InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems

Characterization of Case Study Validated

Case Study 26: Microvariability & Oscillations of Stars (MOST) Satellite Mission: Preservation of Space Telescope Data

Seth Dalby, UBC

Version 2, July 2005

Relevance of the Case Study to InterPARES 2

According to the project proposal, the purpose of the MOST Satellite Mission case study is to "address the issue of the long-term preservation of digital entities that do not belong in the traditional sphere of responsibility of archives and libraries and partake in the characteristics of both."[1] The MOST case study also falls within the research scope of Focus 2 and therefore enables InterPARES to meet its goals in relation to the study of scientific activities that are conducted using experiential, interactive, and dynamic computer technology.

Information about the Creator

The creator is MOST (Microvariability & Oscillations of STars) Satellite Mission, a non-profit research group founded in 2003 and operated jointly by Dynacon Inc., University of British Columbia Physics & Astronomy, UTIAS (University of Toronto Institute for Aerospace Studies) Space Flight Lab, and the University of Vienna. The purpose of the MOST space telescope is to monitor and record variations in the brightness of stars in order to study their structure, age, and evolution, to understand the effects of magnetic fields on our Sun and other stars, and to determine the nature of planets around other stars. Although MOST has no explicitly stated mission, its Web site lists "scientific goals," which may be interpreted as a mission statement of sorts.[2] The main functions of MOST researchers are the collection of data in FITS files and data reduction (analysis) and interpretation.[3]

The research body is managed as a partnership, and although it lacks an official, written organizational structure or explicitly assigned roles and responsibilities, all researchers have specific areas of expertise, and the principal scientist (an Astronomy professor) is accountable to

[1] Case Study 25 Proposal, p. 2.
[2] For further information, see http://www.astro.ubc.ca/MOST/overview.html.
[3] For an exhaustive description of research activities, refer to *MOST Satellite Mission – Model and Definitions* at http://interpares.org/rws/display_file.cfm?doc=cs26_most_model(20050615).pdf.

CSA. The satellite mission is funded by the Canadian Space Agency, which will continue its funding as long as the satellite and cameras function and researchers make use of the data collected. CSA also stipulates that MOST observations are subject to a one-year proprietary period, after which they are to be made accessible to the astronomy community and to the general public. MOST research is undertaken in accordance with academic and astronomy community best practices for research, as well as with guidelines and standards applicable to CSA and Canadian scientific missions.

As related to resources, the mission involves one satellite (owned by Dynacon), which transmits data to three ground stations. The research team at UBC has a large, secured office with separate compartments and several computers for data storage. The main computer and ground storage devices are kept in a locked compartment. A satellite dish is located on the roof of the Astronomy Building at UBC.

Information about the Administration / Management Function

No administrative responsibilities have been formally articulated in writing, although each team member has some responsibilities in this area. The main activities involved in MOST research are the collection of data into FITS files and the subsequent reduction and analysis of them. Documents resulting from MOST activities include:

- Raw data files from satellite and other sources (SDS files)
- Collections and reductions of raw data as FITS files
- Texts with graphs for publication and presentations

According to the Final Report, no formal records management program or written policy has been established to date, due mainly to time constraints and the focusing of priorities elsewhere. A basic record keeping system using Microsoft Windows tools is currently used, and some internal guidelines and written documents concerning the "archiving" of files have been created. However, these guidelines and manuals are the result of experience and perceived best practice rather than archival science and are not rigorously adhered to in practice. Existing preservation activities include the following:

- Official data sets are stored on a computer in the MOST office at UBC and are regularly backed up onto 2 DVDs, which are held in different locations. Files were previously saved on CDs, but these have yet to be migrated to DVD.
- A second 'non-official' set of the data is stored on the computer of one of the MOST researchers.
- Backups are also made of the custom made software programs that are used in the project. Old versions of the programs are preserved whenever anything in the software is altered or updated so that researchers are able to recreate results previous created.
- Files containing scientific data are always preserved even if corrupt or false, as corruptions can later be filtered out without having to delete the corrupt file.

Information about the Digital Entity Being Studied

The digital entities being studied include SDS and FITS files, as well as various metadata. SDS files are those which are transmitted from the satellite to the ground stations and include SDS raw, SDS1, and SDS2 files. SDS2 files contain all critical and some supplementary

data, and SDS1 files are compacted version of SDS2 files, created in fulfillment of the CSA requirement that the satellite be equipped to store data for seven days in case of communication problems. Both SDS1 and SDS2 files are forwarded in the satellite to a processor to form parsed SDS raw files. Once the SDS raw files are transmitted to Earth, they are chopped up into SDS1 and SDS2 files again to create FITS files. These are the basis for all scientific analysis and data. Various metadata are created in the process as well, specifically contextual information about the satellite camera, data about image timing, etc., and orbit information downloaded from the Norad Web site.

The main activities in which researchers engage include data input, transformation of data into new file types, and analysis and publication of data. The documentary byproducts of these activities are a variety of data files (FITS files and other types of records resulting from the analysis of preliminary data).[4] Equipment used during research activities includes a satellite camera, satellite processors, transmission equipment, computers and custom made software.

[4] For a complete list of data files created, their file extensions, and explanations of the function of each, see the Case Study 25 Final Report, pp. 6-7.