The Archaeologist Undeceived: Selecting Quality Archaeological Information from the Internet

Paul Sturges and Anne Griffin
Loughborough University, Loughborough, UK

r.p.sturges@lboro.ac.uk

Abstract

The amount of unreliable information and actual misinformation available via the Internet makes its use problematic for academic purposes, particularly in a data-intensive discipline such as archaeology. Whilst there are many sources for reviews of websites, few apply the type of criteria most appropriate to archaeology. Information and library professionals have developed sets of criteria that can be adapted for the evaluation of archaeological websites. An evaluative tool for archaeological websites, using already-available criteria, was developed and tested on twenty archaeological web sites. It proved capable of allowing its user to make clear distinctions between sites on the basis of quality. Further refining of the evaluative tool is possible on the basis of testing by both archaeologists and information professionals.

Keywords: archaeology, evaluation, Internet, quality, web sites

Introduction

Finding and identifying information of high quality from the Internet is arguably one of the most significant current problems across scholarship, teaching and information work. The Internet has been likened to 'a huge vandalized library' (Gorman, 1995) and a fire hose (Rettig, 1995) gushing with information. Champion (1997) likens the ephemeral 'here today and gone tomorrow' sites, as 'crop marks in the fields of the Internet'. This is clearly unsatisfactory in a data-intensive discipline such as archaeology. Scholars bring a lifetime of immersion in their discipline to any resource, analogue or electronic, but the expenditure of time and mental energy in making the necessary distinctions can be considerable. The student, whether taking a formal programme of study or approaching the subject as an amateur enthusiast, is at a considerable risk from the imperfections and tendency towards misinformation that comes with the Internet. For information professionals (librarians, museum curators or information managers) to facilitate the work of the scholar or direct the student towards information that is reliable and helpful is far from straightforward.

This paper makes a tentative exploration of the potential of evaluation tools to assist both information professionals and end-users of archaeological information from the Internet. The authors are information professionals, both of whom have a long-standing interest in the problems presented by archaeological information and, in the case of one (Griffin), extensive experience of practical archaeology. The paper takes an approach derived from the established practices and traditions of information and library science, on the grounds that these meet the needs of archaeology as a discipline in an appropriate and effective way. The starting point in the argument for this approach is that there are reasons for treating archaeology as a distinct problem-area.

First of all, it is clear that there are not only substantial resources of archaeological information on the Internet, but they have a distinctive character, closely related to the modes of research and communication peculiar to the discipline. Resources derive from a range of providers including university academics, museums, government departments, special interest groups and amateurs. They fall into many different categories, including directories, learning and teaching sites, abstracts and indexes, personal, commercial and university sites. The information may be presented in database form, as a free or subscription journal, a catalogue or as a generic web site. It may contain previously unpublished data and grey literature, such as theses and the content of card indexes, material previously published in print format or prepared specially for Internet distribution, or a synthesis of some of these. Particularly important in a discipline such as archaeology is the presence of 3D reconstructions and modelling, and multimedia applications, such as Quick Time VR, integrating audio, video and animation.

Secondly, the subject is one of those, like health, politics, business and law, which are particularly susceptible to misinformation. The popular appeal of the subject material, coupled with the complexity of the issues, allows those with an agenda other than the discovery of objective truth to spin seductive webs of fantasy and selective presentation of data. A striking example of this is the very selective use of archaeological data to support the argument that astronauts from another world introduced the arts, myths and social organisation of civilisations such as that of ancient Peru. The astronaut theories of Erich von Daniken have been around for over 30 years and are systematically demolished in Websites such as the *Skeptics Dictionary* (Carroll, 2002), but there are other sites that promote the idea and similar attractive eccentricities. More currently, we could take the discovery in 1996 of human remains, subsequently dated at 7300BC, at Kennewick, Washington State, USA. (Piper, 2002) The remains were the subject of a legal struggle between Native Americans seeking to give them a traditional burial and anthropologists who wanted to preserve them for scientific purposes. This is a legitimate dispute, but anyone using search engines for Internet information on the topic would be very likely to discover the *Kennewick Man News* site (New Nation, 2002), which denies the Native American origins of the skeleton in favour of an argument that Europeans were the first true occupants of North America. The site also provides links to various White supremacist resources. At the very least, caution is required when using such material. The question as to how one evaluates archaeological information from the Internet so as to identify that which is as free as possible from both deliberate and inadvertent distortion or deception is thus important at more than one level.

Developing Evaluation Tools

The idea that evaluation of resources is a worthwhile exercise is hardly original. It has a long history in the practice of librarianship, and writers such as Katz (1997) have developed and polished sets of criteria for this purpose. Criteria for evaluating Internet resources exist in some profusion and it is clear that, as argued by Smith (1997), the extension of this tradition to the Internet is a natural role for information professionals. There is, however, virtually nothing directly intended for the evaluation of archaeological resources, except for some work on evaluation of classical studies resources by Merrill (2000). As pointed out by various authors, notably Cooke (1999), the specific needs of communities of users need to be addressed, and only through an understanding of needs can there be effective resource evaluation. The distinctive character of archaeology, from an information professional's point of view, is that it uses scientific technique to achieve humanities outcomes. This means that its practitioners not only need textual materials, but access to a long-established humanities literature, great volumes of data, results of scientific analyses, and sophisticated visual applications. These resources are needed not only for scholars, but there is also an enormous demand for popularly accessible and educational materials in the subject area. The specific objective of the project that this paper describes was to develop an archaeology-specific set of evaluation criteria and carry out some preliminary testing of their effectiveness in practice.

Before proceeding to describe this, the inappropriateness of existing reviewing services needs to be established. Ciolek (1996) calls web site evaluation one of the quests for the electronic grail, and he divides approaches to evaluation into individual work producing checklists, and the larger projects reviewing online material, such as *Infomine* (Mitchell and Mooney, 1996). When looking at the numerous web site guides on the World Wide Web, Collins (1996) points out that 'almost no one seems conscious of the standards carefully developed by the information professional over the last century', and Rettig (1995) points out the prevalence of 'coolness' as a criterion. This amounts to judging a site more on appearance than substantive content and technical characteristics. Of course, some reviewing sites do emphasize content and authority. For example Smith (1997) suggests *The Argus Clearinghouse* (now merged with the *Internet Public Library*) and Rettig mentions *Infofilter* in similar terms. However, major reviewing sites offer little that meets the specific needs of archaeology.

The literature of evaluation does contain indications as to how a more appropriate tool might be developed. The key sources of guidance on choosing criteria for web site evaluation, such as Alexander and Tate (2001) have several main sets of criteria (five: authority, accuracy, objectivity, currency and coverage, in the Alexander and Tate example) each with a sub-set of questions. But Smith (1997) adds key elements to the usual criteria, the new ones being workability, graphic and multimedia design, browsability and organization. He makes the point strongly that different resource genres require different sets of evaluative criteria and suggests a 'toolbox' approach, choosing applicable criteria from a list. Rating performance in a particular category on a pre-determined scale (such as a 1-5 range per criterion that contributes towards an overall average, developed for the *Argus Clearinghouse*) can then be applied to clarify levels of quality. Cooke discusses the pros and cons of checklists and rating tools, pointing out that there is a danger that criteria may sometimes be so restrictive that all web sites would perform badly. She shows that there is also evidence that some of the resources that score highest on a checklist actually have incomplete or inaccurate information. It is precisely this that leads her to the conclusion that 'evaluative tools should be tailored to [specific] subject and user combinations as the greater the coverage, the greater the chances of subjectivity creeping into the equation' (Cooke, 1999 p. 162).

Of course, if automatic classification and evaluation of the publicly accessible World Wide Web sites were possible, the subjective element would be removed leaving a purely quantitative assessment. There is work in progress on this approach, as in the following examples. A software programme to analyse hypertext and content, by statistical clustering, textual analysis, and neural networks has been developed, but a human input is still required (Bauer and Scharl, 2000). An automatic Soft System programme has been used to analyse academics' personal web pages and it was found that that academic rank was not related to the information style, suggesting that information professionals may still be needed to differentiate between recognized scholars and novices (Brown-Syed, 2001). It has also been suggested that second generation, client-side based programs using quantitative indicators, will be able evaluate resources using hypertext metadata (Aguillo, 2000). None of this, or the other quantitative approaches, convinces that a well-devised evaluative tool, based on an intimate knowledge of a subject area is likely to be replaced in the immediate future as the best means of assessing relevant web resources. It is on this basis that the tool described in what follows has been developed.

An Evaluation Tool for Archaeology

Using the principles suggested above, a tool designed to meet the specific needs of archaeology was developed and subjected to some preliminary testing. The criteria were chosen mainly from the *Toolbox of Criteria* (Smith, 1997). They include questions that have been used in the past for print media, and others specifically produced for the WWW. The list naturally has similarities to those proposed by other information professionals, including Cooke (1999) and Tillman (2000). The real difference is that the choice was made specifically for archaeology and, therefore, included questions such as 'Are artefacts depicted with good quality illustrations?' The outline list of criteria for evaluation is as follows:

- scope,
- purpose and audience, reviews,
- content (including; accuracy, authority, copyright, currency, uniqueness, links, quality, and overall quality),
- graphic and multimedia design,
- workability (including; user friendliness, computer environment, searching, browsability and organization, interactivity, connectivity).

The majority of the evaluation questions follow a checklist format (score one for the desirable answer and zero for an undesirable answer). Sometimes, however, a text answer is more appropriate so as to allow richer annotation, which would be displayed with the ratings in any publicly available evaluations using the tool. The numerical rating scale can be expanded where necessary. Thus, the answer to 'Are there contact details, i.e. email and postal addresses for clarification, error correction and new information?' will score one each for the email and postal addresses. In other cases a rating is required, on a scale of one to five, for instance, of the overall quality of the content of the site. A maximum of 64 points was available. When scores had been allocated to a site, they could be totalled and the total used as an overall rating, reflecting four broad classes identified as excellent (58-64), good (42-57), fair (26-41) and poor (0-25). The evaluation tool that produces these ratings takes the following form.

Scope

This calls for assessment of the site's depth (scholarly level) and breadth (subject range), which should be suitable for the proposed audience. Types of material covered should be looked at, including published literature, databases, audio and video clips, and virtual reality. The format of material covered is also relevant, for example, telnet, Gopher and FTP protocols, because accessibility may be restricted by software and hardware considerations. It is usual to ask within this criterion whether coverage is retrospective, but archaeology is by definition a retrospective discipline, so this is taken as given. The questions are as follows.

Is the scope stated (not implied)? Yes = 1, no = 0
Does the scope meet expectations? Yes = 1, no = 0
Breath – how comprehensive is it, what is covered? Text answer
Depth – what audience level is served? Text answer
Are the following resource format types mentioned: audio, video, Telnet, Gopher, FTP? Text answer.

Purpose and Audience

Addressing a specified audience for a particular purpose assists user retrieval of appropriate sites. Statements of aims, objectives, purpose, audience and coverage should be found on home pages, or the 'about this site' or FAQ pages. The subject material should be of an appropriate level or depth. A site counter may indicate popularity, but any visitor comparisons should be made at one time point across various sites and not an accumulation of previous visits, this project does not have the technology to do this.

Is the purpose of the web site clearly defined? Yes = 1, no = 0.
Is the audience of the web site clearly defined? Yes = 1, no = 0.
Does the resource accomplish its purpose as described? Yes = 1, no = 0.
Does the resource suit the intended audience? Yes = 1, no = 0.
Is there a counter on the web site? Yes = 1, no = 0.

Content

Content questions are generally the most important group of criteria for any site and they include matters of accuracy, authority, copyright, currency, uniqueness, links, and quality. It is important to distinguish between content that is factual and that which is opinion. It should also be asked if factual information is stand-alone or is abstracted from elsewhere. Similarly, it should be distinguished whether the site contains original information itself or act as a directory to other sites.

Is the content factual (not opinion)? Yes = 1, no = 0.
Is there original information and/or links? Yes = 1, no = 0.
Does it contain some stand-alone content (not just abstracted from an original source)? Yes = 1, no = 0.

Content – accuracy.

This may have to be inferred rather than measured directly, and may actually be a perception of accuracy. The author's academic record, a clear statement and fulfilment of aims, justification of methods, findings, and conclusions, plus references to appropriate published sources, are reasonable indicators. Other indicators of accuracy include the identification of editors and referees. It needs to be noted from the scope statements or 'about us' pages, if special interest groups whose agenda may suggest some form of bias are associated with the site. Indicators suggesting questionable accuracy may include undated information, obsolete data in fast moving topics, over simplification, exaggeration, emotional and intemperate language, and a stance that does not take opposing views into account.

Is the information likely to be accurate? Yes = 1, no = 0.
Is the information subject to biases, e.g. political, financial, ideological, etc.?
Yes = 0, no = 1.
Does it contain advertising? Yes = 0, no = 1.

Content - authority

Authority is concerned with the reputation, knowledge and expertise of the individual or organizational authors, the identity of these should be easily established. The resource may have its own reputation, and the reputation of other affiliated organizations (sponsors, funding agencies, etc.), may be taken into account. The URL address descriptors may aid identification of the organization type, for example, government department (.gov), academic (.ac, .edu), organization – often a non-profit organization such as a charity or pressure group (.org), company (.co, .com) or an individual's web site within any of the above prefaced by ~. The sources of information should be stated, for example, a bibliography (or webliography) of some description should be supplied to allow verification of the original information (and its suppliers). Full contact details, including a street address, should be available for correction of errors, further questions and the reporting of additional information. It is also pertinent to ask if a professional association or society with expertise in the subject area has reviewed the site, or does it carry some award, badge or kite-mark.

Is the originator a reputable expert or organization, with standing in the field? Yes = 1, no = 1.
Are basic details available on the institution? Yes = 1, no = 0.
Are the sources of information stated? Yes = 1, no = 0.
Can the information be verified? Yes = 1, no = 0.
Are there contact details for clarification, error correction and new information?
    Email address only (score 1).
    Are There Both Email and Postal Addresses (Score 2)?
Has the web site been assessed by information or subject experts, e.g. for listing in a subject gateway or portal? Yes = 1, no = 0.

Has the web site been granted awards, badges or kite marks?
Yes = 1, no = 0.

Content - copyright

Resources are expected to contain the appropriate copyright declarations as a matter of good practice, reassurance of content authority, and identification of the rights holders for potential republication.

Is copyright ownership information available on the site? Yes = 1, no = 0.

Content - currency

Web sites by definition should be more current than printed materials. It is important to know when the information was originally written and more importantly if, and when it was been updated, and if there is a commitment to its revision. The time when the information was first written may be different from when it was first placed in the site. Inclusion of both of these dates, plus the last revision date and some indication that the site is reviewed regularly would be best practice. However, it needs to be remembered that in archaeology the usual criterion of timeliness only partly applies. Material centuries old, such as excavation reports, will still be valuable and should be expected to appear alongside news from current excavations.

Are the resources revised, i.e. not static? Yes = 1, no = 0.
Is revision frequent? Yes = 1, no = 0.
Are the revision dates stated? Yes = 1, no = 0.
Are there revision dates on individual pages, i.e. not just the home page?
Yes = 1, no = 1.
Is there a commitment to regular maintenance and stability?
Yes = 0, no = 1.

Content - uniqueness

The information provided by the site might be unique to that site, that is, not available in printed or other formats. However, if the content is not unique it can be compared to the versions in other formats.

Is the information content available in other forms? Yes = 0, no = 1.
Does the resource complement other resources, e.g. give raw data, update printed materials? Yes = 1, no = 0.
Does the resource have additional features, e.g. audio, video? Yes = 1, no = 0.

Content - links

Useful and effective hypertext links are vital. This includes both inward links, within the same page and to other pages within a site, and outwards links to other sites, or directly to an original reference.

Are there links to other resources? Yes = 1, no = 0.
Do the links have descriptive information? Yes = 1, no = 0.
Are the links appropriate? Yes = 1, no = 0.
Are the links current? Yes = 1, no = 0.

Content – quality

Assessing quality is frankly subjective, and a textual response is required alongside assessment of more practical matters such as grammar, spelling, and typography. The textual response is based on comparison of the site with other similar sites.

Is the web site well written, i.e. does it communicate information clearly? Yes =1, no = 0.

Is the web site well edited, e.g. lacks typographical and spelling errors?
Yes = 1, no = 0.
How does the web site compare with other similar sites? A textual answer, i.e. worse, as good as, better, is required.

Content – overall quality

This asks for a purely subjective impression of the web site, based on the informational content, usefulness to archaeologists as a category of users and usability. It should also reflect the other quality scores given.

Rate the overall quality of the web site's content on a scale of 1-5.

Graphic & Media Design

This section deals with resource organization and presentation. The web site should be logically arranged, with separate pages for separate subject themes, of the appropriate length and presented so the user is not overwhelmed. Pages should have a sensible font size and style, with enough 'white space' to make them easy to read and assimilate. The web site's navigation system should be clear and easy to follow, for example, an index in a separate frame and or a site map, and a search engine is necessary for larger web sites. The navigation system, should have readily identifiable links, backward and forward links to eliminate excessive scrolling, and should not take more than three clicks to reach any page. Images and visual effects, such as moving pictures, should be appropriate, relevant to the subject and add value. Additional multimedia components, such as, audio or video clips or 3-D reconstructions are particularly relevant to archaeology.

Is the web site attractive to look at? Yes = 1, no = 0.
Is the web site well organised? Yes = 1, no = 0.
Do the visual effects improve the web site? Yes = 1, no = 0.
Are artefacts depicted with good quality illustrations? Yes = 1, no = 0.
If present, are special effects appropriate to the resource? Yes = 1, no = 0.
Are navigational aids present? Yes = 1, no = 0.
Are navigational aids effective? Yes = 1, no = 0.

Workability

Workability is a term used by Koopman and Hay (1994), to mean 'ease of use and ease of connection'. It includes a group of criteria including, user friendliness, computer environment, intra-site searching, browsability and organisation, interactivity and connectivity. Factors affecting the consumer's ability to utilise the web site include; design, language, charges, password/registration requirements, help facilities, whether training is required, and if the site is still under construction (as this may lead to disappointment if the advertised pages are unavailable).

Workability - user friendliness

The following questions examine user friendliness.

Is the web site easy to use? Yes = 1, no = 0.
Is the web site design clear? Yes = 1, no = 0.
Is there an easy to use help facility available? Yes = 1, no = 0.

Workability - computer environment

Additional hardware should not be needed: if a web site says 'best viewed with' it is likely that non-standard software is required, so the web site may be in fact be best avoided. If additional software is

necessary to access the information, the requir ements should be explained, and the authors should freely provide downloadable software as plug-ins. Ideally the web site should be accessible on more than just the most popular browsers, such as *Microsoft Internet Explorer* and *Netscape Navigator*.

> Can the resource be accessed *via* standard equipment and software?
> Yes = 1, no = 0.

Workability - searching

An incorporated search engine or tool is a desirable facility in web sites of any size and this is especially true of data-rich archaeological sites. It should be easy to use and incorporate the usual keyword, Boolean searching and truncation for advanced searching.

> Can information be retrieved easily from the resource? Yes = 1, no = 0.
> Does the web site include a helpful search engine or tool? Yes = 1, no = 0.
> Does the search engine include the whole web site? Yes = 1, no = 0.

Workability – metadata

The metadata, found in the HTML header, comprises descriptors, such as, title, subject keywords, author (with or without addresses), affiliation, content, aim, and fo rmat. Metadata is increasingly used to assist information retrieval and evaluation. Using the Dublin Core is not only recommended by the UK Joint Information Systems Committee (JISC, 2001), but is obviously appropriate for an archaeological web site.

> Is appropriate metadata present to assist web site retrieval? Yes = 1, no = 0.
> Is the metadata in the form of Dublin Core descriptors? Yes = 1, no = 0.

Workability - browsability and organisation

The web site should be logically arranged, for example, chronologically or geographically (both of which are particularly appropriate for archaeology), to assist retrieval of information.

> Is the web site organized methodically? Yes = 1, no = 0.
> Is the organization scheme appropriate, e.g. chronological or geographical? Yes = 1, no = 0.

Workability - interactivity

Interactive features include operating via a Common Gateway Interface (CG) so as to allow the acceptance and return of data between the site and the World Wide Web. It can also include quizzes and games to aid understanding, and database query forms. Any of these are likely to enhance the educational value of an archaeology website

> Are there interactive features? Yes = 1, no = 0.
> Do any interactive features present improve the site? Yes = 1, no = 0.

Workability - connectivity

The web site should be stable and dependable, that is, it should not disappear or keep changing its URL. Other desirable features might include: improving downloading speeds by provision of local mirror sites to reduce traffic, the provision of thumbnail images (allowing the user to view a smaller easier loaded version of an image before choosing to enlarge it) and capacity to switch off images to increase the capacity available for downloading.

> Is access reliable? Yes = 1, no = 0.
> Is there a mirror site? Yes = 1, no = 0.
> Do the pages take too long to load? Yes = 0, no = 1.

Testing the Evaluation Tool

Twenty web sites were chosen for evaluation on the basis of the following criteria, which were adopted so as to obtain a representative sample whilst keeping this exploratory exercise comparatively simple:

- English language content.
- Access without payment.
- Content essentially in text and images.
- Both narrow and broad topic ranges (ten of each).
- Obtained via both portal and non-portal sources (ten of each).

The portal chosen for resource selection was the Archaeology directory of the Humbul Humanities Hub (2002). The non-portal websites were found using the Google search engine. The web sites were rated according to the evaluative tool, so as to show whether it permitted the evaluator to make clear distinctions between individual web sites. It was then possible to test if any worthwhile distinction could be detected in the scores achieved by particular categories of site (in this case portal-selected, as against search engine-selected web sites). If this were the case, it might suggest that the evaluation tool was sufficiently sensitive to make distinctions between whole categories of resource (presuming of course that such distinctions existed).

The selected sites (with rating and category) were as follows:

Sites found via Humbul: Archaeology

Narrow themed

A   Anglo-Saxon Cemeteries..49 Good
   http://www.gla.ac.uk/Acad/Archaeology/resources/Anglo-Saxon/cemeteries/index.htm

B   Celtic Inscribed Stones...47 Good
   http://www.ucl.ac.uk/archaeology/cisp

C   Corpus of Writing Tablets From Roman Britain ....41 Fair
   http://www.csad.ox.ac.uk/RIB/RIBIV

D   Duke Papyrus Archive   ....51 Good
   http://scriptorium.lib.duke.edu/papyrus

E   Stone Pages ....39 Fair
   http://www.stonepages.com

Broad themed

F   Ancient Cyprus Web Project ....43 Good
   http://www.ancientcyprus.ac.uk/index.html

G   Digital Egypt for Universities...38 Fair
   http://www.petrie.ucl.ac.uk/digital_egypt/welcome.html

H   Maya Ruins   ....46 Good
   http://mayaruins.com

I   The Saxon Shore ....42 Good
   http://www.pitt.edu/~jegst61/

J   Southampton Archaeology Collection ....43 Good
   http://sccwww1.Southampton.gov.uk/archaeology

Sites found via Google

Narrow themed

K   The Bead Site   ....44 Good
      http://www.thebeadsite.com/welcome.html

L   Celtic Coin Index   ....46 Good
      http://www.writer2001.com/cciwriter2001/index.htm

M   Haida Totem Poles   .... 30 Fair
      http://www.wsu.edu/~1niep/fs12/front.html

N   Roman Inscriptions of Britain   .... 34 Fair
      http://www.bedoyere.freeserve.co.uk/RIB.htm

O   Society for Clay Pipe Research   ....49 Good
      http://www.scpr.fsnet.ci.uk/

Broad themed

P   Archaeology in Arctic North America   ....46 Good
      http://www.arts.uwaterloo.ca/anthro/rwpark/archsuff/arcticintro.html

Q   Archaeology in York   .... 41 Fair
      http://www.york.ac.uk/depts/arch/yccweb/welcome.htm

R   Irish Archaeology   ....42 Good
      http://www.xs4all.nl/~tbreen/ireland.html

S   Maya Archaeology   .... 31 Fair
      http://www.maya-archaeology.org

T   The Nautical Archaeology Society   ....46 Good
      http://www.nasportsmouth.org.uk

From this it is clear that the tool did allow numerical ratings that permit clear distinctions to be made between sites. A range of scores from 30 to 51 out of a possible total of 64 reveals sufficient differences as to help information professionals or users to form an opinion on a site. The four categories (poor to excellent) were much less helpful, as all the sites evaluated were rated good or fair. Possibly with a greater number of sites, the two ends of the scale would have had to be applied more frequently, but, whether this is the case or not, the evaluation tool showed a capacity to make at least a crude distinction between sites. The categories could, of course, be adjusted at will to better reflect the distinctions that emerge from the numerical ratings.

The comparison between the evaluations of portal-selected and search engine-selected sites proved in-conclusive. An unpaired, two-tailed t-test was performed on the scores awarded to web sites retrieved in the two different ways. This revealed no significant difference between the scores awarded to the two sets of sites (at 18 degrees of freedom, significance level of 0.05, t = 1.179). Although this is slightly disappointing, it was never very likely that such a small sample would produce a statistically significant result, and the result almost certainly says more about the group of sites selected than it does about the method of evaluation.

Conclusions

The evaluation tool is, of course, a means for channelling subjectivity. As such, it is the choice of criteria and the values that can be attached to them that affects whether it performs well or not for a user with

no, or few, preconceptions about sites. The evaluation tool developed here reveals a capacity to function reasonably effectively in the context of archaeology. In this it has much in common with other such tools in different subject areas. Indeed it performs better than many seem to have done. For instance, 629 sites containing clinical data (which must be accurate for the most compelling reasons) when tested against benchmarks provided by the Journal of the American Medical Association would have been considered inadequate (Hersh *et al*.,1998). It is unlikely that all of them were indeed as bad as this suggests. In this case, the evaluation tool was almost certainly inappropriately designed and/or calibrated. This leads us back to Cooke's dictum that evaluative tools should be tailored to very specific subject and user combinations (Cooke, 2001).

This calls for the identification of a way in which the evaluation tool described in this paper could be tailored so as to perform even more effectively. In the version used here, a good deal of credit is given for quality under general criteria that might apply quite as well to web sites on most other topics. This suggests that possibly shifting the emphasis in the scoring system aware from these general criteria to the questions with the most obvious significance for archaeology might produce a tool that made sharper distinctions between sites. Thus, a way might be found to shift the scoring weighting towards the graphic and multimedia design category and some of the workability aspects, without neglecting the significance of content. At that stage more rigorous testing of the evaluation tool would be appropriate. This could be done in various ways, including a comparison of the results of an information professional's use of the tool against a completely subjective assessment of the same sites by an archaeologist or, indeed, an archaeologist's use of the tool. Publicly available ratings of web sites using this tool would need to be subjected to regular revision (recognising that change over time is quite usual) and this process could also be used to adjust and recalibrate the tool. At present, however, the authors feel that sufficient progress has been made towards the creation of an archaeology-specific evaluation tool that the usefulness of the enterprise itself is confirmed.

References

Aguillo, I. (2000). A new generation of tools for search recovery and quality evaluation of World Wide Web medical resources. *Online Information Review*, 24, 138-143.

Alexander, J. & Tate, M.A. (2001) Evaluating web resources: checklist for an informational web page. Retrieved June 24, 2002 from http://www2.widener.edu/wolfgram-memorial-library/webevaluation/inform.htm

Bauer, C. & Scharl, A. (2000). Quantitative evaluation of Web site content and structure. *Library Computing*, 19, 134-146.

Brown-Syed, C. (2001). Determining authoritativeness on the Web: an exploration of the content and roles of academics' Web pages. *Library and Archival Security*, 17, 43-70.

Carroll, R.T. (2002). The Skeptics Dictionary. Retrieved November 9, 2002 from http://skepdic.com/vondanik.html

Champion, S. (1997). Archaeology on the World Wide Web: a user's field guide. *Antiquity*, 71, 274. Retrieved July 20, 2002 from http://intarch.ac.uk/antiquity/electronics/champion.html

Ciolek, T.M. (1996) The six quests for the electronic grail: current approaches to information quality in WWW resources. *Revue Informatique et Statistique dans les Sciences Humaines*, 91, 45-71.

Collins, B.R. Beyond cruising: reviewing. *Library Journal*, February, 122-124.

Cooke, A. (1999). *Neal-Schuman authoritative guide to evaluating information on the Internet*. NetGuide Series. New York: Neal-Schuman.

Cooke, A. (2001). *A guide to finding quality information on the Internet: selection and evaluation*. London: Library Association Publishing.

Gorman, M. (1995). The corruption of cataloguing. *Library Journal*, 120, 34.

Hersh, W.R. *et al*. (1998). Applicability and quality of information for answering clinical questions on the Web. *Journal of the American Medical Association*, 280, 1307-1308.

Humbul Humanities Hub (2002). *Archaeology*. Retrieved August 20, 2002 from
    http://www.humbul.ac.uk/output/subout.php?subj=archaeology

JISC (2001). *Metadata and the JISC Website*. Retrieved July 23, 2002 from http://jisc.ac.uk/admin/metadat.html

Katz, W. (1997) *Introduction to reference work*. 7th ed. New York: McGraw Hill.

Koopman, A. & Hay, S. (1994) Swim at your own risk – no librarian on duty: large-scale application of Mosaic in an academic library", in: *Electronic Proceedings of the Second World Wide Web Conference '94: Mosaic and the Web.* Chicago, IL: National Center for Supercomputing Applications, 1994. Retrieved August 10 2002 from
    http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/LibApps/hay/WWWpap.html

Merrill, J. (2000). The Internet and classical civilisation. *Acquisitions Librarian*, 23, 97-115.

Mitchell, S. & Mooney, M. (1996) Infomine: a model Web-based academic virtual library. *Information Technology and Libraries*, 15. Retrieved August 26, 2002 from http://infomine.ucr.edu/pubs/italmine.html

New Nation (2002). Kennewick Man News. Retrieved November 9, 2002 from http://www.newnation.org/NNN-kennewick-man.html

Piper, P.S. (2002). Web hoaxes, counterfeit sites, and other spurious information on the Internet. In Mintz, A.P. ed. *Web of deception: misinformation on the Internet*. Medford, NJ: CyberAge Books. 1-22.

Rettig, J. (1995). Putting the squeeze on the information firehose: the need for neteditors and netreviewers. Retrieved August 22, 2002 from http://www.swem.wm.edu/firehose.html

Smith, A.G. (1997) Testing the surf: criteria for evaluating Internet information resources. *Public Access Computer Systems Review*, 8. Retrieved October 5, 2002 from http://info.lib.uh.edu/pr/v8/n3/smit8n3.html

Tillman, H.N. (2000). Evaluating quality on the net. Retrieved June 24, 2002 from
    http://www.hopetillman.com/findqual.html

Biographies

Paul Sturges is Professor of Library Studies in the Department of Information Science at Loughborough University. He has written on a wide range of topics across the field of information and library science and lectured, delivered conference papers and acted as a consultant in many parts of the world. A special interest in Africa is reflected in *The quiet struggle: information and libraries for the people of Africa* (with Richard Neill) 2nd ed. Mansell, 1998. His most recent work (including consultancy for the Council of Europe) has concentrated on questions of access to information via the Internet, with a particular emphasis on personal privacy, and he is author of *Public Internet access in libraries and information services*, Facet Publishing, 2002. A new edition of the *International encyclopedia of information and library science*, which he has edited with John Feather for Routledge (first ed., 1997) will appear in 2003.

Anne Griffin is a Science graduate of the Open University and took her Masters in Information and Library Studies at Loughborough University 2002. Since 1998 her longstanding practical involvement with archaeological work has been complemented by the archaeology modules she has taken at the University of Surrey.