An Empirical Assessment of a Web-based PPGIS Prototype

Jianfeng Zhao and David J. Coleman

Abstract

Building on earlier Public Participation Geographic Information System research, a GIS-enabled Online Discussion Forum prototype named GeoDF is now being implemented and evaluated. This paper aims at empirically evaluating the usability of the GeoDF prototype to determine: (1) whether or not this new technology will indeed lead to "better" participation; and (2) how non-expert users use and accept the prototype. This paper describes the design of the usability evaluation process in detail to measure the usefulness and ease of use of the GeoDF prototype. On a more general level, the paper presents the usability evaluation design procedures, which aims to contribute to the PPGIS research by providing design guideline for future PPGIS application evaluation. The preliminary evaluation shows that the users are generally satisfied with the GeoDF prototype functionality. Nonetheless, the users' acceptance of PPGIS technology is largely determined by their motivation and other human factors.

INTRODUCTION

A growing body of research called Public Participation Geographic Information Systems (PPGIS) has emerged in the use of GIS and web-based mapping and analysis systems in participatory planning processes (Wong and Chua 2001, Kingston 2002, Craig et.al. 2003, Steinmann et.al. 2004, Keßler and Rinner 2005, Tang 2006). New technologies have emerged to help citizens have a more active role in public service delivery. The social demand for participation in the planning process is on the rise. The goal of the GIS-enabled online Discussion Forum (GeoDF) software prototype originally developed through 2004-2006 by Teresa Tang (2006) is to integrate Web mapping and on-line discussion forum capabilities to help facilitate such participation.

GeoDF (see Figure 1) enables citizens to provide more in-depth feedback to government through the use of Web-based mapping and online collaboration tools (Tang et.al. 2005). The prototype supports the participants to submit and share feedback, as well as to initiate discussions about their concerns; participants can express their views not only with text messages, but also make sketches and annotations on the GIS map. In order to better convey a participant's perspective, the map extent and the map layers that one is viewing is stored by the system and shared among the participants. Moreover, the discussion contributions (i.e., the text messages, GIS map, sketches, and annotations) are organized and presented in a structured way to facilitate the understanding of the evolution of ideas throughout the discussion process. The target user group for the GeoDF prototype is the general public, due to the wide range of possible users in this group of users, which requires that the prototype system must be accessible and easy-to-use for all levels of users. This introduces a motivation for a usability study for the prototype.

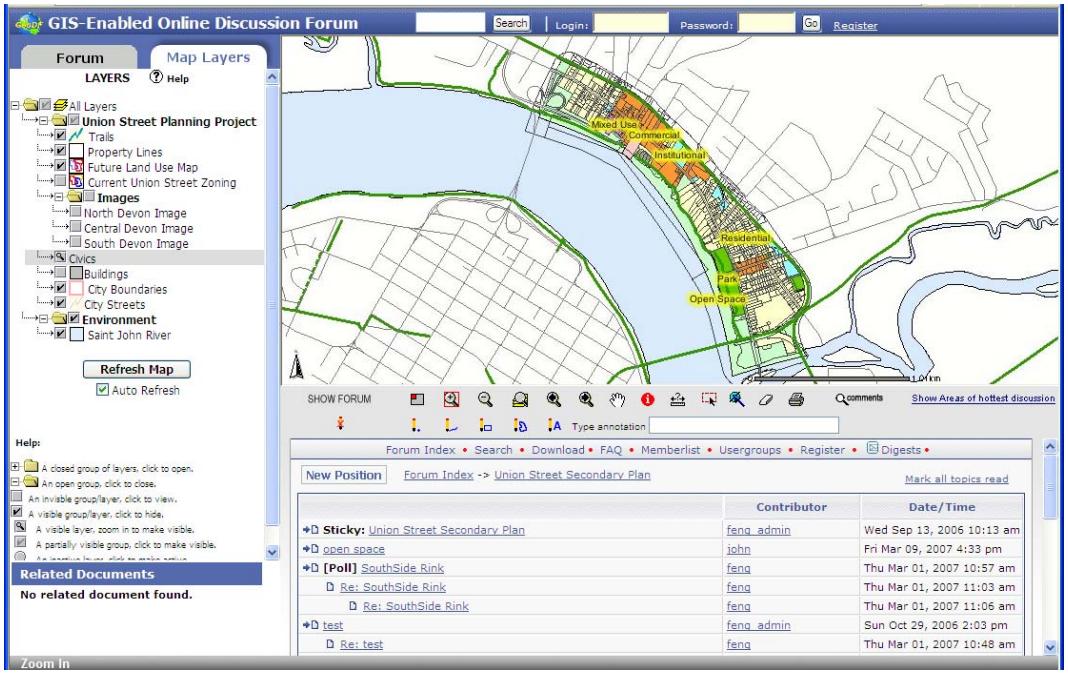

Figure 1 GeoDF prototype with the web-mapping component to the upper right and the discussion forum to the lower right (data source: City of Fredericton)

Human Computer Interaction (HCI) examines the interaction between humans and computers (Preece et.al. 2002). Usability studies form a small part of the larger field of HCI research, investigating the extent to which a computer system supports users to achieve specified goals in an effective, efficient, and satisfactory manner (Nielsen 1993). The experiment performed in this study attempts to empirically evaluate Tang's GeoDF prototype to determine: (1) how non-expert users are using the prototype; and (2) whether or not this new technology will be accepted by the general public and finally result in greater participation. This paper brings together PPGIS research, usability evaluation methods and Technology Acceptance Model (TAM) in HCI, begins with a brief overview of the HCI research, usability evaluation, TAM as well as related work that has been carried out within the PPGIS research. Subsequent sections describe the research methodology, design of the usability test, the proposed test plan and evaluation metrics to evaluate how the GeoDF prototype is understood and used by non-expert users.

HUMAN COMPUTER INTERACTION AND USABILITY EVALUATION FOR PPGIS

Preece et.al. (2002) defined HCI as “a discipline concerned with design, evaluation and implementation of interactive computer systems for human use and with the study of major phenomena surrounding them.” As long as there have been computers, developers have been concerned with how the machine and its software will be used. The major concern in HCI is to facilitate and improve people’s work through the use of computers. HCI studies also aim to meet the needs and requirements of the users with computer systems that are safe, efficient and enjoyable (Preece et.al. 2002). In HCI, "usability evaluation" refers to the process of systematically collecting data on how users use the system for a particular task in a particular environment (Preece etl.al. 2002). The objectives of HCI and usability evaluation are to make computer technology accessible and easy to use for a wider range of users. Usability includes

considerations such as: Who are the users? What do users want or need to do? What is the general background of the users? What is the context in which the user is working? All these usability considerations are intrinsic goals of the PPGIS research – making GIS more accessible and easy to use in order to enhance citizens' roles in decision-making processes.

Using Usability Techniques for PPGIS

Examples of web mapping tools and methods developed and refined for widespread use by the general public have been described by Kingston (2002), Evans et.al. (1999), Keßler (2004), and Voss et.al. (2004), among others. A recent review of PPGIS research and applications literature reveals, however, that more emphasis has been placed on technological aspects involved in developing the applications rather than empirical study of whether the tools are usable or how these tools are being used (Craig et.al., 2002; Haklay, 2003; Steinmann et.al., 2004). Wong and Chua (2001) described four barriers that particularly exist in Web-based PPGIS: cost of interactivity, user diversity, data and copyright costs, and trust and legitimacy. Steinmann et al. (2004) conducted a qualitative expert analysis evaluating twelve PPGIS applications according to their usability, interactivity and visualization, and made comparisons between the US and Europe. Sidlar and Rinner (2006) employed a quasi-naturalistic[1] case study and focused on the general usability aspects such as cost of entry, efficiency, interactivity and connectivity of the Argumentation Maps prototype. Demšar (2007) introduced a low-cost methodology for performing usability evaluation which combines formal and exploratory usability evaluation methods, and then assessed how the participants used the geovisualization tool on a real data set. Among these PPGIS applications, a number of usability experiments have been carried out (Haklay and Tobon, 2003, Sidlar and Rinner 2006, Demšar 2007) and indicated valuable connections between HCI, usability, and PPGIS.

Haklay and Tobon (2003) contend that HCI issues are vital to the success of PPGIS because HCI techniques, including usability evaluation methods, are geared towards understanding how people interact with computer applications within an environment, and therefore would help the PPGIS researchers understand users' expectations as well as the ways in which they use, understand and value the system. Further, they proposed that PPGIS is designed to be used directly by a "general public" possessing potentially a diverse range of world views, cultural backgrounds and knowledge. In such situations, it requires the system to be easily usable and understandable by a broad public audience. Finally, Haklay and Tobon suggested that the quantitative and qualitative data from HCI and empirical studies would provide valuable information on the usage and the role of PPGIS within a wider societal context.

Peng and Tsou (2003) contend that Web-based GIS increases the availability of geospatial data, reduces end-user cost, and offers flexible and customized user experiences through the use of web clients. Steinmann et.al. (2004) however pointed out that the specialized functionality that supports online GIS increases the complexity of a conventional browser experience.

Nontheless, without empirical evidence, it is hard to tell whether or not the advances in technologies will actually benefit and empower the general public, whether or not these PPGIS

[1] Quasi-naturalistic studies use a 'real-world' context but are used with such controls so that both evaluation and collecting of information are easier, and therefore a deeper investigation can be achieved than naturalistic studies.

tools will lead to increased participation, equality and better democracy. It is therefore difficult to measure:

- Are these tools used?
- What are the tools used for?
- How would the stakeholders use these tools?
- How much are these new tools accepted?
- What are the effects of using the PPGIS tools?

Technology Acceptance Model and System Acceptability

The Technology Acceptance Model (TAM) developed by Davis (1989) that offers researchers and practitioners a relatively simple and cost-effective way to predict the ultimate measure of system success—whether or not that system is good enough to satisfy all the needs and requirements of the users. It is an information systems theory that models how users come to accept and use a technology. The goal of TAM is to predict information system acceptance and diagnose design problems before users have any significant experience with a system. Davis included two constructs in TAM (see Figure 2): *Perceived usefulness (PU)*: "the degree to which a person believes that using a particular system would enhance his or her job performance"; and *Perceived ease-of-use (PEOU)*: "the degree to which a person believes that using a particular system would be free from effort". For example, on an exam dealing with the Netscape World Wide Web browser, Morris and Dillon (1997) reported the PU with a reliability of .92, the PEOU with a reliability of .90, and Attitude toward Using with a reliability of .85, they further concluded that the PU is found to have a strong influence on people's intentions, while the PEOU had a smaller but still significant effect that subsided over time.

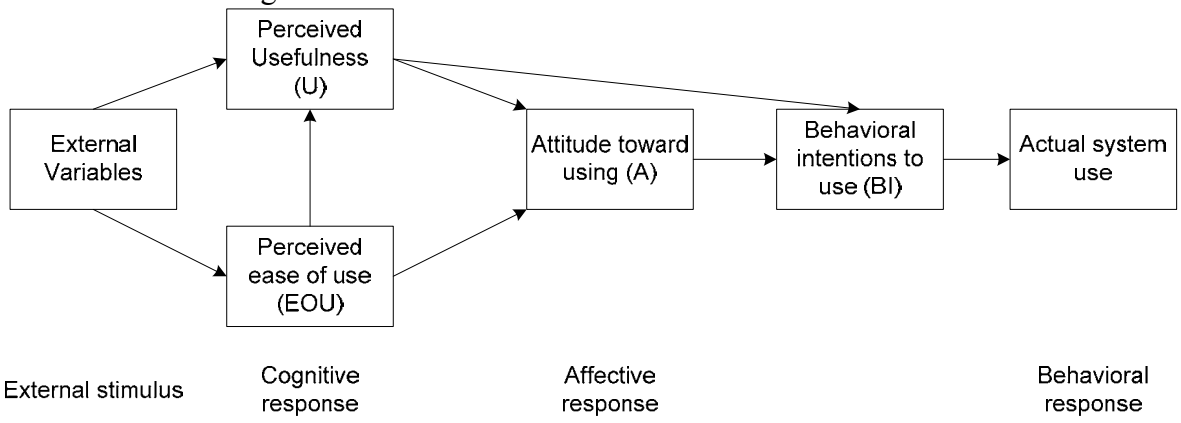

Figure 2 Technology Acceptance Model (after Morris and Dillon 1997)

Nielsen (1993) considered usability as one of the many attributes of *system acceptability* (see Figure 3). The overall acceptability of a system is a combination of its social and practical acceptability. Social acceptability shows if the system is meant for ethical purposes. For example, hacking personal data from a hospital database is not considered socially acceptable. Practical acceptability is the generalization of acceptability of system's cost, compatibility with existing systems, reliability, available support, usefulness and other such considerations. Nielsen (1993) also stated that "usefulness" is the issue of whether the system can be used to achieve some desired goal. Usefulness can be further divided into "utility" and "usability". Utility is the

question wether the functionality of the system in principle can do what is needed and usability is the question of how well users can use that functionality.

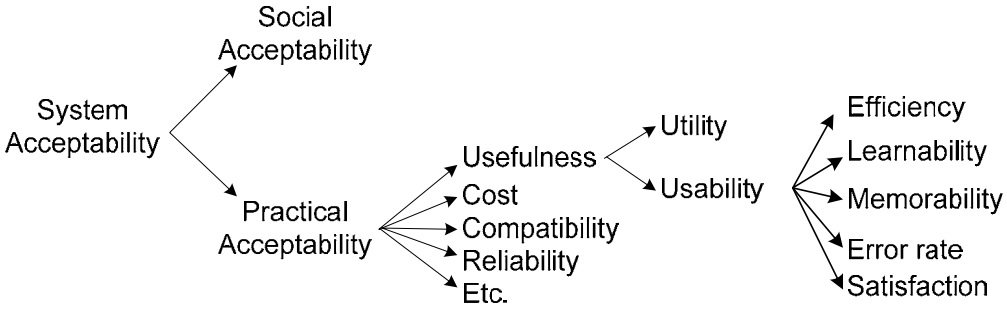

Figure 3 Nielsen's definition of usability as part of acceptability (Nielsen 1993)

Thus, all these elements of system acceptability are very important from the general viewpoint, usability is just one of many elements that should be considered while evaluating the effectiveness of the GeoDF prototype.

Design of Usability Evaluation Criteria

In this paper, the authors attempt to better understand how people use the GeoDF prototype. A series of tests must be carried out in order to investigate whether or not the GeoDF prototype is usable and acceptable. As depicted in Table 1, the evaluation criteria are derived from system acceptability (Nielsen 1993), TAM, user-based usability testing approach (Sweeney et.al. 1993) in order to: (1) examine the system acceptability of the GeoDF prototype in a from several categories, such as usefulness, cost, and reliability; and (2) assess the five main usability factors: "learnability, efficiency, memorability, error rate and satisfaction" (Nielsen 1993). More specifically, the authors are more interested in determining the measurable extent to which improved usability and enhanced social collaboration capabilities in a web-mapping system will result in greater public participation in a planning process.

Evaluation Criteria	Explanation
Usefulness	Usefulness refers to whether the system can achieve the goals of participatory planning. This also refers to the trade-off between providing sophisticated functionality and maintaining simplicity.
Ease of Use	Ease of Use refers to whether the users find the system easy enough to use. This is reflected in the levels of speed, completeness and correctness in the user's performance during the testing.
Cost of Entry	Cost of entry refers to the level of user investment. It is measured by both physical and mental complexity and stress/anxiety of the users.
Intended Users	"Intended users" refers to the demographic background of the different users that are testing the prototype.
Satisfaction	Satisfaction refers to the degree of general positive regard or emotion that the users attribute to the interaction with the GeoDF prototype. This is reflected in the level of positive attitude/opinion which is reported by the users.

Table 1 Evaluation criteria for GeoDF.

Sweeney et al. (1993) classified three main approaches to evaluate human-computer interaction: "user-based approach", "expert-based approach", and "theory-based approach". The user-based approach (see Table 2) is selected for this research as more than one user are involved in the test to complete more than one tasks in an appropriate environment. Detailed test plan and preliminary results are presented in the subsequent sections.

Usability Indicators	Data Collected
Performance (user)	Task time, % completed, Error rates Duration of time in help, Continuance of usage, Range of function used (objective)
Non-verbal behaviour (user)	Eye movement, Orientation duration Frequency of documentation access (objective)
Attitude (user's attitudes and opinions)	Questionnaire and Survey responses, Comments from interviews and ratings, Answers to comprehension questions (subjective)
Cognition	Verbal protocols, Post-hoc comments (objective)
Stress	Galvanic skin response, Heart rate, Event-encephalograms, Rating or comments (objective and subjective)
Motivation	Enthusiasm, willingness and effort (subjective)

Table 2 Usability and Acceptability indicators for GeoDF. (after Sweeney et.al. 1993)

METHODOLOGY

Preece (2002) outlines five categories of usability evaluation methods: analytic, expert, observational, survey, and experimental. Preece described an *analytical evaluation* as using "interface descriptions to predict user performance". An *expert evaluation* uses identified experts in the related field to analyze and evaluate the system. An *observational evaluation* consists of evaluation of the behavior and reactions of users when using the system. A *survey evaluation* solicits users' opinions on the use of the system through the use of a questionnaire or interview. Finally an *experimental evaluation* utilizes the scientific practice of controls to analyze the prototype.

In this line research, the senior author carried out a two-stage usability test, combining analytic, observational and survey evaluation methods. The experiment was performed using very limited resources: the author acted as the developer, coordinator, facilitator and usability observer for both stages. The methodology was developed based on the recommendations of experienced usability expert and usability engineering literature (Nielsen 1993). Both stages were structured in three parts: (1) an introductory session to describe to the participant about the PPGIS concept and logistic of the usability test; (2) a hands-on session where the participants explored the

prototype and completed a set of predefined tasks; and (3) an interview/questionnaire session to gather qualitative data and subjective ratings from the participants. Figure 4 illustrates the approach developed to achieve the research objective.

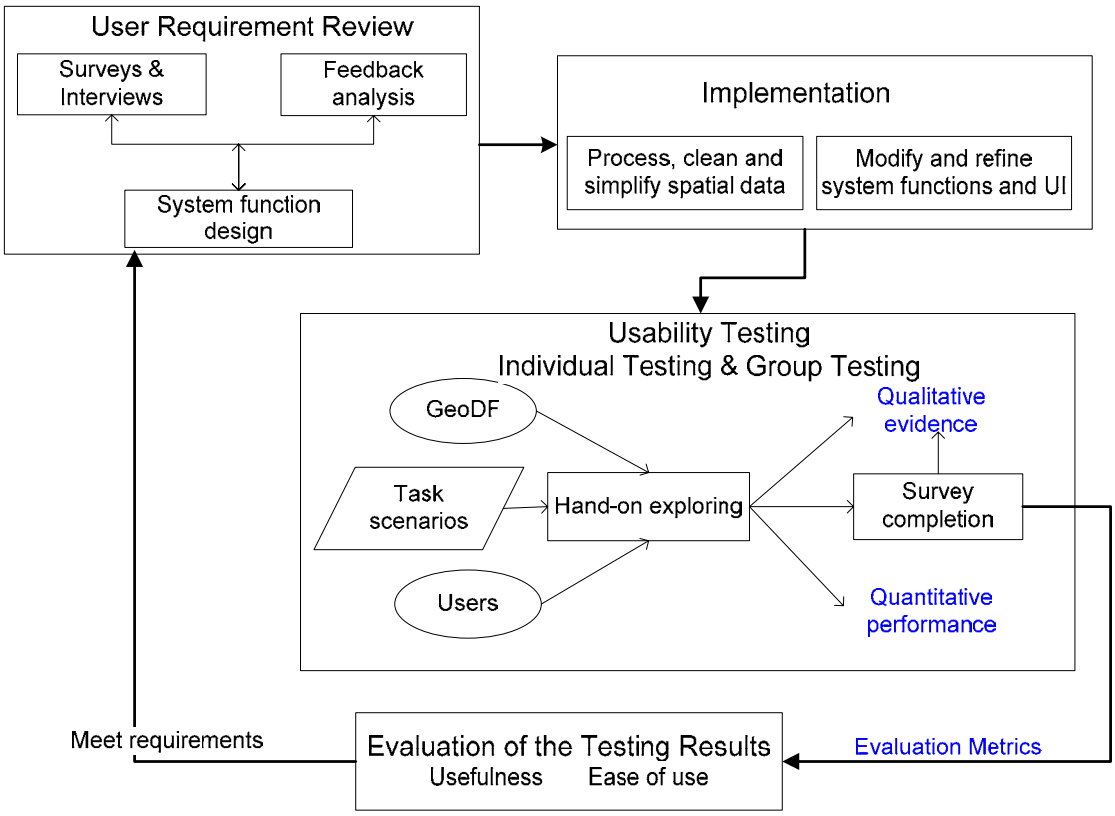

Figure 4 Research Methodology

User Requirement Reviews

Tang (2006) implemented the GeoDF prototype as a proof of concept. She designed the system based on an extensive literature review and implemented the key components and features of the prototype at minimum development costs. In order to analyze GeoDF as a 'real world' application, the author first needed to determine the types and levels of additional functionalities needed in GeoDF to meet user needs and requirements as well as facilitate equal access to information in the 'real world' planning process. An information user requirements review was conducted that included surveys, interviews and focus group meetings with 11 different program managers and technical support staff in the City of Fredericton and the New Brunswick Lung Association – two organizations heavily involved in planning processes involving public consultations.

Implementation

Feedback and comments gathered from these meetings were analyzed and server as the refinement guideline for additional system functionalities. Feedback and comments from those meetings included, for example:

1) *Issues about user sign-up and password protection*.
In order for an input to be valid and prevent from domination of the discussion forum by a small group of malicious users, the City of Fredericton needs to know the identification of the participants. Therefore, user sign-up and password protection mechanism was incorporated into the prototype to check the identification of the users in the registration process.

2) *Concerns over a moderated versus non-moderated discussion list*.
The City of Fredericton requested tools to pre-filter comments to eliminate inflammatory/pornographic comments/personal links. A moderated list was put in place in order for the discussion forum to obtain the most valuable and clean user inputs.

3) *Issues over spatial data source, format and permission*.
Both spatial and aspatial data are required for the set up of GeoDF. Spatial data are the map layers and associated attributes. Spatial data in GeoDF must be in shapefiles (*.shp, *.shx, and *.dbf) format (the only supported file-based format in ArcIMS). City of Fredericton has a Caris GIS environment with Caris NTX format data. Therefore, relevant spatial data files were processed, converted, cleaned and simplified to improve the readability of the map data.

4) *Issues over cross browser compatibility*.
The client software platform is essential for GeoDF to fully function. For proof-of-concept, the prototype system only worked with Internet Explorer 6 and under. Modifications were made for GeoDF to function with the major Internet browsers.

Usability Testing

In order to evaluate the usability and acceptability of GeoDF, the senior author designed a two-stage test plan slated for completion during May/June 2007.

Data capture / Usability Indicators	Observation of users during the experiment	Video-recording of user interaction	Audio-recording of user during interaction	User's post-hoc comments	Self-administered time management	User interview	Questionnaire survey and rating scale.
User's Performance	X	X			X		
Behaviour	X	X					
Attitude			X	X		X	X
Cognition			X	X		X	
Stress			X				X
motivation			X	X		X	X

Table 3 Data Capture Methods used to collect data during the testing.

Data Capture Methods The two rounds of usability testing covers a selection of user-based indicators, these includes: user's performance, behaviour, attitude, cognition, stress and motivation. Sweeney et.al. (1993) offered the most accessible and frequently employed data capture methods to match relevant usability indicators. Data capture methods for this testing included observation of users during the experiment, video-recording of user interaction, audio-recording of user during interaction, user's post-hoc comments, self-administered time management, user interview and user questionnaire survey and rating scale. Table 3 shows the usability indicators and the corresponding data capture methods.

Test Plan The first stage involves meetings with five to eight people individually and aims at detecting most of the usability problems. (Spool (2001) suggests that -- for evaluation purposes -- some website usability engineers believe between five and eight users are all that is needed to detect approximately 85% of the problems present in using a website.) Participants will be recruited by personal invitation from a combination of graduate students, and non-UNB individuals known by the author and her supervisor who will have varying degrees of familiarity with Web-based and/or GIS systems. The users' reaction and operation in the hand-on session and conversation between the author (facilitator) and participants are video/audio-taped. This process is repeated (between five and eight times) until no more new issues are raised in the test. The procedures in Figure 5 will be followed for fixing usability defects that are detected in the individual testing period.

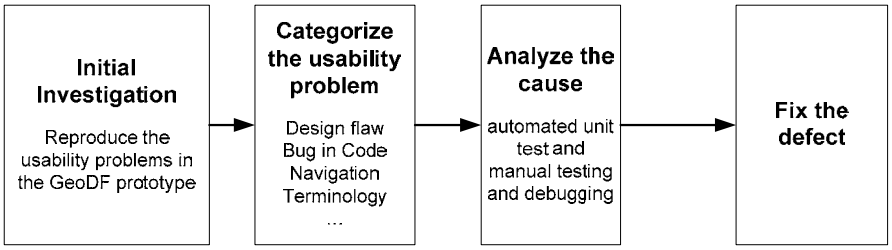

Figure 5 Procedures used to identify and fix the usability defects.

Before starting the second stage, the usability defects will be fixed and modification will also be made to the initial design of the GeoDF prototype according to the feedback in stage one. This improved version, together with the old version of the GeoDF prototype, will be presented to two different sets of participants. Participants involved in the group testing are being recruited using a "snowball[2]" sampling procedure from a combination of graduate student, and non-UNB individuals in order to have various backgrounds. For minimum statistical analysis set, the author will need to recruit two groups of approximately 20 participants each, and then test the improved version against the older version separately in two different workshops.

Usability Results Evaluation

The video and audio tapes of the testing sessions will then be annotated and analyzed to measure how long each task takes for the user to complete, and how many errors the participants make. The evaluation results will be presented in subsequent papers. The users' mental physical and

[2] The "snowball method" refers to that once a contact person has been recruited, s/he will be asked to put us in contact with people who might be interested.

performance demand will be measured subjectively by analyzing the annotated tapes, questionnaires, feedback and comment gathered at the end of each test meeting. In order to identify the most severe usability problems of GeoDF, the problems encountered by each participant during the test meeting will be categorized (See Figure 5). A report summarizing the different categories of usability problems will be created at the end of the first stage.

In the second stage, the author will test the improved version against the older version in two different workshops. The test participants will be asked to fill out the same set of questionnaires. The author will compare the data gathered in these two groups, then measure the change in users' satisfaction and acceptance of the GeoDF prototype quantitatively.

The preliminary evaluation shows that the users are generally satisfied with the functionality. Participants are particularly impressed by the spatially-related discussion that the system supports. Nonetheless, the users' acceptance of PPGIS technology is largely determined by their motivation and other human factors. The preliminary result shows that PPGIS tools can enhance citizens' participation, provided that people are already interested in a given planning issue, and the cost of using such PPGIS tools is relatively low.

CONCLUDING REMARKS

By applying usability analysis methods from HCI for the evaluation of the GeoDF prototype, the authors hope to document how people use and understand a PPGIS tool. This paper focuses on the design of the usability test and metrics for evaluating the usability results. Conducting such an empirical study will provide evidence of how the public would interact with a PPGIS and the public's collaborative behaviour in a group-based environment. This paper also reinforces the link between the research community and the potential users by helping to provide better understanding and cooperation of general public in civic planning and decision-making processes through the usability test. The author is specifically interested in real-world planning practises in conjunction with the local municipality and in the comparison of the usability of different PPGIS tools.

ABOUT THE AUTHORS

Jianfeng Zhao is an international student from China studying M.Sc.E.in Geodesy and Geomatics Engineering at the University of New Brunswick in Canada. Her research focus is on the societal impact of GIS to the general public and usability issue involved in the design and implementation of a Public participation GIS prototype in urban planning procedures.

David J. Coleman is a Professor of Geodesy and Geomatics Engineering and Dean of the Faculty of Engineering at the University of New Brunswick in Canada. Prior to obtaining his Ph.D., he spent 15 years working in project engineering and executive management roles in the Canadian geomatics industry. His research interests deal with policy and economic issues associated with spatial data infrastructure, as well as engineering operations issues associated with spatial data access and management in Web-based environments.

Contact Information:
Department of Geodesy and Geomatics Engineering
University of New Brunswick
P.O. Box 4400, Fredericton, N.B.
Canada, E3B 5A3
Email: j.zhao@unb.ca, dcoleman@unb.ca

ACKNOWLEDGEMENT

The authors would like to thank Dr. Joanna Lumsden of the National Research Council of Canada's Institute for Information Technology for her advice on design of user testing procedures. The advice and cooperation offered by Rob Lunn and Felix McCarthy from the City of Fredericton, and by Eddie Oldfield and Xiaolun Yi of the New Brunswick Lung Association has also been invaluable. Finally, this research is being carried out thanks to the financial support of the Natural Sciences and Engineering Research Council of Canada and the New Brunswick Innovation Foundation.

References

Butler, K. A., 1996, Usability Engineering Turns 10. Interactions, 3(1), 58–75.

Craig, W., T. Harris, and D. Weiner (Eds.), 2002, Community Participation and Geographic Information Systems (London: Taylor and Francis).

Davis, F. D., 1989, Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319-339.

Demšar, U., 2007, Combining Formal and Exploratory Methods for Evaluation of an Exploratory Geovisualization Application in a Low-Cost Usability Experiment. Cartography and Geographic Information Science, 34(1), 29-45(17).

Evans, A., R. Kingston, S. Carver, and I. Turton, 1999, Web-based GIS Used to Enhance Public Democratic Involvement, GeoComputation ’99 conference, July 1999 .
http://www.geog.leeds.ac.uk/papers/99-1/.

Haklay, M., and C. Tobon, 2003, Usability evaluation and PPGIS: towards a user-centred design approach, International Journal of Geographical Information Science, 17(6), 577–592.

Keßler, C., C. Rinner; M. Raubal, 2005, An Argumentation Map Prototype to Support Decision-Making in Spatial Planning, Conference on Association of Geographic Information Laboratories for Europe May, 2005, http://ifgi.uni-muenster.de/~lette/AGILE_2005.pdf.

Kingston, R., 2002, Web Based PPGIS in the United Kingdom. In Craig, W., T. Harris, and D. Weiner (Eds.), Community Participation and Geographic Information Systems (London: Taylor and Francis), 127-138.

Morris, M.G. and Dillon, A., 1997, How user perceptions influence software use, Software, IEEE,14(4), 58-65

Nielsen, J., 1993, Usability Engineering (San Diego: Morgan Kaufmann).

Sidlar, C. and C. Rinner, 2006, Analyzing the Usability of an Argumentation Map as a Participatory Spatial Decision Support Tool, URISA Journal, http://www.urisa.org/Sidlar

Peng, Z. R. and M. H. Tsou, 2003. Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks (NJ: Wiley).

Preece, J., Y. Rogers, and H. Sharp, 2002, Interaction design: Beyond human–computer interaction (New York: John Wiley and Sons, Inc).

Steinmann, R., A. Krek, and T. Blaschke, 2004, Analysis of Online Public Participatory GIS Applications with Respect to the Differences between the US and Europe, 24th Urban Data Management Symposium, http://map3.salzburgresearch.at/docs/udms%20paper_2004_last.pdf.

Tang, Teresa and J. Zhao and D. J. Coleman, 2005, Design of a GIS-enabled Online Discussion Forum for Participatory Planning". Proceedings of the 4th Annual Public Participation GIS Conference, August 2005 (CD-ROM).

Tang, M. Y. T., 2006, Design of a GIS-Enabled Online Discussion Forum for Participatory Planning. Unpublished M.Sc.E. thesis, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, N.B., Canada, 163 pp.

Voss A., I. Denisovich, P. Gatalsky, K. Gavouchidis, A. Klotz, S. Roeder and H. Voss, 2004, Evolution of a participatory GIS, Computers, Environment and Urban Systems, 28(6), 635-65.

Wong, S. and Y.L. Chua, 2001, Data Intermediation and Beyond: Issues for Web-based PPGIS, Cartographica, 38(3&4).