STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE

Working Paper No. 17
English only

CONFERENCE OF EUROPEAN STATISTICIANS

Joint UNECE/EUROSTAT Work Session on Methodological Issues Involving the Integration of Statistics and Geography
(Tallinn, Estonia, 25-28 September 2001)

Topic (ii): New technological solutions, including those based on online data access

DISSEMINATION OF STATISTICAL GEOGRAPHIC INFORMATION BY MEANS OF NEW TECHNOLOGY

Submitted by Statistics Finland[1]

Contributed paper

I. INTRODUCTION

1. The dissemination of geographic information on the Internet can no longer be considered revolutionary for most statistical organisations. Most statistical organisations in the UNECE member states have either already organised delivery of their geographic information, data and services on the Internet, or have on-going projects and are planning to do so. The buzzwords today are LBS (Location Based Services) and DGDS (Dynamic Geographic Data Service), and it is here where our interest is focused. However, in practice, the Internet still represents new technology for many of us, and our next target of interest is mobile technology and its capabilities for statistical geographic information dissemination. In this paper, statistical geographic information consists of statistical data (located by co-ordinates or by regions), region boundaries, thematic maps, GI analyses on statistical data, and services based on the preceding subjects.

2. Challenging pioneering work requiring large investments has often been characteristic of first Internet applications. Even the doubters, if any, among statistical organisations have realised that data and services are bound to be delivered via the Internet sooner or later. Via the Internet, the information can be reached by any number of users, at any time (open 24 hours), independent of platforms used. Faster publishing schedules and information updating are also benefits compared to paper or other media publishing. When we talked about planning dissemination via the Internet some years ago, most of us thought it meant dissemination through a PC screen. Today, we have to plan dissemination via yet another new means - mobile terminals. Using mobile terminals, information can be accessed almost anywhere.

3. This paper is intended to stimulate general discussion about the experiences and basic questions that concern dissemination via new means. It is based on the experiences of the Geographic Information Team at Statistics Finland and on the views and questions the Team has on the dissemination of GI data via new means of technology.

[1] Prepared by Ulla-Maarit Saarinen.

II. INFORMATION DISSEMINATING TO CUSTOMERS – TO WHO ARE WE DISSEMINATING OUR DATA AND PRODUCTS?

4. The Internet seems such an obvious dissemination tool for information that we do not always remember to ask ourselves or our customers whether it is in fact the best way to deliver information. In the 1990s, customers frequently asked when GI information would be made available on the Internet: when could they download, for example, grid square data or regional boundary data from the Net to their own systems. The advantages of the Net were that they could see right away what data were available without having to wait a week, or even weeks, for the data they needed. Instead, they could just pick up the required variables and regions and obtain what they needed. Based on such demand, plans are underway to make all data, products and services available on the Internet.

5. The Internet is the best way to deliver data and thematic maps when, for example, election results are needed quickly for the media. Time is of the essence in this kind of service and for these kinds of customers. Media journalists and graphic designers are accustomed to using the Internet as an easy and rapid way of obtaining what they need. However, we have other kinds of customers, too, who use the Internet but are not familiar with complicated applications. To obtain their information, are they willing to select and then select again from an Internet application, click one button and then yet another and still not necessarily find the desired information? What about the professional GIS customers? Although demand for 24-hour data dissemination service via the Internet already existed in the 1990s, it appears that not all GI data or products are necessarily required to be delivered via the Internet. The Geographic Information Team at Statistics Finland has made some customer surveys to develop its services but although these show that self-service, i.e. service via the Internet, is desirable, not all delivery of data and products is requested this way. There is still demand for specially compiled data and accompanying information through personal contact. We can surely not place all the responsibility regarding what data to pick up and how, or the way they should be used, etc., on our customers? Although we may feel that our customers show a great appetite for increasingly rapid services via the Web, face-to-face discussions about their needs and alternative solutions to these needs are still an essential element of customer service. The conclusion is that our enthusiasm to put everything on the Internet is not necessarily the best or only solution in all cases. This seems to be very clear to everybody in theory, but seems often to be forgotten in practice.

6. Before planning services via the Web, we must recognise our GI customers, as we do when planning any other products. GI customers can be divided into three main groups: i) GIS professionals, ii) users of statistical data (entrepreneurs, researchers, students), and iii) citizens (including students). Citizens and students usually look for free, basic information. Professional users of GIS or statistical data need more in-depth information. Do we know how the customers use the Net and what they use the retrieved information for? Do they buy services or data? All this seems very clear: of course people use the Internet and buy products via Internet, and we know what they use the data for. However, in our enthusiasm to create new GIS services we may forget to discuss these issues thoroughly. If our designed GIS products or organised Internet services are really based on customers' views and needs, we should not have a problem. Specific questions arise if we are aiming for the global market with GI data dissemination. To who are we delivering our data is not the only question we should address when we offer products and services via the Internet. We should also consider the means by which we offer them, as well as the terminals our customers use. Do we want to disseminate information to different terminals and do we have the resources to do so?

7. The pricing of service via the Internet is yet another question. Each organisation has its own policy as to what is free and what is chargeable. Organisations may expect that a service offered via the Internet, e.g. an Internet map service producing thematic maps, should be at least affordable: the investment costs in building a service are often high, but the income from the service does not necessarily cover the costs. Are there enough potential customers who are willing to pay for viewing maps and data on maps or for picking up data by map location? On the other hand, the public may generally expect that everything on the Internet should be free and this puts pressure on the policies of those organisations that

do not disseminate all their data free of charge. GI and maps may be reduced to ”playing the second fiddle” in the planning of services on the Internet. They are not a necessity; they are something extra, meaning that they also generate extra costs.

III. THE STEPS WITH DELIVERY VIA ”NEW TECHNOLOGY”

8. The first step toward proving a service on the Internet has often been a site for presenting GI products and data, or some static image maps and statistical tables which, perhaps, could have been printed out or have been downloadable. This stage does not provide much interactivity for customers or citizens; it rather resembles a show window for products and services. The next step often is the building of an Internet application, in most cases a map interface with more or less interactivity. Other, additional ”services” or sites may also be offered, such as educational material on GIS subjects for schools or chat sites for GIS professionals.

III.1 Map applications

9. Internet map application is probably the most illustrative means for presenting and disseminating geographic data. A map interface can offer facilities for examining statistical data on thematic maps according to the customer’s choice, downloading maps, viewing data concerning specific areas, etc. The most challenging applications allow the user to make queries on personally defined areas, to create buffers and diverse spatial analyses.

10. Technology develops rapidly, as we know. The user interfaces produced a couple of years ago may already seem outdated today, with their inflexible functions.

III.1.1 Database

11. Whether Internet map applications are integrated with other data dissemination via the Internet or not, there is still the problem of how to organise the database. How to solve a situation where the statistical data are not in a relational database, but the map application requires them to be in one. This is not (just) a technical issue. Database solutions in statistical organisations have not always been made with a view to geographical information. Neither have they been made with a view to dissemination via the Internet. The organisation may assume that GIS demands something else, something extra, and why can they not be satisfied with the decisions concerning database solutions that have been made for (other) statistical data dissemination systems in the organisation. The situation gets even more complicated for GIS if the organisation has several database solutions for data dissemination and none of them is relational or is easily convertible to a relational database.

12. The data disseminated on the Internet is usually aggregated data, not based on individual persons. If we are planning an application where the user, for example, has the possibility to define for him- or herself a particular area on a map on which he or she would like to obtain population data, we must consider how appropriate data protection could be executed.

III.1.2 Technical solutions

13. The first map applications are often built with software where all the functions take place on the server. Questions and answers travel between the client and the server. This makes the service slow. However, this problem has been eliminated with the latest software.

14. A problem may also arise if the adopted technical solutions cannot be made compatible with existing ones, or the solution requires a lot of searching or has to be created from scratch (e.g. aforementioned database solutions).

15. New Internet software may offer simple and easy-to-use software for building a map application for the Internet environment, but as they are not designed for presenting statistical data they may require a lot of ”tailoring” before a proper application for displaying statistical data can be made of them.

16. Classifications of regional statistical data have not played a key role in the planning of these software, as they have not been in GIS desktop software, either. As presenting data in the best possible way means using the best possible classification for them, these software still leave a lot to be desired. If data classification method possibilities are limited to quantiles or equal intervals, good results cannot be achieved.

III.1.3 Interface: Usability

17. ”The WWW is a fast medium used by impatient people”, as Menno-Jan Kraak has said. Its users also demand functionality and visualisation. They expect from it the same degree of usability as from the Windows software they use in their PCs. The adopted technical solution may provide the possibility for building into the user interface a host of functions for e.g. area selection, searching, locating, defining, creating thematic maps and classifying them, and for selecting colours or other statistical graphics, or for searching statistical data, tables reports, background data or articles on the selected topic. However, the usability of the interface may get stretched if the offered functions are too plentiful, especially if it has not been carefully designed and tested. As with any other service offered via the Internet, the designing of map interfaces also calls for highly diverse skills in many fields.

18. We can develop our services not only by studying user statistics but also by examining the users' behaviour in the application.

III.1.4 Web cartography

19. The first maps on the Internet have usually been static GIF or JPEG images that have been either scanned from paper maps or converted from maps intended for printing out on paper. However, maps made for printing on paper cannot be used for the Internet environment as such. Screen maps are not the same as paper maps.

20. There are plenty of map interfaces on the Internet and commercial Web services that offer the possibility to use maps. These interfaces/applications allow the users to locate places or search street addresses, etc. They have maps with street names, lakes, rivers, etc. Those who have used these map services expect the same possibilities from statistical organisations' map services. If we build a simple statistical thematic map application with a minimum of functions or without proper street maps, our product may not be regarded as being first quality. Achieving first quality requires financial resources and a statistical organisation may not have the possibilities to compete with commercial ”competitors” if map applications or GIS are not valued enough in the organisation.

21. It is obvious that the Internet will change the concept of a map. Animations, such as overflying views over regions, or 3D or virtual cities will create new ”maps”. The Internet is a medium that is based much more on visualisation than publications are. This imposes demands on visualisation. Professor of Cartography Kirsi Virrantaus (University of Technology, Helsinki) has ”warned” that if cartographers do not make the standards for Web maps, there is a danger that Web graphic designers will make the maps, but do they know how to make cartographically correct maps?

III.2 Positioning system

22. The next step after map applications in the planning of GI dissemination could be consideration of wireless means. The enthusiasm for map visualisation on wireless terminals has, however, been limited by small screens. However, the biggest question mark has been uncertainty about their development and about which media and services the consumers are likely to adopt as their own.

23. Nevertheless, wireless or mobile terminals will be tomorrow's media (as they already are today's). Therefore, we must consider or plan our role in locating systems, location-based services and in mobile terminal services.

24. Mobile terminals have introduced the possibility of locating persons, vehicles and other moving objects. The locating feature of mobile networks is no longer news but the possibilities they offer for purposes other than just for mobile telecommunication network operators' internal use have become topical in recent years. Locating systems have made it possible to geographically pin down emergency calls and other emergency cases, such as lost persons. Locating also makes it possible to give mobile terminal users information about the routes and services (e.g. the nearest restaurant) that would best suit their present position. On the other hand, location-based services suppliers will be able to advertise their services (e.g. a restaurant) to those who are within a certain distance from it and, for instance, offer them benefits.

25. What do these location-aware systems have to offer to statistical information? Location-based mobile networks enable the collection of geographic information in a certain location. The networks also enable mobile terminal users in a given location to retrieve certain information, for instance, information related to their location. All this could happen in real-time. We could collect data on the location of daytime population and night-time population via mobile terminals, or any other location-aware statistical data even by the hour, if most of the population had personal mobile terminals. We could collect data from a certain region at a certain moment in time by sending questionnaires to the mobile terminals of the persons in the region concerned.

26. At present, national legislation defines what kind information can be made available concerning the position of a person and to whom this information may be released. Privacy laws protect the information on an individual's position. For example in Finland, legislation allows the locating of emergency calls. This means that, when required, the operator has to give the position information for emergency calls. In specified cases, the police also have the right to locate the position of a mobile terminal. This right will now be legally extended to the location of missing persons. Extending this right to the collection of position information for reasons other than emergency cases is not likely to happen next year, but it will be a question that needs addressing in the near future. The releasing of information on the position of a person (i.e. his/her mobile terminal) must be based on the voluntary of that person, as Jouni Markkula in his article "Dynamic Geographic Personal Data – New Opportunity and Challenge Introduced by Location-Aware Mobile Networks " has written.

27. Before these kinds of data collections and location-based statistical services are possible, there are many questions to be answered. One question to start with is what kind of statistical geographic data will people or customers need when they are not sitting behind their desks? Could we offer, for example, data about the neighbourhood in which someone with a mobile terminal is standing at the moment? More questions will arise, for example, about the administration and processing of data (e.g. would the information be released by operators as aggregated data without personal identifiers?)

28. More information about this subject is available in an article written by Jouni Markkula from the Information Technology Research Institute at the University of Jyväskylä, Finland. Markkula has studied dynamic geographical personal data collection by location-aware mobile networks and his article will be released in Cluster Computing, 4/2001.

IV. FORTHCOMING DEVELOPMENT

29. Whether we are enthusiastically involved in new technology or consider it as hype, it is difficult to forecast for a period longer than a year or couple of years exactly what the future of disseminating statistical GI will be. We do know that, over the next two years, one half of Web use will be via wireless terminals . . . (probably). We know that, in future, the location of individuals will be easier to determine through the wearing of wristwatches, for example, emitting signals to satellites . . . (probably). On the

other hand, we do not know for sure what kind of terminals people will be using in five years' time. We cannot say whether the development will be the same on all continents. For example, will palms overtake the mobile phones or will smart phones be *the* ones in the next 2-3 years? Will digital television be one way for ordinary citizens to access the Internet in the near future? We do not have an answer for this, but someone has said that if you want to know what media people will use or how they will behave tomorrow, see what the children are doing today.

30. Will XML or GML (Geography Markup Language) be the solution to our many dissemination channels? What will there be and what will we need in the next few years?

31. The future offers plenty of opportunities. The positioning technology of mobile equipment will make location-based services into essential everyday services. The possible implications from this to statistical data tied to locations is an issue most of us are reflecting upon at the moment.

32. Does new technology change our role in the dissemination of statistical GI data and services? Can we compete with the private sector in the latest technology for the offering of services? Will too many technical choices to be managed exist, or will the customers have too many terminals for the acquiring of information? Will there be too many needs? These questions do not only concern the dissemination of GI, but are also relevant as regards the GIS in statistical organisations. Do we have the capabilities (financial, technological, etc.) to build our own services or should we be active in co-operating with other professionals the in GIS field in order to build joint projects that would generate joint end products? Do we have to do everything ourselves? Would it perhaps not be more beneficial to seek active co-operation with others operating in the field? Are our skills and resources sufficient for keeping abreast with the development in this field? Do all the Web services have to be produced in-house or should we co-operate with our customers to build services that suit them. Do statistical organisations have to have all the data? Would a server hotel be the solution? For example, could we use map information (e.g. road network) from the servers of other service suppliers? How do we offer our data, products and services, if the services are centred on portals where the information and services originate from several different sources?

REFERENCES

Markkula, Jouni (2001). Dynamic Geographic Personal Data – New Opportunity and Challenge Introduced by Location-Aware Mobile Networks. Cluster Computing, 4/2001.

Web Cartography – development and prospects (2000). Edited by Kraak, Menno-Jan and Allan Brown. London, Taylor & Francis.