
publish-06: "No such file or directory"; "invalid compressed data"

- **Article Type:** General
 - **Product:** Aleph
 - **Product Version:** 20
-

Problem Symptoms:

In running publish-06 ("Create Tar File for ALEPH Published Records") the following error messages occur, for each record, every time the job runs:

```
tar: XXX01.000123456.PRIMO-FULL.xml: Cannot stat: No such file or directory
tar: XXX01.000123456.PRIMO-FULL.xml: Cannot open: No such file or directory
...
tar: XXX01.000234567.PRIMO-FULL.xml: file changed as we read it
tar: XXX01.000234567.PRIMO-FULL.xml: Cannot unlink: No such file or directory
...
tar: Error exit delayed from previous errors
...
gzip: stdin: invalid compressed data--crc error

gzip: stdin: invalid compressed data--length error
```

And the tar.gz file it produces can not be processed by the Primo pipe.

Cause:

Uncertain.

Resolution:

(The following steps were performed. It's uncertain just which corrected the problem.)

- * set the \$aleph_root/tab/tab_publish_timestamp for PRIMO-FULL back to "201310180559518" (four days ago)
- * stop ue_21
- * delete the bad tar files from /exlibris/aleph/u20_1/xxx01/primo_ongoing_pub/PRIMO-FULL/tar
- * if the the xxx01 util c/1 shows two lib_batch processes running, kill one of them
- * run the xxx01 p_publish_06 with the usual parameters (except, if during the day, with just one process so as to not overload the system)
- * when publish-06 is done, restart ue_21.

The effect of this was to reprocess all of the z00p's created since Oct. 18 -- without errors.

Additional Information

It could be that having ue_21 {which writes z00p ("ALEPH Published Documents") records} running while publish-06 (which reads z00p records) was somehow messing things up.

Other publish-06 error messages:

gzipSkipping to next header

tar: Archive contains obsolescent base-64 headers

tar: Child returned status 1

tar: Error exit delayed from previous errors

Error messages seen in log of Primo pipe:

2013-10-22 12:47:31,647 INFO [t-SplitRecords] [c-StringUtils] - Found non valid xml character:

2013-10-22 12:47:31,647 INFO [t-SplitRecords] [c-StringUtils] - Found non valid xml character:

2013-10-22 12:47:31,647 WARN [t-SplitRecords] [c-OAISplitterCB] - Failed after stripping invalid chars

2013-10-22 12:47:31,647 ERROR [t-SplitRecords] [c-RecordSplitter] - Failed to split records from

XXX01.001081164.PRIMO-FULL.xml

org.xml.sax.SAXParseException: Element type "record" must be followed by either attribute specifications, ">" or "/>".

at org.apache.xerces.util.ErrorHandlerWrapper.createSAXParseException(Unknown Source)

at org.apache.xerces.util.ErrorHandlerWrapper.fatalError(Unknown Source)

at org.apache.xerces.impl.XMLErrorReporter.reportError(Unknown Source)

-
- **Article last edited:** 11/20/2013