

## All-Site bibliography improvements– Request for comments

When we set out to improve the all-site bibliography it was seen as an on-line tool for searching and managing site bibliographies that would be compatible with EndNote (which was the software most in use in the network at that time). Several request for improvements and bugfixes to the current version have been suggested but prior to engaging in any changes we'd like to receive comments on a couple of key issues.

Issue #1 - When a citation is entered into the system it receives a unique ID (this doesn't imply that the citation is unique). This was seen as a good thing by many because it enable citations to be linked to other data resources such as personnel and data catalogs. In addition the system was designed to accept incremental additions so that entire site bibliographies would not have to be uploaded again. There in lies the twist that prompts this request – in reality it is very difficult to detect and act on duplicate bibliography entries without the consensual use of a unique identifier because all of the field used to detect potential duplicates are also the fields most likely to change as a publication moves from 'in press' to published – those being title, authors, and year. We've found repeatedly that it is easier to delete and reload an entire site bibliography than to track down the Type I and Type II error duplicates. We discussed briefly at the Montreal IM meeting the selection of a field in EndNote that could be used to house a unique identifier (his not being a problem for those generating their bibliographies from a database). Don Henshaw and I discussed using Accession Number (%M) because some sites already have one – this would simplify the upload process because it would only be necessary to compare the accession numbers within the 'site' attribute space (essentially creating a compound key). The system-wide unique reference id would continue to be assigned on upload but the accession numbers would be used to discern between records for new inserts and records for updating.

1. Pro or con comments regarding the use of the EndNote/EML Accession number field (%M in the tagged file) as a unique identifier for your site that you assign.

Issue #2 - We currently provide a complete set of interfaces for managing a site bibliography over the web. Inserting, editing, deleting, exporting in EML or EndNote of individual citations can all be accomplished via this web-based interface. We also provide the source code for the search interface that can be customized for your site. The whole package represents a complete solution for a site that wants to provide this service for it's investigators or for general consumption. No commercial package provides this capability at present. This was not one of our original requirements but it co-evolved with the other work because of our needs in testing. Before we get to far down that road of providing a web-services interface to all-site bibliography we would like to know if anyone is actually using these search or management interfaces and to what extent.

2. Pro or con comments on LNO providing the capability to remotely manage a site bibliography vs. being simply an access portal to bibliography data that are managed at the site and uploaded or harvested on a routine basis. Please feel free to carry this discussion over to the personnel database because the same questions will arise there.

Comments should received by me no later than February 1 so that I can condense them into a discussion at the IMEXEC meeting February 7-8.