

INTERNET MAPPING FOR DISSEMINATION OF STATISTICAL INFORMATION*

Gennady Andrienko¹, Natalia Andrienko¹, Hans Voss¹, and Jackie Carter²

¹ GMD - German National Research Center for Information Technology
Germany

² Manchester Computing, University of Manchester, UK

ABSTRACT

City and regional administrations find increasing interest in providing mapping services to their customers: citizens, businesses, non-governmental organizations, etc. The paper discusses possible architectures of implementing such services in the Internet, and then describes our own system, Descartes, in more detail. Descartes can be used in client/server mode, where the interactive client interface can be downloaded as a Java Applet. Descartes is particularly suited for a Web-based analysis and publication of geo-referenced statistical data. The functionality and user interface of Descartes is illustrated by some example applications of the Cities of Bonn and Tilburg and of the UK Census of Population.

1 INTRODUCTION

Geographic Information Systems (GIS) are widely used as tools supporting the management of spatial information and various map-mediated activities. One such activity is thematic mapping, i.e. presentation of spatially referenced attribute data on maps. As city and regional administrations and statistical offices increasingly understand their role as customer-oriented service centers they become interested in providing mapping services to their customers through the Internet. This includes providing wide and easy-to-use access to spatially referenced data, e.g. about demographic or economic indices. Such information is valuable to citizens, local interest groups like environmental organizations, local companies, etc. Services of this kind are also valuable to (yet) non-local "customers" like companies investigating new or additional potential locations, or private persons looking for a location for their new home when securing a new job in the city. Beside the inherent value of having such information at hand the very fact that it is provided is seen as an indication of the positive general attitude of the City (or Region) Administration, and thus fosters the attractiveness of the City for those living there or considering it as a new place to live or to do business.

Section 2 of the paper discusses pros and cons of alternative architectures for Internet mapping services. A general distinction is made with regard to the client interface in terms of local, direct map and data manipulation functions. Section 3 presents some general features of our own solution, Descartes, which is particularly well suited to provision of Internet access to maps and geo-referenced statistical data to a wider public. With its full functionality it may also be used as a Web-based tool for the analysis of such data by professional users. Section 4

* The paper extends the presentations made by the authors on the UDMS '99 (Venice) and TeleGeo'99 (Lyon) conferences

presents further examples¹ of using Descartes, first for analyzing geo-referenced data for the cities of Bonn and Tilburg, and second, applying the ShowMap variant of Descartes for publishing election data of the City of Bonn as interactive images inside of normal Web pages.

The last section of the paper describes application of Descartes to a subset of the 1991 UK Census of Population; a voluminous digital collection of socio-demographic data which is freely available to the UK academic community. Descartes is being employed in order to develop a visualisation gateway to this valuable data resource, in order that users will be able to assess the appropriateness of the data for their use, through visual exploration, before downloading for analysis locally.

2 ORGANIZATION OF MAPPING SERVICES IN THE INTERNET: POSSIBLE SOLUTIONS

Most GIS vendors propose solutions for Internet mapping on the basis of their existing packages. The typical architecture involves a GIS running as a CGI script (Grossley & Boston, 1996, Plewe, 1997) on an Internet server computer (see Fig. 1). For example, versions 2 of the CIESIN Demographic Viewer, Canadian census WWW gateway, and Pennsylvania Statistics² have similar architecture. The interaction between the user and the system is done through CGI forms displayed by a World-Wide Web (WWW) browser on an Internet-connected client computer. CGI forms provide a rudimentary graphical user interface. A user's input by means of a form is transformed into commands to the GIS. In response the GIS generates a map and sends it to the client computer as a raster image, typically in the GIF format. The limitations of this approach are obvious:

- restricted user interface;
- low interactivity (even such a standard operation as zooming is impeded);
- insufficient quality of graphics (limited not only by the speed of the Internet connection but also by some restrictions of the GIF format);
- substantial delays when large raster images are sent from the server to the client in each transaction.

Simple interactivity for server-side mapping can be implemented in the JavaScript language. A script may enable a user to select one out of a set of previously prepared GIF images to be shown on the page. For example, if somebody wants to highlight one of 10 objects presented on a map, he should prepare 11 maps: one without highlighting and 10 highlighting one object each. Once all the images have been downloaded to the client computer the script allows reasonably fast interaction with them on the screen³. In the same way Peterson (1999) implemented simple dynamic mapping and animation. The advantage of this approach is that JavaScript programs can be interpreted by most browsers. The language is very simple and can be used by non-programmers. However, visualizations are limited by raster GIF files, and interaction capabilities are very restricted.

¹ All examples shown in the paper, and many more, are publicly available as demonstrators at <http://allanon.gmd.de/and/java/iris/>. You will also find examples and further information at the home page of Dialogis Software and Services GmbH, a spin-off company of GMD that markets a commercial version of Descartes, cf. <http://www.dialogis.de/>

² URLs <http://atlas.gc.ca/cgi-bin/hsi/gmap/schoolnet/issuemap/> and <http://www.maproom.psu.edu/>

³ See URLs <http://www.geog.fu-berlin.de/bb/themen/politik/brburg/k98/k98bb.html> and <http://maps.unomaha.edu/AnimArt/ActiveLegend/Peterson.html>

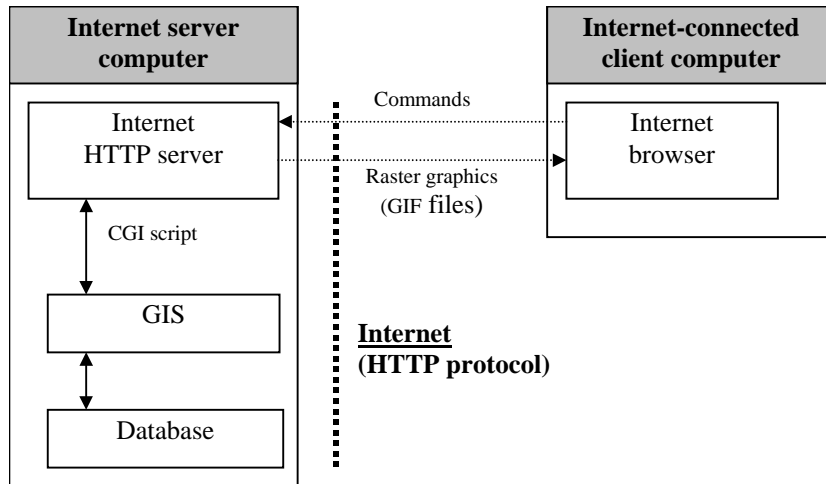
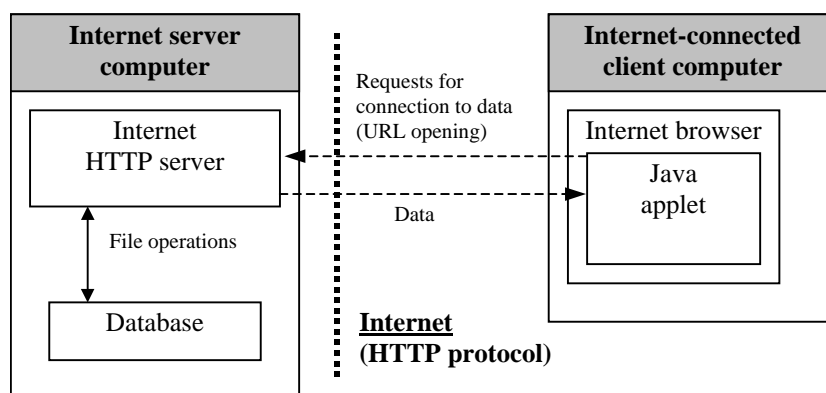


Fig.1. Internet mapping: server-side GIS running as CGI script.

Another solution consists in the use of Java applets (Fig. 2). Java is a programming language that can be interpreted by most Internet browsers on various hardware platforms. Java applets embedded in WWW pages are loaded by a browser from an Internet server computer and then executed on a client computer without requiring any installation. Applets can access data stored on the server through the HTTP protocol. Advantages of this solution are the opportunities for providing an advanced user interface, unlimited user interaction, and generation of high-quality graphics.

This approach also has certain shortcomings. First, a critical issue is the size of the Java applet: downloading of a big applet may require quite a long time. This may seriously inhibit augmentation of the functionality. Second, applets may run rather slowly on some hardware and software platforms. Third, data security is not guaranteed: any other software can access database files accessed by the applet. Although these files can be encrypted, this is not a perfect solution: it is possible to open the encryption mechanism through decompiling the code of the applet. For these reasons most existing mapping applets only provide very basic map-viewing capabilities⁴ and are not used for real problem solving (Sorokine & Merzliakova, 1998).



⁴ See URLs <http://www.lutumtappert.de/java/wahl98.htm> and <http://www.vtt.co.jp/staff/sorokin/mt/>

Fig.2. Internet mapping: client-side GIS Java applet

Sometimes, instead of Java, other interpretable programming languages, such as Tcl/Tk, are used. An example is the system for visual analysis of spatially referenced data named CDV (Dykes, 1997 & 1998). In such a case, to make the system work in the Internet, special software called browser plug-ins must be installed on the client computer⁵. This requirement may actually be a serious limitation because end users may have insufficient competence to carry out the installation or be unwilling to do it.

The available Internet mapping software, like all GIS in general, is rather weak with regard to the cartographic presentation of spatially referenced (statistical) data. Most GIS developers seem to consider this function as marginal. In contrast, our system Descartes is primarily designed to support visual exploration of spatially referenced data through automated cartographic visualization. Correctness of map presentations, i.e. compliance with the established principles of graphic and cartographic design, is very important for proper data interpretation and, as a consequence, for problem solving. As we do not want to presume that all potential end users are familiar with these principles, Descartes incorporates such knowledge about thematic data mapping, and on this basis acts as a professional cartographer providing users with automatically built maps appropriately presenting user-selected data.

The maps built by Descartes are not just static pictures. The system offers a wide range of interactive tools to manipulate the presentations in order to make them more expressive and to create more beneficial conditions for revealing significant relationships, patterns, and regularities.

The services of Descartes are available in the Internet. For this purpose, the functionality of the system is distributed between two parts: the server and the client. The server⁶ provides database access, data manipulation (fulfillment of queries, calculations, transformation of structure of tables etc.), and map design (selection of presentation methods and setting of their parameters using the internal knowledge base on thematic mapping). The client provides the user interface, draws the vector-based maps, and supports interactive manipulation of these maps.

The client is a Java applet implemented in Java 1.0.2. Most Internet browsers such as Netscape versions from 3.0 to 4.5 and Internet Explorer versions from 3.0 to 5.0 support this version of Java. When a user loads the Internet page with the applet, it activates the server part of the system and starts communication with it through a socket connection. This approach allows combining high user interactivity and dynamic data displays on the client side with powerful functionality on the server side while the size of the applet can be kept relatively small. We admit, however, that our approach also may exhibit a drawback: some local networks are protected by firewalls that prohibit socket operations and thus forbid to run Descartes applications across the firewall. A general scheme of Descartes' functionality is presented in Fig. 3.

⁵ The same applies to the latest versions of the Java language that are not supported yet by most Internet browsers. Unfortunately, different versions of the Java plug-in are not completely compatible, and this may cause some trouble for end-users.

⁶ Implemented in C++ and is available for different dialects of Unix and for Windows 95 and NT.

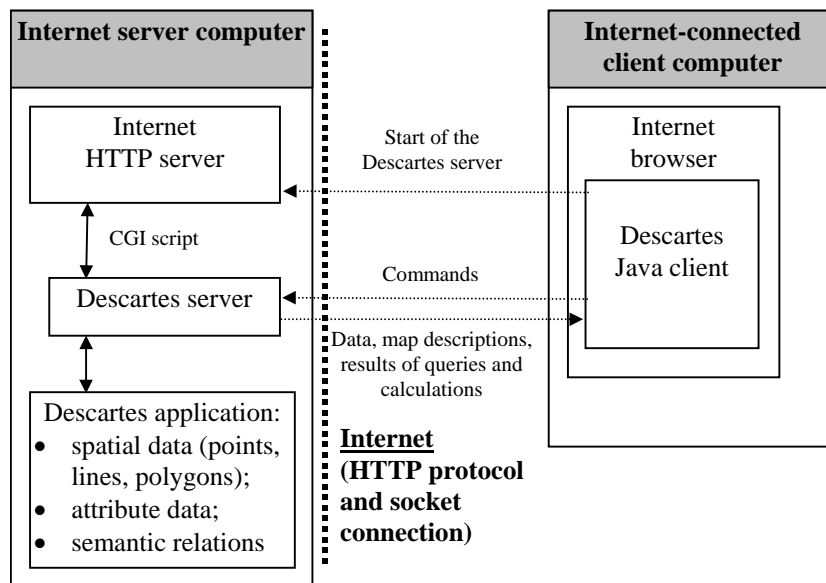


Fig.3. Internet mapping in Descartes: client-server architecture with the client in Java.

3 MAPPING ATTRIBUTE DATA IN DESCARTES

When designing a map Descartes considers several characteristics of the data variables that the user wants to visualize: the number of data variables selected for simultaneous presentation, their types, the number of different values, and semantic relationships (IS-A, PART-OF, and comparability)⁷ between the variables. All these characteristics can be found in the internal database schema or derived from the data.

A table in Descartes is a collection of uniform records (rows) composed of fields (columns). The data must refer to some *geographical objects* listed in one of the columns. A file with coordinates or outlines of these objects in vector format is part of the application. Fig. 4 shows how Descartes displays a table⁸. Note the organization of the table caption: its generation is based on the hierarchical table index supplied to it in the application building process, and it also accounts for relationships between attributes.

⁷ Examples of semantic relationships: “Female population” part-of “Total population”, “Mammals” is-a “Animals”, “Birth rate” comparable-with “Death rate”. These semantic relationships are once described in the application building process by a person familiar with the meaning of the data. An interactive tool called Application Builder supports this and other activities during the process of building a new application.

⁸ This and the following examples refer to a set of demographic data referring to districts of the City of Bonn. The data were kindly provided by the City Administration.

Statistical districts	Change 31.07.96/ 31.12.95		Percentage		Inhabitants per apartment	
	Total	Male	Female	30/06/1996	31/12/1995	
Zentrum-Rheinviertel	-72	46.6	53.4	1.85	1.90	
Zentrum-Muensterviertel	-112	51.6	48.4	1.98	2.05	
Wichelshof	-71	45.6	54.4	1.87	1.89	
Vor dem Sterntor	-150	51.5	48.5	1.76	1.79	
Rheindorfer Vorstadt	-154	47.5	52.5	2.59	2.66	
Ellerviertel	-116	49.0	51.0	2.01	2.05	
Bonn-Gueterbahnhof	25	52.7	47.3	1.99	1.95	
Baumschulviertel	-142	47.0	53.0	1.91	1.95	
Bonner Talviertel	-161	47.8	52.2	1.87	1.90	
Vor dem Koblenzer Tor	-166	51.2	48.8	1.88	1.96	
Neu-Endenich	-40	48.5	51.5	2.17	2.20	

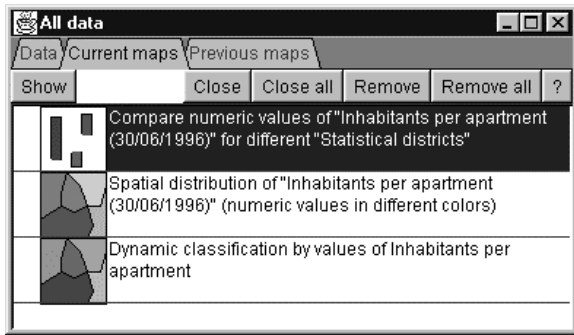
Fig.4. A table with data as it is shown in Descartes.

Having a table on the screen, the user can perform some operations to define the scope of data to be visualized on a map: select table columns; impose a filter on table rows (restrictions on data values); specify an arithmetic formula over the columns and let the system do the corresponding calculations for each record, thus storing the results in a new, dynamically generated column. After the data selection process the user can activate the automated visualization design by pushing the button “Visualize”. In Fig. 4, suppose that the user has selected column “Inhabitants per apartment” by “30/06/1996”. Fig. 5a shows the result as presented to the user. The three icons symbolically represent the three alternative visualization methods that Descartes proposes and applies to build the maps (due to the transformation to grayscale the second and the third icons in Fig. 5a appear to be identical). Note that all names and descriptions of dynamically generated maps and tables are automatically generated by Descartes.

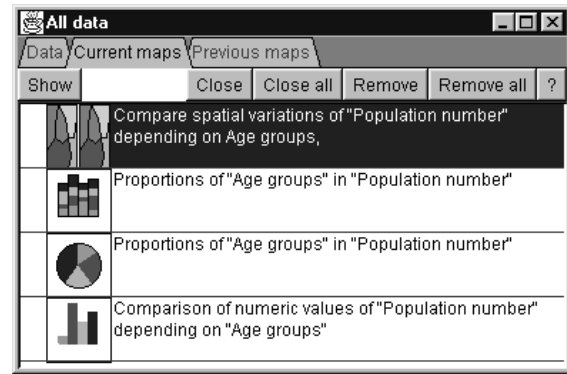
Selection of other attributes will result in different, case-specific proposals of alternative map designs. Fig. 5b presents an example where the user selected several variables, each referring to a different age group of the total population. The variables are thus linked by the relationship “inclusion in a common total”.

To actually view a map, the user selects the corresponding row and presses the “Show” button. Selecting the second item in the list of Fig. 5a will produce the result shown in Fig. 6. Maps in Descartes allow a number of standard GIS operations: zooming and panning, switching on/off geographical layers, and changing colors and sequence of drawing of geographical layers. Descartes also enables interactive access to source data through maps: when the mouse cursor moves over some geographical object, the exact data values associated with this object will be shown in the bottom right frame of the window (cf. Fig. 6).

Map legends provide keys to correct map interpretation and additional information useful for analysis, such as a summary statistics of data variation as shown in Figure 6. The statistics (minimum, maximum, median, and quartiles) are presented graphically by Tukey’s “box and whiskers” plot.



(a)



(b)

Fig.5. (a) A single numeric variable is presented in different ways
(b) several variables related by "inclusion in a common total" are presented

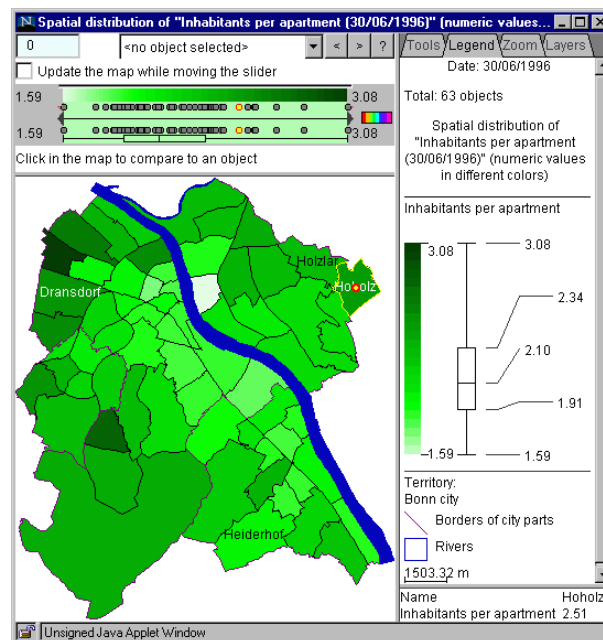


Fig.6. The presentation of a map and its legend in Descartes.

More details about visualization design and many further facilities for interactive manipulation of maps generated by Descartes can be found in (Andrienko & Andrienko, 1998 & 1999a). Here, we only mention one further capability: once a user has found and interactively adapted a specific map of interest this map can be saved for further use. The file of such a map includes a description of the data and the visualization method that was chosen for presenting the data (pie chart, bar chart, or the like). This description can be interpreted by Descartes itself, or by another standalone applet called ShowMap, which can be seen as a functionally reduced variant of Descartes specifically designed for map viewing. In both cases the map is shown to the user in such a way that all interactive operations are enabled: data access, panning and zooming, selection of layers, manipulation of the presentation parameters etc. Unlike the full Descartes system, however, the ShowMap applet can only handle the

given map and data, i.e. no additional attributes or alternative presentation methods can be selected.

Advantages are that due to the reduced functionality operating the ShowMap applet is "easier" to learn, and also its size is reduced, thus loading faster. The presentation of a map saved by Descartes can be embedded into standard HTML pages - just like other graphics. Clicking on the graphics will invoke the ShowMap applet thus making the picture alive. Another significant difference to the full Descartes applet is that the application of ShowMap does not involve any communication with the server, and therefore will not be disabled by firewalls.

4 DESCARTES APPLICATIONS FOR CITIES

4.1 Map analysis: Bonn and Tilburg census data

Fig. 7 compares the absolute numbers of very young citizens (0-3 years) and very old citizens (older than 75 years) in the districts of Bonn. The right frame of the window shows the automatically generated legend. The bottom right frame shows the exact numbers of young and old people for the district under the current mouse position, i.e. Poppelsdorf with values 167 and 374, respectively.

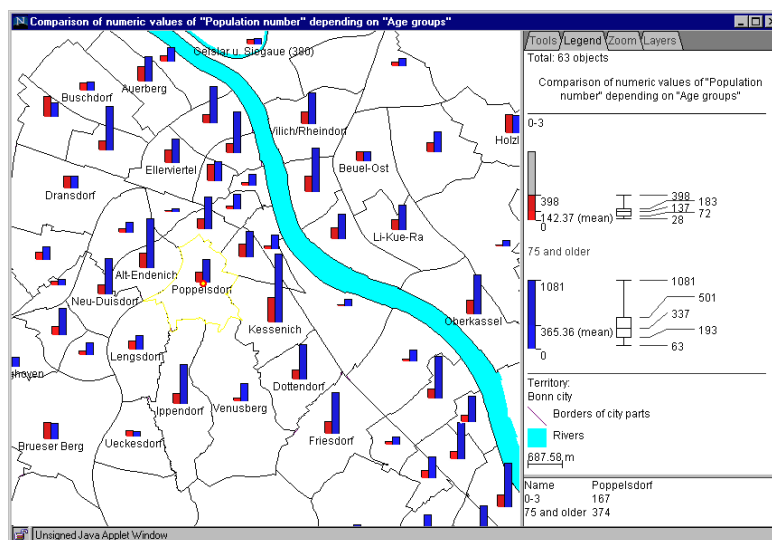


Fig. 7 Analysis of demographic data about city population of Bonn.

Fig. 8 visualizes the population density of the Bonn/Rhein-Sieg/Ahrweiler region. The City of Bonn made this and related maps publicly available (see <http://www.umzug-nach-bonn.de/>). The original motivation was to inform those people about the greater Bonn region who are supposed to change residence from Berlin to Bonn according to the Bonn-Berlin decision of the "Deutsche Bundestag". According to this decision many people working in governmental organizations will move from Bonn to Berlin, but "in exchange" some administrations and organizations will move from Berlin to Bonn. The original data for this application do not directly include population-density as an attribute. Yet this attribute can easily and dynamically be generated from the available source data about area and population of the individual regions. The automatically generated map clearly shows a correlation between population-density and distance from the City of Bonn as the center of the region.

When combined with information about accessibility of the center in terms of minutes by car or public transport, which is also available within the same application, people preferring their home to be in more rural areas could use the map to assist in finding reasonable compromises.

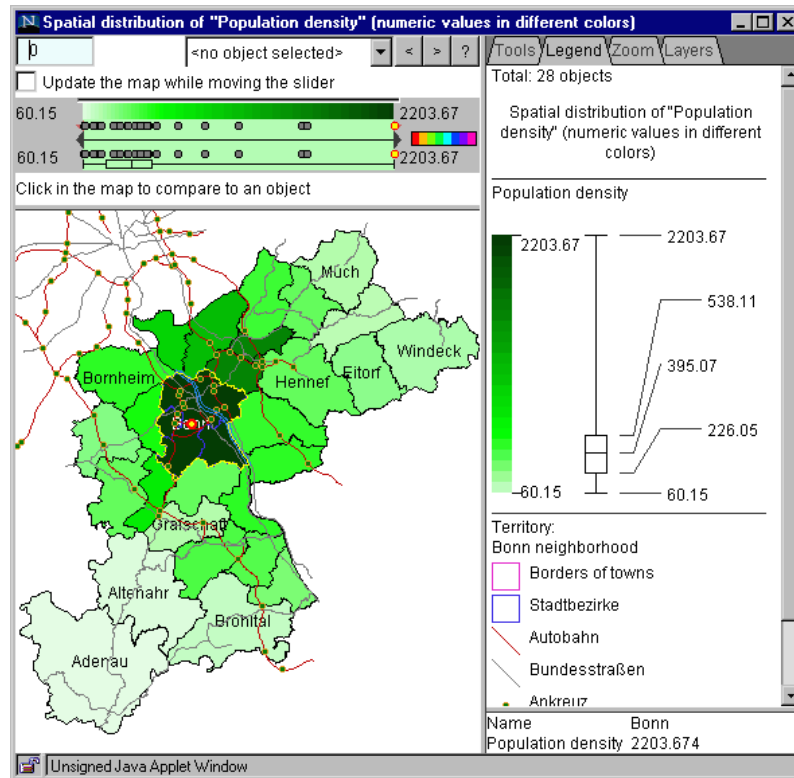


Fig. 8 Population density of the Bonn/Rhein-Sieg/Ahrweiler region.

Fig. 9 provides an example of the cross-classification method of Descartes by indicating relationships between percentages of unemployment and high-education for districts of the City of Tilburg. The left frame of the window shows a scatter-plot with one dot for every district of the city, where increasing percentages of high education and unemployment are shown on x- and y-axes, respectively. Moving the mouse over dots in the plot highlights corresponding districts in the map presentation (middle frame of the window). Vice versa, moving the mouse over districts in the map highlights their position in the scatter-plot. In the original color presentation increasing values along x- and y-axis are shown with darker values of green and red, respectively, and with corresponding mixed colors for inner parts of the plot. The gray-scale figure depicted here cannot adequately convey the potential of this method. In general, cross-classification of two attributes as shown in the figure is a very valuable analytical tool, although in contrast to some other visualization methods it may not be immediately intuitive to people who have not encountered it before. Cross-classification and other methods of similar complexity are therefore primarily intended for professional users, such as planners and decision-makers in different parts of the City Administration.

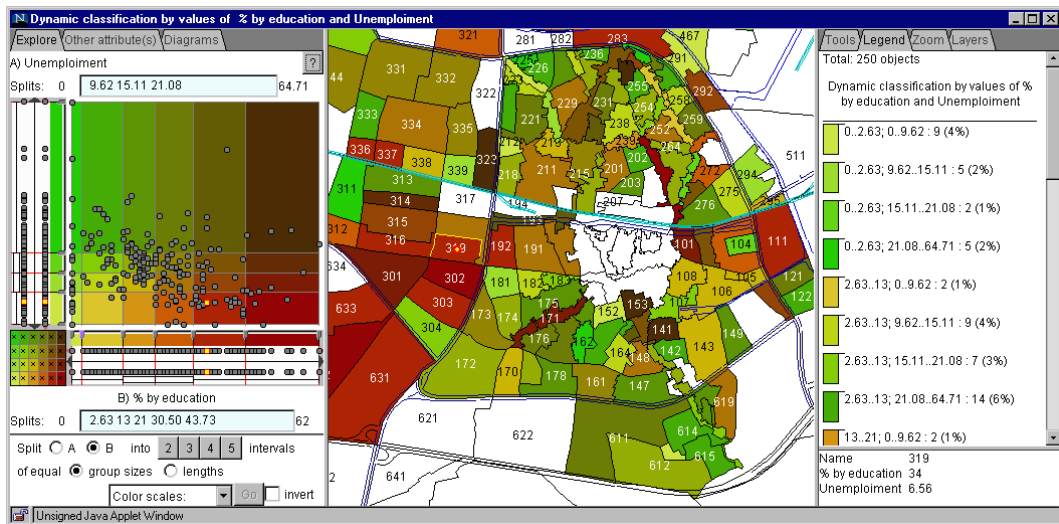


Fig. 9 Cross-classification of percentages of unemployment and high-education for districts of the City of Tilburg.

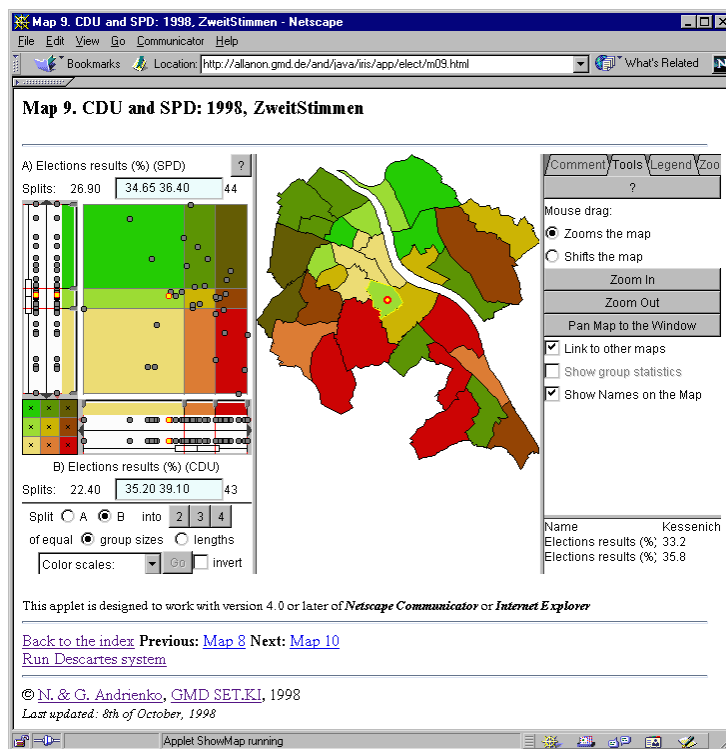


Fig. 10 Cross-classification of percentages of votes for the political parties CDU (x-axis) and SPD (y-axis).

4.2 Map publishing example: Election data

Using the ShowMap applet maps produced with Descartes can be published as interactive images inside of Web pages. The publisher is thus free to design a case-specific graphical and informational context for the interactive maps. Fig. 10 is an element of a series of maps illustrating results of past elections for districts of Bonn (cf.

<http://allanon.gmd.de/and/java/iris/app/elect/indexm.html>). As noted above, the reader of such Web pages can work interactively with the maps but he cannot select other data attributes and he cannot change the presentation method.

The right frame of Fig. 10 shows some zooming and panning features of Descartes, and in the bottom right corner, as usual, shows exact values for the area under the current mouse position.

5 DESCARTES APPLICATIONS FOR UK CENSUS OF POPULATION

5.1 The MIMAS Service and the UK Census of Population

The MIMAS (formerly MIDAS) Service (Manchester InforMation and Associated Services), hosted at the University of Manchester, is a designated national data center serving the UK Higher Education community. One of its key roles is to provide online support for the UK decennial Census of Population which is held in digital form at MIMAS. The main output from the Census is a set of tables of cross-tabulated counts for small geographical areas, known as the Small Area and Local Based Statistics (SAS/LBS) (Cole, 1983). Although only 20 questions were asked in the last Census they generated over 20,000 counts for the area statistics; these are represented as 86 tables in the SAS and 99 tables in the LBS, the main difference being the level of detail available and the geographical level for which the data are aggregated. Thus the SAS has less detail but at a finer geographic resolution than the LBS. These area statistics provide a hugely valuable resource for the study of socio-demographic change in the UK, although the use of the data is subject to licensing restrictions and academics are required to be registered both for the MIMAS Service and for the datasets. There are currently around 1600 registered users of these Census area statistics. In addition to the Census area statistics, digitised boundary data are available under separate licensing restrictions to enable the data to be mapped.

MIMAS is committed to developing its online services to reflect the increasing use of the Web by its academic user base - both for the provision of documentation and more recently, for the delivery of on-line data services (Chruszcz *et al.*, 1998). The KINDS Project⁹ was initiated to increase awareness, accessibility and improved usability of the National Spatial Data Sets held at MIMAS. The first phase of the KINDS Project (1994-1998) delivered a Web-based interface¹⁰ which enabled users to search for and identify spatial datasets of interest, to map the spatial data online and to extract and download the data. Now in its second phase, KINDS II is focusing on the construction of a resource gateway which will bring together spatial data plus derived datasets and teaching and research materials which have been produced by users, thus recycling the knowledge and experience gained within the user community. In another ongoing MIMAS project, a Web-based interface to the Census area statistics, CASWEB¹¹, was launched in September 1998. This enables authorised users to select subsets of the Small Area and Local Based Statistics for the 1991 census datasets and to download these census counts as ASCII files for use locally (Martin *et al.*, 1998).

Whilst the CASWEB interface helps to overcome many of the barriers that were previously encountered by new and inexperienced users of census data, it is likely that users will want to combine the census counts with census boundaries to explore the data in their geographical

⁹ KINDS (Knowledge-based Interfaces to National Data Sets) is a consortium of Manchester Metropolitan University, Manchester Computing and the University of Salford, working with MIMAS.

¹⁰ The KINDS web site is at <http://www.kinds.ac.uk/>

¹¹ CASWEB can be found at <http://census.ac.uk/casweb>

context. Thus CASWEB only partially meets the need for users to be able to engage in full exploratory spatial data analysis online via a Web browser. A separate project was initiated at MIMAS in order to achieve this aim.

5.2 The JANUS Visualization Gateway Project

The JANUS Visualization Gateway Project was conceived in order to create an intelligent, online visualization gateway to the spatial datasets, especially the census and associated digitised boundary datasets, held at MIMAS. The main aims of the JANUS Project have been to firstly create this visualization gateway, through the KINDS interface, and secondly to promote visualization ideas and software tools, such as the CDV (Cartographic Data Visualizer) software, within the UK academic community (Carter and Dykes, 1998; Carter and Bullen, 1999).

The work has progressed in three stages: firstly the CDV software was ported to the MIMAS server and adapted for use as a multi-user, X-windows application¹² together with associated Web pages to educate users about exploratory spatial data analysis techniques; secondly an education programme within the academic community was developed to promote the use of these products; and thirdly software that could enable users to perform exploratory data analysis of census data over the Internet was sought. As noted earlier, whilst it is possible to make applications written in Tcl/Tk (the language used to produce CDV) run over the Internet this has two main drawbacks making it unsuitable for a visualization gateway to copyrighted Census statistics: a plug-in is required and data security cannot be guaranteed. The approach using Descartes' client-server architecture offered the environment required for a Web-based solution. A link was established between MIMAS and GMD which led to the JANUS Project being granted permission to use Descartes to develop the required visualization gateway, and the Descartes software being installed on the MIMAS Unix Server.

As Descartes requires previously prepared applications, following the lead of datasets that had been prepared for use, by registered census users, with CDV, a set of applications was generated to enable users to visually explore a subset of census data for an area of interest to them. Great Britain comprises 55 counties (England and Wales) and 12 regions (Scotland). Within these administrative divisions, the census area statistics are available for further geographic subdivisions; districts, wards and enumeration districts (EDs). The census geography chosen to create the applications was census ward level (for England and Wales) and postcode-sector level (for Scotland) since this level enables geographical patterns to be interpreted at a regional level as well as allowing sufficient detail to be perceived at a local level. Once the applications were created the only additional step was to ensure that they were password protected, in order to comply with the copyright restrictions imposed on the data.

A user wishing to use Descartes to explore an area of interest needs to follow a simple set of steps:

1. They can access the Visualization Gateway through the KINDS interface at <http://www.kinds.ac.uk/> and follow the link to Descartes. This will take them to a Web page giving details about the Visualization Gateway and the use of Descartes for visual exploration. Alternatively they can directly enter the URL <http://www.mimas.ac.uk/descartes/>

¹² CDV can also be downloaded for individual use on a PC

2. If the user is not registered to use the Census area statistics and the associated boundary data they can still use Descartes with a demonstrator dataset for the county of Leicestershire, UK. Permissions were granted under the Argus Project¹³ by the data providers for this data to be freely available for use within the UK academic community.
3. If the user is registered to use the 1991 Census area statistics and the associated digitised boundary data they can now interactively request, from a drop down menu, a county/region of their choice.
4. They will be asked to authenticate themselves by entering their username and password. This step checks that (i) they are registered to use the MIMAS Service and (ii) they are registered to use both the area statistics and the census boundary data.
5. Providing a user is authorized they can now select a pre-prepared table for visual exploration in Descartes.

The table selected represents an application in which Census counts from the 1991 Local Base Statistics have been combined with boundary data. The data available represent some key variables including: number of households, car ownership, overcrowding, limiting long-term illness, gender, ethnic origin, place of birth (UK and Irish Republic) and employment characteristics. Figure 11 provides an example of two linked choropleth maps produced from a pre-prepared Census table for the county of Greater Manchester. The 'Calculate' function in Descartes was used to produce, from the available data, the columns Percent aged 0-24 and Percent aged 60+ in order to visually explore the geographic distribution of young and older people within Greater Manchester's population.

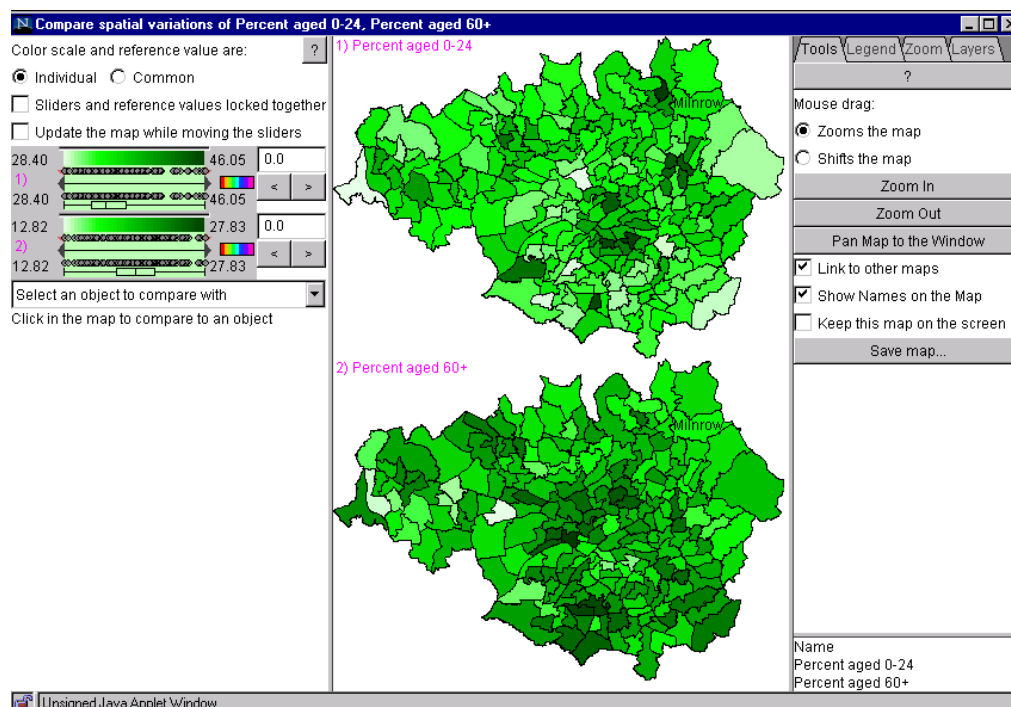


Fig. 11 *Linked choropleth maps of two variables showing Percent in Age ranges 0-24 and 60+ in Greater Manchester, UK*

Census Data © Crown Copyright
Boundary Data: Copyright ED-Line Consortium

¹³ <http://www.mimas.ac.uk/argus/>

It is hoped that the provision of these pre-prepared applications through the Visualization Gateway at MIMAS will further help in overcoming the barriers met by new and inexperienced users when attempting to explore census data. The primary aim is to make the area counts and associated boundary data easier to access and use in teaching and research.

6 CONCLUSIONS AND FURTHER WORK

Thematic mapping services of Descartes have been successfully applied for dissemination of statistical information on city and regional levels in several countries. These applications proved that client-server architecture provides a reasonable compromise between amount of data to be downloaded and functionality available to end-user. Rapid advance in Java technology (in fact, Java applets now run much faster than two years ago) makes it possible to provide Internet users with wide opportunities for data access and analysis.

There are several directions of our further work:

1. Within the EU-funded CommonGIS project (ESPRIT project 28983¹⁴) we plan to add more intelligence to the system. Thus, a rather long sequence of operations is often needed to proceed from source data to a useful presentation. For example, it may be necessary to transform absolute values to percentages, calculate differences or ratios, filter database records, etc. We intend to "wrap" such operation sequences into analysis scripts. The scripts are presented to the user as various analytical tasks formulated in terms of analyzed data and domain notions. These scripts will, first, simplify the acquaintance with the system and release the users from memorizing its capabilities and, second, save time and effort for even experienced users.
2. Exploratory analysis of spatially-referenced data could be combined with the display of complex, multiple-layer vector and raster maps. Current Descartes supports this, but its capabilities are limited by the speed of transfer through the network. Actually, it is impossible to deal with geographic files bigger than 1Mb. To improve this, we started integration with commercial GIS software, Lava/Magma¹⁵, developed by Dutch company Professional Geo Systems (see Andrienko, Andrienko, Voss, & Tuijnman, 1999). This system manages huge volumes of geodata due to sophisticated caching.
3. Many users already have their data loaded into standard GIS software. Quite often they do not want to change the environment they get used to. To support such categories of users, a spin-off company of GMD, Dialogis GmbH¹⁶, is developing an interface module for Descartes from ArcView GIS (Rinner, Schmidt, Andrienko, & Voss, 1999). With this module a user may continue keeping his data in ArcView but solve exploratory tasks and publish interactive maps using Descartes. Transfer of data and activation of Descartes is done automatically.
4. Visual analysis of spatial information could be supported by computational methods of data mining. We started a project for integrating Descartes with general data mining system Kepler (also developed by Dialogis GmbH). In the pilot implementation a user could activate data mining tools (currently feature weighting and classification learning methods) directly from map interface. Graphic displays of both systems are like the maps inside Descartes. This possibility creates good conditions for data exploration (Andrienko & Andrienko, 1999b).

¹⁴ URL <http://commongis.jrc.it/>

¹⁵ URL <http://www.pgs.nl/>

¹⁶ URL <http://www.dialogis.de/>

REFERENCES

1. Andrienko, G. and Andrienko, N. (1998). Intelligent Visualization and Dynamic Manipulation: Two Complementary Instruments to Support Data Exploration with GIS. In *Proceedings Advanced Visual Interfaces AVI'98* (pp.66-75). New York: ACM Press.
2. Andrienko, G. and Andrienko, N. (1999a). Interactive Maps for Visual Data Exploration. *International Journal Geographic Information Science*, 13, pp.355-374.
3. Andrienko, G. and Andrienko, N. (1999b). Knowledge-based visualization to support spatial data mining. In *Proceedings Intelligent Data Analysis IDA'99, Lecture Notes in Computer Science* (to appear). Berlin: Springer.
4. Andrienko, G., Andrienko, N., Voss, H., and Tuijnman, F. (1999). Architecture of the CommonGIS system. In *Proceedings of 5th EC GIS Workshop* (to appear).
5. Carter, J., and Dykes, J. (1998). Developing a Visualization Gateway for the UK Census of Population. In T.Poiker & N.Chrisman, In *Proceeding Spatial Data Handling SDH'98*, (pp.543-555), International Geographic Union.
6. Carter, J. and Bullen, N. (1999). Overcoming barriers to the use of the Census through interactive visualization. Presented at IASSIST Conference 17-21 May 1999, Toronto, Canada. To appear in *IASSIST Quarterly*.
7. Chruszcz, J., Cole, K. and McCombe, A. (1998) The role of the Web in the provision of national data and information services: the MIDAS experience, Paper presented at IASSIST/CSS Conference, New Haven, USA, May 1998.
8. Cole, K. (1993) The 1991 local base and small area statistics, in Dale, A. and Marsh, C. (eds) *The 1991 Census user's guide*. London: H.M.S.O., 201-47.
9. Dykes, J.A. (1997). Exploring Spatial Data Representation with Dynamic Graphics. *Computers & Geosciences*, 23, 345-370.
10. Dykes, J.A. (1998). Cartographic Visualization: Exploratory Spatial Data Analysis with Local Indicators of Spatial Association using Tcl/Tk and cdv. *The Statistician*, 47, 485-497.
11. Grossley, D. and Boston, T. (1995). A Generic Map Interface to Query Geographical Information using the World Wide Web. *World Wide Web Journal*, 1996, 723-735.
12. Martin, D., Harris, J., Sadler, J. and Tate, N. J. (1998) Putting the census on the web: lessons from two case studies. *Area*. 30(4): 311-320.
13. Peterson, M.P. (1999). Active Legends for Interactive Cartographic Animation. *International Journal Geographic Information Science*, 13, 375-383.
14. Plewe, B. (1997). *GIS Online. Information Retrieval, Mapping, and the Internet*. Santa Fe: OnWorld Press.
15. Rinner, C., Schmidt, D., Andrienko, G., and Voss, H (1999). dMaps for Internet Publishing of Statistical Maps, *submitted to ETK'99 meeting*.
16. Sorokine, A. and Merzliakova, I., 1998, Interactive Map Applet for Illustrative Purposes. In *Proceedings of 6th International Symposium on Advances in Geographic Information Systems* (pp.46-51), New York: ACM Press.

COORDINATES

Drs. Gennady and Natalia Andrienko*, Dr. Hans Voss**

GMD - German National Research Center for Information Technology

AiS.KD, Knowledge Discovery Group

Schloss Birlinghoven,

D-53754 Sankt-Augustin,

Germany

Tel +49-2241-14-2329*, - 2532**

Fax +49-2241-142889

e-mail: {Gennady.Andrienko, Natalia.Andrienko, Hans.Voss}@gmd.de

<http://ais.gmd.de/KD/>

Dr. Jackie Carter

MIMAS (Manchester InforMation and Associated Services) (formerly MIDAS)

Manchester Computing

University of Manchester

Manchester M13 9PL

UK

Tel +44 (0)161 275 6725

Fax +44 (0)161 275 6040

Email: j.carter@man.ac.uk

<http://www.mimas.ac.uk/janus/>

<http://www.kinds.ac.uk/>