DR 16-0022 — Shared ML: Escaping strings in ST\_Xstring

Status: Open

Subject: Shared ML: Escaping strings in ST\_Xstring

Qualifier: Request for clarification

Submitter: Charlie Clark Organization: Ecma/Clark Consulting & Research

Contact Information: [charlie.clark@clark-consulting.eu](file:///E:\Std\SC%2034\WG4\DRs\Log\2014\charlie.clark@clark-consulting.eu)

Submitter’s Defect Number: None

Supporting Document(s): None

Date Circulated by Secretariat: 2016-12-06

Deadline for Response from Editor: 2017-02-06

IS 29500 Reference(s): 29500:2016, Part 1, §22.9.2.19, “ST\_Xstring (Escaped String), p. 3797”

Related DR(s): None

Nature of the Defect:

I've got a cell with the value "SW\_x3850\_CPU" in it. How should this be serialised?

§22.9.2.19, “ST\_Xstring (Escaped String)” says:

For all characters that cannot be represented in XML as defined by the XML 1.0 specification, the characters are escaped using the Unicode numerical character representation escape character format \_xHHHH\_, where H

represents a hexadecimal character in the character's value. [*Example*: The Unicode character 8 is not permitted

in an XML 1.0 document, so it must be escaped as \_x0008\_. *end example*]

But it's not clear from this if all such combinations should be escaped?

or just those in the range [001-031]. Excel itself handles such sequences by escaping the first underscore but unfortunately other consumers such as OpenOffice do not remove the escaping so I think this needs clarifying.

I've recently received an SML file that looks like it was created in LibreOffice and I'm having trouble working out how to handle the named styles these are defined as:

<cellStyles count="6">

<cellStyle name="Normal" xfId="0" builtinId="0" customBuiltin="false"/>

<cellStyle name="Comma" xfId="15" builtinId="3" customBuiltin="false"/>

<cellStyle name="Comma [0]" xfId="16" builtinId="6" customBuiltin="false"/>

<cellStyle name="Currency" xfId="17" builtinId="4" customBuiltin="false"/>

<cellStyle name="Currency [0]" xfId="18" builtinId="7" customBuiltin="false"/>

<cellStyle name="Percent" xfId="19" builtinId="5" customBuiltin="false"/>

</cellStyles>

So, there are only six named styles, but there are 20 "master style records". Is this kind of discrepancy intended? What about styles with indices of 1–14?

As with any kind of relationships between elements, it's impossible to write a constraint in the schema, but I understand the relationship between the two should be 1:1. This is important in any situation where you try and combine key-based referencing with positional and any kind of deduplication.

Solution Proposed by the Submitter:

None

Schema Change(s) Needed:

No

**Editor’s Response:**

None

Changes to Part 1: N Part 2: N Part 3: N Part 4: N