
Splitting Services output files into smaller batches for loading

- **Article Type:** General
 - **Product:** Aleph
 - **Product Version:** 15.2
-

Description:

I ran a p_manage_25 service and will now use p_manage_18 to load the records. My manage25 output consists of over 20,000 bib records, and I would like to split the output file into 4 batches to load one at a time.

Resolution:

The site was able to do this with the split command:

```
split -l 150000 yivol096missingnewprint.out yivol
```

There is a minor problem with this method: each line in the input file represents a *field* not a record, so we see that the yivolaa file ends with this:

```
000012534 003 L UPvMLC
000012534 005 L 20000428090245.5
000012534 008 L 770419s1933^^^^^gw^^^^^^^^^^^^^^000^0^ger^d
000012534 035 L $$a(MH)MHAYK91509HU
000012534 035 L $$a(OCOLC)02899654
000012534 035 L $$a(RLG)MAHGAYK91509-B
000012534 035 L $$aMAHGAYK91509-B
000012534 040 L $$aMoSW$$cUPvMLC$$dMH$$dCStRLIN
000012534 050 4 L $$aDD240$$b.F63
000012534 1001 L $$aForsthoff, Ernst,$$d1902-
```

And the yivolab file begins with this:

```
000012534 24514 L $$aDer totale Staat /$$cErnst Forsthoff.
000012534 260 L $$aHamburg :$$bHanseatische Verlagsanstalt,$$c1933.
000012534 300 L $$a48 p. ;$$c22 cm.
000012534 650 0 L $$aTotalitarianism.
<snip>
```

So bib record 000012534 is split across the two files and will not be loaded properly.

Since there were just three records affected in this way (the records at the end of the yivolaa, yivolab, and yivolac files), proceeding with this method and then reloading those three records individually worked.

- **Article last edited:** 10/8/2013