

Software Quality and the Internet Appliance

Summary

Information appliances, which provide simplified, easy access to specific information such as e-mail and Web sites, promise to bring the benefits of computing to a wide customer base, including some computer-averse people who have hitherto avoided buying a computer. Internet appliances are evolving from personal computers, game stations, digital mobile phones, and server technologies. While this allows us to apply well-known quality assurance techniques, including testing techniques, the software quality professional must remember that the risks to product quality are different; the quality bar is higher, especially in terms of usability, robustness, and harmonizing the appliance with the dynamic Internet. Customers will assess the quality of information appliances by the degree to which the appliance reliably, quickly, transparently, and intuitively provides them with access to the desired information, and we expect them to be much less understanding of glitches than the current PC user. Information appliances are gaining wide acceptance—millions will hit the market in the next few years—so many of us who practice software quality professions will spend time working on projects to develop them. Indeed, we expect that information appliances will present tremendous opportunities to those who seek to bring quality to software in the new millenium. This paper presents the test team's findings on one such project.

Rex Black, President/Laurel Becker, Vice President, RBCS, Inc.
31520 Beck Road, Bulverde, TX, 78163, United States of America
T: +1 (830) 438-4830/F: +1 (830) 438-4831/E: rex_black@rexblackconsulting.com

1. Introduction

Recently, a new breed of computer—and the software that animates it—has captured the attention of computer professionals, computer users, and the computer-averse as well: the information appliance. One can define an information appliance as a limited-feature computer with a simplified, intuitive user interface designed to handle particular nuggets of information conveniently. One type, the Internet appliance, promises to bring people the benefits of the Internet without the difficulties of a general-purpose computer. In this paper, we offer the test team's perspective on a successful project to bring the Internet appliance to the customer with a positive experience of quality.

2. The Information Appliance: Revolution or Evolution?

Twenty-five years ago the personal computer era began as hobbyists and entrepreneurs built small computers around an exciting new technology, the microprocessor. The personal computer revolution released software from the control of centralized IS organizations and sophisticated technicians, and brought with it the promise of universally accessible computing. Xerox PARC's advances in the graphic user interface, marketed so successfully first by Apple and now by Microsoft, seemed to adumbrate effortless, ubiquitous computers penetrating every nook and cranny of society.

This technological utopia may yet come, but the general-purpose computer is not the vehicle for delivering on that promise. As one computer usability consultant, Jakob Nielsen, observed at the Twelfth International Software Quality Week conference, computer acceptance has stalled. The PC, with its power and flexibility, is too complicated for most people, and reliability problems can stymie even computer sophisticates. Another human factors expert, Donald Norman, argues that we need a "new paradigm": information appliances that will be "the natural successor to today's [complex computers]" (Norman, 1998).

The information appliance is an evolutionary means to a revolutionary end. The devices themselves, and the software in them, have evolved from personal computers, game stations, and digital mobile telephones. They connect with typical servers for e-mail, content, and Web access. Despite the technological similarities, though, successful information appliances must be more than just stripped-model personal computers or overgrown games stations or phones. Like toasters, ovens, and refrigerators, information appliances perform simple tasks, but must do them in a trivially obvious and totally foolproof fashion. It is the simplicity of function and use, the transparency of the technology—and the new celebrants such easy-to-use appliances will bring to the computer party—that will deliver a great leap forward in applying computer technology to people's lives.

3. An Internet Appliance Case Study

To move beyond generalizations, we introduce our Internet appliance case study. An ideal Internet appliance presents the features of the Internet, including e-mail and the World Wide Web, to the customer in a way that hides complexity, yet provides full access to the fun, fascinating, and convenient aspects. To enable this functionality seamlessly, the Internet appliance in our case study uses a client/server model. The servers at the ISP and the server farm provide e-mail services, customize and deliver news, weather, and other content to the appliance, filter undesirable Web sites, and provide the appliances with software updates. Figure one illustrates the architecture.

We consider this project a success for two reasons. First, our client has received positive reviews from critics, both technical (Peskovitz, 1999) and non-technical (Dreyfus, 1999). Second, the appliance is enjoying significant acceptance in the marketplace. While not perfect, the device is proving good enough for most customers, including many who have sent their first e-mail and surfed to their first Web site with it.

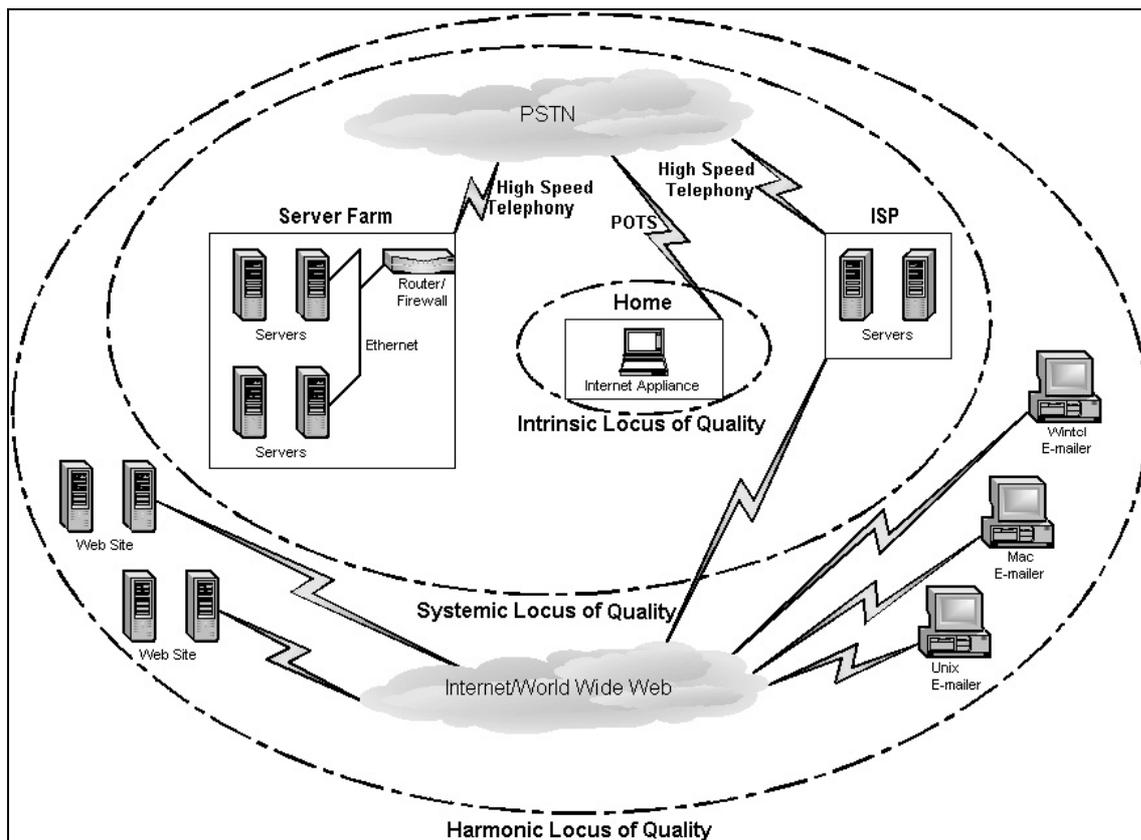


Figure 1: Internet Appliance System Architecture

4. Loci of Internet Appliance Quality

As shown in figure one, we divide Internet appliance quality into three areas, which we call “loci of quality”. We use this phrase because each area contains a set of behavioral points that determine the customer’s experience of quality.

Intrinsic. Those aspects of quality that are strictly a function of the device hardware, the firmware, and the locally hosted software itself.

Systemic. Those aspects of quality that support the functions of the appliance, including the servers, the communications network, and the human aspects of the process.

Harmonic. Those aspects of quality that depend upon how well the otherwise-correct intrinsic and systemic functions interact with the Internet.

Moving from the intrinsic to the systemic to the harmonic locus of quality, the Internet appliance vendor controls fewer factors that influence quality. The vendor determines intrinsic quality just as a PC maker does. Systemic quality, while subject to some influence, depends considerably on the public switched telephony network, the quality of the server components, and the service levels of the ISP. For harmonic quality, the vendor must respond to the official and unofficial standards of the Internet, its Web servers, e-

mail applications, and its netizens.

5. Testing Techniques and Customer Quality Alignment

The testing process for Internet appliances is a straightforward evolution of standard test methodologies. The testing of the appliance itself resembles testing of a laptop computer, while one can test the server farm as one would test an information systems project at a bank. Like both of these situations, the prospective test manager will need to attend carefully to the logistics of the underlying hardware. Custom appliances, especially prototypes developed in off-shore facilities, are hard to come by early in the project as well as being temperamental, subject to failure, and difficult to maintain. Large server farms may contain hundreds of thousands of dollars worth of equipment. The first-time test manager for an Internet appliance project may want to refer to one author's book, *Managing the Testing Process*, for tips on how to handle these challenges (Black, 1999).

To plan our test effort, we started by uncovering the risks to product quality. Working cross-functionally in the organization, we put together a list of about 75 specific types of failure modes—such as “appliance won't boot,” “slow e-mail transfers,” etc.—in the following categories of quality risks:

- Functionality
- Reliability
- Performance
- Security
- Operations
- Error and disaster handling and recovery
- Capacity and volume
- Data flows and data quality
- States and state transitions
- Untested code
- Date and time handling
- Localization
- Appliance configuration options
- Documentation, tutorials, and help screens

Understanding the quality problems that might affect customers, the test engineers then set about developing test cases to provoke these failure modes. The engineers again worked cross-functionally to figure out how the product should behave in certain typical and atypical scenarios. These discussions allowed us to define over 1,000 test conditions.

Two areas of testing presented unique challenges. First, the appliance exchanges datasets with the server farm that may contain executable software such as the operating system, the browser, the address book, or the e-mail applications, non-executable data (called “content”) in the form of e-mail, address book entries, news, or weather, or any combination of these sets of bits. Corruption or loss of any of these items could impair the operation of the device. Indeed, damage to a new OS or application could render the appliance entirely and permanently unusable.

To test this possibility, we put the appliances through extensive update tests. These tests covered storage limitations, power- and connection-failures, deliberately corrupted operating systems and applications, and other error-forcing situations, as well as normal uploads and downloads. We tested tens of thousands of executable (OS and application) update events and data (content, address book entry and e-mail) update events during the System Test effort.

One aspect of this kind of testing we recommend especially is end-to-end testing. For example, if the devices ships directly from the factory with the software installed, make sure the first “boot, connect, and update” sequence will work properly. Errors here will render the appliance dead on arrival, a decidedly negative customer experience of quality.

The second challenge involved usability. While usability is an important quality concern for most PC software and hardware vendors today, they have standards to guide their efforts. Microsoft Windows applications adhere to the Microsoft look and feel, while Apple applications use the Macintosh user interface style. These interfaces are too difficult and complex for our target customer, so we started over. Experienced usability designers worked and reworked the interfaces for the home screen, the content, the browser, the e-mail application, the address book, and the overall navigational paradigms. Early prototypes were taken to locations like airports for target market feedback.

Through an extensive Beta program that included hundreds of users, we obtained more insight from computer-averse users. The Beta program began before System Test, and continued up to system release.

Finally, we made sure that the test team saw the product through customer's eyes. We held frequent discussions about what a computer-inexperienced customer's reasonable expectations would be. The test engineers and test manager sensitized the test technicians to the need to remain critically aware of the user interface as they ran tests, and to verify correct behavior beyond simply the expected results of the test case.

To ensure that we didn't let our previous experience with—and acceptance of—PC quality glitches color our findings, we all adopted an active bias towards reporting problems. Any behavior that might mislead, bewilder, intimidate, or anger a customer, regardless of why it might happen, we reported as a bug.

To implement the tests, we used a combination of manual test suites with automated support tools. Test engineers wrote manual test cases that test technicians keyed in and evaluated on the appliances. We used programs running on load generators to create representative, stress, and peak usage profiles on the server systems. Figure two illustrates the approach, omitting some components to focus on the elements that interfaced with testers or test tools. We had about six to eight appliances active sixteen hours a day, with test technicians and engineers running tests in two shifts, while the load generation programs ran twenty-four hours a day.

For both automated and manual testing, our techniques were primarily behavioral (black box). The development team performed extensive unit and component testing, applying structural (white box) tests to their code. In this way, our independent test team efforts complemented their own testing, resulting in a bug find ratio of 30:70, respectively.

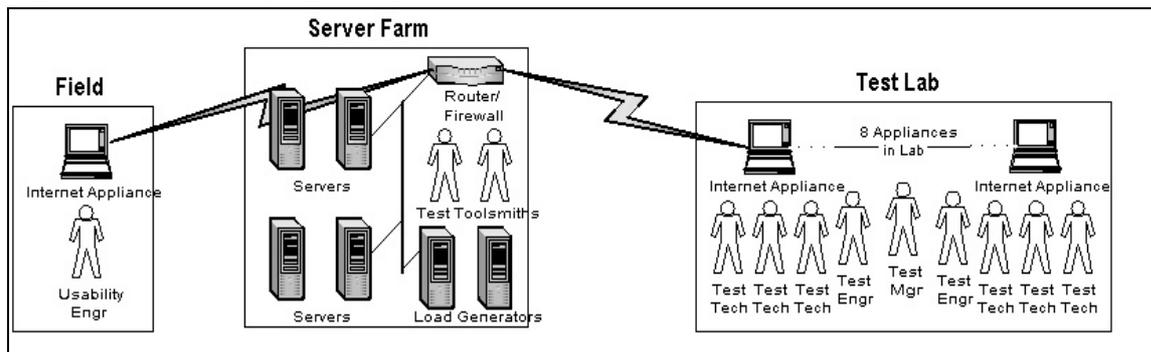


Figure 2: Test Approach

Figure two also illustrates our test team composition. The two in-house test engineers focused on developing and supervising the execution of test cases. The six test technicians ran the test cases, reported results, and handled “pushing” software updates to the devices. One usability test engineer focused on human factors and Internet harmonization issues. Two test toolsmiths developed and ran the automated load and performance tools.

We hired people who understood the customer's outlook. Our test team was composed almost entirely of seasoned behavioral test engineers, non-technical usability professionals, and technical support professionals; only the test toolsmiths and the server test engineer were experienced software developers.

6. Bugs: Where They Arose, What They Affected

In terms of defects (bugs), it is interesting to look at where the faults—the actual code errors—occurred and contrast that with where the failure manifested itself. These analyses are shown in figures three and four. Looking at defect root cause data for both structural and behavioral testing in terms of the culpable subsystem—where the fault was found—we found that about 2/3rds of the bugs resided in the intrinsic locus of quality, 1/4th in the systemic locus, and 1/12th in the harmonic locus. While we do not have new-and-changed function-point or line-of-code metrics available, these proportions fit intuitively with amount of development effort expended on the client software as compared to the less-extensive customization of commercial off-the-shelf systems on the server side.

For the behavioral testing alone—which more closely resembles customer usage—if we look at which locus of quality experienced the failure mode, we see a different picture. Slightly less than half the defects found by the test team occurred during intrinsic tests, and only a little more than 1/5th of the defects surfaced during systemic tests. About three in ten defects were found as part of harmonic testing.

7. Considerations for Internet Appliance Software Quality

The subject of defect data brings us to the lessons learned about achieving quality for an Internet appliance. We present here some interesting and perhaps surprising discoveries about factors that support a seamless customer experience of Internet appliance quality:

- There is such a thing as “enough” CPU, memory, and local storage. Unlike the typical PC, we found that limited-functionality applications both presented a simpler interface and responded with acceptable performance in tight quarters.
- There is, however, no such thing as “too much” bandwidth. Telephone lines, while ubiquitous and

non-threatening to computer-averse customers, provide inconsistent speed and connection capabilities with analog modems, which makes the data communication slow and fickle. Telephone line management features like call waiting, off-hook detection, and central office voice mail systems provide additional challenges.

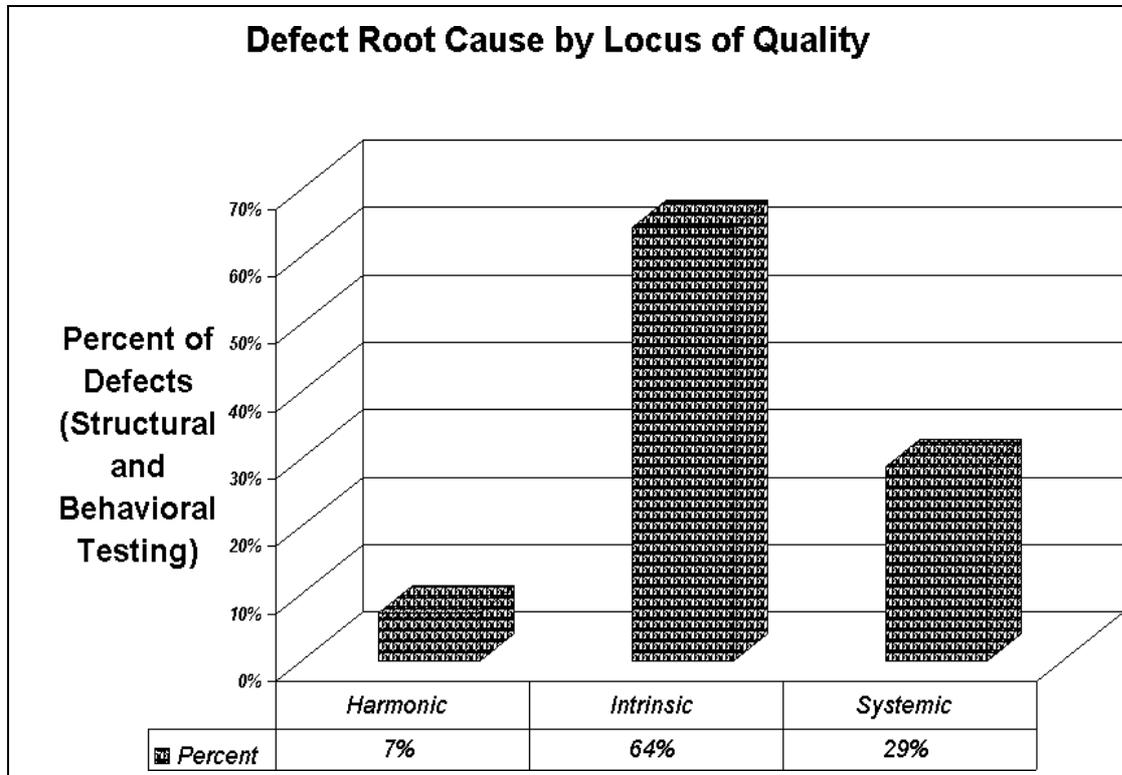


Figure 3: Defect Root Cause (Bug Location) by Locus of Quality

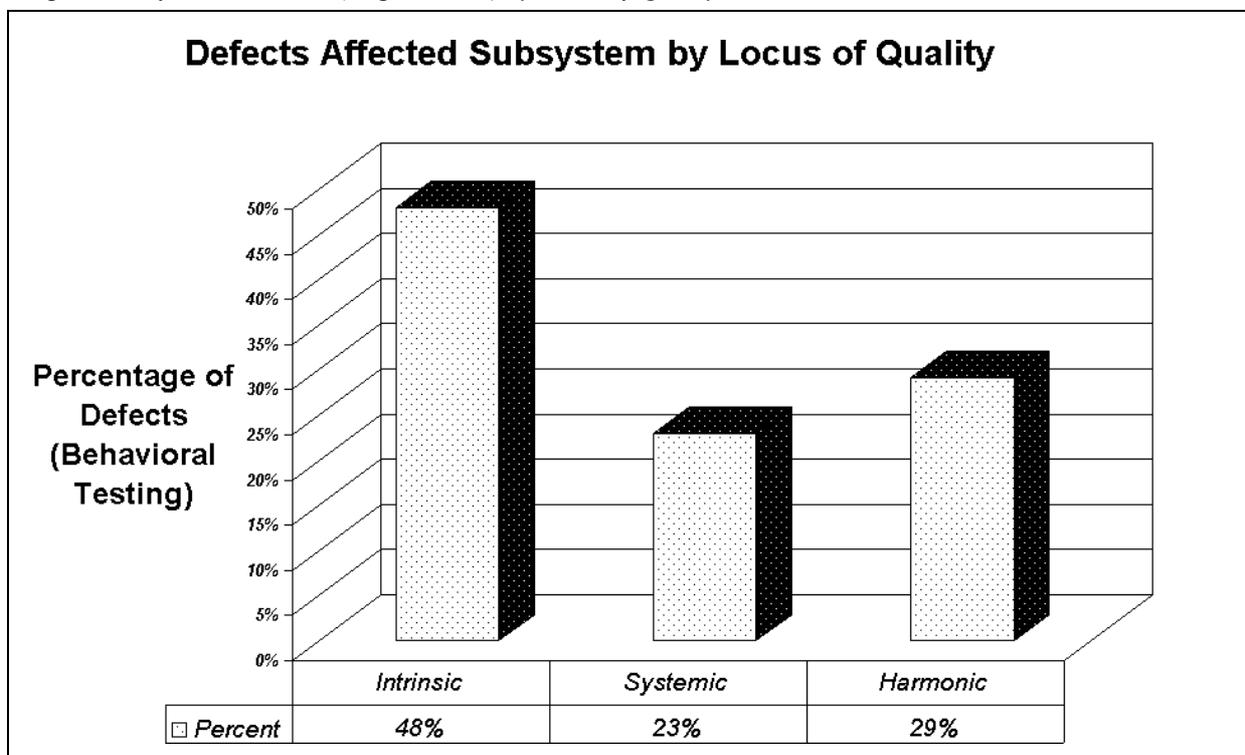


Figure 4: Defect Affected Subsystem (Symptom Location) by Locus of Quality

- Getting the UI right is hard. Usability studies, Beta programs, frequent interface tweaks, and a few complete redesigns of some applications were needed to obtain a simple user interface. The hardware

elements are critical also; the wrong keyboard or mouse can make the most intuitive workflow and screen presentation hard to use.

- Reliability and stability don't happen by accident. An Internet appliance is hardly an "invisible computer" if the browser dies, the e-mail application goes dead, or updates of software or data render the appliance a cute but non-working paperweight. We spent a lot of time testing at boundary conditions to ferret out these kinds of problems.
- Performance modeling and load testing go hand-in-hand. You can't be too careful about making sure your server farm has enough capacity. We started with a combination of performance simulations, based on certain assumptions about usage profiles. During testing, we checked the results of this simulation, and, where we found discrepancies, investigated them. This provided us with a high degree of confidence that unforeseen capacity issues did not await us.
- Localization is strategic and involves more than just standards and certification. A recent report by Forrester Research indicates that Web e-commerce will grow 100% in Europe every year for the next five years (Spiegel, 1999). Beyond just testing the hardware for regulatory compliance, one must make sure that the entire system—including ordering, provisioning, and fulfillment—will accommodate language, time display, time zone, cultural issues, and so forth.
- The Web waits for no one. Web technologies, such as Javascript, Shockwave, and RealAudio, continue to evolve rapidly. Web sites launch these new technologies on the World Wide Web within hours of their advent. One must select a device, an operating system, a browser, and other applications that can harmonize with the Web now and can adapt to new Web technologies. We tested the harmonic aspects of the appliance heavily, and, as our findings show, this was fertile bug-hunting territory.
- Attachment support decisions require trade-offs. There are hundreds of file types floating around on the Internet. Supporting all of them is impossible, even on a PC. The simpler Internet appliance must support a subset of these transparently, then handle the unsupported ones elegantly.
- Filtering software remains imperfect. Any Internet appliance that targets kids or teenagers will need to address the issue of what Internet content can be safely delivered to children. This goes beyond simply sexual imagery. The Web, as we found in our testing, is rife with what many in our society would deem hateful, dangerous, or violent pictures and text. There are also many sites that, while well-intentioned, may provide curious children with information their parents deem inappropriate.
- E-commerce is still flaky. A recent study by Andersen Consulting found that, out of 480 purchases on 100 sites, only 350 were successfully completely (Orr, 1999). Our findings were similar.

To summarize, one can say that quality, for an Internet appliance, is determined by the extent to which the device and the supporting system become invisible to the customer, transparently enabling and simplifying the Internet features he or she wants to use.

8. Beyond the Case Study

Far from being a narrow case study, software quality professionals can generalize our techniques and findings to a wide and rapidly expanding world of Internet and other information appliances. Both the technical and the non-technical press are giving developments in the information appliance arena a lot of ink. For example, two articles about Internet appliances in a popular computer magazine dealt with the use of digital wireless and in-flight phones as Internet appliances (Nash, 1999). One article predicts a growth from 1.1 million Internet-ready phones now to 79.4 million such devices in 2003 (Grotta, 1999). A national news magazine published a major article—including a side-bar on our client's system—about the evolution of the information appliance (Holstein, 1999). Software quality professionals helping to develop these devices and services will face challenges in connection performance and reliability, usability and user interface design, cached content, customer-transparent software upgrades, server load simulation, extensive manual testing, and harmonizing with the Internet.

9. Conclusions

Information appliances not only make it easier for anyone to access information, they can also extend computer technology to people who have rejected the complexity and capriciousness of the general-purpose computer. Our challenge, as software quality professionals, is to understand the key quality factors for such appliances; these factors will differ from those of traditional personal computers, game stations, and digital wireless phones. We believe that, for information appliances, customers will judge quality by the degree to which the appliance reliably, quickly, transparently, and intuitively provides them with access to the desired information. Those testing an information appliance can apply the same concepts and techniques as testers on typical PC and information systems projects, but they must apply these ideas and

methods in the service of different goals, looking for different kinds of defects. The critical importance of harmonizing the device and the overall system with the ever-changing and ever-expanding Internet is especially interesting. We anticipate an explosion of information appliances over the next few years, and many of us in the computer business will work on projects to develop them. The information in this paper will help software quality practitioners who take on ambitious information appliance endeavors.

10. References

Donald Norman: *The Invisible Computer*. Cambridge, Massachusetts: The MIT Press, 1998. ISBN 0-262-14065-9.

David Pescovitz: "Gadget: I-Opener." *The Industry Standard* (Internet Web-site), December 22, 1999.

Joel Dreyfuss: "Want to Go Online? This Appliance Can Take You There." *Fortune* (Internet Web site), December 20, 1999.

Rex Black: *Managing the Testing Process*. Redmond, Washington: Microsoft Press, 1999. ISBN 0-7356-0584-X.

Rob Spiegel: "Europe Closing E-Commerce Gap With U.S." *E-Commerce Times* (Internet Web site), December 21, 1999.

Andrea Orr: "One in Four Online Purchases Thwarted." Reuters Internet newsfeed, December 20, 1999.

Sharon Nash: "Hand-Held Shopping Mall." *PC Magazine*, December 1, 1999, page 32.

Daniel Grotta, et al: "Fly the Web." *PC Magazine*, December 1, 1999, page 12.

William Holstein: "Moving Beyond the PC." *US News and World Report*, December 13, 1999, pages 49 through 58.

11. Authors' Biographies

Rex Black and Laurel Becker are the principal consultants of Rex Black Consulting Services, Inc., an international software and hardware testing and quality assurance consultancy, and have over three decades of professional experience between them. RBCS's clients include Dell, Sun, Hitachi, Motorola, Tatung, Netpliance, General Electric, First USA, Omnipoint, and Clarion, among others. Their work with these clients has taken them to Taiwan, Hong Kong, Japan, the U.K., Canada, Germany, Holland, France, Spain, Switzerland, and Italy, along with locations throughout the United States.

Rex is the author of *Managing the Testing Process*, published by Microsoft Press. He frequently presents original papers at conferences such as Software Testing and Analysis Review, Quality Week, Practical Software Quality Techniques, and the World Congress on Software Quality. Rex also presents training seminars on testing and quality assurance management, including certification courses through the International Institute for Software Testing. Rex holds a B.Sc. in Computer Science and Engineering from UCLA. He belongs to the Association for Computer Machinery and the American Society for Quality.

Laurel, having a strong background in human behavior and psychometric testing, focuses on usability studies, human factors testing, and consulting team dynamics. She helps clients narrow the gap between their products and human behavior and needs. She received her Bachelor of Arts degree in psychology from UCLA and her Masters of Education with a specialty in Counseling and Guidance from the University of Texas at San Antonio. Prior to joining RBCS, Laurel worked in the fields of clinical and research oriented application of human behavior skills. Laurel is a Licensed Professional Counselor and a member of Texas Professional Counselors.