
Does p-file-93 expect UTF-8 or MARC8?

- **Article Type:** General
- **Product:** Aleph
- **Product Version:** 19.01

Description:

When records are loaded using p-file-93, are they loaded in UTF-8 or MARC8? We have a file of records from a vendor that are in UTF-8 and we want them loaded in UTF-8. Is this possible?

Resolution:

Hi, Jenn.

The p-file-93 service will process MARC records with either encoding. It checks the LDR byte to determine if the record encoding is MARC8 or UTF-8. If it's MARC8, it calls the OCLC_TO_UTF character conversion routine to convert to Unicode. If the record encoding is UTF-8, p-file-93 calls the OCLC_UTF_TO_UTF character conversion routine. At most Aleph sites don't have an OCLC_UTF_TO_UTF conversion routine defined, so the records load exactly as they are received. If the records are from OCLC, they will have OCLC's own encoding practices. They are legitimate practices, but not what most Aleph catalogers are used to seeing. Consequently, many sites implement a specific OCLC_UTF_TO_UTF routine so that all encodings are consistent. Developers at Harvard and Univ. of Michigan worked together to analyze OCLC's encoding practices and created an OCLC_UTF_TO_UTF configuration. There is a KB record that explains more about this, and how to implement the OCLC_UTF_TO_UTF conversion if you are interested. It's KB #16384-13233.

Randy

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