The IUCr Electronic Journal

BRIAN MCMAHON AND PETER R. STRICKLAND

International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England

1. Executive summary

The March 1999 meeting of the Finance Committee requested the Chester Office to undertake a study into the provision of online versions of the IUCr journals and provide a demonstration version of an IUCr journal for the Glasgow meeting.

Since March, the Chester Editorial Office has been able to develop a pilot online journals service 'Crystallography Journals Online'. Members of the Finance and Executive Committees wishing to view the service should visit the URL **http://journals.iucr.org/**

The Chester Editorial Office recommends the launch of the service at the Glasgow Congress and immediate introduction of the service to print subscribers of the IUCr journals according to the following model.

1. SGML source files and high-resolution graphical images for each published article are transformed to (*a*) fully-navigable HTML and medium-resolution graphics files; (*b*) PDF page images, both accessible to registered subscribers.

2. Subscribers to the journal must register with the service; manual intervention will be required to validate the registration against Munksgaard subscription records. A beneficial corollary of this is that we will gain up-to-date access to our subscription records.

3. The strategy for managing access should be elaborated, with respect to both the technical capabilities of the server software and the desired business model. The recommended model is that a subscriber for the current year should have free access to *all* online issues of the journal; but is barred completely should the subscription lapse. In addition it is recommended that investigations should begin into the provision of pay-per-view access to individual articles.

4. The public service should be made available through either

   (a) a new Chester server (Sun Ultra 10 workstation, or similar), subject to a substantial upgrade of access to the Internet by way of the existing Internet Service Provider (ISP);

   (b) an identical computer hosted by an external ISP.

   Initial costs are similar; we recommend (*b*) to provide redundancy of service and externalise hardware support issues; but this may need to be balanced against possible other arguments in favour of increasing the bandwidth at Chester.

5. Discussions into collaboration with other electronic publishers have been held. It is recommended that such inquiries should continue with a view to: sharing the administrative workload; providing more flexible access mechanisms; improving hyperlinking to other journals and databases; increasing journal visibility.

6. The online journal resource should be extended from current issues to the complete back run of the journals through the following steps:

   (a) development of a database of metadata on all published articles

   (b) archival-quality digitisation of all back issues

   (c) conversion to web-deliverable PDF files of all scanned articles

   (d) registration of journals within the Digital Object Identifier (DOI) or other such systems

7. A synergistic relationship should be developed with the existing pan-crystallographic web information services. It is recommended that the journals service be named 'Crystallography Journals Online' and that the web services be re-badged as 'Crystallography Online' with similar design and livery.

8. For the long-term success of an IUCr hosted service, it will be crucial to provide a critical mass of crystallographic information that is fully interlinked. Inquiries should proceed into collaboration with other publishers of crystallographic journals and databases on linking and possible inclusion of their information in the service (on a commercial basis).

9. The recommended timescale is

   (a) launch of the service from Chester as a pilot at the Glasgow Congress

   (b) enhancement of functionality, transfer to external server (or installation of extra bandwidth at Chester) by 1 January 2000

   (c) restriction of access to subscribers during 2000, but no later than 1 January 2001

10. Finally, the Editorial Office recommends that the introduction of the electronic service should be seen as an opportunity for the Executive Committee to review its policy for providing low-cost access to our journals for scientists in developing countries.

1.1. Financial implications

The bulk of the costs of R&D and editorial development can be borne within existing expenditure subject to the overall workload; but it is possible that a need for more personnel will become apparent as the project matures. Two particular areas outside our normal sphere of operations that may require staff enhancements are: the provision of end-user support; and the administration of access controls on subscribers.

Well defined costs are

1. New server for development or hosting of the online journals service: Sun Ultra 10 440 MHz workstation, 256 MB memory, 18 GB hard disk: $\sim$ £ 8000
2. (a) Bandwidth upgrade (64 kbps $\rightarrow$ 2 Mbps): installation cost £ 8000; extra annual running cost £ 18 000; enhanced router $\sim$ £ 6000
   (b) *or* facilities-managed hosting service on 10 Mbps backbone: setup cost £ 1500; annual running cost in the range (probably at the upper end) £ 11 000–23 000; the cost depends on the particular configuration: the highest-performance option would employ the same hardware as requested for Chester in item (1) above
3. Additional hard-disk storage for back articles $\sim$ £ 8000
4. Registration of DOI: $ 1000
5. Possible registration of other second-level Internet domain names: $ 50 per year per domain
6. Possible costs for commercial secure web server: *unknown*; of order $ 2000?
7. Registration of trademark logos: $\sim$ £ 250 *per decade* per registered mark in the UK

It may also be beneficial to establish an enhanced promotions budget.

Furthermore, as the experience and stature of the publishing operation at Chester continue to grow, the Finance and Executive Committees should begin to consider the merits and implications of moving responsibility for all aspects of journal publishing (subscription fulfillment, warehousing and distribution of the print edition) under the aegis of the Chester office.

1.2. Possible enhancements

As the project matures, and especially if it proves popular, there will be ample scope to increase its functionality. Some possible future developments with likely financial implications include

1. Hyperlinking to other publishers' online services; to citations and abstracts databases; to related scientific databases.
2. E-commerce delivery of individual articles.

1.3. Acta Crystallographica Section C

Acta Crystallographica Section C does not fit completely into the system developed for the other journals. A study will be undertaken to explore how best to integrate *Acta C*.

2. Introduction

There is a general perception among publishers of scientific literature that their publications should be available to readers and subscribers in an electronic form. Direct pressure from readers, acceptance by librarians of the need to handle non-print media, the availability of the corpus of current titles from a number of established publishers, and competition from new journals launched in electronic-only form, have all contributed to the sense that offering electronic versions of the primary scientific literature is not only desirable, but essential to the survival of the publishing enterprise.

It is also clear that the rapid growth of the Internet and the World Wide Web have shifted the focus of 'electronic' publishing away from the provision of a local archival version on a physical medium such as CD-ROM, to the provision of a network-accessible version fully capable of hyperlinking to related publications. Indeed, so rich is the treasury of knowledge already accessible on the Web that there is a growing feeling among users that they should enjoy instant and uninhibited access to all of human knowledge – an aspiration that shows little regard for the real costs of collecting, organising, storing and redistributing such knowledge.

The IUCr is particularly well placed to accommodate many of the requirements and expectations of readers in this electronic era. It owns a stable of exceptional journals; it has pioneered the use of data exchange mechanisms for the archival, transfer and publication of structure reports; it has close ties with the providers of major structural databases. It has a skilled editorial and technical staff; it exercises the highest standards of quality control; and it already manages a rich and powerful series of web information services.

Experience with using third parties to host electronic versions of the IUCr journals has proved disappointing. The Editorial Office in Chester was asked by the Finance Committee to explore the requirements for itself hosting the electronic journals. This report is the outcome of that exploration.

3. Strategy Review

The exercise of reviewing the requirements of an electronic journal has been very instructive, both in inventorying the extensive progress we have already made towards the goal, and in clarifying the benefits of a distinctive and extensive service

which we can manage and promote under its own identity.

Appendix 1 outlines the major design objectives behind the service we are capable of providing, and consider useful to provide. It summarises the principles behind the pilot implementation hosted at the URL **http://journals.iucr.org**. With the exception of access control, all the design objectives have been met or are in hand.

The development of this prototype was achieved in about three months, largely because of the substantial infrastructure already in place at the Chester offices, and as a result of the systematic and thoughtful development of the computer-based workflow built on top of previous investment in equipment and personnel.

The main components of a fully functional electronic journal project are listed in Appendix 2. The listing covers software, hardware and systems elements, and is not exhaustive. Again, many of the necessary components are already in place.

Appendix 3 describes the appearance and operation of the prototype service.

A number of questions arise in considering how to take the service from the prototype stage into full production. These are outlined in the sections that follow.

4. Connectivity

The prototype service is hosted at the Chester office, which is connected to the Internet by a 64 kbps leased-line connection. It is likely that this bandwidth will prove too low when there is high-volume access to the online journal involving a very large number of HTML and GIF images, and a smaller number of larger graphics files and PDF documents.

It is very difficult to assess in advance the actual throughput of network connections. It has been rightly said that 'The Internet is really just a series of bottlenecks joined by high speed networks.' A 2 Mbps connection might allow much smoother traffic flow to users worldwide than a host computer on an ISP's 10 Mbps backbone; or it might not.

The costs of the two alternatives described above are broadly similar. In detail, the figures run as follows.

1. Line upgrade at Chester 64 kbps → 2 Mbps
   - installation fee: £ 8000
   - annual running costs: £ 29 000 (this represents an *increment* of £ 18 000 per annum over the existing connection, which we would need to keep in any case).
   - Possible router upgrade – hardware (Sun Ultra 5): ∼ £ 4500
   - Possible router upgrade – software (Sun PPP3.0): ∼ £ 1400
2. Facilities managed host. We have been offered a deal hosting a Sun Ultra 10 on a 10 Mbps backbone for about £ 23 000 p.a. They provide daily backups and manage hardware downtime. All software would be installed and managed by us.
   - installation fee: £ 1500
   - annual running costs: £ 23 000

For the online journals project, we recommend the second option (of a facilities managed service) for the following reasons:

- Redundancy – there is nothing to stop us running our own server at Chester for (i) our own use, (ii) possible use by Coeditors or other relevant people for whom it might be more convenient, (iii) redirection in the event of prolonged downtime of the facilities managed host. Probably redundancy of international connectivity is not greatly helped in the event of a major disaster at the central routing stations: most ISPs access the international networks through London docklands sites.

- Security – both of hardware through the facilities management, and unauthorised access through separating the server from our own network. Further, replication of access controls at both sites would ease recovery from a compromised system.

- Guaranteed provision of service – an external provider can undertake 168 hours per week uptime, not possible at Chester owing to limited working hours and persistent local network problems.

- Extensibility – similar deals may be available through affiliated companies on other networks world wide. Leasing of identical server hardware under our direct control will simplify mirroring.

- Lower initial commitment – contracts on an annual basis allow the alternative route of upgrading the current connection at a later stage if demonstrably preferable.

A third approach to improve connectivity is to host the service through a third party. We are also exploring this avenue. In terms of costs directly attributable to bandwidth, it may be a much cheaper option; but there are of course other costs against which that will be offset. It is interesting to note that one of the possible third parties we are considering uses the Pipex facilities-managed service for their own equipment.

5. Subscription model

The technical aspects of establishing an open-access service have been addressed and the prototype service satisfies most operational requirements. The remaining layer is the access control layer, which depends in part on technical capabilities, but also on the policy adopted for permitting access. This policy needs to be determined by the Executive and Finance Committees, perhaps in consultation with the Commission on Journals and Electronic Publishing Committee.

The two obvious approaches are access based on (1) current subscriptions; (2) past subscriptions.

5.1. *Access based on current subscription*

This is the model favoured by bodies such as the American Institute of Physics (AIP) and Institute of Physics Publishing (IoPP). Subscribers with a subscription to the print journal valid for the current year may access *all* online issues of the journal. When their subscription expires, they are unable to access *any* issues.

The benefits are that subscribers can follow hyperlinks to other cited articles in the same journal with no restriction, and that a great deal of added value is associated with the initial subscription. The greatest disadvantage is that subscribers (and libraries in particular) view their acquisitions as of archival importance, and are reticent about an arrangement that is closer in nature to leasing than outright purchase. This concern will be less where online access is a free component of the print subscription, but will be much greater if there is a move to charge extra for online access, or to offer electronic-only subscriptions. A strategy to counter these concerns is to provide an archival CD-ROM containing all the online content at the end of each year (or on shorter timescales, depending on the capacity of the storage media). This is the approach taken by AIP. It may be worth considering the routine provision of such an archival CD-ROM to all subscribers as a component of the electronic subscription. If so, the production cost of the CD-ROM must be added to the production costs of the journal. There will also need to be some R&D work to reconcile the different hyperlinking mechanism from CD-ROM and online media.

5.2. *Access based on past subscriptions*

That is, a subscription for a particular year guarantees access in perpetuity to the contents of the journal published in that year, but *no* access to other years (unless a subscription was also taken out for those years).

The benefit is permanent access to the service acquired – the intellectual product is bought rather than leased. The disadvantages are that the subscriber has no access to articles for unsubscribed years, and that the provision of an online service in perpetuity is an undertaking that is difficult to guarantee. For these reasons, this is not the approach that we recommend.

5.3. *Technical aspects of access control*

At present it appears that Apache web server software can manage either approach. In each case, the directory containing a full article and its associated materials contains an access control file (acf) with lists of groups of users permitted to access those files. The contents of the group files are managed by a database application (which we would probably write in-house, although there might be a case for investigating whether suitable off-the-shelf products exist). Group files would be built for 'current subscribers' or '1999 subscribers', depending on the subscription model supported; and there would also be privileged groups such as 'editorial staff', 'journal coeditors' *etc.*

The major difference between the two models outlined above would be the greater administrative overhead of tracking changes of IP address or password details for subscribers represented in multiple (per-annum) group files as opposed to 'current' subscribers.

5.4. *Dynamic subscriptions*

There will probably be some demand from users for subscriptions extending outside the confines of a calendar year (*e.g.* a new subscriber to the journal in November will not want to pay another large fee in January). So long as the electronic product is coupled strongly to the print journal, one might be able to resist such pressure; but offering the electronic service as a separate product will encourage the requirement for fixed-term access, and not necessarily for 12-month periods. This type of access can probably still be managed through the Apache acf mechanism, but would require more sophisticated database programming to track dynamic subscription data.

5.5. *Pay-per-view*

The option for non-subscribers to access individual articles for a discrete charge should also be studied. Evolving e-commerce software makes the implementation of this approach easier to envisage, though we are still without direct experience of what would be involved. The advantage of this system is that it opens a potentially large market of occasional purchasers; the disadvantage of course is that it reduces the pressure to take a full subscription for journals where it is perceived only occasional items of interest will occur. It is recommended that the implications of moving to such a model be considered very carefully. It is also recommended that there should be no extensive technical development of such an approach until the desirability of adopting it should become clearer.

5.6. *Site licensing*

The Chairman of the Electronic Publishing Committee has commented upon seeing a draft of this report that the pay-per-view subscription model is currently very little favoured especially in America. However, there is a substantial momentum at present for site-licensing and other acquisitive agreements between library consortia and large publishers (again largely in but by no means confined to America) involving whole collections of titles. Whether this business model can be sustained into the future is open to question (there are possible difficulties with anti-competition legislation and with smaller libraries who do not require a complete package).

Nevertheless, it is appropriate for us to monitor and be concerned with these developments. This is another argument for an association with a publishing agent able to include our journals in such negotiated packages. Alternatively, we might begin to investigate such possible arrangements ourselves. Since each such large-scale contract is negotiated individually, the administrative overhead in this approach could be very high.

It is also emphasised that whatever arrangements might be made to incorporate subscriptions to our journals in large-scale package deals, we must safeguard the requirements of subscribers for individual titles.

5.7. Free access

This approach is technically the easiest to manage, but of course affords the greatest threat to income. It is recommended however that it not be rejected out of hand (a surprising number of web-based services are surviving that are free at the point of delivery). Thought should be given to alternative means of income generation (commercial sponsorship; page charges; direct levies on adhering bodies). It should also be considered whether the economics of the process might argue in favour of free access to back issues (*e.g.* those greater than five years old).

5.8. Access control protocol

It is assumed that access will be provided for qualified users (*i.e.* registered subscribers) by allowing access to one of a range of permitted IP addresses; or by the issuing of a username and password-protected account. Support for users who forget their password, or who move to a different IP network *etc.* will have to be provided. At present it is an open question whether this will be primarily an R&D cost (*i.e.* whether it will be possible to write scripts that will help users in most cases), or whether there will be a burden on providing direct end-user support from administrative resources in-house. It also needs to be elaborated whether subscribers should be offered IP *versus* password access on the basis of whether they have an institutional or individual subscription; or whether some other criteria are more appropriate.

These questions should be investigated energetically during the time when access is provided free of restriction.

6. Mirroring

It is likely that there will be complaints about the response time from the server whatever the decision made about its initial hosting and connectivity. It is recommended that the quality of the initial service be assessed during its first year of operation, but from the start we should be thinking about the approach to mirroring the server.

The IUCr web mirrors are not appropriate to this task. The pan-crystallography service has been deliberately engineered to make the minimum possible demands on mirror servers, so there is no guarantee that a national mirror that happily hosts the pan-crystallographic information will be adaptable to the scale and demands of a commercial-quality controlled-access service. In addition, there may be acceptable use restrictions on many of the mirrors, which operate mostly under academic rules.

It is believed that Apache software using access control files for restricting access could be effectively mirrored across a number of Unix workstations identically configured and under our complete control. This is the approach of the 'facilities-managed' service, and we feel it is appropriate for future mirrors on other continents, provided suitable service providers can be identified. It is reasonable to suppose that the costs for each mirror will be of the same order of magnitude as for the facilities-managed solution described for the initial hosting. In practice, telecommunications and hardware costs in locations such as USA and Australia may be lower.

Some R&D work would be required for establishing secure and synchronous mirroring of access control files. Building on the experiences gained in managing the pan-crystallography mirrors, this would hopefully be modest (*i.e.* probably less than a person-month).

Note that if we proceed with a development host at Chester and a production facilities-managed host, the journal contents and access control could be synchronised between the two using the same techniques that would be employed in wider mirroring.

7. Third-party partners

The electronic journal initiative was originally developed to a model that envisaged third-party hosts. Experience with the Blackwell/Munksgaard *Synergy* server proved somewhat disappointing on a number of fronts (whence the impetus to develop our own service). However, in the longer term there may be advantages to integration within or alongside larger-scale players.

Interviews with librarians suggests that large-scale library consortia may wish direct acquisition of the DTD, full-text SGML files and PDF versions of journal articles to implement their own local delivery and information management systems. Since the requirements of different library-scale subscribers may vary, and since they themselves will wish to reduce the number of disparate arrangements they have to make, there is much to argue in favour of collaborating with publishers or agents who carry a lot of influence into the library arena.

In the light of these considerations, we have also been investigating the possibility of working with some third parties.

7.1. Blackwell/Munksgaard

As a large-scale system incorporating dozens, and ultimately hundreds, of journals, the Blackwell model relied on a consistent production model, and this was one of the factors that led to our parting company with their initial proposal. Production of our journal articles to the Blackwell/Munksgaard DTD would have involved either very substantial time in training editorial staff to use a new DTD, or substantial R&D time in writing translation tools. Alternatively, we could have contracted to the service operators through Munksgaard the responsibility for modifying their HTML generating software to take account of our different DTD. This was the approach that was in fact tried, but the quality of translation was poor, and the costs high and rapidly escalating. Beyond the direct costs, the disadvantages of locking in to the Blackwell DTD were seen to be too great: there would be little or no scope for us to extend the functionality of the online product by developing our own DTD. Another major factor with the Blackwell service was the poor opportunity to exhibit the identity of the IUCr in the way in which the journals were presented on the *Synergy* server.

Munksgaard were subsequently asked to tender an offer for hosting a PDF delivery system for page images of the IUCr journals within the *Synergy* system. The cost of this would be high: exact figures will be available for the Finance Committee meeting.

The main advantages that we would have seen in retaining links with the Blackwell/Munksgaard *Synergy* system are

1. the provision of links through author names and keywords to Medline and other databases

2. linking to other full-text journal articles hosted on the Munksgaard server

3. the ability to advance into other electronic publishing areas on the back of bulk export of Blackwell journals to other areas

Of these, (1) is something that we are already beginning to investigate for our own service; and (2) is of rather limited importance given the composition of the Blackwell/Munksgaard journals bundle.

It is really with item (3) – the possible sale of the electronic journal to other markets – that the benefits in retaining links to Munksgaard may lie. Investigation of such a possible benefit is outside the scope of this report.

7.2. Turpin Distribution

We contacted Turpin, who are a company experienced in the worldwide distribution of printed and electronic publications for academic and STM publishers (**http://www.turpin-distribution.com**). They are a wholly-owned subsidiary of the Royal Society of Chemistry, responsible for the distribution of RSC journals, and involved in the RSC electronic journal programme (though RSC itself hosts and manages its own e-journal service).

What they can offer:

1. Above all, flexibility. They are prepared to take full-text SGML and convert to HTML (as with *Synergy*); but they are also prepared to take header SGML and publisher-supplied HTML; solely publisher-supplied HTML; or to act simply as an access gateway to the publisher's own web site.

2. Full-text indexing, search and retrieval, based on the textual content of PDF files. The search engine operates in many modes and can search all journals or groups of journals which they host.

3. They are negotiating links to OCLC, Silver Platter and ISI to enhance interoperability between various electronic information services.

4. Close liaison with publishers, subscription agents and libraries, offering a bridge between us and the STM publishing world generally.

5. Access control to meet the publisher's specification; and end-user customer support.

6. The prospect in the near future of document delivery based on secure online ordering and credit-card payment.

7. Relatively inexpensive (exact figures will be available for the Finance Committee meeting).

The service is hosted on a facilities-managed server on the UUNET/Pipex backbone.

We were impressed by their flexibility, willingness to accommodate trials (they are currently investigating at no cost some of our files to establish what would be involved in tailoring our source material to their system). They have also received strong and supportive trade reviews.

7.3. Institute of Physics Publishing

The publishing wing of the Institute of Physics has a mature electronic journals site hosting full-text (PDF) access to all its journals since 1993. (**http:www.ioppublishing.co.uk/EJ/**) HTML versions of the journals are planned in the near future. The production cycle involves LaTeX source files; HTML will be produced by the Perl *latex2html* software (as used, for example, in our online Teaching Pamphlets). Bibliographic references are parsed from LaTeX source and marked up in XML prior to querying INSPEC and AIP bibliographic databases.

Currently the IoP service is driven from an Informix database that contains all published content; interestingly, they have found that this does not scale satisfactorily, and they are re-engineering their system so that the database component is more concerned with servicing typical database queries, while the content will be broken out into static HTML pages (*i.e.* they will be migrating towards the architecture we have established).

Their service is hosted on a Sun Enterprise 5500 server with a 2Mbps leased-line connection through Pipex.

What they can offer:

1. Full integration into their system. This would be at the level of PDF page images and navigation and article header functionality provided by translation of our SGML source files. This would be similar in some ways to the *Synergy* model but without full-text HTML.

2. Alternatively, they could accommodate our structured full-text HTML file tree, while providing access control and end-user support.

3. The third option is to act as a hosting service only; we would manage access control.

4. In all cases, we could link references to their 'Hypercite' system, which provides links to IoPP and APS journals, to the LANL preprint server and to the INSPEC bibliographic database. This would, however, entail extra development work on our part to tag references in an appropriate manner.

The IUCr would retain extensive control over the look and feel of the pages relating to our own journals. IoPP would be recognised by a small logo or footnote to each page.

IoPP have been asked to quote for providing the various levels of service indicated above.

7.4. American Institute of Physics

American Institute of Physics provides extensive electronic journal services for the journals of the American Physical Society and a number of other learned societies. Their user interface is very comfortable, and early discussions suggest they have substantial technical expertise, and an understanding of how our requirements fit in with, and also differ from, the services they already provide. (**http://www.aip.org/**)

What they can offer:

1. A presence in North America. Transatlantic connectivity is now much more reliable than it was a couple of years ago, so the benefit here may be political as much as technical.

2. Linking to an extensive range of physics titles and to the ISI database.

3. Access control.

AIP were also asked to provide a quotation for hosting our journals. The requirement is that they host our full-text HTML and supplementary materials in the format already available within our prototype service, but apply access control based on our subscription records from Munksgaard. Costings will be available for the Finance Committee meeting.

7.5. Other players

Other parties exist which may in future have a role to play in hosting, redistributing or peering our electronic journals service. They include: other publication service providers such as Dawson, Ovid, Swets, Ebsco, High Wire; library-based services such as BIDS/ingenta and OCLC; online community resource concentrators such as ChemWeb.com and BioMedNet.

We should continue to monitor the role played by each of these suppliers, and be prepared to investigate any that open up a route to increased market penetration or greatly enhanced functionality. However, we feel that the correct approach is to continue to develop our own journal to suit best the needs of the community we serve; and to require collaborative enterprises to retain those features of our product which are of greatest importance.

7.6. Partners or Predators?

When we began discussions with these parties, we felt that they might provide a short cut (and perhaps an economical route) to managing access control and hyperlinking; but as our work on our in-house product has progressed, we are more confident of performing these functions ourselves. While not ruling out the possibility of pursuing these investigations further, we see that they represent longer-term considerations that have more to do with strategic partnerships for mutual gain than necessary technical support.

It also becomes apparent from discussions with such potential partners that their interests in collaborating on electronic journals projects are intimately related to their level of interest in acquiring distribution or advertising rights over the print journals also.

So we are wary of pursuing discussions along these lines outside of a context which includes the forward direction of the entire publishing enterprise. Having said that, we again point out the possible benefits of working with larger players in placing the electronic product in other markets.

8. Timescale

The prototype electronic journal is in place. We recommend that it is launched as a pilot service during the Glasgow Congress. An appropriate announcement could be made during the Managing Editor's presentation to the Commission on Journals Open Meeting. The existence of the service and its links to the existing IUCr information services could be emphasised, and the web services re-launched as 'Crystallography Online', during the Research and Development Officer's presentation in the same session.

Immediate public access should be permitted to the prototype on the Chester server. Some of the functionality of the prototype service is rudimentary or incomplete. It is recommended that these evolving services be developed further with a view to launching a fully functional service on the production server (whether at Chester on a higher-bandwidth leased line or at a facilities-managed site) at the beginning of 2000.

It is recommended that access to the service should continue to be free during the first part of 2000, to allow study and implementation of a subscription model of choice, and to attract potential readers, subscribers and advertisers. The objective would be to have a reliable, secure and fully functional access-controlled service in place no later than the beginning of 2001.

This should also be the target date for full integration of *Acta Crystallographica Section C* within this service.

9. Conclusions

This report confirms the technical feasibility of launching an electronic journal based on the resources of the Chester Editorial Office. It also indicates the need for careful study of the development and growth of the electronic journal. For many years publishers have hesitated to enter the electronic arena at all, for fear of the unknown. Initial steps have now been taken, and we cannot enjoy the luxury of remaining motionless; but the way ahead still leads through tangled undergrowth and twisting paths.

Appendix A.
Design objectives

It has not always been clear what exactly is meant by the phrase 'electronic journal'. Here we specify the objectives towards which we have been working.

A.1. The full contents of the printed journals should be available over the Internet

The World Wide Web is widespread, and would appear to be the natural transport medium. Nevertheless, development should not unnecessarily exclude the use of other Internet protocols in exceptional cases (*e.g.* email, ftp), and should certainly not exclude the adoption of other future technologies.

A.2. High-quality page images should be available

Adobe Portable Document Format (PDF) is a suitable encoding mechanism for page images. It is largely platform independent; is supported by free applications (such as Adobe Acrobat Reader) that permit multiple-page documents to be handled as coherent units; is well suited for web transport, especially with byte-serving server software; is reasonably compact (given the necessarily large file size of any format that displays raster graphics); and is a well-documented – though proprietary – format, already very widely distributed and therefore likely to have a lifetime of at least a decade.

A.3. Full-text articles should be available for web browsers

The product should be able to be read in a web browser, with active hyperlinks to documents of related interest, including but not restricted to: multimedia components of the publication itself; associated supplementary documents supplied by the author; related data sets in public databases; related articles in IUCr journals; related articles published elsewhere.

At the current time, such full-text articles would naturally be marked up in HTML; it is likely that web browsers will evolve to read XML or SGML markup directly, and migration to these formats should be possible (in practice, of course, the source format for the articles is already SGML).

A.4. Access to individual articles should be a subscription service

The principle is that users should pay for access to individual articles, and so access must be controlled at this level. The appropriate charging model is still open to review; but we feel that it is appropriate to design access control on a per-article basis, whether to provide for the possibility of pay-per-article charging, or indeed to make freely available individual articles such as Commission Reports deemed to be suitable for free access by the entire community.

A.5. The contents of individual issues should be freely browsable

Collections of articles online should be organised in the same way as the printed issues – this preserves the unity of *Festschrifte* or other thematic issues, and preserves a link between the online and printed versions of the journal. Permitting unrestricted access acts as an advertisement for the contents of the journals and can be the basis for an email alerting service.

A.6. The contents should be searchable by author and by title, and in due course also by keyword

A.7. Full bibliographic data should be available for all articles

A.8. Supporting (supplementary) materials should be freely available for articles

This is straightforward for materials such as CIF data sets or mathematical appendices in TEX. Digitisation of hard-copy materials (scanning to PDF format) allows the provision on demand of any supplementary document (including existing documents archived as microform).

A.9. Access should be provided to articles already published

The complete digitisation and online supply of every article ever published in IUCr journals would greatly enhance the usefulness and power of the product. Sensibly only scanning and delivery of PDF format page images of earlier articles is realistic; but even this opens new possibilities for thematic reprint collections and educational uses.

A.10. The IUCr should have full control over the look and feel of the online journal

Design decisions for the online journal should complement those for the hard-copy edition.

A.11. The online journals should be integrable with the existing IUCr pan-crystallographic web services

Rapid access to information about the journals and contents of individual issues is provided by the network of mirrors of the IUCr web site. This convenience should be retained. There is therefore a need to allow easy transition from the local mirror of the web services to the active journal delivery site; and as much opportunity as possible to move from the journals site on to the appropriate web services mirror. Articles of general interest from the journals (obituaries, book reviews *etc.*) might also be physically hosted on the information web servers.

Appendix B.
Components of the electronic journals system

To meet or accommodate the design objectives, the following elements of a complete online journals service are identified. This checklist identifies both existing components and elements still to be implemented.

Component	*Source, details*	*Status*
1. Article DTD	Fotek, maintained by IUCr	already in place
2. Issue DTD	developed by IUCr	already in place
3. SGML editing software	ArborText ADEPT	already in place
4. SGML → HTML conversion software	developed by IUCr	already in place
5. Editorial workstations	Sun Sparc5/Ultra 5	already in place; rolling renewal programme
6. Working file storage	Sun Ultra 5 + 90 GB disk	already in place
7. R&D development machine	Pentium III Linux PC	already in place
8. Service hosting hardware	Sun Ultra 10	exact specification may depend on connectivity decision
9. Internet connectivity upgrade	2Mbps leased line or FM service	
10. Coherent workflow system		already implemented
11. Coherent filesystem architecture and file naming conventions		already implemented
12. SGML source files	Typesetter/in-house	already implemented
13. High-resolution graphics files	Typesetter	already implemented
14. Low-resolution graphics files	in-house conversion	already implemented
15. HTML files	in-house conversion	already implemented
16. Search engines	prototype, heuristic, non-systematic	already implemented
17. Relational database management system	YARD	already implemented
18. Published items database (PIDB)	bibliographic metadata	under development
19. Enhanced search software		awaiting development of PIDB
20. Back-issue PDF scanned files	HEDS/in-house conversion	
21. Directory-level access control		under investigation
22. Subscription information handling		under investigation
23. Registration mechanism for subscribers		under investigation
24. Alerting service		to be undertaken
25. Links to databases	PDB, CSD *etc.*	PDB in place; others to be investigated
26. Links to citation databases	Medline, ISI, INSPEC *etc.*	to be investigated
27. Links to other publishers' articles		to be investigated
28. Internet name space registration	iucr.org	already in place; others?
29. Publication name space registration	DOI	to be investigated
30. Protection of identity	Crystallography Journals Online™	to be undertaken
31. Integration with Crystallography Online		to be undertaken

3. Description of the prototype service

The overall service is named **Crystallography Journals Online**. It is recommended that the name and distinctive logo (rendered in staggered black and red Optima typeface) be registered as a trademark in the same jurisdictions under which protection is sought for CIF, mmCIF and other distinctive trade marks or service marks of the IUCr. Similar considerations should be given to other devices identifying the individual journals and other services, *e.g.* 'Foundations of Crystallography Online', 'Crystallography Online'.

The service entry page displays images of the current covers of each of the six journals, together with links for each journal to: the contents of the current issue; the tables of contents of previous issues; reader services; and services of benefit to prospective authors. Tables of contents of back issues provide links to individual papers (for issues in 1999 and subsequently), and to the table of contents of individual issues back to the early 1990s. It will soon be possible to extend the full contents listings back through the entire history of each journal, including the undifferentiated *Acta Crystallographica* of 1948–1967.

Each page has a panel at its top, permitting easy navigation to other parts of the service.

For each article, the navigation bar offers access to the following features: the journal home page; the table of contents of the current issue; a search page; a help page; a PDF version of the current article; any supplementary documents (CIF data sets, unpublished appendices, multimedia *etc.*) associated with the article; and a link to the next article. There is also a mini-table of contents allowing navigation within the current article, implemented as a drop-down menu for browsers capable of executing Javascript, or as a scrollable panel for other browsers.

The body of the article in HTML format has a very rich set

of hyperlinks, allowing the reader to jump immediately to any referenced section, table, figure, footnote or reference; or to retrieve data files in CIF format from the IUCr archive, or from the PDB; or to retrieve bibliographic information about any cited article in an IUCr journal. Files of structural data are served *via* a '3D view' link in a manner allowing three-dimensional visualisation of the molecule described.

Figures are represented on the main article page as captioned thumbnails; each thumbnail is a link to a full-page medium-resolution graphic of high quality. In principle, high-resolution images could also be delivered in the few cases where this might be necessary for full understanding.

Fig. 1. Entry page of the online journals service.

The use of HTML frames allows the full-text HTML, the PDF version of the paper, and any associated supplementary documents or search results to be swapped in and out of the browser window with ease, thus increasing the ease of use and convenience to the reader. Some work will be done to ensure that readers without access to frames-capable browsers may still access the full content of the journal, albeit with less facility.

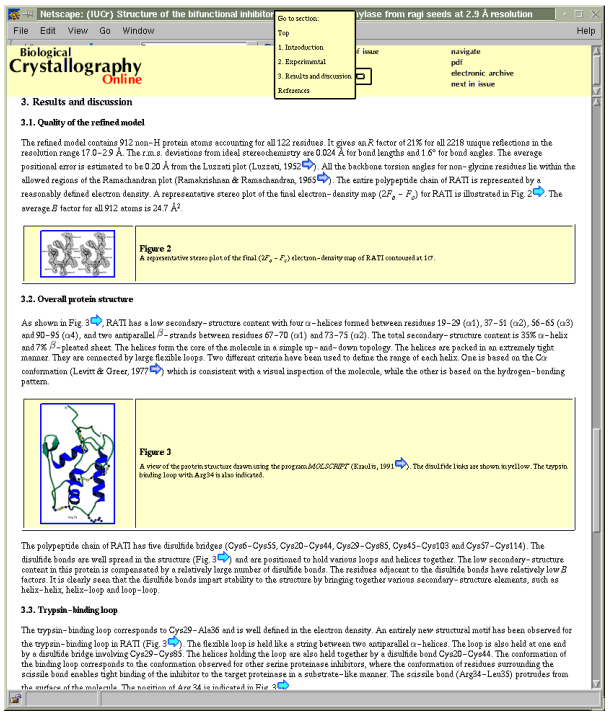

Fig. 2. Typical page of an online journal, showing thumbnail figures, rich hyperlinking of content, and internal navigation menu (at top of frame).