

ISO/IEC JTC 1/SC 34/WG4 N0206

Date: 2011-09-5

Proposed Text for ISO/IEC WD 30114-2

National Body of Japan

Secretariat: JISC

Proposed Text for

Information technology — Extensions of Office Open XML File Formats — Part 2: Character Repertoire Checking

Technologies de l'information — Extensions des formats Office Open XML File — Partie 2: Répertoire de caractères Vérification

Warning

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Document type: International Standard

Document subtype:

Document stage: (20) Preparatory

Document language: E

F:\safeExtensionStandard\charReper\JapaneseProposal.docx STD Version 2.1c2

Copyright notice

This ISO document is a working draft or committee draft and is copyright-protected by ISO. While the reproduction of working drafts or committee drafts in any form for use by participants in the ISO standards development process is permitted without prior permission from ISO, neither this document nor any extract from it may be reproduced, stored or transmitted in any form for any other purpose without prior written permission from ISO.

Requests for permission to reproduce this document for the purpose of selling it should be addressed as shown below or to ISO's member body in the country of the requester:

[Indicate the full address, telephone number, fax number, telex number, and electronic mail address, as appropriate, of the Copyright Manger of the ISO member body responsible for the secretariat of the TC or SC within the framework of which the working document has been prepared.]

Reproduction for sales purposes may be subject to royalty payments or a licensing agreement.

Violators may be prosecuted.

Contents	Page
Foreword.....	iv
Introduction.....	v
1 Scope.....	1
2 Conformance.....	1
2.1 Data Conformance	1
2.2 Application Conformance.....	1
3 Normative references.....	2
4 Terms and definitions	2
5 Additional Requirements	2
5.1 Additional OPC Part.....	2
5.1.1 CREPDL OPC Part.....	2
5.2 Extensions using Markup Compatibility and Extensions	3
5.2.1 Ignorable elements and attributes	3
5.2.2 Alternate Content Blocks.....	4
5.3 Extensions using OOXML-defined extension elements.....	4
Annex A (normative) Schemas.....	5
A.1 NVDL scripts	5
A.1.1 WordprocessingML	5
A.1.2 SpreadsheetML.....	6
A.2 RELAX NG schemas	6
A.3 W3C XML Schema schemas.....	7
Bibliography	8

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 30114-2 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information Technology*, Subcommittee SC 34, *Document Description and Processing Languages*.

This second/third/... edition cancels and replaces the first/second/... edition (), [clause(s) / subclause(s) / table(s) / figure(s) / annex(es)] of which [has / have] been technically revised.

ISO/IEC 30114 consists of the following parts, under the general title *Information technology — Extensions of Office Open XML File Formats*:

- *Part 1: Guidelines*
- *Part 2: Character Repertoire Checking*

Introduction

ISO/IEC 29500 is designed to allow safe extensions. An OOXML extension is safe if OOXML documents containing this extension are guaranteed to conform to ISO/IEC 29500 and are thus guaranteed to be processable by conformant applications of ISO/IEC 29500. Safe extensions of ISO/IEC 29500 may or may not be standardized. ISO/IEC 30114 provides a collection of standardized safe extensions of OOXML.

This part of ISO/IEC 30114 provides a standardized safe extensions for character repertoire checking. For example, it can be used to ensure that some cells in OOXML Spreadsheet documents contain characters in ISO 8859-1 only.

The structure of this part of ISO/IEC 301147 is as follows. ...

Information technology — Extensions of Office Open XML File Formats — Part 2: Character Repertoire Checking

1 Scope

This part of the multi-part standard provides an extension for embedding CREPDL scripts (as specified in 19757-7) as OPC parts within OOXML documents, associating such OPC parts with certain locations in OOXML documents, and examining whether or not text chunks at these locations are in the specified repertoires.

Note This draft covers two locations: sdtPr of WML and c of SML. Should we try to provide more? Comments and sheet names in SML?

2 Conformance

2.1 Data Conformance

An OOXML document shall be conformant to this part of the standard if it satisfies requirements stated in this part.

2.2 Application Conformance

To be supplied.

3 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 19757-2:199x, *General title of series of parts — Part 2: Title of part*

ISO/IEC 19757-4:199x, *General title of series of parts — Part 2: Title of part*

ISO/IEC 19757-7:199x, *General title of series of parts — Part 2: Title of part*

ISO/IEC 29500-1:2011, *General title of series of parts — Part 2: Title of part*

ISO/IEC 29500-1:2011 AMD1:201X, *General title of series of parts — Part 2: Title of part*

ISO/IEC 29500-3:2011, *General title of series of parts — Part 2: Title of part*

ISO/IEC 29500-4:2011, *General title of series of parts — Part 2: Title of part*

ISO/IEC 29500-4:2011 AMD1:201X, *General title of series of parts — Part 2: Title of part*

4 Terms and definitions

For the purposes of this document, **the terms and definitions given in the normative references and the following apply.**

4.1

OPC part

Part as defined in 29500-2

4.2

Alternate Content Blocks

text of the definition

5 Additional Requirements

5.1 Additional OPC Part

5.1.1 CREPDL OPC Part

Content Type:	application/xml
Root Namespace:	http://purl.oclc.org/dsdl/crepdl/ns/structure/1.0
Source Relationship:	To be supplied

An instance of this OPC part type shall be a CREPDL script as specified in ISO/IEC 19757-7.

A package is permitted to contain zero or more CREPDL OPC parts, each of which shall be the target of an explicit relationship in c of SML (see 5.2.1.2) characterRepertoire (see 5.2.1.1) of WML.

5.2 Extensions using Markup Compatibility and Extensions

5.2.1 Ignorable elements and attributes

5.2.1.1 WML

Introduce an ignorable element `cr:characterRepertoire` of the namespace "http://purl.oclc.org/ooxml/extensions/character-repertoire-checking" as a child of `w:sdtPr`. This element shall reference one of the CREPDL OPC parts.

EXAMPLE

```
<w:sdt>
  <w:sdtPr xmlns:mc="http://schemas.openxmlformats.org/markup-
compatibility/2006"
    xmlns:cr="http://purl.oclc.org/ooxml/extensions/character-repertoire-
checking"
    mc:Ignorable="cr" >
    ....
    <cr:characterRepertoire r:id="..."/>
    ....
  </w:sdtPr>
  <w:sdtContent>
    ....
  </w:sdtContent>
</w:sdt>
```

This element shall imply that the content of `w:sdtContent` should be in the character repertoire specified by the referenced CREPDL script.

Issue: `w:sdtPr` appears in `WordprocessingML_Comments`, `WordprocessingML_Endnotes`, `WordprocessingML_Footer`, `WordprocessingML_Footnotes`, `WordprocessingML_Glossary_Document`, `WordprocessingML_Header`, and `WordprocessingML_Main_Document`. Should we allow `characterRepertoire` in all of these OPC parts?.

5.2.1.2 SML

Introduce an ignorable attribute `@cr:id` to elements `c` in SML. This attribute is of the namespace "http://purl.oclc.org/ooxml/officeDocument/relationshipsCharRepertoire" and references to a CREPDL OPC part.

EXAMPLE

```
<sml:c cr:id="..." ...>
  ...
</sml:c>
```

NOTE `@cr:id` is very similar to `@r:id` defined in `shared-relationshipReference.xsd`, but it belongs to a different namespace for making this attribute ignorable.

This attribute shall imply that the content of this cell should be in the character repertoire specified by the referenced CREPDL script.

The content is either directly represented by a child element `is`, or indirectly represented by a child element `v` as an index into the shared string table.

Issue: `c` appears in `SpreadsheetML_Calculation_Chain` and `SpreadsheetML_Worksheet`. Should we allow `characterRepertoire` in both OPC parts?

5.2.1.3 PML

This part of the standard does not introduce ignorable elements or attributes to PML.

5.2.2 Alternate Content Blocks

This part of the standard does not use alternate content blocks.

5.3 Extensions using OOXML-defined extension elements

This part of the standard does not use OOXML-defined extension elements, namely `extLst`.

Annex A (normative)

Schemas

A.1 NVDL scripts

A.1.1 WordprocessingML

An NVDL script wmlExtension.nvdl for the OPC part "WordprocessingML main".

```

<?xml version="1.0" encoding="UTF-8"?>
<rules xmlns="http://purl.oclc.org/dsdl/nvdl/ns/structure/1.0"
  xmlns:a="http://relaxng.org/ns/compatibility/annotations/1.0">
  <anyNamespace>
    <validate schema="WordprocessingML_Main_Document.rnc"
      schemaType="application/relax-ng-compact-syntax">
      <mode>
        <anyNamespace>
          <attach/>
        </anyNamespace>
      </mode>
      <context path="sdtPr">
        <mode>
          <namespace
            ns="http://purl.oclc.org/ooxml/officeDocument/relationshipsCharRepertoire"
            match="elements">
            <validate schema="relationshipsCharRepertoire.rnc"
              schemaType="application/relax-ng-compact-syntax">
              <mode>
                <anyNamespace>
                  <attach/>
                </anyNamespace>
              </mode>
            </validate>
          </namespace>
          <anyNamespace>
            <attach/>
          </anyNamespace>
        </mode>
      </context>
    </validate>
  </anyNamespace>
</rules>

```

A.1.2 SpreadsheetML

An NVDL smlExtension.nvdl script for the OPC part "Spreadsheet Worksheet".

```
<?xml version="1.0" encoding="UTF-8"?>
<rules xmlns="http://purl.oclc.org/dsdl/nvdl/ns/structure/1.0"
  xmlns:a="http://relaxng.org/ns/compatibility/annotations/1.0">
  <namespace ns="http://purl.oclc.org/ooxml/spreadsheetml/main">
    <validate schema="SpreadsheetML_Worksheet.rnc"
      schemaType="application/relax-ng-compact-syntax">
      <mode>
        <anyNamespace>
          <attach/>
        </anyNamespace>
      </mode>
    <context path="c">
      <mode>
        <namespace
          ns="http://purl.oclc.org/ooxml/officeDocument/relationshipsCharRepertoire"
          match="attributes">
          <validate schema="relationshipsCharRepertoireAttribute.rnc"
            schemaType="application/relax-ng-compact-syntax">
            </validate>
          </namespace>
          <anyNamespace>
            <attach/>
          </anyNamespace>
        </mode>
      </context>
    </validate>
  </namespace>
</rules>
```

A.2 RELAX NG schemas

A.2.1

relationshipsCharRepertoireAttribute.rnc (invoked by wmlExtension.nvdl)

```
namespace charCheck =
  "http://purl.oclc.org/ooxml/officeDocument/relationshipsCharRepertoire"

attribute charCheck:id {
  grammar {
    include "shared-relationshipReference.rnc"
    start = r_ST_RelationshipId
  }
}
```

A.2.2

relationshipsCharRepertoire.rnc (invoked by wmlExtension.nvdl)

```
include "shared-relationshipReference.rnc"
start = element characterRepertoireChecking {charCheck_CT_CharacterRepertoireChecking}
charCheck_CT_CharacterRepertoireChecking = r_id
```

A.3 W3C XML Schema schemas**A.3.1**

relationshipsCharRepertoire.xsd

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  elementFormDefault="qualified"
  targetNamespace="http://purl.oclc.org/ooxml/officeDocument/relationshipsCharRepertoire"
  xmlns:r="http://purl.oclc.org/ooxml/officeDocument/relationships"
  xmlns:re="http://purl.oclc.org/ooxml/officeDocument/relationshipsCharRepertoire">
  <xs:import namespace="http://purl.oclc.org/ooxml/officeDocument/relationships"
  schemaLocation="shared-relationshipReference.xsd"/>
  <xs:element name="characterRepertoireChecking">
    <xs:complexType>
      <xs:attribute ref="r:id"/>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

A.3.2

relationshipsCharRepertoireAttribute.xsd

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  elementFormDefault="qualified"
  targetNamespace="http://purl.oclc.org/ooxml/officeDocument/relationshipsCharRepertoire"
  xmlns:r="http://purl.oclc.org/ooxml/officeDocument/relationships"
  xmlns:re="http://purl.oclc.org/ooxml/officeDocument/relationshipsCharRepertoire">
  <xs:import namespace="http://purl.oclc.org/ooxml/officeDocument/relationships"
  schemaLocation="shared-relationshipReference.xsd"/>
  <xs:attribute name="id" type="r:ST_RelationshipId"/>
</xs:schema>
```

Bibliography

- [1] ISO/IEC Directives, Part 2, *Rules for the structure and drafting of International Standards*, 2001
- [2] ISO/IEC TR 10000-1, *Information technology — Framework and taxonomy of International Standardized Profiles — Part 1: General principles and documentation framework*
- [3] ISO 10241, *International terminology standards — Preparation and layout*
- [4] ISO 128-30, *Technical drawings — General principles of presentation — Part 30: Basic conventions for views*
- [5] ISO 128-34, *Technical drawings — General principles of presentation — Part 34: Views on mechanical engineering drawings*
- [6] ISO 128-40, *Technical drawings — General principles of presentation — Part 40: Basic conventions for cuts and sections*
- [7] ISO 128-44, *Technical drawings — General principles of presentation — Part 44: Sections on mechanical engineering drawings*
- [8] ISO 31 (all parts), *Quantities and units*
- [9] IEC 60027 (all parts), *Letter symbols to be used in electrical technology*
- [10] ISO 1000, *SI units and recommendations for the use of their multiples and of certain other units*
- [11] ISO 690, *Documentation — Bibliographic references — Content, form and structure*
- [12] ISO 690-2, *Information and documentation — Bibliographic references — Part 2: Electronic documents or parts thereof*