

# Common Industry Format (CIF) Report Customization for UX Heuristic Evaluation

Llúcia Masip, Marta Oliva, and Toni Granollers

University of Lleida, Jaume II 69, 25001 Lleida, Spain  
{lluciamaar, oliva, tonig}@diei.udl.cat

**Abstract.** The ISO/IEC 25062:2006 standard presents the Common Industry Format template for the user test results. This standard only considers usability results of a specific evaluation technique: the user test. Nowadays, more aspects apart from usability are considered to get a product quality. In addition, other evaluation methodologies are also very used and different types of results can be obtained (formative/summative and qualitative/quantitative). This research proposed how to adapt the Common Industry Format report for the heuristic evaluation methodology and considering more aspects apart from usability.

**Keywords:** CIF, ISO/IEC 25062:2006, customization, heuristic evaluation, UX.

## 1 Introduction

The Common Industry Format (CIF) for Usability Test Reports is a template for reporting usability test that provides a standardized document to be used by companies and User eXperience (UX) practitioners when documenting usability evaluations. This international standard is intended to be used for: (i) Reducing training time for UX practitioners: an individual only needs to learn to use one form regardless of how many companies he or she works for [1]. (ii) Increasing the communication between vendors and purchasing organizations sharing a common language and expectations [1]. (iii) Providing a guide for expert and non-expert organizations [1]. (iv).Providing a comparable report with other similar ones. (v) Presenting a useful template that can be used in the results of all different and specific cases.

Besides CIF template is the best contribution to standardize usability test reports, it presents three challenges in order to be applicable in real context. The first challenge is the applicability of these kinds of reports in more facets than usability. Usability is the most considered facet in the development process of an interactive system, but up to now other facets have been emerging [2]. Therefore, it is necessary to consider in the CIF report the chance of adapting the results of a test according to other facets such as security, emotional, cross-cultural, among other.

Second challenge is concerning the final purpose of the CIF report. As is explained in “Audience” section of the standard [1], the information that the CIF shows is

summative because it tries to present a summary of the results. However, these days, and bearing in mind the competitive sector, much more specific information should be provided to reach more details about the identification of problems. Therefore, the CIF should also include formative information. Furthermore, CIF template has been thought for reporting user tests. But the number of methodologies for improving an interactive system is wide extensive, then CIF template should be customized for reporting more evaluation methodologies. This is the third challenge.

In summary, the main goal of this communication is to present the adaptation of the CIF template to the three challenges introduced above and considering the heuristic evaluation as the methodology to get UX evaluation results. In the following section our methodology to carry out a heuristic evaluation is presented. Then, a comparison between the user test and heuristic evaluation is detailed. Following this section, the state of the art about some CIF customizations are presented. Later, the differences applied in the CIF template updating some sections of the standard report. Afterwards, how to get this new template using our web-based resource called Open-HEREDEUX is described [10]. Finally, conclusions and future work are explained.

### 1.1 Our Methodology for UX Evaluations Based on Heuristics

Heuristic Evaluation (HE) was created by Johnson, Ravden and Clegg [3] but it was promoted by Nielsen and Molich [4] only one year later. The methodology is divided into three main steps: the organization of the evaluation, the evaluation of the heuristics and the extraction of results. In the first step, the administrator of the evaluation should select the best heuristics for the specific interface, should choose who will be the evaluators of the interactive system and he or she should determine the severity factors that evaluators will use to score each heuristic (by default Nielsen's scale [4] is considered in the heuristic evaluation). In the second step, each evaluator scores each heuristic using the severity factors selected previously. Once all evaluators have scored all heuristics, results are taken. Usually in the third step, qualitative and formative results are obtained according to the scores of the evaluators and the observations that they made in the scoring process.

However HE is one of the most used methodologies, it presents some deficiencies that induce a slower and more expensive process. Two are the most relevant: (i) The first deficiency is the non-existence of a repository or library containing specific guidelines (heuristics) saved in. Usually, whoever plans to carry out a heuristic evaluation needs to review the literature, select different sources of heuristic lists and, then, elaborate manually the most appropriate sets of heuristics to determine the best guidelines for every specific interactive system.

(ii) Another relevant deficiency appears at the end of the process, when obtaining the evaluation results. HE traditionally presents qualitative and formative results according to the evaluator's scores and their observations during the scoring process. However, quantitative and summative results are sometimes achieved because the scientific community aims for objectivity [5] which cannot be obtained using qualitative results.

Then, one of our main goals is to solve the heuristic evaluation deficiencies defining a methodology that semi-automates some parts of the heuristic evaluation. Our new proposal includes the repository of information, the heuristic suggestion process, the realization of the evaluation and the extraction of UX results in a standard report.

A repository of information is essential for reaching an optimum process for suggesting heuristics. Therefore, heuristics are the most valuable data that should be related to the components, features, functionalities and UX facets. Using this relation and considering that an interactive system is defined through components, functionalities, features and UX facets, the selection of the most suitable heuristics for a system is evident. [6]

Regarding the suggestion for heuristics, our new proposed methodology solves one of the main deficiencies of the heuristic evaluation methodology. It details for the suggestion of the best heuristics for a specific interactive system using different considerations such as different goals for design or evaluation, different types of heuristics according to the receiver [7], financial constraints (such as the UX degree [8]) and the possibility of documenting conflicting heuristics [9]. Then, these two first stages of the methodology provide the needed conditions to facilitate the evaluation process per se. This stage is not really automatic or optimized but due to the consideration of the rest of the methodology, the scoring of each heuristic becomes a very easy step.

Finally, the last part of the methodology is the extraction of results. Our methodology proposes the automatic process to get the results and also the automatic downloaded report. The most general goals of this automatic downloaded report are: to facilitate the interpretation of these results to improve the specific interactive system, and to compare these results in the possible following evaluations of improved versions and other systems, for instance systems from rivals. We are working on the definition of different UX measures to solve this gap.

In this paper we will focus on the presentation of the UX evaluation results using the heuristic evaluation and presenting the results through the ISO/IEC 25062:2006 standard (Software engineering -- Software product Quality Requirements and Evaluation (SQuaRE) -- Common Industry Format (CIF) for usability test reports). Thus, the next section presents the main difference between user test (technique used in the CIF report) and heuristic evaluation (our base to propose a new UX evaluation methodology based on heuristics).

## 1.2 User Test vs. Heuristic Evaluation

User test and heuristic evaluation are the most used methodologies to improve the product quality [11]. The most important differences among both methods are detailed in the following paragraphs.

Regarding the classification of usability evaluation methodologies, the most used and general classifications are: inquiry, test and inspection evaluation methodologies [12] [13]. User test is a test evaluation methodology where representative users work on typical tasks using the system (or the prototype) and the evaluators use the results

to see how the user interface supports the users doing their tasks. On the other hand, heuristic evaluation is an inspection technique performed by usability specialists that examine usability aspects (from a preselected list of guidelines or heuristics) of a user interface.

Another important difference is concerning the stakeholders of the methodology. User test is prepared by a project manager but it is performed by real end users. In contrast, heuristic evaluation is also prepared by a project manager but it is performed by expert users. Therefore, heuristic evaluation needs expert users understanding those that know very well the heuristic evaluation methodology and/or the interactive system.

In addition, the specific execution of both techniques presents a main difference. User test introduces tasks that users have to perform to get the usability problems of the product. The evaluators of the heuristic evaluation do not need tasks. They usually freely explore the interface for detecting as many problems as possible considering the heuristics of the evaluation.

The execution costs of both techniques are considerably different. Usually, the budget of heuristic evaluation is cheaper than user test, mainly because heuristic evaluation does not need end users. The time to recruit users represents the mainly part of the user test budget. So, only for this reason user test is considered as one of the most expensive techniques. Table 1 summarizes the differences among both methodologies.

**Table 1.** User test vs heuristic evaluation

<i>Methodology</i>	<i>Type</i>	<i>Users</i>	<i>Procedure</i>	<i>Budget</i>
User test	Test	End and real users	Task oriented	Expensive
Heuristic evaluation	Inspection	Expert users	Free navigation	Cheap

## 2 CIF Customization

Certainly, there are few references about how to customize, adapt or modify the CIF for a specific real context of use. The National Institute of Standards and Technology [14] changed some sections and information of the CIF template to adapt it for voting systems. In the same way, another customization done by the same institute is in the electronic health sector [15]. Furthermore, a research to adapt the CIF to the heuristic evaluation methodology is presented in [16]. Using the heuristics evaluation methodology some formative results appear in the CIF report. However, the most informative report is the one presented in [17].

Due to the three of the needed improvements, an adaptation of the CIF report is proposed. Next table presents the current sections of the CIF template and checked with those changed in our proposal. Then more details about the changes done in each section are described.

**Table 2.** The updated sections of the CIF

CIF section		Is it changed?
Executive summary		No
Introduction	Full product description	No
	Test objectives	No
Method	Participants	Yes
	Context of product use in the test	Yes
	Participant's computing environment	No
	Test administrator tools	No
	Experimental design	Yes
	Usability metrics	Yes
Results	Data analysis	Yes
	Presentation of results	Yes

2.1 Method Section

The information regarding the evaluation process is described. The changed sections of the CIF report here detailed are: participants, context of product use in the test, experimental design and usability metrics.

**Participants.** In this section of the CIF template, information about the users is required. Printed version of the standard includes end users whom did the test. Information about their specific profile is asked for. In our proposal, this section needs to include to different user profiles: End users: If the UX would be considered, the end user should be included in the development process of an interactive system. For this reason, the end user who is not an expert in the methodology based on heuristics is considered in the report. Therefore, the novelty is not in the consideration of the end user in the report but the consideration of the end user in the methodology used for achieving formative results. Furthermore, end users will provide information about their likes and feeling to get results for the emotional UX facet. Expert users: Expert users are those users who know the methodology that is used to evaluate the interactive system. For instance, interface designers, UX researchers or project managers. The users are the traditional evaluators proposed in the heuristic evaluation.

In our methodology, this information is obtained from two pre-evaluation questionnaires that each evaluator should fill in before starting the evaluation. The CIF reports this information in two different tables. The table referring to the end users includes information about gender, age, education, computer experience and product or interactive system experience. The table referring to the expert users presents gender, age, education, computer experience, years in HCI and the number of evaluations done.

**Context of Product Use in the Test.** In this section, CIF template includes different subsection. Our focus is in the subsection called “Tasks”. There, many details about the tasks that users have to perform are described.

Tasks subsection was removed in our proposal because heuristic evaluation methodology does not use tasks. End users and expert users use the interactive system freely.

**Experimental Design.** In the same way as the last point, this section has a subsection called “Participant task instructions”. So for the same reason commented above, the subsection called “Participants task instructions” was also removed.

**Usability Metrics.** The CIF template calls this section as “Usability metrics” and metrics for effectiveness, efficiency and satisfaction are proposed.

Bearing in mind our approach towards the UX, our proposal calls this section “UX metrics”. In addition, the metrics for effectiveness, efficiency and satisfaction have been substituted by new ones. (We are working on these UX metrics.)

## 2.2 Results Section

This section describes how the results of the evaluation are provided.

**Data Analysis.** In this section, the CIF template proposes to present enough detail to allow replication of the data analysis methods by another organization if the evaluation is repeated. Specifically, it details the data collection (the differences between the data that was planned to be collected and the data that was actually collected), the data scoring (the mapping between the data values that were collected and the values used in further analysis) and data reduction (the method used to compute the measure of central tendency and to characterize the variation in the data). Finally, the statistical analysis used to analyze the data is explained in detail.

Severity factors used to score each heuristics are presented in our proposal. Heuristic evaluation uses impact and frequency defined by Nielsen as the severity to score the heuristics. Therefore, the results analysis is based on the scorings of each heuristic per each evaluator using these severity factors.

**Presentation of Results.** CIF template proposes specific tables for every usability metric and for every end user. Obviously, if the metrics to analyze the UX are changed, the presentation of the results is also modified.

Qualitative and formative results are presented in our proposal as an improvement list for the interactive system. This improvement list includes the whole set of violated heuristics. Furthermore, quantitative and summative results are shown through the UX measures. These quantitative and summative results enable the comparison of different versions of the same interactive system and it permits comparison among the same types of interactive systems.

## 3 How to Get the CIF Template through Open-HEREDEUX

Then, based on the heuristic evaluation methodology, OPEN-HEREDEUX is presented as a solution to consider the UX in an interactive system design or

evaluation process. OPEN-HEREDEUX is the short name for our proposal: “OPEN HEuristic Resource for Designing and Evaluating User eXperience”. Figure 1 shows the Open-HEREDEUX overview. Open-HEREDEUX is available on this url: [www.grihotools.udl.cat/openheredeux](http://www.grihotools.udl.cat/openheredeux).

Open Repository (Fig 1- 1) is provided with all the necessary information to achieve the set of heuristics as complete and minimum as possible.

Adviser of heuristics is the second component (Fig 1- 2). It intends to be a tool whose objective is to propose, for a specific interactive system, the most appropriate list of heuristics to be used. It is suitable for such usages as recommendation principles in a design phase or as evaluation principles in a UX evaluation based on heuristics. Therefore, heuristic suggestions can be used either as a list of recommendations to design a specific interactive system or as an input for the next, and third, component: Scorer of heuristics, which is in charge of carrying out the realization of the evaluation (Fig 1- 3). Finally, Results Analyzer is the last component. It provides quantitative/qualitative and formative/summative data interpretation (Fig 1- 4).

New CIF template report is obtained by Result Analyzer component. It provides the report in an editable and standard format file. The main tasks that the project manager can carry out using Results Analyzer are: (i) Download the new report. (ii) Upload the previous downloaded (or not) report for saving it in Open Repository. (iii) Download the previous uploaded report.

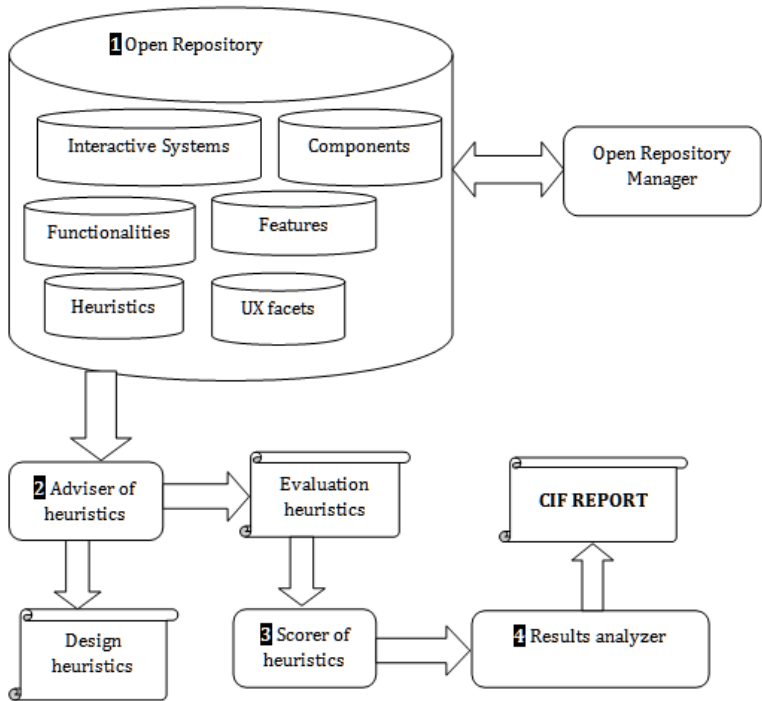


Fig. 1. Open-HEREDEUX overview

In reference to qualitative and formative results, a list of improvements according to the evaluators' observations and heuristics scored incorrectly will be achieved. They are the ones that designers should apply to the interactive system to improve it. Apart from qualitative and formative results, quantitative and summative ones should be presented because these will attempt to show the UX degree for every interactive system [8]. If quantitative and summative results are achieved, UX experts will have a standard method or a possible certification to compare evaluations and see which interactive system provides users with the best experience. Quantitative results will get an objective measure that is impossible to obtain due to the subjectivity of qualitative results.

## 4 Conclusions and Future Work

CIF report is a very useful document to present quality results but it should be customized to enable the better description in different context of use. We proposed the customization of the CIF template in three ways: the adaptation to the heuristic evaluation methodology, the consideration of more aspects apart from the usability and the proposal of summative/formative and qualitative/quantitative results in the report.

Our customized proposal do not change the essence of the CIF template report, it only includes and changes some sections and subsection to provide with an added value to UX result report. Furthermore, the implementation of an automatic generation of this report provides UX practitioners with a very fast way to get evaluation reports. This process could be translated into a certification of the UX evaluation in the near future.

## References

1. ISO/IEC 25062:2006 Software engineering. Software product Quality Requirements and Evaluation (SQuaRE). Common Industry Format for usability test reports (2006)
2. Masip, L., Oliva, M., Granollers, T.: User experience specification through quality attributes. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) INTERACT 2011, Part IV. LNCS, vol. 6949, pp. 656–660. Springer, Heidelberg (2011)
3. Johnson, G.I., Clegg, C.W., Ravden, S.J.: Towards a practical method for user interface evaluation. *Applied Ergonomics* 20(4), 255–260 (1989)
4. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: CHI 1990: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, NY (1990)
5. Neil, J.: Qualitative versus quantitative research: Key points in a classic debate. *Research methods* (2007), <http://wilderdom.com/research/QualitativeVersusQuantitativeResearch.html> (last visit: December 19, 2012)
6. Masip, L., Oliva, M., Granollers, T.: The Open Repository of Heuristics. In: *Proceedings of Interaccion 2012*, Elche, Alicante, Spain, October 3-5. ACM (2012); 978-1-4503-1314-8/12/10 (Copyright 2012)



7. Masip, L., Granollers, T., Oliva, M.: A Heuristic Evaluation Experiment To Validate The New Set of Usability Heuristics. In: Proceedings of 8th International Conference on Information Technology: New Generations, Las Vegas (2011); 978-0-7695-4367-3/11 © 2011 IEEE
8. Masip, L., Granollers, T., Oliva, M.: User Experience Degree and Time Restrictions as Financial Constraints in the evaluation methodology based on heuristics. Accepted paper in 6th Latin American Conference on Human Computer Interaction (2013)
9. Masip, L., Martinie, C., Winckler, M., Palanque, P., Granollers, T., Oliva, M.: A design process for exhibiting design choices and trade-offs in (potentially) conflicting user interface guidelines. In: Winckler, M., Forbrig, P., Bernhaupt, R. (eds.) HCSE 2012. LNCS, vol. 7623, pp. 53–71. Springer, Heidelberg (2012)
10. Masip, L., Oliva, M., Granollers, T.: OPEN-HEREDEUX: open heuristic resource for designing and evaluating user experience. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) INTERACT 2011, Part IV. LNCS, vol. 6949, pp. 418–421. Springer, Heidelberg (2011)
11. “UPA, Salary Survey” Usability Professional’s Association (2009), [http://usabilityprofessionals.org/usability\\_resources/surveys/2009salarysurvey\\_PUBLIC.pdf](http://usabilityprofessionals.org/usability_resources/surveys/2009salarysurvey_PUBLIC.pdf)
12. Nielsen, J., Mack, R.L.: Usability Inspection Methods. John Wiley & Sons, New York (1994) ISBN: 0-471-01877-5
13. Dix, A., Finlay, J., Abowd, G., Beale, R.: Human-Computer Interaction. Prentice Hall (2009)
14. National Institute of Standards and Technology. Voluntary Voting System Guidelines Recommendations to the Election Assistance Commission. Modified CIF template for voting system test laboratories- (August 2007)
15. Schumacher, M., Lowry, S.: Customized Common Industry Format Template for Electronic Health Record Usability Testing. National Institute of Standards and Technology Interagency/Internal Report (NISTIR) – 7742 (November 16, 2010)
16. von Baggio, K., Johnston, L., Burmeister, O.K., Bentley, T.: Common Industry Format: Meeting Educational Objectives and Student Needs? In: Masoodian, M., Jones, S., Rogers, B. (eds.) APCHI 2004. LNCS, vol. 3101, pp. 521–530. Springer, Heidelberg (2004)
17. Formative Test Reports. Journal of Usability Studies 1(1), 27–45 (November 2005)