Web Usability
Enhancing Effectiveness of Methodologies and Improving their Communication Features

A dissertation presented by

Luca Triacca

Supervised by

Prof. Paolo Paolini

Submitted to the

Faculty of Communication Sciences

University of Lugano

for the degree of

Ph.D. in Communication Sciences

September 2005

Board

Supervisor:

Prof.dr. Paolo Paolini, Faculty of Communication sciences, University of Lugano.

Internal examiner:

Prof.dr. Martin Eppler, Faculty of Communication sciences, University of Lugano.

External examiner:

Prof. Tom Brinck, University of Michigan, School of Information.

Abstract

The quality of web sites (or more in general interactive applications) may be improved taking into account the activity of usability evaluation, in which the quality for the end-users is defined and established. One of the most serious problems related to usability field is that usability evaluation methods often were not defined to be effectively reused by the people who did not invent them. The consequence of this lack of reusability is that it is very difficult to promote and disseminate this fundamental activity and consequently it is not enough considered within projects and relative budgets.

The thesis presents a systematic usability method, called MiLE+, which has been developed taking into account the concept of reusability. MiLE+ is the evolution of MiLE (Milano-Lugano Evaluation method) and it tries to improve its forerunner at conceptual level and, as far as reusability is concern, it offers several tools easy-and-ready to use by inspectors.

From the conceptual point of view, one of the most important and innovative contribution is referred to the introduction of the separation between application dependent and independent analysis. This approach allows a usability evaluation more focused on taking into account the nature of the problem and as a consequence it suggests more precise solutions. This distinction influences the three evaluation activities composing the MiLE+ framework: *technical inspection*, *user-experience inspection* and *scenario-based user testing*. The technical inspection is devoted to discover usability problems which are not related to specific application's goals and users, meanwhile the user-experience inspection and the scenario-based user testing are used for identify issues that are strictly connected to the nature of the application under evaluation. All the MiLE+ activities employ several tools developed taking into account the conceptual framework behind the method. These tools allow a deep analysis of the web site, in particular:

- *heuristics for the technical inspection (called technical heuristics)*, that allow the discovery of problems related to application-independent aspects;
- *heuristics for the user-experience inspection and for the scenario-based user testing (called user experience indicators – UEIs)*, that allow finding the problems related to application-dependent aspects;
- *scenarios*, that allows to evaluate the application taking into account the main user profiles and their goals;
- *usability evaluation kits* (U-KITs), that are the toolset which the inspector has to set up before the evaluation.

Another very important feature of the method is its cost-effectiveness. Indeed, in this work it has been highlighted how the MiLE+ activities can be employed considering constraints such as time and resources at disposal.

The reusability of a method is also related to the learning activities that help to communicate and teaching the methodology. So, this work presents in a very detailed manner all the learning activities (courses and modules) and support material (inspector manual) that have been conceived to simplify the understanding and the method's utilization.

To assess and validate the method we present an experiment which, at least in part, proposes an innovative approach to verify the reusability of a usability evaluation method. This approach is based on the concept of agreement on findings among inspectors, which refers to the reliability of the problems' detection obtained by the inspectors themselves. In other words, this approach allows to empirically measuring the ability of the inspectors to produce results that should be comparable and similar.

Table of contents

Acknowledgements	**-ix-**
Chapter 1: Introduction	**-1-**
1.1 Goals of this work	-2-
1.2 Usability for usability methods	-3-
1.3 Research questions	-4-
1.4 Research method and tools	-6-
1.5 Short overview of remaining chapters	-7-
Chapter 2: Usability Foundations	**-9-**
2.1 For a definition of Web usability	-10-
2.2 Quality, Usability and ROI (Return On Investment)	-13-
2.3 Existing usability methods	-16-
2.4 Usability and Accessibility	-30-
Chapter 3: From MiLE to MiLE+ method	**-33-**
3.1 Overview of MiLE model	-34-
3.2 MiLE+ Method: a systematic approach to usability evaluation	-40-
Chapter 4: Teaching MiLE+	**-71-**
4.1 The importance of teaching MiLE+	-72-
4.2 Learning@MiLE+: the modules	-76-
4.3 Learning@MiLE+: the courses	-92-
Chapter 5: Validating MiLE+	**-101-**
5.1 The importance of validating MiLE+	-102-
5.2 The Experiment	-103-
5.3 Results	-111-
5.4 Experiment's conclusions	-127-
Chapter 6: Conclusions and Future Work	**-129-**
References	**-135-**
Annex A: Inspector Manual	**-141-**
Annex B: Teaching Material - Slides Packs	**-259-**
Annex C: Detailed Experiment Results	**-297-**

Acknowledgements

First of all a great thanks to Eugénia, to my parents (Venanzio and Rosangela) and to my brother Matteo, that have supported me during the last years.

A special thanks to Paolo and Lorenzo that have helped me to understand the work of the researcher.

A particular thanks to the TEC-Lab (www.tec-lab.ch) colleagues and friends: Alessandro, Daniele, Davide, Giovanni, Marco and Nicola. We have worked (very) hard, but we have obtained amazing results!

A thanks to the colleagues of Politecnico di Milano, Franca and Nicoletta, for having collaborated actively to my research with very valuable ideas.

A thanks to the University of Lugano which have helped me to study and work in a very proactive and multidisciplinary environment.

Chapter 1:

Introduction

Summary:

The first part of this Chapter introduces the goals and motivations of the entire work, especially focusing on the need of creating a reusable methodology for the usability evaluation.

The second part explores in depth the two research questions and the research method to answer them. The key questions leading this work are:

1. *"Is it possible to engineer and standardize the usability evaluation process proposed by MiLE?";*

2. *"Is it possible to effectively communicate (and teach) the MiLE method?".*

The Chapter ends with the structure of the work.

1.1 Goals of this work

The main objective of this work is to increase the communication aspect and to enhance the reusability of the MiLE methodology. MiLE (Milano-Lugano Evaluation method) is a methodology for evaluating the usability/quality of web sites, fruit of a common research carried on by the Politecnico di Milano and the University of Lugano. The enhancement of usability methods is related to the importance assumed by usability evaluation in the last decade, in particular usability has gained a key strategic role in the internet economy than it had in the past (Nielsen, 1999), since a web site is an "open product", accessible by anyone who navigates in the WWW –World Wide Web (Triacca L., 2003) and for this reason they have to be usable.

The dissemination of the usability evaluation as one of the most important dimension for defining and assessing the quality of a web application (or more in general of interactive application) is hampered from the complexity to communicate how to use different methodologies and techniques. The main problem is that these methods are not *reuse-oriented*, i.e. they have not been defined to be effectively reused by the people who did not invent them. Most of usability techniques are proprietary methods or guru-dependent; in other words, these techniques are difficult to be used by less-experienced evaluators because they do not provide the evaluators with the necessary conceptual tools to gain appreciable results. The problem of the reuse is strongly connected to the difficulty of teaching and communicating the essence of a method in a way that also others can successfully apply it. Project teams are acknowledging the importance of usability evaluation but are still reluctant to make considerable investment in consultancy for an "ad-hoc" evaluation, especially if the web project is at the end and the remaining budget is very limited. Effective reuse of usability knowledge and practices would enhance the adoption of usability techniques by designers and project teams. Indeed, the goal of this work is to propose a methodology which can be used both by usability specialists and by people without a strong background in usability evaluation. Only providing reusable tools for performing the usability evaluation it will be possible to introduce this activity as an essential part of the application's lifecycle.

1.2 Usability for usability methods

Overall this work points out the fact that a website, or in general all the human artefacts, should be usable (easy to use). So we underline many times that usability is one of the most important pillars in order to assess the quality of a product and, as consequence, promote the image of the institution developing and managing the application.

As previously mentioned one of the great issues in the field of usability evaluation of interactive applications is the fact that it is very difficult to identify a usable methodology and this situation affects the possibility to efficacy and efficiently promote the usability evaluation activity.

The reasons why of poor usability of the usability methods should be different:

- *HCI (Human Computer Interaction) is a new field:* since the relationship between man and computer is recent, there are several aspects to be defined. Indeed, usability evaluation of interactive applications is considered a "new science" and methods are either new-fangled or borrowed from other fields.

- *Usability is quite an arduous task:* in general evaluating the usability of human artefacts is a complex activity. For example, if we try to evaluate the usability of a building we have to consider several aspects:
    - the speed of the elevators: a slow one forces people to wait each time they want to go upstairs;
    - the position of emergency exits: if they are located far away from the rooms of the building, they become inaccessible;
    - the existence of architectural features that deny the access to people with disabilities;
    - …

  Evaluate interactive applications is a very complicate activity as well. Indeed, these applications having a lot of interactive features, contents, navigational mechanisms… address to wide and heterogeneous targets. For example the CNN website (www.cnn.com): it has a lot of contents and interactive features (e.g. maps, games, etc.) and it addresses to millions of people around the world. In this case to achieve a high-level usability of CNN.com is an arduous task, as much as its evaluation.

- *Existing Usability methods do not offer reusable tools:* it is very difficult to develop tools easy&ready-to-use for performing this activity. There are a few methods providing tools for the evaluation that can be used by people who do not create them or without a great experience in the field of usability evaluation.

To reach the goal of creating a usable methodology, the main effort of this work has been concentrated on two aspects: the development of "usability evaluation tools" and the creation of "usability learning paths".

1.3 Research questions

The two main research questions that lead the entire work are:

1. **"Is it possible to engineer and standardize the usability evaluation process proposed by MiLE?";**

2. **"Is it possible to effectively communicate (and teach) the MiLE method?".**

The first question needs to be seen on four dimensions:

1. The first one is to identify the best way to improve the reusability in web usability. As stated before, existing methodologies are proprietary methods or guru-dependent and consequently they are difficult to be used by less-experienced evaluators. MiLE presents a general framework that could help increasing the communication of the methodology and therefore growing the reusability of the method itself.

2. The second dimension to investigate is related to the process of systematize the usability evaluation procedure. Without following a clear usability process the evaluation becomes very difficult. Every evaluation process should be stated as clearly as possible both for accelerating the evaluation itself and for communicating in an effective way the overall methodology. In particular, we have increased the communication of the usability evaluation workflow (from the preparatory to the execution phase). This means that for every phase we have identified the principal actors, the activities in which they are involved (e.g. which actor is involved in the scenarios construction) and the way (guidelines) to conduct them.

3. The third dimension is to make the MiLE methodology cost-effective. In term of time and costs the "zoomable" character of MiLE should allow a usability evaluation tailor made to the needs of the client. It is very important to identify

the different activities of MiLE and adapt every version to the client's constraints. This dimension is a consequence of the second one: in fact, once identified the different usability evaluation phase and the actors involved, it is possible to clearly set up the MiLE version to use, always with respect to limitations such as time and budget;

4. The fourth dimension is connected to the necessary expansion of the methodology. Indeed, even though in the last years MiLE has obtained several satisfactory results, it is necessary to re-think several aspects of the methodology with a communication perspective and to re-elaborate some of the results obtained.

For all these reasons a great part of this work is dedicated to present MILE+ (and the related results in using and teaching it), which is the evolution of MiLE.

The second question ("Is it possible to effectively communicate (and teach) the MiLE method?") is the consequence of the first one. Indeed, once answered satisfactorily the first question, it will be possible to identify the process to communicate and teaching MiLE+. To efficiently answer this question the following aspects have been taken into consideration:

- MiLE+ presents different levels of granularity (macroscenarios, scenarios, etc.). As mentioned above, the evaluation could be taught in a very "tailor made" manner considering the needs and the characteristics of the client/students. It is fundamental the importance of creating different learning paths tailor made to the needs of the "users". Indeed, it will be important to clearly establish the workflow for teaching MiLE with respect to the needs of the students and to constraints.

- MiLE+ should be thought as a "cluster of learning modules" (each module corresponds to a specific concept of MiLE+).

To answer to the research questions it had been necessary an interdisciplinary approach. This work is not only concentrated in usability and web design, but it deals with themes connected to multidisciplinary fields such as learning, communication and psychology.

1.4 Research method and tools

The research method will be empirical and theoretical. The research process is iterative and allows the development, assessment and validation the evolution of the current method (from MiLE to MiLE+). The iterative process allows both the possibility to verify every theoretical result with a case study and to use case studies to add new theoretical elements and re-think the methodology.
Moreover, once concretely developed and assessed MiLE+ (from conceptual to practical aspects and tools), we started teaching it for verifying its learnability and reusability.

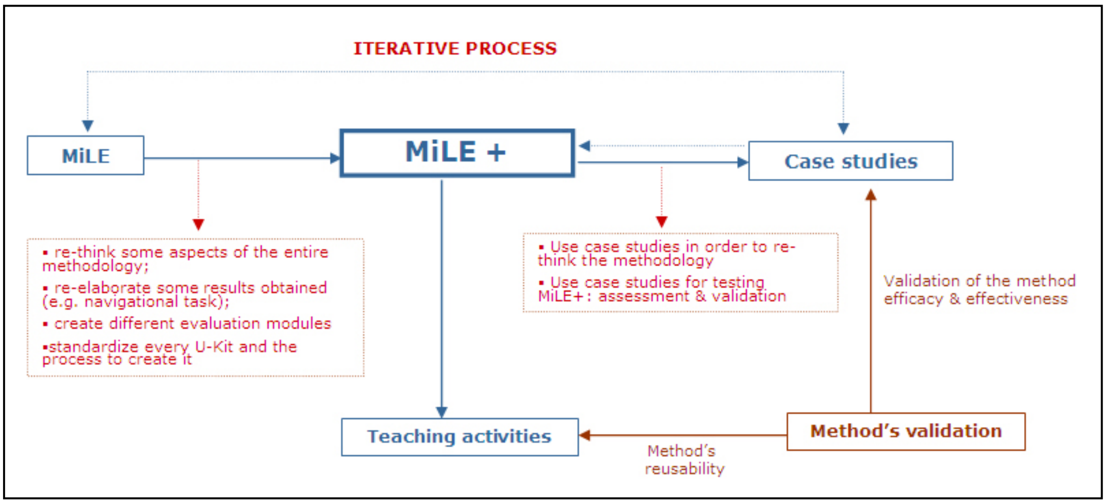


Figure 1: the iterative process followed within this research.

Theoretical development
The theoretical development of the existing methodology will be approached at different levels:

- the survey of the literature in the field of usability inspection techniques and user testing methods. The survey helps to re-think some aspects of the entire methodology. In particular, we will focus our attention on user testing techniques to identify the best way to integrate user testing after the inspection activity. Indeed, at present it not still clear how to integrate the user testing phase in an efficient way;

- the standardization of the MiLE+ process. It will be very useful to reanalyze in depth all the phases of the methodology and to efficiently comment every single concept within a single phase (e.g. the concept of user profile inside the phase of scenario's creation). From one hand this allows to increase the communicability of

MiLE+; to the other hand to identify and isolate the different modules composing MiLE+. The goal of these activities is twofold:

- Identify every evaluation module inside of MiLE+ process;
- Use these modules to create different learning paths to teach MiLE+.

Empirical assessment

As stated above, several cases (applications) will be evaluated using MiLE+ in order to support the theoretical definition (supplying evidence to theory and hypotheses). The results obtained from these case studies will enable to identify the limitations of the first set of theoretical results, thus paving the way for improving the general methodology. Indeed, the aim of this work is not to present every single case study, but the final conceptual results achieved thanks to method's employment in several context.

To verify and validate the learnability and reusability of the method by people who did not invent it we had planned and performed an empirical experiment, which is presented in Chapter 5.

1.5 Short overview of remaining chapters

The rest of the thesis is organized as follows. Chapter 2 presents a short introduction on usability and its importance for the quality of interactive application. Then we present a detailed review of the related works, highlighting key achievements relevant for this research in the field of usability evaluation. As a result, this chapter will point out lacks of the current approaches.

Chapter 3 is dedicated to briefly introduce the strengths and weaknesses of the MiLE method. However, the main goal is to fully present MiLE+. So, first of all we illustrate the innovative conceptual approaches proposed by MiLE+ and its evaluation activities. Then we introduce the suggested evaluation process. The Chapter ends with the explanation of the method's cost-effectiveness and reusability.

Chapter 4 is completely devoted to present MiLE+ learning modules and the courses. For each module it is fully showed its instructional design.

Chapter 5 is reserved to present the experiment which has allowed to empirically verifying the learnability and reusability of the MiLE+ method.

Chapter 6 illustrates some outlooks for future work emerging from the research.

Annex A is dedicated to present the inspector manual; comprising of all the useful tools for the inspector.

Annex B presents the learning material employed during the MiLE+ courses.

Annex C reports the detailed results emerged from the experiment.

Chapter 2:
Usability Foundations

Summary:

The first part of this Chapter introduces the concept of usability and its importance for the quality of human artefacts, and in particular interactive applications. Besides, it highlights the impact of usability evaluation on the ROI (Return On Investment) of the institution which invests part of projects' budget in this fundamental activity.

The second part is devoted to present existing usability methods, divided in two main categories: user-based methods (user testing) and inspection methods.

The Chapter finishes with a brief introduction on the relationship between usability and accessibility.

2.1 For a definition of Web usability

In the '70s, the importance of evaluating the interface of a software product became a crucial activity. It was clear that the efficient use of the interface was heavily dependent on how the interface was designed and communicated to the user.
As more and more software was developed for interactive use, attention to the needs and preferences of end users intensified (Rosson, M.B. et al.: 2002). It is in the early '80s that the usability starts to be a "science". Indeed, in these years the PCs became a familiar object for people and so the interaction with the different applications. It is exactly in this period that a new field of studies was born: the Human-Computer Interaction (HCI). HCI represents the intersection area between computer science and cognitive science. In short terms, Human-Computer Interaction is the study of how people interact with computers and to what extent computers are or are not developed for successful interaction with human beings. In the '90s usability assumed a fundamental role in the development of interactive applications. With the arrival and diffusion of the World Wide Web (1993-1994), the development of a large number of website, and the widespread use of electronic mail it became clear that assessing the usability degree of interactive applications is one of the key factors for the applications' success. Now, in the first years of 21th century, usability concepts are extending towards accessibility. The new challenge in developing usable applications is to develop applications which are accessible by ideally anyone. Accessibility refers to ensuring that content is accessible, ie. ensuring that content can be navigated and read by everyone, regardless of location, experience, the type of computer technology used, or disability. Accessibility is most commonly discussed in relation to people with disabilities, because this group are most likely to be disadvantaged if the principles of accessible Web design are not implemented. Failure to follow these principles can make it difficult or impossible for people with disabilities to access content.
The main goal of the usability evaluation is to detect the most part of the usability problems and breakdowns of a web application, being the usability *"the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in particular environments" (ISO 9241-11).* Usability is therefore a combination of factors that affects the user's experience with the product. These factors, in general, include:

1. *The usefulness of the product*, i.e. the degree to which the product enables a user to do his work and achieve his goals;
2. *Learnability*, i.e. the measure of how rapidly a user can become productive. A measure of how rapidly an infrequent user can re-learn the product after periods of not using it;

3. *Effectiveness*, i.e. the measure of user productivity and how well a user can perform his job;
4. *Efficiency*, i.e the measure of how quickly a user can perform work and the error rate in doing so;
5. *Satisfaction and Attitude*, i.e the degree to which users like the product. Measurement of attitude, perceptions, and feelings about the product;

Nowadays, electronic communication through the Internet is normal in the commercial and the academic fields. Unfortunately, the burden of managing the different aspects of web communication is quite often left in the hands of technical personnel, when mastering website communication requires technical skills as well as dominion of communication science. The creation and management of websites involve both technical and non-technical aspects since a website is a complex and a multidimensional reality. In turn, there are several factors affecting and determining usability (VNET5 Consortium: 2001):

- *What the application is:* analysing the characteristics of the services offered by the application is essential to understand the system under evaluation.
- *Who is using it:* profiles of the potential users of the application have to be carefully taken throughout the process of usability evaluation.
- *What they want to achieve:* the goals and the tasks the users wish to accomplish using the system are the driving concepts necessary to evaluate the actual usefulness and effectiveness of the application.
- *The context:* the access devices and the circumstances of use play an important role in determining how a service is perceived as usable by the potential adopters.

Only if all of these dimensions are taken into consideration, will it be possible to communicate successfully through the Internet. The evaluation of the usability of a web application should consider all these aspects of the application. The concept of website usability may be applied to different aspects of a website but it usually refers directly to its technical dimension. Moreover, usability puts in relation the technical aspect with the users' dimension. At the same time, any usability inquiry, should always take into consideration the communication goals and the different stakeholders[1] (Cantoni et al.: 2003).

Usability has recently assumed a much greater importance in the internet economy than it had in the past (Nielsen J.: 1999), since a web site is an "open product", accessible by anyone who navigates in the WWW (Triacca L.: 2003). This means that both in the design phase and after its launch, it is necessary to assess the real quality of the product.

[1] This includes the people who have an interest in the communication process: owners, promoters, sponsors, visitors, etc.

Usability has therefore become a fundamental issue, in every phase of the design process, from the beginning to the end (Brinck T. et. al.: 2002). In fact, the evaluation helps to ensure that the design is on track to satisfy the goals of the design. The evaluation of the usability should be an activity presents in every phase of the development process of a web application: from the requirements analysis, to launch phase, going through all the intermediate phases (conceptual design, mocks-ups and prototypes and production); this approach is called the Pervasive Usability Process (Brinck T. et. al.: 2002).

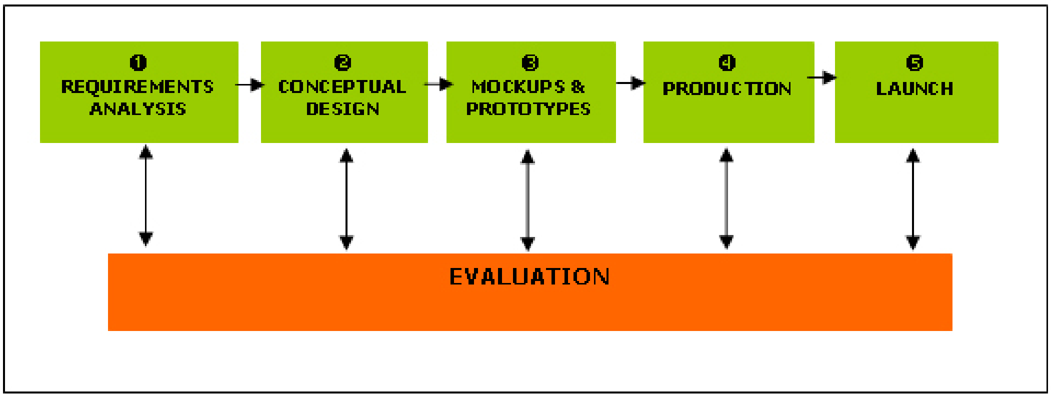


Figure 2: the Pervasive Usability Process (Brinck T. et. al.: 2002)

2.2 Quality, Usability and ROI (Return On Investment)

Complex web applications, whose goal are to communicate information and services to a large number of users, have to pay special attention to their usability, or rather quality. Clearly, this is an arduous task for the designers (and in general for all the stakeholders involved in the development of the application): web applications are of growing complexity, address several targets, deal with complex content, have different communication goals: for all this reasons, they need to be well "usable" and efficient. Evaluating the usability of a web application means to try and answer some crucial questions: e.g., How can we avoid users "getting lost" in the site? How is it possible to improve navigation's effectiveness? What kind of contents shouldn't be missing? How is it possible to know whether the users have learnt anything from the site? The -ambitious-goal is to establish the degree of user satisfaction with the application and consequently a set of guidelines for improving its quality.

The main goal of the usability evaluation of a web application is to reduce the distance and the gap between the system design approach and the user-centered approach. The system design philosophy focus on the technological characteristics of the system and, then, it design the application around the system. At the contrary, the user-centered approach stresses the role of the users themselves, not considered as the weakest part or the periphery of a technological system, but as the center itself and the goal of every technological application. When a system match user needs, the user increases the satisfaction of his user experience.

From an economic point of view, several researches show that improving the usability of general software systems can be highly cost effective. Usability techniques can reduce costs (including development, support, training, documentation and maintenance costs), shorten development time and improve marketability (Donahue G. et. al.: 1999). So it becomes fundamental to dedicate a part of the global project's budget for carrying out the usability. Jakob Nielsen (J. Nielsen: 2003) (well-known usability guru) thinks that development projects should spend 10% of their budget on usability. Following a usability redesign, websites increase usability by 135% on average. The rule of thumb in many usability-aware organizations is that the cost benefit ratio for the usability is $1:$10-$100, that is, for every dollar spent implementing usability techniques, the organization will realize a benefit between $10 and $100 (Gilb, T.: 1998). If this estimate is relative to software systems in general, it is absolutely clear that for the specific domain of web applications, that address a widespread and heterogeneous target, the ROI of the usability evaluation, increase dramatically. Developing high quality web applications

means both enhancing the credibility of the application itself and increasing exponentially the credibility of the organization that manage the web site and the satisfaction of the users/clients.
One of the strategic issues related to the costs of making usability is to determine the key moment for assessing the usability of the application. As previously presented, the usability evaluation is an orthogonal dimension that should influence each step, but the importance to verify the degree of usability (therefore the quality) varies for each step. For example, in the requirement analysis the needed effort for the usability activity is minimal but increases during the development. The key moment for testing the usability degree is the prototyping phase. Testing usability during this phase is essential for time and resources reasons. In fact, the discovery of usability problems early in the creation process reduces dramatically the costs for redesign: modifying the prototype is more cost-effective than changing the final full-fledged application. Besides, during the prototyping phase, it is possible to straightforwardly introduce structural changes (especially for aspects related to navigational strategies and the information architecture); on the contrary, the final application does not allow for structural changes without a large investment in terms of time and resources (Triacca L. et al.: 2004, 2005). Once the application is online the usability becomes a crucial activity for evaluating and monitoring the final quality of the product. Indeed, in this phase it is very important to collect data and feedbacks from the users, for knowing their actual satisfaction.

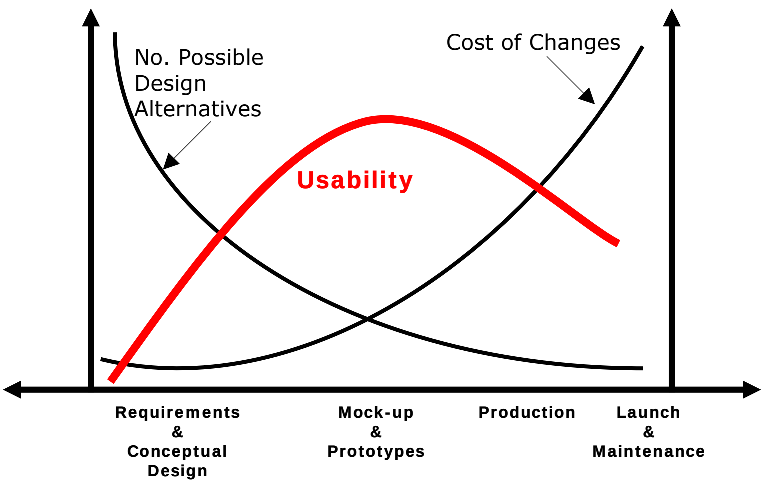


Figure 3: Key moment for assessing usability (adapted from Pressman R.: 1992)

As showed in Figure 3, at the beginning of the development of an application (Requirements Analysis and Conceptual Design) there are a lot of design alternative and the cost of changes is low. As said before, the key moment for evaluating usability is the

mock-up and prototypes phase. Indeed, in this phase is still possible to introduce structural changes and the cost is moderate. One the application is online the cost of chances increases considerably.
Summarizing the consequent benefits of the evaluation activity can be summarised as follows:

- *Gaining higher level of usage:* the services will be easy to use and meeting the expectations of their audience, the more novice users will access and exploit the service offered. As far as the returning customers concerns, usability – intended as a process of on-going improvement - will also retain those who already know the service and use it on a regular basis.

- *Lowering user support cost:* usable services will decrease the cost of helping the customer to search the needed content and accomplish his/her goals with the application. The web application interface, operations and functionality should be self-evident, i.e. understandable to the user without any external intervention or support by other people. Moreover, correctness, relevance and accuracy of the content (besides the expected performance of the network equipment) will likely decrease the probability that customers will complain to the service provider about the quality of service offered.

- *Contributing to trust building:* "Trust Builds From the Customer Experience". Improved usability and professional appearance feels solid; a clear navigation conveys respect for customers and an implied promise of good service. Typos or difficult navigation communicate disregard for the users. These are just few examples showing how usability and a good user experience is crucial to gain trust for the customers. Trust building is the first step to convert users into customers and convert customers into returning clients.

2.3 Existing usability methods

Within the field of usability methods it is possible to identify several approaches for evaluating web usability. Among them, the most commonly adopted are *user-based methods* (or *user-testing* methods) and *usability inspection methods* (or *expert reviews*) (Matera M. et al.: 2002).

2.3.1. User-Based methods (User testing methods)

User-based methods mainly consist of user testing, in which usability properties are assessed by observing how the system is actually used by some representatives of real users (Whiteside J. et al.:1988) (Dix A. et. al.: 1998). User-testing evaluation provides the trustiest evaluation, because it assesses usability through samples of real users. However, it has a number of drawbacks, such as the difficulty to properly select correct user samples and to adequately train them to manage also advanced functions of a web site (Matera M. et al.: 2002). Furthermore, it is difficult, in a limited amount of time, to reproduce actual situation of usage. This condition is called "Hawthorne effect" (Roethlisberger et al.: 1939): if the variable of the experiment are manipulated, it is possible that the productivity of the group observed decreases. Failures in creating real-life situations may lead to "artificial" conclusions rather then realistic results (Lim K.H et al.: 1996). Therefore, user-testing methods are considerable in terms of time, effort and cost. User testing is the main way for evaluating right away the look and feel of the interface, as it is possible to verify at "real-time" the reactions of the users.
Within the category of user-testing methods there are several techniques, the most important are:

- *Thinking aloud;*
- *Contextual inquiry;*
- *Focus group;*
- *Interview.*

Thinking-aloud
Thinking aloud was originally described by Karl Duncker (1945) in his work within experimental psychology were he studied productive thinking. In the field of HCI thinking aloud is one of the most popular techniques. It is often referred to as the usability method and used both in laboratory setting, workshops and field testing (Nielsen J. et al.: 2002). A research conducting by Clemmensen (2002) shows that among HCI practitioners and researchers in Denmark, thinking aloud appeared to be the single most frequently applied technique in usability testing. Other authors (Dix et al.: 1997) promote thinking-aloud for

its simplicity. Jacob Nielsen (1992, 1994) has been tireless in promoting the technique and its benefits (Nielsen J. et al.: 2002). The breakdowns of thinking-aloud are related from one hand to the cognitive load and added strain on users as well as the interruptive role of the observer during the test (Preece, 1994); to the other hand the user probably has some difficulties in speaking when the task is arduous (Preece et al.: 2002).
During the thinking-aloud test, the user should think aloud while performing some specific task with the system. By verbalizing his thoughts, the user allows the observers to know his opinions and feeling about the application. Verbal protocols are recorded concurrently or retrospectively. The subject is probed to verbalise problems that come up. After the recording of verbal protocols, the protocols are encoded according to a previously defined encoding scheme. Verbal reports can be interpreted if the processes by which they were generated are understood. Interpretation is based on the theory that human cognition is information processing (Newell A., Simon H.: 1972). Cognitive processes and their structure account for the results of verbalisations. The accuracy of verbal reports depends on the procedures used to elicit them and the relation between the requested information and the actual sequence of heeded information.
Thinking aloud allows you to understand how the user approaches the interface and what considerations the user keeps in mind when using the interface. If the user expresses that the sequence of steps dictated by the product to accomplish their task goal is different from what they expected, perhaps the interface is convoluted.
Although the main benefit of the thinking aloud protocol is a better understanding of the user's mental model and interaction with the product, you can gain other benefits as well. For example, the terminology the user uses to express an idea or function should be incorporated into the product design or at least its documentation.

Contextual Inquiry

Contextual Inquiry is a specific type of interview for gaining data from the user. This technique aims at understanding the context in which the application is used. Contextual Inquiry (also known as "site visits") is basically a structured technique of observing and interviewing users. It is based on the core principle that understanding the context in which a product (or service) is used (or the work is being performed) is essential for user and customer oriented design. Using contextual inquiry, you visit the workplace of prospective users to see how they work. You observe all aspects that would help define a context for their work - and thus a context for the usage of your product or service.
Contextual Inquiry is adequate in situations where the subject domain is unclear or unfamiliar to the development team, and when the context of work may have a significant effect on the new product or service. For performing a Contextual Inquiry considerable investment of time and effort may be needed in order to elicit sufficient information from the users and the environment to be studied. Contextual Inquiry follows many of the same process steps as field observations or interviews. Contextual inquiry is best done by

a group of researchers who develop a medium- to long-term relationship with a group of organisations who are interested in providing data. According to Holtzblatt and Beyer (Holtzblatt K., Beyer H.: 1996) the relevant steps are the following:

- Identifying the customer: identify the groups that will be using the new technology or are using similar technology, and arrange to access organisations within the groups that give a cross section of the (potential) market.
- Arranging the visit: write to the targeted organisations identifying the purpose of the visit, a rough time-table, and how much of the employees time will be taken up by the exercise. Ensure that some feedback from the day is possible before leaving. Ensure that the participating organisations understand how many visits you intend to make over the time period of the evaluations.
- Identifying the users: a software product will affect many people throughout the organisation, not just the management or the end users. Ensure that you understand the key users in the organisation whose work will be affected by a new system or changes in the current one.
- Setting the focus: select what aspects of the users' work you wish to make the focus of each visit, and write down your starting assumptions. Make a statement of purpose for each visit, and after the visit, evaluate to what extent you have achieved your purpose.
- Carrying out the interview / observation: stay with the selected users until you have managed to answer the questions you have raised in 'setting the focus'. Very often this may involve inviting the user to directly share and comment on your notes and assumptions.
- Analysing the data: the process of analysis is interpretative and constructive. Your conclusions and ideas from one round of observations are input to the next round, and an evaluation of the results so far should be one of the purposes of subsequent visits."

Focus Group

Focus group research has long been a respected method in marketing research (Lazarsfeld P.F: 1972, Merton: 1956, Sullivan S.: 1991). Its hallmark is its "explicit use of the group interaction to produce data and insights that would be less accessible without interaction" (Morgan D.: 1998). Asking a diverse group to give opinions of real or potential products and services quickly clarifies any disagreement among representatives of target markets for products (Sullivan S.: 1991). Morgan claims that "what focus groups do best is produce an opportunity to collect data from groups discussing topics of interest to the researcher (Sullivan S.: 1991). This means that they are informal, but somewhat controlled by questions the researcher posses. When the3y deal with questions that people can discuss in public they give good feedback about why people hold particular opinions, they add a depth to interview research that comes out of that group stimulation. In the field of usability evaluation, the goal of focus group technique is to identify the problems of the application by means of discussions with groups of users. Focus group analysis is an informal technique that can be used to assess user needs and requirements and the satisfaction degree they have using the application. It can be applied at any time in the development process. In a focus group about 6-9 users are brought together over a period of about 2 hours to discuss whatever issues are of interest: new concepts, designs, prototypes, complete application. The moderator running the focus group is responsible for maintaining the focus of the group on the issues of interest following a pre-planned script. One of the main problems is that focus group meetings are demanding in terms of the number of representative users needed. It is preferable to run more than one focus group since the outcome of any single focus group session may not be representative. During the Focus group the moderator presents issues to be discussed in the focus group session. He tries to keep the discussion on track without inhibiting the free flow of ideas and comments from the participants. He ensures that all members of the focus group get to contribute to the discussion.

The focus group tends to highlights only surface-oriented issues. Indeed, during this activity the users concentrate their attention on the look and feel of the interface. Besides, focus group is very useful for collect user feedback on a list of potential features they could include in the next phase of development.

Interview

Interview is an informal technique for the investigation of the users' opinions about the application, e.g. subjective satisfaction, critical incidents, anxieties which are hard to measure objectively. It is a useful method for studying what features of the application users particularly like or dislike. Three types of interviews can be distinguished: unstructured, semi-structured and structured interviews. The type, detail and validity of the collected information vary with the type of interview. The validity of results varies with the experience of the interviewers. The interviewer needs domain knowledge in order to ask the right questions and there is always the risk of bias in what questions the interviewer asks and how the interviewee interprets them. Besides, Interviews are demanding in terms of the number of representative users needed. It is preferable to use questionnaires where possible. Because of the unstructured nature of an interview the result is just a report summing up the comments made by the subject in the interview.

The interview technique is often used to complement laboratory observations of a user’s process (Sullivan P. 1991). In it, the researcher can ask to the users what they found harder and easier, what they remember, or what they preferred, and then check their post hoc responses against their actions in the session. The data collected from the interview helps confirm some findings, shed light on some confusing spots.

According to Kuniavski (2003) the general interview structure is divided in six phases:

1. *Introduction:* in the case of group interview each participant introduces himself; see that it is important to know the other people in the group. This activity emphasizes the similarities between all the participants, including the interviewer. In contrast, an individual interview introduction establishes the role of the interviewer as a neutral, but sympathetic entity.
2. *Warm-up:* the process of answering questions or engaging in a discussion needs everyone to be in an appropriate frame of mind. The warm-up in any interview is designed to get people to step away from their regular lives and focus on thinking about the product and the work of answering questions.
3. *General Issues:* the initial product-specific round of questions concentrates on the issues that surround the product and how people use it. The focus is on attitude, expectation, assumptions and experiences.
4. *Deep focus:* the application is introduced and people concentrate on the details of what it does, how it does it, whether they can use it, and what their immediate experience of it is. For usability testing this phase makes up the bulk of the interview, but for contextual inquiry, where the point is to uncover problems, it may never enter the discussion.
5. *Retrospective:* this phase allows people to evaluate the application in a broader light.
6. *Wrap-up:* this is generally the shortest phase of the interview. It formally completes the interview so that the participants aren't left hanging when the last question is asked, and it brings the discussion back to the most general administrative topics.

2.3.2 Usability Inspection Methods

Usability Inspections methods (called also Expert Reviews) is the generic name for a set of methods based on having expert evaluators inspect or examine usability-related aspects of a user interface (Nielsen J. et al.: 1994). The term *Usability Inspection* born within the fields of software engineering in reference to function and code inspections methods that have been used in software engineering for debugging and improving code (Ackermann A.F. et al.: 1989).
With respect to user-testing evaluation, usability inspection methods are more subjective, having heavy dependence upon the inspector skills (Matera M. et al.: 2002). The focus of usability inspection methods is on the usability related aspects of user-interface of interactive products and services. The objectives of this approach are bounded to the identification of some interface problems in an existing design, and then using these problems to make recommendations for fixing the problems and improving the usability of the design. This means that usability inspections are normally used at the stage in the usability engineering cycle when a user interface design has been generated and its usability (and utility) for users needs to be evaluated (Nielsen J. et al.: 1994).
The main advantage of inspection methods is the relationships between costs and benefits. In fact, performing usability inspection "save users" (Nielsen J. et al.: 1994), (Jeffries R. et al.: 1991) and does not require any special equipment and the inspector alone can detect a wide range of usability problems and possible faults of a complex system in a limited amount of time (Matera M. et al.: 2002). For these reasons, inspection methods have achieved widespread use in the last years, especially in industrial environments (Madsen K.H., 1999). However, current usability inspection methods have a number of drawbacks:
They focus on "surface-oriented" features of the graphical interface (mainly at page level) (Green T.R.G et al: 1996). Only few of them address the usability of the application structure, i.e., on the organization of both information elements and functionality;
They are strictly dependent on the individual know-how, skill and judgment of inspectors, making a subjective process. Domain and application experience may improve the evaluators' performance.

The main inspection usability methods for hypermedia and web applications are:

- *Heuristic evaluation;*
- *Cognitive Walkthrough;*
- *SUE (Systematic Usability Evaluation)*;
- *Content Evaluation.*

Heuristic evaluation

Heuristic evaluation is the most informal method; usability specialists have to judge whether each dialogue element conforms to established usability principles or not (Nielsen J. et al.: 1994). Heuristic Evaluation (created by Jakob Nielsen in 1994) is an inspection method in which one or several evaluators systematically inspect the user interface according to general usability principles (called "heuristics"), which describe the ideal characteristics of a usable interface. The evaluators examine the interface and verify its compliance with these heuristics. In 1990 and 1994, Jakob Nielsen, in collaboration with Rolf Molich, developed a very-well known list of 10 heuristics, which became general principles for user interface design and usability review. One of the main benefit of heuristics inspection – independently from the specific set of heuristics used – is that it provides a "guide" for the evaluators about where and what to look in an application and how to interpret its complexity. In this way, heuristics are useful tools to "force" inspectors analyze the different aspects of the user interface, which are often overlooked without a supporting method at hand. However, some drawbacks should be also noted for heuristics-based inspection. Heuristics enable to carry out a "static" analysis of the application (i.e. to verify if it is compliant with given principles); however, this compliance does not guarantee that the application can effectively support user's goals and tasks. It may seem a paradox that an application with no content (empty pages) is fully compliant with the most known usability heuristics.
The principles given by Heuristic Evaluation (Nielsen: 1994) are fairly broad and can be applied to practically any type of user interface.

The 10 Heuristics provided by Nielsen are:

1. *Visibility of system status:* the system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
2. *Match between system and the real world:* the system should speak the users' language, with words, phrases, and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order;
3. *User control and freedom:* users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.
4. *Consistency and standards:* users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions
5. *Error prevention:* even better than good error messages is a careful design which prevents a problem from occurring in the first place;
6. *Recognition rather than recall:* make objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to

another. Instructions for use of the system should be visible or easily retrievable whenever appropriate;

7. *Flexibility and efficiency of use:* accelerators - unseen by the novice user- may often speed up the interaction for the expert user to such an extent that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions;
8. *Aesthetic and minimalist design:* dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility
9. *Help users recognise, diagnose, and recover from errors:* error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution;
10. *Help and documentation:* even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focussed on the user's task, list concrete steps to be carried out, and not be too large

Cognitive Walkthrough

The Cognitive Walkthrough (CW) grounded on Lewis and Polson’s CE+ theory of explanatory learning (Lewis et al.: 1993; Polson et al. 1992; Wharton et al.: 1994). The CE+ is an information processing model of human cognition that describes human computer interaction in terms four steps (Riedman J. et al.: 1995):

1. The user sets a goal to be accomplished with the system (e.g. “check spelling of this document”)
2. The user searches the interface for currently available actions (e.g. menu items, buttons, command-line…);
3. The user selects the action that seems likely to make progress toward the goal;
4. The user performs the selected action and evaluates the system’s feed-back for evidence that progress is being made toward the current goal.

The CW method has been proposed by Polson (Polson at al.: 1992) as a cheap and quick method for evaluating interface design at an early stage of the system development (clearly CW is also used for the evaluation of final applications).
CW is an inspection method which focuses on the evaluation of the ease of learning of a user interface and learning by exploration. In a cognitive walkthrough an interface design is evaluated in the context of one or more specific user tasks (VNET5 Consortium: 2001). The evaluator(s) acts as if the interface was actually built and he (in the role of a typical

user) was trying to accomplish the tasks. In CW, the evaluators choose a set of representative tasks, and step through the actions performed by an imagined user with certain hypothesized attributes (Ereback A., Höök K.: 1994). The user's goal and the actions expected from the interface are compared. Each step the user (embodied by the inspector) would take is scrutinized: impasses where the interface blocks the "user" from completing the task indicate that the interface is missing something or has some usability problem.

According to Rieman, Franzke and Redmiles (Rieman et al.: 1995) the prerequisites to the CW activity include:

1. A general description of who the users will be and what relevant knowledge they possess;
2. A specific description of one or more representative tasks to be performed with the system;
3. A list of the correct actions required to complete each of these tasks with the interface being evaluated.

SUE (Systematic **U**sability **E**valuation**)**

SUE is an inspection method for hypermedia applications. SUE proposes, instead, that an application must be analyzed at different levels. Interaction and presentation features refer to the most general level, common to all interactive applications. More specific levels address the appropriateness of design with respect to the peculiar nature and purpose of the application. SUE stresses that usability analysis should consider the specific nature of the application to be evaluated, not just broad and general issues common to all the interactive systems (Matera M. et al.: 2002). The more is know about the purpose and the nature of the system being evaluated, the deeper and more effective is the usability evaluation process. An in-depth evaluation of an interactive application is therefore obtained by identifying and focusing on different analysis dimensions, which may be addressed in different evaluation phases. For each analysis dimension, specialized conceptual tools, i.e., Application Model, Usability Attributes and Abstract Tasks must be defined. Following the conceptual tools are shortly explained:

- *Application Model:* a notable feature of SUE is the adoption of design models for describing the application and for precisely identifying and naming the relevant objects of the evaluation. Models also support the organization of concepts and drive the overall evaluation process;
- *Usability Attributes:* identify specific usability properties that an application should satisfy, in order to be usable. They are obtained by decomposing general usability principles into more specialized, fine-grained usability criteria, which specifically address the application features falling into the chosen dimension of analysis.

- *Abstract Tasks;* Abstract tasks can be executed on using a final application, a running prototype, or a set of design specifications. They are "inspection patterns", each one focusing on a specific feature of the application (e.g., the synchronisation of multimedia data, the navigation of a guided tour or a table of content, indexes, etc.). The inspector gathers the usability problems she identifies by performing each abstract task, in order to judge the usability of the overall application, or of some specific aspects. The use of abstract tasks makes the inspection activity more structured, better organised, and helps an organization to standardise and compare the inspection results of different inspectors.

Content Evaluation

For information intensive interactive products, the approach to inspection can also adopt methods of content analysis and communicability evaluation. The objective of content analysis is twofold:

- inspecting the quality of content allows detecting quality breakdowns in the communication;
- content evaluation methods suggest guidelines for designing usable content.

From a communication perspective, the standpoint of methods for content evaluation is focused on the belief that the "happiness" of a communication act must be assessed by a receiver's point of view. Therefore, especially when dealing with content (i.e. coping with the notion of meaning, sense and relevance), the inspector has to take into account that addressee as the starting point and the target of the whole communication effort. Content should not be primarily intended in its technical sense (e.g. image size, length of pages, colour of icons), but it should be addressed as a designed set of ideas and messages conveyed through structured interactive possibilities. The main methods in these fields are:

- *Content Analysis:* Content analysis offers a set of conceptual tools for assessing the effectiveness and the quality of communication of a web application (from navigation to content).
- *Content Evaluation:* Content evaluation of electronic sources relies on the same principles as evaluation of a print source. Content evaluation is performed with a checklist for the five criteria: authority, accuracy, objectivity, currency, and coverage.
- *Criteria for the Evaluation of Internet Information Resources:* The criteria for evaluating Internet information resources is an attempt to amalgamate and assimilate criteria from several sources that can be applied for evaluating and selecting Internet information sources.
- *Internet Information Evaluation Form:* The criteria for evaluating Internet information resources is an attempt to amalgamate and assimilate criteria from

several sources that can be applied for evaluating and selecting Internet information sources.

- *Quality of Internet Information Sources Criteria Questionnaire:* The criteria for evaluating Internet information resources is an attempt to amalgamate and assimilate criteria from several sources that can be applied for evaluating and selecting Internet information sources.

Other inspection methods

There are several other inspection methods.

Formal Usability Inspection

The Formal Usability Inspection method was developed to help engineers to efficiently review the users' potential task performance with a product. The method is based on a formal inspection process consisting of six steps for the detection and description of usability defects. A formal usability inspection consists of one phase where the inspectors work alone. For each defined user profile and task scenario combination the inspectors take the role of the specific user and work through the tasks described in the task scenario. Usability defects are logged on defect logging forms. In addition a task performance model and heuristics are applied to detect defects. Afterwards, all inspectors come together to a logging meeting to aggregate their defects and to find more defects.

Inspection and Design Review

Inspection and Design Review is a general framework for user interface inspections which takes explicitly into account the purpose and the focus of the evaluation. The domain of concern and the depth of the inspection is determined before the inspection starts. Inspections are performed either individually or in groups. The inspection process can be more or less structured. The results are usability problems detected during the inspection and recommendations how to solve them.

Software Inspection

Software Inspection is a technique used to detect defects in software components or finished software products. The objective is to test the minimum requirement: Is the software (or a software component) free of bugs/errors? The domain of concern and the depth of the software inspection are determined before the inspection starts. The procedure for carry out the Software inspection is:

1. The Quality Manager checks if the software is ready for inspection and determines the objectives for the inspection.

2. The moderator plans and prepares the inspection on the basis of instructions received from the Quality Manager. He may also use information from previously executed inspections.
3. During a kick-off meeting the moderator explains the objectives of the inspection to the experts and provides them with the software to be inspected.
4. The experts test the software, log the defects they find, and prepare for the defect logging meeting.
5. During the defect logging meeting the defects found by experts are summarized. The severeness of defects is assessed. Finally, a causal analysis of defects and solutions to prevent the most important defects will be performed.

2.3.3. Two Techniques: Scenario-Based and Heuristic Evaluation

Within these two categories (User Testing and Inspection Methods) the most current usability evaluation techniques for web applications are alternatively based on two main approaches:

- The *heuristic-driven evaluation*, which provide checklists and usability principles for the expert reviewers (Nielsen J.: 1999);
- The *task-driven evaluation*, which provides sets of tasks guiding the user testing, walkthrough and other inspection techniques (Rosson M.B. et.al.: 2002) (Brinck T. et. al.: 2002). Normally, the evaluation based on tasks is used within a scenario. A scenario is the description of a concrete episode of use of the application (Cato, J.: 2001) and help to understand *stories about* use (Carroll J.: 2002).

The two techniques, emerging from different research traditions and usability practices, are often employed separately or alternatively, thus losing the opportunity of gaining a more complete and effective evaluation.
Basically, the main drawbacks of an evaluation based on *heuristics* only are two:

- Usability principles inspiring the reviewer are very good for detecting problems but provide poor design suggestions for the re-design. Actually, appropriate and focused design interventions do not derive directly from the evaluation results;
- Heuristic is very effective for measuring usability qualities of the site but captures very hardly the evaluation of complex scenarios. In fact, the "usability dynamics"

of the web site, i.e. the application behaviour when trying to accomplish real tasks, is often out of scope for a heuristic-based evaluation.

Task-driven usability techniques have on their part a couple of disadvantages:

- Scenario-based approaches can easily detect the feasibility of a task, i.e. whether a task can be actually accomplished or not; however, current techniques do not identify what exactly caused the failure or the success of the task;

- Task failures are hardly mapped on infringements of usability principles. Consequently, the lesson-learned does not enrich coherently the usability body of knowledge available for future evaluation experiences.

One of the main disadvantages shared both by heuristic evaluation and by task-based techniques is that they are not reuse oriented, i.e. they have not been defined to be effectively reused by the people who did not invent them. Most of usability techniques are proprietary methods or guru-dependent techniques; in other words, they are hard to be used by less-experienced evaluators because they do not provide them with the necessary conceptual tools to gain appreciable results. The problem of re-use is strongly connected to the difficulty of teaching and communicating the essence of a method in a way that also other people can apply it successfully.

It is possible to summarize the methods and techniques for evaluating applications as follow:

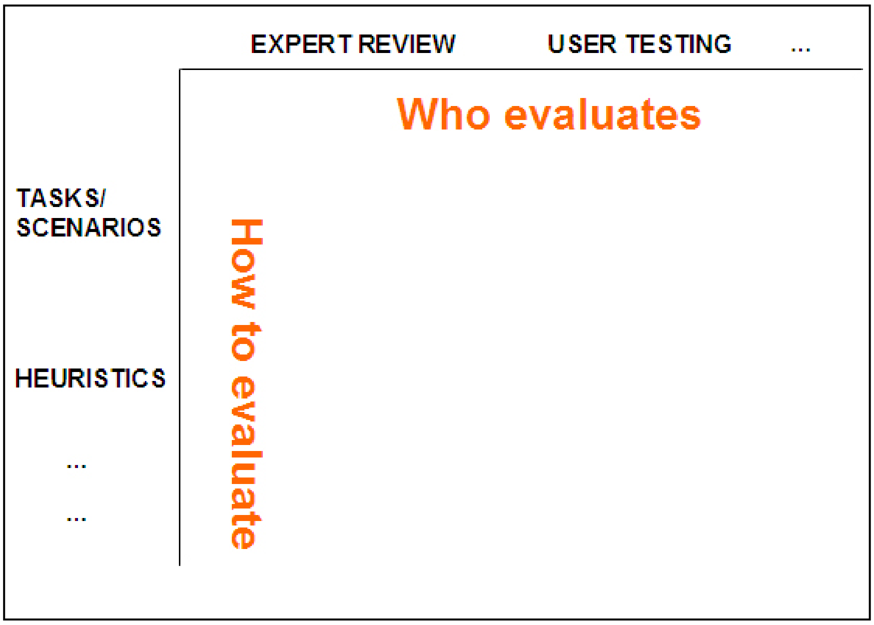


Figure 4: Methods and Techniques for evaluating applications

As before presented there are two main methods for assessing the usability degree of an application: Inspection methods (or Expert Review) and User testing. For carrying out the evaluation using these methods, two techniques can be used: Tasks-driven (or scenario-driven) and Heuristics. Clearly it is possible to combine the different methods and techniques (e.g. performing an Expert review using heuristics and scenarios or carrying out a User testing applying heuristics).

2.3.4. Automatic methods

A third way for evaluating the usability of a web application is representing by automatic methods, which measure the usability by running a user interface specification through evaluation software. The literature (Nielsen J. et al.: 1994) suggests that this approach do not work, for the reason that, until this moment, it is very difficult to create a software that it is able to capture all the usability problems that refer different levels (cognitive, navigation, content…). Most methods for evaluating web site quality assess static HTML according to a number of pre-determined guidelines, such as whether all graphics contain ALT attributes (Ivory, M., Hearst M.: 2002). Another example (Chi et al.: 2000) is represented by a simulation for generating navigation paths for a site based on content similarity among pages, server log data, and linking structure. Neither of these approaches account for the impact of various web page attributes, such as the amount of text or layout of links (Ivory, M., Hearst M.: 2002). In general usability aspects such as consistency and information organization are unaddressed by existing tools. In general, automatic methods are based on several sets of guidelines that are useful to measure page performances, to check the links' quality, for verifying the quality of HTML code, but some experiments (Ratner J. et al.: 1996) have shown that, for example, that HTML guidelines themselves have little consistency. However, automatic methods are a good complement to standard evaluation techniques (inspection methods and user testing) not a substitute.

2.4. Usability and Accessibility

The term "accessibility" has a generic meaning: if we consider the definition taken from the dictionary, accessibility is "the quality of being accessible, or of admitting approach" (Webster Dictionary, 2004). Indeed, an artefact (i.e. a website, a book, a door) is accessible when a user can perceive it and in some way operate with it. For example, a door for being accessible must be first of all perceivable (I must see it), then understandable (I should understand its scope and how use it, i.e. how it can be opened, inwards or outwards) and operable (I should be actually able to open it, i.e. the handle should be easy to reach and grasped). Nowadays, the problem of making "things" accessible is even more important with the advent of ICTs (Information & Communication Technologies): in fact, many of the activities of our contemporary society are based on them. However, disabled people can hardly get access to these applications, since they have not been designed and optimized considering their needs.

Given these premises, Web accessibility is "The power of the Web in its universality". In web communication, "Access by everyone regardless of disability is an essential aspect".

Different kinds of people with special needs can interact with applications in different ways. There are disabilities concerning motor capabilities (for example users unable to use hands and, consequently, an interface pointer like the mouse). There are also disabilities concerning physical capabilities (visually-impaired or hearing-impaired users) or cognitive capabilities (users not able to receive, process and understand complex messages).

In particular, people with visual disabilities (blind or visually-impaired people) have difficulties in using the graphic interfaces of modern web sites. The World Wide Web Consortium (W3C) , that supplies the "strategic" guidelines for the development of web applications, has emanated a standard within the Web Accessibility Initiative (WAI), based on documents prepared by associations of visually impaired people.

The W3C accessibility standard gives simple though important rules. For example, every image has to have a caption and possibly a descriptive text; link names such as "click here" or instructions such us "click on the green button" should be avoided.

The W3C standard for accessibility has made the first fundamental steps to overcome the above problems and guarantee web access to visually impaired users. A set of guidelines have been defined and addressed to designers who want to make their site "accessible" for users with visual disabilities.

For example, a proper alternative text for each image is prescribed (the screen reader reads the alternative text so that a description of the image can be provided), and suggestions for correct contrast between the background and the texts are provided. Guidelines are also defined for designing tables on the web page that might be read by screen readers in a more meaningful way for the user. Besides specific and detailed

indications on in-the-small components of the page, guidelines for effective navigation and layout design are poor and often too vague. Especially with regards to layout and navigation, many of W3C recommendations need to be interpreted and expanded in new, more detailed guidelines, affecting the content and the design of the site.
The W3C standard is an important, though still inadequate, step to solve the problem4.
Anyone who sees a blind person using a screen reader can realize that although the site complies with the W3C standard, it is almost unreadable in practice. There are many reasons for this, the most important of which can be listed as follows:

- *Information Overload:* In the Web, pages are very complex; someone who can see immediately spot the part of the page s/he's interested in, whilst a blind person has to listen to the whole content before s/he can decide whether there's something that interests her/ him.
- *Complex layout:* A Web page doesn't only contain too many "items" of content: its organization also relies on the graphic; again, someone who can see will immediately access information "down on the left side", whilst a visually impaired user will have to wait until the voice of the screen reader reaches that point.
- *Long list of links:* Lists of items are practically unusable since again they rely heavily on the graphic (imagine a voice reading a list of 50 paintings: how can you choose one?).
- *The command "back"*, often used to resume the navigation from a previously-visited page, is very problematic for a user who is faced to listen again the whole.

In general, actually, accessibility does not guarantee usability. In other words, accessibility – as the one suggested by W3C - is a necessary but not at all sufficient condition for usability. Whereas accessibility is often interpreted as making things available and possible to use, usability has to do with supporting user's goals and with user satisfaction. W3C standards focus on having content and navigation "available" to visually-impaired individuals. But how about making content, site structure, and navigation satisfactory and usable for such users? The necessity to rely on the oral channel only (as it is for visually impaired people) deeply modifies and influences the interaction with the website. Accessibility should be defined and treat as a branch of usability: if we say that an application must be usable by all users, then users with disabilities must be included too. If this is deemed to be too difficult, then web designers and developers should clearly and carefully define which "user profiles" are meant to be considered and/or which are not.

Chapter 3:

From MiLE to MiLE+ method

Summary:

The goal of this chapter is to present the MiLE+ methodology. Before explaining in detail MiLE+ an overview of the MiLE model is provided. Indeed, MiLE+ is the evolution of the MiLE method and consequentially, it is necessary to understand its main features and problems. After the overview of MiLE, MiLE+ framework is presented in detail stressing its revolutionary characteristics with respect to its forerunner. In particular, the attention is focused on:

- Separation between application-dependent and application independent analysis;
- Scenarios as drivers of evaluation;
- Heuristics as tools of evaluation;
- Usability Evaluation Kits (UEIs);
- MiLE+ activities: Technical Inspection, User-Experience Inspection and Scenario-based user testing.

The Chapter ends with the explanation of the method's cost-effectiveness and reusability.

3.1 Overview of MiLE Model

In this section we shall briefly outline the main features of MiLE (Milano-Lugano Evaluation Method), i.e. our current approach to evaluation, to be developed and enhanced into MiLE+.

MiLE is the result of a joint cooperation between the Politecnico di Milano and the University of Italian Switzerland. It has already been used in several real-life situations for evaluating websites of different domains (Di Blas et al.: 2002; Bolchini et al.: 2003; Triacca et al.: 2003).

MiLE has been adopted as the basic tool for evaluating cultural websites by the EC funded project MINERVA (MInisterial NEtwoRk for Valorizing Activities in digitization; www.minervaeurope.org), coordinated by the Italian Ministry for Cultural Heritage and Activities. MiLE has also been adopted as a basis for the Thematic Network VNET5 (www.vnet5.org), another EC funded project devoted to providing support for user-centred product creation in Interactive Electronic Publishing.

3.1.1 Conceptual tools for usability evaluation

MiLE tries to combine the features of both user testing and systematic methods. MiLE aims at exploiting the benefits of both approaches, introducing the user testing at the end of the –less expensive- expert review process, guiding the users' inspection towards those aspects of the application that on the previous inspection phase had proved their weakness. MiLE also combines heuristics and tasks, asking the evaluator to "judge" different facets of the task s/he is performing. The tasks are "abstract", i.e. independent from a specific application; this characteristic, a very special feature of MiLE, makes them reusable in different contexts. The evaluation goes through the different levels of which an application is made, artificially separating them: although it is clear that at the end the website offers to its users a blending of all the ingredients of which it is made. It is very different to evaluate the site's structure from the content's quality or the graphic's appeal. Eventually, MiLE has developed a mathematical system to elaborate all the data (through a system of scores and weights) that makes the outcome of the inspection very clear and easily comparable. The distinctive features of MiLE can be summarized as follows:

- Combination of systematic analysis and user testing;
- Task-driven analysis;
- Separation of different levels of analysis (content, navigation, graphics…);
- Scenario-based analysis for the content level (the communicative core of the application);
- Heuristic-based analysis of the tasks;
- Numeric elaboration of the results through scores and weights.

In the following paragraphs we will examine in detail all the above characteristics of MiLE and the research issues to be explored to expand MiLE into MiLE+.

Abstract tasks and Concrete tasks

MiLE adopts the task-driven "philosophy" of analysis, introducing two basic concepts: *Abstract Tasks* (ATs in short) and Concrete tasks (CTs in short).

Abstract Tasks

They are a list of generic actions (generic in that they can be applied to a wide range of applications) capable of leading the inspector through the maze of the different parts and levels of which an application is made, drawing the inspectors' attention toward to the most relevant features of the application.

Concrete Tasks

They are a list of specific actions (specific in that they are defined for a single application), which users are required to perform, while exploring the application during the empirical testing.

Different levels of analysis

MiLE artificially separates different levels of analysis (technology, navigation, content, etc.). Indeed, for example, an application might have a very good structure (navigation level) but, it can be very poor regarding the content. The MiLE levels are:

- *Content:* this level analyzes the quality of the content (in terms of efficacy and quality of the communication);
- *Services*: by services we mean all the functionalities a web site offers to its users; its analysis, for practical reasons, is often combined to the one of the content;
- *Navigation:* within the navigational dimension of a web application we distinguish two basic "movements": (1) the different ways by which a user can reach a specific piece of information; (2) the connections for passing from a specific piece of information to another;
- *Cognitive features of the interface:* the user perceives, understands and remembers the contents and the structure of the application. The choices of the designers create a set of expectations for the user: how does the interface cause these expectations? Are they fulfilled?;
- *Esthetic/graphic level:* this level considers two aspects: the graphic design and the layout. By graphic design we mean colours, type of fonts, images,

etc.; by layout we mean the spatial distribution of the graphic elements in the page;

- *Technology level:* this level analyzes the technological performance of the application, in terms of compatibility with different browsers, the interaction between the web site and the remote database, the level of security of the server that hosts the web site, etc.

For each level a library of Tasks has to be prepared, in order to support the inspection. For some levels (e.g. graphics or navigation) the tasks can be at large independent from the specific application domain; for other levels (e.g. content) we shall have different tasks according to the application domain (i.e., specific tasks for the cultural heritage domain, for the e-commerce domain, and so on).

User scenarios

When it comes to the content's level, where communication issues are stronger, the concept of task is extended by (or rather "included" in) the more comprehensive concept of scenario. The U-KIT (that is, the Usability evaluation KIT) for the content's level consists of a library of scenarios portraying stories about use of the application (Rosson, M.B. et. al.: 2002; Cato, J.: 2001). It is possible to synthesize the concept of user scenario as follows:

User scenario = User profile + User Goal + Tasks

The expert will select those scenarios that are more relevant in relation to the site's goals. An example of very simple and basic scenario for the museum websites domain is "a tourist wants to plan a visit to the museum"; many tasks are implied by this scenario: the tourist will try to obtain information on the opening hours, the means to reach the museum, the ticket's cost, etc. This scenario is particularly relevant if the application's mission is to attract visitors to the real museum; if on the other hand the application's main goal was to be educational, then this scenario would loose its importance with respect to others.
Sketching and "performing" a scenario has many advantages:

- It helps spotting missing pieces of information;
- It helps spotting pieces of content which are irrelevant for all the significant scenarios implied by the site's goal;
- The performance of a scenario's tasks allows verifying its feasibility/ efficacy.

Obviously, it is unfeasible to define all the scenarios needed to cover in detail the whole spectrum of potential tasks that could be performed within the application. The evaluator will have to define the most relevant users' scenarios for the specific application he has to evaluate, trying to elicit the site's goal by interviewing the site's stakeholders, that is, all those who have an interest in the web-site: the client, the designers, the institution, the users, the sponsors, the competitors, etc. S/he will then select from the MiLE's list of scenarios ready-made for the application's domain those that best fit his purpose; if necessary s/he may also create new tasks "specially tailored" for the site s/he's facing. The description of user scenarios could have different levels of granularity, from generic to very detailed. However, a scenario should portrait the type of user, his goal and the task(s) necessary to achieve the goal.

Usability Attributes

In order to make the inspection's results more analytic, tasks are further evaluated through Usability Attributes; they are specific for each level of analysis, although sometimes the same attribute (slightly re-defined in its semantics, according to the new context) can be used for more than one level. Usability Attributes are usability heuristics partly assessed and valid for general interactive applications. Following an example of attributes' list for navigation and content is presented here below:

- *Content's level:* clearness, completeness, conciseness, richness, accuracy, currency;
- *Navigation's level:* effectiveness, orientation, accessibility, self-evidence, predictability, non-ambiguity.

3.1.2 The Process of Usability Evaluation

This section is devoted to briefly present the MiLE evaluation process. Indeed, the complete and detailed version is presented in Section 3.2.4 (*MiLE+ the process guide*).
The MiLE evaluation process is divided into seven main phases:

- *Shaping the Evaluation Usability Kit for the specific application under inspection*;
- *Modeling the application under inspection*;
- *Performing some selected tasks*;
- *Evaluating the tasks through usability attributes*;
- *Weighting the results according to user profiles, communication goals/requirements*;
- *Empirical testing (user testing)*;
- *Reporting the usability evaluation activity*.

Modelling the application under inspection
The inspector draws a high-level mental model – either informally or by adopting a semi-formal model– of the application under inspection. The expected output is represented by a general schema of the most relevant features of the level under inspection; for example, the content structure, the navigational capabilities offered, or the interface elements.

Performing the selected tasks
According to salient user scenarios, the reviewer selects relevant tasks and tries systematically to perform them on the site. For each task, the reviewer assesses whether or not it can be properly accomplished.

Evaluating the tasks through usability attributes
Inspectors score each usability attribute for each task. In this way, tasks are not only evaluated as feasible or infeasible. Tasks are assessed taking into account the different aspect of the application that might have an impact on the user experience. Attributes increase the accuracy of the inspection because they decompose the evaluation of a task in different usability concerns.

Weighting the results according to user profiles and communication goals.
Inspectors weight the score given according to the user profile and the goals of the applications. Low weight means low relevance for the user profile of the scenario; high weight means high relevance. Weights limit the subjectivity of inspection because they balance the general score of the attribute with the needs and expectations of a user profile.

Empirical testing (user testing)
To empirically validate the most critical tasks identified during the inspection a user testing is carried out in a usability lab. The user accomplishes several critical tasks and reports the results obtained. An inspector ensures that the user testing is carried out correctly and gathers the impressions, satisfaction and problems of the users. The expected output is a final usability report that shows the results obtained during user testing.

Reporting the evaluation results
In the final phase the inspector should draw a report, which highlights the problems of the application for each level of analysis; notably issues on the usability attributes and problems in performing the tasks. This document should summarize both inspection and user testing results.

3.1.3 Problems of MiLE

Despite MiLE provides innovative solutions for usability evaluation, it has several problems.

First of all it only investigates in depth content and navigations levels. Conceptually MiLE includes the existence of other levels of analysis (Cognitive features of the interface, Esthetic and graphic, Technology) but in practice it does not provide particular conceptual tool for evaluating these aspects.

The second problem is related to the use of scenarios only for evaluating the content level. In reality, the scenarios can be employed also for evaluating other levels (see 3.2). This conceptual approach is based on a false reasoning which states that only the content development is related too the requirements and goals of the application.

Another great problem is the reusability of navigational abstract tasks: they are too complicated to understand and to use by people who do not invent them. Indeed, they are too much related to the W2000 Design Model (UWA Consortium, 2002) which allows modelling hypermedia and information intensive applications.
In general, MiLE is still to complicate to learn and therefore its reusability should be improved.

The last problem is the lack of consideration towards user testing. Theoretically, MiLE includes the user testing, but practically it does not explain in-depth how to use it and its relationships with the inspection activity.

In general, MiLE can be improved in many ways and within the next sections we present the evolution of MiLE (called MiLE+) which tries to solve the aforementioned problems.

3.2 MiLE+ Method: a systematic approach to usability evaluation

This section is devoted both to present the evaluation activities proposed by MiLE+ and its new conceptual tools. Since MiLE+ is the evolution of MiLE it is important to point out that it keeps the conceptual architecture of MiLE, but it deepens, expands and introduces several concepts.

3.2.1 Separating Application-Independent and Application-Dependent Analysis

The first conceptual innovative feature introduced by MiLE+ with respect to other methodologies is the distinction between the application-independent and the application-dependent analysis. On the one hand, an interactive application can be evaluated from a technical and more "objective" perspective and to the other hand the evaluation can be situated in the context of use of the application. This approach is similar to the one proposed by MiLE, which distinguishes the analysis performed using Abstract tasks and Concrete tasks. However, in MiLE this approach stagnated at an early stage (see section 3.1.3), while in MiLE+ has been developed and systematically adopted. This part of the work is entirely reserved for presenting the reason why it is important to separate these two levels of analysis.

Application Independent Analysis

Every human artefact can be observed, analyzed and evaluated from an objective and generic point of view. For example, if we think about a chair, this should have some technical characteristics for making it usable (for example a comfortable back of chair, a stable bearing…). If we consider the chairs below, the first one (Figure 5) is a usable one from a technical point of view: it has a comfortable back, a stable bearing, and two relaxing arms: from an objective analysis it is not possible to state that it is not usable. On the contrary, the second chair (Figure 6) is not a usable chair: the support is not stable at all (just one leg), there is no a chair back, no arms.

Figure 5: a usable chair **Figure 6: a non-usable chair**

An interactive and multimedia application, in particular websites, can be analyzed from an objective point of view as well as a chair. Clearly they are products having different levels of complexity, but the conceptual approach to the usability evaluation could be similar. Indeed, there are technical usability aspects that can be evaluated independently from the application under analysis (the term technical is used in a broad sense, not only referred to the technology behind the application). Making an Application Independent Analysis means to analyse the features that can be evaluated even without knowing the purposes and the users of the application. There are technical aspects that should comply with general usability parameters (heuristics). In this sense, these types of features are related to design aspects that can be considered without involving the users in the design. In fact, there are several usable design strategies which could be used without thinking about particular users.

Let us see some examples of application-independent features in websites and related usability problems:

– *Background contrast:* independently from the type of website we are using, the contrast between the background and the text should allow the legibility of the textual content.

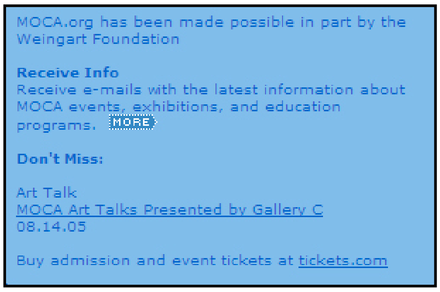


Figure 7: lack of contrast between background colour and font

This example, excerpts from MOCA website (www.moca.org) presents a lack of contrast between background and text. The low legibility of the text is a problem independent from the application we are using.

– *"Go back" (Backward Navigation) in the navigation starting from an index/list: when the user reaches a list of which s/he has to control the navigation while going from the starting index to each element and while going back from one element to the index.*

Figure 8: once the user reaches the painting s/he can not go back to the list

In the case of the Guggenheim Museum website (www.guggenheimcollection.org), once the user reaches the list of artworks now on view and select a painting (e.g. Georges Braque - Landscape near Antwerp) s/he reaches the selected page correctly. When the user tries to return to the list of artworks the backward mechanism is absent. The only navigational mechanism are two links called "Previous Braque work" (❶) and "Next Braque Work" (❷) that allow the navigation within a guided-tour of the Braque's work. Evaluating this navigational feature we do not consider the back of the browser, see that is a stand-alone application and it is not part of the website's design (besides sometimes the back of the browser has anomalous behaviour).

How to evaluate application-independent aspects?

The activity for performing the application independent analysis provided by MiLE+ is called Technical Inspection (detailed explained in section 3.2.3). The aim of MiLE+'s Technical Inspection is the identification of design problems and implementation breakdowns. The output of this evaluation is a number of "technical" problems that are application independent (e.g. the fact that the font size of a text is too small – graphic technical problem – it is a problem independent from the type of application).

Application Dependent Analysis

As previously described, it is possible to analyze the application taking into account the context of use of the application. During the Application Independent Analysis the inspector evaluate the application out of its context. On the contrary when he performs the Application Dependent Analysis he has to situate the evaluation within different scenarios of use (or situations of usage).

If we think on the chairs' example previously explained, it is possible to evaluate them taking into account the scenario of use. Shortly, the scenario of use of the first chair (the office's chair) is a situation where people need a comfortable chair (they have to stay sitting for more then 8 hours), a chair that can easily displace the people within the office, etc. Considering this scenario of use the first chair remains usable. The second chair (a milking stool), which is not usable from a technical point of view, is used in a very particular scenario: a farmer which have to milk several cows. Situating the chair within this scenario it is possible to state that it is a usable chair as well. Indeed, the milking stool allows the farmer to achieve his objectives. Even though the chair still remains lacking of technical usability and it could be improved.

Figure 9: a usable office's chair

Figure 10: a usable milking stool

An interactive application (more than a chair!) which addresses several users should be also evaluated taking into account the scenarios of use (the concept of scenario will be described in depth in the next paragraphs). During the Application Dependent Analysis the inspector has to determine if the user(s) are in the right conditions in order to achieve his (their) goals. Verifying the capability of the user to reach his/her goals means to answer questions such: Do people find the information they need? Are people properly driven and guided to a unexpected content? Is the content relevant to the user(s)? Is the content enjoyable/entertaining for the users?

It is also very important to evaluate if the application can be effectively used in a specific context (while driving, while at home, office, walking, visiting, etc.). Understanding users,

their goals and the contexts of use is essential to evaluate the application dependent usability.
For explaining in depth the features related to the application dependent analysis and relative usability problems, we present some examples:

- *Multilinguisticity:* the content addressing to different type of users speaking difference languages, should be given in more than one language. The multilinguisticity is a feature strictly related to the scenarios of use of the application and to its requirements. It is not possible to state that multilinguisticity is a technical usability feature, because the choice of implementing more then one language in a website is strictly dependent on its target audience.

Figure 11: collection's page of MEN website (www.men.ch)

In the "Musée d'ethnographie de Neuchâtel" (MEN) website (www.men.ch, Figure 7), most of information is provided only in French, even though it is presumable that the audience is not only local, but also an international one (one of the possible target is cultural tourists). The lack of multilinguisticity creates a usability problem related to the contents' fruition for a specific target (cultural tourist).

- *Predictability:* it is the capability of interactive elements (symbols, icons, textual links, buttons, images, etc.) to anticipate the related content and the effects of the interaction. The semantics and semiotics of the interactive elements (e.g. links labels) are strictly related to the type of users that will use the application. For example, if we develop a CD-Rom about Michelangelo addressed to children, the link labels should be understandable for the children (they should be able to anticipate the related content and the page they will reach).

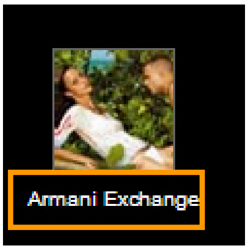


Figure 12: this label is not clear. Which is the content behind it?

Using the Armani website (www.armani.com, Figure 12) one of the link labels is called "Armani exchange". It is not very much clear which is the content behind this label. Only a user who knows Armani in-depth, knows that Armani Exchange is one of the Armani's Collection. From a usability point of view this becomes a problem if, for example, the intended users of the website are not only "Armani fans", but also people who is just curious (they do not have the background for understanding this label). Furthermore, this type of feature and related usability problem are strictly dependent from the type of application.

How to evaluate application-dependent aspects?

The User Experience Inspection and the Scenario-based User Testing are the activities for performing the application-dependent evaluation (they are explained in depth within the in Section 3.2.3). The User Experience Inspection is a scenario-based inspection which allows understanding the existence of application-dependent problems. This means that the evaluator has to imagine stories of use. For this reason, he has to set-up the "User Experience" KIT tailor-made for the application under analysis.

Advantages of separating Application Independent Analysis and Application Dependent Analysis

The necessity of separating the application-independent and the application-dependent analysis is related to the different typology of the problems and consequently to the needed resources for analyzing and correcting them. Performing an application-independent usability evaluation needs less time with respect to the application-dependent evaluation and provides more reliable results. Indeed, most results obtained during the application-independent analysis are almost unquestionable (for example an unreadable text is always a problem for the users independently from the application under evaluation). However, in accordance to the ISO 9241 definition the "real" usability evaluation is made performing the application-dependent analysis (both during the User

Experience Inspection and Scenario-based User Testing). Indeed, during this analysis we take into account particular users, trying to accomplish their goals in an effectively, efficiently and satisfactorily way in particular environment. However, it is important to point out that the evaluation process for analyzing, discovering and solving application-dependent problems, is more complex. Indeed, the problems' analysis and detection needs a preparatory phase for setting all the different tools (e.g. the creation of scenarios, the selection of the User Experience Indicators to use, etc.; in case of user testing the recruiting and the screening of the participants, etc.). These types of problems are strictly connected to the application's nature, its goals, its users and its domain and so the correction of these problems needs a deep work involving not only the development team but also other stakeholders (that's end-users, directors and managers). Therefore, the correction of these problems is a more complicated then the resolution of technical problems and it is more expensive in term of invested resources. Taking again the example of the MEN museum which is only in French (multilinguisticity user experience problem) the process of solving it passes through the director of the museum, the curator, the development team, the translator, etc.: the process needs a lot of resources.

Summarizing, the main advantage of separating application-dependent and application-independent analysis is the possibility to perform the evaluation taken into account two main constraints: resources at disposal (temporal and economical) and the knowledge of the application's domain (see section 3.2.5). It is important to underline that sometimes the results obtained performing the application-independent and the application-dependent analysis could be marked by conflict. Indeed, it could be happen that even though the inspector discovers technical issues, the evaluation of a scenario obtains good results. In this case, the inspector should also communicate the technical issues. In the chairs' example made earlier, even though the milk stool is usable in a particular scenarios, it would be possible to improve some technical features (e.g. the comfort). In the case of web applications even though a scenario is well judged, it could be possible that some technical problems are discovered (e.g. the font size too small). So, the inspector has to manage possible conflicts in the findings and s/he has to separately communicate the results of each activity and suggest the requirements for improvement considering these different aspects.

3.2.2 MiLE+ activities

As previously presented, MiLE+ proposes a specific activity, called Technical Inspection, which aims analysing the application-independent aspects and two evaluation activities for the discovery of application-dependent issues. These two activities are called User-experience inspection and Scenario-based user testing. It is important to highlight that MiLE+ is primarily based on inspection activities and the empirical test is an activity to support and validate the results obtained by inspections.

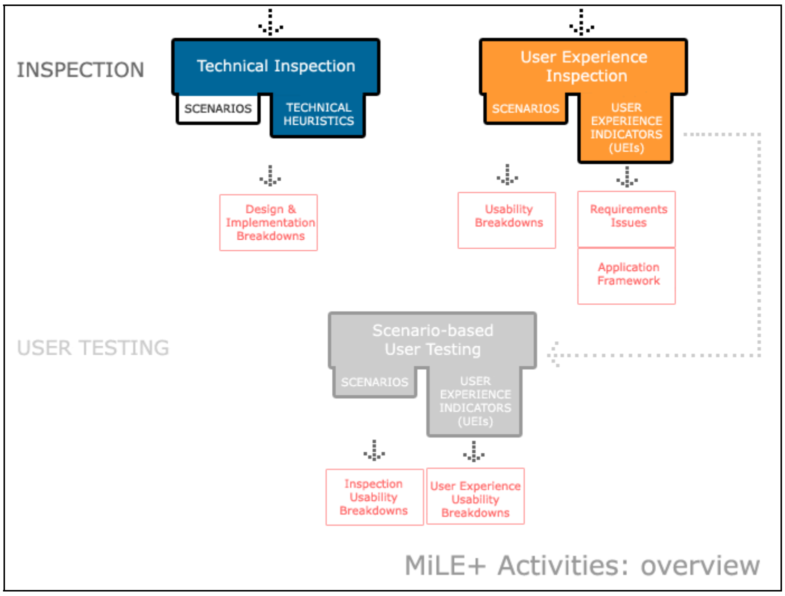


Figure 13: MiLE+ Activities' overview

Figure 13 does not illustrated the MiLE+ activity process (which is explained in the Section 3.2.3), but it only gives an overview of the activities, their relationships and the output. All these aspects are fully explained in this section.

Before going in depth in detailed descriptions of the MiLE+'s activities it is important to introduce its main conceptual tools which represents the "fil rouge" and the skeleton of the methodology. Indeed, there are three main concepts which pass through all the activities. So, it is very important to introduce them as "conceptual glossary" which allows a clear understanding of the activities. These concepts are:

- *Scenarios;*
- *Heuristics;*
- *Usability Evaluation Kits (U-KITs)*

Scenarios: the driver of inspection

As mentioned before, scenarios are "stories about use" (Cato, J.: 2001; Carroll J.: 2002), describing a typical user, one or more goals, and elements of the context of use (place, time, circumstances of use, etc.).
MiLE+ uses scenarios as the driver for usability evaluation because their role is at the heart of an effective usability evaluation. In fact, without a clear understanding of the need of the users and their goals, it becomes really difficult to perform a usability evaluation that may provide useful and in-depth results for a re-design process. The main objective for using scenario-based techniques is to help inspectors envision what could be the intentions and motivations of the final user when interacts with the website and what are the consequences, effects or impact of this interaction.
To this end, knowledge about the domain and the context of use of the application greatly facilitate the evaluators in this activity (Triacca L. et al.: 2004). Evaluators should therefore create scenarios which will use during inspection.
In general a scenario consists of three elements: *user profile, goal* and *tasks*.

User profile:
A User profile identifies a category of users having the same features and goals. Indeed, to define the user profile it is possible to use both socio-demographic features (e.g. age, job, geographic region, etc.) or "webographic"criteria (e.g. Internet knowledge, connexion speed, available technology).

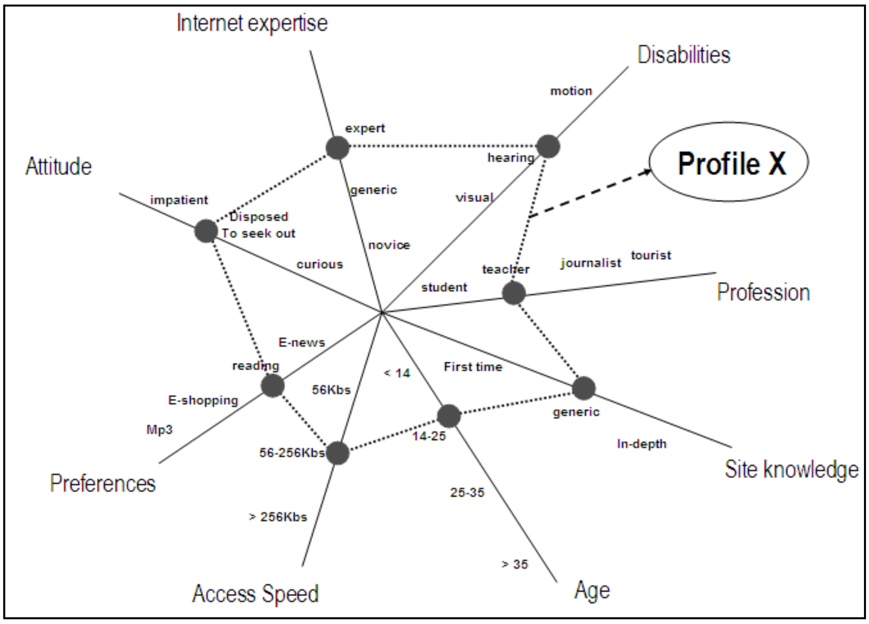


Figure 14 : visual user profile's representation

Goal

It is a high-level target of achievement for the user interaction. In other words, it is the motivation for using the application (e.g. entertainment, study, work, etc.).

A *series of tasks*

They are the actions that the user performs in order to fulfil the scenario. For example, the scenario "a tourist wants to plan a visit to the museum" will imply tasks such as "check the ticket costs", "check the opening days", etc.

Example of scenario for evaluating an e-learning web application:

User Profile	Goal	Tasks
Marc, 26 years old, he would like to take an online course. He uses frequently Internet but he has never used an e-learning application.	Know course conditions	– See course goals – See the course structure – See how to communicate with tutors and peers

Table 1: Example of scenario

Different levels of granularity

The definition of a scenario could have different levels of granularity: from the very high-level scenario (also called macro-scenario) to very detailed one (where goal and tasks are narrower in scope). The choice of the granularity's degree to use is related both to the goals of the inspection and the evaluation constraints (such as time and budget). Therefore, it is possible to associate one or more high-level goals to each user profile thus creating the essential constituents of a user scenario. Since we are here considering very high-level (or macro) goals, we will call these artefacts *macro-scenarios* (Table 2).

Macroscenario A	
User profile	**Student**
Macrogoal	**Plan the learning experience**

Table 2: Example of macro-scenario for an e-learning web application

A typical line of inquiry for eliciting macro-goal may be summarized in the question: "For which reason and motivation would a user use a certain web site?". Macro-scenarios capture a general target of achievement, which may be accomplished through several strategies or (sub)goals. Evaluators may thus identify more detailed scenarios that should take place in order to accomplish a given macro-scenario. This refinement process is usually led by questions such as: "How may the users (e.g. learners) accomplish the macro-goal? What should s/he be able to do to get to the macro-goal?. In this way, evaluators define new lower-level scenarios, which try to anticipate in a structured and organized way the expected user experience (Table 3).

Macroscenario A	Plan the learning experience
Scenario 1	Plan the study
Scenario 2	Know the course conditions
Scenario 3	Know the learning level achieved

Table 3: Refining macro-scenarios into scenario.

Now it is possible to define a series of *tasks* which describe the activities the user should perform on the website for each identified goal. Please, note that this refinement process from high-level goals to detailed tasks is not intended to be complete and exhaustive. Inspectors should identify the critical goals and tasks considered as important according to their evaluation experience.

Macroscenario A	Plan the learning experience	
Scenario 1	**Goal**	**Tasks**
	Plan the study	– Know the time required to take a course – Find the ideal period to take a classroom session – Know the time needed to download a document
Scenario 2	Know course conditions	– See the course goals – See the course structure – See how to communicate with tutors and peers
Scenario 3	Know the learning level achieved	– Make a test in order to verify the level of learning achieved – Verify in which topic there are gaps

Table 4: Refining scenarios into user tasks.

Finally, the result of the scenario definition is a structured set of tasks (and relative goals) associated to each user profile (Table 4).

Heuristics: tools for inspection

While performing each task by traversing and browsing the pages and links, evaluators do not only assess how effectively and efficiently they can complete the tasks and scenarios, but they are also supported by specific heuristics that guide the inspection to focus on the different application aspects that are relevant for the evaluation.
MiLE+ provides two sets of heuristics which should help the evaluation during the usability evaluation, called Technical Heuristics and User-Experience Indicators (UEIs):

Technical Heuristics
Technical Heuristics are a set of heuristics enabling to evaluate application-independent aspects of an application, that is the quality of the design (in all its aspects) and the spotting process of implementation breakdowns. Technical Heuristics are organized into design dimensions (e.g. content, navigation, graphics) and associate each design

dimension to a list of guidelines which help the inspector to analyze each dimension from a "design" perspective. For example, if the evaluator is examining the content of a page, he can use the "content heuristics" to evaluate if the content is well-designed from a technical perspective (e.g. the text is written in short paragraphs, easy to scan, accurate, updated, etc.). Furthermore, the inspector may suspend a purely subjective opinion (e.g. I like this text or not) and is guided in providing comments on specific criteria. MiLE+ provides technical heuristics concerning: Navigation, Content, Technology and Interface Design (comprising of graphics, semiotics and cognitive aspects of the interface). Actually the technical heuristics' library is composed of 36 navigational heuristics, 8 content heuristics, 7 technology/performance heuristics and 31 interface design' heuristics (a total of 82 technical heuristics and they are presented within Annex A_1).

Dimension		**Examples of Technical Heuristics**
Navigation		Consistency of the overall navigation
		Control of a guided-tour
Content		Text accuracy
		Multimedia consistency
Technology/Performance		System reaction to errors of a user
		Operations management
Interface design		
	Cognitive	Information overload
		Scannability
	Graphics	Font size
		Text layout
	Semiotics	Ambiguity of string of characters
		Conventionality of interaction images

Table 5: some examples of Technical Heuristics

User Experience Indicators (UEIs)

They allow evaluating application-dependent aspects. As previously explained, there are aspects of usability which cannot be evaluated by persons without knowing the purposes of the application. Such aspects are highly subjective and heavily dependent on specific user experiences (e.g. understandability, frustration, satisfaction, attractiveness, etc.). User Experience Indicators are the "measure units" to define these and other user experience factors. Therefore, they allow evaluating the quality of each scenario with respect to these user experience characteristics. In other words, User Experience Indicators allow anticipating the potential problems that end-users may encounter during his/her experience with the application. At the moment, the User Experience Indicators library is composed of a total of 20 Indicators (the complete list is presented within Annex A_2).

Categories of interaction	Examples of User Experience Indicators
Content Experience	Completeness
	Relevance
	Comprehensibility
Navigation & Cognitive Experience	Predictability of interactive elements
	Learnability
	Memorability
Interaction Flow Experience	Naturalness
	Engagement
	Recall

Table 6: Examples of User Experience Indicators

Notice that Technical heuristics as well as User Experience Indicators are conceived to guide pro-actively the inspector. Indeed, they "push" him/her to observe the specific facets of the application in a very analytic manner. Other methodologies, such us Nielsen Heuristic Evaluation, tend to be more passive. For example, the Nielsen Heuristics "Match between system and the real world" does not allow the inspector to precisely and quickly evaluate a facet of the application: the inspector each time has to relate and interpret the heuristic to the part of the application under analysis. The pro-activity of MiLE+ heuristics allows to be more precise in the judgment and to gain time (the inspector does not interpret each time the heuristic).

Being MiLE+ a method that can be used flexibly at different levels of granularity – according to the resources available to the evaluators (see 3.2.5)– heuristics may be considered not only necessary when performing detailed tasks. Inspectors may even evaluate goals and macrogoals using a subset of the heuristics (Figure 1). In this case, it is important to note that the time and resources saved have to be balanced with a coarser-grain analysis.

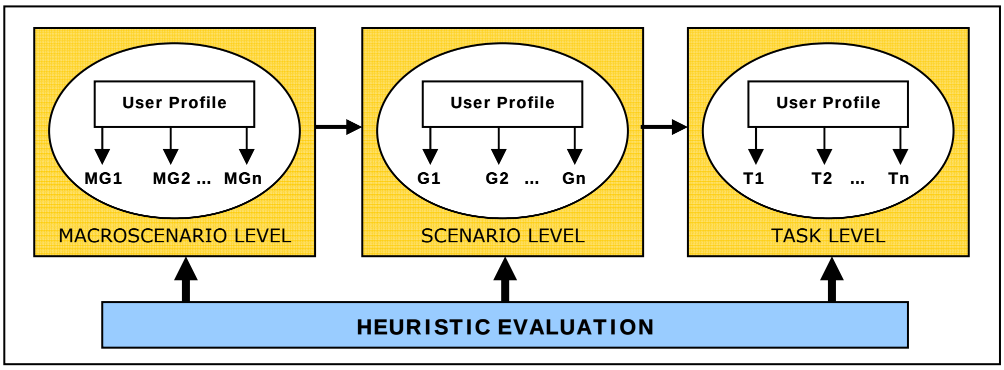


Figure 15: Using heuristics to analytically evaluate each scenario at different levels of detail (MG: Macrogoal, G:.Goal, T: Task)

The Usability Evaluation Kits (U-KITs)

To facilitate the inspection activity, and not to force inspectors to create the evaluation tools each time from scratch, MiLE+ offers a set of reusable evaluation tools (U-KIT, the usability evaluation kit). A U-KIT is a library of specific evaluation tools, which comprises a library of scenarios (User Profiles, Goals and Tasks) related to a specific domain. In addition the U-KIT also includes a library of Technical Heuristics and a library of User Experience Indicators.
In the case that a domain is not covered by the existing library of scenarios the inspector has to create one from scratch.
Note that all MiLE+ libraries are open-source, meaning that each evaluator could create, add or delete some elements with respect to his specific evaluation goals (e.g. it is possible to add a set of heuristics and/or some new User Experience Indicator, to create new library of scenarios for a specific domain, etc.).
Each inspector can set-up the U-KIT with respect to time and budget at his disposal. For example on the one hand, if he has very limited time and money, he could decide to analyze the application only using macro-scenarios and some technical heuristics; on the other hand, if he has a lot of money and a lot of time he could create several scenarios, using the complete library of technical heuristics and the library of User Experience Indicators (see section 3.3.3).

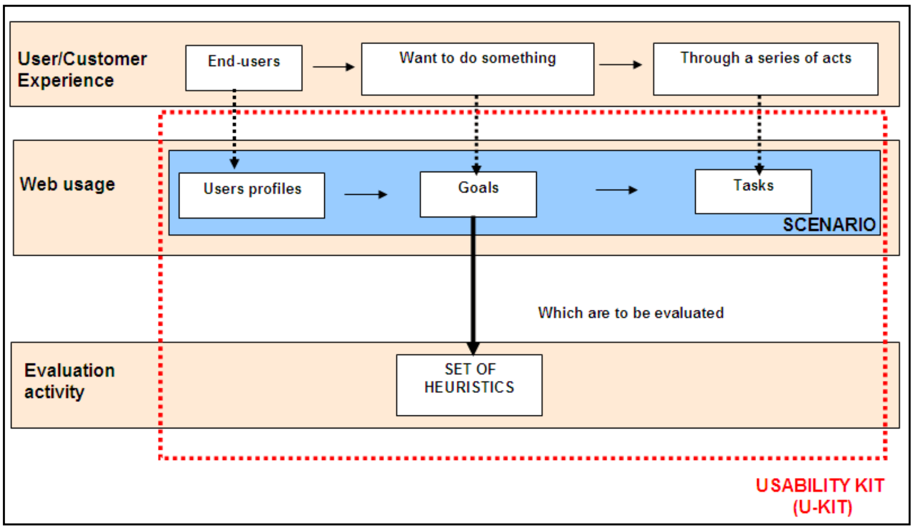


Figure 16 : the Usability Evaluation Kit (U-KIT)

3.2.2.1 Technical Inspection

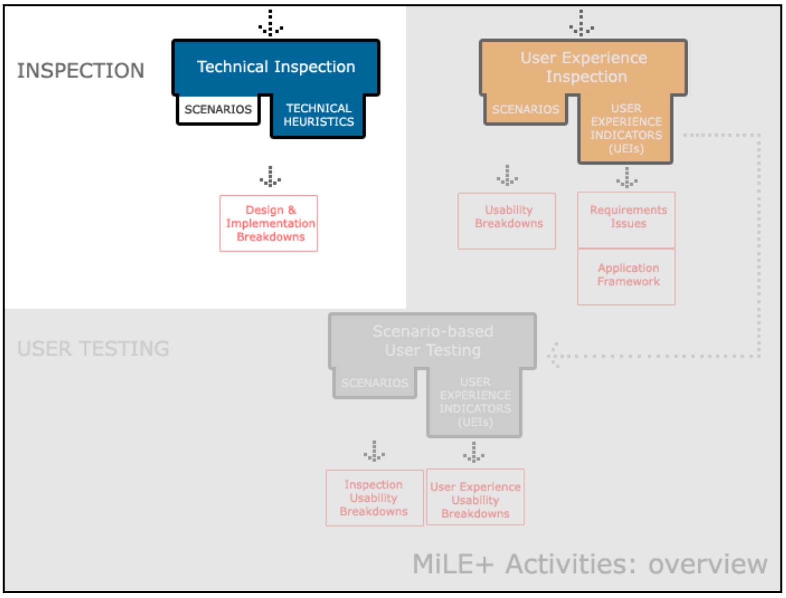


Figure 17 : MiLE+ Technical Inspection

As previously presented, MiLE+'s Technical Inspection is the activity for discovering application-independent problems. It aims at identifying design problems and implementation breakdowns. The output of this evaluation is a number of "technical" problems. During this analysis the evaluator examines the web application taking into account a number of design dimensions, assuming the point of view of the designer and not of particular end-users (like during the User Experience Inspection). Indeed, technical problems are not related to a specific user profile but to a "general user" (they are problems that affect the experience of all people navigating the website).
The design dimensions considered are *Content*, *Navigation, Interface Design (which includes Semiotics, Graphics* and *Cognitive aspects)* and *Technology.*

Content
The content level analyzes the quality of the content in terms of effectiveness of communication) and it verifies if the content and its structure correspond to the expectations of the users.

Navigation
Within the navigational dimension of a web application there are two basic aspects that could be analyzed: on the one hand the different ways for the user to reach a specific

piece of information (also called "access structures"); on the other hand, the connections for passing from a piece of relevant content to a related one.

Interface Design

The design of the interface is a broad dimension that comprises a number of aspects including:

- *Semiotics:* during the interaction with a website the user should easily understand the meanings of the messages proposed; this aspect is related to the content but is different from it. Content concerns the messages and information to design for the user, semiotics deals specifically with how the interface speaks of itself (e.g. link labels and words used by the interface). Having the same content, we can have a completely different semiotic strategy.
- *Graphics:* it studies two aspects: the graphic design and the layout. The graphic design related to choices bounded to colours, type of fonts, icons and other graphic elements on the page; the layout concerns to the spatial distribution of the graphic elements within the page.
- *Cognitive aspects:* observing the interaction with a website, two possible cognitive dimensions should be considered: the cognitive effort for the user while reading a single webpage and the cognitive aspects related to the understanding and memorizing of the information architecture.

Technology

Technology heuristics refers to those aspects related to technology choices and implementation style. The aspects that could be analyzed within this dimension are the formal correctness of the code (the site does not generate errors), the management of critical sections (e.g. operations & transactions), and the reaction of the system to user errors or unexpected user behaviours.

Setting-up the tools (U-KITs) for Technical Inspection

During the Technical Inspection problems are discovered using the heuristics checklists (selected from the library of technical heuristics, see Annex A_1) and scenarios: these two elements compose the U-KIT for Technical Inspection. It is important to point out that the use of scenarios is not mandatory. Indeed, there is no evaluation of the scenarios adequacy. However, they are useful to navigate with clear goals within the application (so the inspector can concentrate his evaluation on the most important parts of the website).

The activity of selection of the different tools composing the U-KIT (e.g. heuristics, scenarios, etc.) is very important as it is in this phase that the evaluator has to decide the tools-set he will use. The use of dimension as aspect under analysis and Technical

Heuristics as "unit of measure" is partially comparable to GQM model (Basili V., 1994), where the Goal is the dimension under analysis (e.g. Content, Navigation…) and Questions are Heuristics to be evaluated (e.g. How does the structural navigation works?). Considering that MiLE+ is a qualitative analysis metrics are not included.
Once the analysis is started the inspector has to complete an evaluation matrix, giving both a score for the selected heuristics (he has to decide the scale) and a comment for each score.

Example of Technical Inspection:

SCENARIO	
USER PROFILE	Art Lover
GOAL	Obtain more information on the museum.
TASK	Find information about the history of museum collection
SCENARIO DESCRIPTION	Joe is an art-lover. He would like to find some information about the history of a particular collection of the museum (e.g. paintings). He wants to know how the museum has acquired some artworks.

Table 7: Example of scenario (not mandatory for performing the Technical Inspection)

Dimension	Heuristic	Score	Comment
Content	Conciseness	3	The text is too long and it is not easy to read.
	Text errors	9	The text does not present errors.
Navigation	Accessibility of different pages in the navigation within the topic "Museum Collection"	9	All the pages of the topic "Museum Collection" are very easy to access.
	Orientation in three navigation	6	Sometimes it happens that if we pass from one section to another, we do not find orientation clues.

Table 8: short example of technical matrix (scale: 3 poor, 6 sufficient, 9: good)

As you can see in this example, during the technical analysis the inspector does not assess the quality and the adequacy of the scenario. In fact, he is concentrated only on the evaluation of technical features. The scenario is used as a tool which helps to concentrate the technical inspection on the most important parts of the application.

3.2.2.2 User Experience Inspection

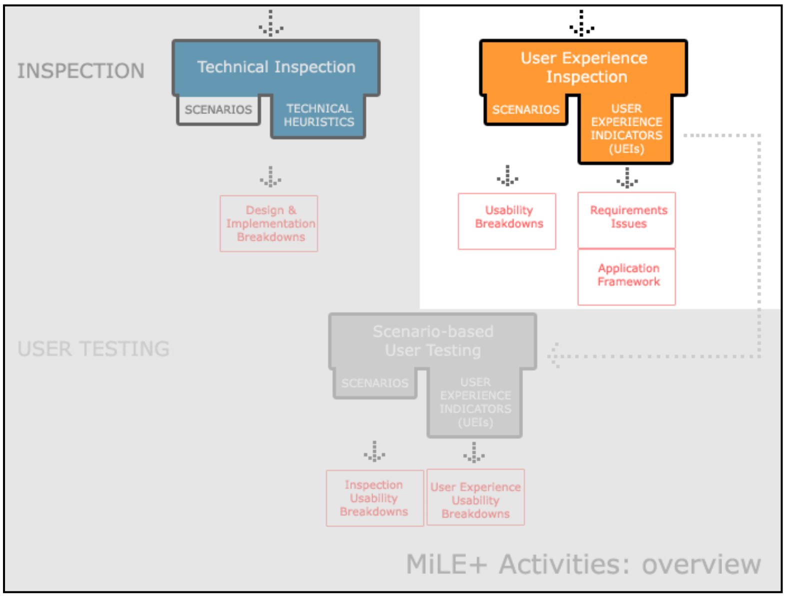


Figure 18: MiLE+ User Experience Inspection

As previously illustrated, the User Experience Inspection and the Scenario-based User Testing are the activities for performing the application-dependent evaluation. The User Experience Inspection is a scenario-based inspection which allows understanding the existence of application-dependent problems without involving end-users. This means that the evaluator has to imagine realistic stories of use. For this reason, he has to set-up the "User Experience" KIT tailor-made for the application under analysis. The KIT is composed by:

- *The scenario library*
- *the library of User Experience Indicators*

Creating and using the scenario library

As before presented, a specific scenario library should be created for each domain (e.g. banking web sites, e-learning web applications, cultural heritage websites, etc.). For creating a domain's library the inspector has to interact with different stakeholders: the client, domain experts, end-users, etc.: s/he has to create an application framework. For example, for creating the library for evaluating a museum websites the inspector should interview the Director of the Museum, he should organize a focus group with art's experts, a focus group with end users, etc.

Another complementary way for creating the library is called the "visioning technique" (Cato, 2001). The inspector has to imagine which ones are the main end-users, their goals and tasks: it is clear that this technique is more superficial (it is very difficult to create libraries without interacting with the stakeholders), but it can still generate reliable results in the case the inspector is an expert of the application's domain. At the end of these activities, the inspector has to refine the results of these interactions, building the scenarios and selecting the most important ones for the evaluation.

During the User Experience Inspection the evaluator has to put himself in the "shoes of the (different) users". This means that he has to examine the relevant scenarios using the User Experience Indicators. The users perceive a website using an "economic" cognitive approach. They understand that a website is composed of dimensions that are not the design dimensions. Indeed, it is not plausible that end-users have the ability of interpreting the website such as a designer, an engineer, etc. (that have a technical point of view). We believe that these dimensions of website's user perception are principally three: the content (e.g. texts, images, videos, etc.), the navigation and cognitive aspects (e.g. labels) and the interaction experience (e.g. satisfaction, engagement using the website). So, the evaluation criteria are divided in three categories corresponding to the different types of user interaction experiences. These categories are:

- *Content Experience Indicators:* measure the quality of user interaction with the content of the application;
- *Navigation & Cognitive Experience Indicators:* allow the measure of how the navigation works and the cognitive aspects of the application meet the cognitive world of the user(s);
- *Interaction Flow Experience Indicators:* allow the measurement of how the interaction with the application is appreciated by the users.

Example of User Experience Inspection

First of all the inspector has to check a list of UEIs concerning the different facets of usability/quality (e.g. richness, completeness, etc.). For each indicator (in relation to a specific scenario or task, it depends from the selected level of granularity), a score must be given. The output of this activity is a scoring matrix which reports the scoring (of each UEIs) and the result obtained by every task.

	UEIs				
Task: Find information about the history of museum collection	Predictability	Understandability	Richness	Comprehensibility	**Global Score for this Task**
Scores	8	8	5	6	6.75 (just average score)

Table 9: Short example of user-experience scoring matrix (scale: 3 poor, 6 sufficient, 9: good)

The second step of the User-experience inspection is called the *weighting phase*. In this phase the inspector has to establish the "real quality" of each critical task with respect to their relevance. After the scoring phase is over, the set of collected scores is analyzed through "weights" which define the relevance of each indicator for a specific user scenario. Weighting allows a clean separation between the "scoring phase" (use the application, perform the tasks, and examine them) from the "evaluation phase" in a strict sense, in which the applications' and the stakeholders' goals are considered. The result is final matrix that shows the overall results obtained by every task. This matrix reports the results according to the goals and the requirements of the application.

	UEIs				
Task: Find information about the history of the museum collection	Predictability	Understandability	Richness	Comprehensibility	**Global Score for this Task**
Scores	8	8	5	6	6.75 (just average score)
Weights	0.1	0.1	0.5	0.3	
Weighted Scores	0.8	0.8	2.5	1.8	5.9 ("weighted average")

Table 10: Short example of user-experience final matrix

Note: the Scale used for completing the analysis is:

- Scores: 0-10 (0: bad, 10 very well done),
- Weights: 0-1 (0: UEI not important; 1 very important. The sum of weights does not have to be more than 1)

The website under evaluation obtained a "pass mark" for this task (5,9/10). Analysing carefully the partial results, it is evident that both the richness and comprehensibility of the information regarding the collection's history should be improved (they are the more important UEIs for this scenarios – weight 0.5 for richness and 0.3 for comprehensibility – and they have obtained a quite negative judgement – 5 and 6).

3.2.2.3 Scenario-based user testing

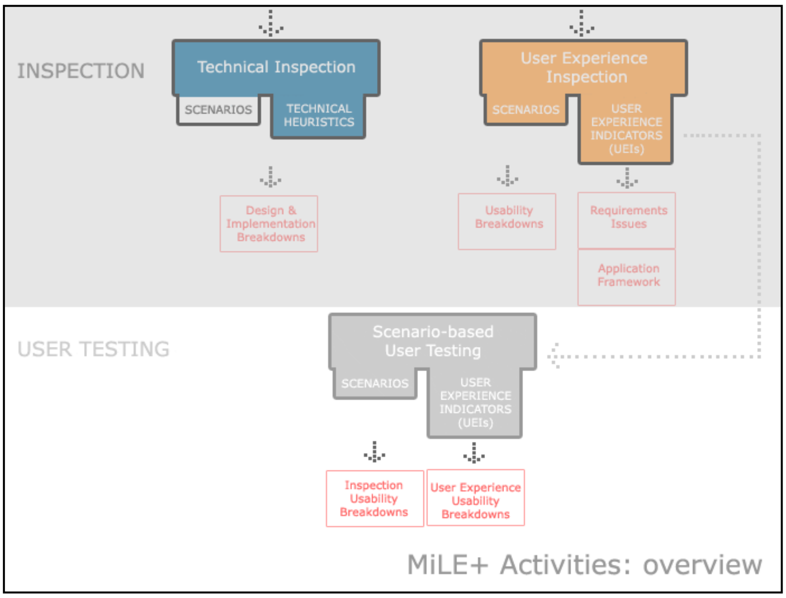


Figure 19: MiLE+ Scenario-based user testing

Even though MiLE+ is an inspection-oriented method, the role of user testing is become more important than in MiLE. In fact, *Scenario-based User Testing* is employed for the evaluation of application-dependent issues as well as the User-experience inspection. Indeed, the main goal of the Scenario-based User Testing is to empirically validate or invalidate the results provided by the User Experience Inspection. Furthermore, it helps reducing the subjectivity of the inspector's judgment. During the test the user accomplishes several tasks belonging to the critical scenarios identified in the User Experience Inspection. A test analyst (s/he could be the inspector who performed the inspections activities) controls that the user testing is carried out correctly and gathers the impressions, the satisfaction and the problems of the users by means of direct observation (recording or taking notes) and debriefing (questions, interviews, etc.).
Moreover, during the user testing the inspector has the possibility to verify the impact of technical problems (emerged during a previous Technical Inspection) on a sample of end-users. For example, if the inspector found a problem with the font readability he could verify the impact of this problem on the users.

For performing the Scenario-based user testing the test analyst has to select between 3 and 6 end-users per scenario (and user profile) under analysis. So, if for example after the Scenario-based inspection the inspector found three critical scenarios and considering that each scenario involves a user profile, the user testing has to involve from 9 to 18 end users. It is clear that this is a very expensive activity. The suggestion to employ between 3 and six users per scenario is based on the adaptation of Nielsen's rule (Nielsen, 1994) which states that 5-6 users are enough for the evaluation of an application.

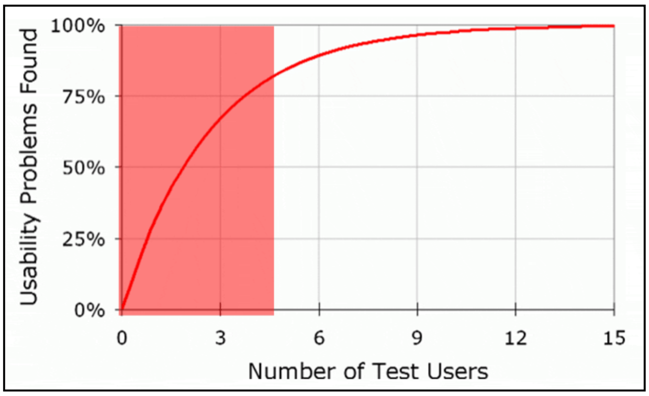


Figure 20: Nielsen's rule for test users' recruiting

There are two main styles for observing the users: *Pure observation* and *Insightful observation*.

Pure observation

During the pure observation the test analyst never intervenes. S/he lets the end-users freely accomplish the tasks and s/he takes some notes and records the test. Throughout the test the end-users have to verbalize their thoughts (thinking aloud).

Insightful observation

The insightful observation is a more interactive approach for conducting a user testing. Indeed, the test analyst is free to ask some questions before a user action (questions such as "I saw you clicked on this button, what do you think it is going to happen?") or after an action (Is that what you expected?). Besides this, the analyst has also the possibility to ask *why questions* (e.g. "Why did you click this button?").

For asking the questions (both written and oral) the analyst has to use an adapted version of the UEIs (which are thought for understanding the quality of the user interaction with the application). Indeed, during the User Experience Inspection the inspector evaluates the task using the UEIs. So for being comparable the User Experience Inspection and the Scenario-based have to use the same "unit of measure". For example if during the inspection we found a problem related to the richness of texts, in the test debriefing it is possible to ask a question with multiple answers, such as:

- *How do you find the texts? ☐ Rich, ☐ Very Rich, ☐ Not reach at all (select only one answer).*

3.2.3 MiLE+ framework at a glance

For concluding the presentation of MiLE+ main features it is possible to summarize its framework as follows:

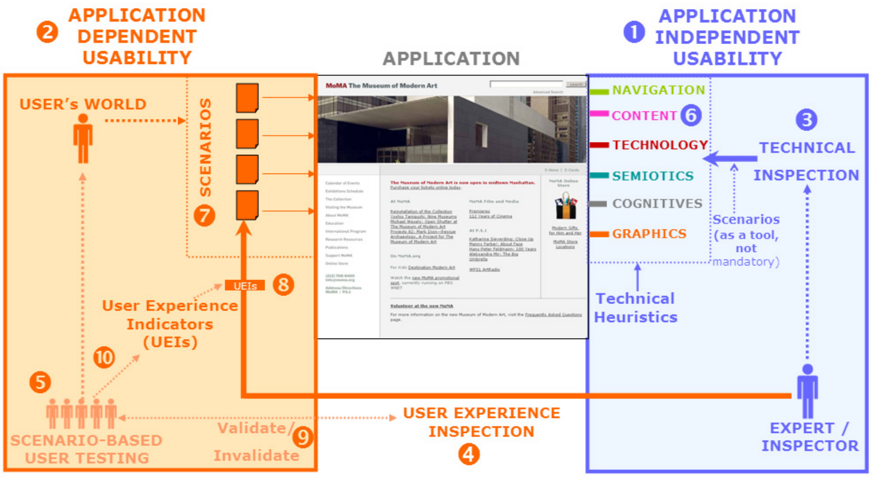


Figure 21: MiLE+ framework at a glance

First of all MiLE+ distinguishes the evaluation of problems related to the application (application-dependent usability ❷) from problems that are not related it (application-independent usability ❶). Starting from this distinction MiLE+ offers three types of evaluation activity, two inspections (called *Technical Inspection* ❸ and *User Experience Inspection* ❹) and a user testing (*Scenario-based user testing* ❺). *Technical inspection*, performed by an expert evaluator, uses technical heuristics, that are organized into the design's dimensions (❻), in order to evaluate the application-independent facets of the application. For carrying out the technical inspection the evaluator can be guided by scenarios (which help him to concentrate on the most important parts of the website) or s/he can navigate randomly (notice that during this inspection the inspector does not evaluate the adequacy of the scenarios). *User-experience inspection* is always performed by an inspector who tries to understand how the website is perceived by particular categories of users (user profiles). For this reason s/he has to put him/her-self in the shoes of the users. For understanding the user's world s/he employs several scenarios of use (❼) and for evaluating them s/he use the UEIs (❽). *Scenario-based user testing* is carried out to validate or invalidate the results of the User Experience Inspection (❾). Throughout the test a sample of end-users has to perform the most critical scenarios identified during the User Experience Inspection. The test analysis is performed adapting the UEIs (❿) both in the debriefing phase and for the problems' reporting (for example if the test analyst records a problem related to absence of texts in a specific language, s/he will report this problem using the UEI multilinguisticity or lack of multilinguisticity).

3.2.4 MiLE+: the process guide

The usability evaluation is a complex activity and there are several approaches and processes to perform it. In this section we present the suggested MiLE+ evaluation process, which is not prescriptive. Indeed, there are a lot of variables that could emerge during the evaluation activity and these variables make almost impossible to provide a rigid process to follow.
However, it is important to point out that the preliminary activities made by the inspector are the choice of the evaluation activities to perform (Technical Inspection, User Experience Inspection and/or Scenario-based user testing) and the creation of the U-Kit which s/he will use. So, the process passes through four main steps: preparatory phase, application-independent analysis (which is structured in sub-activities), application dependent (which is also composed of sub activities) and reporting.

Preparatory phase

In this phase the inspector has to select the tools and create the U-KITS that will use in the evaluation. In particular s/he has to:

- *Setting-up the Usability Kit for the Technical Inspection:* the activity of selection of the different tools composing the Technical U-KIT (in particular the heuristics) is very important since it is in this phase that the evaluator has to decide the tools-set he will use. S/he has also decide if s/he will use scenarios (in this case s/he has to create them).
- *Setting-up the Usability Kit for the User Experience Inspection:* creating the U-KIT for the User-experience inspection means to identify all the variables composing the kit, in particular the user scenarios (user profiles, goals, tasks and relevant usability attributes) and the User Experience Indicators (UEIs) that will be used during the analysis. The scenario should portrait the type of user, their goals and the task(s) necessary to achieve the goal. Sketching the relevant user scenarios for a specific application is a crucial phase; so the inspector has to:
  - *Identify the stakeholders of the application (clients, users, sponsors, etc.);*
  - *Identify their needs and goals;*
  - *Identify a sort of "ranking" of importance for the different goals (prioritize the goals);*
  - *Sketch the most relevant scenarios according to the application's goals.*
- *Setting-up the tools for Scenario-based User Testing:* as preparatory phase for carrying out the user testing it is necessary to prepare the material to recruit the

end-users (e.g. questionnaire), to perform the test (e.g. software, camcorder, questionnaire…) and for gather and analyze the results.

Application-dependent analysis: evaluating technical features

The Technical inspection has two steps:

1. *Performing the (selected) tasks within a scenario (not mandatory) or random inspection;*
2. *Evaluating the application using Technical Heuristics;*

1. Performing the (selected) tasks (not mandatory) or random inspection
The first activity performed by the inspector is to decide if he will perform an inspection using scenarios (which allows concentrating the analysis on the most important areas of the website) or carry out a random inspection (without taking into account particular task and scenarios). In the case s/he decides to use scenarios it is important to set them up in the preparatory phase.

2. Evaluating the application using Technical Heuristics
Once started the analysis the inspector has to fill out a technical evaluation matrix, giving both a score for the selected heuristics (he has to decide the scale) and a comment for each score (for an example of technical matrix see Table 8, pag.55).

Application-dependent analysis: evaluating the scenarios

The Evaluation of each scenario has four steps:

1. *Performing the (selected) tasks*
2. *Evaluating the tasks through User Experience Indicators (UEIs)*
3. *Weighting the results according to user profiles and communication goals*
4. *Performing the Scenario-based user testing*

1. Performing the (selected) tasks
The goal of this activity is to assess the feasibility of some "critical" tasks. According to salient user scenarios, the inspector defines a set of tasks and performs them on the site. For each task, the reviewer assesses whether or not it can be properly accomplished. The result expected for this part is a task list and a two-value mark for each task (YES: it can be accomplished, NO: it is impossible to accomplish it).

2. *Evaluating the tasks through User Experience Indicators (UEIs)*
The inspector should fill out the user-experience scoring matrix (Table 10, pag.58) reporting the score for each UEIs and the overall result obtained by each task.

3. *Weighting the results according to user profiles and communication goals*
This activity provides the final result for the scenarios' evaluation. Indeed, the inspector has to weight each UEIs' score taking into account user profiles and the communication goals of the application. The output of this activity is the final user-experience matrix.

4. Scenario-based user testing
As previously mentioned, this activity is useful to validate or invalidate the inspection's results using a sample of end-users.

- *Selecting and recruiting the users*: this activity is very important because the more the end-users are part of the user profiles previously defined, the more the test will give satisfactory results;
- Carrying out the tests: once selected the users it is possible to carry out the test (normally in a usability lab);
- Gathering and analyzing the results: once performed all the user tests it is very important to collect the results and analyze them;
- Reporting: the expected output of user testing is a report showing the results.

Drafting the final report

The final report should comprise the results of inspections and user testing activities. The final report should also highlight the requirements for the improvement of the application's usability. Normally, the report's structure suggested for the inspector that use MiLE+ is the following (it is possible to find detailed information about how to report usability within Annex A_4):

- *Executive Summary:*
- *Introduction:*
- *Results of usability analysis*
    - *Technical Inspection Results*
    - *User Experience Results*
        - *User-experience inspection*
        - *Scenario-based user testing*
- *Synoptic of results*
- *Requirements for improvement*
- *Conclusions*
- *Annexes*

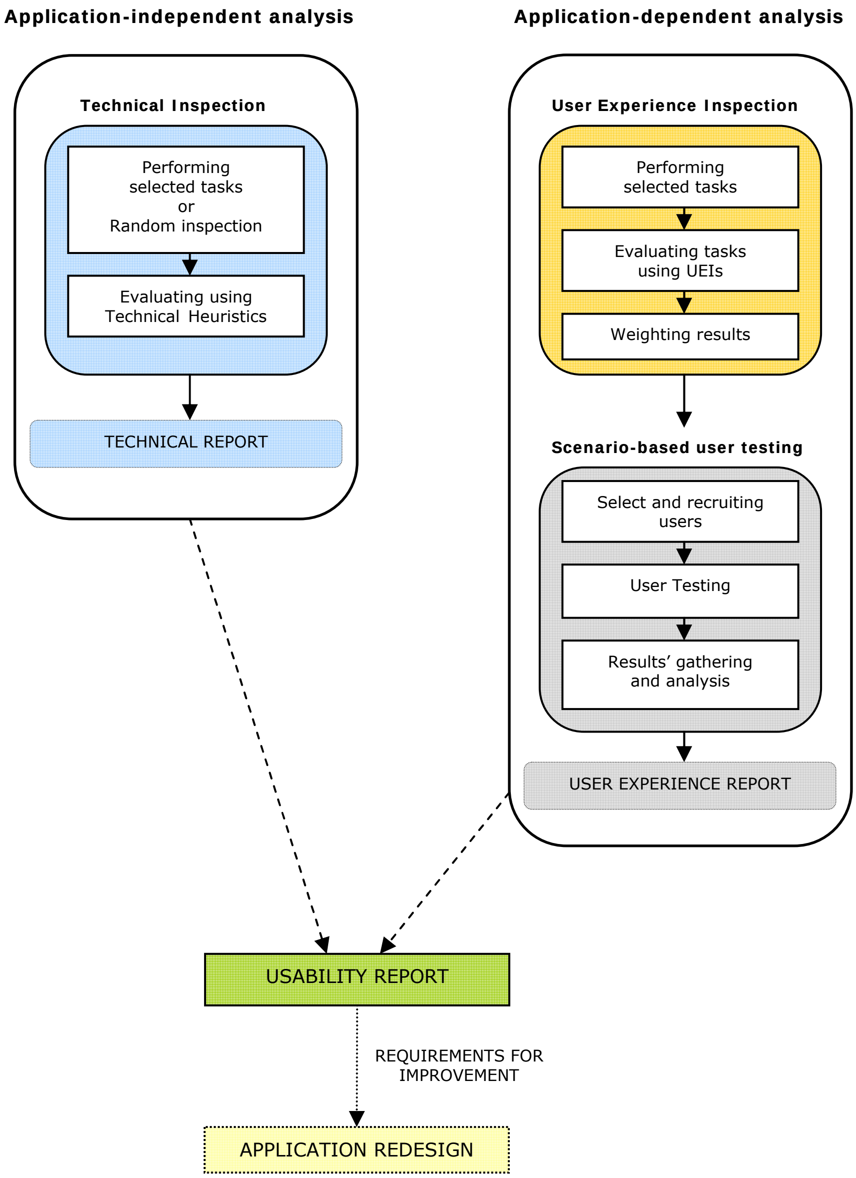


Figure 22: the process for evaluating usability

3.2.5 Cost-effectiveness of MiLE+

This Chapter stresses the fact that MiLE+ is a cost-effectiveness usability evaluation method. To be cost-effectiveness does not mean to be "cheap" as absolute feature or cheaper compared to other methodologies. Cost-effectiveness means to provide a flexible structure which can be adapted to project's constraints, in particular budget and time at disposal.

To represent the high-adaptability of MiLE+ considering these constraints, it has been created and employed an adapted version of the Boston matrix (normally used in the economic field for portfolio management). This matrix is called UMC matrix (Usability Methods' Cost matrix) and can be universally used for the costs of visual representation of usability methods.

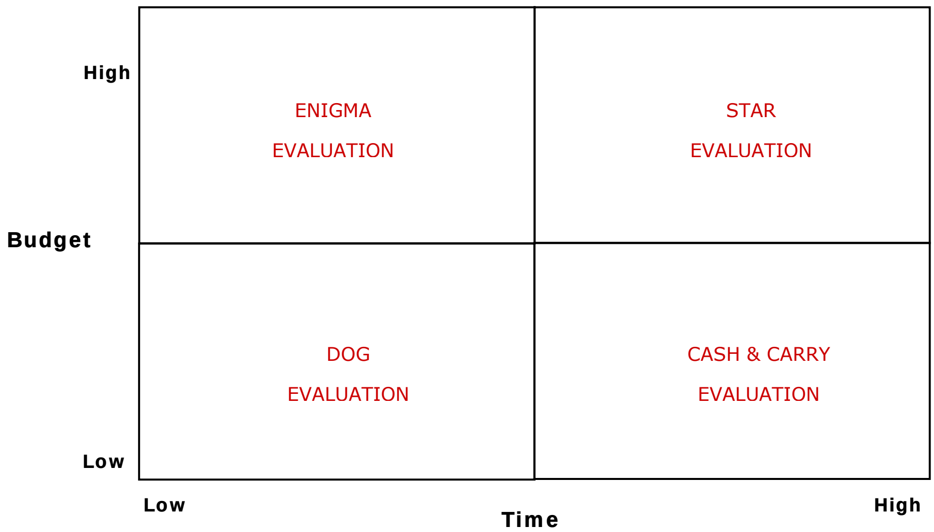


Figure 23: UMC matrix

Using the UMC boxes it is possible to classify all usability methods or activities according to two dimensions:

- on the horizontal axis: the time at disposal in order to perform the evaluation;
- on the vertical axis: the budget allocated for the evaluation.

By dividing the matrix into four areas, four types of cost-approaches can be distinguished:

- *Star evaluation*: it is the best context for performing the usability evaluation of a website, not only because of the fact that there is an high budget allocated and a lot of time, but also because this means that the stakeholders understand and are sensitive to usability and its importance;

- *Cash & Carry evaluation:* it is a difficult situation to manage. The project budget allocated for the usability evaluation is low or fairly low but the time at disposal is high. In this situation the usability expert(s)/consultant(s) has to carefully plan his/her activity. S/he has to produce a good result (s/he the time for providing a high quality service) without employing a lot of economical resources (e.g. to pay the end-users for the user testing).

- *Enigma Evaluation*: it is also a very tricky situation to manage. Indeed, the budget allocated is high but the time at disposal low. Also in this case the activities planning is the key for a satisfactory result. One suggestion would be to increase the human resources employed.

- *Dog Evaluation*: it is the least interesting condition. The budget and the time at disposal are low. The result will be a usability evaluation to macro-scenario level.

Using UMC matrix for representing MiLE+ activities

To show the cost-effectiveness of MiLE+ the activities are mapped within the UMC matrix (Figure 24).

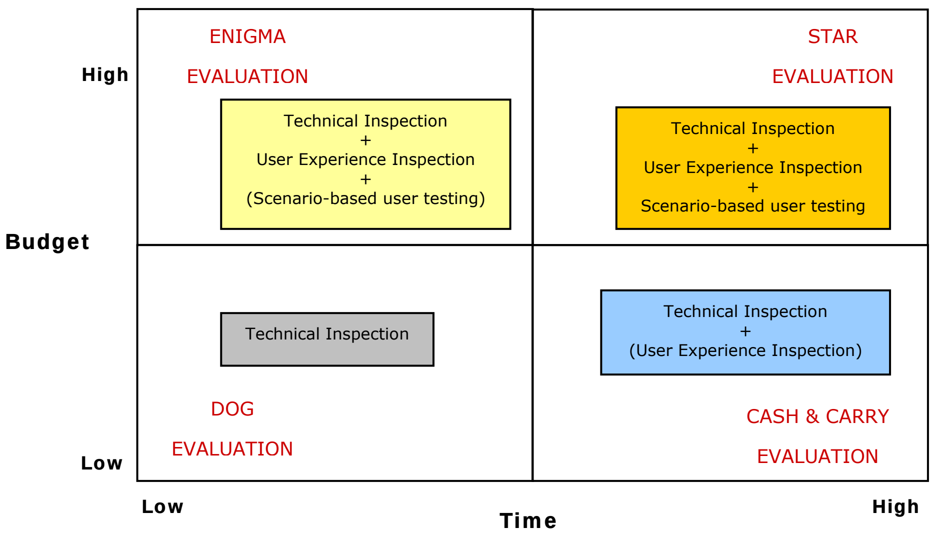


Figure 24: UMC matrix for MiLE+

The extreme cost-effectiveness of MiLE+ is demonstrated by the fact that it provides activities tailored made for the constraints of the project. So it is possible to classify the activities as follows:

- *MiLE+ star:* it is the suitable situation to exploit the efficacy and effectiveness of the method. Considering the time and budget at disposal, it is possible to employ all MiLE+ activities. So the evaluation starts with the technical evaluation, passing through a very detailed User-experience evaluation (using very detailed scenarios) and at the end the Scenario-based user testing. It is important to point out that part of the time should be devoted to the preliminary phase (U-KIT creation, selection of heuristics and UEIs and scenarios' definition);

- *MiLE+ cash&carry:* thinking about the fact that in the cash&carry context the budget is restricted, the suggestion is to perform a very detailed technical inspection and eventually a brief user-experience inspection (at macro-scenario level).

- *MiLE+ enigma:* it is a delicate situation. In this case, it is suitable to perform a very detailed technical inspection (which takes less time compared to the other activities) and a user-experience inspection. The suggestion would be to perform these activities at the same time employing two or three inspectors. It is also possible to carry out a short Scenario-based user testing, focused only on the two or three most critical scenarios. In this case it is important to employ at most three end-users per scenario;

- *MiLE+ dog:* in the case of a dog evaluation the suggestion is to perform only the technical inspection (without using scenarios). Indeed, this activity allows highlighting several application-independent problems in a short time.

Notice that the cost-effectiveness is also given by the ready-to-use tools provided by MiLE+, in particular libraries of scenarios, technical heuristics and UEIs. So the inspector can quickly adapt these tools to the evaluation s/he has to perform.

Chapter 4:
Teaching MiLE+

Summary:

This chapter aims at presenting the learning modules that allow teaching MiLE+ taking into account two main constraints: learners' profile and time at disposal.

The learning modules of MiLE+ are four:

- *Module 1: Introduction on usability;*
- *Module 2: MiLE+ inspections techniques;*
- *Module 3: MiLE+ scenario-based user testing;*
- *Module 4: MiLE + in practice*.

Notice that for each module the instructional design is presented.

The last part of the Chapter is devoted to present four types of MiLE+ courses:

- *MiLE+ full experience:* the complete experience*;*
- *8 MiLE+: usability crash-course:* a full-immersion course of height hours;
- *MiLE+ 4you: usability crash-course:* a four-hour-course to be introduce on the method *;*
- *MiLE+ online course:* the e-learning course.

4.1 The importance of teaching MiLE+

The reusability of a method is also strictly dependent on the teaching structure and instructional materials provided to people interested in using the methodology. Providing the learners with a well-structured course allow communicating and promoting the essence of a method and consequentially transfer the knowledge. Indeed, one of the main goals of this work is to create several *learning modules* and *learning paths* to teach MiLE+ taking into account two dimensions:

- *Learners' features:* in general every teaching activity has to consider the features of the target audience. In teaching a usability method is important to understand the learners' characteristics. Indeed, every usability course could be addressed to a great variety of audiences and therefore it needs to constantly adapt. On the one hand it is possible to have a professional audience, normally consisting of interface designers, engineers, developers, etc. On the other hand the audience could be made up of students at university level.

The Figure 25 shows the main target of the instructional activity.

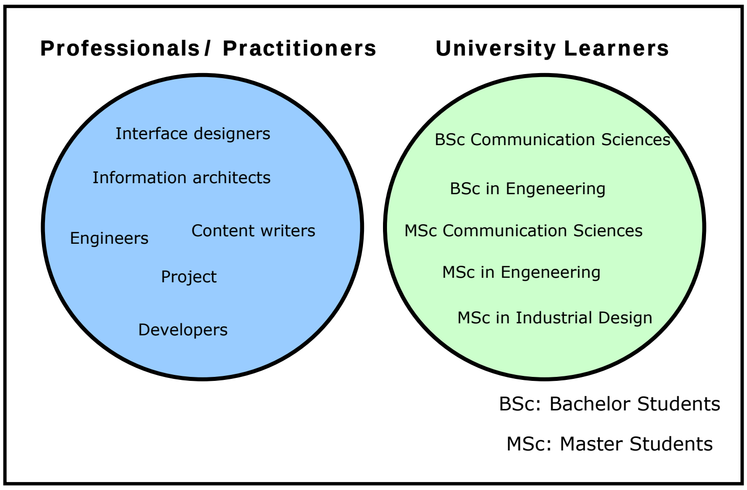


Figure 25: MiLE+ teaching - main targets

- *Time at disposal for learning MiLE+:* in every teaching and learning experience the time at disposal is one of the key variables to plan the activity. In this Chapter is presented a flexible and modular structure for teaching MiLE+. Indeed, the learning modules have been conceived for being adapted to the time at disposal.

The process of creating the learning modules and paths

The main effort to create the learning modules and paths has been concentrated on the development of the Instructional Design behind the different MiLE+ courses. According to Ragan and Smith (1999) it is possible to state that the instructional design is *"…the systematic and reflective process of translating principles of learning and instruction into plans for instructional materials, activities, information resources and evaluation".* The creation process for the MiLE+ instructional design started three years ago following an iterative process. As it often happens, the activity of teaching needs a continuous assessment. In these years, we have tried different teaching and learning styles and we have provided to the learners different types of learning materials. One of the key moments for the re-assessment of the instructional design has been the shift from MiLE to MiLE+ method. Indeed, MiLE+ was born with the main objective to be the reusable and "easy-to-learn" version of the previous method. Moreover, as presented in Chapter 3, MiLE+ introduces several new concepts and activities that should taken into consideration when teaching this new methodology.

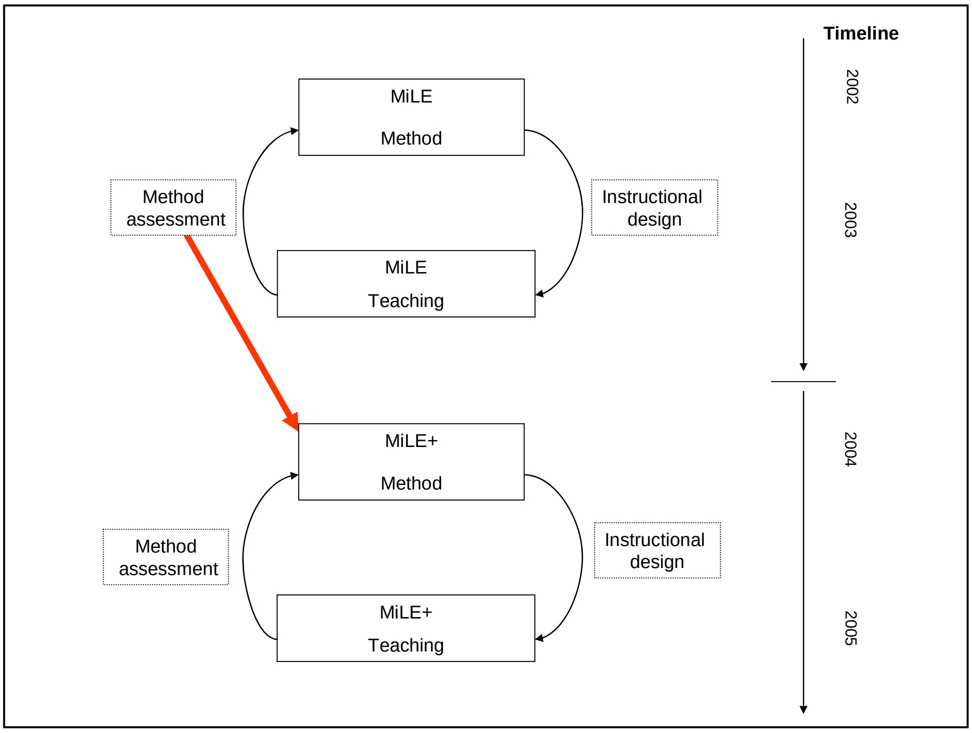


Figure 26: The iterative process for creating MiLE+ learning modules and paths

Figure 26 shows the iterative process followed since 2002. Until the end of 2003, we have been developed MiLE and consequently the method has been taught both to several university classes and practitioners. Every course gave us the possibility to re-think some aspect of MiLE; in particular the interaction with the learners provided us some interesting research issues. Indeed, at the beginning of 2004, after the 4th year's course "New Technologies for companies and institutions" (University of Lugano – Faculty of Communication Sciences), we had research issues emerging and we needed a further development of MiLE: MiLE+ (in particular the need of clearly separate application-dependent and application-independent analysis and consequently the creation of specific activities and tools for performing these analysis – see Chapter 3).

Communicating the Instructional Design

In order to effectively formalize, communicate and represent the instructional design of MiLE+ courses a very simplified version of *E^2ML* method (Botturi, 2003) and the *Quail* model (Botturi, 2004; 2005) has been used.

E^2ML – Educational Environment Modeling Language – is a visual modeling language for the design of educational environments in Higher Education. It is useful for representing the product of the design process: the educational activity or activities performed into an educational environment (Botturi, 2003).

The *Quail* model is a representation device for visualizing learning goals. It is a three dimensional grid representing the *type* of learning outcome, the *level* of knowledge and the *scope* of application.

The *type* of learning goals are taken from Gagné's classification (Gagné et al.: 1992), adding interpersonal skills. Types considered are *factual knowledge* (declarative knowledge, know-that), *concepts* (categories, types of objects, defined concepts, abstractions), *procedures* (steps in a process for accomplishing a task or achieving a goal), *attitudes* (dispositions to behave), *learning strategies* (meta-cognitive strategies, learning to learn), *interpersonal skills* (way of relating to other persons, communication skills).

The *levels of knowledge* are described both according to Lonergan's representation (Lonergan, 1997) and introducing new levels. The levels of knowledge that can be represented using the Quail Model are (Botturi, 2004): *experience*, *understanding commitment* and *action.*

- *Experience*: meeting, considering a possible object of knowledge, and perceiving a correct image of it, which becomes part of the learner's world.
- *Understanding*, which comprises:
    - Inquiry developing an interest and asking "What is it?" or "How is it?" concerning the potential object of knowledge.
    - Insight: understanding a single instance case, grasping the essence (pattern of intelligibility) of the object of knowledge as a single case.
    - Concept: through generalisation, induction and abstraction, conceiving a reusable and articulated formulation of what is understood. The

generation of concepts requires a (verbal) language or means of expression.

- *Commitment*, which includes:
  - Reflection: parallel to inquiry, asking "Is it so?" concerning the new knowledge. Although the word reflection was used for readability, it should be understood as critical reflection.
  - Commitment: assessing the value of the newly acquired knowledge as relevant to the learner's self.
- *Action:* including the new knowledge in the action, as integrated part of the learner's self. This means that after intelligence, freedom (or free will) should be put in motion in order to act the way one has learnt, and to realise the commitment (Botturi, 2004).

The last dimension is the *scope of knowledge*, which describes to what extent the new knowledge is expected to influence the learner in action. The metric is defined according to the performances in Merrill's matrix that are *Remember* (recall knowledge as such),, *Use* (apply knowledge to specific situations) and *Find* (exploit knowledge in order to generate new knowledge).

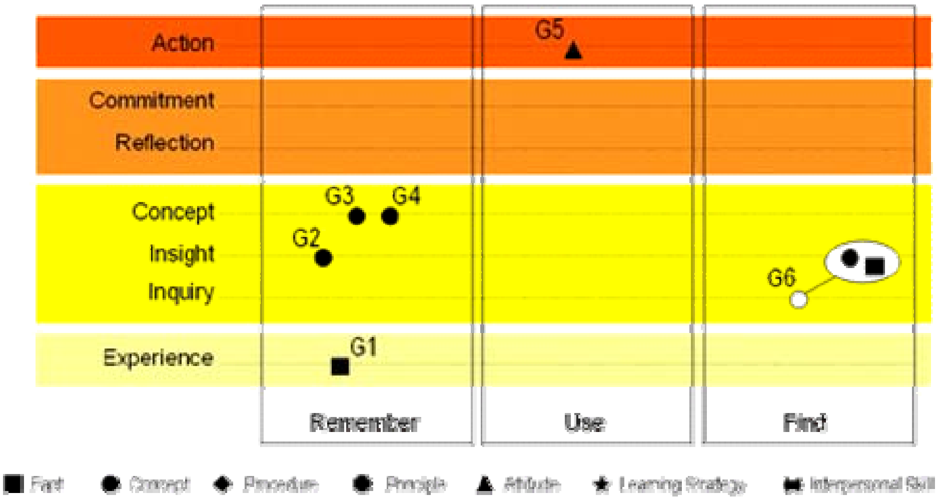


Figure 27: Example of Goals mapping using Quail

4.2. Learning@MiLE+: the modules

This section is dedicated to present the learning modules that allow teaching and learning MiLE+. The architecture and the instructional design for each module are independent from each other. Indeed, it is possible to think each module as a mini course on a specific concept, method or activity related to usability and MiLE+. However, there is a meta-architecture that allows creating a complete course taking into account the type of learners and the time at disposal (examples of these courses are provided on Section 4.3, pag. 90). The modules composing the teaching framework are four:

- *Module 1: Introduction on usability;*
- *Module 2: MiLE+ inspections techniques;*
- *Module 3: MiLE+ scenario-based user testing;*
- *Module 4: MiLE + in practice*.

Each module is presented using the same conceptual structure. In particular, the following structure has been used:

- *Goals Representation;*
- *Activity flow diagram;*
- *Goal mapping;*
- *Relationships Diagram;*
- *Resources list.*

Goals Representation

First of all, the goals of each module have been illustrated in a very detailed manner using an adapted version of the table provided by *E²ML* method.

GOAL STATEMENT						
TAG	**TEACHING STRATEGY ID**	**STATEMENT**	**TARGET**	**APPROACH**	**ASSESSMENT**	**IMPORTANCE**
<Goal tag>	*<ID of learning approach>*	*<statement of the learning goal>*	*<learners target of the learning strategy>*	*<learning approach >*	*<strategy for assessing the goal's achievement>*	*<goal's relative importance>*

Table 10: the detailed table used for presenting the learning goals

Activity flow

By each detailed table the activity flow for teaching and learning the module is presented. The activity flow diagram is the chronological way for teaching the module. It is important to underline that a module could have many flows (it is possible to have different chronological paths in order to teach the module). The diagram also shows the strategies for assessing the goals' achievement.

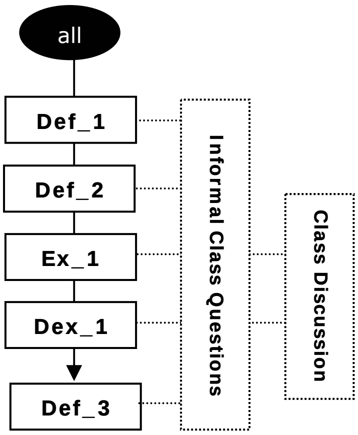


Figure 28: Example of activity flow diagram.

Goal mapping

As presented above, each goal has been mapped using the Quail matrix.

Relationships Diagram

In order to clearly explain the relationships between each module's goals a relationships diagram has been used. This diagram allows showing the influences between goals and the basic logical structure of the modules.

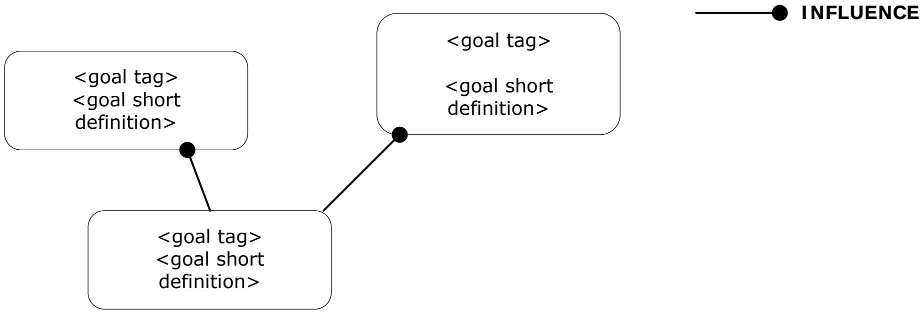


Figure 29: model of relationships diagram

Resources list

For each module the list of the learning resources has been provided.

RESOURCES			
NAME	**DESCRIPTION**	**TYPE**	**USE LOCATION**
<resource's name>	< description of the resource>	<type of resource>	<where to use the resource>

Table 11: the detailed table used for presenting the learning resources

4.2.1 Module 1: Introduction on usability

Goals and learning path

The Module 1 is propaedeutic as it is devoted to introduce the concept of usability and to create the background for understanding the other modules. The Module 1 starts with an excursus about the history and the social impact of usability. After a description on the main usability methods (divided in two parts: inspection and user testing methods), the attention is focused on introducing the differences between usability problems related to the application and its goals and problems that are not strictly dependent on the application under analysis.

To accomplish these objectives, the main learning topics are:

- *Defining usability (G1_M1 - Goal1_Module1);*
- *Introducing main usability methods (G2_M1);*
- *Showing Examples of poor usability (G3_M1);*
- *Presenting application dependent and application-independent analysis (G4_M1);*
- *Introducing MiLE+ (G5_M1).*

GOAL STATEMENT						
TAG	TEACHING STRATEGY ID	STATEMENT	TARGET	APPROACH	ASSESSMENT	IMP.
G1_M1	Def_1	Define Usability Describe the main phases in the history of usability	All	Definitions	Informal Class questions	3
G2_M1	Def_2	Be aware of the main features of main Usability Methods	All	Definitions	Informal Class questions	4
G3_M1	Ex_1	Develop a positive attitude toward usability evaluation and a critical approach to the applications' analysis	All	Examples of poor usability	Informal Class questions + Class discussion	4
G4_M1	Dex_1	Understand the difference between application-dependent and independent analysis	All	Definitions + Examples	Informal Class questions + Class discussion	5
G5_M1	Def_3	Introduce briefly MiLE+ method	All	Definitions	Informal Class questions	5

Table 12: detailed explanation of Module1 goals

Activity Flow

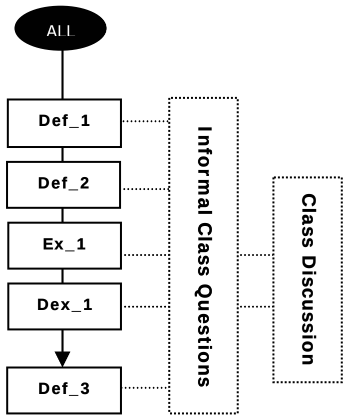


Figure 30: the activity flow of Module 1

The activity flow for Module 1 (and more in general for all MiLE+ modules) is concentrated on using alternatively concepts' definitions and examples for sustaining them. For example, once the instructor introduces the benefits of projecting human artifacts taking into account the usability, it shows examples of poor usability that have a negative impact on the user experience. In general, the assessment of the concepts is made asking informally to the learners some questions after each explanation (informal class questions). For assessing *G3_M4* and *G4_M4* it is possible to have a class discussion in order to develop a positive approach towards the advantages of thinking about usability and for verifying that the learners understand the difference between application-dependent and independent problems (which is the conceptual background necessary for understanding MiLE+).

Goals Mapping

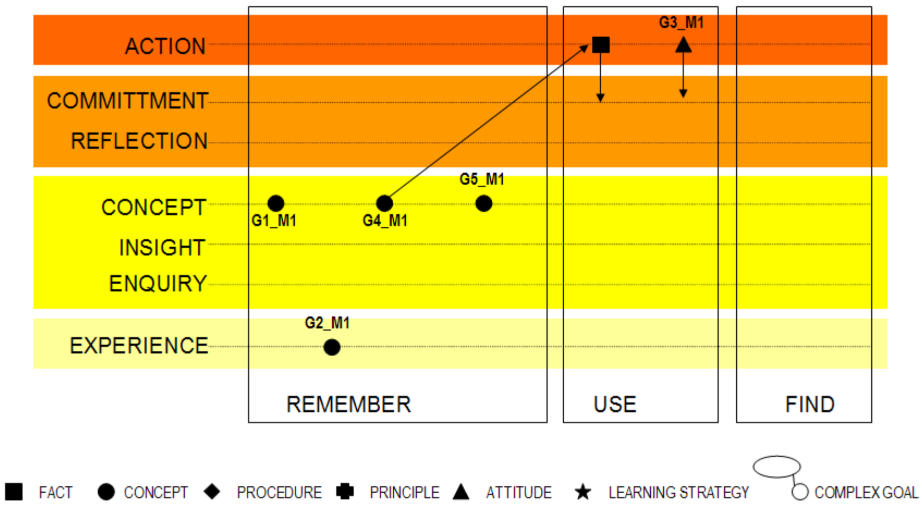


Figure 31: Module 1 goals mapping

Almost all concepts are in the *remember scope*, with the exception of *G3_M1* which has the purpose to commit the audience to the importance of usability. For reaching this goal the instructor and the learners have to comment together a series of examples. Notice the nature of *G4_M1* which first of all introduces a very important concept and then through some examples (also commented and explained by the learners) it should create a new perspective, committing the audience to the new point of view.

Relationships Diagram

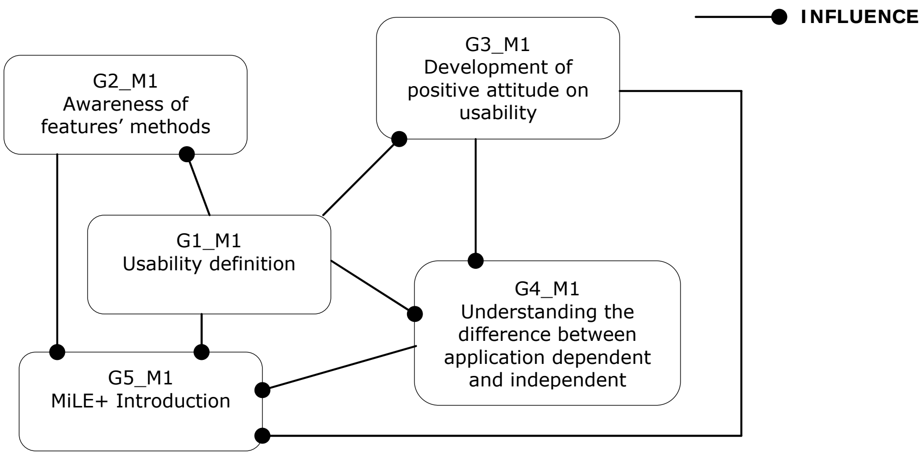


Figure 32: Module 1 relationships diagram

The relationships diagram reveals the importance of *G1_M1* creating the basis for understanding the other goals. Another important role is played by *G4_M1*: indeed, without the full comprehension of the difference between application dependent and independent analysis *G5_M1* could be hardly understood, and, likewise, the concepts and activities of other modules. Moreover, *G5_M1* is also influenced by the understanding of the main usability methods (as explained in Chapter 3, MiLE+ uses some approaches proposed by other methods). It is also important to point out that *G5_M1* is preparatory to the Module 2, which is devoted to explain in detail MiLE+. In general, each module finishes introducing the main concepts that will be the focus of the following module.

Resources list

The resources list for Module 1 includes the Slides pack 1 (see Annex B_1), which is used by the instructor for class lessons and by the learners for revise the concepts. It is also important that the instructor(s) prepares some examples both of general usability problems and issues related to application-dependent and independent analysis. For a concrete understanding of usability background the instructor has to provide the learners with some readings on usability.

RESOURCES			
NAME	**DESCRIPTION**	**TYPE**	**USE LOCATION**
Slides Pack 1	A set of slides presenting: – The definition and the history of usability; – The main usability methods; – The difference between application-independent and application-dependent analysis; – The brief introduction on MiLE+	Content	Classroom + Home
General Examples of poor usability	Several examples of poor usability, not only related to websites, but also excerpts from other domains (general interactive and software applications, households' interfaces…)	Concrete examples	Classroom
Examples of application-dependent problems	Several examples of application-dependent problems extracts from famous websites.	Concrete examples	Classroom
Examples of application-independent problems	Several examples of application-independent problems extracts from famous websites.	Concrete examples	Classroom
Readings on usability	Articles and several chapters of books that introduce the concept of usability. In particular: – Chapter 1-2, Usability for the web, Brinck et al. (2002); – Chapter 1-2, Usability Inspection methods, Nielsen and Mack (2004);	Content	Home

Table 13: Module 1 list of resources

4.2.2 Module 2: MiLE+ inspections techniques

Goals and learning path

Module 2 is dedicated to the complete description of MiLE+ inspections techniques. It starts with the definition of the main concepts (such as U-KITs, scenarios, etc.) and it continues with the explanation of the main inspection techniques (technical inspection and user-experience inspection). Then a great attention is put both on the process for the scenario's creation and on the metrics to evaluate the applications (technical heuristics and user-experience indicators). The "ideal" process for performing the inspections' activities concludes this part of the module. Indeed, the module ends with a brief introduction of the Scenario-based user testing (which is the main topic of the Module 3) and a recapitulation of the MiLE+ framework. So, the main learning topics of Module 2 are:

- *Defining MiLE+ main concepts (G1_M2);*
- *Explaining the main inspection techniques (G2_M2);*
- *Presenting the process for the scenarios creation (G3_M2, G4_M2);*
- *Presenting the metrics for evaluating the applications (G5_M2)*
- *Describing the process for performing inspections (G6_M2);*
- *Introducing the Scenarios-based user testing (G7_M2).*
- *Recapitulating the MiLE+ framework (G8_M2)*

GOAL STATEMENT						
TAG	TEACHING STRATEGY ID	STATEMENT	TARGET	APPROACH	ASSESSMENT	IMP.
G1_M2	Def_4	Define the MiLE+ main concept	All	Definitions	Informal Class questions	5
G2_M2	Dex_2	Assess the MiLE+ main inspection techniques	All	Definitions + Examples	Informal Class questions *and/or* Written Exam	5
G3_M2	Pro_1	Be aware of the process for building scenarios	All	Presentation + Example	Informal Class questions	4
G4_M2	Pra_1	Fixing the process for scenarios' creation	All	Group's Exercise	Group Questions	4
G5_M2	Dex_3	In-depth understanding of evaluation metrics (Technical Heuristics & UEIs)	All	Definitions + Examples	Informal Class questions + Class discussion	5
G6_M2	Gui_1	Be aware of the process for performing MiLE+ Inspections	All	Guidelines presentation	Class discussion	4
G7_M2	Def_6	Introduce briefly the Scenario-based User testing	All	Definitions	Informal Class questions	4
G8_M2	Ass_1	Assess the MiLE+ framework	All	Recapitulation	Informal Class questions *and/or* Written Exam	3

Table 14: detailed explanation of Module2 goals

Activity flow

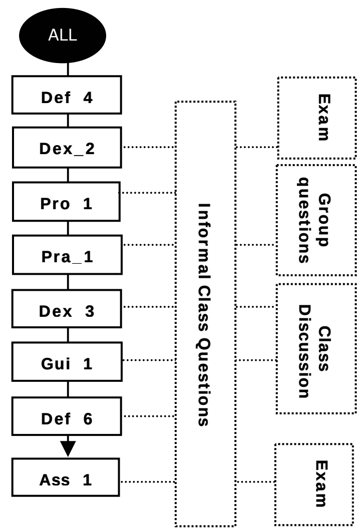


Figure 33: the activity flow of Module 2

The activity flow for Module 2 is quite heterogeneous in its structure. Indeed, after the initial definition of the main concepts the path continues using a combination of conceptual definitions and examples. Moreover, since in this module different processes are presented (e.g. process for scenarios building), it is very important that learners practice them. The last step of the learning path is a deep recapitulation of all the elements composing the MiLE+ framework. The assessment is made first of all using informal class questions and class discussion. It is possible to integrate them making a written exam after *G2_M2* and/or *G8_M2* for verifying if the main concepts are really assessed. Since the learners exercise in the creation of scenarios the instructors and tutors have to discuss with them and answer to the questions in order to assess the process.

Goals Mapping

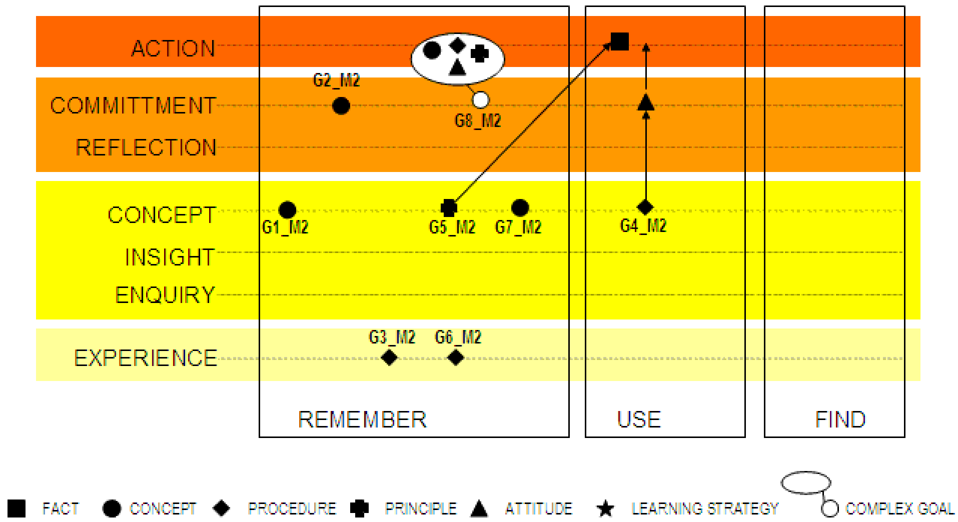


Figure 34: Module 2 goals mapping

Main concepts are in the *remember scope* and it is important to point out the heterogeneity of problems' level. Within this module the learning strategy is based on the alternation of goals that introduce and fix concepts (*G1_M2* and *G2_M2*) and goals that try to create a new attitude within the learners (*G3_M2*, *G4_M2* and *G6_M2*). To develop a new attitude towards usability evaluation the learners work at the experience level (through groups' works and class discussions - *G3_M2* and *G6_M2*). Notice the complexity of *G8_M2* which tries to commit learners through the recapitulation of concepts, procedures, principles and attitudes that lead the MiLE+ framework.

Relationships Diagram

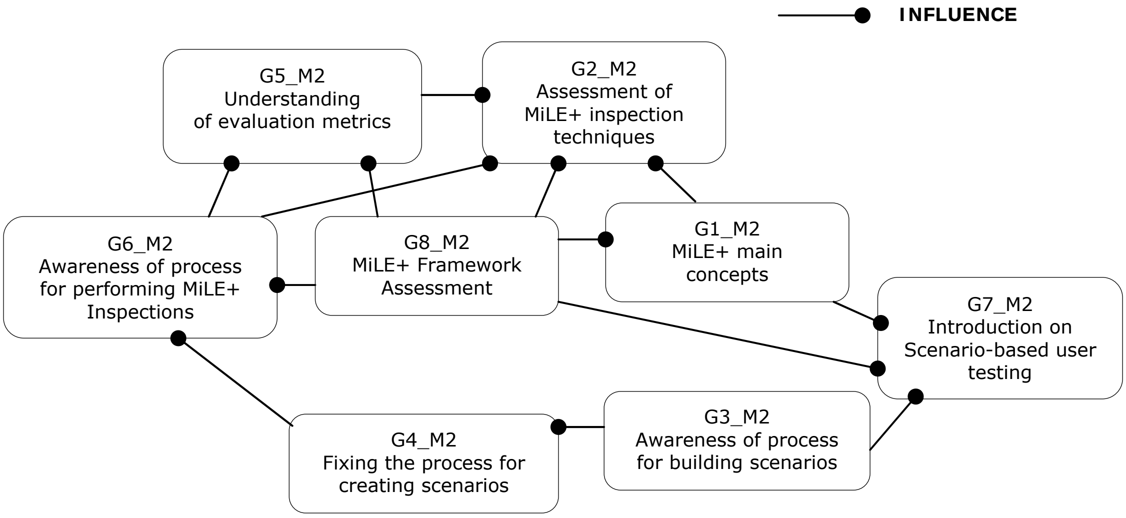


Figure 35: Module 2 relationships diagram

The diagram reveals the high number of influences among the goals. This means that there is a correct planning in deciding the goals of the module. In particular, *G8_M2* plays a fundamental role, although chronologically it is placed at the end of the module. In fact, it helps fixing all the concepts and procedures explained before. Clearly also *G1_M2*, which introduces the module, has an important influence in the understanding of the other goals.

Resources list

The resources list for Module 2 includes the Slides pack 2 (see Annex B_2), which has the main objectives such as the Slides pack 1 (it is used by both the instructor for the class lessons and the learners in order to revise the concepts). It is also fundamental that the instructor(s) prepares some examples of possible problems both of technical and user-experience inspection. In order to understand MiLE+ inspections it is important that learners read the articles explaining it before, during and after the course.

RESOURCES			
NAME	**DESCRIPTION**	**TYPE**	**USE LOCATION**
Slides Pack 2	A set of slides presenting: MiLE+ main concepts (Technical Inspection, User-experience inspection, scenarios, usability kits…); Examples of Technical Inspection Examples of User-experience inspection Process for building scenarios Introduction on Scenario-based user testing	Content	Classroom + Home
Examples of usability problems related to Technical Inspection	Several examples of problems extracts from famous websites related to technical inspection.	Concrete examples	Classroom
Examples usability problems related to User experience inspection	Several examples of problems extracted from famous websites related to user experience inspection.	Concrete examples	Classroom
Exercise 1	An exercise related to scenarios building	Exercise	Classroom *or* Home
Manual	MiLE+ inspector manual	Content	Classroom + Home
Readings on usability	Articles and books about MiLE and MiLE+ method. In particular: - Triacca L, Bolchini D., Botturi L., Inversini A, (2004). MiLE: Systematic Usability Evaluation for E-learning Web Applications. ED Media 04, Lugano, Switzerland (Awarded Paper). - Bolchini D., Triacca L. Speroni M. (2003). MiLE: a Reuse-oriented Usability Evaluation Method for the Web. HCI International Conference, June 2003, Crete, Greece. - Cantoni, L., Di Blas, N., Bolchini, D., (2003). Comunicazione, Qualità, Usabilità. Apogeo (Milano). - Triacca L., Speroni M., Bramani C., Understanding Semiotic issues in Usability Evaluation of Cultural Heritage Websites: the DICE case study, Scoms, University of Lugano. (2005).	Content	Classroom + Home

Table 15: Module 2 list of resources

Notice that the added value is provided by the *MiLE+ Inspector Manual* (see Annex A), which includes the library of heuristics, the library of User Experience Indicators, examples of evaluation and the process guide. Moreover, it includes suggestions and guidelines for reporting the results and it is important to underline that the manual is written to facilitate the readers.

4.2.4 Module 3: MiLE+ scenario-based user testing

Goals and learning path

The Module 3 is completely devoted to explaining the Scenario-based user testing, which is the third evaluation activity proposed by MiLE+. The decision to dedicate a module to this activity is related both to the fact that the user testing has a different nature with respect to inspections and to the complexity in understanding how to performing the test (end-users selection, screening, testing, etc.). With respect to the previous method, Module 3 is less complicated in its organization. It starts with a complete definition of scenario-based user testing and then the instructor illustrates both the guidelines for conducting test and the process for performing it. For achieving these objectives, the main learning topics are:

- *Defining Scenario-based user testing (G1_M3);*
- *Presenting guidelines for conducting the user testing (G2_M3);*
- *Presenting the process for the performing the user testing (G3_M2).*

GOAL STATEMENT						
TAG	**TEACHING STRATEGY ID**	**STATEMENT**	**TARGET**	**APPROACH**	**ASSESSMENT**	**IMP.**
G1_M3	Def_7	Define the Scenario-based User testing	All	Definitions	Informal Class questions	5
G2_M3	Gui_2	Be aware of the guidelines for conducting the Scenario-based user testing	All	Guidelines presentation + Examples	Informal Class questions + Class discussion	4
G3_M3	Proex_2	Be aware of the process for conducting a user testing	All	Definitions + Examples	Informal Class questions and/or Written Exam	4

Table 16: detailed explanation of Module3 goals

Activity flow

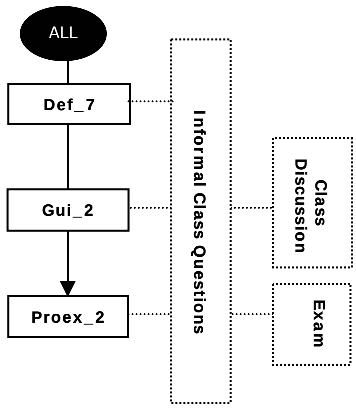


Figure 36: the activity flow of Module 2

The activity flow for Module 2 is simple. Indeed, after the initial definition, for each concept and process' step the instructor illustrates some example. The assessment is made like in previous modules: informal class questions and class discussion. As further assessment it is also possible to introduce a written exam at the end of the module.

Goal Mapping

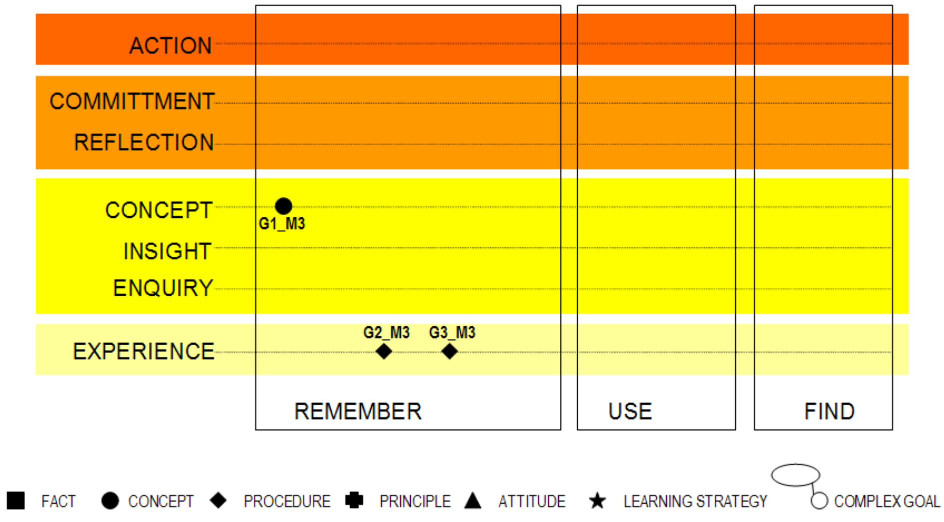


Figure 37: Module 3 goals mapping

Also in the case of Module 3 the goals are in the *remember scope*, as the very practical nature of user testing, *G2_M3* and *G3_M3* are related to procedures and correlated by examples.

Relationships Diagram

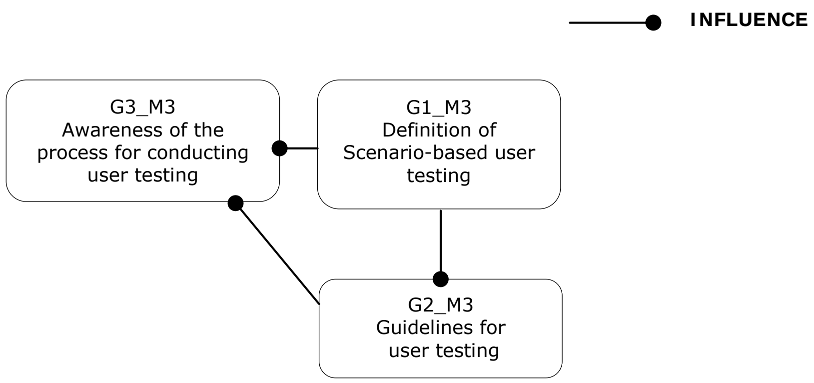


Figure 38: Module 3 relationships diagram

The diagram in this case is very simple and shows the centrality of *G1_M3* which influences the other goals. The guidelines for conducting the user testing could impact the process in terms of performances. For example, if a guideline states "you have to carefully select your end-users", you have to take into account it at the beginning of the test (selection phase).

Resources list

The list comprises the Slides pack 3 (see Annex B_3), which includes slides presenting the definition of Scenario-based user testing, important guidelines for conducting the test and the process. Slides pack 3 is also used both by the instructor for the class lessons and by the learners for revision. Another important learning resources are videos presenting examples of tests. In order to go in depth in details it is very important to provide the learners with readings on user testing.

RESOURCES			
NAME	**DESCRIPTION**	**TYPE**	**USE LOCATION**
Slides Pack 3	A set of slides presenting the Scenario-based user testing, the guidelines and the process for conducting it.	Content	Classroom + Home
Examples of user testing	Videos showing user testing. In particular: - Flash Usability Highlights Video: film clips from user testing of Web-based applications, by J.Nielsen.	Videos examples	Classroom
Readings on user testing	Articles and books' chapters about user testing. In particular: - Kuniavsky, M., (2003). Observing the User Experience. Morgan-Kauffmann. - Molich R., Nielsen J., 230 Tips and Tricks for better Usability Testing, NNGROUP Report - www.nngroup.com/reports/tips/usertest, California.	Content	Home

Table 17: Module 3 list of resources

4.2.4 Module 4: MiLE+ in practice

Goals and learning path

After an introduction dedicated to the "ideal" structure of the usability report, Module 4 is completely reserved to the practice of MiLE+. The learners' practice is gradual: they starts with a simple and individual usability evaluation (normally they analyze part of an application); afterwards they have to perform a real usability evaluation (normally this activity is performed in groups of maximum 3 learners). This means that they have to discover problems, analyze and report them. So, the main learning topics of Module 4 are:

- *Reporting usability: practical tips (G1_M4);*
- *Individual training (G2_M4);*
- *Group training(G3_M4).*

GOAL STATEMENT						
TAG	TEACHING STRATEGY ID	STATEMENT	TARGET	APPROACH	ASSESSMENT	IMP.
G1_M4	Gui_3	Be aware of the "ideal" usability report 's structure	All	Definitions + Guidelines + Examples	Informal Class questions	4
G2_M4	Pra_2	To be trained on MiLE+ Inspections Check the understanding on MiLE+ method	Single Student	Individual case study (a section of a website)	Individual Questions + Tutoring	4
G3_M3	Pra_3	To be trained on a complex usability evaluation using MiLE+ Check the understanding on MiLE+ method	All	Group case study (collaborative)	Group questions + Tutoring + Results Presentation (written & oral)	5

Table 18: detailed explanation of Module4 goals

Activity flow

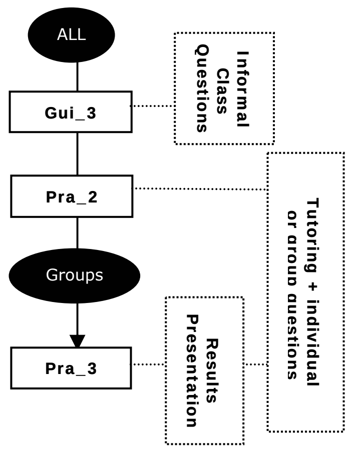


Figure 39: the activity flow of Module 4

In their practice the learners are helped by tutors. Tutors have the role of answering questions and verifying the level of the learners' understanding. The final assessment is provided by a presentation which should be both written (the usability report) and oral (the presentation of the main results obtained).

Goal mapping

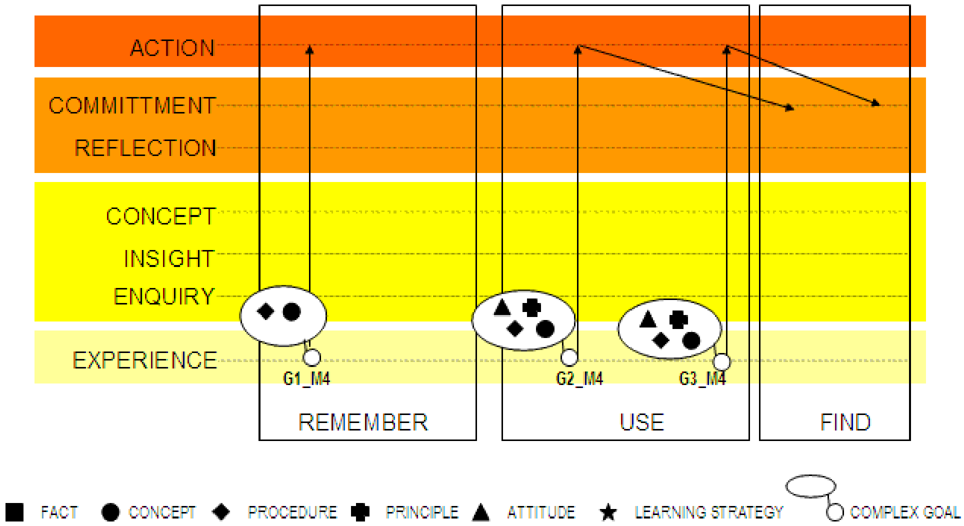


Figure 40: Module 4 goals mapping

As main objective the goals of practicing are to force the learners to use the method applying all the concepts, principles and attitudes learned in the previous modules. They analyze the applications starting from their own experience and through the action (the evaluation) they become committed to the validity both of the method and the learning path. It is important to point out that through the practice the learners generate new knowledge (for example often it happens that the learners discover new heuristics).

Relationships Diagram

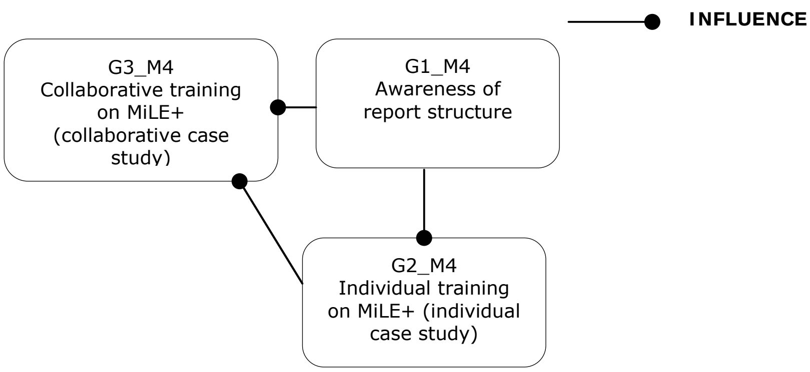


Figure 41: Module 4 relationships diagram

The relationships diagram shows both the influence that *G1_M4* has on practice and, most importantly, the influence of G2_M4 on G3_M4 (the ability of conducting a usability evaluation increases along with the experience of doing it).

Resources list

The list for the Module 4 includes resources to primarily understanding the reports structure. However, it is important that the instructor and the tutors prepare a list of applications to assign both for individual and for collaborative case studies (clearly the learners can choose the application to analyzed).

RESOURCES			
NAME	**DESCRIPTION**	**TYPE**	**USE LOCATION**
Slides Pack 4	A set of slides presenting the "ideal" report structure.	Content	Classroom + Home
Examples of usability reports	Examples of good usability reports.	Content	Classroom + Home
List of applications	A list of applications for individual and groups case studies	Exercise	Home
Manual on usability reporting	The manual describing in details the guidelines for reporting usability issues and the report's structure.	Content	Home
Readings on reporting usability	Articles about usability report and user testing. In particular: - Bolchini D., Colazzo S. (2003), Guidelines for describing usability problems, HCI International Conference, June, Crete, Greece. - Dumas, J., Molich, R., & Jeffries, R., (2004). "Describing usability problems: are we sending the right message?". Interactions, Volume 11, Issue 4 July + August 2004. - Molich R., Nielsen J., 230 Tips and Tricks for better Usability Testing, NNGROUP Report - www.nngroup.com/reports/tips/usertest, California. - Tognazzini B., How to Deliver a Report Without Getting Lynched, ASKTOG Column2001 www.asktog.com/columns/047HowToWriteAReport.html .	Content	Home

Table 19: Module 4 list of resources

4.3 Learning@MiLE+: the courses

This section is dedicated to present four case studies of MiLE+ courses. The target of the courses is different as well as the time at disposal. Notice that the learning goals and the modules used for every course are selected from those proposed within the section 4.2. Also the selection of the topics as well as the type of learning paths have been made taking into account both the learners' typology and time.

The courses have been entitled:

- *MiLE+ full experience;*
- *8 MiLE+: usability crash-course;*
- *MiLE+ 4you: usability crash-course;*
- *MiLE+ online course.*

4.3.1 MiLE+ full experience

The MiLE+ full experience course provides both the conceptual tools and the practice to become a usability evaluator. In taking this course the participants understand in-depth what usability is, its importance and how to use MiLE+. Moreover, they develop a group or personal project which consists of the complete evaluation and reporting of an interactive application. The course length varies from 28 to 36 hours (including ex-cathedra lectures and tutoring); the homework is not included in this range (normally it could vary from 20 to 30 hours pro student, in order to accomplish both the individual case study and the group case study).

Target and situation of usage

MiLE+ full experience course has been conceived for students at a university master level as well as practitioners. Indeed, this type of course needs a certain experience in design and/or in development of interactive applications.

The course has been tested with two master classes:

- *Usability Lab course:* the participants were Students in the first year of TEC-CH Master (Technology-enhanced communication for cultural heritage) at the University of Lugano.
- *Usability of interactive applications:* the participants were Students in the first year of the Master in Computer Science Engineering at the Politecnico of Milan, campus of Como.

Selecting modules, learning goals and activity flow

All the modules, learning goals and paths are used to assemble the *MiLE+ full experience course*. The only variation could be situated in Module 4 (practice) where it is possible to select individual training, the group training or both.

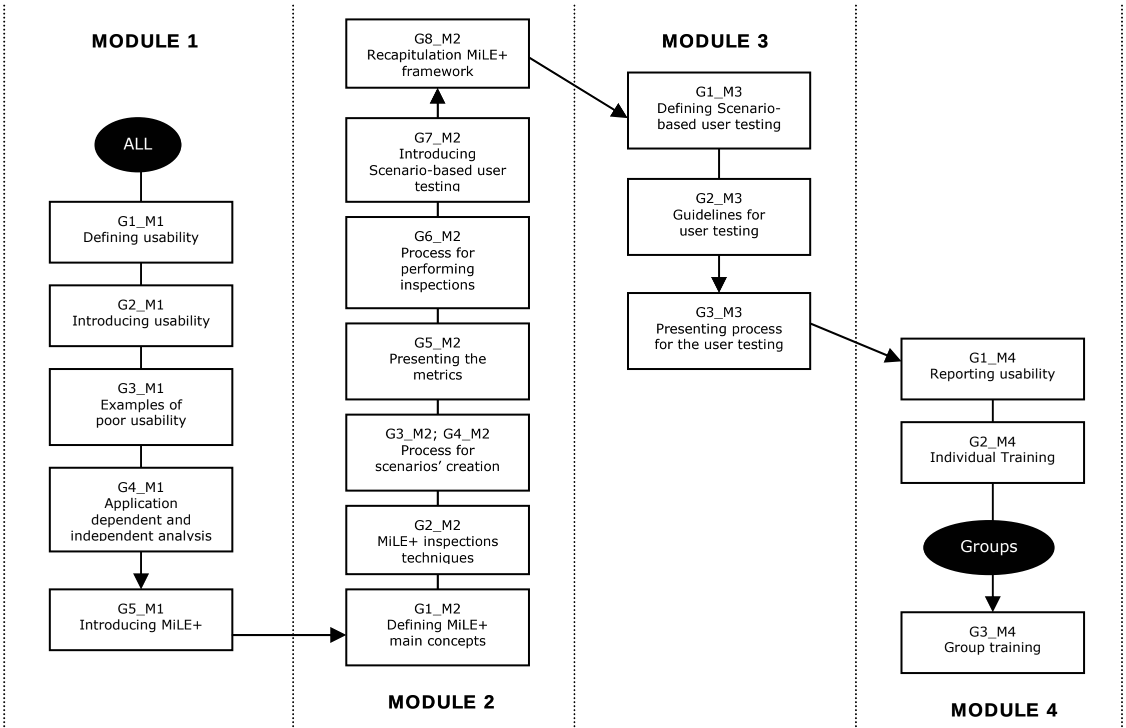


Figure 42: activity flow of *MiLE+ full experience* course.

Scheduling

As said before this course passes through all the modules. Notice that the hours for completing the course can be organized in different ways. For example the course *Usability lab* has been taught in ten days.

The Figure 42 shows that the first four days are devoted to lectures that include learning topics and goals from Module 1 to Module 3. It is important to point out that in parallel the students start with the practice (Module 4), in particular the individual evaluation. During the following three days the students can receive several tutoring sessions and have time for a class discussion. The last day is devoted to the groups' presentations: each group has 15 minutes to present the main results of their evaluation and they have to deliver the usability report.

Date	Schedule	Hours	Activity Type	Theme	Homework
10 Dec.	12.00-16.00	4	Lecture	**Course intro & usability foundations**	○ Assignment *"Informal evaluation" (individual)*
13 Dec.	8.30 – 12.00	4	Lecture	**MILE+ Usability Evaluation: Inspection** ○ **Technical inspection**	○ **Delivery** (in class) *"Informal evaluation"* ○ Assignment *"Usability Evaluation Project" (teamwork)*
14 Dec.	8.30 – 12.30	4	Lecture	**MILE+ Usability Evaluation: Inspection** ○ **User-experience inspection**	
15 Dec.	13.00 – 18.00	5	Lecture	**Scenario-based user testing** **Reporting usability**	
17 Dec.	8.30 – 10.30 14.00 – 16.00	6	Tutoring	*Morning*: **Class Group Tutoring** *Afternoon:* **Tutoring for user testing**	
20 Dec.	9.00 – 11.00	2	Tutoring	**Class questions&answers**	
21 Dec	9.00-11.00	2	Tutoring	**Class Group Tutoring** & course wrap-up	
22 Dec.	13.00 – 15.00	3	Presentation	***Project Group Presentations***	**Delivery** *"Usability Evaluation Project" (teamwork)*

Figure 43: detailed scheduling of *Usability lab* course.

Another possibility for scheduling the *MiLE+ full experience course* is a semester. An example of this program is the course *Usability of interactive applications* here below.

Date	Hours	Theme	Assignment
Fri 18 March	8,30-12,30	Introduction	
Fri 1 April	8,30-12,30	Usability foundations	
Fri 8 April	8,30-12,30	MiLE+: main concepts and technical inspection	**Assignment Group project**
Fri 15 April	8,30-12,30	User-experience inspection and Reporting usability	
Fri 22 April	8,30-12,30	User Testing – examples	
Fri 29 April	8,30-12,30	Questions-Answering	
Fri 6 May	8,30-12,30	*Written exam*	
Fri 13 May	8,30-12,30	User Testing – guidelines	
Fri 20 May	8,30-12,30	Tutoring	
Fri 27 May	8,30-12,30	Tutoring	
Fri 10 June	8,30-12,30	Wrap-up, Questions & Answering	**Delivery of usability report**
Fri 17 June	8,30-12,30	Project presentations	
Fri 24 June	8,30-12,30	Project presentations	

Figure 44: detailed scheduling of *Usability interactive applications* course.

The only substantial difference is the written exam positioned at half-way through the course.

It is important to underline that the participants of this course received all the learning resources presented in section 4.2.

4.3.2 8 MiLE+: usability crash-course

This course is devoted to provide learners with the main concepts and procedures of MiLE+. In a full-immersion day, the participants are trained on the inspection techniques (technical inspection and user-experience inspection) and they can also practice them. At the end of the day they will have the sufficient knowledge for performing a simple application inspection.

Target and situation of usage

In general, the main target of this course are people without experience in usability. Indeed, this course has been organized thinking on bachelor and master students.

The first edition took place the 24th of April 2005 and the participants were 2nd year Bachelor Students, with no experience in design and evaluation of interactive applications.

Selecting modules, learning goals and activity flow

Considering the target and the purposes of this course, all the contents of Module 1 and 2 and a short version of Module 4 are used. The Scenario-based user testing is not introduced, given the complexity of the topic and the lack of time. Consequentially, the full-immersion day is divided in three parts. The first one is dedicated to the Introduction on usability, the second one to the MiLE+ Inspection techniques and the final part to exercise the new knowledge.

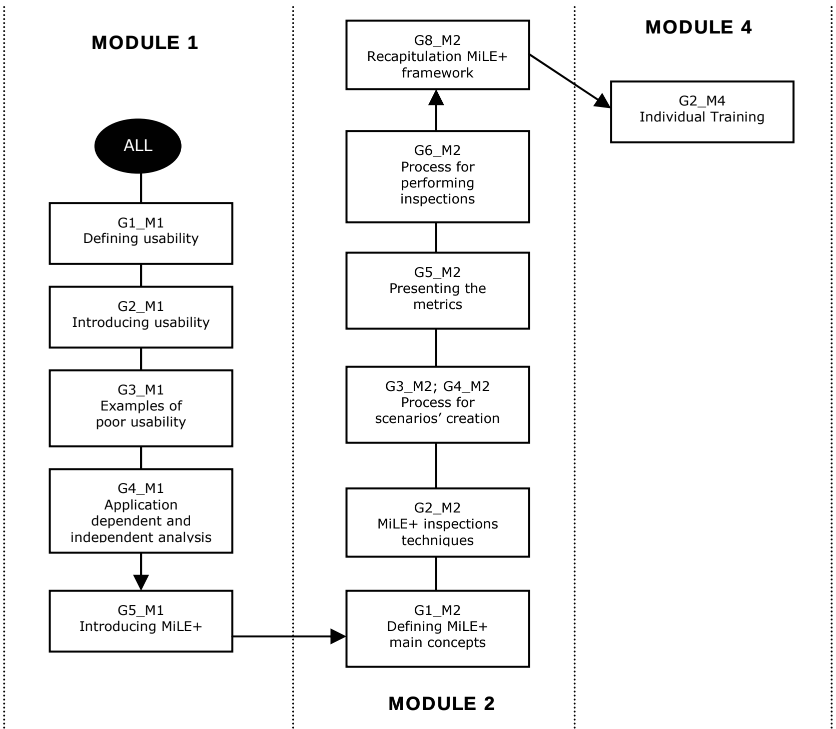


Figure 45: activity flow of 8*MiLE+ usability crash course*.

Clearly, the contents of Module 1 and 2 are described less in detail with respect to the course *MiLE+ full experience*. This choice is also due to the time at disposal for explaining each basic concept.

Scheduling

As previously mentioned, the full-immersion day is divided in three parts. Therefore, the final scheduling takes into account this precondition.

HOUR	THEME	ACTIVITIES	TIME
9.00 - 9.10	Introduction	- Course outline and organization	

9.15 - 11.00
 	
(break at 10)
	Introduction on Usability	- Usability Foundations (Definition, Main Usability methods…)
		- Examples of poor usability
		- Application independent Problems
		  - Examples
		- Application dependent Problems
		  - Examples
		- Brief Presentation of MiLE+ method
			20'
			20'
			25'
			25'
			15'

11.00 - 11.10	coffee break		

11.10 - 13.10
 	
(break at 12.10)
	MiLE+: a systematic method for usability evaluation	- Introduction
		- Main Concepts
		  - Scenarios
		  - Usability Evaluation Kit (U-KIT)
		- Technical Inspection
		  - Examples of Technical Heuristics
		- User Experience Inspection
		  - Examples of User Experience Indicators
		- How to build scenarios
		- Synoptic Picture (MiLE+ recapitulation)
			5'
			30'
			25'
			25'
			20'
			5'

13.10 – 14.30	lunch		
14.30 - 18.15 (break at 16.00)	MILE+ in practice	- Explanation of the main Heuristics and User Experience Indicators - Scenarios Building - Individual Evaluation of a section of a given website	1h 45' 2h
18.15 - 18.30	Discussion & recap	- Overall discussion & recapitulation	15'

Figure 46: *8MiLE+: usability crash-course* detailed scheduling

As the Figure shows, the duration of the modules' contents have been adapted considering the time at disposal. It is important to point out that the participants also received the learning resources, in particular:

- course slides;
- articles both on usability and on MiLE+;
- MiLE+ libraries (technical heuristics and user experience indicators);
- examples of evaluation.

4.3.3 MiLE+ 4you: usability crash-course

This half-day course (4 hours) has the main objective to introduce MiLE+ in a context such as conference workshops, tutorials or seminars. Therefore, considering the target and the time at disposal it is not possible to provide an in-depth course but just an overview of the method. An important element is the interaction between the instructors and the audience. In this context it is always very useful and interesting to discuss the method and its approach to the usability evaluation (it is one of the key moments for assessing and promoting the features of the method).

Target and situation of usage

The target of this course are practitioners and professionals not only in the field of human-computer interaction (e.g. interface designers, developers…) but also project managers that do not have a specific background (e.g. cultural heritage experts). Indeed, this course has been experimented in two circumstances:

- *VAST 2004 Conference tutorial (December 2004)*: it was organized within the 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage. The participants were cultural heritage experts;
- *Swiss Usability Professional Association (SWISSUPA) course (December 2004):* it was organized by TEC-Lab of the University of Lugano and the participants were practitioners in the field of web design, software development and usability evaluation.

Selecting modules, learning goals and activity flow

To create this course a short version of Module 2 including a brief introduction on usability evaluation (excerpts from the Module 1) are used. In particular, we highlight two main concepts:

- *the importance of using scenarios in usability evaluation;*
- *the conceptual difference between technical and user experience inspection.*

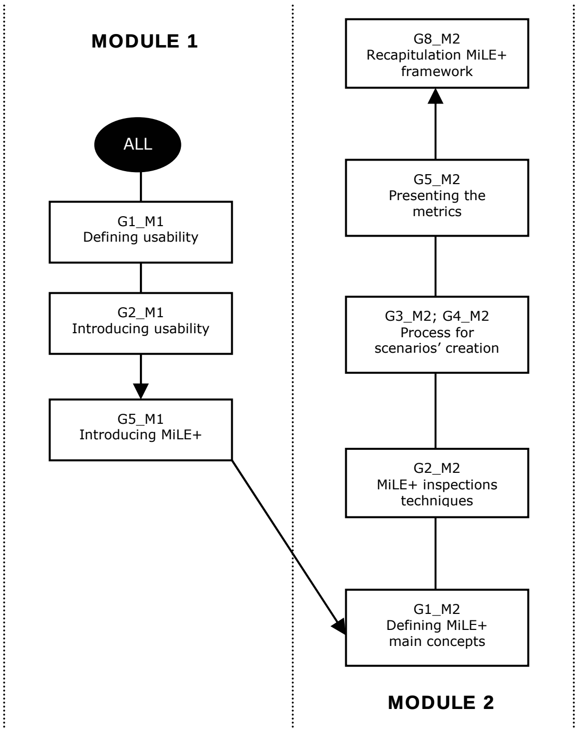


Figure 47: activity flow of *MiLE+4 you usability crash course*.

Scheduling

The four-hour-course scheduling starts with a brief introduction both on usability and on the MiLE+ method. However, the course focuses on the MiLE+ presentation. As said before, in this context it is very important the interaction with the learners. So, after the explanation of each concept the instructors discuss it with the participants.

HOUR	*THEME*	*ACTIVITIES*	*TIME*
9.00 - 9.10	Introduction	- Course outline and organization	
9.15 – 10.00	Introduction on Usability	- Usability Foundations (Definition, Main Usability methods...) - Brief Presentation of MiLE+ method	20' 25'
10.00 - 10.10	coffee break		

10.10 - 12.30
 	
(short break at 12.30)
	MiLE+: a systematic method for usability evaluation	- Introduction
		- Main Concepts
		  - Scenarios
		  - Usability Evaluation Kit (U-KIT)
		- Technical Inspection
		  - Examples of Technical Heuristics
		- User Experience Inspection
		  - Examples of User Experience Indicators
		- How to build scenarios
		- Synoptic Picture (MiLE+ recapitulation)
			10'
			30'
			
			30'
			
			30'
			
			20'
			20'

12.30 - 13.00	Discussion & recap	- Overall discussion & recapitulation	30'

Figure 48: detailed scheduling of *MiLE+ 4you: usability crash-course*.

4.3.4 MiLE+ online course

MiLE+ online course has been developed within the USABLE project (funded by the Swiss Virtual Campus initiative - S-3-002, www.usableproject.net). This project is developing a blended learning course on how to evaluate the usability of web applications. The course will act as a bridge for knowledge transfer from universities and scientific centres (where usability methods are mainly elaborated and validated) to professionals working in the new media industry (user experience designer, communication experts, interaction designers, web developers). In fact, the modular structure of the course will enable not only to organize usability course in a richer and more flexible way in the universities, but also to reuse selected modules for continuing education curricula and courses for professionals in the field of website design and usability evaluation.

The target public are students in computer science and communication sciences (both in the universities and universities of applied sciences) as well as professionals in the new media and electronic publishing industry.

The usable learning modules are seven:

- *Usability foundations*
- *Usability inspection methods;*
- *User testing and empirical testing;*
- *Reporting usability*
- *Usability for cultural heritage websites;*
- *Usability for banking websites*
- *Usability and accessibility*

Each module have the same contents and structure: a short video introducing the module, the content (divided in lesson, slides for revision, references and resources), a questions-answers section (FAQs, Rehearsal questions, Forum), a glossary, learning objective and instructors, the access to the Virtual Usability lab (see Chapter 3) and assignments (individual and collective).

Selecting modules, learning goals and activity flow

MiLE+ method is the core content of the *Inspection methods* and, after an introduction about existing inspection methods, it is fully presented. Therefore, the learning topics are those of MiLE+ Module 2 (from *G1_M2 to G1_M8)* and Module 4 (from *G1_M4 to G3_M4)* . Moreover, the domain analysis reported in the Usable module *Usability for cultural heritage* is entirely based on MiLE+. Also *Usability foundations* and *Reporting usability* have a learning structure similar to the approach presented in this Chapter (in particular *G1_M1* and *G1_M4*). The assessment is included in each Usable module: thanks to the questions-answers sections (in particular the forum) the learners can interact both with other learners and with instructors and tutors.

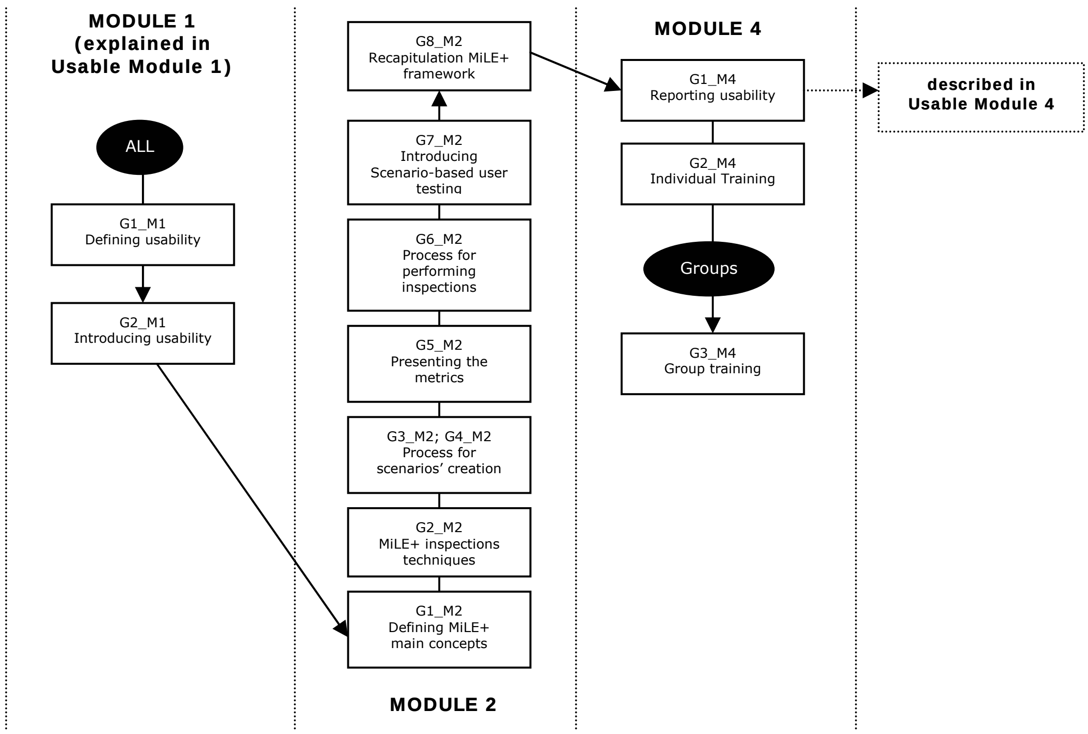


Figure 49: activity flow of *MiLE+ online* course.

Chapter 5:

Validating MiLE+

Summary:

The aim of this Chapter is to validate the MiLE+ method. The goals of the validation are three:

- *assess the learnability and applicability of the method;*
- *verify the effectiveness and the efficiency of the method;*
- *measure the reliability of the method.*

The Chapter describes in a detailed way the experiment carried out in order to achieve the goals above and it is based on three main hypothesises:

- *Effectiveness and efficiency Hypothesis*;
- *Inter-group agreement on findings Hypothesis*;
- *Intra-group agreement on findings Hypothesis.*

The final part of the Chapter is completely reserved for the results of the experiment that will demonstrate the mentioned hypothesises.

5.1 The importance of validating MiLE+

As presented in the previous chapters, MiLE+ has been proposed to help usability evaluators to perform their evaluation in a systematic and efficient way providing them with several usability tools. Moreover, it is important to keep in mind and point out that MiLE+ has been developed to be reusable by different evaluators. To verify and assess the reusability of the MiLE+, a *Validation Experiment* has been carried out, in which two groups of evaluators with different backgrounds have inspected the section "Collection" of the Cleveland Museum of art website.

The main goals of the MiLE+ experiment are:

1. *To assess the learnability and applicability of the method:* it is important to validate the teaching strategies proposed in Chapter 4 and to verify if the support material (see Annex A) is able both to clearly explain the method, and to really help the inspector during the analysis. This goal is a macro-goal that can be verified through goals 2. and 3.
2. *To verify the effectiveness and the efficiency of the method:* a usability method is effective and efficient if it allows the inspector to find the higher possible number of problems as possible in specific conditions (e.g. time at disposal).
3. *To measure the reliability of the method:* the reliability is the characteristic of a result to be reproduced. In the context of validating MiLE+ the goal is to determine if the inspectors performing a usability evaluation inspection at the same conditions (time, website, material at disposal, etc.) are able to produce results that should be comparable and similar. For reaching this goal we have developed a conceptual approach to discover the reliability within a group of inspectors and among different groups. Indeed, the literature is not convincing and it does not offer a suitable approach for the reliability assessment within experiment similar this explained in the Chapter.

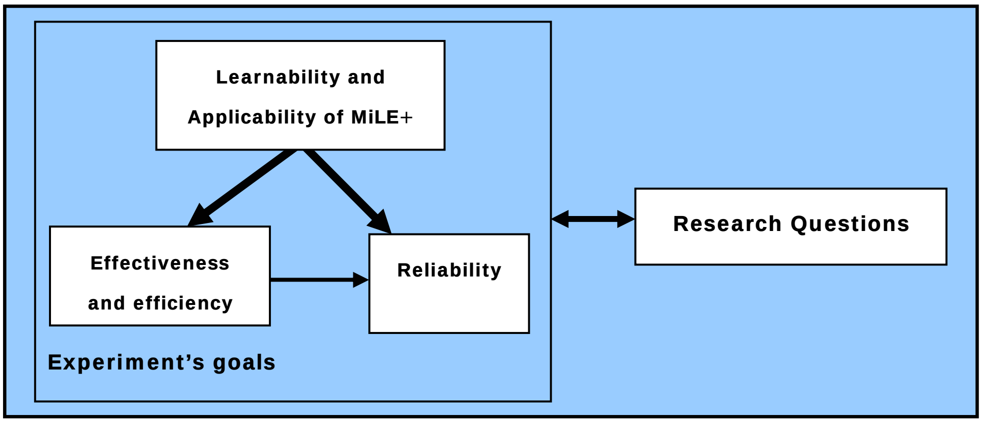


Figure 50: Relationships between experiment's goals and the impact on the Research Questions

Figure 50 shows the relationships between the experiment's goals and the Research Questions. As said before, there is a strong correlation between the verification of the method's learnability and applicability and the other two goals. Moreover, it is possible to measure the reliability of taking into account the problems found by the evaluators (effectiveness) and verifying the degree of agreement in the evaluators' judgments and findings.
Considering the Research question about the possibility to effectively communicate (and teaching) the MiLE+ method the experiment has been conceived to try to answer this crucial question.

5.2 The Experiment

In order to validate the MiLE+ method a comparison study involving a total of 29 students has been conducted. These students coming from two different Universities and they have a different background. The first group, called *Como Group* consisted of 16 students of the Usability Project class: they are at the first year of the Master program in Computer Science Engineering at the Politecnico di Milano (campus of Como). The second group, called *Lugano Group*, consisting of 13 students of the New Media Theory class: they are at the second year of the Bachelor program in Communication Sciences at the University of Lugano. As explained more in detail in Paragraph 5.3, the validation metrics were defined along three major dimensions: *Effectiveness and efficacy*, *Intra-group agreement on findings*, *Inter-group agreement on findings*. As previously presented, the *effectiveness and efficiency* refer to the number of problems found considering several (restrictions such as time at disposal). In other words, it is possible to define the effectiveness as problems recall (that is the proportion of relevant issues that are discovered by the inspectors).
Inter-group agreement on findings refers to the reliability of the results considering the common issues discovered among the groups. *Intra-group agreement on findings* measures the reliability within each group. In other words, the agreement on findings concerns the problems precision which is the degree to which a measurement (problems discovery) is derived from a set of observations having small variation.

For each dimension a specific hypothesis has been tested:

- *Effectiveness and efficiency Hypothesis*: as a general hypothesis we predicted that the *Como Group* should have been more effective compared to the *Lugano Group*. This consideration is related to the students background: the master students of Politecnico di Milano are experienced designers with respect to the bachelor students from Lugano (for an in-depth description of the two students group see paragraph 5.2.1). Before the experiment, we predicted that also within the *Lugano Group* it could be possible to verify a difference in number of discovered problems. Indeed, the Lugano group has been split in three sub-groups (called MiLE+2h, MiLE+4, MiLE+8h): each sub-group received a different training (see paragraph 5.2.1), so it should be possible that inspectors belonging to MiLE+8h group (which received a more complete training) found more problems with respect to MiLE+4h and MiLE+2h inspectors. It is assumable that considering the time at disposal for performing the experiment, the inspectors should have decided to perform only the technical inspection, which is more efficient and effective in a short time. Indeed, in Chapter 3 we have pointed out that the Technical Inspection is the quicker and cheaper activity to identify problems (but it is more superficial).

- *Inter-group agreement on findings Hypothesis*: the agreement's results of Lugano and Como groups on common problems should demonstrate that if a problem is found by an inspector, it is detected by many others (high level of agreement and reliability).

- *Intra-group agreement on findings Hypothesis:* considering the training received it is possible to expect that the Lugano Group's inspectors should be more reliable. In fact, the training for this group has been more structured and, theoretically, the inspectors should have reached a similar level of understanding and applicability of the method.
  Also in this case, there should be a certain difference between the sub-groups of the Lugano Group. However, this difference should not be impressive.

Moreover, we have an additional hypothesis related to the capability of MiLE+ to guide inexperienced evaluators in finding navigational problems (*navigational problems discovery hypothesis*), that are very difficult to detect for people who do not have practice in design of hypermedia applications. In general, other methodologies do not provide sufficient tools and procedures to detect these types of issues. On the contrary, MiLE+ offers several heuristics and guidelines to discover them.

In the following paragraphs is described the experimental method adopted for testing these hypothesises.

Method

Participants and learning paths

As presented earlier in the Chapter, the participants have been selected considering different backgrounds and overall experience especially in design of interactive application. This kind of selection allowed verifying if the method is really learnable and reusable.

Description of the Lugano Group (2nd year Bachelor's students of the New Media Theory class)
The inspectors of this group, composed of 13 students, are inexperienced designers and did not have any practice neither in usability evaluation nor in design of interactive applications. Notice that the students had participated to the *8 MiLE+: usability crash-course* (see Chapter 4) with some change in the learning path. Indeed, this group has been split in three sub-groups called MiLE+2h, MiLE+4h and MiLE+8h. The goal of this further division is to provide usability training at different levels in order to measure the reliability considering different learning paths (see Chapter 4). The first group (MiLE+2h) had two-hour of general overview on usability evaluation, the second group (MiLE+4h) had a four-hour complete training on the MiLE+ inspection techniques and the third group (MiLE+8h) besides the inspection techniques had a four-hour practice on a given website.
All the inspectors have received the same learning material and between the training day and the inspection day they had one week to study it. The learning material consisted of:

- *MiLE+ description:* an article describing the methodology;
- *Library of Technical Heuristics:* the list of technical heuristics, divided into design dimensions (e.g. content, navigation, interface design, etc.) including guidelines and examples for applying them;
- *Library of User Experience Indicators (UEIs):* the list of UEIs explained in details;
- *MiLE+ Examples of use:* two different documents with examples of use both related to the Technical Heuristics and the User Experience Indicators.
- *Hand-outs of the course:* the slides of the course were given to the inspectors;
- *Online Course:* a "special" release of the USABLE online course (http://athena.virtualcampus.ch/webct/logonDisplay.dowebct) has been created for the experiment. This release called Usability Crash-course provided the students with online contents (e.g. video presentations, articles…) on usability foundations and the MiLE+ method.

It is possible to schematize the learning path of Lugano Group as follow:

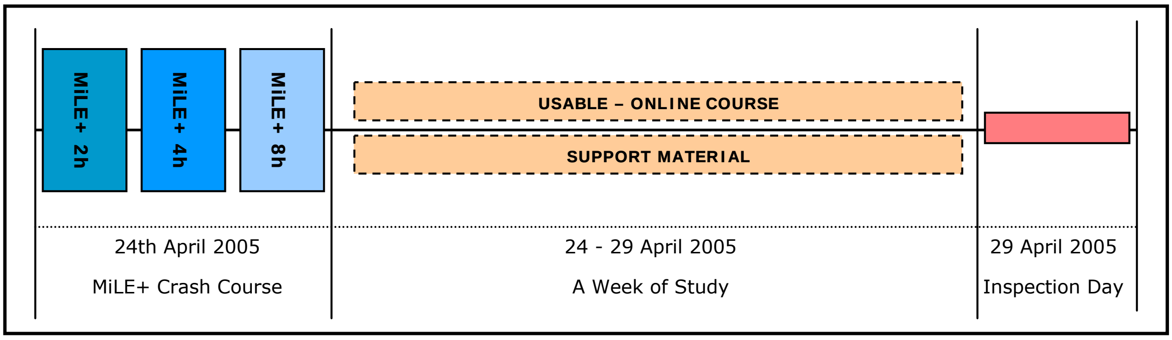


Figure 51: the scheduling for the Lugano Group's training.

Description of the Como Group (students of 1st year of Master in Computer Science Engineering: Usability project Class)

The Como Group was composed of 16 Master's students. They were experienced designers but with no experience in usability evaluation. To assess the "real" potential in learnability of MiLE+ and the usability of the support material (e.g. inspector manuals, see Annex A) this group received a less structured training with respect to the Lugano Group. Indeed, a "virtual training" has been created to help the students in understanding the method. The virtual training was structured as follow:

- *Online Course:* the same release of the USABLE online course (http://athena.virtualcampus.ch/webct/logonDisplay.dowebct) used by the Lugano Group students;
- *Online Forum:* a forum to allow the students interacting with the tutors and other students (http://hoc1.elet.polimi.it/webboard/wbpx.dll/~progettousabilita04-05/login);

All the inspectors had the possibility to meet the tutors once before the inspection day in order to have an in-depth explanation on the method and on the support material (two-hour-session for Questions & Answers). Moreover, the inspectors received the same learning material of the Lugano Group (*MiLE+ description*, *Library of Technical Heuristics, Library of User Experience Indicators, MiLE+ Examples of use).*

It is possible to schematize the learning path of the Como Group as follow:

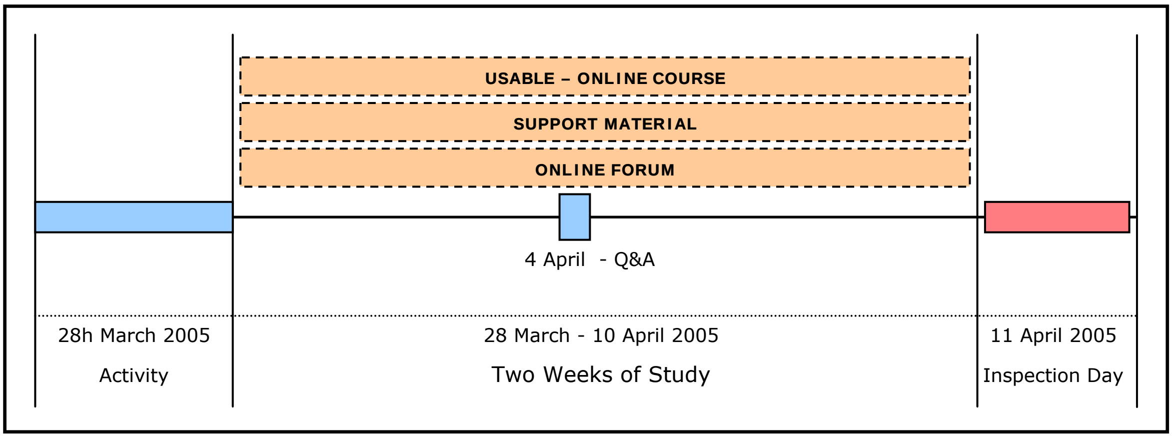


Figure 52: the scheduling for the Como Group's training.

Experiment Procedure and Data collecting

Both groups have followed the same experiment rules. All the inspectors have a maximum of three hours to perform a usability inspection of the section "Collection" of the Cleveland Museum of art website (www.clevelandart.org/index.html). The inspectors principally evaluated the Collections' section even though part of the study was also on the navigational paths from the home page to the Collections' section. The decision to concentrate the evaluation only on this section is due to the time at disposal (they did not have enough time to evaluate the entire application) and to the need of collecting homogeneous results (on the contrary it could have happened that the problems would have been too much fragmented on several website's sections). To facilitate the gathering of the data every inspector received the same report template a couple of days before the experiment. The description of each problem had the following structure:

- *NAME:* the name of the problems;
- *DIMENSION:* the dimension the problem refers to (e.g. Content, Navigation, Semiotics…);
- *DESCRIPTION:* a short description of the problem (maximum three lines);
- *URL:* the location of the problem. If a problem came out at difference instances (e.g. font size too small in different pages), they had to insert minimum two URLs and explain the repetitiveness of the problem.

Example of Problem's presentation:

- *NAME:* Intuitiveness of Homepage's main menu
- *DIMENSION:* Interface Design (Semiotics)

- *DESCRIPTION:* the homepage's main menu is indeed not intuitive at all. It presents 11 images which hided the links to the main section. The problem is that on the one hand the images are not predictable (it is very difficult to understand which the content behind the images is); on the other hand the images change their place randomly.
- *URL:* http://www.clevelandart.org/index.html

It is important to point out that the inspectors did not know the website they would have analyzed before the inspection day. Indeed, 10 minutes before the inspection they received both detailed information on the application (e.g. application's goals, section to analyze…) and the description of a macroscenario. The goal of the macroscenario was *to have an idea of the museum's collection*. Each inspector had also the possibility to refine the macroscenario into several (sub) scenarios for performing a more precise evaluation.
Since the inspectors performed the evaluation at the same place and time (Lugano Group April 29 from 4 pm to 7 pm and Como Group the April 11 from 1 pm to 4 pm) they were severely controlled by an experiment's supervisor.

The application

The Cleveland Museum of Art website presents three main sections: *Collection*, *Search* and *Events*. The section *Collection* exhibits the artworks of the museum. The artworks are organized in 19 sub-collections divided in geographical collections (from the African Art to the Korean art) and works types collections (from Drawings to Textiles). For each artwork an overview and detailed information are provided.
The *Search* is a very heterogeneous section which provides a lot of general information about the Collections, the Special Exhibitions, on how to plan a visit, etc.
The *Events* offers the events calendar and the events divided by type (e.g. Events for families, Events related to the special exhibitions, etc.).
The choice of using the section Collection within the experiment is related to its dimension and to the number of information available. Indeed, the Collection is an information intensive section and it is particularly suitable for being exhaustively analyzed in a limited amount of time. Moreover, this section presents also several interactive features and it gives to the inspectors the possibility to find problems taking into account different dimensions.

Figure 53: the main page of the section collection –
www.clevelandart.org/Explore/

Data Coding

For each participant, the inspection report was analyzed by a usability evaluation expert (called experiment analyst). The experiment analyst has both a strong knowledge of the MiLE+ method and usability analysis/problems reporting. Before explaining in-depth the activities carried out by the analyst, it is important to point out the substantial and conceptual difference in the definition of what a problem is. In this study, for the sake of data analysis, a division between *instance of problem* and *problem* has been made. It is possible to define this difference as follow:

– *Instance of problem:* it is a single and atomic unit or exemplar of problem. For instance, during the usability analysis of a website an inspector could find that in page *a*, in page *b*, *c*…. the font is unreadable (because the font is too small). So, in the evaluation matrix he reports that there is an instance of problem related to the font size in page *a*, in page *b*, *c*, *etc*.

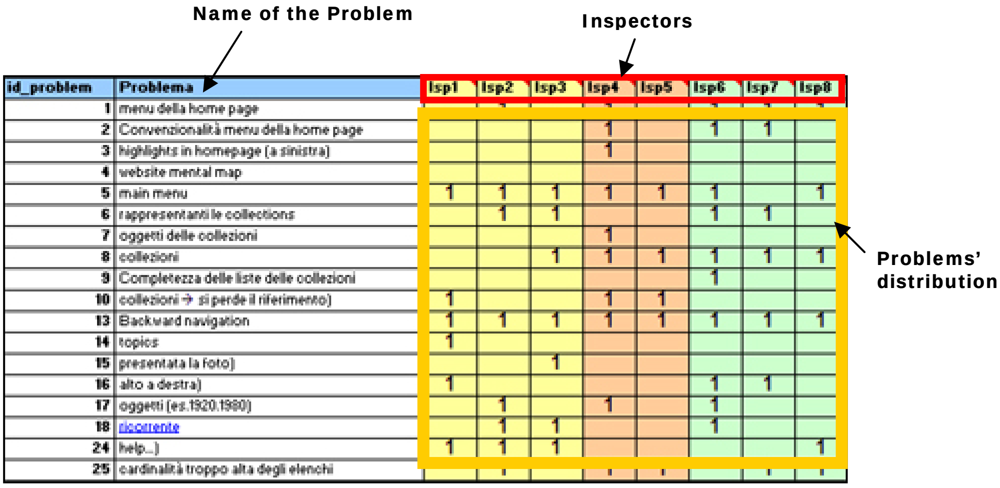
– *Problem:* it can be both an aggregate of instances of problem and a single problem. To be an aggregate means that once the inspector identifies several instances of problem he can homogenize them into the concept of *problem*. Taking again the example of the Font size in page *a*,*b*,*c*, *etc.* above presented, it is possible to uniform it in the problem "Font Size too small". Indeed, they have the same nature and solution. A single problem is a problem which can not instantiated since it is the only one in the entire section under analysis.

The data coding performed by the analyst passed through three main activities.

1. The *first activity* performed by the analyst was the reporting of all the problems' instances. So, starting from the inspectors' evaluation reports, the analyst has completed a matrix inserting all the instances of problems founded by each inspector. The output if this activity was a list of problems' instances.

2. The *second activity* was to homogenize the results of each inspector. The analyst grouped all the instances of problems. The output of this activity was a list of all the problems detected by the inspectors within the three hours of the experiment.

3. The *third activity* was assessing the problems' distribution. The analyst assigned for each problem the values 1 (problem found) or 0 (problem not found) considering each inspector. The output of this activity was a matrix reporting the problems and the distribution for each of them.



id_problem	Problema	Isp1	Isp2	Isp3	Isp4	Isp5	Isp6	Isp7	Isp8
1	menu della home page								
2	Convenzionalità menu della home page				1		1	1	
3	highlights in homepage (a sinistra)				1				
4	website mental map								
5	main menu	1	1	1	1	1	1		1
6	rappresentanti le collections		1	1			1	1	
7	oggetti delle collezioni				1				
8	collezioni			1	1	1	1	1	1
9	Completezza delle liste delle collezioni						1		
10	collezioni → si perde il riferimento)	1			1	1			
13	Backward navigation	1	1	1	1	1	1	1	1
14	topics	1							
15	presentata la foto)			1					
16	alto a destra)	1					1	1	
17	oggetti (es.1920.1980)		1		1		1		
18	ricorrente		1	1			1		
24	help...)	1	1	1					1
25	cardinalità troppo alta degli elenchi								

Figure 54: excerpts of the Como group final matrix

These activities were performed for the analysis of both the Lugano and Como Group.

5.3 Results

5.3.1 Effectiveness and efficiency Hypothesis

To verify the *effectiveness and efficiency* of MiLE+ we consider three measures:

- Average Performance;
- Number of problems found by the inspectors;
- Number of problems detected by a single evaluator (individual effectiveness) and number of problems found in one hour (efficacy).

Average Performance

The first measure considered in our experiment to establish the *effectiveness and efficiency* corresponds to the average performance obtained by the inspectors of the Lugano and Como groups. The performance of a single *i-th* inspector is defined as follow:

$$P_i = \frac{x_i}{N_p}$$

Where:

P_i : performance of *i-th* inspector;

x_i : problems discovered by *i-th* inspector;

N_p : number of problems (which is the total number of problems detected by all the inspectors, see next page "Number of problems found by the inspectors").

So , it is possible to define the average performance each single group as follow:

$$AP_G = \frac{\bar{x}}{N_p}$$

Where:

AP_G : average performance of the group;

$\bar{x}$: average of discovered problems by the group's inspectors;

N_p : number of problems.

The results obtained by the two groups are:

Group	AP
MiLE+ Como group	34,38 %
MiLE+ Lugano group	37,32 %

Table 20: mean of usability problems found by a single inspector

The mean number of usability problems found by a single inspector was respectively 34,38% (Como group) and 37,32% (Lugano Group) of the total number of problems. These results point out the fact that the method should have been understood by the inspectors. Compared to the results obtained by other experiment, in particular Nielsen experiment (Nielsen 1994, pag.44) where the average was 29%, the method seems allowing the inspectors to find in average more problems. It is very interesting to highlight that the mean is similar among the two groups. This means that the method was understood and employed almost in the same way and the effectiveness of the inspectors is comparable.

Number of problems found by the inspectors

The number of problems found by the inspectors is the total of different problems detected in the Collection's section by the inspectors during the three hours of experiment. This number is obtained after the activity of results homogenization made by the analyst.

The numbers of problems are:

Group	# of problems
MiLE+ Como group	42
MiLE+ Lugano group	27

Table 21: total number of problems found by the two groups

There is a great difference between the two groups. The Como inspectors found a total of 15 problems more than the Lugano inspectors. It is difficult to state the reason for this difference, but the most plausible one is that considering their experience in design the Como's inspectors detected more single problems. This means that the method's efficacy and effectiveness is affected by the experience in design of interactive applications. However, 42 and 27 problems detected in only three hours of inspection are remarkable results and underline the pro-activity of MiLE+, which pushes the inspectors to detect the highest possible number of problems considering time constraints. This result confirms the hypothesis that the Como group would have been more effective and efficient with respect to the Lugano Group.

Another interesting analysis concerns the groups' distribution of the problems taking into account the design dimension.

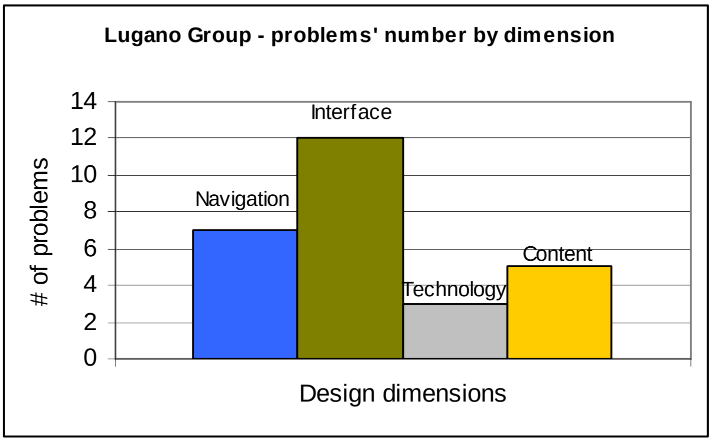


Figure 55: Lugano group – distribution of problems' number considering design dimensions

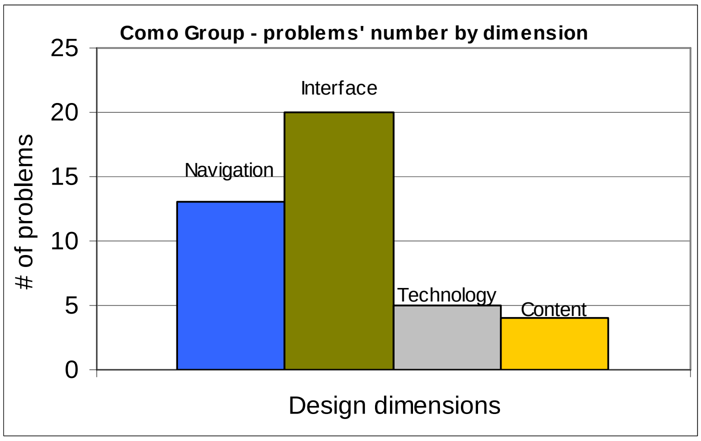


Figure 56: Como group – distribution of problems' number considering design dimensions

Dimension	***Lugano Group - % of problems***	***Como Group - % of problems***
Navigation	25,93	30,96
Interface	44,44	47,61
Technology	11,11	11,90
Content	18,52	9,53

Table 22: % of problems for each dimension

Figure 55, 56 and Table 22 show the distribution of the number of problems considering the design dimension. These distributions are very interesting from different points of view. First of all, they are similar for the two groups. Indeed, most of problems are related to the *interface design* (the Lugano Group found 12 interface problems and the Como Group 20). This is due to the fact that interface design is a broader dimension which includes three sub-dimensions (semiotics, graphics and cognitive aspects) that presents a lot of design elements and therefore they are very complex to manage. As a consequence websites often lack of usability in these aspects and using accurate heuristics it is quite easy to detect problems at the interface level. A first consideration is that, probably, the importance of these dimensions (in particular semiotics) is still underestimated by the project's team. Notice that MiLE+ is one of the first methodologies which point out the importance of semiotic and cognitive design to create usable interfaces (Triacca L. et al.: 2005). Furthermore, it is also important to highlight that the percentage of problems is very similar for the two groups (44,44% for the Lugano Group and 47,61% for the Como Group). This means that both the method really stresses the analysis of the interface and that the inspectors have understood the importance of this design dimension and how to evaluate it.

Navigation is the second dimension for number of detected problems (Lugano Group 7 and Como Group 13 issues). This result is also due to the fact that MiLE+, compared to other methodologies, is strong concentrated in the navigation analysis and provides the inspectors with several tools (heuristics) for its analysis. There is a difference in the percentage of the navigational problems among the two groups (Lugano 25,93% and Como 30,96%): this is probably caused by the different experience in design. However, the result obtained by the Lugano Group is very interesting. Indeed, it shows that first-time inspectors, after a full-immersion-training day, are able to use MiLE+ for the analysis of a very complex dimension such as navigation. This result is the first validation of *navigational problems discovery hypothesis*.

The difference in problems' distribution among the two groups is related to technology and content dimensions. However, the difference is only at distribution level since the number of problems is similar. Indeed, the Lugano Group found 3 technology and 5 content problems; the Como Group detected 5 technology and 4 content issues. This slight difference is referable both to the different background (students in communication are more "sensible" to aspects related to the content, while engineers to those technological) and to the fact that the section could not have usability problems related to these dimensions.

Number of problems detected by a single evaluator (individual effectiveness) and number of problems found in one hour (efficacy)

Number of problems detected by a single evaluator (individual effectiveness)
The measure of *effectiveness* is the number of problems detected by a single evaluator (Individual effectiveness) and the average of the group (Group effectiveness).

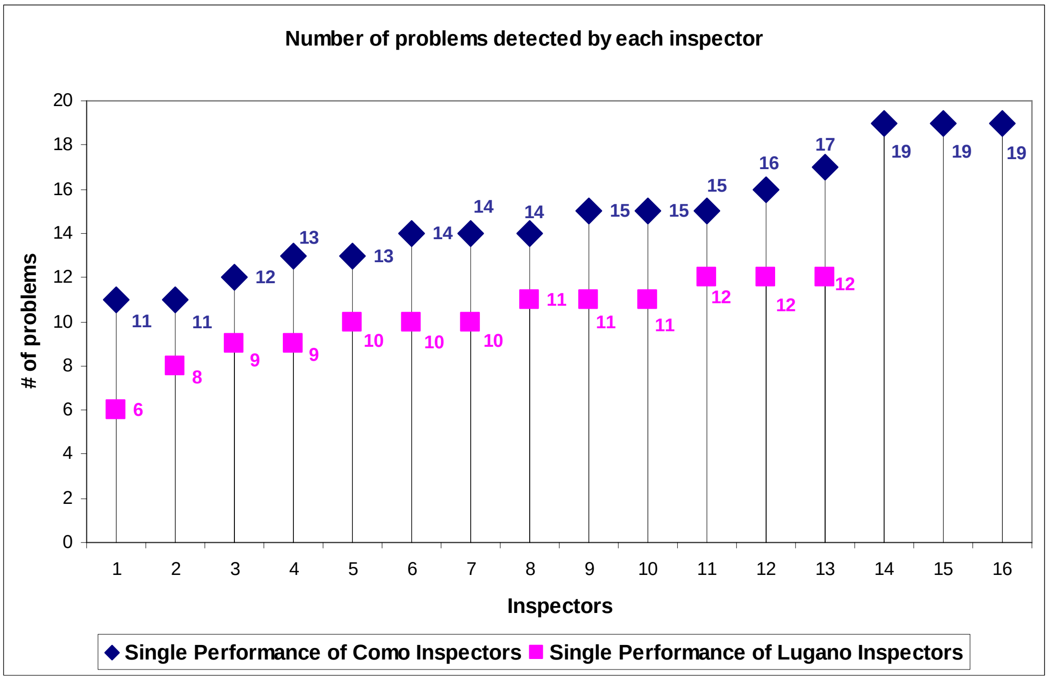


Figure 57: number of problems detected by each inspectors

Average of the Como group: 14,80 problems;
Average of the Lugano group: 10,08 problems.

The effectiveness hypothesis is validated both by the number of problems founded by each inspector and by the groups' averages (14,80 vs. 10,08 problems). As matter of fact, the majority of the Como inspectors has identified more problems than the Lugano inspectors. The more effective inspectors have been *1*, *2*, *3* of the Como group with 19 problems founded in three hours and, for Lugano, inspectors *1*, *3*, *8* with 12 issues detected.

The sub-hypothesis which states that it could be possible to have different results among the three sub-groups of the Lugano group were not verified. In particular, inspectors that had a full-day training (MiLE+8h), should have found more problems compared to those of MiLE+4h and MiLE+2h. Figure 58 shows that two of the "best" three inspectors belonged to MiLE+2h group (they had only a two-hour-training!) and in general there is not a great difference among the groups. Notice that almost all the inspectors performed the technical inspection because of the time at disposal and for the complexity of an in-depth understanding of the User experience inspection. This means that the first two hours of training and the resources available (manuals and online courses) are enough to understand the basic features of MiLE+ Technical inspection and for using it.

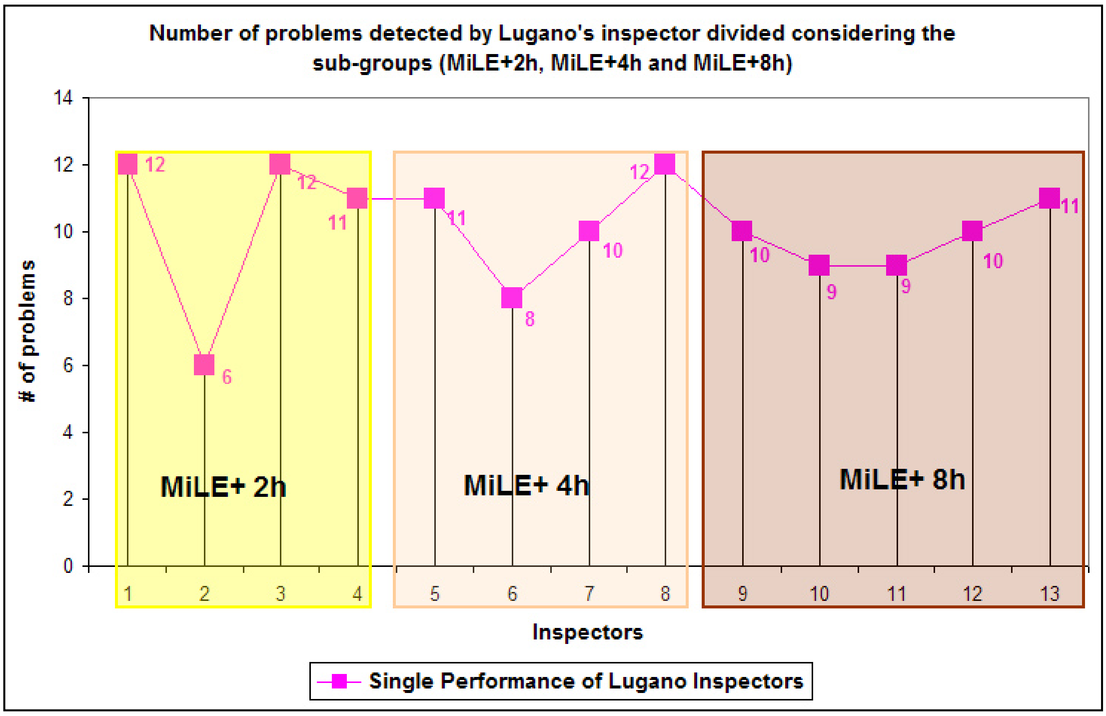


Figure 58: number of problems detected by Lugano's inspectors highlighting the three subgroups

Number of problems found in one hour (efficacy)

The Individual efficiency also refers to the number of problems extracted by a single inspector, in relation to the time spent. It is computed by the following formula (De Angeli A., 2000):

$$Ind\_Efficiency = \frac{P_i}{t_i}$$

where, P_i is the number of problems detected by the *i-th* inspector, and t_i is the time spent to find the problems. The average efficiency of the two groups are:

- Como group: 4,93 problems/hour;
- Lugano group: 3,36 problems/hour.

On average, the Como group inspectors found 4,93 in one hour of inspection, versus the 3,36 of the Lugano Group. These data confirmed the efficient hypothesis: the Como inspectors found more problems in one hour inspection with respect to the Lugano inspectors.

5.3.2 Inter-group and intra-group agreement on findings hypothesises

Considering that in the literature we did not find an appropriate approach for verifying the agreement's hypothesises (and consequentially the reliability among inspectors' results), we had developed a particular approach and a mathematical formula for verifying them.

The formula for the agreement's definition is:

$$AG(Insp_x, Insp_y) = \frac{Insp_x \cap Insp_y}{Insp_x \cup Insp_y}$$

Where:

$AG(Insp_x, Insp_y)$: is the level of agreement among $Inspector_x$ and $Inspector_y$;

$Insp_x \cap Insp_y$: defines the common problems among $Inspector_x$ and $Inspector_y$ (that have been discovered by both the inspectors);

$Insp_x \cup Insp_y$: defines all the problems found by the inspectors (not only those in common).

Appling this formula to every couple of inspectors (*Inspector x; Inspector* y) it is possible to fill the Group Agreement matrix (GA matrix). Each cell of the matrix reports the *agreement value* for every couple.

Inspectors \ Inspectors	1	2	3	4	5	6	7
1		AG1,2	AG1,3	AG1,4	AG1,5	AG1,6	AG1,7
2			AG2,3	AG2,4	AG2,5	AG2,6	AG2,7
3				AG3,4	AG3,5	AG3,6	AG3,7
4					AG4,5	AG4,6	AG4,7
5						AG5,6	AG5,7
6							AG6,7
7							

Figure 59: the GA matrix

Inter-group agreement on findings among Lugano and Como groups

To verify the inter-group agreement of the two groups we unified the problems lists creating a "global" list of problems. Then we selected only those problems in common (which were 9).

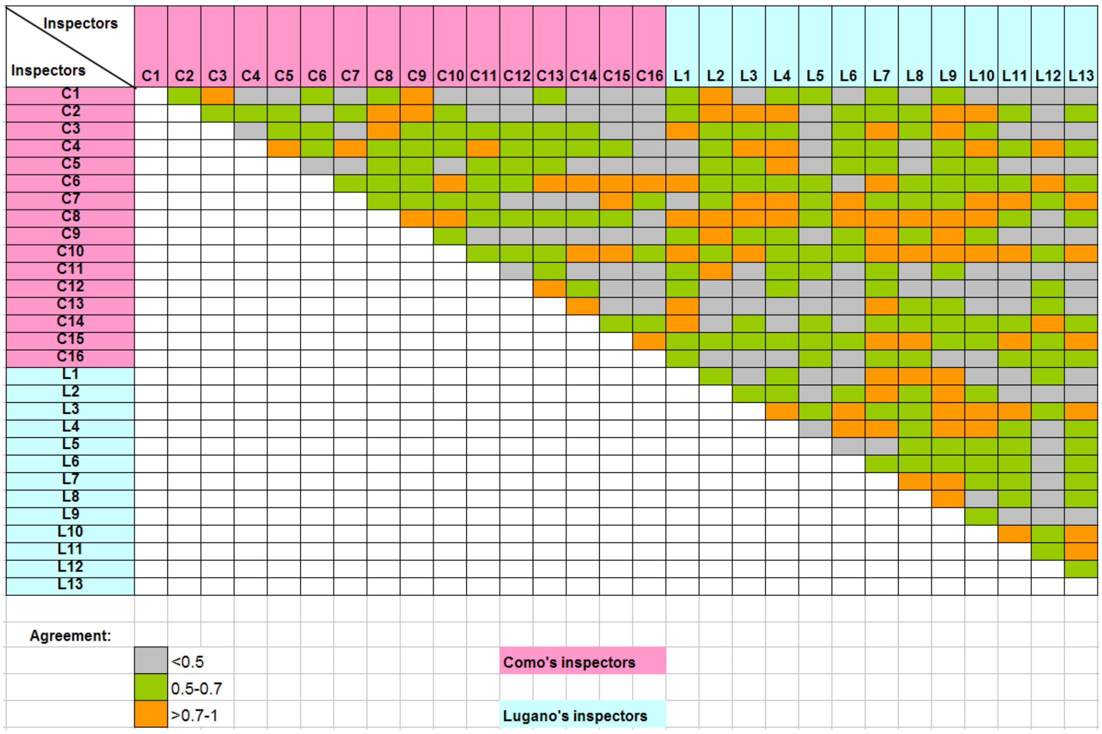


Figure 60: GA matrix showing the agreement on the common problems among Como and Lugano inspectors

Figure 60 shows the agreement among the Lugano and Como inspectors. The mean agreement value is *0.56* (the detailed GA matrix is presented in ANNEX C_1). This means that if a problem was found by an inspector, it was detected by many others. Indeed, the Figure reports that the most part of the agreement values among inspectors' couples is situated between 0.5 and 0.7, with peak values from 0,70 and 1. This result is a first validation of the inter-group agreement on findings hypothesis.

Other interesting agreement's values are the averages of the two groups. Indeed, the average agreement value of the Como group is *0.53* and the one of the Lugano *0.58*. Even if, there is a slight difference, these values demonstrate that the "agreement's behaviour" is similar among the groups.

Intra-group agreement on findings within Como group

The intra-group agreement for the Como group has been calculated considering all the problems found by the inspectors of this group.

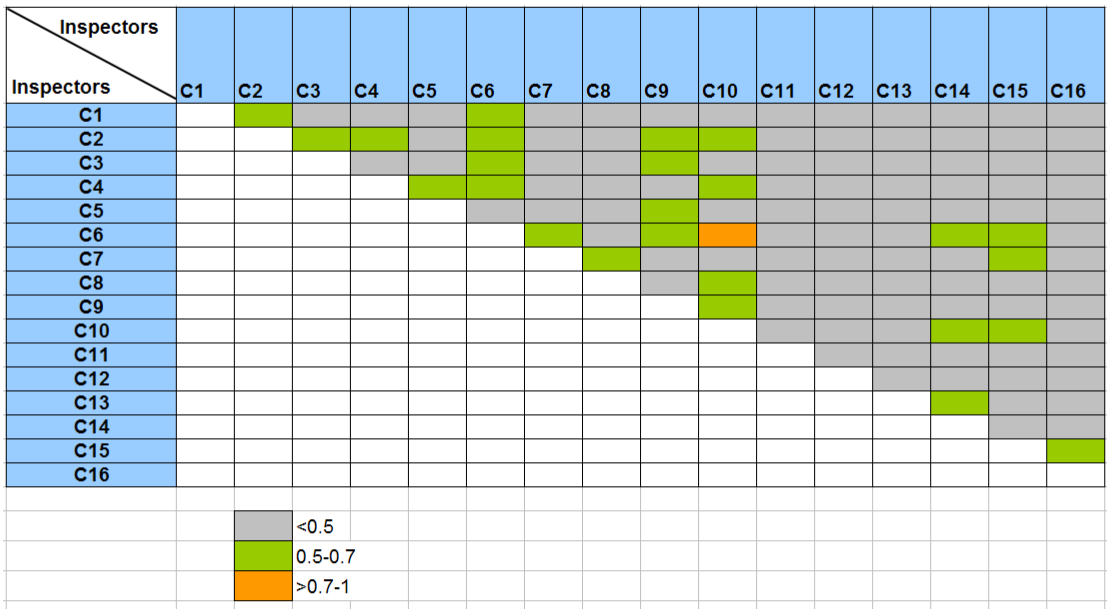


Figure 61: GA matrix for Como group

The agreement within this group is low: the mean agreement value is *0.39* (the detailed matrix is reported in ANNEX C_2). This means that the majority part of the inspectors found problems that are different from those detected by the other inspectors. This could be caused by the unstructured learning of the MiLE+ method. The consequence of this has been that inspectors found several problems (see 5.3.1), but they have been heterogeneous in the findings. Another explanation could be related to the fact that the total number of problem influences the agreement. The more the total number is high, the more the agreement decreases.

Intra-group agreement on findings within the Lugano group

The intra-group agreement for the Lugano group has been calculated taking into account all the problems found by the inspectors of this group.

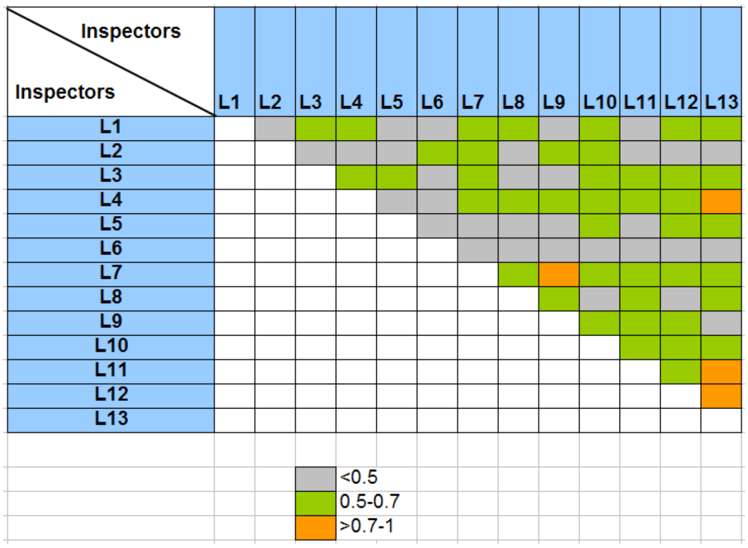


Figure 62: GA matrix for Lugano's group

The mean agreement value obtained by the Lugano group is *0.*53 (for the detailed results, see ANNEX C_3), absolutely higher with respect to the value of the Como group. This is probably due to the more structured training received by the Lugano group, which allows creating a similar behaviour among the inspectors in using the method. It is interesting to notice that this value is similar to the one obtained for Inter-group agreement on findings among Lugano and Como groups.
The most reliable result obtained by the Lugano inspectors with respect to those of Como verifies the intra-group agreement on findings hypothesis.

Intra-group agreement on findings among the Lugano sub-groups

It is possible to verify the agreement among the Lugano sub-groups (MiLE+2h, 4h and 8h) in two ways:

1. assessing the agreement considering the common problems;
2. taking into account all the findings of the Lugano group.

Intra-group agreement on the common problems

The average values of the three sub-groups are:

- MiLE+2h: *0.60*;
- MiLE+4h: *0.58*;
- MiLE+8h: *0.62*.

These results have been obtained coupling each inspector of each sub-group with the other inspectors belonging to the subgroup. For example, the value for MiLE+2h has been found coupling Inspector 1 with inspector 2, 3 and 4; then inspector 2, with 1,3 and 4 and so on for each sub-group (the detailed results are reported in ANNEX C_1).

The values show that within each sub-group there is a satisfactorily agreement among the inspectors. It is possible to state at this point that this is generated by both the method's structure and tools and the teaching.

Intra-group agreement on all the problems

Figure 63 shows the GA matrix for the Lugano group. The average values of the three sub-groups are:

- MiLE+2h: *0.54*;
- MiLE+4h: *0.41*;
- MiLE+8h: *0.59* .

Inspectors \ Inspectors	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13
L1		0,42	0,67	0,67	0,42	0,33	0,58	0,58	0,42	0,58	0,42	0,50	0,50
L2			0,42	0,45	0,36	0,50	0,50	0,33	0,50	0,56	0,33	0,30	0,36
L3				0,67	0,50	0,42	0,50	0,42	0,42	0,67	0,58	0,58	0,67
L4					0,45	0,45	0,64	0,58	0,55	0,55	0,55	0,55	0,73
L5						0,36	0,36	0,42	0,36	0,55	0,45	0,55	0,64
L6							0,40	0,33	0,40	0,44	0,44	0,30	0,45
L7								0,58	0,80	0,60	0,60	0,60	0,55
L8									0,50	0,33	0,58	0,42	0,67
L9										0,50	0,50	0,50	0,45
L10											0,56	0,60	0,55
L11												0,70	0,82
L12													0,73
L13													
	0,58	0,43	0,53	0,62	0,38	0,37	0,45	0,44	0,49	0,55	0,64	0,63	0,64

	average:
MiLE+2h	0,54
MiLE+4h	0,41
MiLE+8h	0,59

Figure 63: GA matrix for the Lugano group which highlights the value obtained by both the single inspector of each sub-group and the average value for each sub-group.

Once more, the values (in particular those of MiLE+2h and MiLE+8h) demonstrates a good level of agreement between sub-groups' inspectors. However, it is important to underline that the mean agreement value of MiLE+4h is considerable lower compared to the other sub-groups. There is not a clear and full explanation to describe the reason for this difference. It could be possible that the four-hour-training, which introduces some new concepts with respect to the two-hours-training, introduces some notions (in particular the user-experience inspection) which are not deepened enough. The consequence of this unclearness could be an heterogeneous and unstructured behaviour in applying the method and this could cause a low agreement among the inspectors.

5.3.3 Navigational problems discovery hypothesis

The aim of this section is to verify the hypothesis which states that MiLE+ guides inexperienced evaluators in finding navigational problems. To validate (or invalidate) this hypothesis we have analyze the distribution of the problems taking into account both the inspectors' attitude to find problems and the level of difficulty to detect the problem. The inspector ability has been defined considering the total number of problems s/he was able to find. The level of difficulty has been established taking into account that if a problem has been found by several inspectors it is an easy problem to detect; on the contrary if an issue has been discovered by few inspectors it was considered difficult to detect. For representing the distribution considering these assumptions we have used a matrix (Figure 64), in which on the abscissa axis the inspector ability to find usability problems are represented and on the ordinate axis the difficulty of problem detection.

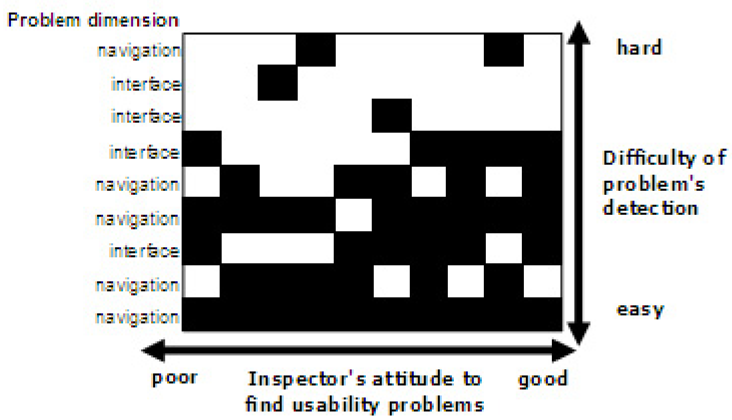


Figure 64: example of the matrix used for representing problems' distribution and the inspectors' attitude .

Each column corresponds to one evaluator, and each row to one usability problem. Each square indicates whether one evaluator found one problems. That square is black if the evaluator assigned to the column found the problem assigned to the row; and white if that evaluator did not find that problem.

Lugano group distribution of problems detection among the inspectors

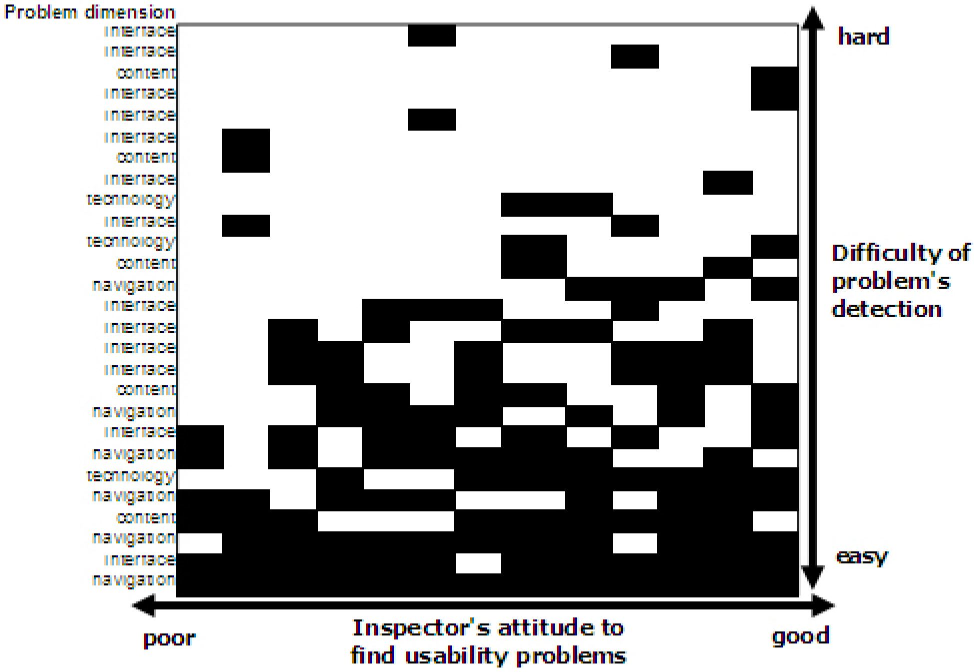


Figure 65: Lugano group problems distribution

Figure 65 shows that the easier problem to find for the Lugano inspector is related to the navigational dimension, followed by one related to the interface and by another navigational problem. Notice that among the five easiest problems three are navigational and among the nine easiest five are related to this dimension. Even if the results presented in 5.3.1 show that most problems are associated to the interface dimension, the easiest to find are navigational.

Another important consideration is related to the problems' distribution: most problems are situated in the bottom section on the right. This could be a symptom that the structured training positively influences the attitude, and therefore allows to find several problems. Considering the fact that a difficult problem is defined taking into account how often it is found by the inspectors, this distribution shows that the training and the tools should facilitate the evaluation activity (for example, it is possible that employing correctly the heuristics to find the problems that normally are hard to discover, the evaluation becomes an "easier" activity).

Como group distribution of problems detection among the inspectors

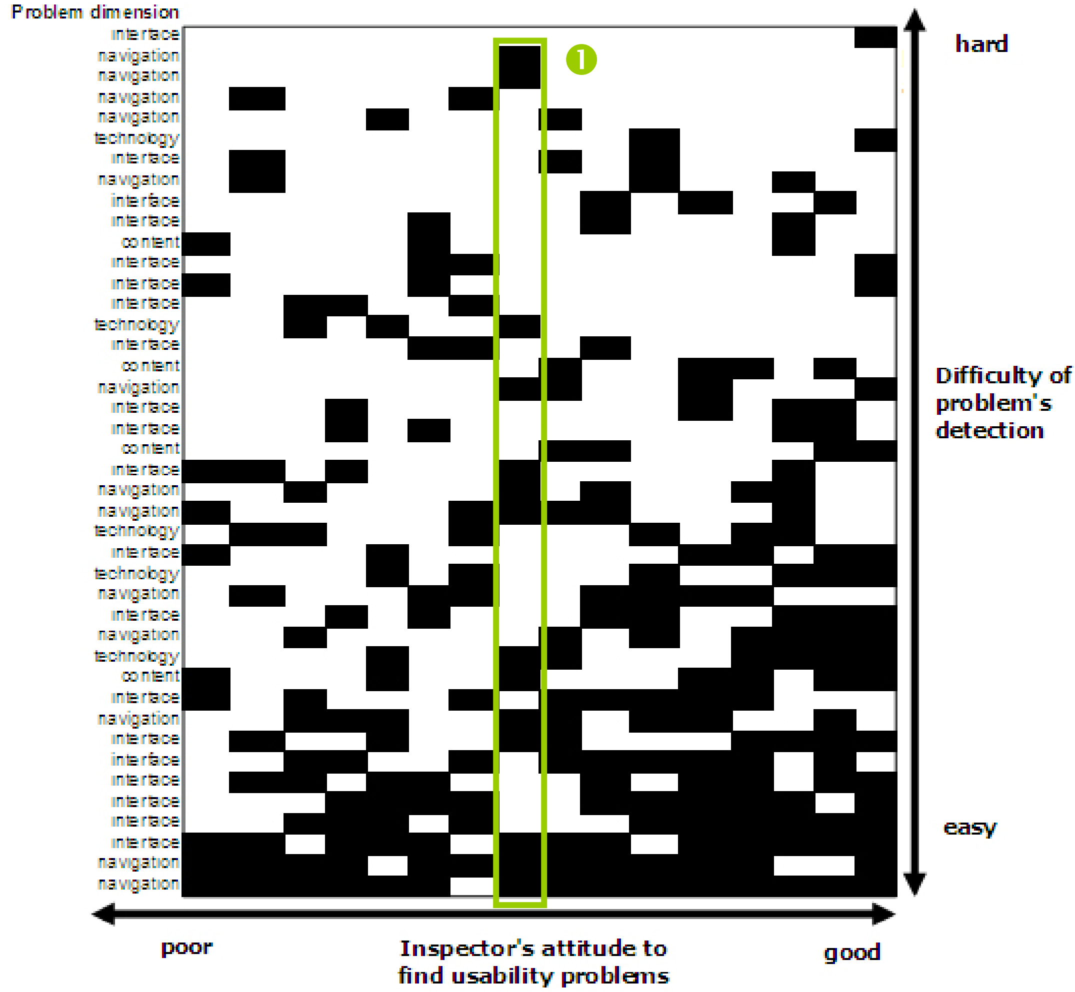


Figure 66: Como group problems distribution

Figure 66 shows that also in the case of the Como group the easiest problems to detect are navigational. However, interface design issues are found with more easiness compared to the Lugano group.

The problems' distribution is less heterogeneous with respect to the Lugano Group. Indeed, without a structured training the inspector's personal attitude emerges even more. The consequences of the unstructured training could be that the inspector does not have the possibility to extend his/her attitude. For example the inspector C5 (❶) has a clear attitude to find navigational problems (indeed s/he found two navigational problems that were not discovered by others).

Lugano and Como distribution of problems detection among the inspectors considering the common findings

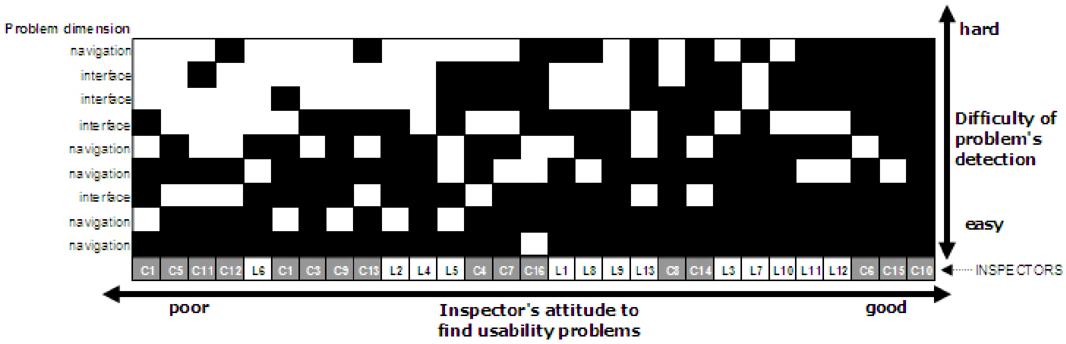


Figure 67: Lugano and Como distribution of problems' detection among considering the common findings

Figure 67 shows that among the common problems the easiest problems to detect are always navigational. In particular, between the five easiest problems four are navigational.
Notice that considering the inspectors' distribution, there is an interesting alternation between Como (C1, C2, C3 etc.) and Lugano (L1, L2, L3 etc.) inspectors, even though Lugano inspectors seem to have a better attitude to find usability problems (these could be caused by the structured training they received): they found less problems but the attitude is better.

5.4 Experiment conclusions

From the analysis of the experiment's results it is possible to draw some interesting conclusions:

- MiLE+ has proved that it is able to conduct inexperienced inspectors to perform an efficient and effective inspection. If we consider that in only three hours of inspection they found a total of 27 (Lugano Group) and 42 (Como group) problems, this means that the structured approach provided by MiLE+ and the tools for learning and applying it really helps the inspectors. Moreover, the single performance of the inspectors is a positive measure on how they used the method. Notice that the Como inspectors had performed the evaluation activity without specific training on MiLE+ and those of the Lugano Group (that are very inexperienced people both in design and in evaluation) received a training which lasted from only two hours to a maximum of height. The fact that they have been able to conduct a usability evaluation with interesting results is absolutely remarkable;

- As assumed, the most part of the inspectors had carried out a technical inspection. This means that they have well-judged (and so understood) the ratio between time at disposal and dimension of the application under analysis. However, the same experiment should be planned focussing in particular on the user-experience inspection. This could be one of the limitations of this experiment which assesses and validates in particular the technical inspection. But, it is important to point out that it could be an arduous and challenging task to plan an experiment for verifying in a quantitative way the user-experience inspection (which for definition investigates aspects that are subjective and depending on the user profiles);

- One of the most interesting considerations is emerged analysing the results of the Lugano group and its sub-groups. It seems that technical inspection is applied almost in the same way by people only having two hours of training compared to those having four and eight hours. This means that the first two hours of the course (with the support of the learning material) create a "zone of proximal development" (Vygotsky, L. S.: 1978) for technical inspection. The learners of the two-hour-course are able to successfully employ the conceptual notions provided: they shift from the conceptual dimension to the action. It is interesting to observe that it seems that the inspectors of four-hour-training are both less effective and reliable. This could be due to the fact that within the four hours, with respect to the two hours, several new concepts are introduced and this could affect the performance of the inspectors. In particular, it could be possible that basic notions related to the user-experience inspection, without practice does not provide enough background for correctly applying them. Moreover, it could be possible that these new notions and the attempt of using them without a sufficient background, affect the overall performance of the inspectors.

– The agreement values (both inter and intra groups), except that of Como, demonstrated that inspectors employing MiLE+ obtained reliable results. All the values are higher than *0.5.* Notice that the time at disposal with respect to the size of the collection's section could have influenced the agreement. Indeed, it is assumable that the agreement's values would increase in proportion to the time at disposal maintaining unchanged the application to analyze. For example, if the inspector had five hours for performing the inspection, it is possible that the agreement (and consequentially the results reliability) intra and inter-groups (and sub-groups) would have been higher than those obtained in three hours (it could happen a convergence phenomenon). However, it should be very interesting to investigate these considerations, repeating the same experiment and varying the time at disposal.

– The low agreement value obtained by the Como inspectors could be ascribable to the unstructured training they received. This shows that the training is one of the most important aspects for obtaining reliable results among the learners. Furthermore, it seems that an unstructured training does not give the possibility to extend the personal attitude of the learners. Another hypothesis to verify is that the total number of problems influences the agreement on findings. In fact, it could be possible that the more the *group A* finds a higher total number of problems with respect to the *group B*; the more the agreement on findings value intra *group A* is lower with respect to this of *group B*.

Chapter 6:

Conclusions and Future Work

Summary:
The first part of this Chapter highlights the key contribution of this work and point out the benefits and limitations of the MiLE+ method.

The Chapter ends with the future work both to further improve and test the method and to validate and assess it.

6.1 Conclusions

Key contributions

The key contributions provided by MiLE+ can be identified within the conceptual model and its reusability.

Conceptual model

One of the most important and innovative contribution is referred to the introduction of the clear separation between application dependent and independent analysis. It is important to underline that this approach is not completely new (Matera et al.2002), but MiLE+ provides the first attempt to systematize it within the usability evaluation activity. The systematization is made providing three different activities (Technical Inspection, User Experience Inspection and Scenario-based user testing) and several tools supporting the evaluation activity. These tools, in particular the technical heuristics' library and the User Experience Indicators' library, have not the pretension to be the definitive and all-inclusive lists of heuristics. On the contrary, they are "open-source", which means that everyone has the possibility to add or delete the heuristics. It is also very important to highlight that MiLE+ proposes a particular and proactive approach to the heuristics definition (see ANNEX A_1).

Another very important feature of the method is its cost-effectiveness. Indeed, within Chapter 3 we have showed (with the help of UMC matrix), how the MiLE+ activities can be employed considering constraints such as time and resources at disposal. The modularity of the method allows a tailor-made analysis taking into account different evaluation's contexts.

Teaching

This work has highlighted the importance of a structured teaching to increase the method reusability. Indeed, teaching is a basic support activity for a successful method's learnability. However, this work points out also the importance of a structured and modular approach to the teaching activity. Within Chapter 4 it has been presented a very detailed instructional design, which is the basis for an efficient and effective teaching of the method. The approach proposes mixes together conceptual definitions with the practice. Indeed, the usability evaluation is an activity which needs practice to be deeply understood. This approach may be overly pragmatic, but after several years of teaching assessment, the results obtained by this didactical strategy are encouraging (it is important to underline that every year these courses are followed by more then 100 University students - both at a bachelor and master level-, researchers and practitioners).

Validating

The experiment presented in Chapter 5 has demonstrated the reusability of the method. Indeed, it has highlighted the ability of inexperience inspectors to apply the method in an affective and efficient way. Moreover, the experiment has pointed out the fact that the level of agreement intra and inter-groups is influenced by the learning activities provided to the learners. The more the learning activity is structured the higher the agreement among the learners. However, it is important to point out that this interesting results have to be verified in further experiments that could establish the reliability of the experiment results.

The interesting research question emerged during the experiment planning, execution and data analysis, is related to the statistical methodology to employ for verifying MiLE+ reusability hypothesis. Considering that within the literature we did not find an exhaustive approach to the problem, we have developed a specific method for the definition of the agreement in the findings among inspectors. Also the reliability of this approach should be tested applying the same approach within other experiments.

Benefits of MiLE+

The adoption of MiLE+ creates several benefits, in particular:

- the systematic approach and the process leads the inspector to a detailed and deep analysis of the web sites;
- the tools (technical heuristics, scenarios, UEIs, etc.) help the inspector to be proactive (they suggest the aspects to observe);
- the use of scenarios as driver for the inspection allow concentrating the evaluation on the most important parts of the application and taking into account the main user profiles and their goals;
- the reusability is given both by the conceptual framework, by the teaching activities and by support material;
- the cost-effectiveness of MiLE+ provides tailored made activities considering the resources at disposal;
- the flexible conceptual structure of the MiLE+ model could be easily adapted for the evaluation of several types of interactive applications (not only web sites).

Limitations of MiLE+

Some limitations of the current version of MiLE+ have also be noted, in particular:

- it does not provide heuristics for the evaluation of the operational aspects of the applications;
- MiLE+ is one of the first methodology which explores in-depth the semiotic dimension, but at present it does not explore enough in detail this aspect;
- even though the conceptual model should allow its application to different types of applications, at the moment it has been adopted in a limited number of cases outside traditional websites' evaluation;

Future research will be devoted to address these and other issues, in order to deliver a better and more usable usability evaluation method.

6.2 Future work

Since several years we have been working to improve MiLE and afterwards MiLE+. We will continue to improve it following three lines of action: *enhancement*, *validation* and *testing*.

Enhancement

Considering the MiLE+'s state-of-the art the enhancement should be concentrate both on the conceptual approach and the tools. The conceptual enrichment should be focussed on the continuous investigation on the nature of the problems, in particular going in depth in understanding the difference between application dependent and independent issues.
The line of action related to the tools should be focussed on the development of technical heuristics taking into account both new design dimensions (e.g. operational) and other types of applications. Indeed, considering the evolution of multichannels applications it should be very important to develop specific libraries of heuristics to evaluate this new kind of applications. Another dimension to be explored more in detail is semiotic design. Indeed, at the moment we are working on a model for tsemiotic analysis of web site: this model could help to improve the library of semiotic heuristics.
Another tools enhancement is related to the U-KITs. At present, we are developing a specific U-KIT for Cultural Applications (e.g. museum websites, digital libraries, etc.) and the future work will focus on the development of U-KITs for other domains (e.g. banks, tourism, educational, etc.).

Furthermore, the inspector manual should be improved. The actual version is the 1.0 and it could be enhanced (even though the actual version has demonstrated to be useful for the inspectors). In particular, the part in which the User Experience Indicators are explained could be improved, presenting good and poor examples of UEIs' employment. Indeed, the actual version explains in details the technical heuristics, but it does not give enough detail about UEIs.

Validation

The validation activity should be concentrated on the replication of the experiment for assessing the reliability and reusability of the method. In particular, the experiment should pay attention on the validation of the user-experience inspection. So, it should be planned taking into account that this activity is more complicated to validate, as it has several subjective variables. Moreover, the experiment should verify the reusability and effectiveness taking into account different categories of inspectors. Indeed, the first version of the experiment has been carried out analysing the impact of the method on inexperienced inspectors; the second version should be performed considering categories such as experience evaluators or type of background (e.g interface designers, engineers, etc.).

Testing

The testing activity should be focussed on several case studies. As mentioned before, at present MiLE+ has been adopted to evaluate web sites. In the future it will be always very important to employ it to evaluate web sites, but it will be interesting to use it for the analysis of other types of interactive applications (e.g. videogames). The employment of MiLE+ out of traditional websites boundaries could be very useful for verifying the validity of its conceptual approach. The final result of the method's expansion could be a "universal" (meta) model for evaluating interactive applications.

References

Ackermann, A.F., Buchwald, L.S., and Lewski, F.H. (1989), *Software inspections: An effective verification process*, IEEE Software 6,3 pp. 31-36.

Basili, V. (1994), *GQM approach has evolved to include models*, IEEE Softw. 11, 1, 8.

Blackmon, M.H., Polson, P.G., Kitajima, M., & Lewis, C. (2002), *Cognitive Walkthrough for the web*, in CHI 2002 Conference on Human Factors in Computing Systems. ACM Press.

Bolchini D., Triacca L., Speroni M. (2003), *MiLE: a Reuse-oriented Usability Evaluation Method for the Web*, HCI International Conference, Crete, Greece.

Botturi L. (2003), *E2ML - Educational Environment Modeling Language*, PhD Thesis, University of Lugano.

Botturi, L. (2004), *Visualizing Learning Goals with the Quail Model*, Australasian Journal of Educational Technologies - AJET 20(2), 248-273.

Botturi, L. (in print, 2005), *E2ML. A Visual Language for the Design of Instruction*, *Educational Technologies Research & Development*.

Brinck, T., Gergle, D., Wood, S.D. (2002), *Usability for the web*, Morgan Kaufmann.

Cantoni, L., Di Blas, N., Bolchini, D. (2003), *Comunicazione, Qualità, Usabilità*, Apogeo, Milano.

Carroll, J. (2002), *Making Use – Scenario-based design of Human-Computer Interactions*, MIT Press.

Cato, J. (2001), *User-Centred web Design*, Addison Wesley.

Chi, E. H., Pirolli P. and Pitkow J. (2000), *The scent of a site: A system for analyzing and predicting information scent, usage, and usability of a website*, In Proceedings of the Conference on Human Factors in Computing Systems, pp. 161- 168, The Hague, The Netherlands, New York, NY: ACM Press

Chi, E.H., Adam S. Rosien, A.S. and Heer, J. (2002), *LumberJack Intelligent Discovery and Analysis of web User Traffic Composition*. In Proc. ACM-SIGKDD Workshop on web Mining for Usage Patterns and User Profiles (webKDD 2002), ACM Press, Edmonton, Canada.

Clemmensen T. (2002), *Interest in theory and method in a Danish community of usability specialists*, Working paper, Copenhagen Business School, Department of Informatics, HCI research group, pp. 1-30.

De Angeli A., Matera M., Costabile F., Garzotto F., Paolini P. (2000), *Validating the SUE inspection technique*, Proceedings of the working conference on Advanced visual interfaces, Palermo.

Di Blas N., Guermandi M.P. Orsini C., Paolini P. (2002), *Evaluating The Features of Museum websites*, Museum & web Conference, Boston.

Dix, A., Finlay, J., Abowd, G., and Beale, R. (1998), *Human Computer Interaction*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall.

Donahue G.M, Weinschenk S., Nowicki J. (1999), *Usability Is Good Business*, Compuware Corporation Report.

Dumas J.S., Redish J.C. (1999), *A practical Guide to Usability Testing*, Intellect 1999.

Dunker K. (1976), *On problem-solving*, in Dashiell, John F.: Psychological Monographs, The American Psycological Association Inc., Washington DC, vol58, pp 1-114 (reprint, Westport, conn.Greenwood).

Ereback A., Höök K. (1994), *Using Cognitive Walkthrough for Evalauting A CSCW Application*, CHI'94, Boston, USA.

Gagné R.M., Briggs R., Wager W. (1992), *Principles of Instructional Design*, 4th ed. Orlando, FL: Harcourt Brace Jovanovich.

Garzotto, F., Matera, M., Paolini, P. (1998), *To Use or Not to Use? Evaluating Usability of Museum Web Sites,* Museums and the Web 1998. Selected Papers from an International Conference, Archives & Museum Informatics, Toronto, CA

Gilb, T. (1998), *Principles of Software Engineering Management*, Addison Wesley, Reading, MA.

Green, T.R.G. & Benyon, D.R. (1996), *The skull beneath the skin; Entity-relationship modeling of information artifacts*, int. J. Hum.-Comput. Stud., vol. 44, no. 6, pp. 801-828.

Holtzblatt, K., & Beyer, H. (1996), *Contextual design, principles and practice*. In D. Wixon & J. Ramey (eds.), Field Methods Casebook for Software Design (pp. 301–333). New York: John Wiley & Sons.

Kuniavski, M. (2003), *Observing the User Experience: A Practitioner's Guide to User Research*, Morgan Kaufmann.

Ivory, M., Hearst, M. (2002), *Statistical Profiles of Highly-Related web Sites*, In Proceedings of CHI Conference, Volume 4, Issue No. 1, pp. 367-374.

Jeffries R., Miller, J., Wharton, C., and Uyeda, K.M. (1991), *User interface evaluation in the real world: A comparison of four techniques*, in Proc. ACM CHI, pp. 119-124.

Kitajima, M., Blackmon, M.H., & Polson, P.G. (2000), *A comprehension-based model of web navigation and its application to web usability analysis*, in S. McDonald, Y. Waern & G. Cockton (Eds.), People and Computers XIV – Usability or Else! (Proceedings of HCI 2000, pp. 357-373)

Krug S. (2000), *Don't make me think. A Common Sense Approach to web Usability*, Que.

Lazarsfeld P.F. (1972), *Qualitative analysis: Historical and critical essays*, Allyn and Bacon, Boston.

Lewis C., Rieman J. (1993), *Task-Centered User Interface Design - A Practical Introduction*, Distributed via anonymous ftp.

Lim, K.H., Benbasat, I. and Todd, P.A. (1996), *An experimental investigation of the interactive effects of interface style, instructions, and task familiarity on user performance*, ACM Trans. Comput. Hum. Interact., vol. 3, no 1,pp. 1-37, January 1996.

Lonergan B. (1957), *Insight: A study of Human Understanding*, London: Longmans, Green and Co. and New York: Philosofical Library.

Madsen, K.H. (1999), *Special issue on the diversity of usability practices*, Commun. ACM, vol. 42, no. 5.

Matera, M., Costable M.F., Garzotto F., Paolini P. (2002), *SUE Inspection: An Effective Method for Systematic Usability Evaluation of Hypermedia*, IEEE Transaction, Vol.32, No. 1, January 2002.

Merton R.K., Fiske M. and Kendall P.L. (1956), *The focused interview: A manual of problems and procedures*, the Free Press, Glencoe.

Morgan D.L. (1988), *Focus groups as qualitative research*, Sage, Newbury Park.

Newell, A., and Simon H.A. (1972), *Human problem solving*, Englewood Cliffs, NJ: Prentice Hall.

Nielsen, J., Mack, R. (1994), *Usability Inspection Methods*, Wiley.

Nielsen, J. (1999), Designing web Usability, New Riders.

Nielsen J. (2003), *Return on Investment for Usability*, www.useit.com/alertbox/20030107.html, Alertbox, January 7, 2003.

Nielsen, J., Clemmensen T., Yssing C. (2002), *Getting access to what goes on in people's heads? - Reflections on the think-aloud technique*, In Proceedings of Nordchi Conference, Denmark.

Preece J. (1994), *Human-Computer Interaction*, Addison-Wesley, England.

Preece J., Roger Y., Sharp H. (2002), *Interaction Design: Beyond Human Computer Interaction*, John Wiley & Sons.

Pressman R.S. (1992), *Software Engineering: A Practicioner's Approach*, McGraw-Hill, New York.

Polson, P. G., Lewis, C., Rieman, J., & Wharton, C. (1992), *Cognitive walkthroughs: A method for theory-based evaluation of user interfaces*, in International Journal of Man-Machine Studies, 36, 741-773.

Ratner, J., Grose, E.M. and Forsythe, C. (1996), *Characterization and assessment of HTML style guides*, In Proceedings of the Conference on Human Factors in Computing Systems, volume 2, pp. 115-116, Vancouver, Canada, April 1996, New York, NY: ACM Press.

Roethlisberger, F.J. and Dickson, W.J. (1939), *Management and the Worker*, Boston, Mass., Harvard University Press.

Rosson, M.B., Carroll, J. (2002), *Usability Engineering*, Morgan Kaufmann.

Smith P.L., Ragan T.J. (1999), *Instructional Design*, Wiley and Sons, New York 1999, pag.2.

Sullivan P. (1991), *Multiple Methods and the Usability of Interface Prototypes: the complementary of laboratory observation and focus groups*, Proceedings of the 9th annual international conference on Systems documentation, Chicago.

Triacca, L., Bolchini D, DiBlas N., Paolini P. (2002), *Wish you were usable! How to improve the quality of a Museum Web site*, EVA 2003, Florence, Italy.

Triacca L, Bolchini D., Botturi L., Inversini A. (2004), *MiLE: Systematic Usability Evaluation for E-learning Web Applications*, ED Media 04, Lugano, Switzerland.

Triacca L., Speroni M., Bramani C. (2005), *Understanding Semiotic issues in Usability Evaluation of Cultural Heritage Websites: the DICE case study*, Studies in Communication Sciences, No. 2., Lugano, 2005 (to appear).

UWA Consortium (2002), *The UWA Approach to Modeling Ubiquitous Web Applications*. Proc. of IST Mobile conference.

Vygotsky, L. S. (1978), *Mind in society: The development of higher psychological processes*, Cambridge, MA: Harvard University Press.

VNET5 Consortium (2002), *Best Practice Manual for User Centred Product Creation in Interactive Electronic Publishing*.

Wharton C., Rieman J., Lewis C, Polson P. (1994), *The Cognitive Walkthrough Method: A Practitioner's Guide*, In Usability Inspection Methods, J. Nielsen and R.L. Mack(Eds.), New York: John Wiley&Sons, pp.105-141.

Whiteside J., Bennet J., and Holtzblatt K. (1988), *Usability engineering: Our experience and evolution*, in Handbook of Human-Computer Interaction, M.Helander, Ed. Amsterdam, The Netherlands, North-Holland, pp.791-817.

Withrow, J., Brinck, T., Speredelozzi, A. (2000), *Comparative Usability Evaluation for an e-Government Portal*, Diamond Bullet Design Report 1-00-2, Ann Arbor, MI. Dec. 2000.

ANNEX A:

Inspector Manual (version 1.0)

Summary:
The Inspector Manual is composed of four documents that help the inspectors both to better understand the tools and the different libraries offered by MiLE+.
These documents are:

- Library of Technical Heuristics (Annex A_1);
- Library of User Experience Indicators (UEIs) – (Annex A2);
- Applying Technical Heuristics and User Experience Indicators: some examples (Annex A_3);
- Guidelines for Reporting Usability Problems (Annex A_4).

MiLE+

(Milano-Lugano Evaluation method)

Library of Technical Heuristics

Index

READING GUIDE .... - 145 -
PROCESS OF USE .... - 146 -

CONTENT .... - 147 -
1a. CONTENT HEURISTICS .... - 148 -
1b. CONTENT ACTIONS .... - 149 -

NAVIGATION .... - 152 -
2.a NAVIGATION HEURISTICS .... - 153 -
2.b NAVIGATION ACTIONS .... - 158 -

INTERFACE DESIGN .... - 169 -
3.a SEMIOTICS HEURISTICS .... - 170 -
3.b SEMIOTICS ACTIONS .... - 172 -
3.c COGNITIVE HEURISTICS .... - 176 -
3.d COGNITIVE ACTIONS .... - 177 -
3.e GRAPHICS HEURISTICS .... - 181 -
3.f GRAPHICS ACTIONS .... - 183 -

TECHNOLOGY/PERFORMANCE .... - 190 -
4.a TECHNOLOGY/PERFORMANCE HEURISTICS .... - 191 -
4.b TECHNOLOGY/PERFORMANCE ACTIONS .... - 192 -

HEURISTICS SYNOPTIC TABLES .... - 197 -

READING GUIDE

The presentation of every Technical Heuristic is structured considering two different levels of learning: explanation (what) and action (how). This distinction is made considering different level of knowledge of the readers: if you are an experienced inspector the description should be enough for understanding and applying the heuristics; on the contrary the actions and some examples should help you in using correctly them. The examples provided are related both to the right application and to infringements of the heuristic *(note: the websites used for the examples have been visited from August until December 2004).*

Moreover, the heuristics are presented in a proactive way: in fact they are related to the feature to observe: so the starting point is the feature (e.g. text) and then are described all the problems related to it. The heuristic is composed by the combination of feature and single problem (e.g. text accuracy)

Example:

Feature to analyzed

Problem

Problem's explanation

Feature	**Text**
Problem	**Accuracy**
Explanation	The accuracy states if a text describes adequately the referenced world, and if it is consistent in itself.
Problem	**Currency**
Explanation	The electronic communication over the web is supposed to be delivered in the precise moment the reader accesses it; thus the offered content must be current as the addressee perceives it, or must clearly show when it was published and the time scope of its validity.

Problem's actions

Feature	**Overall graphic design**
Problem	**Background contrast**
Action	1. Verify if the background used does not obstacle the reading. 2. Verify how it influences the look of pages and the location of all other elements on the screen.
Example	 In this example (www.provincia.potenza.it/museo/default.htm) there is not contrast between the background (green) and the caption of this image (orange). For the user is very difficult to read the text that explain the image.

PROCESS OF USE

There are several ways for using the heuristics, in particular:

- *Random evaluation;*
- *Feature-driven evaluation;*
- *Scenario-driven evaluation.*

Random evaluation

This type of evaluation is useful when the application has small dimensions or for evaluating websites' sections (if it is a narrow analysis). In this situation the inspector should navigate randomly within the application verifying the compliance with the heuristics. This activity should be performed only by expert evaluators that know very well the MiLE+ Technical Heuristics (they are able to identify a problem and link it with the right heuristics).

Feature-driven evaluation

The starting point of this evaluation is a library of heuristics (heuristics list) created by the inspector before the analysis. The process is very easy: the inspector should identify the features (e.g. a text, the overall graphic design, etc.) and verify if there is a problem related to them (e.g. lack of accuracy within the text). It is advisable to use this technique in the case of small website. Once the inspector find a problem related to a feature's instance, s/he should verify if the problem is reiterated. For example if the inspector

Scenario-driven evaluation

For carrying out this activity the first step is the creation of the scenarios' library and heuristics' library (in practice s/he should create the U-KIT). Then the inspector has to use the scenarios for navigating with clear goals within the application (so the inspector can concentrate his evaluation on the most important parts of the website) and for every task or goal (it depends from the level of granularity used for the analysis) s/he applies the heuristics. This technique is very useful for large websites.

CONTENT

1a. CONTENT HEURISTICS

The content level analyzes the quality of the content (in term of efficacy of the communication) and it allows for verifying if the contents and their structure correspond with the expectations of the users.
The goal of the content heuristics is to verify the "technical" quality of the content presented in web applications.

Feature	**Text**
Problem	**Accuracy**
Explanation	The accuracy states if a text describes adequately the referenced world, and if it is consistent in itself.
Problem	**Currency**
Explanation	The electronic communication over the web is supposed to be delivered in the precise moment the reader accesses it; thus the offered content must be current as the addressee perceives it, or must clearly show when it was published and the time scope of its validity.
Problem	**Coverage**
Explanation	The coverage defines the borders of the topics covered by the given website. It must be clear what the text is speaking about and what it is supposed to be covered.
Problem	**Content objectivity**
Explanation	The content objectivity indicates the commitment of the sender with respect to the conveyed content. For example, it must be clear if a message is an advertising or not (if the sender is paid to say something, I do not think that he must be really convinced of what he is saying…).
Problem	**Authority**
Explanation	Authority could be seen under two respects: adequacy of the author to the text (the competence of the author) and adequacy of the author to the reader (the goodwill predisposition of the author towards the reader). The author could be either a person or an institution.
Problem	**Conciseness**
Explanation	People rarely read Web pages word by word: they prefer to read on the screen few lines (15-25 lines). In this sense, conciseness is one of the most important aspects of the art of web-writing. For this reason it is very important to write an effective "short" and concise text.

Feature	**General Communication quality (text, images, …)**
Problem	**Text errors**
Explanation	The written text should not present grammatical errors.
Problem	**Multimedia consistency (images, audio, videos…)**
Explanation	All the multimedia files must be consistent with the subject of the page.

Open set: other may be added, according to the application domain and specific features.

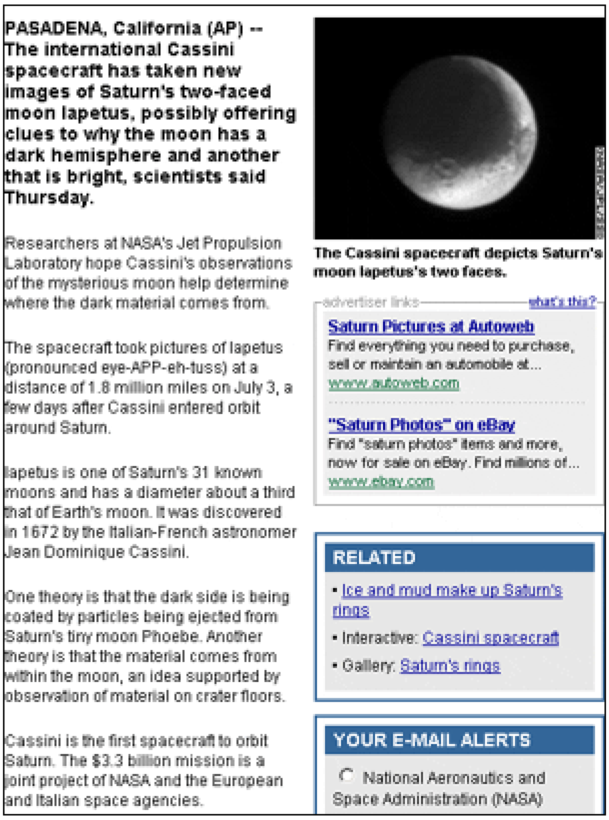
1b. CONTENT ACTIONS

How to use Content Heuristics

The purpose of this document is to explain in an extensive way how to find the usability problems for every content usability feature and to provide a step-by-step action guide for detecting the different problems.

Feature	**Text**
Problem	**Accuracy**
Action	1. Read carefully the text and verify if it: a. describes adequately the referenced world; b. is consistent in itself; c. does not contain errors of any kind.
Problem	**Currency**
Action	1. Try to understand if the text is update or not: a. find the date of the text publication; b. if the date is not reported, try to find other references that could help you to understand the period of publication.
Problem	**Coverage**
Action	1. Read carefully the text and try to answer these questions: a. is it clear what the text is speaking about? b. what it is supposed to cover?
Problem	**Content objectivity**
Action	1. Read carefully the text and verify if the commitment of the sender is clear. Try to understand what type of message you are reading: a. is it a comment? b. is it advertising ? c. is it an investigation about a topic? d. …
Problem	**Authority**
Action	1. Reading the text verify: a. the adequacy of the author (single or institution) to the text (the competence of the author); b. the adequacy of the author to the reader; c. if exists a lack of identification to the reader (responsible of its publications).
Problem	**Conciseness**
Action	1. Count the number of lines of text; 2. If the text is short enough: a. verify if you have understood the main topic presented; b. verify if exists the possibility to download the extensive version of the text (in the case of articles, presentations, … it is useful to allows the user to download the complete version in .PDF format).

Example	 www.cnn.com **Currency** ❶**:** in this example is posted the day of publication of the article. This allows the user to know if the article is updated or not. **Coverage:** only reading the introduction of the article it is clear the topic it is speaking about. **Authority** ❷**:** CNN & REUTERS are reliable sources. Even if the name of the journalist does not appear, the authority is certified. **Conciseness:** in many cases the articles published in CNN.com are to long: however, the application offers an easy to access "Story tool" that allows the user to print immediately the article or to bookmark the page in an external repository (❸). The only lack, from the user's perspective, is that it is not possible to download the article in .pdf version.

<table>
<tr><th>Feature</th><th>General Communication quality (texts, images, flash animations…)</th></tr>
<tr><td>Problem</td><td>Text errors</td></tr>
<tr><td>Action</td><td>1. Read carefully the text and verify that it not contain any grammatical error (you can also, for example, copy and paste the text in a word processor and use the auto-correction tool).</td></tr>
<tr><td>Problem</td><td>Multimedia consistency (images, audio, videos…)</td></tr>
<tr><td>Action</td><td>1. Verify if the multimedia files used for presenting a topic are integrated in a consistency way (e.g. if the text speaks about racism it should be integrated with an image(s), videos, flash animation(s) … that are related to this topic).</td></tr>
<tr><td>Example</td><td>


PASADENA, California (AP) -- The international Cassini spacecraft has taken new images of Saturn's two-faced moon Iapetus, possibly offering clues to why the moon has a dark hemisphere and another that is bright, scientists said Thursday.

Researchers at NASA's Jet Propulsion Laboratory hope Cassini's observations of the mysterious moon help determine where the dark material comes from.

The spacecraft took pictures of Iapetus (pronounced eye-APP-eh-tuss) at a distance of 1.8 million miles on July 3, a few days after Cassini entered orbit around Saturn.

Iapetus is one of Saturn's 31 known moons and has a diameter about a third that of Earth's moon. It was discovered in 1672 by the Italian-French astronomer Jean Dominique Cassini.

One theory is that the dark side is being coated by particles being ejected from Saturn's tiny moon Phoebe. Another theory is that the material comes from within the moon, an idea supported by observation of material on crater floors.

Cassini is the first spacecraft to orbit Saturn. The $3.3 billion mission is a joint project of NASA and the European and Italian space agencies.

The Cassini spacecraft depicts Saturn's moon Iapetus's two faces.

advertiser links — what's this?

Saturn Pictures at Autoweb
Find everything you need to purchase, sell or maintain an automobile at...
www.autoweb.com

"Saturn Photos" on eBay
Find "saturn photos" items and more, now for sale on eBay. Find millions of...
www.ebay.com

RELATED
• Ice and mud make up Saturn's rings
• Interactive: Cassini spacecraft
• Gallery: Saturn's rings

YOUR E-MAIL ALERTS
National Aeronautics and Space Administration (NASA)

www.cnn.com

In this example, the article does not present any error and the image used is strictly connected to the topic.
</td></tr>
</table>

Open set: other may be added, according to the application domain and specific features.

NAVIGATION

2.a NAVIGATION HEURISTICS

Within the navigational dimension of a web application there are two basic aspects that could be analyzed: on one hand the different ways that can be used by a user to reach a specific piece of information; on the other hand, the connections for passing from a content to another content.

This document presents a number of navigational features and for each feature some usability heuristics are described.

The document is divided into three parts:

1. *Basic navigational heuristics;*
2. *Advanced Navigation Heuristics I;*
3. *Advanced Navigation Heuristics II- Navigation Patterns.*

1. BASIC NAVIGATION HEURISTICS

Feature	**Navigation within a topic** (information object, entity)
Problem	**Segmentation**
Explanation	The different information about a topic could be segmented in different pages. For example, if we consider a museum website and the topic *"Author of the painting"*, this topic could be fragmented in different pages (e.g. Biography, Events of his live, More detailed info…). From a navigational point of view, it is important that the user might understand which pages belong to the topic and how the navigation within these pages works.
Problem	**Orientation clues**
Explanation	Within the navigation in a topic it is very important that the user can understand immediately his position within the topic (e.g., "You are in Biography").
Problem	**Accessibility of different pages**
Explanation	It is always essential that all the pages of a topic are easy to access in few clicks.

Feature	**Navigation within a Group of topics** (collection, set of information objects)
Problem	**Introduction list**
Explanation	The introduction list is the starting point for the navigation to a specific topic (e.g. from *paintings of 16th century* to *Venus and Adonis*), therefore it should be clear the strategy used for organizing the list. This strategy could affect the navigation of the user (e.g. if the introduction list is composed of 50 elements organized casually, the user could have some problems for identifying the elements in which he is interested).
Problem	**Orientation clues**
Explanation	It is always important that the user can understand which group of topic s/he is browsing.
Problem	**Accessibility of topics**
Explanation	It should be clear how to get an overview of all topics of the group (how many? If not, which?) and easily reach them.

Feature	**Navigation within a transition (Navigation between topics)**
Problem	**Transition list**
Explanation	The transition list allows the user to navigate across relevant relation between topics that are semantically connected (e.g. from a specific cloth to a particular accessories, the user has to go through a list of accessories); therefore it should be clear the strategy used for organizing the list. This strategy could affect the navigation of the user (e.g. if the transition list is composed of 20 elements - e.g. 20 accessories - randomly organized, the user could have some problems for identifying the elements in which he is interested).
Problem	**Orientation clues**
Explanation	It is always important that the user might understand that s/he is browsing through a transition/relation between two different topics.
Problem	**Accessibility of target**
Explanation	When browsing from a topic to another topic semantically connected, it is basic that the user accesses easily to the target topic.

Feature	**Overall Navigation**
Problem	**Landmarks**
Explanation	The access to the main sections of a web site is given by a number of landmarks. Using the landmarks the user can access easily and quickly all the macro-sections of the application. Therefore, the landmarks should be well highlighted in every page.
Problem	**Consistency**
Explanation	All the web applications have a general navigation architecture that supports the navigation of the user. This navigation has to be consistent among the different parts of the application. In this sense, it is very important that this "general" architecture emerges in a satisfactory way: the user has to comprehend how the general navigation works.
Problem	**Accessibility**
Explanation	Accessibility refers to ensuring that content is accessible, ie. ensuring that content can be navigated and read by everyone, regardless of location, experience, or the type of computer technology used.

Feature	**Tree Navigation**
Problem	**Orientation**
Explanation	Different websites are designed with a tree structure. In this site, the orientation of the user become fundamental both when the user explores a branch (section) of the tree and when he passes from a branch (section) to another. The user should be aware when a change of context happens.
Problem	**Backward navigation**
Explanation	When the user navigates within a tree (in particular when he passes from a section to another) one of the most difficult things to manage is related to the navigation to the previous visited pages. The application should support this action without the use of *back functionality* offered by the browser.
Problem	**Depth anticipation**
Explanation	Often the "tree architecture" of websites is very complex. For this reason, the user could have some problems to have a synoptic both of the website and of each branch.

Open set: other may be added, according to the application domain and specific features.

2. ADVANCED NAVIGATION HEURISTICS I

Feature	**Navigation within a Kind of Topic (Multiple topic)** (information object, entity type)
Problem	**Consistency**
Explanation	The kind of topic (or "multiple topic") is a generic category of topics of interest for the user. The kinds of topics identify the core content of the application. Therefore, all the topics belonging to a kind of topic (e.g. kind of topic *"painting"* → topic: La Gioconda, the Creation of Adam, The return of the prodigal son …) should have the same structure (the same pages, the same navigational strategy…): each topic should be recognizable as an exemplar of a kind.
Problem	**Segmentation**
Explanation	The different pieces of information about a kind topic (and related topics) could be segmented in different pages. For example, if we consider a museum website and the topic *"Author of the painting"*, this topic could be fragmented in different pages (e.g. Biography, Events of his live, More detailed info…). From a navigational point of view, it is important that the user would understand which pages belonging to the topic and how the navigation within these pages works.
Problem	**Orientation clues**
Explanation	Within navigation in a topic it is very important that the user can understand immediately his position within the topic (e.g., "You are in Biography").
Problem	**Accessibility of different pages**
Explanation	It is always essential that all the pages of a topic are easy to access in few clicks.

Feature	**Navigation within a Group of groups of topics** (collection, set of information objects)
Problem	**Introduction list**
Explanation	The introduction list of a group of groups of topics is the starting point for the navigation to a group of topics (e.g. from *paintings by historical period* to *paintings of 16th*), therefore it should be clear the strategy used for organizing the list. This strategy could affect the navigation of the user (e.g. if the introduction list is composed of 10 elements randomly organized, the user could have some problems for identifying the elements in which he is interested).
Problem	**Orientation clues**
Explanation	It is always important that the user would understand which group of group of topics he is browsing.
Problem	**Accessibility of group of topics**
Explanation	The navigation from the introduction list to the different groups of topics should be efficient and, therefore, each group of topics should be reached in few clicks.

Feature	**Backward navigation** (Reference to the past pages or actions)
Problem	**"Go back"** (Note: do not use the back button provided by the browser because the browser is an external application and so its use could not aligned with website behaviour)
Explanation	Some applications offer "go back" functionality allowing the user to go to the previously visited pages. The effect of this "go back" should be take me to the page I just visited before the current one. Be aware that if I reach a page from two different paths the go back should take me to the actual page I come from.
Problem	**History** (Note: do not use the back button functionality provided by the browser because the browser is an external application and so its use could not aligned with website behaviour)
Explanation	The history mechanism allows the user to verify which the visited pages are. The History should support the backtracking of past actions or pages.

Open set: other may be added, according to the application domain and specific features.

3. ADVANCED NAVIGATION HEURISTICS II – NAVIGATION PATTERNS

Feature	**Guided-tour navigation**
Problem	**Orientation clues**
Explanation	The guided-tour provides to the user an "easy-to-use" access to a small group of objects, assuming that user has no reason (or is not able) to select one of them. Considering that the guided-tour consists of a sequence of links among different objects (e.g. topics, pages…) the orientation becomes fundamental for the success of the user navigation (e.g. "you are browsing the photo 10 of 20").
Problem	**Control**
Explanation	The user has to control the navigation through a guided-tour: he should be able to stop, restart reset… the navigation.
Problem	**Navigation strategy**
Explanation	The guided-tour is one of the possible navigation strategies; therefore it is very important to think very well to the goal of the navigation before implementing a guided-tour. Normally, the guided-tour is used for didactical purposes (e.g. a guided-tour of the 20 most important paintings of 16^{th} century) or for promotional reasons (e.g. a tour for presenting the new features of a product).
Problem	**Topology**
Explanation	The order of the elements in a guided-tour is crucial for the success of this navigation strategy.

Feature	**Index navigation**
Problem	**Orientation clues**
Explanation	The index-navigation provides a fast access to a group of objects, for users who are interested to one or more of them, and are able to make a choice. For this reason the user should understand immediately that the object in which he is interested belongs to a specific group of objects.
Problem	**Control**
Explanation	The user has to control the navigation both from the starting index to each element of the index and to go back from one element to the index.
Problem	**Navigation strategy**
Explanation	The index navigation is one of the possible navigation strategies; therefore it is very important to think very well to the goal of the navigation before implementing index navigation (e.g. a photo gallery could be implemented with an index navigation).
Problem	**Topology**
Explanation	The order of the elements in an index navigation is crucial for the success of this navigation strategy.

Feature	**All-to-all navigation**
Problem	**Orientation clues**
Explanation	The all-to-all navigation allows the user to navigate from one page to each other.
Problem	**Control**
Explanation	The user should have the possibility to select every pages linked with the all-to-all navigation.
Problem	**Navigation strategy**
Explanation	The index navigation is one of the possible navigation strategies; therefore it is very important to think very well to the goal of the navigation before implementing an al-to-all navigation.
Problem	**Topology**
Explanation	The order of the elements in an all-to-all navigation is crucial for the success of this navigation strategy.

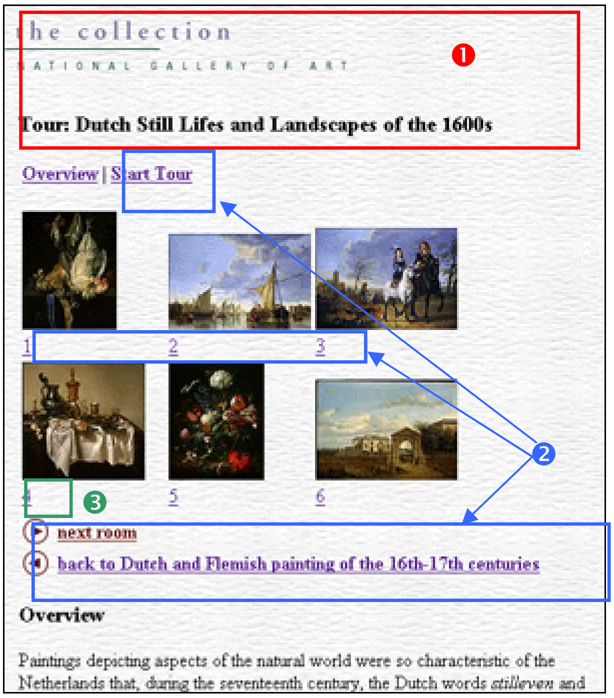
Open set: other may be added, according to the application domain and specific features.

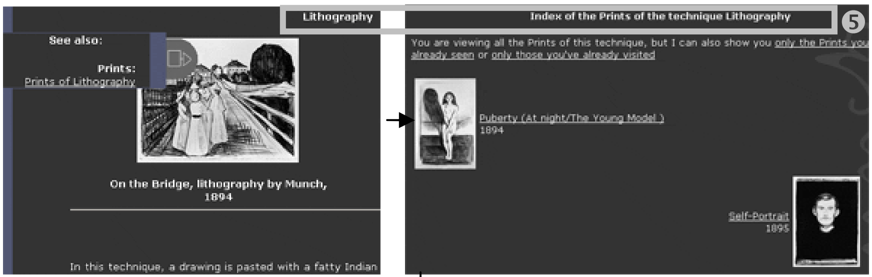
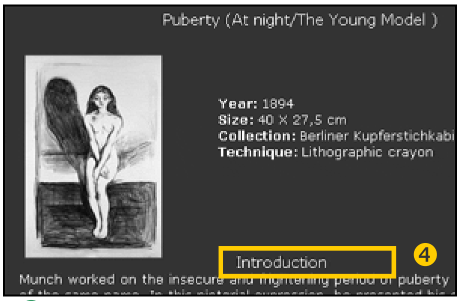
2.b NAVIGATION ACTIONS

How to use Navigation Heuristics

The purpose of this document is to explain in an extensive way how to find the usability problems for every navigational usability feature and to provide a step-by-step action guide for detecting the different problems.

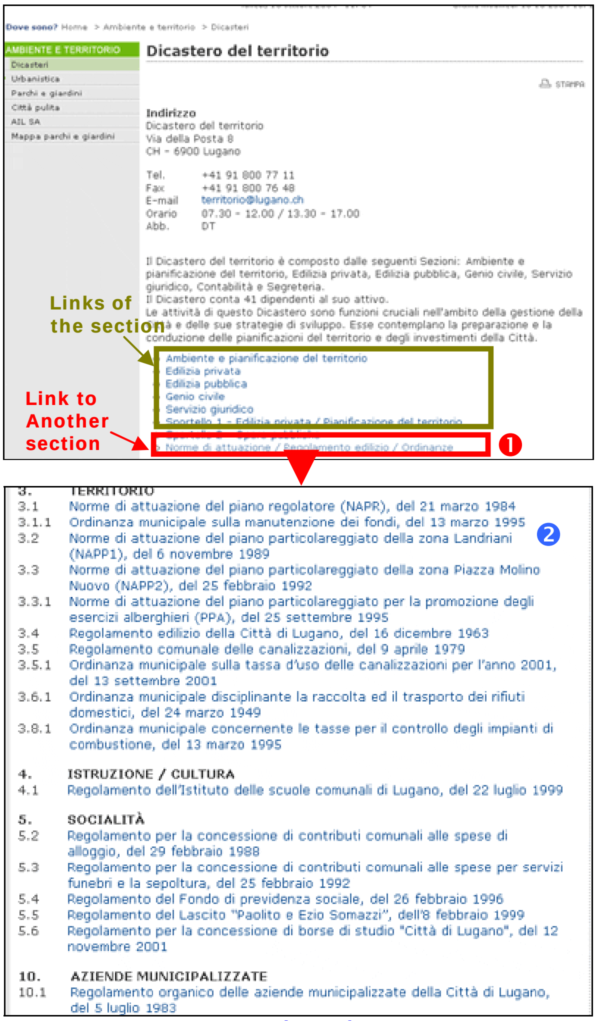
Feature	**Navigation within a Topic** (information object, entity)
Problem	**Segmentation**
Action	1. Identify an instance of topic of the website (e.g. the BMW 3 Series Coupé); 2. Try to understand in how many pages the topic is segmented (e.g. the BMW 3 Series is segmented in 5 pages/sections).
Problem	**Orientation clues**
Action	1. Identify an instance of topic of the website (e.g. the BMW 3 Series Coupé). 2. *Path visibility(where I can go?):* Navigate from the home page to the instance of the selected topic and try to understand the path engaged (where I was?) and/or the paths that it is possible to engage; 3. *Status visibility (where I am?):* Navigate randomly within the topic and try to understand which page is browsing -; 4. *Context visibility:* Navigate randomly within the topic and try to understand the context you are browsing (e.g. the design features of a the BMW 3 Series) .
Problem	**Accessibility of different pages**
Action	1. Identify an instance of topic of the website (e.g. the BMW 3 Series Coupé); 2. Try to navigate from one page to the others and count the number of clicks for accessing these pages.
Example: Instance of topic BMW 3 Series.	 www.bmw.com **Segmentation ❶:** In the case of the BMW 3 Series is clear that the topic is divided into 5 main pages/sections: "Introduction", "Highlights", "Multimedia gallery", "Models and data sheets", "Catalogue". In some cases, these pages/sections are also divided in sub-pages or sections (e.g. Highlights is divided into "Design", "Engines", "Chassis", "Safety"). **Orientation clues ❷:** It is clear both the path engaged (from Home page > Products > Highlights > Design) and the page we are browsing (Design). **Accessibility of different pages:** Once reached the BMW 3 Series all the pages of the topic are very accessible in few clicks.

Feature	**Navigation within a Group of topics** (collection, set of information objects)
Problem	**Introduction list**
Action	1. Navigate from the home page to a group of topics (e.g. from the homepage of the NGA web site– National Gallery of Art – to "Dutch Still Lifes and Landscapes of the 1600s"); 2. Verify if it is understandable the reason why these instances of topics are presented within the group of topics; 3. Try to understand the reason why of the order of the topics instances (e.g. the paintings are organized from most to the less important).
Problem	**Orientation clues**
Action	1. Navigate from the home page to the group of topics and to the group of topics to the instance(s) of topic; 2. Try to understand the path engaged (where I was?) and/or the paths that it is possible to engage (where I can go?) – *Path visibility*; 3. Try to understand which page is browsing - *Status visibility* (where I am?); 4. Try to understand the context you are browsing (e.g. the *paintings of 16th century*).
Problem	**Accessibility of topics**
Action	1. From the home page navigate to an introduction list (e.g. *paintings of 16th century*) and try to access to some instances of topics; 2. During this navigation count the number of clicks necessary to reach the instances of topics.
Example: Group of topics "Dutch Still Lifes and Landscapes of the 1600s"	 www.nga.gov **Introduction list:** Reading the text (Overview) it is possible to understand that all the paintings presented in this group of topics are related to the Dutch Lifes. Besides the web site offers a widening on this group of topics. At the contrary it is not explicit the reason why of the order of these paintings (e.g. it should be clear if they are organized from the most to the less important painting). **Orientation clues:** Navigating from the home page to the group of topic is always present the reference to the collection and to the group of topics (collection) we are browsing (1). In addition different orientation clues allows to understand the path that it is possible to engage (2) and the context we are browsing. **Accessibility:** The accessibility to each topic is not perfectly implemented. In fact, if we clicks on the image (3) we go directly to the painting selected, but clicking the number under the images we go to the "Captions" at the end of the page. There is no consistency in accessibility of the topics.

Feature	**Navigation within a Transition (Navigation between topics)**
Problem	**Transition list**
Action	1. Navigate from an instance of kind of topic to another instance of another kind of topic (e.g. within the MunchundBerlin web site – from the "lithography technique" to the painting called "Puberty"); 2. Within the transition list verify if it is understandable the reason why these instances of topics are presented; 3. Within the transition list try to understand the reason why of the order of the topics instances (e.g. the paintings are organized from most to the less important). 4. Clicking on an instance of the transition list, verify if the target is correctly reached.
Problem	**Orientation clues**
Action	1. Navigate from an instance of kind of topic to another instance of another kind of topic (e.g. within the MunchundBerlin web site – from a technique to a painting created with this technique); *2.* Try to understand which page is browsing (where I am?) - *Status visibility*. a. Verify if you understand when you reach the transition list; b. Verify if you understand when you reach the target of the relation. 3. Try to understand the context you are browsing - *Context visibility:* a. Verify if you understand the context of the transition (e.g. "Index of the Prints of the technique Lithography"); b. Verify if you understand the context of the target of the relation (which is the topic reached).
Problem	**Accessibility of target**
Action	1. Starting from an instance of kind of topic (e.g. the "lithography technique") counts the number of clicks necessary for reaching another instance of another kind of topic that are semantically connected to the source (e.g. the painting called "Puberty").
Example: From "lithography technique" to "Puberty painting"	 1 *Lithografy*: Instance of the kind of topic "Technique" 2 *Transition list* : Index of the Prints of the technique Lithography"  3 *Puberty*: Instance of the kind of topic called "Print" www.munchundberlin.org **Transition list:** within the transition list all the prints realized with the lithography technique are organized from the less to the most recent. **Orientation clues:** navigating through this relation is always clear both the page we are browsing (the status is given by the title of the section 4) and the context (that is given by the main title of the page 5). **Accessibility of target:** from the topic Lithography it is possible to reach the topic target with only two clicks.

Feature	**Overall Navigation**
Problem	**Landmarks**
Action	1. Identify the main landmarks of the website; 2. Using the landmarks try to navigate from one section to the others: once you reach a new section verify if the landmarks are always present; 3. Localize the "service" landmarks (e.g. "privacy policy") 4. Using the "service" landmarks try to: a. navigate from one "service section" to the other "service sections"; b. navigate from one "service section" to one of the main sections (verify if the main landmarks are always present).
Problem	**Consistency**
Action	1. Navigate randomly or taking into account a series of tasks/scenarios (you have to create one or more scenarios); 2. Try to sketch in a formal or semiformal way the main navigation architecture of the website. 3. Navigate once more in the website and verify that the navigational architectural schema is implemented in a consistent way.
Problem	**Accessibility**
Action	1. Create one or more scenarios (define task(s)/goal(s) –e.g. Find information about the new book of John Grisham); 2. Try to achieve the goal(s) of the scenario(s); 3. Count the clicks necessary for achieving the goal(s).
Example: Overall Navigation of Amazon (Books sections)	 www.amazon.com **Landmarks:** the Books' section of AMAZON has a number of landmarks (❶) always present when the user browses this main section; but among them two are not sections of Books ("Magazine", "Corporate Accounts"). In fact, if we are in the sub section "Bestsellers" and we click on "Magazine" we reach another section from which we can not come back directly to "Bestsellers". Besides, AMAZON presents a number of "high-level" landmarks useful for accessing the main functionality of the website (e.g. "View cart", "Wish list"…). AMAZON proposes also different "service landmarks" in every section of the web site (❸). **Accessibility:** in general, using different navigational paths, the content (e.g. the customer reviews of a book) are easy to access in few clicks.

<table>
<tr><th>Feature</th><th>Tree Navigation</th></tr>
<tr><td>Problem</td><td>Orientation</td></tr>
<tr><td>Action</td><td>1. Navigate within the website and try to verify:
a. if the internal links leading to other sections are clear;
b. if, passing from one section to another, orientation clues are given;
c. if in each page orientation clues are given (even if you stay always in the same section).</td></tr>
<tr><td>Problem</td><td>Backward navigation
(Note: do not use the "go back" functionality provided by the browser)</td></tr>
<tr><td>Action</td><td>1. When you reach a page try to navigate back (if this functionality is available) and verify if you reach exactly the previous page.</td></tr>
<tr><td>Problem</td><td>Depth anticipation</td></tr>
<tr><td>Action</td><td>1. Navigate randomly or take into account a scenario in the website and try to answer these questions:
a. Did you know how many branches (sections) has the website?
b. Did you know how much is deep every branch?</td></tr>
<tr><td>Example</td><td>



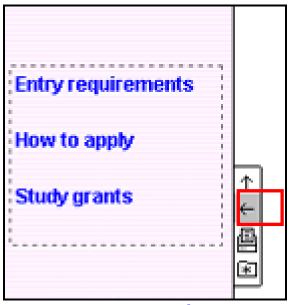
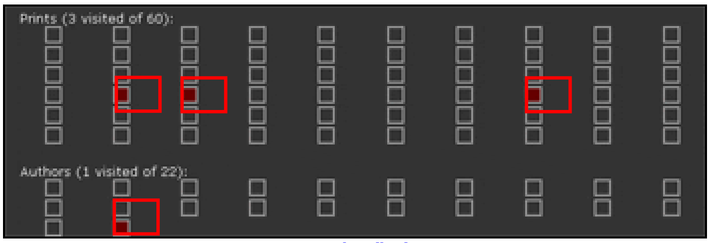
www.lugano.ch

Navigate through the city portal of Lugano sometimes it happen that there are links (❶) that allow to navigate from one section to another. The problem is that the user does not understand the difference between links leading to pages of the main section and links to other sections. Besides, when the user selects one of these links between sections he reaches pages without orientation clues (❷).
</td></tr>
</table>

ADVANCED NAVIGATION HEURISTICS I

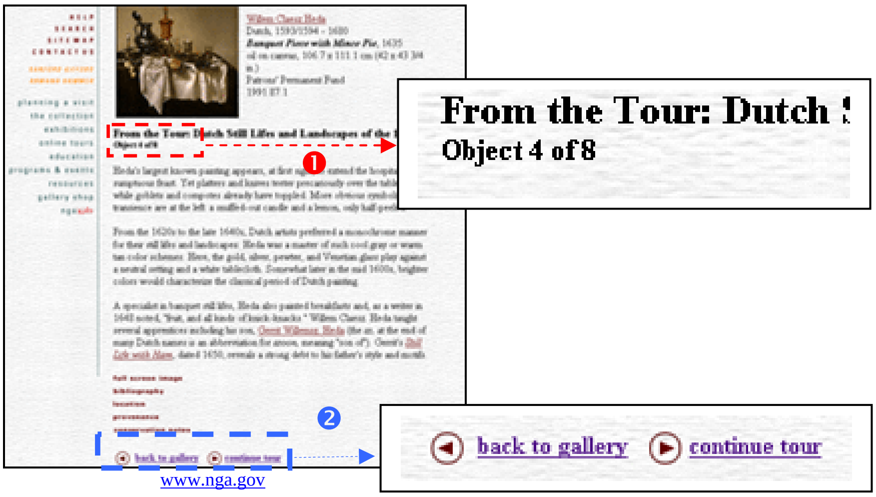
Feature	**Kind of Topic (Multiple topic) Navigation** (information object, entity type)
Problem	**Consistency**
Action	1. Select a number of topics (instances) (4 or 5); 2. Identify a general navigation strategy and the high-level structure of the kind of topic: a. Sketch (in an informal way) the navigation structure of the first topic (e.g. BMW Series 3 Coupé) – In how many pages/nodes the topics is divided; How the navigation works? b. Take the others topics (e.g. Series 5 Touring, Series 3 Sedan, Series 6 Convertible …) and verify if it exist a navigation consistency among the topics.
Problem	**Segmentation**
Action	1. Identify an instance of topic of the website (e.g. the BMW 3 Series Coupé); 2. Try to understand in how many pages the topic is segmented (e.g. the BMW 3 Series is segmented in 5 pages/sections). 3. Verify if al the topics have this segmentation (see Consistency – Action 2). Note: if you has already verify the Consistency you can use the results obtained (e.g. the sketch of the navigation structure) for verify the segmentation.
Problem	**Orientation clues**
Action	1. Identify an instance of topic of the website (e.g. the BMW 3 Series Coupé); 2. Navigate from the home page to the instance of the selected topic and try to understand the path engaged (where I was?) and/or the paths that it is possible to engage (where I can go?) - *Path visibility*; 3. Navigate randomly within the topic and try to understand which page is browsing - *Status visibility* (where I am?); 4. Navigate randomly within the topic try to understand the context you are browsing (e.g. the design features of a the BMW 3 Series) - *Context visibility*.
Problem	**Accessibility of different pages**
Action	1. Identify an instance of topic of the website (e.g. the BMW 3 Series Coupé): for identify a topic you can create a scenario (define goals, tasks); 2. Try to navigate from one page to the others and count the number of clicks for accessing these pages.
Example: Kind of topic "BMW Auto Model"	 www.bmw.com **Consistency & Segmentation ❶:** In the case of the BMW 3 Series is clear that the topic is divided into 5 main pages/sections: "Introduction", "Highlights", "Multimedia gallery", "Models and data sheets", "Catalogue". In some cases, these pages/sections are also divided in sub-pages or sections (e.g. Highlights is divided into "Design", "Engines", "Chassis", "Safety"). This structure (with little differences) is used among all the different BMW models. The navigation is consistency among all the BMW Models: from a section (e.g. Highlights) it is possible to navigate both to the other sections (e.g. "Introduction", "Highlights" …) and to subsections. **Orientation clues ❷:** It is clear both the path engaged (from Home page > Products > Highlights > Design) and the page we are browsing (Design). **Accessibility of different pages:** All the pages of the topic BMW 3 Series are very easy to access with few clicks.

Feature	**Group of groups of topics Navigation** (collection, set of information objects)
Problem	**Introduction list**
Action	1. Navigate from the home page to a group of groups of topics (e.g. within the NGA web site– National Gallery of Art – "Dutch and Flemish 16th-17th centuries"); 2. Verify if it is understandable the reason why of the order of the group of topics presented in the list.
Problem	**Orientation clues**
Action	1. Navigate from the home page to the group of groups of topics ; 2. Try to understand the path engaged (where I was?) and/or the paths that it is possible to engage (where I can go?) – *Path visibility*; 3. Try to understand which page is browsing - *Status visibility* (where I am?); 4. Try to understand the context you are browsing (e.g. *the collections of paintings of "Dutch and Flemish 16th-17th centuries"*).
Problem	**Accessibility of group of groups of topics**
Action	1. Navigate from the home page to a group of groups of topics (e.g. from the homepage of the NGA web site to "Dutch and Flemish 16th-17th centuries" collections); 2. During this navigation count the number of clicks necessary to reach the group of groups of topics topics.
Example: Group of group of topics"	 www.nga.gov/collection/index.shtm **Introduction list:** the group of topics are organized in alphabetical order. **Orientation clues:** the orientation is given by the title of the page and of the paragraph (❶). Note: the title "Paintings" (❶) is not positioned in a visible part of the page: this could be a semiotic and graphic problem that affects also the navigation. **Accessibility:** from the homepage are necessary only two clicks to reach the group of groups of topic "Dutch and Flemish 16th-17th centuries".

Feature	**Backward navigation** (Reference to the past pages or actions)
Problem	**"Go back"** (Note: do not use the "go back" functionality provided by the browser)
Action	1. Navigate randomly or taking into account a series of tasks/scenarios (you have to create one or more scenarios); 2. When you find "Go back", "Previous page"… click it and verify if you really reach the previous page.
Problem	**History** (Note: do not use the "History" functionality provided by the browser)
Action	Verify if exist a history mechanism. If yes: 1. Visit randomly a number of topics and write the topics visited; 2. Using the history mechanism, verify if all the visited topic are reported.
Example	 www.unisi.ch **"Go back":** within the website of the University of Lugano is always present an icon for going back. Trying to use this function several times we have verify that it works well.  www.munchundberlin.org **History:** within the website munchundberlin.org it has been implemented a visual mechanism for tracing the visited topic. During a session we have visited 4 topics (3 prints and 1 author) and the system has correctly traced our session.

Open set: other may be added, according to the application domain and specific features.

ADVANCED NAVIGATION HEURISTICS II – NAVIGATION PATTERNS

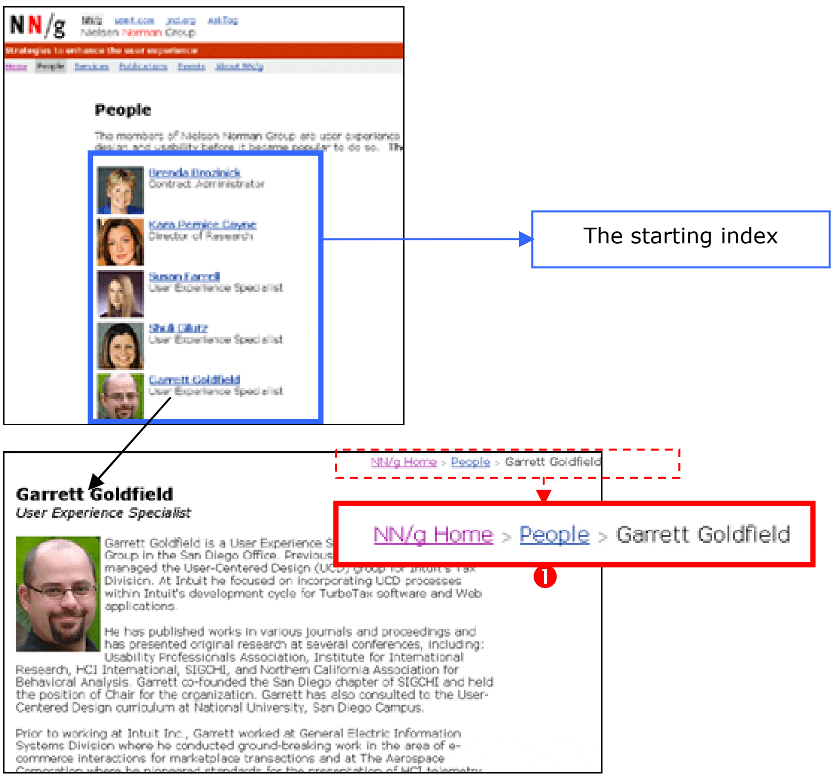
<table>
<tr><th>Feature</th><th>Guided-tour navigation</th></tr>
<tr><td>Problem</td><td>Orientation clues</td></tr>
<tr><td>Action</td><td>1. Identify a guided tour within the website;
2. Navigate within the guided-tour trying to understand:
a. which page is browsing – Status visibility (where I am?);
b. the path engaged (where I was?) the paths that it is possible to engage (where I can go?) - Path visibility;
c. the context you are browsing – Context visibility.</td></tr>
<tr><td>Problem</td><td>Control</td></tr>
<tr><td>Action</td><td>1. Navigating through the elements/members of a guided-tour try to:
a. go "previous" (respectively "next");
b. try to restart the guided tour (respectively “stop” the tour).</td></tr>
<tr><td>Problem</td><td>Navigation strategy</td></tr>
<tr><td>Action</td><td>Evaluate if the pattern guided-tour is suitable for a satisfactory fruition of the content. For doing this:
1. count the pages of the guided-tour,
2. analyse the content of each page.
and answer these questions:
1. do you remember the first page of the guided-tour?
2. do you have a global vision of the guided-tour?;
3. …
Note: the guided-tour is suitable for didactical and promotional purposes.</td></tr>
<tr><td>Problem</td><td>Topology</td></tr>
<tr><td>Action</td><td>1. Start the navigation of a guided-tour and verify if it is understandable the reason why and the order of the elements of the guided tour.</td></tr>
<tr><td>Example</td><td>

Orientation clues ❶: the orientation within this guided-tour is given by title of the tour we are exploring (From the Tour: Dutch Landscapes and Seascapes of the 1600s): this information provides the context. Another orientation clue is the information about the position of the element within the guided-tour (“Object 4 of 8).

Control ❷: the control of the navigation is given only by the possibility to continue the tour or to go back to the gallery (the starting point of the tour). Within this guided-tour is not possible to go “previous”.

Navigation strategy: this guided-tour is composed by 8 objects and the content of each object gives a complete idea about the object itself. Once finished the navigation within the guided-tour it easy to remember the visited objects and their contents. In this case, the strategy of implementing a guided-tour achieves the didactical goal of this part of the application.

Topology: it is not clear the order of the object of the guided-tour.</td></tr>
</table>

Feature	**Index navigation**
Problem	**Orientation clues**
Action	1. from an index (list), select an element of the list; 2. once reached the element of the list verify if it is understandable that this element belong to the starting index (list) - *Context visibility*.
Problem	**Control**
Action	1. starting from an index (list) verify if it is possible to go to each element belonging to the index; 2. verify if from each element reached it is possible to go back to the index.
Problem	**Navigation strategy**
Action	1. evaluate if the pattern index is suitable for a satisfactory fruition of the group of objects in term of similarity of elements (e.g. photos gallery, video gallery, list of people…). 2. counting the number of the elements belonging to the list, verify if the cardinality of the list elements is suitable for the index navigation pattern (if the number is too high – e.g. over 10-15 elements – this strategy is not appropriated).
Problem	**Topology**
Action	1. starting the navigation from an index, verify if it is understandable the reason why and the order of the members belonging to the list.
Example	 www.nngroup.com **Orientation clues:** the orientation is given by the “Status bar” (❶) provided by the application. Using this bar as a clue it should be clear that the selected element belonging to the starting list. **Control:** the “Status bar” (❶) also allows the user to go back (clicking “People”) to the starting list. **Navigation strategy & Topology:** the number of the elements of the list is not so high, so in this case the index navigation is suitable for presenting the members of the staff.

<table>
<tr><th>Feature</th><th>All-to-all navigation</th></tr>
<tr><td>Problem</td><td>Orientation clues</td></tr>
<tr><td>Action</td><td>
1. Identify a topic of the website (e.g. the BMW 3 Series Coupé).
2. Navigate within the topic and for each page reached try to understand:
a. the path engaged (where I was?);
b. the paths that it is possible to engage (where I can go?) - Path visibility;
c. page is browsing - Status visibility (where I am?);
d. which is the context we are browsing - Context visibility.
</td></tr>
<tr><td>Problem</td><td>Control</td></tr>
<tr><td>Action</td><td>1. Within the topic navigation verify if it is possible to navigate from one page to the others.</td></tr>
<tr><td>Problem</td><td>Navigation strategy</td></tr>
<tr><td>Action</td><td>1. Counting the number of pages (nodes) of the topic, verify if the cardinality (the number of the pages) is suitable for the all-to-all navigation pattern.</td></tr>
<tr><td>Problem</td><td>Topology</td></tr>
<tr><td>Action</td><td>1. Verify if it is understandable the reason why and the order of the pages.</td></tr>
<tr><td>Example</td><td>

www.bmw.com

Orientation clues 1: once we navigate within the (instance of) topic BMW 3 Series Coupé are always highlighted both the page we are browsing and the context.

Control: using the contextual menu (1) it is always very easy to navigate from one page to the others.

Navigation strategy: see that the number of the pages is not so high the all-to-all navigation allows the user to reach every page with one click.

Topology: in this case BMW used a very common order of the pages for presenting a product. In fact, they start with a (general) “Introduction” and for going in depth with the presentation they present in succession “Highlights” (with some sub-pages), “Multimedia gallery” … In conclusion, this order is appropriate for an easy and efficient navigation.
</td></tr>
</table>

Open set: other may be added, according to the application domain and specific features.

INTERFACE DESIGN

3.a SEMIOTICS HEURISTICS

During the interaction with a website the user should easily understand the meanings of the messages proposed. In particular, three main semiotic features should be considered:

- *String of characters:* the term(s) used for describing the meaning of a link creates expectations in the user and is the promise that if the user clicks on the link s/he will reach the content s/he is looking for; the terms used for synthesising the content through a title, a heading or a keyword should be clear and representative of the referred content.
- *Interaction images:* the meaning of any non-textual sign or symbol used for navigation purposes or for activating particular operations/services should be clear and intuitive.
- *Macro-areas:* the meaning of a single message often depends on the relation the message has with other messages on the same page: the way they are organised and grouped should help the user in understanding their meaning and the meaning of the whole page.

Feature	**String of characters (labels, titles, headings, etc.)**
Problem	**Ambiguity / Clarity**
Explanation	The term(s) used could be interpreted with different meanings by the user, making her/him confused. The main types of string of characters are: • *Link labels:* they should allow clear navigational choices. • *Headings (captions, subtitles…):* they should synthetize the referred content in an intuitive and familiar way; • *Titles:* they should introduce efficiently the topic of the page; • *Slogans:* they should synthetize the referred content in an intuitive and familiar way; • *Keywords:* it should be clear which the keywords of the content are. • …
Problem	**Labels Overlapping**
Explanation	On the same context there could be different terms/labels having a similar meaning. This could cause indecision in the user to choose the right link or to focus on a particular content.
Problem	**Generality vs. specificity**
Explanation	The term/s used could be either too generic (represent everything and nothing) or too specific, not synthesising exactly the referred content.
Problem	**Information Scent**
Explanation	Beyond the textual string, the user could have some additional content making him/her more conscious in his/her navigational choice. As an example, in an index the label of the link for an item could not be enough for letting the user to understand the meaning of the link: a thumbnail, a short text, a sound could help him in understand better what the textual string stands for.

Feature	**Interaction Images**
Problem	**Conventionality**
Explanation	Symbols and icons used for communicating a particular meaning and having an interaction purpose should be familiar to the user.
Problem	**Intuitiveness**
Explanation	If signs and icons do not follow standards and conventions, their meaning and function should be intuitive and easy.

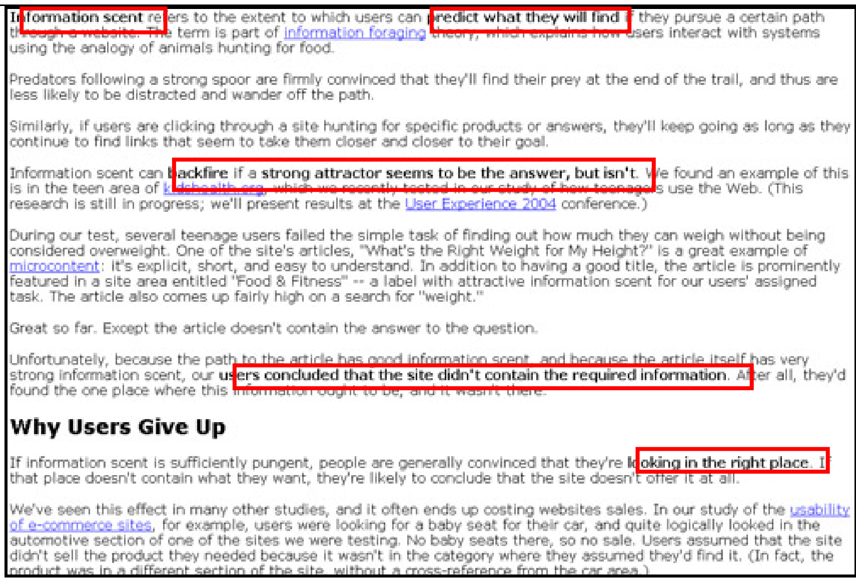
Feature	**Macro-areas**
Problem	**Grouping adequacy**
Explanation	The messages composing a single page can be grouped in macro-areas, that is, in groups of messages having a similar meaning, a content relation or satisfying a common goal/functionality.
Problem	**Position of importance**
Explanation	Each page has a main communicative goal and a main topic to present. The main meaning should be easily recognisable and should be properly grouped with respect to their importance.

3.b SEMIOTICS ACTIONS

How to use Semiotics Heuristics

The purpose of this document is to explain in an extensive way how to find usability problems for every semiotic usability feature and to provide a step-by-step action guide for detecting the different problems.

<table>
<tr><th>Feature</th><th>String of characters (labels, titles, headings, etc.)</th></tr>
<tr><td>Problem</td><td>Ambiguity / Clarity</td></tr>
<tr><td>Action</td><td>A) Actions for testing the links labels
Actions 1: without end- users
a. Identify all the links labels of the page (both labels for the main navigation and those for contextual navigation);
b. Try to anticipate the target of the page (e.g. the label "Shopping bag" means that if we click it, we should reach the shopping bag).
Actions 2: using end-users
a. Ask to a sample of end users the target of the links presented within the page.
and/or
b. Write on a sheet of paper all the links labels of a page and ask to end-users the meaning of each label.
Note: it is possible to combine Actions 1) with Actions 2).
B) Actions for testing Headings (captions, subtitles…)
1. Identify all the Headings of the page;
2. Read the Headings and try to understand their meaning (without read the content);
3. For each Heading read the referred content and verify that they are consistent.
C) Actions for testing Titles
1. Reading the title(s) of the page try to understand the main topic(s) of the page.
2. Read the referred content and verify that it is consistent with its title.
D) Actions for testing Slogans
1. Reading the slogan(s) of the page try to understand the referred content.
E) Actions for testing Keywords
1. Try to identify all the keywords of the page (content);
2. Once identified and isolated the keywords verify if they really summarized the content in an efficient way (Only reading the keywords do you grasp the topic of the page?).</td></tr>
<tr><td>Example</td><td></td></tr>
</table>



Information scent refers to the extent to which users can predict what they will find if they pursue a certain path through a website. The term is part of information foraging theory, which explains how users interact with systems using the analogy of animals hunting for food.

Predators following a strong spoor are firmly convinced that they'll find their prey at the end of the trail, and thus are less likely to be distracted and wander off the path.

Similarly, if users are clicking through a site hunting for specific products or answers, they'll keep going as long as they continue to find links that seem to take them closer and closer to their goal.

Information scent can backfire if a strong attractor seems to be the answer, but isn't. We found an example of this is in the teen area of Kidshealth.org, which we recently tested in our study of how teenagers use the Web. (This research is still in progress; we'll present results at the User Experience 2004 conference.)

During our test, several teenage users failed the simple task of finding out how much they can weigh without being considered overweight. One of the site's articles, "What's the Right Weight for My Height?" is a great example of microcontent: it's explicit, short, and easy to understand. In addition to having a good title, the article is prominently featured in a site area entitled "Food & Fitness" -- a label with attractive information scent for our users' assigned task. The article also comes up fairly high on a search for "weight."

Great so far. Except the article doesn't contain the answer to the question.

Unfortunately, because the path to the article has good information scent, and because the article itself has very strong information scent, our users concluded that the site didn't contain the required information. After all, they'd found the one place where this information ought to be, and it wasn't there.

Why Users Give Up

If information scent is sufficiently pungent, people are generally convinced that they're looking in the right place. If that place doesn't contain what they want, they're likely to conclude that the site doesn't offer it at all.

We've seen this effect in many other studies, and it often ends up costing websites sales. In our study of the usability of e-commerce sites, for example, users were looking for a baby seat for their car, and quite logically looked in the automotive section of one of the sites we were testing. No baby seats there, so no sale. Users assumed that the site didn't sell the product they needed because it wasn't in the category where they assumed they'd find it. (In fact, the product was in a different section of the site, without a cross-reference from the car area.)

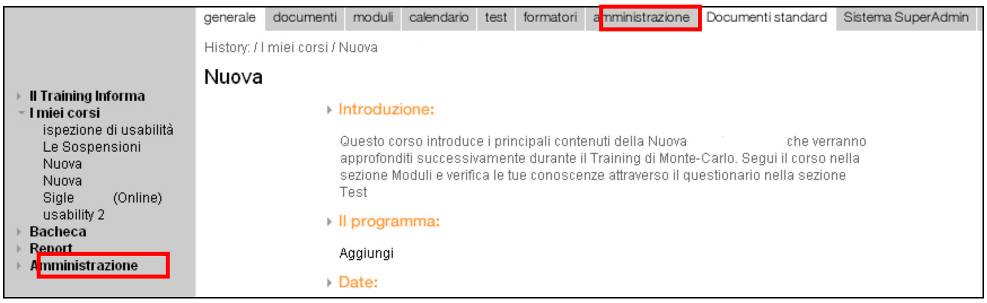
This page of Useit website (www.useit.com/alertbox/20040802.html) presents different keywords that summarize in an efficient way the content's topic. Also the titles (article's and paragraph's titles) help the user to quickly understand both the main topic of the article and the content of each paragraph.

Problem — **Labels Overlapping**

Action

1. Identify all the labels of the page;
2. Verify if there are labels whose respective meanings overlap so significantly to obstacle the decision for the selection.

Example



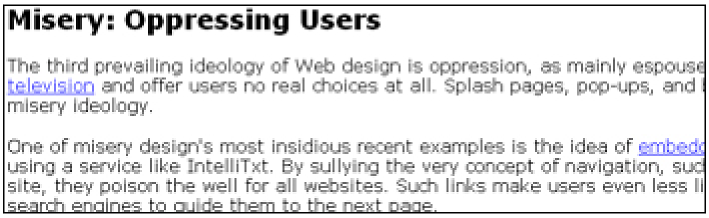
In this e-learning application there are two different sections called "administration" (for privacy reasons it is not possible to communicate the web site address and the company's name) For this motive, it is very difficult to understand immediately the difference between them. The user (especially if not yet familiar with the application) may feel confused seeing two sections with the same name and different functionalities.

Problem — **Generality vs. specificity**

Action

1. Identify and isolate all the types of string characters (labels, titles…);
2. For each category, verify if the string of characters is enough specific with respect to the content it refers to (e.g. a caption should explain specifically the referred image).

Example



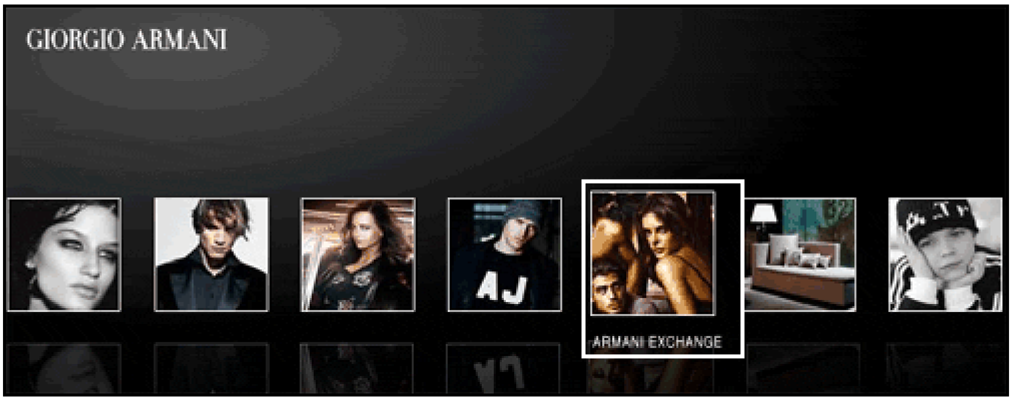
Misery: Oppressing Users

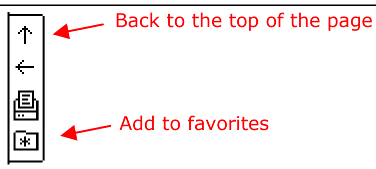
The third prevailing ideology of Web design is oppression, as mainly espouse television and offer users no real choices at all. Splash pages, pop-ups, and misery ideology.

One of misery design's most insidious recent examples is the idea of embedd using a service like IntelliTxt. By sullying the very concept of navigation, suc site, they poison the well for all websites. Such links make users even less li search engines to guide them to the next page.

This paragraph's title (extracts from www.useit.com/alertbox/20040830.html) is a good balance between creativity and specification. In fact, it explain enough specifically the referred content and it is possible to anticipate the related content.

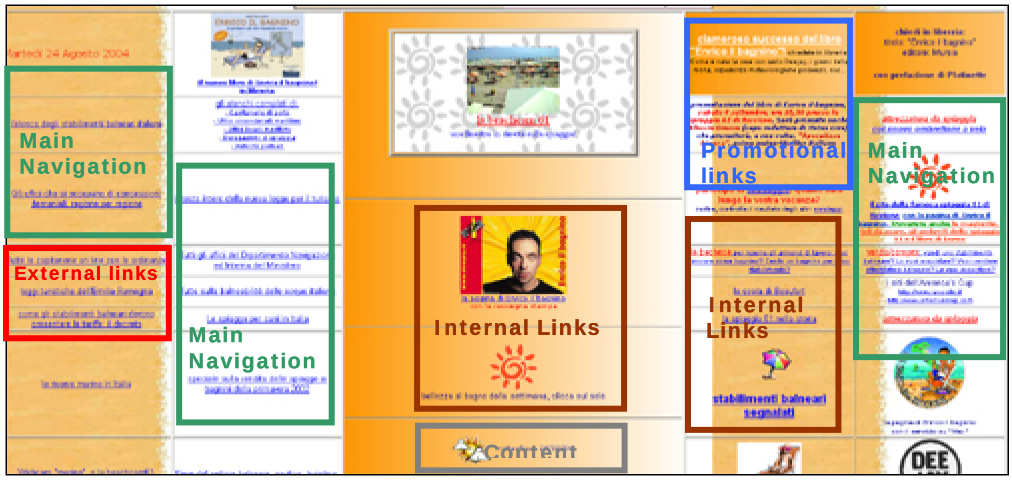
Problem — **Information Scent**

Action	1. Verify if each string of characters anticipates enough the content it refers to or if a thumbnail, a short text, a sound… could help for better explain what the textual string stands for.
Example	 In the case of main menu of the Giorgio Armani's website (www.armani.com) the label "Armani Exchange" does not anticipate enough the referred content. The image used for helping the user to understand the label, does not help too much. In this case there is also an ambiguity problem of the label (see problem above).

Feature	**Interaction Images (icons, photos…).**
Problem	**Conventionality**
Action	1. Identify all the interaction images within the page; 2. Verify if the interaction images (icons, photos…) follow standards and convention familiar to a web user (e.g. if you allow the user to download a .PDF document, used the standard icon -  - for communicate the document format).
Example	 This Icon used within the city portal of Como (www.comune.como.it) serves for informing the user that it is possible to download documents. Using this icon for representing the possibility to download files, could create some problems, in particular the user does not know what kind of file he will open/download.
Problem	**Intuitiveness**
Action	1. Within the page verify if there are interaction images that do not follow standard; 2. If they exist, make sure that they are intuitive for a first-time/web-novice, by means of the following actions: a. select a sample of users and submit them the interaction images (e.g. you can insert the icons in a word document); b. ask to every user the mean of each interaction image.
Example	 Testing this tool-bar with a sample of end-users (nearly 20) we have verified that it is not so intuitive. In particular the end-users do not understand the symbol "Add to favorites" and "Back to the top of the page".

Feature	**Macro-areas**
Problem	**Grouping adequacy**

Action	1. Map the macro-areas of the page:
	  a. map the macro-areas of the homepage;
	  b. map the macro-areas of each type of internal page (e.g. you can have a template for the products page and another for the contacts).
	2. Verify if the information units (for every macro-area) on the page are properly grouped with respect to their meanings, relations, and goals (e.g. if you have a macro-area for the main navigation, verify that all the links of this area lead to the main sections of the website).

Example	 Once mapped the types of messages in the home page of this web site (www.spiaggia61.it), it is possible to count at least 5 types of messages. The problem is that these messages are not properly grouped. For example the main navigation is positioned in three different places within the home page and sometimes the main navigational links are mixed with external or promotional links. In this case, the suggestion is re-think the message grouping.
Problem	**Position of importance**

Action	1. Map the macro-areas of the page:
	  a. map the macro-areas of the homepage;
	  b. map the macro-areas of each type of internal page (e.g. you can have a template for the products page and another for the contacts).
	2. Verify if the information units (for every macro-area) on the page are properly positioned with respect to their importance (the importance depends in the meanings, relations, and goals of the page).

Example	On the homepage above (www.spiaggia61.it) the main navigation as well as presents grouping problems, it is also not properly positioned. In fact, the main navigation is positioned in three places, but not too much highlighted. This design's choice does not allow an easy recognition of the links for navigating to the main sections of the web site.

3.c COGNITIVE HEURISTICS

Observing the interaction with a website two possible cognitive dimensions should be considered: on the one hand, the cognitive effort of the user while reading a single webpage; on the other hand, the cognitive aspects related to the understanding of the information architecture staying behind the web application as a whole, that is, the ground for understanding the whole meaning and structure of the website.

This document presents a number of cognitive problems and for each problem some usability heuristics are described.

The document considers two main features:

- *Cognitive heuristics related to a single page;*
- *Cognitive heuristics related to the Information Architecture.*

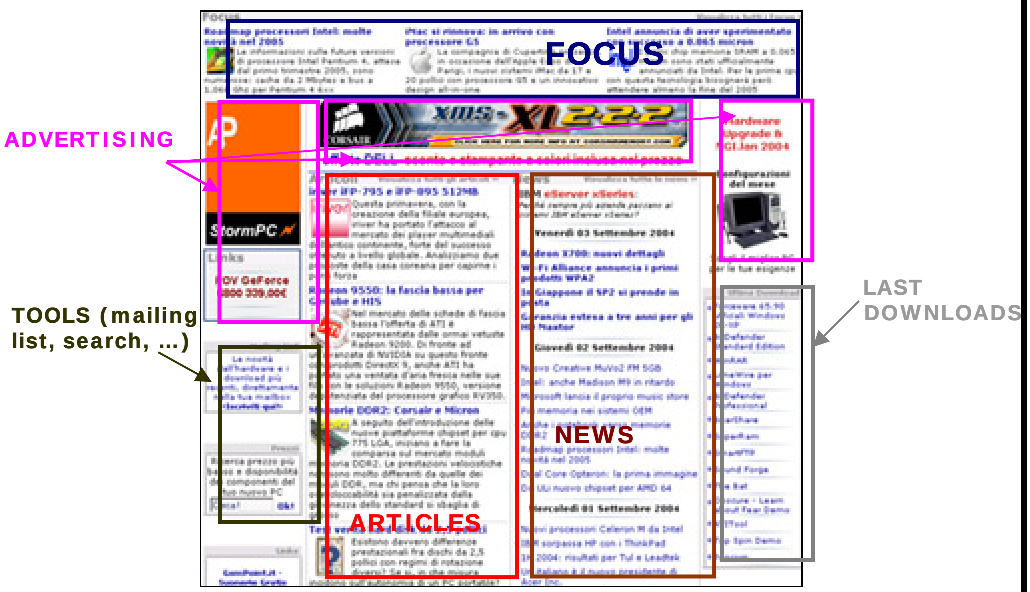
Feature	**Single page** Note: this feature (and related problems) could be verify for: • topic page(s), • group of topics page(s), • transition page(s), • Home page
Problem	**Information overload**
Explanation	A single page is composed by a set of different messages, each having a precise meaning. The quantity of the messages and their degree of heterogeneity could request an excessive effort for a first time/web novice to understand the whole page.
Problem	**Scannability**
Explanation	Users do not "read" the page until they find what they are interested in (a link, a text, an image). First of all, they "scan" it, basing on the structure of the page and how different messages are grouped and organised (in terms of macro areas).
Problem	**Grouping Adequacy**
Explanation	The messages composing a single page ca be grouped in information units, that is, in groups of messages having similar meaning, having a content relation or satisfying a common goal/functionality.

Feature	**Information architecture**
Problem	**Classification adequacy within group of topics and transition lists**
Explanation	The domain that the website describes is split in different information objects. The way these objects are classified within group of topics (e.g. paintings of 15^{th} century) and within transition lists (e.g. paintings painted by an author) deeply influences the user understanding and memorisation of the domain.
Problem	**Separation adequacy within topic pages**
Explanation	The content describing a particular topic of the website (i.e. the content describing a car in a car company website) can be split in more pieces (pages): this separation can help the user to better understand the topic itself (e.g. if we separate the presentation of a car in different pages – "Presentation", "Technical features", "Design" … the user can deeply and better understand the topic).
Problem	**Website Mental map**
Explanation	Users always try to create a mental map of the website, that is, to understand all the different topics described in the website and how they are organised and reachable. The understanding and memorisation of the information architecture positively influences the user experience with the website.

3.d COGNITIVE ACTIONS

How to use Cognitive Heuristics

The purpose of this document is to explain in an extensive way how to find the usability problems for every cognitive usability feature and to provide a step-by-step action guide for detecting the different problems.

Feature	**Single page** Note: these actions could be used for analyse: • Topic page(s), • Group of topics page(s), • Transition page(s), • Home page.
Problem	**Information overload**
Action	1. Try to identify the different messages presented in the page; 2. Count the number of messages; 3. Verify if the quantity of the messages and their meaning on a page is not overwhelming for a first time/web novice: a. enter in the page N-times (e.g. 10) with N different information goals (e.g. from the homepage find the review of a product, find the product X, find the event Y, contact the company…); b. For each goal, verify the needed time in order to understand where the right message is.
Example	 Analyzing a part of this home page (www.hwupgrade.it) it is possible to identify five types of different messages ("Articles/Focus", "News", "Downloads", "Tools" and "Downloads"). It is immediately evident that, even if the messages' categories are not too much, there are too information displayed on the page. The user, in particular the first time user, could have some problems for finding the interesting information.
Problem	**Scannability**
Action	1. Try to understand the meaning of the page in few seconds (e.g. in 5 seconds): a. Verify if the key messages are highlighted (by means of graphical symbols – bullets, icons -, multimedia files, keywords…) and try to understand the meaning of each message they refer to; b. Verify if the main sections in the page are clearly presented (e.g. search area, browsing area, registration area…).

Example

Scanning this page of Amazon.com it is possible to understand easily the meaning of the page. In fact, this page is used for promotional purposes, in particular for presenting the "favorite" (❶) and new books (❷). Besides, other promotional messages are showed (❸) for supporting the idea that this page is used prevalently for promotional purposes.

Feature	**Information architecture**
Problem	**Classification adequacy within group of topics and transition lists**
Action	4. Create one or more scenarios (define task(s)/goal(s) –e.g. Find information about the new book of John Grisham and user profile(s)); 5. Try to achieve the goal(s) of the scenario(s) and verify if the list of the group of topics (e.g. the list of the paintings of 15th century) and/or transition lists (e.g. paintings painted by an author) are organized with respect to the classification belonging to the common ground of the user.
Example	 Navigating within the city portal of Como (www.comune.como.it) - taking into account the scenario "Find the information about army called-up" - to reach this information is quite complicate. In fact, the user should select: "Services for citizens" → "Certificates, Documents,…" → "The Course of life" → "Army Called-up". In this case, there are two problems in classification of this information: 1. it is quite "strange" to classify general information within the sub-section called "Certificates, Documents,…"; 2. it is as much "strange the classification within the sub-section "The Course of life".
Problem	**Separation adequacy within topic pages**
Action	3. Identify an instance of topic of the website (e.g. the BMW 3 Series Coupé); 4. Try to understand in how many pages the topic is split (e.g. the BMW 3 Series is segmented in 5 pages/sections). 5. Verify if the content of the topic has been properly split into pieces (pages/sections) (e.g. in an ecommerce website the description of a running shoe should not be split in 20 pages, seeing that, all things considered, it is a "simple" product – at least for the user).
Example	 www.bmw.com The presentation of the BMW 3 Series Coupé is split in three main sections: "Introduction", "Highlights", "Multimedia Gallery". Besides, the section "Highlights" is divided in four sub-section ("Design", "Engines", "Chassis", "Safety") and each sub-section is also split in some

<table>
<tr><td></td><td>pages (e.g. the sub-section "Safety" is divided in "overview", "Airbags" and "Headlights". The fact of splitting the different information in these sections/pages, it is very useful for understanding the products. Besides, the "split strategy" is consistent with the "real world" (e.g. when we speak about the safety of a car we star with a "general overview" of the topic "Safety" and then we focus our attention on "sub topics" like airbags, headlights…).</td></tr>
<tr><td>Problem</td><td>Website Mental map</td></tr>
<tr><td>Action</td><td>1. Navigate randomly and/or taking into account one or more scenarios (define task(s)/goal(s) –e.g. Find information about the new book of John Grisham);
2. Once navigate through the website, take a sheet and try to draw (also in an informal way):
a. the high level map (main section and sub sections);
b. the contextual map of the different topics;
c. come back the day after and try to reach the same pages previously visited.</td></tr>
<tr><td>Example</td><td>Navigating randomly through the BMW website (www.bmw.com) it is easy for the user to create a mental map of the web site.

<u>High level map- main sections:</u>

- Products - Services - Fascination
The navigation among the sections is all to all (from each section it is possible to go to the others) and also to the "secondary" sections.

<u>High level map- "secondary" sections:</u>

- News - Site assistance - Contact - Careers - Site map - FAQs - Legal disclaimer
The navigation among the sections is all to all (from each section it is possible to go to the others) and also to main sections.

<u>Contextual map of the topic "Product"</u>
Formal representation:
- Products
:: Introduction
:: Highlights
: Design
. Overview
. Powerdome
. Front
. Rear
. Interior
: Engines
: Chassis
: Safety
:: Multimedia gallery
:: Models and data sheets
:: Catalogue
:: Security vehicles

(Very) Informal representation
The topic "Product" is split in different sections (4-5) and some sections have sub sections (e.g. Highlights is split in "Design", "Engines", "Chassis" and "Safety"). In some cases, the sub sections are divided in different pages ("Design" is split in 5 pages).</td></tr>
</table>

3.e GRAPHICS HEURISTICS

This level studies two aspects: the graphic design and the layout. The graphic design refers to choices bounded to colors, type of fonts, icons and other graphic elements on the page; the layout concerns to the spatial distribution of the graphic elements within the page.

Feature	**Overall graphic design**
Problem	**Visual identity**
Explanation	Lack of coordination with the visual identity of the company who run the site (if present).
Problem	**Use of a chromatic code**
Explanation	The correct use of colours in a website is very important for many reasons and helps the users in the navigation: - Colours can identify sections or subsections of the site; - Colours can reinforce the visual identity of the site; - Colours can attract the attention of the users on different elements of the pages (titles, links…); - The set of the colours of the site creates the look and feel of the site.
Problem	**Background contrast**
Explanation	The use of strong colours for the background or not suitable pictures can damage the readability of the contents of the website. Some matches of colours can be very difficult to read especially for people with visual disabilities.
Problem	**Font size**
Explanation	All fonts work at large sizes, problems start at smaller sizes. Text on the screen must be easy to read. Choosing the right font size is important to make it readable.
Problem	**Font colour**
Explanation	The colours used for screen texts must be accurately designed.
Problem	**Font type**
Explanation	Using a readable type of font with a readable size is important to make the reading easier.
Problem	**Text layout**
Explanation	Splitting a long text can simplify the reading. Very long pages (for example, containing an entire chapter) are difficult to scan, and scrolling up and down to refer to different sections of text can be frustrating. Also the wrong use of justification can make it difficult.
Problem	**Anchor identity**
Explanation	Anchors are used to reinforce the presence of a link on the page and it is very important to understand which are the anchors within the pages.
Problem	**Anchor states**
Explanation	When the mouse is over a link or after visiting it buttons and their anchors must communicate visible and well designed changes of state in order to help users in navigation.
Problem	**Icon consistency**
Explanation	Icons are used to represent topics to visit or tasks to do. It is important that the icon set matches with the other graphic elements of the site.
Problem	**Widgets consistency**
Explanation	Widgets are usually used to make up text and split it on the page in order to make it easily found in the text. The widget is a standardized on-screen representation of a control that may be manipulated by the user. Scroll bars, buttons, text boxes, text input area and radio buttons are all examples of widgets.

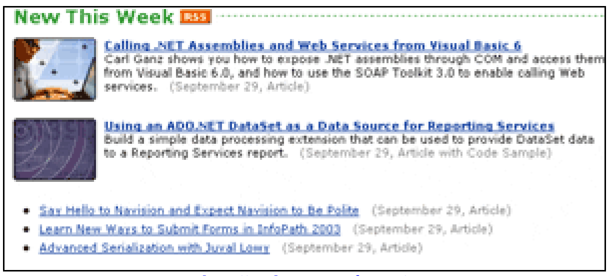
Feature	**Page layout**
Problem	**Position consistency**
Explanation	How objects are arranged on the screen determinates not only how good they look but how easy they are to understand and to use.
Problem	**Layout grid consistency**
Explanation	In the world of print and in the world of web grids give physical reference points to the space on the blank page. The role of the grid is clearest in designs that have a page-like appearance.
Problem	**Layout conventions**
Explanation	Users of western languages are conditioned to: - scan pages from left to bottom right; - assume that larger items are relevant; - assume that something above is more important that something below the page.

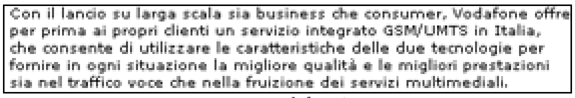
Feature	**Homepage**
Problem	**Redundancy – Overcrowded page**
Explanation	Because the screen has much lower resolution than a paper page, a screen that is filled with text, images, icons and other elements can be much harder to read.
Problem	**Page layout**
Explanation	Home pages have often free layout, this may cause problems in the users to understand the structure of the page.
Problem	**Use of Flash animations**
Explanation	Flash animations are used to make a site dynamic and interactive. Often these animations do not fit with the rest of the site.

3.f GRAPHICS ACTIONS

How to use Graphics Heuristics

The purpose of this document is to explain in an extensive way how to find the usability problems for every graphical usability feature and to provide a step-by-step action guide for detecting the different problems.

Feature	**Overall graphic design**
Problem	**Visual identity**
Action	1. Verify if the visual identity of the site is coordinated with the brand image. For this reason verify: a. if the company/institution logo is "always" correctly displayed; b. if the corporate colours are respected. c. if the overall website graphic style is consistent with the graphical style used for other media (e.g. the promotional brochures, video presentations, …).
Example	 Navigating through the web site of Ferrari (www.ferrari.it) it is possible to identify an overall graphic style that reflects the "heart of the company". Each main section ("Racing", "Cars", and "Corporate") is presented with a specific colour, but the "Ferrari style" is always in "background".
Problem	**Use of a chromatic code**
Action	1. Verify if all the colours of the chromatic set are used for their precise scope: a. verify that all the textual links have the same colour (if more colours are used, verify if it is clear the semantic behind this choice); b. verify that all the texts are written with the same colour (if more colours are used, verify if it is clear the semantic behind this choice); c. verify that all title, subtitle … are written with the same colour (if more colours are used, verify if it is clear the semantic behind this choice); 2. Verify the correct use of colours in order to identify website and/or page sections (e.g. the use of red colour for highlighting the news section); 3. Verify if the colours used in the site are not in conflict with the subject treated in the site (e.g. Black or dark blue for a kids website).
Example	 http://msdn.microsoft.com/ Within the main websites of Microsoft the textual links have the same colour (blue) and style (underlined). The only style difference is that there are bolded and normal links: the reason

	why of this choice is due to the links hierarchy (bolded link are more important).
Problem	**Background contrast**
Action	3. Verify if the background used does not obstacle the reading. 4. Verify how it influences the look of pages and the location of all other elements on the screen.
Example	 In this example (www.provincia.potenza.it/museo/default.htm) there is not contrast between the background (green) and the caption of this image (orange). For the user is very difficult to read the text that explain the image.
Problem	**Font size**
Action	1. Verify if the different types of text are readable (e.g. titles, subtitles, texts…). Research has shown that fonts smaller than 10-11-point elicited slower performance from users. For people over 65, it may be better to use at least 12 or 14 point. *Note: for verifying the font size it is possible to use a sample of users that try to read the content of the page.* 2. Verify if a suitable hierarchy is used among font titles, subtitles and texts, and if this is kept consistent across pages.
Example	Con il lancio su larga scala sia business che consumer, Vodafone offre per prima ai propri clienti un servizio integrato GSM/UMTS in Italia, che consente di utilizzare le caratteristiche delle due tecnologie per fornire in ogni situazione la migliore qualità e le migliori prestazioni sia nel traffico voce che nella fruizione dei servizi multimediali. www.vodafone.it Within many websites the texts are written with font size of 8-9 point without the possibility to enlarge it using the browser's functionalities (the font are fixed by style sheets - .css). In these cases the solution is either to enlarge the size at 10-11 point or to give the possibility to enlarge them.
Problem	**Font colour**
Action	1. Verify if there is an adequate contrast between colour of the text and the background colour. (e.g. Green text over a red background) 2. Verify if the colour of the text is readable on the page. *Note: for verifying the font colour it is possible to use a sample of users that try to read the content of the page.* .
Example	 In this example, the font colour used for designing the contextual menu does not guarantee an adequate contrast. The readability of the entire menu is harmed from this design solution.
Problem	**Font type**
Action	1. Verify if the font is a standard font, verify that the size is readable and if is possible to enlarge font size. For example, it is very important to use sans-serif typefaces such as Verdana for small text of 9 points or less since the low resolution of many monitors means that the detail of a serif font cannot be rendered fully; 2. Verify that the use of bold and underlined text is correct. For example is wrong to use underline text to spot something important because underline text means that we are in the presence of a link.

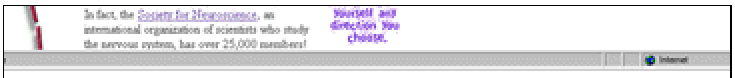
<table>
<tr><td>Example</td><td>


Figure 11. Readable font: Neuroscience for Kids by Eric Chudler

http://ebooks.strath.ac.uk/eboni/guidelines/guideline11.html

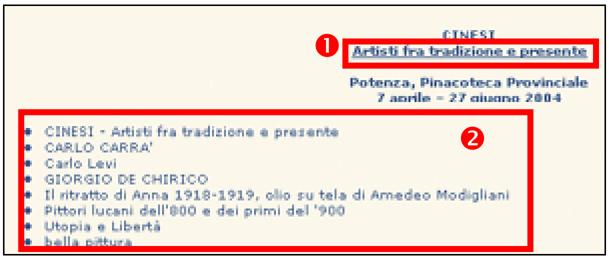
This image caption is written with a readable Font type (Verdana 7.5), even if the use of italic (Neuroscience for kids) could partially obstacle the readability.
</td></tr>
<tr><td>Problem</td><td>Text layout</td></tr>
<tr><td>Action</td><td>
1. Verify if the text is properly split into spaced paragraphs.
2. Verify if the text is not justified but aligned left or at lest centred. Justify text is the worst alignment for screen content.
</td></tr>
<tr><td>Example</td><td>

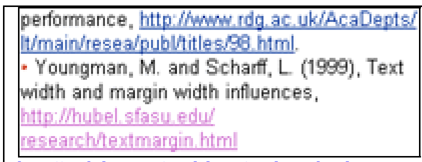

Fort Worth's White Elephant staked its claim to fame as offering the highest quality gambling and food service anywhere in the Southwest for three decades. It was memorable enough to merit respectful mention in the 1902 memoirs of outlaw-lawman James McIntire and the 1907 recollections of lawman-gambler Bat Masterson. It also won a different type of fame as the site of the famous shootout between Luke Short and Timothy I. Courtright.

The Fort Worth establishment began as a simple eatery, opened by F.A. Borocdine in 1884 in the 300 block of Main Street. The food was inferior to that at two nearby places – the Planter's House and the Commercial Restaurant – and the owner's profits remained anemic, so the place was seized by attachment. Within a year, the White Elephant reopened as a "Saloon and Billiard Parlor" with a small restaurant attached. The new ownership consisted of Jewish businessmen Gabriel Burgower, Nathaniel Bornstein and Samuel Berliner, who were not accepted by the local business fraternity. None of the new owners ever put down roots in the community, which did not help business either. Traditionally, saloons were home owned and home operated, and an owner was expected to greet customers and mingle with the crowd. Burgower, who was the on-site manager, split his time between the saloon and his more profitable jewelry business two doors up the block. Bornstein ran the eatery competently but without the flair that would later become a hallmark of the White Elephant's dining room, and Berliner was an absentee partner who also had part interest in San Antonio's White Elephant Saloon.

Burgower and company ran a modest operation built around a bar, some pool tables and a short-order kitchen. It was scarcely the kind of place calculated to appeal to the "sporting fraternity" of professional gamblers, nor did it draw the deep-pocketed amateurs. It took less than a year for the Jewish partners to realize they possessed a white elephant in the pejorative sense of the term.

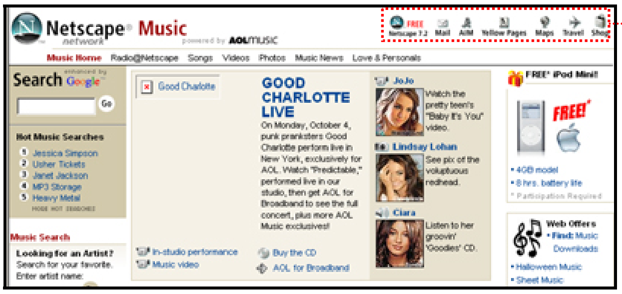
http://historynet.com/we/blwhiteelephantsaloon/

In this example, even if the text is too long, it is at least properly split into spaced paragraph and aligned left. This layout helps the user to better read and scan the page.
</td></tr>
<tr><td>Problem</td><td>Anchor identity</td></tr>
<tr><td>Action</td><td>
1. Verify if the link anchors have clear and conventional visual symbols associated to be distinguished from non-link elements.
</td></tr>
<tr><td>Example</td><td>


In this page, two different styles of textual links are used. For the first link (❶) a bolded and underlined font is used, while for the second (❷) only the blue colour is employed. In this case, it should be better to have a common anchors identity (style): in fact it should be possible that the user does not recognize that ❷ are links.
</td></tr>
<tr><td>Problem</td><td>Anchor states</td></tr>
<tr><td>Action</td><td>
1. Verify if the links/buttons communicate a visible change of state when activate.
</td></tr>
<tr><td>Example</td><td>


http://usability.gov/guidelines/readscan.html#one

In this example, the visited link is differently displayed with respect to non visited link. This design solution helps the user to immediately remember the visited pages.
</td></tr>
</table>

	 www.lugano.ch In this other example, when the mouse is over the link, a well visible change of state appears (the link becomes red and underlined).
Problem	**Icon consistency**
Action	1. Verify if the icons used are consistent and adequate with respect to the colours and the other graphic elements in the site. 2. Verify that icons work both as a group and as independent pieces. 3. Verify that icons are instantly recognizable, simple and cross-cultural.
Example	 In this example, it is clear that the icons' set used for representing the interactive sections or external web site are both coherent with the visual style of the website and instantly recognizable. Besides, it is possible to use them both as a group and as independent pieces. http://channels.netscape.com/ns/music/default.jsp
Problem	**Widgets consistency**
Action	1. Verify if the widgets used are consistent and adequate with respect to the colours and the other graphic elements in the site.
Example	 www.overture.com All the widgets used in Overture website are very adequate with respect to the colours and, more in general, to the overall graphical style of the website.

Feature	**Page layout**
Problem	**Position consistency**
Action	1. Map the macro-areas of the page's types (topics, group of topics, transition pages) and try to identify the elements composing the page layout; 2. Navigate through the website and verify if the elements positions are kept consistently across pages.
Example	

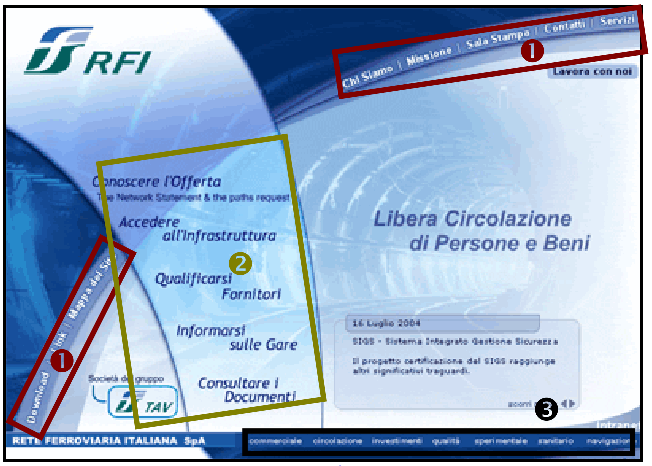
	 www.ubs.com Comparing these two pages of "Our company section" of UBS website, it is clear that the position of fixed elements is maintained consistent among the pages.
Problem	**Layout grid consistency**
Action	For all the different types of pages (topics, group of topics, transition pages): 1. Try to identify the layout grid (template); 2. Verify that all elements composing the template are kept consistently among the different types of pages;
Example	 www.amazon.com The books' pages of Amazon present a layout grid (template) that is kept consistency across the pages. In fact, all the main elements such as "Item information menu" (1), "Recently Viewed Items" menu (2), the "book information" (3), "Ready to buy" box (4) and "More buying choices" box (5) have a consistent position across the pages.
Problem	**Layout conventions**
Action	Verify that: 1. The page is organized for facilitating the user to scan from left to bottom right; 2. The more relevant items are larger than less important; 3. The more relevant items are positioned above and not below the page (for example

<table>
<tr><td></td><td>verify that the main menu items are not positioned in a less visible portion of screen).</td></tr>
<tr><td>Example</td><td>

www.lindt.com
In this example, the position of the link "L'azienda" ("the Company") is left-bottom. In general, this position could be a valid choice when the section presents only a brief history of the company and not dynamic and updated contents. But in this case, the section "contains also the possibility to access to "Investor Relations" section, which is a "sub-website" for finding all information about economical strategy and company's performance.</td></tr>
</table>

<table>
<tr><th>Feature</th><th>Homepage</th></tr>
<tr><td>Problem</td><td>Redundancy – Overcrowded page</td></tr>
<tr><td>Action</td><td>1. Verify it the elements on a page are not redundant;
2. Verify if the page is not overcrowded.</td></tr>
<tr><td>Example</td><td>

www.design.polimi.it
This Homepage provides too much information:
- the highlights of the University (❶);
- the news of the University (❷);
- the links to the useful sections for the students (❸);
- the main menu (❹).
Besides if we click on the link "next page" (❺) we reach another "home page" where more</td></tr>
</table>

	highlights are presented. *Note: this home page is optimized for 800x600 resolution, therefore all the elements are displayed within this (tiny) screen space!*
Problem	**Page layout**
Action	1. Verify if: a. the home page has a recognisable layout; b. the goals of each element of the layout are clear; 2. Verify that the layout respects the characteristics of the entire site.
Example	 www.rfi.it This homepage are not built with a clear page layout. In fact, it is very difficult to understand which the main menu is (❶) and the role of the other menus (❷,❸).
Problem	**Use of Flash animations**
Explanation	Flash animations are used to make a site dynamic and interactive. Often these animation does not fit with the rest of the site.
Action	Verify if the flash animation are coherent with the graphic aspect of the site, especially for icons, colours and graphic elements used.

TECHNOLOGY/PERFORMANCE

4.a TECHNOLOGY/PERFORMANCE HEURISTICS

The technology dimension of a web application is concerned with all those aspects related to technology choices and implementation style. The aspects that could be analyzed within this dimension are the formal correctness of the code (the site do not have to generate errors), the management of critical sections (e.g. operations) and the reaction of the system to errors or unexpected user behaviours.

Feature	**Errors management**
Problem	**System reaction to errors of a user**
Explanation	When some errors occur, the system is blocked and the user cannot go on.
Problem	**Scripting errors**
Explanation	Some Java- VB-Scripts codes could generate errors in particular conditions.
Problem	**Operations management**
Explanation	Hypermedia browsing during a procedure could cause errors or the operation to be cancelled.

Feature	**Browser compatibility**
Problem	**HTML interpretation**
Explanation	HTML is not supported and interpreted in the same way by every browsers (e.g. in visualizing tables and layers).
Problem	**Plug-ins**
Explanation	Installing plug-ins requires administrator permissions on the machine. This should be take into account when the web site used particular plug-in.

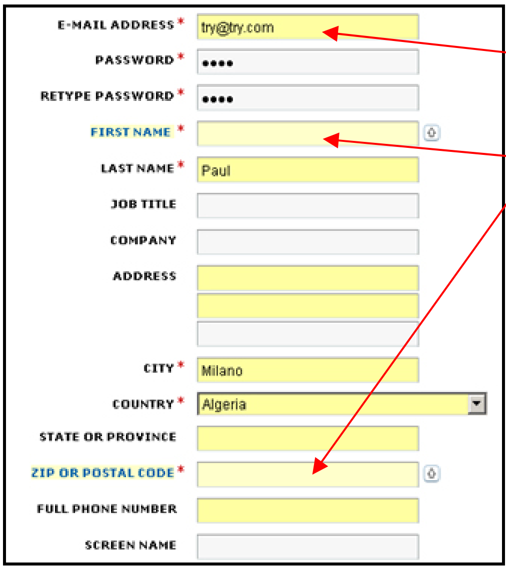
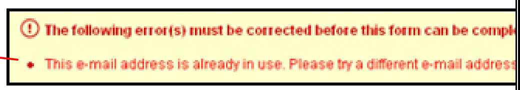
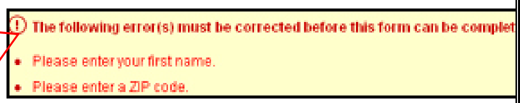
Feature	**Optimization**
Problem	**Page download time**
Explanation	The page has a too big size, the user should wait too much before seeing the content.
Problem	**Media streaming**
Explanation	Streaming audio or video could be not optimized for slow connections.

Open set: other may be added, according to the application domain and specific features.

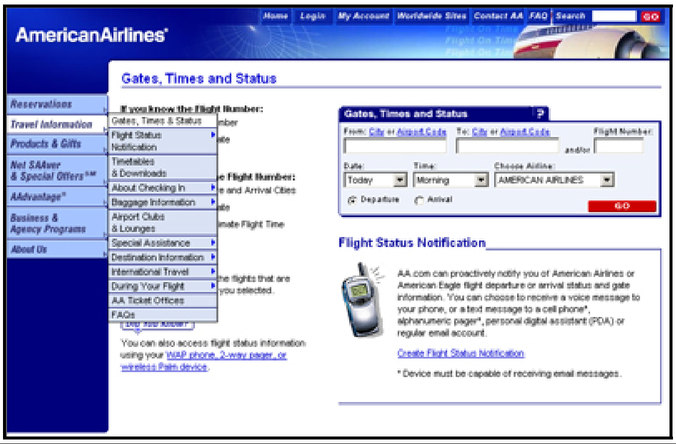
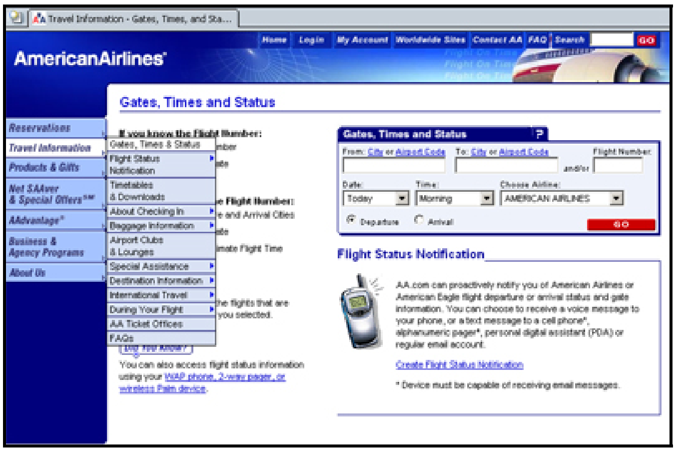
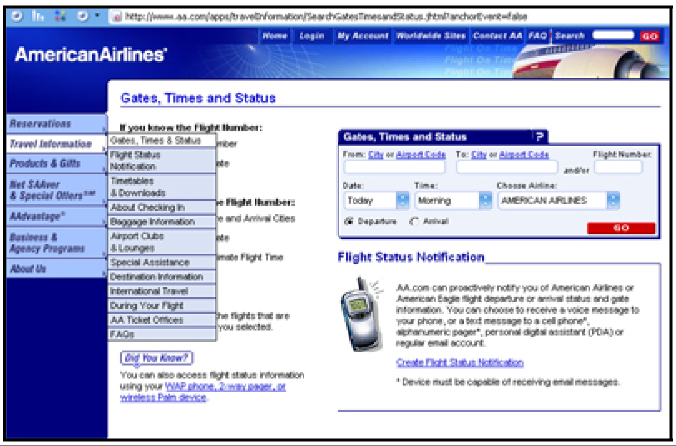
4.b TECHNOLOGY/PERFORMANCE ACTIONS

How to use Technology/Performance Heuristics

The purpose of this document is to explain in an extensive way how to find the usability problems for every technological usability feature and to provide a step-by-step action guide for detecting the different problems.

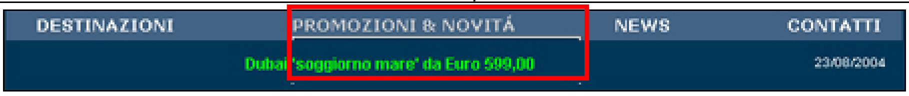
Feature	**Errors management**
Problem	**System reaction to errors of a user**
Action	*Action 1:* 1. Try to fill forms with incorrect information (e.g. for date, emails, countries, etc.); 2. Verify if the system provides an alert about the needed/failed information. *Action 2:* 1. Try to fill form and do not fill some mandatory form fields; 2. Verify if the system provides an alert about the needed/failed information.
Example	   Filling this form (example extracted from Macromedia.com) the system provides an alert for the email field (we have used an email that is already in use) and for mandatory fields (that are not completed). At the contrary, the system does not provide any alert for the fields "City", filled with Milano, and Country, filled with Algeria. It seems clear that the system does not match fields "City" and "Country".
Problem	**Scripting errors**
Action	1. Try to fill forms with special characters (", èàéöäü, ?, etc.) 2. Verify if the system accept and recognize special characters (and allows the user to continue the operation); OR 1. If the system does not accept special characters, verify if an alert is provided by the system.
Example	 Filling this form (extracted from Microsoft.com) we have insert a special character (#): immediately the system alert us about accepted characters and the correct procedure for filling

	the form.
Problem	**Operations management**
Action	1. Start an operation (e.g. buy the book "The Da Vinci Code") and, during the process (e.g. fill a form for buying the book), verify what happen when you browse out using some global navigation menus or the "back" button (e.g. you decide to buy another product and you want to read more information about it); 2. Observe the system reaction when you start a new operation (e.g. add the new product in the shopping bag): verify if the system resumes the previous operations (e.g. it presents a shopping bag with two products).
Example	 In the Amazon website (www.amazon.com) if you start an operation (e.g. "buy this book") and then you stop it and you select another book, the system resumes the previous operations (in this case it shows the previous selected books).

Feature	**Browser compatibility**
Problem	**HTML interpretation**
Action	1. For accomplish this action you have to install the most popular browsers. For example: a. Internet Explorer b. Mozzilla Firefox c. Netscape d. Opera 2. Try to browse the site using these browsers and verify if the website is correctly displayed.
Example	 *Internet Explorer 6*  *Netscape 7.1*  *Opera7.11* This page (extracted from American Airlines website – AA.com) is correctly displayed on the

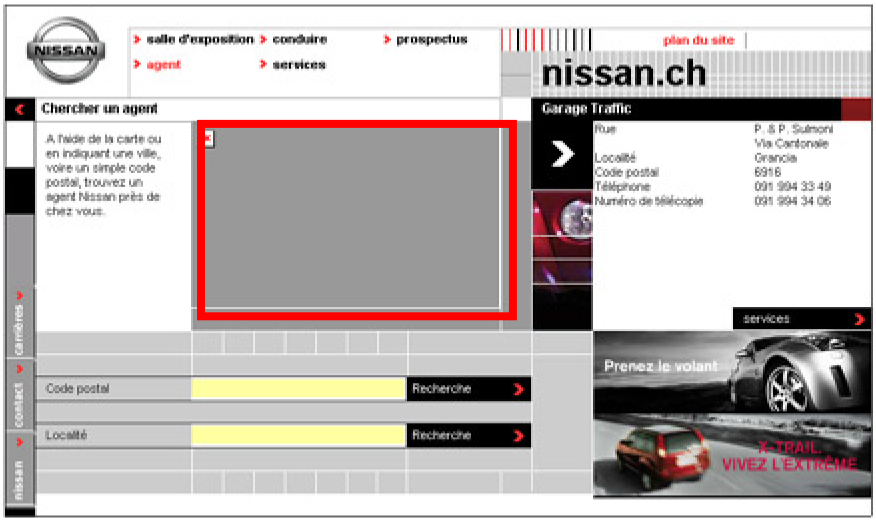
<table>
<tr><td></td><td>
three main browsers (Explorer, Netscape, Opera), even if complex javascripts and .css are used for generating graphical effects.

Internet Explorer 6



Netscape 7.1

Opera7.11

Instead, in this example the sub menu are not correctly displayed both using Netscape and Opera (the mouse-over action is not supported). It is clear that this problem can create serious navigational problems since the user can not select the subsections.
</td></tr>
<tr><td>Problem</td><td>Plug-ins</td></tr>
<tr><td>Action</td><td>
1. Browse the site using a non-administrator account;
2. Try to use every special feature (videos, animations, graphics, etc.);
3. Verify if the features are correctly displayed.
</td></tr>
<tr><td>Example</td><td>


Navigating within the Nissan.ch web site for searching an address of a reseller, the system provides the possibility to use a map for choosing the region and the city. The problem is that for correctly displaying the map a particular plug-in is required, otherwise it is not possible to use it the map (note: for navigating this website we have used a Pentium 4 with Windows XP operation system and all the common plug-ins installed). In this case, the only solution is to use a standard technology for creating these maps (e.g. Flash).
</td></tr>
</table>

<table>
<tr><th>Feature</th><th>Optimization</th></tr>
<tr><td>Problem</td><td>Page download time</td></tr>
<tr><td>Action</td><td>1. Use different types of connection:
a. 56 kb;
b. ADSL 256kb/512kb/ 1Mb
c. …
2. Browse the site and try to download pages (in particular pages with big images or videos) and verify the time needed for displaying every single page (you should not wait more than 10 seconds).</td></tr>
<tr><td>Example</td><td>

Using a connection ADSL 1Mb (Best effort) and browsing the web site of Malta Island (www.visitmalta.com) the pages are very slow to load. The faster pages are displaying in average 15 seconds, but there are sections (e.g. Interactive map) that are loading in more than 30 second. This web site presents a lot of contents and it looks well from the graphical point of view, but the pages' download time produces serious usability problems.</td></tr>
<tr><td>Problem</td><td>Media streaming</td></tr>
<tr><td>Action</td><td>1. Use a slow connection and try to stream a media file from the site;
2. Observe if data are transmitted fluently without scatterings or interruptions.</td></tr>
<tr><td>Example</td><td>

Watching online the Deejay TV (www.deejay.it) it is possible to verify that video and audio are transmitted in a fluent way. In the case of the user utilizes a very slow connection (e.g. 56 Kb/s), the system provides a message that explain that the best view is obtained with connections over 150 Kb/s. However, it is clear that is very difficult to watch online a TV with analogical connections.</td></tr>
</table>

Open set: other may be added, according to the application domain and specific features.

HEURISTICS SYNOPTIC TABLES

NAVIGATION HEURISTICS

HEURISTIC	FEATURE	LEVEL OF COMPLEXITY
Segmentation Orientation clues Accessibility of different pages	**Navigation within a topic**	**BASIC**
Introduction list Orientation clues Accessibility of topics	**Navigation within a group of topics**	
Transition list Orientation clues Accessibility of target	**Navigation within a transition**	
Landmarks Consistency Accessibility	**Overall Navigation**	
Orientation Backward navigation Depth anticipation	**Tree Navigation**	
Consistency Segmentation Orientation clues Accessibility of different pages	**Navigation within a kind of topic**	**ADVANCED I**
Introduction list Orientation clues Accessibility of group of topics	**Navigation within a group of groups of topics**	
"Go Back" History	**Backward Navigation**	
Orientation clues Control Navigation strategy Topology	**Guided-tour navigation**	**ADVANCED II** **NAVIGATION PATTERNS**
Orientation clues Control Navigation strategy Topology	**Index navigation**	
Orientation clues Control Navigation strategy Topology	**All to all navigation**	

CONTENT HEURISTICS

HEURISTIC	FEATURE	LEVEL OF COMPLEXITY
Accuracy Currency Coverage Content objectivity Authority Conciseness	**Text**	**ADVANCED**
Text errors Multimedia consistency	**General Communication quality**	**BASIC**

TECHNOLOGY/PERFORMANCE HEURISTICS

HEURISTIC	FEATURE	LEVEL OF COMPLEXITY
System reaction to errors of a user Scripting errors Operations management	**Errors management**	**ADVANCED**
HTML interpretation Plug-ins	**Browser compatibility**	**BASIC**
Page download time Media streaming	**Optimization**	**BASIC**

INTERFACE DESIGN HEURISTICS
(Cognitive, Semiotics and Graphics Heuristics)

Cognitive heuristics

HEURISTIC	FEATURE	LEVEL OF COMPLEXITY
Information overload Scannability Grouping Adequacy	**Single page**	**ADVANCED**
Classification adequacy within group of topics and transition lists Separation adequacy within topic pages Website Mental map	**Information architecture**	**ADVANCED**

Semiotics heuristics

HEURISTIC	FEATURE	LEVEL OF COMPLEXITY
Ambiguity / Clarity Labels Overlapping Generality vs. specificity Information Scent	**String of characters**	**BASIC**
Conventionality Intuitiveness	**Interaction Images**	**BASIC**
Grouping adequacy Position of importance	**Macro-areas**	**ADVANCED**

Graphics heuristics

HEURISTIC	FEATURE	LEVEL OF COMPLEXITY
Visual identity	**Overall graphic design**	**BASIC**
Use of a chromatic code		
Background contrast		
Font size		
Font colour		
Font type		
Text layout		
Anchor identity		
Anchor states		
Icon consistency		
Widgets consistency		
Position consistency	**Page layout**	**ADVANCED**
Layout grid consistency		
Layout conventions		
Redundancy – Overcrowded page	**Homepage**	**ADVANCED**
Page layout		
Use of Flash animations		

MiLE+

(Milano-Lugano Evaluation method)

Library of User Experience Indicators (UEIs)

Index

Introduction .... - 205 -
1. Content Experience .... - 206 -
2. Navigation & Cognitive Experience .... - 207 -
3. Interaction Flow Experience .... - 208 -

Introduction

This document presents the list of **U**ser **E**xperience **I**ndicator**s** (**UEIs**) useful both for conducting a User Experience Usability Inspection and for the Scenario-based user testing. The UEIs are divided in three dimensions corresponding to the different types of user interaction experiences. These dimensions are:

1. **Content Experience;**
2. **Navigation & Cognitive Experience;**
3. **Interaction Flow Experience.**

1. Content Experience

The Content User Experience Indicators measure the quality of user interaction with the content of the application.

UEIs	Description
Completeness	The user can find all the information required. The content is complete when it presents all principal elements that allow its understanding. The content should be complete both to information and to semantic level.
Richness	The richness refers to the quantity of information that explains the content. Note: The difference between richness and completeness is that the content should have all the needed information elements for explaining it (Richness) and these elements should be complete (Completeness). For example if we present a car the content has to speak about its security system, features, prices … (richness) and all these elements has to be complete (e.g. if we present the prices list, any price should be omit).
Comprehensibility	The comprehensibility is related with the capability of the content to be auto explicative. The main topic(s) of the content should be clear and not ambiguous.
Relevance	Relevance is the relationship between an informational need (which can be an explicit or an implicit question) and the answer(s) which meets it. "Relevance" is different from "truth". Example: given a person who is 32 years old and was born in Münster, if the question (need) is: where was he born?, and the answer is: he is 32 years old, the answer is true but not relevant, vice-versa, if the answer were München, it would be relevant but not true. The ideal case, of course, is when it is relevant and true.
Multilinguisticity	The content addresses to different type of users speaking difference languages, should be given in more than one language.
Multimediality	The use of different multimedia files is helpful for conveying the information. It is clear that these multimedia files should be consistency with the main topic and the goals of the information and they don't should overcrowd the page.
Satisfaction	The capability of the content to satisfy a user means that the information provided meets the desires, the needs and the goals of the user.

2. Navigation & Cognitive Experience

The Navigation & Cognitive User Experience Indicators allow the measure of how the navigation works and how the cognitive aspects of the application meet the cognitive world of the user(s).

UEIs	Description
Self-evidence	Self-evidence is the property of interactive elements used for supporting some elementary operations. These elements should be auto explicative, not ambiguous and consistency with the semantic of the operation.
Predictability	Predictability is the capability of interactive elements (symbols, icons, textual links, buttons, images…) to anticipate the related content and the effects of the interaction.
Learnability	Learnability is the capability of the application to be "clear" for the user. Using the application the user should learn the deep communication strategy supporting the entire system (it should be able to explain how the navigation works, which are the visual strategy for interactive elements, to create a map of the site…).
Information Overload	Information overload refers to the quantity of the message and their degree of heterogeneity. In fact they could request an excessive effort for a first time/web novice to understand the meaning of each message.
Accessibility	Accessibility refers to ensuring that content is accessible, ie. ensuring that Content can be navigated and read by everyone, regardless of location, experience, or the type of computer technology used.
Understandability	Understandability is the degree to which the purpose of the application, the navigation, the content and the interactive elements are clear to the end-user.
Memorability	Memorability refers to the mental faculty of retaining and recalling past experience. When users return to the application after a period of not using it, they should be able to re-establish proficiency the past experiences of use.

3. Interaction Flow Experience

The Interaction Flow Experience Indicators permit the measurement of how the interaction with the application is appreciated by the users.

UEIs	Description
Naturalness	Naturalness is the quality of web application of being natural with respect to the users' *common ground* both referred to the real and to the online world. So the application should present a general semantic that are easy to understand for the user (e.g. the icon representing a home is often used for representing the *go to home page* action).
Effectiveness	Effectiveness is the capability of the user to attain his goals. For reaching his goals the user has to pass through a series of tasks efficiently. The main measure for establishing the degree of effectiveness are: • Success rate of each task and goal; • Number of backtracks (within the execution of the task); • Time for performing the experience.
Engagement	Engagement is the ability of a system to ravish the user. The engagement is normally caused by the quality both of the content and of the overall interaction with the system.
Recall	Recall is the degree of overlapping between the searching space defined by the query and the one covered by the response. Therefore, recall can be view as one of the specific tools for measuring the coverage of content. In this sense, recall is the capability of the system to provide the needed information without much effort for the end-user (in particular if he is a novice user).
Precision	Precision is the purity of retrieval. Precision measures the semantic congruency between the information need (expressed by the query) and the response obtained by the system. Precision is a measurement tool for relevance. In this sense, it is the capability of the system to provide punctual information and not to overload the user with non-desired information.
Satisfaction of the experience	Satisfaction of the user experience means that the user has achieved all his goals. The satisfaction is reached by attaining others user experience indicators such as naturalness, effectiveness, engagement…: in this sense, the general satisfaction of the user experience is a macro-user experience indicator and the goal of the human-computer interaction.

Open set: other may be added, according to the application domain and expected user experience.

Applying Technical Heuristics and User Experience Indicators: some examples

(excerpts from the museum's websites domain).

Note: the websites used for the examples have been visited from September 2004 until February 2005.

Index

1. Applying Technical Heuristics: some examples 211

2. Applying User Experience Indicators: some examples..222

1. Applying Technical Heuristics: some examples

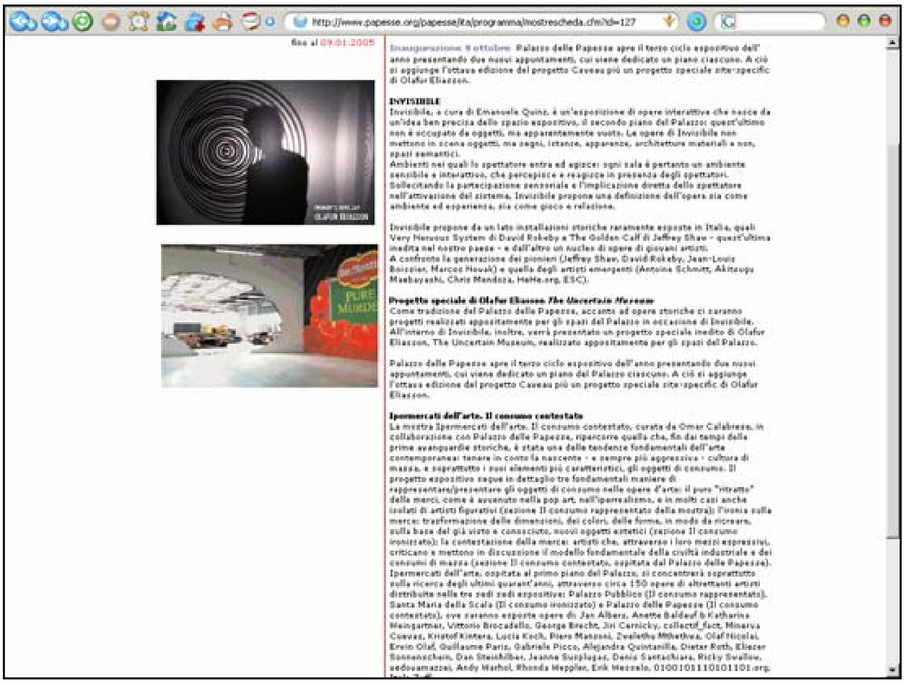
In this Chapter we present some example of infringement to the technical heuristics. As explained before the Technical Heuristics are divided by design dimension: Content, Navigation, Interface design (which includes Semiotics, Cognitive and Graphics) and Technology/Performance.

Content Heuristics

The content heuristics allow us to analyze the quality of the content (in terms of the efficacy of the communication) and for verifying if the contents and their structure correspond with the expectations of the users. The goal of the content heuristics is to verify the "technical" quality of the content presented in web applications.

Text conciseness

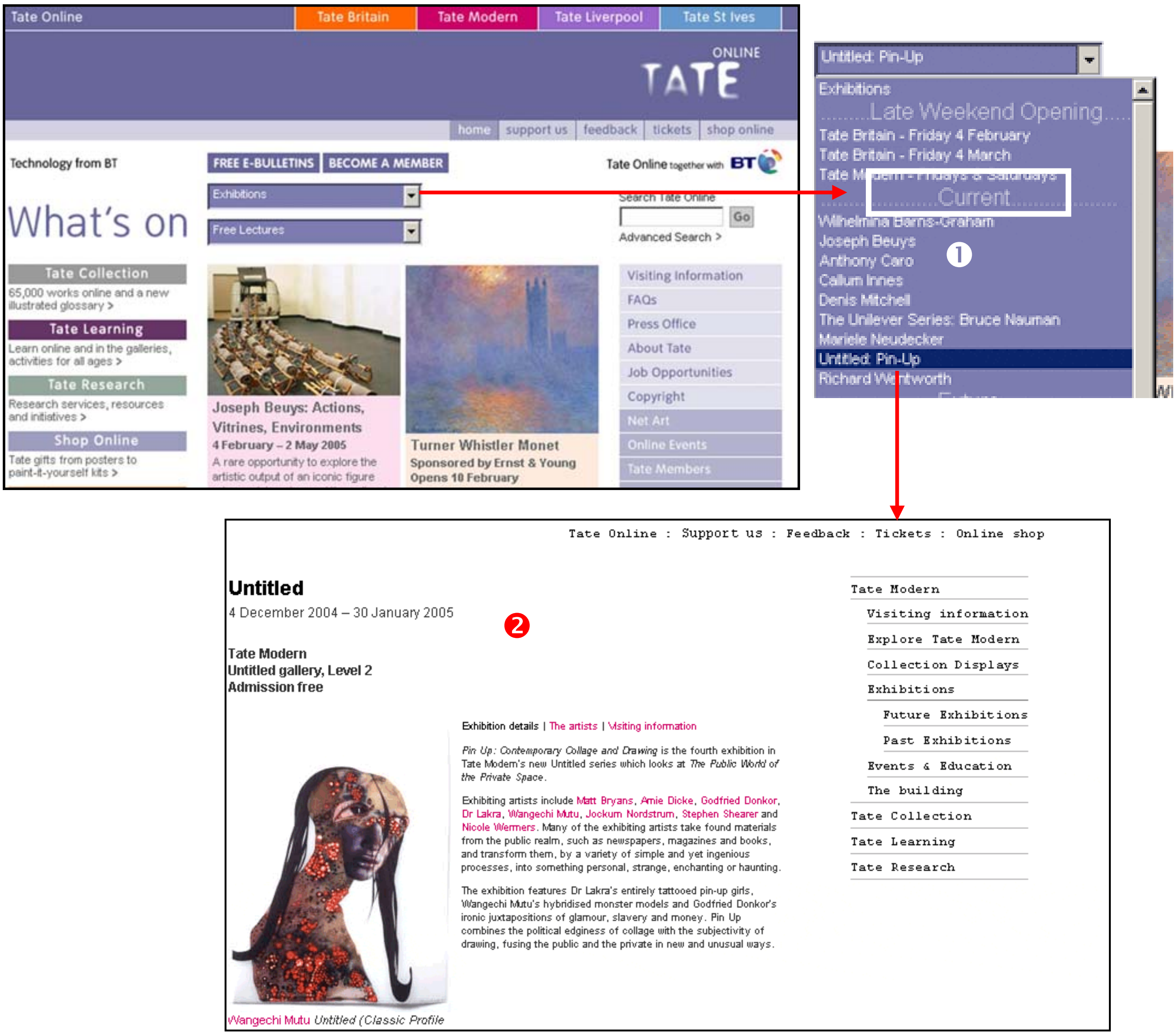
People rarely read Web pages word by word: they prefer to read few lines on the screen (15-25 lines). In this sense, conciseness is one of the most important aspects of the art of web-writing. For this reason it is very important to write an effective "short" and concise text.



This example extracts from the Papesse Museum website (www.papesse.org) shows a text long 50 lines. For the users this "wall of text" it should be very difficult to read. In this case, the web site should present an abstract, a summary of the text and then give to the user the possibility to download the .pdf version.

Currency of information

The electronic communication over the web is supposed to be delivered at the precise moment the reader accesses it; thus the offered content must be as current as the addressee perceives it, or must clearly show when it was published and the time scope of its validity.



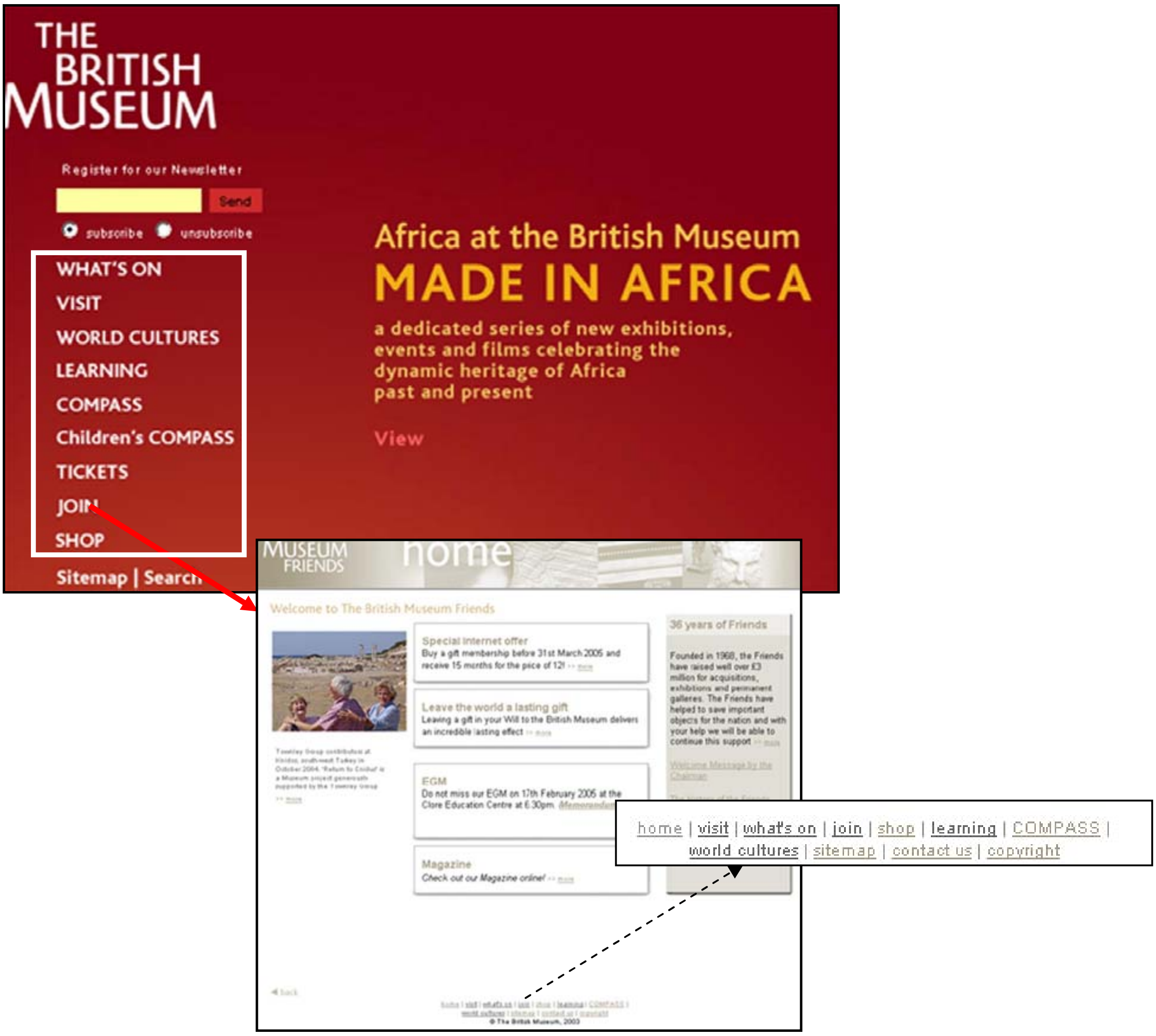
Visiting the Tate Gallery website (www.tate.org.uk) the 9th February 2005, it happens that the exhibition called “Untitled” and ended the 30 January 2005, it is still presented within the section “Current Exhibition” (❶). This lack of currency of the information provided could have a negative impact on the museum’s image.. Indeed, if a user is planning a visit at Tate Museum and s/he is interested in the "Untitled" exhibition and it does not read the end date (❷), s/he will unsatisfy once s/he reaches the museum.

Navigation Heuristics

The Navigational heuristics help to analyze a web application by taking into account from one hand the different ways that can be used by a user to reach a specific piece of information; on the other hand, the connections for passing from one content to another.

Landmarks for overall navigation

The access to the main sections of a web site is given by a number of landmarks. By using the landmarks the user can easily and quickly access all the macro-sections of the application. Therefore, the landmarks should be well highlighted on every page.



In every section of the British Museum website (www.thebritishmuseum.ac.uk) the landmarks for the overall navigation are not highlighted. In particular when the pages are long the landmarks are hidden (the user has to scroll for finding them).

"Go back" (Backward Navigation) in the navigation starting from an index/list
When the user reach a list he has to control the navigation from both the starting index to each element and to go back from one element to the index.

In the case of the Guggenheim Museum website (www.guggenheimcollection.org), once the user reaches the list of art's works now on view and select a painting (e.g. *Georges Braque - Landscape near Antwerp*) he achieves the selected page correctly. When the user tries to return to the list of art's works the backward mechanism is absent. The only navigational mechanism are two links called "Previous Braque work" (❶) and "Next Braque Work" (❷) that allow navigating within a guided-tour of the Braque's work.

Interface Design Heuristics

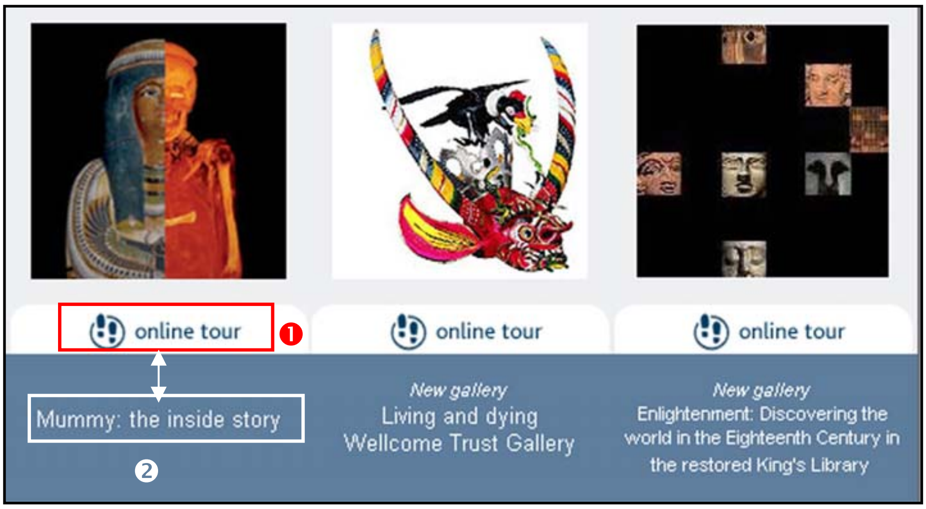
As explained before, the Interface Design is a broad dimension including Semiotics, Cognitive and graphics. These are some examples of heuristics related to these dimensions.

Semiotics Heuristics

During the interaction with a website the user should easily understand the meanings of the messages proposed. Semiotics heuristics help to understand if the messages are understandable for the users.

Ambiguity of Labels

The terms, the symbols, the icons… used could be interpreted with different meanings by the user, making her/him confused.



The main page of the current exhibition of the British Museum (www.thebritishmuseum.ac.uk/whatson/exhibitions/index.html) presents a problem related to the ambiguity of the labels. Indeed, the relation between the different signs (labels) it is not clear: for example, what is the relationship between the title "Online tour" (❶) and the label "Mummy inside story" (❷)?

If we clicked on "Online tour" (not just a title but also a link) and on "Mummy inside story" would we reach the same content?

In actual fact when clicking on "Online tour" (❶) we reach the online tour about "Mummy: the inside story" and if we select the label "Mummy: the inside history" (❷) we reach a section dedicated to a special exhibition about "Mummy".

Position of importance of the macro-areas

The most important areas of the page should be easily recognisable and should be properly grouped with respect to their importance.

Within the Centre Pompidou website (www.cnac-gp.fr/Pompidou/Accueil.nsf/tunnel?OpenForm) there are three macro-areas for main menus and it often happens that the macro-area ❷ and ❸ are not visible (the user has to scroll for using these menus). In particular the macro-area ❸ is used for presenting several links that are important for the users (e.g. *Exhibitions*, *Guided-tours*, Today events…).

Cognitive Heuristics

Cognitive Heuristics help to understand the interaction with a website considering two possible cognitive dimensions: on the one hand, the cognitive effort of the user while reading a single webpage; on the other hand, the cognitive aspects related to the understanding of the information architecture staying behind the web application as a whole, that is, the ground for understanding the whole meaning and structure of the website.

Information overload

A single page is composed by a set of different messages, each having a precise meaning. The quantity of the messages and their degree of heterogeneity could request an excessive effort for a first time/web novice to understand the whole page.

The main page of the Permanent Collection of Met Museum www.metmuseum.org/Works_of_Art/collection.asp presents 32 boxes for highlighting the museum's collections (the picture above represents only a part of the page). It is clear that from a usability point of view this page is information overloaded.

Website mental map

Users always try to create a mental map of the website, that is, to understand all the different topics described in the website and how they are organised and reachable. The understanding and memorisation of the information architecture positively influences the user experience with the website.

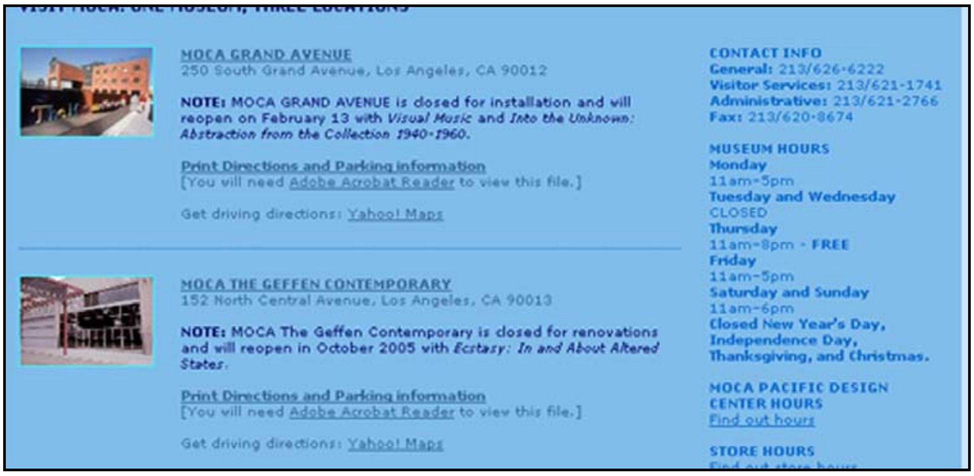
If we navigate for 5 minutes within the Exploratorium website (www.exploratorium.edu) it is very difficult to formalize the mental map. This problem is clearly related to the dimension of the website; however the user should be able to understand how the website is organised and how to navigate within the different sections.

Graphics Heuristics

Graphics heuristics are useful to investigate two aspects: the graphic design and the layout. The graphic design refers to choices bounded to colors, type of fonts, icons and other graphic elements on the page; the layout concerns the spatial distribution of the graphic elements within the page.

Background contrast

The use of strong colours for the background or unsuitable pictures can damage the readability of the contents of the website. Some matches of colours can be very difficult to read especially for people with visual disabilities.



Most pages of the MOCA website (www.moca.org) have a sky-blue background. This stylistic choice creates certain difficulties in reading the different texts; see that they are written with a blue font.

Font size

All fonts work in large sizes, problems start in smaller sizes. Text on the screen must be easy to read. Choosing the right font size is important to make it readable.

The font sizes used for the main menu and the section menu of the Whitney Museum (www.whitney.org) are too small for a comfortable reading.

Technical Heuristics

Technical Heuristics allow the analysis of the technology dimension of a web application, which is concerned with all those aspects related to technology choices and implementation style. The aspects that could be analyzed within this dimension are the formal correctness of the code (the site do not have to generate errors), the management of critical sections (e.g. operations) and the reaction of the system to errors or unexpected user behaviour.

Browser compatibility

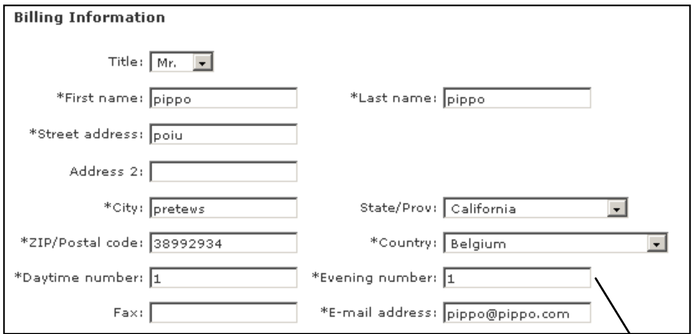
It is notorious that HTML is not supported and interpreted in the same way by every browser (e.g. in visualizing tables and layers).

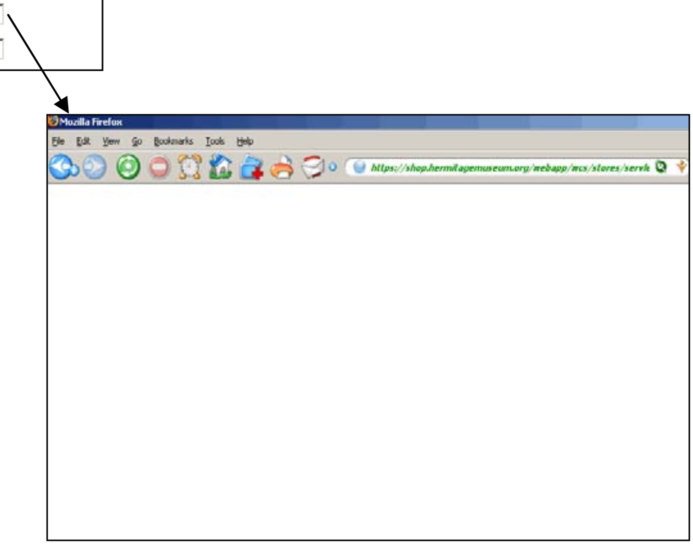
Explorer 6.0 Mozzilla Firefox 1.0

If we use two different browsers for displaying the "Listen section" of the Exploratorium website (www.exploratorium.edu/listen) it is possible to verify that the page is optimized only for Explorer. Indeed, if we use Mozzilla Firefox the content of the page is not correctly displayed.

System Reaction to User's Error(s)

It happens that when some errors occur, the system blocks itself and the user cannot go on.





Filling the form for buying the tickets for visiting the Hermitage museum (www.hermitagemuseum.org) the system does not provide any alert for the fields "State", filled with California, and Country, filled with Belgium. It seems clear that the system does not match fields "State" and "Country". In this case, the user does not understand which error s/he made, seeing that the system displays a blank page.

2. Applying User Experience Indicators: some examples

In this Chapter we present some example of how to use the User Experience Indicators within the User Experience Inspection. As explained before the User Experience Indicators are divided in three dimensions corresponding to the different types of user interaction experience. Indeed, the User Experience Indicators allow the evaluation of the adequacy of Scenarios (which are "Stories about use").
These dimensions are:

1. **Content Experience;**
2. **Navigation & Cognitive Experience;**
3. **Interaction Flow Experience.**

Content Experience Indicators

The Content User Experience Indicators allow the measurement of the quality of user interaction with the content of the application.

Let us take into account the following Scenario on the "Musée d'ethnographie de Neuchâtel" (MEN) website (www.men.ch).

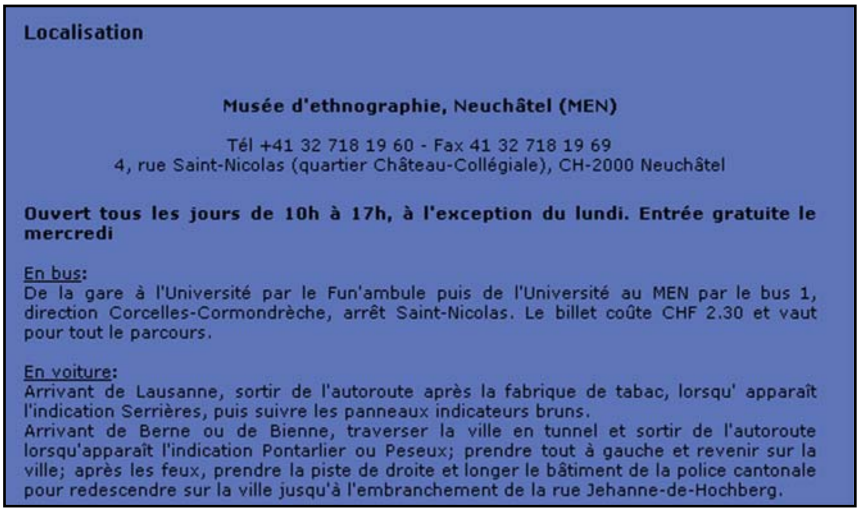
SCENARIO	Well-educated American tourist who knows he will be in town, he wants visit the real museum on December 6th 2004 and therefore he/she would like to know what special exhibitions or activities of any kind (lectures, guided tours, concerts) will take place in that day.
USER PROFILE	**Tourist**
GOAL	**Visit the M useum in a specific day**
TASKS	• **Find the exhibitions occurring on December 6th 2004 in the real museum** • **Find information about the museum's location**

Multilinguisticity
The content which addresses different types of users who speaks different languages, should be presented in more than one language.

Executing the Task1 (*Find the exhibitions occurring on December 6th 2004 in the real museum*) the information about the current exhibition and the events are only in French. This lack of multilinguisticity is a usability problem that has a negative impact on the image of the Museum (that addresses an international target and not only to a regional one).

Satisfaction on provided information

Considering the fact that we haven't found information in English about the collection and the current exhibition. Still we are very interested in the museum and we want to visit it (we try to complete the second Task- Find information about the museum's location). Therefore we need road markings for reaching the museum: also this information is given only in French! We (as American Tourists) are very unsatisfied with the information provided.



Navigation & Cognitive Experience Indicators

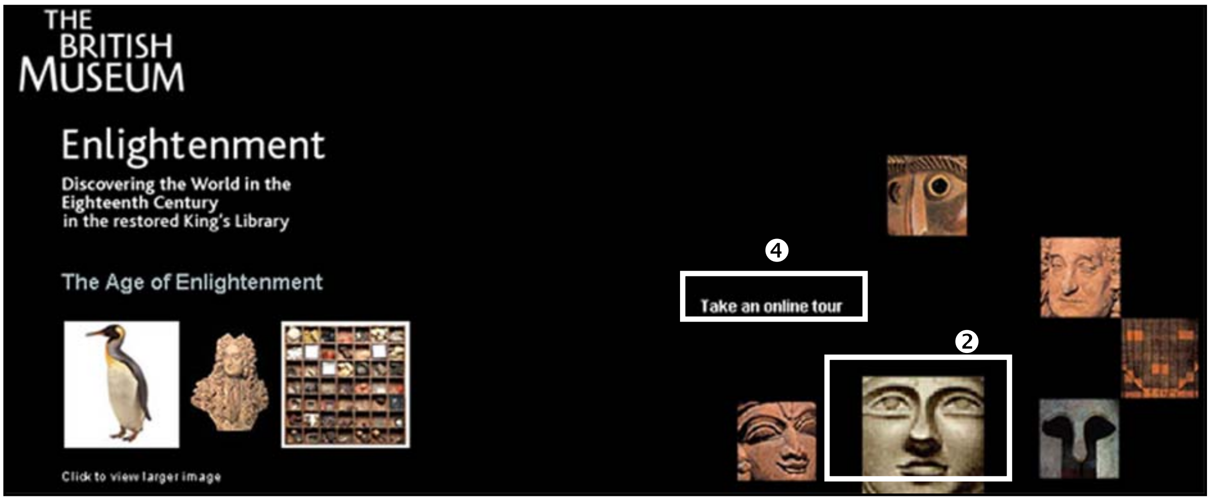
The Navigation & Cognitive User Experience Indicators allows the measurement of how the navigation works and how the cognitive aspects of the application meet the cognitive world of the user(s).

Let us consider the following Scenario on "The British Museum website" (www.thebritishmuseum.ac.uk).

SCENARIO	Marc is looking for some information about Enlightenment period studying at school.
USER PROFILE	Marc, High-school student.
GOAL	To be informed on a specific historical period (e.g. Enlightenment)
TASKS	• Find general information about this period; • Find detailed information about social and religious impact of Enlightenment period.

Links Predictability

Predictability is the capability of interactive elements (symbols, icons, textual links, buttons, images…) to anticipate the related content and the effects of the interaction.



Entering the section of The British Museum website dedicated to the Enlightenment period the user (Marc) has to roll over on the images ❶, ❷ for making visible the links labels *Tours* (❸) and *Take an online tour* (❹). The ambiguity between these labels could create some difficulties for predicting the different target content. This problem becomes worse by the fact that is not possible to visually compare the labels (as said before, the user has to roll over the images for reading the labels).

Memorability of online tours

Memorability refers to the mental faculty of retaining and recalling past experience. When users return to the application after a period of not using it, they should be able to re-establish proficiency the past experiences of use.

When users return to the online tours of British Museum Websites after a period of not using it, they could have some problems for re-establishing proficiency the past experiences of use. This problem is created both by the navigational structure of the online tours and by the semiotic design. It is clear that the main users do not have the capability for understanding the deep navigation structure and the reasons for selecting some symbols, icons and labels. For the users it is only important "to learn" quickly and easily how to use the website (in this case the online tours).

Interaction Flow Experience Indicators

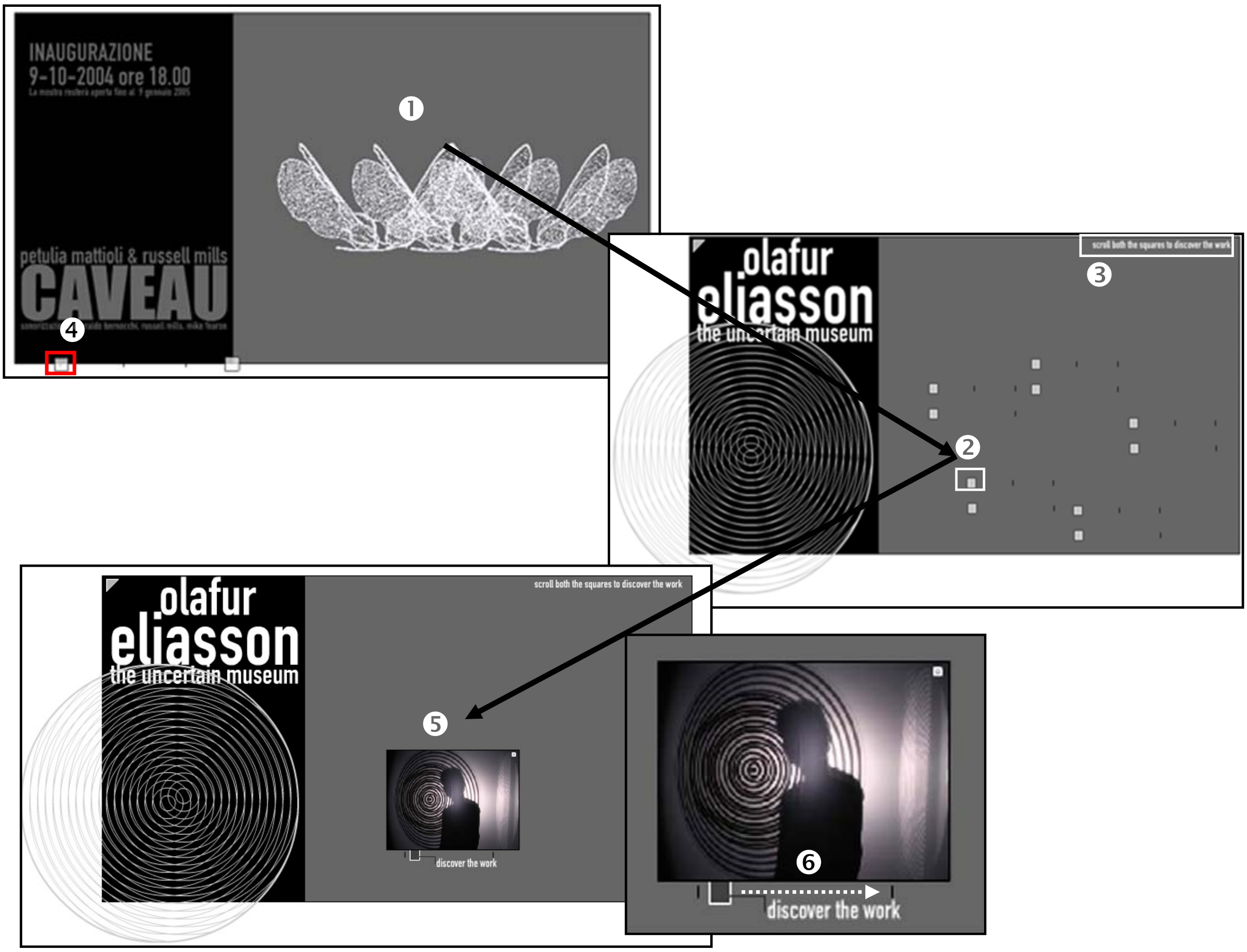
The Interaction Flow Experience Indicators permit the measurement of how the interaction with the application is appreciated by the users.

Let us consider the following Scenario on the "miniwebsite" of Papesse Museum dedicated to a particular exhibition
(www.papesse.org/papesse/minisiti/invisibile/index.htm).
Note: the exhibition ended the 9th January 2005.

SCENARIO	David is looking for some information about special projects actually presented in the museum.
USER PROFILE	Contemporary Art lover
GOAL	To be informed on special project (e.g. Ipermercati dell'Arte-Invisibile) Note: this project includes three sub-projects: *Invisibile*, *The Uncertain Museum* and *Art Hypermarkets. Contesting Consumerism.*
TASKS	• Find general information about special projects; • Find detailed information about a specific project (*Ipermercati dell'Arte-Invisibile)*

Naturalness

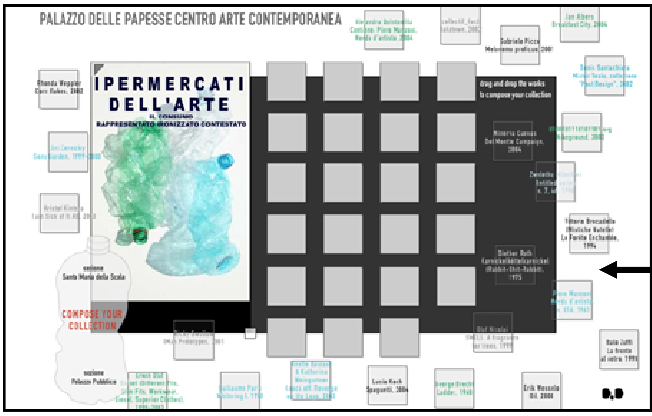
Naturalness is the quality of web application of being natural with respect to the users' common ground both referred to the real and to the online world. So the application should present a general semantic that are easy to understand for the user (e.g. the icon representing a home is often used for representing the go to home page action).

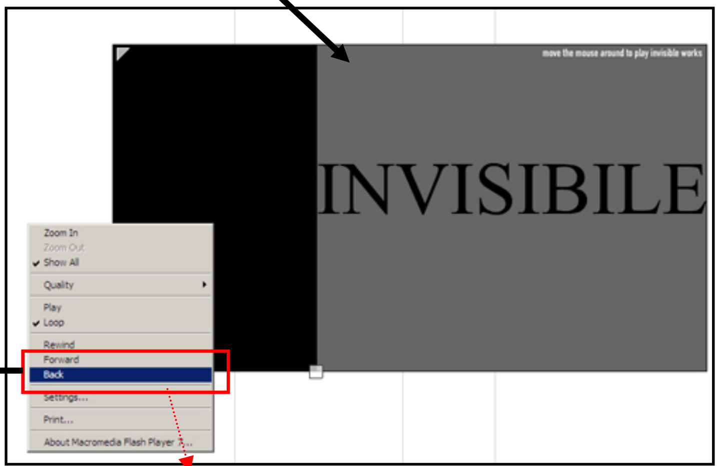


In order to reach the content of the website the user has to click on the images on the center of the home page (❶). A new page with some white squares appears (❷). Now, for reaching the content the user has to scroll the squares from left to right (however this mechanism is explained in the text❸). Once the content appears (e.g. a photo ❺), for zooming the image the user has to scroll the square from left to right ❻. This type of interaction is not kept consistently for all the square: for example for reading the content written by the curator the user has to click on the square ❹.

In this case the designer has created a new paradigm for the interaction that is not natural for the user.

In order to reach the content related to the *Invisibile* project the user has to click the Back button of the Flsh menu on this page.
Note: for activating the Flash menu the user has to click the right-button of the mouse (in windows OS).





In order to reach the content of *Art Hypermarkets. Contesting Consumerism*, the user has to click the Back button of the Flash menu on this page.

This website presents another problem related to the naturalness. Indeed, for achieving the content related to the projects *Invisibile and Art Hypermarkets. Contesting Consumerism* the user has to utilize the default menu provided by Flash. All the interaction for displaying the content of these projects is played on the use of the *Back* button of Flash: this is absolutely unnatural and outside of the Web-standard for the user interaction.

Engagement

Engagement is the ability of a system to ravish the user. The engagement is normally caused by the quality both of the content and of the overall interaction with the system.

Once the users (in particular Contemporary Art lovers) understand the interactive paradigm, they could appreciate (and find engaging!) the fact that the web site is a “virtual work of art” to discover. Even if, the website presents a lot of problems related to the usability (both from technical and user experience point of view)most of these problems should be considered in relation to the goal of the application. Indeed, the website might accomplish the main goals of the stakeholders (“we want a website representing a virtual interactive work of art”). This case study shows the importance in taking into account the goals and the requirements of the application during the usability evaluation. Despite, it is clear that a website which presents an interaction paradigm with this complexity it is usable only by a minority of the users.

Guidelines for Reporting

Usability Problems

Index

1. Introduction 233

2. Guidelines for Describing Usability problems...235

2.1 Separate Concerns 235
2.2 Balance Abstraction Levels...239
2.3 Extendibility: Represent Classes of Problems 241
2.4 Authority: Ground Your Findings 242
2.5 Prioritize and Set Importance 244
2.6 Technicality: Avoid “Usability” Jargon 245

3. Usability Report 246

3.1 Cover 248
3.2 Executive Summary 249
3.3 Table of content 250
3.3 Table of content 251
3.4 Introduction...252
3.5 Results of Usability Analysis 253
3.6 Synoptic of usability problems 255
3.7 Requirements for improvement..256

4. References 257

1. Introduction

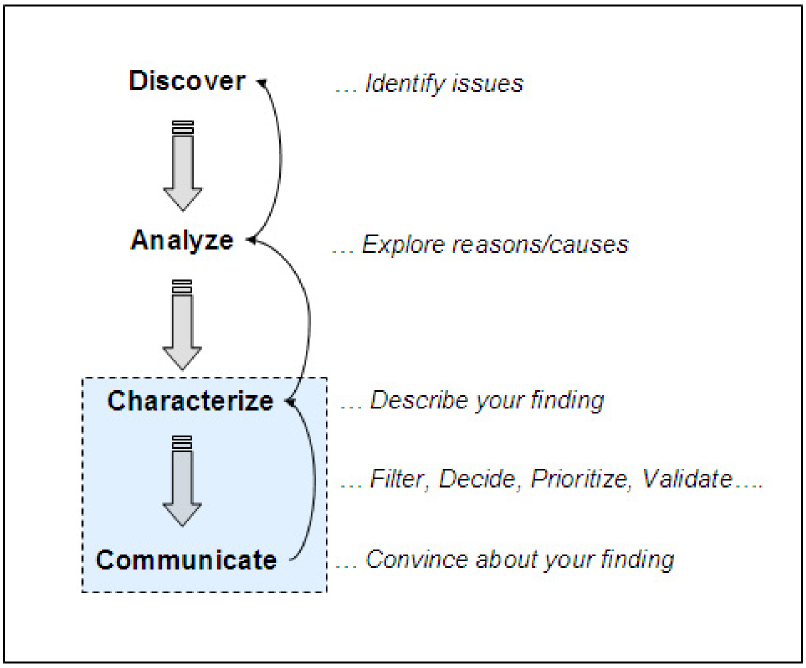
Discovering usability breakdowns is just half of the work in carrying out a usability evaluation. Since the outcome of a usability evaluation has to be reported to people who are not those who carried out the analysis, communicating usability problems effectively to the relevant stakeholders (client, designers, and development team) becomes a crucial concern, since the problems have to be taken into account and the recommendations have to be considered as a useful input for the redesign work (Bolchini D, Colazzo S, 2005).

The key findings of a usability analysis are a set of usability problems. A usability problem is an obstacle to the quality of the user experience". Usability problems should be the key result of any usability analysis, being it performed via inspection methods (Bolchini, Triacca & Speroni, 2003) or user-based methods (Dumas, Molich & Jeffries, 2004).

It is possible to summarize the process of managing usability problems in four fundamental activities:

- ***Discover**:* usability experts should be able to identify those aspects of the user experience which do not work (using one or more methods).
- ***Analyze**:* elaborate ideas, intuition and rough findings gathered from the discovering phase
- ***Characterize**:* means accurately and completely describe the findings and consistently orchestrate the analysis elements emerged so far for shaping coherent problems statements.
- ***Communicate***: means deciding what to say and how to say it according to the circumstance of reporting and to the addressee, how to prioritise and order the presentation of the findings, what to stress more and what to mitigate, which bridges should be built among the different parts of the analysis, and how convincing arguments have to be provided to support the results.

The usability problem lifecycle.



Usability problems are typically communicated in a "Usability Report". A Usability Report is written by usability experts for designers and stakeholders (e.g. marketing managers, product managers…), with the aim of convincing them of the usability problems' encountered relevance and suggesting indications for improvement.

In the first part of this document we focus our attention on several guidelines for "Characterizing" and "Communicating" usability problems, which are important suggestions for creating a successful Usability Report.
The second part is reserved for describing in depth the Usability Report's structure.

2. Guidelines for Describing Usability problems

The guidelines provided in this section are related in particular to the Characterization and Communication of usability problems.

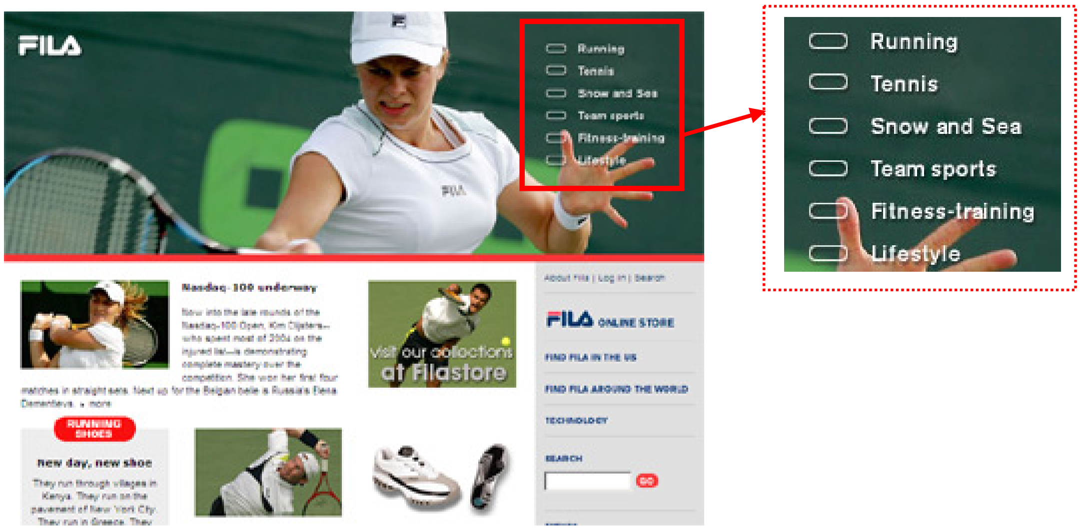
2.1 Separate Concerns

Guideline: Decouple a usability problem into the various "application aspects" it is concerned with (content, navigation, semiotics, graphics, etc.)

Consider the statement: "The navigation is hard in the section X". This consideration is too vague to be considered a usability problem. Indeed, the difficulty of navigation could be related to several causes (e.g. link names, link position and order, content structure, navigation structure…). These problems are thus related to different design dimensions or application aspects (content, navigation, semiotics…) as well, that can also help characterizing the definition of the usability problem.

A proper characterization of the nature of a usability problem facilitates the intervention for fixing it.

Consider the following example excerpts from FILA website (www.fila.com, 27 March 2005):



(1) FILA Home page main navigation bar

(2) FILA main navigation bar

<u>Description of the problem:</u>

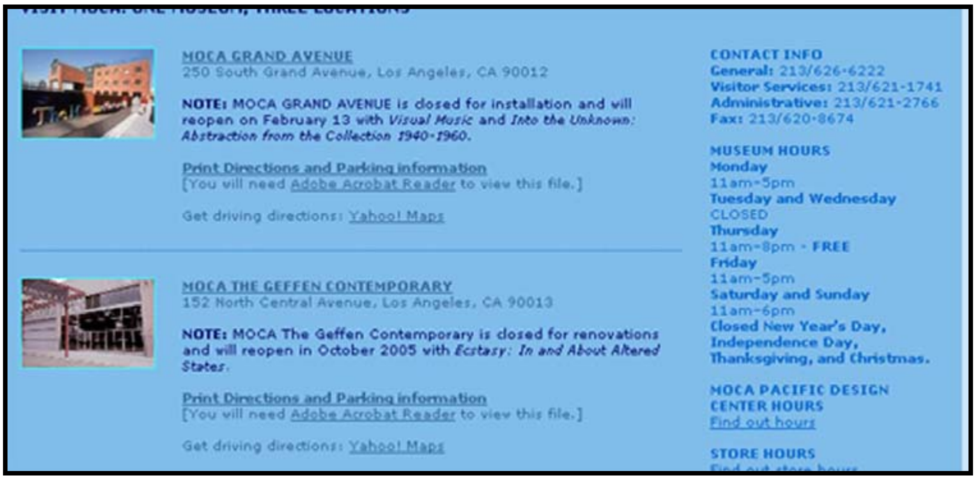
Within the FILA website there is a **semiotic** problems related to the main navigation bar which affects the navigation of the user. In fact, the navigation bar presented in the homepage (1) is clearly understandable; but once we navigate within the internal pages the bar is different (2). The "internal" main menu (2) is not very intuitive for two "semiotics" reasons:

1. Using icons instead of Text labels does not allow to get a quick overview of the main sections (the user has to remember exactly that, for example, "Team sports" button is the fourth from the left);
2. The icons used are not very predictable: it is not intuitive to understand that each rectangle is referred to a site section.

Guideline: Distinguish between problems which are application-independent and problems depending on the purpose of the application (e.g. scenarios supported).

An important separation of problems is between those which strictly concern the design level (also called "technical problems") and problems which are strictly related to the application scenarios (described also in the Module 2 – Inspection Methods).
Technical problems (those concerning navigation, consistency in layout, in link labels, in information architecture, or technological breakdowns) are typically application independent problems, meaning that they can be well detected without knowing the specific purpose and communication goals of the application.

Consider the following example excerpts from MOCA website (www.moca.org, 25 March 2005):



This page shows a typical Technical Problem related to the graphics: the contrast between text and background does not allow one to read the content of the page. The poor contrast is independent from the type of application: it is always a problem! (Being it a museum website, a university website or an e-commerce website).

Other, more crucial problems are those who significantly obstacle the completion of important user scenarios. However, in case that the actual application requirements are not known (strategic objectives, communication goals, specific scenarios that have to be supported), and the analysis of the problems emerged strongly depend on this information, analysts should limit themselves to raising issues about the problems encountered (keeping open the possibility of considering it a real problem or not), providing the stakeholders the elements to reason about their requirements in relation with the issue raised.

Consider the following example extracts from the "Musée d'ethnographie de Neuchâtel" (MEN) website, 22 March 2005):

In this example, one plausible scenario is that foreigner tourist (non French speaking) wants to have more information (in English) about the museum to decide whether to visit it or not. Unfortunately the website is only in French.

This lack of multilinguisticity is a usability problem strictly related to the requirements and the goals of the application: are foreign tourists intended targets of the application? If yes, the lack of multilinguisticity is a usability problem. If not, it is not a usability problem.

2.2 Balance Abstraction Levels

Guideline: Describe usability problems keeping a consistent level of granularity, moving from general to specific and without abruptly mixing details with strategic issues.

Macro-problems should not be confused or intermingled with micro-problems. The difference in granularity may be decided according to the degree of impact of the problem on the overall application. It is important to characterize in depth high level and very general problems before digging into the details of the problems concerning specific features.
Important details such as “difficulties in subscribing to the mailing list via form” should not hide or being intermingled with issues at a higher level of abstraction such as “purchase service declared but not actually possible”.
The level of abstraction of the problem is a good sign of how deep the analysis is and how usability experts master the results delivered. By level of abstraction we mean the degree of detail by which a problem is described. Moving from general to specific is also effective for having stakeholders agree first on the major issues, and then discussing the details. The level of abstraction in which problems are characterized and reported should be kept consistent and balanced, at least for two reasons:

- if analysts start to focus on fine grained issues (not necessarily less important), they risk to lose the “whole picture” of the application;
- stakeholders are facilitated in following your reasoning moving from general concerns to detailed ones.

The same usability problem may be described at different levels of abstraction.

Let us consider the example illustrated in the following schema, excerpted from the usability analysis of the Hermitage Museum (www.hermitagemuseum.org, 10 January 2005).

Services	Information
Shop Online	Ticket Prices
Museum Shop with E-Shop functionality	Telephone and Fax Numbers
Shops inside the museum	Directions
Scholarly Publications	Tours and Lectures
Ordering Images and Photographs	Floor Plans
Education and Technology Center	Contact Information
Info Stands	Publications
Ticket Bookings	Museum Departments
Floor Plans Location of Services	...

Organization of content in section "Services" and "information" in the Hermitage Museum website. Arrows show the similar content of the two sections.

The problem shown in the example may be described at least at three different levels of abstraction:

A1. The criteria used for the information architecture is not always clear. See for example the redundancies in section "Service" and "Information".

A2. Some link names in the "Service" section overlap almost entirely with some link names in the "Information" section. Having this situation, the user is never sure to have consumed all the content available for a given tropic (e.g. museum publications).

A3. The difference between the sections „Information"and „Services"remains unclear, even more so because the contents of these two sections are partially overlapping and repetitive. As the scheme shows, the grey labels are very similar; the white labels are quite similar. The distinction between "Shop online", "Museum shop with E-Shop functionality", "Shops inside the museum" and "Ordering images and photographs" is not quite clear and the orientation is rather complicating than clarifying it. Also the difference between "Telephone and Fax Numbers" and "Contact Information" (consisting of a long list of telephone numbers and addresses), is not clear.

All three statements are true and accurate in describing the problem illustrated in the schema. Not all of them are equally relevant in every communication context. A1 is a very general statement characterizing a problem at the information architecture level. As an example, it refers to the specific section illustrated in the picture, without commenting further on. This quite high level of abstraction is good for overview purposes, such as executive summaries or synoptic, or conclusions, to give the essence of the usability problems and then pointing to more detailed descriptions.

A2 describes the problem in the specific case, interpreting the issue at the semiotic level (link labels) and providing a grounding related to the effect on the user experience. This middle level of abstraction is quietly focused on the details and of course should be complemented by a description of the "information architecture" concern (it is not enough to change the labels to fix the problem). This statement can be considered as a synthetic description of the specific problem.

A3 adds a number of details to the problem, describing what precisely is not clear in which labels and introducing a further concern, which is the degree in which the content sections are overlapping. This is a low abstraction level which is flattened on the details of the specific case and it is good for discussing in-depth the two specific sections at issue, typically as a comment of the table.

2.3 Extendibility: Represent Classes of Problems

Guideline: In case of complex applications, describe each problem as a representative of a class of problems, and characterize each problem by providing a general statement pointing to specific examples

During the usability analysis it often happens that the usability expert does not have enough resources (time and budget) for examining the entire application. So it often happens that problems are recorded and described as they emerge from the analysis, wrongly assuming that the spotted issue is so unique and peculiar. As a consequence the characterization of such a problem is only valid for the incidental context in which it emerged, and the consequent recommendation for improvement just addresses the fixing of that specific situation (Bolchini D. and Colazzo S., 2005).

Let us consider the following problem description (adapted from a problem detected within the Rijksmuseum website, www.rjksmuseum.nl, 8 January 2005):
As entering the category "Jewels", the user can choose among "handmade jewels" but also "bronze", "everyday tools", and "wooden products", whereas the user expects to find only jewel-related objects, or subcategories of jewels.

Clearly, this is an important problem to highlight, but the questions that the expert has to ask himself are:

- Is it an isolated problem?
- What happens in the others categories?

In this case the expert does not have enough resources to analyze in depth all the categories, he has to select randomly 4-5 other categories to verify if the problem is isolated or repeated.
Then, when it comes to precisely characterizing the problem, usability experts should carefully describe it as a general problem (if found more than one categories), pointing to the specific case of the "jewellery" as an example.
In other words, the specific case emerged should be described not as "the" problem to solve but as a representative of a class of problems, which may likely occur in other parts of the application.

Note how the same problem can be characterized according to the explained guideline:
Collection categories present objects, which should not belong in the category that they are in. This may lead to a confusing and disorienting browsing of the categories offered, and make it difficult to the user to locate an object of interest.
Examples:

- **Category "Jewels", contains the subcategories "handmade jewels" but also "bronze", "everyday tools", and "wooden products".**
- **Category "Painting", contains Religious Pieces", "Masks"…**

It is not clear by which criteria objects fall in the categories.

2.4 Authority: Ground Your Findings

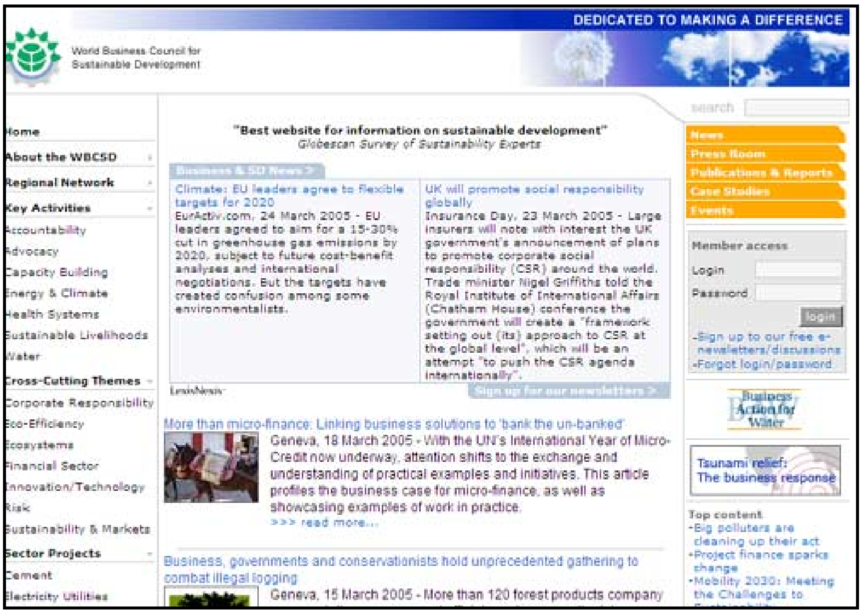
Guideline: Give reason of your findings by drawing to elements which can gain credibility, such as the experience of the analysts, the impact on the user experience and the compliance with the standard and convention.

Authority supports the credibility of the findings and it is necessary to ground them. Communicating usability issues means that it is necessary to highlight and underline the source of credibility.
We can have three main sources or pillars for usability results to be credible:

a) Experience: the analysts is recognized as knowledgeable and experienced.
b) Consequences: usability problems are shown to have an impact on the actual user experience
c) Anomalies: usability problems patently infringed standards, good conventions or common practice in the domain

None of these pillars alone can fully gain the needed authority for usability findings. These elements have to be properly combined, and all have to contribute with different weights to consolidate the credibility of the usability problems. In this way, usability problems can acquire their authority and start to become worth noticing for the stakeholders.

Consider this Example extracts from the World Business Council for Sustainable Development web site, www.wbcsd.org, 25 March 2005 :



Example of Homepage Information Overload.

Comment: *"There is information overload within the Home page".*

Why should this be considered a usability problem? For supporting our finding we can start investigating pillar b) (*Consequences*). In fact, all the information provided on the homepage could distract the user from his task; so he could have some difficulties in understanding which section/part of the website is interesting for him.
This assumption is also grounded on the pillar c) (*Anomalies*). Indeed, it is a good and recognized practice to not overload the pages. From a cognitive point of view the quantity

of the messages and their degree of heterogeneity should not request an excessive effort for the users in understanding the page.
According to the case comment, this may also be founded on the experience of the reviewer (pillar a), if he has long-standing experience in evaluating the usability of websites.

2.5 Prioritize and Set Importance

Guideline: Communicate your findings in order of importance: according to the circumstances, importance may mean gravity for the user experience or estimated effort needed to fix the problem.

It is important to underline that the order of problems is perceived by the addressees of the usability report in an order of importance (say first what is most important).
On one hand, importance may mean the "gravity" of the problem for the user experience ("unclear link labels on the home page" is a more important problem than "having the 'edit quantity of shopping bag item taking 2 seconds to work"). The first problem may absolutely hinder the location of the content, while the second may have the user wait a bit longer for the operation to be completed.
On the other hand, problem importance might also be interpreted as the amount of "burden" or complexity for the designers to fix it. In this case the above priorities should probably be inverted: a semiotic expert can write clearer labelling for the home page in few hours, whereas having to speed up the "edit quantity" operation may need to reconfigure, reinstall or change the business transaction software behind.

	Gravity for the users	**Effort to fix**
Problem 1	10	0.1 person/month
Problem 2	8	0.4 person/month
Problem n	5	0.2 person/month

Example of Table reporting the "gravity" of usability problems.

The numbers reported in the table can be ranked according to any of the two criteria.

2.6 Technicality: Avoid "Usability" Jargon

Guideline: Adapt your concepts and wording to the target audience you are communicating with, in such a way that stakeholders clearly understand the essence of the problems (Dumas, J., Molich, R., & Jeffries).

The right choice of words for communicating the issues emerged during the evaluation is one of the most important aspects of a usability report. It is fundamental to take into account that the addressees should not know the principles of usability and design or even the concepts of the methods that usability experts use.

Example (bad):
Technical heuristics T4 showed that the structural navigation within nodes of the entity type "painting" is inconsistent and not predictable.

This type of statement is unclear for addressees without a design and usability background. The consequence is that they may start considering the usability analysis too obscure and ultimately not interesting for them.

Example (good):
Navigation among the different details of a painting is difficult and disorientating, since links sometimes disappear and their logic is not to easy to understand".

With respect to previous (bad) example, this type of communication style allows the addressees to easily understand the usability problem.

It is fundamental to remember that according to whom the problem communication is addressed (designers, information architects, client, developers, web masters, project managers), proper lexicon should be used to convey the meaning of the problem.

3. Usability Report

The suggestions on how to structure the Usability report are very practical and could help in the creation of a professional and communication-effective document.

a. **Targets of a Usability Report**
   - Developers (engineers, Graphic Designers, Interface Designers…)
   - Product managers
   - Marketing Managers
   - Communication Managers
   - Directors
   - …

b. **Goals of a Usability Report**
   - Communicate usability problems (give an overview of the main problems, rate the emerged issues…);
   - Suggest the requirements for the improvement of the application

c. **Structure of the Usability Report**

   The report normally is structured in 7 main sections.
   - Cover
   - Executive Summary
   - Table of contents
   - Introduction
   - Results of Usability Analysis
   - Synoptic of usability problems
   - Requirements for improvement
   - At the end of the report you can also insert an **appendix (Annex)**, with all the material used and gathered during the evaluation.

Before explaining in depth each section it is important to underline that a usability report **should be as short as possible**. Several usability experts (Molich, Nielsen) fix the "optimal" length in 15-25 pages (include appendices).
It is clear that this assumption should not be prescriptive: indeed the size of the report depends on several factors, in particular:

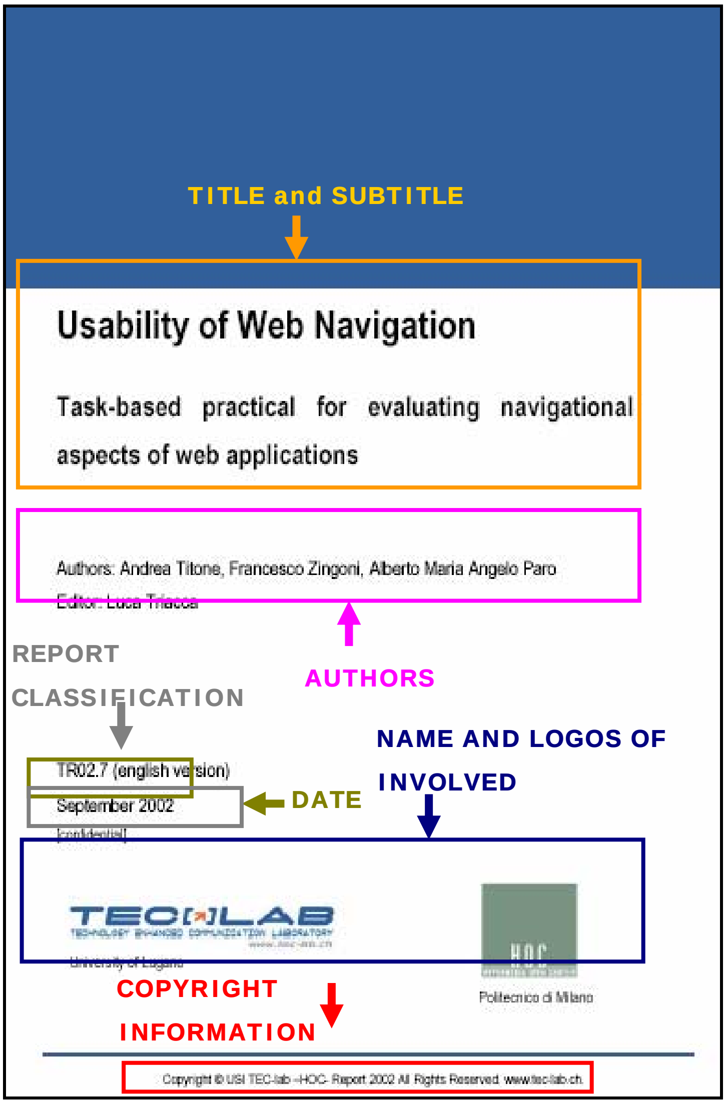
- *The number of problems discovered*: the more problems we find, the more the report becomes longer;

- *The dimension of the website:* the more the website is wide the more the analysis could be complex and articulated;
- *The number of techniques and methods used*: if the inspector(s) uses more then one techniques (for example he blends inspection and user testing) the report should present both results.

3.1 Cover

The cover at least includes at least:

- Title (and subtitle);
- Author(s)
- Name and logo(s) of the Institution(s) performing the evaluation
- Date: day, month and year
- Version of the document;
- Copyright information;
- Other information (e.g. URL of the institutions involved, Addresses, Report number…)



3.2 Executive Summary

Executive summaries are much like other summaries in that their main goal is to provide a condensed version of a longer report's content. The key difference, however, is that executive summaries are written for someone who most likely does not have time to read the original (www.columbia.edu/~ftg1/WRITING%20EXECUT.SUMMARY.html).

The Executive Summary plays a fundamental and strategic role for the communication of the main results of the usability evaluation. Indeed, **the Executive Summary has to report only the main issues emerged during the analysis.** See that, the maximum length of the Summary is 1 page, it is clear that the writing of this part of the report is a great effort in term of conciseness and efficacy: in one page you have to communicate the essence of the whole work.

It is important to remember that the addressees of the Summary are the main stakeholders (e.g. product managers, marketing managers, directors…): for this reason it is called Executive.

Example of Executive Summary from the OPENDRAMA prototype usability evaluation report:

0. Executive Summary

The methodology used for carrying out this usability evaluation is **MiLE** (Milano-Lugano Evaluation method), a new approach for evaluating the usability/quality of web applications. The evaluation of the application has been conducted basically on three design levels:

- **Navigation;**
- **Content;**
- **Graphic/Presentation.**

The major usability problems emerged during the evaluation are the following:

<u>Navigation level</u>

The main navigation problems are related to:

- **Guided-tour context visibility:** when the user starts a guided tour it happens that he loses the context within the tour is activated.

- **Label predictability of the Encyclopedia main menu:** some labels of the Encyclopedia's main menu are not predictable (in particular *More info* and *Now on stage*).

In this case, the user is browsing the operas "Now on stage", but there is not a reference to this context (the application proposes a general OPERAS reference).

<u>Content</u>

The main content problem is the **Structure of content in Opera Synopsis pages.** The *synopsis* page is to long and certain information is hided at the end of the page.

The only suggestion is to re-think the structure of the content or to insert a series of anchors for browsing easily the synopsis page.

<u>Graphic/Presentation</u>

The main problems related to this dimension are:

- **The general navigation bar that leads to the other applications of OPEN OPERA:** the position of this bar (at the top of the page) could create some problem for understanding that this is the tool for navigating to the other OPEN OPERA applications.
- **The main menu of Encyclopedia:** this menu seems "part of" OPEN OPERA and not the main menu of Encyclopedia.

3.3 Table of content

The Table of content should present the main sections and subsections of the document. It is important to keep in mind that the Table of content should be easy and quick to read and for this reason it does not report all the paragraphs; so it should presents only the main pointers. The table of content is not a detailed outline of the report.

The maximum length of the Table of content should be 1 page.

Examples:

Bad Example: this is a detailed outline

Table of Content

1. Executive Summary pag.2
2. Introduction pag.3
   2.1. Motivation pag.3
   2.2. Methodology pag.3
   2.3. Expected Results pag.4
3. Results of usability analysis pag.5
   3.1. Introduction pag.5
   3.2. Technical Inspection results pag.6
      3.2.1. Navigation Problems pag.7
      3.2.2. Content Problems pag.12
      3.2.3. Technology Problems pag.13
      3.2.4. Interface Design Problems pag.13
         3.2.4.1. Semiotics Problems pag.13
         3.2.4.2. Cognitive Problems pag.14
         3.2.4.3. Graphics Problems pag.15
   3.3. User Experience Inspection results pag.15
      3.3.1. Content Experience Problems pag.15
      3.3.2. Interaction Experience Problems pag.18
   3.4. Synoptic of results pag.22
4. Requirements for improvement pag.23
      4.1.1. Navigation requirements pag.23
      4.1.2. Content requirements pag.24
      4.1.3. Technology requirements pag.25
      4.1.4. Interface Design requirementspag.26
5. Conclusions pag.27
6. Annexes pag.28

Good Example: we use only the main pointers

Table of Content

1. Executive Summary pag.2
2. Introduction pag.3
3. Results of usability analysis pag.5
   3.1. Technical Inspection Results pag.6
   3.2. User Experience Results pag.15
   3.3. Synoptic of results pag.22
4. Requirements for improvement pag.23
5. Conclusions pag.27
6. Annexes pag.28

3.4 Introduction

The Introduction should report the goals of the work, the methodology and the tools used and (optional) a brief presentation of the application under analysis.
The maximum length of the Introduction should be 1-1,5 pages.

Example of Introduction from the BMW Learning Community Center (LCC) usability evaluation report:

1. Introduction

After a short reference section about the evaluation methodology, this report describes the problems and proposes suggestions for the improvement of the usability of the application. The scope of this report is therefore usability, i.e. the satisfaction in the user experience – in this case BMW workers as learners and BMW training managers as administrators; consequently, this reports does not include any consideration about the process of production and delivery of online or face to face courses, nor on the organizational impact of the application on the corporate environment.
Intended readers of this report are the application developers (at Alchera), interested in providing an effective, efficient and satisfactory experience to their users.
The main goal of the usability evaluation is to detect the most part of usability problems and breakdowns of a web application, being the usability "the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in particular environments" (ISO 9241-11).

The main goal of the usability evaluation of the **BMW Learning Community Center** is to provide some guidelines for improving the usability and quality of the application. The analysis is divided in two main levels:

- *Content level:* this level analyzes the quality of the content (in terms of communication effectiveness) in order to verify that the contents and their structure match the expectations of the users;
- *Navigation level:* within the navigational dimension of a web application there are two basic aspects: on one hand the different paths that can be used by a user to reach a specific piece of information service; on the other, the connections for passing from a content to another.

The Methodology
The methodology used for carrying out the usability evaluation is **MiLE** (Milano-Lugano Evaluation method), a new methodology for evaluating the usability/quality of web sites (or more in general hypermedia applications) fruit of a common research carried out by the Politecnico di Milano and the University of Lugano. It represents one of the most innovative and efficient approaches for evaluating the quality of a web application.
In fact, in order to actually support the evaluation process, MiLE aims at providing the inspector with a reusable set of evaluation tools (U-KIT, the usability evaluation kit) depending on the specific domain. The U-KIT is a library of macro-scenarios, refined in a number of scenarios, that help to understand *stories about use* (Rosson, M.B et. al.: 2002). Summarizing a U-KIT is composed by:

- Macro-scenarios (every macro-scenario refined in different scenarios);
- User profiles;
- User goals;
- Tasks;
- Attributes of quality (they are the "measure units" to define the quality of tasks).

The usability evaluation activity has been conducted in two different steps. In the first one the front-end of the application has been analysed (the part of the application used by end-users). The second part has evaluated the usability of the back-office (used by administrators and tutors).

3.5 Results of Usability Analysis

The part of the report reserved to Results is clearly one of the most important. In this section all the findings are reported, explained and commented. The Results' section could be divided in subsections or paragraph depending on the type of analysis.

For example if we use both inspection activity and user testing the Results section should be divided into "Results of inspection activity" and "Results of User testing". Moreover if we have carried out both a Technical Inspection and a User Experience Inspection and at the end the User testing, the section could be divided into:

X. Results of usability analysis
- **X.1 Inspection results**
  - **X.1.1 Technical Inspection results**
    - **X.1.1.1 Navigation Problems**
    - **X.1.1.2 Content Problems**
  - **X.1.2 User Experience Inspection results**
    - **X.1.2.1 Content Experience Problems**
- **X.2 User testing results**

Suggestions for improving the communication quality of the results:

1. **Say something nice in the opening paragraph (Tognazzini B, 2001):** Every report you ever write from this day until the day you die should start out by saying something nice about the product or service you are reviewing.
2. **Presenting Problems by dimension: one of the possible ways for reporting the problems is to organize them:**
   a. in design dimensions for the technical inspection (e.g. navigation, content, interface design…),
   b. in dimensions perceived by the user (e.g. interaction, cognitive) for the User Experience Inspection and User Testing.
3. **Use of screenshots:** for presenting efficacy the results of the evaluation activity it is important to use screenshots. Blending written comments and screenshots allows explaining to the readers the findings clearly and effectively and reduce the verbosity of the document.
4. **Insert positive findings:** even if the evaluation activity is centred on the usability problems, it is however important to insert some positive finding. This insertion allows reducing the "tension" between the writer (which "criticized" the

application) and the reader (which is involved in the development of the application).

5. **User testing Results - Distinguish among expert opinions, user opinions and user findings:** opinions may be acceptable within the results of user testing activity if they are clearly marked as such (Molich and Nielsen). Indeed, there is a great difference between the problems' interpretation and rationalisation made by the expert and opinions and findings of the users.
6. **User testing Results – Include quantitative data:** within the presentation of the user testing results it is important to avoid long lists of problems. It is also fundamental to include qualitative data like:
    - **a.** the number of people who successfully completed each task;
    - **b.** the number of people who experienced a particular problem;
    - **c.** a breakdown of problems by experience level.
7. **Express your annoyance tactfully (**Dumas, J., Molich, R., & Jeffries)**:** in describing the problems you should pay attention to the type of words you use. Be careful: do not offend and "attack" the work of others: stay always politically correct. For example, avoid expression like "this is a serious error", "this page is badly conceived"…

3.6 Synoptic of usability problems

The synoptic should present all the usability problems discovered. Indeed, it is the presentation of the issues at a glance. This general overview allows the readers to have a quick understanding of the number, type and gravity of the problems.

The strategy for showing the problems is fundamental, so it is recommendable to organize them **by dimension** (e.g. semiotics, navigation, content…), **by gravity**, **by needed effort** or **combining all these dimensions**.

Example of Synoptic Table for Technical Inspection combining dimension, gravity and needed effort.

Problem	Design Dimension	Gravity	Need Effort
Problem 1	Navigation	High	1 person/month
Problem 2	Navigation	Low	0.1 person/month
Problem 3	Content	Low	0.2 person/month
Problem 4	Semiotics	High	0.1 person/month
Problem 5	Technology	High	0.5 person/month
Problem N	…	…	…

3.7 Requirements for improvement

The last part of the Report is about the requirements for improvement. In this section the inspector has to provide **some positive and encouraging suggestion** on how the main problems could be addressed. The suggestions **should not be mandatory** at all. Indeed, the purpose of the usability evaluation is only to detect problems and usability breakdowns and not to redesign the application. Moreover, using a mandatory communication style could create a certain tension between the usability analyst and the stakeholders (in particular the designer of the application). They could interpret the suggestion as a critique of their work and the reaction could be a total refusal to collaborate.

Example:

Consider the following problem:

"The textual links are not visible at all. Indeed, the textual links have the same colour and style as the text".

Negative Example of Requirement for improvement:

For solving the problem of textual links visibility you have to choose an underlined and bolded style and a different colour (we suggest the blue). Moreover, it is necessary to change the type of font and the size (from Arial 12 to Verdana 10) to enhance scannability of the text.

Positive Example of Requirement for improvement:

For solving the problem of textual links visibility we suggest to choose a different chromatic code. For example as it is common in many websites you could use a blue font and underline it.

4. References

Bolchini D, Colazzo S., *Guidelines for Describing Usability Problems*, Proc. International Conference on Human-Computer Interaction(HCII), Las Vegas, 2005.

Bolchini, D., Triacca, L., & Speroni M., (2003). MiLE: a Reuse-Oriented Usability Evaluation Method for the Web. In *Proc. International Conference on Human-Computer Interaction HCII 2003*, Crete.

Dumas, J., Molich, R., & Jeffries, R. *Describing usability problems: are we sending the right message?*, Interactions, Volume 11, Issue 4 July + August 2004.

Molich R., Nielsen J., *230 Tips and Tricks for better Usability Testing*, NNGROUP Report - http://www.nngroup.com/reports/tips/usertest/ , California.

Sabiescu, A.G., *Usability Evaluation of the Rijksmuseum website*, Master TEC-CH Usability Report Project, USI. Website: http://www.rijksmuseum.nl/index.jsp.

Tognazzini B., *How to Deliver a Report Without Getting Lynched, www.asktog.com/columns/047HowToWriteAReport.html, ASKTOG Column 2001.*

ANNEX B:

Teaching Material –

Slides Packs

Summary:
In this Annex it is possible to find the teaching material used during the different MiLE+ courses. The material is divided in four packs:

- *Slides Pack 1: Introduction on usability;*
- *Slides Pack 2: MiLE+ - a systematic method for usability evaluation;*
- *Slides Pack 3: User testing (Scenario-based user testing);*
- *Slides Pack 4: Reporting usability.*

Slides Pack 1

Introduction on Usability

Usability Foundations: an Introduction

Usability within the lifecycle

- A interactive application is conceived
  - by a variety of *stakeholders*
  - for a variety of objectives
  - addressed to a variety of users
  - enabling them to accomplish a variety of goals
- Proper management of the lifecycle is key to the success of the application.

Usability within the lifecycle

- Requirements management
- Design
- Implementation
- Content management
- Evaluation
- Promotion & Change management
- Maintenance & Enhancement

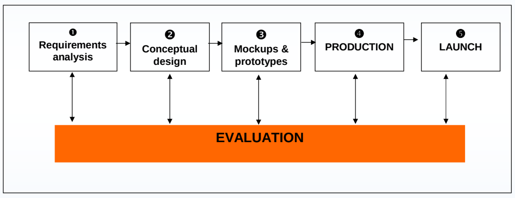
User Experience Lifecycle / 1

- Requirements management → **Understanding user needs motivations and goals**
- Design → **Shaping the user experience**
- Implementation → **Shaping the user experience**
- Content management → **Shaping the user experience**
- Evaluation → **Assessing the user experience**
- Promotion & Change management → **Creating awareness & motivations**
- Maintenance & Enhancement → **Instilling motivations to re-visit**

- SLIDES PACK 1 - 5

User Experience Lifecycle / 2

Pervasive Usability (Brink et al.:2002)



- SLIDES PACK 1 - 6

Our vision on Usability

- **To be successful, Interactive Applications** (being them available on web, palm, kiosks, or a combination of them) **have to be USABLE**.
- Usability is an essential factor for a successful user experience:
  - Being it learning, studying, deciding, visiting, enjoying or being entertained through an interactive application
  - An interactive application is usable if...
- Usability ROI are:
  - Lower customer calls for assistance
  - Brand equity enhancement
  - Trust building
  - . . .
- Example of poor usability:
  - www.lamborghini.com
  - www.giorgioarmani.com

- SLIDES PACK 1 - 7

Understanding Usability

- Application-independent usability aspects:
  - ...understandability....
  - ...navigation quality...
  - ...content accuracy
  - ...consistency
  - ...application status communication
  - ...graphic and layout quality
  - ...interface order....
  - ...compliance with standards and conventions...
  - ......accessibility
- These features can be evaluated even without knowing the purpose and the user of the application

- SLIDES PACK 1 - 8

Understanding Usability

- Application-dependent usability aspects:
  - Users can achieve their goals
    - People find the information they need . . .
    - People are properly driven and guided to unexpected content . .
    - Content is relevant to specific user profiles (kids, local tourists, tourists from abroad, families, curious, ...) . . .
    - Content is enjoyable/entertaining for specific user profiles..
    - ...
  - The application can be effectively used in a specific context (while driving, while at home, office, walking, visiting, ...)
- Understanding users, their goals and the contexts of use is essential to evaluate these features.

- SLIDES PACK 1 - 9

Understanding Usability

- From ISO 9241-11:
  - **Usability is "the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in particular environments".**
    - The effectiveness of application-independent aspects (font, layout, navigation, structure,...) is a necessary condition for usability
    - The effectiveness of application-dependent aspects (meeting user profiles, context, needs and goals) is a necessary condition for usabilty

- SLIDES PACK 1 - 10

Understanding Usability

- Usability evaluation should be done as early as possible in the development cycle:
  - The later errors/problems/flaws are discovered the more is expensive to fix them
  - Anticipate breakdowns and errors on design artifacts
- It should assess the effectiveness of "what is there" and not inquiry about "what is missing"

- SLIDES PACK 1 - 11

State of the Art

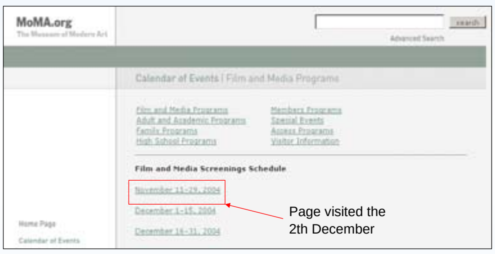
- Approaches to evaluate usability:
  - Who measures usability?
    - **User Testing**
      - Users are the measure of quality
    - **Inspection**
      - Expert review
  - How to measure it?
    - **Task-oriented** (scenario-based)
      - "Doing things" with the application
    - **Heuristics**
      - Verifying compliance with usability principles

- SLIDES PACK 1 - 12

Examples of
Application independent Usability Problems

Content

Currency of Information



www.moma.org/events/film/

Technology/Performance

Browser Compatibility

Explorer 6.0

Mozzilla Firefox 1.0

www.exploratorium.edu/listen/

- SLIDES PACK 1 - 17

Semiotics

- SLIDES PACK 1 - 18

Understandability of the main menu

- SLIDES PACK 1 - 19

Cognitive

- SLIDES PACK 1 - 20

Information Overload

www.metmuseum.org/Works_of_Art/collection.asp

Graphic Technical

Background Contrast & Font size

Are you able to read the different information on the screen?

www.moca.org/museum/visit_home.php

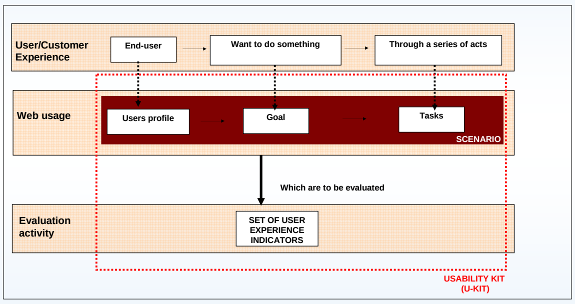
Navigational

Backward Navigation

www.guggenheimcollection.org/site/on_view_now.html

Examples of Application dependent Usability Problems

Concept of scenario



User Experience Inspection / 2

- How to create scenarios
  - Interaction with stakeholders and end-users
  - *Visioning* techniques
- UEIs: tools for evaluating scenarios
  - Three categories of UEIs (corresponding to the different types of user interaction experiences)
    - **Content Experience Indicators**
    - **Navigation & Cognitive Experience Indicators**
    - **Interaction Flow Experience Indicators**

Examples of Content Experience Indicators

SCENARIO	Well-educated American tourist who knows he will be in town, he wants visit the real museum on December 6th 2004 and therefore he/she would like to know what special exhibitions or activities of any kind (lectures, guided tours, concerts) will take place in that day.
USER PROFILE	Tourist
GOAL	Visit the M useum in a specific day
TASKS	· Find the events/exhibitions/lectures occurring on December 6th in the real museum · Find information about the museum's location

UEIs:

- Multilinguisticity
- Satisfaction on provided information

Multilinguisticity

I'm an American tourist.

It does not exist the English version of the current exhibition's description?

And the description of the collection?

www.men.ch/expositions.asp/1-3-583-99-21337-99-32-4-1/

Examples of Navigation & Cognitive Experience Indicators

SCENARIO	Marc looking for some information about Enlightenment period studying at school.
USER PROFILE	Marc, High-school student
GOAL	To be informed on a specific historical period (e.g. Enlightenment)
TASKS	· Find general information about this period; · Find detailed information about social and religious impact of Enlightenment period.

UEIs:

- Links Predictability
- Memorability of online tours

Links Predictability

What happens when I click on the button "Tours"?

And on the button "Take a online tour?"

Which is the difference?

www.thebritishmuseum.ac.uk/enlightenment/theageof.html

Examples of Interaction Flow Experience Indicators

SCENARIO	David looking for some information about special projects actually presented in the museum.
USER PROFILE	Contemporary Art lover
GOAL	To be informed on special projects
TASKS	· Find general information about special projects; · Find detailed information about a specific project (*The Uncertain Museum*) .

UEIs:

- Naturalness
- Engagement

- SLIDES PACK 1 - 33

Naturalness

How the navigation between objects works?

www.papesse.org/papesse/minisiti/invisibile/index.htm

- SLIDES PACK 1 - 34

Engagement

Once understand the interaction strategy of the website, this could be entertaining.

www.papesse.org/papesse/minisiti/invisibile/index.htm

- SLIDES PACK 1 - 35

MiLE+ Method: short introduction

- Developed in cooperation between HOC-Lab (Politecnico di Milano) and TEC-Lab (University of Lugano).
- MiLE+ (Milano-Lugano Evaluation) strikes a healthy balance between heuristic evaluation and task-driven techniques.
- Offers reusable tools and procedures to carry out both inspection and user testing within budget and time constraints.
- Two types of Inspection activities:
  - **Technical Inspection**
    - **For discovering application-independent problems**
  - **User Experience Inspection**
    - **For discovering application-dependent problems**

- SLIDES PACK 1 - 36

Slides Pack 2

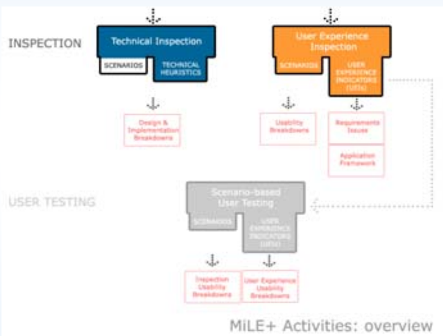
MiLE+: a systematic method for usability evaluation

Introduction and Features / 1

- Developed in cooperation between HOC-Lab (Politecnico di Milano) and TEC-Lab (University of Lugano).
- MiLE+ (Milano-Lugano Evaluation) strikes a healthy balance between heuristic evaluation and task-driven techniques.
- Offers reusable tools and procedures to carry out both inspection and user testing within budget and time constraints.
- Two types of Inspection activities:
  - **Technical Inspection**
    - **For discovering application-independent problems**
  - **User Experience Inspection**
    - **For discovering application-dependent problems**

- SLIDES PACK 2 -

MiLE+ Activities: Synoptic



- SLIDES PACK 2 -

- TECHNICAL INSPECTION

- SLIDES PACK 2 -

Technical Inspection / 1

- **Main goal:** identification of design problems and implementation breakdowns.
- The inspector evaluates the application from the design dimensions' perspective
  - **Content**
  - **Navigation**
  - **Technology**
  - **Interface Design**
    - **Semiotics**
    - **Cognitive**
    - **Graphics**

- SLIDES PACK 2 -

Technical Inspection / 2

- For each dimension we provide a **library of heuristics**:
  - **Usability Heuristics for:**
    - **Content**
    - **Navigation**
    - **Technology/Performance**
    - **Interface Design**
      - **Semiotics**
      - **Graphics**
      - **Cognitive**
- The evaluation using Heuristics can help by scenarios

  but

  We don't examine the adequacy of scenarios...

- SLIDES PACK 2 -

Content

- SLIDES PACK 2 -

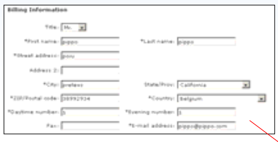
Text Conciseness

www.papesse.org/papesse/ita/programma/mostrescheda.cfm?id=127

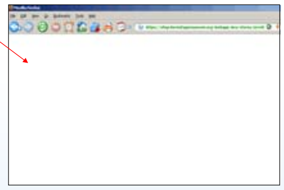
- SLIDES PACK 2 -

Technology/Performance

System Reaction to User's Error(s)



Which is the error?

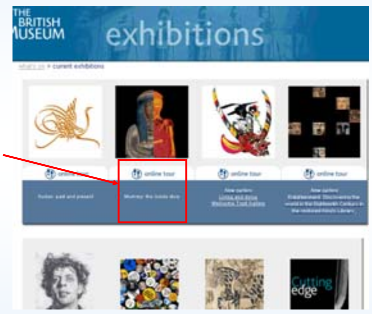


http://shop.hermitagemuseum.org/index.html

Semiotics

Ambiguity of labels

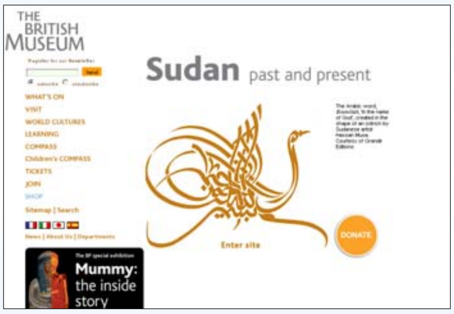
Two labels:
two different
websites



www.thebritishmuseum.ac.uk/whatson/exhibitions/index.html

Cognitive

Web site Mental Map



Navigate within the British Museum Website for 5 minutes.

After 5 minutes are you able to formalize the web site map?

www.thebritishmuseum.ac.uk/

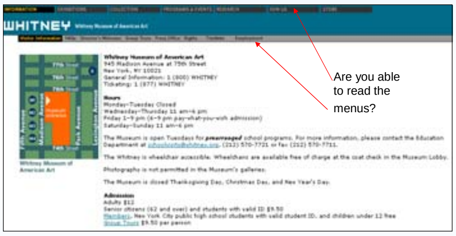
Graphic Technical

Anchor identity & Use of a Link's Cromatic Code

Which are links?

http://purchase.tickets.com/buy/TicketPurchase

Menu Font Size



Are you able to read the menus?

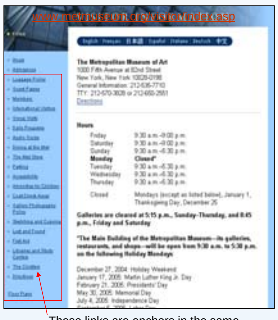
www.whitney.org/information/index.shtml

- SLIDES PACK 2 -

Navigational

- SLIDES PACK 2 -

Consistency within Sections' Navigation Strategy



Links for accessing subsections

www.metmuseum.org/store/index.asp

These links are anchors in the same page.

- SLIDES PACK 2 -

Example of Technical Evaluation Matrix

Dimension	Heuristic	Score	Comment
Content	Currency	3	The day of publication of the article is not highlighted. This does not allow the user to know if the article is updated or not.
	Text errors	9	The text does not present errors.
Navigation	Accessibility of different pages in the navigation within the topic "product"	9	All the pages of the topic Painting are very easy to access.
	Orientation in three navigation	6	Sometimes it happens that if the user passes from one section to another he does not find orientation clues.

- SLIDES PACK 2 -

USER EXPERIENCE INSPECTION

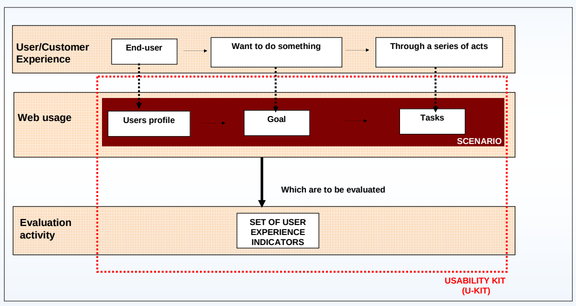
User Experience Inspection / 1

- **Main goal:** evaluating the scenarios (the inspector puts him-self in the shoes of the users)
  - With respect to Technical Inspection (where we evaluate problems application independent) in this inspection the inspector evaluates scenarios.
- The Inspector has to use:
  - **library of scenarios**
  - **library of User Experience Indicators (UEIs)**

User Experience Inspection / 2

- How to create scenarios
  - Interaction with stakeholders and end-users
  - *Visioning* techniques
- UEIs: tools for evaluating scenarios
  - Three categories of UEIs (corresponding to the different types of user interaction experiences)
    - **Content Experience Indicators**
    - **Navigation & Cognitive Experience Indicators**
    - **Interaction Flow Experience Indicators**

Concept of scenario



How to build scenarios / 1

- **Identify the user profiles (or user types): Who will use the application?**

User Type	
	User profiles
Learner	Responsible
	Accepter
	Seller
Instructor	Tutor (1..n)
	Administrator
	Responsible

- **Identify their high-level (or macro) goals: Why will they use the application?**

Macroscenario A	
User profile	**Seller**
Macrogoal	**Plan the learning experience**

User type (or profile) + macro-goal = MACRO-SCENARIO

How to build scenarios / 2

- **Refine the Macro-scenario into scenarios**

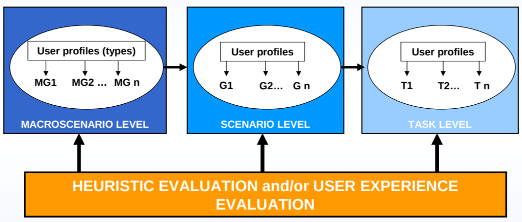
Macroscenario A	Plan the learning experience
Scenario 1	Plan the study
Scenario 2	Know course conditions
Scenario 3	Know the learning level achieved

User profile + Goal = SCENARIO

- **Refine Scenarios into user tasks**

Macroscenario A	Plan the learning experience	
Scenario 1	Goal	Tasks
	Plan the study	– Know the time required to frequent a course – Find the ideal period to frequent a classroom session – Know the time needed to download a document
Scenario 2	Know course conditions	– See course goals – See the course structure – See how to communicate with tutors and peers

Scenarios: different levels of granularity



- **Choose the evaluation level most appropriate to your project.**

Examples of Content Experience Indicators

SCENARIO	Well-educated American tourist who knows he will be in town, he wants visit the real museum on December 6th 2004 and therefore he/she would like to know what special exhibitions or activities of any kind (lectures, guided tours, concerts) will take place in that day.
USER PROFILE	Tourist
GOAL	Visit the M useum in a specific day
TASKS	· Find the events/exhibitions/lectures occurring on December 6th in the real museum · Find information about the museum's location

UEIs:

- **Multilinguisticity**
- **Satisfaction on provided information**

Multilinguisticity

I'm an American tourist.

It does not exist the English version of the current exhibition's description?

And the description of the collection?

www.men.ch/expositions.asp/1-3-583-99-21337-99-32-4-1/

Satisfaction on provided information

I don't have found information in English about the collection and the current exhibition. However I'm very interested in the MEN Museum and I want to visit it. Therefore I need road markings for reaching the museum. But also this information is given only in French !!!

Examples of Navigation & Cognitive Experience Indicators

SCENARIO	Marc looking for some information about Enlightenment period studying at school.
USER PROFILE	Marc, High-school student
GOAL	To be informed on a specific historical period (e.g. Enlightenment)
TASKS	- Find general information about this period; - Find detailed information about social and religious impact of Enlightenment period.

UEIs:

- **Links Predictability**
- **Memorability of online tours**

Links Predictability

What happens when I click on the button "Tours"?

And on the button "Take a online tour?"

Which is the difference?

www.thebritishmuseum.ac.uk/enlightenment/theageof.html

Memorability of online tours

When users return to the online tours of British Museum Websites after a period of not using it, they should be able to re-establish proficiency the past experiences of use?

www.thebritishmuseum.ac.uk/enlightenment/en_tours.htm

- SLIDES PACK 2 -

Examples of Interaction Flow Experience Indicators

SCENARIO	David looking for some information about special projects actually presented in the museum.
USER PROFILE	Contemporary Art lover
GOAL	To be informed on special projects
TASKS	· Find general information about special projects; · Find detailed information about a specific project (*The Uncertain Museum*) .

UEIs:

- **Links Predictability**
- **Memorability of online tours**

- SLIDES PACK 2 -

Naturalness

How the navigation between objects works?

www.papesse.org/papesse/minisiti/invisibile/index.htm

- SLIDES PACK 2 -

How to evaluate scenarios: the phases

1. **Performing the (selected) tasks**
   - assess the feasibility of some "critical" tasks
2. **Evaluating the tasks through User Experience Indicators (UEIs)**
   - For each attribute in relation to a specific task), the inspector has to give a score.
3. **Weighting the results according to user profiles and communication goals**
   - the inspector has to establishing the "real quality" of each critical task with respect to their relevance.

- SLIDES PACK 2 -

Example of evaluation (excerpts from the analysis of the Louvre web site)

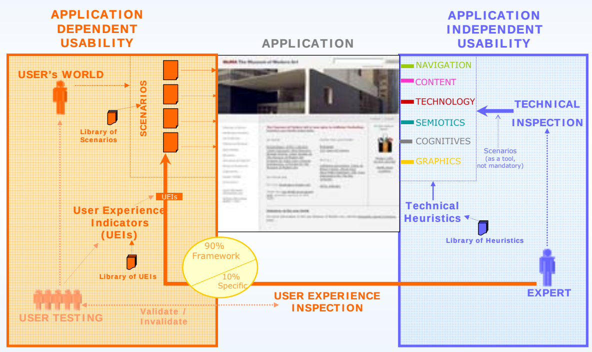
SCENARIO	
USER PROFILE	Art-lover
TASK	Find information about the history of museum collection
SCENARIO DESCRIPTION	Joe is an art-lover. He would like to find some information about the history of a particular collection of the museum (e.g. paintings). He wants to know how the museum has acquired some works of art.

Example of User Experience Matrix

	UEIs				
Task: Find information about the history of museum collection	Predictability	Understandability	Richness	Comprehensibility	**Global Score for this Task**
Scores	8	8	5	6	6.75 (just average score)
Weights	0.1	0.1	0.5	0.3	
Weighted Scores	0.8	0.8	2.5	1.8	5.9 (“weighted average”)

- SLIDES PACK 2 -

MiLE+ Activities Framework (in a picture)



- SLIDES PACK 2 -

Slides Pack 3

User Testing
(Scenario-based user testing)

Outline

- What is a usability test
- When to do a usability test
- How to analyze results
- How to communicate results
- Seven suggestions

Usability Test - 1

- A structured interview with a sample of the “potential” or “actual” target users of the application
- The interview is based on assigning a series of **tasks** to each user to perform on the application.
- Users are observed while they perform tasks and user sessions are recorded.
- Debriefing session with the user elaborate on the experience and investigate issues ex-post.

Usability Test - 2

- Tapes, notes from the observations are analyzed to state:
  - User successes
  - User problems
  - User mistakes
  - User opinions
  - User experience
- Results are compared, the most common issues are listed and illustrated.

Usability Test - 3

- It may be useful for:
  - Uncover problems overlooked by inspector
  - Validate inspection results
    - Confirm
    - Get them worse
  - Invalidate inspection results
  - Capturing the overall user perception of the application
  - Having "fresh eyes" using the application

When to do it

- User testing can be done on:
  - Low-fidelity prototypes
    - Input to rethink design
  - High-fidelity prototype
    - Input to adjust design
  - Partly implemented features
    - Input to fix errors and tune features
  - Running application
    - Input to redesign the next version
- Not once but periodically during development

How to do it

- Planning
- Recruiting users
- Select scenarios
- Elaborate tasks
- Define indicators
- Scripting
- Equipment
- Performance

Planning - 1

- Schedule the test in detail
  - Don't waste user's time
  - Don't waste your time
  - Write down material
  - Ask "why" for each of your choice
  - Discuss the choices within the group of experts
  - Test it before performance

Planning - 2

- Activities to plan
  - Determine user profiles to focus on
  - Start recruiting asap
  - Determine scenarios to focus on
  - Elaborate task
  - Write script to follow
  - Practice the test within the group of experts
  - Perform the test
  - Discuss and collect results
  - Combine results and report findings

- SLIDES PACK 3 -

Recruiting users - 1

- Recruiting the “right” users is crucial for the usefulness of the results
- Users should ideally correspond to user profiles
- Interview users with “screening” questions
  - To check his/her profiling
  - Specify the profile
  - Select tasks

- SLIDES PACK 3 -

Recruiting users - 2

- If you don’t need screening questions before the test, don’t do screening questions
- If yes, user should perceive that:
  - you care about her answers
  - Answers are “somehow” actively recorded (e.g. notes)
  - Answers will make “some” difference wrt the test
- Don’t make user feel she is wasting time and you are wasting time before going to the test.

- SLIDES PACK 3 -

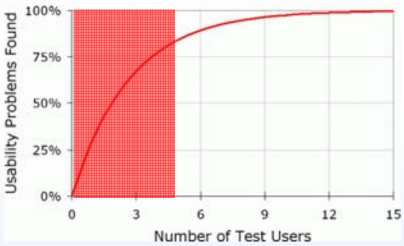
Recruiting users - 3

- Ideal user:
  - someone who is actually going to need the (type of) application in the near future or have use (a type of) it in the past
- Use *profiles* and user *goals may* orthogonally segment your target audience to recruit users.
  - Profiles: Young man, married woman, retired, etc..
  - Goals: Willing to pay bill online, to check account, to evaluate the possibility of buying new produtcs

- SLIDES PACK 3 -

Recruiting users - 4

- Nielsen's rule: 5-6 users (per scenario) are enough



Recruiting users - 5

- Recruit more people than you need:
  - Somebody will show up
  - Ensures minimal downtime in testing
- Users should be willing to use the application
- Don't feel users "to be tested": make them feel valuable for improving the application.

Recruiting users - 6

- Use an *incentive* strategy:
  - Money
  - Free results
  - Visibility
  - Learning
  - ...
- Incentive should become the motivation for accepting
- See material on "Recruiting"

Select scenarios

- Prioritize scenarios on the basis of:
  - **Gravity of usability problems detected during inspection**
  - Relevance to application's mission
  - Relevance to stakeholders
  - Most frequently used
  - Focusing on "new" features
  - Focusing on "highly publicized" features
  - . . .
- Select most important macro-scenarios (2-3) for the user test

Elaborate tasks - 1

- On the basis of the tasks used for inspection (within the selected scenarios), tasks should be:
  - Meaningful & Reasonable
  - Motivating
  - Goal-oriented
  - Not revealing tips
  - Supported
  - In a realistic sequence

- SLIDES PACK 3 -

Tasks should be: meaningful

- Tasks should be “typical” of the kinds of things people will do.
- Be reasonable
- Don’t propose “extreme situations” (difficult to motivate the user)
- NO:
  - Your name is “Y’~d,itr98” and you are buying 150 books each of different genres....
- YES:
  - Your name is <yourname> and you are looking for a couple of latest books about usability.

- SLIDES PACK 3 -

Tasks should be: motivating

- Try to give users good motivations for “visioning” a real situation of use.
- Make use of screening answers to motivate the user (asking things she understands/knows about)
- Users should not look and go for errors, but should try to complete the task
- NO
  - Buy 1 airline ticket for London next friday at 15 PM
- YES
  - Choose your favorite European destination. You’re planning to spend a relaxed weekend there and you prefer to leave next Friday afternoon. Find your flight and buy your e-ticket.

- SLIDES PACK 3 -

Tasks should be: goal-oriented

- Create motivation in terms of **goals** for the users.
- Goals should explain the users “WHY” she will do something
- Goals should be specific (not ambiguous: users should understand what to do) or ill-defined (open-ended: user should be very motivated)
- NO
  - Go buy some books
- YES
  - You want to improve your website. Find the latest book about “web design” and see if it’s worth buying.

- SLIDES PACK 3 -

Tasks should not reveal tips

- Tasks should not contain terms that exist on the interface: they tend to give hints to the user about how to perform the task.
- Be as neutral as possible (but clear!)
- NO:

Find a suitable preschool toy for your nephew

- YES:
  - Your nephew is 3 years old next Monday. For her birthday, you'd like to give her an original present she can have fun with it.

- SLIDES PACK 3 -

Tasks should be supported

- Select tasks that you know are doable and fully supported by the application
  - Undoable tasks:
    - will frustrate the user
    - will not be useful for testing
- "Doable" means
  - The information target of the task exists
  - The content treated in the task exists
  - The operation/transaction target of the task are feasible

- SLIDES PACK 3 -

Tasks should be in a realistic sequence

- Make the user feel the session flow as more realistic as possible
- Assign different tasks in a meaningful order
- To smooth the passage between tasks, insert transition dialogues or narratives not to loose motivation and engagement.
- NO
  - [Buy a toy] [Find airline ticket] [Apply for job position]
- YES
  - [Apply for job position]. *You also see there are discount on fares*. [Find airline ticket for London]. *Now you want to bring her a present.* [find a toy]

- SLIDES PACK 3 -

Tasks: some numbers

- How many tasks?
  - 5-6 tasks per user
- How long does it take?
  - 8-10 mins to execute a task
  - Ca. 1 hour per user

- SLIDES PACK 3 -

Define indicators

- What can be evaluated:
  - Qualitative:
    - Content Completeness
    - Richness
    - Engagement
    - Satisfaction
    - ... (see MiLE+ User Experience Indicators)
  - Quantitative:
    - How many people completed the tasks successfully
    - Completion speed
    - How many "errors" they make
    - "backtrack" moves

- SLIDES PACK 3 -

Scripting

- Write a guide containing:
  - what you are going to tell the user
    - To introduce him/her to the test
    - Explain test rules
    - Screening questions
    - To present tasks
  - what questions you are going to make: for debriefing after the test
    - elaborate questions on the basis of the qualitative indicators you want to investigate)

- SLIDES PACK 3 -

Equipment

- Tools needed for:
  - Recording user sessions (screen)
  - Recording user behavior (face and moves)
  - Recording user's voice
  - Recording expert's voice

- Camtasia (www.techsmith.com)

- Setting
  - different layouts for usability labs.

- SLIDES PACK 3 -

Performance – 1

- Different "styles" for observing users
  - Pure observation
    - Watch & take notes
    - Do not intervene in any moment
    - "Free" Thinking aloud
    - Debriefing
  - Insightful observation
    - Watch & take notes
    - Probe expectations
      - ***Before*** (what do you think it happens?) or ***after*** the action (Is that what you expected?)
      - Make "WHY" questions
      - Suggest tips (to avoid frustration and wasting time)
    - "Solicited" Thinking aloud & remind task

- SLIDES PACK 3 -

Performance – 2

- Meet the user after the test (10-15 minutes)
- Make him/her questions which are **useful** to you to understand:
  - Reasons for his/her behavior
  - User experience indicators
  - Let him/her freely express considerations

How to analyze results

- Collect:
  - Videos
  - Notes
- Watch over the video to note down the problematic situations.
- Relate notes to the video
- Elaborate notes
- Combine the results concerning quantitative and qualitative indicators

Seven Common User Testing Mistakes

Adapted from www.uei.com/articles/usability_testing_mistakes

- Mistake #1: Do you know why you're Testing?
  - User testing is a tool to produce information: it can produce all types of information.
  - You have to know what you want to get out of the test.
  - Remember: user testing is performing after the inspection activity → this is the starting point !

Seven Common Usability Testing Mistakes

- Mistake #2: Not bringing the team together
  - The Game of telephone
  - The team should be involved at every step.
    - Doing the test as near the team as possible (such as local conference room)
    - Giving incentives to participate (food always work ☺).
    - If the team can not attend a specific test. It should be easy for them to see video or get a detailed summary of what happened.

Seven Common Usability Testing Mistakes

- Mistake #3: Not recruiting the right participants
  - Do not concentrate on too general criteria (e.g. age, profession…);
  - The key question to ask yourself: “What attributes will cause one user to behave differently from another?”

- SLIDES PACK 3 -

Seven Common Usability Testing Mistakes

- Mistake #4: Not designing the right tasks
  - The way to you design tasks could have a dramatic outcome on the results.
  - Constantly exploring the “context of use”: “what events or conditions in the world would motivate someone to use this design?”

- SLIDES PACK 3 -

Seven Common Usability Testing Mistakes

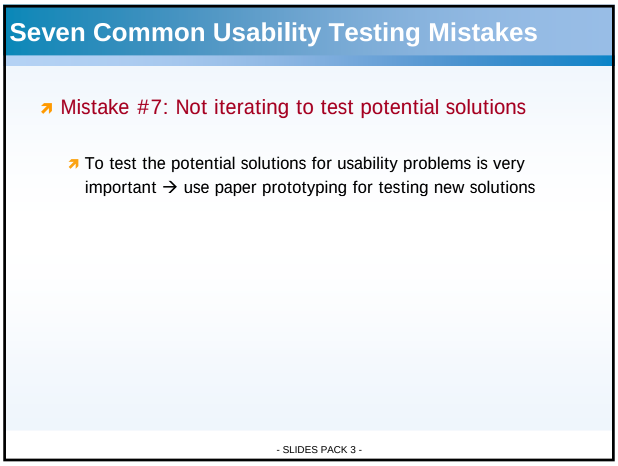
- Mistake #5: Not facilitating the test effectively
  - Conducting a user testing is a learned skill.
  - Goal: to obtain in a limited test time important information → focus on the elements which are important for your team

- SLIDES PACK 3 -

Seven Common Usability Testing Mistakes

- Mistake #6: Not planning how you’ll disseminate results
  - To communicate testing results to the team is fundamental.
    - Reporting results
    - Review sessions
    - …

- SLIDES PACK 3 -


Seven Common Usability Testing Mistakes
Mistake #7: Not iterating to test potential solutions
To test the potential solutions for usability problems is very important → use paper prototyping for testing new solutions
- SLIDES PACK 3 -

Slides Pack 4

Reporting Usability

Outline

1. **Guidelines for describing usability problems**
   - Balance Abstraction Levels
   - Extendibility: Represent Classes of Problems
   - Authority: Ground Your Findings
   - Prioritize and Set Importance
   - Technicality: Avoid "Usability" Jargon
2. **The structure of the Usability Report**
   - Cover
   - Executive Summary
   - Table of content
   - Introduction
   - Results of Usability Analysis
   - Synoptic of usability problems
   - Requirements for improvement

Introduction / 1

- **Goal of the usability report:**

  to communicate to people with an eterogenous background (e.g. marketing managers, engineers, interface designers, product managers...)

- **Key findings:** usability problems

Introduction / 2

- **Process of managing usability problems:**

Discover … *Identify issues*

Analyze … *Explore reasons/causes*

Characterize … *Describe your finding*

… *Filter, Decide, Prioritize, Validate….*

Communicate … *Convince about your finding*

Guidelines for describing usability problems

SEPARATE CONCERN

Guideline: *Decouple a usability problem into the various "application aspects" it is concerned with (content, navigation, semiotics, graphics, etc.)*

Guideline: *Distinguish between problems which are application-independent and problems depending on the purpose of the application (e.g. scenarios supported).*

- SLIDES PACK 4 - 5

Guidelines for describing usability problems

BALANCE ABSTRACTION LEVELS

Guideline: *Describe usability problems keeping a consistent level of granularity, moving from general to specific and without abruptly mixing details with strategic issues.*

- SLIDES PACK 4 - 6

Guidelines for describing usability problems

EXTENDIBILITY: REPRESENT CLASSES OF PROBLEMS

Guideline: *In case of complex applications, describe each problem as a representative of a class of problems, and characterize each problem by providing a general statement pointing to specific examples*

- SLIDES PACK 4 - 7

Guidelines for describing usability problems

AUTHORITY: GROUND YOUR FINDINGS

Guideline: *Give reason of your findings by drawing to elements which can gain credibility, such as the experience of the analysts, the impact on the user experience and the compliance with the standard and convention*

- SLIDES PACK 4 - 8

Guidelines for describing usability problems

PRIORITIZE AND SET IMPORTANCE

Guideline: *Communicate your findings in order of importance: according to the circumstances, importance may mean gravity for the user experience or estimated effort needed to fix the problem.*

- SLIDES PACK 4 - 9

Guidelines for describing usability problems

TECHNICALITY: AVOID "USABILITY" JARGON

Guideline: *Adapt your concepts and wording to the target audience you are communicating with, in such a way that stakeholders clearly understand the essence of the problems.*

- SLIDES PACK 4 - 10

Usability Report – Introduction / 1

- Targets:
  - Developers (engineers, Graphic Designers, Interface Designers...)
  - Product managers
  - Marketing Managers
  - Communication Managers
  - Directors
- Goals:
  - Communicate usability problems (give an overview of the main problems, rate the emerged issues...);
  - Suggest the requirements for the improvement of the application

- SLIDES PACK 4 - 11

Usability Report – Introduction / 2

- Structure:
  - Cover
  - Executive Summary
  - Table of contents
  - Introduction
  - Results of Usability Analysis
  - Synoptic of usability problems
  - Requirements for improvement

- SLIDES PACK 4 - 12

Usability Report

- Cover:
  - Title (and subtitle);
  - Author(s)
  - Name and logo(s) of the Institution(s) performing the evaluation
  - Date: day, month and year
  - Version of the document;
  - Copyright information;
  - Other information (e.g. URL of the institutions involved, Addresses, Report number...)

- SLIDES PACK 4 - 13

Usability Report

- Executive Summary
  - It is addressed to people that does not have time to read the entire report
  - Plays a fundamental and strategic role for the communication of the main results of the usability evaluation

- SLIDES PACK 4 - 14

Usability Report

- Table of content
  - Reports only the pointers to the main sections
  - Is not a detailed outline
- Introduction
  - Reports the goals of the work, the methodology and the tools used

- SLIDES PACK 4 - 15

Usability Report

- Results of usability analysis / 1
  - Is the heart of the report: this section describes, comments and explains the findings
  - The structure of this section depends on the type of analysis

- SLIDES PACK 4 - 16

Usability Report

- Results of usability analysis / 2
  - **Suggestion for improving the communication quality of the results**
    - Say something nice in the opening paragraph (Tognazzini B, 2001)
    - Presenting Problems by dimension
    - Use of screenshots
    - Insert positive findings
    - User testing Results - Distinguish among expert opinions, user opinions and user findings
    - User testing Results – Include quantitative data
    - Express your annoyance tactfully (Dumas, J., Molich, R., & Jeffries):

- SLIDES PACK 4 - 17

Usability Report

- Synoptic of usability problems
  - it is the presentation of the issues at a glance.
  - Organize it by dimension (e.g. semiotics, navigation, content...), by gravity, by needed effort or combining all these dimensions.
- Requirements for improvement
  - Are suggestions not mandatory at all.

- SLIDES PACK 4 - 18

ANNEX C:

Detailed Experiment Results

Summary:
Within this Annex it is possible to find the detailed results of the validation experiment presented in Chapter 5.

Inspectors / Inspectors	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13
C1		0,50	0,80	0,25	0,33	0,50	0,25	0,57	0,80	0,44	0,33	0,33	0,50	0,38	0,33	0,25	0,67	0,80	0,38	0,50	0,50	0,33	0,57	0,43	0,67	0,38	0,22	0,25	0,22
C2			0,57	0,57	0,67	0,44	0,57	0,71	0,83	0,56	0,43	0,43	0,38	0,33	0,44	0,22	0,50	0,83	0,71	0,83	0,43	0,67	0,63	0,50	0,71	0,71	0,50	0,38	0,50
C3				0,38	0,50	0,63	0,38	0,71	0,67	0,56	0,50	0,50	0,67	0,50	0,44	0,38	0,83	0,67	0,50	0,67	0,43	0,50	0,71	0,57	0,83	0,50	0,33	0,38	0,33
C4					0,83	0,56	0,71	0,63	0,50	0,67	0,83	0,57	0,50	0,63	0,56	0,33	0,44	0,50	0,86	0,71	0,38	0,57	0,56	0,44	0,63	0,86	0,63	0,71	0,63
C5						0,33	0,43	0,57	0,50	0,44	0,60	0,60	0,50	0,38	0,33	0,11	0,43	0,50	0,57	0,80	0,13	0,60	0,57	0,43	0,67	0,57	0,38	0,43	0,38
C6							0,56	0,67	0,56	0,89	0,63	0,63	0,75	0,88	0,78	0,75	0,88	0,56	0,67	0,56	0,63	0,44	0,78	0,67	0,67	0,67	0,67	0,75	0,67
C7								0,63	0,50	0,67	0,57	0,38	0,33	0,44	0,75	0,50	0,44	0,50	0,86	0,71	0,57	0,83	0,56	0,63	0,63	0,86	0,86	0,50	0,86
C8									0,86	0,78	0,50	0,50	0,63	0,56	0,67	0,44	0,75	0,86	0,75	0,86	0,50	0,71	0,88	0,75	1,00	0,75	0,56	0,44	0,56
C9										0,56	0,29	0,29	0,43	0,33	0,44	0,22	0,57	1,00	0,50	0,67	0,43	0,50	0,71	0,57	0,83	0,50	0,33	0,22	0,33
C10											0,56	0,56	0,67	0,78	0,89	0,67	0,78	0,67	0,78	0,67	0,56	0,56	0,89	0,78	0,78	0,78	0,78	0,67	0,78
C11												0,33	0,50	0,38	0,33	0,25	0,67	0,80	0,38	0,50	0,50	0,33	0,57	0,43	0,67	0,38	0,22	0,25	0,22
C12													0,80	0,57	0,33	0,25	0,67	0,29	0,38	0,50	0,13	0,33	0,57	0,43	0,43	0,38	0,38	0,67	0,38
C13														0,71	0,44	0,38	0,83	0,43	0,33	0,43	0,25	0,29	0,71	0,57	0,57	0,33	0,33	0,57	0,33
C14															0,67	0,63	0,75	0,44	0,56	0,44	0,50	0,33	0,67	0,56	0,56	0,56	0,56	0,86	0,56
C15																0,75	0,67	0,56	0,67	0,56	0,63	0,63	0,78	0,88	0,67	0,67	0,88	0,56	0,88
C16																	0,63	0,33	0,44	0,33	0,57	0,38	0,56	0,63	0,44	0,44	0,63	0,50	0,63
L1																		0,57	0,44	0,57	0,38	0,43	0,86	0,71	0,71	0,44	0,44	0,50	0,44
L2																			0,50	0,67	0,43	0,50	0,71	0,57	0,83	0,50	0,33	0,22	0,33
L3																				0,86	0,50	0,71	0,67	0,56	0,75	1,00	0,75	0,63	0,75
L4																					0,25	0,80	0,71	0,57	0,83	0,71	0,50	0,38	0,50
L5																						0,43	0,44	0,50	0,50	0,50	0,50	0,38	0,50
L6																							0,57	0,67	0,67	0,57	0,57	0,25	0,57
L7																								0,86	0,86	0,56	0,56	0,44	0,56
L8																									0,71	0,44	0,63	0,33	0,63
L9																										0,63	0,44	0,33	0,44
L10																											0,75	0,63	0,75
L11																												0,63	1,00
L12																													0,63
L13																													
	0,44	0,51	0,54	0,57	0,48	0,46	0,51	0,63	0,52	0,64	0,47	0,47	0,55	0,51	0,54	0,41	0,54	0,51	0,68	0,61	0,44	0,56	0,65	0,60	0,64	0,62	0,59	0,44	0,59
																	0,53	0,58	0,60	0,70	0,46	0,56	0,62	0,67	0,46	0,69	0,70	0,55	0,70

Como

Lugano

Global average: 0,56

av-Lugan	0,58
mile2	0,6
mile4	0,58
mile8	0,62

av-Como	0,53

ANNEX C.1:

Inter-group agreement on findings among Lugano and Como groups

and

Intra group agreement of Lugano sub-groups on the common problems

Inspectors / Inspectors	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
C1		0,56	0,47	0,33	0,39	0,50	0,28	0,44	0,39	0,28	0,28	0,33	0,39	0,39	0,44	0,28
C2			0,63	0,50	0,33	0,61	0,33	0,44	0,56	0,61	0,22	0,22	0,28	0,39	0,39	0,17
C3				0,42	0,37	0,58	0,26	0,32	0,53	0,47	0,32	0,26	0,42	0,37	0,42	0,26
C4					0,53	0,53	0,40	0,33	0,40	0,50	0,33	0,33	0,20	0,47	0,47	0,27
C5						0,41	0,36	0,36	0,50	0,38	0,43	0,36	0,21	0,33	0,20	0,14
C6							0,53	0,35	0,53	0,71	0,35	0,41	0,35	0,53	0,53	0,41
C7								0,50	0,38	0,44	0,25	0,33	0,23	0,40	0,53	0,46
C8									0,38	0,50	0,33	0,33	0,46	0,40	0,47	0,38
C9										0,50	0,31	0,31	0,31	0,27	0,33	0,23
C10											0,38	0,31	0,38	0,56	0,50	0,38
C11												0,36	0,38	0,33	0,20	0,23
C12													0,38	0,33	0,40	0,31
C13														0,60	0,33	0,38
C14															0,47	0,47
C15																0,53
C16																

average	0,39

ANNEX C.2:

Intra-group agreement on findings within Como group

Inspectors \ Inspectors	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13
L1		0,42	0,67	0,67	0,42	0,33	0,58	0,58	0,42	0,58	0,42	0,50	0,50
L2			0,42	0,45	0,36	0,50	0,50	0,33	0,50	0,56	0,33	0,30	0,36
L3				0,67	0,50	0,42	0,50	0,42	0,42	0,67	0,58	0,58	0,67
L4					0,45	0,45	0,64	0,58	0,55	0,55	0,55	0,55	0,73
L5						0,36	0,36	0,42	0,36	0,55	0,45	0,55	0,64
L6							0,40	0,33	0,40	0,44	0,44	0,30	0,45
L7								0,58	0,80	0,60	0,60	0,60	0,55
L8									0,50	0,33	0,58	0,42	0,67
L9										0,50	0,50	0,50	0,45
L10											0,56	0,60	0,55
L11												0,70	0,82
L12													0,73
L13													
	0,58	0,43	0,53	0,62	0,38	0,37	0,45	0,44	0,49	0,55	0,64	0,63	0,64

Global average:	0,53

	average:
MiLE+2h	0,54
MiLE+4h	0,41
MiLE+8h	0,59

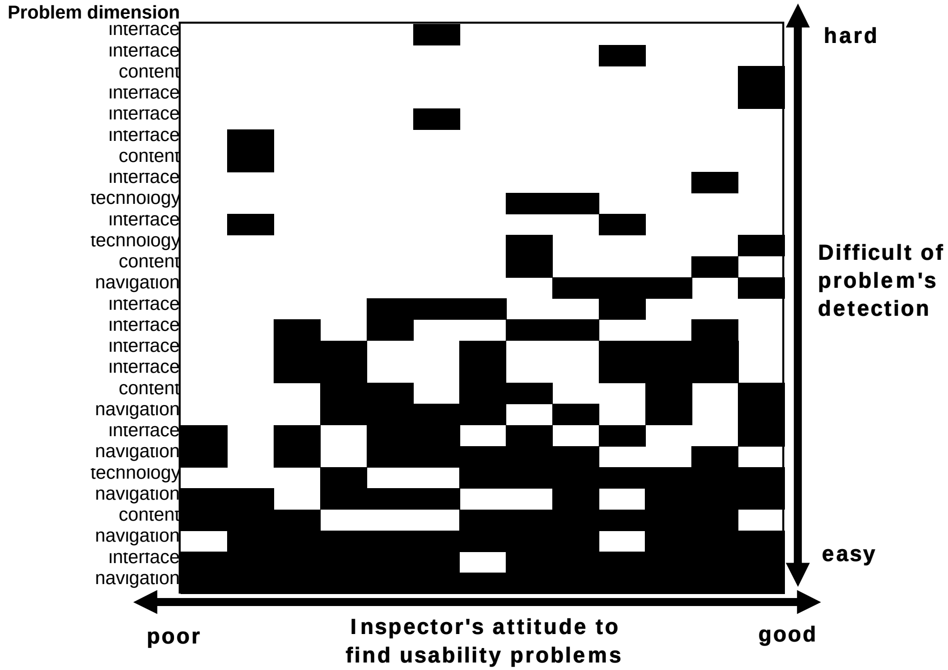
ANNEX C.3:

Intra-group agreement on findings within Lugano group and its subgroups

id_problem	Problem	L2	L6	L10	L11	L7	L9	L12	L1	L4	L5	L13	L3	L8		
3	Elements' classification within hihlights menu in the home page						1								1	interface
15	Accuracy of objects' lists										1				1	interface
17	Comprensibility of objects' technical description (e.g. 1920, 1980)													1	1	content
18	Predictability of links Gallery - 241/2 (recurrent problem)													1	1	interface
26	Ambiguity of some labels for the service' navigation (e.g.Search help e More						1								1	interface
33	Link's predictability for operational activities on the objects		1												1	interface
44	Content currency		1												1	content
58	Predictability del link "provenance"												1		1	interface
16	Link which does not work (decorative images)								1	1					2	technology
24	Text font contrast (serch, help...)		1								1				2	interface
51	Browser compatibility								1					1	2	technology
57	Satisfaction on provided information								1				1		2	content
4	Website mental map									1	1	1		1	4	Navigation
54	Font size too small					1	1	1			1				4	interface
6	Redundance & overlapping icons representing the collections			1		1			1	1			1		5	interface
1	Intuitiveness and information scent of images of the home'page main menu			1	1			1			1	1	1		6	interface
2	Convenzionality of home page's menu			1	1			1			1	1	1		6	interface
9	Completeness of collections' lists				1	1		1	1			1		1	6	content
10	Orientation clues collection menu				1	1	1	1		1		1		1	7	navigation
41	Layout quality	1		1		1	1		1		1			1	7	interface
5	Grouping adequacy & ambiguity of collection main menu	1		1		1	1	1	1	1			1		8	navigation
28	System does not react to the user' s error in filling the forms (e.g. Advanced				1			1	1	1	1	1	1	1	8	technology
25	Accessibility of kind of topics (cardinality of list elements is too high	1	1		1	1	1			1		1	1	1	9	navigation
56	Text conciseness	1	1	1				1	1	1	1	1	1		9	content
8	Topology of objects lists		1	1	1	1	1	1	1	1		1	1	1	11	navigation
43	No font cromatic code	1	1	1	1	1	1		1	1	1	1	1	1	12	interface
13	Backward navigation	1	1	1	1	1	1	1	1	1	1	1	1	1	13	navigation
	TOTAL PROBLEMS	6	8	9	9	10	10	10	12	11	11	11	12	12		

ANNEX C.4:

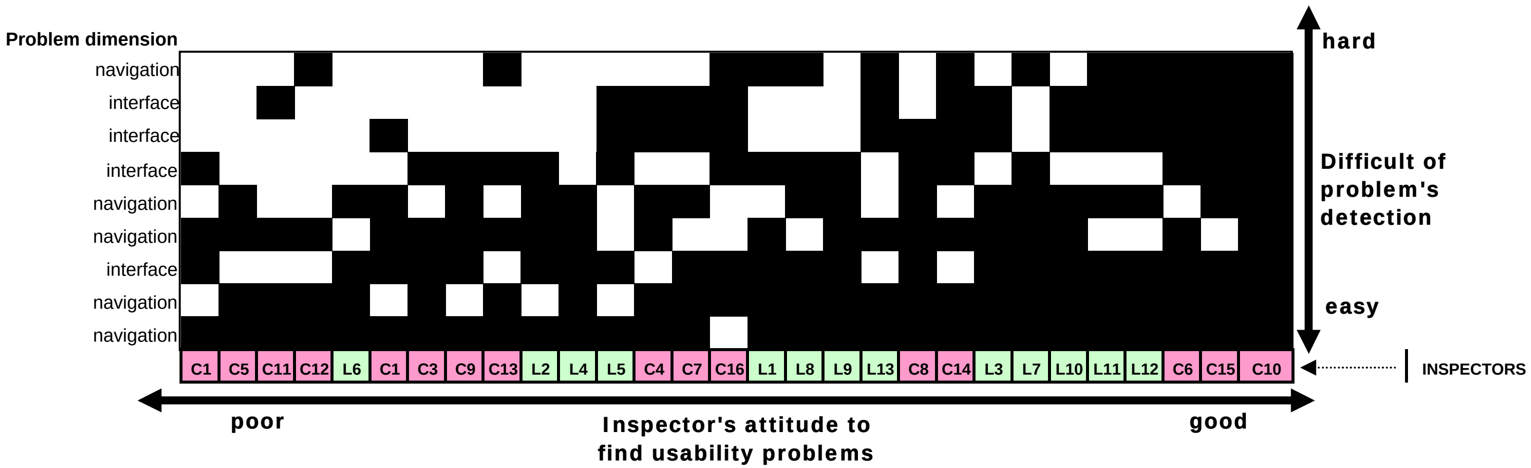
Lugano Group: distribution of problems' detection



ANNEX C.4:

Lugano Group: distribution of problems' detection

id_problem	Problem	C1	C5	C11	C12	L6	C1	C3	C9	C13	L2	L4	L5	C4	C7	C16	L1	L8	L9	L13	C8	C14	L3	L7	L10	L11	L12	C6	C15	C10	
9	Completeness of collections' lists				1					1						1	1	1		1		1		1		1	1	1	1	1	navigation
2	Convenzionality of home page's menu			1									1	1	1	1				1		1	1		1	1	1	1	1	1	interface
1	Intuitiveness and information scent of images of the home'page main menu						1						1	1	1	1				1	1	1	1		1	1	1	1	1	1	interface
41	Layout quality	1						1	1	1	1		1			1	1	1	1		1	1		1				1	1	1	interface
25	Accessibility of kind of topics (cardinality of list elements is too high)		1			1	1		1		1	1		1	1			1	1		1		1	1	1	1	1		1	1	navigation
5	Grouping adequacy & ambiguity of collection main menu	1	1	1	1		1	1	1	1	1	1		1			1		1	1	1	1	1	1	1			1		1	navigation
43	No font cromatic code	1				1	1	1	1		1	1	1		1	1	1	1	1		1		1	1	1	1	1	1	1	1	interface
8	Topology of objects lists		1	1	1	1		1		1		1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	navigation
13	Backward navigation	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	navigation
		4	4	4	4	4	5	5	5	5	5	5	5	6	6	6	6	6	6	6	7	7	7	7	7	7	7	8	8	9	



ANNEX C.5:

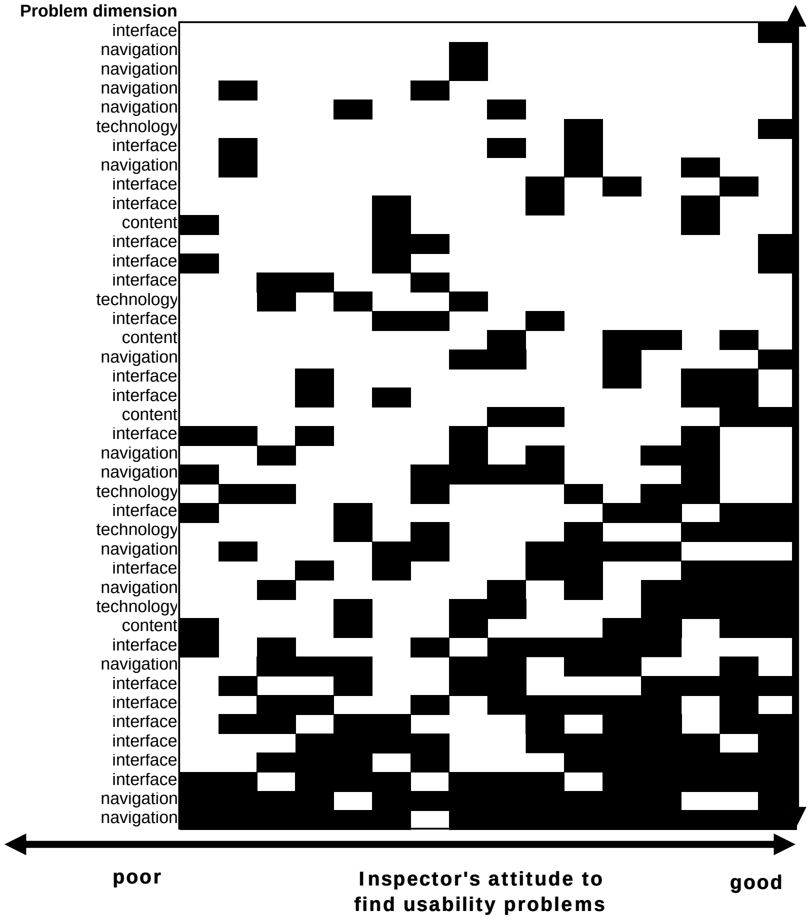
Lugano and Como Groups: distribution of problems' detection considering the common problems

ANNEX C.5:

Lugano and Como Groups: distribution of problems' detection considering the common problems

id_problem	Tipo Problema	11	12	7	8	9	13	16	5	4	14	15	10	Isp6	1	2	3
15	accuratezza liste di oggetti (in alcuni casi non è presentata la foto)																1
30	segmentation del k.of topic Oggetto d'arte								1								
31	Controllo della navigazione ad indice nelle istanze dei k.of.topics								1								
4	website mental map		1					1									
7	Navigation strategy (guided-tour) tra gruppi di oggetti delle collezioni					1				1							
27	efficacia del motore "advanced search" soprattutto per utenti inesperti											1					1
3	Classificazione degli elementi del menu highlights in homepage (a sinistra)		1							1		1					
14	Classification adequacy nelle liste dei group of topics		1									1			1		
26	Ambiguità di alcune etichette die landamrks "secondari" (es.Search help e More info)										1		1			1	
37	predictability del target dei link click here e register as a new user (puntano tutti e due alla stessa pagina!)						1				1				1		
44	Content currency	1					1								1		
47	Visibilità del link Contact us						1	1									1
49	Predictability di see also	1					1										1
50	Icone non chiare del menu delle opere			1	1			1									
51	Browser compatibility			1		1			1								
54	font size troppo piccolo						1	1			1						
17	comprensibilità descrizione "tecnica" degli oggetti (es. 1920.1980)									1			1	1		1	
29	Backward navigation per ritornare all'oggetto una volta mandata l'ecard.								1	1			1				1
40	Sovraccarico informativo della pagina principale della sezione collection				1								1		1	1	
42	Position importance of macroareas				1		1								1	1	
48	Aggiornamento contenut									1	1					1	1
33	predictability dei link per fare "operazioni" sugli oggetti	1	1		1				1						1		
38	problema navigazionale se si usa il link visualizzato dall'opzione big image dell'oggetto			1					1		1			1	1		
10	Orientation clues del menu collection (tra le collezioni → si perde il riferimento)	1						1	1	1	1				1		
16	link non funzionante (immagine decorativa in alto a destra)		1	1				1				1		1	1		
18	Predictability del link Gallery - 241/2 (problema ricorrente	1				1							1	1		1	1
28	sistema non reagisce all'errore dell'utente nel riempire le varie form (ad es. Advanced search (es.access number: se si usano caratteri e non numeri, la form non avvisa dell'errore)					1		1				1			1	1	1
9	Completezza delle liste delle collezioni		1				1	1			1	1	1	1			
24	font contrast del testo (in calce, testo di search help…)				1		1				1	1			1	1	1
32	information overload delle liste di oggetti dispalyed as list			1						1		1		1	1	1	1
35	gestione dell'operazione di mantenimento dell'oggetto inserito nella personal collection					1			1	1				1	1	1	1
46	Completezza delle info su alcuni 'artisti e opere	1				1			1				1	1		1	1
2	Convenzionalità menu della home page	1		1				1		1	1	1	1	1			
25	Accessibilità dei kind of topics impedita dalla cardinalità troppo alta degli elenchi			1	1	1			1	1		1	1			1	
34	learnability del meccanismo di inserimento dell'oggetto della personal collection (difficoltà nell'uso del meccanismo)		1			1			1	1				1	1	1	1
1	Intuitività e information scent delle immagini del menu della home page			1	1			1		1	1	1	1	1		1	
6	Ridondanza & overlapping delle icone rappresentanti le collections		1	1		1	1				1		1	1		1	1
41	Layout quality				1	1	1	1			1	1	1	1	1		1
43	No font cromatic code			1	1	1		1				1	1	1	1	1	1
5	Grouping adequacy & ambiguity of Collection main menu	1	1		1	1	1		1	1	1		1	1	1	1	1
8	Topologia delle liste di oggetti delle varie collezioni	1	1	1	1		1	1	1	1	1	1	1	1			1
13	Backward navigation	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1
	TOTALE TIPI PROBLEMI	**10**	**11**	**12**	**12**	**13**	**13**	**13**	**14**	**15**	**15**	**15**	**16**	**17**	**18**	**18**	**19**

ANNEX C.6:
Como Group: distribution of problems' detection



ANNEX C.6:
Como Group: distribution of problems' detection