

## **Specifying and evaluating usability requirements using the Common Industry Format**

*Four case studies*

Nigel Bevan <i>Serco Usability Services</i> nbevan@ usability.serco.com	Nigel Claridge <i>Scandinavian Usability Associates</i> nigel_claridge@ hotmail.com	Martin Maguire <i>Loughborough University</i> M.C.Maguire@ lboro.ac.uk	Maria Athousaki <i>SIEM</i> siem@otenet.gr
--	--	---	--

**Abstract:** The Common Industry Format for usability test reports has been used to introduce usability into public and private contracts for the development of two web sites, acquisition of a travel management system and acquisition of travel agency software. Four pairs of supplier and consumer organisations worked with usability specialists to establish usability requirements and/or to evaluate whether the supplied system met the requirements.

**Key words:** usability, evaluation, requirements, procurement, standards

### **1. INTRODUCTION**

Incorporating usability requirements in the procurement process can reduce the risk of failure when implementing a newly acquired system and increase ease of use and thus productivity and/or profitability.

- Lack of user performance requirements was a fundamental reason for the expensive costs and delays incurred when new passport issuing software developed by Siemens was installed in the UK (Public Accounts Committee, 1999).
- Two studies have shown that the user success rate in purchasing from current ecommerce web sites is in the range of 25-60% (Nielsen, 2001; Spool & Schroeder, 2001). Small improvements in user performance could lead to substantial increases in revenue.

A new Common Industry Format (CIF) for documenting usability results for use in the procurement process has been developed by a US-based group of companies coordinated by NIST (Bevan, 1999a; Blanchard, 1998). The

format has been approved as American standard (ANSI/NCITS 354), and is intended to be submitted to ISO.

The EU-funded PRUE project (Bevan et al, 2001) has demonstrated the value of using the Common Industry Format in four case studies: public and private contracts for development of a web site, and acquisition of a travel management system and travel agency software.

Serco worked with the Italian Ministry of Justice to introduce usability requirements into the acquisition of a new legal information web site, and Scandinavian Usability Associates (SUA) worked with Ericsson to introduce usability requirements in the procurement of an office software product.

Loughborough University and SIEM used the CIF to evaluate the effectiveness of an existing system to assess its acceptability and the need for improvements as part of a contractual relationship with a supplier. Loughborough University assessed an online shopping website in the UK, and SIEM assessed travel agency software in Greece.

## **2. WHAT IS THE COMMON INDUSTRY FORMAT?**

The Common Industry Format for usability test reports specifies the format for reporting the results of a summative usability evaluation. The most common type of usability evaluation is formative, i.e. designed to identify usability problems that can be fixed. A summative evaluation produces usability metrics that describe how usable a product is when used in a particular context of use (Bevan & Macleod, 1994; Macleod et al, 1997). The CIF report format and metrics are consistent with the ISO 9241-11 definition of usability:

The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.

The type of information and level of detail that is required by the CIF standard is intended to ensure that:

- Good practice in usability evaluation had been adhered to.
- There is sufficient information for a usability specialist to judge the validity of the results (for example whether the evaluation context adequately reproduces the intended context of use).
- If the test is replicated based on the information given, it should produce essentially the same results.
- Specific effectiveness and efficiency metrics must be used, including the unassisted completion rate and the mean time on task.
- Satisfaction must also be measured.

It was envisaged that a supplier would provide a CIF report to enable a corporate purchaser to take account of usability when making a purchase decision. A purchaser could compare CIF reports for alternative products

(particularly if a common set of tasks had been used). The purchaser might specify in advance to a supplier the required values of the usability measures (for example based on the values for an existing product).

The original motivation for the CIF came from usability staff in purchasing companies who were frustrated at purchase decisions made exclusively on the basis of functionality. These companies experienced large uncontrolled overhead costs from supporting difficult to use software (Blanchard, 1998).

The CIF format was agreed by a working group of usability experts from purchasing and supplying companies. It was based on collating good practice from the different companies, and aligning this with ISO 9241-11.

### 3. PURPOSE OF THE TRIALS

Table 1. Summary of trials

	<b>Italian Ministry of Justice + SEMA</b>	<b>Ericsson Telecom + Electrolux</b>	<b>in2netlogic + prezybox.com</b>	<b>Robissa Travel + ExcelSoft</b>
<b>Product</b>	Web site: legal info	Travel expense reporting	Web site: ecommerce	Travel management
<b>Users</b>	Legal staff and public	Office staff	Public	Travel agent staff
<b>Existing system</b>	Traditional dial up	Text based	Existing web site	Existing product
<b>New system</b>	Web based	Web based	Improved version	Improved version
<b>Contractual relationship</b>	The Ministry will use CIF results as part of acceptance criteria	Ericsson will use CIF results as part of acceptance criteria	Basis for negotiation for improvements	Basis for negotiation for improvements
<b>Consumer organisation benefits</b>	Convenient and accurate legal info	Increased productivity and quality of travel info	Increased business	Increased office efficiency

The main purpose of the PRUE project was to evaluate the business value of introducing the CIF into the purchasing and procurement process in a variety of contractual environments. There were four trials (Table 1).

One trial was for the originally envisaged role in the acquisition of corporate software (Ericsson/Electrolux). Another extended this to acquisition of a custom-designed web site (Ministry/SEMA). These trials both focussed on evaluating whether a new web-based system meets requirements established by prior evaluation of the existing text-based system.

The two other trials (in2netlogic/prezzybox.com and Robissa Travel/ExcelSoft) assessed the usefulness of the CIF in negotiating for improvements in a custom-designed system, by evaluating the usability of an existing product to establish requirements for an improved version. The in2netlogic/prezzybox.com trial was in the UK where usability is well-established as an issue, and Robissa Travel/ExcelSoft was in Greece where usability is less well recognised.

#### **4. ITALIAN MINISTRY OF JUSTICE/SEMA: LEGAL WEB SITE**

PRUE has helped the Italian Ministry of Justice introduce usability requirements into the procurement of a new enhanced version of the existing dial-up ITALGIURE-FIND legal information service. This builds on previous work with Italian Public Administrations (Catarci et al, 2000).

SEMA has a contract with the Ministry to produce a new system accessible from internet browsers and mobile phones. No information was previously available on the usability of the existing system, and there were no usability requirements in the contract (although the contract required ISO 13407 to be used).

PRUE planned three activities to introduce usability to the procurement process:

- A stakeholder meeting to identify the importance of usability and the intended context of use.
- An evaluation of the existing system to provide measures of usability as baseline requirements.
- An evaluation of a prototype of the new system to establish whether the requirements have been met.

The stakeholder meeting was attended by senior representatives of the Ministry, SEMA and other stakeholders. Two user groups were identified: expert users with legal knowledge need functionality to support fast and efficient searches, consistent with the natural logic of juridical research. The system will also be accessible to new users without any legal knowledge who should be able to carry out simple searches without training.

In the evaluation of the existing system, five experts and four non-experts attempted to complete different sets of typical tasks. Expert users successfully accomplished 63% of tasks in comparison with 52% of non-expert users. The evaluation also provided the opportunity for SEMA designers to observe real usage of the system.

An important objective was for the new system to at least equal and if possible improve on these success rates. If the measured success rates and user satisfaction scores are not acceptable, the Ministry could negotiate with SEMA to provide an improved system.

The trial has demonstrated several benefits of introducing summative usability testing into the procurement process:

- It provides a concrete benchmark for user performance and satisfaction, thus reducing the risk that the new system is more difficult to use (and therefore less successful) than the existing system.
- It highlights usability problems with the existing system that need to be addressed in the design of the new system.
- It provides specific goals for usability and gives developers the opportunity to become familiar with typical user task scenarios.
- It provides the framework for the more detailed usability work required by ISO 13407.

Summative evaluation results reported in the CIF format have thus been very useful in helping the Ministry understand the needs of different user groups, and in establishing usability requirements for the new system. The user testing has also clear benefits in helping the supplier (SEMA) better understand Ministry of Justice requirements.

## **5. ERICSSON/ELECTROLUX: TRAVEL MANAGEMENT SYSTEM**

In order to support the procurement process, the Common Industry Format (CIF) for usability testing has been used to obtain usability metrics at Ericsson and Electrolux in Sweden. First, the current Ericsson system was tested using the CIF to obtain a baseline value for usability, which resulted in a usability requirements document. The document required information about the way in which the supplier had developed the product and detailed information about user performance and user satisfaction metrics. Secondly, the likely supplier of a new system called WEBRES developed by Electrolux was tested and the results compared to the requirements document.

Ericsson was investigating the purchase of a new interface to the existing travel management system - RES. The current text-based interface no longer supported users task demands and requirements effectively and it was not found easy to use. RES was used by about 25% of Ericsson employees

although the goal was to increase this percentage significantly. It was a complete system that supported business travel management both nationally and internationally.

Ericsson did not wish to make significant changes to the underlying system. They looked to a number of potential suppliers who were able to provide a new interface, which would enhance the efficient and effective use of the existing system. At this time, the prime candidate supplier was Electrolux. They used essentially the same travel management system as Ericsson but had developed in-house interface solution - WEBRES. This interface was a possible replacement to the current RES interface.

At this time, Ericsson had not considered any usability issues during the purchase process. There was no objective data about user performance/subjective assessment using RES and there was no usability requirements document. The PRUE project was able to support the procurement process by conducting summative usability testing on RES using the CIF format. The objective was to obtain user performance and satisfaction metrics of the RES interface through formal user testing. The results from testing, which showed low satisfaction scores and a mean task completion rate of only 55%, led to a usability requirements document against which potential suppliers could be assessed. Subsequently, similar testing using the CIF was conducted on Electrolux's WEBRES. The test results were then compared and contrasted to the usability requirements document prepared by Ericsson.

PRUE has been of great value to Ericsson. Without PRUE, usability issues would not have been included in this procurement process. PRUE has resulted in a usability requirements document for a new interface to RES, which was regarded as a new and positive input to the Ericsson decision process when selecting a supplier organization. The usability requirements were regarded as a good complement to the existing functional and technical specifications. Further, there were benefits of increased understanding of user performance using the current system (poor performance is a large cost in time to Ericsson), an objective understanding of what users thought of the current system and what they require from the new.

From the supplier perspective, Electrolux have been able to understand the performance/ satisfaction related Ericsson requirements for WEBRES. The CIF for usability testing enabled them to assess how WEBRES compared to the Ericsson requirements, and it has provided a clear indication of what improvements need to be made to the interface design.

## 6. PREZZYBOX/IN2NETLOGIC: ECOMMERCE WEB SITE

This study was concerned with an online shopping website called *Prezzybox*. The site, developed by *in2netlogic*, offers a wide range of gift ideas for purchasers to select and have delivered, either to themselves, or to a friend or relative. The website was evaluated by the PRUE partner - the Research School of Ergonomics and Human Factors (RSEHF) at Loughborough University (which incorporates the former HUSAT).

The main objective of the user test was to obtain user performance and satisfaction metrics for Prezzybox along with user comments to highlight any problems with the site. Testing was carried out with 12 users who were all experienced with the Internet and interested in buying from the Prezzybox site. Users were asked to make a real purchase from the site - this being their payment for taking part. Thus the evaluation was designed to test the success of users in completing an online purchase - a crucial success factor for any online shopping site.

The result of the evaluation was a CIF report that documented user performance and satisfaction with the site. This showed that 2 out of the 12 users failed to make a purchase. If the consumer organisation could capture those two users, their sales would increase by 20%!

These performance results were also supported with user satisfaction ratings. The levels of satisfaction recorded were in general around just 'satisfactory' representing clear scope to improve the 'user experience' when using the site.

Evaluator comments and recommendations for change to the site were also included, highlighting features that could be changed to help improve user success in making a purchase.

The CIF benefited the consumer organisation (Prezzybox) by enabling it to:

- Find out how successful consumers will be in making a purchase from their site i.e. what percentage can actually make a purchase?
- Provide a benchmark for user performance and attitude, which can be used for comparison with the shopping site when it is revised.
- Obtain insights into any problems that users face when using the site (to complement the summative results) and to receive suggestions for improving the site.

The CIF benefited the supplier organisation (*in2netlogic*) by enabling it to:

- Obtain objective feedback on the success of the design they produced.

- Identify the most important usability issues that will enable the shopping site to support more successful purchases – and therefore improve the profitability of the site for the consumer organisation.
- Negotiate a new contract to improve the site, based on the test results and the comments and suggestions for improving the site.

Currently the Prezzybox site is being refined based on the feedback. A second round of usability evaluation is planned so that the consumer and supplier organisations can receive concrete evidence that the site has been improved. It is also hoped to show that this usability activity has enhanced the site from the user's point of view.

In summary, the evaluation approach of the CIF is recommended for other online shopping providers in order to test what proportion of users actually make a purchase from their site. It also sets a baseline against which new versions and upgrades to the site can be compared, while also highlighting the problems that need to be fixed if the site is to achieve a greater number of online sales.

## **7. ROBISSA TRAVEL/EXCELSOFT: TRAVEL AGENT SOFTWARE**

In the context of the PRUE Project, SIEM, in cooperation with a software development company and a representative client (Robissa Travel, a travel agency), planned and carried out a usability test for a software application (Global Travel) that supports the management tasks of a travel agency. All the steps of the evaluation procedure, as well as the final results were recorded and reported using the CIF format. The procedure that was followed included the following steps:

1. Definition of the product to be tested
2. Definition of the context of use
3. Specification of the usability requirements
4. Specification of the context of the evaluation
5. Design of the evaluation
6. Performance of the user tests and collection of data
7. Report and analysis of the collected data

The benefits for both participating organizations (i.e., the supplier and the consumer) were considerable:

- The supplier, through a standardised and valid test procedure obtained a concrete and objective measure of the usability of the tested product in order to demonstrate its quality but also spot potential problems.
- Additionally, the evaluation data (usability problems, new user requirements, ideas, etc.) that were collected will be useful as input for the design of a future version of the product.

- The consumer was able to judge the usability of the supplier's product, through the evaluation process as well as the extent to which the specific product caters for his/her particular needs.
- Another expected benefit was that the consumer's opinion and needs, highlighted during the whole process, will be taken into account for the design of the next version of the product.
- Additionally, the process helped the consumer identify his / her real needs, state them to the developers in an organised and understandable way, and produce a document for common reference.
- Finally, since, on the one hand, the supplier will be able to take into account the consumer's needs by using a well-documented and objective procedure, and, on the other hand, the consumer will have an objective process (the CIF report) for judging the extent to which this was accomplished, it is clear that there is a unique opportunity for the creation of a close and mutually beneficial relationship between them. The results of such a relationship can be for the supplier: a loyal customer, and for the consumer: software products of higher quality, productivity and usability.

In conclusion, the study proved that the CIF format can provide real added-value to a software project, because it is a structured and well-defined process. Both the consumer and the supplier considered it as an efficient, effective and worthwhile activity that had positive results for both. The extra resources (time and effort) required are well justified and spent, and both organizations are positive to using again the CIF format in the future. Of course it has to be noted that in order for the whole process to be resource-effective but also valid and productive, an organization that has high expertise in usability engineering is required, since both the wording and the process of the CIF format require a theoretical background, but also a considerable amount of concrete previous experience, in the field.

## 8. CONCLUSIONS

Much early usability work used summative methods (Whiteside et al, 1988), but was not always supported by other user centred design activities. It therefore gained the reputation for being an expensive way to identify problems when it was too late to fix them! So the emphasis moved to formative evaluation (so-called “discount” usability methods) that could be used earlier in development (Nielsen, 1993). It is essential to introduce usability early in the development process, but without subsequent summative testing, it is difficult to judge the effectiveness of the usability work (Bevan, 1999b).

Summative usability testing using the CIF has advantages to both consumer and supplier organisations during procurement. It is one of the most effective ways to enrich the consumer's requirements document with objective user performance and satisfaction metrics (based on the existing system/product used). It provides a platform on which to evaluate potential competitive products from a number of supplier organisations during the procurement of a new system/product. By using the CIF test structure, suppliers are able to demonstrate that their product complies with the usability metrics defined in the requirements document.

The adoption of the format as an international standard should provide a strong case for the wider use of summative testing reported in the Common Industry Format. The PRUE case studies provide strong evidence for the benefits of this approach.

## 9. ACKNOWLEDGEMENTS

The work with the Italian Ministry of Justice was carried out in conjunction with Tiziana Catarci (Universita' di Roma "La Sapienza"), Giacinto Matarazzo (Fondazione Ugo Bordon) and Gianluigi Raiss (AIPA, Rome).

## 10. REFERENCES

- Bevan, N. (1999a) Industry standard usability tests, in S. Brewster, A. Cawsey & G. Cockton (eds.) *Human-Computer Interaction – INTERACT '99 (Volume II)*, British Computer Society, pp 107-108.
- Bevan, N. (1999b) Quality in use: meeting user needs for quality, *Journal of Systems and Software*, **49**(1), pp 89-96.
- Bevan, N. and Macleod, M. (1994) Usability measurement in context. *Behaviour and Information Technology*, **13**(1), pp.32-145.
- Bevan, N., Bogomolni, I. & Ryan, N. (2001) *PRUE: Providing Reports of Usability Evaluation*, [www.usability.serco.com/prue](http://www.usability.serco.com/prue)
- Blanchard, H. (1998) Standards for usability testing. *SIGCHI Bulletin*, **30**(3), pp.16-17.
- Catarci, T., Matarazzo, G., and Raiss, G. (2000) Usability and Public Administration: Experiences of a difficult marriage, in *Proceedings of the 1st ACM international conference. on universal usability*, ACM Press, pp.24-31.
- ISO 9241-11 (1998) *Ergonomic requirements for office work with visual display terminals (VDT)s - Part 11 Guidance on usability*.
- ISO 13407 (1999) *User centred design process for interactive systems*.
- Macleod, M., Bowden, R., Bevan, N. & Curson, I. (1997) The MUSiC Performance Measurement Method. *Behaviour and Information Technology*, **16**, pp 279-293.
- Nielsen J (1993) *Usability Engineering*, Academic Press.

Nielsen J (2001) <http://www.useit.com/alertbox/20010819.html>

Public Accounts Committee (1999) *Improving the Delivery of Government IT Projects*  
<http://www.publications.parliament.uk/pa/cm199900/cmselect/cmpubacc/65/6502.htm>

Spool, J. & Schroeder, W. (2001) Testing Web Sites: Five Users is Nowhere Near Enough.  
*Proceedings of ACM SIGCHI 2001.*

Whiteside, J., Bennett, J. & Holzblatt, K. (1988) Usability engineering: our experience and evolution, in M. Helander (ed.), *Handbook of Human-Computer Interaction*, Elsevier.