

# Archiving of Electronic Journals in Physics and Astronomy: Role of Consortia

By

**Christina Birdie**

*Indian Institute of Astrophysics  
Koramangala  
Bangalore – 560 034*

**Email:** chris@iiap.ernet.in

**Sreeharsha. B.G**

*Indian Institute of Astrophysics  
Koramangala  
Bangalore – 560 034*

**Email:** n\_s\_harsha@yahoo.co.in

**Pallavi. T.G**

*Indian Institute of Astrophysics  
Koramangala  
Bangalore – 560 034*

**Email:** pallavi\_reddy4@rediffmail.com

**Mahantesh. B. Kannur**

*Indian Institute of Astrophysics  
Koramangala  
Bangalore – 560 034*

**Email:** mahantesh\_kannur1@rediffmail.com

## ABSTRACT

Electronic journals unfold their most valuable features when they are inter-linked with other documents, archives and databases. These electronic journals need an archiving system that is based on a global concept. Archiving ensures “continued access indefinitely into the future of records stored in digital electronic form”. Archiving system has to be flexible to adopt the new technological generation. The rapid pace at which technology becomes obsolete makes archiving an extremely expensive undertaking. The librarians are more concerned about the archiving of the electronic journals themselves, or getting access to the archives established by the publishers or the aggregators. This paper focuses on the issues related to the archiving of electronic journals in Physics and Astronomy by the publishers, aggregators and the consortia.

**KEY WORDS:** Archiving of e-journals, Consortia, Physics & Astronomy archives.

## 0. INTRODUCTION

We are in the digital information age, where information is produced, stored and disseminated in the form of bits and bytes. The traditional print journals are now produced in electronic format, often with a different format suitable for various environments. Internet plays an important role in making these electronic journals available to the end users on their desktop computers. There are numerous advantages in publishing journals in the electronic format, both to the publisher and to the reader. While Internet offers lower cost of distribution of electronic journals, the system of reproduction makes the pricing of these journals expensive. The archiving of electronic journals for long-term access will prove to be very expensive if it is not planned properly.

As the popularity of electronic journals is increasing, the librarians as well as the publishers are concerned about archival of electronic journals.

## **1. ARCHIVING**

Archiving of electronic journals assumes greater importance due to the large number of e-journals on the net. Most of the research libraries and publishers are concerned with the archiving of e-journals so that the information can be made available for long-term access to the scholarly community. The main intention of archiving is to preserve the content of e-journals in zeroes and one's, irrespective of the form/kind of display format so that the content will be available for the future usage. Though there is a distinction between archiving, and preservation, these two terms are two faces of a coin. Margaret Phillips [1] in her article has aptly described the two terms as “ Archiving involves the identification, collection and provision of access to digital records in their current format, whilst preservation refers to the management of electronic materials so that they remain accessible as technology changes”.

### **The need for archives**

Though scientific journals are in existence since 15<sup>th</sup> century in the paper form, and well preserved in most libraries, we are now more worried about their archiving because of their new electronic format. When we subscribe to a printed journal, we get a copy in the physical form and we preserve that issue for the future reference. i.e we are archiving that particular journal, whereas when we pay for an e-journal, we are “licensed to access” that journal [2]. In print version, even if we unsubscribe to the journal, we will be having the back volume, whilst this is not in the case of e-journals where we may or may not have rights to access them. The uncertainty of the publisher's policies and also the increased user's requirements has compelled the libraries to invest in archives. Several challenges have to be faced while e-archiving of these journals in terms of the technology change and the resources required for archiving.

## **2. WHO WILL PLAY THE KEY ROLE?**

Till recently, it was the librarian who took the responsibility of archiving of printed journals. Now there is a debate as to, who should play the key role in archiving the electronic journals. Whether it should be the librarian or a Consortium of libraries or the publisher or a third Party Aggregator?

### **2.1 Librarian**

Archiving is not a new concept to the librarians as they have been doing this since the libraries started acquiring printed material. But now they have to deal with the electronic version. The question is “Are librarians competent enough to archive these materials?” because its not an easy task to archive e-journals. The process of archiving requires preserving, storing, organising and providing effective search facilities to the archived data. Hence librarians are expected to have the knowledge of creating and maintaining the databases in addition to the knowledge of computer languages and different features of the software used for archiving purposes when purchased from a vendor. Apart from these there are certain issues which one should consider while planning for archiving, such as;

### **2.2 Cost factor**

Cost is an important factor for creation and maintenance of the database. As technology improves, the cost is bound to rise in terms of upgraded systems, cost of software and its upgradation etc. With the limited and meager budgets, librarians may not be able to afford it [3]

### 2.3 Time factor

Archiving is a time consuming and a long-term process. It requires an additional time for training of the library staff and maintenance of the database.

### 2.4 Technology factor

Computer technology changes every 3 – 4 years making the earlier version obsolete. As and when the technology changes, the librarians need to update their system too. It may not be possible for librarian to take up archiving without additional help.

## 3. CONSORTIA OF LIBRARIES

Libraries in a consortia environment have the advantage of creating a more complete e–archives compared to what the individual libraries can establish. The archive model, which applies to the electronic world, is similar to the cooperative arrangements and inter-library loan of paper materials. Joint venture will facilitate the sharing of expenses incurred, manpower and technical knowledge required for creating and maintaining an archive of electronic journals accessed by the members of a consortium. The Astronomy archive called Astrophysics Data System Abstract Service (ADS) is tightly inter connected with the major journals of astronomy. The contributors to the ADS are publishers of major astronomy journals' astronomers, librarians, and computer professionals. It is a true model of archiving by a consortium, where the members belong to different background. The details of this archiving facility are discussed in Table 1 in this paper.

## 4. PUBLISHERS

The content of any journal is the property of the Publishers. They are the suppliers of e-journals directly to the end users or via third party aggregators. In most of the cases they have the rights to archive and one may consider the publishers as the appropriate party to take up archiving. If they do take up archiving they may have to address the following issues:

**Back issues:** There are very few publishers who are willing to archive the journal from Vol. 1 (1). Others may archive for a specific period or certain volume.

**Cost factor:** Cost is another important point to be considered. Some publishers may charge 10% over the current subscription while others may charge more. Is there a pricing rationale? This is still uncertain since the archiving by publishers keep changing as new technology and policy emerges.

**Cease of Publication of the Journal:** In certain cases where the publication of a particular journal ceases, does one still have the access to the archives?

**Publishers cease:** What would happen if the publishers himself ceased to exist?

**Publication merges:** In case where two or more publishers merge, what would be the policy regarding archives?

**Journal splits:** Suppose if the journal itself splits (As in the case of Journal of the Optical Society of America which split into two to form Journal of the Optical Society of America – A and Journal of the Optical Society of America – B)

**Publication splits:** Who will maintain the archive if the publisher splits?

### **Publisher's Perspective towards Archiving Policy?**

As we have discussed earlier, it is still uncertain, about who should take up archiving. Majority agrees that publisher's should take up archiving. In general it is felt that libraries can trust big publishers like Elsevier, Kluwer, Springer etc.

We present here the policies relating to archiving of some major publishers in Physics and Astrophysics, and some important features of these archives.

Some of publishers whom we have chosen are:

- ◆ Kluwer Academic Publishers
- ◆ Elsevier Science
- ◆ American Institute of Physics
- ◆ Institute of Physics

### **Kluwer Academic Publishers**

Kluwer Publishes over 750 journals and all of them are available online. Back issues of all these titles for the years from 2000 back to 1997 are available as archival access. Over 25 titles in "Full-Text" XML in 2001 and also a powerful Boolean search over all SGML headers and abstracts are available.

**Kluwer Online Archiving:** Perpetual access is granted to holdings for atleast 5 years via Kluwer's server. Access is online via the Kluwer server in the Netherlands via local storage or via gateway or hosting partners.

Libraries may keep one electronic copy for archival / backup purposes. Kluwer's full text of all journals is located on OCLC mirror site. Back volumes include free and free cross-access to joint holdings [4]

## **5. ELSEVIER SCIENCE**

Elsevier Science has taken formal responsibility for digital archiving of journals from 1999, and now it is part of all Science Direct licenses. Science Direct is a world's largest Scientific, Technical and Medical (STM) database, which cover over 1700 journals back to Vol. 1(1).

Elsevier has signed an agreement with National Library of the Netherlands; it is their first external official archive. They have deposited one copy of all the journal files on Science Direct, including all the back files (as they retrodigitize back to Vol.1, No.1) they have accepted to take up the responsibility for the permanent retention and technical migration of data. [5]

Future plans of Elsevier are,

- to establish 1-2 more official archives (similar to National Library of Netherlands). They expect this to be done with another National Library.
- in addition, library customers who subscribe for Science Direct onsite will receive copies of electronic files for local hosting. These sites are not official archives, but they provide a level of backup and reassurance to the library community that multiple copies of the files exist. Elsevier is planning to formalize archiving arrangements with some of these parties. [6]

## 6. INSTITUTE OF PHYSICS

Institute of Physics is digitizing its entire journal archive during 2002; the Institutes journals date back to 1874 with publication of the Proceedings of the Physical Society of London (1874-1967). This massive project involves the digitization of more than 500 Volume-years of journals. The pre-1991, full-text archive will be made freely available during 2002 via electronic journals platform and the archive will be completed in three phases:

<b>Phase I</b>	:	Journal of Physics series, 1968-1990, now complete.
<b>Phase II</b>	:	1968-1990, in progress
<b>Phase III</b>	:	1874-1967, in progress

The latest news on the development of online archive is available in archive update. The publication of the archive will not effect the subscription in anyway. In addition to articles from the current year, institutional subscribers will still be entitled to access a rolling 10-year electronic archive for a particular journal, as at present.

Institute of Physics provides services like Browsing, Navigating, Searching, Reference Linking by hyper cite linking technology. This enables you to link from article references to a number of resources including the INSPEC, COMPENDEX and Page One from Elsevier Engineering Information.

All titles published by Institute of Physics and most other major STM publishers, and papers are held on the Los Alamos server. [7]

As a learned society (not-for-profit) publisher, the Institute of Physics is committed to promoting physics and to serving physicists, worldwide. This latest initiative supports the prime objective and Royal Charter of Institute of Physics to disseminate information in this field.

### American Institute of Physics

As per American Institute of Physics (AIP) Electronic Information archiving and use policies adopted by AIP Executive committee on 12 June 1998

AIP will retain in an archive all electronic information published by the American Institute of Physics. The archive will include all AIP primary publications in its online journal service (OJS) and the supplementary publications material contained in AIPs Electronic Physics Auxiliary Publications Service (E-PAPS).

AIP will hold a primary source material archive and be responsible for the periodic refresh of this archive and its replication to additional archives. The primary source material will not be used for active delivery. At least one complete archive will be maintained outside AIP at a site separate and distant from the primary archive. The archive will be reviewed for refreshing or migration to new information formats at

appropriate intervals. Information will be migrated to new formats when current formats are in danger of becoming obsolete or unsupported, or when new formats provide substantial improvements in features with no loss of content. [8].

AIP is committed to the retention and preservation of scholarly research information published in electronic form within AIPs journals, proceedings and other publications. This policy statement defines the extent of the archiving commitment and the relevant procedures for fulfilling this commitment.

## 7. ARCHIVING THROUGH THIRD PARTY OR AGGREGATORS

One more possibility of archiving of e-journals is third party vendors archiving the E-journal for publishers and disseminating to the end users. These third party vendors collect the journals from different publishers, put them in a database and disseminate to the subscribers. The success of archiving is in the cooperation among the publishers, third party vendors and the librarians. Some of them are JSTOR (<http://www.jstor.org/>), Ovid Technologies (<http://www.ovid.com>), HighWire Press (<http://highwire.stanford.edu>). [9]

**JSTOR:** Journal Storage (JSTOR) was originally conceived by William.G. Bowell, President of Andrew Mellon Foundation. The basic idea of JSTOR was to convert back issues of paper journals into electronic formats. JSTOR was established as an independent non-profit organisation in Aug.1995. JSTOR's mission is to help the scholarly community take advantage of advances in information technology. Initially electronic access to back files of 10 journals in two core fields – Economic and History was started. This pilot project was sponsored by the Mellon foundation. JSTOR's second initiative is to establish convenient links between the archived articles in JSTOR and more recent articles housed on participating publisher's site. JSTOR's future plan is to develop collections in Art History, Education, Law and Music. This second initiative is also funded by Mellon foundation. JSTOR covers over 345 journals, out of these 275 are available online and number of participating publishers is 180. [10].

**Ovid Technologies:** Ovid Technology is one of the major providers of STM journals. All the journals of Ovid can be accessible from Journals@Ovid. It offers perpetual access to its subscribers. Almost 98% of the journals offered through Journals@Ovid has this facility. The archive can be accessed through normal search facility provided and they assure cent percent access to the full text links. The libraries when they unsubscribe will get a final monthly database updated, in the form of a CD-ROM and this will constitute the archive. When Ovid finds a new technology, which can be adopted, for the betterment of service, without any loss of content, it will adapt the technology and certain fee is collected from the subscriber. All the copyright laws will be applicable to the archival data also [11].

**HighWire Press:** Stanford University Libraries' High Wire Press is one, which assures of providing access to 170+ journals. The main policy of High Wire Press is "to provide access in perpetuity to the subscribers".

High Wire provides access to the journals through the journals' web site and assures all the back issues. It provides free access to many journals older than certain period of time and this can be accessed by anyone freely. The High Wire servers are maintained at Stanford University and it will be updated as and when the technology is changed. Funded by National Science Foundation and in Co-operation with Sun Microsystems, High Wire has developed the LOCKSS project. LOCKSS – Lots Of Copies Keep Stuff Safe is a voluntary approach for archiving the journal [12].

## 8. PROLA AND ADS

PROLA and ADS are archives of databases, which cover Physics and Astrophysics exclusively. These two online archive databases cover almost all branches of Physics and Astronomy and they are extensively used by the physicists and astronomers.

**PROLA:** The Physical Review Online Archives covers mainly Physics. Initiated by the American Institute of Physics, PROLA proves to be a very useful Online Archive database for Physicists. However, PROLA full text is available on annual subscription only.

**ADS:** NASA Astrophysics Data System is an online archive developed by NASA. This online archive comprises of four databases. This is suppose to be the biggest online archive related to Astronomy and Astrophysics. ADS is a non-profit archival system, which provides free service and caters to all the needs of a large number of physicists, and astronomers mainly.

The Table 1 covers some of the important features of PROLA and ADS along with their future plans.

*Table 1: Features of PROLA and ADS Archives*

S.N.	Features	PROLA (Physical Review Online Archives)	ADS (Astrophysics Data System)
01	Origin	PROLA was developed, deployed and tested by American Physical Society, the Computer and Applications Group at Los Alamos National Laboratory. PROLA is an e-journal archive system and PROLA covers from 1894 to the present. [13]	The final report of the Astrophysics Data System Dated March 1988, characterized the data environment of the astrophysics community and defined for the future an “Architecture for a data system which will serve the astrophysics community in multi-spectral research through the decade of the 1990’s. [14]  The Primary source or initial data set was from NASA’s Scientific and Technical Information Group (STI), which provides abstracts from 1975 through the middle of 1995. The STI data are estimated to be about 98% complete. The data from journal publishers (mostly 1995 and beyond) should be 100% complete.
02.	Coverage	PROLA covers Physical Review A, B, C, D, E, Physical Review Letters, Physical Review ST AB, Physical Review (Series 1), Reviews in Modern Physics.	ADS covers hundreds of publications, colloquia, symposia, proceedings, NASA reports. It also covers a large number of American Astronomical Journals, Non – American Journals and also general interest journals. [15]
03.	Cost	PROLA full-text article and searching are accessible by subscription only. However browsing tables of contents and abstracts will remain free. Cost of PROLA’s annual subscription is US\$ 470.	ADS is a free service

04.	No.of Articles accessible	Nearly 130,000 articles are accessible through PROLA	ADS provide access to four databases containing more than 2.9 million abstracts.
05.	Basic Services / Functions	Finding things on website with over a million pages is, of course the critical functionality that determines success of an electronic archive, PROLA approaches this problem using four types of actions: Browsing, Searching, Retrieving and Navigating.	Similar to PROLA ADS also provides some of services like Browsing, Searching and Retrieving.  In order to identify the literature in the databases certain unique bibliographic codes are created using standard bibliographic format .The bibliographic code is a 19-digit identifier, which describes the journal article.
06.	Mirror Sites	PROLA's mirror sites are situated in <ul style="list-style-type: none"> <li>• Main PROLA Server (New York City, New York)</li> <li>• Ridge PROLA server (Ridge, Long Island, New York)</li> </ul> Cornell University (Ithaca New York) [16]	ADS bibliographic services are now available from several sites world wide <ul style="list-style-type: none"> <li>• Harvard-Smithsonian Centre for Astrophysics, Cambridge, USA.</li> <li>• Centre de Donnees Astronomiques de Strasbourg, France.</li> <li>• European Southern Observatory, Garching, Germany.</li> <li>• National Astronomical Observatory, Tokyo, Japan.</li> <li>• Pontificia Universidad Catolica , Santiago, Chile.</li> <li>• University of Nottingham, United Kingdom.</li> <li>• Beijing Astronomical Observatory, Beijing, China</li> <li>• Inter-University Centre for Astronomy and Astrophysics, Pune, India</li> <li>• Institute of Astronomy of the Russian Academy of Sciences, Moscow, Russia</li> <li>• Observatorio Nacional, Rio de Janeiro, Brazil</li> <li>• Observatorio Astronomico de Cordoba, Argentina</li> <li>• Korea Astronomy Observatory</li> </ul>
07.	Database	PROLA is a single database online archive but links to the CAS (Chemical Abstract Service) and to the NASA Astrophysics Data System (A.D.S) have been added.	ADS contains data from four databases are <ul style="list-style-type: none"> <li>• Astronomy and Astrophysics</li> <li>• Instrumentation</li> <li>• Physics and Geophysics</li> <li>• Astrophysics Preprints</li> </ul>
08.	Repository	PROLA has an agreement with Library of Congress establishing a repository of all PROLA material	_____

09.	Migration of Data	PROLA also has begun year-by – year migration of recent material. In order to avail new material as well as to offer enhanced functionality, the web server was redesigned	ADS was rebuilt, one major part of the rewrite was to use XML for sorting internal data. This provides flexibility-unified flexibility that was not available before to merge data from various sources into a unified framework. ADS has began migration of data from 1999. [17]
10.	Future Plans	-----	ADS is to provide access to the full text of the historical Astronomical Literature. Another important parts of the literature are observatory publications. ADS is collaborating with the Wolback library at the Harvard-Smithsonian Centre for Astrophysics in a project to microfilm the historical observatory publications. Currently ADS has over 130,000 pages from 30 observatory publications online. [18]

## 9. FUTURE OF ARCHIVING

While the access to archives of e-journals remain as an important requirement for the libraries, the responsibility of creating archives seem to be still at experimental stage in a developing country like India. Even few libraries who have access to e-journals, it will be a burden for them to access the archives established by the publishers/aggregators, which has an additional price tag. Alternatively, can these libraries create their own archives? To some extent it will be a practical solution, provided the individual libraries manage within their allocated library budget, manpower and the required technology. If a consortium of libraries decides to take the responsibility of archiving their journals, it will be an additional advantage in terms of the increased contents for the archive, the shared expenses and also the manpower.

Though the publishers and the aggregators are the ideal group of people who can create archives for the long-term access to e-journals, the libraries and the consortia of libraries can also think in terms of creating their own archives, thus enhancing the awareness and skills of the library professionals, required for archiving purposes.

## REFERENCES

- [01] Phillips, M.E. Ensuring Long – Term Access to Online Publications, JEP - The Journal of Electronic Publishing 4(4) 1999. <<http://www.press.umich.edu/jep/04-04/philips.html>>
- [02] Scott, John.S., Archiving the Online journals, ICSTI forum .No.26,1997. <[www.icsti.org/forum/26/#journals](http://www.icsti.org/forum/26/#journals)>
- [03] Edwards, J., Electronic Journals: problems and panacea? Ariadne – The web version, 15<sup>th</sup> July, 1997. <<http://www.ariadne.ac.uk/issue10/journals>>
- [04] Utlej, R. Kluwer Online 2001. <<http://www.hvar.is/swetsnavigator/kluwer.ppt>>

- [05] Hunter, K., Digital Archiving, 2000. <[http://www.kb.se/bibsam/refgrupper/bref/2000/k\\_hunter.ppt](http://www.kb.se/bibsam/refgrupper/bref/2000/k_hunter.ppt)>
- [06] Private communication with Karen Hunter
- [07] IOP Journal News Update, updated on 4 Nov. 2002. <[http://www.iop.org/EJ/S/0/26798/PdNFviheatagifzfnfDM19Q/help/-topic=freearchive/ejs\\_extra](http://www.iop.org/EJ/S/0/26798/PdNFviheatagifzfnfDM19Q/help/-topic=freearchive/ejs_extra)>
- [08] Archiving and Use of AIP Electronic information, 12 June, 1998 <<http://www.aip.org/journals/archive/arch&use.html>>
- [09] Chaudhary, A.S. and Nanyang, W.Z., Impact of Electronic Journals on Archival Functions of Academic Libraries: A Singaporean Perspective, Paper presented at International Conference on New Missions of Academic Libraries in the 21<sup>st</sup> Century, Beijing, China, Oct 25 – 28, 1998 <[http://library.brandeis.edu/beijing\\_conference/abduschaudhry.doc](http://library.brandeis.edu/beijing_conference/abduschaudhry.doc)>
- [10] < <http://www.jstor.org/about/background.html>>
- [11] <[http://www.ovid.com/support/archive\\_policy.cfm](http://www.ovid.com/support/archive_policy.cfm)>
- [12] Robinson, J. Highwire Press ensures that online Publications don't get lost in cyber space, Stanford News, 2 April 2000. <<http://www.stanford.edu/dept/news/pr/00/000204highwire.html>>
- [13] Thomas, Timothy, Physical Review Online Archives (PROLA): An image Archive for the journal physical review. D-lib magazine June 1998. <<http://www.dlib.org/dlib/june98/06thomas.html>>
- [14] Good, J.C. Overview of the astrophysics Data System (ADS) in Astronomical Society of the Pacific Conference Series Vol.25 pp.35-43, 1992. <[http://adsdoc.harvard.edu/abs\\_doc/help\\_pages/overview.html#use](http://adsdoc.harvard.edu/abs_doc/help_pages/overview.html#use)>
- [15] <<http://prola.aps.org/info/about.html>>
- [16] Eichhorn, G., Accomazzi, A...et al .XML in the ADS. Astronomical Society of Pacific Conference Series Vol.216 pp.255-258.2000.
- [17] Kurtz, M.J... et al., The NASA Astrophysics Data System: Overview Astronomy and Astrophysics Supplement Series.143 (1) pp. 41-59, 2002.

## **BIOGRAPHY OF THE AUTHORS**



**Christina Birdie** is a Librarian working in the Indian Institute of Astrophysics Library for the last two decades. In the year 1999, she was awarded the International membership to SLA by the PAM division of SLA for her contribution towards the establishment of PAM-APF (Asia Pacific Forum). She was also honored as the International Sci-Tech Librarian for the year 1999, by the Sci-Tech division of SLA during the Annual meeting of SLA in the year 1999 at Minneapolis. She is one of the coordinator of the FORSA consortium, a group of eight astronomy libraries in India. Recently she has been elected as a Director of the Asian Chapter of Special Library Association. She has published many papers in national and international and proceedings.



**Sreeharsha. B.G.** completed his MLISc from the University of Mysore in the year 2002 and B.Sc. (Comp. Sci) from the same in 1999. He served JSS Public School as a librarian for nine months and joined Indian Institute of Astrophysics, as a Library Trainee in the month of April-2002. His area of interest includes electronic journals and its related issues, Information Management, Search engines, Semantic Web.



**Pallavi T.G.** obtained her M.L.I.Sc from Bangalore University, Bangalore in the year 2001 and B.Sc (Life Sciences) from the same in 1999. She worked for Institute for Social and Economic Change, Bangalore as Project Assistant for 5 months and joined Indian Institute of Astrophysics, as a library Trainee in April 2002. Her area of Interests includes Search Engines, Web surfing, Environmental Science, etc.



**Mahantesh B. Kannur** obtained his M.L.I.Sc from Karnatak University, Dharwad in the year 2001 and B.Sc. from the same university in 1999. He has served as Librarian in Florence College of Nursing and Physiotherapy, Bangalore from July 2001 to March 2002, and joined Indian Institute of Astrophysics, as a Library Trainee in April 2002. His area of interests includes electronic journals and it's related issues, Knowledge Management, etc.