

## Description of Usability evaluation at KTH (LMS)

### Background usability and evaluation

The procurement of a LMS includes Usability as an evaluation criterion.

Usability is defined as "the extent to which a specified user can use a product to achieve specified goals with effectiveness, efficiency and satisfaction, in a given context of use" (ISO 9241-11 Guidelines for usability).

ISO 9241-11 explains the benefits with measuring usability in the form of user performance and satisfaction. To perform a usability evaluation, you need to know who will use the product, what the user will do and in what situations the product will be used.

### KTH's evaluation of usability – purpose and methods

KTH's evaluation of usability delivered by a service is divided in three different areas using three different methods.

First, in the *usability test*, a representative number of users use the service (perform a number of different scenarios representative for the user's work situation) under as realistic conditions as possible. The purpose of the usability test is to evaluate if the service supports users to achieve specified goals with effectiveness and efficiency (ISO 9241-11).

Second, in connection to the usability test, the users will answer the questionnaire *System Usability Scale (SUS)*<sup>1</sup> individually. The purpose of the survey is to quantify the satisfaction of the service (ISO 9241-11).

Finally, an *expert evaluation* of the service is to be performed on the basis of the dialogue principles of ISO 9241-110. The purpose of expert evaluation is to see how well the service complies with the general guidelines for interface design, e.g. lack of coherence, unclear or confusing interaction and memory load.

### Execution of tests

#### Premises and other conditions

The evaluation is to be conducted at KTH. In the room there should be a workplace where two user can sit and perform the test on a computer, tablet or smartphone together with two observers.

In the premises, it should be possible to set up a camera since the usability test will be recorded.

Both the usability test and the expert evaluation will be based on test data. It is important that the test data in the service is realistic and representative. The scenarios describe the tasks the test subjects shall carry out and are designed so that they are representative for the roles. Detailed description of the test data and scenarios will be sent out in connection with the invitation to the usability evaluation.

KTH will require access to the service the day before the tests starts.

#### Comment [F1]:

Detta document är huvuddokumentet.

Syftet med dokumentet är att förklara bakgrund, syfte, utförande och poängsättning.

Dokumentet kan användas för att ge mer information till de som efterfrågar det.

Dokumentet har inte skickats till leverantören i sin helhet utan olika delar av document är kopierade till dokumenten:

- FFU (Appendix 1 Stage 2 Evaluation of Usability\_changed.docx) och
- slutrapporten (Procurement report LMS.pdf) som delgivits leverantörerna

Kommentarerna nedan visar vilka delar som kopierats till vilka dokument.

#### Comment [F2]: FFU

#### Comment [F3]:

- Slutrapport
- FFU (men inte, för varje area, den text som börjar med "The purpose")

<sup>1</sup> Brooke, John. SUS - A quick and dirty usability scale (1986)

Before the usability test the supplier shall give all participants in the tests an introduction to their service for no more than one hour.

After the introduction, the usability test will be performed without the supplier; it is however recommended that the supplier is available by phone if problems occur.

Total estimated time for the usability evaluation of one supplier is 2 days. The evaluation is performed in three steps described below.

#### Area I: Usability test

The usability test is performed by eight test subjects divided into four pairs. The pairs represent the four roles as shown in the specifications (Teacher, Course Administrator, Course Assistant and Student).

The subjects, in pairs, solve two scenarios.

How the pairs solves the tasks (scenarios) are observed, measured, documented and graded by the observers

The usability test is measuring (according to ISO 9241-11):

- A. *Effectiveness* - "Accuracy and completeness with which users achieve specified goals". In other words, is it possible to carry out the task and what is the quality of the result?
- B. *Efficiency* - "Resources expended in relation to the accuracy and completeness with which users achieve goals." In other words, how long did it take to complete the task?

#### Area II: SUS questionnaire

After the usability tests the test persons will individually fill out an assessment questionnaire (SUS).

SUS is measuring (according to ISO 9241-11) the user *Satisfaction* – "Freedom from discomfort, and positive attitudes towards the use of the product."

#### Area III: Expert evaluation

Finally, two usability experts will go through the service for about four hours (in the same test environment) and assess the service on the basis of the dialogue principles of ISO 9241-110.

### Grading

Comment [F4]: Slutrapport

#### Assessment of usability for the different areas

Comment [F5]: FFU

Area I: Usability test

- A. Usability effectiveness - maximum 260 points
- B. Usability efficiency - maximum 80 points

The fulfillment of effectiveness and efficiency are separately valued as follows:

- The service applies a solution that optimally supports the user's use of the service, i.e. that contains no lack of importance for usability – 100 % of the maximum points above
- The service applies a solution that in a good way endorses the user's use of the service, i.e. that has a useful function, but which in some respects can be improved – 60 % of the maximum points above
- The service applies a solution that supports the user's use of the service, but which in many respects can be improved – 20 % of the maximum points above

Note: if the user can't complete a scenario, the result is set to 0%

#### Area II: System Usability Scale (SUS )

Usability satisfaction - maximum 45 points

The fulfillment of satisfaction is valued<sup>2</sup> as follows:

- The service applies a solution with the best imaginable user experience, i.e. that contains no lack of importance for usability – 100 % of the maximum points above
- The service applies a solution with good user experience, i.e. which in some respects can be improved - 70 % of the maximum points above
- The service applies a solution with poor user experience, i.e. which in many respects can be improved – 40 % of the maximum points above

#### Area III: Expert evaluation

Fulfillment of dialog principles– maximum 140 points

The fulfillment of dialog principles is valued as follows:

- The service applies a solution that optimally supports the user's use of the service, i.e. that contains no lack of importance for usability – 100 % of the maximum points above
- The service applies a solution that in a good way endorses the user's use of the service, i.e. that has a useful function, but which in some respects can be improved – 60 % of the maximum points above
- The service applies a solution that supports the user's use of the service, but which in many respects can be improved – 20 % of the maximum points above

#### Rating of Tenders

The "Percentage in area" will be transformed into the "Tenders result" based on "Maximum points for rating stage 2".

Area	Percentage in area	Maximum points for rating stage 2	Tenders result
I A		260	
I B		80	
II		45	
III		140	
<b>Total</b>		525	

---

<sup>2</sup> Bangor, Aaron, Kortum, Philip & Miller, James. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale, Journal of Usability Studies, Vol. 4, Issue 3, May 2009, pp. 114-123