DR 17-0012 — WML: PAGEREF field

Status: Open

Subject: WML: PAGEREF field

Qualifier: Technical defect

Submitter: Richard ?? Organization: TC45

Contact Information: [p1@originsystems.co.za](mailto:p1@originsystems.co.za)

Submitter’s Defect Number: none

Supporting Document(s): none

Date Circulated by Secretariat: 2017-04-21

Deadline for Response from Editor: 2017-06-21

IS 29500 Reference(s): 29500:2016 Part 1, §17.16.5.45, “PAGEREF”, p. 1235

Related DR(s): none

Nature of the Defect:

Please see the following assertion:

Due to the static nature of the PAGEREF field definition (ECMA Part 1 - 17.16.5.45), it is impossible to build a fully consistent OOXML document containing a table of contents TOC without either:

1. first loading the document in MS Word and manually updating the TOC \*\*; or
2. rewriting the MS Word pagination functionality entirely in order to mimic the document rendering

\*\* where the TOC can be manually updated in the following ways by a user:

1. Via selecting the “Yes” button to the on-load “This document contains fields that may refer to other files. Do you want to update the fields in this document” popup (which is triggered using the w:dirty attribute or w:updateFields element) ; or
2. Via manually selecting the TOC once the document has been opened and right clicking “update TOC

While web commentary I have seen appears to bear the above assertion out, I am really hopeful that it is incorrect. I say this as it seems very much at odds with the concept of an open standard document format that it be impossible to assemble a fully consistent document without subsequent modification via a proprietary application.

Solution Proposed by the Submitter:

None

Schema Change(s) Needed:

No

**Editor’s Response:**

**2017-05-23 Aarti Nankani:**

From MS Experts:

The OOXML format represents the content of the document with complete fidelity and is documented completely for interoperability. The format does *not* represent the fixed layout representation of the document. (Document formats such as XPS or PDF are intended to represent fixed layout.)

A TOC field on OOXML (Part 1, §17.16.5.68) is a field instruction that directs the client application (or user agent) to dynamically (or on demand) calculate a Table of Contents based on the parameters in the field. Since the content of the document can change (either by the client application, or by another application), the results of the TOC can (and often does) get out of sync with the contents of the document. It is the responsibility of the client application to determine if/when to update any cached field results (Part 1, §17.16).

Fields in OOXML also allow the result of a field calculation to be stored in the XML (again, Part 1, §17.16). This allows client applications to persist the results of a field calculation in the file. There is no requirement that a client application persist a field result (other than the syntactic requirement that the field result be present as specified in the standard). If a client application chose to persist an empty field result and always recalculated the field at render time, the resulting file (with an empty field result) would still be a “fully consistent OOXML document.”

Since the layout algorithms used to by Microsoft Word to layout a document created from an OOXML file are unique to Word, the field results for a TOC field that Word stores in the OOXML file are specific to the pagination results calculated by Word’s layout. Moreover, if the content of the document changes after Word calculated the TOC field last, the TOC field results may not include all of the document content. In all of these instances, the document is consistent and valid, but the client application will need to recalculate the TOC field in order to include all document content.

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