ISO 25062 Usability Test Planning for a Large Enterprise Applications Suite

Sean Rice, Jatin Thaker, and Anna M. Wichansky

Oracle USA, 500 Oracle Parkway MS3op3, Redwood Shores, California 94065 USA
Anna.Wichansky@oracle.com

Abstract. In setting out to perform summative usability testing on a new suite of more than 100 enterprise software applications for 400 different user roles, we faced a number of challenges in terms of staffing, scheduling, and resources. ISO 25062 provided a valuable organizing framework to plan, scope, and implement our testing effort. In this paper, we discuss the considerations and steps that we took in test planning and management, including our prioritization strategy and creation of an automated data collection system to minimize impact on staffing resources and the usability engineering workload.

Keywords: ISO 25062, Common Industry Format, summative usability testing, enterprise software, user experience, test automation.

1   Introduction

1.1   Oracle's Next-Generation Business Applications

Oracle is a key provider of software applications that help run enterprises. Enterprise applications are typically used by large organizations, such as industries, governments, and universities, with many employees performing a wide variety of functions. Enterprise software is developed as integrated program suites that enable organizations to perform financial and manufacturing operations, control governance, and manage internal functions, such as human resources. Enterprise software also enables organizations to interface with their customers and with other organizations for procurement, supply chain management, marketing, sales, and other purposes.

Oracle Fusion Applications is a suite of industry-standard enterprise software applications, providing functionality for many different enterprises and geographies. The key features of the applications in this suite include service-oriented architecture, a role-based user experience, and embedded business intelligence. Oracle Fusion Applications is intended to replace older generations of multiplatform applications currently supported by Oracle. The Oracle Fusion Applications design captures the best user experience features of Oracle's current application suites, such as Oracle E-Business Suite, PeopleSoft, Siebel, and JD Edwards. In addition, Oracle Fusion Applications incorporates the latest in technological features, such as social networking, tagging, and superior levels of product integration to create a seamless, productive, and enjoyable experience for enterprise users.

The Oracle Fusion Applications user experience was five years in the making. The development of this suite included an extensive and comprehensive user experience design process: ethnographic research, low-fidelity workflow prototyping, high-fidelity user interface (UI) prototyping, iterative formative usability testing, development feedback and iteration, sales and customer evaluation throughout the design cycle, and summative usability testing. This paper describes how the summative testing was scoped and planned using ISO 25062, Common Industry Format (CIF) for usability test reports [1] as an organizational framework.

1.2   What Is a CIF Usability Test?

CIF refers to the internationally standardized method for reporting usability test findings used by the software industry. The CIF is based on a formal, lab-based test that is used to benchmark the usability of a product in terms of human performance and subjective data. The CIF was developed and is endorsed by more than 375 software customer and vendor organizations led by the National Institute for Standards and Technology (NIST), a US government entity. NIST sponsored the CIF through the American National Standards Institute (ANSI) and International Organization for Standardization (ISO) standards-making processes. Oracle played a key role in developing the CIF.

The CIF report format and metrics are consistent with the ISO 9241-11 [2] 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." Our goal in conducting CIF tests is to measure performance and satisfaction of a representative sample of users on a set of core tasks and to help predict how usable a product will be with the larger population of customers.

1.3   Why Perform CIF Testing?

The overarching purpose of the CIF for usability test reports is to promote incorporation of usability as part of the procurement decision-making process for interactive products. Examples of such decisions include purchasing, upgrading, and automating. CIF provides a common format for vendors to report the methods and results of usability tests to customer organizations. CIF provides a benchmark of usability for the product and enables customers to compare the usability of our software to that of other suppliers. CIF also enables us to compare our current software with previous versions of our software. Our biggest reason for conducting CIF studies is that our customers expect Oracle to provide CIF reports for our products. CIF is often required as part of a Request for Proposal (RFP), especially in Europe.

2   How to Apply CIF Methods to a Large Product Suite

2.1   Project Scoping

Oracle Fusion Applications comprises more than 100 modules in seven different product families. These modules encompass more than 400 task flows and 400 user

roles. Due to resource constraints, we cannot perform comprehensive CIF testing across the entire product suite. Therefore, we had to develop meaningful inclusion criteria and work with other stakeholders across the applications development organization to prioritize product areas for testing. Ultimately, we want to test the product areas for which customers might be most interested in seeing CIF data. We also want to build credibility with customers; we need to be able to make the case to current and prospective customers that the product areas tested are representative of the product suite as a whole.

Our first goal was to identify what a reasonable minimum would be for the number of CIF tests to conduct. Through discussions with internal stakeholders, including people who were most familiar with requests for CIF data that Oracle had received from customers or sales, we proposed CIF testing for 50 percent of the products in the application suite. It was understood that testing would cover only the top use cases for each product. Tests would not be comprehensive across all areas of the products chosen.

The primary activity in the scoping process was to work with the individual product teams to identify the key products and business process task flows in each product to test. We prioritized these products and flows through a series of negotiations among the user experience managers, product strategy, and product management directors for each of the primary product families within the Oracle Fusion Applications suite (Human Capital Management, Supply Chain Management, Customer Relationship Management, Financials, Projects, Procurement, and Setup).

Considerations for CIF study prioritization included:

- The products identified as most important by the strategy directors from a revenue perspective (that is, the company's strategic financial goals)
- Task flows that the strategy team identified as most important within these products
- Level of involvement of the Oracle Fusion Applications User Experience team in UI design (the User Experience team has limited resources and does not have the capacity to provide UI design for every part of the product suite)
- Previous benchmark testing, including CIF tests for previous versions of Oracle E-Business Suite and Total Ownership Experience (TOE) tests for PeopleSoft applications (we wanted to compare the usability of the new products with the previous versions where we had the data)
- Areas for which formative usability studies were conducted during the design cycle (we wanted to collect performance measures for the product areas where we believed that we could improve usability the most)
- The primary user roles for the task flow (if the product task flow was to be used by two user groups, we counted that as two CIF studies for planning purposes due to the increase in the number of participants needed to complete the usability test)

The end result of the scoping exercise was a list of 47 proposed CIF tests for the product suite.

2.2   Resource Considerations

Since 1994, Oracle has established 20 usability labs in nine locations worldwide. Each lab is supported by a dedicated lab administration and participant recruitment team and is available to all Oracle usability engineers and design teams. Because usability testing of previous applications suites was supported at these labs, it was a given that we had appropriate facilities available within the company to perform testing of the scope required for Oracle Fusion Applications.

Fig. 1. Usability engineers in darkened control room at one of Oracle's usability labs in Denver, Colorado, observe testing through a one-way mirror
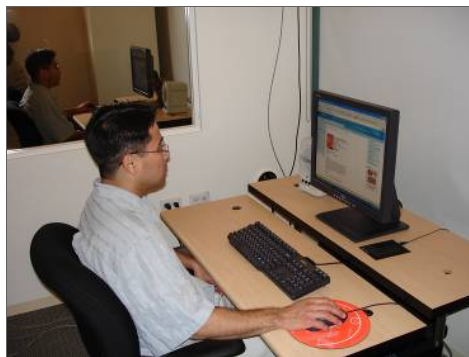

Fig. 2. A test participant completes tasks during a usability test on the lighted side of a one-way mirror in one of Oracle's usability labs

We explored several options for the CIF testing, each with a different impact on time, budget, and other planned user research activities. The staffing options included using existing Oracle usability engineers to conduct the CIF testing, hiring additional

full-time staff, hiring a group of contractors, or pairing interns with usability engineer leads from each product area. Each resourcing option had its pros and cons.

The additional full-time hires option was proposed to management as the ideal option, and the use of interns to test participants was presented as an alternative. Using interns would require a simplified and automated method of data collection and reporting, due to their relative lack of experience with CIF methods and the large number of tests required. Using contractors to conduct the studies was dropped from consideration due to concerns over budget, ramp-up time, and training issues. For all options, we concluded that we would need to hire an additional database administrator to support the product instances (see 2.4 Technical Challenges).

Remote testing (via web and teleconference) was considered as a means to collect participant data. Remote testing can reduce costs and allow for more diverse participant profiling, but can affect product performance due to network lag. There can also be small differences in the way the screens appear via web conference. For instance, we found that the blue color used to indicate table row selection was appearing as a much lighter yellow over web conference.

2.3   Test Protocol Issues

ISO 25062 provides useful guidance on how to standardize many details of the test protocol. Not only does this standardization add a higher degree of experimental control across the wide range of tests, but also it results in improved productivity and economy of test execution.

Here are some areas where we needed to develop policies and procedures not covered in detail by ISO 25062:

- The role of training. Enterprise applications contain complex functionality, and we did not want to measure user performance on a learning curve. Therefore, we had to develop a plan for standardizing the pretest use of in-product user assistance resources, such as help and product overview videos. We also wanted to standardize the familiarization instructions and product orientation time period allowed across all tests.
- The appropriateness of providing assistance. Providing any assistance is strongly discouraged in summative usability testing, due to its effects on the time and completion rate metrics. However, because test participants do get stuck from time to time and are not able to progress to the next step of a task, we adopted our protocol from earlier applications testing and allow two assists. If a third assist is requested, failure and a 0 percent completion rate on that task is noted.
- The use of an expert's benchmark time for comparison. In order to establish a rule for timing out test participants who progress too slowly through a task, in previous testing protocols, we have used expert users' time benchmarks. This expert is often a product manager or other Oracle staff member familiar with the task and software. We generally allow participants three times the benchmark time to complete a task. Thereafter, the user is timed out (and a 0 percent completion rate for that task is noted).

2.4   Technical Challenges

In order to test the Oracle Fusion Applications user experience, an entire integrated suite of related programs from a product family, such as financials, must be installed and configured. Setting up an enterprise application suite requires several pieces of software that work together to provide the suite. This setup requires a technology stack that includes: (1) a database of enterprise data; (2) the middleware, which performs most of the functional calculations and provides the services required for enterprise operations; and (3) the applications being tested. Oracle Fusion Applications runs in a web browser, and customers may use any browser certified by Oracle to work with the applications. The entire integrated suite of applications running on top of the database and middleware layers is known as an instance or environment.

Often the biggest challenge is building an instance that resembles the real world environment at a customer location. This environment must remain stable during the course of the test, and every participant must be presented with an identical environment.

From our previous applications suite usability tests, we determined two approaches to resolve these challenges:

- Setting up a system out of the box (that is, setting up a system from installation media and configuring the system post-installation with appropriate updates and additional set up to mimic the real world environment)
- Creating a clone of a real world environment on the usability labs systems

Both of these approaches require the skills of a database administrator (DBA) to create the applications environment. Creating an applications environment includes setting up the database tier with the appropriate version of the database and ensuring that the requisite database objects and database patches are installed. Following this, the DBA must install and configure the middleware components that are required to service the applications that get deployed into the middleware. The final step in setting up the environment is to deploy the applications that are being tested, and perform any profile tasks that are required to set up the environment to mimic a real world scenario. Profile tasks include setting up users and privileges.

Prior to conducting any testing, the DBA must also ensure that a backup of the database is captured and is available to be used to reset the environment prior to testing each participant. This backup ensures availability of the same exact test environment for all participants.

2.5   Automation to Facilitate Testing

The Oracle usability labs have developed and supported online data collection technologies since 1999. In order to facilitate the large number of tests required for Oracle Fusion Applications, we created a prototype online data collection system using the Oracle database and Oracle Application Express, a tool and development platform to build lightweight applications. Oracle Application Express is shipped along with the Oracle database.

This prototype system interfaces the existing usability lab scheduler with a new usability lab data logger that was built from scratch to automatically produce the CIF test report. The lab scheduler contains information about the logistics of the tests, namely what lab these tests are running in, the characteristics of the lab equipment, the number and type of test participants, and so on. The data logger contains the human performance and subjective data collected during the CIF test. Participants use an internal website to access a licensed version of the Software Usability Measurement Inventory (SUMI) [3]. We are able to capture the necessary parts of the output of these systems on a per test basis and to populate the data into an automated version of the CIF report that contains boilerplate text on the methods and available variable fields for individual and summary data.

All the underlying data is stored in an Oracle database. The information is passed to the applications like the data logger and the lab scheduler on an as-needed basis. This lab scheduler enables the usability engineer to quickly find the appropriate set up required for the usability test, and to look up participant information suited for the study.

The data logger, on the other hand, is used during the course of a study to record task and time information, as well as to record key information about the study. The data logger enables the usability engineer to quickly gather information and use the information recorded during the study to generate an automated CIF report. The automation is handled by integrating the information from the data logger and the lab scheduler.

By building an automated system using Oracle technologies, we can present analysis and metrics-based summaries of the CIF tests clearly and concisely without the usability engineer having to review every segment of the data log and to reinvent report text that is standardized across our labs. The goal is to enhance the productivity of the usability engineers and interns and provide product development teams with the findings and data analysis much sooner than traditional data collection and report creation techniques.

3   Progress and Projections

The automated CIF testing and reporting system has been successfully pilot tested in the Oracle usability labs. Formal CIF testing is expected to be underway for Oracle Fusion Applications before 2012. We project that we will be able to run the 47 CIF tests with a core group of experienced usability engineers supervising interns on up to four Oracle Fusion Applications environments simultaneously. More details will be provided at the conference presentation.

We intend to write a follow-on paper describing the results of this project in terms of the effectiveness of the automation that we have developed for the CIF testing.

4   Conclusions

Applying the CIF standards to summative usability testing for a large suite of software applications that encompass many hundred user roles and task flows is a

challenging usability engineering task. However, economies can be achieved by prioritizing frequently used flows and common user roles to reduce the sheer number of tests required and by automating the testing process itself so that tests can be conducted faster with fewer and less experienced staff members.

References

1. International Organization for Standardization: ISO/IEC 25062:2006(E). Software engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Common Industry Format (CIF) for usability test reports (2006)
2. International Organization for Standardization: ISO 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability (1998)
3. Kirakowski, J., Corbett, M.: SUMI: the Software Usability Measurement Inventory. Brit. J. Ed. Technol. 23, 210–212 (1994)