

USING LOG-FILE ANALYSIS FOR TESTING CARTOGRAPHIC WEBBASED APPLICATIONS

Mette Arleth
GIM - Geoinformation and Mediatechnology
Aalborg University
Fibigerstraede 11
Dk-9220 Aalborg East
fax: +45 98 15 57 75 Email: marleth@i4.auc.dk

Introduction

Testing and evaluating chosen design strategies for cartographic applications - related to the cartographic design as well as the interface design - in an efficient and objective manner, and within a limited budget, remains a problem. On the one hand one often wants to get statistically sound and objective data evaluating the appropriateness of certain design strategies, or comparing the usefulness of two or more different design solutions. On the other hand setting up test-environments and analysing test data is a considerable task, whose applicability and reliability can be discussed. One risk is that the test is arranged in a way that biases the results to the advantage of the result preferred by the designer. Another risk lies in drawing exaggerated conclusions based on simple measurements of "time spent on task". The appropriateness of drawing qualitative conclusions based on quantitative measurements will always be a point in dispute in a science involving humanistic as well as technical theory.

A quite simple method for testing different aspects of the design of cartographic webbased applications is logfile analysis. The logfile is created by the webserver and will usually contain information related to all traffic on the server (Møller-Jensen, 1999). The logfile thereby keeps track of the users' actions on the webpages, which steps have been taken, and how much time has been spent on each page or action. An important task in logfile analysis is obviously to extract the records relevant to the users involved in the test. Afterwards the files can be analysed according to the number and order of actions and time consumption. The logfile thereby constitutes a sort of nonstop observation and registration of the user, an accumulation of data that would need several persons to collect in other ways (Todd, 1994). Obviously the logfile does not contain information about WHY the user performed specific actions or HOW he or she perceived and comprehended the information and different design solutions. Hence the logfile analysis should be combined with either an online questionnaire or an interview, lending room for more qualitative explanations.

In a Ph.D. study concerning screen map design logfile analysis was used together with a questionnaire. Some of the tested assumptions referred to aspects of the specific application - a local authority plan published on the Internet. But the test included assumptions of more general applicability as well:

1. Using the map as an interface for navigating through the information would be

- more effective and more intuitive for the user,
2. Specific cartographic design strategies would result in a more readable and comprehensible map, and thereby in better perception, leading to a faster completion of the task involving the map, and
 3. Specific interface design strategies would encourage the user to explore the application and thereby enhance the understanding of the map.

This paper will report on some problems discovered whilst using the log-file analysis and suggest some solutions. In addition, some results from the test will be presented and discussed.

Performing the test

The test was performed over a three-days period in November 2000. The test persons used a webbased version of the local authority plan to find answers to 25 questions concerning the information in the plan. The test took place in a computer laboratory at Aalborg University, Department of Planning and Development. Each questionnaire was marked with date, time and identification number of the computer used by the test person. This information was used to relate the logfiles generated by the test persons to the questionnaires containing their answers. From the department's webserver log, the logfiles resulting from the actions of the test persons were extracted. The extract contained little more than 2500 records. Afterwards these records were analysed in order to confirm or reject the test assumptions.



Figure 1 Example from the tested application. The users could choose to navigate through the information using the links in the maps or the buttons and the links in the text. It was assumed that using the map for navigation would be the most intuitive and therefore preferred option.

Testing Assumption nr 1

To find the information needed to answer the questions in the questionnaire the test persons had to open specific pages in the application. In the logfile analysis the order in which these pages were opened by the test person was regarded as a sort of path. Obviously an optimal path could be defined, leading to the desired information directly and in the correct order. Comparing the paths constructed by the test persons with the optimal path would result in useful indications of whether the userinterface design was perceived as logical and clear. Moreover, to be able to confirm or reject the first assumption it was necessary to discover whether the test persons moved from page to page by using the hyperlinks in the map or the hyperlinks in the text and on the buttons (see figure 1). In most cases the application supported navigation in text as well as in maps, but as the information required for answering the questions of the test were all geographically related it was assumed that navigating through the hyperlinked maps would be preferred by most users (Arleth 2000).

Testing Assumption nr 2

It was assumed that employing specific designstrategies, particularly regarding the choice of text size, font and style would affect the readability of the map in a positive way. Hence two comparable maps were made part of the application, one assumed to be more readable than the other. The test involved task solution using these two maps. Based on the information in the logfiles the time spent by the test persons on each of the two tasks could be calculated. Provided that the task involving the maps that were assumed to be the more readable was generally performed faster than the other task, the assumption would be regarded as confirmed.

Testing Assumption nr 3

In the tested application the maps serve not only as a navigational tool through the information but also as an interface to supplementary visualisations, such as orthophotos, 3D-models, QTVR-panoramas etc. To access these additional visualisations the user has to activate a function called "The Glasses". When activating the Glasses the user is informed what kinds of visualisations can be accessed through the map currently in display. The basic idea is, that if the user knows that something can be found in the map (something worth looking for), but not where and what, the user's curiosity will be challenged. It is assumed that at least users familiar with the medium will wish to explore the map further. In the logfile analysis it could be discovered how many supplementary visualisations the test persons had looked up. A large number would be indicative of a confirmation of the assumption.

Lack of log (or luck)

Unfortunately the resulting logfile was less informative than expected. Not only did it prove not to include all of the information required for testing the different assumptions, in several cases there had been generated no log at all. In three cases two or more computers had been identified by the log as being the same, thereby mixing the actions

of two or more test persons. In most of the remaining cases the path constructed by the actions of the test persons had “holes” in them, which means that the path shows a jump from one page to another, even though the two pages are not directly linked. These “holes” in the path obviously occurs when pages are stored in the computer’s cache. Next time the page is opened there will be no action on the webserver, and thereby no record in the log. But the cache does not explain all the irregularities.

The irregularities mentioned substantially decreased the number of useful test logs, making any conclusions less sound. Furthermore it had been expected, that the logfile would contain the information required to analyse whether the test persons preferred the text links or the map links when navigating through the plan. This turned out not to be the case. Due to time constraints in the project, the online user test could not be repeated, and to reduce the omissions in the resulting test material, supplementary test methods had to be applied. To be able to confirm or reject the assumption about the preferred navigation, observations of test persons were made, noticing how the test persons moved from page to page. Obviously this test method could only be applied to a limited number of test persons, and hence the results can only serve as an indication. In stead of evaluating the readability of the text on the maps (the second assumption) user preferences about different text styles, fonts and sizes were investigated. This is not directly comparable to testing the readability of the text, as the users may not prefer the style or font which is objectively the most readable one, if they are more accustomed to a different design.

In this way, by applying different testing methods, results were found and conclusions could be made regarding the different assumptions. Had the project period been less limited as regards the time, the online test could have been repeated using more sophisticated logging programs. One possible solution would be to apply a hit counter to the tested application. Several companies offer hit counters for webpages. The hit counter may be visible on the page or hidden, and a logfile of transactions on the webpage can be generated. Most hit counters offers information about number of visitors, search engine, browsers and operating systems etc. But some companies offer more detailed information. One example is the SuperStats from MyComputer.com. Subscribers to SuperStats can benefit from very detailed information about the visitor behaviour on their web-pages. One service that would have been interesting in this project is the Site Path Reports. The Site Path Report seems to contain almost all the information needed in the test just mentioned. However, SuperStats has a drawback; it is NOT free, and thereby breaks with a central premise for the choice of testing method.

Test results

The paths constructed from the logfiles combined with the questionnaires gave a clear indication that the test persons were able to navigate purposefully in the application. The complementary observation of test persons navigating through the application showed that the majority of the test persons chose to use the map links most of the time. A small number of test persons chose not to use the map links at any time; these persons

directed all their attention to the textual information, finding all the answers there. These observations emphasise the importance of supporting the textual approach as well as the map approach to the use of such applications as webbased local authority plans.

To visualise to what extent the test persons were able to follow the ideal path through the pages necessary for the task solution figure 2 was made. The figure shows all the pages of the tested application, but not their mutual connections and hierarchy. Lines are drawn to those pages that the test persons had to look up to solve the tasks in the second half of the test. The two dotted lines indicate that those pages are relevant but not necessary in the navigation through the information. Figure 3 is a visualisation of the paths accumulated by the test persons. The lines vary in colour and width to show which of the pages has the highest hit rate. A comparison of figure 2 and 3 clearly shows that the test persons have been able to look up the relevant pages, giving these the highest hit rates. Figure 3 also shows that most of the pages in the application have been looked up during the test at least one time, but none of the irrelevant pages have been systematically looked up by the test persons.

The purpose of testing the readability of the texts in the maps was to investigate an observation made during the process of designing the maps. When texting screen maps



Figur 2 A visualisation of the ideal path through the application in part of the test. The dotted lines point to pages that were relevant but not necessary for the navigation.



Figure 3 The paths accumulated by the test persons. Though several pages not necessary for the task completion has been looked up, the diagram from figure 2 can be identified, which shows that the test persons have found the right pages.

it is hard to make the texts small and readable at the same time, due to the limited resolution of the screen. Particularly texts that are not placed horizontally suffer from graphic distortions. During the design work it was observed, that text in italic style seemed to be more robust against these distortions than normal style texts (Arleth, 1999). Hence it was found interesting to test, whether the italic styled texts were found more readable by the test persons. As mentioned above this was done by asking the test persons about their preferences in a line of map pairs showing different choices of text font, style and size. Italic style text was not preferred by a majority of the test persons in any of the map pairs. However, it could be found that the italic style texts had twice as many supporters on maps with nonhorizontal text placement, as on maps with horizontal text placement. Even though this can not be taken as a confirmation of the assumption of better readability, it could be seen as a confirmation of the observation of the robustness of the italic style texts.

The test persons were generally content with the possibility of finding complementary visualisations using the Glasses. The log files show that test persons who activated the Glasses looked up at least 3 different visualisations. However, the Glasses were criticised for being less understandable and intuitive than the other parts of the interface.

Particularly the lack of visual clues in the map to indicate the number and placement of the complementary visualisations was mentioned as a problem. The design strategy behind not showing these things was to stimulate the user's curiosity, and make them search round the map for active spots. This solution was inspired by observations of pre-school children working with interactive PC programs. These children merrily point and click anywhere in the applications in a truly explorative manner, expecting something to happen. Apparently adults do not like to work that way; they want visual clues of the possible actions. Making things visible remains one of the most essential and general recommendations in interface design literature e.g. (Norman, 1989), (Norman, 1993), (Carroll et al. 1991), (Nielsen, 1993). The evaluation of the Glasses definitely proves the relevance of this advice - when the intended user is an adult.

Conclusion

Primarily due to communicational problems with the administrator of the network in the department, the online test described in this paper suffered from a line of technical problems, resulting in a flawed and insufficient logfile. This could have been avoided, if programmes generating a more comprehensive logfile had been used. If nothing else, this stresses the importance of good relations to the network administrators, if the success of a test depends on their contribution. In this case it was found necessary to complement the online test with other test methods, making it possible to draw conclusions on the tested assumptions.

References

- Arleth, 2000, Mette Arleth: The use of GI in planning documents published on the Internet, 3rd AGILE conference on geographic information science, Espoo 2000
- Arleth, 1999, Mette Arleth: Problems in screen map design, 19th International Cartographic Conference, Ottawa 1999
- Carroll (ed) 1991, John M. Carroll (ed): Designing interaction, Psychology at the human-computer interface, Cambridge University Press, Cambridge 1991
- Møller-Jensen 1999, Lasse Møller-Jensen: Monitoring user responses to webbased GI interfaces, 2nd AGILE conference on geographic information science, Rome 1999
- Nielsen 1993, Jakob Nielsen: Usability Engineering, Academic Press Inc. 1993
- Norman 1988, Donald A. Norman: The design of everyday things, Currency Doubleday 1988
- Norman 1993, Donald A. Norman: Things that make us smart: defending human attributes in the age of the machine, Addison-Wesley Publishing Company 1993
- Todd 1994, Peter Todd: Process tracing methods in the decision sciences, in Nyerges et. al: Cognitive aspects of human-computer interaction for geographic information systems" Kluwer Academic publishers 1994
- www.mycomputer.com
- The tested application (all text is in Danish) is found on my homepage:
www.i4.auc.dk/marleth/komplan_site