadly associated with the increased risk of injury include income and social status, education, employment conditions, social and physical environments, personal health practices, healthy child development, biologic and genetic endowment, health services, sex, and culture (D'Cunha

2002). Although each of these factors alone plays an important role in the incidence of injury, the complex interrelation of multiple factors may play out differently in specific geographic locations and community contexts. By acquiring, integrating, managing, analyzing, visualizing, and disseminating geospatially referenced information, GIS can support decision-making across a wide variety of applications relevant to reduce the societal burden from injuries (Cusimano et al. 2007).

The present study contributes to enhancing the analysis, utilization, and evaluation of the extensive datasets currently collected on injury in Canada. Key injury-related datasets are currently collected in isolation by a range of stakeholders, and therefore are of limited use for the development of evidence-based decision-making to prevent injuries. A mapping and decision support system that helps with understanding the risk factors that underlie local injury events therefore may provide an effective tool for the development and evaluation of targeted prevention interventions (Eyles 1999).

2.2. Web Mapping and Spatial Data Infrastructures for Public Health

Croner (2003) speculates about Web-based resources for public health decision-making and the integration of health information into spatial data infrastructures. Subsequently, a number of technically isolated Web mapping applications with health-related data contents have been created, including Kamadjeu and Tolentino (2006), Maclachlan et al. (2007), and Cinnamon et al. (2009a). However, the selection of specific Internet map server software often determines the data handling and analytical capabilities available to decision-makers. In an attempt to make geographic data processing interoperable, the Open Geospatial Consortium (formerly: OpenGIS Consortium) was founded in 1994 by the US government and industry partners, and its membership subsequently expanded internationally (OGC 2006). The Consortium developed a number of technical interoperability specifications, including the OpenGIS Web Map Service (WMS) and Web Feature Service (WFS) specifications (published at www.opengeospatial.org).

Among the first published articles on using OpenGIS-conformant technology in public health is a tutorial by Boulos and Honda (2006). The authors highlight the beginning maturity of GIS interoperability standards and open-source GIS software, and provide instructions on how to publish maps of health data on the Web using a WMS. The potential of open specifications such as the Open Geospatial Consortiums interoperability standards for cross-border health issues is emphasized by Moreno-Sanchez et al. (2007). Their Web-based multimedia GIS system for dengue fever monitoring makes use of distributed raster images provided as map layers through a WMS. Arguably the most advanced use of Web mapping standards in public health is presented by MacEachren et al. (2008). In their Web-based cancer atlas for Pennsylvania, the authors assess the potential of the WFS to support a highly interactive online geovisualization tool for public health. For example, the cancer atlas allows users to simultaneously highlight data for a selected county on a map, a cumulative frequency plot, and a data table. Additional functionality includes map comparison and map animation (MacEachren et al. 2008). The research reported in the present paper verifies findings from the cited studies in the context of injury prevention and control. We extend prior work by providing map layers related to the social determinants of injury, which are important to support end-users in public health in understanding injury risk and developing strategies for injury control.

A key idea of GIS interoperability is to keep data holdings with the authoritative source (for example Statistics Canada for Canadian Census data), and provide access to those data through computer networks and standardized software interfaces. The interface specifications make data access independent of the underlying data format, allowing tools from different GIS and web mapping vendors to access the same data. Further, the data provider is enabled to control the scope of access, e.g. whether access to the underlying data values is granted (e.g. WFS interface), or limited solely to graphical rendering of map data (e.g. WMS interface). As an extension of these principles, spatial data infrastructures (or spatial data infrastructures) are organized collections of servers, services, and policies that provide access to a coherent set of geographic data. One such framework is the Canadian Geospatial Data Infrastructure (CGDI), which provides access to numerous government data sets through its discovery portal (geodiscover.cgdi.ca). The CGDI suggests a set of OpenGIS specifications for data and technology to be used by organizations participating in the initiative. Further development of the CGDI aims to expand the list of data providers, and showcase the utilization of available data, technology, and interoperability specifications. The initiative specifically targets four priority areas: public safety, public health, the Aboriginal community, and environment and sustainable development. The research reported here contributes to the CGDI's public health priority area.

3. Methodology

3.1. Data Sources, Characteristics, and Preprocessing

The data used in the Online Injury Atlas for Ontario represent three broad themes: injury rates, demographic and socio-economic risk factors, and orientation and background data. Injury rates were calculated from the Discharge Abstract Database. This database provides a complete account of all hospitalizations in the province of Ontario and is collected by the Ontario Ministry of Health and Long-Term Care. Data were available for the three fiscal years 2000/01, 2001/02, and 2002/03. Each year contained approximately 170,000 records that had been geocoded to the patient's residence by census tracts as the most detailed level of geographic resolution.

The atlas includes maps for the entire Province of Ontario at the level of the local health integration network (LHIN) and municipality (Census subdivision). Further, it includes maps of four urban areas at the census tract level: the Golden Horseshoe (Greater Toronto Area, Hamilton, and Niagara), Ottawa, Windsor, and Sudbury, as well as a map combining these four areas.

The following "event codes" from the International Classification of Disease, Ninth Edition (ICD-9), were used to distinguish broad categories of injury by their causes:

- Motor vehicle and traffic injuries: ICD-9 codes E810-E825
- Fall injuries: ICD-9 codes E880-E889
- Assault injuries: ICD-9 codes E960-E969
- Self-inflicted injuries: ICD-9 codes E950-E959
- Other injuries: all remaining, injury-related ICD-9 codes

Injury rates for these five categories were calculated as an annualized, crude rate per 100,000 persons using population figures from the 2001 Census of Canada. Another set of injury rates were calculated for the following three age groups as an annualized, crude rate per 100,000 persons of the same age group:

- Age 0-14
- Age 15-64
- Age > 64

These age thresholds were chosen to be compatible with Statistics Canada's definitions of working-age and senior population groups, which the intended end-users of the atlas would be most familiar with. The rates were calculated for census tracts and also aggregated to municipal boundaries (census subdivisions). For the larger LHIN boundaries, injury rates by type and age group were calculated due to larger counts that were available.

The injury data from the Discharge Abstract Database are subject to strict confidentiality constraints. Anonymized records were received from the Ministry. Those records were aggregated by Census geographic units and transformed to injury rates. The following sections on cartographic design and the system architecture both discuss additional measures to preserve data confidentiality.

Demographic and socio-economic risk factors were represented by selected variables from the 2001 Census of Canada (www.statcan.ca/english/census01):

- Population density
- % of population who are less than 15 years of age
- % of population who are more than 65 years of age
- % single-parent families
- % recent immigrants (within last 10 years)
- Average household income
- % low income families
- % in labour force who are unemployed
- % over 20 years of age who did not graduate from high school
- % over 20 years of age with a college or university degree

The Census data are subject to the Data Liberation Initiative data use agreement between Canadian universities and Statistics Canada, which gives researchers and students access to population statistics for academic purposes. The Data Liberation Initiative permits the display of Census tract-level static maps over the Internet within the scope of a research project. However, the agreement prohibits the distribution, including the donation, of Census data to users who do not participate in the agreement, when this occurs outside of a research project. For example, online publication of a Census atlas outside of a research project would require special permission from Statistics Canada.

Additional data layers were used to provide geographic orientation and visual reference for the thematic data. These background data comprised the following administrative boundaries, points of interest, transportation networks, and water bodies. The boundary layers included Toronto neighbourhoods sourced from the City of Toronto, and public health units, LHINs, and Census subdivisions (municipalities) from Statistics Canada. The locations of acute-care hospitals were extracted from data provided by the Ontario Ministry of Health and Long-Term Care, and the locations of police stations were acquired from the Internet. Finally, major roads and minor roads from the Atlas of Canada, Natural Resources Canada, were used.

Individual metadata pages were created for each layer in every map in the atlas. These pages specify the data type, source, availability at different geographic levels, the areas covered, variable definitions, reference documents, and information about cartographic representation.

3.2. Software and Development Approach

A survey of interactive Web map sites for health-related data by Cinnamon et al. (2009a) found MapServer (Kropla 2005, Vatsavai et al. 2006, www.mapserver.org) to be a popular software platform for online health mapping. Its use in public health applications was also recommended by Boulos and Honda (2006) and Moreno-Sanchez et al. (2007). MapServer was originally developed at the University of Minnesota and released as open-source software under the GNU general public license. At the time of our development, alternatives such as GeoServer and OpenLayers were not nearly as widely used as MapServer. MapServer is a common gateway interface (CGI) program that uses parameters sent in a request from a client Web page to return a map image. MapServer handles a number of vector and raster input data formats, allows for map configuration, and returns its output as an image or graphics file or as an OpenGIS-compliant Web service.

We found that in addition to MapServer, a number of Web sites used open source GIS and map customization packages built on top of basic MapServer functionality. For the atlas we selected the CartoWeb framework (Kiener and Philipona 2007; www.cartoweb.org), which was also released under the GNU general public license. CartoWeb provides developers with customizable mapping components built on top of MapServer. CartoWeb uses hypertext preprocessor (PHP), Javascript, and Cascading Style Sheets (CSS) techniques along with a library of Web page templates that eases the process of Web page design. For example, it includes a template for map display and navigation that supports the addition of WMS layers and user-specified bookmarks for map views.

The implementation of the Online Injury Atlas that is described in the following section is the result of an iterative process within a longer-term collaborative research program. An earlier description of the extensive injury data collection effort can be found in Cusimano et al. (2007). In Cinnamon et al. (2009a, b) we discuss the implementation and testing of a preliminary online map series for injury rates and socio-economic risk factors that was not based on GIS interoperability standards but contributed to developing the present user interface and cartographic design. The development of the atlas, which is presented in this paper, included the decision on a general system architecture, a prototype implementation, end-user consultations, and the development of a final version with refined cartographic and functional design.

4. Implementation of the Online Injury Atlas

4.1. System Architecture and Prototype Implementation

The atlas attempts to integrate publicly accessible map layers from the Canadian Geospatial Data Infrastructure (CGDI) and sensitive data from the health and social science domain. A service-based architecture was chosen to enable distributed data access at different partner institutions and to utilize data, interoperability specifications, and technology endorsed by the CGDI.

Figure 1 illustrates the atlas architecture. Framework data from the CGDI (left column) are integrated with Census data from a University of Toronto server and injury data from a Ryerson University server. Data integration is achieved through overlay of map layers according to the CGDI-endorsed OpenGIS Web Map Service (WMS) specification (de La Beaujardiere 2006). The Web client interface displays all map layers in the same way so that the end-user does not need to worry about their origins.

Conceptually, there are two levels of access protection. The Census data are served within the atlas project as static map layers from a WMS. The injury data are also served as a WMS, but in addition, they are password-protected and maintained on a highly secured server.

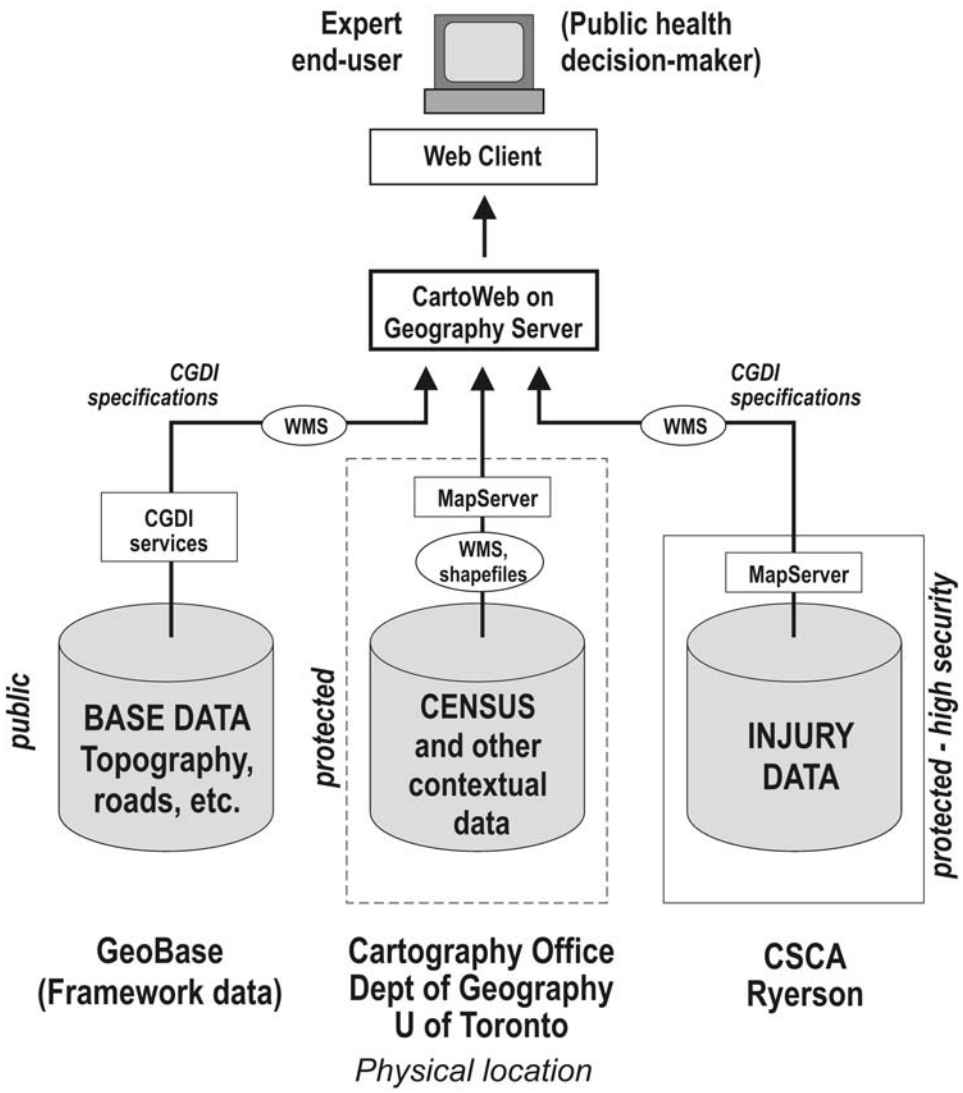

Figure 1: Architecture schema of the Online Injury Atlas for Ontario

A prototype of the Online Injury Atlas was implemented in order to illustrate design and functional options to potential end-users. Customizations that were necessary to the default MapServer and CartoWeb configuration, included:

- side-by-side map window design
- toolbars for mapping purposes
- printing interface and page design
- map extents for zoom functions
- matched map extents function
- help pages

The main limitation of the prototype was that injury data were only available for the Golden Horseshoe Area (Toronto-Hamilton-Niagara, along Lake Ontario shore). In comparison with the final atlas described below, the prototype was also lacking in consistent use of a single colour scheme for thematic mapping and its layer management was not ideally organized.

4.2. End-User Consultation

Following the development of the prototype, a series of end-user tests and evaluations were undertaken to assess the appearance and functionality of the system. Over a four-week period, formal tests were conducted with 14 individual users representing seven different government, non-government, and academic institutions from the health sector. Interviews were scheduled for one or two researchers to visit each end-user at their workplace. Each interview started with a five-minute presentation introducing the appearance and functionality of the atlas. The end-user was then asked to complete five or six tasks using a trial version of the atlas. An additional test phase with five interviews was conducted using an improved prototype. This test included three tasks.

User interaction with the atlas interface, questions asked, time required to complete each task, and major problems encountered were recorded. Task results were compared to reference maps that had been created in a pre-test by members of the research team. A semi-structured questionnaire with closed and open-ended questions was used to assess the strengths and weaknesses of the atlas prototype and to gather suggestions for improvements. The questionnaire contained six sections with four to eleven detailed rating questions on general user aptitude, start-up of the system, tools and controls, map visualization, function groups (tabs), and contents of the atlas. Each of these sections also contained one to three open-ended questions, e.g. “Are there tools that you feel are missing?”. The ratings on the closed questions were collected on a five-point scale from “strongly agree” to “strongly disagree”, including an option called “not applicable / did not try”. In a final section, an overall evaluation of the atlas was requested using eight rating questions and six open-ended questions.

The end-user tests confirmed a number of basic design decisions by the project team, and helped to refine a number of functional objectives for the development of the Online Injury Atlas:

- interactive map format
- split-screen design to compare injury rates and socio-demographic risk factors
- interactive data selection to choose data layers for comparison
- single-variable displays on each map to simplify thematic data representation
- optional background data layers for orientation
- broad selection of injury and socio-demographic data
- most detailed level of geography possible

4.3. Cartographic and Functional Design for the Online Injury Atlas

The cartographic design of the atlas, and decisions on its functionality, depended on two interrelated perspectives. On the one hand, design and functionality should adhere to basic mapping principles. On the other hand, the atlas required consistency between many different maps and their associated functions. Decisions regarding atlas navigation, map window layout, thematic map symbols (classification, colour scheme), and mapping tools are described in the following paragraphs.

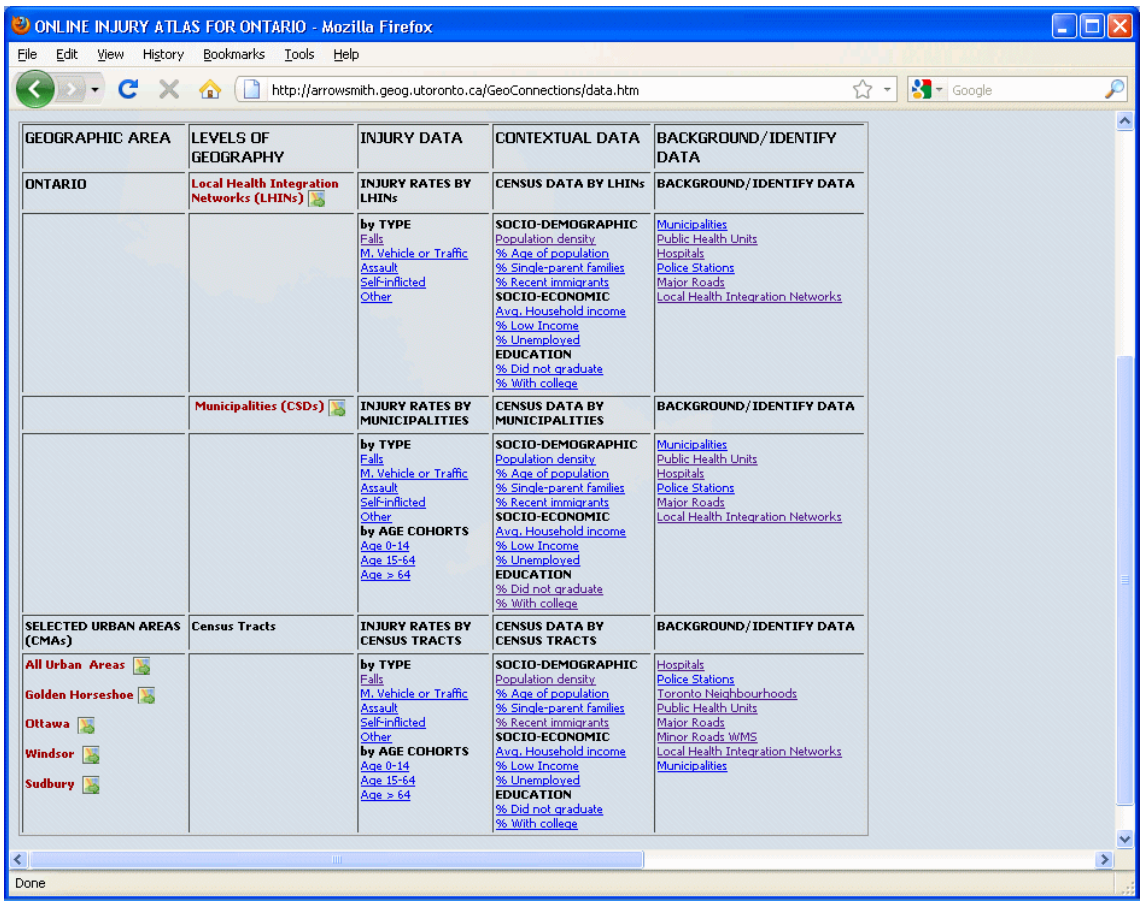

Figure 2: Table of contents with links to interactive maps and metadata pages

In contrast to a single Web map, an online atlas needs a navigational concept and tools to guide viewers to the information they are seeking. There is a continuum of approaches to provide initial access to atlas maps, ranging from a fixed table of contents with links to pre-determined static maps, to a single map with menu-driven selections of geographic area, data layers, and symbolization. Based on its data sets and target audience, the Online Injury Atlas uses an intermediate solution with a table of contents, which lists geographic areas, levels of geographic detail, and groups of data sets (Figure 2). Both, the map window and metadata pages can be accessed from this table. The map window then provides menu-, form-, and tool-driven selections of individual variables and other mapping parameters.

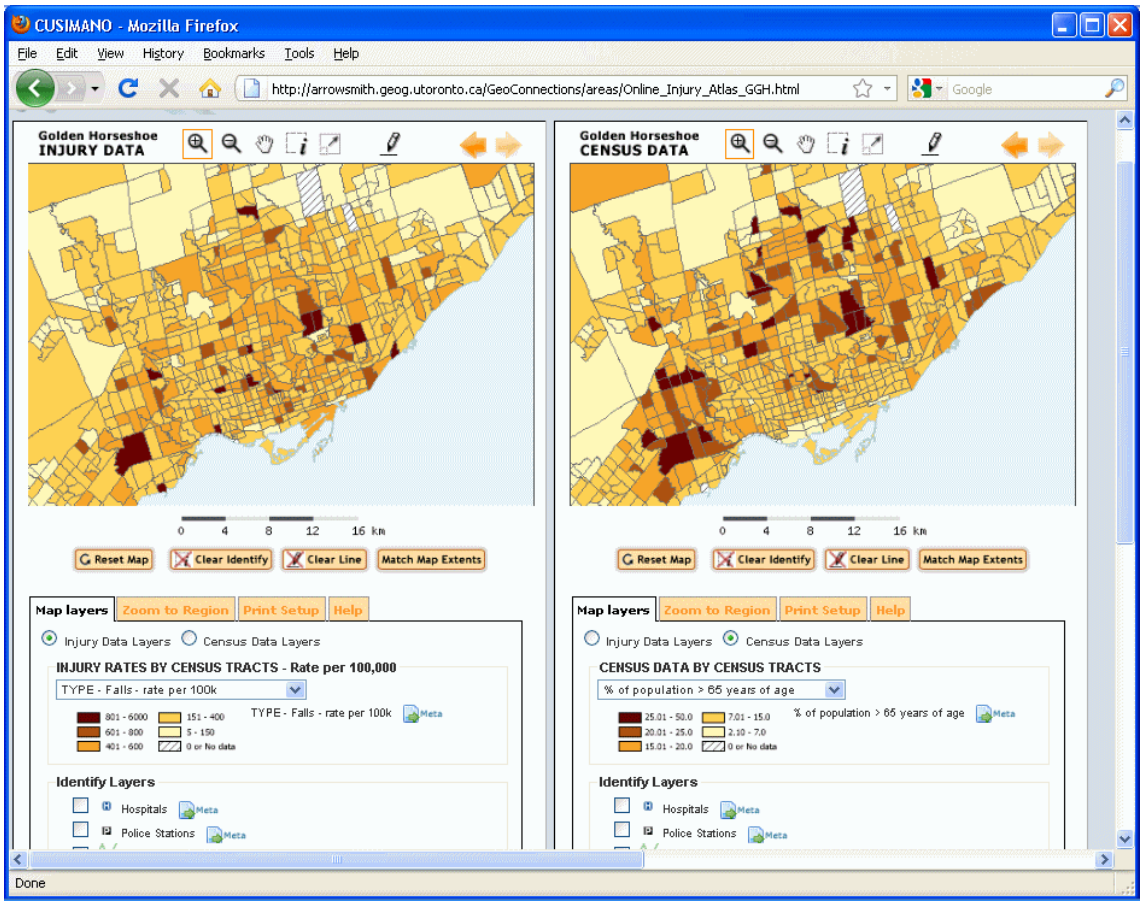

Figure 3: Main map window for side-by-side comparison
(Data source: Ontario Ministry of Health and Long-Term Care; Statistics Canada)

The map window was designed with a side-by-side presentation of two maps (Figure 3) in order to support the visual comparison of spatial patterns in injury rates and socio-demographic variables. This was deemed to be more user-friendly than a visually complex overlay (multi-variate) map. Based on feedback from end-users, the original vertical arrangement of the two maps was changed to horizontal arrangement, and Census data can now be displayed on either of the two maps instead of just one.

Mapping tools were limited to zoom in, zoom out, pan, zoom to full extent, navigate back and forth between previous map extents, and query (identify) functions. Specific functionality in support of the visual comparison includes a tool to match the extent of one map to that of the other map. A basic print layout tab and a help tab can be accessed alternatively to the list of map layers. The landscape orientation in the print layout is another example of functionality added on the basis of end-user feedback.

The selection of the primary thematic variable on each map is facilitated by a drop-down menu with a fixed choice among the injury-related and Census variables described earlier. These variables were mapped using choropleth symbolization. Additional checkboxes are provided for supplementary data layers. These are functionally grouped into (1) “identify” layers, which can

be queried for a particular attribute (e.g. name of hospital); and (2) “background” layers, which are provided in a non-interactive mode for visual reference only (e.g. major road networks).

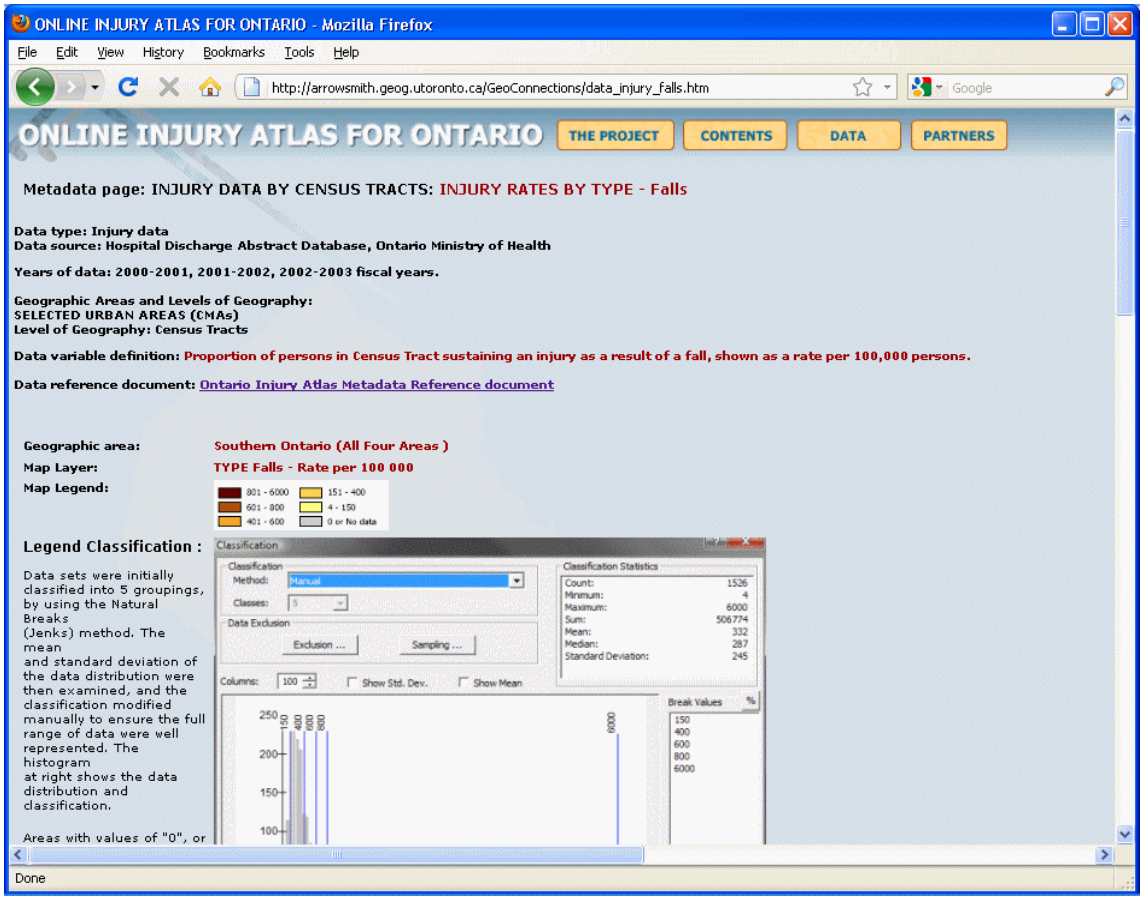

Figure 4: Metadata page documenting data source, thematic symbols, and histogram with class breaks

The method of classification of the thematic data (injury rates, socio-demographic variables) is a critical aspect of the design of choropleth maps, especially when the intention is side-by-side comparison. The issue is that classification imposes a visual structure on the data which may not express the actual spatial distribution; this becomes exacerbated when comparing diverse data sets. This question has been explored in detail by Brewer and Pickle (2002), who reviewed the literature and experimentally tested the previously reported observations and the conventional wisdom on the subject. They concluded that quantile methods perform better than any other in performing the tasks of map analysis and map comparison that they tested. This has become the “default” solution for the design of choropleth maps in series, although many writers agree that for data exploration purposes giving users control over classification is the best solution.

For this atlas, the decision was made to initially classify the data into five classes using “Jenks’ Optimal” classification method (implemented as “Natural Breaks (Jenks)” in ArcGIS 9.2). Upon inspection of the data distribution, the classification was then manually adjusted to represent the full range of data for all areas of the province at all levels of geography, to ensure that the tails of

the distributions (i.e. high and low values including significant outliers) were well captured in the classification, and to achieve easy-to-understand class breaks. Histograms with class breaks are provided on the metadata pages for each variable and geographic area (Figure 4). Areas with zero values or "no data" values due to suppression of data for confidentiality reasons were excluded from the classification process.

There were a number of reasons why this method of classification was chosen for this project. The target audience is public health decision-makers – people inexpert in mapping, but knowledgeable about the data. Brewer and Pickle (2002, p. 672) "sought to evaluate classifications for maps in series in atlases suited to lay interests, so we tested maps with a wide range of questions and suitably generalized response options." In contrast, our audience is interested in the details of the data, especially the high and low ends of the distributions. Quantile classification distorts the representation of these details. This was also the reason for including histograms representing the data distribution. The type of data being represented also played a role. Brewer and Pickle (2002, p.678) state that "Quantiles can also be thought of as converting the mapped data to ordinal rankings", which is most appropriate when applied to standardized data, such as age-adjusted rates. Although injury rates were the key data illustrated, user testing and feedback determined that crude rates rather than age-standardized rates should be shown. In this case, more distribution-specific methods, such as Jenks, are appropriate.

A sequential hue-value colour scheme (yellow-orange-brown) was used on the thematic map layers to support the visual distinction between the five classes. The dark end of the colour scheme was always associated with the "negative" end of the data range, i.e. high injury rates and high values of associated risk factors. In the case of variables that were expected to be inversely related to injury risk, i.e. average income and proportion of college/university graduates, the colour scheme was reversed to use dark colour for the low values (e.g. low income). Zeroes and "no data" values were represented by a line hatch pattern that is distinct from any colour in the thematic symbology.

5. Results and Discussion

The Online Injury Atlas for Ontario was primarily an applied research and development project. As such, it was well received by the participating end-users. Extensive consultation of end-users from organizations with a mandate of injury prevention and injury research ensured that the atlas would be useful for its intended audience. The consultation process was already in place through earlier research projects, when injury-related data were acquired and integrated from a variety of sources with guidance from an external advisory board (Cusimano et al. 2007). Through a preliminary injury Web map development and test (Cinnamon et al. 2009b), a reasonable starting point for prototype development was available. This included the choice of interactive maps (not static or animated maps) from the UMN MapServer and the split-screen design.

The user evaluation of the atlas prototype yielded additional significant improvements. These included additional options in the map printing dialogue (landscape orientation, image resolution, and page layout), improved legibility of icons in the user interface (e.g. metadata link icon), and removal of redundant tools. In the first functional version of the prototype, different colour schemes were used for injury and contextual data, but end-users indicated that they preferred to have the same colour scheme applied to all thematic map layers. The restriction of the hospital

data layer to acute-care hospitals was also suggested by test participants. Generally though, the users tended to request the inclusion of more rather than less data. For example, additional health data such as those on ambulatory care were requested along with separation into more detailed age cohorts. These requests could not be implemented due to the resources required for data processing and due to confidentiality issues with prohibitively small case numbers (counts) within more detailed age groups.

During the development of the atlas, a number of interesting research questions emerged, which are discussed in the following paragraphs. These questions pertained to the cartographic design of an online atlas, its decision support functionality, the handling of data confidentiality and license issues within a spatial data infrastructure, and the use of open-source software to implement the service-based atlas architecture.

The nature of an atlas as a thematically coherent collection of geospatial and associated information influenced the cartographic design of the atlas maps. At the most basic level, different maps were visually related by choosing a single colour scheme for all maps. In addition, the same background data layers were offered for all thematic maps. At a higher level of abstraction, a decision was made to classify data based on the joint distribution across the entire study area and across levels of geography. This approach yields a single set of class breaks on all maps of a theme (e.g. household income), whether restricted to a City or showing the entire Province of Ontario. On the downside, some maps at finer levels of detail do not use the entire colour scheme because larger or smaller data values are geographically clustered in some variables.

The decision to use the “optimal” classification method rather than quantiles or the other methods available, and to manually adjust the classification, was not taken lightly. This method is more time-consuming than a strictly automated method, and manual adjustment opens the design to criticism regarding the objectivity and judgement of the classifier. However it is obvious that no automated method can be ideally suited for every data distribution, and investing the time in visual analysis of the resultant pattern and manual adjustment was deemed preferable and practicable in the context of this project. Ideally for data exploration purposes, the end-user should have a choice of classification methods.

The user interface design visually separates functionality into appropriate places on the display. Basic map navigation functions are always available near the mapped area. Thematic and background layer management is open by default. Functions that are less frequently used, such as map printing, are available through secondary tab selections.

Since the atlas contains sensitive information about the health of people within geographic areas, the data use agreement with the data provider required a number of safeguards for data security and confidentiality of individual patients:

- To provide appropriate physical structures to protect the spaces and computers that housed data (e.g. locked doors with alarms and appropriate institutional security services), electronic firewalls, and privacy contracts with every person that was to access the data;
- To aggregate the data and suppress the release of data for cells with counts of five or fewer individuals through any mode of observation (e.g. maps, data summaries);
- To require end-users to register and sign confidentiality agreements in order to inhibit any secondary uses of identifiable data;

- To prevent any linkage of data gathered in the research with other data about patients, whether those data were contained in public or personal records; and
- Not to identify individuals in any reporting on outcomes of the research project.

Accordingly, the injury rate map service was password protected and made available only to participants in the research project. By comparison, the Census data were subject to an agreement that focused more on data ownership by Statistics Canada than on privacy protection. Therefore, the Census map service could be offered without password protection by cartographically masking the raw Census data. To this end, the OpenGIS WMS was preferred over the WFS. A WMS server provides an image of a map to clients. The map will only show thematic symbols, thereby protecting data values. Accordingly, the identify function of this atlas is available for attributes of some orientation layers (e.g. hospital names), but not for the thematic layers.

In the spirit of spatial data infrastructures, it would be desirable if the data providers would implement authoritative Web services for access to their important datasets from their own servers. For example, a Census data service such as the one implemented within this project, should be offered by Statistics Canada. In the case of health data providers, such services would likely have to be limited to authorized users subject to ethics approvals and protocols.

The atlas also attempted to demonstrate the joint use of protected and public map services. However, at the time of development, several services offered through the CGDI were of limited use for a decision support situation. Some WMS services were not consistently available, or were too detailed and slow to load. This indicates the need for geospatial data providers to monitor and improve the efficiency and flexibility of their services. In the end, the minor roads WMS from the Geobase Web Mapping Service, Natural Resources Canada, was the only CGDI service used in the atlas.

Finally, the choice of open source software tools to build the atlas was considered a success. The advantage of open source software usually being available free of an initial charge and providing full, permanent control over software development was offset by the resources needed to gain an understanding of the tools, in particular CartoWeb, and develop customizations. While the documentation of CartoWeb had gaps and the tool itself is no longer being actively developed, its developer community was available for occasional assistance. Therefore, the overall experience with the use of open source tools for a sustainable implementation of a service-based, online atlas was positive.

6. Conclusion and Outlook

The development of the Online Injury Atlas for Ontario demonstrated the use of map services from a spatial data infrastructure in conjunction with local data layers for decision-making in injury prevention. The WMS specification was selected as means to secure confidential data through static map image overlay. The use of an open source Web mapping framework was successful despite the tradeoff between low initial system cost and higher demands on developer skills. Lessons learned in this project can hopefully be applied to other domains that deal with sensitive data, in particular in other fields within public health.

The iterative design process adopted for the atlas was worthwhile though costly with respect to researcher and developer time. It was essential to have access to injury prevention planners and

practitioners in order to acquire end-user feedback. The functional prototype provides a basis for additional user testing and application in specific injury prevention scenarios. For example, information relating to specific at-risk groups such as children or youth, or regarding specific injury types such as traffic-related injuries, can now be added efficiently to the atlas.

Cartographic design principles such as those referring to data classification and colour choice for choropleth mapping could be translated directly to the online atlas environment. Similarly, the coherence requirement on any atlas was easily translated into the navigation structure and map appearance of the Web atlas. Functionally, the current version of the atlas emphasizes simplicity within the balance of ease-of-use with analytical sophistication. Future development could add geoprocessing services to offer spatial analysis tools for advanced injury prevention research.

Acknowledgements

Partial funding for this research was provided by Natural Resources Canada through the GeoConnections program. The contributions by Mariange Beaudry and Paul Du to data processing and atlas implementation, and Bruce Huang assistance with setting up the Web map services, are gratefully acknowledged. Comments from three anonymous reviewers helped to improve the clarity and focus of the original manuscript.

References

Armstrong, M.P., Rushton, G., Zimmerman, D.L. (1999) Geographically Masking Health Data to Preserve Confidentiality. Statistics in Medicine 18(5): 497-526

Bell, N., Schuurman, N. (2010) GIS and Injury Prevention and Control: History, Challenges, and Opportunities. *International Journal of Environmental Research and Public Health* 7(3): 1002-1017

Bell, N., Schuurman, N., Hameed, S.M. (2008) Are injuries spatially related? Join-count spatial autocorrelation for small-area injury analysis. *Injury Prevention* 14: 346-353

Boulos, M.N., Honda, K. (2006) Web GIS in practice IV: publishing your health maps and connecting to remote WMS sources using the Open Source UMN MapServer and DM Solutions MapLab. *International Journal of Health Geographics* 5:6 (no page numbers)

Brewer, C.A., Pickle, L. (2002) Evaluation of Methods for Classifying Epidemiological Data on Choropleth Maps in Series. *Annals of the Association of American Geographers* 92 (4): 662-681

CIHI (2008) National Trauma Registry: 2007 Injury Hospitalizations Highlights Report. Dated 14 February 2008. Toronto, ON, Canada: Canadian Institute of Health Information. Available at http://www.cihi.ca/cihiweb/en/downloads/ntr_bl_highlights_feb2008_en.pdf [accessed 22 July 2010]

Cinnamon, J., Rinner, C., Cusimano, M.D., Marshall, S.P., Bekele, T., Hernandez, T., Glazier, R.H., Chipman, M. (2009a) Online Map Design for Public-Health Decision Makers. *cartographica* 44(4): 289-300

Cinnamon, J., Rinner, C., Cusimano, M.D., Marshall, S.P., Bekele, T., Hernandez, T., Glazier, R.H., Chipman, M. (2009b) Evaluating web-based static, animated and interactive maps for injury prevention. *Geospatial Health* 4(1): 3-16

Cromley, E. (2003) GIS and Disease. *Annual Review of Public Health* 24(1): 7-24

Croner, C.M. (2003) Public Health, GIS, and the Internet. *Annual Review of Public Health* 24(1): 57-82

Cusimano, M.D., Chipman, M., Glazier, R.H., Rinner, C., Marshall, S.P. (2007) Geomatics in Injury Prevention: The Science, the Potential and the Limitations. *Injury Prevention* 13(1): 51-56

Cusimano, M.D., Marshall, S.P., Rinner, C., Jiang, D., Chipman, M. (2010) Patterns of Urban Violent Injury: A Spatio-Temporal Analysis. *PLoS ONE* 5(1): e8669 (no page numbers)

D'Cunha, C. (2002) 2002 Chief Medical Officer of Health Report. Injury: Predictable and Preventable. Ontario Ministry of Health and Long-Term Care. Toronto, ON, Canada: Queen's Printer for Ontario. Available at http://www.health.gov.on.ca/english/public/pub/ministry_reports/injury_rep02/injury_rep.pdf [accessed 22 July 2010]

de La Beaujardiere, J. (2006) OpenGIS Web Map Server Implementation Specification. Version 1.3.0. Dated 15 March 2006. Available online through http://www.opengeospatial.org/standards/wms [accessed 22 July 2010]

Edelman, L.S. (2007) Using Geographic Information Systems in Injury Research. *Journal of Nursing Scholarship* 39(4): 306-311

Eyles, J. (1999). Health, environmental assessments and population health: tools for a complex process. *Canadian Journal of Public Health* 90 (Suppl.):1-5.

Haynes, R., Reading, R., Gale, S. (2003) Household and neighbourhood risks for injury to 5-14 year old children. *Social Science and Medicine* 57(4): 625-636

Kamadjeu, R., Tolentino, H. (2006) Web-based public health geographic information systems for resources-constrained environment using scalable vector graphics technology: a proof of concept applied to the expanded program on immunization data. International Journal of Health Geographics 5:24 (no page numbers)

Kiener, I., Philipona, C. (2007) Getting started with CartoWeb. Presentation at Free and Open Source Software for Geospatial (FOSS4G) conference, Lab L-07, CartoWeb for Beginners, September 24-27, 2007, Victoria, BC, Canada. Available at http://2007.foss4g.org/labs/L-07/Presentation.pdf [accessed 22 July 2010]

Kim, M.H., Subramanian, S.V., Kawachi, I., Kim, C.Y. (2007) Association between childhood fatal injuries and socioeconomic position at individual and area levels: a multilevel study. *Journal of Epidemiology and Community Health* 61: 135-140.

Kropla, B. (2005) *Beginning MapServer: open source GIS development.* New York: Apress/Springer

MacEachren, A.M., Crawford, S., Akella, M., Lengerich, G. (2008) Design and Implementation of a Model, Web-based, GIS-Enabled Cancer Atlas. *The Cartographic Journal* 45(4): 246-260

Maclachlan, J.C., Jerrett, M., Abernathy, T., Sears, M., Bunch, M.J. (2007) Mapping health on the Internet: a new tool for environmental justice and public health research. *Health and Place* 13: 72-86

McLafferty, S.L. (2003) GIS and Health Care. *Annual Review of Public Health* 24(1): 25-42

Moreno-Sanchez, R., Anderson, G., Cruz, J., Hayden, M. (2007) The potential for the use of Open Source Software and Open Specifications in creating Web-based cross-border health spatial information systems. International Journal of Geographical Information Science 21(10): 1135-1163

Open Geospatial Consortium [OGC] (2006) OGC History (abbreviated). Web page, Open Geospatial Consortium. Dated 2006-07-19. Available at http://www.opengeospatial.org/ogc/history [accessed 22 July 2010]

Peden, M., Oyegbite, K., Ozanne-Smith, J., Hyder, A.A., Branche, C., Fazlur Rahman A.K.M., Rivara, F., Bartolomeos, K. (2008) World report on child injury prevention. World Health Organization, WHO Press, Geneva, Switzerland

Rushton, G. (2003) Public health, GIS, and spatial analytic tools. *Annual Review of Public Health* 24(1): 43-56

Smartrisk (2009). The Economic Burden of Injury in Canada. Toronto, ON, Canada: SMARTRISK. Available at http://www.smartrisk.ca/index.php/burden [accessed 22 July 2010]

UNICEF (2001) A league table of child deaths by injury in rich nations. Innocenti Report Card No. 2, February 2001. The United Nations Children's Fund (UNICEF). Innocenti Research Centre, Florence, Italy

Vatsavai, R.R., Shekhar, S., Burk, T.E., Lime, S. (2006) UMN-MapServer: A High-Performance, Interoperable, and Open Source Web Mapping and Geo-spatial Analysis System. In M. Raubal, H.J. Miller, A.U. Frank, M.F. Goodchild (eds.): Geographic Information Science - Fourth International Conference, GIScience 2006, Münster, Germany, September 2006. Lecture Notes in Computer Science, Volume 4197. Berlin: Springer, pp. 400-417

Wartenberg, D., Thompson, W.D. (2010) Privacy Versus Public Health: The Impact of Current Confidentiality Rules. American Journal of Public Health 100(3): 407-412