ISO/IEC JTC 1 / SC 34 / WGs 1, 4 and 5 in Helsinki 14-17 June 2010

What ISO/IEC JTC 1 / SC 34 / WGs 1, 4 and 5 means?

In short:

- ISO means **International Organization for Standardization** [1]
- IEC means **International Electrotechnical Commission** [2]
- **JTC 1** is Joint Technical Committee 1 of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) [3]
- **SC 34** is the subcommittee 34 of the JTC1, called **Document Description and Processing Languages** [4]
- WGs 1,4 and 5 are the working groups of the subcommittee 34

- WG 1: Markup Languages
- WG 4: Office Open XML
- WG 5: Document Interoperability

I attended these working groups (1,4 and 5) meetings as **a concerned citizen of Finland**, wary of Finnish public sector spending hundreds of millions of Euros on document processing in the near and distant future.

Best Regards,

Jukka Rannila
citizen of Finland

signed electronically

1 http://www.iso.org/
2 http://www.iec.ch/
3 http://www.jtc1.org/
4 http://www.itscj.ipsj.or.jp/sc34/

DISCLAIMERS

This document discusses ODF and OOXML

This document discusses ODF and OOXML, which are highly controversial issues, and therefore several disclaimers are needed in order to keep things less messy, or more clearer.

Also public persecution is always possible for persons, who publish opinions about ODF and OOXML, and therefore several disclaimers are also needed.

Legal disclaimer:

All opinions in this opinion paper are personal opinions and they do not represent opinions of any legal entity I am member either by law or voluntarily. This opinion paper is only intended to trigger thinking and it is not legal advice. This opinion paper does not apply to any past, current or future legal entity. This opinion paper will not cover any of the future changes in this fast-developing area. Any actions made based on this opinion is solely responsibility of respective actor making those actions.

Political disclaimer:

These opinions do not represent opinions of any political party. These opinions are not advices to certain policy and they are only intended to trigger thinking. Any law proposal based on these opinions are sole responsibility of that legal entity making law proposals.

These opinions are not meant to be extreme-right, moderate-right, extreme-centre [5], moderate-centre, extreme-left or moderate-left. They are only opinions of an individual whose overall thinking might or might not contain elements of different sources. These opinions do not reflect past, current or future political situation in the Finnish, European or worldwide politics.

These opinions are not meant to rally for a candidacy in any public election in any level.

Content of web pages:

This text may or may not refer to web pages. The content of those web pages is not responsibility of author of this document. They are referenced on the date of this document. If referenced web pages are not found after the date when this document is dated that situation is not responsibility of the author. All changes done in the web pages this document refers are sole responsibility of those organisations and individuals maintaining those web pages. All illegal content found on the web pages referenced is not on the responsibility of the author of this document and producing that kind content is not endorsed by the author of this document.

Use of broken English

This text is in English, but from a person, whose is not a native English-speaking person. Therefore the text may or may not contain bad, odd and broken English, and can contain awkward linguistic solutions.

5 Based on the Finnish three-party system there is phenomenon called extreme-centre in Finland.

Not a technical advice

This text is not meant to be a technical advice. All technical opinions in this text may contain mistakes, and therefore caution is advised to all readers.

Trademarks

All trademarks are owned by their respective owners, and there is no need to litigate author for misusing trademarks.

AS IS

This text is “AS IS”, i.e. reason is needed when reading the text. Public persecution is always possible, when publishing opinions about ODF and OOXML, and therefore the author asks mercy from the readers.

COPYRIGHT

[The opinion starts on the next page]

1. Prologue, 13 June 2010

I attended the Party [6] congress 11-13 June 2010, and in general people were very **ignorant** of Information Technology.

I had submitted a initiative/proposal to be discussed in the congress.

The initiative/proposal was overruled in the congress, and people were complaining, that I was too technical to the political dummies.

2. Epilogue, 18 June 2010

I attended ISO/IEC JTC 1 / SC 34 / WGs 1, 4 and 5 meetings 14-17 June 2010, and in general people were very **knowledgeable** of Information Technology.

I had the feeling, that I was too non-technical, and I was just a political dummy in a technical meeting.

3. In between, 14-17 June 2010

Seriously? What happened during the meetings in 14-17 June 2010?

THIS story is presented in **chronological order**, and there might be other documents, which presents discussions in other order, e.g. documents by Rex Jaeschke.

4. The WG 1 meeting 14 June 2010

I came early to the meeting, and I was expecting at least ten people to show up, since the conference room was smaller one.

First came Juha Vartiainen of SFS [7] (Suomen Standardisoimisliitto SFS ry), and we discussed generally about the forthcoming week. I have never met Juha before, and we discussed about the Finnish mirror group business (mirror group 306 of SFS, i.e. document formats).

Juha Vartiainen instructed, that people in the WG 1 group have been working together for years.

6 The Party meaning a political party in Finland. It is meaningless to this opinion, which political party the author is supporting, since this opinion is about ODF and OOXML – not about politics in general. No need to offend anybody, if the reader is supporting another party. There is enough offence and defence in ODF and OOXML.

7 http://www.sfs.fi/

Well. After all Murata-san (MURATA Makoto) and Alex Brown arrived, and we had a meeting with four gentlemen.

Juha was right. The discussion started right from the previous face-to-face meeting, and it was rather hard to orient in the discussion without adequate technical knowledge. There was several items in the ballot, and those issues were dismissed. This was so self-evident to the participants, that I did not ask anything about these ballots.

We discussed about the Finnish proposal, which had been distributed for comments. The comments were not that supportive, and we discussed about this proposal. The Finnish proposal had been too sketchy, and therefore it was hard to comment on that.

> The Finnish proposal for document format had been discussed in the SR 306 meetings.
>
> I had distributed my idea of Document-Program to Juha, and most probable way it did not gather much enthusiasm. But it was an idea, and it was sketchy too.

The general conclusion was, that it is easier to persuade national standardisation organisations, when there is more concrete proposal with good introductory texts. The general conclusion is, that totally new standard from scratch is not a feasible way, and there should be something concrete to start with. The Finns are encouraged to work on some real proposal, not with some nice-to-have ambiguous proposal, and then national bodies can give their response to that proposal. It was concluded, that Finns have several national variations to choose from, and the problem is selecting and amending a real proposal.

The next SC 34 plenary was discussed, and the problem is to have a reasonable timetable to all Working Groups (WG).

"RELAX NG Best Practices" was an item for the meeting.

One problem is/was, that programs/programmers use some default values, or even hard-code those default values to documents. Also one problem is that people do not follow complicated rules, and there is no need for more complex rules.

> Personally I understand, why programmers use default values, since document standards are highly complicated per se, and actual implementation is even more complicated task.

"ISO/IEC 19757-2 and ISO/IEC 19757-12" was an item for the meeting.

There was/is the need to keep versions 1 and 2 alive as standards. Therefore there might be need to create ISO/IEC 19757-12 in order to keep versions 1 and 2 as published standards. The ISO/IEC policies of standard versioning and numbering causes the need for possible ISO/IEC 19757-12, since ISO/IEC policy mandates only one active standard version. Then there was discussion about

backward and forward compability of ISO/IEC 19757 versions 1 and 2. Like in all conversions, there are some problems with this. The problem arises, when there are documents complying version 1 and version 2, and validators should distinguish and/or convert different versions.

Information Technology Task Force (ITTF) policies

This issue raised quite a lot of discussion, since that ITTF policy is to have Word 97 or PDF documents. Inside SC 34 there has been a separate/specific technique to document standards, and conversion to Word 97 might cause some problems. There are not much volunteers to create a converter – Yet another converter??

(Juha went to another meeting after lunch....So he was not there in the afternoon)

"Technical report 9573-11:2004 / AMD 1" was an item for the meeting.

What to do? What to do? Should this project be terminated or continued? Alex Brown sent a message to DSDL discussion mailing list during the meeting, and asked persons to send comments on the message.

> I checked the mailing list afterwards, and the enthusiasm is not high to create yet another conversion tool.

Future of WG 1

At the end the need for WG 1 was discussed. Should WG 1 be disbanded, if there are not actual standardisation work items? The issue was discussed. On the other hand, there is some work items, that are still valid.

> My personal opinion is, that non-valid work items should be removed. Even if there is only one work item left, then it would be easy to have meaningful meetings.

After the WG 1 meeting / ZIP format / WG 4

After the formal WG 1 meeting there was general discussion about ZIP format, and possible standardisation of ZIP format. Alex and Murata-san browsed through part 2 of ISO/IEC 29500-2:2008, and showed me some parts of ZIP definitions.

> I checked the ISO/IEC 29500-2:2008 Annex C afterwards with proper time. To my mind, ISO/IEC 29500-2:2008 Annex C raises some fundamental questions about standardisation.

Also it was mentioned that ISO/IEC 26300:2008 contains references to ZIP.

> I checked the ISO/IEC 26300:2008 afterwards with proper time. To my mind, ISO/IEC 26300:2008 and its ZIP references raise some fundamental questions about

standardisation.

It was concluded that there might be some discussion about ZIP standardisation in the WG 4 meeting.

5. The WG 4 meeting 14 June 2010

CJK workshop

Murata-san told about informal CJK workshop. The group gathering is not a formal WG, but its work can implemented, when national bodies decide to incorporate something to standards/proposals.

I realised later, that CJK meant “China-Japan-Korea” workshop.

My personal impression is, that East Asian characters are not well understood, and there has to be better ways to represent them.

The problem was, that during the presentation only Murata-san was the only person in the room representing ideogram languages. Other persons were representing alphabetical languages, i.e. western languages.

Defect reports from JISC

Murata-san presented ten new defect reports from JISC. These defect reports were highly detailed.

Break

“The Byte” / Alex Brown

Alex Brown presented ISO 2382-1 definition of “byte”.

If I understood right, in 29500:2008 it is “octet”, and now it was decided, that byte is used instead of “octet”.

Dates Project Progress Report

Chris Rae presented the date project progress report. Once again there were several defect reports to be handled. Most of the discussion was about procedures.

The politics were discussed too.
What would be the best way to sell the idea of date-related solutions before ballots?

If I understood right, the decision was to have one (big?) amendment (AMD) handling all date-related problems/solutions/proposals.

The pressing issue is to keep defect reports aligned to the new (big?) amendment (AMD)?

The new problems arises, if amendments have their own corrigenda (COR). How to keep things readable, when there are several AMDs and CORs?

Lunch break

There was general discussion about Finland with one group.

One serious discussion was about flexibility of ISO procedures.
One proposal is, that to ISO is submitted material, which is already well-defined.

It was noted, that in September plenary there might be more people in WG 4 meeting.

> As a personal note I propose, that there only well-defined material should be submittted for ISO/IEC JTC1 procedures. The JTC1 procedures demand time.

Defect report maintenance

This discussion was interesting. Everybody seemed to acknowledge, that there will be more defect reports in the future. I did not hear any other statements.

Well. The problem is the amount of defect reports.

There was discussion about new format for submitting defect reports. How should these defect reports be handled by the programs? Everybody seemed to acknowledge, that ISO Livelink is not a working and/or user-friendly system.

> It seemed to me, that there should be some defect handling system (bug tracking), but this has not been used from the beginning. The problem is to set up a defect handling system (bug tracking) when there is already hundreds of bug reports done manually.
>
> I did not propose any defect handling system (bug tracking), but there are several commercial and open-source solutions.

Everybody agreed, that submitting defects should be easy. Also commenting should be easy.

Schema maintenance

This was highly detailed. Murata-san presented schemas in Subversion (assembla.com account). Version tracking in Subversion is rather easy.

This is very technical, but combining schemas is problematic. The problem is, that combining should be manually.

Coffee break

Nothing to report.

Defect reports and 29500:2008 part 2

Murata-san presented at least 22 defect reports related to the part 2. What was recurring several times?

- "not well defined"
- "not specified"
- "not clear"

The problem is also, that there is some straightly copied parts from PKWARE specifications. It is unclear to me, what straightly copied parts will result in the near/distant future.

Should part 2 be rewritten? There was discussion about this.
If all defect reports are gathered together from current version, it might cause several new defect reports.
When the part 2 is rewritten, several non-document defects could be corrected when rewriting the part 2. This could be also faster than collecting all defect reports.

There was also discussion, that possibly other parts should be rewritten also.
It was noted, that possibly part 2 is easiest to rewrite.

Session closed 16.05

Social event / Evening program

During Social event / Evening program there was discussion about following:

- Finnish history / Suomenlinna castle specifics
- photographing
- ODF generally
- a general outsider should be able to read ODF and OOXML standards
- IT procurement of the Finnish government

6. The WG 4 meeting 16 June 2010

New work item proposal

Japanese delegation has prepared new work item proposal, i.e. Safe Extension of Office Open XML. This would mean a new standard, which would have three parts.

There was a lot of discussion about different possibilities.

1) Should ECMA publish a standard? In this way standard would be publicly available.

2) Or should there be a new extension to a current standard? This would mean amendments (AMD).

3) When making with JTC1 rules, the final standard is not publicly available.

If I understood correctly, it was decided to with JTC1 rules, and ITTF is asked to publish standard publicly. On the other hand, there was not wide interest/enthusiasm to go for Fact Track procedures.

Office 2010 Extensions

There is always question of selecting correct extensions to be standardized. Several extensions were presented (WordprosessingML, WordprosessingML, PresentationML)

There was discussion, how to standardize these features.

1) Take as they are.
2) Modify namespaces and identifiers.
3) Modify design of markup.
4) Other options.

The problem for Microsoft would be, that there should be standard and non-standard parts in Microsoft Office documents.

This raised some discussion. For other vendors there should be standard formats. On the other hand, other vendors might have their own extensions. There was discussion about technical notes; Technical notes are standardized in certain ways, and that takes time.

Lunch break, three hours

It was decided, that there is three hours break.
During that break three persons will prepare a presentation about Office 2010 Extensions standardization.

After the break

A lot of discussion raised about Technical Reports (TR) and about registry for extensions. If extensions are handled as TRs, it can take considerable amount of time. Registry has some problems, since registrar should be independent and impartial actor.

Extensions seems inevitable, since different software vendors can extend OOXML. How should promising extensions handled effectively? There is also explosion problem, if there is a huge amount of extensions to OOXML.

It was decided, that there should be some research about registry.

End of the day

At the end of the day Rex Jaeschke read the notes about this day.

7. The WG 5 meeting 17 June 2010

First issue

Klaus-Peter Eckert presented Translation Technical Report. The presented version was working draft (WD 2).

When thinking use cases for document processing, there will be several use cases. Klaus-Peter presented nine different use cases.

The problem might be, that ODF and OOXML have different approaches to save documents. Therefore there can be several ways to compare ODF and OOXML documents.

The next step could be creating a technical report (TR).

Second issue / Open Data Interchange System (ODIS)

Open Data Interchange System (ODIS) was the second issue. Jaeho Lee presented this issue.

In other words clipboard is used for short-term preservation. After all copying parts of documents are rather complicated procedures, e.g. text, figure, metadata.

The proposal was to have Open Data Interchange System (ODIS) in order to standardise clipboard actions.

The main discussion about procedures. It was concluded, that a technical report would be most feasible way to move forward.

> Personally I support a technical report, since there is no guarantee, that software vendors will accept this proposal.

People from South Korea are willing lead this project. The problem is to persuade enough other national bodies to start the project.

Break

CJK workshop

Murata-san presented again the informal CJK workshop results.

There was also a teleconference about ePUB. Murata-san will work on ePUB.

Lunch break

In the lunch break there discussion about the 2008 ballot resolution meeting (BRM), where there was a lot of participants. At the moment there is not that much persons in the Working group 4.

Time to go home

8. General results

I could give following results from the meetings:

Both OOXML and ODF need improvements.

It came quite clear, that both OOXML and ODF need improvements.

ZIP packaging was not discussed after all in the WG4 meeting(s), but that is a very thorny issue both to ODF and OOXML.

Well-defined material to JTC1 procedures

It came quite clear, that only well-defined material should be submitted to JTC1 procedures.

Well-meaning people

The working groups (1, 4, 5) consist of well-meaning people, and the discussion was very polite.

Shrinking amount of real OOXML experts?

I just wonder, if there are enough **real** OOXML experts in the world to handle all OOXML defect reports, amendments and corrigenda.

There was notice, that SC34 plenary should bring more experts to the WG4 meeting.

However, the discussion about defect report system (bug tracking) is indication, that there will be even more defect reports submitted. Who will handle all those new defect reports?

Fog of details

The fog of details is just overwhelming, and for a newcomer understanding all details will take time.

Where is the limit?

Where to draw line for extensions? How much there should be new extensions presented? Current situation is that extensions can be well presented by one corporation.

Complexity

The idea or inevitability of corrigenda for amendments sounds rather complicated. This means more complexity for reading the OOXML standard.

The proposal to rewrite part 2 of the OOXML standard is worth considering, if it reduces complexity.

There should be more simplicity and readability, but that is not the case in the current situation.

Bogged down to JTC1 procedures?

The JTC1 procedures take a considerable amount of time, and all kind of draft phases and ballot times are time-consuming.

9. Ultimate winner: PDF

I have came to the conclusion, that ultimate winner of ODF and OOXML standardisation (hassle) will be PDF.

Most of the documents I receive are PDF files.
Most of the documents I send are converted to PDF.

I have received some OOXML documents, but not a single ODF document. And all those OOXML I have received have been just for reading purposes, not for editing.

Since PDF converters are well developed, they should used extensively, and sending editable files should be as the last option.

10. Runner-up: ePUB

Hallway discussions about ePUB are interesting, and ePUB standardisation should be followed closely. It might present some new solutions, which are not possible with PDF specifications.

11. Old Faithful: Word 97

The mentioning of ITTF policies for Word 97 format is very revealing.

In practise I send most of the editable documents in Word 97 format, since generally people are

knowing nothing about OOXML or ODF. There are several programs, which can read Word 97 format. Since Microsoft's new commitment [8] is to release (all?) information about Word 97 format, there is in principle no hindrance to conform to the Word 97 format.

12. WTO rules

The following LONG text must be read.

"

Agreement on Government Procurement [9] as annex 4(b) to Marrakesh Agreement Establishing the World Trade Organization (WTO).

Article VI: Technical Specifications

1. Technical specifications laying down the characteristics of the products or services to be procured, such as quality, performance, safety and dimensions, symbols, terminology, packaging, marking and labelling, or the processes and methods for their production and requirements relating to conformity assessment procedures prescribed by procuring entities, shall not be prepared, adopted or applied with a view to, or with the effect of, creating unnecessary obstacles to international trade.

2. Technical specifications prescribed by procuring entities shall, where appropriate:

(a) be in terms of performance rather than design or descriptive characteristics; and
(b) be based on international standards, where such exist; otherwise, on national technical regulations(footnote 3), recognized national standards (footnote 4), or building codes.

(footnote original) 3 For the purpose of this Agreement, a technical regulation is a document which lays down characteristics of a product or a service or their related processes and production methods, including the applicable administrative provisions, with which compliance is mandatory. It may also include or deal exclusively with terminology, symbols, packaging, marking or labelling requirements as they apply to a product, service, process or production method.

(footnote original) 4 For the purpose of this Agreement, a standard is a document approved by a recognized body, that provides, for common and repeated use, rules, guidelines or characteristics for products or services or related processes and production methods, with which compliance is not mandatory. It may also include or deal exclusively with terminology, symbols, packaging, marking or labelling requirements as they apply to a product, service, process or production method.

3. There shall be no requirement or reference to a particular trademark or trade name, patent, design or type, specific origin, producer or supplier, unless there is no sufficiently precise or intelligible way of describing the procurement requirements and provided that words such as

8 http://www.microsoft.com/interop/docs/officebinaryformats.mspx
9 http://www.wto.org/english/docs_e/legal_e/gpr-94_01_e.htm

"or equivalent" are included in the tender documentation.

4. Entities shall not seek or accept, in a manner which would have the effect of precluding competition, advice which may be used in the preparation of specifications for a specific procurement from a firm that may have a commercial interest in the procurement.
”

I am just wondering if ODF and OOXML conform to these WTO rules. There is possibilities for determining this.

The following LONG text must be read.

Agreement on Technical Barriers to Trade
Annex 2: Technical Expert Groups [10]

“
The following procedures shall apply to technical expert groups established in accordance with the provisions of Article 14.

1. Technical expert groups are under the panels authority. Their terms of reference and detailed working procedures shall be decided by the panel, and they shall report to the panel.

2. Participation in technical expert groups shall be restricted to persons of professional standing and experience in the field in question.

3. Citizens of parties to the dispute shall not serve on a technical expert group without the joint agreement of the parties to the dispute, except in exceptional circumstances when the panel considers that the need for specialized scientific expertise cannot be fulfilled otherwise. Government officials of parties to the dispute shall not serve on a technical expert group. Members of technical expert groups shall serve in their individual capacities and not as government representatives, nor as representatives of any organization. Governments or organizations shall therefore not give them instructions with regard to matters before a technical expert group.

4. Technical expert groups may consult and seek information and technical advice from any source they deem appropriate. Before a technical expert group seeks such information or advice from a source within the jurisdiction of a Member, it shall inform the government of that Member. Any Member shall respond promptly and fully to any request by a technical expert group for such information as the technical expert group considers necessary and appropriate.

5. The parties to a dispute shall have access to all relevant information provided to a

10 http://www.wto.org/english/docs_e/legal_e/17-tbt_e.htm

> technical expert group, unless it is of a confidential nature. Confidential information provided to the technical expert group shall not be released without formal authorization from the government, organization or person providing the information. Where such information is requested from the technical expert group but release of such information by the technical expert group is not authorized, a non-confidential summary of the information will be provided by the government, organization or person supplying the information.
>
> 6. The technical expert group shall submit a draft report to the Members concerned with a view to obtaining their comments, and taking them into account, as appropriate, in the final report, which shall also be circulated to the Members concerned when it is submitted to the panel.
> ”

My guess is, that there should be a Technical Expert Group to determine ODF and OOXML, especially validity of those standards for government procurement and generally the technical feasibility of those standards for international trade.

13. Final thoughts

These meetings reduced some orthodoxy of my opinions related to ODF and OOXML.

We live in an imperfect world.

May be in the near or distant future we have a situation, when both ODF and OOXML have matured to real interoperability standards.

At the moment we are muddling through somewhere in between – in an unknown speed.