STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS

Working Paper No. 8
English only

Work Session on Geographical Information Systems
(Ottawa, Canada, 5-7 October 1998)

Item 5 of the provisional agenda

STATISTICS AND MAPS, TOGETHER ON THE WEB

Submitted by Statistics Sweden [1]

CONTRIBUTED PAPER

[1] Prepared by Lennart Nordström.

I. STATISTICS AND MAPS ON THE INTERNET

1. Statistics Sweden, Sweden's National Statistical Institute (NSI) and the National Land Survey (NLS) in Sweden have a mutual interest in smooth co-operation and joint development of marketable Geographical Information Systems (GIS) products. Statistical and geographical components are obviously equally important for the two agencies in their GIS development work; to obtain a spatial dimension in statistics is a value-added feature in presentation, analysis and sampling for Statistics Sweden and access to statistics gives the NLS the attributes to the geographical objects.

2. Co-operation between the two agencies has been going on for many years. In the last decade, data has been bought and sold on a cost-recovery basis. The two agencies felt increasingly that focus should be adjusted to the needs of end-users instead of to the needs of the agencies *per se*.

II WHY INTERNET?

3. During 1996, the two agencies decided to look at possibilities of working closer together in developing geo-statistical systems. It was decided that a joint project should look at a marketable service on the Internet, including information such as i) topo-cadastral background maps, ii) direct access to real estate information at object level and c) a wide range of geo-referenceable statistics.

4. Each agency has a high-quality web-site of its own, so why Internet? There was a handful of contributing reasons, some relating to what other media had actually failed to achieve and some to the unique features of the Internet itself.

5. One important experience among the project group members was that, in spite of the excellent pre-conditions in Sweden for implementing GIS, the technology was not really taking off as planned. The fact that GIS software is still too complex for certain user-groups would have to generate alternative solutions for middle-end GIS, i.e. for categories that cannot be labelled as ”doers” but as ”users/occasional users”, or ”viewers”. A GIS without GIS technology would be the solution for these categories.

6. Internet is the *only* interactive mass medium and has seen the development from a ”funny thing” to surf on, to a main electronic highway in a very short time. The medium allows frequent up-dating of information and better information flexibility than any other medium.

7. It was also believed that new potential client-groups could be reached. These potential clients would already be consuming qualified information (but not necessarily statistics or statistics from the NSI), would be computer-literate (but not experienced in the use of GIS) and working with real estate-related issues (but not utilising real estate information or land information sources to their full extent).

8. The project also saw potentially improved cost efficiency and increasing revenue in GIS on the Internet, as information would be paid for only when used and compilation of information would be made directly by the user, without go-betweens from one, or even two, agencies. From a corporate point of view, Internet is furthermore destined to become a major medium for dissemination of statistics, cheaper and more cost-effective than printed or other electronic media for users as well as for producers for many information purposes.

9. The project group also argued that command of Internet technology would provide solutions to in-house information management problems (via Intranet) and bilateral communication network facilities with other agencies (Extranet).

10. The NSI/NLS Internet service which the project group was challenged to develop would have to be an entirely commercial product.

III. PROJECT PROCEEDINGS IN BRIEF

11. The primary target group defined for the pilot test was narrowed down to organisations and institutions for whom the location of statistical objects and information on real-estate at object level would be of primary interest; in other words, insurance companies, real estate-brokers and -companies, banks and credit-rating agencies. From this group, three interested organisations eventually became pilot clients and prototypes for a customised Internet-service were developed. This was done during the period June to August 1997. The test pilots' reactions to the developed products were very favourable and they pushed for a rapid continuation of the development work.

12. The two agencies each invested approximately 200 man-hours in the pilot project. Final reporting was carried out on time in October 1997. A full-scale development of the Internet service was suggested together with tentative budgets and a tight time-schedule for further development. The two agencies signed a three-year agreement in April 1998. A virtual editorial function was established in May 1998 and a steering committee was appointed. The service was eventually baptised *SwedeFacts - Stats on Maps* and its launch date was set for October 1998.

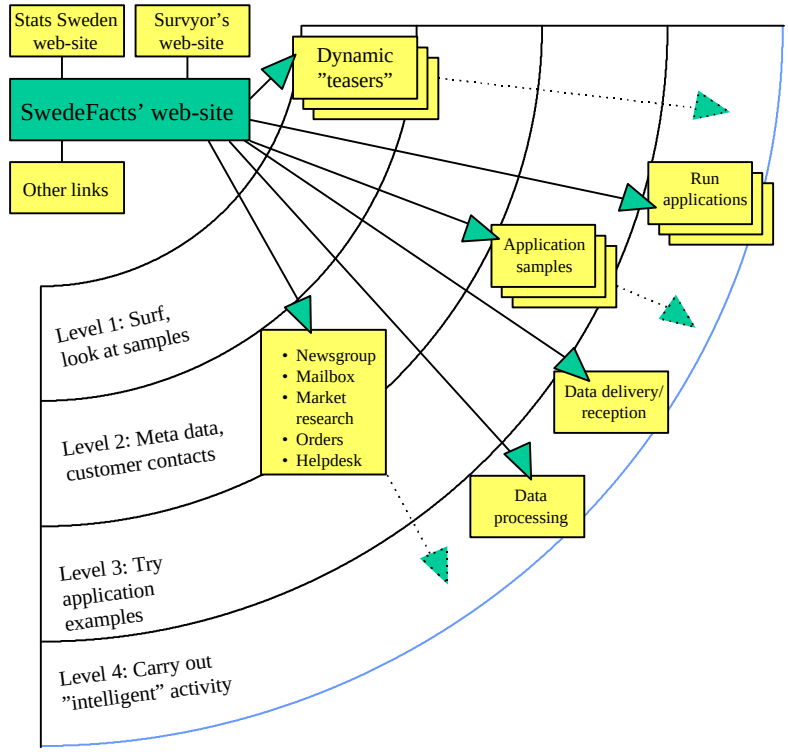
IV. *SWEDEFACTS* - AMBITIONS AND CONTENTS

IV.1 An Overview

13. The graph below was used by the project to illustrate various levels of interaction, from the simplest level where the potential client is exposed to the broad repertoire of services through ready-made illustrative examples of how statistical and geographical information systems can be made to interact, to the most advanced level where clients can access statistical and spatial meta- and macro-databases, even a database with real estate information at object level.

14. The presentation in this paper will focus on level 4 of this graph, the commercial level, towards which all activities in this concept are aimed; as mentioned earlier, *SwedeFacts* is a purely commercial venture.

Graph 1: Geo-statistical Data Levels on the Internet



15. One of the most important features in *SwedeFacts* is that the client *herself/himself* will be able to combine statistical and geographical data at various levels of resolution. This will, in its simplest form, encompass statistics and maps for counties, municipalities and parishes. Smaller administrative areas such as zip-code areas, small NUTS-areas, etc. will be included and the possibility to define buffer zones and hinterlands around point and line features will be provided.

16. Various types of real estate information from the cadastral register at NLS will be included for individual units and for aggregates. Two types of services will be provided:

- a "smorgasbord" of pre-defined information services provided on a "help yourself" basis and charged according to utilisation of data. Certain standardised and special packages can be added to the basic service.

An example of a pre-defined service: thematic maps for population structure in counties and municipalities. A standardised addition to this could, for example, be i) information on individual objects as land parcels or local units or ii) geo-coding of street addresses;

- the creation of a client profile, a customised system with solutions for the individual client, i.e. a corporate GIS on Internet with client-dedicated provision of data and applications. The profile is created in close co-operation with the client.

IV.2 Statistical areas and variables in *swedefacts*

17. The following areas/variables will be included in *SwedeFacts*:

Population: (population by age and sex in broad age groups; population change over the last five years);
Housing: (population in detached/other houses; families in detached/other houses);
Employment: (population aged 16-64, at work, with workplace in area X, aged 16-64, by industry/change over time);
Income: (disposable income by household and by income class; median income/change over time);
Education: (educational level for population 16 years and older);
Local business units: (statistics on local units by branch and size class; later also information on individual objects);
Cadastral information: statistics for land parcels; individual units or aggregates by type code, area, ownership, age/year of construction and changes of ownership over time.

18. At a later stage of development, functions for calculation of distances between buildings, stock by ownership over time, links between land parcel and administrative area and location of productive forest land will be added.

IV.3 Spatial information in *swedefacts*

19. The map repertoire includes topographic maps with administrative boundaries for county, municipality, parish, zip code and grid maps, with 1x1 km grid in rural areas and 500 x 500 metres in urban areas.

20. The digital maps used in *SwedeFacts* are:

Red Map/Röda kartan in 1:250 000, featuring land use, administrative subdivisions, road network, railroad network, urban delineation and hydrography;
the urban area map/Tätort 2000 in 1:25 000, featuring public buildings, parks, hydrography and street network;
cadastral map/ekonomiska kartan in 1:10 000; orthophotos with 1m resolution will be added in a later version.

IV.4 Functions for geographical analysis

21. The generation of various types of thematic maps will be a standard feature in *SwedeFacts*. These thematic maps will be more or less static in their function and their task will be to illustrate distributions of statistics and register information and to provide information on objects on the basically static map.

22. Geographical analysis will be a more flexible and problem-oriented process and will become a gradually developed, additional feature in *SwedeFacts*. Geographical analysis will provide functions for the creation of circle zones, buffer zones along line features, overlay and conditional analysis, analysis of so-called twins, cluster analysis, time-related analysis and analysis of two-point data such as migration and commuting. Functions for population forecasts will also be included.

IV.5 System architecture

23. At the time of writing, the routines for data supply are being developed. The drafted architecture of *SwedeFacts* is the following: *SwedeFacts* will be accessed using a standard web-browser (Netscape or Internet Explorer). The client will communicate with the web server in

HTML code. The web server is linked to the map server, which uses *ArcView Internet Map Server* and/or *MapObjects Internet Map Server*. Analytical functions will be executed in the map server and results and presentations communicated back to the web-server for further transportation in HTML code back to the client.

24. Statistical and cadastral data sets will be copied onto the database servers, while direct connections with the data suppliers is a matter for the future. Geographical data sets will be loaded on to the UNIX database server with Oracle and SDE via the local area network (TCP/IP) at the National Land Survey in Gävle, Sweden.

V. VALUE-ADDED EFFECTS

V.1 Information Contents and Market

25. The implementation of *SwedeFacts* is still underway and it is, of course, too early to evaluate the effects of providing an integrated geo-statistical service on Internet. Statistics Sweden and the National Land Survey, however, expect the effects listed below.

26. The role of the two government agencies has changed over the last decade. Earlier monolithic status in statistics and land information is no longer axiomatic and products and services have to become broader *and* more advanced. The future of the agencies will be dependent on how well they can adjust to the signals from the market.

27. *SwedeFacts* is launched as a complement to other types of commissioned work. Its aim is to expand the market for geo-statistics. Certain GIS products which are too costly to produce with traditional user-paid commissioned work will now become financially feasible for certain client groups.

28. *SwedeFacts* will also become a means of providing geo-statistical information to institutions and organisations who, for various reasons, would not consider investing in GIS technology and competence. Internet is obviously easier to use for non-experts and more widespread than traditional GIS technology and methods.

V.2 Internal Effects

29. For Statistics Sweden, *SwedeFacts* means an alternative use of limited staff and funds. We believe that these scarce resources will be used more efficiently if they focus to a larger extent on building infrastructure for basic geo-statistical needs that can be used by many. Individualised attention to clients would then be of a higher complexicity.

30. Experiences from the development of *SwedeFacts* can be used in a broader context, e.g. for the overall dissemination of statistics. Internet-related technology is comparatively inexpensive. The improved output should be possible to achieve at a lower cost and with potentially higher revenue, making this Internet venture a good business opportunity for users and producers.

31. Last, but not least, *SwedeFacts* has become reality through the close co-operation between Statistics Sweden and the National Land Survey of Sweden; this *entente cordiale* is a very promising pact for the future of geo-statistics.