34

Evaluating Usability of SAGRES Virtual Museum

Considering Ergonomics Aspects and Virtual Guides

Márcia Cristina Moraes, Ana Carolina Bertoletti and Antônio C. Rocha Costa
PPGC/UFRGS, FACIN/PUCRS, Av. Ipiranga, 6681 - prédio 30, POA, RS, Brazil
mmoraes@inf.ufrgs.br
MCT/PUCRS Av. Ipiranga, 6681 - prédio 40, POA, RS, Brazil
ESIN/UCPEL, Félix da Cunha, 412, Pelotas, Brazil

Keywords: home learning, internet, hypermedia, World Wide Web

Abstract: The SAGRES Virtual Museum is an educational environment built on the Web that facilitates the organisation of visits to museums, presenting museums' information bases in a way adapted to the user's characteristics. SAGRES incorporates personal assistants to help the user during the navigation and operation of the system. As an educational environment it is important to evaluate the usability of SAGRES, because products with low usability are generally responsible for frustration, anxiety and low productivity, and they can allow the learner to arrive at mistaken or erroneous conclusions. So we developed some validations. We adopted two methods: tests with users and usability inspection. With these validations we identify some positive and negative characteristics of the system, considering the usability and ergonomics aspects of the interface.

1. INTRODUCTION

Most of the virtual museums on the Web are sites that offer on-line learning resources, inviting the visitor to investigate and to explore the information available. SAGRES is a virtual museum that seeks a cooperation between museums and schools, in order to create a new educational environment that provides continuous education, allowing access to information about the museum to the community in general. SAGRES facilitates the organisation of visits to museums, presenting museums'

D. Watson et al. (eds.), *Networking the Learner*

information bases in a way adapted to the user's characteristics (capacities and preferences). The system determines the group of links appropriate to the users, showing them in a subsequent HTML page (Bertoletti and Costa 1999).

There is a module in SAGRES, the agent's module, responsible for overcoming some limitations of hypermedia systems, such as the lack of initial training and the possibility of the user getting lost during navigation, due to the large amount of links available. This module uses software agents to overcome these limitations. Thus, software agents are used as virtual guides and they have the responsibility to assist the learner during an interaction with the system. These guides are represented by animated characters. According to Hayes-Roth and Doyle (1998) the introduction of characters on the Internet could bring the sensation that people are interacting with 'real characters' in virtual worlds. This sensation is intensified because characters simulate verbal and physical behaviour, similar to the humans, such as happiness, satisfaction and greeting. As people are social beings and interact instinctively with each other, the user's interaction with the interface should become friendlier.

As an educational environment is important to evaluate the usability of SAGRES, because products with low usability are generally responsible for frustration, anxiety and low productivity and they can allow the learner to come to mistaken or erroneous conclusions. This paper presents some validations applied to the SAGRES. According to Nielsen and Mack (1994) it is important to make more than one kind of analysis to validate software. So, we adopted two methods – tests with users and a usability inspection. For the first method, we produced a questionnaire based on the principles of the ISO 9241-10, and for the second we used a ckecklist, ErgoList, developed by the Usability Lab at UFSC in Brazil.

We developed another questionnaire in order to confirm the benefits of personal assistance provided for users and to show that interface personification through animated characters is attractive to users, facilitates the navigation, motivates the accomplishment of tasks and provides a personalised assistance and presentation of information. This questionnaire was based on the criteria presented by Hayes-Roth and Doyle (1998) to validate specifically software agents represented by animated characters.

2. THE SAGRES VIRTUAL MUSEUM

SAGRES is designed for three kinds of users: visitor, teacher and learner. The visitor is responsible for building and managing their visit to the museum. Visitors choose the subjects they are going to study and which

activities they are going to do. The teacher is responsible for building a visit about some subject and makes this available for groups of students. The learner is able to execute a visit previous planned by their teacher.

The agent's module is responsible for generating and managing virtual guides. These guides assist users helping them in navigation and the operation of the system, discarding the need of an initial training and being useful in the analysis and monitoring of the users' actions. The guides are represented by animated characters that use the directed improvisation paradigm (Moraes et al. 1999) to interact with the user. In directed improvisation, animated characters can decide what to do and how to behave during their interaction with the user. To do this they execute instructions that can be in pre-conceived scripts and at the same time they improvise behaviours that are in agreement with the context of the interaction with the user. These behaviours include verbal and physical behaviours, that are similar to those of humans, like vibration and greeting, and can be expressed in different ways (as movement, gestures and conversation). With this representation we intend to provide a friendly interface for the users.

To support the independence between the module agents and SAGRES, the development model was developed in three layers: presentation, business and data. The layers of data and business meet in the server. The data layer stores the SAGRES databases and information related to the agents. The business layer implements the agents' functionality. The presentation layer is located in the client and contains the SAGRES HTML pages and agent's pages. The layers can be seen in Figure 1.

3. CLASSIFYING SAGRES AS EDUCATIONAL SOFTWARE

We found in the literature different forms for classifying educational software. Gamez (1998) presents the following taxonomy.

- *Exercise and Practice*, which has the objective to exercise contents or abilities already known by the learner.
- *Tutorial*, which is responsible for the presentation of contents, using animations, sounds and control of the learner' performance.

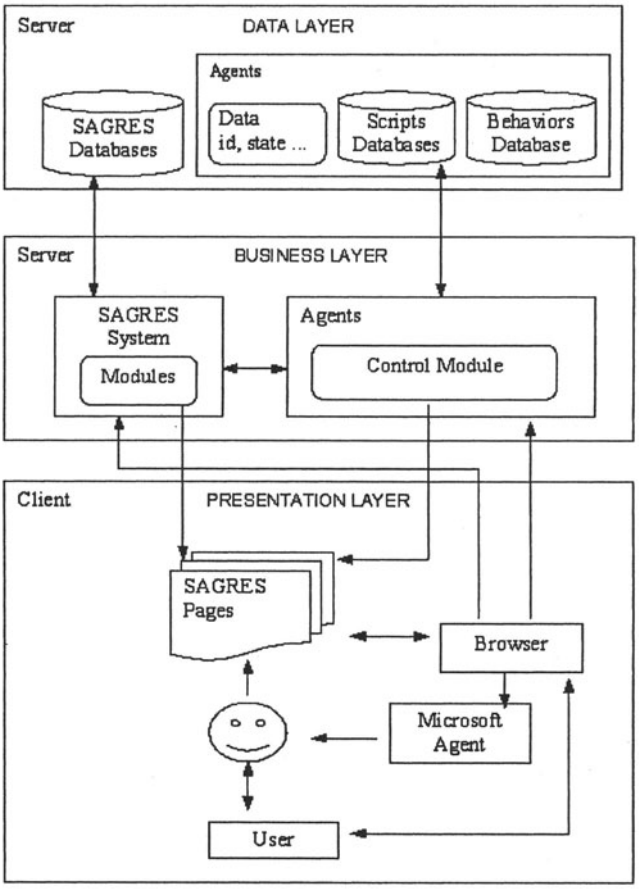

Figure 1. SAGRES Tree Layers Model

- *Intelligent Tutorial System*, which has the objective to consider the knowledge and previous abilities of the learner to choose strategies of learning more adapted for each one.
- *Simulation*, which has the objective to represent or to model parts of the real world as objects, systems and events.
- *Educational Game*, which provides learning trough entertainment;
- *Informative*, which seeks to present information in the form of texts, graphs or tables.

- *Hypertext /Hypermedia*, which is commonly defined as a non lineal form to store and to recover information. The hypertext and the hypermedia can be used as adapted tools to the learning process.

Based on the taxonomy described in Gamez (1998) we can verify that SAGRES can be classified as hypertext/hypermedia, because it is a system developed to the Internet. In this way, it inherits the natural characteristics of the hypermedia. The system also supplies for the teacher, the possibility of elaborating exercises. So, the learner can also do exercises that are additional material. The system also has some characteristics of educational software classified as exercise and practice.

4. USABILITY EVALUATION OF SAGRES AS AN EDUCATIONAL SOFTWARE

For some systems succeed, it is necessary for four characteristics to be present the ability to be useful for a certain public, the ability to be easy to learn, the ability to be easy to manipulate, and the ability to have an attractive interface to get the user's attention. In this sense, the usability evaluation of interfaces is an important stage in the system's life cycle. The main objective of evaluation is to identify problems that can link the user's interaction with the interface. This kind of evaluation is particularly important in the analysis of educational software, because products with low usability are generally responsible for frustration, anxiety and low productivity, and they can allow the learner to come to mistaken or erroneous conclusions.

There are several methods for usability evaluation. These methods are classified as usability inspection methods and empirical tests with the users. Some authors point out that more than one method can be applied to the same interface (Nielsen and Mack 1994, Winckler et al. 2000). We applied two techniques for SAGRES analysis: tests with users through questionnaires, and usability inspection using ErgoList, a specific checklist for ergonomic inspection developed by the Usability Lab at UFSC.

4.1 Tests with users

Medeiros and Cybis (2000) made a study of different ways to do usability evaluation and they verified that one way is through the identification of the users' satisfaction degree, considering the dialogue principles in the part 10 of the ISO 9241 norm.

In order to verify the satisfaction degree and usability of SAGRES, we developed a questionnaire with questions organised into seven categories: suitability for task, self-descriptiveness, controllability, conformity with user expectations, error tolerance, suitability for customisation, suitability for learning (Blanchard 1997). To reduce the time for filling the questionnaire, for each category we defined the maximum of 5 questions. Each question had four options for answers: bad, regular, good and very good, with an associated weighting of one to four.

It is important to note that the objective of the questionnaire was to provide a complete evaluation of the system and each category included certain characteristics of the system. The characteristics related to the virtual guide are present in the suitability for learning and self-descriptiveness. However, this paper also intends to provide a specific evaluation of the virtual guide, in order to verify the learner acceptance and improvements related to its inclusion in the system.

We researched about evaluation approaches for intelligent agents. Hayes-Roth and Doyle (1998) point out that work on intelligent agents inherited the evaluation approaches of Artificial Intelligence and other fields of Computer Science, such as Human-Computer Interfaces. These approaches defined the desirable qualities of systems. However, as virtual guides intend to be more similar to people, traditional approaches need to be adapted or modified. Hayes-Roth and Doyle (1998) present an adaptation of some important categories for the evaluation of animated characters. As virtual guides are represented through animated characters, we could apply these modifications: reliable become variable, predictable becomes idiosyncratic; correct becomes appropriate; complete becomes effective; efficient becomes interesting, and optimal becomes distinctively individual.

Following the ideas presented by Hayes-Roth and Doyle, we also developed a questionnaire to evaluate characteristics related to the agents. Some questions were: Did the guide present appropriated behaviour to each task? Did the guide help you in a personalised way and did he encourage you to do your tasks? Did the guide help you to concentrate on relevant information? Did the information presented by the guide are easy to understand? And, in the future, if you can choose the presentation with or without the guide, which one you choose?

4.2 Usability evaluation

The tool used for the usability evaluation was ErgoList. This tool provides a way to evaluate the facility of use of interactive software. The ergonomic evaluation is accomplished through checklists. Each one is specialised in one aspect or approach that determines the ergonomics of a

human-computer interaction (Project ErgoList 2000). The tool is composed of 18 checklists that seek to evaluate the following characteristics: readiness, grouping for location, grouping for format, feedback, legibility, conciseness, minimum actions, density of information, explicit actions, user's control, flexibility, user's experience, error protection, error message, error correction, consistency, meaning and compatibility. Each one of these checklists comprises of three to twenty-seven subjects, that seek to test the applicability and the conformity, or not, with ergonomic approaches.

4.3 Users identification

The tests were applied to learners from fundamental, medium and superior levels that are partners of the Computer Club at Museum of Sciences and Technology at PUCRS. This selection was made random, and the learners involved are of both genders, with different ability levels and knowledge, all being assiduous users of the system.

4.4 Analysis of the results

The current phase of the research has completed all the usability evaluation tests applied to the visitor module (user's tests and inspection). The usability inspection has been completed for all modules, including teacher and learner modules. And although the tests with users are not fully completed, the data obtained supplies very important information about the usability of the system.

In the data obtained through the tests with users, considering the dialog principles, we verified that some categories have a larger rate: suitability for learning, with average 3,68; self-descriptiveness, with average 3,635; controllability, with average 3,5. Although the other categories have obtained superior average to the good concept, the analysis of the results indicates that some aspects of the system can be improved.

- *Error tolerance*, with average 3,22: the learner should be supplied with immediate feedback. As this system was developed in the Web, the time of request can vary due to the net traffic and the system doesn't inform that a request is being processed.
- *Suitability of task*, with average 3,33: the options vocabulary should be improved, because learner at the fundamental level had difficulties to identify task related to options.
- *Conformity with user expectations*, with average 3,375: a larger information base should be available, including other areas of knowledge.
- *Suitability for customisation*, with average 3,4: it should be clear the available options to adapt the system.

In the analysis of the virtual guide, the average obtained in the questionnaire was 3,68. We verified that users favour the guides' behaviours during the execution of tasks, obtaining an average 3,875. It is important to say that the other aspects obtained averages superior to 3,6. So, the guide successfully reached the approaches proposed by Hayes-Roth and Doyle, displaying normal variability in their choice and manner of executing behaviours, being attractive and interesting to users. Besides, the guide treats the learner in a personalised way, aiding and motivating him to accomplish tasks.

Considering the variables for suitability for learning and self-descriptiveness included the virtual guide reached the most average in previous questionnaire, we can conclude that the learner shows satisfaction with using the guides. This satisfaction is explicit, because all the learner answer 'with' to the question 'In the future, if you can choose the presentation with or without the guide, which one you choose?' In this way we verify that the virtual guide is an important aid to the learner during the operation of the system, executing its main functionality and serving as incentive for future access to the system.

The data obtained through the inspection test, with ErgoList, shows the percentage of conformity, or not, with certain characteristics in the system. Figure 2 shows that most of the applicable approaches obtained percentage of conformity higher than 50%. The best categories are: minimum actions (100%), grouping for location (81,8%), density of information (77,7%) and grouping for format (64,7%). In contrast, error protection and user's experience are the worst categories, because they are not in conformity with ErgoList. This is because certain characteristics are not present in SAGRES such as: confirmation of destructive actions, existence of buttons to undo destructive actions, and existence of dialogues compatible with the users' abilities.

Indeed, several of the approaches are not applicable, because ErgoList is not a checklist developed for Web evaluation. So, the approaches of error correction, user's control, flexibility and explicit actions have questions that do not match with characteristics of Web systems.

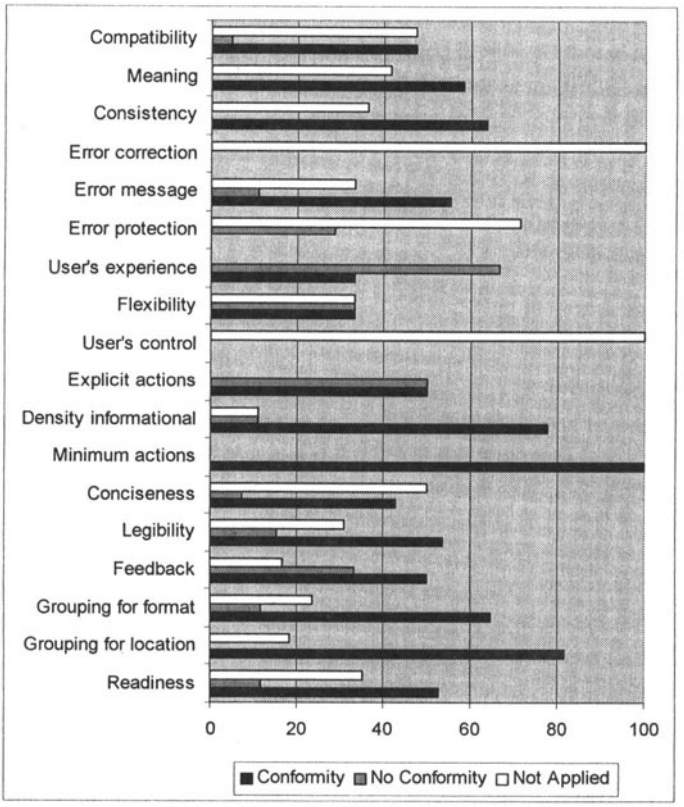

Figure 2. Percentage of Conformity

5. CONCLUSIONS

The evaluation phase presented in this paper presents a usability and practice framework for collecting the degree of users satisfaction, to identify positive and negative points that must be revised and improved, and to validate the satisfaction of users in relation to the virtual guide.

Both methods, tests with users and inspections with ErgoList, show equivalents results, indicating that the ergonomics' aspects of the interface really influence the usability of the system. Considering that error correction, error message and error protection of Ergolist, can be compared to error tolerance in ISO 9241-10, we verify that both present the lowest degree of satisfaction, and must be revised. Additionally, the user's experience of ErgoList, and suitability of tasks of ISO 9241-10, reflect the need to adapt, not only the information presented, but also the dialogues (menus, error

messages and help) within the system. The questionnaire used to probe use of the virtual guide demonstrates that users like the virtual guide. We believe it can facilitate navigation, offering personalised help and demonstrating interesting ways to interact with users, thus achieving its initial goal.

REFERENCES

Bertoletti, A. C. and Costa, A. C. R. (1999) SAGRES – A Virtual Museum. In *Museums and The Web 1999 Conference,* D. Bearman and J. Trant (eds.), New Orleans, Louisiana, March.

Blanchard, H. E. (1997) Standards: User Interface Standards in the ISO Ergonomics Technical Committee. *SIGCHI Bulletin*, vol. 29, n. 1, January.

Gamez, L. (1998) *Ticese: Técnica de Inspeção de Conformidades Ergonômicas em Software Educacional.* MD thesis, Universidade de Minho.

Hayes-Roth, B. and Doyle, P. (1998) *Animate Characters. Autonomous Agents and Multi-Agent Systems.* 1, 195-230, Kluwer Academic Publishers.

Medeiros, M. A. and Cybis, W. A. (2000) Método de Avaliação de usabilidade de *software* a partir da satisfação de usuários e da aplicação de quesitos da norma ISO 9241. In *IHC'2000 – III Workshop sobre Fatores Humanos em Sistemas de Computação,* pp. 93-101.

Moraes, M. C., Bertoletti, A. C. and Costa, A. C. R. (1999) Virtual Guides to Assist Visitors in the SAGRES Virtual Museum. In *International Conference of the Chilean Computer Science Society – SCCC'99,* IEEE Press, Chile.

Mulken, S., André, E. and Müller, J. (1998) The Persona Effect: How Substantial is it? In *XIII Proceedings of the HCI'98,* Berlin, pp. 53-66.

Nielsen, J. and Mack, R. (1994) *Usability Inspection Methods.* John Wiley & Sons, Inc. 413p.

Projeto ErgoList (2000). [www.labiutil.inf.ufsc.br/ergolist/check.htm] 20.10.2000.

Winckler, M., Nemetz, F. and Lima, V. (2000) Interação entre Aprendiz e Computador: Métodos para Desenvolvimento e Avaliação de Interfaces. In *Tecnologia Digital na Educação,* Editora da UFRGS, pp. 7-33.

BIOGRAPHY

Márcia Cristina Moraes, a PhD student in Computer Science at UFRGS, is a professor in the FACIN/PUCRS. Her current research interests are Distributed Artificial Intelligence and Human-Computer Interaction.

Ana Carolina Bertoletti is the Coordinator of Multimedia and Internet Division at MCT/PUCRS. Her current research interests are Computers in Education and Virtual Museum. She has master degree in Computer Science.

Antonio Carlos da Rocha Costa has a Ph.D. in Computer Science and is a lecturer at the ESIN/UCPel. His interest areas are Computers in Education, Distributed Artificial Intelligence, and Theoretical Computer Science.