DOCTOR OF SCIENCES

Methodology for automating web usability and accessibility evaluation by guideline

Beirekdar, Abdo

Award date:
2004

Awarding institution:
University of Namur

Facultés Universitaires
Notre-Dame de la Paix

A Methodology for Automating Guideline Review of Web Sites

Abdo Beirekdar

Thesis submitted in fulfillment of the requirements for the degree of Doctor of Sciences
(Computer Science Option)

- August 30th, 2004 -

Director:	Professor M. Noirhomme-Fraiture
Co-director:	Professor J. Vanderdonckt, Université Catholique de Louvain, Belgium
Jury:	Professor F. Bodart
	Professor J.-L. Hainaut (President)
	Professor Ch. Kolski, Université de Valenciennes, France
	Professor Ph. Palanque, Université Paul Sabatier - Toulouse III, France

Institut d'Informatique
NAMUR

Chapter 8
Conclusions and Perspectives

8.1 Contributions of the presented work

This work tried to overcome major shortcomings of existing evaluation tools. As it seemed that the main cause of these shortcomings is the hard coding of evaluation logic inside the evaluation engine, this work proposed a solution to this problem. The solution consists in a methodology based on separating the ergonomic knowledge from the evaluation engine. This separation enables a dynamic and flexible structuring of this knowledge and the updating of its evaluation logic according to the rapid evolution of Web-related technologies and scientific findings in the fields of ergonomics and human factors.

We have presented the three pillars of the methodology:

- A framework for structuring Web guidelines in a consistent and systematic manner, and we underlined the advantages and benefits of this framework.
- The formal language that we defined to formalize the concepts of the framework and to be the basis of an evaluation tool.
- A prototype of the KWARESMI evaluation tool that we realized to validate the proposed guideline structuring and to define the specifications of a final evaluation tool.

This work makes several key contributions to the advancement of automated Web evaluation, including the following:

- It presents a methodology and tools for automated evaluation of Web sites based on a new evaluation approach.
- It presents one of the first attempts to define a XML formal language to structure evaluation-oriented information about Web usability and accessibility guidelines.
- It presents an automated evaluation tool based on the proposed methodology and aimed at extending the existing tools at many levels:
    - In addition to evaluating well established guidelines used by existing tools, we can evaluate private guidelines (ex. guidelines defined by a company for its own style).
    - Evaluation is more efficient because we focus on needed information only.
    - We can conduct contextual interpretation and evaluation.

8.2 Limits of the presented work

We can classify the limits of the presented work according to their level:

- **Limits related to the approach**: the principal limit is the restriction of the approach to the analysis of HTML code. In fact, the defined approach is based explicitly on the semantics of HTML elements (tag, attribute, scope). It is not possible to apply it on guidelines that need other sources of usability data like examining the content of the Web page or the server log files. However, it is possible to adapt the approach for other languages as soon as the analysis concerns a markup language as we saw in chapter 4 about extending it to CSS. In addition, the approach targets the evaluation of single Web pages only. To be applied to site guidelines, we need to modify some concepts like the syntax of evaluation sets or sets elements as we explained in chapter 4.
- **Limits related to the used guidelines**: there is no set of guidelines proposed in the literature that covers a sufficiently wide rang of ergonomic aspects needed for high quality Web sites. For example, WAI and Section508 guidelines deal with accessibility only. The direct consequence of this limit will surely be the limitation of the scope of the evaluation too itself. Another guidelines related limit, that can appear when evaluating guidelines from multiple sources, is the possibility to have semantically similar, identical or even contradicted guidelines. If the tool does not provide any verification mechanism of these relationships, they will influence the tool efficiency.
- **Limits related to the GDL**: even if guidelines cover wide rang, they will be useless if we can not structure them with our GDL. This means that the GDL must be rich enough to enable the structuring of a wide rang of targeted ergonomic aspects. For example, we can not structure guidelines like ″Error messages must not be humorous″ because it is very difficult to provide a function to verify if a piece of text if humorous or not.
- **Limits related to KWARESMI**: even if we provided a rich set of GDL functions to overcome the previous limits, this does not mean that the systematic and automatic implementation of these functions will be easy or feasible. For example, implementing the humor verification function seems to be less useful and justified than a function for evaluating colors visibility because the later will be used in more cases than the humor function. Therefore, we have some restrictions (limits) related to the cost-effectiveness of implementing some functions.

8.3 Perspectives

We spoke all over this dissertation about applying the proposed methodology on Web usability and accessibility guidelines, but it also applies on all properties that can be expressed logically by a GDL expression. For example, we can evaluate properties like:

- Verify that the number of links in a Web page is <=8.
- Verify that all images are in GIF format.
- Verify that forms have a Reset button.

These are design rules rather than usability or accessibility guidelines.

On the other hand, automated evaluation tools proved to be useful, but this utility is more significant if they are coupled with conception environments [Farenc 1997]. In addition, as the main task of our tools is to evaluate guidelines, we need to manipulate guidelines at many levels, and accomplishing all these tasks by one tool is not trivial and could affect the efficacy of the evaluation process, especially if we target a flexible and efficient tool. For these reasons, a possible solution is to integrate the evaluation tool within a global system that provides additional tools to deal with management of guidelines, leaving the only task of using the guidelines to the evaluation tool. In addition to these evaluation-related tools, the system integrates a Web design environment that can communicate with them to conduct evaluation during conception phase.

8.4 The DESTINE project

Figure 8.1 depicts the global architecture of such a system. In fact, this system will be effectively developed as a project funded by the *Région Wallonne*. The project is called DESTINE (**D**ESIGN & **E**VALUATION **S**TUDIO FOR **IN**TENT-BASED **E**RGONOMIC WEB SITES) and it lasts three years (9/2003 – 9/2006).

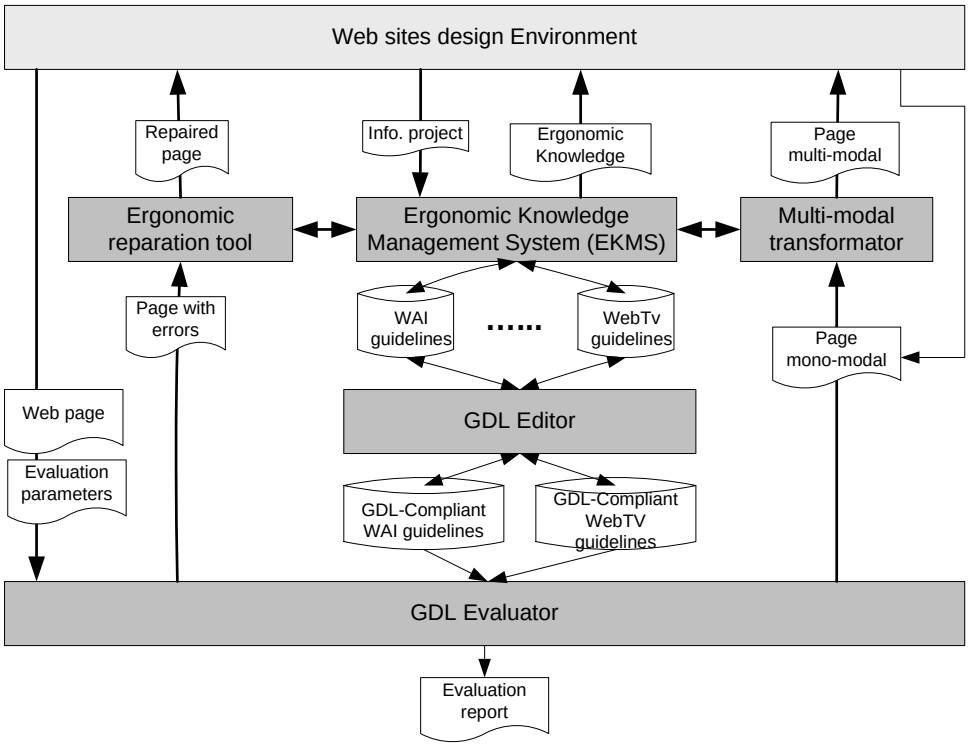

Figure 8.1: Global architecture of the DESTINE Project

DESTINE will be integrated into an environment of Web development by providing the 5 modules of figure 8.1:

1) Ergonomic knowledge management system

This module manages the ergonomic knowledge bases at various levels: creation of a new database of guidelines (W3C's WAI, WebTV, etc.), insertion of the new guidelines in a database, divided and collaborative edition of the existing

guidelines (it will be possible to enrich the database by anyone via the Web), selection of the guidelines corresponding to a given context (targeted user stereotypes, type of site, types of tasks, etc.). in addition to management information about a guideline (source, indexing keys, comments, etc.), a field will contain the specification of this guideline in GDL-compliant form.

2) GDL editor

This tool will make it possible to formally specify a guideline in GDL-compliant form and to store this specification in the database or in an XML file to be exploited later by the evaluation tool.

3) GDL Evaluator

On the basis of some evaluation parameters, this module evaluates the ergonomic quality of a page, a series of pages or a whole site by subjecting it to a set of ergonomic guidelines taken from the databases or XML files. It produces a customizable evaluation report. The pages having ergonomic problems are isolated to be treated by the ergonomic reparation tool. The evaluation tool cannot obviously automate the evaluation of all the guidelines in a complete way (the formal GDL specification will provide necessary information indicating their level of automation: partial, total, with a percentage).

4) An ergonomic reparation tool

This tool repairs any page or series of defective pages where ergonomic problems were identified by the evaluation tool of produces the corrected pages. Reparation can be automatic (the tool corrects all that it can) or assisted (the evaluator controls the correction process).

5) A multi-modal transformer

Even if a site respects some ergonomic guidelines in a given mode of use, it can have lower ergonomic quality in another mode (ex. a graphical interface on a Web navigator). By applying some ergonomic transformation guidelines, this module transforms a page planned for a mode of access into a page accessible by several simultaneous or alternative modes.

We hope that this project will give us the opportunity (time and resources) to overcome the limitations of the current work and to validate our results.