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.

The Protein Data Bank operates a *WAIS* server permitting retrieval of abstracts by index term. The IUCr is investigating similar techniques for accessing its collections of CIF structural data files and biographical information in the *World Directory of Crystallographers*. The technical aspects of such a project are less daunting than the requirements for data security and cost retrieval.

DS-18.03.03 ADVANCES IN GLOBAL DATABASES AND COMMON DATA EXCHANGE FORMATS. By S.R. Hall, Crystallography Centre, University of Western Australia, Nedlands 6009, Australia.

Crystallographers are prolific generators and users of data, and as a consequence their databases are amongst the largest and best organised in science. The rapid growth of databases has been possible because of the modularity of crystal data; the strong computational background of the discipline; and a high level of cooperation within the profession promoted by its national societies, the IUCr and its journals. These factors have enabled databases to be developed and maintained at reasonable cost.

The methods for accessing these databases have, however, changed little in the past two decades, despite rapid changes in computer technology during that time. Certainly data retrieval rates have improved; better graphical interfaces have made data interpretation easier; and remote access facilities are emerging, but the currency of data is still determined largely by the physical distribution of media. The economic and organisational reasons for this are, however, about to change.

Most scientists have access to global computer networks, though these tend to be poorly utilised (except, perhaps, for email use) because of unfamiliarity, cost and inertia. The continuing improvement of national and international network services strengthens the rationale for globally accessible single-site databases. Such facilities provide for immediate online access to the latest data at a reduced overall cost to both the distributor and the user. The IUCr CIF data archive, and the next World Directory of Crystallographers (WDC9), will be available soon as global databases directly accessible from the Chester IUCr office.

Common data exchange protocols, such as the STAR File (Hall, 1991 *JCICS* **31**, 327-333), serve a vital role in the development and access of global databases. This talk will discuss these, and use the Molecular Information File (MIF/SMD), a recent development for exchanging general chemical data, as an example of this interaction.

DS-18.03.04 EXPERIENCE WITH THE PAN- EUROPEAN INFORMATION SERVICE *CONCISE*. By H.D. Flack, Laboratoire de Cristallographie, University of Geneva, 24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland.

CONCISE is a pilot information service, a part of the *COSINE* project (of *RARE*) for electronic communications for European research. *CONCISE* is accessible over the European research networks, over public data networks and using the telephone network. Access is by way of email, interactive session or file transfer. At the present time (25th Feb. 1993) use of *CONCISE* is free of charge. From the readers' and information-providers' point of view, *CONCISE* may be viewed as the electronic equivalent of a newsletter structured on a (mainly) hierarchical arrangement of items. The ordinary user may read and copy these news items. However contributions must be sent to an information provider (editor) for them to be included on *CONCISE*. The European Crystallographic Committee has arranged to have a special interest group for crystallography on *CONCISE* and the secretary is currently acting as an information provider for crystallographic news in Europe to *CONCISE*. The news items included are announcements of national and European meetings, meeting reports, indexes of *Acta Crystallographica* and the *Journal of Applied Crystallography*, personalia, news from the International Union of Crystallography and the European Crystallographic Committee, etc. Certainly the most flexible way of consulting *CONCISE* is by way of an interactive session. However for those who have difficulty obtaining an interactive connection - this is particularly true for the countries of East and Central Europe - access to *CONCISE* by email is a satisfactory alternative. For the information provider it turns out that submission of information by way of email is the easiest option, as an item is prepared in the form of an ASCII file with a few bookeeping commands added.

DS-18.03.05 PROMISE AND PITFALLS IN ELECTRONIC INFORMATION. By E.N. Maslen*, Crystallography Centre, University of Western Australia, Nedlands, Western Australia 6009

New communications technology has the potential to provide scientists with information on the latest scientific developments rapidly, and cheaply. The dream of a desk-top supercomputer is approaching reality, and technical means for linking that device to comprehensive data banks of scientific information should be with us by the year 2000. We will soon be able to transform scientific data into the form which we find easiest to assimilate, with minimal effort. This also has the potential to drive into bankruptcy those sectors of the scientific publishing industry unwise or unfortunate enough to mis-calculate its consequences. Scientists regard many aspects of management as peripheral to the public well-being, with good reason. Nevertheless the challenges in managing electronic publishing of scientific information are as demanding as the technical challenges whose resolution brought this new age. It is not just a question of anticipating which competing technologies will survive, but timing their introduction optimally. The nature of the advances in electronic information that are expected will be outlined, and the management challenges posed by those advances will be described.

DS-18.03.06 **ELECTRONIC MAIL: STIMULATING COMMUNICATION.** By M.M. Teeter, Department of Chemistry, Boston College, Chestnut Hill, MA 02167. USA

The advent of electronic mail has greatly impacted scientific research worldwide. Communication has sped up so much that conversations are possible via email nearly in real time. Many forums for communication across the network now exist and these can stimulate research collaborations as well as career development.

The first ingredient for communication is an *address list*. For crystallographers, this is currently available both by ftp and by fileservers via both the Protein Data Bank and EMBL. The address list enables direct communication for information or research collaboration as well as providing a mailing list of email accessible crystallographers.