DR 17-0029 — SML: Quoting of Worksheet Names Containing Periods

Status: Further Consideration Required

Subject: SML: Quoting of Worksheet Names Containing Periods

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: 2017-10-13

Deadline for Response from Editor: 2017-12-13

IS 29500 Reference(s): 29500:2016, Part 1, §18.17.2.3, “Cell References”, pp. 2049–2053

Related DR(s): None

Nature of the Defect:

Should worksheet names with periods be quoted in cell references?

I recently come across problems relating to the quoting from sheet names.

According to the specification these need quoting in cell references if they contain the following special characters:

**§18.17.2.3 Cell References**

sheet-name-special=

sheet-name-start-character-special,

[ [sheet-name-characters-special], sheet-name-end-character-special] ; sheet-name-start-character-special=

character (apostrophe | "\*" | "[" | "]" | "\" | ":" | "/"

| "?") ;

(\* any character, including operator, except ', \*, [, ], \, :, /, or ? \*) sheet-name-end-character-special= sheet-name-start-character-special ; sheet-name-characters-special= sheet-name-character-special, {sheet-name-character-special} ; sheet-name-character-special= apostrophe, apostrophe | character (apostrophe | "\*" | "[" | "]" | "\" | ":" | "/"

| "?") ;

(\* any character, including operator, except \*, [, ], \, :, /, or ? \*)

However, it seems that Excel at least also quotes sheet names with periods in them: Sheet.1!$A$1 becomes 'Sheet.1'!$A$1.

Does this mean that the specification should be updated to include "." as sheet-name-character-special?

Furthermore, it looks like Excel struggles with some sheet titles, for example you cannot add charts to sheets named "an12.103.1". Is this documented anywhere?

Solution Proposed by the Submitter:

None

Schema Change(s) Needed:

No

**Editor’s Response:**

**2017-11-06 Aarti Nankani:**

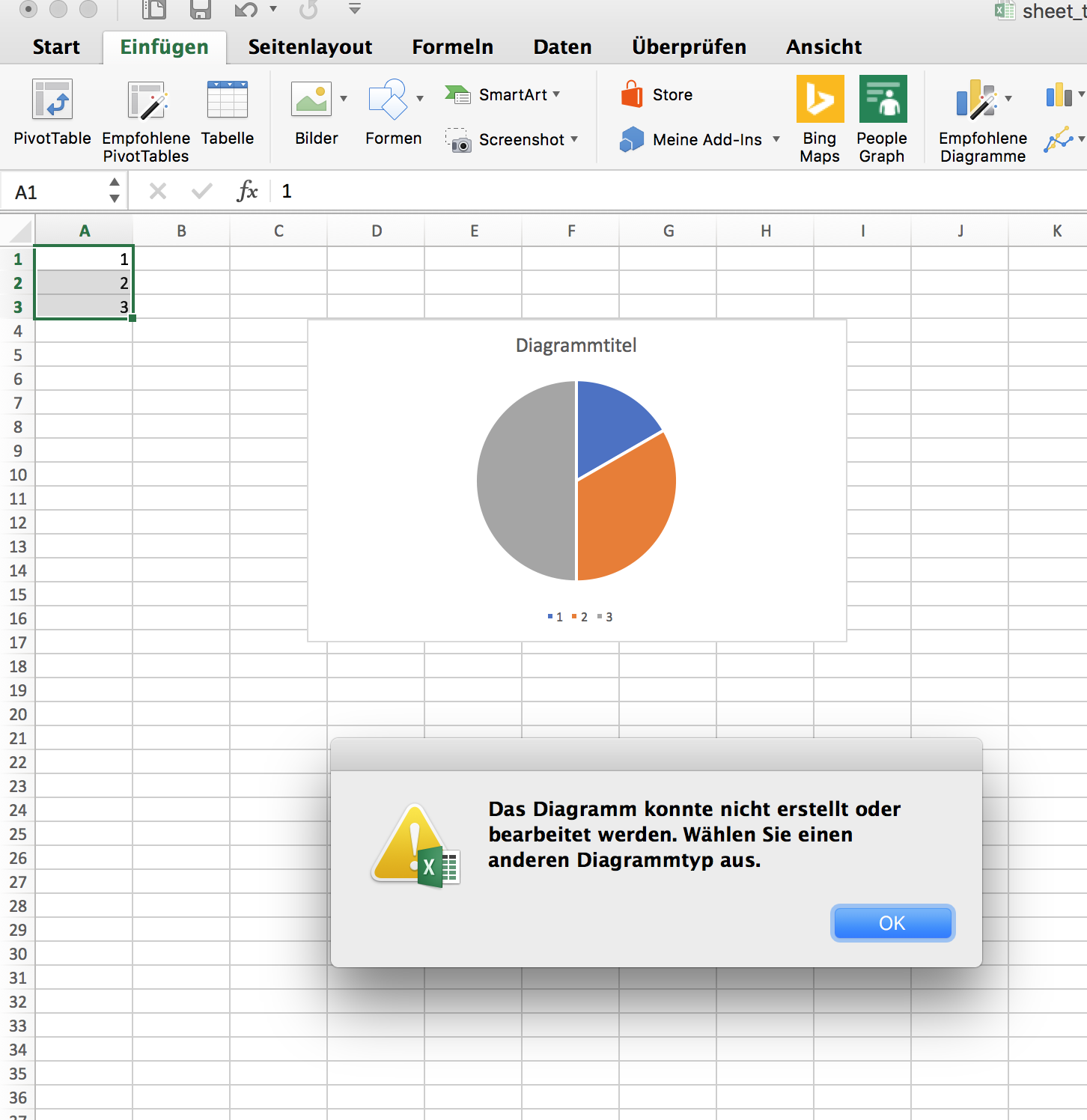
Response from our product team:

Excel requires quotes on sheet names that have '.' or 0-9 as the first character of the name. We do not add quotes the name when those characters appear in any other position within the name which is different from what submitter claims our behavior to be. We tried but were unable to reproduce the behavior described in the two assertions by the submitter (1) that quotes are added by the product and 2) that chart can't be added to sheets with special names). Both scenarios worked as expected and as documented, so we would need more specific reproducible steps to investigate further.

**2017-11-14 Charlie Clark:**

I can do a screen capture for this. It reliably doesn't work. I'm enclosing a screenshot of the source of the enclosed file in which the sheet title was changed \*after\* the chart was created. Quotes are clearly in use even though the sheet title starts with letters.





Below is a screencast (also uploaded to <https://www.dropbox.com/s/9w1dt1btuxchubz/excel_sheet_title.mp4>).



I hope it's self-explanatory. The sheet renaming shows that there is some code in Excel for handling the quotes and also preventing the creation of objects with sheet references in it that doesn't know about the quoting.

Confused? Not half!

**2017-11-14/16 Geneva Meeting:**

Assigned to Aarti.

**2018-05-09 Charlie Clark:**

Here my update on the question:

Within source XML cell references which characters must be quoted or require quoting of the worksheet name. Is the current definition sufficient? Are there other object names where this applies? Ideally it would be possible to provide a regex for testing this and I would be happy to contribute to this.

Also, [for] interoperability would it be possible for Microsoft to provide a list of reserved names that should be avoided for worksheets, defined names, tables, etc.? This could be extended by other vendors as required.

The example originally provided that caused problems in Excel 2016 seems to have been resolved, but not necessarily for other versions. I cannot find a reference to this in any of the release notes such as

<https://support.office.com/en-us/article/release-notes-for-office-2016-for-mac-ed2da564-6d53-4542-9954-7e3209681a41?ui=en-US&rs=en-US&ad=US#bkmk_nov2017>

It would be very useful if Microsoft could identify which version contains the fix. This is not a request for any internal or privileged information, but would be useful to point users to. It would also document the important work of the working group and Microsoft's commitment to it.

**2018-09-04 MS Experts:**

For the new questions in Charlie’s last comment, it would be better to open a new request since they are broader than the original request and represent an increase in scope.

In his most recent comment, Charlie mentioned the original ask has been resolved. One clarification that can be made is to the EBNF to indicate that the first character of an unquoted name should be further restricted to also exclude ‘.’ and 0-9.

*[Ed: There were additional questions from Charlie, but we’re still waiting for feedback on those.]*

**2018-09-05 Charlie Clark:**

I agree that a detailed discussion of release notes could be a separate issue. But it remains relevant to this issue as to whether it can be acknowledged that a particular release solved the problem: was the issue as described a problem with the XML or MS Excel 2016?

Thanks very much for this, but, unfortunately it seems, alas, still incomplete. For example, if the % character is used, the worksheet name also seems to require quoting. If the specification is still incomplete, then perhaps the solution is to require quoting for all worksheet names and remove additional distinctions.

Also, I may have missed it, but shouldn't the use of EBNF be explicit?

Does a sample implementation and test suite exist for this? I think this issue highlights the importance of being able to validate such rules, preferably at the schema level. Indeed, given the use of XML schema it seems odd to have additional requirements that cannot be checked against the schema. The EBNF might make sense at the formal level but should these not be accompanied where possible by relevant regular expressions which can be enforced in the schema?

Also, who edits and validates the EBNF?

Furthermore, there has been no information provided about reserved names for worksheets, i.e., names that while formally correct, may cause problems between implementations and which should, therefore, be avoided.

**2018-09-06 Charlie Clark:**

I've done some more reverse engineering and extended the list of characters that seem to need escaping. This is probably not exhaustive, but I think having a working regex may allow us to make progress on this.

^(\d|\.)|([-%'(){};\s\"]

**2018-09-12 MS Experts:**

Our intent by “names should be further restricted to also include ‘.’ and 0-9” is that those items need to be added to the list of characters that require quotation marks, not that they replace the existing list in the documentation. The existing list in MS-OI29500 §2.1.861 does already include %. To the question around which version the issue existed in and was resolved in, since we have been unable to reproduce the issue internally, we are unable to comment on which versions experienced the behavior.

**2018-09-14 Charlie Clark:**

Looks like some confusion: I didn't expect [0-9] and period to replace the existing list of characters that mean the worksheet title needs quoting, but to extend it. The same with % and the other characters I discovered.

As to where the list is kept then I don't think it should be in the implementers notes as the specification is extremely precise. My preference would actually be to mention the need to quote in the sheet element of the workbook object with a link to the rules in formulae. I would suggest that the narrative prose include the list of characters that mean the worksheet name needs quoting, and whether quoting can be generally used (might be the most pragmatic approach). The regex could be useful here as anyone can run with it rather than implementing the EBNF, but I think the format also needs updating.

The error is still present in Excel 2010 for Windows and Excel 2011 for Mac, and if I still had the install image for Office 2016 for Mac I could probably reproduce it. I was able to demonstrate on November 2017 and I run the "office slow" updates and it was fixed when we looked in March, so basically a build from October 2017 should demonstrate it.

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