

DocSII Symposium 2004, Bangkok

Today and Tomorrow of Document Style Specifications Language Libraries for Multi-lingual Document Interchange

This text or its updated version is posted on
<http://www.y-adagio.com/public/committees/docsii/refs/2004/papr/sslib.htm>.

Yushi KOMACHI

Panasonic Communications, Shimomeguro, Tokyo Japan
email: komachi@y-adagio.com

2004-10-04

Abstract

1. [Field research in Mongolia](#)
 2. [Field research in Sri Lanka](#)
 3. [Amendments to ISO/IEC TR 19758 \(DSSSL library\)](#)
 - 3.1 [Amd.1 - Extensions to Basic Composition Styles and Tables](#)
 - 3.2 [Amd.2 - Extensions to Multilingual Compositions \(South-East Asian Compositions\)](#)
 - 3.3 [Amd.3 - Extensions to Multilingual Compositions \(North and South Asian Compositions\)](#)
 - 3.4 [Review/comments on the latest texts of the amendments](#)
 - 3.5 [JTC1/SC34 meetings](#)
 4. [Requirements for extension to document style specification languages](#)
 - 4.1 [Extension to DSSSL](#)
 - 4.2 [Extension to XSL](#)
 5. [Related activities](#)
 - 5.1 [XSLT library](#)
 - 5.2 [XSL sample-codings](#)
 6. [Conclusion](#)
- [References](#)
-

Abstract

Overview of the DocSII (Asian Document Style Standardization for Information Interchange) activities taken in Oct. 2003 through Sept. 2004 and related activities on document style specification libraries are dealt with. Members of DocSII visited to Mongolia and Sri Lanka and studied document styles and layouts in published or distributed texts in those countries. Being based on the studies, amendments to ISO/IEC TR 19758: DSSSL library were developed and submitted to ISO/IEC JTC1/SC34. They are now in their balloting stages according to the JTC1 rules. DocSII is trying to develop XSL sample-codings and an Implementation Guide for Document Style Processing as well. Those topics will be discussed in the DocSII Symposium 2004 in Bangkok.

1. Field research in Mongolia

After the last DocSII Symposium 2003, some DocSII members visited to The National Museum of Mongolian History, National Library, Mongolian University of Science and Technology, and book shops in Ulaan Baatar, to make a research on document styles/layouts in written, printed and published texts. The Mongolian specific document styles/layouts in Mongolian texts were extracted and reported in

- Report on the Field Research in Mongolia^[1] (in Japanese)
- Style/layout in Mongolian documents written by Cyrillic characters^[2] (in Japanese)
- Style/layout for a Mongolian textbook in vertical composition^[3] (in Japanese)

The reports include the following style/layout features:

- heading
- indentation
- table of contents
- ordered/unordered list
- enclosure
- multilingual mixture
- line spanning
- etc.

2. Field research in Sri Lanka

In Jan. 2004, DocSII member were dispatched to Colombo to study document styles/layouts that we can usually see in the published documents in Sri Lanka. They visited to University of Colombo, Department of Government Printing, National Library and Documentation Services Board, the Associated Newspapers of Ceylon, National Museum, Department of National Archives, and bookshops in Colombo. Some interesting examples of Sri Lankan specific document styles/layouts were reported in

- Document Style Examples from Sinhalese documents^[4]
- Preliminary study on Sinhalese Document Styles^[5]
- Report on the Field Research in Sri Lanka^[6] (in Japanese)

The reports include the following style/layout features:

- palm-leaf specific style
- clause separator
- ordered/unordered list
- multilingual mixture
- no indentation of the 1st para
- etc.

NOTE: Detailed reports on the field research in Mongolia and Sri Lanka are included in the Annual Report of DocSII 2003 (in Japanese), published by CICC in March 2004.

3. Amendments to ISO/IEC TR 19758 (DSSSL library)

ISO/IEC TR 19758^[7] provides the following style specification libraries:

- basic composition style
- headline
- page number
- note
- inlinenote
- emphasizing mark
- superscript/subscript
- word-length adjustment
- character space adjustment
- clause
- list
- table
- heading
- ruby
- paragraph indentation
- score
- rule

- inline

After the discussion in DocSII Symposium and related domestic meetings in Japan, draft amendments to ISO/IEC TR 19758 were developed and submitted to ISO/IEC JTC1/SC34 for international review and balloting.

3.1 Amd.1 - Extensions to Basic Composition Styles and Tables

This amendment^[8] provides additional DSSSL libraries for basic composition styles and table compositions. The libraries include:

- parameter specification for content-driven basic composition styles
- table composition specification with rounded corner, table header column/row, multiple strings in table cells, and word wrapping.

The programme of work for this Amd.1 is shown in Table 3.1.

Table 3.1 Programme of work for the Amd.1

Committee and WG	SC34/WG2					
Project Number	JTC1.34.19758.01 (ID: tr19758a1)					
Document Number	ISO/IEC TR 19758:2003/Amd.1					
Stages (text submission)	NP	CD	FCD	DIS/FDIS	IS	Note
Target Dates	----.--	2003.12	----.--	2004.05	2004.12	
Actual Dates	----.--	2003.12	----.--	2004.04	----.--	DAM1 is WG2 N161 (JTC1 N7441)

3.2 Amd.2 - Extensions to Multilingual Compositions (South-East Asian Compositions)

This amendment^[9] provides additional DSSSL libraries for multilingual compositions. The style elements supported in this amendment are based on the requirements for South-East Asian documents.

The libraries include the following style elements:

- composition style extension for multilingual composition
- extended emphasizing mark
- first paragraph without indentation
- two em indentation
- open quote as a paragraph initiator
- interrupted underline/overline
- enclosure
- drop cap
- line spanning
- letter space for emphasis (Sperrsatz)
- unordered list with special bullet

The programme of work for this Amd.2 is shown in Table 3.2.

Table 3.2 Programme of work for the Amd.2

Committee and WG	SC34/WG2					
Project Number	JTC1.34.19758.02 (ID: tr19758a2)					
Document Number	ISO/IEC TR 19758:2003/Amd.2					
Stages	NP	CD	FCD	DIS/FDIS	IS	Note
Target Dates	----.--	2003.12	----.--	2004.05	2004.12	
Actual Dates	----.--	2003.12	----.--	2004.04	----.--	DAM2 is WG2 N163

In accordance with the summary^[20] of the last DocSII Symposium:

1. DocSII provides the ISO/IEC JTC1/SC34 project of Amendment to ISO/IEC TR 19758 and the related ISO members with the information discussed in this symposium, so that the project may draft a PDAM (Proposed Draft Amendment) text.

DocSII document (DocSII N20, 21, 33, 23, 25, 26, 28, 34, 30, 35, 31, 37) were distributed in the SC34/WG2 members with WG2 document number (N136 through N147 respectively).

In the Philadelphia meeting of WG2, the following draft amendments were created and submitted^[14] to SC34:

- PDAM1 for SC34 balloting
- PDAM2 for SC34 balloting
- working draft of Amd.3 for NB review.

3.5.2 SC34 plenary and WG2 meeting in Amsterdam, April 2004

Just before the Amsterdam meeting, PDAM1 and PDAM2 texts were approved by the balloting of SC34 members. Project editor created the disposition of the comments on PDAM1/PDAM2 texts and the following texts^[15]:

- DAM1 for JTC1 balloting
- DAM2 for JTC1 balloting
- PDAM3 for SC34 balloting

3.5.3 SC34 plenary and WG2 meeting to be held in Washington D.C., Nov 2004

Project editor is planning to work on drafting of

- the final text of Amd.1 for publication
- the final text of Amd.2 for publication
- DAM3 for JTC1 balloting

which will include the additional comments provided or to be provided by DocSII members.

In addition, SC34/WG2 is planning to have an opportunity to review the 1st working draft of "Implementation Guide for Document Style Processing".

4. Requirements for extension to document style specification languages

4.1 Extension to DSSSL

Amd.2^[16] to DSSSL is now on development. The FPDAM2 text is on balloting with closing date of Aug. 2004. The project editor will include additional requirements and comments, which are and will be clarified in our works of DSSSL libraries.

4.2 Extension to XSL

Actual implementation of an XSL processor indicated some inconvenience of style specification by existing XSL 1.0. An XSL expert group in Japan submitted requirements to W3C requesting some corrections and extensions to the XSL 1.0. Those requirements are listed on W3C web:

Comments on XSL1.0

- *This message:* [Message body] [Respond] [More options]
- *Related messages:* [Next message] [Previous message]

From: <komachi@y-adagio.com>
Date: Sat, 17 Jan 2004 23:00:31 +0900
Message-ID: <000601c3dd02\$5a7117a0\$020ba8c0@KOMACHIST>
To: <xsl-editors@w3.org>
Cc: <epw1@jsa.or.jp>, <sca@us.ibm.com>, <alrb@us.ibm.com>, <komachi@y-adagio.com>

Dear XSL-Editors,

Attached you will find the JSA(Japanese Standards Association)/EPCOM(Committee for Electronic Publishing)'s submission on
- a comment on XSL1.0
- requirement for extension to XSL1.0

Best Regards,
Yushi Komachi (Convener of EPCOM/WG1)

Yushi Komachi
Panasonic Communications Co., Ltd.
2-3-8 Shimomeguro, Meguro-ku, Tokyo 153-8687 Japan
Phone +81 3 5434 7053, Fax +81 3 5445 3663
Email komachi@y-adagio.com

- text/plain attachment: [comm_on_XSL.txt](#)
- application/x-zip-compressed attachment: [req_extension_to_XSL.html](#)

Received on Saturday, 17 January 2004 20:55:51 EST

5. Related activities

5.1 XSLT library

Some XSLT experts in Japan developed XSLT libraries for generating table of contents and indexes for XML documents. It was published as TR X 0059:2002^[17]. Responding to user requirements, the TR X 0059 are being revised for including additional functionalities:

- character string processing (searching, replacing and deleting a string)
- tag processing (replacing and deleting a tag)
- attribute processing (replacing an attribute name, and deleting an attribute)
- Japanese specific processing (conversion between Katakana and Hiragana characters, and conversion between em and en characters)

The revised text^[18] will be published as TS X 0059 after some discussion and an approval in JISC.

5.2 XSL sample-codings

An XSL description is more difficult to configure a library than a DSSSL description. Some style elements in ISO/IEC TR 19758 and its amendments, therefore, are described by XSL as sample-codings^[19]. There will be a presentation on the sample-codings in the DocSII Symposium 2004. The sample-codings are expected to grow more and more rich during our discussion in the Symposium.

Referring to the sample-codings, an actual XSL specification particularly for Asian XML documents can be much more easily developed.

5.3 Implementation Guide for Document Style Processing

The requirements for "Implementation Guide for Document Style Processing" were proposed^[20] in the DocSII Symposium 2003 held in Ulaan Baatar, Sept 30 and Oct. 1, 2003.

As shown in the working draft^[21] of "Implementation Guide for Document Style Processing", some document processing and interchanging environment requires relaxation of style specification and therefore some style specifications are incorporated in formatting and rendering systems. For minimizing the difference between the rendered page images of sending system and receiving system in such an environment, Implementation Guide for Document Style Processing becomes essential.

This topic will also be discussed in the DocSII Symposium 2004.

6. Conclusion

DocSII developed some DSSSL libraries, XSL sample-codings and other documents in its three-year activities. The DSSSL libraries (Amd.1 and Amd.2 to ISO/IEC TR 19758) are on the final processing for publication. Using those libraries and sample-codings, style specifications for XML documents written in Asian languages can more easily be described than usual. The style specifications are indispensable for rendering the XML documents to be printed or displayed.

An important and expected application of the style specification libraries could be e-Government, where a number of documents are XMLised and distributed or stored. Most of government documents are human-readable ones that have to be rendered for visible representation. Therefore, style specifications have to be assigned for those documents.

DocSII expects the libraries will contribute to the document processing and promote the related markets particularly in Asian countries.

References

- [1] DocSII N44, Report on the Field Research in Mongolia, 2003-10
- [2] DocSII N45, Style/layout in Mongolian documents written by Cyrillic characters, 2004-01
- [3] DocSII N49, Style/layout for a Mongolian textbook in vertical composition, 2004-01
- [4] DocSII N41, Document Style Examples from Sinhalese documents, 2003-08
- [5] DocSII N42, Preliminary study on Sinhalese Document Styles, 2004-01
- [6] DocSII N51, Report on the Field Research in Sri Lanka, 2004-01
- [7] ISO/IEC TR 19758:2003, DSSSL library for complex compositions, 2003-04
- [8] JTC1 N7441, ISO/IEC TR 19758/DAM1: Extensions to Basic Composition Styles and Tables, 2004-05
- [9] JTC1 N7442, ISO/IEC TR 19758/DAM2: Extensions to Multilingual Compositions (South-East Asian Compositions), 2004-05
- [10] SC34 N515, ISO/IEC TR 19758/PDAM3: Extensions to Multilingual Compositions (North and South Asian Compositions), 2004-04
- [11] JTC1 N7523, Summary of voting for JTC1 N7441, ISO/IEC TR 19758/DAM1, 2004-08
- [12] JTC1 N7524, Summary of voting for JTC1 N7442, ISO/IEC TR 19758/DAM2, 2004-08
- [13] SC34 N531, Summary of Voting on JTC1/SC34 N515, ISO/IEC TR 19758/PDAM3, 2004-07
- [14] SC34 N465 (SC34/WG2 N149), Recommendations of Dec. 2003 Meeting of ISO/IEC JTC1/SC34/WG2 in Philadelphia, 2003-12
- [15] SC34 N504 (SC34/WG2 N169), Recommendations of Apr. 2004 Meeting of ISO/IEC JTC1/SC34/WG2 in Amsterdam, 2004-04
- [16] SC34 N510, ISO/IEC 10179/FPDAM2, Extensions to multilingual and complicated document styles, 2004-04
- [17] TR X 0059:2002, XSLT Library, 2002-09
- [18] Draft TS X 0059:2004, XSLT Library, 2004-08
- [19] DocSII N54, Sample of Multilingual Document Style, 2004-08
- [20] DocSII N37, Summary of DocSII Symposium 2003, 2003-10
- [21] DocSII N52, Implementation Guide for Document Style Processing (1st WD), 2004-08