

Sources: What types of encoding does Muse use?

We unify all the information from all the sources into Unicode using UTF-8, and we pass it from ICE and all of our Bridges as UTF-8 and XML-encoded entities. However, we are not limited to extracting data just from UTF-8 bytes, and we should not extract bytes as UTF-8 when it is not appropriate. All our extraction in Muse is to UTF-8 but from a wide range of encodings (including UTF-8).

We accommodate this large number of encodings, detailed in the documentation, when we extract data; only then is it converted into Unicode (with UTF-8 encoding) on our side.

Unique solution ID: #1051

Author: Administrator

Last update: 2013-09-02 11:34