Position paper on IUCr response to the *Global Information Commons for Science Initiative*

John R. Helliwell
IUCr representative to ICSTI
School of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, England and CCLRC Daresbury Laboratory, Daresbury, Warrington, WA4 4AD, England

Peter R. Strickland
Managing Editor, IUCr publications
International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU

Brian McMahon
IUCr representative to CODATA
International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU

12 September 2006

Summary
In response to a multi-stakeholder proposal for a Global Information Commons for Science Initiative (GICSI), the International Union of Crystallography (IUCr) has prepared a response, included as the Appendix to this paper. This paper provides the background to the IUCr's position, and summarises the IUCr's own initiatives in working towards the ideal of a global science information commons. The IUCr is broadly supportive of the GICSI initiative, but emphasises the need to realise this in a way that embraces established information providers committed to the ideal of universal and equitable access to scientific information, but funded through a variety of economic systems.

Executive Summary

The Global Information Commons for Science Initiative (GICSI) is sponsored by CODATA, ICSTI, INASP, WDC, ICSU in collaboration with IAP, TWAS, OECD, UNESCO and Science Commons, and aims to: (1) improve understanding and increase awareness of the societal benefits of easy access to and use of scientific data and information; (2) promote adoption of successful methods and models for sustainable open availability and facilitate re-use of scientific data and information, as well as cooperative sharing of research materials and tools among researchers; and (3) encourage and coordinate the efforts of those trying to devise and implement effective means to achieve these objectives. The initiative places particular emphasis on developing countries and the deployment of the results of publicly-funded research. All three goals will be promoted through the construction of an online 'open-access knowledge space', and specifically targeted projects carried out in collaboration with its participating stakeholders.

The IUCr endorses the goals of the Initiative, and supports the proposed implementation, subject to a sound business model and effective management by its participating stakeholders.

The IUCr recognises the Initiative's strong emphasis on the principle of open availability as an ideal medium of dissemination for publicly-funded research findings and data. However, the IUCr itself is one among many organizations that provide high-quality value-added publishing and data services, historically funded through subscriptions, direct sales and other commercially-mediated mechanisms. Although the IUCr is experimenting with various routes toward open availability, it is not clear that this can always be achieved, or that it is invariably the best solution.

The IUCr considers it essential for the good of science that there be full and equitable access to scientific data and information. *Full* means that no bar is imposed on grounds of race, creed, political opinion, economic status, disability, gender *etc.*; *equitable* means that the costs of providing the access be distributed according to principles of fairness and respect for individuals' ability to pay.

Introduction

The Global Information Commons for Science Initiative (http://www.codata.org/wsis/GICSI-prospectus.html) is sponsored by the Committee on Data for Science and Technology (CODATA), the International Council for Scientific and Technical Information (ICSTI), the International Network for the Availability of Scientific Publications (INASP), the World Data Centers (WDC) and the International Council for Science (ICSU), in collaboration with the InterAcademy Panel on International Issues (IAP), the Academies of Science in Developing Countries (TWAS), the Organisation for Economic Co-operation and Development (OECD), the United Nations Educational, Scientific, and Cultural Organization (UNESCO), and Science Commons. It is a multi-stakeholder initiative arising from the second phase of the World Summit on the Information Society in Tunis in November 2005. It has the following goals: (1) improved understanding and increased awareness of the societal benefits of easy access to and use of scientific data and information, particularly those resulting from publicly funded research activities; (2) wide adoption of successful methods and models for providing open availability on a sustainable basis and facilitating re-use of publicly funded scientific data and information, as well as cooperative sharing of research materials and tools among researchers; and (3) encouragement and coordination of the efforts of the many stakeholders in the world's diverse scientific community who are engaged in efforts to devise and implement effective means to achieve these objectives, with particular attention to developing countries. It is intended that all three goals will be promoted through the construction of an online 'open-access knowledge space', as well as specifically targeted projects carried out within the Initiative and in collaboration with its participating stakeholders.

As a matter of principle, the IUCr will strongly support any initiative that furthers the universal and equitable dissemination of scientific knowledge. The very nature of the scientific endeavour is that it seeks to unlock the secrets of nature through open, critical analysis of objective data, and it can achieve this only through unhindered sharing of data, observations, previous results and publications. Individual scientists and their professional organizations have always been active in protecting basic rights of travel, communication and freedom of thought, and the sense of common involvement in the scientific process transcends national, ethnic, religious or other sectional interests. In this spirit, the IUCr warmly commends the objectives of GICSI in lowering the barriers to dissemination and use of scientific information.

However, we temper this with the remark that scientific observations, data and literature are collected and managed under a variety of economic conditions and models, and urge that the understandable desire to establish rights of common use not threaten existing archiving and dissemination systems for which there is no current alternative.

The IUCr is unusual among scientific unions in its active involvement in four particular types of information resources of importance to crystallographers and structural scientists: in primary research reports and communications published in its journals; in authoritative reference compilations (*International Tables for Crystallography*); in curated archival databases of three-dimensional molecular structures; and in structure solution and refinement packages, an area not discussed in detail in the current GICSI prospectus. We consider each of these in turn, and discuss how we respond to the call for an information commons in terms of our current policies and, indeed, financial obligations to the global crystallographic science community.

First, however, we make some general comments on economic models. Much of the argument underlying the GICSI proposal rests on the attractive idea that the results of publicly funded research should be made available without further cost to the public that has provided the initial funding. In practice, however, governments do not generally enforce centralized handling of the output of research. Many public agencies seek to maximize value for money by contracting out services that are felt to be an unfair burden on central taxation. The extent to which this is done varies from country to country, and from administration to administration. To give one example: in the UK the national cartographic service, the Ordnance Survey, which was once an official body, has been reconstituted as an 'arms-length' trading agency, which must support itself through direct sales of its services. This is in contrast, for example, to the federally funded US Geological Service.

It may well be argued that scientific publishers and databases, which add value to the results of scientific research through the processes of peer review, editing, printing, distribution, annotation, curation and archiving, represent contracted-out services operating on a similar basis to a self-financing Ordnance Survey. This would seem, for example, to be the attitude of the current UK government (http://www.publications.parliament.uk/pa/cm200405/cmselect/cmsctech/249/249.pdf).

Having said that, we recognise that some changes in policy are occurring. UK Government funding agencies, and the major medical research funding charity The Wellcome Trust, have recently allowed their grant holders to include direct costs of publication, such as reasonable rates of open-access fees, in research grant proposals. This important change will allow the public at large free and open access to research outputs, a significant fraction of which relate to matters of major public concern such as health and medical matters and also environmental issues.

Here, our purpose is not to defend or argue for one particular approach to such value-added services; rather, to point out that different models, many well established for decades or longer, already exist in practice. Certainly, the publishing and database activities of the IUCr have grown up in the context of such a decentralized model of value-adding services. They have formed part of the market economy providing services generally to the community of practising scientists. However, this is not a completely free market economy. Distortions are introduced by the prevalent centralized funding of research activity, so that certain purchasing funds are inelastic to the available supply. We are familiar with the problem of library budgets being consistently capped or even reduced, while the volume of published research literature continues to grow.

We note that many of the information initiatives in the academic community nowadays seek to reintegrate much of the value-adding functionality of publishers and database providers into the centralized funding ethos of national research programs. Such initiatives include institutional repositories, preprint servers and open-access journals and data stores. However, based on our experience of almost 60 years as an information provider with the highest standards of peer review, data quality assessment and control, we wonder whether they have underestimated the real costs of ensuring the necessary level of quality. Certainly, publicly funded resources of the highest calibre do exist. The Protein Data Bank (PDB) is a shining example in our own area of interest. But such resources are far from cheap to run.

We also raise a note of caution about long-term reliance on public funding. Although governments allocate tax dollars to research programs on the basis of perceived need for the public good, there are few assurances of long-term support for individual projects if political objectives change (or tax revenues fall). While it can be argued that different patterns of funding merely reflect different (perceived) public needs, nonetheless the long-term management of the records of science needs to be secured independently of short-term funding decisions. We believe that this argues in favour of a diversity of funding strategies, and the persistence of some independent or privatised guardians of the scientific record. We do not believe this negates the goals of GICSI, but we need to emphasize that we seek to realise those goals within our own context as an independently financed scientific organization.

1. The journals of the IUCr

Acta Crystallographica was created in 1948, following the (temporary) demise of *Zeitschrift für Kristallographie* during the war, as an international journal of the highest standard of peer review and technical quality. It is a journal produced by the IUCr to meet the needs of its entire global community. It was published from the outset in technical and accounting partnership with commercial publishers (Cambridge University Press, then Munksgaard, now Blackwell Publishing), and paid for by subscription (although with early grants from UNESCO). As is almost universally the case for academic research publications, the success of the journals relies heavily on the time and commitment provided by individual scientists as peer reviewers and members of editorial boards. The IUCr operates on a not-for-profit basis, and is not liable to tax. Trading surpluses from subscription sales are reinvested in journals development, or allocated to the IUCr's other activities in education or support for crystallography in the international scientific community. The IUCr's stable of journals has now grown to eight titles, of which two are electronic-only; the other six are still distributed as print + electronic journals, and their web versions are popular.

The electronic medium has seen many innovations and value-added features in these web editions. As well as richly-hyperlinked articles available to subscribers, the journals offer free and seamless access to supporting materials, including primary data sets. The funding model at launch was based on the existing two-tier (institutional and personal) subscription models to the print journals. Traditionally, the journals had supplied supporting documents free of charge upon request; the electronic medium simply facilitated this access (and by automating it, removed the associated administrative costs). In an effort to use the electronic dissemination medium to help the community as much as possible, certain categories of article were made freely available. These included education papers, software application papers relevant to the IUCr's own exchange standard, and reports of the IUCr's scientific Commissions (including nomenclature recommendations).

For some years now, the open-access movement has been putting pressure on the traditional journal subscription model. There are various motivations behind the open-access movement. A prime motivational driver is one of principle, that the public should have easy, indeed free, access to the results of research already paid for through taxes. Another motivation is a simple feeling that value for money is lacking, especially in the light of (apparent) cost savings through electronic dissemination. (It must be remembered that the 'first-page' cost – *i.e.* the cost of peer reviewing, technical editing and typesetting an article – accounts for the bulk, perhaps as much as 95%, of a journal's production costs.) Another motivation is a reaction to continuing caps on library budgets in the tide of ever-increasing volume of publications

(sometimes coupled to 'Big Deal' arrangements with publishers that offer access to more titles but with an associated reduction in the freedom to direct the library's overall holdings); still another motivation focuses on intellectual property rights issues. Nevertheless, whatever the motivation of the advocates of open-access publishing, it is an aspiration that fits well within the framework of a global information commons.

Like many publishers, the IUCr has found it difficult to plan an economic transition to an open-access publishing model, given the long-established practice of funding scientific publication through subscriptions. A substantial number of new publications are appearing as open-access, typically funded by an 'author-pays' model, but few existing titles have converted to such a model. We have begun to experiment with transitioning to open access by introducing a hybrid system where an author may choose to pay the production costs for an article, in return for which it is made freely available. This experiment has been assisted by a generous grant from JISC, which has covered all articles from authors in UK universities and Higher Education establishments for three years. It is clear that open access increases the visibility of authors' work (as measured by download statistics), although the impact on citation still needs to be assessed. While the UK grant has provided a large number of open-access articles, it is noteworthy that the take-up of the author-pays model remains quite low from other authors. Of course, during this experimental phase access to the journals as a whole remains through institutional subscriptions, so the attraction of additional fees for authors may be limited.

Overall, the initial experiments have encouraged us to believe that open-access funding may be practical at least in some circumstances, and we are considering the feasibility of transitioning one or more existing titles to this method. The younger, electronic-only, titles seem the obvious candidates.

One consequence of such a change would be to shift the burden of payment from library to author, and in many cases to increase the cost to a contributing author beyond what would have been incurred as a subscription fee. Perhaps this will be a true test of the ideal of an information commons, when individuals incur higher costs to secure an overall greater common good. Such a burden may be particularly difficult for authors in developing countries, where the level of research funding is relatively low. It is noteworthy that *Acta Cryst. E* has achieved its unusual growth rate by attracting a disproportionately high number of articles from countries with developing economies.

The IUCr already takes account of the financial difficulties of developing nations by administering its own Journal Grants Fund, and by participating through its commercial partner in the activities of INASP. It seems appropriate to consider subsidization of 'author-pays' contributions by developing-nation authors to at least the same level as implied by current INASP programs and the Journal Grants Fund if we are to improve the access from developing nations to the information commons.

Another approach to changing the funding model for journals is to attract direct sponsorship, sufficient to cover production costs. That this cannot be neglected as a potentially viable funding source is illustrated by the findings of a survey carried out by the Association of Learned and Professional Society Publishers and published in 2005 (http://www.alpsp.org/publications/FAOAoverviewREV.pdf). However, success is likely to vary with the subject area of the journal. Sponsorship from bioscience companies has been sought for *Acta Cryst. F*, but with only modest results. There may also be concerns over

longevity and editorial freedom of a publishing model based on sponsorship, although it remains potentially a helpful source of additional funds, perhaps to defray the cost to contributors of an 'author-pays' open-access model.

2. *International Tables for Crystallography*

The *International Tables* series currently comprises eight large volumes of reference data and material essential for the practising crystallographer. There is also a low-cost *Teaching Edition* of extracts from the volume on space-group symmetry. An electronic version of the contents of this series is under development, with plans to provide it through the online service managed by a commercial partner (SpringerLink) in the latter part of 2006. Again, the economic model for such a new product is still evolving, and is likely initially to be based upon existing practice in the domain of printed publications. Hence, access will be sold to electronic editions; the access to the specific edition purchased will be maintained indefinitely, but access to a new edition must be purchased anew. An alternative model is one of subscription, where users pay a recurring fee to allow current access to the latest version of the online content. This means that users would recognise the specific value of up-to-date access, and thus explicitly fund the ongoing development of a *service*. We feel that this is particularly appealing as the online medium allows development of new databases, query engines, hyperlinked resource discovery and visualization. These both use the online environment effectively and are decoupled from a traditional commissioning/authoring/revision cycle. However, there is not yet a viable market for subscriptions to online book products, and we shall proceed initially with the edition-purchase model, although with a feeling that this makes it more difficult to integrate the reference information in the *Tables* with the growing pressure for open access and information commons.

It is also worth considering the investment of resources required for such reference works. Each volume represents the dedicated and considered efforts of experts in authoring, reviewing and editing authoritative compilations, followed by very detailed and labour-intensive technical editing of the highest calibre. An individual volume may take a decade to compile and a year to produce. The existing sales model generates income to pay for this labour (and to generate a modest surplus). If the contents were made open-access, the level of income required from other sources to support the product would be substantial. We are not aware of examples of sponsorship or other indirect income sources funding similar ventures in an open-access environment.

Nevertheless, we are also experimenting with opening access to certain types of reference information by developing an online encyclopaedic dictionary of crystallography. This will be a voluntary effort, involving members of various IUCr Commissions and other bodies under the guidance of an academic editor, and using collaborative web-based authoring and dissemination software. It is intended that the results will be freely accessible. This project will provide valuable experience on the feasibility of commissioning such authoritative resources, and an idea of how the real costs can be distributed amongst the participants.

3. Primary research data and databases

The primary research data in the field of crystallography are structure-factor files, reduced sets of data from raw experimental output that provide all the necessary information to solve, refine and validate a crystal structural model (*i.e.* to generate sets of atomic positional

coordinates and atom displacement parameters). There is a long-standing acceptance within crystallography of the principle that such primary data sets should be freely available for sharing and re-use (with appropriate credit) within the structural science community. Sometimes, when the actual solution of a structure is difficult, there is a case for allowing the original researchers to hold back disclosure of the structure factors for a reasonable period of time (*e.g.* six months). Likewise, certain commercial purposes are served by publication of the results of a structure determination while withholding full disclosure of structure factors. However, it is the policy of IUCr journals to require that authors deposit full sets of structure factors and atomic coordinates in support of any published determination of small-unit-cell structures, and other crystallographic journals and databases have similar policies. As a matter of policy, these atomic coordinates and structure factors are freely available for the use of individual researchers from the IUCr journal web site, and therefore already represent a key element of the information commons.

Given access to the primary data, it is possible and increasingly straightforward in most cases for a scientist to reconstruct the crystal structural model; and therefore it is only logical to disseminate the derived data that describe a structure according to the same policy. Since 1991 the IUCr journals have been collecting structural data sets describing the three-dimensional arrangement of atoms in a crystal (and their triaxial components of movement within the lattice), in a standard exchange format known as the Crystallographic Information Framework (CIF); and all such data sets are made freely available from the IUCr journals web site in support of the published articles that carry the scientific discussion of the significance of the results.

Similar principles of open availability of structural data are held by many of the curated databases of crystallographic information. The Protein Data Bank (PDB), a publicly funded database of protein and nucleic acid structures, provides coordinate data sets and structure factors (subject to any time-limited embargoes requested by depositors). The PDB is the primary depository for biological macromolecular structural data, and in this role supports journals that publish such structures, as well as serving the needs of the biostructural research community at large.

The Cambridge Crystallographic Data Centre (CCDC) is a not-for-profit organisation that curates coordinate data for organic and metal-organic structures from the research literature as well as from private depositions. The CCDC creates a comprehensive, validated and value-added database known as the Cambridge Structural Database (CSD), and the complete CSD System (CSDS) includes the CSD itself and associated software for search, visualisation and analysis of stored information. Access to the full CSDS is not open: users take out annual subscriptions which generate revenue to sustain the organization's activities, but subscription levels for academics are heavily discounted. Further discounts, via charitable grants-in-aid, are available to academics in developing countries. However, individual data sets are provided free of charge upon request to *bona fide* researchers. This represents a valuable service to the community, especially in the case of structures published in older journals or for journals that do not themselves have a policy of deposition and distribution for data sets. The CCDC also makes software for CIF validation and structure visualisation freely available to the crystallographic and chemical communities (see below).

Inorganic and metal crystal structures are administered by the Inorganic Crystal Structure Database (ICSD) and Metals Database respectively. Powders represent an interesting category

where mixtures of compounds form a sample and the Powder Diffraction Database holds these. All three of these databases operate on a fee basis incorporating academic discounts.

In all of these ways, crystallography has already established an effective information commons for structural data. However, a corollary of the ease of transfer of electronic data files has been the growing recognition of the need for well-defined practices to support and respect the intellectual property rights that these data sets embody. Thus, authors of IUCr journal articles have traditionally transferred copyright in their publication to the IUCr (which has licensed back to the authors the normal rights of academic re-use); but it has not always been clear to what extent this also included supporting documents, including structural data sets. In a similar way, the CCDC feels honour bound to redistribute individual data sets upon request, since it does not hold copyright to the data; but the CSD as a whole is a value-added research tool with its own software interface that is much more than just an unstructured collection of data sets, and the CCDC holds copyright in the Cambridge Structural Database as an entity.

In a relatively new development, crystallography laboratories and service facilities are now beginning to advertise their structural data sets on open-access web sites. Metadata describing the holdings of such data are freely disseminated through the OAI-PMH exchange protocol (or by other technical means such as RSS feeds); and the data sets themselves are accessible (subject to the agreement of the relevant rights holders) for download. It is intended that these data sets will be harvested by databases and harvested or linked to by journals. However, the assignment of intellectual property rights to the data may not always be an easy or obvious procedure.

We therefore recognise that a consensus formalism that describes such rights, or at least the licensed secondary rights to dispose of the data, will play an important role in the handling, dissemination and re-use of these data. For this reason we have been following with interest the efforts of the Science Commons movement to define machine-readable licences that will facilitate and legitimise secondary uses. Although the standard exchange formalism in crystallography (CIF) cannot support the Science Commons mechanisms directly, we undertake to develop mechanisms that express exactly those Science Commons licences that can be adopted in the crystallographic domain.

Finally, as part of the GICSI initiative, the IUCr undertakes to work with the crystallographic databases to try to ensure that individual researchers in the least developed countries can get access to the databases.

4. Crystallographic software

An important component of a science information commons, and one that is not considered in detail so far by the proponents of GICSI and other stakeholders, is the software used in the various stages of the data life cycle. The current situation in crystallography involves many commercial and proprietary, closed-source, systems that: store and encode image and other raw data; perform data reduction based on commercial equipment parameters; and solve and refine the resultant crystal structures. There are also several program systems developed by the academic community that perform many of these functions (although perhaps mostly at the solution/refinement stage); and some academic/commercial collaborations that expose only some of the computational machinery to public view. Within the corpus of software developed in the academic community, there is a variety of attitudes towards exposing and

sharing software code. Relatively few crystallographic software projects have been developed under open-source or related software licences (*e.g.* the GNU GPL), although the number is increasing. We can mention examples varying in scale and scope, such as the UK Crystallographic Computing Project 4 (*CCP4*) for protein crystallography (www.ccp4.ac.uk), the *PHENIX* software suite (http://www.phenix-online.org/), the long-established *Xtal* suite (http://xtal.sourceforge.net) or the *RasMol* molecular visualisation program (http://www.openrasmol.org/).

As with other aspects of the data/information life cycle, we recognise the right to proper intellectual property and exploitation rights of the intellectual effort invested in developing complex software systems. Furthermore, we recognise also that the quality of the results of a crystal structure determination depend very much upon the details of the algorithms used, and we encourage the development of open documentation of algorithms, increasing open-source release of software, and the development of community projects that allow cross-validation of results obtained from different packages.

The IUCr already has some technical means to assist this objective. Its Crystallographic Information Framework defines thousands of data items very precisely, so that output data from different programs can be guaranteed to have the same meaning. Individual programs such as *PLATON* have been developed that give independent assessments of the internal consistency of results generated by other packages. Continuing developments within CIF are working towards machine-readable declaration of specific algorithms expressing and validating data relationships. We also note that the PDB (see above) makes freely available the software it develops for the management and validation of structural data in CIF and other standard formats; likewise, the CCDC makes freely available the *enCIFer* program for CIF format validation and editing, and the *Mercury* package for structure visualisation.

Nevertheless, the inner workings of some programs are still not fully exposed to public scrutiny, and we encourage more open documentation of proprietary algorithms, together with escrow deposit of fully documented non-disclosable code, to provide some measure of long-term security of access to the specific calculations and manipulations that underlie the derived data informing scientific discussion.

Data description standards

A prerequisite for unambiguous communication of data in the electronic era, for interoperability and for long-term storage of data and information, is adherence to agreed standards of information representation. The IUCr has invested considerable time and resources in the development of such a standard for crystallographic data, the Crystallographic Information Framework (CIF). CIF is designed to complement other exchange mechanisms – it transforms trivially to relational database representations, and efficiently to a variety of XML and other syntactic representations. There are implementations of the CIF ontology in bioinformatics middleware (a CORBA applications programming interface for macromolecular structure has been approved by the Object Management Group) and in chemical structure markup (Chemical Markup Language (CML) of Murray-Rust & Rzepa). Efforts are under way to harmonize image representations in CIF format (imgCIF) with other image representation standards. Likewise, the CIF ontology will inform ongoing efforts to define metadata for the description and characterization of crystal structure data sets in open-archive repositories.

All aspects of CIF as a data exchange standard are fully documented on the IUCr's web site and its publications.

The IUCr also seeks to work with existing standards in other areas of information representation. Journal articles are marked up with Standard Generalized Markup Language (SGML) conforming to the ISO 12083 standard for scientific articles; web content conforms to W3C standards and recommendations as far as possible.

We believe that it is important to continue to promote the adoption of common standards and to work towards interoperability of new standards in order to realise most effectively the information-exchange requirements of a global information commons.

ANNEX: Co-operative sharing of research materials and tools

The GICSI prospectus includes among its goals 'cooperative sharing of research materials and tools among researchers'. This is not an area we have considered in detail in putting together this paper, which addresses mainly the data and literature aspects of the information commons. However, research materials and tools legitimately fall under the aspirations of a scientific commons, in the wider sense of the term ('any sets of resources that a community recognizes as being accessible to any member of that community' – *Wikipedia*) being promoted by GICSI.

In the domain of structural sciences, developing countries could form a very important constituency of scientists who could harness the opportunity to use synchrotron radiation and neutron facility data that did not lead to publication by the users who originally proposed the work (*cf.* the rule adopted by ESA in microgravity crystallization flights that data from an experiment should be released into the public domain after 24 months if no direct publication is in train). There are also opportunities to provide more widespread access through 'telepresence' and internet-based instrument control to synchrotron-radiation and neutron data collection facilities. This would easily allow greater direct participation and use of expensive research equipment to scientists in developing countries, but would obviously involve careful criteria to ensure that the access granted in this way was considered 'equitable'.

While the development of these ideas fall outside the scope of this paper, we recommend that the IUCr consider whether it has a formal part to play in encouraging this component of the proposed scientific commons through its Commissions on Synchrotron Radiation and Neutron Scattering.

APPENDIX: IUCr response to the *Global Information Commons for Science Initiative*

The Global Information Commons for Science Initiative (GICSI) is a multi-stakeholder initiative arising from the second phase of the World Summit on the Information Society in Tunis in November 2005, with an overall goal of accelerating the development and scaling up of open scientific data and information resources on a global basis.

The International Union of Crystallography (IUCr) acknowledges the importance to the scientific endeavour of full and equitable access to scientific information. It therefore warmly endorses the goals of the Initiative, and supports the proposed implementation, subject to a sound business model and effective management by its participating stakeholders.

The IUCr recognises the Initiative's strong emphasis on the principle of open availability as an ideal medium of dissemination for publicly funded research findings and data. However, the IUCr itself is one among many organizations that provide high-quality value-added publishing and data services, historically funded through subscriptions, direct sales and other commercially mediated mechanisms. Although the IUCr is experimenting with various routes towards open availability, it is not clear that this can always be achieved, or that it is invariably the best solution.

The IUCr's continuing and proposed contributions to fostering a global information commons through its publishing and data activities include:

1. Continuing the practice of supplying free of charge the primary research data, machine-readable sets of atomic coordinates and structure factors, and other supplementary documents supporting the primary research results reported in its scientific journals.
2. Continuing the practice of providing open access to education papers, software application papers and Commission reports published in those journals.
3. Continuing to provide means such as the IUCr Journal Grants Fund and involvement with a commercial partner in INASP programmes to maximise the availability of low-cost subscriptions to developing nations.
4. Continuing its current hybrid open-access/subscription model to provide authors with an opportunity to pay production costs for their articles and thereby allow them to be accessed freely.
5. Continuing studies and projects towards the introduction of full open-access publication under an appropriate economically sustainable model.
6. Investigating the possibility of commissioning and implementing high-quality open-access reference resources through collaborative Web tools.
7. Formalising a machine-readable set of licensing terms stipulating re-use and redistribution of primary and derived data sets modelled on Science Commons paradigms.
8. Continuing to offer the *checkCIF* service as an objective tool for assessing the quality of crystal structural data, and encouraging cross-checking of data-based results through a diversity of publicly documented programs.
9. Working with the crystallographic databases to try to ensure that individual researchers in the least developed countries can get access to the databases.
10. Promoting the development of open-source collaborative computational projects in crystallography.
11. Promoting the open documentation of algorithms and computational techniques employed by software in reducing, analysing and transforming crystallographic data, even for closed-source applications.
12. Encouraging sharing of research material and tools across the international scientific community.

The IUCr proposes to work in support of the organizations on which it is directly represented (ICSU, CODATA, ICSTI) to realise the goals of GICSI in a manner which is sustainable within the IUCr's approach to publishing, data collection, organisation, curation and archiving.