The Six Quests for the Electronic Grail: Current Approaches to Information Quality in WWW Resources

T. Matthew CIOLEK

Résumé. Cet article analyse les approches logicielles, procédurales, structurelles, bibliographiques, évaluationnelles et organisationnelles de la qualité de l'information en ligne. Un progrès rapide dans tous ces domaines est indispensable pour faire de l'Internet un medium fiable pour la publication de textes scientifiques.

Keywords: Electronic publishing of scholarly works, WWW, information quality.

Mots-clés : Édition électronique de travaux scientifiques, WWW, qualité de l'information.

1. The untrustworthiness of the WWW

The untrustworthiness and mediocrity of information resources on the World Wide Web (WWW), that most famous and most promising offspring of the Internet, is now a well recognised problem (Ciolek: 1995a, Treloar: 1995, Clarke: 1996). It is an inevitable conclusion to anyone who has worked, even briefly, with the 'Information' areas of the Web, the other two being 'Exchange' and 'Entertainment' (Siegel: 1995).

The problems with the Web are many. WWW documents continue to be largely un-attributed, undated, and un-annotated. As a rule, information about the author and publisher is either unavailable or incomplete.

✉ Coombs Computing Unit, Research School of Social Sciences; Australian National University; Canberra, ACT 0200 (Australia).
Fax: +61 (0)62 57 18 93 E-mail: tmciolek@coombs.anu.edu.au

Frequently, the rationale for placing a document on-line and information about how it relates to other materials is not explicitly stated. It has also been observed that the Web remains a place in which far too many resource catalogues seem to chase far too few original or non-trivial documents and data sets (Ciolek: 1995a).

Simultaneously, there are no commonly accepted standards for the presentation of on-line information. Instead, there is an ever-growing proliferation of publication styles, page sizes, layouts and document structures. Moreover, links to other Web resources tend to be established promiscuously, that is without much thought for the target's relevance or quality. There is also a pronounced circularity of links. This means that many Web pages carry very little information, apart from scantily annotated pointers to some other equally vacuous index pages that serve no other function apart from pointing to yet another set of inconclusive indices and catalogues (Ciolek: 1995a). Finally, emphasis continues to be placed on listing as many hypertext links as possible—as if the reputation and usefulness of a given online resource depends solely on the number of Web resources it quotes. In practice this means that very few such links can be checked and validated on a regular basis. This leads, in turn, to the frequent occurrence of broken (stale) links.

The whole situation is further complicated by the manner in which Web activities are managed by people whose daily activities fuel the growth of the Web. The truth is, there is very little, if any, systematic organization and coordination of work on the WWW. The project started five years ago by a handful of CERN programmers has now been changed beyond recognition. With the advent of *Mosaic*, the first user-friendly client software (browser) in September 1993, WWW activities have literally exploded and spread all over the globe. Since then Web sites and Web documents continue to grow, proliferate and transform at a tremendous rate. This growth comes as a consequence of intensive yet un-reflective resonance and feedback between two powerful forces, innovation and adaptation.

The first of these is a great technological inventiveness (often coupled with an unparalleled business acumen) of many thousands of brilliant programmers (Erickson: 1996). The second consists of many millions of daily, small scale, highly localised actions and decisions. These are taken on an ad hoc basis by countless administrators and maintainers of Web sites and Web pages. These decisions are made in response to the steady flow of

new technical solutions, ideas and software products. They are also made in reaction to the activities, tactics and strategies of nearby WWW sites.

The Web, therefore, can be said to resemble a hall of mirrors, each reflecting a subset of the larger configuration. It is a spectacular place indeed, with some mirrors being more luminous, more innovative or more sensitive to the reflected lights and imagery than others. The result is a breathless and ever-changing 'information swamp' of visionary solutions, pigheaded stupidity and blunders, dedication and amateurishness, naivety as well as professionalism and chaos. In such a vast and disorganised context, work on simple and low-content tasks, such as hypertext catalogues of online resources, is regularly initiated and continued at several places at once. For example, there are at least four separate authoritative 'Home Pages' for Sri Lanka alone (for details see Ciolek: 1996c). At the same time, more complex and more worthwhile endeavours, such as the development of specialist document archives or databases, are frequently abandoned because of the lack of adequate manpower and funding (Miller: 1996).

There is a school of thought, represented in Australia most eloquently by T. Barry (1995), which suggests that what the Web needs is not so much an insistence that useful and authoritative online material be generated, but rather one that intelligent yet cost-effective 'information sieves and filters' be developed and implemented.

Here the basic assumption is that the Web behaves like a self-organising and self-correcting entity. It is indeed so, since as the online authors and publishers continue to learn from each other, the overall quality of their networked activities slowly but steadily improves. Such a spontaneous and unprompted process would suggest that with the passage of time, all major Web difficulties and shortcomings will eventually be resolved. The book or a learned journal as the medium for scholarly communication—the argument goes—required approximately 400 years to arrive at today's high standards of presentation and content. Therefore, it would be reasonable to assume that perhaps a fraction of that time, some 10–15 years, might be enough to see all the current content, structural and organizational problems of the Web diminish and disappear.

This might be so, were it not for the fact that the Web is not only a large-scale, complex and very messy phenomenon, but also a phenomenon which happens to grow at a very steep exponential rate.

2. The urgency of the Web repair tasks

In January 1994 there were approximately 900 WWW sites in the entire Internet. Some 20 months later (August 1995), there were 100,000 sites and another 10 months later (June 1996) there were an estimated 320,000 sites (Netree.com: 1996). It is as if the Web was unwittingly testifying to the veracity of both Moore's and Rutkowski's Laws. Moore's Law proposes that electronic technologies are changing dramatically at an average of every two years; while Rutkowski says that in the highly dynamic environment of the Internet, fundamental rates of change are measured in months (Rutkowsk: 1994a). As the consequence, problems which are soluble now, when the Web pages can still be counted in tens of millions of documents, will not be soluble in the near future since the Web will be simply too massive. There is no doubt that the World Wide Web is running out of time. The WWW is facing an ungraceful collapse, a melt-down into an amorphous, sluggish and confused medium. This is a transformation which would undoubtedly place Web's academic credibility and useability on a par with that of countless TV stations, CB radios and USENET newsgroups.

Thus, we seem to be confronted with a curious paradox. On the one hand the WWW appears to offer a chance (Rutkowski: 1994b), the first real chance in humanity's long history (Thomas: 1995, Anderson *et al.*: 1995), for a universally accessible, 'flat' and democratic, autonomous, polycentric, interactive network of low-cost and ultra-fast communications and publication tools and resources. This 'people's network' is now beginning to connect individuals, organizations and communities regardless of their disparate physical locations, time zones, national and organizational boundaries, their peculiar cultures and individual interests. On the other hand, the very creative processes which are responsible for bringing the Internet and the Web into existence simultaneously appear to threaten it with disarray, wasteful repetition and a massive inundation of trivia.

Structurally, this is a situation of perfectly mythological proportions. It is closely akin to the 12th century legend of the Holy Grail and the Knights of the Round Table (Matthews: 1981). In the legend, a great and proud realm is governed by a wise and noble king. The king, however, is stricken down by an illness with many psychological and physical manifestations. His ill health is not limited to his person only, but is inextricably linked to the wilting of all that surrounds the monarch—his faithful people are troubled and uneasy, rare animals are declining, the trees bear no fruit and the fountains are unable to play. The mythological parallels between the story about the Fisher

King and the current predicament of the Internet as a whole and the malaise of its foremost ruler, the World Wide Web, are obvious.

Throughout this paper I shall list and describe some of the major attempts to overcome the current shortcomings of the Web. Whenever possible I shall refer to examples drawn from the widely defined social sciences, including Asian Studies, and the humanities. There appear to be at least six quests for better Web information resources. These quests have been embarked on almost simultaneously by many people, both individually and in concert with each other. Some of these quests are carried out as a series of short bursts of activity, while others are part of long-term, systematic and carefully planned research projects. They all seem to be striving to reach the same goal, although their paths, their adventures, their difficulties and their individual narratives may be dissimilar. They all appear to be focused on the notion of information quality.

These quests for quality—that Grail-like object of intense admiration and longing—that elusive but utterly essential ingredient of all our electronic enterprises, are conducted along six partially different, partially overlapping paths. These are: (1) programming, (2) procedural, (3) structuring, (4) bibliographical, (5) evaluative, and finally (6) organizational approaches to the rescue and repair of the Web information resources.

3. WWW repair

3.1. Programming approaches

The prevailing philosophy here is that once the online publishers are given a wide range of flexible tools for generation and manipulation of hypertext documents the Web will, willy nilly, become a home for the expression of complex, meaningful and elegant thoughts. This, in turn, should provide the necessary stimulus for the widespread acceptance of the Web as an academically acceptable tool for scholarly (Bailey: 1995) and technical publications.

The technical or software engineering approaches are stimulated by the work of T. Berners-Lee and R. Cailliau (creators of the original HTML language and of the original WWW server/client software), as well as that of L. Wall (creator of the Perl language), J. Gossling (creator of the Java language), M. Andreessen (creator of the enhanced HTML; designer of the

contemporary business-strength server/client WWW software), H.W. Lee and B. Bos (creators of the Cascading Style Sheet (CSS1) mechanism) and many other, often anonymous, people. Their programming work, carried out on a number of simultaneous fronts (W3C: 1996a), is focused on broadening and refining functionalities and capabilities the Web documents and their constituent parts.

One of the main areas of programming activities is concerned with expansion and refinement of the hypertext markup language (HTML), including handling of mathematical and scientific equations and formulae as well as of non-latin languages and character sets. Attention is also paid to the future use of the Web for general SGML applications as opposed to dummying SGML down to a simpler HTML format (W3C: 1996b). Important progress has also been made on creation of a Style Sheet Language. That language aims at separating the HTML code, structure and content from the form and appearance of documents. Once implemented, the Style Sheet Language would offer a powerful and manageable way for authors, artists and typographers to create the visual effects (e.g. fonts, colours, spacing) they want without sacrificing the device-independence of their work or adding new HTML tags (W3C: 1996c). Another area of intensive programming work focuses on the Platform for Internet Content Selection (PICS). This is a software tool which would provide content labelling, rating systems for and access control to Web information resources. The self-rating capability offered by PICS enables content providers to describe and label the material they create and distribute. Simultaneously the PICS third-party rating would permit multiple, independent labelling services to associate additional labels with content created and distributed by online authors (W3C: 1996d). Finally, there is intensive work the CGI (Common Gateway Interface) and Java scripts. These programs provide users with the means to create data-input pages for online collection of corrections, feedback and other reader supplied information (Barry: 1996). They also permit construction of advanced interactive graphics, data-processing, data-display, and data-dissemination networked tools (Tessier: 1996).

All these engineering approaches seem to aim at the construction of a series of interlocking modular, intelligent or quasi-intelligent software agents. The idea is to use the software to organise, channel and guide publishing and communication activities on the Web just as roads, tunnels and railway tracks channel, guide and safeguard the flow of wheeled vehicles. This programming approach supposes that such channelling and guidance

will greatly reduce the scope for common errors and blunders and that it will help to make the Web a more comfortable work environment.

3.2. Procedural approaches

The procedural approaches commence with an assumption that people's Web publishing activities are learnable and improvable skills. It appears that all site administration and Web document design, production and maintenance procedures are can be documented, analysed, streamlined, re-organised, taught, and always improved on. An important ingredient of this approach is a belief that through the documentation and careful analysis of the best practices and the most efficient ways of accomplishing a given task, one will be able to progress from a private "realm of art, guess-work and intuition" into the public "realm of craft, routine decision-making and logic" (*cf.* Ciolek: 1995a, 69).

To this end, a number of technical publications (Liu: 1994) and Web sites have sprung up, featuring electronic collections of operating procedures, manuals, templates and 'cookbooks' as well as ample discussion of the Network ethics and Netiquette.

Firstly there are procedures, or sets of instructions, documenting step by step, sequences of minor tasks leading to the successful completion of a major task or activity. Details of complex interactions between the maintainers of a given site, the subtle characteristics of the electronic information they manipulate, as well as their understanding of Web behaviour and structure all need to be recorded and accounted for. Moreover, whether this constitutes a minor or major task depends entirely on the degree of precision with which 'a fractal edge of praxis' is to be handled. As our knowledge of Web operations grows, the amount of detail which begs exact and careful coverage also increases. An example of the sequence of tasks one needs to undertake when setting up a Web page is given in Ciolek (1995b). Each of these major steps, described as 'Data Acquisition', 'Data Preparation', 'Data Formatting', 'Document Naming', 'Directory Naming', 'Document Installation', 'Setting Ownership and Protection Levels', 'Updating Web Indices and Catalogues', 'Connecting Installed Documents to the Web', and finally, 'Document Maintenance', can be further broken down into detailed sub-procedures. It is assumed that, ultimately, it is possible to specify an exact and complete sequence of operations (in short, an algorithm) one needs to invoke in order to perform a particular range of Web publishing

activities. Once this is accomplished there seems a real possibility that an appropriate site/page maintenance-automation software or even a simple CGI/Java helper tool can be written and used on a regular basis.

Another path to the quality of Web resources leads through the adherence to publication guidelines and templates. These documents specify styles and preferred presentation standards for online materials produced by a given team or institution. The guidelines may address such issues as HTML compliance, standard formats, language style, length of documents, use of graphics, titles, links to other documents, backgrounds and other browser-specific extensions, typography, header and footer templates (UMCP Libraries Web Editorial Board: 1996). Instructions may vary quite considerably in terms both of the detail and exactness with which stylistical and editorial decisions are to be handled. Sometimes guidelines make a distinction between mandatory and recommended features of a Web document (Electronic Library Access Committee: 1995). At other times, instructions are to be followed and the templates are expected to be emulated with absolute fidelity. On the whole, the more detailed the instructions, the more unified and elegant is the appearance of a given site. However, such uniformity and precision are usually attained at a cost. The speed with which new materials can be added to the existing collection, as well as the speed with which technological or procedural innovations are adopted, is usually lost.

A third area of improvement to the Web concerne the Netiquette (Network Working Group: 1989), or the 'traffic rules' for ordering and facilitating the interactions between large numbers of strangers working with each other in Cyberspace. Undoubtedly, the great bulk of such notes and 'savoir-vivre' observations applies to the recommended conduct across the USENET news groups and various mailing lists (Gargano: 1989, Berleant and Liu: 1995). However, there also seems to be a slowly developing body of elementary rules for publication and use of the Web-based resources. In late May 1996 the main WWW netiquette points, derived from a guideline by Rinaldi (1996) and also from the work of Ciolek (1996b), clearly separate the responsibilities of readers and authors.

For instance, readers may be advised not to treat the Web as a frivolous playground, avoid impulse-surfing and to conserve the bandwidth by disabling 'graphics load' options on their browsers. They are also advised to notify a page maintainer about errors/mistakes present in his/her document and when doing so to provide complete URLs of the page in question and of the dead link itself.

At the same time publisher and author are urged that when moving a document from one URL to another, they should always leave on the old URL a complete redirection message for the period of at least a few weeks. They are also reminded that links leading to large volumes of data (text, images, video or voice) should also include an indication of their size in Kb. Another set of suggestions is to keep URL naming standards simple and parsimonious in switching between upper and lower case, and to include the option of text links in documents with a large number of graphics. Finally, authors are reminded not to infringe copyright laws or publish obscene, harassing or threatening materials. They are also urged to remember that authors of WWW documents are ultimately responsible for what they allow users worldwide to access.

In the final analysis, the procedural approach suggests that the health and well-being of networked resources is the joint responsibility of its publishers and readers (*i.e.* not of the network owners or various official regulatory bodies regardless how much they would like to exercise such responsibility). Both publishers and readers need to cooperate and guard each other against blunders and abuses which disrupt the system and threaten its long-term viability (Network Working Group: 1989). This sentiment is echoed by A. Rutkowski, who, as the president of the Internet Society, remarked recently that "The Internet is a creature of the unregulated, highly dynamic computer networking field—not the traditional regulated monopoly telecom environment. The Internet does best where the environments are subject to little or no [centralised – tmc] regulation of any kind" (Rutkowski: 1994b).

3.3. Structuring approaches

The 'structuring' approach proposes to cope with the problems of the evergrowing volumes and complexity of rapidly changing networked materials through a system of electronic labels, annotations and meta-data tags (Crossley: 1994, Text Encoding Initiative: 1996a, Rosenfeld: 1996b). Therefore a common encoding scheme is sought for complex textual structures in order to reduce the diversity of existing encoding practices, simplify processing by machine, and encourage the sharing of electronic texts (Sperberg and Burnard: 1991).

Firstly, data-location tags are devised to provide a reader with a means of discovering where information exists and how it might be obtained

or accessed on the network. For instance, the Text Encoding Initiative (1996b) bibliographic tagging captures the intricate distinctions required by most bibliographic systems by establishing at least 27 different fields, such as 'address', 'annote', 'author', 'booktitle', 'chapter', 'date', 'edition', 'editor', 'editors', 'fullauthor', 'fullorganization', 'howpublished', 'institution', 'journal', 'key', 'meeting', 'month', 'note', 'number', 'organization', 'pages', 'publisher', 'school', 'series', 'title', 'volume' and, finally, 'year'.

Secondly, current work on contextual annotation provides a way of placing a given document or database record within a larger corpus of related materials, as well as within the context of an institution responsible for placing it online. For example, an organization may recommend (Australian Department of Defence: 1996) that documents published from its WWW server include the following comment fields:

(i) Identification (must be unique within the system and last for the life of the document);
(ii) Description (author or originator, title, version, date, time of creation, owner or document manager, originating organization, date and time of receipt);
(iii) Responsibility (organizational unit responsible, date and time of registration, template, compound document links, language, format, media, standard used, file number, index or thesaurus terms);
(iv) Status (draft/final? security classification);
(v) Retention/Disposal information (retention period, disposal authority number, disposal status, disposal date).

Thirdly, work is also proceeding on data-filtering. Labels and annotations are increasingly used to describe resources' content, structure and overall characteristics and so give an indication of their fitness for use (Armstrong: 1994). Online resources with annotations offering detailed information such as consistency of data, accessibility/ease of use; coverage/scope; timeliness of updates; error rate/accuracy; integration across documents and records; supported output formats; documentation, and value to cost ratio certainly allow for better or quicker data filtering. This is a feature especially useful when dealing with masses of materials floating on the WWW system which since January 1991 has grown from a couple of hundreds of hypertext documents mounted on a handful of machines, to a collection which in mid-May 1996 comprised 30 million pages found on 225,000 servers—and in which the search on the keyword "Dalai Lama" returned over 3,000 unique records (Altavista: 1996).

Finally, a detailed markup of the document can be embarked on in order to provide multiple ways of viewing and analysing information contained in a given collection of documents. For instance, work conducted by U. App and C. Wittern in the field of ancient and medieval Chinese Buddhist texts consists of several mark-up 'sweeps' done for each of the documents (Mohr: 1996). First of all, there is a basic structural markup, aimed at separating and annotating the logical divisions and elements of the document. The second stage of work is focused on content markup. Specialist tags in a document are created to mark all occurrences of personal names (e.g. names and titles of Ch'an/Zen teachers, monks and government officials); place-names (e.g. cities, rivers, lakes, mountains, temples and monasteries); names of documents (sutras, collections, biographies of famous monks etc.); dates; philosophical and religious concepts, and so forth.

This stage of work often takes weeks or months and needs to allow for easy customization and gradual addition of new tags depending on ones' familiarity with the content and context of the tagged materials. Finally, upon completion of the content-markup the document may be passed to data specialists who carry out overall SGML markup using dedicated editing software. The SGML encoding (Goldfarb: 1990) allows for great flexibility in providing texts for world-wide network delivery. This is possible since the mark-up separates presentation and formatting information from structure and content information, and facilitates display on different devices. Also, the 'grainy' or structured nature of the marked-up documents allows for their transmission as fragments rather than as entire text. SGML tagging also allows for more focused information retrieval operations on a given corpus of texts.

Documents and resources, when fully tagged and marked up according to the best of the SGML and Text Encoding Initiative standards (TEI: 1996a), can then be turned into specialist knowledge systems. These systems would be able to offer multiple and increasingly complex views of the same electronic corpus of information to users. The original body of texts would be seamlessly linked with supporting raw-data, additional documentation, commentaries and annotations, bibliographies, as well as with external calls to various databases, interactive maps, and finally to appropriate sound and video-resources. A first step in that direction has been already taken in the form of ZenBase CD1 (App: 1995).

3.4. Bibliographical approaches

There is also a growing consensus among the users of the World Wide Web that unless there are adequate, consistent and simple means for academic referencing of the whole range of the networked information, the Web resources will not be awarded full recognition within academic discourse. A. Greenhill and G. Fletcher (1995) wrote: "Unless corrected, the significance of this oversight will be exacerbated as more academic journals become available on-line and more computer literate students enter tertiary study. Furthermore, the status of researchers who have published in this medium will be affected and universities may deprive themselves of the staff best equipped to meet the challenges of the information economy."

Within the last decade or so, a number of new tools for delivery of scholarly or factual information have been developed. These are: E-mail messages, FTP (File Transfer Protocol) files, FTP Mailserv files, Gopher files, Listserv messages, Online databases and records, Standalone databases and records, Synchronous Communications (MOOs, MUDs, IRC, etc.) transactions, Telnet sites and files, USENET news, Web files as well as a variety of specialist computer programs (applications). How great an interest there is in proper bibliographical referencing of those sources of information can be gauged from the fact that one of the online guides to Citations of Electronic Sources (Walker: 1995) was accessed not less than 3,811 times within a 15 day period 2–17 May 1996.

While computer-mediated tools greatly increase the range and speed with which data are delivered, they also display a number of characteristics not usually found in conventional paper publications.

Firstly, many online resources are highly unstable and changeable. Secondly, electronic materials frequently lack the complete set of data about their author and the document itself. Moreover, many sources of information, like e-mail, listserv or IRC messages, frequently do not have a fixed abode on the network. Another complication arises from the fact that electronic documents, as a rule, do not possess the pagination structure so typical of paper publications. Finally, electronic documents are extremely sensitive to slightest typographical changes to their addresses (URLs). Also, it is a common practice for users of online documents to copy-and-paste the listed URLs (as opposed to more time-consuming retyping) and incorporate them into their own web pages. This means that the URLs need to be published in citations verbatim, that is without any embellishments and paper-style 'packaging', such as brackets, quotes or full stops.

Even at this early stage in the development of the Internet, numerous schemes have already been proposed to tackle the issue of scholarly referencing of online materials (Walker: 1995, Lee and Crane: 1993, Lee and Crane: 1996). A systematic attempt to collate information on the complete range of these approaches has been already undertaken by A. Greenhill (1996). An overview of existing works aimed at the development of common reference standards suggest that the world of electronic citations seems to be governed by three strongly interacting forces:

(a) the body of existing conventions developed for the realm of paper-based publications;
(b) the body of emerging conventions for keeping track of network-based publications;
(c) the pragmatics of the reader's behaviour, always focused on the directness, ease and speed with which tasks can be performed.

Since the realm of the networked publication is largely a product of the grass-root and user-driven decisions and developments, one may conclude that none of the elaborate, detailed and highly embellished citations schemes, such as the ones proposed by Page (1996) will attract much following. By the same token, minimalistic conventions involving just a handful of simply presented fields (*i.e.* 'surname', 'name', year', 'document title' and 'url'), such as those developed by Li and Crane (1996 – APA Style) or Greenhill and Fletcher (1995) do indeed have a chance of becoming a *de facto* standard.

3.5. Evaluative approaches

The evaluative approaches to the Web start with an assumption that the networked information resources, however dissimilar they might be, share in fact a number of common characteristics or features, and that they can be graded or rated in terms of 'scores' received for each of those features. Ideally, such evaluations should be a simple procedure, so that they could be automated and carried out by a piece of software. However, at present only labour intensive, and often idiosyncratic, manual processes are being devised.

The steadily growing interest in the techniques suitable for assessing and comparing Web resources has lead recently to the creation of specialist sites monitoring practical and methodological developments in this area

(Smith: 1996a, Auer: 1996). The 'evaluative' activities seem to form two main streams: (a) individual work on creation of checklists or "toolboxes" of criteria that enable WWW information sources to be assessed, and (b) commercial, long-term projects aimed at the periodical reviews and gradings of large volumes of online material. In the first case, the emphasis is on finding how the overall quality of the networked resources can be meaningfully discerned, analysed and compared. In the second case, the emphasis is on a quick separation of potentially popular materials from the rest of the Web so that a site providing such rudimentary 'filtering' services can attract Internauts and draw them towards the site's fee-based operations.

The first group of approaches is represented by the works of Caywood (1995), Ciolek (1996c), Smith (1996b), Tillman (1996) and Grassian (1996) who attempt to specify and enumerate the essential ingredients, or features, of a 'good' or 'high quality' or 'useful' Web resource. Thus the proposed indicators of quality of the Net resources involve summaries of characteristics such as:

(i) 1. ease of access, 2. good design, 3. good content [a three point synthesis of the 27 items long check-list (Caywood: 1995)].

(ii) 1. uniqueness of information, 2. ease of finding it on the net, 3. ease of access, 4. good structure and organization, 5. good formatting and presentation, 6. usefulness and trustworthiness, 7. ease of resource maintenance [a seven point synthesis of the 115 item-long check-list (Ciolek: 1996c)].

(iii) 1. scope (breadth, depth, time, format [type of resources covered]), 2. content, 3. accuracy, 4. currency, 5. authority, 6. format and appearance, 7. audience, 8. purpose, 9. uniqueness, 10. workability (user-friendliness, search facilities, connectivity), 11. cost [an eleven point synthesis of the 54 item-long check-list (Smith: 1996b)].

(iv) 1. ease of determining the resource's scope, 2. ease of identifying the meta-data (the authority of authors, the currency of information, the last update, the nature of the updates), 3. stability of information, 4. ease of use [a four point synthesis of the 10 item-long check-list (Tillman: 1996)].

(v) 1. content and evaluation, 2. source and date, 3. structure, 4. other issues [a four point synthesis of the 44 item-long check-list (Grassian: 1996)].

Each of those criteria is based, in turn, on a series of more detailed questions and sub-questions. For instance, Cayman's (1995) criterion of the 'ease of access' relies on answers to the following checkpoints: "Is the site

still useful with an ASCII browser like *Lynx*? Is it written in standard HTML, or have proprietary extensions been used? Does it use standard multimedia formats? Do parts of it take too long to load? Is it usually possible to reach the site, or is it overloaded? Is it stable, or has the URL changed? Is it open to everyone on the Internet, or do parts require fees? Are any rules for use stated up front?"

The commercial approaches are best represented by work initiated by McKinley/Magellan site (1996) and, independently, by the Point Corporation (1996). Magellan is an online guide to the Internet that includes a directory of tens of thousands rated and reviewed Internet sites and a vast database of yet-to-be-reviewed sites. Magellan covers Web sites, FTP and gopher servers, newsgroups, and Telnet sessions. An excerpt from the 'Frequently Asked Questions' file (McKinley: 1996) says:

> Q: What kinds of sites does Magellan review?
> A: We aim for a lively mix of sites, from familiar Internet favourites to the newest of the new, in all of our subject areas [...] Magellan does not review sites relating to pornography, paedophilia, or hate groups.

The rating procedure, adopted by commercial sites, is simple. Magellan reviewers evaluate each of the selected Web sites, awarding from one to ten points in three criteria: 'Depth' (= is the site comprehensive and up-to-date?); 'Ease of Exploration' (= is the site well-organized and easy to navigate?), and finally 'Net Appeal' (= is the site innovative? Does it appeal to the eye or the ear? Is it funny? Is it hot, hip, or cool? Is it thought-provoking? Does it offer new technology or a new way of using technology?). The final result of these operations is an overall rating of one to four Magellan stars, depending on the number of points awarded to a given resource: one star (1–12 points), two stars (13–21 points), three stars (22–27 points), and four stars (28–30 points). A similar procedure is adopted by the Point Corporation which aims "to point out the good stuff, save you time, and help you to achieve 100% pure surfing pleasure" (Point: 1996) and which evaluates the sites on a scale from one to 50 points. Their three criteria are: 'Content' (= how broad, deep, and thorough is the information? Are there good links? Good clips? Is it accurate? Complete? Up-to-date?); 'Presentation' (= is the page beautiful? Colourful? Easy to use? does it lead readers through the information nicely? Does it break new ground?); and 'Experience' (= Is this fun? Is it worth the time? Would you recommend it to your friends?).

The work on evaluative approaches, scholarly and commercial alike, has barely started. It raises, however, a number of methodological questions.

Firstly, the selection criteria used in the reviewed evaluative procedures tend to be very general indeed. Concepts such as 'ease of access' or 'user-friendliness' or 'crisp page layout', 'detailed meta-data' seem to be applied to the online materials in a very general fashion, as if all documents and all resources were written in the same natural language, had the same complexity, same structure, and served the same purpose. Can one really use the same vague, impressionistic concept to compare a single document with a collection of research papers, and finally, with a large-scale electronic archive? One would think not.

Furthermore, the operational meaning of each of the employed criteria remains unclear. Does the notion of 'workability' (Smith: 1996b) refer to the same phenomenon identified as 'ease of finding, ease of access, good formatting and presentation' by Ciolek (1996c); or that referred to as the 'stability of information and ease of use' by Tillman (1996)? Also, how does one go about measuring the breadth, depth or thoroughness of information? Moreover, what does it mean that a page or a graphic image may 'take too long to load'? How many seconds, and under what circumstances, are considered to be an acceptable waiting time? Another, and related problem, is that of the intra- and inter-evaluator consistency of the rating procedure. Ideally, one should expect that the same material, when evaluated at different times, will invariably receive the same score. Similarly, various judges, while using the same checklist of questions, should give the same site overwhelmingly similar scores.

It can be seen, therefore, that what we tend to call summarily 'the quality' or 'the value' of an information resource is, in fact, a product of complex dynamic interactions between a large number of variables.

For example, if we talk about electronic information in general, we could start by enumerating such resources as the FTP, WWW, Gopher, Telnet, E-mail, Listserv, IRC and so forth. On the other hand, if we focus on the Web information resources alone, then we would do well if we listed such facilities as data-files, on-line papers, e-journals, resource-guides, and home pages of various research projects. Finally, if we concern ourselves solely with the WWW-based e-journals then we should be able to make a number of distinctions between the journal's title-page, its masthead and section on editorial policies, the table of contents of the entire journal, the

table of contents for a given issue, individual articles of the journal, and so on.

There are also many aspects of each of these types of information. The most obvious aspects are:

(1) Language of the online information (e.g. text-based information may be expressed in English, German, Sanskrit, Korean etc.). Some of these languages depend on a simple set of 26 Latin characters, others require the use additional accented characters, others still are based on double-byte codes necessary for accurately mapping the tens of thousands of ideographs;
(2) Encoding (e.g. ASCII, Big5, Unicode, number formats, date formats, etc.);
(3) Accuracy of the information, or its relationship to that which it attempts to represent (e.g. completeness of the data; presence/absence of spelling and typing errors; the handling of accents, macrons and diacritics, etc.);
(4) Size of the information (e.g. measured in number of characters, number of computer screens or in kilobytes and megabytes);
(5) Structure (e.g. division into chapters, sections and paragraphs; organization of documents into linear sequences, circles, hub and spokes, trees, lattices and random access [database] systems);
(6) Layout (e.g. arrangement and placement of the information on the screen), and finally
(7) Presentation (e.g. choice of typography, font sizes, use of colour, use of decorative material, etc.).

Finally, it appears that one can distinguish at least five levels of networked information:

(1) Pointer—the address of a unit of information (e.g. hypertext link (URL), details of the subdirectory path and filename; database name and unique keyword combination; bibliographical references, film and frame number, record and track ID number). The pointer seems to consist of the actual address and any number of associated labels and annotations which comment on the object targeted by the address.
(2) Item—the minimal addressable unit of information (e.g. a line in a document, a paragraph, a chapter, a table of contents, a graph, a table with statistical data). An item usually consists of a body of text with or without a certain number of item-specific pointers.
(3) Document—a coherent collection of information items (e.g. FTP document, Web page, database record, e-mail message, letter, memorandum,

article, slide, photograph, video-clip, sound track). A document is a mosaic constructed from information items and a number of document-specific pointers.

(4) Resource—a coherent, annotated collection of documents (e.g. FTP archive, database, WWW publication, journal, book, telephone directory, video cassette, LP record, CD, audio cassette). In other words, a resource is a complex mosaic constructed from several interrelated documents as well as from resource-specific information items and pointers.

(5) Information system—a coherent, well-annotated, indexed and cross-referenced collection of resources (e.g. a virtual library, encyclopaedia, photo-archive, video-library, sound archive, record library) and the interconnecting pointers.

A brief glimpse of how all these variables may interact with each other is offered by Table 1. A closer look at the table suggests the following remarks and comments.

The exact meaning of a given aspect of information, such as 'Size' or 'Presentation', depends on the level at which it is applied. Thus the notion of optimal size in terms of a 'Document' (e.g. article in a journal) is not identical with the optimal size for a 'Resource' (e.g. journal itself). Furthermore, each of the above matrix cells, formed by an interaction between aspect and level of information is capable of generating a large number of detailed queries. For instance, the cell "size × pointer" inevitably leads to a discussion of not only of the maximum acceptable number of characters within a URL itself, but also of their maximum number within any label attached to it, as well as within any annotations and commentary fields accompanying a given URL. Also, it is advantageous if questions are specific and practical, and responses to them are as detailed and factual as possible (e.g. 'under 60 chars'). This is important if a given practical solution is to be evaluated, revised and improved upon. Also, detailed, practical specifications are easier to work with, even if they are initially erroneous, than more general ones (e.g. 'professional feel', or 'short loading time').

Another observation is that the presence of occasional un-answered questions (marked with "???") indicates that at least one of the variables involved was couched in a too general a fashion and the wording needs to be re-cast in practical terms (e.g. "what is the most suitable layout for a help file (as opposed to a meta-data file) in a WWW-based large scale information system?" or "where exactly should the hypertext pointers be located on a TOC page of an online journal?").

Table 1

Quality issues and concerns in the WWW scholarly publications

LEVEL / ASPECT	Pointer	Item	Document	Resource	Info. System
Language	ease of writing and copying of the URL	legibility	legibility of headers and footers	universal legibility of TOCs and indices	universal legibility of TOCs, indices help files
Encoding	universal legibility ASCII code	good handling of symbols and local characters	handling of local language and English	uniformity within a resource	good handling of all possible encoding systems
Accuracy	absolute freedom from errors	good handling of numbers and accents	completeness of data	timeliness of data	completeness of coverage
Size	brevity URL size < 60 characters	several fast loading items	fast loading document size < 10–15 kb	fast switching to another document	fast switching to another resource and help files
Structure	URL + label + annotations	logical sequence navigation	good intra-document navigation	good inter-document navigation	good inter-resource navigation
Layout	Navigation links in standard locations	crispness, legibility, clarity	main info at the top of document	main info at the top page	???
Presentation	clarity and cripness	understated, professional feel	short loading time, good taste	consistency uniformity of 'feel'	differentiation between resources

Finally, Table 1 suggests that simpler levels of information appear to be less redundant, and less tolerant of errors, mistakes, and shortcomings, than the higher levels of information. A single typing error at the level of a pointer is more detrimental to the useability of on-line materials than the identical transposition of characters at the level of information item (e.g. footnote) or at the level of a whole document (e.g. research paper). Similarly, a single typing error within a pointer at the level of a URL is more critical than an identical transposition of characters within a label or an annotation accompanying that URL.

In sum, in order to speak intelligently about such a general concept as 'the quality of information' we have to undertake a careful and detailed analysis involving two simultaneous procedures. Firstly the complete range of performance criteria such as 'stability of information', 'ease of navigation', 'currency of information' or 'net-appeal' needs to be systematically mapped onto a detailed matrix. This matrix is formed by the intersection of all the variables comprising the types, aspects and organizational levels of information. Secondly, results of such mapping have to be related to the overall context of users' (readers') network hardware, software as well as their previous experiences, expectations and knowledge. None of these tasks is easy or can be carried out in a mechanical fashion.

Clearly further intensive work needs to be done in this area. It might, perhaps be furthered by a closer involvement of various professional associations and learned societies, so that comparisons and ratings of various Internet sites and resources are not only done on a regular basis, but are done in a replicable and meaningful manner.

3.6. Organizational approaches

Finally, the organizational approaches assume that the prevailing chaos, methodological shortcomings and scattering of efforts can be overcome through energetic and 'competitive cooperation' (Rutkowski: 1994a) between various individuals and institutions with a stake in the Web. The good-natured competitiveness among the players is assured and reinforced by the continued adherence to the heterogeneous, distributed, polycentric model of Internet activities. Cooperation, on the other hand, is founded on voluntary agreement to share and circulate relevant information and to delineate spheres of activity in order to avoid any major encroachment on a colleague's field of expertise.

Among the activities concerned with the organizational adjustments of the Web there are three projects which deserve special praise. These are: the "WWW Virtual Library Project (WWW VL)" (Secret: 1996) initiated in 1991 by T. Berners-Lee; "The Clearinghouse for Subject Oriented Guides to the Internet", created in 1993 by L. Rosenfeld (1996a); and the "Special Interest Networks (SINs)".

One of the best examples of successful cooperation on the Web is provided by the Special Interest Networks (SINs), which were first proposed in 1994 by D. Green and J. Croft (1994). This idea draws extensively on the examples set both by the WWW VL and "The Clearinghouse". It also broadens them into a widely cast network of expert sites that collaborate to provide a complete range of information activities for a given subject area (Green: 1995). According to Green-Croft's approach, SINs should combine the roles of information suppliers, distributors and users. They should be able to act as the Web equivalents of professional and scholarly societies and as the electronic counterparts to the traditional libraries. Therefore, the SIN nodes (specialist WWW sites) need to be dedicated to:

(i) promotion of communication among the networked scholars,
(ii) development of research tools and resources through maintenance of specialist virtual libraries and stable repositories of knowledge and information; and
(iii) speedy dissemination of research results through online publication of data and analyses.

In addition, SINs could also offer a fourth function:

(iv) provision of expert information services to governmental and commercial clients.

Such SIN nodes are expected to provide the necessary organization (= ensuring that users can obtain information easily and quickly), stability (= ensuring that sources remain available and that links do not go 'stale'), quality (= ensuring that the data are accurate and up-to-date) and standardization (= ensuring common, regular format for collection and interchange of the data and documents) to the network-based information. The coordination between activities of constituent SIN sites is achieved through their logical design as well as through well-planned division of responsibilities. It is also attained through automation of data maintenance tasks, systematic mirroring of each other's data collections, adherence to jointly developed standards, and observance of the uniform quality control measures. To be

successful, SINs should strive to provide reliable and authoritative online information services, to encourage participation among qualified researchers, and to accommodate the inevitable growth both in data holdings and in the scale of Internet operations.

The subsequent elaborations of the theme (Green: 1995) also suggest that the Special Interest Networks are well equipped to handle the explosion of the networked information, its sheer volume, the rapid turnover and change (especially the need to maintain information up-to-date), and the proliferation of its forms (paper, microfilm, CD-ROM, off-line computer files, and online documents). According to D. Green (1995, 17) "the SIN model provides a user-driven solution, in which groups of people interested in a particular topic organise and index information in the ways they find most useful. The twenty-first century will surely become the era of the knowledge web. SINs, in whatever form they may take, will play a major role in its organization."

4. Concluding remarks

In the medieval legend, of the many who set out on the long and arduous quest, few ever catch more than a fleeting glimpse of the elusive Grail, and only three (Galahad, Parcifal and Bors) succeed in finding it and bringing it to the ailing King.

The contemporary legend of the Internet is even more complex and more demanding than its illustrious predecessor. The long-term viability of the Web as a medium for scholarly publications requires that the best techniques, best practices, and the best methodologies be not only searched for, and not only glimpsed but also that they be found, documented and widely disseminated. This has to be accomplished on the widest possible scale and swiftly, before the Web dissolves into an amorphous mass of repetitive, indifferent, and dubious informational snippets.

As this author wrote in January 1996 (Ciolek: 1996a, 108): "the WWW system has reached a cross-roads. Since its inception in 1991 [...] the WWW-based information, tracked by dozens of Web Crawlers and Harvesters, continues to grow exponentially without much thought for guidelines, safeguards and standards concerning the quality, precision, trustworthiness, durability, currency and authorship of this information. This

situation is untenable. Unless serious and energetic remedial steps are taken [...] the system currently known as the WWW may need to be redesignated as the Multi-Media Mediocrity, or the MMM, for short." Half a year later, in mid-1996, the urgency of the repair tasks has grown even stronger.

Acknowledgements

I am indebted to Philipp Kelly and Allison Ley for critical comments on the first draft of this article. This paper is a result of the many rewarding years of work with my colleagues from the ANU Coombs Computing Unit team, especially with Rob Hurle and Sean Batt, and with such exemplary and inspiring Internet friends as Irena Goltz, Maureen Donovan, and Arthur Secret.

Bibliography

Altavista: 1996, "The Internet's Home Page",
http://altavista.digital.com/

ANDERSON (Robert H.) *et. al.*: 1995, "Universal Access To E-mail-Feasibility and Societal Implications", *RAND Report MR-650-MF.*
http://www.rand.org:80/publications/MR/MR650/

APP (Urs): 1995, ed., *ZenBase CD1* (Kyoto: International Research Institute for Zen Buddhism), ISBN 4-938796-18-X. See also
http://www.iijnet.or.jp/iriz/irizhtml/irizhome.htm

ARMSTRONG (Chris, J.): 1994, "Databases and Quality: Why not try 'What You See Is What You Get'?", *Managing Information*, Nov./Dec., 1,
gopher://ukoln.bath.ac.uk:7070/00/BUBL_Main_Menu/H/H2/H2C/H2C28/ ✂
H2C28007_-_Managing_Information_article_reprint

AUER (Nicole): 1996, "Evaluation of Internet information resources—a bibliography",
http://www.vuw.ac.nz/~agsmith/evaln/poll.htm#auer

Australian Department of Defence: 1996, *Interim Guidelines for Establishing a WEB Information Service*,
http://www.adfa.oz.au/DOD/aboutgd1.html

BAILEY (Charles W.): 1995, "Network-Based Electronic Publishing of Scholarly Works: A Selective Bibliography", *The Public-Access Computer Systems Review* 6, no. 1. (Version 21: 5/24/96).
http://info.lib.uh.edu/pr/v6/n1/bail6n1.html

BARRY (Anthony): 1995, *NIR in not enough*,
http://snazzy.anu.edu.au/CNASI/pubs/Questnet95.html

BARRY (Anthony): 1996, *Libraries, the Web, Interactive Forms and CGI Scripts*, http://snazzy.anu.edu.au/CNASI/pubs/vala96.html

BERLEANT (Daniel) and LIU (Byron): 1995, "Robert's Rules of Order for e-mail meetings", *IEEE Computer*, 28, 11 http://info.computer.org/pubs/computer/kiosk/11/kiosk.htm

CAYWOOD (Carolyn): 1995, *Library Selection Criteria for WWW Resources*, http://duckdock.acic.com/carolyn/criteria.htm

CIOLEK (T. Matthew): 1995a, "Ensuring High Quality in Multifaceted Information Services", in *Proceedings of the AUUG'95 and Asia-Pacific WWW'95 Conference*, Sept. 17–21 1995, Sydney, Australia, pp. 68–75. http://coombs.anu.edu.au/SpecialProj/QLTY/EnsuringQlty.html

CIOLEK (T. Matthew): 1995b, "Procedures for Document Publication on the Coombsweb WWW Server", http://coombs.anu.edu.au/SpecialProj/QLTY/PROC/proc-WWW.html

CIOLEK (T. Matthew): 1996a, "Today's WWW–tomorrow's MMM? The specter of multi-media mediocrity", *IEEE Computer*, 29, 1, pp. 106–108. http://www.computer.org/pubs/computer/kiosk/01/kiosk.htm

CIOLEK (T. Matthew): 1996b, ed., *Asian Studies WWW Virtual Library*, http://coombs.anu.edu.au/WWWVL-AsianStudies.html

CIOLEK (T. Matthew): 1996c, *Quality Info. Systems—Catalogue of Potent Truisms*, http://coombs.anu.edu.au/SpecialProj/QLTY/QltyTruisms.html

CLARKE (Roger): 1996, *Net-Etiquette—Mini Case Studies of Dysfunctional Human Behaviour on the Net*, http://www.anu.edu.au/people/Roger.Clarke/II/Netethiquettecases

CROSSLEY (David): 1994, "WAIS through the Web—Discovering Environmental Information", *Proceedings of the The Second International WWW Conference (WWW Fall 94) Mosaic and the Web – Chicago, USA (17–20 October, 1994)*. http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/Searching/crossley/paper.html

Electronic Library Access Committee (ELAC): 1995, *UW-Madison Campus Libraries Web Page, Standards and Guidelines*, http://www.library.wisc.edu/help/tech/Web_standards.html

ERICKSON (Jonathan): 1996, "Excellence in Programming Awards", *Dr. Dobb's Journal*, 245, March, pp. 16–17.

GARGANO (Joan): 1989, *A Guide to Electronic Communication & Network Etiquette*, gopher://leviathan.tamu.edu/00h/internet/etiq.txt

GOLDFARB (Charles F.): 1990, *The SGML Handbook* (Oxford: Oxford University Press).

GRASSIAN (Esther): 1996, *Thinking Critically about World Wide Web Resources*, http://www.ucla.edu/campus/computing/bruinonline/trainers/critical.html

GREEN (David G.): 1995, "From Honeypots to a Web of SIN—Building the World-Wide Information System", in *Proceedings of the AUUG'95 and Asia-Pacific WWW'95 Conference*, Sept. 17-21 1995, Sydney, Australia, pp. 11–18.
http://www.csu.edu.au/special/conference/apwww95/papers95/ ✂
dgreen/dgreen.html

GREEN (David G.) and CROFT (Jim): 1994, "Proposal for Implementing a Biodiversity Information Network", in *Linking Mechanisms for Biodiversity Information*, (Proceedings of a Workshop for the Biodiversity Information network, Base de Dados Tropical, Campinas, Sao Paulo, Brasil).
http://www.ftpt.br/bin21/proposal.html

GREENHILL (Anita) and FLETCHER (Gordon): 1995, *A Proposal for Referencing Internet Resources*,
http://www.gu.edu.au/gwis/hub/hub.acadref.html

GREENHILL (Anita): 1996, ed., *Electronic References & Scholarly Citations of Internet Sources*,
http://www.gu.edu.au/gint/WWWVL/OnlineRefs.html

LI (Xia) and CRANE (Nancy): 1993, *Electronic Style: A Guide to Citing Electronic Information* (Westport: Meckler).

LI (Xia) and CRANE (Nancy): 1996, *Bibliographic Formats for Citing Electronic Information*,
http://www.uvm.edu/~xli/reference/estyles.html

LIU (Cricket) et al.: 1994, *Managing Internet Information Services* (Sebastopol, Ca.: O'Reilly & Associates, Inc.).

McKinley (The ~ Group, Inc.): 1966, "How are the sites rated?",
http://www.mckinley.com/mckinley-txt/250.html

MATTHEWS (John): 1981, *The Grail—Quest for the Eternal* (London: Thames & Hudson).

MILLER (George): 1996, ed., *Indonesian Serials Database—Database Majalah Indonesia*,
http://coombs.anu.edu.au/SpecialProj/AJC/IND/Indonesia-jrnls.html

MOHR (Michel): 1996, *The Taipei meeting of the Electronic Buddhist Text Initiative—Impressions by Michel Mohr*,
http://www.iijnet.or.jp/iriz/irizhtml/ebti/taipei.htm

Netree.Com: 1996, "Netree Internet Statistics—Estimated"
http://www.netree.com/netbin/internetstats

Network Working Group: 1989, *Ethics and the Internet*,
`ftp://ds.internic.net` in the directory `/rfc/rfc1087.txt`

PAGE (Melvin E.): 1996, *A Brief Citation Guide For Internet Sources in History and the Humanities*,
http://h-net.msu.edu/~africa/citation.html

Point Communications Corporation: 1996, *Frequently Asked Questions about Point*,
http://www.pointcom.com/gifs/welcome/

RINALDI (Arlene H.): 1996, *The Net: User Guidelines and Netiquette*,
http://www.fau.edu/faahr/netiquette.html

ROSENFELD (Louis B.): 1996a, ed., *Clearinghouse for Subject Oriented Guides to the Internet*,
http://www.lib.umich.edu/chouse/chhome.html

ROSENFELD (Louis B.): 1996b, "Label Laws: Some rules for clearly identifying content" *Web Architect*, 29 March 1996,
http://www.gnn.com/gnn/wr/96/03/29/webarch/index.html

RUTKOWSKI (Anthony-Michael): 1994a, *Today's Cooperative Competitive Standards Environment for Open Information and Telecommunication Networks and the Internet Standards-Making Model*,
http://info.isoc.org/papers/standards/amr-on-standards.html

RUTKOWSKI (Anthony-Michael): 1994b, *The Present and Future of the Internet: Five Faces*,
http://info.isoc.org/speeches/interop-tokyo.html

SECRET (Arthur): 1996, ed., *World Wide Web Virtual Library*,
http://www.w3.org/hypertext/DataSources/bySubject/Overview.html

SIEGEL (David): 1995, *The Balkanisation of the Web*,
http://www.dsiegel.com/balkanization/intro.html

SMITH (Alastair): 1996a, ed., *Evaluation of Information Resources*,
http://www.vuw.ac.nz/~agsmith/evaln/evaln.htm

SMITH (Alastair): 1996b, *Criteria for Evaluation of Internet Information Resources*
http://www.vuw.ac.nz/~agsmith/evaln/index.htm

SPERBERG-MCQUEEN (C.M) and BURNARD (Lou): 1991, eds., *Guidelines for the Encoding and Interchange of Machine-Readable Texts*. (Chicago, Oxford: Text Encoding Initiative).

Text Encoding Initiative (TEI): 1996a, *TEI Guidelines for Electronic Text Encoding and Interchange (P3)*.
http://etext.virginia.edu/TEI.html

Text Encoding Inititative (TEI): 1996b, *Bibliographic Citations and References*,
http://etext.virginia.edu/bin/tei-tocs?div=DIV2&id=COBI

TESSIER (Tom): 1996, "Using JavaScript to Create Interactive Web Pages", *Dr. Dobb's Journal*, 245, March, pp. 84–89.

THOMAS (Hugh): 1995, *An Unfinished History of the World* (London: Papermac & Macmillan General Books).

TILLMAN (Hope N.): 1996, *Evaluating Quality on the Net*, http://challenge.tiac.net/users/hope/findqual.html

TRELOAR (Andrew): 1995, "Scholarly Publishing and the Fluid World Wide Web", in *Proceedings of the AUUG'95 and Asia-Pacific WWW'95 Conference*, Sept. 17–21 1995, Sydney, Australia, pp. 326–332. http://www.csu.edu.au/special/conference/apwww95/papers95/ ✂ atreloar/atreloar.html

Umcp Libraries Web Editorial Board: 1996, *Style Guide for Authors of Web Pages*, http://www.itd.umd.edu/UMS/UMCP/BOARD/style_guide.html

WALKER (Janice R.): 1995, *MLA-Style Citations of Electronic Sources*, http://www.cas.usf.edu/english/walker/mla.html

W3C (World Wide Web Consortium): 1996a, *W3C Activity areas*, http://www.w3.org/pub/WWW/Consortium/Prospectus/ActivityList

W3C (World Wide Web Consortium): 1996b, *Hypertext Markup Language (HTML)*, http://www.w3.org/pub/WWW/MarkUp/Activity.html

W3C (World Wide Web Consortium): 1996c, *Web Style Sheets*, http://www.w3.org/pub/WWW/Style/

W3C (World Wide Web Consortium): 1996d, *Platform for Internet Content Selection (PICS)*, http://www.w3.org/pub/WWW/PICS/