Google position on DIS-29500 (OOXML)

Background

On February 25th to 29th 2008 a Ballot Resolution Meeting (BRM) was held in Geneva to attempt to finalize the text of the DIS-29500/OOXML ECMA proposed ISO standard for editable office documents.

Although the thirty-two National Bodies present spent commendable time and effort trying to create fixes and review the proposed changes to correct the flaws in the proposed standard, the large number of reported problems (over one thousand) meant it was impossible to adequately discuss them in the time allowed. In the words of the US Head of Delegation, Frank Farance:

"Eighty percent of the changes were not discussed…It's like if you had a massive software project and 80% of it was not run through QA. …It's a big problem…I've never seen anything like this, and I've been doing this for 25 years."
— **Frank Farance** (US Head of Delegation) in Computerworld[1]

The Malaysian National Standards body took the unusual step of issuing a press release stating:

"*Malaysia's Department of Standards (STANDARDS MALAYSIA) recently found the Draft ISO standard, ISO/IEC DIS 29500: Office Open XML (OOXML) specification for electronic document formats, had the majority of its technical issues still not addressed satisfactorily.*"

which may be read in full at their web site.[2]

Large numbers of corrections to DIS-29500/OOXML were not discussed at all, but merely voted on as a blanket "approval", in contradiction to JTC1 (Joint Technical Committee 1, Information Technology) Directives, which should have been followed in the BRM. These blanket approvals make substantial changes to the text of the proposed standard, without proper review.

It is clear to Google that the "Fast Track" process, attempting to rapidly approve DIS-29500/OOXML, has been a failure and we feel the proposed standard should be re-submitted in the normal standards process to allow adequate time for correction and harmonization with the existing ISO standard, ISO/IEC 26300 (ODF).

Google respectfully requests that National Standards Bodies vote "no" on the fast-tracking of DIS-29500 as an International Standard. Standards Bodies that previously voted "yes" or "abstain" are still able to change their votes to "no" in this critical period.

The following provides a brief overview of why we believe it is important to National Standards Bodies to vote "no" on the fast-tracking of DIS-29500/OOXML.

DIS-29500/OOXML fails to meet its design goals

1 http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=9065958
2 http://www.standardsmalaysia.gov.my/press%20release%204%20March%202008.pdf

The design goal of DIS-29500/OOXML has not been reached. Quoting from the ECMA white paper "Office Open XML Overview": "*OpenXML was designed from the start to be capable of faithfully representing the pre-existing corpus of word-processing documents, presentations, and spreadsheets that are encoded in binary formats defined by Microsoft Corporation.*"[3]

This goal has not been met. The DIS-29500/OOXML text fails to specify a canonical mapping between the binary formats and DIS-29500/OOXML. Requests to provide such a mapping were made by multiple National Standards Bodies, both before and during the BRM, but this has not been provided by ECMA or Microsoft.

Without such a mapping as part of DIS-29500/OOXML there is only one implementor who will be able to "*faithfully represent .. binary formats defined by Microsoft Corporation.*" and that is Microsoft Corporation. Without the canonical binary mapping, every DIS-29500/OOXML implementor will end up with a different representation when translating the same binary document.

DIS-29500/OOXML has weak conformance clauses

Conformance clauses determine when an implementation meets the definition of the standard. As such they are essential to provide effective interoperability. Without well-written conformance clauses a wide range of applications that minimally conform to the standard can claim full conformance, leading to customer confusion.

This becomes particularly relevant when standards provide mechanisms for extension. Interoperability depends upon the documentation of such extensions. For this reason, other standards mandate documentation of extensions as part of their conformance clauses, e.g. ISO/IEC 9899:1999 for the C programming language, which states: "*An implementation shall be accompanied by a document that defines all implementation-defined and locale-specific characteristics and all extensions.*"

By contrast DIS-29500/OOXML lacks requirements to document vendor-specific extensions, and has many "implementation defined" features, some of which affect what a user would see on viewing or printing a document. These holes in the specification allow "conforming" applications to produce a wide range of incompatible files with no guarantee of fidelity of representation for customers.

The conformance clauses of DIS-29500/OOXML regulate conformance in name only. The lack of effective conformance clauses makes DIS-29500 unsuitable as an International Standard.

DIS-29500/OOXML retains binary formats

The BRM split DIS-29500/OOXML into "strict" and "transitional" schema's. The "transitional" schema is intended to be used when translating binary documents, the "strict" schema for new documents. The intent is that "transitional" documents will be replaced over time. However, as many DIS-29500/OOXML documents will be created by editing the large number of existing binary documents the word "transitional" is misleading. Such "transitional" documents are likely to live as long as the DIS-29500/OOXML specification is in use, propagating the idiosyncrasies of the legacy translations that DIS-29500/OOXML was meant to replace.

In addition, both the "transitional" and "strict" schema's allow embedded binary elements that are of

3 http://www.ecma-international.org/news/TC45_current_work/OpenXML%20White%20Paper.pdf), under "*Purposes for the Standard*"

“application defined” type, including existing binary Microsoft Word Documents or Microsoft Excel Spreadsheets. These documents are embedded directly into a DIS-29500/OOXML document in legacy binary format by the only existing implementation of DIS-29500/OOXML, Microsoft Office. Such documents with embedded binary format documents would still meet the “strict” schema definition of DIS-29500/OOXML, which makes the term meaningless.

The ISO standardization of DIS-29500/OOXML will not mean an end to document interoperability problems, but would perpetuate them long into the future.

Conclusion

In our previous papers posted here :

http://www.odfalliance.org/resources/Google%20XML%20Q%20%20A%20(2).pdf

http://www.odfalliance.org/resources/google-response.pdf

Google already concluded that DIS-29500/OOXML is unnecessary, application specific, and of insufficient quality to be fast-tracked as an International Standard. Despite the hard work of National Standards Bodies in the BRM we still believe this to be the case, due to the inappropriateness of submitting a six thousand page unfinished specification into the “Fast Track” process.

Other issues remain with DIS-29500/OOXML that were out of scope for the BRM, such as the legal uncertainty around Microsoft's Open Specification Promise and the ability for code implementing DIS-29500/OOXML to be released under a Free Software/Open Source license.

The large number of changes approved without discussion at the BRM means National Standards Bodies will be unable to see a finished specification before the final deadline for revising their vote. Rather than approving an International Standard sight-unseen, Google would recommend a “no” vote on the fast-tracking of DIS-29500/OOXML.