Mr. John Hoyle
Chairman, LSS Advisory Review Panel
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

Re: Header Working Group Report "Recommended Fields for LSS Header Records"

Dear Mr. Hoyle:

We have some comments on the ARP Header Working Group recommendations which we feel should be addressed prior to any votes being taken on header elements.

A primary concern is that, on four separate issues, de facto policy is being established.

1.) The Abstract/Summary field has been identified, and we agree that it should be one of the fields. Deciding which categories of documents require abstracts is a fundamental information management question for the LSS. Thus, the precise guidelines for when this field is going to be required, and a detailed description of the style to be used, both need to be promulgated as part-and-parcel of designating this as an LSS header field. This will allow participants to begin including this information, where required, for all LSS-relevant records processed henceforth. Likewise, the volume of records involved and the size of the abstract each figure into the sizing of the LSS header record files.

2.) In the discussion of Editing of headers by LSSA on page 3, a recommendation has been made that LSSA staff will review submitted data against quality control standards and LSSA staff will correct entries.

By one reading, the recommendation as worded sounds like it anticipates that records would bypass the Capture System. All records must flow through the Capture System operations before an LSS system load disk is created; the conceptual design has never included an additional review by OLSSA prior to data load because the LSS Capture Systems will be responsible for meeting quality standards and are all to be operated under the strict processing procedures put in place by OLSSA.

An alternative reading of this recommendation suggests that the output of the Capture System process will be so deficient as to require an additional review by LSSA staff prior to database loading. Tightly controlling the Capture System processes and

procedures obviates the need for subsequent review except as already outlined in the <u>LSS Capture System Detailed Design Document</u>. If such a second review was anticipated, it is our opinion that OLSSA staff again reviewing the output of a Capture System installation will ultimately prove to be an unworkable strategy because of the sheer number of records. Likewise, the Capture Station configurations to be operated directly by the LSSA were not sized for this volume of re-review and scrutiny. Finally, it is redundant of work that should have been done either by the submitter in preparing records for submission to the LSS or by the Capture System processing.

A third alternative reading of this recommendation suggests that data generated from feeder systems such as DOE's RIS would, of necessity, require extensive scrutiny and rework within the Capture System environment. For DOE and NRC, with 90% of the information, rigorous data capture procedures should be instituted and audited in both the feeder systems such as DOE's RIS and in any co-located Capture System which supports conversion to the LSS header format. For the parties with smaller volumes of submissions, the LSSA can more easily check, edit, and add information to headers than control standardized entry procedures for the other parties' feeder systems. Hence, more or less rework may be required by the Capture System depending on who the submitter is, but all of the correction work and additive cataloging is via the Capture System.

Perhaps we are belaboring the point, but, all other elements of the DOE program will be performed under rigorous QA procedures and it is the adherence to these procedures that gets continually audited. We feel that this is the model that should be used for LSS data submission as well. An optimal environment is one where the quality standards that will be acceptable are defined well in advance, already implemented in internal procedures, where the OLSSA dedicates resources for continuing audits of submitters' adherence to processing procedures (both in and outside of the Capture System environment), and, where batches of submitted data not meeting quality standards are returned to the submitter for cleanup. We cannot support massive reprocessing by LSSA staff. When batches are returned wholesale, direction is provided to remediate the submitters preprocessing until it conforms to the stated quality standards.

3.) Page 2, notes that an issue to be resolved by ARP is the updating of a header record when two participants submit different headers for the same document and they characterize some information differently, for example, the title/description. Should all the information be merged into one header or does the first header prevail? The recommendation of the working group is to append the subject information, from a subsequent submission that is different, to the respective fields of the original header.

We have a number of concerns about this recommendation:

3/243

A.) The concept of continually revisiting and updating header records raises more fundamental questions. To what purpose? When will the updates end? This recommendation invites changes to an LSS record once it has been submitted and "locked". DOE's records and headers are those of the license applicant, and the Rule says that each party is responsible for submission of all of its own relevant materials. Should we be designing a system which allows anyone to editorialize on the license applicant's (or anyone else's) submitted header? Will DOE still be responsible for the contents of such a changed header record?

Is the OLSSA authorized to be more than the custodian of the LSS, and is OLSSA ready to accept that responsibility?

B.) Any _created_ title is just that, and will _always_ be subjective. In a system providing text search capabilities and a controlled vocabulary, will a superior title promote retrieval any more effectively than a merely adequate one? Our recommendation is to define the standards for a created title and ensure that the submitter complies with the standard.

C.) This scenario is most likely with created titles, identifiers, descriptors, and abstracts. The recommendation to add data values to a textual field such as a title or an abstract could cause horrific database administration problems depending on the DBMS used, e.g.; reloads of indexes on gigabytes of data.

Also, there could be auditability and integrity problems.

4.) Appendix B, discusses the Related Documents field. For the submitter, it will be used to store relationships between submitted cataloging units, such as parent/child, superseding versions, etc., so that this can be identified during the submission of records to a capture station. Then, the LSS administrator (Capture System operator) is to convert this data into LSS acceptable pointers in the LSS environment, where all duplicates are filtered out and pointers set to existing versions.

The submitters conceptualization of linkages may not track exactly to the nature of the LSS linkages. What happens if there is no LSS equivalent to the submitter's relational statement? The LSSA will erase the submitter's non-analogous statement. Does this violate the participants use of the LSS as its records system? No, document linkages are still available.

But, this approach forces each submitter to commit to the LSS design and configuration. This is a policy decision for which we should be eliciting up-front commitment. The fact that we are requesting this commitment should be made explicit.
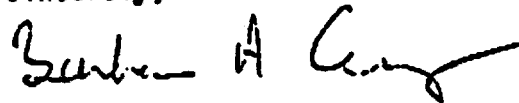
4/243

In addition, we have comments that are of less critical import, but should still be addressed prior to acceptance of the recommended fields.

5.) Page 3, and continuing on page 4, recommends a code field for the location of non-text/non-bit-mappable data. It should be added to the list as part of the submitter header and submitted in a non-code format so that the control list can be developed. We request this be added to the submitters' fields list.

6.) Appendix B, page 5, in discussing administrative and process tracking fields, suggests additional data be maintained in the LSS header. Most of the items on this list are not header data, indeed they are processing tracking data, and items a-f on this list will already be available in the process control databases maintained in the Capture System processing. Why duplicate the data in the LSS header, too? If the systems administration staff needs the data, they could mount the history file of the process control data files from the Capture Systems.

We recommend that these comments be presented to the members of the ARP in advance of the upcoming meeting and that the members of the ARP give them due consideration before any call for a vote on the recommended list of fields.

Sincerely,

Barbara A. Cerny, Director
Information Resources Management
    Division
Office of Civilian Radioactive Waste
    Management