

# ISO/IEC JTC 1/SC 34 N 1716

DATE: 2011-10-18

**ISO/IEC JTC 1/SC 34**  
**Document Description and Processing Languages**  
Secretariat: [Japan \(JISC\)](#)

<b>DOC. TYPE</b>	Meeting Report						
<b>TITLE</b>	Minutes of the Busan Meeting of ISO/IEC JTC 1/SC 34/WG 5, 2011-09-29						
<b>SOURCE</b>	WG 5 Convenor						
<b>PROJECT</b>							
<b>STATUS</b>	This document is circulated to the SC 34 members for information.						
<b>ACTION ID</b>	FYI						
<b>DUE DATE</b>							
<b>DISTRIBUTION</b>	P, O and L Members of ISO/IEC JTC 1/SC 34 ; ISO/IEC JTC 1 Secretariat; ISO/IEC ITTF						
<b>ACCESS LEVEL</b>	Open						
<b>ISSUE NO.</b>	232						
<b>FILE</b>	<table><tr><td><b>NAME</b></td><td>1716.pdf</td></tr><tr><td><b>SIZE (KB)</b></td><td></td></tr><tr><td><b>PAGES</b></td><td>11</td></tr></table>	<b>NAME</b>	1716.pdf	<b>SIZE (KB)</b>		<b>PAGES</b>	11
<b>NAME</b>	1716.pdf						
<b>SIZE (KB)</b>							
<b>PAGES</b>	11						

Secretariat ISO/IEC JTC 1/SC 34 - IPSJ/ITSCJ (Information Processing Society of Japan/Information Technology Standards Commission of Japan)\* Room 308-3, Kikai-Shinko-Kaikan Bldg., 3-5-8, Shiba-Koen, Minato-ku, Tokyo 105-0011 Japan \*Standard Organization Accredited by JISC  
Telephone: +81-3-3431-2808; Facsimile: +81-3-3431-6493; E-mail: [kimura@itsci.ipsj.or.jp](mailto:kimura@itsci.ipsj.or.jp)

---

# Minutes of the Busan Meeting of ISO/IEC JTC 1/SC 34/WG5, 2011-09-29

**Jaeho Lee (jaeho@uos.ac.kr)**

**2011-09-29**

## 1. Opening remarks

The meeting started at 09:20 on 2011-09-20. The convener, Jaeho Lee, welcomed everyone to the face-to-face meeting of WG5 in Busan.

## 2. Roll call of delegates

The following members were present during part or all of the meeting:

Name	Affiliation	Employer/Sponsor
Sam Oh	SC 34 Chair	Sungkyunkwan University
Jaeho Lee	WG5 Convener, KR HoD	University of Seoul
Dong-Sun Nam	KR	Hancom
Jung-Koo Kwon	KR	Hancom
SungKyu Park	KR	IBM
Do-Hyun Kim	KR	IBM
SangBeom Ham	KR	Microsoft
Ning Li	CN HoD	Beijing Information Technology Institute
Hongyun An	CN	RedFlag Chinese 2000 Software Co. Ltd.
Xinsong Wu	CN	CESI
Mario Wendt	DE HoD	Microsoft
Klaus-Peter Eckert	DE	Fraunhofer Fokus
Makoto Murata	WG4 Convener, JP HoD	International University of Japan
Francis Cave	GB HoD	Francis Cave Digital Publishing

Name	Affiliation	Employer/Sponsor
Alex Brown	GB	Griffin Brown Digital Publishing Ltd.
Doug Mahugh	ECMA	Microsoft
John Haug	ECMA	Microsoft
Jim Thatcher	ECMA	Microsoft

Present were 18 people representing 5NBs and 1 liaisons.

### 3. Adoption of the agenda

The agenda (SC 34/WG5 N 1655) was adopted by unanimous consent with the following agenda item being added:

- Discussion on document feature identification proposed by Alex Brown

### 4. Disposition of Comments for DTR 29166, "OpenDocument Format (ISO/IEC 26300) / Office Open XML (ISO/IEC 29500) Translation"

Dr. Klaus-Peter Eckert presented draft disposition on comments. All comments are accepted except the following comment, NZ-1.

*NZ-1: IS26300:2006 and IS29500:2008 are the same approved ISO documents referenced by N10617. As yet no further versions have passed through the ISO/JTC1 approval process. There will be later versions of each document passing through JTC1 in the future and for reasons of legacy support they will build on the current contents of 26300 and 29500, so the content of N10617 remains useful. We vote approve and request that JTC1/SC34 commit to keeping N10617 up to date as the two standards that it refers to evolve.*

NZ-1 is rejected because the maintenance of the TR is desirable but it is only possible in case sufficient resources can be provided.

WG5 will ask ITTF to publish the Document SC 34 N 1690, "OpenDocument Format (ISO/IEC 26300) / Office Open XML (ISO/IEC 29500) Translation"

## **5. Discussion of Study Period for "Reference Model for Identification and Description of Document Fragments"**

WG5 convener presented the draft report on Study Period for "Reference Model for Identification and Description of Document Fragments." The draft is not completed as planned to circulate for review by SC34 members and thus request to extend the deadline to January 2nd, 2012.

## **6. Discussion on document feature identification**

Alex Brown presented a proposed project for SC 34, "Document Feature Identification" for discussion. The presentation includes. See Attachment A for details.

As the result of discussion participants agreed on the need for further study of uses cases and applications and asked to present the result in the next meeting in Prague.

## **7. Action Items**

- WG5 requests SC 34 to accept Disposition of Comments on DTR 29166 prepared by WG 5 in Document SC 34 N 1690 and to instruct its Secretariat to submit revised text prepared by Project Editor to ITTF for publication.
- WG5 requests the convener to circulate a draft interim report on Study Period activities to SC34 WG5 members by December 12<sup>th</sup> for review.
- WG5 requests the convener to submit the interim report on Study Period activities to SC34 Secretariat by January 2<sup>nd</sup>, 2012 for review by SC34 members.

## **8. Other Business**

We thanked our meeting host KATS and Ms. Kim and Ms. Choi for their assistance, and sponsor Microsoft, KSA, and Greater Busan CVB for the dinner on Tuesday night and the bus trip and meals on Wednesday.

## **9. Future meetings**

### **Face-to-Face Meetings:**

1. 2012-02-6/9 , Prague, CZ
2. 2012-06-25/29, Brasilia, BR (with the SC 34 plenary)
3. 2012-10-15/19, London, UK (with the SC 34 plenary)

## **10.Adjournment**

Adjourned by unanimous consent at 16:30 on 2011-09-29.

Attachment A

# Document Feature Identification

A proposed Project for SC 34

Alex Brown & Francis Cave  
2011-09-29

## The problem

- Estimating effort for conversions of document sets (not “guesswork followed by bullying”)
- Knowing whether documents contain feature sets which impact interoperability (e.g. embedded objects, non-standard extensions)
- Procurement specification (e.g. formally ruling-out non-standard specifications from office document producers)

## Related Work

- Patrick Durusau's work on PSIs for document structures
- Fraunhofer's work on feature identification
- TR 29166
- Application profiles in ISO/IEC 29500

## What is a “document feature”?

- Technically:
  - Nodes in an XML document, or set of documents, that can be tested for by an XPath query
- Practically:
  - Those nodes may connote some higher-level abstract concept (e.g. a repeating table header / a hyperlink ...)
  - Or they may be more format-specific and syntactic in nature (e.g. identifying that a style attribute is used in an XHTML document)
  - We don't care which!



## Identifying features

- Features are unambiguously identified by IRIs
  - E.g.  
<http://xmlopen.org/features/tables/repeating-header>
- They are tested for by applying XPath queries to documents
- Different document formats require different XPath queries to test for features

## The Technology

- Features are expressed as a collection of IRIs which accompanying annotation
- Features are tested for using Schematron (ISO/IEC 19757-3) schemas, particular to the format of documents being tested
- Features are reported in a simple XML format enumerating the IRIs of features found (a transform from Schematron SVRL output)

## Out-of-scope

- Identifying the relationship between document features
  - Though this may naturally may be done (using e.g. Topic Maps or OWL) as a separate exercise
- Creating mappings between formats based on feature equivalence
- In short: in SC 34 tradition keep it small, simple, and focused on answering user need

## Next Steps

- Proposed to be an IS; recommended steps:
  - Prototype software
  - Socialization of idea
  - NWIP
  - Project Start

Any Questions?