A Usability Framework for the Design and Evaluation of an Exploratory Geovisualization Environment

Etien L. Koua; Menno-Jan Kraak
International Institute for Geoinformation Science and Earth Observation (ITC), P.O. Box 6
7500 AA Enschede, The Netherlands
Email: koua@itc.nl; kraak@itc.nl

Abstract

The exploration of large geospatial data for finding patterns and understanding underlying processes is one of the challenges in geovisualization research. New methods are needed for effective extraction of patterns and appropriate visualization tools are necessary to support knowledge construction throughout the exploration process. Based on an approach to combine visual and computational methods, a visualization environment has been developed to support visual data mining and knowledge discovery tasks. This environment integrates non-geographic information spaces with maps and other graphics that allow users to explore patterns and attribute relationships. The development of the tool intends to facilitate knowledge construction using a number of steps that underline data mining and knowledge discovery methodology. In order to investigate the effectiveness of the design concept, an empirical usability testing is planed to assess the tool's ability to meet user performance and satisfaction. In this test, different options of map-based and interactive visualizations of the output of a Self-Organizing Map (SOM) are used to explore a socio-demographic dataset. The study emphasizes the knowledge discovery process based on exploratory tasks and visualization operations. This paper describes the usability framework used to guide the design, and examines key aspects of the evaluation of such visual-computational environment.

***Keywords**: Exploratory visualization, Visual data mining, Knowledge discovery, Self-Organizing Map.*

1. Introduction

Information visualization techniques are increasingly used in combination with exploratory data analysis techniques to explore large geospatial data. The integration of feature extraction tools with appropriate user interfaces, is important to support user's understanding of underlying structures and processes in the data. However the design of such interactive representation forms still lacks a delineation of fundamental operations that user might apply to an interactive map or related graphic, and guidelines for their appropriate application [1]. Design and usability issues for such integrated geovisualization tools are an important step to their success.

Usability of GIS products and specifically geovisualization tools has received considerable attention in recent years. The need to assess the usefulness and usability of geovisualization tools is increasing as new types of interactions emerge [9]. This interest has recently linked the HCI field, cognitive science and information science in a number of studies [4-7]. There is however a lack of evaluation methodologies and particularly task specifications for user based testing in exploratory geovisualization tool [8]. The map use studies [10] usually conducted in the field of cartography are not necessary fully applicable in new interactive visualizations that involve new representational spaces and advanced user interfaces. Information visualization techniques are often applied in dynamic and interactive maps designed in cartography [11]. These interactive visual geospatial displays allow to explore data, generate hypotheses, develop problem solutions and construct knowledge [12]. This is by definition what geovisualization is all about; an active process that uses advanced user interfaces to allow users to highlight, filter, and sort data as they search for patterns and relationships.

We present a usability framework for design and evaluation of an exploratory geovisualization environment based on an approach to combine visual and computational methods for knowledge discovery. The design integrates non-geographic information spaces with maps and other graphics that allow exploring patterns and

IEEE COMPUTER SOCIETY

attribute relationships. The tool intends to facilitate knowledge construction using a number of steps provided in data mining and knowledge discovery methodology. The development of the tool is based on the Self-Organizing Map (SOM) neural network algorithm, and relates data mining and knowledge discovery methods for the extraction of patterns. Some graphical representations are used to portray extracted patterns in a visual form in order to support the understanding of the structures and the geographical processes. In order to investigate the effectiveness of the design concept, an empirical usability testing is planed to assess the tool's ability to meet user performance and satisfaction. The evaluation involves different options of map-based and interactive visualizations of the output of the SOM for the exploration of a socio-demographic dataset. The evaluation study emphasizes the knowledge discovery process based on exploratory tasks and visualization operations. In this paper, we describe the usability framework for design as well as the evaluation methodology structured to investigate the effectiveness of the design concept, by examining visual exploration tasks. The methodology is based on an understanding of a number of knowledge discovery activities, visualization operations and a number of steps in computational analysis used to visualize patterns in the data.

2. A framework to support exploratory visualization

The proposed framework explores ways to effectively extract patterns in the data using data mining techniques and to represent the results using graphical representations for visual exploration. This framework is based on current understanding of effective application of visual variables for cartographic and information design, developing theories of interface metaphors for geospatial information displays, and previous empirical studies of map and information visualization effectiveness [7]. The framework guided initial design decisions presented here and is used to structure subsequent user studies. In the next sub section, we outline the main components of the approach including a description of the computational analysis step (data mining), and an examination of tasks and operations that can suggest requirements for design.

2.1. Visual data mining and knowledge discovery for understanding geographical processes

One approach to analysis of large amount of data is by using data mining and knowledge discovery methods. The algorithms used in data mining are often integrated into Knowledge Discovery in Databases (KDD), a larger framework that aims at finding new knowledge from large databases. Recent effort in data mining and knowledge discovery in databases (KDD) has provided a window for geographic knowledge discovery. This framework has been used in geospatial data exploration [13-17] to discover and understand structures and patterns in complex geographical data. The main goal of data mining is identifying valid, novel, potentially useful and ultimately understanding patterns in data [18]. A number of applications of data mining in spatial and spatio-temporal research was summarized in a survey by Roddick and Lees [19]. The promises inherent in the development of data mining and knowledge discovery processes for geospatial analysis include the ability to yield unexpected correlation and causal relationships.

Since the dimensionality of the dataset is very high, it is often ineffective to work in such high dimension space to search for patterns. Because the SOM adapts its internal structures to structural properties of the multidimensional input such as regularities, similarities, and frequencies, it can be used to search for structures in the multidimensional input. We use the SOM algorithm as a data mining tool to project input data into an alternative measurement space based on similarities and relationships in the input data that can aid the search for patterns. It becomes possible to achieve better results in such similarity space rather than the original attribute space [20]. Graphical representations are then used to enable visual data exploration allowing the user to get insight into the data, evaluate, filter, and map outputs. This allows several variables and their interactions to be inspected simultaneously, and receive feedback from the knowledge discovery process [21] by means of interaction techniques that support the process.

2.2. Integrating computational analysis and visualization

The first level of the computation analysis described above provides a mechanism for extracting patterns from the data. The output of this process is depicted using graphical representations (information spaces) to facilitate human perception and cognitive processes [10, 22], by offering visualizations of the general structure of the dataset (clustering), as well as the exploration of relationships among attributes. The design of the visualization environment incorporates several graphical representations that provide ways for representing similarity (patterns), relationships, including a distance matrix representation, 2D and 3D projections, 2D and 3D surfaces, and component planes visualization. They highlight different characteristics of the computational solution and integrate them with other graphics into multiple views to allow brushing, linking, zooming, and rotation for exploratory analysis and knowledge discovery purposes, and to enhance exploration (see figure 1). These multiple views are used to present interactions between several variables over the space of the SOM, maps and

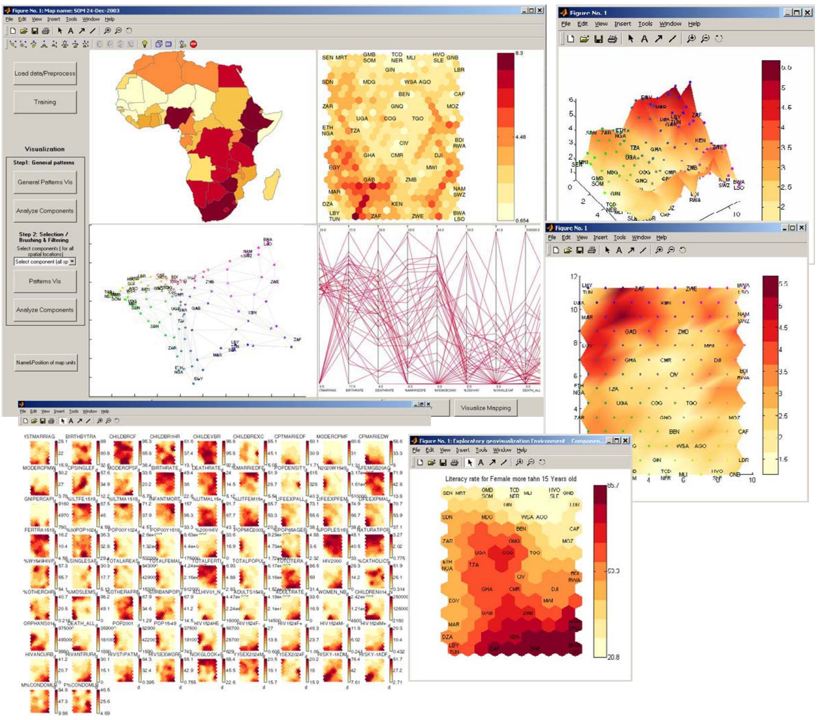

Figure 1. The user interface for the exploratory geovisualization environment in multiple views with the visualization of component planes (bottom left) and map unit labels (bottom right)

parallel coordinate plots, and to emphasize visual change detection and the monitoring of the variability through the attribute space. These alternative and different views on the data can help stimulate the visual thinking process, which is characteristic for visual exploration. The component planes show the values of the map elements for different attributes and how each input vector varies over the space of the SOM units. Comparatively with the maps, patterns and relationships among all the attributes can be easily examined in a single view. Since the SOM represents the spatial clustering of the multivariate attributes, the visual representation becomes more accessible and easy for exploratory analyses to help in identifying causes and correlates. Such map overlays have been important hypothesis-generating tools in research and policy-making [23]. We base the exploratory visualization and knowledge discovery process on four aspects:

- Patterns discovery (through similarity representations)
- Correlations and relationships exploration for hypothesis generation
- Exploration of the distribution in the dataset, on the map
- And the detection of irregularities in the data.

2.3. General requirements for geovisualization design

Based on the analytical aspect of spatial analysis, some general requirements were identified for the design such geovisualization. They include accuracy in results, flexibility in use (scaling, rotation, querying, brushing, browsing, effective navigational tools), adaptability (appropriate for the task and applicable for different situations), exploratory, multiple (alternative) views to consolidate knowledge construction (with maps and other graphics). Other issues such as detection of irregularities (unusual and predictable behavior), and knowledge discovery can also be included.

3. Usability evaluation methodology for geovisualization

Usability evaluation is central in Human-Computer Interaction (HCI), to ensure that the design of a user interface meets the user requirements. HCI has a strong emphasis on User-Centered Design, a design approach that views knowledge about users and their involvement in the design as a central concern, and includes users in testing and evaluations in an interactive design process. In new designs of interfaces for geovisualization, this link between usability testing and user-centered design is becoming more prominent [5-7]. Choosing a usability evaluation method requires consideration of methodology issues and the objectives of the evaluation. In early design stages, inspection techniques such as cognitive walkthrough [24], and usability review or heuristic evaluation [25, 26] are often used for identifying usability issues, for the validation of design decisions, and to get feedback on key aspects of the functionality, interface, and the overall navigation. Given the exploratory nature of geovisualization environments, particularly the visual-computational environment for which this assessment methodology is developed, a user-based evaluation (usability testing) is certainly the most suitable approach to assess usability and usefulness. Usability testing involves assessing the tool's ability to meet user performance and satisfaction objectives, and is conducted based on a number of representative user tasks, for which a certain number of usability factors are measured. Since the design of the tool is based on a user-centered approach early involvement of users was made in a preliminary interface feature inspection [7], in which several aspects of the representation forms, graphics, and color schemes, were presented to users for analysis.

3.1. Conceptual goals, exploration tasks and visualization operators

One way to examine exploration and knowledge discovery support in the visualization environment is by assessing user performance for a number of defined tasks and goals. The main goal of geospatial data analysis is to find patterns and relationships in the data that can help answer questions about a geographic phenomenon or process. The geographical analysis process can be viewed as a set of tasks and operations, needed to meet the goals of the data exploration. Examples of basic tasks in this process include examining spatial positioning of elements of interest, in order to verify spatial proximity amongst different elements, verifying their spatial density, and obtaining an overview of how a target value measured at one particular spatial location, or at various neighboring locations, varies for different attributes. These tasks involve a number of activities and operations that users will perform during the exploration process described in figure 2. Several authors have suggested taxonomies of visualization operations [2, 3, 27, 28]. The most comprehensive list [2, 29] includes: identify, locate, distinguish, categorize, cluster, distribution, rank, compare, associate, and correlate. A delineation of some of these operations for visualization and analysis of spatial data was provided by [3], and include selection, association, and grouping. From these taxonomies of visualization goals described above, three key exploratory tasks for knowledge construction can be identified:

1. *Categorize and classify*: users must be aware of the different clusters that were found in the data. The different clusters can be viewed in different perspectives, 2D and 3D space, rotation, …
2. *Compare*: Users can categorize and review relationships, perceive commonalities and distinctions.
3. *Reflect* (evaluate, integrate, generalize): after completing most activities, users can reflect on the patterns they observe, and the general rules that can be constructed.

3.2. Defining user tasks for usability evaluation

The evaluation of the graphical representations and interfaces needs to be grounded in a task model that can focus more on the user's goals and the tasks he needs to perform, than the interface side. The task model intends to support the development of the experimental setup for the evaluation to cover the different levels of analysis included in the use of visualization tools. A task be can be seen as a sequence of necessary steps and is comprised of objectives, the definition of the problem, and methods necessary for the resolution of the problem. The conceptual goals, the different steps of the exploration and knowledge discovery process, and the key visualization tasks and operators described above, are used as the basis for defining a low level taxonomy of operational tasks that users need to perform to meet the conceptual goals. This is achieved by a decomposition of the basic visualization operators: analyzing task structures of real world visualization problems, representing the collection of subtasks, corresponding taxonomic or classification as well as a set of semantic relationships among these concepts, and other entities necessary to perform the task. A set of representative tasks are derived and structured in a task scenario for the evaluation study. Task scenarios provide a task-oriented perspective on the interface and represent a structure and flow of goals and actions that participants are supposed to evaluate. Such scenarios ensure that certain interface features are evaluated. Our view of exploration is based on the fact that users have initial access to the global views or results from the data mining process. The tasks of exploration are then based on the following goals: selection of objectives, data, views and tools for completing the required tasks, and finally interpretation and evaluation of the results of the exploration.

3.3. Evaluation plan

The usability evaluation proposed here is goal-oriented. Appropriate usability indicators are drawn from the goals and corresponding measurements are set from the goals described above. The objectives of the usability evaluation need to be clearly defined as well methods that are best suited for capturing the necessary data. The evaluation intends to assess the visualization tool's ability to meet user performance and satisfaction, with regards to the general goal of exploring patterns and relationships in data. For example the percentage of users that will be able to complete representative tasks within a certain time or

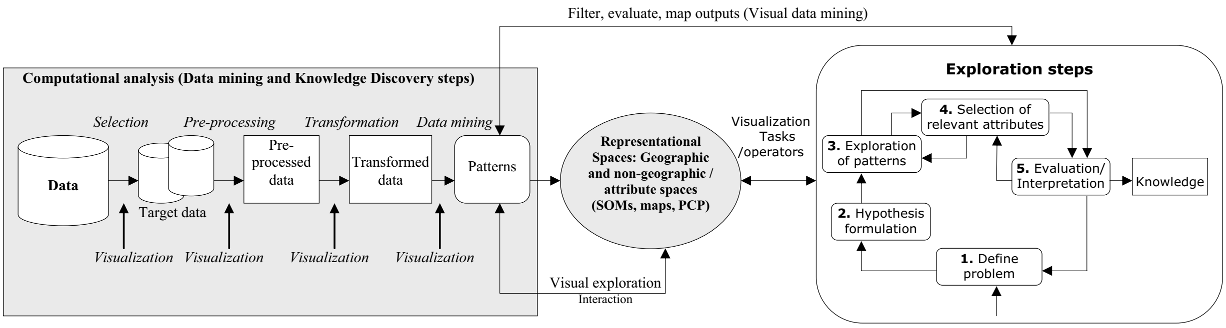

Figure 2. Data mining, exploratory visualization and knowledge discovery processes. The first part of this process consists of the general data mining and knowledge discovery steps (computational analysis). Each of the steps of the computational analysis can allow for visualization. Patterns extracted as a result of the computational process can be explored using graphical representations (geographic and non-geographic information spaces).

without requiring assistance, or the percentage of users that will be satisfied with the usability of the tool. The evaluation is designed to assess the tool's functionality, flexibility for exploratory tasks, and most importantly the ability to support knowledge construction process. Participants are selected to represent the target population of people who are the likely users. The user group includes geographers, demographers, environmental scientists, epidemiologists, and others who have experience in data analysis and the use of GIS. They are domain specialists who have knowledge about the data and have both the motivation and qualifications to do proper interpretation of the analysis. Test participants are involved in the evaluation of the different graphical representations in terms of specific end-user tasks related to the operational visualization tasks described above. They are encouraged to interact with the interface. While completing a number of tasks, users are asked to visually examine the representations, respond to questions, and report their preferences and viewpoints about the representation forms. The objective is to measure and compare their performance scores and the user's level of understanding for the different representation forms and tasks. The kind of evidence the evaluation intends to provide includes responses to specific questions focused on use of features and representations to perform specific tasks, and how users interpret and understand and use the basic visualization features and representation forms. The proposed assessment methodology includes three criteria:

1. *Effectiveness* focuses on the tool functionality and examines the user's performance for the tasks. This can be measured by the time spent for completing tasks, percentage of completed tasks [30-32], correctness of outcome of task performance and response, the success and accuracy (error rate and error types), duration of time spent for help and questions, frequency of documents access, range of function used and the level of success, the ease of use or level of difficulty, level of user guidance and support.
2. *Usefulness* refers to appropriateness of the tool's functionality and assesses whether the tool meets the needs and requirements of the users when carrying tasks, the extend to which users view the tools as supportive for their goals and tasks, user's level of understanding and interpretation of the tool's results and processes. It includes, flexibility, compatibility in relation to user's expectations. This is gathered through task performance, verbal protocols, post-hoc comments and responses on questionnaire.
3. *User reactions* refer to user's attitude, opinions, subjective views, and preferences. It can be measured using questionnaires and survey responses, comments from interviews and ratings.

4. Conclusion

In this paper we have presented an evaluation strategy for assessing the usability and usefulness of a visual-computational analysis environment designed to contribute to the analysis of large volumes of geospatial data. The design of the tool focuses on the effective application of computational algorithms to extract patterns and relationships in geospatial data, and visual representation of derived information to facilitate knowledge construction. The evaluation method emphasizes exploratory tasks and knowledge discovery support. New representation forms used to visualize geospatial data such as the Self-Organizing Map use new alternative techniques to represent the attribute spaces. An important step in the design of such visualization tools will rely on understanding the way users make interpretations of the information spaces. The choice for a representation metaphor is crucial for successful use of the tool. The link between the attribute space visualization tools and maps, in multiple views can provide multiple perspectives for exploration, evaluation and interpretation of patterns and ultimately support for knowledge construction. The methodology presented here will be used for the usability testing of the visual-computational environment, to examine the effectiveness of the representations for exploratory tasks and knowledge discovery support.

5. References

1. MacEachren, A.M., *An evolving cognitive-semiotic approach to geographic visualization and knowledge construction.* Information Design Journal, 2000. **10**(1): p. 26–36.
2. Keller, P. and M. Keller, *Visual clues: Practical Data Visualization*. 1992, Los Alamitos, CA: IEEE Computer Scociety Press.
3. Qian, L., M. Wachowicz, D.J. Peuquet, and A.M. MacEachren. *Delineating operations for visualization and analysis of space-time data in GIS*. in *GIS / LIS*. 1997. Cincinnati.
4. MacEachren, A.M. and M.J. Kraak, *Research challenges in geovisualization.* Cartography and Geoinformation science, 2001. **28**(1).
5. Haklay, M. and C. Tobon, *Usability evaluation and PPGIS: toward a user-centered design approach.* International Journal of Geographical Information Science, 2003. **17**(6): p. 577-592.
6. Fuhrmann, S., P. Ahonen-Rainio, R. Edsall, S.I. Fabricant, E.L. Koua, C. Tolon, C. Ware, and S. Wilson, *Making useful and useable Geovisualization: design and evaluation issues*, in *Exploring Geovisualization*, J. Dykes, A.M. MacEachren, and M.J. Kraak, Editors. 2004, Elsevier: Amsterdam.
7. Koua, E.L. and M.J. Kraak, *Evaluating Self-organizing Maps for Geovisualization*, in *Exploring Geovisualization*, J. Dykes, A.M. MacEachren, and M.J. Kraak, Editors. 2004, Elsevier: Amsterdam.
8. Slocum, T.A., C. Blok, B. Jiang, A. Koussoulakou, D.R. Montello, S. Fuhrmann, and N.R. Hedley, *Cognitive and usability issues in geovisualization: a research agenda.* Cartography and Geographic Information Science, 2001. **28**(1): p. 61-76.
9. Muntz, R.R., T. Barclay, J. Dozier, C. Faloutsos, A.M. MacEachren, J.L. Martin, C.M. Pancake, and M. Satyanarayanan, *IT road map to a Geospatial Future, report of the Committee on Intersections Between Geospatial Information and Information Technology*. 2003, National Academics Press: Washington, D C.
10. MacEachren, A.M., *How maps work: representation, visualization, and design*. 1995, New York: The Guilford Press.
11. Kraak, M.J., *About maps, cartography, geovisualization and other graphics.* Geoinformatics journal, 2000. **3**.
12. MacEachren, A.M., *Visualization in modern cartography: setting the agenda*, in *Visualization in modern cartography*, D.R.F. Taylor, Editor. 1994, Pergamon: Oxford, UK. p.1-12.
13. Openshaw, S., A. Cross, and M. Charlton, *Building a prototype geographical correlates machine.* International Journal of Geographical Information Systems, 1990. **4**(4): p. 297-312.
14. MacEachren, A.M., M. Wachowicz, R. Edsall, D. Haug, and R. Masters, *Constructing knowledge from multivariate spatiotemporal data: integrating geographical visualization with knowledge discovery in databases methods.* International Journal of Geographical Information Science, 1999. **13**(4): p. 311-334.
15. Miller, H.J. and J. Han, *Geographic data mining and knowledge discovery*. 2001, London: Taylor and Francis.
16. Wachowicz, M., *The role of geographic visualization and knowledge discovery in spatio-temporal modeling.* Publications on Geodesy, 2000. **47**: p. 27-35.
17. Gahegan, M., M. Harrover, T.M. Rhyne, and M. Wachowicz, *The integration of geographic visualization with Databases, Data mining, Knowledge Discovery Construction and Geocomputation.* Cartography and Geographic Information Science, 2001. **28**(1): p. 29-44.
18. Fayyad, U., G. Piatetsky-Shapiro, and P. Smyth, *From data mining to knowledge discovery in databases*, in *Artificial Intelligence Magazine*. 1996. p. 37-54.
19. Roddick, J.F. and B.G. Lees, *Paradigms for Spatial and Spatio-Temporal Data Mining*, in *Geographic Data Mining and Knowledge Discovery*, H.J. Miller and J. Han, Editors. 2001, Taylor and Francis: London. p. 33-49.
20. Strehl, A. and J. Ghosh, *Relationship-based clustering and visualization for multidimensional data mining.* INFOMS Journal on Computing, 2002. **00**(0): p. 1-23.
21. Keim, D.A., *Information Visualization and Visual Data Mining.* IEEE transactions on Visualization and Computer Graphics, 2002. **7**(1): p. 100-107.
22. Card, S.K., J.D. Mackinlay, and B. Shneiderman, *Readings in Information Visualization. Using Vision to Think*. 1999, San Francisco: Morgan Kaufmann Publishers.
23. Croner, C., L. Pickle, D. Wolf, and A. White, *A GIS approach to hypothesis generation in epidemiology*, in *ASPRS/ACSM technical papers*, A.W. Voss, Editor. 1992, ASPRS/ACSM: Washinton, DC. p. 275-283.
24. Polson, P.G., C. Lewis, J. Rieman, and C. Wharton, *Cognitive walkthroughs: A method for theory- based evaluation of user interfaces.* International Journal of Man-Machine Studies, 1992. **36**: p. 741-773.
25. Nielsen, J., *Heuristic evaluation*, in *Usability Inspection Methods*, J. Nielsen and R.L. Mack, Editors. 1994, John Wiley & Sons: New York, NY.
26. Nielsen, J. and R. Molich. *Heuristic evaluation of user interfaces*. in *ACM CHI'90*. 1990. Seattle, WA.
27. Zhou, M.X. and S.K. Feiner. *Visual task characterization for automated visual discourse synthesis*. in *Computer Human Interaction*. 1998. Los Angeles, CA.
28. Ogao, P.J. and M.J. Kraak, *Defining visualization operations for temporal cartographic animation design.* International Journal of Applied Earth Observation and Geoinformation, 2002. **4**: p. 11-22.
29. Wehrend, S. and C. Lewis. *A problem-oriented classification of visualization techniques*. in *IEEE Visualization*. 2000.
30. Sweeney, M., M. Maguire, and B. Shackel, *Evaluating user-computer interaction: a framework.* International Journal of Man-Machine Studies, 1993. **38**.
31. Fabricant, S.I., *Evaluating the usability of the scale metaphor for querying semantic information spaces*, in *Spatial Information Theory: Foundations of Geographic Information Science*, D.R. Montello, Editor. 2001, Springer Verlag: Berlin, Germany. p. 156-171.
32. Rubin, J., *Handbook of usability testing: How to plan, design, and conduct effective tests*. 1994, New York: John Wiley & Sons, Inc.