EXPERT TESTING REPORT

519141-LLP-1-2011-1-ES-KA3-KA3MP

1. INTRODUCTION

This report summarizes the methodology, execution and results of the expert testing carried out in TRAILER project. The expert testing is defined before the pilots in order to assure the proper work of the workflow and system components and also to detect possible issues in an early stage and before real users interact with the system. Therefore this testing is applied in order to guarantee the proper work of the system.

The tested systems are the ILC, Portfolio, institutional environment and its integration. For the test and to gather informal learning activities the bookmarklet is used and for future testing the game, the widgets will be included.

The document is structured as follows. In the second section the methodology applied is presented, taking into account the different type of tests carried out. After that, in section 3, how the methodology is applied is described. Later on (Section 4) the results are shown, and finally some conclusions are added.

2. METHODOLOGY

During the project different usability tests are needed. The idea in this stage is to test the concept and how it is possible to put it into practice through the implemented systems. This means that the main focus will not be the specific functionalities or interfaces but how they facilitate the process.

Taking this into account the first stage will be an expert test in which all the partners will be involved. From each partner at least two persons (if possible with different familiarity with the framework) should test the system and for the testing different activities will be carried out to measure different usability issues.

The methodology proposed is based in inspecting techniques, direct and indirect techniques.

The inspecting technique used is the Cognitive Walkthrough - CW (Polson, Lewis, Rieman, & Wharton, 1992). It is a evaluation technique in which expert reviewer define scenarios to complete the tasks in an early prototype. Once the scenario is defined the reviewer act as a final user working with the system looking for errors that can affect to the final goal of the system. They should act as if the interface were complete and taking into account each step that the user should carry out (Sidar, 2009). With this methodology it is possible to check the concept by using a prototype from the experience of the expert, and with the data recovered re-design the system to solve possible errors. To support these techniques programs to record the screen and voice will be used.

The CW results will be compared with a direct technique known as Think Aloud - TA (Lewis, 1982). The primary concern is to support the development of usable systems by identifying design deficiencies (Boren & Ramey, 2000). Think Aloud protocols involve participants thinking aloud as they are performing a set of specified tasks. In this Project it will be supported with programs that record screen and voice.

In addition, some indirect some indirect techniques such as surveys are used to gather the users perception of the system and a System Usability Scale (SUS) form (Brooke, 1996) to know the final user satisfaction; and also to gather the perception of ease of use (PEOU) by following a Venkatesh and Bala application of TAM3 (Venkatesh & Bala, 2008).

Moreover the expert testing implies several tasks. For each of them a survey was defined, and through them it is possible to gather the experts opinion about the components. Such information was evaluated by using qualitative techniques [26].

3. APPLICATION

In order to carry out the expert testing, what to test and a set of specific tasks that should be carried out by the experts should be defined.

The idea was to validate the methodology through the system and find possible design errors, that is, "to tag, classify, organize and evaluate the informal learning activities that a user carried out and that can be taken into account from the institution".

To do so there were three available components by now: the ILC, the Portfolio and the Institutional Environment. These three elements were integrated to facilitate the activity and they were tested individually and all together. This was done through a set of tasks related to each component. After they were completed the experts should answer a survey, fulfil a SUS and PEOU tests. The task were very simple because in this way it was easier to figure out the problems that a user can find in an specific part of the component.

For the ILC the tasks proposed were:

- T1.1. Gather an informal activity with the bookmarklet, tag and classify it and send it to the Portfolio.
- T1.2. Add in the ILC an informal activity manually, tag and classify it and send it to the Portfolio.

For the Portfolio the tasks proposed are:

- T2.1. Complete the information of an ILA that has an associated competence publishing the ILA to the institutional context. The competence extra information should be also completed.
- T2.2. Complete the information of an ILA and associate to it an existing competence. It is necessary to publish the ILA to the institutional context. The competence extra information should be also completed.
- T2.3. Complete the information of an ILA and associate to it a new user competence. It is necessary to publish the ILA to the institutional context. The competence extra information should be also completed.
- T2.4. Define a showcase with ILA and profile information and share it with others.

The information to gather was the same than for the previous.

To test the institutional system an installation for each institution will be done. The proposed tasks are:

- T3.1.Create an institutional competence.
- T3.2. Update the institutional competence.
- T3.3. Validate a user-defined competence.

From these items it was needed to gather: the steps taken to carry this out and a voice description of them, the time employed to complete the task, the breakdowns found and the mistakes experts commit till they complete the activity. In addition the experts can describe the things they miss and would improve in the ILC, the Portfolio and the Institutional Environment. In addition experts' opinion about the whole system was requested.

The surveys for each task can be find at:

- ILC:
  - T1.1 https://docs.google.com/spreadsheet/viewform?formkey=dGdXLUhPUnhhWi1oUWh0a3RvS3pSWUE6MQ
  - T1.2 https://docs.google.com/spreadsheet/viewform?formkey=dDl6LV9rZm9YZk9iSk5tOVlqdGtWQUE6MA
- Portfolio:
  - T2.1 https://docs.google.com/spreadsheet/viewform?formkey=dEk4NTdTRXdycnd1UFVNRDZodGtrN1E6MA
  - T2.2 https://docs.google.com/spreadsheet/viewform?formkey=dE9EdUt5X2xXaHV4NnhuWU15UnhUTnc6MA
  - T2.3 https://docs.google.com/spreadsheet/viewform?formkey=dEpzbU8ySlRFWU9paW9HTmx6VXUydmc6MA
  - T2.4 https://docs.google.com/spreadsheet/viewform?formkey=dC1RYnprQk9xOGFhODBpQWZPMkJqLWc6MA
- Institutional Environment:
  - T3.1 https://docs.google.com/spreadsheet/viewform?formkey=dEtrdUJTTEZINE9WdFRDSk9nelFrOVE6MA
  - T3.2 https://docs.google.com/spreadsheet/viewform?formkey=dDZPd0tDSjNRNm5VbzFBTjlGblRqNnc6MA
  - T3.3 https://docs.google.com/spreadsheet/viewform?formkey=dGxDdFB3VmFmaTdBeV9UV05NZnU0Tmc6MA

- Satisfaction, Perception of Ease of Use and Opinion test: https://docs.google.com/spreadsheet/viewform?formkey=dGhfeTRkSlZIbmNjT1VjdnZnSGRkU0E6MQ

4. RESULTS

From the surveys and the recordings different errors were found. Amongst the challenges of analysing the data produced by this method, distinguishing particular types of errors and determining the severity of those errors was the priority. In terms of the former, patterns of recurrent 'breakdowns' (Winograd and Flores, 1986) in practice were identified through an analysis of the videos. The degree of severity of breakdown was determined by the extent to which the breakdown caused a disruption to the flow through the CW. In effect, what is produced is a probability distribution of breakdown moments in the experience of the expert user-testing group.

Taking into account the tasks to do and the people involved in the experiment (14 experts on the problem domain). The 91,7% of the tasks could be completed. From them by using the CW and the TA application 66 moments of breakdown were identified. The indication of severity level followed the Nielsen classification (Nielsen, 1994).

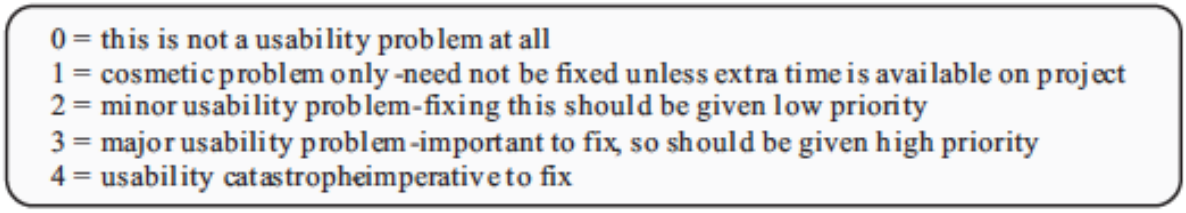

Figure 1. – Nielsen's severity rating for usability problems

Below is a list of the issues detected per each component. To distinguish the severity of the issue different colours have been used. 0 severity level in green, 1 in blue, 2 in grey, 3 in red and 4 in black.

ILC

- Authorization is not clear, it should be documented.
- User and pass is not friendly.
- Competences and Tags are not shown properly .
- More simple, something like delicious.
- More messages explaining thing.
- Improvement in the layout.
- Scroll error – Multiselect and Search.
- Communication problems.
    - Problems with ILAs sent to the portfolio
- Problems with back button. When you are editing a pending competence and click on cancel it does not return to the list of pending competences.
- Course affiliation is something confusing.
- URL is not a existing competence
- Problems with login when password is forgotten
- Add button next to add new tags, or clarify the message about how to add tags.
- Add a link to the portfolio.
- oAuth first message can be confusing, specially because the two contexts look quite different.
- **Portfolio – ILC integration should be transparent as transparent for the user as possible.**
- Left block can confuse the user
- Too much information in the screen make the ILC not entirely clear
- A bit confusing that some information can be edited in the ILC and in the portfolio (clarify this) and a link from one to the other.
- Clarify where the what are working áreas
- The tags shown in the combo box are not specific for this institution
- Content field is not clear
- It is important to clarify that the ILC Works manually and as a kind of delicious
- Tips in the ILC should be more explicit (in example it is necessary to push in the ? to know how to introduce the words).

PORTFOLIO

- Problems to associate the selected competences.
- Problems with experience and EQF fields when editing a competence.
- Some tips for the users would be helpful.
- Undo Button (the possibility to disassociate a competence from an ILA)
- Some information about ILAs is missed (content or comments are not shown).
- HTML symbols.
- Comments and content are mandatory and it has no sense, this should be advice.
- Report when actions are or not succeeded.
- Direct information about the tags associated for an ILA is not shown.

- Remove testing data.
- Lot of free space on the 7ight side of the screen.
- Difficulty to find a competence there is no search area.
- My competence is not very understandable there is no a direct link to the existing list of competences (It should be explained better).
- **Association of persons with institutions.**
- "The user fields in the dark" (No explanation, difficult to use). II
- Manage competences button is not clear (the general interaction with competences).
- Some steps should be combined.
- List of competences is not refreshed nor when adding a new competence.
- The list of associated competences to an ILA is not updated when you add a new one.
- Number of competences is not the same of the competences you have.
- From the ILA editor you can't add a new competence you have to go to the my competences, add it and later mange it from the ILA editor
- Difficult to find competences
- There is no possibility to delete an ILA.
- Not clear if there is a necessity to save a showcase
- When saving a showcase other are overwrigthed
- When sharing not only the showcase is shared but more information
- Labels missed in the showcase
- When accessing to manage showcases the initial list of showcases is not shown but the edition of the last one accessed.
- The down button on an activity does not work.
- The pdf exportation fails.
- The save button is not intuitive III.
- Buttons such as add competence in the my-competences section should be positioned in other place because if the list is to long and depending on the resolution of the screen you should scroll to see the button.
- When adding a new competence and go to my competences the competence does not appear, maybe mycompetences section should be separted from competences catalogue or better explained.
- It is not clear why some fields appear in the showcase when adding an ILA to it, maybe it should be good to explain a bit more the fields.

INSTITUTIONAL ENVIRONMENT

- Concepts should be explained (Working area, Institutional competence),
- The catalogue when using the method getCompetenceList does return all the competences and not only those validated by the institution.
- Align labels to boxes.
- Add information about what a * means near to a field.
- Include a top menu

- Spacing of the form field
- Problems when unselect areas and tasks.
- Check when validating a user competence if the competence is previously added to the institutional environment or no.t
- When pushing in the dropdown combo box of an area it hides the other options so the user can think that he/she has finished when not.

The distribution of the problems taking into account Nielsen classification by component can be seen in Table 1.

Table 1. – Distribution of erros following Nielsen classification

Component/Severity	Level 0	Level 1	Level 2	Level 3	Level 4
ILC	1	8	11	2	1
Portfolio	5	8	17	3	1
Inst. Environment	1	4	4	0	0

Whilst many breakdown moments did not severely disrupt from the flow of the task (i.e. they were at Levels 1 and 2), it is fair to say that the video evidence showed that the cumulative effect of these was more disruptive in terms of user disposition to the tasks overall (this was captured in the surveys described below). Some software issues were more serious and did cause significant interruption in the task flow (i.e. at levels 3 and 4). These results regarding the flow of experience were fed back to development teams leading to a redesign of parts of the process. Level 0 errors should be taken into account because the fact that they were not considered usability errors does not mean that they were not important.

This data was triangulated with data from the surveys. The survey data captured a more general level of satisfaction. Possibly as a result of the low-level breakdowns in experience, satisfaction levels were 18.4% below the acceptable satisfaction level of 68% described by Sauro (2011). A summary of the results per each expert and the items of the SUS survey can be seen in Table 2.

Tabla 3. – SUS score results

Participant	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	SUS Score
Expert 1	4	2	4	2	3	3	2	3	2	2	**57,5**
Expert 2	2	3	2	2	2	4	2	4	2	3	**35,0**
Expert 3	2	4	3	4	2	5	2	2	3	3	**35,0**
Expert 4	3	2	4	1	3	4	4	1	4	3	**67,5**
Expert 5	1	4	3	3	2	4	1	4	2	5	**22,5**
Expert 6	5	1	5	1	4	2	5	1	4	2	**90,0**
Expert 7	3	4	4	2	4	4	3	2	4	3	**57,5**
Expert 8	3	3	3	5	4	3	2	2	1	4	**40,0**
Expert 9	4	4	3	2	3	4	3	4	3	2	**50,0**
Expert 10	2	5	1	2	1	5	4	5	2	4	**22,5**
Expert 11	2	5	2	4	2	4	1	5	2	5	**15,0**
Expert 12	5	3	3	4	2	5	4	5	3	5	**37,5**
Expert 13	3	4	4	2	3	2	3	2	3	2	**60,0**
Expert 14	4	3	2	1	2	3	3	3	3	2	**55,0**
Average											**49,6**

In order to validate reliability of the distinction-making in the survey analysis, a Cronbach's Alpha was calculated on the variance between user responses to the questionnaire questions. The alpha coefficient showed a value of 0.934 indicating a high level of consistency in the user experience across the sample of expert users. The questions relating to the Perception of Ease of Use showed a netrual value of 4, which shows that the general perception of ease of use is not bad (the statistics for the results can be seen in Table 3).

Table 3. – Statistics for the Items of PEOU

	Average	Standard Deviation	N
Item1	4,21	1,251	14
Item2	4,29	1,684	14
Item3	4,00	1,569	14
Item4	4,00	1,754	14

Regarding with the opinions of the expert two matrix of results were defined (Table 4 and Table 5). The former gathers opinions related with each of the components while the latter classify their opinion on different themes.

Table 4. Qualitative analysis of the opinion of experts classified by the component

	ILC	**Portfolio**	**Inst. Comp**
Expert 1	-	Complex, more description, explanation needed	-
Expert 2	Confusing at the beginning	Complex, difficult interaction	Works good
Expert 3	No intuitive improve interaction	Simple	-
Expert 4	pretty straightforward, improve interaction	Straightforward	Easy
Expert 5	Clear, improve interface	Intuitive	Simple and Intuitive
Expert 6	Ok	-	Fairly easy
Expert 7	Simple and straightforward	-	-
Expert 8	Not user friendly	-	Easy
Expert 9	Not easy to use	Better description and help needed	-
Expert 10	-	-	-
Expert 11	Clear and Simple	Clear	Simple
Expert 12	Easy	Improve interaction	-
Expert 13	Need to simplify	Over-complex	-
Expert 14	Crude	Complex	More explanation to clarify concepts

From this table it is possible to see that although the ILC is for most of the experts quite simple and straightforward, it is necessary to simplify the steps and clarify what means some fields of the form, moreover more tips and a review of the interface is needed. With regard to the portfolio, it is seen as a complete tool but quite complex and with a not very intuitive, it is necessary to improve navigation, change interface and clarify what each section means. If possible it would be interesting to reduce some steps. Regarding to the institutional environment it is seen as a quite simple

tool but it would be interesting to write a document describing each concept managed with this tool.

As commented above general opinions about the system are also gathered. These opinions are classified in the different themes they deal with, Integration, Training and Improvements. The results can be seen in Table 5.

Table 5. Qualitative analysis of the opinion of experts about the whole system classified by theme.

	Integration	**Training**	**Improvements**
Expert 1	Is not correct	A workshop is needed	Debugging is needed to work properly, need to be more intuitive
Expert 2	Issues should be solved		Need to be more intuitive, simplification of interaction
Expert 3	Integration should be solved	Teach the users	Integration in a seamless way
Expert 4	Does not look as integrated product		Task analysis and alignment with the process
Expert 5	Is correct	Not necessary	
Expert 6	Correct	Not necessary	Simplify some options
Expert 7		Some training could be helpful	Explain concepts, improve interface
Expert 8	Too much changes between contexts		Adaptation of the tools to the user
Expert 9	-	-	Need to simplify
Expert 10	Integration is not correct	Training needed	Make clearer what component do
Expert 11	-	-	No improvements needed
Expert 12	-	-	More clarity in the forms
Expert 13	-	Training needed	-
Expert 14	Integration is not crude	-	Seamless integration

From this table it is possible to see that Integration is not working properly and it can be a problem for the project. The final user should not perceive a change of context, so both integration and look and feel should be improved. Regarding to the necessity of a special training, although the system is quite simple, it would be interesting to carry out a workshop with the final users to facilitate them the use of the system.

Other of the most relevant improvements is to simplify the steps to complete the process.

Taking into account these results and those of the other techniques applied, the challenge of finding a technical solution to the bridging of formal and informal learning activities still requires significant work in order to appear seamless and natural to users. However, it should be noted that developments to date have created the opportunity for exploring these experiences in more detail and addressing issues of breakdown in experience. Furthermore, the data reveals that the breakdowns resulted from practical problems rather than conceptual difficulties in understanding what the TRAILER tools were attempting to do. Further work will reveal whether improvements in the tools can produce a more smoothly aligned match between the ideals of TRAILER and the practice.

5. CONCLUSIONS AND FUTURE STEPS

To be defined in Bolton.