

Morpho - Bug #1009

Problem with non-ascii characters

03/11/2003 02:25 PM - Dan Higgins

Status:	Resolved	Start date:	03/11/2003
Priority:	Normal	Due date:	
Assignee:	Dan Higgins	% Done:	0%
Category:	morpho - general	Estimated time:	0.00 hour
Target version:	Postpone	Spent time:	0.00 hour
Bugzilla-Id:	1009		
Description			
<p>One of the students entering data added a bounding geographic lat/long value as deg/minutes/seconds, using the special PC character for degrees. This is not the format described for lat/long (one should use fractional degrees), but the use of the special degree character (which has a character code above 127) results in a UTF-8 error of 'illegal one character unicode character' when the saved data is parsed.</p> <p>It is not clear just where the problem occurs. The degree symbol is in the data that is successfully submitted to metacat, but problems occur later when editing. (Maybe the problem is with Xalan?)</p>			
Related issues:			
Is duplicate of Morpho - Bug #1015: problems with 'special' characters		Resolved	03/20/2003

History

#1 - 04/02/2003 11:45 AM - Dan Higgins

- Bug 1015 has been marked as a duplicate of this bug. ***

#2 - 04/02/2003 11:54 AM - Dan Higgins

The problem was traced to Windows characters outside the normal ASCII range pasted into Morpho fields. The characters have the upper bit set (i.e. are in the range of 128-255).

The fix was to change the 'normaize()' method which is used to check for illegal XML characters in many places within Morpho. This method replaces illegal characters like '&' with entities like %amp; The method checks for characters with code greater than 127 and replaces the text with equivalent entities.

Note that the 'normalize()' method was defined in 8 places within Morpho! A new utility class (XMLUtil) was created and used in the various other classes.

#3 - 03/27/2013 02:15 PM - Redmine Admin

Original Bugzilla ID was 1009