
SIS error: "Invalid byte 1 of 1-byte UTF-8 sequence"

- **Product:** Alma
-

Question

The SIS job ("Users SYNCHRONIZE using profile Student Information System") fails with error: "Invalid byte 1 of 1-byte UTF-8 sequence.", or something like, "An invalid XML character (Unicode: 0x0) was found in the element content of the document".

Answer

Unicode characters in general needs to be output in UTF-8 encoding.

Check the input file for special characters (non-Latin diacritics). Open the file using Notepad++ and validate with the XML Plugin, or thanks to another XML validation tool.

When a file is uploaded with special characters in XML, the special characters should be escaped as they are in HTML. This is also the case when uploading a bibliographic file via an import profile.

For example,

Instead of → (Unicode Character 'RIGHTWARDS ARROW' (U+2192)), use HTML `→`

Examples and further information is in the Blog "[Accented characters in SIS import and export](#)".

Additional Information

After fixing the XML file, zip it and place on the FTP directory. Wait for the next SIS job to run, or run manually:

1. Navigate to: Configuration > General > Integration Profiles
2. Edit the profile for: Student Information System
3. Select the "Actions" tab
4. In the "SYNCHRONIZE" section click: Run.

-
- **Article last edited:** 19-JAN-2022