In SIS we allow unicode characters. We understand that FDSN StationXML uses ASCII as a standard. If in our system the string is "Puerto Peñasco" and we want to write it out in FDSN StationXML that could be converted to dataless can we write it in the FDSN StationXML out as:

Puerto Peñasco

And the converter write it as whatever is appropriate in dataless?
I believe dataless does not accept & -- but at least users of FDSNStationXML would be able to get back to unicode format.

Please advise.

Ellen

---

**Subtasks:**

Bug # 407: 8bits ascii char in dataless  
Closed

---

**History**

#1 - 07/20/2015 11:33 AM - Yazan Suleiman
Can you please provide an example xml file

#2 - 07/20/2015 11:54 AM - Chad Trabant
  - Assignee set to Yazan Suleiman

Hi Ellen,

As you state, the non-ASCII characters in StationXML description fields are not allowed. That being said, I would expect this to be a common issue for conversion.

The safest thing for SIS would be to ASCII-ify the strings before giving them to the converter so you do it the way you want.

We will investigate how to map the UTF-8 strings to ASCII when creating dataless SEED, in order to try and avoid making invalid SEED and warn the user of any conversion.

I do not see how we can ever reverse that process without making the ASCII version of the string very ugly by including escaped Unicode in the strings.