18-Data Bases and Communications

following which it has been discussed widely at several international conferences. Also, various respective organisations and the Directors of SR laboratories have been consulted. These consultations have echoed the feeling that a Journal dedicated to SR is urgently required.

OCM-18.02.04JOURNAL OF APPLIED CRYSTALLOGRAPHY. By A.M. Glazer, Clarendon Laboratory, Department of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, UK.

The Journal of Applied Crystallography is now in its 26th year, and over this time it has been continuously evolving. In recent times, a number of special features have been added to the journal, such as the incorporation of Lead Articles, where an author is invited to submit a paper on a subject of general interest to the readers. Such papers are refereed in the normal manner, but are given prominence once accepted. In addition, JAC has come to serve the crystallography community by acting as the main source of publications of computer programmes, new commercial products, and laboratory notes. From time to time, JAC also publishes Fast Communications.

Manuscripts are normally submitted to the Editor or one of the coeditors, who individually arrange for at least two referees. Each member of the team is responsible in the end for acceptance or rejection of a paper, with the Editor having the final say in case of a dispute, which, in the experience of this Editor, is very rare. The editorial team members communicate with each other regularly, principally by sending around monthly status reports to each other to ensure that everyone knows which papers have been submitted and their present position. From time to time, this procedure has picked up cases of authors submitting the same manuscript to more than one editor. The status reports are also sent to the Technical Editor in Chester who keeps an overall eye on all of the journals. Without the efficient running of the Chester offices, journals such as JAC could not successfully appear with the high quality that the crystallography community has come to expect. In this talk, present and future policy of the Journals Commission regarding JAC will be discussed.

OCM-18.02.05 CHESTER PUBLISHING OPERATIONS. By M. H. Dacombe, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

The main function of the technical-editing office of the IUCr in Chester is the production of the journals *Acta Crystallographica* and *Journal of Applied Crystallography*. Over the past few years there have been considerable developments in the way in which these journals are produced.

In particular, papers intended for publication in Section C of *Acta* are now submitted direct to the Technical Editor and various checks are carried out on each structure (these checks include symmetry, geometry and previous publication) before the paper and the results of these checks are sent to a Co-editor.

Section D of *Acta* was launched in January 1993 and the possibility of starting another new journal on synchrotron radiation is under active investigation.

Machine-readable submissions to all the Union's journals is encouraged and much material is now prepared in Chester as camera-ready copy - *Fast Communications*, regular submissions, indexes and, particularly, *Regular Structural Papers* in Section C. The whole procedure for Section C has been facilitated by the use of CIFs and a by-product is that a CIF file for each structure is archived in Chester and is also available for onward transmission to the established databases. Additionally, the ninth edition of the World Directory of Crystallographers will be held in Chester as a database in CIF format. The database in Chester is accessible via email and it is intended to make it available on-line in due course. Various other publications are handled in Chester. The most important of these is the series *International Tables for Crystallography*.

OCM-18.02.06 **COMPUTERIZATION OF PUBLISHING AND RELATED DATA ARCHIVING** By B. McMahon, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England

Since the last congress of the IUCr there has been considerable investment in high-performance computing equipment in the Technical Editor's office. Six Unix-based graphics workstations and a file server reside on an Ethernet local area network; there are also network connections to IBM PC and Apple Macintosh-type machines. The network provides for data transfer between the computers and data input/output equipment: floppy disk drives supporting various formats; cartridge and open-reel tape units; CD-ROM and optical disk drives; and laser printers. There are also full network connections to the UK national academic network JANET and to the global Internet.

Structural data in papers submitted for publication to *Acta Crystallographica* are checked using a wide range of standard crystallographic software, freely donated by the community. The validated or corrected data may then be typeset without the risk of typographic errors being introduced by manual keyboarding. When these data arrive in Chester in hard-copy form, they are input to the local computer system as Crystallographic Information Files (CIFs). Authors may also submit CIFs electronically; where this is done in the case of *Regular Structural Papers* in Section C of the journal, the entire paper may be typeset automatically. The acceptance of CIF as the data transfer standard by the IUCr has ensured that these procedures can be developed for optimum efficiency.

It is the *efficiency* of the new procedures that render them cost-effective in terms of journal production. Other techniques for converting different word-processed input to the requirements of the journal are also being developed; but without exception they require more manual intervention and consequently have a less benign impact on the publishing process.

Strategies are being developed for making available to the crystallographic community the structural data stored in machine-readable form.

OCM-18.02.07
THE MACROMOLECULAR CIF DICTIONARY.
By Keith D. Watenpaugh*, Upjohn Laboratories, 301 Henrietta St., Kalamazoo, MI 49007, USA; Helen M. Berman, Chemistry Department, Rutgers University, Piscataway, NJ 08855, USA; Philip E. Bourne, Department of Biochemistry and Molecular Biology, Columbia University, 630 W. 168th Street, New York, NY 10032, USA; Paula M. D. Fitzgerald, Merck Research Laboratories, P. O. Box 2000, Rahway, NJ 07065, USA.

A dictionary of Crystallographic Information File (CIF) data items to describe both the macromolecular crystallographic experiment and the structural results of the crystallographic experiment has been developed. A CIF contains crystallographic data stored in a subset of the STAR format (S. R. Hall (1991). *J. Chem. Information* **31**, 326-333). The CIF format is designed to facilitate archiving and exchange of crystallographic data in a standardized manner, with the particular goal of automating the electronic publication of crystallographic results.

The International Union of Crystallography (IUCr) has sponsored this initiative, and is actively encouraging the adoption of CIF as a standard for the field. A core dictionary of data names sufficient to describe the small molecule crystallographic experiment (S. R. Hall, F. H. Allen, and I. D. Brown (1991). *Acta Cryst.* **A47**, 655-685) was adopted by the IUCr at the 1990 International Congress in Bordeaux.

Subsequently, the IUCr commissioned a working group to extend the CIF definitions to cover the macromolecular case. The members of the working group are Drs. Paula Fitzgerald (Chair), Enrique Abola, Helen Berman, Philip Bourne, Eleanor Dodson, Arthur Olson, Wolfgang Steigemann, Lynn Ten Eyck and Keith Watenpaugh. This working group is presenting this macromolecular extension of the core dictionary to the IUCr for formal approval at the 1993 Congress in Beijing. Following the recommendations of its executive advisory board, the Protein Data Bank will be adopting CIF as the new format for the crystallographic information distributed by the PDB.

The working group is actively seeking input from the crystallographic community concerning the macromolecular CIF dictionary. The completeness and accuracy of the definitions along with the completeness and clarity of the accompanying documentation is an on-going project.

OCM-18.02.08

THE CIF DDL DICTIONARY AND ITS ROLE IN CIF SOFTWARE APPLICATIONS. By S.R. Hall, Crystallography Centre, University of Western Australia, Nedlands 6009, Australia.

The Crystallographic Information File (Hall, Allen & Brown, *Acta Cryst.* A47, 655-685) was adopted by the IUCr in 1990 as a primary method for sending manuscripts and data to *Acta Crystallographica C* electronically. This approach has proved successful and, with the widespread availability of CIF generating software in standard crystallographic packages, it is now being applied for other data exchange and database purposes. The CCDC (Cambridge), ICDD (Philadelphia) and PDB (Brookhaven) databases are expected to adopt CIF's as their preferred data exchange medium. Other major data bases will follow, especially when the checked CIF data becomes available directly from the *Acta* archives in Chester.

The existence of precise data definitions in the form of the CIF dictionary has been an important factor in the rapid acceptance of the CIF approach. The first electronic DDL version of a CIF dictionary was released in 1991 as 'cifdic.C91'. This contained the definitions of data items commonly used in small molecule studies. At this conference the definitions of *powder* and *macromolecular* data, coordinated by working groups headed by Brian Toby and Paula Fitzgerald respectively, will be put forward for ratification by the IUCr CIF Committee. Following this, the DDL dictionary files 'cifdic.P93' and 'cifdic.M93' will be released for applications in these fields.

This talk will highlight the properties of the DDL CIF dictionary and how it is applied by computer programs such as *CYCLOPS, CIFtbx,* and *Star_Base.* All future software applications which read, write or manipulate CIF data will need to invoke these dictionaries to check the conformance and identity of data.

18.03 – Electronic Diffusion of Information

DS-18.03.01 COMPUTER NETWORKS: THEIR USE AND LIMITATIONS. By Y. Epelboin, Laboratoire de Minéralogie-Cristallographie, URA 009 CNRS, Universités P.M. Curie et Paris VII, 75252 Paris Cedex 05, France

The establishment of a computerised World Directory which is now under way has lead to a general thinking about the use of networks in information exchanges between scientists. In this talk we will describe how the networks can be used and the limitations arising from their topology and technology.

The most widely used technology, nowadays, is an extension of the local area network protocol TCP/IP better known (but improperly) as Ethernet networking. Its development is linked to the diffusion of Unix. The second one is an IBM technology used in the famous Bitnet network. ISO protocols such as X25 have been installed in United Kingdom (Janet) and are developing in other European countries. In other communities such as astronomy (SPAN) a Digital technology is used everywhere. This diversity restricts the exchanges between the different networks. Electronic mail is the only flow of data which circulates easily between networks and other information (software, data ...)are often encapsulated in messages which severely restricts the possibilities of communication.

However more and more people have a transparent link to the TCP/IP world which allows various communications which will be discuss in this talk:
- e-mails which pass through all networks and which may be used to address a query to a server. The world directory will be available to everybody by this means, sending queries in messages and receiving the answers from the server in the same manner.
- lists servers where people are registered on a list and receive all the information, moderated or not, which is addressed as messages to the list. The use of lists servers is compatible through different networks.
- bulletin boards where people establish a connection to a server and consult a given information menu by menu. It is restricted to the feasibility of establishing a connection to the server and retrieving a piece of data may be limited since file transfers are limited or impossible between different networks.
- direct connections and use of dedicated software such as access to a database. The user works directly on the remote server. The restrictions are the same as for bulletin boards. Scientists will have access to the World Directory by this means.
- anonymous ftps, mainly restricted to the TCP/IP world, which allow to retrieve character and binary files.

A new technique is developing, based on a client server protocol where part of the software is resident on the user's machine, part on the server.

Other limitations appear when exchanging formatted documents which contain text, formulae and drawings. However one may foresee the day when the IUCr Newsletter will be available through the network.

The advantages, possibilities and limitations in the use of networks will be discussed explaining the choices for the World Directory of Crystallographers.

DS-18.03.02 DISTRIBUTED NETWORK INFORMATION SERVERS By B. McMahon, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England

It is almost a commonplace in the modern world that valuable and useful data may be accessed and retrieved from databases, file transfer archives and electronic bulletin boards. Crystallographic structural data are now archived on the IUCr Editorial computer system, thus making feasible the concept of an electronic journal of crystallography. The logical extension of these ideas is the provision of an electronic library, a facility for retrieving diverse information from any of a collection of sites providing data.

Already, the use of standard protocols across the global Internet has given rise to several tools for drawing together the library resources of various archive sites. *gopher*, developed at the University of Minnesota, provides a hierarchy of menus allowing structured access to data categories. Each entry in the menu may represent data stored on a different machine, anywhere in the world. The casual user may choose a menu item without knowing or needing to know the location of the information sought.

The *WAIS* (Wide-Area Information Server) system is a standard way of indexing the textual content of files stored at an archive site. A *WAIS* request can select from a document collection all files containing a requested word or term; it can rank the selected files in order of likely relevance by applying a heuristic test based on the number of occurrences of the target term; and it is in principle possible to perform searches based on context. *WAIS* is often used to supply index services to a data collection initially accessed through *gopher*.

A third global information dissemination system, *WWW* (World-Wide Web, developed at CERN), provides for formatted documents to be read on-screen as though typeset. Hypertext links within the document allow cross-references to be followed as the document is perused. As with the other tools, the links in the hypertext chain may reside on geographically remote computers.