



# **Organization Model for usability-testing in ITRACT.**

"We intend to develop and test innovative tools for efficient, user- and environment-friendly transport networks across the NSR." [1]

This is one of the main sentences about the goals of ITRACT. In WP5 the goal is defined by:

"The aim of WP5 is to test and evaluate the newly developed solutions for sustainable, userfriendly transport management."

One of the first steps in ITRACT is to create an organization model for testing these new and user-friendly applications.

In *Gablers Wirtschaftslexikon* user-friendliness is defined as an attribute of software-quality. It is the character of a software-product especially of its interface and dialog-system which has to be adjusted to the user-requirements. [2]

User-friendliness software means software that is easy to use or which is "usable". So software-usability is one of the key topics in WP5.

#### Usability:

Usability is a well studied field that leads to the DIN EN ISO 9241 standard. Part 110 describes seven dialog principles.

These seven principles are [3]:

- suitability for the task (the dialogue should be suitable for the user's task and skill level);
- self-descriptiveness (the dialogue should make clear what the user should do next);
- controllability (the user should be able to control the pace and sequence of the interaction);
- conformity with user expectations (it should be consistent);
- error tolerance (the dialogue should be forgiving);
- suitability for individualization (the dialogue should be able to be customized to suit the user);
- suitability for learning (the dialogue should support learning).

Ben Shneiderman also made researches on that field and formulated eight golden rules [4]:

- strive for consistency,
- enable frequent users to use shortcuts,
- offer informative feedback,
- design dialogs to yield closure,
- offer error prevention and simple error handling,
- permit easy reversal of actions,
- support internal locus of control,





reduce short-term memory load,

Jacob Nielsen found ten heuristics [5]:

- visibility of system status,
- match between system and the real world,
- user control and freedom,
- consistency and standards,
- error prevention,
- recognition rather than recall,
- flexibility and efficiency of use,
- aesthetic and minimalistic design,
- help users recognize, diagnose, and recover from errors,
- help and documentation.

We combined all this principles, rules and heuristics and reduced them to a checklist developers should recognize while working out the applications. We present some methods which show, how the checkpoints can be evaluated.

Some methods must be used before the implementation begins. Others go along during the implementation and some can be done in a last step of the implementation.

### Prearrangements:

#### Personas:

In ITRACT the transport companies defined some personas. These personas are typical users of the transport system in the belonged region.

#### **Definition of target groups:**

For every application target groups should be defined. The main question is: Who shall use the application? It is not compulsory that the application reaches all defined personas.

## **Test- Methods for conception and implementation phases**

#### Use cases

A use case is a description of how users will perform tasks on your application. They are sequences of actions that the system can perform while interacting with the actor. Actors can be described by personas.

This method is a method that should be used before the implementation starts.

Each use case should capture following questions:





- Who is using the Website? => given by personas and target groups.
- What does the user want to do?
- What is the user's goal?

Use cases can be written in an easy-to-understand narrative. This makes it understandable for all engaged project members [6].

Edward Kenworthy [7] outlines eight steps to develop use cases:

- 1. Identify who is going to be using the Website.
- 2. Pick one of those actors.
- 3. Define what that actor wants to do on the site. Each thing the actor does on the site becomes a use case.
- 4. For each use case, decide on the normal course of events when that actor is using the site.
- 5. Describe the basic course in the description for the use case. Describe it in terms of what the actor does and what the system does in response that the actor should be aware of.
- 6. When the basic course is described, consider alternate courses of events and add those to "extend" the use case.
- 7. Look for commonalities among the use cases. Extract these and note them as common course use cases.
- 8. Repeat the steps 2 through 7 for all other actors.

#### Card Sorting

Card Sorting is a helpful method to design and evaluate the structure of the application, the navigation and the wording used by the application. A detailed process is given in "Card sorting: a definitive guide" by Spencer and Warfel [7].

- 1. Divide the content and the structure / navigation in singular information units.
- 2. Write the information units on cards.
- 3. Find out the proband expectations by questions like:
  - a. What content do you expect under the navigation term....?
  - b. Which term would you expect for content about...?
- 4. In a next step ask the proband to sort the cards by similarity. So you can find out the possible structure of the application.

Card Sorting is possible as an open or a closed sort.

- Open Sort: Users are asked to sort items into a group and make up their own groups and give them a name.
- Closed Sort: Users sort items into previously defined category names.

#### Cognitive Walkthrough

This method proves the suitability of learning. Usability experts put themselves in the position of the user and "walk through" the application. By this method the typical user-problems can





be identified. But it must be said that the cognitive walkthrough appears to detect far more potential problems than actually exist [9].

The cognitive walkthrough is a time reducing and low cost method because it is not necessary to find a couple of test persons. This method should be used several times during the implementation process.

### **General Test-Criteria**

General test criteria are various, but most of them can be done during the realization of the application. These tests should be repeated in fixed time intervals. Diverse literature describes many different tests [11],[12],[13],[14]. The most important tests that are easy to handle are:

- Look after the right spelling of the text and error messages.
- Pay attention to good error messages. They should be relevant, helpful, informative, clear, easy to understand, truthful and complete [15].
- Investigate the error rate.
- When forms must be filled out, the logic of the order and clarity of fields should be reviewed, so that wrong inputs can be avoided.
- Test the reaction time of the application.

Within these tests, smaller problems can be solved directly. Further these tests are simple with only slightly costs.

## **Test Methods with participants**

For the following methods participants should be engaged. It is necessary to consider that the participant should be persons of the specified target groups.

It is important that participants from all target groups are involved.

Jacob Nielsen describes that 80% of the problems can be revealed by only five participants [16].

### Focus-Groups

In ITRACT the main target groups are elderly people and pupils. This circumstance has been revealed by the definition of the personas. The main problem of the target groups in ITRACT could be the contradictions between the target groups. The method "focus groups" is a good possibility to detect these contradictions. Normally its goal is to collect ideas, understand the





5

### Usability tests with participants

Usability testing is a technique to evaluate the applications by testing it with representative users. In the test, users will try to complete typical tasks while observers (developers and business experts) watch, listen and take notes.

The goals are

- to evaluate if participants are able to complete identified routine tasks successfully and how long it takes to do that,
- to find out how satisfied participants are by using your application,
- to identify changes required to improve user performance,

Following points should be considered:

- Let the participants try to complete typical tasks.
- The tasks should be embedded in a context that provides useful information to users.
- Ask the participants to think out loud.
- Test the application, not the participants.
- Keep notes of the behavior and thoughts of the participants.

#### **Eye-tracking**

Eye-tracking is an improved usability test [17]. With an eye-tracking tool the order of the observation of objects in the application can be determined. Also the intensity of the observation of singular objects can be measured.

By eye-tracking it is possible to get information about the subconscious perception and information processing.

Mainly following questions can be answered:

- What elements of my site are perceived by users and which are completely overlooked?
- Are navigation elements recognized as such?
- What texts are read and which are only scanned?
- Will users guide effectively to the content that is relevant to them?
- How fast decides a user to use a navigation point?
- How fast recognizes the user important information?





## Test-Methods for the pilot phase

## A/B testing and multivariate testing

While A/B testing will test different content for one visual element on a page, multivariate testing will test different content for many elements across one or more pages to identify the combination of changes that yields the best result.

Multivariate testing is often used after publishing an application [18].

Every variant should be supported by hypotheses. Otherwise the number of variants is too large to evaluate them all.

Multivariate testing can find the optimized appearance of:

- Headings: Try different text, size, color.
- Images: Try different sizes, different images, different positions on a page.
- Buttons: Try different positions on a page, different sizes, colors, labels on the buttons.
- Forms: try different length of fields, different fieldnames, different order of fields.
- Especially for websites: try different background colors, different sizes of headlines, positions of logos, position of login, search fields, navigationbars.

The use of software like Google Website Optimizer (freeware) or similar tools is advised.

### Surveys

Surveys can be very different. From multiple choice questions up to scaling systems or open text answers - everything is possible. To create a questionnaire or opinionaire is a complex task.

For fast and essential testing it may be adequate to use standardized questionnaires like the System Usability Scale (SUS) or the Computer System Usability Questionnaire (CSUQ).

The SUS, developed by Brooke [19], reflects a strong need in the usability community for a tool that could quickly and easily collect a user's subjective rating of a product's usability. Brooke named the SUS a quick and dirty method, but it is an often used and accepted usability test method [20].

Ten questions have to be answered by a couple of users during the pilot phase.

1. I think that I would like to use this system frequently





- 2. I found the system unnecessarily complex
- 3. I thought the system was easy to use
- 4. I think that I would need the support of a technical person to be able to use this system
- 5. I found the various functions in this system were well integrated
- 6. I thought there was too much inconsistency in this system
- 7. I would imagine that most people would learn to use this system very quickly
- 8. I found the system very cumbersome to use
- 9. I felt very confident using the system
- 10. I needed to learn a lot of things before I could get going with this system

Every question can be answered on a scale from 1 to 5 points "I strongly disagree" up to "I strongly agree".

#### Scoring:

For odd items: subtract one from the user response.

- For even-numbered items: subtract the user responses from 5
- This scales all values from 0 to 4 (with four being the most positive response).
- Add up the converted responses for each user and multiply that total by 2.5. This converts the range of possible values from 0 to 100 instead of from 0 to 40.

#### Results:

- 100 Points correspond to a perfect System without any usability problems.
- Values greater the 80 points correspond to a good usability.
- Values between 60 and 80 points are satisfactory.
- Values lower than 60 indicate significant problems.

The CSUQ developed by Lewis [21] is a questionnaire with 19 questions and a scale of seven points to answer [22].

The SUS or CSUQ questionnaire could be implemented in the pilot applications. An environment for the analysis must be worked out by the developers.

## Closing Methods and certification

### DAkks (Deutsche Akkreditierungsstelle GmbH)

The DAkks is a national accreditation agency which develops standardized procedures for usability tests. The procedures are based on the international standard DIN EN ISO 9241. It







contains well defined different steps. The guidelines are trackable at the homepages of DAkks [10].

A certification by DAkks would be very valuably for an application, but nevertheless a certification by DAkks may be charged.

## The usability testing procedure in ITRACT

The planned applications are very different in functionality and they also run under different operation systems and hardware infrastructure. In addition, the applications will be developed in various locations throughout Europe.

As seen above, the testing is not a one-time process, but a frequently repeated, accompanying process.

Most of the usability tests can easily be done by the developers. The checklist attached to this document supports the developers.

Nevertheless, it is reasonable to check the new application by an eye-tracking tool. The Jade Hochschule owns an eye-tracking system and would like to test up to ten different applications.





The European Regional Devi

## The Checklist

Criteria	<b>Evaluation method</b>	State of work
Effectiveness		
Identify users goals	<ul> <li>Target groups</li> <li>Personas</li> <li>Use cases / scenarios</li> <li>Focus groups</li> </ul>	undone in process done
Provide precise information and extensive help	<ul> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Usability tests</li> <li>Eyetracking</li> <li>Surveys</li> </ul>	undone in process done
Create a good information structure	<ul><li>Card sorting</li><li>DAkks test method</li></ul>	undone in process done
Offer useful and constructive functions	<ul> <li>Target groups</li> <li>Personas</li> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Surveys</li> </ul>	undone in process done
Efficiency		
Perform a task analysis	<ul> <li>Target groups</li> <li>Personas</li> <li>Use cases / scenarios</li> <li>Focus groups</li> <li>Surveys</li> </ul>	undone in process done
Reduce workload	<ul> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> </ul>	undone in process done
Offer effective functions	<ul> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability-Tests</li> <li>Surveys</li> </ul>	undone in process done

9

TRACT Provide the second secon	The Interreg IVB North Sea Region Programme	European Union
Criteria	Evaluation method	State of work
Guarantee orientation	<ul> <li>Card sorting</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Eyetracking</li> </ul>	undone in process done
The most important first	<ul> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Eyetracking</li> </ul>	undone in process done
Appropriateness of tasks		
Seclusion of dialogues	<ul> <li>Cognitive walkthrough</li> <li>DAkks test method</li> <li>General test criteria</li> <li>Usability tests</li> </ul>	undone in process done
Offer a self-contained user-interface	<ul> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Eyetracking</li> <li>Multivariate tests</li> </ul>	undone in process done
Definition of terms	<ul> <li>Card sorting</li> <li>Cognitive walkthrough</li> <li>Web analysis</li> <li>General test criteria</li> <li>Focus groups</li> <li>Usability tests</li> <li>Eyetracking</li> <li>Multivariate tests</li> </ul>	undone in process done
Guarantee adequate response time for each target group	<ul> <li>Target groups</li> <li>Personas</li> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Surveys</li> </ul>	undone in process done
Give feedback	<ul> <li>Cognitive walkthrough</li> <li>DAkks test method</li> <li>General test criteria</li> <li>Usability tests</li> </ul>	undone in process done









11

Criteria	Evaluation method	State of work
Confirmation		
Give feedback for every step	<ul> <li>Cognitive walkthrough</li> <li>DAkks test method</li> <li>General test criteria</li> <li>Usability tests</li> </ul>	undone in process done
Provide clear feedback	<ul> <li>Cognitive walkthrough</li> <li>Usability tests</li> <li>Multivariate tests</li> <li>Surveys</li> </ul>	undone in process done
Adapt type and extend of a feedback to the task	<ul> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Multivariate tests</li> <li>Surveys</li> </ul>	undone in process done
Give personal feedback	<ul> <li>Personas</li> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Multivariate tests</li> <li>Surveys</li> </ul>	undone in process done
Give acoustic or visual feedback	<ul> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Eyetracking</li> <li>Multivariate tests</li> <li>Surveys</li> </ul>	undone in process done
Controllability	-	
Set up control functions	<ul> <li>Personas</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> </ul>	undone in process done









The European Regional Devi

Criteria	<b>Evaluation method</b>	St	ate of work	
Offer emergency exits	<ul> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>DAkks test method</li> <li>Usability tests</li> <li>Eyetracking</li> </ul>	undone	in process	done
Support explorative learning	<ul> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Eyetracking</li> <li>Multivariate tests</li> </ul>	undone	in process	done
Suggestibility of speed	<ul> <li>Personas</li> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Usability tests</li> </ul>	undone	in process	done
Opportunity to choose between different work equipment	<ul> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> </ul>	undone	in process	done
Support experienced users	<ul> <li>Personas</li> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Usability tests</li> </ul>	undone	in process	done
Consistency				
Consistency to provide fixed rules and certainty	<ul> <li>Target groups</li> <li>Personas</li> <li>Card sorting</li> <li>Cognitive walkthrough</li> <li>DAkks test method</li> <li>General test criteria</li> <li>Focus groups</li> <li>Usability tests</li> </ul>	undone	in process	done









Criteria	Evaluation method	State of work
Provide expectation compliant information structure	<ul> <li>Personas</li> <li>Use cases / scenarios</li> <li>Card sorting</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Eyetracking</li> </ul>	undone in process done
Mind design standards and conventions	<ul> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Eyetracking</li> <li>Multivariate tests</li> </ul>	undone in process done
Consistency and conformity with user expectations of terms	<ul> <li>Card sorting</li> <li>Cognitive walkthrough</li> <li>General test criteria</li> <li>Usability tests</li> <li>Eyetracking</li> <li>Multivariate tests</li> </ul>	undone in process done
Predictable performance of tasks	<ul> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>DAkks test method</li> <li>Usability tests</li> <li>Eyetracking</li> </ul>	undone in process done
Design of a complex and detailed style guide	Focus groups	undone in process done
Fault talaranas		
Perfect error-prone functions for the target group to avoid mistakes	<ul> <li>Target groups</li> <li>Personas</li> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> </ul>	undone in process done
Permit minimal correction work	<ul> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> </ul>	undone in process done





European	Union	
		_

The European Regional Devi

Criteria	Evaluation method	State of work
Give constructive error messages	<ul> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>General test criteria</li> <li>Usability tests</li> <li>Multivariate tests</li> </ul>	undone in process done
Expectation compliant design of errors	<ul> <li>Personas</li> <li>Cognitive walkthrough</li> <li>Usability tests</li> <li>Eyetracking</li> <li>Multivariate tests</li> </ul>	undone in process done
Customizability Offer individual and relevant information	<ul> <li>Target groups</li> <li>Personas</li> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability test</li> <li>Eyetracking</li> <li>Surveys</li> </ul>	undone in process done
Application adaptable to users characteristics	<ul> <li>Personas</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> </ul>	undone in process done
Application adaptable to previous knowledge	<ul> <li>Personas</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Eyetracking</li> <li>Multivariate tests</li> <li>Surveys</li> </ul>	undone in process done
Offer conventional shortcuts	<ul> <li>Personas</li> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>Focus groups</li> <li>Usability tests</li> <li>Eyetracking</li> </ul>	undone in process done



TTRACT Windowski and Alexandre	The Interreg IVB North Sea Region Programme	European Union
Criteria	<b>Evaluation method</b>	State of work
Provide minimalist design and relevant information	<ul> <li>Cognitive walkthrough</li> <li>Usability tests</li> <li>Eyetracking</li> <li>Multivariate tests</li> </ul>	undone in process done
Use concise language	<ul> <li>Card sorting</li> <li>Cognitive walkthrough</li> <li>General test criteria</li> <li>Usability tests</li> <li>Eyetracking</li> <li>Multivariate tests</li> </ul>	undone in process done
Aesthetics		
Collaboration of designers, users and developers	<ul><li>Personas</li><li>Focus groups</li></ul>	undone in process done
Mind the laws of perception	<ul> <li>Cognitive walkthrough</li> <li>General test criteria</li> <li>Eyetracking</li> <li>Multivariate tests</li> </ul>	undone in process done
Create pleasant color spaces	<ul> <li>Cognitive walkthrough</li> <li>Eyetracking</li> <li>Multivariate tests</li> </ul>	undone in process done
Mind the laws of typography	<ul> <li>Cognitive walkthrough</li> <li>General test criteria</li> <li>Eyetracking</li> <li>Multivariate tests</li> </ul>	undone in process done
Consider different display devices	<ul> <li>Personas</li> <li>Use cases / scenarios</li> <li>Cognitive walkthrough</li> <li>DAkks test method</li> <li>General test criteria</li> <li>Focus groups</li> <li>Eyetracking</li> <li>Surveys</li> </ul>	undone in process done





17

## References

[1] Homepage of ITRACT; *www.itract.eu;* (Retrieved Jan. 2013).

[2] Gabler Wirtschaftslexikon: *Definition;* 

http://wirtschaftslexikon.gabler.de/Archiv/75615/benutzerfreundlichkeit-v5.html; (Retrieved Feb. 2013).

[3] International Organization for Standardization: *DIN EN ISO 9241 Part 110;* http://www.iso.org; (Retrieved Feb.2013).

[4] Shneiderman, B.: Designing the user interface: Strategies for effective human-computer interaction (3rd ed.); Addison-Wesley Publishing.; (1998).

[5] Nielsen, J.: *Heuristic evaluation.* In Nielsen, J., and Mack, R.L. (Eds.), *Usability Inspection Method;* John Wiley & Sons; (1994).

[6] Carroll, J.M.; Rosson, M.B.: Usability Engineering. Scenario-Based Development of Human-Computer Interaction.; Morgan Kaufmann. (2001).

[7] Kenworthy, E.: *Use case modeling: Capturing user requirements.;* http://www.zoo.co.uk/~z0001039/PracGuides/pg\_use\_cases.htm; (1997).

[8] Spencer, D. and Warfel, T.; *Card Sorting: a definitive guide*; http://boxesandarrows.com/card-sorting-a-definitive-guide/; (2004).

[9] Wharton, C., Rieman, J., Lewis, C., and Polson, P.; *The cognitive walkthrough method: A practitioner's guide.* In Nielsen, J., and Mack, R. (Eds.), *Usability inspection methods.;* John Wiley & Sons, Inc.; (1994).

[10] Homepage DAkks: Leitfaden Usability; http://www.dakks.de; (1994).

[11] Courage, C. & Baxter, K.: Understanding Your Users: A Practical Guide to User Requirements Methods, Tools, and Techniques.; Morgan Kaufmann. (2005).

[12] Barnum, C. M.: Usability testing essentials; Elsevier Inc.; (2011).

[13] Albers, M., Still, B. (Eds.): Usability of complex information systems; CRC Press; (2011).

[14] Tullis, T., Albert, B.: *Measurement the user experience;* Elsevier/Morgan Kaufmann; (2008).

[15] Grice, H. P.: Logic and Conversation. In Martinich, A.P. (Ed).: *Philosophy of Language*. Oxford University Press; (1975).







[16] Nielsen, J.: *Website;* http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/; (Retrieved Feb. 2013).

[17] Nielsen, J., Pernice, K.: Eyetracking web usability; New Riders; (2010).

[18] Eberhard-Yom, M.: Usability als Erfolgsfaktor; Cornelsen Verlag; (2010).

[19] Brooke, J.; "SUS: a "quick and dirty" usability scale". In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland: Usability Evaluation in Industry.; Taylor and Francis; (1996).

[20] Sauro, J.: *Measuring Usability with the System Usability Scale (SUS);* http://www.measuringusability.com/sus.php; (2009).

[21] Lewis, J. R.: *IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use.* International Journal of Human-Computer Interaction, 7:1, 57-78.; (1995).

[22] Human-Computer Interaction Resources: *The Computer System Usability Questionnaire;* http://hcibib.org/perlman/question.cgi; (Retrieved Feb.2013).