

Towards a Corpus of Validated Web Design Guidelines

Jean Vanderdonckt

▶ **To cite this version:**

Jean Vanderdonckt. Towards a Corpus of Validated Web Design Guidelines. 4th ERCIM Int. Workshop on " User Interfaces for All " UI4All'1998, Oct 1998, Stockholm, Sweden. hal-01319625

HAL Id: hal-01319625
https://hal.science/hal-01319625

Submitted on 21 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

in Proc. of 4th ERCIM Int. Workshop on « User Interfaces for All » (Stockholm, 19-21 octobre 1998.), thème spécial : « Towards an Accessible Web », C. Stephanidis & A. Waern (éds.), ERCIM, 1998, pp. 101–116.

Towards a Corpus of Validated Web Design Guidelines

Jean Vanderdonckt

Louvain School of Management (LSM), Université catholique de Louvain (UCL)
Place des Doyens, 1 – B-1348 Louvain-la-Neuve (Belgium)
Phone : +32-(0)10/47.85.25 – Fax : +32-(0)10/47.83.24
E-mail : jean.vanderdonckt@uclouvain.be, vanderdoncktj@acm.org
URL : http://www.uclouvain.be/jean.vanderdonckt

Abstract. Numerous guidelines are available to support designing accessible and usable web sites. However, most of these guidelines come from multiple guidelines sources with various trust levels so that their application does not guarantee any improvement of neither the accessibility nor the usability of web sites. A corpus of validated web design guidelines is therefore expected to insure such an improvement. To reach a corpus of guidelines designers can trust in, five steps were performed : firstly, guidelines sources were categorised according to a source model; secondly, the most representative and interesting guidelines sources were selected on a basis of selection parameters; thirdly, guidelines belonging to these selected sources were categorised in turn according to a general purpose guideline model and subsequently gathered to form an initial corpus of guidelines; fourthly, particular guidelines were selected by properties to be subject to a dedicated validation; fifthly, these guidelines were empirically validated by analysing the user satisfaction score. A first corpus of validated web design guidelines was consequently divided into three parts : text-only sites, graphical sites, and framed sites. This corpus is itself transformed in a web site with a design relying on its own contents.

1. INTRODUCTION

The incredible expansion of the World Wide Web (web) sites and the fascination they generate throughout the world lead us to consider the accessibility and the usability of web sites for all potential kinds of users. This wish belongs to the “User Interfaces for All” vision where potential users particularly include users who are less well equipped, users living in regions whose geographic situation or social context could be a hindrance, users who are less trained, users with special needs (e.g., people with disabilities) [Stephanidis 98]. These users are confronted with several difficulties such as

- an *appropriation difficulty* : users connecting to a web site for the first time experience many trouble in efficiently visiting the web site (e.g., they have little or no knowledge on how to directly manipulate the windows and other interaction objects, they do not follow any precise navigation strategy); moreover, since they are not accustomed to these novel interaction techniques, their habits are largely overthrown [Akoumianakis 98];
- a *learning difficulty* : design varies from one web site to another, often forcing users to learn a new access method each time a new site is accessed; when this design remains mostly similar in principle, the similarities and discrepancies are long to identify for users;
- an *adaptation difficulty* : some web sites are not malleable enough to be able to afford various interactive situations, various types of users and various cultural or contextual accesses (e.g., users participating to a virtual campus are forced to access the web site from the same workstation because their log-in is not mobile); this lack of malleability can decrease the maximal potential user of web sites by all users;
- an *accessibility difficulty* : users equipped with modest or limited hardware and software cannot access the web site properly or can only partially access the contents [Gappa 97, Lowney 96] (e.g., a user is not able to see a sophisticated graphic animation due to far remote access and very limited, yet expensive, bandwidth, a learner cannot complete an exercise due to an incompatible web browser); user with special needs are not supported

(e.g., visually impaired users are not allowed to user their speech synthetiser or any other transcription material if the web site is not designed for this purpose) [Hadadj 97];

- a *usability difficulty* : the ergonomic quality of web sites highly depends of the designer's openness and worry about human factors and the ability to address them in the site design.

Designers today hold a higher social responsibility in making everything possible not to hinder all these users who cannot correctly access and use web sites due to knowledge, equipment, training limitations. Waiting for the reduction of these limitations, we could consider, study and share resources that contribute to make web sites more accessible and usable. One possible concretisation of these resources regards the manipulation of web design guidelines coming from Human-Computer Interaction (HCI). Because HCI is proved to be useful for improving the usability of traditional interactive applications, we investigated the ability of applying HCI guidelines on web sites. Since these guidelines have been developed with traditional interactive applications in mind, only some of them are relevant in the context of web sites while others are not applicable. So, only the suitable guidelines have to be selected to maximise the impact of web sites usability. In parallel, guidelines have been specifically introduced for web sites (e.g., accessibility guidelines to make a web site more reachable, usable, affordable while still being understandable and intelligible).

Our research aims to develop a corpus of such design guidelines that can be used by any designer to build a user interface for a web site accessible and usable for all. For this purpose, the guidelines need not only be assembled together in an appropriate way but also be validated so that they can be of generic use for user interfaces for all.

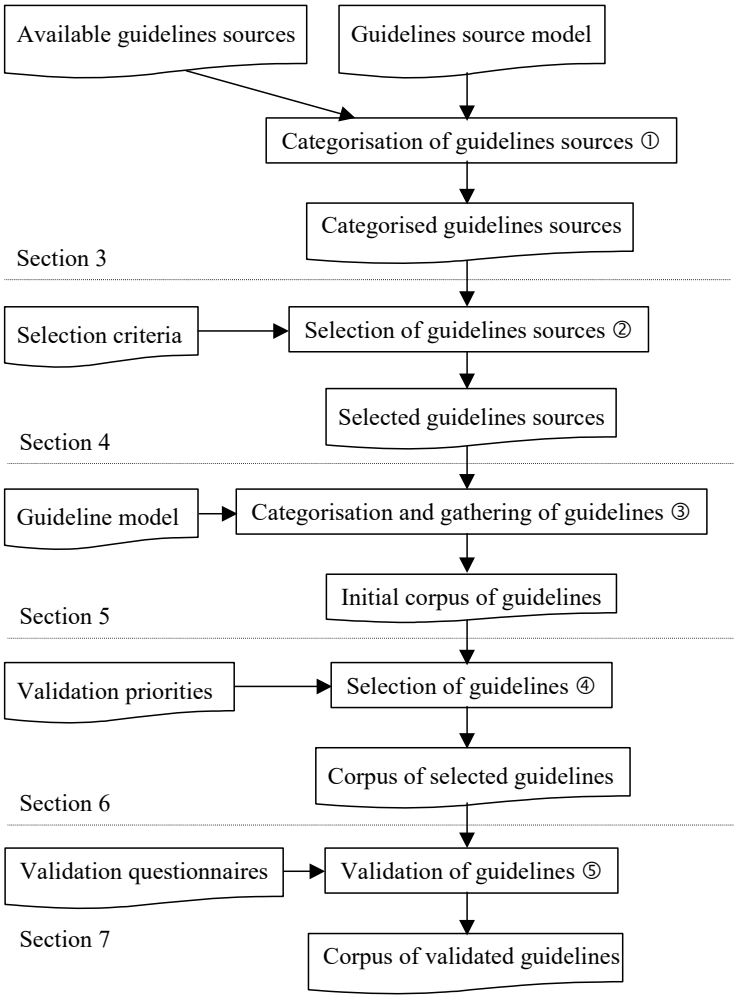


Fig. 1. The five steps of the complete process.

In the next section (section 2), the question whether or not designers feel indeed a need to use a corpus of validated guidelines is raised. This question is further decomposed into an examination of the need for guidelines, the need for organising them into a corpus and the need for validation, which are linked to related work. In the subsequent section, we present the complete process we followed to reach a first version of the intended corpus. This process is composed of five steps which are interdependent (Fig. 1) : only the combination of these steps can result in a successful corpus as described above. Each step is detailed into a single section as depicted in the graphical table of contents of Fig. 1 : the categorisation of guidelines sources (section 3), a selection of guidelines sources (section 4), a categorisation and gathering of guidelines (section 5), a selection of guidelines (section 6) and a validation of these guidelines (section 7). Finally, some conclusions are drawn from this experience to highlight how the place of such a corpus in the web site design process might be determined (section 8).

2. THE NEED FOR A CORPUS OF VALIDATED WEB DESIGN GUIDELINES

2.1. The need for guidelines

Web site design does not assume to use a methodology leading to a unique solution. This design is not indeed an exact process : as the methodology progresses in the development, designers are rapidly faced to a multitude of design options involving many different aspects and allowing multiple solutions for solving a single problem, for deciding a single design option. Because web site design allows alternative ways to decide each design option, there are no strict governing rules that guarantee an accessible and usable web site. The absence of such rules for design makes this task harder for designers. However, the manipulation of guidelines can restrict the scope of these alternative ways and can shed light on the consequences of every decided design option as pointed out by Comber [Comber 95].

Through a reported example of a professionally designed tourist web site for Scotland and other classes of multimedia sites, Johnson [Johnson 98] argues that there is indeed a problem when designers do not manipulate guidelines : several accessibility and usability errors can be detected either manually or automatically. Through several experiments, Borges, Morales & Rodriguez [Borges 96, Borges 97, Rodriguez 97] demonstrated that web site designers could effectively improve the usability by applying seventeen fundamental guidelines. By measuring the accomplishment time for five basic tasks carried out on three sites, with and without guidelines, the authors observed some substantial time reduction (from 16% to 82%).

2.2. The need for a corpus

Guidelines rarely appear in isolation since they can mainly be extracted from five types of guidelines sources [Scapin 90] :

1. *design rules* consist in a series of functional and/or operational rules governing design for a family of web pages. They are written to no longer require any designer interpretation, thus reducing the scope of alternative ways to a unique one. They typically consider the physical page layout, its format, its size;
2. *sets of ergonomic rules* contain recommendations serving for a large spectrum of web sites. Each recommendation is presented with a statement, illustrated with examples, explained with some comments, etc. Each recommendation comes from a human consensus between the different parts using them, this consensus being theoretical, empirical or both;
3. *standards* propose a sequence of functional and/or operational specifications standardising web site design. They are generally promulgated by official organisations (e.g., [W3C 98]);

4. *style guides* prefer ergonomic rules insuring some consistency between a defined set of web pages (e.g., an organisation's web site, a family of web pages for services);
5. *ergonomic algorithms* try to systematise, to operationalise ergonomic rules into software tools that support designers in the design and evaluation of their web sites. For instance, Bobby [CAST 97] automatically evaluates a submitted web page for accessibility guidelines (e.g., "Provide alternative text for all images"), compatibility guidelines (e.g., "NATURAL SIZE FLAG clause in IMG tag is not compatible with HTML 3.2") and for downloading guidelines (e.g., "The total download time of a page should not exceed 20 sec.").

We have argumented the need for a corpus by reporting on ten difficulties; among them, the variety and the number of guidelines sources are probably the most determinant causes for organising guidelines into a comprehensive and unique repository. Ratner, Grose & Forsythe compared 21 so-called web guidelines sources containing a sum of 357 unique guidelines and 5 general guidelines sources containing a sum of 270 unique guidelines [Ratner 96]. They found that only 53 guidelines (20%) are devoted to the web whereas 75% of these guidelines are concentrated in one guidelines source. These figures also show reasons why design guidelines specific to the web require a separate and appropriate organisation. This organisation can be founded on ergonomic criteria as introduced by Bastien & Scapin [Bastien 95]. They have proved that the manipulation of guidelines for traditional interactive applications could benefit from an organising scheme based on seven empirically validated ergonomic criteria. With Leulier [Bastien 96], they extended this proof to the general ISO 9241 standard and are today conducting research to validate a taxonomy of ergonomic criteria for the Web.

2.3. The need for validation

One important step in the complete process of building a corpus as outlined in the previous subsection consists in validating guidelines. Validating a guideline is aimed to quantify, to assess the impact of applying a particular guideline on some accessibility and/or usability aspects for which this guideline is intended. When web sites appeared for the first time, no design guidelines were available at that time; most of them appeared in 1995. Despite that the amount of available design guidelines is still growing, it is time to worry about this validation to contradict the lack of interest in validating guidelines. Since guidelines came from various origins (i.e. person, group, organisation), their scope usually reflects the purpose for which they have been introduced :

- some guidelines are sometimes seen as the expression of a common design practice and the corpus, as a memorandum of these practices ;
- some guidelines have been introduced without proving anything on their effective impact : providing such guidelines and propagating them in guidelines sources induce a dangerous risk : an invalidated guideline may be considered as de facto mandatory after a certain period of time because more and more designers rely on this guideline (guidelines that are not validated can be propagated as quickly as the guidelines that are not validated) ;
- since the vision « User Interfaces for All » encompasses more and more users' categories, assessing previously introduced guidelines with respect to these new categories becomes important ;
- some guidelines are completely contradictory so that designers are faced to a dilemma : they are uncertain about which guideline to apply. For instance, the guideline "Try to keep default link colors" is stated on one hand as "Do not change the default colors for visited and unvisited links. An unvisited link is generally blue and a visited link is purple" and one the other hand as "Visited links should appear in red on your pages, unvisited links should appear in blue" ;

- some guidelines highly rely on the latest hardware and software technology : since both dramatically evolve together during the last years, applying related guidelines remains unpredictable so that imagining that a guideline could remain eternal becomes an illusion. This kind of guideline requires a regular and continuous validation. For instance, the guideline "Try to make the horizontal measurement of information not exceed 480 pixels" is explained with the comment "Most current personal computers monitors display 640x480 pixels on 13- to 15-inch screens. The Macintosh and Windows versions of Netscape and Mosaic both default to a window size that limits horizontal area of web pages to about 500 pixels" ;
- some guidelines may appear arbitrary to the designer's eyes who would like to see a web site certified, authenfied by official guidelines, to highlight a quality label. Without validated guidelines, designers may be confused and no longer encouraged towards this goal ;
- some guidelines are embodied in ergonomic algorithms implemented in software tools that automatically perform syntactic analysis of the web site. If the tools incorporate invalidated guidelines, they could provide a certification based on false guidelines.

For all these reasons, a corpus of guidelines needs to be validated primarily because its recent availability (more or less since 1995) is not always accompanied with such validation results bringing confidence in its contents. A corpus of guidelines needs to be validated long before they broadcast unsure results that could demolish the accessibility and the usability rather than improving them. We now present the complete process we followed towards a first corpus of validated web design guidelines. The sequential presentation of its steps does not necessarily subsume that the steps are performed one after another; instead, they can be partially parallel.

3. CATEGORISATION OF GUIDELINES SOURCES (STEP ①)

Guidelines sources belonging to the five types identified in section 2.2 are really numerous and continuously growing so that it becomes almost impossible to provide a complete listing of them at a given time. Yet, some studies, for instance [Bastien 96, Grose 98], are providing abundant references of guidelines sources that can be merged to references found in general search web sites (e.g., Alta Vista, Yahoo, Lycos) or in dedicated web sites (e.g., Usable Web [UsableWeb 98], Use It [UseIt 98]). The close examination of these sources lead us to conclude that their diversity did not allow almost any type of comparison. To be able to survey the range of these guidelines sources, a domain specific categorisation model has to be applied to each guidelines source. Information that is categorised using such a model can then be retrieved and compared. The contents of this categorisation model can be interpreted as meta-information about the range of guidelines sources. The intended categorisation model results, of course, from analysing existing and foreseen attributes allowing source identification. This model, adapted from previous work, consists of the following attributes :

- title : name of the guidelines source ;
- contents description : summary of the underlying structure, like a table of contents ;
- reference : physical reference of the guideline source (e.g., book, report, URL) ;
- organisation : name and type of the responsible organisation (e.g., a university, a private company, a research centre, a single person, an authoritative committee) ;
- source type : type of the guidelines source as defined in section 2.2 (i.e., design rules, set of ergonomic rules, standards, style guides, ergonomic algorithms) ;
- publication date : date of the publication along with last modification date ;
- elaborateness : approach followed by the responsible organisation to build the source (e.g., survey results, free thinking, personal preference, empirical study, scientific experiments) ;

- guideline type : category to which the guidelines found in the source are related (e.g., site specific, page contents related, navigation specific, intended for graphics, requiring Java applets, exploiting 3D animations) ;
- amount of guidelines : total of guidelines found in the source ;
- guideline example(s) : one or many examples of guidelines found in the guidelines source.

- **Title :** Yale C/AIM Web Style Guide
- **Contents description :** style guide dedicated to web site design structured into a logical way : general philosophy, general user interface design, site design, graphics, multimedia, appendices.
- **Reference :** http://info.med.yale.edu//caim/manual/contents.html
- **Organisation :** Yale-New Haven Medical Center (academic research center)
- **Source type :** style guide
- **Publication date :** January 1997
- **Elaborateness :** elaboration based on long research (since 1991)
- **Guideline type :** site design, page design, navigation, graphics, animations, multimedia
- **Amount of guidelines :** more than 60
- **Guideline example :** *Interlaced GIF*

The conventional (non-interlaced) GIF graphic downloads one line of pixels at a time, and Web viewers like Netscape display each line of the image as it gradually builds on the screen. In interlaced GIF files the image data is stored in a format that allows Netscape (and other viewers that support interlaced GIFs) to begin to build a low-resolution version of the full-sized GIF picture on the screen while the file is still downloading. The "fuzzy-to-sharp" animated effect of interlacing is visually appealing, but the most important benefit of interlacing is that it gives the reader a quick preview of the full area of the picture. This preview effect can be misleading interlaced graphics are not faster-loading than non-interlaced graphics, they just look as if they download faster because the rough preview comes up faster.

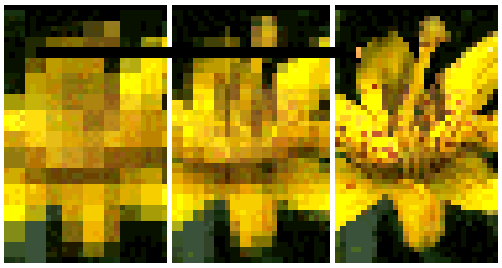


Fig. 2. Example of the categorisation of a guidelines source.

Delving into literature on human factors, guidelines, ergonomic rules, etc, a first set of 30 guidelines sources was considered as vitally relevant because they each target specific aspects of web site design. These 30 guidelines sources were then catagorised according to the above categorisation model. Fig. 2 illustrates the categorisation of the Yale C/AIM Web Style Guide [Yale 97].

4. SELECTION OF GUIDELINES SOURCE (STEP ②)

Once guidelines sources are categorised (section 3), they can be explored equally, their contents and the corresponding quality can be judged according to selection criteria that restrict the range of acceptable guidelines sources. These criteria are expressed in terms of the model attributes :

- the title is only provided as an identifier, thus it does not convey information relevant for selecting a guidelines source;
- the contents description does not always represent the quality of the contents, thus it is not considered;
- the reference wears little or no influence on selecting a guidelines source;
- the organisation name and type are often appropriate indicators of the domain where they work in; we can keep organisation showing experience and competence in the domain of human factors, software ergonomics while forgetting those coming from private or personal initiatives;
- the source type is considered as relevant since a guideline expressed as a design rule does not possess the same scope and confidence as a guideline expressed in a style guide;

- the publication date and the last modification date are generally significant of the attention towards the guidelines source : guidelines sources published or lastly modified before 1995 were removed;
- the elaborateness also indicates a certain trust level : sources based on survey results, empirical studies or scientific experiments guarantee such a level;
- the guideline type often mirrors the main goals of the guidelines source : since we would like to avoid hardware- or technology-oriented guidelines (e.g., for Java applets, for FLC animations), guidelines sources only limited to this type are left out in favour of sources related to high-level design options (e.g., for global site design, page layout, navigation, use of graphics, use of frames);
- while guidelines sources for traditional interactive applications may display a huge collection of guidelines (from 200 up to 2700), guidelines sources for web sites rarely hold more than 30 guidelines. We consider sources including at least 50 guidelines to be significant;
- since copious examples can facilitate the interpretation of applying guidelines, guidelines sources having guidelines without explanation and illustration are rejected.

The criteria for selecting guidelines sources need to be ranked by order of importance. The following order is preferred : the guideline type, the guidelines examples, the elaborateness, the publication/last modification date, the organisation name and type, the amount of guidelines. By applying these selection criteria, the original range of 30 guidelines sources was reduced to 7 :

1. Ameritech Graphical User Interface Standards and Design Guidelines [Detweiler 96],
2. User Interface Design for Sun Microsystem's Internal Web [Levine 96],
3. Apple Web Design Guide [Apple 97],
4. IBM Web Design Guidelines [IBM 97],
5. Yale C/AIM WWW Style Manual [Yale 97],
6. Web Site Usability: Design Guidelines [WebSite 97],
7. A Jakob Nielsen's Alertbox : Guidelines for Multimedia on the Web [Nielsen 95].

It is interesting to note that the main reason for not selecting a guidelines source resides in the lack of information about the poor guideline statement (second criteria) and their elaborateness (third criteria).

5. CATEGORISATION AND GATHERING OF GUIDELINES (STEP ③)

5.1. Categorisation of guidelines

Though selected guidelines sources all satisfy the selection criteria, they do not necessarily present guidelines consistently. Having various guidelines structures, a previously introduced general model for categorising guidelines was tried in order to facilitate their future manipulation. After several trials, the proposed way to transform general, informal, unstructured or differently structured guidelines into uniform, categorised guidelines was proved unsuccessful : even after having selected sources containing extensive guidelines, most attributes remained unfilled due to their poor statement (e.g., linguistic level, ergonomic criteria, usability and utility factors, rank of importance) [Muller 97]. Filling these blank attributes by substituting ourselves to the source's author(s) is a long and hazardous process which is hard to render objective : ignoring the context in which guidelines have been reported can lead us to falsify their expression. Even with this wish in mind, it is almost impossible to complete them uniformly and in accordance with the source goals. Rather than adapting provided guidelines to the general guideline model, a specific-purpose guideline model is adapted to the sum of information available in guidelines. This model is detailed below :

- identifier : number that univocally determines a guideline;
- title : short sentence summarising the main goal of a guideline;
- statement : complete sentence describing the guideline application;
- warrant : explanation of the guideline foundation;
- positive examples : one or many concrete examples where the guideline is respected : these examples are accompanied with techniques and methods that effectively illustrate how the guideline is applied[1] (e.g., HTML pieces of code, files);
- negative examples : one or many concrete examples where the guideline is transgressed : these examples are accompanied with techniques and methods that effectively illustrate how the guideline is not applied[1];
- bibliographic references : list of all guidelines sources where a single guideline is mentioned;
- nature : *syntactic* if the guideline can be applied, tested, verified by an automata (e.g., a software tool) or *semantic* if applying, testing, verifying a guideline only rests on human appreciation (e.g., experts judgement).

- **Identifier : R4.6**
- **Title :** *For graphics incorporating text, the text should be antialiased.*
- **Statement :** Where possible, convert graphics with anti-aliasing software (for instance, Photoshop) to reduce step-like jaggedness often seen in bit-mapped graphics.
- **Warrant :** These "jaggies" are particularly evident and visually distracting on diagonal lines and the edges of curves and circles. Antialiase the text for better viewing.
- **Positive example :**



To see better the difference : http://www.westlake.com/gdclass/6-antialiasing/text-solid.html
In left you will see the letter "a" without the antialiasing and in right, the antialiasing letter.

- **Negative example :**



- **Bibliographic reference :** AMERITECH95 [Detweiler 96]
- **Nature :** semantic

Fig. 3. Example of the categorisation of a guideline.

Fig. 3 depicts an example of a guideline categorised according to this model. All guidelines contained in selected guidelines sources are expressed analogously.

5.2. Gathering of guidelines

Gathering of guidelines basically consists in deciding a starting structure for the corpus and in feeding this structure with categorised guidelines.

5.2.1. Structuring the corpus

Although graphics are important resources in web sites, they can be inaccessible by some web browsers (e.g., Lynx). GVU study claims however that only 1% of web sites visitors are still using them [Georgia 98]. Another accessibility reason for motivating the support of text-only information is that many users with special needs are enabled to exploit text-only information by using adaptation hardware and software (e.g. visual replication for auditory-impaired users or auditive feedback for visually-impaired users).

[1] Such examples along with their techniques and methods are notably rare to find out in guidelines sources. Thus, one important work consisted in looking for examples that are expressive and precise enough to locate the guideline application or unapplication and in documenting them in the guideline model.

A well known accessibility problem is that formatting text into columns may confuse users : the left table of Fig. 4 is translated into a senseless or absurd meaning at the right. The related accessibility guideline here would be : “Provide alternatives to tables and frames”. Tables with transparent borders are a common trick to keep a page layout appearing equally.

There is 30% chance that	Weather temperature
this event	is still increasing
will occur	tomorrow

➔

There is 30% chance that weather temperature this event is still increasing will occur tomorrow

Fig. 4. Example of accessibility problem.

Moreover, graphics embedded in pages slow down the downloading : if this waiting time becomes too important, it is likely that the user will become impatient and zap to another site. 65% of users have a major concern about the downloading time as observed by GVU [Georgia 98]. Some other users ten to prefer to get an information as quick as they can so that graphics become unimportant in search tasks. To support these users, a text-only version of any graphical web site is required. Most style guides also recommend this version. The corpus will therefore be divided into two parts : a part related to building a textual version of a web site and a part related to building a graphical version of a web site. We can conclude that a textual version of a web site is always required and that a graphical version remains optional. The third part of the corpus is intended to deal with frames. Some experts estimate the usability of frames is to be disputed :

- not all browsers support frames,
- frames induce potential navigation problems (e.g., the Back button of Netscape Navigator V2.0 does not work properly);
- when users insert a bookmark for a web page, only the web home page is saved because the bookmark records the URL of all displayed frame (i.e. FRAMESET), but not the URL containing the page for data entry;
- the model for referencing web pages by their URL is no longer valid in catalogues, since incomplete;
- some browsers are unable to print information contained in a specific frame; when they are able, they are unable to print all frames in one user action;
- frames rapidly generate scroll bars when screen display is limited, thus reducing screen real estate.

Other experts believe that framed web sites are allowed provided experienced developers design them. Without closing the debate, we do not want to forbid frames to responsible designers. The corpus is consequently divided into three parts : guidelines for text-only sites, guidelines for graphical sites and guidelines for framed sites. This division seems to be original since it does not appear in any of the seven guidelines sources. This division is motivated by the following reasons :

- guidelines valid for both text-only and graphical versions are duplicated (e.g., R.1.1 “At minimum, it should exist a home page” is duplicated in the graphical version in R.2.1);
- developers designing a particular version can only refer to the needed part (e.g. a text-only site does not need to consider guidelines on graphical version);
- guidelines that are unnecessary for a particular version can be omitted (e.g. the guideline “Provide Alternative Text for Images” is omitted in the text-only version)
- guidelines inducing a potential risk for cognitive overloading are eliminated (e.g., the guideline “If you use images for navigation, make sure you also provide redundant text links” is eliminated);
- guidelines that are common (e.g., for site design, page design, navigation) can be separated from guidelines that are specific (e.g., for graphics, for frames).

5.2.2. Feeding the corpus

To feed the corpus, gathered guidelines are incorporated by performing multiple operations :

- *insertion* : if a guideline does not belong to the current pool, it is added to the relevant sections; if a subsection refines the structure, it is created (e.g., search, user inputs); all model attributes are added whenever possible;
- *concatenation* : if a guideline already exists in the current pool, bibliographic references are updated[2]; each new model attribute missing in the pool is added; each existing model attribute is compared with the current one : if no difference lies between the two values, the existing value is reinforced; if a conflict separates the two values, the difference is reported : other examples are added whenever possible;
- *specialisation* : if a guideline is expressed at a too high level, it might be further decomposed into several individual guidelines (e.g., by concretising a property with multiple cases);
- *generalisation* : if a guideline is expressed at a too low level, it might be partially abstracted into a more general guidelines (e.g., by abstracting one property).

6. SELECTION OF GUIDELINES (STEP ④)

After an initial corpus of guidelines was set up, guidelines were then submitted to a global critique. We were lead to distinguish between guidelines that can be considered as admitted and guidelines that still remain totally open and initiate a scientific debate. It was likely that this approach will generate a certain extent of convergences and divergences, as well as shadow areas in the current research or in our existing scientific and statistic devices. The existence of practically no definitive rules about what should or should not be taken into account for accepting or rejecting a convergence or a divergence made the selection of guidelines a hard task. The absence of consensus on how to approach it made this task even harder, especially when considering the need for validating guidelines as pointed out in subsection 2.3. Therefore, several conventional, thus arbitrary, validation properties were enumerated in order to select guidelines subject to a validation[3] :

- every guideline holding a contradiction needs to be clarified;
- every guideline that is technology-dependent is to be tested;
- every guideline with the warrant attribute unspecified needs to be validated;
- every guideline referring to a well established theory will not be validated[4];
- every guideline based on a proved principle or derived from will not be validated[4];
- every guideline which is estimated to hard or to subjective to be submitted to a validation will be discarded[4].

Fig. 5a shows a guideline based on the strong user preference for fast page access as measured in the GVU study [Georgia 98]. Having this validated, two other guidelines attaching a file type to optimal compression techniques (Fig. 5b and Fig. 5c) might be validated by derivation. All guidelines contained in the initial corpus were then marked as subject to validation or not. Marked guidelines were then selected for validation in the next step.

---

[2] On one hand, some guidelines are typically replicated from one guidelines source to another without explicit link : this situation can so frequently occur that the real origin of a guideline can be lost in a chain of circular quotations. On the other hand, a high replication frequency might sometimes reveal their possible importance.

[3] Performing an exhaustive validation of all guidelines contained in the initial corpus would become a highly-resource consuming activity whose ambition goes far beyond our limited understanding of guidelines scope and our fixed time to devote to validation.

[4] Though such a validation would be interesting.

- **Identification number :** **R4.1.1**
- **Title :** *Use the smallest graphics as possible.*
- **Justification :** More the graphics'size is small and more the time to download these graphics is short.

↓

- **Identification number :** **R4.1**
- **Title :** *Use the GIF format for line art or images with few colours.*
- **Justification :** Not to slow down the downloading.

↓

- **Identification number :** **R4.2**
- **Title :** *Use the JPEG format for complex images such as photographs.*
- **Justification :** Not to slow down the downloading.

Fig. 5. Example of guidelines derivation.

7. VALIDATION OF GUIDELINES (STEP ⑤)

Guidelines selected for validation were finally submitted to a validation method. The results of applying this method are presented and discussed hereafter. Since the main goal of this paper is to present the followed approach towards a corpus of validated web design guidelines, we only outline briefly the main characteristics of the survey in which the validation method was applied. For a detailed presentation and discussion, we refer to [Badot 98].

7.1. Method

A validation method is considered as *theoretical* when the guideline is motivated by theoretical evidence coming from sciences like cognitive psychology, theory of human factors, software ergonomics. For instance, "The right number of different items in the navigational aid is 7±2". Since sometimes various theoretical reasons can compete one with another and since real users' preferences are not always compatible with theoretical deduction, another validation method can be considered as *empirical* when the guideline is approved on basis of acquired figures (e.g., experimental measures).

In this kind of method, usability questionnaires that report experiences on measurable variables are frequently used. Some of them are of general purpose such as Software Usability Measurement Inventory (SUMI) [Cork 97] or Questionnaire for User Interface Satisfaction (QUIS) [HCIL 96] while others are of specific purpose for web sites such as Web site Analysis and Measure Ment Inventory (WAMMI) [Nomos 97]. These questionnaires were intended to test the usability of a user interface for a traditional interactive application or for a web site. In order to validate the corpus of selected guidelines, we choose to use validity questionnaires instead of usability questionnaires for the following reasons :

1. usability questionnaires are more intended to test whether a particular user interface is usable or not. In order to obtain a guideline, this test needs to be reproduced with other user interfaces, under varying circumstances, thus requiring intensive web site re-development and testing;
2. validity questionnaires are more intended to directly let users estimate whether a guideline is valid or not. In this sense, they are closer to our validation goal;
3. in the five empirically measurable quality criteria proposed by Shneiderman [Shneiderman 97], i.e. time to learn, speed of performance, rate of errors by users, subjective satisfaction, retention over time, time to learn and speed of performance are the only criteria used up to now to validate web design guidelines as observed in such studies [Borges 97, Muller 97]. Thus, we would like to take another viewing angle by concentrating on the level of subjective user satisfaction to which they lift guidelines;

4. taking another viewing angle can logically confirm or infirm guidelines that have been validated by the speed of performance criteria because performance does not necessarily agree with user performance.

Basically, seven questionnaires were developed and submitted to users :

- questionnaire #1 collects demographic data (e.g., sex, age, country, background, experience level) and optional personal information (e.g., e-mail) allowing receiving the survey results;
- questionnaire #2 validates the corpus structure by asking relevant question;
- questionnaire #3 asks questions about the page components and links to graphics (i.e. thumbnails);
- questionnaire #4 asks questions about the page components, the interaction objects and graphical links (i.e. graphical buttons);
- questionnaire #5 asks questions about conveying information through graphics, graphical links (i.e. image maps) and navigational aid;
- questionnaires #6 and #7 ask questions about global and local navigation, page size and layout, page format for printing/saving purposes.

Fig. 6 depicts some typical examples of question raised in the questionnaires. Each question is identified by a three-digit number : the first digit corresponds to the questionnaire number, the second digit is related to a questionnaire section, the third digit locates the question within this section. For instance, 423 stands for question 3 in section 2 related to page components in questionnaire #4 about graphical web sites. For reference, the full electronic questionnaire is available on line at http://www.info.fundp.ac.be/~vesale/PVWG/Quest. These questions are formulated so that the following requirement are met : questions should be clear, questions should not be directive, questions should be neutral, questions should not be embarrassing, questions should not be hypothetical, questions should minimise the prestige biais, the learning effort. Moreover, some questions hold some subjective, open and qualitative character. This seems appropriate since the validation questionnaire is aimed to collect users' opinion (hence, subjective) that is representative (hence, open) about guideline (hence, qualitative).

- Questions 313-413-513-613-713 : What do you think about a web site accessible in both graphical and text-only versions? The idea behind this question is to identify what the user satisfaction is for having various web site versions.
- Questions 314-414-514-614-714 : What do you think about a web site having two different versions at the same time : One with "frames" and another without "frames" ? The idea behind this question is to identify what the user satisfaction is for having framed and unframed versions.
- Questions 323-423 : What do you think about having a link to a home page?
  The idea behind this question is to try to know the average importance of a link to a home page on a scale with 7 levels (0=unuseful, 3=practical, 6=essential)
- Question 551 : without using mouse (unless for scrolling), how many clickable regions do you see in the following image maps?
  Image map 1 : …
  Image map 2 : …
  Image map 3 : …
  Image map 4 : …
  The idea behind this question is to submit to the user 4 rather differently designed image maps and to discover the recognition criteria.
- Question 552 : In the above image maps, where are the clickable regions easily perceived and why?

Fig. 6. Example of questions raised in the questionnaire.

7.2. Results and discussion

Survey participants were solicited as follows : announcements on newsgroup, announcements made to different mailing lists, to some organisation (e.g., members of the Computer Science

Department of Loughborough University of Technology, UK, and students of Institut d'Informatique (FUNDP, Belgium) and by paper mail (all members of FUNDP). 164 people responded to the survey over a one-month period from May 15, 1998, to June 30, 1998. Like most surveys on the Internet, the sampling suffered from non-random selection and autoselection. Respondents responded via the electronic questionnaires on the Internet, through e-mail for those who subscribed to the mailing lists and postal mail for some organisation in order to include people that do not frequently use Web or e-mail.

Most respondents came from Europe and USA (USA+Canada=36%, Europe-Belgium=31%, Belgium=27%, Other=6%). Their average age was 33 years old and their professional background was directly related to Computer Science (82%) versus non-Computer Science (18%). Their computer experience level is relatively high (μ=4.8 on a 0-6 scale) as well as their Internet proficiency (μ=4.6 on a 0-6 scale). The responses provided by participants to questionnaires were examined in order to deduce and motivate a validation score for each concerned guideline as follows :

$$\text{score} = \begin{cases} 0 \text{ if nothing can be deduced about the validity from responses} \\ 1 \text{ if some elements came in favour of the guideline} \\ 2 \text{ if many elements came in favour of the guideline} \\ 3 \text{ if almost all responses concur in favour of the guideline} \end{cases}$$

Each guideline is examined across responses from several questions and each question provided responses to examine several guidelines. In order to give a flavour of the survey results, we detail here four examples : one about the corpus structure, one guideline extracted from the text-only version, one extracted from the graphical version, and one for frames.

7.2.1. Validation of the corpus structure

Through questions 313, 414, 513, 613, 713, 7% of respondents judged important that a same web site enables users to choose between a text-only version and a graphical version; 41% of respondents evoked excessive downloading times and 17% reported several accessibility trouble concerning the graphical version. Through questions 314, 414, 514, 614, 714, 82% of respondents found useful a choice between a graphical version and a framed version; 28% of participants are really concerned about frames (see reported problems above) and suck about them. Therefore, a large totality of users seem to adhere to the suggested repartition into three parts (validation score = 3) : a text-only version is always required, a graphical version is not forbidden, but an in-depth reflection should be initiated about frames.

7.2.2. Validation of a text-only guideline

This validation concerns the R.2.1.2.1 guideline entitled "The WWW page should include the authors' name. Through questions 323 and 423, the utility (μ=0,413) of providing the site author's name is not considered as fundamental. The presence of the author's name is consequently considered as practical, nothing more (validation score=1).

7.2.3. Validation of a graphical guideline

This validation concerns the R.3.1.2.3 guideline entitled "For image maps, clearly delineate the clickable region". In question 551, 4 images maps were submitted to users' appreciation :

1. image map #1 represented a country map with 12 red region names and 2 blue labels : 32% of respondents identified 12 clickable regions, 52% identified from 10 to 15 clickable regions, while 32% did not perceive any region;
2. image map #2 represented a world map : 52% did not perceive any region where 36% identified between 5 and 7 regions (probably, the continents);

3. image map #3 showed a picture of the Rushmore Mounts where presidents' heads are sculpted with their names : 80% did recognise the clickable regions while 12% did not perceive anything;
4. image map #14 depicted a navigation bar equipped with 13 navigation buttons and a logo : 92% were able to discern 13 or 14 clickable regions, 84% identified the buttons, while only 4% did not perceive anything.

Through question 552, 72% of respondents quoted borders as a clear way to recognise a clickable region (in image map 4), 28% spoke about a label (in image map 3) and 28% identified the colour code and the format (in image map 1). Through question 553, 64% of respondents believed borders are the best way to delineate a clickable region and 20% prefer colour and style combination to delineate it. The clickable regions are consequently vital to delineate (validation score=3), for example with borders, colour, style by order of preference.

7.2.4. Validation of a frame-based guideline

This validation concerns the R.1.1 guideline entitled "Use frames if you really need them". Through questions 314, 414, 514, 614, 714, 29% of respondents claimed that frames stay a real web accessibility hindrance while 12% believed frames improve page layout provided they are properly implemented. Since one third of visitors are allergic to frames, the use of frames was partially recommended (validation score=2). By repeating this process on selected guidelines, they have been attached to validation scores as indicated in Table 1.

Table 1. Results by validation score

Validation score	Amount of guidelines
3	38
2	26
1	8
0	8

8. CONCLUSION

Throughout the five steps discussed in sections 3 to 7, 64 guidelines out of 80 have been proved valid (validation score=2 or 3), where 16 guidelines were proved invalid (validation score=0 or 1). Validating guidelines with such scores increases the certainty and the trust in the way these guidelines are applied; after been applied, they can be certified with a minimum expected user satisfaction level. This level does not necessarily meet results and conclusions provided by measuring user performance. It is therefore important to compare these two series of results to highlight differences that might arise between preference and performance.

As mentioned before, ergonomic algorithms embodied in software tools should now take the validation scores into account in order to evaluate a web site across a checklist of guidelines : settings should be allowed to specify the scope of evaluation (preference versus performance). A software tool cannot unfortunately encompass all guidelines validated in this approach since some of them are of semantic nature. Only syntactic guidelines can be considered. This is why it seems important to us

- to confirm/infirm guidelines that have been validated or invalidated in this survey with other studies (e.g. [Vora 98]);
- to extend other potential user categories as encouraged in the "User Interface for All' vision to detect discrepancies as user profiles change;

- to integrate the current corpus into any set of newly validated guidelines not to avoid any possible enriching;
- to open the structure and the contents to any advance in research and development.

Anyway, the approach we described lead us to consider a first sketch of corpus containing web design guidelines whose structure and contents were partially validated together. This corpus has been itself used to check the electronic version of the corpus both in French and in English. This web site is available at http://www.info.fundp.ac.be/~vesale/PVWG/Guide.

Acknowledgments

The author would like to thank Pierre Badot and Valery Detez for conducting the complete user survey that has been here summarised, Prof. James Alty, Loughborough University of Technology for his advice and all respondents for taking the time to respond to the survey and for providing insightful comments and appreciation.

REFERENCES

[Akoumianakis 97] D. Akoumianakis, C. Stephanidis, Human Factors Knowledge Repository for Designing Accessible User Interfaces, in *Proc. of Annual Conf. of the Ergonomics Society*, Grantham, April 15-17, 1997, pp. 504–509.

[Apple 97] Apple Web Design Guide, Apple Computers Inc., February 18, 1997. Accessible at http://applenet.apple.com/hi/web/intro.html

[Badot 98] P. Badot, V. Detez, Towards a Generic Corpus of Validated Ergonomic Rules for Creating Web Sites, Master of Computer Science thesis, Institut d'Informatique, FUNDP, Namur, September 1998.

[Bastien 95] J.M.C. Bastien, D.L. Scapin, Evaluating a User Interface with Ergonomic Criteria, *International Journal of Human-Computer Interaction,* 7, 1995, pp. 105–121.

[Bastien 96] J.M.C. Bastien, D.L. Scapin, C. Leulier, Looking for Usability Problems with the Ergonomic Criteria and with the ISO 9241-10 Dialogue Principles, in Companion Proc. of CHI'96, Vancouver, April 13-16, 1996, ACM Press, New York, 1996, pp. 77–78. Accessible at http://www.acm.org/sigchi/chi96/proceedings/intpost/Bastien/jmcb_txt.htm

[Borges 96] J.A. Borges, I. Morales, N.J. Rodríguez, Guidelines for Designing Usable World Wide Web Pages, in *Companion Proc. of CHI'96*, Vancouver, April 13-16, 1996, ACM Press, New York, 1996, pp. 277–278. Accessible at http://www.acm.org/sigchi/chi96/proceedings/shortpap/Rodriguez/rn_txt.htm

[Borges 97] J.A. Borges, I. Morales, N.J. Rodríguez, Page Design Guidelines Developed through Usability Testing, in [Grose 98].

[CAST 97] Bobby Version 2.0, HTML Web analyzer, Center for Applied Special Technology (CAST), National Universal Design Laboratory, Peabody, Massachusetts, 1997. Accessible at http://www.cast.org/bobby.

[Comber 95] T. Comber, Building Usable Web Pages : An HCI Perspective, in *Proc. of AusWeb'95 Australian WWW Conf. "Innovation and Diversity"*, Ballina, April, 30-May, 2, 1995, Norsearch Ltd, Ballina, 1995, pp. 119–124. Accessible at http://elmo.scu.edu.au/sponsored/ausweb/ausweb95/papers/hypertext/comber/

[Cork 97] Software Usability Measrument Inventory (SUMI), Human Factors Research Group, University of Cork, Ireland, 1997. Accessible at http://www.ucc.ie/hfrg/questionnaires/sumi/index.html

[Detweiler 96] M.C. Detweiler, R.C. Omanson, Ameritech Web Page User Interface Standards and Design Guidelines, Ameritech Corp., 1996. Accessible at http://www.ameritech.com/corporate/testtown/library/standard/std-guix.html

[Gappa 97] H. Gappa, R. Opperman, M. Pieper, Certifying Web Accessibility for the Handicapped by ISO 9241 Conformance Testing, in *Proc. of 3rd ERCIM Workshop on "User Interfaces for All"*, Obernai, November 3-4, 1997.

[Georgia 98] GVU's WWW User Surveys, Graphic, Visualization, and Usability Center, University of Atlanta, Atlanta, 1998. Accessible at http://www.gvu.gatech.edu/user_surveys

[Grose 98] E.M. Grose, Ch. Forsythe, J. Ratner (eds.), *Human Factors and Web Development*, Lawrence Erlbaum Associates, Mahwah, 1998. Table of contents accessible at http://www.erlbaum.com/2704.html

[Hadjadj 97] D. Hadjadj, R. Agro, D. Burger, Customising HTML by Filtering : Techniques for Making the Internet Accessible to the Visually Handicapped, in *Proc. of 3rd ERCIM Workshop on "User Interfaces for All"*, Obernai, November 3-4, 1997), 1997. Accessible at http://www.ics.forth.gr/ercim-wg-ui4all/UI4ALL-97/proceedings.html

[HCIL 96] Questionnaire for User Interface Satisfaction (QUIS), Human-Computer Interaction Laboratory, University of Maryland, 1996. Accessible at http://www.lap.umd.edu/QUISFolder/quisHome.html

[IBM 97] IBM Web Design Guidelines, International Businness Machines, 1997. Accessible at http://www.ibm.com/ IBM/HCI/guidelines/web/web_guidelines.html

[Johnson 98] P. Johnson, Towards Principles for the Design and Evaluation of Multimedia Systems, in *Proceedings of HCI'98*, Sheffield, September 1-4, 1998), Cambridge University Press, Cambridge, 1998.

[Levine 97] R. Levine, User Interface Design for Sun Microsystem's Internal Web, Sun Microsystems Inc., 2 August 1996. Accessible at http://www.sun.com/styleguide.

[Lowney 96] G. Lowney, Design Guidelines to Make Sites Accessible, in *Proc. of 3rd Conf. on Human Factors and the Web*, October 1996. Available at http:// ?

[Muller 97] J.-L. Muller, Y. Brène, Towards a First Basic Corpus of Ergonomic Rules for Designing Web Pages, Master of Computer Science thesis, Institut d'Informatique, FUNDP, Namur, September 1997.

[Nielsen95] J. Nielsen, Guidelines for Multimedia on the Web, A Jakob Nielsen's Alertbox, December 1995. Accessible at http://www.useit.com/alertbox/9512.html

[Nomos 97] Website Analysis and MeasureMent Inventory (WAMMI), Nomos Management AG, 1997. Accessible at http://www.nomos.se/wammi/index.html

[Ratner 96] J. Ratner, E.M. Grose, E.M., Ch. Forsythe, Characterization and Assessment of HTML Style Guides, in *Companion Proc. of CHI'96*, Vancouver, April 13-16, 1996, ACM Press, New York, 1996. Accessible at http://www.acm.org/sigchi/chi96/proceedings/intpost/Ratner/ rj_txt.htm

[Rodríguez 96] N.J. Rodríguez, Guidelines for designing usable WWW pages, 1996. Accessible at http://exodo.upr.clu.edu/ rumhp/hci/wwwdesign.html

[Rodríguez 97] N.J. Rodríguez, J.A. Borges, I. Morales, Page Design Guidelines for Improving WWW Navigation, in *Proc. of HCI International'97*, San Francisco, July 1997, Elsevier, Amsterdam, 1997.

[Scapin 90] D.L. Scapin, Organizing Human Factors Knowledge for the Evaluation and Design of User Interfaces, *International Journal of Man-Machine Studies*, 2(3), 1990, pp. 203–229.

[Shneiderman 97] B. Shneiderman, *Designing the User Interface*, Addison-Wesley, Reading, 1997.

[Stephanidis 98] C. Stephanidis, G. Salvendy, D. Akoumianakis, N. Bevan, J. Brewer, P.L. Emilianis, A. Galetsas, S. Haataja, I. Iakovidis, J. Jacko, P. Jenkins, A. Karshmer, P. Korn, A. Mareus, H. Murphy, Ch. Stary, G. Vanderheiden, G. Weber, J. Ziegler, Towards an Information Society for All : An International Research & Development Agenda, *International Journal of Human-Computer Interaction*, 10(2), 1998, pp. 107–134.

[Usable 98] Usable Web. Accessible at http://www.usableweb.com, http://usableweb.com/item/chi.html

[UseIt 98] Use It. Accessible at http://www.useit.com

[Vora 98] P.R. Vora, Designing for the Web : A Survey, *Interactions*, 5(3), May-June 1998, pp. 13–30.

[W3C 98] WAI Accessibility Guidelines – Page Authoring, Web Accessibility Initiative, World Wide Web Consortium, Geneva, 1998. Accessible at http://www.w3.org/TR/1998/WD-WAI-PAGEAUTH-0203

[Web 97] Web Site Usability : Design Guidelines, January 2, 1997. Accessible at http://infoweb.magi.com/~asd/guidelines/guide.htm

[Yale 97] Yale C/AIM Web Style Guide, Yale-New Haven Medical Center, January 1997. Accessible at http://info.med.yale.edu/caim/manual/contents.html