

Date: 2008-09-17

**ISO/IEC JTC1/SC2/WG2
Coded Character Set
Secretariat: Japan (JISC)**

Doc. Type: Input to ISO/IEC 10646

Title: Introduction for the ISO/IEC 10646:2008 Working Draft

Source: Project Editor

Project: JTC1 02.10646

Status: For review by WG2

Date: 2008-09-17

Distribution: WG2

Reference: WG2 N 3274, 3276, 3362, N3408, N3509, N3510

Medium:

Summary: This document introduces the Working Draft (WD) for the new edition of ISO/IEC 10646. The WD was created following resolution WG2 M52.24 (WG2 meeting #52, Redmond USA) and is itself made of two documents:

N3509: ISO/IEC 10646:2008 WD text part

N3510: ISO/IEC 10646:2008 WD partial charts

N3509 contains all textual parts, including annexes, which were part of the previous edition, and amendment 1 to 6. Because Amendment 5 and 6 are still under ballot, the content related to these amendments is still subject to change. Note however that Amendment 5 is in its FDAM ballot phase which precludes any technical changes.

In addition, following suggestions made in document N3274 and updated in N3362, the new edition has its own set of technical change to be better aligned with the Unicode Standard, especially on terminology. This was already described in a preliminary version of the WD available as WG2 N3276.

The main differences between N3276 and N3509 are as follows:

- Addition of new terms: combining class, ill-formed CC-data-element subset, and minimal well-formed CC-data-element. This improves synchronization with Unicode 5.1.
- New sub-clause in the clause 20 'Combining character' describing combining class and canonical ordering. These are concepts related to normalization already part of the standard.
- Update of the description for CJK Unified Ideographs as proposed by document WG2 N3408.
- Removal of annotation from list of character names.
- Named UCS Sequence Identifiers are now identified and referenced by a link to an external document.
- Update of clause 26 and 26 to represent latest additions as added by amendments 1-6.

- Annex P is now reserved to additional information concerning CJK Unified Ideographs. Other information is now directly part of the name part of the charts. Therefore the size of Annex P has been significantly reduced.

N3510 is the chart section of the standard, with the new format which has been used since Amendment 5. It does not contain any charts for CJK Unified Ideographs which now all require multi-column layout. It does however include all other repertoires, including content under ballot with Amendment 6. These repertoires contain additional information on characters that were until now only available in the Unicode Standard code charts. This made the character annotation concept unnecessary.

The first page of N3510 contains additional information about the code charts, such as known errata.

Note that N3510 is a very large document. It is likely that eventually the code charts will be split as done in the previous edition.

The Working Draft does not include any linked data files because they are strictly identical to the result of the amendments. These files will be added in the FCD to be delivered later this year.

WG2 experts should already start to study the WD content before the next WG2 meeting in Hong Kong (October 2008). It is expected that the WD will be the base for a FCD to be under ballot in November 2008.
