

Main differences between PDF 1.7 and PDF 2.0 Standard Structure Elements

Common Tag	Difference between PDF 1.7 and PDF 2.0	References
Annot (Annotation)	PDF 1.7: categorized as <i>Inline only</i> PDF 2.0: categorized as <i>Grouping, Block, or Inline</i> . If multiple annotations, then all annotations must be of a single Subtype.	ISO 32000-1:2008, Table 338 and §14.8.4.4.3 ISO 32000-2:2020, Table 368
Caption	PDF 1.7: categorized as <i>Grouping for tables, lists or figures</i> PDF 2.0: categorized as <i>Grouping or Block, appropriate for any type of content</i>	ISO 32000-1:2008, Table 333 and 336 ISO 32000-2:2020, §14.8.4.8.4 and Table 372
Div (Division)	PDF 1.7: a generic block-level element or group of elements PDF 2.0: a grouping <i>orthogonal to the document's semantic structure</i>	ISO 32000-1:2008, Table 333 ISO 32000-2:2020, Table 365
Figure	PDF 1.7: categorized as <i>Inline</i> , unless Placement is set to <i>Block</i> PDF 2.0: categorized as <i>Grouping, Block, or Inline</i>	ISO 32000-1:2008, Table 340 ISO 32000-2:2020, Table 373
Form (Widget Annotation)	PDF 1.7: categorized as <i>Inline</i> , unless Placement is set to <i>Block</i> PDF 2.0: categorized as <i>Grouping, Block, or Inline</i>	ISO 32000-1:2008, Table 340 ISO 32000-2:2020, Table 368
Formula	PDF 1.7: categorized as <i>Inline</i> , unless Placement is set to <i>Block</i> PDF 2.0: categorized as <i>Grouping, Block, or Inline</i>	ISO 32000-1:2008, Table 340 ISO 32000-2:2020, Table 374
H_n (Heading level <i>n</i>)	PDF 1.7: heading levels where <i>n</i> is an <i>integer between 1-6</i> (inclusive) only PDF 2.0: heading levels where <i>n</i> is an <i>integer ≥ 1</i>	ISO 32000-1:2008, Table 335 ISO 32000-2:2020, Table 366
L (List)	PDF 1.7: categorized as <i>Block only</i> PDF 2.0: categorized as <i>Block or Inline</i>	ISO 32000-1:2008, Table 334 and §14.8.4.3.3 ISO 32000-2:2020, Table 370
Lbl (Label)	PDF 1.7: categorized as <i>Block only</i> PDF 2.0: categorized as <i>Inline only</i> for content that distinguishes it from other content inside the same parent element.	ISO 32000-1:2008, Table 334 ISO 32000-2:2020, Table 368
Link (Link Annotation)	PDF 1.7: categorized as <i>Inline only</i> PDF 2.0: categorized as <i>Grouping, Block, or Inline</i>	ISO 32000-1:2008, Table 338 and §14.8.4.4.2 ISO 32000-2:2020, Table 368
Part	PDF 1.7: a large-scale division of a document (e.g. for grouping articles or sections) PDF 2.0: a grouping of structure elements <i>without hierarchy</i>	ISO 32000-1:2008, Table 333 ISO 32000-2:2020, Table 365
Sect (Section)	PDF 1.7: a container for grouping <i>related</i> content elements PDF 2.0: a grouping of structure elements <i>with hierarchy</i>	ISO 32000-1:2008, Table 333 ISO 32000-2:2020, Table 365

PDF 1.7 tags <i>not</i> in PDF 2.0 standard structure namespace
Art
BlockQuote
BibEntry
Code
Index
Note
Private
Quote
Reference
TOC / TOCI

Structure element introduced in PDF 2.0	Suggested role mapping to PDF 1.7 tag
Aside	Div
Artifact	Private
DocumentFragment	Div
Em	Span
FENote	Note
H_n (where <i>n</i> ≥ 7)	P or H6
Strong	Span
Sub	Div or Span
Title	H1 or P

See ISO 32000-2:2020, Annex M.

PDF Association Cheat Sheet – Standard Structure Elements

PDF 1.7 Standard Structure Elements

All Tagged PDFs (document catalog **MarkInfo** dictionary, **Marked** key is **true**) must start with a **Document** structure element with the **StructTreeRoot** as its parent (**P** key) as the start of the structure hierarchy.

	Structure Type (s key)	Tables ¹	Description	Version	
Grouping	Art	333	Article. A relatively self-contained body of text as a single narrative or exposition.	1.4	
	BlockQuote	333	One or more paragraphs of text by someone other than the author of the surrounding text.	1.4	
	Caption	333	Brief portion of text describing a table, list or figure.	1.4	
	Div	333	Division. Generic block-level element or group of elements.	1.4	
	Document	333	A complete document.	1.4	
	Index	333	Sequence of entries containing identifying text accompanied by reference elements.	1.4	
	NonStruct	333	Grouping element having no inherent structural significance; it serves solely for grouping purposes.	1.4	
	Part	333	Large-scale division of a document.	1.4	
	Private	333	Grouping element containing private content belonging to the application producing it.	1.4	
	Sect	333	Section. A container for grouping related content elements.	1.4	
	TOC	333	Table of Contents. A list made up of TOCI and/or other nested table of contents (TOC) entries.	1.4	
	TOCI	333	Table of Contents Item. An individual member of a table of contents (TOC).	1.4	
Block	H	334, 335	Heading level determined by nesting depth. Typically the first child of a division.	1.4	
	H_n (<i>n = 1-6 only</i>)	334, 335	Heading. Explicitly numbered heading levels 1-6.	1.4	
	L	334, 336	List. Children are optional caption (Caption) followed by one or more list items (LI).	1.4	
	Lbl	334, 336	Label. Name or number that distinguishes an item from others in the same list or group of like items.	1.4	
	LBody	334, 336	List body. The descriptive content of a list item.	1.4	
	LI	334, 336	List item. Children may be one or more labels (Lbl), list bodies (LBody), or both.	1.4	
	P	334, 335	Paragraph of text.	1.4	
	Table	334, 337	2D logical structure of table cells in table rows.	1.4	
Inline	Annot	338	An association between a portion of the ILSE's content and a corresponding PDF annotation.	1.5	
	BibEntry	338	Bibliographic entry identifying the external source of some cited content. May contain a label (Lbl).	1.4	
	Code	338	Fragment of computer code.	1.4	
	Figure	340	Encloses graphical content.	1.4	
	Form	340	A widget annotation representing an interactive form field.	1.4	
	Formula	340	Encloses a mathematical formula.	1.4	
	Link	338	Association between Link structure element and link annotation(s).	1.4	
	Note	338	Note of explanatory text such as footnote or endnote. May contain a label (Lbl).	1.4	
	Quote	338	Inline portion of text attributed to someone other than the author of the surrounding text.	1.4	
	Reference	338	A citation to content elsewhere in the document.	1.4	
	Span	338	Generic inline portion of content with no inherent characteristics.	1.4	
		Ruby	338, 339	Contains entire Ruby assembly as one RB followed by either one RT or the 3-elements RP , RT , RB .	1.5
		Warichu	338, 339	Contains entire Warichū assembly. May contain WP and WT .	1.5
Ruby / Warichū	RB	339	Ruby base text. Only valid parent is Ruby .	1.5	
	RP	339	Ruby punctuation. Only valid parent is Ruby .	1.5	
	RT	339	Ruby annotation text. Only valid parent is Ruby .	1.5	
	WP	339	Warichū punctuation. Only valid parent is Warichu .	1.5	
	WT	339	Warichū text. Only valid parent is Warichu .	1.5	
		TR	337	Table row containing table header (TH) or table data (TD) cells.	1.4
Table	THead	337	Table header row group (TR) constituting the table's header.	1.5	
	TBody	337	Group of rows (TR) constituting the table's body.	1.5	
	TFoot	337	Table footer row group (TR) constituting the table's footer.	1.5	
	TH	337	Table header cell for one or more rows or columns. Only valid parent is TR .	1.4	
	TR	337	Table row containing table header (TH) or table data (TD) cells.	1.4	
	TD	337	Table data cell. Only valid parent is TR .	1.4	

1. PDF 1.7 structure element Table references are to ISO 32000-1:2008.

BLSE = Block-Level Structure Element

ILSE = Inline-Level Structure Element

PDF 2.0 Standard Structure Elements

All Tagged PDFs (document catalog **MarkInfo** dictionary, **Marked** key is **true**) must start with a **Document** structure element with the **StructTreeRoot** as its parent (**P** key) as the start of the structure hierarchy.

	Structure Type (s key)	Table ¹	Description	Version
Document	Document	364	Encloses a logical document.	1.4
	DocumentFragment	364	Logical document fragment, extracted from an original complete document.	2.0
Grouping	Aside	365	Encloses distinct content from other content in parent struct element.	2.0
	Div	365	Grouping orthogonal to the semantic structure.	1.4
	NonStruct	365	Grouping that lacks inherent structural significance.	1.4
	Part	365	Grouping of structure elements without hierarchy.	1.4
	Sect	365	Grouping of structure elements with hierarchy.	1.4
Block	H_n (n ≥ 1)	366	Explicitly numbered heading levels. No upper limit for <i>n</i> .	1.4
	FENote	366	Footnotes and endnotes.	2.0
	H	366	Heading. Typically, the first structure element in its parent.	1.4
	P	366	Paragraph.	1.4
	Title	366	Title of a document.	2.0
Inline	Sub	367	Sub-division of a block element.	2.0
	Annot	368	Contains PDF annotations (except Link and widget annotations) and associated content, if any.	1.5
	Em	368	Emphasis.	2.0
	Form	368	Contains widget annotations and associated content, if any.	1.4
	Lbl	368	Label.	1.4
	Link	368	Association between Link struct element and a link annotation.	1.4
	Span	368	Generic inline portion of content.	1.4
Ruby / Warichū	Strong	368	Strong importance, seriousness, or urgency.	2.0
	Ruby	369	Contains entire Ruby assembly as one RB followed by either one RT or 3-elements RP , RT , RB .	1.5
	RB	369	Ruby base text. Only valid parent is Ruby .	1.5
	RP	369	Ruby punctuation. Only valid parent is Ruby .	1.5
	RT	369	Ruby annotation text. Only valid parent is Ruby .	1.5
	Warichu	369	Contains entire Warichū assembly. May contain WP and WT .	1.5
List	WP	369	Warichū punctuation. Only valid parent is Warichu .	1.5
	WT	369	Warichū text. Only valid parent is Warichu .	1.5
	L	370	List (Block or inline).	1.4
	LI	370	List item.	1.4
	LBody	370	List item body.	1.4
	Table	Table	371	2D logical structure of table cells in table rows (Block).
TBody		371	Group of TR structure elements constituting the table's body.	1.5
TD		371	Table data cell.	1.4
TFoot		371	Group of TR structure elements constituting the table's footer.	1.5
TH		371	Table header cell.	1.4
Thead		371	Group of TR structure elements constituting the table's header.	1.5
TR		371	Row of table header (TH) or table data (TD) cells.	1.4
	Artifact	375	Artifact content (Grouping, Block or Inline).	2.0
	Figure	373	Encloses graphical content (Grouping, Block or Inline).	1.4
	Formula	374	Encloses a formula (Grouping, Block or Inline).	1.4
	Caption	372	Encloses a caption (Grouping or Block). Must be first or last child in the semantic parent element.	1.4
	math	§14.8.6.3	MathML 3.0 structure elements with NS namespace http://www.w3.org/1998/Math/MathML	2.0

1. PDF 2.0 structure element Table references are to ISO 32000-2:2020 including resolved errata at <https://pdf-issues.pdfa.org>.

BLSE = Block-Level Structure Element

ILSE = Inline-Level Structure Element

ISO TS 32005 defines containment rules for child/parent nesting of all PDF 1.7 and PDF 2.0 structure elements in PDF 2.0 files:

<https://pdfa.org/resource/iso-ts-32005-hierarchical-inclusion-rules/>