

University Libraries

University Libraries Staff Publications

University of Iowa

10-15-2001

Caption Pattern Records: The 853
Process

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

Caption Pattern Records: The 853 Process
Annual Meeting, International Consortium of Aleph Users
October, 2001

Donna Hirst
Head, Library Automation
University of Iowa, Iowa City, IA
Donna-hirst@uiowa.edu

Lighten the topic:

Eight men are sitting at a table. The boss says: We can confuse them even more if we bury the index in the middle of the May issue with text on its verso. Page it as part of the May issue, skip the June issue, and so on.

We will probably all agree that computer systems have huge potential for assisting us with the mass of serials we receive daily. However, it's also true that even the computer can't cope with some of the permutations that our publishers dream up.

Aleph Factoids

A factoid is basically a slang word. It means a "little fact". This discussion will not be a review of Aleph serials control, but rather just a few quick comments to insure that we are all at the same starting point.

Publication schedules may be created in Aleph in two ways: 1) first, through the MARC21 853 and 853 X fields of the ADM record and 2) second, through the Serials client by using the Aleph Publication Schedule Form.

The Aleph 853 field is used to record:

- 1) Captions that identify the enumeration levels
- 2) Captions that identify the chronological levels
- 3) Codes that define the publication pattern
- 4) Punctuation

The 853 field includes the date when the Schedule was created but not the date when the title began or the date when the subscription began.

Aleph has modified the MARC21 definitions to support system functionality. In MARC21 the parallel fields of 853 and 863 are defined. Similar parallel fields of 854 with 864 and 855 with 865 exist. However, in Aleph the fields are defined as 853/853X pairs, 854/854X pairs, and 855/855X pairs. In MARC 21 these serial fields are stored in a separate holdings record. However in Aleph these serial fields are stored in the ADM record.

The claim interval is in the Subscription Record. This claim interval is added to the date generated from the 853X\$\$\$ to create the "Expected" date, which is picked up by the Issue claim report.

You will hear the term “Publication Pattern”. The publication patterns are the codes (both alphabetic and numeric) which reside in the 853 field which define how the serial behaves.

A typical, straight forward publication pattern might be:

853 \$a v. \$b no. \$u 12 \$v r \$ w m

Subfield a is the first level of enumeration, in this case volumes. Subfield b is the second level of enumeration, in this case number. Subfield u records the number of units in this next higher unit. Subfield u must be numeric and in this case indicates 4 issues per volume. Subfield v contains one of two codes. A code of r indicates the numbering restarts at the completion of the unit and a code of c indicates it is continuous. This subfield r indicates whether the numbering begins again with each volume. The subfield w is usually an alphabetic code and in this case the m indicates monthly. This is a very simple illustration of a caption pattern field. There are additional subfields that allow for more complex statements.

As mentioned there are two alternative ways of creating serial receiving patterns. So far I have been talking about the use of 853+ fields in the ADM record. The alternative procedure is to use the Publication Schedule Form in the Serials Client. The Publication Schedule Form is easy to use but it is not as flexible as the 853 process. The other significant functional advantage of the 853 fields is that the data is created through the application of a national standard (MARC 21) and thus extract programs exist which can allow you to share the data with other libraries.

Rationale for using 853 +

The University of Iowa considered reasons for using prediction with Aleph’s 853 implementation versus the Publication Schedule Form in the Serials Client. Since the 853 mechanism is more closely aligned with the US MARC standard, the 853 process was preferred. We have found the functionality of the Aleph 853 process to be good, and we believe this process is more efficient than using the Publication Schedule Form. Regular maintenance of the 853 data is straight forward. The system is very flexible thus making the function easy to work with.

Nature of Iowa’s file

When the University of Iowa was determining how to create our Caption Pattern Database, we had a few alternatives. At that point we could create our file from Notre Dame’s file, or we could create the file from scratch. We decided to do the latter because we believed that there would not be a large percentage of matches with Notre Dame and there were some variations in practice between the two libraries. When we began creating our file a special administrative library, IOW52, was created so that we could work continuously on caption patterns through the cutover period. When our cutover was complete, Ex Libris copied all the IOW52 data into our production IOW50 Administrative Library. Activating our production Caption Pattern File took 1-2 days after cutover.

In preparation for creating our pattern records we created various Aleph templates, but we discovered as we proceeded that the templates did not prove to be especially helpful. The local templates (based on work with our previous online system) should not be confused with the templates in Aleph, of which we made extensive use in creating our caption pattern database. In our original plan, we attempted to define a number of categories of publication patterns, assigned standardized caption patterns to them and then copied those caption patterns into the records defined in IOW52. This procedure proved somewhat cumbersome and still required individual attention to actually create the schedule on which issues could be received, to confirm its accuracy and appropriateness and to set the 853X\$\$3 correctly. What we actually did with the majority of titles was to create a number of basic templates for a range of enumeration and chronology labels and languages and, using the “expand from template” function, insert them into the ADM record. We created patterns for the most frequent titles first, and searched for the bibliographic record using the system number from our previous system. We assigned 4 people to creating Caption Patterns. From the creation of the initial caption pattern record to the completion of all of our active patterns, staff took 1 to 1.5 months prior to cutover and then another 6 months after cutover. It proved to be very helpful that we began our work with the most frequent titles. One of our East Asian catalogers spent approximately 1 full month creating 1,000 Chinese and Japanese caption pattern records for our collection. Our completed caption pattern file included 11,600 records. With 7-8 months of work and up to four people working on the file, we did approximately 1,500 titles a month.

During the early creation of our caption pattern records prior to cutover, we had to determine how to handle the Issue Date. The Issue date is the date when the issue is expected to arrive and which is controlled by the \$\$3 subfield of the 853X field. We ended up setting the "Issue date" at an arbitrary date that had nothing to do with the real dates that the issues would be expected to arrive.

Our plan was that we would have to check each title individually, anyway, in the procedure that we called "Initial check-in." In that step, we would create the publication schedule and verify that the correct caption pattern had been chosen and the volume and year were correct, as well as resetting the \$\$3 subfield to reflect the publication pattern of the most recent volume or year. Unfortunately, we did not place enough emphasis in the workflow instructions on choosing a "real world" date for the \$\$3 subfield, so we must review them now as part of our claiming procedure. In retrospect, we certainly should have paid more attention to the \$\$3 subfield date.

Language Support

As with a number of processing functions, Aleph is strong in its ability to handle processing in multiple languages. In serials receiving, language for enumeration labels in the vernacular, such as Bd, tom, etc., can be explicitly entered into caption pattern records, and computer assisted language processing can be controlled through tables interacting with the ADM record. The language code in the 008 field of the ADM record controls the language of the labels “month”, “season” and “ordinal abbreviations”.

Ex Libris has created caption pattern language tables for 30 separate languages, from Dutch to Polish to Swedish. You can view these tables at Tab 853-chrono and Tab 853-numbering. Typically serials staff will not be able to see these tables online from their workstations, but systems staff can access these tables.

Tab 853-chrono contains multiple versions of the words “month” and “year” so that library staff do not have to standardize multiple forms of foreign labels. For this table the language code in the 008 of the ADM record triggers the use of the correct label during receiving. The University of Iowa found most of the Aleph tables to be adequate for our processing needs, but there were no Tab 853-chrono entries for Chinese or Japanese. One of our East Asian catalogers created the content for the Chinese and Japanese extensions to the Aleph table, and one of our Systems Librarians added the necessary information online through the tables editor.

Tab 853-numbering works similarly to Tab 853-chrono and supplies Ordinal suffixes for various languages. Column 1 includes the language, column 2 has the number, and column 3 contains the suffix. In release 14.2, English and French are the only languages built into this table, however there is potential for Ex Libris to expand this table or an individual library could do so as well.

Current workflow

As mentioned earlier, the University of Iowa has found the use of the 853 caption patterns to be very efficient. Approximately nineteen out of twenty of our serials have caption pattern records and reliance on prediction has resulted in faster throughput in the serials department and has reduced backlogs.

It probably won't be a surprise to you, but some serials don't lend themselves to prediction. Some examples of problem cases are:

- Serials with 3 issues a month.
- Serials that don't have a fixed number of issues a year.
- If serials are very infrequent, then it's not worth a prediction record.
- Some Elsevier titles have a very irregular pattern. The first example is Brain research, which publishes an indefinite number of volumes each year, two issues each, sometimes combined and sometimes individually, at frequent but slightly irregular intervals. We expect that we will never be able to handle this publication "pattern" using prediction.
- An instance of a title that is potentially predictable is Arthritis and rheumatism, which publishes two volumes per year concurrently. That is, it publishes one volume as a monthly with twelve issues. It also publishes a second volume, with a separate section title, as a bimonthly, six issues/year, but with volume numbers in the same sequence. We could set this title up using the fields for supplements and indexes respectively, but this would violate the spirit of the MARC rules for 854/854X or 855/855X fields. We have been reluctant to permit free-form use of these fields, being concerned that Ex Libris may depend on those field tags to be interpreted literally as Supplements and Indexes in programming for ASHS, the Automatic Summary Holdings Statement. Even if we chose to ignore this potential problem, the patterns wouldn't currently work for prediction because there is no way to tell the system to skip volume

numbers. We would end up with two volumes published concurrently but numbered in parallel, not in sequence.

When a new subscription begins at Iowa, the Caption Pattern Record is created online by a staff member. We don't use worksheets. One person does all this work and it takes 2-3 minutes to create a new caption pattern record. We currently create several new caption patterns a week in the receiving unit. Typically we create new caption pattern records before the title is cataloged. Approximately 30-40 new titles are processed a month.

We have not had a problem in keeping up with maintenance on caption pattern records. When a pattern changes, we hold the maintenance and wait for the next issue to arrive. We do approximately 12 maintenance transactions a day on caption pattern records.

Particular questions have arisen regarding caption pattern creation. We do not create caption pattern records for retrospective pieces; these records are strictly to support receiving. However, we might create caption pattern records if we were to purchase back issues for the first time.

We have not found it to be helpful to use the 854 and 855 fields to create patterns for supplements and indexes. Often these special pieces are received by just adding an issue. In the serial client you can request the Issue Form from the Publication Schedule.

The North American Aleph community has been observing an interesting phenomenon among libraries using the 853 patterns. When a library encounters an especially problematic serial, someone from the library describes the serial title and its abnormalities on the NAAUG listserv (this is the listserv for the North American Aleph Users Group). So far someone else within the Aleph community has already solved the problem, or, more often, has recognized that it cannot be predicted, and sends out the pattern solution to the NAAUG community through the list. This is an interesting but less than ideal way to share cataloging copy. However since another mechanism for sharing these complex records does not exist, this listserv process seems to be working for these particularly difficult cases.

Current ICAU Usage

For the most part, use of the 853 Caption Patterns has been somewhat limited among the worldwide Aleph institutions. Most North American Libraries are using this prediction process and it is beginning to be more popular around the world..

The libraries in Switzerland including IDS and ZVG are using 853 fields.

IDS

Staatsverweltun St. Gallen

ZVG

Four German libraries including KOBV and Volkswagen use 853's.

KOBV

Senatsbibliothek Berlin

Volkswagen

Konrad-Zuse-Zentrum fur Informationstechnik Berlin

Nanterre in France plus other academic libraries use 853's.

In Brazil Pontificia Universidade use caption patterns.

Pontificia Universidade Catolica do Riio Grande do Sul

The University of Westminster in the United Kingdom prefers 853's but also uses Publication Schedules.

Issues related to Import and Export:

As indicated some non-North American libraries are using 853's, but I have been told that no libraries outside North America have used the Downloading process to build their initial files. To date four North American libraries have contributed their Caption Pattern Files for other libraries to use. The process of Ex Libris downloading the data is non-intrusive and very fast.

Current available files for downloading include:

Boston	6742
McGill	9097
Notre Dame	7754
Iowa	11,600

The process of matching the available file with your data involves some fairly sophisticated matching algorithms.

It is possible to get a report of the matches before a download occurs so that a library can determine the number of matches that will occur ahead of time. It is possible to refine the matching algorithm over multiple preliminary runs before the final Caption Pattern load occurs.

In terms of building a local file from another libraries' efforts, it is often helpful to know general characteristics of the file of the donating library. For Iowa approximately 15.5% percent of our 11,600 active serial records are not in English. Somewhat less than 50% of our active serial titles are published outside the United States.

MIT considers its use of the Ex Libris serial pattern database for migration purposes to be one of the most notable aspects of their implementation. This was the first time that Ex Libris capitalized on this growing database as an integral part of a system migration, and its use enabled MIT to go live with serial prediction patterns for about 55% of its active serials. Kim Maxwell, Serials Librarian at MIT, commented that the main advantage of using patterns culled from other ALEPH sites has largely been an education and training bonus. The patterns are there in the correct format, and we have been able to study and learn from these examples in context. For titles without converted patterns, we have been able to quickly build patterns from scratch by example. Our overall match rate was about 55%, though in some subject areas it was as high as 80%. For these titles, preparing the title for check-in is very fast, since the pattern has only to be tweaked rather than created anew.

Managing the import file

As indicated, Ex Libris is currently doing Caption Pattern data loads from four separate contributed files. I have been told that this winter these files will be merged and the duplicates removed so that a maximum number of records could be available efficiently. A process for keeping this data, which is stored at Ex Libris, current will also need to be developed. Iowa's caption pattern file included 11,600 records as of mid-May, 2001 when the file was exported to Ex Libris. As of mid-August, just 3 months later, the file had grown by 737 titles or over 6%. These 737 records are not in Ex Libris' file at the moment. Some process to refresh Ex Libris' file would be desirable.

I have learned that Harvard University has 40,000 caption pattern records, and if these records become available to the Aleph community this will be a very rich resource for new libraries.

What's happening nationally & internationally?

Sharing within the Aleph community is not enough. The United States CONSER Task Force on Publication Patterns and Holdings, managed by the US Library of Congress and OCLC, is progressing toward its goal of standardizing the use of the MARC Format for Holdings Data by system vendors so that libraries can share this data through CONSER records and can also transport the data when changing systems. The cooperative nature of the task forces' approach is evidenced by the participation thus far of fourteen libraries and ten system vendors. Ex Libris participates in these meetings. An FAQ on the project is available at: <http://lcweb.loc.gov/acq/conser/faqpph.html> , home of the online publication, CONSERline.

Aleph Enhancements

A number of Aleph Enhancements are in the development pipeline. Some of these are in the planning stage, and some have already been started. I don't have firm dates of release for any of these.

1. An 'on the fly' Merge option for Fast Check-in.
2. Check-in Loader for multiple subscriptions by vendors such as Ebsco. With this function, check-in is done at the source or by the vendor.
3. Move prediction patterns from the ADM to the HOL record. This enhancement is seen as very desirable by Iowa, and will solve some of our serial control functional problems.
4. Linking the 853/853X patterns to issues and items.
5. Adding \$o to 854 and 855 patterns. This enhancement will put the Ex Libris process in sync with the MARC standard.
6. Using 853 data to automatically populate the issue Description field.

7. Adding alternative chronology using the \$m. This enhancement will also be a subfield tied to MARC21.

8. Linking items to their source patterns (853/854/855).

9. An Expand routine to create a Z39.71 statement from 853/863 fields in MARC 21 BIB or HOL records.

I do not know all the details regarding these enhancements. This list was given to me by Doron Greenspan, Serials expert at Ex Libris, and I'm sure he could provide additional information on any of the items.

So it seems I've taken you hither and yon. It's time to close up. Think about the end of a periodical. Also the end of my talk. I'd like to close with a quotation from T.S Elliot's Four Quartets.

What we call the beginning is often the end
And to make an end is to make a beginning.
The end is where we start from.