

Preserving for Future; Technology Issues

Christina Birdie

Indian Institute of Astrophysics

Bangalore - 34

Methods of Preservation

- Preservation
 - Traditional
 - Reprography
 - Digital



IIA Archives Display

- Indian Institute of Astrophysics has more than 5000 items of archival material in the form of handwritten manuscripts, correspondence, printed documents, published Scientific papers, photographs, films, and glass plates, which are subjected to special process of traditional preservation in the last two years.
- While the originals are preserved carefully, IIA is in the process of microfilming the fragile and important documents and papers, especially the annual reports of 18th century which helps us in the process of duplication and digitization for better use without handling the originals.
- Simultaneously we are digitizing all these contents using the Minolta PS 7000, CCD Scanner to convert the analog information into digital in various formats like TIFF, GIFF, JPEG, TXT, RTF, PDF, HTML. They are stored in DVDs and dedicated Server. Those contents which are out of copyright are uploaded in full text form in our open access repository created using open source software Dspace for access in the internet.

Cultural Heritage

Cultural heritage – is defined as objects of historical, scientific, cultural, social, technological, or other value which, regardless of their physical location, are valuable to be preserved for future generations and , using the most expedient technologies must be accessible for public use. Digital cultural heritage is the part of the cultural heritage that consists of objects that are converted to digital format or created in digital format.

(UNESCO Guidelines for the Preservation of Digital Heritage, 2002; <http://unesdoc.unesco.org>)

Digital Preservation

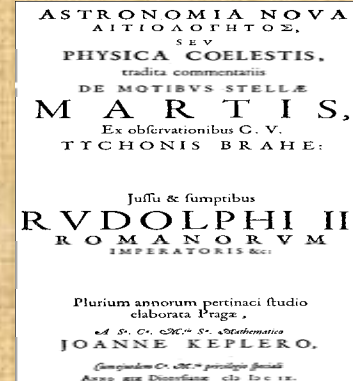
A process which embraces all actions that can be taken with the aim of ensuring the current and long-term survival and accessibility of the physical form, informal content and relevant metadata of archival records, including actions taken to influence records creators prior to acquisition or selection.

(National Archives of Australia (2001))

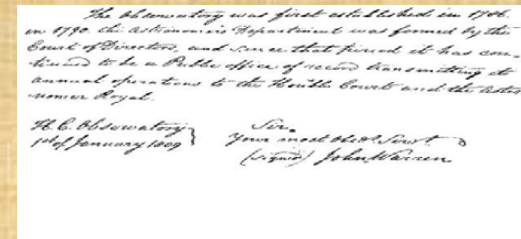
Document Heritage

- Document Heritage
 - Textual documents
(Manuscripts, posters, books, news papers, journals)
 - Non-textual documents
(Drawings, maps, prints, music, tapes etc...)
 - Audiovisual documents
(Photographs, films, discs, tapes etc...)
 - Electronic data of all types in analogue or digital form
 - Web pages, E-Mail messages

Astronomia Nova by Kepler, 1609(Oldest Book in the IIA Library)



Early Manuscript which confirmed the year of establishment of Madras Observatory as 1786



Hand-drawn sketch of Madras Observatory in 1792



Administration of Digital Conversion

- Cost factor for Digital Conversion
 - Cost of digitization depends on the magnitude of the project
 - Hardware
 - Software
 - Networked Access
 - Personnel
 - Dedicated Space
 - Long Term Access
- Technical requirements for Digital Conversion

- Initially digitization in IIA was undertaken as part of the Million Book Project (MBP). The cost involved in digitizing the collection in IIA was absorbed as part of the Library budget, which included the salary for the trainees. The project MBP, donated a Minolta PS 7000 CCD scanner to the library to carry out the digitization programme.

- We have chosen the file formats TIFF, GIFF, JPEG, TXT, RTF, PDF, HTML to save the information and stored them in DVDs. The backup is also stored in the dedicated servers.

Digitization is no doubt the issue that most fascinates and haunts preservation managers in archives and libraries at the moment. The possibilities seem limitless, the advantages are obvious, and from all sides there is pressure to exploit the new medium for preservation purposes, sometimes to the extent that funds are earmarked for digitization that might previously have been allocated to microfilming or conservation...

(<http://www.clir.org/pubs/reports/digpres.html>)

Media Type	Conversion Method	Resolution	Archive File Format	Screen Presentation Format	Print Presentation Format
Black & White Text Document	Flatbed Scanner or Digital Camera	1-bit, 600 dpi	TIFF w/CCITT Fax 4 Compression	GIF, 4-bit, 120 to 200 dpi	Acrobat (PDF), 1-bit, 300 or 600 dpi
Illustrations, Maps, Manuscripts, etc	Flatbed Scanner or Digital Camera	8-bit grayscale or 24-bit color, 200 to 300 dpi	TIFF	Multiple JPEG, 24-bit, 512x768, 1024x1536, 2048x3072, Quality Level 50	JPEG, 24-bit, 2048x3072, Quality Level 50-100
35mm Black&White & Color slide or negative	PhotoCD or Slide Scanner	24-bit, 2048x3072	PhotoCD or TIFF	Multiple JPEG, 24-bit, 512x768, 1024x1536, 2048x3072, Quality Level 50	JPEG, 24-bit, 2048x3072, Quality Level 50-100
Medium to Large Format photograph, slide, negative, transparency or color microfiche	ProPhotoCD or Drumm Scanner	24-bit, 4096x6144	PhotoCD or TIFF	Multiple JPEG, 24-bit, Quality Level 50	JPEG, 24-bit, 4096x6144, Quality Level 50-100
Black & White Microfilm	Microfilm Scanner	1-bit 600 dpi	TIFF w/ Fax 4	GIF, 4-bit, 120 to 200 dpi	PDF, 1-bit, 300 or 600 dpi
		8-bit, 300 dpi	TIFF	GIF, 8-bit 120 to 200 dpi	PDF, 8-bit, 300 or 600 dpi

Metadata

- Metadata
 - The metadata is defined as data about data or information about the digitized object.
 - Categories of Metadata
 - Administrative metadata
 - Technical metadata
 - Metadata on access ability
 - Descriptive metadata

The screenshot displays the Indian Institute of Astrophysics Repository website. The browser window title is "Indian Institute of Astrophysics : National Workshop on Preserving our Scientific Heritage - Po - Microsoft Internet Explorer". The address bar shows the URL: http://prints.iap.res.in/handle/2248/4206?mode=full&subnk_simple>Show+Full+Item+Record. The page content includes a search bar, navigation links, and a metadata table.

Search IAP Repository
Advanced Search

Browse
Communities & Collections
Issue Date
Author
Title
Subject
Statistics

Sign on to:
Receive email updates
Login IAP Repository authorized users
Edit Profile
Help
About DSpace

Indian Institute of Astrophysics >
Archival Collections >
Indian Institute of Astrophysics >
Research from IIA Archives >
Research from IIA Archives >

Please use this identifier to cite or link to this item: <http://hdl.handle.net/2248/4206>

Full metadata record

DC Field	Value	Language
dc.contributor.author	IIA Archives	-
dc.date.accessioned	2008-12-29T15:18:28Z	-
dc.date.available	2008-12-29T15:18:28Z	-
dc.date.issued	2007-12	-
dc.identifier.citation	IIA Newsletter, Vol. 12, No. 4, pp. 11	en
dc.identifier.uri	http://hdl.handle.net/2248/4206	-
dc.description	Open Access	en
dc.language.iso	en	en
dc.publisher	Indian Institute of Astrophysics	en
dc.relation.uri	http://www.iap.res.in/research/report/newsletter	en
dc.rights	© Indian Institute of Astrophysics	en
dc.subject	National Workshop - Poster	en
dc.subject	Preservation and Conservation - Workshop	en
dc.title	National Workshop on Preserving our Scientific Heritage - Poster	en
dc.type	Other	en

Appears in Collections: Research from IIA Archives

Files in This Item:

File	Description	Size	Format	
December 2007.pdf	Open Access	75.59 kB	Adobe PDF	View/Open

Show simple item record

Items in IAP Repository are protected by copyright, with all rights reserved, unless otherwise indicated.

Administration of Digital Preservation

- Cost for Digital Preservation
- Media chosen for preservation
- Standards identified for preservation
 - Open Archival Information System Reference Model (OAIS RM)
 - PREMIS (Preservation Metadata: Implementation Strategies) – a joint collaborative project of OCLC and Research Library Group
 - XML and Schema
 - Persistent Identifier
 - ISO standards for Digital Archiving: ISO 14721:2003 OAIS and RM
- Managing copyright issues
- Consortium of heterogeneous players

- At IIA we have chosen the Open Source Software Dspace to create the OA Repository which includes all the archival digitized contents.

<http://prints.iiap.res.in/handle/2248/707> .

- We are flexible and open to the process of migration of contents, as and when the technology is upgraded.

- Since IIA holds the copyright for many of the archival items, we have taken the initiative to digitize most of them and also upload them in the repository. Wherever the contents are protected by copyright we have created only metadata for those items.

Digital Preservation Options

- Preservation Methods
- Technology obsolescence protection
 - Refreshing
 - Migration
 - Hardware migration: refreshment
 - Software migration
 - Migration at point of access
 - Normalization
 - Emulation
 - Mirroring
- Analogue Backups

Initially at IIA we had stored the digitized contents in floppy disks which could be read in the earlier generation computers.

As this became obsolete we switch over to magnetic tapes for storing larger volume of data.

As the technology changed we have migrated to CDs and DVDs as the storage medium now.

Physical Storage of Plates, Films & CCDs at Kodaikanal Observatory

Plate Vault Storage



Plates, films & CCDs

Data	Plate		Film		CCD	
	From	To	From	To	From	To
Whitelight	August 1903	September 1975	September 1975	Continuing	March 2008	Continuing
Hα	April 1911	April 1977	April 1977	July 2007		
Ca+K	October 1904	February 1989	February 1989	October 2007	March 2008	Continuing

(Courtesy: S. Muneer, IIA, Bangalore)

DIGITIZERS AT KODAIKANAL OBSERVATORY

- Digitizers are installed at Kodaikanal Observatory to digitize the Photoheliograms, Ca⁺K and H α spectroheliograms of the Sun obtained from 1904 onwards. The main components of the digitizers are:

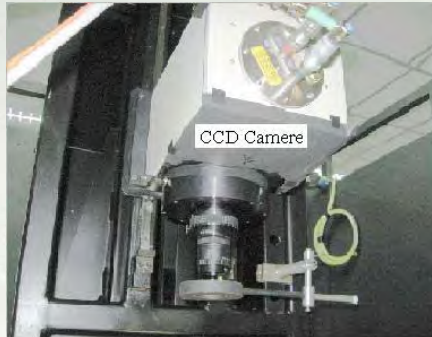
- CCD camera: Format 4K \times 4K
- Pixel size 15 micron
- Read out 16 bit
- 4 port read out @500 kHz
- Temperature of operation:
– 100°C (Cryo cooling)
- Room conditions: Temperature, Humidity and dust controlled

Number of Plates to be Digitized at Kodaikanal Observatory

- White light Images: ~ 44000
- Ca-K line Images: ~ 41000
- H-alpha Images: ~ 38000
- Prominence Images: ~ 34000

(Courtesy: Jagdev Singh et al. Poster displayed during the IIA in-house meeting 15-16 Apr. 2008)

DIGITIZERS AT KODAIKANAL OBSERVATORY



Research projects with Digitized data

- Long term variations in the photospheric and chromospheric rotation rate
- Variation of tilt angle with the solar cycle phase and its implication to the helicity and solar dynamo.
- Variation in the size of the Sun with the phase of solar cycle.
- Variation of background flux with the phase of solar cycle
- Variation in polar regions with the phase of the solar cycle
- Archival of large data base in digital form

(Courtesy: S.Muneer, IIA, Bangalore)

Solar Data Digitization & Administration

- Data is in ‘astronomy’ FITS format file system
- Individual file name is `yymmdd_hhmm.fit` where `yy` is the last two digits of the year of observation, `mm` is that of month, `dd` the day, `hh` is hour of observation and `mm` is minutes.
- Digitization project will yield 5TB of data
- Data storage is presently done in hard drives and DVDs
- Centrally managed Backups and storage systems are being planned
- Data Access will be enabled from web

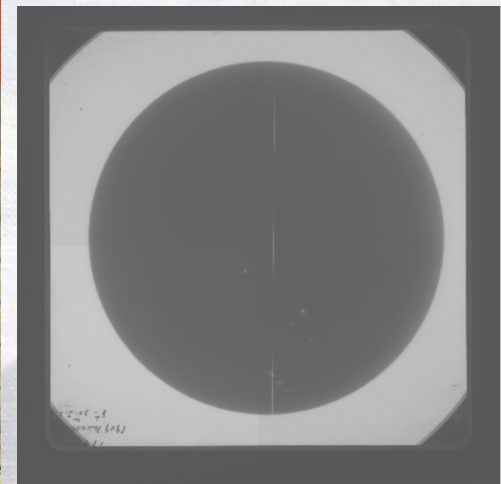
File Header

A typical file header looks like this:

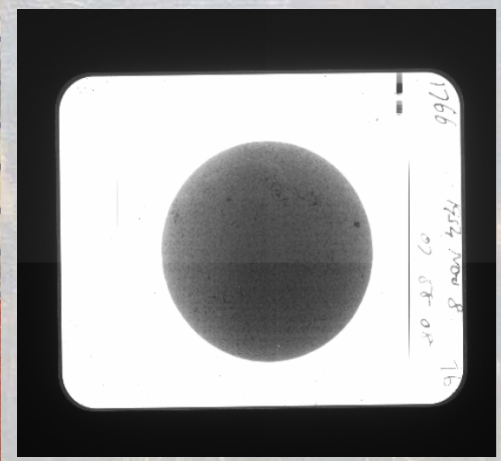
```
980605_0755.fit[4096,4096][short]:  
No bad pixels, min=0., max=0.  
(old)  
Line storage mode, physdim  
[4096,4096], length of  
user area 284 s.u.  
Created Thu 21:49:40 15-Jan-  
2009, Last modified Thu  
21:49:15 15-Jan-2009  
Pixel file "980605_0755.fit" [ok]  
FITSDATE = '2008-7-30' /  
DATE FITS FILE WAS  
CREATED  
DATE-OBS = '2008-7-30' /  
DATE OF EXPOSURE  
TIME-OBS = '15:32:45' /  
TIME OF EXPOSURE  
EXPTIME = 5.000000 /  
EXPOSURE TIME IN  
SECONDS  
INSTRUME = 'SATURN'  
/ DETECTOR USED  
SERIALNO = '06370054' /  
DETECTOR SERIAL  
NUMBER  
PIXELSIZ = 15 /  
DETECTOR PIXEL SIZE  
IN MICRONS
```

Physical Logbook Entry & Digitized Image of the Sun

Serial Number.	Date.	Time in L. S. T.			Duration.	Plate.	Slitwidth.	Settings.	Remarks.	Photographed by	Beginning.	Mean.	End.	F	E	S	M	D		
		Beginning.	Mean.	End.																
		15.10.1975									5.4.77.									
		Eulien cloudy.																		
934	Hf	04.04.05	04.04.42	04.05.18	1.13	23050	-1	3.5	Hegy	SV	02.46.50	02.49.10	02.49.29	0.39					Photo-Technical - 4.6 Thick SMA.	
935	Kf	04.04.13	04.04.27	04.10.00	2.47	15R	-1	4.6	Blue Thick	"	02.55.38	02.55.57	02.56.15	0.37					Kodak Film 2498 - 4.7 " "	
		8K became V. Thick									04.55.40	04.55.55	04.56.10	0.30					Photo-Technical - 5.0 " "	
		16.10.75									05.22.20	05.22.36	05.22.51	0.31					Zenith Astronomical 5.0 " "	
		No observation - cloudy									Summary of H α trials: -									
		17.0.75									For H α spectroheliograms, Agfa Photo-technical Panchromatisch Platte, Kodak 2485, Kodak 2498, Geipan 36 and Zenith Astronomical Bedford were tried									
		Cloudy 8K Day									Of them Agfa Photo-technical Panchromatisch is found to be identical in performance of ISRP (Sensitized for H α) These plates are to be given top most priority for H α spectro-heliograms. There are 29 djs. in stock and this will go through for a year. Next best is Kodak 2485 film (70mm). This can be used for both for H α and Kf.									
		18.10.75									All the rest were found not responding to the H α trials. They are not to be used for the purpose concerned.									
936	Kf	03.10.38	03.11.00	03.11.22	0.44	15R	-1	4.6	V. Thick	PST.										
937	Kf	03.13.11	03.13.33	03.13.55	0.44					"										
938	Kf	03.22.22	03.24.26	03.26.30	4.08	15R	+3	1.2		"										
939	Kf	03.30.41	03.32.48	03.34.44	4.13					"										
		19.10.75																		
940	Kf	01.51.45	01.52.08	01.52.31	0.44	15R	-1	4.6	Blue	PMS.										
941	Kf	02.00.40	02.01.03	02.01.26	0.46					"										
942	Kf	02.13.44	02.15.59	02.17.13	2.29	15R	+3	1.5		"										
943	Kf	02.20.56	02.21.51	02.23.36	3.30	15R	+3	1.5		"										
		20.10.75																		
944	Kf	01.38.09	01.38.33	01.38.57	2.48	23056	-1	4.5	Thin haze	PST.										
945	Kf	01.41.25	01.41.45	01.42.05	0.40	15R	-1	4.7		"										
946	Kf	01.46.37	01.48.23	01.50.33	4.20					"										
947	Kf	02.01.47	02.02.56	02.04.05	4.18					"										
948	H α	02.15.43	02.16.04	02.16.25	0.42	SRP	-	4.8		"										
949	H α	02.19.07	02.19.27	02.19.47	0.40					"										
		21.10.75																		
		Eulien lot of P.C.																		
950	H α	04.23.03	04.23.35	04.23.40	4.43	SRP	.	4.5	Blue Jap.	PMS.										
		22.10.75																		
951	Kf	03.22.05	03.22.27	03.22.44	23056	-1	4.5			PST.										
952	Kf	03.31.35	03.31.53	03.32.11	0.36	15R	-1	5.0		"										
953	Kf	03.35.30	03.37.30	03.39.30	4.00					"										
954	H α																			



Digitized Image (White Light)



Digitized Image (Ca+K line)
(Courtesy: S.Muneer, IIA, Bangalore)

Challenges to Digital Preservation

- Storage Media problem
- Hardware Obsolescence
- Software and format Obsolescence
- Web Archiving, unknown files, community data sets

Future Strategies

- Access to data tomorrow requires decisions concerning preservation today
- Adequate funding to support persistent digital access
- Well-defined roles, responsibilities of the stakeholders
- Corporate/public partnership with Government initiatives