

The evolving use of SGML in electronic journals

Judith C. Wusteman

Department of Library and Information Studies

University of Dublin

Ireland

e-mail: j.c.wusteman@ucd.ie

ABSTRACT

The application of SGML in e-journal production and presentation is steadily increasing. This paper discusses some current experiments in the use of SGML for the representation of both bibliographic information and full article text. It describes the 'standard' journal DTDs currently available in the STM domain and discusses their use.

1. Introduction

Although the SGML standard has been in existence since 1986, its use in both paper and electronic journals has become widespread only since the mid-1990s. Initial experiments by publishers and others are raising a raft of issues which need to be resolved. Section 2 of this paper introduces examples of the application of SGML in current e-journal projects. Some projects, such as *Infobike/JournalsOnline*, are using SGML solely for the representation of bibliographic information. Others, such as *SuperJournal*, CLIC and the Illinois DLI project, are using SGML for the full article text. Section 3 describes the 'standard' Document Type Definitions (DTDs) available for header information and full article text in the Science, Technical and Medical (STM) fields. Comments are made on the use of these DTDs in the projects listed in Section 2. The paper concludes with a discussion on the appropriate design of DTDs for e-journals.

2. Projects involving SGML

Despite the increasing prominence of SGML in the journals arena, only three of the sixty-plus projects funded by the UK's eLib (Electronic Libraries) programme [1] use SGML to any appreciable extent. These three, *SuperJournal* [2], *Infobike/JournalsOnline* [3] and CLIC [4], are some of the larger projects in the eLib programme. The majority of eLib e-journal projects involve the use of HTML or PDF. The Programme Director, Chris Rusbridge, has suggested that HTML should be regarded as a short-term solution 'with perhaps a medium term aim of migrating to SGML and a suitable DTD' [5].

Both *Infobike/JournalsOnline* and *SuperJournal* use SGML as an intermediate format for the provision of bibliographic information by publishers to the system. Various commercially-based systems, for example BH Blackwell's *Electronic Journal Navigator* [6], use SGML in a similar way. The use of SGML to represent bibliographic information is discussed in Section 3.1. The *SuperJournal* and CLIC projects are experimenting with the use of SGML markup of full article texts; in both cases, the SGML article files are converted to HTML for display. This use of SGML is discussed in Section 3.2.

The US Digital Libraries Initiative (DLI) programme has taken an alternative approach to eLib: instead of funding a large number of relatively small projects, it is funding six major projects. It is noticeable that this has resulted in a much greater use of SGML, presumably because of the initial start-up costs involved in its use. One project in particular, the Illinois

DLI project [7], is developing a large testbed of full-text journal articles in SGML. The issues raised in this project will become increasingly relevant to publishers in the UK as they begin to offer full article SGML provision.

Various other projects are also experimenting with the use of SGML for e-journal articles. WebDOC [8], for example, is a joint project between the Dutch libraries foundation, Pica, and the US-based Research Libraries Group. A central catalogue, WebCAT, is being developed which provides access to a large collection of electronic documents from international publishers and from libraries in the Netherlands, Germany and the USA. SGML is being used to represent bibliographic information.

Another example is ELSA (Electronic Library SGML Applications) [9], a collaboration between Jouve S.A., Elsevier Science and De Montfort University. ELSA is part-funded by the Commission of the European Community under the Library Programme of DG XIII [10]. It involves the development of a prototype system for the storage and retrieval of journal articles in SGML from approximately 100 Elsevier Science journals. The 5,000 articles are tagged using the publicly available Elsevier Science Article DTD version 3.0.0. Jouve are developing a proprietary SGML viewer which will be hard-wired for the Elsevier DTD [11].

3. Standard DTDs

When SGML first came to prominence in journal publishing, it was widely assumed that it would be possible to develop and use one standard DTD, or set of DTDs, to represent all journals and their constituent articles. The International Organisation for Standardisation (ISO) 12083 DTDs of 1994 were an attempt at such a standard. Despite being promoted by several publishing organisations, ISO 12083 has not been adopted as widely as had been hoped. (This is further examined in Section 3.2.) In fact, current journal DTDs, be they header or full-article, vary significantly between publishers.

'Standard' header DTDs in the STM domain include the MAJOUR (Modular Application for JOURnals) DTD [12] and the more recently developed SSSH (Simplified SGML for Serial Headers) DTD [13]. The ISO 12083 DTD set incorporates tags for header information as well as full article text. The only other freely available DTD for full STM articles is the Elsevier Science Article DTD, mentioned in Section 2. The following sections describe these DTDs; the projects introduced in Section 2 are used to illustrate the issues arising from their use.

3.1 Header DTDs and Their Application

The MAJOUR (Modular Application for JOURnals) Header DTD of 1991 [12] was designed by the European Workgroup for SGML (EWS) primarily as an exchange format for bibliographic information for scientific articles. Springer Verlag, for example, uses MAJOUR and Blackwell Science uses a modified version of it. Several years after the publication of the MAJOUR DTD, the OASIS data set [14] was initiated by a group of publishers. This is a minimum set of SGML tags for journal headers and is an attempt to overcome some of the drawbacks of MAJOUR. In early 1996, the OASIS recommendations were implemented as a DTD: the Simplified SGML for Serial Headers (SSSH) DTD [13]. SSSH was developed by PIRA International [15] for Book Industry Communications (BIC).

It was initially hoped that SSSH would be approved as MAJOUR version 2, but this has proved impossible to date, due to some of the changes from MAJOUR. However, SSSH has been designed to permit simple migration from MAJOUR to SSSH. The most important changes between MAJOUR and SSSH have included making more of the elements optional, reducing the levels of hierarchy, including additional article types, and allowing alternative

models for authors' affiliations. Version 1 of SSSH introduced an element for the Serial Item and Contribution Identifier (SICI) code [16]. However, the area of article identifiers is currently in a state of flux and it appears that multiple identifiers will need to be supported for the foreseeable future. Version 2 of SSSH, therefore, allows identifiers using notations other than SICI.

The Infobike/Journals Online Project

The *Infobike* project is a collaboration between BIDS, ICL, several publishers and university libraries and the University of Kent. UKC provides advice on SGML and electronic publishing, as well as performing technical audit and evaluation.

The aim of the Infobike project is to provide electronic ordering and delivery of journal articles. Infobike has recently been combined with a similar project, JournalsOnline. Through the UK National Site Licence (HEFCE) agreement [17], JournalsOnline will provide access to Blackwell Science, Blackwell Publishers and Academic Press journals. The articles will initially be made available in Adobe Portable Document Format (PDF).

The bibliographic information for each article is provided to the Infobike/JournalsOnline system via SGML-tagged files, the components of which are then stored in ICL's ODB II object database. This database permits fielded searching of the articles' bibliographic details. It was decided that the Infobike/JournalsOnline project should support delivery of bibliographic information via any of the following journal header DTDs:

- Those of Academic Press and Blackwell Science (because they were initial partners in the project)
- The SSSH DTD
- The MAJOUR DTD
- A DTD that conforms to the 'Infobike DTD'. (This DTD defines the minimum bibliographic details required by the system.)

Use of alternative DTDs would require further conversion routines to be developed and therefore increase potential costs for the publisher involved.

Both MAJOUR and SSSH are supported because, although MAJOUR is better established, SSSH incorporates improvements to the MAJOUR DTD, as explained in Section 3.1. In addition, both DTDs are freely and publicly available. The Elsevier Science Article DTD was also considered, as it was being used by a publisher considering joining the project. However, a decision was made not to support it, as it does not conform to Infobike requirements. Although the Elsevier DTD includes article-level bibliographic information, such as article author and title, it does not include journal-level information, such as journal title, ISSN, publisher information and so on.

The SuperJournal Project

The aim of the *SuperJournal* project is to develop e-journals in order to investigate the requirements of readers, authors and libraries, and to consider the implications of processing large volumes of journal articles [18]. *SuperJournal* currently accepts data from publishers in the following formats: an SGML header with the full article in PDF or the header and full article in SGML.

To deal with header information, a 'generic' DTD has been developed. This DTD will accept header data marked up according to the DTDs used by any of the twenty society, university press and commercial publishers participating in the project. A routine has been defined to

down-convert from this generic DTD to a '*SuperJournal* header DTD'. The components of the latter are stored in a database, as in the *Infobike/JournalsOnline* project. According to the Technical Project Manager, Ross MacIntyre, the definition of the generic header DTD was possible because of the 'inherent similarity' of the content of the SGML header files. He comments that 'The DTDs looked very different but what they describe is probably 80 percent the same' [19].

After discussions with UKC, the generic DTD was extended to accept input from the SSSH DTD; although SSSH was not in use by any of the participating publishers at the time, it was recognised as a potential future standard. It has since been adopted by several publishers in the project. BH Blackwells also decided to recommend SSSH for use in their *Electronic Journal Navigator*. Details of how *Superjournal* are dealing with full article SGML can be found in Section 3.2.

3.2 Full Article DTDs and Their Application

An increasing number of publishers are considering the use of SGML for the full text of articles. In the long term, journal articles may be stored in SGML and delivered either in a second format, possibly HTML, or in SGML itself using an SGML viewer. Whether the option of using an SGML viewer becomes popular will depend on how SGML on the Web evolves, in particular, how viewing software develops. As described in Section 3.2, the Illinois DLI project has highlighted problems with SoftQuad's SGML viewer, Panorama [20]. An earlier SGML-based viewer was OCLC's Guidon [21], designed specifically for e-journals. This used SGML files to create its indexes and to generate TEX DVI files for display. Guidon was developed before recent Internet advances. It has not been able to compete with the popularity of HTML-based Web browsers and OCLC announced its withdrawal in 1996. The new OCLC system uses standard Web browsers to display HTML files converted from the SGML [22].

The World Wide Web Consortium have recently been promoting a simple dialect of SGML called Extensible Markup Language, or XML [23]. The goal of XML is 'to enable generic SGML to be served, received, and processed on the Web in the way that is now possible with HTML.' XML browsers are beginning to appear and may well be suitable platforms for future display of e-journals.

ISO 12083 DTDs

There is currently less choice of standard full article DTDs than of those describing header information only; ISO 12083 and the Elsevier Science Article DTD are the only widely used and publicly available examples at present in the STM field. Until recently, publishers were not particularly interested in pursuing standards in this area as they could not envisage a reason for exchanging full text. Discussions concerning the development of MAJOUR into a full-text DTD, for example, floundered because of this; such views are slowly changing.

The ISO 12083 DTDs of 1994 are based on the earlier ANSI/NISO Z39.59 DTD of 1988, originally developed under the auspices of the Association of American Publishers. The standard comprises DTDs for serials and their constituent articles and for mathematics and books. The maths DTD can be embedded in the article DTD which, in turn, can be embedded in the serial DTD.

ISO 12083 was used* in the Illinois DLI project (See Section 3.2.). It was also used in the production of the American Institute of Physics (AIP) journal, *Applied Physics Letters Online* (APLO). This journal was originally made available using the OCLC Guidon system, as illustrated in Figure 1. According to Tim Ingoldsby of AIP, writing in 1994 [24], ISO 12083

was being 'embraced by the international physics publishing community'. But other comments from the physics community were not quite so positive. For example, Prof Franck Laloe, chairman of the publications committee of the European Physical Society, commented in 1996 that 'ISO 12083 is not yet a good DTD, in particular for maths, which would explain why no one uses this DTD' [25]. Since Guidon has been discontinued, APLO is now available on the Web via the new Online Journal Service of the AIP which provides HTML tables of contents and PDF articles.

The University of Michigan Press also experimented with ISO 12083. Until early 1997, SGML test versions of articles from JEP, the *Journal of Electronic Publishing* [26], were made available on the Web for viewing using Panorama [20]. The articles have since been withdrawn but another of the Press's publications, the *Bryn Mawr Reviews*, is now in the process of migrating to SGML. It has been decided, however, that ISO 12083 'would not be adequate' for this journal [27].

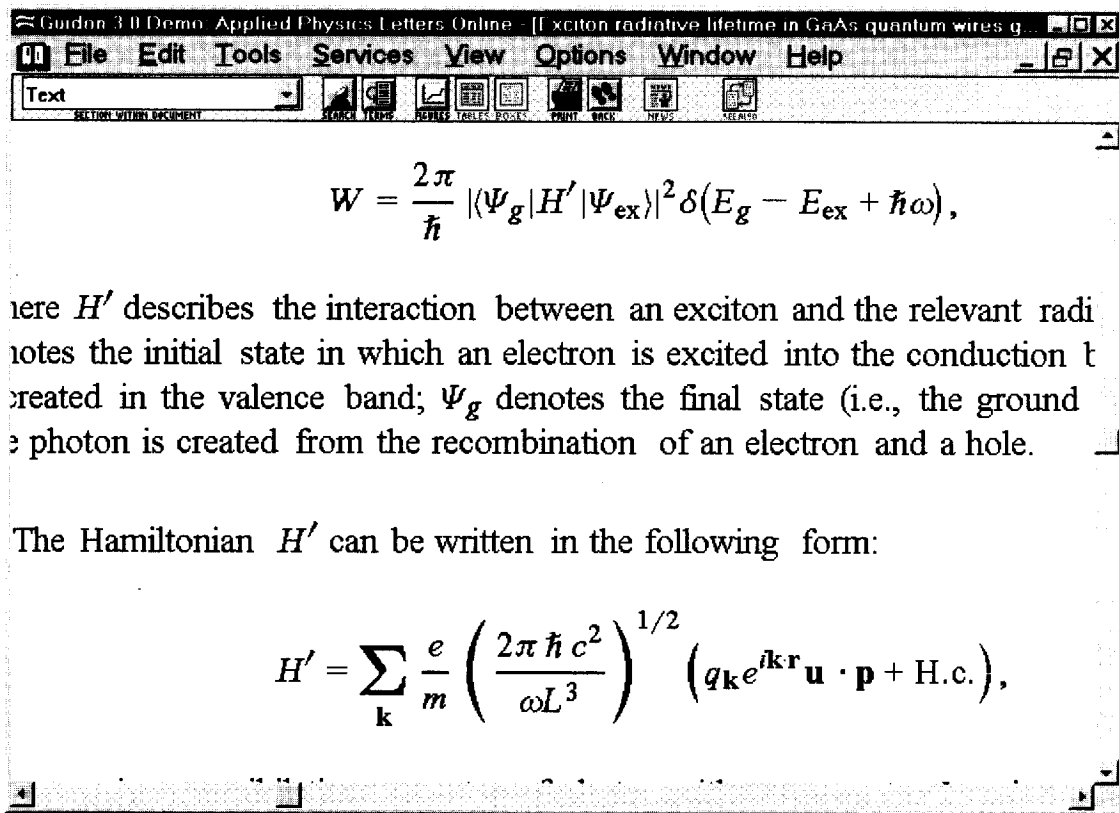


Figure 1. An article from *Applied Physics Letters Online* viewed using Guidon

Several publishers have developed their own proprietary DTD using ISO 12083 as a base. Others have used the MAJOUR header DTD as a base. But publishers invariably make modifications to whichever DTD is chosen in order to customise it to local needs.

There is criticism of all aspects of ISO 12083; some feel that it was foolish for it to have been published as a standard rather than being allowed to evolve through use. With hindsight, the calls for the ISO 12083 standard to be adopted could be interpreted as a recognition of the importance of the use of standards in general, rather than an endorsement of this standard in particular.

The Illinois DLI Project

The four-year, four million dollar University of Illinois at Urbana-Champaign DLI project ('Federating Repositories of Scientific Literature') [7] was begun in September 1994. It involves the rendering and display of SGML using SoftQuad's Panorama with Netscape Navigator. The Open Text search engine is used as a retrieval engine for the SGML documents. Sample articles are publicly available for viewing via Panorama [28]; an example is shown in Figure 2.

Some of the problems encountered in the first two years of the Illinois DLI project are follows:

- The labour-intensive nature of SGML coding.
- The non-session based nature of the Web which makes complex searching difficult.
- The fact that indexes of SGML full-text articles tend to be larger and more complex than their HTML equivalent.
- The shortcomings of Panorama, in particular, problems in presenting maths, equations and diacritics.
- The significant differences between the DTDs used by different publishers; this has a 'dramatic effect on indexing and retrieval'.

Various proposals were made to overcome these problems, including

- The development of a standard for document meta-data. (This is currently an area of frenetic research activity with developments such as the Dublin Core Element Set [29].)
- Facilitation of the adoption of ISO 12083 for journal articles.
- More standard methods of coding mathematics to be adopted among publishers.

Very few publishers actually represent mathematics in SGML. Many maths journals are typeset using TEX. Within the final SGML or HTML article, the equations are often represented as bitmaps or simply in the original unprocessed TEX markup. There was disappointment when the proposed MATH tag of HTML 3.0 was abandoned in HTML 3.2; this had been regarded by some as a possible solution to the problem of displaying maths in HTML. After over a year of discussion, the MATH tag has now reappeared; the W3C HTML MATH Working Group have agreed on a core proposal for HTML maths and the final proposal is due in May 1997.

Guidon, based on processed TEX files, handled both display and inline maths adequately, as may be seen in Figure 1, although, even here, errors sometimes occurred if a 'stray composition code or unusual character string' was happened upon [30].

As mentioned in Section 3.2.1, the ISO 12083 Maths DTD has come in for criticism. It has been suggested, in some quarters, that its best chance of making an impact in the short term is for it to be adapted to automatically accept formulae written in TEX.

In the longer term, there is talk of a 'semantic' maths DTD to take the place of the current generation of 'presentational' DTDs that simply describe the visual structure of maths. But publishers of maths-intensive journals need a solution now. The semantic DTD may take a long time to arrive - if it ever does; a semantic DTD is, for some, 'as impossible for maths as it is for natural language' [31].

As a result of the initial findings of the Illinois project, a publisher consortium is to fund the development, by SoftQuad, of enhanced equation display in Panorama.

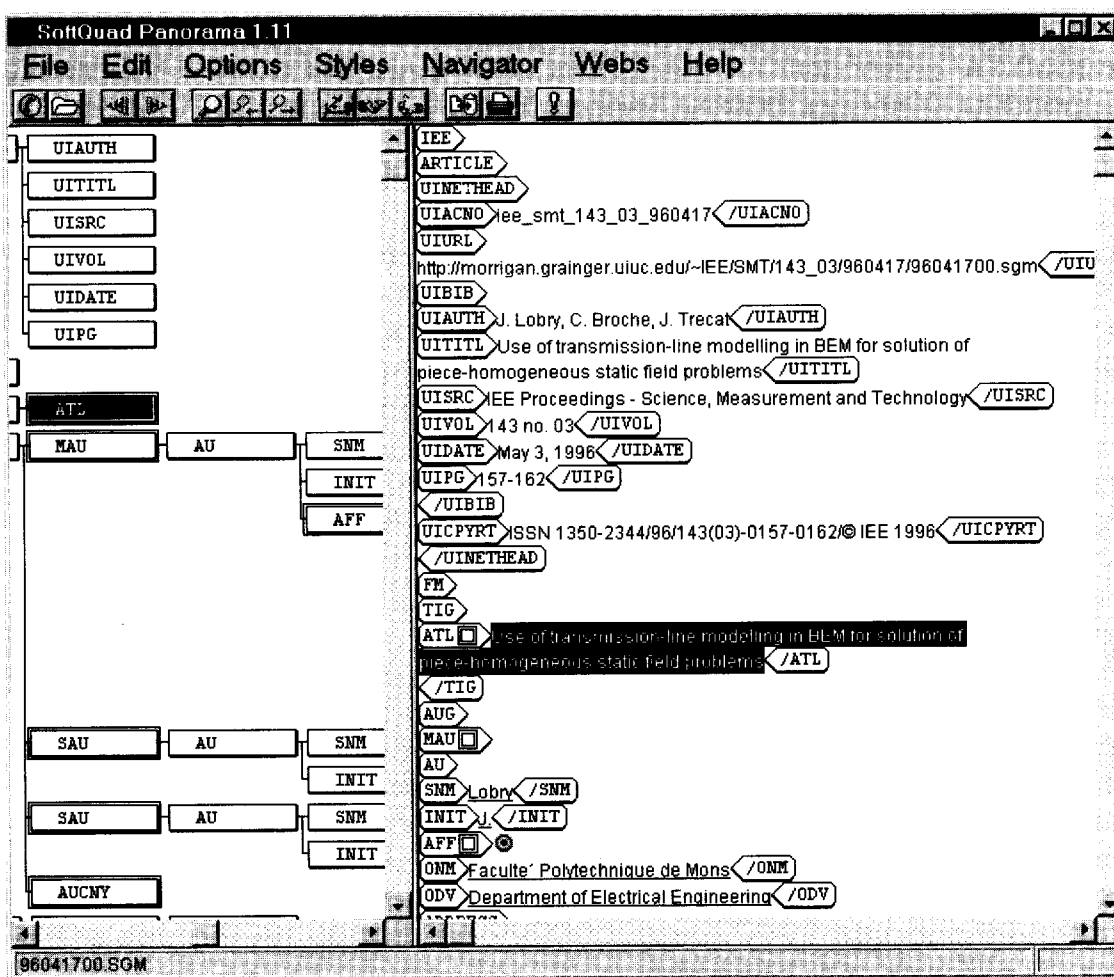


Figure 2. Sample article from the Illinois DLI Project

CLIC

The objective of the CLIC consortium is to 'introduce and set standards for fundamentally new ways of communicating molecular science' [32]. The project involves the creation of an electronic version of the journal *Chemical Communications*, to be delivered via the Web.

The eventual aim is to enable on-the-fly conversion to HTML from an SGML database. However, because of the 'inadequate specification' of the DTD used to mark up the journal articles, this conversion still requires significant human intervention and thus a parallel HTML archive.

The DTD being used to mark up the articles is a modified version of ISO 12083, referred to as the RSC (Royal Society of Chemistry) DTD. Here too, ISO 12083 has been found to be inadequate in some aspects. In fact, Henry Rzepa of the CLIC project has commented 'I cannot bring myself to believe that a monolithic DTD such as perhaps 12083 will ever serve a specialist community such as chemistry particularly well' [33]. The current phase of the CLIC project includes attempts to add the required functionality in order to 'produce a chemistry-specific, semantically rich SGML Document Type Definition for the RSC's journals, starting from the standard ISO-12083 DTD and developing along the lines of Peter Murray-Rust's Chemical Markup Language' [34].

Members of the CLIC consortium are involved in the development of Chemical Markup Language (CML) a 'generic tool for management of molecular and technical information, especially geared to Inter- and Intra-net use'. CML is an application of SGML, but an implementation in Java has recently been developed.

SuperJournal

Unsurprisingly, dealing with multiple full article DTDs in SuperJournal has proved to be more complex than simply dealing with the SGML headers. The wide variation in DTDs means that a 'generic' DTD, as developed for the header information, is not feasible. A full text *Superjournal* DTD has been defined; all SGML files will be down-converted to conform to this DTD and separate article conversion modules are currently being defined for the different DTDs involved. This will mean that only one set of filters is required for conversion to subsequent formats such as HTML. Considerable pre-processing and re-formatting may be required to allow conversion to the target SuperJournal DTD [19].

4. Future proofing

With the arrival of SGML in the 1980s, the myth arose that simply coding information in SGML meant that the information was future-proofed. Now that this myth is being debunked, a new assumption is emerging: that the use of standard DTDs future-proofs information. But this is not true either; the use of information-poor or badly-structured DTDs can result in loss of information.

DTDs that are acceptable in some contexts can be information poor when used in others. The Elsevier Science Article DTD, mentioned in Section 3.1, for example, has caused problems for projects such as ELSA and Infobike/JournalsOnline because it covers only article-based information and not journal-wide information.

An illustration of poor DTD structuring is to use SGML to describe the physical appearance of documents rather than their logical structure. For example, many journal DTDs take the following approach to defining emphasis:

- `<em1>` defines bold
- `<em2>` defines italics
- `<em3>` defines underline, and so on

The assumption appears to be that, because the physical descriptions bold, italics and underline are not actually used as SGML tags, the resulting DTD describes logical structure and not physical appearance; this is obviously not the case. Unless logical objects, such as components of bibliographies - for example, author name, title and so on - are identified as such and not simply as components to be underlined or presented in bold type, some of the potential advantages of SGML are ignored. These advantages include the ability to perform fielded searches, to easily convert from one DTD to another and so on. Now that DSSSL (Document Style Semantics Specification Language) [35] has been accepted as a standard for SGML and XML style sheets, and as it becomes more widely adopted in practice, it should be easier to design SGML systems in a way that dissociates style from logical structure.

5. Conclusions

'For both authors and publishers, it would be advantageous to agree on one standard DTD for the encoding of journal papers.' This quote is taken from a paper entitled 'Standard DTDs and Scientific Publishing', written in 1992. It is interesting to note that the first author works for Elsevier, which did not adopt the ISO standard DTD, but developed its own. The application of SGML helps to future-proof e-journals, but its use alone will not achieve future-proofing. The use of standard DTDs is vital, but these standards must be well-structured and appropriately designed if they are to be adopted.

References

- [1] JISC Electronic Libraries Programme (eLib), Owner/maintainer: John Kirriemuir. Available from: <http://ukoln.bath.ac.uk/elib/> [March 17 1997]
- [2] Pullinger, D. and Baldwin, C., *Superjournal: A Project in the UK to Develop Multimedia Journals*, D-Lib Magazine, Project Briefings and Updates, January 1996. Available at URL <http://www.dlib.org/dlib/january96/briefings/01super.html>
- [3] InfoBike project, Owner/maintainer: Bath Information Data Services (BIDS). Available at: <http://www.bids.ac.uk/elib/infobike/homepage.html>. BIDS JournalsOnline Service demo available at: <http://www.journalsonline.bids.ac.uk/JournalsOnline> [February 26 1997]
- [4] CLIC Consortium Electronic Journal Project, Owner/maintainer: Henry Rzepa. Available from <http://www.ch.ic.ac.uk/clic/> [February 26 1997], JISC Electronic Libraries Programme
- [5] Rusbridge, C. 18 Jul 1996, Re: HTML standards. Mailed to lis-elib@mailbase.ac.uk (eLib list),
- [6] Blackwell's Electronic Journal Navigator, Owner/maintainer: Blackwell's. Available from: <http://mktdev1.blackwell.co.uk/es/ejn.htm> [March 17 1997]
- [7] Digital Libraries Initiative (DLI) project at the University of Illinois at Urbana-Champaign, Owner/maintainer: University of Illinois at Urbana-Champaign. Available from <http://dli.granger.uiuc.edu/> [February 26 1997]
- [8] WebDOC: Access to electronic documents, Owner/maintainer: PICA. Available from: <http://www.pica.nl/docs/en/webdoc/webproj.html>. [March 19 1997]
- [9] ELSA : Electronic Library SGML Applications, Owner/maintainer: Elsevier Science. Available from: <http://www.elsevier.nl/homepage/about/resproj/elsa.html> [March 19 1997]
- [10] Web links to Telematics for Libraries projects, Owner/maintainer: DG XIII-E-4 Telematics for Libraries, Ian M. Pigott. Available from: <http://www2.echo.lu/libraries/en/lib-link.html> [March 19 1997]
- [11] Private communication from David Houghton of International Institute for Electronic Library Research, Division of Learning Development, De Montfort University, March 20 97

- [12] Majour Header DTD, Owner/maintainer: Springer Verlag. Available from: <http://tick.ntp.springer.de/author/sgml/help-sgml.html> [February 26 1997]
- [13] Simplified SGML for Serial Headers (SSSH), Owner/maintainer: PIRA. Available from: <http://www.pira.co.uk/sssh/> [February 26 1997].
- [14] The OASIS data set, Available from: <gopher://gopher.sagepub.co.uk:70/00/OASIS> [March 1997]
- [15] PIRA International, Owner/maintainer: PIRA. Available from: <http://www.pira.co.uk/> [March 21 1997]
- [16] Serial Item and Contribution Identifier Standard ANSI/NISO Z39.56-1996 Version 2. Owner/maintainer: SunSITE Manager, Berkeley. Available from: <http://sunsite.Berkeley.EDU/SICI/> [March 21 1997]
- [17] UK National Site Licence (HEFCE) agreement, HEFCE Council Briefing, July 1996 No. 3, Owner/maintainer: HEFCE. Available from: <http://back.niss.ac.uk/education/hefce/pub96/cb396.html#briefing> [March 21 1997]
- [18] Superjournal Project, Owner/maintainer Ann Apps, SuperJournal Project, Manchester Computing. Available from: <http://www.superjournal.ac.uk/sj/> [March 19 1997]
- [19] MacIntyre, R., *Superjournal* Technical Project Manager, Personal communication, 19 Feb 1997
- [20] SoftQuad Panorama Viewer, Owner/maintainer: SoftQuad. Available from: <http://www.sq.com/panorama/viewer/> [February 26 1997]
- [21] Applied Physics Letters Online Now Available, Owner/maintainer: OCLC. Available from: <http://www.oclc.org/oclc/press/950120.htm> [March 21 1997]
- [22] OCLC Electronic Journals Online, Owner/maintainer: OCLC. Available from: <http://www.oclc.org/oclc/menu/ejo.htm> [March 21 1997]
- [23] Extensible Markup Language (XML), W3C Working Draft, Owner/maintainer: Bray, T. (Textuality), Sperberg-McQueen, C.M. (University of Illinois at Chicago). Available from: <http://www.textuality.com/sgml-erb/WD-xml.html> [February 26 1997]
- [24] Ingoldsby, T., AIP'S Applied Physics Letters Online: Coming In January, Owner/maintainer: American Institute of Physics. Available at: <http://www.aip.org/epub/ingold.html> [March 20 1997], Reprinted with permission from *Computers in Physics*, Vol.8, No. 4, pp. 398-401, July/August 1994.
- [25] Digest of an email discussion between Franck Laloe and Eric van Herwijnen. Posted on the Web in April 1996., Available at ISO SC4/WG6 Electronic Publishing - Workitem Mathematics at <http://mish161.cern.ch/sc4wg6/math/math.htm>

[26] The University of Michigan Press, Owner/maintainer: University of Michigan Press. Available from: <http://www.press.umich.edu/> [February 28 1997],

[27] Price-Wilkin,J. University of Michigan Press, Personal communication,14 March 1997

[28] Digital Libraries Initiative project at the University of Illinois at Urbana -Champaign, Off-Campus Demonstration, Owner/maintainer: University of Illinois at Urbana-Champaign. Available from: <http://morrigan.grainger.uiuc.edu/ieedemo/> [March 20 1997]

[29]Some Metadata Resources, Owner/maintainer: John Kirriemuir. Available from: <http://ukoln.bath.ac.uk/ukoln/other/metares.html> [March 22 1997],

[30]Walter,M., AIP Case Study: Physics Society Takes Its Journal Online, Seybold Report on Publishing Systems, Vol 25 No 8, December 1995,

[31]Poppelier,N.A.F.M. and van Herwijnen,E. and Rowley,C.A., Standard DTDs and Scientific Publishing, Available from: <http://www.elsevier.nl/homepage/about/sgml/>. [March 19 1997], 1992

[32] A Flagship Chemistry Electronic Journal: The CLIC Consortium (Press Release), Owner/maintainer: Henry Rzepa. Available from <http://www.ch.ic.ac.uk/clic/Flyer.html> [March 18 1997],

[33] Rzepa,H., 4 November 1995, Re: ISO 12083 vs TEI. Newsgroup: Mailed to liselib@mailbase.ac.uk (eLib list)

[34] Chemistry Markup Language, Owner/maintainer: Peter Murray-Rust. Available from: <http://www.venus.co.uk/omf/cml/doc/index.html> [March 21 1997]

[35] Prescod,P., Introduction to DSSSL, Owner/maintainer: Prescod,P., March, 1997. Available from: <http://itrc.uwaterloo.ca/papresco/dsssl/tutorial.html> [March 22 1997]

Judith Wusteman has been a lecturer in computer science since 1990. Before that, she completed a PhD in artificial intelligence. Her research interests are in electronic publishing, particularly electronic journals and SGML, and in the social implications and applications of computing. Judith directed operations at UKC for the audit of the eLib Infobike project until September 1996. She has been involved in the development of SGML DTDs for both e-journals and encyclopedias, and is a committee member of the British Computer Society Special Interest Group on Electronic and Multimedia Publishing.