

PubMed Central: A phenomenal advance in electronic publishing

C. Anil Kumar and V. Vishnu

PubMed Central (PMC) the new web-based repository being developed by the National Centre for Biotechnology Information (NCBI) at the US National Library of Medicine (NLM) and funded by the US National Institutes of Health (NIH), will archive, organize and distribute peer-reviewed reports from biomedical journals (<http://pubmedcentral.nih.gov>). The distinguishing characteristic of PMC is that it offers the full text of peer-reviewed articles free to users, by acting the role of a world-class library. BioMed Central is the main supplier of electronic journals to PMC¹.

Many publishers are already making their journals freely available on the web within a year of publication. Even though this is the scenario in this digital age, the launching of the PMC was aimed at certain other objectives. Free access to the full text of scientific literature would be a phenomenal advance in scientific electronic publishing and certainly will benefit science and clinical medicine and patient care. Data from diverse sources are stored in a common format in a single repository and presented to the user in a uniform style, while still clearly maintaining the identity of each journal¹. It will disseminate research better and more cheaply than it is done now. A time will come when libraries could stop subscribing to journals and thereby releasing money back to researchers, if this free access to information would continue.

PMC follows in the footsteps of other highly successful and useful services that NCBI has developed like GenBank, the genetic sequence data repository, and PubMed, the database of citations and abstracts to biomedical and other life sciences journal literature. For PMC to succeed, it should work in unison with PubMed, MEDLINE, BioMed Central (BMC), etc.

PubMed, MEDLINE, BMC and PMC

PubMed was developed by the NCBI at the NLM, located at the NIH. It was developed in conjunction with publishers of biomedical literature as a search tool

for accessing literature citations and linking to full-text journals at web sites of participating publishers. PubMed provides access to over 11 million citations from MEDLINE and other life sciences journals. If the publisher has a web site that offers full text of its journals, PubMed provides links to that site as well as sites to other biological data, sequence centres, etc. User registration, a subscription fee or some other type of fee may be required to access the full-text of articles in some journals. PubMed also provides access and links to the integrated molecular biology databases included in NCBI's Entrez retrieval system.

MEDLINE is NLM's premier bibliographic database covering the fields of medicine, nursing, dentistry, veterinary medicine, health care system and pre-clinical sciences. MEDLINE contains bibliographic citations and author abstracts from more than 4000 biomedical journals. The file contains over 11 million citations dating back to 1966. Coverage is worldwide, but most records are from English language sources or have English abstracts.

BMC is a commercial publisher of online biomedical journals, which provides free access to articles in its site. BMC also deposits its articles in PMC when they are published.

PMC, operated by NCBI, a division of the US NLM, is an electronic archive for full-text journal articles offering unrestricted access to its contents. Every full-text article in PMC has a corresponding entry in PubMed. PMC, by storing data from diverse sources in a single repository with a common format, makes the data more accessible and easier to use and opens the door to greater integration with related resources. PMC is not a publisher. Participation by publishers in PMC is voluntary and the participating journals must meet certain editorial standards.

Participation in PMC

Publishers will have to sign participation agreement with PMC to set up a schedule for the regular addition of journal issues

to PMC. A participating journal is expected to include all its peer-reviewed primary research articles in PMC. A publisher, at its discretion, may also deposit other contents such as review articles, essays and editorials. Review journals and similar publications that have no primary research articles, are also invited to include their contents in PMC. However, primary research papers without peer review are not accepted. PMC and the journal agree on a presentation style for the journal in PMC. The journal can supply complete data for all back issues for inclusion in PMC. Contribution to PMC must meet certain editorial standards and PMC will accept material from any life sciences journal, including plant sciences, which meets one of the following criteria:

- (1) It is covered by a major abstracting or indexing service such as MEDLINE, Agricola, Biosis, Chemical Abstracts, EMBASE, PsycINFO or Science Citation Index; or
- (2) It has at least three editorial board members who are currently principal investigators on research grants from major funding agencies in the US or abroad.

PMC does not accept material from individual authors.

For a journal that is not yet covered by any of the major indexing services PMC needs brief letters from three editors who are listed on the journal masthead and who meet the PMC requirement of being a principal investigator on a research grant from a major funding agency. Each letter should confirm the writer's receipt of such funding (identify the granting agency and grant number or similar specifics) and his or her editorial involvement with the journal.

Submission of data in PMC

Required data format

A journal must supply the full text of articles to PMC in a SGML (Standard Generalized Markup Language) or XML (eXtensible Markup Language) format,

which conforms to any established DTD (Document-Type Definition) for journal articles. Figures should be supplied as high-resolution (TIFF or encapsulated postscript) images. Supplementary material in the form of video, audio or data files may also be submitted. PMC does not accept HTML data. A new PMC XML DTD will be made public soon, for those wishing to submit their data in that form. A journal does not have to use this DTD, if it already generates data conforming to another DTD. Other DTDs supported by PMC include Blackwell, BMC, Keton, OUP and Ovid.

A journal may deposit its material in PMC and make it available for public release as soon as it is published, or it may delay release for a specified period after initial publication. A journal is guaranteed access to a copy of its deposited data, at no cost, at any time in the future. PMC does not claim copyright on any material deposited in the archive and it rests with the journal publisher or with the individual authors, whichever is applicable. Now, full text of articles pub-

lished in the *British Medical Journal* (which was the first journal to join PMC) are also available free from PMC on day of publication¹.

Evaluation of data

PMC reviews a sample of data from the journal to ensure that a publisher can routinely supply journal data of sufficient quality to generate complete and accurate articles online. PMC normally requires a minimum sample of three issues or 50 articles for evaluation. The journal should provide one issue from the current year and one issue from each of the preceding years. If a journal expects to start with less than three years' worth of issues, it should provide three sample issues from the available range of issues. If the minimum number of issues defined above comprises less than 50 articles, PMC will require additional issues to make up the 50-article minimum. If a journal is brand new and has published fewer than three issues or 50 articles, PMC still needs to have a reasonable

sample, something in the vicinity of 10 articles.

Conclusion

PMC is a project masterminded by the US NIH to make the results of original research in the life sciences freely available to everyone via the Internet. Material is presented to the user in a uniform format after receiving data from various journals. Participation in PMC will improve the quality of the journal's electronic archival record.

1. Delamothe, T. and Smith, R., *Br. Med. J.*, 2000, **321**, 192; 2001, **322**, 1-2.

C. Anil Kumar, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011, India; and **V. Vishnu***, Regional Cancer Centre, Thiruvananthapuram 695 011, India. *For correspondence (e-mail: rccvm@md2.vsnl.net.in).

Techno Market Survey Reports from TIFAC

The Technology Information, Forecasting and Assessment Council (TIFAC) is a Registered Society of the Government of India, Department of Science and Technology and has been on the scene since 1988. It has published over 250 Techno Market Survey Reports so far. These are specialized reports and cover a wide variety of topics of economic relevance to India. According to TIFAC, they cover areas 'from agriculture to advanced sensors and from materials to biotechnology'.

V. S. Ramamurthy (Secretary, Department of Science and Technology) said that the Techno Market Survey Reports are neither technical reports nor market surveys but are a combination of both. The reports are concise, consolidated and unbiased and give to the end user well-researched information on options in various technologies.

For example, TIFAC has made a Techno Market survey of 'Traffic Management Streamlining Technologies' that covers the whole gamut of status and profile of Traffic in India with special reference to Kanpur and Delhi. The Traffic report studies viability of Traffic Management Technology and future investments forecast in this area for some sample cities, and recommends an action plan. This is just one example of the 250 Reports available as of now with several more in the pipeline. The forthcoming Reports are on topics such as Biochips, Transgenic Animal Models, Human Body Implants, Fuel Cells and Technologies for retrofitting existing structures to make them earthquake resistant.

Among the other Techno Market Survey Reports available, mention must be made of strategy reports on Non-ferrous metals. These pertain to Aluminium,

Copper, Magnesium and Titanium. TIFAC reports on Refractory metals, Rare Earths etc. are expected by the end of the year. Many organic acids that are currently imported like Itaconic, Gluconic, Glutamic, Lactic and L-Lysine can be produced using molasses as raw material (by-product of the sugar industry). The technology for production, assessment of demand, supply, import export details, etc. are available in five volumes. A comprehensive picture of Science and Technology in India in the form of several reports are also available. For more on this and other TIFAC activities log on to their website www.tifac.org.in.

Nirupa Sen, 1333 Poorvanchal Complex (Old), J.N.U. New Campus, New Delhi 110 067, India (e-mail: nirupasen@vsnl.net).