

Encoding and mis-encoding: latin (iso-8859-1) and utf-8

Dobbel-errors

Letter	latin mis-encoded as-utf	mis-encoded utf shown-as-latin	Explanation	Reference
æøå/ÆØÅ	= 0xFFFD	ï¿½ = 0xEF 0xBF 0xBD	illegal UTF-8 sequences.	stackoverflow.com 1
æ				
ø	(0xFFFD)	ï¿½ (0xEF 0xBF 0xBD)		
å				
Æ				
Ø		ï¿½		
Å				

Letter	utf mis-encoded as-latin	mis-encoded latin shown-as-utf	Explanation	Reference
æøå/ÆØÅ	2 bytes often Ã and ...	ï¿½ = 0xEF 0xBF 0xBD	UTF-8 multi-byte string is interpreted with a single-byte encoding	stackoverflow.com 2
æ				
ø	Ã,			
å	Ã¸			
Æ				
Ø				
Å				