

Evaluating e-travel sites with the use of design patterns

Petros Georgiakakis¹, Symeon Retalis¹, Yannis Psaromiligkos², Yannis Arapoglou¹

¹University of Piraeus, Department of Technology Education and Digital Systems
80 Karaoli & Dimitriou, 185 34, Piraeus
Tel: 0030 210 414 2746
{geopet, retal}@unipi.gr & johnrapoglou@yahoo.gr

²Technological Education Institute of Piraeus
General Department of Mathematics
Computer Science Laboratory
250, Thivon & P. Ralli, 122 44 Athens, Greece
Tel. 0030 210 5381193, Fax. 0030 210 5381351
jpsa@teipir.gr

Abstract

This paper presents an innovative expert usability evaluation method called DEPTH (usability evaluation method based on DESign PaTterns & Heuristics criteria). DEPTH is a method for performing expert-based scenario driven heuristic usability evaluation of e-sites. DEPTH focuses on the functionality of e-sites and emphasizes on usability characteristics within their context. This can be done not only by examining the availability of functionality of an e-site but also the usability performance of the supported functionality according to specific context related usage scenarios. Scenario driven expert based usability evaluation/inspection is resource demanding. Moreover, it is not easy to find usability experts for specific types/genres of e-sites. Thus, the main motivational idea for proposing DEPTH have been: i) to minimize the preparatory phase of a usability evaluation process by re-using inspection scenarios and ii) to assist a novice usability evaluator (one who is not necessarily familiar with the genre of the e-site) by showing best design practices. Thus, DEPTH proposes the (re)use of expert's knowledge captured in design patterns. This paper briefly describes the DEPTH method and presents findings from applying it to a usability evaluation study of e-travel sites.

Keywords: Heuristic evaluation, design patterns, reuse of design expertise.

1. Introduction

The world wide web has grown tremendously with a vast amount of information and services offered to users. Users have been turned into consumers of services and products from various types of e-sites. However the developers of the e-sites do not

help the users by designing badly. User experiences are not always positive. Usability evaluation though necessary is often neglected. It is estimated that that projects which do not undertake usability studies early in the development process will typically spend between 13% and 46% of their budget on correcting the design at later stages [UsabilityByDesign.COM, (2002)].

Usability evaluation of e-sites is not an easy task and requires a lot of effort [Nielsen, (1998)]. One approach is the use of usability experts which raises the cost for the organization undertaking the task [Dix et. Al. (2003)]. It is often difficult to find a usability expert who will be able to perform his/her tasks and pinpoint a lot of usability problems which stem from the general usability heuristics as well as to successfully determine usability problems which have to do with the specific context of use for the e-site. Not only is it difficult to find usability experts [Nielsen, (2000)], but it is even harder to find type/genre specific e-sites usability experts. Digital genres are described as a classification system for kinds and types of digital products [Schmid-Isler, (2000)]. During the last years several digital genres of e-sites have been studied such as online newspapers, e-shops, e-learning sites, etc.

Thus, a practical approach for solving the problem of finding usability experts for the specific genre of an e-site under evaluation could be to, accurately and efficiently, help a typical novice usability evaluator (NUEs) in performing usability evaluation for that genre of e-sites. This can be achieved by transferring the expert knowledge to the NUEs and guiding them to perform an e-site evaluation with the aid of systematic approaches and supported toolkits.

Such an approach is the DEPTH method (usability evaluation based on DDesign PaTterns and Heuristics criteria). DEPTH is an expert based scenario driven evaluation method. It eliminates the difficulties of expert based evaluation described above and provides an integrated framework where the NUE can find and (re)use expert knowledge for better performing the evaluation tasks of genres of e-sites. The innovative ideas behind the DEPTH approach are: i) the reuse of expert knowledge in the form of design patterns during the evaluation process. A design pattern describes a problem, a solution to it in a particular context, and the benefits or drawbacks from using that solution [Alexander, (1996); Gamma et. Al. (1994)]; ii) the use of scenarios of genres specific of e-sites.

In this paper we describe the DEPTH method for designing and executing usability evaluations of e-sites giving as an exemplar case the evaluation study of the genre of e-travel sites. The rest of the paper is organized as follows: Section 2 describes DEPTH in detail. Section 3 describes the application of DEPTH on several sites of the e-travel genre in order to clarify the main points of the method. Finally, in section 4 we discuss the current status of the method, as well as our future plans.

2. The DEPTH approach

According to DEPTH the evaluation process of an e-site should focus on three dimensions: functionality according to genre, usability performance of the functionality according to context related usage scenarios, and general usability performance according to heuristics criteria.

We present the whole process in Figure 1 using an activity diagram which depicts the general activities and responsibilities of the elements that make up our method. The basic aim of our method is to provide a framework where an evaluator can find and (re)use expert knowledge in order to perform an evaluation that supports the above dimensions.

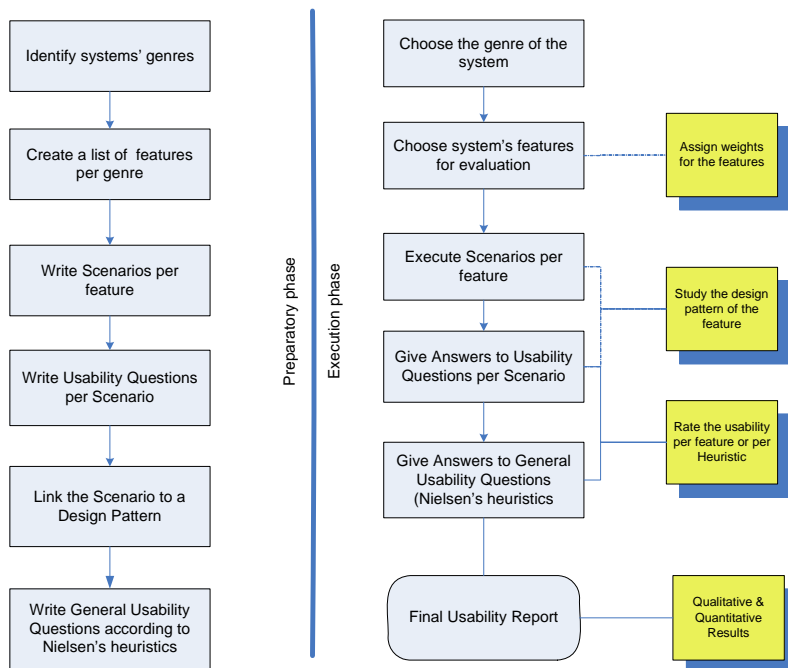


Figure 1. The preparatory and execution phases of the DEPTH method.

The left column in Fig.1 presents the general steps/actions of the evaluator in order to prepare the usability evaluation study according to DEPTH.

The right column in Fig 1, shows the steps need to be made by a NUE (execution phase). According DEPTH a usability evaluation study should start by first selecting the specific genre of the e-site under evaluation. Having as input the analytical table of the functionality of the specific genre of the e-site under evaluation, the evaluator can easily perform the next step, which is a simple checking to ensure whether the e-

site supports the underlying functionality. This step provides the first deliverable of our method which is a functionality report. This report describes the features supported by the selected e-site. The evaluator can also assign weights to the various features selected thus indicating which features are critical according to the specific context of use and need to be well designed.

At the same step the evaluator has to decide which of the supported features will be further inspected. As we have already mentioned, the production of the functionality table alone is not enough for someone to select the right e-site. We may have e-sites of similar genre, like e-travel sites, which may contain the same set of features but vary in usability. In other words, “It is not only the features of the applied technology but especially the way of implementation of the technology”, as [Lehtinen et. Al. (1999)] says for different genre of e-sites.

Moreover, inspecting the usability of the e-site involves two tasks: (a) evaluation in the small, i.e. inspecting the usability at the specific context of use, and (b) evaluation in the large, i.e. evaluating the general usability conformance to well-defined heuristics criteria. For the first task our method suggests the (re)use of domain knowledge through the design patterns and the underlying design pattern languages. Such a language can adopt issues from HCI design patterns since usability is of prime importance, while at the same time will take into account the particularities of the type of genres under evaluation, and so forth [Van Welie et. Al. (2004)]. So, the evaluator, for each specific feature selected, performs walkthrough inspection of the e-site following the tasks of a proposed scenario. The evaluator may also decide to modify the proposed scenario to better serve the needs of the specific context of use for the e-site under evaluation. During the e-site inspection, the evaluator is able to see the ideal design solution as it has been recorded in related pattern(s). By having access to a design pattern related to the specific feature under inspection, the evaluator can complement his/her findings about its usability and he/she can state the identified usability problems as well as suggest possible solutions.

After having finished the inspection of all features of the e-site, the evaluator is called to “evaluate at large”, i.e. to give his/her opinion about the general usability performance of the e-site according to the Nielsen’s heuristic criteria. As a consequence, a final evaluation report is constructed which has two parts: a context specific part and a general part. The first is comprised of the evaluator’s opinion based on the walkthrough inspection of the e-site according to the scenarios given, while the second presents the general usability performance according to the Nielsen’s criteria.

3. Evaluating the DEPTH method

3.1 Scope of evaluation study

In order to evaluate our method we conduct an experiment with non expert usability evaluators. Thirty postgraduate students of our Department after having completed an introductory course on Human - Computer Interaction, have been asked to evaluate e-sites of a specific genre. In order to make an experiment that would be closer to their interests we proposed them to evaluate two travel e-sites. The travel e-sites allow the user to search and book air tickets, hotels, renting a car, etc. Some of the e-sites had been carefully picked up since various problems had been identified during user experiments previously organised by our group.

Students had average knowledge and experience with such a genre of e-sites, and none of them claimed to be expert in using (nor designing) such e-sites. Actually only two of them had used one of the e-sites under evaluation, 23 of them had used less than 3 times an e-site for booking hotels and none of them had used any other feature. Students were quite accustomed to the concept of design patterns. They had already been taught about them. They had also been learned to apply design patterns to e-site development projects.

We used DEPTH only from the novice evaluator's point of view since we wanted to focus in this specific perspective. The main research questions of this evaluation study were: Can the DEPTH method help novice usability evaluators identify usability problems (especially complex ones)? Can the DEPTH method make NUEs improve their ability to propose solutions to the identified usability problems? Is the DEPTH method easy to apply? Does the DEPTH method make the NUEs' evaluation process easier, more flexible and enjoyable? Does the DEPTH method make NUEs feel confident that they performed a good evaluation study? Do the NUEs appreciate the added value of Design Patterns for usability evaluation?

3.2 Evaluation process

According to the preparatory phase of the DEPTH method, several e-sites of the e-travel genre have been thoroughly examined with respect to the features they provide. Thus a superset of their features had been created, as shown in Figure 2.

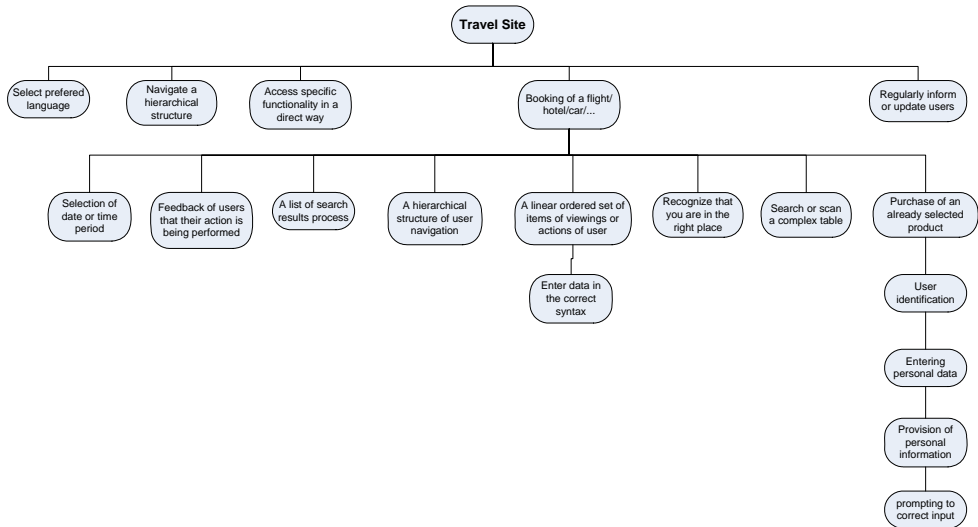


Figure 2. E-travel site functionality

In order to relate the features of the e-travel site to design patterns, we selected design patterns (DP) from Martijn van Welie’s web design patterns repository (<http://www.welie.com/>) as well as from Yahoo’s Design Pattern Library (<http://developer.yahoo.com/ypatterns/>). So, we created a table where each feature is related to one or more design patterns as shown in Table 1.

Table 1. Examples of relations between Features and Design Patterns.

FEATURES (F)	DESIGN PATTERNS (DP)
F1. Purchase of an already selected product	DP1. http://www.welie.com/patterns/showPattern.php?patternID=purchase-process
F2. User identification	DP2. http://www.welie.com/patterns/showPattern.php?patternID=login
F3. A list of search results process	DP3. http://www.welie.com/patterns/showPattern.php?patternID=search-results
F4. A hierarchical structure of user navigation	DP4a. http://developer.yahoo.com/ypatterns/pattern_breadcrumbs.php or DP4b. http://www.welie.com/patterns/showPattern.php?patternID=crumbs

For each of these features we created a related usage scenario. For example for the functionality “F1: Booking of a flight/hotel/car/” the scenario we assigned the scenario “S1: Booking” as shown in Table2.

Table 2. Task Scenario and Questions for a specific functionality.

(S1)Description:	Booking
Task:	You are looking for the price of a ticket in association with date and time factors. Determine the item you wish to book or buy among probably some other options given. Select town or even airport you wish to leave from and going to. If you don't know the names or the abbreviations find out from information given from the site. Choose the departing date, the time, the booking class, and the number of people traveling with you. Because you are looking for the best available price and you don't mind to be precise to the dates you've given choose these dates to be flexible but near to what you want. Submit your choices.
Questions:	<ul style="list-style-type: none"> • Did you easily find the searching area where you can place your options? • In the area where you can choose the departure and the arrival airport or town did you have to write down to the form the name or the abbreviation or you could select from a list given? • If you should write down was there any help to find out the correct abbreviation or spelling? • Were there any error messages in case something wrong has been typed? • In the departing and returning date forms was there the help of a calendar or you should type yourself the exact date? • Could you select flexible dates for example ± 3 days from what you've chosen? • When you selected the persons traveling was it obvious what they meant with the term Adult or Senior, or Children? • Could you search for a one way trip if you wish? • When you submitted your choices for a query at their databases were you informed that some action was taken place? • Did some results come out?

The NUES had to conduct all the tasks of the proposed scenarios while having the ability to look at the related design pattern that the DEPTH method proposes. After having fulfilled the inspection of the e-travel site, they had to express their overall opinion about the e-site according to Nielsen heuristic criteria [Nielsen, (1993)] that are: Visibility of system status, match between system and the real world, user control and freedom, consistency and standards, error prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and minimalist design, help users recognize, diagnose, and recover from errors, help and documentation. Finally a

report in which all the answers of the questions proposed from the scenario and Nielsen's heuristic criteria is automatically generated.

Not only did we analyze the reports written by the NUEs but we conducted focus group interviews (in teams of three students) to get better insight of their opinion about DEPTH. The major advantage of conducting a focus group interview [Krueger, et. Al. (2000)] was the ability to obtain detailed information through group cooperation. Due to space limitations of this paper, we are presenting the key findings in the next section.

3.3 Evaluation findings

Using, throughout this experiment, NUEs helped us verify what we intended to prove, i.e. the DEPTH method can actually enable NUEs perform reliable and high quality evaluations. After having performed the tasks of the proposed inspection scenarios, they were able to identify simple usability problems, while at the same time they were also assisted in identifying complex context specific problems, which could not be easily spotted if scenarios and design pattern have not been given. The NUEs clearly stated that the design patterns helped them in realizing the good design practices concerning the various features of an e-travel site.

As an unexpected outcome, many of the evaluation reports that we received showed that the DPEs were also suggesting solutions for each of the problems identified. We, as reviewers of the experiment, wanted to know where this kind of knowledge came from. When we asked our students how they got these references, all of them mentioned the added value of the design pattern that accompanied each feature. By considering the solution given from the pattern and customizing it to the context of the specific e-site, they were able to offer clear context specific solutions for the usability problems. Among other positive remarks, it was also mentioned that i) DEPTH can be used in evaluating isolated areas of interest by simply choosing only few features and ii) DEPTH helped students learn effectively about the use of patterns for interaction design.

However the method has some limitations. Design patterns are not that many. So it is difficult to find mature pattern languages to support the variety of e-sites genres. That became obvious from the collection of design patterns we proposed as we deliberately chose some that are not pretty matured. Even if we assume that the pattern language is there, pretty matured, will there always be a design pattern to validate all areas of interest in a digital genre? Luckily, there are some guidelines which are written in a form similar design patterns [Shneiderman, (2006)]

Another major issue related to the evaluation of the DEPTH method is the creation of genre dependent scenarios. Who should be the creator of those? Will the scenarios be highly scripted or loosely defined? What will the granularity of each scenario be? The need of experts in the creation of these task scenarios is meaningful. We may

want to define scenarios that are very descriptive, or we may want to use scenarios that are more general. We need to have several scenarios, of different granularities, for each feature and let the user decide between cost and efficiency and choose the one that is most appropriate to the case of study.

4. Conclusions

In this paper we provided an overview of DEPTH, which is an innovative method for performing scenario-based expert heuristic usability evaluation for e-sites.

A similar example of expert evaluation using design patterns has been performed by Van Welie for evaluating the general usability of web museums without user involvement. He analyzed the major sections of the museum websites in order to recognize patterns that had been applied. He commented that the evaluation is easier when the evaluator knows the pattern collection available [Van Welie et. Al. (2004)]. The idea of re-usable scenarios have been proposed by the eLSE method [Lanzilotti, et Al. (2006)].

Our method is innovative since it uses the added value of design patterns in a very systematic way within the usability evaluation process. Our method can be easily used by NUEs. When DEPTH was applied by non usability expert for the evaluation of e-travel sites, the results were satisfactory. The expert knowledge embedded in the form of design patterns and usage scenarios was readily available to the novice evaluators, thus enhancing their inspection methods and improving their judgments about the usability of the functionality tested. As the field of design patterns grows and matures, this method seems very promising and highly applicable. Our next step is to apply this method to other genres, such as learning object repositories and develop a toolkit for better supporting the method.

5. Acknowledgments

This work has been partially funded by «METER» project of the EU project Leonardo da Vinci Community Vocational Training Programme (ref num EL/06/B/F/PP-148206), Commission of the European Communities Directorate-General Education and Culture.

6. References

- Alexander, C. (1996).The Origins of Pattern Theory: the Future of the Theory, And The Generation of a Living World, Keynote speech at the 1996 ACM Conference on Object-Oriented Programs, Systems, Languages and Applications (OOPSLA), retrieved from <http://www.patternlanguage.com/archive/ieee/ieeetext.htm>
- Dix A., Finlay J., Abowd G., Beale R., (2003): Human-Computer Interaction (3rd Edition). Prentice Hall, ISBN 0-13-046109-1

- Gamma, E., Helm, R., Johnson, R., Vlissides, J. (1994). Design Patterns – Elements of reusable object oriented software, Addison –Wesley.
- Georgiakakis P., Tzanavari A., Symeon R., and Psaromiligkos Y., (2005) Evaluation of Web applications Using Design Patterns, INTERACT 2005, Rome Italy, 13 Sept. 2005.
- Krueger, R. A. and Casey M. A. (2000). Focus Groups: A Practical Guide for Applied Research. 3rd Edition. Thousand Oaks, CA: Sage Publications
- Lanzilotti, R., Ardito, C., Costabile, M. F., and De Angeli, A. (2006). eLSE Methodology: a Systematic Approach to the e-Learning Systems Evaluation. Educational Technology & Society, 9 (4), 42-53.
- Lehtinen, E., Hakkarainen, K., Lipponen, L., Rahikainen, M., & Muukkonen, H. (1999). Computer supported collaborative learning: A review of research and development (The J.H.G.I. Giesbers Reports on Education No. 10). Nijmegen, the Netherlands: University of Nijmegen, Department of Educational Sciences.
- Martín G. Y. (2005). Wiki tools for collaborative learning environments, Final project thesis, Telecommunications Engineering, Universidad de Valladolid, August, 2005
- Nielsen, J. (1993). Usability Engineering, Academic Press, London.
- Nielsen, C., (1998). Testing in the field. In: Werner, B. (Ed.), Proceedings of the third Asia Pacific Computer Human Interaction Conference, IEEE Computer Society.
- Nielsen, J. (2000). Designing Web Usability: The Practice of Simplicity, New Riders Publishing, Indianapolis.
- Sartzetaki M., Psaromiligkos Y., Retalis S. and Avgeriou P. (2003). Usability evaluation of e-travel sites based on design patterns and heuristic criteria, 10th International Conference on Human - Computer Interaction, Hraklion Crete, June 22-27, 2003.
- Schmid-Isler, S. (2000). The Language of Digital Genres: a Semiotic Investigation of Style and Iconology on the World Wide Web, Proceedings of the 33rd Hawaii International Conference on System Science, Hawaii, IEEE Press, CD-ROM
- Shneiderman B. (2006), Research-Based Web Design & Usability Guidelines, U.S. Government Printing Office, ISBN 0-16-076270-7
- Stefani, A., Xenos, M. 2001. A model for assessing the quality of e-travel systems: PC-HCI 2001 Conference, 2001.
- Van Welie, M. and Klaassen, B. (2004) ‘Evaluating museum websites using design patterns’, Technical Report: IR-IMSE-001, Available at: <http://www.welie.com/articles/IR-IMSE-001-museum-sites.pdf>.
- Usability by Design.COM (2002), eBusiness survey, March 2002, Available at: http://www.usability.uk.com/resources_surveys.htm