The Development and Implementation of the SeaWiFS Records Management System (RMS)

Elaine R. Firestone
General Sciences Corporation
Laurel, Maryland 20702

ABSTRACT: A records management system (RMS) has been developed in support of the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Project of the National Aeronautics and Space Administration's (NASA) Goddard Space Flight Genter (GSFC). SeaWiFS is an ocean color sensor scheduled for launch in 1994. The development, implementation, and maintenance of this RMS is especially needed in light of the Project's role as a facet of NASA's Mission to Planet Earth effort, of which the Earth Observing Satellite (EOS) system, scheduled for deployment in the late 1990s, is a part. The RMS is Macintosh-based and uses a bibliographic database built under the Pro-Cite® package developed by Personal Bibliographic Software, Inc. (PBSI). The RMS seeks to provide easy access to current and archival documentation of the SeaWiFS Project activities. The documents serve as a historical record for future missions and its predecessor, the Nimbus-7 Coastal Zone Color Scanner (CZCS). This documentation is necessary to ensure that critical information related to the Project is available to the oceanographic, marine science, and scientific communities. The need for a system of keeping mission-related data was demonstrated by the CZCS during its period of operation from 1978 through 1986. For CZCS, the development of documentation and the resultant cataloging and archiving was not a priority. Because of this, unless CZCS data was published, some of the valuable information gleaned from this project was lost. So that this does not happen to SeaWiFS data, a customized RMS was developed. The concepts described herein are applicable to a wide range of applications and are not specific to SeaWiFS, NASA, marine science, or other disciplines.

1. INTRODUCTION

The Sea-viewing Wide Field-of-view Sensor (SeaWiFS) is an ocean color instrument scheduled for launch on the SeaStar satellite in 1994 (Hooker et al., 1992). Due to the size, scope, duration, and historical significance behind the mission itself, a Records

Management System (RMS) is required for tracking the numerous Project documents that have been, and will be, generated in conjunction with this endeavor.

In academia, as well as the government and private industry, small-to-medium sized groups and projects exist (5-45 people) that have records management problems. The problems can exist in different forms, such as the following:

1. A group that used to consist of 3 people rapidly grows to 20 and the individual record keeping systems are no longer sufficient for the group's needs.

2. Two or more very small departments with their own record keeping systems get reorganized and merged into one department that needs one centralized system.

3. An existing group has a large collection of one or two types of documents, e.g., article reprints in specific topics, that are no longer manageable without a formal system.

In the case of the SeaWiFS Project, Situation (described above) existed. The Project used to consist of 5 people and then quickly expanded to 30. Everyone on staff had their own records that they kept in their offices. The records consisted of reprints, conference proceedings, historical mission data, photographs, transparencies, etc., and no one person on the Project had complete records of the Project from its inception. Furthermore, although someone with a records management background was consulted to implement a system, it was not clear in the beginning, who would be responsible for the maintenance of it. For that reason, it had to be facile enough so that someone without this type of background, e.g., scientists or secretaries, could use it.

As can be seen in these instances, records management systems can be simple ones for managing a hundred of one or two types of documents, or more complex systems involving hundreds of documents or other media types. This paper describes the system that was developed and implemented for the SeaWiFS Project at the National Aeronautics and Space Administration's (NASA) Goddard Space Flight Center (GSFC). Note that the concepts illustrated here are not specific to NASA, marine science, or any other discipline, With small modifications, they can be applied to other efforts, and therefore, can be used virtually anywhere an RMS is needed.

2. SCOPE OF EFFORT

The system that needed to be developed for the Project had to be able to encompass a wide range of documentation, which run the gamut of complexity and length from one page informational memoranda to several hundred page Critical Design Review presentations. In between these two extremes, are other types of presentations, conference proceedings, reprints, journal articles, and memoranda in which information not published elsewhere is given. The RMS had to be able to accommodate these diverse documentation types, as well as to have the capability to expand and accommodate other

types of documentation. In addition, it needed to be flexible enough to accommodate other types of informational media, e.g., photographs, 35 mm slides, and transparencies, as the needs of the Project change.

3. POSSIBLE ALTERNATIVES

There had been three primary alternatives suggested instead of the proposed customized system. They were:

1. Do not have a formal system; leave things status quo. At the time, the status quo included all scientists having their own libraries, mostly consisting of documents in their particular area of interest and expertise, and subsequently, their own methods for filing and retrieving the information. Although this was workable for the individual, no central repository existed for all Project documents, regardless of topic.

2. Use the GSFC library as the repository and subsequent retrieval mechanism for SeaWiFS documents.

3. Have a greatly scaled-down version of a system where only journal articles and memoranda, etc., are housed and tracked.

Each of these alternatives is discussed below.

3.1 Do Not Have a System

This was the situation when the idea of an RMS was conceptualized. It was difficult to determine what information was in a given area, or even who might have it, thereby making document retrieval difficult and time consuming.

3.2 Use the GSFC Library

This alternative would not work at all and would not even be possible. The GSFC library has an extensive holdings list. However, since it is trying to serve the needs of the entire GSFC community, it cannot possibly meet the documentation storage and retrieval requirements of every effort currently in progress at Goddard.

The GSFC library has specific hours of operation, and a limited number of research librarians and computer terminals on which literature searches can be executed. This is not to say that the SeaWiFS staff should not, or could not, make use of the library. The library would be a very useful resource as an adjunct to an RMS for locating journal articles, etc., but not documents generated by the Project. With the Aerospace Research

Information Network (ARIN) and NASA's Research Connection (RECON) databases at the disposal of any authorized user, literature and bibliographic searches can be executed from a remote terminal. These databases are also invaluable tools for performing interdisciplinary research, since over 3,000 periodicals and journals are listed, which are directly available through the GSFC library in conjunction with the inter-library loan program with other NASA facilities.

In a mission such as SeaWiFS, however, the scientific requirements are such that information is updated and modified on a regular basis, sometimes daily. This new information must be readily obtainable whenever the staff requires it. The GSFC library has neither the mechanism for cataloging and updating records of this type, nor the inclination to store multiple iterations of a given document prior to publishing, if in fact the document is destined for publication; indeed, no library does. Only after actual publication is it considered for inclusion in the library holdings, at which time it is cataloged according to the Library of Congress system. Also, publication is not a viable end-product of many of the types of documents that the staff needs typically to access.

3.3 Have a Scaled-Down System

This smaller version of a system would entail archiving only journal articles and conference proceedings of interest to the Project, and memoranda. Saving journal articles and conference proceedings would, in many cases, duplicate the holdings of the GSFC library. The library, however, would not archive memoranda.

Due to the large volume of documents which the staff needs to access, a smaller version of the original system would not be adequate for its needs. It would not accomplish the stated goal of having a method by which mission-critical SeaWiFS documents can be readily retrieved in the smallest amount of time. These documents are *not* in the form of journal articles, conference proceedings, etc., even though they are important to the Project and should be part of the more extensive customized RMS, discussed in the next section. These holdings are primarily: Preliminary Design Reviews, Critical Design Reviews, configuration control documents, and other types of presentations that are required to demonstrate the preparedness of mission-critical facilities and other elements of the Project which support the effort starting at launch.

A smaller version would, however, be an ideal way for a small department in a university, or other small group, to formally track specific types of documents, such as reprints of journal articles, papers written for conference proceedings, and the like.

The most important thing to remember when designing an RMS on any scale, is that the specific needs of the department, project, and individuals involved have to be considered.

4. A CUSTOMIZED RMS

4.1 Introduction to Approach

The RMS that was designed for the SeaWiFS Project, as described in the following pages, was built on an existing commercial-off-the-shelf (COTS) software package, which had been determined to be one of the best available packages for this purpose (see Section 4.3). This, by its very nature, made the effort inexpensive, and thereby affordable for the Project.

In an RMS developed for a mission of this type, a number of features had to be integrated so that it meets the requirements of the user(s) both now and in the future. As such, it must be expandable, duplicatable, and sortable.

Expandability enables the RMS to grow with the Project. Since the current needs are naturally less expansive than those in 1994 when the satellite is due to be launched, the RMS needs to have the capacity to grow and adapt as the needs of the mission grow.

Duplicatability simply means that the same cataloging system used for Critical Design Reviews, can also be used to catalog journal articles, diskettes, slides, presentations, memoranda, etc.

Sortability is necessary to meet the needs of various searches required, i.e., one time a search of the RMS by the document's first author may be required, but another time, by the year it was written and the type of document. As a point of information, this sorting capability is not available through the GSFC library. If the ARIN database is used, only a reverse chronological (most recent holding first) listing is given.

4.2 The Cataloging Scheme

The whole idea of a Project-specific RMS was to have documents readily retrievable, whether or not they are available through the GSFC library. The RMS should have a unique cataloging scheme that is meaningful to the user(s).

4.2.1 Catalog (Call) Number Structure

The elements of the catalog *number* are as follows:

1. Year (YY): A numeric field, which for this RMS is the last two digits of the year, as in 93. Because of the way many hardware and software systems are designed, other RMSs designed now should use the full four digits of the year, i.e., YYYY to account for the change in the century.

2. Type of Documentation: An alphabetic field which consists of one or two letters (in this system). Presently, the types of documentation of the SeaWiFS RMS includes and their respective letter codes are: documents (D); presentations (PR); papers, including journal articles, conference publications, etc., (P); memoranda, including one page information letters, minutes of meetings if applicable, and/or white papers (M); transparencies (T); electronic media (EM); books (B); and 35mm slides (S). Note, although capital letters are used here, they are for aesthetic purposes only since the software used is not case sensitive.

3. Month (MM): Uses the numerical month equivalent, if applicable. If this information is unknown, or the document in a state of flux, set this field to double zero, 00.

4. Source Organization (SSS): For this RMS, a numeric field is used, whereby each source organization, e.g., GSFC, Jet Propulsion Laboratory (JPL), other NASA facilities, NASA contractors, etc., is assigned a three digit number. Using this method, up to 999 sources can be cataloged, however, because of the software used (Section 4.3), if the number of sources exceeds this number, one can simply enter higher numbers, such as, 1,000 or more, without any modifications to the database structure.

5. Item or accession number (NNNN): A numeric field which is assigned sequentially, and makes the document, or item, unique.

Using these elements, a sample catalog number for a document written by GSFC in May of 1991 might be as follows:

91-D-05-001-0001

Where 91 is the year; D indicates the item is a document; 05 is the numerical symbol for the month of May; 001 is the assigned number for GSFC, i.e., the source, or where the document originated; and finally, 0001 is the item, or accession number.

If other documents are written in the same month and year by GSFC, the item number (the last number in the grouping, i.e., 001) will change thereby giving each holding in the RMS a unique number. Since the individual elements of the above number are entered separately in their own fields within the database, *in addition to the call number written as a whole*, the elements are sortable on any of these fields, that is, if a sort is required for all holdings in the RMS written by GSFC in 1991, it can be done.

4.2.2 Questions for Consideration

Throughout the process of developing and implementing the RMS, both for the electronic database and the filing system for the hard copies of the documents, questions arose regarding the structure of the call number. The first question voiced was: "Why not use an alphabetic abbreviation instead of a number to denote the source

organization? It would be easier to locate documents within the files." This idea had been considered but not implemented for a number of reasons. First, although it initially would be easier to locate documents with an alphabetic source indicator, e.g., GSFC instead of 001, it would become increasingly more difficult as sources were added. Another, although secondary, aspect to having this as a numeric field, is that it makes the hard copy files a little more difficult to use for unauthorized users. People who are unfamiliar with the system, but who might need access to the documents would have to see the cataloger or other authorized user for help in finding required documents.

The second question was: "Why not put the year and month together instead of separated by document type, i.e., 91-05 instead of 91-D-05?" In point of fact, this was the structure of the number in the initial stage of the RMS. After approximately 25 documents had been cataloged, it was apparent that although this system made sense on paper, it did not when put into actual practice. Since authorized users would be working primarily from the hard copies of the documents held, and not have immediate access to the electronic RMS, the hard copy filing system had to fulfill their needs for easy access. Most requirements for document retrieval involve a year and document type, i.e., a presentation given in 1991, 91-PR. The month is secondary to this information. The structure of the call number was changed to reflect this and the cataloged documents were renumbered.

4.2.3 Other Information for Cataloging

Other information to be cataloged, if applicable, includes:

1. The full date of publication, e.g., January 7, 1992.

2. Author(s).

3. Authors' affiliation, e.g., GSFC, JPL, etc.

4. Description, or abstract, of the document.

5. Cross-reference numbers (when appropriate)
   a) NASA Contract Number, if the document was written by a subcontractor.
   b) NASA Technical Memorandum Number
   c) ISBN number, if the document has been published
   d) GSFC Library Number (Library of Congress system, if document has been published)
   e) Contractor Document Number, i.e., the number given by the contractor's technical publication department, if applicable

6. Subject or Keywords.

Fig. 1 is a partial screen in the electronic database which shows this information.

4.3 The Software Used for the System

Prior to the idea and subsequent development of a full scale RMS, the SeaWiFS Project had already purchased a copy of a software package known as Pro-Cite® for use in a scaled-down version of the RMS. This package, currently in use by the American Library Association, is available through Personal Bibliographic Software, Inc. (PBSI), in Ann Arbor, Michigan. After reviewing the literature from the company, it was determined for a number of reasons that this would, in fact, be an excellent package to use for the full-scale customization effort (PBSI 1991). Some of the salient features of the package are as follows:

A. 100,000 records can be contained in each database, with each record containing up to 16 pages (32,000 characters).

B. Entire records or specific fields can be searched with Boolean operators.

C. Help screens can be accessed at any time.

D. Records can be sorted on up to six levels and a Save and Restore Search Expressions feature allows for using the same search expression at a later time.

E. Customized output, in the form of various character styles including subscripts and superscripts, can be created.

F. Customized bibliographies can also be developed, to meet specific needs.

G. Records can be edited with global commands to change many records at one time, or selectively edit one record at a time.

H. Duplicate records are detected automatically and, once detected, the duplicate record can be deleted in full or in part, retaining just the abstract, as an example.

I. Unlike other database packages, the individual fields do not have character types (as in date, numeric, currency, alphabetic fields) or size limitations (for example, 10 characters long). With this feature, if the designer of the fields initially determined that a field should be three characters long but the cataloger needs four, change is made by simply entering in the extra characters.

J. The software is versatile enough to allow for customization of fields.

K. The package is available in either Macintosh or IBM versions and files created in one environment can be easily ported to the other using Biblio-Link®, another PBSI product.

4.4 Supplies Required for Implementation

Implementation of the RMS was started immediately after the implementation plan was approved (Firestone 1992), using the aforementioned call number system, even though the hardware had not yet been purchased. This was accomplished by manually cataloging all documents first on index cards. The reasons for this step included the following:

1. The initial step is labor and time intensive. When done this way, a computer does not have to be tied up for long periods of time while a document is being studied for the appropriateness of the cataloging scheme. (This involves verification that the individual elements that make up the catalog number are correct, as well as the retrieval of missing information that also goes into the database.)

2. Since the process is manual, the effort was not impacted by not having the hardware.

3. Since only common office supplies were needed, no large outlay of money was needed.

4. When documents are cataloged manually, an actual hard copy "card catalog" exists at every point in the cataloging process and can be utilized immediately. Even if hardware problems developed, a card catalog would still be available and could be utilized. Too many times, in technologically advanced libraries without card catalogs, no searching can be done when the computers go down.

If both the appropriate hardware and software are available at the onset, then this manual step can be eliminated and the items to be archived can be cataloged directly into the database, being careful to make frequent backups and printouts of the bibliography.

5. Issues to be Addressed

5.1 Accessibility

At the time the RMS was conceptualized, it was assumed that the primary user of the system and files would be one of the Project scientists. However, questions arose as to whether that person would be the only user, or would there be secondary users also? Where would the files be housed? If they would be in the scientist's office, how would secondary users get access to them if he is not there? If there are secondary users, where would the computer be housed? Would the secondary users have access to the system?

5.2 Maintenance and Upgrades

Who will catalog future entries both in hard copy form and then enter them into the system? when this effort was being considered at the Project level, the Project secretary was considered for this effort. However, when scope of the system changed to be a full-scale RMS, it was believed that this would have to be revisited.

Other considerations were:

- Who would be responsible for updates and enhancements to the database itself, if and when needed?
- Who would train secondary users, if required, on the system?
- Who would be responsible for tracking outgoing documents and subsequent refiling when they are returned?
- Would literature searches be required on the system? If so, who would do them?
- Would literature searches be required from the GSFC library through ARIN or RECON? If so, who would do them? The Project staff had been dependent on the GSFC library for support in this area?
- Who would decide which documents and other items would be entered into the system?

5.3 Personnel Requirements

It was suggested that a full or part-time position be created for a Records and Information Manager (RIM), especially if secondary users would be using the system. This position would entail all of the responsibilities listed in Sections 5.0-5.3, above. It was further recommended that this person would obtain a password to use the ARIN database and be trained in the use of RECON to assist in literature and bibliographic searches for Project staff. The flow chart in Fig. 2 illustrates the primary avenues for how an item is entered.

6.0 The Current System

The system, as described, was approved and implementation of it began in February 1992. A part-time RIM was hired for this effort; and the system resides on the RI's Macintosh computer. To date, there are over 120 holdings in the system. Due to the volume of holdings and space constraints at GSFC, efforts are now underway to investigate alternative filing systems which would take up less space than the traditional filing cabinets presently in use.

What had started as a system for one scientist, is now the official repository of all Project-specific documentation for the entire SeaWiFS Project. Although there were opponents to this full-scale RMS, no one now doubts the usefulness of it. Almost everyone on the Project, or associated with it, has made use of the RMS at one time or another.

7.0 Future Plans

It is the intention of the RIM, within the next year, to catalog the individual collections of marine science articles, in the form of reprints, from the scientists on the SeaWiFS Project. In addition to having these papers electronically searchable for future use by the Project, the citations will be published in a separate volume of the SeaWiFS Technical Report Series, *NASA Technical Memorandum Number 104566*, for which the RIM is the Technical Editor. It will be disseminated throughout the marine science and oceanographic communities as a definitive source for information in this emerging area of science.

Glossary

ARIN	Aerospace Research Information Network
COTS	Commercial-Off-The-Shelf (software)
CZCS	Coastal Zone Color Scanner
EOS	Earth Observing Satellite
GSFC	Goddard Space Flight Center
IBM	International Business Machines
JPL	Jet Propulsion Laboratory
NASA	National Aeronautics and Space Administration
PBSI	Personal Bibliographic Software, Inc.
RECON	Research Connection (NASA)
RIM	Records and Information Manager
RMS	Records Management System
SDPS	SeaWiFS Data Processing System
SeaWiFS	Sea-viewing Wide Field-of-View Sensor

References

Firestone, E.R., 1992: *A Plan for the Implementation of the SeaWiFS Records Management System,* General Sciences Corporation, written under NASA Goddard Space flight Center Contract Number NAS5-31733, 7 pp.

Hooker, S.B., W.E. Esaias, G.C. Feldman, W.W. Gregg, and C.R. McClain, 1992: An Overview of SeaWiFS and Ocean Color. *NASA Tech. Memo. 104566, Vol. 1,* S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 24pp., plus color plates.

Personal Bibliographic Software, Inc., 1991: *Pro-Cite® for the Macintosh,* Ann Arbor, Michigan, 580 pp.