**UTF-8 Special Character Handling**

**May 2020 Resource Management**

Alma offers a new method of handling UTF-8 special characters (with diacritics). UTF-8 special characters may be represented in both the composed or decomposed version of the character in bibliographic or authority records. Now, you have the option to configure your system with normalize on save always to save the composed version of special characters. This may be especially useful to implement to avoid the use case where multiple records are changed due to their conversion to composed representation. Such records are marked for preferred-term correction (PTC) and cause heading updates (the only difference is the composed/decomposed nature of a special character).

**Implications for Discovery**

The previous handling of UTF composed/decomposed representation caused unnecessary heading updates. The PTC kept updating headings whose only difference was the composed/decomposed characters. When nothing was changed in a GND authority other than changing a decomposed letter to a composed letter, the change triggered the following redundant steps:

* Preferred term correction
* Publishing to Aleph & Primo

Once this development is active in the system, if the correction only concerns the change of character format from decomposed to composed, another publishing will take place, but this will be a one-time publish.

With this new development, we reduce the number of records that are redundantly published and the number of heading updates. Once this new development is active, slowly the number of records with composed characters will grow. This will reduce the amount of unnecessary records that the Authority Task Control List needs to handle, and remove the unnecessary load on the system.

**Implications for Authority Control**

When multiple records are marked for preferred-term correction and the only difference is the composed / decomposed nature of a special character, then the system will again create the linkage between the bibliographic and the authority record (same as clicking F3, but here it will happen automatically, without user involvement). However, the user will see no change in the representation on screen, and will not know that such took place.

**Management**

If you are interested in this new method of handling UTF-8 special characters, contact Support to configure this for your system. In a consortia, each individual IZ needs to submit a request to Ex Libris to configure the relevant system customer parameters to have the special character always saved in its composed format. Note that when this is implemented, it applies to all new and modified records.

Two implementation options exist:

* Each time a new or modified bibliographic or authority record is saved in Alma, then any decomposed special character is replaced with and saved as the composed version of the special character. This applies to all workflows for saving records such as manually, via import, with preferred-term correction, and so forth.
* You can indicate a particular vocabulary, so that any decomposed special character is replaced with and saved as the composed version of the special character for a new or modified record when the record is saved, but only for the vocabulary specified in this parameter.