Summary of the Aleph Record Weeding Project, Phase 1
Presented to the Technical Services Standing Committee, January 27, 2010
Prepared by Jennifer Jubinville, Harvard University Library Office for Information Systems

Project Summary
Aleph Record Weeding was prioritized by OIS in 2008 as a project to remove unneeded records from the Aleph system. Since the conversion to Aleph/HOLLIS Classic in 2002, only small subsets of records (authorities and some patrons) have been weeded in any systematic way. The implementation of an ongoing record weeding maintenance routine will produce technical, functional, and financial benefits, reducing the overall size of the Aleph database and improving usability and performance by removing clutter.

Options for weeding include:
- deleting Aleph records entirely
- deleting records from Aleph and archiving the full record outside of Aleph (most likely in the Library Reporting System)
- abbreviating records into stubs in Aleph and archiving the full record outside of Aleph.

Protection of active and/or useful data from removal or change is also an important consideration.

This initial phase of the project involved working with the OIS Aleph Support Team and other University librarians to identify sets of records that do not need to be retained in Aleph as well as records that should not be changed or removed.

Future work includes:
- production of policy recommendations to support and complement weeding activities
- close evaluation of the following recommendations in the context of the University Record Retention Schedule
- evaluation of the cost effectiveness versus effort of implementing certain weeding proposals
- drafting more detailed functional and technical specifications for a weeding maintenance routine
- planning for the addition of archiving functionality to the Library Reporting System
- evaluation of record sets whose status as weeding candidates is unclear
- manual cleanup of particularly troublesome records

Feedback from Librarians
In Fall 2009, the following Aleph Working Groups and a Coordinating Committee were introduced to the project and were asked to identify record sets that could be weeded from Aleph, as they relate to the charge of their particular group and area of professional expertise. The Technical Services Managers Affinity Group was also introduced to the project and members of that group were invited to participate.

- Aleph Acquisitions Working Group
- Aleph Cataloging and Authorities Working Group
- Circulation, Privileges, and ILL Coordinating Committee
- Bibliographic Standards Working Group

What follows are the recommendations of these groups and members of the OIS Aleph Support Team.
**Weeding Recommendations**

The following types of records should be deleted from the Aleph database; estimated record counts, as of 1/25/2010, are included for selected record sets. Careful consideration of the relationship to associated records in Aleph and other HUL systems must be made, which may require retention of individual records otherwise suitable for weeding.

To implement these recommendations, they must be prioritized and discussed in greater detail in order to define functional and technical specifications for an automated record removal routine.

**Bibliographic Records**
- Active records that have not been edited in 90 days and are not linked to holding, item, or order records [66,000]
- Deleted records [174,000]

**Holdings Records**
- Active records lacking 001, 852, and OWN fields [34,000]
- Deleted records [416,000]

**Patron Records**
- Employees and Summer School Students:
  - Records with an expiration date older than 5 years and no loans or associated cash transactions
- Students:
  - Records with an expiration date older than 5 years with no loans and with or without associated cash transactions
- Summer School Faculty:
  - All records with no associated loans
- Proxies, HCL Special Borrowers, Sub-Library Special Borrowers, and Pseudopatrons:
  - Continue manual management by OIS

**Cash Transactions**
- Remove when associated patron is deleted

**Circulation Events**
- Records with a Circulation System Number of ‘000000000’ [309]

**Acquisitions Vendor Records**
- Records created more than 2 years ago and never used (not associated with any Subscriptions, Orders, or Invoices, including Order Unit vendor records)

**Authority Records**
- Deleted records

**Archiving Recommendations**

The following types of records should be deleted from Aleph and archived in and made accessible through the Library Reporting System; estimated record counts, as of 1/25/2010, are included for selected record sets. Careful consideration of the relationship to associated records in Aleph and other HUL systems must be made, which may require retention of individual records otherwise suitable for archiving.
To implement these recommendations, the infrastructure and support of the Library Reporting System must be evaluated and enhanced to provide perpetual access to records no longer present in Aleph.

Circulation Events
- Records with an Event Date older than 10 years [0]

Loan History Records
- Records with a Return Date older than 15 years [75]

Budget Records
- Records older than 4 fiscal years and not associated with any orders

Order Records
- Orders older than 4 fiscal years:
  - Records with a status other than New, Ready to Send to Vendor, or Sent to Vendor [119,000]
  - Records with a status of Sent to Vendor and which have been completely arrived [127,000]

Subscription Records
- Records closed more than 4 fiscal years ago [9,000]

Acquisitions Vendor Records
- Records older than 4 fiscal years and with an inactive status [350]

Records to be Converted to Stubs and Archived
The following types of records should be abbreviated in the Aleph database, while the full record will be archived in and made accessible through the Library Reporting System. Careful consideration of the relationship to associated records in Aleph and other HUL systems must be made, which may require retention of complete individual records otherwise suitable for abbreviation and archiving.

Invoice Records
- Records older than 4 fiscal years, approved, and in batches that have been released [137,000]

Records to be Protected from Weeding or Archiving
The following types of records should remain in the Aleph database.

Bibliographic Records
- Active records linked to holdings, orders, and/or items
- Active records edited within 90 days
- Suppressed records

Holdings Records
- Active records linked to a bibliographic record
- Suppressed records

Items and Item History Records
- All records

Loan Records
- All records
Topics Requiring Further Evaluation

Suppressed Holding Records
As of 1/25/2010, there are 778,259 suppressed holdings in Aleph. The Aleph Cataloging and Authorities Working Group and Bibliographic Standards Working Group jointly recommend further study of deaccessioning workflows. The consensus in both groups is that suppression rather than deletion has been widely used for deaccessioned materials. A policy or practice is needed to recommend deletion during deaccessioning.

Suppressed Bibliographic Records
There exist suppressed bibliographic records with active holdings. As these records create problems in the public catalog, the Aleph Cataloging and Authorities Working Group and Bibliographic Standards Working Group jointly recommend a project to evaluate and clean up such records.

There exist suppressed bibliographic records with active orders. The Acquisitions Working Group recommends a project to evaluate and clean up such records.

Reserves Records
CPILL conditionally recommends weeding reserves records that have been inactive for 5 years, but the impact of the removal of such records on other systems and on item records associated with reserves records otherwise suitable for weeding must be examined before an official recommendation can be made.

Paid Invoices (Released or Frozen Status)
The Acquisitions Working Group recommends further evaluation of whether archiving budget records would adversely affect invoice records that have a status of Released or Frozen.

Orders with a New Status
There exist order records with a New status where the order has been sent to the vendor. Changing the status from New to Sent to Vendor will result in inaccurate order date information. The Acquisitions Working Group recommends evaluation and cleanup of these records so that they can be archived with other orders older than 4 years that have actually been sent to vendor and are completely arrived.

Orders with a Delayed Status
The Acquisitions Working Group recommends a project to evaluate and clean up such records.

Active Orders for Closed Libraries
The Acquisitions Working Group recommends a project to evaluate and clean up such records.

Multiple Vendor Records for a Vendor
There exist multiple vendor records for the same vendor at the same address. The Acquisitions Working Group recommends a project to evaluate and clean up such records.

Budgets Older than Four Fiscal Years
The OIS Aleph Support Team recommends further evaluation of the need to weed budgets and budget encumbrances in tandem.

Resource Files
The OIS Aleph Support Team recommends further evaluation of the need to weed the Resource File libraries (HVD02 and HVD03) in the Aleph system.
Electronic Book Plates
A planned project to implement electronic book plates may have an impact on how certain financial data are weeded or archived. Once the implementation plan for this feature is finalized, weeding or archiving routines will need to factor in its record retention requirements. Members of the OIS Aleph Support Team will address this consideration.

Future Steps
As described elsewhere in this document, more analysis is necessary before actual removal or archiving of records can begin. Planning for the next phase would be enhanced by feedback from the Technical Services Standing Committee about the above recommendations and by support from the Committee in identifying or allocating resources to undertake the projects described in the Topics for Further Evaluation.