

GIS for Everyone: the CommonGIS project and beyond

Gennady Andrienko, Natalia Andrienko, and Hans Voss

Fraunhofer AIS - Fraunhofer Institut Autonome Intelligente Systeme

Schloss Birlinghoven, Sankt-Augustin, D-53754 Germany

Tel: +49-2241-142329, -2532

Fax: +49-2241-142072

E-mail: {Gennady.Andrienko, Natalia.Andrienko, Hans.Voss}@ais.fraunhofer.de

URL: <http://www.ais.fraunhofer.de/SPADE/>

Abstract. The project CommonGIS had the motto “GIS for everyone”. Its objective was to devise approaches and methods for making modern techniques for map-based exploratory data analysis accessible to a broad community of potential users. Besides technical accessibility, which was attained by applying the Java applet technology, it was important to ensure that the techniques could be utilized appropriately and effectively even by users with no or little expertise in cartography or geographical information systems. This task was especially challenging because of the novelty of the exploratory tools. Our experiments showed that, without previous instruction, people only made limited use of these tools. Therefore, it was necessary to find proper ways of providing intelligent software-based assistance that would encourage and help people to use the tools.

In the beginning, a knowledge-based system was designed, which automated the process of selecting appropriate visualization methods depending on data characteristics in accord with the principles of cartography and graphics design. While this proved valuable, the need in taking into account particular goals of analysis was realized. Moreover, it was necessary to assist in the usage of the selected instruments in accord with the goals. The chapter will show how the knowledge base was extended for this purpose.

1 Introduction

The CommonGIS project began in November 1998 and was funded by the European Commission (Esprit project N 28983). The work was completed in June 2001. The primary goal was to devise and elaborate approaches and methods for making modern map-based facilities for exploratory data analysis accessible to and usable by a broad community of potential users.

Within the project, a software system, called CommonGIS, was developed. The system runs over the Internet by means of a standard web browser. It can be viewed both as an interactive Web-GIS because it provides some standard GIS functionality, and as a tool for exploratory analysis of geographically referenced statistical data. To support the analysis, CommonGIS provides highly interactive and dynamically transformable maps and other types of data displays, in accord with the concept of “geographic visualization” (MacEachren 1994, MacEachren and Kraak 1997). The interactivity of the data displays on a web client is achieved by Java applet technology.

While it is generally believed that map interactivity greatly supports data exploration, it was found that the geovisualization facilities included in CommonGIS were too unusual for users to be understood and utilized without previous introduction (we are convinced that this is not specific to CommonGIS but would apply to any novel tool or paradigm). Since in the situation of accessing tools over the web no direct teaching is possible, alternative methods were sought for acquainting users with novel techniques and instruction in how to utilize them. Thus an idea of intelligent (knowledge-based) user guidance arose.

This chapter describes how this idea was realized in CommonGIS. Prior to that, the existing literature on knowledge-based user support in data visualization is briefly reviewed.

2 Knowledge-based user support in data visualization

In order to visualize data properly, one needs to obey the established principles of graphical and, for spatial data, cartographic representation (see, for example, Bertin 1983). The survey of Murray (1994) describes a number of software systems that incorporate these principles in order to help users in graphical presentation of information. Mackinlay (1986) was the first who showed the feasibility of automated knowledge-based graphics design. His system, APT, could select visual primitives, such as position, size, color, or shape, depending on characteristics of data components, and appropriately combine them. Later systems such as SAGE (Roth and Mattis 1990) and VISTA (Senay and Ignatius 1994) applied and further advanced the ideas introduced by Mackinlay. In SAGE the graphics design was sensitive to the user's information seeking goals expressed as generic tasks of the kind "accurate lookup of separate data values", "comparison of values of two attributes", "study of the distribution of values", or "revealing of functional correlation among attributes".

Casner (1991) considers more precisely specified tasks constructed from primitive logical operators such as a query for a property of an object. However, the perceptual operators considered by Casner refer only to the elementary level of reading (Bertin 1967/1983). Therefore the resulting graphics can only serve as good data inventories rather than support exploratory analysis.

While all the above-mentioned authors dealt with non-spatial data, Zhan and Battenfield (1995) describe an expert system that recommends appropriate methods for cartographic representation. The VIZARD system (Jung 1995) not only selects suitable methods but also automatically generates thematic maps. Like SAGE, VIZARD attempts to suit the presentations to user's goals expressed as generic tasks: "read values", "locate values on the map", "compare values", or "detect spatial trend".

A common feature of all aforementioned systems for automated design of graphical data displays is that they generate static, non-interactive images that can be considered as fully analogous to graphics and maps on paper. Our knowledge-based system, Descartes (Andrienko and Andrienko 1999), which existed before the beginning of the CommonGIS project, could automatically produce highly interactive, transformable maps suitable for visual exploration of spatial data. Cartographic representation methods were selected depending on characteristics of data, but the system did not account for user's analytical goals.

In CommonGIS, it was initially planned that Descartes would be made sensitive to user's goals by means of reusing the knowledge from VIZARD. However, in the course of the project this idea greatly evolved.

3 User's goals in visualization design and data analysis

3.1 Initial vision and its evolution

The knowledge base of VIZARD consisted of two parts: "expressiveness knowledge" that accounted for data characteristics and "effectiveness knowledge" that specified for each cartographic representation method its degree of suitability for different task types. The effectiveness degrees were determined by experiments; the results are summarized in Table 1 (translated from Jung 1998).

Table 1. Effectiveness of cartographic representation methods for different tasks

	Read values	Locate values	Compare values	Detect trend
Bar charts	certainly good	good	good	bad
Graduated circles	average	certainly bad	good	bad
Choropleth map	bad	slightly good	certainly bad	good
Dot map	average	slightly good	terrible	average

Here the verbal formulations of the effectiveness values stand for fuzzy numeric coefficients actually specified in the knowledge base. VIZARD used fuzzy logic and fuzzy arithmetic for determining the effectiveness of displays combining several elementary representation methods (Jung 1996).

While on a first view it seemed quite attractive to just reuse the effectiveness knowledge of VIZARD and also its logical inference mechanism, it was later realized that the set of general tasks considered in VIZARD was too abstract and limited. One may distinguish, for example, a variety of types of comparison tasks, rather than just a single "compare values". Thus, the task "compare birth rates in Germany and France" is essentially different from the task "estimate the position of Germany among the countries of Europe with respect to birth rate" (or, in other words, compare Germany with all other countries). The two tasks need to be fulfilled in different ways, and a representation method such as bar charts may be good for one of them and bad for the other. Therefore it was necessary not simply to reuse existing knowledge but to elaborate it.

However, even this vision proved to be too simplistic. It eventually turned out that reusing the effectiveness knowledge of VIZARD was not a valuable approach because this knowledge was only valid for static cartographic images. It may be observed that the interactivity and transformability of computer map displays dramatically changes the characteristics of presentation methods in relation to the goals. Thus, in CommonGIS the task of reading values becomes rather trivial, irrespective of the representation method used: the exact values are immediately

shown when moving the mouse cursor over a geographical object. Hence, there is no reason to prefer bar charts when the user needs to read values. Another example is the interactive technique of visual comparison (Andrienko and Andrienko 1999) that radically changes the properties of traditional choropleth maps. In particular, it makes them well suited for the task “compare values”, contrary to the knowledge represented in VIZARD.

Consequently, a need arose for a considerable update of the effectiveness knowledge base taking into account the new properties of interactive displays in comparison to their static counterparts. It soon became clear that it would be inappropriate to simply change the values of the effectiveness coefficients. There are several reasons for this. First, some interactive techniques, such as the interactive access to exact values, can be applied irrespectively of the representation method used. The knowledge base would need some extensions to correctly handle such techniques. Second, and more serious, VIZARD requires the selection of a single task type, and it then optimizes the design according to this selected type. In a real data analysis process one rarely addresses one task at a time. Tasks of different types intermingle, and it would be very inconvenient for the user to get a separate display for tasks of each type, especially when the overall set of task types would be much more extended, as it was planned. Third, it is generally difficult and unnatural for the user to decompose analytical activities into primitive operations such as “read values”, “locate values”, and “compare values”. The user would certainly prefer to see goals formulated in problem- and domain-oriented terms, for example, “to investigate variation of population structure throughout Europe”. The conclusion was that a completely new knowledge base had to be developed instead of a modification of the existing one.

Another issue that arose during the project was that the tools intended for a wide community of users were not just static maps but novel and interactive map analysis tools. Communication with potential users and usability tests revealed that many people had difficulties in understanding the new tools. The results of the usability tests are briefly reported below.

3.2 Tests of usability of interactive maps in CommonGIS

Software development within the project was done in two stages. On the first stage an early prototype was created. It was then tested for possible flaws regarding its general design and the user interface. In the second stage the final version of the system was developed taking into account the results of prototype testing.

During the tests of the prototype the subjects realized that CommonGIS offered something completely new and unusual compared to traditional GIS. The general opinion was that only advanced GIS users would be able and willing to use these novel features. Therefore the test participants suggested separating the interactive map manipulation tools from the traditional GIS functions. Moreover, they would prefer these tools to be initially hidden and only appear on demand.

General conclusions that could be drawn from the prototype tests are the following:

- Users cannot grasp the existence and meaning of new interactive tools just from their appearance without previous introduction, and, consequently, cannot guess how to use them.

- Users feel uncomfortable when encountering unfamiliar features with unclear purpose. They would rather prefer not to find such features in the system.
- Users are accustomed to static maps and tend to adhere to traditional mapping principles. They may have no idea that a map can be interactively manipulated to facilitate data analysis.

Since it became clear that users needed some background in how to utilize the exploratory tools, it was reasonable to check how much teaching was required and in what form. For this purpose another series of tests was conducted with a specific focus on the use of the interactive facilities. A detailed description of the tests can be found in (Andrienko et al. 2002); here a brief summary is given.

The tests of the exploratory tools took place close to the end of the project. At that time the final version of the system was mostly ready, and it was used for the experiments. Instead of the free exploration of the system that took place in prototype evaluation, the users were given a set of specific tasks requiring application of the tools. This allowed us to assess tool learnability using a numeric measure, error rate, that is, the percentage of wrong answers to the total number of answers. Besides fulfilling the tasks, the subjects were asked to complete questionnaires in which they could express their subjective satisfaction and opinions concerning particular techniques and the system in general: understandability, ease or difficulty of use, attractiveness, etc.

The tests were done in three rounds. The first two rounds were performed on-site using a local installation of the system while the third round was conducted over the Internet. Participants of the first round (9 persons) were given a lecture demonstrating the interactive facilities of CommonGIS that they needed to apply in the test. The presentation took about 30 minutes. The subjects did not practice using the tools by themselves before starting the test. The second round took place one month later with participation of the same users as in the first round. The purpose was to check tool memorability, that is, whether the users could maintain the acquired knowledge and skills about the interactive techniques and re-apply them after a recess. The third round (Internet-based tests) involved more than 200 subjects who did not get any prior introduction and could rely only on written explanations and illustrated examples. The purpose was to test whether people accessing CommonGIS on the web were able to master its interactive exploratory facilities without being directly taught.

The results of the first two rounds were rather encouraging. In the first round, the subjects were quite able to use the interactive tools and demonstrated solid performance in fulfilling the requested tasks. The results of the second round in terms of fulfilling the tasks were even better than those of the first round. This means that people, once having understood the new tools and trained to use them, could preserve and reuse the knowledge and skills acquired. Moreover, the subjective satisfaction data showed that the subjects liked the new tools, and many very favorable comments were received. Some expressed their wish to use the system in their professional activities.

In the Internet-based tests, however, the overall performance of the subjects was much worse than in the first two rounds. A probable conclusion is that the written explanations did not adequately substitute for the introductory lecture. It was detected that many subjects did not read the explanations at all or viewed them for just a few seconds (the testing software recorded the times when an explanation page was

opened, if at all, and when the next action took place). It should be noted that ignoring written explanations is quite a typical behavior for casual users.

However, it would be wrong to say that the reluctance of the subjects to read the instructions was the only reason for the poor overall performance. Many of the test participants spent enough time for viewing the explanations but still made a lot of errors in fulfilling the tasks. At the same time, some subjects who did not read any explanations performed poorly at the beginning of the test but greatly improved by the end. This demonstrates that users may be psychologically rather diverse, and there are people who prefer and are able to master new tools independently using the “trial and error” method.

The main overall conclusion from the usability studies is that users are, in principle, able to understand and adopt the new ideas concerning map interactivity and manipulability. However, these ideas need to be appropriately introduced; people can hardly grasp them just from the appearance of the maps and controls. While a proper design of the user interface is extremely important, it is insufficient just to apply the established human interface design principles (see, for example, Macintosh 1995). The central idea of these principles is to base upon people’s previous knowledge of the world and experiences with other software and to keep consistency with people’s expectations. In a case of such novel concepts as interactive maps and graphs there are often no analogous situations in typical people’s experience that could be utilized as metaphors. It may be also impossible to implement unconventional features using only standard user interface elements with standard behaviors. Consistency with user’s expectations is also hard to keep: the very idea of a map dynamically changing its appearance conflicts with typical expectations since most people are mainly accustomed to static paper maps. Hence, most users will not be able to master the novel tools of CommonGIS just relying on their previous experiences, and teaching becomes indispensable.

As it was seen from the first two rounds, an introductory demonstration may be sufficient for understanding the purposes of interactive tools, and short training enables people to use the tools. In situations when no direct teaching is possible, for example, in the case of first time users encountering the tools on the web, adequate substitutes need to be provided. It was seen that on-line help is not always effective. Training users by means of interactive on-line tutorials may be more appropriate but still not ideal. Many people will not invest time and effort for such training without being sufficiently convinced of the considerable benefits of the new tools.

Ideally, the new techniques should be introduced to users when they perform their own analysis, that is, each technique should be introduced exactly at the moment when it would be useful. This means that an intelligent guide is needed, which understands the users’ goals, proposes the appropriate instruments, and helps in using them.

4 Intelligent user guidance in CommonGIS

4.1 Requirements to the guidance

The original idea of knowledge based map design that takes into account user’s goals was transformed into the idea of providing intelligent guidance throughout the whole

process of data analysis. The guiding system, or guide for short, needed to satisfy the following requirements: the guide should

- propose relevant analysis instruments (such as maps, non-cartographic displays, and interactive techniques to manipulate them) at the moments when they become useful. For this purpose the guide needs to recognize the current goal of the user, or, in other words, the data analysis task to be fulfilled;
- not request the user to explicitly decompose goals to the level of primitive operations. Instead, the guide should communicate with the user about the goals in terms of high-level analysis tasks.
- not propose a separate display for each primitive operation. Instead, the guide should provide for each high-level task a complex of tools that fit together, can be used in parallel and in arbitrary order, and collectively support all the primitive operations involved.
- allow the user to decide whether to use any of the proposed tools or not. The decision should be supported by information about the instruments: how they work and for what purpose they can be used.
- supply the user with explanations necessary for understanding how to use the instruments in the context of each task (one and the same instrument can be useful for multiple tasks but be used differently in each case).

Within the project a prototype of the guide was implemented as a new type of on-line, task- and situation-specific help system.

4.2 Data analysis scenarios

One of the key problems was the recognition of the user's current goal. In seeking a solution to this problem the idea appeared of generic data analysis scenarios, that is, very high-level tasks expressed in general terms such as "investigate statistical data referring to territory division" or "select an appropriate place (or places) according to multiple criteria". It was hypothesized that there exists a finite and rather limited number of such high-level geography-related tasks, or that at least a constrained number of tasks would be sufficient to cover most needs of the potential CommonGIS user community. Each scenario is composed of smaller tasks that, in their turn, may consist of subtasks, and so on recursively. Often individual tasks inside one scenario must be performed in a given sequence, but it is also possible that the order of some tasks is irrelevant.

As an example, consider the task hierarchy corresponding to the scenario "select an appropriate place (or places) according to multiple criteria". This scenario may stand for various specific selection tasks such as "select a house to buy", "select a place to go on vacation", or "select a location for a new waste disposal factory".

1. Explore the opportunities
 - 1.1. Survey the territory
 - 1.2. Define the options to select from

- 1.3. Investigate spatial distribution of the options
 - 1.3.1. Observe spatial distribution of the options, detect patterns
 - 1.3.2. Explain observed patterns of distribution of the options
 - ...
- 1.4. Investigate variation of thematic characteristics of the options
 - 1.4.1. Observe spatial distribution of values of attributes, detect patterns and trends
 - 1.4.2. Explain observed peculiarities of distribution of attribute values
 - ...
 - 1.4.3. Reveal correlations among the attributes characterizing the options
- 1.5. Investigate variation of neighborhood of the options
 - 1.5.1. Investigate variety of objects around the options, detect presence or absence of objects of a particular kind
 - 1.5.2. Investigate variation of thematic characteristics of objects in neighborhood
2. Select suitable options
 - 2.1. Exclude inappropriate options
 - 2.1.1. Filter options on the basis of their thematic characteristics
 - 2.1.2. Filter options on the basis of the neighborhood
 - ...
 - 2.2. Rank or classify the appropriate options according to the degree of preference
 - 2.2.1. Define criteria
 - ...
 - 2.2.2. Rank or classify the options according to the criteria

Note that in any particular case not all of these tasks may be applicable. For example, in the selection of a house the options will usually be already defined: these are the houses proposed for sale. Therefore the task 1.2 “Define the options to select from” is in this case irrelevant, while in the task “select a place to go on vacation” it may be significant. The task 1.3 “Investigate spatial distribution of the options” makes sense only when the options are rather numerous and spread over a wide territory.

The generic scenarios are formally described in a knowledge base. The guide may thus look up the knowledge base and find out which scenarios actually exist and what smaller tasks they consist of. For each terminal task (that is, a task without descendants in the hierarchy) the knowledge base specifies appropriate instruments of analysis. When a user wishes to utilize this support system, the guide will propose to select the scenario that matches most of the user’s goals. Then the user may follow the given hierarchy of tasks. When a terminal task is reached, the guide proposes the corresponding analytical tools together with instructions on how to apply each of them in the context of this task. The user does not have to attend to every task

described for the scenario but may skip any one that is deemed irrelevant. On the other hand, the proposed list of potentially relevant tasks will serve as a reminder that helps not to overlook important steps of analysis, or encourages considering steps that were originally not deemed relevant.

Whenever possible, the user may decide in what order to fulfill the tasks. However, some tasks may use results of other tasks and, hence, require the latter to be done earlier. For example, the task 2.2.2 “Rank or classify the options according to the criteria” uses the results of the task 2.2.1 “Define criteria”. In such cases the guide should ensure that the user adheres to the required order.

Note that the tasks are not detailed to a level of primitive operations such as “read values”, “locate values”, or “compare values” as this would be inconvenient for users. Instead, the knowledge base enumerates the available analytical instruments and specifies for each of them which basic operations are covered by its use. For example, the interactive technique “dynamic query” supports such operations as “locate values”, “filter objects on the basis of attribute values”, and “detect patterns in spatial distribution of values”. For each operation there is an instruction addressed to the user about how the operation can be done with the respective instrument. Further on, for each terminal task of a scenario it is specified which of the instruments are appropriate for it and which of the basic operations supported by these instruments are relevant in the context of this task. Thus, the dynamic query can be used for the tasks 1.4.1 “Observe spatial distribution of values of attributes” and 2.1.1 “Filter options on the basis of their thematic characteristics”. The relevant operation in the former case is “detect patterns in spatial distribution of values” and in the latter “filter objects on the basis of attribute values”.

When a user chooses a terminal task, the guide proposes the appropriate instruments, according to the specification in the knowledge base. For each instrument a brief description is available. Only when the user activates one of the instruments, the guide shows the instructions concerning the relevant basic operations. Thus, when the user employs the dynamic query tool in the context of task 1.4.1, the guide will instruct the user how to apply it for detecting patterns in spatial distribution of values. When the tool is activated in the context of task 2.1.1, the guide explains how to use it for filtering objects on the basis of attribute values.

4.3 Knowledge base for task support

The knowledge base for task support is defined in the XML language, and consists of two major parts: definition of scenarios and description of available instruments. Each scenario is a hierarchy of general tasks. Currently, the knowledge base includes only two generic scenarios: “investigate demographic data referring to territory division” and “select an appropriate place (or places) according to multiple criteria”. The instruments include:

- cartographic representation methods such as unclassified and classified choropleth maps or maps with bar charts;
- map manipulation techniques associated with particular representation methods (for example, visual comparison is associated with unclassified choropleth maps);

- other interactive techniques not restricted to a particular map type, such as the dynamic query tool or the tool for accessing exact attribute values;
- non-cartographic displays such as a scatterplot or a plot of parallel coordinates;
- analysis tools based on data transformations and calculations, for example, a tool for automated ranking of options on the basis of multiple criteria implementing one of algorithms of multi-criteria decision support.

As already mentioned, for each terminal task the appropriate instruments are specified. Also, for any task the knowledge base may specify requirements regarding the data that need to be met so that this task becomes relevant. For example, the task 1.4 “Investigate variation of thematic characteristics of the options” makes sense only if there are some thematic characteristics available. Such knowledge allows the guide to display only those tasks that can be executed with the available data.

Each task and each instrument has a name and a brief explanation of its meaning or purpose. Although currently only English versions of these texts are provided, the design and implementation of the knowledge base foresees that text for the interaction with users can be stored and displayed in different languages.

When appropriate, inputs and outputs of an instrument are described. For example, the dynamic query tool needs as input a set of objects with attributes that are used for the specification of query constraints. The output of this tool is a subset of the initial set of objects. There are also instruments that do not produce any output, for example, the tool for visual comparison or for viewing attribute values.

Tasks with inputs and/or outputs may be linked to other tasks through these inputs and/or outputs. For example, the output of task 2.1 “Exclude inappropriate options” is a subset of the initial set of options. It is used as an input to task 2.2 “Rank or classify the appropriate options according to the degree of preference”. However, the task 2.1 is optional, and a user may skip it. The knowledge base specifies which alternative input shall be taken for task 2.2 in this case, namely the whole set of options. Similarly, task 1.2 “Define the options to select from” provides an input to all the following tasks. The knowledge base may state that, if task 1.2 is skipped, the user needs to specify which of the available map layers contains the options to select from.

Outputs of tasks result from the application of the corresponding instruments, therefore the knowledge base links outputs of tasks to outputs of the instruments. When several instruments having outputs need to be associated with a task, this can be done according to one of the following schemes:

1. The instruments are declared as alternative. Then, if the user applies one of them after another, the results of the earlier applied instrument will be discarded, and only the results of the last applied instrument are taken as the output of the task.
2. The instruments are applied in a prescribed order. The output of each instrument is taken as an input for the next instrument in the sequence. The output of the last instrument is the result of the task.

- Each instrument can be applied independently of others. The results of several instruments are combined using either the operation of intersection or the operation of union. This scheme is applicable only when the instruments produce homogeneous results.

The same schemes may be applied to subtasks of a task. For example, the subtasks of task 2.1 “Exclude inappropriate options” can be executed independently of each other (scheme 3). Their results (subsets of the initial set of options) are combined using the intersection operation that defines the result of task 2.1. The same scheme applies to the subtasks of task 2.2.1 “Define criteria”, but here the results (i.e., criteria defined) are combined through the union operation.

4.4 The user interface of the task-supporting guide

The guide may be activated on the user’s initiative at any time. At the beginning, the guide will ask which of the scenarios should be applied. Additionally, the user must tell the guide which data sets to use. For example, in the scenario “explore demographic data referring to territory division” the guide needs to know which of the available map layers contains the territory division and what attributes the user is interested in. This information is called the **context** of the scenario. After definition of the context the guide will display the hierarchy of potentially relevant tasks for the selected scenario (Figure 1).

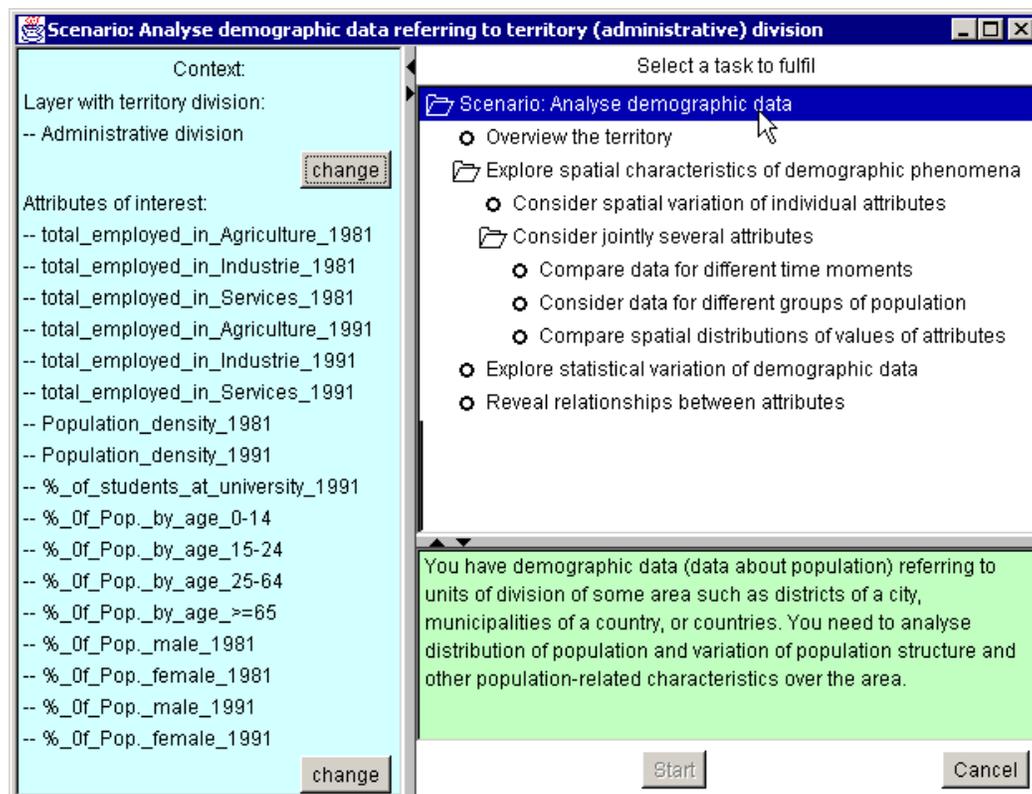


Figure 1. The guide displays the hierarchy of potentially relevant tasks for a user-selected generic scenario.

The left part of the window shows the currently defined context. The user may change this by pressing the “change” button. The upper-right part of the window displays the task hierarchy. Circles denote terminal tasks and “folder” symbols mark higher-level tasks. The lower-right part contains an explanation of the meaning of the currently highlighted task. By traversing the task hierarchy, the user may view explanations of all tasks.

When the user selects a terminal task for execution, the guide replaces the frame displaying the task hierarchy by a frame devoted to this particular task (Figure 2). As before, on the left the context is shown. On the right the system displays a general instruction how to use the map in order to do the task. Below this instruction the recommended instruments are listed. For each instrument the user can get a brief explanation of its meaning. The explanation is displayed in a popup window when the user moves the mouse cursor over the name of the instrument.

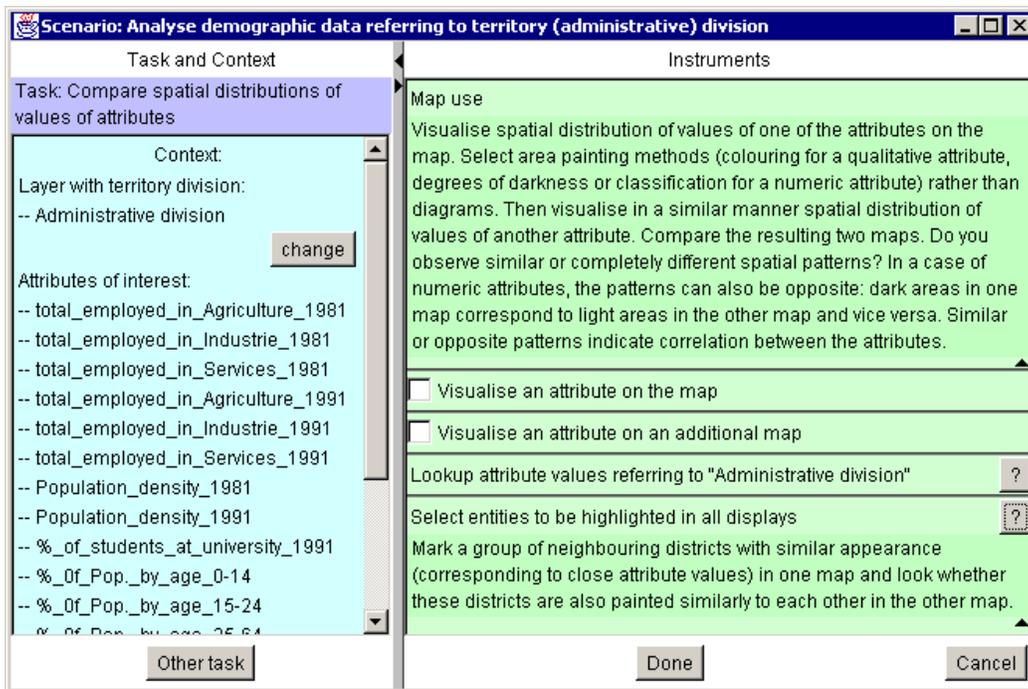


Figure 2. The guide displays a map use instruction and a list of instruments appropriate for the task “Compare spatial distributions of values of attributes”.

Instruments appropriate for a task are of two kinds. One kind is system functions or tools that must be specially activated. The other kind, such as access to attribute values or highlighting of objects, are always available without special activation. The guide marks instruments of the first kind with checkboxes, which may be clicked for activating them. Instruments of the second kind are also listed among the recommended instruments in order to at least remind the user that they may be helpful in fulfilling the task. The buttons labeled by question marks provide access to the appropriate pages of the system’s on-line manual where the user can find more extended explanations concerning the use of the tools. If provided in the knowledge base, the guide can additionally display task-specific instructions.

When the user activates some instrument by checking the corresponding checkbox, the guide may ask to specify the input for the instrument. Thus, for the instrument “Visualize an attribute on the map” the user needs to select an attribute. When the input is specified, the guide launches the appropriate tool. At the same time information about the tool is inserted into the guide’s window (Figure 3). This information consists of two parts: 1) general explanation of how the instrument works (for example, for a graphical display – how data are represented by graphical symbols); 2) task-specific instruction, that is, how to use the instrument for the current task. The latter is constructed from the instructions for the particular basic operations supported by the instrument (these instructions are specified in the knowledge base). The guide will show only instructions for those operations that are relevant for the current task.

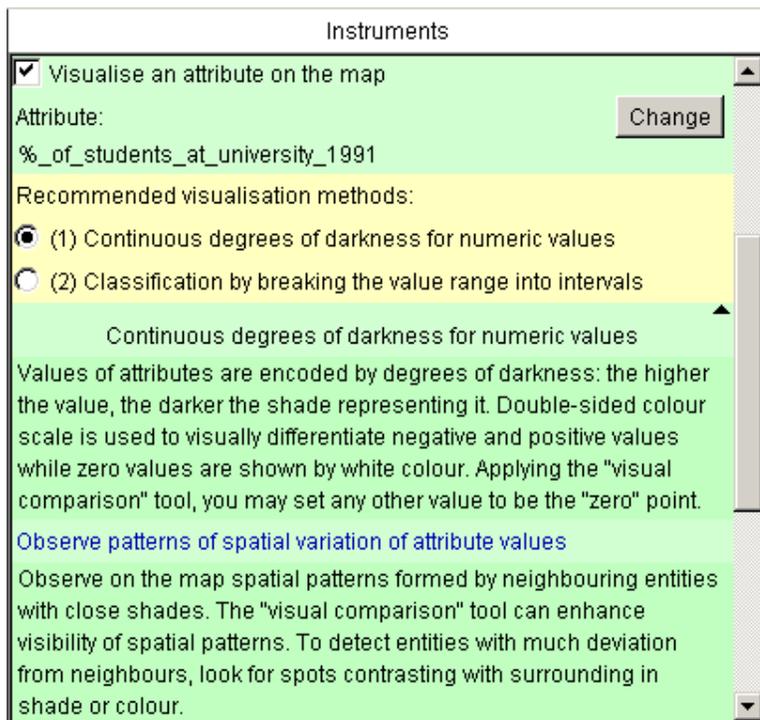


Figure 3. The guide explains how numeric values are represented using a continuous scale of degrees of darkness.

For the visualization of attributes on maps the guide only proposes cartographic representation methods that are appropriate with regard to the selected attributes and the current task. When the user selects one of the methods the data are immediately represented on the map by means of this method, and appropriate map manipulation tools are attached to the map. At the same time the guide inserts information about this method in the window for task support (Figure 3). Information about the manipulation tools is also provided.

To avoid overloading the user with masses of text, all the explanations and instructions appear initially in a “folded” state. A triangular mark indicates presence of a text. Clicking on the mark opens the text for reading. The user can “fold” the text again when it is no more needed.

In summary, the user interface of the guide serves the following purposes:

- Remind the user about potentially relevant subtasks of a task, and recommend an appropriate sequence of subtasks;
- Suggest analysis instruments when they are useful, that is, appropriate for the subtask being currently fulfilled;
- Activate the instruments when the user decides to apply them, that is, visualize data, create corresponding controls, etc.;
- Explain how to use the instruments that are relevant for the current subtask.

4.5 Directions for further work

The current design and implementation of the guide is a prototype to be further developed in several directions. First, the guide must be extended in order to process input-output links between instruments, from instruments to tasks, and between tasks. On this basis the guide must be enabled, when necessary, to induce a logical sequence of performing tasks. Second, the knowledge base, which currently only contains knowledge for two general scenarios, needs to be extended to a broader spectrum of scenarios. Third, the usability of the guide needs to be tested. In particular, the following hypotheses need to be validated:

- The concepts of general scenarios, tasks, and task hierarchies are understandable and appropriate for users, that is, users are able to relate their specific tasks to the general tasks the system deals with;
- The guide encourages the use of available interactive techniques, that is, more users will actively apply more of the interactive techniques;
- The instructions provided by the guide are effective, that is, people can and want to learn how to apply the new exploratory instruments without previous oral presentation and live demonstration;
- The user interface of the guide is understandable and convenient.

Regarding the user interface, we are in particular worried about the large amounts of text provided by the guide. On the one hand, it was learned that common users are reluctant to read any manuals or explanations, even when they are available on-line. On the other hand, it is necessary not only to prompt people into using the novel exploratory techniques but also to explain how these techniques work and how to apply them. This is hard to do this without text. Since the guide shows text in relatively small portions and only when applicable, there is a hope that the users will not feel overloaded. Of course, this needs validation.

It is also planned to pursue an idea of complementing textual information with visual methods that demonstrate the work of the tools. For example, when CommonGIS generates a choropleth map, the guide could automatically play an animation demonstrating how this map can be changed through the visual comparison tool. Such a visual demonstration may be sufficient for an advanced user to understand how to use the tool, and in this case there would be no further need in reading the corresponding texts. At the same time, a visual mini-demonstration could help casual

users in understanding the textual explanations. Besides, it might raise interest and willingness to learn more about the tool.

5 Implementation notes

The system CommonGIS is implemented in the Java language and can be used in two ways: as an applet running in a standard Java-enabled Web browser and as a local application, after being installed on a user's computer. The system has a component-based architecture and allows various configurations of the components, according to specific user's needs. For example, if it is not planned to use particular methods for cartographic visualization, statistical graphs, or computation modules, the respective components may be removed from the system. The same applies to the components for reading specific data formats.

Although the CommonGIS project was finished, the development of the system is continuing. Thus, it was not planned in the project to implement the user interface of the system in multiple languages. By the end of the project, the UI of the system was only in English as well as the texts in the on-line help. Since the system excited interest of potential users in Germany, it became necessary to create a variant of CommonGIS with a German UI and, in general, allow its easy extension to other languages (Peterson 1993, 1998). Recently the system has been modified so that all the texts used in the UI have been separated from the program code and put in so-called resource files. The resource files in English can now be substituted by similar files in German or any other language, and no changes of the program code are required. Currently the system can communicate with users in English or German. The on-line help is also available in two languages.

The post-project development of CommonGIS also includes enhancement of the system's functionality, improvement of the user interface, and implementation of customization facilities.

The system can be downloaded from the web site <http://www.CommonGIS.de/> and used free of charge for research and educational purposes. The downloading package includes instructions for the installation of CommonGIS on a local computer and on a web server, the program documentation, and the description of the system's API (application programmer interface) for those who would like to configure and extend the system according to their needs.

6 Conclusion

The overriding goal of the CommonGIS project was to find approaches and methods for making modern map-based exploratory facilities accessible to a broad community of users. From the technical perspective, this goal was attained by applying the Java applet technology. However, it was also necessary to ensure that the technical tools and functions could be utilized appropriately and effectively. Besides professional cartographers, the target user community included ordinary people using thematic maps for practical needs (such as selection of a house to buy) or for self-education, and working people from various fields that might occasionally use thematic maps in their profession. These potential casual users typically have no specific expertise in cartography, principles of graphical representation of information, and geographical information systems. CommonGIS was designed and realized as a knowledge-based

system that incorporates this expertise and applies it for three complementary purposes. First, it assists users to represent data on maps and other graphical displays in accord with principles of cartography and graphics design. Second, it suggests exploratory instruments suitable to the goals of analysis. Third, it assists in utilizing these instruments by context-specific instructions about how to operate them. Such instructions are especially important because many of the instruments are novel even for experienced GIS users.

Although we have no doubt that such knowledge-based services can be of great use, further tests will be needed regarding the usability of the specific implementation of these services.

Acknowledgements

We thank all our CommonGIS project partners for fruitful collaboration in the course of the project. We are grateful to the project officer Peter Diry and the reviewers Prof. Werner Kuhn and Prof. Gerhard Joubert for their constructive criticism and valuable comments and advice. We would also like to thank all participants of the usability tests conducted within the project.

References

- Andrienko, G. and Andrienko, N., 1999. Interactive maps for visual data exploration. *International Journal Geographical Information Science* 13(4): 355-374
- Andrienko, N., Andrienko, G., Voss, H., Bernardo, F., Hipolito, J., and Kretchmer, U., 2002. Testing the Usability of Interactive Maps in CommonGIS, *Cartography and Geographic Information Science*, accepted
- Bertin, J., 1983. *Semiology of Graphics. Diagrams, Networks, Maps*. Madison: The University of Wisconsin Press
- Casner, S.M., 1991. A Task-analytic Approach to the Automated Design of Graphic Presentations. *ACM Transactions on Graphics* 10(2): 111-151
- Jung, V., 1995. Knowledge-based Visualization Design for Geographic Information Systems, in *Proc. 3rd ACM Workshop on Advances in GIS*. ACM Press, pp. 101-108
- Jung, V., 1996. Fuzzy effectiveness evaluation for intelligent user interfaces to GIS visualization, in *Proc. 4th ACM workshop on Advances in GIS*. ACM Press, pp.155-162
- Jung, V., 1998. *Integrierte Benutzerunterstützung für die Visualisierung in Geo-Informationssystemen*, Fraunhofer IRB Verlag, Fraunhofer-Institut für Graphische Datenverarbeitung IGD, Darmstadt
- MacEachren, A. M., 1994. Visualization in modern cartography: Setting the agenda. In: *Visualisation in modern cartography*. New York, New York: Elsevier Science Inc. pp.1-12.
- MacEachren, A.M., and Kraak, M.-J., 1997. Exploratory cartographic visualization: Advancing the agenda. *Computers and Geosciences* 23(4): 335-44.
- Macintosh, 1995. *Macintosh Human Interface Guidelines*. Reading, Massachusetts: Addison-Wesley Publishing Company

- Mackinlay, J., 1986. Automating the Design of Graphical Presentation of Relational Information. *ACM Transactions on Graphics* 5(2): 110-141
- Murray, B.S., 1994. Intelligent Information Presentation Systems. *The Knowledge Engineering Review* 9: 269-286
- Peterson, M.P., 1993. Improving the International User Interface in Computer Cartography, in *Proceedings of the 16th International Cartographic Conference 1993*, pp.870-880.
- Peterson, M.P., 1998. Improving the Design and the International User Interface of Maps on the Internet, in *Proceedings of WebNet '98 - World Conference of the WWW, Internet and Intranet 1998*, pp. 1155-1156.
- Roth, S.M. and Mattis, J., 1990. Data Characterization for Intelligent Graphics Presentation, in *Proc. SIGCHI'90: Human Factors in Computing Systems*. ACM Press, pp. 193-200
- Senay, H. and Ignatius, E., 1994. A Knowledge-based System for Visualization Design. *IEEE Computer Graphics and Applications* 14 (6): 36-47
- Zhan, F.B. and Battenfield, B.P., 1995. Object-oriented Knowledge-based Symbol Selection for Visualizing Statistical Information. *International Journal of Geographical Information Systems* 9: 293-315