INTERNATIONAL STANDARD

ISO/IEC
24754

First edition
2008-08-15

Information technology — Document description and processing languages — Minimum requirements for specifying document rendering systems

Technologies de l'information — Description des documents et langages de traitement — Exigences minimales pour spécifier les «document rendering systems»

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 24754 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 34, *Document description and processing languages*.

Information technology — Document description and processing languages — Minimum requirements for specifying document rendering systems

1 Scope

When a structured document is interchanged between an originator and a recipient, the recipient refers to the style specifications that the originator provides to reconstruct the presentation. However, when the recipient does not have sufficient rendering functionality, it may fail to reconstruct the presentation output as the originator expected. In order to preserve presentation output in the course of interchange, the originator and recipient need to negotiate over functionalities referring to the specifications of document rendering systems. To satisfy this requirement, this standard provides the minimum requirements for specifying document rendering systems. The document specification format shall have the same minimum requirements described here.

This International Standard can apply to the document processing environment, where a document is given in a logically structured format which is expressed by a structure markup language, and the visual representation of the document is described by means of the external style and layout specifications which a style and layout specifications language provides. The visual representation of the given document is generated when the style and layout specifications are applied to the logical structure by a document rendering system.

This International Standard provides an abstract list of the features that a rendering system for an authored document may have. The list provides a frame of reference, against which the user and implementor can compare the features of document rendering systems. However, this International Standard does not specify a concrete interchange syntax or direct how each document rendering system shall behave.

This International Standard provides the minimum requirements to specify the features that a document rendering system which transforms formatting objects to rendering output. It may be used as a frame of reference, against which the user, implementer, or software agent may compare the features of a document rendering system. According to these requirements, the user may express what he or she expects of a document rendering system, the implementer may describe the functionality and capability of the document rendering system that he or she implements, and the software agent may negotiate a minimum set of functionality and capability that are shared across different document rendering system implementations.

2 Terms and definitions

2.1 document rendering system

software agent which takes in formatted document and generates rendered output

2.2 rendering

transforming of a structured document from a certain set of constraints to another set of constraints

NOTE: In most cases, a rendered document has stronger constraints than a pre-rendering document does.

2.3 feature

characteristic which a document rendering system supports in order to process a document, particularly at times when the document rendering system negotiate the rendering content with physical boundaries, e.g. line break or page break, of the output media

2.4 paged media

form of media, which is divided into one or more smaller repetitive areas, each of which has its own independent geometric attributes

2.5 region

geometric part of a page that has static position and can appear at the same position on sequential pages

2.6 area

arbitrary part on a page that may have a floating position or be handled as a flowed object

2.7 header

a set of objects that are positioned at the beginning of each page of a document, which may include such objects as page number, title of document, etc.

2.8 footer

objects that are positioned at the end of each page of a document, which may include such objects as page number, title of document, etc.

2.9 operator

a person who interacts with a document or with the document rendering system in order to print or display the document

3   Features of a document rendering system

3.1   General description of a document rendering system

3.1.1   Supported output device

The output device(s) that the document rendering system supports are described. E.g. printing systems and browsers.

3.1.2   Supported document formats

The document formats, e.g. XML, RTF, which the document rendering system supports to express the logical structure of a given document are described.

3.1.3   Supported style languages

The style specification languages, DSSSL and other style specification languages, that the document rendering system supports are described.

3.1.4   Rendering media dimension

The dimension of the rendering media, whether it is paged media or non-paged media, is described here. If the rendering media is paged, supported page sizes are described, e.g. A4, Letter, Legal.

3.1.5   Colour support

The level of support for colour is described, referring to whether colour is supported by the document rendering system, and if colour is supported, what colour profiles are supported. Colour system, e.g. RGB and CMYK, and depth of colour, e.g. 8-bit (256 colours) and 24-bit (16 million colours).

3.1.6   Dynamic content generation

The support for dynamically-generated content, meaning content calculated by and supplied by the stylesheet either as fixed value or conditionally-generated based on the presence or absence of data at rendering time.

3.1.7 Supported coded character sets and encoding schemes

Supported coded character standards and encoding schemes are described. E.g. UTF-8, SHIFT_JIS, KS-5601.

3.1.8 Composition of combined characters

The level of support for the composition of combined syllabic characters, whether the document rendering system supports the stacking of combined characters and environmental glyph alteration are supported, is described. E.g. Thai, Arabic.

3.1.9 Directionality of text

Support for the inherent directions of the progression of text (from left to right, right to left or top to bottom) and lines (from top to bottom, right to left or left to right) is described.

3.1.10 Rendering of white-space characters

The interpretation of white-space characters, e.g. tab, space, carriage return and line feed, is described. A string of white-space characters may be collapsed or may be interpreted literally. Additionally, white-space characters may be removed before or after line breaks.

3.1.11 Font substitution

Support for font substitution mechanism, e.g. font substitution based on ISO/IEC 9541-1, is described.

3.1.12 Font-resource portability

Capability of embedding and retrieving a font resource is described. If the document rendering system supports font-resource portability, it can extract an embedded font resource from the document and use it for rendering.

3.1.13 Embedding of non-standard characters and glyphs

Support for the embedding of non-standard characters and glyphs, e.g. glyphs registered by ISO/IEC 10036, is described.

3.2 Description of pagination and rendering of page layout

3.2.1 Page geometry

3.2.1.1 Simple page layout

The attributes, e.g. page width and height, and margin and padding on top, bottom, left and right, of a page that the document rendering system supports are described.

3.2.1.2 Columnar pagination

The level of support for columnar pagination, i.e. whether columns are supported, and whether the column gap can be specified, etc. is described.

3.2.1.3 Region setting and positioning

Basic attributes, e.g. width, height, margin and padding on top, bottom, left and right, and positions in the page, of the regions that the document rendering system can specify are described.

3.2.2 Page-geometry sequencing

Support for different types of page-geometry sequencing is described. Example types for page-geometry sequencing are single page ordering, repeatable page ordering, and conditional page ordering. Example types for conditional page ordering are page parity and page position.

3.2.3　Page flow

Support for static content, paginated content and flow maps is described. E.g. document title, page header, page footer, sidebars and multiple flows.

3.3　Layout independent formatting

This subsection provides a list of features that are independent from the layout of the document.

3.3.1　Z-index

Support for the specification of the z-index depth level of an area is described.

3.3.2　Relatively-positioned areas

Support for relatively-positioned areas, or floats, for graphics and other objects is described. Treatments available for surrounding text are described. E.g. tight, box and arbitrary shape.

3.3.3　Absolutely-positioned areas

Support for absolutely-positioned areas for graphics and other objects is described. The available bases for the absolute measurements are described. Treatments available for surrounding text are described. E.g. tight, box and arbitrary shape.

3.3.4　Annotation

Support for footnote, endnote and side note, the note body and the reference to the note is described.

3.3.5　Cited content

Support for references to cited content that will be replaced at rendering time is described.

3.3.6　Dynamic content alteration

Support for content alteration that is triggered by the rendering environment, operator interaction or other conditions is specified.

3.4　Tables

This subsection provides a list of features that relates to the rendering of tables.

3.4.1　Table geometry

Supported table geometry, e.g. width, height, padding and margin, is described.

3.4.2　Table caption

Support for table captions is described.

3.4.3　Table header

Support for the table header and its behaviour is described. Support for the orientation of headers and iteration of headers after the page break when the table spans across pages, is described.

3.4.4　Table footer

Support for the table footer and its behaviour is described. Support for the orientation of footers and iteration of footer before the page break when the table spans across pages, is described.

3.4.5   Table background

Support for the types of background for a table, e.g. the placement of an image or the colour, is described.

3.4.6   Complex table compositions

The level of support for complex table compositions is described here, based on the structure set forth in ISO/IEC TR 19758.

3.5   Lists

This subsection provides a list of features that relates to the rendering of lists.

3.5.1   List item dimensions

Support for the geometry of the list, e.g. widths, heights, paddings and margins of the list item label and the list item body, is described here.

3.5.2   List item identification

Support for the identification of list items is described, including techniques such as enumeration and itemisation, and the types of symbols that can be supplied as list item labels.

3.6   Block formatting

This subsection provides a list of features that relates to the rendering of blocks.

3.6.1   Block geometry

Support for the geometry of a block of lines, e.g. width, height, margins, paddings, line leading and line-stacking strategy, is described here.

3.6.2   Wrapping of characters

Support for the wrapping of characters is described, e.g. the wrapping of characters that often takes place for punctuation characters in East Asian scripts, such as Chinese, Japanese and Korean. This is an equivalent feature of Japanese 'Kinsoku'.

3.6.3   Wrapping of words

Support for the wrapping of words, e.g. whether the last word in a line is wrapped to the next line, or kept in the same, is described. Wrapping of words often takes place in English and other European scripts.

3.6.4   Hyphenation

Support for the hyphenation at the line end is described. The principles by which the document rendering system hyphenates and breaks words, e.g. rule-based, dictionary-based and manual.

3.6.5   Wrapping of lines

Support for the wrapping of lines, e.g. controlling widows and orphans, is described.

3.6.6   Wrapping of blocks

Support for the wrapping of blocks, e.g. whether a block which only contains a small number of lines, such as a chapter heading, shall be wrapped to the next column or page, is described.

3.6.7 Indentation

Support for indentation, its values and its use, is described.

3.6.8 Content alignment

The types of content alignment, e.g. lines in a block, areas in the flow, areas in marginalia, are described.

3.7 Decoration

This subsection provides a list of features that relates to the decoration of a document.

3.7.1 Graphics support

Support for the rendering of instream and/or external graphics, e.g. the supported formats, scaling, positioning, is described.

3.7.2 Leader

Support for leaders, e.g. the leader length, truncation of the leader, alignment of the leader, is described.

3.7.3 Rule

Support for rules, e.g. orientation, length, types of rule, is described.

3.8 Inline-level formatting

This subsection provides a list of features that relates to the rendering of inline constructs.

3.8.1 Inline object

Support for the control of inline areas in a line, e.g. baseline shift, line height, is described.

3.8.2 Bidirectional override

Support for the override of the inherent directionality of characters is described.

3.9 Character-level formatting

This subsection provides a list of features that relates to the rendering of characters.

3.9.1 Character spacing

Support for the types of spacing between characters, e.g. specified behaviours, automatic behaviours, is described.

3.9.2 Word spacing

Support for the types of spacing between words, e.g. specified behaviours, automatic behaviours, is described.

3.9.3 Character variation

Support for character variants, e.g. drop initial cap, small caps, is described.

3.9.4 Character decoration

Support for character decorations, e.g. strike-through, underline, overbar, reverse, is described.

3.9.5   Font selection

Support for selection of fonts, e.g. font family, font style, font weight, is described.

3.10 Linking

Support for the specification of the association between areas of a document and other areas or external locations, e.g. for operator interaction, is described.

Annex A
(informative)

Example value descriptions

This informative annex offers examples of values one might use when specifying the features of a document rendering system according to this international standard. These values are not meant to be exhaustive but only documentary to give the reader consideration of the kinds of values that might apply for each feature. Some values are given in reference to international and other standards, which is not mandatory. This international standard explicitly does not standardize the values for specifying the features in order that users of this standard are not constrained by any set of values that might be incomplete. The onus is on the user of this standard to describe a particular feature sufficiently well for the reader of their features to understand the capabilities offered by a document rendering system.

Note the convention of suffixing the annex section titles with the corresponding body section titles is for guidance to the normative text associated with an item.

A.1   General description of a document rendering system (3.1)

A.1.1   Supported output device (3.1.1)

Examples:

—   laser printer

—   inkjet printer

—   dot matrix printer

—   LCD monitor

—   CRT monitor

—   NTSC (TU-R BT.470-6)

—   PDF/A (ISO 19005-1)

—   PDF/X-1a (ISO 15930-1)

—   JPEG file (ISO/IEC 10918-1)

—   Macintosh PICT (Inside Macintosh: Imaging with QuickDraw, Appendix A)

A.1.2   Supported document formats (3.1.2)

Examples:

—   HTML 4.0

—   SPDL (ISO/IEC 10180)

—   PostScript Level 2

—   PDF 1.4

— PCL 4

— rich text format version 1.6

A.1.3 Supported style languages (3.1.3)

Examples:

— CSS 2.1

— XSL-FO 1.1

— DSSSL (ISO/IEC 10179)

A.1.4 Rendering media dimension (3.1.4)

Examples:

— paged media

— non-paged media

— A4

— US-letter

— US-legal

A.1.5 Colour support (3.1.5)

The colour support describes the native colour space and colour depth that are available on the device correctly. There are several popular standards for display devices that specifies colour space range, colour encoding and colour depth in detail. Simple reference these standards is sufficient description of colour support for display devices.

The available colour space and colour depth for printing devices are not standardized in detail. The printing device may accept the device independent colour specification and render the requested colour by approximated colour for the device. For example, some monochrome printers may accept the colour document data and convert it to monochrome data internally and the output is monochrome. Although it is helpful to describe the acceptable colour spaces and colour matching methods, the description on acceptable colour spaces is not essential. The native colour support or available subset shall be described. Some standards of printing colour chart (e.g. ISO 12642-1 initial set of 182 colours) can be referred as available colour space. If there is no appropriate device-independent standard to describe available colour space, the device dependent colour space families in ISO 10180 may be used for to avoid undefined keywords.

Examples:

— IEC 61966-2-1 Default colour space (sRGB, 8bit per R, G, B components)

— IEC 61966-2-2 Extended RGB colour space (scRGB, 16bit per R, G, B components)

— ISO 12642-1 colours by CMYK inks

— ISO/IEC 10180 DeviceCMYK (8bit for C, M, Y components and 10bit for K component)

— ISO/IEC 10180 DeviceGray (16bit)

A.1.6 Dynamic content generation (3.1.6)

Examples:

— table of contents generation

— list numbering

— figure numbering

— example numbering

— table numbering

A.1.7 Supported coded character sets and encoding schemes (3.1.7)

Considering the fact that some character sets are designed for specific character encodings, each character sets provides the list of acceptable encoding schemes. For the specification of multilingual coded character set "ISO/IEC 10646" is not recommended to avoid the unclear coverage. The specification by subsets in ISO/IEC 10646 Annex A is better.

Examples:

— ISO/IEC 8859-1 coded by ISO-2022, UTF-8 or UTF-16

— ISO/IEC 10646 collection 1 coded by UTF-8

— ISO/IEC 10646 collection 1, 2, 3, 4, 30 coded by UTF-8, UTF-16 or UTF-32BE

A.1.8 Composition of combined characters (3.1.8)

The compositions of combined characters are expected to be described by the character set and encoding schemes to give source characters and the coverage of composed glyph or characters. Considering the fact that the coverage of required composition is not standardized exactly except of Latin, Cyrillic, Greek and IPA characters, it is possible that the coverage of composed glyph or characters are unspecified.

Examples:

— ISO/IEC 6937

— ISO/IEC 10646 collection 2, 3, 30 composed by ISO/IEC 10646 1, 4 implementation level 3 coded by UTF-8

— ISO/IEC 10646 collection 14, 15, 64, 68 implementation level 2 coded by UTF-8 (Arabic)

— ISO/IEC 10646 collection 16 implementation level 2 coded by UTF-8 (Devanagari)

— ISO/IEC 10646 collection 29, 53 implementation level 3 coded by UTF-8 (Hangul Jamo)

A.1.9 Directionality of text (3.1.9)

This description is not the direction of text against paper or printing. When each glyph is rendered as isolated graphics, it is possible to line up the glyph for any direction. This description shall provide the information on the control of implicit pen position moving and selection of glyph shape which is appropriate for direction during the rendering of coded text. Most scripts have their default inline direction (direction from a glyph to next logical glyph), and the selection of glyph shape in their default inline direction is required. If the rendering in non-default inline direction is supported, the pair of character set and the method to specify the non-default inline direction shall be described. On many systems, the method to specify the inline direction is different from the method to select the appropriate glyph shape for the specified inline direction. The glyph shape selection for non-default inline direction is strongly dependent with the font resource; it shall be as described in 3.1.12 and 3.1.13.

Examples:

— Left-to-right

— ISO/IEC 8859-6:1999, G.1

— ISO/IEC 10646 collection 1, 2, 3, 4, 7, 14, 15, 64, 68 (left-to-right, right-to-left by ISO/IEC 10646 collection 202)

— ISO/IEC 10646 collection 93, 94, 95, 96 (left-to-right, bottom-to-top by XSL writing-mode property)

— ISO/IEC 10646 collection 380 (top-to-bottom, right-to-left by XSL writing-mode property)

A.1.10 Rendering of white-space characters (3.1.10)

There are two groups of white-space characters. One group is the characters that the white-space with fixed size is inserted (e.g. ISO-8859-1 0x20 space, ISO-6429 0x09 horizontal tab, ISO 10646 U+3000 ideographic space). Another group is the control characters that the rendering system cannot reflect the designed features and displayed by white-space (e.g. ISO-6429 0x0A line feed, 0x0B vertical tab). The white-space characters in the first group are often repeated to adjust the text layout. The rendering system acting as simple typewriter displays the repeated white spaces literally, but some intelligent rendering system collapses the repeated white-space to single space (e.g. HTML). The collapse of repeated white-space may be restricted to the white-spaces before new line sequence, because the effect of collapse is expected to be negligible in many cases.

Examples:

— ISO/IEC 10646 U+3000: always rendered in literal.

— ISO/IEC 10646 U+0020: collapsed before new line sequence.

— ISO/IEC 10646 U+0009: 8 times width of U+0020, collapsed before new line sequence.

— ISO/IEC 10646 U+000A: always collapsed and displayed by single ISO 10646 U+0020 space.

A.1.11 Font substitution (3.1.11)

A missing font resource may be substituted by another font resource which properties are matching with original font resource. The properties to be matched are defined in several standards (ISO/IEC 9541-1, ISO/IEC 14496-22 4.2.7 OS/2 table, ISO/IEC 14496-22 4.6.6 PCLT table, CSS 2.1, XSL 1.1), but some of them are missing the definition of matching algorithm. If the algorithm used in the document rendering system is not standardized and difficult to described, the list of referred properties is helpful.

Examples:

— CSS 2.1 font matching algorithm

— ISO/IEC 10180 FontReference::MatchRules::SameIfSpecified matched by FontReference::Satisfaction::NameOrProps

— ISO/IEC 14496-22 Annex B. Font Class parameter

— ISO/IEC 14496-22 4.2.7.17 Panose

A.1.12 Font-resource portability (3.1.12)

This entry provides the information on the methods to collect the fonts required in rendering. If the input document format is designed to include embedded font resources in various formats, it is recommended to list acceptable formats. For example, CSS 3 web fonts have the tags to indicate the font formats used in CSS, some of them are unpublished format. Also the method of retrieval and the supported font formats shall be described. This entry does not describe the font embedding feature of device output. For example, this entry does not describe if the rendering system generates ISO 19005 PDF/A data with embedded font or all texts are rendered as graphics. The font-resource portability can be described as a collection of the entries.

Examples:

— ISO/IEC 19005 PDF/A embedded Type1 font and Type42 font

— CSS 3 web fonts (ISO/IEC 9541-3 Type1 font and ISO/IEC 16500-6 TrueDoc PFR) retrieved by URI via file and http protocol

A.1.13 Embedding of non-standard characters and glyphs (3.1.13)

This description gives information on the method to embed non-standard glyph identifier in the coded text: a glyph cannot be specified by character code (e.g. CJK ideographic variant, glyph of punctuations in specific inline direction), a ligature or combining character coded by multiple character code points. Except of glyph identifier, the description on glyph selection mechanism is helpful to guide the access of non-standard characters and glyphs. ISO/IEC 14496-22 Open Font Format is a standard of a font format including the predefined glyph substitution table, although the specification does not guarantee the availability of required glyphs.

This description does not provide information on the data format which embeds font files of these glyphs into the document. It shall be as described in 3.1.12.

Examples:

— ISO/IEC 10036 glyph identifier

— Unicode IVD Adobe-Japan1

— ISO/IEC 14496-22 OFFS, glyph substitution by vert and vrt2 for "hani" scripts.

A.2   Description of pagination and rendering of page layout (3.2)

A.2.1   Page geometry (3.2.1)

A.2.1.1   Simple page layout (3.2.1.1)

Examples:

— physical page dimensions

— logical region placement

— logical region orientation

— logical region dimensions

— padding

— margin

A.2.1.2   Columnar pagination (3.2.1.2)

Examples:

— evenly-sized columns

— equally-spaced columns

— arbitrarily-sized columns

— arbitrarily-spaced columns

— arbitrary column flow

— simple serpentine columns

A.2.1.3   Region setting and positioning (3.2.1.3)

Examples:

— region placement

— region dimensions

— region margins

— region padding

— region identification

A.2.2   Page-geometry sequencing (3.2.2)

Examples:

— single page

— repeating page

— conditional page

— page parity

— page position

— page blankness

A.2.3   Page flow (3.2.3)

Examples:

— flowed paginating content

— static repeating content

— content associated with named regions

— multiple flows of paginating content

A.3   Layout independent formatting (3.3)

A.3.1   Z-index (3.3.1)

Examples:

— layering of content coincident on a page

A.3.2   Relatively-positioned areas (3.3.2)

Examples:

— floating graphics

— tight text flow

— box text flow

— arbitrary-shape text flow

A.3.3   Absolutely-positioned areas (3.3.3)

Examples:

— floating graphics

— tight text flow

— box text flow

— arbitrary-shape text flow

A.3.4   Annotation (3.3.4)

Examples:

— footnotes

— endnotes

A.3.5   Cited content (3.3.5)

Examples:

— footnote citation

— title citation

— citing arbitrary content

A.3.6   Dynamic content alteration (3.3.6)

Examples:

— operator-directed alteration

— alterations triggered by formatting conditions

A.4   Tables (3.4)

A.4.1   Table geometry (3.4.1)

Examples:

— table dimensions

— table margins

— padding

A.4.2   Table caption (3.4.2)

Examples:

— caption content

— caption positioning

A.4.3 Table header (3.4.3)

Examples:

— header orientation

— header iteration

— header suppression

— header annotation

A.4.4 Table footer (3.4.4)

Examples:

— footer orientation

— footer iteration

— footer suppression

A.4.5 Table background (3.4.5)

Examples:

— colour in table background

— image in table background

— repetition of image in table background

— anchoring of image in table background

— scaling of image in table background

A.4.6 Complex table compositions (3.4.6)

Examples:

— complex table compositions by TR 19758:2002

A.5 Lists (3.5)

A.5.1 List item dimensions (3.5.1)

Examples:

— list label width

— list label height

— list label padding

— list label margin

— list body width

— list body height

— list body padding

— list body margin

A.5.2   List item identification (3.5.2)

Examples:

— item identification by itemization

— item identification by use of symbols

— resumption of list item identification

A.6   Block formatting (3.6)

A.6.1   Block geometry (3.6.1)

Examples:

— block width

— block height

— block margin

— block padding

— line leading

— fixed size line stacking

— accomodating size line stacking

A.6.2   Wrapping of characters (3.6.2)

Examples:

— wrapping of characters

— wrapping of characters by JIS X 4051

A.6.3   Wrapping of words (3.6.3)

Examples:

— engaging wrapping of text to new lines

— disengaging wrapping of text to new lines

— adjustable count of words being wrapped on last line of a block