

SUMMARY OF THE “FUTURE PROSPECTS” and DISCUSSION SESSION Friday 3rd December 1999

J. Poole, CERN, Geneva, Switzerland

Abstract

The session was composed of two parts: a talk by Michel Goossens entitled “Publishing on the Web Today and Tomorrow” and a review of JACoW activities in an open forum including the definition of some plans for the future.

1 WEB PUBLISHING

1.1 *Using HTML*

Michel Goossens explained that HTML has many problems like browser vendor specific extensions which can be invalid or no-longer supported and because the basic grammar of the language is fixed there are no ways to adapt or extend it. These are typical problems with fast developing technologies before adequate standards are established.

It can be seen from present use of HTML that meta-data is playing an increasingly important role in web documents but there is only limited support for it. Finally, although the WWW passes across international barriers, there are still difficulties to exchange data because there are language-specific aspects which are not solved. The HTML language has evolved since its introduction but HTML 4 has been frozen and subsequently abandoned. The next development in the pipeline is XHTML which is a modular application of HTML using XML. The future of Web publishing is therefore within the XML (Extensible Markup Language) domain.

1.2 *XML*

XML is a kind of self describing system which is based on SGML and it is easy to learn. It has the great advantage that it can be adapted to the users requirements like using native languages etc. The basic language of XML is Unicode and therefore internationalisation is built in from the beginning. Since 1996 an international committee has been looking at the development of a system for defining, validating and sharing document formats on the web. This committee came up with the recommendation for the XML 1.0 standard in February 1998. Since that time a number of further standards have been introduced (XSL, XPointer, XLink, MathML) and work continues on a large number of others.

All of the major players in the internet have adopted the XML standard and although Netscape and Explorer cannot interpret XML yet, development has started. The logical relations between the elementary parts of a document are described in the Document Type Definition (DTD) and

various scientific and commercial communities are developing “standard” DTD’s. Examples of such developments are as diverse as Chemical Markup Language (CML), Music Markup Language (MusicML) and Weather Observation Markup Format (OMF). There are also of course, many E-commerce applications being developed.

The use of Cascading Style Sheets (CSS) will increase since they allow the decoupling of presentation and content. The use of Object Technology via the Document Object Model definition in DTD’s indicates a more rigorous approach to Web publication.

1.3 *Impact for Our Community*

Clearly the Web publication of scientific data will become much easier. Goossens estimates that soon HTML-only browsers will be a thing of the past and most Internet tools will support XML natively. A key feature will be the ability to develop our own language to suit our specific purposes. Furthermore XML documents will be portable between systems (Web, WORD, \LaTeX etc.).

However, XML is still in its infancy and it will be some years before it will be widely used in the accelerator community. At CERN the portability between XML, \LaTeX and FrameMaker/SGML will be investigated during the coming year. The next stage in this development will be to include WORD, Wordperfect and so on. In Europe, the Esprit project known as *TIPS* (Tools for Innovative Publishing in Science) will look at possible solutions.

In the long term one can imagine that tools like WORD and FrameMaker will allow authors to store their documents in XML as can be done from \TeX now. In this environment one can imagine that the process of publishing conference proceedings on the Web will become much easier. The part of the publishing process devoted to making the indices and wrapper will however, not really benefit very much from these developments.

2 JACOW DEVELOPMENTS

There was an open discussion about what the delegates thought JACoW should be doing in the future and how we could achieve this.

The first area which was debated was the software and procedures that could be shared amongst the various conferences. The most recent conference was PAC’99 and a large development effort went into the Web interfaces for abstract and paper submission and for the programme committee’s input. This work was done by BNL and LANL and

they agreed to make the source code available for anyone wishing to use it.

It was also decided that an E-mailing list for conferences was a useful object and it was agreed that EPAC's list could be updated with information from other conferences. The updated list will be available for any one wishing to use it.

The JACoW site contains keywords and it was agreed that the list should be updated according to the experience at each conference. A number of acronyms which do not appear on the list need to be added.

It was also agreed that the basic instructions and help which are applicable to all conferences should move to the JACoW site and that the list of Frequently Asked Questions (FAQ) which was compiled for PAC should be added.

Other software which has already been shared are the scripts and programmes which have been developed to automate the publishing process (Web page generation from database, page numbering and so on). These scripts are passed from conference to conference and are updated to follow systems evolution.

3 FURTHER WORK

A number of areas where further work was desirable were also identified. For example, the database structures and schema should, in principle, be useful to any conference. At present they are Oracle specific, but the preparation of some meta data describing the structures would be of benefit.

The Acrobat software continues to be improved and it delivers greater performance for each conference. The parameters for the distiller are not well understood and it would be useful to really investigate the possibilities and obtain an optimum set.

Enfocus PitStop software has also developed rapidly and the full potential for conference paper processing has not been fully exploited. This should be investigated with a view to using it to do some of the checking of submitted papers.

A generic postscript driver geared to the needs of publication on A4 wide, US letter high paper would be a large advantage. It was agreed that this possibility would be followed up.

It has been seen that training of personnel in the processing office has not been done very well so far. The procedures may be mature enough now that some better defined training programme can be established. It was agreed that this topic should be reviewed after EPAC 2000.

There were some problems for which no solution could be identified and these remain open questions. The most important of these being how to deal with special characters in names and text across the various media (Web/E-mail and FTP abstract submission, papers via FTP and diskette etc.).

The question of how to achieve a similar look and feel across the different CD's and Web presentations was raised. After lengthy discussions it was agreed that it is more im-

portant for the conferences to maintain their personal look and feel than to try and find some common ground in this respect.

The JACoW team plays an important role in the publication of all conference's proceedings because of the close collaborations and exchanges between the conferences. The field of electronic publication is developing fast and it was agreed that it is necessary to review the workings of JACoW at least once per year.

4 CONCLUSIONS

XML is a long term development which will revolutionise Web publishing. However, it is not something which will affect JACoW in the medium term (5-6 years). The JACoW collaboration has been a great success and the website is greatly appreciated by the accelerator community. The collaboration between the conference series has proven to be very effective and will continue to expand as more conferences join JACoW. The number of paper volumes is in steep decline and it was agreed that it would not be long before they were no longer needed and it was agreed that JACoW encourages this trend. The suggestion that the CD-ROM was no longer required was rejected.

It was agreed that it would probably not be necessary to repeat the workshop for a few years but that a meeting of the team should be held at least once per year. In the team meeting there should be sufficient time to study and evaluate problems and the timing should be such that it does not clash with the peak activities in any of the conferences. It was also agreed that the meetings held at the conferences (only EPAC and PAC to date) provide an important channel for feedback to the organising committees and that they should be maintained. The delegates expressed their satisfaction with the workshop which they felt had allowed the knowledge and experience from nearly five years of electronic publication to be successfully shared.