Planning and Implementing ALEPH System Purges

Donna L. Hirst
University of Iowa, donna-hirst@uiowa.edu

Purging ALEPH Order Data
June 6, 2006

Questions:
Who is already purging Aleph data?
How many librarians are in the audience?
How many programmers or technical staff are in the audience?

Scope of Purge Discussions

- Iowa has been regularly purging the **Patron File** on an annual basis.
- Iowa has just completed a comprehensive purge of all **acquisitions and fund files**.
- We are in the process of purging and archiving **Circulation History files and Event files**.
- This session is primarily about purging acquisitions data, but if time permits, we could discuss some of our additional experiences with purges.

Overview of Purges at Iowa

- Iowa went into production with Aleph in August 2000.
- Prior to implementing Aleph, Iowa had a NOTIS system.
- As part of our initial Aleph implementation, we purged virtually all of our NOTIS files. We involved unit heads and functional groups in the decisions of how much data to purge.
- When we built our initial Aleph system we carried over a lot of historical data into ALEPH Acquisitions and Circulation files.
- Ex Libris told us that the system could handle up to 3 years of acquisitions data. We carried forward 3 years of data knowing that the files would need to be purged relatively soon.
- We did experience response time problems in acquisitions. We applied a few Band-Aids here and there. We started talking about purges but didn't get serious until 2004.

Rationale

- Response time in the acquisitions client was slow
- The large amount of acquisitions data created complications during general system upgrades especially version upgrades.
- Staff had problems efficiently identifying current data in the midst of extensive legacy data
- Approximate record counts for primary acquisitions files:

Orders	Z68	600,000
Order Logs	Z71	6.6 million
Invoice Line	Z77	580,000
Invoice Header	Z75	180,000
Budget Transactions	Z601	900,000
Budget	Z76	2,200
Vendor	Z70	20,000

<u>Planning for Acquisitions Purges</u>

- Decision making was very collaborative and cross departmental.
- We had three basic groups of people involved in decision making
  Library functional experts
  Library systems office
  Programming staff
- Within the Libraries there was involvement from:
  - University Library – Monographic Acquisitions, Serials Acquisitions, and Government Publications
  - Law Library
  - Finance

Slide of staff

Who were the Players?

Library Automation Office
Head (General overview)
Acquisitions/Serials Subcommittee
Head of Finance
Head of Acquisitions, Main Library
Head of Acquisitions, Law Library
Support staff from Monographic and Serials Acquisitions – 3 people
Resource People:
Government Publications representatives
Additional staff to review serials issues
Computer Center:
Head of Programming Team (Analysis) and Programmer

- As the analysis of needs began, we quickly determined that different units had different functional requirements. Sydney will be discussing this later.
- We were able to make quite a few compromises whereby units ended up with decisions which they believed were less than most desirable for their group but which served the library effort as a whole. There were also a number of situations where the systems and programming staff agreed to be all things to all people, and thus jumped through complex technical hoops in order to meet individual unit needs.

<u>Processes and Procedures</u>

- A document detailing library needs was created during 2004. Library groups had input plus the systems office and the programming staff. The document was finalized in September, 2004 and turned over to the programming staff.
- Numerous reports were requested to assist in analysis of acquisitions data and would be used to refine the specification.
  - Some of the reports were new reports written for this project.
  - Some reports which were needed for purge analysis were existing reports which needed minor modifications.

  - Some requests were for pre-existing reports to be run on a special schedule.
- The document contained the initial decisions that were made regarding which records were to be purged.

<u>Testing Phase (Test Database)</u>

- Produce Reports
- Analyze Reports
- Run Purges
- Refine Specs
- Retest

<u>Implementation Phase</u>

- Run reports against the Production Database. Some were run early in the process and some were run immediately preceding the purge. Some reports were run after the purge.
- Run Purges against the Production Database

All related work took place from the summer of 2004 to late spring 2005.

Since I think that Sydney and John may answer many of your questions as they proceed with their talks, we should hold questions until the end.