Website Testing

Abridged extract from
The Website Manager's Handbook
(Chapter 3) by *Shane Diffily*

"Genuinely useful for helping think through the key issues of website management"
Gerry McGovern. Author of *Killer Web Content.*

Website Testing

Website Testing is a process for evaluating the conformance of a site to an agreed set of guidelines. The purpose of testing is to ensure a website is capable of operating to a minimum acceptable standard in order to meet the Goals that have been set for it.

Unfortunately, some organisations view this phase of development as an unwelcome delay that can prevent their project finishing on time. Judging by the number of sites that are launched with such basic errors as broken links or missing images, second-rate testing appears to be the norm. This is in stark contrast to the often rigorous sign-off procedures that are followed for other media.

For example, no business would ever dream of issuing a printed brochure before thoroughly checking it for errors in spelling, imagery or layout. Yet, many websites are launched after only the most cursory of testing. It is simply taken on trust that everything will be OK. The trouble with this is that site visitors are left to pick up the pieces when things go wrong. Inevitably this can damage the perceived trustworthiness of an organisation.

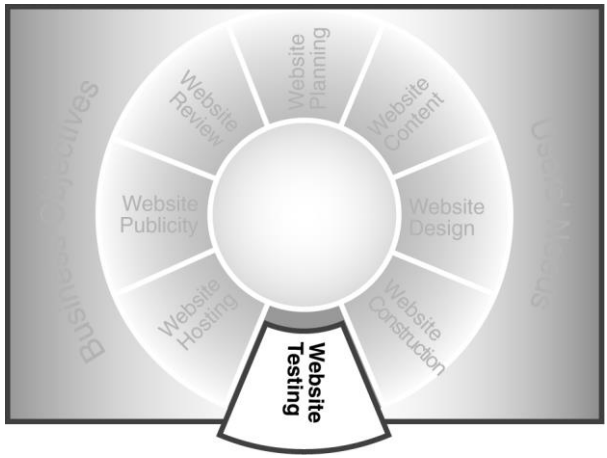

Figure 1. Website Testing as a phase of the Website Development Cycle.

What is needed is a change in mindset—away from one that sees testing as an obstacle, towards one that sees it as a facilitator of site Goals. A possible way to achieve this is to demonstrate the value that testing can add to a site. For example, the web guru Jakob Neilsen has established that by spending 10% of a project budget on usability testing, the quality of a visitor's online experience can improve by up to 135%[1]! Imagine applying this to a website whose revenue relies on credit card transactions, e.g. Amazon.com. The easier the site is to use, the more money can be collected.

The Website Testing Catalogue

Yet, usability is only part of the story. Website Testing encompasses many other areas—ranging from simple spell checking to full security reviews. For convenience, these can be grouped into a catalogue that lists all appropriate methodologies.

[1] Useit.com "Return on Investment for usability!
http://www.useit.com/alertbox/20030107.html January 2003. Accessed December 2005.

Test Method	Description
Code Testing	This tests that all languages conform to accepted code standards.
Design Testing	This tests that all pages conform to the website's preferred layout and design.
Spelling Testing	This tests that HTML and other code has been inserted in an optimal manner.
Hyperlink Testing	This tests that all links to all documents and assets resolve correctly.
Page Weight Testing	This ensures that all pages conform to the maximum allowed page weight.
Browser Testing	This tests that the website displays correctly across target browsers and Operating Systems.
Usability Testing	This ensures that the website conforms to appropriate practice in the area of usability.
Accessibility Testing	This ensures that the website conforms to the stated level of accessibility outlined in the organisation's Web Accessibility Policy.
Security Testing	This tests that the website operates with minimum risk in a secure environment.
Functional Testing	This tests that the website operates as expected under normal and error inducing conditions.
Performance Testing	This tests the responsiveness of the website to user actions.
Website Standards Review	This reviews the website against the organisation's Website Standard.
Operational Monitoring	This puts in place procedures for the ongoing monitoring of the site.

Figure 2. Website Testing Catalogue.

The overall co-ordination of these activities is the duty of the Development Team Leader. On her shoulders rests responsibility for ensuring everything is in proper working order. She may also carry out various aspects of testing herself, notably the Website Standards Review.

However, in most circumstances testing is performed by specialists from within the Development Team.

The Skills and Resources Needed for Testing

For example, Functional Testing may be undertaken by Developers, and Performance Testing by technical personnel. Where a team is large enough, it is advisable to get hold of staff who have not been directly involved in a project and ask them to carry out such assessments. This ensures familiarity does not lead to errors being overlooked.

It may also be desirable to seek external assistance for specialist disciplines like Accessibility and Security. This is particularly necessary where in-house skills are not good enough. Some areas of testing may even entail the participation of site visitors, e.g. usability, where data about user experience is important.

Finally, the procedures of site testing themselves require an assortment of technology in order to occur. This can include anything from simple office stationery (pen and paper) to specialist evaluation software. The degree to which these are needed depends mainly on the scale of the site to be assessed. For example, a large Transactional website is likely to need more technology than a small Basic Site. Of course, budget constraints also set limits to what can be provided.

Now that we understand what is needed for Website Testing, we can start to explore the processes and procedures by which it is carried out.

Code Testing

As the first task in the assessment catalogue, Code Testing ensures that the basic components of a site are in conformance with accepted standards. This includes:

- MarkUp (HTML/XHTML)
- StyleSheets (CSS)
- Client-Side Scripting (ECMAscript/JavaScript)
- Server-Side Scripting (PHP, ASP, JSP, PERL, etc)

An assessment like this is required because improperly authored code can lead to problems of presentation and functionality on some user agents, notably smartphones and screen readers.

Several tools are available to assist this review. For example, the W3C provides an online validator for assessing MarkUp and StyleSheets. To use this validator:

- Select the address of the page you wish to validate
- Visit http://validator.w3.org
- Insert the address of the web page you wish to validate and click 'check'

The page is then checked against the appropriate standard. If it fails to comply, a list of issues is displayed that can then be used for correcting errors. In circumstances where the file you wish to evaluate is offline (i.e. it has not yet been published on the internet) it can be checked by uploading the code to the validator.

Needless to say, this method of page-by-page assessment can be very tedious when a large website is involved. Thankfully, many of the Quality Assurance tools reviewed in Chapter Two (page **Error! Bookmark not defined.**), provide functionality that speed it up. For example, WebQA from Watchfire and Website Monitor from HiSoftware include modules for evaluating compliance with Web Standards. Similarly, the authoring tools Adobe Dreamweaver and Expression Web Designer contain reporting functions for validating MarkUp.

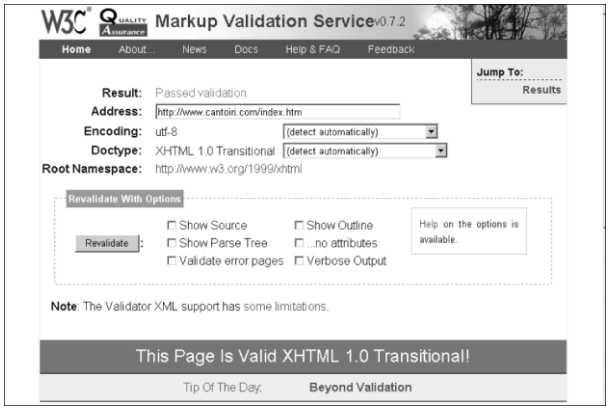

Figure 3. The MarkUp validation service from W3C.

The assessment of Client-Side and Server-Side Scripting languages require more specialist programs. Known as Integrated Development Environments (IDE), these are used to identify and fix bugs. Two popular examples are Microsoft Visual Studio and Eclipse.

Design Testing

The purpose of a Design Test is to ensure that each page on a website is in conformance with the templates agreed for it during development. This includes Information Architecture, Navigation, Interaction, Interface, Information and Visual Design. The basic procedure is to review a site and try to locate unplanned changes in structure or appearance, e.g. distended layout, missing images or inappropriate colours. These can then be corrected as necessary.

The execution of a Design Test is intensively manual, simply because it requires every page to be analysed individually. Technology is no match for the human eye in this regard!

That said, some technology (in the form of user agents) is needed in order to mimic the experience of site visitors (desktop computer, smartphone, PDA, etc). By employing such devices a Designer can be sure that she is viewing a site in the same way as an ordinary user. This

means that any observations can be considered accurate (notwithstanding the issue of browser compatibility which is explored below).

Spelling and Grammar Testing

Nothing on the web appears more amateur than carelessly written text. As such, a detailed focus on language is essential for maintaining a professional appearance.

The bulk of responsibility for this activity lies with the Website Editor. As seen in Chapter Two, the resources needed to assist this task include dictionaries, thesauri and grammar guides. Word processing programs such as Microsoft Word also provide useful functionality. As a result, it is usually possible to prevent poorly written content going online.

However, no process is flawless and bad spellings can sometimes escape notice. In this case, website Quality Assurance tools like those created by HiSoftware or Watchfire (page **Error! Bookmark not defined.**) can be useful. These include spell checking capability, as do many Website Content Management Systems. By scheduling a regular review with such a device, any errors that were overlooked can be corrected.

Hyperlink Testing

The humble hyperlink is probably the main reason for the overwhelming success of the World Wide Web. For example, the ease with which fragmented information can be linked together has revolutionised knowledge sharing. Yet, while hyperlinks are a cause for celebration, they can also be a source of considerable frustration when they point to pages that no longer exist!

Fortunately, a substantial array of Quality Assurance software is available to help detect such faults. Many of these were explored in

Chapter Two (page **Error! Bookmark not defined.**). In addition, several web authoring programs (e.g. Adobe Dreamweaver) and Website Content Management Systems have similar capability. In fact, because of the technology used within WCM, manufacturers claim it is impossible for broken links to arise. This is because a WCM system can automatically detect and delete links to pages that have been removed.

Page Weight Test

A Page Weight Test seeks to ensure that anything placed online conforms to a maximum allowed filesize (usually expressed in kilobytes). For example, the maximum recommended weight for any page primarily used for navigation on the web (e.g. a homepage) is 60 Kilobytes (kB). This limit is based on the time it takes to download a file of this size over a standard 56.6 Kilobits-per-second (kilobit/s) modem connection[2]. 60kB takes about 8 to 10 seconds, which research has shown is the limit of patience for the majority of web users[3].

Non-navigation pages, such as those with paragraphs of text, can safely extend up to 100kB. This is because visitors can begin to read a long page while the rest of it is loading. As the uptake of high-speed broadband grows, this limit will gradually increase.

It should be noted, however, that the figure for page weight must be calculated from all the files being viewed—not just the basic HTML. As such, a web page consisting of a 45kB HTML file, two images of 15kB each and a StyleSheet of 3kB, has a total weight of 78kB.

As with many aspects of Website Maintenance, page weights can be measured using Web Quality Assurance software. Similarly, Content

---

[2] 56 Kilobits (56×10^3) equals about 56,000 bits.

[3] Useit.com "The need for speed". http://www.useit.com/alertbox/9703a.html Accessed January 2006.

Management Systems can also set restrictions on the size of files permitted to be published.

Browser Compatibility Test

As previously discussed, browser compatibility remains an active issue because of the variety of devices that now come with web capability. Indeed, the diversity of internet enabled user agents is forecasted to increase dramatically in the coming years as more and more gadgets go online, e.g. GPS systems, video game players, iPods, etc. The purpose of Browser Compatibility Testing is to ensure that a site can display and function in a useable way on all such appliances. At present, this mainly encompasses desktop computers and (increasingly) mobile devices.

To assist Browser Compatibility Testing, some organisations create 'Test Labs' in which a variety of user agents can be assessed. These user agents are chosen to reflect the devices that are preferred by website visitors. This might include desktop computers like a Microsoft Windows PC, an Apple MAC and a Linux machine (and perhaps even a UNIX or Sun OS box, if a technical audience is targeted). The lab may also contain a range of PDAs, smartphones and Web TV, Playstation or Xbox devices. Each of these is then loaded with the browser software used by visitors.

The following list shows the range of devices and browsers that could be included in such a lab.

Windows XP, 2000, NT, ME, 98, 95	Apple	Other, i.e. Linux, Unix, Sun, PDA, Mobile Phone, Web TV.
Browsers include: Internet Explorer Netscape Firefox Opera Mozilla AOL browser	Browsers include: Apple Safari Internet Explorer Netscape Firefox Opera Camino	Devices & browsers include: Firefox Konqueror Mozilla Netscape Chimera Opera Smartphone/PDA Internet Explorer PDA Thunderhawk PDA Blackberry™ Web TV Windows Ultra-Mobile PC Playstation® Xbox® Jaws® Screen Reader

Figure 4. Equipment to be included in a compatibility test.

Because many organisations cannot afford such comprehensive test suites, other help has become available. For example Openwave.com has downloadable software that allows mobile devices to be mimicked and tested on a desktop computer. Similarly, we have already learned about Browsercam.com (page **Error! Bookmark not defined.**) which allows the appearance of a website to be evaluated on many different platforms via the internet. Facilities like these are very useful when development resources are restricted.

The procedure of compatibility testing itself involves a review of site content for conformance against a set of design and functional specifications, across all selected browsers.

As this takes place, it may be noticed that content is presented well on some systems, but poorly on others. That is, a site may appear exactly as planned in the Mozilla Firefox browser, but be less than optimal in Internet Explorer. While the challenge for Developers is to ensure a consistent online experience, this is not always possible. As such, a mechanism that allows a site to 'degrade gracefully' is required.

'Graceful degradation' is a concept that declares as long as a visitor can read content and use applications properly, the lack of pixel-perfect layout may be overlooked. In this sense, browsers that are planned for 'graceful degradation' should be those that are least used by visitors.

Usability Test

A Usability Test is the measure of the quality of a visitor's experience when interacting with a website. The web guru Jakob Neilsen has defined usability as encompassing five factors. These are[4]:

- Ease of learning
- Efficiency of use
- Memorability
- Frequency of errors
- Personal level of satisfaction

Because of the variety of issues involved, there is no single test that can be defined as **the** usability test for a website. Rather this area

[4] Useit.com "Usability 101: Introduction to Usability" http://www.useit.com/alertbox/20030825.html August 2003. Accessed July 2005.

encompasses a range of assessment techniques that together seek to improve overall performance. We have already become familiar with some of these. For example, Card Sorting (page **Error! Bookmark not defined.**) is a usability technique for building an Information Architecture.

It should also be clear that usability testing does not commence only when the construction of a website is complete. Rather, it occurs in tandem with the Development Cycle itself. Some of the most common procedures used during production are explored below:

Website Planning Phase

Expert Review

An expert review engages an experienced usability consultant to assess a website against the parameters of good design practice. Expert Reviews are often used as a starting point when initiating a redesign project. Such a review might also be undertaken when a website is created for the first time. In this circumstance, the focus is on competitor sites, in order to gather lessons about what makes them so successful.

Personas

As we have seen, a Persona is a profile of an imaginary user who encompasses all the characteristics of a target audience. If several audiences exist, several Personas may be needed. For example, the Personas for a business listed on the stock-exchange may include professional investors, as well as ordinary customers. Personas are a proven way of keeping a Development Team focussed on users' needs.

Survey

A survey is a useful way of gathering opinions from a website audience about their expectations for a site. A survey can also be used to collect views within a business for an intranet development project.

Focus Group

A focus group uses many of the same techniques as a survey, though an invited audience takes the place of a random sample.

Hallway Surveys

A Hallway Survey is a technique used for evaluating intranet designs within an organisation where the target audience is collected together. In this method, a design is displayed in a public area, e.g. a canteen, under the supervision of a manager. Passing staff are then asked to participate in simple task assessments or asked for their opinions based on a series of preplanned questions.

Design Phase

Card Sorting

As we have seen, card sorting is very effective for building an Information Architecture. Not only can it be carried out without any technology, it is also useful with small groups who are representative of a wider audience.

LoFi Task Assessment Exercise

A LoFi task assessment exercise is a technique for evaluating design assumptions. In this method, an outline Website Design is created on paper and presented for assessment by users, e.g. wireframe.

HiFi Task Assessment

A HiFi assessment mimics the approach of a LoFi assessment, however the design is presented in a more sophisticated manner, e.g. as full colour graphics or in simple HTML. As with the LoFi assessment, the objective is to ensure that planned tasks can be successfully completed.

Final Expert Review

A final expert review of a website may be conducted to identify any last minute usability improvements that can be made.

Usability Lab

Many of the resources required for usability testing are very cheap. For example, both Card Sorting and LoFi assessments require nothing more than pen and paper. The most expensive resource in such circumstances is the time needed to host the sessions. However, more sophisticated techniques require additional expenditure. Indeed, some large organisations choose to invest in specialised usability labs explicitly for this purpose.

A usability lab is a room that comes equipped with all the facilities needed to carry out comprehensive usability testing. This normally includes several internet user agents, a video camera (to record user experiences) and observation points where design staff can watch tests without interfering in them.

Usability software from firms such as TechSmith® (from $200) can assist the monitoring of such tests. Programs of this type track mouse movements on-screen via video. This allows user actions and expressions to be evaluated together. It also means that sessions can be recorded and played back at any time.

For organisations where usability is a key aspect of success, e.g. an online bank, a lab like this represents a sensible investment.

Web Accessibility Test

The evaluation of a website for accessibility is relevant only if this feature was stipulated as a Deliverable at the site planning stage. However,

given everything that we have learned about Web Standards, the law and the benefits of accessibility, this should be treated as given.

The purpose of an Accessibility Test is to evaluate the compliance of a site to established standards. These standards may be expressed in law (as in the UK and USA), or refer to international guidelines like those of the WAI. The actual task of evaluation is carried out by Developers, though there are significant advantages to employing specialist evaluation firms for this work.

Specialist Accessibility Assistance

For example, as we saw on page **Error! Bookmark not defined.**, WCAG 1.0 is self-accrediting. That is, you decide for yourself if your site is compliant or not. Needless to say, this can lead to the temptation to award compliance even if some issues have been missed. Similarly, the methodologies by which accessibility is assessed are constantly evolving. Only experts in the field can know which are acceptable to the disabled community. Finally, some aspects of evaluation require specialist tools to be implemented effectively. A dedicated service provider is much more likely to have such resources at hand.

Although the cost of hiring an accessibility specialist may be prohibitive, independent confirmation can be taken as proof that your organisation is serious about supporting users with disabilities[5]. This in itself may be useful as a marketing tool.

Accessibility Review Process

However, even if the work of evaluation is implemented externally, it is still worthwhile understanding the process to be followed. In this regard, the recommendations of the WAI are particularly beneficial. While these

[5] The Irish utility company ESB validates compliance with WCAG 1.0 by employing an independent evaluator, http://www.esb.ie/main/home/accessibility.jsp

are only intended to ensure adherence to the WCAG 1.0 standard, they are useful for a general review.

Step 1. Identify the standard with which the website aims to comply

For the WAI, this means compliance with WCAG 1.0 Level A, Level AA or Level AAA (though the standard might also be stipulated in law, as in Section 508 or the UK Disability Discrimination Act).

The specific criteria to be adhered to are available as a series of checkpoints that can be used by Developers when constructing a site. The checkpoints for WCAG 1.0 are available online at www.w3c.org/WAI.

Step 2. Identify the pages that will comply with the standard

Sometimes it is not feasible for an entire website to be compliant with an accessibility standard. For example, legacy content may be very expensive to convert. In this regard, the WAI allows sections to be excluded from compliance, as long as such exclusions are clearly notified to website visitors.

Step 3. Use an automatic evaluation tool to gauge compliance

Some of the most widely used accessibility evaluation tools include Bobby from Watchfire, Wave from WebAIM and A–prompt from the University of Toronto. In general, these work by trawling a site and assessing each page against the WAI standard. This includes text equivalents (alt tags) for images, the coding of data-tables and document declarations. Watchfire also includes aspects of Bobby technology in the Quality Assurance suites WebXM and WebQA. In addition, the web authoring packages, Dreamweaver from Adobe and Web Designer from Microsoft incorporate basic accessibility reporting tools. Some Content Management Systems can also be configured to check for compliance.

Step 4. Undertake a manual evaluation of website content

Several aspects of website accessibility are quite subjective, meaning that tools such as Bobby can incorrectly label good content as inaccessible. As such, a manual review is necessary before compliance can be finally certified.

A manual review requires the use of a user agent, such as a desktop web browser. The purpose of the review is to mimic the experience of a visitor with a disability. For example, older people find small text hard to read. Therefore, a simple check of accessibility is to establish if text can be increased in size.

For those with more profound impairments, e.g. blindness, a more thorough evaluation is required. In this case, the computer screen could be turned off and the mouse unplugged. The objective in this case is to establish if it is possible to navigate and read the website in the same way as someone with no vision. To assist this, it is also necessary to invest in a screen reader.

A **screen reader** is an assistive technology that allows people with visual impairments to browse the web. A screen reader works by dictating text on a web page aloud to visitors. JAWS by Freedom Scientific (www.freedomscientific.com) is a leader in this area.

When evaluating a website with a screen reader, the aim is to establish if it is possible to navigate and read the site by using aural clues and keyboard movements alone. This is because a person with blindness cannot use many of the tools or clues available to sighted persons, e.g. a mouse. Some of the items to be checked for include:

- Is information presented in a meaningful order when spoken, e.g. are headlines presented before body text?
- Is it possible to navigate and input details to a web form without recourse to a mouse?

- Are plain text descriptions provided for all images that are central to the understanding of content?
- Do suitable titles appear for hyperlinks that change the onscreen environment, e.g. that open a new window or application?
- Does the website still work when scripting is disabled in the browser? (This is because some screen readers cannot interpret Client-Side Scripting.)

Once any issues have been rectified, a final review of the site against the guidelines of the WCAG 1.0 can be completed. If all the requirements have been met, the site can be awarded compliance status and the appropriate logo displayed.

Of course, new content must also comply with this standard. As such, it is recommended that a complete website accessibility assessment be carried out at least every six months.

Security Test

A key threat to the ongoing development of the World Wide Web is concern about online security. An endless series of viruses and data infiltrations have caused significant disruption to the internet, as well as increased costs for development and hosting. This is because of the extra security equipment that is now necessary. Yet, the most serious consequence of all this activity is that it is undermining public confidence in the web.

As we saw earlier, trust is a key factor for determining the success or failure of an online venture. The same holds true at a global level—if the public do not trust the internet as a secure means of communication, they will not use it. This is particularly problematic for Transactional sites that depend on credit card payments. For example, one-third of consumers

say they would increase their online spending if they felt more secure about privacy[6]. As such, safeguarding the notion of trust and maintaining good security need to be top priorities for any Development Team.

Although web security encompasses a wide variety of disciplines, the fundamentals that underlie it can be expressed in just three concepts. These are:

- Confidentiality
- Integrity
- Availability

A website that fails to uphold each of these not only threatens its own business, but also exposes customers to risk.

Confidentiality

Confidentiality is the idea that information should only be available to those who are authorised to use it. For example, visitors may be given permission to download information from a website, but not upload it. The most common means for controlling such access is a 'Firewall'.

A **Firewall** is a software program that regulates traffic between 'zones of trust' on a computer network. For example, the internet is considered a zone of 'zero trust' because of the many viruses and other security problems that originate from it. In contrast, the computer upon which a website resides is 'high trust' because it can be tightly managed by a Technical Support Team. A key aim of website security is to allow connections between both these zones while also minimising risk.

---

[6] Forrester Research "Online Privacy Concerns: More than Hype". March 2004.

To achieve this, a Firewall can be configured to limit the type of traffic that is acceptable, e.g. uploading or downloading.

Yet, there may be instances where it is desirable for access to be extended. For example, a bank may wish to grant customers the ability to manage their accounts online. The challenge in this instance is to open the Firewall, whilst also limiting entry to approved persons only. In most cases this can be facilitated by some form of 'access control'.

Access Control

Access control means restricting the right of entry to a network to a limited audience. For example, a website that contains valuable research may only allow people who have paid a subscription fee to see their information. The most common means of doing so is via a Username/Password combination.

A Username/Password works by requiring two matching pieces of information to be entered into a site. These can then be compared against a database record. If they agree, access is granted.

Figure 5. Access control on www.gartner.com

However, this simple combination may not always be enough. Criminals know that most people use terms like their children's names as

passwords, and that more complex words are often written down as memory aids. Such carelessness is a leading cause of identity theft.

Identity Theft

Identity theft occurs where a criminal obtains data about an individual and attempts to pass themselves off as that person for fraudulent purposes. In 2005, 55 million Americans were exposed to identity theft[7].

While the eradication of identity theft is probably impossible, some simple rules can minimise its impact. For example, website users should not enter personal details into a site about which they have any doubts. Similarly, they can be advised to avoid passwords based on personal or family history, and not to share them with anyone. Other helpful guidelines include:

- Select passwords of at least eight characters.
- Include a mix of alphabetic, numeric, special (e.g. asterisk or hyphen) and uppercase characters.
- Select a word from a foreign language.
- Deliberately mis-spell the word.

A good password could be the German word "**zeitgeist**" (spirit of the age), rendered as "**seit-gei5T**".

In some circumstances, further levels of authentication may be needed to protect customers' data. For example, a website may request a secret PIN number or pose a 'Challenge Question' (e.g. your mother's maiden name) before granting access.

---

[7] USA Today "2005 worst year for breaches of security" http://www.usatoday.com/tech/news/computersecurity/2005-12-28-computer-security_x.htm December 2005. Accessed January 2006.

Physical Authentication

In extreme cases, it may even be necessary to limit access to visitors who are equipped with a physical authentication device. These are now being used to facilitate access to corporate extranet applications and banking systems.

An authentication device (such as those manufactured by RSA Security and Vasco [8]) is a piece of equipment that generates random PIN numbers. To access a secured site, a site visitor must use the currently displayed PIN together with their own username and password. Because the PIN is synchronised with the source website, it can easily be established if the number entered is valid or not.

Figure 6. Examples of RSA SecurID Token Cards.

Hackers and Crackers

Evaluating the resilience of access controls is a key procedure for testing site confidentiality. This is because many organisations will at some stage gain the attention of a Hacker or Cracker.

A **Hacker** is someone who wishes to break into a secure system, although they generally do not wish to undertake any type of illegal activity. In fact, Hackers may often be benign and simply seek to highlight inadequate security to website owners. That said, the phenomenon of Hacktivism can result in a site being penetrated in order to deface or vandalise it—perhaps for a political purpose.

---

[8] RSA Security, www.rsasecurity.com. Vasco, www.vasco.com

A **Cracker**, on the other hand, has malicious intent and may attempt to steal or corrupt data.

The Open Web Application Security Project is an organisation "dedicated to finding and fighting the causes of insecure software". In pursuit of this they maintain a list of the 'Top Ten Most Critical Web Application Security Vulnerabilities[9]' commonly exploited by Crackers. This list is compiled by a variety of security experts and represents a consensus on the most critical issues facing Developers. As of June 2006, these included:

- Unvalidated Inputs
- Broken Access Controls
- Broken Authentication and Session Management
- Cross Site Scripting (XSS) Flaws
- Buffer Overflows
- Injection Flaws
- Improper Error Handling
- Insecure Storage
- Denial of Service
- Insecure Configuration Management

Testing these vulnerabilities must form part of any security assessment. Additional tests for website confidentiality include ensuring authentication software is correctly configured and that all known loopholes are closed. Some firms go so far as to hire professional Hackers to conduct 'Penetration Tests' on their sites. These reports can be used as a means of tightening up access.

---

[9] The OWASP Foundation, "The Top Ten Most Critical Web Application Vulnerabilities." Copyright Open Web Application Security Project (OWASP) http://www.owasp.org/documentation/topten.html Accessed March 2006.

Specialist security review software is another useful tool for assessing possible vulnerabilities[10]. Packages like AppScan™ from Watchfire and Web Vulnerability Scanner from Acunetix can test for many of the issues on the OWASP list (where product cost is dependent on website scale).

Finally, the website of the Computer Emergency Response Team (CERT®) is an excellent resource for monitoring general internet security developments. CERT (www.cert.org) is a US government-funded institute that publishes advisories and incident reports about online threats.

Integrity

"Integrity" is a concept that seeks to prevent data being interfered with in an unexpected way, especially when being transferred over a network.

On a closed network, e.g. an email system within a university, the risk of unauthorised interference is low because all users are known to the Technical Support department. However, no such certainty is available on the internet. This means transferring details like credit card numbers over the web is inherently more risky. In such circumstances, the best way to manage the integrity of data is by way of encryption.

Encryption

Encryption is a system that uses mathematical algorithms to modify data so that it is unintelligible to anyone without a decryption key. Secure Sockets Layers (SSL) is the currently accepted standard for encrypting web transactions and can be used to protect data to a strength of 128 bits.

A bit is a unit of information, i.e. 0 or 1. As such, a 2-bit encryption key has four possible values: 00, 01, 10, and 11. As the number of bits increases, the amount of possible permutations grows exponentially. This

[10] For more visit, http://www.watchfire.com/securityzone/product/appscansix.aspx

means a 128 bit key has over **300 trillion trillion** combinations. If a Cracker attempted to decipher such a code, it could take years of work on the world's most powerful computers to find the right answer. Therefore, to all intents and purposes, SSL transactions are fully secure.

All the latest desktop browsers come preconfigured with 128 bit SSL capability. Such browsers also display various visual clues to help internet users find out if a page they are visiting is secure. For example, Microsoft Internet Explorer displays a 'padlock' symbol. The address of the web page also changes from 'http' to 'https'—the 's' indicates that the page is secure.

Figure 7. Close-up of Internet Explorer browser version 7 (Beta) showing address bar and 'padlock' symbol indicating a SSL secured webpage (from www.lulu.com).

Creating an SSL encrypted website is quite straightforward. All that is required is a certificate of identity and an encryption key from an approved vendor, for example Verisign or Thawte. SSL can then be enabled by the Technical Support Team.

Quantum Cryptography

While 128-bit encryption is adequate for current needs, the demand for even safer means of communication is growing all the time. Developments in cryptography are advancing to the point that, in the near future, it may be impossible to decode certain exchanges. 'Quantum

Cryptography' as it is termed, involves encoding information onto particles of light. The laws of physics ensure that if such data is interfered with, any attempted intrusion can be detected. Quantum cryptography may start to be introduced for video transmissions by 2007.

The most common test procedure for site Integrity, involves checking that all SSL certificates are up-to-date and that the host computer can handle secure transactions in the correct manner. This is especially critical for Transactional sites that rely on credit-card submissions for revenue.

Availability

The concept of availability requires information to be available to those who want it, when they want it, without interruption. For most sites this translates as a need for content to be up and running 100% of the time. Such high availability is particularly important for sites that engage in eCommerce. This is because online retailers are unable to collect revenue whenever a site is down. In this regard, one of the biggest threats to Transactional sites is the so-called Denial of Service (DoS) attack.

Denial of Service Attack

A DoS attack occurs when a site is bombarded with traffic from a malicious source. In such an event, the infrastructure of the website is unable to cope with such high levels of activity and effectively shuts down.

Unfortunately, these attacks are increasingly common simply because they are so easy to carry out. They have even been used as a means of extorting money[11]. For example, several online gambling firms have been threatened with a DoS attack if they did not pay blackmail to a criminal

[11] For example, BBC News Online "Online Service Foils Ransom Plot" http://news.bbc.co.uk/2/hi/technology/4579623.stm Accessed May 2005.

gang. For many sites, it proved cheaper to pay the money than to suffer the loss of revenue that would result from a security incident.

While such attacks are difficult to prevent, some basic screening can be carried out to block traffic from suspicious sources. However, even such fundamental measures can be ineffective because it is so difficult to distinguish between legitimate and criminal activity on the internet. The only fallback for most sites is to continuously monitor traffic and disconnect suspicious visitors before (or as) an attack occurs. To assist this, software known as an Intrusion Detection System (IDS) can be useful for screening visits and highlighting unusual goings-on, e.g. unexpected surges. Some well known IDS programs include Cisco Systems® Secure IDS and Top Layer IPS 5500 (licence cost is dependent on the scale of a website).

Of course, it should be realised that not every surge in traffic constitutes a DoS attack. It may simply be that many legitimate customers are visiting the website at once—perhaps as a result of a promotional campaign. As such, before taking any action, unusual instances should be checked to ensure that they really are a threat.

In general terms, the procedure for testing the availability of a site is to implement software for continuous monitoring. Options in this regard are explained below in the test process 'Performance Testing'.

Other Security Issues

Alongside such software issues, the security of website hardware should not be overlooked. The basic assessment procedure here is to ensure hardware is stored in a secure location—perhaps in a locked room to which only approved individuals have access. Many hosting companies provide such services, with the promise that their facilities are protected against destruction by vandalism, fire, flooding, etc.

However, if the worst does happen and a physical infrastructure is compromised, it is sensible for a standby solution to be in place. Some website hosting companies also provide this as part of their service.

(An organisation that manages its own infrastructure needs to have a similar set-up. Options in this regard are explored in Chapter Five– Website Infrastructure.)

Functional Test

Functional Testing is a process for evaluating whether a website can operate as expected under normal and error-inducing conditions. That is, does a site do what it is supposed to, even if a visitor makes a mistake when interacting with it?

For example, if the objective of a website is to allow people to book a hotel room, can such transactions be completed successfully? If errors occur, are they recoverable or is all information lost?

Everything else being equal, the only way to establish whether a site is functionally sound is to test each step of every application and record the outcome. Inevitably, this can be incredibly time consuming for a website that includes many features.

Fortunately, experienced Developers can write scripts that automate such tests. Furthermore, some Website Quality Assurance tools include the ability to record and run assessments, e.g. Web QA from Watchfire. Additional services are available in more advanced 'Performance Management' software from companies like Keynote Systems, BMC Software and Mercury.

As well as testing if applications work correctly, it is also necessary to see how a site reacts when things go wrong. For example, if an online form to book a hotel room includes a field for credit card numbers, what happens if a user attempts to submit a wrongly formatted number? If the

application has been well designed, the form should display an error message advising of the problem and indicating how it may be resolved.

HTTP Errors

Another common test is to consider the effect of a visitor entering an address for a page that does not exist, or to which they do not have authorised access. In such circumstances, a system generated error message should appear. The most common examples of such pages include:

"HTTP 404. Page Not Found"

This page appears where a user enters an address for a standard HTML page that does not exist. For example, if I am looking for
www.website.com/products.HTML
and I mistype it as
www.website.com/prodcts.HTML
I will receive a HTTP 404 message.

"HTTP 500. Internal Server Error"

This error appears where a user enters an address for a Server-Side Scripted page that does not exist. For example, if I am looking for
www.onlineshop.com/dvd-pricelist.php
and I mistype it as
www.onlineshop.com/dvd-priceleest.php
I will receive a HTTP 500 message.

"HTTP 403. Unauthorised Access"

This page appears where a user attempts to enter an address to which they do not have authorised access.

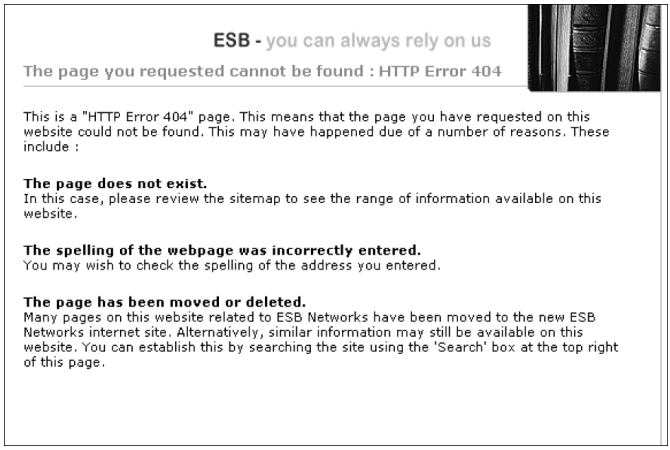

Figure 8. A customised HTTP 404 error message (from www.esb.ie).

Because these pages are automatically generated, the error messages they contain are usually not very helpful. As such, it is recommended that a series of customised messages be shown in their place. While the normal HTTP 404 page may merely state what has happened in a very general way, the customised page could include text that directs the visitor to a Search Engine or sitemap. It could also be wrapped within the site's design template.

Performance Test

The purpose of a Performance Test is to gauge the responsiveness of a website under normal and exceptional operating conditions. For example, on an average day a small website may receive a few hundred visitors. But what happens if a promotional campaign attracts thousands more? Will the Website Infrastructure be able to cope with the increase in load, or will it crash?

Performance Testing aims to establish the 'happy values' within which a site can operate. Many aspects of such testing can be accomplished by executing scripts that mimic real life scenarios. These include:

Load Testing

This is a test that mimics standard activity on a website and identifies the limits of acceptable performance. That is, based on an average number of visitors, do response times remain within acceptable limits?

Load testing can also consider contingencies in the event that activity on a website increases dramatically. For example, can additional processing power be made available if traffic increases over time? If not, users may experience a poor response which could damage the business. Typical recommendations for spare capacity range from 25% to 50% of average loads (bearing in mind that traffic peaks are often three times greater than average figures).

Stress Testing

As the name implies, Stress Testing pushes a website to the edge to establish how well it reacts in extreme circumstances. A test of this nature could be used to determine the maximum number of visitors a site can handle at any one time. It can also allow a Technical Support Team to plan how it would respond in circumstances where heavy traffic is received (perhaps by prioritising some traffic over others).

Endurance Testing

This test evaluates what happens in the event that heavy loads are sustained for long periods. Can the computer that hosts the site continue to deliver content effectively, as well as manage its own internal systems, e.g. memory caching? If not, what mechanisms are in place for reducing activity in a measured way, e.g. by deliberately cutting-off some visitors?

Spike Testing

Finally, Spike Testing can be used to establish what happens in the event of a sudden dramatic increase in activity that lasts only a few seconds.

Will the website crash or can it be configured to process requests in an orderly manner?

The tasks of Performance Testing are carried out by a Technical Support Team, sometimes with the support of software from companies like Keynote Systems, BMC Software and Mercury Interactive (licence cost is dependent on website scale).

Website Standard Review

This review seeks to evaluate a site against the guidelines in a Website Standard. A Website Standard is a document that details an organisation's approach to every aspect of site management and construction (see page **Error! Bookmark not defined.** for more).

The evaluation process involves comparing each item in the Standard against the site (in the manner of a checklist) and ticking them off as necessary. The main items to focus on are those that encompass development practices specific to the organisation itself.

For example, a Development Team may have its own preferences for the naming of files, linking to external websites and the use of pop-up windows. Because these rules can change from organisation to organisation, they need to be tested for prevailing circumstances.

Signing Off Successful Website Testing

Once all testing has been completed, the Team Leader is in a position to decide if the site can go live. Such a decision is essentially a judgement about whether she believes the site conforms to a minimum acceptable standard.

If it does, then there is no need to delay—the site can go live immediately.

However, most Website Testing will uncover at least a few problems that require attention. Of these, there may be a small number that cannot be resolved in time for an agreed date. The challenge for the Team Leader is to determine whether to launch the site as it is (complete with errors) or insist on a delay to allow remedial action.

Needless to say, any decision of this type is inherently thorny. For example, going live too early could result in bad press if some key applications do not work properly. Yet, delaying a launch could antagonise stakeholders who want the site to be made public.

Show Stoppers versus Nice to Change

To allow a Team Leader to arrive at a sensible conclusion, it is useful to categorise problems into one of four groups. These are:

- **Show stopper**. This indicates a problem that could seriously impede the integrity of the site, e.g. a security review finds that the Firewall is intermittently failing, leaving the site open to attack.
- **Highly disruptive**. An error of this kind implies that a core design or development requirement has not been satisfied. For example, a key application may have failed a Functionality Test.
- **Somewhat disruptive**. This category encompasses problems that are not considered overly serious. For example, a Discussion Forum that does not accept postings containing HTML.
- **Nice to Change**. Issues in this category typically include incidental suggestions that may improve the general performance of a site, e.g. marginal results from a series of usability tests.

By labelling problems according to the disruption they cause, these categories provide an objective means for assessing the impact on site quality. They can then be used by the Team Leader to arrive at a sensible

decision. For example, problems classified as 'Show Stoppers' or as 'Highly Disruptive' usually mean the site cannot golive until a fix has been put in place. In such circumstances, a renegotiation of a launch date will be necessary.

In contrast, problems classified as 'Somewhat Disruptive' or as 'Nice to Change', indicate it is probably OK to release the site—even though some minor issues may arise.

Ongoing Assessment

Once the site is live, the Development Team Leader (in conjunction with the Maintenance Team) should set out a programme for continuous Operational Testing. A programme like this indicates when various activities of the Website Testing Catalogue need to be repeated. For example, the calendar might specify that 'Link Testing' should be reviewed each week (as seen in Chapter Two–Website Maintenance), whereas Accessibility Testing may only be needed twice a year. The benefit of such a schedule is that it ensures a site can continue to conform to the high standard achieved when first released.