

## Preparing for a Phase Change

**Franck Laloë** discusses why the EPS workshop *Electronic Publishing in Physics* (Paris, 2-3 March 1995) focussed on the ways the physics community can prepare itself for the electronic age.

Most agree that 1995 will be a year of phase change for physics publications because, after a long period of latency, several important journals will come on-line. However, when one discusses electronic publishing, the point of view of physicists, that is of authors and readers, may differ from those of bodies such as publishing houses and libraries. Most physicists nowadays use computers at work, for communicating, and in writing articles, and therefore directly produce electronic documents which in principle can be used for a variety of purposes. But we have not yet reached the situation where the inherent flexibility of electronic documents is satisfactorily exploited. For instance, much re-keying is still done, often for minor reasons such as to cope with the different layout styles used by journals. This is especially so in Europe which publishes a large variety of physics journals.

As authors we want our articles, once typed and ready, to be sent to: (1) colleagues; (2) preprint databases and servers — including local servers now being set up by some institutes for their staff's publications; (3) journals to be printed and eventually placed on a CD-ROM for storage or in a server delivering electronic journals. Information in an article may be extracted (4) and entered into a non-bibliographic databases using key words, abstracts, tables of numbers, etc. This is less frequent in theoretical physics where preprint databases seem to be sufficient. But elaborate databases are used constantly, and often at great cost, in more technical and applied areas, in astrophysics, and in disciplines such as chemistry, biology and pharmacy where there exists much numerical and quantitative data.

### Standardised Structure Needed

Many of the physicists' needs in (1) and (2) are satisfied by the almost universal use

**Franck Laloë** (laloë@physique.ens.fr) chairs the EPS Action Committee for Publications. A theoretical physicist in condensed matter, he is a CNRS Directeur de Recherche in the Physics Department, Ecole Normale Supérieure, 24, rue Lhomond, F-75231 Paris. He received his Ph.D. in 1970 while working with the CNRS, which he joined in 1968, and has worked as a visiting scientist at Imperial College, London, Sussex University (UK) and MIT (USA). Formerly a Co-editor of *Europhysics Letters*, Professor Laloë is a currently the Vice-President of France's Conseil Supérieur des Bibliothèques and member of the Co-ordination Committee of the Haut Conseil de l'Information Scientifique et Technique.



*E. van Herwijnen of NICE Technologies speaking at the EPS Workshop on Electronic Publishing in Physics. He headed the EPS group whose work led to the ISO 12083 maths and physics DTD.*

of  $T_{E}X$  or  $L_{A}T_{E}X$ . The situation is less clear for (3) and (4), and we are far from an ideal world where, after the initial typing of the manuscript by the author, everything would be done automatically. The major reason is that the initial manuscript does not include sufficient structure. If we, authors, put in the necessary structure at the beginning we will be able to profit more from our manuscripts and to exert a better control on dissemination of information. We have to move away from the persuasive WYSIWIG paradigm of "what you type is what you get". Our job is not to produce the best typography in the world! It is much more important for us to ensure flexibility for multiple presentation styles. Most of us also wish to store manuscripts in a format that does not change every year, yet contains enough structure to be converted automatically into, for instance, the different generations of the Hyper-Text Markup Language (HTML) used over the Internet's World-Wide Web. Indeed, a strong argument for adopting SGML is that texts could be automatically converted into the various versions of HTML.

$T_{E}X$  word processors (more correctly, pre-processors) or some future sort of  $T_{E}X$  will probably remain essential as their remarkable typographic qualities make them the best candidates for operations such as printing. Physicists also have large reservoirs of  $T_{E}X$  documents which they wish to continue to use.  $T_{E}X$ , moreover, can evolve towards more structure, and indeed  $L_{A}T_{E}X$  already contains an elementary form of structure ( $L_{A}T_{E}X^3$ , if it comes, will have more).

For instance,  $L_{A}T_{E}X$  includes the notion of "authors" so that it can easily be used to automatically extract all articles by Einstein,

while excluding those that mention his name in the main text. But we need more structure, such as the notion of affiliation (e.g., "on leave from") if we wish to feed and access databases more efficiently and to avoid excessive dependence on commercial companies offering proprietary solutions.

There are at present two main candidates for standardising the necessary structure of physics articles. The first is to tag documents using the Standard Generalised Markup Languages combined with the ISO 12083 maths and physics Document Type Definition that has been approved by both EPS and the American Physical Society — an agreement that owes much to the past activity of SGML specialists within the EPS Action Committee for Publications. The other is based on standardising the  $T_{E}X$  or  $L_{A}T_{E}X$  macros commands and/or other methods to define a detailed structure.

The former approach makes directly a big step forward. Its great advantage is that it addresses the problem as a whole; the latter approach is progressive and maybe easier to implement. As always, the market (in the present case, we the authors) will decide. The past years have shown that the pressure for change comes primarily from the authors and not from journals. Publishers cannot impose a standard if it is not widely accepted. It is also fair to say that until now authors have not sought structured documents. But this could change rapidly when the appropriate authoring tools become available. Indeed, several software companies have started to release SGML tools (see insert).

### Undreamt-of Scenarios

Books may not change very much in the electronic era, except for technical details that we all know about (automatic indexing, hypertext links, etc.). Review journals which have an obvious intellectual appeal, will probably serve a useful function using a system that is not too different from what we now have. Much more radical changes are in store for most other physics journals. There are two extremes, namely the "paral-



*J.-F. Abramatic (on the right) from INRIA who heads the European World-Wide Web WEBCORE Project talking to H. van der Rijst of Elsevier Science.*

#### SGML AUTHORING TOOLS

ADEPT Publisher	ArborText
Author for Word	Microsoft
Cyberleaf	Interleaf
HotMetal	SoftQuad
Intellitag	WordPerfect
Quaterdeck	
SGML Editor	Grif
Tag Wizzard for Word	NICE Technologies

#### ON-LINE DOCUMENT DELIVERY SYSTEMS

RightPages	AT&T/Springer
GUIDON	OCLC
WEBZ	OCLC
WWW	CERN...
Hyper-G	Graz University

#### PORTABLE DOCUMENT SOFTWARE

Acrobat	Adobe
Common Ground	No Hands Software
Replica	Fallon
Envoy	WordPerfect

lel" and the "perpendicular" scenarios (one could say evolutionary and revolutionary).

In the parallel scenario, the electronic journal only exists in parallel with the usual paper version, as will be the case for *Physics Review Letters* from this July. This has advantages in terms of shortening communication delays (the electronic form generally appears before the paper version), automatic searching for information, etc. The storage problems of libraries call for the extensive use of CD-ROMs for archival, and the access to information by scientists in underdeveloped countries may become easier.

In the perpendicular scenario, the traditional journals disappear, or become so different that their new wave function has a very little overlap with what is was before (hence the term). For instance, an individual laboratory or a group of laboratories may set up servers with their own rules for ensuring scientific quality, safety of archival, etc. Our imagination is in fact the only limit because the nature of present journals does not have to be transposed. Take scientific evaluation which remains an absolute necessity. One can conceive of systems of databases with different levels — those where all articles are accepted and remain forever, those for articles after they have been selected by qualified readers, etc.

We have to be ready to cope with and analyse numerous possibilities. This is the main challenge today, and in addressing it the workshop reviewed the state-of-the art and the various approaches before discussing how, as physicists, we should act in order to be reasonably prepared whatever the outcome. The problem is not only to develop better software for authoring and viewing, but also to address legal issues such as copyright. Perhaps the most important aspects concern efforts to harmonize developments across national borders and scientific disciplines.

## Scientific Publishing without Publishers

Jean Zinn-Justin argued at the EPS workshop *Electronic Publishing in Physics* (Paris, 2-3 March 1995) that electronic documents will allow essential changes to be made in the way physicists publish articles.

It is becoming increasingly obvious that electronic publishing is not paper publishing with electronic means. The combination of electronic production of scientific documents, T<sub>E</sub>X, electronic mail, Internet and now the Web, and the tremendous success of Los Alamos-type preprint database servers, is resulting in a true revolution. The whole publishing process has therefore to be re-examined.

As there are many types of publications playing a variety of different roles, in discussing the implications of electronic publishing it is necessary to focus on certain types. Scientific contributions for which authors expect to be paid, and reviews and books for which the paper form in bound volumes is especially useful, will remain in the near future in the hands of traditional publishers.

In focussing on scientific articles and journals one needs to analyze the role they play in general scientific activity, and to better understand their properties and the new possibilities. Since the electronic revolution will directly affect traditional publishers, it raises a natural question: will the future see scientific publishing without publishers? I shall argue, in particular, that electronic publishing gives us an opportunity to survive the slow collapse of the refereeing system — one of the principle cornerstones of paper publication.

### Crisis in the Referee System

Although evaluation and selection are among the most critical items in publishing an article, publishers also undertake to:

– Gather scientific information, select (with the help of scientists) contributions worth publishing, and organize a discussion between referees and authors, which sometimes helps improve the content. Note that the paper medium imposes an artificial coupling of two different tasks, namely immediate evaluation of a piece of work and a

means for scientists to communicate the results of their research.

- Edit and format articles.
- Disseminate and archive information.
- Help organize information, although often only in minor ways (by creating topical letters, comment sections, indexes, etc.).

A few empirical remarks are appropriate.

### Increased productivity

The number of published pages is increasing exponentially, but alas not the scientific content, or at least not at the same rate. I can only guess the reasons: increase in the number of physicists; increase in the physicists' "productivity" owing to external pressure (the feeling that a healthy number of published papers is essential for a scientific career); improved production tools (e.g., computers).

Referees are no longer able to control the flood. This is obvious to physicists who have been in the editorial business for some time (those who are not have generally stopped reading journals anyway). Referees receive too many papers which are poorly written and highly specialized. Assessing relevance and novelty therefore becomes an almost impossible task. Moreover, the less interesting papers are the more demanding in terms of the referee's time. If a referee becomes too selective the author generally fights back, asks for another referee, then goes to another journal. The author is really unlucky if a paper is not published in the end, and several referees eventually discover that their time has been totally wasted. Finally, if existing journals become too selective, new journals are created. Moreover, in a world in which the success of a journal is measured by the number of published pages, it cannot be expected that publishers themselves will try to discourage authors.

### Shortcomings

Paper publication has two serious shortcomings. While a conscientious referee may have written a long and argued report, the final outcome can only be "yes" or "no", even when the report would prefer "maybe". Moreover the answer "no" means depriving, perhaps on subjective grounds, a physicist



Jean Zinn-Justin (zinn@amoco.saclay.cea.fr), seen here speaking at the EPS *Electronic Publishing in Physics* workshop, is a theoretical physicist with CEA Saclay. He has served as the Editor-in-Chief of *J. de Physique* and is currently Honorary Editor of *J. Phys. A*, Associate Editor of *Nucl. Phys. B*, Editor of *Fortschritt der Physik*, and Co-editor of *Proc. Les Houches Summer Schools*.