

Accessible Tables using Tagged PDF

TUG-2022, 22–24 JULY 2022 — ROSS MOORE



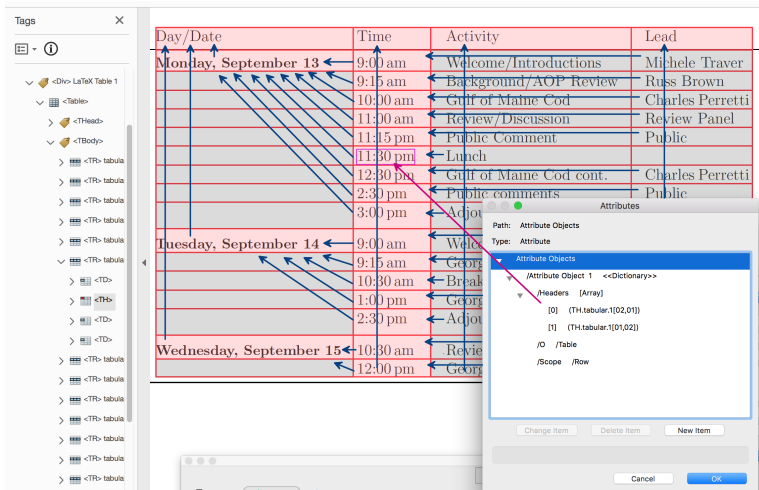
What makes a table Accessible?

This page gives a good summary of what is needed in an electronic document: <https://www.w3.org/WAI/tutorials/tables/>

- ▶ header and data cells are marked as such;
- ▶ each data cell needs to be associated with the relevant header cells that tell what the data means;
- ▶ the scope of each header cell is given; that is, to which rows or columns, or groups thereof, does it refer;
- ▶ a caption and/or summary to give a brief overview.

Relying on visual cues alone is not sufficient to create an accessible table. With structural markup, headers and data cells can be programmatically determined using software ...

Example: Calendar/schedule



The image shows a LaTeX Beamer presentation slide with a calendar table. On the left, a 'Tags' sidebar displays the document structure, including a table with columns for Day/Date, Time, Activity, and Lead. The table content is as follows:

Day/Date	Time	Activity	Lead
Monday, September 13	9:00 am	Welcome/Introductions	Michele Traver
	9:15 am	Background/AOP Review	Russ Brown
	10:00 am	Gulf of Maine Cod	Charles Perretti
	11:00 am	Review/Discussion	Review Panel
	11:15 pm	Public Comment	Public
	11:30 pm	Lunch	
	12:30 pm	Gulf of Maine Cod cont.	Charles Perretti
	2:30 pm	Public comments	Public
	3:00 pm	Adjou	
Tuesday, September 14	9:00 am	Welc	
	9:15 am	Georg	
	10:30 am	Break	
	1:00 pm	Georg	
	2:30 pm	Adjou	
Wednesday, September 15	10:30 am	Revie	
	12:00 pm	Georg	

The 'Attributes' dialog box is open, showing the following structure:

```
Attributes
Path: Attribute Objects
Type: Attribute
Attribute Objects
  /Attribute Object 1 <<Dictionary>>
    /Headers [Array]
      [0] (TH.tabular:1[02,01])
      [1] (TH.tabular:1[01,02])
    /O /Table
    /Scope /Row
```

Examples: more table heads

Table 1: Comparison of reference points estimated in an earlier assessment and from the current update. The overfishing threshold is the $F_{MSY_{proxy}}$ ($F_{0.95}$). The biomass target, ($SSB_{MSY_{proxy}}$) w long-term stochastic projections of fishing at the $F_{MSY_{proxy}}$. Median recruitment reflects the median age-1 recruitment from 1982–2017. Intervals shown reflect the 5P and 95P percentiles.

	2019 $M=0.2$	2019 M_{ramp}	$M=0.2$	M_{ramp}
F_{MSY}	0.149	0.145	0.143	0.175
SSB_{MSY} (mt)	42,692 (27,916, 62,785)	63,867 (46,144, 84,098)	39,912 (25,353, 59,589)	60,916 (41,916, 84,098)
MSY (mt)	7,540 (4,863, 1,366)	11,201 (8,149, 5,268)	7,111 (4,462, 11,023)	10,472 (7,439, 14,929)
Median recruits (age-1) (000s)	4.377 (1,161, 4,434)	8.464 (2,853, 5,934)	4.677 (1,064, 6,392)	9.219 (2,129, 18,434)
Overfishing	Yes	Yes	Yes	No
Overfished	Yes	Yes	Yes	No

The image shows a screenshot of a LaTeX editor interface with a table and a right-hand panel. The table is highlighted with a red border. The right-hand panel shows the 'Attributes' section for the selected table cell, with the 'ID' field set to 'TD.tabular:1[4,3]'.

The image shows a screenshot of a LaTeX editor interface with a table and a right-hand panel. The table is highlighted with a red border. The right-hand panel shows the 'Attributes' section for the selected table cell, with the 'ID' field set to 'TD.tabular:1[4,3]'.

Year	Catch (mt)	SSB (mt)	F_{Full}	Catch (mt)	SSB (mt)	F_{Full}
2020	409	2,635	0.162	409	3,925	0.119
2021	523	3,500	0.127	523	4,750	0.113
2022	821	4,508	0.123	949	5,251	0.175
2023	959	5,488	0.123	1,417	5,701	0.175
2024	1,244	7,279	0.173	1,306	6,802	0.175

The image shows a screenshot of a LaTeX editor interface with a table and a right-hand panel. The table is highlighted with a red border. The right-hand panel shows the 'Attributes' section for the selected table cell, with the 'ID' field set to 'TD.tabular:1[4,3]'.

Examples: hierarchical table heads

Catch and status table for Gulf of Maine Atlantic cod. All weights are in (mt), recruitment is the fishing mortality on fully selected ages. Note terminal year SSB and F_{full} is not re.

	2012	2013	2014	2015	2016	2017	2018	2019
Data								
Recreational harvest	0.00	1.05	1.51	1.68	1.54	1.71	1.60	1.11
Recreational landings	3.45	5.94	7.96	11.16	11.16	11.16	11.16	4.3
Commercial harvest	0.0	5.4	9.7	14.1	16.9	16.9	16.9	7
Commercial landings	2.50	4.54	8.24	10.7	13.0	13.0	13.0	3.5
Catch for Assessment	4.95	10.38	16.90	25.87	27.07	27.07	27.07	10.7
Model Results (I=0.2)								
Spawning Stock Biomass	10.04	18.36	24.5	18.1	17.9	23.0	20.4	28.3
F _{full}	1.66	2.16	2.7	3.19	3.50	3.61	3.72	3.16
Recruits (age-1)	110.6	46.7	33.0	30.1	5.0	0.6	0.14	2.08
Model Results (I=0.7)								
Spawning Stock Biomass	11.71	28.8	65.5	85.0	185	277.0	276.0	223
F _{full}	1.46	1.85	2.1	2.3	2.4	2.4	2.4	1.7
Recruits (age-1)	192.5	148.1	179	169	102.4	17.7	11.0	181

Diagram illustrating the hierarchical structure of the table head. Red arrows point from the table cells to the corresponding nodes in the tree view on the left and the object properties/attributes panels on the right.

Table 6: Overzicht van de concentraties en hun standaarddeviaties van pesticiden in mg/kg in de (n)behandelde wortelen en de meelwormen gekweekt op deze wortelen. Bij NA lag de concentratie onder de detectielimiet van de LCMS-MS.

Pesticiden	Concentraties pesticiden (mg/kg)			
	Wortel	Oelbuisveld	Gebruikelijk	Behandeld
Wortel	na nit-bodding	na nit-bodding	na nit-bodding	na nit-bodding
Meelwormen	na nit-bodding	na nit-bodding	na nit-bodding	na nit-bodding
Wortel	na nit-bodding	na nit-bodding	na nit-bodding	na nit-bodding
Meelwormen	na nit-bodding	na nit-bodding	na nit-bodding	na nit-bodding

Diagram illustrating the hierarchical structure of the table head. Red arrows point from the table cells to the corresponding nodes in the tree view on the left and the object properties/attributes panels on the right.

Examples in HTML

REAL-WORLD TABLES — FISHERY DATA

Normal Table

The basic style adopted for tabular material is in accordance with `tbl`. Supported rules are `\toprule`, `\midrule`, `\bottomrule` which can inherit instead of being for aesthetic appearance only. This extra semantics takes `!`

- `\toprule` indicates the start of a (THead) block of rows;
- `\midrule` indicates the end of a block, with the following row starting;
- `\bottomrule` indicates the end of the tabular; closing the (TBody) block

Table 1: Comparison of reference points estimated in an earlier assessment and from update. The overfishing threshold is the $F_{MSY_{proj}}$ ($F_{40\%}$). The biomass target, (S): long-term stochastic projections of fishing at the $F_{MSY_{proj}}$. Median recruitment reference-1 recruitment from 1982–2017. Intervals show reflect the 5th and 95th percent

	2019 $M=0.2$	2019 $M=0.2$	$M=0.2$
F_{MSY}	0.173	0.175	0.173
SSB_{MSY} (mt)	42,692 (27,916–62,785)	63,867 (46,144–84,098)	39,912 (25,472–59,589)
MSY (mt)	7,580 (4,853–11,366)	11,420 (8,149–15,268)	7,171 (4,462–11,023)
Median recruits (age-1) (000s)	4,377 (1,161–14,434)	8,464 (2,353–15,934)	4,677 (1,064–16,392)
Overfishing	Yes	Yes	Yes
Overfished	Yes	Yes	Yes

Tagged tabulars from Chapter 5

Alignment using preamble commands

§5-2-1	<table border="1"><tr><td>A</td><td>B</td><td>C</td></tr><tr><td>100</td><td>10</td><td><i>I</i></td></tr></table>	A	B	C	100	10	<i>I</i>	§5-2-2	<table border="1"><tr><td>A</td><td>B</td><td>C</td></tr><tr><td>100</td><td>10</td><td><i>I</i></td></tr></table>	A	B	C	100	10	<i>I</i>	<table border="1"><tr><td>100</td></tr><tr><td>A</td></tr></table>	100	A
A	B	C																
100	10	<i>I</i>																
A	B	C																
100	10	<i>I</i>																
100																		
A																		

Checking alignments:

§5-2-3	<table border="1"><tr><td>1 1 1 1</td><td>2 2 2 2</td><td>3 3 3 3</td></tr><tr><td>1 1 1 1</td><td>2 2 2 2</td><td></td></tr><tr><td>1 1 1 1</td><td></td><td></td></tr></table>	1 1 1 1	2 2 2 2	3 3 3 3	1 1 1 1	2 2 2 2		1 1 1 1			§5-2-4	<table border="1"><tr><td>1 1 1 1</td><td>2 2 2 2</td><td>3 3 3 3</td></tr><tr><td>1 1 1 1</td><td>2 2 2 2</td><td></td></tr><tr><td>1 1 1 1</td><td></td><td></td></tr></table>	1 1 1 1	2 2 2 2	3 3 3 3	1 1 1 1	2 2 2 2		1 1 1 1		
1 1 1 1	2 2 2 2	3 3 3 3																			
1 1 1 1	2 2 2 2																				
1 1 1 1																					
1 1 1 1	2 2 2 2	3 3 3 3																			
1 1 1 1	2 2 2 2																				
1 1 1 1																					

§5-2-5	<table border="1"><tr><td>1 1 1 1</td><td>2 2 2 2</td><td></td></tr><tr><td>1 1 1 1</td><td>2 2 2 2</td><td>3 3 3 3</td></tr><tr><td>1 1 1 1</td><td>2 2 2 2</td><td>3 3 3 3</td></tr></table>	1 1 1 1	2 2 2 2		1 1 1 1	2 2 2 2	3 3 3 3	1 1 1 1	2 2 2 2	3 3 3 3	§5-2-13	one two three 1 2 3
1 1 1 1	2 2 2 2											
1 1 1 1	2 2 2 2	3 3 3 3										
1 1 1 1	2 2 2 2	3 3 3 3										

Using array package features:

§5-2-6	<table border="1"><tr><td>1 2 3 4 5 6</td><td>1 2 3 4 5 6 7 8</td></tr><tr><td>7 8 9 0 1 2 3 4</td><td>9 0 1 2 3 4 5 6</td></tr><tr><td>5 6 7 8 9 0</td><td>7 8 9 0</td></tr></table>	1 2 3 4 5 6	1 2 3 4 5 6 7 8	7 8 9 0 1 2 3 4	9 0 1 2 3 4 5 6	5 6 7 8 9 0	7 8 9 0	§5-2-7	<table border="1"><tr><td>$10^{10!}$</td><td>a big number</td></tr><tr><td>10^{-999}</td><td>a small number</td></tr></table>	$10^{10!}$	a big number	10^{-999}	a small number
1 2 3 4 5 6	1 2 3 4 5 6 7 8												
7 8 9 0 1 2 3 4	9 0 1 2 3 4 5 6												
5 6 7 8 9 0	7 8 9 0												
$10^{10!}$	a big number												
10^{-999}	a small number												

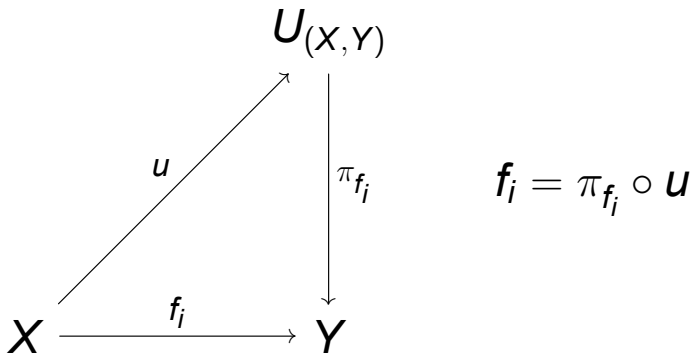
Super-consciousness is a	Possibilités et espérances	Mogelijkheden en hoop	Super-consciousness is a	Possibilités espérances
--------------------------	----------------------------	-----------------------	--------------------------	-------------------------

website: <http://science.mq.edu.au/~ross/TaggedPDF/TUG2022/>

more tagged PDFs: <http://science.mq.edu.au/~ross/TaggedPDF/>

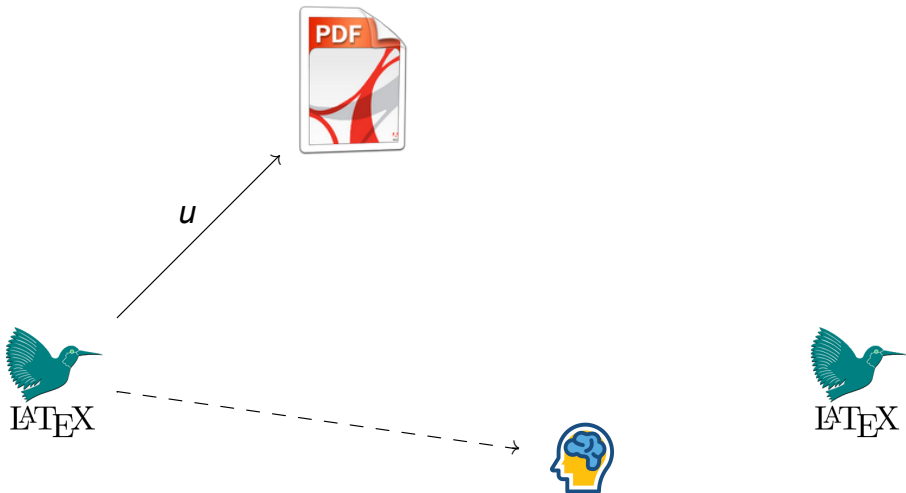
Algebra: Universal Morphism

What's in a name? PDF/UA *Universal Accessibility*



Any structure preserving map $f_i : X \rightarrow Y$ factors through the 'Universal Object' $U_{(X,Y)}$, via a projection π_{f_i} .

PDF/UA: Universal Accessibility



Visualising how a Tabular works

Template: generated from the `<cols>`

```

... # ... & ... # ... & ...   ...   & ... # ... \cr
  ↑      ↑                       ↑
  └──┬──┘ └──┬──┘               └──┬──┘
  ...   ... & ...   ... & ...   ...   & ...   ... \l

```

User data:

```

...   ... & ...   ... & ...   ...   & ...   ... \l
...   ... & ...   ... & ...   ...   & ...   ... \l
\end{tabular}

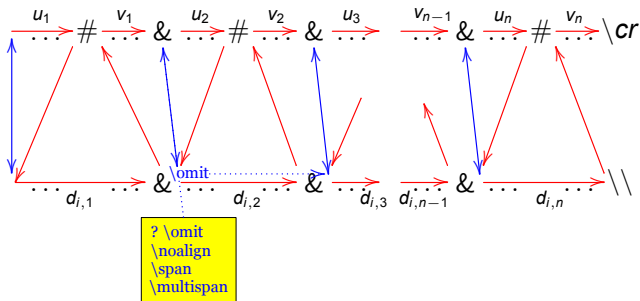
```

We conceptualise a table as merging user-supplied data into the preamble/template; both as an author coding the data, and as a reader visualising some or all of the overall layout.

However, this is *not* the way \TeX processes these token-streams!

Switching, not merging

Template: generated from the $\langle \text{cols} \rangle$



User data:

Summary: $\text{T}_{\text{E}}\text{X}/\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$ processes tokens for row $_i$ in the order

$$u_1 d_{i,1} v_1 u_2 d_{i,2} v_2 u_3 d_{i,3} \dots d_{i,n-1} v_{n-1} u_n d_{i,n} v_n$$

with the ability to omit some portions of the template, and to include extra non-aligned material.

A L^AT_EX preamble

`p{3.75cm} p{4.5cm} p{7.75cm}` produces the following preamble:

```
\unhcopy\@arstrutbox \hskip\col@sep \setbox\z@ \hbox\bgroup\bgroup\CT@everycr
{} \vtop\@startpbox {3.75cm} \ignorespaces \@sharp \unskip \relax\@endpbox
\do@row@strut \egroup\egroup\begin@group\CT@setup \CT@row@color \CT@cell@color
\CT@do@color \endgroup\@tempdima \ht\z@ \advance\@tempdima
\minrowclearance\LTX@vrule \@height \@tempdima \@width \z@ \unhbox\z@
\hskip\col@sep & \hskip\col@sep \setbox\z@ \hbox\bgroup\bgroup\CT@everycr
{} \vtop\@startpbox {4.5cm} \ignorespaces \@sharp \unskip \relax\@endpbox
\do@row@strut \egroup\egroup\begin@group\CT@setup \CT@row@color \CT@cell@color
\CT@do@color \endgroup\@tempdima \ht\z@ \advance\@tempdima
\minrowclearance\LTX@vrule \@height \@tempdima \@width \z@ \unhbox\z@
\hskip\col@sep & \hskip\col@sep \setbox\z@ \hbox\bgroup\bgroup\CT@everycr {}
\vtop\@startpbox{7.75cm} \ignorespaces \@sharp \unskip \relax\@endpbox
\do@row@strut \egroup\egroup\begin@group\CT@setup \CT@row@color \CT@cell@color
\CT@do@color \endgroup\@tempdima \ht\z@ \advance\@tempdima
\minrowclearance\LTX@vrule \@height \@tempdima \@width \z@ \unhbox\z@
\hskip\col@sep \tabskip\z@ \cr
```

Notes on the preamble

- ▶ `\@sharp` becomes `#` during processing of the table's rows; during the creation of the preamble, it is set to `\relax`, to be inert under repeated expansion.
- ▶ Notice that `\ignorespaces` occurs immediately before the data tokens for a cell's contents, as does `\unskip` come immediately after. But as these tokens are commonly used with input, and can be redefined within various packages, they are *not useful* as hooks to begin and end content tagging.
- ▶ Instead, the `\tabular` macro is adjusted to perform tagging tasks before calling up its usual \LaTeX expansion.

The next page illustrates how tagging is started this way, and propagated through the cells and rows, by adjusting the `&` token.

Preamble with some tagging aspects

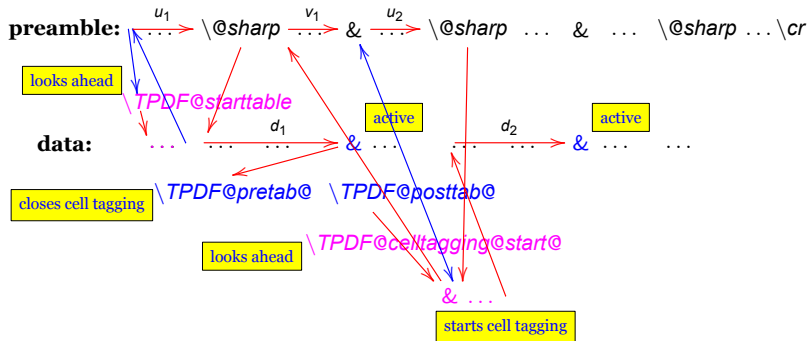
`|>\{\large}c|>\{\large\bfseries}l |>\{\itshape}c|` produces the following preamble:

```
\unhcopy\@arstrutbox {\vrule\@width \arrayrulewidth \TPDF@settabcol@prop{0}{9}{>} \hskip\col@sep
\TPDF@settabcol@prop {0}{0}{c} \setbox\z@ \hbox\bgroup\bgroup\CT@everycr{} \hskip\z@ plus.5fill \relax
\kern\z@ \dollarbegin \advance\TPDF@thiscolcount \@ne \large\ignorespaces\@sharp \unskip\relax
\dollarend \do@row@strut \hskip\z@ plus.5fill \relax\egroup\egroup\begin@group\CT@setup \CT@row@color
\CT@cell@color \CT@do@color \endgroup\@tempdima \ht\z@ \advance\@tempdima \minrowclearance\vrule\@height
\@tempdima \@width \z@ \unhbox\z@ \hskip\col@sep {\vrule\@width \arrayrulewidth}
\TPDF@settabcol@prop {0}{9}{>} & \hskip\col@sep \TPDF@settabcol@prop {1}{0}{1} \setbox\z@
\hbox\bgroup\bgroup \CT@everycr{} \dollarbegin \advance\TPDF@thiscolcount \@ne \large\bfseries
\ignorespaces\@sharp \unskip\relax\dollarend \do@row@strut \hfill\egroup\egroup\begin@group\CT@setup
\CT@row@color \CT@cell@color \CT@do@color \endgroup\@tempdima \ht\z@ \advance\@tempdima
\minrowclearance\vrule\@height \@tempdima \@width \z@ \unhbox\z@ \hskip\col@sep {\vrule\@width
\arrayrulewidth} \TPDF@settabcol@prop {0}{9}{>} & \hskip\col@sep \TPDF@settabcol@prop {0}{0}{c}
\setbox\z@ \hbox\bgroup\bgroup \CT@everycr{} \hskip\z@ plus.5fill \relax\kern\z@ \dollarbegin
\advance\TPDF@thiscolcount \@ne \itshape\ignorespaces\@sharp \unskip\relax\dollarend \do@row@strut
\hskip\z@ plus.5fill\relax\egroup \egroup \begin@group\CT@setup \CT@row@color \CT@cell@color \CT@do@color
\endgroup\@tempdima \ht\z@ \advance\@tempdima \minrowclearance\vrule\@height \@tempdima \@width \z@
\unhbox\z@ \hskip\col@sep {\@width \arrayrulewidth}\tabskip\z@ \cr
```

The macro `\TPDF@settabcol@prop` here records the cell's text-alignment, to be passed to HTML as an attribute via CSS rules. When a `<cols>` entry represents vertical alignment, or some semantic information, then this can be handled also.

Tagging a Tabular layout

\LaTeX 's usual `\tabular` is replaced with `\TPDF@tabular` which, after reading the `\langle cols \rangle`, does many tagging-related tasks including making `&` into an 'active' character. A macro `\TPDF@starttable` is inserted immediately before the user data being read from the file-source. This initiates a 'look-ahead' to see what is coming next; e.g., `\toprule` or `\hline`, `\multicolumn` or `\multirow`, a blank cell or content for a normal cell; taking appropriate actions for each.



Further cells are tagged, using the active `&`. When a row is done, its tagging is closed; there is a 'look ahead' to commence tagging the next row's cells, following any `\hrule`, `\midrule`, etc.