

Final Report on Electronic Archiving Project for the North American Serials Interest Group, Inc.

LIBRARY AUTOMATION CONSULTANTS FINAL REPORT
JANUARY 5, 2001

SUBMITTED BY
MARK MCFARLAND, MLIS
2012 SINGING BROOK
AUSTIN, TX
(512) 926.8146
(512) 926.1560 (FAX)
M.MCFARLAND@MAIL.UTEXAS.EDU

Summary

The main criteria to use in the decision about how best to manage electronic records should be based on the following factors:

- volume of electronic-only information being produced by the organization
- need to continue producing printed documents
- the purpose of the electronic-only information
- the resources and organizational structure available to manage this information
- cost associated with maintaining electronic archive

Given the fact that NASIG does not have the resources or the need to host a digital library environment coupled with the fact that most of the information produced by NASIG is intended for use in print formats it seems logical for NASIG to follow a "print and save" strategy for its electronic information when possible. However, due to the preponderance of source files from which most printed information is generated it is also recommended that NASIG establish a set of guidelines for its members who are involved in creating and maintaining electronic source files for digital-only information and for printed information. Keeping the digital copy of an electronic record "permanently" implies regular cycles of data refreshing and migration in the context of regular cycles of hardware, operating system, and application upgrades. With these routine maintenance cycles will come ongoing costs. In addition to routine costs there may be unexpected costs associated with new technologies that may be attractive and even essential to the preservation effort. For these reasons, the decision to maintain digital records is a very significant decision that will have impacts throughout the organization. Finally, NASIG should identify that group of electronic documents that needs a separate retention period, decide what those retention periods should be, then establish a process by which those documents are disposed of. Again, since NASIG is outsourcing its information systems management to vendor this will require cooperation between the NASIG archivist and the vendor.

The appraisal of electronic records should involve answering two questions:

- Does a digital record have archival value?
- If so, must it be retained in digital form?

NASIG should not attempt to embark on the development of a locally operated digital archival repository due to the resources that would be required to implement and maintain such a repository.

Working Assumptions

We have been making the following assumptions with respect to your request for a report that responds to your stated concerns and issues: 1- that your need for printed information is not likely to change in the next 12 to 18 months; and 2- that you are not likely to change your desire for a remotely hosted/managed website in the next 12 to 18 months; 3- that the NASIGWEB functions as a de facto digital archive for a variety of organizational data.

Responses to your questions and concerns are based on these assumptions.

Responses to Questions and Concerns

What follows is a series of responses to the specific issues raised by NASIG with respect to the archiving of electronic information.

1. What should NASIG do with regards to its electronic documents and communications in terms of archival strategies, given the all-volunteer and transitory nature of our organization?

Our recommendation is that NASIG should incorporate existing electronic information into the current records management process. Furthermore, NASIG should not undertake a project to transform current practice with respect to creation and dissemination of official organizational information. As stated above, the "all-volunteer and transitory nature" of your organization is an important factor in weighing how best to proceed with respect to incorporating electronic documents into your records management processes. In terms of a strategy for archiving electronic information a "print and save" model combined with an effort to bring your electronic-only materials under control is our primary recommendation.

The NASIG archivist should first identify which electronic documents should be subject to formal management processes and then determine the retention periods of those documents.

We would caution against a policy that would too readily justify the permanent retention of an electronic record. The criteria for making the determination that an electronic record be permanently retained should be stringent and pertain only to information that is critical to the ability of the organization to fulfill its stated mission. NASIG should direct its resources to the preservation of its information to the degree that is commensurate with the value of that information to the organization.

The electronic documents owned by NASIG can be categorized as follows: 1-information that is produced, stored and managed in electronic formats but intended for print; 2-information that is designed and intended for electronic use only; 3-printed or other non-electronic materials acquired by NASIG.

The documents in the first category (the group of source files used primarily to produce printed documents) should be printed, saved and disposed of in accordance with your existing retention schedule. These electronic source files should be maintained in text-based formats when possible to insure that the information in these files can be moved forward from application to application and machine to machine over time. Files that are not viewed frequently (once a month or more) but that have been deemed of value to the organization should be viewed and copied at least annually then printed saved to insure availability.

The information in the second category (the group of digital-only files) should be viewed and copied annually according to their predetermined value to your existing retention schedule. It is not always possible or desirable to produce printed output from information that was designed to be used as digital-only information. Many databases, for example, do not always lend themselves to print as a faithful output format (a library OPAC, for example). The usable information stored in electronic databases often cannot be easily reproduced outside the software architecture that renders the data in all its various "views".

The information in the third category (non-digital information which might lend itself readily to a digitization process) should not be converted as a means of preserving the original information encoded on paper. Digital media should not be used as a preservation solution unless a commitment can be made to the ongoing refreshment and migration of all necessary hardware, software and information content.

2. What position(s) in the organization should be responsible for electronic documents and communications in terms of an archival strategy? Where should it be done? Do we reconsider the role of archivist; have 2 archivists, or assign electronic documents management to the Electronic Communications Committee?

The NASIG archivist should retain responsibility for all tasks associated with the appraisal and collection of NASIG's electronic records. However, given the existence of the NASIG website and its variety of web-based information there is a need to augment the traditional archivist's responsibilities in a way that will bring the materials on the website under the purview of the archivist. We do not recommend the creation of separate positions to manage print and electronic materials separately. The current archivist should incorporate several new tasks and share the responsibility of liaising with the website hosting company with either the NASIG web manager or the Electronic Communications Committee. If there is a NASIG webmaster this person should work with the archivist to develop a set of guidelines that govern the means by which information is placed on the website and how that information will be maintained over time. Since NASIG has outsourced the maintenance and operation of the technical infrastructure required to make its electronic information available it is important for the archivist to maintain contact with either the hosting company or with the person at NASIG who is responsible for managing the information on the server. It is important that contact with the hosting company involves information exchange about how the materials on the server are being managed by the hosting company and that this information is documented. By outsourcing the website (which is a very reasonable and practical solution for NASIG) you are, in effect, outsourcing the management of your digital repository. It is likely that the NASIGWEB is the main collecting point (in a physical sense) for a large variety of information produced by NASIG; as such, the NASIGWEB is serving as a quasi-archival repository for your electronic information. It is reasonable, therefore, to extend the management tasks associated with maintaining electronic documents to NASIGWEB.

It is important to know the answers to the following questions:

- how frequently is the website backed up?

- is there a copy of your web content backed up to tape or to optical media?

- is a copy of the backup media being kept offsite?

- what is the process for retrieving a copy of a file from the backup system?

- how quickly can a copy of a backed up file be retrieved?

- how would the hosting company provide copies of the information on the website should NASIG decide to move to another company?

- what is the company's disaster recovery plan with respect to your information?

- what is the company's policy with respect to moving your data forward as they change hardware platforms or operating system environments?

It is important that the archivist maintain a relationship with the information resources provider such that there is a good channel of communication for discussing issues pertaining to the applications and information operated on the remote host.

3. Make recommendations on what is needed to establish a records management process for electronic documents.

As far as this consultant has been able to determine there are a small number of electronic-only records being produced by NASIG. These electronic-only records consist of the correspondence generated by the electronic mail distribution lists sponsored by various committees and the award applications produced by the Awards and Recognition committee. This group of electronic documents would form the content of the electronic documents series in your current retention schedule. It would be the task of the archivist to determine the retention schedule for these materials and determine the process by which obsolete files are disposed of. A practical approach to managing these two very different types of information would be to employ a print and save policy on the awards forms and an electronic-only solution on the correspondence.

A print and save strategy is appropriate when the large majority of the information produced is in print. Printing out forms and saving them in print format simplifies the process of managing the electronic file by establishing a safe copy of the data in a printed file. However, printing an electronic document that was produced to be used as an electronic-only file can become cumbersome for frequently updated documents or even impossible for complex electronic documents that contain hyperlinks that form substantial portions of the content, particularly when portions of the hyperlinked content are not produced or maintained by the organization responsible for the main document.

Guidelines for Creating a Records Retention Policy for Electronic Files on Desktop Computers

If an electronic-only file stored on a personal computer consists of the only copy of a document and is deemed to be an "official" document (i.e. relating to the organization's business) then these files should not be destroyed except as scheduled. If a record is merely a source file for an "official" printed document that source file (the electronic file) should not be considered an "official" record and would not meet criteria for scheduled disposition or management. However, these source files should not be maintained longer than necessary for the production of current version of the printed record. Staff or committee members responsible for producing and maintaining electronic source files should review their documents frequently and be encouraged to delete obsolete electronic files (i.e. files no longer necessary for producing the official record) from the fixed disks on their desktop systems as well as from all other storage media (offline or nearline). This would include removal of the file from any magnetic or optical storage media (floppy disks, zip disks, backup tapes, or cd-rom).

However, electronic mail does not lend itself to a print and save strategy and is best managed in the electronic environment in which it is created and distributed. The criteria for determining how long to save email and whether or not to archive it should be established by the NASIG archivist in consultation with those responsible for the management of the distribution lists themselves. The resources required to manage, dispose of, or archive electronic mail are, obviously, of primary importance and are jointly controlled by the list owners and by the technical staff at the hosting company. Since the tasks in the process of managing electronic mail are distributed among the list owners, technical staff at the hosting company and the archivist it will be necessary to communicate the policy for governing electronic mail very clearly. The policy should state the purpose of the electronic correspondence and why it is of value, the length of time the correspondence will be retained and that it will be regularly backed up during its active status prior to deletion as prescribed by your retention schedule.

Guidelines for Creating a Records Retention Policy for Electronic Mail

The organization's records management policy for electronic mail should state that this form of expression is considered to be a businesslike expression and that the character of these communications should be professional in nature. The retention period for email will vary depending on the needs and requirements of the organization. Many organizations determine that a 30 day retention period is all that is necessary for email. However, mail generated by subscribers to distribution lists is often archived and made available to subscribers as a means of mining messages for useful information. This is a decision that will have to be made by an individual organization.

4. Recommend what should be captured and archived electronically (identification, policy and security issues).

The decision to convert existing print-based (or otherwise analog encoded) information to digital formats has many ramifications. There are two primary reasons to undertake a conversion: 1- to provide wider access to a particular information object; 2- to incorporate information into a digital collection or aggregation of information to accomplish specific purposes. This implies that the purpose for creating a digital archive of printed materials should not, at this time, be for the purpose of permanent retention, but, instead, for the purpose of providing better access to information content. Therefore, the process of identifying information to be converted to digital formats should be based on the assumption that you are attempting to provide better access to the information content for end users (however large or small the population of those users may be).

The policy issues, in the context of a records management strategy, would be based on the status of those documents as "official" records and the retention schedule would reflect their disposition.

The security issues pertaining to electronic information are varied and numerous. There are issues pertaining to the accuracy and integrity of information. This aspect is best understood in terms of traditional scholarly publishing and the "controls" that publishers place on the process of making viable scholarship available to the research community. Another aspect of security is availability. Availability in the sense of how stable and robust are the information systems that support and provide access to materials. Are these systems secure against common threats from hackers or hardware failure? Finally, there are issues pertaining to the physical surroundings in which the hardware is located. Is the hardware managed in safe, protected area that is not subject to physical abuse? Is the staff capable of managing the hardware, etc??

We would not recommend that any materials in the extant printed archive be converted to digital formats and archived electronically without a specific purpose or purposes that would justify the expense of preparation for conversion, conversion, and ongoing maintenance. Again, moving the contents of the current printed archive is not recommended due to the complexity and expense of ongoing maintenance. However, if NASIG were to move in a direction that obviated the need for printed materials it would make sense, at that point, to seriously consider a conversion project for the printed archive.

5. Make recommendations on how access to these materials will be handled.

Since we are not recommending that NASIG embark on a conversion project to migrate to an electronic records management system there is little need to address recommendations on access to such a digital archives repository. It is assumed that the primary and de facto digital repository is the website maintained by Bee.Net. We could not find information pertaining to setting up a search feature on the Bee.Net website. It would be useful for you to provide a search interface through the website to all content on the website. This is a responsibility the service provider should be willing to speak with you about.

6. Who will have access to these archives and how accessible are they?

In general, access to your materials should be decided based on who needs access to particular information based on their work duties or on services that have been purchased.

Conclusion

NASIG's consideration of the question of whether or not to implement an electronic records management system seems to be driven by what appears to be a widespread movement away from printed information and toward electronic information. Many organizations are struggling with similar questions and there are a wide variety of approaches to the problem. However, it does not appear, at this time, to be in the best interest of NASIG to embark on such a project. Instead, we are recommending a very modest approach to dealing with the growing number of electronic files produced and managed by NASIG. This approach, hopefully, has been clearly outlined in the preceding pages in the responses to stated concerns and questions. Basically, the recommended approach is to print and save digital-only files that readily lend themselves to printed formats; print and save electronic source files used to produce printed documents; avoid converting printed materials from the existing archive to digital formats unless there is a very good reason for doing so.

It is also recommended that the NASIG archivist work with the Electronic Communications Committee and Bee.Net to assemble some basic information about how NASIGWEB is managed and maintained by technical staff at Bee.Net.

In developing the project report it became clear that the question posed in the RFP is a fundamental question for many organizations at this particular time due to the meteoric rise of the World Wide Web and the rapid growth in the distribution of so much digital information through

the global Internet. Even the midst of a digital information explosion there is still a need for organizations and end-users to consume information in print-based and other analog formats and this need will continue until such time as technology developers can deploy solutions that meet all the needs of information consumers ? not just those needs that pertain to information that is best-suited for use over networked computer systems. For the time being, at least, it still makes better sense for NASIG to continue a conservative approach to electronic records management by continuing to focus most of its resources on adapting traditional records management strategies to electronic information. Even at this, however, NASIG recognizes the implications of attempting to manage a growing volume of electronic, network-based information and that traditional records management is not adequate as the volume of electronic records grows.